Single – Anchored Soft Bandwidth Allocation System with Deflection Routing for Optical Burst Switching

Timucin Ozugur, Farid Farahmand and Dominique Verchere

ALCATEL Research and Innovation Center, TX, USA
Corresponding Author: *Tim.Ozugur@alcatel.com*

Abstract

In Optical Burst-Switched (OBS) networks, the most important problem is the collision problem, which occurs when burst packets contend for the same outgoing interface. TCP-based retransmission is a poor solution since it complicates the ingress Edge Router due to sequencing, and requires excessive amount of memory. We propose a single-anchored soft bandwidth allocation system for OBS networks to provide 100 percent reliability without retransmissions. The anchor node, which is the OBS node that blocks the burst, stores the burst, and informs the ingress Edge Router by a Negative Acknowledgment message. The anchor also provides a bandwidth reservation between itself and the egress OBS node using Probe – Probe-ACK handshake. This handshake provides SOFT ALLOCATION state along the path. The LSPs may be in "anchor mode" or "non-anchor mode". SOFT ALLOCATION does not allow other "anchor mode" LSPs to use the allocated bandwidth, but "non-anchor mode" LSPs may use with low priority. The high priority is reserved for the "anchor mode" LSP. Soft Bandwidth Allocation system provides reliability. Other advantage is that no retransmission is needed in this mechanism.

1. Introduction

Optical Burst Switching, which allows switching of data channels entirely in the optical domain by doing resource allocation in the electronic domain, provides a high utilization than the optical wavelength switching with a moderate hardware implementation complexity. The OBS can be seen as a viable solution to terabit backbone networks by allowing switching of data channels entirely in the optical domain and doing resource allocation in the electronic domain, [1 – 4]. The OBS control packet and corresponding data burst packet, which precedes the control packet, are launched from the edge router at time instants separated by an offset time. The control packet contains information required, such as label, the length of the burst and the offset time, to route the data burst through the optical core backbone. The control packet is sent through out-of-band in-fiber control channels. The control packet is processed electronically at each of the optical cross-connects (OXC) controller to make the routing decisions, such as outgoing fiber and wavelength. The OXCs are configured to switch the data burst, which is expected to arrive after an offset time. The data burst is then switched entirely in the optical domain, thereby removing the electronic bottleneck in the end-to-end data path between the edge routers. Figure 1 shows the OBS network, where burst packet and burst header packet (BHP) follow different sets of channels, Data Channel Group (DCG) and Control Channel Group (CCG), respectively. The channels of a DCG as well as its corresponding CCG could be physically carried on the same fiber or on different fibers. The Edge Routers support the legacy interfaces, such as packet over SONET (PoS), IP over ATM or Gigabit Ethernet, etc.

In an Optical Burst-Switched (OBS) network, the most important problem is the collision problem, which occurs when burst packets contend for the same outgoing interface at each node, [1, 2]. If none of the Fiber Delay Lines (FDL) or wavelengths are available, one burst succeeds, the rest of the bursts are dropped. This probability is referred to as *Burst Blocking Probability (BBP)*. If another wavelength is available, the burst packet is converted to this wavelength using the wavelength converters. It is shown that the BBP may be over 20 percent for fully – utilized OBS network depending on the number of wavelengths at each interface without FDLs, [1]. FDLs may help to ease the burst dropping. However, gain is not very significant due to limited FDL resources.

When a burst belonging to an LSP is blocked in the OBS network, the only way to recover the burst is TCP retransmission. However, TCP does not provide fast recovery due to its host-to-host behavior and time-out mechanism. Moreover, other bursts belonging to this flow are discarded due to Go-Back-N mechanism. If a TCP retransmission mechanism is deployed between ingress-egress Edge Router pair, it complicates the ingress Edge Router due to sequencing, and requires excessive amount of memory since ingress buffers all bursts in its memory space until they are ACKed. The receipts of Acknowledgment (ACK) and Negative Acknowledgment (NACK) are restricted by timers, which complicate the ingress side. If no ACK or NACK receives during a certain time period, ingress side is timed out, and starts re-transmitting the bursts. Selective Repeat only retransmits selectively, however in practice it needs infinite buffering at the egress side.

