BURST SEGMENTATION: A NEW APPROACH TO REDUCE PACKET LOSS IN OPTICAL BURST SWITCHED NETWORKS

by

Vinod M. Vokkarane, B.E. in CS

THESIS
Presented to the Faculty of the Graduate School of
The University of Texas at Dallas
in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

THE UNIVERSITY OF TEXAS AT DALLAS
December 2001

Dedicated to my Parents

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor Dr. Jason Jue for giving me the guidance that have been instrumental in shaping my career. But for his invaluable advice and assistance it would not be possible for me to complete this thesis.

I would also like to express my gratitude to Dr. Balaji Raghavachari and Dr. Gopal Gupta for giving me the opportunity to work with them as a research assistant for the past year, in the Telecommunication Laboratory. I would also like to express my gratitude to Dr. Balaji Raghavachari and Dr. S. Q. Zheng for serving on my committee. Also a sincere thanks to all the members of Telecommunication Lab for their encouragements and advice. And most importantly to my parents, my brother and friends, for staying beside me at all times and helping me realize my ambitions.

BURST SEGMENTATION: A NEW APPROACH TO REDUCE PACKET LOSS IN OPTICAL BURST SWITCHED NETWORKS

Publication No. ______________

Vinod M. Vokkarane, M.S
The University of Texas at Dallas, 2001

Supervising Professor: Jason Jue

The current fast-growing Internet traffic is demanding more and more network capacity each day. The concept of wavelength-division multiplexing has provided us an opportunity to multiply network capacity. Currently optical switching technologies allow us to rapidly deliver the enormous bandwidth of WDM networks. Optical burst switching offers all-optical, high-speed data rate, format transparent switching, which are essential characteristics needed to support future networks supporting different classes of data.

In this report, we analyses a critical issue involved in the optical burst switched networks, namely contention resolution. In existing contention resolution schemes for optical burst switched networks, when contention between two bursts cannot be resolved through other means, one of the bursts will be dropped in its entirety, even though the overlap between the two bursts may be minimal. For certain applications, which have stringent delay requirements but relaxed packet loss requirements, it may be desirable to lose a few packets from a given burst rather than losing the entire burst.

We introduce a new approach called *Burst Segmentation*, to reduce packet loss during contention resolution. Burst segmentation is also analysed in conjunction with deflection routing to reduce packet loss during contention. We also introduce two ways of resolving contention namely, Deflect-First policy and Segment-First policy. In burst segmentation, rather than dropping the entire burst during contention, the burst may be broken into multiple segments, and only the overlapping segments during contention are dropped. Through simulation, it is shown that segmentation policy reduces packet loss substantially when compared to the standard policy of dropping the contenting burst in the event of a contention.

With deflection, burst segmentation can achieve a significantly reduced packet loss rate as compared to the deflect and drop policy, where in we deflect the contending burst to an alternate free port. In case all the alternate ports are busy, the burst is dropped.

We have also looked at handling prioritized data traffic. Prioritized routing and burst segmentation techniques for resolving contention show significant reduction in packet loss over the corresponding standard resolution policies. By selectively applying these techniques to bursts with different QoS requirements, we can offer a range of differentiated services at the optical layer. The simulation results for the new policies have been significantly better than their standard counterparts.

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

CHAPTER 1

INTRODUCTION TO OPTICAL BURST SWITCHING

The amount of raw bandwidth available on fiber optic links has increased dramatically with advances in dense wavelength division multiplexing (DWDM). The increase of bursty Internet traffic as opposed to non-bursty voice traffic, makes it important to handle this increasing class of traffic. In order to efficiently utilize this bandwidth, an all-optical transport method, which avoids optical buffering while handling bursty traffic, and which supports fast resource provisioning and asynchronous transmission of variable sized packets, must be developed. Optical Burst Switching (OBS) is one such method for transporting traffic directly over a bufferless optical WDM network [1].

Circuit and packet switching have been used for many years for voice and data communications respectively. Burst switching [2, 3], on the other hand, is less common. Switching techniques primarily differ based on whether data will use *switch cut-through* or *store and forward*. In circuit switching, a dedicated path between two stations is necessary. A call has to be established, the data is transferred and the call is disconnected. Resource reservation is done for the duration of the call. In packet switching, the data is broken into small packets and transmitted. The resources can be shared by different sources. End stations can send/receive data at their own speed. The individual packet can be individually switched or a virtual circuit can be set up. In the first case, the routing decision is done at a packet level while in the later, it is on a virtual channel level. Individual routing may lead to out-of-order message delivery.

Circuit switching is advantageous when we have a constant data rate (fixed delays) on the network like voice traffic; however, it is not suitable under bursty traffic conditions, or when circuits are idle [4]. Packet switching works well with variable rate traffic like data traffic, and can achieve higher utilization. Prioritization of data can also be incorporated in packet switching; however, it is difficult to give QoS assurances (best effort service), and packets can have variable delays [5].

Optical burst switching was introduced only recently for optical (WDM) networks, and is thus not as well understood as optical circuit and packet switching. Circuit switching uses two-way reservation schemes that have a large round trip. While packet switching has a large buffer requirement and complicated control and strict synchronization issues. OBS is designed to achieve a balance between the coarse-grained circuit switching and the fine-grained packet switching. As such, a burst may be considered as having an intermediate “granularity” as compared to circuit and packet switching. OBS uses one-way reservation schemes with immediate transmission, in which the data burst follows a corresponding packet without waiting for an acknowledgement [6, 7, 8, 9, 10, 11].

1.1 OBS Network Architecture

In optical burst switched networks, bursts of data consisting of multiple packets are switched through the network all-optically. A control message (or header) is transmitted ahead of the burst in order to configure the switches along the burst’s route. The data burst follows the header without waiting for an acknowledgement for the connection establishment.

Fig. 1.1 shows a OBS network. It consists of edge nodes (or routers) and core nodes

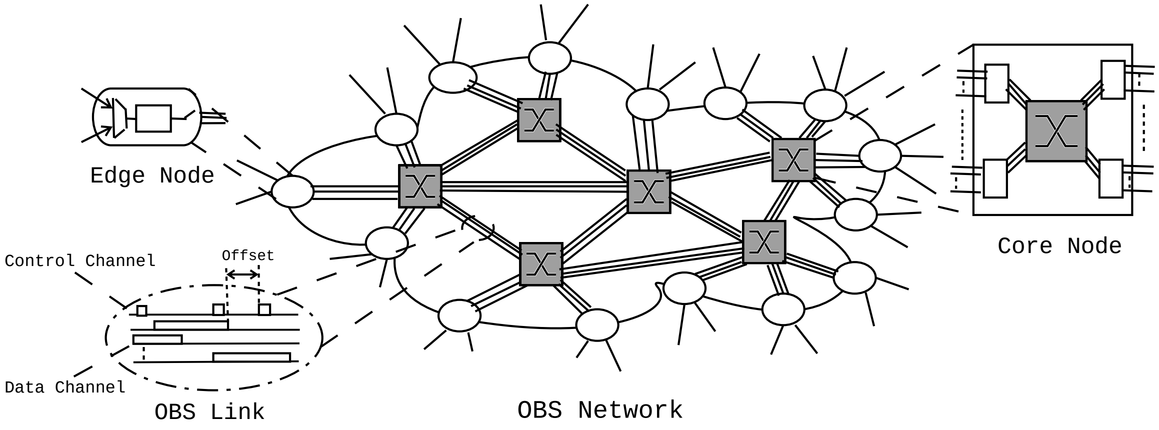

Figure 1.1: OBS Network Architecture

(or routers). An OBS network consists of optical burst switches interconnected with WDM links. An optical burst switch 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 may be fixed to carry one or more IP packets.

An generic switching node comprises of the following:

- *Input interface:* header and data burst reception and header conversion to electrical signals [12].
- *Switching control unit:* header interpretation, scheduling, collision detection and resolution, forwarding table lookup, switching matrix control, header rewrite and wavelength conversion control.
- *Wavelength converters* and *optical delay lines (ODLs):* the delay lines are used as a buffer to store the burst for a delay period.

- *Optical switching unit:* space switches for switching the data burst.

The *edge nodes* (either ingress or egress router), will have additional functionality of burst creation by aggregation and de-aggregation. Different policies, such as having a threshold or a timeout can be used to aggregate bursty data packets to create an optical burst and to send the burst into the network [13] . The *core nodes* will have WDM receivers, WDM transmitters, WDM multiplexers, WDM de-multiplexers, node amplifiers, switch control units, wavelength converters, delay lines and space division switches.

1.2 Offset Time and Reservation Schemes

Optical Burst switching techniques differ based on how and when the network resources like bandwidth, is reserved and released.

