Traffic Shaping Schemes for Improving QoS in Data Networks

Tamer Dag
Department of Computer Engineering
Kadir Has University, Istanbul, Turkey
tamer.dag@khas.edu.tr

Abstract

As the typical characteristics of traffic in data networks are bursty which is difficult to handle, traffic shaping schemes are used for regulating the average rate and burstiness of a flow of data that enters into the network. While smoothing the traffic of time sensitive applications at the user premises can reduce the losses due to buffer overflows within the network, it may have an adverse effect on the losses due to the deadline violations due to the delay introduced by the smoothing process. In this paper, these loss factors are investigated and some simple metrics that capture the potential for buffer overflow and deadline violation losses under a traffic shaping scheme are introduced. Such metrics can be useful for the design of effective traffic shaping schemes that balance the adverse effects of buffer overflow and deadline violation losses.

1 Introduction

The real-time Variable Bit Rate (rt-VBR) applications [1, 2, and 3] such as voice and video require an end-to-end timing relationship among the communicating applications. Usable bandwidth is allocated to a rt-VBR connection according to the associated source traffic descriptors negotiated at the connection set up. Based on these descriptors, resources are allocated to guarantee the QoS demands during the connection's duration.

Factors such as the propagation delay, network utilization, VBR subframe lengths have significant impact on the delivered QoS service and have been considered in the past [4, 5]. In addition to these factors, the way in which traffic is shaped before delivered to the network impacts on the overall performance.

Traffic smoothing reduces the burstiness of the traffic delivered to the network by, in essence, spreading the transmission time of the cells over a transmission interval.

By this way, the induced cell losses due to buffer overflows are reduced as a result of the increased statistical multiplexing gain.

Although this approach can be useful for non-real-time VBR (nrt-VBR) applications which do not have stringent delay requirements, delay sensitive rt-VBR applications may suffer from the smoothing process as a result of the reduction of the delay tolerance of the transmitted cells after the smoothing process. As a result, smoothing would reduce the deadline margin of the cells and lead to an increased cell losses within the network due to deadline violations. This latter effect has been neglected in the past.

In this paper, the rt-VBR traffic model is described along with some traditional traffic shapers. Some metrics capturing the potential for buffer overflow and deadline violation losses are introduced. It is also presented how these metrics can help design traffic shapers that balance effectively the impact of cell losses due to buffer overflows and deadline violations. An example of such traffic shaper - the linear envelope traffic shaper - is introduced and some numerical results illustrating the effectiveness of the proposed traffic shaper design approach are presented.

2 Traffic Shapers for rt-VBR Sources

A VBR source, depending on the coding scheme used [6, 7, 8], generates traffic in a continuous manner at varying rates. Typically, a rt-VBR application generates cells in cycles of variable lengths and the amount of cells generated per cycle varies primarily due to the employed coding scheme. Through the rest of the paper, these cycles will be referred to as subframes.

Let S_k denote the k^{th} subframe of length S_k slots and A_k denote the average cell rate over the given subframe S_k. Assume that for this subframe, the VBR source generates N_k cells as shown in Figure 1. Thus, the average cell rate over a subframe of a VBR source is, $A_k = N_k/S_k$.

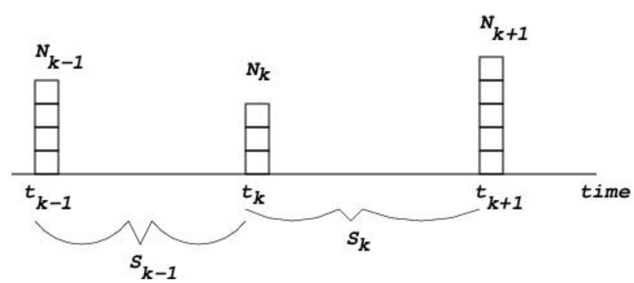

Figure 1: An example of a rt-VBR traffic generation process

Let t_k denote the time instant where the $\boldsymbol{k^{th}}$ VBR subframe begins and $\boldsymbol{D_k}$ denote the deadline of the cells generated at time $\boldsymbol{t_k}$. Thus, the cells generated at the $\boldsymbol{k^{th}}$ subframe should be serviced at their destination by $\boldsymbol{t_k + D_k}$. Otherwise, the deadline of the cells will be violated and they will be dropped.

