Prioritized Buffer Management in Photonic Packet Switches for Synchronously Arriving Fixed-length Packets

Hiroaki Harai† **and Masayuki Murata**‡

† Network Architecture Group, Communications Research Laboratory
4-2-1, Nukui-kita, Koganei, Tokyo 184-8795, Japan
Phone: +81-42-327-5418
E-mail: harai@crl.go.jp

‡ Cybermedia Center, Osaka University
1-32, Machikaneyama, Toyonaka, Osaka 560-0043, Japan
Phone: +81-6-6850-6615
E-mail: murata@cmc.osaka-u.ac.jp

Abstract

We investigate a photonic packet switch architecture that enables a high node throughput and provides priority services. We describe PBSO (partial buffer sharing with overwriting) method that allows control of an optical fiber-delay-line buffer by prioritized buffer management under conditions of synchronous arrival of fixed length optical packets at a packet switch. The PBSO method is based on a single queue and its complexity is $O(p)$, where p is the number of priority classes. We first present photonic buffer architectures which can support the PBSO method. We also develop an analytical method for PBSO where $p = 2$. Through analysis and simulation results, 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 actually applied to 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 that 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 other words, 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 methods.

Key words Photonic packet switch, Optical fiber delay line buffer, Electronic buffer management, Priority queueing, Analysis, Fixed-length packet

1 Introduction

With the growth of multimedia systems such as WWW and high-speed access services such as ADSL (asymmetric digital subscriber line), the Internet must handle an increasingly large amount of traffic. The growth 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 in the Internet, it is necessary to improve the node throughput (i.e., the number of packets or bits processed by a node per unit time), as well as to increase the link capacity. Currently, most solutions rely upon electronic processing to implement packet forwarding. The node throughput can be improved by advances in LSI technology known as Moore's law and by large-scale distributed or pipelined processing. Link capacity is easily increased by bundling optical fibers, but integration and pipelined processing may limit the increase in node throughput. As it is necessary to increase the line speed and the number of ports at nodes, the electrical limitations motivate us to introduce other technologies.

One promising way to increase the node throughput is to use MPLS (Multi-Protocol Label Switching) technology [1] over photonic technology. While MPLS requires the establishment of a closed domain to utilize new, lower-layer technology, it can be incorporated with photonic technology in order to build a very high-speed Internet. For this purpose, GMPLS (Generalized MPLS) [2], which can overlay a lightpath topology [3], is being developed. However, there are problems with deploying GMPLS over lightpath networks. The most difficult problem is the capacity granularity; a unit of bandwidth between edge node pairs of the GMPLS domain is a full wavelength capacity. This capacity may be too large to accommodate the traffic between the node pairs.

Photonic packet switching technology under MPLS is more promising to solve the coarse granularity problem. Photonic packet switching has an inherently finer granularity in the optical domain. The functions of packet switches are roughly divided into five groups: address/label lookup (i.e., forwarding), switching, buffer management, buffering, and routing (see Fig. 1). To transfer very high-speed data at rates such as 160Gbps without opto-electronic and electro-optic conversion, switching and buffering must be handled in the optical domain. In transferring extremely large numbers of packets in a short period of time, however, electronic memory access is a bottleneck for packet forwarding. It is thus desirable for address lookup to be done optically. Address lookup [4, 5], switching [6, 7], and buffering [8, 7] can be handled in the optical domain. Buffer management, however, still requires electronic processing for selecting an appropriate fiber delay line (FDL) in the optical buffer, because optical logic and optical RAM (random access memory) are still impractical. It is therefore important to use a less complex buffer management algorithm, because complex algorithms cannot handle a large number of packets arriving simultaneously. Thus, the use of a complex algorithm may result in unexpected degradation of photonic packet switch performance.

Diversification of applications is another evident characteristic of the Internet. Diversification implies the need for priority services rather than best-effort services to provide high-quality services. DiffServ (Differentiated Services) [9] is a typical example. Priority queueing is a practical method for DiffServ. State-of-the-art LSI technology can support multiple queues such as those for CBQ (class-based queueing) [10] and DRR (deficit round robin) [11], in which only one packet at a time is dequeued from one of multiple queues. Such solutions are currently used for DiffServ implementations. However, we cannot expect the use of such sophisticated and complicated queue management schemes in photonic packet switches because lack of optical RAM makes it difficult to manage multiple queues. Only a vital method to treat the packet queueing in the optical domain is to use FDLs, by which a simple mechanism is necessary to realize DiffServ services.

In this paper, we investigate prioritized buffer management for photonic packet switches. We assume that optical packets arrive at photonic packet switches synchronously and that they are of fixed length. IP packets are fragmented. We describe a prioritized buffer management scheme called PBSO (partial buffer sharing with overwriting) method. PBSO provides different levels of drop precedence described in DiffServ Assured Forwarding [12]. 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) [13, 14] does. Although HOL (head-of-the-line priority queueing) [15, 16] and PO (push out) [13] methods provide a smaller mean packet loss probability than PBSO, their complexities are $O(B)$, where B is the maximum queue length. Since the typical number of priority classes necessary for drop precedence in Assured Forwarding is two or three, and it is much smaller than the maximum queue length, PBSO is more suitable for the management of an optical buffer than is PO or HOL. A main contribution of this paper is to propose a realization method of PBSO in an optical domain to support DiffServ in the photonic network. PBSO can be implemented via the FDL buffer presented in this paper.

We also develop an analytical method for PBSO. A number of buffer management methods for optical buffers have been analyzed [17, 18, 19, 20]. For example, in [17], buffer management based on FDLs and wavelength conversion is described. In [19], buffer management for variable-length packets is analyzed. Buffer management based on FDLs and wavelength conversion is also described in [20]. Since variable-length packets arrive at a packet switch in [20], the wavelength selection policy is different from that described in [17]. Those studies, however, do not analyze prioritized buffer management. Our analysis 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 including buffer 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 2 shows the photonic packet switch architecture. Our buffer management is applied to the optical buffers in this architecture. The $N \times N$ packet switch consists of N $1 \times N$ bufferless packet switches followed by N $N \times 1$ 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 [4, 5]. The architecture thus provides ultra-high node throughput to the packet switch. The $N \times 1$ buffers attached to the packet switch are used to avoid packet collision and to reduce 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 [6, 7, 8, 17, 18, 19, 20, 21, 22]. 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, but 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 provide a priority service, we need additional buffer management, which is described in the following subsection.

2.2 PBSO: Partial Buffer Sharing with Overwriting

Now, consider a case of two priority classes ($p = 2$). Under the PBSO method, the buffer manager allows every arriving packet to enter the queue when the queue length is shorter than some predefined threshold, as in a PBS method [13]. When the queue length is equal to or greater than the threshold, an arriving class-2 packet (low-priority packet) is allowed to enter the tail of the queue. An arriving class-1 packet (high-priority packet) is allowed to enter the queue at the position behind the class-1 packet that entered the queue most recently if the class-1 packet that is already in the queue 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, which is then regarded as a discarded packet.

