Slotted Optical Burst Switching (SOBS) Networks

Zhenghao Zhang, Lin Liu, and Yuanyuan Yang
Dept. of Electrical & Computer Engineering
State University of New York at Stony Brook
Stony Brook, NY 11794, USA
{zhhzhang, linliu, yang}@ece.sunysb.edu

Abstract

In this paper we study Optical Burst Switching (OBS) networks. Since OBS still suffers high loss ratio due to the lack of buffer at the OBS core routers, we study methods to reduce the loss without using optical buffers. We consider time-slotted OBS called Slotted OBS (SOBS), where routers are synchronized and only send fixed length bursts at the beginning of time slots. Our simulation results show that SOBS reduces the packet loss probability significantly. Moreover, we show that SOBS can be implemented with little or no additional cost and has other advantages such as the better supporting of Quality of Service (QoS).

Index Terms—**Optical burst switching (OBS), slotted optical burst switching (SOBS), optical networks, optical switches, packet loss probability.**

I. Introduction

All optical networking with Wavelength Division Multiplexing (WDM) is widely regarded as the most promising candidate as the backbone network for future high speed communications because of its huge bandwidth [3]. However, its deployment has been slow, which is partly due to the lack of some fundamental technologies, especially the all-optical memories. Although there has been some recent advance in optical buffering [4] which might lead to controllable optical memory, it is unlikely that such technology will become mature in the very near future to provide optical memories of comparable size as the electronic buffers equipped in today's routers. Thus the network engineers have to face the challenge of using a network that has very large link bandwidth and yet no or very few buffers in the core switches to provide the best possible service to satisfy the ever growing need of the users.

Today, researchers have proposed three types of optical networks based on WDM, namely, the Wavelength Routed (WR) network, the Optical Packet Switching (OPS) network, and the Optical Burst Switching (OBS) network. A WR network is like the circuit switching network, in which to send information from one place to another, a *lightpath* has to be established. Apparently, its disadvantage is that it cannot well adapt to dynamic traffic because the entire bandwidth on the lightpath is reserved by the source-destination pair and cannot be used by others even when there is no traffic between the pair. An OPS network is like the electronic packet switching network and can better adapt to dynamic traffic. However, it is difficult to implement because it needs optical memories at the switches that can only be emulated by expensive Fiber Delay Lines (FDL) with today's technology. The OBS network proposed in [5], [6] appears to be a better choice since it does not waste bandwidth when there is no traffic and it does not need a large optical buffer in the switches. Thus the OBS network will be the main focus of this paper.

An OBS network is shown in Fig.1. In an OBS network, there are two types of routers, the edge router and the core router. The edge router aggregates packets to the same destination as a burst, and before the burst, a header burst is sent through a different channel to reserve the bandwidth on the links on the path from the source to the destination. Upon receiving the control burst, the core routers will configure the switching fabrics at the arrival time of the data burst to direct the data burst to the correct destination. In this way, the data burst does not have to be buffered at the core routers and "cut through" the core routers without delay. However, if two bursts want to access to the same output wavelength at the same time, one of them, or at least the overlapping part of the one of the bursts will have to be dropped.

At first glance, OBS should have better link usage efficiency than the WR network. However, this may not be the case. In OBS, because the core routers have no buffer, the burst loss probability is very high as compared to electronic routers. As a simple example, suppose the burst loss probabilities of the OBS routers are all 0.01 and are independent of each other. If a burst needs to travel 10 hops, then with a probability 0.1, it will be dropped. Note that when a burst is dropped, all the bandwidth it used is wasted and since a

Fifth IEEE International Symposium on Network Computing and Applications (NCA'06)
0-7695-2640-3/06 $20.00 © 2006 IEEE

