The HKU Scholars Hub   The University of Hong Kong   香港大學學術庫

Title	**Novel resource reservation schemes for optical burst switching**
Author(s)	**Li, CY; Lit, GM; Wai, PKA; Li, VOK**
Citation	**Ieee International Conference On Communications, 2005, v. 3, p. 1651-1655**
Issued Date	**2005**
URL	**http://hdl.handle.net/10722/45802**
Rights	**Creative Commons: Attribution 3.0 Hong Kong License**

Novel resource reservation schemes for optical burst switching

C.Y. Li*, G.M. Li†, P.K.A. Wai*, and Victor O.K. Li†

*Photonics Research Center and Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hong Kong, China
Tel: +852 2766-4094, Fax: +852 2362-8439, E-mails: {enli, enwai}@polyu.edu.hk
†Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, China
Tel: +852 2857-8425, Fax: +852 2559-8738, E-mails: {gmli, vli}@eee.hku.hk

Abstract— We propose to improve the throughput performance of optical burst switching by using regional controller nodes and window-based reservation. Both methods increase the information available to the intermediate nodes during scheduling decisions. Simulations show that the proposed reservation schemes provide significant improvement in the throughput performance compared with the original optical burst switching when the network is heavily loaded.

I. INTRODUCTION

Optical packet switching is one of the solutions proposed to satisfy the rapidly increasing communication bandwidth demand. In *true* all optical packet switching, both packet header processing and packet switching/forwarding are carried out optically to avoid the electro-optical bottleneck [1]. However, because of the lack of feasible optical logic devices and practical optical buffers, hybrid electro-optical approaches have been proposed in which packet switching/forwarding is still done optically but the packet headers are converted to the electrical domain for processing. As header processing techniques become more sophisticated, services such as traffic classification and priority routing will be possible in such networks. Among the different hybrid approaches, optical burst switching (OBS) has attracted much attention recently because of the potential of implementation using current technology [2]–[10].

OBS networks use one-way reservation to reserve the required resources for data transmission [3]. When a packet arrives at a node, it is stored in an electronic buffer until the number of packets with the same destination reaches a threshold value or the first packet in the batch exceeds the storage time limit. The node then sends out a control packet to the destination. After an *offset time*, a *data burst* containing the new packets are sent out following the path of the control packet. The control packet reserves the resources at the nodes on the path for the data burst. No acknowledgment is sent back to minimize electronic buffering and the data burst waiting time. If the reservation by the control packet is successful, the data burst will pass through all nodes on the path from source to destination without any processing and optical-electrical-optical (O/E/O) conversions. Thus no optical buffers are required and even slow optical switches can be used. Large propagation delay between nodes is no longer a major system performance concern. Testbeds of OBS using technology similar to wavelength division multiplexing have been demonstrated recently [4].

In its original form, OBS simply drops a data burst if the latter encounters output contention at an intermediate node. This may occur when the control packet of the data burst fails to reserve the necessary resources. Retransmission, if any, will be carried out by higher layers. The probability of loss in OBS increases rapidly when the system loading increases. Resource utilization in OBS networks is similar to that of random multi-access networks, *e.g.,* Aloha, in multihop environment [11]. In these networks, the throughput decreases rapidly when the system loading exceeds a threshold value because retransmissions can lead to a positive feedback and result in even more losses.

In order to improve network performance in OBS, many methods have been proposed to resolve contentions such as using optical buffers [3], deflection routing [5], and burst segmentation [6]. In traditional packet-switched networks, packets in contention are stored and scheduled according to their assigned priorities. This store-and-forward approach is not feasible in OBS networks because of the lack of practical optical buffers. For a 10 μs data burst which corresponds to 100 kbits of data in a 10 Gbps channel, an optical buffer constructed with fiber delay line (FDL) will be two kilometers long. Besides being bulky, a FDL based optical buffer also has fixed delay. Making practical FDL optical buffers for variable size packets remains a challenge, not to mention using them to schedule the contending data bursts. Deflection routing is not very useful in OBS networks either. In deflection routing, the network is typically synchronized and slotted. The number of inputs is equal to the number of outputs at a node. Therefore we can deflect (intentionally misroute) the contending packets to available outputs, which eliminates or minimizes the use of optical buffers [12]. Usually transit packets are given higher priority than new packets. Since the burst length is variable and the transmission is asynchronous in OBS networks, a transit

data burst may be blocked at a node by a new data burst that arrives slightly earlier [5]. The node may either preempt the transmission of new data bursts or store the transit data burst in optical buffers. Neither method is attractive. Recently, segmentation has been proposed for contention resolution in OBS networks [6]. A data burst may contain multiple packets. Transmitting the non-overlapping portions of the contending data bursts to the destinations may reduce retransmissions but the reduction is achieved at the expense of increased signaling complexity and the need for optical processing.