In a general case, that is, for p priority classes, the behavior of PBSO can be described by p variables, $X_1, X_2, \ldots, X_p$, which represents positions at which packets enter the buffer, and $p-1$ thresholds $B_2, B_3, \ldots, B_p$. The thresholds are used in order to realize prioritized buffer management for p priority classes. As will be described in more detail later, the position at which an arriving packet is allowed to enter the buffer depends on the priority class that the packet belongs to, and is determined by using the queue length and the threshold values. Hereafter, we assume that the packets have a fixed length and that the packets arrive 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.

The buffer manager has the following information for the PBSO method.

- Variables $X_1, X_2 \ldots, X_p$, which respectively represent positions at which arriving class-1, class-2, ..., class-p packets are allowed to enter the queue. An arriving packet that belongs to class-i is stored in position X_i.
- $p-1$ thresholds: $B_2, B_3, \ldots, B_p$. For descriptive simplicity, we introduce the notation B_1. The condition $B = B_1 \geq B_2 \geq B_3 \geq \ldots \geq B_p$ must be satisfied.
- Queue length, which is identical to X_p.

We next describe the PBSO algorithm as follows.

(1) Initialization. Variable X_k is set to 0 for each k and thresholds are set appropriately.

(2) In each time slot, the following procedure is executed.

(2-1) For $n = 1, 2, \ldots, N$, the following procedure is executed according to round robin fashion.

(2-1-1) If a packet arrives at input port n and the packet belongs to class-i, the buffer manager allow the packet to enter the buffer at position X_i. However, an arriving class-i packet is discarded if $X_i = B$. If the packet enter the buffer, the next step (2-1-2) is executed. Otherwise, the next step is skipped.

(2-1-2) After the packet has been handled, $X_1, X_2, \ldots, X_p$ are changed as follows.

- If the queue length (X_p) 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_j ($1 < j \leq p$) and it is shorter than B_{j-1} and if priority class i of the arriving packet is less than or equal to j (i.e., $i \geq j$), then set $X_k \leftarrow X_k + 1$ for each $k = j, j+1, \ldots, p$.
- If the queue length is equal to or greater than threshold B_j ($1 < j \leq p$) and it is shorter than B_{j-1} and if priority class i of the arriving packet is greater than j (i.e., $i < j$), then set $X_k \leftarrow X_k + 1$ for

each $k = 1, 2, \ldots, j-1$. The positions for packets whose priority classes are j or lower are then determined so that these packets are not stored ahead of packets whose class is higher than j; that is, set $X_l \leftarrow \max(X_{j-1}, X_l)$ for each $l = j, j+1, \ldots, p$.

- If $X_k = B$ ($k = 1, 2, \ldots, p$), the corresponding variables are not changed.

(2-2) After all packets arriving at 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 length is equal to or shorter than threshold B_p, set $X_k \leftarrow \max(X_k - 1, 0)$ for each $k = 1, 2, \ldots, p$.
- Otherwise, if the queue length is a value between two adjacent thresholds B_i and B_{i-1} ($1 < i \leq p$), then 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 $l = i, i+1, \ldots, p$.

The buffer manager for PBSO is simple. First, it does not require that the number of packets stored in the buffer for each priority class. It also does not require the position of the packets in the buffer for each priority class, while PO [13] does. It only requires p variables, that is, $X_1, X_2, \ldots, X_p$. The complexity of handling a packet is $O(p)$, and depends on the number of renewals of the variables $\{X_1, X_2, \ldots, X_p\}$ in Step (2-1-2). The actual processing time can be minimized by applying parallel processing. An attractive feature of PBSO is that for each priority class, the arrival order of packets at each port is identical to the departure order, except for discarded packets. The consistency of the arrival and departure orders improves performance for higher layer applications (e.g., TCP).

Figure 3 illustrates the behavior of PBSO where the number of priority classes $p = 2$, the buffer length $B = 9$, and the threshold value $B_2 = 5$. Figure 3(a) represents the relationship between the maximum queue length (B) and the threshold value (B_2). Consider a queue whose length is four, where one packet has been served and three packets are waiting as in Fig. 3(b). Variables X_1 and X_2 indicate 4, that is, arriving packets are allowed to enter fourth position of the buffer regardless of the priority class. Assume that five packets (classes 1, 1, 2, 2, and 1 in the scheduled order) arrive at the packet switch in the next slot and all the packets are designated to the same output port. At the end of the time slot, X_1 and X_2 are changed to 3 as shown in Fig. 3 (c) according to Step (2-2) in the PBSO algorithm.

The five arriving packets are handled in the following way. The first two class-1 packets are stored as in Fig. 3(d), because the queue is shorter than the threshold at this time. After the class-1 packets have been handled, the variables are changed to $X_1 = X_2 = 5$ and the queue reaches the threshold. The next class-2 packets are then allowed to enter the tail of the queue in order. Figure 3(d) shows the queue status after the first four packets have been handled. X_2 is changed to 7 as shown in the figure, because only X_2 is changed according to Step (2-1-2). Note that only variable X_2 is changed, because the queue reaches threshold. The remaining class-1 packet enters the queue at the position shown in Fig. 3(e). Consequently, one class-2 packet is overwritten by one class-1 packet.

2.3 Photonic Buffer Architecture

An $N \times 1$ optical buffer that supports PBSO is shown in Fig. 4. The buffer has two optical switches of which switching times are on the sub-nanosecond level, such as $LiNbO_3$ switches. The switches are controlled by a PBSO buffer manager. Allowing a class-i packet to enter position X_i in a queue means that the buffer manager assigns delay $X_i T$ to the class-i packet. After the buffer manager assigns the delay, it then controls the first optical switch. The packet is transferred to a fiber delay line d_{X_i}. When the corresponding packet reaches the second optical switch, the buffer manager controls this optical switch. When a packet is regarded as an overwritten one, the packet is discarded by the buffer manager that is controlling the second optical switch. All other packets are switched to output.

PBSO can be implemented with broadcast-and-select-based optical buffer as well as the above-described space-switch-based optical buffer, as in Fig. 4. These buffers are capable of PBS as well as PBSO, and are described in [23].

3 Analysis

In this section, we develop an approximation analysis of PBSO buffer management in which two priority classes are used ($p = 2$). We first describe a system model for our analysis for deriving packet loss probability, and then present equations for calculating the packet loss probability for each priority class. A detailed description is found in Ref. [24].