In this paper, we consider the label-switched OBS networks using Generalized Multi-Protocol Label Switching control plane (GMPLS), [5, 6, 8, 9]. When a burst belonging to a Label-Switched Path (LSP) is blocked in the OBS network, deflection routing cannot be used for label-switched bursts since destination IP address and port is transparent to the bursts and their headers. We propose Soft Bandwidth Allocation system to provide 100 percent reliability without retransmissions. Since there is no retransmission in the proposed method, the ingress side doesn't buffer any burst after it is transmitted. It doesn't use any timers or sequencing mechanism either. In a low-utilized network, there is no blocking, so there is no control messages forwarded. The OBS system takes action only when there is a burst blocking. The burst blocking OBS node stores the burst first. This node becomes an ANCHOR node for this LSP. The ANCHOR node informs the ingress node by a NACK message. Therefore, the NACK control messaging takes place only when there is a burst-blocking event. The NACK message provides the Anchor address to the ingress Edge Router. New burst packets are forwarded to the Anchor node through the LSP, if it not possible they use IP routing (destination = Anchor) through deflection routing.

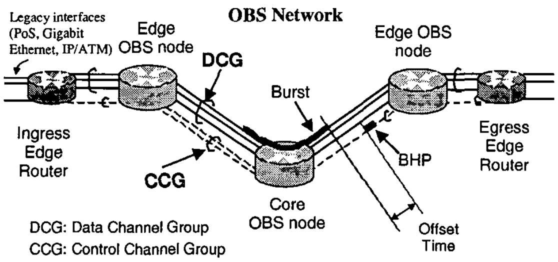


Figure 1. An optical burst-switched network.

When the anchor transmits the NACK message, it also transmits a PROBE message to the egress OBS node at the same time. The Anchor node inserts available time periods of outgoing wavelengths into the PROBE message. When a downstream OBS node receives the PROBE message, it takes out unavailable time periods for its own outgoing wavelengths from the PROBE message (the OBS node considers the propagation delay while calculating the available times). Therefore, the PROBE message only covers time periods for which both nodes (anchor and downstream node) are available. This procedure continues until the PROBE message reaches the egress OBS node. At the receipt of the PROBE message by the egress OBS node, The Probe message shows all available times for all the hops between the Anchor and egress OBS nodes.

The egress OBS node transmits a PROBE ACK message when it receives the PROBE message. The PROBE ACK message is similar to "Allocated Label" object in GMPLS. The PROBE ACK message allocates the certain available time periods in a specific wavelength to this specific LSP throughout the OBS nodes between the Anchor and egress OBS nodes. The allocation amount is based on the negotiated peak burst size of the LSP and bandwidth demand based on the LSP contract. This allocation is a SOFT ALLOCATION. SOFT ALLOCATION doesn't allow other "anchor mode" LSPs to use the allocated bandwidth, but "non-anchor mode" LSPs may use with low priority. High priority is reserved for the "anchor mode" LSP. The PROBE-PROBE ACK handshake provides a reliable transmission in the system without the need of retransmission and possibility of blocking in the downstream. If the system employs a retransmission mechanism between ingress-egress Edge router pair, this provides a path reliability by complicating the ingress Edge router with excessive memory spaces. Soft Bandwidth Allocation system provides 100 percent reliability with a sub-path SOFT ALLOCATION without complicating the ingress or egress Edge Router.

The outline of the paper is as follows: Section 2 describes the new Soft-Bandwidth Allocation system architecture. In Section 3, we give the simulation results. Section 4 proposes the future works and concludes the paper.

