

EFFICIENT MPEG VIDEO TRAFFIC SHAPING FOR THE NEXT GENERATION INTERNET

Mohammad F. Alam
 Electro-Optics Program
 University of Dayton
 Dayton, Ohio 45469-0245, USA
 ferdous@ieee.org

Mohammed Atiquzzaman
 Department of Electrical and Computer Engineering
 University of Dayton
 Dayton, Ohio 45469-0226, USA
 atiq@ieee.org

Mohammad A. Karim
 Department of Electrical and Computer Engineering
 University of Tennessee
 Knoxville, Tennessee 37996-2100, USA
 karim@utk.edu

Abstract

The Internet Engineering Task Force (IETF) has defined the Guaranteed Service (GS) with firm delay and bandwidth guarantees in the Integrated Services model for supporting real-time services. Before setting up a GS flow, traffic parameters of the source have to be specified in terms of a number of token bucket traffic descriptors. The source must conform to these descriptors by shaping its traffic. In this paper we study a novel traffic shaper for efficient transmission of MPEG video streams over GS. We develop an analytical model of the traffic shaper, and also carry out numerical simulation of the transmission performance to verify the analytical model. Our study also provides a methodology to determine the token bucket traffic descriptors that have to be specified during flow set up.

1. Introduction

The Internet Engineering Task Force (IETF) has defined two new services on the Internet Protocol (IP) networks which are collectively called *Integrated Services*: the Guaranteed Service (GS) and the Controlled Load (CL) Service. The Guaranteed Service [1] provides guarantees of bandwidth as well as end-to-end-delay, while the Controlled Load Service [2] provides a service which closely approximates the behavior visible to applications receiving best-effort service under lightly loaded conditions of the network.

While setting up a flow over GS, network resources have to be reserved. The reservation process is dependent on the characteristics of the traffic to be generated by the end application and the required level of service guarantees. The Motion Pictures Experts Group (MPEG) standard for video coding [3] has received worldwide acceptance for storage and transmission of compressed video. A traffic shaping mechanism is required at the transmitting end for MPEG video transmission over GS so that the sender can transmit traffic bursts up to its allowed limit which is negotiated at flow setup time while ensuring that the traffic is conformant to the negotiated traffic specification. Although a token bucket itself can be used at the source to shape MPEG traffic [4], the token bucket algorithm is a complicated algorithm from implementation point of view [5]. Thus, a simpler traffic shaper is desirable that will ensure that the traffic is conformant to the token bucket traffic policing elements in the network.

Most of the previous studies on MPEG transmission are based on the Asynchronous Transfer Mode (ATM) networks [6-11], although some studies on Transmission Control Protocol (TCP) based IP networks have been reported [12]. The IETF has specified the connectionless User Datagram Protocol (UDP) as the layer-4 protocol for fast transport of MPEG GS streams instead of the reliable but slow connection-oriented TCP protocol.

Recent studies on video transmission over GS have assumed either a constant-bit-rate traffic [13] or a simple leaky-bucket traffic shaping [14] at the transmitting end. Depending on the shaper parameters chosen, such leaky-bucket or constant-bit-rate traffic shapers either fail to utilize the full burst-handling capability of GS, or can not ensure that the IP packets sent out to the Internet is conformant to the pre-negotiated traffic specifications. A leaky-bucket traffic shaper reduces or removes the burstiness of the MPEG data stream, and may introduce large amount of delay, which can result in a large buffer requirement at the traffic shaper.

In this paper, we propose a novel traffic shaper of IP packets at the MPEG video source. Methods for choosing parameters for the traffic shaper are discussed, and then the effect of the shaper parameters on delay and buffer requirement are studied. We develop an analytical model for studying the effect of the traffic shaper parameters on delay and buffer requirement. We verify the model by comparing the delay and buffer requirements obtained from simulation of several different MPEG video sequences. The simulation results show excellent agreement with the theoretical model. Relationship between the traffic shaper parameters and traffic specifications are also studied.

