uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

u Ottawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Atousa Vali Sichani
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Electrical Engineering)
GRADE / DEGREE

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

Signaling Protocols for Survivable All-optical Networks

TITRE DE LA THÈSE / TITLE OF THESIS

Hussein Mouftah
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

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

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Halima Elbiaze

Oliver Yang

Chang Cheng Huang

Amiya Nayak

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Signaling Protocols for Survivable All-optical Networks

Atousa Vali Sichani

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies
In conformity with the requirements for the degree of
Doctor of Philosophy
In Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
School of Information Technology and Engineering
University of Ottawa

January 2006

Your file Votre référence
ISBN: 978-0-494-18610-7
Our file Notre référence
ISBN: 978-0-494-18610-7

TO MY FATHER,

MOHAMMAD VALI SICHANI

AND MY BABY,

GORGEEN

Abstract

Survivable mechanisms, including fault localization and connection restoration, are achieved by means of electronically analyzing overhead bits in opaque optical networks. However, unlike opaque networks, transparent all-optical networks are unable to examine signal quality at transponders. Consequently, locating and isolating failures and attacks (intentional failures) in transparent optical networks are much harder than in opaque networks. Likewise, connection restoration and path rectification are more complex. Therefore, previously defined survivable terminologies, models, and techniques should be evolved and re-designed with respect to the dissimilarities to be applicable to all-optical networks.

Accordingly, this thesis proposes several fault detection protocols, including Fault Localization Signaling Protocol, Rolling-back Signaling Protocol, and Limited-perimeter Vector Matching Protocol, and a number of connection restoration and re-routing techniques, such as Open Link Restoration Protocol, Minimum Weighted-path Restoration Protocol, and Signaling Nested Reservation Protocol, which are feasible in all-optical communication networks.

The proposed protocols are all implemented by signaling in the supervisory channels of the overlay model. Since these protocols use controlling signals only as triggering signs, they are not involved in substantial processes and delays. As a result, they are compatible to high-speed all-optical networks. However, the signaling mechanisms embedded in these protocols utilize additional bandwidth from the control network.

Acknowledgements

I am grateful to my thesis supervisor, Professor Hussein Mouftah, for his constant consultation, collaboration, cooperation, and compassion. The freedom that he has given me during the course of this thesis has been my great responsibility and is appreciated.

I am thankful those professors who have taught me their knowledge. I would like to explicitly acknowledge Dr. H. T. Mouftah, Dr. A. Safwat, Dr. C. Yeh, and Dr. S. Gazor for their contributions to my success.

I am also thankful to my fellows both at Queen's university and at the University of Ottawa for their friendships and companionships. I would like to particularly name Jun Zheng, Baoxian Zhang, Bin Zhou, Ali Nouroozifar, Quan Hum, Nabil Naas, Abdel Hamid Taha, Abdel Hamid Eshoul, Shafiq Hashmi, Shoib Siddiqui, Richard Pramod, Hongxia Wang, Tun Hu, and Hamid Reza Salighehrad.

The persons who deserve lots of my thanks are my sons, Arsham and Gorgeen who have been affected the most by my study. Their persistent love and understanding have been endless and no words could express how much I worship and thank them.

I am also obliged to my husband, Dr. Saeed Gazor who has been very considerate, supportive, and encouraging these years. I could never follow my dreams and achieve my goals without him.

Table of Contents

List of Figures

List of Acronyms and Symbols

AON	All Optical Network
APS	Automatic Protection Switch
ATM	Asynchronous Transfer Mode
BDI	Backward Defect Indicator
BER	Bit Error Rate
BLPR	Bidirectional Link Protected Ring
BRP	Backward Reservation Protocol
BRS	Backward Restoration Scheme
BWS	Bubble-based Wave-guide Switch
CAC	Connection Admission Control
CI	Confidence Interval
DLS	Dynamic Lightpath Establishment
DP	Diagnosis Phase
E/O	Electronic to Optical Conversion
EDFA	Erbium-Doped Fiber Amplifier
FDI	Forward Defect Indicator
FEC	Forward Error Correction
FRP	Forward Reservation Protocol
HFD	Heuristic Depth First Searching
IGP	Interior Gateway Protocol
GL	Generalized Loop-back Protocol
ITSA	Interactive Two-Step-Approach
JET	Just Enough Time
JIT	Just in Time
LAN	Local Area Network
LCS	Liquid Crystal Switch
LM	Link Multiplexing

LVM	Limited-perimeter Vector Matching Fault Localization Protocol
LS	Link State
MAN	Metropolitan Area Network
MEMS	Micro Electro-Mechanical Switch
MWRP	Minimum Weighted-path Restoration Protocol
MRT	Maximum Retransmission Time
NACK	Negative Acknowledgement
NRP	Nested Reservation Protocol
NRUNN	Network Restoration Using Neural Networks
O/E	Optical to Electronic Conversion
OBS	Optical Burst Switching
OCh	Optical Channel Layer
OMS	Optical Multiplex Section
OSC	Optical Supervisory Channel
OSNR	Optical Signal to Noise Ratio
OSPF	Open Shortest Path First
OTS	Optical Transmission Section
OUT	Optical Transport Unit
OXC	Optical Cross-Connect
PCP	Pre-Computational Phase
PLC	Planar Lightwave Circuit
PM	Path Multiplexing
PROB	Probe Signal
PRP	Progressive Reservation Protocol
QoS	Quality of Service
RA	Raman Amplifier
RAID	Redundant Array of Inexpensive Disks
RES	Reservation Signal
REL	Release Signal
RSVP	Resource Reservation Protocol

