Burstiness Bounds for Some Burst Reducing Servers *

Steven Low

AT&T Bell Labs
600 Mountain Ave
Murray Hill, NJ 07974
slow@research.att.com

Pravin Varaiya

EECS Department
University of California
Berkeley, CA 94720
varaiya@helios.berkeley.edu

Abstract

In [18], we propose a framework to study a stream of traffic or message as it is transferred over an ATM connection. A message is modeled as a deterministic fluid flow, and an ATM node is modeled as a server which allocates bandwidth among concurrent messages. The key concept is that of the burstiness curve $b(\mu)$ of an incoming message which gives the buffer size needed if the message is served at rate μ. It is shown there that the fixed rate, the leaky bucket, and the affine server are burst reducing. This paper presents the burstiness bound for each of these servers. We also relate a sequence of identical affine servers to a fixed rate server, and suggest how 'quality of service' parameters may usefully be based on the burstiness curve.

1 Introduction

For our purposes, an ATM network provides a one-way connection from source to destination over a route spanning several nodes and transmission links. A connection is used to transport a traffic stream or *message.* Before transmission, a message is segmented into small, fixed-size cells. The bandwidth, or the number of cells per second (cps), allocated to a connection can vary at each link in its route and over time. Whenever the instantaneous rate of a message arriving at a node exceeds the allocated bandwidth, the excess cells are buffered.

We model a message m as a deterministic fluid flow with rate $m(t)$ cps, $0 \leq t \leq T$. As m arrives at a node it is processed and transformed into another fluid flow $m' = \Phi(m)$. The node is thus a traffic shaping device viewed abstractly as a function Φ which transforms arriving messages into outgoing messages. We neglect propagation and node processing delays, and so m' is the message which arrives at the next node in the route. If a message m originating at the source is processed in sequence by nodes $\Phi_1, ..., \Phi_k$ along its route, the message reaching the destination is

$$m' = \Phi_k \circ ... \circ \Phi_1(m)$$

and the entire route is simply another node $\Phi = \Phi_k \circ ... \circ \Phi_1$. The number of cells buffered at time t by node Φ is

$$b(t) = \int_0^t [m(s) - \Phi(m)(s)]ds$$

and so the node must provide buffers for $\max_t b(t)$ cells for processing m.

In [18], we define the *burstiness curve* $b_m(\mu)$ of a message m as the maximum number of cells that must be buffered at a node which transmits m at the fixed rate μ cps. Modeling a node Φ as a server which allocates bandwidth ϕ_m to each message m, we show there that two popular servers—the fixed rate and leaky bucket servers—are burst-reducing, and that the 'affine' server, $\phi_m(t) \equiv \lambda + \alpha(m(t) - \lambda)$, $\lambda \geq 0$, $0 \leq \alpha \leq 1$, is the only burst reducing memoryless server. We also show that a sequence of fixed rate servers is equivalent to a single bottleneck server.

In this paper, we extend these earlier results. In §2 we present the burstiness bound for each of these servers, and a monotonicity property of leaky bucket servers. In §3 we relate a sequence of identical affine servers to a fixed rate server. In §4 we suggest how 'quality of service' parameters may usefully be based on the burstiness curve. Proofs are omitted without comment and can be found in [21, 20].

Several sets of published work are relevant to the issues explored here. First, there are several clever schemes that implement service disciplines which enable nodes to allocate a specified amount of bandwidth

*Research supported by Pacific Bell, the MICRO program, and the NSF Grant IRI-9120074.

to each connection, despite variations in the burstiness and number of concurrent connections [14, 33, 11, 8, 13, 28, 31]. These schemes justify our characterization of a node as a server which allocates bandwidth.

Second, there is the work which models a message as a Markov-modulated fluid [3, 12, 15, 29, 22, 26, 9]. The concern is to study buffer requirements at a node serving one or several messages. The complexity of these models generally precludes study of what happens to a message as it travels over a sequence of nodes, as we are able to do in §3 (and in [18]) with the simpler, deterministic fluid model. Moreover, this work does not permit a direct study of the tradeoff between buffers and bandwidth as substitutable resources for serving a connection. Our approach is designed for such a study [17, 19].

