

Service Differentiation in Optical Burst Switched Networks

Chi-Hong Loi, Wanjiun Liao*, and De-Nian Yang

Department of Electrical Engineering
National Taiwan University
Taipei, Taiwan

Abstract—This paper proposes a preemptive multiclass wavelength reservation protocol to provide differentiated service for Optical Burst Switched (OBS) networks. Unlike existing approaches, which may degrade to classless schemes or which may suffer from low wavelength utilization, our mechanism is robust, efficient, and supports an incremental deployment of QoS support. We maintain a usage profile for each class at the router, and implement a preemptive wavelength reservation mechanism to ensure QoS. We derive an analytical model and conduct simulations to evaluate the performance. The result shows that our approach performs the best in terms of lower blocking probability and higher resource utilization, making it an excellent QoS mechanism for OBS networks.

I. INTRODUCTION

Optical Burst Switching (OBS) [1-2] is a promising solution to provide terabit optical routing and to build an all-optical WDM layer for the optical Internet. Based on the concept of "burst" switching, OBS groups several IP packets with the same network egress address and common attributes like quality of services (QoS) into a burst and forwards the burst through the network as a single entity. A burst consists of a control header and a burst payload. OBS uses physically separate wavelengths (or channels) to transmit data bursts (i.e., payloads) and their headers, with the burst header transmitted slightly ahead in time. This allows optical core routers to process the headers electronically for the establishment of an end-to-end optical path, and to switch data bursts optically. Based on a one-way reservation protocol, OBS precedes a data burst with a control header in a predefined offset time, without waiting for an acknowledgment before the beginning of the data transmission.

Current IP only provides best effort service to deliver variable length packets. The future Internet may demand differentiated services for multimedia applications. For the optical Internet to be truly ubiquitous, one must address, among other important issues, how the WDM layer supports differentiated service. There are many mechanisms that implement QoS in the literature [3-5], mostly using buffers and scheduling. These approaches, however, incur high processing overhead at intermediate nodes and mandate a certain amount of buffers for switching. To date, no efficient optical buffer is available. The use of electronic

buffers necessitates opto-electro-optic (O/E/O) conversions, which must be avoided in an all-optical network where data is kept in the optical domain at all intermediate nodes. This calls for new QoS mechanisms which do not require buffers at the WDM layer in OBS networks. [6] proposed a new approach that assigns different offset times to different service classes so as to provide differentiated services in terms of burst loss probability for classes of different priorities (we call it "JET QoS" in the rest of the paper). JET QoS, however, requires a homogenous reservation mechanism, in which the burst length and the offset time of each class should be set to the same values for all routers in an OBS domain; otherwise, it may degrade to a classless approach like the original JET [2]. The offset times, however, may be not maintainable at the routers due to possible longer rerouting paths and congestion in control channels (consumed for extra processing time of the control header), rendering the mechanism unable to provide QoS. Furthermore, JET QoS reduces the loss rates of high priority traffic at the expense of an increase in the loss rates of lower priority traffic. [7] pointed out that this treatment is particularly unfair to long bursts of low priority. In order to cope with the unfairness, a proportional QoS scheme is studied in [7]. The loss rate of each class is maintained in a predefined proportion according to the priority. An arrival packet will be dropped if its predefined loss rate is violated regardless of whether there is an idle channel. The intentional dropping gives more and longer free periods of wavelengths in the output link to admit high priority bursts. This approach, however, always causes excessive dropping. As a result, it has higher overall blocking probability and worse wavelength utilization as compared to classless approaches like JET. Moreover, [7] makes no suggestion on the setting of the proportion factor. Improper setting may lead to high blocking probabilities of lower classes, or may be unable to enforce service differentiation to different classes of traffic.