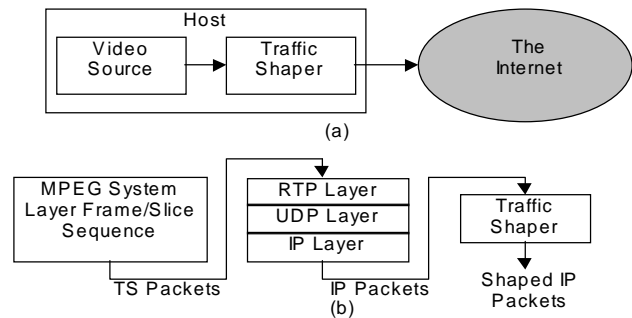


Figure 1: (a) The system block diagram, and (b) the network protocol layers for MPEG transmission over the Guaranteed Service IP network.

2. Guaranteed Service and MPEG Transmission

A real-time stream using any of the Integrated Services on the Internet is established when an end application reserves bandwidth from the network using the Resource ReSerVation Protocol (RSVP) [15]. IETF has also defined the real-time transport protocol (RTP) for delivering real-time traffic on IP networks with timing information for reconstruction as well as feedback on reception quality. The encapsulation mechanism of MPEG-1 and MPEG-2 video streams in IP packets using the RTP protocol has also been specified [16] by IETF. Figure 1(a) shows the system block diagram for MPEG video transmission, while Fig. 1(b) shows the protocol stack for transporting MPEG video transport stream (TS) packets over an IP network. Figure 2 shows a typical IP packet structure for RTP-based applications.

IP Header IPv4=20 Bytes IPv6=40 Bytes	UDP Header 8 Bytes	RTP Header 12 Bytes	Payload (Variable)
---	-----------------------	------------------------	-----------------------

Figure 2: An IP packet structure using the RTP protocol for real-time MPEG stream transmission.

While setting up a flow over GS, traffic specifications (TSPEC) for the flow has to be specified in advance in terms of a number of token bucket traffic descriptors. The traffic descriptors include: bucket size b , the average (or token generation) rate r and the peak bucket rate p . Figure 3 shows the token bucket parameters schematically.

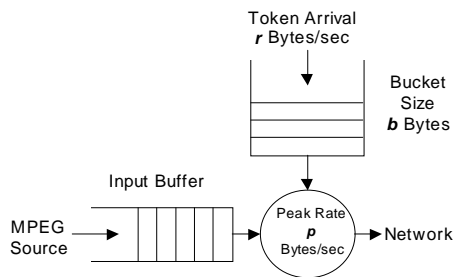


Figure 3: Schematic diagram of the token bucket traffic policing arrangement and the TSpec token bucket parameters (r , b , p) specified in the Guaranteed Service.

An MPEG video sequence comprises of three types of frames: Intra (I), Bi-directional (B) and Predictive (P). At the beginning of a group of pictures (GOP), an I-frame is transmitted. After the I-frame, a number of B- and P-frames are transmitted. A typical sequence of frames, for example, is IBB-PBB-PBB-PBB. I-frames are much larger in data content than other frames. For the initial I-frame in a GOP, usually a large burst is present at the beginning of every GOP. Thus, it is logical to design a traffic shaper that will transmit a burst of high data rate at the beginning of every GOP period. We propose a traffic shaper that transmits at the beginning of a GOP at a high data rate P for a fraction of the GOP period (a burst) for rapidly transmitting the I-frame, followed by a period of data transmission at a low rate Q . Figure 4(a) is the state transition diagram, and Fig. 4(b) is the system block diagram of the traffic shaper. A timer is required for switching the output data rate between P and Q . Figure 4(c) shows the data output rate as a function of time.

