Proceedings of the 35th
Conference on Decision and Control
Kobe, Japan • December 1996

Leaky Buckets : Sizing and Admission Control [1]

Vidyadhar G. Kulkarni
Department of Operations Research
The University of North Carolina
CB # 3180, Smith Bldg.,
Chapel Hill, NC 27599-3180 USA
vg_kulkarni@unc.edu

Natarajan Gautam
Department of Operations Research
The University of North Carolina
CB # 3180, Smith Bldg.,
Chapel Hill, NC 27599-3180 USA
gautam@or.unc.edu

Abstract

In this paper we propose a new unified approach to address simultaneously the issues of traffic policing (using leaky buckets), admission control and network dimensioning. First we compute the effective bandwidth of the output of the leaky bucket. We find a surprising result that effective bandwidth of the output of a leaky bucket is independent of the token pool size. Furthermore, the output effective bandwidth exhibits discontinuous behavior as the token pool size approaches infinity. We use these results in an optimization model to find the "optimal" leaky bucket parameters. We explain how this optimization program can be used to do network dimensioning if the input traffic characteristics are known, or to do connection admission control if the network parameters are fixed.

1 Introduction

The emerging high speed networks are expected to carry a broad range of traffic (video, audio and data). In the networks using asynchronous transfer mode (ATM), each traffic-source is described by its stochastic characteristics, and is assured a quality of service (QoS), as measured by cell-loss probability, delay, delay-jitter, etc. The network dimensioning, traffic policing and connection admission control is done with the aim of guaranteeing a specified quality of service (QoS) to all the users.

The concept of effective bandwidth (also known as equivalent capacity) and its use in admission control for the statistical multiplexing of bursty sources is now well-documented and accepted (see [1], [8], and [3].) The effective bandwidth is a number associated with a traffic-source such that if the sum of the effective bandwidths of all the sources multiplexed onto a buffer is less than the output rate of that buffer, then the QoS is satisfied for each source.

The above methodology assumes that the source traffic will conform to its stated statistical properties. However, in practice we need a policing mechanism to make sure that a non-conforming source does not affect the network adversely. Due to the high data transmission speed of the network, one needs to use rate based open loop traffic policing mechanisms, rather than the "more optimal" feedback based mechanisms (see [4]). One common method of policing the user traffic is called the Leaky Bucket.

A leaky bucket is essentially a credit management mechanism that controls the traffic entering the network using two parameters γ and M. See Section 3 for details. We study the buffered leaky bucket in this paper, and refer the reader to [11] for the analysis of the unbuffered leaky bucket.

One of the important questions is how to set the parameters of the leaky bucket, viz., M and γ, and the network parameters, viz., buffer sizes and link speeds so that a quality of service to network users is assured. The leaky bucket sizing has been addressed in the literature by many researchers, see [13], [7], etc. Bandwidth allocation in general has also received considerable attention, see [2], [5], etc.

It is clear from the above literature that the problems of network design, bandwidth allocation and call admission control are treated separately so far. Here we present a combined approach that treats these problems simultaneously, thus continuing the approach taken in [7] by incorporating the more recent results on effective bandwidths.

In Section 2, we recapitulate some of the recent results on effective bandwidths and study the output from a leaky bucket in Section 3. These results are then used in an optimization model that is presented in Section 4. An efficient algorithm to solve the optimization problem is then described in Section 5. This section also presents numerical results illustrating many different uses of the optimization problem.

[1]This work was partially supported by NSF Grant No. NCR-9406823

2 Preliminary Results

In this section we recapitulate some of the more recent results on effective bandwidths from [1], [3], [8] and [9].

Consider a single buffer fluid model driven by a random environment process $\{Z(t), t \geq 0\}$. When the environment is in state $Z(t)$, the fluid enters the buffer at rate $r(Z(t))$. Let $X(t)$ be the amount of fluid in the buffer at time t. The buffer has infinite capacity and is serviced by a channel of constant output rate c.

