Performance Analysis of Prioritized Buffer Management in Photonic Packet Switches for DiffServ Assured Forwarding

Hiroaki Harai † and Masayuki Murata ‡

† Network Architecture Group, Communications Research Laboratory
Koganei-shi, Tokyo 184-8795, Japan

‡ Cybermedia Center, Osaka University
Toyonaka-shi, Osaka 560-0043, Japan

ABSTRACT

We investigated the photonic packet switch architecture that enables a high node throughput and provides priority services. Previously, we proposed PBSO (partial buffer sharing with overwriting) method that enables prioritized buffer management to control an optical buffer [1]. The PBSO method is based on a single queue and its complexity is $O(p)$, where p is the number of priority classes. PBSO can be used to provide different levels of drop precedence in DiffServ (Differentiated Services) Assured Forwarding. In this paper, we propose an analytical method of PBSO where $p = 2$. We assess the accuracy of the analytic method. We show that PBSO improves the packet loss probability in each priority class more than the existing PBS (partial buffer sharing) does, and that it can be used for prioritized buffer management of an optical buffer. PBSO is especially effective when the arrival rate of higher priority class packets is much lower than that of lower priority class packets. In this case, PBSO dramatically improves the performance of higher priority class packets while the degradation in the performance of lower priority class packets is small. In PBSO, a larger number of higher priority class packets can be accepted at a given packet loss probability than in PBS or non-priority method.

Keywords: Photonic packet switch, Optical fiber delay line buffer, Electronic buffer management, Priority queueing, Analysis, Packet loss probability, Acceptable rate

1. INTRODUCTION

With the spread of multimedia systems such as WWW and high-speed access services such as ADSL (asymmetric digital subscriber line), the Internet must handle a greater amount of traffic than before. The spread of the access networks affects the design of the core network. In order to build a core network capable of handling a tremendous amount of traffic, it is necessary to improve the packet forwarding capability (i.e., node throughput) at a node as well as to increase the link capacity. Currently, we rely on electronic processing for packet forwarding. In this case, the node throughput can be improved with the advancements in the LSI technology according to Moore' s Law and the use of large-scale distributed/pipelined processing. While the link capacity is easily increased by bundling optical fibers, the integration and pipelined processing may limit the increase in the node throughput. As the line speed and the number of ports at the node are increasing, electrical limitations motivate us to introduce optical technology into packet forwarding.

Several researchers investigated photonic packet switches. The functions of packet switches are roughly divided into the following five groups of functions: address/label lookup (i.e., forwarding), switching, buffer management, buffering, and routing functions. To transfer very high-speed data such as 160Gbps without OEO conversion, switching and buffering must be handled in optical domain. To transfer extremely large amount of packets in a short time, electronic memory access is a bottleneck for packet forwarding. Address lookup is expected to be operated optically. Address lookup [2,3], switching [4,5], and buffering [6,5] can be handled in the optical domain. In contrast, buffer management, in which the delay in the optical buffer is determined, still requires electronic processing since optical logic and optical RAM (random access memory) are still impractical. It is therefore important to use a less complex buffer management algorithm since a complex algorithm cannot handle a large number of packets arriving simultaneously. That is, the use of a complicated algorithm may result in unexpected degradation of the photonic packet switch performance.

Further author information: (Send correspondence to H. Harai)
H. Harai: E-mail: harai@crl.go.jp, Tel: +81-42-327-5418, Fax: +81-42-327-6680

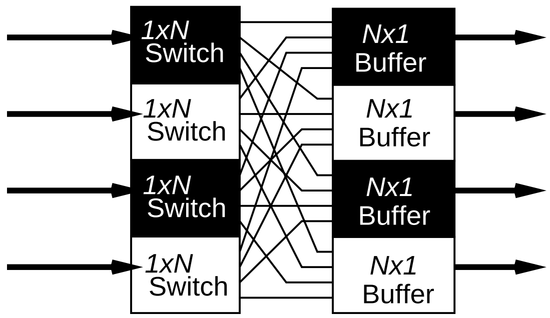

Figure 1. $N \times N$ photonic packet switch architecture ($N = 4$)

On the Internet, we can easily notice the diversification of applications as well as an increase in the amount of traffic. The diversification of applications implies the need for priority services in order to provide high-quality services rather than best-effort services. DiffServ (Differentiated Services) [7] is a typical example. Priority queueing is a practical method for DiffServ. Since advances in the LSI technology have alleviated constraints on the buffer size in the electrical domain, multiple queues such as those in CBQ (class-based queueing) [8] and DRR (deficit round robin) [9] are currently used for DiffServ. We cannot expect the use of such elegant and complicated queue management schemes in photonic packet switches because lack of optical RAM makes management of multiple queues difficult. But we can still implement the DiffServ capability based on a single queue in photonic packet switches that have optical fiber delay line buffer.

In [1], we proposed PBSO (partial buffer sharing with overwriting) method that enables prioritized buffer management in a photonic packet switch. PBSO provides different levels of drop precedence described in DiffServ Assured Forwarding [10]. It is based on a single queue, and its complexity is $O(p)$, where p is the number of priority classes. This method provides a smaller packet loss probability than PBS (partial buffer sharing) [11] does. Although HOL (head-of-the-line priority queueing) [12] and PO (push out) [11] methods provide a smaller packet loss probability than PBSO, their complexities are $O(B)$, where B is the maximum queue length. Since in PBSO, p is two or three for drop precedence in Assured Forwarding and it is much smaller than the maximum queue length, PBSO is more suitable for the management of an optical buffer than PO and HOL. PBSO itself can be applicable to more general way but it is more suitable for management of optical buffer in the photonic packet switch rather than that of electrical memory buffer.

In this paper, we propose an analytical method of PBSO. A number of buffer management methods for optical buffers have been analyzed [13–16]. For example, in [13], buffer management based on fiber delay lines and wavelength conversion is described. In [15], buffer management for variable-length packets is analyzed. Buffer management based on fiber delay lines and wavelength conversion is also described in [16]. Since variable-length packets arrive at a packet switch in [16], the wavelength selection policy is different from that in [13]. However, the above studies did not analyze prioritized buffer management. By using our analysis, we can determine the packet loss probability depending on the priority class.

