A Novel Scheme to Reduce Burst-Loss and Provide QoS in Optical Burst Switching Networks

Ashok K. Turuk and Rajeev Kumar

Department of Computer Science and Engineering,
Indian Institute of Technology Kharagpur,
Kharagpur, WB 721 302, India
{akturuk, rkumar}@cse.iitkgp.ernet.in

present Email akturuk@nitrkl.ac.in

Abstract. Burst loss due to contention is a major issue in optical burst switching networks. In this paper, we propose a contention resolution scheme that uses a offset time different from that of conventional optical burst switching (OBS) to reduce burst loss and to provide QoS in optical burst switching networks. The proposed scheme can be tuned to both prioritized traffic and delay constraint traffic by changing the offset time. For selecting a data-channel, we propose three channel selection algorithms, namely Least Recently Used (LRU), First Fit (FF), and Priority Set (PS) algorithms. We simulate and compare proposed scheme with the preemptive priority just-enough-time (PPJET) contention resolution scheme. We consider bursty traffic in our simulation. It is found that our scheme outperforms PPJET in burst-loss.

1 Introduction

There has been a phenomenal increase in the number of Internet users and the variety of Internet applications in recent years. This has resulted in exponential growth of Internet traffic, demanding a huge bandwidth at the backbone network. To meet this growing demand for bandwidth, wavelength division multiplexing (WDM) network has become the de-facto choice for the backbone network. IP over WDM networks have drawn much attention among researchers, and many integration schemes between IP and WDM layers have been proposed [1].

To carry IP traffic over WDM networks three switching technologies have been studied: optical circuit switching, packet switching and burst switching. Optical circuit switching and packet switching have their own limitations when applied to WDM networks. Circuit switching is not bandwidth efficient unless the duration of transmission is greater than the circuit establishment period. It is shown that establishment of circuits (lightpaths) in optical networks is an NP-hard problem [2]. Many heuristics and approximation algorithms exist for establishing lightpaths in optical networks e.g., see [3] and the references therein. Packet switching is hop-by-hop store and forward scheme and, needs buffering and processing at each intermediate node. It is flexible and bandwidth efficient. However, technology for buffering and processing in optical domain is yet to mature for this scheme to commercialize. Fiber delay lines proposed in literature provide limited buffer capability and are suitable for delays of fixed duration only.

L. Bougé and V.K. Prasanna (Eds.): HiPC 2004, LNCS 3296, pp. 309–318, 2004.

In this context optical burst switching (OBS) is emerging as the new switching paradigm for the next generation optical networks [4, 5]. It combines features of both circuit and packet switching. As such there exists no formal definition of OBS, the features defined by Yoo and Qiao [4] for OBS have become de-facto standards. The burst-size granularity (which lies between circuit and packet switching), separation of control and data bursts, one-way (for most cases) or two-way reservation scheme, and no optical buffering are important characteristics of the OBS paradigm.

Some major issues in optical burst switching networks are: (i) contention resolution, (ii) burst assembly, and (iii) quality-of-service (QoS) support. In a buffer-less OBS network contending burst is lost. Therefore, burst-loss should be minimized in OBS networks, is the key design parameter. A few approaches to contention resolution used in OBS are: buffering [6], deflection routing [7], burst segmentation [8, 9] and window based technique [10]. Burst assembly is the process of aggregating and assembling IP packets into bursts[11]. With increase in variety of Internet applications, different applications such as voice-over-IP (VoIP), video-on-demand, video conferencing etc. demand different QoS requirements. To meet the QoS requirements of different applications, IETF proposed IntServ and DiffServ schemes. However, such conventional priority schemes are defined for electronic domain which trivially mandates the use of buffers at intermediate nodes. Such schemes cannot be used directly to support QoS in buffer-less OBS networks. Thus, any scheme to support *differential* QoS requirements in OBS networks should not mandate the use of buffers at intermediate nodes.

