Analysis of a dynamic assignment queueing model with Poisson cluster arrival processes*

Toshihisa OZAWA* and Takuya Asaka**
*Department of Business Administration, Komazawa University
1-23-1 Komazawa, Setagaya-ku, Tokyo 154-8525, Japan
E-mail: toshi@komazawa-u.ac.jp
**NTT Multimedia Networks Laboratories

1999.10.20

Abstract

We consider a dynamic assignment queueing model with multiple packet classes, which has a number of queues, each with its own server. This model arises from the output buffer control of an ATM-based packet switching system, which is connected to another system via multiple links. Each packet is divided into cells and transmitted by cell-by-cell transmission through the links. Such packet arrival processes can be modeled as Poisson cluster arrival processes. An arriving packet is assigned to one of the queues according to a dynamic packet assignment scheme, which is a variation of the shortest queue policy and tries to assign buffer space and/or transmission bandwidth fairly to each class when the system is congested. We derive an approximation of the packet loss probability by using a decomposition method and an asymptotic of the cell loss probability. Its accuracy is examined in comparison with simulation results. The results of this paper can be used for dimensioning the buffer sizes of the ATM-based packet switching systems.

1 Introduction

Recent developments in computer communication require higher speed interconnection services between local area networks (LANs) over wide geographical areas. One way to provide these services is to build a packet overlay network on top of an asynchronous transfer mode (ATM) backbone network by using ATM-based packet switching systems, whose typical examples are connectionless servers (CLSs) [3, 4, 13, 17]. In such a packet overlay network, ATM-based packet switching systems are interconnected via permanent virtual circuits (PVCs). Each packet is divided into cells of 53 bytes in data length including the headers and is transmitted by cell-by-cell transmission through the PVCs. If ATM adaptation layer (AAL) type 3 or type 4 is used in the network, the first and last cells of each packet can be detected by using the values of the segment types (STs) of cells, and if AAL type 5 is used, those cells can be detected by using the values of the payload types (PTs) of cells. Hence the ATM-based packet switching systems can identify individual packets without reassembling the packets. Furthermore, by reading packet destination addresses directly from the first cells of the packets, they can also route and forward the packets by high-speed cell-by-cell processing without reassembling the packets.

When we attempt to build such a packet overlay network, we must sometimes use multiple PVCs to connect ATM-based packet switching systems because of technological and/or economical restrictions [5, 7] (Fig. 1). In this case, a packet assignment scheme for multiple PVCs is necessary to construct an output buffer control mechanism used in the ATM-based packet switching system. Here, assigning a packet to a queue corresponds to assigning all the cells of the packet to the same queue[1]. Such an output

---

[1]Cell assignment schemes of this type make it possible for ATM-based packet switching systems to identify each packet, whose cells are transmitted through a common PVC, by using the values of the STs or the PTs of cells. Cell assignment schemes that can assign the cells of each packet to different queues are the other type, and they require ATM-based packet
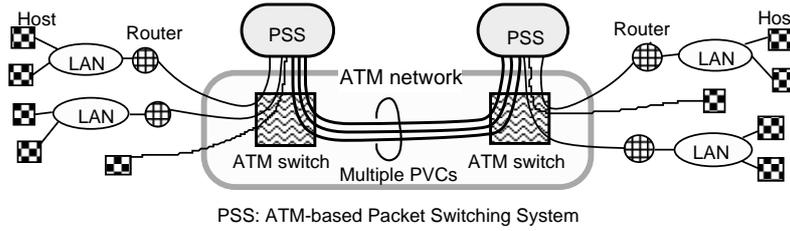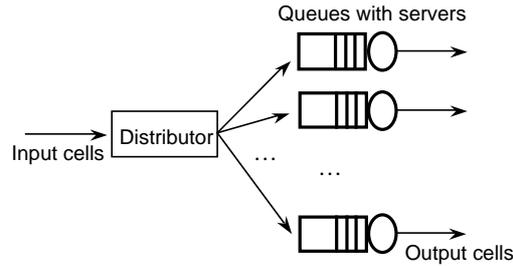
Figure 1: A packet overlay network.

Figure 2: A parallel server model.

buffer control mechanism together with multiple PVCs can be modeled as a parallel server model, where servers correspond to the PVCs and each server has its own cell buffer (Fig. 2). One of the typical packet assignment schemes is the shortest queue policy [19], which is a dynamic packet assignment scheme and randomly assigns an arriving packet to one of the shortest queues. Note that in this paper the queue length is counted including the cell being served at the server. This policy intends to use the servers more efficiently and to make the packet loss probability smaller.

However, when the shortest queue policy is used and one user sends many packets at a time, packet delays for other users may become longer and their many packets may be lost. Hence we consider the case of multiple packet classes, each corresponding to a user, and a variation of the shortest queue policy in order to assign buffer space and/or transmission bandwidth fairly to each class when the system is congested. Here each class can be identified by, for example, the source addresses (SAs) of the packets, their destination addresses (DAs), and the SA-DA pairs, and it can also be identified by the virtual channel identifiers (VCIs), the virtual path identifiers (VPIs), and/or the multiplexing identifiers (MIDs) of the cells of the packets [1, 2]. In our packet assignment scheme, at the arrival of a packet if there exist no idle servers and there exist packets of the same class in the system, the arriving packet is assigned to the queue to which the latest packet of the same class was assigned. The scheme is composed of the following cell-based packet assignment rules.

(i) At the arrival of the first cell of a packet if there exist no cells of the same class in the system, the cell is randomly assigned to one of the shortest queues.

(ii) At the arrival of the first cell of a packet if there exist cells of the same class in the system and if there are no idle servers, then the cell is assigned to the queue to which the latest packet of the same class was assigned.

(iii) At the arrival of the first cell of a packet if there exist cells of the same class in the system and if there exist idle servers, then the cell is randomly assigned to one of the queues of the idle servers.

(iv) A cell which is not the first cell of a packet is assigned to the same queue as the first cell of the packet.

---

switching systems to identify the individual packets whose cells may be transmitted through different PVCs [7]; This makes it difficult to construct dynamic assignment schemes of the latter type. Hence we consider only those of the former type in this paper.

Rule (i) corresponds to the shortest queue policy. Rule (ii) makes packets of each class apt to be assigned to just one queue when the system becomes congested, and these rules lead to fairness among classes. Rule (iii) contributes to using servers more efficiently. Rule (iv) makes this cell assignment scheme a packet assignment scheme.

The input process of each class can be modeled as a Poisson cluster arrival process [14]. Packets, which are called clusters in the Poisson cluster arrival process, arrive at the system via a Poisson process, and each packet is decomposed into a certain number of cells, which are called customers in that arrival process. These cells of the packet are usually assumed to arrive at the system with constant intervals[2] and are sent out through a common PVC. We approximately analyze this parallel server model with finite buffers and derive an approximation for the packet loss probability. Here a packet is assumed to be lost when at least one cell of the packet is lost in the network, since that packet cannot be assembled at the destination.

The rest of the paper is constructed as follows. In Section 2 we describe our parallel server model in detail. In Section 3 we introduce an approximate model for a single queue of the original model. In Section 4 we analyze the approximate model and obtain an approximation for the packet loss probability. In Section 5 the accuracy of the approximation is discussed through numerical experiments, and an improvement of the approximation of the packet loss probability is proposed.

2   Model description

Here we present a rigorous definition of our parallel server model with multiple packet classes. Our model is described as follows (see Fig. 2).

There are S queues. Each queue behaves like a $\cdot/D/1/K$ model, where service times are equal to a constant $\Delta \, (> 0)$, which corresponds to the time taken to transmit one cell, and K is the buffer size including the service position. In each queue cells are served according to the first-in-first-out (FIFO) discipline.