This paper is organized as follows. In Section 2, we briefly summarize the photonic packet switch architecture and PBSO buffer management method. In Section 3, we describe our analytical method of PBSO. In Section 4, we assess the accuracy of the proposed analysis. In Section 5, we investigate the performance of the PBSO method and show its effectiveness. We present our conclusions and directions for future work in Section 6.

2. PHOTONIC PACKET SWITCH ARCHITECTURE

2.1. Overview of the Photonic Packet Switch

Figure 1 depicts the photonic packet switch architecture to which our management of optical buffer is applied. The $N \times N$ packet switch consists of N pieces of $1 \times N$ bufferless packet switches followed by N pieces of $N \times 1$ buffers. In each pair, a $1 \times N$ switch and an $N \times 1$ buffer are optically interconnected in a fully meshed manner. The $1 \times N$ bufferless packet switches enable fast address lookup by using photonic address lookup functions [2,3]. As a result, the architecture provides an ultra-high node throughput to the packet switch. The $N \times 1$ buffers are attached to the packet switch. Those are used to avoid packet collision and to improve the packet loss probability. Each buffer is allocated to a different output of the $N \times N$ packet switch.

The existing optical buffer consists of multiple FDLs, each of which has a different length [4–6,13–18]. Such buffers can avoid internal collision of packets with the help of buffer management that assigns each packet to a different FDL. This is similar to packet buffering in an electronic RAM buffer, which can avoid internal contention by allocating multiple data to different entries. However, these two buffers are significantly different. The electronic RAM buffer can dequeue packets at any time, while in the optical FDL buffer, the departure time of a packet is determined by the length of each FDL in the buffer. To avoid packet collision, we need a buffer management method in which the appropriate FDL is selected for each arriving packet before the packet enters the FDL buffer. Also, to obtain a priority service, we need additional buffer management, which we will describe in the following subsection.

2.2. PBSO: Partial Buffer Sharing with Overwriting

Now, let us consider a case of two priority classes ($p = 2$). In this case, a threshold is introduced. Under the PBSO strategy, the buffer manager allows every arriving packet to enter the queue when the queue is shorter than the threshold as in the PBS strategy. When the queue length is equal to or greater than the threshold, an arriving class-2 packet (low priority class packet) is allowed to enter the tail of the queue. An arriving class 1 packet (high priority class packet) is allowed to enter the queue at the next position of the class-1 packet that entered the queue most recently if the entered class-1 packet is waiting behind the threshold. Otherwise, the class-1 packet enters the queue at the threshold. In these cases, the class-1 packet may overwrite a class-2 packet. If a class-2 packet is waiting after the threshold, the class-1 packet overwrites the class-2 packet. The overwritten class-2 packet is regarded as a discarded packet. The behavior of PBSO can be represented by two variables, X_1 and X_2. In general, The behavior of PBSO can be represented by p variables, $X_1, X_2, \ldots, X_p$ and $p - 1$ thresholds are introduced.

Hereafter, we assume that a packet is of a fixed length and that each packet arrives at a photonic packet switch synchronously. A maximum of N packets arrive at the optical buffer and the buffer manager must handle all the packets within the time equivalent to the packet length, T.

Now, we will describe the behavior of packets in the PBSO strategy.

- The buffer manager handles packets from multiple ports by following round robin scheduling. It takes time T to handle all the ports.
- The buffer manager introduces $X_1, X_2 \ldots, X_p$ to represent where class-1, class-2, ..., class-p packets are stored in the queue, respectively. A packet with class-i is stored in position X_i.
- The buffer manager introduces ($p - 1$) thresholds: $B_2, B_3, \ldots, B_p$ ($B \geq B_2 \geq B_3 \geq \ldots \geq B_p$).
- When a packet arrives at an input port, the buffer manager stores the packet according to $X_1, X_2, \ldots, X_p$. However, an arriving class-1 packet is discarded if $X_1 = B$, an arriving class 2 packet is discarded if $X_2 = B$, and so on. After the packet has been handled, $X_1, X_2, \ldots, X_p$ change as follows.
  - If the queue is shorter than threshold B_p, set $X_k \leftarrow X_k + 1$ for each $k = 1, 2, \ldots, p$, regardless of the priority class of the arriving packet.
  - If the queue length is equal to or greater than threshold B_i ($1 < i \leq p$) and it is shorter than B_{i-1} and if a packet whose priority class is equal to or lower than i (i.e., classes $i, i+1, \ldots, p$) arrives, set $X_k \leftarrow X_k + 1$ for each $k = i, i+1, \ldots, p$. Here, $B_1 = B$.
  - If the queue length is equal to or greater than threshold B_i ($1 < i \leq p$) and it is shorter than B_{i-1} and if a packet whose priority class is higher than i (i.e., classes $1, 2, \ldots, i-1$) arrives, set $X_k \leftarrow X_k + 1$ for each $k = 1, 2, \ldots, i-1$. The positions for packets whose priority classes are equal to or shorter than i are then determined so that these packets are not stored ahead of packets whose class is higher than i, that is, set $X_l \leftarrow \max(X_{i-1}, X_l)$ for each $l = i, i+1, \ldots, p$. Here, $B_1 = B$.
  - If $X_k = B$ ($k = 1$ or 2 or ... or p), the corresponding variables do not change.
- After the packets of all ports have been handled, the queue length decreases by one. This is because one packet departs from the queue. At this time:
  - If the queue is equal to or shorter than threshold B_p, set $X_k \leftarrow X_k - 1$ for each $k = 1, 2, \ldots, p$.
  - Otherwise, if the queue length is equal to or greater than threshold B_i ($1 < i \leq p$) and it is shorter than B_{i-1}, set $X_k \leftarrow \max(X_k - 1, B_i)$ for each $k = 1, 2, \ldots, i-1$ and $X_l \leftarrow X_l - 1$ for each $l = i, i+1, \ldots, p$.