This paper proposes a preemptive wavelength reservation protocol to provide differentiated service for OBS networks without requiring buffers at the WDM layer. We maintain a usage profile for each class at the

* This author is also with the Graduate Institute of Communications Engineering, Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan. Email: wjliao@cc.ee.ntu.edu.tw

router, and implement a preemptive wavelength reservation algorithm to ensure QoS. Unlike JET QoS, which is unfair and may degrade to a classless mechanism, our protocol supports incremental QoS deployment and cooperates well with other "best-effort" reservation mechanisms like Horizon [1] and the original JET [2]. Furthermore, our approach provides service differentiation without the excessive dropping problem suffered by [7]. We derive an analytical model and conduct simulations to evaluate the performance of the proposed mechanism. The results show that our approach performs the best in terms of lower blocking probability and higher resource utilization, making it an excellent QoS mechanism for OBS networks.

The rest of the paper is organized as follows. Section II describes the proposed preemptive multiclass wavelength reservation protocol. Section III presents an analytical model and shows the simulation results to evaluate the performance of the proposed mechanism. Finally the concluding remarks are included in Section IV.

II. PREEMPTIVE RESERVATION MODEL

This section describes the proposed service differentiation with a preemptive wavelength reservation mechanism for the OBS network.

A. Protocol Fundamentals

Suppose that a switch has a total of m wavelengths per output link to serve data bursts. Considering the characteristics of traffic, we classify bursts into k different classes of service, say c_1, c_2, \dots, c_k , and assign each class a service priority. Without loss of generality, classes c_1, c_2, \dots, c_k are assumed to have priority in an ascending order. The higher the priority, the lower the blocking probability. Since the blocking probability in a switch is inversely proportional to the number of idle wavelengths, we let classes of higher priorities use more resources (i.e., wavelengths). In other words, if class c_j has priority over class c_i , class c_j bursts are allowed to use more resources than class c_i . Each class is associated with a usage limit, defined as a proportion of system resources the class is allowed to use. Let p_i be the usage limit assigned to class c_i . $0 \leq p_1 < p_2 < \dots < p_k \leq 1$, and

$$\sum_{i=1}^k p_i = 1.$$

In our protocol, each class is associated with a predefined usage limit. A switch maintains a usage profile for a class per output link, and monitors the current usage of each class. Based on the profile table, the switch can determine if there is an *eligible* wavelength in the

outgoing link to schedule a new request (i.e., control headers). A wavelength is eligible to a request, say R , if it is not assigned to any other request during the burst duration of request R .

The usage profile of each class records a predefined usage limit, a current usage, and a list of scheduled requests in the same class with the following triple: burst duration, outgoing wavelength, and a predefined timer. The burst duration can be calculated from the burst header, which carries an offset time (i.e., the time difference between the control header and the data burst) and a burst length. Let l , s , and e be the burst length, the start time and the end time of a data burst, respectively. The start time s is equal to the current time plus the offset time carried in the control header, and $e = s + l$. The burst duration is maintained in the format of (the start time, the end time) of the data burst. The outgoing wavelength may be an eligible or preempted wavelength scheduled to transmit the data burst in the burst duration. The predefined timer records the maximum tolerable time to wait for the receipt of the data burst, in an attempt to cope with network faults and the preemption incurred by our protocol.

B. In-Profile Verification

A class of traffic is said to be in profile if its current usage does not exceed a predefined usage limit; otherwise, the class is out of profile. Intuitively, the usage of class c_i is defined as the total duration of the scheduled bursts of class c_i over the total duration of the scheduled bursts of all classes. Considering the bursty nature of traffic, the class usage is monitored in a short time period, say τ , which is called the monitoring timescale in [7]. The current usage of class c_i is defined as follows:

$$\rho_i = \frac{\sum_{j=1}^{n_i} l_j}{\sum_{i=1}^k \sum_{j=1}^{n_j} l_j},$$

where k is the number of classes, n_j is the

number of scheduled requests of class c_j in $(t, t + \tau)$, and l_j is the burst length of request R_j .