RWA	Routing and Wavelength Assignment
SAN	Storage Area Network
SDH	Synchronous Digital Hierarchy
SHR	Self Healing Ring
SLE	Static Lightpath Establishment
SNR	Signal to Noise Ratio
SNRP	Signaling Nested Reservation Protocol
SFLP	Signaling Fault Localization Protocol
SOA	Semiconductor Optical Amplifier
SONET	Synchronous Optical Network
SPEM	Shortest Path Eulerian Matching
SPRING	Shared Protection RING
TC	Tandem Connection
TDM	Time Division Multiplexing
TI	Trace Identifier
UPPR	Unidirectional Path Protected Ring
VoIP	Voice over Internet Protocol
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing
MPλS	Multi-Protocol Lambda Switching
$A(t)$	System availability
$A_i(t)$	Element ith availability
A_{Pk}	Path availability
$C^{(1)}$	Cycles with at least one edge in $\bigcup_{k=1}^{n} F^{(k)}$
$C^{(n)}$	$C^{(n-1)} \bigcup \{\text{cycles with no edges in } F^{(1)} \text{ and any edge in } F^{(n)}\}$
D	Network diameter
$E[d_{sd}]$	Average distance between two nodes in a matrix-formed mesh network
$E\,[UB]$	Average unused bandwidth in the BRP protocol
$E\,[SB]$	Average saving bandwidth in the SNRP protocol

$E[T]$	Mean Time To Failure (MTTF)
$E[T_R]$	Mean Time To Repair (MTTR)
$E[T_B]$	Mean Time Between successive Failures (MTBF).
$E^{(i)}$	All graph edges excluding a set of i purged edges, ($E^{(i)} = E - \bigcup_{k=1}^{i} F^{(k)}$)
F	Number of failures
$F^{(i)}$	Edge ith in graph $G^{(n)}$
$G^{(n)}$	An at least 2-edge connected graph with $V^{(n)}$ vertices and $E^{(n)}$ edges
l_{p_k}	Number of links
L	The average length of a real time data packet in the SNRP protocol
L_p	Hop weight
N_{sd}	Number of hops between a source-destination pair
N_d	Number of hops between a severed connection and the destination
p_k	Path kth comprising links $\left(i_{1,k}, i_{2,k}, \cdots, i_{l_{p_k},k}\right)$
$P(n)$	Distribution of source-destination pairs
$R(t)$	System reliability
$R_i(t)$	Element ith reliability
S_p	Switching weight
S_{LVM}	LVM space complexity
S_X	LVM space complexity when the fault detection area was not limited
$S^{(n)}$	A set of all cycles in $G^{(n)}$
T_{LVM}	LVM time complexity
T_X	LVM time complexity when the fault detection area was not limited
T_{BRS}	Time taken to restore a path by the BRS scheme
T_{RBSP}	Time taken to restore a path by the RBSP protocol
T_d	Time taken for the destination to detect a failure

T_r	Time taken for a node to compute a new route (routing delay)
T_s	Time taken for a node to configure a switch (switching delay)
T_c	Time taken for a node to check signal loss condition
T_p	Propagation delay between two adjacent nodes
u_p	Number of added (extra) switches in the MWPR protocol
x_i	Length of a non-real time packets in the SNRP protocol
$V^{(n)}$	Vertices of graph $G^{(n)}$
W	Network multiplexing degree
η	Improved average efficiency in the SNRP protocol
δ	Hop factor in the MWPR protocol
ϕ	Switching factor in the MWPR protocol
α	Average propagation delay between two adjacent nodes in the LVM protocol
β	Computational cycle

Chapter 1. Introduction

1.1 Background on Optical Networks

OPTICAL communication systems date back two centuries ago, to the invention of "optical telegraph" in the 1790s [Hec99]. However, factual optical communication networks emerged when the basic elements of optical communication such as semiconductor lasers and low attenuation optical fibers were invented in the late 1960 [Gof02]. Although the fundamentals of optical systems were completed by this time, the significant mismatch between the peak electronic and optical rates (5 orders of magnitude) created a bottleneck in the way of high-bandwidth pure optical communications. To solve the problem, Delange introduced Wavelength Division Multiplexing (WDM) in October 1970 [Del70]. Even though WDM leads optical transporting systems to a new era, other related issues, including routing, management, manufacture, accounting, protection, and performance analysis, were soon appeared [Muk00]. Since then, enormous time and efforts have been dedicated by the researchers to address these issues.

Heuristic Static and Dynamic Lightpath Establishment (SLE, DLE) [Zan00a] techniques have been proposed to solve Routing and Wavelength Assignment (RWA) problem proven to be a NP-complete problem [Ram94]. A variety of control mechanisms such as Centralized, Distributed, Hybrid, and Hierarchical [Muk96] have been developed to manage the networks. Several switching techniques such as Just-In-Time (JIT), Just-Enough-Time (JET) [Bal02], Multi Protocol Lambda Switching (MPλS) [Ali01], and Reservation Protocols [Yua99] have been devised to set up lightpaths. Different types of optical switches such as Optical Cross-Connects (OXCs), Liquid Crystal Switches (LCSs), 2D/3D Micro Electro-Mechanical Switches (MEMSs), Planar Lightwave Circuit (PLC) and Bubble-based Waveguide Switches (BWSs), with various switching speed, have been built to connect wavelengths [App04]. Various network models namely the Overlay, Augmented, Integrated, and Peer-to-peer have been developed to craft optical

structural designs [Rap01]. Security measures [Med97] and signaling techniques have been developed to filter attacks out [Pat01]. High technological amplifiers such as Erbium-Doped Fiber Amplifiers (EDFAs) and Raman Amplifiers (RAs) [Ram02] have been designed to optically amplify the signal. High-performance monitoring equipments including Photo-detectors and Pin Photo-detectors [Amr97] have been invented to bypass electronic monitoring. Different layers were established within the optical layer to define explicit controlling tasks. Incorporation mechanisms were defined to handle diverse-domain problems [Ago96]. Finally, many standards have been registered to put these technological/operational developments into practice and shape current optical networks with wider bandwidth, faster speed and more reliable service than other wired or wireless networks to end users.

