

# Design of Optical Burst Switches based on Dual Shuffle-exchange Network and Deflection Routing

Man-Ting Choy  
Department of Information  
Engineering, The Chinese  
University of Hong Kong  
[mtchoy1@ie.cuhk.edu.hk](mailto:mtchoy1@ie.cuhk.edu.hk)

Yun Deng

Department of Information  
Engineering, The Chinese  
University of Hong Kong  
[ydeng1@ie.cuhk.edu.hk](mailto:ydeng1@ie.cuhk.edu.hk)

Tony T.Lee

Department of Information  
Engineering, The Chinese  
University of Hong Kong  
[tlee@ie.cuhk.edu.hk](mailto:tlee@ie.cuhk.edu.hk)

## Abstract

*In this paper, we propose a novel approach to implement Optical Burst Switching (OBS) using the Dual Shuffle-exchange Network (DSN) as the core switching fabric. DSN possesses the self-routing property which allows major simplifications on the complex crossbars setup mechanisms. In addition, its asynchronous and buffer-less natures are highly preferable in the optical environment. We also show that with an appropriate error-correcting routing algorithm, the output wavelength contentions can be reduced by means of internal deflection routing.*

## 1. Introduction

The optical burst switch (OBS) [1], [2] has been highly regarded as a viable solution on providing terabit switching in the near future because of its easy implementation, high bandwidth utilization and flexibility. In optical burst switching, multiple packets are aggregated into a larger burst at the source before sending to the destination. Bandwidth is reserved in each intermediate node by one-way protocols in which data burst are sent after its control packet without waiting for the acknowledgment. As the control and data are sent separately, no buffering in the intermediate nodes are needed to store data temporarily while the control packet is being processed. This is highly preferable as optical RAM development is still in its early stage. In fact, the OBS compromises the circuit and the packet switching schemes. The overhead caused by connection setup in circuit switching and the headers in packet switching are remedied by means of the one-way reservation and the burst aggregation, respectively. Therefore, the OBS scheme can also provide circuit or packet switching when

necessary, simply by adjusting the reservation scenario and the burst length of the scheme.

A challenging issue on OBS is to reduce the blocking caused by output contention at each switching node. Output contention occurs when there are too many input bursts destined for the same output fiber simultaneously. If no buffering is provided, the bursts which have lost contention must be dropped. There has been a great deal of research work addressing this issue in recent years [3] – [5], and most focused on the channel scheduling and the signaling protocol. Both approaches show satisfactory improvement on the blocking performance, but these enhancements come at a price of high complexity in terms of both hardware architecture and scheduling algorithm. On the other hand, the design of core switching fabric is less mentioned and explored. Two switching fabric architectures for optical burst switching have been discussed in [6]. The first one associates the tunable wavelength converters with the optical crossbars, passive multiplexers and demultiplexers. The second design replaces the optical crossbars of the first design with passive wavelength grating routers (also known as arrayed waveguide grating multiplexers or AWGMs). However, these architectures require complicated scheduling algorithms that make them less appealing especially in high-speed optical networks.

In this paper, we propose a novel approach to implement the OBS switch fabric by the dual shuffle-exchange network (DSN) [7]. The DSN was originally developed to support packet switching. It uses the idea of deflection routing to solve the problem of packet contention. As a result, no buffer is needed to queue the packets at the input or inside the network. Packets can be properly routed through the DSN simply by a routing tag. This implies that the network does not need a central controller and thus is highly scalable. It can be shown that the complexity of DSN is of the order  $N \log N$ , which is above the Shannon's lower bound on switch complexity.

We discovered that DSN's buffer-less and asynchronous natures are highly preferable in the OBS environment. Its self-routing property can substantially reduce the system complexity while still achieving low blocking probability and high throughput.

We further improve the switch using a multi-plane architecture, which consists of  $k$  DSNs in parallel. Assuming that the arriving bursts at the switch can be evenly distributed among the planes, the burst deflection in each plane can then be greatly alleviated and thus the required number of stages for the error correction can be significantly reduced. We derive the optimal value of  $k$ , as a function of the loss rate bound, such that the overall complexity of the switch is minimized.

The rest of the paper is organized as follows. In Section 2, we give a brief overview on the basics of Dual Shuffle-exchange Network. We present our OBS architecture and discuss some implementation issues in Section 3. In Section 4, we model the blocking in two designs of the switching modules inside the DSN switch and present simulation results on the blocking characteristics of different switch designs. In Section 5, we analysis the performance of the different parallel  $k$  planes and give some results of simulation. Finally, we conclude our work in Section 6.