2. Soft Bandwidth Allocation System Architecture

In this study, we assume that OBS employs GMPLS to route the burst packets through Label Switched Paths (LSPs). It is obvious that the core of the OBS network is more prone to the burst blocking. We can argue that more LSPs have cut-through at the core part of the OBS than the edge parts of the OBS network. Hence, the core part experiences more burst blocking than the edge parts. For this reason, it is more likely to have an anchor node in the core part of the OBS network.

2.1 Methodology

In this section, we explain the methodology of the Single-Anchored Soft Bandwidth Allocation system. Figure 2(a) shows the OBS system, where an LSP is established from ingress to egress Edge Router. The ingress Edge Router sends the burst packets asynchronously through this path. Assume that a burst packet is dropped at OBS node A. The OBS node A becomes an "**Anchor**" node. The anchor node is given in Figure 2(b). In today's network, the optical data transmission requires regeneration, and optical regeneration is not mature. Therefore, data regeneration is electronic, which requires O-E-O conversion. Electronic regenerators may be incorporated into single-anchored system by allowing only electronic regenerator nodes to be an anchor. If an OBS node, which drops the burst, is not an electronic regenerator, then the blocked node resends the burst data to first upstream electronic regenerator node. Then, this upstream electronic regenerator node becomes the "Anchor".
The anchor node performs the following actions:

1. The burst packet, which will suppose be dropped, goes through optical-electronic (O/E) conversion, and is stored at this Anchor node. An LSP can only have single Anchor node. If there is no burst dropping for the LSP, then there is no Anchor node.
2. After storing the burst, the Anchor node sends two messages at the same time: (1) Negative Acknowledgment (NACK) destined for ingress Edge Router, and PROBE message destined for egress OBS node. PROBE message is processed hop by hop, and RSVP or CR-LDP signaling protocols may be used for this purpose.
3. Egress OBS node returns a PROBE ACK message destined for the Anchor node. This message is also processed hop by hop. Contents of the messages will be explained in the later.

Note that ingress Edge Router doesn't need to deploy any retransmission technique in the Single-Anchored system. Figure 3 shows the signaling messages, NACK, PROBE and PROBE ACK. In this given example, ingress Edge Router receives the NACK message after 3-hop processing time. The egress OBS node receives the PROBE message after 2-hop processing time. Finally, the Anchor node receives the PROBE ACK message in 4-hop processing time after PROBE (and NACK) message is initiated. Therefore, total elapsed time for this example is 4-hop processing time.

In OBS, the control packet and corresponding data burst packet, which precedes the control packet, are launched from the edge router at time instants separated by an offset time. The control packet contains information required, such as label, the length of the burst and the offset time, to route the data burst through the optical core backbone. The control packet is sent through out-of-band, in-fiber control channels. The control packet is processed electronically at each of the optical cross-connects to make the routing decisions, such as outgoing fiber and wavelength. The optical cross-connects are configured to switch the data burst which is expected to arrive after an offset time. By the arrival of the control packet, the OBS node knows whether burst will be blocked or not. Therefore, the OBS node points itself as an Anchor node, and initiates the NACK and PROBE messaging before the corresponding burst arrives.

2.2 Signaling Messages

The actual signaling for setting up, tearing down, and maintaining LSPs can be done either using label distribution protocols (LDPs) or via Resource Reservation Protocols (RSVP), [8, 9]. In the RSVP control plane, the NACK and PROBE-ACK objects are encapsulated into the Resv, and PROBE object is encapsulated into ResvErr messages propagating upstream and downstream, respectively. Note that same objects can be used in LDP-based GMPLS control plane. The messages may also be deployed as independent of the RSVP or LSP control planes. In the following sections, we define the message types for the NACK, PROBE and PROBE-ACK messages.