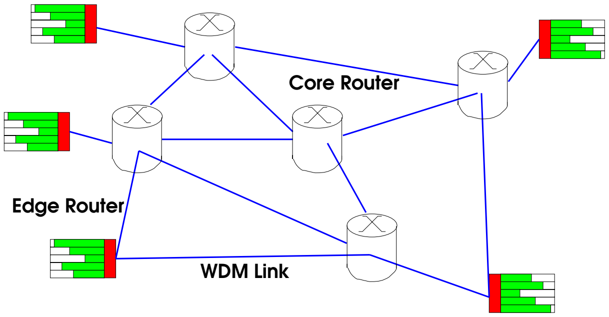

Fig. 1. An OBS network consists of edge routers and core routers connected by WDM links.

dropped burst will eventually be retransmitted, invoked by upper layer protocols such as TCP. Thus on average 10% of the bandwidth is wasted in this example. Therefore, to make OBS a viable choice for high speed networking, it is crucial to reduce the burst loss probability. In electronic networks, to reduce the loss probability, one can always add more buffers to the core switching routers. However, this cannot be done in optical networks, at least in the near foreseeable future due to the lack of optical memory, therefore other methods have to be adopted.

In this paper we will study methods for reducing the loss in OBS networks without resorting to buffering. We will mainly consider the time-slotted OBS called Slotted OBS (SOBS), where routers are synchronized and send only fixed length bursts at the beginning of time slots where a time slot is the length of a burst. This is because it has a significantly lower loss ratio than the non-slotted OBS and can be implemented with little or no additional cost. The reason for the reduction of loss ratio is that in SOBS the routers are better cooperating with each other than in the non-slotted OBS, much like the throughput of the slotted ALOHA is much higher than the non-slotted ALOHA. SOBS also has other important features such as better supporting Quality of Service (QoS) than the non-slotted OBS. Thus we expect SOBS to be the design choice of future OBS networks.

We are fully aware that we are not the first to propose the Slotted OBS network, as it has been considered, for example, in [1], [2]. However, we believe there are certain features of the SOBS that have not been fully discussed, which motivates the writing of this paper. For example, it has not been mentioned that SOBS will actually improve the overall link efficiency because less link bandwidth is wasted due to burst dropping and the void between consecutive bursts. Also, to the best of our knowledge, we are the first to point out the unique advantage of SOBS in supporting Quality of Service (QoS). We also propose a novel low-cost method for packet aligning at the core routers.

We will also explore other methods that have been shown to be effective in reducing the loss in OBS networks and incorporate some of them into SOBS. We will consider balancing the load on network links that was studied by [8] [16], and give a new load balancing algorithm. We will also consider *Burst Emission Control* at end routers, a concept first introduced by [7], and argue that it cannot be practically implemented to improve the performance of OBS. It should also be mentioned that *Burst Segmentation* [9] which is very effective in reducing loss in non-slotted OBS networks, is not needed in SOBS because in SOBS bursts either do not overlap or overlap completely, thus there is no segment of bursts.

The rest of the paper is organized as follows. Section II gives the motivations for moving from non-slotted OBS to slotted OBS and analyzes the cost of implementing SOBS. Section III gives a new algorithm for finding load balanced routes in OBS. Section IV discusses Burst Emission Control. Section V gives the simulation results. Finally, Section VI concludes the paper.

II. Slotted Optical Burst Switching – SOBS

In this section we give arguments in supporting the using of Slotted OBS and discuss its implementations.

A. Why Slotted

The reasons that lead us to considering the Slotted OBS are the following.

First, from the point of view of reducing the loss, burst should be of the same length. Note that bursts are dropped only when overlapped with other bursts. Therefore, comparing to longer bursts, shorter bursts should have a smaller chance of being dropped since they are less likely to overlap with other bursts. In the extreme case, if burst is infinitely short, no collision will occur. Thus, from a game theory point of view, to minimize the loss of burst, every edge router will choose to send packets in the shortest length burst. However, the length of the burst cannot be infinitely small because of limitations of the network such as the set up time of the core routers, and thus eventually the length of the burst will converge to a minimum value and will be the same in all edge routers.

