

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600

UMI<sup>®</sup>



H

**Optical Networking and Real-Time Provisioning;  
An Integrated Vision For a  
Next Generation Survivable Optical Internet**

by

**Chadi Assi**

A dissertation submitted to the Graduate Faculty in Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy

The City University of New York

2003

UMI Number: 3103079

Copyright 2003 by  
Assi, Chadi M.

All rights reserved.

UMI<sup>®</sup>

---

UMI Microform 3103079

Copyright 2003 by ProQuest Information and Learning Company.  
All rights reserved. This microform edition is protected against  
unauthorized copying under Title 17, United States Code.

---

ProQuest Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

© 2003

**Chadi Assi**

**All Rights Reserved**

This manuscript has been read and accepted for the Graduate Faculty in Engineering in satisfaction of the dissertation requirements for the degree of Doctor of philosophy.

6-30-2003

Date

  
\_\_\_\_\_

Prof. Mohamed A. Ali  
Chair of Examining Committee

6-30-2003

Date

  
\_\_\_\_\_

Dean Mumtaz K. Kassir  
Executive Officer

Prof. Samir Ahmed (EE Dept. CCNY, CUNY)

Prof. Gorgios Ellinas (EE Dept. CCNY, CUNY)

Prof. Barry Gross (EE Dept. CCNY, CUNY)

Dr. Andrew Wallace (AT&T)

Supervisory Committee

The City University of New York

## Abstract

# **OPTICAL NETWORKING AND REAL TIME PROVISIONING; AN INTEGRATED VISION FOR A NEXT GENERATION SURVIVABLE OPTICAL INTERNET**

**By**

**Chadi Assi**

**Advisor: Professor Mohamed A. Ali**

Recently, there has been a dramatic increase in data traffic, driven primarily by the explosive growth of the Internet. At the same time, the rise of optical networking, first with wavelength-division multiplexing (WDM) transmission technology and more recently with optical multiplexers, and optical cross-connects (OXC)s devices, is moving us towards the vision of creating an intelligent “all-optical” Internet; a vision that offers the promise of satisfying the bandwidth requirement of the Internet infrastructure.

This thesis addresses the important problem of how to dynamically provision, maintain, and, in the event of a network failure, restore both high-speed connection requests (full wavelength) and low-rate traffic streams (sub-lambda connections) in future IP/MPLS-centric WDM-based optical networks. It focuses on the implementation,

modeling, and analysis of various protocols and algorithms associated with this next generation of optical networks.

The main body of this work is divided into two main overlapping phases. The first phase addresses dynamic routing and signaling associated with provisioning, maintaining, and restoring full optical channels/lambda (lightpaths) at the physical layer. We propose two constraint-based dynamic algorithms for real time provisioning of optical channels: the semi-adaptive, where partial information is available for the routing algorithm to make its selection decision and the full adaptive routing where global knowledge of network resources is made available through optical extensions to existing IP-based routing algorithms. We develop analytical models to evaluate the performance of the proposed provisioning algorithms and we validate their accuracy through numerical comparisons with simulation results.

Moreover, due to the high bandwidth involved, any link failure in the form of fiber cut will have catastrophic results unless protection and restoration schemes form an integral part of the network design and operation strategies. Protection mechanisms (dedicated or shared) involve the pre-provisioning of backup resources during the call admission; alternatively, restoration is more dynamic and does not involve any beforehand reservation. Clearly, the latter approach exhibits longer restoration times and cannot guarantee complete service recovery; thus protection mechanisms are more attractive. Here, fault tolerant routing algorithms are required to compute and select link/node-disjoint paths between each s-d pair. Working and protection capacity are then reserved, through signaling protocols, on each path. We develop efficient provisioning

algorithms for shared path protection that exploits the sharability property to improve the overall network performance. We compare the performance of different decentralized selection algorithms, namely the Distributed Path selection with Global and Local Information (DPGI, DPLI respectively). Simulation models are used in conjunction with analytical modeling for our numerical comparison.

Fault management protocol deals with recovering the affected connections after the failure occurrence. While dedicated protection results in faster service recovery, the overhead associated with the excessive waste of resource cannot be tolerated. On the other hand, in shared protection switches along the protection path are only reserved and they need to be configured upon fault detection. Here, message propagation and switch configuration times are the main source of delays; we propose a *novel method* that eliminates these components combined together, and as a result bounding the restoration time to a minimal value.

The second phase addresses dynamic routing and signaling associated with provisioning, maintaining, and restoring low-rate traffic streams (sub-lambda) at the logical (IP/MPLS) layer. In this scenario, the IP/ MPLS routers are clients of the Optical Transport Network (OTN) and are connected to their peers over dynamically switched lightpaths spanning potentially multiple OXCs. The set of light paths built on top of the physical topology represents the logical connectivity between the IP/MPLS routers connected to the OTN. That is, two IP/MPLS routers are considered to be adjacent if they are connected by a direct lightpath. To dynamically provision low-rate traffic streams at

the IP/MPLS layer, we develop several dynamic routing algorithms that provision connections on the logical topology built on top of the underlying physical connectivity connecting IP/MPLS routers. In this case, our simulation results have demonstrated that low-rate traffic streams may best be dynamically provisioned at the IP/MPLS layer. Moreover, to provision restorable connections, we design a survivable logical topology on top of the underlying physical connectivity; we have developed a novel integrated protection scheme to dynamically allocate restorable-bandwidth guaranteed paths for designing IP/MPLS over WDM networks that can protect against single optical link/node failures.

***This thesis is dedicated  
to my parents for their continuous support.  
I am so grateful for them.***

## ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who contributed and gave me the possibility to complete this thesis. First, I am deeply indebted to my supervisor Professor Mohamed Ali for all his support, guidance, discussions and exhaustive editing of my publications and thesis. I have to admit that without his continuous support, it would have been impossible to put up together all of this work.

I am also indebted to many people in the electrical engineering department; particularly, Professor Samir Ahmed, who has been a source of enthusiasm and encouragement over many years; Professor Barry Gross for all the stimulating discussions we had. I also would like to thank Professor George Ellinas and Dr. Andrew Wallace for serving on my PhD committee.

Much appreciation is given to Abdallah Shami for all the discussions and cooperation and for all the wonderful time we have shared. I also wish to thank people with whom I have had the opportunity to work and discuss with during the past few years, including Yinghua Ye, Ahmed Khalil, Zhensheng Zhang, Nasir Ghani, Sudhir Dixit, Prof. Tarek Saadawi, and Prof. Ibrahim Habib.

I am very grateful to my parents, my brothers and my sister for their constant support and encouragement. Finally, I would like to thank Laura for all the support, patience and love she had shown throughout my studies.

1. Introduction.....	1
1.1 Introduction.....	1
1.2 Thesis Statement .....	9
2. Real-Time Provisioning in WDM Mesh Networks .....	12
2.1 Introduction.....	12
2.2 IP over Optical network architectures.....	15
2.2.1 Interconnection Models: .....	16
2.2.2 The Overlay Model: .....	17
2.2.3 The Peer Model:.....	18
2.2.4 The Augmented Model: .....	18
2.3 Routing and Wavelength Assignment (RWA) .....	19
2.3.1 Overview of RWA: .....	20
2.3.2 Dynamic Lightpath Computation: .....	23
2.4 Analytical Model .....	25
2.4.1 Network and traffic model assumptions: .....	26
2.4.2 Fixed routing:.....	28
2.4.3 Fixed alternate routing: .....	31
A- Fixed Alternate Routing with first fit wavelength assignment.....	31
B- Fixed Alternate Routing with Trunk Reservation.....	32
2.5 Dynamic Provisioning .....	33
2.5.1 Algorithm I “Full Adaptive Routing”:.....	36
2.5.2 Algorithm II “Semi-Adaptive Routing”: .....	40
2.6 Performance evaluation .....	42

2.6.1 Accuracy of the analytical model: .....	42
2.6.2 Performance of the proposed dynamic algorithms: .....	48
2.7 Conclusion .....	50
2.8 References .....	51
3. Survivability in WDM-based Networks .....	55
3.1 Introduction.....	55
3.2 Background on Survivability .....	58
3.2.1 Survivability in Non-WDM Networks: .....	58
3.2.1.1 Pre-designed Protection .....	58
3.2.1.2- Dynamic Restoration:.....	62
3.2.2 Survivability in WDM optical mesh networks: .....	63
3.2.2.1- Link-Based Protection: .....	64
3.2.2.2- Path-Based Protection: .....	66
3.3 Routing in survivable WDM mesh networks.....	68
3.3.1- Background and Motivation: .....	70
3.3.2- Efficient Path Selection Algorithms: .....	71
3.3.2.1- Distributed Path selection with Global Information, DPGI: .....	72
3.3.2.2- Distributed Path selection with Local Information, DPLI:.....	73
3.4 Analytical model for shared path protection.....	75
3.4.1- Network and traffic assumptions:.....	76
3.4.2- End-to-end path blocking: .....	78
3.4.2.1- Property 1: .....	79
3.4.2.2- Computation of $\alpha_j$ .....	80

3.5 Performance evaluation .....	82
3.6 Conclusion .....	88
3.7 References.....	89
4. Connection Management for Survivable WDM Networks .....	91
4.1 Introduction.....	91
4.2 Connection Management in WDM networks .....	92
4.3 Connection management in survivable WDM mesh networks.....	98
4.3.1- Background and motivations:.....	98
4.3.2 – The network model: .....	100
4.3.3- Protocol description.....	101
4.3.3.1- 1:1 Dedicated Protection: .....	103
4.3.3.2- Shared Path Protection: .....	104
4.3.4- Wavelength assignment.....	107
4.4 Fault management and fast restoration algorithms .....	109
4.4.1- Source-based restoration (SBR): .....	110
4.4.2- Destination-based restoration (DBR): .....	112
4.4.3- Enhanced SBR:.....	113
4.5 Performance Evaluation.....	117
4.6 Conclusion .....	125
4.7 References:.....	126
5. Real-Time Provisioning in IP-Over-WDM Networks.....	128
5.1- Introduction .....	128
5.2- IP over WDM Network Architecture .....	132

5.3- Sub-wavelength Path Provisioning.....	135
5.4. Constraint-Based Routing Algorithms.....	140
5.4.1 General Approach .....	140
5.4.2 The First proposed Algorithm:.....	141
5.4.3 The Second Proposed Algorithm:.....	142
5.4.4 The Two Adaptive Constraint-based Routing Schemes: .....	143
5.4.4.1- Least Loaded Routing (LLR): .....	145
5.4.4.2- Most Loaded Routing (MLR):.....	146
5.5- Fault Tolerant Routing in IP over WDM Network: .....	146
5.5.1- The Non-Information Routing:.....	147
5.5.2- The Complete Information Routing: .....	148
5.5.3- Sharing Restoration Capacity:.....	148
5.6- Performance Evaluation .....	153
5.6.1- Evaluation of the different constraint-based routing algorithms:.....	153
5.6.2- Evaluation of the integrated-protection algorithms:.....	157
5.7- Conclusion.....	161
5.8 References.....	162
6. Conclusions and Future Work .....	165
6.1 Real-Time Provisioning in WDM Optical Mesh Networks .....	165
6.2 Survivability in WDM-based Networks .....	167
6.3 Connection Management for Survivable WDM Networks .....	167
6.4 Real-Time Provisioning in IP-Over-WDM Networks.....	168
6.5- Future Work.....	169

**Bibliography.....172**

## List of Figures

Fig. 2.1 Optical Networking Model.....	17
Fig. 2.2 The sample network model, and its multi-layered graph approach.....	36
Fig. 2.3 Illustration Examples.....	40
Fig. 2.4 FR with FF wavelength assignment.....	45
Fig. 2.5 FR with FF wavelength assignment (Ring Network).....	45
Fig. 2.6 FR with different wavelength assignment schemes .....	46
Fig. 2.7 FAR and FR with FF wavelength assignment.....	46
Fig. 2.8 FAR with Trunk Reservation (FARwTR).....	47
Fig. 2.9 Blocking probability vs. Load, for different routing schemes .....	49
Fig. 2.10 Comparison of different selection schemes for adaptive routing case.....	49
Fig. 3.1 Automatic protection switching.....	60
Fig. 3.2 The self-healing ring .....	61
Fig. 3.3 Link-based protection .....	66
Fig. 3.4 Path-based protection .....	68
Fig. 3.5 Simulated Network Topologies .....	83
Fig. 3.6 Comparison of Analytical model and simulation.....	84
Fig. 3.7 Comparison between different path selection algorithms.....	85
Fig. 3.8 Impact of path selection on sharability gain .....	86
Fig. 3.9 Impact of path selection on average hop count.....	87
Fig. 4.1 Lightpath Setup in WDM Networks.....	97
Fig. 4.2 Network Architecture.....	101
Fig. 4.3 Failure Detection/Notification with SBR .....	111

Fig. 4.4 Failure Detection/Notification with DBR .....	113
Fig. 4.5 SBR with Offset Time Based Restoration.....	115
Fig. 4.6 SBR with Enhanced Offset Time Based Restoration.....	117
Fig. 4.7 The 16-Nodes NSF Network.....	118
Fig. 4.8 Blocking probability vs. load for dedicated and shared protection.....	119
Fig. 4.9 First Fit vs. MSW wavelength assignments.....	120
Fig. 4.10 CST for dedicated (1:1) and shared protection vs. Load.....	121
Fig. 4.11 Restoration time vs. Load for different restoration algorithms.....	122
Fig. 4.12 Impact of the SCT on the RT.....	124
Fig. 5.1 IP network over WDM network.....	130
Fig. 5.2 IP over WDM node architecture.....	133
Fig. 5.3 Logical connectivity between IP/MPLS routers.....	134
Fig. 5.4 Service Provisioning/Aggregation at the Router Level.....	138
Fig. 5.5 Label binding for sub-wavelength provisioning at the IP-layer.....	139
Fig. 5.6 Example of a Survivable IP over WDM network.....	150
Fig. 5.7 Different routing algorithms.....	154
Fig. 5.8 Impact of adaptive routing.....	155
Fig. 5.9 Network performance vs. nb of wavelengths vs. nb. of transceivers.....	157
Fig. 5.10 Comparison of all protection schemes.....	158
Fig. 5.11 Impact of restoration capacity sharing on the network performance.....	159
Fig. 5.12 Blocking probability vs. Router Processing Speed.....	161

## Chapter 1

# 1. Introduction

## 1.1 Introduction

The change in the fundamental character of backbone network traffic, as demonstrated by the current shift in the telecommunications industry from traditional voice-centric TDM/SONET (Time Division Multiplexing/Synchronous Optical Networks) optimized circuit-switched networking paradigm to data-centric packet-optimized networking paradigm, is leading to revolutionary changes in the traditional concepts of how networks are constructed. The primary reason for the paradigm shift in network design is the nature of the traffic crossing today's long-haul backbones. Specifically, Internet Protocol (IP) applications are the fastest growing segment of a service provider's network traffic. This growth is expected to continue well into this century.

Fortunately, the Internet's birth coincided with several technological advances that have enabled carriers to cope with such explosive growth in data traffic. The most important development, of course, is the rise of optical networking, first with wavelength-division multiplexing (WDM) and more recently with optical amplifiers, and optical cross-connects (OXC's) - with the ability to add, drop, and in effect construct wavelength-switched and wavelength-routed networks. Another important

development was the proliferation of data-centric equipment, such as high performance IP/MPLS routers. These network elements supporting OC-12, OC-48, and OC-192 interfaces, could be interfaced directly to WDM transport systems, thus eliminating intermediate SONET multiplexing equipment.

These advances are setting the stage for the creation of an “All Optical Internet”, one that is free from all bandwidth and scalability restrictions, as well as distance limitations. High-performance routers (IP layer) plus an intelligent optical transport (DWDM) layer equipped with a new breed of photonic networking equipment hold the keys to the two pieces of the puzzle that will comprise the next-generation Optical Internet. Of course, there are several visions of how best to marry these two technologies. The key question appears to be whether carriers should control their next-generation data-centric networks using only routers, or some combination of routers and OXC equipment.

The realm in which we are likely to see the first real deployments of high-performance routers interconnected by intelligent optical core networks that utilizes optical networking technology will likely be within the long-haul backbone. Several carriers have already completed field trials of all-optical systems, and many others have their trials in process. The key question is how will the migration to this new data-centric optical networking paradigm evolve? The key ingredient to the success of any migration strategy is that the new networking paradigm must coexist with the aging TDM circuit switching infrastructure, offer the network reliability inherent in

legacy SONET systems and at the same time meet the surging demand for packet traffic. In other words, how will the legacy SONET-based infrastructure currently in place make a graceful transition to the next generation data-centric networking paradigm is the real challenge?

### **1.1 Today's Problems**

The fundamental problem facing both carriers and service providers is that Internet traffic is already skyrocketing, and look certain to accelerate even further as broadband access technologies such as DSL (digital subscriber line) and cable modems are rolled out to the masses. Service providers are faced with daunting challenges that span both the access and backbone transport network requirements. They must offer service breadth—IP, Frame Relay, ATM, Ethernet and TDM-based voice – plus the flexibility to add and modify new services as market conditions warrant. While transporting data services, the new data-centric networking paradigm must coexist with the voice infrastructure and offer the network reliability inherent in legacy SONET/SDH systems.

Let us first consider the trouble with today's most widely used standard, SONET. SONET infrastructure, that has been the standard for wide area networking for the past 15 years, ideally designed to handle circuit-switched voice and point-to-point TDM connections, began to feel burdened by growing data traffic demands and the need to deliver real-time circuit provisioning. This is because SONET was

designed before the explosion of the asynchronous and unpredictable nature of Internet-based traffic profiles.

The asynchronous and unpredictable nature of Internet-based traffic profiles introduces a great deal of unpredictability into the network, which makes forecasting for growth all the more difficult, and the issue of scalability a major concern. Unfortunately, SONET was never designed to deal with these problems. SONET networks are always built under the assumption that transit/receive traffic is always balanced and predictable, and growth was easy to forecast.

SONET's inability to handle the huge influx of data traffic, and the associated unpredictable traffic patterns, signaled the end of this standard's long reign in the backbone transport network. While long-haul architectures based solely on SONET will be far too expensive and complex to meet the needs of tomorrow's backbones, some of the functionality SONET encompasses will live on in different, more efficient vehicles. The emerging data-centric networking paradigm "a purely Optical Internet", by contrast, will bypass SONET altogether, providing for a backbone without capacity restrictions or distance limitations. So, while some SONET functionality is still needed, a separate SONET layer is not.

That is not the end of the traditional network's troubles. Another major issue facing today's service providers is the static and cumbersome process of provisioning end-to-end circuits. Provisioning a cross-country OC-48 service is extremely manual,

involves complex network planning and rollout activities, and generally takes several months to accomplish. The rollout of DWDM has led to another problem; current digital cross-connects simply are not big enough to handle the number of connections running across backbones. Carriers can foresee needing huge switch fabrics with thousands of ports, and right now, they do not exist.

## **1.2 Tomorrow's Promises**

All of this is about to change. Several dramatic advances in the optical-networking arena in recent years have emerged to challenge both the traditional view of networking (e. g., routing, switching, provisioning, protection, and restoration). First, the abundance of bandwidth propelled by the explosion of DWDM poses, for the first time, a new challenge to network architects. Whereas in the past, IP, ATM and related routing protocols have focused on managing the scarcity of bandwidth, the new challenge is managing the abundance of bandwidth. Second, rapid advances in DWDM technology have presented an attractive opportunity to evolve DWDM technology toward an optical networking infrastructure with transport, multiplexing, switching, routing, survivability, bandwidth provisioning and performance monitoring, supported at the optical layer.

Harnessing the newly available bandwidth is difficult because today's core network architecture - and the networking equipment it is built on - lack both the scale and functionality necessary to deliver the large amounts of bandwidth required

in the form of ubiquitous, rapidly scalable, multi-gigabit services. Today's core network architecture model has four layers: IP and other content bearing traffic, over ATM for traffic engineering, over SONET for transport, over WDM for fiber capacity. This approach has functional overlap among its layers, contains outdated functionality, and will not be able to scale to meet the exploding volume of future data traffic, which make it ineffective as the envisioned architecture for next generation data-centric networking paradigm.

A simplified, two-tiered architecture that require two types of sub-systems will set the stage for a truly optical Internet: service delivery platforms that enforce service policies; and transport platforms that intelligently deliver the necessary bandwidth to these service platforms. If IP can be mapped directly onto the WDM layer, some of the unnecessary network layers can be eliminated, opening up new possibilities for the potential of collapsing today's vertically layered network architecture into a horizontal model where all network elements work as peers to dynamically establish optical paths through the network.

This simplified two-tiered architecture Optical Internet is defined here as mapping IP traffic directly over the optical layer as there is no underlying SONET/SDH and ATM transport layer/protocol. This architecture is commonly referred to as the IP-over-WDM problem. The IP-over-WDM problem has two fundamental sub-problems: (i) the user-plane protocol stack problem, and (ii) the routing problem. The user-plane protocol stack problem consists of defining the

protocol stack for carrying IP packets over multi-wavelength optical links, and the corresponding network architecture. Candidate protocol stacks include: (1) IP over AAL/ATM over SONET/SDH over WDM, (2) IP over PPP/HDLC over SONET/SDH over WDM, (3) IP over PPP/HDLC over WDM, and (4) IP/MPLS directly over WDM. In accordance with the two-tiered architecture definition of an Optical Internet described above, this work considers the two-layer protocol stack (IP directly over WDM) as the only viable candidate model for building the next-generation Optical Internet.

Now that the basic building blocks are available for building such a “network of networks”, the key innovations will come from adding intelligence that enables the inter-working of all the network elements (IP/MPLS Routers, ATM switches, DWDM transmission systems and OXCs). A critical issue for realizing such intelligent optical networks is how to provide the desired features of rapid provisioning/restoration and automated capabilities between the optical layer and the client layers. It is widely accepted that the best way to achieve this is to adapt the IP topology self-discovery and routing capabilities to the optical network environment.

Current research focuses on the use of distributed management schemes such as multi-protocol label switching (MPLS) to provide the control plane necessary to ensure automated, intelligent provisioning and maintaining connections and managing network resources. In this type of application the label is the wavelength of the incoming signal; hence, the term multi-protocol lambda switching (MP $\lambda$ S) is

more commonly used. Industry organizations like the OIF and IETF are rushing to extend MPLS-framework (Generalized-MPLS) to support not only devices that perform packet switching, but also those that perform switching in time, wavelength, and space. The main goal of these initiatives is to provide a framework for real-time provisioning of optical channels, through combining recent advances in Multi-protocol label switching (MPLS) traffic-engineering control plane with emerging optical switching technology in a hybrid IP-centric optical network.

The following are some of the salient features and associated requirements that we have taken as guidelines during the course of this work: We strongly believe that these features are the underpinning en-route towards implementing the envisioned “Optical Internet”:

- (1) All conventional aspects of networking functionality have to be reconsidered; the most basic assumptions of how networks are constructed are being called into question. At a minimum, service providers, carriers and enterprise network managers will be forced to fully rethink routing, switching, and traffic-engineering for the emerging data-centric networking paradigm.
- (2) SONET transport should give way to optical transport; fast restoration (50 ms) is necessary; without it, a significant availability and reliability benchmark set by SONET would be lost.
- (3) Rapid provisioning of high-speed connection requests (full wavelength) as well as low-rate traffic streams (sub-lambda) must be an integral part of the new architecture.
- (4) ATM cell granularity and traffic engineering are made obsolete by the use of MPLS at the IP level and wavelength traffic engineering at the optical transport layer.

(5) Restoration will take place at the optical transport layer in a fast, scalable manner. The service layer (IP/MPLS) will be informed of the event only if the transport layer cannot restore due to lack of transport resources. Then, the service layer will be advised to perform its restoration functions. The two will perform in a non-overlapping, predictable, and scalable manner.

## 1.2 Thesis Statement

This thesis addresses the important problem of how to dynamically provision, maintain, and, in the event of a network failure, restore both high-speed connection requests (full wavelength) and low-rate traffic streams (sub-lambda connections) in future IP/MPLS-centric WDM-based optical networks. It focuses on the implementation, modeling, and analysis of various protocols and algorithms associated with this next generation of optical networks.

The main body of this work is divided into two main overlapping phases. The first phase addresses dynamic routing and signaling associated with provisioning, maintaining, and restoring full optical channels/lambda (lightpaths) at the physical layer. We propose two constraint-based dynamic algorithms for real time provisioning of optical channels: the semi-adaptive, where partial information is available for the routing algorithm to make its selection decision and the full adaptive routing where global knowledge of network resources is made available through optical extensions to existing IP-based routing algorithms. We develop analytical

models to evaluate the performance of the proposed provisioning algorithms and we validate their accuracy through numerical comparisons with simulation results.

Moreover, due to the high bandwidth involved, any link failure in the form of fiber cut will have catastrophic results unless protection and restoration schemes form an integral part of the network design and operation strategies. Protection mechanisms (dedicated or shared) involve the pre-provisioning of backup resources during the call admission; alternatively, restoration is more dynamic and does not involve any beforehand reservation. Clearly, the latter approach exhibits longer restoration times and cannot guarantee complete service recovery; thus protection mechanisms are more attractive. Here, fault tolerant routing algorithms are required to compute and select link/node-disjoint paths between each s-d pair. Working and protection capacity are then reserved, through signaling protocols, on each path. We develop efficient provisioning algorithms for shared path protection that exploits the sharability property to improve the overall network performance. We compare the performance of different decentralized selection algorithms, namely the Distributed Path selection with Global and Local Information (DPGI, DPLI respectively). Simulation models are used in conjunction with analytical modeling for our numerical comparison.

Fault management protocol deals with recovering the affected connections after the failure occurrence. While dedicated protection results in faster service recovery, the overhead associated with the excessive waste of resource cannot be tolerated. On

the other hand, in shared protection switches along the protection path are only reserved and they need to be configured upon fault detection. Here, message propagation and switch configuration times are the main source of delays; we propose a *novel method* that eliminates these components combined together, and as a result bounding the restoration time to a minimal value.

The second phase addresses dynamic routing and signaling associated with provisioning, maintaining, and restoring low-rate traffic streams (sub-lambda) at the logical (IP/MPLS) layer. In this scenario, the IP/ MPLS routers are clients of the Optical Transport Network (OTN) and are connected to their peers over dynamically switched lightpaths spanning potentially multiple OXCs. The set of light paths built on top of the physical topology represents the logical connectivity between the IP/MPLS routers connected to the OTN. That is, two IP/MPLS routers are considered to be adjacent if they are connected by a direct lightpath. To dynamically provision low-rate traffic streams at the IP/MPLS layer, we develop several dynamic routing algorithms that provision connections on the logical topology built on top of the underlying physical connectivity connecting IP/MPLS routers. In this case, our simulation results have demonstrated that low-rate traffic streams may best be dynamically provisioned at the IP/MPLS layer. Moreover, to provision restorable connections, we design a survivable logical topology on top of the underlying physical connectivity; we have developed a novel integrated protection scheme to dynamically allocate restorable-bandwidth guaranteed paths for designing IP/MPLS over WDM networks that can protect against single optical link/node failures.

## Chapter 2

# 2. Real-Time Provisioning in WDM Mesh Networks

## 2.1 Introduction

Recently, there has been a dramatic increase in data traffic, driven primarily by the explosive growth of the Internet as well as the proliferation of virtual private networks (VPNs). At the same time, the rise of optical networking, first with wavelength-division multiplexing (WDM) transmission technology and more recently with optical multiplexers, and optical cross-connects (OXC) devices, is moving us towards the vision of creating an “all-optical” Internet. In particular, these technologies yield the ability to add, drop, and in effect construct wavelength-routed networks, heralding a new era in which bandwidth is relatively abundant and inexpensive. To some, a key realization of this vision will occur when lightpaths can be provisioned automatically to create bandwidth between end-users, with timescales on the order of minutes or seconds [1,2].

To better understand and appreciate the provisioning issue, we need to look into how circuits are provisioned in a typical network today. Provisioning a cross-country SONET service today requires several steps. First, connectivity from the customer premise to the carrier’s POP must be established for each end of the circuit. Second, a physical path must be mapped out between the many physical hubs in a carrier’s

network between the two points. Each path must be checked for fiber/ring bandwidth availability. Terminating equipment must be ordered and installed on each end of each fiber path, and each interconnect point must have capacity on the optical cross-connect system. All of the cross-connects and physical interconnects must then be made and each segment documented and tested. This process is extremely manual and generally takes several months to accomplish. WDM complicates this process even further because tens and soon hundreds of wavelengths are supported on individual fiber strands. Clearly, an automated optical routing layer will facilitate much faster provisioning.

Before this vision can be realized, however, networks need to slim down. Today's core network architecture model has four layers: IP and other content bearing traffic, over ATM for traffic engineering, over SONET for transport, and over WDM for fiber capacity. This approach has significant functional overlap among its layers and typically suffers from the lowest common denominator effect where any one layer can limit the scalability of the entire network. When first conceived, this layering made sense, but as IP and WDM evolve, a more efficient inter-working is called for, i.e., one that exploits the complimentary features of each domain. In effect, high-performance routers plus a smart optical transport layer equipped with a new breed of Photonic networking components and subsystems together are setting the foundation for the next-generation networking paradigm.

The solution, many believe, is to layer IP directly over the optical substrate [2]. If IP can be mapped directly onto the WDM layer, some of the functional overlap can be eliminated, potentially collapsing today's vertically layered network

architecture into a horizontal model where all network elements work as peers to dynamically establish optical paths. To bring the IP and WDM layers together, however, new capabilities must be added to both layers. A framing standard is needed for carrying packets directly over lambdas. Signaling standards are needed so that IP devices can control optical resources [3]. More importantly, with the conventional multi-layered architecture out of the way, automated provisioning systems will gain direct access to WDM resources, and dynamic lightpath provisioning will become easier and more practical to implement.

This chapter considers the problem of real-time provisioning of optical channels in a hybrid IP-centric WDM-based networking model. Provisioning of connections requires algorithms for route selection, and signaling mechanisms to request and establish connectivity within the network along a chosen route. In particular, the problem of route selection in such wavelength-routed networks is referred to as the “routing and wavelength assignment (RWA) problem [4]. Here, we present a review of RWA schemes and also compare the performance of two different constraint-based routing/RWA algorithms for dynamic provisioning of the optical channels. These algorithms are examples of approaches that might be used to simplify the complex problem of dynamic lightpath computation. Methodologies and associated algorithms for dynamic lightpath computation are outlined. We present an overview of the emerging architectural alternatives of the two-layer model, referred to in the literature as “the interconnection models”, for IP over optical networks, namely the overlay, the peer, and the augmented models [5]. Finally, we examine the implications of

implementing the proposed RWA schemes for the lightpath provisioning aspects for each of the three emerging interconnection models.

The remainder of this chapter is organized as follows. Section 2 presents an overview of the emerging architectural alternatives of the two-layer model. Overview of RWA is presented in Section 3. In Section 4, we present an analytical model for solving the RWA problem; dynamic lightpath provisioning is presented in section 5 and section 6 presents the simulation results. Finally, Section 7 offers summary and conclusion.

## **2.2 IP over Optical network architectures**

In the network model considered here, clients (e. g., IP/MPLS routers) are attached to an optical core network, and connected to their peers over dynamically switched optical paths (lightpaths) spanning potentially multiple OXCs. The interaction between the client and the optical core is over a well-defined signaling and routing interface, referred to as the User-Network Interface (UNI). Meanwhile, the optical core network consists of multiple OXCs interconnected by optical links in a general mesh topology. This network may be multi-vendor, where individual vendor OXCs constitutes sub-networks. Each sub-network itself is assumed to be mesh-connected. The interaction between the sub-networks is over a well-defined signaling and routing interface, referred to as the Network-Network Interface (NNI), (see Figure 2.1). Each OXC is assumed to be capable of switching a data stream from a given input port to a given output port. This switching function is controlled by appropriately configuring a cross-connect table.

### 2.2.1 Interconnection Models:

One approach for transporting IP traffic over WDM networks is to use a multi-layered architecture comprising of IP/MPLS layer over ATM over SONET over WDM. If an appropriate interface is designed to provide access to the optical network, multiple higher layer protocols can request lightpaths to peers connected across the optical network. This architecture has four management layers. Another approach is to use a packet over SONET approach, doing away with the ATM layer, by putting IP/PPP/HDLC into SONET framing. This architecture has three management layers. The fact that both approaches support multiple protocols increases complexity for IP-WDM integration because of various edge-interworkings required to route, map and protect client signals across WDM sub-networks [6].

The two-layer model, which aims at a tighter integration between IP and optical layers, offers a series of advantages over the current multi-layer architecture model. MPLS [7] and its extension G-MPLS [8] have been proposed as the integrating structure between IP and optical layers. Nevertheless, routing in non-optical and optical parts of hybrid IP networks needs to be coordinated. To examine the architectural alternatives for the two-layer model (IP-over-optical network), it is important to distinguish between the data plane and control planes over the user-network interface (UNI). The IP-over-optical network architecture is classified according to the organization of the control plane, i.e., e., whether there is a single integrated or separate independent monolithic routing and signaling protocol

spanning the IP and the optical domains. Several models have been proposed including overlay, augmented, and peer-to-peer models [5].

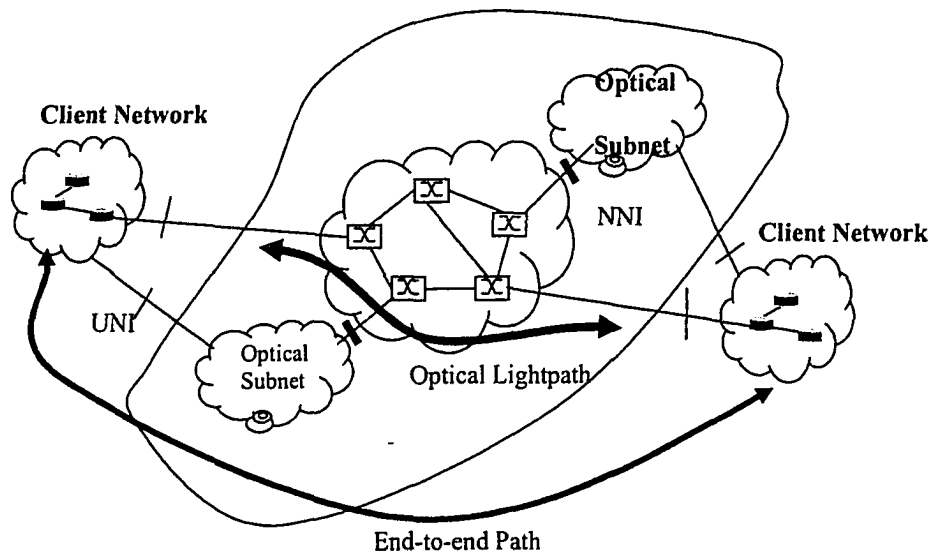


Figure 2.1- Optical Networking Model

### 2.2.2 The Overlay Model:

Under the overlay model, IP domain is more or less independent of the optical domain, that is IP domain acts as a client to the Optical domain. The IP/MPLS routing and signaling protocols are independent of the routing and signaling protocols of the optical layer. Thus, the topology distribution, path computation, and signaling protocols would have to be defined for the optical domain. In this model, the client routers request high-bandwidth connections (lightpaths) from the optical network through the UNI. The client routers are provided with no knowledge of the optical network topology or resources. In this scenario, the optical network provides point-to-point connection to the IP domain. The overlay model may be statically provisioned using a network management system or may be dynamically provisioned.

### **2.2.3 The Peer Model:**

In the peer model, the two layers are collapsed into a single integrated layer managed and traffic engineered in a unified manner. In this regard, the OXCs are treated just like any other router (IP/ MPLS routers and OXCs act as peers) and there is only a single instance of a routing protocol spanning an administrative domain consisting of the core optical network and the surrounding edge devices (IP/MPLS routers, ATM switches). Thus, from a routing and signaling point of view, there is no distinction between the UNI, the NNI (network-network-interface), and any other router-to-router interface. This allows the SP edge devices to have a full access to the topology of the core network. A common IGP like OSPF or IS-IS may be used to exchange topology information. The assumption in this model is that all the optical switches and the routers have a common addressing scheme.

### **2.2.4 The Augmented Model:**

In the augmented model, the IP and optical domains can be functionally separated, each running its own routing protocol, but exchanging full reachability information across the UNI using a standard protocol. For example IP addresses could be assigned to optical network elements and carried by optical routing protocols to allow reachability information to be shared with the IP domain to support some degree of automated discovery. This model combines the best of the peer and overlay interconnection models; it is relatively easy to deploy compared to the peer model in the near term. Also, this is a convenient solution, since it allows implementation of

both provisioning and restoration procedures for optical sub-networks independent of the client network routing. In addition, this approach supports the common scenario where the optical network and client networks are administered by different entities.

The central issue in this model is how the routing information is exchanged at the IP-optical UNI. There are two possibilities for this. The first is to consider the interdomain IP routing protocol, BGP, which may be adopted for exchanging routing information between IP and optical domains. The second is to consider the use OSPF areas (OSPF supports a two-level hierarchical routing scheme through the use of OSPF areas) to exchange routing information across the two domains [9]. On the other hand, running a protocol like BGP across the UNI may be considered too involved, at least for initial implementations of the UNI. A simpler approach would be to limit the reachability information passed through the optical network [9].

### **2.3 Routing and Wavelength Assignment (RWA)**

Provisioning of connections requires algorithms for route selection, and signaling mechanisms to request and establish connectivity within the network along a chosen route. The problem of route selection in such wavelength-routed networks is referred to as the “routing and wavelength assignment (RWA) problem”, and consists of two sub-problems. The first is the routing problem, which determines the path along which the connection can be established. The second problem is to assign a wavelength (or a set of wavelengths) on each link along the selected path (wavelength assignment problem). Real-time provisioning implies that both the path and wavelength should be chosen/assigned dynamically (dynamic RWA), depending on

the network state. In general, all networking models described above, regardless, require route computation and wavelength assignment to provision a lightpath, i.e. e., and dynamic RWA engine.

### 2.3.1 Overview of RWA:

Given a set of connections, the problem of setting up lightpaths by routing and assigning a wavelength to each connection is called the routing and wavelength assignment (RWA) problem [4]. Typically, connection requests may be of three types: static, incremental and dynamic [10]. With static traffic, the entire set of connections is known in advance, and the problem is then to set up lightpaths in a global fashion while minimizing network resources such as the number of wavelengths or the number of fibers in the network. Here, the RWA problem for static traffic is known as Static Lightpath Establishment (SLE) and can be formulated as mixed-integer linear program [11]. In the incremental-traffic case, connection requests arrive sequentially, a lightpath is established for each connection, and the lightpath remains in the network indefinitely.

For the case of dynamic traffic, a lightpath is setup for each connection request as it arrives, and the lightpath is released after some finite amount of time. The objective in the incremental and dynamic traffic cases is to setup lightpaths and assign wavelengths in a manner that minimizes the amount of connection blocking [10]. This problem is referred to as the Dynamic Lightpath Establishment (DLE). Generally, the DLE is more difficult to solve, and therefore, heuristics methods are

generally employed. Heuristics exist for both the routing sub-problem and the wavelength assignment sub-problem.

For the routing sub-problem, there are three basic approaches that can be found in the literature: fixed routing, fixed-alternate routing, and adaptive routing [10]. Fixed routing is one variant of the “static routing” in which routing decisions do not vary with time. Moreover, in fixed routing the same fixed route for a given source-destination pair is always selected. The fixed alternate routing approach considers multiple routes between a source-destination pair and each node in the network maintains an ordered list of a number of fixed routes to each destination node. When a connection request arrives, the source node attempts to establish the connection on each of the routes from the list in sequence, until a route with a valid wavelength assignment is found. Conversely, in adaptive routing [10, 12, 13, 14], the route from a source node to a destination node is chosen dynamically, depending on the network state. Adaptive routing requires extensive support from the control and management protocols to continuously update the routing table at the node. An advantage of adaptive routing is that it results in lower connection blocking than fixed and fixed-alternate routing.

Meanwhile, for the wavelength assignment sub-problem, a number of heuristics have been proposed [15, 16, 17]. These heuristics are Random Wavelength Assignment, First-Fit, Least-Used, Most-Used, Min-Product, Least-Loaded, MAX-SUM, Relative Capacity Loss, Wavelength Reservation, and Protecting Threshold. In [13], the authors propose an adaptive unconstrained routing (AUR), which incorporates network state information into route computation and channel allocation.

Currently, The algorithms that offer the best performance are Relative Capacity Loss (RCL) [17], and Distributed relative Capacity Loss (DRCL) [10]. RCL calculates the Relative Capacity Loss for each path on each available wavelength and then chooses the wavelength that minimizes the sum of the relative capacity loss on all the paths. DRCL is proposed in [10] and is based on RCL but it is more efficient in a distributed environment. For a tutorial review on the RWA problem, we refer the reader to [10].