The level of smoothing that will be applied for the transmission of the cells generated at the subframe boundaries will have impact on both overall network effectiveness and the resulting QoS delivered to the particular rt-VBR application.

Let $\boldsymbol{f(t)}$ describe the shaping function employed at the rt-VBR sources. Thus, $\boldsymbol{f(t)}$ determines how the $\boldsymbol{N_k}$ cells are spread over the subframe of length $\boldsymbol{S_k}$ depending on the deadlines $\boldsymbol{D_k}$.

One way to implement the traffic shaping function $\boldsymbol{f(t)}$ is by using a leaky bucket (LB) [9,10, 11, and 12] scheme. In this scheme, the traffic shaping function $\boldsymbol{f(t)}$ produces tokens and these tokens are placed in a token pool. The rt-VBR source cells generated at subframe boundaries can then be transmitted only if there is a token in the token pool. If the token pool is empty, the rt-VBR source cell has to wait for the generation of another token. Otherwise, it is immediately transmitted and the number of tokens in the token pool is decreased by one.

In the next two subsections, two traffic shapers are described by introducing the appropriate traffic shaping function $\boldsymbol{f(t)}$. The impact on cell losses due to buffer overflows and deadline violations are discussed.

2.1 The One Envelope Traffic Shaping Function

Under the One Envelope traffic shaper, the VBR source cells are transmitted in such a way to maximize their delay tolerances. Thus, if $\boldsymbol{N_k}$ cells are generated at the $\boldsymbol{k^{th}}$ subframe $\boldsymbol{S_k}$, these cells will be transmitted during the first $\boldsymbol{N_k}$ slots of the subframe.

The traffic shaping function $\boldsymbol{f(t)}$ for the one envelope traffic shaper can be formulated as shown below:

$$f(t) = \begin{cases} 1 & t_k \le t \le t_k + N_k \\ 0 & t_k + N_k \le t \le t_{k+1} \end{cases}$$

According to this function - and referring to the LB mechanism - one token is produced for each time slot starting with a subframe boundary until the entire cells generated at that subframe are transmitted.

Under the one envelope traffic shaping, the rt-VBR applications are expected to have minimal losses due to deadline violations, as the VBR source cells are transmitted as quickly as possible after they are generated. Since the VBR sources will be transmitting with their peak rates for the first $\boldsymbol{N_k}$ slots of each subframe $\boldsymbol{S_k}$, the resulting traffic would be very bursty and buffers will tend to fill very quickly when statistical multiplexing is employed. Thus, the cell losses due to buffer overflows are expected to be high under this scheme. The rt-VBR applications are expected to perform well in terms of losses due to the deadline violations under the one envelope as the VBR source cells are transmitted as quickly as possible once they are generated. However, the cell losses due to the buffer overflows may be high when VBR sources are statistically multiplexed.

2.2 The Constant Envelope Traffic Shaping Function

The constant envelope traffic shaper is the traditional traffic shaper used for nrt-VBR applications. The purpose of the constant envelope traffic shaper is to distribute the generated VBR cells uniformly over some time period of $\boldsymbol{D_k}$; $\boldsymbol{D_k}$ does not represent a deadline in the case of nrt-VBR applications but a smoothing horizon. Thus, if $\boldsymbol{N_k}$ cells are generated at the $\boldsymbol{k^{th}}$ subframe $\boldsymbol{S_k}$, one cell will be transmitted for every $\boldsymbol{D_k/N_k}$ slots.

The traffic shaper function $\boldsymbol{f(t)}$ for the constant envelope shaper can be formulated as shown below.

$$f(t) = \begin{cases} \dfrac{N_k}{D_k} & t_k \le t \le t_k + D_k \\ 0 & t_k + D_k \le t \le t_{k+1} \end{cases}$$

On the contrary to the one envelope traffic shaper, the constant envelope traffic shaper is expected to perform well in terms of the buffer overflows within the network due to the extensive smoothing applied. However, if this mechanism is employed for the traffic shaping of rt-VBR applications with deadline $\boldsymbol{D_k}$, the induced losses due to the deadline violations are expected to be high compared to the losses occurring under the one envelope traffic shaper. This is due to the fact that cells can be transmitted very close to their deadlines under the constant envelope scheme.