3.1 System Model

Assume that packets arrive at the packet switch synchronously, we focus on one output port of a $N \times N$ photonic packet switch and 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 simplify the analysis, we introduce the following model. When the queue length reaches the threshold, the scheduled orders of the remaining arriving packets are sorted; that is, the arriving class-1 packets enter the queue first and are followed by the arriving class-2 packets. Recall the example in Fig. 3, in which five packets arrive in a slot. In PBSO, the last three packets, which were handled after the queue reached the threshold, entered the queue in round-robin order. Accordingly, the first class-2 packet was overwritten. See Fig. 3 (e). On the other hand, Fig. 5 (a) shows the packets stored according to the analytical model. In the analytical model, the last arriving class-1 packet enters the queue after the first two class-1 packets, and is followed by the two class-2 packets. Accordingly, no packet is overwritten by a packet that arrives at the same time. However, class-2 packets behind the threshold are overwritten if class-1 packets arrive at the system after the next slot. Assume that a class-1 packet arrives in the next slot. Packets in the buffer move ahead and variables X_1 and X_2 are changed as shown in Fig. 5 (b). In this case, the class-2 packet with mark “x” is overwritten by the arriving class-1 packet.

3.2 Derivation of Packet Loss Probability

We use the following notation to present equations for calculating the packet loss probability for each priority class.

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.

r The total arrival rate of packets to each input port. Packets arrive at each port according to a Bernoulli process. The packets are destined for output ports according to a uniform distribution.

r_1 The arrival rate of class-1 packets to each input port.

r_2 The arrival rate of class-2 packets to each input port.

q_{i_1,i_2} Steady-state probability that $X_1 = i_1$ and $X_2 = i_2$.

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 class-1 packet loss probability, PL_1, is calculated from the probability that class-1 packets have entered the system, $\overline{PL}_1$. We consider three different cases: (1) the queue length does not reach the threshold, (2) the queue length is shorter than the threshold before the transition, but the queue length is equal to or greater than the threshold after the transition, and (3) the queue length remains 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 (r_1), we can derive probability $\overline{PL}_1$.

$$\begin{aligned}\overline{PL}_1 = & \frac{1}{r_1}\sum_{i=0}^{B_2-1} q_{i,i}\left\{\left(\sum_{n_1=1}^{B_2-i}\sum_{n_2=0}^{B_2-i-n_1} b_{n_1,n_2} n_1\right)\right. \\ & \left.+\sum_{n_2=0}^{N}\sum_{n_1=B_2-n_2-i+1}^{N-n_2} b_{n_1,n_2}\sum_{x=0}^{\min(n_1,B_2-i)}\left(\frac{\binom{B_2-i}{x}\binom{n_1+n_2-(B_2-i)}{n_1-x}}{\binom{n_1+n_2}{n_1}}\times\min(n_1,x+B-B_2)\right)\right\} \\ & +\frac{1}{r_1}\sum_{i_1=B_2}^{B-1}\sum_{i_2=i_1}^{B-1} q_{i_1,i_2}\left(\sum_{n_1=1}^{N}\sum_{n_2=0}^{N-n_1} b_{n_1,n_2}\min(n_1,B-i_1)\right).\end{aligned} \tag{1}$$

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 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 packets that enter and are not overwritten in a slot by the arrival rate of class-2 packets (r_2), we can derive probability $\overline{PL}_2$.

$$\begin{aligned}\overline{PL}_2 = & \frac{1}{r_2}\sum_{i=0}^{B_2-1} q_{i,i}\left\{\left(\sum_{n_1=0}^{B_2-i}\sum_{n_2=1}^{B_2-i-n_1} b_{n_1,n_2} n_2\right)+\sum_{n_2=1}^{N}\sum_{n_1=B_2-n_2-i+1}^{N-n_2} b_{n_1,n_2}\right. \\ & \left.\times\sum_{x=0}^{\min(n_2,B_2-i)}\left(\frac{\binom{B_2-i}{x}\binom{n_1+n_2-(B_2-i)}{n_2-x}}{\binom{n_1+n_2}{n_2}}\times\min\Big(n_2,\max(x,B-(n_1+i))\Big)\right)\right\}\end{aligned} \tag{2}$$

$$+ \frac{1}{r_2} \sum_{i_1=B_2}^{B-1} \sum_{i_2=i_1}^{B-1} q_{i_1,i_2} \left(\sum_{n_1=0}^{N} \sum_{n_2=1}^{N-n_1} b_{n_1,n_2} \times \left\{ \min\Big(n_2, B - \max(n_1+i_1, i_2)\Big) - \min(n_1, i_2 - i_1) \right\} \right).$$

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

$$PL_1 = 1 - \overline{PL}_1, \qquad PL_2 = 1 - \overline{PL}_2. \tag{3}$$

4 Assessment of Accuracy

We assess the accuracy of our analytical model for PBSO by simulation. The presentation of our results focuses on an output port of an $N \times N$ photonic packet switch. The analytical results obtained by using the PBS method and by a best-effort method (i.e., non-priority queueing) are shown. These are based, respectively, on analytical model developed in [13] and on our analytical model for the case of $B_2 = B$. 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 6 and 7 present the results for different ratios of class-1 packets to class-2 packets: $1 : 1$, and $1 : 3$, respectively. The values in parentheses represent threshold B_2. The results above the case of non-priority queueing represent the performance for class-2 packets, while those below represent the performance for class-1 packets. The horizontal axis is the arrival rate of packets for each input port of the packet switch (i.e., $r = r_1 + r_2$). The vertical axis is the packet loss probability. The lines are analytical results while the symbols show simulation results. From these figures, we can see that all analytical results are in good agreement with the simulation results.

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 6 and 7 also show the performance of PBSO. We selected two threshold values for PBSO ($B_2 = 11$ and 12); one is the same value as that for PBS and the other is lower by 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 much more clear in the case where the ratio of class-1 packets to class-2 packets is smaller. When we use $B_2 = 11$ for PBSO and focus on Fig. 7, the class-1 packet loss probability is improved by about three orders of magnitude over the non-priority queueing.

Let us focus on the difference between PBS and PBSO, when the same threshold, $B_2 = 12$. Figure 7 indicates that PBSO has higher performance than PBS for class-2 packets. On the other hand, performance of PBSO is inferior to that of PBS for class-1 packets. This is explained as follows. In contrast to the PBS method, PBSO allows class-2 packets to enter the buffer even if the queue length is equal to or longer than the threshold. Some packets are not overwritten if class-1 packets do not arrive before class-2 packets go ahead the threshold. This probability, which is labelled "Non-overwritten" in Fig. 8, is much larger than the probability that packets are overwritten (labelled "Overwritten"). This indicates that PBSO performs better than PBS for class-2 packets. On the other hand, once class-2 packets that are not overwritten go ahead of the threshold, they cannot be overwritten by class-1 packets. If many class-1 packets arrive at the buffer at that time, the packets are likely to be discarded more easily than that in the PBS method. Thus, the performance of PBSO for class-1 packets is not better than that of PBS for the same threshold condition. PBSO does, however, improve the total packet loss probability, as shown in Fig. 9.