Many schemes, in recent years, have been suggested to support priority based QoS in OBS networks. All the proposed schemes have tried to reduce the burst loss. It is not the burst-loss only but also the number of packet-loss that matters. For example, consider three bursts b_1, b_2 and b_3 of size 10, 20 and 50 number of packets each. A loss of any of the bursts indicates 33% of the burst loss. However, if we consider loss in terms of packets, the packet-loss comes out to be 14%, 28% and 64% respectively. Therefore, it is desirable that a contention resolution scheme should take care of the losses calculated in terms of packets. Consider Fig. 1(a) and assume both requests r_1 and r_2 have the same priority and arrive at a node at the same time. In OBS, the burst b_2 is always dropped. However, if the burst-size is taken into account the larger burst b_2 could have succeeded and the smaller burst b_1 is dropped. This will result into larger number of packets transmitted and higher resource utilization.

Next, consider Fig. 1(b) having two bursts of the same priority. In OBS, both the bursts are dropped. However, taking the burst-size or the number of hops traversed into contention resolution scheme one of the bursts succeeds. This gives rise to lower burst loss and larger number of packets transmitted. Thus, if we consider two more parameters – burst-size and number of hops traversed – in resolving contention, this will guarantee that at least one of the bursts succeeds and larger number of bits transmitted.

End-to-end delay is another key parameter for QoS provisioning. All delay sensitive applications demand that end-to-end delay is bounded by the delay constraints imposed by the respective application. Contention should be resolved by considering the delay factor too.

In this paper, we present a flexible algorithm for contention resolution to support a larger set of QoS parameters in OBS networks. We consider packet loss and number of

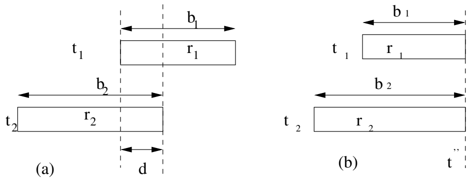

Fig. 1. Illustrations for Contention Resolution : (a) two requests are partially overlapped for a period d, (b) two requests have the same reservation instance

hops traversed, in addition to priority, for resolving contention. Our scheme is generic and can easily be adapted to satisfy delay-constraints. The aim is to reduce blocking probability of the bursts arising due to resource contention at intermediate nodes as well as to meet the delay constraints of the delay sensitive traffic. The proposed scheme guarantees that at least one of the bursts succeeds when contention occurs; the contention should be resolved in accordance with satisfaction of QoS parameters. To select data-channel, we propose three channel selection algorithms – (i) *Least Recently Used* (LRU), (ii) *First Fit* (FF), and (iii) *Priority Set* (PS). Channel selection algorithms run at the ingress routers to select a data-channel for reservation and for subsequent transmission. We evaluate the proposed scheme with the above channel selection algorithms.

The rest of the paper is organized as follows. Section 2 explains the contention resolution technique; few assumptions and notations used are described in Sub-section 2.1. Channel selection algorithms are explained in Section 3. Simulation results are presented in Section 4 and compared with PPJET. Finally, some conclusions are drawn in Section 5.

2 Proposed Contention Resolution Scheme

2.1 Assumptions and Notations

We model an optical network by means of a undirected graph $G(V, E)$ where V is the set of vertices (nodes) and E represents the set of links/edges in the network. Two types of nodes (here after, we use the terms node and router interchangeably) are identified: edge routers and core routers. Every edge router has $(n_e - 1) \times P$ electronic buffers where n_e is the number of edge routers, P is the number of priority classes supported in the system. Each buffer belongs to a specific pair of priority class and an egress router. The core router has no buffer; this is a desirable feature of the optical burst switching networks. A core router acts as a transit router for data-traffic. Thus, the data-traffic remains in optical domain from ingress to egress router. Propagation delay, t_p, between every pair of adjacent vertices in graph G is assumed to be the same. Processing delay of the control packet at each router is assumed to be δ. We use the following notations in rest of the paper:

$H_t^{sd}(r)$: Number of hops for the request r between source - destination pair (s, d),
$H_i^{sd}(r)$: Remaining number of hops for the request r between source - destination pair (s, d) at node i.

$Original\ burst$: A burst for which resources are already reserved at the core router, $Contending\ burst$: A burst whose reservation request has resulted in resource contention at the core router.

