Report Number: WUCS-91-22

1991-02-01

Performance Evaluation of a Traffic Control Mechanism for ATM Networks

Andreas D. Bovopoulos

Future ATM networks will be required to support a plethora of services, transaction types and cell sequence behaviors with performance guarantees. Before this goal can be realized, however, some basically problems related to bandwidth allocation and traffic control in the ATM layer must be resolve. Such problems will in all likelihood defy solution as long as they studied in isolation without a unifying traffic characterization and traffic control framework. This work presented in this paper is part of an ongoing effort directed at the development of an integrated traffic characterization and control infrastructure for ATM networks. In this paper... **Read complete abstract on page 2.**

Performance Evaluation of a Traffic Control Mechanism for ATM Networks

Andreas D. Bovopoulos

Complete Abstract:

Future ATM networks will be required to support a plethora of services, transaction types and cell sequence behaviors with performance guarantees. Before this goal can be realized, however, some basically problems related to bandwidth allocation and traffic control in the ATM layer must be resolve. Such problems will in all likelihood defy solution as long as they studied in isolation without a unifying traffic characterization and traffic control framework. This work presented in this paper is part of an ongoing effort directed at the development of an integrated traffic characterization and control infrastructure for ATM networks. In this paper a traffic control mechanism capable of monitoring and controlling a rich family of incoming traffic behaviors is introduced. The family of support traffic behaviors is clearly characterized. Some possible roles for the traffic control mechanism in a traffic control infrastructure are presented along with its detailed performance analysis. Finally, some new results on the stochastic behavior of a bursty source are presented.

Performance Evaluation of a Traffic Control Mechanism for ATM Networks

Andreas D. Bovopoulos

WUCS-91-22

February 1991

Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
Saint Louis, MO 63130-4899

This work was supported by the National Science Foundation under grant NDC-9110183 and an industrial consortium of Bellcore, BNR, DEC, Italtel SIT, NEC and SynOptics.

Presented in part at INFOCOM '92.

Performance Evaluation of a Traffic Control Mechanism for ATM Networks†

Andreas D. Bovopoulos

Department of Computer Science
and
Computer and Communications Research Center
Washington University in St. Louis

ABSTRACT

Future ATM networks will be required to support a plethora of services, transaction types and cell sequence behaviors with performance guarantees. Before this goal can be realized, however, some basic problems related to bandwidth allocation and traffic control in the ATM layer must be resolved. Such problems will in all likelihood defy solution as long as they are studied in isolation without a unifying traffic characterization and traffic control framework.

The work presented in this paper is part of an ongoing effort directed at the development of an integrated traffic characterization and control infrastructure for ATM networks. In this paper a traffic control mechanism capable of monitoring and controlling a rich family of incoming traffic behaviors is introduced. The family of supported traffic behaviors is clearly characterized. Some possible roles for the traffic control mechanism in a traffic control infrastructure are presented along with its detailed performance analysis. Finally, some new results on the stochastic behavior of a bursty source are presented.

† This work was supported by the National Science Foundation under grant NCR-9110183, and an industrial consortium of Bellcore, BNR, DEC, Italtel SIT, NEC, NTT, SynOptics.

Address: P.O. Box 1045, Washington University, St. Louis, MO 63130-4899, USA. e-mail: andreas@patti.wustl.edu, FAX: (314) 935-7302

1. Introduction

Research concerning ATM networks is currently focused in the study of teletraffic problems appearing in an ATM environment [3, 7, 14]. The fundamental problem of the ATM technology stems from the fact that the performance of an ATM statistical multiplexer depends on the time behavior of the incoming traffic. As a result, without sufficient traffic control provisions, an ATM network may fail to provide grade of service (GOS) guarantees to end users. Each connection in the ATM layer results in a cell sequence (CS) that can be both analyzed and controlled at one or more of the following time scales: call, burst, and cell [2, 8, 9]. At each of the time scales for which a CS control is provided, a resource allocation scheme can be introduced and classified as corresponding to the call, burst, or cell time scale.

In order to design an integrated traffic characterization and control infrastructure capable of guaranteeing a GOS and providing that GOS at the minimum cost, the following are required: (i) the design of a traffic control and resource allocation infrastructure capable of efficiently handling a wide variety of traffic behaviors, (ii) given a particular traffic control and resource allocation scheme, the determination of the GOS that can be provided to an incoming CS, (iii) given a particular GOS requirement, the determination of the traffic control and resource allocation scheme required to provide to an incoming CS the desired GOS at the minimum cost. The work presented in this paper is part of an ongoing effort directed towards satisfying these requirements.

In this paper, a traffic generator capable of modeling a broad spectrum of incoming traffic behaviors is introduced. A traffic control mechanism (TCM) capable of supporting a number of different cell time scale control and resource allocation schemes is described. For each of a variety of incoming CS time behaviors and a given control and resource allocation scheme, the GOS that can be provided by the TCM is determined.

In Section 2, the traffic control mechanism (TCM) and its control capabilities are described. In addition, the TCM is compared with alternative traffic control mechanisms described in the literature. In Section 3, a traffic model capable of describing a rich family of traffic behaviors is introduced. In Section 4, the performance analysis of the TCM under the assumption that the cell service time is zero is presented. In Section 5, the performance analysis of the TCM for the case in which the cell service time is assumed to be a constant positive number is presented. In Section 6, the generality of the traffic model introduced in Section 3 is demonstrated. In addition, some new results regarding the behavior of a bursty source are presented. The paper ends with examples and the discussion of a number of research problems that must be solved for the further development of the traffic characterization and control infrastructure.

2. The Traffic Control Mechanism (TCM)

In this section the TCM is described. The TCM (Fig. 1) has a cell buffer of size K and a token pool of size M. Time is divided into units called slots, where the time interval corresponding to a slot equals the service time of a cell. Incoming cells are served by the TCM in the order of their arrival. Tokens act like permits for cell service. The cell at the front of the buffer is served at the beginning of a slot if a token exists, in which case a token is removed from the token buffer. If no token is available at the beginning of a slot, no cell is serviced during that time slot. Incoming cells are stored in the cell buffer as long as the buffer is not full; if the buffer is full, cells are discarded.

With the control and resource allocation just described, the TCM provides cell level control and resource allocation; i.e. control and bandwidth allocation decisions are made at the cell time scale. In addition to this cell level control, burst or call level controls and resource allocation schemes can be utilized. The focus of this paper, however, is on cell level control and resource allocation. Issues relating to burst level control and its interaction with cell level control will be the subject of subsequent papers.

For reasons explained below, tokens are generated *with a deterministic and periodic pattern*. Specifically tokens are generated in accordance with a T-state, cyclic, deterministic Markov chain. The number of tokens generated in state l of the Markov chain is equal to s_l for $0 \leq l \leq T-1$. Once generated, a token is either

(i) immediately utilized by a cell requiring service, (ii) stored in the token pool if there is no cell requiring service and the token pool is not full, or (iii) immediately discarded if the token cannot be immediately utilized by a cell and if the token pool is full. The time between two transitions of the Markov chain is D slots, and therefore the period of the deterministic and periodic pattern is $D \times T$ slots. The token patterns that can be generated in this fashion are deterministic and thus predictable and yet at the same time rich enough to implement a wide variety of control policies. The TCM control scheme can be easily implemented using two counters, one which records the state of the token pool and one which keeps track of the state of the token Markov chain.

Generating tokens in the manner just described is desirable for the following reasons. By appropriately selecting D and $T \times D$, the size of the cell buffer K, the size of the token pool M, and the number of tokens generated at each transition of the Markov chain, s_l, for $0 \leq l \leq T-1$, a wide variety of GOSs can be provided to an incoming CS. Because tokens are generated according to a deterministic and periodic pattern, a TCM can optimally support deterministic and periodic incoming CSs; a deterministic and periodic incoming CS controlled by a TCM results in a departing CS which is also deterministic and periodic. Further by using a deterministic control scheme, an open loop control scheme can be implemented when desired.

Recently, attention has been given to a hierarchical family of protocols [5] based on the periodic exchange of information related to the transmission of bursts. Note that a burst is a group of packets, and a packet is the fundamental information block at the transport layer. The structure of the TCM makes the TCM the ideal mechanism through which hierarchical type protocols can be implemented at the cell level. In such an environment the TCM parameters remain fixed for the duration of a burst and may be reset from burst to burst.

The richness of the capabilities of the TCM can be better appreciated by comparing the TCM with the leaky bucket [13, 4] and virtual clock [15] control mechanisms. The leaky bucket mechanism is a special case of the TCM, which can emulate the behavior of the leaky bucket if the cell buffer is removed, the token pool is full when the first cell of the CS arrives, the rate with which tokens are generated equals the service rate of the leaky bucket controller, and $T = D = 1$. The major difference between the TCM and the virtual clock mechanism is that while the virtual clock mechanism attempts to guarantee first order statistics to incoming CSs and to regulate burst behaviors, the TCM can better regulate more complicated traffic behaviors. Furthermore, whereas the virtual clock mechanism attempts to eliminate bursty transmissions, the TCM can support some types of services resulting in bursty CSs via the generation of bursty token patterns. To summarize, the TCM can support a much richer family of traffic behaviors and control patterns than either the leaky bucket or virtual clock schemes. Whereas the leaky bucket is envisioned as a good mechanism for admission control and the virtual clock as a good mechanism for transmission control, the TCM can be used to provide control both inside the network as well as at network boundaries (Fig. 2).

Efficient utilization of the TCM in a bandwidth allocation and traffic control infrastructure requires that methodologies be developed to determine the GOS that can be provided by the TCM to an incoming CS. The development of such methodologies is the subject of the remainder of this paper.

Performance analyses of variations of the TCM have been presented in the literature. In [11] for a Poisson source, the Laplace transform of the waiting time distribution and the inter-departure distribution were computed for the case in which the TCM parameters were $T = 1$, $s_0 = 1$, and $K = \infty$. The throughput was also computed for the case of a finite K. In [12], the throughput of the TCM was evaluated for the special case in which K is infinite, $T = 1$, and $s_0 = 1$. In [1, 2] a variation of the TCM was considered as a rate control throttle. For a Markovian arrival process source and an independent renewal token source, it was shown that the throughput was a function of $K + M$. The subject of the comparison of alternative traffic policing mechanisms was the subject of investigation of [6, 10].

The TCM is a new traffic control mechanism. It differs from the variations of the leaky bucket mechanism that have appeared in the literature in the fact that it is integrated into the traffic characterization and control infrastructure of the network and in the fact that it can monitor and control a much richer family of traffic behaviors. The analysis presented in this paper is complete and informative for performance parameters including loss distribution, waiting time distribution, token loss distribution, cell buffer occupancy and token buffer occupancy.

The TCM performance analysis presented in this paper is developed in the following manner. First, a traffic generator capable of modeling a broad spectrum of traffic behaviors is introduced in Section 3. This traffic generator is used in all subsequent performance studies. Next the TCM performance is evaluated under the assumptions that the service time of a cell is zero and that a slot is not defined with respect to the cell service time, i.e. a slot is simply a time interval of unit length. These assumptions are used in [1, 2, 11] as well to simplify analysis. A TCM operating under the zero cell service time assumption is symbolized by TCMZ (TCM Zero) and analyzed in Section 4. Later, the simplifying assumptions are removed, and the TCM performance is evaluated assuming a constant positive cell service time equal to a unit of time referred to as a slot. A TCM operating under the constant positive cell service time assumption is symbolized by TCMP (TCM Positive) and analyzed in Section 5.

3. The Traffic Source Generator Model

The token source Markov chain changes state every D slots. The state of the TCM is observed at each transition of the token Markov chain. The transition times of the token source are defined as *renewal epochs*. The model of the traffic source considered in the present paper is of the following form: At the renewal epochs, the behavior of the cell source is described by an embedded S-state homogeneous, irreducible and aperiodic Markov chain, with states $\{\sigma_0, \sigma_1, \cdots, \sigma_{S-1}\}$. Let σ and $\hat{\sigma}$ be any two states of the embedded Markov chain. Let $p_{\sigma\hat{\sigma}}$ be the probability that the source is in state $\hat{\sigma}$ at a renewal epoch, given that its state at the previous renewal epoch was σ. In addition let p_σ be the equilibrium probability that the source is in state σ at a renewal epoch.

Let ${}^\sigma A(u)$, for every u, $u \geq 0$, be the total number of arrivals in the u units of time immediately following a renewal epoch, given that the state of the source at the renewal epoch was σ. Further, let ${}^\sigma\alpha_k(u) \stackrel{\text{def}}{=} P[{}^\sigma A(u) = k]$, for every u, $u \geq 0$. For simplicity the notation ${}^\sigma\alpha_k$ will be used to refer to ${}^\sigma\alpha_k(D)$. Given that the state of the source at the renewal epoch is σ, let ${}^\sigma A_n(u)$, for $0 \leq u \leq 1$, be the total number of arrivals in the first u units of time of the n^{th} slot immediately following the renewal epoch, where $0 \leq n \leq D$. Notice that ${}^\sigma A_n(u) = {}^\sigma A(n+u) - {}^\sigma A(n)$. Let ${}^\sigma\alpha_k^n(u) \stackrel{\text{def}}{=} P[{}^\sigma A_n(u) = k]$. For simplicity the notation ${}^\sigma A_n$ will be used to refer to ${}^\sigma A_n(1)$, and the notation ${}^\sigma\alpha_k^n$ will be used to refer to ${}^\sigma\alpha_k^n(1)$.

The traffic load is the mean number of arriving cells in a slot time interval and is given by

$$Load \stackrel{\text{def}}{=} \frac{1}{D}\sum_{\forall\sigma}\sum_{k=1}^{\infty} k \times {}^\sigma\alpha_k \times p_\sigma = \frac{1}{D}\sum_{\forall\sigma}\sum_{n=0}^{D-1}\sum_{k=1}^{\infty} k \times {}^\sigma\alpha_k^n \times p_\sigma \quad . \tag{3.1}$$

The traffic source model described above is rich yet amenable to performance analysis. In Section 6, a number of well known source models are shown to be special cases of this model.

4. Performance Analysis of a TCMZ

4.1. Expected Throughput of a Traffic Source Controlled by a TCMZ

As in [11], let q be the number of cells in the cell buffer and w be the number of tokens in the token pool. Notice that the probability that both q and w are positive is zero, since any available tokens are instantaneously utilized by cells awaiting service. The state of the cell buffer and token pool is described by the sum of the number of tokens required to service all cells in the cell buffer and the number of tokens required to fill the empty positions in the token pool. Therefore when $q = 0$, the state of the cell buffer and token pool is $M - w$, for $0 \leq w \leq M$. When $w = 0$, the state of the cell buffer and token pool is $q + M$, for $0 \leq q \leq K$.

