ABSTRACT

JONNADULA, VENKATA R. Performance Analysis of Congestion Control Schemes in OBS Mesh Networks. (Under the direction of Professor Harry G. Perros).

Optical Burst Switching (OBS) is a novel idea offering promising solutions for all-optical WDM networks. It combines the best of optical packet switching and wavelength routing while taking into account the limitations of the current all-optical technology. OBS achieves high traffic throughput and high resource utilization by aggregating multiple packets into a single burst. A control packet sent prior to the burst transmission sets up an optical path between the sender and the receiver. After a pre-determined offset time, the burst itself is transmitted without waiting for an acknowledgment from the receiver. Although OBS networks achieve high resource utilization, burst loss due to contention in the bufferless core can degrade the overall performance of an OBS network significantly. Wavelength conversion (partial and full), optical buffering using fiber delay lines (FDLs) and deflection routing are some approaches for resolving contention. In wavelength conversion, a burst is shifted to a free wavelength on the same link. In buffering, fiber delay lines are used to delay a burst when the outgoing link is busy and in deflection routing, the bursts are deflected to alternate ports in the event of contention. Contention avoidance schemes such as policing the traffic at the source or carefully routing traffic so as to minimize contention can also be used to minimize burst loss in an OBS network.

In this thesis, we analyze and compare the performance of various congestion control schemes in an OBS mesh network using simulations. Partial and full wavelength conversion, deflection routing using FDLs, path restoration using feedback messages and policing at the source node are some of the suggested schemes in the literature for resolving and avoiding contention. These schemes have been studied and evaluated separately in the past. We analyze and compare the performance of some of these congestion control schemes in a single study using simulations, and assuming the following three different traffic classes.

Three traffic classes have been identified for this study. The first class of traffic (Class 1 or Real Time) is a variable bit rate traffic with stringent end-to-end delay constraints, the second class (Class 2) is variable bit rate with no delay constraints, and the third class (Class 3) is non-real-time variable bit rate best effort traffic. The results are compared with respect to the number of real time streams (Class 1) originating at each

source node.

The results show that wavelength conversion is a better approach for congestion control in the OBS networks. Specifically, burst loss can be greatly reduced with just a few converters at each node and with restricted converter capability. Moreover, the results show that excessive deflection routing can degrade the network performance significantly. Simple link deflection does not achieve significant benefit and hence one has to use intelligent path restoration techniques (either partial restoration or complete path restoration) for reaping the benefits of deflection routing. Finally, we observed that OBS mesh networks are suitable for real time traffic given a suitable congestion control mechanism.

Performance Analysis of Congestion Control Schemes in OBS Mesh Networks

by

Venkata R. Jonnadula

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial satisfaction of the
requirements for the Degree of
Master of Science

Department of Computer Science

Raleigh

2004

Approved By:

Dr. Mihail Devetsikiotis

Dr. Rudra Dutta

Dr. Harry G. Perros
Chair of Advisory Committee

To my loving parents . . .

Biography

Venkata Jonnadula was born on November 13, 1981 in Vijayawada, Andhra Pradesh, India. He received his M.Sc(Tech) degree in Information Systems from Birla Institute of Technology & Science, Pilani, India in the summer of 2003. He has been a masters student at the North Carolina State University, in the Department of Computer Science, since August 2003.

Acknowledgements

My foremost thanks to Dr. Harry G. Perros, my advisor, for his constant guidance and support. I am indebted to him for his innumerable ideas about the research work. Without his help and guidance, this work would have never been completed.

I feel very thankful to Dr. Mihail Devetsikiotis and Dr. Rudra Dutta for agreeing to be on my thesis committee.

I thank Vishwas Puttasubbappa and everyone in Dr. Perros' research group for their help on the simulations and for the lively discussions. My gratitude is due to Maneesha, my dearest friend, for being with me at all times and helping me in every aspect. I thank my roommates, Srinath and Mahesh, for all their useful criticisms and for their amazing company.

My thanks are due to my sister, Chandana, and my brother, Venkat, for helping me in every aspect. I am also grateful to my brother, Satya, for his love and friendship. Without their support, I would not be what I am today.

Finally, I reserve my highest gratitude to my motherland India and to my parents. I am forever indebted to my parents, to whom I dedicate this thesis.

Contents

List of Figures

Chapter 1

Introduction

1.1 Optical Switching

The Wavelength Division Multiplexing (WDM) transmission technology has been deployed widely over the last few years in the Internet infrastructure. WDM can become the solution of choice for faster networks that can meet the ever increasing internet traffic. Several different technologies to transfer data over WDM have been developed,such as Wavelength Routed Networks (WRNs), Optical Packet Switching Networks (OPSNs) and Optical Burst Switching Networks (OBSNs), of which, only WRN has been implemented so far. However, WRNs, which employ circuit switching, may not be suitable for different applications that will use the emerging optical internet due to the limited number of wavelengths per fiber. Optical Packet Switching (OPS) has been proposed as a solution to transporting packets over an optical network. The problem with OPS, however, is the lack of practical optical buffer technology. Optical Burst Switching (OBS) is a promising new technology and it can be viewed as lying between OPS and Wavelength routing. The separation of control and data planes allows for higher flexibility and greater network manageability in OBS.

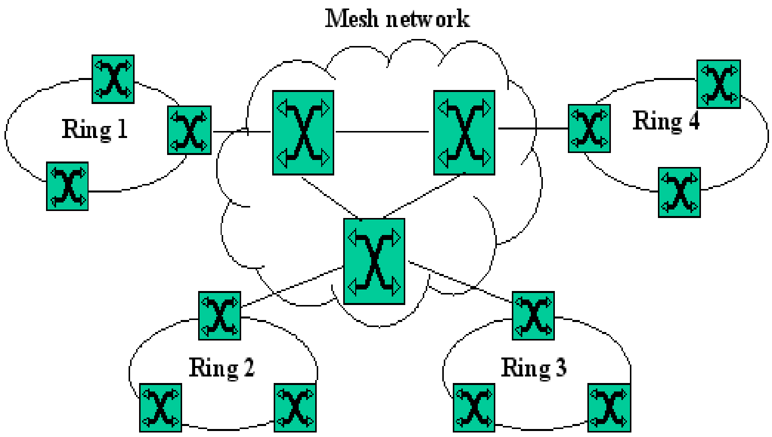

Figure 1.1: An Example of an Optical Network

1.1.1 Wavelength Routed Networks

The wavelength routed networks are built on the concept of *lightpaths*. A wavelength routing network, shown in 1.1, consists of *optical cross-connects (OXCs)* connected by a set of optical fibers. A lightpath, an all-optical path which spans over multiple links in the optical network, is setup between a source and destination for transferring data. The OXCs provide the switching and routing functions for supporting the setup of lightpaths between clients. The main constraint with this approach is the limited number of wavelengths per fiber which makes it difficult to create a full mesh of lightpaths between all end users in a large WRN.

Establishing lightpaths in a wavelength routed network takes at least a round-trip delay which accounts for a poor wavelength utilization if the connection holding time is very short. The bursty nature of internet traffic also leads to poor wavelength utilization. Traffic grooming has to be employed to support statistical multiplexing of data from different users for better wavelength utilization.

1.1.2 Optical Packet Switching

Optical Packet Switching is capable of transporting packets over an optical network. In OPS, the packets along with the control information are transferred over optical links. At each intermediate node, the control information is extracted and processed in the electronic domain, after which, the packet is switched to the next node. The user data is transmitted in optical packets, which are switched within each optical packet switch entirely in an optical domain. Thus, the user data remains as an optical signal for the entire path from source to destination, and no optical-to-electrical and electrical-to-optical conversions are required.

The main problem with this technology is the lack of optical buffers which are necessary to resolve output port contention. An optical packet switch, and in general any packet switch requires buffering. Currently, there are no random access optical buffers. Instead, optical fibers are used to delay packets for a fixed time. Another problem is the delay involved in configuring the switch for each incoming packet.

1.1.3 Optical Burst Switching

Optical Burst Switching is a technique for transmitting bursty traffic. It was based on *ATM block transfer* (ABT), an ITU-T standard for burst switching in ATM networks. Optical Burst Switching [16, 19, 22] has received considerable attention recently as an alternative to optical packet switching. The unit of transmission is a *burst*, which may consist of several IP packets, ATM cells, an HDTV frame or even raw bit streams. Due to the variability in burst size, OBS can be viewed as lying between Wavelength Routing and Optical Packet Switching. Prior to transmitting a burst, a control packet is sent into the network in order to set up a bufferless optical connection all the way to the destination. After an offset delay time, the data burst is transmitted optically. The connection is set up uniquely for the transmission of a single burst, and it is torn down after the burst has been transmitted.

All proposed OBS schemes in literature have some common features described below.

- Granularity: The transmission unit size of OBS is between the optical circuit switching and optical packet switching
- Separation of Control and Data: Control information is transmitted on a separate wavelength (reserved only for control messages)
- Variable Burst Length: The size of the optical burst can vary

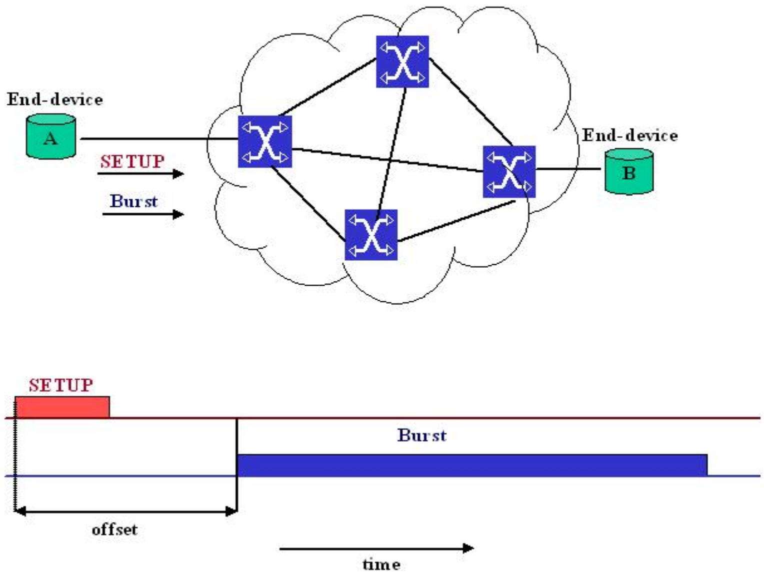

Figure 1.2: An Example of Optical Burst Switching

- No Optical Buffering: The intermediate nodes in the network do not require optical buffers. Bursts go through the intermediate node without any delay

1.2 Congestion Control

A network is said to be congested when the users of the network collectively demand more resources than the network has to offer. Congestion in a network leads to delayed packet arrivals, packet droppings and even denial of service for certain hosts. Congestion control can be achieved using preventive and/or reactive schemes. In preventive congestion control, bandwidth is allocated to a connection at setup time, so that a given QoS is achieved. In reactive congestion control, the rate at which the end systems inject traffic into the network can be regulated or the routing of traffic inside the network can be altered so as to minimize packet contention at the intermediate nodes.

Different congestion control schemes have been proposed for OBS networks, such as,

- Wavelength Conversion: If two bursts are destined to go out of the same output port at the same time, they can still be transmitted but on two different wavelengths. Wavelength converters can be used to convert one of the incoming bursts on to a different wavelength.

- Optical Buffering: Optical buffering can be implemented using Fiber Delay Lines (FDL). An FDL can delay a burst for a specified amount of time, which is related to the length of the delay line. Delay lines are not commercially viable presently.

- Deflection Routing: In this scheme, when there is a conflict between two bursts, one will be routed to the correct output port, and the other will be routed to any available output port. However, the deflected burst may follow a longer path to the destination. As a result, the end-to-end delay for a burst may be unacceptably high. Also, bursts may be delivered out of order at the destination, and therefore, they may have to be reordered.

- Burst Segmentation: When contention occurs, instead of dropping the entire burst, a node breaks the burst into multiple segments, and only the overlapping segments are dropped [20]. In this way, even though the burst loss probability does not change, the loss probability of the data contained in the burst is reduced.

This thesis analyzes two of the above congestion control schemes, namely deflection routing and wavelength conversion. We study the effect of each of these schemes on the burst loss probability with real time and non-real time traffic.

1.3 Thesis Organization

This thesis aims at analyzing and comparing the performance of some congestion control schemes in OBS networks. Chapter 2 presents the concept of Optical Burst Switching and the various congestion control schemes. The chapter also provides a brief introduction to OBS ring networks. We also review the literature on congestion control in OBS networks.

Chapter 3 describes the simulation model used in the thesis. The traffic classes and the network are explained. An event-based simulation model is used to simulate the events in the OBS network. Simulations are used for modelling full and partial wavelength conversion and deflection routing.

In chapter 4, we present and discuss the results that we have obtained. Graphs have been plotted for each of the congestion control schemes. The results provide some new insights on congestion in OBS networks.