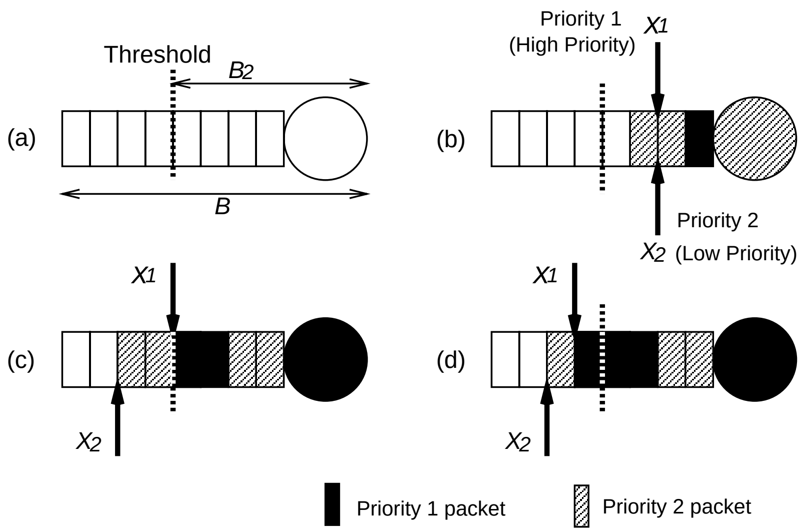

Figure 2. Packet behavior in PBSO

Figure 2 illustrates the behavior of PBSO where $p = 2$. Figure 2(a) represents the relationship between the maximum queue length (B) and the threshold value (B_2). Let us consider a queue whose length is four, that is, one packet has been served and three packets are waiting as in Fig. 2(b). Five packets (classes 1, 1, 2, 2, and 1 in the scheduled order) arrive at the packet switch and all the packets are designated to the same output port. Variables X_1 and X_2 for the new packets are as in the figure since the tail of the packets shown in Fig. 2(b) moves ahead after time T.

Now, we will describe how the five arriving packets are handled. The first two class-1 packets are stored as in Fig. 2(c) since the queue is shorter than the threshold at this time. When the next two class-2 packets are handled, the queue reaches the threshold. The class-2 packets are allowed to enter the tail of the queue in order. Figure 2(c) shows X_1 and X_2 in the transient state. The remaining class-1 packet enters the queue at the position shown in Fig. 2(d). Consequently, one class-2 packet is overwritten by one class-1 packet.

Figure 3 shows an $N \times 1$ optical buffer that supports PBSO. The buffer has two optical switches, which are controlled by a PBSO buffer manager. Allowing a class-i packet to enter positions X_i in a queue indicates that the buffer manager assigns delay X_i to the class-i packet. After the buffer manager assigns the delay, it then controls the first optical switch. The packet is transfered to a fiber delay line "+$X_i T$". When the corresponding packet reaches the second optical switch, the buffer manager also controls this optical switch. When a packet is regarded as an overwritten one, the packet is discarded by the buffer manager controlling the second optical switch. All other packets are switched to output.

3. ANALYTICAL MODEL

3.1. System Model

We develop an analytical model of PBSO buffer management where two priority classes are used ($p = 2$). We use the following notations to represent photonic packet switches and their behavior.

- N The number of input/output ports of a switch. An $N \times N$ packet switch has N input ports and N output ones (see Fig. 1).
- B The number of fiber delay lines in each output port. Each fiber delay line generates unique delay, $0, T, \ldots, (B-1)T$ (see Fig. 3).
- T The time equivalent to the length of a packet. The packet is of a fixed length.
- B_2 The threshold. When a class-2 packet is waiting behind the threshold in a queue, the packet may be overwritten by an arriving class-1 packet.
- ρ The arrival rate of packets to each input port. Packets arrive at each port according to the Bernoulli process. The packets are destined for output ports according to a uniform distribution.
- ρ_1 The ratio of class-1 packets to the total number of packets.
- ρ_2 The ratio of class-2 packets to the total number of packets. Here, $\rho_2 = 1 - \rho_1$.

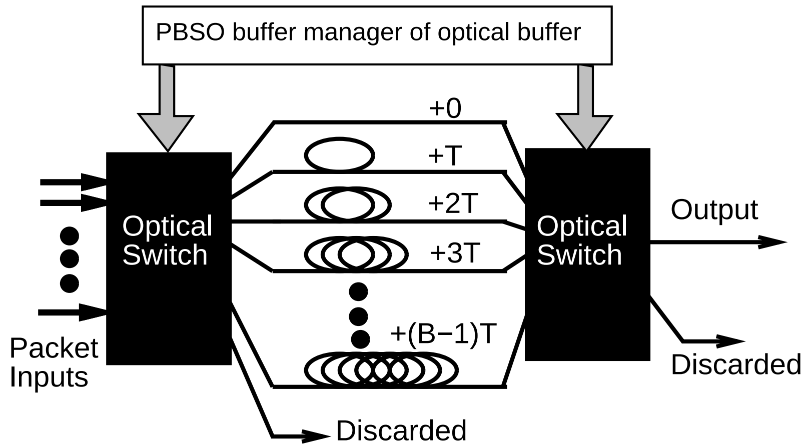

Figure 3. Space-switch-based single-stage optical buffer architecture for PBSO

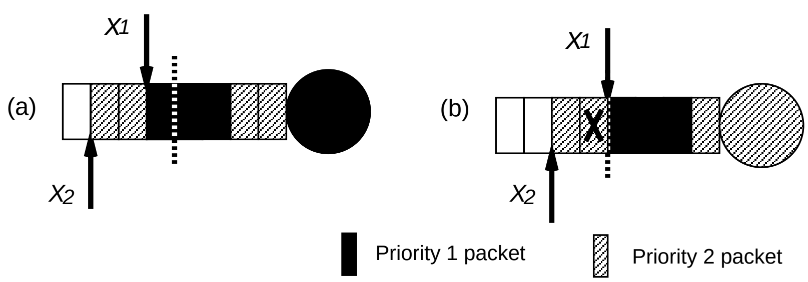

Figure 4. Packet behavior in analytical model

As described in the previous section, we assume that packets arrive at the packet switch synchronously. Focusing on one output port, we consider a discrete-time N-channel queueing system with finite buffer B. A maximum of N packets arrive at the system at the same time. To make analysis simple, we introduce the following model. When the queue length reaches the threshold, the remaining arriving packets are sorted, that is, the arriving class-1 packets enter the queue first followed by the arriving class-2 packets. Let us consider the example in Fig. 2. In PBSO, the last three packets enter the queue in the round robin order. Accordingly, the first class-2 packet is overwritten. In the analytical model, the last arriving class-1 packet enters the queue after the first two class-1 packets, followed by the two class-2 packets. Accordingly, no packet is overwritten by a packet that arrives at the same time (See Fig. 4 (a)). If a class-1 packet arrives at the system in the next slot, a class-2 packet is overwritten (See Fig. 4 (b). The packet with "x" is overwritten by the arriving class-1 packet).

