Alcom-FT Technical Report Series
ALCOMFT-TR-03-11

Admission control of leaky bucket regulated sources in a queueing system with priority

Vincent Dumas[a], Fabrice Guillemin[b] and Philippe Robert[c*]

[a]Deceased, Vincent Dumas' work was done when visiting INRIA Rocqencourt in 2001

[b]France Telecom R&D, F-22300 Lannion, e-mail: Fabrice.Guillemin@francetelecom.com

[c]INRIA Rocquencourt, F-78153 Le Chesnay, e-mail: Philippe.Robert@inria.fr

A new admission control algorithm for leaky bucket regulated traffic sources is developed in this paper by analyzing the $M^{[X]}/M/1/K$ queue. A time priority level is taken into account by introducing a Poisson stream, which has time priority over the other sources. Effective bandwidths are then derived on the basis of exact upper bounds for the loss probability in the $M^{[X]}/M/1/K$ queue with priority. These effective bandwidths are then compared with other effective bandwidths, which have been proposed earlier in the technical literature. It turns out that the effective bandwidths obtained in this paper, although conservative, offer better performances when the peak to mean ratio is high and the mean rate is small when compared to the link transmission capacity.

1. INTRODUCTION

One popular approach to guaranteeing quality of service in broadband integrated networks consists of negotiating traffic parameters between the network and the user prior to any data transfer. The most commonly used parameter set is composed of the peak rate π and the mean rate ρ together with the bucket size σ; these three parameters describe to some extent the statistical behavior of a flow of packets (a connection in ATM networks or an IP flow in the Internet). These parameters are controlled at network access point by a dual leaky bucket (one for the peak rate and another with bucket size σ for the mean rate ρ) in order to prevent any quality of service degradation due to malicious usage. Such sources are referred to as (σ, ρ, π)- regulated sources and are the central objects of the so-called network calculus [1].

Resource allocation and admission control for (σ, ρ, π)-regulated sources have given rise over the past ten years to a huge amount of work. Several admission control algorithms have been proposed in the technical literature and implemented in switches and routers available today on the market. Most admission control algorithms have been developed by considering a single FIFO buffer and rely on refined Chernoff bounds for estimating the loss probability (see for instance [2,3]).

*Partially supported by the contract 00-1B320 with France Telecom and the Future and Emerging Technologies programme of the EU under contract number IST-1999-14186 (ALCOM-FT)

Admission control in queueing systems with priority has received less attention than the FIFO case in the literature. Most algorithms rely on the so-called reduced service rate (RSR) approximation (see [4,5]), which claims that everything happens as if the service rate of the low priority buffer were reduced up to the input rate of the high priority traffic. More recently, admission control relying on Chernoff bounds have been developed for priority systems, for instance by Reisslein [6] and Shakottai and Srikant [7].

In this paper, we address the problem of admission control by considering a queueing model, which stochastically dominates the original model composed of a single buffer fed with the superposition of (σ, ρ, π)-regulated traffic sources. This dominating model is the $M^{[X]}/M/1/K$ queueing system. We derive exact upper bounds for the loss probability in this system and we analyze what happens when a time priority is given to a Poisson stream. This allows us to derive the effective bandwidth for a (σ, ρ, π)-regulated source in the presence of a time priority without invoking the RSR approximation. The effective bandwidth obtained in this paper is then compared to other effective bandwidths, namely those introduced by Elwalid *et al* [2] and Gibbens and Hunt [3], when there is no priority traffic.

The organization of this paper is as follows: The motivation for introducing in the framework of admission control the $M^{[X]}/M/1/K$ queueing system is presented in Section 2. Upper-bounds for the loss probability in this queueing system are established in Section 3. Analysis of the $M^{[X]}/M/1/K$ with priority and the computation of effective bandwidths are carried out in Section 4. Comparisons with other effective bandwidths, when there is no priority traffic, are performed in Section 5. Concluding remarks are presented in Section 6.

2. PROBLEM FORMULATION

We consider in this paper statistical multiplexing of traffic sources, which are regulated by a dual leaky bucket. One leaky bucket controls the peak rate π of a source with a stringent tolerance. The second leaky bucket enforces the mean rate ρ of a source with a bucket size σ.