Second, to reduce the loss, the receiving and the sending of the bursts should be synchronized. Note that if the edge routers send out bursts at random time instants, bursts will arrive at the core routers at random time instants. This will cause high burst loss since the overlapping time of bursts cannot be controlled. In fact it could happen that a burst is dropped only because it overlaps with other bursts for a tiny amount of time, as shown in Fig.2. (This is not true if

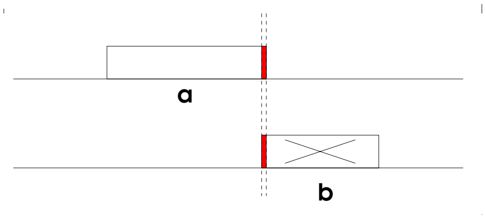

Fig. 2. Burst b is dropped because it overlaps with burst a.

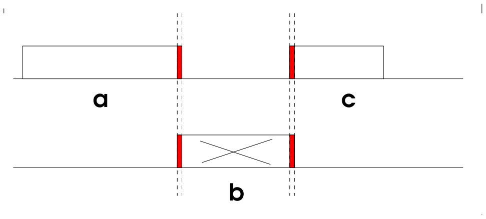

Fig. 3. Burst b is dropped because it cannot be fitted in the void between burst a and burst c.

burst segmentation is used, in which only the overlapping part of the one of the bursts is dropped. However, burst segmentation will cause some other problems such as more control traffic to inform the downstream routers about the new burst length, and should be regarded as the last resort.) On the other hand, if bursts arrive at the core routers at the beginning of time slots and are of the same length, burst will either not overlap or overlap completely. In this case, a burst is dropped only because it has to be dropped, i.e., every part of it overlaps with some other bursts.

Another important reason for supporting the aligning is that it will improve link utilization. Note that if burst arrives at random time instants, the void between two successive bursts are also random and cannot be utilized in some cases. As shown in Fig.3, burst b cannot be fit in the void between burst a and burst c even if its length only slightly exceeds the length of the void period. This results in poor link utilization. On the other hand, if all bursts are of the same length and bursts are aligned, the void between two bursts must be long enough to carry a burst.

B. Implementing SOBS

The implementation of SOBS is actually not very different from non-slotted OBS. To synchronize the arrival time at the core routers, FDLs of appropriate length can be appended to each link to make the link length multiples of a time slot, as shown in Fig.4, such that bursts will arrive at the core routers at the beginning of time slots. It should be mentioned that only the data burst needs to go through the appended FDL, while the control burst modulated on a separate wavelength can be filtered out and send to the core router directly. The biggest challenge in implementing SOBS is to synchronize all routers. This can be achieved by using high precision clocks at the routers and constantly adjusting the phase of the clock based on the difference between the arrival time of the bursts and the beginning of local time slots.

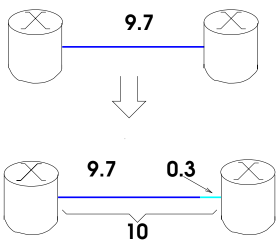

Fig. 4. After appending an FDL of length 0.3 to a link of length 9.7, the link length becomes 10, which is a multiple of the burst length.

The only major additional cost of the SOBS as compared to the non-slotted OBS is the additional FDLs appended to each link. However, this increase of cost could be very small, since the length of the FDLs is short if the length of a time slot is short, which is usually the case. The total number of FDLs is also small since only one FDL is needed for each *link*. In addition, as will soon be seen, other cost could be reduced in SOBS.

C. Other Benefits of SOBS

There are many other advantages of SOBS. The most celebrated one is that it can better support Quality of Service (QoS). OBS encounters great difficulties in supporting QoS because when bursts arrive asynchronously, low priority bursts could arrive earlier and could be allocated with resources which will block high priority bursts that arrive later. To solve this problem, many methods have been proposed, such as increasing the time offset between the control bursts and the data bursts of higher priority bursts such that the control bursts can be sent out earlier to increase the chance of successful resource reservation. However, this will cause longer delay time of the packets in the higher priority bursts. Note that, without any difficulty, this problem is completely solved in SOBS since all bursts arrive at the same time and the scheduler can always choose to serve the high priority bursts first.