Although we treat all nodes along a route in the same way, in practice the first node through which a source accesses the network may have a special traffic shaping or 'policing' function. The leaky bucket server once was proposed as a traffic shaping device to facilitate buffer allocation in packet-switched networks [27]. Subsequently, it was given a policing function in SMDS and Frame Relay networks. One justification of this server is that the interdeparture times of the cell stream has smaller variance or coefficient of variation than the interarrival times [10, 25, 16]. (See [18, 2, 21] for a different justification.) Several papers in *IEEE J. Selected Areas in Communications*, vol 9(3), April 1991, discuss the policing function.

The approach in [7, 6] also models a message m as a deterministic fluid; and m is said to satisfy the burstiness constraint (σ, ρ) if

$$\sup_{s\leq t} \int_s^t [m(r) - \rho]dr \leq \sigma \qquad (1)$$

This characterization was used to derive performance bounds [7, 6], and to analyze leaky bucket [5], round-robin polling systems [23, 24], and routing algorithms [5, 4]. The connection between the (σ, ρ) characterization and ours is apparent if we treat $\sigma = \sigma(\rho)$ as a function of ρ,

$$\sup_{s\leq t} \int_s^t [m(r) - \rho]dr = \sigma(\rho)$$

Then $\sigma(\rho)$ is the burstiness curve of m. The two parameter characterization (σ, ρ) corresponds to one point on the burstiness curve.

2 Burstiness Bounds

A *message* is a bounded, nonnegative function $m(t)$, $0 \leq t \leq T$, where $m(t)$ is the instantaneous rate in cells per second or cps, and $T < \infty$ is its duration. Suppose m is served by an infinite buffered server at a constant rate μ cps. The buffer is initially empty. The number of cells buffered at time t is

$$b_m(t) = \sup_{s\leq t} \int_s^t [m(r) - \mu]dr$$

and so the maximum buffer occupancy is

$$b_m(\mu) = \sup_{s\leq t\leq T} \int_s^t [m(r) - \mu]dr \qquad (2)$$

If the message name, m, is clear from the context, we write $b(\mu)$ instead of $b_m(\mu)$.[1]

A function $b(\mu)$, $\mu \geq 0$, which is nonnegative, convex, and strictly decreasing for $\mu < M$, with $b(M) = 0$, and $-db/d\mu(0) = T < \infty$, is called a burstiness curve. Let B be the set of all such curves. From [18, Proposition 1], $b_m(\mu) \in \mathsf{B}$ for all messages m, with $M := \sup_t m(t)$. We call $b_m(\mu)$ the burstiness curve of m.

The following construction proves to be useful in establishing the tightness of burstiness bounds below. Let $b \in \mathsf{B}$. Then

$$\tau(\mu) := -\frac{d}{d\mu}b(\mu), \ \mu \geq 0 \qquad (3)$$

is a nonnegative, decreasing function, and

$$b(\mu) = \int_\mu^M \tau(\lambda)d\lambda = \int_\mu^\infty \tau(\lambda)d\lambda, \ \mu \geq 0 \qquad (4)$$

Define the unimodal message[2]

$$m(t) = \inf\{\mu \mid \tau(\mu) \leq t\}, \ 0 \leq t \leq T \qquad (5)$$

Proposition 1 ([18, 21]) *Let $b \in \mathsf{B}$. The unimodal message m in (5) has $b_m(\mu) \equiv b(\mu)$.*

Our node model is an idealization of an output buffered ATM switch. Terminating at a node are several input and output links. A node has two parts. The first part is the switch fabric which, for each connection, routes cells coming over an input link to the

[1] We abuse notation and use $b(t)$ to denote the buffer occupancy at time t and $b(\mu)$ to denote maximum buffer occupancy at rate μ.

[2] A message m is *unimodal* if it has a single peak, i.e. $\{t \mid m(t) \geq \mu\}$ is an interval for every μ.

appropriate output link. The second part is the node processor. It operates as follows. The cells coming out of the switch are buffered separately for each output link. Each buffer is then read out according to the node's service discipline which determines how the link's bandwidth is shared among concurrent connections. Thus, each output link has its own buffers and bandwidth, and can be analyzed separately.