Finally, a summary of research contributions and suggested future work are given in chapter 5.

Chapter 2

Optical Burst Switching: Background and Literature Review

2.1 Introduction

In a wavelength routing network, a connection has to be set up before data is allowed to be transmitted. The resources remain allocated to this connection even when there is no traffic to be transmitted. In view of this, the utilization of the connection may be low when the traffic is bursty. *Optical Burst Switching (OBS)* is a different optical networking scheme which is better suited for the transmission of bursty traffic. The name of this technique is attributed to the fact that the data is transmitted optically in bursts.

2.2 OBS Networks

OBS has not as yet been standardized, but it is regarded as a viable solution to the problem of transmitting bursty traffic over an optical network. OBS was based on the *ATM block transfer (ABT)*, an ITU-T standard for a class of service in ATM networks.

Over the last few years, research in OBS networks has rapidly progressed from purely theoretical investigations to prototypes and proof-of-concept demonstrations. Battestelli and Perros [4] provide a detailed survey on OBS and its variations. An OBS network consists of OBS nodes interconnected with WDM fiber in a mesh topology. As OBS node is an OXC and it consists of amplifiers, multiplexers/demultiplexers, a switch fabric and an electronic control unit. The OBS node can switch an optical signal on wavelength λ_i of an input fiber to the same wavelength of an output fiber. If it is equipped with converters, it can switch the optical signal of the incoming wavelength λ_i to another free wavelength of the same output fiber, should wavelength λ_i of the output fiber is in use. (We assume here full conversion, and each converter can convert an optical signal to any other wavelength). Unlike wavelength routing networks where a connection may stay up for a long time, the switch fabric of an OBS node has to have an extremely short configuration time.

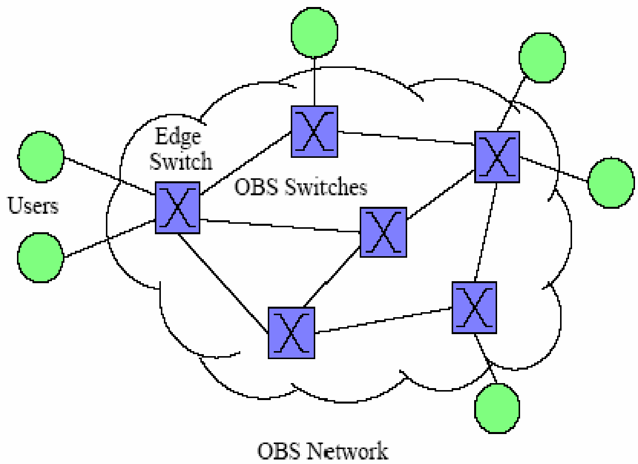

Figure 2.1: An OBS Network

An OBS end-device (IP routers, ATM switches or Frame Relay switches equipped with an OBS interface) collects traffic from various electrical networks, such as ATM, IP and Frame Relay, and then transmits it to the destination OBS end-device optically through the OBS network. The collected data is sorted based on a destination OBS end-device address and it is assembled into larger size units, called *bursts*. In order to transmit a burst, the end-device first transmits a control packet, and after a delay, known as the *offset*, it transmits its burst. The control packet contains information such as the burst length and the burst destination address, and it is basically a request to setup a connection, i.e., a lightpath, end-to-end. After the transmission is completed, the connection is torn down. Up to W bursts per fiber may be simultaneously transmitted if each link between two adjacent OBS nodes carries W wavelengths. An end-device may also have the ability to transmit W bursts to its ingress OBS node simultaneously, or it may have only one wavelength available on which to transmit one burst at a time.

The control packets may be transmitted optically on a designated signaling wavelength or electronically over a packet switched network, such as an IP or ATM network. In either case, the control packet can only be processed by each OBS node electronically. There exists a separation of control and data, both in time and physical space. It facilitates efficient electronic control while it allows for a greater flexibility in the format and transmission rate of the user data since the bursts are transmitted entirely as optical signals which remain transparent throughout the network,

The time it takes for the connection to be setup depends on the end-to-end propagation delay of the control packet, the sum of all the processing delays of the control packet at all the intermediate OBS nodes, and configuration delays. (Depending upon the propagation delay between adjacent OBS nodes, the configuration delay at each node may overlap to some extent). Now, the time it takes for a burst to reach the destination end-device is equal to the end-to-end propagation delay, since it is transmitted as an optical signal that traverses the OBS switches without any processing or buffering delays. In view of this, the transmission of a burst is delayed by an offset so that it always arrives at an OBS node, after the control unit of the node had the chance to process the control packet associated with the burst and configure the OXC. Let t_{proc} be the time it takes to process a control packet at a node, t_{conf} be the time it takes to configure an OXC, and N the total number of OBS nodes along the path of the burst. Then a good upper bound for the offset can be obtained using the expression: $Nt_{proc} + t_{conf}$, see figure 2.2.

If the offset is not long enough, then the burst may arrive at an OXC before the control packet, or before the OXC has a chance to configure its switch. Hence, the OXC cannot be configured for the incoming burst and hence has to be dropped, see figure 2.3.

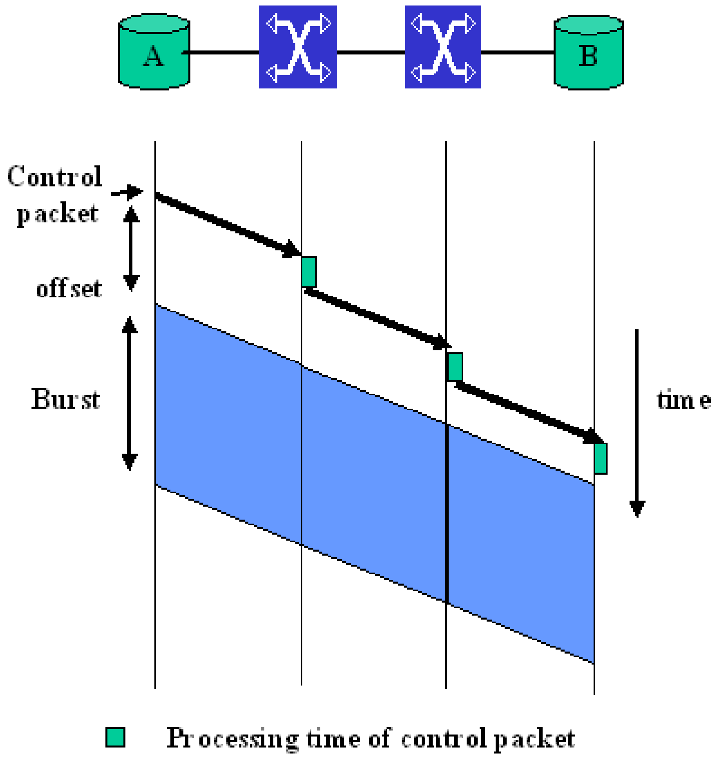

Figure 2.2: The Offset for a Control Packet

2.2.1 Burst assembly

Each end-device maintains one queue per destination end-device. Packets arriving at the end-device are placed accordingly into the various destination queues, from where they are transmitted over the OBS network in bursts. The size of the burst is controlled by a timer. When the timer associated with a destination queue expires, all the data in the queue are grouped into a single burst which is then transmitted out according to the OBS scheme. The burst size has to be less than a maximum since large bursts tend to occupy the network resources for long periods of time thus blocking the transmission of other bursts. Also, the size of the burst has to be greater than a minimum because of the signaling overheads for setting up a connection. Therefore, when the timer expires, a burst is assembled if its size is greater than the required minimum. Also, a burst can be assembled before the timer expires if the data in the queue reaches the maximum burst size.

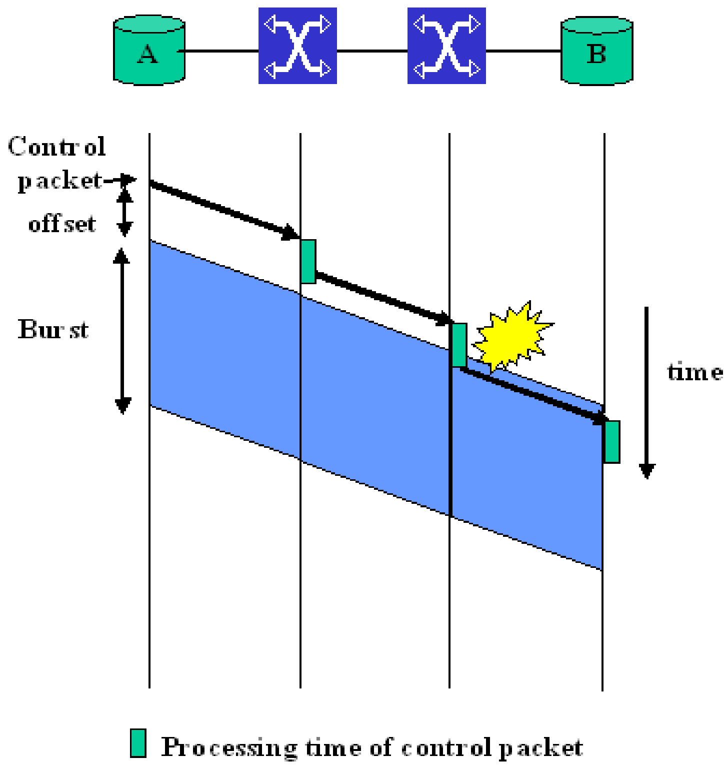

Figure 2.3: A Low Offset for a Control Packet

The duration of the timer, and the maximum and minimum burst sizes can be estimated using modelling techniques, such as simulation.

The end-device can also introduce priorities when transmitting bursts. Specifically, each destination queue may be further sub-divided to a number of quality-of-service queues. The arriving packets are grouped into these queues, which are served according to a scheduler. In addition, different timers and maximum/minimum burst sizes can be used for different queues.

2.2.2 Connection Setup Schemes

Two different schemes can be used to setup a connection, namely *on-the-fly connection setup* and *confirmed connection setup*.

On-the-fly Connection Setup In the on-the-fly connection setup scheme, a burst is transmitted after an offset without any knowledge of whether the connection has been successfully established end-to-end. Since the source does not wait for the acknowledgement, the offset size is between the transmission time of the control packet and the round-trip delay of the control packet.

In *just-enough-time (JET)* proposed by Qiao and Yoo [16], and in *just-in-time (JIT)* proposed by Wei and McFarland [22], the offset is selected in a manner that takes into account the processing delays of the control packet at the intermediate switches. The switches are already configured when the burst arrives. At the destination node, the burst arrives just after it has been configured. Yoo et al. [27] also described a priority scheme for JET to support multi-class traffic. Specifically, a burst with a higher class is assigned an additional offset. They also analyzed the performance of this priority scheme assuming no class interference. Dolzer et al. [9] proposed an iteration algorithm to analyze approximately this priority scheme considering the class interference, and they found that the burst inter-arrival time distribution has only a small impact on the burst loss probability, while the burst length distribution and the ratio of mean burst length of the classes have a great impact on the burst loss probability.

Verma et al. [7] [8] presented a traffic shaping scheme which randomizes the offset to reduce the burst loss probability.

Confirmed Connection Setup In the confirmed connection setup scheme, see figure 2.4, a burst is transmitted after the end-device receives a confirmation from the OBS network that the connection has been established. It is very close to optical circuit switching in that it incurs a round-trip delay to setup the transmission, and since the control packet reserves resources, delivery of the burst is guaranteed. The optical burst switching scheme called *Tell And Wait (TAW)* [23] belongs to this class. The long offset time causes long data delay making this scheme unsuitable for real time traffic.

In another variant called the *Tell And Go setup (TAG)* [23], the source transmits the setup message and immediately after, it transmits the optical burst. However, there is no time for the OBS node to process the setup message and to configure its switch fabric on time so that it can transmit the incoming burst to its destination output port. This

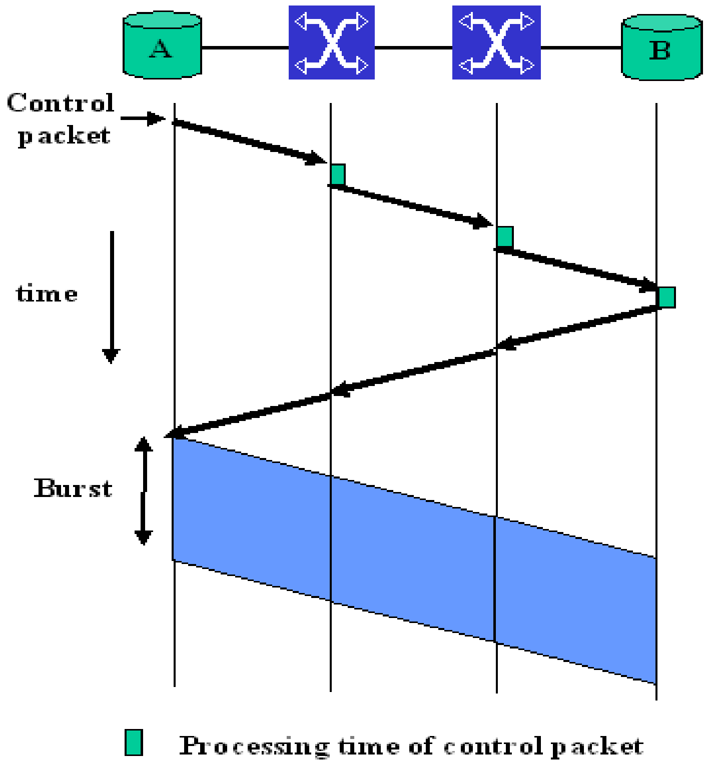

Figure 2.4: Confirmed Connection Setup for OBS

scheme is practical only when the switch configuration and switch processing time of a control packet are very short. In order to implement this scheme, some optical buffering may be required to hold the burst until the node can process the setup message.

