

# SUPPORTING MULTIPLE CLASSES OF SERVICES IN IP OVER WDM NETWORKS \*

Myungsik Yoo † and Chunming Qiao ‡

† Department of Electrical Engineering ‡ Department of Computer Science and Engineering  
University at Buffalo (SUNY), Buffalo, New York 14260

† myoo@eng.buffalo.edu ‡ qiao@computer.org

## Abstract

In this paper, we consider the next generation Optical Internet, where direct communication channels between IP entities are provided through an all-optical WDM layer. Since current IP provides only best effort service, supporting Quality of Service (QoS) has become a crucial issue. We propose a simple, yet efficient priority scheme to support basic QoS at the WDM layer of the Optical Internet. Unlike existing approaches that depend on buffers (queues) and scheduling algorithms (e.g., Fair Queuing), our scheme does not mandate any buffer in the WDM layer in order to differentiate the services. The proposed scheme is based on the concept of using an offset time between a control (set-up) packet and the corresponding data burst as in optical burst switching (OBS) [1]. More specifically, by assigning different offset times to different classes, it is possible for a higher priority class to be isolated from lower priority classes. We quantify the offset time required for class isolation, and analyze the blocking probability of each class.

## 1 Introduction

For the past couple of decades, the Internet has been a dominant communication infrastructure for transporting data traffic through Internet Protocol (IP), which provides *best effort* service when delivering variable length packets. In recent years, the amount of Internet traffic is doubling every 3 to 6 months, and seems to keep growing exponentially. Such an unprecedented explosion of the traffic demand calls for the next generation Internet using Wavelength Division Multiplexing (WDM), which can provide huge deliverable bandwidth.

\*This research is sponsored in part by a grant from NSF under contract number ANIR-9801778

IP over ATM [2,3] has been considered as one of the possible ways of building the next generation Internet mainly because ATM has well-defined Quality of Service (QoS) support. However, using ATM to transport IP incurs a high ATM *cell tax* because of the ATM header [4] and complex ATM signaling protocols. An alternative would be IP over SONET/SDH [4] which provides fast restoration in the event of a fiber cut or failure. However, the cost of SONET/SDH equipment at a bit data rate at/above 10 Gbps is high, and in addition, the restoration capability is inherent in the Internet. Therefore, IP over WDM (or so-called the *Optical Internet*) seems to be an efficient and economical way to implement the future Internet since the overhead of SONET/SDH and ATM can be eliminated. In the next generation Optical Internet, direct communication channels between IP entities can be provided by an all-optical WDM layer.

Given that current IP provides only *best effort* service, supporting QoS becomes a crucial research issue in the Optical Internet since the future integrated Internet will demand differentiated services. For example, applications such as Internet telephony and video conferencing require a higher QoS than electronic mail and general web browsing.

Up to date, most literatures [5-7] have proposed to implement QoS using buffers (queues) and scheduling algorithms (as in IP over ATM) such as fair queuing (FQ), weighted fair queuing (WFQ), frame-based fair queuing (FFQ), and etc. Such scheduling algorithms introduce a high processing overhead and/or hardware cost. In addition, in order for existing QoS mechanisms to differentiate services, all intermediate nodes should have a certain amount of buffer, which is often partitioned into small portions, one for each priority. However, the use of electronic buffer necessitates O/E and E/O conversions, thus sacrificing the data transparency. On the other hand, no optical buffer (RAM) is

available and the use of fiber-delay lines (FDLs), which can provide a limited delay, should also be avoided as much as possible in the WDM layer of the all-Optical Internet.

The proposed QoS scheme to be described mainly differs from existing ones in that it does not mandate any buffer (or FDL). Instead, it extends the concept of a "base" offset time introduced in [1] to support QoS. Specifically, each burst is assigned different "extra" offset time according to its class. When a burst is generated at a source node, its corresponding control packet is sent to reserve the bandwidth (wavelength) using delayed reservation (DR) (see Figure 1). The data burst, after waiting at the source for the offset time (base plus extra), follows the path that has been set up by the control packet without waiting for an acknowledgement from its destination. Since the extra offset time, which can provide an end-to-end priority (as to be discussed later), is determined at the source node, end-to-end QoS can be easily accomplished.

