On the Performance of an ATM Switch with Multichannel Transmission Groups

by

Arthur Y. M. Lin and John A. Silvester

Technical Report No. CENG 89-35

On the Performance of an ATM Switch
with Multichannel Transmission Groups

Arthur Y. M. Lin and John A. Silvester

Department of Electrical Engineering - Systems
University of Southern California
University Park, Los Angeles, CA 90089-0781

ABSTRACT

A routing architecture applying the concept of multichannel transmission group (MCTG) for ATM systems is proposed. We present a queueing analysis of an internally nonblocking ATM switch employing this MCTG concept with partially shared output buffers. The analysis is based on the discrete-time $D^{[A]}/D/c/B$ queueing model. We consider both bulk input traffic with constant bulk inter-arrival time (D) and general bulk-size distribution (A) and deterministic traffic $(D_1 + \cdots + D_N)$. The impact of switch speedup on the performance is also taken into account. It is shown that the MCTG architecture yields better performance in terms of delay and cell loss probability than its single channel using fixed-path routing counterpart. We also find that the switch speedup required to closely approximate the optimal performance obtained by having the switch fabric run N times as fast as the input and output channels where N is the size of the switch is rather small compared with N. This makes the practical realization of the proposed switch architecture feasible.

Contents

1 INTRODUCTION

Historically, communication networks are individually developed for specific applications. Examples include the Circuit-Switched Telecommunication Network (CSTN) for voice service, various packet-switched networks (e.g. ARPANET) for data traffic, and the cable television network (CATV) for TV signal distribution. However, the desire, from the user as well as the network provider, for new services and applications, convenience, economy, flexibility, and efficiency together with the prevalence of fiber optics and new switching technologies will lead us to a new generation of communication system generally referred to as Broadband Integrated Service Digital Network (B-ISDN). A wide range of communications capabilities will be integrated in B-ISDN. The key objective of B-ISDN is to be able to support and accommodate diverse and dynamic user demands such as data, voice, graphics, video, images, and their possible combinations in a flexible and cost-effective manner, both at the level of the individual interface and the system as a whole [25]. B-ISDNs will support both switched and nonswitched connections. B-ISDNs will support services requiring both circuit-mode and packet-mode information transfer capabilities [30]. It will provide an integrated user-network interface allowing simultaneous accesses to a wide variety of network services and applications.

Asynchronous Transfer Mode (ATM) will be the target transfer mode (i.e., multiplexing and switching techniques) solution[1] for implementing B-ISDN [30]. ATM is a specific packet-oriented transfer mode using asynchronous (statistical) time division multiplexing and switching mechanism, providing integrated network access as well as an unique basic network transport service. Information is transported in small fixed-length "packets" called *cells* [2] prefixed with a header containing a label that

---

[1]Synchronous Optical Network (SONET) favored by T1S1 and/or asynchronous time division (ATD) favored by some in ETSI will provide the target transmission infra-structure.

[2]Most recently, the size of ATM cell has been standardized by CCITT to have a 48-octet infor-

1

uniquely identifies the virtual channel[3] (VC) over a User-Network Interface (UNI) or Network-Node Interface (NNI). ATM is a connection-oriented technique allowing the network to support both connection-oriented as well as connectionless services [8], [30]. The main feature of ATM is its flexibility in bandwidth allocation and variable user bit rate support.

The tremendous increase in transmission speed provided by fiber optics has not been matched by a corresponding increase in switching and processing capacities. Photonic switching and computing systems are still in their early stages and far away from real wide-spread deployment. Consequently, the "long"-term presence of the optic-electronic boundary is envisioned. This results in a movement of network bottlenecks: from *links* to *nodes*. In addition, the characteristic of conventional packet-switched networks of normally high error-rate is no longer representative of ATM/B-ISDNs with the introduction of optical fibers[4]. Because of these fundamental changes, neally all per cell processing is pushed outside of the communication subnetwork in ATM systems. Only relatively simple functions, such as self-routing, error detection, buffer management, can be accomplished with every switching node in the network; whereas more complex network control protocols, such as flow control, error recovery, should be performed on an end-to-end or edge-to-edge basis [2], [23].

In this paper we propose a network routing architecture applying the concept of a multichannel transmission group (MCTG) for ATM networks. We then study the performance of an ATM switch employing this MCTG architecture and compare with that using a conventional single channel (SCTG) approach. The effect of switch speedup on the performance is also taken into account.

---

mation field plus a 5-octet header.

[3]also called virtual connection, virtual circuit, or virtual link by others.

[4]The transmission bit-error-rate of optical fibers can be as low as 10^{-12} or even lower in current technology.

In ATM/B-ISDNs, two adjacent communication nodes (UNI or NNI) are commonly connected by several parallel communication channels. Because of the connection-oriented nature of the ATM, two basic approaches exist for routing cells: *fixed-path* and *fixed-node*. Fixed-path routes all cells, during the information transfer phase of a VC, along a fixed sequence of nodes and channels while fixed-node specifies only the nodes. The fixed-node routing can be extended by subdividing the channels between any pair of adjacent nodes into a set of *transmission groups* (MCTGs). Information moving through the network is distributed among parallel channels within a MCTG, permitting more efficient use of network resources (bandwidth, switching and processing capacities, and buffers) and providing better reliability/availability. The MCTG routing architecture provides the additional advantage of super-rate switching which can naturally support services and applications with peak bandwidth requirement exceeding the capacity of any single link.

The MCTG concept has been considered in some packet-switched network architectures and protocols including Transmission Groups in System Network Architecture (SNA) of IBM [11], [14], and the multilink procedure (MLP) in the X.25 and X.75 link layer protocols. While enjoying performance and reliability advantages, the MCTG architecture in conventional packet-switched networks suffers from the sequencing problem caused by having variable length packets, or different transmission rates associated with different channels in a MCTG[5]. To restore the (end-to-end) ordering of packet transfer for any given call, a *resequencing* mechanism which might be implemented on a hop-by-hop and/or an end-to-end basis is required [1], [37]. However, this problem is not present in ATM systems since the cells are fixed length and it is reasonable to assume that all channels in a given MCTG have the same transmission rate in light of the fact that the data rates of both UNI and NNI are subject to standardization (sequencing problems due to errors are inherent and are properly resolved on an end-to-end basis in ATM networks due to the very low error

---

[5]Transmission errors on one of the channels may also cause out-of-sequence packet delivery.

rates found in optical fibers) [22].