2.2.1 NACK Message

The NACK message includes the following information:

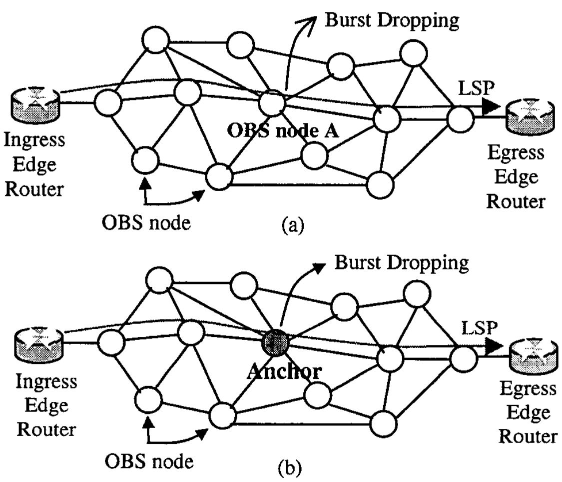


Figure 2. (a) Burst dropping at OBS node A, (b) Anchor node for an LSP.

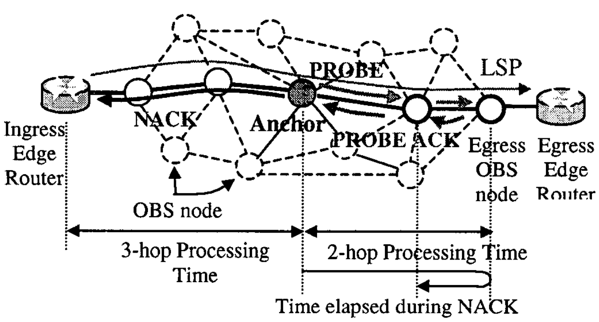


Figure 3. Signaling procedures and messages for Anchor.

<LSP Label, Anchor IP address>

After the receipt of the NACK message, ingress Edge Router transmits burst packets belonging to this LSP to the Anchor node. In the control packet, there should be a bit referring the "Anchor mode" transmission or "Non-Anchor mode' transmission. The control packet should also have an address field for the Anchor node. In the "non-Anchor mode", this field is all zero.

2.2.2 PROBE Message

The PROBE message is similar to "Label Set" object in GMPLS. The PROBE message includes the following information:

<Available Time in all wavelengths>

The Anchor node inserts available time periods of outgoing wavelengths into the PROBE message. When a downstream OBS node receives the PROBE message, it takes out unavailable time periods for its own outgoing wavelengths from the PROBE message (the OBS node considers the propagation delay while calculating the available times). Therefore, the PROBE message only covers time periods for which both nodes (anchor and downstream node) are available. This procedure continues until the PROBE message reaches the egress OBS node. At the receipt of the PROBE message by the egress OBS node, remaining time periods are all available for all the hops between the Anchor and egress OBS nodes.

2.2.3 PROBE ACK Message

The PROBE ACK message is similar to "Allocated Label" object in GMPLS. The PROBE ACK message allocates the certain available time periods in a specific wavelength to this specific LSP throughout the OBS nodes between the Anchor and egress OBS nodes. The allocation amount is based on the peak burst size of the LSP and bandwidth demand based on the LSP contract. This allocation is a SOFT ALLOCATION. Other anchors cannot reserve the same time periods. However, if there is a burst belonging to a non-anchor mode LSP, this burst may use this time period if the anchor mode LSP doesn't have any burst packet at the time being. Therefore, the PROBE ACK message includes the following information:

<soft allocated wavelength, start time, periodic interval>

Figure 4 depicts the SOFT ALLOCATION scheme. Node A is the upstream OBS node, and node B is the downstream node within the LSP. Assume that node A is the Anchor, and node B is the egress OBS node. Node A issues a PROBE message, and advertises the available time intervals. Node B compares its own available time intervals with the one in the PROBE message. Node B finds out that it may allocate the peak burst size of this LSP within wavelength $\lambda 3$. Node B also calculates the periodic interval for the SOFT ALLOCATION. Then, node B advertises this information to node A within a PROBE ACK message. The PROBE ACK can be viewed as a downstream control packet that setting up the allocation tables.