Previously, we have conducted a preliminary study of a two-class scheme [8]. In this paper, we propose and evaluate a model which can support an arbitrary number of classes. The paper is organized as follows. Section 2 describes the proposed priority scheme for QoS. Section 3 presents the analytic model for the proposed scheme. Section 4 shows the numerical results from analysis and simulation, followed by the conclusion in Section 5.

## 2 The Proposed Priority Scheme

In order to describe the proposed priority scheme in a simple way, we assume that there is no buffers available at any switch (i.e., a bufferless WDM network), and first consider two classes of service: namely class 0 and class 1, where class 1 has priority over class 0. In Section 3, the number of classes will be extended to  $n$ , where  $n \geq 2$ . We also assume that the link has only one wavelength for data (and an additional wavelength for control). In the proposed priority scheme, to give class 1 a higher priority for wavelength reservation, an extra offset time, denoted by  $t_o^1$ , is given to class 1 traffic (but not to class 0, i.e.,  $t_o^0 = 0$ ). For simplicity, we assume that the base offset time is negligible as compared to the extra offset time, and will refer to the latter as simply the offset time hereafter.

Figure 1 illustrates why a class 1 request that is assigned  $t_o^1$  obtains a higher priority for reservation than a class 0 request. For  $i = 0, 1$ , let  $t_a^i$  and  $t_s^i$  be the arrival time and the service-start time for a class  $i$  request

denoted by  $req(i)$ , respectively, and let  $l_i$  be the burst length requested by  $req(i)$ .

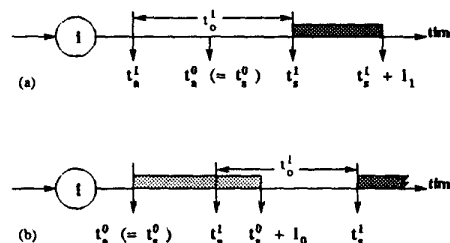


Figure 1: Offset time for service differentiation

Consider the following two situations where contentions among two classes of traffic are possible. In the first case as illustrated in Figure 1(a),  $req(1)$  comes first and reserves wavelength using DR, and  $req(0)$  comes afterwards. Clearly,  $req(1)$  will succeed, but  $req(0)$  will be blocked if  $t_a^0 < t_s^1$  but  $t_a^0 + l_0 > t_s^1$ , or if  $t_s^1 < t_a^0 < t_s^1 + l_1$ . In the second case,  $req(0)$  arrives, followed by  $req(1)$  as shown in Figure 1(b). When  $t_a^1 < t_s^0 + l_0$ ,  $req(1)$  would be blocked had no offset time been assigned to  $req(1)$  (i.e.,  $t_o^1 = 0$ ). However, such a blocking can be avoided by using a large enough offset time so that  $t_s^1 = t_a^1 + t_o^1 > t_s^0 + l_0$ . Given that  $t_a^1$  may equal to  $t_a^0 + \sigma$ , where  $\sigma > 0$  can be very small,  $t_o^1$  needs to be larger than the maximum burst length over all bursts in class 0 in order for  $req(1)$  to completely avoid being blocked by  $req(0)$ . With that much offset time, the blocking probability of (the bursts in) class 1 becomes only a function of the offered load belonging to class 1, that is, independent of the offered load belonging to class 0. On the other hand, the blocking probability of class 0 will be determined by the offered load belonging to both classes. In the next Section, we quantify the offset time required to achieve a given degree of class isolation, and analyze the blocking probability of each class when the number of classes to be supported is an arbitrary integer  $n \geq 2$ .

