Packet Multiplexers with Adversarial Regulated Traffic

Srini Rajagopal, Martin Reisslein, and Keith W. Ross
Department of Systems Engineering
University of Pennsylvania
Philadelphia, PA 19104
{sriniraj, reisslei, ross}@eniac.seas.upenn.edu

June 1997

Abstract

We consider a finite-buffer packet multiplexer to which traffic arrives from several independent sources. The traffic from each of the sources is *regulated*, i.e., the amount of traffic that can enter the multiplexer is constrained by known regulator constraints. The regulator constraints depend on the source and are more general than those resulting from cascaded leaky buckets. We assume that the traffic is adversarial to the extent permitted by the regulators. For lossless multiplexing, we show that if the original multiplexer is lossless it is possible to allocate bandwidth and buffer to the sources so that the resulting segregated systems are lossless. For lossy multiplexing, we use our results for lossless multiplexing to estimate the loss probability of the multiplexer. Our estimate involves transforming the original system into two independent resource systems, and using adversarial sources for the two independent resources to obtain a bound on the loss probabilities for the transformed system. We show that the adversarial sources are not extremal on-off sources, even when the regulator consists of a peak rate controller in series with a leaky bucket. We explicitly characterize the form of the adversarial source for the transformed problem. We also provide numerical results for the case of the simple regulator.

1 Introduction

Consider a finite-buffer packet multiplexer to which traffic arrives from several independent sources. For the multiplexer to provide quality of service (QoS) guarantees, such as limits on packet loss probabilities, it must have some knowledge about the traffic characteristics of the sources. Because the reliability of statistical models of traffic is questionable for many source types, in recent years there have been several studies on the performance of packet-switched nodes that multiplex *regulated traffic*, e.g., traffic which conforms to known constraints imposed by leaky buckets. These studies suppose that the traffic from the sources is *adversarial* to the extent permitted by the regulators [2] [1] [5] [6] [4] [3] [8] [7] [11] [13] [10]. Some of these studies assume that the multiplexer provides deterministic QoS guarantees (e.g., no packet loss) whereas others assume that multiplexer provides the less stringent probabilistic QoS guarantees (e.g., a limit on packet loss probability).

In a recent paper, LoPresti *et al.* [8] examine a packet-switched node with regulated traffic. Motivated by earlier work of Elwalid *et al.* [3], LoPresti *et al.* consider both deterministic QoS guarantees and probabilistic QoS guarantees. They assume that each source is regulated by a *simple regulator*, namely, a regulator that consists of a peak-rate controller in series with a leaky bucket. For deterministic QoS, LoPresti *et al.* show that if the multiplexer has sufficient link bandwidth and buffer capacity to provide lossless multiplexing, then the multiplexer's buffer and bandwidth can be allocated among the sources so that the resulting segregated systems are lossless. For probabilistic QoS, they develop a new approach to estimate the loss probability. Specifically, they transform the two-resource (bandwidth and buffer) allocation problem into two independent single-resource allocation problems; they then analyze these simpler, independent resource problems, taking on-off periodic sources for their adversarial sources.

Although the simple regulator is a popular policing mechanism within several standards bodies, it has been observed that it can often be a poor characterization of a source's worst-case traffic. A tighter and more powerful characterization is given by a more general regulator consisting of a cascade of multiple leaky buckets [5] [4]. For example, when the sources are VBR video sources, it is often possible to admit significantly more connections by replacing the simple regulator with cascaded-leaky-bucket regulators [5]. It is therefore desirable to extend the important work of [8] and [3] to the case of more general regulators.

In this paper we reexamine the model of [8] in the context of *generalized regulators*, which are even more general than cascaded leaky buckets. We first reexamine the lossless multiplexer of

LoPresti *et al.*, and extend their lossless results to generalized regulators. Using elementary tools from calculus, we show that if the original multiplexer is lossless, then it is possible to allocate bandwidth and buffer to the sources so that the resulting segregated systems are also lossless. We determine the optimal resource allocations and show that the buffer-bandwidth tradeoff curve is convex for generalized regulators. We also show that the segregation result does not necessarily hold for the delay-based QoS metric, even when the regulators are the simple regulators.

We then examine the multiplexer for probabilistic loss guarantees. We use our results for lossless multiplexing to estimate the loss probability of the multiplexer. As in [8], our estimate involves the following three steps: (*i*) choose a point on the buffer-bandwidth tradeoff curve and transform the original system into two independent resource systems; (*ii*) use adversarial sources for the two independent resources to obtain a bound on the loss probabilities for the transformed system; (*iii*) minimize the bound by searching over all points on the buffer-bandwidth tradeoff curve. Our principle contribution for probabilistic loss guarantees is an explicit characterization of the adversarial source for the transformed problem in Step (*ii*). Importantly, *the most adversarial source is not a periodic on-off source for the transformed problem consisting of two independent resources.* In fact, even in the case of simple regulated sources as studied in [8], the most adversarial source is not a periodic on-off source. Thus, in addition to generalizing the theory in [8] to the case of general regulated sources, we provide the true adversarial source for the case of the simple regulator. We also provide an algorithm to calculate the estimate of loss probability, assuming the truly adversarial sources.

We mention here that in [3] the original multiplexor problem is transformed into a bufferless multiplexer problem, and then the loss probability is bounded with the Chernoff bound. In this case, the worst-case adversarial sources are indeed on-off periodic sources. But when the original problem is transformed into a problem consisting of two independent resources, one bufferless resource and one buffered resource, the worst-case sources are no longer on-off periodic sources, even for simple regulators.

This paper is organized as follows. In Section 2 we define the model and the generalized regulators. In Sectin 3 we address lossless multiplexing. In Section 4 we address lossy multiplexing. In Section 5 we provide numerical results for lossy multiplexing of simple regulators, i.e., regulators consisting of a peak rate controller in series with a leaky bucket.

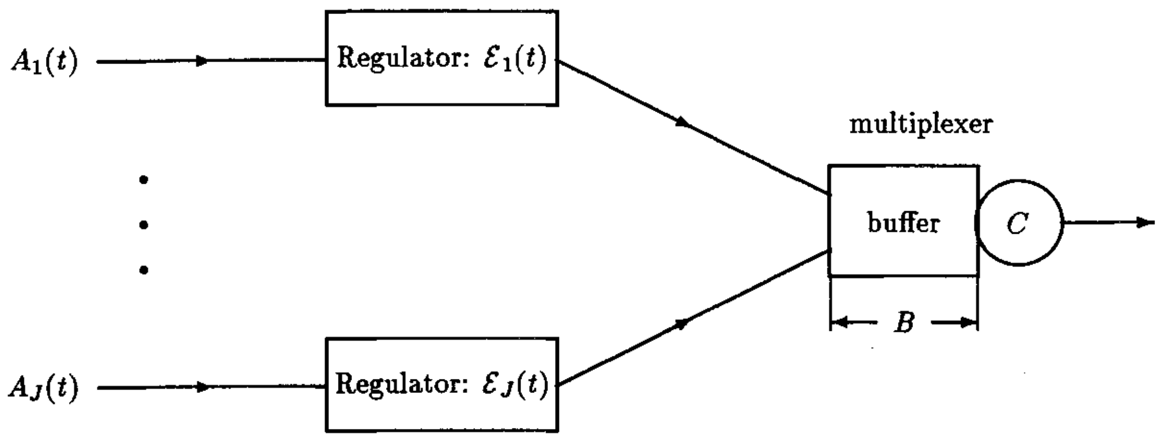

Figure 1: Link of capacity C, buffer of capacity B, and J regulators.

2 Regulated Traffic

We consider a link of rate C which is preceded by a finite buffer. Let J be the number of sources that send traffic to the buffer, and let $j = 1, \ldots, J$ index the sources. Each source j has an associated regulator function, denoted by $\mathcal{E}_j(t)$, $t \geq 0$. The regulator function constrains the amount of traffic that the jth source can send over an time interval of length t to $\mathcal{E}_j(t)$. More explicitly, if $A_j(t)$ is the amount of traffic that the jth source sends to the buffer over the interval $[0, t]$, then $A_j(\cdot)$ is required to satisfy

$$A_j(t+\tau) - A_j(\tau) \leq \mathcal{E}_j(t) \text{ for all } \tau \geq 0, \quad t \geq 0.$$

Figure 1 illustrates a multiplexer consisting of a link of rate C, a buffer of capacity B, and J sources with regulated traffic functions, $\mathcal{E}_j(t)$, $j = 1, \ldots, J$.

A popular regulator is the simple regulator, which consists of a peak-rate controller in series with a leaky bucket; for the simple regulator, the regulator function takes the following form:

$$\mathcal{E}_j(t) = \min\{\rho_j^1 t \,, \sigma_j^2 + \rho_j^2 t\}.$$

For a given source type, the bound on the traffic provided by the simple regulator may be loose and lead to overly conservative admission control decisions. For many source types (e.g., for VBR video [5]), it is possible to get a tighter bound on the traffic and dramatically increase the admission region. In particular, regulator functions of the form

$$\mathcal{E}_j(t) = \min\{\rho_j^1 t, \sigma_j^2 + \rho_j^2 t, \ldots, \sigma_j^{L_j} + \rho_j^{L_j} t\}$$

are easily implemented with cascaded leaky buckets and can lead to improved admission regions (see [5]).

In this paper we shall consider extremely general regulator functions, which include as special cases the forms mentioned above. To avoid certain trivialities, however, we shall always assume that $\mathcal{E}_j(0) = 0$, $\mathcal{E}_j(t)$ is non-decreasing in t, and that $\mathcal{E}_j(t)$ is subadditive in t (i.e., $\mathcal{E}_j(t_1 + t_2) \leq \mathcal{E}_j(t_1) + \mathcal{E}_j(t_2)$ for all t_1 and t_2). Also, unless explicitly mentioned otherwise, we shall assume that each $\mathcal{E}_j(t)$ is concave in t. Let

$$\mathcal{E}(t) = \sum_{j=1}^{J} \mathcal{E}_j(t)$$

be the aggregate regulator function. Due to the concavity of the $\mathcal{E}_j(t)$'s, the aggregate regulator function $\mathcal{E}(t)$ is also concave.

