Deflection routing protocol for burst switching WDM mesh networks

Xi Wang*, Hiroyuki Morikawa, and Tomonori Aoyama

School of Engineering, The University of Tokyo
7-3-1 Hongo Bunkyo-ku Tokyo 113-8656 JAPAN

ABSTRACT

Given the significant progress made and the continuing advances expected in the optical networking technology, it becomes attractive to build a future Optical Internet that natively supports bursty IP datagrams. Burst switching WDM optical networks are touted as suitable network architectures for future Optical Internet backbones. However, the lack of optical processing capabilities results in increased burst blocking probability, which in turn leads to very limited network performance. Efficient contention resolution method is therefore necessary. Based on discussions of the state of the art of recent optical technologies, a deflection routing protocol for burst switching WDM mesh networks is proposed. The idea of this approach is to use idle optical links as fiber delay lines for contention resolution. Simulation results show that the proposed protocol is available solution for effectively reducing the blocking probability and increasing the performance of burst switching WDM optical networks.

Keywords: Optical Internet, deflection routing, optical burst switching, WDM

1. INTRODUCTION

The Internet traffic is definitely different from the traditional voice traffic not only for its exponentially growing speed, but also its well-known burst characteristic.[1-3] The "Optical Internet" employing the concept of wavelength division multiplexing (WDM) and wavelength routing has been touted as the Internet backbone of the future.[4] In such a network, the network link layer connections are "dedicated" wavelengths on a WDM optical fiber directly connected to a high performance network router. For the purposes of effectively utilizing WDM multi-wavelengths and eliminating the bottleneck of electronic processing, large-scale efforts are underway to design a single, ubiquitous optical access layer with tight IP interworking to eliminate one or more of the intermediate layers (e.g., SONET/SDH, ATM). Meanwhile, as a longer-term solution, WDM subsystems consist of optical routing nodes and optical fiber links can be used to switch individual wavelengths optically and establish cut-through lightpaths (optical transmission paths between nonadjacent nodes in the optical networks domain),[5] thereby bypassing intermediate protocol processing and allowing for even closer IP-WDM layer integration. In order to realize such IP-over-WDM paradigms, optical switching function and its control protocols in optical routing nodes become indispensable between IP layer and underlying optical transport layer. Among various optical switching paradigms, Optical Burst Switching (OBS) [6-8] shows advantages in terms of switching efficiency for bursty IP traffic and optical hardware feasibility. We will use the term "burst optical networks" to refer to wavelength routing optical networks that apply OBS as their switching paradigm. Studies on OBS also show that burst optical networks are quite suitable network architectures for future Optical Internet backbones.[6-8]

In optical burst switching, the transmission links carry multiple WDM channels, which can be dynamically assigned to user data bursts. One channel on each link is reserved for control information. This separation of control and data simplifies the data path implementation, facilitating greater use of optical switching technologies. A *Setup* control packet is sent on the control channel to announce an upcoming burst. The control packet is then followed by a burst of data after a short delay. The *Setup* control packet is interpreted by the control portion of the switching systems, while the data channels are switched through transparently with no examination or interpretation of the data. Since data bursts are sent out without waiting for the acknowledgement replies from receivers to setup the path (one-way reservation paradigm), the burst could be blocked in an intermediate node due to resource contention, in which case, the burst has to be dropped. In order to reduce dropping probability, a contention resolution mechanism is required for an optical burst switch.

---

[*] Correspondence: Email: xi@mlab.t.u-tokyo.ac.jp; WWW: http://www.mlab.t.u-tokyo.ac.jp/~xi; Telephone: +81 3 5841 6710; Fax: +81 3 5841 6776

Optical buffering, wavelength conversion and deflection routing are technologies that can be used in an optical burst switch to resolve contention.[9] For example, fiber delay lines/circulating loop buffers can be used to temporarily store the bursts until the resources become available. However, as current optical buffers are typically limited to providing a few tens of μs delay,[10] it is not easy to store longer optical burst. Moreover, implementing optical buffers involves a great amount of hardware and complex electronic controls. Wavelength conversion is another key functionality that ensuring contention resolution and blockage removal in WDM networks by converting the data on one wavelength to another wavelength. The performance improvement on wavelength conversion gain has been extensively studied with analytical and simulation-based methods in the literature.[11-13] However, the problem of the wavelength conversion is the immaturity of optical converter technology.[14] Although all-optical conversion technologies without the aid of opto-electronic and electro-optic conversion are desirable and have been developed over a number of years, various issues such as the performance and cost implications still keep them far from practical use. Eliminating wavelength conversion can significantly simplify switching fabric and reduce the cost. Therefore, we assume that the switches are incapable of wavelength conversion in this work. The same wavelength should be assigned on all links along the route that an optical burst passes.[15-16] This restriction induces high blocking rate when the number of links a burst traverses becomes big. Compared to optical buffering and wavelength conversion, it is relatively easy to optically switching burst data.[17] Accordingly, another technique to control traffic flow and resolve contention in burst optical networks becomes attractive - "deflection routing". [18-20]