Optical networks can also pose added wavelength continuity constraints [4], and these may require the use of wavelength conversion (also referred to as wavelength translation or wavelength changing). A wavelength converter is a device, which takes at its input a data channel modulated onto an optical carrier with a wavelength  $\lambda_{in}$ , and produces at its output the same data channel modulated onto an optical carrier with a different wavelength  $\lambda_{out}$ . If wavelength converters are included in the OXCs in WDM networks, connections can be established without the need to find an unoccupied wavelength, which is the same on all the links making the route. This means that networks with wavelength converters are equivalent to traditional circuit switched networks. Wavelength converters thus result in improvements in network performance. On the other hand, it has been shown that a careful wavelength assignment in wavelength-continuous network can lead to improved performance; thus, reducing the benefits of wavelength converters [18, 19]. In [18], the authors investigate the benefits of limited wavelength conversion for ring and mesh-torus topologies with fixed shortest path routing. The authors of [20] used a hypercube network to study limited conversion with fixed shortest path routing and a first-fit wavelength selection algorithm. It is shown that limited wavelength conversion (25

%) achieves the same performance improvement as full wavelength conversion [18,20].

Additionally, many other constraints can also serve to complicate the RWA process, especially in all-optical networks. Specifically, besides wavelength continuity requirements, these include analog attenuation effects and power limitations. For example, adequate signal-to-noise ratios (SNRs), cross-talk, and dispersion effects caused by subsystem components and fiber links can be computed along candidate paths. This information can be incorporated into route resolution strategies by defining new cost functions [21]. In [21], the authors have extended the routing and wavelength assignment problem to account for the power degradation of a routed signal due to non-ideal behavior of optical components such as multiplexers, de-multiplexers, taps, and fiber links.

### **2.3.2 Dynamic Lightpath Computation:**

Dynamic computation of a lightpath involves the implementation of two traffic engineering components: an information distribution mechanism that provide knowledge of the relevant attributes of available network resources, and a path selection process that uses the information distributed by the dynamic link-state advertisement algorithm to select a path that meets the specific requirements of the traffic flow. In a fully distributed IP-over-optical network implementation, these are:

- 1) **An information distribution mechanism-** that provides knowledge of network's topology and the available resources. This component is implemented by defining relatively simple extensions to the interior gateway protocol (IGP), e. g. open

shortest path first (OSPF) so that link attributes are included as part of each router's link-state advertisement. Some of the traffic-engineering extensions that need to be added to the IGP link-state advertisement include maximum link bandwidth, maximum reservable link bandwidth, current bandwidth reservation, current bandwidth usage, and link coloring. These extensions capture optical link parameters and any constraints specific to optical networks. Such topology and link state information is then flooded to all nodes via updates. Another important component is to define naming and addressing convention for different elements of the physical plant hierarchy. This algorithm is capable of periodically updating and advertising all of the above link attributes. The link state updates can be triggered, for instance, based on a given threshold of the number of available wavelengths per fiber, below which the updates can be triggered. Once each node has a representation of the full physical network topology and the available resources on each link, a path selection algorithm is required, i.e., dynamic RWA.

- 2) **A path selection process-** that uses the information distributed by the dynamic link-state advertisement algorithm to select an explicit route that meets the specific requirements of the traffic flow. This process can be performed either off-line or on-line using a constraint-based routing calculation. The source router (peer model), or the border OXC/central management node (augmented/overlay models) are basically responsible for computing the complete path all the way to the destination through the optical domain, and then initiating path setup using the signaling protocol (e. g., CR-LDP or RSVP). The route may be specified either as

a series of nodes (routers/ OXCs), or in terms of the specific links used (as long as IP addresses are associated with these links).

## 2.4 Analytical Model

Analytical models for analyzing the performance of all-optical wavelength-routed networks have been proposed in [22, 23, 16, 13, 24]. In [22], the author uses a generalized reduced load approximation to compute the end-to-end blocking probabilities for fixed routing and least loaded routing. Random wavelength assignment scheme was used throughout the analysis due to its simplicity. The model is shown to give good results only for fairly small networks and its complexity grows exponentially with the hop length. In [24], a new analytical technique, based on the inclusion-exclusion principle from combinatorics, was proposed for the analysis of all-optical networks with no wavelength conversion and random wavelength assignment. The authors propose two models of low complexity. The first model improves the model proposed in [25] in that the complexity of calculation is independent of the hop length and scales only with the capacity of the link. The second model shows that it is accurate for sparse networks. The authors also extend their models to analyze fixed routing and least loaded routing both with random wavelength assignment scheme.

Analytical models for the First Fit wavelength assignment have been proposed in [23, 16, 13]. Here, the layered-graph approach proposed in [25] is used to simplify the lightpath establishment. In this approach, each layer corresponds to a single wavelength and the number of layers corresponds to the number of wavelengths. This can alleviate the difficulties incurred by the wavelength continuity constraint by

simultaneously considering the routing and wavelength assignment on each layer. Traffic, which is blocked on one layer, overflows into the second layer and versions of overflow traffic model are used. In [23], the authors assume the arrival process on each link to be a BPP (Binomial-Poisson-Pascal) distribution and they also model the overflow traffic to follow a BPP distribution. The authors of [13] developed a new analytical model to compute the blocking probabilities for fixed routing and fixed alternate routing with first fit wavelength assignment by adopting the layered-graph approach in their analysis assuming the arrival and the overflow processes to follow Poisson distribution.

#### 2.4.1 Network and traffic model assumptions:

In this section we state the assumptions about the network and the traffic used in our models for calculating the path blocking probability of an all-optical network with no wavelength converters. A network with  $N$  nodes and  $J$  links can be represented by a directed graph  $G(N, J)$ . Each link has the same number of wavelengths ( $W$ ). Calls arrive according to a Poisson process and the duration of each call is exponentially distributed with unit mean.

Let  $\lambda$  be the total traffic load offered to the network,  $a_r$  be the offered load to the  $r^{th}$  s-d connection pair,  $R = \{r_1, r_2, \dots, r_{N(N-1)}\}$  be the set of all routes in the network,  $E = \{1, 2, \dots, J\}$  be the set of all links, and  $W = \{\lambda_1, \lambda_2, \dots, \lambda_W\}$  is the set of all wavelengths.

In our approach the problem of computing path blocking on a particular wavelength is reduced to computing approximate blocking probability in a circuit switched

network [26, 27]. We extend the layered-graph approach [25], to develop an approximate analytical model for the path blocking. By using this layered approach, the RWA problem is reduced to only finding an end-to-end path on a wavelength between the s-d pair. The search for a path and a wavelength may be viewed as a search over a sequence of  $W$  logical paths, where a logical path is a combination of a physical path and a particular wavelength. If the logical path under consideration is available then the connection is established, otherwise the connection request overflows to the next logical path. Here, we use the moment matching techniques, specifically the Equivalent Random Method (ERT) [28] to model the overflow traffic between different wavelength layers.

Let:

$\alpha_r^{\lambda_k}$  : The arrival rate on route  $r_i$  on wavelength layer  $\lambda_k$ .

$B_j^{\lambda_k}$  : The link blocking probability on link  $j$  on wavelength  $\lambda_k$ .

$\alpha_j^{\lambda_k}$  : The traffic load on link  $j$  on wavelength  $\lambda_k$ .

$L_r^{\lambda_k}$  : The path blocking probability on route  $r_i$  on wavelength layer  $\lambda_k$ .

If we consider a network with  $F$  fibers per link, the number of idle circuits on link  $j$  on wavelength  $\lambda_k$  can be viewed as a birth-and-death process given that the traffic arrival follows a Poisson distribution.

$$P_m = \text{Prob} \{m \text{ circuits are used}\} = \frac{(\alpha_j^{\lambda_k})^m}{m!} \left[ 1 + \sum_{n=1}^F \frac{(\alpha_j^{\lambda_k})^n}{n!} \right]^{-1} \quad (1)$$

$$B_j^{\lambda_k} = \text{Prob} \{F \text{ circuits are used on link } j \text{ on wavelength } \lambda_k\} = P_F$$

If the number of fibers per link is one, then:

$$B_j^{\lambda_k} = \frac{\alpha_j^{\lambda_k}}{1 + \alpha_j^{\lambda_k}} \quad (2)$$

On the other hand, if the traffic offered to a link does not follow a Poisson distribution, it is defined by its peakedness  $Z = \frac{m}{v} \neq 1$  where  $m$  is the mean and  $v$  is the variance of the offered traffic. Thus, if  $(A, N, Z)$  characterizes the system, where  $A$  is the offered traffic and  $N$  is the number of trunks per link; using the ERT, we transform this system into the following system  $(\frac{A}{Z}, \frac{N}{Z}, 1)$  and thus the link blocking is computed using the Erlang B system [28].

#### 2.4.2 Fixed routing:

##### A- Fixed Routing with First fit Wavelength Assignment

The traffic rate on link  $j$  on wavelength  $\lambda_k$  can be approximated [26, 27] by:

$$\alpha_j^{\lambda_k} = \sum_{r: j \in r} a_r^{\lambda_k} \cdot \prod_{m \in r - \{j\}} (1 - B_m^{\lambda_k}) \quad \forall r \in R \quad (3)$$

Here, the idea of approximation is simple to explain [26]: suppose a Poisson stream of rate  $a_r^{\lambda_k}$  is thinned by a factor  $(1 - B_m^{\lambda_k})$  at each link  $m \in r - \{j\}$  on wavelength graph  $\lambda_k$  before being offered to link  $j$ . If this thinning could be assumed independent from link to link and over all routes passing through link  $j$ , then the traffic offered to link  $j$  would be Poisson at rate  $\alpha_j^{\lambda_k}$ .

The path blocking on route  $r_i$  on a wavelength  $\lambda_k$  can be computed by using the link distribution:

$$L_{r_i}^{\lambda_k} = 1 - \prod_{j \in r_i} (1 - B_j^{\lambda_k}) \quad (4)$$

Note that  $L_{r_i}^{\lambda_k}$  is the blocking probability on route  $r_i$  on wavelength layer  $\lambda_k$  when the offered traffic to route  $r_i$  is  $a_{r_i}^{\lambda_k}$ . Here, a fraction of  $a_{r_i}^{\lambda_k}$  overflows to wavelength  $\lambda_{k+1}$  with probability  $L_{r_i}^{\lambda_k}$ . Hence,

$$a_{r_i}^{\lambda_{k+1}} = a_{r_i}^{\lambda_k} \cdot L_{r_i}^{\lambda_k} \quad (5)$$

The total blocking on route  $r_i$  is then approximated by:

$$L_{r_i} = \prod_{k=1}^W L_{r_i}^{\lambda_k} \quad (6)$$

Next we show the numerical solution to this problem. Due to the non-linear relation between equations (1) and (3), an iterative procedure is used to solve this system.

<b>1. Initialization:</b>	
• $\tilde{L}_{r_i}^{\lambda_k} = 0$	for $i = 1 \dots N(N-1)$ , $k = 1 \dots W$
• $a_{r_i}^{\lambda_k} = 0$	for $i = 1 \dots N(N-1)$ , $k = 2 \dots W$
• $a_{r_i}^{\lambda_1} = \frac{\lambda}{N(N-1)}$	for $i = 1 \dots N(N-1)$
• $\alpha_j^{\lambda_k} = 0$	for $j = 1 \dots J$ , $k = 2 \dots W$
• $\alpha_j^{\lambda_1} = \sum_{r_j \in r} a_r^{\lambda_1}$	$\forall r \in R$
<b>2. Repeat for all wavelengths:</b>	
2.1- determine $B_j^{\lambda_1}$	for $j = 1 \dots J$ from (1)
2.2- determine $\alpha_j^{\lambda_1}$	for $j = 1 \dots J$ from (3)
2.3- determine $L_{r_i}^{\lambda_1}$	for $i = 1 \dots N(N-1)$ from (4). If $\max_i  L_{r_i}^{\lambda_1} - \tilde{L}_{r_i}^{\lambda_1}  < \epsilon$ then compute $a_{r_i}^{\lambda_{k+1}}$ for $i = 1 \dots N(N-1)$ from (4), compute $\alpha_j^{\lambda_{k+1}} = \sum_{r_j \in r} a_r^{\lambda_{k+1}}$ $\forall r \in R$ , for $j = 1 \dots J$ , and go back to step 2. If not converge, then $\tilde{L}_{r_i}^{\lambda_k} = L_{r_i}^{\lambda_k}$ for $i = 1 \dots N(N-1)$ and go back to 2.1.

Table 2.1- Calculation of  $L_R$  for fixed routing

### B- Fixed Routing with Random Wavelength Assignment

In this section, we extend our model to analyze all-optical wavelength-routed networks for fixed routing and Random Wavelength Assignment.

In Random wavelength assignment, each wavelength is equally likely to be selected by any incoming call request; thus, we define the probability of selecting a wavelength from a set of  $W$  wavelengths to be:

$$P(\lambda_i \text{ is selected}) = \frac{1}{|W|}, \forall \lambda_i \in W$$

If  $\lambda$  is the total traffic in the network, then the traffic per s-d pair on a given wavelength graph  $\lambda_k$  is initially given by:

$$a_{r_i}^{\lambda_k} = \frac{\lambda}{N(N-1)} \times P(\lambda_k \text{ is selected}) = \frac{\lambda}{N(N-1)} \times \frac{1}{|W|}.$$

Note that, a blocked call on a wavelength layer  $\lambda_k$  overflows to any wavelength

$$\lambda_{k'} \in W - \{\lambda_k\} \text{ with equal probability, } p = \frac{1}{|W|-1}, \text{ if } |W| > 1$$

One can estimate now the traffic rate on any route  $r_i$  on any wavelength layer  $\lambda_k$  using the following recursive equation:

$$a_{r_i}^{\lambda_k, (l)} = a_{r_i}^{\lambda_k, (l-1)} \times (1 - L_{r_i}^{\lambda_k, (l-1)}) + \sum_{m \in W - \{\lambda_k\}} a_{r_i}^{\lambda_m, (l-1)} \times p \times L_{r_i}^{\lambda_m, (l-1)} \quad (7)$$

The first half of the sum in equation (7) accounts for the remaining traffic on route  $r_i$ , while the second half equates the overflow traffic from the remaining wavelengths in the set. Note that here, traffic on a route  $r_i$  on a wavelength  $\lambda_{k'} \neq \lambda_k$  overflows with probability  $p$  to route  $r_i$  on wavelength  $\lambda_k$ .

A similar numerical procedure to the ones shown in Fig.1 is followed; here, the route traffic is computed each time using equation (7).

### 2.4.3 Fixed alternate routing:

#### *A- Fixed Alternate Routing with first fit wavelength assignment*

In Alternate Routing, each source-destination pair is assigned a set of paths. The set of paths may be searched in a fixed or adaptive order. Alternate routing improves the performance of the network by allowing a call blocked on a route to try an alternate route with sufficient resources.

Let  $L$  be the number of edge-disjoint paths between an s-d pair. On each path, the wavelength set is searched in a fixed order to find a suitable wavelength to establish the connection. In alternate routing, the search for a path and a wavelength may be viewed as a search over a sequence of  $LW$  logical paths, where a logical path is the combination of a particular physical route and a particular wavelength. If the logical path under consideration is available, then the connection is established; otherwise the connection overflows to the next logical path.

Based on the same underlying approximation (link blocking independence [28]), we can extend equation (3) to compute the link traffic:

$$\alpha_j^{\lambda_i} = \sum_{l=1}^L \sum_{r^{(l)}: j \in r^{(l)}} a_{r^{(l)}}^{\lambda_i} \prod_{k \in r^{(l)} - \{j\}} (1 - B_k^{\lambda_i}) \quad (8)$$

Where,  $a_{r^{(l)}}^{\lambda_i}$  is the traffic intensity on route  $r^{(l)}$  on wavelength  $\lambda_i$

The traffic blocked on the first route overflows to the alternate route, and the traffic intensity on the alternate route on wavelength  $\lambda_i$  is approximated by:

$$a_{r_i}^{\lambda_i} = a_{r_i}^{\lambda_w} \times L_{r_i}^{\lambda_w} \quad , i = 2, \dots, L \quad (9)$$

$$a_{r_i}^{\lambda_i} = a_{r_i}^{\lambda_{i-1}} \times L_{r_i}^{\lambda_{i-1}} \quad , i = 2, \dots, W \quad (10)$$

$L_{r_i}^{\lambda_i}$  is the path blocking on the  $i^{\text{th}}$  alternate route on wavelength  $\lambda_i$ ;  $L_{r_i}^{\lambda_i}$  is approximated as in equation (4). The overall blocking probability for an s-d pair is:

$$L_{r_i} = \prod_{l=1}^L \prod_{k=1}^W L_{r_i}^{\lambda_k} \quad \forall r_i \in R \quad (11)$$

### ***B- Fixed Alternate Routing with Trunk Reservation***

Under this scheme, blocking of alternately routed calls is determined by the number of free wavelengths in the alternate path, which is defined as the minimum number of free wavelengths on all the links along the path [29]. In particular, a number “r” is chosen such that if the alternate path has fewer than “r” free wavelengths, all the alternately routed traffic will be banned on that path [29]. In other words, the last “r” free wavelengths are always reserved for the directly routed traffic. The underlying principle of this scheme is to block alternately routed traffic more often on the busy links or paths and to let it be carried by the alternate paths that are not so busy. This actually helps in equalizing the traffic load over the network so that fewer calls will be lost due to local traffic congestion. Furthermore, this scheme helps to protect the directly routed traffic from being blocked due to excessive loading of the alternately routed traffic, because the last “r” free wavelengths are always reserved only for the directly routed traffic.

To find the end-to-end blocking probability, we use an approach similar to that of the alternate routing case. In this approach, a link with W wavelengths is seen as W

virtual links where each one corresponds to a wavelength along the link. Thus, we need to find the load offered to each individual virtual link given a trunk threshold “ $\Gamma$ ”.

If we consider the case where only one alternate route is allowed ( $L=2$ ), then the offered load could be approximated as follows:

$$\alpha_j^{\lambda_i} = \alpha_j^{\lambda_i,d} + \alpha_j^{\lambda_i,a} \quad \text{for } 0 \leq i \leq W - r \quad (12)$$

$$\alpha_j^{\lambda_i} = \alpha_j^{\lambda_i,d} \quad \text{for } W - r + 1 \leq i \leq W$$

Where  $\alpha_j^{\lambda_i,a}$  and  $\alpha_j^{\lambda_i,d}$  are the offered load to link  $j$  on wavelength  $\lambda_i$  from both direct and alternate traffic and they are given by:

$$\alpha_j^{\lambda_i,a} = \sum_{r^{(1)}:j \in r^{(1)}} a_{r^{(1)}}^{\lambda_i} \prod_{k \in r^{(1)} - \{j\}} (1 - B_k^{\lambda_i}) \quad (13)$$

$$\alpha_j^{\lambda_i,d} = \sum_{r^{(2)}:j \in r^{(2)}} a_{r^{(2)}}^{\lambda_i} \prod_{k \in r^{(2)} - \{j\}} (1 - B_k^{\lambda_i}) \quad (14)$$

The link blocking probability and the end-to-end blocking probability are given by equations (2) and (4) respectively.

## 2.5 Dynamic Provisioning

In this section we present some combined RWA algorithms for dynamic provisioning of optical lightpaths. Specifically, these algorithms integrate and collapse both the routing and wavelength assignment sub-problems into a single dynamic constraint-based routing problem. Thus, the emphasis here is on the adaptive routing problem, rather than focusing on the wavelength-assignment problem. It has been shown that the routing scheme has much more of an impact on the overall

network performance than the wavelength-assignment scheme. Moreover, both algorithms are also shown to be amenable to fully distributed implementations.

The network is viewed as a multi-layered graphs (Fig. 2.2) each corresponding to a specific wavelength. For a connection request and on a given wavelength, Dijkstra shortest path algorithm that is suitably modified for WDM networks is used for computing a constrained path. This is achieved by associating each link in the network with a specific weight function that incorporates WDM specific information such as the number of available wavelengths and the total wavelengths. This means that the algorithm might compute (on-line)  $W$  Paths, each corresponding to one of the  $W$  wavelengths. Then, one of these paths is selected according to a global selection criterion. Thus, the problem of wavelength-assignment is totally mitigated and both the routing and wavelength assignment sub-problems are now integrated and collapsed into a single dynamic constraint-based routing problem. This in contrast to the work reported in [14], where a single path is first calculated; then a wavelength is assigned to the path by propagating a wavelength request to all the OXCs along the path. Thus, such algorithm avoids the overhead associated with such a wavelength request (probe message).

The algorithm is implemented as per the following [1]:

1. First consider a multi-fiber IP-centric WDM-based network whose physical topology consists of multiple OXC interconnected via point-to-point WDM links in an arbitrary mesh topology.

2. Assume that none of the OXCs has wavelength conversion capability. Hence, to meet a connection request, a lightpath, that uses the same wavelength on all the links along the entire route from source-to-destination, has to be set up.
3. Both algorithms are based on a fully distributed implementation in which all nodes maintain a synchronized and identical topology and link state information (traffic-engineering database, TED).
4. Assuming that  $W$  is the number of wavelengths per fiber, the network is represented by  $W$  identical graphs, each conforming to the physical topology and a particular wavelength. In view of this multi-graph model, each physical link is now represented by  $W$  virtual links (channels), each corresponding to one of the wavelength graphs. Figure 2.2 illustrates the concept of the multi-graph approach for a simple network with four nodes, four physical links, and  $W = 2$ .
5. For a given connection request, a constrained route is calculated on each of the wavelength graph, throughout the entire network from source to destination, typically using a shortest path algorithm but with the link weights adjusted to attain some sort of local resource optimization. Clearly, there are at most  $W$  paths that can be calculated, each corresponding to a given wavelength, provided that each path can meet the given routing constraint. As a result, we get the vector  $V = \langle \text{Path}_i, \text{Wavelength}_i \rangle, i = 1, \dots, W$ , where the number of entries stored in  $V$  may vary from no entries at all (request is blocked), to a possible maximum of  $W$  entries. Finally, to globally optimize the network resources, provided that the number of entries stored in  $V$  is more than one, an entry  $(\text{Path}_i, \text{Wavelength}_i)$  out of all the other possible entries, has to be selected. Thus, by virtually separating

wavelengths, both the routing and wavelength assignment sub-problems are now reduced into a single dynamic constraint-based routing problem (see Fig. 2.2).

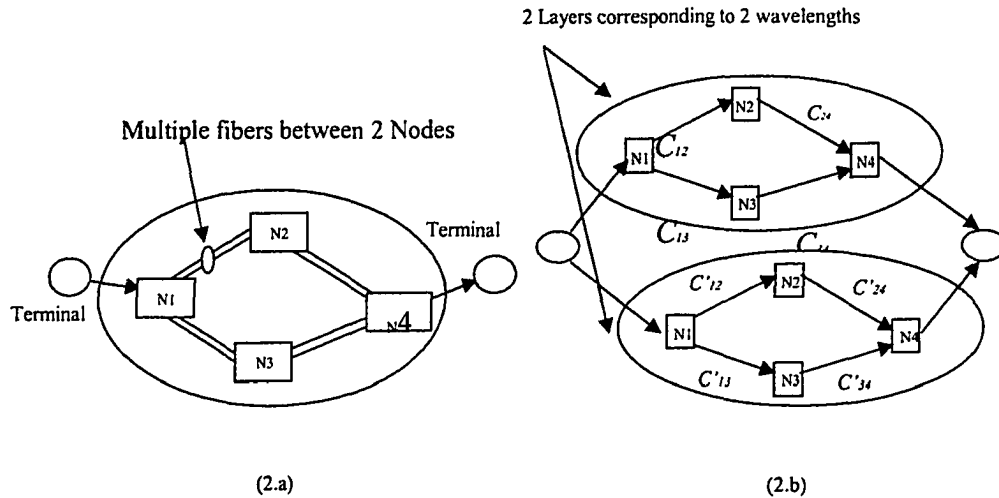


Fig. 2.2. The sample network model (2.a), and its multi-layered graph approach (2.b).

### 2.5.1 Algorithm I “Full Adaptive Routing”:

The implementation of this algorithm is as follows:

1. For a given wavelength graph  $\lambda_i$ , each virtual link, in each of the  $W$  wavelength graphs, is assigned a cost. Basically, the cost of a link at a given wavelength  $\lambda_i$ , is defined here as the inverse of the number of available channels over that particular link. Hence, initially the cost of a given link throughout the entire network is set =  $(1/F)$ , where  $F$  is the number of fibers (per link) connecting two adjacent OXCs. In general, the cost of link  $L_j$  at wavelength  $\lambda_i$ ,  $C(L_j^{\lambda_i})$ , is given by:

$$C(L_j^{\lambda_i}) = \begin{cases} \frac{1}{F - N(L_j^{\lambda_i})} & \text{if } N(L_j^{\lambda_i}) < |F| \\ \infty & \text{if } N(L_j^{\lambda_i}) = |F| \end{cases} \quad (15)$$

Where  $N(L_j^{\lambda_i})$  is the number of occupied (unavailable)  $\lambda_i$ 's on link  $L_j$ .

2. For a given wavelength  $\lambda_i$ , we associate each path throughout the entire network with a total cost,  $C_{sd}^{\lambda_i}$ , which is defined here as the summation of the costs of all individual links spanning the entire path from source to destination.

$$C_{sd}^{\lambda_i} = \sum_{j=1}^n C(L_j^{\lambda_i}) \quad (16)$$

Where  $\langle L_1, L_2, \dots, L_n \rangle$  is the set of  $n$  links that comprise the path.

3. For a given connection request, run the Dijkstra algorithm on the first wavelength graph  $\lambda_1$  to find the shortest path (the path with minimum  $C_{sd}^{\lambda_1}$ ). Store the calculated path along with its corresponding wavelength  $\lambda_1$  as the first entry of the vector  $V$ . Note that the calculated local path, for a given wavelength graph  $\lambda_i$ , is not necessarily the path with the minimum number of hops.
4. Repeat step 3 for each of the remaining  $W-1$  wavelength graphs. Note that the vector  $V$  might have now up to  $W$  entries.
5. Examine the contents of the vector  $V$  and perform one of the following instructions:
  - a) If the vector  $V$  has no entries at all, reject the connection request; otherwise go to step # b.
  - b) If the vector  $V$  has only one entry, select this entry as the combination (Path <sub>$i$</sub> , Wavelength <sub>$i$</sub> ) that satisfies the connection request. After assigning the path, update the weights associated with all links along the entire path (just on the corresponding wavelength graph  $\lambda_i$ ) by basically decrementing the number of

the available  $\lambda_i$ 's (channels), on every link along the selected path, by one; otherwise go step # c.

- c) If the vector V has more than one entry, select one of the those entries combination that satisfy one of the following global path selection schemes:

### I. Total Cost-Based selection

In this scheme, a total cost,  $C_{sd}^{\lambda_i}$ , is associated with each computed path within the vector V (<Path<sub>i</sub>, Wavelength<sub>i</sub>>), given by:

$$C_{sd}^{\lambda_i} = \sum_{j=1}^n C(L_j^{\lambda_i}) \quad (17)$$

The path with the minimum total cost  $C_{sd}^{\lambda_i}$  is selected and assigned to the connection. Note that this selection criterion skews the conventional shortest path search to favor less utilized network resources. This is illustrated in Figure 2.3a where the path with fewer hops is not selected.

### II. Balanced Cost-Based selection

In this case, each path is assigned a balanced cost,  $C_{sd}^{B\lambda_i}$ , defined by:

$$C_{sd}^{B\lambda_i} = n \times \sum_{j=1}^n C_{L_j}^{\lambda_i} \quad (18)$$

Where n is the total number of links along the path.

The path with the least balanced cost  $C_{sd}^{B\lambda_i}$  is selected and assigned to the connection. Note that this selection criterion strikes a balance between the minimum cost and the minimum number of hops. The main objective of this selection criterion

is to avoid assigning long paths to a connection and route connections over the healthy part of the network. This is illustrated in Figure 2.3b where the connection is routed here on the path with wavelength  $\lambda_j$  that has a balanced cost ( $2*1$ ) less than that with wavelength  $\lambda_i$  [ $3*(11/12)$ ], thus avoiding assigning the connection to the longer path ( $\lambda_i$ ). Note, however, that the total cost of the path with wavelength  $\lambda_j$  (1) is higher than that with  $\lambda_i$  (11/12).

### III. Future Cost-Based selection

This scheme uses the same total cost  $C_{sd}^{\lambda_i}$  of Eq. (17), but with the individual link cost  $C_{L_j}^{\lambda_i}$  of Eq. (15) redefined as:

$$C(L_j^{\lambda_i, Future}) = \frac{1}{F-1-N(L_j^{\lambda_i})} \quad \text{if } N(L_j^{\lambda_i}) < |F| \quad (19)$$

Thus, the total future cost of this scheme is given by:

$$C_{sd}^{\lambda_i, Future} = \sum_{j=1}^n C_{L_j}^{\lambda_i, Future} \quad (20)$$

The path with the minimum future cost is selected and assigned to the connection. If all paths have the same future cost of infinity (i.e., along a path at least one link has one last channel available), select the path with the least cost using the original definition of individual link cost  $C(L_j^{\lambda_i})$  of Eq. (15). Note that this selection criterion strongly favors assigning paths with the least current/future-utilized resources. This is illustrated in Figure 2.3c where the path with infinity future cost is not selected.

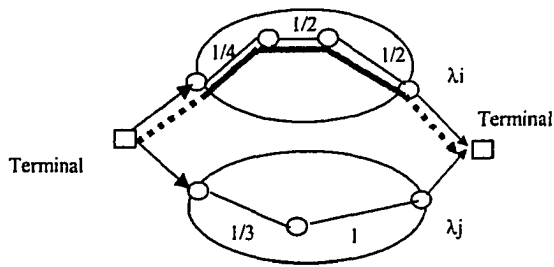


Fig. 2.3.a (Total Cost Path Selection,  $\lambda_i$ )  
 Fig. 2.3.a (Total Cost Path Selection,  $\lambda_i$ )

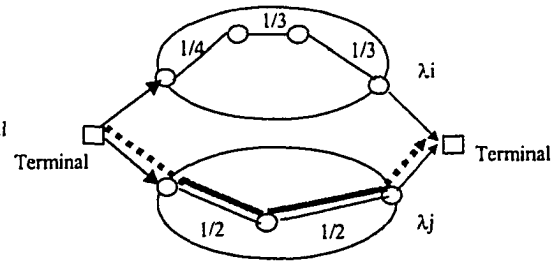


Fig. 2.3.b (Balanced Cost Path Selection,  $\lambda_j$ )

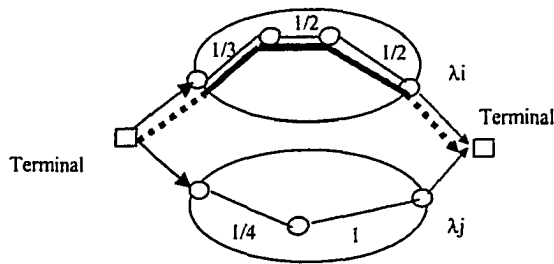


Fig. 2.3.c (Future Cost Path Selection,  $\lambda_i$ )

Figure 2.3. Illustration Examples

Two comments are in order here. First, for any of the three-selection schemes described above, once a path is assigned, the weights associated with all links along the selected path should be updated as indicated in step 5b above. Second, all link state updates must be propagated and advertised to all other core nodes throughout the network. Note that the frequency of the link state update per unit time is proportional to the number of accepted/released calls. In other words, the link state update is triggered once for every accepted/released call. As a result, the signaling overhead associated with a high volume calls may be excessive.

**2.5.2 Algorithm II “Semi-Adaptive Routing”:**

This algorithm adopts the same implementation procedures developed for the full adaptive algorithm described above, except for the following fundamental differences:

1. A shortest path algorithm (Dijkstra algorithm) is initially run off-line to calculate the shortest path (just minimum number of hops) between every source-destination node pair (routing tables) throughout the entire network. These off-line computed routing tables are stored at each node in every wavelength graph. Thus, the initial routing tables are identical for all  $W$  wavelength graphs.
2. For an initial connection request, the ingress node at every wavelength graph consults its own routing table for the shortest path. As a result, similar to algorithm I, we may get as much as  $W$  paths, where one of them can then be selected according to the selection schemes described above.
3. For all the consecutive connection requests, the routing tables remain unchanged; so that step 2 is repeated until the cost of a link  $L_j$  in a given wavelength graph  $\lambda_i$  goes to infinity (no more available  $\lambda_i$ 's). In this case, link  $L_j$  is removed from wavelength graph  $\lambda_i$  and the routing tables are calculated for each node again. Note that the Dijkstra algorithm is run this time on-line to find the shortest path (the path with minimum  $C_{sd}^{\lambda_i}$ ).

Note that the signaling overhead associated with the link state updates for this algorithm is considerably less than that of algorithm I (link state updates is only triggered when the cost of the link goes to infinity). In addition, the time associated

with computing a path for the semi-adaptive algorithm is less than that of the full adaptive one, since the path is directly read of the routing table.

Finally, in the case where the lightpath is wavelength continuous as this work has assumed, optical non-linearities, chromatic dispersion, amplifier spontaneous emission and other factors together limit the scalability of an all-optical network. Routing in such networks will then have to take into account noise accumulation and dispersion to ensure that lightpaths are established with adequate signal qualities.

## 2.6 Performance evaluation

The accuracy of the proposed analytical model is presented in this section and the performance of the proposed dynamic RWA algorithms is also evaluated via simulation of the mesh-based NSF network [4]. The NSF network consists of 14 nodes and 21 physical links. Each adjacent node pair is connected through a bi-directional physical link that consists of  $F$  fibers, where each fiber is assumed to have the same number of wavelengths ( $W$ ). We use a dynamic traffic model in which call requests arrive at each node according to a Poisson process with a network arrival rate  $\lambda$ . An arrival session is equally likely to be destined to any node in the network. The session holding time is assumed to be exponentially distributed with mean  $1/\mu$ . The blocking probability is the metric used to evaluate the network performance. In each simulation run, a large number of requests are generated one after the other, and the results are averaged over many simulation runs.

### 2.6.1 Accuracy of the analytical model:

We conduct a numerical comparison of the analytical technique against simulation results to study the accuracy of the proposed model. The number of

wavelengths per fiber is set to be 4 and 8. Figure 2.4 shows the comparison between the analytical model, presented in section 3, and simulation results for the NSF network. The first fit wavelength assignment is used. The figure shows the accuracy of the model when the number of wavelengths is 4. On the other hand, when the number of wavelengths becomes 8, the analytical model underestimates the blocking probability. This is due to the fact that the more the number of wavelengths is, the more the traffic overflows between the wavelength layers; and hence the less accurate the overflow traffic is approximated. Nevertheless, the moment matching technique gives a closer estimation of the overflow traffic than the Poisson assumption.

On the other hand, Figure 2.5 shows the same comparison using a different architecture, the ring network. The analytical model here overestimates the blocking probability (the computed approximate blocking probability is less than the blocking probability that was obtained through the simulation); this is due to the fact that in ring architecture, routes usually tend to have more hops; hence, the assumption of link-independency of our model is not as accurate as in mesh networks.

Different wavelength assignment schemes were compared analytically and through simulation and results are shown in figure 2.6. Here we see that our model is also accurate for the random wavelength assignment not only for small networks as opposed to [22] but also for more complex networks like NSF. From the simulation, it is obvious that the FF slightly outperforms the Random selection (RS) as opposed to the analytical model where at lower load the RS slightly outperforms the FF.

Figure 2.7 shows the results of the fixed alternate routing (FAR) versus the fixed routing (FR). As shown, fixed alternate routing improves the network

performance at light loads due to the flexibility in the routing algorithm. However, as the load increases, alternate routing degrades the blocking performance and the improvement of FAR is very limited; this is due to the fact that alternate routing will occupy more transmission resources and thus, at higher loads will prevent direct traffic from being routed and leads to instability. By limiting the flow of the alternately routed traffic, one can overcome the instability problem. Trunk reservation (TR) is one technique to improve the performance of alternate routing at higher loads. Figure 2.8 shows the analytical comparison between the FR, the FAR and the FAR with TR. The figure shows that at lower loads, Fig.8.a, FAR routing outperforms the other routing techniques and gives a lower bound on the blocking. Also shown on the same figure the improvement on blocking performance by using TR. By decreasing the trunk threshold one can improve the network performance by allowing more resources on the alternate routes to be used by blocked connections. However, Fig.8.b shows that alternate routing will degrade the performance at higher loads (upper bound on blocking), whereas FAR with TR can better manage the network resources by selecting the right trunk threshold ( $r = 5$  in this case).