The state of the TCMZ at a renewal epoch is described by (j,l,σ), where j is the state of the cell buffer and token pool immediately before the renewal epoch, l is the state that the token Markov chain enters at the renewal epoch, and σ is the state of the cell source at the renewal epoch. For any real number x, with $x = x^+ - x^-$, $x^+ = \max\{0,x\}$, and $x^- = \max\{0,-x\}$,

$$q = (j-M)^+ = (M-j)^- \tag{4.1}$$

and

$$w = (M-j)^+ \quad . \tag{4.2}$$

Let

$$w(j,l,\sigma) \stackrel{\text{def}}{=} min\{(w-q+s_l)^+, M\} = min\{((M-j)^+ - (M-j)^- + s_l)^+, M\} \quad .$$

Then,

$$w(j,l,\sigma) = min\{(M-j+s_l)^+, M\} \quad . \tag{4.3}$$

Similarly, let

$$q(j,l,\sigma) \stackrel{\text{def}}{=} (q-s_l)^+ = (j-M-s_l)^+ \quad . \tag{4.4}$$

Notice that if (j,l,σ) is the state of the system at a renewal epoch, $w(j,l,\sigma)$ and $q(j,l,\sigma)$ are the total number of available tokens and cells respectively immediately after the renewal epoch. Let

$$\zeta(j,l,\sigma) \stackrel{\text{def}}{=} w(j,l,\sigma) - q(j,l,\sigma) \quad .$$

Note that $w(j,l,\sigma)$ and $q(j,l,\sigma)$ cannot both be positive. Then $\zeta^+(j,l,\sigma) = w(j,l,\sigma)$ and $\zeta^-(j,l,\sigma) = q(j,l,\sigma)$.

The state of the cell and token buffers immediately before a renewal epoch is completely described by the parameter j. Given that the state of the cell and token buffers immediately before a renewal epoch is j, let ${}^\sigma_l b_j^{j^*}$ be the probability that the state of the two buffers immediately before the following renewal epoch is j^*.

$${}^\sigma_l b_j^{j^*} \stackrel{\text{def}}{=} \begin{cases} {}^\sigma\alpha_{j^*-(j-s_l)^+}, & \text{if} j^* - (j-s_l)^+ \geq 0 \text{ and } 0 \leq j^* < K+M \\ 1 - \sum_{n=0}^{K+M-1} {}^\sigma_l b_j^n, & \text{if } j^* = K+M, \\ 0, & \text{otherwise} \end{cases} \tag{4.5}$$

for $0 \leq j \leq K+M$, $0 \leq l \leq T-1$, and $\sigma \in \{\sigma_0, \sigma_1, \cdots, \sigma_{S-1}\}$.

Let the square matrix ${}^\sigma_l\mathbf{P}$ be defined such that ${}^\sigma_l\mathbf{P}(j,j^*) \stackrel{\text{def}}{=} {}^\sigma_l b_j^{j^*}$ for $0 \leq j, j^* \leq K+M$. In short,

$${}^\sigma_l\mathbf{P} \stackrel{\text{def}}{=} \left[{}^\sigma_l b_j^{j^*}\right] \quad . \tag{4.6}$$

With ${}^\sigma_l\pi_j$ defined as the equilibrium probability that the system is in state (j,l,σ) at a renewal epoch, let ${}^\sigma_l\pi \stackrel{\text{def}}{=} [{}^\sigma_l\pi_0,\ {}^\sigma_l\pi_1,\ \cdots,\ {}^\sigma_l\pi_{K+M}]$, and ${}^\sigma\pi \stackrel{\text{def}}{=} [{}^\sigma_0\pi\ \ {}^\sigma_1\pi\ \ \cdots\ \ {}^\sigma_{T-1}\pi]$. In Appendix I, an iterative algorithm is presented for the computation of the equilibrium probabilities.

The throughput, i.e. the expected number of cells passing through the TCMZ per unit time, is symbolized by $E\gamma$ and given by

$$E\gamma = \sum_{\forall\,\sigma} \sum_{l=0}^{T-1} \sum_{j=0}^{K+M} \frac{{}^\sigma_l\pi_j}{D} \Big\{ \big(K + \zeta(j,l,\sigma)\big) + \sum_{n=0}^{K+\zeta(j,l,\sigma)-1} {}^\sigma\alpha_n \times \Big(n - \big(K + \zeta(j,l,\sigma)\big)\Big) \Big\} \quad . \tag{4.7}$$

The quantity $K + \zeta(j,l,\sigma)$, the definition of state, and the matrix ${}^{\sigma_i}_l\mathbf{P}$ for all l, $0 \leq l \leq T-1$ and for all σ_i, $0 \leq i \leq S-1$, depend on K and M only through the sum $K+M$. Therefore *the equilibrium probabilities and throughput of the TCMZ depend on K and M only through the sum K+M.*

4.2. Loss Distributions Associated with a Traffic Source Controlled by a TCMZ

Let L be the random variable describing the number of cells lost between two successive renewal epochs. Then,

$$P[L = k] = \begin{cases} \sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{j=0}^{K+M} {}_l^\sigma\pi_j \times \left\{\sum_{n=0}^{K+\zeta(j,l,\sigma)} {}^\sigma\alpha_n\right\}, & \text{if } k = 0 \\ \sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{j=0}^{K+M} {}_l^\sigma\pi_j \times {}^\sigma\alpha_{K+\zeta(j,l,\sigma)+k}, & \text{if } k \geq 1 \end{cases} . \tag{4.8}$$

From the last expression, the expected number of cells lost in a slot time interval can be computed.

$$E[L] = \frac{1}{D} \sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{j=0}^{K+M} {}_l^\sigma\pi_j \times \left\{ \sum_{n=K+\zeta(j,l,\sigma)}^{\infty} \Big(n - (K + \zeta(j,l,\sigma))\Big) \times {}^\sigma\alpha_n \right\} . \tag{4.9}$$

The total number of cells arriving at the TCMZ in a slot time interval is given by

$$E[L] + E\gamma = \frac{1}{D} \sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{j=0}^{K+M} {}_l^\sigma\pi_j \times \left\{\sum_{n=0}^{\infty} n \times {}^\sigma\alpha_n\right\} = \frac{1}{D} \sum_{\forall\sigma} p_\sigma \left\{\sum_{n=0}^{\infty} n \times {}^\sigma\alpha_n\right\} = Load . \tag{4.10}$$

From the above expressions, the probability that an arriving cell is lost, symbolized by P_L, is given by

$$P_L = \frac{E[L]}{Load} . \tag{4.11}$$

During any time interval in which the cell buffer is full, arriving cells are lost. Each lost cell can be thought as being a member of a group of consecutive lost cells. The last cell in the group is always the last cell lost before the renewal epoch at which new tokens are generated, cells in the cell buffer are serviced, and positions in the cell buffer are freed. The distribution of the number of cells in the group of consecutive lost cells is now computed.

Let the function $f(l)$, where l is the state of the token generator Markov chain, be defined as follows. $f(l)$ takes on the value 0 if $s_l = 0$; else, $f(l)$ equals the number of state transitions of the token generator Markov chain starting with state l and ending with the state at which additional tokens are generated. That is,

$$f(l) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{if } s_l = 0, \\ n, & \text{if } s_l \neq 0,\ s_{(l+1)\ mod\ T} = 0, \cdots, s_{(l+n-1)\ mod\ T} = 0, s_{(l+n)\ mod\ T} \neq 0, \end{cases} \tag{4.12}$$

for $0 \leq l \leq T-1$.

Now let R_b be the random variable describing the number of cells lost in the time interval occuring between two successive points in time at which tokens are generated. Then the probability that $R_b = k$ is given by

$$P[R_b = k] = \begin{cases} \sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{j=0}^{K+M} {}_l^\sigma\pi_j \times \left\{\sum_{n=0}^{K+\zeta(j,l,\sigma)} {}^\sigma\alpha_n\big(f(l)D\big)\right\}, & \text{if } k = 0 \\ \sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{j=0}^{K+M} {}_l^\sigma\pi_j \times {}^\sigma\alpha_{K+\zeta(j,l,\sigma)+k}\big(f(l)D\big), & \text{if } k \geq 1 \end{cases} . \tag{4.13}$$

Given one or more cells are lost, each lost cell belongs to a group of consecutive lost cells of size greater than or equal to one. Let L_b represent the size of a group of consecutive lost cells. Then,

$$P[L_b = k] = \frac{P[R_b = k]}{1 - P[R_b = 0]} , \tag{4.14}$$

for all k, $k \geq 1$.

Given the dependence of the equilibrium probabilities on K and M only through sum $K + M$, the distributions of the random variables L, R_b, and L_b associated with the TCMZ depend on K and M only through the sum K+M.

4.3. Loss Distribution Associated with a TCMZ Token Generator

Let L_T be the random variable that describes the number of lost tokens at a renewal epoch. If the TCMZ state is (j, l, σ), the value of L_T is $(s_l - j)^+$.

$$P[L_T = k] = \sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{j=0}^{K+M} {}_l^\sigma\pi_j \times 1(s_l - j = k) \quad , \tag{4.15}$$

for $k \geq 0$, where $1(s_l - j = k)$ is the indicator function which takes the value one if the expression in paretheses is true and otherwise takes the value zero. Further,

$$E[L_T] = \sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{j=0}^{K+M} {}_l^\sigma\pi_j \times (s_l - j)^+ \quad . \tag{4.16}$$

Given that the equilibrium probabilities depend on K and M only through the sum $K + M$, the distribution of the random variable L_T corresponding to the TCMZ depends on K and M only through the sum $K + M$.

4.4. The Waiting Time Distribution Associated with a Traffic Source Controlled by a TCMZ

Let (j, l, σ) be the state of the TCMZ at a renewal epoch. A cell arriving u units of time after the renewal epoch is tagged, where $u \in [0, D)$. If $1 \leq {}^\sigma A(u) \leq \zeta^+(j, l, \sigma)$, the tagged cell is served instantaneously. If $1 + \zeta^+(j, l, \sigma) \leq {}^\sigma A(u) \leq K + \zeta(j, l, \sigma)$, then the tagged cell must wait until enough tokens have been generated to serve all cells arriving before the tagged cell as well as the tagged cell itself. If $K + \zeta(j, l, \sigma) < {}^\sigma A(u)$, the tagged cell is lost. The probability that the tagged cell arriving at time u is not lost is given by $P[{}^\sigma A(u) - \zeta(j, l, \sigma) \leq K]$.

If $1 + \zeta^+(j, l, \sigma) \leq {}^\sigma A(u) \leq K + \zeta(j, l, \sigma)$, the number of renewal epochs required to generate a sufficient number of tokens is symbolized by $G_{{}^\sigma A(u)}(j, l, \sigma)$, where

$$G_{{}^\sigma A(u)}(j, l, \sigma) = \min_n \{n | \sum_{m=1}^{n} s_{(l+m) \bmod T} \geq {}^\sigma A(u) - \zeta^+(j, l, \sigma)\} \quad ,$$

and the waiting time of the tagged cell is $D - u + (G_{{}^\sigma A(u)}(j, l, \sigma) - 1)D = G_{{}^\sigma A(u)}(j, l, \sigma)D - u$.

Let $W_k^u(l, j, \sigma)$ be the random variable describing the waiting time of the k^{th} accepted cell arriving u units of time after the last renewal. With (l, j, σ) referring to the state of the TCMZ at the most recent renewal epoch, if $1 + \zeta^+(j, l, \sigma) \leq {}^\sigma A(u) \leq K + \zeta(j, l, \sigma)$,

$$P[W_k^u(l, j, \sigma) > x] = P[G_k(j, l, \sigma)D - u > x] = 1\Big(G_k(j, l, \sigma)D - u > x\Big) \quad , \tag{4.17}$$

for any $x \geq 0$. If $1 \leq k \leq \zeta^+(j, l, \sigma)$, $P[W_k^u(l, j, \sigma) = 0] = 1$.

With σ referring to the state of the traffic source at the time of the last renewal, let $F_k^\sigma(u)$ refer to the distribution of the random variable u, where u is the arrival time of the k^{th} non-lost cell arriving since the last renewal. Further, let W be the random variable describing the waiting time distribution of an arriving cell which is not lost. The probability that the waiting time is greater than x units of time, given that the state of the TCMZ at the most recent renewal epoch was (l, j, σ), is given by

$$P[W > x | (j, l, \sigma)] = \frac{\sum_{k=\zeta^+(j,l,\sigma)+1}^{K+\zeta(j,l,\sigma)} \int_0^D P[W_k^u(l, j, \sigma) > x] dF_k^\sigma(u)}{\sum_{k=1}^{K+\zeta(j,l,\sigma)} \int_0^D dF_k^\sigma(u)} = \frac{\sum_{k=\zeta^+(j,l,\sigma)+1}^{K+\zeta(j,l,\sigma)} \int_0^{U_k(x,j,l,\sigma)} dF_k^\sigma(u)}{\sum_{k=1}^{K+\zeta(j,l,\sigma)} \int_0^D dF_k^\sigma(u)}$$

for any $x \geq 0$, where $U_{A(u)}(x,j,l,\sigma) \stackrel{\text{def}}{=} min\{D, (G_{^\sigma A(u)}(j,l,\sigma)D - x)^+\}$. Therefore,

$$P[W > x|(j,l,\sigma)] = \begin{cases} \frac{\sum_{k=\zeta^+(j,l,\sigma)+1}^{K+\zeta(j,l,\sigma)} P[^\sigma A(U_k(x,j,l,\sigma)) \geq k]}{\sum_{k=1}^{K+\zeta(j,l,\sigma)} P[^\sigma A(D) \geq k]}, & \text{if } \sum_{k=1}^{K+\zeta(j,l,\sigma)} P[^\sigma A(D) \geq k] > 0 \\ undefined, & \text{otherwise} \end{cases} \tag{4.18}$$

where

$$P[^\sigma A(u) \geq k] = 1 - \sum_{k=0}^{k-1} {^\sigma\alpha_k(u)}$$

for $x \geq 0$, $0 \leq u \leq D$, and $k \geq 1$.

Let

$$\bar{P}[W > x|(j,l,\sigma)] = \begin{cases} P[W > x|(j,l,\sigma)], & \text{if } \sum_{k=1}^{K+\zeta(j,l,\sigma)} P[^\sigma A(D) \geq k] > 0 \\ 0, & \text{otherwise} \end{cases} \tag{4.19}$$

for $x \geq 0$, $0 \leq u \leq D$, and $k \geq 1$. Similarly,

$$\bar{P}[W \geq 0|(j,l,\sigma)] = \begin{cases} 1, & \text{if } \sum_{k=1}^{K+\zeta(j,l,\sigma)} P[^\sigma A(D) \geq k] > 0 \\ 0, & \text{otherwise} \end{cases} . \tag{4.20}$$

The waiting time distribution of any accepted cell is

$$P[W > x] = \frac{\sum_{l=0}^{T-1} \sum_{j=0}^{K+M} \sum_{\forall\sigma} {^\sigma_l\pi_j} \bar{P}[W > x|(j,l,\sigma)]}{\sum_{l=0}^{T-1} \sum_{j=0}^{K+M} \sum_{\forall\sigma} {^\sigma_l\pi_j} \bar{P}[W \geq 0|(j,l,\sigma)]} , \tag{4.21}$$

for any $x \geq 0$. $P[W = 0] = 1 - P[W > 0]$.

Because the lower bound of the summation appearing in the numerator of (4.19) $\zeta^+(j,l,\sigma) + 1$ is not a function of the sum $K + M$ but depends rather on the individual K and M values, *the waiting time distribution corresponding to any accepted cell depends explicitly on the individual K and M values.*

5. Performance Analysis of a TCMP

5.1. Expected Throughput of a Traffic Source Controlled by a TCMP

The state of the TCMP immediately before the beginning of the n^{th} slot since the last renewal, for all n, $0 \leq n \leq D-1$, is described by (j_c, j_t, l, σ), where j_c and j_t are the total number of cells and tokens respectively in the TCMP at the end of the previous slot, and l and σ are the states of the token source and cell source respectively at the beginning of the 0^{th} slot. Let

$$s_l^n = \begin{cases} s_l & \text{if } n = 0 \\ 0 & \text{otherwise} \end{cases} ,$$

for $0 \leq n \leq D-1$, and $0 \leq l \leq T-1$.