## 2. Basics of Dual Shuffle-exchange Network

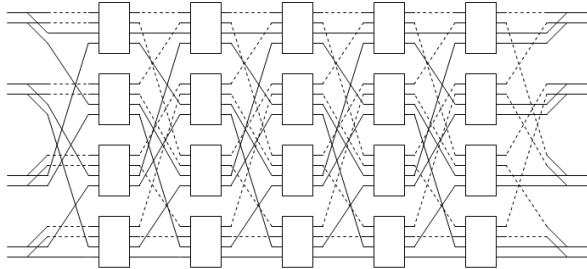


Figure 1. A  $8 \times 8$  Dual Shuffle-exchange Network.

Figure 1 shows an  $8 \times 8$  Dual Shuffle-exchange Network with 5 stages. We define  $N$  as the number of input ports of the switch and  $L$  be the number of stages,  $N=8$  and  $L=5$  in this case. The  $4 \times 4$  switching modules inside the network are interconnected in this way such that packets from any input ports can self-route to its destination port in  $n=\log_2 N$  stage if there is no packet contention. This connection pattern also permits an error-correcting routing algorithm. When contention occurs inside a  $4 \times 4$  switching module, the loser packet will be deflected to one of the idle output ports available. A one-stage routing instruction will be added to this packet based on which output port this packet is deflected. By successfully following this routing instruction in the next stage, the deflected packet can return to the state where it was deflected and resume its routing. Successive

deflections can also be corrected by this algorithm. Figure 2 shows the state-transition diagram of this error-correcting algorithm. Each state represents the number of remaining stages the packet still have to go through until it reaches its destination. Theoretically, all loss probability requirements ( $P_{\text{loss}}$ ) can be achieved by means of increasing  $L$  and it has been shown [7] that  $L$  is analytically bounded by

$$L \leq 2.793n - 3.554 \ln(n+1) + 3.554 \ln P_{\text{loss}}^{-1} + 1.162$$

Since each stage consists of  $N/2$  switch modules, the complexity of the DSN for a given  $P_{\text{loss}}$  is therefore  $N \log N$ .

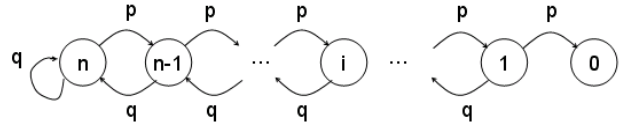
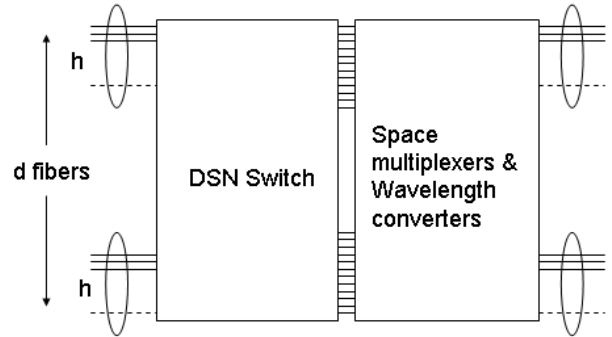


Figure 2. Markov Chain for bounding  $L$ .

## 3. Proposed Architecture

### 3.1. General Architecture.

(a)



(b)

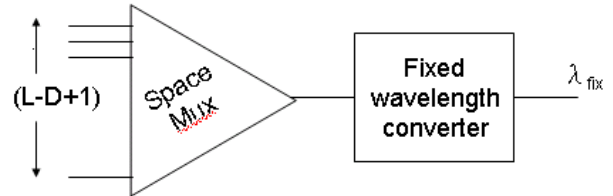


Figure 3. (a) Block diagram of the proposed architecture. (b) One module in the space multiplexer and wavelength converter unit.

Figure 3(a) shows our proposed switching node architecture. There are  $d$  incoming and  $d$  outgoing optical fibers, each contains  $h+1$  wavelength channels. One of the channels is reserved for control signals. The central switching fabric is a  $dh \times dh$  DSN switch. According to the JIT [8] signaling protocol, the basic switching

modules inside the switch are directly reserved for the incoming burst immediately after the arrival of the request message, and remain until the arrival of a release message. With the deflection routing property of the DSN, burst can exit from the network beginning from the  $D^{\text{th}}$  ( $D=\log_2 d$ ) stage. Therefore for each fiber, we have  $h(L-D+1)$  output ports but only  $h$  output wavelength channels available. For simplicity reason, we propose to have  $h$  space multiplexers for each fiber and each space multiplexer would be serving  $(L-D+1)$  input ports, as shown in Figure 3(b). Whenever a fully-routed burst finds its corresponding multiplexer is being occupied, this burst would have to be deflected as if it is deflected due to contention inside the network. When this burst is fully-routed again, it can try another multiplexer. Bursts leaving the multiplexers would go through fixed wavelength conversion at the end of the switch.