Fix attention on one output link occupied, say, by n concurrent connections. Then the service discipline of this link defines a map Φ,

$$\Phi(m_1, ..., m_n) = (m'_1, ..., m'_n) \quad (6)$$

where the messages m_i and m'_i represent the traffic on the ith connection coming into and leaving from the node. We assume that the switch fabric at each node has a sufficiently large throughput so that it is not a scarce resource. The only scarce resources are the buffers and bandwidth. The node is then effectively described by a collection of maps, $\{\Phi\}$, one per output link.

In this paper we study a single connection, so $n = 1$ in (6), and a node is a map Φ from an input message m to an output message $m' = \Phi(m)$. More concretely, Φ is determined by the bandwidth allocation $\phi(t) = \phi_m(t)$, $t \geq 0$, which may depend on the input message. The buffer occupancy $b(t) = b_m(t)$, $t \geq 0$, and $m' = \Phi(m)$ are given by

$$b(t) = \sup_{s \leq t} \int_s^t [m(r) - \phi(r)]dr \quad (7)$$

$$m'(t) = \begin{cases} \phi(t) & \text{if } b(t) > 0 \\ m(t) & \text{if } b(t) = 0 \end{cases}$$

We will interchangeably use Φ or ϕ to refer to a node or its service discipline.

We distinguish nodes which reduce bursts since they lead to better utilization of bandwidth and buffers. Let $b_i = b_{m_i}$, $i = 1, 2$. We say that b_1 is *less bursty* than b_2, or m_1 is *less bursty* than m_2, and we write $b_1 \prec b_2$ or $m_1 \prec m_2$, if

$$b_1(0) = b_2(0) \text{ and } b_1(\mu) \leq b_2(\mu), \ \mu \geq 0$$

Let M be the set of all messages. $\Phi : \mathsf{M} \rightarrow \mathsf{M}$ is *burst reducing* if for every m, $\Phi(m) \prec m$ or $b_{\Phi(m)} \prec b_m$. For any server $\Phi : \mathsf{M} \rightarrow \mathsf{M}$, define its *burstiness bound*, $\bar{\Phi} : \mathsf{B} \rightarrow \mathsf{B}$ by

$$\bar{\Phi}(b)(\mu) = \sup_{m \in M(b)} b_{\Phi(m)}(\mu)$$

where $M(b)$ is the set of all messages m with $b_m \prec b$. $\bar{\Phi}(b)$ is called a burstiness bound for b because any message m less bursty than b is burst-reduced after passing through the server Φ. Moreover, the burstiness of the output message $\Phi(m)$ is upper bounded by $\bar{\Phi}(b)$. The bound is said to be ***tight*** if for every $b \in \mathsf{B}$ there is $m \in M(b)$ such that $b_{\Phi(m)} \equiv \bar{\Phi}(b)$. We study three servers.

2.1 Fixed rate server

Proposition 2 *The fixed rate server, $\phi_m(t) \equiv \lambda$, is burst reducing. Its tight burstiness bound $\bar{\Phi}(b)$ is the largest function in* B *such that $\bar{\Phi}(b) \prec b$ and $\bar{\Phi}(b)(\lambda) =$* 0.

For each $b \in \mathsf{B}$, the unimodal message constructed in (5) achieves the burstiness bound. The burstiness bound is illustrated in Figure 1.

2.2 Leaky bucket server

The *leaky bucket* server is specified by two numbers $\lambda > 0$ and $\Gamma \geq 0$ [18]. It operates as follows. The source has a credit $\gamma(t)$, incremented at rate λ cps up to a ceiling of Γ cells, and decremented at the output rate. If $\gamma(t) > 0$, the output rate $m'(t)$ equals $m(t)$ and the buffer occupancy $b(t) = 0$. But if $\gamma(t) = 0$, the output rate is at most λ, and buffers will fill up if $m(t) > \lambda$.

Proposition 3 *The leaky bucket server $\Phi = (\lambda, \Gamma)$ is burst reducing. Its tight burstiness bound $\bar{\Phi}(b)$ is the largest function in* B *such that $\bar{\Phi}(b) \prec b$ and $\bar{\Phi}(b)(\lambda) \leq$* Γ.

