

Traffic Reshaping in Packet-Switched Virtual-Circuit Fixed-Packet Networks

Sergey Gorinsky* Sanjoy Baruah*[†] Alexander Stoyenko*

Real-Time Computing Laboratory
Department of Computer and Information Science
New Jersey Institute of Technology
Newark, NJ 07102, USA
{gorinsky,sanjoy,alex}@rtlab.njit.edu

Abstract

Impact of traffic reshaping on providing deterministic guarantees of timely data delivery in packet-switched virtual-circuit fixed-packet networks of an arbitrary topology is investigated. Two types of traffic smoothing are studied: global reshaping (when traffic on all network connections is smoothed) and local reshaping (when the traffic specification is changed only for a single connection). The paper demonstrates under which conditions and to which degree reshaping should be applied and derives optimal values of the traffic model parameters.

1 Introduction

As multimedia network applications are quickly growing in their variety and resource requirements, timely data delivery in computer networks gains larger and larger importance. Since even an insignificant loss of transferred information can be unacceptable for such applications, providing deterministic guarantees of network behavior became a subject of intensive research [6]. The general solution is to use connection admission control: based on the user descriptions of traffic on the new and already established connections, the network decides whether it is able to provide requested guarantees to all the connections; if it is not, then the new connection is rejected. It has been noticed that presence of bursty traffic is disadvantageous from timeliness guarantees' point of view. This paper investigates *reshaping* of traffic inside the network and tries to answer the question whether making traffic smoother to conform to another traffic specification can yield better delay bounds. Knightly and Rossaro [5] show that reshaping cannot improve delays for one-hop connections, but are likely be beneficial for multi-hop connections. Georgiadis et al [2] prove that reshaping allows Earliest Deadline First (EDF) scheduling discipline to provide better performance than Generalized Processor Sharing (GPS). Exact analysis is usually very complicated even for relatively simple traffic models, and attempts to experimentally study an impact of reshaping are being undertaken [4].

This paper examines traffic reshaping in packet-switched virtual-circuit fixed-packet networks and is organized as follows. Section 2 describes the network model. Global and local reshaping are investigated in Sections 3 and 4 correspondingly. The results are summarized in Section 5.

*Real-Time Computing Laboratory, Department of Computer and Information Science, New Jersey Institute of Technology, Newark, NJ 07102, USA. Research is supported in part by the U.S. ONR Grant N00014-93-1-1047 and by the NSF Grant CCR-9402827.

[†]Department of Computer Science and Electrical Engineering, The University of Vermont, Burlington, VT 05405, USA. Research is supported in part by the NSF Grant CCR-9596282. Dr. Baruah is on leave from the Real-Time Computing Laboratory at NJIT.

2 The Network Model

Packet-switched virtual-circuit fixed-packet networks of an arbitrary topology are studied in this paper. Our model can be viewed as a generalization of ATM (Asynchronous Transfer Mode) networks. Network nodes are interconnected by links. Different links may be characterized by different bandwidths. Point-to-point communication is supported by means of connections: data are transferred from a source node to a destination node along some fixed route in the network. The leaky bucket model is used for description of traffic on all connections. Each connection is characterized by a requested end-to-end delay bound D and a pair of parameters (σ, ρ) : a maximum burst of size σ is allowed over the maximum average rate ρ . Connection establishment takes place before transmission of user's data and succeeds only if requested delay bounds can be provided. Traffic entering a node is demultiplexed and shaped to conform the connection specification, and each packet is transferred to the queue of an appropriate output link (or an application process at the destination node). Every queue is associated with a First-Come First-Served (FCFS) scheduler which decides in which order packets are sent over the link. Using the same approach as in [2, 7], one can easily prove that shaping does not increase end-to-end delay bounds; thus, we do not consider shaping delays in the analysis below. As a general rule, the switching delay (time to transfer a packet from an input link to an output link queue) is negligible in comparison with the queueing delay (waiting time in the queue) and the transmission delay (transmission time of the packet) [1, 8]. In this study, we therefore assume that delay experienced by a packet at a node comprises only queueing and transmission delays. Packets are assumed to arrive at a queue instantaneously, and this arrival is aligned to the start of transmission of the next packet. In practice, it means that a packet can experience an additional (up to time taken by one packet to be transmitted over the link) *alignment* delay at a node. For instance, if packets are 53 bytes long and the link bandwidth is 155 Mb/s, then the alignment delay equals to 2.7 μ s. Since alignment delays are not affected by traffic reshaping, we do not take them into account in this paper.