Nevertheless, optical networks still suffer from service disruptions caused by environment damages, human errors/attacks and limited component life-cycles. Whatever the reason, due to the high bandwidth, a single failure in optical networks leads to the loss of several terabits per second. Thus, keeping these networks running within a very limited downtime is an imperative issue that needs to be carefully addressed. Nowadays, a common requirement for optical networks is that a connection to be available 99.999% of the time [Ram98]. The required availability corresponds to a connection downtime of less than 5 minutes in average per year. The practical way to obtain this availability is to make the network, survivable.

Protection and restoration techniques are the major tools for ensuring service continuation or network survivability. In general, protection refers to a system involving reserved/redundant resources, using the physical layer Automatic Protection Switches (APSs) to switch traffic from failed to protection facilities. Numerous protection techniques have been proposed during the last two decades for optical networks, exclusively for Synchronous Optical Network /Synchronous Digital Hierarchy (SONET/SDH) networks. Examples of these techniques in the literature are 1+1, 1:1, 1: N, Self Healing Rings (SHRs), Shared Protection RINGs (SPRINGs) [Ram02], *p*-Cycle [Gro02] and ER-*p*-Cycle [Zha04] which we survey numerous of them in this thesis.

On the other hand, restoration refers to a system based on reconfigurable cross-connects and spare resources, where a restoration layer scheme is used to reroute affected traffic to alternate paths. In this regard, having survivable topologies and mechanisms, which allow source-destination (s-d) pairs to be connected by alternate routes, is one of the principal conditions for maintaining survivability in this type of networks. Surrballe [Sur84], Grover [Gro87], NETSPAR [chn93], Generalized Loop-back [Med99a] and Max 2-Route Flow protocols [Kar03] are some of the introduced restoration protocols for optical networks.

Although the introduced dynamic restoration/rerouting mechanisms are all bandwidth-efficient and work well for optical networks, they are not compatible to transparent all-optical networks. This is primarily due to the nature of all-optical components which lack the capability of monitoring and analyzing signals optically [Kob96]. This incapability of examining signals at intermediate nodes makes fault detection and restoration procedures completely different from other types of optical networks. Thus, the current optical network terminologies, models and protocols should be customized to consider the latest concept and features of transparent all-optical networks [Mas03].

This thesis considers transparent all-optical network as the most promising communication networks for constructing core networks and based on this consideration, proposes new restoration/rerouting protocols applicable to all-optical systems. Also, looking at the optical layer as the most efficient layer for restoration purposes, it proposes novel fast fault detection/localization mechanisms that operate within this layer.

1.2 Motivation

In particular, keeping transparent all-optical networks running unaffected, in the event of node and link failures, is a crucial task. Accordingly, reliable recovery services, including high-efficient fault detection/localization and restoration/rerouting mechanisms similar to those provided by SONET, are required to protect transparent optical networks. Fault detection/localization procedure in SONET is accomplished by examining frame overhead bytes and also through signal monitoring in the electronic layer. Similar techniques are neither efficient nor attainable in the transparent all-optical networks. In

general, whenever a link is severed or a channel is disconnected in these networks, the edge routers may be able to detect the existence of a fault, but the exact location of the fault cannot be indicated. Although, electronic monitoring devices could be also deployed in all-optical networks, inserting these tools removes the goodness of pure optical communications. To our best survey, no suitable all-optical fault detection/localization technique exists at this time. Hence, it is important to develop compatible and reliable fault detection/localization methods to promptly pinpoint faults and consequently, facilitate restoration procedures of transparent optical networks. In addition, to make a better use of the available bandwidth and compensate the fault detection/localization delays, more resourceful and speedy rerouting/restoration techniques should be designed.

1.3 Problem

Even though protection and restoration mechanisms for optical networks have been developed and enhanced particularly during the last decade, the same techniques for transparent all-optical networks are not always applicable and have not been mature yet. For instance, certain issues such as Routing and Wavelength Assignment (RWA) and wavelength conversion make WDM SHRs different from those of SONET. In contrast to SONET, which electronic monitoring is possible through Digital cross Connect (DXCs) to detect low Signal to Noise Ratio (SNR) and Bit Error Rate (BER), electronic monitoring is not possible in all-optical networks using all-optical switches/ Optical Cross Connects (OXCs). The monitoring shortage makes it impossible to identify the location of failures. As a result, available rerouting/restoration mechanisms could not reach their goal and ultimate level of efficiency. Another problem, which emerges in the absence of fault localization techniques, is caused by diverting traffic blindly through alternate paths. Rerouting a set of affected paths in this case (without amending the routing tables) might result in sudden chaos or cause further blocking (in case of examining a restoration path over the disconnected link). All the above-mentioned shortages and drawbacks in the optical layer have led to the use of a higher or a combination of higher layers restoration mechanisms. However, these higher layer mechanisms are opaque and too slow for the all-optical functions. In addition, their

employments require coordination between layers. Therefore, to preserve the transparency and high bandwidth of all-optical networks, the stated problems need to be solved through improving technological aspects, such as the deployment of advanced optical monitoring devices and/or be resolved by introducing enhanced survivable algorithms and mechanisms.