When multiplexing such (σ, ρ, π)-regulated sources in a FIFO queue with finite capacity, admission control algorithms are generally developed by taking into account the so-called worst-case traffic pattern associated with a source. Such a pattern is periodic and composed of bursts with size $b = \lfloor 1+(\sigma-1)\pi/(\pi-\rho) \rfloor$ at the peak rate, the distance between the starting times of two consecutive bursts is b/ρ. The duration of a burst is b/π and the silence period between two bursts is $b(\pi-\rho)/(\pi\rho)$.

However, it has recently been shown by Kesisidis and Konstantopolous [8] that in a fluid flow context, the periodic On/Off process is not the worst case when considering the loss probability. Such a result is in line with the counterexamples exhibited by Doshi [9], who showed that in a buffered system, the worst case is not the periodic On/Off process. In the case of a single input process, the traffic pattern, which maximizes the loss probability, is in fact periodic and composed of bursts at the peak rate, followed by activity periods at the mean rate ρ and followed in turn by silent periods so as to respect the mean rate constraint. It is nevertheless easy to check that the worst case probability given in [8] is less than the loss probability obtained when batches with size σ arrive according to a

Poisson process with intensity ρ/σ.

This result is true for a single input process, but it is reasonable to conjecture that it holds when considering several input processes. This leads us to replace each individual (σ, ρ, π)-regulated flow with a batch Poisson process; the batch size is equal to σ and the intensity of the Poisson process is ρ/σ. Now, when considering statistical multiplexing of different traffic sources, we are led to analyze a FIFO queue fed with a batch Poisson process, which corresponds to the superposition of the individual batch Poisson processes. Finally, taking into account the discrete nature of packets, we assume that packet service times are exponentially distributed. This is clearly not the case for ATM. However, since the exponential distribution is more variant than the deterministic distribution, the $M^{[X]}/M/1/K$ queue stochastically dominates the $M^{[X]}/D/1/K$ queue. Once again, this fact can rigorously be proved via classical stochastic ordering arguments. Hence, by considering the $M^{[X]}/M/1/K$ queue for admission control, we derive conservative equivalent bandwidths.

Note that the $M^{[X]}/M/1/K$ system has already been considered by Gravey *et al* [10] to determine the maximal burst size for a bandwidth allocation on the basis of the mean rate only. In this paper, we do not search for this maximal value but we allow for heterogeneous burst sizes. The counterpart is that we have an over-allocation factor on the sustainable cell rate, which can be significant in some situations.

The real motivation for introducing the $M^{[X]}/M/1/K$ system is in that we can easily account for a priority level and carry out exact computations. When considering a priority level, we assume that the higher priority packets have preemptive resume priority over the lower priority packets. This is clearly a distortion of the real system, since the service of a packet cannot be interrupted. Nevertheless, if the size of packets is small (and this is precisely the case for ATM cells), the difference between the preemptive resume priority and the real system will not be significant. This is why we neglect this border effect in the following.

3. LOSS PROBABILITY IN A $M^{[X]}/M/1/K$ QUEUE

The $M^{[X]}/M/1/K$ queue has been analyzed in details by Tijms and Van Ommeren in [11]. Our goal in this section is to derive simple upper bounds for the loss probability in that queueing system. These bounds will then be used to develop a new admission control algorithm in the following sections.

Throughout this section, we consider a single-server queue with capacity $K < \infty$ fed with batches of packets, which arrive according to a Poisson process with rate λ. The batch size has a general probability distribution $\{\beta_j,\ j = 1,\ 2,\ \ldots\}$ with generating function $\beta(z) = \sum_{j=1}^{\infty} \beta_j z^j$ and finite mean $\bar{\beta}$. Define the load of the system as $\rho = \lambda\bar{\beta}\bar{S}$, where $\bar{S}$ is the mean service time. It is assumed that $\rho < 1$ so that the system with infinite capacity is stable.