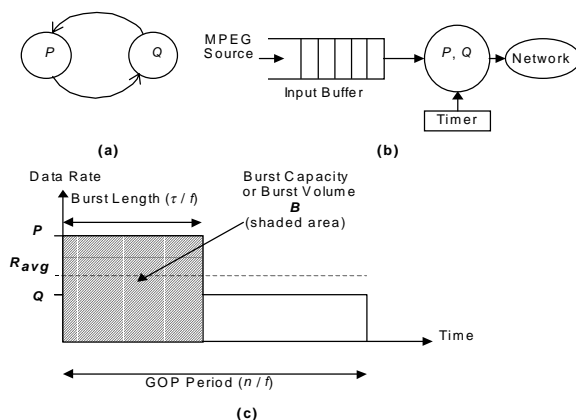


Figure 4: (a) The state transition diagram, (b) the schematic diagram, and (c) the time versus data-rate (shaping function) of the traffic shaper.

Our proposed traffic shaper fully utilizes the traffic burst transmission guarantees provided by GS on the Internet, and ensures that the generated traffic is conformant. As far as the authors are aware, *no other simple traffic shaping scheme has been proposed which utilizes the burst-handling capability of the Guaranteed Service.*

3. Analytical Model

MPEG frames are arranged into Group of Pictures (GOP). The GOP structure is usually described as $MmNn$ where n is the total number of frames in a group of pictures and m is the I-P or P-P

frame interval. For example, an $M3N12$ sequence is IBB-PBB-PBB-PBB. We assume that the I-frame data rate is R_I , the B-frame data rate is R_B , and the P-frame data rate is R_P . The average data rate for an $MmNn$ GOP sequence, R_{avg} , can be written as

$$R_{avg} = \frac{R_I + (n - n/m)R_B + (n/m - 1)R_P}{n} \quad (1)$$

Let us assume that the number of frames per second is f , the peak rate of transmission is P , and the burst volume, which is the area under the peak-rate transmission (at rate P) in the shaper, is B (see Fig. 4). Also, we define the *normalized burst length* τ as:

$$\tau = \frac{\text{Burst length}}{\text{GOP period}} = \frac{B/P}{n/f} \quad (2)$$

which is the fraction of time the traffic shaper transmits at the peak rate compared to the time taken to transmit one GOP. In addition, we also define a *normalized burst volume* β as:

$$\beta = \frac{\text{Burst volume}}{\text{I frame data size}} = \frac{B}{R_I/f} \quad (3)$$

which is the ratio of the burst volume to an I-frame data size in the MPEG video sequence.

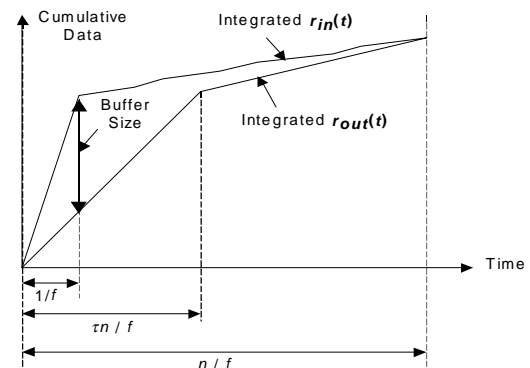


Figure 5: Input data rate $r_{in}(t)$ and output data rate $r_{out}(t)$ integrated to find the maximum difference between the input cumulative data and output cumulative data for calculating the maximum buffer size requirement of the traffic shaper in a GOP.

3.1 Calculation of Buffer Size

Let us assume that $r_{in}(t)$ is the input instantaneous data rate, and $r_{out}(t)$ is the instantaneous output data rate at time t (see Fig. 5). The buffer requirement S is given by the maximum value of the difference between the total data input and output within one GOP period as

$$S = \max \left[\int_0^{\tau n/f} [r_{in}(t) - r_{out}(t)] dt \right] \quad (4)$$

The buffer size S can be evaluated from Eq. (4), as explained in the Appendix, for the case when the burst length is at least equal to the I-frame transmission time, i.e., $\tau > 1/n$ as

$$S = K \frac{R_I}{f} \left[1 - \frac{\beta}{\tau n} \right] \quad (5)$$

where K is a factor which represents the maximum-to-average buffer size ratio.

