A SIMPLE THEORY OF TRAFFIC AND RESOURCE ALLOCATION IN ATM

S. Low and P. Varaiya *
University of California, Berkeley CA 94720

August 19, 1991

Abstract

We propose a simple framework to study the tradeoffs between bandwidth and buffer in an ATM network. Traffic is modeled as a deterministic fluid. The key concept is the burstiness curve $b(\mu)$ which gives, for a connection, the maximum number $b(\mu)$ of cells that must be buffered if the connection is served at a fixed rate of μ cells per second. Two popular service disciplines - the fixed rate discipline and the leaky bucket discipline - are shown to be burst reducing. The characterization of a class of practical service disciplines that are burst reducing is obtained. We exhibit the tradeoffs between buffer and bandwidth along a connection and the options for resource allocation among simultaneous connections.

1 Introduction

For our purposes an ATM network provides a one-way connection between two end points over a virtual circuit. A virtual circuit is a path between its end points spanning several switching nodes. The virtual circuit is used to transport a message from one end point to the other. The message is packaged into small, uniformly sized units called cells. The bandwidth, i.e. number of cells per second or cps, allocated to a connection varies from node to node and over time. If at any time the rate of incoming cells exceeds the bandwidth allocated to the connection, the node buffers the excess cells. The bandwidth and buffers at a node are shared among the active connections according to the allocation rule at the node. The rule for allocating bandwidth is called the service discipline or the rate control policy; the rule for buffers is the buffer management scheme. This paper is concerned with bandwidth and buffer allocation.

Earlier papers were concerned with buffer management schemes that promote the savings from sharing while preventing a bursty connection from degrading performance [11, 8]. Typically, a node shared by n connections is modeled as n $M/M/1$ queueing systems that share a finite storage capacity. Buffer management schemes, ranging from complete partitioning to complete sharing, can be compared in terms of their blocking probability and server utilization. A recent work [2, 3] investigated buffer allocation in a large Jackson network and showed that allocating buffers in inverse proportion to the logarithm of the effective service rate at a node is close to optimal.

Attention has also been paid to the design of service disciplines that guarantee a fixed bandwidth to a connection regardless of changes in the number and burstiness of active connections, [10, 15, 13]. (See [14] for a recent survey.) This is motivated by the need to support real-time services [7]. A mathematical framework for deterministic traffic analysis is developed in [5], and the burstiness characterization there can be considered a special case of ours. [12] also adopts a deterministic approach to analyse transient behavior of virtual circuits; however, burstiness is not characterized.

Most previous papers do not focus as we do on the tradeoff between bandwidth and buffers as substitutible resources for providing service. The paper is organized as follows.

A message $m(t), t \geq 0$, is a non-negative function of time. In §2 we characterize the burstiness of a message as a function $b_m(\mu)$ that measrues its buffering requirement.

Consider a message m travelling over a virtual circuit through nodes $1, \cdots, k$. The service discipline at any node defines a mapping Φ from incoming message to outgoing message and so the end-to-end behavior is

*Address: EECS Dept., University of California, Berkeley, CA 94720, slow@eclair.berkeley.edu, varaiya@helios.berkeley.edu.
Work supported by Pacific Bell and the MICRO program.

given by the composite map $\Phi_k \circ \cdots \circ \Phi_1$ where Φ_i is the map defined by node i. In §3, we characterize service disciplines that are burst reducing and show that the fixed rate discipline and the leaky bucket discipline are burst reducing. §4 illustrates the tradeoff between bandwidth and buffers along a virtual circuit.

§5 exhibits the potential for gains from sharing the bandwidth at a node among several connections. The simplest bandwidth allocation rules allocate a fixed amount of resources to each connection. The most complex rules assume advanced knowledge of all connections. Savings in resources increase as we go from the simplest to the most complex rules.