Optical burst switching is an adaptation of an International Telecommunication Union-Telecommunication Standardization Sector (ITU-T) standard for burst switching in asynchronous transfer mode (ATM) networks, known as ATM block transfer (ABT). There are two versions of ABT: ABT with delayed transmission and ABT with immediate transmission. In the first case, when a source wants to transmit a burst, it sends a packet to the ATM switches on the path of the connection to inform them that it wants to transmit a burst. If all the switches on the path can accommodate the burst, the request is accepted and the source is allowed to go ahead with its transmission. Otherwise, the request is refused, and the source has to send another request later. In ABT with immediate transfer, the source sends the request packet, and then immediately following the request, without receiving a confirmation, the source transmits its burst. If a switch along the path cannot carry the burst due to congestion, the burst is dropped. These two techniques have been adopted to optical networks.

In optical burst switched networks, the control header and the data burst are separated at the source, as well as subsequent intermediate nodes, by an offset time, as shown in Fig. 1.2. The offset time allows for the header to be processed at each node while the burst is buffered electronically at the source; thus, no fiber delay lines are necessary at the intermediate nodes to delay the burst while the header is being processed. The control message may also specify the duration of the burst in order to let a node know when it may reconfigure its switch for the next burst, a technique known as Delayed Reservation (DR) [1].

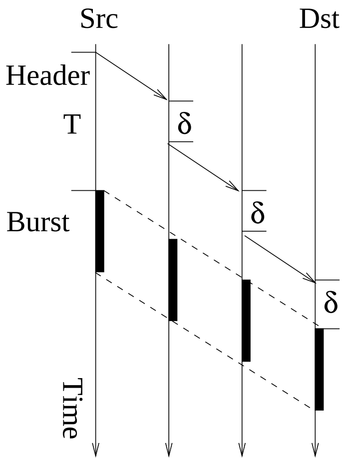

Figure 1.2: The use of offset time in OBS.

The tell-and-go (TAG) scheme [14, 15] is similar to the ABT with immediate transmission, and the tell-and-wait (TAW) scheme [15] is similar to ABT with delayed transmission. An intermediate scheme known as Just Enough Time (JET) was proposed in [1].

In the TAG scheme, the source transmits the control packet and then immediately transmits the optical burst. In this scheme, it may be necessary to buffer the burst in the optical burst switch until its control packet has been processed. In the JET scheme there is a delay between transmission of the control packet and transmission of the optical burst. This delay can be set to be larger than the total processing time of the control packet along the path. This way, when the burst arrives at each intermediate node, the control packet has been processed and a channel on the output port has been allocated. Therefore, there is no need to buffer the burst at the node. This is a very important feature of the JET scheme, since optical buffers are difficult to implement. A further improvement of the JET scheme can be obtained by reserving resources at the optical burst switch from the time the burst arrives at the switch, rather than from the time its control packet is processed at the switch. In [1] a variation of JET was proposed which supports quality of service. Specifically, two traffic classes were defined: real-time and non-real-time. A burst belonging to the real-time class is allocated higher priority than a burst belonging to the non-real-time class by simply using an additional delay between transmission of the control packet and transmission of the burst.

1.3 Research Objectives

In this work, we investigate the OBS network architecture, its advantages, issues, reservation policies and related concepts. Our research is concentrated on Contention Resolution in OBS networks.

We introduce a new concept for reducing packet loss during contention called *Burst Segmentation*. In burst segmentation, when there is a contention only the overlapping packets of the burst is dropped instead of the entire contending burst. If we consider deflection as an option, then the remaining part of the segmented burst can be deflected to an alternate port. This significantly improves the packet loss compared to the standard policy of just dropping the contending burst or deflecting the contending burst. There are two ways of implementing burst segmentation with deflection namely, *Segment-First policy* and *Deflect-First policy*. In the Segment-First policy, the original burst is segmented and it's tail is deflected. While in the case of Deflect-First policy the contending burst is deflected if an alternate port is free, otherwise the original burst is segmented and its tail is dropped. We study the performance of both the policies with and without deflection and observe that policies with deflection outperform the standard dropping policy (with and without deflection).

We also investigate prioritized data traffic and address QoS issues. We develop new schemes to handle contention in the case of prioritized data. Simulations are done to test the performance of the policies under various data ratios, loads, etc. The results indicate that the new burst segmentation policy works well even in the case of providing differentiated services.

1.4 Overview

This report consists of five chapters. This chapter has outlined a brief introduction to optical burst switching as well as the research objectives. Chapter 2 covers prior art on contention resolution in optical networks. Chapter 3 proposes the concept of burst segmentation for contention resolution. Chapter 4 addresses the QoS issues while implementing burst segmentation with deflection. Both chapters 3 and 4 introduce many new contention resolution policies in combination with deflection and prioritized routing, and also give the simulation results for each of the policies. Chapter 5 concludes the thesis and identifies areas of future work.

CHAPTER 2

CONTENTION RESOLUTION IN OPTICAL NETWORKS

2.1 Contention Resolution

Contention resolution is necessary in order to handle the case where more than one packet (or burst) are destined to go out of the same output port at the same time. This is a problem that commonly arises in packet switches, and is known as *external blocking*. External blocking is typically resolved by buffering all the contending packets, except one that is permitted to go out. In an optical packet switch, techniques designed to address the external blocking problem include optical buffering, wavelength conversion, and deflection routing. Whether these will prove adequate to address the external blocking problem is still questionable. Below we look at each of these solutions.

2.2 Optical Buffering

Typically, contention in traditional electronic packet-switching networks is handled through buffering; however, in the optical domain, it is more difficult to implement buffers, since there is no optical equivalent of random-access memory. Instead, optical buffering is achieved through the use of fiber delay lines [16, 17]. 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. Some papers have investigated approaches for designing larger buffers without a large number of delay lines [18, 19]. In [18], the buffer size is increased by cascading multiple stages of delay lines. In [19], the buffer size is increased by utilizing so called non-degenerate buffers in which the length of the delay lines may be greater than the number of delay lines in the buffer. This approach yields lower packet loss probabilities, but does not guarantee the correct ordering of the packets. Note that, 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. To delay a single packet for 5 μs requires over a kilometer of fiber. Because of this size limitation of optical buffers, a node may be unable to effectively handle high load or bursty traffic conditions. Wavelength controlled fiber loop buffer and wavelength routing based photonic packet buffers are described in [20, 21].

In addition to buffering packets optically, it is also possible to buffer packets electronically. Electronic buffering can be accomplished by sending the packets up to the electronic switching or routing layer. The disadvantage of such an approach is that the network loses transparency, and each node must have electronic switching or routing capabilities, resulting in higher network costs and also electronic memories which must keep up with the speeds of optical networks. Furthermore, a greater load will be placed on the processing capabilities of the electronic switch or router. An alternative would be to implement electronic buffers directly as a part of the optical switch itself. In this case each node would still require additional transmitters and receivers, and would need to be aware of the transmission format of the packets; however no additional electronic routing or switching capability would be required. Delay lines may be acceptable in prototype switches, but are not commercially viable.

2.3 Wavelength Conversion

In WDM, several wavelengths run on a fiber link that connects two optical switches. This can be exploited to minimize external blocking as follows. Let us assume that two packets are destined to go out of the same output port at the same time. Both packets can be still transmitted, but on two different wavelengths. This method may have some potential in minimizing external blocking, particularly since the number of wavelengths that can be coupled together onto a single fiber continues to increase. For instance, it is expected that in a year there will be as many as 200 wavelengths/fiber.

Wavelength conversion is the process of converting the wavelength of an incoming channel to another wavelength at the outgoing channel. Wavelength convertors are devices that convert an incoming signal's wavelength to a different outgoing wavelength thereby increasing the reuse factor. Wavelength convertors offer a 10%-40% increase in reuse values when wavelengths availability is small [22].

Despite such expectations and some promising experimental reports, the wavelength conversion technologies are as yet immature. The following are the different categories of wavelength conversion:

- *No conversion:* No wavelength shifting.
- *Full conversion:* Any wavelength shifting is possible and so channels can be connected regardless of their wavelengths.
- *Limited conversion:* Wavelength shifting is restricted so that not all combination of channels may be connected.

- *Fixed conversion:* Restricted form of limited conversion that has 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.

There are many wavelength conversion algorithms and algorithms to minimizing the number wavelength converters [23, 24, 25]. The wavelength conversion techniques suffers from the high cost involved in building wavelength convertors.

2.4 Deflection Routing

Deflection routing or hot-potato routing is ideally suited to switches that have little buffer space. This approach of resolving contention is to route the contending packets to an output port other than the intended output port [26, 27, 28]. 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. Also, packets will have to be re-ordered at the destination since they are likely to arrive out of sequence. Below, we examine various optical packet switch architectures that have been proposed in the literature. While deflection routing is generally not favored in electronic packet-switched networks due to potential looping and out-of-sequence delivery of packets, it may be necessary to implement deflection in photonic packet-switched networks, where buffer capacity is very limited, in order to maintain a reasonable level of packet losses. However, before attempting to deploy deflection in photonic packet-switched networks, a comprehensive study is required in order to identify potential methods for overcoming some of the limitations of deflection, and to determine whether or not these methods, along with the potential benefits of deflection, are sufficient to justify implementation.