When another PROBE message comes in from a different LSP, the OBS nodes considers these allocated slots as HARD ALLOCATED. However, when a burst is coming from a "non-Anchored mode" LSP, the node may try to allocate this slot to this burst if nothing is available. If the "Anchored mode" LSP doesn't send any burst at this slot, the other burst is successfully allocated. Otherwise, the burst, which belongs to the "Anchored mode LSP, wins this slot, and the burst, which belongs to the "non-anchored mode " LSP is stored in this node, and this node becomes an Anchor for this LSP. Since the burst belonging to an "Anchor mode" LSP has a priority over any other burst, this Single-Anchored system architecture provides a reliable transmission between the anchor and egress Edge Router for this LSP. None of the bursts belonging to the "Anchor mode" LSP drops between the Anchor and the egress Edge Router.

2.3 Deflection Routing for Second Burst Blocking Node

The upstream hops of the Anchor node have still unreliable transmission type. In other words, a burst belonging to an "Anchor mode" LSP may still be blocked at an upstream node. Control packet of a burst packet belonging to the "Anchor mode" LSP follows the LSP until it reaches an OBS node that blocks the burst due to no available resources. Then, this OBS node performs deflection routing for the control packet. The burst packets are destined for the Anchor node through the deflection-routing path. All burst packets should follow the same deflection routing path to prevent sequencing problem. The deflection routing path can be calculated as the shortest path using the OSPF.

2.4 "Anchor mode" Tear Mechanism

Several different schemes may be used to tear down the "Anchor mode". We give two different options in this disclosure:

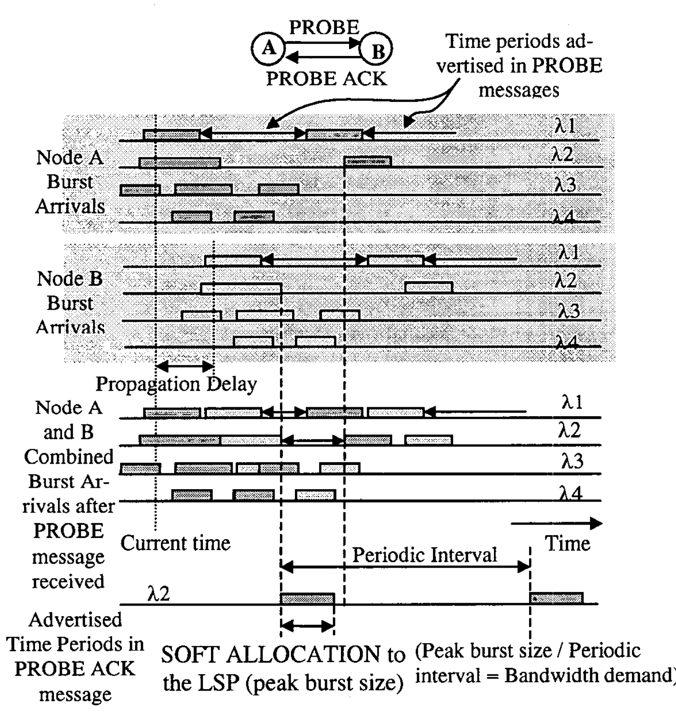


Figure 4. SOFT BANDWIDTH ALLOCATION procedure.

- The Anchor node, itself, is timed out if it doesn't receive any burst packets for this specific LSP. The Anchor sends a "TEAR DOWN" message to the ingress Edge Router. Until this "TEAR DOWN" message reaches the ingress Edge Router, there are propagation and processing delays based on the number of hops. The Anchor node estimates this delay, and it still processes the burst packets of this LSP during this delay. At the elapsed of this delay, the Anchor sends a "TEAR DOWN" message towards downstream to clear the "SOFT ALLOCATION" states.
- The egress Edge Router is timed out if it doesn't receive any burst packets for this specific LSP at the "SOFT ALLOCATED' time slots. It sends a "TEAR DOWN" message to the Anchor, and the Anchor forwards it to the ingress Edge Router. Until this "TEAR DOWN" message reaches the ingress Edge Router, there are propagation and processing delays based on the number of hops. Each node estimates this delay, and they still process the burst packets of this LSP during this delay. At the elapsed of this delay, all the nodes clear the "SOFT ALLOCATION" states without forwarding any control packets.

