uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Salim Youssef Said
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A.Sc. (Electrical Engineering)
GRADE / DEGREE

School of Information Technology and Engineering
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Control and Survivability Issues for Optical Burst Switched Networks
TITRE DE LA THÈSE / TITLE OF THESIS

H. Mouftah
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

H. Elbiaze
CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

A. El Saddik
C. Huang

Gary W. Slater
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

Control and Survivability Issues for Optical Burst Switched Networks

by

Salim Youssef Said

Thesis

Presented to the School of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the degree
Master of Applied Science, Electrical Engineering

Ottawa Carleton Institute for Electrical and Computer Engineering
School of Information Technology and Engineering
University of Ottawa
Ottawa, Ontario, K1N 6N5
Canada

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-494-11401-0
Our file Notre référence
ISBN: 0-494-11401-0

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Dedicated

The memory of my grandmother,

Alia Toufic Saade

(1924-2004)

And to my family…

Abstract

Optical burst switching (OBS) is the most promising all-optical architectures for the next generation Internet. OBS is still being developed and several issues still need to be addressed. Several studies have stressed the importance of the choice of the *offset time* and its criticality in influencing the burst loss phenomenon in OBS networks. The main problem arises in estimating the initial delay that is compatible with the optical burst switching architecture implemented.

This thesis aims to provide a mechanism to calculate the *offset time* and to study its effect on the network performance and refining QoS levels. The study is extended to networks under failure to investigate the influence of the delay factor on the restoration process especially in compensating failed-network performance degradations. The simulation studies validated the proposed mechanism and formed an annex that can be integrated on any OBS restoration mechanism. A new QoS parameter is defined to meet service differentiation with respect to the data losses and delays.

Acknowledgements

I would like to thank my supervisor Professor Hussein Mouftah for his academic supervision and financial support throughout my master studies, which without, this work could have been impossible. I would also like to thank Professor Halima Elbiaze for here valuable comments and close follow up of my work.

Special thanks for OPNET for supporting me with a free license for OPNET Modeler. The simulator provided is a very powerful tool that helped in validating the work presented in this thesis.

Not to forget the support of my family and friends that has allowed me to reach this stage. Especially, I would like to thank my grandparents for their hospitality and encouragement which made me feel just like home.

Proverbs 18:15 –

An intelligent mind acquires knowledge, and the ear of the wise seeks knowledge.

Table of Contents

Abstract iii

Acknowledgement iv

Table of Contents v

List of Figures viii

List of Tables x

Acronyms xi

Chapter 1: Introduction

1.1 Background 1

1.2 Motivation and Objective 3

1.3 Thesis Contribution 4

1.4 Thesis Outline 5

Chapter 2: State of-Art OBS and Survivability Techniques

2.1 Introduction 7

2.2 OBS Architecture 9

2.3 Offset Time 13

2.3.1 Fixed Offset Time 13

2.3.2 Statistical Offset Time 14

2.3.3 Adaptive Offset Time 15

2.4 Network Survivability...16

2.4.1 Protection 18

2.4.2 Restoration 19

2.5 OBS Survivability 21

2.5.1 OBS Protection 21

2.5.2 OBS Restoration 24

Chapter 3: Offset Time and Design Issues

3.1 Introduction 27

3.2 Overview and Problem Statement 28

3.3 Offset Time Calculation Methodologies..31

- 3.3.1 Wavelength Routed OBS Network Offsets 31
- 3.3.2 Statistical Offsets 32
- 3.3.3 Fixed Offset 32

3.4 Just-In-Time Offset Time Calculation...33

3.5 Restoration of Link Failures 36

3.6 Failure Scenario 38

3.7 Remakes 42

3.8 The QoS-based Restoration Scheme 44

Chapter 4: OBS Network Modeling and Simulation

4.1 Introduction 46

4.2 OPNET Modeler Overview..47

4.3 Optical Burst Switching Network Model 48

- 4.3.1 Model Assumptions 49
- 4.3.2 Packet Objects 51
- 4.3.3 Link Object 53
- 4.3.4 Node Object 54
- 4.3.5 Operation Scenarios 61

Chapter 5: Simulation Results and Discussion

5.1 Simulation Assumptions 63

5.2 Burst Blocking Definitions..66

5.3 Simulation Result for ARPANET...67

- 5.3.1 Offset Time Extension Effect Under No Failure 68
- 5.3.2 Offset Time Extension Effect Under Failure 71

5.4 Simulation Result for Inter-Connected Ring Network 76

5.4.1 Offset Time Extension Effect Under No Failure 77
5.4.2 Offset Time Extension Effect Under Failure 78
5.5 Simulation Result for QoS Restoration Scheme 83

Chapter 6: Conclusion and Future Research

6.1 Summary and Concluding Remarks 85
6.2 Future Research...87

References 89
List of Publications 96

List of Figures

Chapter 2: State of-Art OBS and Survivability Techniques

Figure 2.1: Offset time in OBS 8

Figure 2.2: Generic OBS Network Architecture 9

Figure 2.3: OBS Functional Diagram 11

Figure 2.4: Offset Time 13

Figure 2.5: Restoration Time according to ITU-T recommendation M.495 20

Figure 2.6 Recovery vs. Performance 21

Chapter 3: Offset Time and Design Issues

Figure 3.1: OBS control and switching planes 28

Figure 3.2: Delays encountered at intermediate nodes 33

Figure 3.3: Burst losses versus time 38

Figure 3.4: OBS core network 39

Figure 3.5: Simple check algorithm 41

Figure 3.6: Different scenarios representing the position of the control packet and the data burst upon a link failure 43

Chapter 4: OBS Network Modeling and Simulation

Figure 4.1: OPNET Modeler Hierarchy 47

Figure 4.2: OBS Control Packet Structure 51

Figure 4.3: Teardown Message Structure 52

Figure 4.4: Correspondence between an optical link and its channels 53

Figure 4.5: Optical Burst Switching Core Node Model 55

Figure 4.6: OBS Source Process Model 57

Figure 4.7: OBS Source Child Process 57

Figure 4.8: OBS Source GChild Process 59

Figure 4.9: Sink Process Model 59

Figure 4.10: OBS Edge OSC Process Model 60

Chapter 5: Simulation Results and Discussion

Figure 5.1: ARPANET Topology 67
Figure 5.2: Overall BBP vs. network load at different offset time extensions 69
Figure 5.3: Overall BBP vs. time extension factor at different network load 69
Figure 5.4: Overall BBP vs. time extension factor at different OXC time 70
Figure 5.5: Burst Blocking due to Insufficient Time vs. Extension Factor 72
Figure 5.6: Global BBP vs. Restoration phase at Different Extension Values 73
Figure 5.7: Local BBP vs. Restoration phase at Different Extension Values 73
Figure 5.8: Overall BBP under failure vs. network load at different offset time extensions 75
Figure 5.9: Overall BBP under failure vs. time extension factor at different network load 75
Figure 5.10: Inter-Connected Ring Network's Topology 76
Figure 5.11: Overall BBP vs. network load at different offset time extension in the interconnected ring topology 77
Figure 5.12: Burst Blocking due to Insufficient Time vs. Extension Factor in an interconnected ring topology 78
Figure 5.13: Burst Losses under link failure versus the restoration phase at different network loads 79
Figure 5.14: Overall BBP under failure vs. network load at different offset time extensions when explicit setup scheme is used 80
Figure 5.15: Global BBP vs. Restoration phase at Different Extension Values when explicit setup scheme is used 81
Figure 5.16: Burst Losses under link failure versus the restoration phase at different network loads when explicit setup scheme is used 81
Figure 5.17: Link Utilization in absence of failures
    (a) Link Utilization at node "0" under no failures 82
    (b) Link Utilization at node "5" under no failures 82
Figure 5.18 Link Utilizations under link failure
    (a) Link Utilization at node "0" 83
    (b) Link Utilization at node "5" 83

List of Tables

Chapter 2: State of-Art OBS and Survivability Techniques

Table 01: Classification of Outage Time Impacts 17

Chapter 3: Offset Time and Design Issues

Table 3.1: QoS Level Computation Table 45

Chapter 5: Simulation Results and Discussion

Table 5.2: QoS Level Computation Table for ARPANET 84
Table 5.3: QoS Level Computation Table for an interconnected ring network ...84

Acronyms

ACK	Acknowledgement
BB	Burst Blocking
BBP	Burst Blocking Probability
BSN	Burst Switching Node
CS	Centralized Server
CP	Control Packet
DWDM	Dense Wavelength Division Multiplexing
FIS	Failure Indication Signal
FIP	Fault Indication Packet
FNM	Fault Notification Message
FSM	Finite State Machine
JET	Just-Enough-Time
JIT	Just-In-Time
LOBS	Labeled Optical Burst Switching
LSP	Labeled Switched Paths
LOS	Loss of Signal
MPLS	Multi-protocol Label Switching
OBS	Optical Burst Switching
OCS	Optical Circuit Switching
OXC	Optical Cross Connects
OPS	Optical Packet Switching
O/E/O	Optical-to-Electric-to-Optical
QoS	Quality of Service
RE	Routing Entity
SD	Signal Degradation
TM	Teardown Message
WROBS	Wavelength Routed Optical Burst Switching

Chapter 1

INTRODUCTION

1.1 Background

In the past decade, the Internet traffic has grown in an exponential manner, where data network capacities has surpassed voice network capacities [LIS00]. The optical technology was the viable transmission media pliable to that tremendous growth, especially after the advances in the Dense Wavelength Division Multiplexing (DWDM), where the transmission capacity per fiber is expected to exceed 1 Tbps in the near future.

Several measurements and statistical studies have shown that Internet traffic is self-similar where it exhibits a bursty nature at all time scales [LEL94]. Existing optical circuit switching (OCS) is not suitable for supporting bursty traffic. The bandwidth in circuit switching cannot be efficiently utilized if the data transmission does not have a long duration relative to the set-up time of the lightpath. The alternative to optical circuit switching is optical packet switching (OPC).

Optical packet switching is more efficient than OCS and it can be used to support IP and ATM augmented traffic. The switching operation in OPS is performed optically, where all the information stays in the optical form throughout the transmission. The user's data are transmitted by means of optical packets, in which a wavelength is only allocated to a packet when it is transmitted, and it can be reused by others after the transmission. However, OPS requires practical, cost-effective, and scalable implementations of optical buffering, optical header processing, and synchronization. With immature optical and photonic technology, these requirements are the major problems of this approach which are still several years away.

Optical burst switching (OBS) is the most promising all-optical architectures for the next generation Internet. Optical burst switching serves as a compromise between the current state of technology and the demands and fast growth of current networks. In OBS, the data bursts follow directly their control headers without waiting for reservation acknowledgments. Control headers are processed electronically at intermediate nodes; reserving resources for the coming data bursts and thus evading the need for optical buffers.

Leveraging the strengths of optical switching technologies, optical burst switching is a potential method by which future optical networks may use the available optical resources more effectively. OBS is still being developed and it has not been standardized yet [JUE05]. Several issues still need to be addressed as well as the assessment of the architectural aspects of implementing OBS.

1.2 Motivation and Objective

A momentous issue for optical burst switched networks is the ability to support a wide range of services, those being supported by IP networks today, and to cope with network failures. A single failure in optical networks can disrupt services for a large number of users, cause significant revenue loss for service providers, and reduce significantly the Quality of Service (QoS) in the network.

In order to support QoS in OBS networks, the network must be able to provide guarantees as well as differentiation with respect to data losses and delays. Several studies have stressed the importance of the choice of the *offset time* and its criticality in influencing the burst loss phenomenon in OBS networks.

The main problem arises in estimating the initial delay that is compatible with the optical burst switching architecture implemented. This thesis aims to provide a mechanism to calculate the *offset time* in Just-In-Time architecture and to study its effect on the network performance and refining QoS levels. The study will be extended to networks under failure to investigate the influence of the delay factor on the restoration process especially in compensating failed-network performance degradations. A new QoS parameter that meets certain burst blocking probabilities during a network failure is presented.

1.3 Thesis Contribution

The major goal of this thesis was to study the importance of the "*offset time*" value on the network operation and its effect on the network performance under normal and failure scenarios. An enhanced restoration mechanism and a new QoS parameter for OBS networks under link failures is proposed. The contributions of this thesis are summarized as follows:

- A refined offset time estimation for Just-In-Time architecture protocol.
- A study of the effect of the delay parameters, mainly the switch fabric configuration time, on the network's data loss.
- A basis to ease defining relative QoS levels in OBS networks.
- Verification of the importance of the offset time value on the restoration scheme used and a presentation of a methodology to best estimate the offset time to compensate any link failures.
- An improved restoration mechanism that takes into account the offset-time factor in compensating failed-network performance degradations by the effective and efficient use of resources.

- A new QoS-based restoration scheme that ensures specific levels of burst blocking probability upon a link failure.

1.4 Thesis Outline

This thesis is organized as follows. Chapter 2 presents an introduction to optical burst switching, describing its operation and the various current architectures. The concept of network survivability is discussed, describing protection and restoration mechanisms. Then a literature survey is conducted with the up-to-date techniques in achieving survivable optical burst switched networks.

Chapter 3 presents various techniques in finding the offset time, a time interval that separates the transmission of the data burst and its corresponding control packet. The effect of the offset time value on the network performance in terms of blocking probability and meeting QoS parameters is studied. A new methodology to calculate the offset time and present a structure to define QoS levels in Just-In-Time architectures under normal network operation is proposed. The study is extended to networks under link failures and shows the effect of the offset time value on several restoration mechanisms proposed in literature and its role in the recovery process. An enhanced restoration mechanism for OBS networks under link failures and a new QoS parameter that meets certain data loss levels is presented.

Chapter 4 describes the simulator used and provides an overview of the model built to simulate the core of an optical burst-switched network, describing the modeling assumptions and the detailed behavior of the modeled system. In Chapter 5, we present the results conducted in the simulations and validate the proposal described in chapter 3. Chapter 6 concludes this thesis and identifies suggestions for future work.

Chapter 2

STATE-OF-ART OBS AND SURVIVABILITY TECHNIQUES

2.1 Introduction

Optical burst switching (OBS) is the ultimate balance between optical circuit switching (OCS) and optical packet switching (OPS). OBS combines the advances of both techniques and overcome the current technology limitations. OBS is capable of transporting variable-sized units, called bursts [QIA99], over the optical infrastructure, ranging from long bursts as in OCS to short bursts as in OPS. Before the transmission of a burst, a correspondent control packet is sent on a separate channel to reserve the resources along the path to the destination. The control packet goes optical-to-electric-to-optical (O/E/O) conversions at the intermediate nodes taking advantage of the fast electronic processing and refuting the need for optical processing needed in OPS.

As shown in Figure 2.1, the time lag between data burst and the control packet, defined as the *offset* time, should be enough to process the control packet and configure the optical cross connects (OXC) along the path. This will ensure that the data burst will not need to be buffered at intermediate nodes. As the control packet approaches the destination, the *offset* time duration becomes smaller along with the separation time between the control packet and the data burst.

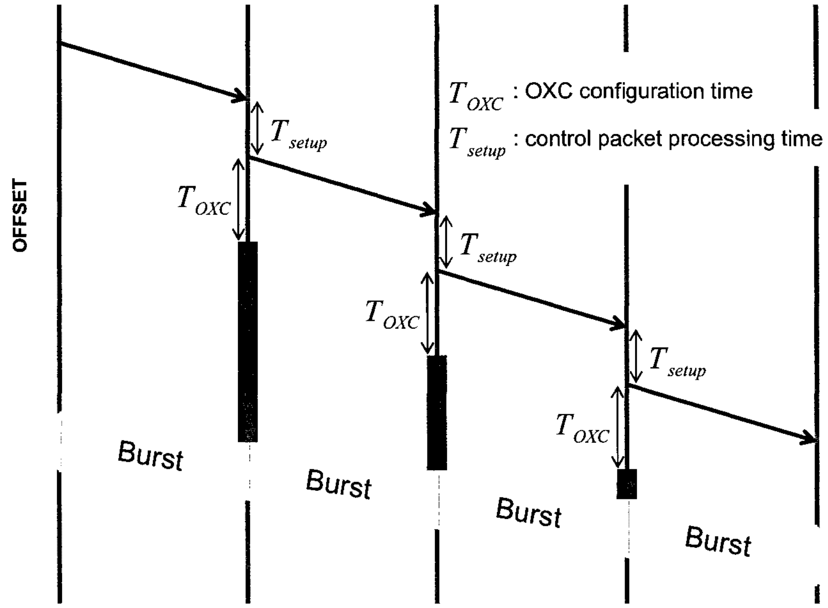


Figure 2.1: Offset time in OBS

The one-way reservation used in OBS overcome the extra delay encountered in OCS and provides a faster and more efficient solution with a higher degree of statistical multiplexing [ILI02]. OBS varies based on how the network resources are reserved and released. The next section review the OBS architectures proposed in literature.

2.2 OBS Architecture

An optical burst switching network consists of burst switching nodes (BSN) interconnected by fiber links. Each link consists of a number of wavelength channels, some reserved for exchanging control information and others for transmitting data. Note that the control channels can be physically on the same fiber link or on separate links. The BSN can be either an edge or a core node. As shown in Figure 2.2, an edge node connects various network elements such as IP routers, frame relay switches, ATM switches, etc. to core nodes in an OBS network frame. As a result, the OBS network architecture can be portioned into two parts, an OBS network edge and an OBS network core.

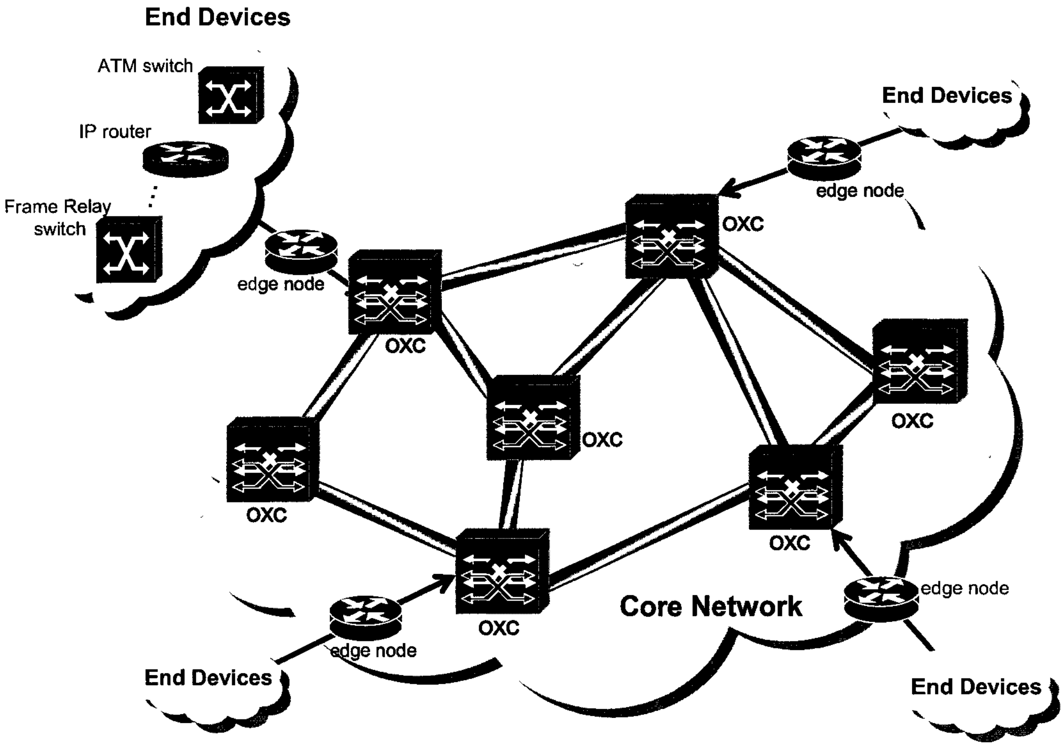


Figure 2. 2: Generic OBS Network Architecture

At the edge of the OBS network, an edge node collects the incoming data packets such as IP packets, ATM cells, etc. and stores them in electronic buffers according to their destination and class of service. A burst assembly mechanism places these packets into bursts based on an assembly policy. Assembly algorithms are characterized by three parameters: preset timer or threshold, maximum burst length and minimum burst length.