On-the-fly connection setup can achieve better throughput than the confirmed connection setup scheme because the bursts can be transmitted without waiting for positive acknowledgments from the network. Confirmed connection setup incurs a round-trip delay to set up the transmission, and the delivery of the burst is guaranteed.

2.2.3 Reservation and release of resources in an OXC

Baldine et al. [3] described four different types of wavelength reservation schemes based on the amount of time a burst occupies a path inside the switching fabric of an OBS node. There are two schemes that can be used to configure an OXC, namely, *Immediate* *setup* and *estimated setup*. In the immediate setup scheme, the control unit configures the OXC to switch the burst output port immediately after it has processed the control packet. In the estimated setup scheme, the control unit calculates the time of arrival $t_{arrival}$ of the burst at the node, and it then waits to configure the OXC at $t_{arrival}$.

Two different schemes also exist as to when the control unit will instruct the OXC to release the resources for switching the burst, namely, *estimated release* and *explicit release*. In the estimated release scheme, the control unit calculates when the burst will completely go through the OXC, and when the time occurs, it instructs the OXC to release the allocated resources. This requires knowledge of the burst duration. An alternative scheme is the explicit release scheme, where the transmitting end-device sends a message to inform the OBS nodes along the path of the burst that it has finished its transmission. A control unit instructs the OXC to release the connection when it receives this message.

Combining the two setup schemes with the two release schemes gives the following four different mechanisms.

- Immediate setup with explicit release (Scheme 1): In this scheme, the reservation of the wavelength starts immediately after the OXC receives the control packet and ends when the release message arrives, refer figure 2.5. Configuration of the wavelength for this scheme is easy as the OXC needs to maintain a single on/off bit to note the busy and free states of the wavelength.

- Immediate setup with estimated release (Scheme 2): The control packet contains the offset and the duration of the burst. The OXC calculates an estimated time of when the burst transmission would be completed. The reservation of wavelength starts when the OXC receives the control packet. The end of reservation is based on the estimated time of completion.

- Estimated setup with explicit release (Scheme 3): In this scheme, the control packet contains the offset of the burst, but not the duration of the burst. The reservation starts at the beginning of the burst minus the time to configure the switch, which is calculated using the offset of the burst, and ends after the release message arrives.

- Estimated setup with estimated release (Scheme 4): In this scheme, the control packet contains both the offset and the duration of the burst. The reservation starts at the beginning of the burst minus the time to configure the switch, which is calculated

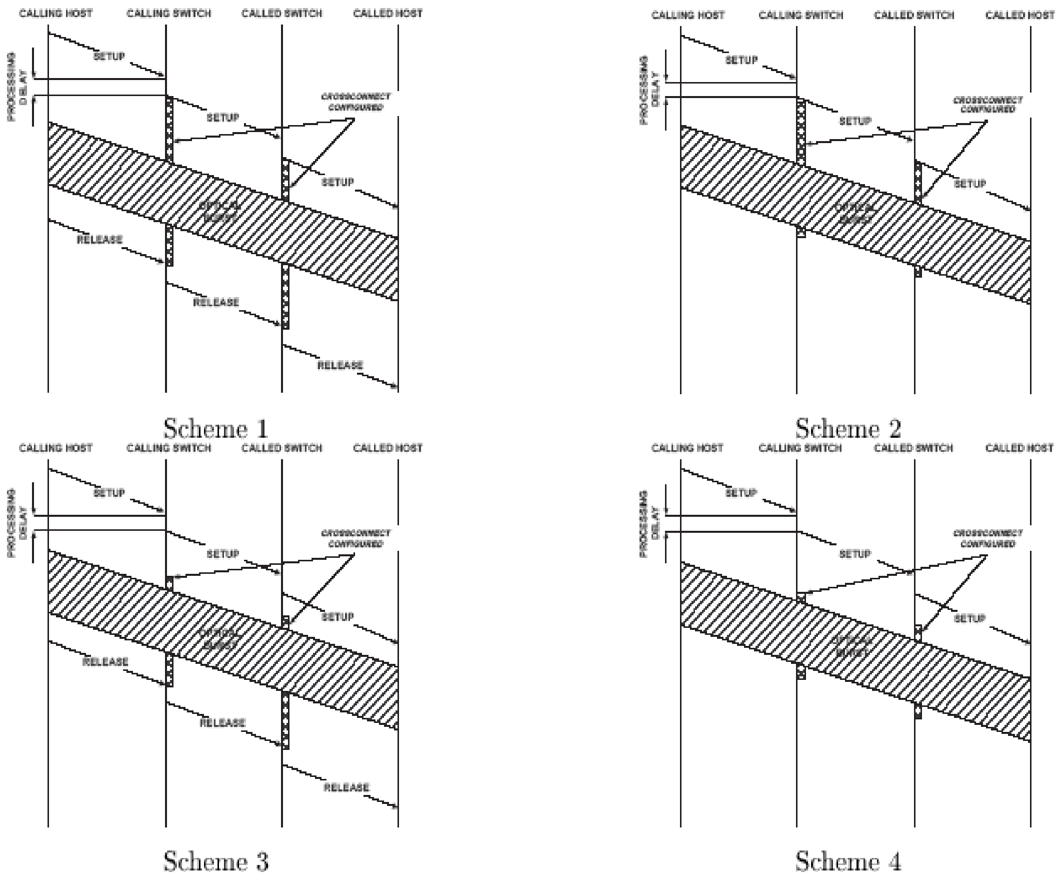

Figure 2.5: The Four Reservation Schemes

using the offset of the burst, and ends at the end of the burst, which is calculated using the duration of the burst. Each wavelength is associated with a vector of time period, which includes the period the wavelength is busy.

Just-In-Time (JIT) proposed by Wei and McFarland [22] belongs to scheme 1. The Horizon scheme proposed by Turner [19] for a scalable burst switch architecture belongs to scheme 2. *Just-Enough-Time* (JET) proposed by Qiao and Yoo [16] belongs to scheme 4. Guager et al.

Dolzer et al. [9] compared the burst loss probabilities of schemes 1, 2 and 4. They found that with variable offset time, scheme 4 achieves the lowest burst loss probability, followed by scheme 2, and scheme 1 achieves the highest burst loss probability.

2.2.4 The Jumpstart Project

The Jumpstart project [1] was an ARDA-supported research project between NCSU and MCNC, addressing the design, specification, performance evaluation, and hardware implementation of a signaling protocol for OBS networks. The signaling protocol follows the just-in-time (JIT) approach (explicit setup and explicit release), and part of this signaling scheme has been reported by Baldine et al [3]. The Jumpstart project [1] is designed to support scheme 1 and scheme 2, unicast and multicast, short bursts and light-paths, persistent path setup and on-the-fly path setup, and best effort and quality-of-service.

The Jumpstart architecture is based on the immediate setup with timed or explicit release and it uses both on-the-fly and confirmed connection setup methods. A single high-capacity signaling channel is assumed for each WDM fiber. The signaling channel can be one of the W wavelengths of the fiber, or it can be implemented in the electric domain using a packet-switched network. In the Jumpstart prototype network, it was implemented using an ATM network. The signaling messages associated with the establishment and tearing down of connections for burst transmission are processed in hardware in order to assure fast connection establishment. Other control messages are processed in software. The bursts are transmitted through the intermediate OBS nodes transparently without any electro-optical conversion. In view of this, optical digital signals of different formats and modulations, as well as analog signals can be transmitted. No global synchronization scheme is required between the OBS nodes.

The following are some of the signaling messages that were defined:

- SETUP
- SETUP ACK
- KEEP ALIVE
- RELEASE
- CONNECT
- FAILURE

Jumpstart uses a hierarchical addressing scheme with variable length addresses similar in spirit to the NSAP address format. Each OBS node consists of an OXC which is a 2D MEMS switch fabric, and a control unit, which is known as the *JITPAC Controller* or simply the *JITPAC*. The JITPAC controllers communicate with each other via an ATM network. Because of the nature of optical burst switching, the signaling messages are associated with bursts and they follow the same path as their bursts. In both the control plane and the data plane, the forwarding of a burst or a control message may be done using either the next hop information as in the IP network, or using labels as in the MPLS domain.

2.3 OBS Ring Networks

The survey on OBS networks remains incomplete without the mention of OBS ring networks. Xu et al. [24] proposed an OBS ring architecture consisting of N OBS nodes connected by optical fibers, with $N+1$ wavelengths. The ring can be a metropolitan are network (MAN) serving as the backbone that interconnects a number of access networks as shown in figure 2.6. Each of the OBS nodes has a fixed transmitter, set to one of the N wavelengths called the *home wavelength*, and a tunable receiver so that it can receive bursts along the transmission wavelengths of the other nodes. In addition, each of the OBS nodes in equipped with a secondary pair of a fixed transmitter and a fixed receiver, set to the separate control wavelength, in order to communicate control information along the ring. The data waiting for transmission is organized into queues according to their destination. The transmit queues are served in a Round-Robin fashion.

In this OBS ring architecture, it is possible for two OBS nodes to send bursts, overlapping in time, toward the same destination. Consequently, these bursts will contend for the tunable receiver of the destination node and one of them will have to be dropped. Xu et al. [24] proposed various access protocols to alleviate burst contention problem and analyzed their performance in terms of throughput, packet delay, throughput fairness and delay fairness.

Vishwas et al. [15], extended these access protocols to the case where different types of traffic such as HDTV, SAN data and best effort data is transported over the ring. Traffic classes like real-time variable bit rate (class 1), variable bit rate with no stringent

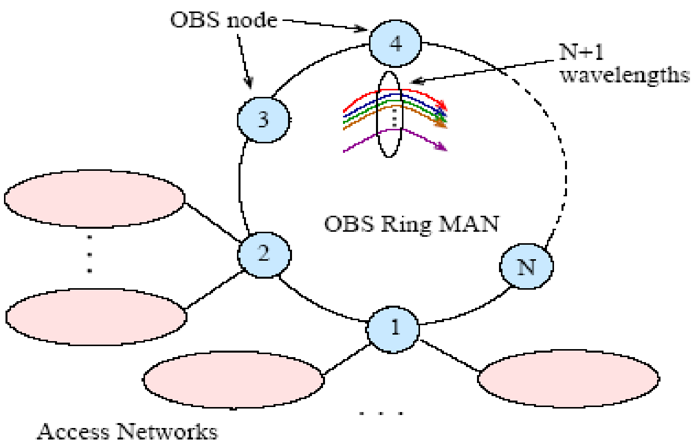

Figure 2.6: An OBS Ring Network

end-to-end delay constraints (class 2) and best effort traffic (class 3) have been used for simulations. The access protocols serve to minimize burst contention in the OBS ring. In *Destination-Reservation Free Protocol*, the nodes transmit bursts without making any reservations at the receiver node. Hence, there is no guaranteed acceptance of transmitted bursts. Tokens are used to resolve receiver collisions for class 1 bursts in *Token Protocol*. All the nodes can transmit bursts only if they possess the token to transmit. A request and acknowledgement mechanism is employed in *Ack Protocol* for class 1 traffic. *Token-Token Protocol* is a collision free protocol for class 1 and class 2 bursts. Nodes use the token mechanism to ensure that the bursts belonging to class1 or class 2 categories are received without receiver conflict. *Ack-Ack Protocol* ensures guaranteed reception both for class 1 and class 2 bursts through acknowledgements.

The simulation results indicate that token based protocols are more scalable with ring size and number of nodes in the ring. Ack and Token protocols provided guaranteed delivery to class 1 bursts but do not ensure zero burst loss for classes 2 and 3. Nevertheless, the simulation results indicate that OBS ring networks are capable of supporting real time traffic through careful selection of access protocols.