Performance of PBSO for class-1 packets can be improved by proper setting of the threshold B_2. We focus on the case when the PBSO threshold is lower than the PBS threshold by 1 in Fig. 7 (i.e., $B_2 = 11$ in PBSO and $B_2 = 12$ in PBS). The packet loss probability for class-1 packets in PBSO is lower than that for PBS. The packet loss probability for class-2 packets and the mean packet loss probability in PBSO are also better than in PBS.

We find similar characteristics in other conditions. Figure 10 shows the performance of PBSO for other thresholds for the case of $B = 15$. The results above the case of non-priority queueing represent the performance for class-2 packets, while those below represent the performance for class-1 packets. When we set the threshold for PBSO to $B_2 = 10$, which is the same as that for PBS in the figure, the performance of PBSO for class-1 packets is not better than that of PBS. With the threshold set to $B_2 = 9$, which is lower by 1 than that for PBS, PBSO provides better performance for class-1 packets than PBS does. Moreover, in $B_2 = 8$, the difference in packet loss probability for class-1 packets is larger than one order of magnitude. In spite of such a large difference, the performance of PBSO for class-2 packets is almost the same as that of PBS. Figure 11 shows performance of PBSO for the case of $B = 20$. By setting the threshold for PBSO to a value lower by 1 or 2 than that for PBS, PBSO provides better performance for class-1 packets than PBS does.

We now investigate the dependence of performance on the ratio of class-1 packets to the total number of packets. Figures 12 and 13 show the packet loss probability as a function of the ratio of class-1 packets to the total number of packets r_1/r. The total arrival rates are fixed at $r = 0.5$ and 0.8 in Figs. 12 and 13, respectively. The vertical axis is the packet loss probability in each priority class. The lines above the case of non-priority queueing represent the performance for class-2 packets, while those below represent the performance for class-1 packets. The horizontal axis is the ratio of class-1 packets to the total number of packets, r_1/r. We find that when the ratio is small, the improvement in the performance for 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 for class-2 packets is small. In PBS, when the ratio of class-1 packets to the total number of packets is small, the performance for class-1 packets is improved at the expense of class-2 packets, for which the performance degrades. We find similar characteristics in other conditions such as $B = 20$, as shown in Figs. 14 and 15.

Finally, we investigate what the acceptable rate of class-1 packets should be to meet a certain packet loss probability. Figure 16 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, r_2. We find that PBSO can accept more class-1 packets than can non-priority queueing. The more class-2 packets that arrive, the clearer becomes the difference in the acceptable rate between the two methods. PBSO performs better than PBS when the arrival rate of class-2 packets is higher.

6 Concluding Remarks

We investigated optical buffer management for prioritized packet services for the case of fixed-length packets with synchronous arrival. The PBSO scheme presented achieves $O(p)$ complexity, where p is the number of priority classes, a moderate overhead. We presented photonic buffer architectures which can support the PBSO method. Since priority-based buffer management for optical buffers has not been analyzed, we developed an analytical method for PBSO. We assessed the accuracy of our analysis by simulation. From the analytical results, we found that 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 for class-1 packets while degradation in the performance for class-2 packets is small; and 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 fixed length packets arrive at the packet switch synchronously. However, the more general case of variable-length packets and asynchronous arrivals also needs to be examined. We are currently extending this work to consider the key provisions, see [25].

References

[1] E. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol label switching architecture," *RFC 3031*, January 2001.

[2] E. Mannie *et al.*, "Generalized multi-protocol label switching (GMPLS) architecture (draft-ietf-ccamp-gmpls-architecture-03.txt)," *IETF Internet Draft (Work in Progress)*, August 2002.

[3] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: An approach to high bandwidth optical WAN's," *IEEE Transactions on Communications*, vol. 40, pp. 1171–1182, July 1992.

[4] N. Wada, H. Harai, W. Chujo, and F. Kubota, "Photonic packet routing based on multi-wavelength label switching using fiber Bragg gratings," *ECOC 2000 Technical Digest (26th European Conference on Optical Communication)*, vol. 4 (No. 10.4.6), pp. 71–72, September 2000.

[5] K. Kitayama and N. Wada, "Photonic IP routing," *IEEE Photonic Technology Letters*, vol. 11, pp. 1689–1691, December 1999.

[6] D. K. Hunter and I. Andonovic, "Approaches to optical Internet packet switching," *IEEE Communications Magazine*, vol. 38, pp. 116–122, September 2000.

[7] 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*, vol. E83-B, pp. 2304–2311, October 2000.

[8] D. K. Hunter, M. C. Chia, and I. Andonovic, "Buffering in optical packet switches," *IEEE/OSA Journal of Lightwave Technology*, vol. 16, pp. 2081–2094, December 1998.

[9] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, “An architecture for Differentiated Services,” *RFC 2475*, December 1998.

[10] S. Floyd and V. Jacobson, “Link-sharing and resource management models for packet networks,” *IEEE/ACM Transactions on Networking*, vol. 3, August 1995.

[11] M. Shreedhar and G. Varghese, “Efficient fair queueing using deficit round robin,” *ACM Computer Communication Review*, vol. 25, pp. 231–242, October 1995.

[12] J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski, “Assured Forwarding PHB groups,” *RFC 2597*, June 1999.

[13] 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*, vol. 9, pp. 1524–1536, December 1991.

[14] F. Callegati, G. Corazza, and C. Raffaelli, “Exploitation of DWDM for optical packet switching with quality of service guarantees,” *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 190–201, January 2002.

[15] L. Kleinrock, *Queueing systems Volume II: Computer applications*. A Wiley-Interscience Publication, 1976.

[16] C. Develder, M. Pickavet, and P. Demeester, “Assessment of packet loss for an optical packet router with recirculating buffer,” *The 6th IFIP Working Conference on Optical Network Design and Modelling (ONDM 2002)*, February 2002.

[17] 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*, vol. 15, pp. 219–227, February 1997.

[18] J. Diao and P. L. Chu, “Analysis of partially shared buffering for WDM optical packet switching,” *IEEE/OSA Journal of Lightwave Technology*, vol. 17, pp. 2461–2469, December 1999.

[19] F. Callegati, “Optical buffers for variable length packets,” *IEEE Transactions Letters*, vol. 4, pp. 292–294, September 2000.

[20] M. Murata and K. Kitayama, “Ultrafast photonic label switch for asynchronous packets of variable length,” *Proceedings of IEEE INFOCOM 2002*, pp. 371–380, June 2002.

[21] 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*, vol. 16, pp. 1725–1736, October 1998.