Second, in SOBS, the cost of the switching fabric can be greatly reduced. In SOBS, since all bursts are the same length and arrive at the switch at the same time, switching fabric needs only to be *rearrangeably-nonblocking*. On the other hand, in OBS, the switching fabric has to be *strictly-nonblocking*, i.e., has to be able to connect any input to any output without disturbing the existing connections under any legitimate state. The cost of the rearrangeably-nonblocking switch can be much less than strictly-nonblocking switch. For example, in the three-stage Clos network, the former uses only half of the middle stage switches of the latter. The saving in the switching fabric can likely compensate for the cost of the additional FDLs, or, even make the overall cost of SOBS less than that of the OBS.

Third, since the data bursts are delayed by FDL but the control bursts are not, the scheduler has more time and can use more complex algorithms that will in turn improve the network performance.

Fourth, in SOBS, the control bursts can be much simpler and shorter, since it does not need to carry the information of the length of the data burst, or the exact time of the arrival of the data burst. All that is needed is the source-destination identification and the time slot of the arrival of the data burst. Using shorter control burst will reduce the probability of control burst collision and improve the reliability of the network.

III. Load Balancing Algorithm

In this section we will study another method for reducing the loss in OBS networks, which is to balance the load of the links. The majority works about OBS treat OBS much like a packet switching network and focus on the design and scheduling in core OBS routers. The edge routers only assemble packets into bursts and send it out to a predetermined route, often assumed as the shortest path to the destination. This "local greedy" approach lacks global coordination, and could lead to highly unbalanced link load which increases the burst loss probability. Therefore, to reduce the burst loss probability it is also important to find ways to balance the load of the links.

To the best of our knowledge, this was first realized by [16], in which two distributed load balancing algorithms were given. However, since centralized algorithms have better performance than distributed algorithms and it is not difficult to choose a node in the OBS network as the administrator, we focus on centralized algorithms. In [8], a centralized algorithm for load balancing was given in which the problem was formalized as an Integer Programming problem. However, the method in [8] has a very high complexity, because the number of variables in the integer programming problem is at least $O(n^4)$ where n is the number of nodes in the network, and the running time of their algorithm could be even higher. Also, since there is no polynomial time algorithm for solving integer programming problems, the solution given by [8] is not optimal, though may be very close. Since OBS networks is designed to support highly dynamic traffics, it is likely that the traffic intensities between node pairs change relatively quickly and thus the load balancing algorithm will have to be run faster. In this section we will give an algorithm that balances the traffic load and runs much faster than the algorithm given in [8].

The basic idea of our algorithm is simple. We assign costs to links which represent the amount of congestion of the links. The more the congestion, the higher the cost. And, for a source-destination pair, the Dijkstra's Algorithm is run to find a shortest path which is effectively the path which adds minimum amount of congestion to the network.

The cost of a link is calculated according to

$$C = (1 + \sum \rho_i)^{1.5} \tag{1}$$

where ρ_i is the traffic intensity of the i_{th} route that passes this link. Initially all links have cost 1. The algorithm will pick a source-destination pair and run the Dijkstra's Algorithm and find the shortest path that connects them. After the route is determined, the cost of the links will be updated. Then the algorithm picks another source-destination pair until all source-destination pairs have been assigned routes.

In our algorithm, longer distance source-destination pairs are picked earlier than shorter distance source-destination pairs, where the distance between a pair of nodes is defined as the fewest number of hops to reach from one node to another that can be found by a Breadth First Search (BFS). This is because after a link is involved in a route its cost will increase and it will become unlikely to be chosen by the Dijkstra's Algorithm later. If longer distance pairs are picked later than the shorter distance pairs, while the routes for the shorter distance pairs will likely remain short, the routes for the longer distance pairs could become even longer if some of their preferred links have been used by shorter distance pairs. Therefore, while the burst loss for shorter distance pairs will still be small, the burst loss for longer distance pairs could increase since the longer the burst has to travel, the more likely it will be lost. This unfairness toward longer distance pairs can be partly solved by finding routes for them first, such that their length of their routes will not increase. Of course in this case the length of the routes between the shorter distance pairs will increase. However, the overall effect will be a more balanced burst loss probability among all source-destination pairs.

