Comparison of Shaping and Buffering for Video Transmission

György Dán and Viktória Fodor
Royal Institute of Technology,
Department of Microelectronics and Information Technology
P.O.Box Electrum 229, SE-16440 Kista, Sweden
{gyuri,viktoria}@it.kth.se

Abstract

Video communication over the Internet requires performance guarantees in terms of limited packet loss probability and end to end delay. This paper compares two possible network scenarios for transmitting video streams, source shaping combined with small buffers at the network nodes and delay limited buffering in the network. It is shown that due to its simplicity and performance comparable to buffering source shaping can provide the solution for efficient video transmission in the Internet.

1 Introduction

The transmission of video traffic over large packet switched networks like the Internet is still a fundamental problem of network design. First, video applications require quality of service guarantees from the network, in terms of limited packet loss, end to end delay and delay variation. The acceptable packet loss probability depends on the video coding scheme and is in the range of 10^{-5}–10^{-3}. The delay limitation depends on the application. While in the case of computer to human applications delay can be up to several seconds, in the case of human to human applications delay should be kept below 150 ms. Second, the provisioning of the required quality at a network utilization acceptable for the operator is a complex issue due to the characteristics of coded video streams. Considering the widely used MPEG coding, the transmission rate is changing in short and long time scales. Moreover, as a result of data compression the distribution of the packet losses affects the percepted visual quality: multiple packet loss in a single frame decreases the visual quality of the corresponding picture, and the effect of a packet loss in an I frame propagates to P and B frames.

Research and developement efforts today address the problem of finding efficient traffic control solutions that support visual communication and can be introduced with acceptable cost [3, 7]. One of the essential questions is whether to design networks with large buffers at the network nodes or rather use small buffers and source shaping when transmitting delay and loss sensitive video traffic (E.g., [8, 9]). Intuitively, using large buffers low packet loss probability can be provided even at high network utilization, but the control of end to end delay and delay variation has to be solved. On the other hand, source shaping with small buffers at the nodes bounds the delays but at the price of increased packet loss probability.

Considering the feasibility of the two solutions, shaping has its advantages, as it can be introduced gradually, according to the individual applications' needs and does not require any support from the network. On the other hand, in networks with large buffers delay aware scheduling and/or jitter compensation has to be introduced at all the network nodes to limit the end to end delay and avoid the increase of burstiness of the traffic streams.

In [9] the performance of source shaping and buffering is compared for networks providing strict end to end delay bounds. The paper compares two solutions. In one of them, source shaping with the maximum acceptable delay is applied and nodes are equipped with small buffers, performing so called packet scale buffering [1]. In the other solution the maximum acceptable delay is divided among the network nodes thus nodes perform burst scale buffering with buffer size defined by the per node maximum delay. Nodes apply jitter compensation for each video stream to control the burstiness of the stream. It is shown that in the terms of packet loss traffic shaping is less efficient in the case of a single node multiplexer, but in the case of long transmission paths, when the maximum end to end delay has to be split among many nodes, traffic shaping outperforms buffering.

In this paper we further evaluate the performance of these two solutions, comparing the distribution of losses and their effect on the visual quality of the received video stream. The paper is organized as follows. The next section describes the two network scenarios considered, section 3 presents simulation results and discusses the loss characteristics of the scenarios, finally in section 4 we conclude our work.

2 System description

We consider the following networking scenario. On-line video streams have to be transmitted through a large packet switched network - specifically, the Internet. We assume, that the network has DiffServ [2] capabilities, and the transmission of video streams is not disturbed by best effort traffic.

We assume that call admission control is introduced for the video streams to provide performance guarantees. Measurement based end-node call admission control solutions that limit the packet loss probability of the accepted streams without any network signaling or per stream processing at the nodes are proposed in e.g.,[3, 5].

While packet loss is limited by the admission control, end to end delay is limited by the network architecture, specifically by the buffer sizes at the source shapers and network nodes.

The considered system is shown in figure 1, consisting of the source and the destination of the video stream, the source coder and decoder, the transmission and reception control units that can contain source shaper and playout buffer and error control coding and decoding.