Class c_i is said to be in profile if $\rho_i \leq p_i$; otherwise, the class is out of profile.

C. Burst Preemption

Suppose that a switch has newly received R_x , an in-profile class request, with start and end times of s_x and e_x , respectively. Let C_O be a set of candidate classes to be preempted, defined as

- (1) $C_O = \{c_i \mid \rho_i > p_i, i = 1, 2, \dots, k\}$
- (2) $\forall c_i \in C_O, \exists R_v \in c_i, s_v - \varepsilon_s \leq s_x \leq e_x \leq e_v + \varepsilon_e$,

where ρ_i and p_i are the current usage and the usage limit of class c_i , respectively; ε_s and ε_e are the gaps to the previous and next scheduled requests, respectively, on the same wavelength as request R_v , a scheduled, out-of-profile request; s_v and e_v are the start time and end time of request R_v , respectively. These two conditions imply that every class in C_O must be out-of-profile (by condition (1)), and must, at least, include a request previously scheduled and its burst duration plus the two gaps to the previous and next scheduled requests on the same wavelength can cover the newly received request (by condition (2)).

Suppose that there is no eligible wavelength on an outgoing link to schedule the in-profile request R_x . The preemption process is triggered and proceeds as follows. The switch preempts a wavelength from a scheduled, out-of-profile request in the set of candidate classes C_O , starting from the class with the lowest priority to the highest priority. Request R_p will be the victim to be preempted if the preempted wavelength is eligible to serve request R_x (i.e., the burst duration of request R_p , plus the two gaps to the previous and the next requests, are eligible to request R_x). In other words, the time period of a wavelength can be used to schedule R_x if R_p is preempted.

D. Operation Overview

A switch keeps monitoring its usage table. Upon receiving a class c_i request, the switch first attempts to identify a wavelength eligible to the request. If the attempt succeeds, the request is scheduled, and the usage profile of class c_i is updated; otherwise, the following takes place. The switch examines if the class to which the request belongs is in profile, using the in-profile verification algorithm described in Sec. II-B. If it is in profile, a previous scheduled request of an "out-of-profile" class is preempted using the burst preemption algorithm described in Sec. II-C; otherwise, the request is rejected and the data burst is simply dropped. Once a wavelength is preempted, the switch updates the current usage of both requests' classes accordingly.

A switch may schedule a request either with an eligible wavelength, or a preempted wavelength. To prevent a scheduled request from being preempted in any switch of the data path, thereby wasting resources due to inconsistent scheduling, a switch should be able to detect preemption on previously scheduled bursts in any switch. We discuss this issue from two aspects:

- (1) A burst is preempted during the transmission.

In this case, the switch truncates the burst and sends out a Guard-E signal in the physical layer as described in [8] to end the burst. Thus, all switches in the channel can learn the preemption.

- (2) A burst is preempted before the start of the transmission.

Recall that each pair of (burst duration, outgoing wavelength) is associated with a predefined timer in the usage profile. The timer is activated either at the requested start time of a burst, or in the middle of any burst transmission when a data packet is not received in time. On expiry of the timer, the switch assumes an occurrence of a fault (either a physical fault or a preemption) if no data burst has been received.

In either case (i.e., timeout or on receipt of a Guard-E signal), the switch will remove the switching information of the associated burst, and makes the wavelength available to other requests.

III. PERFORMANCE EVALUATION

A. Analytical Result

This section analyzes the performance of the proposed preemptive mechanism. To simplify the analysis, we assume that both burst length and arrival time are exponentially distributed with an average of L . The offset time between the burst header and payload is assumed constant. We consider K classes of services in the system and m wavelengths in a switch.