In [26], hot-potato routing is compared to store-and-forward routing in a ShuffleNet. [27] and [28] compare hot-potato and deflection routing in ShuffleNet and Manhattan Street Network topologies. Since both the ShuffleNet and Manhattan Street Network are two-connected (each node has an outgoing degree of two), the choice of the deflection output port is obvious. When the nodal degree is greater than two, a method must be developed to select the alternate outgoing link when a deflection occurs. In [29], deflection routing is studied in irregular mesh networks. Rather than choosing the deflection output port arbitrarily, priorities are assigned to each output port, and the ports are chosen in the prioritized order. In [30], deflection is studied together with optical buffering in irregular mesh networks with variable-length packets. The nodes at which deflection can occur, as well as the options for the deflection port, are limited in such a way as to prevent looping for the given network; however, a general methodology for selecting deflection options to avoiding looping in any arbitrary network is not given.

CHAPTER 3

BURST SEGMENTATION: AN APPROACH FOR REDUCING PACKET LOSS DURING CONTENTION

3.1 Introduction

A major concern in optical burst switched networks is contention, which occurs when multiple bursts contend for the same link. Contention in an optical burst switched network is particularly aggravated by the highly variable burst sizes and the long burst durations. Furthermore, since bursts are switched in a cut-through mode rather than a store-and-forward mode, optical burst-switched networks generally have very limited buffering capabilities. While existing contention resolution schemes for photonic packet networks, such as deflection and buffering, may be utilized in optical burst switched networks, additional schemes may also be necessary in order to combat high contention rates and to achieve high network utilization.

In [31], an offset scheme was proposed for isolating classes of bursts, such that low-priority bursts do not cause contention losses for high-priority bursts; fixed and variable fiber delay line buffers were also utilized to further reduce blocking. In [31] and [8] contention is reduced by utilizing additional capacity in the form of multiple wavelengths. In both cases, optical wavelength conversion was assumed. While optical wavelength conversion has been demonstrated in laboratory environments, the technology is not yet mature, and the range of possible conversions is somewhat limited.

While optical wavelength conversion has been demonstrated in laboratory environments, the technology is not yet mature, and the range of possible conversions is somewhat limited. Most of the current literature deals with approaches to minimize burst losses rather than packet losses. In existing contention resolution schemes for optical burst switched networks, when contention between two bursts cannot be resolved through other means, one of the bursts will be dropped in its entirety, even though the overlap between the two bursts may be minimal. For certain applications, which have stringent delay requirements but relaxed packet loss requirements, it may be desirable to lose a few packets from a given burst rather than losing the entire burst. In this chapter, we will introduce a new contention resolution technique called *burst segmentation*, in which only those packets that overlap with a contending burst will be dropped [33].

The chapter is organized as follows: Section 3.2 introduces the concept of burst segmentation and describes the segment dropping policies. Section 3.3 discusses segmentation with deflection. Section 3.4 compares the simulation results for different contention resolution policies in a specific network topology, and Section 3.5 concludes the chapter.

3.2 Burst Segmentation

In burst segmentation, the burst is divided into basic transport units called segments. Each of these segments may consist of a single packet or multiple packets, and the segments define the possible partitioning points of a burst when the burst is in the optical network. Each segment/packet will have additional header information like, its length, checksum, etc. All segments in a burst are initially transmitted as a single burst unit. The out-of-band header will have information like, the length of the burst, the offset time, etc.

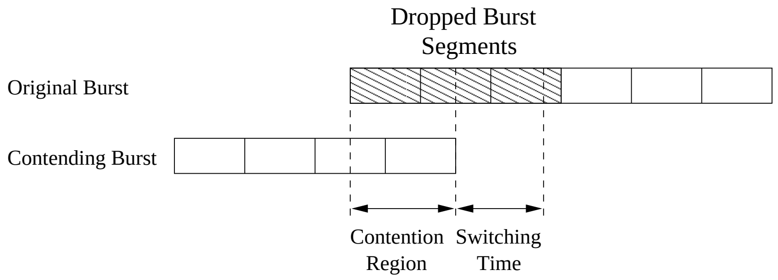

Figure 3.1: Selective segment dropping for two contending bursts.

When contention occurs, only those segments of a given burst which overlap with segments of another burst will be dropped, as shown in Fig. 3.1. If switching time is non-negligible, then additional segments may be lost when the output port is switched from one burst to another.

There are two approaches for dropping burst segments when contention occurs between bursts. The first approach is to drop the tail of the first burst (Fig. 3.1), and the second approach is to drop the head of the contending burst. A significant advantage of dropping the tail segments of bursts rather than the head segments is that there is a better chance of in-sequence delivery of packets at the destination, assuming that dropped packets are retransmitted at a later time.

One issue that arises when the tail of a burst is dropped is that the header for the burst, which may be forwarded before the segmentation occurs, will still contain the original burst length; therefore, downstream nodes may not know that the burst has been truncated. If downstream nodes are unaware of a burst's truncation, then it is possible that the previously truncated tail segments will contend with other bursts, even though these tail segments have already been dropped at a previous node. These contentions may result in unnecessary packet loss.

If a tail-dropping policy is strictly maintained throughout the network, then the tail of the truncated burst will always have lower priority, and will never preempt segments of any other burst. However for the case in which tail dropping is not strictly maintained, some action must be taken to avoid unnecessary packet losses. A simple solution is to have the truncating node generate and send out a trailing control message to indicate when the truncated burst ends. In this policy, the offset between the trailer packet and the end of the truncated burst is similar to the offset between the header and the start of the burst.

In a head-dropping policy, the head segments of the contending burst will be dropped. A head-dropping policy will result in a greater likelihood that packets will arrive at their destination out of order. Also, the control message of the contending burst would need to be modified and delayed. The advantage of head-dropping is that it ensures that, once a burst arrives at a node without encountering contention, then the burst is guaranteed to complete its traversal of the node without preemption by later bursts.

In this chapter, we consider a modified tail-dropping policy when determining which segment to drop. In this policy, the tail of the original burst is dropped only if the number of segments in the tail is less than the total number of segments in the contending burst. If the number of segments in the tail is greater than the number of segments in the contending burst, then the entire contending burst is dropped. This approach reduces the probability of a short burst preempting a longer burst and minimizes the number of packets lost during contention.

There are a number of additional issues and challenges which arise when implementing burst segmentation in practical systems:

1. Switching time: Since the system does not implement buffering or any other delay mechanism, the switching time is a direct measure of the number of packets lost during reconfiguring the switch due to contention. Hence, a slower switching time results

in higher packet loss. While deciding which burst to segment, we consider the remaining length of the original burst, taking the switching time into account. By including switching time in burst length comparisons, we can achieve the optimal output burst lengths for a given switching time.

2. Segment boundary detection: In the optical network, segment boundaries of the burst are transparent to the intermediate nodes that switch the burst segments all-optically. At the network edge nodes, the burst is received and processed electronically. Since the burst is made up of many segments, the receiving node must be able to detect the start of each segment and identify whether or not the segment is intact. If each segment consists of an Ethernet frame, detection and synchronization can be performed using the preamble field in the Ethernet frame header, while errors and incomplete frames can be detected by using the CRC field in the Ethernet frame.

3. Trailer creation: The trailer has to be created electronically at the switch where the contention is being resolved. The time to create the trailer can be included in the header processing time, δ, at each node.

3.3 Segmentation with Deflection

A basic extension of burst segmentation is to implement segmentation with deflection. Rather than dropping segments of a burst, we can either deflect the entire burst or deflect segments of the burst to an output port other than the intended output port. This approach is referred to as deflection routing or hot-potato routing [26, 27, 28]. Implementing segmentation with deflection (Fig. 3.2) increases the probability of the burst reaching the destination and hence improves the performance. One problem which may arise is that a burst may encounter looping or may be deflected multiple times, thereby wasting network bandwidth. In order to avoid these problems, when the hop-count of the burst reaches a threshold, the

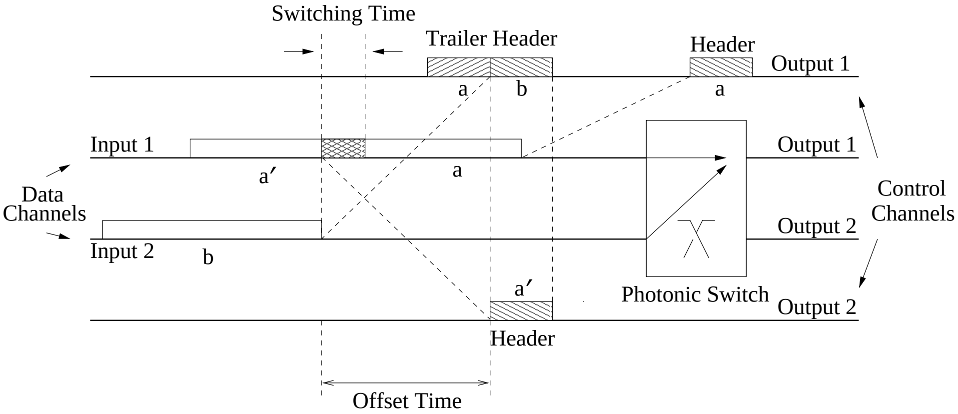

Figure 3.2: Segmentation with deflection policy for two contending bursts.

burst is dropped. This limitation on hop-count also ensures that the offset time maintains a reasonable value. Also, deflection increases blocking probability in high load conditions [30].

