

2005

New contention resolution techniques for optical burst switching

Kishore Reddy Koduru

Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses



Part of the [Electrical and Computer Engineering Commons](#)

Recommended Citation

Koduru, Kishore Reddy, "New contention resolution techniques for optical burst switching" (2005). *LSU Master's Theses*. 580.

https://digitalcommons.lsu.edu/gradschool_theses/580

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

**NEW CONTENTION RESOLUTION
TECHNIQUES
FOR
OPTICAL BURST SWITCHING**

A Thesis

**Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering**

in

The Department of Electrical Engineering

**By
Kishore Koduru
Bachelor of Engineering (Computer Engineering), BIET, 2002
May 2005**

Dedicated to my dear parents

Acknowledgements

I would like to thank my advisor Dr. Ahmed El. Amawy for the ideas that led to this work, for his timely comments, guidance, support and patience throughout the course of this work. I thank Dr. Jerry Trahan and Dr. Hsiao-Chun Wu for being on my defense committee. I thank all the members of the Optical Networks research group for assisting me with their ideas during Friday-meetings. I thank Mr. Moc Elgoog for his extended support during my entire stay at LSU. Last but not the least, I thank all my friends who have remotely helped in the successful completion of this work.

Table of Contents

ACKNOWLEDGEMENTS.....	iii
LIST OF FIGURES.....	vi
ABSTRACT.....	viii
CHAPTER 1: INTRODUCTION.....	1
1.1 Optical Transmission System.....	3
1.2 Optical Fiber.....	4
1.3 Wavelength Division Multiplexing (WDM).....	6
1.4 Components of WDM Optical Networks.....	7
1.4.1 Wavelength Multiplexer.....	7
1.4.2 Wavelength Add/Drop Multiplexer.....	8
1.4.3 Optical Crossconnects (OXC).....	9
1.4.4 Optical Amplifier.....	11
1.5 WDM Network Architectures.....	12
1.5.1 Broadcast-and-Select Networks.....	12
1.5.2 Wavelength Routed Networks.....	14
1.6 Routing and Wavelength Assignment.....	15
1.6.1 Route Selection.....	16
1.6.2 Wavelength Selection.....	17
1.7 Problem Formulation and Layout of Thesis.....	18
CHAPTER 2: OPTICAL BURST SWITCHING (OBS).....	20
2.1 OBS Network Architecture.....	22
2.2 Reservation Schemes in OBS.....	24
2.2.1 Tell-and-Go (TAG).....	25
2.2.2 Just-In-Time (JIT).....	26
2.2.3 Just-Enough-Time.....	27
2.3 Contention Resolution.....	29
2.3.1 Optical Buffering.....	30
2.3.2 Wavelength Conversion.....	32
2.3.3 Deflection Routing.....	34
CHAPTER 3: NEW DEFLECTION-BASED CONTENTION RESOLUTION SCHEMES.....	37
3.1 Backtrack on Deflection Failure.....	38
3.1.1 Routing Protocol.....	38
3.1.2 Backtrack on Deflection Failure with Increase in Initial Offset.....	40
3.1.2.1 Routing Procedure.....	40
3.1.3 Backtrack on Deflection Failure with Open Loop Reservation.....	42
3.1.3.1 Routing Procedure.....	43

3.1.4 Example for Backtrack on Deflection Failure.....	44
3.2 Bidirectional Reservation on Burst Drop for Retransmission Bursts.....	47
3.2.1 Offset Calculation.....	49
CHAPTER 4: SIMULATION & RESULTS.....	52
4.1 Simulator Setup.....	52
4.2 Assumptions.....	53
4.3 Simulation Parameters.....	55
4.4 Results.....	55
CHAPTER 5: CONCLUSION AND FUTURE SCOPE.....	72
BIBLIOGRAPHY.....	74
VITA.....	76

List of Figures

1.1 Optical transmission system.....	3
1.2 Fiber optic cable.....	5
1.3 Different modes of propagation.....	5
1.4 Wavelength division multiplexing.....	6
1.5 Diffraction grating for separation of wavelengths.....	8
1.6 Wavelength add/drop multiplexer.....	9
1.7 Wavelength cross-connect.....	10
1.8 Erbium-doped amplifier.....	12
1.9 Broadcast-and-Select network.....	13
1.10 Wavelength routed network.....	14
2.1 Comparison of optical switching schemes.....	21
2.2 OBS network architecture.....	23
2.3 Use of delayed reservation.....	24
2.4 JIT scheme.....	27
2.5 JET scheme.....	28
2.6 Contention resolution using FDL.....	32
2.7 Wavelength conversion.....	34
2.8 Deflection routing.....	35
3.1 Information fields in a control burst.....	39
3.2 Backtrack on deflection failure.....	45
3.3 Decision table for time period (0-4).....	45

3.4 Decision table for case 1	46
3.5 Decision table for case 2	47
3.6 Control burst drop due to reservation failure.....	49
3.7 Bidirectional reservation on control burst drop.....	50
4.1 NSFNET with 14 nodes.....	54
4.2 USA long haul network with 28 nodes.....	54
4.3 Packet loss probability at low loads.....	56
4.4 Packet loss probability at high loads.....	57
4.5 Packet loss probability versus low load (BDF-BR)	58
4.6 Packet loss probability versus high load (BDF-BR)	59
4.7 Blocking probability versus fixed offset.....	60
4.8 Average number of hops traveled versus intended number of hops.....	62
4.9 Average number of hops traveled versus load.....	63
4.10 Poisson traffic versus bursty traffic (low loads).....	64
4.11 Poisson traffic versus bursty traffic (high loads)	65
4.12 Link utilization versus load.....	66
4.13 Packet loss probability with different number of wavelengths per fiber.....	67
4.14 Throughput versus offered load.....	68
4.15 Throughput versus offered load (single retransmission).....	69
4.16 Average transmission count versus offered load.....	70
4.17 Packet loss probability versus load (USA long haul network).....	71

Abstract

Optical burst switching (OBS) is a technology positioned between wavelength routing and optical packet switching that does not require optical buffering or packet-level parsing, and it is more efficient than circuit switching when the sustained traffic volume does not consume a full wavelength. However, several critical issues still need to be solved such as contention resolution without optical buffering which is a key determinant of packet-loss with a significant impact on network performance.

Deflection routing is an approach for resolving contention by routing a contending packet to an output port other than the intended output port. In OBS networks, when contention between two bursts cannot be resolved through deflection routing, one of the bursts will be dropped. However, this scheme doesn't take advantage of all the available resources in resolving contentions. Due to this, the performance of existing deflection routing scheme is not satisfactory. In this thesis, we propose and evaluate three new strategies which aim at resolving contention.

We propose a new approach called *Backtrack on Deflection Failure*, which provides a second chance to blocked bursts when deflection failure occurs. The bursts in this scheme, when blocked, will get an opportunity to backtrack to the previous node and may get routed through any deflection route available at the previous node. Two variants are proposed for handling the backtracking delay involved in this scheme namely: (a) Increase in Initial Offset and (b) Open-Loop Reservation. Furthermore, we propose a third scheme called *Bidirectional Reservation on Burst Drop* in which bandwidth reservation is made in both the forward and the backward directions simultaneously. This scheme comes into effect only when

control bursts get dropped due to bandwidth unavailability. The retransmitted control bursts will have larger offset value and because of this, they will have lower blocking probability than the original bursts.

The performance of our schemes and of those proposed in the literature is studied through simulation. The parameters considered in evaluating these schemes are blocking probability, average throughput, and overall link utilization. The results obtained show that our schemes perform significantly better than their standard counterparts.

Chapter 1

Introduction

In recent years, explosive demand for network bandwidth has become a major challenge for network engineers due to increasing global popularity of the Internet and the increased applications it affords. A continuous demand for networks of high capacities at low cost exists. Optical data communication has been acknowledged as the best solution for meeting the present bandwidth requirements of the users and for supporting future network services. This is because; in theory each optical fiber has the ability to support bandwidth demand of up to 50 THz [1]. Apart from this, optical fibers are inexpensive and provide extremely low bit-error rates (typically 10^{-12}) [2]. The optical fiber is less bulky than other cables. Optical signals travel clearly for longer distances and are immune to electrical interferences. Furthermore, fiber cables are much more difficult to tap than copper wires, so in addition there is a security advantage [2]. All these factors make optical data networks the networks of the future.

Optical Networks may be classified as:

1. First Generation Optical Networks: These optical networks involved replacing copper cables by optical fibers as the medium of transmission. The switching and processing of bits were, however, handled in the electronic domain as before. Optical fibers were preferred for bit rates greater than 10 Mbps. Examples of first generation optical networks are SONET/SDH networks that form the core telecom networks in North America, Europe and Asia [3, 5]. Other examples include the FDDI-based enterprise networks. From a network layering point of view, the impact of first generation optical networks was felt primarily in the *physical layer*.

From hereon, there were primarily two fundamental ways of increasing the speed in networks; either increase the electronic processing speeds by improved time division multiplexing (TDM) techniques or increase the capacity by using multiple carrier wavelengths in the fiber at the same time [7].

2. Second Generation Optical Networks: These networks were made capable of using multiple carrier wavelengths that were multiplexed onto a single fiber thus offering increased bandwidth. The technique is called *Wavelength Division Multiplexing (WDM)* [3, 4, 10]. The primary improvement of second generation optical networks over their first generation counterparts from a technological point of view was in incorporating the switching and routing functionality in the optical domain and allowing for the transparency of data format, protocol and bit rates. It thus allowed for smaller electronic load on a node by ensuring the need to terminate the traffic intended only for that node while allowing the other traffic to cut right through the node in the optical domain. In first generation networks, a node would have to terminate all the optical signals (irrespective of whether they are intended for itself or not), convert them to electronic signals, process them and then regenerate the traffic not intended for itself into optical signals and send them on the appropriate outgoing links. The second generation optical switches are called *Optical Cross-connects (OXC)*s. These switches may be configured to switch optical signals from any incoming port to any outgoing port.

The next-generation optical networks will involve optical packet switching and All-Optical Networks (AON). In an AON, all network-to-network interfaces are based on optical transmission, and all user-to-network interfaces use optical transmission on the

network side of the interface. All buffering, switching and routing within AON network nodes are performed optically. In these networks, it is envisioned that the DWDM based dynamic optical network elements such as optical cross-connects and optical add/drop multiplexers will have full control of all wavelengths [5]. In addition, they are expected to have full knowledge of the traffic carrying capacity and the status of each wavelength. With such intelligence, these networks are envisioned as being self-connecting and self-regulating.

1.1 Optical Transmission System

Today's low-loss glass fiber optic cable offers almost unique advantages over all previously developed transmission media. The basic point-to-point fiber optic transmission system consists of three basic elements: the optical transmitter, the fiber optic cable and the optical receiver as shown in Figure 1.1.

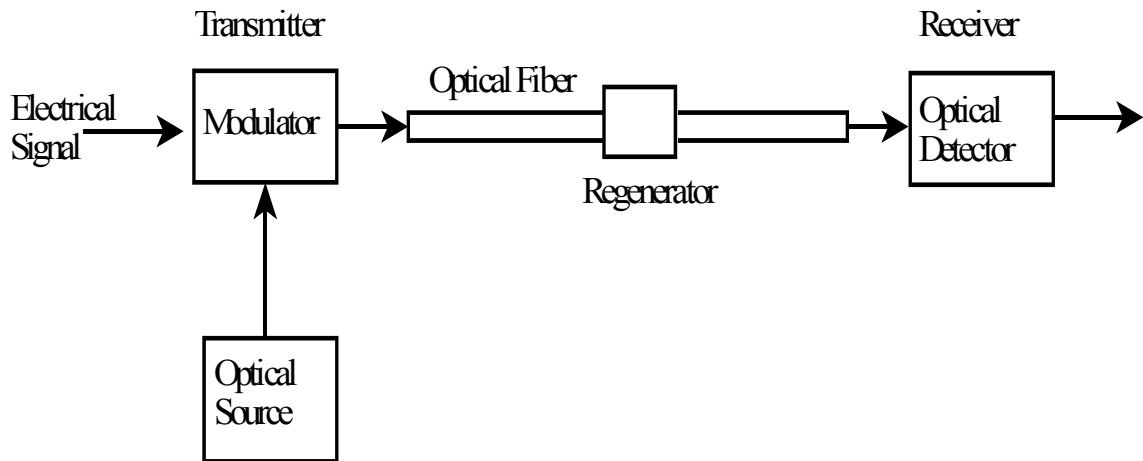


Figure 1.1 Optical transmission system

The transmitter converts an electrical analog or digital signal into a corresponding optical signal. The source of the optical signal can be either a light emitting diode, or a solid state laser diode. The light source can be modulated according to an electrical input signal to produce a beam of light which is transmitted into the transmission medium [2]. The optical fiber is the transmission medium. When the optical information reaches the receiver, the on/off light signals are converted back to electrical signals by an optical detector. In this system, the information undergoes electronic-optical-electronic conversion. The transmission characteristics of an optical fiber are usually given in terms of attenuation for a given wavelength over a given distance (length of the fiber). As the distance traveled by the signal increases, the attenuation also increases. When the signal becomes weak, the information carried cannot be retrieved from the signal. In order to prevent excessive attenuation, regenerators are used to boost the signal power and to restore the shape of the signal.

1.2 Optical Fiber

The main purpose of an optical fiber is to guide light waves with minimum attenuation (loss of signal). Optical fibers are composed of fine threads of glass in layers, called the core and cladding that can transmit light at about two-thirds the speed of light in vacuum. Though admittedly an oversimplification, the transmission of light in optical fiber is commonly explained using the principle of total internal reflection [6]. With this phenomenon, 100 percent of light that strikes a surface is reflected. Light is either reflected or refracted depending on the angle of incidence (the angle at which light strikes the interface between an optically denser and an optically thinner material).

The core has a higher refractive index than the cladding, allowing the beam that strikes that surface at less than the critical angle to be reflected.

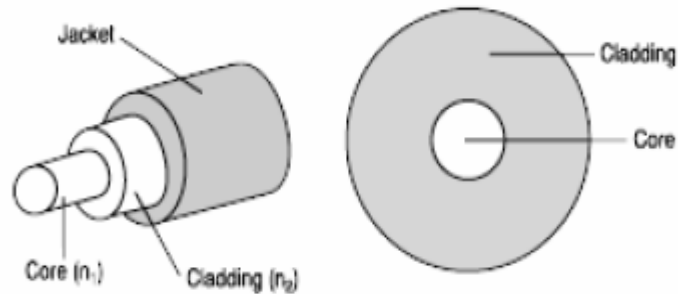


Figure 1.2 Fiber optic cable

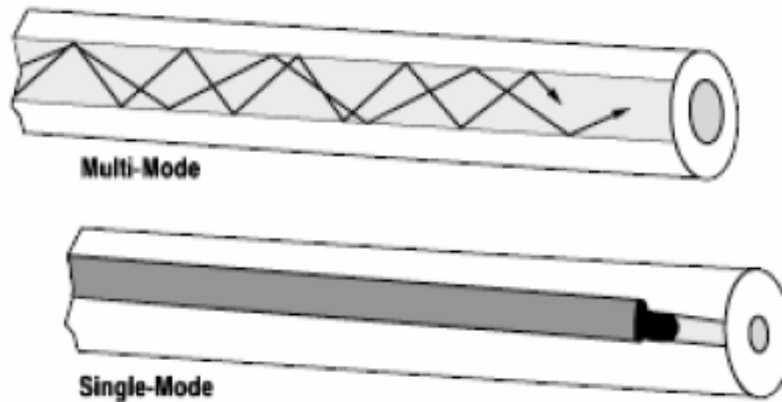


Figure 1.3 Different modes of propagation

There are two modes of fiber propagation known as multi-mode and single-mode. The single-mode fiber optic cable provides better performance but at a higher cost. The multimode fiber has a graded refractive index profile, due to which many rays of light can bounce at different angles [7]. Each ray is said to have a different mode, hence, the name multimode fiber. If a stepwise refractive index is used, the fiber will act like a waveguide and the light will travel in a straight line along the center axis of the fiber.

Such fibers are known as single mode fibers. A single mode fiber has lower attenuation and less time dispersion. However it is more expensive than the multimode fiber. These fibers are used mainly in Wide Area Networks [7].

1.3 Wavelength Division Multiplexing

In order to fully exploit the offered bandwidth of a fiber, the bandwidth is divided into a number of channels on different wavelengths. This method of sending many light beams of different wavelengths simultaneously on the same fiber is referred to as “Wavelength division multiplexing” (WDM) [3]. This method exploits the huge opto-electronic bandwidth mismatch by requiring that each end-user’s equipment operate only at electronic rate. But multiple WDM channels from different end-users may be multiplexed on the same fiber.

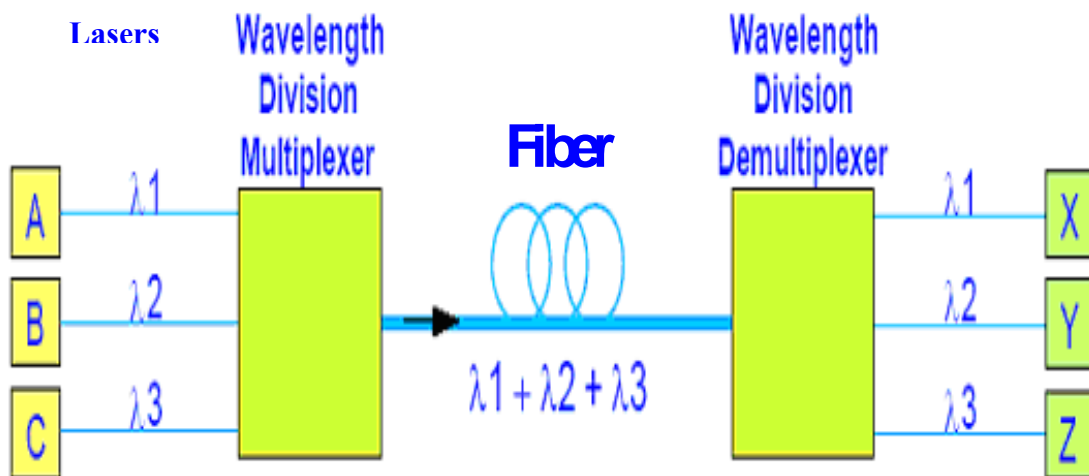


Figure 1.4 Wave division multiplexing

In a simple WDM system shown in Figure 1.4, the transmitting side has a series of fixed-wavelength or tunable light sources, each of which emits signals at a

unique wavelength. A multiplexer is used to combine these optical signals into a continuous spectrum of signals and to couple them onto a single fiber. Within the optical link, there will be various types of optical amplifiers. At the receiving end, a demultiplexer is used to separate the optical signals into appropriate detection channels for signal processing. The WDM systems are classified into dense wavelength division multiplexing (DWDM) systems and coarse wavelength division multiplexing (CWDM) systems. In DWDM, the bandwidth of the fiber is divided into more than 8 wavelengths. CWDM refers to the systems where the fiber bandwidth is divided into less than 8 wavelengths.

1.4 Components of a WDM Optical Network

Some of the major modules contained in a WDM optical network include wavelength multiplexers, optical crossconnects, optical amplifiers, and wavelength add/drop multiplexers. The following subsections discuss these components and their functions in a WDM network.

1.4.1 Wavelength Multiplexers

The function of this device is to combine independent signal streams operating at different wavelengths onto the same fiber and to separate them at the receiver. In principle, any demultiplexer also can be used as a multiplexer [8]. For simplicity, the word “multiplexer” is used as a general term that refers to both the combining and separating functions. The technologies used in these devices are include thin-film filters, arrayed waveguide gratings, Bragg fiber gratings, diffraction gratings, and interleavers. Among these, diffraction grating is the tool of choice for spatially separating different wavelengths contained in a beam of light [6]. The grating technique

is shown in Figure 1.5. The device consists of a set of diffracting elements, such as narrow parallel slits or grooves, separated by a distance comparable to the wavelength of light. These diffracting elements can be either reflective or transmitting. With this method, separating and combining wavelengths is a parallel process.