Let N be the number of packet classes. We only consider a symmetric case, where the arrival processes of individual classes are mutually independent and subject to a common stochastic law. Packets in each class arrive at the system via a homogeneous Poisson process with intensity λ_0. The sizes (the number of cells) of packets in all classes are i.i.d. random variables and we denote by X_0 a generic random variable representing the packet size. Let the time between consecutive cell arrivals in the same packet be equal to a constant $\delta \, (> 0)$. Here $1/\delta$ corresponds to the cell transmission speed of PVCs connecting user routers to ATM-based packet switching systems, while $1/\Delta$ corresponds to that of PVCs interconnecting the ATM-based packet switching systems. The former speed is usually less than or equal to the latter one; However we sometimes use PVCs of the same speed to build a packet overlay networks. Hence we assume that they have the same value (i.e. $\Delta = \delta$). We will refer to an arrival process of this type as a Poisson cluster arrival process [14] and denote by $M^{B(X,D)}$, where $B(X,D)$ represents the characteristic of each packet as in [11]; X means that the distribution of the packet (cluster) size is general and D means that the cell (customer) arrival intervals in the packet are determinate. Packets (and hence cells) are assigned to one of the queues according to the shortest-queue-type packet assignment scheme described in the previous section. A cell is lost if it arrives at the system when the buffer of the assigned queue is full. A packet is considered to be lost if at least one cell of the packet is lost.

For convenience of description, in the following sections, Δ and δ are set equal to the unit time length (i.e. $\Delta = \delta = 1$).

3   Decomposition of the model

The original model introduced in the preceding section is very complicated, and it seems difficult to get exact results for performance measures. In order to derive an approximation for the packet loss probability in the next section, here we shall decompose the original model into S independent $M^{B(X,D)}/D/1/K$ models each of which approximately describes the stochastic behavior of the corresponding queue. Hereafter we assume that all processes considered in this section are in steady states.

---

[2]In [14], these interarrival times are assumed to be i.i.d. random variables.

For $j = 1, 2, ..., S$, let $\tilde{A}_j(t)$ be the number of packets that are assigned to queue j during the interval $(0, t]$, $\tilde{\alpha}_j(t)$ be the stochastic intensity of $\tilde{A}_j(t)$ at time t, and $L_j(t)$ be the queue length of queue j, or equivalently the number of cells being served or waiting in queue j, at time t. First, we derive the conditional expectation of $\tilde{\alpha}_j(t)$ given that $L_j(t) > 0$; $\mathrm{E}[\tilde{\alpha}_j(t) \mid L_j(t) > 0]$ is decomposed as follows:

$$
\begin{aligned}
&\mathrm{E}\left[\tilde{\alpha}_j(t) \mid L_j(t) > 0\right] \\
=\ & \mathrm{E}\left[\tilde{\alpha}_j(t) \,\middle|\, L_j(t) > 0, \prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 0\right] \mathrm{Pr}\left(\prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 0 \,\middle|\, L_j(t) > 0\right) \\
&+ \mathrm{E}\left[\tilde{\alpha}_j(t) \,\middle|\, \prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 1\right] \mathrm{Pr}\left(\prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 1 \,\middle|\, L_j(t) > 0\right).
\end{aligned} \tag{1}
$$

The conditions "$L_j(t) > 0$" and "$\prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 0$" mean that, at time t, there exist empty queues but queue j is not empty. According to Rules (i) and (iii), the expectation in the first term on the right side of equation (1) is, therefore, zero. The expectation in the second term is given as

$$
\mathrm{E}\left[\tilde{\alpha}_j(t) \,\middle|\, \prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 1\right] = \frac{N}{S}\lambda_0\,, \tag{2}
$$

because the summation of the left side of equation (2) on j is equal to $N\lambda_0$ and S queues are symmetric. Substituting (2) into (1), the conditional expectation of $\tilde{\alpha}_j(t)$ is given as

$$
\mathrm{E}\left[\tilde{\alpha}_j(t) \mid L_j(t) > 0\right] = \frac{N}{S}\lambda_0 \mathrm{Pr}\left(\prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 1 \,\middle|\, L_j(t) > 0\right). \tag{3}
$$

Note that, since Rule (ii) is not used for deriving equation (3) explicitly, another expression for the expectation of $\tilde{\alpha}_j(t)$ seems to be needed to approximate the packet arrival rate under the condition where the system is congested, in other words, where the value of $L_j(t)$ is large. This will be discussed again in Subsection 5.2.

The conditional probability on the right side of formula (3) is represented as

$$
\mathrm{Pr}\left(\prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 1 \,\middle|\, L_j(t) > 0\right) = \frac{\mathrm{Pr}\left(\prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 1\right)}{\mathrm{Pr}(L_j(t) > 0)}. \tag{4}
$$

Since S queues are symmetric, the denominator on the right side of this equation is given by

$$
\mathrm{Pr}(L_j(t) > 0) = (1 - b_C^{(K)})\frac{N}{S}\lambda_0 \mathrm{E}[X_0], \tag{5}
$$

where $b_C^{(K)}$ denotes the cell loss probability. The numerator is the probability that all the servers are busy, and its approximation is derived as follows. Consider a modification of the parallel server model, where each packet is regarded as one customer and his service time is equal to the size of the packet. In this modified model, customers of each class arrive via a Poisson process with intensity λ_0, an arriving customer is assigned to one of the queues according to the same scheme, and each queue behaves like a $\cdot/G/1/K$. Here we consider a case where the capacities of the buffers are infinite. Since, in the original model, cell arrival intervals in each packet are equal to the constant service time, the probability that all the servers are busy in the original model is equal to that probability in the modified model. Hence we approximate that probability of the original model by using the probability that all the servers are busy in an $M/G/S/\infty$ model, where the arrival rate is equal to $(1 - b_C^{(K)})N\lambda_0$ and the service time distribution is the same as that of X_0. From [9], this probability of the $M/G/S/\infty$ model can be approximated by using the corresponding probability of an $M/M/S/\infty$ model, where the arrival rate is the same as that of the $M/G/S/\infty$ model and the mean service time is equal to $\mathrm{E}[X_0]$. As a result, we obtain the following approximation:

$$
\mathrm{Pr}\left(\prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 1\right) \approx \frac{S\tilde{a}_0^S \tilde{p}_0}{S!\,(S - \tilde{a}_0)}, \tag{6}
$$

4

where

$$\tilde{a}_0 \equiv (1 - b_C^{(K)}) \, N \, \lambda_0 \, \mathrm{E}[X_0], \tag{7}$$

$$\tilde{p}_0 \equiv \left(\sum_{j=0}^{S-1} \frac{\tilde{a}_0^j}{j!} + \frac{\tilde{a}_0^S}{S!} \frac{S}{S - \tilde{a}_0} \right)^{-1}. \tag{8}$$

From formulas (4), (5), and (6), the following approximation is obtained.

$$\Pr\left(\prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 1 \,\middle|\, L_j(t) > 0 \right) \approx \frac{S\tilde{a}_0^{S-1}\tilde{p}_0}{(S-1)! \, (S - \tilde{a}_0)}. \tag{9}$$

A decomposed $M^{B(X,D)}/D/1/K$ model is given as follows: let λ be the packet arrival intensity in the model. From formulas (3) and (9), λ is given by

$$\lambda \equiv \frac{N\lambda_0\tilde{a}_0^{S-1}\tilde{p}_0}{(S-1)! \, (S - \tilde{a}_0)}. \tag{10}$$

The distribution of packet sizes and the value of service times of cells in the decomposed model are the same as those in the original model.

$\mathrm{E}[\hat{\alpha}_j(t) \,|\, L_j(t) > 0]$ is the conditional mean arrival intensity of packets given that $L_j(t) > 0$; however, we use the same value for the case when $L_j(t) = 0$. The reason is as follows. From Remark 1 below, both the cell loss probability and the packet loss probability are independent of the packet arrival intensity during idle periods because the packet arrival processes are assumed to be Poissonian and the beginning point of each busy period is a regeneration point for the cell arrival processes.