There could be one or many alternate deflection ports. The alternate deflection port(s) could be allotted ahead of time using fixed port assignment policy or using the second shortest path algorithm. A load balancing approach, which is based on the current link utilizations, could also be used so that the burst is deflected to an under-utilized link. In this chapter, we consider only one alternate deflection port, and choose the port which results in the second shortest path to the destination.

Selection of which burst (or burst-segments) to deflect during contention could be done in one of the following ways. Firstly, the burst with shorter remaining length (taking switching time into account) could be deflected to the alternative port, or dropped if the alternate port is busy (Fig. 3.2). Secondly, we could incorporate priorities into the burst, so that in case of contention the lower priority burst is deflected or segmented based on the underlying policy.

Now combining segmentation with deflection, we have two approaches for ordering the contention resolution policies, namely, *segment-first* and *deflect-first*. In the segment-first policy, if the remaining length of the original burst is shorter than the contending burst, then the original burst is segmented and its tail is deflected, otherwise the contending burst is deflected. In case the alternate port is busy, the deflected part of the original burst is dropped. In the deflect-first policy, in case of contention, the contending burst is deflected if the alternate port is free. If the alternate port is busy and if the remaining length of the original burst is shorter, then the original burst is segmented and its tail is dropped. If the contending burst was found to be shorter, then the original burst is dropped. In this chapter, we consider the segment-first policy.

An example of the segmentation-deflection scheme is shown in Fig. 3.2. Initially when the header for *burst a* arrives at the switch, it is routed onto output port 1. Once the header of the *burst b* arrives at the switch we have a contention. Since the offset time is common to all the bursts, the header indicates when and where the bursts will contend. So taking the switching time into consideration, and based on the segment-first policy, one of the bursts is deflected (or segmented and deflected) to the alternate port if it is free and is dropped otherwise. Here the remaining length of *burst a* is less than the length of *burst b*. Hence *burst a* is segmented and its tail is deflected to the alternate port as a new burst. A header is created for the deflected new burst and sent on the output 2 control channel. This new header generation is done at the time the header of *burst b* is processed. A trailer is created for the segmented *burst a* and is sent on the control channel of output 1. Packets of the burst to be segmented are lost during the reconfiguration of the switch. Hence faster switching time improves the performance. In the segmentation with deflection policy, the processing time δ (Fig. 1.2) at each node includes the time to create a header for the new burst segment in case of contention. Hence the offset time is same as in the case of standard optical burst switching.

A possible side-effect of segmentation with deflection is that, when there is contention, the shorter remaining burst will get segmented and will be deflected as a new burst. Creating these new short bursts may lead to burst fragmentation. The newly created short burst may contend with other bursts in the network, leading to additional fragmentation. Fragmentation is not a major issue, as the policies for deflection and dropping tend take care of the smaller burst. Every time a burst is segmented, the lengths of the two colliding bursts are compared and the smaller of the contending burst or the remaining part of the first burst is deflected or segmented respectively. Thus, the short, fragmented bursts will have lower priority and will not significantly hinder other bursts.

Another issue when implementing segmentation and deflection is how to handle long bursts which may span multiple nodes simultaneously. If a long burst passing through two or more switches experiences contention from two or more different bursts at different switches, then, based on the timing of these contentions, the contentions are resolved in the following manner:

If an upstream node segments the burst first, then the downstream nodes are updated by the trailer packet to eliminate unnecessary contentions. On the other hand, if the contention occurs at the downstream node before the upstream node, and if the burst's tail is deflected at the downstream node, then the upstream contentions will not be affected. If the downstream node drops the tail of the burst, then the upstream node will not know about the truncation and will continue to transmit the tail. The downstream node may send a control message to the upstream node in order to reduce unnecessary contentions with the tail at the upstream node.

3.4 Simulation Results

In order to evaluate the performance of the segmentation and deflection schemes, we develop a simulation model. The following have been assumed to obtain the results:

- Burst arrivals to the network are Poisson with rate λ.
- Burst length is exponentially distributed with rate μ.
- Load is measured in Erlang.
- Transmission rate is 10 Gbps.
- Packet length is 1500 bytes.
- Switching time is 10 μs.
- There is no buffering or wavelength conversion at nodes.
- Each node handles both bypassing and locally generated or terminated bursts.
- Bursts are uniformly distributed over all sender-receiver pairs.
- Dijkstra shortest path routing algorithm is used to find the path between all node pairs.

Figure 3.3 shows the 14-node NSFNET on which the simulation was implemented. The distances shown are in km. We have compared four different policies for handling contention in the OBS network, they are:

- *Drop Policy (DP):* Drop the entire contending burst.

- *Deflect and Drop Policy (DDP):* Deflect the contending burst to the alternate port. If the port is busy, drop the burst.
- *Segment and Drop Policy (SDP):* Segment-first policy without deflection.
- *Segment, Deflect and Drop Policy (SDDP):* Segment-first policy with deflection.

Figure 3.4 plots the total packet loss probability versus the load for the four different contention resolution policies. An average burst length of $1/\mu$ = 100 ms is assumed. We observe that SDP performs better than DP in all load conditions, and the two policies with deflection namely, DDP and SDDP perform better than the corresponding policies without deflection at low loads. Also, at low loads DDP performs better than SDDP since there is no loss due to switching time in DDP; whereas, at high loads, SDDP is better than DDP. A logical explanation would be that, in segmentation, on average only half of the packets from one of the bursts are lost when contention occurs. Also, at low loads, there is a greater amount of spare capacity, increasing the chance of successful deflection. At high loads, deflection may add to the load, increasing the probability of contention, and thereby increasing loss.

Figure 3.5 shows the packet-loss performance at very high loads. SDDP performs the best when the load is under 50 Erlang, after which SDP performs better. DDP is good only at low loads, while at very high loads DP fares better than DDP. We observe that, at very high loads, policies without deflection perform better then the policies with deflection. This is because deflection increases the effective arrival rate within the network, which may lead to more contentions.

Figure 3.6 shows the average number of hops versus load. For the deflection policies, the number of deflections increase as the load increases, resulting in increasing average hop distance at low loads. As the load increases further, those bursts which are further

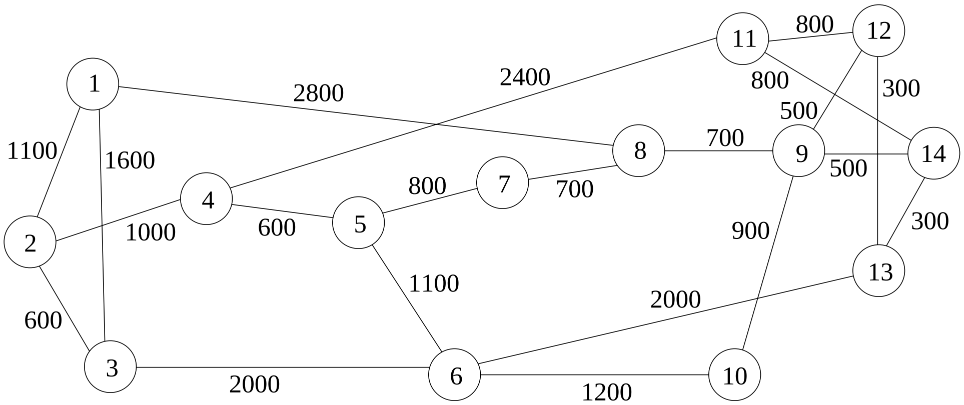

Figure 3.3: Picture of NSFNET with 14 nodes.