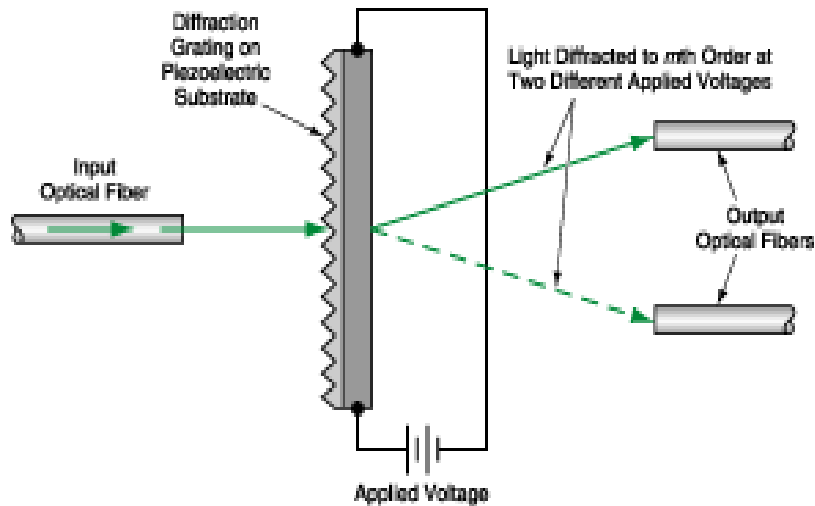


Figure 1.5 Diffraction grating technique for separating wavelengths

1.4.2 Wavelength Add/Drop Multiplexer

A wavelength add/drop multiplexer (WADM) allows the insertion or extraction of a wavelength from a fiber at a point between terminals. A WADM can operate either statically or dynamically. WADM consists of a de-multiplexer, followed by a set of 2×2 switches, one for each wavelength. The switches are followed by a multiplexer. The switches are managed electrically. They control which of the incoming wavelengths flow through the WADM and which are dropped locally. If some incoming wavelengths are dropped locally in WADM a new data stream can be added on to the same wavelength at this WADM location. More than one wavelength can be dropped and added if the WADM interface has the necessary hardware and processing capabilities [4].

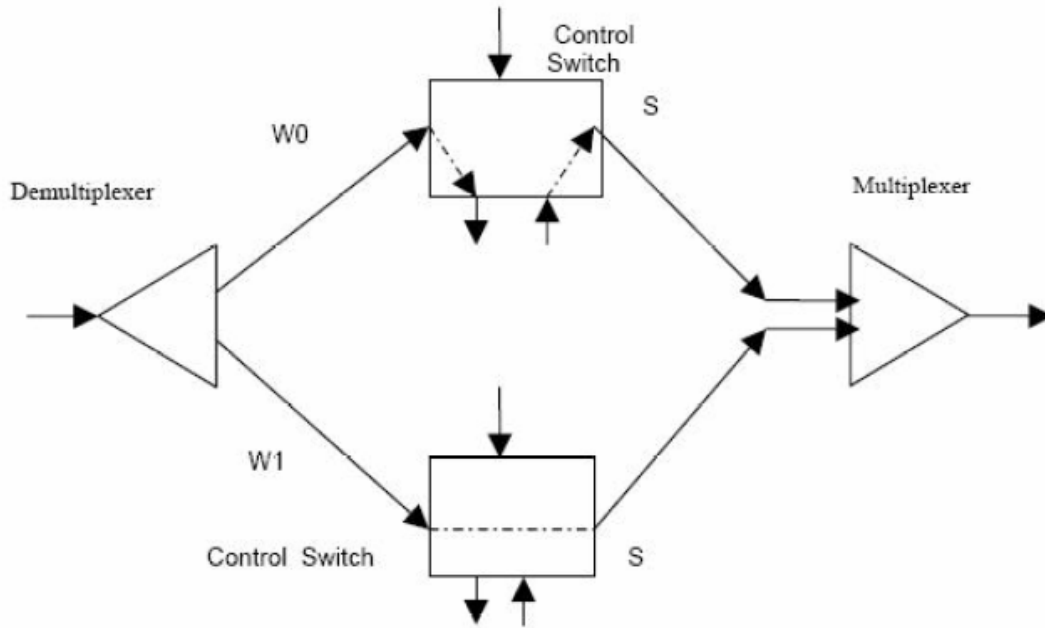


Figure 1.6 Wavelength add/drop multiplexer

1.4.3 Optical Crossconnects (OXC)

Optical crossconnects are used to route wavelengths between input ports and out ports. The main function of the OXC is to dynamically reconfigure the network at the wavelength level for restoration or to accommodate changes in bandwidth demand. OXC systems are expected to be the cornerstone of the photonic layer providing carriers more dynamic and flexible options in building network topologies with enhanced survivability. The architecture of an OXC is shown in Figure 1.7. The typical OXC capabilities are

- **Fiber switching:** the ability to route all of the wavelengths on an incoming fiber to a different outgoing fiber.
- **Wavelength switching:** the ability to switch specific wavelengths from an incoming fiber to multiple outgoing fibers.

- **Wavelength conversion**—the ability to take incoming wavelengths and convert them (on the fly) to other optical frequencies on the outgoing ports; this is necessary to achieve strictly non-blocking architectures when using wavelength switching.

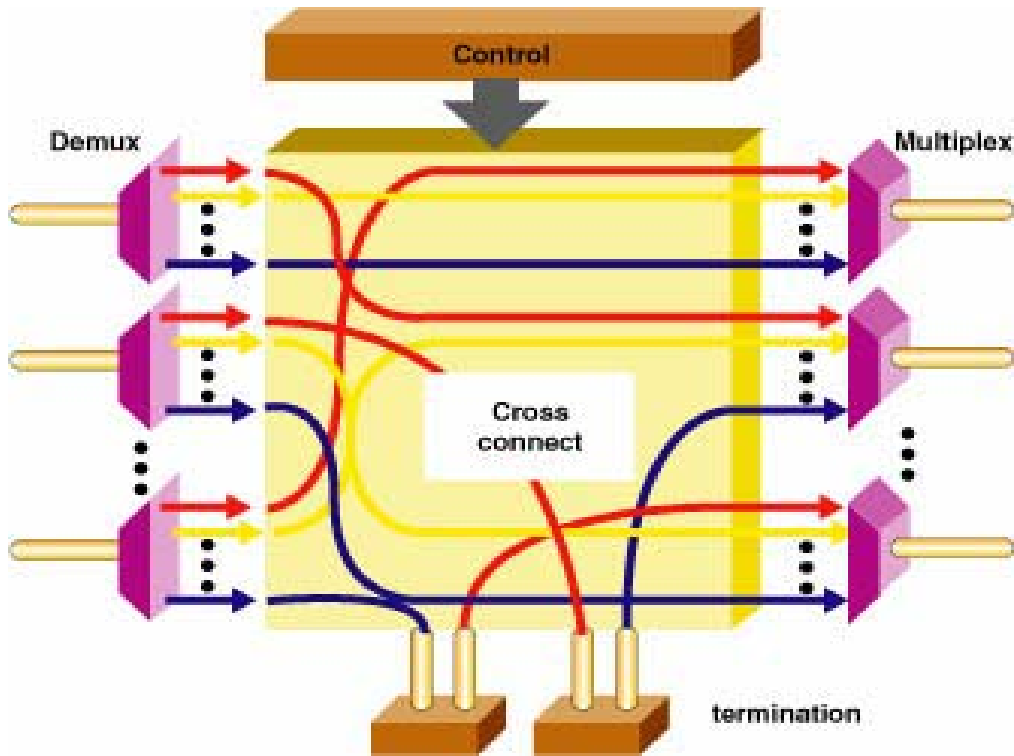


Figure 1.7 Wavelength cross-connect

OXC's can be divided into the following classes [3]:

- The fiber switch cross-connect (FXC)
- The wavelength selective cross-connect (WSXC)
- The wavelength interchanging cross-connect (WIXC)

A fiber switch cross-connect switches all of the wavelength channels on one input fiber to an output fiber, in effect acting as an automated fiber patch panel. FXC are less complex, and thus expected to be less costly, than a wavelength selective or wavelength

interchanging cross-connect. A wavelength selective cross-connect can switch a subset of the wavelength channels from an input fiber to an output fiber. Functionally, they therefore require de-multiplexing (in the frequency spectral domain) of an incoming wavelength multiplex into its individual constituting wavelengths. This cross-connect type offers much more flexibility than an FXC, allowing the provisioning of wavelength services, which in turn can support video distribution, distance learning, or a host of other applications. A wavelength interchanging cross-connect is a WSXC with the added capability to translate or change the frequency (or wavelength) of a channel from one frequency to another. This feature reduces the probability of not being able to route a wavelength from an input fiber to an output fiber because of wavelength contention. WIXC offers the most flexibility for restoration and provisioning of services. The WIXC may not be very cost effective since some circuits may not always need wavelength conversions. One effective method is to share wavelength converters.

1.4.4 Optical Amplifier

Optical amplification is required to compensate for various losses such as fiber attenuation, coupling and splitting loss in the star couplers, as well as coupling losses in the wavelength routers. The advent of a fiber optic repeater device called the Erbium doped fiber amplifier has enabled WDM to be a cost-effective technology. An Erbium doped fiber amplifier (EDFA), is an optical or IR repeater that amplifies a modulated laser beam directly, without opto-electronic and electro-optical conversion [5]. Some of the important properties which have led to using EDFAs in large numbers in optical transmission systems are high power conversion efficiency, high gain, low noise, and low polarization dependence and temperature sensitivity [3].

The structure of a typical EDFA is shown in Figure 1.8. The device uses a short length of optical fiber doped with the rare-earth element Erbium. When the signal-carrying laser beams pass through this fiber, external energy is applied, usually at IR wavelengths. This so-called pumping excites the atoms in the Erbium-doped section of the optical fiber, increasing the intensity of the laser beams passing through.

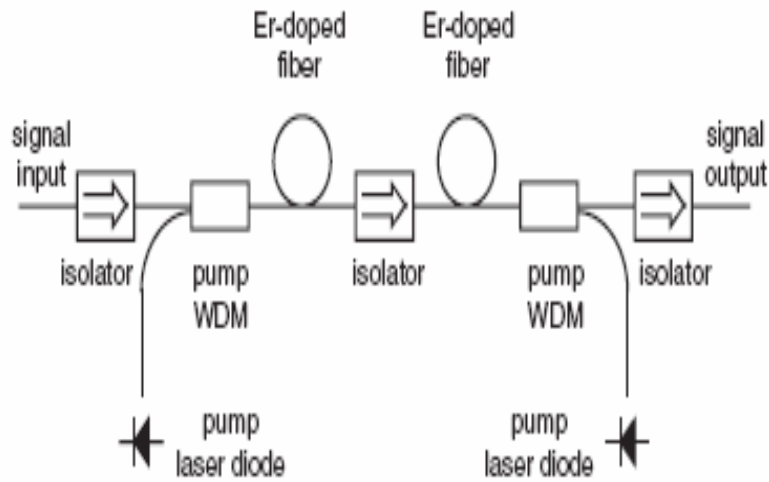


Figure 1.8 Erbium doped fiber amplifier structure

1.5 WDM Network Architectures

The most common classes of WDM network architectures are: Broadcast-and-select (local-area) networks and Wavelength routed (wide-area) networks. The following sections deal with these network architectures.

1.5.1 Broadcast-and-Select Networks

In a broadcast-and-select network, a passive coupler is connected to all the nodes in the network as shown in Figure 1.9. Each node in the network has a set of tunable optical transmitters and tunable optical receivers. A node sends its information to the star coupler on one of the available wavelengths using the tunable laser which produces optical information stream. The information from multiple sources is optically

combined by the star and the signal power of each stream is equally split and broadcast to all of the nodes. An optical filter is used by the destination nodes' receiver to extract the required wavelength stream from the received broadcast. When one node sends information, it is received by all the nodes in the network and only those nodes which need that information will tune their receivers to the desired wavelength. Thus, the network provides multicast capability. In this model, when a node failure occurs, the rest of the network can still function without any problems. Hence, the passive-star model enjoys a fault-tolerance advantage over some other distributed switching networks [4, 7].

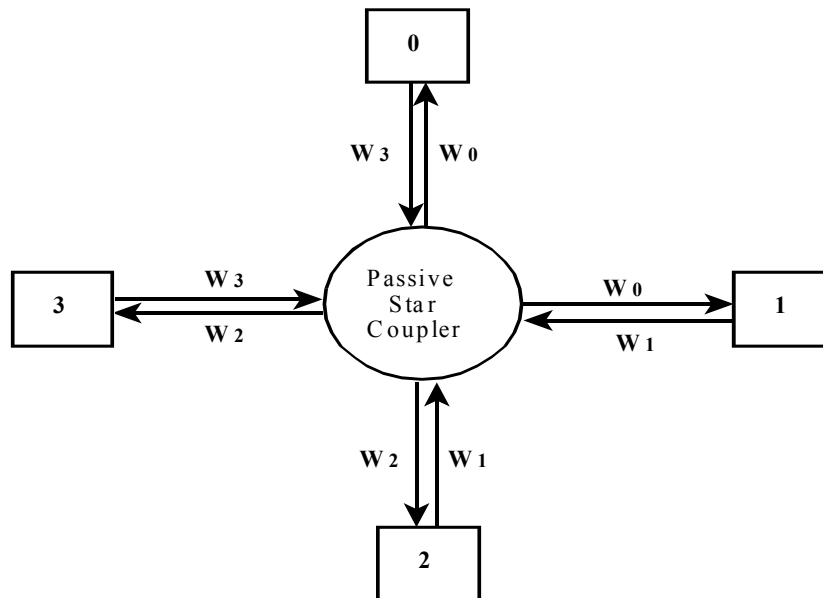


Figure 1.9 Broadcast-and-select network

However, broadcast and select networks have certain limitations. They require a large number of wavelengths, typically at least as many as the number of nodes in the network. Thus the networks are not scalable beyond the number of available wavelengths [5]. Since the transmitted power is split among the various nodes of the

network, the signal will not be able to span long distances. Because of these limitations, this model is suitable only for local area networks.

1.5.2 Wavelength Routed Networks

Wavelength routed networks have the potential to avoid the problems associated with the broadcast-and-select networks. They avoid the power splitting loss due to broadcast and they can be scalable to wide area networks. A wavelength routed network consists of wavelength cross connects (active switches) interconnected by point-to-point fiber links to form an arbitrary physical topology. Each node in the network is equipped with a set of transmitters and receivers, both of which may be tunable. Each end user is connected to the active switch by a fiber link. The combination of end user and its corresponding active switch is referred to as a node.

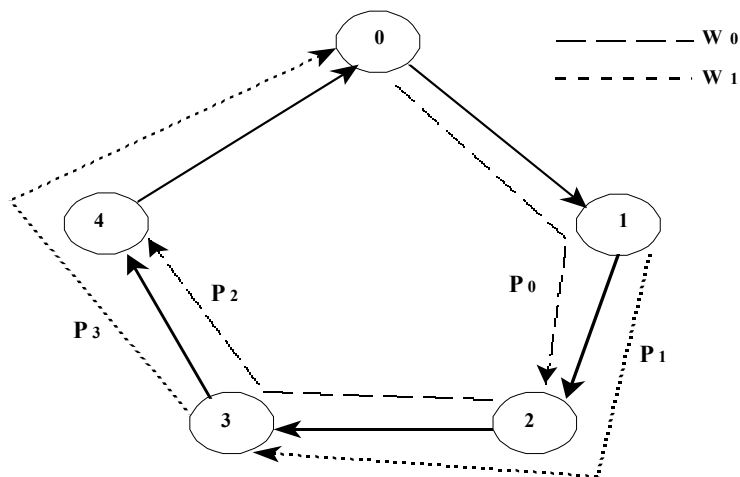


Figure 1.10 Wavelength routed network

In wavelength routed networks, the communication mechanism is called a *Lighpath*. A *lighpath* is an all-optical wavelength continuous path which is established between two nodes in the network. It may span more than one fiber link and is created by allocating the same wavelength throughout the path [6]. A message is sent from one node

to another node using a lightpath without requiring any optical-electronic-optical conversion or buffering at the intermediate nodes. The requirement that the same wavelength be used on all the links of the path between two nodes is called the *wavelength continuity constraint* [2] [6]. No two lightpaths can have the same wavelength on any common fiber. This is known as the *distinct wavelength assignment constraint* [6]. The wavelength continuity may not be necessary if the network is equipped with wavelength converters which have the ability to convert the information stream from wavelength to another wavelength without electronic conversion.

A typical wavelength routed network is shown in Figure 1.10. The network has five nodes and two wavelengths. Lightpaths need to be established between node pairs $\langle 0, 2 \rangle$, $\langle 1, 3 \rangle$, $\langle 2, 4 \rangle$ and $\langle 3, 0 \rangle$. The figure shows the lightpath establishment for those node pairs with out any problem. The lightpaths p_0 and p_2 use wavelength w_0 and lightpaths p_1 and p_3 use wavelength w_1 . Suppose we need to establish another lightpath between node pair $\langle 4,1 \rangle$. The route for this is $4 \rightarrow 0 \rightarrow 1$. Wavelength w_0 is available on the link $4 \rightarrow 0$ and wavelength w_1 is free on link $0 \rightarrow 1$. Though bandwidth is available along the path, a lightpath cannot be established because of wavelength continuity constraint.

1.6 Routing and Wavelength Assignment (RWA)

Routing and wavelength assignment is the fundamental control problem in WDM wavelength routed networks. In WDM wavelength routed optical networks, lightpaths need to be established before any communication takes place between the nodes. In order to establish a lightpath between two nodes, two decisions have to be made. The first is the selection of the path from the source node to the destination node

and the second is the selection of wavelength to be assigned to the path. Many problems in wavelength routed networks have RWA as a sub problem.

Depending on the traffic in the network, the RWA problem is classified into static and dynamic. In case of static traffic demand, the connection requests are known in advance. The traffic demand may be provided in terms of source-destination pairs. The objective is to assign routes and wavelengths so as to maximize the number of demands satisfied. In dynamic traffic demand, the connection requests arrive and depart randomly. The established lightpaths will remain only for a finite time. Since the traffic is dynamic, the network has no knowledge of future connection requests. Because of this, the dynamic RWA algorithms perform poorly when compared to the static RWA algorithms [10]. A dynamic RWA algorithm processes the connection requests strictly in the order of connection arrival time, whereas a static RWA algorithm processes the connection requests in the order decided by some heuristic. The RWA problem can be divided into route selection and wavelength selection.

1.6.1 Route Selection

Route selection algorithms can be classified into three types: fixed routing (FR), alternate routing (AR), and exhaust routing (ER).

Fixed routing: For each node pair in the network, a fixed route is assigned. These routes are calculated offline and they do not change with the changing network conditions. The performance degrades as the offered load increases.

Alternate routing: For each node pair in the network, a set of candidate routes are computed offline. When a connection request arrives, the route is selected from among only those in the set of candidate routes assigned for that node pair.

Exhaust routing: In exhaust routing, when a connection request arrives for a node pair, all the possible routes between the node pair are considered and one among them is selected. A conventional shortest path algorithm is typically used to find the best possible route.

1.6.2 Wavelength Selection

The wavelength selection algorithms can be classified into most-used, least-used, fixed-order, random-order, and round-robin.

Most-used: This algorithm gives preference to the wavelength which is used on the largest number of links in the network. The wavelengths are searched in descending order of their use. The main idea behind this algorithm is to pack the lightpaths tightly so that future connection requests will have many available wavelength continuous routes. In order to know the wavelength usage, the global state information of the network is to be known.

Least-used: In this case, the wavelength which is used on the least number of links in the network will be selected. This scheme attempts to distribute the load on the wavelengths uniformly across the entire network.

Fixed-order: All the wavelengths in the network are indexed. This algorithm searches for wavelength in a fixed order and the first free wavelength will be selected.

Random-order: All the wavelengths are indexed and the selection is done randomly. Each wavelength has equal probability of being selected.

Round-robin: This method tries to distribute the load on the wavelengths equally by assigning the wavelengths in a round-robin fashion from the pool of available wavelengths.

Another important issue in WDM wavelength routed networks is the connection blocking probability. It is a measure of how likely a connection request will get blocked because of unavailable network resources. The wavelength continuity constraint increases the blocking probability of connections with larger hop counts when compared to connections with smaller hop counts [10] [3]. Fairness and admission control algorithms are used to regulate network traffic and to provide fairness among connection requests.

1.7 Problem Formulation and Layout of Thesis

Wavelength division multiplexing technology on optical fiber communication has produced tremendous amount of raw bandwidth. Nowadays, bursty internet traffic is consuming most of the available bandwidth as opposed to non-bursty voice traffic [15]. This bursty internet traffic, which is increasing day by day, has to be handled with proper technology. An all-optical transport protocol has to be developed to utilize this bandwidth efficiently and to avoid optical buffering while handling bursty traffic.

Circuit switching and packet switching have been used for many years. However, these technologies are mainly used with voice and data traffic, respectively [12] [13]. Though optical packet switching can handle internet traffic more efficiently, the optical hardware technology has not been developed well enough to afford this. Optical burst switching (OBS) is a scheme which has been viewed as a viable option for handling the bursty traffic until optical packet switching technology becomes a reality [9][12][13]. OBS has been designed to achieve a balance between the coarse-grained circuit switching and fine-grained packet switching. In this work, we

investigate the various issues related to optical burst switching technologies. Our main interest is concentrated on contention resolution techniques in OBS networks which play a great part in reducing packet loss and congestion in the network. This thesis presents and studies the performance of three new techniques for reducing packet loss in OBS schemes. The results obtained through simulations show that our schemes exhibit low blocking probabilities when compared to other techniques proposed in the literature.