Jong [12] proposed several access protocols for multicasting in an OBS Ring network. A new architecture called the LightRing has been proposed by Fumagalli and Krishnamoorthy [10] with multi-token protocol to prevent contention among bursts. Each node can transmit on any of the wavelengths as long as it has the token associated with that particular wavelength. Several *Burst Assembly and Transmission (BAT)* strategies which deal with simultaneous assembly and scheduling of bursts are proposed. Packets from different flows can be assembled into the same burst so as to achieve lower latency of real-time packets. Bouabdallah et al. [5] proposed a collision avoidance MAC protocol for a metropolitan bus-based optical access network. Analytical models were developed to calculate the mean access delay of each node in such a shared-medium system. Fairness issues were also investigated.

2.4 Congestion Control in OBS Networks

Congestion in a network can be identified by several inter-related congestion indicators.

- Number of bursts dropped at an OXC exceeds a certain threshold
- Buffer size at a user exceeds a pre-defined threshold
- The input queue in a JITPACK (Jumpstart) exceeds a pre-defined threshold

Burst loss is an important performance measure of an OBS network, and it occurs when the on-the-fly connection setup is used. When a control unit receives a control packet, it extracts the burst's destination from the packet and it then makes a decision as to whether its OXC can switch the burst or not. If it cannot switch the burst, it will not forward the control packet to the next OBS node, and it will send a negative acknowledgment back towards the transmitting end-device in order to release any resources already allocated. However, the end-device may have already transmitted the burst by the time it receives the negative acknowledgment, which means the burst will get lost upon arrival at the node.

A small number of FDLs can be used in order to reduce the burst loss rate, though as in the case of the optical packet switching, FDLs are not commercializable. An alternative solution is to use converters. In this case, if a burst is coming into an OXC on a wavelength which is in use at the destination output fiber, the burst can be converted to another free wavelength.

Finally, deflection routing can be used to divert a burst that it would be otherwise lost, to an alternative path towards the destination end-device. The alternative path, however, may have more hops than the original one and therefore the offset may not be long enough to prevent burst loss. A possible remedy is to delay the deflected burst in an FDL so that its offset from the control packet is appropriately long. Another solution is to use an offset which is the upper bound of all offsets within the autonomous system. This simple solution avoids the need for FDLs, at the expense of delaying the transmission of some of the bursts more than necessary. Finally, deflected bursts impose an extra load on the alternative links, and this has to be taken into account when planning the capacity of the OBS network.

2.4.1 Deflection Routing

Deflection routing is perhaps the most studied congestion control scheme in OBS networks. Although deflection routing reduces burst losses up to some degree, there are several factors that need to be taken care while using this scheme.

- Deflection requires re-calculation of the offset
- The bursts have to be delayed appropriately
- Alternate routes are required to be calculated
- Implementation of deflection routing is dependent on whether we use FDLs or not

When using FDLs for deflection routing, we can either choose to use common set of FDLs per port or per OXC. The burst is routed to the FDLs, assuming that there is space. If there is no space, the burst is dropped. Another approach is to use an FDL for each incoming wavelength. The FDL will delay the burst for as long as it takes to process a control packet message. This method eliminates the need for offsets totally.

Without using FDLs for deflection routing, the only available solution is to use a large offset, which is adequate for all possible paths within the OBS network. But if the OBS network is large, the throughput may be affected with this large offset since each burst has to wait for an extra delay before being sent into the network.

Deflection routing can lead to some interesting problems in OBS networks. One important observation is that excessive deflection may have negative effects on the network's performance. Excessive deflection causes most of the links to be busy which are otherwise free to transmit non-deflected bursts. Another problem with deflection routing is that the bursts may be delivered out of order and this requires re-ordering in the receiver's electronic domain.

Different types of deflection schemes are possible in an OBS network.

- One-hop Deflection: The routing algorithm calculates for each OXC a *primary, secondary*, and may be a *tertiary* path to each destination. It then downloads the primary, secondary, and the tertiary next-hop to each OXC. When congestion occurs on a link, the OXC diverts the incoming burst on its secondary next-hop for that destination, figure 2.7. If secondary next-hop is over a congested link, then it uses the tertiary next hop. Note that one-hop deflection incurs the problem of routing loops in the network which is indicated by a dotted arrow in the given figure.

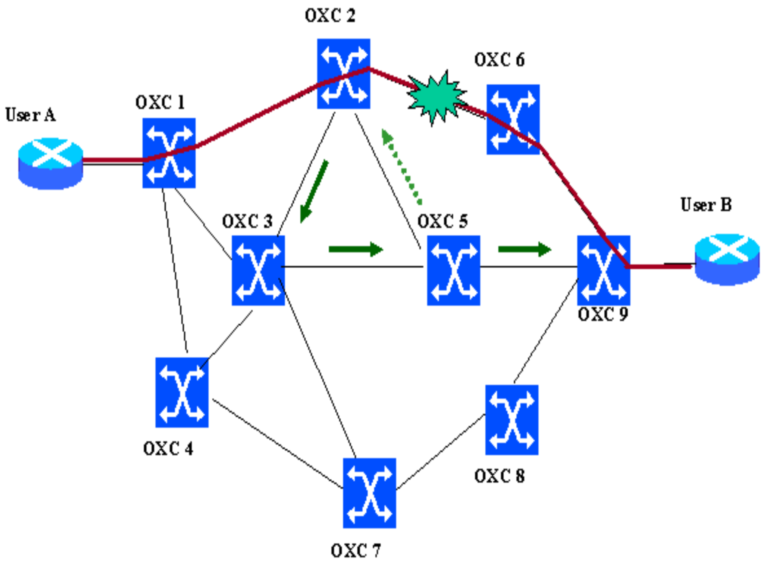

Figure 2.7: One-Hop Deflection in an OBS Network

- Path Deflection: The routing algorithm calculates for each OXC a *primary, secondary*, and may be a *tertiary* path to each destination. When a link of an OXC becomes congested, the OXC diverts an incoming burst to its secondary path (or tertiary path if secondary path is occupied). This path has to be followed by all the OXCs along the secondary path, figure 2.8. End-to-end path restoration can be used as an extreme measure and can be triggered by feedback messages from the OXCs.

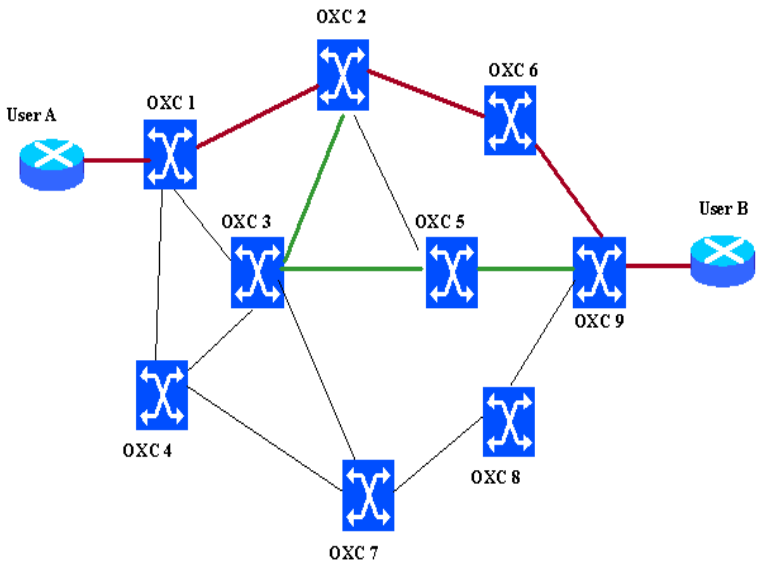

Figure 2.8: Path Deflection in an OBS Network

Wang et al. [21] propose a deflection routing protocol which includes four basic functions. The *Deflection function* allows an OBS switch to deflect the burst to any idle link when the default output link is in use. Although short term loops are possible because of this approach, a *ttl* field in the control header prevents infinite loops. The *Sender Check Function* prevents a sender from deflecting its own burst. The sender does not send out a burst until the default link becomes idle. In the case that a burst is deflected back to the sender, the sender does not perform the deflection forwarding again. With this check function, limited buffering becomes available for contention resolution. In *Sender Retransmission Function*, if the default output link is in use, the sender will attempt to retransmit the burst after a certain interval time which depends on the burst length. This function increases the average retransmission interval exponentially which suppresses traffic explosion and reduces the blocking probability. Lastly, the *Burst-length Decision Function* updates the average burst length for each burst switch is dynamically updated based in the link speed and the average RTT. For a 4 x 4 grid network and 16 wavelengths, the above deflection protocol produces significant improvement over no deflection case when the traffic load is under 40%. For higher traffic load ($\geq$ 40%), the improvement is not very significant although the proposed technique always has a lower blocking probability than the no deflection case.

Hsu et al. [11] have done a performance analysis of deflection routing in optical burst-switched networks based on the Just-Enough-Time (JET) signaling. They analyzed the impact of deflection routing on JET-based OBS networks by analysis and simulations. A queuing model has been proposed to estimate the burst loss probability. Deflection routing will not succeed without enough offset time or buffered delay. *Extra offset time*, *Delayed-at-previous-hops*, *Delayed-at-congested-node* and *Delayed-at-next-hop* are four techniques for solving this problem. The results indicate that deflection routing brings significant performance gain, especially with fewer wavelengths and under light load. Excessive deflection induces highest burst loss because of immoderate usage of resources.

Shortest Path Prioritized Random Deflection Routing (SP–PRDR), see Craig [6], combines standard Internet Protocol next-hop forwarding with prioritized random deflection for control packets and corresponding bursts. Craig et al. [6] show through analysis and simulation, that by deflecting otherwise blocked bursts to a random unused output link and reclassifying these deflected bursts with a strictly lower and preemptable priority label, the burst loss rate for all ranges of input loads is strictly decreased compared to the no-deflection case.

2.4.2 Wavelength Conversion

Wavelength conversion, see figure 2.9, is seen as important alternative to deflection routing. The use of converters at an output fiber permits a better utilization of the fiber's wavelengths. A few converters can greatly reduce burst loss in an OBS network. When a burst arrives at an OXC and the wavelength on the output port is busy, a converter is capable of converting the burst on to a different free wavelength on the same output port. Wavelength conversion can be either full or partial. In *full wavelength conversion*, there are as many converters as the number of wavelengths. In *partial wavelength conversion*, there are fewer converters than wavelengths. In general, a converter has a limited conversion range d. That is, it can convert the wavelength of an incoming burst λ to any other wavelength within the range $(\lambda\text{-}d,\ \lambda+d)$. As an example, consider a converter with a conversion range of *2*. If a burst arrives on a wavelength λ, the converter can convert this burst on to any of the *5* wavelengths in the range λ-*2* to λ+*2*. Though partial conversion offers only a limited conversion capability, it can achieve very good results.

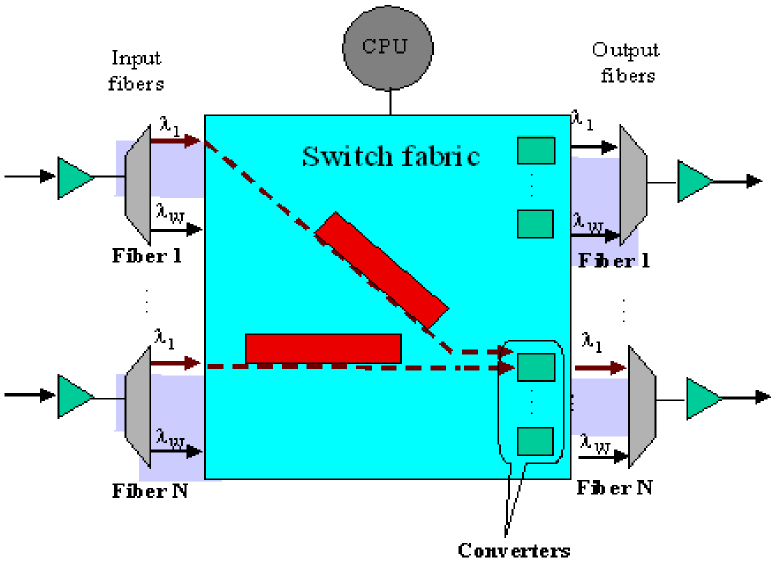

Figure 2.9: Wavelength Conversion in an OBS Node

A combination of two or more congestion control schemes can also be used sometimes to achieve better performance in OBS networks. In our studies, we have simulated wavelength conversion and deflection routing together.

2.4.3 Related Work

Yoo et al [26] analyze policies to handle blocked bursts such as drop, retransmission, deflection routing and buffering in the multi-hop very high performance Backbone Network Service (vBNS). The performance of the above policies is compared in terms of the average wavelength utilization, the average wavelength efficiency and the average end-to-end extra delay. They show that the buffering policy is useful with scarce network resources (e.g., bandwidth), but the dropping policy in conjunction with the offset-time-based scheme is good enough with abundant resources. The simulations also reveal that limited FDL buffer can greatly improve the QoS performance of real-time applications with scarce network resources.