### 3.2. Implementation issues on the $4 \times 4$ switching module.

For each  $4 \times 4$  switching module, four output links are connected to the next stage while another four output links are connected to the space multiplexers. Therefore, 4 additional  $1 \times 2$  switches are needed as illustrated in Figure 4.

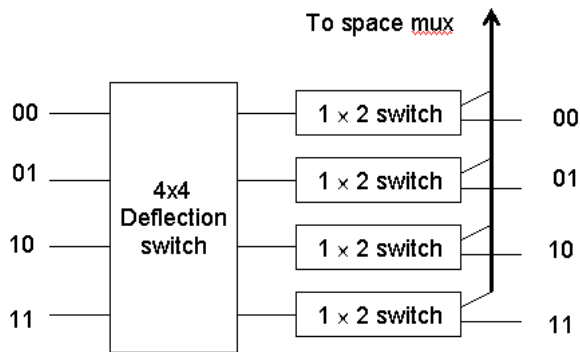


Figure 4. Block diagram of a  $4 \times 4$  switch module.

If we implement the  $4 \times 4$  switching module using a non-blocking crossbar switch, a total of 16 crosspoints are needed. [7] shows that in the packet switching environment, a two-stage banyan switch as shown in Figure 5(a), is a possible alternative. Granted that a burst may be deflected because of “internal conflict” even when there is no contending burst for the same “external output” (see Figure 5(b)), the increase in the overall deflection probability is actually quite small. The non-blocking and banyan design alternatives will be compared in more detail in the section 4.

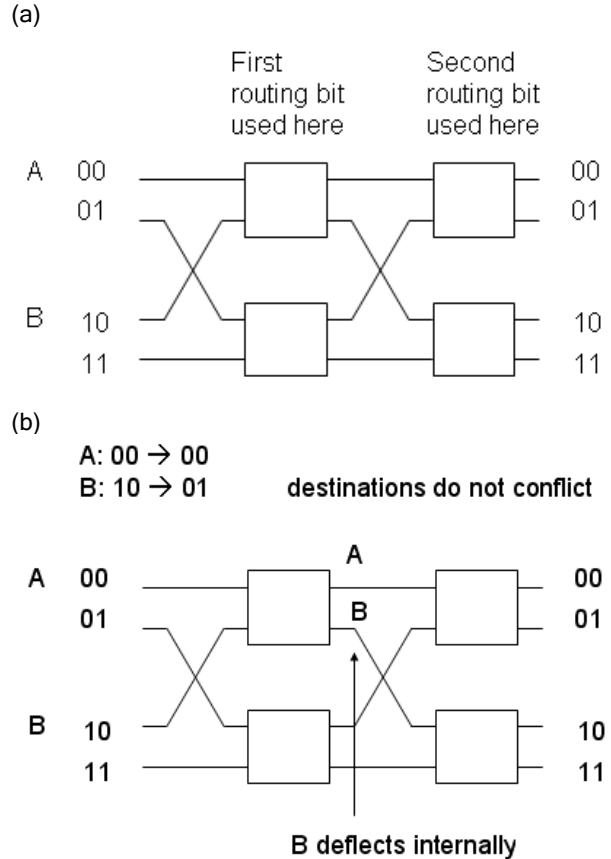


Figure 5. (a) A  $4 \times 4$  banyan deflection switch; (b) An example of internal conflict when there is no output conflict..

## 4. Analysis and Simulation on single-plane architecture.

This section compares the deflection probability between non-blocking and banyan deflection switch using an appropriate analysis. Simulation results verifying the validity of the analysis and the effects of  $L$  on the blocking probability are presented.