In practice we want to bound the overflow probability of the finite buffer of size B by ϵ. This can be achieved if the infinite buffer content process satisfies

$$\lim_{t\to\infty} P(X(t) > B) < \epsilon. \tag{1}$$

Let $A(t)$ be the total amount of fluid input from the source to the buffer in time t. Thus

$$A(t) = \int_0^t r(Z(u))du.$$

The asymptotic log moment generating function (ALMGF), $h(v)$, is defined to be

$$h(v) = \lim_{t\to\infty} \frac{1}{t} \log E\{\exp(vA(t))\}. \tag{2}$$

In the asymptotic region
$B \to \infty$ and $\epsilon \to 0$, such that $-\log(\epsilon)/B \to \delta > 0$,
the inequality (1) is satisfied if

$$h(\delta)/\delta < c. \tag{3}$$

The quantity $h(\delta)/\delta$ is called the *effective bandwidth* of the input source.

We recapitulate the results for the ALMGF of a general on-off source using [9]. The successive on and off times (generically denoted by U and D, respectively) are independent. The total traffic generated during on time is rU (not necessarily at a constant rate), while no traffic is generated during off time. Let

$$e(u, v) = E\{e^{-u(U+D)+vrU}\} = \tilde{U}(u - rv)\tilde{D}(u),$$

where $\tilde{U}$ and $\tilde{D}$ are the Laplace Stieltjes Transforms (LSTs) of the on and off time distributions respectively. Define, for $v \geq 0$,

$$e^*(v) = \sup_{\{u>0: e(u,v)<\infty\}} \{e(u, v)\},$$

$$u^*(v) = \inf\{u > 0 : e(u, v) < \infty\}.$$

Then from [9], we get

(a) if $e^*(v) \geq 1$, $h(v)$ is a unique solution to $e(h(v), v) = 1$,

(b) if $e^*(v) < 1$, $h(v) = u^*(v)$.

We shall use the results of this section to study the ALMGF of the output of the leaky bucket in the next section.

3 Output of a Single Buffered Leaky Bucket

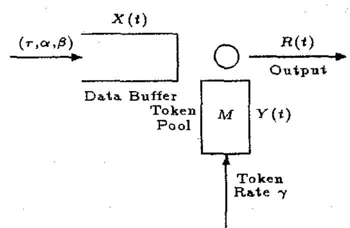

Figure 1.

Consider an exponential on-off source with on time parameter α and off time parameter β. Let $Z(t)$ be the state of the source at time t. When the source is on ($Z(t) = 1$), it generates traffic at rate r and at rate 0 when off ($Z(t) = 0$). The traffic from this source is input to the data buffer of a leaky bucket as shown in Figure 1. There is a token pool of size M into which tokens are generated continuously at a fixed rate γ. (The new tokens are discarded if the token pool is full.) Let $Y(t)$ be the number of tokens in the token pool at time t ($Y(t) \leq M$). If there is a token in the token pool, an incoming packet takes one and enters the network. If the token pool is empty, packets wait in the infinite capacity data buffer for tokens to arrive.

Let $X(t)$ be the amount of traffic in the data buffer at time t and define

$$W(t) = X(t) + M - Y(t).$$

A typical sample path of $\{W(t), t \geq 0\}$ is shown in Figure 2 (see [4]).

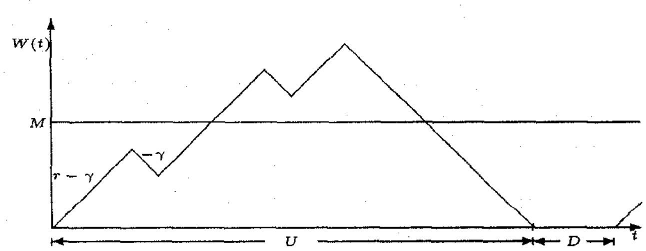

Figure 2.