We define the following three situations that can occur when an intermediate router receives a reservation request:

- *No contention* (NC): When no contention for resources occurs at the intermediate core router.
- *Contention resolved* (CR): When a contention occurs at an intermediate core router i and for at least one of the requests $H_i^{sd}(r) > H_t^{sd}(r)/2$.
- *Contention-not-resolved* (CNR): When contention occurs at intermediate core router(s) and for none of the request $H_i^{sd}(r) > H_t^{sd}(r)/2$.

2.2 Proposed Scheme

OBS is based on either *one way* or *two way* reservation protocol. The minimum latency in one way reservation protocol is $P + \delta \cdot H$ where the minimum latency in two way reservation protocol is $2P+\delta \cdot H$. Our proposed scheme is a one way reservation protocol however it differs from other OBS schemes in two aspects - one, the offset time, and second, the methods adopted for contention resolution. In other OBS schemes, the offset time is $\delta \cdot H$ where δ is the processing delay of control packet at each node and, H is the number of hops between source-destination pair. In our scheme, we take the offset time to be $P + \delta \cdot H$ where P is the propagation delay between source-destination pair. The need for the additional P units of time is explained subsequently. The minimum latency of burst in other OBS schemes, is $P + \delta \cdot H$ which is same if a burst is sent along with control packet in optical packet switching. The minimum latency in optical circuit switching is $3P + \delta \cdot H$. In the proposed scheme, the minimum latency of a burst is $2P + \delta \cdot H$. Thus, we can say the minimum latency in our scheme is identical to OBS with $two\ way$ reservation protocol. However, the proposed scheme is a $one\ way$ reservation protocol where each burst experiences an additional delay of P units. The scheme is also tunable to delay sensitive traffic. For delay sensitive traffic the offset time in the proposed scheme is taken to be $\delta \cdot H$ which is the same as that in OBS. However, this offset can be made adaptive to the application needs. In the scheme, if a contention occurs and the situation is a *CR* one as mentioned in Section 2.1 then a burst is further delayed for the contention period. However, this delaying technique of our scheme is not applicable in case of delay sensitive traffic. For delay sensitive traffic if the required resources are not available within that amount of time, the burst is dropped.

Secondly, the proposed scheme differs from OBS in the method adopted for contention resolution. In OBS, the resource conflict is resolved on the basis of request priority and the time instance for which request is made. In addition to the above two parameters, we take burst-size and the number of hops traversed to resolve contention. A higher priority request is given a priority. However, for the same priority requests, the one that has traversed the maximum number of hops, is accepted for better resource utilization. The request that has traversed the maximum number of hops have more resources reserved on the path. Accepting this request will give rise to higher resource utilization. For same priority and the equal number of hops traversed the burst that has larger burst-size is accepted. For same priority, equal number of hops traversed and the same burst-size their instance of reservation is taken for resolving the conflict. Thus, the tie in contention resolution is resolved in order of priority, number of hops traversed, burst-size and delay.

Next, we explain the basis of having P additional units of delay in offset time with the help of timing diagrams illustrated in Figures 2 and 3. The total delay encountered by a control packet for source-destination pair (s, d) is no greater than $\varDelta = \delta \times H_t^{sd}(r)$. The offset-time, T, in OBS is taken to be at least $\varDelta$. In Fig. 2 the number of hops between source - destination pair (s, d) is 4. Therefore, the offset-time T in OBS is 4δ. In OBS, if a contention occurs say at node A or at node C then the burst is dropped at A or at B as shown in Figures 2(b) and 2(c) respectively. With this offset time a contending burst cannot be further delayed.