1.4 Thesis Objectives

Our objective is to contribute to the design of survivable all-optical networks, with the focus on the design and performance evaluation of fault localization and restoration mechanisms. Taking into account the shortage of all-optical networks in monitoring and analysis and attempting to cope with it have set the objective of this thesis as follows.

- To devise advanced, transparent fault detection/localization mechanisms within the optical layer. In particular, we intend to speed up the pace of restoration techniques that use fault localization techniques as perquisites.
- To propose compatible and competent restoration techniques for all-optical networks.
- To enhance the bandwidth usage by introducing more bandwidth efficient rerouting protocols.
- To investigate the performance of the proposed protocols in comparison with other introduced techniques based on numerical results from analytical models and/or by studying simulation outcomes.

1.5 Thesis Contributions

Restoration is achieved through different layers in optical networks, but it is preferable to recover all-optical networks in the optical layer to preserve the full transparency and maintain the high speed/bandwidth. The challenge, therefore, is to devise fault localization protocols which operate in the optical layer, instead of relying on other higher layers to provide protection and restoration mechanisms such as SONET, IP, and ATM. To do so, we proposed several fault localization protocols [Sic04a], [Sic04b], and

[Sic05f] that detect and localize failures by means of disseminating control signals in the control channels of the overlay model.

Rerouting affected traffic can be accomplished through redundancy either by expanding the network resources or by narrowing down the network bandwidth and saving part of the bandwidth for protection usage. However, some of the networks structural designs either do not allow adding more elements to existing infrastructures or become uneconomical. Therefore, it is essential to develop efficient rerouting and bandwidth allocation techniques that dynamically discover and utilize the spare resources in the networks. In order to contribute to the progressive research in this area, we have introduced a number of novel bandwidth allocation protocols [Sic03a], [Sic03b], [Sic03c], and [Sic05a]. We have also proposed [Sic05c], [Sic05d], and [Sic05e] restoration mechanisms for more efficient recovery.

1.6 Thesis Outline

This thesis has been organized in eight chapters. Chapter 2 discusses background material on survivable optical networks. It also surveys fault detection/localization mechanisms, protection techniques, and restoration protocols along with associated advantages and disadvantages in the literature. In Chapter 3, we introduce a novel bandwidth allocation technique. Next, a new dynamic restoration protocol is introduced in Chapter 4. Chapter 5 also proposes a high-efficient restoration scheme. Chapters 6 and 7 introduce new fault localization mechanisms. Finally, Chapter 8 concludes the thesis with concluding remarks and future research.

Chapter 2. Signaling Protocols for Survivable Networks

2.1 Introduction

IN this chapter we survey signaling protocols devised for survivable optical networks to guarantee service connectivity in the presence of failures. These signaling protocols include failure localization, protection, and restoration mechanisms. This survey comprises those signaling techniques that we came across during the course of this thesis.

2.2 Signaling Fault Detection/Localization Protocols

Moy [Mac00] introduced one of the earliest fault detection protocols in Open Shortest Path First (OSPF) protocol. Fault detection in OSPF is achieved by transferring/receiving detection signals called "Hello messages". Figure 2.1 shows the Hello signals communicating between neighboring routers to obtain the line/node status. These signals are propagated by a predetermined small interval namely, the Hello interval. A smaller Hello interval permits faster connectivity detection. However, smaller Hello intervals consume significantly more bandwidth than longer ones. Furthermore, using a long Hello interval for an extended network to different domains/partitions, will delay the responsiveness of the network in determining the connectivity status. Thus, the number of network partitions is a deciding factor for choosing an optimal Hello interval. In addition, a Hello interval should be chosen in correlation to another parameter known as the router dead interval [Wol95]. Normally, the default router dead interval is four times the chosen Hello interval. By standard, the Hello signals are transferred every 10 sec. After four missing Hello messages, a failure is detected. As a result, it takes at least 40 sec. for the

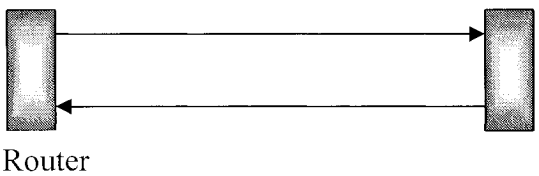

Figure 2.1: Fault detection in the OSPF protocol.

OSPF Hello protocol to detect the occurred failure. Although this method is practical, its speed is far below the required speed of fault detection in high bit rate networks especially optical communication networks.

Similar technique to the OSPF Hello protocol is applied for Voice over Internet Protocol (VoIP). In this network, short and low-bandwidth control packets, called *keep-alive* packets, monitor the network frequently to help rapid switchover of calls [Kar04].

Fault detection also accomplished by means of a master node called dispatcher in communication networks. This method monitors network nodes and associated links by means of parallel centralized signaling as it is illustrated in Figure 2.2. One of the known protocols in this type is *ping-pong* protocol mainly used in submarines. This technique is also called *master-slave* and *pingger-pongger* [Lu-04]. However, this method is neither scalable nor efficient because of employing a centralized dispatcher. Furthermore, returned low power signals are sometimes wrongly interpreted as disconnections.