Based on these discussions of the state of the art of recent optical technologies, in this paper, we present a deflection routing protocol for WDM mesh optical networks. The idea of this approach is to use idle optical links as fiber delay lines for contention resolution. The paper is organized as follows. In the next section, we elaborate on the motivation for deflection routing in burst optical networks, and also describe the drawbacks of normal deflection routing algorithm. In Section 3, we present the deflection routing protocol in detail. The performance of proposed protocol is evaluated via computer simulation and the results are shown in Section 4. We conclude the paper in Section 5.

2. MOTIVATION FOR DEFLECTION ROUTING

2.1. Basic concept

A conceptual view of deflection routing is given in Fig. 1. Both senders ***A*** and ***B*** are sending bursts to receiver ***E*** (we denote their bursts as ***b(A, E)*** and ***b(B, E)*** hereafter). Before sending bursts, senders ***A*** and ***B*** send control packets (denoted as ***c(A, E)*** and ***c(B, E)***) on their out-of-band control channels for announcement. Since ***c(B, E)*** arrives at Node ***C*** earlier than ***c(A, E)***, the output link of Node ***C*** towards Node ***E*** is reserved for ***b(B, E)***. When ***c(A, E)*** arrives at node ***C***, the link between ***C*** and ***E*** is still in use by ***b(B, E)***. Node C then checks other output links and selects the idle link between C and D to deflect ***b(A, E)***. Node ***D*** forwards ***b(A, E)*** via link between ***D*** and ***E*** based on its routing table. Since every node performs deflection routing in this manner, the deflected burst arrives at its destination with some extra propagation delay, i.e., traverses several additional nodes than the shortest path. The idle optical links can be considered as fiber delay lines for "buffering" the blocked bursts. The bursts in the congested part of the network are then distributed to other underused parts, thus the overall link utilization and network performance can be improved.

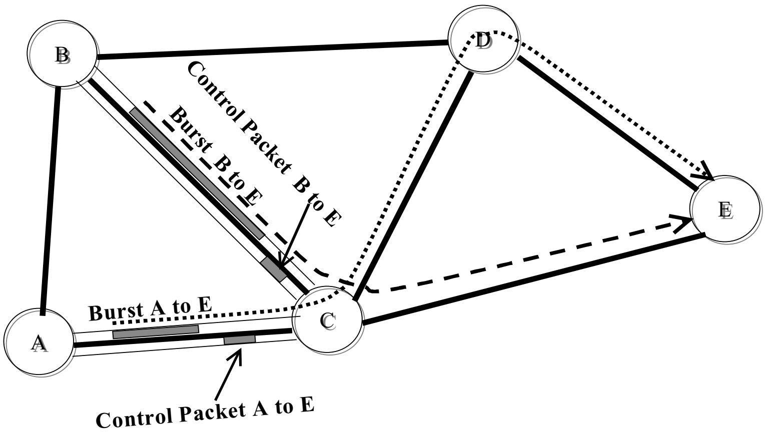


Figure 1. A conceptual view of deflection routing for optical bursts.

2.2. Benefits of deflection routing

The benefits of deflection routing for burst optical networks are discussed in the following. In traditional burst optical networks, if an intermediate node along the path a burst is traversing fails to reserve the resource (WDM channels), the burst has to be dropped and retransmitted again from the sender. In such a case,

a. A dropped burst wastes the bandwidth on the partially established path. If the burst data has been injected into the network, the network should do the best to forward it to the destination, more than simply drop it. For example, as shown in Fig. 2, when the receiver node is 6 hops away from the sender and the burst is dropped at the 5th hop, it has to be retransmitted and the total hop distance rises to (5 + 6 = 11) hops. If deflection is available, (5 + deflection hop count) is enough, which in most case is less than the total number of hops with the case of retransmission.