One can reduce the number of contentions by improving system bandwidth utilization. For example, one may use adaptive multi-path routing to smooth the traffic distribution among the links. In this paper, we propose to enhance the performance of OBS by improving the resource reservation mechanism. We will keep the optical switching/forwarding of the data burst simple but allow the complexity of the electronic processing of the control packets to increase so that the proposed schemes can still be implemented with current technology. In Section II, we review the data burst contention problem. We observe that the lack of reservation information of neighboring nodes and the difference in the arrival sequence between the control packets and data packets at a node lead to inefficient resource reservation in OBS. We propose to use regional controller and window-based reservation in Sections II-A and II-B, respectively, to improve the network performance. We study the performance of the two proposed schemes in Section III and find that they can improve system utilization when compared to that of the original OBS even if the system is heavily loaded. We conclude in Section IV.

II. OUTPUT CONTENTION AND RESOURCE RESERVATION

Output contention is not new in packet switched networks but the problem is especially serious in OBS networks because the size of the data bursts is variable, the nodes are not synchronized, and there is no coordination between data bursts from different nodes. In the following, we show that system utilization can be greatly increased by improving the resource reservation mechanism in OBS.

A. Regional controller nodes

In order to provide optical packet switching services with current technology, OBS decouples the resource reservation and the data transmission processes. This is similar to the common channel signaling (CCS) that has been used to separate the call setup control and the voice transmissions paths in traditional telephone networks. CCS uses two-way reservation [13]. The source waits for the reservation acknowledgment before commencing communication. The decoupling of control and transmission in CCS is more complete than that of OBS. The CCS call setup control signals can be transmitted and processed in dedicated networks that are completely separated from the networks transmitting the voice/data. An OBS control packet and its corresponding data burst are transmitted at an offset time apart but both share the same routing path. Thus CCS can make more *intelligent* decisions than OBS.
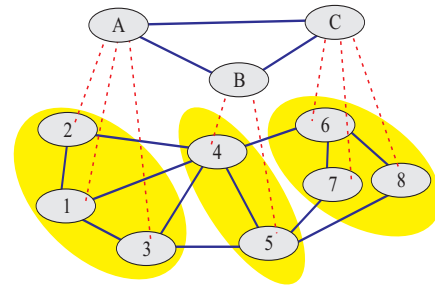

Fig. 1. A communication network and its control network. The communication nodes in the sets {1,2,3}, {4,5}, {6,7,8} are under the supervision of control nodes A, B, and C, respectively. Each control node has complete information and control of the nodes in its own region.

CCS however becomes bandwidth inefficient if the end-to-end propagation delay is much larger than the communication period. OBS on the other end has no significant performance degradation under the same condition.

Figure 1 shows a communication network and its control network that is dedicated for the handling and transportation of the control signals. The communication nodes are labeled by numerals and the control nodes by letters of the alphabet. The communication nodes in the sets {1,2,3}, {4,5}, {6,7,8} are under the supervision of control nodes A, B, and C, respectively. Each control node has complete information and control of the nodes in its own region. In OBS, the reservation is carried out node by node. The status of the next node in the path is not available. Optimization of the routing path is difficult. For example, a control packet from Node 1 to Node 8 may still try to reserve channels in link $(1, 3)$ even though all the outputs of Node 3 are not available. In this case, the data burst will be dropped when it reaches Node 3. With the control network, since each control node is aware of the status of the communication nodes in its own region, another routing path $1 \rightarrow 4 \rightarrow 6 \rightarrow 8$ can be used instead. Using a control network, the status of all nodes in the same region is considered during reservation. The routing path can be optimized among the nodes inside a region. System resource utilization can be improved even if one-way reservation is used. Services such as quality of service (QoS) is also simpler to implement.