3 Global reshaping

First, let us consider a simplified model in which all connections are characterized by the same burstiness σ .

According to the schedulability conditions given in [3], each connection employing link k is guaranteed the following minimum delay bound d_k of sending data over the link:

$$d_k = \frac{1}{l_k} \sum_{m=1}^{M_k} \sigma = \frac{M_k \sigma}{l_k}$$

where l_k is a bandwidth of link k and M_k is the number of connections using this link.

Let a connection C be characterized by the requested end-to-end delay bound D_c and the maximum average rate ρ_c . Assume that the links traversed by connection C are numbered as 1, 2, 3, ..., $N_c - 1$, N_c . Then the minimum end-to-end delay bound that can be guaranteed to the connection is:

$$d_c = \sum_{k=1}^{N_c} d_k = \left(\sum_{k=1}^{N_c} \frac{M_k}{l_k} \right) \sigma \tag{1}$$

Since the connection has been established,

$$d_c \leq D_c \tag{2}$$

Now let us globally reshape traffic on each connection to decrease its burstiness down to σ' where $\sigma \geq \sigma' \geq 0$. What is a new value of the minimum end-to-end delay bound that can be guaranteed to connection C ?

Smoothing delay at the entrance to the network is bounded by

$$d_{sm} = \frac{\sigma - \sigma'}{\rho_c}$$

Then the minimum end-to-end delay bound for connection C becomes:

$$d_c' = d_{sm} + \left(\sum_{k=1}^{N_c} \frac{M_k}{l_k}\right)\sigma' = \frac{\sigma - \sigma'}{\rho_c} + \left(\sum_{k=1}^{N_c} \frac{M_k}{l_k}\right)\sigma'$$

Taking (1) into account, we can rewrite this expression as:

$$d_c' = d_c + \frac{\sigma - \sigma'}{\rho_c} + \left(\sum_{k=1}^{N_c} \frac{M_k}{l_k}\right)(\sigma' - \sigma) = d_c + (\sigma - \sigma')\left(\frac{1}{\rho_c} - \sum_{k=1}^{N_c} \frac{M_k}{l_k}\right)$$

This result is somewhat surprising since it shows that the degree of reshaping, i.e. parameter $(\sigma - \sigma')$, does not affect the fact whether the reshaping decreases or increases the delay bound; this parameter just determines how much the bound is improved or worsened. The qualitative effect of reshaping is decided by the term $\left(\frac{1}{\rho_c} - \sum_{k=1}^{N_c} \frac{M_k}{l_k}\right)$. If

$$\rho_c \geq \frac{1}{\sum_{k=1}^{N_c} \frac{M_k}{l_k}}, \quad (3)$$

the reshaping is beneficial, and to obtain the minimum possible delay bound, one should minimize the value of allowed bursts, i.e. make $\sigma' = 0$. If

$$\rho_c < \frac{1}{\sum_{k=1}^{N_c} \frac{M_k}{l_k}}, \quad (4)$$

the reshaping increases the delay bound and therefore should not be applied.

To sum it up, depending on network load, connection traffic either should be made Constant Bit Rate (CBR) or should not be reshaped at all.

Note that the global reshaping is advantageous for all the connections only if condition (3) holds for every connection. Otherwise, delay bounds for some connections are increased. For some applications, an increase of the end-to-end delay bound is acceptable as far as the bound does not exceed the value requested by the user. Such increase can be desirable, since reducing burstiness of a connection diminishes delays on other connections. To preserve the promised delay guarantees, it is necessary to ensure that for each connection C satisfying condition (4), the end-to-end delay bound d_c' is not larger than the requested bound D_c , i.e.,

$$\frac{\sigma - \sigma'}{\rho_c} + \left(\sum_{k=1}^{N_c} \frac{M_k}{l_k}\right)\sigma' \leq D_c$$