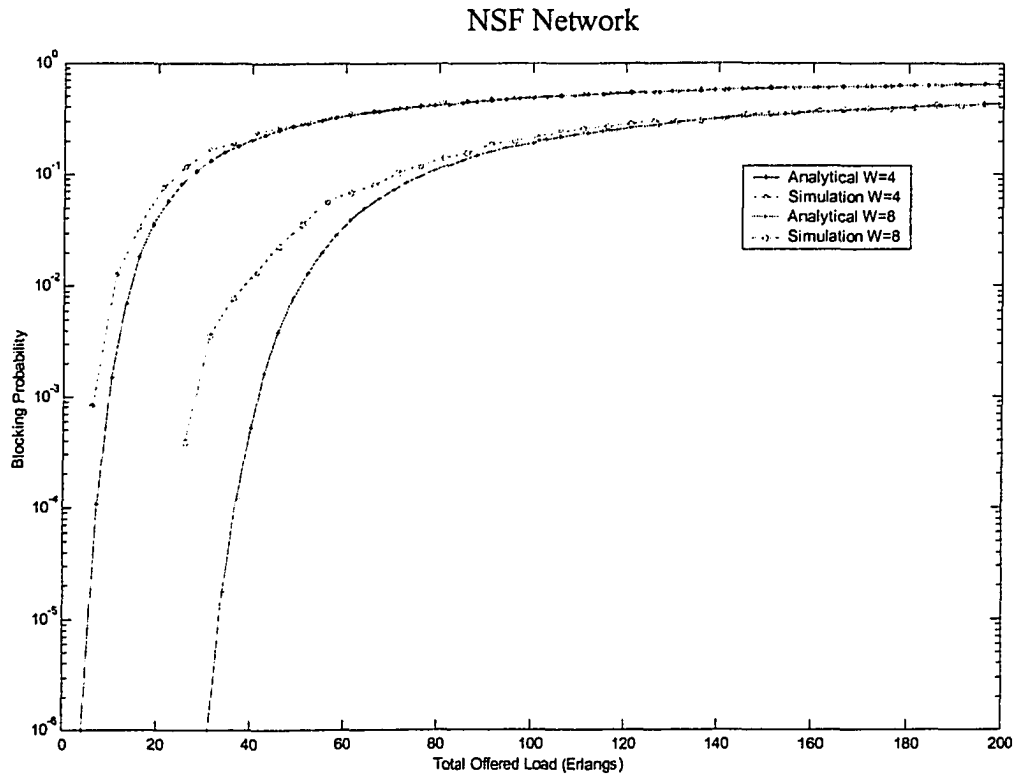


Figure 2.4- FR with FF wavelength assignment

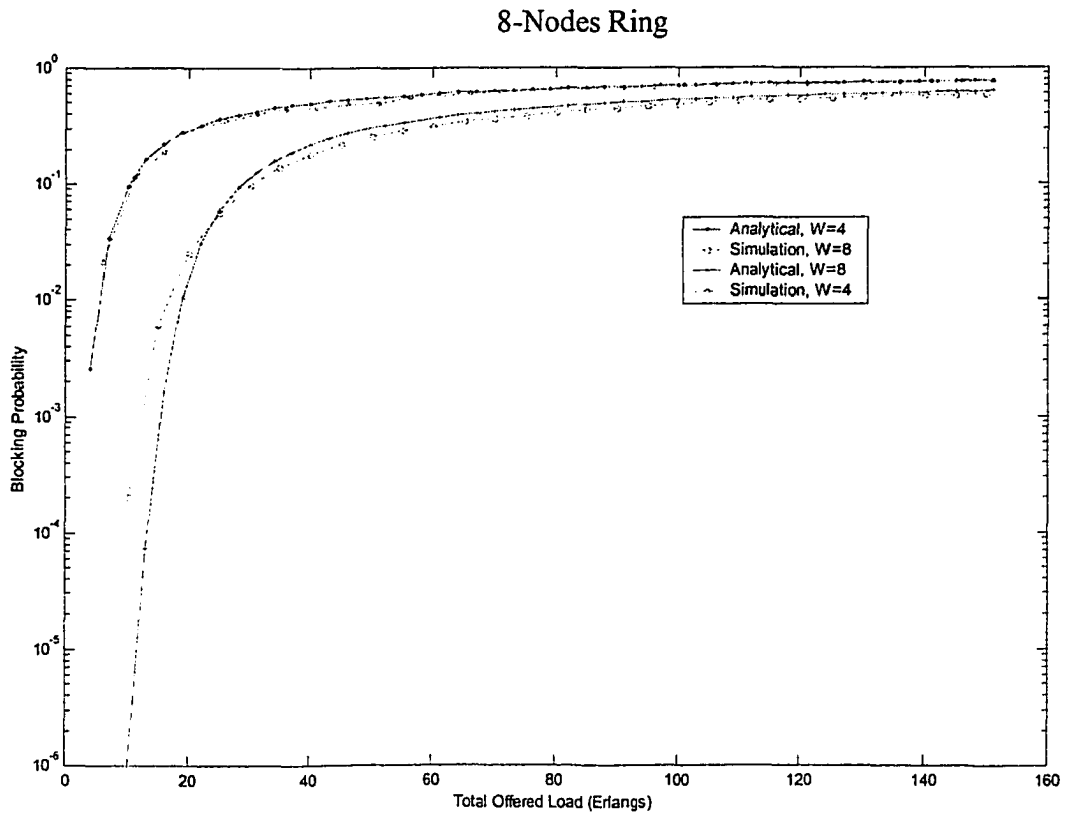


Figure 2.5- FR with FF wavelength assignment (Ring Network)

NSF Network, W = 4 wavelengths

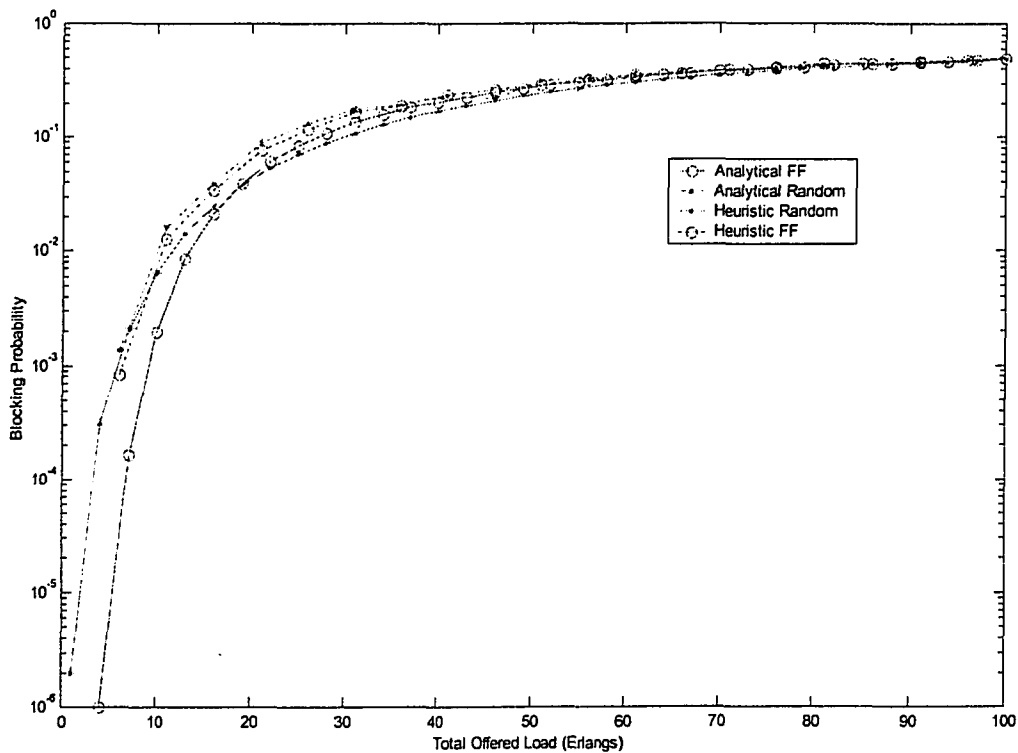


Figure 2.6- FR with different wavelength assignment schemes

NSF Network

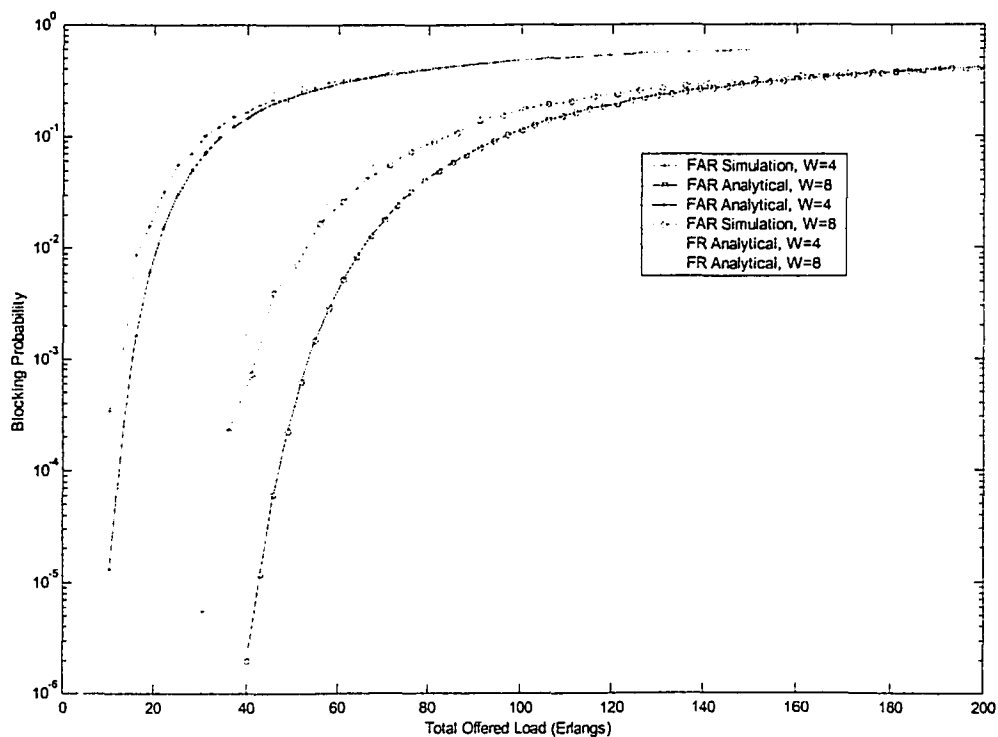
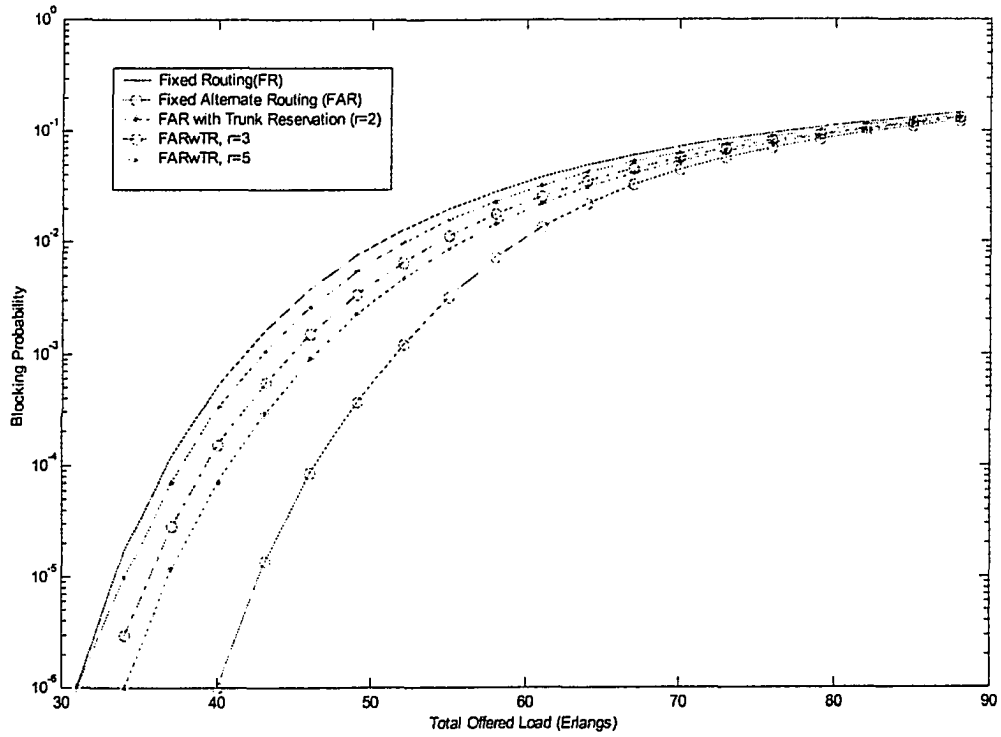


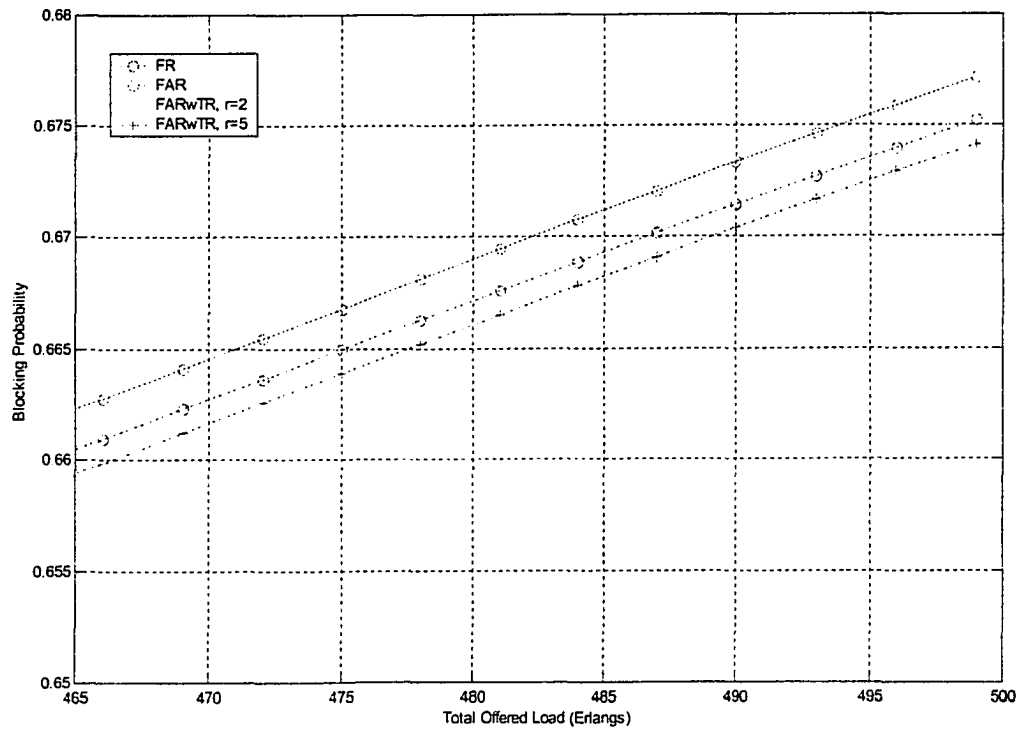
Figure 2.7- FAR and FR with FF wavelength assignment

NSF Network, W=8 wavelengths



(a)

NSF Network, W=8 wavelengths



(b)

Figure 2.8- FAR with Trunk Reservation (FARwTR)

### 2.6.2 Performance of the proposed dynamic algorithms:

The blocking probability is the metric used to evaluate the performance of the proposed dynamic routing algorithms. In each simulation run, a large number of requests are generated one after the other, and the results are averaged over many simulation runs. If at any time, a connection request cannot be satisfied according to the algorithms developed above, the connection request is dropped.

Fig. 2.9 shows the simulated blocking probability vs. the calls arrival rate for both algorithms when the total cost-based path selection criterion (scheme I) is used. The simulated blocking probability is also shown in the figure for the conventional static RWA scheme used in most algorithms. As expected, it can be seen from the Figure, the performance of both dynamic algorithms is significantly better than that of the static one. Note, however, that the performance of the full-adaptive algorithm is slightly better than that of the semi-adaptive. These results always hold, independent of which path selection scheme described above (I, or II, or III) is used.

Fig. 2.10 shows the simulated blocking probability vs. the calls arrival rate of the full- adaptive algorithm for all the three path selection schemes described above. As can be seen from the Fig., the path selection process based on future cost performs the best, followed by the path selection process based on the balanced cost, then by the path selection process based on total cost. Also included in the figure is the adaptive First Fit (FF) algorithm [10]. Note that the performance of the future cost algorithm is significantly better than that of the FF. However, the performance of the balanced cost is almost the same as that of the FF (same results were also obtained in [10] with the DRCL algorithm), and both of them outperform the total cost scheme.

NSF Net, 2 Fibers/Link

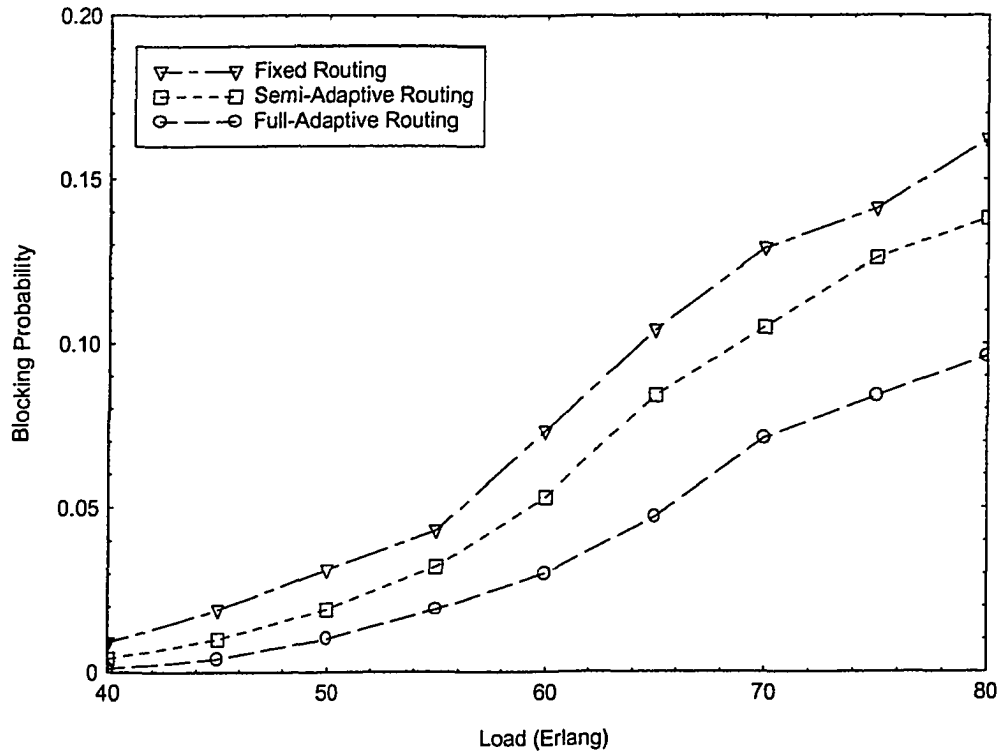


Figure 2.9 Blocking probability vs. Load, for different routing schemes

NSF Net, 2 Fibers/Link

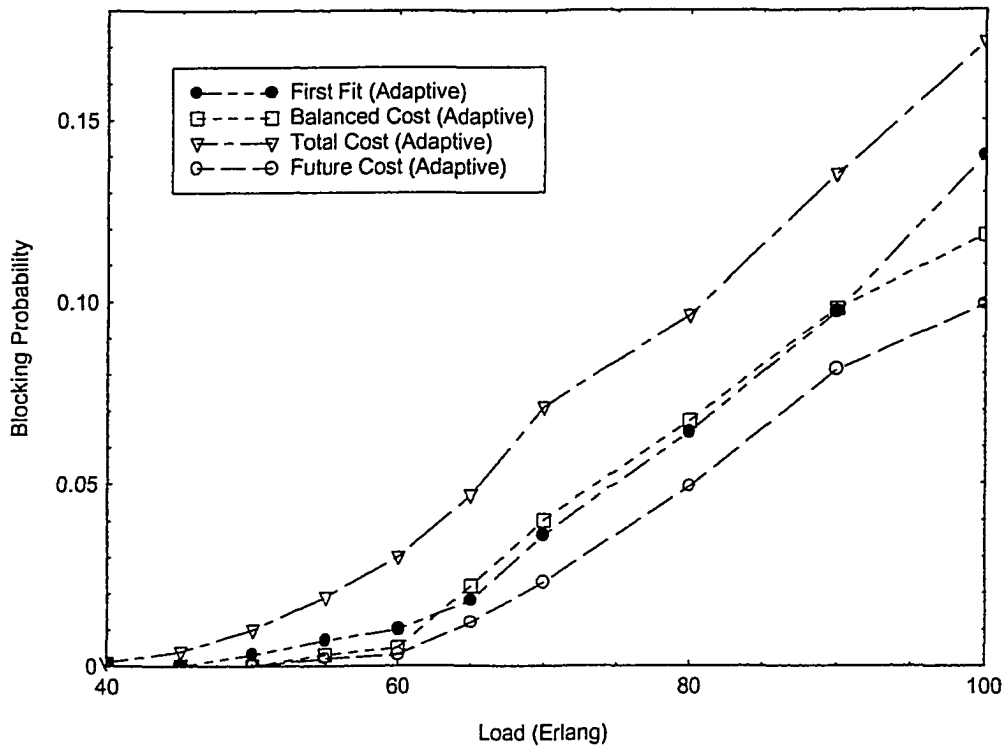


Figure 2.10 Comparison of the different selection schemes for the adaptive routing case.

## 2.7 Conclusion

Real-time provisioning of optical channels in a hybrid IP-centric WDM-based networking model has been presented in this chapter. Specifically, the work presented here has addressed the implementation issues of the path selection component of the traffic-engineering problem in such a network. Methodologies and associated algorithms for dynamic lightpath computation were outlined.

Here, provisioning of connections in a WDM based network is defined as the process of dynamically selecting efficient end-to-end routes and assigning wavelengths, i.e., routing and wavelength assignment (RWA) and involves the implementation of two traffic engineering components: an information distribution mechanism that provides knowledge of the relevant attributes of available network resources, and a path selection process that uses the information distributed by the dynamic link-state advertisement algorithm to select a path that meets the specific requirements.

In this study, we presented the dynamic RWA problem as a key requirement for lightpath provisioning. We analytically modeled the conventional RWA problem with different wavelength assignment schemes. Fixed alternate routing (FAR) with/without trunk reservation (TR) is one approach to improve the network performance; analytical models for alternate routing w/o trunk reservation are also presented. We validated the accuracy of the proposed models through numerical comparison with simulation results of the RWA. We also showed the efficiency of alternate routing in improving the performance. However, as the traffic load increases, alternate routing degrades the blocking performance and the improvement

of FAR is very limited; this is due to the fact that alternate routing will occupy more transmission resources and thus, at higher loads will prevent direct traffic from being routed and leads to instability. By limiting the flow of the alternately routed traffic, one can overcome the instability problem. Trunk reservation (TR) is one technique to improve the performance of alternate routing at higher loads and can better manage the network resources by selecting the right trunk threshold.

We have also presented and compared the performance of two different constraint-based routing and wavelength assignment (RWA) algorithms, for dynamic provisioning of the optical channels and three path selection schemes have also been proposed for each algorithm. Both routing algorithms are based on a fully distributed implementation. The performance of both algorithms is then compared with that of the conventional static RWA algorithm. It is shown that the dynamic full-adaptive algorithm outperforms the semi-adaptive one, and both of them significantly outperform the conventional static algorithm. It is also shown that the Future Cost-Based Selection scheme outperforms both the total-based and the balanced selection schemes.

## 2.8 References

- [1] C. Assi, A. Shami, R. Kurtz, and M Ali "Optical Networking and Real-Time Provisioning; An Integrated Vision For The Next-Generation Internet", IEEE Network, July/August, 2001.
- [2] N. Ghani et al., "On IP-over-WDM Integration", IEEE Communications magazine, March 2000.

- [3] N. Ghani, "Lambda-Labeling: A framework for IP-over-WDM using MPLS", *Optical Networks Magazine*, Vol. 1, PP45-58, April 2000.
- [4] B. Mukhejee, *Optical Communications Networks*, McGraw-Hill, NY, 1997.
- [5] B. Rajagopalan et al, "IP over Optical Networks: A Framework", Work in Progress, Internet Draft, draft-many- optical-framework-02.txt
- [6] N. Chandhok et al, "IP over Optical Networks: A Summary of Issues", Work in Progress, Internet Draft, draft-osu-ipo-mpls-issues-02.txt.
- [7] R. Callon, "A Framework for Multi-protocol Label Switching" Work in Progress, Internet Draft, Sept 1999.
- [8] P. Ashwood-Smith et al, "Generalized MPLS, Signaling Functional Description" Work in Progress, Internet Draft, draft-ietf-mpls-generalized-signaling-01.txt.
- [9] D. Pendarakis, B. Rajagopalan, and Debanjan Saha, " Routing information exchange in Optical networks", Work in Progress , Internet Draft, draft-prs-optical-routing-01.txt.
- [10] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," *Optical Networks Magazine*, vol. 1, pp. 47-60, Jan. 2000.
- [11] R. Ramaswami and K. N. Sivarajan, "Dynamic Allocation in All-Optical Ring Networks," *Proc., IEEE ICC '97*, Montreal, Quebec, Canada, vol. 1, pp. 432-436, June 1997.
- [12] S. Ramamurthy and B. Mukherjee "Fixed-Alternate Routing and Wavelength Assignment in Wavelength Routed Optical Networks", *proc. IEEE Globecom'98*, vol. 4, pp. 2295-2302, Nov. 1998.
- [13] A.Mokhtar and M. Azizoglu "Adaptive wavelength routing in all optical networks", *IEEE/ACM trans. Net*, Vol. 6, pp. 197-206, April 1998.

- [14] Tibor Fabry-Asztlos, Nilesh Bhide and Krishna M. Sivalingam "Adaptive Weight Functions for Shortest Path Routing Algorithms for Multi-Wavelength Optical WDM Networks", Proc. IEEE ICC2000, New Orleans, LA, June 2000.
- [15] I. Chlamtac, A. Ganz, and G. Karmi, "Purely Optical Networks for Terabit Communication," Proc., IEEE INFOCOM '89, Washington, DC, vol. 3, pp. 887-896, April 1989.
- [16] E. Karasan and E. Ayanoglu, "Effects of wavelength Routing and Selection Algorithms on Wavelength Conversion Gain in WDM Optical Networks," IEEE/ACM Transactions on Networking, vol. 6, no. 2, pp. 186-196, April 1998.
- [17] X. Zhang and C. Qiao, "Wavelength Assignment for Dynamic Traffic in Multi-Fiber WDM Networks", proc. 7<sup>th</sup> ICCCN, pp. 479- 485, Oct. 1998
- [18] J. Yates, P. Rumsewicz, and J. Lacey, "Wavelength converters in dynamically reconfigurable WDM networks" IEEE communication Survey, Q2, 1999.
- [19] C. Assi, A. Shami, Z. Zhang, M. Ali, "Impact of wavelength converters on the performance of optical networks" Optical Network, vol.3, issue 2, March/April 2002.
- [20] H. Harai, M. Murata, and H. Miyahara, "Performance of All-Optical Networks with limited range wavelength conversion" Proc., IEEE ICC, Montreal Canada, 1997, pp. 416-421, April 1997.
- [21] M. Ali et al, "Routing and Wavelength Assignment (RWA) with Power Considerations In All-Optical Wavelength-Routed Networks", Proc. IEEE GLOBECOM '99, Rio De Janeiro, Brazil), pp. 1443-1437, Dec. 1999.
- [22] A. Birman. "Computing Approximate Blocking Probabilities For A Class Of All-Optical Networks," IEEE Journal of Selected Areas of Communication, volume 14, pages 852-857, June 1996.
- [23] H. Harai et al. "Performance Of Alternate Routing Methods in All-Optical Switching Networks," Proceedings, IEEE INFOCOM '97, pages 16-524.

- [24] A. Sridharan and K. Sivarajan. "Blocking in All-Optical Networks," IEEE INFOCOM 2000, pp. 990-999.
- [25] C. Chen and S. Banerjee, "A new model for optimal routing and wavelength assignment in wavelength division multiplexed optical networks" IEEE INFOCOM 96, pp. 148-155. 1996
- [26] F. P. Kelly, "Loss Networks," The Annals Of Applied Probability, Volume 1, Issue 3, pages 319-378, August 1991.
- [27] F. P. Kelly, "Blocking Probabilities In Large Circuit-Switched Networks," Advances in Applied Probability, volume 18, pages 473-505, 1986.
- [28] A. Girad. "Routing and Dimensioning in Circuit-Switched Networks," Addison-Wesley, Reading, Massachusetts, 1990.
- [29] T. K. G. Yum, M. Schwartz, "Comparison of Routing Procedures for Circuit Switched Traffic in Non-Hierarchical Networks" IEEE Trans. on Comm., Vol 35, May 1987, pp. 534-544.

## Chapter 3

### 3. Survivability in WDM-based Networks

#### 3.1 Introduction

Recently, Wavelength Division Multiplexed (WDM) mesh networks have received much attention and emerged as a very promising transport solution for next-generation data networks. Given the scale of these architectures, where each fiber can potentially carry terabits of data, service survivability [1,2,3] becomes a critical requirement for network planning and management.

WDM transport networks support different higher-layer services, such as SONET connections, asynchronous transfer mode (ATM) virtual circuits, and IP-switched datagram traffic. According to the layered structure of a network, survivability can be offered at the WDM layer or higher layers. Some of the higher-layer services, such as SONET and ATM, actually have their own protection mechanisms, while some may not have recovery mechanisms incorporated in the protocols. Under this situation, the WDM layer should be able to offer them. However, WDM layer survivability cannot protect against failures at higher layers, and some survivability must be provided at higher client layers as well.

Incorporating survivability mechanisms at multiple layers leads to the issues of assigning functions to each layer and coordinating the layers in effecting recovery

from a fault. The set of rules describing the point of origination of the fault recovery process and the interaction between the various layers is called an *escalation* or *interworking* strategy. Two escalation strategies have been proposed [4] based on the layer at which the fault recovery process is initiated. In the *bottom-up* strategy, recovery starts at the layer closest to the failure, and escalates upward upon expiration of a holdoff timer. This timer allows the lower layers time to recover from a fault (if possible) before triggering recovery mechanisms at a higher layer. This strategy ensures very quick activation of the recovery process. In the *top-down* strategy, recovery always starts at the uppermost layer and escalates downward. Holdoff timers are not necessary in this strategy, but a disadvantage is the potentially large number of traffic streams that must be restored at the higher layers. A third strategy would start the recovery process at some intermediate layer, and escalate either upward or downward based on the received alarms and survivability statistics. A cost-performance comparison of the escalation strategies, reported in [4] for an ATM-over-synchronous digital hierarchy (SDH) network, found that the bottom-up approach was better in terms of both equipment cost and recovery time. However, a main attraction of the top down strategy is that it can provide differentiated QoS for survivability to different users.

On the other hand, providing survivability functionality at the optical layer has many advantages [5, 6]:

- Speed: Recovery at the WDM layer is much faster because the nodes can act quickly upon the occurrence of failures and do not have to wait for higher-layer indication signals.

- **Simplicity:** It needs less coordination than recovery at higher layers.
- **Effectiveness:** Optical restoration makes more efficient use of restoration capacity because of resource sharing among different service layers.
- **Transparency:** The wavelength routing protection technique is independent of the protocols used in higher layers.

Thus, in this chapter we only focus on survivability techniques provided at the optical layer. Service survivability at the optical layer stipulates that upon the failure of any network element, all affected connections be rerouted within a short time interval using spare capacity reserved on alternate paths. Alternatively, if spare capacity is not reserved a-priori, the network must dynamically search/compute alternate routes and reserve resources to restore affected connections. However, since this latter approach exhibits longer restoration times and cannot guarantee complete service recovery, it may not be very feasible. Hence, the work herein only considers the case where spare capacity is already reserved during connection setup, referred to hereafter as path protection. The rest of this chapter is organized as follows. A background on survivability in WDM and non-WDM networks is presented in section 3.2. In section 3.3 we present routing algorithms in survivable optical networks; different distributed path selection algorithms are also presented and their respective complexities and scalabilities compared. Subsequently, previous analytical work on WDM network performance evaluation is reviewed in section 3.4 and a new analytical model for shared path protection proposed is derived. Section 3.5 compares the accuracy of the derived analytical models against detailed simulation

results and also compares the different path selection schemes. Final conclusions are presented in section 3.6.

### **3.2 Background on Survivability**

In order to design a survivable optical network, one must lay out the possible failures under which the network must be survivable. The basic types of network failures generally considered are link and node failure. Link failure usually occurs because of cable cuts; node failure is due to equipment failure at network nodes. Besides node and link failures, which are common failure situations in any communication network, channel failure is also possible in WDM optical networks. A channel failure is usually caused by the failure of transmitting and/or receiving equipment operating on that channel (wavelength). Initial work on survivability in WDM optical networks has focused mostly on the recovery from a single link or node failure. In this section, we discuss several survivability techniques used in optical networks and we present a classification of survivability techniques commonly used.

#### **3.2.1 Survivability in Non-WDM Networks:**

Two general categories for survivability in non-WDM optical networks can be classified: pre-designed protection and dynamic restoration:

##### **3.2.1.1 Pre-designed Protection**

Here recovery from network failures is based on preplanned schemes and it usually relies on resources (fibers, wavelengths, switches, etc.) dedicated to protection purposes. Resources are reserved for recovery from failures at either connection setup or network design time, and kept idle when there is no failure. From

this point of view, the use of capacity is not very efficient, but on the other hand, the level and speed of recovery from a failure can be guaranteed. The most common pre-designed protection schemes currently used in non-WDM optical networks are Automatic protection switching (APS) and self-healing ring (SHR).

Typically APS is used to handle *link* failures and it has three main architectures: 1 + 1, 1:1, and 1:N. The difference between the three architectures is the assignment of protection resources. In 1 + 1 APS (Fig. 3.1a), a protection link is provided for every working link. The source node transmits the information signal on both the working and protection links. The receiver at the destination node compares the two signals and chooses the better one (e.g., the less noisy one). In case of a link failure, the destination node is still able to receive the signal on the operational link. In 1:1 APS, shown in Fig. 3.1b, every working link has a protection link (under normal condition the protection link could be either idle or used to carry low-priority traffic), but the source and destination nodes switch to the protection link only when a failure on the working link is detected. Figure 3.1c shows how a 1:N APS system works. In this scheme,  $N$  working links share a single protection link, thereby providing protection against the failure of any one of the  $N$  working links. But, unlike in 1:1 APS, the traffic switched to the protection link must be switched back to the working link after it is repaired so that the protection link is available for any future working link failures. In general,  $m:n$  protection refers to an APS scheme in which  $m$  protection links are shared among  $n$  working links. When more than  $m$  of the  $n$  working links fail simultaneously, the traffic routed to the protection links can be decided according to pre-assigned priorities.

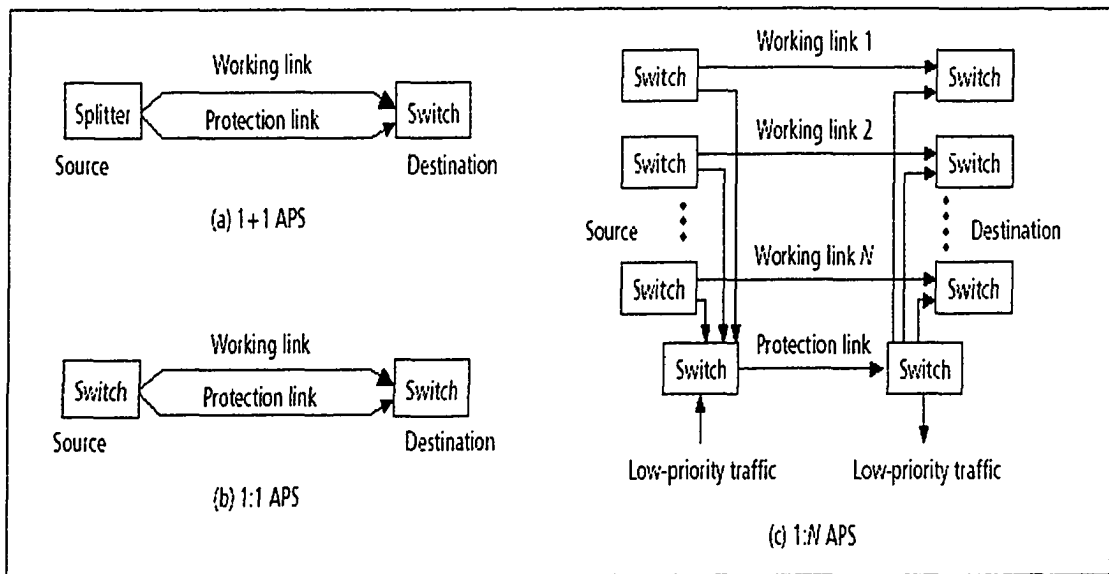


Figure 3.1- Automatic protection switching

On the other hand, SONET SHR is a very successful technique for survivable optical networks in which networks are designed to have ring architectures. SHR is more flexible than APS in that it can handle both link and node failures. The developments of high-speed add/drop multiplexing (ADM) technology and the simple control mechanism used in SHR have made it a very attractive way to provide survivability. Unidirectional SHR (USHR) and bi-directional SHR (BSHR) are two types of SHRs in SONET systems. The difference between these two categories is the direction of the traffic flow under normal operation. In USHR, Fig.2a, the normal traffic flow goes around the ring in one direction. Any traffic routed to the protection ring because of a failure is carried in the opposite direction. In BSHR, working traffic flows in both directions.

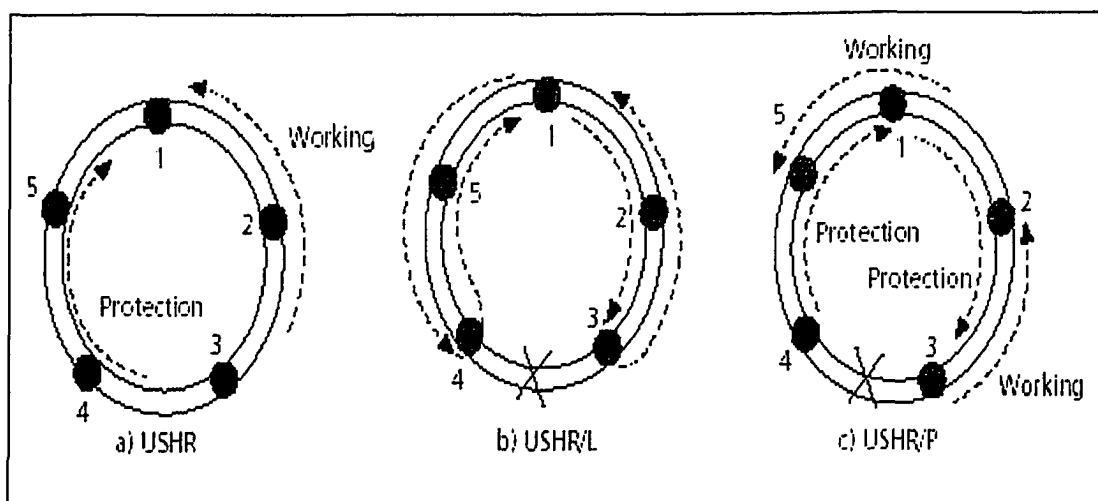


Figure 3.2- The self-healing ring.

USHR protection can be done in two different ways: line protection switched USHR (USHR/L) and path protection switched USHR (USHR/P). In USHR/L, also known as *loopback*, the two nodes adjacent to a failure are responsible for switching the affected traffic to the protection ring (Fig.3.2b). On the other hand, USHR/P is basically a 1 + 1 protection scheme because the signals for every connection are transmitted on both rings. When a failure occurs and affects one of the signals, the ADM at each node decides which signal is still good and then chooses it. USHR/P is the fastest SHR scheme because no switching of signals is needed.

The typical architectures of BSHR are two- and four-fiber line protection BSHR/2 and BSHR/4. In BSHR/2, half of the capacity on each ring is reserved for protection. When a failure occurs, the two nodes adjacent to the failed site will loop the affected traffic using the reserved capacity on both rings. In BSHR/4, two fibers are dedicated as working fibers, so the capacity on them can be completely utilized. Upon failure, the nodes adjacent to the failure site simply loop the affected traffic from the working to the protection fibers.

### 3.2.1.2- Dynamic Restoration:

Dynamic restoration implies the discovery of spare capacity dynamically in the network to restore the affected services; that is, the resources used for recovery are not reserved at the time of connection establishment, but are chosen from available resources when the failure occurs. Typically this is more efficient than pre-designed protection from the viewpoint of resource utilization. However, the restoration time is usually longer and 100 percent service recovery cannot be guaranteed because sufficient spare capacity may not be available at the time of failure. In SONET networks, intelligent digital cross-connect systems (DCSs) and controllers are used as the main components to realize dynamic restoration. The controllers may be centralized or distributed to achieve a suitable balance between recovery time and complexity of recovery algorithms. But no matter which ones are used, recovery time and complexity are much more than for pre-designed protection schemes.

Due to the large traffic carried in fiber optic systems, recovery time is a very important parameter to meet quality of service (QoS) requirements. Dynamic restoration techniques are at a disadvantage from this perspective. Research has shown that both pre-designed protection and dynamic restoration schemes have advantages depending on network topology. For example, in a point-to-point link, APS is the best solution, while SHR offers the most benefits in a ring topology [7]. Most of today's communication networks use these preplanned protection mechanisms rather than dynamic restoration methods. When a large mesh network is to be made survivable, APS and SHR covering may require an inordinate amount of additional capacity. In this situation, dynamic restoration seems to be a practical way

to do the job, but much research remains to be done in designing rapid restoration schemes.

### 3.2.2 Survivability in WDM optical mesh networks:

As WDM system deployment advances, mesh topologies using optical cross-connects (OXC) are likely to emerge. Research has recently focused on studying WDM mesh network survivability and most of the studies so far have considered only single-link failures. In WDM networks, the failure of even a single link can cause the failure of several channels simultaneously, a much more serious situation than in non-WDM systems. Furthermore, fiber cuts are among the most common failure scenarios.

Considering the single-link failure scenario in mesh networks, a simple way to provide survivability is the dedicated fiber scheme. Here, each link in the network has its dedicated backup link. Upon a link failure, traffic is simply routed over its backup link, as shown in the example of Fig. 3.3a in which the link between nodes 5 and 6 has failed. Because most link failures are due to fiber cable cuts, the backup fiber is required to be diversely routed. This is a complicated task for network design and realization, and is obviously a waste of capacity. Other solutions that can use the capacity more efficiently have been proposed and, as mentioned earlier, fall under two categories: *pre-designed protection* and *dynamic restoration*. Dynamic restoration schemes are carried out using intelligent OXC and controllers in the network. As mentioned before, dynamic restoration algorithms are usually complicated, and restoration time is much longer than for pre-designed protection

algorithms. Pre-designed protection schemes are by far the most studied for WDM optical networks. Because of the multi-channel traffic, the design algorithms used in a WDM network are more complicated than those used in non- WDM systems. In the following discussion, we will mostly discuss the proposed pre-designed protection approaches. There are two main pre-designed protection techniques against single-link failures in WDM networks. One is link-based protection, the other, path-based protection. We present these in detail below.

### 3.2.2.1- Link-Based Protection:

The basic idea of link-based protection is that a protection path is reserved for each link, and when the link fails, traffic is rerouted (looped back) around the failed link. As an example, in Fig. 3.3b, after a link failure between nodes 5 and 6, the affected traffic is rerouted through the backup path 5–2–6. Here, the end nodes of the failed link (i.e., nodes 5 and 6) are responsible for recovery. In a WDM network, each link carries many channels, and the failure of a single link causes the failure of all the channels on the link. In link-based protection, each working channel has a protection wavelength path. The protection paths used for different working wavelengths on the same link may use different paths and/or different wavelengths. For example, Fig. 3.3b shows two different protection paths (5–2–6 and 5–1–2–3–6) for the same link 5–6. Link-based protection schemes can be further classified as dedicated or shared link protection.

In *Dedicated link protection*, a protection wavelength path is dedicated to a working channel on a particular link. Therefore, if the protection paths for (some wavelengths on) two different links overlap, different wavelengths must be assigned

to the protection path on the overlapping portion even if the working wavelengths on the two links are the same. As an example, consider Fig. 3.3c. Let  $\lambda_1$  on path 5–2–6 (labeled protection path 1) be the protection wavelength path for a working channel on link 5–6, and the protection path for a working channel on link 1–2 be 1–5–2 (labeled protection path 2). Then a different wavelength, say  $\lambda_2$ , must be assigned to protection path 2, even if the working wavelengths on links 5–6 and 1–2 are the same, say  $\lambda_1$ . Note that this requires wavelength conversion if link 1–2 fails. The above example indicates the difficulty in designing efficient protection schemes in large networks. Efficient design is especially difficult if wavelength conversion facilities are unavailable. On the other hand, dedicated link protection may offer protection against the failure of multiple links. For example, in Fig. 3.3c both working channels can be recovered if both links 1–2 and 5–6 fail simultaneously. However, note that recovery of working channel 5–6 is not possible if both links 5–2 and 5–6 fail at once.