### 4.1. Analysis: Non-blocking versus banyan.

Let us first derive the deflection probability of an internally non-blocking switching module. When a new burst arrives at this non-blocking switching module, it will see at least one output port is idle, or equivalently 0 to 3 output ports are being occupied. Assuming each burst is equally likely to destine to one of the four output ports, the burst length is exponential with mean  $\mu$  and idle time is exponential with mean  $\lambda$ , figure 6 depicts a markov chain for this  $4 \times 4$  non-blocking switching module. The states of the markov chain are the number of busy output ports seen by a new arriving burst. After obtaining the

limiting probabilities  $\pi_i$ , the deflection probability for non-blocking switching module can be calculated as follows

$$q_n = 1 - \left( \pi_0 + \frac{3}{4}\pi_1 + \frac{2}{4}\pi_2 + \frac{1}{4}\pi_3 \right)$$

$$= 1 - \frac{\lambda(\lambda + \mu)^3}{(\lambda + \mu)^4 - \mu^4}$$

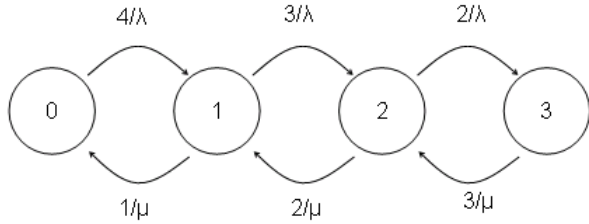


Figure 6. Markov Chain for Non-blocking switch.

Similarly, the markov chain in Figure 7 depicts one of the crosspoints in Figure 5. With the same set of parameters as above, we obtain the deflection probability for the whole banyan switch as follows

$$q_b = 1 - \left( \pi_0 + \frac{1}{2}\pi_1 \right)^2$$

$$= 1 - \left[ \frac{\lambda(\lambda + \mu)}{(\lambda + \mu)^2 - \mu^2} \right]^2$$

By differentiating the difference between  $q_n$  and  $q_b$ , we obtain the maximum absolute difference is no bigger than 0.115 at  $\lambda/\mu = 2.487$ .

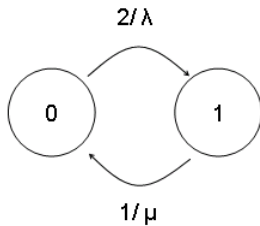


Figure 7. Markov Chain for a 2 × 2 crosspoint.

#### 4.2. Simulation.

In this section, we validate our analytic results by simulation. Using 16 fibers, 64 wavelength channels and input loading of 0.5, Figure 8 plots the simulated loss probability as a function of the number of available stages of the switch. From the graph, we see that the number of stages needed for a fixed  $P_{loss}$  in the banyan design is slightly higher than that in the non-blocking design. Besides, it is observed that both curves drop significantly

at the middle while turn flat at the end. The former is due to the fact that some bursts have already left the switch in some previous stages, resulting in smaller traffic load and therefore smaller deflection probability at the back of the switch. The latter, on the other hand, is due to bursts that suffer from output contentions, that is more than 64 inputs are destined to the same output fiber. This problem requires more number of stages to tackle. Figure 9 plots the loss probability versus the input traffic load using 8 fibers and 64 wavelengths for  $L = 16, 32$  and  $64$  respectively..

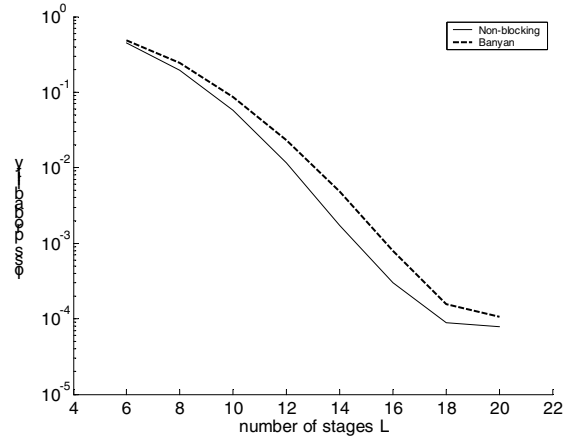


Figure 8. Loss probability versus number of stages for banyan and non-blocking switch.

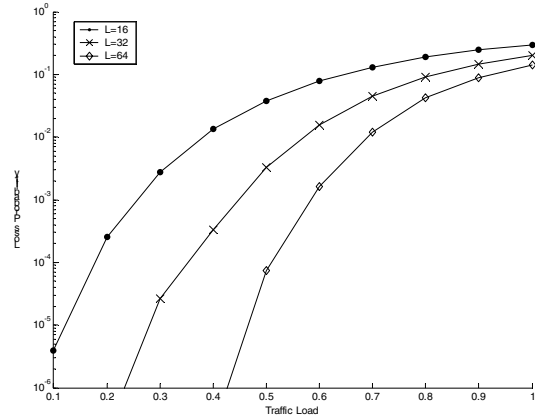


Figure 9. Loss probability versus input traffic load for various  $L$ .