Before preceding with our analysis of the lossless systems, it is convenient at this point to introduce some notation and state a few technical facts. Let $\mathcal{E}_j^+(t)$ denote the right derivative for $\mathcal{E}_j(t)$ and $\mathcal{E}_j^-(t)$ denote the left derivative for $\mathcal{E}_j(t)$. Let $\mathcal{E}_j'(t)$ denote the derivative of $\mathcal{E}_j(t)$ whenever the derivative exists at t. Similarly define $\mathcal{E}^+(t)$, $\mathcal{E}^-(t)$ and $\mathcal{E}'(t)$. We will make use of the following fact: If $\mathcal{E}(t)$ is differentiable at t^*, then all of the $\mathcal{E}_j(t)$'s are differentiable at t^* (due to the concavity of the $\mathcal{E}_j(t)$'s).

3 Guaranteed Lossless Service and Optimal Segregation

It is well known [2] that the amount of traffic in the buffer does not exceed $B_{\min}$, where

$$B_{\min} = \max_{t \geq 0}\{\mathcal{E}(t) - Ct\} \ . \tag{1}$$

(To avoid trivialities we assume that the maximum is attained in (1).) Furthermore, due to subadditivity, it is possible to define traffic functions $A_j(t)$, $j = 1, \ldots, J$, such that the buffer contents will attain $B_{\min}$. Thus the minimum buffer size that will guarantee lossless operation is $B_{\min}$. Throughout the remainder of this section we assume that the multiplexer is lossless, i.e., we assume that the multiplexer buffer B satisfies $B \geq B_{\min}$.

It will be useful to write (1) in a more convenient form. If $\mathcal{E}(t)$ is differentiable then from (1) we have

$$B_{\min} = \mathcal{E}(t_{\max}) - Ct_{\max} , \tag{2}$$

where $t_{\max}$ is any solution to $\mathcal{E}'(t) = C$. More generally, there exists a $t_{\max}$ such that

$$\mathcal{E}^+(t_{\max}) \leq C \leq \mathcal{E}^-(t_{\max}) , \tag{3}$$

and any $t_{\max}$ which satisfies (3) also satisfies (2). Throughout the remainder of this section, fix a $t_{\max}$ that satisfies (3) (and therefore (2) as well).

We now address the following question: Is it possible to allocate bandwidth and buffer to the J sources so that each of the resulting segregated systems is also lossless? We shall see that the answer to this question is yes, but depends critically on the concavity of the $\mathcal{E}_j(t)$'s.

To address this issue, consider a new system which consists of a link of rate c preceded by a finite buffer. Suppose only the traffic from source j is sent to this system. The minimum buffer size that will ensure lossless operation is

$$B_{\min}(j, c) = \max_{t \geq 0}\{\mathcal{E}_j(t) - ct\} . \tag{4}$$

We say that a collection of J positive numbers $c_1, \ldots, c_J$ is a *bandwidth allocation* if $c_1 + \cdots + c_J = C$. For a given bandwidth allocation, we create J segregated systems, with the jth segregated system having link rate c_j and receiving traffic only from source j.

Theorem 1 1. *For all allocations* $B_{\min} \leq \sum_{j=1}^{J} B_{\min}(j, c_j)$.

2. *If one or more of the* $\mathcal{E}_j(t)$*'s is not concave then we may have* $B_{\min} < \sum_{j=1}^{J} B_{\min}(j, c_j)$ *for all allocations* $c_1, \ldots, c_J$.

3. *If each* $\mathcal{E}_j(t)$ *is concave then* $B_{\min} = \sum_{j=1}^{J} B_{\min}(j, c_j^*)$ *where* $c_j^* = \mathcal{E}_j'(t_{max})$ *if* $\mathcal{E}(t)$ *is differentiable at* $t = t_{\max}$ *and where*

$$c_j^* = \mathcal{E}_j^+(t_{\max}) + \alpha[\mathcal{E}_j^-(t_{\max}) - \mathcal{E}_j^+(t_{\max})]$$

with

$$\alpha = \frac{C - \mathcal{E}^+(t_{\max})}{\mathcal{E}^-(t_{\max}) - \mathcal{E}^+(t_{\max})}$$

if $\mathcal{E}(t)$ *is non-differentiable at* $t = t_{\max}$.

Proof. The proof of the first claim follows from (1) and (4):

$$\begin{aligned} B_{\min} &= \max_{t\geq 0}\{\mathcal{E}(t) - Ct\} = \max_{t\geq 0}\sum_{j=1}^{J}\{\mathcal{E}_j(t) - c_j t\} \\ &\leq \sum_{j=1}^{J}\max_{t\geq 0}\{\mathcal{E}_j(t) - c_j t\} = \sum_{j=1}^{J} B_{\min}(j, c_j) \end{aligned}$$

For the second claim, we offer the following counterexample with $J = 2$, $C = 1$. The envelope function for the first source is:

$$\mathcal{E}_1(t) = \begin{cases} t & \text{if } 0 \leq t \leq 1 \\ 1 & \text{if } 1 \leq t \leq 3 \\ 1 + (t-3) & \text{if } 3 \leq t \leq 4 \\ 2 & \text{if } t \geq 4\,. \end{cases}$$

The envelope function for the second source is:

$$\mathcal{E}_2(t) = \begin{cases} 2t & \text{if } 0 \leq t \leq 2 \\ 4 & \text{if } t \geq 2\,. \end{cases}$$

It is easily seen that $B_{\min} = 3$ whereas $B_{\min}(1, c_1) + B_{\min}(2, c_2) \geq 10/3$ for all allocations. Note that both $\mathcal{E}_1(t)$ and $\mathcal{E}_2(t)$ are non-decreasing and sub-additive. However, $\mathcal{E}_1(t)$ is not concave.

For the third claim, we first show that $c_1^*, \ldots, c_J^*$ is a feasible allocation. Suppose that $\mathcal{E}(t)$ is differentiable at $t_{\max}$. Due to the concavity assumption, this implies that each of the $\mathcal{E}_j(t)$'s is differentiable at $t_{\max}$. Thus

$$\begin{aligned} \sum_{j=1}^{J} c_j^* &= \sum_{j=1}^{J} \mathcal{E}_j'(t_{\max}) \\ &= \mathcal{E}'(t_{\max}) = C\,. \end{aligned}$$

If $\mathcal{E}(t)$ is not differentiable at $t = t_{\max}$, then it is easy to show directly from the definition of the c_j^*'s that $c_1^* + \cdots + c_J^* = C$. It remains to show that $B_{\min} = \sum_{j=1}^{J} B_{\min}(j, c_j^*)$. For a fixed transmission rate c, the concavity of the $\mathcal{E}_j(t)$'s and (4) imply

$$B_{\min}(j, c) = \mathcal{E}_j(t^*) - ct^*\,,$$

where t^* is any t that satisfies

$$\mathcal{E}_j^+(t) \leq c \leq \mathcal{E}_j^-(t)\,. \qquad (5)$$

By the definition of c_j^*,

$$\mathcal{E}_j^+(t_{\max}) \leq c_j^* \leq \mathcal{E}_j^-(t_{\max})\,.$$

Thus, $t_{\max}$ is a t that satisfies (5) for $c = c_j^*$. Therefore,

$$B_{\min}(j, c_j^*) = \mathcal{E}_j(t_{\max}) - c_j^* t_{\max} ,$$

which in turn implies

$$\begin{aligned}\sum_{j=1}^{J} B_{\min}(j, c_j^*) &= \sum_{j=1}^{J}[\mathcal{E}_j(t_{\max}) - c_j^* t_{\max}] \\ &= \mathcal{E}(t_{\max}) - C t_{\max} = B_{\min} .\end{aligned}$$

■

From Theorem 1 we know that it is possible to allocate bandwidth and buffer so that the resulting segregated systems are lossless, provided that the regulator functions are concave. This result generalizes a result in [8], in which all regulators were assumed to be simple regulators. This result also provides a motivation for the approach we take in Section 4 when we study probabilistic QoS.

Theorem 1 also gives fairly explicit formulas for these optimal allocations. In the following subsection we outline an efficient algorithm for calculating the allocations.

3.1 Algorithm to Calculate Allocations

In this subsection suppose that each of the regulator functions takes the form of cascaded leaky buckets:

$$\mathcal{E}_j(t) = \min\{\rho_j^1 t, \sigma_j^2 + \rho_j^2 t, \ldots, \sigma_j^{L_j} + \rho_j^{L_j} t\} .$$

Without loss of generality we may assume that

$$0 = \sigma_j^1 < \sigma_j^2 < \cdots < \sigma_j^{L_j} \tag{6}$$

and

$$\rho_j^1 > \rho_j^2 > \cdots > \rho_j^{L_j} . \tag{7}$$

Let

$$T_j^l = \frac{\sigma_j^{l+1} - \sigma_j^l}{\rho_j^l - \rho_j^{l+1}}, \quad l = 1, 2, \ldots, L_j - 1 .$$

In order to avoid trivialities we assume that

$$T_j^1 < T_j^2 < \cdots < T_j^{L_j - 1} . \tag{8}$$

With these assumptions, $T_j^1 < T_j^2 < \cdots < T_j^{L_j-1}$ are the breakpoints of $\mathcal{E}_j(t)$.

Here is an efficient algorithm for determining the optimal allocations $c_1^*, \ldots, c_J^*$ defined in Theorem 1. First sort $T_j^l, l = 1, \ldots, L_j, j = 1, \ldots, J$, in increasing order. Number them as $T_1, T_2, \ldots, T_L$. These points are the break points of $\mathcal{E}(t)$. Let k be the maximum l such that $\mathcal{E}^-(T_l) \leq C$. Note that to calculate $\mathcal{E}^-(T_l)$ it suffices to calculate $\mathcal{E}_j^-(T_l)$ for $j = 1, \ldots, J$; and to calculate $\mathcal{E}_j^-(T_l)$, we can determine the l_j such that $T_j^{l_j} \leq T_l < T_j^{l_j+1}$ and then set $\mathcal{E}_j^-(T_l) = \rho_j^{l_j+1}$ if $T_j^{l_j} < T_l$ and set $\mathcal{E}_j^-(T_l) = \rho_j^{l_j}$ if $T_j^{l_j} = T_l$.