There exist algorithms for finding routes in WR networks that are similar to ours. The difference, however, is how the costs of the links are calculated. In a WR network the link

TABLE I. Load Balancing Algorithm

```
Input : Set of node pairs $S$ and network topology.
Output: The routes found for all node pairs in $S$.
Use BFS to find the distances between all node pairs in $S$.
while there is still a node pair in $S$
    Let $s - d$ be the pair with longest distance.
    Run the Dijkstra's Algorithm, find the shortest path between $s$ and $d$.
    Let this path be the route for $s - d$ and update the cost of the involved links according to Equation (1).
    Remove $s - d$ from $S$.
end while
```

cost is usually in the form of $\frac{1}{k-l}$, where k is the number of wavelengths per fiber and l is the number of lightpaths that uses this link. Note that the cost will go to infinity when $l = k$, which essentially means that when there are k paths going through that link, the link is completely occupied and cannot be used by others. However, in our algorithm, when the link traffic grows, the cost will increase, but will not increase to infinity. This is because that the OBS network regards traffic as random, thus, when link traffic is large, though small, there is still a chance for a burst to pass the link on the event that all other nodes happen to have no burst at this particular time. This is the statistical multiplexing gain and is actually the major advantage of OBS network over WR network.

The load balancing algorithm is shown in Table I. The input to the algorithm is the network topology and the set of node pairs with non-zero traffic intensities, denoted as S. The output of the algorithm are the routes found for all node pairs in S.

IV. Burst Emission Control

In the effort to reduce the burst loss probability, we have also tried a technique called the *Burst Emission Control*. Burst Emission Control is to reduce the burst emission rate, i.e., the rate of of sending burst into the network, and control the timing of the emission, when the edge routers have detected high burst losses. It is very similar to the congestion control in the TCP, where TCP reduces packet sending rate by reducing the window size when some packets are not received by the other end. Burst Emission Control and its similarities with TCP congestion control were first realized in [7] and was shown effective in reducing the burst loss probabilities. In this section, however, we will present reasons to challenge the Burst Emission Control and argue that it is not an viable technique for improving the performance of OBS networks.

In electronic networks the core routers usually have large size buffers. The data streams sent by different TCP sessions are first written into the buffer to be queued and will exit the router when they have reached the head of the queues. Packet loss occurs only when the buffer overflows. Thus, once a packet loss is detected, TCP can respond with a reduced packet sending rate which will reduce the amount of buffer occupation in the core routers and thus effectively reduce the packet loss ratio. It is important to note that only the *rate*, not the *timing*, of the packet sending that matters, because the packets will be accumulated at the buffers of the core routers and one data stream affects other data streams only through the *accumulated* amount of packets it has sent during a certain period. This makes it relatively simple to implement congestion control, because the appropriate sending rate can be determined relatively easily by probing the network at only one end of the TCP session. However, in optical networks where there is no or very little buffer, to avoid collision the edge routers not only need to choose a suitable sending rate, but also have to choose the correct time to send the burst, because even if the sending rate is correct, if two routers happen to send their bursts at the same time (and the delays are also the same), the two bursts will still collide. Therefore, to make burst emission control successful, the edge routers will have to have full knowledge of when other routers are going to send their bursts in order to determine when to send its own, which requires a large amount of information exchange in real time and is impossible in any network of practical size.

However, as [7] shows, reducing the emission rate when burst loss is high can indeed reduce the burst loss ratio. Note that the burst loss ratio is reduced mainly because less bursts are sent: when every router chooses to send less bursts, the chance of collision will almost surely be reduced. However, this reduction in burst loss is actually not what is needed by the network, because it is at the price of a reduced link usage and a smaller overall throughput. When there are heavy traffic demands, the network should not ask all edge routers to send less bursts only to reduce the burst loss ratio, rather, it should find ways to utilize all the bandwidth and time slots available on the links and fill them with data.