3.2 Calculation of Delay

The delay experienced by each frame within the same GOP in the shaper is maximum for the I-frame due to its large size. Assuming that the transmission of the I-frame begins exactly at the same instant as the shaper starts its transmission at the peak rate P , the maximum delay experienced by the frames is given by the delay experienced by the I-frame. The maximum delay t_{delay} of the

frames is calculated by the elapsed time between the completion of the arrival and the completion of the departure of an I-frame at the traffic shaper. As explained in the Appendix, this time difference is the maximum delay experienced by all the frames and is defined as the delay through the traffic shaper and is given by

$$t_{delay} = \begin{cases} \frac{1}{f} \left[\tau n - 1 + R_f \frac{n(1-\beta)(1-\tau)}{nR_{avg} - \beta R_f} \right] & \beta < 1 \\ \frac{1}{f} \left[\frac{\tau n}{\beta} - 1 \right] & \beta \geq 1. \end{cases} \quad (6)$$

4. Relationship Between Shaper Parameters And TSpec

The traffic generated by the proposed traffic shaper is to be policed by a token bucket traffic policing element in the network. The geometrical shape of the traffic shaper output can be used for determining the appropriate token bucket parameters (b, r, p) as follows:

$$p = P \quad (7)$$

$$b = (P - R_{avg})\tau / f = B \frac{P - R_{avg}}{P} \quad (8)$$

5. Simulation

A simulation program was written which simulated the behavior of a first-in-first-out (FIFO) queue with output data transmission rate control according to the characteristics of our proposed shaper. MPEG trace data from several different movie clips were used [20] as input to the simulator. The program takes as input β and τ . The simulation program calculates the delay experienced by each frame and also keeps track of instantaneous buffer occupancy as it runs. Statistics on delay and buffer queue length are gathered and saved. The simulation is performed for different ranges of values of β and τ .

6. Results

First, results from the analytical model developed in Section 3 are presented. In Fig. 6, the average delay experienced by a GOP as a function of the burst volume as given by Eq. (6) is shown in broken lines. The average delay is normalized to unit GOP period and the burst volume is normalized to the I-frame data size. The normalized delay is plotted for $0.2 < \tau < 1.0$.

For a constant τ , the delay is reduced for increased burst volume. For constant burst volume, increasing τ also increases the delay considerably. $\tau = 1.0$ represents the case where the traffic shaper transmits at a constant data rate equal to the average data rate for the full GOP period. $\tau = 1.0$ represents a constant-bit-rate transmission as well as a leaky-bucket traffic shaper. This traffic shaping delay has been neglected in [14] although the delay experienced by an MPEG stream through a leaky-bucket traffic shaper is comparable to the end-to-end delay. Figure 6 also shows that significant reduction in traffic shaping delay can be obtained by choosing a lower value of τ .

Figure 7 shows the buffer size requirement as a function of the normalized burst volume as given by Eq. (5) in broken lines. The buffer size is also normalized to the I-frame data size. For Figs. 6 and 7, the analytical model parameters are taken from the DINO sequence, a sequence from the movie Jurassic Park. In Fig. 7, for a constant τ , the buffer size requirement is reduced for increased burst volume of the shaper. For a constant burst volume, decreasing the value of τ also decreases buffer size requirement because of the faster initial data rate available when τ is decreased.

Next, simulation results are presented in Figs. 6 and 7 where the normalized average delay and the normalized maximum buffer requirement, respectively, are shown for the DINO sequence in

solid lines. The buffer requirement is normalized to the average I-frame data size. Here τ varies between 0.2 and 1.0.

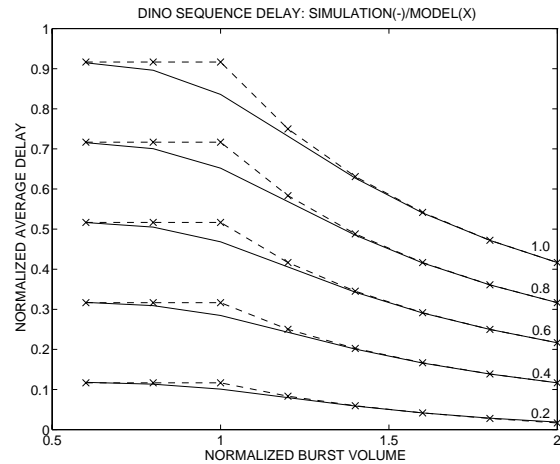


Figure 6: Normalized average delay for a traffic shaper for the DINO sequence as a function of the normalized burst volume. The solid lines represent simulation results while the broken lines (with \times mark) represent analytical results from Section 3.2. Five different curves are plotted for values of τ in the range $0.2 < \tau < 1.0$.