The $t_{\max}$ in Theorem 1 is T_k. Once having determined k, find k_j such that $T_j^{k_j} \leq T_k < T_j^{k_j+1}$ and set $c_j^* = \rho_j^{k_j+1}$ if $T_j^{k_j} < T_k$ or set $c_j^* = \rho_j^{k_j} + \alpha(\rho_j^{k_j+1} - \rho_j^{k_j})$ if $T_j^{k_j} = T_k$, where α is defined in Theorem 1 and can also be determined directly from the ρ_j^l's.

3.2 The Buffer-Bandwidth Tradeoff Curve

For a given link rate C let $B_{\min}(C)$ be the maximum buffer contents defined by (1).The function $B_{\min}(C)$ is called the buffer-bandwidth tradeoff curve. For a probabilistic analysis in the next section, it will be useful to understand the behavior of the buffer-bandwidth tradeoff curve. To this end, for each fixed C let $t(C)$ be a value of $t_{\max}$ that satisfies (3). It is easily seen that $t(C)$ is non-increasing in C.

Theorem 2 *$B_{\min}(C)$ is non-increasing and convex in C.*

Proof. We first show that $B_{\min}(C)$ is non-increasing. Let $h > 0$. From (2) we have

$$B_{\min}(C) - B_{\min}(C+h) = \mathcal{E}(t(C)) - \mathcal{E}(t(C+h)) + t(C+h)(C+h) - t(C)C\,. \tag{9}$$

From the concavity of $\mathcal{E}(t)$ we have

$$\mathcal{E}^-(t(C)) \leq \frac{\mathcal{E}(t(C)) - \mathcal{E}(t(C+h))}{t(C) - t(C+h)}\,. \tag{10}$$

From (3) we have

$$C \leq \mathcal{E}^-(t(C))\,. \tag{11}$$

Combining (9)-(11) gives

$$\begin{aligned} B_{\min}(C) - B_{\min}(C+h) &\geq \mathcal{E}^-(t(C))[t(C) - t(C+h)] + t(C+h)(C+h) - t(C)C \\ &\geq C[t(C) - t(C+h)] + t(C+h)(C+h) - t(C)C \\ &= t(C+h)h \geq 0, \end{aligned}$$

which proves the first statement.

For the convexity of $B_{\min}(C)$, let $C_1 \le C_2$ and let $h > 0$. We must show

$$B_{\min}(C_2 + h) - B_{\min}(C_2) \ge B_{\min}(C_1 + h) - B_{\min}(C_1) . \tag{12}$$

By (2) it is equivalent to show

$$\begin{aligned} &\mathcal{E}(t(C_2 + h)) - \mathcal{E}(t(C_2)) + \mathcal{E}(t(C_1)) - \mathcal{E}(t(C_1 + h)) \\ &\ge t(C_2 + h)(C_2 + h) - t(C_2)C_2 - t(C_1 + h)(C_1 + h) + t(C_1)C_1 . \end{aligned} \tag{13}$$

Using the arguments in the proof of monotonicity, we have

$$\frac{\mathcal{E}(t(C_2)) - \mathcal{E}(t(C_2 + h))}{t(C_2) - t(C_2 + h)} \le C_2 + h \tag{14}$$

and

$$\frac{\mathcal{E}(t(C_1)) - \mathcal{E}(t(C_1 + h))}{t(C_1) - t(C_1 + h)} \ge C_1 . \tag{15}$$

Combining (13), (14) and (15) we obtain (12). ■

From Theorem 2 we know that $B_{\min}(C)$ is a decreasing convex function of C. If each of the regulator functions $\mathcal{E}_j(t)$ is piecewise linear, then it is easily shown that $B_{\min}(C)$ is a decreasing convex piecewise-linear function. Using the arguments in the proof of Theorem 2, it is straightforward to show that the optimal allocation c_j^* for the jth segregated system is increasing in C and that the buffer requirement for the jth segregated system, $B_{\min}(j, c_j^*)$, is decreasing in C.

3.3 Delay Metric

In Subsection 2.1 we showed how to allocate bandwidth so that, for lossless operation, the collective buffer requirements of the segregated system is equal to the buffer requirement of the multiplexed system. In other words, for the *buffer metric* we can find a bandwidth allocation such that the segregated system performs as well as the multiplexed system. In this subsection we briefly consider a natural *delay metric*. We show that it is not generally true that the segregated system performs as well as the multiplexed system for the delay metric.

For the multiplexed system the maximum delay is $d := B_{\min}/C$. For the jth segregated system with bandwidth c_j the maximum delay is $d(j, c_j) := B_{\min}(j, c_j)/c_j$. For a given allocation, we define the maximum delay of the collective segregated system to be the maximum of the maximum delays of the individual segregated systems, that is,

$$d_{seg} := \max_{1 \le j \le J} d(j, c_j) .$$

The following theorem draws comparisons between the maximum delay of the multiplexed system, d, and the maximum delay of the collective segregated system, d_{seg}.

Theorem 3 *1. For all allocations* $d \leq \max_{1\leq j\leq J} d(j, c_j)$.

2. There exist concave $\mathcal{E}_j(t)$*'s such that* $d < \max_{1\leq j\leq J} d(j, c_j)$ *for all allocations.*

3. If $\mathcal{E}_1(t) = \cdots = \mathcal{E}_J(t)$ *(homogeneous regulator functions), then* $d = \max_{1\leq j\leq J} d(j, c/J)$.

Proof. From Theorem 1 we have

$$B_{\min} \leq \sum_{j=1}^{J} B_{\min}(j, c_j) .$$

Dividing both sides of the above by $C = c_1 + \ldots + c_J$ and using the inequality

$$\frac{x_1 + \cdots + x_J}{y_1 + \cdots + y_J} \leq \max_{1\leq j\leq J} \frac{x_j}{y_j}$$

we obtain

$$\begin{aligned} d &\leq \frac{\sum_{j=1}^{J} B_{\min}(j, c_j)}{\sum_{j=1}^{J} c_j} \leq \max_{1\leq j\leq J} \{\frac{B_{\min}(j, c_j)}{c_j}\} \\ &= \max_{1\leq j\leq J} d(j, c_j), \end{aligned}$$

which establishes the first claim.

For the second claim we offer the following example: $C = 1$, $J = 2$, $\mathcal{E}_1(t) = 10$ for all $t \geq 0$ and

$$\mathcal{E}_2(t) = \begin{cases} 2t & \text{if } 0 \leq t \leq 5 \\ 10 & \text{if } t \geq 5 . \end{cases}$$

From (1) we have $d = 15$, $d(1, c_1) = 10/c_1$, and $d(2, c_2) = 10/c_2 - 5$. It is easily seen that for all allocations $\max(10/c_1, 10/c_2 - 5) > 15$.

The third statement follows directly from (1) and the definitions of d and $d(j, C/J)$. ∎

For the remainder of the paper we will use the original buffer metric.

4 Statistical Multiplexing with Small Loss Probabilities

For VBR sources the admission region can typically be made significantly larger by allowing loss to occur with minute probabilities, e.g., loss probabilities on the order of 10^{-6}. In this section we use our results of Section 3 to derive the worst-case loss probabilities for the multiplexer with regulated traffic.

We consider the same system defined in Section 2: The multiplexer consists of a link of rate C which is preceded by a finite buffer. There are J sources and the jth source has an associated regulator function, denoted by $\mathcal{E}_j(t)$, $t \geq 0$. In this section we suppose that the system resources are not sufficient to provide guaranteed lossless service. In other words, we assume $B < B_{\min}(C)$, so that there exists arrival processes which meet the regulator constraints but which cause the buffer to overflow. Let P_{loss} denote the expected fraction of time during which the buffer overflows. Our goal is to determine a bound for P_{loss} that holds for *all* combinations of arrival processes which meet the regulator constraints. To this end, we follow the methodology in [8] (which in turn is inspired by the paper [3]).

Let $a_j(t)$ be the rate at which source j transmits traffic at time t. We view $\{a_j(t),\ t \geq 0\}$ as a stochastic process. Our goal is to find independent rate processes $\{a_j(t),\ t \geq 0\}$, $j = 1, \ldots, J$, which maximize the loss probability over the class of all rate processes that meet the regulator constraints. To simplify the analysis, however, we only consider rate processes of the form

$$a_j(t) = b_j(t + \theta_j),$$

where $b_j(t)$ is a deterministic *periodic* function with some period T_j, and θ_j is a random variable, uniformly distributed over $[0, T_j]$. We assume that the phases $\theta_1, \ldots, \theta_J$ are independent, which implies that the rate processes $\{a_j(t),\ t \geq 0\}$, $j = 1, \ldots, J$, are also independent. We refer to $b_j(t)$ as a *source-j rate function.*

We say that a source-j rate function $b_j(t)$ is *feasible* if

$$\int_{\tau}^{t+\tau} b_j(s)ds \leq \mathcal{E}_j(t) \text{ for all } \tau \geq 0, \quad t \geq 0. \tag{16}$$

Note that for a given rate function $b_j(t)$ and phase θ_j the amount of source-j traffic sent to the multiplexer over the interval $[0, t]$ is

$$A_j(t) = \int_0^t b_j(s + \theta_j)ds.$$

Thus the regulator constraint

$$A_j(t + \tau) - A_j(\tau) \leq \mathcal{E}_j(t) \text{ for all } \tau \geq 0, \quad t \geq 0$$

is satisfied if and only if $b_j(t)$ is a feasible rate function.

As in [8], our derivation of a bound for P_{loss} involves the following three steps: (i) choose a point on the buffer-bandwidth tradeoff curve and transform the original system into two independent resource systems; (*ii*) use adversarial rate functions for the two independent resources to obtain a bound on the loss probabilities for the transformed system; (*iii*) minimize the bound by searching over all points on the buffer-bandwidth tradeoff curve. LoPresti *et al.* use an on-off rate function for their worst case rate function. Our approach differs from that of [8] in two respects. First, we allow for generalized regulators as opposed to simple regulators. Second, we derive the true adversarial rate functions, and employ these true adversarial rate functions in the bound for P_{loss} for both simple and generalized regulators.