Localized contention resolution techniques do not address adequately the more fundamental problem of congestion under heavy load. Hence, network level contention avoidance using load balanced routing techniques, see Jason [17], has been studied as an alternative. Jason et al. [17] proposed two techniques to avoid contention. The first is a congestion-based dynamic route selection technique using fixed alternate shortest path routes. The second is a least-congested dynamic route calculation technique with different weight functions. In both the techniques, the core nodes in the network gather the load information on their output links and send feedback to all the edge nodes so as to enable the edge nodes to balance the load. The congestion information gathered at each core node is the offered load in a link. The simulation results show that the proposed contention avoidance techniques improve the network utilization and reduce the packet loss probability significantly.

Yoo et al. [27] studied the effect of fiber optic delay line on the burst loss probability, and they found that optical delay lines can reduce the burst loss probability, and the length of the optical delay line required to achieve a given burst loss probability can be significantly reduced as the number of wavelengths increases. Kim et al. [13] presented a deflection routing algorithm. Besides the normal offset time, each burst is assigned a "routing offset time" to accommodate the extra burst delay of the deflection routing. They found that by introducing a 10% routing offset time, the burst loss probability could be significantly reduced.

An elaborate list of all publications on Optical Burst Switching can be found at Vinod [2].

Chapter 3

Simulation Setup

An event-based simulation model was constructed with a view to analyzing the performance of various congestion control schemes in OBS mesh networks. The simulation model was constructed following the typical event-based simulation techniques, see Perros [14]. In the simulation model, the actual data plane is not simulated. The equations for traffic classes have been borrowed from Vishwas [15].

3.1 Traffic Classes

Three different types of traffic classes have been implemented in the simulation model.

- Class 1: The class 1 traffic is assumed to be HDTV traffic. To this effect, for each node i, i=1,2,...,N, a number of HDTV streams are setup at the beginning of the simulation. Each of these streams originate at node i and terminate at destination node j. In each stream, frames are generated at a rate of 60 frames per second giving an inter-frame arrival time of 16.667 milliseconds. We assume that the frames follow the MPEG 2 Group of Pictures (GOP) structure of IBB PBB PBB PBB. The size of each frame is generated using the auto-regressive model, S(t) - S(t-12) = e(t) - 0.69748 x e(t-3), where S(t) is the size of frame t, and e(t) N(0,σ^2) with σ^2=4849.5. This autoregressive model produces an average bit rate of 20 Mbps. A trace of frame sizes generated over a period of time is plotted in figure 3.1.

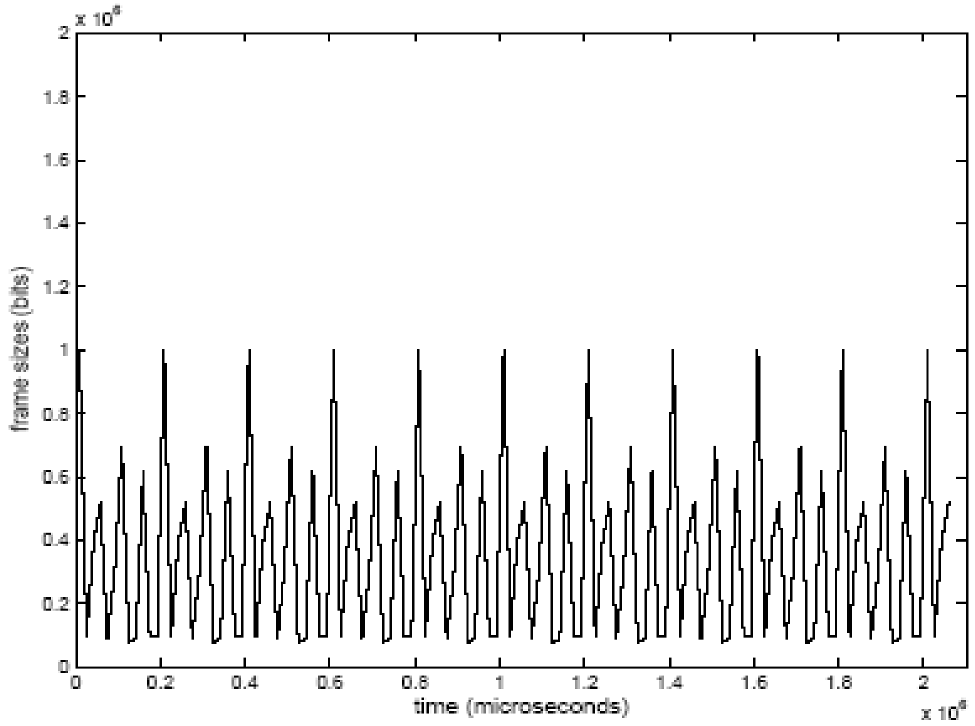

Figure 3.1: Frame Sizes of stream 8 generated in Node 5

- Class 2: The class 2 source is a variable bit rate source with no end-to-end time constraints. The packets are assumed to be generated from a storage area network (SAN), with the following packet-size distributions: 44% of 64Kbytes, 18% of 56K, 21% of 40K, 4% of 32K, 4% of 24K and 6% of 8Kbytes, see Trevitt [18].

- Class 3: The class 3 traffic arrival process is best effort traffic and it is modelled as in Xu et al. [25] by a modified Interrupted Poisson Process. The ON and OFF periods are exponentially distributed. Packets arrive back to back during the ON period at the rate of 2.5 Gbps. The last packet that arrives when the ON period ends is truncated. During the OFF period, no packets are generated. The mean packet size is 500 bytes and any packet size above 5000 bytes is truncated to 5000 bytes. To calculate the ON and OFF periods, we use the coefficient of variation c^2, defined as the ratio of the variance of the packet inter-arrival time divided by the squared mean of the packet inter-arrival time. c^2 indicates the burstiness of the arrival process.

$$c_{IPP}^2 = 1 + \frac{2\lambda\mu_1}{(\mu_1 + \mu_2)^2} \tag{3.1}$$

where $1/\lambda = (500\text{bytes})/(2.5\text{Gbps}) = 1.6\mu\text{s}$, and $1/\mu_1$ and $1/\mu_2$ are the mean times of the ON and OFF periods. The arrival process of class 2 traffic is completed by the following equation:

$$AverageArrivalRate = 2.5Gbps \cdot \frac{\mu_1}{\mu_1 + \mu_2} \tag{3.2}$$

3.2 Simulation Model and Network Setup

The following sections explain the structure of the simulation.

3.2.1 The OBS Network Under Study

We have used the nation-wide IP backbone network topology shown in figure 3.2 for our simulation study. This is a modified version of an IP backbone topology. The total number of nodes in the network are 24 with an average nodal degree of 3.583 and an average hop distance of 2.992. Each of the edge nodes in the network are assumed to be connected to a single user node.

Each of the user nodes is connected to the edge node through a single fiber link each capable of carrying n wavelengths. Each pair of adjacent nodes in the network are also connected by a single link with n wavelengths. There are 15 user nodes in total attached to each of the edge nodes in the network.

Each node in the network has knowledge about the entire OBS network including the number of hops required to reach each destination. The routing table at each node contains the next hop information for each destination and the number of hops to each. Typically, each node maintains the following information at all times of the simulation.

- Routing Table: This information is required to route the bursts to the appropriate output port. A list of primary, secondary and tertiary links are also stored in the

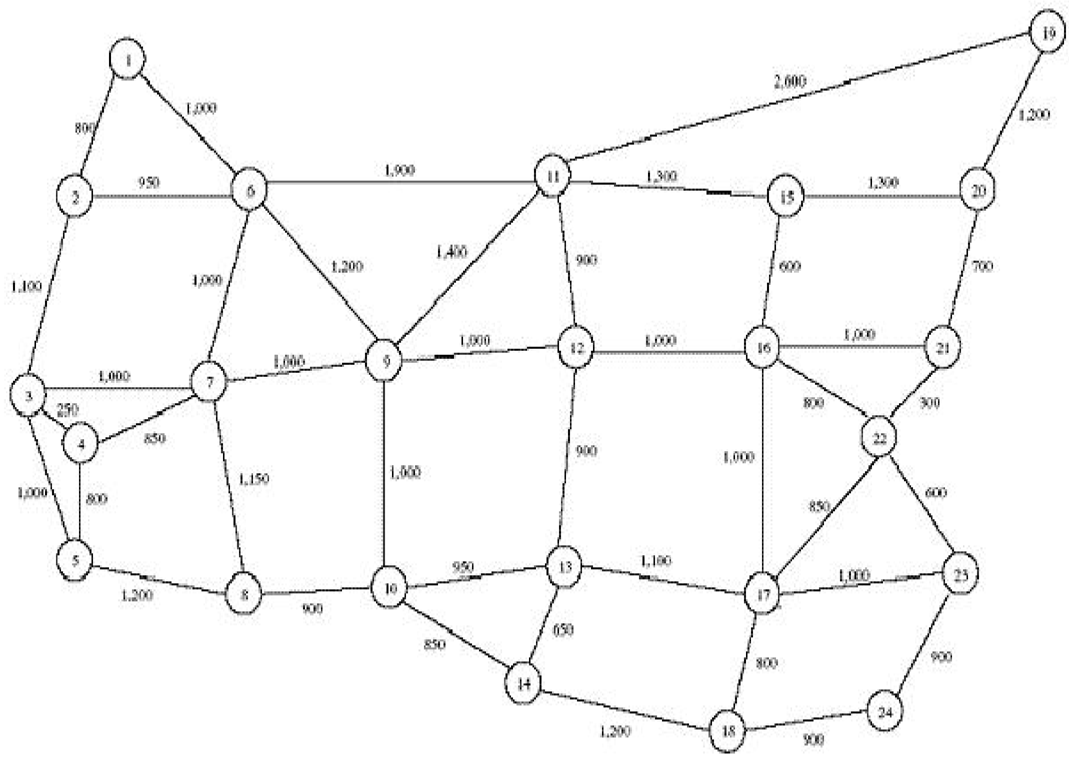

Figure 3.2: OBS Network Used for Study

routing table.

- Wavelength Array: This vector contains a list of the wavelengths that are currently occupied and the burst number that has reserved this wavelength. It also includes the time when this wavelength has to be made free.

- Converter Array: This information is only when converters are assumed to be present. This is used to keep a list of the busy and free converters at each node.

- Data Array: This is used to store the results for the simulations. The number of bursts arrived, number of bursts dropped and number of frames sent etc. are examples of the values stored at each node.

3.2.2 Simulation Model

A typical event based simulation [14] has been used for this thesis. Only the control plane is simulated which is sufficient for gathering results for the OBS network.

The following events are identified for the simulation.

- TIMER_EXPIRED: This event is used to identify the end of the expiration of the timer. Each node maintains a local timer which expires every 10 μs. When the timer expires, we check if the node has data to be sent that is greater than the minimum burst size. If yes, the data is packed into a burst and a control packet is sent to the destination. Otherwise, the timer is renewed and inserted in the event list.
- RECV_CTRL_FRAME: Receive control frame event is used to identify the receipt of a control packet at an OXC. When a control packet is received, the OXC checks to see of the wavelength on the outgoing port is free. If the wavelength is free, resources are reserved for this burst. Otherwise, the congestion control scheme is used to find a free wavelength on the same output port (wavelength conversion) or a free wavelength on a different output port (deflection). If no free wavelength is available, the burst is dropped.
- RELEASE_RESOURCES: This event is used to signal the end of the transmission of a burst on a link. The resources like wavelength and converters if any are freed in this event. This makes it possible for other bursts to reserve the resources later.
- RECEIVE_FEEDBACK: Whenever a burst is dropped, a feedback message is sent to the upstream node. The resources which have been reserved for the dropped burst are freed when a feedback message is received and the message is sent to the upstream node of the present node. The feedback message proceeds all the way to the sender of the burst.

The following input variables are used in the simulation.

No of Streams Each user node maintains a set of real-time HDTV streams. Each stream has a unique destination which does not change throughout the simulation. The destination of each stream is decided at the beginning of the simulation.

No of Wavelengths The number of wavelengths per fiber can be changed. But for most of the simulations, this number has been kept at a value of 4.

No of Converters and Conversion Range These values are used for the wavelength conversion scheme. The conversion range decides the conversion capability of each converter.

Burst Timer The burst timer value is used to identify the end of the timer event. A timer value of 10 μsec is used for the simulation.

Other Variables There are several other variables in the simulation. The number of batches the simulation ran, the number of bursts that each user node has to send in each batch, the minimum and maximum size of class 2 and class 3 bursts are some of them.

The simulation results are plotted with 95% confidence interval estimated by the method of batch means, see Perros [14]. The time unit for the simulation is μsec. There are thirty batches in the simulation and each batch is completed when all the nodes generate 10,000 bursts. The confidence intervals are very tight and they are not discernible in the graphs.

3.2.3 The Algorithms

The sections below explain some of the basic algorithms used in the simulations.