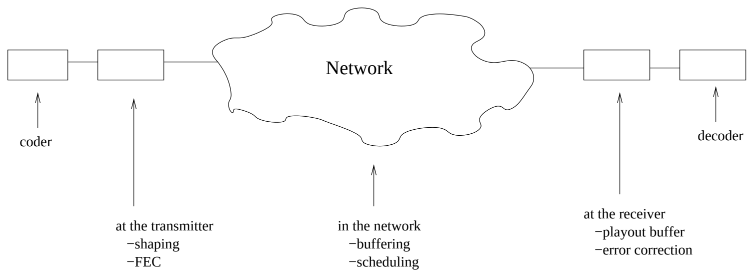

Figure 1: Considered network model

Source coding

The sources generate MPEG coded streams, the most commonly used encoding scheme for the storage and transmission of video information. In the MPEG stream information is stored as a sequence of frames, corresponding to a sequence of pictures in a video, generated with fix time intervals. Compression is achieved by eliminating the spatial and temporal redundancy of the information in the frames. Spatial redundancy is decreased by intraframe coding of the individual frames, while temporal redundancy is reduced by interframe coding between subsequent frames. Thus the sequence of frames consists of intraframe coded frames (I frames), and interframe coded predicted (P frames) and bidirectionally predicted frames (B frames). The subsequent frames between two consecutive I frames form a group of picture (GOP). The GOP structure of the streams can be different, depending on the required quality. A typical example for the sequence of frames is IBBPBBPBBPBB.

The stream is highly bursty, with fluctuations on two timescales. The intraframe coding compresses complex scenes with less efficiency, and consequently, the frame sizes change on the long term at the scene changes. The interframe coding leads to short term frame size fluctuation, since I frames are usually significantly larger than P frames, and P frames are larger than B frames. The scale of the fluctuation is about a factor of 3 on the long, and a factor of 10 on the short term.

As a consequence of the coding scheme, information loss in the three frame types has different effect on the percepted visual quality. The loss of data in an I frame propagates forward through the next GOP and backward to the last P frame (affecting up to 14 frames if the number of frames in an open GOP is 12). Meanwhile, the loss of data in a B frame only affects that particular frame.

Source shaping

Shapers used at the sources decrease the frame to frame fluctuation of the coded video stream. The shaper we use in this work is a single buffer leaky bucket, as it is proved to be optimal for networks with small buffers [8]. Frames leaving the encoder are stored in the shaper buffer and are transmitted with a given transmission rate, which is adjusted to provide lossless, delay limited shaping. Shaper algorithms for on-line video streams are proposed in [6]. In this paper, we apply a low complexity solution based on the shaper buffer content, as described in [4].

The algorithm to control the shaper transmission rate r aims to minimize the maximum and the variance of the transmission rate, is efficient for large range of end to end delay limit and assists real time operation as it has low computational complexity and does not require knowledge on the GOP structure.

The algorithm assumes that the shaper can detect the type of the arriving frame and applies the following rules to set the shaper rate.

The shaper rate can be changed at any frame arrival n. The lowest acceptable shaper rate $r_{min}(n)$ is calculated based on the amount of data in the buffer, $b(n)$ and the delay limit d, as:

$$r_{min}(n) = \frac{b(n)}{d}.$$

The shaper rate is increased at any frame arrival if $r_{min}(n) > r(n-1)$. The shaper rate is decreased if $r_{min}(n) < r(n-1)$, the new frame is of type I and the buffer is empty before the frame arrival. This rule is based on the assumption that small P or B frames do not indicate intensity change in the video stream and is restricted since the delay limit of individual frames in the buffer can not be ensured. To decrease fluctuation, the shaper rate is decreased gently,

$$r(n) = \frac{r_{min}(n) + r(n-1)}{2}.$$

P and B frames entering the shaper when the buffer is empty are transmitted with a rate such that the frame leaves the buffer before the new frame arrives, i.e., in one frame time, in order to prevent the shaper from keeping data before larger I and P frames arrive.

If shaping is not applied the source sends the individual frames smoothed over one frame time i.e., 40 ms if the frame rate is 25 frames per second.

Node architecture

Packet streams leaving the source nodes are multiplexed at the network nodes. Two different node architectures are considered.

1. The output buffers at the nodes provide buffering for simultaneously arriving packets only (packet scale buffering). In this case the buffer size is in the range of the ratio of the link speed and the peak rate of the streams.

2. The output buffers are large to provide buffering for bursts (burst scale buffering). The buffer size in this case is limited by the maximum acceptable end to end delay and the number of hops on the transmission path. Specifically, it means maximizing the buffer capacities by $B_i = \frac{d}{N}c_i$ where d is the maximum end to end delay, N the number of nodes on the transmission path, and c_i the link capacity at node i [9]. It is assumed that nodes apply jitter compensation, which means that the delay at the networking nodes is exactly d/N.