### 5. Analysis and Simulation on multi-plane architecture.

From the results of Section 4, we notice that a slight reduce of traffic load may results in a significant drop in the blocking probability, especially when the applied traffic load is already small. We believe that by dividing

the input traffic into  $k$  different planes, the blocking probability inside each plane would be largely reduced and so is the number of stages required. A  $N \times N$  switching architecture with the multi-plane fabric is shown in Figure 10. If the loading of each input port is  $\rho$ , after evenly divided by demultiplexers, the input loading of each  $N \times N$  switching plane  $\rho' = \rho/k$ .

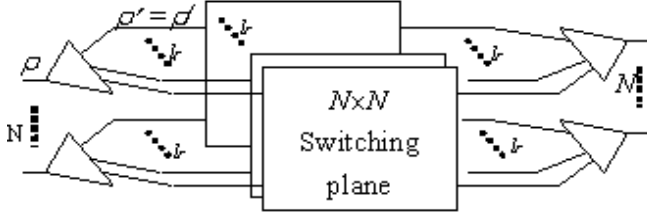


Figure 10. The architecture of multi-plane fabric

### 5.1 Performance analysis of multi-plane architecture

In our burst traffic model, we assume burst time-length is exponential distribution with mean  $\square$  and idle time-length is exponential distribution with mean  $\lambda$ . We define the loading of each input of the module  $\rho = \mu/(\mu + \lambda)$ .

As presented in Section 4.1, the deflection probability of an internally non-blocking switching module is

$$q = 1 - \frac{\lambda(\lambda + \mu)^3}{(\lambda + \mu)^4 - \mu^4} = \frac{\rho(1 + \rho + \rho^2)}{(1 + \rho^2)(1 + \rho)}$$

To analyze the delay and complexity of multi-plane fabric compared with single-plane switch, we let  $T_i$  denote the expected additional number of stages would have taken by a burst in state  $i$  before reaching its destination. The expected number of stages required by the initial state  $n$  ---  $T_n$  is the average delay of burst in one switch node and it can represent the required complexity of switch to some extent. As illustrated in Figure 2, we have

$$T_n = \frac{n}{p-q} - \left( \frac{1}{p-q} - \frac{1}{p} \right) \left[ \frac{1 - (q/p)^n}{1 - q/p} \right]$$

$$\text{with } q = \frac{\rho(1 + \rho + \rho^2)}{(1 + \rho^2)(1 + \rho)}, p = \frac{1}{(1 + \rho^2)(1 + \rho)}$$

Notice that  $q$  should be less than 0.5 as  $p$  is always larger than  $q$ . With a large  $n$  and we can have an approximate value

$$T_n \approx \frac{n}{p-q} = \frac{n(1 + \rho + \rho^2 + \rho^3)}{1 - \rho - \rho^2 - \rho^3}$$

In multi-plane fabric, input loading of each plane  $\rho' = \rho/k$ , we can have average delay of burst

$$T_n' \approx \frac{n(1 + \rho' + \rho'^2 + \rho'^3)}{1 - \rho' - \rho'^2 - \rho'^3} = \frac{n(k^3 + k^2\rho + k\rho^2 + \rho^3)}{k^3 - k^2\rho - k\rho^2 - \rho^3}$$

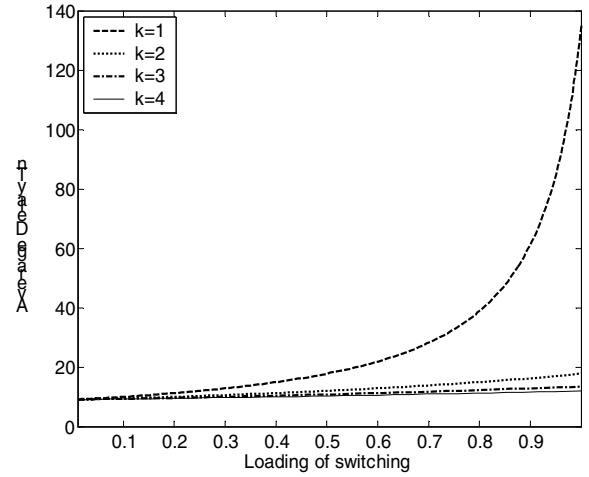


Figure 11. The average delay of Dual Shuffle-exchange Network with  $n = 9$

In the Figure 11, we can find  $k=2$  is good enough to reach a small average delay and  $k=3$  will introduce too much complexity with little gain.