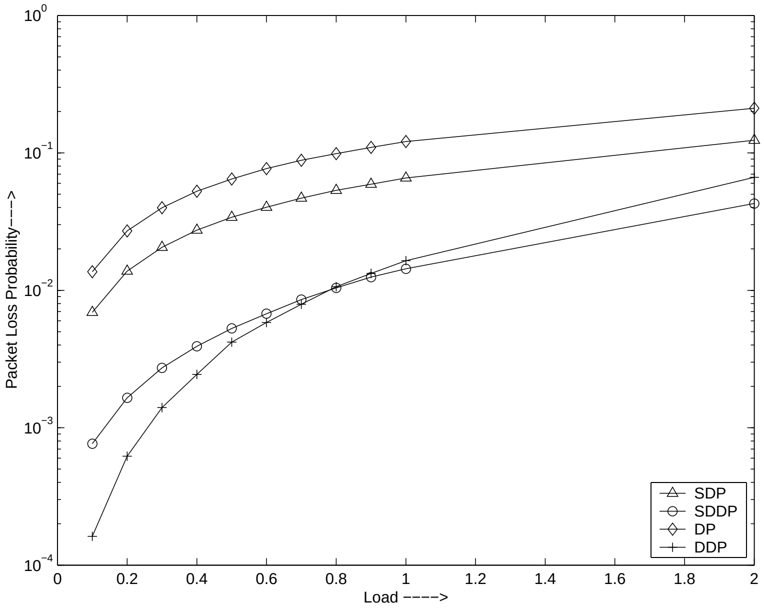

Figure 3.4: Packet loss probability versus load for NSFNET at low loads with $\frac{1}{\mu} = 100$ ms.

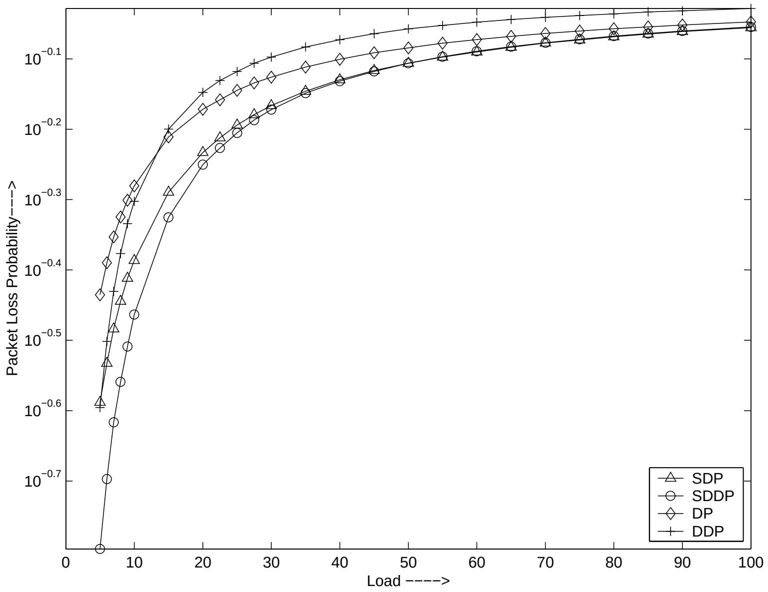

Figure 3.5: Packet loss probability versus load for NSFNET at high loads with $\frac{1}{\mu} = 100$ ms.

from their destination will experience more contention than those bursts which are close to their destination. Thus, bursts with higher average hop count are less likely to reach their intended destination, and the average hop distance will decrease as load increases.

Figure 3.7 shows the simulation results of the average output burst size versus the load for SDP and SDDP. The output burst size is measured over both dropped and successfully received bursts. Initially, the burst size decreases with increasing load, as there are more segmentations with the increasing number of contentions. As the load increases further, the segmented bursts encounter more contentions, and because the segmented bursts have smaller size (lower priority), they are dropped. The values for DP and DDP are constant for different values of load as the size of the burst is never altered.

The packet loss probability versus load for different values of switching time is shown in Fig. 3.8. As the switching time increases, the performance of SDDP decreases as a greater number of packets are lost during the re-configuration of the switch. On the

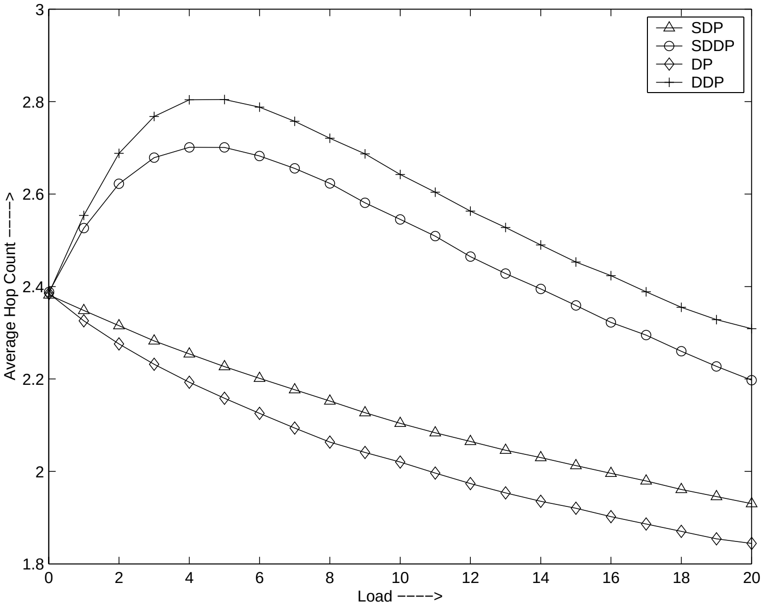

Figure 3.6: Average number of hops versus load for NSFNET with $\frac{1}{\mu} = 100$ ms.

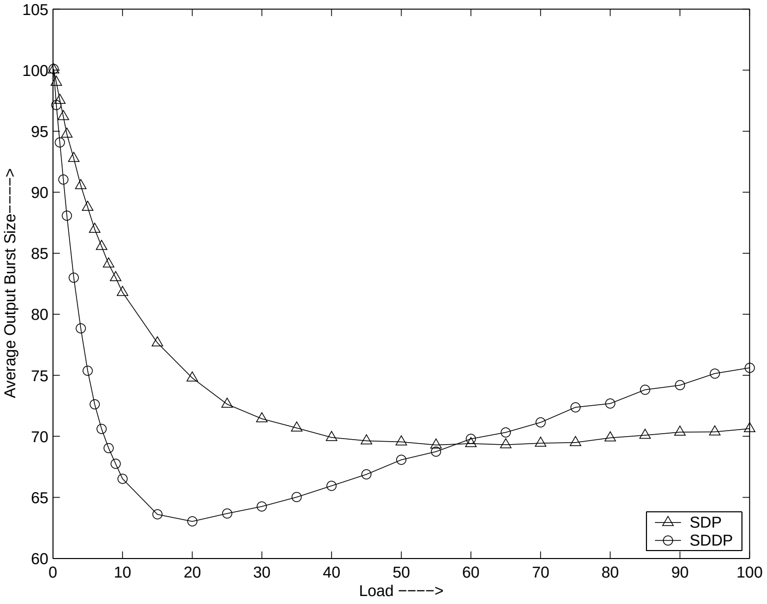

Figure 3.7: Average output burst size versus load for NSFNET with $\frac{1}{\mu} = 100$ ms.

other hand, DDP is not affected by the switching time and is almost constant. At low switching times, the results show that SDDP is better than the standard DDP. While at higher switching times, the standard DDP is better than the new SDDP because of the loss of packets during the switching time.

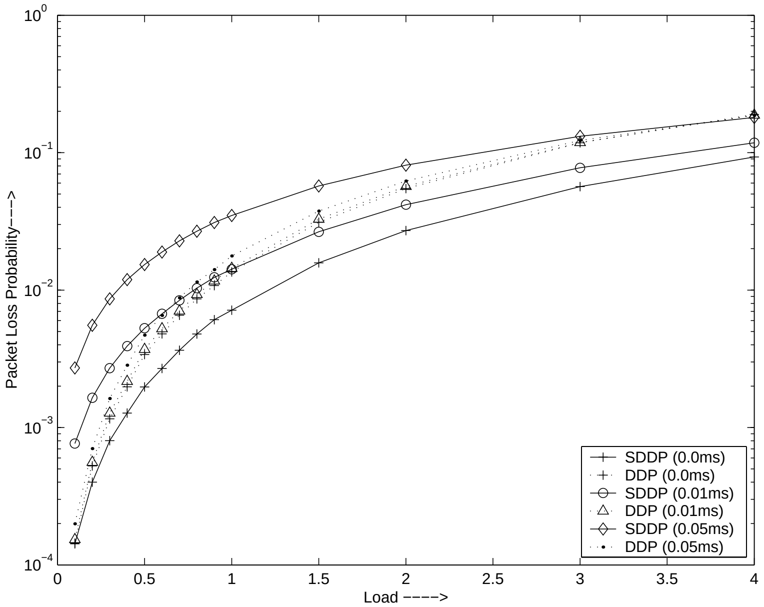

Figure 3.8: Packet loss probability versus load at varying switching times for NSFNET with $\frac{1}{\mu} = 100$ ms.