The remainder of the thesis is organized as follows: Chapter 2 discusses the details of optical burst switching including various reservation techniques proposed in the literature. It also covers the traditional contention resolution techniques used in OBS. Chapter 3 proposes our new packet loss reduction techniques along with their signaling protocols and timing diagrams. Chapter 4 studies the performance of our schemes using simulations. It also compares our results with those for previously known schemes. Chapter 5 concludes the thesis and identifies areas of future work.

Chapter 2

Optical Burst Switching

Optical burst switching (OBS) is a promising new technique which attempts to address the problem of efficiently allocating resources for bursty internet traffic. Circuit switching and packet switching have been used for many years for voice and data communication, respectively. OBS can combine the best of the coarse-grained circuit-switching and the fine-grained packet-switching paradigms while avoiding their shortcomings, thereby efficiently supporting bursty traffic generated by upper level protocols or high-end user applications directly [11, 13, 16]. OBS differs from circuit and packet-switching primarily in whether *cut-through* or *store-and-forward* is used and in how bandwidth is reserved (and released).

In circuit switching, a dedicated path has to be established between two nodes before any data transmission takes place [19]. The time taken for establishing such path is equal to the round trip delay. The reserved resources stay idle for the entire path setup time and account for poor resource utilization. The benefit of Optical Burst Switching (OBS) over conventional circuit switching is that there is no need to dedicate a wavelength for each end to end connection [13]. In addition to this, the path setup time is much less than the round-trip delay. In packet switching, the data is broken into small packets and transmitted. The data is transmitted using “store and forward” technique. The resources can be shared by different sources. End stations can send/receive data at their own speed [11, 12]. The individual packet can be individually switched or a virtual circuit can be set up. Packet switching has large buffer requirement and complex control and sync issues. For the optical domain, packet switching is

not yet feasible because of optical hardware limitations. Optical RAMs do not exist yet to meet the high buffer requirements of packet switching. In addition, optical burst switching seems to be more viable than optical packet switching since burst data does not need to be buffered or processed at intermediate nodes. This allows the strengths of optical switching technologies to be leveraged effectively and the problem of buffering in the optical domain to be circumvented. OBS combines the advantages of both circuit and packet switching and ensures efficient bandwidth and resource utilization [11, 15].

Optical Switching (paradigm)	Bandwidth Utilization	Latency (set-up)	Optical Buffer
Circuit	low	high	not required
Packet/Cell	high	low	required
Burst	high	low	not required

Figure 2.1 Comparison of optical switching schemes

Optical burst switching is based on the separation of the control plane and the data plane [14]. The basic switching entity in OBS is called a burst which is a (digitized) talk spurt or a data message. In optical burst switching data packets are aggregated into much larger bursts before transmission through the network. This allows amortization of the switching overhead across multiple packets. The data burst (DB) is preceded in time by a control burst (CB), which is sent on a separate control wavelength. The control burst requests resource allocation at each switch. At each intermediate node, the CB is processed electronically and the time taken for processing a CB is known as the “processing time”. After processing, the CB reserves a wavelength on an outgoing link for the DB. This reservation will be for a time

period starting from the time the DB is expected to arrive to the time the DB is transmitted completely. The reservation time and duration can be calculated using the offset and the DB length. If no reservation can be made, then the CB is dropped. If the reservation is successful, the CB is forwarded to the next node along the path to the destination. The offset is chosen in such a way that the reservation is already made at each node before the DB arrives at that node. OBS uses one-way reservation schemes with immediate transmission, in which the data burst follows a corresponding control burst after waiting for a short offset time without waiting for an acknowledgement [13]. The offset time gap between the CB transmission and the DB transmission is generally used for aggregating the data packets into a data burst [12, 16].

2.1 OBS Network Architecture

An OBS network consists of optical core nodes and electronic edge nodes connected by WDM links. Packets are assembled into bursts at network ingress, which are then routed through the OBS network and disassembled back into packets at network egress to be forwarded to their next hops [14]. Edge nodes provide burst assembly/disassembly functions. A core node is mainly composed of an optical switching matrix and a switch control unit.

An OBS node is built from optical and electronic components besides optical receivers and optical transmitters. The optical components include multiplexers (Mux), demultiplexers (Demux) and an optical switching network (OSN). The electronic components include input modules (IM), output module (OM), a control burst router (CBRT), and a scheduler [2]. An optical burst switch control unit transfers a burst coming in from an input port to its destination output port. Depending on the switch architecture, it may or may not be equipped with optical buffering. The fiber links carry multiple wavelengths, and each wavelength can be seen as a

channel. The control packet associated with a burst may also be transmitted in-band over the same channel as data, or on a separate control channel. The burst size may be fixed to carry one or more IP packets.

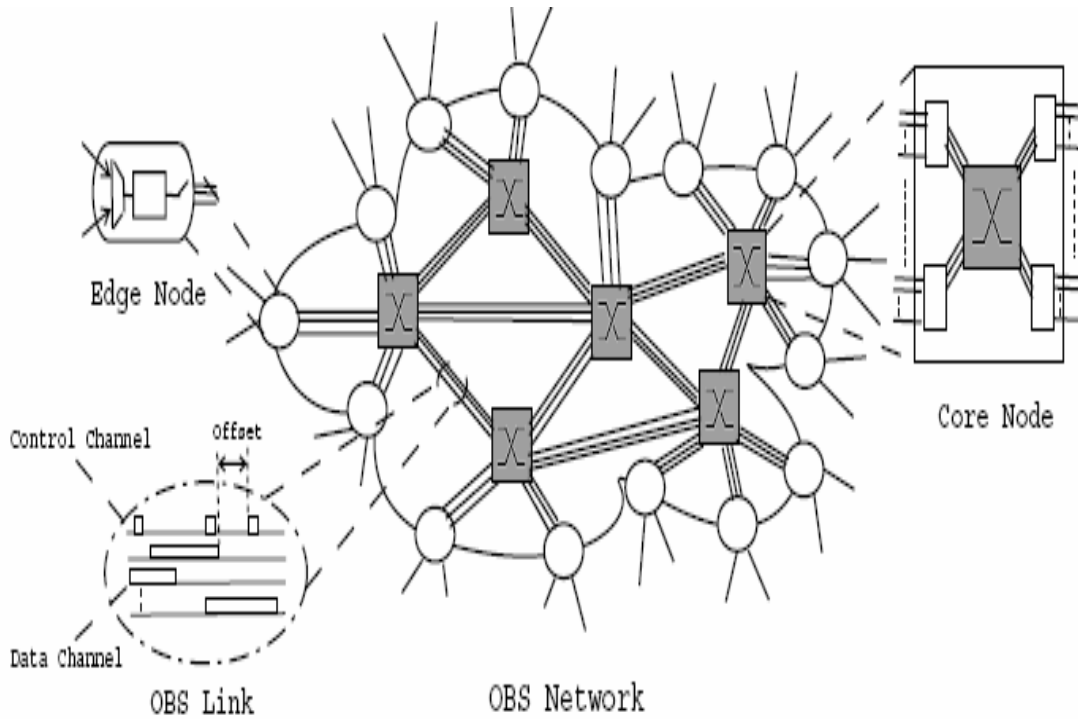


Figure 2.2 OBS network architecture

When an edge node intends to transmit a data burst, it first sends a control burst on the control wavelength to the nearest core node. At the core node, the CB on the control wavelength is input to the corresponding IM, which converts the CB into electronic form by the receiver. The control fields are extracted from the CB. The CBRT uses these control fields to determine the next outgoing fiber for the corresponding DB by consulting a routing table maintained locally. The CB is scheduled for transmission onto the selected outgoing link by the scheduler and the CB is buffered until the scheduled time. The scheduler maintains a CB queue. The scheduler also reserves wavelength on the determined links for the upcoming DB. The CB is then forwarded onto

the OM, which updates its control fields and transmits it to the selected outgoing fiber using the optical transmitter. Just before the DB arrives, the switching element in the node is configured to connect the input port to the corresponding output port for the entire duration of the burst transmission. If the CB is unable to reserve the wavelength for its corresponding DB, then the CB will be dropped as well as its DB.

2.2 Reservation Schemes in OBS

Optical Burst switching schemes differ based on how and when the network resources like bandwidth, are reserved and released. Optical burst switching is an adaptation of burst switching technique in asynchronous transfer mode (ATM) networks, known as ATM block transfer (ABT) [17]. There are two versions of ABT: ABT with delayed transmission and ABT with immediate transmission.

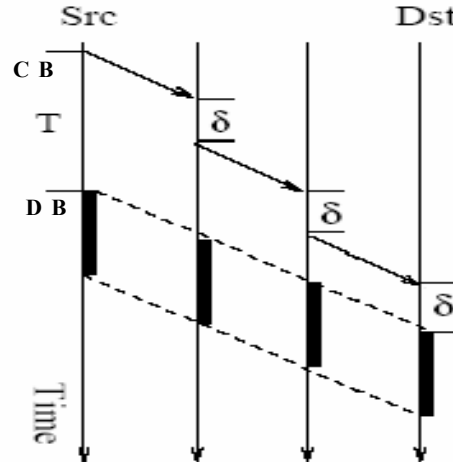


Figure 2.3 Use of offset delayed reservation

In case of an immediate reservation scheme, an output wavelength is reserved for a data burst immediately after the arrival of the corresponding control burst; if a wavelength cannot be

reserved at that time, then the setup message is rejected and the corresponding burst is dropped [16]. In a delayed reservation scheme, the CB and the DB are separated in time by an offset value in order to accommodate the processing of the CB. An output wavelength is reserved for a burst just before the arrival of the first bit of the burst. If, upon arrival of the setup message, it is determined that no wavelength can be reserved at the appropriate time, then the setup message is rejected and the corresponding burst is dropped [16]. These two techniques have been adopted in OBS. Depending on bandwidth reservation, offset time and control management, three schemes for OBS implementation have been proposed: Tell-and-go (TAG) [16], Just-in-time (JIT) [12][13] and Just-enough-time (JET) [14].

2.2.1 Tell-And-Go (TAG)

This is an immediate reservation scheme. In TAG, the CB is transmitted on a control channel followed by a DB, which is transmitted on a data channel with zero or negligible offset. The CB reserves the wavelength and buffer (FDL) at each intermediate node along the path for the DB. When the DB reaches an intermediate node, it is buffered using the reserved FDL until the CB processing is finished. Then the DB is transmitted along the reserved channel. If no wavelength is available for reservation, the burst is dropped and a negative acknowledgement (NAK) is sent to the source. The source node sends another CB after transmitting the DB for releasing the reserved wavelengths along the path. Here, the burst size is not fixed in advance. FDLs are expensive and they can only buffer data optically for a very short time. Optical buffering is the main drawback of this scheme. Furthermore, if the “release” CB which is sent to release the reserved bandwidth along the path is lost, then these wavelengths will not be released and this creates bandwidth wastage [13, 14].

2.2.2 Just-In-Time (JIT)

This scheme also comes under immediate reservation. Here, an output wavelength is reserved for the upcoming burst as soon as the CB processing is finished. The source transmits the DB after an offset time which is greater than the total CB processing time. If the wavelength is not available, the burst is dropped. The difference between JIT and TAG is that the buffering of the DB at each node is eliminated by inserting a time gap between the CB and the DB. Since the bandwidth is reserved immediately after the CB processing, the wavelength will be idle from the time the reservation is made till the first bit of the DB arrives at the node. This is because of the offset between the CB and the DB. Since the offset value decreases as the CB gets closer to the destination, the idle time also decreases. An in-band-terminator is placed at the end of each burst which is used by each node to release the reserved wavelength after transmitting the DB [12, 19].

Wavelength reservation in JIT at an intermediate node is shown in Figure 2.4. Let t be the time a CB arrives at some OBS node along the path to the destination. Let T_{setup} be the amount of time it takes an OBS node to process the control burst. Let T_{offset} be the offset value of a burst. This is equal to the time gap between the CB and the DB transmission. The offset value depends on (1) the wavelength reservation scheme, (2) the number of nodes the burst has already traversed, and (3) other factors, such as whether the offset is used for service differentiation [11]. T_{oxc} is the amount of time it takes the OXC to configure its switch fabric to set up a connection from an input port to an output port. Once the processing of the CB is complete at time $t + T_{\text{setup}}$, a wavelength is immediately reserved for the upcoming burst, and

the operation to configure the OXC fabric to switch the burst is initiated. When this operation completes at time $t + T_{\text{setup}} + T_{\text{OXC}}$, the OXC is ready to carry the burst.

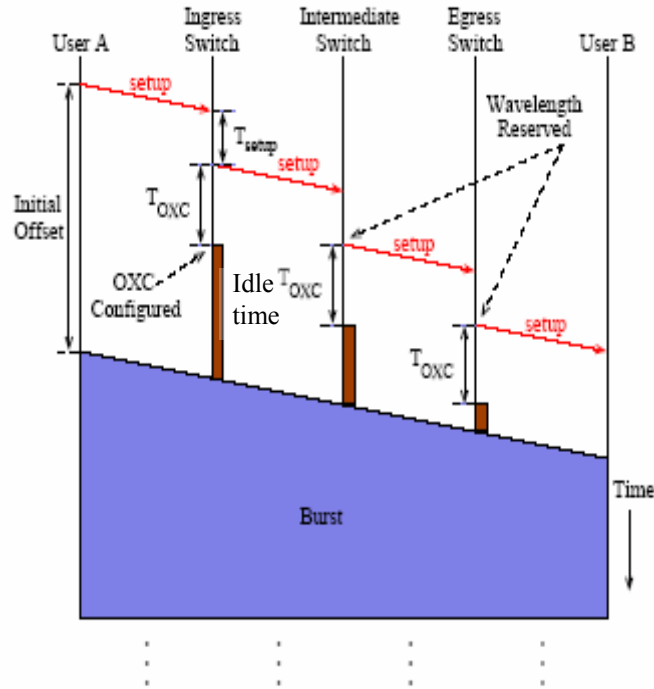


Figure 2.4 JIT scheme

Note that the burst will not arrive at the OBS node under consideration until the time is $(t + T_{\text{offset}})$. As a result, the wavelength remains idle for a period of time equal to $(T_{\text{offset}} - T_{\text{setup}} - T_{\text{OXC}})$. Also, since the offset value decreases along the path to the destination, the deeper inside the network an OBS node is located, the shorter the idle time between the instant the OXC is configured and the arrival of the burst [16].

2.2.3 Just-Enough-Time (JET)

This is a delayed reservation scheme. Here, the size of the burst is decided before the CB is transmitted by the source. The offset between CB and DB is also calculated based on the hop count between the source and destination. At each node, if bandwidth is available, the CB

reserves wavelength for the upcoming burst for a fixed duration of time. The reservation is made from the time when the first bit of DB reaches the node till the last bit of DB is transmitted to the output port. This eliminates the wavelength idle time which is the main difference between JET and JIT. Since the wavelength is reserved for a fixed duration, there is no need for explicit signal for releasing the reserved wavelength along the path. Since there is no wastage of bandwidth in this scheme, the network utilization for this scheme is higher than with the other schemes. But, this scheme involves complex scheduling when compared to other schemes.

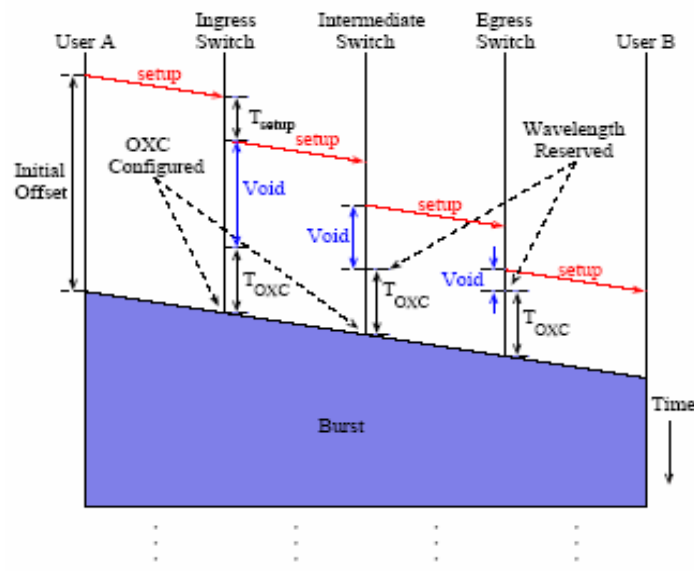


Figure 2.5 JET scheme

The operation of delayed reservation in JET is shown in Figure 2.5. Let us again assume that a control burst arrives at an OBS node at time t . Let the offset be T_{offset} and let the length of the DB be Δ . The first bit of the corresponding burst is expected to arrive at time $t+T_{offset}$. After processing the CB, the node reserves a wavelength for the DB starting at time $t_1 = t + T_{offset} - T_{OXC}$ and ending at time $t_1+\Delta$. At time t_0 , the OBS node instructs its OXC fabric

to configure its switch elements to carry the data burst, and this operation completes just before the arrival of the first bit of the burst. Thus, whereas immediate reservation protocols only permit a single outstanding reservation for each output wavelength, delayed reservation schemes allow multiple setup messages to make future reservations on a given wavelength (provided of course, that these reservations, do not overlap in time). A void is created on the output wavelength between time $(t + T_{\text{setup}})$, when the reservation operation for the upcoming burst is completed, and time $(t_1 = t + T_{\text{offset}} - T_{\text{OXC}})$, when the output wavelength is actually reserved for the burst. In an attempt to use the voids created by the earlier setup messages, void filling algorithms are employed in JET [16].

TAG and JIT schemes are significantly simpler than JET since they do not involve complex scheduling or void-filling algorithms. On the other hand, previous studies have shown that JET performs better than either JIT or TAG in terms of burst loss probability [14] [16].

2.3 Contention Resolution Schemes

Contention resolution is necessary for handling certain cases where two or more bursts try to reserve the same link and the same wavelength for the same time. This is called external blocking. In packet switching, this is avoided by buffering the contending packets. In OBS, when two or more bursts contend for the same wavelength and for the same time duration, only one of them is allotted the bandwidth. In such case, one or a combination of the following three major options for contention resolution can be applied in addition to the option of dropping the unsuccessful bursts.

Wavelength domain: By means of wavelength conversion, a burst can be sent on a different wavelength channel of the designated output line [18].

Time domain: By utilizing an FDL buffer, a burst can be delayed until the contention situation is resolved. In contrast to buffers in the electronic domain, FDLs only provide a fixed delay and data leave the FDL in the same order in which they entered [18].

Space domain: In deflection routing, a burst is sent to a different output link of the node and consequently on a different route towards its destination node. Space domain can be exploited differently in case several fibers are attached to an output line. A burst can also be transmitted on a different fiber of the designated output line without wavelength conversion [18].

When there is no available unscheduled channel, and a contention cannot be resolved by any one of the above techniques, one or more bursts must be dropped. The policy for selecting which bursts to drop is referred to as the *soft contention resolution policy* and is aimed at reducing the overall burst loss rate, BLR, and consequently, enhancing link utilization [9]. Several soft contention resolution algorithms have been proposed and studied in earlier literature, including the shortest-drop policy [25] and look-ahead contention resolution [26]. In burst segmentation, only that part of the burst which is involved in a reservation conflict will be dropped [16]. The contention resolution policies are considered as *reactive* approaches in the sense that they are invoked after contention occurs. An alternative approach to reduce network contention is by *proactively* attempting to avoid network overload through traffic management policies [9].

2.3.1 Optical Buffering

Optical buffering is achieved through the use of *fiber delay lines* (FDL). Due to the lack of optical random access memory, FDL is currently the only way to implement optical buffering. By implementing multiple delay lines in stages [16] or in parallel [17], a buffer may

be created that can hold a packet for a variable amount of time. In any optical buffer architecture, the size of the buffers is severely limited, not only by signal quality concerns, but also by physical space limitations. The FDLs are bulky. To delay a single packet for $5\mu\text{s}$ would require over a kilometer of fiber [17]. Because of this size limitation of optical buffers, a node may be unable to effectively handle high load or bursty traffic conditions. Furthermore, signal dispersion and attenuation are some of the limitations of FDLs. Because of these drawbacks, delay lines may be acceptable in prototype switches, but are not commercially viable.