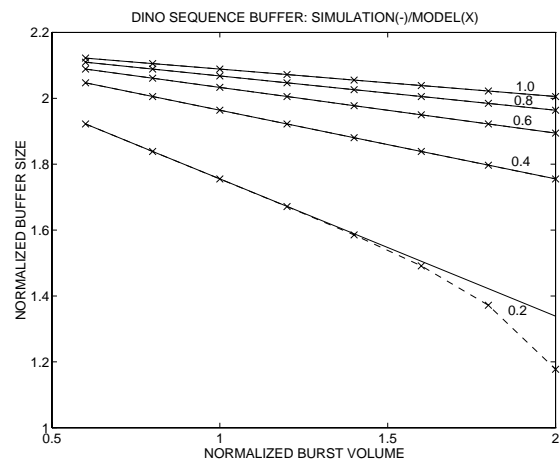


Figure 7: Normalized buffer size for a traffic shaper for the DINO sequence as a function of the normalized burst volume. The solid lines represent simulation results while the broken lines (with \times mark) represent analytical results from Section 3.1. Five different curves are plotted for values of τ in the range $0.2 < \tau < 1.0$. The solid and broken lines overlap completely for $\tau = 0.4, 0.6, 0.8$ and 1.0 .

We observe that there is excellent agreement between the analytical model and the simulation results in Fig. 6 for the normalized average delay. The analytical model result for delay represents the delay experienced by a GOP of average GOP size. Thus, the statistical variations of the GOP sizes and frame sizes cause the simulation result curves to become smoother than the analytical results. This can explain the deviation of the analytical results from simulation results around $\beta = 1$. For buffer size, excellent agreement between analytical results and numerical simulation is observed in Fig. 7.

In Fig. 8, we compare the delay as a function of the burst volume for four different video sequences: DINO (Jurassic Park),

STAR2 (Star Wars), TERM (Terminator 2) and MTV. The burst length for all the sequences has been set for $\tau = 0.2$. Average delay in each sequence is normalized to unit GOP period while the burst volume is normalized to the average I-frame data size of each of the respective sequences. The close match between four different types of video sequences in Fig. 8 suggests that normalized delay and normalized burst volume are related by a curve which is almost identical for a wide range of video sequences.

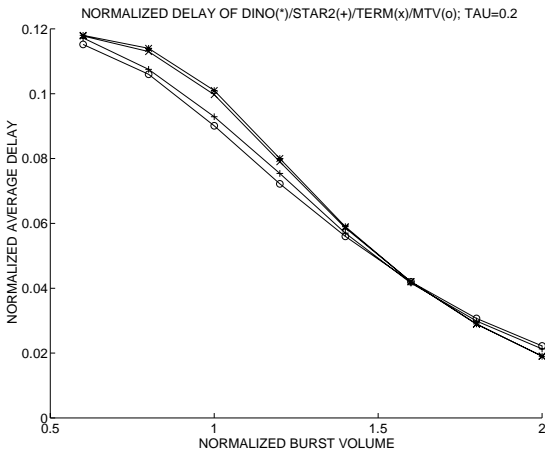


Figure 8: Normalized average delay of four different video sequences as a function of normalized burst volume for the case $\tau = 0.2$. The sequences are DINO(*), STAR2(+), TERM(x), and MTV(o).