Hence, for every such connection the following must be satisfied:

$$\sigma' \geq \frac{\frac{\sigma}{\rho_c} - D_c}{\frac{1}{\rho_c} - \sum_{k=1}^{N_c} \frac{M_k}{l_k}}$$

Because $\sigma \geq \sigma' \geq 0$ and, according to (1) and (2), $D_c \geq \left(\sum_{k=1}^{N_c} \frac{M_k}{l_k}\right)\sigma$, the minimum value that may be selected for σ' is equal to:

$$\sigma' = \text{door}\left(\max_{\rho_c < \frac{1}{\sum_{k=1}^{N_c} \frac{M_k}{l_k}}} \left\{ \frac{\frac{\sigma}{\rho_c} - D_c}{\frac{1}{\rho_c} - \sum_{k=1}^{N_c} \frac{M_k}{l_k}} \right\}\right)$$

where *door* function is defined as follows: $\text{door}(x) = x$ if $x \geq 0$ and $\text{door}(x) = 0$ when $x < 0$.

The next section considers local reshaping when the traffic specification is changed only for one connection.

4 Local reshaping

In this section, we allow different connections to have different burstiness.

According to the schedulability conditions given in [3], the minimum end-to-end delay bound that can be guaranteed to connection C without reshaping is:

$$d_c = \sum_{k=1}^{N_c} \left(\frac{1}{l_k} \sum_{m=1}^{M_k} \sigma_{k,m} \right) \quad (5)$$

where $\sigma_{k,m}$ is the allowed burst of the m -th connection on link k .

Now let us reshape traffic on connection C by reducing its burstiness from σ to σ' , where $\sigma' \geq 0$, while keeping the burstiness of the other connections unchanged. According to [3], the minimum delay of sending data over link k for connection C equals to:

$$d'_k = \frac{1}{l_k} \left(\sum_{m=1}^{M_k} \sigma_{k,m} - \sigma + \sigma' \right) = \frac{1}{l_k} \sum_{m=1}^{M_k} \sigma_{k,m} - \frac{\sigma - \sigma'}{l_k}$$

Therefore, the the minimum end-to-end delay bound becomes:

$$\begin{aligned} d_c' &= d_{sm} + \sum_{k=1}^{N_c} d'_k = \frac{\sigma - \sigma'}{\rho_c} + \sum_{k=1}^{N_c} \left(\frac{1}{l_k} \sum_{m=1}^{M_k} \sigma_{k,m} - \frac{\sigma - \sigma'}{l_k} \right) \\ &= \sum_{k=1}^{N_c} \left(\frac{1}{l_k} \sum_{m=1}^{M_k} \sigma_{k,m} \right) + (\sigma - \sigma') \left(\frac{1}{\rho_c} - \sum_{k=1}^{N_c} \frac{1}{l_k} \right) \end{aligned}$$

Taking (5) into account, we can rewrite this expression as:

$$d_c' = d_c + (\sigma - \sigma') \left(\frac{1}{\rho_c} - \sum_{k=1}^{N_c} \frac{1}{l_k} \right)$$

Once again, the qualitative result of reshaping does not depend on the degree of reshaping.

If $\rho_c \geq \frac{1}{\sum_{k=1}^{N_c} \frac{1}{l_k}}$, the reshaping is advantageous, and to minimize the end-to-end delay bound, $\sigma' = 0$ should

be selected, i.e., connection traffic should be transformed to conform the CBR specification. Apart from improving the guarantees given to connection C , this will reduce delay bounds for the other connections by

$$\sigma \sum_{k=1}^{N_c} \frac{1}{l_k}.$$

If $\rho_c < \frac{1}{\sum_{k=1}^{N_c} \frac{1}{l_k}}$, reshaping increases the delay bound and therefore should not be applied.

Thus, depending on the requested rate, amount and bandwidths of links constituting the virtual circuit, connection traffic either should be made CBR or should not be reshaped at all.