2 The burstiness curve of a message

A message is a bounded, non-negative function $m(t)$, $0 \leq t \leq T$, where $m(t)$ is the instantaneous rate of the message at time t, in cells per second or cps, and $T < \infty$ is its duration. Suppose the cells in the message are stored in an infinite buffer read out by a server at a fixed rate μ cps. The buffer is initially empty. The buffer occupancy or backlog at time t in cells is

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

and so the maximum buffer occupancy is

$$b_m(\mu) = \sup_{0 \leq s \leq t \leq T} \int_s^t [m(r) - \mu] dr$$

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

Proposition 1 *$b(\mu)$ is nonnegative, convex, and strictly decreasing for $\mu < M = \sup_t m(t)$. Also, $b(\mu) = 0$ for $\mu \geq M$.*

It will be useful to distinguish functions with properties described in this proposition.

Definition 1 *A function $b(\mu)$, $\mu \geq 0$, which is nonnegative, convex, and strictly decreasing for $\mu < M$, with $b(M) = 0$, and $-\frac{d}{d\mu}b(0) = T < \infty$, is called a burstiness curve. Let $\mathcal{B}$ be the set of all such curves.*

By Proposition 1, $b_m \in \mathcal{B}$ for every message m. The next result is a sort of converse. A message m is *unimodal* if it has a single peak, i.e., if $\{t \mid m(t) \geq \mu\}$ is an interval for every μ.

Proposition 2 *Let $b \in \mathcal{B}$. There is a unimodal message m with $b_m(\mu) = b(\mu)$, for all μ.*

Messages can be partially ordered according to their burstiness.

Definition 2 *Let $b_i = b_{m_i}$, $i = 1, 2$, be two burstiness curves. b_1 is less bursty than b_2, or m_1 is less bursty than m_2, denoted $b_1 \prec b_2$ or $m_1 \prec m_2$, if*

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

The next result shows that multiplexing saves resources.

Proposition 3 *Let $b_i = b_{m_i}$, $i = 1, 2, \ldots$. Let $\alpha_i > 0$, with $\sum \alpha_i = 1$, $m(t) = \sum \alpha_i m_i(t)$ and $b = b_m$. Then*

$$b \prec \alpha_1 b_1 + \alpha_2 b_2 + \ldots$$

For any message $\bar{m}$ with burstiness curve $\bar{b}(\mu)$, consider the set M of all messages less bursty than $\bar{m}$,

$$M = \{m \mid m \prec \bar{m}\}.$$

Every m in M contains the same number of cells, namely $\bar{b}(0)$. M is a subset of the Banach space $\mathcal{L}^1[0, \infty)$. $m \in M$ is an *extreme* point of M if it cannot be expressed as a proper convex combination

$$m = \alpha m_1 + (1 - \alpha) m_2,\ 0 < \alpha < 1$$

of two distinct points m_1 and m_2 in M. If b_1 and b_2 are in $\mathcal{B}$ with $b_1(0) = b_2(0)$, then $b_1 \wedge b_2$ denotes the largest b in B such that $b \prec b_1$ and $b \prec b_2$.

Theorem 1 *To each burstiness curve b is associated a convex set of messages*

$$M(b) = \{m \mid b_m(0) = b(0),\ b_m \prec b\}.$$

An extreme point m of $M(b)$ is either an unimodal message with $b_m \equiv b$ or is such that $m(t)$ equals 0 or $M := \sup m(t)$. Every m in $M(b)$ can be expressed as a sequential limit $m = \lim m_n$ in the weak $\mathcal{L}^1$-topology, where the m_n are finite convex combinations of the extreme points of $M(b)$. Finally, $M(b_1 \wedge b_2) = M(b_1) \cap M(b_2)$.

3 A Node

A node consists of a switch followed by output buffers. Each outgoing link has its own buffers and bandwidth, and so its service discipline may be specified and analysed separately.

The traffic on a virtual circuit arriving at a node over an incoming link is some message m. The traffic on the same virtual circuit leaving the node over an outgoing link is message m'. Fix attention on a particular outgoing link and suppose n virtual circuits occupy this link. Then the service discipline of that link can be abstractly defined as a mapping Φ,

$$\Phi(m_1, ..., m_n) = (m'_1, ..., m'_n) \tag{1}$$

where the messages m_i and m'_i represent incoming and outgoing messages on the ith virtual circuit. The node is described by a collection of service discipline maps, Φ, one for each outgoing link.