The loss probability $\pi_{\text{loss}}(K)$ in the system is defined as the long term fraction of packets that are lost. Tijms and Van Ommeren [*ibid.*] showed that the generating function $Q(z)$ of the number of packets in the system in the stationary regime is given for $|z| \leq 1$

by

$$Q(z) = \frac{(1-\rho)\beta^*(z)}{(1-\rho z\beta^*(z))} \quad \text{where} \quad \beta^*(z) = \frac{1-\beta(z)}{(1-z)\bar{\beta}};$$

$\beta^*(z)$ is the generating function of the distribution $\{\beta_j^* = \sum_{i>j} \beta_i/\bar{\beta},\, j = 0,1,2,\ldots\}$. The loss probability is then given by:

$$\pi_{\text{loss}}(K) = \frac{(1-\rho)\sum_{j\geq K} q_j(\infty)}{1-\rho\sum_{j\geq K} q_j(\infty)}. \tag{1}$$

Let R denote the convergence radius of $\beta^*(z)$. In the case $\rho R\beta^*(R) > 1$, it can be showed (see [11] for details) that there exists a unique $z_0 \in (1, R)$, which is also the solution with the smallest module to the equation

$$\rho z_0 \beta^*(z_0) = 1. \tag{2}$$

The asymptotic behavior of the queue length probability distribution function depends on whether $\rho R\beta^*(R) > 1$ or not. By using Kingman's upper bound [12], we can specifically prove the following result (see [**?**] for details).

Proposition 1 *The loss probability is such that*

$$\pi_{loss}(K) \leq \begin{cases} \beta^*(z_0) z_0^{-K} & \text{if } \rho R\beta^*(R) > 1, \\ \beta^*(R) R^{-K} & \text{if } \rho R\beta^*(R) \leq 1. \end{cases} \tag{3}$$

4. THE $M^{[X]}/M/1/K$ QUEUE WITH PRIORITY

4.1. Equivalent single source model

Let us consider statistical multiplexing of (σ, ρ, π)-regulated sources in a buffer with capacity $K < \infty$. Assume that there are N source types and x_i sources with type i, characterized by the triplets $(\sigma_i, \rho_i, \pi_i)$, $i = 1, \ldots, N$. As stated in Section 2, in view of admission control, we assume that each individual stream of type i generates batches of σ_i packets according to a Poisson process with rate ρ_i/σ_i. Finally, we assume that the bucket sizes σ_i, $i = 1, \ldots, N$ cannot exceed a given constant κK, with $\kappa \in (0, 1)$, and we set $K' = (1-\kappa)K$.

Independently from the above N source types, high priority packets arrive at the single-server channel according to a Poisson process with rate λ_0. Their service times have a general probability distribution function $F_0(x)$ with Laplace transform $F_0^*(s) = \int_0^\infty e^{-sx} dF_0(x)$, $\Re(s) \geq 0$, and finite mean $\bar{S}_0$. This new stream has preemptive resume priority over the other traffic sources, and and priority packets are queued in a separate buffer with infinite capacity. Letting $\rho_0 = \lambda_0 \bar{S}_0$, it is assumed that the total offered load $\rho_0 + \sum_{i=1}^N x_i \rho_i$ is less than one.

As far as the original sources are concerned, all they see is that service is interrupted during the busy periods of the high priority $M/G/1$ queue. The time between two consecutive busy periods is exponentially distributed with parameter λ_0; the durations of the busy periods have a probability distribution function $G_0(x)$ with Laplace transform given by $G_0^*(s) = \int_0^\infty e^{-sx} dG_0(x)$ for $\Re(s) \geq 0$; it is well-known (see for example Feller [13], Chap. XIV.4) that G_0^* can be characterized as the smallest root of the equation $G_0^*(s) = F_0^*(s + \lambda_0(1 - G_0^*(s)))$; the mean duration of a busy period is $\bar{B}_0 = -(G_0^*)'(0) = \bar{S}_0/(1-\rho_0)$.

Furthermore, it is easily seen that the overflow probability in the $M^{[X]}/M/1/K$ queue will remain unchanged if we contract each busy period in one point and gather all the packets arrived meanwhile from the original sources in one single batch. Elementary calculations show that the generating function and the mean value of this batch size are

$$\beta_0(z,x) = G_0^*\left(\sum_{i=1}^N x_i\lambda_i(1-\beta_i(z))\right) \quad \text{and} \quad \bar{\beta}_0(x) = \bar{B}_0 \sum_{i=1}^N x_i\lambda_i\bar{\beta}_i, \tag{4}$$

respectively.

This procedure is equivalent to replacing the higher priority $M/G/1$ queue with one additional source of type, say, type 0, which generates batches of packets according to a Poisson process with rate λ_0; the batch size has a probability distribution with generating function $\beta_0(z,x)$ and finite mean $\bar{\beta}_0(x)$. Packet service times are independent random variables having the usual exponential distribution with finite mean $\bar{S}$.