4.1 The Virtual Segregated System

Fix a point (C_ν, B_ν) on the buffer-bandwidth tradeoff curve, and consider a lossless multiplexer with total amount of bandwidth C_ν and buffer space B_ν. Because the system resource pair (B, C) lies below the buffer-bandwidth tradeoff curve, we must have either $C_\nu > C$ or $B_\nu > B$ or both. For this lossless system we use Theorem 1 to allocate bandwidths $c_1^\nu, \ldots, c_J^\nu$ from C_ν and buffers $b_1^\nu, \ldots, b_J^\nu$ from B_ν such that each of the corresponding J segregated systems is lossless. This collection of J segregated systems is called the virtual segregated system [8].

For each $j = 1, \ldots, J$, fix a feasible rate function $b_j(t)$. Each rate function generates a stochastic arrival process

$$A_j(t) = \int_0^t b_j(s + \theta_j) ds.$$

For this arrival process, let U_j be a random variable that corresponds to the steady-state utilization of the jth segregated system; similarly, let V_j be the random variable that corresponds to the steady-state buffer contents of the jth segregated system. Because the θ_j's are independent across the sources, $U_1, \ldots, U_J$ are independent of each other and $V_1, \ldots, V_J$ are independent of each other.

For these fixed rate functions it can be argued [8] that

$$P_{\text{loss}} \leq P_\nu(\sum_{j=1}^{J} U_j > C) + P_\nu(\sum_{j=1}^{J} V_j > B)\,. \tag{17}$$

(The argument in [8] is for a simple regulator. It can be easily extended to our generalized regulators.) The equation (17) is the starting point of our own analysis.

Using the Chernoff bound we get

$$P_{\text{loss}} \leq \min_{\alpha \geq 0} \left\{ \frac{\prod_{j=1}^{J} M_{U_j}^\nu(\alpha)}{e^{\alpha C}} \right\} + \min_{\alpha \geq 0} \left\{ \frac{\prod_{j=1}^{J} M_{V_j}^\nu(\alpha)}{e^{\alpha B}} \right\} \tag{18}$$

where $M^{\nu}_{U_j}(\alpha)$ and $M^{\nu}_{V_j}(\alpha)$ are the moment genrating functions of U_j and V_j repectively, i.e., $M^{\nu}_{U_j}(\alpha) = E[e^{\alpha U_j}]$ and $M^{\nu}_{V_j}(\alpha) = E[e^{\alpha V_j}]$. Since (18) is valid for all points (C_ν, B_ν) on the buffer-bandwidth tradeoff curve, we have

$$P_{\text{loss}} \le \min_{(C_\nu, B_\nu)} \left[\min_{\alpha \ge 0} \left\{ \frac{\prod_{j=1}^{J} M^{\nu}_{U_j}(\alpha)}{e^{\alpha C}} \right\} + \min_{\alpha \ge 0} \left\{ \frac{\prod_{j=1}^{J} M^{\nu}_{V_j}(\alpha)}{e^{\alpha B}} \right\} \right] \tag{19}$$

We emphasize that the right-hand side of (19) depends on the fixed feasible rate functions. In order to give a bound that holds for *all* feasible rate functions we need to maximize the right-hand side of (19) over the set of all feasible rate functions. To this end, we introduce the notion of a *source-j adversarial rate function.*

Corresponding to each choice of (ν, α), we say that a source-j rate function is *adversarial* if (i) it is feasible, and (ii) it has the largest value of $M^{\nu}_{U_j}(\alpha)$ and $M^{\nu}_{V_j}(\alpha)$ among all feasible source-j rate functions. Now suppose that we can find the source-j adversarial rate functions for each choice of (ν, α); let U^*_j, V^*_j, $j = 1, \ldots, J$, be the corresponding steady-state random variables. We then have the following bound on P_{loss}:

$$P_{\text{loss}} \le \min_{(C_\nu, B_\nu)} \left[\min_{\alpha \ge 0} \left\{ \frac{\prod_{j=1}^{J} M^{\nu}_{U^*_j}(\alpha)}{e^{\alpha C}} \right\} + \min_{\alpha \ge 0} \left\{ \frac{\prod_{j=1}^{J} M^{\nu}_{V^*_j}(\alpha)}{e^{\alpha B}} \right\} \right] \tag{20}$$

Note that by using $M^{\nu}_{U^*_j}(\alpha)$ and $M^{\nu}_{V^*_j}(\alpha)$, which corresponds to the source-j adversarial rate function, we have obtained in (20) a bound on P_{loss} that is valid for all combinations of feasible arrival functions. We now proceed to characterize the adversarial rate functions.

4.2 Adversarial Sources

Throughout this subsection fix a ν, α and j. We now focus on determining a feasible rate function which maximizes both $M^{\nu}_{U_j}(\alpha)$ and $M^{\nu}_{V_j}(\alpha)$ over the set of feasible rate functions. We assume that the regulator functions have the form

$$\mathcal{E}_j(t) = \min\{\rho^1_j t, \sigma^2_j + \rho^2_j t, \ldots, \sigma^{L_j}_j + \rho^{L_j}_j t\} .$$

Note that $\mathcal{E}_j(t)$ is non-decreasing, concave, piecewise-linear and sub-additive. (The analysis that follows can easily be extended to the case of more general $\mathcal{E}_j(t)$ which are non-decreasing, concave and sub-additive.) Without loss of generality we also assume that (6), (7) and (8) hold. Note that the manner in which the allocations $(c^{\nu}_1, \ldots, c^{\nu}_J)$ are chosen (see Theorem 1) ensures that $\rho^{L_j}_j \le c^{\nu}_j \le \rho^1_j$ for all $j = 1, 2, \ldots, J$.

For a given feasible rate function $b_j(t)$ with period T_j, the arrival rate at time t is $a_j(t) = b_j(t+\theta_j)$ where θ_j is uniformly distributed over $[0, T_j]$. Corresponding to this $a_j(t)$ arrival rate process, let $v_j(t)$ be the buffer contents and $u_j(t)$ be the link utilization at time t. Note that $v_j(t)$ and $u_j(t)$ are periodic with period T_j. Also the steady-state random variables corresponding to $v_j(t)$ and $u_j(t)$ have distributions

$$P(V_j \leq x) = \frac{1}{T_j} \int_0^{T_j} 1(v_j(s) \leq x) ds$$

and

$$P(U_j \leq x) = \frac{1}{T_j} \int_0^{T_j} 1(u_j(s) \leq x) ds \, .$$

Note that these distributions do not depend on the phase θ_j and are completely determined by the rate function $b_j(t)$ and the link rate c_j^{ν}.

Throughout the remainder of this subsection we treat the case $c_j^{\nu} > \rho_j^{L_j}$. In the following subsection we deal with the simpler case $c_j^{\nu} = \rho_j^{L_j}$. Let

$$\delta_j = \max\{t > 0 : \frac{\mathcal{E}_j(t)}{t} \geq c_j^{\nu}\} \, . \tag{21}$$

Note that since $\rho_j^{L_j} < c_j^{\nu} \leq \rho_j^1$ and since $\mathcal{E}_j(\cdot)$ is an increasing concave function, δ_j is a uniquely defined, finite and strictly positive number. We now define an important class of rate functions. Let T_{off} be such that $0 < T_{\text{off}} \leq \delta_j$ and let

$$T_j = \frac{\mathcal{E}_j(T_{\text{off}})}{\rho_j^{L_j}}.$$

Now consider a rate function $b_j(t)$ with period T_j defined as follows:

$$b_j(t) = \begin{cases} \mathcal{E}_j^{+}(t) & 0 \leq t \leq T_{\text{off}} \\ 0 & T_{\text{off}} \leq t \leq T_j \end{cases}$$

Such a rate function is pictured in Figure 2.

This rate function is completely characterized by the parameter T_{off}. Note that the average arrival rate for this rate function is simply $\rho_j^{L_j}$. Let S_j be the collection of all rate functions of this form. Each rate function in S_j is identified through its T_{off} parameter.

We will show that the set S_j has the following important properties:

1. Each member of S_j is a feasible source-j rate function.

2. All members in S_j have identical $M_{U_j}^{\nu}(\alpha)$, and the members of S_j maximize $M_{U_j}^{\nu}(\alpha)$ over the set of all feasible source-j rate functions.

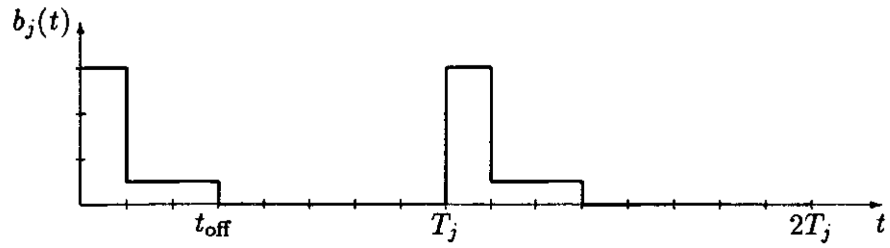

Figure 2: Example of a rate function in Set S_j when $t_{\text{off}} = 3$ and $\mathcal{E}_j(t) = \min\{3t, 2.5 + 0.5t\}$.

3. The member in S_j which has the largest $M^{\nu}_{V_j}(\alpha)$ has, in fact, the largest $M^{\nu}_{V_j}(\alpha)$ among all feasible source-j rate functions.

Hence, we will have shown that in order to find the source-j adversarial rate function corresponding to each choice (ν, α) we need only consider the rate functions in the set S_j. Further, since the rate functions in S_j are characterized by a single parameter, T_{off}, this essentially involves a single-parameter optimization problem. We now proceed to formally state and prove the properties listed above.