Results from stochastic are also used to indicate the location of a failure. This notion of discovery has been described by terms such as *knowledge extraction, database exploration, data-pattern processing, information harvesting and data dredging*. The process of fault detection applied to large quantities of data, finding implicit knowledge hidden in the information. Many machine-intelligence and information techniques including *pattern analysis*, *induction decision trees*, *clustering*, and *time series analysis* are used to facilitate data processing [Tso04]. These methods are usually applied to power systems subjected to disturbance. However, they are considerably time-consuming and not always reliable. As a result, they cannot be tuned into very high bandwidth

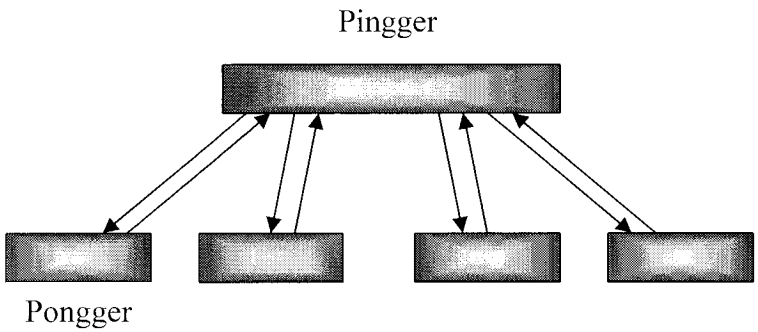

Figure 2.2: Fault detection in the Pingger-pongger protocol.

networks. Over the last two decades, several uncertain and probabilistic reasoning frameworks have also been introduced, however their context is out of the scope of this thesis. Interested readers can find a comprehensive review on them in [Laz92].
Although the introduced protocols have been put in practice for decades, none of them is compatible to optical networks and satisfies the required speed and reliability level. Thus, specific survivable signaling protocols were devised during past decades for fast fault localization in optical networks.

2.2.1 Fault Detection in SONET

Fault detection/localization in Synchronous Optical Networks (SONET) is achieved by overhead inspection. From an end-to-end perspective, the signal carries overhead information but client does not see this overhead. The role of this overhead is to ensure that the connection is flawlessly managed before any disruption and can be recovered afterward. Some of the SONET overhead sections are as follows:

- Trace Identifier (TI) is used to verify that the signal is transported to the correct destination. This term would be particularly important in WDM networks, as the signal may be converted from one wavelength to another.
- Forward Defect Indicator (FDI) conveys failure information to nodes downstream. A single failure may cause multiple alarms to be generated and consequently multiple/improper actions to be taken in response to the reported failure. Thus, it is essential to recognize the root-cause alarm and suppress the remaining alarms. This can be accomplished by using a special signal. FDI signal is used to suppress false alarms produced by downstream nodes. In all-optical networks the FDI is difficult to use in overhead because all-optical nodes are incapable of examining the overhead. In this type of networks FDI should be carried by the control network (OSCs), which has the ability to read the overhead. The FDI signal is sometimes referred to as the Alarm Indication Signal (AIS).
- Backward Defect Indicator (BDI) is a response by the end terminal acknowledging the notification of a fault condition in networks. It is used in conjunction with FDI.

Similar to FDI, it should traverse through Optical Supervisory Channels (OSCs) in case of WDM networks.

It is notable that separate FDI and BDI signals are needed for different sublayers in the optical layer; for instance, to distinguish between fiber failures and channel failures, or to differentiate between the failure of a section of the link between amplifier locations and that of the entire link. Another major reason to use the fault indicators is to prevent triggering other nodes from protection switching. For example, nodes adjacent to a failure detect the failure and may invoke a protection switch to reroute traffic around the failure. At the same time, other nodes downstream or upstream may decide to reroute traffic through other alternate paths. Receiving FDI ensures a node that the task of rerouting traffic is taking care of by a particular node. This way other nodes will relax.

2.2.1.1 Associated Issues

In SONET when a failure condition is detected, it is communicated to the network management entity responsible within the domain. It is also communicated to the neighboring nodes to inhibit them from false failure detection and reporting. For example, when a link in a path is severed, it will impact all downstream nodes by the interruption of data flow. The network management entity initiates a remote testing, fault correction, fault localization, isolation, and restoration procedures. The disconnected node takes proper action and informs the management of the results. Therefore, an effective WDM transport network should consider some critical factors for taking proper actions [Ram02].

- Influences due to nonlinearities, dispersion, and wavelength-to-wavelength interactions are not falsely reported as an interruption of service.
- Interoperability, so that as the optical signal is transported from one provider domain to another, monitoring different signal parameters is done with the same accuracy.
- Payload transparency, so that clients use the same network for a variety of payload types.
- Signal survivability, so clients could rely on the provided service.