With regional controller nodes in OBS, the control packets of the data bursts are sent to the network formed by the regional controller nodes. In principle, the resource reservation can be similar to that of the original OBS but now are done controller node by controller node instead. Furthermore, the controllers can be physical or logical. A logical controller is formed by a set of nodes with relatively small propagation delay between them. The time required for exchanging status information and commands between the nodes forming the logical controller will be negligible when compared to the control packet processing time. Due to page limitations, issues such as the optimal control network topology, routing of control packet, and implementation of logical regional controllers will not be further discussed here. In Section III, we use simulations to study the performance of OBS using logical regional controller nodes. A control packet is assumed to be
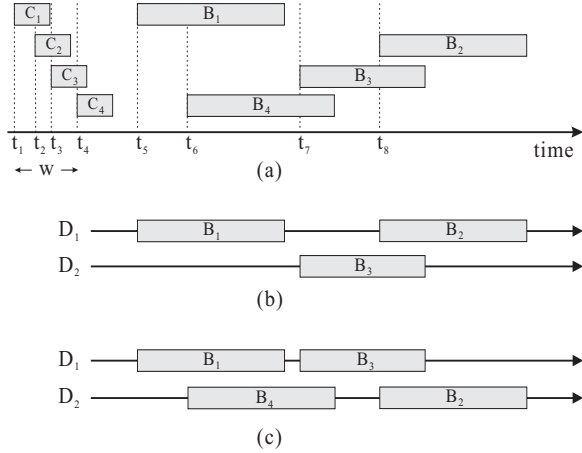
Fig. 2. (a) Arrivals of control packets C_1 to C_4 and the corresponding data bursts B_1 to B_4. (b) The channel assignment according to the order of arrival of the control packets. (c) The channel assignment according to the order of arrival of the data bursts.

processed only once at the boundary nodes of a region to determine the routings inside the region.

B. Window-based reservation

In the original OBS networks, after the control packet is sent, a data burst has to wait for an offset time at the source node before transmission because of the processing delay of the control packets at intermediate nodes. The minimum offset time that has to be set at the source node is Hc_x, where H is the total number of intermediate nodes on the path and c_x is the required processing time for the control packet at a node. As the control packets and data bursts pass the intermediate nodes, the offset times between them varies. The minimum offset time between a pair of control packet and data burst reduces to $(H - h)c_x$ if the control packet has passed the h-th intermediate node. Sometimes, the offset time may be set to a value larger than that of the minimum to gain more control on the data bursts. For example, the offset time has been lengthened in order to provide QoS classes to the data bursts [7]. As a consequence, the variation of the offset times can be large, and the arrival sequence of the data bursts at a node can often be different from that of the control packets, leading to inefficient resource scheduling.

Figure 2(a) shows the arrivals of four control packets and their corresponding data bursts. We assume that these four data bursts request the same output D with wavelength channels D_1 and D_2. The control packets are marked C_1 to C_4 in the order of their arrivals. The data bursts are marked B_1 to B_4 following their corresponding control packets. The arrival sequence of the data bursts is assumed to be B_1, B_4, B_3, and B_2. We assume that both wavelength channels are idle when the first control packet arrives. At time t_1, control packet C_1 arrives at the node and is assigned channel D_1 for its data burst B_1 with starting time t_5. To minimize the idle periods between transmissions, the data burst B_2 is assigned channel D_1 at time t_2 with starting time t_8 [2], [8]. Consequently, data burst B_3 is assigned channel D_2 at time

t_3 with starting time t_7 but contention problem occurs for the data burst B_4. As shown in Figure 2(b), data burst B_4 has to be dropped if there is no other contention resolution scheme. If the channel assignment is based on the order of arrival of the data bursts rather than that of the control packets, the arrangement shown in Figure 2(c) will be obtained. Data bursts B_1 and B_3 are assigned to channel D_1 while channel D_2 is for the data bursts B_2 and B_4. Data burst B_4 will not be dropped.

Figure 2 shows us an example of burst dropping caused by insufficient information. One way to solve the problem is to make the arrival sequence of the data bursts the same as that of the control packets. This can be achieved by requiring the data bursts to go through at each intermediate node a fixed delay that is at least the maximum processing time for the control packet at the intermediate nodes [2], [3]. Another approach is to delay the channel assignments so that more data burst arrival information is available. In Figure 2, the node will correctly assign the channels if all the scheduling decisions can be delayed w seconds, where $w = t_4 - t_1$ is the time window required for such delayed scheduling decisions. Examples of such approach include pipeline buffering [9] and virtual fixed offset time [10]. Adding a time window W will cause the control packets to stay longer at a node, i.e., $c_x + W$ instead of just the processing time c_x. The common assumption is that sufficient FDLs have been installed at the node to delay the data bursts such that they will not overtake the control packets at the intermediate nodes [9], [10]. With FDLs installed at every nodes, a data burst will have a total buffering delay of HW from source to destination. The original OBS used the offset time between the control packets and data bursts to eliminate the use of optical buffers at intermediate nodes. Following the same spirit in window-based reservation scheme, we can avoid increasing the optical hardware complexity at the intermediate nodes by simply increasing the offset time. Hence, the minimum offset time between a pair of control packet and data burst under the proposed window-based reservation becomes