Recently, Turner proposed a similar concept (called link group according to his terminology) to allow dynamic load distribution across a set of communication links by cyclically selecting them at each input port of a fast packet switching network [33]. Buffers are dedicated at each input port of the switch fabric. Therefore, a link group is functionally equivalent to a set of single-server queues in parallel. A detailed simulation study is provided in [3]. Hui used the trunk group concept and developed a three-phase algorithm for optimally allocating network resources in ATM networks in [13]. The analysis in his paper concentrated on evaluating the blocking probability within a trunk group at the burst level using the central limit theorem and large deviation theory. In [28], Pattavina introduced a similar concept of a channel group to address the problem of bandwidth allocation in a broadband packet switch and showed a feasible hardware implementation in a Batch-banyan switch with input queueing. The channel group can be functionally modeled as a multi-server queue with shared buffers. Some performance results obtained using simulation and simplified analysis were also provided.

We analyze the performance of applying the MCTG architecture in an internally nonblocking ATM switch[6] with *partially* shared output buffers using the proposed $D^{[4]}/D/c/B$ queueing model in the present paper. An N by N (i.e., N inputs and N outputs) switch is considered. The switch fabric runs S times as fast as the input and output channels, where $1 \leq S \leq N$. A typical system model is shown in figure 1. Notice that buffers are dedicated to each individual output MCTG. However, the output queue for a MCTG is *logically* shared and may not be implemented in a common memory accessible by all output channels of the MCTG. This is justified from a practical point of view in a high-speed network setting. One promising buffer

---

[6]An interconnection network which can perform all possible connections between inputs and outputs without conflicts in the use of links/nodes inside the network is called a *nonblocking network* (e.g., crossbar switch, Knockout switch [34], Batcher-banyan network [12]).

4

organization scheme can be found in [34].

The network traffic can be roughly classified into two main categories: deterministic (or Continuous Bit stream Oriented (CBO)) and bursty. They can be further sub-divided according to various attribute values such as bandwidth requirement, burstiness, etc.. Deterministic traffic (e.g. voice, FAX, file transfer) can be modeled as the superposition of a set of periodic arrival processes, while more general traffic can be modeled as a bulk arrival process with constant bulk inter-arrival time and general bulk-size distribution. It is assumed that the time axis is equally divided into time slots (discrete-time) and all arrivals and departures[7] of the switch are synchronized with slot boundaries. This assumption can be justified by the synchronous operation of standardized underlying SONET frame [31] and the fixed cell size of ATM. The transmission time of a cell is one time slot. Without loss of generality, we focus on a tagged output MCTG queue. Both infinite and finite buffer systems (for the case of periodic arrival, these correspond to no-loss and loss systems respectively) are considered. The resulting queueing system with bulk arrivals, deterministic inter-bulk arrival time (D), general bulk-size distribution (A), multi-server (c), deterministic service (D), and buffers of size B is then generally called the discrete-time $D^{[A]}/D/c/B$ queueing system. Note that the buffer size B is defined in this paper as the maximum number of cells, *excluding* those in transmission, that the system can hold.

Much work has appeared in literature on the performance evaluation of broadband packet switches with the majority devoted to the single-channel (SCTG $- c = 1$) system (or, correspondingly, the fixed-path routing architecture). Karol, Hluchyj, and Morgan [18] compared the performance of input versus output queueing for a time-slotted space-division fast packet switch. In their analytical model for output queues,

---

[7]Assume that an arriving cell at the switch is transmitted immediately without any delay if there is a free channel. The concept of virtual cut-through is applied here [19].

they solved the system with infinite buffer and assuming binomial bulk size. The finite buffer case was studied later by Karol and Hluchyj in [10]. In addition, they considered the problem of using a packet switch for circuit-switched traffic and addressed the tradeoff between the size of buffer required in a switch and the probability that an existing circuit connection is disrupted. Their analysis was based on a discrete-time queueing system with periodic arrival processes. The system was solved by considering all possible arrival patterns (i.e., exhaustive search) [16], [17]. Eckberg, in his early work [6], derived an efficient algorithm for computing the exact delay distribution for the *continuous-time* single server queue with periodic arrivals, deterministic service time and infinite buffer size. More recently, Eckberg and Hou investigated the effects of output buffer sharing on buffer requirements in an ATM packet switch using a discrete-time $D^{[A]}/D/1$ queueing model with binomial distributed bulk size [7]. In [32], Tran-gia and Ahmadi solved a discrete-time $G^{[X]}/D/1/S$, according to their terminology, queueing system, motivated by its applications in packet-switching systems. Note that for the case of a single-channel with finite buffer their analysis is somehow more general by incorporating general bulk inter-arrival times. Morris proposed an algorithmic method to analyze a discrete-time $D/D/1$ queue with infinite as well as finite buffer applied to packet-switching applications [26]. Recently, efficient recursive computational algorithms were developed by Cidon and Sidi in [5] for analyzing the performance of a single-channel ATM system with periodic arrivals. Lin and Silvester proposed to apply the MCTG routing architecture in ATM networks and studied the performance of an MCTG ATM switch for both general and deterministic traffic [24], [22].

All of the work mentioned above assumed the switch operates at the same speed ($S = 1$) or N times as fast ($S = N$) as the input and output channels for simplicity. Most recently, Oie et al. considered the effect of speedup (S) on delay and loss performance for the case of an arbitrary S such that $1 \leq S \leq N$ in an internally nonblocking packet switch with conventional SCTGs [27]. In addition to the case

6

that buffers are only placed at outputs, they also studied the case with both input as well as output buffers through simulations.

While most of the previous work has emphasized on the single-channel system, we analyze a more general case — multi-channel system. Both general and deterministic traffic are considered. For input traffic with general bulk-size distribution, this paper studies infinite as well as finite buffer cases. For periodic input traffic, we look at no-loss as well as loss systems. The analysis given in this paper is for an arbitrary value of S, where $1 \leq S \leq N$.