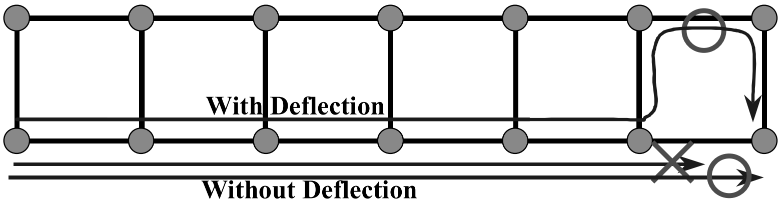


Figure 2. The effect of deflection routing in bandwidth gain.

b. The delay becomes very large when retransmitting a blocked burst in long-distance links. Different from the traditional concept that mainly concerns about the processing delay in every switching node, the transmission delay becomes dominant value in high-speed, broad-bandwidth optical networks due to RTT and big-fat-pipe effect. For example, the duration of the burst at 10 Gbps is only 0.8 ms for l MB bursts. On the other hand, the transmission delay over a 100 km optical fiber link would be 0.55 ms. If the destination is 4 hops away from the sender and a 1 MB burst is dropped at the 3rd hop, it at least takes (0.55 ms x 3 hops x 2 + 0.55 ms x 4 hops + 0.8 ms= 6.3) ms with one time of retransmission to make the burst totally reach its destination. However, if deflection is performed at the 3rd node, the total transmission time will be (0.55 ms x 3 hops + 0.55 ms x deflection hop count + 0.8 ms). If the deflection hop count is under 7 (in real case the number is actually much smaller), the total transmission delay will be reduced. Accordingly, deflection routing decreases the waste of bandwidth and the retransmission delay by eliminating the probability of burst dropping, yielding performance improvement.

2.3. Limitations of normal deflection routing

Deflection routing also has its disadvantages. For example, normal deflection routing is only efficient when the traffic load of the whole network is relatively low. Fig.3 shows the blocking probability improvement versus traffic load in normal deflection routing algorithm, where the dashed line indicates the case without deflection routing while other solid lines indicate various cases with deflection routing for different link *propagation delay / burst duration* ratio. We learn from the graph that when the traffic load is relatively low, deflection routing can greatly decrease the burst blocking probability. However, when the traffic load grows, the effect of deflection decreases and eventually induces even higher blocking probability than the case of no deflection. The rapid performance degradation of deflection routing is due to its "indiscriminate" deflection procedure. Deflection routing is based on the assumption that if the default output link is in use, most of other links are idle and available for deflection use. However, when the traffic increases, this assumption begins to break down because the number of idle links for deflection use decreases. Moreover, deflected traffic further lower the network capacity to process newly generated bursts. The blocking probability increases rapidly and the network throughput collapses completely when the load exceeds a certain threshold. In order to suppress this effect, we add two sender control functions in our burst optical deflection routing protocol. These functions reduce unnecessary deflection routing and limit the traffic increase when detecting the network congestion.

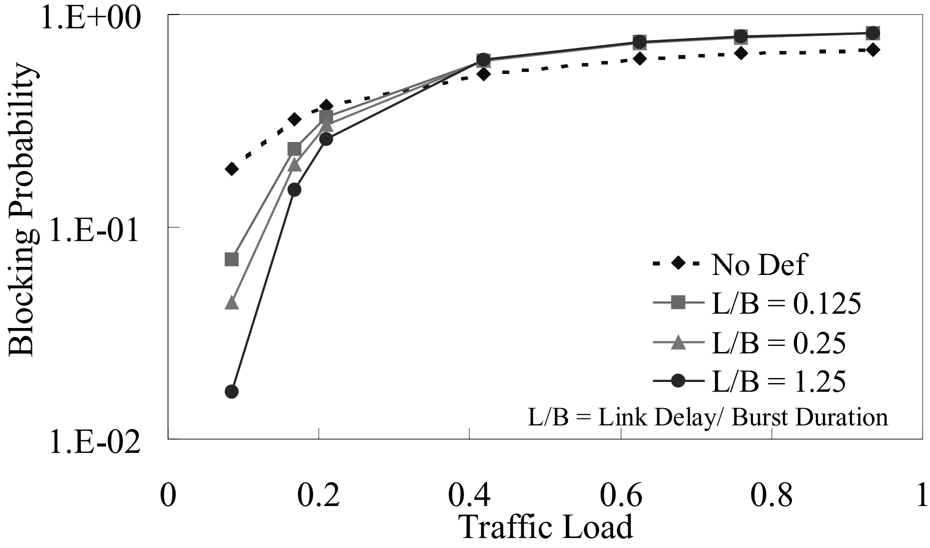


Figure 3. Blocking probability vs. traffic load in normal deflection routing algorithm.

3. DEFLECTION ROUTING PROTOCOL

3.1. Node functions

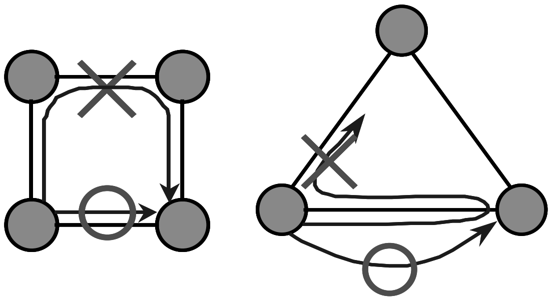
The proposed deflection routing protocol is realized by combining three basic functions together: Deflection function, sender check function and sender retransmission function.