[**Remark 1**] Consider a $G/GI/c/K$ model. Let $\hat{A}(t)$ be the number of customers that arrive at the system during an interval $(0, t]$ and $\hat{V}_n$ be the service time of the nth arriving customer. Suppose that the beginning point of each busy period is a regeneration point for $\hat{A}(t)$. Let $\hat{B}(t)$ be the number of customers that overflow from the system during an interval $(0, t]$. Since $\hat{B}(t)$ can be generated from $\{\hat{A}(s), \, s \le t\}$ and $\{\hat{V}_n, \, n \le \hat{A}(t)\}$, the beginning point of each busy period is also a regeneration point for $\hat{B}(t)$. Therefore, from the property of renewal-reward processes [8], the overflow probability $\hat{b}$ is given by

$$\hat{b} = \lim_{t \to \infty} \frac{\hat{B}(t)}{\hat{A}(t)} = \frac{\mathrm{E}[\hat{B}_0]}{\mathrm{E}[\hat{A}_0]},$$

where $\hat{A}_0$ is the number of customers that arrive during a busy cycle (a busy period and the subsequent idle period), and $\hat{B}_0$ is the number of customers that overflow during the busy cycle. Because all the arrivals and overflows in each busy cycle occur in the busy period, both $\hat{A}_0$ and $\hat{B}_0$ are independent of the probability law of $\hat{A}(t)$ during the idle period of the busy cycle. Therefore, $\hat{b}$ is also independent of the probability law of $\hat{A}(t)$ during idle periods. □

4 An approximate analysis of a discrete-time version of the $M^{B(X,D)}/D/1/K$ model

In this section, we consider a discrete time version of the $M^{B(X,D)}/D/1/K$ model; cells arrive and depart only at discrete time points $\{0, 1, 2, \dots\}$. We further assume that departures occur before arrivals at the same time point. For the case where the distribution of X_0 is geometric, the cell loss probability was derived in [11]. Here we assume that X_0 is bounded above and do not assume that its distribution is geometric. An approximation of the packet loss probability is derived in the following three steps: (1) An upper bound of the cell loss probability is obtained. (2) Using this upper bound, an approximation of the cell loss probability is derived. (3) The packet loss probability is approximated using the cell loss probability.

4.1 Upper bound of the cell loss probability

Let $a_P(n)$ be the number of packets arriving at time n. Then $\{a_P(n)\}$ are i.i.d. random variables subject to a Poisson distribution with mean λ. Let X_j be the size (the number of cells) of the jth arriving packet and assume that each X_j has the same upper bound g_{max}. Packets are numbered sequentially, with the first packet that arrives after time zero having the number zero. Let ρ be the traffic intensity defined as $\rho \equiv \lambda \operatorname{E}[X_0]$. Let $a_C(n)$ be the number of cells that arrive at time n. When the jth packet arrives at time n, its cells arrive sequentially at n, $(n+1)$, ..., $(n+X_j-1)$. Let $I_P(n)$ be the number of active packets that have cells arriving after time n and $R_C(i,n)$ be the number of residual cells of the ith active packet for $i = 1, 2, ..., I_P(n)$; for example, if a packet of size X_k arrives at time n when the system is empty, then $R_C(1,n) = X_k - 1$, $R_C(1, n+1) = X_k - 2$, and so on. Let $Z(n)$ be a vector indicating the number of active packets and the number of residual cells:

$$Z(n) \;=\; (I_P(n); \, (R_C(1,n), \, R_C(2,n), \, \dots, \, R_C(I_P(n), n)\,). \tag{11}$$

Since $\{Z(n)\}$ forms a Markov chain and $a_C(n)$ satisfies equation

$$a_C(n) \;=\; a_P(n) + I_P(n-1), \tag{12}$$

$\{a_C(n)\}$ can be considered as a discrete-time Markov Aadditive process (MAP) with the underlying Markov chain $\{Z(n)\}$.

We first analyze an $M^{B(X,D)}/D/1/\infty$ model, which has the same arrival process and service times as the $M^{B(X,D)}/D/1/K$ model, but the buffer capacity is infinite. For each $n \geq 0$, consider a queueing process that began at time $-n$, and let $L_0(n)$ be the number of cells in the system at time zero. We introduce the following notations:

$$M_0(n) \;\equiv\; L_0(n) - a_C(0) \quad \text{for } n \geq 0, \tag{13}$$
$$u(n) \;\equiv\; a_C(-n) - 1 \quad \text{for } n \geq 0, \tag{14}$$
$$U(n) \;\equiv\; \sum_{j=1}^{n} u(j) \quad \text{for } n \geq 1, \quad U(0) = 0. \tag{15}$$

From the definition of the process $\{u(n)\}$, this process becomes a discrete-time MAP with the underlying Markov chain $\{Z(n)\}$. Renumbering the states of $\{Z(n)\}$ such that the state space becomes $\mathcal{Z}^+$, the kernel of the MAP and its transform are defined by

$$P_{ij}(k) \;\equiv\; \Pr(Z(n+1) = j, \, u(n+1) = k \mid Z(n) = i) \;=\; f_{ij}(k)\, p_{ij}, \tag{16}$$
$$\tilde{P}_{ij}(\theta) \;\equiv\; \sum_{k=-c}^{\infty} e^{\theta k} f_{ij}(k) p_{ij} \;=\; \tilde{f}_{ij}(\theta) p_{ij}, \tag{17}$$

where

$$f_{ij}(k) \;\equiv\; \Pr(u(n+1) = k \mid Z(n+1) = j, Z(n) = i), \tag{18}$$
$$p_{ij} \;\equiv\; \Pr(Z(n+1) = j \mid Z(n) = i), \tag{19}$$
$$\tilde{f}_{ij}(\theta) \;\equiv\; \sum_{k=-c}^{\infty} e^{\theta k} f_{ij}(k). \tag{20}$$

$L_0(n)$ and $M_0(n)$ are represented by $U(\cdot)$ in the following manner:

$$\begin{aligned} L_0(n) &= [u(1) + [u(2) + \cdots + [u(n)]^+ \,]^+ \,]^+ + a_C(0) \\ &= \max_{0 \leq j \leq n} U(j) + a_C(0), \end{aligned} \tag{21}$$
$$M_0(n) = \max_{0 \leq j \leq n} U(j). \tag{22}$$

Under some suitable conditions for $\{a_C(n)\}$, the stationary versions of $L_0(n)$ and $M_0(n)$, L_0 and M_0, are obtained as follows:

$$M_0 = \lim_{n \to \infty} M_0(n) = \sup_{n \in \mathcal{Z}^+} U(n), \tag{23}$$
$$L_0 = \lim_{n \to \infty} L_0(n) = M_0 + a_C(0). \tag{24}$$

6

We further introduce the following notations:

$$\Lambda(n;\theta) \equiv \log \mathrm{E}[e^{\theta U(n)}] \quad \text{for } n > 0 \text{ and } \theta \in \mathcal{R}, \tag{25}$$

$$\Lambda(\theta) \equiv \lim_{n\to\infty} n^{-1}\Lambda(n;\theta) \quad \text{for } n > 0 \text{ and } \theta \in \mathcal{R}, \tag{26}$$

$$\theta^* \equiv \sup\{\theta \mid \Lambda(\theta) \leq 0\}. \tag{27}$$

Then, using the results of [6], an upper bound of the tail distribution of M_0 is given by the following theorem.