The remainder of this paper is organized in the following fashion. In section 2, we present a queueing analysis of the performance of an ATM switch employing the proposed MCTG routing architecture for both infinite and finite buffer cases using the general discrete-time $D^{[A]}/D/c/B$ queueing model. Next, in section 3, we extend the recursive algorithms developed by Cidon and Sidi in [5] to a multi-channel system with N superposed periodic arrival processes $(D_1 + \cdots + D_N/D/c/B)$ taking into account the cell loss due to $S < N$ to analyze the performance under the deterministic traffic condition. Numerical results for comparing the MCTG approach with the SCTG and investigating the performance impact of switch speedup on an internally nonblocking $N \times N$ ATM switch are provided in section 4. Finally, we conclude the paper in section 5.

2 GENERAL TRAFFIC

We first introduce the notation used in this paper. For a discrete-time random variable (r.v.) X, we define the following:

- probability mass function (pmf) : $x(k) \triangleq P[X = k]$

7

- mean : $\bar{X} \triangleq E[X]$

- variance : $\sigma_X^2 \triangleq Var(X)$

- coefficient of variation : $c_X \triangleq \sigma_X / \bar{X}$

- probability generating function (p.g.f.) : $X(z) \triangleq \sum_{k=-\infty}^{\infty} x(k)z^k$

The switch we consider consists of N inputs and N outputs, where the outputs is divided into M ($M \le N$) MCTGs as shown in figure 1. Without loss of generality, we focus on a tagged MCTG queue, say m ($1 \le m \le M$), with $c^{(m)}$ channels and buffers of size $B^{(m)}$ excluding the cell transmitting register. The MCTG index is dropped below unless otherwise mentioned.

2.1 Infinite Buffer

We shall assume that the arrival process to the tagged MCTG consists of the superposition of N independent arrival sources. Cells from a particular input, say n ($1 \le n \le N$), arrive at the system according to a random process[8]

$$\mathbf{A}^{(n)} = \{A_i^{(n)} : 1 \le i < \infty\}$$

where $A_i^{(n)}$ is the r.v. of the number of arrivals to the system at the beginning of time slot i from input n. For any n ($1 \le n \le N$), the random variables $\{A_i^{(n)} : 1 \le i < \infty\}$ are assumed to be independent and identically distributed (i.i.d.). For time-invariant inputs, the pmf of $A_i^{(n)}$ is $a^{(n)}(k) \triangleq P[A_i^{(n)} = k]$ and has finite mean, $\lambda^{(n)} = E[A_i^{(n)}]$, and variance, $\sigma_{A_i}^{(n)^2} = Var(A_i^{(n)})$. The overall arrival process to the system can thus be modeled as a bulk (or batch) arrival process with deterministic bulk inter-arrival

---

[8]Here we adopt the convention of using subscripts to indicate evolution in time and superscripts to indicate the particular input port.

time (i.a.t.) equaling to one slot and general bulk-size distribution $\{A_i : 1 \le i < \infty\}$. The total number of cells arriving to the system at the beginning of slot i is given by

$$A_i = A_i^{(1)} + A_i^{(2)} + \cdots + A_i^{(N)}$$

Again the random variables $\{A_i : 1 \le i < \infty\}$ are also i.i.d., with pmf $a(k) \triangleq P[A_i = k]$ resulting from

$$a(k) = a^{(1)}(k) * a^{(2)}(k) * \cdots * a^{(N)}(k)$$

where the symbol $*$ denotes the discrete convolution operator, and have finite mean, $\lambda = E[A_i]$ and variance, $\sigma_{A_i}^2 = Var(A_i)$.

If the arrival processes at each input port of a switch are i.i.d., the input traffic of the switch is *uniform* or *balanced*. The output traffic of a switch is *uniform* if the probability that an input cell destined for a particular MCTG, say m $(1 \le m \le M)$, is proportional to the capacity $(c^{(m)})$ of the MCTG. It is assumed in this paper that both input and output traffic are uniform.

We define the r.v. $\hat{A}_i$ as the number of cells actually entering the tagged output MCTG queue with pmf $\hat{a}(k) \triangleq P[\hat{A}_i = k]$. Let Q_i be the r.v. of the number of cells waiting in the queue (i.e., the backlog) at the end of slot i. For the moment let us assume infinite buffer. The evolution of the system can thus be described by the following equation:

$$Q_i = [Q_{i-1} + \hat{A}_i - c]^+ \tag{1}$$

$$\hat{A}_i = \min\{A_i, S\} \tag{2}$$

where we employ the notation $X^+ = \max(X, 0)$. This queueing system can be modeled as a discrete-time Markov chain (DTMC). For a stable system (i.e., $\lambda \le c$), the state probability $q_i(k)$ of Q_i converges to an equilibrium (steady state) pmf $q(k)$.

We now introduce the generating functions:

$$A(z) = \sum_{k=-\infty}^{\infty} a(k) z^k$$

9

$$\hat{A}(z) \triangleq \sum_{k=-\infty}^{\infty} \hat{a}(k)z^k$$

$$= \sum_{k=0}^{S-1} a(k)z^k + \sum_{k=S}^{\infty} a(k)z^S$$

$$= \sum_{k=0}^{S-1}(z^k - z^S)a(k) + z^S \qquad (3)$$

$$Q(z) = \sum_{k=-\infty}^{\infty} q(k)z^k$$

where $A(z)$, $\hat{A}(z)$, and $Q(z)$ are analytic in the unit disk $|z| < 1$.

We first multiply equation (2) by z^k and sum over the index k assuming the system at steady state (i.e. $i \to \infty$). After some manipulation, this leads to

$$Q(z) = \frac{\displaystyle\sum_{m=0}^{c-1} r_m(z^c - z^m)}{z^c - \hat{A}(z)} \qquad (4)$$

Now we are in a position to get rid of the c unknown terms, $\{r_m\}_{m=0}^{c-1}$, in the numerator of equation (4). This can be done by invoking the requirement that $Q(z)$ is analytic in $|z| < 1$ and the fact that $Q(1) = 1$. Using Rouche's theorem, it can be shown that the equation

$$z^c - \hat{A}(z) = 0 \qquad (5)$$

has exactly $(c - 1)$ complex zeros inside $|z| \leq 1$ while the cth is clearly $z = 1$; we denote these zeros by $\theta_1, \ldots, \theta_{c-1}$ and $\theta_c = 1$ [15], [21][9]. The analyticity of $Q(z)$ requires that all the zeros of the denominator within and on the unit circle must also make the numerator vanish. That is, $\theta_1, \ldots, \theta_{c-1}$ and $\theta_c = 1$ are the coincident roots of the denominator and numerator in (4). Therefore, applying the fact that $Q(1) = 1$ and carrying out the necessary manipulation, we rewrite equation (4) as

$$Q(z) = \frac{(c - \hat{\lambda})(z - 1)}{z^c - \hat{A}(z)} \prod_{m=1}^{c-1} \frac{z - \theta_m}{1 - \theta_m} \qquad (6)$$

[9]The analysis for $M/D/c$ is shown in [9].