3.1.1. Deflection function

Every intermediate optical burst switch uses idle links as Fiber Delay Lines (FDLs) for contention resolution. If the default output link is in use, it deflects the burst data to any of other idle links. Note that a burst may be deflected back to the sender again. For the purpose of utilizing links as optical buffers, short-term loops are not prohibited in our deflection routing method. We add a *TTL* field to burst control packets to prevent infinite loops.

3.1.2. Sender Check Function

A sender does not deflect its own bursts. Although an intermediate node has to deflect a burst immediately when its default link is occupied, a sender is not subject to this constraint. Since a sender keeps the data burst until it receives an *ACK* control packet for the corresponding data burst, it can always send the burst with the default output links. Compared to sender retransmission, deflection routing is efficient for saving delay and bandwidth consumption of bursts with long hop count. However, deflecting one-hop-bursts induces greater delay and bandwidth consumption than simply resend them from a sender. By restricting sender deflection, unnecessary deflection can be eliminated and the performance degradation of deflection routing is alleviated. A sender check function contains two phases: I. A sender does not send out a burst until its default link becomes idle (See Fig.4 a). II. If a burst is deflected back to its sender, the sender does not perform deflection forwarding again (See Fig.4 b). When forwarding burst data, each burst switch checks the sender field of every burst. If the switch finds a burst's sender is itself and the default output link for that burst is idle, it can forward the burst immediately. Otherwise, the switch waits until the default link becomes idle and then retransmits the burst.



a. b.

Figure 4. Sender check function (a. Sender does not perform deflection. b. Sender does not deflect its own bursts).

3.1.3. Sender Retransmission Function

When a sender is sending a burst and the default output link is in use, the sender will attempt to retransmit the burst after a certain interval time. The interval time T is calculated by

$$T = Burst\ Length \times n,$$
$$where \quad n = RAN[0, 2^k),$$
$$[k = \min(m, 10) \quad m : Attempt\ Count].$$

In normal deflection routing, a sender attempts to send out burst data immediately after its link becomes idle. There is no control mechanism for limiting the amount of traffic injected into the network. Accordingly, the monotonic traffic increase results in high blocking probability in normal deflection routing. By using sender retransmission algorithm above, the average retransmission interval increases exponentially, effectively suppressing traffic explosion and reducing the blocking probability.

3.2. Routing Procedures

In a burst optical network, each node handles both bypassing and locally generated/terminated bursts (i.e., either a core or a edge node of the burst optical network domain). All data bursts are fully controlled by the control packets. We define three kinds of control packets: *SETUP*, *ACK*, and *NAK*. The information fields are listed in Table 1.

Table 1. The information fields of the control packet.

Information	**Description**
Packet Identifier	The kind of control packet (*SETUP*, *ACK* and *NAK*)
Receiver Address	Burst receiver address
Sender Address	Burst sender address
Burst Identifier	Sequence number of the burst
Offset Time	Offset time between control packet and burst
Absolute Time	The departure time of the burst at each node
Burst Duration	The length of burst in time domain
Deflection Flag	Set when the burst is deflected
Time To Live	Prevent endless loop
Failure Identifier	"All link busy" or "TTL=0"

The routing procedure of an intermediate node in the burst optical network is shown in Fig. 5. On receiving a *SETUP* control packet, a node performs following process:

- Check the destination node address, then look up the routing table to decide the proper output link.
- Check whether the wavelength of that link has been reserved by other bursts. If not, reserve bandwidth for the upcoming data burst and forward *SETUP* control packet to the next node.
- If the wavelength has been reserved by other bursts, after executing sender check and sender retransmission procedure, select any available link and deflect the setup control packet to the link.
- If the assigned wavelength on every output link is not available, a *NAK* packet is sent back to its upstream node to indicate the burst dropping.
- If at least one output link is available for the upcoming data burst, setup the switch to build cut-through path and transmit it to the output link.
- If a burst has reached its destination, an *ACK* packet is sent back to the sender to inform the successful delivery.