2.2.2 Digital Wrapper

The SONET frame is divided into two fields, one that comprises overhead and one that contains data. The SONET overhead is also partitioned into two sections, line and path related overhead. In optical networks, there is a relationship between Signal to Noise Ratio (SNR) with Bit Error Rate (BER). In general, the more degraded the OSNR, the higher the BER. Therefore, to ensure that the quality of signal at the receiver will be at the exact level, a strong error correction code is added. Such powerful error correction code is the Forward Error Correction or FEC. FEC is based on Reed-Solomon error correction code. This code has a total length of 255 bytes, 238 bytes of which are data. The code is able to detect 16 errors and corrects 8 errors. In optical transport networks (long-haul), the fundamental philosophy of SONET frame has been adopted but with a more powerful suitable protocol for high WDM rates, known as *digital wrapper.* The importance of digital wrapper lies in FEC, Tandem Correction (TC) function, path level protection, and high bandwidth communication channels. Each digital wrapper frame consists of 16 subframes called an Optical Transport Unit (OTU). A subframe includes three blocks: a 1-byte block for operations, administrations, and maintenance of the optical channel; a 283-byte block of data in which the client data is mapped to higher layers (SONET, ATM, and IP); a 16-byte block FEC code. Like SONET/SDH, OTU requires the long sequences of "1"source or "0"source be avoided. This is achieved by deploying a scrambler. The digital wrapper defies three sublayers; the path (OCh-P), the tandem connection (OCh-TC), and the section (OCh-S). The OCh-P includes features for optical channel ID, fault indicators (FDI/BDI), signal quality monitoring, and backward quality indication (SQM/BQI). The last functions provide the capability of isolating the source of degraded performance and verifying quality of service. However, current optical line systems protect the optical multiplexed section layer (that is when all wavelengths fail due to fiber cut) and there is no protection for one or a group of wavelengths. This fact excludes certain topologies such as ring with Optical Add and Drop Multiplexers (OADMs).

The OCh-TC includes features for tandem connection management (ID, FDI/BDI, TCM, and SDM/BQI). A mismatch between the incoming and outgoing trace bytes causes an alarm. The trace's ID provides the ability to verify connectivity through connected basics, such as OADM, OXC. Tandem Connection Maintenance (TCM) maintains a channel in a subnetwork that does not include the OCh termination elements. The OCh-S processes associated functions include features for framing, FEC, performance monitoring. To execute FEC, an OTU is interleaved into 16 subframes and FEC is operated on each subframe and over all bytes (255 bytes). Although, digital wrapper greatly improves the BER, it also consumes bandwidth by approximately 7% . Figure 2.3 illustrates the optical transport unit in SONET.

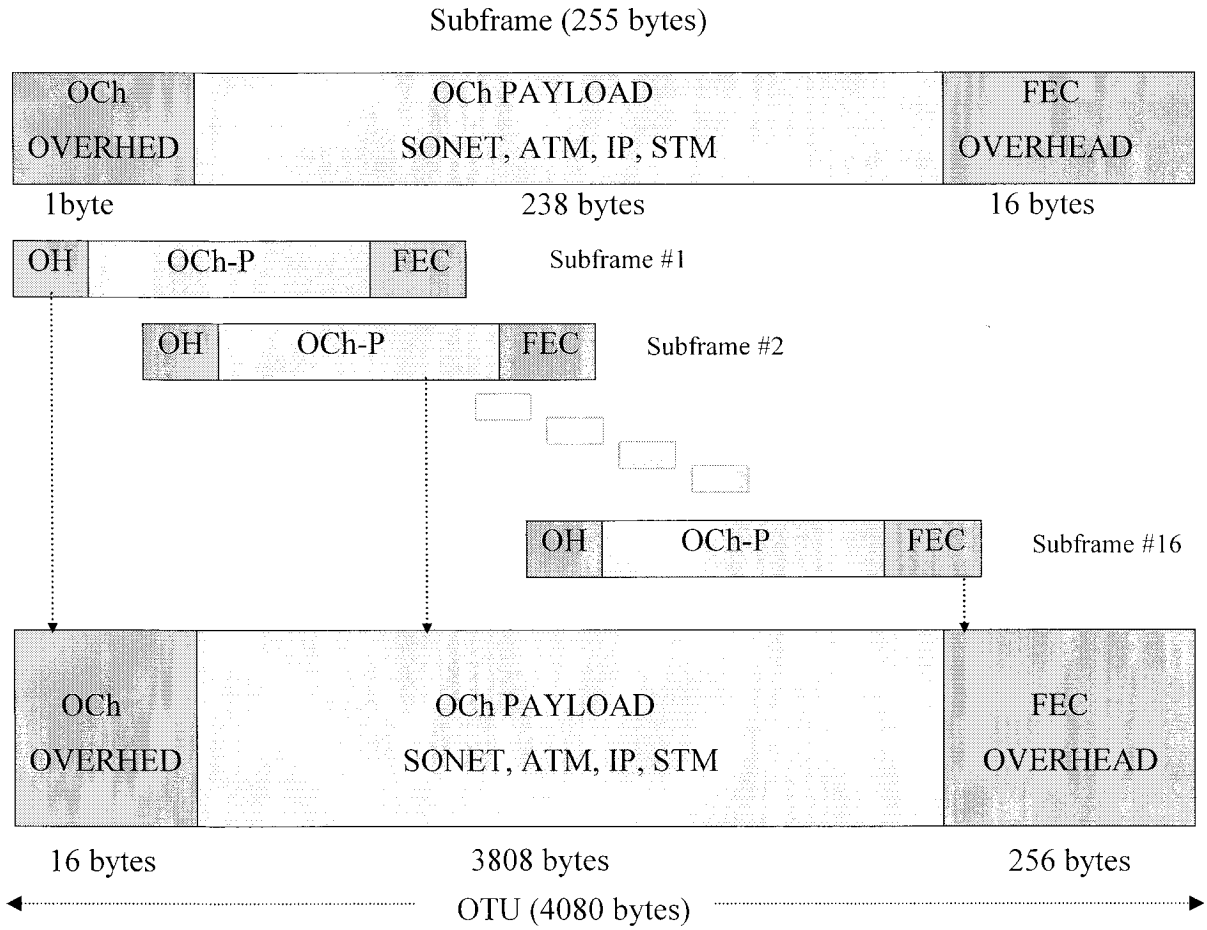

Figure 2.3: The optical transport unit in SONET consists of 16 interleaved subframes.