10

where the successful arrival rate, $\hat{\lambda}$, is given by

$$\hat{\lambda} = \left.\frac{d\hat{A}(z)}{dz}\right|_{z=1} = \lambda(1-\psi)$$

and ψ, the loss probability due to small switch speedup $(S < N)$ is given by

$$\psi = \frac{1}{\lambda}\sum_{k=S+1}^{\infty}(k-S)\cdot a(k) \tag{7}$$

The queue size distribution, $\{q(k)\}_{k=0}^{\infty}$, may now be obtained by

$$q(k) = P[Q = k] = \left.\frac{d^k Q(z)}{dz^k}\right|_{z=0}$$

although this frankly gets to be rather difficult as c gets large. With some tedious computations, we can, however, find the mean and variance of the queue size:

$$\bar{Q} = \frac{\hat{A}''(1) - c(c-1)}{2[c-\lambda(1-\psi)]} + \sum_{m=1}^{c-1}\frac{1}{1-\theta_m} \tag{8}$$

$$\sigma_Q^2 = \left\{\frac{2}{3} + \frac{2[\hat{A}''(1) - c(c-1)]}{c-\lambda(1-\psi)}\right\}\sum_{m=1}^{c-1}\frac{1}{1-\theta_m}$$

$$+\frac{\hat{A}'''(1) - c(c-1)(c-2)}{3[c-\lambda(1-\psi)]} + \frac{[\hat{A}''(1) - c(c-1)]^2}{2[c-\lambda(1-\psi)]^2} + \bar{Q} - (\bar{Q})^2 \tag{9}$$

where

$$\hat{A}''(1) = \left.\frac{d^2\hat{A}(z)}{dz^2}\right|_{z=1} = E[\hat{A}_i^2]$$

$$\hat{A}'''(1) = \left.\frac{d^3\hat{A}(z)}{dz^3}\right|_{z=1} = E[\hat{A}_i^3]$$

The average waiting time of cells in the queue is obtained by applying the well known Little's result[10]

$$\bar{W} = \frac{\bar{Q}}{\lambda(1-\psi)} \tag{10}$$

In addition, we are now interested in deriving the waiting time pmf. Consider a tagged cell that arrives at the system with $n-1$ other cells (i.e., bulk size is n) in slot

[10]All moments of the queue size and arrival bulk-size may be obtained from (6) and $A(z)$, which, in turn, enable us to find all moments of the waiting time via the extension of Little's result [9].

11

i, provided that the backlog from the previous slot is l, then the probability that the tagged cell falls in a batch of size n is given by $n \cdot \hat{a}(n) / \lambda(1 - \psi)$ [4]. It is assumed that all cells in the ith bulk are transmitted before the first cell in the $(i+1)$th bulk and the order of transmission within a bulk is random. The waiting time of the tagged cell thus consists of the delay of transmitting those cells ahead of it arriving in the same batch and those backlogged from previous slot. Consequently, the waiting time of an arbitrary cell can be expressed as follows:

$$w(k) = \sum_{n=1}^{S} \frac{n \cdot \hat{a}(n)}{\lambda(1 - \psi)} \cdot \frac{1}{n} \sum_{m=0}^{n-1} \sum_{l=kc-m}^{(k+1)c-1-m} q(l) \tag{11}$$

$$= \frac{1}{\lambda(1 - \psi)} \sum_{n=1}^{S} \hat{a}(n) \sum_{m=0}^{n-1} \sum_{l=kc-m}^{(k+1)c-1-m} q(l) \tag{12}$$

where $q(l) = 0$ if $l < 0$.

Example: single-channel with binomial bulk size and $S = N$

Now, considering a single-channel (SCTG $- c = 1$) system, $Q(z)$ becomes

$$Q(z) = \frac{(1 - \lambda)(z - 1)}{z - A(z)}$$

If it is assumed that each input of the switch forms a Bernoulli process with arrival probability p per slot. Thus $\hat{A}_i = A_i$ (for any output) has the binomial pmf

$$a(k) = P[A_i = k] = \binom{N}{k} \left(\frac{p}{N} \right)^k \left(1 - \frac{p}{N} + \frac{p}{N} z \right)^{N-k} \qquad 0 \le k \le N$$

and its corresponding p.g.f.

$$A(z) = \left(1 - \frac{p}{N} + \frac{p}{N} z \right)^{N}$$

After some algebra, we find the mean and variance of queue size for $c = 1$ case from (8) and (9) as follows:

$$\bar{Q} = \left(\frac{N-1}{N} \right) \frac{p^2}{2(1-p)} \tag{13}$$

$$\sigma_Q^2 = \left[\frac{(N-1)(N-2)}{N^2} \right] \frac{p^3}{3(1-p)} + \left(\frac{N-1}{N} \right) \frac{p^2}{2(1-p)}$$
$$+ \left(\frac{N-1}{N} \right)^2 \frac{p^4}{4(1-p)^2} \tag{14}$$

12

These are identical to the results in [7] and [18].

2.2 Finite Buffer

In this subsection, we consider the same queueing system except the buffer space (per output MCTG) is now assumed to be finite with size B $(B < N - c)$. The system evolves according to the following equation:

$$Q_i = \min\{[Q_{i-1} + \hat{A}_i - c]^+, B\} \tag{15}$$

The system state Q_i constitutes a stationary finite-state DTMC with state transition probabilities $p_{kl} \triangleq P[Q_{i+1} = l | Q_i = k]$ given by

$$p_{kl} = \begin{cases} \sum_{n=0}^{c-k} \hat{a}(n) & l = 0, \ 0 \le k \le \min\{c, B\} \\ \hat{a}(c + l - k) & 1 \le l \le B - 1, \ 0 \le k \le \min\{c + l, B\} \\ \sum_{n=B+c-k}^{S} \hat{a}(n) & l = B, \ B + c - S \le k \le B \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

Both transient and steady (equilibrium) state probabilities can be obtained [20]. In this paper, we focus on the steady state probabilities which we find by numerical solution techniques such as LU decomposition together with forward substitution and backsubstitution, or Gaussian elimination with backsubstitution [29]. The average queue length, $\bar{Q}$, can then be determined by

$$\bar{Q} = \sum_{k=1}^{B} k \cdot q(k)$$

where $q(k)$ is the steady state probability found above.