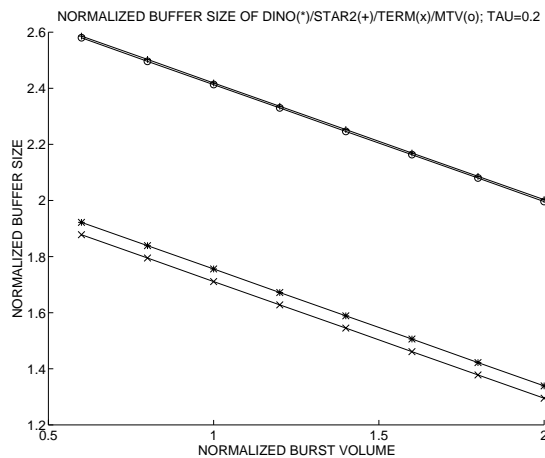


Figure 9: Normalized buffer size of four different video sequences as a function of normalized burst volume for the case $\tau = 0.2$. The sequences are DINO(*), STAR2(+), TERM(x), and MTV(o).

In Fig. 9, we compare the normalized buffer requirement for the same four video sequences as a function of the normalized burst volume where $\tau = 0.2$. Wide variation in the normalized buffer requirement is observed for different types of video sequences. In Fig. 10, we compare the average delay for the same video sequences with $\tau = 0.6$ and observe the same close match between all of the curves. Compared to Fig. 8, increase in delay is observed due to the higher value of τ chosen. In Fig. 11, the buffer requirement for the four video sequences are compared for $\tau = 0.6$. Again, wide variation in the normalized buffer size is observed for different video sequences, suggesting that specifying buffer requirement requires careful design.

For a given delay and normalized burst length τ , the burst volume B and the TSpec bucket size b can be calculated. From these values, the peak rate p can also be calculated using Eqs. (7)-(8). The token generation rate is calculated from the maximum GOP data size. The maximum GOP data size and the average I-frame data sizes are determined by the simulation program. For the four sequences, these values are listed in Table I.

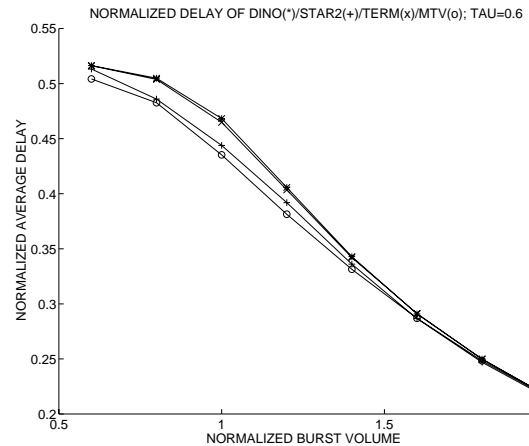


Figure 10: Normalized average delay of four different video sequences as a function of normalized burst volume for the case $\tau = 0.6$. The sequences are DINO(*), STAR2(+), TERM(x), and MTV(o).

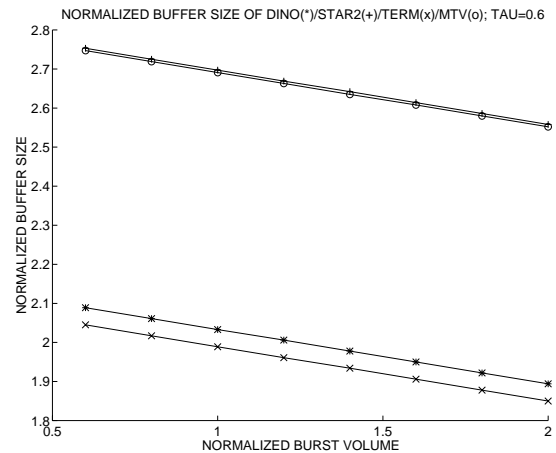


Figure 11: Normalized buffer size of four different video sequences as a function of normalized burst volume for the case $\tau = 0.6$. The sequences are DINO(*), STAR2(+), TERM(x), and MTV(o).

Table I. Normalization factors for the four different video sequences.