The preset timer or threshold is responsible for triggering the creation of the data burst. This parameter is very important since it controls the characteristic of the data burst arrival to the network. The maximum and minimum burst size shapes the size of the data bursts. These parameters are usually calculated based on the processing times at intermediate nodes and the ratio of the control channels to the data channels in the fiber [GE00]. In case there is not enough data to be aggregated, bit-padding is used to assemble the minimum burst size [BAT04]. After the burst assembly process, the data burst is ready for transmission to the desired destination. At egress edge nodes, a dissembler is used to break down the burst into the original packets. Afterwards, the packets are forwarded to the desired destinations.

Another edge node task is the routing and wavelength assignment. Three major approaches are used to route traffic in OBS networks. The first approach uses hop-by-hop routing, as in IP networks, where each node maintains a forwarding table. The second approach maps the technique used in MPLS [QIA00] while the third approach uses pre-calculated routes stabled via RSVP or CR-LDP [BAT03].

OBS is characterized as a one way signalling protocol, i.e. the user does not wait for a positive acknowledgment before sending the data. To ensure proper functioning, a control packet should be sent ahead of the data burst to reserve the optical resources along the path. The signalling and reservation mechanism can be either centralized or distributed.

Centralized OBS signalling protocol, which is known as wavelength routed optical burst switching (WR-OBS) [DUE01] [DUS02], is a two-way reservation mechanism. A centralized server (CS) is responsible for resource allocations and burst transmission scheduling. Upon the arrival of a data burst, the edge node sends a control packet to the CS, which has a global knowledge of the network. After control packet processing, the centralized server allocates a route and a free wavelength channel for the incoming request and acknowledges the edge node. Upon the arrival of the positive ACK, the edge node transmits the burst.

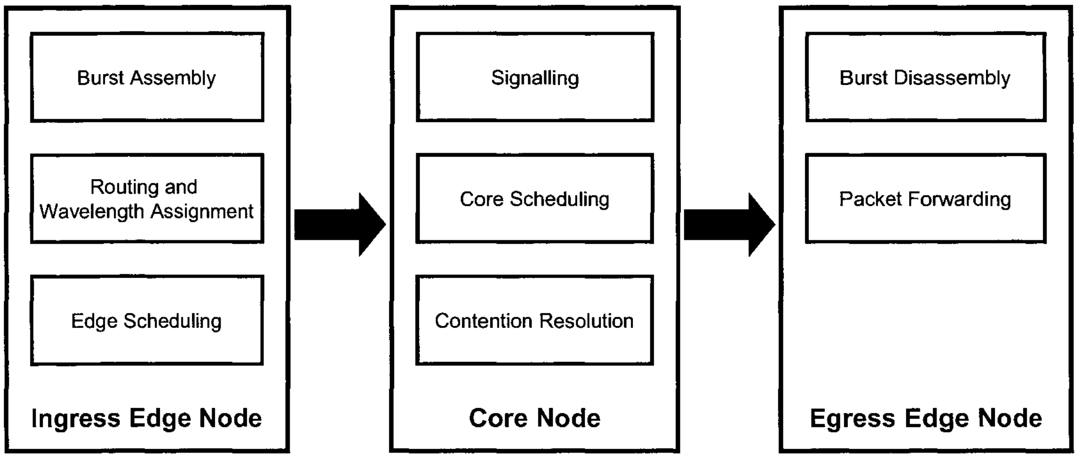


Figure 2.3: OBS functional diagram [JUE05]

In distributed OBS networks, the core node takes the responsibility of signalling and resource management. The functionalities of the OBS network elements is summarised in Figure 2.3. Several reservation schemes have been proposed for OBS networks, such as Just-Enough-Time (JET) [QIA99], Horizon[TUR00], Just-In-Time (JIT) [WEI00] [BAL02].

The variants of the signalling schemes differ in how soon the resources are allocated and released about the occurrence of the data burst. Some schemes, such as JIT, use explicit or immediate setup mechanism. Upon the arrival of the control packet at the core node, the appropriate wavelength is reserved and necessary port configurations are done. On the other hand, in estimated setup mechanisms (as in JET and Horizon) the control packet carries information about the burst length and the *offset* time. This information is necessary to estimate the arrival time of the first bit of the data burst. The core-node's scheduler performs reservation actions based on the expected arrival time. Estimated setup schemes outperform explicit setup schemes in terms of burst loss probability, but they require more complex schedules and are harder to implement. Similarly, resource release can be either explicit or estimated. In explicit release schemes, a teardown message is sent after the burst to release the reserved wavelengths. On the other hand, estimated release schemes approximate the end of the data burst and free resources after the transmission of the burst's last bit. Unlike estimated release schemes, explicit release schemes results in lower bandwidth utilization and increased message complexity.

2.3 Offset Time

An important trait of OBS networks is the separation of the control and data planes. One dimension of this separation is "time". As shown in Figure 2.4, an *offset* time separates the transmission of the control packet and its corresponding data burst. This separation is very critical in terms of insuring proper functioning of OBS networks. During this time lag, all intermediate core nodes will have the time to process the control packet and perform necessary resource reservations. As the data burst arrive, it can cut through without the need for any buffers.

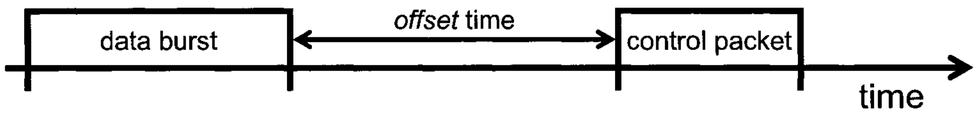


Figure 2.4: Offset Time

As the OBS architecture and signaling mechanism changes, the offset time calculation rules vary. These variations can be classified into three categories: fixed *offset* time, statistical *offset* time, and adaptive *offset* time.

2.3.1 Fixed Offset Time

The commonly used method to calculate the offset time was proposed with the Just-Enough-Time (JET) OBS control protocols. Each intermediate core node is characterized by a set of delays. These delays represent control packet queuing and processing delays and the optical cross connect configuration time [QIA99] [WEI00] [BAL02]. Obviously, the offset time should be greater than or equal to the sum of all the delays encountered at intermediate core nodes between the source and the destination. Depending on the routing mechanism used, the number of hops between the source and destination could not be exactly known. Therefore, the ingress edge node has to estimate the number of hops with some relaxation factor to ensure enough time is granted before the arrival of the data burst.

In the centralized WR-OBS architecture, and as discussed in Section 2.2, a centralized request server is responsible for the control/signaling mechanisms. The *offset time* is calculated as the sum of the time the request waits in the queue, the computation time of the routing and wavelength allocation algorithm, and the round trip propagation delay between the ingress node and the centralized request server. The propagation delays vary with respect to the position of the node relative to the request server. In order to control the *offset* time value, the queuing delay for a connection request is bounded by a maximum threshold, t_{sched}, which is the maximum time allowed for a connection request to remain in the queue. This value is set by the network administrator depending on the traffic level in the network.

2.3.2 Statistical Offset Time

This scheme was proposed in OBS networks using MPLS [QIA 00]. A leaky bucket regulator is responsible for generating tokens according to some random process where no token buffering is allowed. As soon as a burst is assembled, its corresponding control packet is immediately launched, and the burst is delayed by an *offset time* equals to the time difference between the current time and the arrival time of the next token according to the underlying random process. Note that if the transmission time of a previous burst overlaps with the arrival time of a coming token, the token can not be grabbed. This variable *offset time* imposes a statistical domination property and thus regulates the average rate at which data bursts are released into the network. It has been shown in [BLU 99, AWD 99] that this *offset time* is useful to handle traffic engineering and QoS provisioning purposes.

2.3.3 Adaptive Offset Time

Adaptive *offset* time mechanism takes advantage of the OBS network state to acclimatize the value of the *offset* time. Under high traffic loads and in the absence of buffering capabilities and wavelength conversion at core nodes, incoming bursts may compete for the same resources. In such cases, the resource is contended and one of the bursts will be dropped. One possible solution is to deflect one of the contending bursts on an alternative route. This will result in a change in the route length and congestion level in the network. As a result, the *offset* time needs to be adjusted to reflect the bandwidth utilization of the links and the network's node traffic state.

A sliding window is used to measure the average change in either the route length between all possible source-destination pairs or the average load change at the active network nodes. During a window frame, the change level is recorded and suitable time extensions are added to the *offset* time value.

2.4 Network Survivability

With the continuous growth in the volume of traffic transported by the network infrastructures, network reliability is becoming a serious issue in the network design. A network failure may cause fatal impairment on the service integrity to innumerable end users [MOU02]. It is often revenue loss and business disruption that are first in mind, but harmful complications from a number of network dynamic effects also could occur. Table 2.1 summarizes these effects, based on [GRO03].

Today and upon designing a network, operators must develop a survivability concept. A network is considered survivable if it can maintain service continuity to the end users [ZHO00]. The included survivability strategies must be able to cope with current and future network sizes. Two possible failures can occur in a network, node and link failures. Node failures are rare to happen, since optical cross connects are robust and equipped with duplicates to overcome any possible hardware breakdowns. On the other hand, link failures are more likely to happen due to man-made errors, and uncontrollable natural phenomena.

Network survivability can be achieved through protection and restoration. Protection is an operation performed in advance of a failure to defend the network against any possible disruption. On the contrary, restoration is conducted after the failure occurs to recover the affected traffic and reinstates service continuity.

Target Range	Duration	Main Effects / Characteristics
Protection Switching	< 50 ms	No outage logged: system reframes, service "hit", 1 or 2 error seconds (traditional performance spec for APS systems), TCP recovers after one corrupted frame, no TCP fallback. Most TCP sessions see no impact at all.
1	50 ms – 200 ms	< 5% voice-band disconnects, signaling system (SS7) switch-over, SMDS (frame-relay) and ATM cell-re-routing may start.
2	200 ms – 2s	Switched connections on older channel banks dropped (CGA alarms) (traditional max time for distributed mesh restoration), TCP/IP protocol backoff.
3	2s – 10s	All Switched circuit services disconnected. Private line disconnects, potential data session / X.25 disconnects, TCP session time-outs start, webpage not available errors. Hello protocols between routers begin to be affected.
4	10s – 5 min	All calls and data sessions terminated. TCP/IP application layer programs time out. Users begin attempting mass redials / reconnects. Routers issuing LSAs on all failed links, topology update and resynchronization beginning network wide.
"Undesirable"	5 min – 30 min	Digital switches under heavy reattempts load, "minor" societal/business effects, noticeable Internet "brownout".
"Unacceptable"	> 30 min	Regulatory reporting may be required. Major societal impacts. Headline news. Service Level Agreement clauses triggered lawsuits risks: 911, travel booking, educational services, financial services, stock market all impacted.

Table 2.4.1: Classification of Outage Time Impacts

2.4.1 Protection

Protection is a lower layer mechanism that provides a first level of defense against common faults, such as fiber cuts. Protection is topology and technology specific and offers fast recovery. A fixed amount of capacity is dedicated for protection purposes in order to make a fast transfer of traffic from failed links to the protection links. Depending on how this pre-assigned capacity is used, one can distinguish between dedicated or shared protection mechanisms.

There exist three types of protection: 1+1, 1:1, and 1:N. In 1+1 protection, the traffic to be protected is simultaneously sent over two parallel paths. During normal operation, the destination receives the two traffic streams and selects one of them. In case of failure among the chosen path, the destination simply switches onto the other path. No protection signaling is required because the destination can handle a failure by itself, and the source node does not have to do anything else but always copying the traffic onto the alternative path. This makes 1+1 protection very simple to implement, and the achievable restoration time is very short. The disadvantage of 1+1 protection is the waste of bandwidth.

In 1:1 protection scheme, two parallel paths are also used. However during normal operation, there is no traffic sent across the alternative path. Only in the case of failure along the working path, both the source and the destination switch to the protection path. The disadvantage of 1:1 protection is the overhead needed for signaling. As a result, it will lead to a slower restoration time than that in 1+1 protection. In order to achieve higher network utilization, low priority traffic is transmitted on the protection path. In the case of failures, this traffic is dropped to the advantage of the high priority traffic.

The last protection technique is the 1:N. This technique can only handle a single failure. In case of multiple failures, the signaling protocol switches the traffic of the higher priority to the protection path. This scheme can be extended to M:N, where M represents the working paths and N is the number of protection paths.

2.4.2 Restoration

Restoration provides protection against network failure in the second step. Typically, restoration can handle not only link failures but also node or multiple concurrent failures, as opposed to protection. Restoration might be implemented in a centralized or distributed approach. In both cases, a network failure must be detected locally and then must be propagated to the control element controlling the restoration procedure. Distributed protection restores failed services faster than centralized protection.

Restoration speed is a key design criterion when developing a survivability concept. The restoration time is defined as the time elapsed between failure detection and connection restoration. The ITU-T recommendation M.495 "Maintenance International Transmission Systems" specifies how the restoration time is calculated as shown in Figure 2.5.

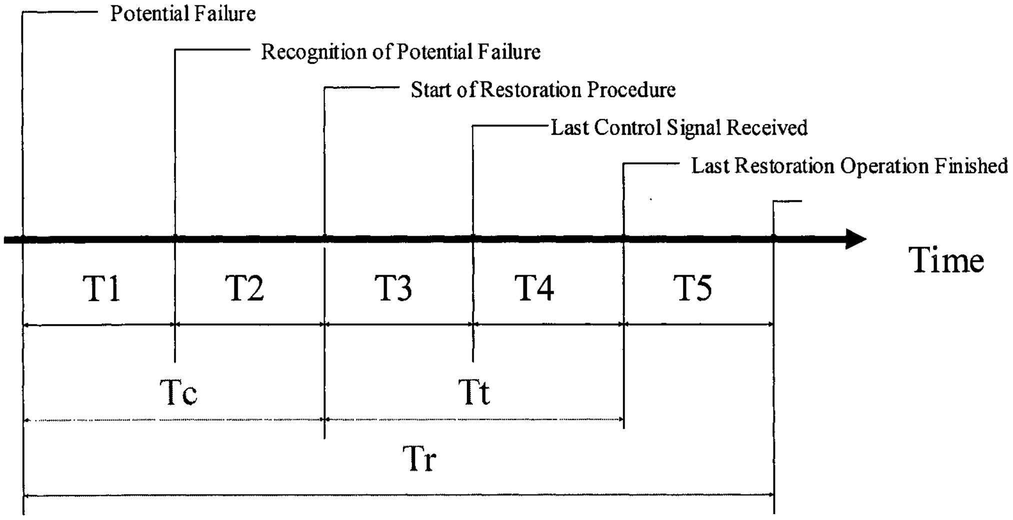


Figure 2. 5: Restoration Time according to ITU-T recommendation M.495

In the case of failure, the network node next to the failure detects the failure by either Loss of Signal (LOS) or Signal Degradation (SD) and triggers the restoration process after the failure is confirmed. During the restoration procedure, control signals are transmitted to all nodes in the network. Upon global knowledge, the protection switching operation or some re-synchronization is initiated after being verified.

High speed recovery time usually comes at the cost of more equipment use and thus higher network cost. The other dimension is the flexibility of our design. The three dimensions are of cross-conflicting nature, as shown in Figure 2.6, and represent a challenge to network designers.

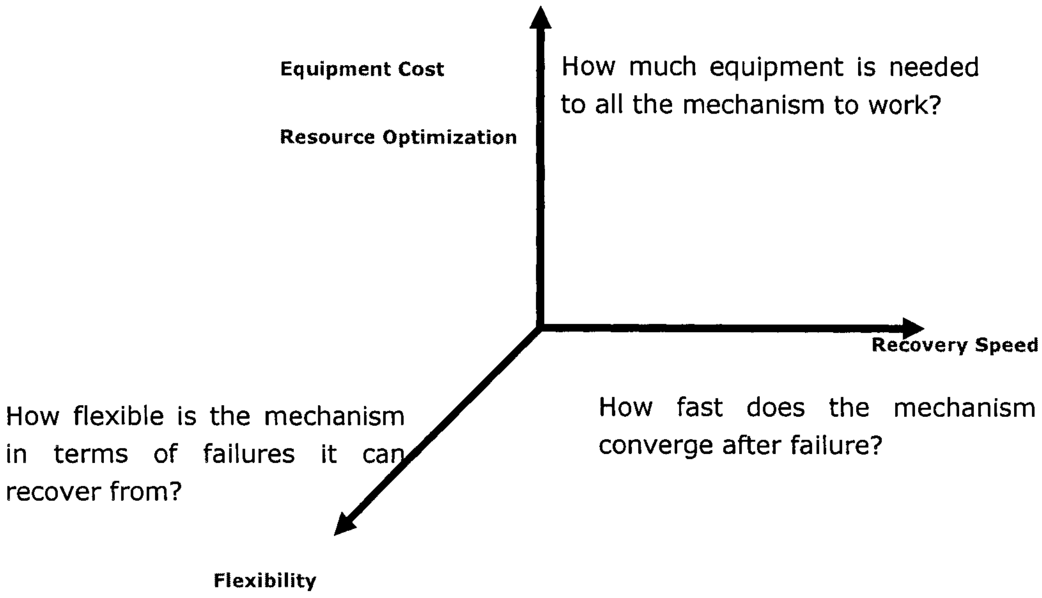


Figure 2.6: Recovery vs. Performance

2.5 OBS Survivability

As discussed in Section 2.4, an important issue in OBS networks design is survivability. Upon a network failure, large amount of data will be lost and degradation in the network. Few studies have considered the issue of survivability in optical burst switching networks. This section presents a literature review of the work done in this area.

2.5.1 OBS Protection

Protection schemes provide a higher degree of survivability, but consumes greater amount of resources. There is little work done on OBS protection schemes

[QIA00] [QIU05] [GRI03]. The protection schemes are based on methods used in multi-protocol label switching (MPLS) networks.

Protection schemes in labeled optical burst switching (LOBS) networks are provided through the establishment of redundant labeled switched paths (LSP) for each working path [QIA00]. A source node initially sends bursts over only one of the LSPs. When a failure occurs, the source node is notified of the failure and begins sending bursts over the backup LSP. In this scheme, no additional resources are required in the network other than the LSP entries at the label-switching routers along the backup path. Idle resources for backup routes can be used to carry lower-priority traffic and thus improving the network level utilization. This technique can also be extended to other protection schemes such as 1+N and N:1. In the 1+N case, a primary LSP is protected by multiple backup LSPs, where each backup LSP carries a faction of the traffic from the original LSP.

In [GRI03] the authors propose a 1+1 protection architecture for OBS networks with JIT architecture. In the 1+1 OBS protection, two disjoint LSP routes are determined for each burst session. The ingress node duplicates incoming packets and sends one copy on each of the label-switched paths. Thus, if a link fails on one LSP, the burst will continue to be received on the other LSP. The authors discuss the delay mismatch between the two disjoint LSP paths and propose constraint routing to reduce the mismatch and buffering at the ingress node of the leading path.

The advantage of this 1+1 protection is that no additional actions are required in the OBS network in order to recover from a failure. The disadvantage is that the scheme uses at least twice as many resources as the unprotected case. Also, the destination must be able to eliminate redundant bursts. The concern of this scheme is the fault detection and localization. The rigid data framing format with associated overhead monitoring bytes used in SONET/SDH networks is not appropriate for labeled OBS networks due to its bursty nature. A proposed solution is the monitoring of the control channel since the control packets are terminated at each hop. This solution is only good for fiber cuts and not applicable to channel failures.

[QIU05] presents a survivable OBS network using a multiple ring approach. The network is divided into rings such that the traffic between any node pair is restricted within one ring only. Under normal network operation, Keep Alive packets are sent periodically between the communicating entities. Upon a link failure, the downstream node detects the failure by Loss of Light while the upstream node detection relies on the absence of the Keep Alive packets. The upstream node sends a Fault Indication Packet (FIP) around the ring to announce the failure which upon all the traffic is switched to the protection path on the ring.

This method takes into account the difference in the number of hops between the working and protection path and adapt the control packet fields (wavelength, offset) to these changes. The disadvantage of this method is the complexity in finding the different rings in the mesh network and the route within the ring may not be optimal. Also this method will lead to high losses under high loads and in the absence of wavelength converters.

2.5.2 OBS Restoration

With the advances in the optical burst switching techniques, shorter burst sizes will be overwhelming. As a result, protection schemes will be inapt to be implemented due to the low resource utilization and incapability to support QoS requirements. Restoration is more effective technique in terms of resource utilization, but requires longer restoration time. Restoration offers better scalability in terms of fault management.

All the restoration schemes presented in the literature considers single link failures. These schemes can be categorized based on the OBS control architecture implemented.