3.2. Numerical Algorithm

Let X_1 and X_2 be random variables that indicate where class-1 and class-2 packets are stored in the queue, respectively (see Subsection 2.2). The corresponding steady-state probability of the queueing system is denoted as

$$q_{i_1,i_2} = P(X_1 = i_1, X_2 = i_2), \qquad i_1 = 0, \cdots, B-1, \quad i_2 = 0, \cdots, B-1. \tag{1}$$

To make the description simple, we introduce a vector consisting of steady state probabilities:

$$\boldsymbol{Q} = \{q_{00}, q_{01}, \ldots, q_{0,B-1}, q_{10}, q_{11}, \ldots, q_{1,B-1}, \ldots, q_{B-1,0}, q_{B-1,1}, \ldots, q_{B-1,B-1}\}, \tag{2}$$

The associated state transition matrix, $\boldsymbol{P}$, consisting of state transition probabilities is represented by

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{p}_{00} \\ \vdots \\ \boldsymbol{p}_{0,B-1} \\ \vdots \\ \boldsymbol{p}_{B-1,0} \\ \vdots \\ \boldsymbol{p}_{B-1,B-1} \end{bmatrix}, \tag{3}$$

where $\boldsymbol{p}_{i_1,i_2} = \{p_{i_1,i_2,0,0}, \ldots, p_{i_1,i_2,0,B-1}, \ldots, p_{i_1,i_2,B-1,0}, \ldots, p_{i_1,i_2,B-1,B-1}\}$.

We derive the packet loss probability by the following procedure.

1. Determine $\boldsymbol{P}$ from Eqs. (6) through (15) described in Subsection 3.3.
2. Determine $\boldsymbol{Q}$ by solving stationary equation $\boldsymbol{Q} \cdot \boldsymbol{P} = \boldsymbol{Q}$.
3. Determine each packet loss probability from Eqs. (16) through (19) described in Subsection 3.4.

3.3. Derivation of State Transition Matrix $\boldsymbol{P}$

To represent each element of state transition matrix $\boldsymbol{P}$, we use the following notations.

a_k probability that k packets arrive at the system in a slot.
b_{n_1,n_2} probability that n_1 class-1 packets and n_2 class-2 packets arrive at the system in a slot.

The two probabilities are given by:

$$a_k = \begin{cases} \binom{N}{k} \left(\frac{\rho}{N}\right)^k \left(1 - \frac{\rho}{N}\right)^{N-k}, & \text{if } 0 \le k \le N \\ 0, & \text{otherwise,} \end{cases} \tag{4}$$

$$b_{n_1,n_2} = \begin{cases} a_{n_1+n_2} \binom{n_1+n_2}{n_1} \rho_1^{n_1} \rho_2^{n_2}, & \text{if } 0 \le n_1 + n_2 \le N \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

In PBSO, we consider the following five different cases, each of which is dependent on the queue length and the number of arriving packets. In each case, we determine the state transition probability, p_{i_1,i_2,j_1,j_2}, in which $(X_1, X_2) = (i_1, i_2)$ moves to (j_1, j_2).

(1) The queue length does not reach the threshold before and after the transition.

(1-a) $i_1 = 0$ or $1, i_1 = i_2, j_1 = j_2 = 0$. In this case, the queue becomes or stays empty. The sum of the current queue length and the number of arriving packets must be 0 or 1.

$$p_{i_1,i_2,0,0} = \sum_{k=0}^{1-i_1} a_k. \tag{6}$$

(1-b) $0 \le i_1 \le B_2, i_1 = i_2, 0 < j_1 < B_2, j_1 = j_2$. After the transition, the queue is not empty and it is shorter than the threshold. In this case, the sum of the current queue length and the number of arriving packets is $j_1 + 1$.

$$p_{i_1,i_2,j_1,j_2} = a_{j_1-i_1+1}. \tag{7}$$

(2) The queue is equal to or longer than the threshold before and after the transition (except for case (4)).

(2-a) $B_2 \le i_1 \le i_2 < B, \max(B_2, i_1 - 1) \le j_1 \le j_2 < B - 1$. In this case, the queue stays equal to or becomes larger than the threshold and it is not full. Additionally,

- $i_2 - 1 \le j_1 < B - 1$. Arriving class-1 packets overwrite all class-2 packets waiting behind the threshold. In this case, arriving class-2 packets enter the queue at the tail of the arriving class-1 packets. When the state transits from (i_1, i_2) to (j_1, j_2), the corresponding equations are given by:

$$p_{i_1,i_2,j_1,j_2} = \begin{cases} b_{1,j_2-B_2} + b_{0,j_2+1-B_2} & \text{if } i_1 = i_2 = B_2 \\ b_{j_1-i_1+1,j_2-j_1} & \text{otherwise.} \end{cases} \tag{8}$$

- $\max(B_2, i_1 - 1) \le j_1 < i_2 - 1$. Arriving class-1 packets do not overwrite all class-2 packets waiting behind the threshold. In this case, arriving class-2 packets enter the queue at the tail of the waiting class-2 packets. When the state transits from (i_1, i_2) to (j_1, j_2), the corresponding equations are given by:

$$p_{i_1,i_2,j_1,j_2} = \begin{cases} b_{1,j_2-i_2+1} + b_{0,j_2-i_2+1} & \text{if } i_1 = j_1 = B_2 \\ b_{j_1-i_1+1,j_2-i_2+1} & \text{otherwise.} \end{cases} \tag{9}$$