Sequence	Largest GOP data size (bits)	Average I-frame data size (bits)
DINO	10166368	55076
TERM	8068200	37388
MTV	18231552	69862
STAR2	7052696	44012

7. Conclusion

In this paper, we have analyzed the characteristics of a simple traffic shaper for efficient MPEG video transmission over the Guaranteed Service of the next generation Internet. The proposed traffic shaper is simpler than the complicated token bucket traffic

shaper. Both analytical modeling and numerical simulation were used for studying the proposed traffic shaper. The analytical results are in excellent agreement with simulation results. Our study shows that the traffic shaper can keep the traffic conformant to the negotiated traffic specification while utilizing the full burst-handling capability of the Guaranteed Service. For larger burst volumes, the average delay and buffer size are reduced. Reducing the burst length while keeping the burst volume constant reduces the delay and buffer size requirement for the shaper.

Our study also shows that the average delay as a function of the burst volume follows almost identical curves for different video sequences. This general characteristic of the shaper can be utilized for choosing appropriate traffic shaper parameters. The buffer size shows wide variations for different types of video sequences.

These results can be utilized for specifying a suitable TSpec parameter while setting up a Guaranteed Service flow for a pre-specified average delay through the shaper.

Appendix

Let us assume that the traffic shaper transmits at an average rate R_{avg} , and the data transmission rates are P and Q , respectively, during high and low data transmission rates (see Fig. 4). Each GOP period is (n/f) seconds long and during this period, transmission at rate P continues for (B/P) seconds and at rate Q continues for $(n/f)-(B/P)$ seconds.

The total data transmitted during a GOP period can be written as,

$$R_{avg} \frac{n}{f} = B + Q \left(\frac{n}{f} - \frac{B}{P} \right) \quad (A.1)$$

from which we get

$$Q = \frac{R_{avg} - Bf/n}{1 - B/P} \quad (A.2)$$

Now, eliminating B from Eqs. (2) and (3), we get,

$$P = \frac{\beta R_l}{\tau n} \quad (A.3)$$

From Fig. 5, the buffer requirement is maximum at time $t = 1/f$, provided that $R_l > P > R_p > R_b$. Under this assumption, the required buffer size can be written as,

$$S = (R_l - P) / f = (R_l / f) \left[1 - \frac{\beta}{\tau n} \right] \quad (A.4)$$

where Eq. (A.3) is used to arrive at the last step. Here, the buffer size represents the average buffer size. However, we need to multiply this buffer size by a constant K to arrive at Eq. (5) such that K represents the ratio of the maximum-to-average buffer size. The value of K is computed from MPEG trace data.

For $\beta \geq 1$, an I-frame is completely contained inside the burst volume of the shaper. A complete I-frame arrives at the input buffer of the shaper at time $t = 1/f$, and leaves the shaper at time $t = R_l / fP$. Hence, the delay, which is the time difference between the arrival and the departure of the end of the I-frame, is given by

$$t_{delay} = \frac{R_l}{fP} - \frac{1}{f} = \frac{1}{f} \left[\frac{\tau n}{\beta} - 1 \right] \quad (\beta \geq 1) \quad (A.5)$$

where Eq. (A.3) is used for eliminating R_l . When $\beta < 1$, a complete I-frame is not contained inside the burst volume of the shaper. The burst length (B/P) can also be written as $(\tau n/f)$ from definition of τ (Eq. 2). Let us assume that the I-frame is transmitted at rate P for the burst length, and also transmission at rate Q is required for an additional duration of time t' . Thus, the total amount of data transmitted is equal to the data size of the I-frame, and can be written as the sum of the data transmitted at rate P and at rate Q as,

$$P \frac{\tau n}{f} + t' \frac{(nR_{avg}/f) - B}{(n/f) - (B/P)} = \frac{R_l}{f} \quad (A.6)$$

from which we get

$$t' = \frac{R_l}{f} (1 - \beta) \frac{n(1 - \tau)}{nR_{avg} - \beta R_l} \quad (A.7)$$

where B and P are eliminated using Eqs. (A.3) and (3). The total delay for the case $\beta < 1$ can thus be written as the burst length plus t' minus the time of arrival of the I-frame as,

$$t_{delay} = \left(\frac{\tau n}{f} + t' \right) - \frac{1}{f} \quad (\beta < 1) \quad (A.8)$$

which can be written, using (A.7) as,

$$t_{delay} = \frac{1}{f} \left[\tau n - 1 + R_l \frac{n(1 - \beta)(1 - \tau)}{nR_{avg} - \beta R_l} \right] \quad (\beta < 1). \quad (A.9)$$