Again, for each $b \in \mathsf{B}$, the unimodal message constructed in (5) achieves the burstiness bound. The burstiness bound is illustrated in Figure 2. When $\Gamma = 0$, a leaky bucket server reduces to a fixed rate server, and Proposition 3 reduces to Proposition 2.

Leaky bucket servers can be ordered according to their credit rate λ and ceiling Γ. The smaller is λ or Γ, the greater is the burst reduction and the larger is the backlog at the leaky bucket, as shown by the following result. A similar result in a stochastic setting on the monotonocity in Γ appears in [2].

Proposition 4 *Consider Figure 3 where $\Phi_i = (\lambda_i, \Gamma_i)$, $i = 1, 2$. If $\lambda_1 \leq \lambda_2$ and $\Gamma_1 \leq \Gamma_2$, then $c_1(t) \leq c_2(t)$ and $b_1(t) \geq b_2(t)$ for all m, t.*

From the proposition, $b_{\Phi_1(m)} \prec b_{\Phi_2(m)}$ for all m.

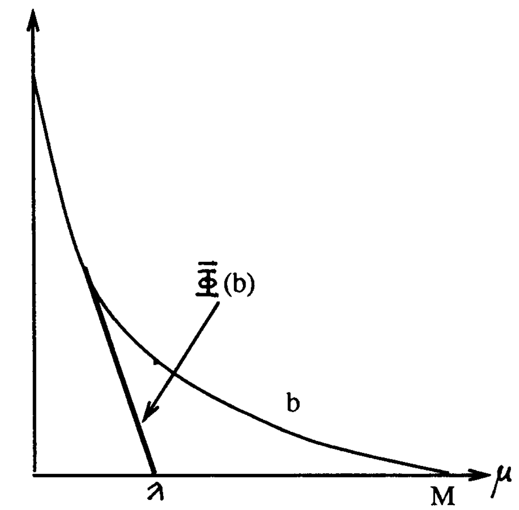

Figure 1: Tight burstiness bound for fixed rate server

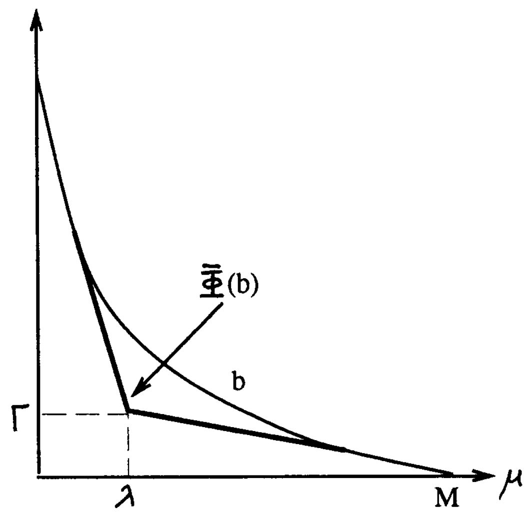

Figure 2: Tight burstiness bound for leaky bucket server

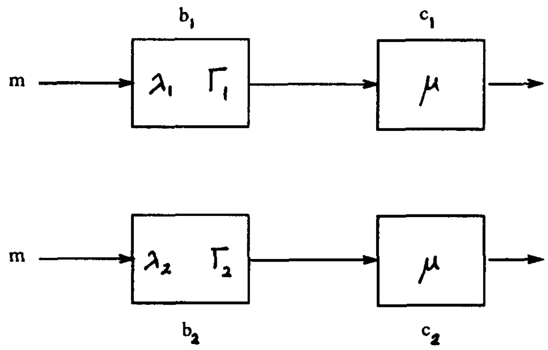

Figure 3: Ordering of leaky bucket servers

2.3 Affine Server

We say that ϕ or Φ is *memoryless* if $\phi_m(t) = \phi(m(t))$ for all $m, t \geq 0$. The fixed rate, but not the leaky bucket, server is memoryless. A more general memoryless server is the affine server ϕ^α,

$$\phi^\alpha(m) \quad = \quad \lambda + \alpha(m - \lambda) \qquad (8)$$

for some $\lambda \geq 0, 0 \leq \alpha \leq 1$. ($\phi^0$ is fixed rate.) An affine server may be implemented by basing the allocated bandwidth on an estimate of $m(t)$ obtained, for instance, by counting the cell arrivals in a short time interval as in Frame Relay networks [1] or as proposed in [32, 14].