[Theorem 1] [6]   If $\mathrm{E}[\,u(1)\,] = \rho - 1 < 0$, then

$$\Pr(M_0 \geq k) \leq \varphi(\theta)\,e^{-\theta k} \quad \text{for } k > 0 \text{ and } \theta \in [0,\theta^*], \tag{28}$$

where $\varphi(\theta)$ is the essential supremum

$$\varphi(\theta) \equiv \sup\left\{ x \,:\, \Pr\left(\frac{1(u(n) > 0)}{v_{Z(n)}(\theta)} > x\right) > 0 \right\}, \tag{29}$$

and $(v_i(\theta))$ is the right eigenvector of the matrix $(\tilde{P}_{ij}(\theta))$ corresponding to the simple maximal eigenvalue $e^{\Lambda(\theta)}$. The vector $(v_i(\theta))$ is positive and bounded above. $\qquad\square$

Since $[L_0 - 1]^+ = [M_0 + u(0)]^+ \overset{\mathrm{d}}{=} M_0$, an upper bound of the tail distribution of L_0 is given as

$$\begin{aligned}
\Pr(L_0 \geq k) &= \Pr([L_0 - 1]^+ \geq k - 1) \\
&= \Pr(M_0 \geq k - 1) \\
&\leq \varphi(\theta)\,e^{-\theta(k-1)} \quad \text{for } k > 1 \text{ and } \theta \in [0,\theta^*].
\end{aligned} \tag{30}$$

Let $b_C^{(K)}$ be the cell loss probability in the $M^{B(X,D)}/D/1/K$ model. Then we have the following theorem.

[Theorem 2]

$$b_C^{(K)} \leq \rho^{-1}\Pr(L_0 \geq K + 1). \tag{31}$$

(Proof)   See Appendix. $\qquad\square$

Theorem 2 and (30) directly lead to an upper bound of $b_C^{(K)}$ as

$$b_C^{(K)} \leq \rho^{-1}\,\varphi(\theta)\,e^{-\theta K} \quad \text{for } \theta \in [0,\theta^*]. \tag{32}$$

4.2   Approximation of the cell loss probability

On the right side of inequality (32), θ^* is easily calculated according to Proposition 1 below; however, it is difficult to calculate $\varphi(\theta)$ because the eigenvector $(v_i(\theta))$ should be calculated. Hence we try to approximate the cell loss probability. From [6], the asymptotic decay rate of the distribution of M_0 is $-\theta^*$, i.e. $\lim_{K\to\infty} K^{-1}\log\Pr(M_0 \geq K) = -\theta^*$. From this and Theorem 2, it is proved that the asymptotic decay rate of $b_C^{(K)}$ with respect to K is bounded by $-\theta^*$, as follows:

$$\begin{aligned}
\lim_{K\to\infty} K^{-1}\log b_C^{(K)} &\leq \lim_{K\to\infty} K^{-1}\log(\rho^{-1}\Pr(L_0 \geq K + 1)) \\
&= \lim_{K\to\infty} K^{-1}\{-\log\rho + \log\Pr(M_0 \geq K)\} = -\theta^*.
\end{aligned}$$

From this result we propose the following approximation:

$$b_C^{(K)} \approx b_C^{(1)}\,e^{-\theta^*(K-1)}. \tag{33}$$

$b_C^{(1)}$ is the cell loss probability of the $M^{B(X,D)}/D/1/1$ model, and it can be obtained by using the fluid flow approximation as follows. Consider an $M/G/\infty$ model, where the arrival rate is equal to λ and the service time distribution is the same as that of X_0. Let $L_\infty(t)$ be the number of cells in the system at time t. We assume that, in this $M/G/\infty$ model, work is lost at a rate $(L_\infty(t) - 1)$ when $L_\infty(t) > 1$, and let the volume of lost work correspond to the number of lost cells in the $M^{B(X,D)}/D/1/1$ model. As a result, $b_C^{(1)}$ is approximately given by the following formula:

$$
\begin{aligned}
b_C^{(1)} &\approx \lim_{t \to \infty} \frac{\int_0^t [L_\infty(s) - 1]^+ \, ds}{\int_0^t L_\infty(s) \, ds} \\
&= \frac{\mathrm{E}[L_\infty(\infty) - 1]^+}{\lambda \, \mathrm{E}[X_0]} \\
&= 1 - \rho^{-1}(1 - e^{-\rho}).
\end{aligned}
\tag{34}
$$

The value of θ^* can be calculated according to the following proposition.

[Proposition 1] Assuming $\rho = \lambda \, \mathrm{E}[X_0] < 1$, the value of θ^* is obtained as the unique positive root of the equation

$$
-(\theta + \lambda) + \lambda \, \xi(e^\theta) = 0,
\tag{35}
$$

where $\xi(z) \equiv \mathrm{E}[z^{X_0}]$. The root can be easily calculated using usual numerical methods, such as Newton's method or the binary search method.

(Proof) Let us consider a discrete-time version of an $M^X/D/1/\infty$ model, where all the cells of each packet arrive at the same time, not with regular intervals. The packet arrival process, packet sizes, and service times are the same as the discrete-time version of the $M^{B(X,D)}/D/1/K$ model. For this $M^X/D/1/\infty$ model, let $a_C^0(n)$ be the number of cells that arrive at time n, and let $u^0(n)$, $U^0(n)$, $\Lambda^0(n;\theta)$, and $\Lambda^0(\theta)$ be defined in the same manner as $u(n)$, $U(n)$, $\Lambda(n;\theta)$, and $\Lambda(\theta)$ in the previous subsection, respectively. This time $\{a_C^0(n)\}$ are i.i.d. random variables subject to a compound Poisson distribution.

Since the packet sizes are bounded above by g_{max}, we have

$$
U^0(n) - U^0(g_{max}) \leq U(n) \leq U^0(n + g_{max}) \qquad \text{for } n \geq g_{max}.
\tag{36}
$$

Hence the following inequality holds:

$$
\begin{aligned}
\log \mathrm{E}[e^{\theta U^0(n)}] + \log \mathrm{E}[e^{-\theta U^0(g_{max})}] &\leq \log \mathrm{E}[e^{\theta U(n)}] \\
&\leq \log \mathrm{E}[e^{\theta U^0(n)}] + \log \mathrm{E}[e^{\theta U^0(g_{max})}].
\end{aligned}
\tag{37}
$$

Dividing each term by n and letting n tend to infinity, the relation $\Lambda(\theta) = \Lambda^0(\theta)$ is obtained. Therefore $\Lambda(\theta)$ is given by

$$
\Lambda(\theta) = \Lambda^0(\theta) = -\theta + \lambda(1 - \xi(e^\theta)).
\tag{38}
$$

$\square$

4.3 Approximation of the packet loss probability

Consider the $M^{B(X,D)}/D/1/K$ model and assume that it is ergodic. Let $A_P(n)$ and $A_C(n)$, respectively, be the number of packets that arriving during interval $[0, n)$ and that of cells arriving during the same interval; and let $B_P^{(K)}(n)$ and $B_C^{(K)}(n)$, respectively, be the number of packets lost during interval $[0, n)$ and that of cells lost during the same interval. A packet is assumed to be lost when at least one cell of the packet is lost. Let $b_P^{(K)}$ and $b_C^{(K)}$ be the packet loss probability and the cell loss probability, respectively. In order to represent $b_P^{(K)}$ in terms of $b_C^{(K)}$, we introduce the mean number of lost cells in a lost packet, κ. This is given by $\kappa \equiv \lim_{n \to \infty} B_C^{(K)}(n)/B_P^{(K)}(n)$. Using κ, $b_P^{(K)}$ is given by

$$
b_P^{(K)} = \lim_{n \to \infty} \frac{B_P^{(K)}(n)}{A_P(n)} = \lim_{n \to \infty} \frac{\frac{B_C^{(K)}(n)}{A_C(n)} \frac{A_C(n)}{A_P(n)}}{\frac{B_C^{(K)}(n)}{B_P^{(K)}(n)}} = \frac{b_C^{(K)} \mathrm{E}[X_0]}{\kappa},
\tag{39}
$$

where the upper and lower bounds of κ are given by

$$1 \le \kappa \quad \le \quad \lim_{n \to \infty} \frac{B_C^{(K)}(n)}{B_P^{(K)}(n)} \le \lim_{n \to \infty} \frac{\sum_{j=0}^{B_P^{(K)}(n)} X_j}{B_P^{(K)}(n)} = \mathrm{E}[X_0]. \tag{40}$$