V. Simulations

To compare the performance of the proposed schemes, we have conducted simulations on the well known NSF network shown in Fig.5. In the simulations the links in the network are assumed to be bidirectional with 8 channels, each operating at 10 Gbps. Uniform traffic is considered, i.e., the traffic statistics between all node pairs are assumed to be the same. The traffic arriving at one node destined for another node is assumed to be Poisson. For simplicity, in

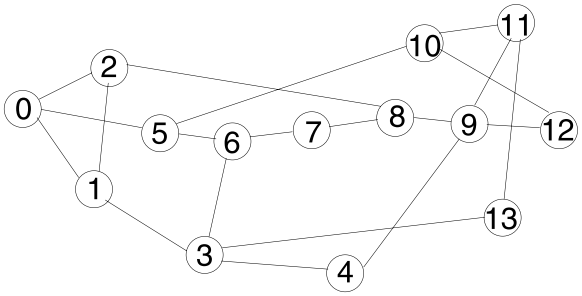

Fig. 5. The NSF network.

the simulations, the packets are assumed to be of fixed size and the size is chosen to be 10 Kb. The packets are assumed to arrive at the nodes in bursts, i.e., arrive in packet trains, and the length of the burst, L, is assumed to be a random variable following exponential distribution: $P(L=l)=2^l$, where l is a positive integer.

We have implemented OBS under three different schemes, namely the non-slotted OBS with shortest-path routing, the non-slotted OBS with load-balanced routing, and the slotted OBS with load-balanced routing, which will be referred to as SPOBS, LBOBS, and LBSOBS, respectively. In all schemes, there is a queue for each possible destination at each node. In both SPOBS and LBOBS, a burst is sent out immediately when the size of the queue exceeds 10 packets, or when a timer times out. In LBSOBS, however, a burst is sent out only at the beginning of a time slot when the size of the queue exceeds 10 packets, or when a timer times out. The performance of these schemes is measured by the *packet loss ratio* which is defined as the ratio of lost packets over all transmitted packets. We did not use the *burst loss ratio* because packet loss ratio is the more fundamental measure seen by the users.

Fig.6 shows the packet loss ratios of SPOBS, LBOBS, and LBSOBS as a function of the traffic intensity (number of packets arrived per second) between a node pair. It can be seen that both LBOBS and LBSOBS outperforms the SPOBS. The margin between the curves of SPOBS and the LBOBS is the gain of using the load balanced algorithm, however, it is somewhat surprising that the gain is very small. After examining the data we found that because the NSF network is such a well designed network, the routes found by the shortest path algorithm and the load balancing algorithm are quite similar, which results in the small gain. However, note that the LBSOBS performs significantly better than the SPOBS. In fact, the loss ratio is reduced by 40% to 70% under all traffic intensities, which supports our use of the slotted scheme.

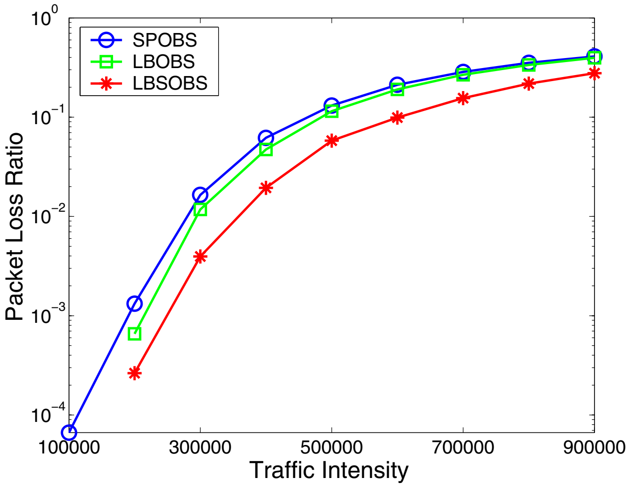

Fig. 6. Packet loss ratios of Shortest Path OBS (SPOBS), Load Balanced OBS (LBOBS), and Load Balanced Slotted OBS (LBSOBS).

VI. Conclusions

In this paper we studied the Slotted Optical Burst Switching (SOBS), where the network operates in a time-slotted manner and the bursts are fixed length. We have presented reasons in support of using SOBS and conducted simulations which show that the loss ratio of SOBS is much smaller than non-slotted OBS. We believe SOBS should be the design choice of future OBS networks.