In centralized OBS networks, [KOZ03] presents a restoration scheme for WR-OBS networks. Upon a link failure, the centralized request scheduler is notified and the routing tables are updated using adaptive shortest path routing such that the consecutive bursts are routed around the failed link. As discussed in Section 2.2, WR-OBS request schedulers are associated with a timer that controls the time the call request wait in the queues. In the case of a link failure, the timer is incremented by a failure-compensating delay to adapt to the failure. This schemes satisfies QoS end-to-end delay requirements and tend to improve the overall burst loss probability.

In distributed control OBS networks, the restoration schemes use deflection routing to coop with link failures. [XIN04a] [XIN04b] presents a fast restoration scheme based on deflection routing. When a link fails, the nodes adjacent to the failure have to notify the element(s) responsible for routing in order to update the global routing tables. Meanwhile, these nodes will deflect the traffic around the failed link. The authors investigate between local and distributed deflection schemes. In local deflection, the node adjacent to the failure takes the initiative to deflect the incoming burst on a pre-calculated alternative route. In distributed deflection the node adjacent to the failure sends a fault notification message (FNM) to all its neighbouring nodes. The FNM contains the destination information for all the primary routes passing through the faulty link. This scheme distributes the re-routed traffic all over the network and thus decreases the chances of network congestions.

This scheme suffers at high network from deterioration of the congestion level in the network. In addition, the capacity on the links between the node adjacent to the failure and its neighbouring nodes is not utilized. To overcome this problem, [XIA04b] proposed a deflection ratio, α, to determine the portion of affected burst that should be deflected at the adjacent nodes. A preliminary and important factor that was not taken into consideration is the offset time issue. In cases where the alternative route is longer than the primary route, in spite of any restoration mechanism used, the burst might be dropped since the control packet will not have enough time to reserve resource ahead of the data burst.

[LEE04] tackles this problem partially. The authors present a hybrid restoration scheme employing a combination of sub-path and path restoration schemes to recover single link failures in OBS networks. Upon a link failure, the upstream node adjacent to the failure knows first, and thus takes up the initial step and employs a sub-path restoration. Afterwards, the upstream node sends a failure indication signal (FIS) to the source node. As the FIS reaches the source node, it takes over the responsibility from the upstream node and employs path restoration.

During the sub-path restoration state, the node adjacent to the failure checks the offset time of the control packets on the failed link. If the offset time is long enough for a new control packet to be generated, this replacement packet can use the backup path. If the offset time is very short, incoming data burst will be lost. This scheme only considers a solution to restore the data bursts that have lost their control packets, and does not discuss the offset time issue for control packets that arrive before their ingress node is aware of the failure. In this case all the bursts using longer alternative routes will be dropped. Another disadvantage of this scheme is the need to record the offset time of all the control packets that traverses a particular node.

Chapter 3

OFFSET TIME AND DESIGN ISSUES

3.1 Introduction

The purpose of this work is to study the importance of the "*offset time*" value on the network operation and its effect on the network performance under normal and failure scenarios. In the previous chapter, the state-of-the-art of optical burst-switched networks and the various OBS survivability approaches proposed in literature was presented. This chapter presents different approaches in calculating the offset time based on the OBS architecture, with the focus on the JIT-based optical burst-switched networks. The importance of the *offset time* estimation mechanism on the proper functioning of the network is discussed plus its significance in supporting service differentiation with respect to loss and delay, and thus defining relative QoS levels in optical burst switched networks. This study is extended to networks under failures, and an improved restoration mechanism that takes into account the *offset-time* factor in compensating failed-network performance degradations is proposed.

3.2 Overview and Problem Statement

An intrinsic feature of the OBS is the separation of the control and data planes, which facilitates the electronic processing of the control messages at optical core routers and provides an end-to-end transparent optical path for transporting data bursts. The core of an optical burst switched network, as shown in Figure 3.1, can be envisioned as two networks coupled together: a pure optical network consisting of optical switching fabrics responsible for transferring information in light-form, and a control network consisting of smart routing elements responsible of all the control and management issues in the network.

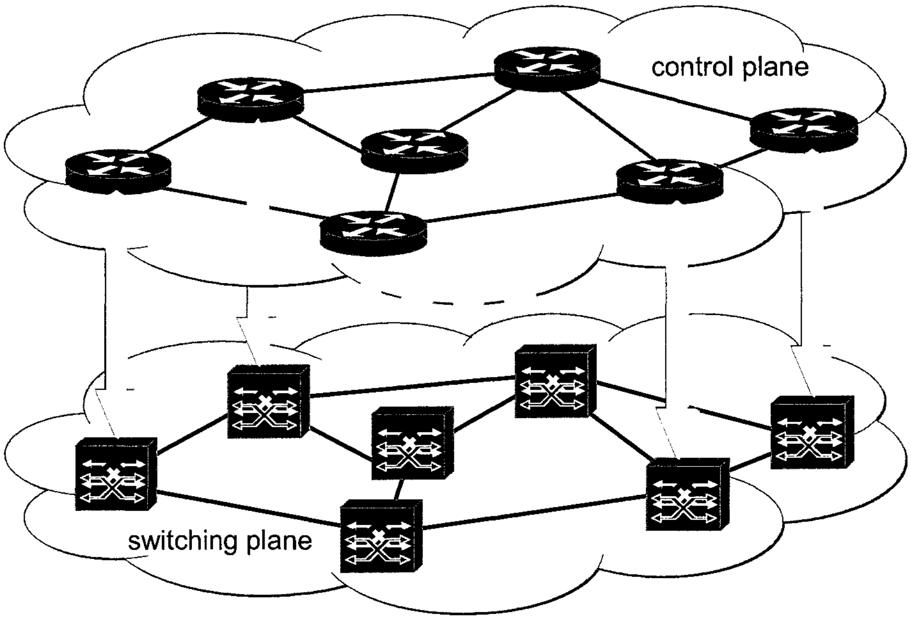


Figure 3.1: OBS control and switching planes

As described in Chapter 2, the transfer of data across the network is achieved by sending a control message ahead of the data burst in order to reserve the resources along the path. Each control message stops at intermediate nodes to be processed, and appropriate commands are given to the switching entity to be configured so that the data burst can cut-through without the need for intermediate buffering. Having enough time between the transmission of the control message and its corresponding data bursts allows the network to function properly.

The time lag between the control message and the data burst is persuaded by a series of delays necessary in the OBS operation. These delays can be classified according to the network planes, as in Figure 3.1; control-plane delays and optical-plane delays.

The main sources of delay the control packet encounters at the intermediate controlling entities are the queuing and processing time delays. The queuing delay represents the time the control packet has to wait in a queue in order to be processed. Usually this value is connected to the level of congestion in the network, and in most cases is not considered, and stills an open area of research. The processing delay is the time necessary to handle the control packet on the network system, where the control packet fields are read and necessary information is extracted from them. The value of the processing delay in JIT-based architecture is in the range of 12 μsec and is expected to reach 50 nsec in the near future.

The only delay encountered in the optical plane is the optical-cross-connect (OXC) configuration time. This time represents the time necessary to stabilize the optical cross connect on a specific wavelength between the appropriate input-output ports after it receives a command from the controlling entity. This time is technology dependent, currently ranging in milli-seconds and is expected to decrease dramatically (to micro-seconds) with the advance in photonic switching techniques.

Other sources of delays are the transmission time and the propagation time delays. These delays are of an insignificant effect on the *offset time* calculations. Usually the size of the control message is relatively small to the size of the data burst, and thus the transmission time (in the micro-second range) of the control header is neglected and assumed to be zero. The control message and the data burst traverses the same links from given source to any desired destination, therefore, these packets encounter the same propagation delays and thus it is irrelevant to enter this parameter in the *offset time* calculations.

The main problem arises in estimating the initial delay that is compatible with the optical burst switching architecture implemented. While the exact mechanism to perform and refine this estimation remains a topic for further investigation, we will study the effect of the *offset time* value on the network performance and refining QoS levels. In the next section, a summary of the various techniques used to determine the *offset time* between the transmission of the control headers and the data bursts, based on the architecture used is presented.

3.3 Offset Time Calculation Methodologies

There are variations in the OBS literature on how exactly to determine the pre-transmission offset time. This section describes the various methodologies used, mainly in wavelength routed optical burst switched networks, OBS networks using MPLS, and Just-Enough-Time based architectures.

3.3.1 Wavelength Routed OBS Network Offsets

In the WR-OBS architecture, a centralized request server is responsible for the control/signaling mechanisms. The ingress node sends a request to the request server for a connection. The connection request is queued until its turn to be processed, where a route is calculated from the source to the corresponding destination and the requested number of wavelengths is reserved [DUS02]. After the ingress node receives a confirmation message from the request server, it begins data transmission. The *offset time* is calculated as the sum of the time the request waits in the queue, the computation time of the routing and wavelength allocation algorithm, and the propagation delay on the way to and the way back between the ingress node and the centralized request server. Note that the propagation delays vary with respect to the position of the node to the request server. In addition, the queuing delay for a connection request is bounded by a maximum threshold, usually referred to t_{sched}, which is the maximum time allowed for a connection request to remain in the queue. This value is set by the network administrator depending on the traffic level in the network.

3.3.2 Statistical Offsets

This method imposes a statistical domination property and regulates the average rate at which data bursts are released into the network. As discussed in Chapter 2, as soon as a burst is assembled, its corresponding control packet is launched, and the burst is delayed by an *offset time* equals to the time difference between the current time and the arrival time of a token that is generated by a leaky bucket according to some random process.

3.3.3 Fixed Offsets

The commonly used scheme, which was proposed with the Just-Enough-Time (JET) OBS control protocols, computes an *offset time* equals to the sum of the total delays encountered at the intermediate nodes between the source-destination [QIA99]. It is assumed that the exact number of hops in the network is given by the routing protocol at the edge of the network. Thus the *offset time* equals the total processing delays at all the intermediate OBS hops plus the switch fabric configuration time of the egress OBS node. Usually the processing and OXC configuration time at each node to be identical, but in practice these time durations may vary from node to node depending on the technology used (processor's speed, switching technology, etc.) and the switching entity hardware implementation.

3.4 Just-In-Time Offset Time Calculation

Define the *offset time*, *δ*, as the time necessary to reserve resources along the path between the source and the destination before the arrival of the first bit of the data burst. In JIT-based OBS networks, the main sources of delay are the processing time (T_{proc}), the optical cross-connect configuration time (T_{oxc}), and the queuing delay the control messages encounter at the intermediate nodes (T_{queue}), as shown in Figure 3.2.

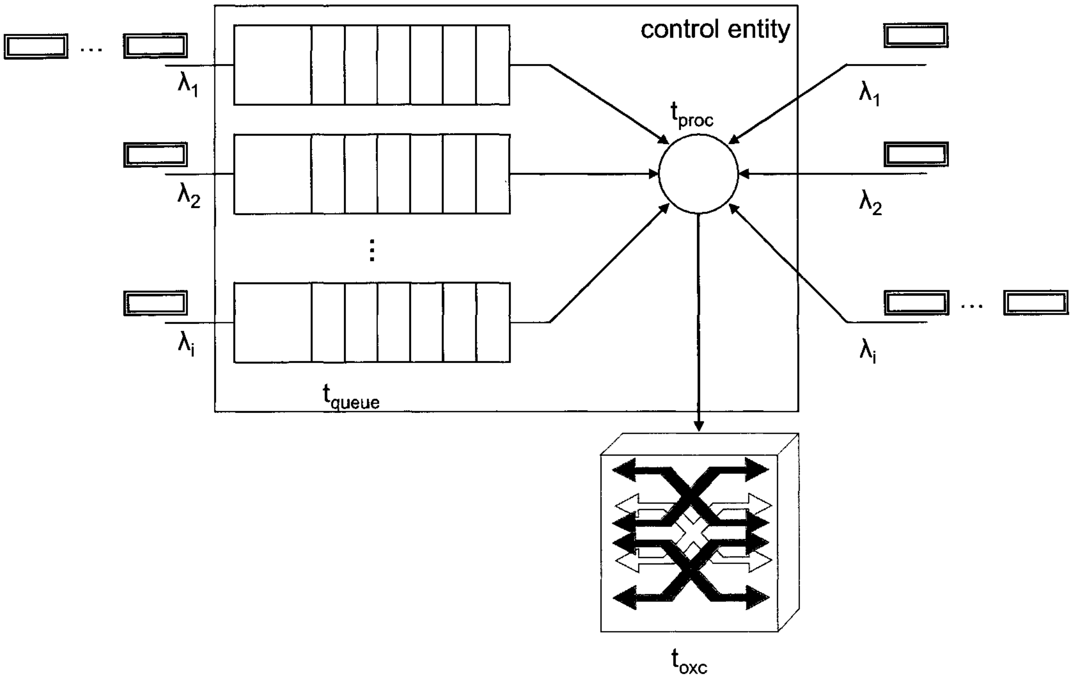


Figure 3.2: Delays encountered at intermediate nodes

As discussed in the previous subsection,

$$\delta \geq \varepsilon + \sum_{i=1}^{v} \varphi(i) \quad (3.1)$$

where $\varphi(i) = T_{proc} + T_{oxc} + T_{queue}$;

ν = is the total number of hops along the path.

ε = safety guard to compensate the uncertainty in the T_{proc} and T_{oxc} values

Let $\bar{T}_{queue}$ be the average queuing delay each controlling entity encounter in the network. Assuming that the processing and the switch fabric configuration time are identical on all the nodes in the network, we can rewrite Equation 3.1 as follows:

$$\delta \geq (\nu * \overline{\varphi}) + \varepsilon \quad \text{.....(3.2)}$$

where $\overline{\varphi}(i) = T_{proc} + T_{oxc} + \overline{T}_{queue}$;

Note that averaging the queuing delay $\overline{T}_{queue}$, will not have any influence on the value of the offset time, since $\overline{T}_{queue} = \frac{1}{\nu}\sum_{i=1}^{\nu} T_{queue}$.

Define two new parameters δ_o, the necessary time to reserve resources, which is the exact number of hops along the path between the source and the destination multiplied by the average delay at the intermediate nodes, and δ_+, the extra time added to the offset time, representing the greater-or-equal sign in Equation 3.2.

$\delta_o = (\nu * \overline{\varphi})$

$\delta_+ = \sigma * \delta_{del}$,

where σ is the *extension factor,* $(\sigma \geq 0)$

δ_{del} is the delay level

The delay level δ_{del}, measured in units of time, is a generic form that can be defined to represent the extra unit of delay to be added to the *offset time*. For example, δ_{del} can represent a QoS delay level defined in the network, or an average node delay $\overline{\varphi}$. The extension factor, σ, is an integer multiplied to the delay level to represent the total extra duration added to the *offset time*. Having different values for σ can define different QoS levels in the network for example. Introducing δ_+ helps to refine and adapt the offset time factor to the network needs from the control point of view and to any technological changes.

Equation 3.3 shows the flow of the *offset time* factor definitions that will be referred to in the rest of this chapter.

$$\delta \geq \varepsilon + \sum_{i=1}^{v} \varphi(i) \quad \leftrightarrow \quad \delta \geq (v * \overline{\varphi}) + \varepsilon \quad \leftrightarrow \quad \delta = \delta_o + \delta_+ + \varepsilon \;................(3.3)$$

3.5 Restoration of Link Failures

The routing mechanism implemented in JIT architectures is similar to IP networks, where each node has a forwarding table that contains the next hop to the desired destination. The routing protocol computes two paths between every source-destination (s-d) pair; a primary route which is calculated based on the shortest number of hops, and an alternative route that is link disjoint from the primary route. Then the routing tables node "*i*" is updated with the identification of the next hop on both the primary and alternative route to every destination in the network, plus the total number of hops along the calculated routes (both primary and alternative routes) as shown below:

- primary route = (d, n_i, v_i);
- alternative route = (d, n_i', v_i')

  where d = destination,

  n_i = next hop,

  v_i = number of hops on the route (primary or alternative)

This study considers only link failures. It has been shown that multiple links failures are very rare to happen [TON94] [VER02]. The restoration mechanism relies on implements deflection routing, i.e. the burst stream will be deflected away from the primary route to the pre-calculated alternative route.

Upon a link failure, the nodes adjacent to the failure detect the failure and localize it, and notify the rest of the nodes in the network about the failure. The fault notification process can be either centralized or distributed depending on the routing mechanism used. In centralized approaches, a centralized routing is notified while in distributed approach, nodes exchange information in a way similar to link-state protocols. During this phase, the nodes adjacent to the failure will deflect the traffic from the primary to the alternative route. Upon a global routing update, a new set of routes is determined reflecting the current state of the network.

The objective of any restoration procedure used is to minimize the burst blocking probability between the instant of failure till we restore the failed link. Achieving such goal can be done by either a fast detection/notification procedure or by the efficient use of spare network resources. Fast detection/notification procedure is beyond the scope of this thesis and will not be discussed.

Before pursuing the discussion, we differentiate between different levels of burst blocking probabilities (BBP) as shown in Figure 3.3 [XIN04]. The time duration between t_o and t_l represents the network under no failures. This period maintain a burst blocking (BB) level of b_o. At time t_l, a link fails and the blocking increases to a level equal b_l until the adjacent nodes detect the failure and start deflecting the data burst on the alternative routes. This mechanism reduces the BB to level b_2. After the global routing table updates, the BB decreases to a lower level but does not achieve the BB levels of the network under no failures. The reason behind that is the lower connectivity present in the network. Note that highly connected networks may achieve b_3 levels close to b_o.

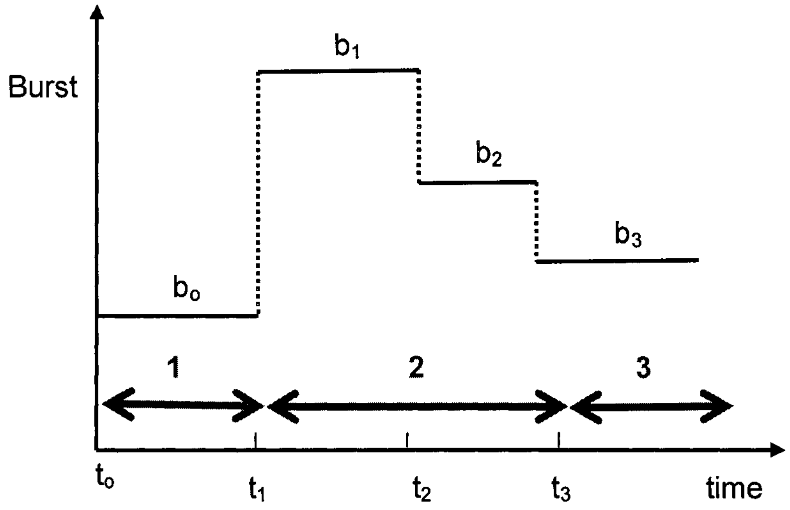


Figure 3.3: Burst losses versus time [XIN 04]

In the next section we will consider a failure scenario and discuss the influence of the offset time on the BB in the interval [t_2, t_3].

3.6 Failure Scenario

Upon of a link failure, bursts are deflected away from the failed link onto the alternative route. In most cases, the alternative route is longer than the primary route in terms of the number of hops. As the difference in the number of hops increases between the alternative and original routes, the pre-calculated *offset time* will be insufficient to ensure the proper path setup before the burst arrival. This will result in burst dropping toward the end of the route, and will have an effect on the overall burst blocking probability during the fault-recovery period.

Consider Figure 3.4 below. The primary route between the source–destination nodes (S-D) passes along nodes 1 and 2. Upon a link failure between nodes 1 and 2, and as soon as node 1 detects the failure, the incoming data burst destined to D will be re-routed on the alternative route between node 1 and destination D. It is obvious that the primary route between (1-D) is not active due to the link failure present. The new route along nodes 3-4 (in red) surpasses the primary route by one extra hop.