On the other hand, *Shared link protection* allows different protection paths to share a wavelength on the overlapping portion if the corresponding working channels are on different links. Shared link protection utilizes capacity more efficiently than dedicated link protection, and can provide 100 percent recovery from single link failures. Figure 3.3d shows an example of shared link protection. Protection paths 1 and 2 (used to protect a working channel on links 5–6 and 1–2, respectively) can share wavelength 1 on link 5–2. Note, however, that a different wavelength must be used to protect a different working channel on link 5–6 if protection path 1 is used for that working channel.

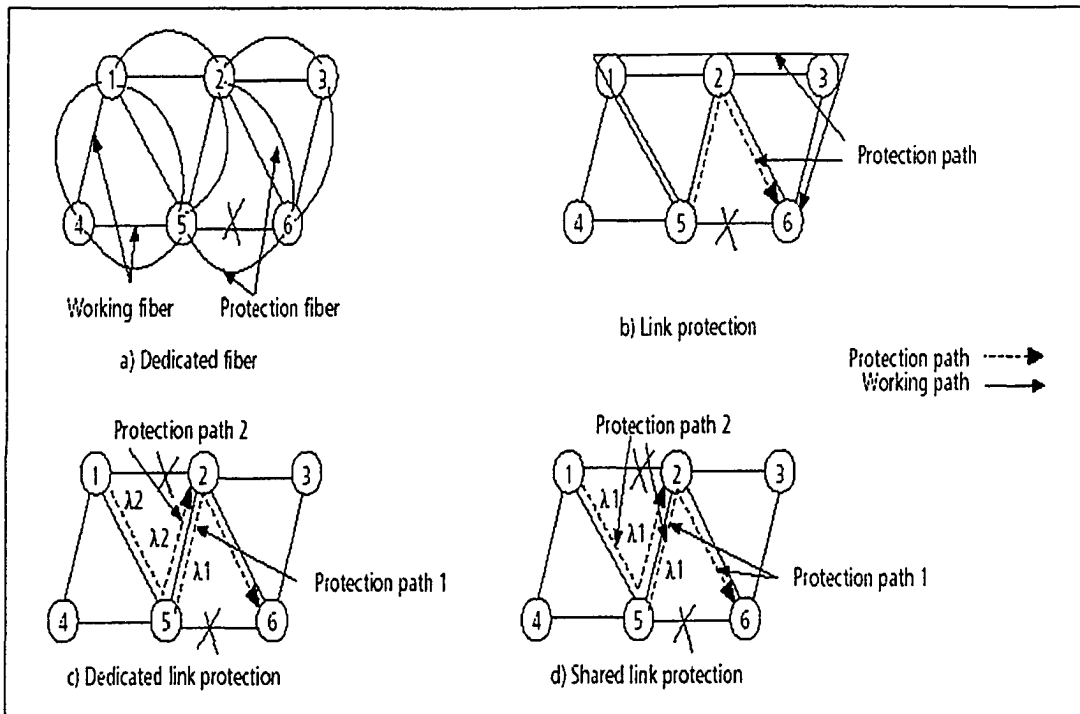


Figure 3.3- Link-based protection

### 3.2.2.2- Path-Based Protection:

In WDM systems, path-based protection refers to the reservation of a protection path and wavelength for each working wavelength path and each link failure. Upon failure of a link, the source and destination nodes of each affected connection switch to the corresponding protection wavelength paths. As opposed to link-based protection, which involves only the nodes adjacent to the link failure, path-based protection needs a mechanism to notify the affected connection end nodes of the failure. This requires the cooperation of several network nodes, and may not be easily achievable. The protection wavelength paths for every link failure are usually reserved at connection setup, and should be disjoint with the failed link. Upon link failure, the wavelength paths reserved for this failure scenario are activated. As a

special case, when a protection wavelength path is disjoint with every link of the working path, the same wavelength path can be used to restore a connection upon *any* single-link failure along the working path. Note that in this case, the identification of the failed link is not required to initiate recovery. An example of the special case is shown in Fig. 3.4a, where the working path is 4–5–6. When the link between nodes 5 and 6 fails, nodes 4 and 6 switch the connection to the protection path 4–1–2–6. The wavelength used on the protection path can be the same as or different from the working wavelength. Also, the protection paths used for different connections using the same working path can be different. Similar to link-based protection, path-based protection can be dedicated or shared. In *dedicated path protection*, the backup wavelength on the links of a protection path is reserved for a specific working connection. This implies that two overlapping protection paths must use different wavelengths even if the working paths do not overlap. For example, Fig. 3.4b shows two working paths, 4-5-6 and 1-2-3, both using  $\lambda_1$ . The protection wavelength path for connection 1 is  $\lambda_2$  on 4–1–2–6 ( $\lambda_1$  is a working wavelength on link 1–2 and cannot be used for protection). The protection wavelength path for connection 2 is 1–5–2–6–3. Since these two protection paths have the common link 2–6, and  $\lambda_2$  is assigned to protection path 1, protection path 2 has to be assigned a different wavelength (e.g.,  $\lambda_1$ ). Dedicated path protection requires a large amount of extra capacity for protection purposes, and when there is no failure, the protection resources are kept idle. The positive aspect is that it is able to provide recovery from not only single-link failures, but also *some* multilink failures.

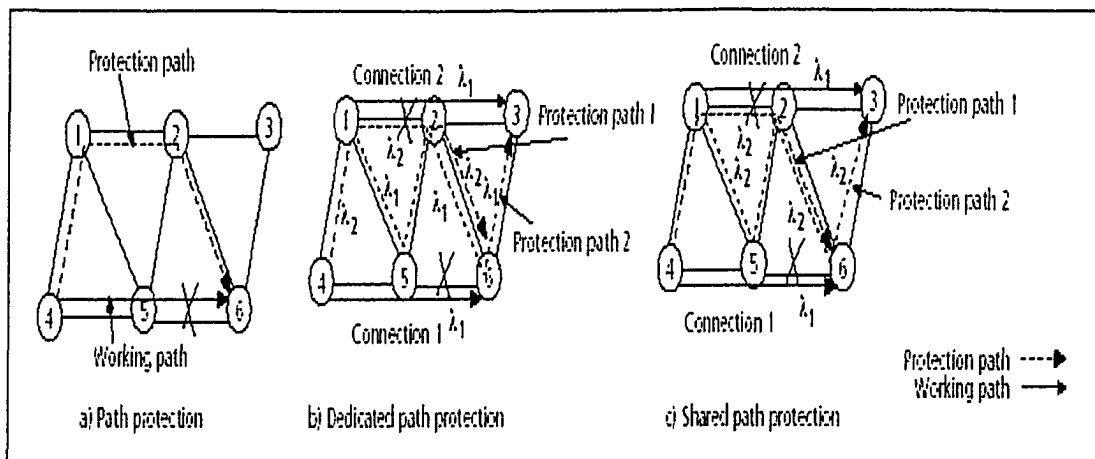


Figure 3.4- Path-based protection

*Shared path protection* allows the use of the same wavelength on a link for two different protection paths if the corresponding working paths are link-disjoint. Thus, it is possible to utilize the capacity more efficiently, while still achieving 100 percent recovery from single-link failures. An example of shared path protection is given in Fig. 3.4c. The two backup paths can now share  $\lambda_2$  on link 2–6. Therefore, only one wavelength on this link has to be reserved for protection, as opposed to two for dedicated path protection.

### 3.3 Routing in survivable WDM mesh networks

The recent advances in optical technologies have enabled the migration of optical networks from point-to-point and ring-based topologies into wavelength routed mesh based architectures. WDM-based mesh network infrastructures are likely to provide high capacity and cost effective transport network. Given the scale of these architectures, where each fiber can potentially carry terabits of data, service

survivability [2,3] becomes a critical requirement for network planning and management [1]. Service survivability stipulates that upon the failure of any network element, all affected connections be rerouted within a short time interval using spare capacity reserved on alternate paths. Alternatively, if spare capacity is not reserved a-priori, the network must dynamically search/compute alternate routes and reserve resources to restore affected connections. However, since this latter approach exhibits longer restoration times and cannot guarantee complete service recovery, it may not be very feasible thus we only consider the case where spare capacity is already reserved during connection setup, referred to hereafter as path protection.

A variety of optical path protection schemes have been proposed using concepts such as disjoint dedicated paths (1+1, 1:1), or shared backup paths [1]. In 1+1 protection, traffic is transmitted simultaneously on both working and protection paths, and in case of failure, the destination is required to switch between working and protection paths, i.e., *non-signaled* switchover. Although such recovery is very fast, it is very inefficient due to the inherent resource redundancy. Meanwhile in 1:1 path protection, both working and protection paths are provisioned simultaneously, but data is only rerouted to the working path. As a result, protection paths can be used to transmit low priority *pre-emptable* traffic during non-failure conditions. Namely, in case of a failure along the working path, the source node is notified through *signaling* messages and data is transmitted on the protection path. As in 1+1 protection, recovery here is relatively fast, but efficiency is somewhat improved. Shared path protection builds upon this concept by further increasing resource efficiency. Here, backup resources can be *shared* amongst many lightpaths if they fail independently.

Now since resources along the protection path are shared, switch nodes along the protection path are not configured at connection setup. Instead, upon a failure, source nodes (of failed paths) are notified and generate signaling messages to configure the switches along the protection path. In this work we focus on shared path protection as it yields fast restoration/high efficiency for mesh WDM networks and we address the online path selection algorithm of restorable connections as a key issue pertaining to shared mesh protection.

### 3.3.1- Background and Motivation:

We consider the problem of online provisioning of mesh-restored optical connections, where distributed path selection is used to compute optimal pairs of primary and backup paths [8, 9]. In this scenario, where a source node has to make a routing decision without the knowledge of future requests, a common routing objective is to select potential diverse paths (primary and backup) while allocating the minimum number of resources, i.e., maximize the use of reserved sharable channels. Another objective is also to select backup paths that provide fast recovery for the affected connections in case of link failure. Now since resources on shared backup paths are only configured *after* failure occurrences, this requires the selection of shorter backup paths to ensure rapid service recovery. However, these two optimality criteria typically conflict most of the time, i.e., to maximally exploit the gain of wavelength sharability, one needs to allow longer backup paths, however shorter paths are required for fast service restorability. Also note that these sophisticated path

selection algorithms require detailed knowledge of network state, as derived from extensions to existing link state routing protocols [10, 11].

Recent work on service survivability has focused on implementing capacity design algorithms for fast connection recovery in WDM networks. For example, in [12], the authors formulate the problem as a generalization of the maximum concurrent flow problem. Resources on backup paths are pre-configured to ensure fast service recovery, however restrictions on backup path sharing are imposed. On the other hand, the authors of [13] present a stochastic approach to compute shared mesh restorable connections to improve network scalability. Here, less information is deemed sufficient to determine the sharability of channels with remarkable accuracy. Meanwhile, a distributed path selection algorithm for a GMPLS-based shared mesh restoration scheme is presented in [9]. Here, a Full Information Restoration (FIR) algorithm is proposed, using signaling protocol extensions to distribute and collect additional link state information without relying upon centralized setups. On the other hand, [14] addresses the problem of efficient shared path protection under distributed control with only partial information knowledge.

### **3.3.2- Efficient Path Selection Algorithms:**

We propose two distributed path selection schemes, the first one relies upon global information (Distributed Path selection with Global Information, DPGI) at each node to determine link sharability and compute optimal routes. Meanwhile, the second approach only maintains local information (Distributed Path selection with

Local Information, DPLI) and determines link sharability in a distributed manner.

The details of both algorithms are next presented:

### 3.3.2.1- Distributed Path selection with Global Information, DPGI:

In DPGI, upon arrival of a connection request, the source node computes a set of  $K$  shortest paths for the primary,  $P = \{p_1, p_2, \dots, p_K\}$ . Then for each  $p_n \in P$ , a link-disjoint secondary path and a wavelength,  $s_n^{\lambda_i}$ , are selected simultaneously ( $s_n^{\lambda_i}$  does not necessarily have to be the shortest link disjoint path). Consider the following notation, where a network is represented as a directed graph  $G(N, J, W)$ , with  $N$  being the set of nodes,  $J$  the set of links, and  $W$  the set of wavelengths. This graph can be also represented by a set of  $W$  virtual graphs,  $G^{\lambda_i}(N, J)$   $i = 1, \dots, W$ , with each one corresponding to a specific wavelength. Then using the network state information available at the source node, the links weights on each graph are adjusted as follows:

$$C^{\lambda_i}(j) = \begin{cases} 0 & \text{if } \lambda_i \text{ on link } j \text{ is sharable} \\ 1 & \text{if } \lambda_i \text{ on link } j \text{ is not used} \\ \infty & \text{if } j \in p_n \text{ or } \lambda_i \text{ is not sharable} \end{cases} \quad (1)$$

where  $C^{\lambda_i}(j)$  is the weight assigned to a link  $j$  on  $G^{\lambda_i}$ . Note that the source node can determine the sharability (or non-sharability) of a link by checking its global information database. If the primary path of the connection in consideration is link-disjoint with the primary paths of all connections protected by this link, then a wavelength is said to be sharable. A regular shortest path algorithm (e.g. Dijkstra) is then used to compute the shortest path on each graph. This procedure is repeated for all virtual graphs, and hence up to  $W$  link disjoint paths (or less) with  $p_n$  can be

found. From this, the path with the minimal total cost is selected as a candidate disjoint path:

$$\text{Min}_i (C_n^{\lambda_i}), \text{ where } C_n^{\lambda_i} = \sum_{j \in s_n^{\lambda_i}} C^{\lambda_i}(j) \quad (2)$$

After computing all pairs  $(p_n, s_n^{\lambda_i})$   $n = 1, \dots, K$ , the pair with the minimal combined cost is selected. Here the cost of the primary path is the number of hops along the path, while the cost of the secondary path is as defined in (2). Once  $p_n$  is identified, a wavelength is assigned for the path. Note that any wavelength selection scheme can be used, but for simplicity, we use the First Fit approach. After selecting the primary/backup path pair and their corresponding wavelengths, the source node triggers its signaling protocol to simultaneously reserve resources along both paths. Subsequently, after receiving two positive acknowledgement (ACK) signaling messages, the source starts transmitting data on the primary path.

Note that besides the complexity associated with this algorithm, a large amount of information needs to be stored and frequently updated to reflect any network-wide changes. These updates can become a severe bottleneck if connections setup/takedown frequencies are high. Moreover, as network size increases, the scalability of this protocol becomes another concern.

### 3.3.2.2- Distributed Path selection with Local Information, DPLI:

Here, we propose an alternative distributed protocol which exploits the sharability property of reserved channels yet only requires that local information be maintained at each node. This information is maintained in a local database to reflect the status of

each wavelength on each outgoing link. Namely, if a wavelength on an outgoing link is reserved for protection, a list of all primary paths of connections protected by this wavelength is also stored. Each node also maintains up to  $K$  shortest-paths and their corresponding link-disjoint paths.

Upon receiving a connection request, the source node selects a pair of disjoint paths from the arranged fixed set of routes. This path pair could be the first one in the list or the shortest pair. Subsequently the source node sends a *Probe* message on each route toward the destination to collect available resources on each path. This message carries the set of free wavelengths along the first path(s) link to be established. Additionally, the *Probe* message sent along the backup path also has an additional field that contains the physical route of the working path of the connection (another one-bit flag is also attached to the message to indicate the type of the path to be setup). When an intermediate node receives this message, it determines the set of available wavelengths on its outgoing link (via the local database) and generates a new wavelength vector to be attached to the *Probe* message. This message is then forwarded to intermediate downstream nodes. These nodes determine the sharability of the wavelength by retrieving the route of the connection's working path, say  $p$ , from the received *Probe* packet and checking to see if it is link-disjoint with the routes,  $p1$ ,  $p2$ , etc., of the connections protected by this wavelength.

Meanwhile, upon receiving a *Probe* message on the primary path, the destination node selects a wavelength from the attached set of available wavelengths using any wavelength selection scheme. This node then generates a *Reserve* message containing the selected wavelength and forwards it back upstream towards the source

to configure intermediate switches. Each intermediate node checks to see if this wavelength is available and if so, reserves it by cross-connecting its switch and updating its local database. Otherwise, the intermediate node sends a negative acknowledgement (*NACK*) message to the source and a *Release* message back to the destination to indicate the setup failure. On the other hand, when receiving a *Probe* message on the secondary path, a wavelength is selected and a *Reserve* message is sent back to the source. Here, the *Reserve* message only updates the local information database at each intermediate node without cross-connecting the switch. Upon receiving both *Reserve* messages, the source deems the connection request successful and begins to transmit data on the primary path. Otherwise, if at least one *NACK* message is received, the connection setup fails and any reserved resources must be released (for further details on this protocol, please refer to next chapter). Unlike the previous protocol, the complexity of this algorithm is minimal since path computation is done off-line. This improves scalability with network size since only local information is maintained at each node.

### **3.4 Analytical model for shared path protection**

In WDM network with path protection, a connection request is successful if two diversely routed paths can be found between a source-destination (s-d) pair and commensurate resources (wavelengths) can be allocated for them. Given the fact that both paths are diversely routed, blocking on one path is independent from the blocking on the second path. Hence the success probability of one connection is defined to be:

$$\begin{aligned}
S^{s-d} &= \text{prob}(\text{success on primary and success on secondary}) \\
&= \text{prob}(\text{success on primary}) * \text{prob}(\text{success on secondary}) \\
&= [1 - \text{prob}(\text{blocking on primary})] * [1 - \text{prob}(\text{blocking on secondary})].
\end{aligned}$$

The blocking for a connection between an s-d pair is computed as follows:

$$P^{s-d} = 1 - S^{s-d} = 1 - (1 - p_1)(1 - p_2) = p_1 + p_2 - p_1 p_2 \quad (3)$$

In other words, one needs to compute the blocking on both the primary ( $p_1$ ) and on the secondary ( $p_2$ ) for each connection so that overall blocking can be approximated. More detailed treatments for network traffic modeling and path blocking are now considered:

#### 3.4.1- Network and traffic assumptions:

In this section we state the assumptions about network characteristics and traffic behaviors, which are used in our models for calculating path-blocking probability without wavelength converters. Let a network with  $N$  nodes and  $J$  links be represented by a directed graph  $G(N, J)$ , with each link having the same number of wavelengths ( $W$ ). Connection call requests arrive according to a Poisson process and the duration of each call is exponentially distributed with unit mean. Furthermore, let  $\lambda$  be the total traffic load offered to the network,  $a_r$  be the offered load to the  $r^{\text{th}}$  s-d connection pair,  $E = \{1, 2, \dots, J\}$  be the set of all links, and  $W = \{\lambda_1, \lambda_2, \dots, \lambda_W\}$  is the set of all wavelengths.

For analytical simplicity and tractability considerations, we pre-compute the possible set of primary and secondary routes between any s-d pair.

Let  $R^{(1)} = \{r_1^{(1)}, r_2^{(1)} \dots r_{N(N-1)}^{(1)}\}$  be the set of all primary routes in the network, and  $R^{(2)} = \{r_1^{(2)}, r_2^{(2)} \dots r_{N(N-1)}^{(2)}\}$  be the set of the secondary routes, link-disjoint with their corresponding primary routes. Here, the pair  $(r_i^{(1)}, r_i^{(2)})$  is the pair of primary and secondary routes between the  $i^{\text{th}}$  s-d pair.

In our approach, the problem of computing path (primary, secondary) blocking on a particular wavelength is reduced to computing approximate blocking probability in a circuit switched network. Specifically, we extend the layered-graph approach [15], to develop an approximate analytical model for the path blocking. By using this layered approach, the RWA problem is reduced to only finding a wavelength on an end-to-end path between the s-d pair. The search for a wavelength on a path may be viewed as a search over a sequence of  $W$  logical paths, where a logical path is a combination of a predefined physical path and a particular wavelength. If the logical path under consideration is not available then the connection request overflows to the next logical path. When we find the two logical paths under consideration, the connection request is established. Otherwise, the connection is blocked. Here, we use moment matching techniques, specifically the Equivalent Random Method (ERM) [16] to model the overflow traffic between different wavelength layers.

Further consider the following variable definitions:

$a_{r_i^{(1)}}^{\lambda_k}$  : The arrival rate on primary route  $r_i^{(1)}$  on wavelength layer  $\lambda_k$  .

$a_{r_i^{(2)}}^{\lambda_k}$  : The arrival rate on secondary route  $r_i^{(2)}$  on wavelength layer  $\lambda_k$  .

$B_j^{\lambda_k}$  : The link blocking probability on link  $j$  on wavelength  $\lambda_k$  .

$\alpha_j^{\lambda_k}$ : The traffic load on link  $j$  on wavelength  $\lambda_k$ .

$L_{r_i^{(1)}}^{\lambda_k}$ : The primary path blocking probability on primary route  $r_i^{(1)}$  on wavelength layer  $\lambda_k$ .

$L_{r_i^{(2)}}^{\lambda_k}$ : The secondary path blocking probability on secondary route  $r_i^{(2)}$  on wavelength layer  $\lambda_k$ .

If we consider a network with  $F$  fibers per link, the number of idle circuits on link  $j$  on wavelength  $\lambda_k$  can be viewed as a birth-and-death process given that the traffic arrival follows a Poisson distribution:

$$P_m = \text{Prob} \{m \text{ circuits are used}\} = \frac{(\alpha_j^{\lambda_k})^m}{m!} \left[ 1 + \sum_{n=1}^F \frac{(\alpha_j^{\lambda_k})^n}{n!} \right]^{-1} \quad (4)$$

$$B_j^{\lambda_k} = \text{Prob} \{F \text{ circuits are used on link } j \text{ on wavelength } \lambda_k\} = P_F \quad (5)$$

### 3.4.2- End-to-end path blocking:

The traffic load on link  $j$  on wavelength  $\lambda_k$  is determined to be:

$$\alpha_j^{\lambda_k} = \alpha_j^{\lambda_k,(1)} + \alpha_j^{\lambda_k,(2)} \quad (6)$$

where  $\alpha_j^{\lambda_k,(1)}$  is the traffic load on link  $j$  from the set of primary routes passing by link  $j$ , and  $\alpha_j^{\lambda_k,(2)}$  is the traffic load on link  $j$  from the secondary routes passing by link  $j$ .

Note that the traffic load on link  $j$  from a connection routed on a path  $r$  (whether primary or secondary) can be approximated as follows [16]:

$$\alpha_j = a_r \times \prod_{f \in r - \{j\}} (1 - B_f) \quad (7)$$

where  $a_r$  is the offered traffic on route  $r$ .

Obviously, one can estimate  $\alpha_j^{\lambda_k, (1)}$ :

$$\alpha_j^{\lambda_k, (1)} = \sum_{r^{(1)}, j \in r^{(1)}} a_{r^{(1)}}^{\lambda_k} \cdot \prod_{m \in r^{(1)} - \{j\}} (1 - B_m^{\lambda_k}) \quad \forall r^{(1)} \in R^{(1)} \quad (8)$$

On the other hand, the traffic load on link  $j$ , due the set of secondary routes cannot be approximated by (8) if shared path protection is used. However, if dedicated path protection is being considered, similar to (8) one can compute:

$$\alpha_j^{\lambda_k, (2)} = \sum_{r^{(2)}, j \in r^{(2)}} a_{r^{(2)}}^{\lambda_k} \cdot \prod_{m \in r^{(2)} - \{j\}} (1 - B_m^{\lambda_k}) \quad \forall r^{(2)} \in R^{(2)} \quad (9)$$

In the case of shared path protection, we develop the following model to compute the link load.

#### 3.4.2.1- Property 1:

Let  $s_1$  and  $s_2$  be two secondary paths passing through a link  $j$  and let  $p_1, p_2$  be their corresponding primary paths. Additionally, let  $\alpha_{j,1}$  and  $\alpha_{j,2}$  be the corresponding traffic load on link  $j$  from both secondary paths independently (using Eq. 7). Then the aggregated link load can be computed as follows:

$$\begin{aligned} \alpha_j &= \text{Max}(\alpha_{j,1}, \alpha_{j,2}) \quad \text{if } p_1 \cap p_2 = \Phi \\ &= \alpha_{j,1} + \alpha_{j,2} \quad \text{if } p_1 \cap p_2 \neq \Phi \end{aligned} \quad (10)$$

Using this property one can generalize to the case where link  $j$  is shared by  $m$  secondary paths,  $S^j = \{s_1, s_2, \dots, s_m\}$  where their corresponding set of primary paths is  $P^j = \{p_1, p_2, \dots, p_m\}$ . However, the difficulty here stems from the fact that some of the corresponding primary paths of this set of secondary paths (passing through link  $j$ ) might themselves be link disjoint amongst each other while others could overlap in at least one link. In other word, if  $p_1 \cap p_2 \cap \dots \cap p_m = \Phi$ , then one

can write  $\alpha_j = \text{Max}(\alpha_{j,1}, \alpha_{j,2}, \dots, \alpha_{j,m})$ ; however, if  $p_1 \cap p_2 \cap \dots \cap p_m \neq \Phi$  does not imply  $\alpha_j = \alpha_{j,1} + \alpha_{j,2} + \dots + \alpha_{j,m}$ , rather,  $\alpha_j = f(\alpha_{j,1}, \alpha_{j,2}, \dots, \alpha_{j,m})$  and needs to be determined.

### 3.4.2.2- Computation of $\alpha_j$

In this section we are interested in computing the aggregated load on link  $j$  from a set of  $m$  secondary paths passing by this link. We define  $F = E - \{j\}$ , the set of all links in the network except link  $j$ .

We further define the array  $T[N(i)]$ , where each element  $N(i)$  is:

$$N(i) = \{p_k / i \in p_k, k = 1, \dots, m\} \quad \forall i \in F$$

$T[N(i)]$  is an array of size  $|F|$ , where an element at index  $i$  represents the subset of all primary paths from the set  $\{p_1, p_2, \dots, p_m\}$  passing by link  $i$ .

Next we compute independently the traffic load on link  $j$  from each secondary route passing by link  $j$  (using Eq. 7),  $\alpha_{j,1}, \alpha_{j,2}, \dots, \alpha_{j,m}$ . From this we then compute:

$$\begin{aligned} \beta(i) &= f\{N(i)\} = \sum_{k=1}^{|N(i)|} \alpha_{j,k} \quad \text{if } |N(i)| \neq 0 \\ &= 0 \quad \text{otherwise} \end{aligned} \quad (11)$$

$\beta(i)$  represents the traffic load on link  $j$  from the subset of the secondary paths in the set  $\{s_1, s_2, \dots, s_m\}$  where their respective subset of primary paths passes through link  $i$ .

After computing  $\beta(i)$ ,  $\forall i \in F$ , one can compute the overall traffic load on link  $j$ :

$$\alpha_j^{\lambda_k(2)} = \text{Max}_i \{\beta(i)\} \quad (12)$$

Using both equations (8) and (12), one can compute the total traffic load on a link  $j$  using equation (6). Equations (5) and (6) can be solved iteratively to compute the exact values for the link blocking.

On the other hand, given the link blocking on a given wavelength layer, one can compute the primary path and secondary path blocking between each s-d pair on this wavelength layer as:

$$L_{r_i^{(1)}}^{\lambda_k} = 1 - \prod_{j \in r_i^{(1)}} (1 - B_j^{\lambda_k}), \forall r_i^{(1)} \in R^{(1)} \quad (13)$$

$$L_{r_i^{(2)}}^{\lambda_k} = 1 - \prod_{j \in r_i^{(2)}} (1 - B_j^{\lambda_k}), \forall r_i^{(2)} \in R^{(2)} \quad (14)$$

Additionally, all traffic blocked on wavelength  $\lambda_k$  overflows to wavelength  $\lambda_{k+1}$ :

$$a_{r_i^{(1)}}^{\lambda_{k+1}} = a_{r_i^{(1)}}^{\lambda_k} \cdot L_{r_i^{(1)}}^{\lambda_k} \quad (15)$$

$$a_{r_i^{(2)}}^{\lambda_{k+1}} = a_{r_i^{(2)}}^{\lambda_k} \cdot L_{r_i^{(2)}}^{\lambda_k} \quad (16)$$

Blocked traffic keeps overflowing until the last wavelength, where it is considered to be fully blocked. Hence the overall primary path blocking (respectively the secondary) is computed as:

$$L_{r_i^{(1)}} = \prod_{k=1}^W L_{r_i^{(1)}}^{\lambda_k}, L_{r_i^{(2)}} = \prod_{k=1}^W L_{r_i^{(2)}}^{\lambda_k} \quad (17)$$

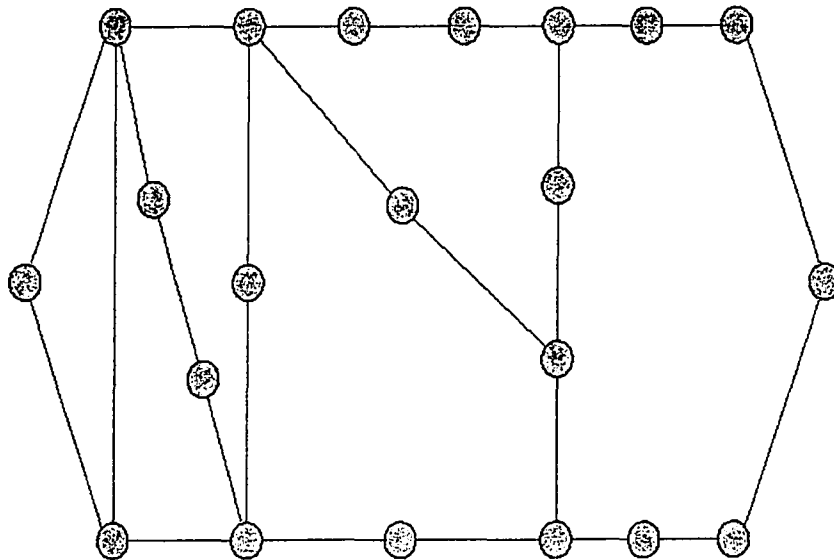
Namely, (17) computes the blocking on route  $r_i$  by computing the product of the route blocking on each wavelength correspondingly.

Having  $L_{r_i^{(1)}}$  and  $L_{r_i^{(2)}}$ ,  $\forall r_i^{(1)} \in R^{(1)}$ ,  $\forall r_i^{(2)} \in R^{(2)}$ , one can compute the connection blocking using equation (3) and then the overall network blocking can be computed:

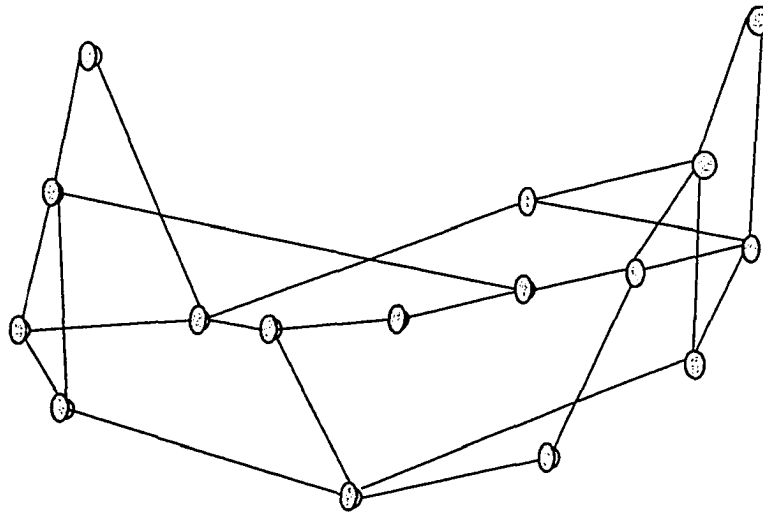
$$P = \frac{\sum_{s-d} P^{s-d}}{N(N-1)} \quad (18)$$

### 3.5 Performance evaluation

In this section we use simulation results to demonstrate the accuracy of our derived analytical model. We also compare the performance of the different path selection algorithms, in terms of blocking probability, sharability gain, and connection average hop count. Here, two network topologies are used: the 16-node NSF network and 21-node DARPA network (Fig. 3.5). The number of wavelengths per fiber is set to 4 and we consider only one fiber per link. Connection requests arrive at each node according to a Poisson process with arrival rate  $\lambda$  and the connection holding times have exponential distribution with unit mean.



(a) DARPA Network



(b) NSF Network

Figure 3.5- Simulated Network Topologies

First, we show in figure 3.6 the comparison between the proposed analytical model and simulation results for dedicated and shared path protection. We solve the model proposed in section II to determine the end-to-end blocking probability (Eq. 18). First, the path blocking (primary and secondary) on each wavelength graph is computed using equations (13) and (14), where the link blocking is calculated by iteratively solving (5) and (6). Namely, traffic blocked on a given wavelength overflows onto the next wavelength (Eq. 15 and 16) and so on until the last wavelength. The total path (primary and secondary) blocking is computed using (17), and the connection blocking is computed using (3). From this, the overall network blocking is computed using (18). We use the 16-node NSF network to evaluate the performance of this analytical model. Although larger networks can also be evaluated, related convergence times for the numerical solutions become prohibitive.

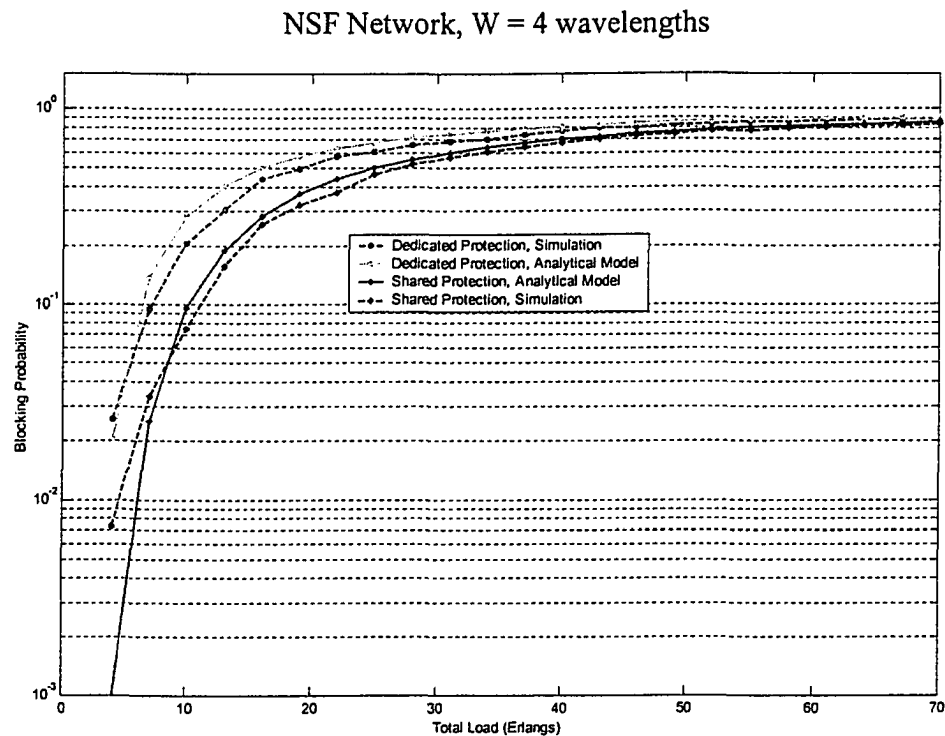


Figure 3.6- Comparison of Analytical model and simulation

As Figure 3.6 indicates, our methodology for estimating link loading in shared path protection is fairly accurate, and matches the simulation results closely. However, since backup paths usually tend to have slightly longer routes, the analytical model slightly overestimates the blocking probabilities observed in the simulations.

Next, we compare the performance of the path selection algorithms presented in section III. The first one relies upon global information maintained at each node, DPGI, while the second only uses local information, DPLI. Overall, our results confirm that the path selection model has a significant impact on network performance. Here, three plots are shown on the figure (Fig. 3.7), with the first one representing DPGI and the other two plots representing DPLI. As mentioned earlier, when using DPLI, the source node maintains an ordered list of  $K$  shortest paths. In DPLI1 (Figure 3.7), the source node always selects the first pair of paths (which

corresponds to the shortest primary and its shortest link-disjoint backup) in the ordered list. Meanwhile DPLI2 (Figure 3.7) selects the shortest pair. The results indicate that DPGI outperforms the other schemes since it can better exploit the sharability of reserved backup channels by favoring the selection of protection routes with more reserved sharable channels. However, the drawback of DPGI is the high complexity involved in the path selection procedure. Moreover, as connections setup/takedown frequencies increase, excessive updates are required to ensure up-to-date resource usage information. Clearly, this poses a notable scalability problem and mandates larger databases as network loadings increases. On the other hand, DPLI1 slightly outperforms DPLI2, which is due to the fact that the former scheme selects longer protection paths, thus allowing for more channel sharability.

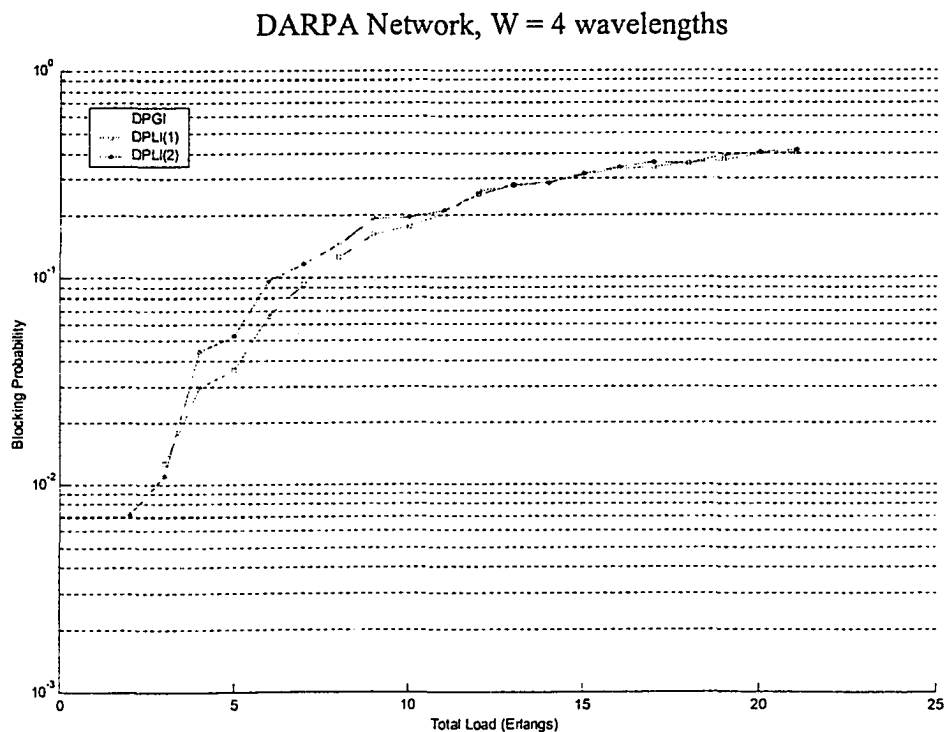


Figure 3.7- Comparison between different path selection algorithms

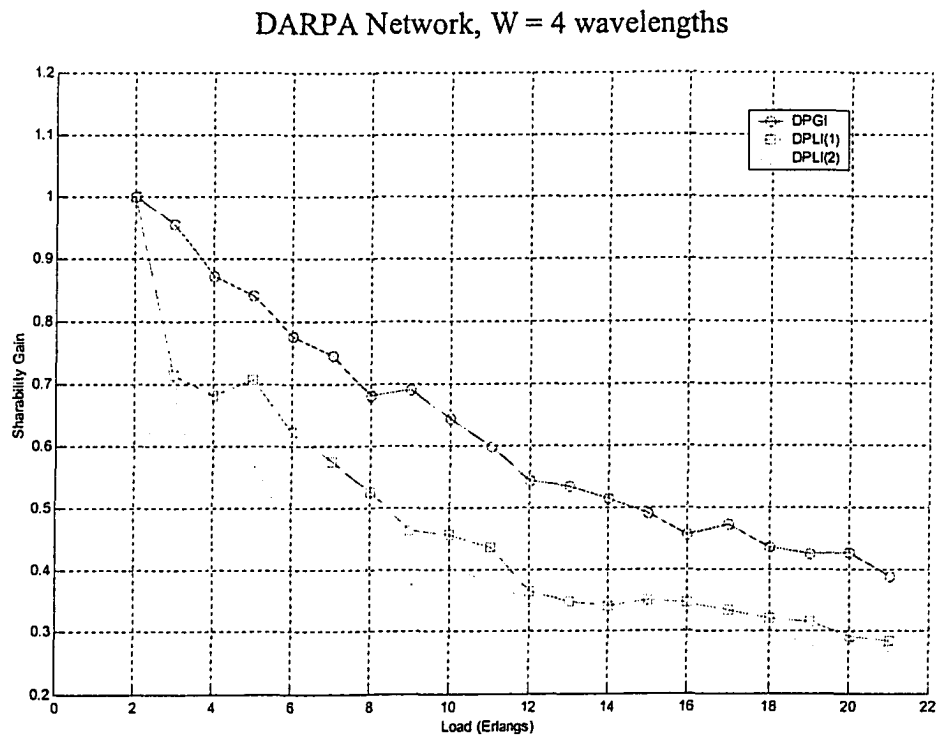


Figure 3.8- Impact of path selection on sharability gain

Another key performance comparison is the *sharability gain* of the path selection algorithms (Figure 3.8). This value is defined as the improvement the algorithm yields over conventional dedicated path protection. As expected, DPGI exhibits better sharability gain, making it more favorable if the network utilization is the key criterion to be optimized. The figure shows that DPGI improves the sharability gain by 15-25% with respect to the other path selection schemes.