$$T_h = (H - h)(c_x + W), \qquad (1)$$

if the control packet has passed h intermediate nodes, and $T_0 = H(c_x + W)$ is the minimum offset time that should be set at the source node. The time window W can be integral multiples of c_x but its optimal value is network dependent. The required optical hardware complexity will be similar to that of the original OBS, i.e., no FDLs are required. This approach increases the resource reservation flexibility but it also increases the variation of the offset times of the data bursts arriving at a node. We shall study the performance of this window-based reservation approach in OBS using simulations with different c_x and W in Section III.

III. PERFORMANCE EVALUATION

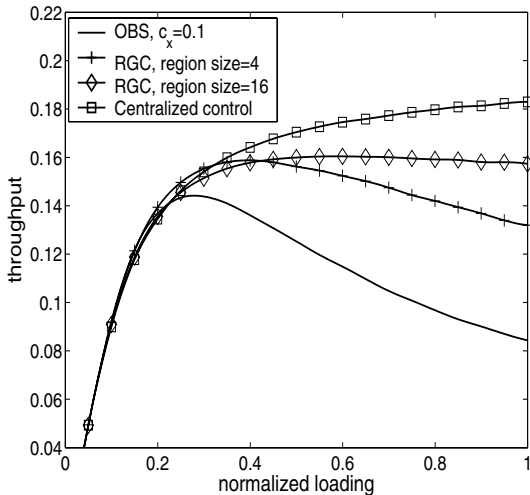We use simulations to investigate the performance of the two proposed OBS reservation schemes: the regional controller

Fig. 3. The throughput of the original OBS and OBS with regional controllers (RGC) of an 8×8 MSN. The processing time of a control packet c_x for the original OBS is equal to a tenth of data burst transmission time, *i.e.,* 0.1 time unit. Those for RGC with region sizes of 4, 16, and 64 (centralized control) are 0.2, 0.4, and 0.8 time units, respectively.

Fig. 4. The throughput of the original OBS and OBS with regional controllers (RGC) of an 8×8 MSN. The processing time of a control packet c_x for the original OBS is equal to a data burst transmission time, *i.e.,* one time unit. Those for RGC with region sizes of 4, 16, and 64 (centralized control) are 2, 4, and 8 time units, respectively.

(RGC) and the window-based reservation (WBR). Similar to that in [3], we use an 8×8 Manhattan Street Network (MSN) [14]. The traffic arrival is in the form of data bursts. The arrival of data bursts is a Poisson process. When a new data burst arrives at a node, it randomly chooses a destination from the rest of the nodes in the network. The length of each data burst (transmission time) is an exponentially distributed random variable with unit mean. The propagation delay of a link is assumed to be 62.5 time units, similar to that used in [3]. The normalized load offered to a node is the ratio of the average data burst length to the inter-arrival time normalized by the number of channels per link. For comparison, we also include the results of OBS with just-enough-time (JET) reservation [3]. In the simulations, we assume four channels per link and all nodes receive the same offered load. All simulations are run sufficiently long such that the 95% confidence intervals are smaller than 1% of the results.