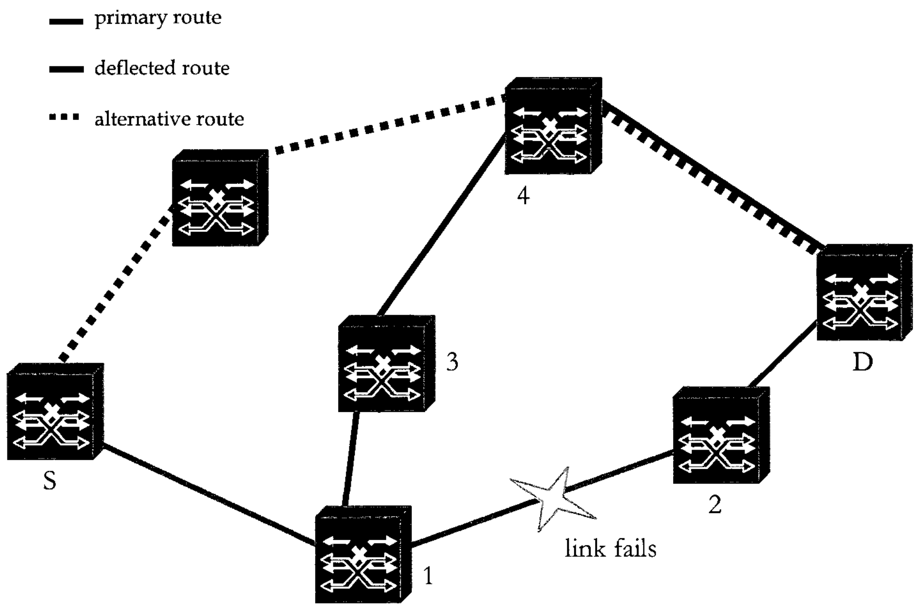


Figure 3.4: OBS core network

The pre-calculated *offset time* at the source S will not be sufficient to reserve the resources along the new route between nodes S-D. Therefore, the data bursts will arrive at node 4 together with the control packet, with no enough time to configure the switching entity. In the absence of buffering capabilities at the intermediate nodes, the burst will be dropped or incorrectly delivered. As a result, and in spite of the re-routing done, all the bursts traversing this path will be dropped.

To overcome this phenomenon, the *offset time* must be adjusted to cease burst dropping due to un-configured resources. The result of a link failure will cause two different types of blocking probabilities:

- $\boldsymbol{B_L}$: burst blocking due to unavailable resources, mainly free wavelengths
- $\boldsymbol{B_T}$: burst blocking due to insufficient time available to reserve the resources along the alternative route before the arrival of the data burst.

B_L plus B_T represents the burst blocking value of b_2 in Figure 3.5.1. The objective is to minimize the values of B_L and B_T and thus utilize the available resources more efficiently.

Two possible solutions helps to reduce the value of B_L are the increase the number of wavelengths and deployment of wavelength converters in the network. This will serve in increasing the chances for incoming requests to find free resources.

In order to reduce B_T, the value of the *offset time* has to be extended. Assign the delay level δ_{del} , to the value of the average node delay $\overline{\varphi}$. With appropriate choice of the extension factor, σ, this will insure that the offset time will hold σ -

extra hops added to the primary route. The extended offset time is shown in Equation 3.4:

$$\delta = \delta_o + \delta_+ + \varepsilon = (\nu * \overline{\varphi}) + (\sigma * \overline{\varphi}) + \varepsilon = \overline{\varphi} * (\sigma + \nu) + \varepsilon \ldots\ldots\ldots\ldots\ldots (3.4)$$

At the node adjacent to the failure, and before re-routing the burst, the algorithm presented in Figure 3.5 must be performed.

```
if (next hope is y & destination is d)
    if ($\delta \geq \Delta * \overline{\varphi}$ )  // $\Delta = v_x' - v_x$
        forward burst to alternative route
    else
        drop the burst
```

Figure 3.5: Simple check algorithm

This algorithm will ensure that no bursts will be re-routed unless their corresponding control packet has enough time to reserve the resources toward the destination. This guarantees that no unnecessary reservations of resources will occur, with a prior knowledge that any re-routing will not lead to the successful burst deliver to the desired destination and thus will increasing the level of congestion in the network resulting in higher burst blocking probability. It is important to note that the deflected bursts might be dropped later on due to unavailable free wavelengths.

Deploying this mechanism with the restoration scheme will reduce the value of B_T down to zero.

The challenge of this algorithm is the appropriate choice of the extension factor σ. This will be studied during the simulation phase, where we will define apposite values of σ. Key issues that heavily affect the value of σ are the network topology and network diameter, largest number of nodes along the shortest paths in the network.

3.7 Remarks

Figure 3.6 presents four scenarios of the control packet and its corresponding data burst positions. In the first two scenarios, where the control packet is on or has passed the failed link, the corresponding data bursts will be dropped definitely since the node adjacent to the failure is not able to regenerate another control packet.

It is very import to note in the second scheme that the resources will be reserved on the other side of the failed link waiting for the data burst to come. In explicit release schemes, this will lead to resource reservations for an undetermined amount of time. A viable solution is to have a timer at the intermediate nodes with a value greater than the *offset time*. If the timer expires before the burst arrival, the reserved resources will be released.

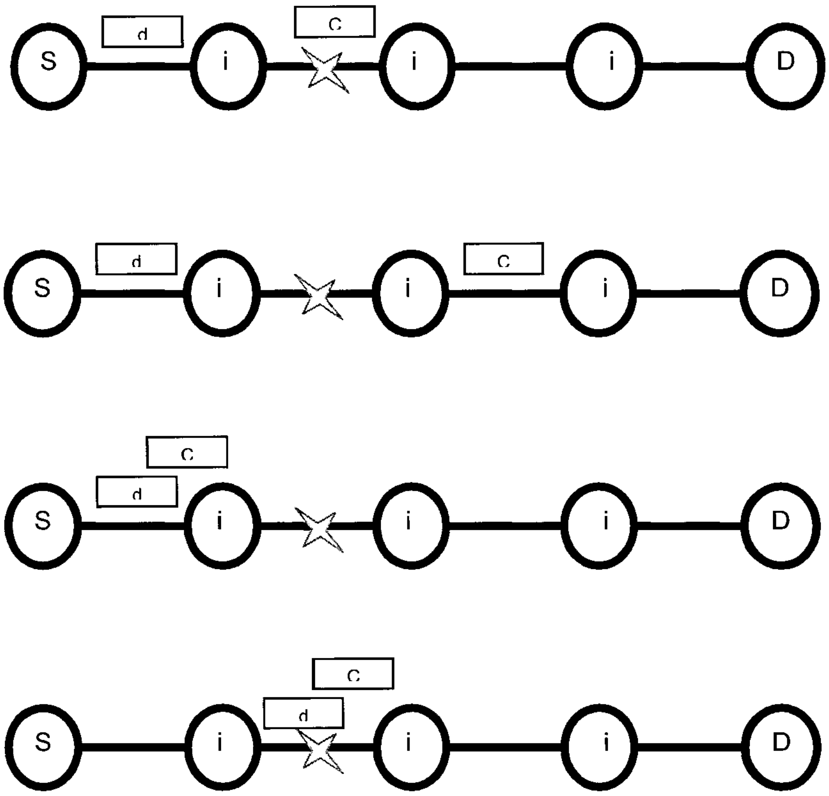


Figure 3.6: Different scenarios representing the position of the control packet and the data burst upon a link failure

In the third scenario, the link fails before the arrival of the control packet. In this case, we will be able to perform the proposed enhanced restoration mechanism.

A fourth scenario might occur when both the control packet and the data burst are on the link when it fails. This scenario is most likely to happen at the rare end of the path, where the offset time value has elapsed and very close to zero. This will lead to the loss of both packets.

3.8 The QoS-based Restoration Scheme

As discussed earlier, it is a major issue for OBS networks to support a wide range of services for different kinds of applications. In this section we will introduce a novel relative QoS restoration scheme for OBS networks.

Adding different extension values to the *offset time* enables the creation of different classes of service correspondent to the data bursts. This service differentiation guarantees a certain level of data loss and delays in OBS networks under failure.

Through proper knowledge of the level of BBP each extension-factor value increment introduces, one can define different classes of service relevant to the value of σ. It is also known that each increment of σ introduces a delay equal to the total hold-ups at a single node plus the propagation delay (t_{prop}) on the extra link traversed. In other words, each value of σ_i corresponds to a different value of burst dropping B_i and to an extra delay of $\sigma_i * (t^i_{prop} + \bar{\varphi})$. This relation is shown in Equation 3.8.1.

$$\sigma_i \leftrightarrow B_i \leftrightarrow f(\sigma_i) = \sigma_i * (t^i_{prop} + \bar{\varphi}) \quad \ldots\ldots (3.8.1)$$

With the compensation between the level of burst loss and the delay encountered during a restoration process or either parameter alone, one can define a QoS dimension fitting the desired needs.

Consider Table 3.1 below. It is desired to define, for example, five different QoS levels depending on the burst dropping in the network. Each BBP value is attained with a specific increment of σ_i and results in an extra delay of $\sigma_i *(t^i_{prop} + \bar{\varphi})$. Note that the values of σ_i are gotten from either simulation studies or measurements performed by network administrators.

		BBP (%)				
i	**Delay (sec)**	0-20	20-40	40-60	60-80	80-100
0	0	$\sigma_i = 0$				
1	$1 *(t_{prop} + \bar{\varphi})$		$\sigma_i = 1$			
2	$2 *(t_{prop} + \bar{\varphi})$			$\sigma_i = 2$		
3	$3 *(t_{prop} + \bar{\varphi})$			$\sigma_i = 3$		
.	.					
.	.					
n	$n *(t_{prop} + \bar{\varphi})$					$\sigma_i = n$

Table 3.1: QoS Level Computation Table

Extracting the values from Table 3.1, one can define the five QoS classes desired. Note that $\sigma_i = 2$ and $\sigma_i = 3$ belong to the same BBP interval, but they result in different delays. If the user want to define QoS classes based on burst dropping only, it is preferred to used the σ_i with the smaller value.

Chapter 4

OBS NETWORK MODELING AND SIMULATION

4.1 Introduction

Optimum **Net**work performance simulation tool, OPNET, provides a comprehensive development environment that supports the modeling of communication networks and distributed systems. The OPNET environment incorporates tools for all study phases, including model design, simulation, data collection, and data analysis. With the offered flexibility and scalability, OPNET Modeler was chosen to perform the simulation studies on the work presented in this thesis. This chapter provides an overview of the model built to simulate the core of an optical burst-switched network, describing the modeling assumptions and the detailed behavior of the modeled system.

4.2 OPNET Modeler Overview

OPNET Modeler is based on a series of hierarchal editors that directly parallel the structure of a real network. As shown in Figure 4.1, this hierarchy starts at the "Network Level" where the topology of a communication network is represented graphically. Networks consist of node models interconnected via link objects. Each node model consists of a group of functional elements, called modules, interconnected among each other. Modules are capable of generating, sending, and receiving packets from other modules to perform a certain function within a node.

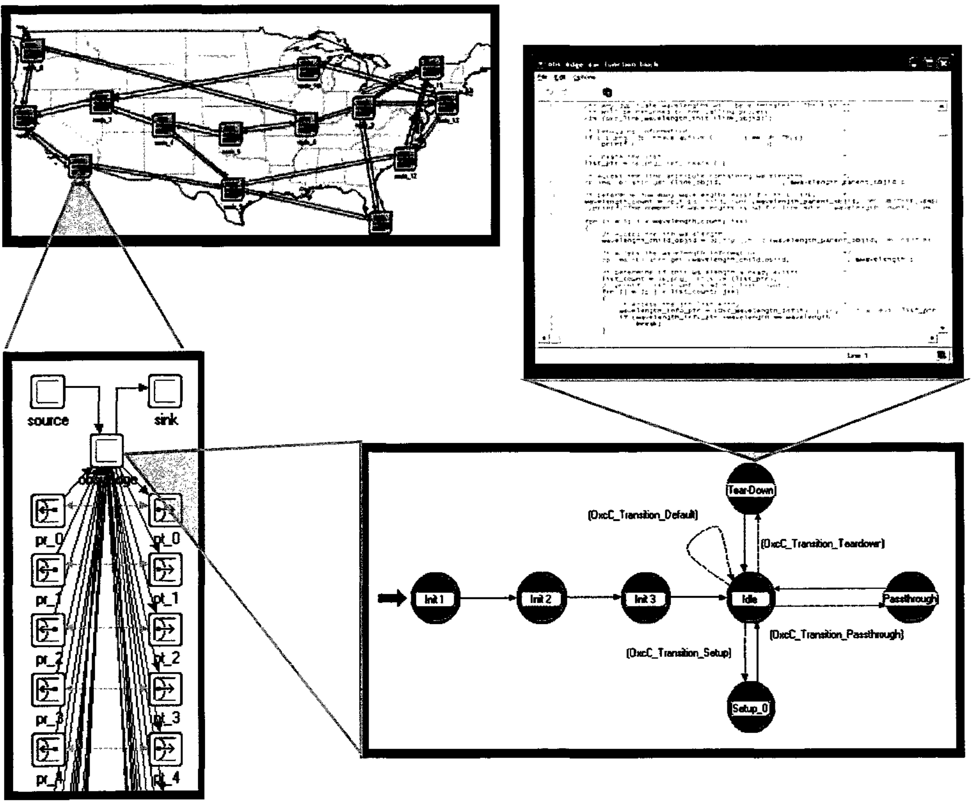


Figure 4. 1: OPNET Modeler Hierarchy

Modules are assigned process models at the “Process Level”. Process models reflect the behavior of individual objects using the powerful Finite State Machine (FSM) approach. Each state and transition defines the progression of a process in response to events. At the lowest level, each state consists of a C/C++ source code reflecting the characteristics of the modeled system’s behavior.

4.3 Optical Burst Switching Network Model

The core concept of modeling is to build a model that is equivalent to a real system. However equivalency is a subjective term which must be defined more precisely. Clearly, equivalency means that the model behaves in some sense like the real system. Nevertheless, for practical reasons, models are usually limited to representing only certain aspects of the system of interest with the following objectives kept in mind:

- The model should answer all the questions of interest
- The model should have the desired level of accuracy
- The model should support validation, i.e. building confidence in the results produced.
- The model should accommodate the necessary range of operating conditions (changing traffic, new patterns, etc.)

To model an optical burst switched network core, we created different objects that modeled all network elements starting from the assorted packet-formats generated to optical links and optical burst switching nodes. In the next subsections, a list of the model assumptions and a description of each object model behavior is presented.

4.3.1 Model Assumptions

Just-In-Time (JIT) architecture is deployed in a bufferless network. Resources in the network are reserved and released explicitly by exchanging setup (control) and teardown packets between the communicating entities, as described in Chapter 3. Therefore, no schedulers are required at the intermediate nodes.

Due to the fact that the core of an optical burst switched network is simulated, an assumption that no burst aggregation is necessary at the edge of the network is made. Each node is capable of generating call requests according to user-defined distribution during a predefined time interval. These requests are presented by sending control, data burst, and a teardown messages consequently. The length of the bursts follows a user-defined distribution measured in seconds. Note that the lower and upper bounds of the burst size should reflect the values used in regular burst assembly algorithms. Each node experiences delays while processing control packets and configuring the optical cross connect (OXC). These delays are set by the user depending on the technology specification. The destination choice follows a uniform distribution between all the registered nodes in the network.

Signaling assumptions:

- Signaling is done out of band
- A wavelength is dedicated for exchanging control packets back and forward between the nodes.
- Infinite queues are used to queue control packets at intermediate nodes.
- Control packets are neither acknowledged neither re-transmitted.
- The control plane is 100% reliable.

Routing assumptions:

- A routing entity (RE) is responsible for computing routing tables for all nodes in the network.
- A primary and its link-disjoint alternative route are calculated between every source-destination (s-d) pair using Dijkstra's algorithm.
- Each routing table consists of the next hop address on the primary and alternative routes and the total number of hops toward the destination respectively.
- Upon a failure, a notification message is sent to the RE, where the routing tables are updated to reflect the network failure after a predefined restoration time.
- Random wavelength assignment algorithm is used under the wavelength continuality constraint, i.e. no wavelength conversion is deployed in the network.

4.3.2 Packet Objects

Three types of packets that are transmitted between communicating entities is defined: the control, teardown, and data packets. Each packet type is an object that contains formatted information defined as a set of fields.

The control packet (CP) originates from the requestor of a connection to setup a path along the network to the desired destination. The control packet, shown in Figure 4.2, is a formatted packet that consists of five fields. The fields represent the connection identification number, input wavelength identifier, next hop address, offset time, and the destination address respectively and their corresponding size. Other fields such as the protocol version, message type, cyclic redundancy check, and layering information [WEI00] are not included in the CP fields because of their irrelevancy on the performance of the signaling protocol under the simulation assumptions took.

connection ID	wavelength color	next hop	OFFSET time	destination
4 octets	2 octets	8 octets	4 octets	8 octets

Figure 4.2: OBS Control Packet Structure

A unique connection identification number is given for each connection request that finds an available wavelength on the source's outgoing link. The corresponding wavelength color is tagged by the source node in the CP and reflects the desired wavelength to be reserved along the path to the destination. In case wavelength conversion is deployed in the network, the wavelength identifier field is updated by the intermediate nodes on a hop-by-hop basis to reflect the changes in the wavelength color.

The *offset time* field reflects the lag in time between the control packet and the data burst. This field is decremented at each intermediate node by the amount of delay encountered locally at the node. Usually, such delays are the same at all nodes in the network. In the explicit setup schemes, the only role of the "*offset time*" field is to reveal the difference in the number of hops between the primary and alternative routes. In estimated setup schemes, the *offset time* is also used by the scheduler in the controlling entity to estimate the arrival time of the data burst and perform the necessary resource reservations before the expected arrival.

The teardown message (TM) structure is shown in Figure 4.3. The fields represent the connection identification number, next hop address, and the destination address. The connection identification number is the same as the number tagged on the correspondent control packet. At the intermediate nodes, this number is matched with control packet ID number and the appropriate action is taken (either forward the teardown to the next hop or drop it).

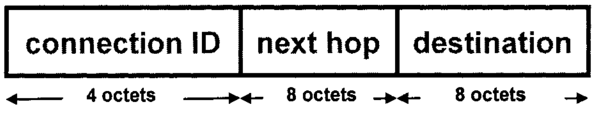


Figure 4.3: Teardown Message Structure

The third packet represents the data burst. The size of the burst can vary from one connection to other. In the case of explicit setup schemes, the user can ignore communicating these packets, and depend solely on connection duration defined in the source node's attributes.

4.3.3 Link Object

An optical link was modeled using the Link Editor. As shown in Figure 4.4, a link consists of different channels, each representing a unique wavelength color. Each channel incorporates a simple first-in-first-out queue on its transmitting side to ensure that only one packet is being transmitted at a given time.

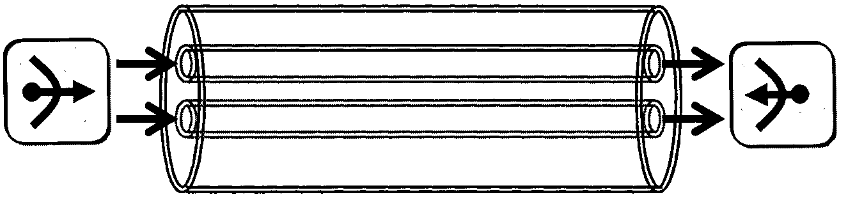


Figure 4.4: Correspondence between an optical link and its channels

Different link attributes are defined to model the physical and logical characteristics of an optical link. Beside other graphical attributes, the following attributes are described:

- **Cost:** This attribute represents the cost of the link, which is primarily intended as a routing parameter. It is set to "1" to allow the routing paradigm to calculate the routes based on the smallest hop count.
- **Bit error rate:** This attribute is used to specify the probability of bit errors in packets which are transmitted over a link. It is set to "0" to ensure zero physical channel losses.
- **Data rate:** This attribute specifies the speed of data transmission over the link. Different data rates were defined (OC1, OC3, OC12, OC48, T1, T3...) to provide flexibility in changing parameters during the simulation run phase.
- **Delay**: This attribute specifies the propagation delay which will be incurred by packets sent over the link.
- **Packet formats:** This attribute specifies the type of packets that are supported over the link. We added the three packet types discussed in the previous section.
- **Wavelength:** This attribute defines the number of wavelengths to be used in the network. Each wavelength color is represented by a distinctive number.

4.3.4 Node Object

An optical burst switched core node, as shown in Figure 4.5, consists logically of two planes: a control plane and a switching plane. The control plane is responsible for generating connection requests, processing control and teardown messages, and configuring the switching entity accordingly. On the other hand, the switching plane performs space and wavelength switching. Keeping the same performance of the control and switching planes, we integrated the *control entity* and *oxc* modules into a single module; *obs_edge_oxc*, during implementation to reduce the complexity of the model.