The reservation scheme involving optical buffer contention resolution consists of two phases: *wavelength reservation* in the output port and *FDL reservation* in the optical buffer [19]. During the wavelength reservation phase, the scheduler checks the required wavelength at the output port first. If the required wavelength will be idle at $t + \Delta$ and the idle duration is long enough to accommodate the DB, this wavelength is reserved immediately. If the wavelength is not available for that particular period of time, then the minimum waiting time W for reserving the wavelength is computed. If $W > D$ (fiber delay), the DB has to be discarded, since no FDL can provide such a delay. In the case of $W \leq D$, FDL reservation is performed. The wavelength reservation is made for the latest available time and until then the DB will be buffered through the reserved FDL. The DB will be transmitted from the FDL onto to the reserved output wavelength as soon as the waiting time equals W . In case, both the required wavelength and the FDL are not available, then the burst will be dropped. Optical buffering is generally used in combination with the other contention resolution schemes such as wavelength converters and deflection routing to improve performance. However, they are not feasible for large scale deployment.

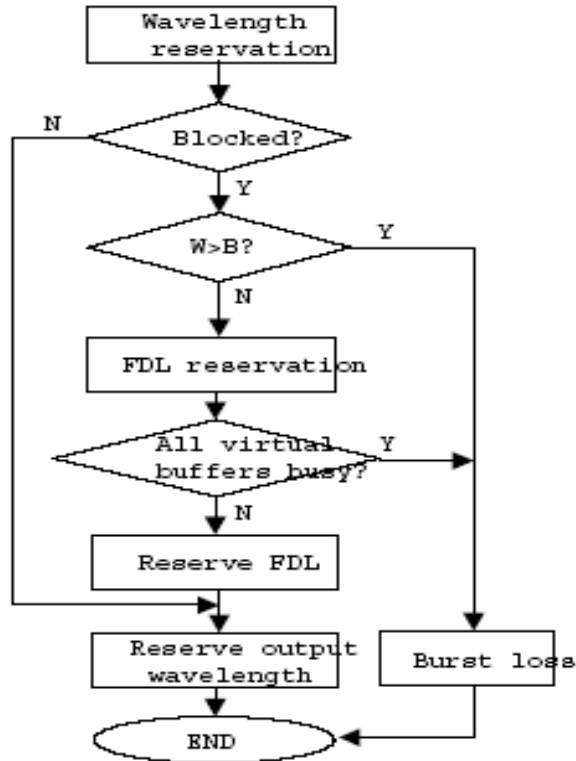


Figure 2.6 Contention resolution using FDL

2.3.2 Wavelength Conversion

In wavelength routed networks, lightpaths are required to carry messages. The wavelength continuity constraint has to be satisfied for successful communication. If a route is free but no common wavelength is available on it, then it cannot be used for setting a lightpath. This results in the blocking of the connection, even though the bandwidth is available. All such connections would have been successful if there were no wavelength continuity constraint.

Wavelength conversion is the process of converting a wavelength on an incoming channel to another wavelength on the outgoing channel [2, 4]. A wavelength converter is a device that is capable of converting an incoming signal's wavelength to a different outgoing wavelength. The wavelength continuity constraint can be relaxed by the use of wavelength

conversion. The wavelength conversion is classified into: optical-electronic conversion and all-optical conversion. The disadvantages of optical-electronic-optical conversion (such as complexity and large power consumption) have increased the interest on to all-optical conversion [19].

The following are the different categories of wavelength conversion:

Full conversion: Any wavelength shifting is possible. Channels can be connected regardless of their wavelengths.

Limited conversion: Wavelength shifting is restricted so that not all combinations of channels may be connected.

Fixed conversion: A restricted form of limited conversion such that, for each node, each channel maybe connected to exactly one pre-determined channel on all other links.

Sparse wavelength conversion: Networks are comprised of a mix of nodes having full and no wavelength conversion capabilities; i.e. only a subset of nodes in the network have conversion capability.

The concept of wavelength conversion is shown in Figure 2.7. Assume that connections are required to be established between node pairs (C, D) and (A, D). Both connections will select the wavelength W_1 for lightpath establishment. At node B, both connections try for wavelength W_1 on link BD. Only one of the connections can be accepted. Let that be the connection (C, D). Wavelength W_2 is available on the link BD. Since the connection (A, D) is unable to satisfy the wavelength continuity constraint, it would be dropped. But, by converting the wavelength of connection (A, D) from W_1 to W_2 , the connection can be routed onto link BD. Thus, the connection will be successful by using the wavelength conversion capability.

Wavelength converters offer a 10%-40% increase in reuse values when wavelengths availability is small [15]. There are many wavelength conversion algorithms and algorithms to minimizing the number wavelength converters. Despite the high expectations and some promising experimental reports, wavelength conversion technologies are as yet immature and are highly expensive for deployment in real networks.

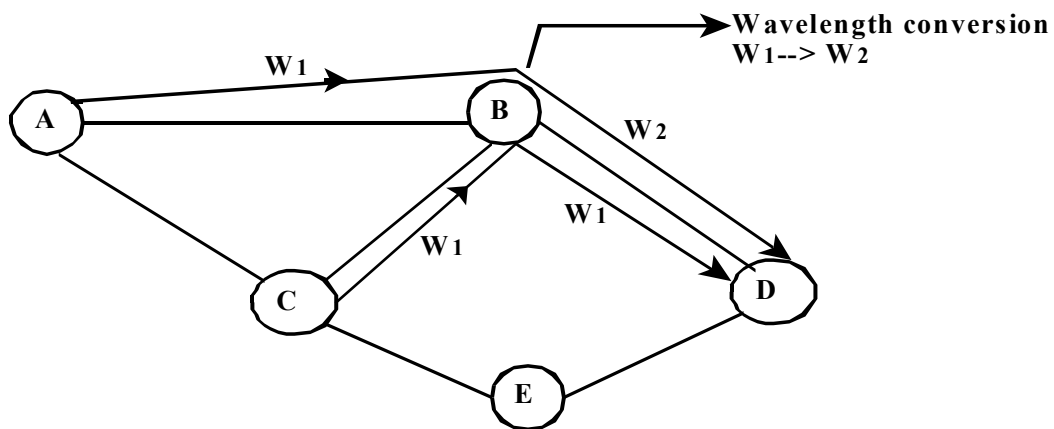


Figure 2.7 Wavelength conversion

2.3.3 Deflection Routing

Deflection routing is the approach of resolving contention by routing a contending packet to an output port other than the intended output port [22, 23, 24]. However, the deflected packet may end up following a longer path to its destination. As a result, the end-to-end delay for a packet may be unacceptably high. Deflection routing is generally not favored in electronic packet-switched networks due to potential looping and out-of-sequence delivery of packets. In WDM optical networks where buffer capacity is very limited and wavelength conversion is not feasible, implementation of deflection routing may be necessary in order to maintain a reasonable level of packet losses.

An example of deflection routing in WDM networks is given in Figure 2.8. Both nodes A and B are sending bursts to node E. Before sending bursts, nodes A and B send control packets (denoted as C (A, E) and C (B, E)) on their out-of-band control channels for bandwidth reservation for their respective data bursts. Let's say, C (B, E) arrives at Node C earlier than C (A, E). In this case, the output link CE is reserved by C (B, E). When C (A, E) arrives at node C, the link CE is not available. Without deflection, this burst will be dropped. But, Node C checks other output links and selects the deflection link CD which is idle, to deflect B (A, E). Node D forwards B (A, E) via the link between D and E based on its routing table. Since every node performs deflection routing in this manner, the deflected burst arrives at its destination with some extra propagation delay, i.e., it traverses several additional nodes than the shortest path. The idle optical links can be considered as fiber delay lines for “buffering” the blocked bursts. The bursts in the congested part of the network are then distributed to other underused parts, thus overall link utilization and network performance can be improved. If the burst cannot be deflected, then it will be dropped. Such an instance will be referred to as “normal deflection failure” in this thesis.

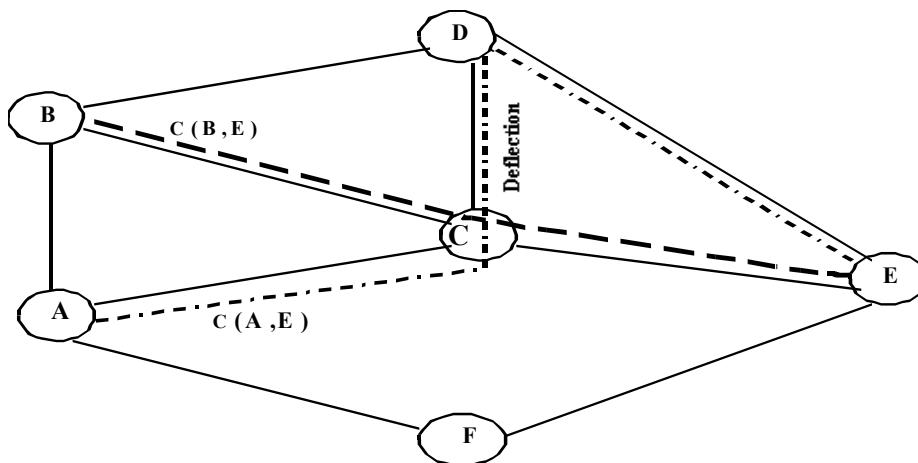


Figure 2.8 Deflection routing

Deflection routing implementation in OBS has many benefits. When a burst is dropped, it wastes the bandwidth on the partially established path. If the burst data has been injected into the network, the network should do the best to forward it to the destination, rather than simply drop it. Also, when a retransmission of the dropped burst is done, the total transmission delay will be the sum of the delay of the dropped burst and the delay of the retransmitted burst. This delay becomes very large when retransmitting a blocked burst in long- distance links. By applying suitable algorithms like *limited deflection* [18], burst looping can be reduced. In JET, deflection routing coupled with optical buffering (FDL) tends to reduce the problem of insufficient offset time [19].

Chapter 3

New Deflection-Based Contention Resolution Schemes

As of now, optical wavelength conversion and optical buffering technologies are very immature. It seems that the most viable option for reducing burst loss caused by contentions is deflection routing. However, the traditional deflection routing scheme doesn't consider all the available resources in resolving a contention. Due to this, the degree of success in contention resolution gained by the deflection routing scheme is not satisfactory. In this chapter, three new schemes are proposed for handling contentions that aim to improve on the existing schemes.

Consider Figure 2.8 in the previous chapter. As the control burst C (A,E) reaches node C, the status of the outgoing links at node C is checked. If both the links CE and CD are unavailable for reservation, then the burst is dropped. Basically, at each node, each burst has only one chance for deflection. If no idle bandwidth is available at any node, the burst is dropped without getting a second chance. For instance, suppose that node C is congested. If the control burst fails to reserve bandwidth at node C even on the deflection route, then the control burst will be dropped. Since the complete network state is not known to all the nodes, nodes will try to send their bursts through C till they realize the congestion at C after losing some bursts.

We propose a new scheme called **“Backtrack on deflection failure”** which provides a second chance to a blocked burst when a deflection failure occurs. Two variants are proposed to handle the backtracking delay involved in this scheme. Furthermore, we propose a third scheme called **“Bidirectional reservation on burst drop”** in which bandwidth reservation is made in both the forward and the backward directions at the same time. This

scheme comes into effect only when a control burst gets dropped due to bandwidth unavailability. In the following sections, we describe how these new schemes work.

3.1 Backtrack on Deflection Failure

In this scheme, at any node, if the control burst fails to reserve a wavelength on any of its primary or deflection routes, the burst will not be dropped as in earlier deflection schemes. Instead, the burst in this scheme will get a second chance to backtrack to the previous node and may get routed through any deflection route available at the previous node. Due to backtracking, each burst will face an increase in the lightpath setup delay. The additional setup delay will be equal to twice the propagation delay between the two nodes involved in deflection failure and backtrack and twice the control burst processing time. This is because, on deflection failure, the burst will backtrack to the previous node and no reservations will be made during this round trip time period. When this happens, it will reduce the offset gap between the control burst and its corresponding data burst. Therefore, the chances of the data burst reaching a node before a reservation is made will increase. To avoid such an event, the extra delay created through backtracking should be properly accommodated. This is part of the tradeoff involved in providing a second chance for any burst. In order to accommodate the extra delay that may be caused by deflection failure and subsequent backtracking, two approaches are proposed. They are **“Increase in the initial offset”** and **“Open loop wavelength reservation”**.

3.1.1 Routing Protocol

The lightpath setup mechanism involves four types of control bursts: primary control burst (PCB), backtrack control burst (BCB), probe burst (PB) and backtrack probe burst (BPB) in addition to a negative acknowledgement (NAK). Whenever a burst drop occurs,

a NAK is sent to the source to inform it about the burst drop. A primary control burst is used for wavelength reservation along the path to the destination. This control burst contains all the information required for lightpath setup. Whenever a normal deflection failure occurs, a BCB is created which is a copy of the corresponding PCB. The BCB is used for backtracking to the previous node on normal deflection failure. A probe burst is similar to a PCB except that no reservation will be made by a probe. The main function of PB is to probe the network and to inform the previous node about congestions. BPB is similar to BCB. When PB encounters contention and deflection failure, a BPB is created and sent to the previous node to inform it about the contention. The most common information fields and their descriptions in a typical control burst are shown in Figure 3.1.

Information	Description
<i>Packet identifier</i>	Kind of control burst (PCB,BCB,PB,BPB,ACK,NAK)
<i>Sender address</i>	Source node address of the burst
<i>Receiver address</i>	Destination node address of the burst
<i>Burst number</i>	Sequence number of the burst
<i>Offset time</i>	Time gap between control and data bursts
<i>Absolute time</i>	Departure time of CB at each node
<i>Burst length</i>	Duration of the data burst
<i>Timeout</i>	Time for the burst to live in the network to prevent looping
<i>Backup flag</i>	Set when both primary and deflection routes are available
<i>Deflection flag</i>	Set when the burst deflects or backtracks

Figure 3.1 Information fields in a control burst

3.1.2 Backtrack on Deflection Failure with Increase in Initial Offset

In order to provide backtracking capability for a control burst, the extra delay that will be caused by backtracking should be considered. One solution for this problem is to increase the initial offset. Initial offset of a burst is generally equal to the total processing time of the control burst from its source to the destination. This should be increased to accommodate the backtrack delay. The backtrack delay is equal to the sum of the round trip propagation delay between any two nodes in the network and two times the control processing delay. No reservation is made during this entire time duration. Thus, for providing backtracking capability, the total initial offset should be greater than the sum of the total processing delay and the backtrack delay.

3.1.2.1 Routing Procedure

The routing procedure for each intermediate node in an OBS network is as follows:

When a node receives a control burst, it is processed. Depending on the status of the outgoing links and the status of the information fields in the CB, the node takes an appropriate decision. A routing algorithm for the above scheme which describes all possible routing decisions is given below.

1. Begin

2. If (BURST IDENTIFIER=PCB) then

3. If (both the primary and deflection routes are available) then

4. Make reservation on the primary link. Forward a PCB along the reserved path.

5. Convert a copy of PCB into a PB and send it along the deflection link.

6. end-if

7. If (only one among the primary and deflection links is available) then

8. Make reservation on the available link and forward the PCB on that link.
9. **If** (neither the primary nor the deflection link is available) **then**
10. **If** (Backup flag is set) /*it denotes that the burst can backtrack to the previous node in the path and can take the deflection route available*/ **then**
11. Convert the PCB into a BCB and send it to the previous node.
12. **Else**
13. Drop the burst and send a NAK to the source.
14. **end-if**
15. **end-if**
16. **If** (BURST IDENTIFIER=BCB) **then**
17. Delete the reservation made on the primary link for this burst (because the burst faced deflection failure at the next node on that link).
18. **If** (deflection route is available at the present node) **then**
19. Make reservation on that link. Convert the BCB into a PCB and forward it on the deflection link.
20. **Else**
21. Drop the burst and send a NAK to the source.
22. **end-if**
23. **If** (BURST IDENTIFIER=PB) **then**
24. **If** (both the primary and deflection routes are available) **then**
25. Forward the probe burst on the primary output link. No reservation is made.
26. **If** (only one among the primary and deflection links is available) **then**
Forward the probe burst along the available output link.

27. If (neither primary nor deflection route is available) then

Transmit the probe burst to the previous node on its traveled path after changing the burst identifier to a BPB (backtrack probe burst).

28. end-if

29. If (BURST IDENTIFIER=BPB) then

30. Mark the link, between the present node and the node from which the BPB has backtracked, as unavailable for that particular burst. /* Due to this, if the primary CB corresponding to that probe burst reaches this node, it will not make reservation along this path which has a deflection failure on the next node in the path. The PCB will opt for an alternate available route. This avoids backtracking of the primary control burst due to unavailability of bandwidth on the next node which had been probed by its PB*/.

31. end-if

32. End

3.1.3 Backtrack on Deflection Failure with Open Loop Reservation

When a deflection failure occurs at a node, the control burst backtracks if it has an available deflection route at the previous node. In the previous scheme, in order to provide backtracking capability, the offset has been increased. Let's say that only 20% of the bursts get blocked due to normal deflection failure and hence will utilize backtracking capability. The remaining 80% of the bursts will thus be successful without backtracking. Even though these bursts don't use the backtracking capability, they face an extra delay due to the increase in initial offset which provides the backtracking capability. This may be considered a drawback. By reserving available bandwidth on the backtrack link, this initial offset increase can be

eliminated. The scheme we describe here uses this idea while backtracking, and thus overcomes the drawback associated with the “*Increase in initial offset*” approach.

When a control burst faces a normal deflection failure and if it has an available deflection route on the previous node in the path, then the burst checks for bandwidth availability on the link from the present node to the previous node. If bandwidth is not available on that link, the burst is dropped. If it is available, it is reserved. This creates an open loop in the path reserved for the data burst. The control burst then backtracks to the previous node and tries the deflection route at that previous node. By reserving bandwidth on the backtrack link, the length of the path for the upcoming data burst is increased as it creates a loop between those nodes involved in backtracking and deflection. This will accommodate the backtrack propagation delay. In short, the purpose of this approach is to keep the data burst far enough in time behind its control burst while providing the backtracking capability to the control burst without increasing the initial offset.

3.1.3.1 Routing Procedure

The only situation when this protocol differs from the one in section 3.1.2 occurs when the Primary Control Burst (PCB) faces a normal deflection failure. The following algorithm details the routing decisions taken by the intermediate node in such situations.

- 1. Begin**
- 2. If (BURST IDENTIFIER=PCB) then**
- 3. If (both primary and deflection routes are not available) then**
- 4. If (Backup flag is set) then**
- 5. Check for bandwidth availability on the backtrack link and reserve it, if available. Convert the PCB into a BCB and send it to the**

previous node. If bandwidth on the backtrack link is unavailable, then drop the burst and send a NAK to the source.

6. **If** (Backup flag is not set) **then**
7. Drop the burst and send a NAK to the source.
8. **end-if**
9. **end-if**
10. **If** (BURST IDENTIFIER=BCB) **then**
 /* The reservation made on the primary link for this burst will not be deleted
 unlike in the previous scheme. This creates an extra delay for the data burst
 along the backtrack route */.
11. **If** (deflection route is available at the present node) **then**
 Make reservation on that link. Convert the PCB into a BCB and
 forward it on the deflection link.
12. **Else**
13. Drop the burst and send a NAK to the source.
14. **end-if**
15. **end-if**
16. **End**

3.1.4 Example for Backtrack on Deflection Failure

Consider the network in Figure 3.2. Node S is the sender and node D is the destination for a burst. For simplicity, let the propagation delay between any two nodes be equal to unity. Let the CB processing delay be negligible. The routing decisions in the network based on “Backtrack on deflection failure” scheme are explained as follows.