Acknowledgements

This research work was supported in part by the U.S. National Science Foundation under grant numbers CCR-0073085 and CCR-0207999.

References

[1] J. Ramamirtham and J. Turner, “Time Sliced Optical Burst Switching,” *IEEE INFOCOM 2003*, San Francisco, CA, USA, Mar. 2003.

[2] F. Farahmand, V. M. Vokkarane and J. P. Jue, “Practical Priority Contention Resolution for Slotted Optical Burst Switching Networks,” *The First International Workshop on Optical Burst Switching (WOBS)*, Dallas, Texas, USA, Oct. 2003.

[3] L. Xu, H. G. Perros and G. Rouskas, “Techniques for optical packet switching and optical burst switching,” *IEEE Communications Magazine*, pp.136-142, Jan. 2001.

[4] C. J. Chang-Hasnain, Pei-cheng Ku, Jungho Kim, and Shun-lien Chuang, "Variable optical buffer using slow light in semiconductor nanostructures," *Proceedings of the IEEE*, vol.91, no.11, pp.1884 - 1897, Nov. 2003.

[5] M. Yoo and C. Qiao "A new optical burst switching protocol for supporting quality of service," *SPIE Proc. of Conf. All-optical Networking*, vol.3531, pp.396-405, 1998.

[6] C. Qiao and M. Yoo, "Optical burst switching (OBS) - a new paradigm for an optical Internet," *Journal of High Speed Networks*, 8(1):69-84, 1999.

[7] S.Y. Wang, "Using TCP congestion control to improve the performances of optical burst switched networks," *IEEE ICC '03*, May 2003, Anchorage, Alaska, USA.

[8] J. Zhang, B. Mukherjee, S. Wang, Y. C. Kim, K. Zhu, and D. Datta, "Pre-planned global rerouting for fault management in labeled optical burst-switched WDM networks," *IEEE Globecom '04*, Nov. 2004, Dallas, Texas, USA.

[9] V. Vokkarane, J.P. Jue, and S. Sitaraman, "Burst segmentation: An approach for reducing packet loss in optical burst switched networks," *IEEE ICC '02*, vol.5, pp. 2673-2677, 2002.

[10] J. Y. Wei and R. I. McFarland Jr., "Just-in-time signaling for WDM optical burst switching networks," *Journal of Lightwave Technology*, vol.18, no.12, pp. 2019-2037, Dec. 2000.

[11] R. Ramaswami and K. N. Sivarajan, *Optical Networks: A Practical Perspective*, 1st Ed., Academic Press, 2001.

[12] Q. Zhang, V. M. Vokkarane, B. Chen, and J. P. Jue, "Early drop and wavelength grouping schemes for providing absolute QoS differentiation in optical burst-switched networks," *IEEE Globecom '03*, pp.2694-2698, Dec. 2003.

[13] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-over-WDM networks," *IEEE Journal on Selected Areas in Communications*, vol.18, no.10, pp. 2062-2071, Oct. 2000.

[14] Z. Zhang and Y. Yang, "Optimal scheduling in WDM optical interconnects with arbitrary wavelength conversion capability," *IEEE Trans. Parallel and Distributed Systems*, vol.15, no.11, pp. 1012-1026, Nov. 2004.

[15] Z. Zhang and Y. Yang, "Scheduling in buffered WDM packet switching networks with arbitrary wavelength conversion capability," *IEEE INFOCOM 2004*, Hong Kong, March 2004.

[16] G.P.V. Thodime, V.M. Vokkarane, and J.P. Jue, *IEEE Globecom '03*, vol.5, pp.2628-2632, Dec. 2003.

[17] X. Yu, C. Qiao, and Y. Liu, "TCP implementations and false time out detection in OBS networks," *IEEE INFOCOM 2004*, Hong Kong, March 2004.

Fifth IEEE International Symposium on Network Computing and Applications (NCA'06)
0-7695-2640-3/06 $20.00 © 2006 **IEEE**