Class-Based Dynamic Buffer Allocation for Optical Burst Switching Networks[†]

Jingxuan Liu and Nirwan Ansari
Advanced Network Laboratory
Department of Electrical and Computer Engineering
New Jersey Institute of Technology
Newark, NJ 07102-1982, USA

Abstract-- **Optical Burst Switching (OBS) is a promising technology to facilitate IP-Over-WDM. In order to improve the quality of services (QoS) in OBS transmission networks, an adaptive method based on linear predictive filtering is proposed to dynamically allocate the limited optical buffers to different classes of traffic. Optical buffers are allocated on a per-class and per-label basis, according to the incoming traffic intensity and priority. This buffer management strategy decreases traffic loss caused by resource contention, and gives the loss-sensitive traffic precedence to access buffers over the loss-tolerant traffic. A new system model is introduced, and simulations show that our proposed adaptive-filter-based dynamic buffer allocation method provides better QoS compliance to traffic with higher priority, and achieves a lower system weighted loss than a fix-sized buffer system does.**

Index Terms-- Optical burst switching, WDM, dynamic buffer allocation, adaptive filter, traffic priority

I. INTRODUCTION

Optical Burst Switching (OBS) is one of the most promising mechanisms for the next generation optical Internet to support IP over WDM. It combines advantages of both the wavelength switching and optical packet switching (OPS), with high wavelength utilization and low buffering requirement for optical packets.

An OBS system can implement transparent transmission of data bursts. Each data burst is preceded by a control header, which is transmitted in a different optical channel from those for data traffic. While the control header needs to be OEO (Optical-Electrical-Optical)-converted at intermediate nodes to process the routing information, data bursts never go through any OEO conversion, nor are buffered in the electronic domain. Since less coupling between a control header and its coordinating payload is needed, an OBS system requires much less buffers than an OPS system does.

The switching granularity in an OBS network is finer than that in a wavelength switching system (which has the granularity of a single wavelength). In an OBS system, multiple IP packets with the same destination and quality of service (QoS) request may be routed as an entity, namely, a burst. A wavelength can be shared by bursts with different Source-Destination (DS) pairs. Such strategy enables an OBS system to have a better bandwidth utilization in term of wavelength, as discussed in [1,2,3].

Although OBS outperforms OPS and wavelength switching with less OEO conversion overhead and better bandwidth utilization, many OBS-specific issues have to be addressed before this technology becomes practical and efficient. Two among many are resource contention and burst loss. Resource competition at a switching node may occur due to the unavailability of optical buffers or output path; e.g., when the optical buffers are not enough to accommodate bursts while their coordinating control headers are being electronically processed, or when there are no more data channels available for bursts heading for the same output path simultaneously. Such resource competitions may result in data burst loss. Data burst loss, especially those with higher QoS requirement, is critical because each data burst may contain a very large amount of data. To support next generation Internet applications and IP over WDM, QoS requirements must be considered in the system design. Various aspects, such as offset management, wavelength conversion, and system scheduling, can be employed to provide QoS [4,5].

Buffer-based methods, in which FDLs (Fiber Delay Lines) are typically used for contention resolution, are an alternative strategy. Quite a few studies have been conducted by using optical buffers for either synchronized or asynchronous optical packet systems, both based on the static buffer allocation consideration [6,7]. Due to the limited availability of optical buffers, however, we believe optical buffers should be effectively utilized. Optical buffers should be managed in such a way that QoS performance for different classes of traffic should not only be facilitated, but the buffer resource utilization should also be improved.

In this paper, we propose a dynamic buffer allocation strategy, which aims to achieve near-optimal utilization of limited resources (optical buffers), and to minimize burst loss resulted from the resource contention. The dynamic buffer allocation is performed on a per-label and per-class basis. Bursts with the same QoS requirement belong to one class. A "higher" class has priority to get more buffers. This strategy guarantees better service to the more loss-sensitive traffic, which is usually classified in a higher class.