(2-b) $B_2 \le i_1 \le i_2 < B, \max(B_2, i_1 - 1) \le j_1 < B - 1, j_2 = B - 1$. The queue becomes or stays full and a class-2 packet exists at the tail of the queue. Unlike in (2-a), in this case, some class-2 packets are discarded. Additionally,

- $i_2 - 1 \le j_1 \le B - 1$. Arriving class-1 packets overwrite all class-2 packets waiting behind the threshold. When the state transits from (i_1, i_2) to (j_1, j_2), the corresponding equations are given by:

$$p_{i_1,i_2,j_1,B-1} = \begin{cases} \displaystyle\sum_{n_2=B-1-B_2}^{N-1} b_{1,n_2} + \sum_{n_2=B-B_2}^{N} b_{0,n_2} & \text{if } i_1 = j_1 = B_2 \\ \displaystyle\sum_{n_2=B-1-j_1}^{N-(j_1-i_1+1)} b_{j_1-i_1+1,n_2} & \text{otherwise.} \end{cases} \tag{10}$$

- $\max(B_2, i_1 - 1) \le j_1 < i_2 - 1$. Arriving class-1 packets do not overwrite all the class-2 packets waiting behind the threshold. In this case, arriving class-2 packets are enqueued at the tail of the waiting class-2 packets. When the state transits from (i_1, i_2) to (j_1, j_2), the corresponding equations are given by:

$$p_{i_1,i_2,j_1,B-1} = \begin{cases} \displaystyle\sum_{n_2=B-i_2}^{N-(B_2-i_1+1)} b_{B_2-i_1+1,n_2} + \sum_{n_2=B-i_2}^{N} b_{0,n_2} & \text{if } i_1 = j_1 = B_2 \\ \displaystyle\sum_{n_2=B-i_2}^{N-(j_1-i_1+1)} b_{j_1-i_1+1,n_2} & \text{otherwise.} \end{cases} \tag{11}$$

(3) The queue is shorter than the threshold before the transition and the queue length is equal to or greater that the threshold after the transition (except case (4)). In this case, we take into account the number of arriving class-1 and class-2 packets handled in a slot before the threshold is reached. This is because packets stored before the threshold is reached are not sorted but those after the threshold is reached are sorted.

(3-a) $0 \le i_1 < B_2, i_1 = i_2, B_2 < j_1 \le j_2 < B - 1$. The queue is not full. When X_1 transits from i_1 to j_1, $(j_1 + 1 - B_2)$ class-1 packets are enqueued after the threshold is reached. In $j_1 = B_2$, we also have this transition when no class-1 packet is enqueued after the threshold is reached. The corresponding transition probabilities are given by:

$$p_{i_1,i_2,j_1,j_2} = \begin{cases} \displaystyle\sum_{n_1=1}^{B_2+1-i_1} b_{n_1,n_2} \frac{\binom{B_2-i}{n_1-1}\binom{n_1+n_2-(B_2-i)}{1}}{\binom{n_1+n_2}{n_1}} + \sum_{n_1=0}^{B_2-i_1} b_{n_1,n_2} \frac{\binom{B_2-i_1}{n_1}}{\binom{n_1+n_2}{n_1}} & \text{if } j_1 = B_2 \\ \displaystyle\sum_{n_1=j_1+1-B_2}^{j_1+1-i_1} b_{n_1,n_2} \frac{\binom{B_2-i}{n_1-(j_1+1-B_2)}\binom{n_1+n_2-(B_2-i)}{j_1+1-B_2}}{\binom{n_1+n_2}{n_1}} & \text{if } j_1 > B_2, \end{cases} \tag{12}$$

where

$$n_2 = j_2 - i_1 + 1 - n_1. \tag{13}$$

(3-b) $0 \le i_1 < B_2, i_1 = i_2, B_2 < j_1 < B-1, j_2 = B-1$. The queue becomes full and a class-2 packet exists at the tail of the queue. Unlike (3-a), in this case, some class-2 packets are discarded. In $j_1 = B_2$, we also have this transition when no class-1 packet is enqueued after the threshold is reached. The corresponding transition probabilities are given by:

$$p_{i_1,i_2,j_1,B-1} = \begin{cases} \displaystyle\sum_{n_1=1}^{j_1+1-i_1} \sum_{n_2=B-i_1-n_1}^{N-n_1} b_{n_1,n_2} \frac{\binom{B_2-i}{n_1-1}\binom{n_1+n_2-(B_2-i)}{1}}{\binom{n_1+n_2}{n_1}} \\ \displaystyle\quad + \sum_{n_1=0}^{j_1-i_1} \sum_{n_2=B-i_1-n_1}^{N-n_1} b_{n_1,n_2} \frac{\binom{B_2-i}{n_1}}{\binom{n_1+n_2}{n_1}} & \text{if } j_1 = B_2 \\ \displaystyle\sum_{n_1=j_1+1-B_2}^{j_1+1-i_1} \sum_{n_2=B-i_1-n_1}^{N-n_1} b_{n_1,n_2} \frac{\binom{B_2-i}{n_1-(j_1+1-B_2)}\binom{n_1+n_2-(B_2-i)}{j_1+1-B_2}}{\binom{n_1+n_2}{n_1}} & \text{if } j_1 > B_2\ . \end{cases}$$

(4) $j_1 = j_2 = B-1$. After the transition, the queue becomes or stays full and a class-1 packet is at the tail of the queue. When $X_1 = B-1$, $X_2 = B-1$. Since the sum of steady state transition probabilities in each line of matrix $\boldsymbol{P}$ is equal to 1, we have the following equation.

$$p_{i_1,i_2,B-1,B-1} = 1 - \sum_{j_1'=0}^{B-2} \sum_{j_2'=0}^{B-1} p_{i_1,i_2,j_1',j_2'}\ . \tag{14}$$

(5) Others. These conditions do not occur.