Theorem 5 *Let ϕ be memoryless and nondecreasing. Then ϕ is burst reducing if and only if*

$$\phi(m) \quad = \quad \lambda + \alpha(m - \lambda) \qquad (9)$$

for some $\lambda \geq 0$, $0 \leq \alpha \leq 1$. Furthermore, ϕ given by (9) has the non-tight burstiness bound $\bar{\Phi}(b)$ which is the largest function in B such that $\bar{\Phi}(b) \prec b$ and $\bar{\Phi}(b)(\mu) \leq \alpha b(\mu)$, $\mu \geq \lambda$.

The following proposition is used to prove Theorem 7.

Proposition 6 *Consider the arrangement in Figure 4 where $\phi^\alpha(m) = \lambda + \alpha(m - \lambda)$. For $\mu \leq \lambda$, $b(t) + b_1(t) \equiv b_2(t)$ and $m_1(t) \equiv m_2(t)$.*

3 Single multi-hop connection

Consider a connection spanning nodes $\Phi_1, ..., \Phi_k$ in sequence as in Figure 5. We neglect propagation and processing delays. Then the overall map from source to destination is $m' = \Phi(m)$, where $\Phi = \Phi_k \circ ... \circ \Phi_1$. The burstiness bounds in §2 can be used to obtain worst case buffering requirement at each node when a message is transported across a sequence of nodes.

We consider a connection spanning k identical affine servers, $\phi_i(m) = \lambda + \alpha(m - \lambda)$. By Theorem 5, $\Phi \equiv \Phi_k \circ \circ \circ \Phi_1$ is burst reducing. Theorem 7 says that Φ is 'faster' than a single fixed rate server $\phi \equiv \lambda$, since $\sum b_i(t) = (1 - \alpha^k)b(t) < b(t)$.

Theorem 7 *Consider Figure 5 where $\phi_i(m) = \lambda + \alpha(m - \lambda)$, $i = 1, \ldots, k$, and $\phi_s(t) \equiv \lambda$.*

1. *$b_i(t) = (1 - \alpha)\alpha^{i-1}b(t)$, $i = 1, \ldots, k$, and*
2. *$m_{k+1}(t) \equiv (1 - \alpha^k)m''(t) + \alpha^k m_1(t)$.*

When $\alpha = 0$, Theorem 7 reduces to [18, Theorem 2] for fixed rate servers. It is in this sense that we say a sequence of fixed rate servers is equivalent to a bottleneck server.

4 Quality of service

We suggest an approach to service definition. We think of a service as a contract between the user and the network operator. A contract consists of two sets of 'quality parameters'. User parameters $(\bar{\eta}, \bar{b})$ bound the average rate and burstiness of a message; the network parameter Δ bounds the end-to-end delay.

We say that a user message $m(t), 0 \leq t \leq T$, is conformant if

$$\frac{1}{T}\int_0^T m(r)dr \leq \bar{\eta}, \text{ and } b_m(\mu) \leq \bar{b}(\mu), \mu \geq \bar{\eta} \qquad (10)$$

The first condition says that the average message rate is at most $\bar{\eta}$ cps. The second condition implies that if