Let ${^\sigma_l b^{j_c^*,j_t^*}_{j_c,j_t}}(n)$ be the transition probability that the TCM will move from state (j_c, j_t, l, σ) immediately before the beginning of the n^{th} slot to a state with j_c^* cells and j_t^* tokens immediately before the beginning of the following slot. Using the notation $\lceil x \rceil^M \stackrel{\text{def}}{=} \min\{x, M\}$, and $\lfloor x \rfloor_M \stackrel{\text{def}}{=} \max\{x, M\}$, the transition probabilities are defined for $0 \leq j_c, j_c^* \leq K$, $0 \leq j_t, j_t^* \leq M$, $0 \leq n \leq D-1$ and $0 \leq l \leq T-1$, as follows:

$${^\sigma_l b^{j_c^*,j_t^*}_{j_c,j_t}}(n) = \begin{cases} {^\sigma\alpha^n_{j_c^*-j_c+1(j_t+s_l^n>0,j_c>0)}} & \text{if } j_c^* < K,\ j_t^* = \lceil j_t + s_l^n - 1(j_t + s_l^n > 0, j_c > 0) \rceil^M \\ 1 - \sum_{j=\lfloor j_c-1 \rfloor_0}^{K} {^\sigma\alpha^0_{j-j_c+1(j_t+s_l^n>0,j_c>0)}} & \text{if } j_c^* = K,\ j_t^* = \lceil j_t + s_l^n - 1(j_t + s_l^n > 0, j_c > 0) \rceil^M \\ 0 & \text{otherwise} \end{cases}$$

Let the square matrix ${}^{\sigma}_{l}\mathbf{P}_n$ be defined such that ${}^{\sigma}_{l}\mathbf{P}_n(j_t*(K+1)+j_c, j_t^**(K+1)+j_c^*) \stackrel{\text{def}}{=} {}^{\sigma}_{l}b^{j_c^*,j_t^*}_{j_c,j_t}(n)$ for $0 \le j_c, j_c^* \le K$, $0 \le j_t, j_t^* \le M$, and $0 \le n \le D-1$. In short,

$$ {}^{\sigma}_{l}\mathbf{P}_n \stackrel{\text{def}}{=} \left[{}^{\sigma}_{l}b^{j_c^*,j_t^*}_{j_c,j_t}(n)\right] \quad . \tag{5.1}$$

With ${}^{\sigma}_{l}\pi^{(n)}_{j_c,j_t}$ defined as the equilibrium probability that the system is in state (j_c, j_t, l, σ) immediately before the beginning of the n^{th} slot since the last renewal epoch, let ${}^{\sigma_i}_{l}\pi^{(n)} \stackrel{\text{def}}{=} [{}^{\sigma_i}_{l}\pi^{(n)}_{0,0}, {}^{\sigma_i}_{l}\pi^{(n)}_{1,0}, \cdots, {}^{\sigma_i}_{l}\pi^{(n)}_{K,M}]$ for every l, $0 \le l \le T-1$, for every n, $0 \le n \le D-1$ and for every i, $0 \le i \le S-1$. In Appendix II, an iterative algorithm is presented for the computation of the equilibrium probabilities ${}^{\sigma_i}_{l}\pi^{(n)}$ for all n, $0 \le n \le D-1$, for all l, $0 \le l \le T-1$ and for every i, $0 \le i \le S-1$. Notice that $\sum_{i=0}^{S-1}\sum_{l}^{T-1} {}^{\sigma_i}_{l}\pi^{(n)}\mathrm{e} \; = \; 1/D$, for all n, $0 \le n \le D-1$.

If j_c and j_t are the total number of cells and tokens respectively immediately before the beginning of the n^{th} slot, a departure will occur at the end of the n^{th} slot if $j_c > 0$ and $j_t + s_l^n > 0$. Let ${}^{\sigma}_{l}\bar{\mathbf{P}}_n$ be the column vector whose element $j_t*(K+1)+j_c$ is one if $j_c > 0$, $j_t + s_l^n > 0$ and is zero otherwise, for $0 \le j_c \le K$ and $0 \le j_t \le M$.

The throughput, i.e. the expected number of cells passing through the TCMP per unit time, is symbolized by $E\gamma$ and given by

$$ E\gamma \; = \; \sum_{\forall\sigma}\sum_{l=0}^{T-1}\sum_{n=0}^{D-1} {}^{\sigma}_{l}\pi^{(n)}\,{}^{\sigma}_{l}\bar{\mathbf{P}}_n \quad . \tag{5.2}$$

5.1.1. The Loss Distribution in a Slot Time Interval

Let L_s be the random variable describing the number of cells lost in a slot time interval. Further, let ${}^{\sigma}_{l}Y^{(k;n)}_{j_c,j_t}$ be the probability that k cells are lost during the n^{th} slot since the last renewal epoch, where $0 \le n \le D-1$, given that the system is in state (j_c, j_t, l, σ) immediately before the beginning of the n^{th} slot. Then for $0 \le n \le D-1$,

$$ {}^{\sigma}_{l}Y^{(k;n)}_{j_c,j_t} \stackrel{\text{def}}{=} \begin{cases} \sum_{j_c^*=\lfloor j_c-1\rfloor_0}^{K} {}^{\sigma}\alpha^{0}_{j_c^*-j_c+1(j_t+s_l^n>0,j_c>0)} & \text{if } k=0 \\ {}^{\sigma}\alpha^{n}_{K+k-j_c+1(j_t+s_l^n>0,j_c>0)} & \text{if } k \ge 1 \end{cases} \quad .$$

If ${}^{\sigma}_{l}\mathbf{Y}^{(n)}_{k} \stackrel{\text{def}}{=} [{}^{\sigma}_{l}Y^{(k;n)}_{0,0}, {}^{\sigma}_{l}Y^{(k;n)}_{1,0}, \cdots, {}^{\sigma}_{l}Y^{(k;n)}_{K,M}]^T$ for $0 \le n \le D-1$, then for $k \ge 0$, the probability that k cells are lost in a slot time interval is given by

$$ P[L_s = k] \; = \; \sum_{\forall\sigma}\sum_{l=0}^{T-1}\sum_{n=0}^{D-1} {}^{\sigma}_{l}\pi^{(n)}\,{}^{\sigma}_{l}\mathbf{Y}^{(n)}_{k} \quad . \tag{5.3}$$

5.1.2. The Loss Distribution Between Two Successive Renewal Epochs

Let ${}^{\sigma}_{l}h^{j_c^*,j_t^*}_{j_c,j_t}(k,n)$ be the probability that k cells are lost during the n^{th} slot, given that immediately before the slot there are j_c and j_t cells and tokens respectively in the TCMP and given that at the end of the slot there are j_c^* and j_t^* cells and tokens respectively in the TCMP. Then

$$ {}^{\sigma}_{l}h^{j_c^*,j_t^*}_{j_c,j_t}(k,n) \; = \; \begin{cases} {}^{\sigma}\alpha^{n}_{j_c^*-j_c+1(j_t+s_l^n>0,j_c>0)} & \text{if } k=0,\; j_t^* = \lceil j_t + s_l^n - 1(j_t+s_l^n>0, j_c>0)\rceil^M \\ {}^{\sigma}\alpha^{n}_{K+k-j_c+1(j_t+s_l^n>0,j_c>0)} & \text{if } k \ge 1,\; j_t^* = \lceil j_t + s_l^n - 1(j_t+s_l^n>0, j_c>0)\rceil^M \\ 0 & \text{otherwise} \end{cases} \quad ,$$

for $0 \le j_c \le K$, $0 \le j_t \le M$ and for $0 \le n \le D-1$. Let

$$ {}^{\sigma}_{l}\mathbf{H}^{(n)}_{k} \stackrel{\text{def}}{=} \left[{}^{\sigma}_{l}h^{j_c^*, j_t^*}_{j_c, j_t}(k,n) \right] \quad , \tag{5.4} $$

for $0 \le k$ and $0 \le n \le D-1$.

If

$$ {}^{\sigma}_{l}\mathbf{H}_{k} \stackrel{\text{def}}{=} \begin{cases} {}^{\sigma}_{l}\mathbf{Y}^{(0)}_{k} & \text{if } D = 1 \\ \sum_{k_0+\cdots+k_{D-1}=k} {}^{\sigma}_{l}\mathbf{H}^{(0)\sigma}_{k_0} \times \cdots \times {}^{\sigma}\mathbf{H}^{(D-2)\sigma}_{k_{D-2}\,l}\mathbf{Y}^{(D-1)}_{k_{D-1}} & \text{if } D \ge 2 \end{cases} \quad , \tag{5.5} $$

the probability that k cells are lost between two successive renewal epochs is given by

$$ P[L_s = k] = \frac{\sum_{\forall\sigma}\sum_{l=0}^{T-1} {}^{\sigma}_{l}\pi^{(0)\sigma}_{\ l}\mathbf{H}_k}{\sum_{\forall\sigma}\sum_{l=0}^{T-1} {}^{\sigma}_{l}\pi^{(0)}\mathbf{e}} = D\sum_{\forall\sigma}\sum_{l=0}^{T-1} {}^{\sigma}_{l}\pi^{(0)\sigma}_{\ l}\mathbf{H}_k \quad , \tag{5.6} $$

for all k, $k \ge 0$.

5.1.3. The Distribution of the Lost Tokens at the TCMP

Let L_T be the random variable describing the number of lost tokens at a renewal epoch. Given that $\sum_{i=0}^{S-1}\sum_{l}^{T-1} {}^{\sigma_i}_{l}\pi^{(0)}\mathbf{e} = 1/D$, if the TCMP state is (j_c, j_t, l, σ) immediately before a renewal epoch, the value of L_T is $\left(j_t + s_l - 1(j_c > 0) - M\right)^+$, with

$$ P[L_T = 0] = D\sum_{\forall\sigma}\sum_{l=0}^{T-1}\sum_{j_c=0}^{K}\sum_{j_t=0}^{M} 1\left(j_t + s_l - 1(j_c > 0) - M \le 0\right) {}^{\sigma}_{l}\pi^{(0)}_{j_c,j_t} \quad , \tag{5.7} $$

and

$$ P[L_T = k] = D\sum_{\forall\sigma}\sum_{l=0}^{T-1}\sum_{j_c=0}^{K}\sum_{j_t=0}^{M} 1\left(j_t + s_l - 1(j_c > 0) - M = k\right) {}^{\sigma}_{l}\pi^{(0)}_{j_c,j_t} \quad , \tag{5.8} $$

for all k, $k \ge 1$. Similarly,

$$ E[L_T] = D\sum_{\forall\sigma}\sum_{l=0}^{T-1}\sum_{j_c=0}^{K}\sum_{j_t=0}^{M} \left(j_t + s_l - 1(j_c > 0) - M\right)^+ {}^{\sigma}_{l}\pi^{(0)}_{j_c,j_t} \quad . \tag{5.9} $$

5.2. The Waiting Time Distribution of a Traffic Source at the TCMP

A cell arriving u time units after the beginning of the n^{th} slot following the last renewal is tagged, where $0 \le u < 1$. Let the TCMP state be (j_c, j_t, l, σ) immediately before the beginning of the n^{th} slot since the last renewal, for all n, $0 \le n \le D-1$. If $j_c + {}^{\sigma}A_n(u) \le K$, then the cell is accepted; otherwise it is rejected. Accepted cells are served in the order of their arrival. In what follows a methodology is introduced for the computation of the waiting time distribution of an accepted cell.

All slots are enumerated according to their proximity to the renewal epoch at which the token source most recently entered state $l = 0$. Each time slot is assigned a number between 0 and $DT - 1$. According to the enumeration rule, the first slot following the renewal epoch at which the token source most recently entered state $l = 0$ is numbered 0, the next slot is numbered 1, and so on.

Using the enumeration rule just described, let y_i be the number of tokens generated at the beginning of the i^{th} slot, for $0 \le i \le DT-1$.

$$y_i \stackrel{\text{def}}{=} \begin{cases} s_{\frac{i}{D} \bmod T} & \text{if } \lfloor \frac{i}{D} \rfloor = \frac{i}{D} \text{ and } 0 \le i \le DT, \\ 0, & \text{elsewhere.} \end{cases}$$

Let j_c and j_t be the number of cells and tokens respectively in the TCMP at the end of the i^{th} slot. Let $g(i, j_c, j_t)$ be the additional time spent by the j_c^{th} cell in the TCMP either in the cell buffer or in service. If $g(i, 0, j_t) = 0$, for $0 \le i \le DT-1$ and $j_t \ge 0$, then for $j_c \ge 1$ and all i, $0 \le i \le DT-1$,

$$g(i, j_c, j_t) = \begin{cases} 1 + g((i+1) \bmod DT, j_c, 0) & \text{if } j_t + y_{(i+1) \bmod DT} = 0 \\ 1 + g((i+1) \bmod DT, j_c - 1, \lceil j_t + y_{(i+1) \bmod DT} - 1 \rceil^M) & \text{if } j_t + y_{(i+1) \bmod DT} > 0 \end{cases} .$$

From the previous relations, $g(i, 1, j_t)$ can be computed for every j_t, $0 \le j_t \le M$. Based on the $g(i, 1, j_t)$ computations, $g(i, 2, j_t)$ can be computed for every j_t, $0 \le j_t \le M$. In this fashion, $g(i, j_c, j_t)$ can be computed for $0 \le i \le DT-1$, $j_c \ge 0$ and $j_t \ge 0$.

Let $W_k^n(u, j_c, j_t, l, \sigma)$ be the random variable describing the waiting time of the k^{th} accepted cell arriving u units of time after the beginning of the n^{th} slot since the last renewal, for $0 \le n \le D-1$ and $0 \le u < 1$, given that the TCMP state immediately before the beginning of the n^{th} slot is (j_c, j_t, l, σ). If $K_{max}^n(j_c, j_t, l, \sigma) \stackrel{\text{def}}{=} K - j_c + 1(j_c > 0)1(j_t + y_{lD+n} > 0)$,

$$W_k^n(u, j_c, j_t, l, \sigma) = 1 - u +$$
$$+ \begin{cases} g(lD+n, j_c+k, 0) & \text{if } j_t + y_{lD+n} = 0 \\ g(lD+n, k, \lceil j_t + y_{lD+n} \rceil^M) & \text{if } j_t + y_{lD+n} > 0,\ j_c = 0 \\ g(lD+n, j_c+k-1, \lceil j_t + y_{lD+n} - 1 \rceil^M) & \text{if } j_t + y_{lD+n} > 0,\ j_c > 0 \end{cases}$$

for $0 \le n \le D-1$, and $1 \le k \le K_{max}^n$.