$$p_{i_1,i_2,j_1,j_2} = 0\ . \tag{15}$$

3.4. Derivation of Packet Loss Probability

We now derive the packet loss probability for each priority class. The class 1 packet loss probability, PL_1, is calculated from the probability that class-1 packets entered the system, $\overline{PL}_1$. As in determining the state-transition probability, $\boldsymbol{P}$, we consider three different cases: (1) the queue length does not reach the threshold, (2) the queue is shorter than the threshold before the transition, and the queue length is equal to or greater than the threshold after the transition, and (3) the queue length stays or becomes greater than the threshold. By calculating the expected number of class-1 packets that successfully enter the queue in a slot in each case and dividing it by the arrival rate of class-1 packets ($\rho\rho_1$), we can derive probability $\overline{PL}_1$.

$$\begin{aligned} \overline{PL}_1 &= \frac{1}{\rho\rho_1} \sum_{i=0}^{B_2-1} q_{i,i} \Bigg\{ \left(\sum_{n_1=1}^{B_2-i} \sum_{n_2=0}^{B_2-i-n_1} b_{n_1,n_2} n_1 \right) \\ &\quad + \sum_{n_2=0}^{N} \sum_{n_1=B_2-n_2-i+1}^{N-n_2} b_{n_1,n_2} \sum_{x=0}^{\min(n_1,B_2-i)} \left(\frac{\binom{B_2-i}{x}\binom{n_1+n_2-(B_2-i)}{n_1-x}}{\binom{n_1+n_2}{n_1}} \times \min(n_1, x+B-B_2) \right) \Bigg\} \\ &\quad + \frac{1}{\rho\rho_1} \sum_{i_1=B_2}^{B-1} \sum_{i_2=i_1}^{B-1} q_{i_1,i_2} \left(\sum_{n_1=1}^{N} \sum_{n_2=0}^{N-n_1} b_{n_1,n_2} \min(n_1, B-i_1) \right). \end{aligned} \tag{16}$$

As for class-2 packets, we need to take into account the existence of overwritten packets. The class-2 packet loss probability, PL_2, is calculated from the probability that class-2 packets enter the system and were not overwritten by class-1 packets, $\overline{PL}_2$. As in the derivation of $\overline{PL}_1$, we consider the three cases described above. Packets are not overwritten in (1). In (2), packets

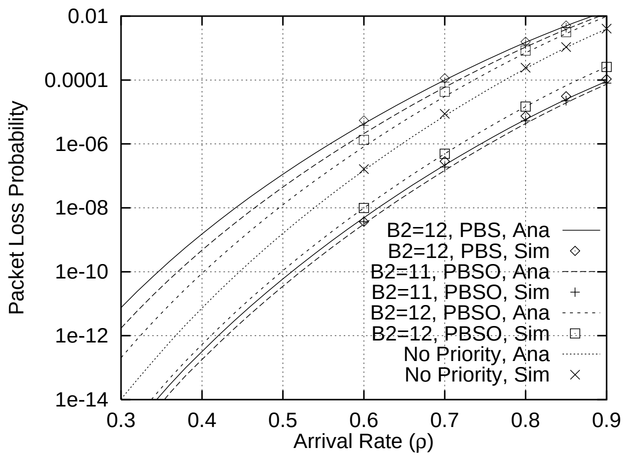

Figure 5. Comparison of analysis and simulation results ($B = 15, \rho_1 : \rho_2 = 1 : 1$)

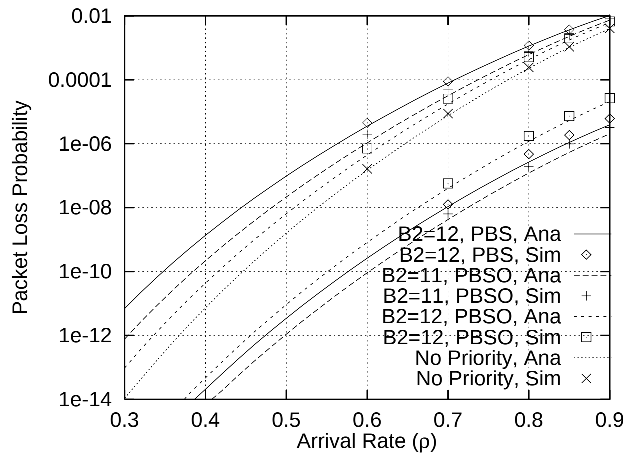

Figure 6. Comparison of analysis and simulation results ($B = 15, \rho_1 : \rho_2 = 1 : 3$)

are not overwritten but class-2 packets behind the threshold are enqueued after class-1 packets. In (3), $\min(i_2 - i_1, n_1)$ class-2 packets that have already been in the queue are overwritten by the n_1 number of arriving class-1 packets, where i_1 and i_2 represent the positions of the enqueued class-1 and class-2 packets, respectively. By dividing the expected number of entering and not overwritten packets in a slot by the arrival rate of class-2 packets ($\rho\rho_2$), we can derive probability $\overline{PL}_2$.

$$
\begin{aligned}
\overline{PL}_2 = & \frac{1}{\rho\rho_2} \sum_{i=0}^{B_2-1} q_{i,i} \Bigg\{ \left(\sum_{n_1=0}^{B_2-i} \sum_{n_2=1}^{B_2-i-n_1} b_{n_1,n_2} n_2 \right) + \sum_{n_2=1}^{N} \sum_{n_1=B_2-n_2-i+1}^{N-n_2} b_{n_1,n_2} \\
& \times \sum_{x=0}^{\min(n_2, B_2-i)} \left(\frac{\binom{B_2-i}{x}\binom{n_1+n_2-(B_2-i)}{n_2-x}}{\binom{n_1+n_2}{n_2}} \times \min\Big(n_2, \max(x, B-(n_1+i))\Big) \right) \Bigg\} \\
& + \frac{1}{\rho\rho_2} \sum_{i_1=B_2}^{B-1} \sum_{i_2=i_1}^{B-1} q_{i_1,i_2} \left(\sum_{n_1=0}^{N} \sum_{n_2=1}^{N-n_1} b_{n_1,n_2} \times \Big\{ \min\Big(n_2, B - \max(n_1+i_1, i_2)\Big) - \min(n_1, i_2 - i_1) \Big\} \right).
\end{aligned}
\tag{17}
$$