3 Metrics for Measuring the Efficiency of VBR Traffic Shapers

This section proposes two metrics which can capture the potential of traffic shapers for inducing cell losses due to the buffer overflows and deadline violations. These metrics measure the average remaining deadline and the level of smoothness of the traffic shaped by a traffic shaping function. Using these two metrics, it is possible to compare various traffic shaping functions and gain some insight into the relative values of cell losses expected to be induced under those traffic shaping functions.

3.1 The Remaining Deadline ($\overline{D}$) Metric

The remaining deadline metric, $\overline{D}$ measures the average delay tolerance of a traffic shaper for a given traffic shaping function ***f(t)***. The average transmission delay of the cells, ***T*** for a given value of deadline D_k, subframe length S_k and number of cells to be transmitted N_k can be expressed by the following equation,

$$T = \frac{1}{N_k}\int_0^{S_k} tf(t)dt \ .$$

Thus, the remaining deadline metric D is equal to,

$$\overline{D} = D_k - T = D_k - \frac{1}{N_k}\int_0^{S_k} tf(t)dt$$

Let $\overline{D_1}$ and $\overline{D_c}$ denote the remaining deadline metrics for the one envelope and constant envelope traffic shaping functions, where

$$\overline{D_1} = D_k - \frac{N_k}{2}$$

$$\overline{D_c} = \frac{D_k}{2}$$

Thus, the following relationship holds for the remaining deadline metrics of the traffic shapers considered.

$$\overline{D_1} > \overline{D_c}$$

From the above result, it can be concluded that the potential for cell losses due to deadline violations for the constant envelope traffic shaping function is larger than that of the one envelope traffic shaping function. Using the above argument, it may be possible to compare any two traffic shaping functions in terms of the cell losses due to deadline violations.

3.2 The Smoothness ($\overline{M}$) Metric

Since the constant envelope traffic shaping provides for the smoothest possible traffic profile, smoothness metric could be defined to represent some type of distance between a traffic shaping function from the constant envelope traffic shaping function. For this reason, the smoothness metric $\overline{M}$ is defined to be the L2 distance of a traffic shaping function to the constant envelope traffic shaping function. Thus, the $\overline{M}$ metric can provide some information about the variability of a traffic shaping function. Since the statistical multiplexing gain and the cell losses due to buffer overflows depend on the level of smoothness of the traffic delivered to the network, $\overline{M}$ metric can be a relative measure of the expected cell losses due to buffer overflows.

The smoothness metric $\overline{M}$ of a traffic shaping function ***f(t)*** for a given value of deadline D_k and number of cells to be transmitted N_k is defined by the following equation.

$$\overline{M} = \frac{1}{N_k}\int_0^{D_k} (f(t) - \frac{N_k}{D_k})^2 dt$$

Let $\overline{M_1}$ and $\overline{M_c}$ denote the remaining deadline metrics for the one envelope and constant envelope traffic shaping functions, where

$$\overline{M_1} = 1 - \frac{N_k}{D_k}$$

$$\overline{M_c} = 0$$

The following relationship holds for the smoothness metrics of the traffic shapers considered.

$$\overline{M_1} > \overline{M_c}$$

From the above result, it can be concluded that the potential for cell losses due to buffer overflows for the one envelope traffic shaping function is larger than that of the constant envelope traffic shaping function. Using the above argument, it may be possible to compare any two traffic shaping functions in terms of the cell losses due to the buffer overflows.

4 Linear Envelope Traffic Shaping

The linear envelope traffic shaper - proposed in this section - can be considered as an example traffic shaper which tries to balance the impact of cell losses due to buffer overflows and deadline violations. The linear envelope traffic shaping function allows more smoothing regarding the transmission of the cells than the one envelope traffic shaping function. It also allows for more cell delay tolerance than the constant envelope traffic shaping function.

As a result, the amount of cell losses under the linear envelope traffic shaping function due to buffer overflows is expected to be lower than that under the one envelope traffic shaping function. On the other hand, the amounts of cell losses due to deadline violations are expected to be lower than that under the constant envelope traffic shaping function.