Given that the state of the source at the last renewal epoch was σ, let $F_k^{(\sigma,n)}(u)$ be the distribution of the arrival time of the k^{th} cell during the n^{th} slot following the last renewal epoch, for $0 \le u < 1$ and $0 \le n \le D-1$. Then,

$$P[W^n(j_c, j_t, l, \sigma) > x] = \frac{\sum_{k=1}^{K_{max}^n(j_c,j_t,l,\sigma)} \int_0^1 1(W_k^n(u, j_c, j_t, l, \sigma) > x) dF_k^{(\sigma,n)}(u)}{\sum_{k=1}^{K_{max}^n(j_c,j_t,l,\sigma)} \int_0^1 dF_k^{(\sigma,n)}(u)} .$$

Notice that $W_k^n(u, j_c, j_t, l, \sigma)$ is a *decreasing* function of u for all k, $k \ge 0$. Let $U_k^n(x, j_c, j_t, l, \sigma)$ be the maximum value of u in the interval $[0, 1)$ for which $W_k^n(u, j_c, j_t, l, \sigma) > x$. Then,

$$P[W^n(j_c, j_t, l, \sigma) > x] = \begin{cases} \frac{\sum_{k=1}^{K_{max}^n(j_c,j_t,l,\sigma)} P[{}^\sigma A_n(U_k^n(x,j_c,j_t,l,\sigma)) \ge k]}{\sum_{k=1}^{K_{max}^n(j_c,j_t,l,\sigma)} P[{}^\sigma A_n \ge k]}, & \text{if } \sum_{k=1}^{K_{max}^n(j_c,j_t,l,\sigma)} P[{}^\sigma A_n \ge k] > 0; \\ undefined, & \text{otherwise,} \end{cases} \tag{5.10}$$

for all n, $0 \le n \le D-1$.

Let

$$\bar{P}[W^n(j_c, j_t, l, \sigma) > x] = \begin{cases} P[W^n(j_c, j_t, l, \sigma) > x], & \text{if } \sum_{k=1}^{K_{max}^n(j_c,j_t,l,\sigma)} P[{}^\sigma A_n \ge k] > 0; \\ 0, & \text{otherwise,} \end{cases} \tag{5.11}$$

for all n, $0 \le n \le D-1$. Let the column vector ${}^\sigma_l\bar{\mathbf{P}}[W^n > x]$ be defined such that its $(j_t * (K+1) + j_c)^{th}$ element is $\bar{P}[W^n(j_c, j_t, l, \sigma) > x]$, for all possible values of j_c and j_t. Then, if $P[W \le x]$ is the probability that any incoming cell will be delayed at most x units of time, then for x, $x \ge 0$,

$$P[W > x] = \frac{\sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{n=0}^{D-1} {}^\sigma_l\pi^{(n)} {}^\sigma_l\bar{\mathbf{P}}[W^n > x]}{\sum_{\forall\sigma} \sum_{l=0}^{T-1} \sum_{n=0}^{D-1} {}^\sigma_l\pi^{(n)} {}^\sigma_l\bar{\mathbf{P}}[W^n > 0]} . \tag{5.12}$$

6. Study of Different Traffic Sources

In order for a traffic source generator to be useful in an ATM environment, it must be capable of modeling a rich variety of traffic behaviors. A performance analysis based on such a traffic source generator is useful because it provides a uniform methodology through which a control infrastructure can be tested under diverse traffic conditions. In this section, a number of different traffic source models appearing in the ATM literature are presented and shown to be special cases of the traffic source model described in Section 3 and utilized throughout this paper. By showing how a wide variety of continuous and discrete time traffic source generators can be modeled using the traffic source model described in Section 3, it is demonstrated that the traffic generator on which the performance analysis presented in this paper is based is indeed capable of modeling a rich variety of traffic behaviors.

6.1. Continuous Time Traffic Source Models

6.1.1. An Alternating Renewal Process Traffic Source

In an ATM environment the majority of traffic sources will have bursty characteristics. In this subsection, the behavior of a bursty source modeled as an alternating renewal process is studied. The source alternates between two states, the idle state i and the active state a. During each visitation to state i, the distribution of the time that the source spends in that state i is given by $F_i(u)$. Similarly, during each visitation to state a, the distribution of the time that the source spends in that state a is given by $F_a(u)$. Let $F_i^{(k)}(u)$ (similarly $F_a^{(k)}(u)$) be the k^{th} convolution of $F_i(u)$ (similarly $F_a(u)$). Further, let ${}^iT_a(u)$ be the random variable describing the total time that the source spends in state a in the time interval $[0, u)$, given that at time 0 the source is in state i. The random variables ${}^iT_i(u)$, ${}^aT_a(u)$ and ${}^aT_i(u)$ are defined similarly.

Proposition 1: *The distribution of the total time that the source spends in the active state in the time interval [0,u], given that the state at time 0 was the idle state, is given by*

$$P[{}^iT_a(u) \leq x] \;=\; \sum_{k=0}^{\infty} F_a^{(k)}(x) \times \{F_i^{(k)}(u-x) \;-\; F_i^{(k+1)}(u-x)\} \quad , \tag{6.1}$$

for every x, $0 \leq x \leq u$. In particular $P[{}^iT_a(u) \leq u] \;=\; 1$ and $P[{}^iT_a(u) = 0] \;=\; 1 \;-\; F_i(u)$.
The distribution of the total time that the source spends in the active state in the time interval [0,u], given that the state at time 0 was the active state, is given by

$$P[{}^aT_a(u) < x] \;=\; 1 \;-\; \sum_{k=0}^{\infty} F_i^{(k)}(u \;-\; x) \times \{F_a^{(k)}(x) \;-\; F_a^{(k+1)}(x)\} \tag{6.2}$$

for every x, $0 \leq x \leq u$. In particular $P[{}^aT_a(u) = u] \;=\; F_a(u)$ and $P[{}^aT_a(u) = 0] \;=\; 0$.

Proof: (6.1) is proven first. Let τ_a^k (respectively τ_i^k) be the total time that the source spends in the active (respectively idle) state when it visits the active (respectively idle) state for the k^{th} time, for all k, $k \geq 1$. Let

$$N_i(u) \;\stackrel{\text{def}}{=}\; \begin{cases} 0, & \text{if } \tau_i^1 > u \\ sup\{k : \; \sum_{n=1}^{k} \tau_i^n \leq u\}, & \text{otherwise} \end{cases} \quad .$$

Then

$$P[{}^iT_a(u) \leq x] \;=\; 1 \times P[N_i(u-x) = 0] \;+\; \sum_{k=1}^{\infty} P[\sum_{n=1}^{k} \tau_a^n \leq x] \times P[N_i(u-x) = k] \quad .$$

Therefore,

$$P[{}^iT_a(u) \leq x] \;=\; \sum_{k=0}^{\infty} F_a^{(k)}(x) \times \{F_i^{(k)}(u-x) \;-\; F_i^{(k+1)}(u-x)\} \quad ,$$

and the proof of (6.1) is complete. Notice that

$$P[{}^aT_a(u) < x] = P[u - {}^aT_i(u) < x] = P[{}^aT_i(u) > u - x] = 1 - P[{}^aT_i(u) \leq u - x]. \tag{6.3}$$

Equations (6.3) and (6.1) prove (6.2).

Equation (6.2) can be also proven as follows:

$$\begin{aligned} P[{}^aT_a(u) < x] \;&=\; \int_{[0,x)} \{\sum_{k=0}^{\infty} F_a^{(k)}(x-y) \times \{F_i^{(k)}(u-x) \;-\; F_i^{(k+1)}(u-x)\}\}dF_a(y) \\ &=\; \sum_{k=0}^{\infty} F_a^{(k+1)}(x) \times \{F_i^{(k)}(u-x) \;-\; F_i^{(k+1)}(u-x)\} \\ &=\; \sum_{k=0}^{\infty} F_a^{(k+1)}(x) \times F_i^{(k)}(u-x) \;-\; \sum_{k=0}^{\infty} F_a^{(k+1)}(x) \times F_i^{(k+1)}(u-x) \\ &=\; \sum_{k=0}^{\infty} F_a^{(k+1)}(x) \times F_i^{(k)}(u-x) \;-\; \sum_{k=0}^{\infty} F_a^{(k)}(x) \times F_i^{(k)}(u-x) \;+\; 1 \quad . \end{aligned}$$

Therefore,

$$P[{}^aT_a(u) < x] \;=\; 1 \;-\; \sum_{k=0}^{\infty} F_i^{(k)}(u \;-\; x) \times \{F_a^{(k)}(x) \;-\; F_a^{(k+1)}(x)\} \quad .$$

■

6.1.2. A Two-State Markov Process Traffic Source

If the alternate renewal process studied in the previous subsection is a two state Markov process, let λ be the transition rate from state i to state a; similarly let μ be the transition rate from state a to state i. With this notation, $F_i(u) \;=\; 1 \;-\; e^{-\lambda u}$, and $F_a(u) \;=\; 1 - e^{-\mu u}$ for $u \geq 0$. The behavior of a two state Markov process is determined by the generator

$$\mathbf{Q} \;=\; \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \tag{6.4}$$

If $\mathbf{P}(u) = [p_i(u), p_a(u)]$ is the 1×2 vector describing the probability that the Markov process is in the idle and active states at time u, then $\mathbf{P}(u) \;=\; \mathbf{P}(0)e^{\mathbf{Q}u}$,

$$p_i(u) \;=\; \frac{\mu}{\mu+\lambda} \;+\; \{p_i(0) \;-\; \frac{\mu}{\mu+\lambda}\}e^{-(\lambda+\mu)u} \quad , \tag{6.5}$$

and

$$p_a(u) \;=\; \frac{\lambda}{\mu+\lambda} \;+\; \{p_a(0) \;-\; \frac{\lambda}{\mu+\lambda}\}e^{-(\lambda+\mu)u} \quad . \tag{6.6}$$

For this class of traffic sources, the following proposition holds.

Proposition 2: *The distribution of the total time that the source spends in the active state in the time interval [0,u], given that the state at time 0 was the idle state, is given by*

$$P[{}^iT_a(u) \le x] \;=\; 1 \;-\; \lambda e^{-\mu x} \int_0^{u-x} e^{-\lambda y} I_0(2\sqrt{\lambda\mu x y})dy \quad , \tag{6.7}$$

for every x, $0 \le x \le u$. *In particular* $P[{}^iT_a(u) \le u] \;=\; 1$ *and* $P[{}^iT_a(u) = 0] \;=\; e^{-\lambda u}$.

The distribution of the total time that the source spends in the active state in the time interval [0,u], given that the state at time 0 was the active state, is given by

$$P[{}^aT_a(u) < x] \;=\; \mu e^{-\lambda(u-x)} \int_{[0,x)} e^{-\mu y} I_0(2\sqrt{\lambda\mu(u-x)y})dy \quad . \tag{6.8}$$

In particular $P[{}^aT_a(u) = u] \;=\; e^{-\mu u}$ *and* $P[{}^aT_a(u) = 0] \;=\; 0$.

Proof: With $F_i(u) \;=\; 1 \;-\; e^{-\lambda u}$ and $F_a(u) \;=\; 1 - e^{-\mu u}$ for $u \ge 0$, (6.1) becomes:

$$\begin{aligned}
P[{}^iT_a(u) \le x] \;&=\; \sum_{k=0}^{\infty} e^{-\lambda(u-x)} \frac{(\lambda(u-x))^k}{k!} \sum_{n=k}^{\infty} e^{-\mu x} \frac{(\mu x)^n}{n!} \\
&= \sum_{n=0}^{\infty} e^{-\mu x} \frac{(\mu x)^n}{n!} \sum_{k=0}^{n} e^{-\lambda(u-x)} \frac{(\lambda(u-x))^k}{k!} \\
&= \sum_{n=0}^{\infty} e^{-\mu x} \frac{(\mu x)^n}{n!} \{1 \;-\; \sum_{k=n+1}^{\infty} e^{-\lambda(u-x)} \frac{(\lambda(u-x))^k}{k!}\} \\
&= 1 \;-\; \sum_{n=0}^{\infty} e^{-\mu x} \frac{(\mu x)^n}{n!} \int_0^{u-x} \lambda e^{-\lambda y} \frac{(\lambda y))^n}{n!} dy \\
&= 1 \;-\; \lambda e^{-\mu x} \int_0^{u-x} e^{-\lambda y} I_0(2\sqrt{\lambda\mu x y})dy
\end{aligned}$$

where $I_0(y) \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} \frac{(\frac{y}{2})^{2n}}{(n!)^2}$ is the modified Bessel function of zero order. Therefore,

$$P[{}^iT_a(u) \le x] \;=\; 1 \;-\; \lambda e^{-\mu x} \int_0^{u-x} e^{-\lambda y} I_0(2\sqrt{\lambda\mu x y})dy \quad , \tag{6.9}$$

for every $x \in [0, u)$. Similarly (6.1) becomes

$$P[{}^aT_a(u) < x] \;=\; 1 \;-\; \sum_{k=0}^{\infty} e^{-\mu x} \frac{(\mu x)^k}{k!} \sum_{n=k}^{\infty} e^{-\lambda(u-x)} \frac{(\lambda(u-x))^n}{n!} \quad . \tag{6.10}$$

After some algebraic manipulation, the previous relation becomes

$$P[{}^aT_a(u) < x] \;=\; \mu e^{-\lambda(u-x)} \int_{[0,x)} e^{-\mu y} I_0(2\sqrt{\lambda\mu(u-x)y})dy \tag{6.11}$$

for every x, $x \in [0, u)$.

∎

Two special cases of a two-state Markov process traffic source are now further analyzed. The first case is referred to as the deterministic case, with the second case referred to as the interrupted Poisson case.

A Two-State Markov Process: The Deterministic Case

The information generated by a traffic source during an active period is packetized into cells. Each cell contains the information generated during the previous $1/c$ units of activity period, with one exception: immediately before a renewal epoch, any information generated since the last cell was formed is packetized into a cell. The following equations represent the packetization process:

$$^{\sigma}\alpha_k(u) \;=\; P[\frac{k}{c} \leq {}^{\sigma}T_a(u) < \frac{k+1}{c}]$$

and

$$^{\sigma}\alpha_k \;=\; P[(\frac{k-1}{c})^{+} < {}^{\sigma}T_a(D) \leq \frac{k}{c}] \quad ,$$

for $0 \leq u < D$, $k \geq 0$ and $\sigma \in \{i, a\}$. These expressions can be computed using (6.7) and (6.8). Notice that in every interval between two successive renewal epochs, at most $\lceil cD \rceil$ cells can be generated.

If $p_{\sigma\hat{\sigma}}(n)$ is the probability that the source is in state $\hat{\sigma}$ at the beginning of the n^{th} slot given that at the preceding renewal its state was σ, for $\sigma, \hat{\sigma} \in \{i, a\}$, then

$$p_{ii}(n) \;=\; \frac{\mu}{\mu+\lambda} \;+\; \frac{\lambda}{\mu+\lambda}e^{-(\lambda+\mu)n} \quad ,$$

$$p_{ai}(n) \;=\; \frac{\mu}{\mu+\lambda}(1 \;-\; e^{-(\lambda+\mu)n}) \quad ,$$

$$p_{aa}(n) \;=\; \frac{\lambda}{\mu+\lambda} \;+\; \frac{\mu}{\mu+\lambda}e^{-(\lambda+\mu)n} \quad ,$$

and

$$p_{ia}(n) \;=\; \frac{\lambda}{\mu+\lambda}(1 \;-\; e^{-(\lambda+\mu)n}) \quad ,$$

for $0 \leq n \leq D-1$.