The output rate from the leaky bucket at time t is given by

$$R(t) = \begin{cases} \gamma & \text{if } W(t) \geq M \\ r & \text{if } W(t) < M \text{ and } Z(t) = 1 \\ 0 & \text{if } W(t) < M \text{ and } Z(t) = 0. \end{cases} \tag{4}$$

Define the following :

$$U = \inf\{t > 0 : W(t) = 0 | W(0) = 0, Z(0) = 1\} \tag{5}$$

See Figure 2 for an illustration of U. Let V be the total amount of traffic output from the leaky bucket in time U. It is clear that $W(t) > 0$ and token pool is non-full during $(0, U)$. Hence the tokens enter the token pool at rate γ during $(0, U)$. Since the token pool is full at 0 and U, it is clear that the total number of tokens removed from the pool over $(0, U)$ must be the same as the total number of tokens that entered the pool over $(0, U)$. Hence we get

$$V = \gamma U. \tag{6}$$

Using this we compute the ALMGF of the output of the leaky bucket.

Theorem 1 *Let $D(t)$ be the total output from the leaky bucket over $[0, t]$. The ALMGF of the output of the leaky bucket*

$$h_D(v) = \lim_{t\to\infty} \frac{1}{t} \log E\{\exp(vD(t))\}$$

is given in terms of the ALMGF of the input, $h_A(v)$, as

$$h_D(v) = \begin{cases} h_A(v) & \text{if } 0 \le v \le v^* \\ h_A(v^*) - \gamma v^* + \gamma v & \text{if } v > v^*, \end{cases} \tag{7}$$

where

$$v^* = \frac{\beta}{r}\left(\sqrt{\frac{\gamma\alpha}{\beta(r-\gamma)}} - 1\right) + \frac{\alpha}{r}\left(1 - \sqrt{\frac{\beta(r-\gamma)}{\gamma\alpha}}\right)$$

and

$$h_A(v) = \frac{1}{2}\left(rv - \alpha - \beta + \sqrt{(rv-\alpha-\beta)^2 + 4\beta rv}\right).$$

Proof : The output process can be seen to be a general on-off source (as discussed in Section 2) with on time U (as in Eq. (5)), off time D (an exp(β) random variable), and total output during on time γU (see Eq. (6)). (Note that the output rate is not constant during on time). Let $\tilde{U}$ and $\tilde{D}$ be the LSTs of U and D respectively. Using the results from [12] we get $\tilde{D}(w) = \beta/(\beta + w)$, and

$$\tilde{U}(w) = \begin{cases} \frac{w+\beta+\gamma s_0(w)}{\beta} & \text{if } \bar{d} \ge 0 \\ \infty & \text{otherwise} \end{cases}$$

where $\bar{d} = b^2 + 4w(w + \alpha + \beta)\gamma(r - \gamma)$,
$s_0(w) = (-b - \sqrt{\bar{d}})/(2\gamma(r - \gamma))$, and
$b = (r - 2\gamma)w + (r - \gamma)\beta - \gamma\alpha$.
Hence from Section 2 and [10], we have the ALMGF of the output from the leaky bucket in terms of the ALMGF of the input as given in the theorem. ♠

Thus the ALMGF of the leaky bucket is identical to that of the output from a queue (with output rate γ), and is independent of M! This means that M does not play any role as far as reducing the effective bandwidth, but acts strictly as a policing device that prevents arbitrarily large peak-rate bursts from entering the network.

We would like to comment upon a curious discontinuous behavior at this point. Although the ALMGF of the output is related to that of the input as stated in Theorem 1 for all $M < \infty$, we have $h_D(v) = h_A(v)$ if $M = \infty$. This is because the $\{(W(t), Z(t)), t \ge 0\}$ process is transient if $M = \infty$, thus making the above analysis inapplicable. However, in that case, the leaky bucket is transparent and $D(t) = A(t)$ for all $t \ge 0$, thus making the two ALMGFs identical. In practice using $M = \infty$ is never a good idea, and hence this discontinuity will not bother us in the later analysis.