The traffic shaping function ***f(t)*** for the linear envelope shaper can be formulated as shown below.

$$f_l(t) = \begin{cases} \frac{-2N_k}{D_k^2}(t - t_k) & t_k \le t \le t_k + D_k \\ 0 \quad t_k + D_k \le t \le t_{k+1} \end{cases}$$

Using the traffic shaping function $f_l(t)$ of the linear envelope traffic shaper, the remaining deadline ($\overline{D_l}$) and the smoothness ($\overline{M_l}$) metrics can be shown to be equal to,

$$\overline{D_l} = \frac{2D_k}{3}$$

$$\overline{M_l} = \frac{N_k}{3D_k}$$

A comparison of the linear envelope traffic shaping function with one and constant envelope traffic shaping functions can be performed by using the corresponding remaining deadline and smoothness metrics. The following relationships holds for the two metrics considered:

$$\overline{D_1} > \overline{D_l} > \overline{D_c}$$

$$\overline{M_1} > \overline{M_l} > \overline{M_c}$$

Based upon these observations the linear envelope traffic shaper is expected to be more efficient than the constant envelope traffic shaper in terms of the cell losses occurring due to deadline violations. Similarly, it is expected to be more efficient than the one envelope traffic shaper in terms of the cell losses occurring due to buffer overflows. Thus, for the overall efficiency (deadline + buffer overflow), the linear envelope traffic shaper can be a suitable candidate for applications which are delay sensitive and require low cell losses.

5 Numerical Results

In this section, some numerical results are presented in order to illustrate the impacts of various VBR source traffic shapers on cell losses due to buffer overflows and deadline violations.

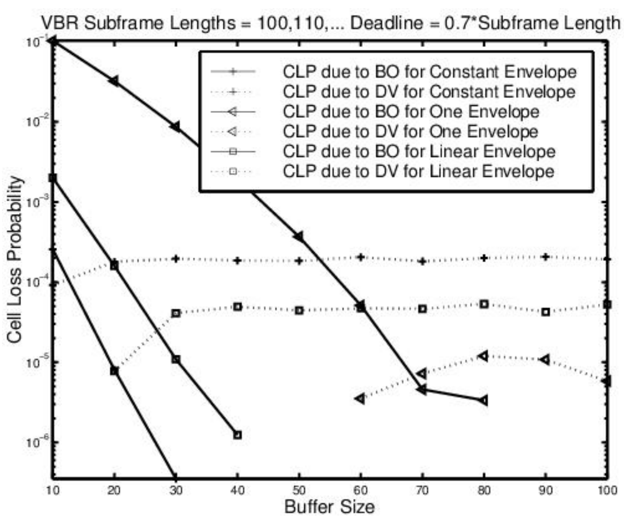

Figure 2: BO and DV CLPs as a function of buffer size for VBR subframe lengths = 100,110... 190 and deadline = 0.7*subframe length

For the simulations, a transmission link shared by 10 VBR sources is implemented. For each VBR source, a fixed subframe length S_k is used. The average transmission rates of each VBR source for a subframe of length S_k is 0.09, 0.06 or 0 cells/slot.

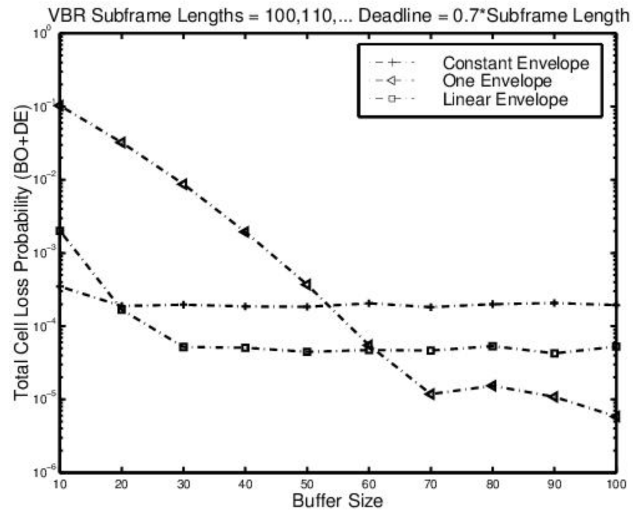

Figure 3: Total CLPs as a function of buffer size for VBR subframe lengths = 100,110... 190 and deadline = 0.7*subframe length

Figure 2 presents the cell loss probabilities due to buffer overflows (BO) and deadline violations (DV) as a function of the buffer size. The VBR subframe lengths for each source are constant with the following values. S_k=100 slots for source 1, S_k=110 slots for source 2... S_k=190 slots for source 10. For each source, the cell deadlines are assumed to be equal to 70% of its corresponding subframe length, that is $D_k = 0.7S_k$.