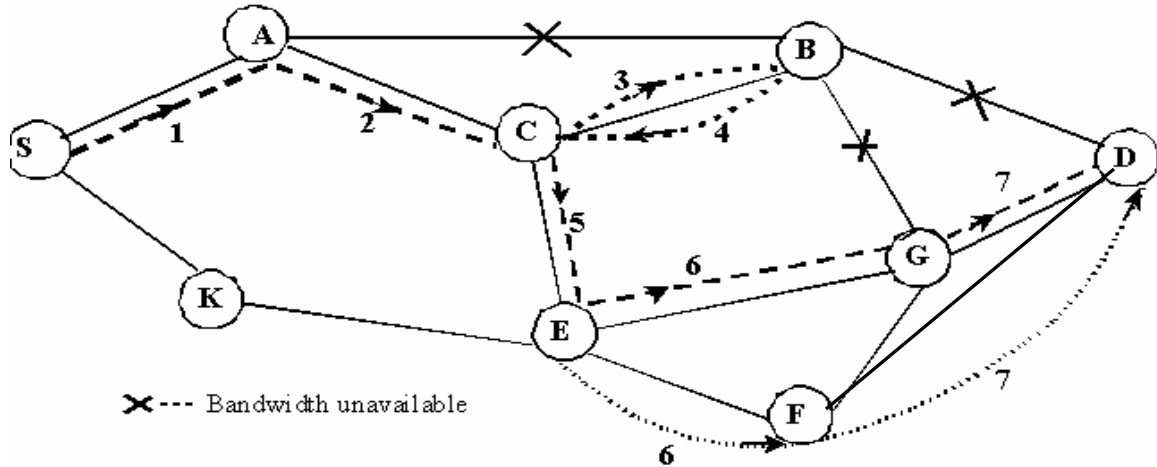


Figure 3.2 Backtrack on deflection failure

Time	0	1	2	3	4
CB	S (S---A) (S---K)	A (A-x-B) (A---C)	C (C---B) (C---E) Backup flag=1	B (B-x-D) (B-x-G) Backtrack	C (C---E)
Probe	P1(S---K)		C P2(C---E)	E P2(E---G)	G P2(G---D)

Figure 3.3 Decision table for time period (0-4)

The shortest route from S to D is along S-A-B-D. At time 0, the CB is processed by the source S. At S, both the primary link (S—A) and the deflection link (S—K) are available. Bandwidth is reserved on the primary link and the CB is forwarded on the link (S—A). A probe burst P1 is created and is forwarded on the link (S—K). At time 1, the CB reaches node A. At node A, bandwidth on the primary link (A—B) is not available. The deflection link (A—C) is available and the reservation is made on the deflection link. The CB is forwarded on link (A—C). At time 2, the CB reaches node C. At node C, both the primary link (C—B) and the deflection link (C—E) are available. Reservation is made on the primary link. The backup flag in the CB is set to unity and the CB is sent to the next node on link (C—B). A new probe P2 is created and

forwarded on the link (C—E). At time 3, the CB reaches node B. Neither the primary link (B—D) nor the deflection link (B—G) are available at node B. In such a case, the burst would be dropped in the earlier deflection schemes. In our scheme, the CB backtracks to the previous node and tries to use an alternate path to reach the destination. Since the backup flag is set to unity in the CB, node B sends the CB to the previous node C. In the mean time, the probe P2 reaches node E. At the node E, both the primary link (E—G) and the deflection link (E—F) are available. P2 is forwarded on the primary link to node G. At time 4, the CB backtracks from the node B and reaches C. The deflection link (C—E) is available and the reservation is made on that link. The CB is then sent to the node E. The probe P2 reaches node G. To understand the idea behind the use of probe bursts, let us consider the following two different cases.

Time	5		6		7
CB	E	(E---G) (E---F)	G	(G---D)	DEST D
Probe	P2(DEST D)				

Figure 3.4 Decision table for Case 1

Case 1: If the link (G—D) is available, the probe P2 is forwarded on that link to the destination D. At time 5, node E receives the CB. After processing the CB, reservation is made on the available primary link (E—G) and the CB is forwarded to the next node G. The probe P2 reaches the destination D at time 5. At time 6, node G receives the CB and sends it on the available primary link (G—D) after making the wavelength reservation. At time 7, the CB reaches the destination. The corresponding data burst will follow the reserved route **S-A-C-E-G-D**.

Case 2: If both the primary link (G—D) and the deflection link (G—B) are unavailable, the probe P2 is sent back to the previous node E. At time 5, node E receives P2 and the CB. Node E processes P2 and marks the link (E—G) as “unavailable” for the corresponding CB even though

the link is available. This is because, P2 has backtracked from G to E due to bandwidth unavailability. If the CB is sent on this link to node G, it will also have to backtrack to E due to bandwidth unavailability at G. This will cause an extra round trip delay in the reservation process. After processing the CB, the bandwidth is reserved on the deflection link (E—F). The main purpose of the probe burst is to minimize the chances of backtracking by the corresponding CB due to unavailable bandwidth and in turn reduce the total route reservation time. At time 6, node F receives the CB and sends it to the next node D after making reservation on link (F—D). At time 7, the CB reaches the destination.

Time	5		6		7
CB	E	(E---G) (E---F)	F	(F---D)	DEST D
Probe	E	P2(Drop) (E-x-G)			

Figure 3.5 Decision table for Case 2

3.2 Bidirectional Reservation on Burst Drop for Retransmission Burst

In OBS, when a control burst encounters a deflection failure, the burst is dropped and a NAK is sent to the sender. After receiving the NAK, the sender sends a new control burst for reservation of bandwidth. However, the probability of this control burst getting blocked will be the same as the probability of the previously failed control burst. Also, even though the data burst is ready to be sent into the network, it has to wait for certain amount of time equal to the offset before it can be retransmitted.

In “**Bidirectional reservation on burst drop**” scheme, when a control burst is blocked at any intermediate node, the node calculates the total time a control burst will consume to reach the sender from the present node. It also calculates the total propagation

delay for a data burst to reach the present node from the sender. This total time will be taken as the offset and tries to reserve bandwidth on the output link. If the output link is not available for that time, then the link will be reserved starting at time it is available which is greater than the offset and the offset value will be changed accordingly. The node then creates two new control bursts one for forward reservation and one for backward reservation.

With the value of the offset decided, the forward control burst is sent on the output link towards the destination. At each node on its path, the forward control burst reserves available bandwidth and tries to reach the destination. In parallel, the backward control burst will be sent to output link which leads to the source. At each intermediate node en route to the source, the backward control burst tries to reserve bandwidth on the output link leading to the previous node in its path to the source. The reservation time will be calculated based on its offset. If either of the control bursts encounters a block (while going forward or backward), that control burst is dropped and a NAK will be sent to the source as well as the destination. The NAK will inform the intermediate nodes about the unsuccessful lightpath setup and the intermediate nodes will remove any reservations made for the corresponding data burst. This way, the route between source and destination is split into two parts and the reservation is made concurrently in both directions.

In existing OBS schemes, the size of the offset is typically set to a value equal to the total control burst processing delay. But in our third scheme, for the retransmission burst, the offset will include the propagation delay in addition to the total processing delay. Propagation delay is assumed to be much greater than the processing delay (generally, propagation delay is in milliseconds and processing delay is in microseconds). All those bursts which get blocked will have this extra offset time. The first time bursts will have normal offset

which will be much smaller than the retransmitted bursts. Hence, the contention if occurs will be either among retransmission bursts or among first time bursts. This should tend to increase the probability of success for a retransmission burst. Also, the sender can transmit the data burst as soon as it receives the backward control packet by examining its offset.

3.2.1 Offset Calculation

Consider a multi-node optical network shown in Figure 3.6. Let the diameter of the network be N hops. Let T_p be the CB processing time, T_d be the propagation delay between any two nodes in the network and L be the length of the data burst. In general, $T_p \ll T_d$.

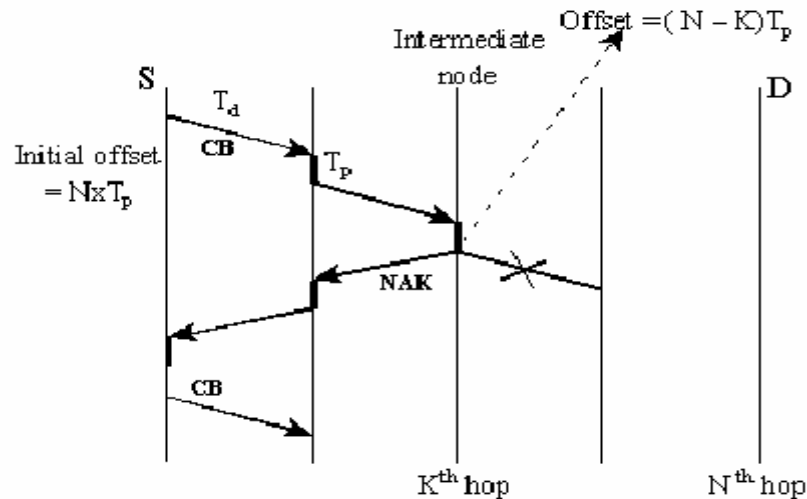


Figure 3.6 Control burst drop due to reservation failure

Let us consider a burst whose source is S and destination is D and the distance between them is N hops. The minimum offset value to guarantee that the data burst will arrive at the destination immediately after the control burst has been processed is equal to:

$$T_{\text{offset}} = N * T_p$$

At time t_0 , the source sends the CB into the network for bandwidth reservation. Let us suppose that the CB reaches an intermediate node k after k hops and that the node finishes processing the CB at time t_k . Burst drop due to normal deflection failure is shown in Figure 3.6.

The time at which the first bit of the data burst reaches the node $k = t_k + (N-K) T_p$

The time at which the last bit of the data burst leaves the node $k = t_k + (N-K) T_p + L$

Hence, the reservation is to be made for the time period $[t_k + (N-K) T_p, t_k + (N-K) T_p + L]$. If the bandwidth is unavailable for the calculated time period and if the deflection is not possible, then the control burst will be dropped by node K and a NAK will be sent to the source.

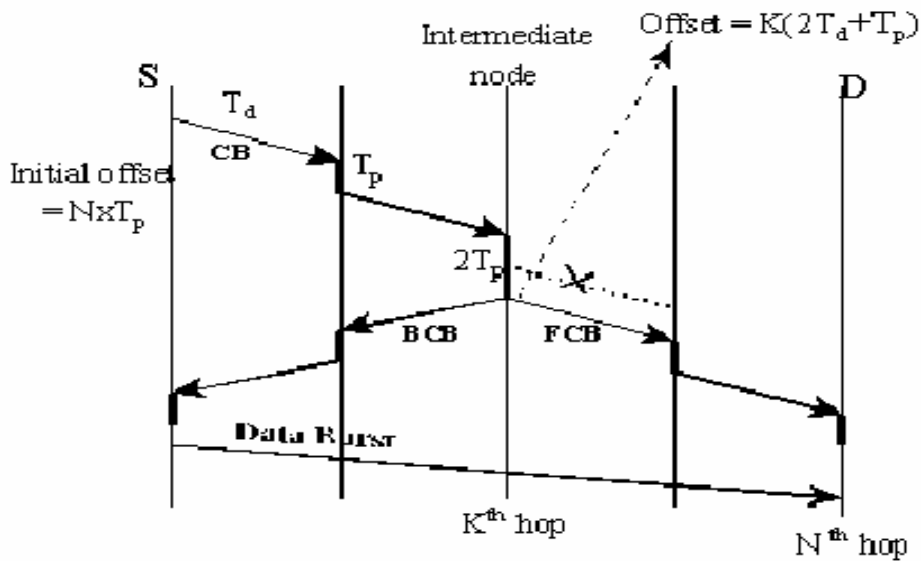


Figure 3.7 Bidirectional reservation on control burst drop

In “Bidirectional reservation on burst drop” scheme shown in Figure 3.7, a NAK will not be sent to the source. Instead, two new control bursts, a forward control burst (FCB) and a backward control burst (BCB), are created at that node (node K in our example). The FCB is sent forward to the destination and the BCB is sent backward to the source. These two control bursts try to reserve the bandwidth for the data burst whose corresponding control burst has been dropped. Before sending the new control bursts into the network, offset value is to be determined.

The offset for the newly created control bursts is calculated as follows:

The number of hops between the present node and the source is k .

The time required for the BCB to reach the source from the present node = $K (T_d + T_p)$

The time required for the DB to reach the present node from the source = $K T_d$

The total time required for the DB to reach the present node from the source = $K (2 T_d + T_p)$

This will be the new minimum offset for the new control bursts. Certain processing time is required for creating new CBs and for calculating the offset. Let us suppose that the bandwidth reservation is made at time t_{k+1} . The node will reserve the bandwidth for the time period $[t_{k+1} + K (2 T_d + T_p), t_{k+1} + K (2 T_d + T_p) + L]$. If the reservation is unavailable for the required time period, then the earliest available time greater than $[t_{k+1} + K (2 T_d + T_p)]$ is selected and the reservation is made accordingly. The new offset is increased based on the starting time of the bandwidth reservation.

After the reservation is made at node k , the FCB is sent in the forward direction and the FCB performs the bandwidth reservation for the DB along its path to the destination. The BCB backtracks towards the source reserving the bandwidth along its path. The intermediate nodes which receive these control bursts use the stored offset value for calculating the time period of bandwidth reservation.

Since the new offset in bidirectional reservation is much larger than the normal offset, the probability of such bursts getting dropped will be much less than the dropping probability for the bursts with normal offset. Thus, the probability of success for retransmitted burst is expected to be higher than the probability of success for 'first time' bursts. This is the hypothesis behind this third approach of ours. The next chapter presents the results for our proposed approaches which were obtained through simulations.

Chapter 4

Simulation & Results

In this chapter, we present the results obtained through simulations for our proposed approaches and for the schemes presented in the literature. We compare the results and show that our schemes perform better than those previously proposed in the literature. Two previous schemes “No deflection” and “Deflection” are compared with our proposed schemes which include the two versions of “Backtrack on deflection failure” and “Bidirectional reservation on burst drop”.

4.1 Simulator Setup

In order to evaluate the performance of the new contention resolution schemes, we designed a new simulator. The simulator was developed in the C language. The sample networks used in the evaluation are NSFNET and USA long haul network. A separate Poisson traffic generator has been used to provide traffic input for the network. The simulator accepts input parameters such as burst size, duration of simulation, burst arrival rate, number of wavelengths per fiber, propagation delay, processing delay, and offset choice. The lightpath establishment requests are processed in the order in which they are received. Simultaneous requests received at different nodes are processed in parallel. If multiple reservation requests are received at the same time by any node, then they are answered in random order. During the running of the simulations, if the requests are not satisfied successfully due to wavelength unavailability, then the corresponding bursts are dropped. The output of the simulation gives the burst loss rate for the specified parameters. Burst loss rate is the ratio of the total number of blocked bursts to the total number of bursts that arrived at all the nodes in the network. The

burst loss rate is also called as the blocking probability. While calculating the blocking probability, the bursts which are blocked at their respective sources without using any of the network resources are not taken into consideration.

4.2 Assumptions

The NSFNET with 14 nodes is shown in Figure 4.1 is used for simulation of the routing protocols. The USA long haul network with 28 nodes is shown in Figure 4.2. The results of NSFNET simulations are compared with the results of USA network simulations in order to study the topology-dependent performance. Using Dijkstra's algorithm, shortest paths and alternate shortest paths are calculated between all the nodes in the network. A Poisson traffic generator is used to generate traffic required for the simulations. These pre-computed paths are stored in the routing tables. The following are the assumptions used in the simulations:

- All the links in the network have the same length.
- Each fiber has the same number of wavelengths.
- No wavelength converters and no optical buffering are available.
- Burst generation is done using a Poisson distribution.
- Burst length is exponentially distributed.
- Wavelength is randomly assigned by the sender for each burst.
- Control bursts are transmitted through separate control channels.
- The control burst is never blocked due to unavailability of control wavelength.
- Network traffic is evenly distributed.

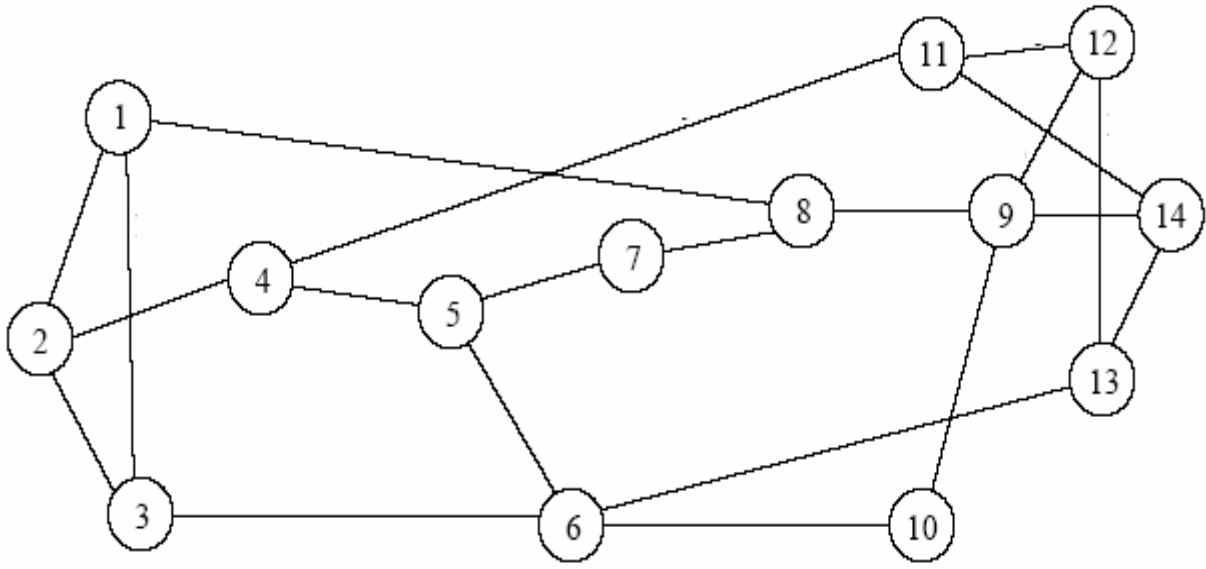


Figure 4.1 NSFNET with 14 nodes

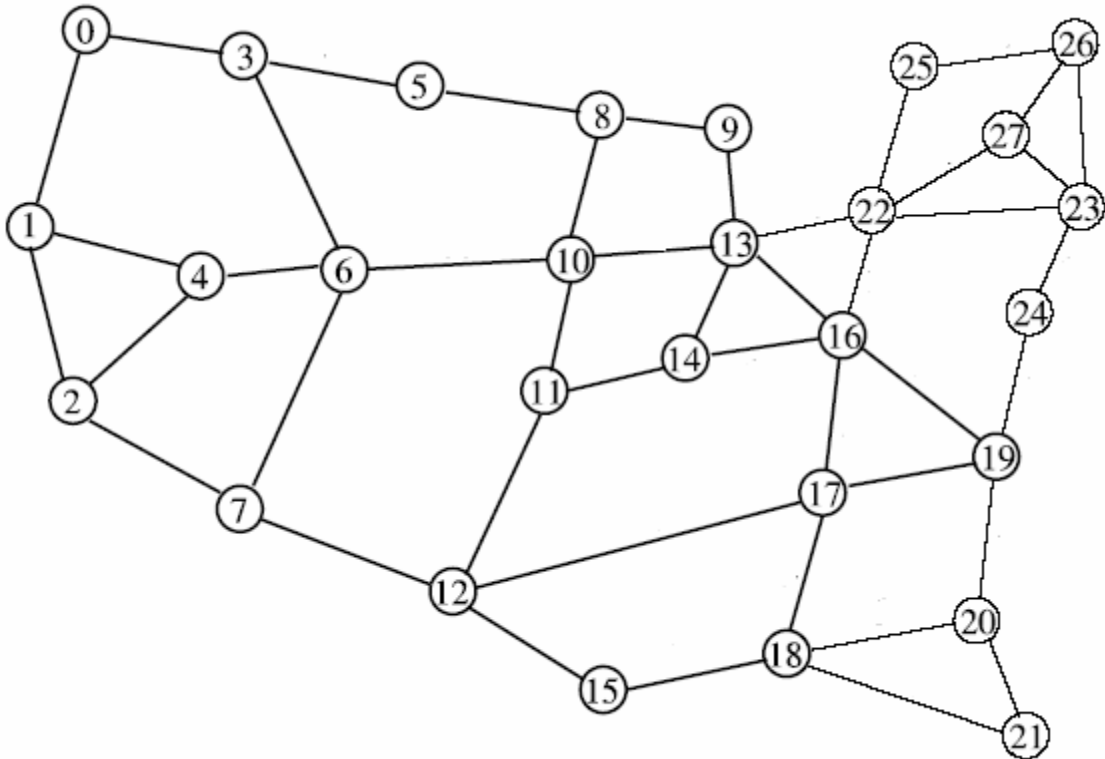


Figure 4.2 USA long haul network with 28 nodes

4.3 Simulation Parameters

Some of the parameters used in the simulation are

- Wavelengths= 4-16 per fiber
- Average burst length= 100 μ sec
- Control burst processing time= 2.5- 5 μ sec
- Switching time = 10 μ sec
- Propagation delay on a link = 0.1 to 1 millisecond.