## 3 Analysis

To simplify our analysis, let us assume that the burst length is exponentially distributed with an average of  $L_{i-1}$  for a given class  $i-1$ . Then, the percentage of the bursts in class  $i-1$  whose length is no longer than a given value,  $t$ , is obtained by the following probability distribution function (PDF)

$$PDF = 1 - e^{-\mu_{i-1} \cdot t} \quad (1)$$

Note that Equation 1 gives the minimum probability for a given class  $i$ , denoted by  $R_{i,i-1}$ , that  $req(i)$  will not be blocked by  $req(i-1)$ . For example, let  $t_{i,j}$  denote the difference in the offset times assigned to classes  $i$  and  $j$ , where  $i > j$ . If  $t_{i,i-1} = t_o^i - t_o^{i-1}$  is set to  $L_{i-1}$ , then  $R_{i,i-1} = 0.6321$  since the probability that the length of a burst in class  $i-1$  is shorter than  $L_{i-1}$  (and thus  $t$ ) is  $1 - e^{-1} = 0.6321$ .

We may rewrite Equation 1 by replacing PDF with  $R_{i,i-1}$ , and burst length  $t$  with  $t_{i,i-1}$  as follows,

$$t_{i,i-1} = t_o^i - t_o^{i-1} = -\frac{\ln(1 - R_{i,i-1})}{\mu_{i-1}} = -\ln(1 - R_{i,i-1}) \cdot L_{i-1} \quad (2)$$

Table 2 shows some of the corresponding values of  $t_{i,i-1}$  and  $R_{i,i-1}$ . Note that  $R = 1$  can be achieved only with  $t = \infty$  (or  $t_o^i = \infty$ ). However, as can be seen in Table 1, when  $t_{i,i-1} = 5 \cdot L_{i-1}$ ,  $R_{i,i-1}$  is over 99%, which means that class  $i$  is virtually free of blocking by class  $i-1$ . In fact, even when  $t_{i,i-1} = 3 \cdot L_{i-1}$ ,  $R_{i,i-1} = 0.95$  which is still close to 1. In addition, if class  $i-1$  is also free of blocking by class  $i-2$  with  $R_{i-1,i-2} = 0.95$  (i.e., when  $t_{i-1,i-2} = 3 \cdot L_{i-2}$ ), then class  $i$  is virtually free of blocking by class  $i-2$  because  $t_{i,i-2} = 3 \cdot L_{i-1} + 3 \cdot L_{i-2}$  and when  $L_{i-1} = L_{i-2}$ , we have  $R_{i,i-2} > 0.99$ . Similarly, class  $i$  is virtually free of blocking by all other lower priority classes such as  $i-3, i-4, \dots, 0$ .

$t_{i,i-1}$	$0.4 \cdot L_{i-1}$	$L_{i-1}$	$3 \cdot L_{i-1}$	$5 \cdot L_{i-1}$
$R_{i,i-1}$	0.3296	0.6321	0.9502	0.9932

Table 1: Offset time and R

We now analyze the blocking probability assuming a complete class isolation (i.e.,  $R_{i,i-1} = 1.0$ ). In this analysis, we consider two cases, namely the *classless* and the *prioritized*. In the classless case, no offset time will be assigned to any class (i.e., no priority as in current IP's best effort service), whereas in the prioritized case, a longer offset time will be assigned to a higher priority class. In order to simplify our analysis, we consider a bufferless switch with a single output link, and assume that there are  $n$  classes, each of which generates traffic with the exponentially distributed arrival rate and service rate. Let  $\lambda_i$  be the average burst arrival rate and  $\mu_i = 1/L_i$  be the average service rate in class  $i$ , where  $L_i$  is the average burst length in class  $i$ . Accordingly, the offered load of class  $i$  will be given by  $\rho_i = \lambda_i / (\mu_i \cdot k) = r_i / k$ , where  $r_i = \lambda_i / \mu_i$ , and the total offered load will be  $\rho = \sum_{i=0}^{n-1} \rho_i$ .