Shaping of the sources has some advantages compared to buffering inside the network. i) Shaping can be introduced gradually according to the applications' needs and nodes inside the network are not affected. Buffering with delay limit requires updating of all the network nodes. ii) Furthermore, buffering with delay limit requires per stream delay and jitter control at the nodes, with stream specific limits, an thus is a source of network scalability problems. In the case of source shaping the end nodes keep track of the delay control and network nodes do not perform per stream processing.

3 Performance evaluation

To compare shaping and delay limited buffering we compare the average packet loss probabilities, the packet loss distribution among I, P and B frames, the probability of consecutive packet losses and losses in small blocks of packets. The presented results are simulation results, the simulation time was between 20000 to 60000 seconds to have enough loss events even in the case of loss probabilities in the order 10^{-5}.

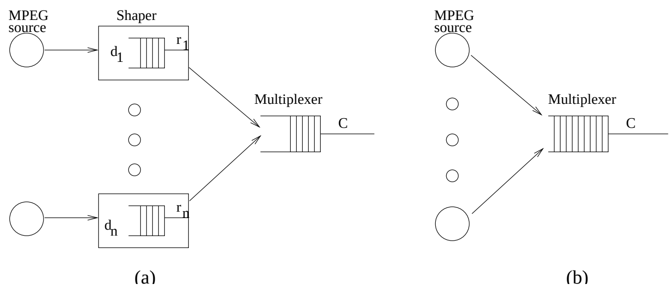

Figure 2: The considered network model, (a) with shapers, (b) with buffering

The simulated network model is shown in figure 2. The system includes traffic sources, source shapers and a multiplexing node with a single output link. Two scenarios are considered. In scenario (a) the nodes apply source shaping with delay limited by the maximum acceptable end to end delay and the multiplexer provides a small buffer for packet scale buffering. In scenario (b) the sources are not shaped and the acceptable per node delay is allocated to the multiplexer buffer.

We present results for an MPEG-4 video trace, a talk show with an average rate of 540 kbps. The trace is approximately 2700 seconds, thus 67000 frames long. The frames of the MPEG trace are packetized to 188 bytes, as given for the transport stream in the MPEG-2 standard [IEC61883]. The capacity of the output link is 22.5 Mbps. The size of the shaper buffer is determined by the considered end to end delay limits, 20 ms and 40 ms, acceptable for real-time communication. When shaping is used the buffer at the multiplexer stores up to 10 packets to provide packet scale buffering. When, instead, delay limited buffering is applied at the network nodes, we assume a transmission path length of 10 nodes, resulting buffer capacity for 38 and 66 packets for delays of 20 ms and 40 ms respectively.

Average packet loss probabilities

First we investigate the average packet loss probability of the multiplexed streams as a function of the average load at the multiplexer, defined by the ratio of the sum of the mean rates of the streams to the link transmission capacity.

Figure 3 shows the average packet loss probability as a function of the load for end to end delays of 20 ms and 40 ms for both the shaped and the buffered traces. Buffering results in a lower average packet loss probability, the difference is less than one order of magnitude. Considering an acceptable packet loss probability of 10^{-5}, the difference of the maximum network load with shaping and buffering is 3–8% depending on the delay limit.

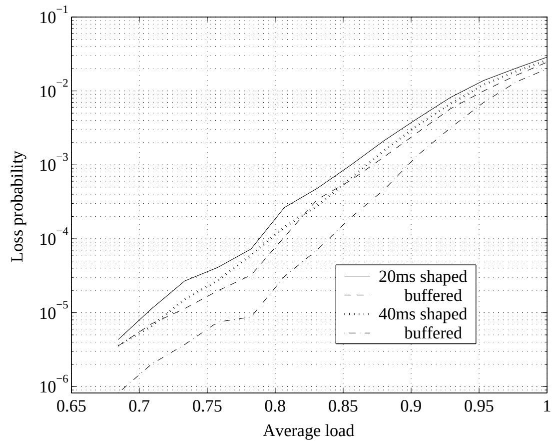

Figure 3: Average packet loss probability of the shaped and buffered streams.

Packet loss probabilities in I, P and B frames

In addition to the average packet loss probability of the streams it is worthwhile to evaluate the packet loss probability in individual frame types, since, as a consequence of the coding scheme it affects the percepted visual quality.