Another characteristic of our system is that optical buffers are allocated according to the real-time workload, or traffic intensity. This is implemented by an adaptive-filter-based algorithm. The system keeps track of the traffic flow of each

---

[†] This work has been supported in part by the New Jersey Commission on Higher Education via the NJI-TOWER project, and the NJ Commission on Science and Technology via the NJ Center for Wireless Telecommunications Center.

class, and predicts the incoming traffic amount in the next buffer update interval. This prediction, combined with the traffic class priority, is then used to calculate the new optical buffer size needed for each class. Simulation results confirmed the effectiveness of our algorithm, showing that a system with our proposed adaptive-filter-based dynamic buffer allocation strategy substantially reduces the class-based loss ratio as compared to the results from a system using the static buffer allocation method. Weighted loss ratio, which will be explained in Section II, is used as a figure of merit for system performance.

The rest of the paper is organized as follows. Section II describes the system model. Section III explains the principle of traffic prediction, and presents our proposed dynamic buffer allocation. Simulation results and analysis are presented in Section IV, and conclusions are drawn in Section V.

II. THE SYSTEM MODEL

We assume MultiProtocol Label Switching (MPLS) [1] is used to setup a lightpath and to reserve bandwidth for each source-destination pair (SD). Each Label Switching Path (LSP) is identified with a label, and can be transmitted via more than one channel. Label merging is assumed possible, i.e., traffic from different sources may share the same label, provided that they are destined for the same egress.

We shall first list the following notations that are widely used throughout the paper:

l_i: the LSP label of an incoming burst, $i \in \{0,\dots,m-1\}$, where m is the maximum number of labels allowed in the system;

c_j: the class of an incoming burst, $j \in \{0,\dots,n-1\}$, where n is the maximum number of classes allowed in the system. Each incoming burst is identified by a tuple (l_i, c_j), which specifies the label and the class it belongs to;

B^{l_i} : the size of the optical buffers assigned to label l_i. B^{l_i} is shared and dynamically allocated among all classes with label l_i. This buffer size is fixed for each label, and is dynamically allocated among all classes.

$B^{l_i}_{c_j}(k)$: the portion of B^{l_i} allocated to class c_j at the k-th buffer update operation, subjected to . $\sum_{j=0}^{n-1} B^{l_i}_{c_j} = B^{l_i}$

$L^{l_i}_{c_j}$: The total length of lost bursts of the (l_i, c_j) traffic flow;

$D^{l_i}_{c_j}$: The total length of incoming bursts of the (l_i, c_j) traffic flow;

α_{c_j}: *weight factor*. Each class has a weight factor, representing the cost of losing a burst of c_j class. A class with a larger α value has more stringent loss requirement.

For the rest of this paper, we address the buffer allocation with respect to only one label, and thus, for notational simplicity, we will either omit the referencing of this label, or simply refer to it as l.

Figure 1 shows the system architecture, in which four labels (l_0, l_1, l_2, l_3) are carried in each input fiber (f_0, f_1). We suppose that after switching, l_0 and l_2 are carried on output fiber f_0'; l_1 and l_3 are carried on output fiber f_1'. The system architecture is similar to the one in [8]. However, here, we employ an output buffering scheme. Optical buffers (FDLs) at each output path are dynamically allocated among all classes. The SCU (switch control unit) has to implement the algorithm proposed in this paper.

To evaluate the system performance, two figures of merit, the burst loss ratio and the system weighted loss, are defined as follows:

$$Loss\ Ratio = \rho^l_{c_j} = \left(\frac{L^l_{c_j}}{D^l_{c_j}}\right)\cdot 100\% \qquad (1)$$