2.2.3 Mesh Optical Networks

For mesh optical networks fault detection/localization is accomplished in different layers. Related research papers are very limited in this area. One of the first works detects hard failures within the physical layer. Optical monitoring components (e.g., power meters) generate alarms when signal loss is detected. This work tries to optimize the number of monitoring components. Then, it formulates the problem and finds an algorithm to solve it using an alarm matrix. Finding an optimal alarm matrix is equivalent to removing a set of redundant monitors from the network [Sta02].

A work in [Hai04] proposes to partition a network into sub-networks called "islands" and discovers a faulty situation "a node or link failure" with *island-by-island restoration protocol*. Island identification is an off-line procedure, executed during network planning, and occasionally updated when the network topology is changed. Three different partitioning methods namely minimal island, shortest path island, and 2-stage island are defined. However, the protocol is not able to explicitly define the failure location.

Another work that operates in the optical layer suggests decomposing a network into a set of cycles. All nodes and links of the given network should be covered at least by one cycle. These cycles are called *monitoring cycles*. A transceiver is assigned to one node in each monitoring cycle. Therefore, a loop-back dedicated supervisory channel is set up. Network performance in terms of optical power, optical spectrum, optical SNR, high BER can be measured within each monitoring cycle. Once a fault occurred in any element, it triggers an alarm in the affected cycle. Two algorithms, Heuristic Depth First Searching (HDFS) and Shortest Path Eulerian Matching (SPEM) are proposed to find monitoring cycles. The proposed methods decrease overall cost of monitoring equipments. However, they both increase the computational complexities [Zen04].

A fault detection/localization technique has been developed in [Mas00] called Fault Location Algorithm (FLA). To speed up the fault localization protocol, the fault localization process is decomposed into two phases. In the first phase based on the established paths, a binary tree is generated. Since this phase is executed independent of receiving any alarms, in advance, it is called Pre-Computational Phase (PCP). In the

second phase or Diagnosis Phase (DP), binary tree is traversed based on the alarms received from the active components. The protocol is able to detect multi-failure occurrences. However, the protocol could not detect failures associated with empty leaves. In addition, the protocol model and component definitions are not applicable to transparent WDM networks [Mas03]. Other related works using spectrum analyzers and photodiodes could be found in [Kob96], [Amr97], [Med98], and [Pat01].

In addition to those protocols, a finite state machine method is proposed in [Li-97] but its complexity for large-scale and dynamic networks impedes its deployment. Also, ITU is currently in the process of setting standard using GMPLS for fault detection on all-optical networks.

2.3 Optical Layer Protection Techniques

Optical layer protection schemes are similar to SONET techniques. However, their implementation is substantially different. The optical layer consists of the Optical Channel (OCh) layer also known as path layer, the Optical Multiplex Section (OMS) layer (line layer), and the Optical Transmission Section (OTS) layer.

SONET protection techniques cover either line layer such as BLSR or path layer like UPSR. Optical protection schemes handle the OCh and OMS layers. The OCh protection scheme deals with lightpath restorations, while the OMS protection technique restores the entire group of affected wavelengths comprising a link. Figure 2.4 illustrates different

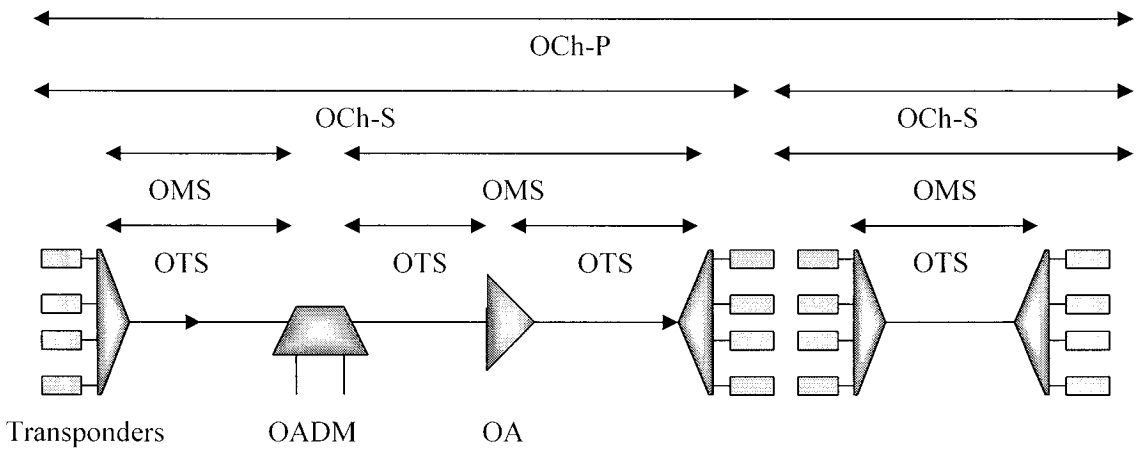

Figure 2.4: The OCh, OMS, and OTS layers within the optical layer.

layers that protection techniques can function within them. In SONET the cost of processing the path layer compared to the line layer is not significant. In contrast, there is a significant cost-difference between the OCh layer schemes and relative OMS layer schemes in WDM. In fact, the cost of OCh protection grows linearly with the number of channels that need to be protected. On the other hand, in SONET Linear Terminal Elements (LTE) and Optical Add, Drop Multiplexers (OADM), and Digital Cross-connects (DXCs) perform the protection in the network layer while Optical Cross-connects (OXCs) play the protection role in linear, ring, and mesh networks at the wavelength level.