The simulation results of different contention resolution policies for a 15-node network Fig. 3.9, gives similar results.

3.5 Conclusion

In this chapter, we introduced the concept of burst segmentation for contention resolution in optical burst switched networks, and we investigated a number of different policies with and without segmentation and deflection. The segmentation policies perform better than the standard dropping policy, and offer the best performance at high loads. The policies

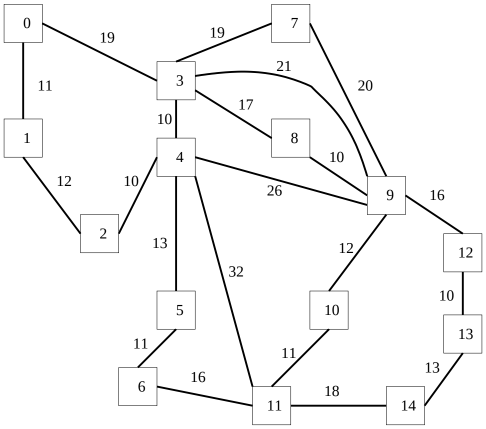

Figure 3.9: 15 Node PACnet

which incorporate deflection tend to perform better at low loads.

CHAPTER 4

PRIORITIZED ROUTING AND BURST SEGMENTATION FOR QUALITY OF SERVICE IN OPTICAL BURST-SWITCHED NETWORKS

4.1 Introduction

In this chapter, we will refer to the burst which arrives to the switch first as the *original burst*, and the burst which arrives to the switch later as the *contending burst*. One approach to resolving contention is to use *deflection routing*, in which the contending burst is sent to an alternate output port. *Prioritized routing* is an approach in which the choice of packet/burst to deflect is made based on priority [32].

Another important issue in optical burst-switched networks is how to provide differentiated service in order to support the various quality of service (QoS) requirements of different applications. In [31], an offset scheme was proposed for isolating classes of bursts, such that low-priority bursts do not cause contention losses for high-priority bursts; fixed and variable fiber delay line buffers were also utilized to further reduce blocking.

In this chapter, we will combine burst segmentation with prioritized deflection routing in order to offer differentiated services at the optical layer. We will assume that there are two priority classes, and that a high-priority burst is one which has low delay and loss tolerance while a low-priority burst has higher level of delay and loss tolerance.

4.2 Burst Segmentation And Deflection

To overcome some of the limitations of optical burst switching, burst segmentation can be used to minimize packet loss during contention. When contention occurs, a burst is divided into multiple segments, and only those packets of a given burst which overlap with segments of another burst will be dropped. If switching time is non-negligible, then additional packets may be lost when the output port is switched from one burst to another. Segmentation is used primarily to minimize loss of high priority bursts. There are two approaches for segmenting a burst when contention occurs. The first approach is to segment the tail of the original burst, and the second approach is to segment the head of the contending burst. A significant advantage of segmenting the tail of bursts rather than segmenting the head is that there is a better chance of in-sequence delivery of packets at the destination, assuming that dropped packets are retransmitted at a later time. In this chapter, we will assume that the tail of the original burst will be segmented when segmentation takes place (Fig. 3.1). When a burst is segmented, its control message is updated accordingly. For the case in which segmentation occurs in the middle of a packet, the fractional packet is lost.

Burst segmentation can also be implemented with deflection. Rather than dropping the tail segment of the original burst, we can either deflect the entire contending burst, or we can deflect only the tail segment of the original burst. Implementing segmentation with deflection increases the probability that a burst's packets will reach the destination, and hence improves performance. At each node, one or more alternate deflection ports can be specified for each destination. The order in which the alternate deflection ports are attempted is determined by a shortest-path policy.

The QoS scheme is implemented by selectively choosing which burst (original or contending) to segment or deflect during contention. We define the following policies for handling contention between two bursts:

- *Segment First and Deflect Policy (SFDP):* The contending burst wins the contention. The original burst is segmented, and its segmented tail may be deflected if an alternate port is free, otherwise the tail is dropped.
- *Deflect First and Drop Policy (DFDP):* The contending burst is deflected to an alternate port, if no alternate port is free the contending burst is dropped.
- *Deflect First, Segment and Drop Policy (DFSDP):* The contending burst is deflected to a free port if available, otherwise the original burst is segmented and its tail is dropped, while the contending burst is transmitted.

These approaches minimize the loss probability of a high priority burst during contention. The Deflect First and Drop Policy (DFDP) is a greedy approach, which performs better than SFDP at low loads, while at high loads it is out performed by SFDP.

The decision of which of these three policies to implement when a contention occurs is based on the priorities of the two bursts involved, as well as the lengths of the bursts, as shown in Table 1.

The rows of Table 1 show the various contention situations in a two priority network. *Longer Remaining Burst* indicates which of the two contending burst has a greater number of packets remaining to be transmitted from the point of contention. If all else is equal, the burst with fewer remaining packets will be given slightly lower priority in order to minimize the number of packets lost or deflected.

Table 4.1: QoS policies for various contention situations.

Condition	Original Burst Priority.	Contending Burst Priority.	Longer Remaining Burst	Scheme 1	Scheme 2
1	High	High	Contending	DFSDP	SFDP
2	High	Low	Contending	DFDP	DFDP
3	Low	High	Contending	SFDP	SFDP
4	Low	Low	Contending	DFSDP	SFDP
5	High	High	Original	DFDP	DFDP
6	High	Low	Original	DFDP	DFDP
7	Low	High	Original	SFDP	SFDP
8	Low	Low	Original	DFDP	DFDP

To better understand the table, let us consider Condition 2 [Fig. 4.2]. The original burst has high priority while the contending burst has low priority, and the contending burst is longer than the original burst. This contention is resolved by DFDP, where we first attempt to deflect the low-priority burst, and then drop the low-priority burst if no deflection ports are available. On the other hand, in Condition 7 [Fig. 4.4], a high-priority contending burst contends with a low-priority original burst, and the number of packets in the high-priority burst is less than the number of packets in the tail of the low-priority burst. In this case, we segment the tail of the low-priority burst and attempt to deflect the tail if a deflection port is available; otherwise, the tail is dropped (SFDP).

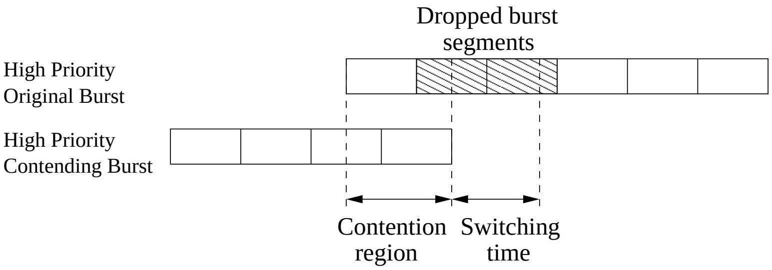

Figure 4.1: Condition 1 of table for two contending bursts.

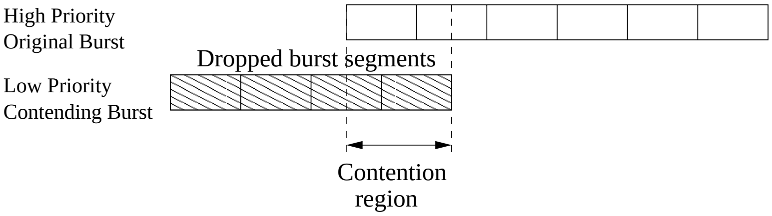

Figure 4.2: Condition 2 of table for two contending bursts.

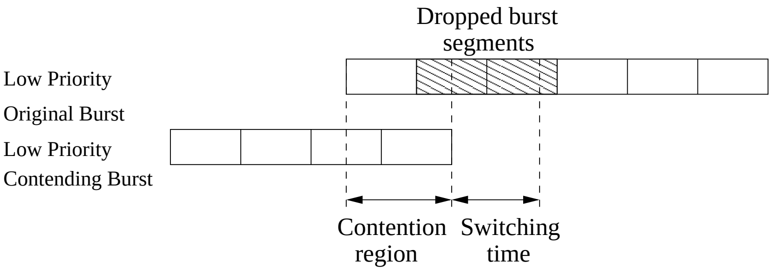

Figure 4.3: Condition 4 of table for two contending bursts.

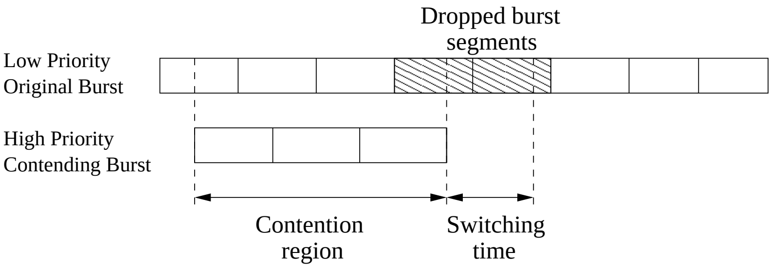

Figure 4.4: Condition 7 of table for two contending bursts.