To obtain the probability of cell loss due to buffer overflow, ϕ, we observe a tagged cell contained in an arriving batch of size n. The probability of the tagged cell being

13

in a bulk of size n is $n \cdot \hat{a}(n)/\lambda(1 - \psi)$. The conditional probability of cell loss given that the queue size is k upon arrival, ϕ_k, is thus

$$
\begin{aligned}
\phi_k &= \sum_{n=B+c+1-k}^{S} P[\text{the tagged cell is lost} \mid \text{arriving batch-size} = n] \\
&\qquad \cdot P[\text{the tagged cell in an arriving bulk of size } n] \\
&= \sum_{n=B+c+1-k}^{S} \frac{(n+k-B-c)}{n} \cdot \frac{n \cdot \hat{a}(n)}{\lambda(1-\psi)}
\end{aligned} \tag{17}
$$

Then, by the theorem of total probability, we have

$$
\phi = \frac{1}{\lambda(1-\psi)} \sum_{k=0}^{B} q(k) \sum_{n=B+c+1-k}^{S} (n+k-B-c) \cdot \hat{a}(n) \tag{18}
$$

The overall cell loss probability, φ, is then given by

$$
\varphi = \psi + \phi - \psi\phi \tag{19}
$$

So the (effective) waiting time, $\bar{W}^*$, averaged over cells actually admitted into the queue is given by

$$
\bar{W}^* = \frac{\bar{Q}}{\lambda^*}
$$

where

$$
\lambda^* = \lambda(1-\psi)(1-\phi)
$$

λ^* is the effective arrival rate of the system defined as the rate of cells successfully entering the system. The results in [10] and [22] for the SCTG and MCTG systems, respectively, under the assumption of $S = N$ are special cases of the results presented here.

3 PERIODIC TRAFFIC

In this section, we consider the same queueing system but with periodic arrival process which can then be modeled as a discrete-time $D_1 + \cdots + D_N/D/c/B$ queueing system.

14

An efficient computational algorithm for analyzing the performance of the queueing system is presented by extending the methods developed in [5] to multiple channels and taking the cell loss caused by $S < N$ into account.

3.1   No Loss System

Following the approach in [5], a *frame* is defined to be any F consecutive time slots. The arrival process consists of N $(D_1, \cdots, D_N)$ independent periodic arrival sources and each source generates exactly one arrival per frame. For a stable system, the obvious condition $N \leq c \cdot F$ must hold. In this subsection, the system is assumed to be capable of storing $N - c$ cells (excluding those cells in transmission) to guarantee no loss.   Let A_i be the number of arrivals at the beginning of slot i and Q_i be the number of cells waiting in the queue at the end of slot i. For a given set of arrival sources, the N cells arrive in the same pattern and can be described by [5], [16],

$$A_i = A_{jF+i} \qquad 1 \leq i \leq F, \quad j = 1, 2, \cdots$$

Consequently,

$$\hat{A}_i = \hat{A}_{jF+i} \qquad 1 \leq i \leq F, \quad j = 1, 2, \cdots$$

As in the last section, the system equation is:

$$\begin{aligned} Q_0 &= 0 \\ Q_i &= [Q_{i-1} + \hat{A}_i - c]^+ \end{aligned} \qquad (20)$$

The resulting queue size can be shown to be periodic by extending lemma 2 and corollary 1 of [5]. Therefore, we only need to calculate q_i's for $F \leq i \leq 2F - 1$ using the system evolution equation given in (20). In addition, it can be easily proved that the system reaches the steady-state periodic behavior for any given set of arrival sources after at most F time slots (one frame). Applying the approach presented in

15

[5], we construct an alternate system by enforcing $Q_{j-F} = 0$ for an arbitrary slot j ($F \leq j \leq 2F - 1$). Based on theorem 1 of [5], we can solve the system using the alternate (disconnected) system instead of the original cyclic system.

In this work, we generalize the algorithms given in [5] to the multi-channel case. It is assumed that the N arriving cells are uniformly distributed over a frame of length F. Let $a_{F,N}(k)$ be the probability that k out of the N cells arrive at the beginning of any slot in a frame of size F and $q_{F,N}(k)$ be the probability that a typical slot in the system has a queue size of k. We thus have

$$a_{F,N}(k) = \binom{N}{k} \left(\frac{1}{F}\right)^k \left(1 - \frac{1}{F}\right)^{N-k} \qquad 0 \leq k \leq N \tag{21}$$

The corresponding $\hat{a}_{F,N}$ and $\psi_{F,N}$ can be defined and determined in a similar way as in previous section.

The set of recursive computational equations for calculating the queue size pmf of a no-loss $D_1 + \cdots + D_N / D / c / B$ queueing system is then given by

$$q_{F,0}(k) = 0 \qquad F \geq 1,\ k \geq 1 \tag{22}$$

$$q_{F,N}(k) = 0 \qquad F \geq 2,\ k \geq 1,\ 1 \leq N \leq k \tag{23}$$

$$q_{1,N}(k) = \begin{cases} 1 & k = N - c,\ N > c \\ 1 & k = 0,\ 0 \leq N \leq c \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

$$q_{F,N}(k) = \begin{cases} 1 & k = 0,\ N \leq c \\ 0 & k \neq 0,\ N \leq c \end{cases} \tag{25}$$

$$q_{F,N}(0) = \sum_{n=0}^{c} \hat{a}_{F,N}(n) \left[\sum_{k=0}^{c-n} q_{F-1,N-n}(k) \right] \qquad F \geq 2,\ N \geq c \tag{26}$$

$$q_{F,N}(k) = \sum_{n=0}^{k+c} \hat{a}_{F,N}(n) q_{F-1,N-n}(k - n + c) \qquad F \geq 2,\ 1 \leq k \leq N - c \tag{27}$$