The value of κ depends on how the cells of a packet are consecutively lost. In [10], it was shown that the lengths of the periods of loss are independent of the buffer size. From this and numerical experiments in the next section, we conjecture that the value of κ is almost invariant with respect to the buffer size K if the value of K is sufficiently large. However, it is difficult to calculate κ except for the case when $K = 1$. Therefore, we use the value of κ in that case ($K = 1$) as an approximation of κ. Letting this value of κ be $\hat{\kappa}$, the following approximation of $b_P^{(K)}$ is obtained from (33) and (39).

$$b_P^{(K)} \quad \approx \quad \frac{\mathrm{E}[X_0]}{\hat{\kappa}} b_C^{(1)} e^{-\theta^*(K-1)}. \tag{41}$$

$\hat{\kappa}$ is approximately obtained as follows. Consider the $M^{B(X,D)}/D/1/1$ model, and let α be the probability that a tagged packet is not lost. Using α, $\hat{\kappa}$ is given by

$$\hat{\kappa} \quad = \quad \frac{\rho \, b_C^{(1)}}{\lambda \, (1 - \alpha)} = \frac{\mathrm{E}[X_0] \, b_C^{(1)}}{1 - \alpha}, \tag{42}$$

where $b_C^{(1)}$ is given by formula (34). We approximate the value of α by using the probability that a tagged packet arrives when the system is empty and no other packets arrive at the system until all the cells of the tagged packet arrive. Since a tagged packet arriving when the system is not empty may not be lost, this approximation underestimates the value of α. Therefore, $\hat{\kappa}$ is also underestimated by its approximation. Let Y_0 be the random variable representing a time interval between consecutive packet arrivals. The distribution of Y_0 is exponential with mean λ^{-1}. The approximation of α is given by

$$\begin{aligned} \alpha \quad &\approx \quad (1 - \rho(1 - b_C^{(1)})) \, \mathrm{Pr}(X_0 - 1 \le Y_0) \\ &= \quad 1 - \rho + \rho \, b_C^{(1)}) \sum_{j=1}^{\infty} e^{-\lambda(j-1)} \, \mathrm{Pr}(X_0 = j) \\ &= \quad (1 - \rho + \rho \, b_C^{(1)}) \, e^{\lambda} \, \xi(e^{\lambda}), \end{aligned} \tag{43}$$

where we assume that $\mathrm{Pr}(X_0 = 0) = 0$, and $b_C^{(1)}$ is given by formula (34).

5   Numerical experiments and an improvement of the approximation

This section explains the accuracy of the approximations for the cell loss probability, the packet loss probability, and the mean number of lost cells in a lost packet κ by comparing them with simulation results; it also presents an improvement of the approximation of the cell loss probability for a parallel server model with the shortest-queue-type packet assignment scheme. Hereafter we use two types of packet size distributions in common: one is a unit distribution, whose mean is equal to 35, i.e. $\mathrm{E}[X_0] = 35$; the other is a uniform distribution, whose minimum and maximum values are equal to 1 and 35, and whose mean is equal to 18, i.e. $\mathrm{E}[X_0] = 18$. The reason why we use the packet size of 35 as the maximum packet size is that the maximum packet length in the Ethernet is 1518 bytes and a packet of this size is divided into 35 cells when AAL type 3 or 4 is used.

5.1   An $M^{B(X,D)}/D/1/K$ model

Figures 3 and 4 show comparisons between approximation results and simulation results for the cell and packet loss probabilities and κ in an $M^{B(X,D)}/D/1/K$ model. In each case, packet arrival intensity λ is determined to make the traffic intensity $\rho \, (= \lambda \, \mathrm{E}[X_0])$ be equal to a given value.

Cell loss probability

From Figs. 3(a) and 4(a), we see that the approximation of the cell loss probability is sufficiently accurate for all cases.

The mean number of lost cells in a lost packet: κ

From Figs. 3(b) and 4(b), we see that, for all cases, κ is underestimated by its approximation when the value of the buffer size K is near one, and it is overestimated by its approximation when the value of K is large; the reason for the former was explained in Subsection 4.3. From these experiments, we propose the following conjecture about κ: the value of κ is rapidly decreasing as K is increasing, and it converges to some value. If we obtain the value that κ converges to, it is expected to improve the approximation.

Packet loss probability

From Figs. 3(a) and 4(a), we see that the approximation of the cell loss probability is also sufficiently accurate for all cases. Since the approximations of the cell loss probability and κ make the approximation of the packet loss probability, the approximation errors of the packet loss probability result from those of the cell loss probability and κ. From these experiments, the errors of κ do not affect the accuracy of the approximation of the packet loss probability so much.
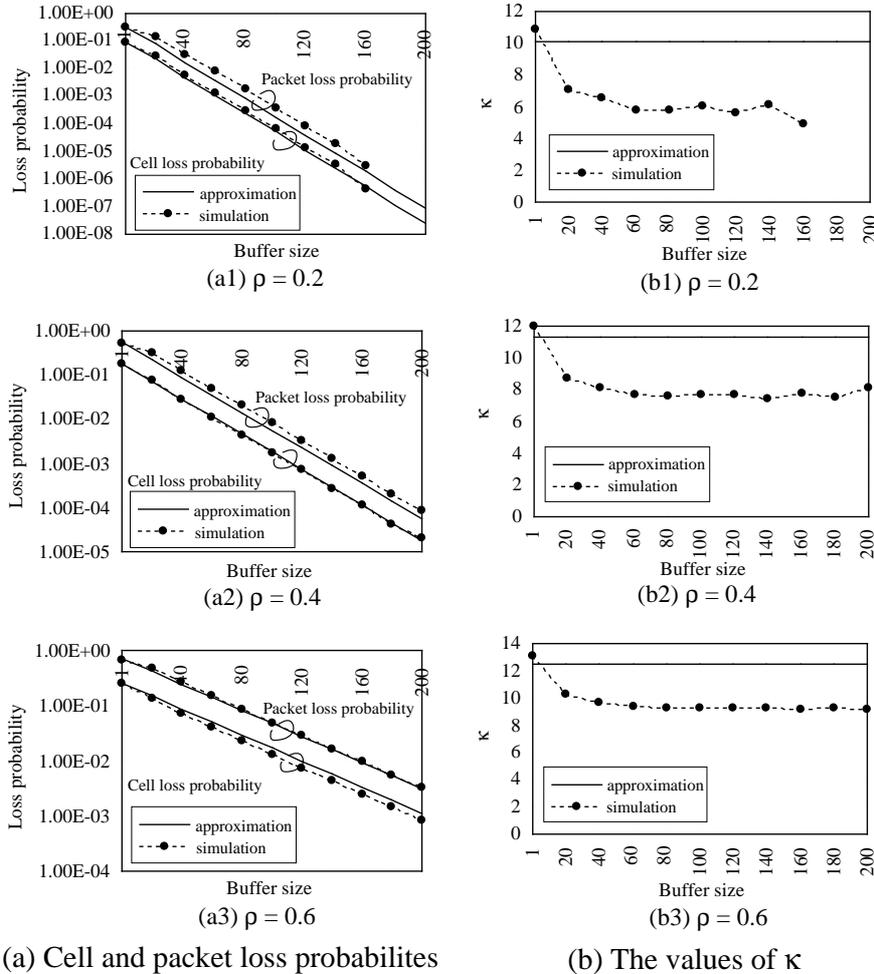

(a) Cell and packet loss probabilites   (b) The values of κ

Figure 3: An $M^{B(X,D)}/D/1/K$ model ($X_0 = 35$: constant).