3. Performance Results

In this section, we simulate the PACNET network model with and without the soft bandwidth allocation (SBA) system. The PACNET network model is given in Figure 5. The PACNET consists of 15 nodes. The unit of length in 1 is 10 km. We assume that each link is bi-directional and has one wavelength. Data rate on each link is 10G. We simulate the network in two different ways: (1) Each station is the source of one LSP. So, the network has 15 LSPs. Destination is chosen randomly with a uniform distribution; therefore source-destination (S-D) pair is uniformly distributed. (2) The network has 12 LSPs, where each LSP has a source from stations 0 to 11. These LSPs randomly select a destination from stations 12, 13 or 14. We refer this simulation scenario bottleneck (S-D) distribution. The bottleneck links are between station 11 and 14, and stations 9 and 12. The duration of each LSP is 10 sec. After the elapse of 10 sec., we select a new destination for the LSP. The burst arrival is exponentially distributed. The average arrival rate ranges from 1 to 500 bursts per sec. The (S-D) routes are calculated using fixed-shortest-path.

In Figure 6(a), we plot burst-blocking rate for each scenario. Referring to the figure, the burst-blocking rate is higher for uniform (S-D) distribution at low burst arrival rates. As the burst arrival rate increases, burst-blocking rate of Scenario 2 (bottleneck (S-D) distribution) exceeds the burst-blocking rate of Scenario 1 (uniform (S-D) distribution). The reason is simple: An increase in the arrival rate has a greater impact on the bottleneck links. As the arrival rate increases, the bottleneck links experience more severe burst blocking. In Figure 6(b), we plot burst-blocking rate for the scenarios with soft bandwidth allocation. Referring to the figure, the uniform (S-D) distribution experiences less burst blocking than the bottleneck (S-D) distribution. In the simulations, we assume the following: After the anchor is established, if the burst is blocked at the non-reliable segment, we don't use the deflection routing, and we drop the burst. That is the reason why the burst-blocking rate is not zero. We assume that the cost of the deflection routing is high in this PACNET network. In the future studies, we will simulate the soft bandwidth allocation system with the deflection routing. We also plan to compare the results with TCP implementation for the burst recovery. For example, PACNET without SBA gives 40 percent of burst blocking for uniform (S-D) distribution at arrival rate of 500 burst per sec. PACNET with SBA gives 20 percent of burst blocking at the same rate. It means that the single anchor only saves 50 percent of the blocked bursts.

In Figure 7, we plot the average burst transmission delay from source to destination versus the good throughput. Good throughput represents the successful burst transmission. As the arrival rate increases, the average delay of bottleneck scenario exceeds the delay of the uniform scenario as expected. It is due to the fact that most of the LSPs establish their anchor at node 9 or 11 because burst blocking occurs generally at bottleneck links right after these links. Note that the average burst transmission delay is approximately 1.18 and 1.56 msec. for uniform (S-D) and bottleneck (S-D) scenarios without the SBA, respectively. The difference is due to the fact the routes in bottleneck scenario is longer than the uniform scenario.