Note that $q_{F,N}(k) = 0$ whenever $N > c \cdot F$. Now the average queue size, $\bar{Q}_{F,N}$, is

simply

$$\bar{Q}_{F,N} = \sum_{k=1}^{N-c} k q_{F,N}(k)$$

$$= \sum_{n=0}^{N} \hat{a}_{F,N}(n) \bar{Q}_{F-1,N-n} + \frac{F}{N(1-\psi_{F,N})} - c$$

$$+ c \left[\sum_{n=0}^{c-1} \hat{a}_{F,N}(n) \sum_{k=0}^{c-n-1} q_{F-1,N-n}(k) \right] \qquad (28)$$

We can apply Little's result to obtain the average waiting time in the queue, $\bar{W}_{F,N}$, as

$$\bar{W}_{F,N} = \frac{F}{N(1-\psi_{F,N})} \bar{Q}_{F,N}$$

To find the waiting time pmf, we consider a tagged cell arriving at the system that has frame size F and $N-1$ arrivals together with N_1 $(N_1 \le S)$ other cells (i.e., bulk size$=N_1+1$) out of $(N-1)$ arriving cells in slot F. The position of the tagged cells in this arriving bulk is assumed to be uniformly distributed over the bulk of size N_1+1. The tagged cell has to wait for the transmission of those cells ahead of it within the same bulk, N_2, plus all cells backlogged from the previous slot, Q_{F-1}. If $Q_{F-1}+N_1 \le c-1$, there is no waiting for the tagged cell. Otherwise, the tagged cell will wait for k slots provided that $kc \le Q_{F-1}+N_2 \le (k+1)c-1$ where $1 \le k \le \lfloor \frac{N-1}{c} \rfloor$. The pmf of the waiting time, $w_{F,N}(k)$, of a typical cell can then be found as follows:

$$w_{F,N}(k) = \sum_{N_1=0}^{N-1} \frac{a_{F,N}(N_1)}{N_1+1} \sum_{N_2=0}^{N_1} \sum_{l=kc-N_2}^{(k+1)c-1-N_2} q_{F-1,N-N_1-1}(l) \qquad (29)$$

3.2 Loss System

We now proceed to consider the case of a loss system in which the maximum number of cells (excluding those cells in transmission), B (buffer size), that may be stored is less than $N-c$. When an arriving cell from a particular source is lost, all cells from this source will be lost [35], [36]. In this paper, we are interested in determining

17

the loss probability due to buffer overflow, $\phi_{F,N}$, of a typical cell for a specific set of arriving sources.

The analysis of the loss system follows the same procedure as the no-loss system presented in last subsection except equation (24) becomes

$$
q_{1,N}(k) = \begin{cases} 1 & k = 0, \ 0 \le N \le c \\ 1 & k = N - c, \ c \le N \le c + B \\ 1 & k = B, \ N \ge c + B \\ 0 & \text{otherwise} \end{cases} \tag{30}
$$

and equation (27) is still valid for $k < B$. In addition, we have

$$
\begin{aligned}
q_{F,N}(B) &= 0 \qquad F \ge 2, \ N \le c + B \\
q_{F,N}(B) &= \sum_{n=c}^{B+c-1} \hat{a}_{F,N}(n) q_{F-1,N-n}(B + c - n) + \sum_{n=B+c}^{N} \hat{a}_{F,N}(n) \qquad F \ge 2
\end{aligned} \tag{31}
$$

Adopting the same argument as for computing the waiting time pmf in the previous subsection, we are now in a position to evaluate the loss probability due to buffer overflow, $\phi_{F,N}$, as follows:

$$
\phi_{F,N} = \sum_{l=0}^{B} \sum_{n=l+c}^{N-1} \frac{(n + 1 - l - c)\hat{a}_{F,N-1}(n)}{n + 1} q_{F-1,N-n-1}(B - l) \tag{32}
$$

The effective arrival rate, λ^*, is thus

$$
\begin{aligned}
\lambda^* &= \lambda(1 - \psi_{F,N})(1 - \phi_{F,N}) \\
&= \frac{N}{F}(1 - \psi_{F,N})(1 - \phi_{F,N})
\end{aligned}
$$

So the average waiting time in the queue, $\bar{W}_{F,N}^*$, averaged over those cells admitted into the system is given by:

$$
\bar{W}_{F,N}^* = \frac{\bar{Q}_{F,N}}{\lambda^*}
$$

where

$$
\bar{Q}_{F,N} = \sum_{k=1}^{B} k \cdot q_{F,N}(k)
$$

18

4 NUMERICAL RESULTS

Some numerical results for various performance measures of interest are shown in this section. We first compare the performance of an MCTG ATM switch with a conventional SCTG counterpart under the assumption of $S = N$. We then exhibit the performance impact of switch speedup in terms of average delay and cell loss probability for both SCTG and MCTG.

4.1 Performance Comparison: MCTG v.s. SCTG

The system model of an N by N internally non-blocking ATM switch with partially shared output buffers employing the MCTG routing architecture is shown in figure 1. For simplicity, it is assumed that all channel groups have the same capacity (i.e., number of output channels) $c^{(m)} = c$ $(m = 1, \ldots, M)$ and both input and output traffic are uniform. As mentioned before, the uniform input traffic is defined to be the traffic pattern in which every input port has i.i.d. arrival process. The uniform output traffic is defined to be the traffic pattern in which the probability that an input cell destined for a particular MCTG is proportional to the capacity ($c^{(m)}$) of the MCTG. Due to the equal capacity assumption in this presentation, each input cell is destined for any output MCTG with equal probability $1/M$ where M is the number of MCTGs in the switch. Note that for the conventional SCTG architecture we have $c = 1$ and $M = N$. We then focus on a tagged output MCTG. For the random arrival process, a binomial distributed bulk-size is assumed[11]. Each input generates a cell in any time slot with probability p (i.e., offered load). For the moment, we assume that the switch under consideration is operated N times as fast as the input and output channels.

---

[11]More general bulk-size distribution and the effects of traffic burstiness can also be taken in account without much difficulty.

We start by considering the infinite buffer case. Figure 2 shows the average delay time as a function of the offered load (p) for various sizes of MCTGs (c) and a switch of size $N = 32$. The scaling advantage favoring a larger c is clearly shown. The delay time is dramatically reduced as c increases.