4 Sizing of Leaky Buckets and Admission Control

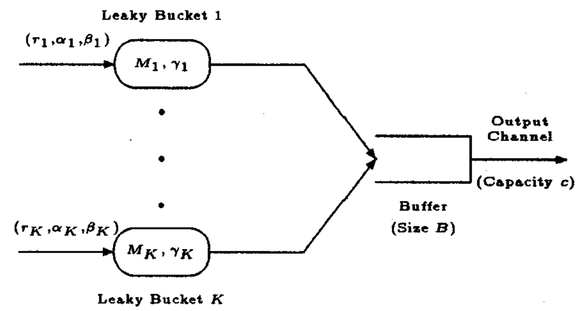

Figure 3.

We assume that there are K (a fixed positive integer) sources of traffic. Let r_i be the peak rate and m_i the mean rate of the i^{th} source. We assume that the i^{th} input source is policed by a leaky bucket with parameters γ_i and M_i. The output from the K leaky buckets is multiplexed onto a single buffer of size B and constant output rate c.

4.1 The Optimization Problem

We now conceptually formulate the problem of selecting the parameters M_i, γ_i, $(1 \le i \le K)$. We consider the following issues arising from the contract between the sources and the network:

- As long as the i^{th} source adheres to its agreed upon characteristics, the fraction of the traffic that faces a delay of more than a fixed amount d_i^* is bounded above by ζ_i.
- The fraction of the traffic from the i^{th} source that is discarded by the network due to buffer overflows is bounded above by ϵ_i. (This is called the cell loss probability constraint.)

Now consider the following optimization problem:

P: Minimize $\sum_{i=1}^{K} \gamma_i$

Subject to:

1. **Waiting time constraint at the leaky bucket,**
2. **$m_i < \gamma_i \leq r_i$, for $1 \leq i \leq K$.**
3. **$\sum_{i=1}^{K} M_i \leq M^*$.**

The objective function and constraint (3) need some explanation. The parameter γ_i of the i^{th} leaky bucket serves the following function: No matter how badly the source behaves, the data rate in the arbitrarily long bursts that it can send into the network is bounded above by γ_i. Thus if all sources were to simultaneously misbehave, the network will get traffic at maximum rate $\sum_{i=1}^{K} \gamma_i$. Hence it make sense to ensure that this worst case situation is kept the best possible. Similarly, the parameter M_i can be thought of the largest instantaneous burst that the leaky bucket will allow from the i^{th} source. Thus if M^* is the largest burst that the network can handle, (for example we may set $M^* = B$ or $M^* = B/2$.) then it makes sense to add constraint number (3). It will be seen that this turns out to be a very crucial constraint.

Constraint (1) is self explanatory, arising out of the contract stipulations. Constraint $m_i < \gamma_i$ is needed for stability, and $\gamma_i \leq r_i$ is needed to keep the the leaky bucket operation non-trivial. The quantitative expression for constraint (1), to be derived later, is valid only in the range $m_i < \gamma_i \leq r_i$. Note that the packet loss constraint in the network is not explicitly included in the above optimization problem. We show next how this constraint can be handled.

4.2 Uses of the Model

• Determining the channel capacity

Suppose we want to design a network to handle a given set of sources. We first choose B to satisfy budget constraint or the maximum delay constraint. Then we solve **P** for these given sources and this buffer size B, and obtain the optimal values of γ_i and M_i for $1 \leq i \leq K$.