We model the number of wavelengths used by each class in a switch as a continuous-time Markov chain. Let λ_i and μ_i be the arrival rate and the service rate of class c_i , respectively. Let $\bar{n} = (n_1, n_2, \dots, n_k)$ denotes the state of the Markov chain, where n_i is the number of wavelengths used by class c_i (i.e. the current usage of class c_i is $\rho_i = n_i/m$). In our analysis, we assume τ is equal to 0. In other words, whether a request is accepted depends on the state at the moment that the request arrives plus the offset time. The state space of the Markov chain can be derived as

$$S = \left\{ \bar{n} \mid \sum_{i=1}^K n_i \leq m, 0 \leq n_i \leq m, 1 \leq i \leq K \right\}.$$

Let $P_{\bar{n}, \bar{n}}(j^+)$, $P_{\bar{n}, \bar{n}}(j^-)$, $P_{\bar{n}, \bar{n}}(i^+, j^-)$ be the transition rate from \bar{n} to $\bar{n}(j^+)$, $\bar{n}(j^-)$, and $\bar{n}(i^+, j^-)$, where

$$\bar{n}(j^+) = (n_1, \dots, n_{j-1}, n_j + 1, n_{j+1}, \dots, n_k),$$

$$\bar{n}(j^-) = (n_1, \dots, n_{j-1}, n_j - 1, n_{j+1}, \dots, n_k), \text{ and}$$

$$\bar{n}(i^+, j^-) = \begin{cases} (n_1, \dots, n_{i-1}, n_i + 1, n_{i+1}, \dots, n_{j-1}, n_j - 1, n_{j+1}, \dots, n_k) & \text{if } i < j \\ (n_1, \dots, n_{j-1}, n_j - 1, n_{j+1}, \dots, n_{i-1}, n_i + 1, n_{i+1}, \dots, n_k) & \text{if } i > j \end{cases}$$

Subject to the limited number of outgoing wavelengths, the transition rate can be described in two cases.

(i) Case 1: when $\sum_{i=1}^K n_i < m$

$$P_{\bar{n}, \bar{n}(j^+)} = \lambda_j, \text{ and } P_{\bar{n}, \bar{n}(j^-)} = \begin{cases} n_j \mu_j, & \text{if } n_j > 0 \\ 0, & \text{otherwise} \end{cases}$$

(ii) Case 2: when $\sum_{i=1}^K n_i = m$

$$P_{\bar{n}, \bar{n}(i^+, j)} = \begin{cases} \lambda_i, & \text{if } n_i < r_i, j = \min\{h | n_h > r_h\} \\ 0, & \text{otherwise} \end{cases}$$

$$P_{\bar{n}, \bar{n}(j^-)} = \begin{cases} n_j \mu_j, & \text{if } n_j > 0 \\ 0, & \text{otherwise} \end{cases}$$

where r_i denotes the number of wavelengths class c_i is allowed to use (i.e. the usage limit of class c_i is $p_i = r_i/m$). Because the Markov chain is aperiodic, finite, and irreducible, the stationary probability $p(\bar{n})$ exists. From the stationary probability, we can derive the following performance matrix:

(i) The blocking probability of a new arrival burst of class c_i is given by $P_{block}(i) = \sum_{\bar{n} \in \left\{ (n_1, \dots, n_k) \mid \sum_{j=1}^k n_j = m, n_i \geq r_i \right\}} p(\bar{n})$.

(ii) The preempted probability of an accepted burst of class c_i is given by

$$P_{preempted}(i) = \frac{N_{preempted}(i)}{N_{accepted}(i)} = \sum_{j=1}^K \sum_{\substack{\bar{n} \in \left\{ (n_1, \dots, n_k) \mid n_j < r_j, i = \min\{h | n_h > r_h\} \right\} \\ \sum_{s=1}^k n_s = m}} p(\bar{n}) \times \frac{\lambda_i}{\lambda_j} \times \frac{1}{1 - P_{block}(i)}$$

where $N_{preempted}(i)$ is the number of class c_i bursts preempted and $N_{accepted}(i)$ is the number of class c_i bursts accepted.