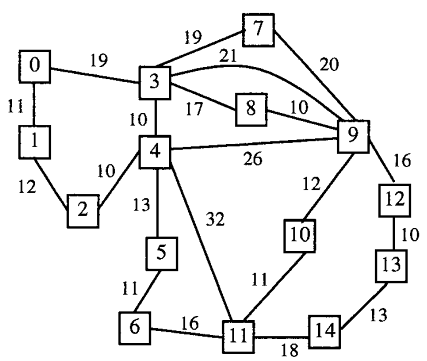
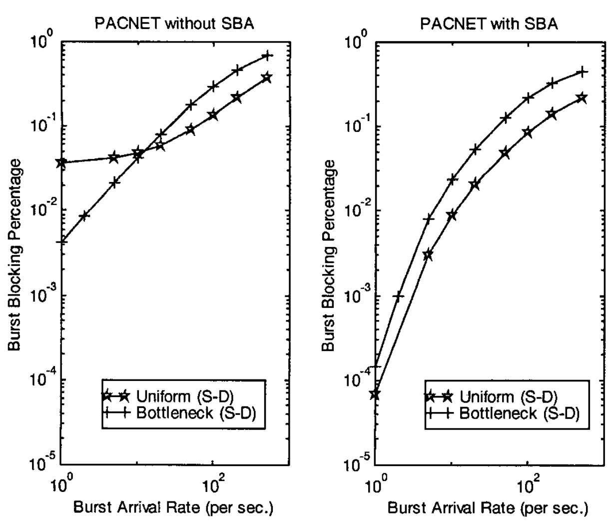


Figure 5. Physical topology of the PACNET network.



(a) (b)

Figure 6. Burst Blocking Rate, (a) without SBD, (b) with SBD.

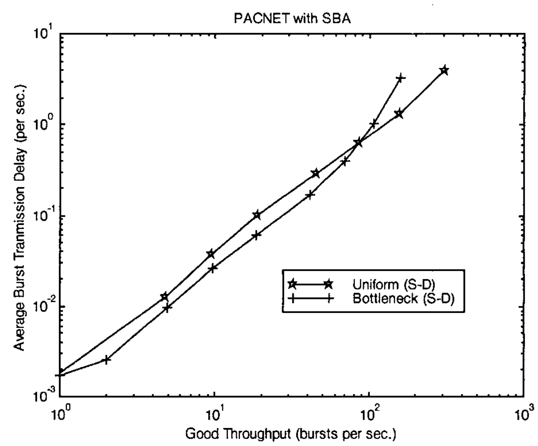


Figure 7. Average Burst Transmission Delay.

4. Conclusions

In this paper, we propose a Single-Anchored Soft Bandwidth Allocation system for OBS networks. The proposed system deploys negative acknowledgment and deflection routing to provide reliable. The Single-Anchored system has the following advantages: (1) No ACK or NACK message unless there is a burst blocking. In retransmission mechanism, there is always ACK and NACK messages even if there is no burst blocking. (2) No need for excessive buffering at the ingress and egress Edge routers. The ingress Edge Router doesn't have any sequencing mechanism or any timers related to the retransmission mechanism. (3) It makes the ingress and egress Edge Routers much simpler. (4) The messaging is very simple, and can be encapsulated into the existing signaling control planes (RSVP/CR-LDP).

References

[1] S. Verma et al., "Optical Burst Switching: A Viable Solution for Terabit IP Backbone," *IEEE Network*, p.48-53, November 2000.

[2] M. Yoo et al., "Optical Burst Switching for Service Differentiation in the Next Generation Optical Internet," *IEEE Commun. Mag.*, p.98-104, February 2001.

[3] M. Yoo, C. Qio, "Optical Burst Switching (OBS) – A New Paradigm for an Optical Internet," *Int'l. J. High Speed Networks*, Vol. 8, No. 1 1999.

[4] Y. Xiong et al., "Control Architecture in Optical Burst Switched WDM Networks," IEEE JSAC, Vol.18, No.10, p.1838-1851, October 2000.

[5] "Generalized MPLS - Signaling Functional Description", draft-ietf-mpls-generalized-signaling-04.txt.

[6] "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", draft-ietf-ccamp-gmpls-architecture-00.txt.

[7] C. Qiao, "Labeled Optical Burst Switching for IP-over-WDM Integration," IEEE Communications Magazine, p.104-114, September 2000.

[8] "Generalized MPLS Signaling - RSVP-TE Extensions", draft-ietf-mpls-generalized-rsvp-te-03.txt.

[9] RFC 2205, "Resource Reservation Protocol (RSVP) - Version 1 Functional Specification", September 1997.