Now, let $D_i(t)$ be the total output from the i^{th} leaky bucket over time $[0, t)$. Using the optimal parameters γ_i and M_i, use the results of Section 3 to obtain the ALMGF $h_i(\delta)$ of $D_i(t)$. Consider the case when all $\epsilon_i = \epsilon$. (This means all sources require the same Quality of Service.) Then it is known that the QoS criterion is satisfied if

$$\sum_{i=1}^{K} \frac{h_i(\delta)}{\delta} < c$$

where $\delta = -\log(\epsilon)/B$. (See [1], [8].) Then the capacity needed to satisfy the packet loss constraints constraints is greater than

$$c^* = \sum_{i=1}^{K} \frac{h_i(\delta)}{\delta}.$$

• Call admission control

Suppose the capacity c and the buffer size B are given. Suppose we have K sources requesting service. We first solve **P** and compute the optimal parameters γ_i and M_i. Using these we compute $h_i(\delta)$ as in the previous paragraph. Let

$$c^* = \sum_{i=1}^{K} \frac{h_i(\delta)}{\delta}.$$

If $c^* < c$ then we can admit all the sources, otherwise some will have to be denied access. Furthermore, if we have already admitted K sources, and a new $K+1$ st source arrives, we resolve **P** and compute the new optimal parameters M_i and γ_i ($1 \leq i \leq K+1$), and compute

$$c^* = \sum_{i=1}^{K+1} \frac{h_i(\delta)}{\delta}.$$

If $c^* < c$ we can admit the new source and reset the leaky bucket parameters of all the existing sources. Note that this dynamic change should be transparent to the users, since their service contract will continue to be honored.

5 Algorithms and Numerical Results

In this section we study the optimization problem **P** for the leaky bucket under the assumption that the i^{th} source is an exponential on-off source with on time parameter α_i and off time parameter β_i. The peak rate is r_i and the mean rate $m_i = \frac{r_i \beta_i}{\alpha_i + \beta_i}$. We assume that $d_i^* = 0$, for all i. Then delay constraint at the input buffer can be stated using the steady state analysis of the $\{(W(t), Z(t)), t \geq 0\}$ process as (see [7])

$$e^{-\theta_i M_i} \leq \zeta_i, \text{ for } i = 1, 2, \ldots, K,$$

where

$$\theta_i = \frac{r_i(\gamma_i - m_i)\alpha_i}{(r_i - \gamma_i)(r_i - m_i)\gamma_i}, \tag{8}$$

and ζ_i is the expected fraction of traffic from source i that experience a non-zero delay. Then **P** can be restated as

$$\textbf{PB:} \qquad \min \left\{ \sum_{i=1}^{K} \gamma_i \right\},$$

subject to the constraints,

$$\begin{aligned} e^{-\theta_i M_i} &\leq \zeta_i, \text{ for } i = 1, 2, \ldots, K, \\ M_1 + M_2 + \ldots + M_K &\leq B, \\ \frac{r_i \beta_i}{\beta_i + \alpha_i} < \gamma_i &\leq r_i, \text{ for } i = 1, 2, \ldots, K. \end{aligned}$$

Note that the constraint

$$\frac{r_i \beta_i}{\beta_i + \alpha_i} < \gamma_i$$

would be automatically satisfied as $\gamma_i = m_i$ will imply that $M_i = \infty$ which is not possible. Hence we can drop the above constraint. Also note that the larger the M_i, the smaller the γ_i and hence the constraint

$$M_1 + M_2 + \ldots + M_K \leq B$$

will be binding. Therefore the optimality problem can be equivalently stated as follows:

$$\textbf{PB1:} \qquad \min \left\{ \sum_{i=1}^{K} \gamma_i \right\},$$

subject to the constraints,

$$\begin{aligned} e^{-\theta_i M_i} &\leq \zeta_i, \text{ for } i = 1, 2, \ldots, K, \\ M_1 + M_2 + \ldots + M_K &= B, \\ \gamma_i &\leq r_i, \text{ for } i = 1, 2, \ldots, K. \end{aligned}$$