- Timer Expiration: Whenever a timer expires, the user node checks to see if a burst can be created. Bursts are created on a priority basis. The following summarizes how a control packet is created.
  - When the timer expires, check if a burst can be created from any of the burst priority queues, see figure 3.3, with real-time queue having the highest priority. Only one burst, if possible, is created for every 10 μsecs.
  - If a burst can be created, check if there is a free wavelength available at the edge node to send this burst
  - Then, check if the transmitter is not busy transmitting other bursts at this time
  - Create a control packet and send it to the edge OBS node
  - Create a timer expiration event and put it in the event queue

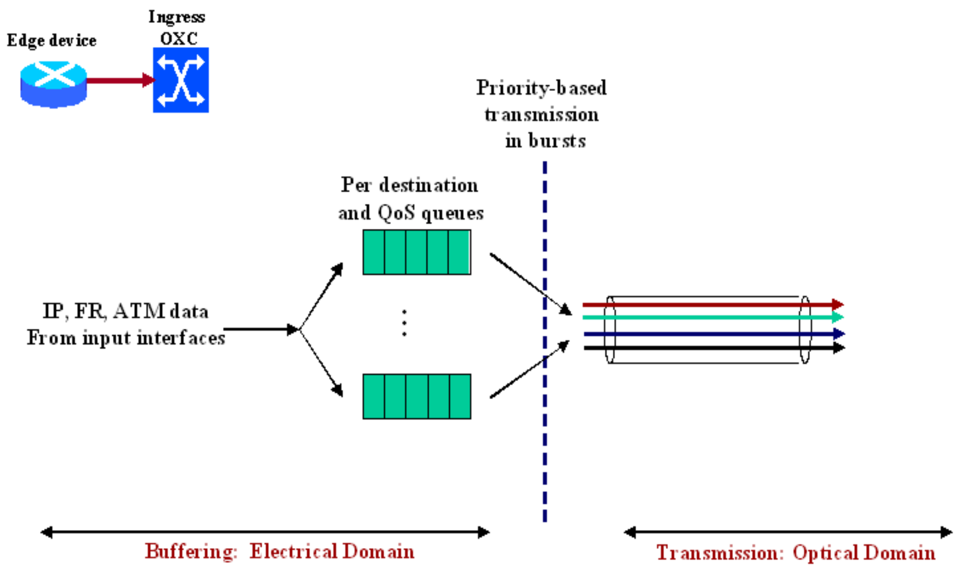

Figure 3.3: Buffering in an Edge Node

- If any of the above checks fails, simply create a timer expiration event and check at the next timer event

- Deflection Routing: The following algorithm illustrates the simulated deflection routing strategy.
  - Check if there is an output port which has a wavelength free and same as the incoming burst wavelength
  - If yes, check if there is free FDL to be allocated to this burst
  - If yes, allocate the FDL and the wavelength on this port to the incoming burst
  - Create a RECV_CTRL_FRAME event and send it to the next downstream node
  - if FDL is not free or there is no free wavelength on any other port, then the burst is dropped and a feedback message is sent to the upstream node

Temporary loops are not prohibited in this simulation but the *ttl (time to live)* field in the control packet takes care os avoiding permanent loops. Also, if a control packet arrives on the same port as it has previously arrived, the control packet is dropped thus arriving looping of control packets.

The deflection is at the link level which means that each burst is deflected to a different port independent of the choices made by other nodes. In path deflection, each node has to follow the pre-assigned deflected path to reach the destination.

- Wavelength Conversion: We have simulated both partial wavelength conversion and full wavelength conversion. Full wavelength conversion is easy because there are as many converters as there are wavelengths at each link. In partial wavelength conversion, there are fewer converters and a free wavelength has to be within the conversion capability of the converter. In both these schemes, it is mandatory for the burst to reserve a free converter. If there are no free converters, then the burst is dropped and a feedback message is sent back to the upstream node.

- Conversion and Deflection: In this scheme, real time bursts are given the highest priority and can be either converted or deflected to a different port or wavelength. Class 2 traffic can only be converted to a different wavelength while class 3 traffic is just dropped if the wavelength on the output port is already reserved for another burst.

Priorities are also employed at the destination node. If a node is busy receiving a lower class burst when a higher class burst arrives, then the current burst is preempted and the higher class burst is received. In this way, the real time frames are given higher priority than the class 2 and class 3 bursts. At the sender node, if real time data is waiting to be sent out from the queue when the timer expires, it is given a higher priority than the other traffic. This is because, real-time traffic has a strict end-to-end delay requirement and hence, has to be sent out from the queue as early as possible.

Chapter 4

Simulation Results

The simulation results are plotted with 95% confidence interval estimated by the method of batch means, see Perros [14]. Thirty one batches are simulated and the results from the first one are ignored due to the initial conditions. Each batch is completed when each node generates 10,000 bursts. The confidence intervals are very tight and they are not discernible in the graphs. The following sections present the graphs and a brief explanation on the results obtained.

4.1 Simulation Inputs and Outputs

At the beginning of the simulation, the destination node of each stream at each node is randomly selected. Specifically, the same number of HDTV streams originate at each node. Each stream contributes an average of 20 Mbps of the total traffic. For each class, the transmission queue in a node was assumed to have a buffer size of 1 MB. Each wavelength was assumed to have a bandwidth of 2.5 Gbps with 4 wavelengths per link. The average class 2 arrival rate is fixed at *0.5* and the average arrival rate of class 3 traffic to *0.5* for all the results obtained. The total average arrival rate is the sum of the average arrival rates of the 3 traffic classes.

The simulation results were obtained by varying the number of HDTV streams originating at each edge device while keeping the for class 2 and class 3 traffic fixed. In all the graphs, the x axis represents the number of HDTV streams originating at each edge device unless otherwise specified.

Given below is a list of the different schemes that we have simulated:

No Conversion and No Deflection No congestion control is used in this scheme. The bursts are simply dropped when the outgoing wavelength is busy. This is typically the base case scenario against which all the other results are compared. Curves marked "Drop" belong to this scheme.

Link Deflection In case of congestion, the burst is deflected to a different output port on the same wavelength in this scheme. Curves marked "Deflection" belong to this scheme.

Partial Conversion One wavelength converter is used per output fiber. Each of the converters has a conversion range of *2*. Curves marked "PartialConversion" belong here.

Full Conversion The number of converters is equal to the number of wavelengths on each output fiber. Curves marked "FullConversion" belong to this scheme.

Deflect and Convert In this scheme, an attempt is made to convert the incoming burst on to a different wavelength in case of congestion. If conversion fails due to lack of wavelengths on the same output fiber, the burst is deflected to a different output port. Curves marked "Deflect&Convert" belong to this scheme.

Priority Deflection and Conversion In this scheme, real-time traffic is both converted or deflected in case of congestion. Class 2 traffic is only converted and class 3 traffic is neither converted nor deflected. Curves marked "Priority" belong here.

4.2 Burst Timer

The burst timer is an important aspect in an OBS network. A too low burst timer results in smaller bursts being transmitted while a large burst timer increases the delay experienced by each burst which could effect real-time bursts. The following graph shows the optimum value that we have obtained for our simulations.

Figure 4.1 plots the percentage hit ratio versus the burst timer in μsecs. Hit ratio is defined as the total number of real-time bursts (i.e HDTV frames) received in time by all

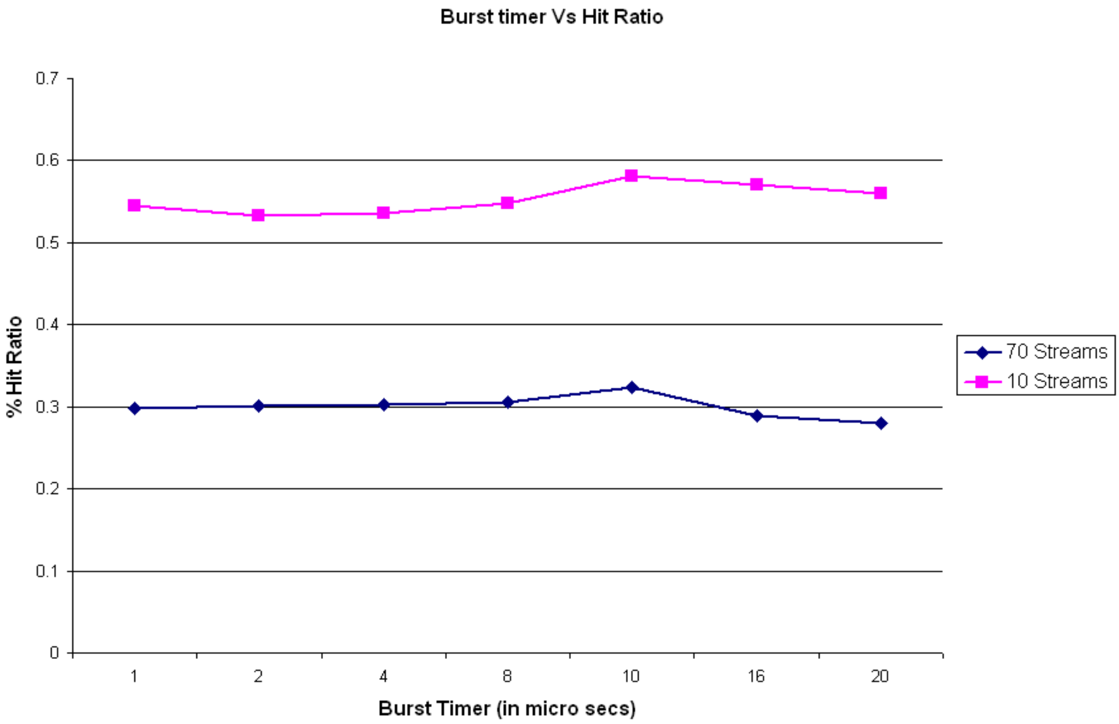

Figure 4.1: Burst Timer Vs Hit Ratio

the nodes, divided by the total number of frames sent by all the nodes. A frame is received in time if it arrives within 17 milliseconds of receiving the previous frame from the same stream. If a frame arrives after 17 milliseconds of receiving the last frame, then the frame is late and is a miss. A higher hit ratio for a larger number of streams per node is the desired performance criterion. From the figure, we can see that a burst timer value of 10 μsecs gives a higher hit ratio than all other values both for lower and higher streams per node. Hence, all our simulations have been run with a burst timer value of 10 μsecs to achieve optimum results.

4.3 Utilization

The % Bandwidth utilization is defined as the amount of time a link is busy transmitting bursts. The following graph (fig 4.2) plots the % bandwidth utilization for the link between node 1 and the edge OBS node. The % bandwidth utilization for class 2 and class 3 traffic is constant because of the constant class 2 and class 3 arrival rate.

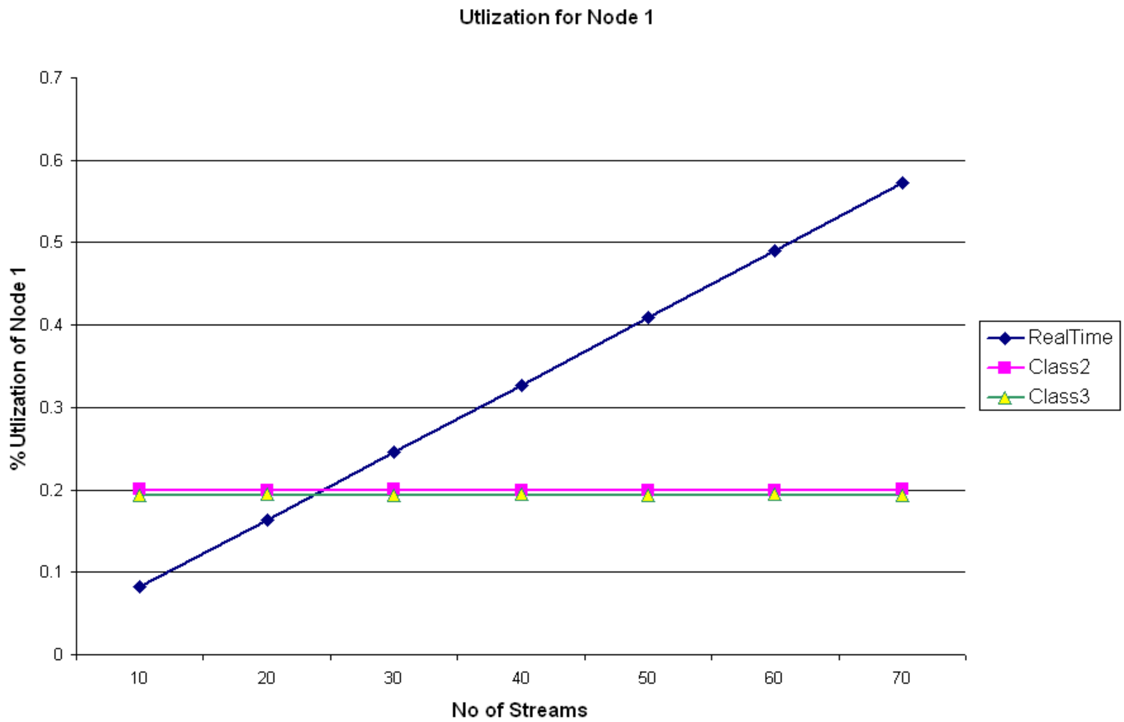

Figure 4.2: % Bandwidth Utilization for Node 1

The real-time traffic streams increase in each round of the simulations. Hence, the utilization of the link for real-time traffic also increases accordingly. For 10 streams, the total utilization of the link is approximately 50% which increases to 95% 70 streams. Each link carries 4 wavelengths in our simulations. Hence, on an average, the utilization of each wavelength is approximately 25%.