4.4 Results

We have compared five different policies for contention resolution in OBS networks. They are

- No Defl: No deflection; Drop the contending burst.
- Defl: When a CB is blocked, deflect the contending bursts to an alternate port; Drop the burst on normal deflection failure.
- BDF-O (BDF): Backtrack on deflection failure with increase in initial **O**ffset;
- BDF-R: Backtrack on deflection failure with open loop **R**eservation.
- BDF-BR: Bidirectional reservation on burst drop for retransmission

Figure 4.3 plots the total packet loss probability versus the load for four different contention resolution policies. The offered load ranges from low to medium. We observe that our proposed BDF schemes perform better than Defl and No Defl schemes at low loads. As the load increases, the performance gap between our schemes and existing ones decreases. A logical explanation would be that, when a burst faces a normal deflection failure, both the Defl and No Defl schemes drop that burst. Whereas in our BDF schemes, the burst will try to

backtrack and utilize any free links available for successful completion of bandwidth reservation. Also, at low loads, there will be larger number of free links available. As the load increases, the number of free links decreases. This explains the reduction in the performance gap between our schemes and the other schemes. We also observe that BDF-O performs better than BDF-R in this load range. This is because, in BDF-O, the offset size of the bursts is increased to accommodate the backtrack delay and hence they will not face any problems while backtracking. On the other hand, in BDF-R scheme, the bursts rely on the availability of the backtrack link and the bursts will backtrack only if bandwidth reservation is successful on the backtrack link. Otherwise the bursts will get dropped.

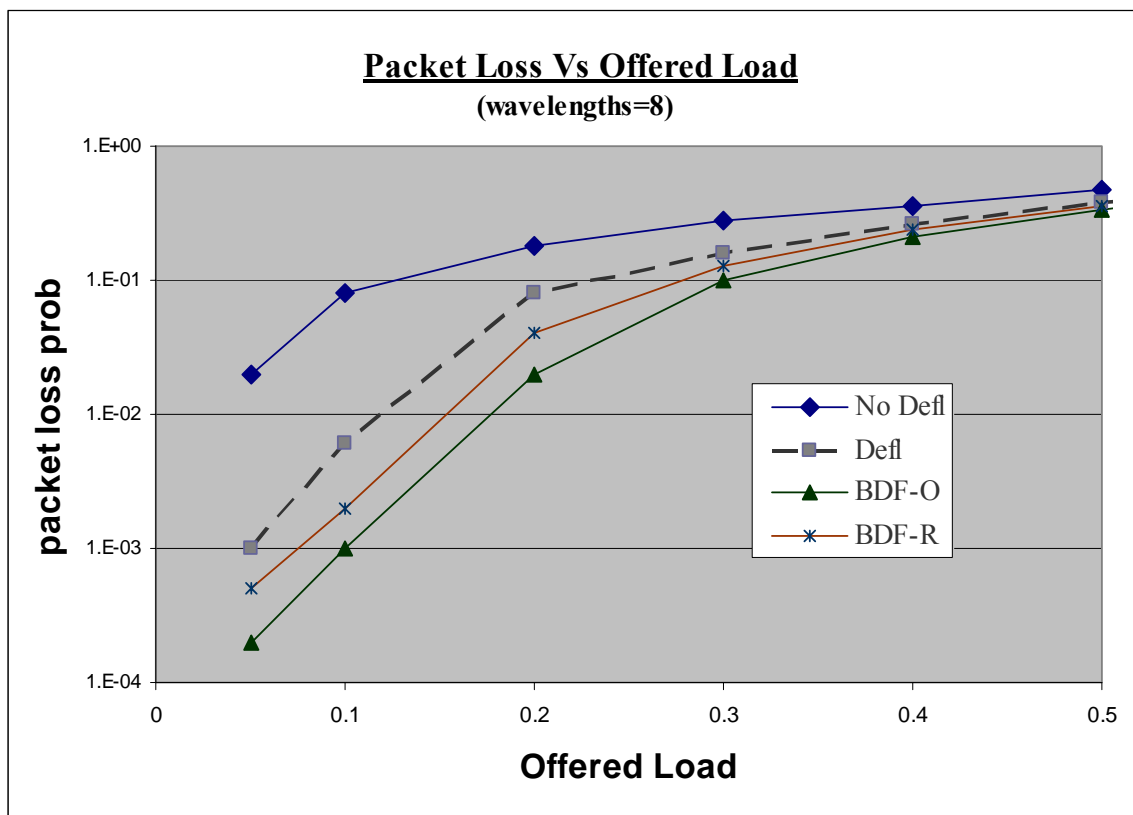


Figure 4.3 Packet loss probability at low load

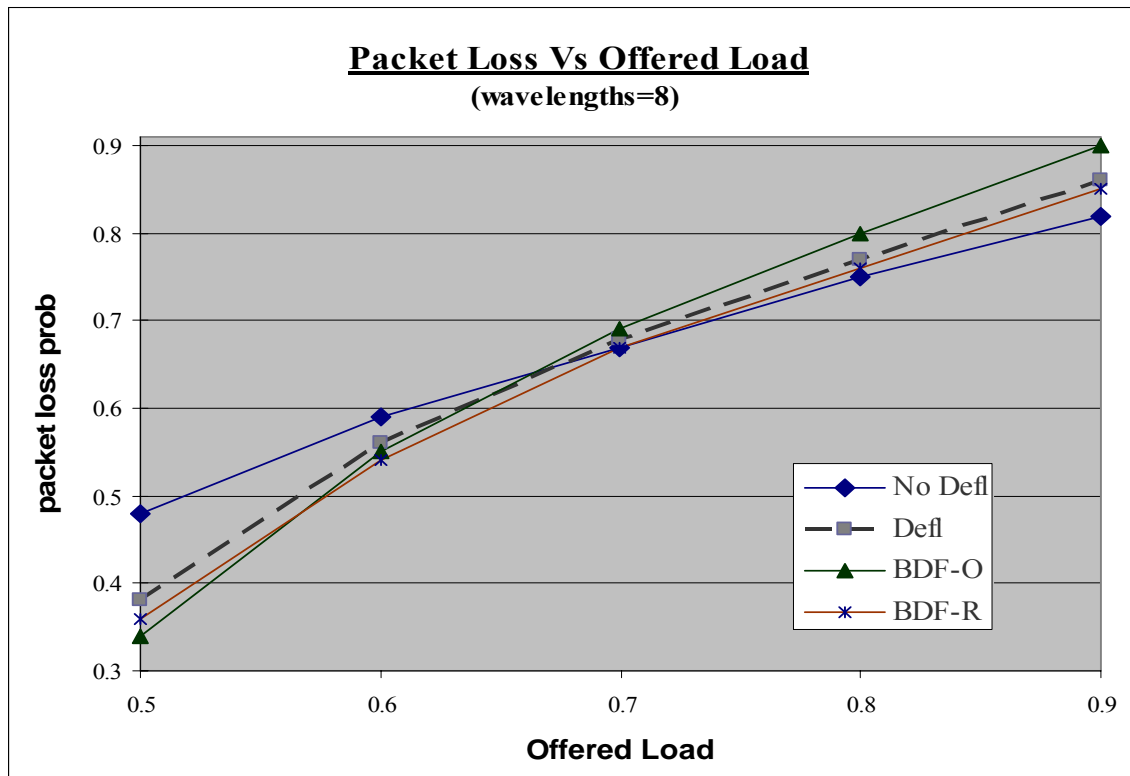


Figure 4.4 Packet loss probability at high load

Figure 4.4 shows the packet loss performance at high loads. We observe that the performance of BDF schemes deteriorates as the load increases from medium to high. The conventional schemes perform better than BDF schemes at high loads. The reason for this can be explained as follows. As the load increases, more and more bursts will encounter normal deflection failure and all those bursts will try to backtrack and use any available resources. This will create more traffic and more contentions and will tend to increase the blocking probability. Defl scheme is a subset of BDF scheme in terms of routing options. The contentions created through deflection will be fewer when compared to those of BDF. In case of No Defl scheme, when a reservation failure occurs, the bursts are dropped instantly and no new contentions will be created due to the dropped bursts. Hence the No Defl scheme performs

better than all other schemes at high loads. We can also observe that the performance of BDF-R is better than that of BDF-O at high loads. This is because, bursts in BDF-R have a greater probability of getting blocked before backtracking due to the unavailability of backtrack link bandwidth. All such bursts which fail to backtrack will be dropped and no further resources will be used. Due to this, the overall contention creation will be less in BDF-R when compared to that of BDF-O. This is the reason for the change in their performance.

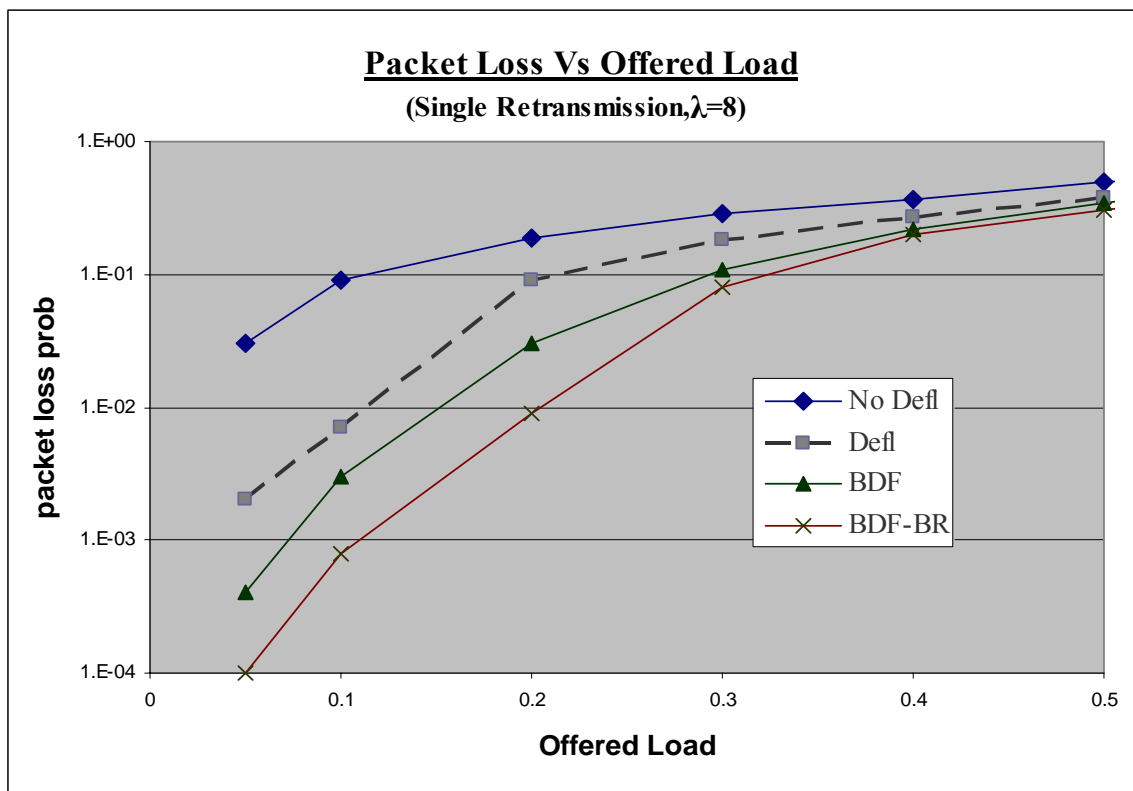


Figure 4.5 Packet loss probability at low load

Figure 4.5 and Figure 4.6 show the packet loss probability at various loads. Here, the blocked bursts will have a chance for a single retransmission. The BDF-BR scheme is an extension of BDF scheme and the actual bidirectional reservation process comes into effect only at the time of bandwidth reservation for the retransmitted bursts. The BDF-O outperforms

BDF-R in most of the situations as observed from the previous plots. Therefore the BDF-O scheme is considered as the BDF scheme in this simulation. In the plot, we observe that BDF-BR performs better than all the other schemes at all the load levels. This can be explained as follows.

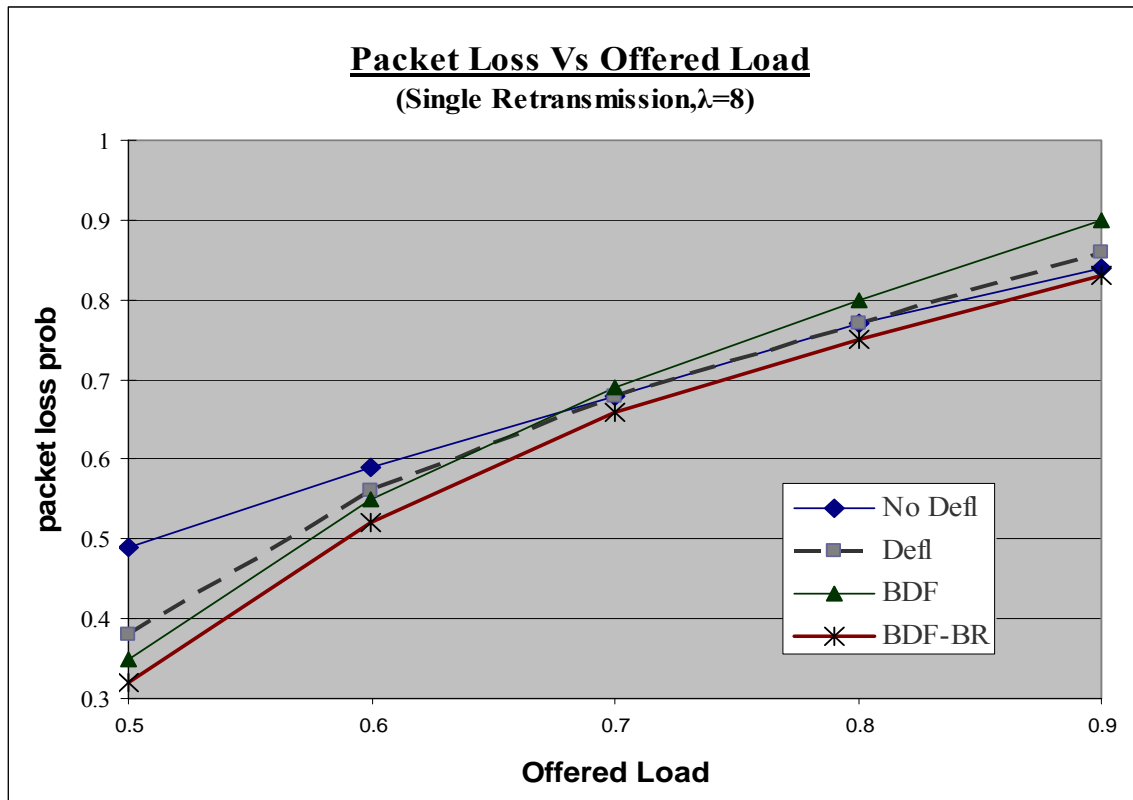


Figure 4.6 Packet loss probability at high load

In Defl, No Defl and BDF schemes, when a burst is dropped, a NAK is sent to its source. The source on receiving the NAK will retransmit the control burst after waiting for some random time. The Offset for this retransmission burst will be the same as that of the previously failed burst. But in BDF-BR, when a burst is dropped, the bandwidth reservation for the retransmitted burst is done immediately with high offset value which includes the propagation delay. Reservation is done in parallel in both forward and backward directions

from the node where the control burst is dropped. We know that bursts with high offset will have less probability of getting blocked than bursts with low offset. Since the retransmitted bursts in BDF-BR have larger offsets, the probability of success for these bursts is high. We also observe that the performance of BDF-BR drops as the load increases. This is because of the fact that BDF-BR involves BDF scheme for the “first time” bursts. We know that the performance of BDF scheme degrades at high loads due to extensive contention creation. This will have an effect on the overall performance of the BDF-BR scheme.

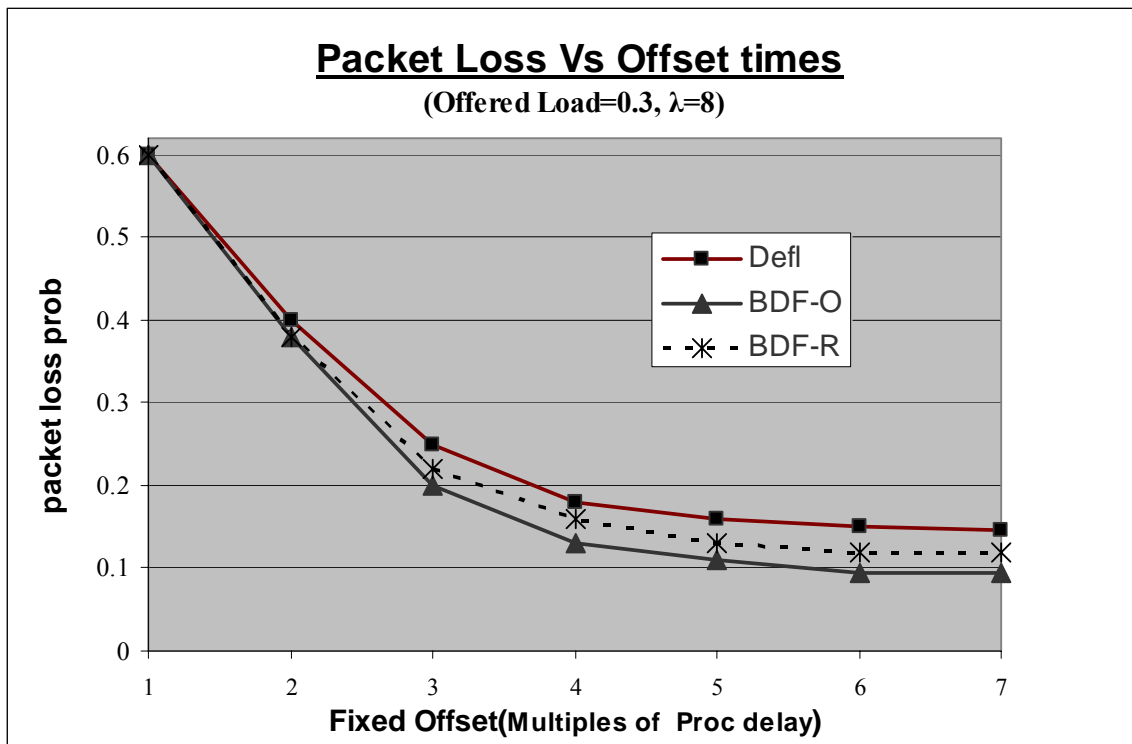


Figure 4.7 Packet loss probability versus fixed offset

Figure 4.7 shows the packet loss probability versus fixed offset. The offset value for all the bursts generated is fixed in this simulation. The distance between the source and the destination of a burst will be referred as its hop count. Generally, the offset for each burst depends on its hop count. The offset is calculated in terms of the processing delay (hop count).

Typically, in OBS, the offset of a burst is equal to hop count times the processing delay. In the plot, we observe that, at low offsets, the blocking probability is very high. This is because, at some point in time all the control bursts whose hop count is greater than the fixed offset value will fail to reserve bandwidth before their corresponding data bursts arrive. The data bursts will reach the nodes where the reservation is not yet made and will be dropped. As the fixed offset value increases to 5 and above, the loss rate decreases drastically and becomes stable for all the three schemes. This is because, the bursts have their offset set to values greater than the diameter of the network which is 4 for NSFNET. Under these conditions, control bursts are more likely to reach their respective destinations before their corresponding data bursts. This will reduce packet loss due to insufficient offset. Also, as offset size increases further there will not be any change in the loss rate due to the fact that network resources are limited. We also observe that BDF-O performs better than BDF-R and BDF-R performs better than Defl.

Figure 4.8 shows the average hops traveled versus the intended number of hops while considering only the successful bursts. The intended number of hops for a burst is nothing but the distance between source and destination of the burst. This is referred as hop count. When the intended hop count is one, the average hops traveled remains around one for all the schemes. When the hop count increases, the number of hops traveled increases to a larger extent in the BDF and Defl schemes than in No Defl scheme. Clearly, the bursts with larger hop counts have higher chance of getting blocked than bursts with smaller hop counts. As the hop count increases, the chance of encountering a reservation failure increases. When the bursts face such situations, they will either deflect or they will backtrack. Bursts in the Defl and BDF schemes tend to take alternate routes when they face a reservation failure and such

routes are not always shortest routes. Hence, the average number of hops traveled by the bursts is nearly five in case of BDF when the intended hop count is equal to 4.