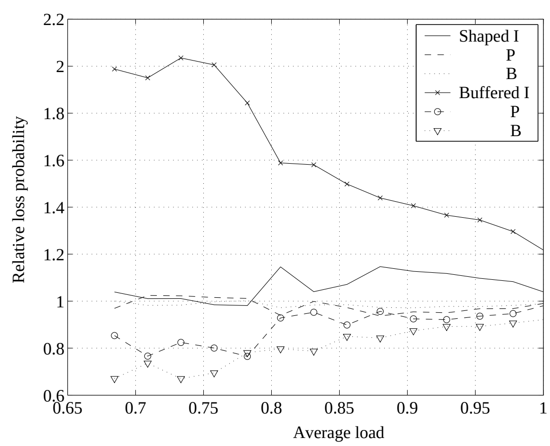

Figure 4: Relative packet loss probability in I,P,B frames of the shaped and buffered streams for d = 40 ms.

Figure 4 shows the packet loss probability in I, P and B frames relative to the average loss probability for the shaped and buffered sources for an end to end delay of 40 ms. The figure shows that while in the case of unshaped sources the loss probability in the I frames is the highest, up to 100% above the average loss probability and that in the B frames is the lowest, in the case of shaped sources the loss probabilities in the individual frame types are roughly the same. In the I frames the decrease of loss probability is around 60%. Consequently, as losses in the I frame have a significant effect on the visual quality, traffic shaping improves the distribution of the losses among the frame types.

Consecutive packet loss

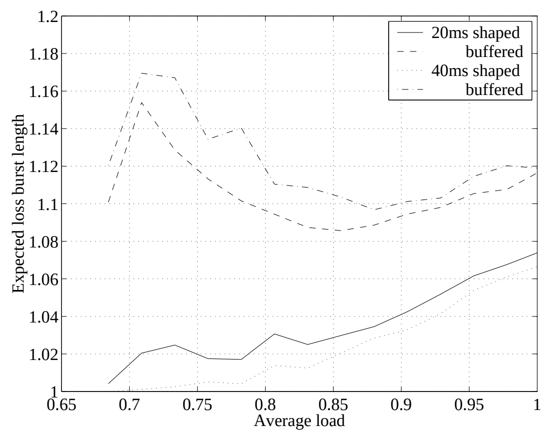

Figure 5: Expected number of consecutive packet losses.

Consecutive packet losses might have a degrading effect on the visual quality, as larger parts of a picture have to be recovered. Figure 5 shows the expected number of consecutively lost packets. In the case of the shaped sources the expected loss burst length is 1 for a load up to 0.7-0.8, meaning, that single packet losses happen in most of the cases, and increases as the load increases. In the case of buffered sources the graph of the expected burst length has a U shape, reaching a minimum of 1.08 consecutive packet losses at an average load of approximately 0.85. The shape of the graphs for the buffered case can be explained by the burstiness of the sources. In the case of low average load a bursty stream sending at a high bitrate is probable to cause a congestion period itself, and loose all its packets during that period, thus losses tend to occur in bursts. In the high load region the congestion periods get longer what explains the increase in the expected loss burst length for both the shaped and the buffered streams.

Considering the probability of consecutive packet losses, the result indicates that in the case of buffering the probability of two or more consecutive packet losses is orders of magnitude higher than in the case of shaping at low or modest average load.

Losses in packet blocks

This part evaluates the average number of packets lost in small blocks of packets. The results help to investigate whether forward error correction solutions, based on block coding of a number of packets can improve the quality of the transmission.

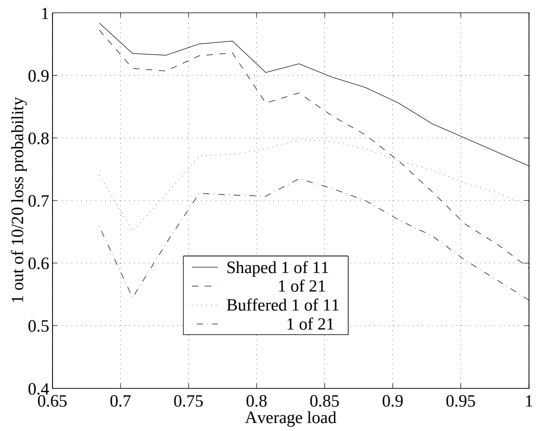

Figure 6: Probability of 1 packet loss out of 10 and 20 packets for d = 20 ms.