Assuming that there are  $k$  wavelengths on the output link, and the switch is capable of wavelength conversion, the blocking probability in the classless case can be calculated using the following *Erlang's loss formula* (M/M/k/k), where  $r = \rho \cdot k$ .

$$B(k, \rho) = \frac{1/k! \cdot r^k}{\sum_{m=0}^k 1/m! \cdot r^m} \quad (3)$$

Without loss of generosity, assume that class  $i$  has a higher priority over class  $j$  if  $i > j$ , where  $0 \leq i, j \leq n-1$ . The priority between classes can be accomplished by assigning class  $i$  an offset time of  $t_o^i$ , where  $t_o^i > t_o^{i-1}$  for  $i = n-1, n-2, \dots, 1$ , and  $t_o^0 = 0$ .

We now determine the blocking probabilities of each class, denoted by  $pb_i$ , where  $i = 0, 1, \dots, n-1$ , in the prioritized case. Let  $\rho_{n-1,j}$  be the partial sum of offered load and  $PB_{n-1,j}$  be the partial overall blocking probability from class  $n-1$  through class  $j$ , respectively, then we have  $\rho_{n-1,j} = \sum_{i=j}^{n-1} \rho_i$  and  $PB_{n-1,j} = B(\rho_{n-1,j}, k)$ . Assume that the *conservation law* holds, that is, the overall performance (i.e., blocking probability and throughput) stays the same regardless of the number of classes and the priority assigned to them, then we have  $PB_{n-1,0} = B(\rho, k)$ . In addition, when  $c_i = \rho_i / \rho$ , which is the ratio of the offered load in class  $i$  over the total offered load, the overall blocking probability can be obtained by summing the blocking probabilities of each class, that is,  $PB_{n-1,0} = \sum_{i=0}^{n-1} c_i \cdot pb_i$ .

To obtain each  $pb_i$ , we start with the highest class,  $n-1$ , and move down to the lowest class, 0. Since class  $n-1$  has the highest priority, and is completely isolated from any other classes with  $R_{n-1,n-2} = 1.0$ ,  $pb_{n-1} = B(\rho_{n-1}, k)$ . According to the conservation law, we have  $PB_{n-1,n-2} = B(\rho_{n-1,n-2}, k)$ , and from the class blocking probabilities,  $PB_{n-1,n-2} = \sum_{i=n-2}^{n-1} c_i \cdot pb_i$ . By equating the two, we have  $pb_{n-2} = \frac{B(\rho_{n-1,n-2}, k) - c_{n-1} \cdot pb_{n-1}}{c_{n-2}}$ . This procedure can be applied recursively, and accordingly, a general blocking probability for a given class  $j$ , except for class  $n-1$ , is

$$pb_j = \frac{B(\rho_{n-1,j}, k) - \sum_{i=j+1}^{n-1} c_i \cdot pb_i}{c_j} \quad (4)$$

## 4 Results and Discussion

In this section, numerical results from both simulation and analysis of a bufferless single-node system are presented. These results are obtained assuming that all sources have equal average arrival rate ( $\lambda_0 = \lambda_1 = \dots = \lambda_{n-1} = \lambda/n$ ) and average service rate

( $\mu_0 = \mu_1 = \dots = \mu_{n-1} = \mu/n$ ), i.e., generate equal amount of offered load ( $\rho_0 = \rho_1 = \dots = \rho_{n-1} = \rho/n$ ).

Figure 2 and Figure 3 show the blocking probabilities as a function of the offered load,  $\rho$ , when the number of classes,  $n$ , is 4 and the number wavelengths,  $k$ , is 8. For simulation, class  $i$  is assigned the offset time of  $i \cdot 3 \cdot L$  (i.e.,  $t_o^i = i \cdot t_o^1$ , where  $t_o^1 = 3 \cdot L$ ), where  $i = 0, \dots, 3$  and  $L$  is the average burst (or packet) length of each class. In Figure 2, the dotted and solid lines show the analysis and simulation results, respectively. As can be seen, analysis and simulation results are closely matched. The service differentiation can be observed in this Figure in that the blocking probability decreases as the priority increases.