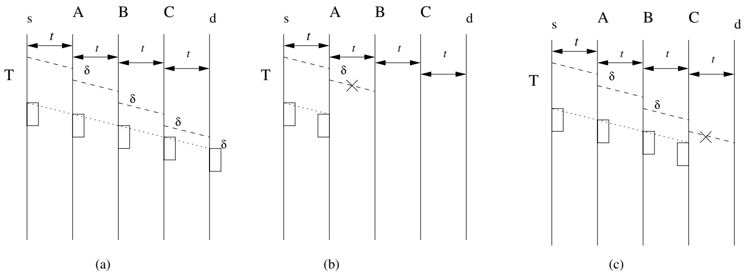

Fig. 2. Timing Diagram of burst switching network: (a) no contention occurs at intermediate nodes, (b) contention occurs at node A and (c) contention occurs at node C

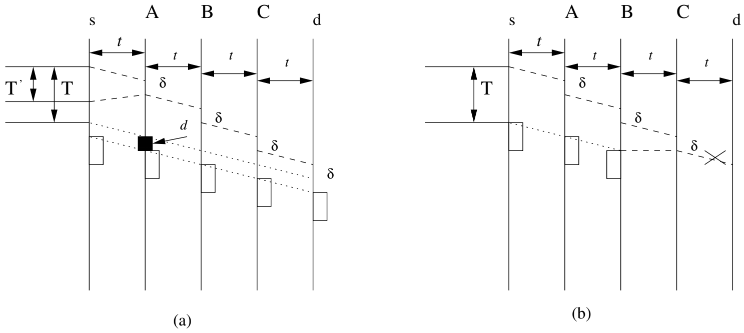

Fig. 3. Timing Diagram for the proposed scheme: (a) contention at node A is resolved, and (b) contention at node C but the burst is dropped at node B

In our proposed scheme, the offset, T, between source-destination pair (s, d) is taken to be $(t + \delta)H_t^{sd}(r)$. For the above example the offset time between source-destination pair (s, d) is $4(t + \delta)$. Let us consider Fig. 2(b) where contention has occurred at node A and, d be the duration of the contention period. The control packet has taken *one* hop to reach the node A from the source s. If a message is sent from node A to the source s to delay the transmission of burst for the contention period d it will reach s at $T^{'} = 2(t + \delta)$ after the source s has sent the control packet (Fig. 3(a)). The offset-time $T > T^{'}$ i.e., source s will receive the message to delay the transmission before the expiry of offset-time. Hence, the transmission of the burst is delayed and is not dropped at node A as shown in Fig. 3(a).

Let us consider Fig. 2(c) where contention has occurred at node C and, d be the duration of the contention period. If a message is sent from node C to the source s to delay the transmission of burst for the contention period d it will reach s at $T^{'} = 6(t + \delta)$. The offset-time $T < T^{'}$, i.e., source s will receive the delay message after it has transmitted the burst and the burst is dropped at node C. Therefore, instead of sending a delay message if a resource-release message is sent from node C, the message will release the resources reserved at node B before the burst arrives at node B and is dropped at node B rather than at node C. This gives rise to better utilization of the resources on link BC which was earlier occupied by the request.

Thus, in the proposed scheme, a contention occurs at node i and $H_i^{sd}(r) > H_t^{sd}(r)/2$ (this is the CR situation as described in Section 2.1), a message is sent to delay the transmission of the burst for duration of the contention period. For $H_i^{sd}(r) <= H_t^{sd}(r)/2$ (this is the CNR situation as described in Section 2.1), a message is sent to release the reserved resource.

We illustrate below possible cases of contention and the way contention is resolved in the scheme. For all the cases we refer to Fig. 1(a). In Fig. 1(a), the value of t_1 and t_2 indicates the time of arrival of requests r_1 and r_2 respectively, at a core router i. Burst-size of the requests r_1 and r_2 is indicated by b_1 and b_2 respectively. Below, we give interpretation for Case 1; the rest of the cases are interpreted in the same way. In Case 1 the contention has occurred due to the arrival of requests r_1 and r_2 at the core router i at the same time ($t_1 = t_2$). The remaining number of hops to be traversed for request r_1 is $H_i^{sd}(r_1) > H_t^{sd}(r_1)/2$ and, for r_2 is $H_i^{mn}(r_2) > H_t^{mn}(r_2)/2$. We have assumed that the contention has occurred at the core router i.