(a1) ρ = 0.2   (b1) ρ = 0.2

(a2) ρ = 0.4   (b2) ρ = 0.4

(a3) ρ = 0.6   (b3) ρ = 0.6

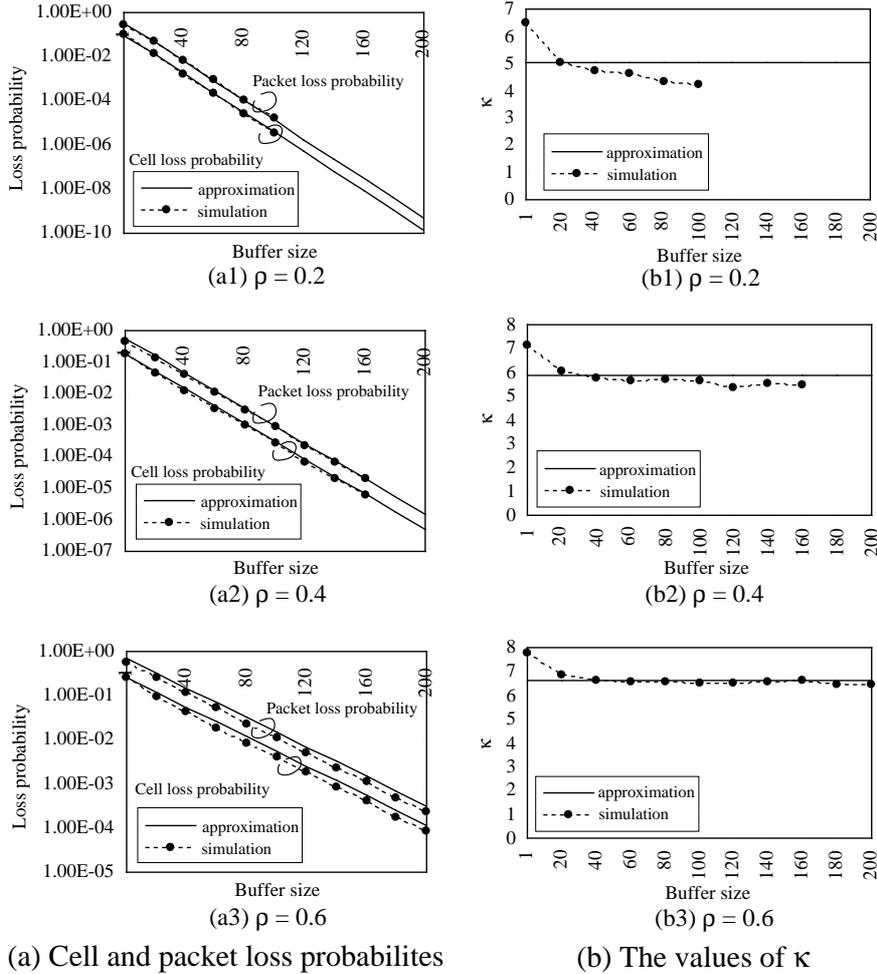(a) Cell and packet loss probabilites   (b) The values of κ

Figure 4: An $M^{B(X,D)}/D/1/K$ model (X_0 is uniformly distributed between 1 and 35).

5.2 A parallel server model and an improvement of the approximation

Figure 5 shows comparisons between approximation results and simulation results for the cell and packet loss probabilities in a parallel server model with the shortest-queue-type packet assignment scheme. Since $\tilde{a}_0$ in formula (10) contains the cell loss probability $b_C^{(K)}$, we here use an iteration to obtain the approximation results, as follows; Initially, set $b_C(K)$ in formula (7) to be zero and calculate an approximate value of the cell loss probability of the parallel server model by using formula (10) and the results of section 4; Next, set $b_C(K)$ in formula (7) to be this approximate value and calculate a new approximate value of the cell loss probability; Continue this procedure until the difference between a new approximate value of the cell loss probability and the previous one becomes sufficiently small. From this figure we see that our approximations are accurate only when the buffer size K is near one. The reason for this is considered as follows: if the number of cells in queue j is greater than one, at least two packets have recently been assigned to the queue and at the time point when the latest packet was assigned to the queue all the servers in the system were busy. This means that the next packet of the same class as the latest packet is probably assigned to the same queue according to Rule (ii). In such a situation, it is considered that the conditional expectation of $\tilde{\alpha}_j(t)$, $\mathrm{E}\left[\tilde{\alpha}_j(t) \,|\, L_j(t) > 0\right]$, can be represented as the following formula more accurately than formula (3).

$$\mathrm{E}\left[\tilde{\alpha}_j(t) \,|\, L_j(t) > 0\right] \;\approx\; \lambda_0 + \frac{N-S}{S}\lambda_0 \Pr\left(\prod_{\ell=1}^{S} 1(L_\ell(t) > 0) = 1 \;\middle|\; L_j(t) > 0\right)$$

11

$$= \lambda_0 + \frac{(N-S)\lambda_0 \tilde{a}_0^{S-1} \tilde{p}_0}{(S-1)!(S-\tilde{a}_0)}. \qquad (44)$$

On the right side of the first line of the equation, the first term represents that at least one class is assigned to queue j; This means that if an arriving packet of another class is assigned to queue j, at least one of other classes is assigned to each queue. Hence the second term on the right side of the first line contains the factor $(N-S)/S$ instead of N/S. Using this formula, we propose a hybrid approximation, in which the packet arrival intensity is given by formula (10) when the value of K is small, and it is given by formula (44) when the value of K is large. Let λ_1 be the packet arrival intensity given by formula (10) and λ_2 be that given by formula (44). When λ is given by λ_1, let θ^* and $b_C^{(1)}$ be denoted by θ_1^* and $b_{C,1}^{(1)}$; when λ is given by λ_2, let θ^* be denoted by θ_2^*. A new approximation of the cell loss probability is given by

$$b_C^{(K)} \approx b_{C,1}^{(1)}(1 - b_{C,2}^{(1)}/b_{C,1}^{(1)}) e^{-\theta_1^*(K-1)} + b_{C,2}^{(1)} e^{-\theta_2^*(K-1)}, \qquad (45)$$

where $b_{C,2}^{(1)}$ is the solution of equation

$$b_{C,1}^{(1)} e^{-\theta_1^*(K_0-1)} = b_{C,2}^{(1)} e^{-\theta_2^*(K_0-1)}, \qquad (46)$$

and $K_0 \equiv \log \beta / \theta_2^*$; β is a parameter to determine the value of K at which the term of $e^{-\theta_2^*(K-1)}$ becomes superior to the term of $e^{-\theta_1^*(K-1)}$.
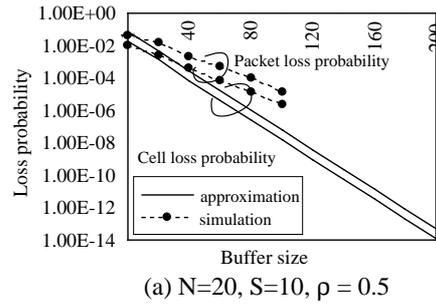

(a) N=20, S=10, $\rho = 0.5$

Figure 5: A parallel server model with Scheme B': the first versions of the approximations for the cell and packet loss probabilities ($X_0 = 35$: constant).

Figures 6 and 7 show comparisons between approximation results and simulation results for the cell and packet loss probabilities in a parallel server model with the shortest-queue-type packet assignment scheme, where the approximation results are calculated by using formula (45). The value of β is equal to 10 for all cases. From these figures, we see that the approximations of the cell and packet loss probabilities are rather accurate except for the cases where the packet size distribution is the uniform distribution and the value of ρ is relatively large.