After a data burst arrives at a node, a control packet is sent out immediately to reserve the required resource on the path. The data burst is then transmitted after an offset time. In the original OBS, the offset time is equal to the total processing time of the control packet at intermediate nodes. The offset time in RGC OBS is the total time that the control packet spends at the controllers. We assume logical regional controllers with square regions of 4, 16, and 64 nodes. Thus we have RGC OBS with 16, 4, and 1 regional controllers, respectively, in the network. A network with a single controller means centralized control. Since a packet often has d intermediate nodes when it passes a square region of d^2 nodes, we assume that the processing time in a controller with square region of d^2 nodes is $d \times c_x$, *i.e.,* $2c_x$, $4c_x$, and $8c_x$ for RGC OBS with 16, 4, and 1 controllers, respectively. Hence, the average offset time of RGC OBS is larger than that of the original OBS. For WBR OBS, the offset time is initially
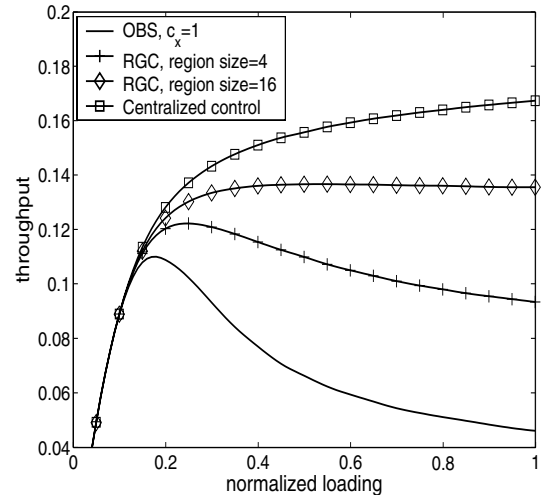
set to $H(c_x + W)$ and is decreased node by node according to Eq. (1). In this simulation, shortest path routing is used for both control packets and data bursts.

Figures 3 and 4 show the throughput performance of RGC OBS with different region sizes and control packet processing times. The processing time of a control packet c_x in the original OBS is equal to one tenth and one data burst transmission time, *i.e.,* 0.1 and 1 time units. Those for RGC OBS with region sizes of 4, 16, and 64 (centralized control) are 0.2, 0.4, and 0.8 time units in Figure 3, and 2, 4, and 8 time units in Figure 4, respectively. The throughput of the original OBS and the RGC OBS with region sizes of 4, 16 and 64 (central control) are plotted in solid line, and lines with pluses, diamonds, and squares, respectively. Figures 3 and 4 show that the performance of the original OBS (JET reservation without any contention resolution) degrades rapidly in the 8×8 MSN if the system is heavily loaded. One reason is that the small output degree of MSNs results in congestion. Simulations using other network topologies show that the throughput can be stabilized and does not decrease when system loading increases. Figures 3 and 4 show that RGC OBS has better throughput performance than the original OBS when the normalized loading is larger than about 0.2 in Figure 3 and 0.15 in Figure 4. As shown in the figures, RGC OBS with different region sizes have similar performance in lightly loaded systems. For the same normalized loading in heavily loaded systems, a larger region size has better throughput performance. The processing time however also increases accordingly. Similar results have been observed in simulations with different processing times and network topologies.

Figures 5 and 6 give the throughput performance of *simple* WBR OBS using different window sizes and control packet processing times. In the simulations, an OBS node with WBR checks all control packets that arrive during the time
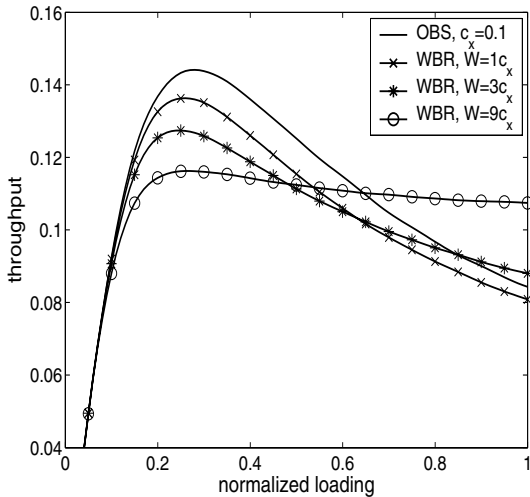
Fig. 5. The throughput of the original OBS and OBS using window based reservation (WBR) with window size $W = 1$, 3, and 9 times of the packet processing time c_x of the original OBS. c_x is equal to 0.1 data burst transmission time.

Fig. 6. The throughput of the original OBS and OBS using window based reservation (WBR) with window size $W = 1$, 3, and 9 times of the packet processing time c_x in normal OBS. c_x is equal to the transmission time of one data burst.

window. Then the node assigns the channels first-come-first-served (FCFS) according to the estimated data burst arrival sequence. Significant performance improvement can be obtained if sophisticated optimization techniques are used. In Figures 5 and 6, the lines with crosses, asterisks, and circles are the throughputs for WBR OBS with window size W = 1, 3, and 9 times the control packet processing time c_x. The solid lines are the results of the original OBS and included for comparison. We observe that WBR is more suitable for systems with large processing time and heavy loading. We find that the performance of the original OBS degrades only slightly with the increase in offset time if the offset time is sufficiently high. WBR increases the offset time variation and therefore causes lower throughput in lightly loaded systems. However, the advantage of delaying the scheduling decision becomes significant when the system loading increases. From Figures 5 and 6, the optimal window size W depends on the network loading as well as the control packet processing time. An adaptive algorithm is required to apply WBR in practical situations.