B. Simulation Result

This section presents the simulation results that compare the performance of the proposed mechanism with classless (i.e., best effort), JET QoS [6], and Drop [7]. The classless mechanism may be Horizon [1] or the original JET [2]. The JET QoS is the original JET with different extra offset times assigned to different classes of bursts. Drop refers to the intentional dropping approach proposed in [7].

We consider bufferless switches with m wavelengths in each output link. Each switch is assumed to be capable of full wavelength conversion. We assume there are k classes, all of which generate bursts with an exponential inter-arrival time and exponential burst duration. To simplify the computation and without loss of generality, the simulation is based on the assumption that all sources

have the same arrival rates (i.e., $\lambda_1 = \lambda_2 = \dots = \lambda_k = \lambda$) and service rates (i.e., $\mu_1 = \mu_2 = \dots = \mu_k = \mu$).

We first consider two classes only, namely, classes 1 and 2. We let class 2 have priority over class 1, and assign the usage limits of 0.0 and 1.0 to classes 1 and 2, respectively. Thus, class 2's traffic can preempt class 1's traffic when necessary. Figs. 1 to 3 show the blocking probabilities of different mechanisms as a function of the offered load in a single bufferless WDM switch. The

offered load here is defined as $\frac{\sum \lambda_i}{m\mu}$, where m is the number of wavelengths in each link, λ_i is the burst arrival rate of class c_i and μ is the service rate of each burst.

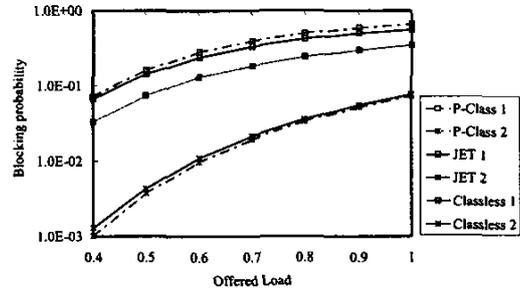


Figure 1. Blocking probabilities of the three mechanisms

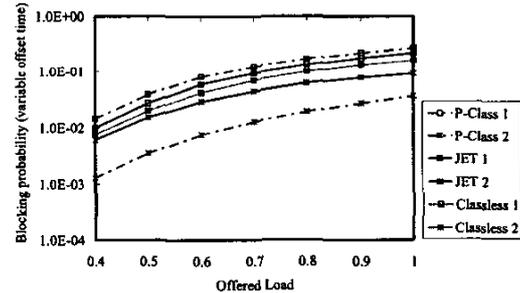


Figure 2. Blocking probabilities of the three mechanisms with network congestion

Fig. 1 shows that the blocking probabilities of the three approaches when each output link has 8 wavelengths. The blocking probabilities increase as the offered load increases. Both JET QoS and the proposed preemptive approach (denoted P in the figure) provide service differentiation for multiclass traffic. Fig. 2 shows the blocking probabilities of the three approaches when the offset delay time becomes invalid due to network congestion. Our approach still provides differentiated service for classes 1 and 2 traffic, while JET QoS may degrade into a classless scheme. Fig. 3 shows the system utilization of the three mechanisms. It can be seen that our approach has the highest system utilization, followed by

JET QoS, and Drop is the worst. Thus, from Figs. 1 to 3, our approach performs the best in terms of better differentiation and higher system utilization.