To solve **PB1**, define the Lagrangian

$$L(\gamma_1, \ldots, \gamma_K, M_1, \ldots, M_K) = \sum_{i=1}^{K} \lambda_i (e^{-\theta_i M_i} - \zeta_i)$$

$$+\mu(\sum_{i=1}^{K} M_i - B) + \sum_{i=1}^{K} \nu_i(\gamma_i - r_i) + \sum_{i=1}^{K} \gamma_i,$$

where λ_i's, μ and ν_i's are associated Lagrange multipliers. Using the standard Lagrangian approach, it is easy to reduce the first order necessary conditions for optimality to

$$\sum_{i=1}^{K} -\log(\zeta_i)/\theta_i = B,$$

$$\left\{ \begin{array}{c} \frac{m_i(r_i - m_i)^2}{r_i \alpha_i (\gamma_i - m_i)^2} + \frac{r_i - m_i}{r_i \alpha_i} = \frac{-1}{\mu \log(\zeta_i)}, \nu_i = 0, \gamma_i < r_i, \\ \text{or} \\ \lambda_i = 0, M_i = 0, \gamma_i = r_i, \mu_i \nu_i < 0. \end{array} \right. \tag{9}$$

Using the above conditions, we derive the following algorithm to solve **PB1**. First note that as γ_i varies from m_i to r_i in Eq. (9), μ varies from 0 to $-\alpha_i/\log(\zeta_i)$. Thus for a fixed $\mu \in (0, \max_i\{-\alpha_i/\log(\zeta_i)\})$ one can solve Eq. (9) to get

$$\gamma_i = \hat{\gamma}_i = m_i + \sqrt{\frac{x}{y}}$$

where

$$x = \frac{m_i(r_i - m_i)^2}{r_i \alpha_i}, \text{ and}$$

$$y = -\frac{r_i - m_i}{r_i \alpha_i} - \frac{1}{\mu \log(\zeta_i)}.$$

Let

$$\gamma_i(\mu) = \min\{\hat{\gamma}_i, r_i\}.$$

It can be seen that $\gamma_i(\mu)$ is a monotone function of μ. This fact is used in the following algorithm

1. Using binary search over $\mu \in (0, \max_i\{-\alpha_i/\log(\zeta_i)\})$, solve
$$\sum_{i:\gamma_i(\mu) < r_i} -\log(\zeta_i)/\theta_i = B,$$
where θ_i is computed by using $\gamma_i = \gamma_i(\mu)$ in (8). Let the final value of μ be μ^*.
2. Set $\gamma_i^* = \gamma_i(\mu^*)$, and $M_i^* = -\frac{\log(\zeta_i)}{\theta_i^*}$, where θ_i^* is computed by using $\gamma_i = \gamma_i(\mu^*)$ in (8).

It is easy to show that the optimal solution corresponds to the objective function being minimized. We illustrate the algorithm by means of two numerical examples.

Example 1: Consider the case of K *iid* sources with common parameters α, β and r. It is easy to observe that $M_i^* = M^*$ and $\gamma_i^* = \gamma^*$ for all i, as given below:

$$M^* = \frac{B}{K},$$

$$\gamma^* = \frac{-a_2 + \sqrt{a_2^2 - 4a_1 a_3}}{2a_1},$$

where $a_1 = r - m$, $a_2 = \frac{\alpha r B}{-K\zeta} - r(r - m)$ and $a_3 = m\alpha Br/K \log(\zeta)$.
Therefore, using γ^* and M^*, the ALMGF $h(\delta) = h_i(\delta, \gamma^*, M^*)$ can be obtained from Theorem 1. The QoS criteria for cell-loss probability is satisfied if

$$K\, h(\delta)/\delta < c. \tag{10}$$

Thus in the design problem we should choose c to satisfy the above; while in the call admission problem we should keep admitting calls as long as the number of calls K in the system satisfies (10).