Finally, our new model is equivalent to the single-source model of Section 3, where batches arrive at rate $\lambda(x) = \lambda_0 + \sum_{i=1}^N x_i\lambda_i$, have generating function and mean value

$$\beta(z,x) = \frac{\lambda_0}{\lambda(x)}\beta_0(z,x) + \sum_{i=1}^N \frac{x_i\lambda_i}{\lambda(x)}\beta_i(z) \quad \text{and} \quad \bar{\beta}(x) = \frac{\sum_{i=1}^N x_i\lambda_i\bar{\beta}_i}{\lambda(x)(1-\rho_0)},$$

respectively. The offered load $\rho(x)$ defined by $\rho(x) = \lambda(x)\bar{\beta}(x)\bar{S} = \sum_{i=1}^N x_i\rho_i/(1-\rho_0)$, is less than one as expected.

4.2. Loss probability in the presence of the high priority queue

Denoting by $\pi_{\text{loss}}(K,x)$ the overflow probability, we intend to use the upper bounds of Section 3, but we have to determine whether $\rho(x)R(x)\beta^*(R(x),x) > 1$ or not, where $R(x)$ is the radius of convergence of the series $\beta^*(z,x) = (1-\beta(z,x)/[(1-z)\overline{\beta}(x)]$. First note that

$$\rho(x)z\beta^*(z,x) = \lambda(x)\bar{\beta}(x)\bar{S}z\frac{1-\beta(z,x)}{(1-z)\bar{\beta}(x)} = \rho_0(x)z\beta_0^*(z,x) + \sum_{i=1}^N x_i\rho_i z\beta_i^*(z),$$

where $\rho_0(x) = \lambda_0\bar{\beta}_0(x)\bar{S}$.

According to equation (4), we have

$$\rho_0(x)z\beta_0^*(z,x) = \frac{\rho_0}{1-\rho_0}\overline{G}_0^*\left(\sum_{i=1}^N x_i\lambda_i(1-\beta_i(z))\right)\sum_{i=1}^N x_i\rho_i z\beta_i^*(z),$$

where $\overline{G}_0^*(s) = (1-G_0^*(s))/(s\bar{B}_0)$ is the Laplace transform of the probability density function $(1-G_0(x))/\bar{B}_0$. We thus obtain

$$\rho(x)z\beta^*(z,x) = \frac{\bar{S}z}{z-1}f_0\left(\sum_{i=1}^N x_i\lambda_i(\beta_i(z)-1)\right) \quad \text{with } f_0(s) = s\left(1 + \frac{\rho_0}{1-\rho_0}\overline{G}_0^*(s)\right). \tag{5}$$

Note that the function f_0 only depends on the characteristics of the high priority $M/G/1$ queue.

At this stage we have to make an assumption on the tail behavior of $F_0(x)$. Define

$$S = \sup\{s : F_0^*(-s) < \infty\} \quad \text{and} \quad T = \sup\{t : G_0^*(-s) < \infty\}. \tag{6}$$

In the following, we assume that $S > 0$ and, to avoid trivialities, that $x \neq 0$.

Since the power series $\beta_i(z)$, $1 \leq i \leq N$, have infinite convergence radii, equation (5) clearly shows that the convergence radius $R(x)$ of $\beta^*(z, x)$ (or equivalently of $\beta(z, x)$) is the unique real number $z > 1$ such that:

$$\sum_{i=1}^{N} x_i \lambda_i (\beta_i(z) - 1) = T \Rightarrow \rho(x) R(x) \beta^*(R(x), x) = \frac{\bar{S} R(x)}{R(x) - 1} f_0(T) \tag{7}$$

where we have used equation (5) in the last step.

In the following, we set $C = \bar{S}(1/\varepsilon)^{1/K'} f_0(T)/((1/\varepsilon)^{1/K'} - 1)$. To use the upper-bounds for the loss probability established in Section 3, we have to consider the two cases $C > 1$ and $C < 1$ separately. We first consider the case $C < 1$, which entails that $\bar{S} f_0(T) < 1$. In the following, we formulate the results in an RSR approximation style. As it will clearly appear from the results, the equivalent bandwidths obtained by using the upper bounds derived in Section 3 are closed to those obtained via an RSR approximation.