Theorem 4 *Every member of S_j is a feasible rate function.*

Proof. Fix a T_{off} and let $b_j(t)$ be the corresponding member of S_j. It follows immediately from the definition of $b_j(t)$ that

$$\int_0^t b_j(s)ds \leq \mathcal{E}_j(t) \text{ for all } 0 \leq t \leq T_j. \tag{22}$$

We can, in fact, show that

$$\int_0^t b_j(s)ds \leq \mathcal{E}_j(t) \text{ for all } t \geq 0. \tag{23}$$

To see this consider any arbitrary $t = nT_j + s$, where n is some non-negative integer and $0 \leq s \leq T_j$.

$$\begin{aligned} \int_0^t b_j(s)ds &= \int_0^{T_j} b_j(s)ds + \ldots + \int_{(n-1)T_j}^{nT_j} b_j(s)ds + \int_{nT_j}^{nT_j+s} b_j(s)ds \\ &\leq nT_j\rho_j^{L_j} + \mathcal{E}_j(s) \\ &\leq (\mathcal{E}_j(nT_j + s) - \mathcal{E}_j(s)) + \mathcal{E}_j(s) \\ &= \mathcal{E}_j(t)\,. \end{aligned}$$

The first inequality follows from (22) and from the fact that the average rate of $b_j(t)$ over any period is $\rho_j^{L_j}$. The second inequality follows because the slope of $\mathcal{E}_j(t)$ is never less than $\rho_j^{L_j}$.

Also, because $b_j(t)$ is non-increasing over each of its periods, we have

$$\int_{\tau}^{t+\tau} b_j(s)ds \leq \int_0^t b_j(s)ds \text{ for all } \tau \geq 0, \quad t \geq 0. \tag{24}$$

Combining (23) and (24) gives the desired result. ∎

Theorem 5 *Each member of S_j maximizes $M_{U_j}^{\nu}(\alpha)$ over the set of all feasible rate functions.*

Proof. Each rate function in S_j leads to the following form for $u_j(t)$, the utilization of the jth segregated system: $u_j(t)$ is periodic with period T_j; and

$$u_j(t) = \begin{cases} c_j^{\nu} & 0 \leq t \leq D_{\text{on}} \\ 0 & D_{\text{on}} \leq t \leq T_j \end{cases}$$

where $D_{\text{on}} = \frac{\mathcal{E}_j(T_{\text{off}})}{c_j^{\nu}} = (\frac{\rho_j^{L_j}}{c_j^{\nu}})T_j$.

The corresponding steady-state random variable is

$$U_j = \begin{cases} c_j^{\nu} & \text{with probability } \frac{\rho_j^{L_j}}{c_j^{\nu}} \\ 0 & \text{with probability } (1 - \frac{\rho_j^{L_j}}{c_j^{\nu}}) \end{cases} \tag{25}$$

Note that $E[U_j] = \rho_j^{L_j}$.

For any feasible source, the steady state rate at which traffic leaves the jth segregated system, U_j' (say), must have a peak value less than or equal to c_j^{ν}. Further, because the segregated system is lossless, the long-run average rate at which traffic departs the jth segregated system must equal the long-run average rate at which traffic enters the system, which is at most $\rho_j^{L_j}$. Hence, we must have $E[U_j'] \leq \rho_j^{L_j}$. Among all random variables which have a peak value less than or equal to c_j^{ν} and a mean value less than or equal to $\rho_j^{L_j}$, U_j as defined in (25) has the highest moment generating function, $M_{U_j}^{\nu}(\alpha)$. This is shown in the following argument (adapted from [9]). Let U_j' be any non-negative random variable with distribution $F_{U_j'}(x)$ with a peak value $c' \leq c_j^{\nu}$ and mean value $\mu' \leq \rho_j^{L_j}$. Then, since $\alpha \geq 0$,

$$\begin{aligned} M_{U_j}^{\nu}(\alpha) - M_{U_j'}^{\nu}(\alpha) &= (\frac{\rho_j^{L_j}}{c_j^{\nu}})e^{\alpha c_j^{\nu}} - \frac{\rho_j^{L_j}}{c_j^{\nu}} + 1 - \int_0^{c'} e^{\alpha x} dF_{U_j'}(x) \\ &\geq (\frac{\mu'}{c_j^{\nu}})e^{\alpha c_j^{\nu}} - \frac{\mu'}{c_j^{\nu}} - \int_0^{c'} (e^{\alpha x} - 1) dF_{U_j'}(x) \\ &= \frac{1}{c_j^{\nu}} \int_0^{c'} [x(e^{\alpha c_j^{\nu}} - 1) - c_j^{\nu}(e^{\alpha x} - 1)] dF_{U_j'}(x) \\ &\geq 0\,. \end{aligned}$$

■

Let $b_j^*(t)$ be a rate function in S_j that has the largest $M_{V_j}^{\nu}(\alpha)$.

Theorem 6 *$b_j^*(t)$ maximizes $M_{V_j}^{\nu}(\alpha)$ among all feasible rate functions.*

Proof. Consider any feasible source-j rate function $b_j(t)$ with period T_j. The actual arrival rate at time t is $a_j(t) = b_j(t + \theta_j)$ where θ_j is the random phase. Here, we are concerned only with the steady-state distributions of the buffer contents and the utilization rate of the jth segregated system which are independent of the phase. Hence, in the rest of the proof, we will, without loss of generality, set the phase to zero and consider $b_j(t)$ to be the arrival rate at time t. The corresponding buffer contents process, $v_j(t)$, is also periodic with period T_j.

In general, both $b_j(t)$ and $v_j(t)$ can have rather complicated forms with several intervals within a period where each is non-zero. However, we will first show the desired result for feasible rate functions that give a buffer content process of the form $v_j(t) > 0$ for $0 < t < \tau_j$ and $v_j(t) = 0$ for $\tau_j \leq t \leq T_j$, for some $0 < \tau_j < T_j$. For rate processes of this form we have

$$v_j(t) = \begin{cases} \int_0^t b_j(s)ds - c_j^{\nu}t & 0 \leq t \leq \tau_j \\ 0 & \tau_j \leq t \leq T_j \end{cases}$$

Note that, since $v_j(t) > 0$ for all $0 < t < \tau_j$, we must have

$$\tau_j \leq \delta_j \; . \tag{26}$$

We show next that $M_{V_j}^{\nu}(\alpha)$ corresponding to such a feasible rate function is smaller than that corresponding to $b_j^*(t)$. We do this by showing that there is a rate function in set S_j, $\bar{b}_j(t)$, with steady-state buffer contents $\bar{V}_j$ which is stochastically larger than V_j and which, hence, has a larger MGF (moment generating function).

Let T_{off} be such that $\mathcal{E}_j(T_{\text{off}}) = c_j\tau_j$. From (26) and (21) we get, $\mathcal{E}_j(T_{\text{off}}) < \mathcal{E}_j(\tau_j)$ if $\tau_j < \delta_j$ and $\mathcal{E}_j(T_{\text{off}}) = \mathcal{E}_j(\delta_j)$ if $\tau_j = \delta_j$. Hence, since $\mathcal{E}_j(\cdot)$ is non-decreasing and δ_j is uniquely defined, $T_{\text{off}} \leq \tau_j \leq \delta_j$. By definition, the rate function in S_j corresponding to this T_{off} is periodic with period $\bar{T}_j = \frac{\mathcal{E}_j(T_{\text{off}})}{\rho_j^{L_j}}$ and has the form

$$\bar{b}_j(t) = \begin{cases} \mathcal{E}_j^+(t) & 0 \leq t \leq T_{\text{off}} \\ 0 & T_{\text{off}} \leq t \leq \bar{T}_j \; . \end{cases}$$

The corresponding buffer contents at time t, $\bar{v}_j(t)$, is given as

$$\bar{v}_j(t) = \begin{cases} \mathcal{E}_j(t) - c_j^{\nu} t & 0 \leq t \leq T_{\text{off}} \\ \mathcal{E}_j(T_{\text{off}}) - c_j^{\nu} t & T_{\text{off}} \leq t \leq \tau_j \\ 0 & \tau_j \leq t \leq \bar{T}_j \end{cases}$$

Denote the corresponding steady-state random variable as $\bar{V}_j$.

Clearly, $v_j(t) \leq \bar{v}_j(t)$ for all $0 \leq t \leq T_{\text{off}}$. Note, also, that we cannot have $v_j(t) > \bar{v}_j(t)$ for any $T_{\text{off}} \leq t \leq \tau_j$ since that would require $v_j(t)$ to decrease at a rate strictly faster than c_j^{ν}, in order for both $v_j(t)$ and $\bar{v}_j(t)$ to be zero at τ_j. Hence, we get

$$v_j(t) \leq \bar{v}_j(t) \text{ for all } 0 \leq t \leq \tau_j \ . \tag{27}$$

Also, we can show that

$$T_j \geq \bar{T}_j \ . \tag{28}$$

To see this, note that the utilization rate of the jth segregated system with arrival rate $b_j(t)$ is c_j^{ν} whenever $v_j(t)$ is non-zero. Hence, $P(U_j = c_j^{\nu}) \geq \frac{\tau_j}{T_j}$. Also, since the average utilization rate must be equal to the average arrival rate, which in turn is smaller than $\rho_j^{L_j}$,

$$\rho_j^{L_j} \geq E[U_j] \geq c_j^{\nu} P(U_j = c_j^{\nu}) \geq c_j^{\nu} \frac{\tau_j}{T_j}$$

and so,

$$T_j \geq \frac{c_j^{\nu} \tau_j}{\rho_j^{L_j}} = \frac{\mathcal{E}_j(T_{\text{off}})}{\rho_j^{L_j}} = \bar{T}_j \ .$$

Equations (28) and (27) imply that

$$P(V_j > x) \leq P(\bar{V}_j > x) \text{ for all } x \geq 0 \ .$$

We have thus shown that V_j is stochastically smaller than $\bar{V}_j$ and hence has a smaller MGF. It is immediate from the definition of $b_j^*(t)$ that $M_{V_j}^{\nu}(\alpha)$ is smaller than that corresponding to $b_j^*(t)$.