As expected, under all traffic shapers, the cell losses due to buffer overflows decrease as the buffer size increases. However, this is not true for the cell losses due to deadline violations. As the buffer size increases, the VBR cells tend to spend more time in the buffer which in turn increases the losses due to deadline violations. The one envelope traffic shaper results in the lowest number of cell losses due to deadline violations, while the constant envelope results in the highest number of cell losses. The reverse is true for the cells losses occurring due to buffer overflows. In terms of the efficiency, the linear envelope traffic shaper is in the middle both for losses due to buffer overflow and deadline violations.

Figure 3 presents the total cell loss probabilities (cell loss probabilities due to buffer overflows and deadline violations) as a function of buffer size. Although, the one and constant envelope traffic shapers are more efficient in terms of cell losses due to deadline violations and buffer overflow respectively, the linear envelope traffic shaper seems to be the optimal case for buffer sizes between 20 and 60 cells. Since for very large buffer sizes, the losses due to buffer overflows become insignificant compared to the losses due to deadline violations, the one envelope scheme becomes the optimal scheme. Similarly, for small buffer sizes, the cell losses due to buffer overflows dominate the cell losses due to the deadline violations and the constant envelope scheme becomes optimal scheme.

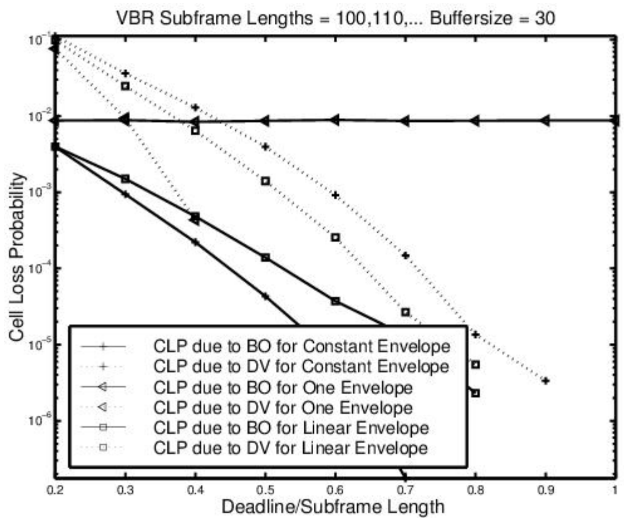

Figure 4: BO and DV CLPs as a function of deadline duration for VBR subframe lengths = 50, 60... 140 and buffer size = 30

Figure 4 illustrates the cell loss probabilities as a function of deadline duration for VBR subframe lengths of ***Sk***=50 slots for source 1, ***Sk***=60 slots for source 2... ***Sk***=140 slots for source 10. As expected, when the deadline durations increase, the cell losses due to deadline violations decrease. Except for the one envelope traffic shaper, increasing the deadline duration decreases the cell losses due to buffer overflows. This is due to the increased statistical multiplexing gain with increased deadline durations. As cell transmissions spread over a longer period of time for the constant and linear envelope traffic shapers by increasing deadline durations, the cell losses due to buffer overflows decrease.

Due to the constant amount of cell losses for different values of deadline durations under the one envelope traffic shaper, the linear envelope traffic shaper becomes the optimal scheme when the total cell losses are considered as shown in Figure 5.