Theorem 1 *Assume that* $C < 1$ *and let* $K' = (1 - \kappa)K$. $\pi_{loss}(K) \leq \varepsilon$ *if* $\sum_{i=1}^{N} x_i a_i(\varepsilon) \leq 1 - \rho_0$, *where the quantities* $a_i(\varepsilon)$, *referred to as equivalent bandwidths, are given for* $1 \leq i \leq N$ *by*

$$a_i(\varepsilon) = (1 - \rho_0) \frac{\lambda_i}{T} \left(\beta_i \left((1/\varepsilon)^{1/K} \right) - 1 \right). \tag{8}$$

Proof. Under the assumption $C < 1$, $R(x) > (1/\varepsilon)^{1/K'}$ and $\rho(x) R(x) \beta^*(R(x), x) < 1$. From Proposition 1, $\pi_{\text{loss}}(K) \leq \varepsilon$. From the definition of the convergence radius and equation (7), the condition $R(x) \geq (1/\varepsilon)^{1/K'}$ is equivalent to

$$\sum_{i=1}^{N} x_i \lambda_i \left(\beta_i \left((1/\varepsilon)^{1/K'} \right) - 1 \right) \leq T.$$

This completes the proof. □

Let us now consider the case when $C > 1$. Before stating a result similar to Theorem 1, let us note the following property whose proof is given in [**?**]

$$\sum_{i=1}^{N} x_i \lambda_i \left[\beta_i \left(\frac{1}{1 - \bar{S} f_0(T)} \right) - 1 \right] \leq T \Rightarrow \sum_{i=1}^{N} x_i \lambda_i \left[\beta_i \left(\frac{1}{1 - \bar{S} f_0(T)} \right) - 1 \right] \leq T \tag{9}$$

This allows us to state the following result.

Theorem 2 *Assume that $C > 1$ and $(1/\varepsilon)^{1/K'} < 1/(1 - \bar{S}f_0(T))^+$. $\pi_{loss}(K, x) \leq \varepsilon$ if $\sum_{i=1}^N x_i a_i(\varepsilon) \leq 1 - \rho_0$, where the effective bandwidths $a_i(\varepsilon)$ are given for $1 \leq i \leq N$ by*

$$a_i(\varepsilon) = \lambda_i(1-\rho_0)\frac{\beta_i\left((1/\varepsilon)^{1/K'}\right) - 1}{f_0^{-1}\left((1-\varepsilon^{1/K'})/\bar{S}\right)}. \tag{10}$$

Proof. Assume that $C > 1$. If the convergence radius $R(x)$ is such that

$$R(x) < 1/(1-\bar{S}f_0(T))^+ \Leftrightarrow \sum_{i=1}^N x_i\lambda_i\left[\beta_i\left(\frac{1}{(1-\bar{S}f_0(T))^+}\right) - 1\right] > T, \tag{11}$$

then $\rho(x)R(x)\beta^*(R(x), x) > 1$. From Proposition 1, the loss probability $\pi_{\text{loss}}(K) \leq \varepsilon$ if $z_0 \geq (1/\varepsilon)^{1/K'}$, which is equivalent to the condition $\rho(x)(1/\varepsilon)^{1/K'}\beta^*((1/\varepsilon)^{1/K'}, x) \leq 1$. In view of equation (5), this condition is equivalent to

$$f_0\left(\sum_{i=1}^N x_i\lambda_i(\beta_i((1/\varepsilon)^{1/K'}) - 1)\right) \leq \frac{(1/\varepsilon)^{1/K'} - 1}{\bar{S}(1/\varepsilon)^{1/K'}} = \frac{1-\varepsilon^{1/K'}}{\bar{S}}.$$

The function f_0 is continuous and increasing on $[0, T]$, with $f(0) = 0$ and $f_0(T) > (1 - \varepsilon^{1/K})/\bar{S}$ (by assumption). Hence, the above condition is equivalent to

$$\sum_{i=1}^N x_i\lambda_i(\beta_i((1/\varepsilon)^{1/K'}) - 1) \leq f_0^{-1}\left((1-\varepsilon^{1/K'})/\bar{S}\right). \tag{12}$$

Thus, we have proved that under the assumption $R(x) < 1/(1-\bar{S}f_0(T))^+$, $\pi_{\text{loss}}(K) \leq \varepsilon$ if inequality (12) is satisfied.