We now extend this argument to the case of a general feasible rate function $b_j(t)$. Assume, without loss of generality, that the corresponding buffer content process $v_j(t)$ has m (some positive integer) non-zero portions within a single period, identified by $v_j^1, v_j^2, \ldots, v_j^m$ in the following manner:

$$v_j(t) = \begin{cases} 0 & 0 \le t \le t_j^1 \\ v_j^1(t - t_j^1) & t_j^1 \le t \le t_j^1 + \tau_j^1 \\ v_j^2(t - t_j^2) & t_j^2 \le t \le t_j^2 + \tau_j^2 \\ \vdots & \\ v_j^m(t - t_j^m) & t_j^m \le t \le t_j^m + \tau_j^m \\ 0 & t_j^m + \tau_j^m \le t \le T_j \end{cases}$$

where $\tau_j^i > 0$, $i = 1, 2, \ldots, m$, and $t_j^i \ge t_j^{i-1} + \tau_j^{i-1}$, $i = 2, \ldots, m$. Here, t_j^i and $t_j^i + \tau_j^i$ represent the endpoints of the ith non-zero portion.

We can express each non-zero portion $v_j^i(t)$ as a periodic function, with period T_j, of the following form:

$$v_j^i(t) = \begin{cases} \int_{t_j^i}^{t_j^i + t} b_j(s)ds - c_j^\nu t & 0 \le t \le \tau_j^i \\ 0 & \tau_j^i \le t \le T_j \,. \end{cases}$$

Let V_j^i denote the corresponding steady-state random variable with MGF $M^\nu_{V_j^i}(\alpha)$.

It is easily seen that

$$V_j = \begin{cases} V_j^1 & \text{with probability } \left(\frac{\tau_j^1}{\sum_{i=1}^m \tau_j^i}\right) T_j \\ \vdots & \\ V_j^m & \text{with probability } \left(\frac{\tau_j^m}{\sum_{i=1}^m \tau_j^i}\right) T_j \end{cases}$$

and hence,

$$M^\nu_{V_j}(\alpha) = \sum_{l=1}^m \left(\frac{\tau_j^l}{\sum_{i=1}^m \tau_j^i}\right) M^\nu_{V_j^l}(\alpha) \,. \tag{29}$$

Now, the ith non-zero portion, when viewed in isolation, has the simple form assumed in the earlier part of the proof, and can be viewed as the buffer contents at time t of the jth segregated system subject to the following arrival rate:

$$b_j^i(t) = \begin{cases} b_j(t_j^i + t) & 0 \le t \le \tau_j^i \\ 0 & \tau_j^i \le t \le T_j \,. \end{cases}$$

Note that $b_j^i(t)$ is also a feasible source-j rate function with period T_j. Hence, from our earlier argument, we know that $M^\nu_{V_j^i}(\alpha)$ is smaller than the MGF that corresponds to $b_j^*(t)$. Hence, from (29), we get that $M^\nu_{V_j}(\alpha)$ is also smaller than that corresponding to $b_j^*(t)$. We have thus shown that $b_j^*(t)$ maximizes $M^\nu_{V_j}(\alpha)$ over the set of all feasible source-j rate functions. ∎

From Theorems 5 and 6 the following corollary is immediate.

Corollary 1 *There exists a rate function belonging to S_j which maximizes both $M^{\nu}_{U_j}(\alpha)$ and $M^{\nu}_{V_j}(\alpha)$ over the set of all feasible source-j rate functions. This rate function is the required source-j adversarial rate function corresponding to (ν, α).*

Thus, when $c^{\nu}_j > \rho^{L_j}_j$, in order to find the source-j adversarial rate function corresponding to any choice of (ν, α) we need only consider the rate functions in set S_j.

4.3 The Case of $c^{\nu}_j = \rho^{L_j}_j$

We now deal with the special case of $c^{\nu}_j = \rho^{L_j}_j$. When $c^{\nu}_j = \rho^{L_j}_j$ it is easily seen that the adversarial source-j rate function has the following form:

$$b_j(t) = \mathcal{E}^{\dagger}_j(t) \quad \text{for all } t \geq 0.$$

Clearly, this rate function satisfies (16). We will drop the requirement of periodicity for this special case and consider this rate function to be feasible. (Alternatively, we could consider this rate function to be trivially periodic with a period of $+\infty$.) This rate function leads to the following degenerate form of the corresponding steady-state random variables:

$$U^*_j = c^{\nu}_j \quad \text{with probability 1}$$
$$V^*_j = b^{\nu}_j \quad \text{with probability 1}$$

with corresponding MGFs

$$M^{\nu}_{U^*_j}(\alpha) = e^{\alpha c^{\nu}_j}$$
$$M^{\nu}_{V^*_j}(\alpha) = e^{\alpha b^{\nu}_j}$$

which are clearly the largest possible values for these quantities.

In the next section we consider input sources that are constrained by simple regulators and describe a heuristic procedure to efficiently compute P_{loss} for this case.

5 Simple Regulators

In the last section we showed that for each segregated system there exists a rate function in S_j which is adversarial to the greatest extent possible permitted by the regulator constraint $\mathcal{E}_j(t)$. The set S_j includes the extremal periodic on-off rate functions studied in LoPresti *et al.* [8]. It is therefore natural to pose the following question: Is the extremal periodic on-off rate function adversarial?

Rate Function	T_{on}	T_{off}	T	D_{on}	D_{off}
1	$\frac{\sigma}{P-\rho}$	$\frac{\sigma}{\rho}$	$\frac{\sigma P}{\rho(P-\rho)}$	$\frac{\sigma P}{c(P-\rho)}$	$\sigma\frac{1-\rho+P(1-1/c)}{\rho(P-\rho)}$
2	$\frac{\sigma}{c-\rho}$	$\frac{\sigma}{\rho}$	$\frac{\sigma c}{\rho(c-\rho)}$	$\frac{\sigma}{c-\rho}$	$\frac{\sigma}{\rho}$
3	u	$\frac{\sigma}{\rho}$	$u+\frac{\sigma}{\rho}$	$\frac{\sigma+u\rho}{c}$	$\frac{(u\rho+\sigma)(c-\rho)}{\rho c}$

Table 1: On and off times of rate functions and corresponding segregated systems.

In this section we focus our attention on simple regulators $\mathcal{E}_j(t) = \min\{\rho_j^1, \sigma_j^2 + \rho_j^2 t\}$. We first show that the adversarial rate function in S_j *is not* the extremal on-off rate function used in LoPresti *et al.* This implies that the use of on-off rate functions, as in LoPresti *et al.*, can lead to overly optimistic admission regions. We then present an algorithm for calculating P_{loss} using the adversarial rate functions for each of the sources. This involves, for each source j, a search to find the T_{off} that leads to the most adversarial behavior.

5.1 Sub-Adversariality of On-Off Rate Functions

Fix a segregated system j. For ease of notation, let $P_j = \rho_j^1$, $\rho_j = \rho_j^2$ and $\sigma_j = \sigma_j^2$; the traffic constraint function is thus given by $\mathcal{E}_j(t) = \min(P_j t, \sigma_j + \rho_j t)$. We study 3 different rate functions, all complying with the imposed traffic constraint function. All these rate functions belong to S_j. Figure 4a gives the plots of the traffic constraint function, $\mathcal{E}_j(t)$, and the actual arrivals, $A_j(t)$, of the studied rate functions. Figure 4b depicts the arrival rate function $b_j(t)$. Figure 4c gives the link utilization $u_j(t)$. Figure 4d shows the buffer contents of the segregated system. Note that traffic leaves the segregated system at rate c_j whenever the buffer is nonempty. For the remainder of this section, we remove the subscript j from all notations.

Rate function 1 is the extremal on–off rate function used by Elwalid *et al.* [3] and LoPresti *et al.* [8]. It transmits at peak rate P for $T_{\text{on}_1} = \sigma/(P-\rho)$, at which time the token pool is completely emptied. The rate function then turns off and waits for $T_{\text{off}_1} = \sigma/\rho$, allowing the token pool to be refilled with σ tokens. The rate function then transmits the next burst of size PT_{on_1} at peak rate. The buffer is filled at rate $P-c$ while the source transmits at rate P. The maximum buffer contents is therefore $b = (P-c)T_{\text{on}_1}$. After the source has turned off, the buffer is drained at rate c. The utilization of the segregated system is c for $D_{\text{on}_1} = T_{\text{on}_1} + b/c$ and 0 for $D_{\text{off}_1} = T_{\text{off}_1} - b/c$. Rate Function 1 along with the other two rate functions are summarized in Table 1.

Rate function 2 transmits at peak rate P for T_{on_1}, it then continues sending traffic at rate ρ