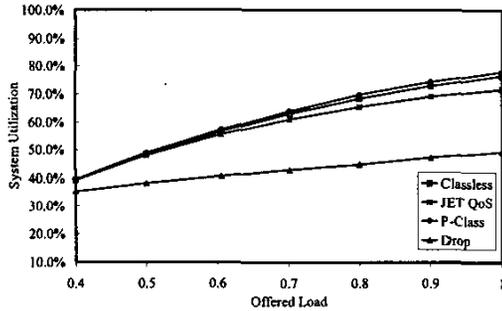


Figure 3. System utilization of the four mechanisms

C. Analysis vs. Simulation

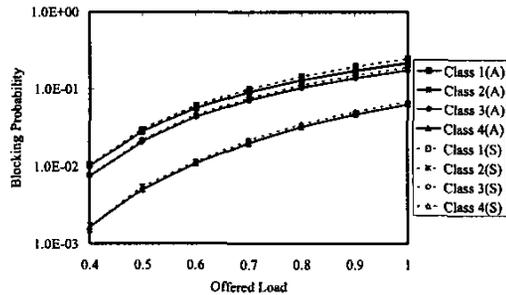


Figure 4. Blocking probability of new bursts

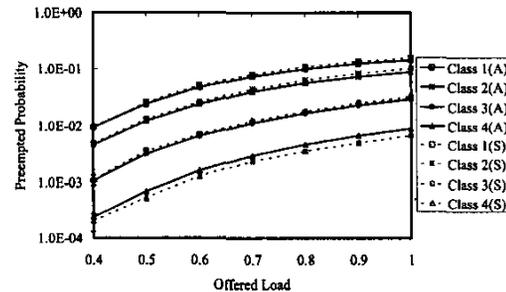


Figure 5. Preempted probability of an accepted call

This section compares the analytical results with the simulation. We consider four classes, i.e., class 1, class 2, class 3, and class 4, with ascending priorities. The usage limits of classes 4 to 1 are 0.5, 0.25, 0.125, and 0.125, respectively. In other words, the used wavelengths of each class at time t are 4, 2, 1 and 1 respectively. Fig. 4 shows the blocking probabilities of new coming bursts and Fig. 5 depicts the preempted probabilities of accepted bursts. In both figures, the solid lines (indicated by A) are from the analytical results, and the dashed lines (indicated by S) are from the simulation. The analytical results match the simulation results very well. The latter is obtained with 98% confidence interval within ± 0.001 of

the mean value.

IV. CONCLUSION

In this paper, we have described a new bufferless mechanism using a preemptive wavelength reservation mechanism to differentiate services in optical burst switched networks. From the analytical model and simulation, we show that our approach performs best in terms of lower blocking probability and higher resource utilization, making our approach an excellent QoS mechanism for OBS networks.

ACKNOWLEDGEMENT

This work was supported in part by the National Science Council, Taiwan, under Grant Number 91-2213-E-002-067, and in part by the Communications Software Technology project of the Institute for Information Industry (III) and sponsored by Ministry of Economic Affairs (MOEA), R. O. C.

REFERENCES

- [1] J. Turner, "Terabit burst switching," *Journal of High Speed Networks*, vol. 8, pp. 3-16, 1999
- [2] 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.
- [3] H. Zhang, "Service Disciplines for Guaranteed Performance service in Packet Switching Networks," *Proc. of the IEEE*, vol. 83, pp. 1374-1396, Oct. 1995.
- [4] A. Varma and D. Stilladis, "Hardware Implementation of Fair Queuing Algorithms for Asynchronous Transfer Mode Networks," *IEEE Communications Magazine*, vol. 35, pp. 74-80, 1997.
- [5] U. Briem et al., "The Traffic Management for an ATM Switch with per-VC Queuing: Concept and Implementation," *IEEE Communications Magazine*, vol. 36, pp. 88-93, Jan. 1998.
- [6] M. Yoo and C. Qiao, "Supporting Multiple Classes of Services in IP over WDM Networks," *Proc. IEEE Globecom '99*, pp. 1023-1027, 1999.
- [7] Y. Chen, M. Hamdi and Danny H.K. Tsang, "Proportional QoS over OBS networks", session FPT03, Proc. IEEE GLOBECOM 2001.
- [8] Y. Xiong, M. Vandenhouete, and H. C. Cankaya, "Control Architecture in Optical Burst-Switched WDM Networks," *IEEE JSAC*, vol. 18, no. 10, pp. 1838-1851, Oct. 2000.