### 5.2 Simulation results

We do the simulation based on 8 fibers, 64 wavelengths switch, and the input loading of whole switch is 0.72 and 0.36 with  $k=1,2,3$  or 4 planes DSN.

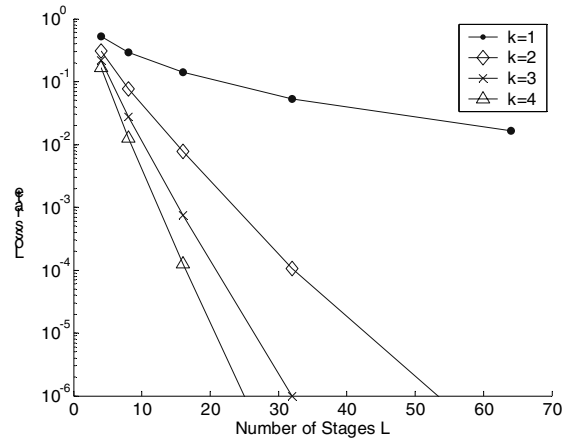


Figure 12. Loss rate of DSN versus number of stages with different  $k$  when loading is 0.72

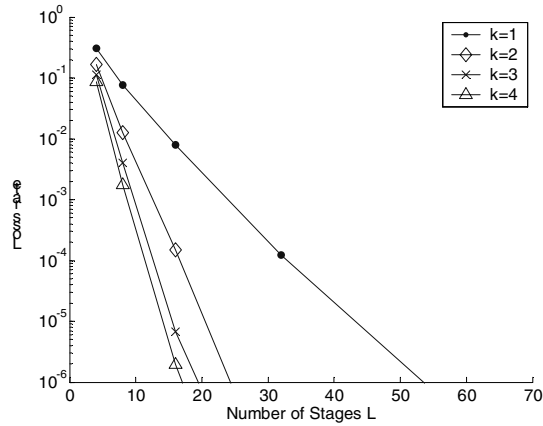


Figure 13. Loss rate of DSN versus number of stages with different k when loading is 0.36

From the simulation results of Figure 12 and Figure 13, we find that with a small  $n$ , by using  $k=2$  multi-plane fabric, we can make the switch with least complexity, that is, the switch with best scalability.

## 6. Conclusion

This paper has described a Dual Shuffle-exchange Network that can be employed in Optical Burst Switching. This switching architecture provides a significant reduction on complexity. We also propose a multi-plane architecture based on DSN. We find that with 2 DSN switches operated in parallel, we can further reduce the complexity of switch. Some implementation considerations are presented and analyzed. Simulation results on the performance are also given.

## 7. References

- [1] Jonathan S. Turner, "Terabit Burst Switching," *Journal of High Speed Networks*, 1999.
- [2] C.Qiao and M.Yoo, "Optical Burst Switching (OBS) – A New Paradigm for an Optical Internet," *J. High Speed Networks*, vol. 8, no. 1, 1999, pp.69-84.
- [3] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of Optical Burst Switching in IP-Over-WDM networks," *Journal on Selected Areas in Communications*, Oct 2000.
- [4] Detti, A.; Eramo, V.; Listanti, M., "Performance evaluation of a new technique for IP support in a WDM optical network: optical composite burst switching (OCBS)," *IEEE J. Lightwave Technology*, Feb 2002 , pp.154-165
- [5] Jin-Bong Chang; Chang-Soo Park, "Efficient channel-scheduling algorithm in optical burst switching architecture," *High Performance Switching and Routing, 2002. Merging Optical and IP Technologies. Workshop on* , 2002, pp. 194 -198
- [6] J. Ramamirtham and J. Turner. Design of wavelength converting switches for optical burst switching. In *proceedings of INFOCOMM*, volume 1, pages 362-370, 2002.
- [7] Liew, S.C.; Lee, T.T., "NlogN Dual Shuffle-Exchange Network with Error-correcting Routing," *IEEE Communications.*, vol. 42, pp. 754-766, Feb-Apr 1994.
- [8] J.Y. Wei and R.I. McFarland, "Just-in-time Signaling for WDM Optical Burst Switching Networks," *J.Lightwave Tech.*, vol.18, no. 12, Dec. 2000, pp. 2019-37.