Then considering the finite buffer case, we compare the MCTG of capacity $c = 8$ with the single channel one. Results are presented in figure 3 for the loss probability for several comparable buffer sizes under a 40×40 switch. We can again see the great performance improvement that can be achieved by using MCTGs. In figure 4, we compute the loss probability as a function of the buffer and switch sizes for the SCTG and the 4-channel MCTG under an 80% offered load condition. The results consistently show that the MCTG architecture can significantly outperform its SCTG counterpart and the performance improvement increases as the capacity of MCTGs increases.

For periodic input traffic, figures 5 and 6 plot the average waiting time and the cell loss probability, respectively, as a function of the normalized buffer size B (defined as the average size of buffers per channel) for a fixed frame size, $F = 10$, and various c under the same loading condition (80%). Note that in figure 5 the waiting time becomes fixed and there is no loss for $B > N - c$. We find similar performance advantages by applying MCTGs as in the random input traffic case. For a given loss probability, the buffer size required in the MCTG architecture is much less than that in classical SCTG one.

Generally speaking, as the capacity of MCTGs increases, the performance improvement increases because of the statistical advantage gained by sharing switch resources. A more detailed treatment of the comparison between the MCTG routing architecture and the SCTG can be found in [24] and [22].

4.2 Effect of Speedup

So far we have assumed $S = N$. In this subsection, we consider the performance impact of switch speedup by having arbitrary values of S, where $c < S \leq N$. Note that the speedup must be at least as great as the capacity of MCTGs. Otherwise there is no delay for cells actually admitted to an output MCTG queue to be transmitted because the number of cells that can enter an output queue in any slot is always less than the capacity of an MCTG and most of cells are lost at inputs. Figure 7 shows the loss probability caused by $S < N$ as a function of switch speedup S for the 2-channel MCTG case under various loading conditions and the same system environment as in figure 2. The speedup loss probability, ψ, becomes negligible when $S \geq 6$. Therefore, in figure 8, we plot the waiting time against normalized effective offered load (defined as the average offered load successfully entering the output MCTG queue per channel) and see that the waiting time is very close to that of the $S = N$ case when $S \geq 6$ while being fairly different from the ideal case ($S = N$) when $S < 6$ because of a large number of cell losses occurring at inputs of the switch. Note that the curve for $S = N$ (ideal case) is plotted for the purpose of comparison.

For the finite buffer case, we first show the loss probabilities due to $S < N$, ψ, and buffer overflow, ϕ, versus offered load in figures 9 and 11 and the total loss probability, φ, in figures 10 and 12 for different values of S for the SCTG and the 8-channel MCTG systems, respectively, for a system configuration identical to that of figure 3 with $B = 3$ and $B = 24$. We then show the loss probabilities, ψ and ϕ, and the total loss probability, φ, as a function of buffer size B for various S in figures 13, 15 and figures 14, 16, respectively, using the same system parameters as in figure 4 and $N = 40$. In all cases we see that the performance measures converge to those of the ideal case with a speedup value only slightly larger the capacity of MCTGs. This is because the loss due to switch speedup (ψ) becomes negligible with a rather small value of S.

The loss probabilities, ψ, ϕ, and the total loss probability, φ, are plotted against switch speedup S for a SCTG and a 4-channel MCTG switch of size $N = 40$ with buffers of size $B = 5$ per channel group under a 90% load condition in figures 17 and 18. All the results presented in this subsection consistently show the fact that the speedup required to closely approximate the ideal case ($S = N$) is relatively small making it easier to physically implement the proposed MCTG scheme in an ATM switch with partially shared output queues.

5   CONCLUSIONS

In this paper, we have proposed a routing architecture applying the concept of MCTG for ATM systems and analyzed the performance of an internally nonblocking ATM switch employing MCTGs based on the discrete-time $D^{[A]}/D/c/B$ queueing model. For input traffic with general bulk-size distribution, we consider both infinite and finite buffer cases. For a periodic arrival process we extended the results of [5] to a multi-channel system for both loss and no-loss systems. The performance impact of switch speedup has also been taken into account.

As indicated and quantified in [10] and [18], both output queueing and completely shared buffering schemes achieve the optimal delay-throughput performance in an internally nonblocking ATM switch with conventional SCTGs using fixed-path routing because cells are delayed or lost only when two or more cells from different inputs simultaneously arrive at the same output under the assumption that the switch fabric runs N times as fast as the input and output channels (i.e., $S = N$). This delay and/or loss probability can, however, be further reduced by using the proposed MCTG architecture in an ATM switch with partially shared output buffers. Performance improvements gained by the MCTG approach are clearly shown by means of

various numerical examples for both random and deterministic traffic conditions. The MCTG routing architecture has the additional advantages of supporting super-rate switching which can naturally support services with peak bandwidth requirement exceeding the capacity of any single channel, providing better reliability/availability, reducing the potential buffering requirement, and allowing efficient resource allocation for ATM/B-ISDNs. In addition, the sequencing problem associated with such approach in conventional packet-switched networks can be easily solved in the ATM environment.

With regards to the effect of switch speedup, in all cases we find that the speedup required to closely approximate the ideal performance obtained by having $S = N$ is rather small compared with N. This makes the physical realization of the proposed switch architecture easier.

References

[1] F. Baccelli and A. M. Makowski. "Queueing models for systems with synchronization constraints". *Proc. of the IEEE*, vol. 77, no. 1, pp. 138-161, Jan. 1989.

[2] A. Bhargava, J. Kurose, D. Towsley, and G. V. Leemput. "Performance comparison of error control schemes in high speed computer communication networks". *IEEE J. Select. Areas Commun.*, vol. SAC-6, no. 9, pp. 1565-1575, Dec. 1988.

[3] R. Bubenik and J. S. Turner. "Performance of a broadcast packet switch". Technical report, WUCS-86-10, Comput. Sci. Dept., Washington Univ., June 1986.

[4] P. J. Burke. "Delays in single-server queues with batch input". *Oper. Res.*, vol. 23, pp. 830-833, July-Aug. 1985.

[5] I. Cidon and M. Sidi. "Performance analysis of asynchronous transfer mode (ATM) systems". Technical report, RC-14395 (#64197), IBM Research Division, T. J. Watson Research Center, Yorktown Heights, N.Y., Jan. 17 1989.