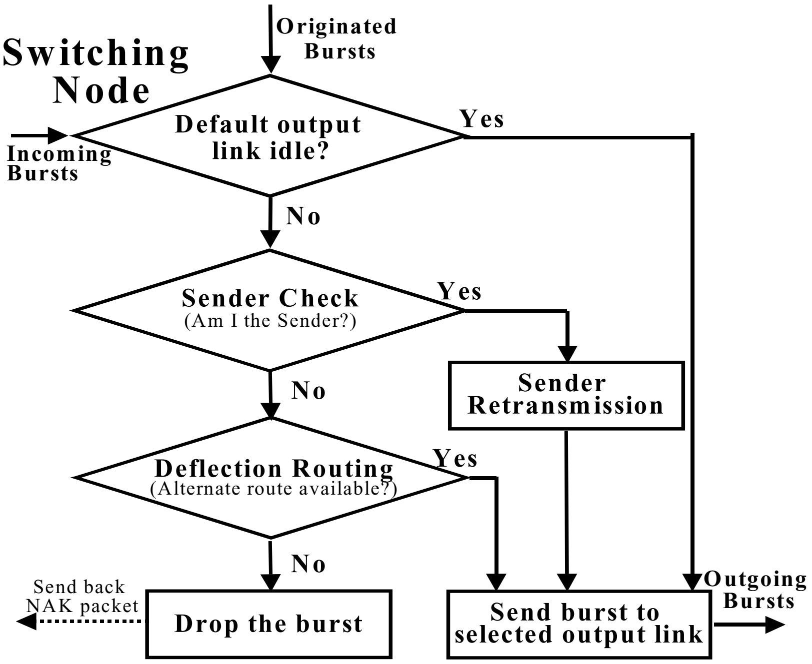


Figure 5. Routing procedures in the optical switch.

4. SIMULATIONS

4.1 Assumptions & Network Models

The assumptions and network models used in the simulation are listed below:

- The network consists of nodes and links interconnected in a mesh interconnection pattern.
- Each single fiber link has the same number of wavelengths.
- Each node handles both bypassing and locally generated/terminated bursts.
- There is no wavelength conversion in the burst switch.
- There is no buffer in the burst switch.
- Optical bursts are generated at the ingress node of a burst optical network domain.
- Bursts are randomly generated for arbitrary sender-receiver pairs.
- Burst duration time is exponentially distributed.
- Wavelength is randomly assigned at a sender.
- Dijkstra routing protocol is applied. routes are randomly chosen in case of the same distance (hop count).

Note that although we assume there is no optical buffer and wavelength conversion in switching nodes, our protocol does not necessarily exclude the use of these capabilities. There are several possible combinations of optical buffering and wavelength conversion with deflection routing. Extending the proposed protocol to support these options are our future work.

The simulation parameters are shown in Table 2.

Table 2. Simulation parameters.

Parameter	Value
Network size	4 x 4 – 16 x 16 mesh
Wavelengths	4 - 16
Link distance	20 - 5000 kilometers
Link bandwidth	10 Gbps
Average burst length	1 MB
Traffic load	0.04-0.9
Deflection TTL	12-60

4.2 Simulation Results

We began by addressing the effect of blocking probability reduction realized by our proposed deflection routing method. Fig. 6 shows the result of a set of simulations in which a part of the deflection routing functions is turned on and off. In this case, the network size is 4 x 4 and 16 wavelengths are available in each link. From these plots, we can see the effect of deflection routing. The bold line shows the result when no deflection routing is applied. In the case of normal deflection routing (shown as "Def Only"), the performance is improved when the traffic load is under 0.4, however, in the case of higher traffic load (> 0.4), the blocking probability becomes higher than the case of no deflection.

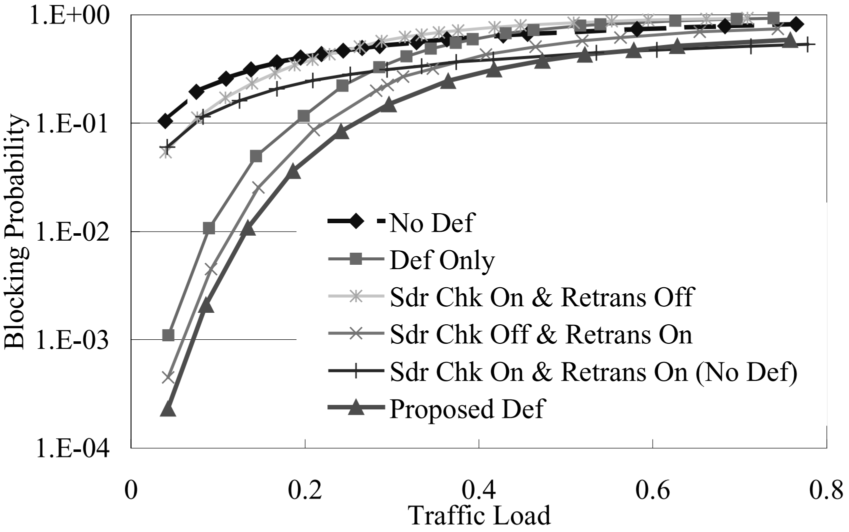


Figure 6. The effect of proposed method in terms of blocking probability reduction.