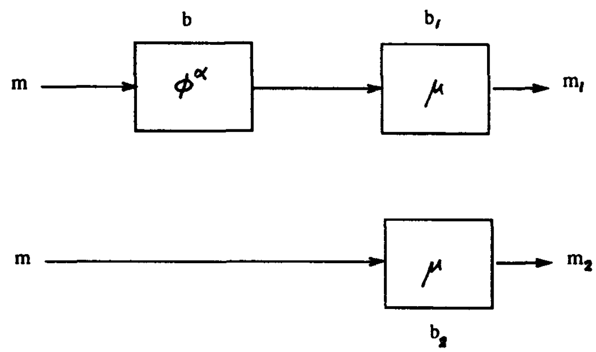

Figure 4: Affine server

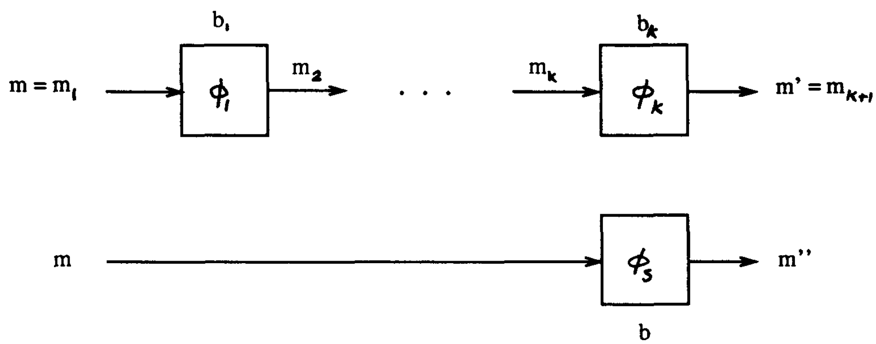

Figure 5: Single multi-hop connection over affine servers

1a.1.6

m is served at a bandwidth $\mu \geq \bar{\eta}$, then at most $\bar{b}(\mu)$ cells need to be buffered.

Suppose a conformant message m is transported by the network over a connection involving fixed rate buffers. Suppose the bottleneck rate is $\mu_s \geq \bar{\eta}$. From [18, Theorem 2], the end-to-end delay is less than Δ if

$$\Delta - \text{'propagation and processing delays'} \geq \frac{b_m(\mu_s)}{\mu_s} \quad (11)$$

and so the connection is conformant if this condition is satisfied.

Condition (10) can be enforced given the user's burstiness curve. The network can determine whether there is enough spare bandwidth and buffers to meet (11) at the time of connection setup. Further implications in this direction are studied in [17, 19].

5 Conclusion

[18] proposes a framework in which a message is a deterministic fluid flow and a node is a map which transforms an incoming message into an outgoing message. Such a map is more concretely defined as a service discipline which allocates bandwidth among concurrent messages. The usefulness of the framework is demonstrated by a study of the performance of fixed rate, leaky bucket, and affine servers, and, more interestingly, by the performance of a sequence of such nodes. This paper extends earlier results to present the burstiness bounds for these servers, and relate a sequence of identical affine servers to a fixed rate server.

The concepts of a message's burstiness curve, and of a node's burstiness bound are used to exhibit the tradeoff between bandwidth and buffers that is possible while ensuring that no cell loss occurs due to buffer overflow. Some applications, however, may be able to tolerate a certain amount of cell loss. A valuable extension of the proposed framework should exhibit the tradeoff between bandwidth, buffers and loss. An attempt in this direction is reported in [30].

References

[1] American National Standard for Telecommunications. *Addendum to T1.606 - Frame Relaying Bearer Service - Architectural Framework and Service Description*, 1990. ANSI T1S1/90-175R1.

[2] V. Anantharam and P. Konstantopoulos. Burst reduction properties of leaky buckets in ATM networks. *Proc. 29th Ann. Allerton Conf. Commun., Contr., and Computing*, October 1991.

[3] D. Anick, D. Mitra, and M. M. Sondhi. Stochastic theory of a data-handling system with multiple sources. *Bell System Technical Journal*, 61:1871–1894, 1982.

[4] M. C. Chuah. *Analysis of networks of queues via projection techniques.* PhD thesis, University of California, San Diego, June 1991.

[5] M. C. Chuah and R. L. Cruz. Approximate analysis of average performance of (σ, ρ) regulators. *Proceedings of Infocom'90*, pages 874–880, June 1990.

[6] Rene L. Cruz. A calculus for network delay Part II:network analysis. *IEEE Transactions on Information Theory*, 37(1):132–141, January 1991.

[7] Rene L. Cruz. A calculus for network delay Part I:network elements in isolation. *IEEE Transactions on Information Theory*, 37(1):114–131, January 1991.

[8] Alan Demers, Srinivasan Keshav, and Scott Shenker. Analysis and simulation of a fair queueing algorithm. *Proceedings of ACM SIGCOMM'89*, pages 3–12, 1989.

[9] Jr. E. G. Coffman, B. M. Igelnik, and Y. A. Kogan. Controlled stochastic model of a communication system with multiple sources. to appear in IEEE Transactions on Information Theory.