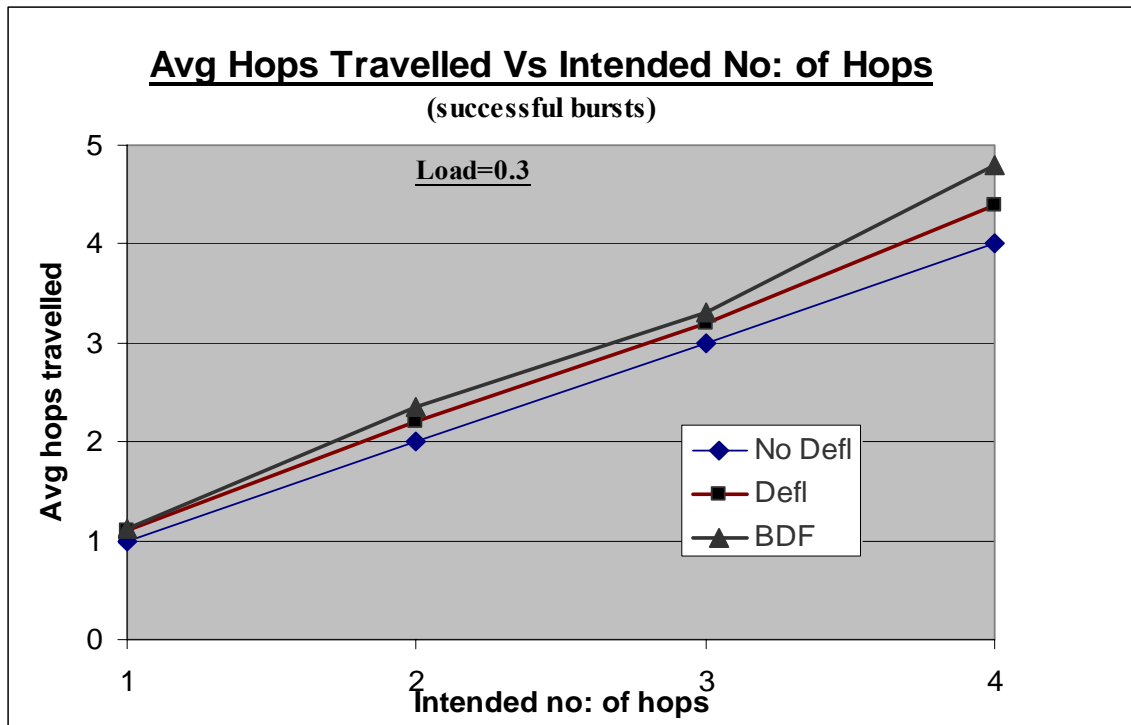


Figure 4.8 Average number of hops traveled verses intended number of hops

Figure 4.9 shows the average hop count versus load. In this plot, the average hop count implies the average number of hops traveled while considering both successful and unsuccessful bursts. We observe that the average number hops traveled decreases as the load increases in case of the No Defl scheme. A logical explanation is that, as the load increases, the blocking probability increases and more and more bursts will be blocked and dropped before reaching their respective destinations.

The behavior of Defl, BDF-O and BDF-R schemes is explained as follows. At very low loads (such as 0.1), most of the bursts are successful without using deflection or backtracking. Hence the number of hops traveled by such a burst will be equal to its intended

number of hops which is the shortest route hop count. As the load increases to 0.2, the number of bursts which encounter primary reservation failure increases. Since the traffic is still not so heavy, most of these bursts will either backtrack or will deflect to other alternate routes (which are not always the shortest routes) and will be successful. Hence the average hop count in these three schemes increases as the load increases from 0.1 to 0.2. As the load increase further from 0.2 to 0.5, the average hop count decreases gradually and as the load approaches 0.5, the average hop count value decreases more drastically. This behavior is explained next.

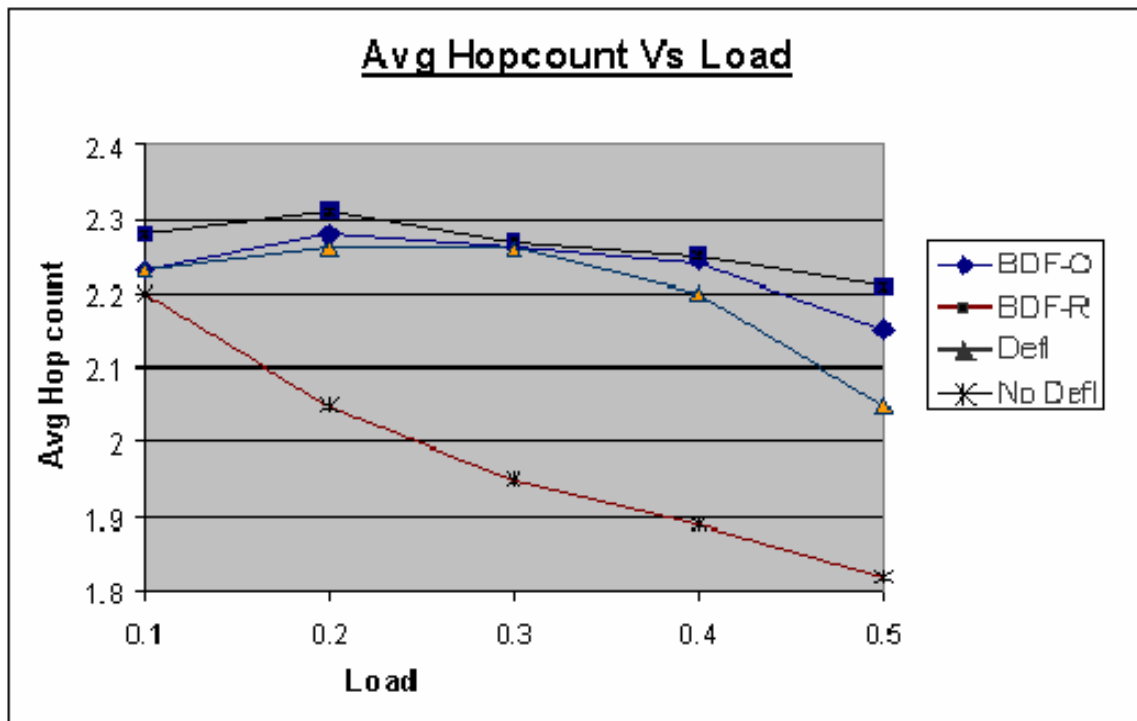


Figure 4.9 Average hops traveled versus load

As the load increases, the number of bursts which either deflect or backtrack will increase. More of these bursts will fail to reach their destinations. The average number of hops traveled by the unsuccessful bursts will be less than the average number of hops traveled by successful bursts. As the load increases, the number of unsuccessful bursts also increases and this leads to

a gradual decrease in the average hop count. BDF schemes have higher hop count than other schemes because of the fact that bursts in the BDF schemes travel larger number of hops due to backtracking than the bursts in other schemes. The BDF-R has higher hop count than BDF-O at all load levels. The reason for this is that a successfully backtracked data burst in BDF-R travels one extra hop than such burst in BDF-O due to backtrack link reservation in BDF-R.

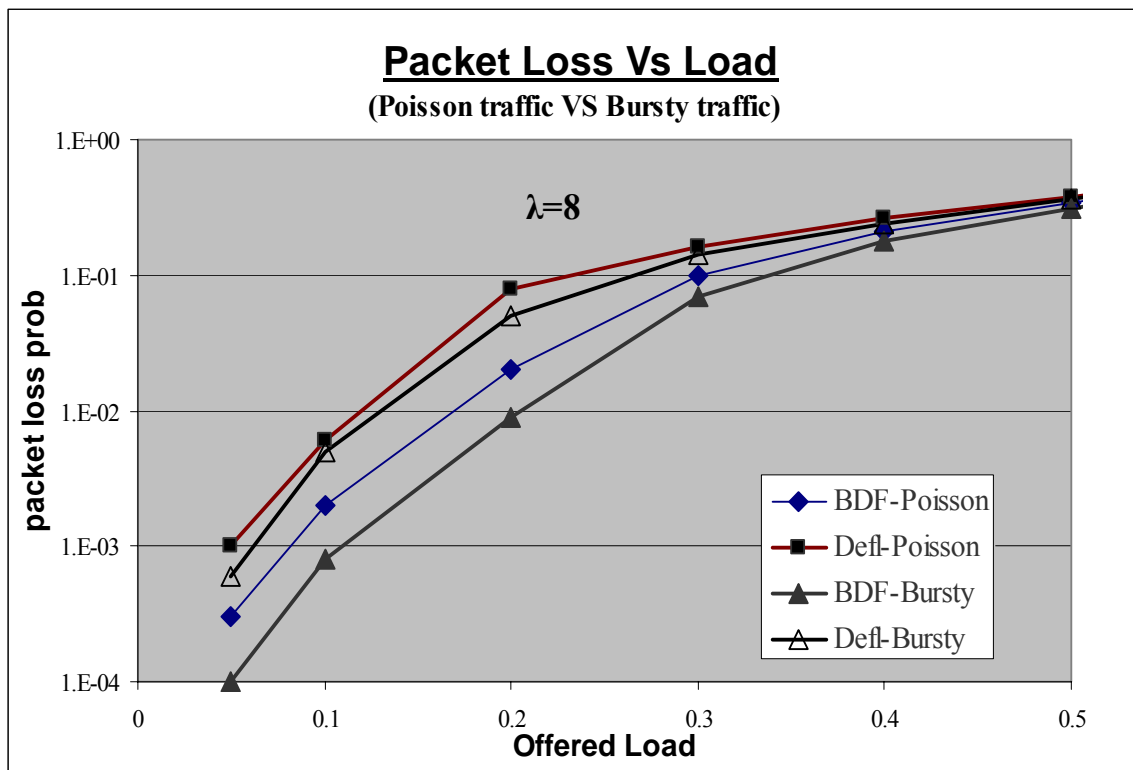


Figure 4.10 Poisson traffic versus bursty traffic (low loads)

Figures 4.10 and 4.11 show packet loss probability versus offered load. A poisson model and a bursty traffic model are considered for traffic generation. Internet traffic is typically bursty in nature [22]. In the bursty traffic model, traffic generation is not continuous at any node. At each node, traffic is generated for some duration of time called the busy time followed by some duration of time called the idle time where no traffic is generated. In the

overall traffic generation process, the busy and idle times repeat alternately. At any time, not all the nodes in the network will be busy or idle.

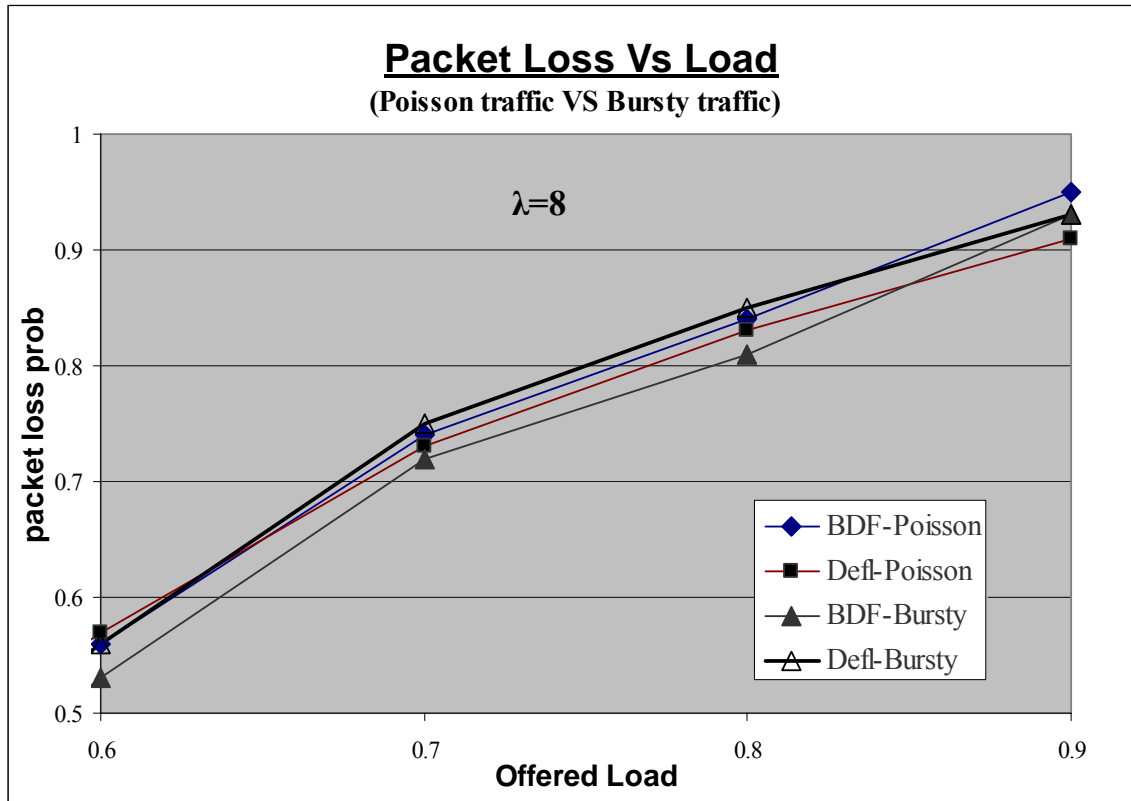


Figure 4.11 Poisson traffic versus bursty traffic (high loads)

In this plot, the performance of the Defl and BDF (BDF-O) schemes are compared for both the poisson and the bursty traffic models. Both these schemes perform better with bursty traffic than with poisson traffic. A logical explanation for this is that when bursts get blocked at nodes which are busy, they will either deflect or backtrack to the nodes which are idle. The probability of the bursts getting blocked is higher at busy nodes than at idle nodes. Hence, more such bursts will be successful with bursty traffic than with poisson traffic where (statistically) all the nodes will have similar traffic at all times. We also observe that the BDF scheme outperforms the Defl scheme at all levels of load and traffic.

Figure 4.12 plots link utilization versus offered load. Link utilization is calculated as the ratio of the traffic load on the link to the link's capacity. It can also be calculated as the ratio of the link busy time to the total time. From the figure, we notice that the BDF scheme has higher link utilization than the other schemes. This can be explained as follows.

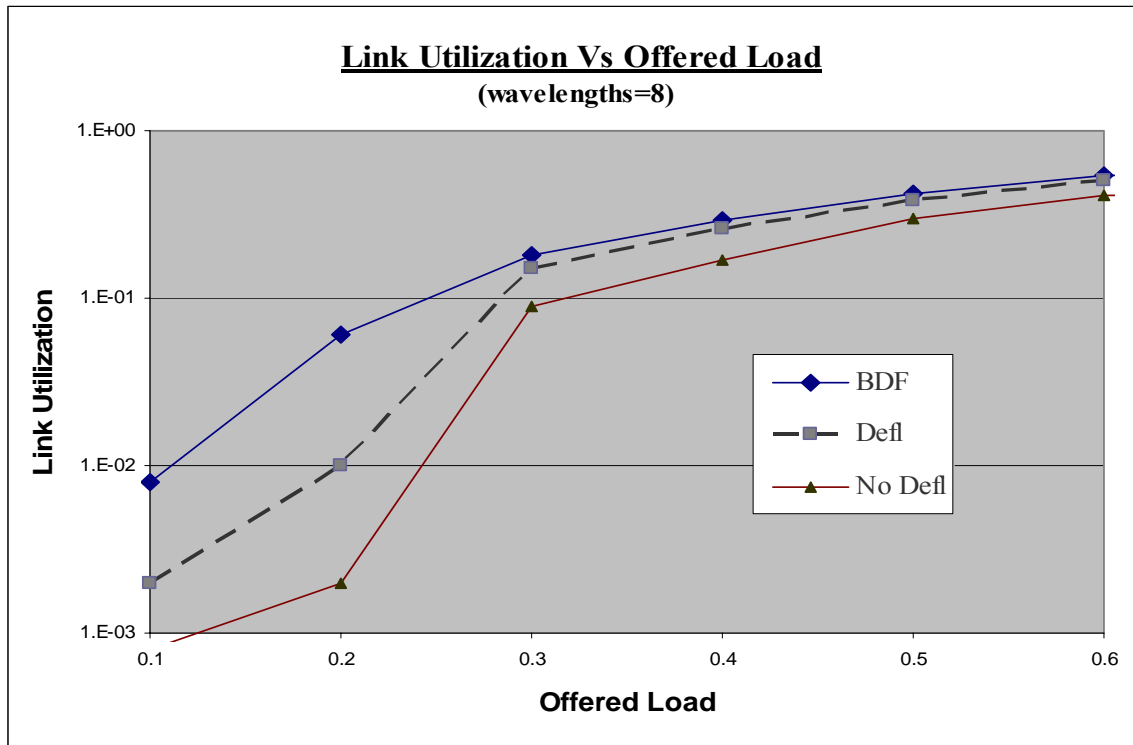


Figure 4.12 Link utilization versus load

When a burst encounters a normal deflection failure, it is dropped in both the Defl and No Defl schemes. But in the BDF scheme, such burst will try to utilize any available resources for bandwidth reservation by using the backtracking method without getting dropped. This is the reason for having better overall link utilization with BDF. Also, at normal loads, link utilization and blocking probability are inversely proportional.

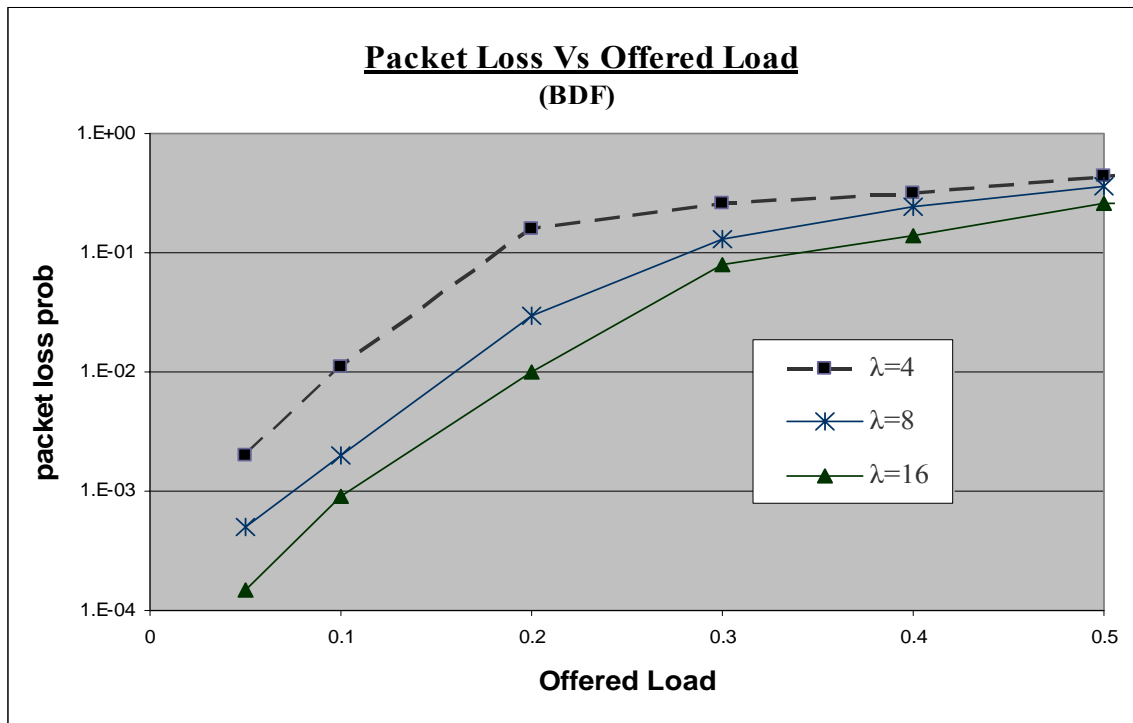


Figure 4.13 Packet loss rate with different number of wavelengths/fiber

Figure 4.13 shows the packet loss rate versus the load for different number of wavelengths per fiber. The BDF scheme is used in these simulations. We observe that packet loss decreases as the number of wavelengths per fiber increases. An explanation for this is that the probability of bursts contending for the same wavelength decreases as the number of wavelengths increases. The difference in packet loss rate between the three schemes decreases as the load increases. The reason for this is that the amount of contention will increase as the load increases and more bursts will get dropped. Thus the effect of having more wavelengths decreases as the load increases further.

Figure 4.14 plots average throughput versus offered load. Throughput is calculated as the ratio of the total number of packets successful to the total number of packets sent during a fixed time frame. From the figure, we notice that the BDF-O scheme performs better than the

other schemes. This is attributed to the backtracking capability of the bursts in BDF-O scheme and to the increased initial offset. All the schemes have similar throughput at low load and as the load increases the performance gap between our schemes and the existing schemes widens. At higher load, the performance gap decreases. The reason for this behavior can be explained as follows. At low loads, the success rate for the bursts in all the schemes is high and the blocking probability is very low due to availability of bandwidth. At intermediate loads, the bursts in our schemes utilize the backtracking capability when normal deflection failure occurs. All such bursts will get an extra chance for reaching the destination. But the bursts in the other schemes will be dropped on encountering normal deflection failure. The network will be congested at high load. More bursts will get blocked and the effect of backtracking decreases due to congested backtrack nodes. Hence the performance drops.