So far, our comparison has only focused on the merits of the DPGI scheme versus other path selection algorithms. However, although DPGI is more resource-efficient, i.e., since it favors longer backup paths to maximally exploit sharability, it generally has slower recovery timescales. For example, Figure 3.9 shows the average path (primary, secondary) hop counts for DPGI, DPLI1 and DPLI2. As expected, DPGI yields longer backup paths (approximately 8 hops), whereas DPLI2 yields

shortest backup paths (approximately 5.5 hops). Meanwhile, DPLI1 performance is found to lie in between the above two schemes (approximately 7 hops).

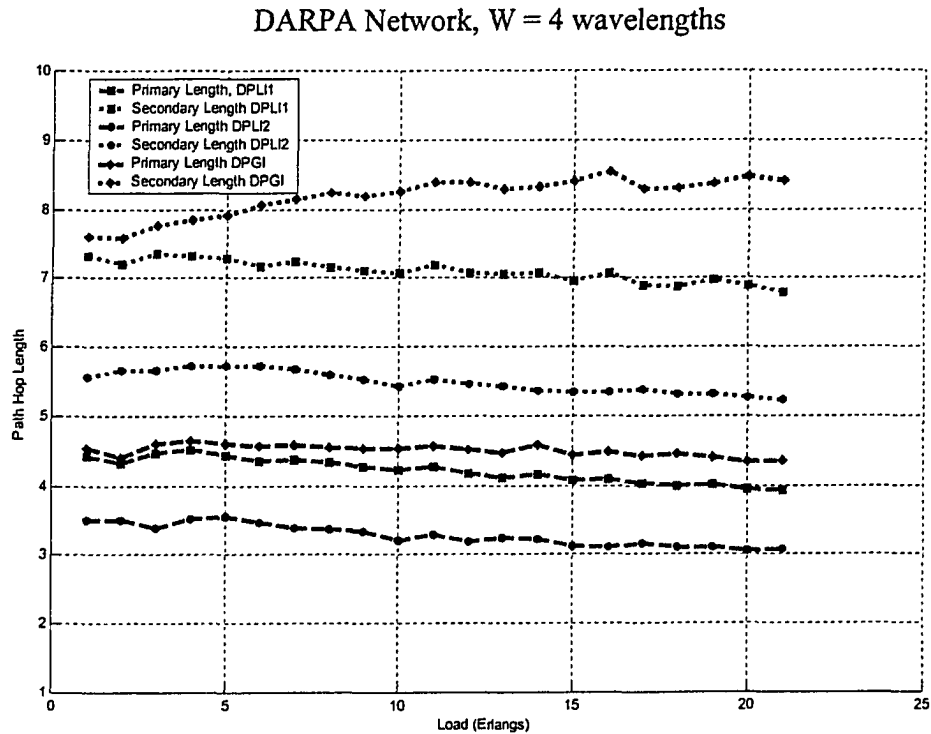


Figure 3.9- Impact of path selection on the average hop count

These findings have a major impact on recovery time analysis. Upon detecting any network element failure, the source node of each affected connection is notified via signaling messages. Subsequently, the notified node initiates the backup path setup by sending a Failure Recovery Message along the protection route to cross-connect the switches at intermediate nodes. Meanwhile, the source has to wait for an ACK from the destination before it can start transmitting data on the backup route. Therefore, at least one round trip delay on the backup path is incurred, in addition to the message processing time at each intermediate node and the time taken to

configure each intermediate switch. Clearly, recovery times also depend very much upon the length of the backup path, as well as message processing and switch configuration times. When the propagation time is a dominant factor (e.g., a 1000-mile fiber span implies 5.36 ms propagation delay), the use of longer backup paths may not be feasible (e.g., a 5000 miles cross-country path will incur 53.64 ms of only round trip propagation delay aside from per-node message processing time and cross connection times). Hence, short backup path selection is crucial when the restoration time is a key requirement. Overall, the use of DPGI is unfavorable herein, and instead schemes such as DPLI2 are more suited. For further details on recovery time analysis and the impact of related signaling procedures, see chapter 4.

### **3.6 Conclusion**

In this chapter, we addressed the path selection problem as a key requirement in the design of efficient mesh restorable networks. Two path selection algorithms were presented, namely the DPGI that relies on global information maintained at each node and DPLI, where only local information is necessary. We showed that path selection algorithms intended to optimize the resource utilization by maximally exploiting the sharability of reserved channels do not necessarily result in fast restoration. Rather, these two optimality criteria tend to be conflicting. Analytical model for evaluating the performance of shared path protection is presented and its accuracy is validated through numerical comparison with simulation results.

### 3.7 References

- [1] B. Doshi et al, "Optical Network Design and Restoration" Bell Labs Tech. J., vol.4, no.1, Jan-Mar 1999, pp. 58-84
- [2] G. Mohan, C.S.R. Murthy "Lightpath Restoration in WDM Optical Networks", IEEE Network, November/December 2000.
- [3] D. Zhou, and S. Subramanian "Survivability in Optical Networks", IEEE Network, November/December 2000.
- [4] P. Demeester et al., "Resilience in Multilayer Networks," IEEE Communication Magazine, vol. 37, no. 8, Aug. 1999, pp. 70-76.
- [5] O. Gerstel, R. Ramaswami, and G. H. Sasaki, "Fault Tolerant Multi-wavelength Optical Rings with Limited Wavelength Conversion," *IEEE JSAC*, vol. 16, no. 7, Sept. 1998, pp. 1166–78.
- [6] O. Crochat and J. Y. Le Boudec, "Design Protection for WDM Optical Networks," *IEEE JSAC*, vol. 16, no. 7, Sept. 1998, pp. 1158–65.
- [7] T. H. Wu, Fiber Network Service Survivability, Artech House, 1992.
- [8] C. Assi, A. Khalil, N. Ghani, and M Ali, "Efficient Shared Path Protection in Mesh WDM Networks" The Journal of Networks, Software Tools, and Applications, Cluster computing, to appear, 2004.
- [9] G. Li, W. Wang, C. Kalmaneck, and R. Doverspike, "Efficient Distributed Path Selection for Shared Restoration Connections" *IEEE INFOCOM 2002*, NY.
- [10] K. Kompella et al. "OSPF extensions in support of Generalized MPLS" IETF Internet Draft, work in progress, Feb 2001.
- [11] K. Kompella et al. "IS-IS extensions in support of Generalized MPLS" IETF Internet Draft, work in progress, Feb 2001.

- [12] O. Hause, M. Kudialam and T. V. Lakshman, "Capacity Design of Fast Path Restorable Optical Networks" *IEEE INFOCOM* 2002, NY.
- [13] E. Bouillet, J. Labourdette, G. Ellinas, R. Ramamurthy, S Chaudhuri, "Stochastic Approaches to Compute Shared Path Mesh Restored Lightpaths in Optical Network Architecture" *IEEE INFOCOM* 2002, NY.
- [14] C. Qiao, and D. Xu, "Distributed Partial Information Management (DPIM) Schemes for Survivable Networks-Part I" *IEEE INFOCOM* 2002, NY.
- [15] C. Chen and S. Banerjee, "A New Model for Optimal Routing and Wavelength Assignment in WDM Optical Networks" *IEEE INFOCOM* 1996.
- [16] A. Girard, *Routing and Dimensioning in Circuit-Switched Networks*, Addison Wesley, 1990.

## Chapter 4

# 4. Connection Management for Survivable WDM Networks

### 4.1 Introduction

Recent advances in optical networking technology, first with wavelength-division multiplexing (WDM) and more recently with optical cross-connects (OXCs) along with the wide deployment of high-speed IP/MPLS routers is setting the foundation for the next-generation data-centric networking paradigm. In this scenario, the role of synchronous digital hierarchy/optical network (SDH/SONET) will diminish, and future IP networks will evolve towards a model comprising high-performance IP/MPLS routers interconnected by intelligent optical core networks (IP-over-WDM) that will directly provide a global transport infrastructure for legacy and new IP services. A major driver for realizing this evolution is the potential ability of such networks to provide fast automatic setup and teardown of lightpaths across the optical network, with the capability of supporting diverse client signals on the paths. Provisioning of lightpaths requires control and management protocols to perform routing and wavelength assignment (RWA) functions, as well as to exchange signaling information and to reserve resources along the provisioned paths.

Equally important to the process of dynamically provisioning lightpaths in mesh-based wavelength-routed networks is the reliability offered by the network to the services and lightpaths it supports. This requires the development of the appropriate protection and restoration schemes, which minimize the data loss when a link failure occurs [1, 2, 3, 4].

This chapter will focus on the implementation of network control and management protocols in survivable wavelength routed WDM networks. This includes routing and signaling protocols to setup and take down protected lightpaths, and fault detection and fast restoration algorithms associated with the network management plane. The rest of the chapter is organized as follows. In section 4.2 we discuss the various issues related to the dynamic control and management of lightpaths in WDM optical networks and review some of the protocols proposed in the literature. In section 4.3, we propose a hybrid distributed fault-management protocol that attempts to combine the best of both the link state and distributed-routing approaches. Fault detection and fast Restoration Algorithms are presented in section 4.4. In section 4.5 we present numerical examples and finally we conclude in section 4.6.

## **4.2 Connection Management in WDM networks**

The implementation of wavelength-routed networks requires the implementation of control and management protocols to perform routing and wavelength assignment functions, as well as to exchange signaling information and to reserve resources [5]. A control mechanism can be either centralized or distributed. In a centralized approach, a central controller in the network maintains a database,

which contains information such as the state of existing connections, the network topology, and the wavelength usage on each link. When a connection request arrives, the controller selects a route and a wavelength based on the information contained in the database, and notifies each node along the selected route to setup this new connection by cross connecting the appropriate switch on the corresponding wavelength. Here, no coordination is required between nodes in the networks, and the central controller can make optimal routing and wavelength assignment decisions. However, the drawback of this centralized approach is the scalability as the network size becomes larger and the traffic demand becomes higher; hence, distributed routing algorithms are proposed. In a distributed approach, each node maintains partial or full information about the network-wide state and existing connections. This information is exchanged and updated if necessary among the nodes to reflect changes in the network state. Upon the arrival of a connection request, a route and a wavelength are selected locally at the specific node. Different route selection criteria can be used and these include fixed shortest path routing, fixed alternate routing and adaptive routing. In fixed shortest path routing, the same fixed route is always used for any given source-destination (s-d) pair. In fixed alternate routing, the route is chosen from a fixed set of predetermined static routes. In adaptive routing, the route is determined dynamically at the connection setup time, based on either global or local information. Link-state routing and distance-vector routing are two examples of adaptive routing [5, 6, 7].

When the connection is established in a distributed manner, control messages traveling from the source to the destination (forward reservation) are used to reserve

resources along the connection route; similarly, control messages traveling from the destination back to the source node (backward reservation) can also be used. In distributed approaches where global information is maintained, the source node will have enough information on the wavelengths availability on each link along the selected route; thus the source node may choose the desired wavelength and make the resource reservation in the forward direction. However, if only local information (or no information) is maintained at each node, then the source node may be unable to choose an appropriate wavelength for the entire path. One option is to use forward reservation where the control message traveling along the selected route will reserve all available wavelengths on each link along the route. Once the control message is received by the destination, the latter selects one wavelength from the wavelengths reserved along the entire path and releases the remaining wavelengths. The drawback of this approach is that network resources are being over reserved for a short period of time, which may lead to the blocking of subsequent connection requests and lower network utilization. Backward reservation is an alternative approach for reserving resources along a route. Here, the source node sends a control message(s) to the destination without reserving any resources. The control message(s) will collect information about wavelength usage along one or more paths, and the destination, using this information, will then decide on a route and a wavelength. Subsequently, the destination sends a reservation message to the source node to reserve the appropriate network resources.

Connection management protocols for WDM networks have been extensively studied. In [5], a management scheme based on the Link State Protocol is proposed,

where each node maintains the complete network topology and state information. Upon the arrival of a connection request, the source node utilizes this information to select a route and a wavelength. The authors of [6] presented a distributed routing approach in which a routing table is established by employing a distributed Bellman-Ford algorithm. For every incoming connection request the source node will choose a wavelength that results in the lowest cost to the destination. The authors compared the performance of their approach with the one in [5]. While the Link State Protocol has lower stabilizing time than the Bellman-Ford based distributed routing, the latter approach has a lower connection setup time, slightly lower blocking probability, and better scalability.

On the other hand, in networks with only local information [7, 8], distributed signaling protocols are needed to find end-to-end routes and reserve resources. Two types of signaling protocols have been investigated in the literature, namely forward and backward signaling protocols as mentioned previously [7]. Another alternative is to maintain global network connectivity information (made available to each node through the Link State Protocol, i.e. OSPF) where information about network resource usage is maintained locally. Here, upon the arrival of a connection request, the source node computes an explicit end-to-end path to the destination. Once the route has been found, the source initiates a *Probe* message containing an Explicit Route Object (*ERO*), which records the route that has been determined, and a Wavelength Set Object (*WSO*) which contains all available wavelengths on the outgoing link to the downstream node. The Probe message is then forwarded to the downstream node. Upon receiving the message, the intermediate node reads off the

*ERO* the next hop, generates a new *WSO* by intersecting the received *WSO* with the set of available wavelengths on the outgoing link. If the *WSO* is found empty, a *NACK* message is sent back to the source and the connection request is blocked. Otherwise, once the probe message is received at the destination, a wavelength is selected and a  $Resv(\lambda)$  message is sent on the backward path to configure the appropriate switches and update the corresponding databases. Here, due to the nature of this destination based wavelength selection, contention might occur while configuring the switches [9]. Two connections might contend for the same wavelength and a First Come First Serve (*FCFS*) scheme is employed to resolve this contention. A *NACK* message is sent upstream toward the source of the blocked connection and a *Release* message is sent downstream toward the destination to free any reserved resources.

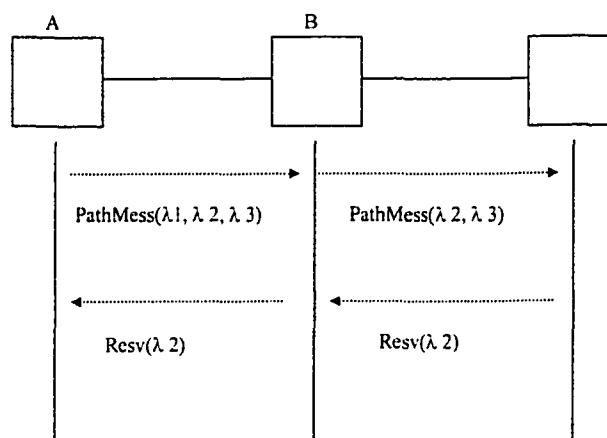


Fig 4.1 (a)

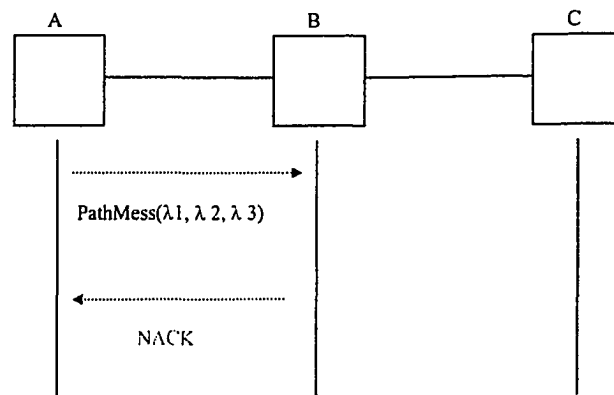


Fig. 4.1 (b)

Figure 4.1 shows an example of how a lightpath is set up in a WDM network using this approach. In Figure 4.1(a), a Probe message, or Path Reserve Message, is sent downstream toward the destination. B, upon receiving the message and reading off the next hop, generates a new WSO and forwards the message downstream. Edge node C, after recognizing itself as a destination, selects a wavelength and sends a Reserve message back to the source to configure the intermediate cross-connects. In Fig.4.1 (b), however, it is shown that the WSO is found empty and hence a NACK message is sent back to the source indicating the connection setup failure. On the other hand, in Fig. 4.1 (c), contention occurred while reserving the wavelength at node B, and a NACK message is forwarded upstream toward the source while a Release message is sent toward the destination to free the allocated resources.

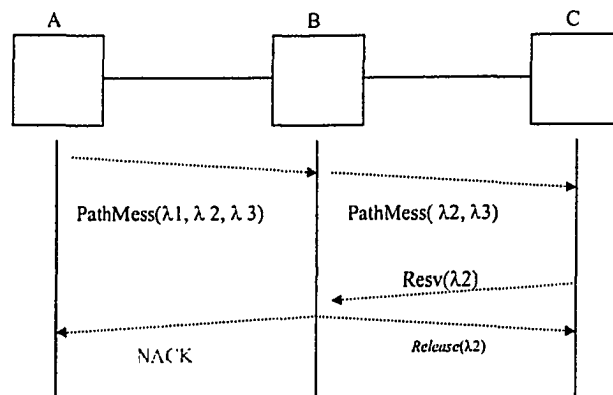


Fig 4.1 (c)

Figure 4.1- Lightpath setup in WDM networks

### 4.3 Connection management in survivable WDM mesh networks

#### 4.3.1- Background and motivations:

In WDM networks, where each fiber is carrying data in the order of terabits per second, service restorability [1] becomes a critical requirement for network planning and management. Service survivability stipulates that upon the failure of any network element, all affected connections be rerouted within a short time interval using spare capacity reserved on alternate paths. We focus on path protection in this section and we develop a distributed management protocol to setup working and protection paths and to determine resource sharability. In path protection, two alternately routed paths (working and protection) are provisioned for each connection request. Data is transmitted along the working path, while the protection path is used upon any network element failure affecting the working path. The protection path can be either dedicated (1:1, 1+1) to each connection or shared amongst multiple connections if their working paths do not fail simultaneously. In 1+1 protection, traffic is

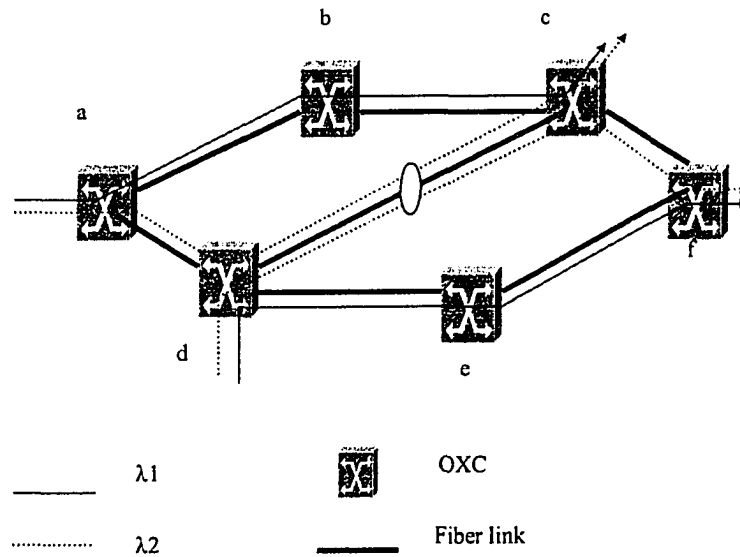
transmitted simultaneously on both working and protection paths. The destination receives data on the working path and in case of a failure it switches to the protection path. This type of recovery is very fast but resources are not used efficiently due to the inherent redundancy associated with the scheme. Meanwhile, in 1:1 path protection, both working and protection paths are provisioned simultaneously and data is transmitted along the working path while the protection path can be used to *transmit some low priority pre-emptable traffic*. In case of a failure along the working path, the source node is notified through signaling messages and data is transmitted on the protection path. As in 1+1, restoration in 1:1 is relatively fast but resources are not used efficiently. Shared path protection utilizes resources more efficiently [1-4]. Here, since backup resources are shared amongst many lightpaths, the switches on the protection path are not configured at the connection setup time; rather, upon a failure, the source node of the failed lightpath is notified and generates a signaling message to configure the switches along the protection path. This scheme obviously utilizes network resources more efficiently, however, it exhibits longer restoration times.

Provisioning working and protection paths and determining resource sharability in WDM networks requires management protocols. We propose a management protocol that is an extension of the protocols proposed in the previous section; it is a hybrid protocol that attempts to combine the best of both the link state and distributed-routing approaches. Specifically, the proposed hybrid approach combines Link State Protocol to disseminate and update information only about the physical connectivity of the network and a distributed local information-based signaling algorithm for connection and fault management. The purpose of using a hybrid

approach is two vantages: 1)- reducing the signaling overhead associated with the global information-based Link State Protocol by using a distributed approach where only local information is maintained at each node; 2)- eases the implementation of the routing protocol where physical constraints, such as link/node diversity, are imposed.

#### **4.3.2 – The network model:**

In the network modeled here, each node consists of an optical switch that can perform wavelength switching and an electronic controller that controls the optical switch [5]. The controller maintains global information of the physical connectivity of the network; it also maintains local information about wavelengths usage on the outgoing links of each node. Figure 4.2 shows a typical scenario of a wavelength routed WDM network architecture where 2 connections are setup between nodes “a-c”, and “d-f” respectively. Two diverse routes are setup for each connection (working and protection). A wavelength on a link can be in one of the following states: “0” if the wavelength is free, “1” if the wavelength is used (up) and “2” if it is reserved by a protection path. Each optical node is supposed to have enough transceivers so that blocking will be due only to insufficient wavelengths along the route or contention over the same wavelength.



Local Information Database

Connection-ID	In Port, $\lambda$	Out Port, $\lambda$	Type
1	1, $\lambda 2$	3, $\lambda 2$	Protection
2	0, $\lambda 1$	2, $\lambda 1$	working
2	0, $\lambda 2$	3, $\lambda 2$	protection

<Out Port, $\lambda$ >	<Connection ID, Physical route for working>
3, $\lambda 2$	1, <a, b, c> 2, <d, e, c>

Sharability Database

Figure 4.2- Network Architecture

### 4.3.3- Protocol description

The hybrid control and management protocol for provisioning working and protection paths and reserving resources and updating network state works as follows:

- 1) Each node in the network is required to maintain a routing table that contains an ordered list of a number of fixed shortest routes and their corresponding link-disjoint

paths to each destination node. Other than the static routing table, each node will maintain the following two databases: (a) a local information database that reflects the local resource usage at that node, e. g., the status of wavelength usage on its own outgoing links; and (b) a local sharing database that maintains information about the lightpaths whose backup paths traverse that node. This information is required in the case of shared protection to assist the signaling protocol to determine whether a wavelength on a link is shareable or not. Both the local information and sharing databases at node “d” are also shown in Fig. 4.2. As can be seen,  $\lambda_2$  on the outgoing port “2” is being shared by both connections. Note that the connection ID for each connection is stored along with the physical route of its working path. The routes  $\langle a, b, c \rangle$  and  $\langle d, e, f \rangle$  are the corresponding working paths for both connections protected by  $\lambda_2$  on link “d-c”; both routes are link-disjoint.

2) Upon the arrival of a connection request, the source node uses its routing table to select a route for the working path and a link-disjoint route for the protection route. If both paths are found, a distributed signaling protocol (to be described below) to simultaneously allocate resources along both routes is triggered. 3) If a link failure occurs, the link state protocol is triggered to disseminate information about this topological change (only the physical connectivity) to all nodes in the network so that future connections will not be miss-routed. Note that the link state protocol is only triggered upon a failure occurrence.

#### 4.3.3.1- 1:1 Dedicated Protection:

The algorithm used for the distributed signaling is a local information-based probing [10] in which the source node attempts to simultaneously establish the connection along each of the link-disjoint (working/protection) routes. Upon receiving a connection request, the source node sends a *Probe* message on each route toward the destination. This *Probe* message carries the set of free wavelengths along the first link of the path(s) to be established. When the next hop receives the message, it retrieves the received set of wavelengths from the message and intersects it with its own free wavelength set, and forwards the result to the next hop. If the final set is not empty, the last hop-router (destination) must pick one free wavelength from the resulting set, by using a wavelength assignment algorithm such as first-fit (used for both the working and protection paths), configures its local node, and sends a reserve message (*Reserve* ( $\lambda$ )) back to the previous hop with the chosen wavelength. Upon receiving a reservation message on a link  $L$ , the previous hop checks if the desired wavelength is available; if it is, configures its local node and passes the message to its previous hop, until the reservation message is received by the first hop (source node).

The *Reserve* message also updates the local information database maintained at each node by adding information about the new connection, such as connection ID, connection type (1 for working and 2 for protection), and changing the status of the reserved wavelength from “idle” to “up” if working or “reserved” if protection. After receiving both *Reserve* messages on both routes at the source node, the setup succeeded and data transmission starts on the working path. Note that, due to the nature of this destination based wavelength selection, contention might occur on one

(or both) path (s). In this case a *Release* message is sent back to the destination to free the allocated resources and a *NACK* message is sent to the source to indicate the setup failure. After receiving the *NACK* message from one path, the source node checks for the setup status along the second path. If succeeded, a *Release* message is sent to the destination along that path to free the allocated resources and the call is blocked. If not, the source keeps the received *NACK* message in case the setup succeeds on the second path.

Note also, while the connection request (*Probe*) moves from one node to another, the generated wavelength vector might be empty on both or either paths and the connection cannot be setup. The intermediate node drops the *Probe* packet and sends a *NACK* message back to the source indicating the setup failure; the *NACK* message is handled similar to the way described earlier in this section. Overall, a connection succeeds if both *Reserve* messages on both routes are received.

#### **4.3.3.2- Shared Path Protection:**

The basic signaling components described in the previous section for the 1:1 dedicated path protection is also applicable to the shared path protection. However, in shared protection, the complexity stems from the fact that a channel on a given link on the protection path of a given connection can be shared with other connections. Thus, the issue of determining the resource sharability is of paramount interest in this section.

In distributed routing where only local information about resource usage is maintained at each node, a sharability database maintained at each node is also needed to reflect the sharability information about local resources.

Upon the arrival of a connection request, the node selects a pair of paths to the destination, and then a distributed signaling protocol is triggered to allocate resources along both paths. The source node simultaneously probes both paths (*Primary\_Probe*, *Secondary\_Probe*) to determine the resource availability. Initially, *Primary\_Probe* contains: 1)- a vector of the available wavelengths on the first outgoing link on the primary path, 2)- the path on which the connection is routed, 3)- the connection identifier of the request. A wavelength is said to be available if it is not used by any other connection (primary or protection). Besides this information, *Secondary\_Probe* carries the physical route along which the primary path is routed. Here a wavelength is said to be available if it is not used by any connection or if it is a protection sharable wavelength; the sharability of a wavelength is determined locally at the node by checking its sharability database whether the new connection can use this wavelength without violating the sharability condition (i.e., all primary paths protected by this wavelength have to be link-disjoint). Each probe message is then forwarded along its corresponding path to the downstream node. Upon receiving a probe message, an intermediate node determines the next hop along the path and computes the set of available wavelengths on the outgoing link. This set is then intersected with the received vector in the probe message and a new vector comprising only the common available wavelengths on the previous links is attached to the probe message and then forwarded to the next hop until it reaches the

destination. If the new vector happens to be empty (i.e., no common wavelengths are available), the node then drops the probe message and sends a negative acknowledgment (*Primary\_NACK* or *Secondary\_NACK*) to the source to indicate the failure of the path setup. At the destination, when a probe message is received, a wavelength is selected and a reservation message is sent back to the source node on the same path to allocate the appropriate resources. Note that, due to the distributed nature of this protocol, contention might occur while allocating resources in the backward direction. In this case, a node detecting a contention must release whatever resources have been reserved and send a *Primary\_NACK* (or *Secondary\_NACK*) to the source indicating the failure of the path setup.

While allocating resources, we differentiate between resource allocation on the primary path and the protection path. A *Primary\_Reserve* message will reserve the corresponding wavelength and cross-connect the switch at the corresponding node. The message also will update the local information database maintained at the node to reflect this change in the local resource usage. On the other hand, a *Secondary\_Reserve* message will only reserve the wavelength without changing the status of the switch fabric and it will update the local information database as well as the local sharability database (see figure 1). Finally, upon receiving both reserve messages, the source deems the connection request successful and begins to transmit data on the primary path. Otherwise, if at least one negative acknowledgment message is received, the connection setup fails and any reserved resources must be released.

Tearing down connections, on the other hand, requires also signaling protocols to de-allocate any reserved resources and modify the local and sharability database at each node along the path. Here, *Primary\_Release* and *Secondary\_Release* messages are transmitted by the source on both paths and forwarded towards the destination. Each node receiving a *Primary\_Release* message will free the reserved wavelength and update this change into the local information database, and then forward the message downstream towards the destination. On the other hand, a node, upon receiving the *Secondary\_Release* message, first checks the local sharability database. If this connection is the only one using this wavelength, then the wavelength is returned to the Idle status and the local information as well as sharability databases are modified accordingly. Otherwise, if other connections are protected by the same wavelength, then the node searches the sharability database for the connection ID of the lightpath being released and it takes off from its database this entry and forwards the message downstream without any further actions. The destination will drop both messages once they are received.

#### **4.3.4- Wavelength assignment**

Wavelength assignments are required in WDM networks when provisioning lightpaths [11]. In this work, we assume the wavelength continuity constraint, i.e. the same wavelength has to be assigned on all links along the path of a connection. The first fit (FF) [11] wavelength assignment is used for assigning wavelengths along the primary path. However, on the protection path we propose a new assignment scheme that aims into locally maximizing the usage of shared protection wavelengths, thus

improving the overall resource utilization. Here, for each element in the wavelength vector attached to the *Secondary\_Probe* message, we append a *counter* that reflects the number of hops along the route where this wavelength is found sharable. Initially at the source node, the wavelength, which is found sharable, has its counter initialized to 1 whereas all other available wavelengths have their counters initialized to 0. Each intermediate node, upon receiving the probe message, will increment the counter associated with a specific wavelength if this wavelength is found sharable; if not, then the counter keeps its old value. In this approach, the control message collects information regarding the sharability condition of each wavelength along the route and then the most shared wavelength is selected at the destination. To better illustrate this approach, let  $\vec{V}_\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$  and  $\vec{V}_c = \{c^{\lambda_1}, c^{\lambda_2}, \dots, c^{\lambda_m}\}$  be the received wavelength vector and wavelength counter at the destination respectively. Let  $P = \{l_1, l_2, \dots, l_k\}$  be the set of links along the protection path. Then, an element  $c^{\lambda_j}$  will take the following value:

$$c^{\lambda_j} = \sum_{i=1}^k \phi_i^{\lambda_j} \quad (1)$$

where:

$$\phi_i^{\lambda_j} = \begin{cases} 0 & \text{if } \lambda_j \text{ is not sharable on link } l_i \\ 1 & \text{otherwise} \end{cases}$$

A wavelength  $\lambda_n$  is selected such that:

$$n = \max_{j=1 \dots m} (c^{\lambda_j}) \quad (2)$$

We refer to this assignment scheme as the Most Shared Wavelength (MSW) assignment. This assignment scheme favors the allocation of protection wavelengths that are already reserved, while deferring the idle wavelengths for other use (e.g., primary path allocation for a different connection). Note that, the tradeoff here is the signaling overhead required from the signaling protocols.

#### **4.4 Fault management and fast restoration algorithms**

Rapid restoration algorithms that minimize the service interruption time are critical for the design of a reliable and cost effective network. Recovering from any network element failure requires failure detection and notification mechanisms, and signaling protocols for configuring the intermediate nodes along the pre-planned restoration paths.

In this section, we focus on the signaling process required for configuring the nodes along the pre-planned restoration paths in a shared mesh network. As mentioned before, resources are reserved during the provisioning phase; however cross-connects along the restoration paths are not configured since switch ports and wavelengths can be shared among multiple connections to achieve better network resource utilization. Hence, the restoration procedure usually involves sending recovery messages along the restoration path and configuring the optical cross connects along the path. This in turn will result in delaying the recovery process and increasing the service interruption time. Thus, we propose a novel restoration signaling, that minimizes the service interruption time in a shared mesh-based framework. The significant contribution of the new algorithm is that the connection restoration time is

independent of the restoration path length (i.e., eliminating the propagation delay) and also it is independent from the accumulation of the switch configuration times. Next, we present two basic restoration signaling, namely the source based and destination based restoration, then we propose our rapid restoration algorithm.

#### 4.4.1- Source-based restoration (SBR):

In SBR, upon detecting a failure, the source node of each affected lightpath is notified through a Failure Notification Message (*FNM*) and it initiates the protection path setup process by generating a Failure Recovery Message (*FRM*) to be transmitted along the protection path. Each intermediate node receiving the *FRM* cross-connects the corresponding switch and forwards the message to its downstream node. After receiving the *FRM*, the destination sends a *Recovery-ACK* message to the source to indicate the beginning of data transmission. Figure 4.3 shows a sample example where SBR is employed. Upon detecting the failure of link b-c, the node detecting the failure (each link is carrying two fibers one in each direction, thus any end-node of the link can detect the failure) sends a *FNM* towards the source node. Node “a”, after receiving the notification message, initiates the protection path setup by sending a *FRM* message towards the destination on the protection path. Each intermediate node receiving this message cross-connects its switch and forwards the message to its downstream node. At the destination “f”, an *ACK* message is sent back towards the source and the traffic is switched for transmission on the protection path.

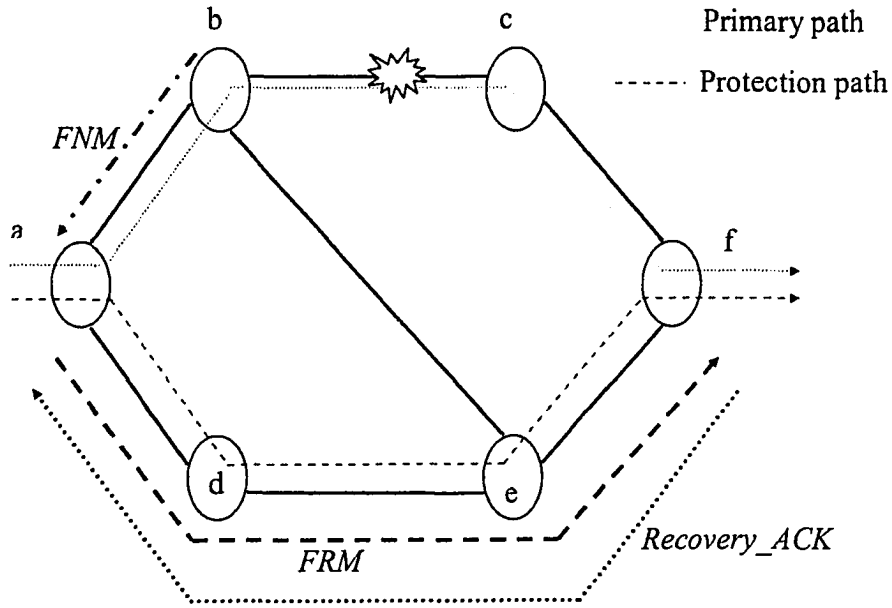


Figure 4.3- Failure Detection/Notification with SBR

The overall restoration time is estimated as follows:

$$RT^{SBR} = T_d + T_N + T_1 + T_2 \tag{3}$$

Where,  $T_d$  and  $T_N$  are the failure detection time and the failure notification time (time it takes to notify the source of the failed lightpath), respectively:

$$T_N = \sum_{i=1, j=1}^m (T_p^i + T_D^j) , \text{ where } T_p^i \text{ is the message processing time at node } i \text{ and } T_D^j \text{ is the}$$

propagation delay on fiber  $j$ , and  $m$  is the number of hops between the node detecting the failure and the source.

$$T_1 = \sum_{i=1, j=1}^{n-1} (T_p^i + T_{sc}^i + T_D^j) \text{ is the time it takes the control message to setup all switches}$$

on the backup path, where  $T_{sc}^i$  is the switch configuration time at node  $i$  and  $n$  is number of hops on the protection path.

$T_2 = \sum_{i=1, j=1}^{n-1} (T_p^i + T_D^j)$  is the time it takes the ACK message to travel back to the source.

#### 4.4.2- Destination-based restoration (DBR):

In DBR, the destination node of each affected connection is notified and it initiates the process of configuring the protection path. Here, the destination node sends a FRM towards the source on the protection path to configure the intermediate switches. The source, upon receiving this message, switches the traffic from the affected lightpath onto the protection path, see figure 4.4.

Note that this approach improves the restoration time when compared to SBR by eliminating the one-way propagation delay. The restoration time is then measured as follows:

$$RT^{DBR} = T_d + T_N + T_3 \quad (4)$$

$T_d$  and  $T_N$  are the failure detection time and the failure notification time (time it takes to notify the destination of the failed lightpath), respectively:

$$T_N = \sum_{i=1, j=1}^m (T_p^i + T_D^j), \text{ where } T_D^j \text{ is the propagation delay on fiber } j, \text{ and } m \text{ is the number}$$

of hops between the node detecting the failure and the destination.

$$T_3 = \sum_{i=1, j=1}^n (T_p^i + T_{sc}^i + T_D^j) \text{ is the time it takes the control message to travel towards the}$$

source and setup all switches on the backup path.

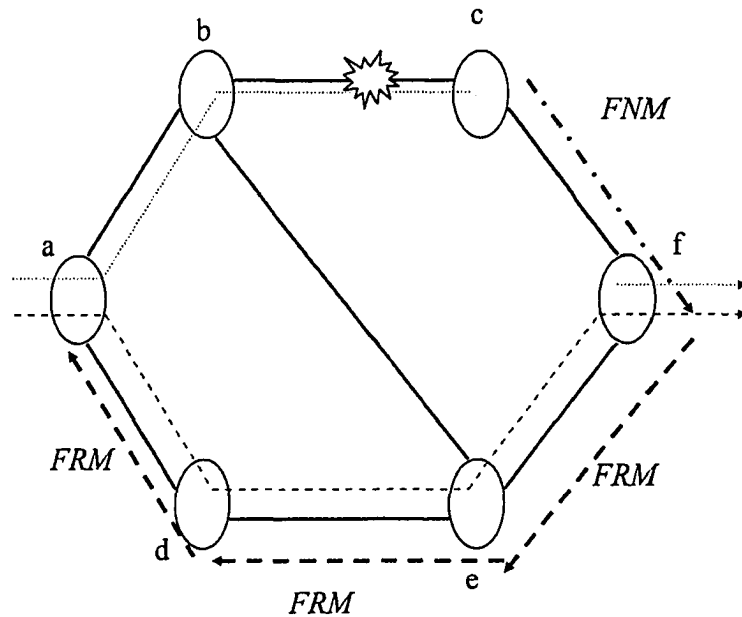


Figure 4.4-Failure Detection/Notification with DBR

#### 4.4.3- Enhanced SBR:

Given that backup resources are reserved a-priori, they only need to be configured upon the failure detection. Here, while resources are being configured, contention is unlikely to occur; thus, the source node of each failed connection does not need to wait for an ACK from the destination to start its data transmission. Rather, and unlike the 2-way messaging protocol described in section 4.4.1, the source needs to properly compute an offset time (guard time) after which data transmission can take place. While the offset time should be selected such that the intermediate switches along the path are configured before data arrival (i.e. large enough), the offset time should also be small enough to guarantee rapid restoration. A straightforward method to determine the offset time is to compute the time the control message takes until the last switch along the protection path is configured. Although

this approach does not rely on receiving an ACK message to start the transmission, the restoration time still depends on the forward propagation delay. This can be further improved if the offset time is chosen independent of the propagation delay; here, similar to the JIT proposed in [12], the source computes just enough time after which it starts data transmission. This time should be chosen such that any intermediate switch has its cross-connect (XC) configured to provide a fast cut-through for the data right before its arrival. A reasonable value will be the sum of the message processing times, the XC switching times (SCT) plus some small value epsilon to ensure that the destination has its XC configured by the time the data is received, thereby eliminating the impact of the message propagation delay along the protection path.