In this section we are concerned with a single connection, so $n = 1$ in (1) and a node involved in this connection is defined as a mapping Φ from m to $m' = \Phi(m)$. Φ is determined by the service allocation $\phi(t) = \phi_m(t)$, $t \geq 0$, which may depend on the incoming message, and the corresponding buffer occupancy $b(t) = b_m(t)$, $t \geq 0$, given by

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

The outgoing traffic $m' = \Phi(m)$ is then given by

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

From the viewpoint of resource conservation it is desirable that the outgoing traffic be less bursty than the incoming traffic. Formally, a mapping Φ is *burst reducing* if $\Phi(m) \prec m$ for every message m.

Proposition 4 *The fixed rate scheme $\phi_m(t) \equiv \lambda$ for all m is burst reducing.*

The *leaky bucket* discipline is specified by two numbers $\lambda > 0$ and $\Gamma \geq 0$. The input source is given a credit $\gamma(t)$ incremented at the fixed rate λ cps up to a ceiling of Γ cells and decremented at the output rate. Thus, if $\gamma(t) > 0$, the output $m'(t)$ equals the input $m(t)$ and the buffer occupancy $\beta(t) = 0$. But if $\gamma(t) = 0$, the output rate is at most λ, and buffers will fill up if $m(t) > \lambda$. Since $\gamma(t)$ and $\beta(t)$ can never simultaneously be strictly positive, we can track both quantities by the single real variable

$$b(t) = \beta(t) - \gamma(t).$$

Using this variable, the leaky bucket discipline is given by

$$\begin{aligned} \frac{d}{dt}b(t) &= [m(t) - \lambda]1[b(t) \geq -\Gamma] \\ m'(t) &= m(t)1[b(t) \leq 0] + \lambda 1[b(t) > 0] \end{aligned}$$

Proposition 5 *The leaky bucket scheme is burst reducing.*

A mapping ϕ or Φ is *memoryless* if $\phi_m(t) = \phi(m(t))$ for all m, t. The output $\Phi(m)(t)$ may depend on the entire m, but the bandwidth *allocation* depends only on the instantaneous input rate $m(t)$. A generalization of fixed rate servers is the class of memoryless affine servers ϕ^α defined by

$$\phi^\alpha(m) = \alpha(m - \lambda) + \lambda \tag{2}$$

for some $\lambda \geq 0$ and $0 \leq \alpha \leq 1$. Hence ϕ^0 is a fixed rate server.

Proposition 6 *Define ϕ^α by (2). For any m*

1. *$b^\alpha(t) \equiv (1 - \alpha)b^0(t)$, where where $b^\alpha(t)$ and $b^0(t)$ denote the backlog at servers ϕ^α and ϕ^0.*
2. *$\Phi^\alpha(m)(t) \equiv (1 - \alpha)\Phi^0(m)(t) + \alpha m(t)$.*

The following theorem characterizes the class of memoryless service disciplines that are burst reducing.

Theorem 2 *Let ϕ be memoryless and nondecreasing. Then ϕ is burst reducing if and only if it is affine.*

Several proposed service disciplines [1, 9, 4, 6] allocate a minimum bandwidth to a connection with additional bandwidth given when there is spare capacity. This latter feature utilizes bandwidth more fully but, as Theorem 2 suggests, it may increase burstiness.

4 A virtual circuit

The behavior of a virtual circuit spanning nodes $\Phi_1, ..., \Phi_k$ is given by the map $\Phi = \Phi_k \circ ... \circ \Phi_1$ where

$$m_1 \equiv m; \ m_{i+1} = \Phi_i(m_i), 1 \leq i < k; \ m' = m_k.$$

The following theorem illustrates the connection between service disciplines and buffers along a virtual circuit.