To illustrate how the framework developed in Sections 3,4, and 5 is applicable for the deterministic case, the source parameters utilized in Sections 3,4, and 5 are evaluated for this particular traffic source model. In the notation introduced previously, $S \;=\; 2$, $\sigma_0 = i$, $\sigma_1 = a$, $p_{ii} \;=\; p_{ii}(D)$, $p_{ai} \;=\; p_{ai}(D)$, $p_{aa} \;=\; p_{aa}(D)$, and $p_{ia} \;=\; p_{ia}(D)$.

$$^{\sigma}\alpha_k(u) \;=\; P[\frac{k}{c} \leq {}^{\sigma}T_a(u) < \frac{k+1}{c}] \quad ,$$

for $\sigma \in \{i, a\}$, $k \geq 0$, and $0 \leq u < D$. Furthermore,

$$^{\sigma}\alpha_k \;=\; P[(\frac{k-1}{c})^{+} < {}^{\sigma}T_a(D) \leq \frac{k}{c}] \quad ,$$

for $\sigma \in \{i, a\}$ and $k \geq 0$. Finally,

$$^{\sigma}\alpha_k^n(u) \;=\; P[(\frac{k-1}{c})^{+} < {}^{\sigma}T_a(n+u) - {}^{\sigma}T_a(n) \leq \frac{k}{c}]$$

$$= \int_{\tau\in(0,1]} P[(\frac{k-1}{c})^{+} < {}^{\sigma}T_a(n+u) - {}^{\sigma}T_a(n) \leq \frac{k}{c} | {}^{\sigma}T_a(n) \;=\; \tau] dP[{}^{\sigma}T_a(n) \leq \tau] \quad ,$$

for $\sigma \in \{i, a\}$, $0 \leq u \leq 1$, $k \geq 0$, and $0 \leq n \leq D-1$.

Because of the memoryless property the last equation can be simplified as follows:

$$^{\sigma}\alpha_k^n(u) \;=\; p_{\sigma i}(n)\,{}^{i}\alpha_k(u) \;+\; p_{\sigma a}(n)\,{}^{a}\alpha_k(u) \quad ,$$

for $\sigma \in \{i, a\}$, $k \geq 0$, $0 \leq n \leq D-1$, and $0 \leq u \leq 1$.

A Two-State Markov Process: The Interrupted Poisson Case

With the interrupted Poisson case, during the active state, cells are generated according to a Poisson distribution with rate c. Using the notation utilized in Sections 3,4, and 5, $S = 2$, $\sigma_0 = i$, $\sigma_1 = a$, $p_{ii} = p_{ii}(D), p_{ai} = p_{ai}(D), p_{aa} = p_{aa}(D)$, and $p_{ia} = p_{ia}(D)$.

$$^{\sigma}\alpha_k(u) = \int_{\tau\in(0,u]} e^{-c\tau}\frac{(c\tau)^k}{k!} dP[^{\sigma}T_a(u) \le \tau] \quad ,$$

for $\sigma \in \{i, a\}$, $u \ge 0$, $0 \le n \le D-1$, and $k \ge 0$. Finally,

$$^{\sigma}\alpha_k^n(u) \;=\; \int_{\tau\in(0,1]} e^{-c\tau}\frac{(c\tau)^k}{k!} dP[0 < {}^{\sigma}T_a(n+u) - {}^{\sigma}T_a(n) \le \tau] \quad ,$$

for $\sigma \in \{i, a\}$, $0 \le u \le 1$, and for $0 \le n \le D-1$.

Because of the memoryless property the last equation can be simplified as follows:

$$^{\sigma}\alpha_k^n(u) \;=\; p_{\sigma i}(n)\,{}^{i}\alpha_k(u) \;+\; p_{\sigma a}(n)\,{}^{a}\alpha_k(u) \quad ,$$

for $\sigma \in \{i, a\}$, $k \ge 0$, $0 \le n \le D-1$, and $0 \le u \le 1$.

6.1.3. An Unobserved Multi-State Markov Process Traffic Source

A multi-state traffic source that is not observed at the renewal epochs can be described as a single state traffic source with $S = 1$, $p_{00} = 1$ and $p_0 = 1$. Furthermore, if $T_a(u)$ be the total time that the source spends in the active state in $[0, u)$, the probability that k cells are generated by the traffic source during the first u units of time following the last renewal epoch, is given by

$$\alpha_k(u) \;=\; P[A(u) = k] \;=\; \sum_{\forall\sigma} P[^{\sigma}A(u) = k]p_{\sigma} \quad .$$

Finally,

$$\alpha_k^n \;=\; \sum_{\forall\sigma} {}^{\sigma}\alpha_k^n p_{\sigma} \quad .$$

for $0 \le n \le D-1$, and $k \ge 0$.

As an example, if the source is an unobserved two-state Markov process traffic source, then

$$P[T_a(u) \le x] \;=\; \frac{\lambda}{\mu+\lambda}P[^{a}T_a(u) \le x] \;+\; \frac{\mu}{\mu+\lambda}P[^{i}T_a(u) \le x] \quad . \tag{6.12}$$

Therefore for $k \ge 0$, the unobserved traffic behavior is described as a single state traffic source with

$$\alpha_k(u) \;=\; \frac{\lambda}{\mu+\lambda}{}^{a}\alpha_k(u) \;+\; \frac{\mu}{\mu+\lambda}{}^{i}\alpha_k(u) \quad ,$$

for $0 \le u \le D$, and

$$\alpha_k^n(u) \;=\; \alpha_k(u) \quad ,$$

for $0 \le u \le 1$, and for $0 \le n \le D-1$.

6.1.4. A Poisson Process Traffic Source

A traffic source following a Poisson distribution with rate λ can be described as follows: $S = 1$, $p_{00} = 1$ and $p_0 = 1$.

$$\alpha_k(u) \;=\; e^{-\lambda u}\frac{(\lambda u)^k}{k!} \quad ,$$

for $u \geq 0$, $0 \leq n \leq D-1$, and $k \geq 0$. Finally,

$$\alpha_k^n(u) \;=\; \alpha_k(u) \quad ,$$

for $0 \leq u \leq 1$, and for $0 \leq n \leq D-1$.

6.2. Discrete Time Source Models

6.2.1. Multiplexing of N Independent and Identically Distributed Two-state Bursty Sources

A bursty source is modelled as a two-state Markov chain. The Markov chain makes a transition every $1/c$ units of time. The idle and active states are represented by 0 and 1 respectively. When the Markov chain enters state 1, a cell is generated with probability p_{gen}. While in state 0, no cells are generated. Let q_{01} be the probability that the source is currently in the idle state and that it will be in state 1 immediately after the transition. The transition probabilities q_{00}, q_{10}, and q_{11} are defined similarly. Notice that $q_{00} + q_{01} = 1$ and $q_{10} + q_{11} = 1$. Let

$$\mathbf{Q} \;\stackrel{\text{def}}{=}\; \begin{pmatrix} q_{00} & q_{01} \\ q_{10} & q_{11} \end{pmatrix} \quad .$$

Let ${}^{\sigma}q(0,n)$ (respectively ${}^{\sigma}q(1,n)$) be the probability that after n transitions, the two-state Markov chain will be in state 0 (respectively 1), given that the state of the Markov chain at the last renewal epoch was σ. Also let ${}^{\sigma}q(n) \stackrel{\text{def}}{=} [{}^{\sigma}q(0,n)\ {}^{\sigma}q(1,n)]$. If $\sigma = 0$, then ${}^{0}q(0) \stackrel{\text{def}}{=} [1\ 0]$, and if $\sigma = 1$, then ${}^{1}q(0) \stackrel{\text{def}}{=} [0\ 1]$. ${}^{\sigma}q(n) \;=\; {}^{\sigma}q(0)\mathbf{Q}^{\mathbf{n}}$, for $n \geq 0$ and $\sigma \in \{0,1\}$. The probability that k cells are generated in m consecutive transitions of the Markov chain is given by:

$$\sum_{k_1=0}^{1} \cdots \sum_{k_m=0}^{1} \prod_{n=1}^{m} {}^{\sigma}q(k_n, n)p_{gen}^{k} 1(\sum_{j=1}^{m} k_j \;=\; k) \quad .$$

For this type of source model, $S = 2$, $\sigma_0 = 0$, $\sigma_1 = 1$, $p_{00} \;=\; {}^{0}q(0,D)$, $p_{01} \;=\; {}^{0}q(1,D)$, $p_{10} \;=\; {}^{1}q(0,D)$, and $p_{11} \;=\; {}^{1}q(1,D)$.

$${}^{\sigma}\alpha_k(u) \;=\; \begin{cases} \sum_{k_1=0}^{1} \cdots \sum_{k_{\lfloor uc \rfloor}=0}^{1} \prod_{n=1}^{\lfloor uc \rfloor} {}^{\sigma}q(k_n,n)p_{gen}^{k} 1(\sum_{j=1}^{\lfloor uc \rfloor} k_j \;=\; k) & \text{if } 1 \leq \lfloor uc \rfloor \\ 1 & \text{if } k = 0, \text{ and } 1 > \lfloor uc \rfloor \\ 0 & \text{otherwise} \end{cases} \quad ,$$

for $\sigma \in \{0,1\}$, and $0 \leq u \leq D$, and $k \geq 0$. Finally,

$${}^{\sigma}\alpha_k^n(u) \;=\; {}^{\sigma}q(0, \lfloor nc \rfloor)\,{}^{0}\alpha_k(\lfloor (n+u)c \rfloor - \lfloor nc \rfloor) \;+\; {}^{\sigma}q(1, \lfloor nc \rfloor)\,{}^{1}\alpha_k(\lfloor (n+u)c \rfloor - \lfloor nc \rfloor) \quad ,$$

for $\sigma \in \{0,1\}$, $k \geq 0$, $0 \leq n \leq D-1$, and $0 \leq u \leq 1$.

When N of the above sources are multiplexed, the state of the multiplexed sources is the number of active sources.

$$ {}^{\sigma}\alpha_k^n(u) = \sum_{k_0+\cdots k_{N-1}=k} {}^{\sigma_0}\alpha_{k_0}^{n,0}(u) \otimes \cdots \otimes {}^{\sigma_{N-1}}\alpha_{k_{N-1}}^{n,N-1}(u) 1(\sum_{j=0}^{N-1} \sigma_j = \sigma) \quad , $$

for all $k \geq 0, k_0 \geq 0, \cdots k_{N-1} \geq 0$, $0 \leq n \leq D-1$, and $0 \leq u \leq 1$. If $\hat{p}_{\sigma_1\sigma_2}$ is the probability that the state of the multiplexed sources makes a transition from state σ_1 to σ_2,

$$ \hat{p}_{\sigma_1\sigma_2} = \sum_{k_1=0}^{\sigma_1} \sum_{k_2=0}^{N-1-\sigma_1} \binom{N}{\sigma_1-k_1, k_1, N-\sigma_1-k_2, k_2} p_{11}^{k_1} p_{10}^{\sigma_1-k_1} p_{01}^{k_2} p_{00}^{N-\sigma_1-k_2} 1(k_1+k_2=\sigma_2) \quad . $$

for all $\sigma_1, \sigma_2 \in \{0, 1, \cdots, N-1\}$.

6.2.2. Binomial Traffic Source Model

The binomial source model is a special case of the two-state Markov chain source model. In particular, if a cell is generated every $1/c$ units of time with probability p, then $p_{gen} = 1$, and

$$ \mathbf{Q} = \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix} \quad . $$

In addition ${}^{\sigma}q(1,n) = p$, and ${}^{\sigma}q(0,n) = 1-p$, for every $n \geq 0$ and $\sigma \in \{i, a\}$.

The binomial traffic source model can also be seen as a traffic source with $S = 1$, $p_{00} = 1$ and $p_0 = 1$.

$$ \alpha_k(u) = \binom{\lfloor uc \rfloor}{k} p^k (1-p)^{\lfloor uc \rfloor - k} \quad , $$

for $0 \leq u \leq D$, and $k \geq 0$. Finally,

$$ \alpha_k^n(u) = \binom{\lfloor (n+u)c \rfloor - \lfloor nc \rfloor}{k} p^k (1-p)^{\lfloor (n+u)c \rfloor - \lfloor nc \rfloor - k} \quad , $$

for $0 \leq n \leq D-1$, $0 < u \leq 1$, and $k \geq 0$.

6.2.3. The Constant Bit Rate Traffic (CBRT) Source

The CBRT source is another example of a single state traffic source. One cell is generated every $1/c$ units of time. Let D be an integer multiplier of $1/c$. For this type of traffic source, $S = 1$, $p_{00} = 1$ and $p_0 = 1$.

$$ \alpha_k(u) = P[\frac{k}{c} \leq u < \frac{k+1}{c}] = P[\lfloor uc \rfloor = k] = 1(\lfloor uc \rfloor = k) \quad , $$

for $0 \leq u \leq D$, and $k \geq 0$. Finally,

$$ \alpha_k^n(u) = P[\lfloor (n+u)c \rfloor - \lfloor nc \rfloor = k] = 1(\lfloor (n+u)c \rfloor - \lfloor nc \rfloor = k) \quad , $$

for $0 \leq n \leq D-1$, $0 < u \leq 1$, and $k \geq 0$.

Notice that this last source model can be seen as a special case of a binomial source model for $p = 1$.

7. Examples, Discussion, and Direction of Future Research

7.1. Examples

In this section, a number of examples illustrating the results derived in this paper are presented. In Section 6, some examples of discrete and continuous time traffic source models analyzed in this paper were presented. In this section, results corresponding to four of these traffic source models are presented. The first traffic source model used in the examples is the Poisson source presented in Subsection 6.1.4. The second source model is the CBRT source presented in Subsection 6.2.3. These two source models have been selected for further study because they represent widely used non-bursty source models. The third source model utilized is the deterministic two-state continuous time Markov process appearing in Subsection 6.1.2. The fourth source is the discrete time Markov chain process presented in Subsection 6.2.1. The third and fourth models are traffic generators used to model bursty sources. For each figure, a complete specification of the TCM and incoming CS parameters is given. In addition the token generation pattern is specified by the notation $< s_0, s_1, s_2, \cdots, s_{T-1} >$.

One important aspect of this paper is the detailed study of the distribution of the activity period of a two-state continuous time Markov process. In Figure 3, for every x, $0 \leq x \leq u$, $P[{}^iT_a(u) \leq x]$ is illustrated by the dotted line, and $P[{}^aT_a(u) < x]$ is illustrated by the solid line. An important characteristic of the curves shown in Figure 3 is the linear behavior of both plotted distributions for the cases in which the mean active time plus the mean idle time is comparable to or exceeds u.

In Figures 4a through 4g, the effect of the cell buffer size on the performance of a two state discrete Markov chain bursty source is studied. Detailed results about the throughput, the expected loss, the cell loss probability, the cell loss distribution in a slot's time interval, the occupancy distribution of the cell and token buffers, the utilization of the generated tokens, and the waiting time distribution of the cells are presented. Notice that the size of the token buffer is ten. Furthermore at each renewal epoch, the number of tokens generated alternates between twenty and ten. When twenty tokens are generated, if the token buffer is empty and the cell buffer is not empty, nine tokens are lost, one token is used for the service of a cell in the cell buffer, and another ten tokens are used to fill the token buffer. When ten tokens are generated, if the token buffer is empty and the cell buffer is not empty, no tokens are lost. These observations explain the shape of Figure 4f. From Figures 4e and 4g, notice that while increasing the cell buffer size beyond twenty has very little effect on the cell loss probability, it results in a considerable increase in the cell waiting time.