The class-1 and class-2 packet loss probabilities are finally determined by:

$$PL_1 = 1 - \overline{PL}_1, \tag{18}$$

$$PL_2 = 1 - \overline{PL}_2. \tag{19}$$

4. ASSESSMENT OF ACCURACY

We assess the accuracy of our analytical model for PBSO by using simulation. We show our results focusing on an output port of an $N \times N$ photonic packet switch. We also show analytical results obtained by using the PBS method and a best-effort method. These are based, respectively, on [11]'s analytic model and on our $B = B_2$ model. In this assessment, the number of fiber delay lines at each output port is set at $B = 15$. The number of ports is set at $N = 16$. Several threshold values are used: $B_2 = 12$ for PBS, $B_2 = 11$ and 12 for PBSO. The simulation was run at a packet generation time of 10^9.

Figures 5 and 6 present the results for different ratios of class-1 packets to class-2 packets: $1 : 1$, and $1 : 3$, respectively. The results that are higher than in the case of the non-priority queueing show the performance of class-2 packets. The lower ones show the performance of class-1 packets. The horizontal axis is the arrival rate of packets to each input port of the packet switch. The vertical axis is the packet loss probability. The lines are analytical results while the plots are simulation results. From these figures, we can see that all analytical results are in good agreement with the simulation results.

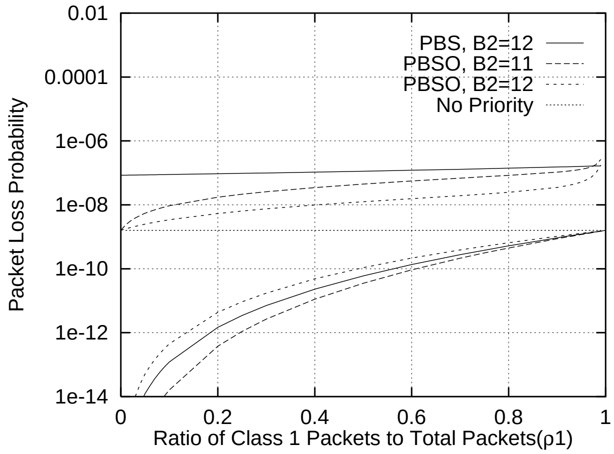

Figure 7. The packet loss probability as a function of ratio of class-1 packets to the total number of packets ρ_1 under constant arrival rate $\rho = 0.5$ ($B = 15$)

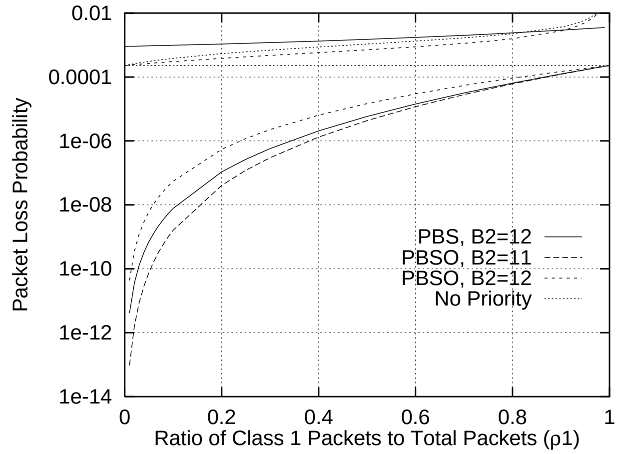

Figure 8. The packet loss probability as a function of ratio of class-1 packets to the total number of packets ρ_1 under constant arrival rate $\rho = 0.8$ ($B = 15$)

5. PERFORMANCE OF PBSO

We now investigate the performance of PBSO derived from our analysis. We first show that PBSO can provide different levels of packet forwarding based on priority classes. We then show that PBSO improves the packet loss probability of packets of both priority classes more than PBS does. We show what the acceptable rate of class-1 packets should be to meet a certain packet loss probability (e.g., 10^{-6}) when the arrival rate of class-2 packets is fixed.

Figures 5 and 6 also show the performance of PBSO. We used the same thresholds as in [1]. From these figures, we can easily see that PBSO provides different levels of packet forwarding based on priority classes and improves the class-1 packet loss probability more than non-priority queueing does. This is clearer when the ratio of class-1 packets to class-2 packets is smaller. When we use $B_2 = 11$ for PBSO and focus on Fig. 6, the class-1 packet loss probability is improved by about three orders of magnitude compared with that in the non-priority queueing.

Let's focus on the case when the PBSO's threshold is lower by 1 than the PBS's threshold (i.e., $B_2 = 11$ in PBSO and $B_2 = 12$ in PBS). The packet loss probability for packets of each priority class in PBSO is improved more than in PBS. On the other hand, when we employ the same thresholds in PBS and PBSO ($B_2 = 12$), PBSO improves the class-2 packet loss probability more than PBS does but it degrades the class-1 packet loss probability. This is true even for smaller arrival rates. We can get better performance by properly determining threshold B_2.

We now investigate the performance depending on class-1 packets to the total number of packets ratios. Figures 7 and 8 show the packet loss probability as a function of the ratio of class-1 packets to the total number of packets ρ_1. The total arrival rates are fixed at $\rho = 0.5$ and 0.8 in Figs. 7 and 8, respectively. The vertical axis is the packet loss probability in each priority class. The lines that are higher than those for non-priority queueing show the results for class-2 packets. The lower ones show the results for class-1 packets. The horizontal axis is the ratio of class-1 packets to the total number of packets, ρ_1. We find that when the ratio is small, the improvement in the performance of class-1 packets compared to that in non-priority queueing is clearer than when this ratio is large. In this case, we can see that the degradation in the performance of class-2 packets is small. In PBS, when the ratio of class-1 packets to the total number of packets is small, the performance of class-1 packets is improved at the expense of class-2 packets, whose performance degrades.