By adopting the sender check function and the sender retransmission function separately, the degree of deflection's performance degradation is mitigated. By combining sender check and sender retransmission function, the blocking probability is greatly reduced. The result of proposed deflection routing method is shown as "Proposed Def". The blocking probability is greatly reduced when the traffic load is relatively low. Notice that the blocking probability of proposed method is always lower than that of no deflection for overall traffic load. In addition, it is also always lower than any other cases.

Next, the relationship between the average burst length and the average link propagation delay for various traffic loads was studied. In Fig. 7, three lines show the results of proposed deflection routing method for traffic load 0.084, 0.167 and 0.21 respectively. For reference, the blocking probability without deflection routing is shown as single dots (because the blocking probability is the same when the *propagation delay / burst duration* ratio is changed).

From the simulation result, we see that if the link *propagation delay / burst duration* ratio is smaller than 1, the effect of deflection begins to degrade. However, when the ratio is larger than 1, there is not significant performance improvement attained. Accordingly, the ratio around 1 is thought to be sufficient. For example, in order to achieve better deflection routing effect in a burst optical network with 1 ms average propagation delay (200 kilometers links), the average burst duration time should be set to lower than 1 ms (with burst length of 1.25 MB in 10 Gbps links).

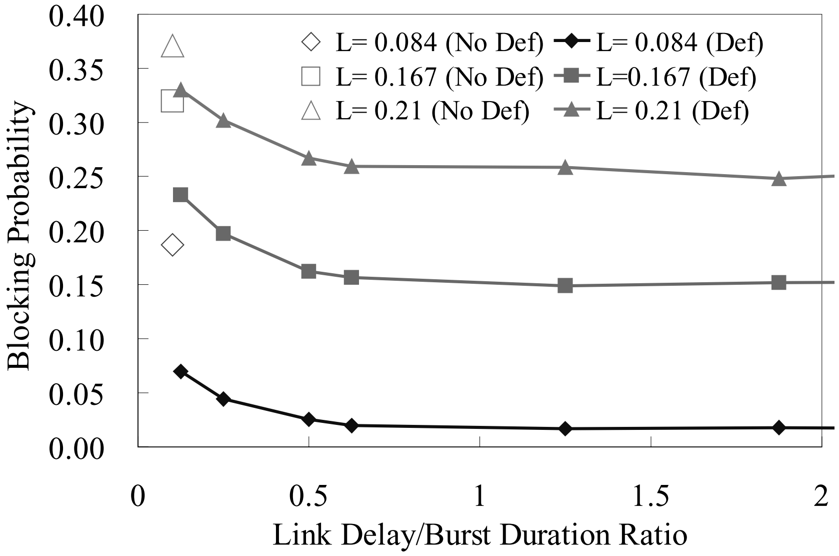


Figure 7. The relationship between average burst length and average link propagation delay.

We also studied the fairness issue of bursts with different hop count. Fig. 8 shows the blocking probabilities of bursts with different hop count. Comparison was made between the case of with and without deflection under various traffic loads. Note that the blocking probability of bursts of 1 hop is excluded because one-hop-burst blocking probability is always 0 in our protocol. For light loaded cases, (i.e., load < 0.1), notable blocking probability reduction is observed for bursts with all kinds of hops count. In the case of higher load (load=0.41), the effect of deflection routing decreases. The blocking probabilities of bursts with various hops count are almost the same as the case of no deflection. Accordingly, we see from the figure that the blocking probability reduction for bursts with different hop count is almost the same for both cases at various traffic loads. These results indicate the proposed deflection routing protocol can evenly (fairly) decrease the blocking probability of the bursts with different hop count under various traffic loads.

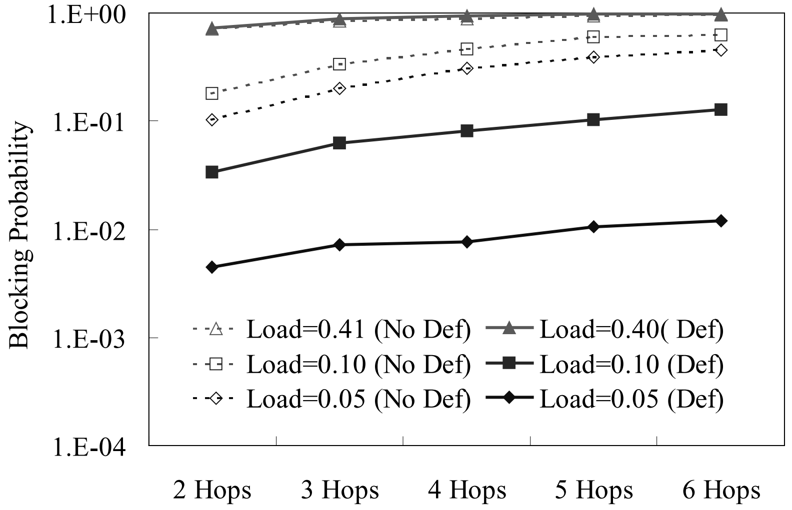


Figure 8. Blocking probability of bursts with different hop count.