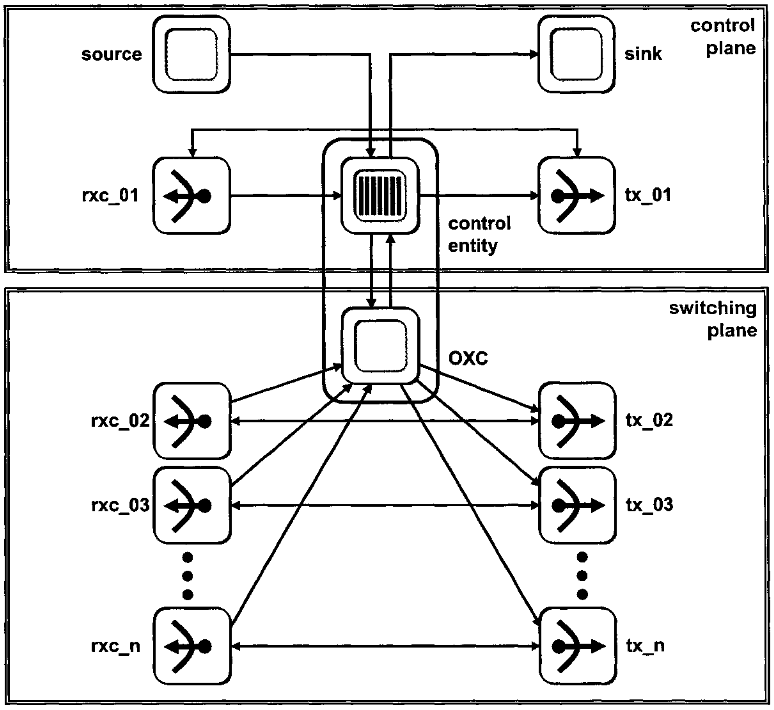


Figure 4.5: Optical Burst Switching Core Node Model

Each OBS node is associated with a set of attributes defined by the user to provide control on the node's operation. These attributes allows node modifications to be done externally and offers the node generality and flexibility when modeling different network situations. Below is the set of the defined attributes with their functionalities:

- Destination: this attribute defines the destination for the calls generated at the node. It can be either set randomly or explicitly.
- Start time: this attribute defines the start time for call generations. This attribute allows offline calculations to be done before the start of the network operation.
- Inter-arrival time: this attributes represents the time that separates two consecutive generated calls. This attribute can follow a wide set of defined distributions such as normal, exponential, binomial, etc.
- Duration: this attribute defines the call duration. This duration is equivalent to the length of the data transmitted measured in unit of time.
- End time: this attribute defines the time in which calls will no longer be generated.

The transceivers shown in Figure 4.5 serve as an interface between the packet streams inside the node and the optical links outside the node. The transceivers collect/distribute packets over the corresponding channels within the communication link. The *source*, *sink*, and *oxc* modules are associated with processes describing their behavior.

(*i*) Source Process

The source process model has a single forced state as shown in Figure 4.6. This state is responsible for registering the node's connection parameters (start time, inter-arrival time, duration, etc.) defined in the node's attributes. These parameters are used later by the source's child processes to generate connection requests according to the parameters defined. This process is associated with two children, the *obs_source_child* and the *obs_source_gchild.*

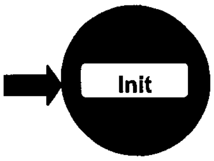


Figure 4.6: OBS Source Process Model

(ii) Source's Child Process

The first child process, shown in Figure 4.7, is invoked by the source process model. This child process creates a list that holds the generated child processes, confirm the start and end time are correctly specified, and schedule their generation. After this stage, it obtains the current simulation time and schedule an event for the start of the first connection request.

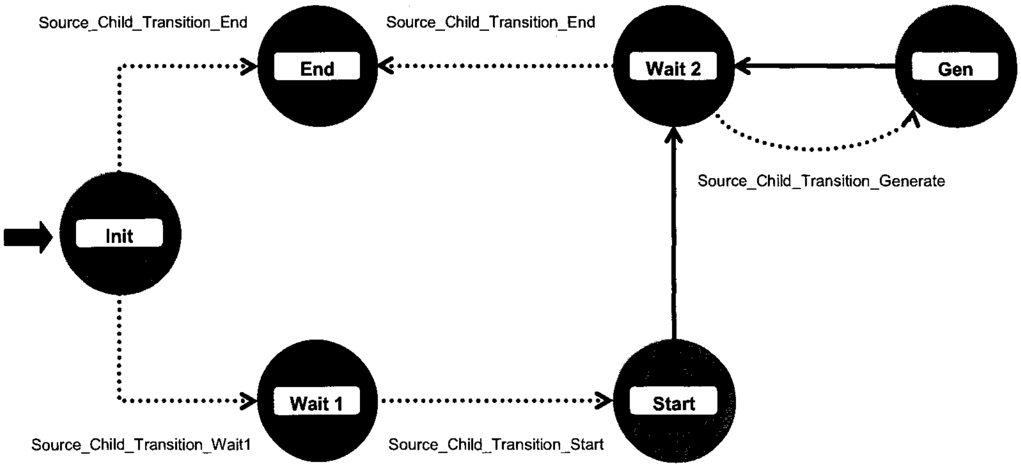


Figure 4.7: OBS Source Child Process

Upon scheduling, the *generate* state allocates memory to hold source generation information, obtains a unique connection id number, determines the connection duration, and sets the destination address according to a uniform distribution.

(iii) Source's GChild Process

The second child process is invoked by the source child process model. This process consists of four forced states and two unforced states, as shown in Figure 4.8. The *setup* state is responsible for creating setup packets and setting the appropriate packet fields. The *offset* state gets the number of hops along the route toward the destination from the *obs_edge_oxc* process and calculates the offset time and generates an interrupt to send the data burst. The *burst* state generates a data burst and schedule an interrupt to invoke the *teardown* state after a period equals to the length of the burst measured in seconds. The *teardown* state generates a teardown packet and fills the packet fields with the necessary connection information and transits to the *end* state where the *gchild* process is destroyed.

(*iv*) Sink Process

The *idle* unforced state is responsible for destroying the control, data, and teardown packets if the node is the sought destination.

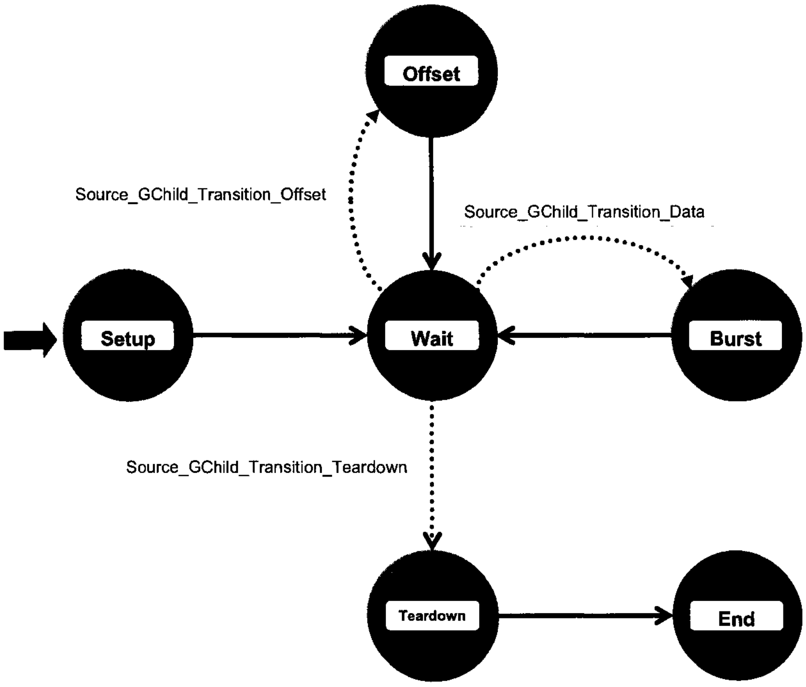


Figure 4.8: OBS Source GChild Process

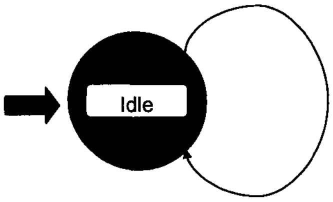


Figure 4.9: Sink Process Model

(*v*) OBS Edge OXC Process

The obs_edge_oxc is the most fundamental process. It performs the functions of the control and the switching plane conjunctionally. It consists of seven states as shown in Figure 4.10.

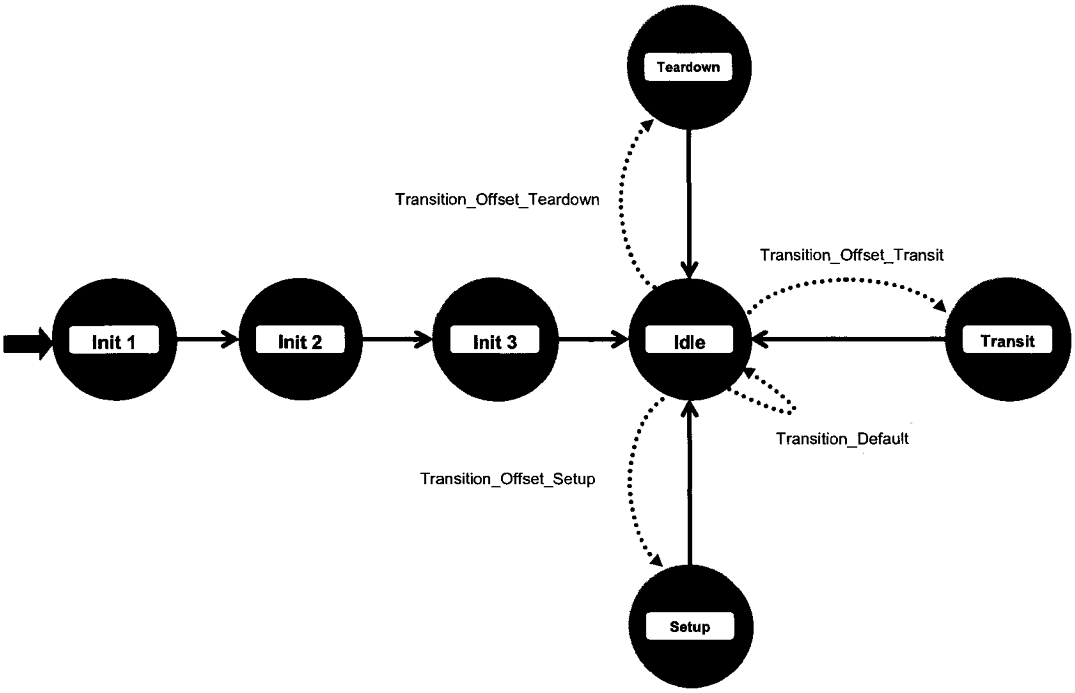


Figure 4.10: OBS Edge OSC Process Model

The *init 1*, *init 2*, and *init 3* states perform registration and initialization tasks. These tasks are performed offline, i.e. before the node starts generating or receiving packets. The *init 1* unforced state registers the node's information and initialize all state variables and statistics. It also registers the adjacent link information such as the link's condition (enabled, disabled), link's cost, etc. The *init 2* unforced state determines all the accessible wavelengths through the node and the entire destination node's information. *Init 3* state is responsible for process registry allocation to allow exchange of information among different processes, such as the exchange of number of hops between s-d pairs.

4.3.5 Operation Scenarios

A node can operate as either a source node, intermediate node, or a destination node depending on the packet it is processing. The three different scenarios are described below.

(*i*) Scenario 1: Source Node

The *setup* state receives the control packet generated at the source's gchild process and processes the CP fields. It first checks for the destination's address and pick up the next hop address from the node's routing table. Then it tries to reserve a wavelength on the corresponding outgoing link. If the reservation succeeds, the wavelength color is tagged in the control packet, and the corresponding CP is forwarded to the next hop, otherwise the connection request is blocked and local statistics are updated.

After the offset timer expires, a teardown packet is received at the *teardown* state. The source node checks for the next hop based on the destination address of the packet and forward it to the next node releasing the reserved wavelength on the outgoing link.

(ii) Scenario 2: Intermediate Node

Upon arrival of a packet, the intermediate node checks the packet type. If the packet is of type control; the controlling entity at the node checks for the next hop and its corresponding outgoing link and tries to reserve the wavelength tagged in the control packet. If the process succeeds, the CP packet is forwarded to the next hop and the wavelength table is updated. Otherwise, the CP is dropped and local statistics is updated. In the case of a teardown packet, the same operation is performed but wavelengths are released instead.

(iii) Scenario 3: Destination Node

The packet's destination address is checked. If the destination address matches the node's address, the arriving packet (control or teardown) is moved to upper layer where it will be destroyed. Local statistics are updated to reflect a successful connection request establishment.

Chapter 5

SIMULATION RESULTS AND DISCUSSION

In this chapter we present the results conducted from the simulations performed using OPNET Modeler. We discuss the results obtained from simulating two different network topologies, and validate the ideas presented in Chapter 3.

5.1 Simulation Assumptions

A Just-In-Time (JIT) OBS network is simulated using the model described in Chapter 4. The network core nodes are assumed to have no optical storage capabilities but can queue the communicated control and teardown messages in electronic buffers after the appropriate Optical-to-Electric conversion. It is assumed that all the requests are generated randomly with a Poisson distribution at the core OXCs and are uniformly distributed among all the nodes in the network. The Poisson distribution is used since it reflects a realistic environment for any network environment where it meets three important conditions [BCB00]:

- Requests come one at a time,
- The inter-arriving time between two consecutive events is exponentially distributed,
- Incoming requests are completely independent from previous requests.

The load in the network is calculated based on the Erlang rho (ρ) formula shown in Equation 5-1. The load is varied between 0.2 and 1.0 Erlang with increments of 0.2 Erlang.

$$\rho = \lambda\left(\frac{1}{\mu} + \overline{T}_{OFFSET}\right) \qquad \text{5-1}$$

where

- λ is he arrival rate

- μ is the service time

- $\overline{T}_{OFFSET}$ is the average offset time, calculated based on the average number of hops in the network

Each active optical cross connect handles both bypassing and locally generated or terminated bursts and is characterized by a set of attributes that reflects its behavior during the simulation runs. These attributes are:

- Start time: is set to 100 sec in all the simulation runs performed. This time interval allows all offline computations to be done before the start of the run.
- Inter-arrival time: is set to an exponential distribution with a variable mean. The value of the mean was set differently in each run to reveal the load desired in the network.
- Duration: is set to $\frac{1}{\mu} + \overline{T}_{OFFSET}$ in order to reflect the length of the data transmitted measured in unit of time. The value of $\frac{1}{\mu}$ is fixed to 40 msec.
- End time: is set to end of the simulation. Each simulation run time was set to ensure that the output reaches a steady state. This was achieved by generating 500,000 simulation events whose run-time varied depending on the inter-arrival rate of the call requests.

Signaling is assumed to be done out of band on a dedicated wavelength. Infinite queues are used to queue control packets at intermediate nodes that are neither acknowledged nor re-transmitted. As discussed in Chapter 4, a routing entity (RE) is responsible for computing routing tables for all nodes in the network. Each routing table consists of the next hop address on the primary and the link-disjoint alternative route in addition to the total number of hops toward the destination respectively. Random wavelength assignment algorithm is used under the wavelength continuality constraint, i.e. no wavelength conversion is deployed in the network.

The optical links used exhibits a zero bit error rate to eliminate any chances of physical channel losses. Each link is assigned a cost of “1” to allow the routing paradigm to calculate the routes based on the smallest hop count. The transmission rate on the link is 10 Gbps over 32 wavelengths.

5.2 Burst Blocking Definitions

A set of burst blocking probabilities (BBP) that will be used throughout this chapter is defined below.

- B_N: burst blocking probability under no failures
- B_F: burst blocking probability between the instant a link fails and the time of its detection. This value depends on the detection; notification speed of the restoration mechanism used.
- B_R: burst blocking probability upon the detection of the failure until the global routing update is performed. B_R represents the restoration phase losses.
- B_L: burst blocking during the restoration phase due to unavailable resources, mainly free wavelengths
- B_T: burst blocking during the restoration phase due to insufficient time available to reserve the resources along the alternative route before the arrival of the data bursts.
- B_U: burst blocking probability after global routing update is done.

Under no failures $B_N = B_L$, since the only cause of burst losses is the unavailability of free wavelengths to transport the traffic between various s-d pairs. Under link failures, the value of B_R is the sum of the B_T plus B_L ($B_R = B_T + B_L$).

5.3 Simulation Results for ARPANET

This section will investigate the performance of the enhanced restoration scheme discussed in Chapter 3 on the Advanced Research Project Agency Network (ARPANET).

The network topology of ARPANET is shown in Figure 5.1. ARPANET consists of 14 nodes with an average nodal degree equals to 3, connected by 21 bi-directional links. The average hop distance in the network is 2.14286 using Dijkstra's shortest path algorithm.

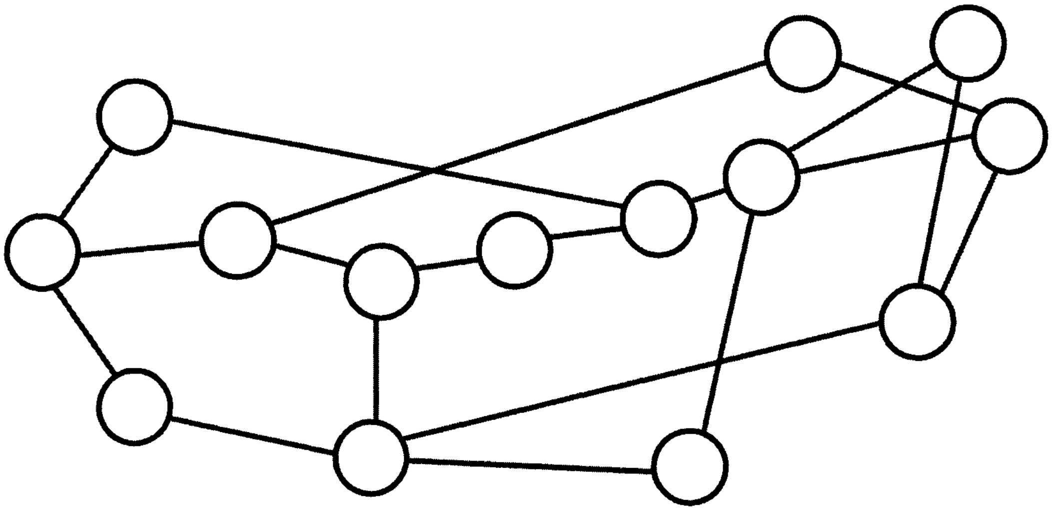


Figure 5.1: Advanced Research Project Agency Network (ARPANET) Topology

5.3.1 Offset Time Extension Effect Under No Failures

In JIT OBS network architecture with estimated setup schemes, the effect of extending the offset time will not have any affect on the network's burst blocking probability. Since the reservation is done just before the arrival of the data burst, network resources will not be reserved for a longer duration and are not affected by any increase in the *offset* time factor. On the other hand, the end-to-end latency is increased by the value of the extension time.

On the other hand, in JIT OBS network architecture with explicit setup schemes, the network resources are reserved directly after the arrival of the control packet and are released after the transmission of the last bit of the data burst. As discussed earlier, the *offset* time represents the time gap between the control packet and the data bust. As a result, any change in the value of the *offset* time will have a direct impact on the reservation time duration of the network resources and thus on the network's overall burst losses.

Figure 5.2 compares the effect of the offset time extension on the network performance in terms of burst losses. The OXC configuration time t_{oxc}, is set to 1 millisecond and the processing time t_{proc}, is set to 50 microseconds. The value of B_N increases as the extension factor σ increases. B_N increases by 6.8% at network load of 0.2 Erlang and maintains almost the same behaviour at different network loads (6.5% at 1.0 Erlang). This is shown more clearly in Figure 5.3, where the curve slopes is close to zero.