Finally, we investigate what the acceptable rate of class-1 packets should be to meet a certain packet loss probability. Figure 9 shows the acceptable rate for packet loss probabilities of 10^{-6} and 10^{-9}. The packet loss probabilities are only applied to class-1 packets. The vertical axis shows this rate. The horizontal axis shows the arrival rate of class-2 packets, $\rho\rho_2$. We find that PBSO can accept more class-1 packets than non-priority queueing. The more class-2 packets arrive, the clearer the difference in the acceptable rate between the two methods. PBSO provides better performance than PBS when the arrival rate of class-2 packets is higher.

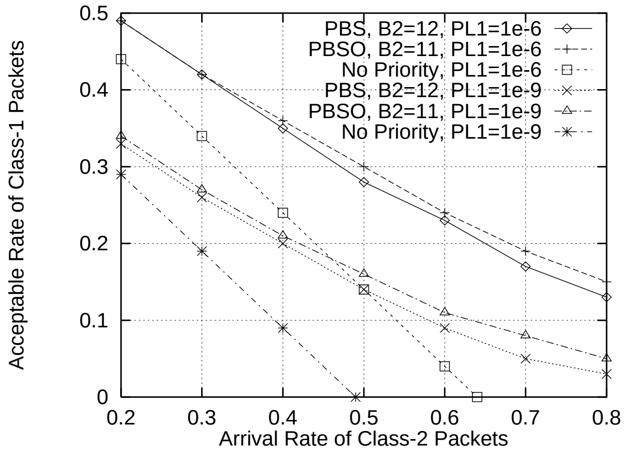

Figure 9. The acceptable rate of class-1 packets to meet a certain packet loss probability ($B = 15$)

6. CONCLUDING REMARKS

We investigated the management of an optical buffer providing priority services. PBSO can be used for optical buffer since its complexity, $O(p)$, is small, where p is the number of priority classes. Since priority-based buffer management for optical buffers has not been analyzed, we developed an analytical method of PBSO. We assessed the accuracy of our analysis by using simulation. From the analytical results, we found the following: (1) PBSO improves the packet loss probability in each priority class more than the existing PBS does; (2) When the ratio of class-1 packets (i.e., higher class packets) to the total number of packets is small, PBSO dramatically improves the performance of class-1 packets while degradation in the performance of class-2 packets is small; (3) Compared to PBS and non-priority queueing, PBSO can handle a larger number of class-1 packets at a given packet loss probability.

In this paper, we assumed that packets are of a fixed length and that they arrive at the packet switch synchronously. However, we need to examine a more general case such as one with variable-length packets and asynchronous packet arrival. This will be the topic of our future work.

REFERENCES

1. H. Harai and M. Murata, "Prioritized buffer management in photonic packet switches for DiffServ Assured Forwarding," in *The 6th IFIP Working Conference on Optical Network Design and Modelling (ONDM 2002)*, February 2002.
2. N. Wada, H. Harai, W. Chujo, and F. Kubota, "Photonic packet routing based on multi-wavelength label switching using fiber Bragg gratings," in *Proceedings of ECOC 2000 (26th European Conference on Optical Communication)*, vol. 4 (No. 10.4.6), pp. 71–72, September 2000.
3. K. Kitayama and N. Wada, "Photonic IP routing," *IEEE Photonic Technology Letters* **11**, pp. 1689–1691, December 1999.
4. D. K. Hunter and I. Andonovic, "Approaches to optical Internet packet switching," *IEEE Communications Magazine* **38**, pp. 116–122, September 2000.
5. K. Habara, H. Sanjo, H. Nishizawa, Y. Yamada, S. Hino, I. Ogawa, and Y. Suzaki, "Large-capacity photonic packet switch prototype using wavelength routing techniques," *IEICE Transactions on Communications* **E83-B**, pp. 2304–2311, October 2000.
6. D. K. Hunter, M. C. Chia, and I. Andonovic, "Buffering in optical packet switches," *IEEE/OSA Journal of Lightwave Technology* **16**, pp. 2081–2094, December 1998.
7. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for Differentiated Services," *RFC 2475* , December 1998.
8. S. Floyd and V. Jacobson, "Link-sharing and resource management models for packet networks," *IEEE/ACM Transactions on Networking* **3**, August 1995.

9. M. Shreedhar and G. Varghese, “Efficient fair queueing using deficit round robin,” *ACM Computer Communication Review* **25**, pp. 231–242, October 1995.
10. J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski, “Assured Forwarding PHB groups,” *RFC 2597* , June 1999.
11. Y. Lin and J. A. Silverster, “Priority queueing strategies and buffer allocation protocols for traffic control at an ATM integrated broadband switching system,” *IEEE Journal on Selected Areas in Communications* **9**, pp. 1524–1536, December 1991.
12. L. Kleinrock, *Queueing systems Volume II: Computer applications*, A Wiley-Interscience Publication, 1976.
13. S. L. Danielsen, B. Mikkelsen, C. Joergensen, T. Durhuus, and K. E. Stubkjaer, “WDM packet switch architectures and analysis of the influence of tunable wavelength converters on the performance,” *IEEE/OSA Journal of Lightwave Technology* **15**, pp. 219–227, February 1997.
14. J. Diao and P. L. Chu, “Analysis of partially shared buffering for WDM optical packet switching,” *IEEE/OSA Journal of Lightwave Technology* **17**, pp. 2461–2469, December 1999.
15. F. Callegati, “Optical buffers for variable length packets,” *IEEE Transactions Letters* **4**, pp. 292–294, September 2000.
16. M. Murata and K. Kitayama, “Ultrafast photonic label switch for asynchronous packets of variable length,” in *to be presented at IEEE INFOCOM 2002*, June 2002.
17. D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, A. Franzen, and I. Andonovic, “SLOB: A switch with large optical buffers for packet switching,” *IEEE/OSA Journal of Lightwave Technology* **16**, pp. 1725–1736, October 1998.
18. H. Harai, N. Wada, F. Kubota, and W. Chujo, “Contention resolution using multi-stage fiber delay line buffer in a photonic packet switch,” in *Proceedings of IEEE ICC 2002*, May 2002.