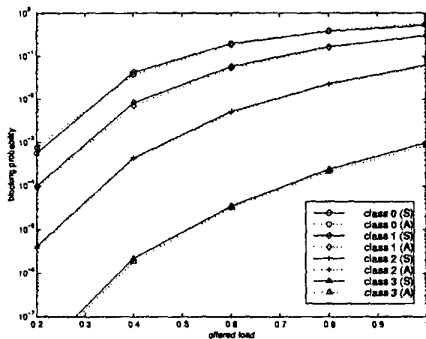


Figure 2: Class blocking probability of class 0 ~ 3

Figure 3 shows the overall blocking probabilities as  $t_o^1$  (and  $t_o^i$ ) changes. It is shown that the overall blocking probabilities stay the same regardless of the  $t_o^i$ 's, which shows that the assumed conservation law holds. Hence, one can conclude that the proposed priority scheme effectively achieves service differentiation while the overall performance remains the same as that of the classless case. It is also interesting to observe the bounds of the blocking probabilities of each class by comparing both Figure 2 and Figure 3. When  $t_o^i = 0$ , it becomes the classless case, where all classes share the high blocking probability (Figure 3) because all bursts equally compete for the resource. As  $t_o^i$  increases, blocking probabilities of each class move to the pattern as in Figure 2, where low priority classes move upwards, while high priority classes move downwards.

Figure 4 shows the effect of the number of wavelengths,  $k$ , when the offered load,  $\rho$ , is 0.8 and the number of classes,  $n$ , is 4. The blocking probability of the classless case is  $10^{-2}$  ( $10^{-3}$ ) even when  $k = 64$  ( $k = 128$ ). However, in the prioritized case,  $pb_1, pb_2, pb_3$  have much lower value than the classless case. As an extreme,  $pb_3$  is as low as  $10^{-25}$  ( $10^{-48}$ ) when  $k = 64$

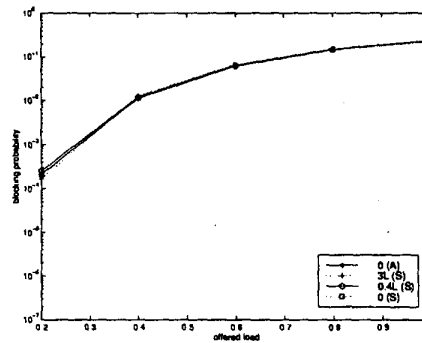


Figure 3: Overall blocking probability

( $k = 128$ ), which is more than twenty (forty) orders of magnitude reduction when compared to the classless case. As  $k$  increases, the blocking probability of higher priority classes decreases rapidly. Note that service differentiation costs only small increase in  $pb_0$ , which is very close to the blocking probability of the classless case.

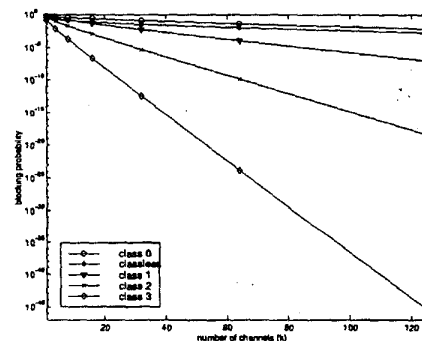


Figure 4: Blocking probability vs. number of channels

Figure 5 plots the blocking probabilities as a function of the number of classes,  $n$ , when the number of wavelengths,  $k$ , is 16 and the offered load,  $\rho$ , is 0.8. The horizontal line shows blocking probability in the classless case, while the top and bottom curves show  $pb_0$  and  $pb_{n-1}$ , respectively. Under the conservation law, the overall blocking probability of the prioritized case is the same as that of the classless case, which remains the same (around  $10^{-1}$ ) regardless of  $n$ . However, as  $n$  increases,  $pb_{n-1}$  decreases rapidly ( $10^{-12}$  when  $n = 10$ ). This is because even though the overall offered load,  $\rho$ , remains the same, the effective offered load of class  $n - 1$ ,  $\rho_{n-1}$ , decreases when  $n$  increases. This means that higher priority classes receive better service when more classes exist.