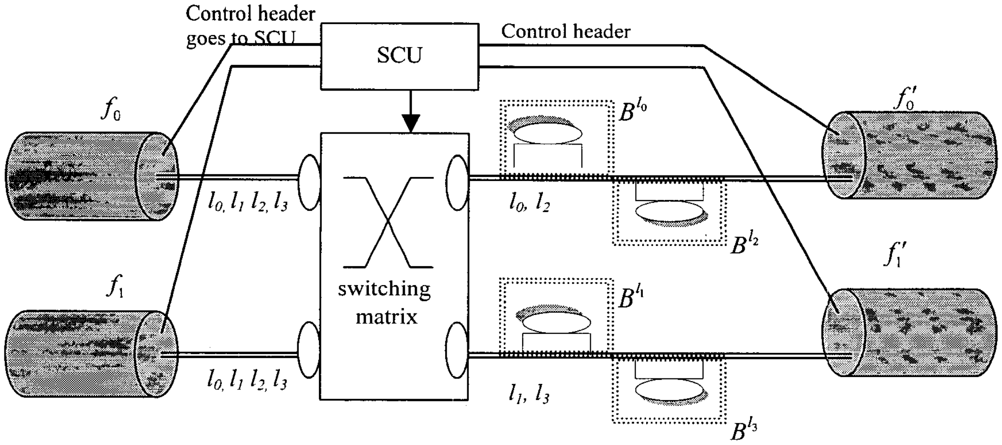

Figure 1. System architecture

$$\text{System Weighted Loss} = \eta = \frac{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(L^{l_i}_{c_j}\cdot\alpha_{c_j})}{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(D^{l_i}_{c_j}\cdot\alpha_{c_j})}\cdot 100\% \qquad (2)$$

Eq. (1) considers the system performance on a per-label and per-class basis, while Eq. (2) evaluates on a system scale. A good buffer management strategy should yield low burst loss ratio for traffic with higher priority, and keep the weighted loss--a measure of the total loss caused by all traffic classes--at an acceptable level.

III. TRAFFIC PREDICTION AND DYNAMIC BUFFER ALLOCATION

As stated in Section I, our buffer allocation is based on traffic prediction. However, the high-speed nature of optical networks makes it difficult, if not impossible, to perform dynamic allocation calculation for each individual incoming burst. To circumvent this problem, the update interval Δt must be chosen appropriately. The value of Δt depends on the SCU's processing power and the total buffering capacity of B^l. Too small the Δt introduces too large processing overhead for updating the buffer allocation; too large the Δt renders the updating irresponsive to traffic pattern changes (i.e., the estimated buffer size becomes outdated). In our system, we suggest Δt varies according to the traffic load as follows:

$$\Delta t \approx \frac{200}{\lambda},$$

where λ is the average burst arrival rate. Such an update interval renders better traffic prediction performance.

Figure 2 shows the traffic prediction algorithm. An P-order Linear Predictive Filter (LPF) is used to predict the next incoming traffic of class c_j in label l. Let $X^l_{c_j}(k)$ be the total traffic amount (in *microsecond*) of burst stream (l, c_j) during the k-th update interval $[T_k, T_k+\Delta t]$, and let $\hat{X}^l_{c_j}(k+1)$ be the predicted length of traffic in the next interval $[T_{k+1}, T_{k+1}+\Delta t]$. $\hat{X}^l_{c_j}(k+1)$ is estimated by the previous traffic amount according to

$$\hat{X}^l_{c_j}(k+1) = \sum_{i=1}^{P} w^l_{c_j}(i)\cdot X^l_{c_j}(k-i+1), \qquad (3)$$

where $w^l_{c_j}(i)$, $i \in \{1,\dots,P\}$, is the coefficient of the adaptive filter. The coefficient is updated by the normalized LMS algorithm as follows [9]:

$$(W^l_{c_j})^{k+1} = (W^l_{c_j})^k + \frac{\mu\cdot e^l_{c_j}(k)\cdot(X^l_{c_j})^k}{\|(X^l_{c_j})^k\|^2}, \qquad (4)$$

where:

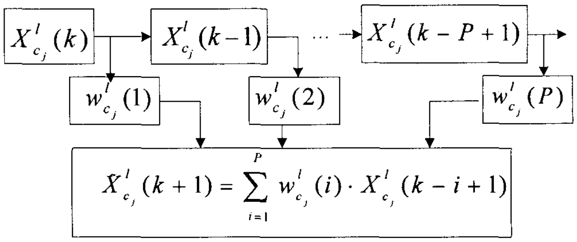

Figure 2. Traffic prediction with LPF

$$(W^l_{c_j})^k = \begin{bmatrix} w^l_{c_j}(1) \\ w^l_{c_j}(2) \\ \vdots \\ w^l_{c_j}(P) \end{bmatrix}^k, \quad (X^l_{c_j})^k = \begin{bmatrix} X^l_{c_j}(k) \\ X^l_{c_j}(k-1) \\ \vdots \\ X^l_{c_j}(k-P+1) \end{bmatrix},$$

and μ is an adjustable parameter of the LPF, determined by the actual traffic pattern. $e^l_{c_j}(k)$ is the error between the predicted and actual traffic length in the k-th update interval ($e^l_{c_j}(k) = X^l_{c_j}(k) - \hat{X}^l_{c_j}(k)$).