[22] H. Harai, N. Wada, F. Kubota, and W. Chujo, “Contention resolution using multi-stage fiber delay line buffer in a photonic packet switch,” *Proceedings of IEEE ICC 2002*, pp. 2843–2847, May 2002.

[23] H. Harai and M. Murata, “Prioritized buffer management in photonic packet switches for DiffServ Assured Forwarding,” *The 6th IFIP Working Conference on Optical Network Design and Modelling (ONDM 2002)*, pp. 231–246, February 2002.

[24] H. Harai and M. Murata, “Performance analysis of prioritized buffer management in photonic packet switches for DiffServ Assured Forwarding,” *Proceedings of SPIE vol. 4874 (OptiComm 2002)*, pp. 298–309, July 2002.

[25] H. Harai and M. Murata, “Photonic buffer architecture to support prioritized buffer management for asynchronously arriving variable-length packets,” *Proceedings of ONDM 2003 (The 7th IFIP Working Conference on Optical Network Design and Modelling)*, pp. 1103–1118, February 2003.

Author Biography

Hiroaki Harai

Organization: Network Architecture Group, Communications Research Laboratory

Position: Senior Researcher

Address: 4-2-1, Nukui-kita, Koganei, Tokyo 184-8795, Japan

E-mail: harai@crl.go.jp

Phone: +81-42-327-5418

Hiroaki Harai received the M.E. and Ph.D. degrees in information and computer sciences from Osaka University, Osaka, Japan, in 1995 and 1998, respectively. In April 1998, he joined Communications Research Laboratory, Tokyo, Japan. He is currently a Senior Researcher of Communications Research Laboratory. His current research topic is to develop photonic networks, especially photonic packet-switched networks. His research interest also includes routing and wavelength assignment problem and bandwidth management in IP over WDM networks. He is a member of IEEE and IEICE.

Masayuki Murata

Organization: Cybermedia Center, Osaka University

Position: Professor

Address: 1-32, Machikaneyama, Toyonaka, Osaka 560-0043, Japan

E-mail: murata@cmc.osaka-u.ac.jp

Phone: +81-6-6850-6615

Masayuki Murata received the M.E. and D.E. degrees in Information and Computer Sciences from Osaka University, Japan, in 1984 and 1988, respectively. In April 1984, he joined Tokyo Research Laboratory, IBM Japan, as a Researcher. From September 1987 to January 1989, he was an Assistant Professor with Computation Center, Osaka University. In February 1989, he moved to the Department of Information and Computer Sciences, Faculty of Engineering Science, Osaka University. From 1992 to 1999, he was an Associate Professor in the Graduate School of Engineering Science, Osaka University, and from April 1999, he has been a Professor of Osaka University. He moved to Advanced Networked Environment Research Division, Cybermedia Center, Osaka University in April 2000. He has more than two hundred papers of international and domestic journals and conferences. His research interests include computer communication networks, performance modeling and evaluation. He is a member of IEEE, ACM, The Internet Society, IEICE and IPSJ.

List of Figures

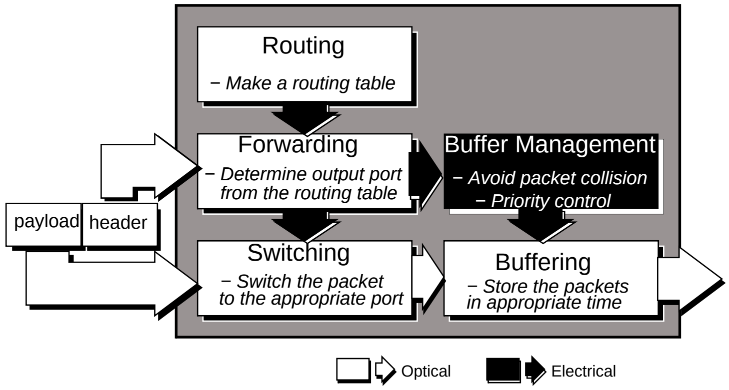

Figure 1: Internal functions in a packet switch

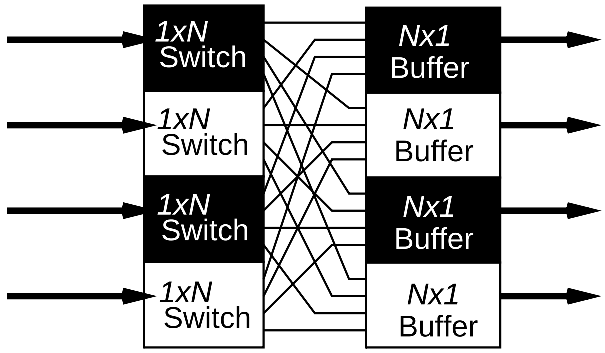

Figure 2: $N \times N$ photonic packet switch architecture ($N = 4$)