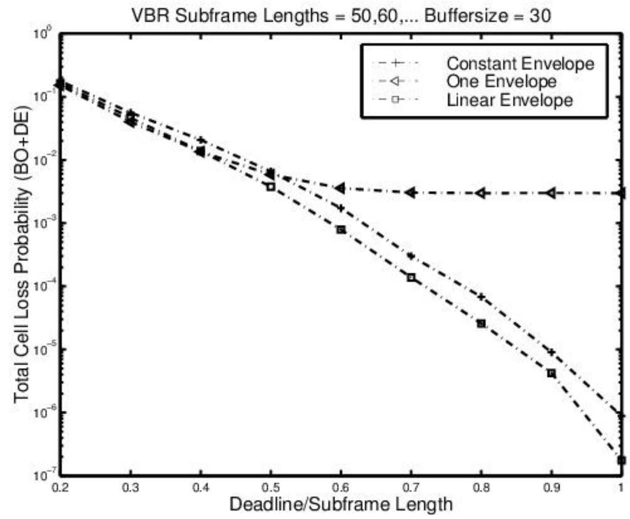

Figure 5: Total CLPs as a function of deadline duration for VBR subframe lengths = 50, 60... 140 and buffer size = 30

6 Conclusions

In this paper, the impact of VBR traffic shapers on the cell losses due to buffer overflows and deadline violations is considered and remaining deadline and smoothness metrics are introduced for a comparison of the effectiveness of the traffic shapers.

Most of the network parameters have significant impact on the performance of the applications. Since rt-VBR applications are delay sensitive and require very low cell losses, their performance highly depends on these parameters such as the utilization of the network, the size of the buffers. In addition to such parameters, the VBR source transmission schemes affect their performance as well. While transmitting cells in such a way to maximize their remaining deadlines can be quite efficient in terms of cell losses due to deadline violations, it increases the cell losses due to buffer overflows. The reverse argument is true for a traffic shaper which tries to transmit the source cells as smooth as possible.

As an example, a linear envelope traffic shaper which balances the adverse effects of buffer overflow and deadline violations is proposed in this paper. It is observed that the linear envelope traffic shaper becomes a very efficient scheme when the overall performance, thus the cell losses due to buffer overflows and deadline violations, are considered.

In order to examine various traffic shapers in terms of buffer overflows and deadline violations, the proposed metrics can be used for a general idea. By the simple interpretation of the traffic shaping functions, the remaining deadline metric and smoothness metric can be calculated and new token generator functions according to the specific needs of the applications can be produced.

References

[1] ATM Forum, "Traffic Management Specification Version 4.0", ATM Forum/95-0013R10, Feb. 1995.

[2] Raif O. Onvural, "Asynchronous Transfer Mode Networks (Performance Issues)", Artech House, Second Edition, 1995.

[3] Mischwa Schwarz, "Broadband Integrated Networks", Prentice Hall, 1996.

[4] T. Dag and I. Stavrakakis, "Study of the Impact of Network and Source Time Scales on Feedback Based Flow Control", SPIE'96, Boston, Nov. 1996.

[5] T. Dag and I. Stavrakakis, "Evaluation of ABR Traffic Management Under Various System Time Scales", Computer Networks and ISDN, pp. 2103-2117, Feb. 1998.

[6] T. Chen and T. R. Hsing, "Digital Coding Techniques for Visual Communications," J. Visual Comm. and Image Representation, Vol. 2-1, pp. 1-16, 1991.

[7] T. Chen, "Adaptive Transform Coding via Quadtree-Based Variable Block Size DCT," IEEE Proc. ICASSP, pp. 1858-1867, Glasgow, Scotland, 1989.

[8] M. Kaneko, Y. Hatori, A. Koike, "Improvement of TransformCoding Algorithm for Motion-Compensated Interframe Prediction Errors - DCT/SQ Coding," IEEE JSAC, Vol. 5, pp. 1068-1078, 1987.

[9] H. Ahmadi, R. Guerin and K. Sohraby, "Analysis of a rate-based access control mechanism for high speed networks", IEEE Globecom'90, 1990.

[10] K. Sohrady and M. Sidi, "On the performance of bursty and correlated sources subject to Leaky Bucket rate-based access control scheme",IEEE Infocom'91, 1991.

[11] E. Rathgeb, "Modeling and performance comparisons at policing mechanisms for ATM networks",IEEE J. Sel. Areas Commun. SAC-9, Apr. 1991.

[12] M. Butto, E. Cavallero and A. Tonietti, "Effectiveness of the Leaky Bucket policing mechanism in ATM networks",IEEE J. Sel. Areas Commun. SAC-9, Apr. 1991.