[6] A. E. Eckberg. "The single server queue with periodic arrival process and deterministic service times". *IEEE Trans. Commun.*, vol. COM-27, no. 3, pp. 556-562, March 1979.

[7] A. E. Eckberg and T.-C. Hou. "Effects of output buffer sharing on buffer requirements in an ATDM packet switch". In *Proc. IEEE INFOCOM '88*, pp. 459-466, 1988.

[8] P. Gonet, P. Adam, and J. P. Coudreuse. "Asynchronous time-division switching: the way to flexible broadband communication networks". In *Proc. Int. Zurich Semi. Digit. Commun. '86*, pp. 141-148, 1986.

[9] D. Gross and C. M. Harris. *"Fundamentals of Queueing Theory"*. John Wiley & Sons, New York, NY, 1974.

[10] M. G. Hluchyj and M. J. Karol. "Queueing in high-performance packet switching". *IEEE J. Select. Areas Commun.*, vol. SAC-6, no. 9, pp. 1587-1597, Dec. 1988.

[11] W.N. Hsieh and I. Gitman. "Routing strategies in computer networks". *IEEE Comput. Mag.*, vol. 17, no. 6, pp. 46-56, June 1984.

[12] A. Huang and S. Knauer. "Starlite: a wideband digital switch". In *Proc. IEEE GLOBECOM '84*, pp. 121-125, 1984.

[13] J. Hui. "Resource allocation for broadband networks". *IEEE J. Select. Areas Commun.*, vol. SAC-6, no. 9, pp. 1598-1608, Dec. 1988.

[14] J. M. Jaffe, F. H. Moss, and R. A. Weingarten. "SNA routing: past, present, and possible future". *IBM Syst. J.*, vol. 22, no. 4, pp. 417-434, 1983.

[15] M. Kac, P. Boudreau, and J. Griffen. "An elementary queueing problem". *Amer. Math. Monthly*, vol. 69, no. 8, pp. 713-724, 1962.

[16] M. J. Karol and M. G. Hluchyj. "Using a packet switch for circuit-switched traffic: a queueing systems with periodic input traffic". *IEEE Trans. Commun.*, vol. COM-37, no. 6, pp. 623-625, June 1989.

[17] M. J. Karol and M. G. Hluchyj. "Using a packet switch for circuit-switched traffic: a queueing systems with periodic input traffic". In *Proc. ICC '87*, pp. 1677-1682, 1987.

[18] M. J. Karol, M. G. Hluchyj, and S. P. Morgan. "Input versus output queueing on a space-division packet switch". *IEEE Trans. Commun.*, vol. COM-35, no. 12, pp. 1347-1356, Dec. 1987.

[19] P. Kermani and L. Kleinrock. "Virtual cut-through: a new computer communication switching technique". *Computer Networks*, vol. 3, pp. 267-286, 1979.

[20] L. Kleinrock. *"Queueing Systems, Volumn I: Theory"*. John Wiley & Sons, New York, NY, 1975.

[21] A. G. Konheim and B. Meister. "Service in a loop system". *JACM*, vol. 19, no. 1, pp. 92-108, Jan. 1972.

[22] A. Y. M. Lin and J. A. Silvester. "Fixed-node routing architecture and its performance in ATM networks". *to appear in Proc. IEEE INFOCOM '90*.

[23] A. Y. M. Lin and J. A. Silvester. "A critical analysis of some design issues in very high speed integrated networks (VHSIN): recommendations for the next generation of communication systems". Technical report, Networks and Distributed Systems Laboratory, Department of EE-Systems, Univ. of Southern California, April 1988.

[24] A. Y. M. Lin and J. A. Silvester. "Queueing analysis of an ATM switch with multichannel transmission groups". Technical report, CRI-89-25, Networks and Distributed Systems Laboratory, Dept. of EE-Systems, Univ. of Southern California, Sept. 1989.

[25] S. E. Minzer. "Broadband ISDN and asynchronous transfer mode (ATM)". *IEEE Commun. Mag.*, vol. 27, no. 9, pp. 17-24, Sept. 1989.

[26] R. J. T. Morris. "An algorithmic technique for a class of queueing models with packet switching applications". In *Proc. ICC '81*, pp. 41.2.1-41.2.6, 1981.

[27] Y. Oie, M. Murata, K. Kubota, and H. Miyahara. "Effect of speedup in non-blocking packet switch". In *Proc. ICC '89*, pp. 410-414, 1989.

[28] A. Pattavina. "Multichannel bandwidth allocation in a broadband packet switch". *IEEE J. Select. Areas Commun.*, vol. SAC-6, no. 9, pp. 1489-1499, Dec. 1988.

[29] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *"Numerical recipes in C: the art of scientific computing"*. Cambridge University Press, Cambridge, 1988.

[30] T1S1 Technical Sub-Committee. "Broadband aspects of ISDN". Working Draft, May 1989.

[31] ANSI T1.105-1988. "American National Standard for Telecommunications: Digital Hierarchy - Optical Interface Rates and Formats Specification". March 1988.

[32] P. Tran-gia and H. Ahmadi. "Analysis of a discrete-time $G^{[X]}/D/1-S$ queueing system with applications in packet-switching systems". In *Proc. IEEE INFO-COM '88*, pp. 861-870, 1988.

[33] J. S. Turner. "Design of a broadband packet switching network". *IEEE Trans. Commun.*, vol. COM-36, no. 6, pp. 734-743, June 1988.

[34] Y. S. Yeh, M. G. Hluchyj, and A. S. Acampora. "The Knockout Switch: a simple, modular architecture for high-performance packet switching". *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 8, pp. 1274-1283, Oct. 1987.

[35] C. Yuan. *"On the performance of protocols to support integrated voice and data services"*. PhD thesis, Dept. of Electrical Engineering - Systems, Univ. of Southern California, Dec. 1988.

[36] C. Yuan and J. A. Silvester. "Queueing analysis of delay constrained voice traffic in a packet switching system". *IEEE J. Select. Areas Commun.*, vol. SAC-7, no. 5, pp. 729-738, June 1989.

[37] T.-S. P. Yum and T.-Y. Nagi. "Resequencing of messages in communication networks". *IEEE Trans. Commun.*, vol. COM-34, no. 2, pp. 143-149, Feb. 1986.