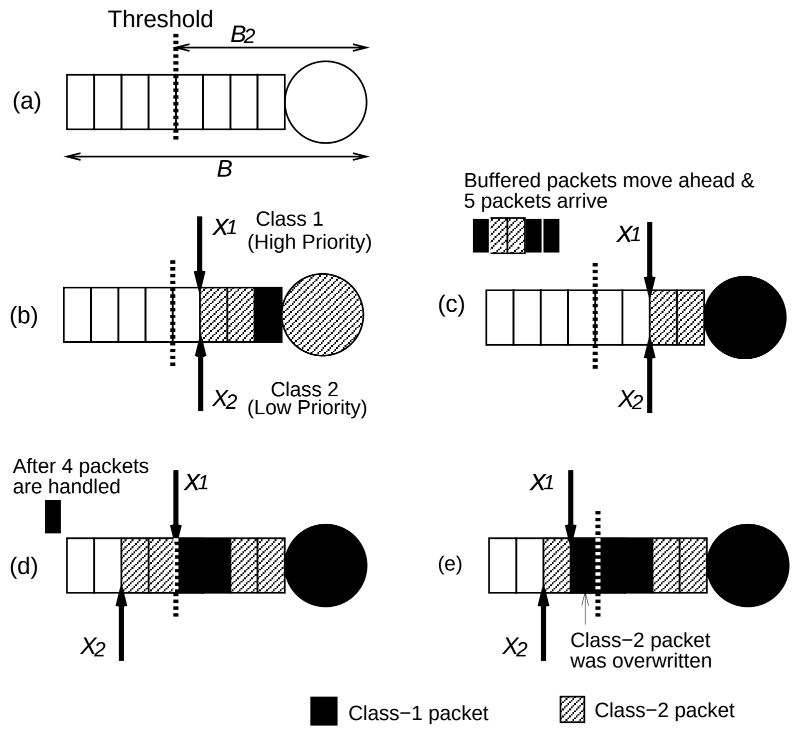

Figure 3: Packet behavior in PBSO

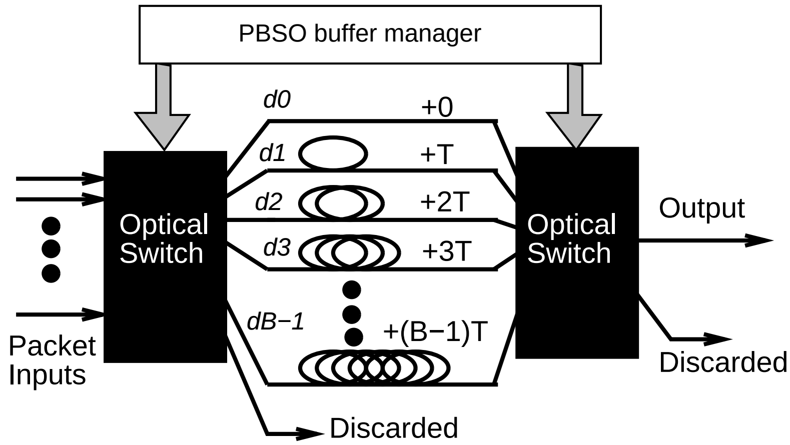

Figure 4: Space-switch-based single-stage optical buffer architecture for PBSO

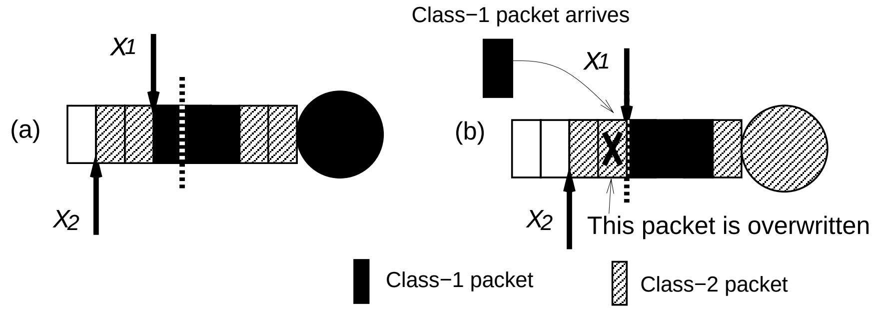

Figure 5: Packet behavior in analytical model

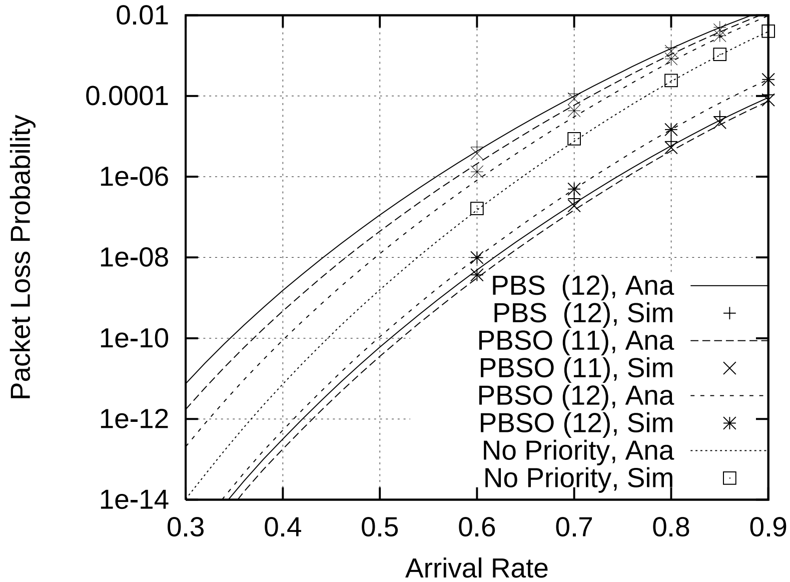

Figure 6: Comparison of analysis and simulation results ($B = 15, r_1 : r_2 = 1 : 1$)

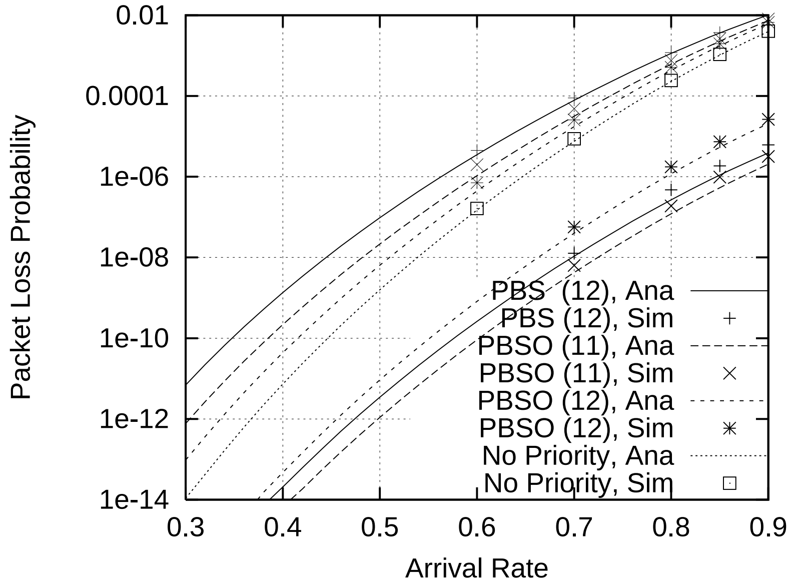

Figure 7: Comparison of analysis and simulation results ($B = 15, r_1 : r_2 = 1 : 3$)

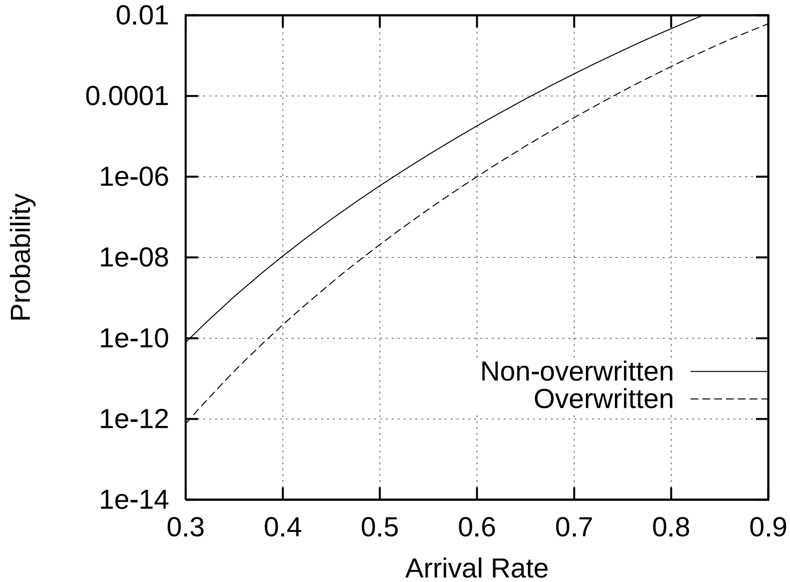

Figure 8: The probabilities that class-2 packets are overwritten and that class-2 packets that have entered behind threshold are not overwritten ($B = 15, B_2 = 11, r_1 : r_2 = 1 : 3$)