Assume now that $\bar{S}f_0(T) < 1$ and $R(x) > 1/(1-\bar{S}f_0(T))$. In that case, the convergence radius is such that $\rho(x)R(x)\beta^*(R(x), x) < 1$. Since $C > 1$, then $R(x) > 1/(1-\bar{S}f_0(T)) > (1/\varepsilon)^{1/K'}$ and hence, by Proposition 1, $\pi_{\text{loss}}(K) < \varepsilon$. From the definition of the convergence radius and equation (7), the condition $R(x) > 1/(1-\bar{S}f_0(T))$ is equivalent to inequality (9). But, from relation (9), this latter inequality implies inequality (12). Hence, this latter inequality is a dummy condition to ensure that $\pi_{\text{loss}}(K) \leq \varepsilon$ when $C > 1$ and $R(x) > 1/(1-\bar{S}f_0(T))$. This completes the proof. □

To conclude this section, let us consider the case when all packets have exponential service times, with mean $\bar{S}_0$ (resp. $\bar{S}$) for high (resp. low) priority packets. From the definition of the variables S and T, it is a matter of simple analysis to obtain that

$$\overline{G}_0^*\left((1-\rho_0)t/\bar{S}_0\right) = \frac{2}{1+t+\sqrt{(1+t)^2+4\rho_0 t/(1-\rho_0)}},$$

and

$$T = \frac{(1-\sqrt{\rho_0})^2}{\bar{S}_0}, \quad f_0(T) = \frac{1-\sqrt{\rho_0}}{\bar{S}_0}, \quad C = \frac{\bar{S}}{\bar{S}_0}\frac{1-\sqrt{\rho_0}}{1-\varepsilon^{1/K}}. \tag{13}$$

Trite calculations eventually lead to the surprisingly simple expression $f_0^{-1}(u) = u(1 - \rho_0/1-\bar{S}_0 u)$ for $u \leq (1-\sqrt{\rho_0})/\bar{S}_0$.

4.3. Admission control algorithm

In this section, we consider statistical multiplexing of leaky bucket regulated traffic sources in a buffer with finite capacity and we use the results of the previous sections to develop an admission control algorithm, which ensures that the packet rejection probability is less than a given threshold ε. We specifically assume that x_i sources of type i, characterized by the traffic parameters $(\sigma_i, \rho_i, \pi_i)$, $i = 1, \dots, N$, are multiplexed into a buffer with capacity $K < \infty$. The bucket sizes σ_i, $i = 1, \dots, N$ are upper bounded by the constant κK for some $\kappa \in (0,1)$.

Moreover, we assume that this buffer is served only when a high priority buffer is empty. This high priority buffer is fed with packets arriving according to a Poisson process. This Poisson assumption is supported by the fact that in real implementations, time priority is given to real time packet flows, which are usually with constant bit rates. The Poisson process is then a good approximation for the superposition of a large number of such real time flows.

Motivated by applications to ATM networks, we assume that all packets (high and low priority packets) have the same mean service time $\bar{S}$. In that case, the condition $C > 1$ is satisfied if $\varepsilon^{1/K'} > \sqrt{\rho_0}$. Since ε takes typical values of the form 10^{-n} for $n \in \{4, 5, .., 10\}$ and the reduced buffer capacity K' is several thousands of packets large, $\varepsilon^{1/K'}$ is close to one. Then, the condition $\varepsilon^{1/K'} > \sqrt{\rho_0}$ naturally holds. For instance, for $K = 3000$, $\kappa = 10\%$, and $\varepsilon = 10^{-7}$, this condition translates into $\rho_0 < 98.8\%$, which is a very loose requirement. Hence, in view of Theorem 2, we have the following result for defining effective bandwidths associated with the (σ, ρ, π)-regulated sources.