Case 1: $t_1 = t_2, H_i^{sd}(r_1) > H_t^{sd}(r_1)/2, H_i^{mn}(r_2) > H_t^{mn}(r_2)/2.$
Accept the high priority request and send a message to the ingress router of low priority request to delay the transmission for the contention period d. For same priority of both the requests, accept the request that has traversed the maximum number of hops and send a message to the ingress router of the other request to delay the transmission for the contention period d. For same priority and the equal number of hops traversed, accept the request that has larger burst-size and send a message to the ingress router of other request to delay the transmission for the contention period d.

Case 2: $t_1 = t_2, H_i^{sd}(r_1) \leq H_t^{sd}(r_1)/2, H_i^{mn}(r_2) > H_t^{mn}(r_2)/2.$
Accept the high priority request. If the low priority request is r_1 then it is dropped else a message is sent to the ingress router of r_2 to delay the transmission for the contention period d. For same priority of both the requests, accept the one that has traversed the maximum number of hops. Other request is processed as explained. For same priority

and equal number of hops traversed, accept the one with higher burst-size. Other request is processed as explained earlier.

Case 3: $t_1 = t_2$, $H_i^{sd}(r_1) > H_t^{sd}(r_1)/2$, $H_i^{mn}(r_2) \leq H_t^{mn}(r_2)/2$.
Requests are processed as explained in Case 2. Here, the request that is to be dropped is r_2.

Case 4: $t_1 = t_2$, $H_i^{sd}(r_1) \leq H_t^{sd}(r_1)/2$, $H_i^{mn}(r_2) \leq H_t^{mn}(r_2)/2$.
Requests are processed as in Case 1. Here the request that is not accepted is dropped.

Case 5: $t_1 < t_2$, $H_i^{mn}(r_2) > H_t^{mn}(r_2)/2$.
In this case the request r_1 has arrived before r_2 and resources are already reserved for the request r_1. For request r_2 a message is sent to the ingress router to further delay the transmission of burst for the contention period d.

Case 6: $t_1 < t_2$, $H_i^{mn}(r_2) \leq H_t^{mn}(r_2)/2$.
As in Case 5 resources are already reserved for the request r_1. Request r_2 is dropped.

Case 7: $t_1 > t_2$, $H_i^{mn}(r_2) > H_t^{mn}(r_2)/2$.
In this case request r_1 has arrived at a later point of time than r_2 and is contending with request r_2. Requests are processed similar to Case 5.

Case 8: $t_1 > t_2$, $H_i^{sd}(r_1) \leq H_t^{sd}(r_1)/2$.
As in Case 7, request r_1 has arrived at a later point of time than r_2 and is contending with request r_2. Requests are processed similar to Case 6.

In the above cases, cases 4, 6 and 8 are *CNR* situations and the rest are *CR* situations as defined in Section 2.1.

3 Channel Selection Algorithms

In this section, we describe three channel selection algorithms called (i) *Least Recently Used* (LRU), (ii) *First Fit* (FF), and (iii) *Priority Set* (PS) algorithms used in channel selection for our proposed contention resolution scheme (Section 2.2). The channel selection algorithms are run only at the edge routers to find the data-channels for which reservation request is to be made and subsequently transmit the data-burst. In LRU, a data-channel which is idle for the maximum duration is selected. In FF, data-channels are searched from the lowest index and the one that is available first, is selected. For example, consider Fig. 4, LRU channel selection algorithm selects the data-channel 2 as it is idle for the maximum duration where as FF channel selection algorithm selects the data-channel 0.

In PS approach, we decompose the set of data-channels, S, into P sub-sets, S_i, of data-channel. P is the number of priority classes supported. $S = S_0 \cup S_1 \cup \cdots \cup S_{P-1}$. A priority class i selects the data-channel from the set S_i. If no data-channel is available in the set S_i then it selects from the set S_{i-1} and if not available then from the set S_{i-2}. This process is iterated till the lowest priority set S_0 is searched. If no data-channel is available in the set S_0 then the burst is dropped at the ingress router. The number of data channels in the set S_i is in proportion to the traffic of priority class i.