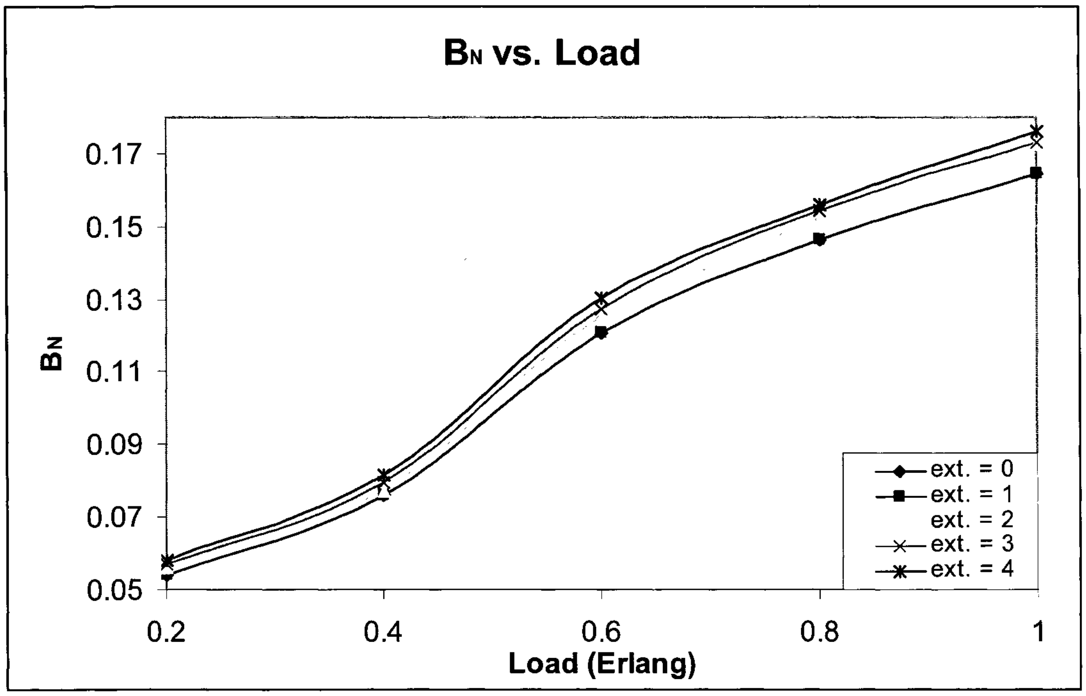


Figure 5.2: Overall BBP vs. network load at different offset time extensions

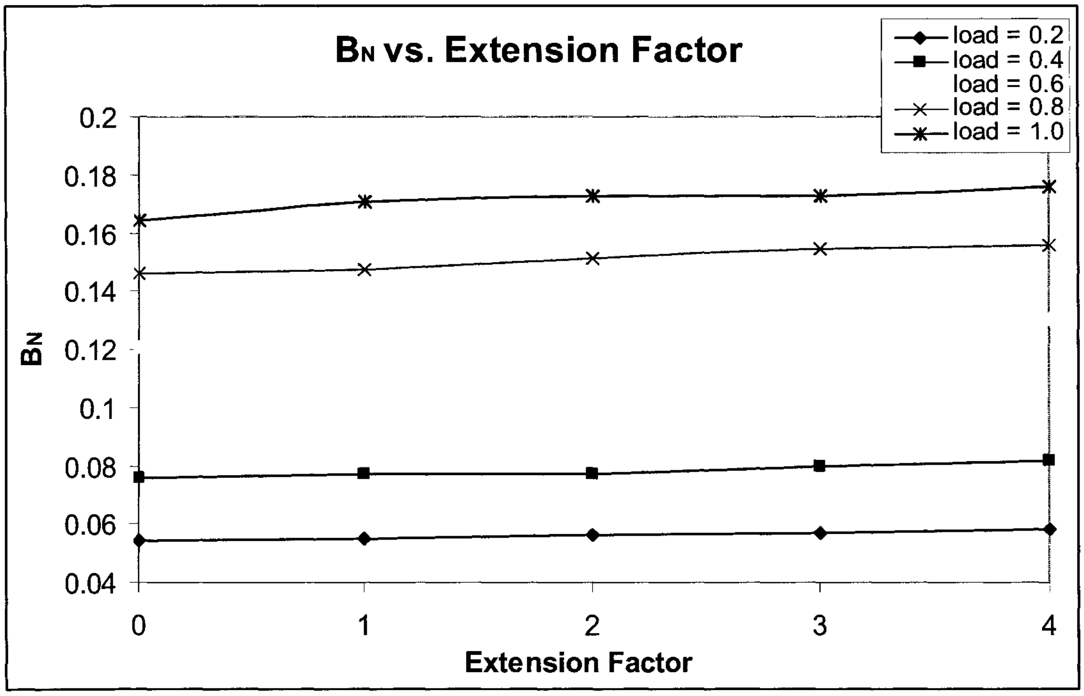


Figure 5.3: Overall BBP vs. time extension factor at different network load

This slight change in the value of B_N is due to the fact that the delay encountered at the intermediate nodes is relatively small compared with the burst duration (1.05 ms to 40 ms). Figure 5.4 presents burst blocking probability under no failures B_N versus different increments of the extension factor σ at different OXC configuration values. The network load is set to 0.8 Erlang. As the value of t_{oxc} increases, the change in the burst loss increases more dramatically. For example, at $t_{oxc} = 10$ ms and $\sigma = 4$, the value of B_N increases form 0.1559 to 0.2097 with a relative increase of 25.6%. The extra time added to the offset time is equivalent to the burst duration time (~ 40 ms) and thus the effect B_N is more significant, while the loss increase is less (only 10.2%) at $\sigma = 2$.

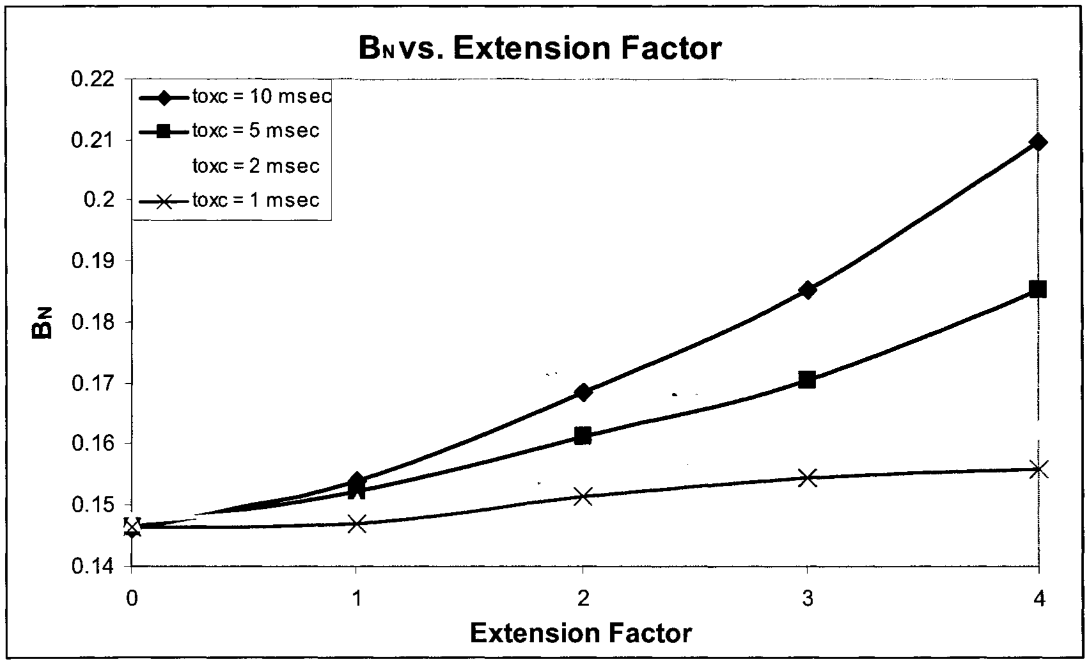


Figure 5.4: Overall BBP vs. time extension factor at different OXC configuration time

5.3.2 Offset Time Extension Effect Under Failures

This section studies the effect of the offset time extension on the network performance under link failures. All the results shown assume 40 ms-long bursts, OXC configuration time t_{oxc} = 1 ms, and processing time t_{proc} = 50 μs. It is assumed that estimated setup is used unless otherwise specified.

Recall the three phases defined in Chapter 3. Phase 1 represents the network under no failures. Phase 2 represents the network under a link failure. This phase lasts from the moment a link fails until the failure is recovered. The "recovery" term used refers to the global routing table update and the awareness of all the nodes in the network about the failed link. This is represented by phase 3. The fault detection time is assumed to be 15 ms with a 200 ms delay to update the nodes' routing tables with the new computed routes.

Figure 5.5 shows the average burst blocking due to insufficient time to reserve network resources, B_T, versus the extension factor σ. All the links in the ARPANET were failed one at a time. The correspondent B_T values and their σ increments were recorded respectively. For example, at $\sigma = 1$, 76% of the rerouted bursts will be dropped toward the end of the route due to un-configured resources. It was found that at $\sigma = 4$, there will be no burst losses due to insufficient time ahead of the data burst to reserve the resources along the deflected alternative routes. Note that some links exhibited smaller σ values to reach $B_T = 0$.

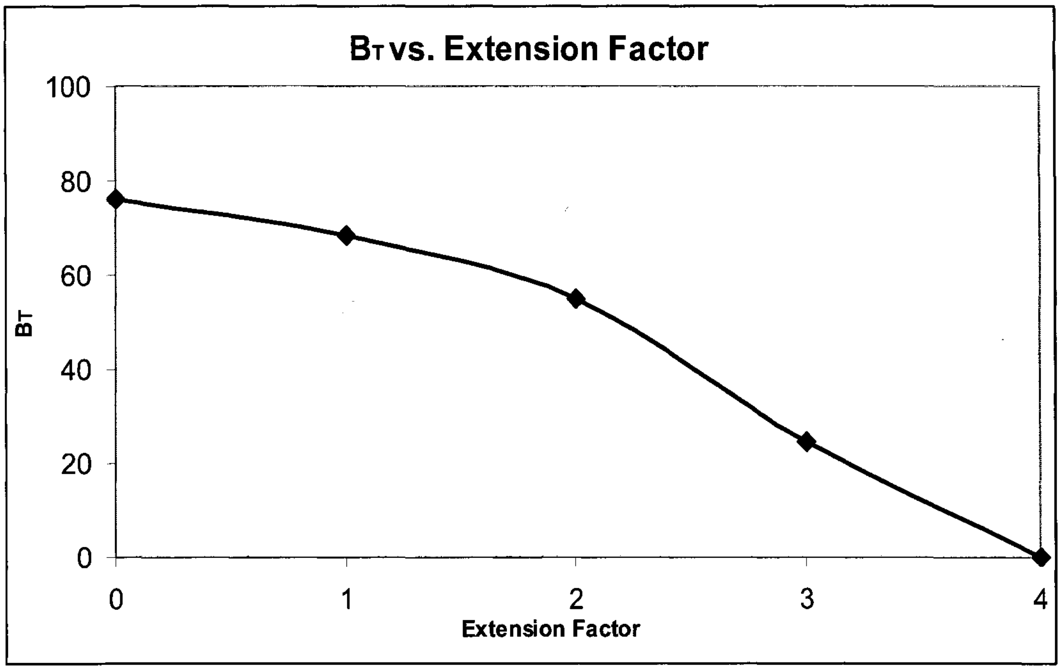


Figure 5.5: Burst Blocking due to Insufficient Time vs. Extension Factor

With the maximum value of "effective" σ in the ARPANET being known, Figure 5.6 compares the global burst blocking levels represented by B_R at the three phases with different increments of the extension factor σ. The results are computed at a network load of 0.8 Erlang. The global burst loss probability is 0.213 when the offset time factor is not considered. The B_R value decreases to [0.2082, 0.2077, 0.1838, 0.1701] with σ increments of [1, 2, 3, 4] respectively. The result is an improvement in the value of B_R by [2%, 5.5 %, 13.5%, 20%]. This shows the importance of the offset time factor in the restoration mechanism deployed and its role in reducing the overall network burst blocking probability. The value of B_R at phase 3 is slightly bigger than that at phase 1 due to the fact that the network connectivity has decreased by one link and thus the computed new routes are longer, resulting in a higher probability of used wavelengths and thus higher BBP.

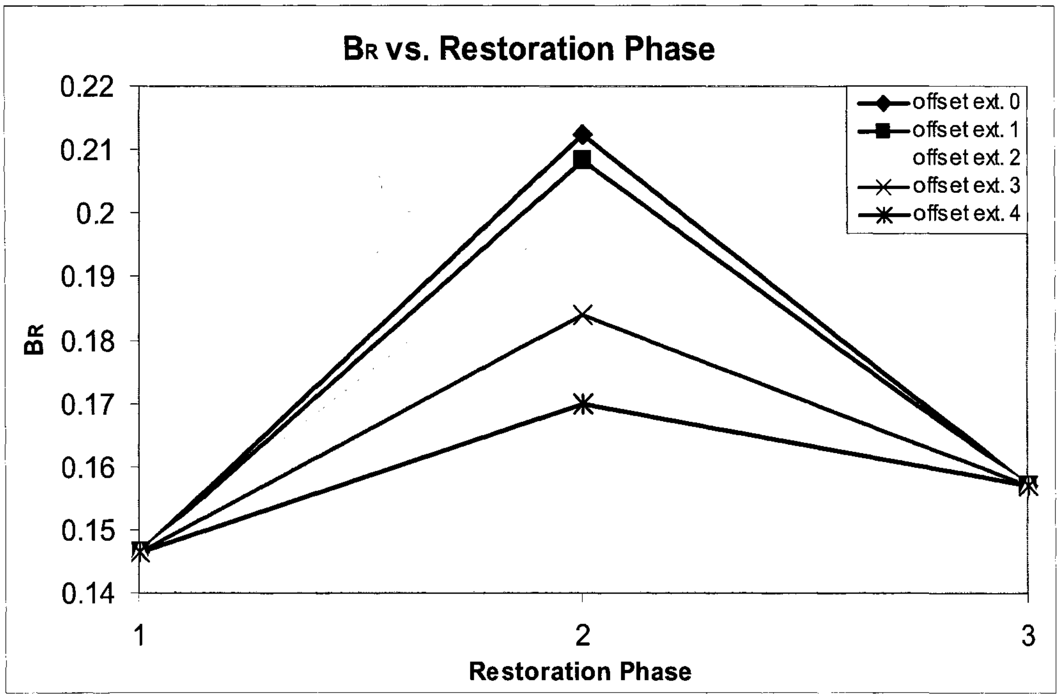


Figure 5.6: Global BBP vs. Restoration phase at Different Extension Values

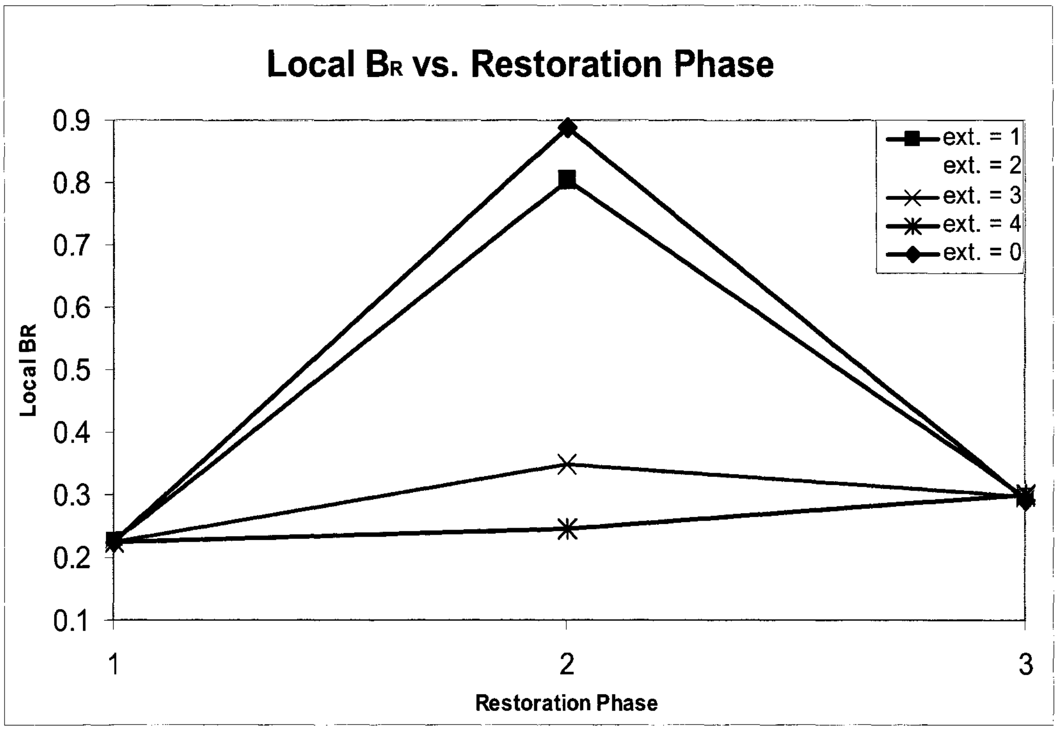


Figure 5.7: Local BBP vs. Restoration phase at Different Extension Values

Figure 5.7 shows the local B_R values during the 3 phases when the link between nods 8-9 fails. As the extension factor σ increases, the *offset* time will be sufficient to re-route the bursts on the alternative routes and thus less bursts will be dropped locally. In spite of this fact, the value of B_R is still higher at phase 2 than that at phase 1, even when $\sigma = 4$, because all the burst will be dropped until nodes 8 and 9 detects the failure.

Note that the different values of B_R at phase 2 reflect the B_T values in Figure 5.5. For example, the value of B_T at $\sigma = 2, 3$ drops from 55% to 26% by almost 30%. This shows clearly in Figure 5.7 with the decrease of B_R from 0.616 to 0.347 by almost 30%. This is due to the fact that most of the burst dropping is due to the algorithm implemented. In addition, the nodal degree of the nodes adjacent to the failure is higher that 2 and thus there will always exit a secondary link(s) to re-route the bursts. On the other hand, if the failure happens between nodes 4 and 6, all the incoming bursts at node 6 will be dropped since the nodal degree is 2.

Figure 5.8 compares the effect of the offset time extension on the network performance in terms of burst losses during link failures. The value of B_R decreases as the extension factor σ increases. Note that as the network load increases, the σ effect on the B_R value becomes more noteworthy. Figure 5.9 shows more clearly the improvement in the burst loss level as we extend the *offset* time. The range of enhancement varies from 17% to 20% with loads ranging from 0.2 Erlang to 1.0 Erlang at $\sigma = 4$.