Theorem 3 *Consider that type i traffic sources, characterized by the traffic parameter sets $(\sigma_i, \rho_i, \pi_i)$, $i = 1, \dots, N$, are statistically multiplexed into a buffer with capacity $K < \infty$ in the presence of time priority buffer fed with a Poisson packet stream offering a load ρ_0. Under the assumptions that i) all packets have the same mean service $\bar{S}$, ii) the bucket sizes $\sigma_i \le \kappa K$ for $i = 1, \dots, N$ and some $\kappa \in (0,1)$, and iii) $\sqrt{\rho_0} < \varepsilon^{1/K'}$, the admissibility region $\mathcal{A}^p_\varepsilon$ (i.e., the set of N tuples $(x_1, \dots, x_N)$ such that $\pi_{loss}(K, x) \le \varepsilon$) is given by*

$$\mathcal{A}^p_\varepsilon = \left\{ (x_1, \dots, x_N) \in \mathbb{N}^\mathbb{N} : \sum_{i=1}^N x_i e_i(\varepsilon) \le 1 - \rho_0 \right\},$$

where the effective bandwidths $e_i(\varepsilon)$ are given by $e_i(\varepsilon) = \min\{\pi_i, a_i(\varepsilon)\}$ with

$$a_i(\varepsilon) = \frac{1-\rho_0}{1-(1/\varepsilon)^{1/K'}\rho_0} \rho_i (1/\varepsilon)^{1/K'} \beta_i^* \left((1/\varepsilon)^{1/K'} \right). \tag{14}$$

Since for typical values of K' and ε, $(1/\varepsilon)^{1/K'}$ is very close to one, we see from the above result that $a_i(\varepsilon)$ is equivalent to the effective bandwidth $a^r_i(\varepsilon)$ obtained by using the RSR approximation, namely by taking $\rho_0 = 0$ in the expression of $a_i(\varepsilon)$. Hence, the admissibility regions $\mathcal{A}^p_\varepsilon$ defined in Theorem 3 and $\mathcal{A}^r_\varepsilon$ corresponding to the effective bandwidths $a^r_i(\varepsilon)$ are almost equal.

5. COMPARISON WITH OTHER ALGORITHMS

In this section, we compare the admission control algorithm based on Theorem 3 when $\rho_0 = 0$ (i.e., there is no high priority traffic) with other admission control algorithms, which have been so far proposed in the literature for ATM networks. Thus, we consider VBR connections of different types $i = 1, \dots, N$; those of type i are characterized by a peak cell rate PCR_i, a sustainable cell rate SCR_i and a maximum burst size MBS_i (expressed in cells). We assume that the above VBR connections are statistically multiplexed in a buffer with size K and that for all $i = 1, \dots, N$, $\sigma_i \leq \kappa K$ for some $\kappa \in (0, 1)$. Finally, let $K' = (1 - \kappa)K$.

Algorithm by Elwalid, Mitra and Wentworth [2]. This algorithm defines an effective bandwidth $\tilde{e}_i$ according to the procedure described in the paper by Elwalid *et al* [2]. Some comparisons between the over-allocation factors $\tilde{\eta}_i = \tilde{e}_i / SCR_i$ and $\eta = a_i(\varepsilon)/SCR_i$ obtained when using the equivalent bandwidths $\tilde{e}_i$ and $a_i(\varepsilon)$, respectively are reported in Table 1. (The data are taken from the paper by Elwalid *et al* [2]). It turns out that the admission control developed by Elwalid *et al* performs better when the peak to mean ratio is rather small and the bucket size is large. However, when the peak to mean ratio is “sufficiently” large (say, greater than 10) and bucket sizes are “sufficiently” small when compared with the buffer capacity, then the admission control designed in this paper leads to smaller over-allocation factors.

		SCR_i	PCR_i	σ_i	$\tilde{e}_i$	$\tilde{\eta}_i$	η
service rate $C = 150$, $\varepsilon = 10^{-9}$,	type 1	0.15	6.0	24	0.24	60%	9%
$K = 3333$, and $\kappa = 10\%$	type 2	0.15	1.25	220	0.27	80%	135%
service rate $C = 90$, $\varepsilon = 10^{-9}$,	type 1	0.15	6.0	24	0.28	86%	16%
$K = 2000$, and $\kappa = 11\%$	type 2	0.15	1.25	220	0.32	113%	369%
service rate $C = 150$, $\varepsilon = 10^{-9}$,	type 1	0.5	3	83	0.83	66%	22%
$K = 5000$, and $\kappa = 10\%$	type 2	1.0	5	80	1.55	55%	21%

Table 1
Comparison between the over-allocation factors $\tilde{\eta}_i$ and η in different situations.