As a numerical example, consider the following input parameters: $\alpha = 1$, $\beta = 0.4$, $r = 1.2$, $\epsilon = 10^{-7}$, $\zeta = 0.003$, and $K = 27$. Then the optimal leaky bucket parameters are $\gamma^* = 1.1711$ and $M^* = 0.1695$, which yields $h(\delta)/\delta = 0.7360$. Thus Eq. (10) is satisfied if $c > 27 * 0.7360 = 19.87$. Notice that the optimal leaky bucket parameters, and hence the value of $h(\delta)/\delta$ will change with K.

Example 2: Suppose there are two types of sources. There are K_1 *iid* type 1 sources with $\alpha_1 = 2.4$, $\beta_1 = 0.4$, $r_1 = 2.0$ and $\zeta_1 = 10^{-5}$, and K_2 *iid* type 2 sources with $\alpha_2 = 1$, $\beta_2 = 0.4$, $r_2 = 1.2$ and $\zeta_2 = 0.003$. The network buffer has capacity $B = 10$, $\epsilon = 10^{-7}$ for both types of sources. It is clear that the optimal leaky bucket parameters will depend only on the type of the source. For a given (K_1, K_2), we can obtain the optimal γ_1^*, γ_2^*, M_1^* and M_2^* by solving **PB1**. The effective bandwidth of the output of a type i source will be $h_i(\delta)/\delta$. The QoS criterion is then

$$K_1 \frac{h_1(\delta)}{\delta} + K_2 \frac{h_2(\delta)}{\delta} < c.$$

If this is satisfied we say that the pair (K_1, K_2) is feasible. Table 1 gives the values of γ_1^*, γ_2^*, M_1^*, M_2^*, $h_1(\delta)/\delta$, and $h_2(\delta)/\delta$ for the pairs $\{(K_1, K_2) : 1 \leq K_1 \leq 5, 1 \leq K_2 \leq 6\}$. They are obtained using the algorithm mentioned in the previous subsection. This table can be used for both the design problem as well as admission control problem as follows.

K_1 / K_2	1		2		3		4		5	
1	0.86	0.79	1.17	1.04	1.36	1.19	1.51	1.2	1.6	1.2
	7.01	2.99	4.50	1.00	3.31	0.06	2.50	0	2.0	0
	0.68	0.67	0.80	0.74	0.84	0.74	0.85	.74	0.85	.74
2	0.99	0.89	1.21	1.07	1.36	1.19	1.51	1.2	1.6	1.2
	5.85	2.07	4.22	0.78	3.30	0.05	2.50	0	2.0	0
	0.73	0.71	0.81	0.74	0.84	0.74	0.85	.74	0.85	.74
3	1.07	0.95	1.24	1.09	1.37	1.19	1.51	1.2	1.6	1.2
	5.23	1.59	4.04	0.64	3.29	0.04	2.50	0	2.0	0
	0.76	0.73	0.81	0.74	0.84	0.74	0.85	.74	0.85	.74
4	1.12	1.00	1.26	1.11	1.37	1.19	1.51	1.2	1.6	1.2
	4.85	1.29	3.92	0.54	3.29	0.03	2.50	0	2.0	0
	0.78	0.73	0.82	0.74	0.84	0.74	0.85	.74	0.85	.74
5	1.16	1.03	1.28	1.12	1.37	1.19	1.51	1.2	1.6	1.2
	4.60	1.08	3.83	0.47	3.28	0.03	2.50	0	2.0	0
	0.79	0.74	0.82	0.74	0.84	0.74	0.85	.74	0.85	.74
6	1.19	1.05	1.29	1.13	1.37	1.19	1.51	1.2	1.6	1.2
	4.41	0.93	3.76	0.41	3.28	0.03	2.50	0	2.0	0
	0.80	0.74	0.82	0.74	0.84	0.74	0.85	.74	0.85	.74

Legend :

γ_1	γ_2
M_1	M_2
$\frac{h_1(\delta)}{\delta}$	$\frac{h_2(\delta)}{\delta}$

Table 1.