6   Conclusions

This paper derived an approximation of the packet loss probability for a parallel server model with a dynamic packet assignment scheme, which is a variation of the shortest queue policy. Numerical examples showed the accuracy of the approximations. This approximation is useful for dimensioning the buffer sizes of ATM-based packet switching systems.
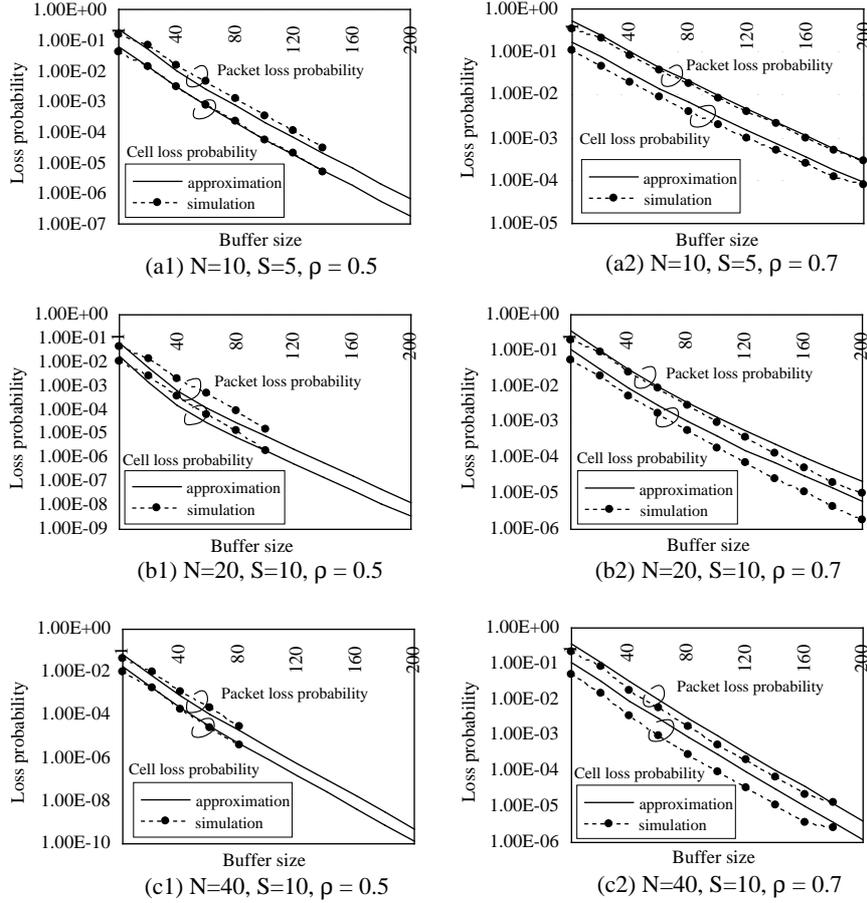
Figure 6: A parallel server model: cell and packet loss probabilities ($X_0 = 35$: constant).

A Proof of Theorem 2

In this appendix, an upper bound of the cell loss probability is obtained using the queueing model with an infinite buffer. Approximations for cell loss probability have been obtained in other works [12], [15], and [16].

We consider a general discrete-time queueing model denoted by $D^X/D/c/K$: arrivals and departures only occur at discrete time points $\{n : n \in \mathcal{Z}^+\}$, and it is assumed that the departures occur before the arrivals at the same time point. c denotes the number of servers and K denotes the buffer size including the service positions. Let a_n be the number of cells that arrive at time n. The arrival process, $a \equiv \{a_n, n \in \mathcal{Z}^+\}$, is general; that is, a_n may depend on $\{a_j, j \leq n-1\}$. Let $L_n^{(K)}$ be the number of cells in the system at time n. The queueing process $\{L_n^{(K)}, n \in \mathcal{Z}^+\}$ is given by the following recursive formulas:

$$L_{n+1}^{(K)} = \min\{K, [L_n^{(K)} - c]^+ + a_n\} \text{ for } n \in \mathcal{Z}^+ \quad \text{and} \quad L_0^{(K)} = a_0. \tag{47}$$

Before proving Theorem 2, the following proposition is presented:

[**Proposition A1**] ([18], Theorem 9) Consider two queueing models, $D^X/D/c/K_1$ and $D^X/D/c/K_2$, which have the same arrival process a. For these models, the following formula is obtained:

$$K_1 \leq K_2 \quad \Rightarrow \quad L_n^{(K_1)} \leq L_n^{(K_2)}, \quad \text{for all } n \in \mathcal{Z}^+. \tag{48}$$
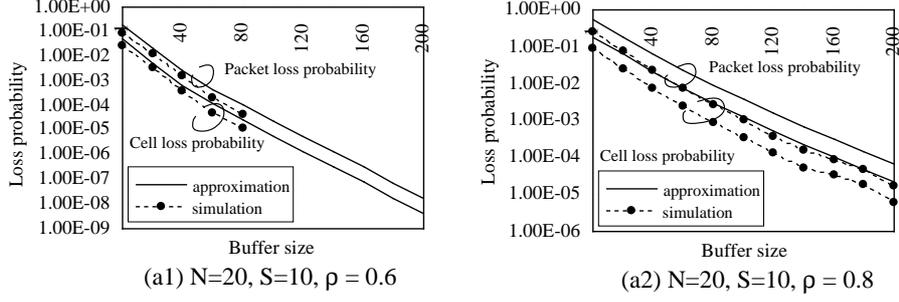
□

13

Figure 7: A parallel server model: cell and packet loss probabilities (X_0 is uniformly distributed between 1 and 35).

Let $B_n^{(K)}$ be the number of cells lost at time n. $B_n^{(K)}$ where $n \in \mathcal{Z}^+$ is given by

$$B_0^{(K)} = \max\{0, \ a_0 - K\}, \tag{49}$$
$$B_{n+1}^{(K)} = \max\{0, \ [L_n^{(K)} - c]^+ + a_{n+1} - K\}, \quad n \in \mathcal{Z}^+. \tag{50}$$

The cell loss probability $b^{(K)}$ is given by $b^{(K)} \equiv \lim_{n \to \infty} \sum_{\ell=0}^{n-1} B_\ell^{(K)} / \sum_{\ell=0}^{n-1} a_\ell$, and the traffic intensity ρ is given by $\rho \equiv \frac{1}{c} \lim_{n \to \infty} \sum_{\ell=0}^{n-1} a_\ell / n$. Let $(p_j^{(\infty)})$ be the limiting distribution of the number of cells in the system for the $D^X/D/c/\infty$ model with the same arrival process a. This is given by $p_j^{(\infty)} \equiv \lim_{n \to \infty} \sum_{\ell=0}^{n-1} 1(L_\ell^{(\infty)} = j)/n$.

[**Theorem 2**] (Generalized version) Assuming that the $D^X/D/c/\infty$ model is stable (i.e. $\lim_{n \to \infty} L_n^{(\infty)} < \infty$ w.p.1) and $\sum_{\ell=0}^{\infty} a_\ell = \infty$ w.p.1, an upper bound of $b^{(K)}$ is obtained as follows:

$$b^{(K)} \leq \frac{1}{\rho} \sum_{j=K+1}^{\infty} p_j^{(\infty)} \quad \text{w.p.1.} \tag{51}$$

(**Proof**) Let $\hat{Q}_n$ be defined by $\hat{Q}_n \equiv [L_n^{(\infty)} - K]^+$ where $n \in \mathcal{Z}^+$ and $\hat{a}_n$ by

$$\hat{a}_n \equiv \max\{0, \ a_n - \max\{0, \ K - [L_{n-1}^{(\infty)} - c]^+\}\}$$
$$= \max\{0, \ \min\{a_n, \ L_n^{(\infty)} - K\}\}. \tag{52}$$