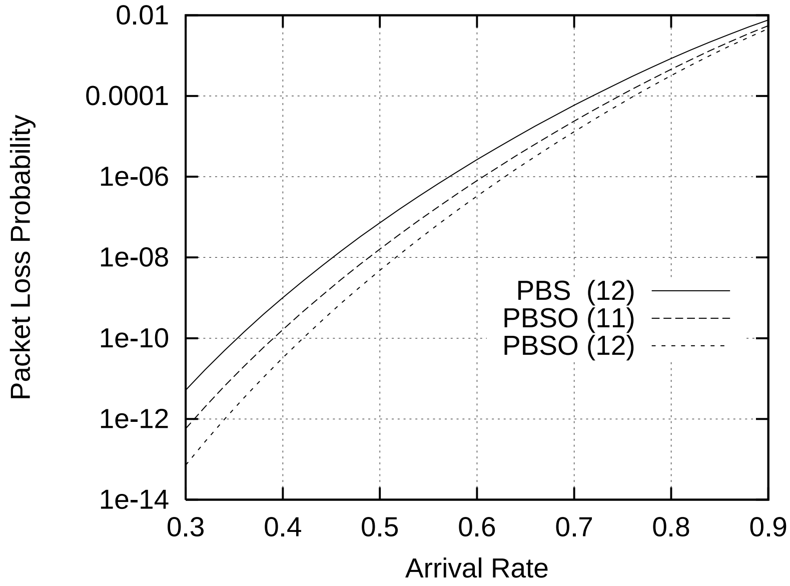

Figure 9: Mean packet loss probability ($B = 15, r_1 : r_2 = 1 : 3$)

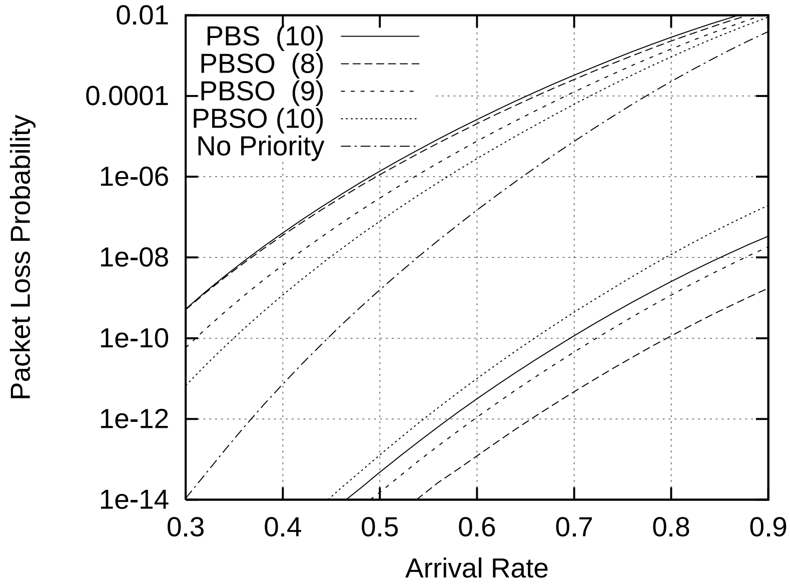

Figure 10: Packet loss probability in each class ($B = 15, r_1 : r_2 = 1 : 3$)

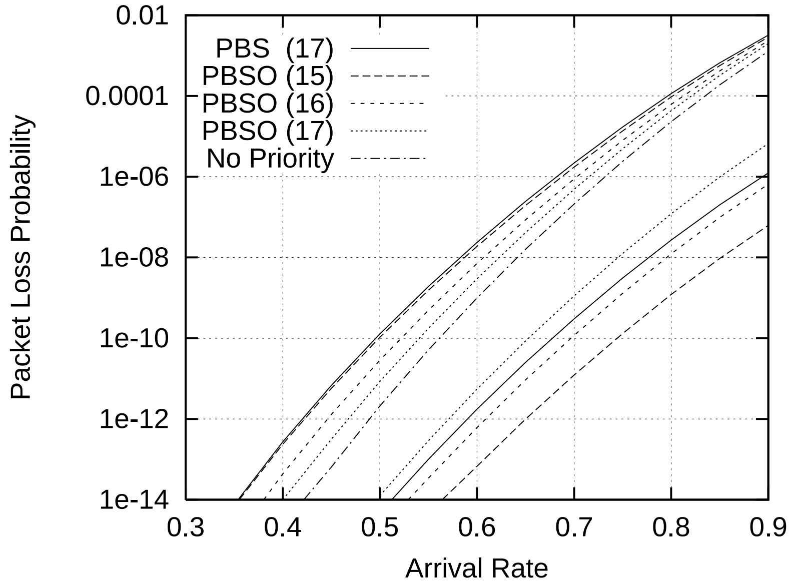

Figure 11: Packet loss probability in each class ($B = 20, r_1 : r_2 = 1 : 3$)

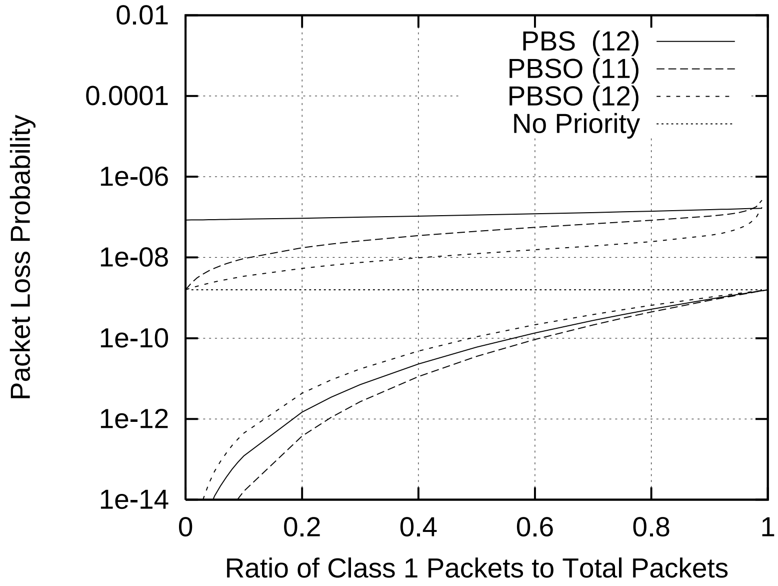

Figure 12: The packet loss probability as a function of ratio of class-1 packets to the total number of packets r_1/r under constant arrival rate $r = 0.5$ ($B = 15$)