For the priority class 0, if no data-channel is available in the set S_0 then the burst is dropped at the ingress router. To illustrate the working of Priority Set approach, we consider two priority classes 0 and 1; class 1 has higher priority than class 0. We divide the available data-channel as shown in Fig. 5 in two sets $S_0 = \{0, 1\}$ and $S_1 = \{2, 3\}$. Let a burst of class 1 arrive at t_a and it is to be transmitted at t_s after the base-offset time t_{offset}. Since all the data-channels in the set S_1 are busy at t_s, channel 0 from the set S_0 is selected.

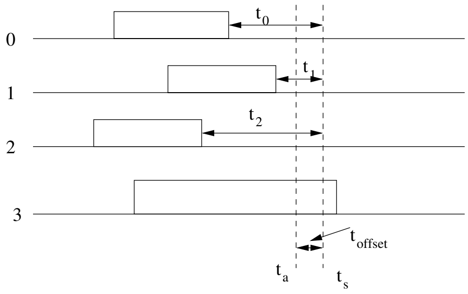

Fig. 4 Illustration for selection of data-channel in LRU and FF algorithms

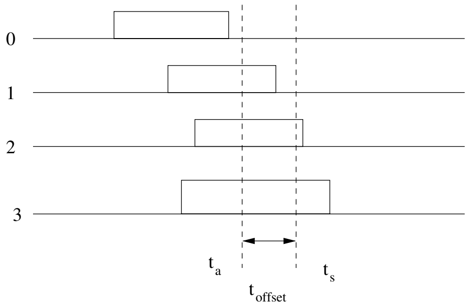

Fig. 5 Illustration for selection of data-channel in PS algorithm

4 Simulation Results

We assume the following time-units for different tasks to carry out the simulation. The propagation delay, t_p, between any two adjacent nodes in the burst switching network is assumed to be $1ms$. The processing time of control packet at the router is assumed to be $2\mu s$. We assume there is no wavelength conversion and there exists no optical buffer in the switch. For simplicity and without loss of generality, we consider two classes of traffic: class 0 (low priority) and class 1 (high priority). We generate high priority traffic with a probability of 0.4. Traffic is generated only at the edge router and, the load is measured in Erlang.

We compare the simulation results of our proposed scheme with PPJET [12]. We consider burst blocking probability as the performance metric for comparison. We have taken number of wavelengths available on each link to be *seven*. Traffic in the Internet is reported to be bursty in nature [13]. We consider bursty traffic with Pareto ($\alpha = 1.1$) distributed burst length and Pareto ($\alpha = 1.1$) distributed inter-arrival time.

We include the overall burst loss for the proposed scheme with three channel selection algorithms, in Fig. 6 and compare with PPJET. It is observed from Fig. 6 that the overall burst loss in our scheme is lower than that in PPJET. Of the proposed channel selection algorithms, LRU algorithm gives lower overall burst loss, and PS gives higher. The higher overall burst loss in PS is due to the higher low priority burst loss. We generated many more results through simulation by varying various parameters; the detailed results will be presented during the conference.

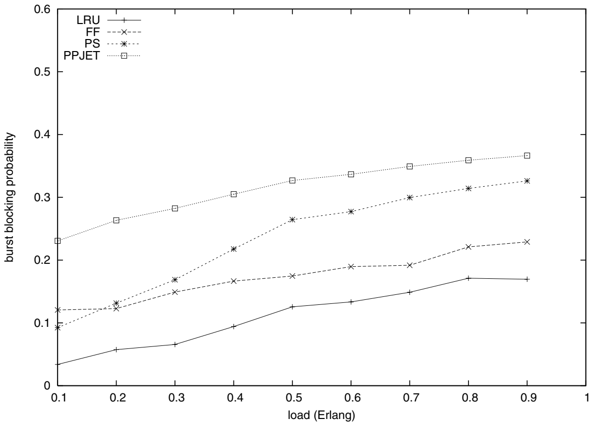

Fig. 6. Overall burst loss in the proposed scheme with different channel selection algorithms and PPJET. Pareto distributed burst-size and Pareto distributed inter-arrival of burst is considered

5 Conclusions

In this paper, we proposed a contention resolution scheme for OBS networks. The scheme takes the following three parameters – priority, number of hops traversed and burst-size - into account to resolve contention. The proposed scheme is adaptable to both prioritized and delay constraint traffic. We also proposed three channel selection algorithms called, LRU, FF and PS algorithms to select data-channel at the ingress router for the proposed scheme. We simulate our scheme with each of the channel selection algorithms and compare the results with PPJET. We consider bursty traffic in our simulation. Simulations were carried out for prioritized traffic. We observed that LRU channel selection algorithm gives lower overall burst loss. In addition, Priority Set channel selection algorithm gives the lowest high priority burst loss.