The guard time is then:

$$T_g^{(1)} = \sum_{i=1}^{n-1} (T_p^i + T_{sc}^i) + \varepsilon \quad (5)$$

where  $n$  is the number of hops along the backup path,  $T_p^i$  and  $T_{sc}^i$  are the message processing time and the switch configuration time at node  $i$ , respectively. The restoration time is then measured as follows:

$$RT^{(1)} = T_d + T_N + T_g^{(1)} \quad (6)$$

where,  $T_d$  and  $T_N$  are the failure detection time and the failure notification time (time it takes to notify the source of the failed lightpath), respectively.  $T_N = \sum_{i=1, j=1}^m (T_p^i + T_D^j)$

where  $T_D^j$  is the propagation delay on fiber  $j$ , and  $m$  is the number of hops between the node detecting the failure and the source.

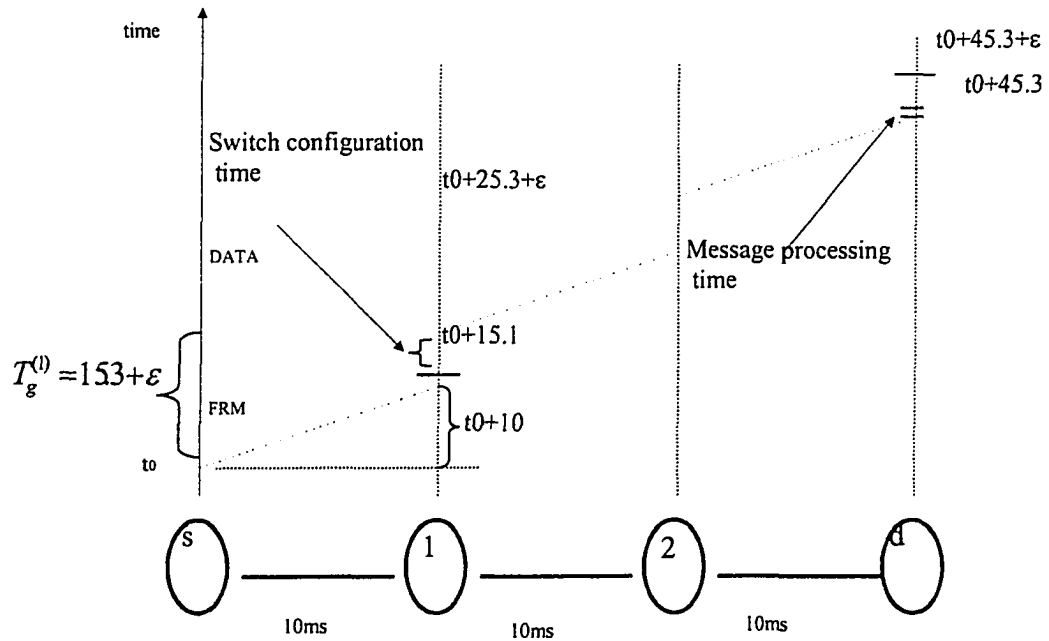


Figure 4.5- SBR with offset time based restoration

Figure 4.5 illustrates how the guard time based restoration protocol works. Here, the source of the failed lightpath is notified at time  $t_0$ . The control message is then forwarded along the backup path (s-1-2-d) to configure the intermediate switches and an offset time is computed after which the source resumes its data transmission. Assuming the propagation delay on each fiber to be 10 ms, the switch configuration time would be 5 ms and the message processing time would be 100  $\mu$ s. Then the control message is received at node 1 at time  $t_0+10$ , the local switch is configured at time  $t_0+15.1$  and the message leaves the node at  $t_0+15.1$ ; the message keeps propagating towards the destination. At  $t_0+40.2$  the control message is received by the destination and the switch is configured at  $t_0+45.3$ . Now let us follow the transmission of data. According to Eq.5, after an offset time =  $3*(0.1+5) + \epsilon = 15.3 +$

$\epsilon$  ms, the source resumes transmission. At  $t_0+25.3+\epsilon$ , node 1 had already configured its switch ( $t_0+15.1$ ) and thus data is switched immediately to the next hop until it reaches its destination. At the destination, data is received at  $t_0+45.3+\epsilon$  where the local switch has been configured at  $t_0+45.3$ .

Clearly, this approach eliminates the impact of the message propagation delay on the overall restoration time. However, the restoration time is still dependent on the accumulation effect of the SCT, see Eq.5. This could lead to slower restoration times if the SCT is large [1] ( $\sim$  tens of ms) and could eventually deteriorate the performance of the restoration protocol; thus one needs to choose a guard time independent of the SCT accumulation effect as well. Pipelining technique was proposed in [12] to study the signaling for optical burst switching networks and it is an effective technique to speed up the restoration process when the switch configuration time is considerably large as shown in [1]. Here, upon receiving the control message, the intermediate node forwards the message immediately to its downstream node and simultaneously starts configuring its local switch. This way the source node only needs to wait an offset time that is equal to the sum of the message processing times plus one switching time then it starts its transmission; thereby eliminating the cumulative switching effect along the protection path:

$$T_g^{(2)} = \sum_{i=1}^{n-1} (T_p^i) + T_{sc} + \epsilon \quad (7)$$

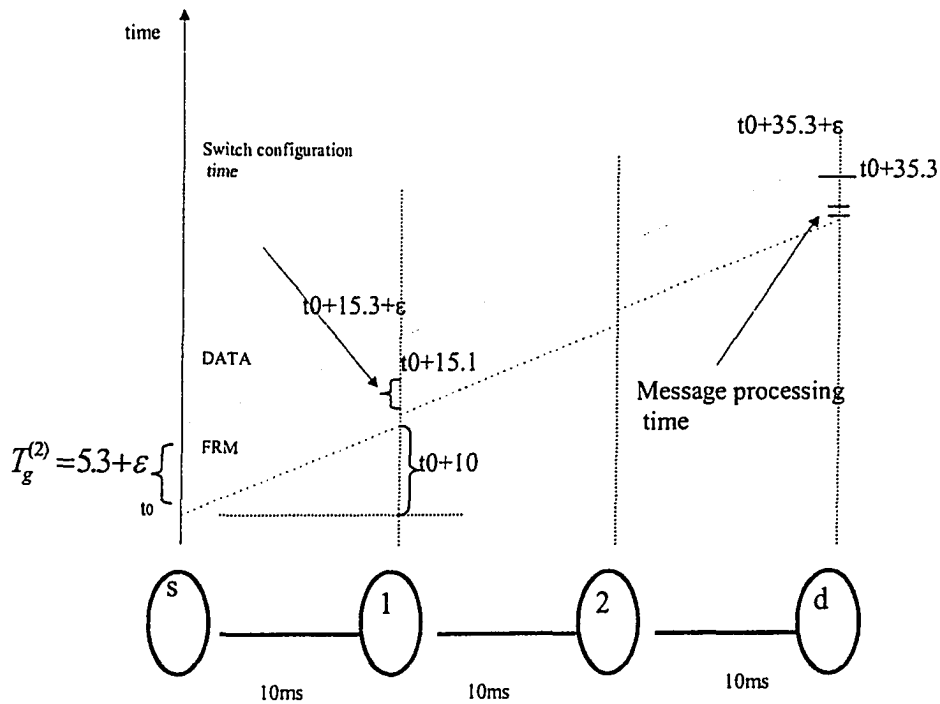


Figure 4.6- SBR with enhanced offset time based restoration

Figure 4.6 depicts the effectiveness of the use of guard time and pipelining techniques in achieving rapid service restoration. Here, at time  $t_0 + 35.3 + \epsilon$ , the connection is entirely restored and traffic is switched to the backup path; conversely, using the first approach (Fig. 4.5), at  $t_0 + 45.3 + \epsilon$  the connection is restored. This shows that pipelining reduces the restoration time by 10ms.

#### 4.5 Performance Evaluation

The performance of the proposed hybrid connection management protocol and fast restoration algorithm is evaluated via simulation of the mesh-based 16 nodes NSF network shown in figure 4.7. The numbers on the links represent link distances in tens of kilometers. We evaluate the performance of our protocol using the following metrics: blocking probability, restoration time, and data loss.

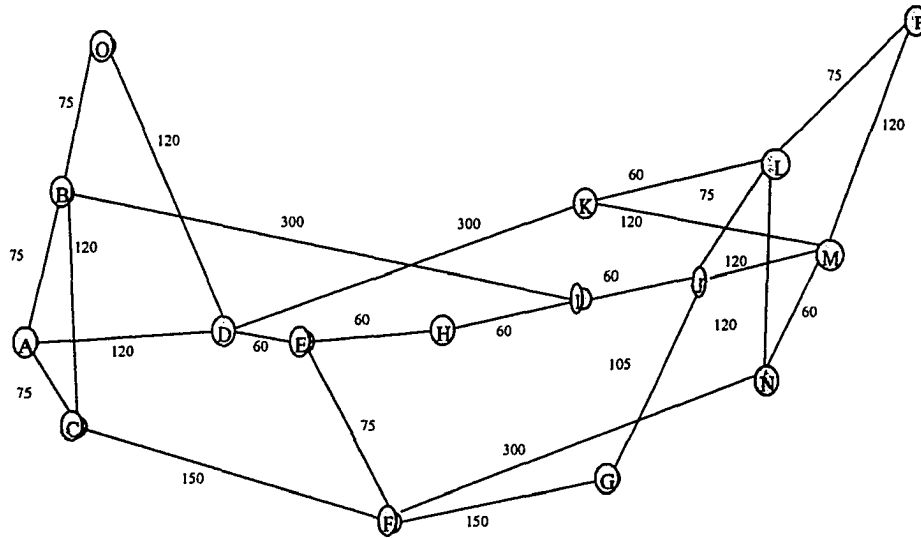


Figure 4.7- the 16-nodes NSF network

Following are the parameters used in our simulation:

- 1)- Connection requests arrive at each node as a Poisson process with mean arrival rate  $\lambda$  arrivals/*ms*.
- 2)- Connection requests are uniformly distributed among all source-destination pairs.
- 3)- The connection-holding time is exponentially distributed with mean  $1/\mu = 100$  ms.
- 4)- The number of wavelengths on each link,  $W = 8$ .
- 5)- The message processing time is assumed to be  $100\mu\text{s}$ .
- 6)- The switch configuration time is  $0.5$  ms.
- 7)-  $\varepsilon = 50\mu\text{s}$ .
- 8)- Fault detection time  $T_d = 100\mu\text{s}$ .

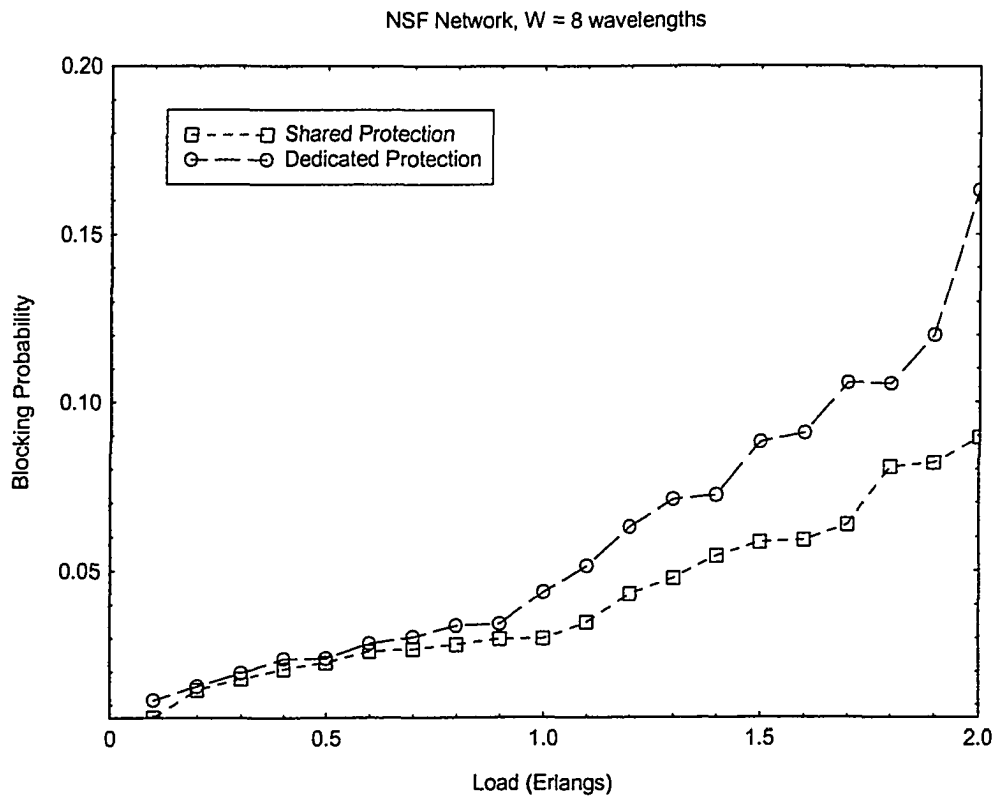


Figure 4.8- Blocking probability vs. load for dedicated and shared protection

Figure 4.8 shows the blocking probability vs. load for both dedicated and shared path protection. The figure shows that sharing backup resources is a versatile method that improves the network performance by eliminating network redundancy. The First Fit (FF) wavelength selection is used to assign wavelengths on both working and protection paths for both and shared protection. The figure shows that when the load is lighter, both and shared protection perform similarly; this is due to the fact that the network has sufficient resources to satisfy almost all connections. On the other hand, at higher load, shared path protection is able to manage resources more efficiently by sharing backup resources.

Next we compare the performance of different wavelength assignments. Figure 4.9 shows a comparison between the FF and MSW wavelength assignments. As mentioned previously, the MSW locally maximizes the use of sharable wavelengths, in which the selection of the most sharable wavelength along a path is favored. The figure shows that MSW performs slightly better than FF. However, as the load increases, both selection schemes perform similarly. This is due to the fact that blocking at higher loads is due to insufficient resources rather than poor resource management. Note also that MSW selection criterion requires more information to be propagated, leading to additional overhead on the management protocol

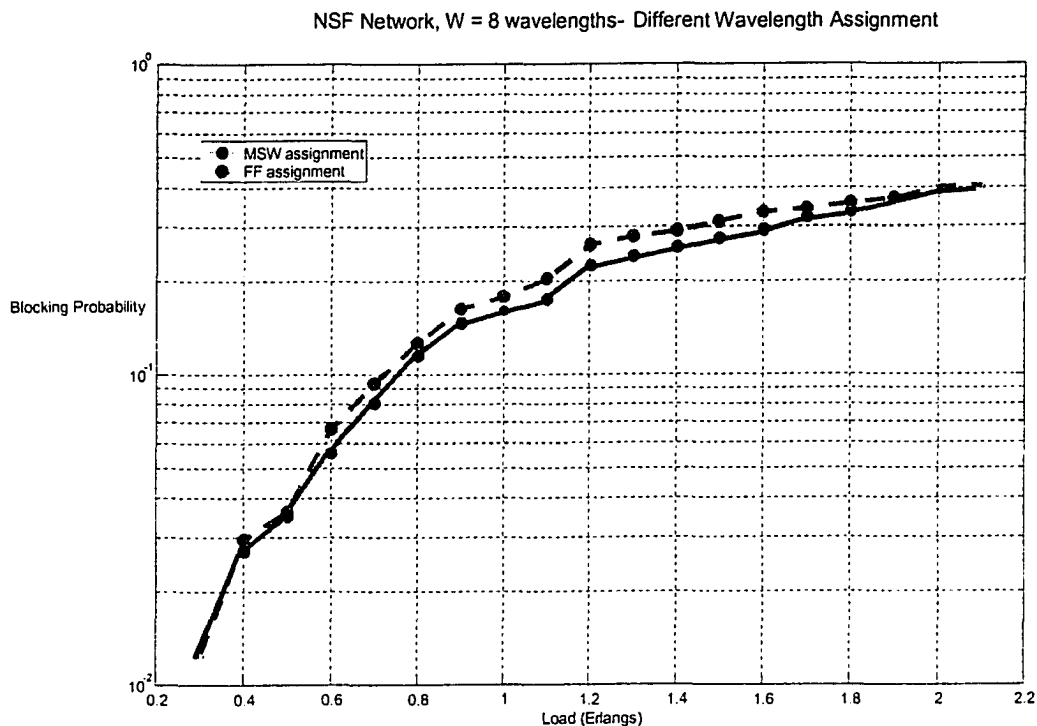


Figure 4.9- First Fit vs. MSW wavelength assignments

Connection setup time (CST) is another key comparison between both protection schemes. The higher setup delay in dedicated protection comes from the

switch configuration time and propagation delay of signaling messages when setting up backup paths. On the other hand, setting up a backup path in shared protection doesn't involve the cross-connect configuration. Rather switches are configured after a network element failure. This results in reduced connection setup delays as shown in figure 4.10. Note that at higher loads, the CST decreases, this is due to the fact that longer paths at higher loads tend to be more blocked.

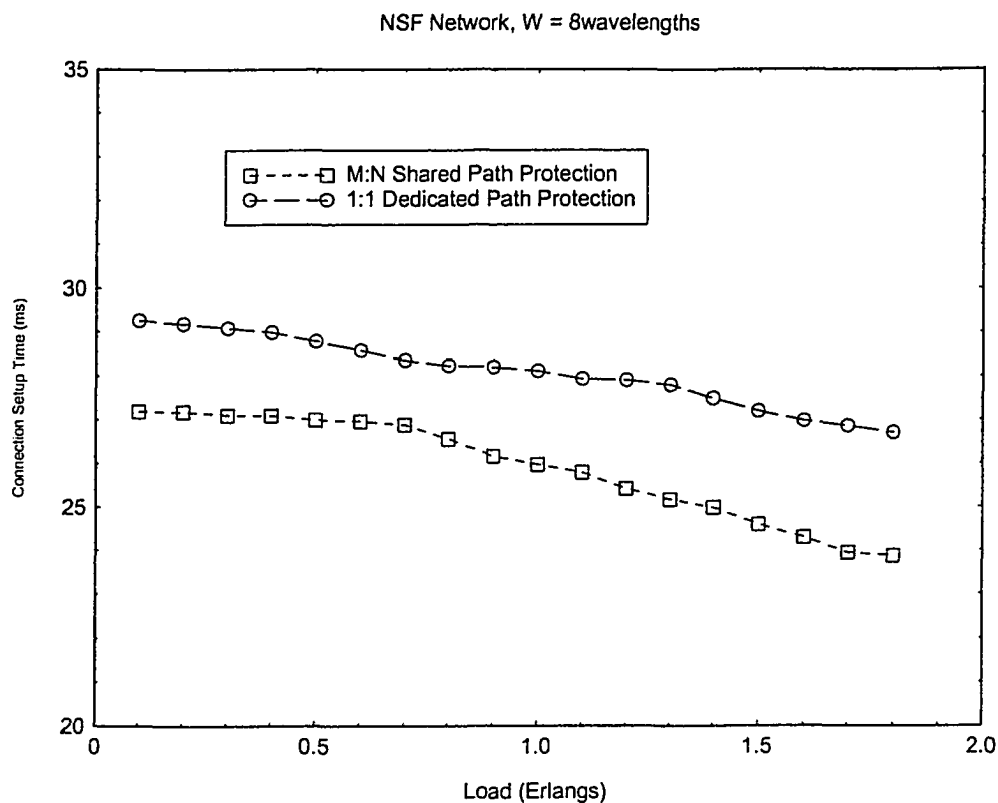


Figure 4.10- CST for dedicated (1:1) and shared protection vs. Load

Next we compare different restoration algorithms (RAs). We compare the conventional SBR, (RA1), and DBR (RA2) and SBR with the straightforward offset time (RA3) and finally SBR with offset time and pipelining, (RA4). The metric of comparison is the restoration time. Here, after routing all connections (at a particular

load) in the network, at random we take down one link and we run the restoration algorithm to restore the affected connections; we repeat the same experiment with the same seed using other algorithms.

Figure 4.11 shows that RA3 and RA4 considerably reduce the restoration times when compared to RA1 (by eliminating a round trip delay of the control messages along the protection path) and RA2. Note also that RA4 slightly outperforms RA3 given that  $T_g^{(2)} < T_g^{(1)}$  and that the switch configuration time is small and its accumulation effect is relatively minimal.

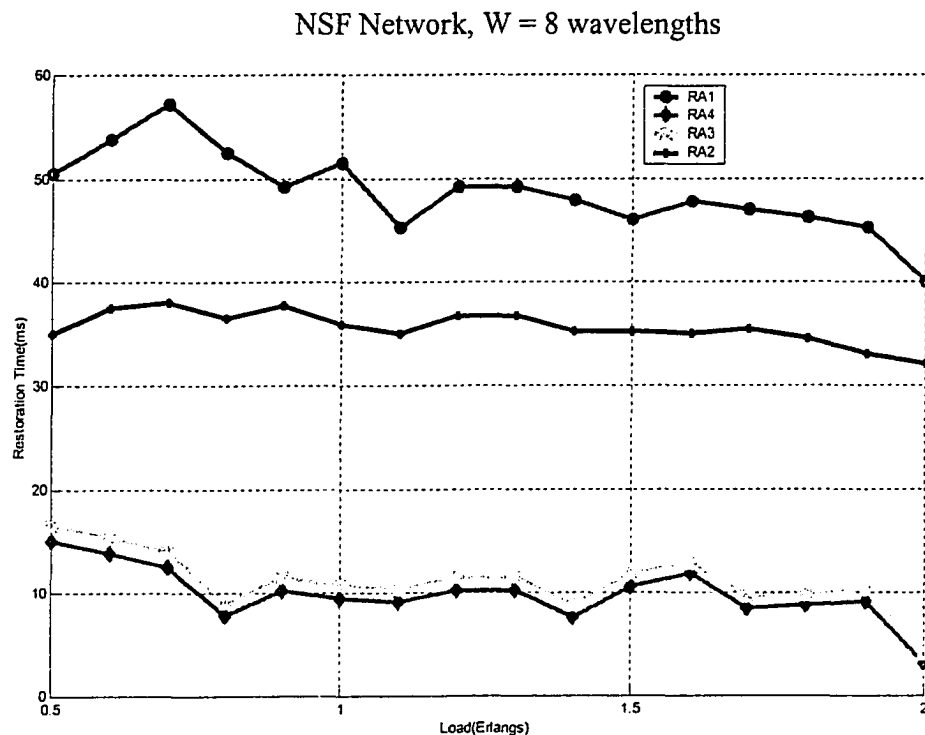


Figure 4.11- Restoration time vs. Load for different restoration algorithms

The accumulation of the SCT along the backup path could have a major impact on the restoration time if the SCT is fairly large. RA4 eliminates this effect by deploying pipelining techniques and, as shown in equation 7, the guard time only

depends on one SCT. However, RA1 and RA3 suffer from this accumulation. To see clearly the impact of the accumulation of the SCT for the three restoration algorithms, we need to look at the RT vs. SCT plot for each algorithm. Theoretically, one can find the slope (i.e. impact of the SCT accumulation effect on the RT) of each plot as follows:

For RA1:

$$\begin{aligned}
 RT^{SBR} &= T_d + T_N + T_1 + T_2 \\
 &= T_d + \sum_{i=1, j=1}^m (T_p^i + T_D^j) + \sum_{i=1, j=1}^{n-1} (T_p^i + T_{SC}^i + T_D^j) + \sum_{i=1, j=1}^{n-1} (T_p^i + T_D^j) \\
 &= T_d + \sum_{i=1, j=1}^m (T_p^i + T_D^j) + \sum_{i=1, j=1}^{n-1} (T_p^i + T_D^j) + \sum_{i=1, j=1}^{n-1} (T_p^i + T_D^j) + (n-1)T_{SC} \quad \text{given that}
 \end{aligned}$$

$$T_{SC}^i = T_{SC}, \forall i$$

The slope then becomes:

$$\frac{\partial RT^{SBR}}{\partial T_{SC}} = n-1 \quad (8)$$

Similarly, for RA3 and RA4:

$$\begin{aligned}
 RT^{(2)} &= T_d + T_N + T_g^{(2)} \\
 &= T_d + T_N + \sum_{i=1}^{n-1} (T_p^i) + T_{sc} + \varepsilon
 \end{aligned}$$

$$\frac{\partial RT^{(2)}}{\partial T_{SC}} = 1 \quad (9)$$

$$\begin{aligned}
 RT^{(1)} &= T_d + T_N + T_g^{(1)} \\
 &= T_d + T_N + \sum_{i=1}^{n-1} (T_p^i + T_{sc}^i) + \varepsilon
 \end{aligned}$$

$$= T_d + T_N + \sum_{i=1}^{n-1} (T_p^i) + \varepsilon + (n-1)T_{SC}, \text{ given that } T_{SC}^i = T_{SC}, \forall i$$

$$\frac{\partial RT^{(1)}}{\partial T_{SC}} = n-1 \quad (10)$$

Equations 8 and 10 show that the increase in the restoration time vs. the cross connects switching time depends on the accumulation of the SCT along the protection path (RT  $\propto$  number of hops on the protection path); alternatively, RA2 is not affected by this accumulation, and the RT only depends on one value of the SCT, Eq.9.

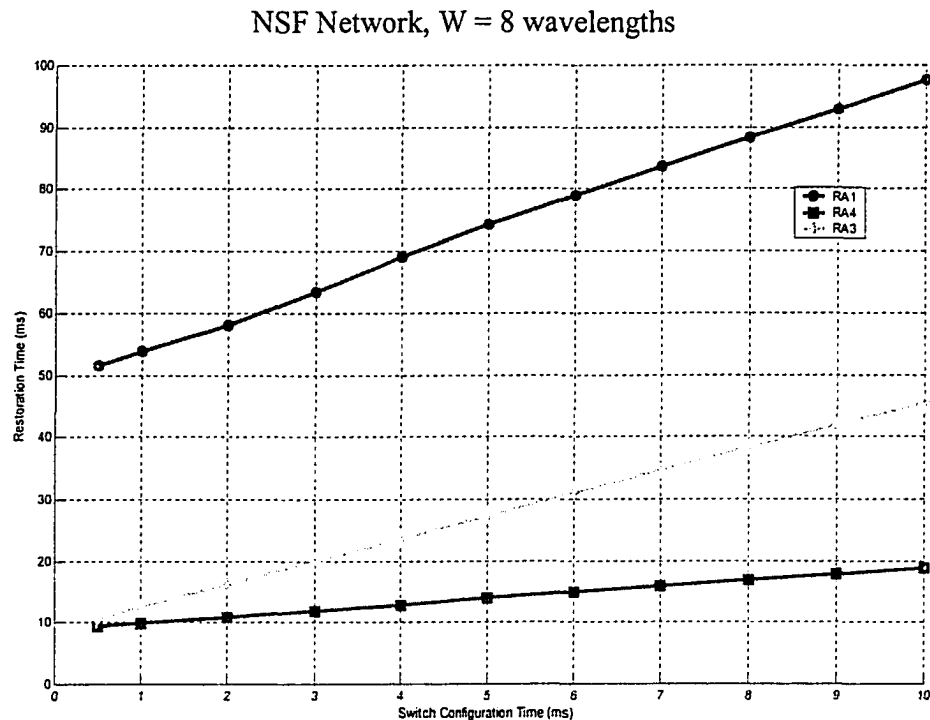


Figure 4.12- Impact of the SCT on the RT.

This impact of the switch configuration time (SCT) on the restoration time (RT) is presented in figure 4.12.

Figure 4.12 shows the simulation results of the increase in the restoration time when the SCT increases. However, this impact is negligible on RA4. From the figure, it is obvious that the slope of the RT for RA4 is very close to one; however, it is equal to almost  $n-1$  ( $n$  is number of hops along the backup path) in RA1 and RA3. This means that longer backup paths become unfavorable in RA1 and RA3 if SCT is large. Figure 4.12 shows that if  $SCT = 10$  ms, RA1 and RA3 achieve a restoration time of 98ms and 45 ms, respectively, whereas RA4 keeps this value minimal, 19ms.

#### 4.6 Conclusion

In this chapter we presented a novel hybrid distributed fault-management protocol for combating Single-fiber failures in mesh-based WDM optical networks. The proposed hybrid approach combines Link State Protocol to disseminate and update information only about the physical connectivity of the network and a distributed local information-based signaling algorithm for connection management. We have evaluated the performance of the proposed hybrid approach by comparing both the dedicated 1:1 path protection and the shared path protection in terms of blocking probability, connection setup time and restoration time under failure assumptions. Also, we presented different restoration algorithms as a key requirement for designing shared restorable mesh networks. We proposed two restoration algorithms, namely the source based restoration (SBR) and the destination based restoration (DBR). We also proposed a new method to achieve rapid restoration. This method builds upon the SBR by introducing offset time after which the source can start transmitting data on the protection path. Efficient offset time selection combined

with pipelining techniques was shown to achieve rapid restoration that is independent on the protection path length.

#### 4.7 References:

- [1] M. Goyal, G. Li, and J. Yates, "Shared Mesh Restoration: A Simulation Study", OFC'02, March 2002.
- [2] B. Doshi et al, "Optical Network Design and Restoration" Bell Labs Tech. J., vol.4, no.1, Jan-Mar 1999, pp. 58-84
- [3] G. Austin, et. al., "Fast, Scalable, and Distributed Restoration in General Mesh Optical Networks" Bell Labs. Tech. Journal, vol. 6, nb. 1, pp. 67-81, 2001, Jan-June 2001.
- [4] C. Assi, Y. Ye, A. Shami, S. Dixit, and M. Ali, "Performance Evaluation of Efficient Path Selection and Fast Restoration Algorithms for Shared Restorable Mesh WDM Networks" Proc. IEEE ICC, Anchorage, Alaska, 2002.
- [5] R. Ramaswami and A. Segall, "Distributed Network Control for Wavelength Routed Optical Networks," Proceedings, IEEE INFOCOM '96, San Francisco, CA, pp. 138-147, Mar. 1996.
- [6] Hui Zang et al, "Dynamic lightpath establishment in wavelength-routed WDM networks," IEEE Comm. Magazine, Vol. 39 No. 9, PP 100-108, Sept. 2001.
- [7] X. Yuan, R. Melhem, R. Gupta, Y. Mei, and C. Qiao, "Distributed Control Protocols for Wavelength Reservation and Their Performance Evaluation," Photonic Network Communications, vol. 1, no. 3, 1999, pp. 207-218.
- [8] J. P. Jue and G. Xiao, "An Adaptive Routing Algorithm with Distributed Control Scheme for Wavelength-Routed Optical Networks," proceedings, Ninth International Conference On Computer Communications and Networks, Las Vegas, NV, Oct. 2000.

- [9] I. Widjaja, A. Elwalid, " Study of GMPLS Lightpath Setup over lambda-Router Networks ", Proc. of IEEE, ICC'02.
- [10] C. Assi, Y. Ye, A. Shami, S. Dixit, and M. Ali, "A Hybrid Distributed Fault-Management Protocol for Combating Single-Fiber Failures in Mesh-based DWDM Optical Networks" IEEE GLOBECOM 2002, Taipei, Taiwan.
- [11] C. Assi, A. Shami, and M. Ali, "Optical networking and real-time provisioning; An integrated vision for the next-generation Internet", IEEE Network, July/Aug. 2001, Vol. 15 No. 4, PP. 36-45.
- [12] J.Y. Wei, R.I. McFarland, "Just-in-time signaling for WDM optical burst witching networks", J. of Lightwave Technology, Vol. 18 no. 12, pp. 2019-2037, Dec. 2000.

## Chapter 5

# 5. Real-Time Provisioning in IP-Over-WDM Networks

### 5.1- Introduction

The evolving next generation Internet transport infrastructure is moving towards a model of high-speed routers interconnected by intelligent, re-configurable optical core networks that will directly provide a global transport infrastructure for legacy and new Internet Protocol (IP) services [1,2]. In this model, clients (e.g., IP/MPLS routers) are attached to an optical core network, and connected to their peers over dynamically switched optical paths (lightpaths) spanning potentially multiple optical cross-connects OXCs (figure 1). The optical core network consists of multiple optical cross-connects interconnected by optical links in a general mesh topology. This transition is driven by the deployment of high-speed IP routers and ATM/MPLS switches. Specifically, these data-centric network elements (NEs) provide the essential bandwidth management functions including multiplexing up to OC-12, OC-48, OC-192, GbE (Gigabit Ethernet), and 10GbE rates. The use of interfaces with such rates renders intermediate SONET/SDH multiplexing NEs as potentially obsolete. Reducing network overlay and eliminating SONET/SDH multiplexing and

associated stand-alone NEs is also accompanied by removing SONET-layer traffic grooming and survivability functions and replacing them with management at higher (IP, ATM) or lower (WDM) layers. In this scenario, the function of multiplexing traffic onto wavelengths may be passed onto the IP/MPLS routers.

A critical issue for realizing such intelligent optical networks is how to provide the desired features of rapid provisioning/restoration and automated capabilities between the optical layer and the client layers. It is widely accepted that the best way to achieve this is to adapt the IP topology self-discovery and routing capabilities to the optical network environment [1]. Current research focuses on the use of distributed management schemes such as multi-protocol label switching (MPLS) to provide the control plane necessary to ensure automated provisioning and maintaining connections and managing network resources. In this type of application the label is the wavelength of the incoming signal; hence, the term multi-protocol lambda switching (MP $\lambda$ S) [3,4] is more commonly used. Industry organizations like the OIF and IETF are rushing to extend MPLS-framework (Generalized-MPLS) [5] to support not only devices that perform packet switching, but also those that perform switching in time, wavelength, and space.

In current network scenarios, the traffic is dynamic and heterogeneous where the networks are required to provide dynamic services to the user at a rate that is much lower than the full wavelength capacity. These sub-rate streams are very common and can comprise smaller TDM channels (e.g., SONET/SDH 155 Mb/s OC-3 or 622 Mb/s OC-12), or a variety of storage area network (SAN) protocol interfaces (e.g., 200 Mb/s Escon, 100/1000 Mb/s Ethernet, 1.0 Gb/s Fiber Channel, etc), or even

arbitrary packet flows. In order to achieve maximum efficiency, one would need to bundle these low-rate traffic streams efficiently onto high capacity lightpaths so the number of wavelengths that have to be processed at each router is minimized. This is known in the literature as the traffic-grooming problem [6,7]. The main objective is to reduce the number of IP ports needed on the routers (one per wavelength added/dropped at the router) as well as to reduce both the total switching capacity of the routers and the number of wavelengths required to achieve full connectivity.

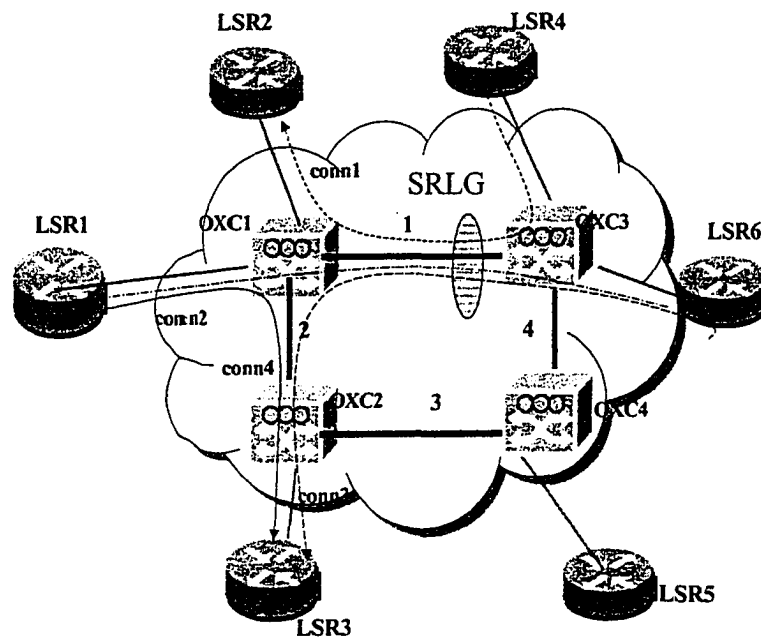


Figure 5.1- IP network over WDM network

Most early work on traffic grooming has focused on SONET rings [8], where traffic is often static and known in advance. This is appropriate because today's backbone transport infrastructures are organized in rings. However, as networks evolve to become more IP-centric, grooming for IP traffic will become an important area for future work. Recently, the traffic grooming in WDM mesh networks has been addressed. In [6], the authors investigate the problem of grooming lower-speed

traffic streams into high capacity lightpaths in a WDM-based optical mesh network. A mathematical formulation is presented and several fast heuristics are proposed and evaluated. Meanwhile, in [7], the authors proposed two heuristics to route lower-speed traffic streams over a logical topology and derive analytical models for the network blocking probability. The performance of WDM networks with sparse and constrained grooming capabilities have also been studied in [9]. Here, the authors specify an analytical model to calculate the blocking performance based upon link and wavelength independence assumptions.

To address traffic aggregation in IP layer, one needs to extend the conventional dynamic provisioning schemes [9], where the bandwidth of a connection request is assumed one wavelength, to support the provisioning of “sub-lambda” connection requests. In this chapter, we propose two GMPLS-based routing algorithms to provision end-to-end paths for low-rate traffic-stream connections by leveraging MPLS based explicit routing algorithms and optical layer routing and wavelength assignment. The performance of each algorithm is then evaluated for three different constraint adaptive selection schemes, namely, conventional shortest path selection scheme, least loaded and most loaded selection schemes. In this study, the IP network routing, topology distribution, and signaling protocols are assumed to be independent of the corresponding protocols in the optical domain, but relevant routing information are exchanged between two network domains. To further augment the Integrated routing algorithms, this chapter examines a new integrated protection scheme to

dynamically allocate restorable-bandwidth guaranteed paths for designing IP/MPLS over WDM networks that can protect against single optical failure [10, 11].

The rest of the chapter is organized as follows. Section II describes the IP over WDM network architecture; an overview of network survivability is summarized in section III, the proposed constraint-based routing algorithms are introduced in section IV. Section V presents a framework for the design of a survivable IP over WDM network. Performance evaluation is presented in section VI and section VII concludes the paper.

## **5.2- IP over WDM Network Architecture**

The optical network model considered here consists of multiple Optical Cross-connects (OXCs) interconnected by optical links in a general topology [2, 10] (referred to as an "optical mesh network"). Each OXC is assumed to be capable of switching a data stream from a given input port to a given output port. This switching function is controlled by appropriately configuring a crossconnect table. A "lightpath" from an ingress port in an OXC to an egress port in a remote OXC is established by setting up suitable crossconnects in the ingress, the egress and a set of intermediate OXCs such that a continuous physical path exists from the ingress to the egress port, (see figure1). Even though the introduction of optical switches into the core network has alleviated electronic bottlenecks at intermediate routing nodes, appropriate wavelength routing algorithms are required to efficiently allocate channels for user

requests. Assigning network resources to successfully carry connection requests (lightpaths) is well known as the routing and wavelength assignment problem (RWA) [11] and has been extensively studied over the past few years.

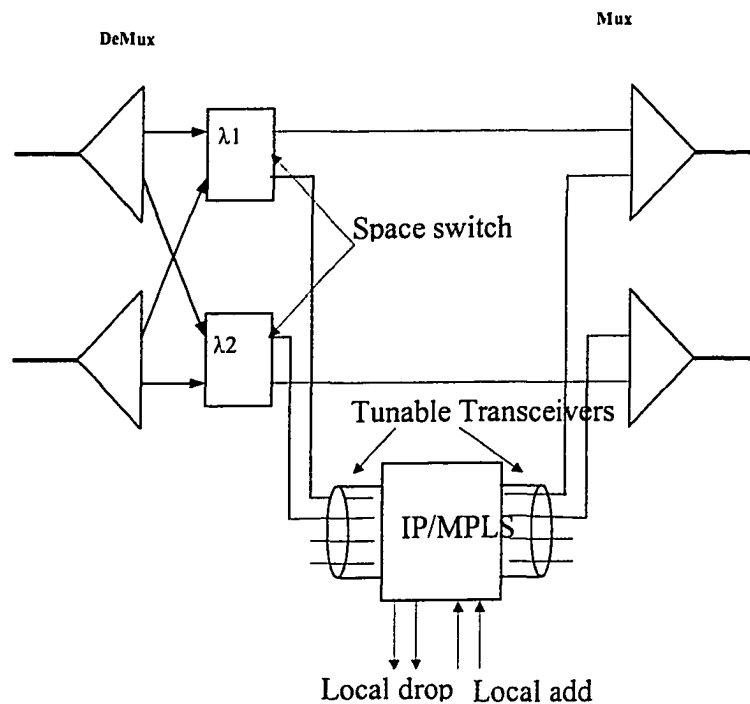


Figure 5.2. IP over WDM node architecture

In this proposed network model, IP routers are attached to the optical core network and connected to their peers over dynamically switched lightpaths. Figure 5.2 shows the IP over WDM node architecture for 2 wavelengths. The node architecture is composed of two components (figure 5.2): Wavelength routing switch (WRS) and IP/MPLS router. The WRS performs wavelength routing and enables space switching of wavelengths from one port to another. On the other hand, the IP/MPLS router is attached to the optical switch through an array of tunable transceivers (4 in the example shown in figure 5.2) and has its own instance of

routing algorithm and helps on provisioning label switched paths (LSPs) and multiplex them efficiently into high capacity lightpaths and setting up new lightpaths to build adjacencies with other routers. With this, a signal is either switched from an input port to an output port without undergoing O/E/O (optical-electrical-optical) conversion, or converted into electrical at the local IP/MPLS router and dropped if this is the final destination or converted into electrical and statistically multiplexed with another traffic stream and sent out at the corresponding output port.