Sometimes, an increase of the end-to-end delay bound is acceptable as far as the bound does not exceed the value D_c requested by the user. Since reshaping improves performance guarantees for the rest of connections, such increase can be even desirable. To preserve the promised delay bound, it is necessary to ensure that for connection C (for which $\rho_c < \frac{1}{\sum_{k=1}^{N_c} \frac{1}{l_k}}$), the end-to-end delay bound d_c' is not larger than the requested bound

D_c , i.e.,

$$\sum_{k=1}^{N_c} \left(\frac{1}{l_k} \sum_{m=1}^{M_k} \sigma_{k,m} \right) + (\sigma - \sigma') \left(\frac{1}{\rho_c} - \sum_{k=1}^{N_c} \frac{1}{l_k} \right) \leq D_c$$

Because $\sigma \geq \sigma' \geq 0$ and, according to (2) and (5), $D_c \geq \sum_{k=1}^{N_c} (\frac{1}{l_k} \sum_{m=1}^{M_k} \sigma_{k,m})$, the minimum value that may be selected for σ' is equal to:

$$\sigma' = \text{door}(\sigma - \frac{D_c - \sum_{k=1}^{N_c} (\frac{1}{l_k} \sum_{m=1}^{M_k} \sigma_{k,m})}{\frac{1}{\rho_c} - \sum_{k=1}^{N_c} \frac{1}{l_k}})$$

5 Conclusions

In this paper we investigated impact of traffic reshaping inside the network on providing deterministic guarantees of timely delivery of data. Our model can be viewed as a generalization of ATM networks.

In Section 3, we examined global reshaping for the case when all connections were characterized by the same burstiness. First, we applied traffic reshaping to lessen end-to-end delay bounds for every connection. We proved that, depending on network load, traffic on the connections either should be made CBR or should not be reshaped at all. Second, we considered the case when reshaping augmented the end-to-end delay bounds of some connections. Sometimes, such an increase is acceptable as far as the bounds do not exceed the values promised to the users. At the same time, the other connections may benefit from the reshaping. We derived the minimum value of burstiness on the connections which preserves the guaranteed timely behavior.

Section 4 considered local reshaping. Different connections were allowed to have different burstiness. First, we tried to reshape traffic on a single connection to decrease its end-to-end delay bound while keeping the burstiness for the rest of the connections unchanged. We demonstrated that (depending on the requested rate, amount and bandwidths of links constituting the virtual circuit) traffic on the connection either should be made CBR or should not be reshaped at all. Then, we considered the case when reshaping increases the end-to-end delay bound of the connection and obtained the minimum burstiness which preserves the delay bound requested by the user.

Even though conclusions based on a case study do not necessary remain valid for a generalized problem, our experiments with more complicated models allow us to suggest that traffic reshaping may be advantageous for providing timeliness guarantees in packet-switched virtual-circuit fixed-packet networks and that the optimal values of the traffic model parameters may be derived. In future, we plan to study impact of traffic reshaping in the context of more generic types of networks.

References

- [1] I. Cidon, I. Gopal, G. Grover, and M. Sidi, "Real-Time Packet Switching: A Performance Analysis," IEEE Journal on Selected Areas in Communications, 6(9):1576-1586, 1988.
- [2] L. Georgiadis, R. Guerin, V. Peris, and K.N. Sivarajan. "Efficient Network QoS Provisioning Based on Per Node Traffic Shaping," IEEE/ACM Transactions on Networking, 4(4):482-501, August 1996.
- [3] S. Gorinsky, S. Baruah, T. J. Marlowe, and A. D. Stoyenko. "Exact and Efficient Analysis of Schedulability in Fixed-Packet Networks: A Generic Approach," Technical Report CIS-96-26, Department of Computer and Information Science, NJIT, 1996.
- [4] R. Guerin. Personal correspondence.
- [5] E. Knightly and P. Rossaro. "Improving QoS through Traffic Smoothing," Proceedings of IFIP Workshop on Quality of Service, March 1996.
- [6] C. Partridge. "Gigabit Networking," Addison-Wesley, Reading MA, 1993.
- [7] H. Zhang and D. Ferrari. "Rate-Controlled Service Disciplines," Journal of High Speed Networks, 3(4):389-412, 1994.
- [8] Q. Zheng and K. G. Shin. "On the Ability of Establishing Real-Time Channels in Point-to-Point Packet-Switched Networks," IEEE Transactions on Communications, 42(2/3/4):1096-1105, February/March/April 1994.