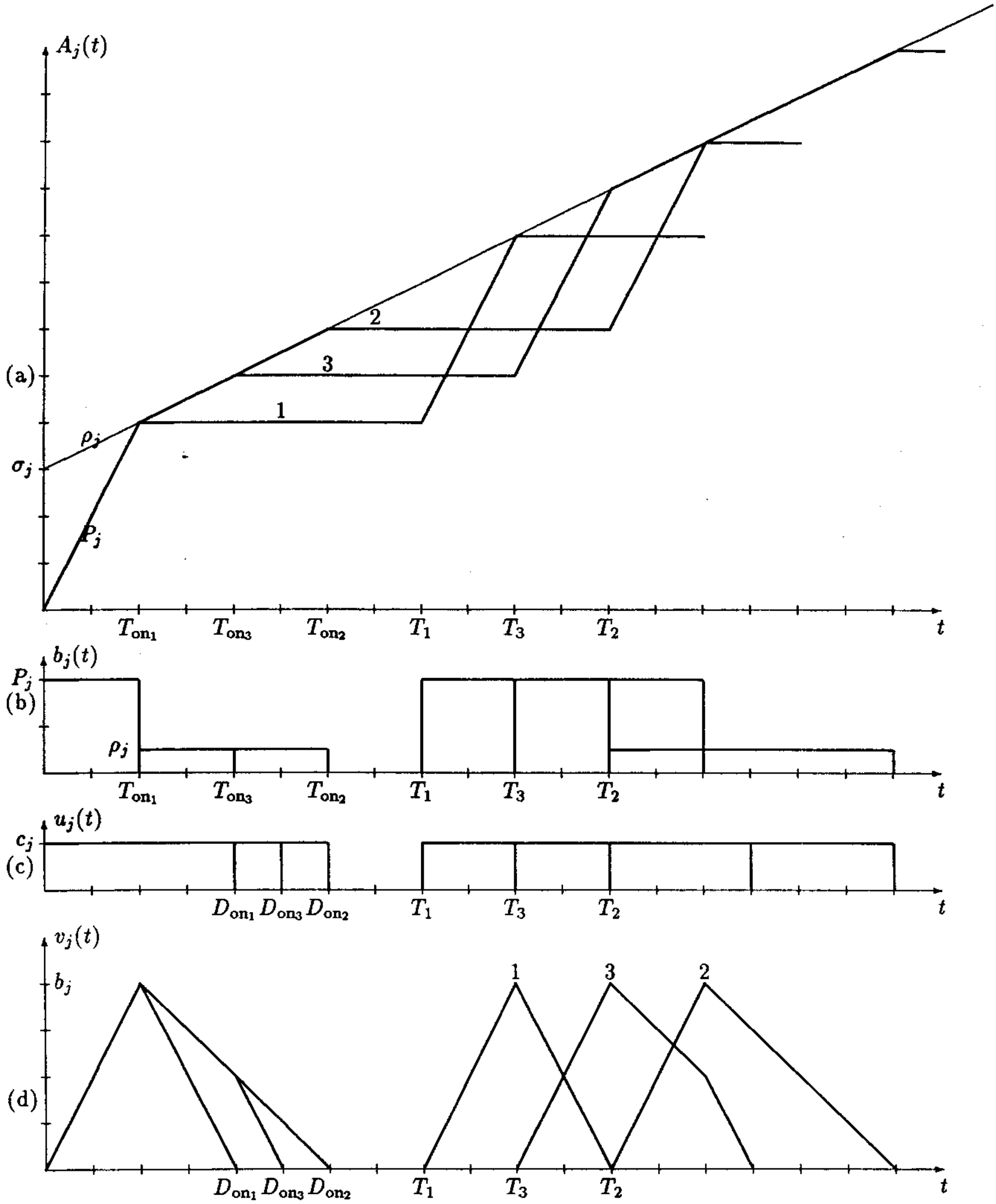

Figure 3: Illustration of rate functions 1, 2 and 3. (a) Amount of traffic arriving to the segregated system $A_j(t)$. (b) Arrival rate process $b_j(t)$. (c) Utilization process $u_j(t)$. (d) Buffer content process $v_j(t)$.

into the segregated system until the corresponding buffer process hits zero. As is the case for rate function 1, the buffer is filled up to b at rate $P - c$; it is now however drained at rate $c - \rho$. The source transmits therefore greedily for $T_{\text{on}_2} = T_{\text{on}_1} + b/(c - \rho)$. It then shuts off, waits until the token pool is replenished and repeats the described transmission pattern.

Rate function 3 generalizes the rate function behaviors discussed so far. It transmits greedily for u, $T_{\text{on}_1} \leq u \leq T_{\text{on}_2}$, that is, it transmits at rate P for T_{on_1} and then at rate ρ for $u - T_{\text{on}_2}$. The corresponding buffer process is depicted in Figure 4d. The buffer is filled to b at rate $P - c$. It is then drained at rate $c - \rho$ during the interval $[T_{\text{on}_1}, u]$. Let $v(u)$ denote the buffer contents at time u; clearly, $v(u) = \sigma + u(\rho - c)$. The remaining traffic $v(u)$ is drained at rate c. Loosely speaking, rate function 3 lies between the extremes of rate function 1 and rate function 2: it is equivalent to rate function 1 for $u = T_{\text{on}_1}$ and is equivalent to rate function 2 for $u = T_{\text{on}_2}$.

We now turn our attention to the buffer processes of the described rate functions. Let V_1, V_2 and V_3 be random variables denoting the buffer contents corresponding to rate function 1, 2 and 3. It can be easily verified that V_1 and V_2 have identical distribution functions:

$$P(V_1 \leq x) = P(V_2 \leq x) = 1 - \omega + x\frac{\omega}{b} \quad 0 \leq x \leq b, \tag{30}$$

where $\omega = \rho/c$ is the long run probability that the segregated system is busy. The distribution function of V_3 is given by

$$P(V_3 \leq x) = \begin{cases} 1 - \omega + x\frac{\omega}{b}\frac{P\sigma}{(P-\rho)(\rho u+\sigma)} & \text{for } 0 \leq x \leq v(u) \\ 1 - \omega + x\frac{\omega}{b}\frac{c\sigma}{(c-\rho)(u\rho+\sigma)} + \frac{u\rho^2}{(c-\rho)(u\rho+\sigma)} - \frac{\rho^2}{(c-\rho)c} & \text{for } v(u) \leq x \leq b. \end{cases}$$

Next we show that V_3 is strictly stochastically larger than V_1 and V_2 whenever $T_{\text{on}_1} < u < T_{\text{on}_2}$. First, note that

$$\frac{P\sigma}{(P-\rho)(\rho u+\sigma)} < 1 \quad \text{for } u > T_{\text{on}_1}.$$

Furthermore, it can be shown that

$$x\frac{\omega}{b}\frac{c\sigma}{(c-\rho)(u\rho+\sigma)} + \frac{u\rho^2}{(c-\rho)(u\rho+\sigma)} - \frac{\rho^2}{(c-\rho)c} < x\frac{\omega}{b}$$

for $u < T_{\text{on}_2}$ and $x < b$. Hence,

$$P(V_3 \leq x) < P(V_1 \leq x) \quad \text{for } 0 \leq x < b. \tag{31}$$

Thus V_3 is strictly stochastically larger than V_1 and V_2. This implies that the moment generating function of V_3 is larger than that of V_1 and V_2. The loss probability computed with rate function 3 is therefore larger than that corresponding to rate functions 1 and 2. Rate function 1, which is used in LoPresti *et al.* , can therefore lead to overly optimistic admission decisions.

$M_{V_3}(\alpha)$	$1-\omega+\frac{\omega\sigma}{\alpha b(\rho u+\sigma)}\{\frac{\rho(c-P)}{(P-\rho)(c-\rho)}e^{\alpha v(u)}+\frac{1}{1-\omega}e^{\alpha b}-\frac{P}{P-\rho}\}$
$\frac{\partial M_{V_3}(\alpha)}{\partial\alpha}$	$\frac{\omega\sigma}{\alpha^2 b(\rho u+\sigma)}\{\frac{\rho(c-P)}{(P-\rho)(c-\rho)}[\alpha v(u)-1]e^{\alpha v(u)}+\frac{1}{1-\omega}(b\alpha-1)e^{\alpha b}+\frac{P}{P-\rho}\}$
$\frac{\partial^2 M_{V_3}(\alpha)}{\partial\alpha^2}$	$\frac{\omega\sigma}{\alpha^3 b(\rho u+\sigma)}\{\frac{\rho(c-P)}{(P-\rho)(c-\rho)}[\alpha^2 v^2(u)-2\alpha v(u)+2]e^{\alpha v(u)}+\frac{1}{1-\omega}(\alpha^2 b^2-2\alpha b+2)e^{\alpha b}-\frac{2P}{P-\rho}\}$
$\frac{\partial M_{V_3}(\alpha)}{\partial u}$	$\frac{\omega\sigma}{\alpha b(\rho u+\sigma)^2}\{\frac{\rho(P-c)}{(P-\rho)}[\alpha\rho u+\alpha\sigma+\frac{\omega}{1-\omega}]e^{\alpha v(u)}-\frac{\rho}{1-\omega}e^{\alpha b}+\frac{P\rho}{P-\rho}\}$

Table 2: The moment generating function of the buffer process V_3 and its derivatives

5.2 Finding the most adversarial Rate Function

In this subsection we espouse the problem of finding the most adversarial rate function among the rate funcions fitting the template of rate function 3. Toward this end we need to find the on–time u that maximizes the moment generating function of V_3. The moment generating functin of V_3, defined as $M_{V_3}(\alpha) = E[e^{\alpha V_3}]$, and its derivative with respect to u are given in Table 2. The table gives furthermore the first and second derivative of $M_{V_3}(\alpha)$ with respect to s. These expressions are needed for the computation of P_{loss} (see Section 5.3).

Setting $\partial M_{V_3}(\alpha)/\partial u$ to zero, we obtain

$$(\alpha\rho u + \alpha\sigma + \frac{\omega}{1-\omega})e^{-\alpha(c-\rho)u} = \frac{(P-\rho)e^{\alpha b} - P(1-\omega)}{(1-\omega)(P-c)e^{\alpha\sigma}}. \tag{32}$$

It can be shown that (32) has exactly one solution in $[T_{\text{on}_1}, T_{\text{on}_2}]$. It can be computed efficiently with Newtons method [12] using $(T_{\text{on}_1} + T_{\text{on}_2})/2$ as initial solution. We observed in our numerical investigations that $(T_{\text{on}_1} + T_{\text{on}_2})/2$ provides in many cases a good approximation of the solution of (32). Rate function 3 with $u = (T_{\text{on}_1} + T_{\text{on}_2})/2$ may therefore be used as an approximation of the most adversarial rate function.

5.3 Numerical Examples

In this subsection we report on some numerical investigations with the most adversarial rate function. For the computation of P_{loss} we essentially follow the numerical procedure outlined in LoPresti *et al.* [8]. In addition to the computations conducted by LoPresti *et al.*, however, we solve (32) in order to find the most adversarial rate function.

We compare our approach with that of Elwalid *et al.* [3] and LoPretsi *et al.* in Figure 4. We use the same two source classes (see Table 3) as LoPresti *et al.* in [8, Fig. 15]. They in turn use the same parameters as Elwalid *et al.* in [3, Fig. 13]. The bandwidth and buffer size are $C = 45$ Mbps

class	ρ(Mbps)	P(Mbps)	σ(cells)
1	0.15	1.5	225
2	0.15	6	24.4

Table 3: Leaky Bucket parameters of sources.