In Figure 4h, the effectiveness of the TCM is evaluated for sources having long idle and active periods. In the five curves appearing in Figure 4h, the load is kept constant and equal to .375 cells/slot. At the same time, the duration of the idle and active periods is progressively increased. In particular the mean active period is increased from 2 slots to 4 slots, 8 slots, 16 slots, and 160 slots. Notice that the cell buffer size increase has a diminishable effect on the achieved cell loss probability. The above results demonstrate that for sources with long idle and active periods, the TCM, in order to be effective, must generate tokens with the rate at which the source generates cells during an activity period. If the TCM generate tokens at such a rate, however, the effect is a dramatic increase in wasted bandwidth (*i.e.* tokens) during the inactive periods. Therefore, for sources with long idle and active periods, an alternative type of control capable of ceasing the generation of tokens while the source is inactive and resuming the generation of tokens when the source is reactivated is desirable. A modification of the TCM capable of handling such traffic behaviors will be presented and analyzed in a subsequent paper.

In Figure 5, similar results are presented but for the case in which the source follows the Poisson distribution. Of particular interest is the linear shape of the cell loss distribution curve in Figure 5a.

In Figures 6 and 7, the curves are derived under the assumption that the service time of a cell is zero. The results presented in Figures 6 and 7 differ substantially from those presented in Figures 4 and 5. This comparison demonstrates that the cell level behavior cannot be accurately predicted under the widely used assumption of a zero cell service time [1,2,11].

In Figure 8, the effect of the cell buffer size on the waiting time distribution of a CBRT source is studied. In this example no cell loss is observed.

Notice that a multiplexer is a special case of a TCM module in which the token buffer is eliminated and one token is generated at the beginning of each slot, *i.e.* $M = 0$ and $s_l^n = 1$, for all values of n and l. If a token is not used immediately at the time of its generation it is lost. The remaining examples illustrate the application of the results derived in this paper on the problem of multiplexing a number of independent and identically distributed (i.i.d.) bursty sources. A bursty source is modeled by the two-state discrete Markov chain model introduced in Section 6.2.1. Let

$$Activity = Prob[\text{Source is active}] = \frac{p_{01}}{p_{01} + p_{10}} \quad .$$

Further, let

$$Burstiness = \frac{\text{Maximum Source Rate}}{\text{Mean Source Rate}} = \frac{\frac{1}{c}}{\frac{1}{c} \times P_{gen} \times Prob[\text{source is active}]} \quad .$$

Therefore,

$$Burstiness = \frac{1}{Activity \times P_{gen}} \quad .$$

The mean burst length is $\frac{1}{p_{10}}$ cells, and the mean silence length is $\frac{1}{p_{01}}$. Further, notice that

$$\frac{1}{c} = \frac{\text{Output Link Rate}}{\text{Maximum Source Rate}} \quad .$$

In Figure 9, the buffering required for the multiplexing of a number of i.i.d bursty sources under a specific cell loss probability GoS is computed. As expected, the increased number of multiplexed sources leads to a linear increase in the traffic activity of the multiplexer, which in turn results in a nonlinear increase in the buffering required if the same cell loss probability is to be maintained.

In Figure 10, it is demonstrated that the output rate of the multiplexer divided by the peak input traffic rate is a fundamental quantity in the characterization of the cell loss probability. Observe that for buffer sizes less than this ratio, each source has at most one cell in the buffer, whereas for buffer sizes greater than this ratio, each source may have multiple cells in the buffer. The shape of the cell loss probability versus the buffer size raises the question of whether buffering can be considered an effective method for the solution of the congestion problem. It appears that buffering should be used only for a temporary surge of traffic.

7.2. Discussion

The results presented in this paper are complete in the sense that all quantities of interest can be derived for any source which can be described by the traffic model introduced in Section 3. Further, all quantities of interest can be derived by simple additions and multiplications. Current experience suggests that the computational procedures introduced in this paper are stable and efficient. Examples with 60,000 states have been evaluated without difficulty.

We are currently trying to identify invariant parameters which would uniquely describe the relationship between the source behavior, the TCM bandwidth and buffer parameters, and the provided GOS without having to run the algorithms presented in this paper. Such results would be necessary in order to make a real time evaluation of the performance that could be guaranteed to a particular incoming CS by a particular TCM and to make a real time calculation of the TCM buffer and bandwidth requirements needed in order to achieve a particular GOS for a particular incoming CS.

We are also utilizing the analysis presented in this paper in the study of the behavior of a TCM when the incoming CS is composed of a number of multiplexed, independent, not identically distributed bursty sources.

One of the limitations of the current version of the TCM is its inability to handle traffic sources with long active and idle periods. For such sources, enhancement of the TCM control capability is necessary in order to do resource reservation not only at the cell level but at the burst level as well. A desirable TCM enhancement currently being developed is the capability of deallocating buffer space and bandwidth (i.e. tokens) from idle sources and reallocating this buffer space and bandwidth to sources currently in need of resources. This approach aims at guaranteeing QoS at the cell level of bursts in progress. The QoS at the burst level is guaranteed through call and/or burst admission control.

8. Conclusions

For the commercialization of ATM technology, the solution of the bandwidth allocation and traffic control problems is critical. Such problems will in all likelihood defy solution, however, as long as they are studied in isolation without a unifying traffic characterization and traffic control infrastructure for ATM networks. The work presented in this paper is part of an ongoing effort directed at the development of such an integrated traffic characterization and control infrastructure.

In this paper, a traffic control mechanism (TCM) capable of monitoring and controlling a rich family of traffic behaviors is introduced. Conceived as part of a network-wide resource allocation and control scheme being developed for an ATM environment, the TCM has characteristics that enable its use for both global and local resource allocation. In particular, the periodic generation of tokens makes the TCM a prime candidate for a traffic control involving the periodic exchange of information between loosely coupled systems.

In addition to the description of the TCM, an exhaustive performance analysis of the TCM is presented, and a number of new results concerning the behavior of a bursty source are derived. The performance methodology presented has a number of desirable features: (i) it is exhaustive, (ii) it is computationally efficient, (iii) it is performed exclusively in the time domain, and (iv) it introduces and utilizes a decoupling between traffic source behavior and the TCM.

Acknowledgments

Many of the ideas as well as the structure of this paper were enhanced through numerous discussions with my wife Patti. A substantial portion of the program used to evaluate the TCMZ performance was written by Einir Valdimarsson. Their contributions are greatly appreciated.

9. References

[1] Berger, A. W.,"Performance Analysis of a Rate Control Throttle Where Tokens and Jobs Queue," *Proceedings of INFOCOM'90,* San Francisco, California, June 5-7, 1990, pp. 30-38.

[2] Berger, A. W.,"Overload Control Using a Rate Control Throttle: Selecting Token Bank Capacity for Robustness to Arrival Rates," *IEEE Transactions on Automatic Control,* Fall, 1990.

[3] Boyer, P., "A Congestion Control for the ATM," Proceedings of the 7th ITC Seminar, New Jersey, October 1990, paper 4.3.

[4] Butto, M., Cavallero, E., and Tonietti, A., "Effectiveness of the "Leaky Bucket" Policing Mechanism in ATM Networks," *IEEE Journal on Selected Areas in Communications,* Vol. SAC-9, No. 3, April 1991, pp. 355-342.

[5] Doshi, B.T., Johri, P.K., Netravali, A.N. and Sabnani, K.K., "Error and Flow Control Performance of a High Speed Protocol," preprint, 1991.

[6] Dittmann, L., Jacobsen, S., and Moth, K., "Flow Enforcement Algorithms for ATM Networks," *IEEE Journal on Selected Areas in Communications,* Vol. SAC-9, No. 3, April 1991, pp. 343-350.

[7] Eckberg, A.E., Luan, D.T., and Lucantoni, D. M., "Meeting the Challenge: Congestion and Flow Control Strategies for Broadband Information Transport," *Proceedings of IEEE GLOBECOM'89*, November 27-30, 1989.

[8] Hui, J, "Resource Allocation for Broadband Networks," *IEEE Journal on Selected Areas in Communications,* Vol. SAC-6, No. 9, December 1988.

[9] Ohnishi, H., Okada, T. and Noguchi, K., "Flow Control Schemes and Delay Loss Tradeoff in ATM Networks," *IEEE Journal on Selected Areas in Communications,* Vol. SAC-6, No. 9, December 1988.

[10] Rathgeb, E. P., "Modeling and Performance Comparison of Policing Mechanisms for ATM Networks," *IEEE Journal on Selected Areas in Communications,* Vol. SAC-9, No. 3, April 1991, pp. 325-334.

[11] Sidi, M., Liu, W. Z., Cidon, I., and Gopal, I., "Congestion Control Through Input Rate Regulation," *Proceedings of GLOBECOM'89,* Dallas, Texas, November 27-30, 1989, pp. 1764-1768.

[12] Sohraby, K, and Sidi, M., "On the Performance of Bursty and Correlated Sources Subject to Leaky Bucket Rate-Based Access Control Schemes" *Proceedings of IEEE INFOCOM'91,* April 9-11, 1991, pp. 426-434.

[13] Turner, J. S., "New Directions in Communications (or Which Way to the Information Age?)," *IEEE Communications Magazine,* Vol. 24, No. 10, October 1986, pp. 8-15.

[14] Vakil, F. and Singh, R. P., "Shutter: A Flow Control Scheme for ATM Networks," Proceedings of the 7th ITC Seminar, New Jersey, October 1990, paper 4.5.

[15] Zhang, L., "VirtualClock: A New Traffic Control Algorithm for Packet Switching Networks," *Proceedings of SIGCOM'90 Symposium,* Philadelphia, Pennsylvania, September 24-27, 1990, pp. 19-29.

10. Appendix I

Iterative Algorithm for the Computation of the Equilibrium Probabilities of the TCMZ

```
for all l, 0 ≤ l ≤ T − 1
for all i, 0 ≤ i ≤ S − 1
for all j, 0 ≤ j ≤ K + M
    σi_l π̂_j = 1/(T*S*(K+M+1));

do
{
    for all l, 0 ≤ l ≤ T − 1
    for all i, 0 ≤ i ≤ S − 1
        σi_l π = σi_l π̂;
    for all i, 0 ≤ i ≤ S − 1
    for all l, 0 ≤ l ≤ T − 1
    {
        l2 = (l + 1) mod T;
        σi_l2 π̂ := Σ_{j=0}^{S−1} σj_l π σj_l P p_{σj σi};
    }
} while (Σ_{i=0}^{S−1} Σ_{l=0}^{T−1} |σi_l π̂ − σi_l π|e > ε);

for all l, 0 ≤ l ≤ T − 1
for all i, 0 ≤ i ≤ S − 1
    σi_l π = σi_l π̂;
```

∎

11. Appendix II

Iterative Algorithm for the Computation of the Equilibrium Probabilities of the TCMP

for all l, $0 \le l \le T-1$
for all i, $0 \le i \le S-1$
for all n, $0 \le n \le D-1$
for all j_c, $0 \le j_c \le K$
for all j_t, $0 \le j_t \le M$
  ${}^{\sigma_i}_{l}\widehat{\pi}^{(n)}_{j_c,j_t} = \frac{1}{T*S*D*(K+1)*(M+1)}$;

do
{
  for all l, $0 \le l \le T-1$
  for all i, $0 \le i \le S-1$
  for all n, $0 \le n \le D-1$
    ${}^{\sigma_i}_{l}\pi^{(n)} = {}^{\sigma_i}_{l}\widehat{\pi}^{(n)}$;
  if (D == 1)
  {
    for all l, $0 \le l \le T-1$
    for all i, $0 \le i \le S-1$
    {
      $l_2 = (l+1) \ mod \ T$;
      ${}^{\sigma_i}_{l_2}\widehat{\pi}^{(0)} \ := \ \sum_{j=0}^{S-1} {}^{\sigma_j}_{l}\pi^{(0)}\,{}^{\sigma_j}_{l}\mathbf{P}_0 p_{\sigma_j\sigma_i}$;
    }
  }
  else if (D > 1)
  {
    for all l, $0 \le l \le T-1$
    for all i, $0 \le i \le S-1$
    {
      ${}^{\sigma_i}_{l}\widehat{\pi}^{(1)} \ := \ {}^{\sigma_i}_{l}\pi^{(0)}\,{}^{\sigma_i}_{l}\mathbf{P}_0$;
      for all n, $2 \le n \le D-2$
        ${}^{\sigma_i}_{l}\widehat{\pi}^{(n)} \ := \ {}^{\sigma_i}_{l}\pi^{(n-1)}\,{}^{\sigma_i}_{l}\mathbf{P}_{n-1}$;
      $l_2 = (l+1) \ mod \ T$;
      ${}^{\sigma_i}_{l_2}\widehat{\pi}^{(0)} \ := \ \sum_{j=0}^{S-1} {}^{\sigma_j}_{l}\pi^{(D-1)}\,{}^{\sigma_j}_{l}\mathbf{P}_{D-1} p_{\sigma_j\sigma_i}$;
    }
  }
} while ($\sum_{n=0}^{D-1}\sum_{i=0}^{S-1}\sum_{l=0}^{T-1} |{}^{\sigma_i}_{l}\widehat{\pi}^{(n)} - {}^{\sigma_i}_{l}\pi^{(n)}|\mathrm{e} > \epsilon$);

for all l, $0 \le l \le T-1$
for all i, $0 \le i \le S-1$
for all n, $0 \le n \le D-1$
  ${}^{\sigma_i}_{l}\pi^{(n)} = {}^{\sigma_i}_{l}\widehat{\pi}^{(n)}$;

■

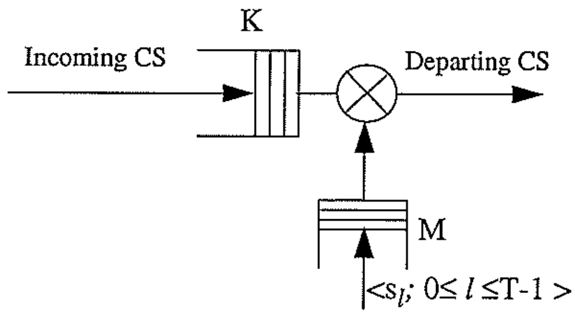
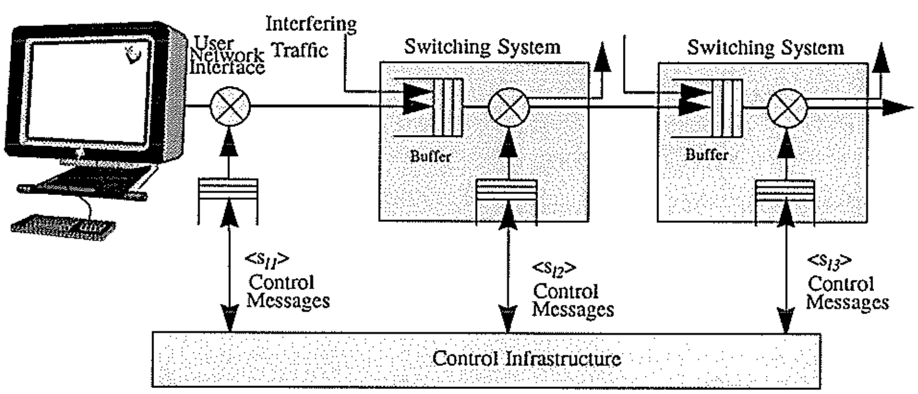