Finally, we examine the applicability of proposed deflection method to larger network scales. Fig. 9 and Fig. 10 show the results of the deflection routing performance in variouse network sizes. Grid networks form 4 × 4 (16 nodes) to 16 ×16 (256 nodes) are tested. In each figure, left vertical axis indicates the percentage of blocking probability in logarithm form; right vertical axis indicats the ratio between the blocking probability of non-defletion and deflection-applied cases in logarithm form. Note in ether case, the sender check and sender retransmission function are applied. In Fig. 9, since there is only slight traffic load (0.04), deflection routing greatly decreases the possiblity of blocking, especially for small networks, the improvement ratio is as high as several hundred order. With the increase of network size, the blocking probability in both cases increases and the value of improvement ratio decrease. However, this ratio still remains higher than 10 even in 16 × 16 networks. On the other hand, as shown in Fig. 10, when the traffic load is increased to 0.2, the benefit of deflection routing decreases.When the network size grows, the performance improvement ratio becomes near to 1, which indicates that there is little benefit of deflection routing in large networks. We learn from the result that the decision of proper network size is also important on designing a burst optical network for the purpose of acquiring the best performance.

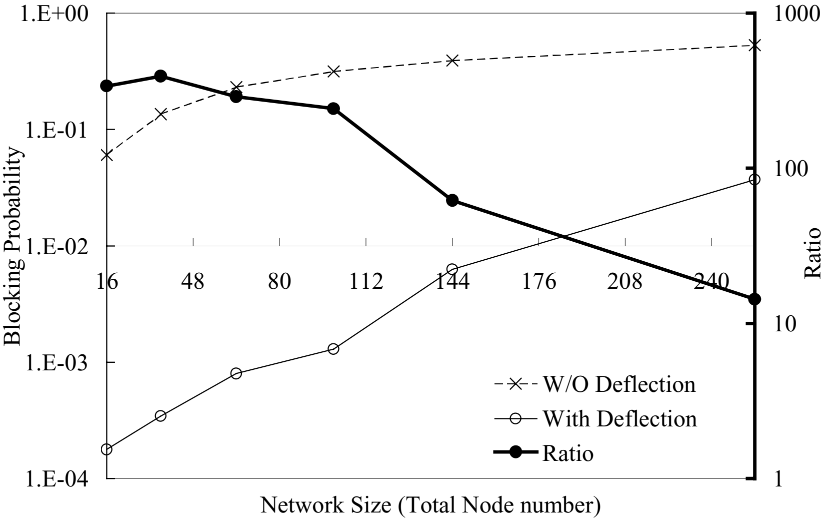


Figure 9. Deflection Effect vs. Network Size (Traffic Load = 0.04).

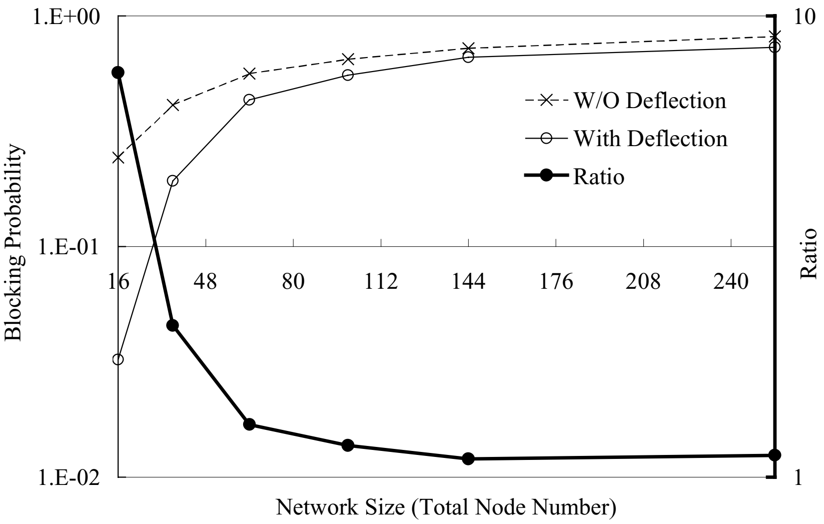


Figure 10. Deflection Effect vs. Network Size (Traffic Load = 0.2).