We define two different QoS schemes that differ in the handling of Conditions 1 and 4, [Fig. 4.1, 4.3] i.e., when the contention is between two bursts of the same priority, and the contending burst is longer than the original burst. Scheme 1 will first attempt to deflect the contending burst, and will segment the original burst only if no deflection port is available (DFSDP). Scheme 2 will always segment and attempt to deflect the tail of the original burst (SFDP).

4.3 Simulation Results

In order to evaluate the performance of the segmentation and deflection schemes, we develop a simulation model. The following have been assumed to obtain the results:

- Burst arrivals to the network are Poisson with rate λ.
- Burst length is exponentially distributed with average burst length of $1/\mu$ = 100 ms.
- Load is measured in Erlang.
- Transmission rate is 10 Gbps.
- Packet length is 1500 bytes.
- Switching time is 10 μs.
- There is no buffering or wavelength conversion at nodes.
- Each node handles both bypassing and locally generated or terminated bursts.
- Bursts are uniformly distributed over all sender-receiver pairs.
- Dijkstra shortest path routing algorithm is used to find the path between all node pairs.

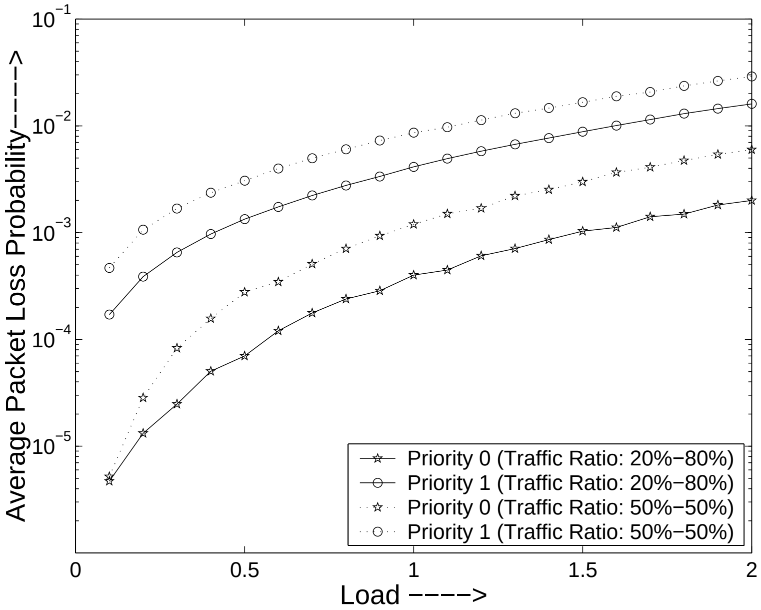

Figure 4.5: Packet loss probability versus load for different traffic ratios using Scheme 1

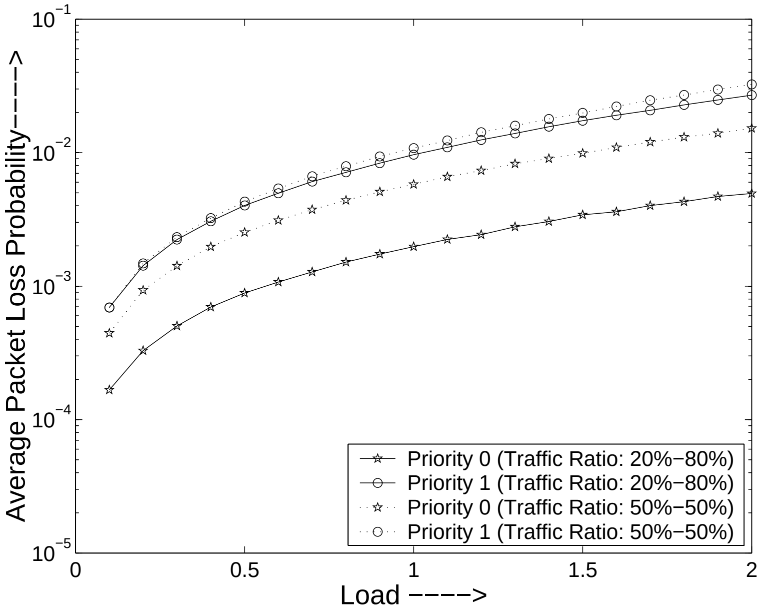

Figure 4.6: Packet loss probability versus load for different traffic ratios using Scheme 2

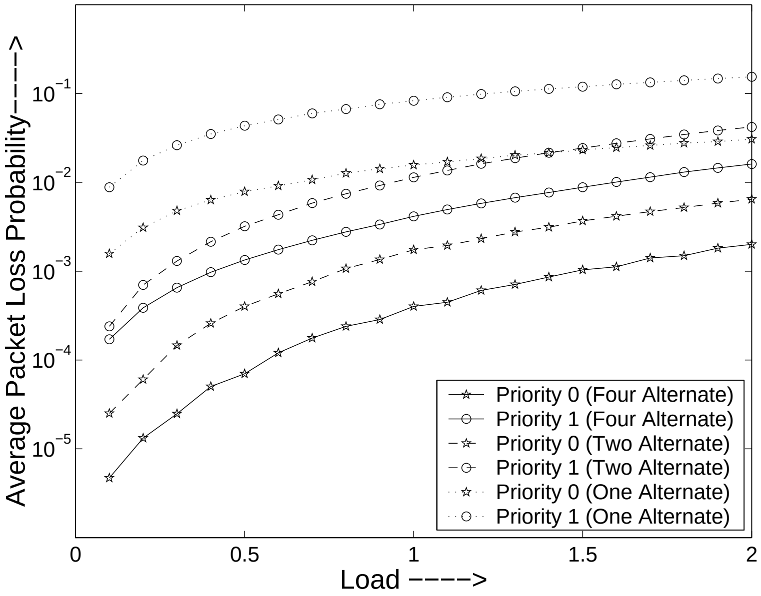

Figure 4.7: Packet loss probability versus load Scheme 1 with different number of alternate deflection ports.

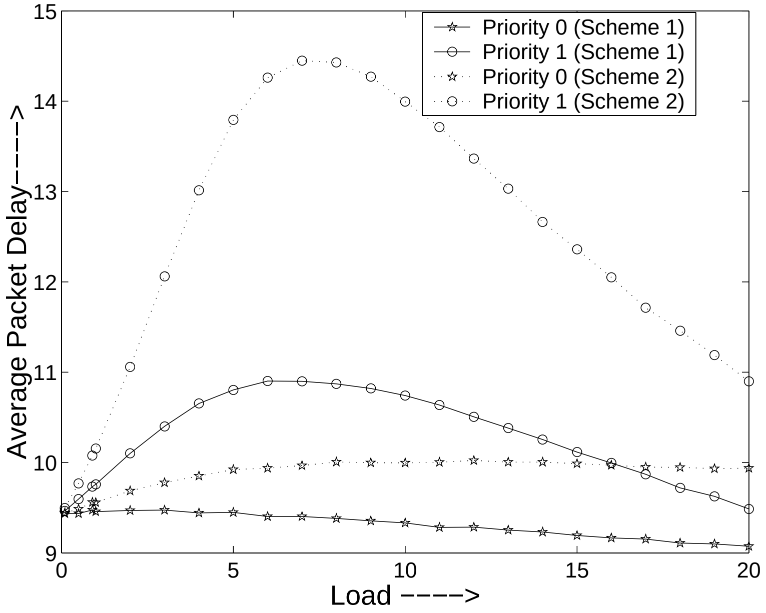

Figure 4.8: Average packet delay versus load using Scheme1 and Scheme2.

Figure 3.3 shows the 14-node NSFNET on which the simulation was implemented. The distances shown are in km. Figures 4.5 and 4.6 plot the packet loss probability versus load for high-priority (priority 0) and low-priority (priority 1) packets, using Scheme 1 and Scheme 2 respectively. Each shows packet losses for the case in which there is an equal amount of high-priority and low-priority traffic, and the case in which 20% of the traffic is high priority and 80% of the traffic is low priority. We observe that the loss of high-priority packets are lower than that for low priority packets. We also observe that Scheme 1 performs better than Scheme 2 at these loads; thus, at low loads, it is better to attempt deflection before segmentation when two bursts are of equal priority.

Figure 4.7 plots total packet loss probability versus load for different number of alternate deflection ports with 20% of high-priority and 80% of low-priority traffic. We observe that there is a significant improvement when we use two alternate deflection ports instead of one alternate ports, while there is less improvement from two to four alternate deflection ports. This result is due to the low nodal degree of NSFNET (Figure 3.3) and may differ for other networks.