Algorithm by Gibbens and Hunt [3]. Gibbens and Hunt have used a fluid model with heterogeneous On/Off sources to define a closed form effective bandwidth $\hat{e}_i$ in [3], which can be written as $\hat{e}_i = (1 + \hat{\eta}_i)SCR_i$. The over-allocation factor $\hat{\eta}_i$ is compared, as a function of the bucket size σ, with the over-allocation factor η_i obtained when using the effective bandwidth $a_i(\varepsilon)$, for different values of the mean to peak rate ratio y_i, when $K = 3000$ and $\kappa = 10\%$. It turns out that Gibbens and Hunt admission algorithm performs slightly better when the ratio y_i is relatively large (e.g., $y_i = 0.5$). But, our admission control becomes better when y decreases, i.e., when the peak rate is much larger than the sustainable rate.

6. CONCLUSION

We have proposed in this paper a new admission control algorithm for allocating resources in broadband packet networks to packet streams, which are regulated by dual leaky buckets. This new algorithm relies on the analysis of the $M^{[X]}/M/1/K$ queueing system and can account for a priority level, which may be used for instance by real time traffic. It is easily checked that the same approach can be used for a system with more than two priority levels. Several numerical experiments have been carried out in order

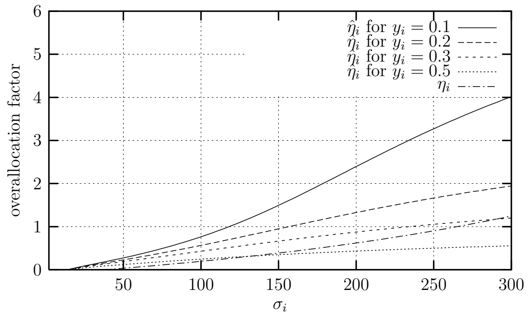

Figure 1. Comparison between the over-allocation factors η_i and $\hat{\eta}_i$ as a function of the bucket size σ_i and different values of the mean to peak rate ratio y_i – $K = 3000$, $\kappa = 10\%$.

to compare the algorithm presented in this paper and other admission control algorithms published earlier in the literature. It turns out that the present algorithm yields similar results, which are even a little bit better when the peak to mean ratio is sufficiently large and the mean rate is small when compared to the link transmission capacity. Yet, from a practical point of view, the two key advantages of the proposed algorithm when compared with others are in that (1) the effective bandwidths have simple closed form expressions and (2) the admissibility region is delimited by a hyper-plane.

REFERENCES

1. R.L. Cruz. A calculus of network delay, Part I : network elements in isolation. *IEEE Trans. Information Theory 37(3)*, January 1991.
2. 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 J. Sel. Area Commun.*, 13(6):1115–1127, August 1995.
3. R. Gibbens and P.J. Hunt. Effective bandwidthss for the multi-type UAS channel. *Queueing Systems*, pages 17–28, 1991.
4. A.W. Berger and W. Whitt. Effective bandwidths with priorities. *IEEE/ACM Trans. on Networking*, 6(4):447–460, 1998.
5. A.W. Berger and W. Whitt. Extending the effective bandwidth concept to networks with priority classes. *IEEE Communications Magazine*, 36(8):78–84, 1998.
6. M. Reisslein, K.W. Ross, and Srini Rajagopal. Guaranteeing statistical QoS to regulated traffic: the single node case. In *Proc. IEEE Infocom'99*, pages 1060–1071, New York, mar 1999.
7. S. Shakkottai and R. Srikant. Many sources delay asymptotics with applications to priority queues. *Queieng Systems*, 39:183–200, 2001.
8. G. Kesidis and T. Konstantopoulos. Worst case performance of a buffer with independent shaped arrival processes. *IEEE Comm. Lett.*, 4(1):26–28, 2000.
9. B. T. Doshi. Deterministic rule based traffic descriptors for broadband ISDN : worst case behavior and connection acceptance control. In *Proc. Globecom'93*, December 1993.
10. A. Gravey, J. Boyer, K. Sevilla, and J. Mignault. Resource allocation for worst case traffic in ATM networks. *Performance Evaluation*, 1-2:19–43, 1997.
11. H. Timjs and J.V. Ommeren. Asymptotic analysis for buffer behavior in communication systems. *Probability in the Engineering and Informational Sciences*, 3:1–12, 1989.
12. J.F.C. Kingman. Inequalities in the theory of queues. *Journal of the Royal Statistical Society B*, 32:102–110, 1970.

13. W. Feller. *An introduction to probability theory and its applications*, volume II. John Wiley & Sons Ltd, 1971.