Here the set of lightpaths connecting IP routers represents the logical IP-connectivity (figure 5.3) between the routers and based on this connectivity, IP routers are capable of configuring their routing tables, exchanging information regarding their connectivity, SRLG attributes, bandwidth usage of individual lightpaths, etc. This allows traffic-engineering procedures to be employed at the IP layer to effectively map the traffic into the corresponding lightpaths and load-balance the traffic across the network.

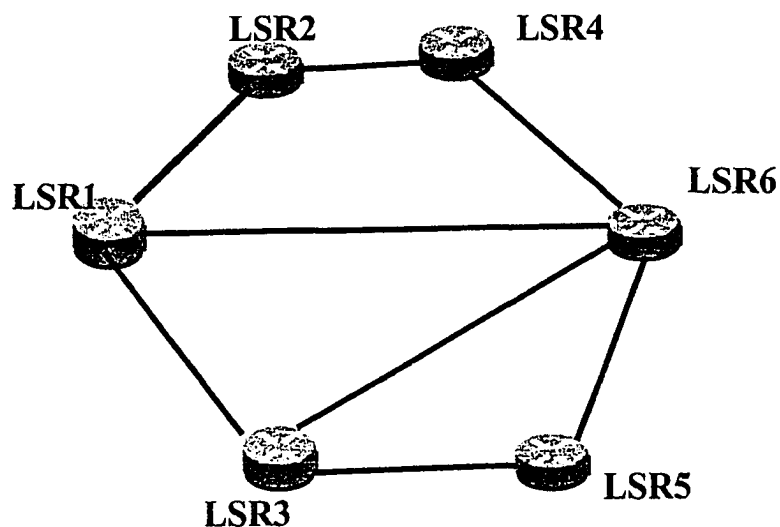


Figure 5.3. Logical connectivity between IP/MPLS routers

### 5.3- Sub-wavelength Path Provisioning

In addition to provisioning end-to-end lightpaths, the optical control plane should have its signaling protocols modified to allow provisioning sub-wavelength bandwidths across the lightpath instead of provisioning full discrete bandwidths [12, 13]. Moreover, the control plane should also allow the provisioning of different type of services, e.g., SONET, GbE, Multimedia (with video, audio and storage services), thus extending the use of it to network services. The services can be mixed on one or more wavelengths in different combinations, limited only by the lambda bandwidth. These services might have different QoS or priorities, such as constant bit rate (CBR), guaranteed bandwidths (GBW) and best effort service (BE). This all packet switching allows cross-connecting services from one lambda to another by using packet switching techniques, see figure 5.4. Here, bandwidth aware OSPF-like protocols are used to calculate the best available feasible route. These protocols require not only traditional topology discovery but also traffic engineering extensions.

An IP router maintains information about its logical connectivity with other routers [13], and based on this connectivity a source router is capable of provisioning an end-to-end path for a connection setup request by running its own instance of routing algorithm. A path on the IP-layer is a set of consecutive lightpaths, or a direct lightpath, with available bandwidth. IP traffic is then forwarded on this new IP-connection. If the path for a connection request spans more than one lightpath, traffic sent along the path will be dropped at intermediate nodes, which are connecting-points of different lightpaths along the path, and statistically multiplexed with other

traffic on the outgoing lightpath (figure 5.4). In this approach, information regarding the bandwidth availability along the logical links (lightpaths) and the electronic forwarding speeds of intermediate IP routers, are necessary and crucial for the IP-routing algorithms to compute paths to setup connections. The adaptive traffic engineering mechanism to effectively route and load-balance the traffic across the network and fast- rerouting are two basic features that the constraint-based routing algorithms should support. In this regard, extensions to existing routing algorithms to support traffic engineering and the provisioning of feasible paths subject to certain constraint should be implemented.

As mentioned earlier, in sub-wavelength traffic provisioning, traffic flow to reach its destination may be routed on a multi hop route (e.g., passing a set of successive lightpaths between s-d pair), and each lightpath may carry packets with different final destinations. Hence, a packet level processing is carried out by the IP/MPLS router attached to the optical node where the lightpath is terminated. For example, in figure 5.4, lightpath between router “a” and “c” carries packets, which are destined to router “c” and “d” respectively. At “c”, traffic undergoes an O/E (Optical to Electrical) conversion and a packet distiller module separates the transit packets from the packets to be dropped at “c” by reading off their labels. Transit packets then are forwarded to the label swapping module and new labels are then attached to those packets destined to “d”. Packets generated at “c” and destined to “d” go through a label encapsulation procedure and get aggregated with the transit packets to be mapped onto the lightpath setup between these two routers. Attaching

labels to packets will speed up the forwarding mechanism; an intermediate node by reading off the labels will immediately know whether to drop locally the packets or forward them into an output port/wavelength. To complete the provisioning of paths for sub-wavelength connections, bandwidths need to be reserved along the paths by using signaling protocols (e.g. RSVP extension). A label switched path, (LSP), for a particular flow of data between s-d pair needs to be established before the data transmission can begin. The signaling protocol again follows a “downstream-on-demand” mode whereby an upstream node makes an explicit label request and a downstream node assigns a particular label in response to that request. Moreover, a node can only assign a label to its upstream neighbor if it is the destination node, or if it already receives a label from its next hop (downstream node). Thus, the signaling phase consists of a label request traversing hop-by-hop from the source to the destination, followed by a label assignment traversing in the opposite direction back to the source. The concept of label here is different from the one used for the optical domain, and the signaling is also independent of the optical signaling.

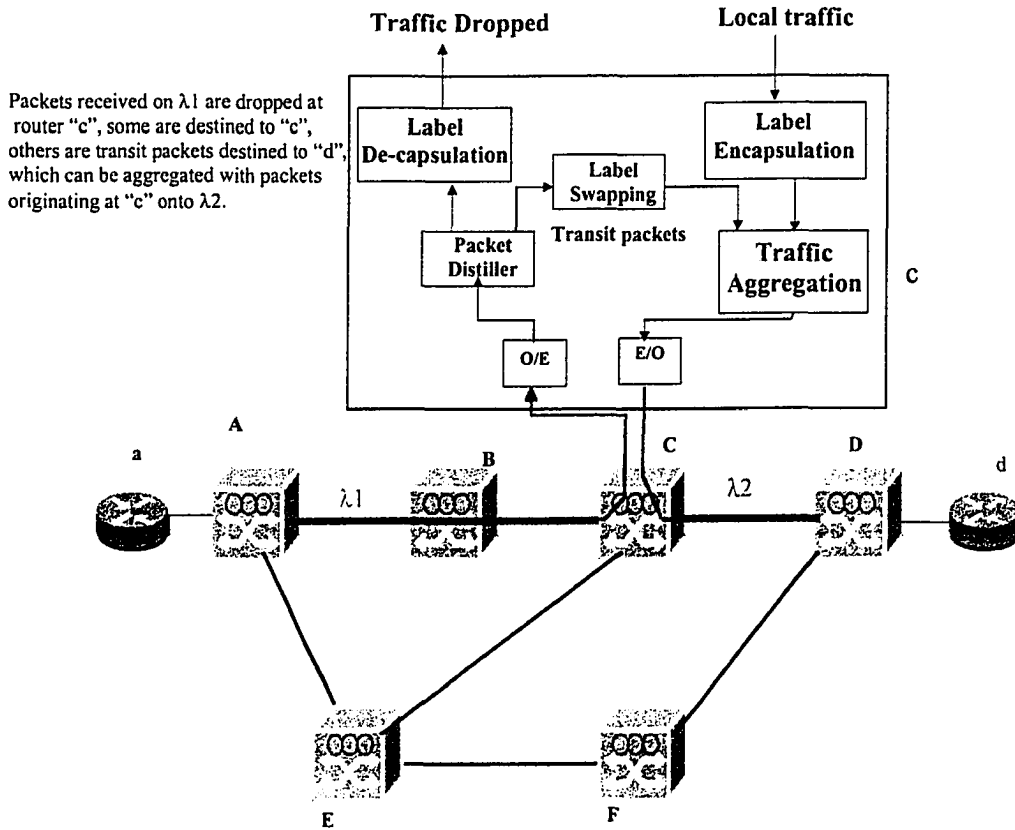


Figure 5.4- Service Provisioning/Aggregation at the Router Level

We show the difference through an example in Figure 5.4. Assume that three calls arrive sequentially: the first call between "a" and "c" requesting a bandwidth for 1 Gbps, the second between "c" and "d" requesting a bandwidth of 3 Gbps, and the third between a and d requesting a bandwidth of 2 Gbps, furthermore, the channel capacity is assumed to be 10 Gbps. The first connection can be satisfied by setting up a lightpath between "a" and "c" on the physical topology. The same thing goes for the second call request between "c" and "d". However, the story changes for the third request between "a" and "d". Router "a" is able to provision a path at the IP-layer without setting a lightpath between "a" and "d" (path is "a-c", "c-d"). Here, packets going from "a" to "d" are mapped to the first lightpath (a-c) and then a packet level

switching takes place at the IP/MPLS node “c” to forward the packets into the next lightpath. Clearly, a mechanism to forward the packets from one lightpath to another lightpath is needed to successfully deliver packets to their ultimate destination, and differentiate those packets from the packets dropped at the intermediate node or forwarded into another lightpath because their destinations are different. Here, MPLS signaling protocols can be used to request/bind labels into packets that need to be forwarded at intermediate nodes.

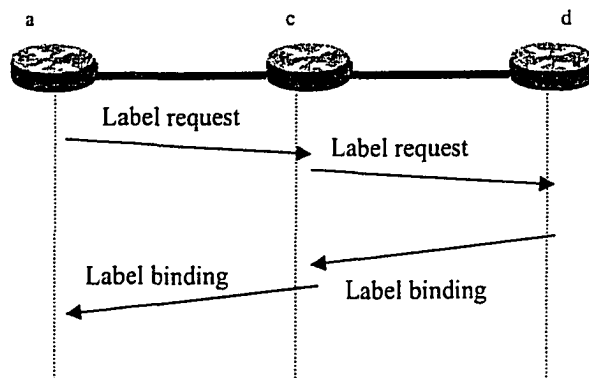


Figure 5.5- Label binding for sub-wavelength provisioning at the IP-layer

Figure 5.5 shows the process of establishing an LSP between “a” and “d” to deliver packets destined to “d”. The lightpath “a-c” can carry some packets, which could be dropped at “c”, and the rest of the packets are statistically multiplexed and forwarded onto the lightpath “c-d” until they reach the ultimate destination. Labels attached to those packets will help the intermediate router “c” to decide to which output port the packets should be forwarded. Note that the computed path between “a-d” should be a feasible path conforming to the TE requirement.

## 5.4. Constraint-Based Routing Algorithms

### 5.4.1 General Approach

To carry connection requests, lightpath connections may be established between pairs of nodes. A connection request may traverse a single-hop route or multi-hop route. In single-hop case, a connection is allowed to traverse a single lightpath (same wavelength is used throughout the entire route from source-to-destination), which means that only end-to-end traffic grooming (multiplexing) is allowed. In multi-hop case, a connection is allowed to traverse multiple lightpaths, i. e., a connection can be dropped/terminated at an intermediate node and multiplexed with other low-speed connections on different lightpaths (wavelengths) before it reaches its destination. In the case of a single-hop route, only calls with the same source-destination pair can be multiplexed onto one lightpath. Clearly, this will yield low lightpath utilization and further lower overall network throughput also. On the other hand, in the case of a multi-hop route, calls with different source-destination pairs can be multiplexed into the same lightpath.

Here, we first start by initializing the virtual topology to empty and hence all IP routers' routing tables have no entries. We then calculate all shortest routes (min. number of hops) for every source-destination pair and store them at each corresponding OXC. Border routers supply large capacity and relatively aggregated flow to the optical domain. IP routers, based on their connectivity are capable of finding potential routes to carry the traffic streams by running their own instance of routing algorithm on the virtual topology.

Depending on the algorithm to be used, if no available single/multi hop route were found on the logical topology, the source node sends a request to the ingress OXC to setup a lightpath, on the physical topology, to the destination. Then, the ingress OXC invokes its routing and wavelength assignment (RWA) algorithm to setup a lightpath to the destination, where the first fit (FF) [11] wavelength assignment scheme is used. Note that the OXC has to disseminate this information to other core OXCs and to the border IP router that is connected to it. Subsequently, the source IP router that had set up this lightpath adds an entry to its routing table that identifies this lightpath. It also maintains information about the bandwidth allocated along this lightpath. This information is then disseminated to other border routers through signaling channels.

#### **5.4.2 The First proposed Algorithm:**

The detailed implementation of the algorithm used by the source client node to process an incoming request is described in Table 1.

*Step1: Search whether a direct lightpath between s-d pair with available bandwidth exists, if yes, go to step6, and else continue.*

*Step 2: Search whether there is pending direct lightpath setup which may be triggered by a previous call, if yes go to Step 5, else continue.*

*Step3: Run IP constraint-based routing on the IP connectivity with  $h=2$  ( $h$  is the number of hops) to find the path with available bandwidth. If successful, go to step 6, else continue.*

*Step4: Buffer the incoming request, and invoke the dynamic RWA to compute a direct lightpath between s-d pair. If there is no resource available for direct lightpath, block the call and stop, else use the GMPLS based signaling mechanisms to establish the new lightpath, and continue.*

*Step5: Buffer the incoming request, and wait for lightpath setup successful acknowledgement, if unsuccessful, block the call and stop, otherwise, update the logical topology at each node by broadcasting the message which contains new lightpath information, and continue,*

*Step6: Reserve the bandwidth along the related lightpaths for the incoming request or/ requests in buffer, update the IP layer resource usage information, thus the connection request is satisfied.*

Table 1

The above algorithm attempts to establish multi-hop connections on the logical topology for blocked calls; clearly this improves fiber utilization. However, routing on the logical topology is restricted to two logical hops since considering too many hops cannot improve the network throughput. Moreover it increases the required electronic speed at intermediate hops [13] and may result in longer end-to-end delays.

### 5.4.3 The Second Proposed Algorithm:

The objective of the second approach is to exploit optical resources as much as possible. In this approach, for every connection request, we first setup a lightpath on the physical topology for each traffic stream if a direct lightpath couldn't be found, and if the request is blocked, we resort to routing on the IP connectivity. This

approach clearly tries to increase the connectivity of the virtual topology connecting IP routers by setting up direct lightpaths first. This in turn further increases the options for the routing algorithm on the virtual topology. The steps for this algorithm are similar to the first one, except that steps 3 and 4 are conversed. Note, it has been assumed that the path selection scheme used in both of the above algorithms is the conventional shortest path routing where the constraint imposed is the minimum number of hops.

#### **5.4.4 The Two Adaptive Constraint-based Routing Schemes:**

Note that in both of the above algorithms, even though routing traffic flows over the logical topology is dynamic it is still not fully adaptive and the only constraint imposed was the number of hops. To add more constraints, we incorporate link-state information into the path selection decision and make the routing more adaptive. Specifically, in adaptive routing schemes, the algorithm operates from a fixed set of potential connection paths for each stream, with no pre-determined order. Values are given to the paths in the set, and the path with the highest value at call-arrival time is selected. Now one variant of adaptive routing is the *residual capacity adaptive routing* (RCAR) scheme, which was first proposed for conventional circuit switched networks [14] and it uses occupancy information on all network trunk groups. It is assumed that this information is periodically updated via network probing mechanisms running on the respective nodes. Herein, we assume such information is also available for the routing algorithm and instead focus on the performance objectives. The overall goal of the routing algorithm is to choose routes subject to

specific constraints, for instance a route that increases the rate of successful connection setups, while maximizing the utilization of network resources [14]. Hence, we associate each link on the logical topology with a parameter  $C_r^l$  that directly reflects the amount of the residual capacity along the link “ $l$ ” between node  $i$  and node  $j$ . If  $P = \{p_1^{ij}, p_2^{ij}, \dots, p_n^{ij}\}$ , with respective residual capacities  $\{C_{p_1}^{ij}, C_{p_2}^{ij}, \dots, C_{p_n}^{ij}\}$ , is the set of direct lightpaths between nodes  $i$  and  $j$ , then  $C_r^l$  equals to either  $\text{Min}\{C_{p_1}^{ij}, C_{p_2}^{ij}, \dots, C_{p_n}^{ij}\}$  or  $\text{Max}\{C_{p_1}^{ij}, C_{p_2}^{ij}, \dots, C_{p_n}^{ij}\}$ , or  $f\{C_{p_1}^{ij}, C_{p_2}^{ij}, \dots, C_{p_n}^{ij}\}$  according to the different path selection criteria.

We define a logical link “ $l$ ” to be feasible if  $C_r^l$  is greater than or equal to  $C_{Req}$ , where  $C_{Req}$  is the bandwidth required by the connection request.

Next, we show how the different constraint selection schemes will choose paths to meet certain selection criteria. Consider a path  $R^{s-d}$  between s-d pair on the logical topology,  $R^{s-d}$  can be a single hop (direct feasible lightpath,  $R^{s-d} = \{l\}$ ) or multi-hop (sequence of consecutive feasible lightpaths,  $R^{s-d} = \{l_1, l_2, \dots, l_m\}$ ) and we associate a path with a combined cost ( $C_R^{s-d}$ ) defined to be:

$$C_R^{s-d} = \begin{cases} \sum_{i=1}^m \frac{1}{C_r^{l_i}} & \text{if } R = \{l_1, l_2, \dots, l_m\} \\ \frac{1}{C_r^l} & \text{if } R = \{l\} \end{cases} \quad (\text{Equation 1})$$

This combined cost is used to evaluate each path between s-d pair individually; we have chosen this function into our selection scheme since it dynamically balance the impact of the hop count and the path load.

Since adaptive routing operates from a set of alternative routes, the routing algorithm should compute all available feasible routes between an s-d pair, let  $S = \{R_1^{s-d}, R_2^{s-d}, \dots, R_k^{s-d}\}$  be the set of all available feasible routes between s-d pair and  $C = \{C_{R_1}^{s-d}, C_{R_2}^{s-d}, \dots, C_{R_k}^{s-d}\}$  be their respective associated combined costs. Next we present two different constraint selection criteria to select a route amongst a set of potential available routes.

#### 5.4.4.1- Least Loaded Routing (LLR):

The least loaded routing algorithm is a well-known algorithm that attempts to evenly distribute the load among the alternative routes between a given source-destination node pair. Here a connection is always routed on a path with the highest residual capacity (i.e., highest combined residual capacity). We propose to use the LLR concept into our path selection procedure to select routes on the virtual topology where a connection is always routed on a path with the lowest combined cost. Here the cost associated with a virtual link "l"  $C_r^l$  is equal to  $Max\{C_{p_1}^j, C_{p_2}^j, \dots, C_{p_n}^j\}$  and the IP/MPLS routers keep the value  $C_r^l$  and the associated lightpath information. Amongst the set of different available routes, this scheme selects the path  $R_i^{s-d}$  with cost  $C_{R_i}^{s-d} = Min(C_{R_1}^{s-d}, C_{R_2}^{s-d}, \dots, C_{R_k}^{s-d})$ . Table 2 illustrates a dynamic algorithm to select the desired path for each coming call between source node and destination node, and used to substitute Step1 through Step 3 in table 1.

*Step1: Compute  $K$  candidate routes if possible with  $R_k$  representing  $k^{\text{th}}$  route ( $1 \leq k \leq K$ ) on the IP-connectivity.*

*Step2: For  $R_k, 1 \leq k \leq K$ , calculate  $C_{R_k}^{s-d}$  by using equation 1.*

*Step3: Search for  $i$  such that  $C_{R_i}^{s-d} = \text{Min}(C_{R_k}^{s-d})$  for ( $1 \leq k \leq K$ ).*

*Step4: Select route  $R_i$  as the path to carry the request traffic stream.*

Table 2

#### 5.4.4.2- Most Loaded Routing (MLR):

The most loaded routing algorithm attempts to pack connections on alternative routes by favoring the most utilized trunks [14]. Thus, it attempts to leave other alternative routes very lightly loaded and increase the chance of admitting future incoming connections with larger bandwidth requirements. This technique tends to pack traffic streams onto paths that are highly utilized in order to maximize the utilization of individual lightpaths. Here, the set of possible paths are arranged in descending order of their combined cost values and searched sequentially starting with the first one. Similar to LLR, we compute the residual capacity of a virtual link “ $l$ ”,  $C_r^l$  to be equal to  $\text{Min}\{C_{p_1}^{ij}, C_{p_2}^{ij}, \dots, C_{p_n}^{ij}\}$ . This scheme selects the path  $R_i^{s-d}$  with cost  $C_{R_i}^{s-d} = \text{Max}(C_{R_1}^{s-d}, C_{R_2}^{s-d}, \dots, C_{R_k}^{s-d})$ . The algorithm is similar as the one tabulated in table 2, except that we are looking for the route with the maximum combined cost.

#### 5.5- Fault Tolerant Routing in IP over WDM Network:

The second component to augment the above-developed constraint-based routing algorithms is to enhance them to provide fault tolerance, i.e., dynamically

provisioning restorable bandwidth guaranteed paths [15]. In this regard, additional constraints are needed to allow the routing algorithm to compute diverse routes such as link/node disjoint and SRLG-disjoint routes. Here, the RWA should be enhanced to compute the first  $k$  diverse paths (if possible) between any source-destination pair [16]. This will allow the ingress OXC to provision both a primary and a backup path to carry the traffic flow whenever requested from the IP edge router. On the other hand, if the IP router were to provision the path on the logical connectivity, it should be capable of computing a primary path for the active LSP with available capacity and another physically diverse route for the backup LSP. Note that to reduce the overhead of backup traffic streams through a logical link and improve the overall network throughput, bandwidth-sharing techniques can be used, also known as backup multiplexing [18]. Here, note that whether sharing backup capacity can be used or not depends on the information available to the routing algorithms. Hence, we propose two different techniques where the first one doesn't allow backup sharing and called the non-information routing and the second is the full information routing and permits the best sharing:

#### **5.5.1- The Non-Information Routing:**

Under this approach, the only information available to the routing algorithm is the residual capacity throughout the network [15, 17], and the physical paths of the virtual links. The first parameter can help in computing paths with available capacity while the second allows the provisioning of physically diverse routes. However, no

further information is provided to tell whether a backup bandwidth along a virtual link could be shared or not.

### 5.5.2- The Complete Information Routing:

In contrast to the previous one, here more information is provided from the routing algorithm (such as the primary and backup routes of all connections routed on each virtual link, the reserved working bandwidth and backup bandwidth) which allows to make decision whether the restoration capacity could be shared or not.

### 5.5.3- Sharing Restoration Capacity:

As mentioned earlier, sharing restoration capacity can yield to a better overall network throughput, however the payoff here comes from the amount of information that is needed by the routing algorithm to make the decision. Here, we assume that this information is given and we provide a mechanism to estimate the total shareable restoration capacity along a virtual link. Let “l” be a virtual link (lightpath) of capacity  $C$  connecting two remote IP routers. Let  $P = \{p_1, p_2, \dots, p_n\}$  be the set of primary LSPs going through that link and let  $c_{p_i}$  be the bandwidth required by a primary LSP  $i$ . Moreover, let  $B = \{b_1, b_2, \dots, b_m\}$  be the set of backup LSPs going through the same lightpath with  $c_{b_i}$  be the required restoration capacity for a backup LSP  $i$ .  $C_R$  is the residual capacity along the lightpath and can be computed as  $C_R = C - C_p - C_B$  where  $C_p$  and  $C_B$  are the total primary capacity and backup capacity along the lightpath.  $C_p$  can be estimated to be  $C_p = \sum_{i=1}^n c_{p_i}$ . If restoration bandwidth sharing techniques are not used then  $C_B$  can be estimated as  $C_B = \sum_{i=1}^m c_{b_i}$ .

However, if sharing techniques are used then we estimate  $C_B$  as in the following example:

Let  $B = \{b_1, b_2\}$  be the set of backup LSP along a lightpath and  $p_1, p_2$  be their corresponding primary LSPs. If  $p_1$  and  $p_2$  are link-disjoint then  $C_B = \text{Max}(c_{b_1}, c_{b_2})$ , otherwise,  $C_B = c_{b_1} + c_{b_2}$  and the following property is true  $C_R + C_P + C_B \leq C$ .

The following are the basic properties of the proposed fault-tolerant routing with traffic-grooming capability:

1)- if the working paths of two low-rate traffic stream connections are link-disjoint, they could share the same backup bandwidth on the same link-disjoint backup path.

2)- if the working paths of two traffic stream connections are not link-disjoint, they could be protected by the same link-disjoint path (link-disjoint of both working paths), but they can't share their backup bandwidth. In other words, each working path has to be allocated its own backup bandwidth provided that there exists a sufficient bandwidth on the same link to accommodate both backup bandwidths.

3)-the bandwidth allocated for each of a primary and a backup path for different connections can be groomed onto the same wavelength.

4)- if a working path for a traffic stream connection spans two lightpaths, or more, ( $\{w\} = \{w_1\} \cup \{w_2\}$ ), and if the backup path is routed on two different lightpaths, or more, ( $\{b\} = \{b_1\} \cup \{b_2\}$ ), they should satisfy the following property to ensure fault tolerance:  $\{b_1\} \cap \{w\} = \emptyset$  and  $\{b_2\} \cap \{w\} = \emptyset$ , where  $\{p\}$  represents the set of physical links along a path  $p$ .

5)- A backup LSP is said to be feasible if its route is physically diverse with the route of its primary path and if the summation of the sharable capacity and the residual capacity along the path is greater than or equal to the required protection capacity. Note that if the restoration capacity is not sharable, it is enough to have the residual capacity along the path be greater than or equal to the required protection capacity.

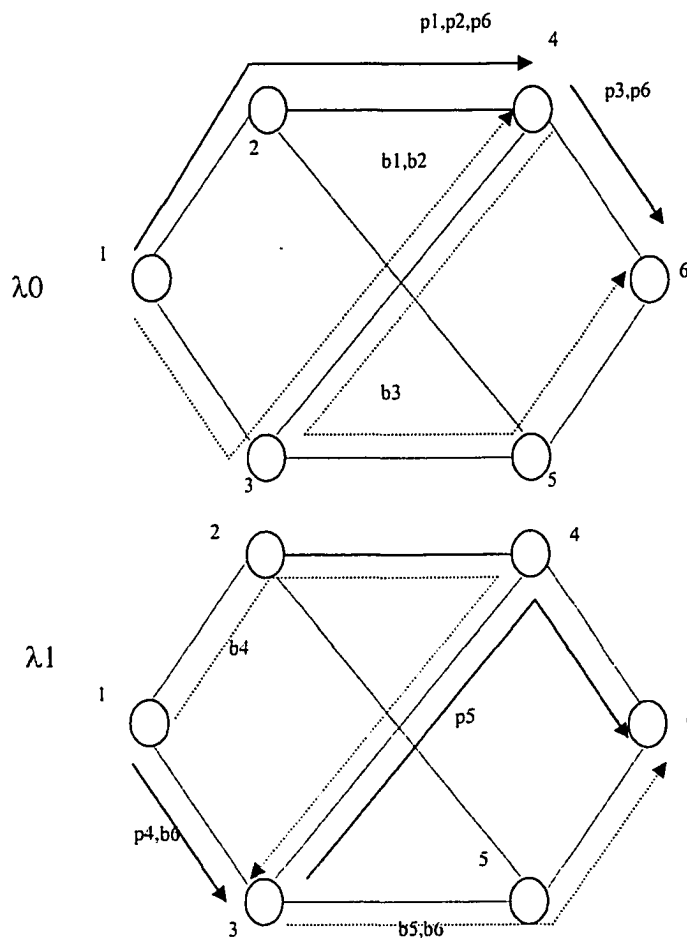


Figure 5.6. Example of a Survivable IP over WDM network

Figure 5.6 shows an example on how to successfully route guaranteed low-rate traffic stream connections. The example shows the network represented as a

layered graph with 2 wavelengths. We assume that 6 traffic-stream requests (LSPs) arrive in order and each request is setup before the next traffic stream arrives. The first request between nodes 1 and 4 can be setup on wavelength  $\lambda_0$  along the links 1-2 and 2-4 ( $p_1$  refers to primary for first request) and the backup path ( $b_1$  refers to the secondary or backup path) is setup on the link-disjoint path (1-3-4) on the same wavelength and each request reserves the corresponding bandwidth. Request 2 comes between the same source and destination and can be setup on the same route if there is bandwidth available. Here since  $p_1$  and  $p_2$  are not link disjoint,  $b_1$  and  $b_2$  cannot share the same backup bandwidth; instead we need to reserve the corresponding bandwidth for each one on this backup path. The same goes for request number 3, 4 and 5. Request number 6 comes between nodes 1 and 6. Here node 1 is able to find a 2-logical hops path to the destination ( $p_1$ - $p_3$ ) and reserves the corresponding bandwidth to route the connection if both lightpaths have available bandwidth. Now for the protection path, node 1 finds a two logical hops path ( $p_4$ ,  $b_5$ ) to reach the destination and backup bandwidth is reserved along this route if the available bandwidth permits. Note that since  $p_5$  and  $p_6$  are not link disjoint, the backup bandwidth on  $b_5$  cannot be shared with  $s_6$ , instead an additional bandwidth needs to be reserved on the lightpath between nodes 3 and 6 to assure 100% recovery of all affected services along the failed path. Here we also show that on the same lightpath, both working connections and backup connections can be multiplexed ( $p_4$  and  $b_6$  are both multiplexed onto the same lightpath).

Next we present the steps for the fault tolerant routing of sub-wavelength connection requests:

*Step1: compute a path with available bandwidth for the primary LSP on the IP-connectivity between the source and destination; if found, update SRLG associated with primary LSP, and go to step3, else continue.*

*Step2: invoke the RWA at the ingress OXC to compute a lightpath from source to destination; if fail, block the call and stop, else update SRLG information associated with lightpath and continue,*

*Step3: remove the lightpaths which contain the primary LSP and the links without enough residual capacity in SRLGs, compute a feasible path for the backup LSP on the IP-connectivity, if successful, go to step 5, else continue.*

*Step4: invoke the RWA at the ingress OXC to compute a lightpath with SRLGs constraints of the primary LSP: if fail, go to step7, else continue.*

*Step5: if the lightpath(s) for primary LSP and backup LSP need to be setup, use GMPLS signaling mechanisms to establish the lightpath(s), and wait for ACK, if get NACK go to step 7, else, update the logical topology at each node by broadcasting the message which contains new lightpath information, and continue*

*Step 6 reserve corresponding bandwidth on both primary and backup paths, update resource usage information at IP connectivity and broadcast the updated SRLGs information over the network, thus the call is accepted.*

*Step 7: drop the call, release the resources of primary LSP, and update SRLG information at source node resource table, and stop.*

Table3

## 5.6- Performance Evaluation

In this section we evaluate the performance of the above-mentioned algorithms by simulations. We consider the NSF network where nodes are interconnected by bi-directional fiber-optic links. Each network node consists of an all-optical OXC for spatial-switching of wavelength entities, and an IP/MPLS router attached to it. All lower-rate traffic streams are generated randomly using a normal distribution, with flow arrivals following a Poisson distribution with mean rate  $\lambda$  and exponential holding time distribution with mean  $1/\mu$ . Furthermore, every flow is equally likely to be destined to any destination node in the network, i.e., even network loading. In the following, the overall blocking performance is used as the metric for comparison.

### 5.6.1- Evaluation of the different constraint-based routing algorithms:

Figure 5.7 shows the blocking probability versus the total load for both routing algorithms, defined in sections 4.2 and 4.3 respectively. In algorithm 1, we observe that the probability for a connection to be routed on two logical-hops path is relatively high. This in turn consumes twice as much bandwidth to setup the connection, as compared to setting up on a direct path (which in turn decreases the overall network utilization). However, the second algorithm attempts to increase the connectivity of the virtual topology as much as possible, thereby yielding more options for IP routing. The overall results confirm that the second algorithm outperforms the others for the above-mentioned reasons.

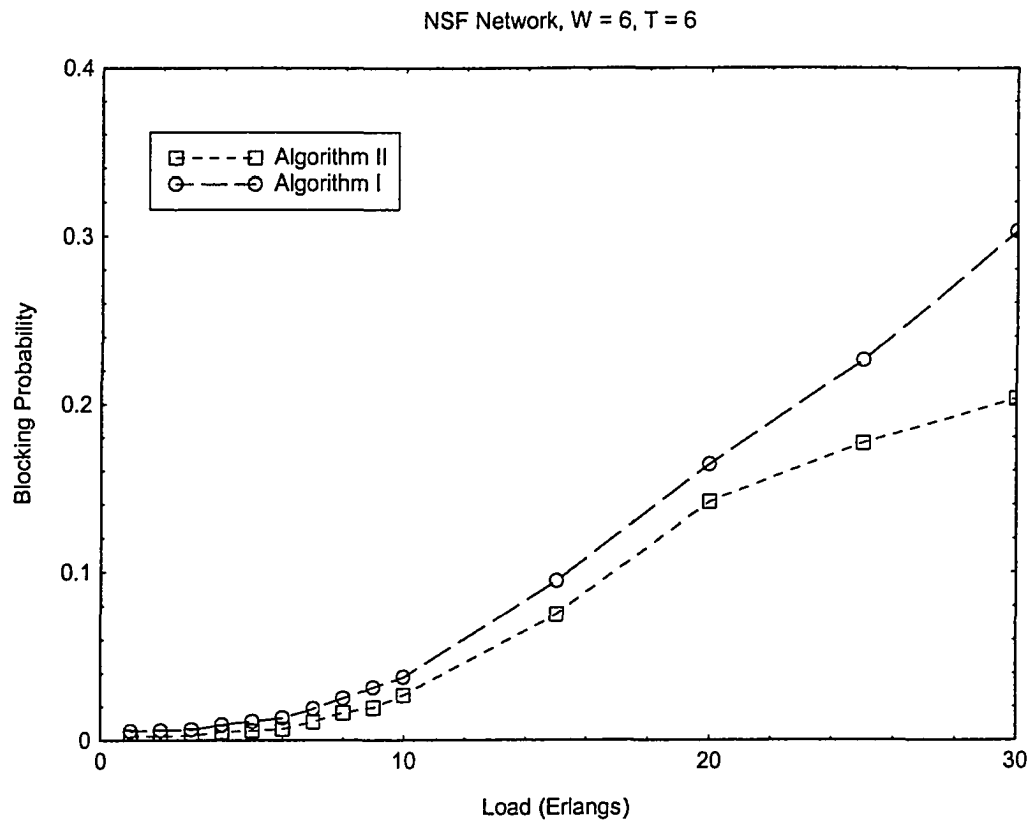


Figure 5.7. Different routing algorithms

Figure 5.8 shows the improvement in network performance introduced by applying adaptive routing where the least-loaded (LLR) and the most-loaded (MLR) routing schemes have used in these calculations. Also included in the figure a reference path selection, first available (FA), is used where the first feasible path available between s-d pair is chosen. In all cases the LLR selection scheme outperforms the other selection schemes since connections tend to be routed on paths with the largest expected number of free trunks.

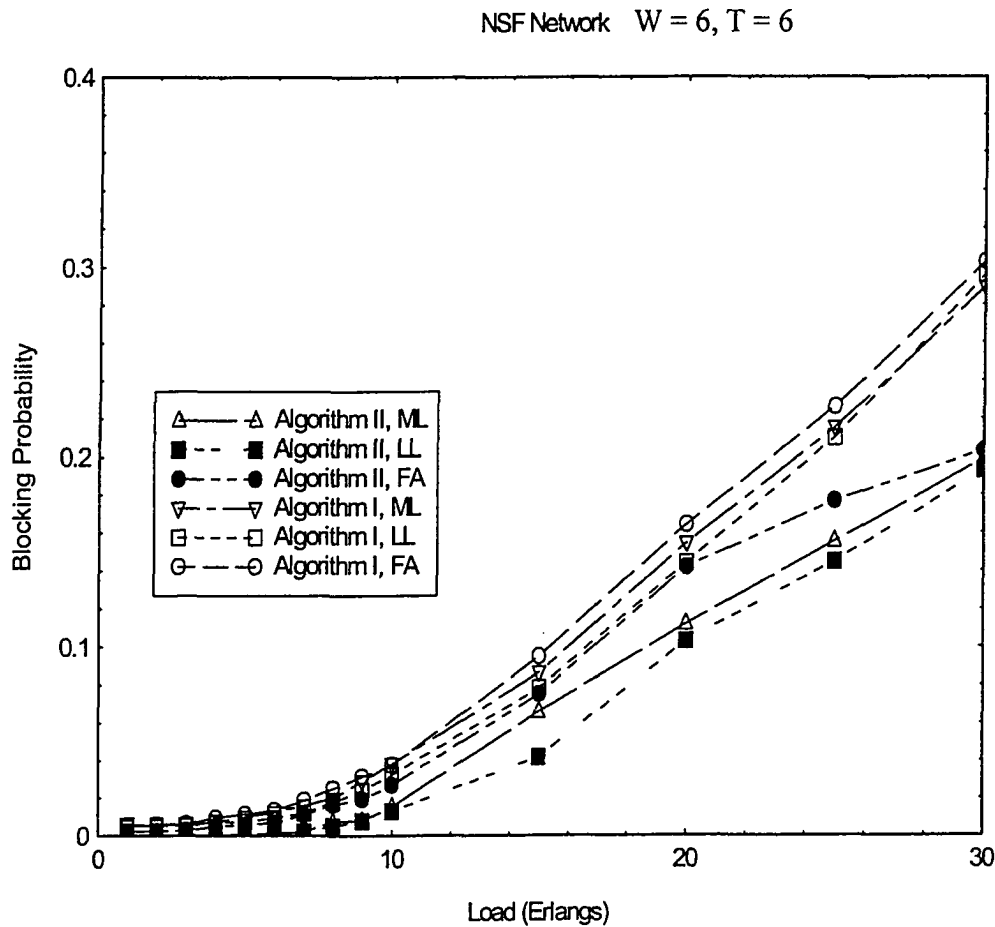


Figure 5.8. Impact of adaptive routing

Next we study the impact of number of transceivers ( $T$ ) and the number of wavelengths ( $W$ ) on the performance of the traffic grooming algorithms. Figure 5.9 shows the blocking probability versus the number of transceivers versus the number of wavelengths in the NSF topology for the previous traffic pattern. Here, the grooming algorithm used is the one in section 5.4.3 and the path selection scheme is the FA. The figure indicates that when the number of transceivers is small, increasing the number of wavelengths does not improve the network performance since there are not enough ports to setup direct lightpaths between the network nodes. Here, the

wavelengths resources end up being underutilized and also connections tend to be routed on multi-hop paths between s-d pairs since direct lightpaths don't exist. On the other hand, when the number of transceivers at each node is increased, the network throughput increases significantly. This is because more lightpaths have been established (giving there are enough wavelengths in the network) to carry the connection requests. However, when most of the links have fully utilized the available wavelengths, increasing the number of transceivers will not help to improve the network throughput and will result in lower transceiver utilization. The figure also indicates that when the number of wavelengths is small, increasing the number of transceivers is of little benefit since there are not enough wavelengths to setup more lightpaths, which leaves the logical topology lightly connected (e.g.,  $W = 2$ ,  $T = 4, 5, 6$ , etc.). On the other hand, a better performance is achieved if available resources (wavelengths, ports) exist; however, an optimum number of resources ( $W_{opt}$ ,  $T_{opt}$ ) is always needed to achieve certain performance for a given traffic load, after which adding more resources will not improve the network throughput and result in decreasing the resources' utilization.