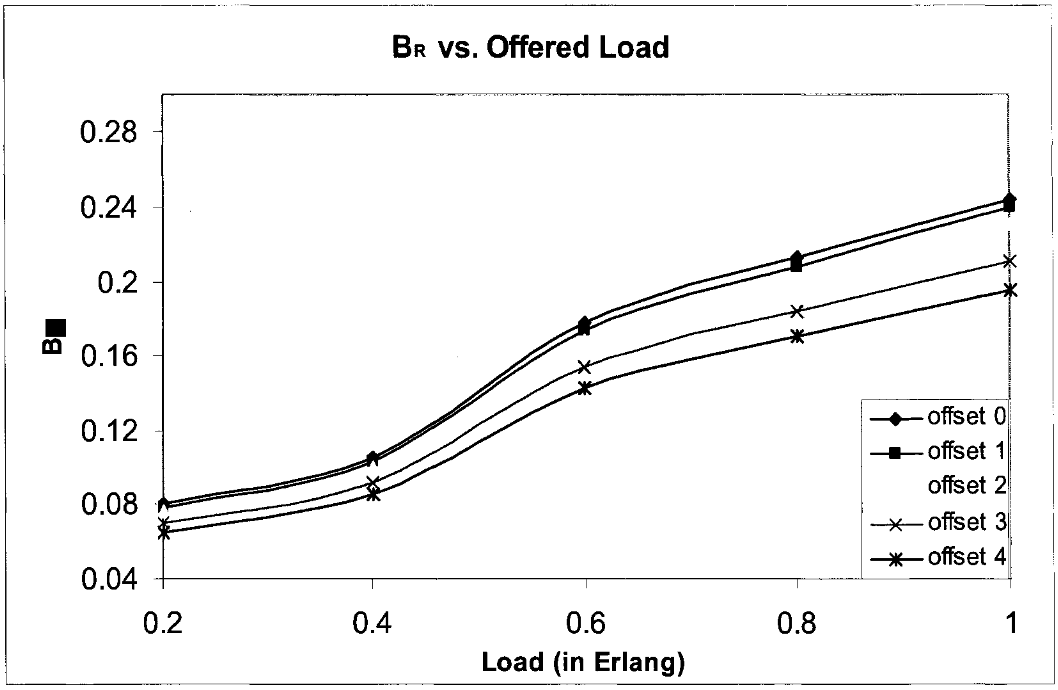


Figure 5.8: Overall BBP under failure vs. network load at different offset time extensions

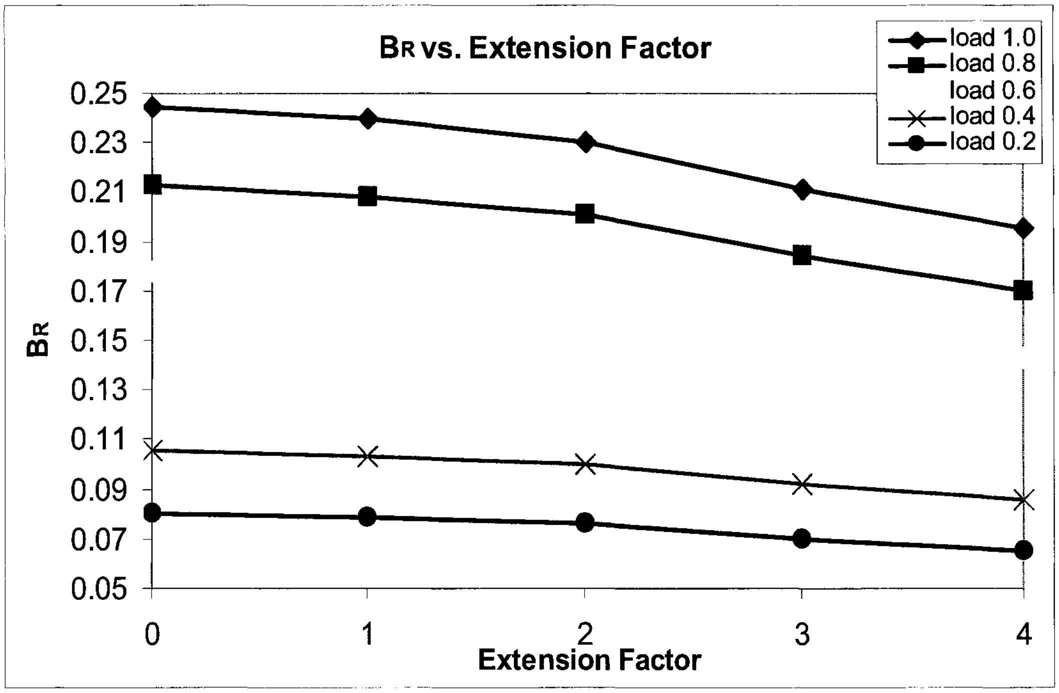


Figure 5.9: Overall BBP under failure vs. time extension factor at different network load

5.4 Simulation Results for an Inter-Connected Ring Network

This section will extend the study of the performance of the enhanced restoration scheme discussed in Chapter 3 on an inter-connected ring topology and corroborate extensively results shown in Section 5.3.

The topology of inter-connected ring network is shown in Figure 5.10. ARPANET consists of 15 nodes with an average nodal degree equals to 2.67, connected by 20 bi-directional links. The average hop distance in the network is 2.41 using Dijkstra's shortest path algorithm.

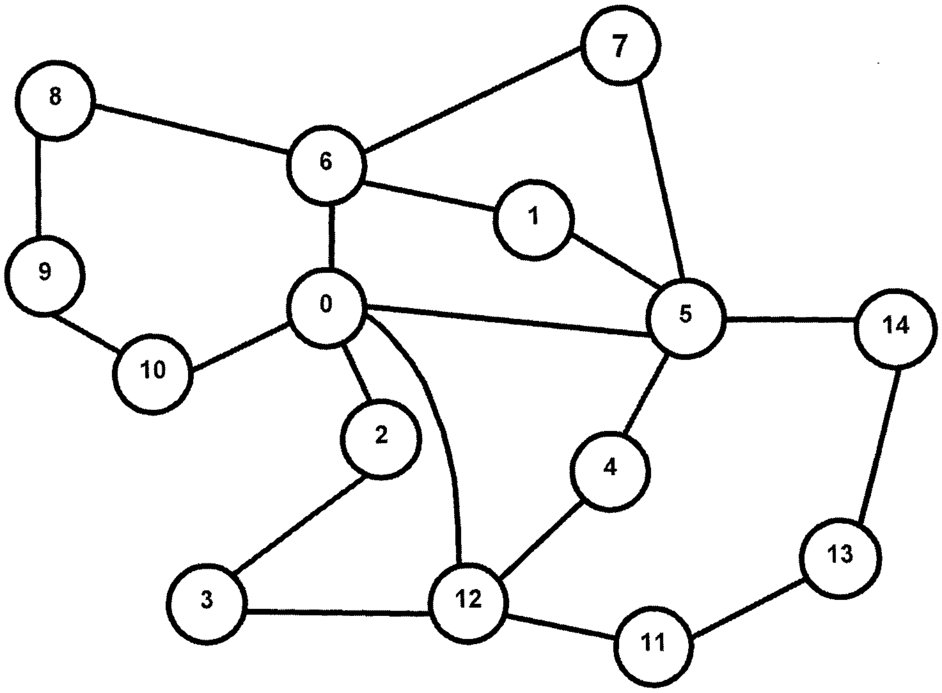


Figure 5.10: Inter-Connected Ring Network's Topology

5.4.1 Offset Time Extension Effect Under No Failures

As discussed earlier, JIT architectures with explicit setup schemes are affected by the change in the value of the *offset* time duration. This effect is reflected on the network's burst loss probability, as shown in Figure 5.11. With OXC configuration time of 5 ms and processing time of 50 μs, the increase in the *offset* time duration is conjunctionally reflected in a rise in the B_N value. The average increase when extending the offset time by a factor of 2 is around 12.5 %.

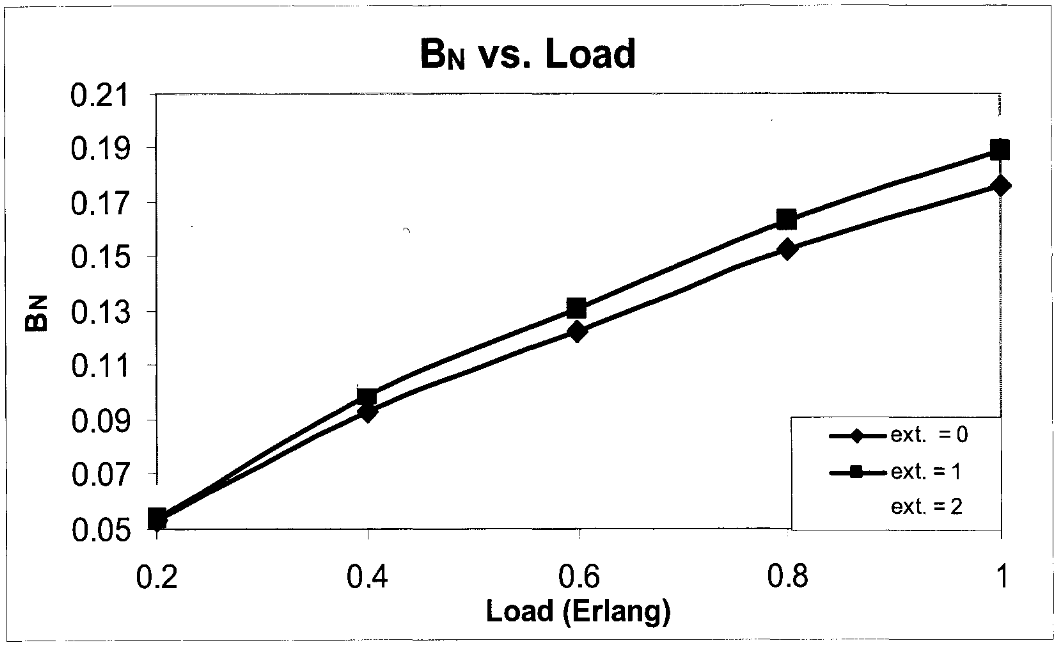


Figure 5.11: Overall BBP vs. network load at different offset time extension in the interconnected ring topology

5.4.2 Offset Time Extension Effect Under Failures

This section shows the effect of the offset time extension on the network performance under link failures. All the results shown assume 40 ms-long bursts, OXC configuration time t_{oxc} = 5 ms, and processing time t_{proc} = 50 μs.

Figure 5.12 shows the average burst blocking due to insufficient time to reserve network resources, B_T, versus the extension factor σ. It is found that at an extension value of 2, the burst loss due to insufficient time to reserve the resources on the alternative routes is zero. It is noted that with $\sigma = 1$, the B_T value does not improve that much since the average alternative route lengths is greater that the primary routes by more than one hop. This will result in and average burst loss close to that with non-extended *offset* time.

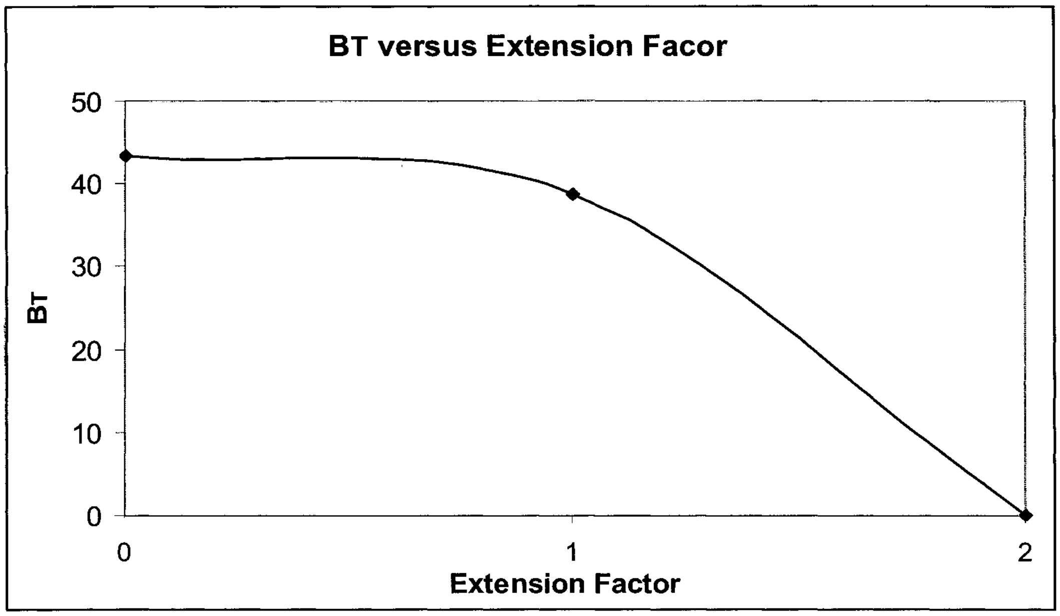


Figure 5.12: Burst Blocking due to Insufficient Time vs. Extension Factor in an interconnected ring topology

Note that figure 5.12 is independent of the value of t_{oxc} or t_{proc}, and is only dependant on the network physical connectivity and the routing algorithm used. For example, if the routing algorithm used computes the link-disjoint shortest cycle (i.e. obeying the link disjoint constraint, the primary and its alternative route form a cycle that is shortest in length, but not necessary that the primary path is the shortest route possible), the maximum value of σ might be larger.

Figure 5.13 shows the overall burst losses B_R with different increments in the *offset* time at different network loads. It is noted that the value of B_R does not improve at σ = 1 and maintains a close level as of when σ = 0. At σ = 2, the value of B_R improves by 17% and thus reduces the overall burst loss level.

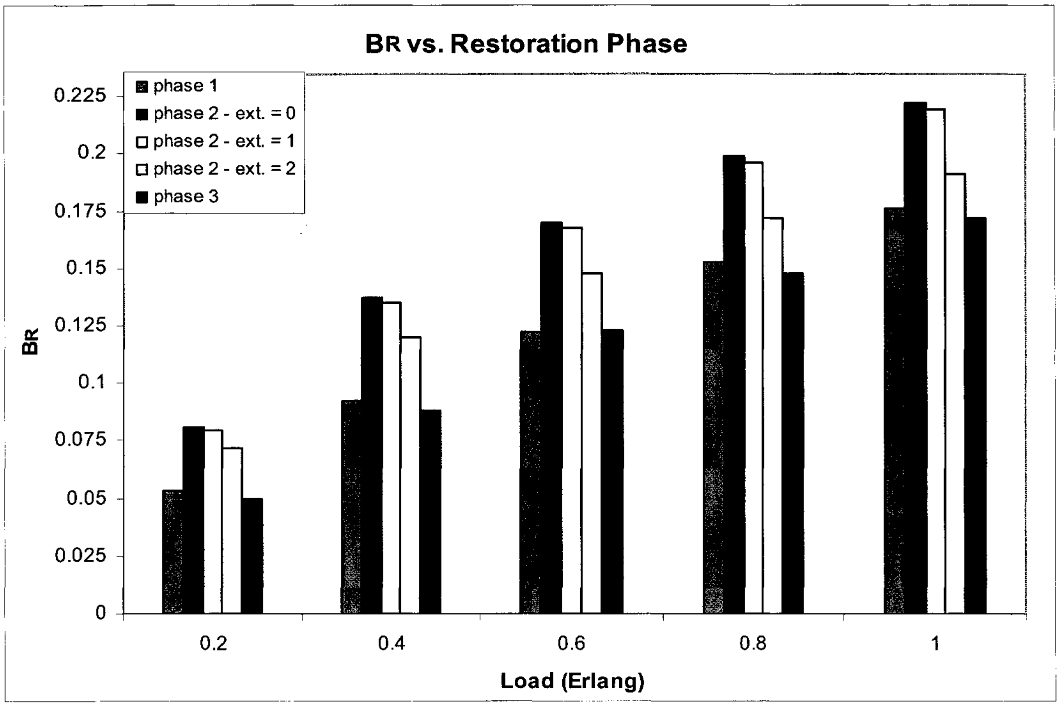


Figure 5.13: Burst Losses under link failure versus the restoration phase at different network loads

When explicit setup scheme is used, the behavior changes as the extension factor increases. At $\sigma = 1$, the value of B_R is greater than that at $\sigma = 0$. The reason behind that is the small effect σ have in reducing the value of B_T, as previously shown in Figure 5.12, and the extended time it introduce while reserving resources. These factors lead to higher burst losses and degradation in the network performance. Inversely, at $\sigma = 2$, the B_T value decreased to zero, and the overall burst losses during restoration was improved by deploying the enhanced restoration algorithm.

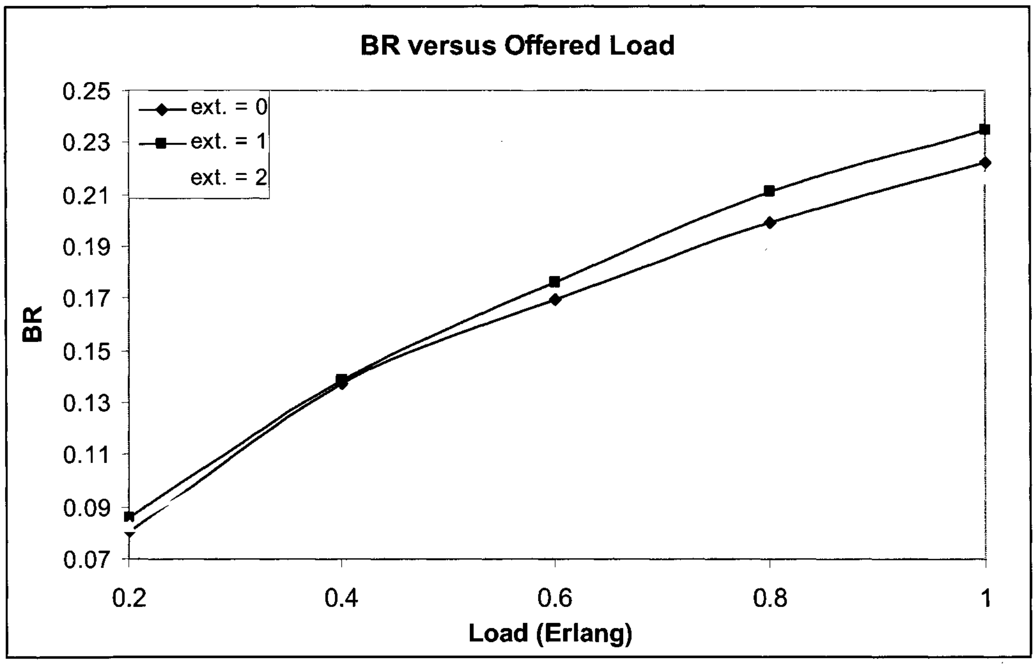


Figure 5.14: Overall BBP under failure vs. network load at different offset time extensions when explicit setup scheme is used

The same effect can be seen in Figure 5.14. An important thing to note down is the σ effect on the burst loss performance at phases 1 and 3. It is very importance to balance between the network loss improvement at phase 2 and the degradations at phases 1 and 3.

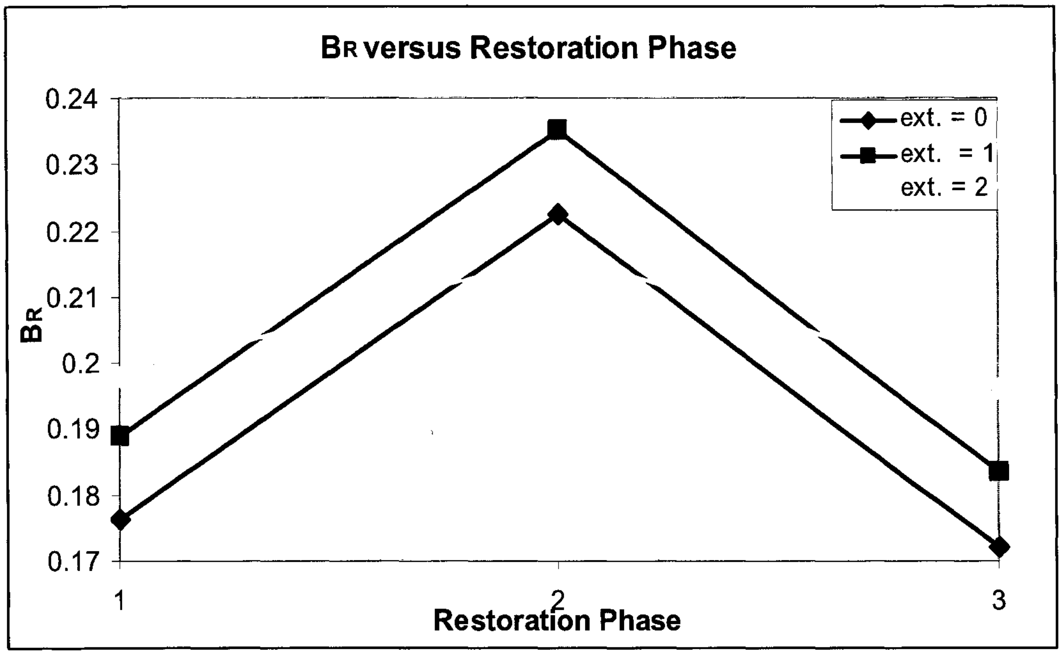


Figure 5.15: Global BBP vs. Restoration phase at Different Extension Values when explicit setup scheme is used