[10] A. I. Elwalid and Debasis Mitra. Analysis and design of rate-based congestion control of high speed networks, I: stochastic fluid models, access regulation. *Queueing Systems*, 9(1-2):29–63, October 1991.

[11] D. Ferrari and D. Verma. A scheme for real-time channel establishment in wide-area networks. *IEEE Journal on Selected Areas in Communications*, 8(3):368–379, April 1990.

[12] D. P. Gaver and J. P. Lehoczky. Channels that cooperatively service a data stream and voice messages. *IEEE Transactions on Communications*, COM-30(5):1153–1162, 1982.

[13] S. J. Golestani. Congestion-free communication in high-speed packet networks. *IEEE Transactions on Communications*, 39(12):1802–1812, December 1991.

[14] C. Kalmanek, H. Kanakia, and S. Keshav. Rate controlled servers for very high speed networks. *Proceedings of Globecom'90*, pages 12–20, December 1990.

[15] L. Kosten. Stochastic theory of a data-handling systems with groups of multiple sources. In H. Rudin and W. Bux, editors, *Performance of Computer-Communication Systems*, pages 321–331. Elsevier, Amsterdam, 1984.

[16] Lei Kuang. On the variance reduction property of buffered leaky bucket. Submitted to IEEE Transaction on Communications, 1991.

[17] S. Low and P. Varaiya. Optimal bandwidth and buffer allocations in ATM networks. *Proc. 29th Ann. Allerton Conf. Commun., Contr., and Computing*, October 1991.

[18] S. Low and P. Varaiya. A simple theory of traffic and resource allocation in ATM. *Proceedings of Globecom'91*, pages 1633–1637, December 1991.

[19] S. Low and P. Varaiya. Service provisioning in ATM networks. Submitted for publication, 1992.

[20] S. Low and P. Varaiya. Traffic shaping in ATM networks. Submitted for publication, 1992.

[21] Steven Low. *Traffic Management of ATM Networks: Service Provisioning, Routing, and Traffic Shaping*. PhD thesis, UC Berkeley, May 1992.

[22] D. Mitra. Stochastic theory of a fluid model of producers and consumers coupled by a buffer. *Adv. Appl. Prob.*, 20:646–676, 1988.

[23] Galen Sasaki. Input buffer requirements for round robin polling systems. *Proc. 27th Ann. Allerton Conf. Commun., Contr., and Computing*, pages 397–406, September 1989.

[24] Galen Sasaki. Input buffer requirements for round robin polling systems with nonzero switchover times. *Proc. 29th Ann. Allerton Conf. Commun., Contr., and Computing*, October 1991.

[25] Moshe Sidi, Wen-Zu Liu, Israel Cidon, and Inder Gopal. Congestion control through input rate regulation. *Proceedings of Globecom'89*, pages 1764–1768, December 1989.

[26] T. E. Stern and A. I. Elwalid. Analysis of a separable Markov-modulated rate models for information-handling systems. *Advances in Applied Probability*, 23(1):105–139, March 1991.

[27] Jonathan S. Turner. New directions in communications (or which way to the information age?). *IEEE Communications Magazine*, pages 8–15, October 1986.

[28] Dinesh C. Verma, Hui Zhang, and Domenico Ferrari. Guaranteeing delay jitter bounds in packet switching networks. *Proceedings of TriComm'91*, April 1991. Chapel Hill, North Carolina.

[29] Alan Weiss. A new technique for analyzing large traffic systems. *Adv. Appl. Prob.*, 18:506–532, 1986.

[30] Michael Wong and Pravin Varaiya. A deterministic fluid model for cell loss in ATM networks. *Proceedings of Infocom'93*, April 1993.

[31] Hui Zhang and Srinivasan Keshav. Comparison of rate-based service disciplines. *Proceedings of SigComm'91*, pages 113–121, Sept. 1991.

[32] Lixia Zhang. *A new architecture for packet switching network protocols*. PhD thesis, MIT, 1989.

[33] Lixia Zhang. Virtualclock: A new traffic control algorithm for packet-switched networks. *ACM Transactions on Computer Systems*, 9(2):101–124, May 1991.