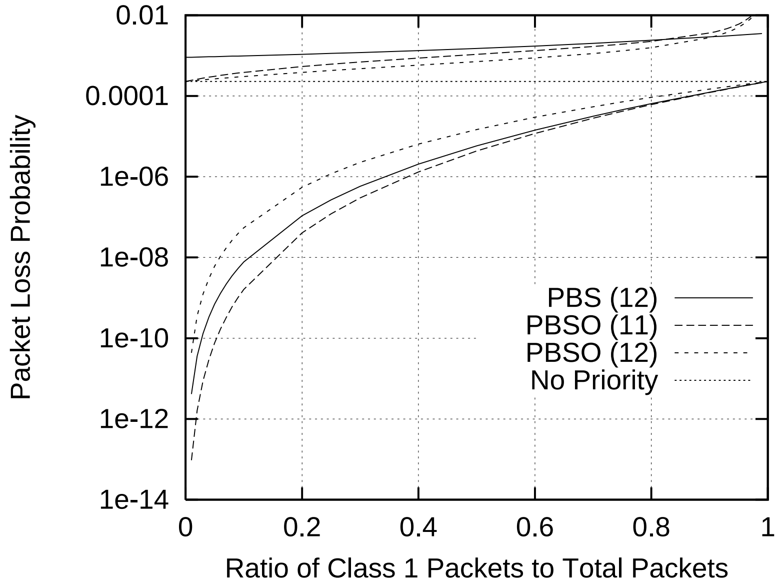

Figure 13: The packet loss probability as a function of ratio of class-1 packets to the total number of packets r_1/r under constant arrival rate $r = 0.8$ ($B = 15$)

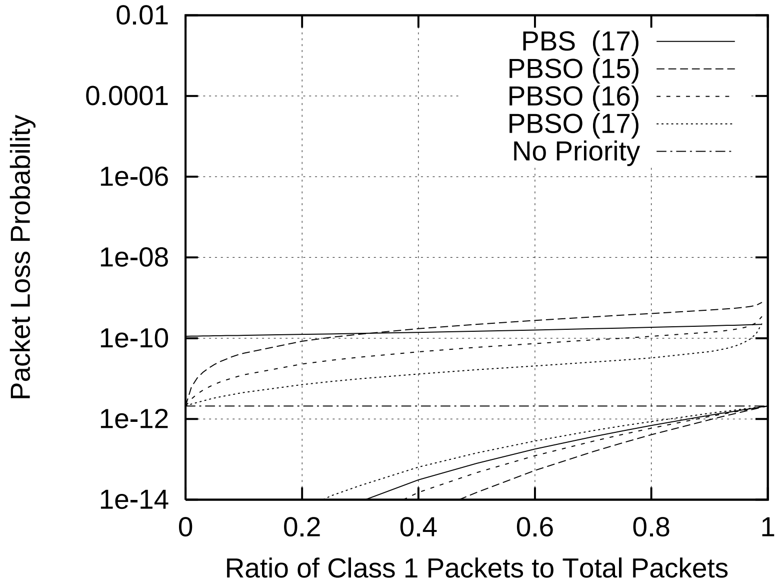

Figure 14: The packet loss probability as a function of ratio of class-1 packets to the total number of packets r_1/r under constant arrival rate $r = 0.5$ ($B = 20$)

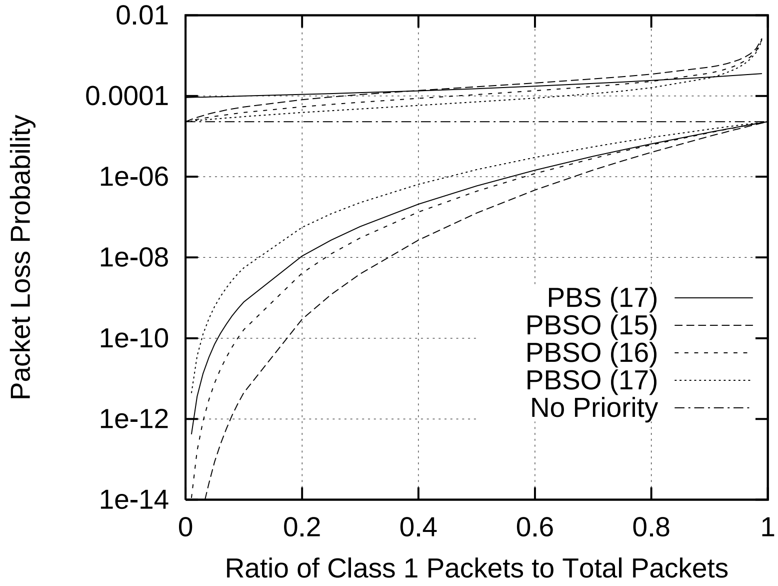

Figure 15: The packet loss probability as a function of ratio of class-1 packets to the total number of packets r_1/r under constant arrival rate $r = 0.8$ ($B = 20$)

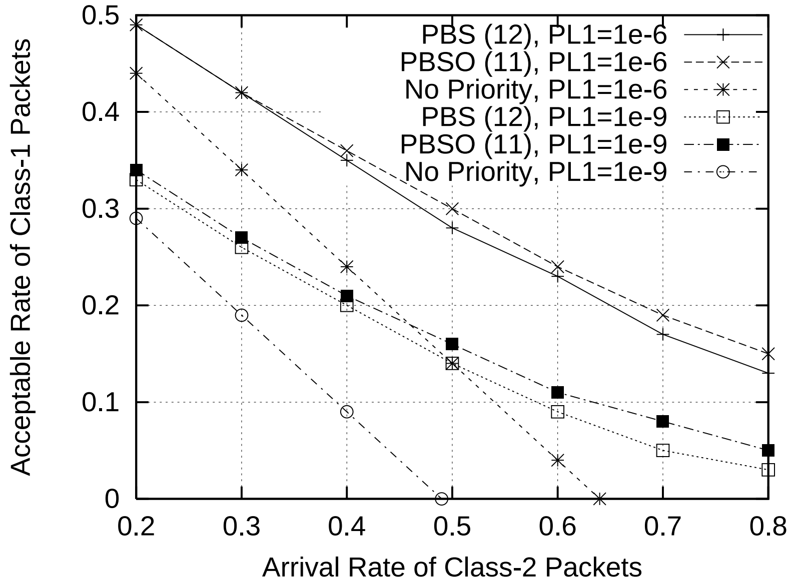

Figure 16: The acceptable rate of class-1 packets to meet a certain packet loss probability ($B = 15$)