To prove that $\{\hat{Q}_n\}$ is the queueing process for a $D^X/D/c/\infty$ model with arrival process $\{\hat{a}_n\}$, $\hat{Q}_0 = \hat{a}_0$ and $\hat{Q}_{n+1} = [\hat{Q}_n - c]^+ + \hat{a}_{n+1}$, $n \in \mathcal{Z}^+$ are verified as follows:

$$\hat{a}_0 = \max\{0, \ \min\{a_0, \ a_0 - K\}\}$$
$$= \max\{0, \ a_0 - K\} = \hat{Q}_0,$$

and

$$[\hat{Q}_n - c]^+ + \hat{a}_{n+1}$$
$$= \max\{0, \ \max\{0, \ L_n^{(\infty)} - K\} - c\} + \hat{a}_{n+1}$$
$$= \max\{\hat{a}_{n+1}, \ \hat{a}_{n+1} + L_n^{(\infty)} - K - c\}$$
$$= \max\{ \ [\min\{a_{n+1}, \ L_{n+1}^{(\infty)} - K\}]^+, \ [\min\{a_{n+1}, \ L_{n+1}^{(\infty)} - K\}]^+ + L_n^{(\infty)} - K - c\}$$
$$= \max\{0, \ \min\{a_{n+1}, \ L_{n+1}^{(\infty)} - K\}, \ L_n^{(\infty)} - K - c,$$
$$\min\{a_{n+1}, \ L_{n+1}^{(\infty)} - K\} + L_n^{(\infty)} - K - c\},$$

where

$$L_n^{(\infty)} < K + c \ \Rightarrow \ L_{n+1}^{(\infty)} - K = [L_n^{(\infty)} - c]^+ - K + a_{n+1} < a_{n+1}$$
$$\Rightarrow \ [\hat{Q}_n - c]^+ + \hat{a}_{n+1} = \max\{0, L_{n+1}^{(\infty)} - K\} = \hat{Q}_{n+1},$$

14

and

$$L_n^{(\infty)} \geq K + c \quad \Rightarrow \quad L_{n+1}^{(\infty)} - K = L_n^{(\infty)} - c - K + a_{n+1} \geq a_{n+1}$$
$$\Rightarrow \quad [\hat{Q}_n - c]^+ + \hat{a}_{n+1} = L_n^{(\infty)} - K - c + a_{n+1}$$
$$= [\,[L_n^{(\infty)} - c]^+ + a_{n+1} - K]^+ = \hat{Q}_{n+1}.$$

For the $D^X/D/c/\infty$ model with arrival process $\{\hat{a}_n\}$, let $\hat{d}_n$ be the number of cells that departed at time n. Since $\hat{d}_{n+1} = \min\{c, \hat{Q}_n\}$, $\hat{Q}_n = 0 \Rightarrow d_{n+1} = 0$, and $d_{n+1} \leq c$, the following inequality is obtained:

$$1(\hat{Q}_n > 0) \quad \geq \quad \frac{\hat{d}_{n+1}}{c}, \quad n \in \mathcal{Z}^+. \tag{53}$$

Using $1(\hat{Q}_n > 0) = 1([L_n^{(\infty)} - K]^+ > 0) = 1(L_n^{(\infty)} > K)$, we get

$$\sum_{\ell=0}^{n-1} 1(L_\ell^{(\infty)} > K) \quad = \quad \sum_{\ell=0}^{n-1} 1(\hat{Q}_\ell > 0)$$
$$\geq \quad \sum_{\ell=0}^{n-1} \frac{\hat{d}_{\ell+1}}{c}$$
$$= \quad \frac{1}{c}(\sum_{\ell=0}^{n-1} \hat{a}_\ell - [\hat{Q}_{n-1} - c]^+)$$
$$= \quad \frac{1}{c}(\sum_{\ell=0}^{n-1} \hat{a}_\ell - [L_{n-1}^{(\infty)} - K - c]^+). \tag{54}$$

From this formula and Proposition A4.1, we get

$$c \cdot \sum_{\ell=0}^{n-1} 1(L_\ell^{(\infty)} > K) + [L_{n-1}^{(\infty)} - K - c]^+ \quad \geq \quad \sum_{\ell=0}^{n-1} \max\{0, \min\{a_\ell, [L_{\ell-1}^{(\infty)} - c]^+ + a_\ell - K\}$$
$$\geq \quad \sum_{\ell=0}^{n-1} \max\{0, \min\{a_\ell, [L_{\ell-1}^{(K)} - c]^+ + a_\ell - K\}$$
$$= \quad \sum_{\ell=0}^{n-1} \max\{0, [L_{\ell-1}^{(K)} - c]^+ + a_\ell - K\}$$
$$= \quad \sum_{\ell=0}^{n-1} B_\ell^{(K)}. \tag{55}$$

Using this formula and the assumptions, the proof is completed as follows:

$$b^{(K)} \quad = \quad \lim_{n\to\infty} \frac{\sum_{\ell=0}^{n-1} B_\ell^{(K)}}{\sum_{\ell=0}^{n-1} a_\ell}$$
$$\leq \quad \lim_{n\to\infty} \frac{\sum_{j=K+1}^{\infty} \sum_{\ell=0}^{n-1} 1(L_\ell^{(\infty)} = j)/n}{\frac{1}{c}\sum_{\ell=0}^{n-1} a_\ell/n} + \lim_{n\to\infty} \frac{[L_{n-1}^{(\infty)} - K - c]^+}{\sum_{\ell=0}^{n-1} a_\ell}$$
$$= \quad \frac{1}{\rho} \sum_{j=K+1}^{\infty} p_j^{(\infty)} \qquad \text{w.p.1.} \tag{56}$$

$\square$

References

[1] ITU-T Recommendation I.363: B-ISDN ATM Adaptation Layer (AAL) specification (1994).

[2] ITU-T Recommendation I.364: Support of the broadband connectionless data bearer service by the B- ISDN (1994).

[3] G. J. Armitage and K. M. Adams: Packet reassembly during cell loss. *IEEE Network*, September (1993) 26–34.

[4] G. Boiocchi, P. Crocetti, L. Fratta, M. Gerla, and M. A. Marsiglia: ATM connectionless server: performance evaluation. In H. Perros, G. Pujolle, and Y. Takahashi (eds.): *Modelling and Performance Evaluation of ATM Technology* (Elsevier Science Publishers B. V., North-Holland, 1993), 185–195.

[5] C. B. S. Traw and J. M. Smith: Striping within the network subsystem. *IEEE Network*, July/August (1995) 22–29.

[6] N. G. Duffield: Exponential bounds for queues with Markovian arrivals. *Queueing Systems*, **17** (1994) 413–430.

[7] E. Gustafsson and G. Karlsson: A literature survey on traffic dispersion. *IEEE Network*, March/April (1997) 28–36.

[8] D. P. Heyman and M. J. Sobel: *Stochastic models in operations research Vol. I* (McGraw-Hill Book Company, New York, 1982).

[9] P. Hokstad: Approximations for the $M/G/m$ queue. *Oper. Res.*, **26**(3) (1978) 510–523.

[10] S. Q. Li: Study of information loss in packet voice systems. *IEEE Trans. on Commun.*, **37**(11) (1989) 1192–1202.

[11] M. Miyazawa and G. Yamazaki: Loss probability of a burst arrival finite queue with synchronized service. *Prob. in Engineering and Information Sciences*, **6** (1992) 201–216.

[12] M. Miyazawa and H. C. Tijms: Comparison of two approximations for the loss probability in finite-buffer queues. *Prob. in Engineering and Information Sciences*, **6** (1993) 19–27.

[13] T. Ozawa and T. Asaka: Approximation analysis of packet loss probability for a shared output-buffer. *Symposium on performance models for information communication networks*, Kyoto (1996) 335–346 (in Japanese).

[14] K. Sohraby: Delay analysis of a single server queue with Poisson cluster arrival process arising in ATM networks. *GLOBCOM '89* (1989) 611–616.

[15] H. Sakasegawa, M. Miyazawa, and G. Yamazaki: Evaluating the overflow probability using the infinite queue. *Management Science*, **39**(10) (1993) 1238–1245.

[16] H. C. Tijms: Heuristics for finite-buffer queues. *Prob. in Engineering and Information Sciences*, **6** (1992) 277–285.

[17] S. Ushijima and K. Noritake: A study on virtual LAN service architecture over ATM based connectionless data networks. *Technical Report of IEICE*, **SSE94-135** (1994) 45–50 (in Japanese).

[18] W. Whitt: Comparing counting processes and queues. *Adv. Appl. Prob.*, **13** (1981) 207–220.

[19] W. Winston: Optimality of the shortest line discipline. *J. Appl. Prob.*, **14** (1977) 181–189.