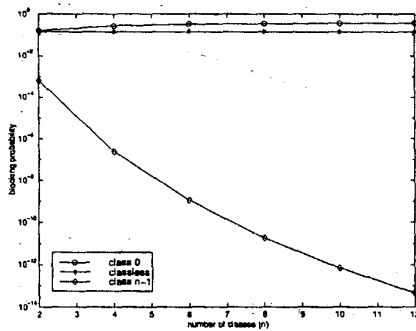


Figure 5: Blocking probability vs. number of classes

Finally, for delay-sensitive applications, it is imperative to discuss the end-to-end delay. We first note that, the use of a base offset time does not increase the end-to-end delay of a burst because the offset time merely substitutes for the total processing delay to be encountered by the corresponding control packet (see [1]). As compared to circuit switching, the end-to-end delay of the proposed scheme is  $2P$  shorter, where  $P$  is the end-to-end propagation delay (typically tens of msec coast-to-coast).

To assess the impact of using an extra offset time, let the total processing delay be  $\Delta$  (typically tens of  $\mu\text{sec}$ ) and the average burst length be  $l$  (typically a few  $\mu\text{sec}$  or a few kilobytes at 10 Gbps). In addition, assume that in a real-time application, each byte (or bit) generated at the source has to reach its destination in  $D$  msec, which is often large compared to  $\Delta$  or even  $P$ . For example, for today's voice and video communications, it may be acceptable for  $D$  to be as large as a few hundreds of msec. Accordingly, if the extra offset time is equal to a few tens times of the average burst length, the increase in the end-to-end delay may not be significant (for example, when  $n = 10$  and  $R_{n-1, n-2} = 0.95$ ,  $t_{n-1, n-2} = 9 \cdot 3 \cdot l$ , which is only a few tens to hundreds of  $\mu\text{sec}$ ), although the reduction in the blocking probability of higher priority classes can be significant.

## 5 Conclusion

In this paper, a novel priority scheme has been proposed, which can efficiently support differentiated services in the next generation Optical Internet. Unlike the existing priority schemes, the proposed scheme does not mandate the use of any buffer at the intermediate nodes, and hence is especially suitable for the WDM layer as the technology for optical buffer (e.g., fiber de-

lay lines or FDLs) is not as mature as the electronic counterpart. We have compared the performance of the prioritized case with that of the classless case. It has been shown that in the prioritized case, a service differentiation can be achieved by simply taking advantage of the extra offset time, while the overall performance remains the same as in the classless case. As a future work, we will extend the performance evaluation work from a single node system to a network to identify the delay characteristics of high priority classes, and determine the network resources (e.g., the number of wavelengths and amount of FDLs) needed to meet the delay and loss requirements of certain applications.

## References

- [1] C. Qiao and M. Yoo, "Optical burst switching (OBS) - a new paradigm for an Optical Internet," *Journal of High Speed Network*, vol. 8, no. 1, pp. 69-84, 1999.
- [2] M. Laubach, "Classical IP and ARP over ATM," *RFC 1577*, Jan. 1994.
- [3] P. White, "ATM switching and IP routing intergration: The next stage in Internet evolution," *IEEE Communications Magazine*, vol. 36, pp. 79-83, Apr. 1998.
- [4] J. Manchester et al., "IP over SONET," *IEEE Communications Magazine*, vol. 36, pp. 136-142, May 1998.
- [5] H. Zhang, "Service disciplines for guaranteed performance service in packet-switching networks," *Proceedings of the IEEE*, vol. 83, pp. 1374-1396, Oct. 1995.
- [6] A. Varma and D. Stiliadis, "Hardware implementation of fair queuing algorithms for asynchronous transfer mode networks," *IEEE Communications Magazine*, vol. 35, pp. 74-80, Dec. 1997.
- [7] U. Briem et al., "Traffic management for an ATM switch with per-VC queuing: Concept and implementation," *IEEE Communications Magazine*, vol. 36, pp. 88-93, Jan. 1998.
- [8] M. Yoo and C. Qiao, "A new optical burst switching (OBS) protocol for supporting quality of service," in *SPIE Proceedings, All optical Communication Systems: Architecture, Control and Network Issues*, vol. 3531, pp. 396-405, Nov. 1998.