Once the incoming traffic during the next time interval has been predicted, it is combined with the loss weight factor (α_{cj}) to decide the buffer size for each class c_j. The estimated buffer size is:

$$\hat{B}^l_{c_j}(k+1) = \left(\frac{\alpha_{c_j}\cdot\hat{X}^l_{c_j}(k+1)}{\sum_{j=0}^{n-1}\alpha_{c_j}\cdot\hat{X}^l_{c_j}(k+1)}\right)\cdot B^l. \qquad (5)$$

The complexity of the algorithm increases linearly as the order of the adaptive filter (P) and the number of classes (n) increases. According to Eq. (4) and Eq. (5), the complexity of the algorithm is $O(n*P)$.

IV. SIMULATION RESULTS AND ANALYSIS

We conducted several simulations to study the behavior of the system based on the adaptive predictor. The simulator was OPNET 8.0 on SUN Blade 1000 platform. The traffic was modulated as Poisson processes. An order-12 adaptive filter was employed.

We simulated two incoming fibers (f_0, f_1), each supports two labels (l_0, l_1). Two classes of traffic were considered: loss-sensitive real-time traffic (c_0), and loss-tolerant nonreal-time traffic (c_1). Class c_0 has higher priority than c_1. For the output path, we observe the behavior of the one with label l_0, and compare the performance of our system with that resulted from a static buffer allocation system, in term of loss ratio (ρ) and system weighted loss (η) (Eq. (1) and Eq. (2)). In the static buffer allocation system, the total buffer B^{l_0} is initially allocated to each class according to their priority:

$$B_{c_j}^{l_0} = \frac{\alpha_{c_j}}{\alpha_{c_0} + \alpha_{c_1}} \cdot B^{l_0}, \quad j \in \{0,1\}$$

This size is fixed throughout the simulation. Denote

ρ_0 : loss ratio for class 0 in the dynamic buffer system;

ρ_1 : loss ratio for class 1 in the dynamic buffer system;

η : system weighted loss of the dynamic buffer system;

$\tilde{\rho}_0$: loss ratio for class 0 in the static buffer system;

$\tilde{\rho}_1$: loss ratio for class 1 in the static buffer system;

$\tilde{\eta}$: system weighted loss of the static buffer system.

Figure 3 shows the loss ratio and system weighted loss versus the available buffer size. The traffic load for label l_0 is 0.38. The x-axis indicates the buffer capacity of label l_0 in term of the ***maximum*** number of bursts allowed; i.e., B^{l_0} divided by the average burst length. $\alpha_{c_0} : \alpha_{c_1}$ is set to be 2.5:1.

As shown in the figure, our dynamic buffer allocation system offers much better service (specially, less loss ratio) to the "higher" class traffic ($\rho_0 < \rho_1$). Bursts of the "higher" class suffer less loss ratio than those of the "lower" class, because the "higher" class traffic is given more buffers owing to its larger loss weight factor, thus providing better QoS.

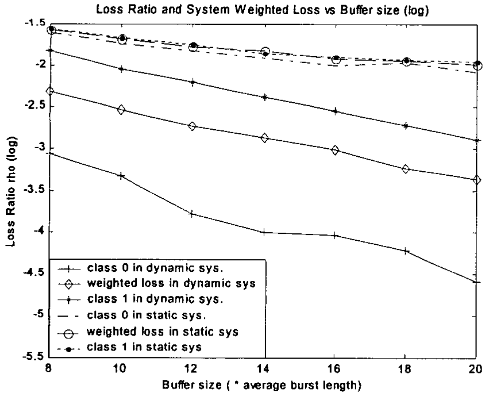

Figure 3. Loss ratio vs buffer size

Our dynamic method improves buffer utilization greatly. To achieve the same loss ratio, our system requires much less buffers than the static buffer allocation system does. This merit lies in the flexibility of our algorithm. Buffer sizes are not fixed for each class. Instead, when the workload of one class becomes heavier for some time intervals, it is allowed to get more buffers than it is initially assigned. This method makes better use of the limited buffers, and improves the performance on a system scale.