Figure 1



A simplified queueing model for ***resource allocation*** and ***traffic control.***

Figure 2

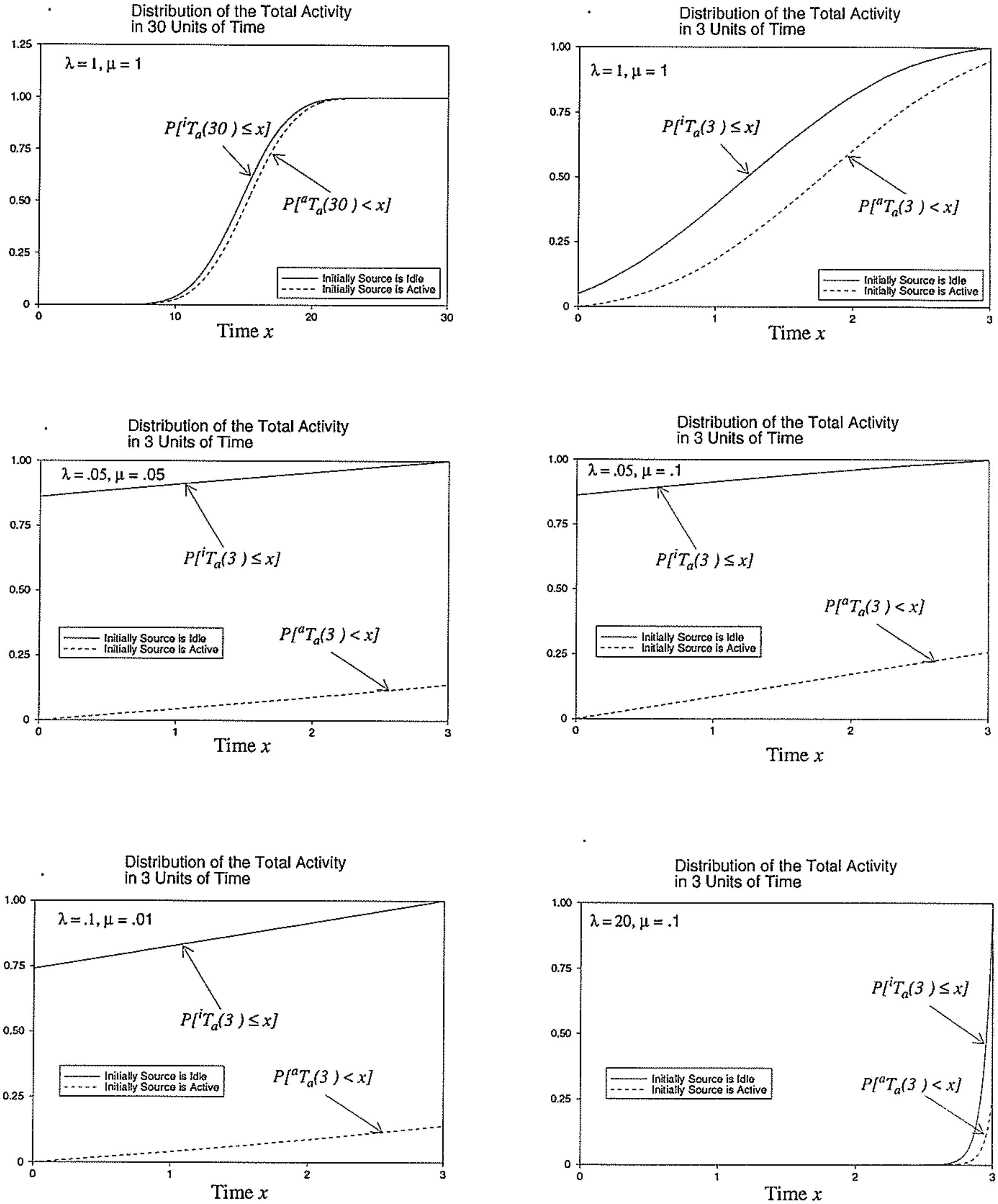

Distribution of the Total Activity
in 30 Units of Time
λ = 1, μ = 1
P[ⁱTₐ(30) ≤ x]
P[ᵃTₐ(30) < x]
Initially Source is Idle
Initially Source is Active
1.25
1.00
0.75
0.50
0.25
0
0
10
20
30
Time x
Distribution of the Total Activity
in 3 Units of Time
λ = 1, μ = 1
P[ⁱTₐ(3) ≤ x]
P[ᵃTₐ(3) < x]
Initially Source is Idle
Initially Source is Active
1.00
0.75
0.50
0.25
0
0
1
2
3
Time x
Distribution of the Total Activity
in 3 Units of Time
λ = .05, μ = .05
P[ⁱTₐ(3) ≤ x]
P[ᵃTₐ(3) < x]
Initially Source is Idle
Initially Source is Active
1.00
0.75
0.50
0.25
0
0
1
2
3
Time x
Distribution of the Total Activity
in 3 Units of Time
λ = .05, μ = .1
P[ⁱTₐ(3) ≤ x]
P[ᵃTₐ(3) < x]
Initially Source is Idle
Initially Source is Active
1.00
0.75
0.50
0.25
0
0
1
2
3
Time x
Distribution of the Total Activity
in 3 Units of Time
λ = .1, μ = .01
P[ⁱTₐ(3) ≤ x]
P[ᵃTₐ(3) < x]
Initially Source is Idle
Initially Source is Active
1.00
0.75
0.50
0.25
0
0
1
2
3
Time x
Distribution of the Total Activity
in 3 Units of Time
λ = 20, μ = .1
P[ⁱTₐ(3) ≤ x]
P[ᵃTₐ(3) < x]
Initially Source is Idle
Initially Source is Active
1.00
0.75
0.50
0.25
0
0
1
2
3
Time x

Figure 3

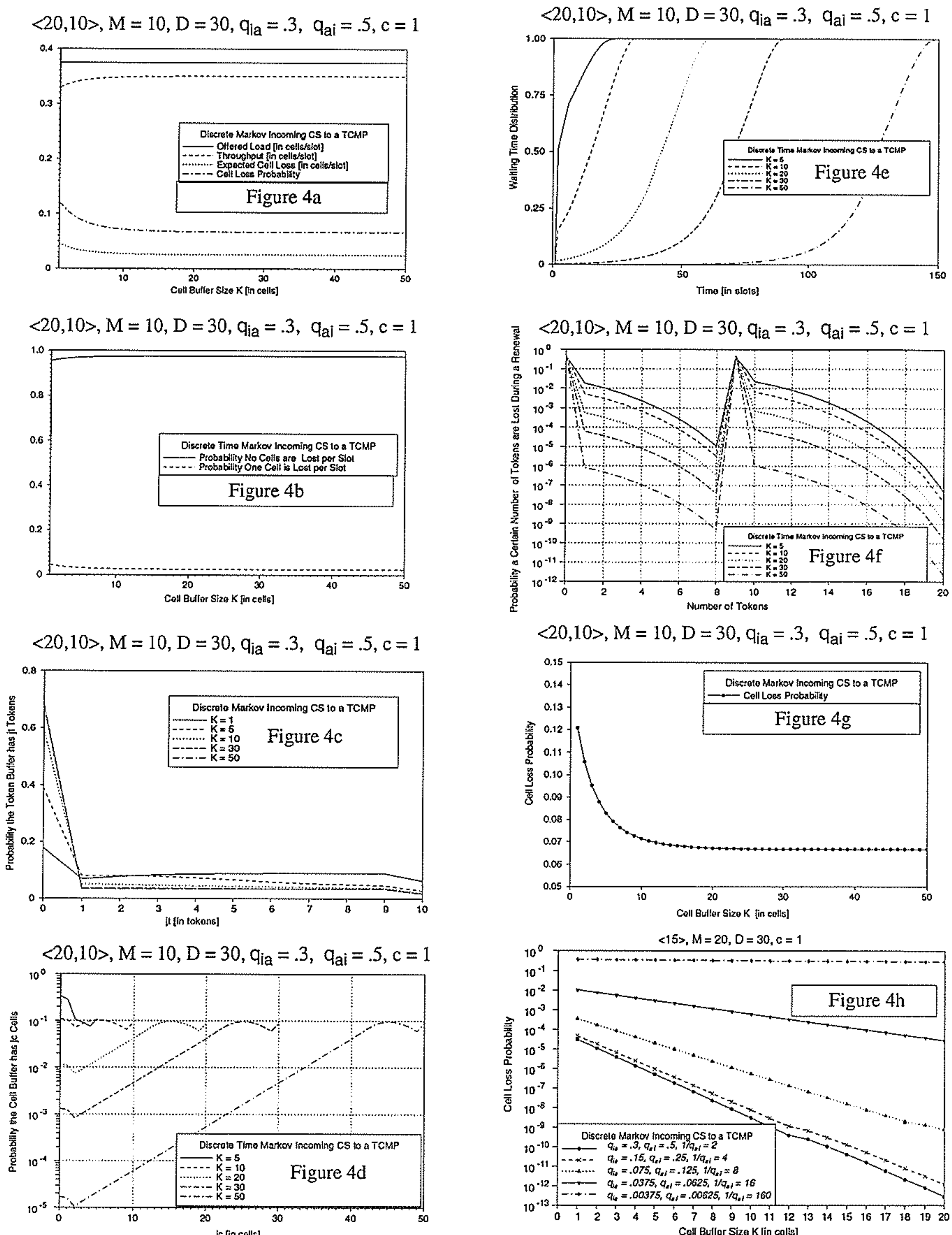

<20,10>, M = 10, D = 30, q_{ia} = .3, q_{ai} = .5, c = 1
Discrete Markov Incoming CS to a TCMP
Offered Load [in cells/slot]
Throughput [in cells/slot]
Expected Cell Loss [in cells/slot]
Cell Loss Probability
Figure 4a
Cell Buffer Size K [in cells]
<20,10>, M = 10, D = 30, q_{ia} = .3, q_{ai} = .5, c = 1
Discrete Time Markov Incoming CS to a TCMP
K = 5
K = 10
K = 20
K = 30
K = 50
Figure 4e
Waiting Time Distribution
Time [in slots]
<20,10>, M = 10, D = 30, q_{ia} = .3, q_{ai} = .5, c = 1
Discrete Time Markov Incoming CS to a TCMP
Probability No Cells are Lost per Slot
Probability One Cell is Lost per Slot
Figure 4b
Cell Buffer Size K [in cells]
<20,10>, M = 10, D = 30, q_{ia} = .3, q_{ai} = .5, c = 1
Probability a Certain Number of Tokens are Lost During a Renewal
Discrete Time Markov Incoming CS to a TCMP
K = 5
K = 10
K = 20
K = 30
K = 50
Figure 4f
Number of Tokens
<20,10>, M = 10, D = 30, q_{ia} = .3, q_{ai} = .5, c = 1
Probability the Token Buffer has jt Tokens
Discrete Markov Incoming CS to a TCMP
K = 1
K = 5
K = 10
K = 30
K = 50
Figure 4c
jt [in tokens]
<20,10>, M = 10, D = 30, q_{ia} = .3, q_{ai} = .5, c = 1
Discrete Markov Incoming CS to a TCMP
Cell Loss Probability
Figure 4g
Cell Loss Probability
Cell Buffer Size K [in cells]
<20,10>, M = 10, D = 30, q_{ia} = .3, q_{ai} = .5, c = 1
Probability the Cell Buffer has jc Cells
Discrete Time Markov Incoming CS to a TCMP
K = 5
K = 10
K = 20
K = 30
K = 50
Figure 4d
jc [in cells]
<15>, M = 20, D = 30, c = 1
Figure 4h
Cell Loss Probability
Discrete Markov Incoming CS to a TCMP
q_{ia} = .3, q_{ai} = .5, $1/q_{ai}$ = 2
q_{ia} = .15, q_{ai} = .25, $1/q_{ai}$ = 4
q_{ia} = .075, q_{ai} = .125, $1/q_{ai}$ = 8
q_{ia} = .0375, q_{ai} = .0625, $1/q_{ai}$ = 16
q_{ia} = .00375, q_{ai} = .00625, $1/q_{ai}$ = 160
Cell Buffer Size K [in cells]