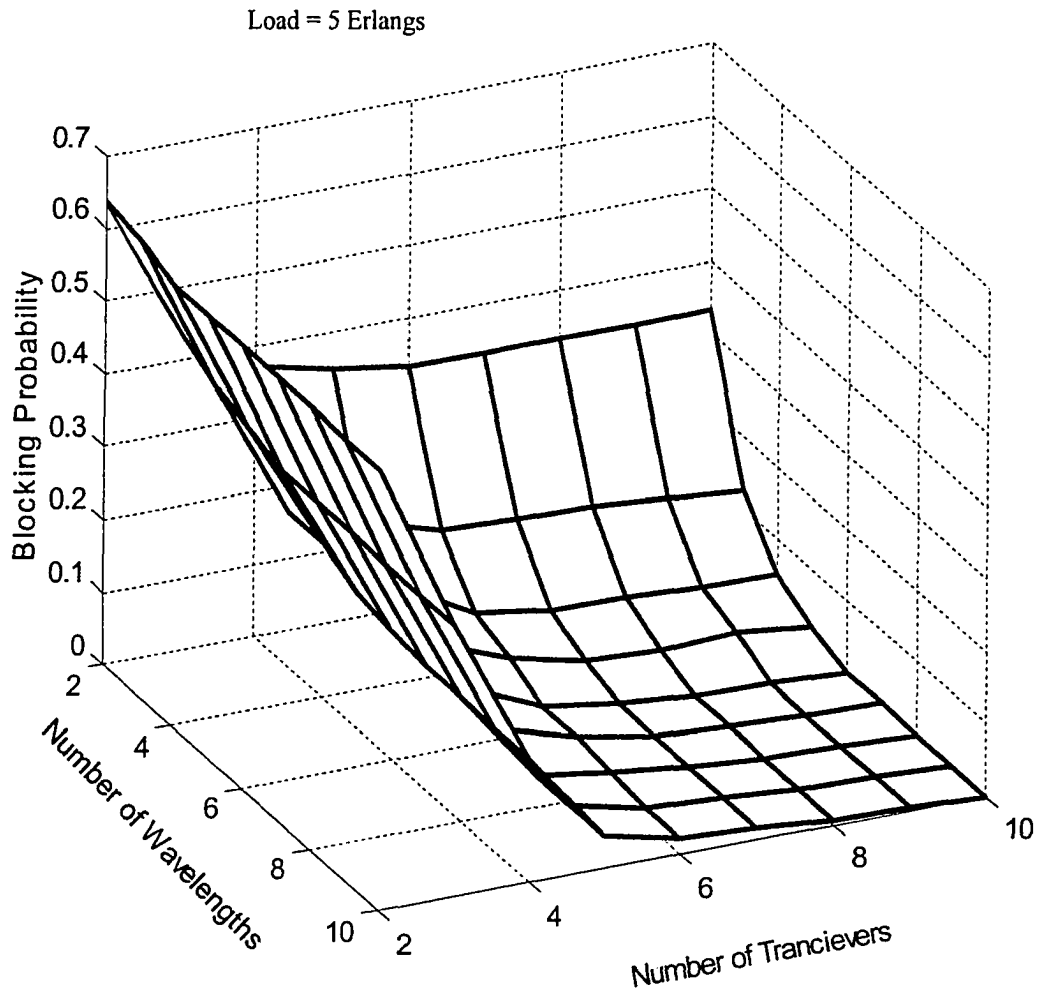


Figure 5.9. Network performance vs. number of wavelengths v. number of transceivers.

### 5.6.2- Evaluation of the integrated-protection algorithms:

Figure 5.10 shows the blocking probability vs. the arrival rate for three different schemes. The first one is the reference WDM scheme where a primary LSP is setup on a direct lightpath and another backup LSP is setup on another direct link-disjoint/SRLG-disjoint lightpath. The second is the proposed integrated IP over WDM protection where it is assumed that working routes are not allowed to undergo

O/E/O at any intermediate logical hop between any s-d pair (only single-hop is allowed for the primary working paths), however protection routes are allowed to multi-hop on the logical topology. The third is the modified IP over WDM protection scheme where it is assumed that both working and protection routes are allowed to multi-hop on the logical topology. As can be seen from the figure, the third scheme outperform both schemes since the algorithm tends to explore multi-hop routes on the existing logical topology and pack low-rate traffic connections more efficiently onto existing lightpaths.

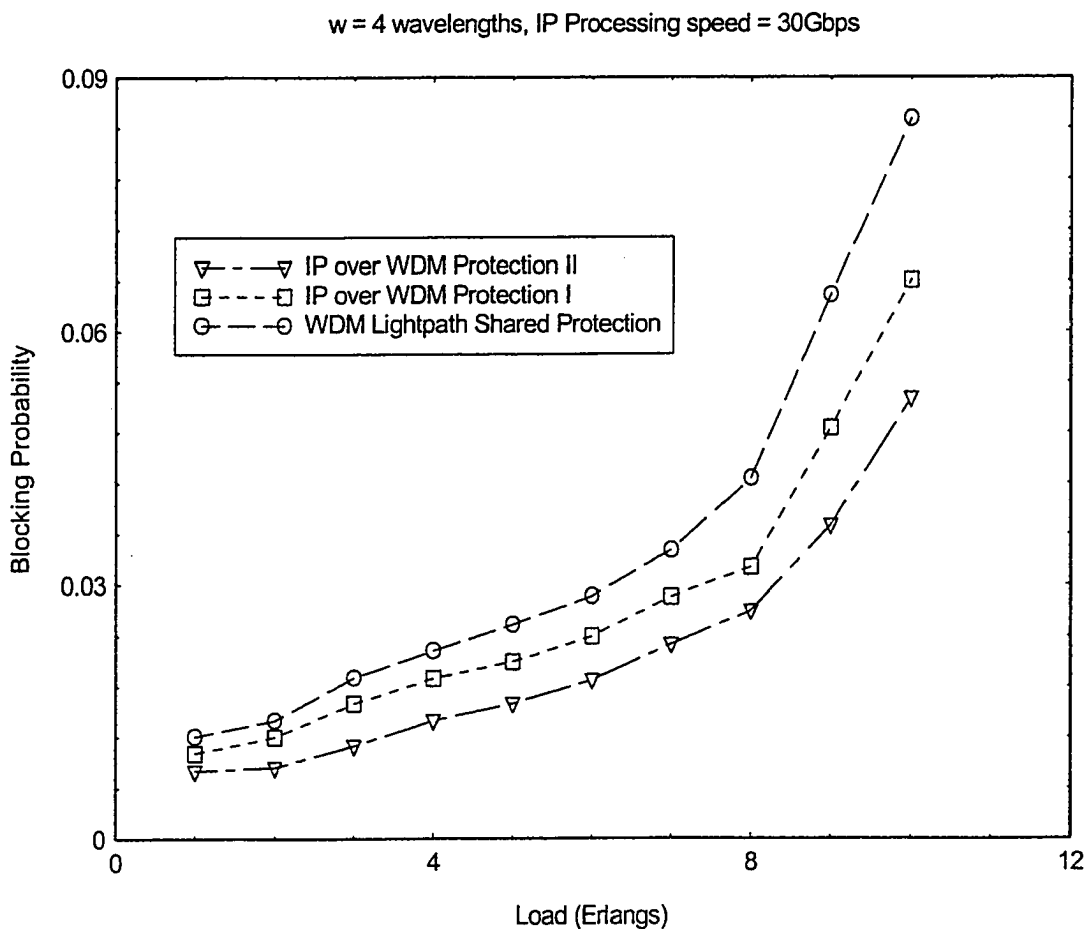


Figure 5.10: comparison of all protection schemes

Based on the network information availability, there are two different protection approaches: the dedicated protection path scheme, which is suited for no-global network information available scenario. If the network status is known at each node, the shared protection approach can be used. The network blocking probability of the shared protection and that of the dedicated one are also compared for our proposed integrated approach and the results are shown in figure 5.11. As can be seen from the figure, sharing backup resources is a versatile approach in a survivable network design. It optimizes the network resource utilization so as to accommodate more future demands in the network. However, appropriate traffic-engineering extensions to existing routing protocols have to be considered to allow for sharing backup resources.

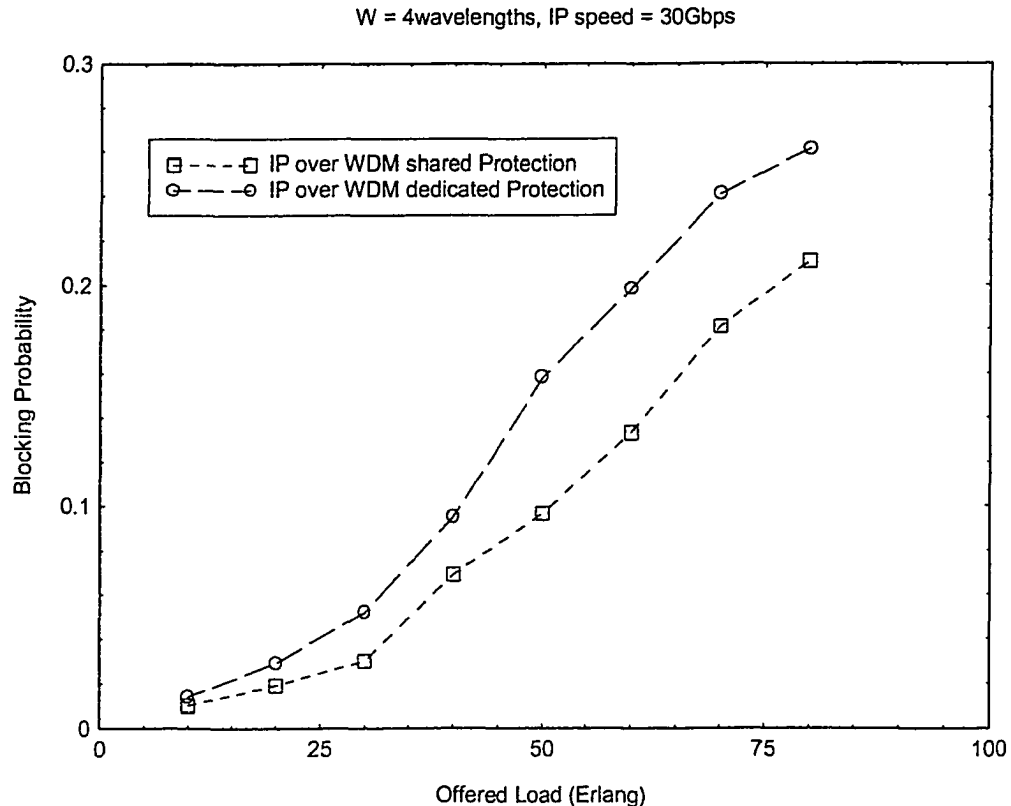


Figure 5.11- Impact of restoration capacity sharing on the network performance

Since multi-hop routing could be used for working/protection path allocation, the IP processing speed becomes another possible bottleneck in providing survivability. Figure 5.12 shows the effect of IP electronic processing speed on network blocking probability for the two proposed IP over WDM protection schemes. As expected, the higher the speed of the IP router, the better is the performance, but up to a certain threshold value at which the performance stays almost the same; since after that value, the wavelength capacity becomes the bottleneck. The simulation results also show that increasing the number of wavelengths per fiber is one way to scale with the IP router processing speed; another way is to increase the capacity of the optical channel. The figure also shows that when the processing speed is relatively small, increasing the number of wavelength will not improve the overall network performance since the bottleneck becomes the electronic speed. On the other hand, the figure shows also the benefit of multi-hop routing capability. While at lower speed, intermediate nodes prevent connections to be multiplexed onto lightpaths, the benefit is clear at high processing speed.

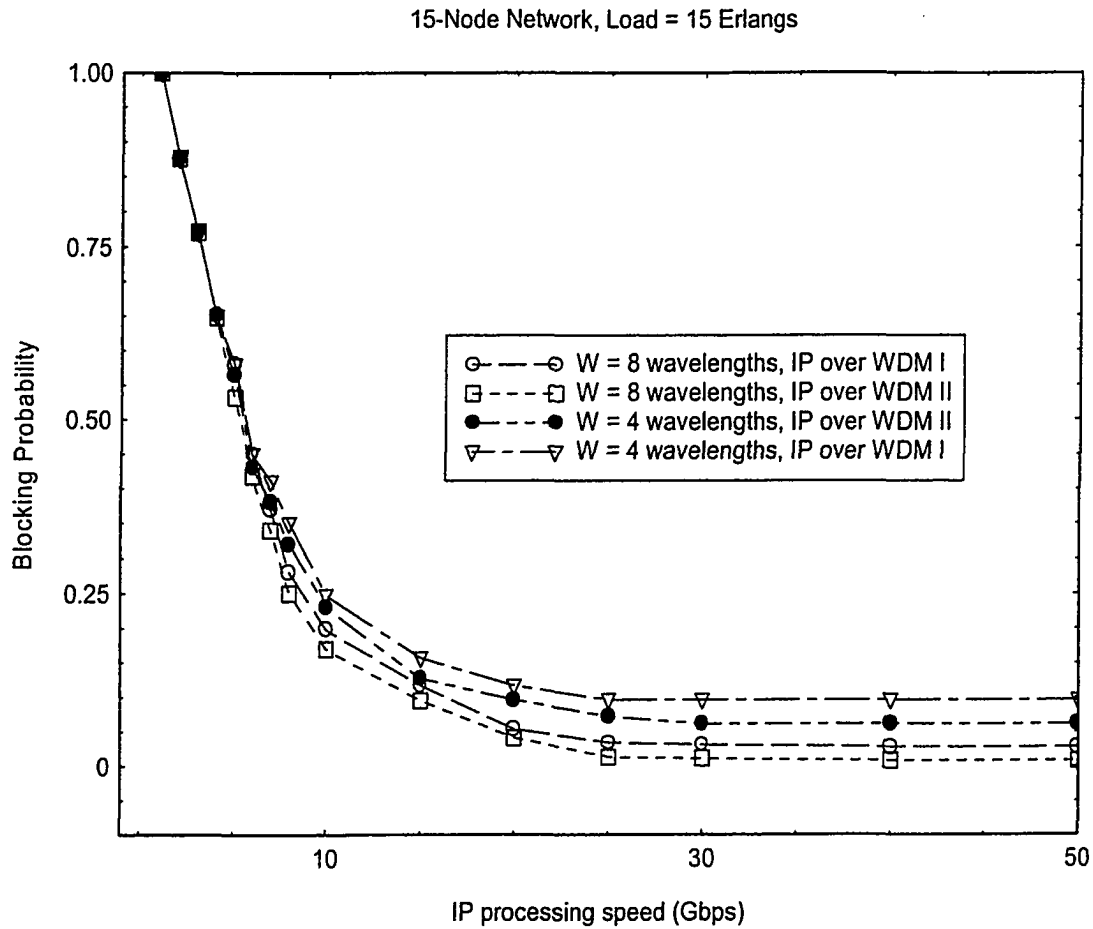


Figure 5.12- Blocking probability vs. Router Processing Speed

### 5.7- Conclusion

This work has developed and implemented comprehensive, unified constraint-based routing algorithms within the generalized MPLS framework (GMPLS) to provision “sub-wavelength” circuits (low-rate traffic streams). Constraint routing is further augmented in this work by dynamically routing both an active and another alternate link/node-disjoint backup path at the same time in order to provision a given connection request. Two different routing algorithms to dynamically route low-rate traffic streams in future IP-centric optical networks have been presented. One algorithm attempts to first route connection requests on the logical topology

interconnecting the IP routers, then, if blocked, on the underlying WDM physical topology by setting up a lightpath if possible. The second algorithm attempts to route connection requests on the physical topology if a direct path couldn't be found, then, if blocked, on the logical topology. This approach increases the connectivity of the virtual topology. The performance of each algorithm is then evaluated for three different constraint-based adaptive selection schemes, namely, conventional shortest path selection scheme, least loaded selection scheme, and most loaded selection scheme. To further augment these algorithms, we have also developed a novel integrated protection scheme to dynamically allocate restorable-bandwidth guaranteed paths for designing IP/MPLS over WDM networks that can protect against single optical link/node failures.

## 5.8 References

- [1] M.A.Ali, A. Shami, C. Assi, Y.Ye, R Kurtz, "Architecture Options for Next-Generation Networking Paradigm: Is Optical Internet the Answer", PNET, vol. 3, nb ½, Jan-Jun 2001.
- [2] B. Rajagopalan et al, "IP over Optical Networks: Architecture Aspects", IEEE Communication Magazine, PP.94-102, September 2000.
- [3] D. Awduche et al, "Multi protocol Lambda Switching, Combining MPLS Traffic Engineering with Optical Cross connects ", Work in Progress, Internet Draft, draft-awduche-mpls-te-optical-01.txt
- [4] N. Ghani, "Lambda-Labeling: A framework for IP-over-WDM using MPLS", Optical Networks Magazine, Vol. 1, PP45-58, April 2000.
- [5] E. Mannie, editor, "GMPLS Architecture", IETF Internet Draft, draft-many-gmpls-architecture-00.txt, March 2001.

- [6] K. Zhu, B. Mukherjee, "Traffic Grooming in an Optical WDM Mesh Network", *IEEE ICC'01*, June 2001.
- [7] Y. Sun, J. Gu, "Traffic Grooming in All-Optical Networks", *IEEE ICC'01*, June 2001.
- [8] O. Gerstel, R. Ramaswami, G. Sasaki, "Cost Effective Traffic Grooming in WDM Rings", *IEEE INFOCOM'98*, pp. 69-77.
- [9] S. Thiagarajan, A. Somani, "Performance Analysis of WDM Optical Networks with Grooming Capabilities", *Proc. SPIE Intl. Symp. on Voice, Video and Data Comm.*, Nov. 2000.
- [10] C. Assi, A. Shami, R. Kurtz, and M. A. Ali, "Optical networking and real-time provisioning; An integrated vision for the next-generation Internet", *IEEE Network Magazine*, July/Aug. 2001, Vol. 15 No. 4, PP. 36-45.
- [11] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks", *Optical Networks Magazine*, vol. 1, pp. 47-60, Jan. 2000.
- [12] C. Assi, A. Shami, Y. Ye, S. Dixit, and M. Ali "On The Merit of IP/MPLS Protection/Restoration in IP over WDM Networks" *IEEE GLOBECOM 2001*, San Antonio, Texas.
- [13] C. Assi, A. Shami, Y. Ye, S. Dixit, and M. Ali, "Integrated Routing Algorithms for Provisioning "Sub-Wavelength" Connections in IP over WDM Networks" *Journal of Photonic Networks*, 2002.
- [14] A. Girard, *Routing and Dimensioning in Circuit-Switched Networks*, Addison Wesley, 1990.
- [15] M. Kodialam, T.V. Lakshman, "Dynamic Routing of Bandwidth Guaranteed Tunnels with Restoration", *Proc. of Infocom, 2001*.
- [16] B. Doshi et al, "Optical Network Design and Restoration", *Bell Labs Tech. J.*, vol.4, no.1, Jan-Mar 1999, pp. 58-84.

- [17] C. Assi, Y. Ye, A. Shami, S. Dixit, I. Habib, and M. A. Ali, "Designing a survivable IP-over-WDM network", OptiComm 2001, PP. 1-11, Denver, Colorado, Aug. 01.
- [18] G. Mohan, C.S.R. Murthy "Lightpath Restoration in WDM Optical Networks", IEEE Network, November/December 2000.

## Chapter 6

### **6. Conclusions and Future Work**

This dissertation studied the problem of provisioning in optical networks and made few important contributions to the body of knowledge in the design and analysis of scalable, reliable, and cost-effective next-generation survivable optical networks. The ideas and techniques developed in this dissertation can be utilized in designing and implementing next-generation WDM optical networks. This chapter summarizes the main results and contributions made in this dissertation.

#### **6.1 Real-Time Provisioning in WDM Optical Mesh Networks**

Real-time provisioning of optical channels in a hybrid IP-centric WDM-based networking model has been presented. Specifically, the work presented here has addressed the implementation issues of the path selection component of the traffic-engineering problem in such a network. Methodologies and associated algorithms for dynamic lightpath computation were outlined.

Here, provisioning of connections in a WDM based network is defined as the process of dynamically selecting efficient end-to-end routes and assigning wavelengths, i.e., routing and wavelength assignment (RWA) and involves the implementation of two traffic engineering components: an information distribution mechanism that provides knowledge of the relevant attributes of available network

resources, and a path selection process that uses the information distributed by the dynamic link-state advertisement algorithm to select a path that meets the specific requirements.

In this study, we presented the dynamic RWA problem as a key requirement for lightpath provisioning. We analytically modeled the conventional RWA problem with different wavelength assignment schemes. Fixed alternate routing (FAR) with/without trunk reservation (TR) is one approach to improve the network performance; analytical models for alternate routing w/o trunk reservation are also presented. We validated the accuracy of the proposed models through numerical comparison with simulation results of the RWA. We also showed the efficiency of alternate routing in improving the performance. However, as the traffic load increases, alternate routing degrades the blocking performance and the improvement of FAR is very limited; this is due to the fact that alternate routing will occupy more transmission resources and thus, at higher loads will prevent direct traffic from being routed and leads to instability. By limiting the flow of the alternately routed traffic, one can overcome the instability problem. Trunk reservation (TR) is one technique to improve the performance of alternate routing at higher loads and can better manage the network resources by selecting the right trunk threshold.

We have also presented and compared the performance of two different constraint-based routing and wavelength assignment (RWA) algorithms, for dynamic provisioning of the optical channels and three path selection schemes have also been proposed for each algorithm. Both routing algorithms are based on a fully distributed implementation. The performance of both algorithms is then compared with that of

the conventional static RWA algorithm. It is shown that the dynamic full-adaptive algorithm outperforms the semi-adaptive one, and both of them significantly outperform the conventional static algorithm. It is also shown that the Future Cost-Based Selection scheme outperforms both the total-based and the balanced selection schemes as well as the first fit (FF) being widely studied in the literature.

## 6.2 Survivability in WDM-based Networks

In this chapter, we addressed the path selection problem as a key requirement in the design of efficient mesh restorable networks. Two path selection algorithms were presented, namely the DPGI that relies on global information maintained at each node and DPLI, where only local information is necessary. We showed that path selection algorithms intended to optimize the resource utilization by maximally exploiting the sharability of reserved channels do not necessarily result in fast restoration. Rather, these two optimality criteria tend to be conflicting. Analytical model for evaluating the performance of shared path protection is presented and its accuracy is validated through numerical comparison with simulation results.

## 6.3 Connection Management for Survivable WDM Networks

In this chapter we presented a novel hybrid distributed fault-management protocol for combating Single-fiber failures in mesh-based WDM optical networks. The proposed hybrid approach combines Link State Protocol to disseminate and update information only about the physical connectivity of the network and a distributed local information-based signaling algorithm for connection management

(setting up/tearing down resources). We have evaluated the performance of the proposed hybrid approach by comparing both the dedicated 1:1 path protection and the shared path protection in terms of blocking probability, connection setup time and restoration time under random failure assumptions. Also, we presented different restoration algorithms as a key requirement for designing shared restorable mesh networks. We proposed two restoration algorithms, namely the source based restoration (SBR) and the destination based restoration (DBR). We also proposed a new method to achieve rapid restoration. This method builds upon the SBR by introducing offset time after which the source can start transmitting data on the protection path. Efficient offset time selection combined with pipelining techniques was shown to achieve rapid restoration that is independent from the protection path length.

#### **6.4 Real-Time Provisioning in IP-Over-WDM Networks**

This work has developed and implemented comprehensive, unified constraint-based routing algorithms within the generalized MPLS framework (GMPLS) to provision “sub-wavelength” circuits (low-rate traffic streams). Constraint routing is further augmented in this work by dynamically routing both an active and another alternate link/node-disjoint backup path at the same time in order to provision a given connection request. Two different routing algorithms to dynamically route low-rate traffic streams in future IP-centric optical networks have been presented. One algorithm attempts to first route connection requests on the logical topology interconnecting the IP routers, then, if blocked, on the underlying WDM physical topology by setting up a lightpath if possible. The second algorithm attempts to route

connection requests on the physical topology if a direct path couldn't be found, then, if blocked, on the logical topology. This approach increases the connectivity of the virtual topology. The performance of each algorithm is then evaluated for three different constraint-based adaptive selection schemes, namely, conventional shortest path selection scheme, least loaded selection scheme, and most loaded selection scheme. To further augment these algorithms, we have also developed a novel *integrated protection scheme* to dynamically allocate restorable-bandwidth guaranteed paths for designing IP/MPLS over WDM networks that can protect against single optical link/node failures.

## 6.5- Future Work

In mesh networks with shared protection, rerouting upon failures generally involves sending restoration messages along restoration paths to reconfigure appropriate cross-connects along those paths. Since each fiber might carry up to hundreds of connections in each direction, the failure of a link might affect a large number of connections. In chapter 4 of this thesis, we addressed the issue of connection and management protocols in survivable optical networks. We presented different restoration algorithms and proposed to use an offset time-based restoration combined with pipelining techniques to achieve fast network restoration times (RT). It was considered that the main parameters that contribute to the RT are the switch configuration times and message propagation delays; we proposed to minimize their impact and derived an expression of the RT that is only dependent on the message processing time (MPT). The effect of the MPT was neglected in our study because

the number of wavelengths per fiber was relatively small (8-16 wavelengths); hence the number of connections that fail simultaneously is also considered small. However, the situation changes when the number of failed connections increases. In this new situation, many restoration messages and OXC configuration commands may need to be processed at the network nodes. This in turn introduces queueing delays and may become a significant component of the RT, unlike what was reported in chapter 4. More importantly, these queueing delays may make it difficult to provide restoration time guarantees to connections which have pre-determined RT requirements. Thus, it is critical to address extensions to the work presented in chapter 4 and investigate the impact of message queueing delays on the overall RT. Moreover, mechanisms to effectively coordinate the re-routing process of connections that fail simultaneously will be required.

The provisioning algorithms studied throughout this thesis have only considered the single network element failure (chapter 3). Multiple element failures are considered to be an important topic for future research. To the best of our knowledge, the only work that has been done in this area only considered the static provisioning; i.e. given the set of traffic demand, the network can be dimensioned to withstand double link failures. Typically, this is not the case. Network traffic tends to be dynamic and hence dynamic online provisioning of a survivable network that withstands multiple element failure is of paramount importance. Our very preliminary investigations led us to believe that the use of hybrid link/path protection with limited resource sharability mechanisms could be best suited for survivable provisioning. We

intend to further study the problem of provisioning with the assumption of multiple element failures. This study will consider the routing as well as the control and management sub-problems.

The problem of provisioning sub-wavelength connections has gained a lot of research interest lately, and we addressed it in chapter 5 where different architectures and routing algorithms have been proposed; however some issues still remain unresolved. For example, while provisioning on the logical layer has definitely its merits, provisioning at the optical layer could also yield better performance. Hence, intelligent, adaptive, QoS routing that dynamically decides on which layer the provisioning should be done are subject for further investigation. Another issue that this thesis did not address is the virtual topology reconfiguration. While we provided a method for the virtual (or logical) topology design in chapter 5, this virtual topology may not be the optimal topology. Studies showed that heuristics could be used to compute the optimal virtual topology (one that maximizes the network throughput) and methods to migrate the current virtual topology to the optimal are required. One objective of this study will be to maximize the benefit (increase throughput) while minimizing the cost (connection disruption times) of the reconfiguration process.

## Bibliography

### Chapter 2:

- [1] C. Assi, A. Shami, R. Kurtz, and M Ali "Optical Networking and Real-Time Provisioning; An Integrated Vision For The Next-Generation Internet", IEEE Network, July/August, 2001.
- [2] N. Ghani et al., "On IP-over-WDM Integration", IEEE Communications magazine, March 2000.
- [3] N. Ghani, "Lambda-Labeling: A framework for IP-over-WDM using MPLS", Optical Networks Magazine, Vol. 1, PP45-58, April 2000.
- [4] B. Mukhejee, Optical Communications Networks, McGraw-Hill, NY, 1997.
- [5] B. Rajagopalan et al, "IP over Optical Networks: A Framework", Work in Progress, Internet Draft, draft-many- optical-framework-02.txt
- [6] N. Chandhok et al, "IP over Optical Networks: A Summary of Issues", Work in Progress, Internet Draft, draft-osu-ipo-mpls-issues-02.txt.
- [7] R. Callon, "A Framework for Multi-protocol Label Switching" Work in Progress, Internet Draft, Sept 1999.
- [8] P. Ashwood-Smith et al, "Generalized MPLS, Signaling Functional Description" Work in Progress, Internet Draft, draft-ietf-mpls-generalized-signaling-01.txt.
- [9] D. Pendarakis, B. Rajagopalan, and Debanjan Saha, " Routing information exchange in Optical networks", Work in Progress , Internet Draft, draft-prs-optical-routing-01.txt.
- [10] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," Optical Networks Magazine, vol. 1, pp. 47-60, Jan. 2000.

- [11] R. Ramaswami and K. N. Sivarajan, "Dynamic Allocation in All-Optical Ring Networks," Proc., IEEE ICC '97, Montreal, Quebec, Canada, vol. 1, pp. 432-436, June 1997.
- [12] S. Ramamurthy and B. Mukherjee "Fixed-Alternate Routing and Wavelength Assignment in Wavelength Routed Optical Networks", proc. IEEE Globecom'98, vol. 4, pp. 2295-2302, Nov. 1998.
- [13] A. Mokhtar and M. Azizoglu "Adaptive wavelength routing in all optical networks", IEEE/ACM trans. Net, Vol 6, pp197-206, April 1998.
- [14] Tibor Fabry-Asztalos, Nilesh Bhide and Krishna M. Sivalingam "Adaptive Weight Functions for Shortest Path Routing Algorithms for Multi-Wavelength Optical WDM Networks", Proc. IEEE ICC2000, New Orleans, LA, June 2000.
- [15] I. Chlamtac, A. Ganz, and G. Karmi, "Purely Optical Networks for Terabit Communication," Proc., IEEE INFOCOM '89, Washington, DC, vol. 3, pp. 887-896, April 1989.
- [16] E. Karasan and E. Ayanoglu, "Effects of wavelength Routing and Selection Algorithms on Wavelength Conversion Gain in WDM Optical Networks," IEEE/ACM Transactions on Networking, vol. 6, no. 2, pp. 186-196, April 1998.
- [17] X. Zhang and C. Qiao, "Wavelength Assignment for Dynamic Traffic in Multi-Fiber WDM Networks", proc. 7<sup>th</sup> ICCCN, pp. 479- 485, Oct. 1998
- [18] J. Yates, P. Rumsewicz, and J. Lacey, "Wavelength converters in dynamically reconfigurable WDM networks" IEEE communication Survey, Q2, 1999.
- [19] C. Assi, A. Shami, Z. Zhang, M. Ali, "Impact of wavelength converters on the performance of optical networks" Optical Network, vol.3, issue 2, March/April 2002.
- [20] H. Harai, M. Murata, and H. Miyahara, "Performance of All-Optical Networks with limited range wavelength conversion" Proc., IEEE ICC, Montreal Canada, 1997, pp. 416-421, April 1997.

- [21] M. Ali et al, "Routing and Wavelength Assignment (RWA) with Power Considerations In All-Optical Wavelength-Routed Networks", Proc. IEEE GLOBECOM '99, Rio De Janeiro, Brazil), pp. 1443-1437, Dec. 1999.
- [22] A. Birman. "Computing Approximate Blocking Probabilities For A Class Of All-Optical Networks," IEEE Journal of Selected Areas of Communication, volume 14, pages 852-857, June 1996.
- [23] H. Harai et al. "Performance Of Alternate Routing Methods in All-Optical Switching Networks," Proceedings, IEEE INFOCOM '97, pages 16-524.
- [24] A. Sridharan and K. Sivarajan. "Blocking in All-Optical Networks," IEEE INFOCOM 2000, pp. 990-999.
- [25] C. Chen and S. Banerjee, "A new model for optimal routing and wavelength assignment in wavelength division multiplexed optical networks" IEEE INFOCOM 96, pp. 148-155. 1996
- [26] F. P. Kelly, "Loss Networks," The Annals Of Applied Probability, Volume 1, Issue 3, pages 319-378, August 1991.
- [27] F. P. Kelly, "Blocking Probabilities In Large Circuit-Switched Networks," Advances in Applied Probability, volume 18, pages 473-505, 1986.
- [28] A. Girad. "Routing and Dimensioning in Circuit-Switched Networks," Addison-Wesley, Reading, Massachusetts, 1990.
- [29] T. K. G. Yum, M. Schwartz, "Comparison of Routing Procedures for Circuit Switched Traffic in Non-Hierarchical Networks" IEEE Trans. on Comm., Vol 35, May 1987, pp. 534-544.

**Chapter 3:**

- [1] B. Doshi et al, "Optical Network Design and Restoration" Bell Labs Tech. J., vol.4, no.1, Jan-Mar 1999, pp. 58-84
- [2] G. Mohan, C.S.R. Murthy "Lightpath Restoration in WDM Optical Networks", IEEE Network, November/December 2000.
- [3] D. Zhou, and S. Subramanian "Survivability in Optical Networks", IEEE Network, November/December 2000.
- [4] P. Demeester et al., "Resilience in Multilayer Networks," IEEE Communication Magazine, vol. 37, no. 8, Aug. 1999, pp. 70-76.
- [5] O. Gerstel, R. Ramaswami, and G. H. Sasaki, "Fault Tolerant Multi-wavelength Optical Rings with Limited Wavelength Conversion," *IEEE JSAC*, vol. 16, no. 7, Sept. 1998, pp. 1166-78.
- [6] O. Crochat and J. Y. Le Boudec, "Design Protection for WDM Optical Networks," *IEEE JSAC*, vol. 16, no. 7, Sept. 1998, pp. 1158-65.
- [7] T. H. Wu, *Fiber Network Service Survivability*, Artech House, 1992.
- [8] C. Assi, A. Khalil, N. Ghani, and M Ali, "Efficient Shared Path Protection in Mesh WDM Networks" *The Journal of Networks, Software Tools, and Applications, Cluster computing*, to appear.
- [9] G. Li, W. Wang, C. Kalmanek, and R. Doverspike, "Efficient Distributed Path Selection for Shared Restoration Connections" *IEEE INFOCOM 2002*, NY.
- [10] K. Kompella et al. "OSPF extensions in support of Generalized MPLS" IETF Internet Draft, work in progress, Feb 2001.
- [11] K. Kompella et al. "IS-IS extensions in support of Generalized MPLS" IETF Internet Draft, work in progress, Feb 2001.

- [12] O. Hause, M. Kudialam and T. V. Lakshman, "Capacity Design of Fast Path Restorable Optical Networks" *IEEE INFOCOM 2002*, NY.
- [13] E. Bouillet, J. Labourdette, G. Ellinas, R. Ramamurthy, S Chaudhuri, "Stochastic Approaches to Compute Shared Path Mesh Restored Lightpaths in Optical Network Architecture" *IEEE INFOCOM 2002*, NY.
- [14] C. Qiao, and D. Xu, "Distributed Partial Information Management (DPIM) Schemes for Survivable Networks-Part I" *IEEE INFOCOM 2002*, NY.
- [15] C. Chen and S. Banerjee, "A New Model for Optimal Routing and Wavelength Assignment in WDM Optical Networks" *IEEE INFOCOM 1996*.
- [16] A. Girard, *Routing and Dimensioning in Circuit-Switched Networks*, Addison Wesley, 1990.

#### **Chapter 4:**

- [1] M. Goyal, G. Li, and J. Yates, "Shared Mesh Restoration: A Simulation Study", OFC'02, March 2002.
- [2] B. Doshi et al, "Optical Network Design and Restoration" Bell Labs Tech. J., vol.4, no.1, Jan-Mar 1999, pp. 58-84
- [3] G. Austin, et. al., "Fast, Scalable, and Distributed Restoration in General Mesh Optical Networks" Bell Labs. Tech. Journal, vol. 6, nb. 1, pp. 67-81, Jan-June 2001.
- [4] C. Assi, Y. Ye, A. Shami, S. Dixit, and M. Ali, "Performance Evaluation of Efficient Path Selection and Fast Restoration Algorithms for Shared Restorable Mesh WDM Networks" Proc. IEEE ICC, Anchorage, Alaska, 2002.
- [5] R. Ramaswami and A. Segall, "Distributed Network Control for Wavelength Routed Optical Networks," Proceedings, IEEE INFOCOM '96, San Francisco, CA, pp. 138-147, Mar. 1996.

- [6] Hui Zang et al, "Dynamic lightpath establishment in wavelength-routed WDM networks," IEEE Comm. Magazine, Vol. 39 No. 9, PP 100-108, Sept. 2001.
- [7] X. Yuan, R. Melhem, R. Gupta, Y. Mei, and C. Qiao, "Distributed Control Protocols for Wavelength Reservation and Their Performance Evaluation," Photonic Network Communications, vol. 1, no. 3, 1999, pp. 207-218.
- [8] J. P. Jue and G. Xiao, "An Adaptive Routing Algorithm with Distributed Control Scheme for Wavelength-Routed Optical Networks," proceedings, Ninth International Conference On Computer Communications and Networks, Las Vegas, NV, Oct. 2000.
- [9] I. Widjaja, A. Elwalid, " Study of GMPLS Lightpath Setup over lambda-Router Networks ", Proc. of IEEE, ICC'02.
- [10] C. Assi, Y. Ye, A. Shami, S. Dixit, and M. Ali, "A Hybrid Distributed Fault-Management Protocol for Combating Single-Fiber Failures in Mesh-based DWDM Optical Networks" IEEE GLOBECOM 2002, Taipei, Taiwan.
- [11] C. Assi, A. Shami, and M. Ali, "Optical networking and real-time provisioning; An integrated vision for the next-generation Internet", IEEE Network, July/Aug. 2001, Vol. 15 No. 4, PP. 36-45.
- [12] J.Y. Wei, R.I. McFarland, "Just-in-time signaling for WDM optical burst witching networks", J. of Lightwave Technology, Vol. 18 no. 12, pp. 2019-2037, Dec. 2000.

## Chapter 5:

- [1] M.A.Ali, A. Shami, C. Assi, Y.Ye, R Kurtz, "Architecture Options for Next-Generation Networking Paradigm: Is Optical Internet the Answer", PNET, vol. 3, nb ½, Jan-Jun 2001.
- [2] B. Rajagopalan et al, "IP over Optical Networks: Architecture Aspects", IEEE Communication Magazine, PP.94-102, September 2000.

- [3] D. Awduche et al, "Multi protocol Lambda Switching, Combining MPLS Traffic Engineering with Optical Cross connects ", Work in Progress, Internet Draft, draft-awduche-mpls-te-optical-01.txt
- [4] N. Ghani, "Lambda-Labeling: A framework for IP-over-WDM using MPLS", Optical Networks Magazine, Vol. 1, PP45-58, April 2000.
- [5] E. Mannie, editor, "GMPLS Architecture", IETF Internet Draft, draft-many-gmpls-architecture-00.txt, March 2001.
- [6] K. Zhu, B. Mukherjee, "Traffic Grooming in an Optical WDM Mesh Network", *IEEE ICC'01*, June 2001.
- [7] Y. Sun, J. Gu, "Traffic Grooming in All-Optical Networks", *IEEE ICC'01*, June 2001.
- [8] O. Gerstel, R. Ramaswami, G. Sasaki, "Cost Effective Traffic Grooming in WDM Rings", *IEEE INFOCOM'98*, pp. 69-77.
- [9] S. Thiagarajan, A. Somani, "Performance Analysis of WDM Optical Networks with Grooming Capabilities", *Proc. SPIE Intl. Symp. on Voice, Video and Data Comm.*, Nov. 2000.
- [10] C. Assi, A. Shami, R. Kurtz, and M. A. Ali, "Optical networking and real-time provisioning; An integrated vision for the next-generation Internet", *IEEE Network Magazine*, July/Aug. 2001, Vol. 15 No. 4, PP. 36-45.
- [11] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks", *Optical Networks Magazine*, vol. 1, pp. 47-60, Jan. 2000.
- [12] C. Assi, A. Shami, Y. Ye, S. Dixit, and M. Ali "On The Merit of IP/MPLS Protection/Restoration in IP over WDM Networks" *IEEE GLOBECOM 2001*, San Antonio, Texas.

- [13] C. Assi, A. Shami, Y. Ye, S. Dixit, and M. Ali, "Integrated Routing Algorithms for Provisioning "Sub-Wavelength" Connections in IP over WDM Networks" *Journal of Photonic Networks*, 2002.
- [14] A. Girard, *Routing and Dimensioning in Circuit-Switched Networks*, Addison Wesley, 1990.
- [15] M. Kodialam, T.V. Lakshman, "Dynamic Routing of Bandwidth Guaranteed Tunnels with Restoration", *Proc. of Infocom*, 2001.
- [16] B. Doshi et al, "Optical Network Design and Restoration", *Bell Labs Tech. J.*, vol.4, no.1, Jan-Mar 1999, pp. 58-84.
- [17] C. Assi, Y. Ye, A. Shami, S. Dixit, I. Habib, and M. A. Ali, "Designing a survivable IP-over-WDM network", *OptiComm 2001*, PP. 1-11, Denver, Colorado, Aug. 01.
- [18] G. Mohan, C.S.R. Murthy "Lightpath Restoration in WDM Optical Networks", *IEEE Network*, November/December 2000.