Figure 4 presents the loss ratio versus the traffic load. In this simulation, we assume the available buffer for each label (both classes) is sixteen times that of the average burst length. $\alpha_{c_0} : \alpha_{c_1}$ is set to be 2:1.

This figure shows that even when the traffic load is heavy (>0.8), our method provides much better service to the "higher" class traffic, because class c_0 may take advantage of its larger weight factor and gets more buffers over c_1. Although this results in ρ_1 being a little bit higher than $\tilde{\rho}_1$, the method grants better service to the class that is more sensitive to burst loss, and meanwhile it achieves a lower system weighted loss ($\eta < \tilde{\eta}$). When the traffic load becomes lighter (<0.7), both classes get lower loss ratio in our system ($\rho_i < \tilde{\rho}_i$, *i=0, 1)*.

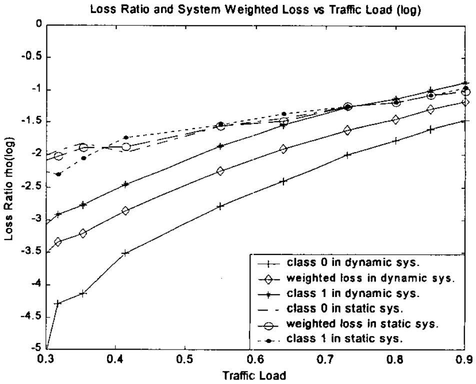

Figure 4. Loss ratio vs traffic load

Figures 5 and 6 show that the buffer size assigned to each class depends on both the traffic class and traffic intensity. We conducted this simulation by setting the traffic load of c_0 to be

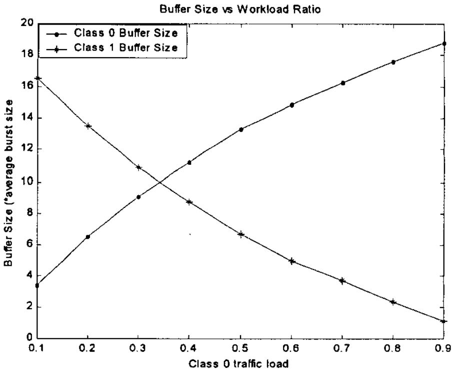

Figure 5 Buffer size vs traffic load of c_0

$\{0.1,\ldots,0.9\}$ of the total offered traffic with label l_0, with the total available B^{l_0} being 20 times the average burst length, and $\alpha_{c_0} : \alpha_{c_1}$ being $2{:}1$.

As shown in Figure 5, when the two classes have the same traffic density ($c_0 = 0.5$), the higher priority class receives more buffers ($B^{l_0}_{c_0} > B^{l_0}_{c_1}$). When the traffic intensity of the lower priority class grows large enough ($c_0 < 0.3$), however, c_1 gets more buffers ($B^{l_0}_{c_0} < B^{l_0}_{c_1}$).

Figure 6 shows the loss ratio versus the traffic intensity. When the traffic load is the same for both classes ($c_0 = 0.5$), the higher priority class suffers less loss ratio, since larger buffers are assigned to it.

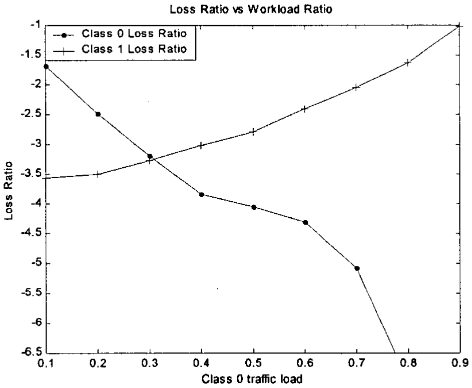

Figure 6 Loss ratio vs traffic load of c_0

The performance of the adaptive filter can be measured by the loss ratio. There is a tradeoff between the performance and the complexity. Simulation results show that a 12-order adaptive filter can predict the incoming traffic very well, as shown in Figure 7.