Figure 4.8 plots total delay versus load with 20% of high-priority and 80% of low-priority traffic for the two QoS schemes. Scheme 2 has lower delays compared to Scheme 1, as Scheme 2 follows the segment-first approach rather than the deflect-first approach. The delay for high-priority bursts remains in a consistent range, while the low-priority bursts have higher delay due to multiple deflections. At very high load, bursts which are farther from their destination are less likely to reach their destination compared to those bursts which are close to their destination; thus, the average delay will eventually decrease as load increases.

4.4 Conclusion

In this chapter, we apply the concepts of burst segmentation and priority-based deflection routing algorithms to provide differentiated services in optical burst-switched networks. The high-priority bursts have significantly lower losses and delay then the low-priority bursts, and the policies which incorporate deflection tend to perform better than the policies with limited deflection or no deflection.

CHAPTER 5

CONCLUSION

Optical burst switches are the engines for the high-speed Internet transport on optical networks. Optical burst switching combines the advantages of packet switching and circuit switching in a single network. Data and control information are sent through different wavelength channels in a WDM system. When bursts and headers are sent separately on different channels, new protocols are necessary to avoid burst loss. We looked into different switching techniques, examined the optical burst switched network architecture and the different reservation policies.

We also looked into the prior art of contention resolution. We described the policies like optical buffering, wavelength conversion, and deflection routing (or hot-potato routing) to resolve contention.

In Chapter 3, we introduced the concept of burst segmentation for contention resolution in optical burst switched networks, and we investigated a number of different policies with and without segmentation and deflection. The segmentation policies perform better than the standard dropping policy, and offer the best performance at high loads. The policies which incorporate deflection tend to perform better at low loads, while deflection is not as effective at high loads.

In Chapter 4, we apply the concepts of burst segmentation and priority-based deflection routing algorithms to provide differentiated services in optical burst-switched networks. The high-priority bursts have significantly lower losses and delay then the low-priority bursts, and the policies which incorporate deflection tend to perform better than the policies with limited deflection or no deflection.

Areas for future work include developing an analytical model for an OBS networks with burst segmentation. Introducing the wavelength dimension and buffering into the simulation model will provide more options for contention resolution. The creation of the burst, i.e., burst aggregation, at the ingress routers is an interesting area of future work, wherein we can vary parameters like the burst length, base on a threshold or the burst aggregation time, base on a timeout, to shape the traffic into the network. Dynamic load balanced routing will be a welcome addition to the new policies.

To effectively evaluate the quality of service offered by various priority policies, a retransmission scheme for dropped packets could be implemented in order to measure end-to-end delay. A reasonable approach would be to implement a TCP layer on top of the optical burst switched layer. In such an implementation, it would also be useful to evaluate how TCP layer congestion control schemes react to and interact with various contention resolution schemes.

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] E. Haselton, “A PCM frame switching concept leading to burst switching network architecture,” *IEEE Communications Magazine, vol. 21, pp. 13-19, June 1983.*

[3] S. Amstutz, “Burst switching - an introduction,” IEEE Communications Magazine, vol. 21, pp. 36-42, Nov. 1983.

[4] D. K. Hunter et al., “WASPNET: A Wavelength Switched Packet Network,“ *IEEE Commun. Mag.*, Mar. 1999, pp. 120-29.

[5] S. Yao, B. Mukherjee, and S. Dixit, “Advances in Photonic Packet Switching: An Overview,” *IEEE Commun. Mag.*, Feb. 2000, pp. 84-94.

[6] M. Yoo and C. Qiao, “A Novel Switching Paradigm for Buffer-Less WDM Networks,” *IEEE Communications Magazine*, 1999.

[7] 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.

[8] J.S. Turner, “Terabit Burst Switching,” *Journal of High Speed Networks*, vol. 8, no. 1, pp. 3-16, 1999.

[9] C. Qiao, “Labeled Optical Burst Switching for IP-over-WDM Integration,” *IEEE Communications Magazine*, vol. 38, no. 9, pp. 104-114, Sept. 2000.

[10] M. Yoo and C. Qiao, “Supporting Multiple Classes of Service in IP over WDM Networks,” *Proceedings, IEEE Globecom ’99*, pp. 1023-1027, Dec. 1999.

[11] R. Ramaswami and K.N.Sivarajan, *Optical Networks: A Practical Perspective*, Morgan Kaufmann Publishers, 1998.

[12] C. Guillemot et al., “Transparent Optical Packet Switching: The European ACTS KEOPS Project Approach,“ *J. Lightwave Tech.*, vol. 16, no. 12, Dec. 1998, pp. 2117-33.

[13] A. Ge, F Callegati, L. Tamil “On Optical Burst Switching and Self-Similar Traffic,” *IEEE Communications Letters*, vol. 4, no. 3, March 2000.

[14] E. Varvarigos and V. Sharma, “The Ready-to-Go Virtual Circuit Protocol: A Loss-Free Protocol for Multigigabit Networks Using FIFO Buffers,“ *IEEE/ACM Trans. Net.*, vol. 5, Oct. 1997, pp. 705-18.

[15] I. Widjaja, “Performance Analysis of Burst Admission Control Protocols,“ *IEEE Proc. Commun.*, vol. 142, Feb. 1995, pp. 7-14.

[16] I. Chlamtac, A. Fumagalli, L. G. Kazovsky, et al., “CORD: Contention Resolution by Delay Lines,” *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 5, pp. 1014-1029, June 1996.

[17] Z. Haas, “The ‘Staggering Switch’: An Electronically Controlled Optical Packet Switch,” *IEEE Journal of Lightwave Technology*, vol. 11, no. 5/6, pp. 925-936, May/June 1993.

[18] D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, et al., “SLOB: A Switch with Large Optical Buffers for Packet Switching,” *IEEE Journal of Lightwave Technology*, vol. 16, no. 10, pp. 1725-1736, Oct. 1998.

[19] L. Tancevski, G. Castanon, F. Callegati, and L. Tamil, “Performance of an Optical IP Router Using Non-Degenerate Buffers,” *Proceedings, IEEE Globecom ’99*, Rio de Janeiro, Brazil, pp. 1454-1459, Dec. 1999.

[20] G. Bendeli et al., “Performance Assessment of a Photonic ATM Switch Based on a Wavelength Controlled Fiber Loop Buffer,“ *OFC’96 Technical Digest*, OFC, 1996, pp 106-7.

[21] W. D. Zhong and R. S. Tucker, “Wavelength Routing Based Photonic Packet Buffers and Their Applications in Photonic Packet Switching Systems,“ *J. Lightwave Tech.*, vol. 16, no. 10, Oct. 1998, pp. 1737-45.

[22] R. Ramaswami and K.N. Sivarajan, “Routing and Wavelength Assignment in All-Optical Networks,” *IEEE/ACM Transactions on Networking*, vol. 3, no. 5, pp. 489-500, Oct. 1995.

[23] G. Xiao and Y. Leung, “Algorithms for allocating wavelength converters in all-optical networks”, IEEE/ACM Trans. on Networking, vol. 7, no. 4, pp. 545-557.

[24] H. Zang, J.P. Jue, and B. Mukherjee, “A Review of Routing and Wavelength Assignment Approaches for Wavelength-Routed Optical WDM Networks,” Optical Networks Magazine, vol. 1, no. 1, Jan. 2000.

[25] S. L. Danielsen et al., “Analysis of a WDM Packet Switch with Improved Performance Under Bursty Traffic Conditions Due to Tunable Wavelength Converters,“ *J. Lightwave Tech.*, vol. 16, no. 5, May 1998, pp. 729-35.

[26] 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.

[27] 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.

[28] 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.

[29] G. Castanon, L. Tancevski, and L. Tamil, “Routing in All-Optical Packet Switched Irregular Mesh Networks,” *Proceedings, IEEE Globecom ’99*, Rio de Janeiro, Brazil, pp. 1017-1022, Dec. 1999.

[30] S. Yao, B. Mukherjee, S.J.B. Yoo, and S. Dixit, “All-Optical Packet-Switched Networks: A Study of Contention Resolution Schemes in an Irregular Mesh Network with Variable-Sized Packets,” *Proceedings, SPIE OptiComm 2000*, Dallas, TX, pp.235-246, Oct. 2000.

[31] M. Yoo, C. Qiao, and S. Dixit, “QoS Performance of Optical Burst Switching in IP-Over-WDM Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no.10, pp. 2062-2071, Oct. 2000.

[32] S. Yao, B. Mukherjee, and S.J.B. Yoo, “A Comparison Study Between Slotted and Unslotted All-Optical Packet-Switched Network with Priority-Based Routing,” *Proceedings, OFC 2001*, TuK2, Anaheim, CA, March 2001.

[33] V. Vokkarane, J. Jue, and S. Sitaraman, “Burst Segmentation: An Approach For Reducing Packet Loss In Optical Burst Switched Networks,” *UTD Technical Report UTDCS-20-01*, September 2001.