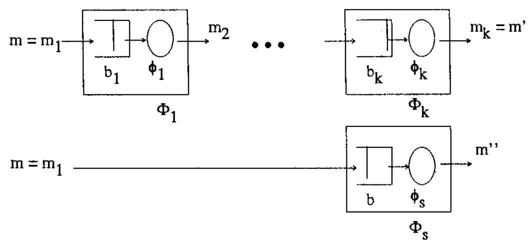

Figure 1: Tradeoff between bandwidth and buffers along a virtual circuit

Theorem 3 *Consider the virtual circuits in figure 1. Suppose* $\phi_i \equiv \mu_i$, $i = 1, \cdots, k$, *and* $\phi_s \equiv \mu_s$ *where* $\mu_s := \min\{\mu_i \mid i = 1, \cdots, k\}$. *Then, for all* m, $\sum b_i(t) \equiv b(t)$ *and* $m'(t) \equiv m''(t)$.

In standard queueing analysis the nodes along a virtual circuit are often modeled as independent $M/M/1$ systems, so the buffer requirements at different nodes are independent. In contrast, Theorem 3 asserts that the sum of the backlogs is a constant. The theorem suggests that we may define the delay of a virtual circuit spanning several nodes to be $\frac{b(\mu_s)}{\mu_s}$, provided that each node is a fixed rate server and μ_s is the smallest service rate.

5 Resource sharing at a node

Recall that a link processor serving n virtual circuits is defined as a mapping Φ,

$$\Phi(m_1, ..., m_n) = (m_1', ..., m_n').$$

Let $\phi(t) = (\phi_1(t), \cdots, \phi_n(t))$ be the bandwidth allocated to $m_1, \cdots, m_n$ at time t. Suppose $\sum \phi_i(t) \le \mu$, where μ is the maximum bandwidth on the output link. Let

$$b_i(\phi) \quad = \quad \sup_{s \le t} \int_s^t [m_i(r) - \phi_i(r)] dr$$

be the maximum buffer occupancy by connection i under ϕ, and $b(\phi) := [b_1(\phi), \cdots, b_n(\phi)]$.

Let $B_G := \{b \in \Re^n | \exists \phi, \sum \phi_i(t) \le \mu, b(\phi) \le b\}$ be the set of achievable maximum backlogs under general service disciplines. Let $B_F := \{b \in \Re^n | \exists \phi(t) :\equiv (\mu_1, \cdots, \mu_n), \sum \mu_i \le \mu, b_i(\mu_i) \le b_i\}$ be the set of achievable maximum backlogs under fixed rate service disciplines. Then, we have

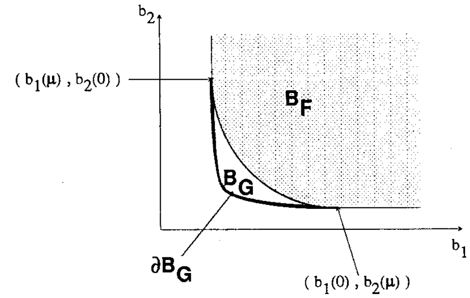

Figure 2: Resource sharing at a node

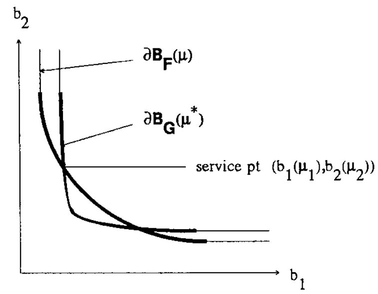

Figure 3: Saving bandwidth

Proposition 7 *Both* B_F *and* B_G *are convex. Furthermore,* $B_F \subseteq B_G$.

For $n = 2$, the situation is illustrated in figure 2. Note that to achieve the boundary $\partial B_G(\mu)$ of $B_G(\mu)$ may require advance knowledge of all m_i.

Proposition 7 illustrates the potential for gains in sharing resources among the virtual circuits. Suppose each virtual circuit i subscribes to $(\mu_i, b_i(\mu_i))$. Then each output link requires $\sum \mu_i = \mu$ bandwidth and $\sum b_i(\mu_i)$ buffers in the fully segregated case. Since $B_G(\mu') \subseteq B_G(\mu)$ if $\mu' < \mu$, we may use the same number of buffers but save bandwidth by using, instead of μ,

$$\mu^* \quad := \quad \inf\{\mu' | \exists \phi, \sum \phi_i(t) \le \mu', b_i(\phi) \le b_i(\mu_i)\}.$$

The situation for $n = 2$ is illustrated in figure 3.