5. CONCLUSION

In this paper, a deflection routing protocol for contention resolution in burst switching WDM mesh networks is proposed. By combining sender check and sender retransmission functions with the deflection routing, simulation results show that proposed protocol can greatly decrease the blocking probability especially for low traffic load networks. At the same time, the blocking probability is always kept lower for overall traffic load than the case without deflection routing. These results indicate the proposed deflection routing protocol can be regarded as a viable solution for contention resolution in burst optical networks. With the concept and method of optical-burst deflection routing presented in this paper, it is feasible to greatly simplify optical burst switch fabric and build reasonable burst optical networks for future Optical Internet - which is the goal we are pursuing.

6. REFERENCES

1. V. Paxson and S. Floyd, “Wide Area Traffic: The Failure of Poisson Modeling,” *IEEE/ACM Trans. Networking*, **3**, pp. 226-244, 1995.
2. V. Paxson, “Towards a Framework for Defining Internet Performance Metrics,” in Proc. *INET '96*, Montreal, 1996.
3. R. Jain, S. Routhier, “Packet Trains: Measurements and a New Model for Computer Network Traffic,” *IEEE J. Select. Areas Commun.*, **4**, pp. 1162-1167, 1986.
4. T. W. Chung et al., “Architectural and Engineering Issues for Building an Optical Internet,” *http://www.canet2.net/c3/architectural.pdf*
5. Chlamtac, A. Ganz, and G. karmi, “Lightpath Communications: An Approach to High Bandwidth Optical WAN’s,” *IEEE Trans. Commun.*, **40**, pp. 1171-1182, 1992.
6. C. Qiao and M. Yoo, “Optical Burst Switching (OBS): A New Paradigm for an Optical Internet,” *J. High Speed Networks*, **8**, pp. 69-84, 1999.
7. Chaskar et al., “A Framework to Support IP over WDM Using Optical Burst Switching,” *Wksp. Opt. Networking*, Dallas, 2000.
8. J. S. Turner, “Terabit Burst Switching,” *http://www.arl.wustl.edu/~jst/ projects/TerabitBurst.html*
9. R. Ramaswami and K. N. Sivarajan, *Optical Networks A Practical Perspective*, Morgan Kaufmann, San Francisco, 1998.
10. F. Masetti, et al., “Fiber Delay Lines Optical Buffer for ATM Photonic Switching Applications,” in *Proc. INFOCOM’93*, pp. 935-942, San Francisco, 1993.
11. E. Karasan and E. Ayanoglu, “Effects of Wavelength Routing and Selection Algorithms on Wavelength Conversion Gain in WDM Optical Networks,” *IEEE/ACM Trans. Networking*, **6**, pp. 186-196, 1998.
12. Kovacevic and A. Acampora, “On Wavelength Translation in All-optical Networks,” in Proc. *INFOCOM’95*, pp. 413-422, Boston, 1995.

13. S. Subramaniam and R. Barry, “Wavelength Assignment in Fixed Routing WDM Networks,” in Proc. *ICC’97*, pp. 406-415, Montreal, 1997.
14. J. M. H. Elmirghani and H. T. Mouftah, “All-Optical Wavelength Conversion Technologies and Applications in DWDM Networks,” *IEEE Commun. Mag.*, pp. 86-92, Mar. 2000.
15. R. Ramaswami, K. N. Sivarajan, “Routing and Wavelength Assignment in All-Optical Networks,” *IEEE/ACM Trans. Networking*, **3**, pp. 489-500, 1995.
16. Birman, A. Kershenbaum, “Routing and Wavelength Assignment Methods in Single-Hop All-Optical Networks with Blocking,” in Proc. *INFOCOM’95*, pp. 431-438, 1995.
17. T. Kato, et al., “10Gb/s Photonic Cell Switching with Hybrid 4x4 Optical Matrix Switch Module on Silrca based Planar Waveguide Platform,” in Proc. *OFC'98*, post deadline paper, San Jose, 1998.
18. D J. Blumenthal, et al., “Photonic Packet Switches: Architecture and Experimental Implementations,” *Proc of the IEEE*, **82**, pp. 1650-1667, 1994.
19. Jae-Hyun Park, et al., “The Deflection Self-Routing Banyan Network: A Large-Scale ATM Switch Using the Adaptive Self-Routing and its Performance Analyses,” *IEEE/ACM Trans. Networking*, **7**, pp. 588-604, 1999.
20. E. A. Varvarigos, et al., “A Virtual Circuit Deflection Protocol,” *IEEE/ACM Trans. Networking*, **7**, pp. 335-349, 1999.