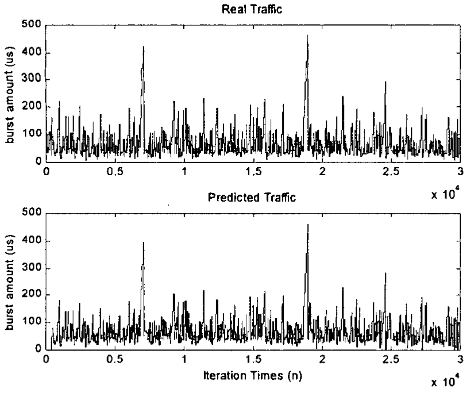

Figure 7. Predicted traffic vs. real traffic (order = 12)

V. CONCLUSIONS

OBS is an attractive technology to realize the next generation of optical networks (IP over WDM). To meet the situation that different classes of traffic (data, voice, video, etc) are converging into the same network but each requiring a different level of QoS, class-based QoS must be supported by an OBS network. This paper presents a simple but novel algorithm that dynamically allocates the limited optical buffer to increase the resource utilization and improve the system performance. This adaptive method lowers burst loss probability, without causing excessive delays for the high priority traffic.

In contrast to the offset-based QoS, the buffer-based method presented in this paper is not required to calculate the offset end to end before transmitting the bursts. The point-to-point (or distributed) feature results in simple routing and signaling algorithms. Its full impact on the network performance, along with more analytical work, such as the delay and update interval will be reported in the near future.

ACKNOWLEDGEMENT

The authors wish to acknowledge the help and insightful discussions with Yuanqiu Luo, Binghu Zhang, and Hong Zhao, on OPNET simulations, practical issues of optical networks, and adaptive filters.

REFERENCES

[1] R. Xu, G. Qian, P. Ye, "A novel IP with MPLS over WDM-based broad-band wavelength switched IP network," *IEEE J. Lightwave Tech.*, Vol. 19, No. 5, May 2001, pp. 596-602.

[2] G. Li, R. Simha, "On the wavelength assignment problem in multifiber WDM star and ring networks," *IEEE/ACM Trans. Networking*, Vol. 9, No. 1, Feb 2001, pp. 60-68.

[3] B. Li, A. Ganz, C.M. Krishna, "An in-band signaling protocol for optical packet switching networks," *IEEE JSAC*, Vol. 18, No. 10, Oct. 2000, pp. 1876-1884.

[4] H. Zang, J.P. Jue, B. Mukherjee, "Capacity allocation and contention resolution in a photonic slot routing all-optical WDM mesh network," *IEEE J. Lightwave Tech.*, Vol. 18, No. 12, Dec. 2000, pp. 1728-1741.

[5] M. Yoo, C. Qiao, and S. Dixit, "QoS Performance of Optical Burst Switching in IP-Over-WDM Networks," *IEEE JSAC*, Vol. 18, No. 10, Oct. 2000, pp. 2062-2071.

[6] L. Tancevski, A. Ge, and G. Castanon, "Optical Packet Switch with Partially Shared Buffers: Design Principles," *IEEE OFC*, Vol. 2, March 2001, pp. Tuk3-1-Tuk3-3.

[7] M. Yoo, C. Qiao, and S. Dixit, "The Effect of Limited Fiber Delay Lines on QoS Performance of Optical Burst Switched WDM Networks," *Proc. ICC'01*, Vol. 2, pp.974-979.

[8] Y. Xiong, M. Vandenhoute, H.C. Cankaya, "Control Architecture in Opical Burst -Switched WDM Networks," *IEEE JSAC*, Vol. 18, No. 10, Oct. 2000, pp. 1838-1851.

[9] M. L. Honig and D. G. Messerschmitt, *Adaptive Filters Structures, Algorithms, and Applications*, Klumer Academic Publishers, 1990.