IV. CONCLUSION

One of the advantages of optical burst switching is that it can provide optical packet switching services with currently available technology. Due to the lack of feasible optical processing devices, data burst contention resolution in the optical domain is difficult. One can reduce data burst contentions by better utilizing the system resource. We propose to improve the resource reservation mechanism in OBS by using regional controllers or window-based reservations. Both approaches increase the complexity of the electronic processing of control packets at OBS nodes but not the optical switching/forwarding of the data bursts. Simulations using a Manhattan Street Network show that the proposed methods can improve the
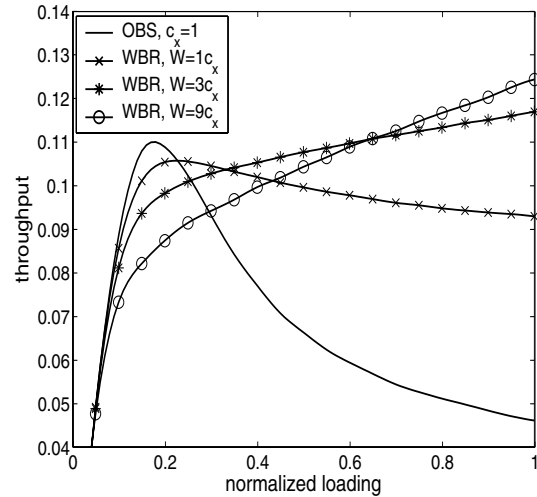
throughput performance of OBS even if the system is heavily loaded.

REFERENCES

[1] X.C. Yuan, V.O.K. Li, C.Y. Li and P.K.A Wai, "A novel self-routing scheme for all-optical packet switched networks with arbitrary topology," *IEEE Journal of Lightwave Technology*, Vol. 21, No. 2, pp. 329-339, 2003.

[2] J.S. Turner, "Terabit burst switching," *Journal of High Speed Networks,* Vol. 8, pp. 3-6, 1999.

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

[4] Y. Sun, T. Hashiguchi, N. Yoshida, X. Wang, H. Morikawa, and T. Aoyama, "Architecture and design issues of an optical burst switched network testbed," *Proceedings of OECC/COIN 2004,* pp. 386-387, 2004.

[5] C.F. Hsu, T.L. Liu, and N.F. Huang, "Performance analysis of deflection routing in optical burst-switching networks," *Proceedings of IEEE Infocom 2002,* pp. 1803-1812, 2002.

[6] A. Detti, and M. Listanti, "Impact of segments aggregation on tcp Reno flows in optical burst switching networks," *Proceedings of IEEE Infocom 2002,* pp. 1803-1812, 2002.

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

[8] Y. Xiong, M. Vandenhoute, and H.C. Cankaya, "Control architecture in optical burst-switched WDM networks," *IEEE Journal on Selected Area in Communications,* Vol. 18, No. 10, pp. 1838-1851, 2000.

[9] H.L. Li, H.M. Neo, and T.L.J Ian, "Performance of the implementation of a pipeline buffering system in optical burst switching networks," *Proceedings of Globecom 2003*, Vol. 5, pp. 2503-2507, 2003.

[10] J.K. Li, C. Qiao, J.H. Xu, and D.H. Xu, "Maximizing throughput for optical burst switching networks," *Proceedings of INFOCOM 2004,* Vol. 3, pp.1853-1863, 2004.

[11] F.A. Tobagi, "Multiaccess protocols in packet communication systems," *IEEE Transactions on Communications,* Vol. 28, No. 4, pp. 468-488, 1980.

[12] C.Y. Li, P.K.A. Wai, X.C. Yuan, and V.O.K. Li, "Deflection routing in slotted self-routing networks with arbitrary topology," *IEEE Journal of Selected Area on Communications*, Vol. 22, No. 9, pp. 1912-1822, 2004.

[13] A.R. Modarressi, and R.A. Skoog, "An overview of Signaling No. 7," *IEEE Proceedings,* Vol. 80, No. 4, pp. 590-606, 1992.

[14] N.F. Maxemchuk, "Routing in Manhattan Street network," *IEEE Transactions on Communications*, Vol. 35, No. 5, pp. 503-512, 1987.