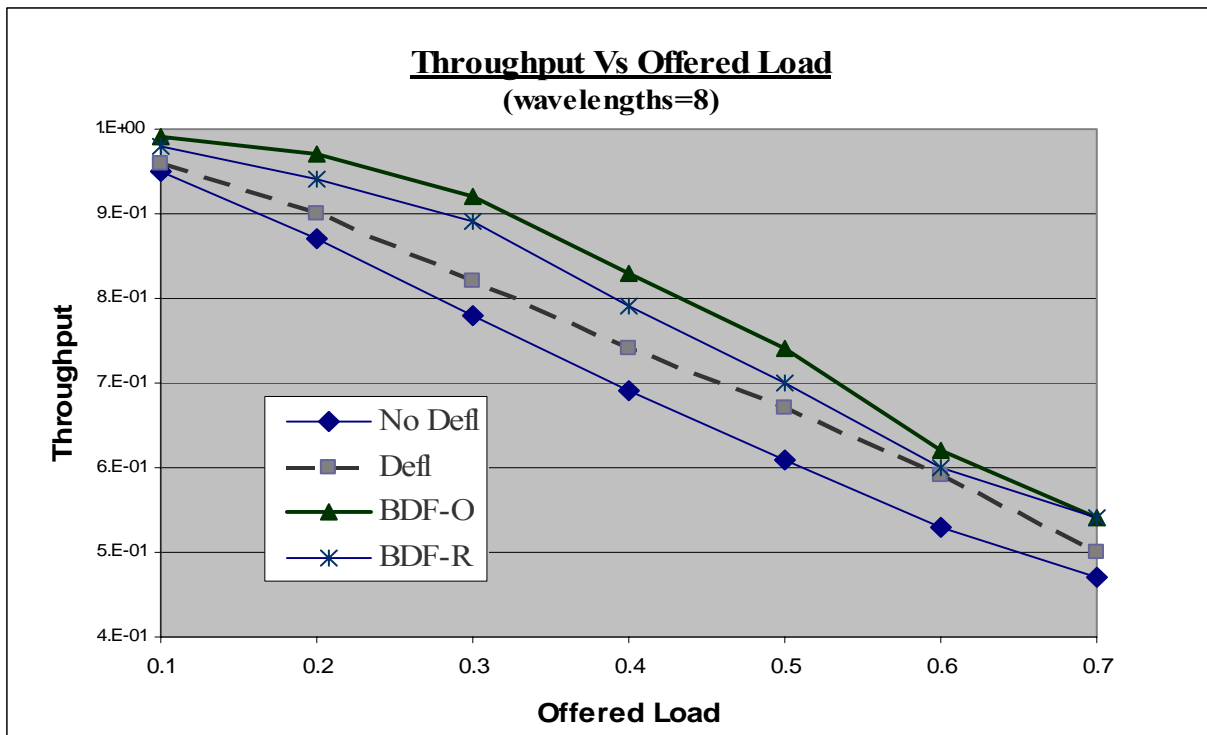


Figure 4.14 Throughput Vs Offered load

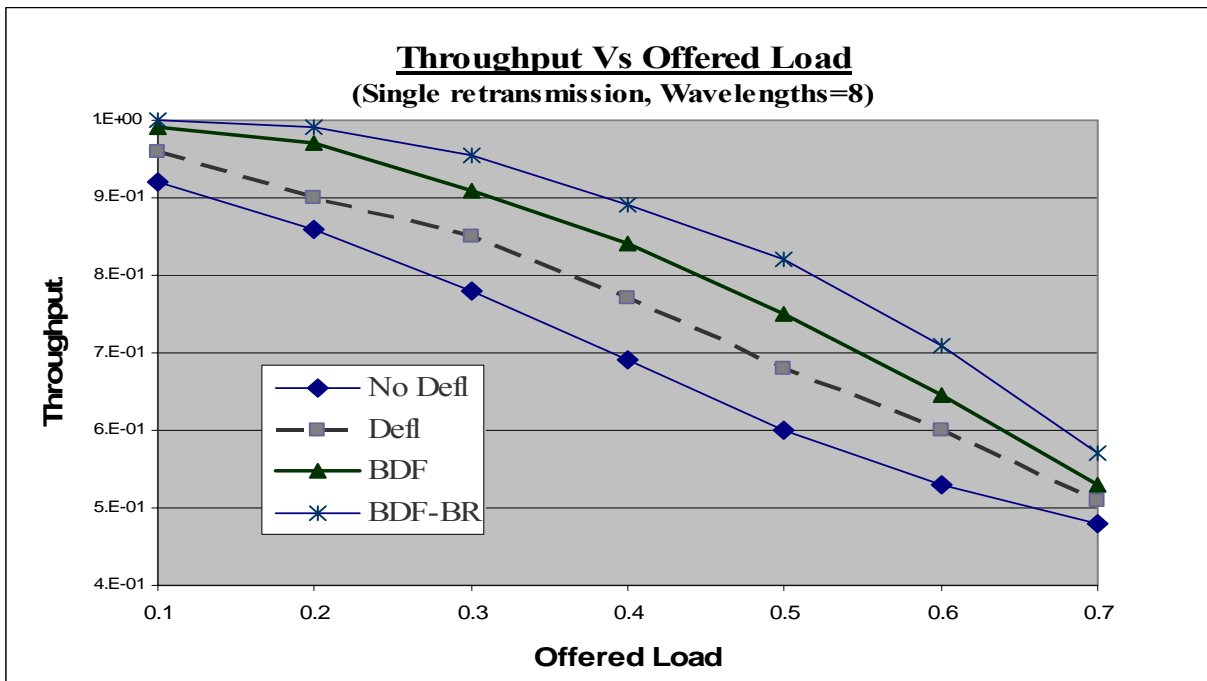


Figure 4.15 Throughput Vs Offered load (Single retransmission)

Figure 4.15 plots the throughput versus the offered load. The BDF-BR scheme is compared with the other schemes. Each blocked burst in this simulation will be retransmitted once. From the figure, we notice that the BDF-BR scheme performs better than all other schemes at all load levels. The BDF-BR and the BDF schemes perform similarly at low loads. The reason for this is that the BDF scheme is a subset of BDF-BR scheme. The bidirectional reservation takes place only when a burst is completely blocked. As the load increases, more bursts will get blocked. In the schemes other than BDF-BR, during retransmission of such bursts, the offset has the same value as all the other bursts. Hence the probability of these bursts getting blocked will be same as that for first time bursts. However, in BDF-BR, the offset for the retransmitted bursts will be increased to accommodate the propagation delay as explained in the previous chapter. These bursts will have high probability of being successful because of the increased offset value. This explains the behavior of the BDF-BR scheme.

Figure 4.16 plots the average transmission count versus the offered load. The transmission count is the total number of times a burst is transmitted by its sender before it reaches the destination. A blocked burst will be retransmitted until it becomes successful. The average transmission count (ATC) is the average of the transmission counts of all successful bursts during the entire simulation time.

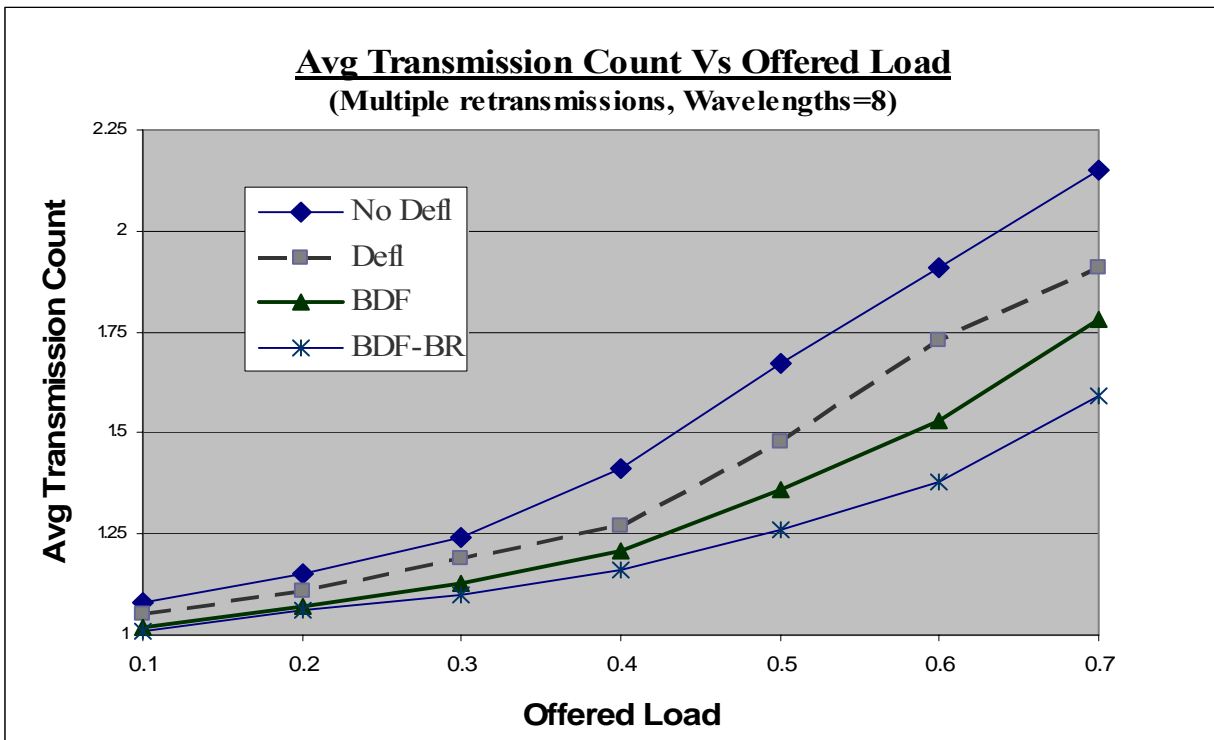


Figure 4.16 Average transmission count Vs Offered load

In the figure, we notice that the BDF-BR scheme has the least transmission count at all load levels. The bursts in the other schemes have higher transmission count and the ATC increases rapidly as the load increases from medium to high. The reason for this behavior is that the probability of success decreases as the load increases. When a blocked burst is retransmitted, it has the same offset as that of the first time burst. The blocking probability of the retransmitted burst will remain unchanged. At higher loads, the blocking probability is high and this causes

multiple retransmissions before a burst gets successful. In case of BDF-BR, the retransmitted bursts will have larger offset values and the blocking probability of these bursts decreases and the chance of getting successful increases. This is the reason for low ATC in BDF-BR.

Figure 4.17 shows the packet loss rate versus the offered load in the USA long haul network. The overall behavior of the schemes presented is similar to that in the case of NSFNET. We observe that the packet loss rate for all the schemes is slightly higher in this network than in NSFNET. This is because, the average node degree in the Long haul network is less than the average node degree in NSFNET. A node which has a degree of three or more will have deflection capability. Also, there are more nodes in the USA network with no deflection and backtracking capability than in the NSFNET. This explains the performance drop in this network.

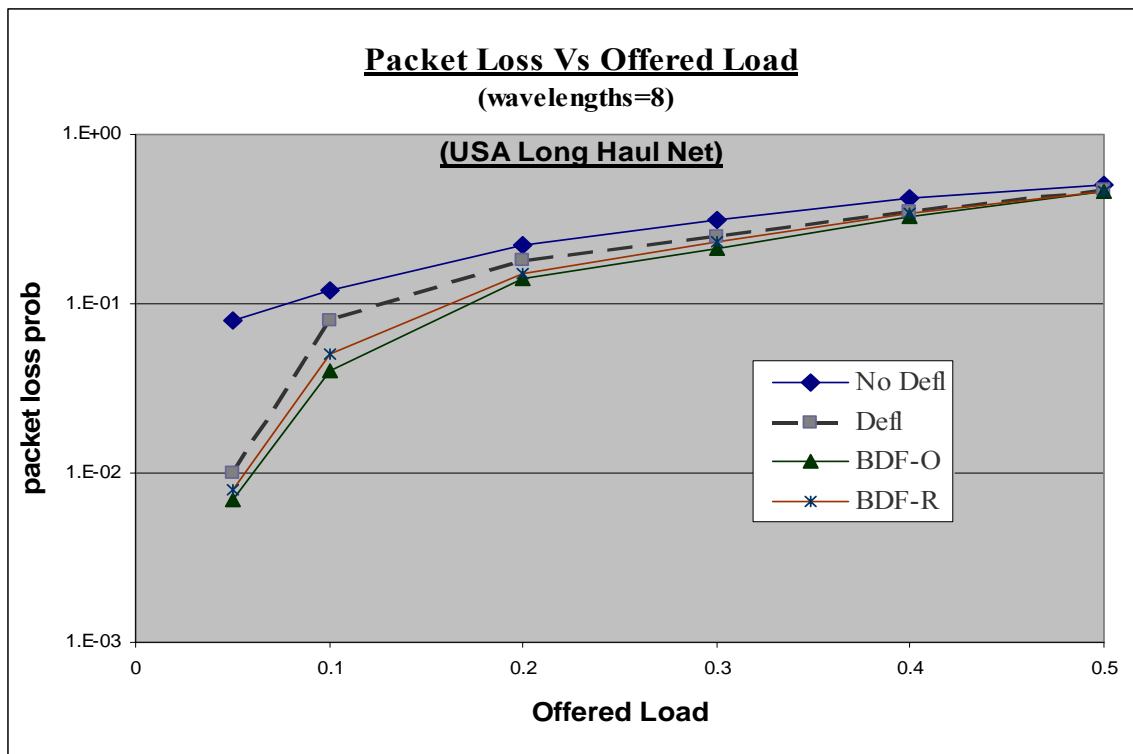


Figure 4.17 Packet loss versus load (USA long haul network)

Chapter 5

Conclusion and Future Scope

Optical burst switching lies somewhere between optical packet switching and optical circuit switching. Data and control information are sent through different wavelength channels in an OBS system. In this report, we reviewed different switching techniques and examined the optical burst switched network architecture and the different reservation policies associated with it. We also summarized the art of contention resolution and the resolution schemes including optical buffering, wavelength conversion, and deflection routing (or hot-potato routing).

We introduced the novel concept of backtracking for contention resolution in optical burst switched networks, and proposed two different backtracking policies one with and one without increase in initial offset. Through simulations we have shown that our proposed backtracking policies perform better than the standard deflection policy, and offer the best performance at medium loads. However, policies which incorporate deflection are not as effective at high loads. For that reason, we have also proposed the concept of bidirectional reservation on burst drop. The primary objective of this method has been to provide contention-free bandwidth reservation (as much as possible) for retransmitted bursts. In this scheme, when a control burst is blocked because of bandwidth unavailability, two new control bursts are created and the reservation is done concurrently in both forward and backward directions from the blocked node. The offset value for these bursts is much greater than that of normal bursts and this approach can potentially reduce the blocking probability considerably for retransmitted bursts. Our simulation results support this hypothesis. They show that

the blocking probability with bidirectional reservation is much smaller than that obtained with any other scheme.

Areas for future work include utilizing the wavelength conversion and optical buffering to provide more options for contention resolution. Dynamic load balanced routing could be a welcome addition to the new policies. While we have considered only single-hop backtracking in this work, future work can incorporate multi-hop backtracking and backup wavelength reservation. Implementing quality of service (QOS) with these schemes is another interesting area for future research.

Bibliography

- [1] C. Qiao and M. Yoo, "Optical Burst Switching (OBS) - A New Paradigm for an Optical Internet," *Journal of High Speed Networks*, vol. 8, no.1, pp. 69-84, Jan. 1999.
- [2] L. Xu, H.G. Perros, and G Rouskas, "Techniques for optical packet switching and optical burst switching" *IEEE Communications Magazine*, pp.136-142, Jan. 2001.
- [3] International Engineering Consortium, "The Direction of Optical Networking Market".
- [4] Mukherjee Biswanath, WDM Networking Evolution, 26.6.1997 [Referenced 10.1998].
- [5] Adar Shtainhart, Ronen Segal, Aviad Tsherniak, "Wavelength Division Multiplexing".
- [6] S. Yao, B. Mukherjee, and S. Dixit, "Advances in Photonic Packet Switching: An Overview," *IEEE Commun. Mag.*, Feb. 2000, pp. 84-94.
- [7] Sateesh Chandra, "Survivable Multicasting in WDM Optical Networks", A Master's Thesis, Department of Electrical Engineering, Louisiana State University, 2004.
- [8] B. Mukherjee, *Optical Communication Networks*, McGraw-Hill, New York NY, 1997.
- [9] F. Farahmand, Q. Zhand, J. P. Jue, "A Feedback-Based Contention Avoidance Mechanism for Optical Burst Switching Networks", 2004.
- [10] R. Ramaswami, K.N. Sivarajan, "Routing and Wavelength assignment in All-Optical Networks," *IEEE/ACM Transactions on Networking*, October 1995, vol.3, pp. 489-500.
- [11] M. Yoo and C. Qiao, "A Novel Switching Paradigm for Buffer-Less WDM Networks," *IEEE Communications Magazine*, 1999.
- [12] 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.
- [13] S. Verma, H. Chaskar, and R. Ravikanth, "Optical burst switching: A viable solution for Terabit IP backbone.," *IEEE Network*, pages 48-53, Nov/Dec 2000.
- [14] M. Yoo and C. Qiao, "Just-Enough-Time (JET): A High Speed Protocol for Bursty Traffic in Optical Networks" In proceeding of *IEEE/LEOS Conf. on Technologies for a Global Information Infrastructure*, pp. 26-27, Aug. 1997.
- [15] L. Xu, H.G. Perros, and G Rouskas, "Techniques for optical packet switching and optical burst switching" *IEEE Communications Magazine*, pp.136-142, Jan. 2001.

- [16] Jing Teng, and George N. Rouskas, "A Comparison of the JIT, JET, and Horizon Wavelength Reservation Schemes on A Single OBS Node", 2003.
- [17] Vinod M. Vokkarane and Jason P. Jue, "Burst Segmentation: an Approach for Reducing Packet Loss in Optical Burst-Switched Networks," Special Issue on Engineering the Next Generation Optical Internet, SPIE/Kluwer, Optical Networks, vol. 4, no. 6, pp. 81-89, Nov./Dec., 2003.
- [18] S. Lee, K. Sriram, H. Kim, J. Song, "Contention-based Limited Deflection Routing in OBS Networks" in proceedings of Globecom 2003, pages 2633-2637.
- [19] C. Qiao and M. Yoo, "Choices, Features and Issues in Optical Burst Switching", Optical Networks Magazine, Vol. 1, No. 2, pp. 36-44, 2000.
- [20] S. Yao, B. Mukherjee, and S. Dixit, "Advances in Photonic Packet Switching: An Overview", IEEE Commun. Mag., Feb. 2000, pp. 84-94.
- [21] M. Listanti, V. Eramo, and R. Sabella, "Architectural and technological issues for future optical Internet networks", IEEE Communications Magazine, vol. 38, issue 9, pg: 82-92, Sept. 2000.
- [22] A. S. Acampora and I. A. Shah, "Multihop Lightwave Networks: A Comparison of Store-and-Forward and Hot-Potato Routing", IEEE Transactions on Communications, vol. 40, no. 6, pp. 1082-1090, June 1992.
- [23] F. Forghieri, A. Bononi, and P. R. Prucnal, "Analysis and Comparison of Hot-Potato and Single-Buffer Deflection Routing in Very High Bit Rate Optical Mesh Networks", IEEE Transactions on Communications, vol. 43, no. 1, pp. 88-98, Jan. 1995.
- [24] A. Bononi, G. A. Castanon, and O. K. Tonguz, "Analysis of Hot-Potato Optical Networks with Wavelength Conversion", IEEE Journal of Lightwave Technology, vol.17, no. 4, pp. 525-534, April 1999.
- [25] Sungchang Kim, Namook Kim, Minho Kang, "Contention Resolution for Optical Burst Switching Networks Using Alternative Routing", Proceedings IEEE, ICC 2002, New York, NY, April-May 2002.
- [26] Ching-Fang Hsu, Te-Lung Liu, Nen-Fu Huang, "Performance Analysis of Deflection Routing in Optical Burst-Switched Networks", Proceedings IEEE, INFOCOM 2002, New York, NY, June 2002.
- [27] K. Lu, J. P. Jue, G. Xiao, and I. Chlamtac, "Intermediate-Node Initiated Reservation (IIR): A New Signaling Scheme for Wavelength-Routed Networks," IEEE Journal on Selected Areas in Communications, vol. 21, no. 8, pp. 1285-1294, October 2003.

VITA

Kishore Reddy Koduru was born in India in 1979. He completed his high school from LRG School, Anantapur, India, in 1995 and then obtained his Bachelor of Engineering degree in computer science and engineering from BIET, Davangere, India, in 2002. In fall 2002, he arrived at Louisiana State University to pursue higher education. He is expected to graduate with his Master of Science in Electrical Engineering degree in the spring of 2005.