At the other extreme, we may use the same amount

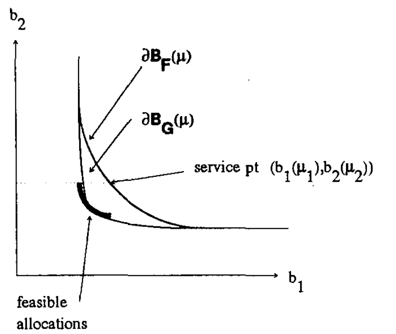

Figure 4: Saving buffers

of bandwidth but save buffers by using, instead of $\sum b_i(\mu_i)$,

$$b^* \ := \ \inf\{\sum b_i(\phi) | \exists \phi, \sum \phi_i(t) \le \mu, b_i(\phi) \le b_i(\mu_i)\}.$$

The situation for $n = 2$ is illustrated in figure 4.

6 Conclusion

We have presented a framework to study the tradeoffs between bandwidth and buffers in an ATM network. We characterize traffic burstiness by its buffering requirement. Two popular service disciplines - the fixed rate discipline and the leaky bucket discipline - are burst reducing, and we have obtained a characterization of a class of service disciplines that are burst reducing. We apply our results on a single server to investigate the behavior of a virtual circuit spanning several nodes. Finally we illustrate the potential for gains from sharing bandwidth at a node among several connections.

References

[1] Jr. A. E. Eckberg, D. T. Luan, and D. M. Lucantoni. Meeting the challenge: congestion and flow control strategies for broadband information transport. *Proc. GLOBECOM'89*, December 1989.

[2] V. Anantharam. The optimal buffer allocation problem. *IEEE Transactions on Information Theory*, 35(4), July 1989.

[3] V. Anantharam and A. Ganesh. Correctness within a constant of an optimal buffer allocation rule of thumb. Submitted to IEEE Transactions on Information Theory, 1991.

[4] Krishna Bala, Israel Cidon, and Khosrow Sohraby. Congestion control for high speed packet switched networks. *INFOCOM'90*, June 1990.

[5] Rene L. Cruz. A calculus for network delay Part I,II. *IEEE Transactions on Information Theory*, January 1991.

[6] A. I. Elwalid and Debasis Mitra. Analysis and design of rate-based congestion control of high speed networks, I: Stochastic fluid models, access regulation. *To appear in Queueing Systems*, 1991.

[7] D. Ferrari and D. Verma. A scheme for real-time channel establishment in wide-area networks. *IEEE Journal on Selected Areas in Communications*, 8, April 1990.

[8] G. J. Foschini and B. Gopinath. Sharing memory optimally. *IEEE Transactions on communications*, March 1983.

[9] G. Gallassi, G. Rigolio, and L Fratta. ATM: bandwidth assignment and bandwidth enforcement policies. *Proc. GLOBECOM'89*, December 1989.

[10] C. Kalmanek, H. Kanakia, and S. Keshav. Rate controlled servers for very high speed networks. *Proceedings of Globecom 1990*, December 1990.

[11] Farouk Kamoun and Leonard Kleinrock. Analysis of shared finite storage in a computer network node environment under general traffic conditions. *IEEE Transactions on communications*, July 1980.

[12] Samar Singh, Ashok K. Agrawala, and Srinivasan Keshav. Deterministic analysis of flow and congestion control policies in virtual circuits. preprint, 1990.

[13] Dinesh C. Verma, Hui Zhang, and Domenica Ferrari. Delay jitter control for real-time communication in a packet switching network. *Proceedings of TriComm'91*, March 1991.

[14] Hui Zhang and Srinivasan Keshav. Comparison of rate-based service disciplines. *to appear in SIGCOMM'91*, Sept. 1991.

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