We compared our scheme with another contention resolution scheme called PPJET. We found lower overall blocking probability in our proposed scheme using LRU than

PPJET for all load. The proposed scheme using PS channel selection algorithm gives the lower blocking for high priority traffic than PPJET. Thus we can conclude that if a lower overall burst loss is required then our scheme with LRU selection algorithm can be used. If a low blocking of high priority traffic is desired then the proposed scheme with PS algorithm may be the choice.

The lower blocking in our scheme comes with an additional delay. In PPJET an incoming burst is delayed for an amount of time which is equal to the total processing time of the control token at each node. However, in our scheme an additional delay which is equal to the propagation time between source to destination, is involved.

References

1. Yao, S., Yoo, S.B., Mukherjee, B., Dixit, S.: All-Optical Packet Switching for Metropolitan Area Networks: Opportunities and Challenges. IEEE Communications Magazine **39** (2001) 142 – 148
2. Chlamtac, I., Ganz, A., Karmi, G.: Lightpath Communications: An Approach to High Bandwidth Optical WANs. IEEE Transactions on Communications **40** (1992) 1171 – 1182
3. Dutta, R., Rouskas, G.: A Survey of Virtual Topology Design Algorithm for Wavelength Routed Optical Networks. Optical Network Magazine **1** (2000) 73–89
4. Yoo, M., Qiao, C.: Optical Burst Switching (OBS) - A New Paradigm for an Optical Internet. Journal of High Speed Network **8** (1999) 69 – 84
5. Qiao, C., Yoo, M.: Choices, Features and Issues in Optical Burst Switching. Optical Network Magazine **1** (2000) 36 – 44
6. Yoo, M., Qiao, C., Dixit, S.: QoS Performance in IP over WDM Networks. IEEE Journal on Selected Areas in Communications, Special Issues on Protocols for Next Generation Optical Internet **18** (2000) 2062 – 2071
7. Kim, H., Lee, S., Song, J.: Optical Burst Switching with Limited Deflection Routing Rules. IEICE Trans. Commun. **E86-B** (2003)
8. Vokkarane, V.M., Zhang, Q., Jue, J.P., Chen, B.: Generalized Burst Assembly and Scheduling Techniques for QoS Support in Optical Burst-Switched Networks. In: Global Telecommunications Conference, 2002, GlOBECOM'02. Volume 3. (2002) 2747 – 2751
9. Zhang, Q., Vokkarane, V.M., Chen, B., Jue, J.P.: Early Drop Scheme for Providing Absolute QoS Differentiation in Optical Burst-Switched Networks. In: Workshop on High Performance Switching and Routing, 2003, HPSR. (2003) 153 – 157
10. Farahmand, F., Jue, J.P.: Look-ahead Window Contention Resolution in Optical Burst Switched Networks. In: Workshop on High Performance Switching and Routing, 2003, HPSR. (2003) 147 – 151
11. D*ü*ser, M., Bayvel, P.: Analysis of a Dynamically Wavelength-Routed Optical Burst Switched Network. Journal of LightWave Technology **20** (2002) 574 – 585
12. Kaheel, A., Alnuweiri, H.: A Strict Priority Scheme for Quality-of-Service Provisioning in Optical Burst Switching Networks. In: Proceedings of the Eight IEEE International Symposium on Computers and Communication (ISCC'03). (2003) 16 – 21
13. Paxson, V., Floyd, S.: Wide Area Traffic: The Failure of Poisson Modeling. IEEE/ACM Transaction on Networking **3** (1995) 226 – 244