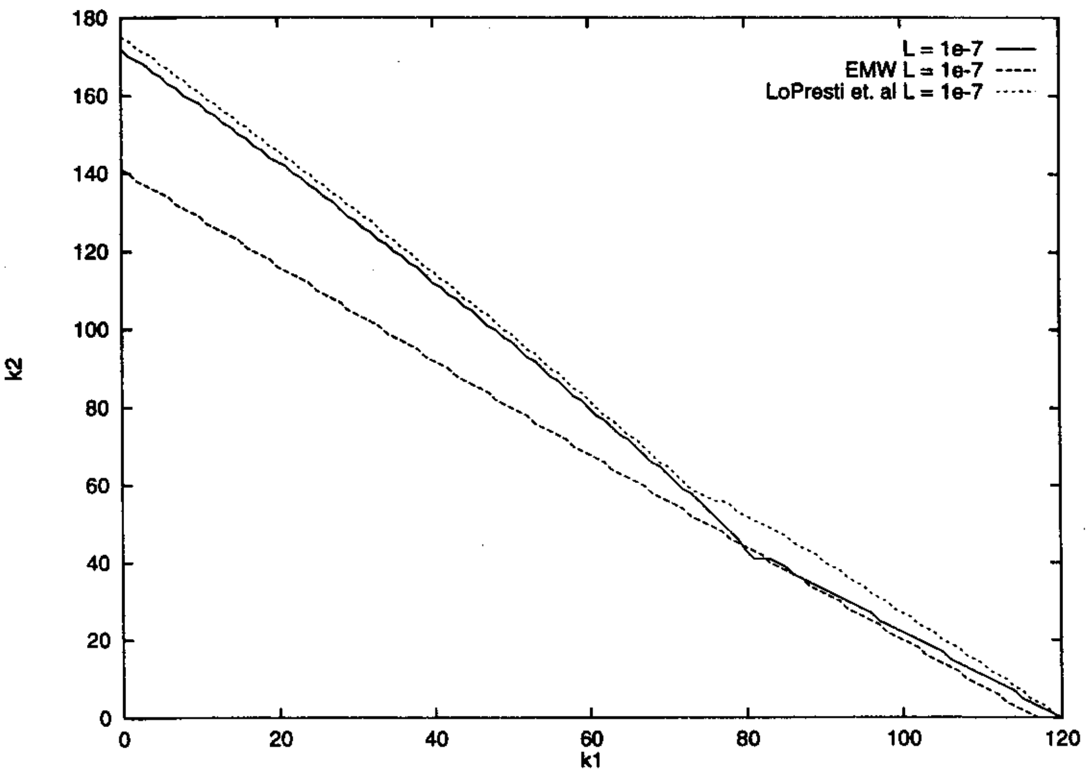

Figure 4: Comparison of our approach with Elwalid *et al.* [3] and LoPresti *et al.* [8].

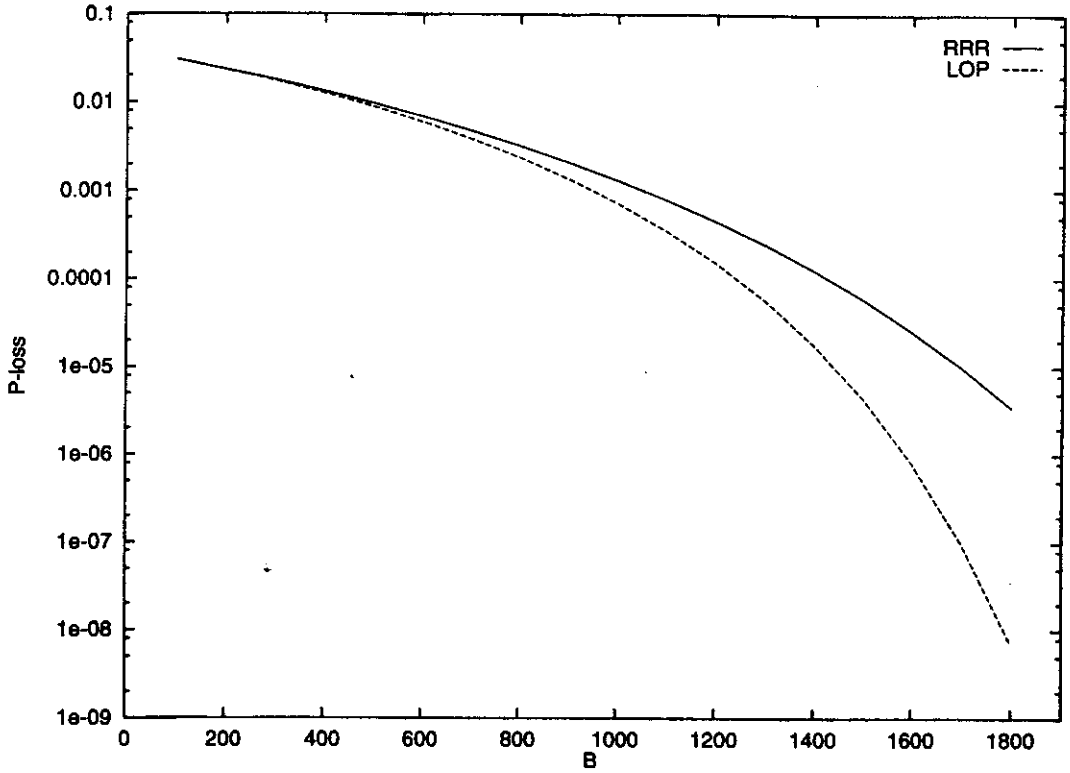

Figure 5: P_{loss} as a function of buffer size B.

and $B = 1000$ cells (1 cell = 53 bytes) in this example. The figure depicts the admission region corresponding to the admission control criterion $P_{\text{loss}} \leq 10^{-7}$. We observe that employing the truly adversarial rate function results in an admission region that lies generally between that of Elwalid *et al.* and LoPresti *et al.*. Because we are using the truly adversarial sources, our approach has a smaller admission region than LoPresti *et al.*. Our approach admits slightly less connections than the approach of LoPresti *et al.* in the range $0 \leq k_1 \leq 75$. For $k_1 = 0$, we admit 172 connections of class 2 while LoPresti *et al.* allow 175 connections. The gap between the two approaches widens for $k_1 > 75$. This is due to the fact that the optimal resource allocation according to Theorem 1 allocates $c_2^\nu = \rho_2$ in this region. Rate function 3 degenerates to the form described in Section 4.3 for this allocation. The moment generating function of this rate function is significantly larger that that corresponding to rate function 1, resulting in a noticably smaller admission region for our approach. The gap is at its widest for $k_1 = 81$. Our approach admits 41 connections of class 2 wile LoPresti *et al.* admit 51 connections.

In Figure 5 we consider a single source class with $P = 5$ cells/sec, $\rho = 2.5$ cells/sec and $\sigma = 20$ cells. (This choice of paramters is inspired by Oechslin [10].) We consider transmitting the traffic of 200 of these sources over a link of capacity $C = 575$ cells/sec. The figure shows P_{loss} computed according to our approach (RRR) and LoPresti *et al.* as a function of the buffer size B. We observe that both approaches give about the same loss probability for buffers smaller than 800 cells. For large buffers, however, the approaches differ greatly. For $B = 1400$ cells the loss probability computed according to LoPresti *et al.* is about one order of magnitude smaller that that computed with the most adversarial rate function. For $B = 1700$ cells the gap widens to roughly two orders of magnitude. We conclude from the figure that the approach of LoPresti *et al.* can significantly underestimate the loss probability.

References

[1] C.-S. Chang. Stability, queue length and delay of deterministic and stochastic networks. *IEEE Transactions on Automatic Control*, 39(5):913–931, May 1994.

[2] R. Cruz. A calculus for network delay, Part I: Network elements in isolation. *IEEE Transactions on Information Theory*, 37:114–131, 1991.

[3] A. Elwalid, D. Mitra, and R. H. Wentworth. A new approach for allocating buffers and bandwidth to heterogeneous regulated traffic in an ATM node. *IEEE Journal on Selected Areas in Communications*, 13(6):1115–1127, August 1995.

[4] L. Georgiadis, R. Guerin, V. Peris, and K.N. Sivarajan. Efficient network QoS provisioning based on per node traffic shaping. In *Proceedings of IEEE Infocom '96*, San Francisco, CA, April 1996.

[5] E. Knightly, J. Liebeherr, D. Wrege, and H. Zhang. Deterministic delay bounds for VBR video in packet-switching networks. *IEEE/ACM Transactions on Networking*, 4(3):352–362, June 1996.

[6] E. W. Knightly. H-BIND: A new approach to providing statistical performance guarantees to VBR traffic. In *Proceedings of IEEE Infocom '96*, San Francisco, CA, April 1996.

[7] K. Kvol and S. Blaabjerg. Bounds and approximations for periodic on/off queue with applications to ATM traffic control. In *Proceedings of IEEE Infocom*, pages 487–494, 1992.

[8] F. LoPresti, Z. Zhang, D. Towsley, and J. Kurose. Source time scale and optimal buffer/bandwidth trade-off for regulated traffic in an ATM node. In *Proceedings of IEEE Infocom*, Kobe, Japan, April 1997.

[9] D. Mitra and J. A. Morrison. Multiple time scale regulation and worst case processes for ATM network control. In *Proceedings of 34th Conference on Decision and Control*, pages 353–358, 1995.

[10] Ph. Oechslin. Worst case arrival of leaky bucket constrained sources: The myth of the on-off source. In *Proceedings of the fifth IFIP International Workshop on Quality of Service*, pages 67–77, Columbia University, New York, May 1997.

[11] A. Parekh and R. Gallager. A generalized processor sharing approach to flow control in integrated services networks: The single node case. *IEEE/ACM Transactions on Networking*, 1(3):344–357, June 1993.

[12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C, The Art of Scientific Computing.* Cambridge University Press, Cambridge, MA, second edition, 1992.

[13] J. W. Roberts, B. Bensaou, and Y. Canetti. A traffic control framework for high speed data transmission. In *Proceedings of IFIP Workshop Modelling, Performance Evaluation, ATM Technology*, pages 243–262, 1993.