References

- [1] S. Shenker, C. Partridge, and R. Guerin, "Specification of guaranteed quality of service," Internet Engineering Task Force, RFC 2212, September 1997.
- [2] J. Wroclawski, "Specification of the controlled-load network element service," Internet Engineering Task Force, RFC 2211, September 1997.
- [3] The MPEG family of standards includes MPEG-1, MPEG-2 and upcoming MPEG-4, formally known as ISO/IEC-11172, ISO/IEC-13818 and ISO/IEC-14496.
- [4] M. F. Alam, M. Atiquzzaman, and M. A. Karim, "Effects of Source Traffic Shaping on MPEG Video Transmission over Next Generation IP Networks," *1999 International Conference on Computer, Communications, and Networks*, Boston, MA, USA, October 11-13, 1999.
- [5] C. Partridge, "Token bucket with leaky bucket rate control," *Gigabit Networking*, Addison-Wesley, Reading, MA, pp. 262-263, 1994.
- [6] B. Zheng and M. Atiquzzaman, "Traffic Management of Multimedia over ATM Networks," *IEEE Communications Magazine*, vol. 37, no. 1, pp. 33-38, January 1999.
- [7] J. S. Kim and J. K. Kim, "Adaptive traffic smoothing for live VBR MPEG video service," *Computer Communications*, vol. 21, no. 7, pp. 644-653, 1998.
- [8] W. Zhu, Y. Wang, and Y. Q. Zhang, "Jitter smoothing and traffic modeling for MPEG-2 video transport over ATM networks," *International Journal of Imaging Systems and Technology*, vol. 9, no. 5, pp. 332-339, 1998.
- [9] J. D. Salehi, Z. L. Zhang, J. Kurose, and D. Towsley, "Supporting stored video: Reducing rate variability and end-to-end resource requirements through optimal smoothing," *IEEE/ACM Trans. Networking*, vol. 6, no. 4, pp. 397-410, 1998.
- [10] B. Zheng and M. Atiquzzaman, "Multimedia Over High Speed Networks: Reducing Network Requirements With Fast Buffer Fillup," *IEEE GLOBECOM'98*, Sydney, Australia, November 8-12, 1998.
- [11] P. Cuenca, B. Caminero, A. Garrido, F. Quiles, and L. Orozco-Barbosa, "QoS and statistical multiplexing performance of VBR MPEG-2 video sources over ATM networks," *1998 11th Canadian Conference on Electrical and Computer Engineering*, vol. 1, pp. 33-36, Toronto, Canada, May 24-28, 1998.
- [12] S. Jacobs and A. Eleftheriadis, "Streaming video using dynamic rate shaping and TCP congestion control," *Journal of Visual Communication and Image Representation*, vol. 9, no. 3, pp. 211-222, 1998.
- [13] K. Van der Wal, M. Mandjes, and H. Bastiaansen, "Delay performance analysis of the new internet services with guaranteed QoS," *Proc. IEEE*, vol. 85, no. 12, pp. 1947-1957, December 1997.
- [14] H. Naser and A. Leon-Gracia, "Performance evaluation of MPEG2 video using guaranteed service over IP-ATM networks," *Multimedia Computing and Systems Conference*, Austin, Texas, USA, June 28 - July 1, 1998.
- [15] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource ReSerVation Protocol (RSVP) - Version 1 functional specification," Internet Engineering Task Force, RFC 2205, September 1997.
- [16] D. Hoffman, G. Fernando, V. Goyal, and M. Civanlar, "RTP payload format for MPEG1/MPEG2 video," Internet Engineering Task Force, RFC 2250, January 1998.
- [17] O. Rose, "Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems," Report No. 101, Institute of Computer Science, University of Würzburg, February 1995.