4.4 Burst Loss Percentage

Overall burst loss rate is defined as the total number of bursts (real-time, class 2 and class 3) lost because of collisions divided by the total number of bursts transmitted by all the nodes. Figure 4.3 plots the *overall burst loss rate* with varying number of HDTV streams per node.

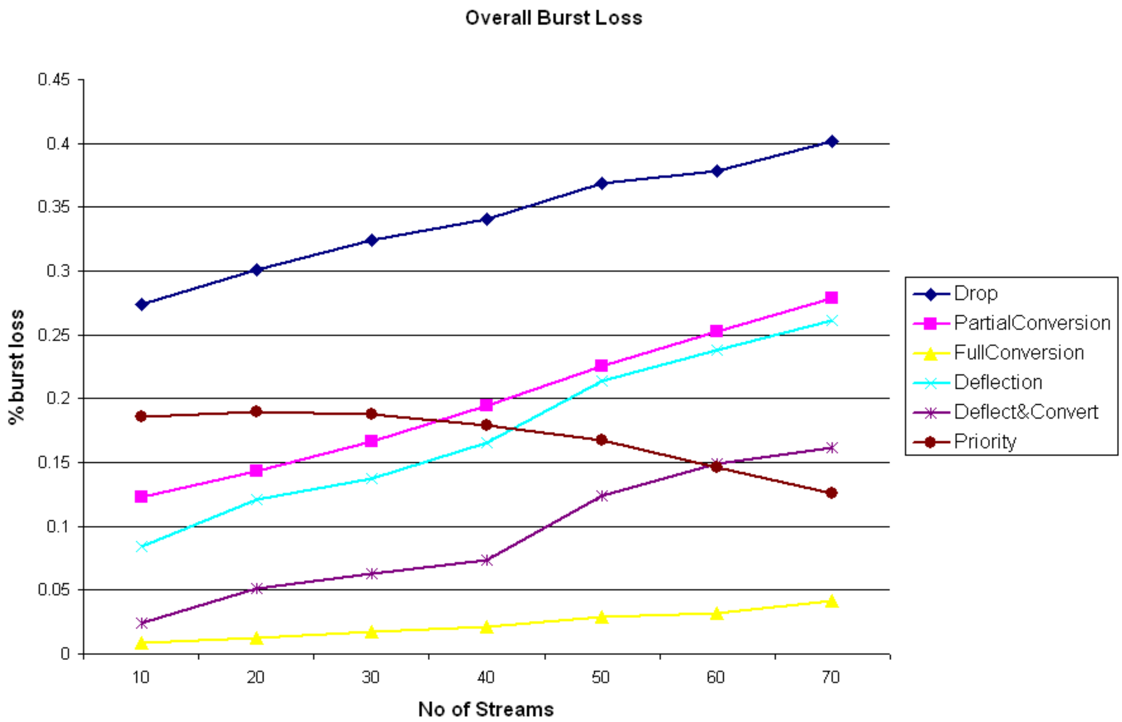

Figure 4.3: Overall Burst Loss Rate

Using no conversion or deflection gives rise to very high burst loss. This is evident from the graph, see figure 4.3, which shows that over 40% of the bursts are lost due to collisions when no congestion control schemes are used. On the other hand, use of conversion and/or deflection reduces the overall burst loss significantly. The *Full Conversion* scheme gives the best results because it utilizes all the wavelengths on a link. We can also notice that *Deflection* also achieves better results but the burst loss still remains at a little over 25% for 70 real-time streams. An important observation that can be made from the figure is the effect of *partial wavelength conversion*. One converter with a conversion range of 1 can achieve results very close to that of deflection routing. A combination of partial wavelength conversion and deflection routing has a lower overall burst loss rate because of the better utilization of links and wavelengths. The priority scheme shows heavy losses for low utilization due to dropping of all the class 3 bursts. As the traffic load increases, the ratio of class 3 bursts dropped to the over all burst drop decreases and hence the drop in the curve.

Figure 4.4 plots the *class 2 burst loss rate*. It can be observed that the results are similar to the *overall burst loss rate*. *Full conversion* is capable of achieving a very low burst loss rate. In fact, with 70 streams, this scheme only suffers around 6% burst loss. The other schemes also exhibit similar burst loss rates as in the case of overall burst loss rates. The priority scheme shows a lower class 2 burst loss due to lesser occupancy of wavelengths by class 3 traffic.

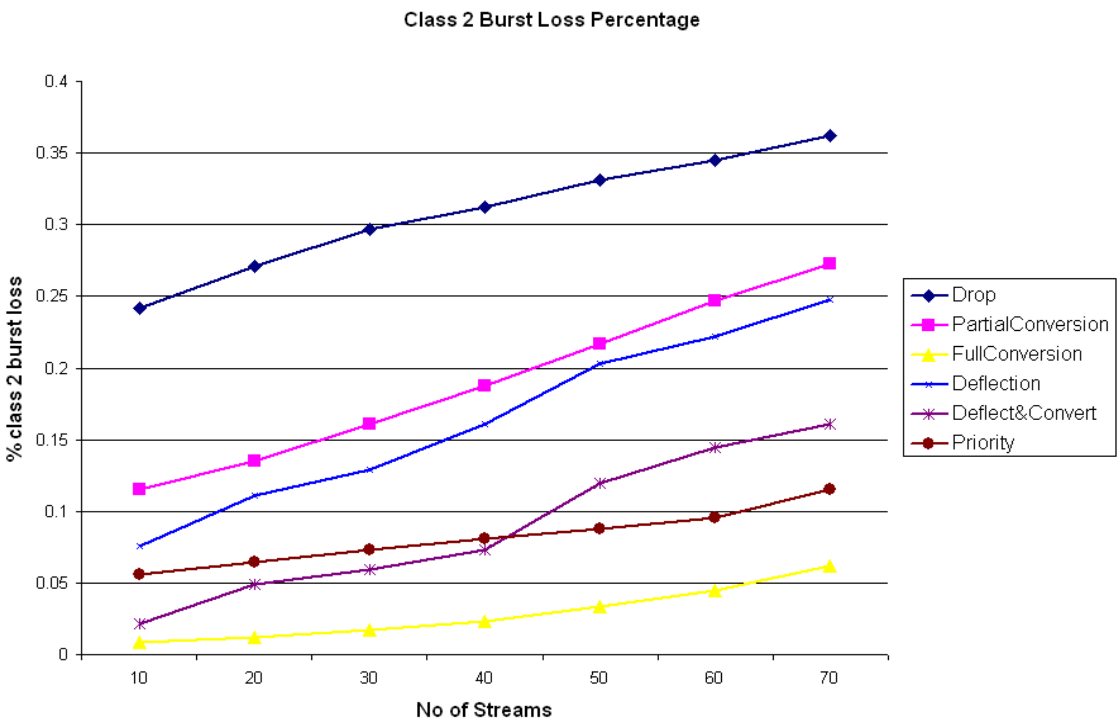

Figure 4.4: Class 2 Burst Loss Rate

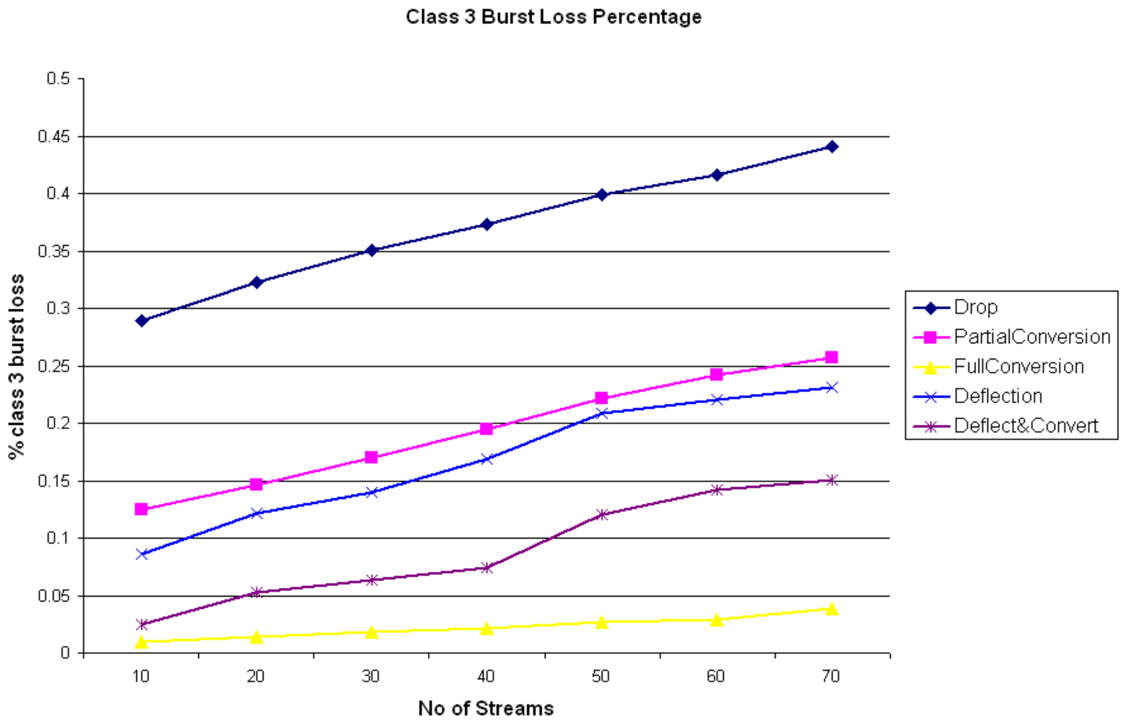

Figure 4.5: Class 3 Burst Loss Rate

Figure 4.5 plots the *class 3 burst loss rate*. The results are similar to *class 2 burst loss rate* because of the similarity in utilizations of the class 2 and class 3 traffic,refer fig 4.2.

4.5 Real-Time Traffic Analysis

Figure 4.6 plots the *hit ratio* for all the congestion control schemes. This is an important performance metric that describes how well OBS can support real-time traffic. Hit Ratio is defined as the total number of real-time bursts (each burst contains a single HDTV frame) received in time by all the nodes, divided by the total number of bursts sent by all the nodes. As mentioned earlier in the chapter, a frame is received in time if it arrives within 17 milliseconds of receiving the previous frame from the same stream. The

Full Conversion scheme has a higher hit ratio than any other congestion control mechanism. The results obtained are close to 95% hit ratio. All the other protocols have a sudden fall in the hit ratio after a particular number of streams per node. The priority scheme shows the highest hit ratio close to 99.5% because of the highest priority given to real-time bursts. Since class 3 bursts are dropped during congestion, this bandwidth is utilized by real-time bursts to achieve high hit ratio.

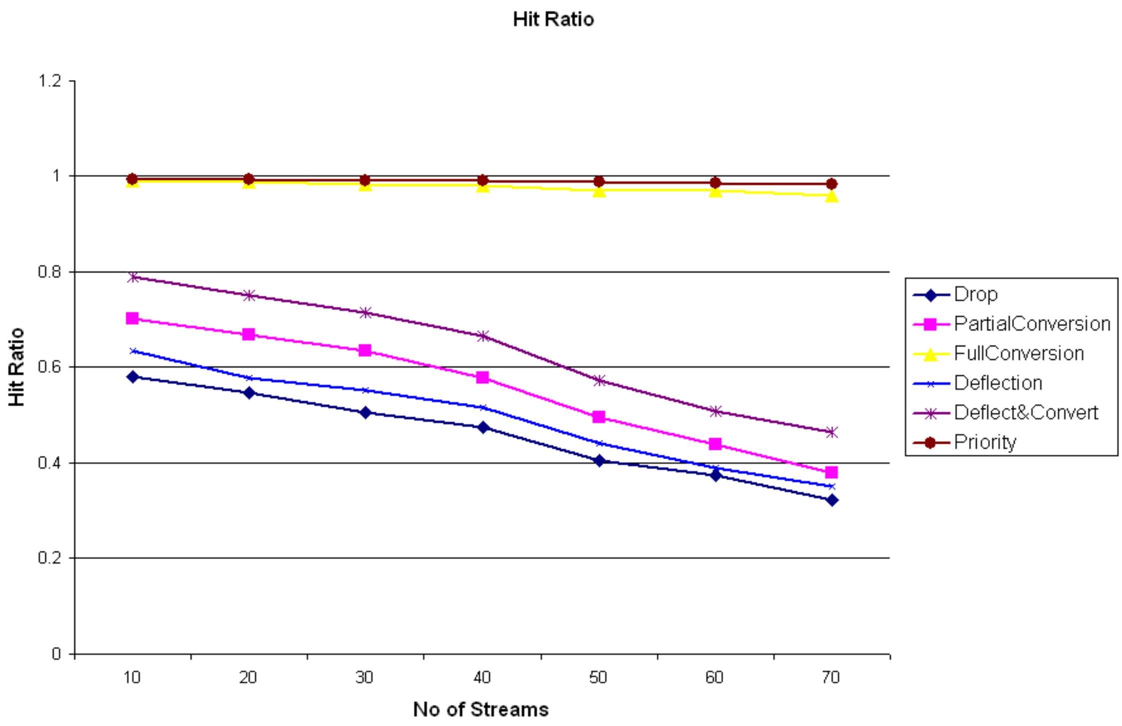

Figure 4.6: Hit Ratio for Real-Time Traffic

Deflection routing performs more than the other schemes. The hit ratio drops down to below 40% for 70 streams per node or when the utilization of each wavelength is close to 25%. All the schemes achieve a good hit ratio for low utilization in the network. The hit ratio ranges between 60% and 80% for all of the schemes.