For example, suppose we want to be able to handle 4 sources of type 1 and 5 of type 2. Then for the pair $(4, 5)$ we see that the sum of the output effective bandwidths is $4 * 0.85 + 5 * 0.74 = 7.10$. Hence we must choose $c > 7.10$ in order to handle this traffic. On the other hand suppose $c = 6.2$ is given. Then the pair $(3, 4)$ is feasible if we use the optimal parameters from Table 1, however the pair $(3, 5)$ is infeasible. Thus the call admission can be done using a table like this. As a final example, all feasible pairs (K_1, K_2) are shown in the region R (including the boundary) of Figure 4 for the case of $c = 18.1$.

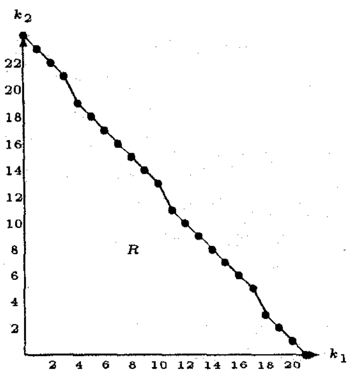

Figure 4.

References

[1] C. S. Chang and J. A. Thomas. *Effective Bandwidth in High-speed Digital Networks.* IEEE Journal on Selected areas in Communications, 13(6), 1091–1100, 1995.

[2] I. Cidon, I. Gopal and R. Guerin. *Bandwidth Management and Congestion Control in plaNET* IEEE Commun. Mag., 29, 54–63, 1991.

[3] G. de Veciana, G. Kesidis and J. Walrand. *Resource Management in Wide-Area ATM Networks Using Effective Bandwidths.* IEEE Journal on Selected Areas in Communications, 13(6), 1081–1090, 1995.

[4] A.I. Elwalid and D. Mitra. *Analysis and Design of Rate-Based Congestion Control of High Speed Networks, part I: Stochastic Fluid Models, Access Regulation.* Queueing Systems, Theory and Applications, Vol.9, 29–64, 1991.

[5] L. Gün and R. Guerin. *Bandwidth Management and Congestion Control Framework for the Broadband Network Architecture.* Computer Networks and ISDN Systems, 26, 61–78, 1993.

[6] R. Guerin, H. Ahmadi and Naghshineh. *Equivalent Capacity and its Application to Bandwidth Allocation in High-speed Networks.* IEEE J. Sel. Areas Commun.,SAC-9, 968–981, 1991.

[7] L. Gün, V.G. Kulkarni and A. Narayanan. *Bandwidth Allocation and Access Control in High-speed Networks.* Annals of Operations Research, 49, 161–183, 1994.

[8] G. Kesidis, J. Walrand, and C.S. Chang. *Effective bandwidths for Multiclass Markov Fluids and Other ATM sources.* IEEE/ACM Transactions on Networking, 1(4), 424–428, 1993.

[9] V. G. Kulkarni. *Effective Bandwidths for Markov Regenerative Sources.* Submitted to Queueing Systems, Theory and Applications, 1996.

[10] V. G. Kulkarni, and N. Gautam. *Admission Control of Multi-Class Traffic with Service Priorities in High-Speed Networks.* To appear in Queueing Systems, Theory and Applications, 1996.

[11] V. G. Kulkarni, and N. Gautam. *Leaky Buckets : Sizing and Admission Control.* Technical Report No. 95-10, Dept of Operations Research, University of North Carolina, Chapel Hill, NC 27599, 1996.

[12] A. Narayanan and V. G. Kulkarni. *First Passage Times in Fluid Models with an Application to Two-priority Fluid Systems.* To appear in Proceedings of the IEEE International Computer Performance and Dependability Symposium, 1996.

[13] K. Sohraby and M. Sidi. *On the Performance of Bursty and Modulated Sources Subject to Leaky Bucket Rate-based Access Control Schemes.* IEEE Trans. on Commun., Vol. 42, No. 2/3/4, 1994.