Figure 6 shows the probability that out of 10 or 20 packets only 1 packet, if any, gets lost for an end to end delay of 20 ms. Figure 7 shows the same probability for an end to end delay of 40 ms. For the shaped sources the probability of loosing 1 packet out of 10 or 20 is close to 1 at a load up to 0.8, and decreases as the load increases. The graphs showing the 1 out of 10 and 20 packet loss probability for the buffered sources however have an upside down U shape, in accordance with figure 5. Thus the probability of loosing more than 1 packet out of 10 or 20 first increases as the average load increases to reach its maximum of 0.73 at an average load of 0.85, then decreases as the average load further increases.

The results show that with source shaping maximum 5% of the losses are multiple losses in packet blocks up to a load of 0.8, compared to 20–30% with buffering. It indicates that that while in the case of buffering the high probability of multiple losses makes error correction coding inefficient, in the case of shaping forward error correction might be useful to improve transmission quality in terms of residual packet loss probability.

4 Conclusion

In this paper we evaluated the feasibility and efficiency of source shaping versus delay limited buffering for the transmission of delay and loss sensitive video transmission over the internet.

The extensive performance analysis, based on simulation, provided the following results.

- The average packet loss probability can be up to one order of magnitude lower in the case of buffering at the network nodes.
- The packet loss distibution among frames is uneven if source shaping is not applied, with high packet loss probability in the I frames. Source shaping equalizes the per frame packet loss probabilities.
- The probability of consecutive packet losses can be orders of magnitude lower in the case of source shaping, depending on the average loss probability.

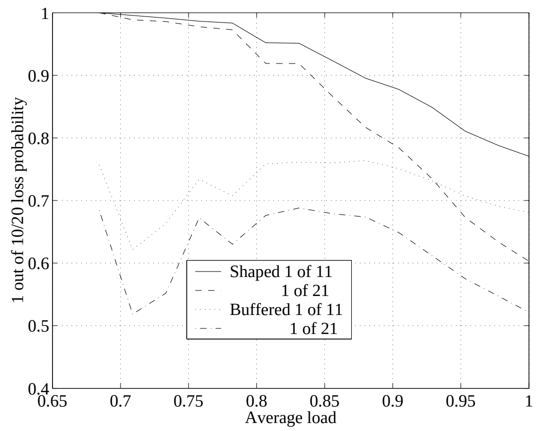

Figure 7: Probability of 1 packet loss out of 10 and 20 packets for d = 40 ms.

- The probability of multiple losses in blocks of packets is 4–6 times lower with shaping than with buffering, giving a fair chance of efficient forward error correction.

The above results indicate that while the average packet loss might be higher in the case of packet scale buffering combined with traffic shaping, due to the change in the packet loss distribution, the percepted visual quality can be close to the one in networks with large buffers. Considering the feasibility of the two solutions, we believe that source shaping together with small buffers at the network nodes can provide the solution for transmission of delay sensitive video traffic in the Internet.

References

[1] "Broadband Network Teletraffic, Final Report of Action COST 242," Springer, 1996.

[2] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An Architecture for Differentiated Services. RFC 2475", December 1998.

[3] L. Breslau, E. W. Knightly, S. Shenker, I. Stoica, H. Zhang, "Endpoint Admission Control: Architectural Issues and Performance," in *Proc. of ACM SIGCOMM 2000*, 28 Aug. – 1. Sept., 2000, pp. 57-69.

[4] Gy. Dán, V. Fodor, "On The Efficiency of Shaping Live Video Streams", *SPECTS'02*, July 14–18, 2002.

[5] V. Elek, G. Karlsson, R. Ronngren, "Admission Control Based on End-to-end Measurements," in *Proc. of IEEE INFOCOM 2000,*, March 26–30, 2000, pp. 623–630.

[6] S. S. Lam, S Chow, D. K. Y. Yau, "An Algorithm for Lossless Smoothing of MPEG video," *ACM SIGCOMM Computer Communication Review*, vol. 24, no. 4, Oct. 1994, pp. 281-293

[7] A. Mankin, F. Baker, B. Braden, S. Bradner, M. O'dell, A. Romanow, A. Weiurib, L.Zhang, "Resource ReSerVation Protocol - Version 1 Applicability Statement Some Guidelines on Deployment. RFC 2208," September 1997.

[8] M. Reisslein, K. W. Ross, S. Rajagopal, "Guaranteeing Statistical QoS to Regulated Traffic: The Multiple Node Case," in Proc. *IEEE Decision & Control'98*, pp. 531-538, 1998.

[9] T. Wu, E. W. Knightly, "Buffering vs. Smoothing for end-to-end QoS: Fundamental issues and comparison," in Proc. *IEEE Performance'99* Aug. 1999.