Another performance metric that can give a good idea about the number of streams each node can support is the *mean frame delay*. The mean frame delay is defined as the sum of the end-to-end delay experienced by all the frames received, divided by the number of

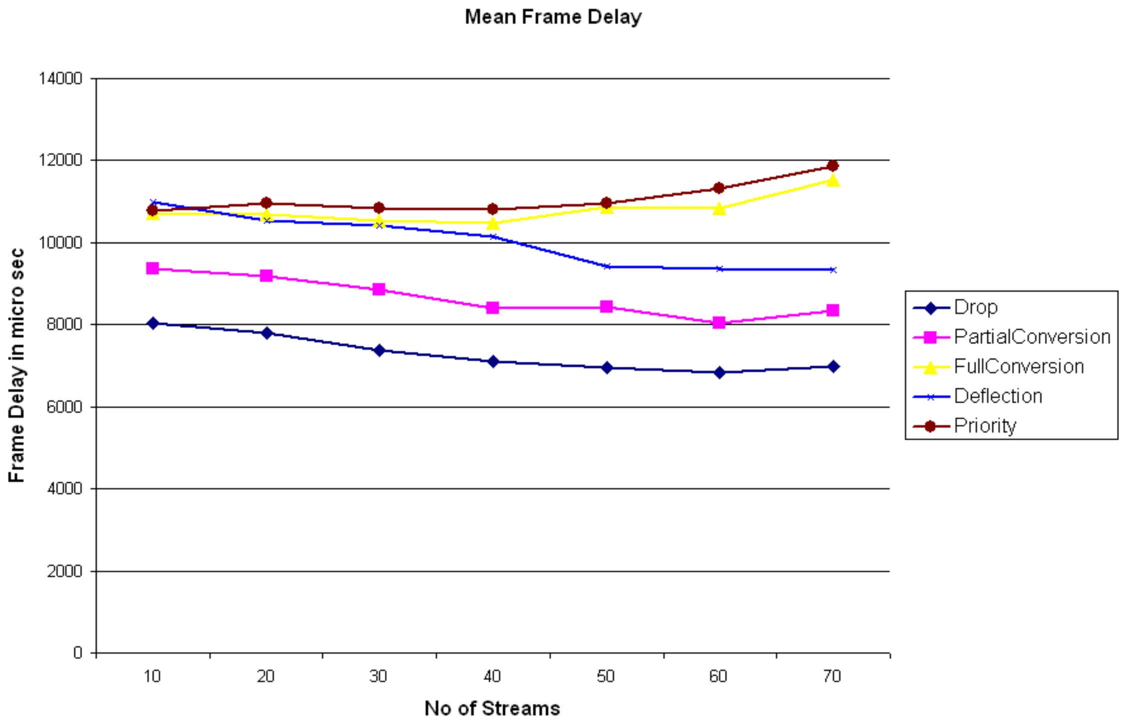

Figure 4.7: Mean Frame Delay

frames received by all the nodes. The end-to-end delay comprises of the queueing delay that a frame experiences at the transmitting node from the moment it arrives to the moment it is transmitted out, and the propagation delay. Figure 4.7 plots the mean frame delay versus the real-time traffic. The partial wavelength conversion scheme shows a lower mean frame delay than the other protocols. This implies that it can support more streams per node than any other protocol. Deflection schemes have much higher delays to the inherent delay involved in following the extra deflected links for the bursts. This is evident from the deflection and the priority schemes. The priority scheme has the highest mean frame delay because os the high occupancy of all wavelengths and links in the network.

4.6 Concluding Remarks

We studied the congestion control problem in OBS networks and analyzed the performance of deflection routing and wavelength conversion. The following conclusions can be drawn from the simulation results obtained.

- In order to have about 1% burst drop rate, the utilization per wavelength (at the source) should be kept to as low as 12% and full wavelength conversion is required. This means that the network resources are heavily under-utilized. Increasing the utilization increases the burst drop rate significantly.
- In order to have about 5% burst drop rate, the wavelength utilization (at the source) should be kept at 25% with full wavelength conversion.
- Full Conversion scheme performs better than deflection routing. In fact, partial wavelength conversion with very few converters has burst drop rates similar to deflection routing. Deflection routing scheme induces higher drop rates when the load in the network increases significantly. Wavelength conversion schemes perform better under heavy network load.
- Much of the focus in the study is on real-time traffic. From the graphs shown in the above sections, we can see that OBS networks are capable of carrying real-time traffic. Wavelength conversion schemes achieve a hit ratio close to 95% when the utilization of each wavelength is about 25%.
- The priority scheme can be used to achieve low drop rates for real-time bursts. Hit ratio close to 99.5% is possible with this scheme.

Overall, the simulations results indicate that wavelength conversion schemes perform better than deflection routing. Partial wavelength conversion with just one converter at each output port can achieve performance close to that of deflection routing. The full wavelength conversion scheme has been shown to perform the best over all other schemes. The availability of wavelength converters in the future can enhance the OBS network performance to a great extent.

Although OBS networks are capable of achieving low burst drop rates, the performance degrades heavily when the network load increases beyond 40%. Hence, OBS will not be used in this IP-network like form, i.e., transmit when you want. It requires a low utilization in order to maintain an acceptable burst loss. Present congestion control schemes (deflection routing, conversion, rate control) cannot completely eliminate burst loss.

Chapter 5

Summary and Future Research

In this thesis, we analyze and compare the performance of two congestion control schemes in an OBS network. First, we present a detailed literature review for deflection routing congestion control scheme and analyze its performance in an OBS network which carries real-time and non-real time traffic. The second scheme is wavelength conversion which comprises of both partial and full wavelength conversion. From the results obtained from simulation, we find that, under reasonable assumptions regarding the network parameters, wavelength conversion is capable of achieving low burst drop rate and is also capable of achieving a very good hit ratio.

5.1 Future Work

Our work can be extended in the following direction:

- Feedback based congestion control schemes can be analyzed. Feedback messages can be used to either achieve path deflection or a complete end-to-end path restoration in case of congestion. We believe we can have a better performance using path deflection than the simple link deflection currently used in the simulations. Complete end-to-end path restoration can be used as an extreme measure to route bursts through non-congested links in the network. Labels can be used to implement path deflection and path restoration.

- This thesis only analyzes a few reactive congestion control schemes. In future, we plan to implement proactive congestion control schemes like bandwidth allocation at setup time in our simulations and analyze their performance.

- Increasing the number of wavelengths on each link can reduce congestion in the network. One could analyze the effect of the number of wavelengths on the burst loss rate in an OBS network.

Bibliography

[1] The jumpstart project. Available from URL `http://jumpstart.anr.mcnc.org`.

[2] Obs publications list. Available from URL `http://www.utdallas.edu/~vinod/obs.html`.

[3] I. Baldine, G. N. Rouskas, H. G. Perros, and D. Stevenson. Jumpstart: A just-in-time signaling architecture for wdm burst-swithced networks. *IEEE Communications*, 40(2):82–89, February 2002.

[4] T. Battestilli and H. G. Perros. An introduction to optical burst switching. *IEEE communications optical magazine*, 41:510–515, 2003.

[5] N. Bouabdallah, A. L. Beylot, and G. Pujolle. Collision avoidance and fairness issues in metropolitan bus-based optical access networks. *Technical Report, Paris University*, 2003.

[6] C. Cameron, A. Zalesky, and M. Zukerman. Shortest path prioritized random deflection routing in optical burst switched networks. In *3rd International Workshop on Optical Burst Switching (WOBS)*, October 2004.

[7] H. M. Chaskar, S. Verma, and R. Ravikanth. Optical burst switching: A viable solution for terabit ip backbone. *IEEE Network*, pages 48–53, November/December.

[8] H. M. Chaskar, S. Verma, and R. Ravikanth. A framework to support ip over wdm using optical burst switching. In *IEEE/ACM/SPIE Optical Network Workshop*, January 2000.

[9] K. Dolzer and C. Guager. On burst assembly in optical burst switching networks – a performance evaluation of just-enough-time. In *Proceedings of 17th International Teletraffic Congress*, pages 149–161, September 2001.

[10] A. Fumagalli and P. Krishnamoorthy. A low-latency and bandwidth-efficient distributed optical burst switching architecture for metro ring. In *Proceedings of IEEE ICC*, volume 2, pages 1340–1344, 2003.

[11] C. Hsu, T. Liu, , and N. Huang. Performance analysis of deflection routing in optical burst-switched networks. In *Proceedings of IEEE, INFOCOM 2002, New York, NY*, June 2002.

[12] J. Jong. Multicast access protocols in an optical burst switched wdm ring network. Master's thesis, NC State University, 2002.

[13] S. Kim, N. Kim, and M. Kang. Contention resolution for optical burst switching networks using alternative routing. In *IEEE ICC 2002*, April 2002.

[14] H. G. Perros. Computer simulation techniques: The definitive introduction. Available from free download from URL `http://www.csc.ncsu.edu/faculty/perros/simulation.pdf`.

[15] V. S. Puttasubbappa and H. G. Perros. Access protocols to support different service classes in an optical burst switching ring. In *Networking 2004, Athens*, May 2004.

[16] C. Qiao and M. Yoo. Optical burst switching (obs) - a new paradigm for an optical internet. *Journal of High Speed Networks*, 8(1):69–84, January 1999.

[17] G. P. V. Thodime, V. M. Vokkarane, and J. P. Jue. Dynamic congestion-based load balanced routing in optical burst-switched networks. In *GLOBECOM 2003 – IEEE Global Telecommunications Conference*, volume 22, pages 2628–2632, December 2003.

[18] S. Trevitt. Traffic patterns in fibre channel, May 2002. Available from URL `http://www.mcdata.com`.

[19] J. S. Turner. Terabit burst switching. *Journal of High Speed Networks*, 8(1):3–16, January 1999.

[20] V. Vokkarane, J. Jue, and S. Sitaraman. Burst segmentation: an approach for reducing packet loss in optical burst switched networks. In *Proceedings IEEE, International Conference on Conference (ICC) 2002, New York, NY,*, April 2002.

[21] X. Wang, H. Morikawa, and T. Aoyama. A deflection routing protocol for optical bursts in wdm networks. In *Proceedings of the Fifth Optoelectronics and Communications Conference (OECC 2000), Makuhari, Japan*, pages 94–95, 2000.

[22] J. Y. Wei and R. I. McFarland. Just-in-time signaling for wdm optical burst switching networks. *journal of Lightwave Technology*, 18(12):2019–2037, December 2000.

[23] I. Widjaja. Performance analysis of burst admission control protocols. *IEEE Proceedings of Communications*, 142:7–14, February.

[24] L. Xu, H. G. Perros, and G. Rouskas. A simulation study of optical burst switching and access protocols for ring networks. In *Proceedings of Networking, Pisa, Italy*, May 2002.

[25] L. Xu, H. G. Perros, and G. N. Rouskas. A simulation study of optical burst switching access protocols for wdm ring networks. *Journal of Computer Networks*, 41(2):143–160, January 2003.

[26] M. Yoo, C. Qiao, and S. Dixit. A comparative study of contention resolution policies in optical burst switched wdm networks. In *SPIE Proc. of Conference on Terabit Optical Networking Architecture, Control and Management issues, Boston, MA*, volume 4213, pages 124–135, November 2000.

[27] M. Yoo, C. Qiao, and S. Dixit. Qos performance of optical burst switching in ip-over-wdm networks. *Journal on Selected Areas in Communications*, 18(10):2062–2071, 2000.