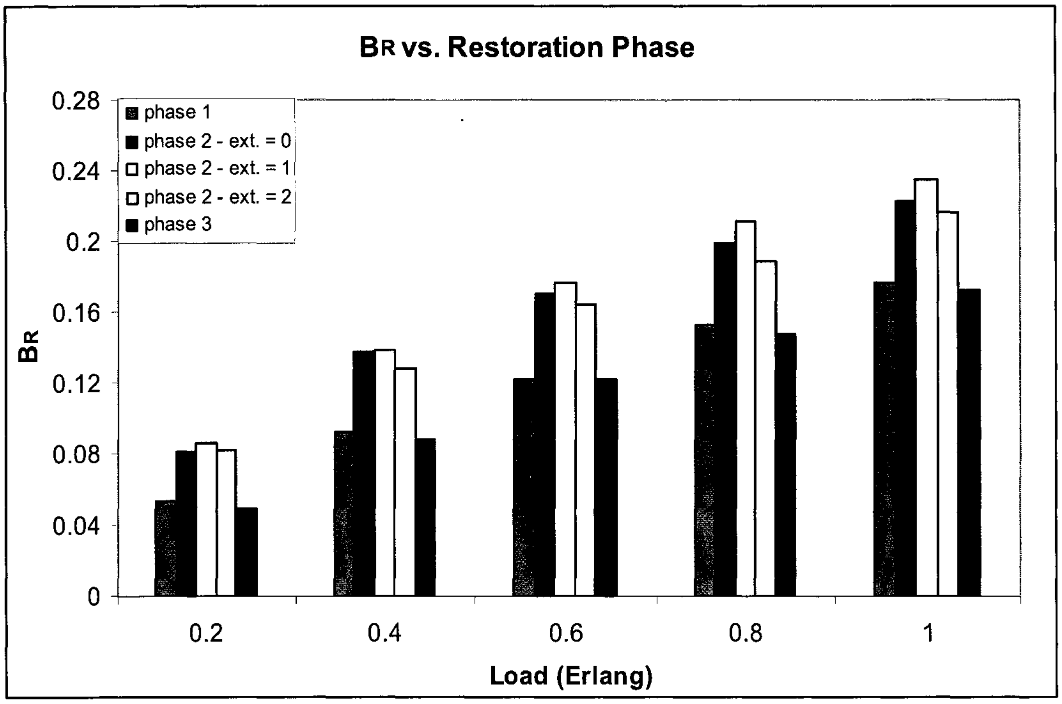
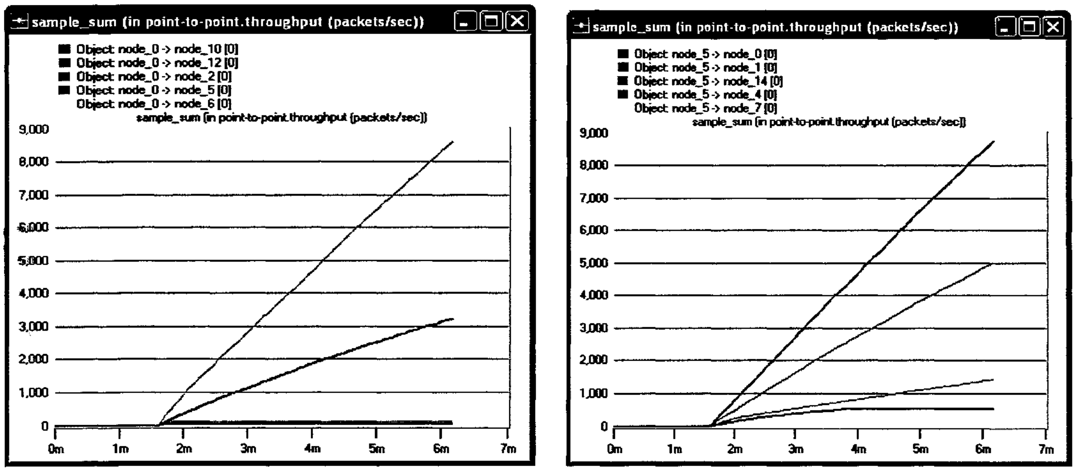


Figure 5.16: Burst Losses under link failure versus the restoration phase at different network loads when explicit setup scheme is used

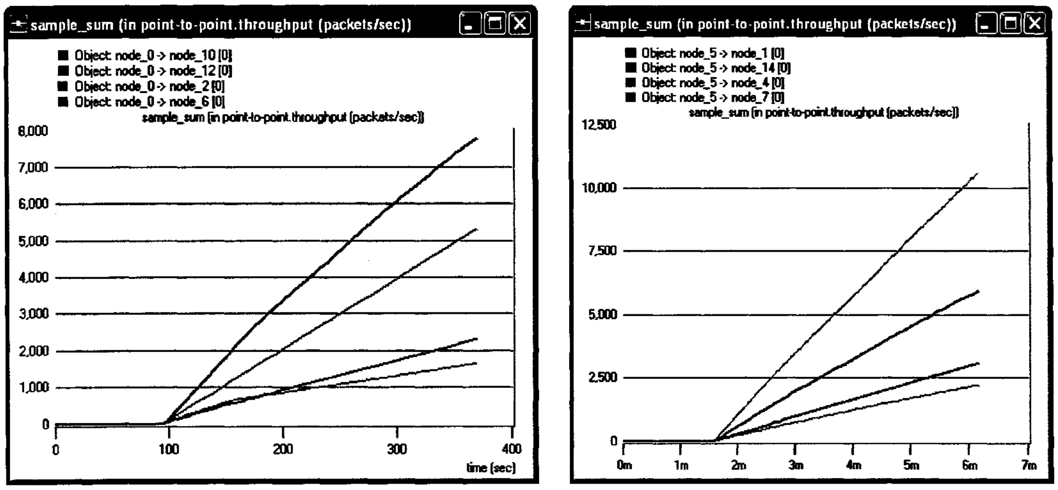
Figure 5.16 shows the different phases at different network loads. The burst losses decreases at $\sigma = 2$ and improves the loss performance by 14.3 %. It is noted that the value of B_R at phase 3 is less than that at phase 1. After viewing the network's link utilization, it was found that some links are under-humiliated by other links. As a result, and upon a link failure and global routing updates, the load in the network will be spread more equally along the links.

Figure 5.17 shows the utilization level on all the links connected to nodes "0" and "5". In the absence of link failures, Figure 5.17 (a),(b) shows that links between nodes 0-2, 0-6, 0-10, 5-1, 5-7, and 5-14 are under utilized. When the link between nodes "0" and "5" fails, the utilization on the links mentioned increases, as shown in Figure 5.18 (a),(b), and thus more resources are available for use. As a result, the overall burst loss probability decreases.



(a) (b)

Figure 5.17: Link Utilization in absence of failures
(a) Link Utilization at node "0" under no failures
(b) Link Utilization at node "5" under no failures



(a) **(b)**

Figure 5.18 Link Utilizations under link failure
(a) Link Utilization at node "0"
(b) Link Utilization at node "5"

5.5 Simulation Results for QoS Restoration Scheme

The novel QoS Restoration Mechanism presented in Chapter 3 will be discussed in this section. Several QoS levels can be defined in correspondence with the extension factor σ and the level of loss they presume. The result will guarantee a service differentiation with respect to the data losses and delays.

Figures 5.5 and 5.12 presented the values of B_T with respect to different extension values, σ. Each σ value corresponds to a certain level of B_T. Getting these values, one can build a table similar to Table 3.8.1 presented in Chapter 3. Assume one wants to create 4 levels of QoS ranging from 0% loss to 100% loss. The corresponding QoS table is shown Table 5.1.

	ARPANET	B_T (%)			
QoS Level	**Delay (sec)**	0-25	25-50	50-75	75-100
1	0	$\sigma_i = 4$			
2	$1 *(t_{prop} + \bar{\varphi})$		$\sigma_i = 3$		
3	$2 *(t_{prop} + \bar{\varphi})$			$\sigma_i = 2$	
3	$3 *(t_{prop} + \bar{\varphi})$			$\sigma_i = 1$	
4	$4 *(t_{prop} + \bar{\varphi})$				$\sigma_i = 0$

Table 5.1: QoS Level Computation Table for ARPANET

Note that both σ = 1,2 define the same service level with respect to the burst loss probability. On the other hand, at σ = 2, the delay is double that at σ = 1.

Similarly, Table 5.2 shows only 2 levels of QoS. Since at σ = 1, both the burst loss and the extra delay encountered are larger that those at σ = 0, one can ignore σ = 1, an only counts on σ = 0, 2.

	Inter-Ring	B_T (%)	
QoS Level	**Delay (sec)**	0-25	50-75
1	0	$\sigma_i = 2$	
2	$2 *(t_{prop} + \bar{\varphi})$		$\sigma_i = 0$

Table 5.2: QoS Level Computation Table for an interconnected ring network

Note that each network can offer up to a maximum ***n*** QoS levels based on the maximum effective extension factor σ (n = σ). This is dependant on the network physical connectivity and routing algorithm used.

Chapter 6

CONCLUSION AND FUTURE RESEARCH

6.1 Summary and Conclusive Remarks

The major goal of this thesis was to study the importance of the "offset time" value on the network operation and its effect on the network performance under normal and failure scenarios.

In the absence of any network failures, extending the offset time in JIT explicit setup schemes showed an increase in the burst loss. This behavior was expected, since as the *offset* time increases, the network resources will be reserved for longer durations, and thus less free available wavelengths will be present leading to a higher blocking probability. On the contrary, in JIT estimated setup schemes, the offset time is independent of the reservation duration and thus will not have any influence on the network's burst loss. The simulation studies explored the effect of the switching time on the network losses. As the switching technology advance, the configuration time will be smaller and thus the degradation in the loss performance will decrease.

The study was then extended to OBS networks under single link failure. In JIT estimated setup schemes, extending the *offset* time duration helped in dipping the burst loss B_T down to zero. This was reflected in reducing the overall network loss probability by 17% in the interconnected ring network and by 20% in the ARPANET. As a result, the enhanced restoration scheme proposed met the expected improvement levels desired.

On the other hand, in JIT explicit setup reservation schemes, it was found that as the extension factor increases the improvement in the overall burst loss was not as significant as that with estimated setup schemes especially after coupling its consequences on the network's BBP during phases 1 and 3. The reason behind that is the high extension time added. This effect was abridged as the OXC configuration time decreased. It is also important to note that the relation between the length of the data burst and the delay encountered at each node plays an important role on the BBP. Therefore, with relatively longer burst duration to the offset time value, one can expect good improvements in the burst loss levels. As a result, our enhanced restoration scheme will improve network performances for JIT explicit setup schemes as switching technology evolves.

Link failures in OBS networks caused higher burst losses and thus reduced significantly the Quality of Service (QoS) in the network. Therefore, a novel QoS Restoration Mechanism was proposed to resolve this issue. Several QoS levels were defined in correspondence with the extension factor σ and the level of loss they presume. The result was a guaranteed service differentiation with respect to the data losses and delays.

6.2 Future Research

Offset time and its effect on the performance on OBS networks was first studied in this thesis. There are still many issues that need further investigation and development. This section discusses some of these issues.

- Fault detection and notification plays an important role in reducing the BBP upon a network failure. This is still an open research issue.
- Latency performance studies upon extending the offset time factor should be further investigated.
- It was found that the maximum effective value of σ is independent of the network load or ant other network parameter. It is directly related to the network physical connectivity and the routing algorithm used. A future research portion is to find the mathematical relation between these parameters.
- The studied performed tried to reduce the value of B_T upon a link failure. Reducing the value of B_L could be further investigated by appropriate

wavelength converter placement or increasing network resources such as deploying more wavelengths of optic fibers.

References

Books

[GRO03] W. Grower, "Mesh-Based Survivable Transport Networks: Options and Strategies for Optical, MPLS, SONET, and ATM Networking", Prentice Hall 2003, ISBN: 013494576.

[JUE05] J. Jue and V. M. Vokkarane, "Optical Burst Switched Networks", Springer 2005, ISBN: 0-387-23756-9.

[MOU02] H.T. Mouftah and P.H. Ho, "Optical Networks: Architecture and Survivability", Kluwer academic Publishers 2002, ISBN: 1402071965.

Journal and Conference Papers

[BAL02] I. Baldine, G. N. Rouskas, H. G. Perros, and D. Stevenson, "JumpStart: A Just-In-Time signaling Architecture for WDM Burst-switched Networks", IEEE Communications Magazine, February 2002, vol. 40, no. 2, pp. 82-89.

[BAL03] I. Baladine, M. Cassada, A. Gragg, G. Karmous-Edwards, D. Stevenson, "Just-In-Time Optical Burst Switching Implementation in the ATDnet All Optical Networking Testbed", Proceedings to GLOBECOM 2003.

[BAT03] T. Battestilli and H. Perros, "An Introduction to Optical Burst Switching", IEEE Communications Magazine, August 2003, vol. 41, no. 8, pp. S10-S15.

[BAT04] T. Battestilli and H. Perros, "Optical Burst Switching for Next Generation Internet", IEEE Potentials, January 2005, vol.23, no.5, pp.40-43.

[CHE04] Y.Chen, C. Qiao, and X.Yu, "Optical Burst Switching: A New Area in Optical Networking Research", IEEE Network, June 2007, vol.18, no.3, pp.16-23.

[DET02] A. Detti, V. Eramo, and M. Listanti, "Optical Burst Switching with Burst Drop (OBS/BD): An Easy OBS Improvement", Proceedings to ICC 2002, April-May 2002, vol.5, no.28, pp. 2687-2691.

[DUE01] M. Dueser and P. Bayvel, "Bandwidth Utilization and Wavelength Re-use in WDM Optical Burst-Switched Packet Networks," Proc. IFIP 5th Working Conf. Opt. Net. Design and Modeling, February 2001, vol.1, pp. 23-24.

[DUS02] M. Duser and P. Bayvel, "Analysis of a Dynamically Wavelength-Routed, Optical Burst Switched Network Architecture", Journal of Lightwave Technology, April 2002, vol.20, no.4, pp.574-585.

[GAU01] C. Gauger, K. Dolzer, J. Späth, and S. Bodamer, "Service Differentiation in Optical Burst Switching Networks", Proceedings to ITG-Fachtagung Photonic Networks, March 2001.

[GE00] A. Ge, F. Callegati, and L. Tamil, "On Optical Burst Switching and Self-SimilarTraffic", IEEE Communiaction Letters, March 2000, vol.4, no.3, pp.98-100.

[GRI03] D. Griffith and S. Lee, " A 1+1 Protection Architecture for Optical Burst Switched Networks", IEEE Journal on Selected Areas in Communications, November 2003, vol. 21, no. 9, pp. 1384-1398.

[JIN04a] T. Jing and G.N. Rouskas, “On Wavelength Assignment in Optical Burst Switched Networks”, Proceedings of the First International Conference on Broadband Networks, 2004, pp.24-33.

[JIN04b] Z. Jing, W. Shuang, Z. Keyao, D. Datta, K. Young-Chon, and B. Mukherjee, “Pre-Planned Global Rerouting for Fault Management in Labeled Optical Burst-switched WDM Networks”, Proceedings to GLOBECOM 2004, vol.3, pp.2004-2008.

[KIM05] S.Kim, J.S. Choi, and M. Kang, “Providing Absolute Differentiated Services for Optical Burst Switching Networks: Loss Differentiation”, IEE Proceedings, August 2005, vol,152, no.4, pp.439 – 446.

[KOR03] E. Korlovski and P. Bayvel, “Link Failure Restoration in Wavelength-Routed Optical Burst Switched (WR-OBS) Networks”, Proceedings to OFC 2003, vol.1, pp.222-223.

[KOZ03] E. Kozlovski, “Survivability of Wavelength-Routed Optical Burst-Switched Networks with Guaranteed IP Services”, Proceedings to ICTON 2003, pp.148-153.

[LEE04] H. Lee, K. Song, W. So, J. Zhang, D. Datta, B. Mukherjee, and Y.C. Kim, “A Hybrid Restoration Scheme Based on Threshold Reaction Time in Optical Burst-Switched Networks”, Proceedings to ICCSA 2004, pp. 994–1003.

[LIA04] W. Liao and C. Loi, “Providing Service differentiation for Optical-Burst-Switched Networks”, Journal of Lightwave Technology, July 2004, vol.22, no.7, pp.1651– 1660.

[LIU02] D.Q. Liu and M.T. Liu "Optical Burst Switching Reservation Process Modeling and Analysis", Proceedings to ICCS 2002, November 2002, vol.2, pp.925-928.

[MAA04] A. Maach, G. Bochmann, and H. Mouftah, "Robust Optical Burst Switching", Telecommunications Network Strategy and Planning Symposium, NETWORKS 2004, June 2004, pp. 447-452

[OHK02] S. Oh, Y. Kim, M. Yoo, and H. Hong, "Survivability in the Optical Internet Using the Optical Burst Switch", ETRI Journal, April 2002, vol.24, no.2, pp.117-130.

[QIA00] C. Qiao "Labeled Optical Burst Switching for IP-over-WDM Integration", IEEE Communication Magazine, September 2000, vol.38, no.9, pp.104-114.

[QIA03] C. Qia and Y. Chen, "The Potential of Optical Burst Switching (OBS)", Proceedings to OFC 2003, vol.1, pp.219-220.

[QIA99] C. Qiao and M. Yoo, "Optical burst switching (OBS)-A new paradigm for an optical Internet", Journal of High Speed Networks, January 1999, vol.8, no.1, pp.69-84.

[QIU05] Y. Qui, Y. Ji, and D. Xu, "A New Survivable OBS Network Using a Multiple Ring Approach", Proceedings to SIPE 2005, vol. 5625, pp. 950-958.

[QUI04] Y. Qiu, Y. Ji, D. Xu, and Z. Gao, "Optical Burst Switching for Metropolitan Area Networks", IEEE Communications Magazine, September 2004, vol.2, no.29, pp.527-530.

[ROD04] J. Rodrigues, N.M. Garcia, and P. Lorenz, "Object-Oriented Modeling and Simulation of Optical Burst Switching Networks", IEEE Communication Society, GLOBECOM 2004 Workshop.

[SCH04] M. Schlosser, H. Woesner, and A. Schroth, "Simulation Model for an Optical Burst Switched Network", Local and Metropolitan Area Networks, April 2004, pp.103- 108.

[TON94] M. To and P. Neusy, "Unavailability Analysis of Long-Haul Networks," IEEE Journal on Selected Areas in Communications, January 1994, vol.12, no.1, pp100-109.

[TUR00] J. S. Turner, "Terabit Burst Switching", Journal of High Speed Networks, January 1999, vol.8, no.1, pp.3-16.

[VER00] S. Verma, H. Chaskar, and R. Ravikanth, "Optical Burst Switching: A Viable Solution for Terabit IP Backbone", IEEE Network Magazine, Nov.-Dec. 2000, vol.14, no.6, pp.48-53.

[VER02] A.J. Vernon and J. D. Portier, "Protection of Optical Channels in All-Optical Networks," 18th Annual National Fiber Optic Engineers Conference (NFOEC 2002), Dallas, TX , September 2002, pp. 1695-1706.

[VUZ02] H. L. Vu and M. Zukerman, "Blocking Probability for Priority Classes in Optical Burst Switching Networks", IEEE Communications Letters, May 2002, vol.6, no.5, pp214-216.

[WEI00] J. Y. Wei and R. I. McFarland, "Just-in-time Signaling for WDM Optical Burst Switching Networks" Journal of Lightwave Technology, December 2000, vol.18, no.12, pp.2019-2037.

[XIN04a] Y. Xin, J. Teng, G.Karmous-Edwards, G. Rouskas, and D. Stevenson, "Fault Management with Fast Restoration for Optical Burst Switched Networks", Proceedings to BroadNets October 2004.

[XIN04b] Y. Xin, J. Teng, G.Karmous-Edwards, G. Rouskas, and D. Stevenson, "A Novel Fast Restoration Mechanism for Optical Burst Switched Networks", Proceedings to BroadNets, October 2004.

[XIO00] Y. Xiong, M. M. Vandenhoute, and H. Cankaya, "Control Architecture in Optical Burst-Switched WDM Networks", Proceedings to IEEE JSAC, October 2000, vol.18, no. 10, pp.1838-1851.

[XUQ03] J. Li J. Xu, C. Qiao, and G. Xu, "Efficient Channel Scheduling Algorithms inOptical Burst Switched Networks," Proceeding to INFOCOM, April 2003.

[YOO00] 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, October 2000, vol. 18, no. 10, pp2062-2071.

[YOO01] M. Yoo, C. Qiao, and S. Dixit, "Optical Burst Switching for Service Differentiation in the Next-Generation Optical Internet", IEEE Communications Magazine, February 2001, vol.39, no.2, pp.98-104.

[YU03] F.Yu, D. Wang, R. Sinha, G. Li, B. Doverspike, and C. Kalmanek, "Hybrid Centralized/Distributed Approach to Optical Network Restoration", Proceedings to OFC 2003, vol. 2, pp. 778-779.

[ZHA04] J. Zhang, S. Wang, K. Zhu, D. Datta, Y. Kim, and B. Mukherjee, "Pre-planned Global Rerouting for Fault Management in Labeled Optical Burst-

Switched WDM Networks", Proceedings to GLOBECOM 2004, vol.3, no.29, pp.2004-2008.

[ZHA04] Q. Zhang, V.M Vokkarane, J.P. Jue, and C. Biao, "Absolute QoS Differentiation in Optical Burst-Switched Networks", IEEE Journal on Selected Areas in Communications, November 2004, vol.22, no.9, pp.1781-1795.

List of Publications

[1] S.Y. Said, H. Elbiaze, and H.T. Mouftah, "*A Novel Offset-Based Restoration Mechanism for Optical Burst Switched Networks*", submitted to ICC2006, Istambul, Turkey.

[2] S.Y. Said, H.T. Mouftah, and H. Elbiaze "*A QoS-Based Restoration Mechanism for Optical Burst Switched Networks*", submitted to OFC2006, Anaheim, California, USA.