Figure 4

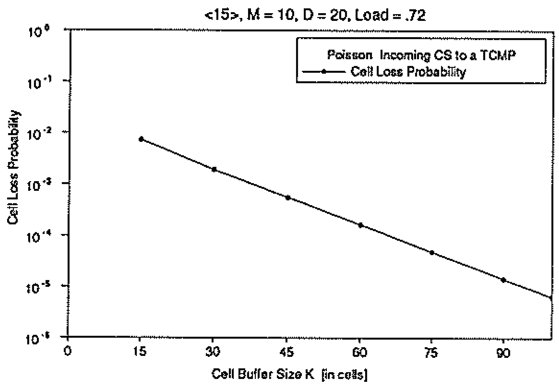


Figure 5a

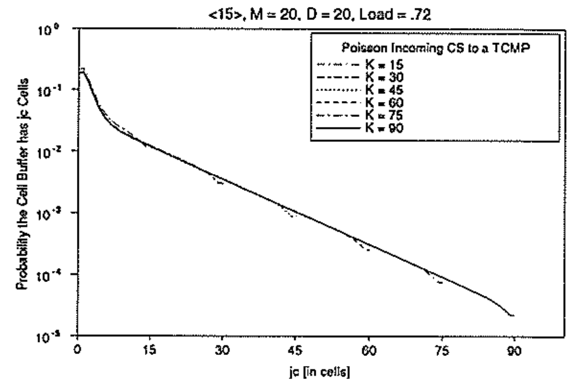


Figure 5d

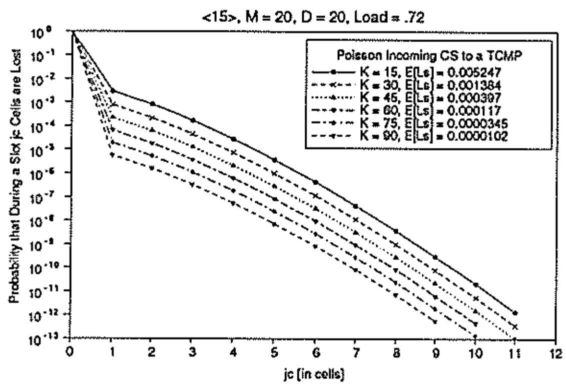


Figure 5b

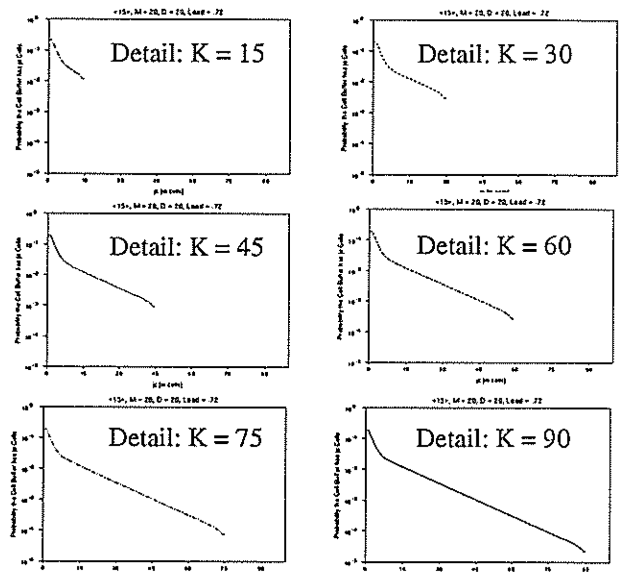


Figure 5e: A detailed representation of Figure 5d

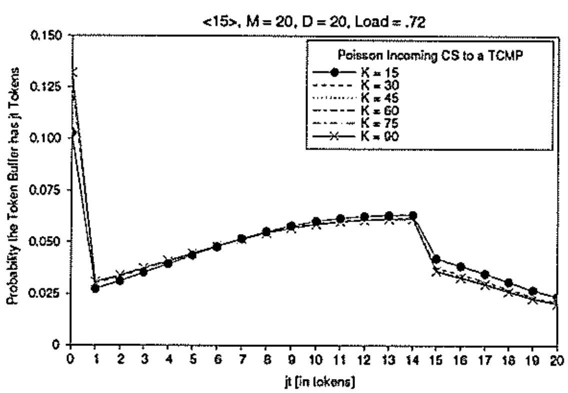


Figure 5c

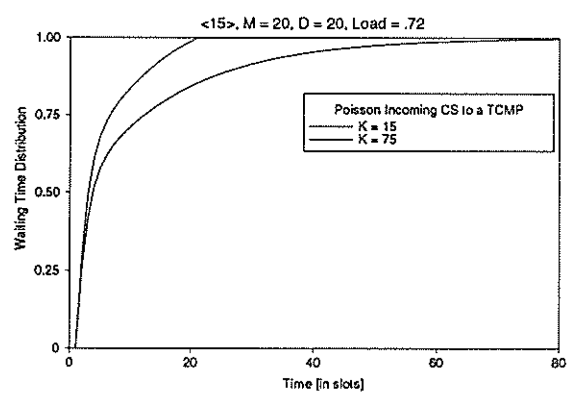


Figure 5f

Figure 5

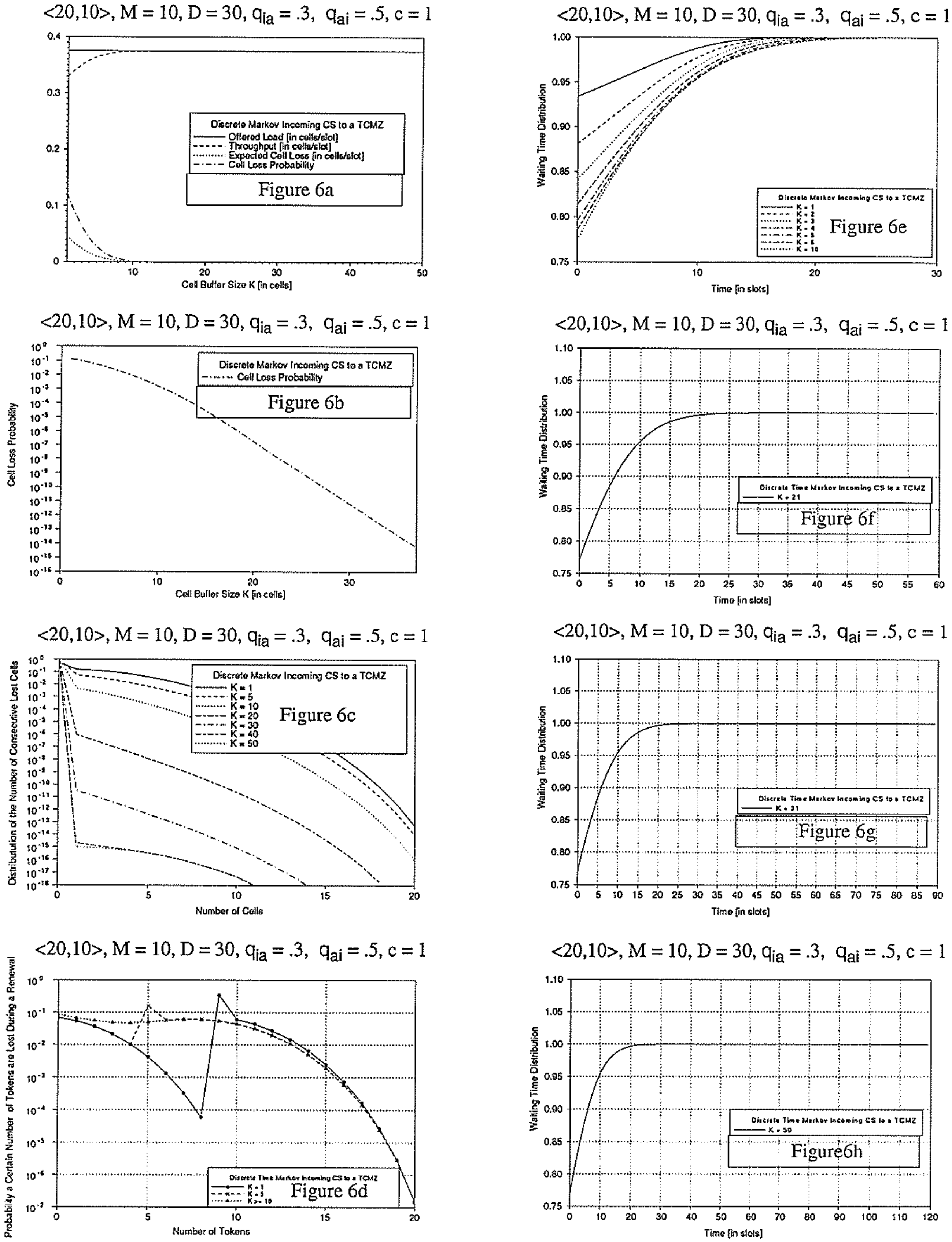

<20,10>, M = 10, D = 30, qia = .3, qai = .5, c = 1
Discrete Markov Incoming CS to a TCMZ
Offered Load [in cells/slot]
Throughput [in cells/slot]
Expected Cell Loss [in cells/slot]
Cell Loss Probability
Figure 6a
Cell Buffer Size K [in cells]
<20,10>, M = 10, D = 30, qia = .3, qai = .5, c = 1
Waiting Time Distribution
Discrete Markov Incoming CS to a TCMZ
K = 1
K = 2
K = 3
K = 4
K = 5
K = 6
K = 10
Figure 6e
Time [in slots]
<20,10>, M = 10, D = 30, qia = .3, qai = .5, c = 1
Cell Loss Probability
Discrete Markov Incoming CS to a TCMZ
Cell Loss Probability
Figure 6b
Cell Buffer Size K [in cells]
<20,10>, M = 10, D = 30, qia = .3, qai = .5, c = 1
Waiting Time Distribution
Discrete Time Markov Incoming CS to a TCMZ
K = 21
Figure 6f
Time [in slots]
<20,10>, M = 10, D = 30, qia = .3, qai = .5, c = 1
Distribution of the Number of Consecutive Lost Cells
Discrete Markov Incoming CS to a TCMZ
K = 1
K = 5
K = 10
K = 20
K = 30
K = 40
K = 50
Figure 6c
Number of Cells
<20,10>, M = 10, D = 30, qia = .3, qai = .5, c = 1
Waiting Time Distribution
Discrete Time Markov Incoming CS to a TCMZ
K = 31
Figure 6g
Time [in slots]
<20,10>, M = 10, D = 30, qia = .3, qai = .5, c = 1
Probability a Certain Number of Tokens are Lost During a Renewal
Discrete Time Markov Incoming CS to a TCMZ
K = 1
K = 5
K >= 10
Figure 6d
Number of Tokens
<20,10>, M = 10, D = 30, qia = .3, qai = .5, c = 1
Waiting Time Distribution
Discrete Time Markov Incoming CS to a TCMZ
K = 50
Figure6h
Time [in slots]

Figure 6

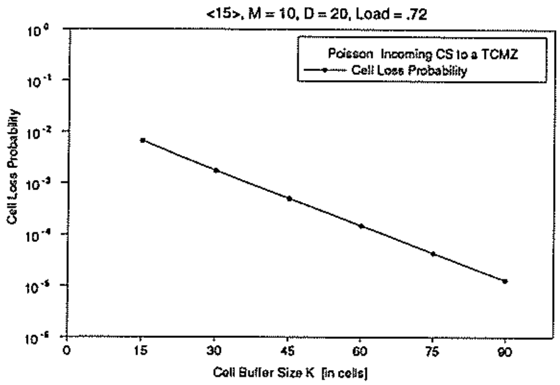


Figure 7a

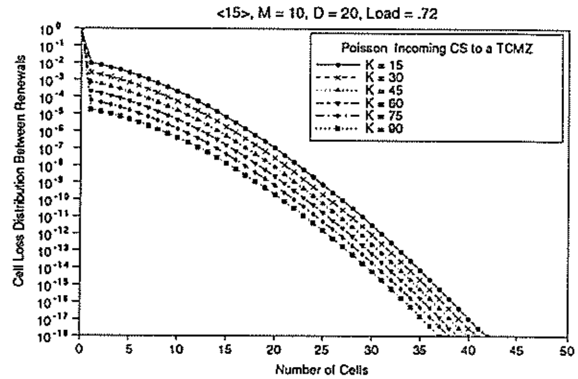


Figure 7c

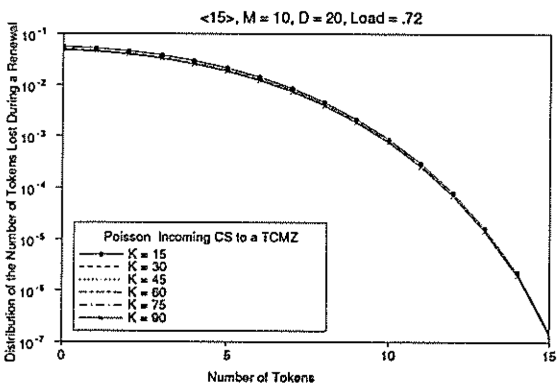


Figure 7b

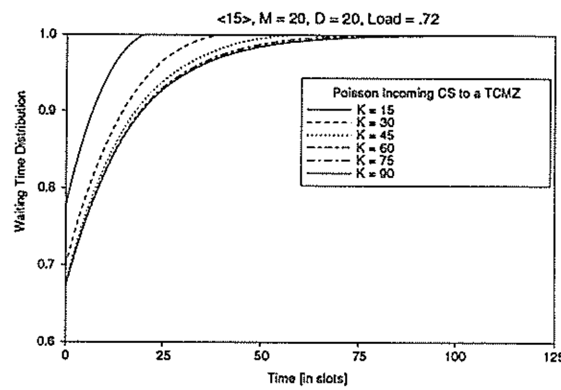


Figure 7d

Figure 7

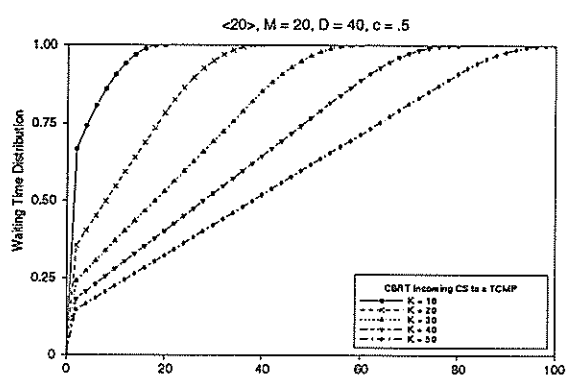


Figure 8

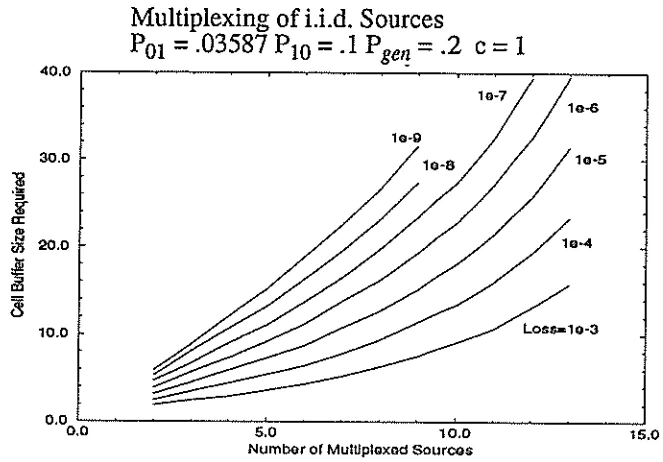
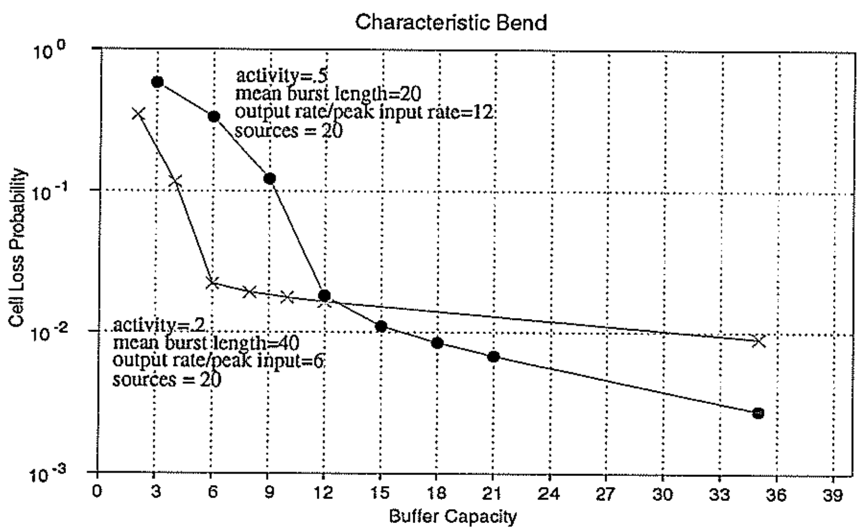


Figure 9: Buffering requirement for the multiplexing of a number of i.i.d sources at a specific cell loss probability.



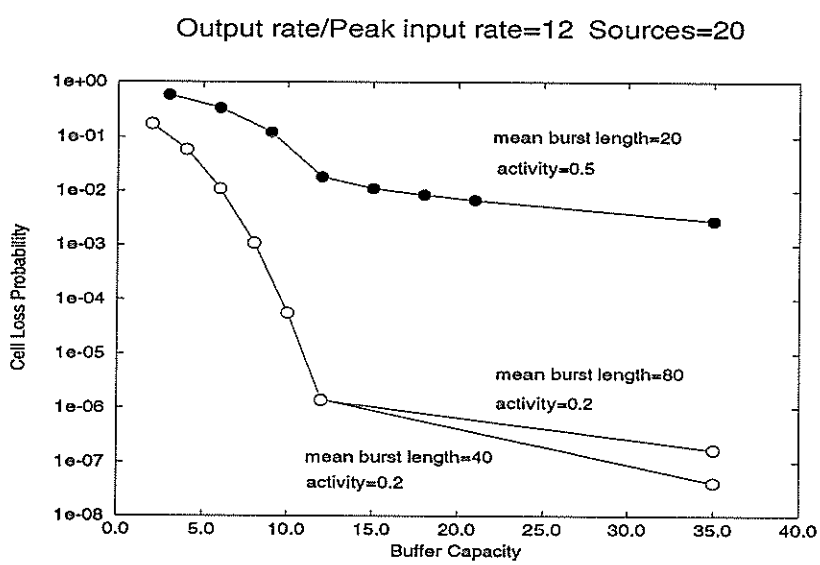


Figure 10