2.3.1 Advantages and Disadvantages of the Optical Layer Restoration

There are extensive protection mechanisms in the SONET layer as well as in higher layers such as IP and ATM. These layers are all designed to work independently and not rely on protection/restoration mechanisms supported by other layers. Nevertheless the provided protection mechanisms exist in the higher layers, there are several advantages for WDM networks to be protected by the optical layer.

- The optical layer protection offers fast protection functions, which IP and ATM networks are unable to sustain at the same level. IP traffic is supported only for "best effort" traffic protection. In addition, ATM protection techniques are time-consuming.
- The optical layer protection is more cost-effective. This is because of sharing a single protection wavelength for a group of working wavelengths. However, this technique can only handle a single link cut at the interruption time. Of course it is more efficient to have fiber cuts protected by the optical layer, since a single switch then could restore all the channels.
- The optical layer protection makes more efficient use of protection/restoration bandwidth because of resource sharing among different service layers [Zho00].
- The optical layer protection eliminates part of the mapping computational complexities by taking actions as the first recovery layer [Str01].

- Without the optical layer protection, a single fiber cut generates a large number of alarms and restoration requests that the network management needs to deal with. Instead, if the optical layer were to restore the failure, fewer entities have to be rerouted. Therefore, the process becomes faster and simpler.
- The optical layer protection can be used to support an additional degree of resilience in a network.
- The optical layer protection reacts faster in alarming circumstances than other layers, which removes *race condition* (if the overall network is not carefully engineered, in the event of a failure occurrence, all layers will activate their protection mechanisms simultaneously) [Gre00].
- With the optical layer protection a variety of protection/restoration schemes are being developed for different topologies. This is in contrast with the SONET network, which is based on ring protection.
- Employing the optical layer offers multiple classes of service.

However, the optical layer has some limitations including:

- Not all failures can be handled by the optical layer particularly those related to client equipments.
- The optical layer may unable to distinguish between loss of light and low signal-to-noise ratio. For example, a transparent network can only monitor the presence or absence of power.
- Protection routes may be longer than the working routes. Also protection paths may pass through different networks owned individually, which requires extra budget for interworking and coordination.

2.3.2 (1+1) Optical Protection

Optical protection can be implemented in two different layers, namely OMS and OCh layers. With (1+1) OMS protection layer, service is only protected from link failure. This scheme uses an optical splitter to bridge the optical signal and dual feed it on diversely routed working and protection fibers. The protection switching is entirely performed in

the optical domain. Because there are no duplicated electronics, a failure of the transmitter or the receiver leads to loss of service.

On the other hand, (1+1) OCh layer protection can protect individual wavelengths. The client signal is divided into two disjoint lightpaths at the source. The working and protection lightpaths carry the same signal. The destination chooses the better signal. Both (1+1) protection schemes do not need any signaling, which makes them simple to implement. However, they are not bandwidth-efficient because protection bandwidth is not shared among multiple connections.

2.3.3 1:1 Optical Protection

In (1:1) optical protection (see Figure 2.5), the client signal is not bridged permanently to both the working and protection fibers. Instead a switching occurs once a failure is detected. An Automatic Protection Switch (APS) signaling channel is used to coordinate the protection switching in the event of a failure. In (1:1) optical protection approach, no communication path exits over the protection fiber because optical layer does not accommodate duplicate electronics. Whenever a failure happens, the receiver sends a message to the transmitter over the protection fiber. As a result, it does not receive any reply back. When data is sent through protection fiber, the service is restored. This

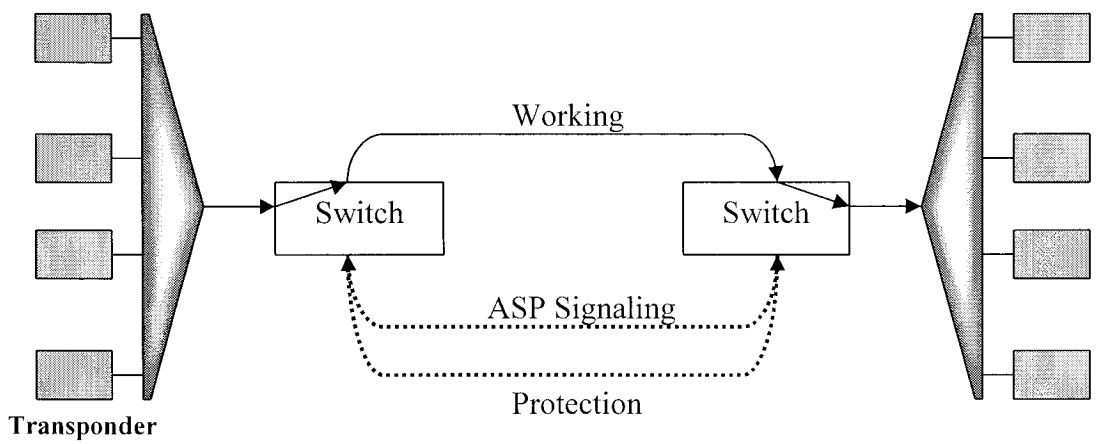

Figure 2.5: (1:1) protection in the OMS layer.

scheme is more cost-effective than (1+1) optical protection mechanism. Furthermore, if a separate "out-of-band" channel is provided or additional signaling is applied, it can support low-priority traffic simultaneously.

2.3.4 (1: N) Transponder Protection

Both OMS and OCh protection schemes can restore the interrupted service by using a spare fiber/wavelength for every N fibers/wavelengths. Of course, in this type of network only the highest priority traffic will be protected once a failure occurs. Although theses techniques can handle link/path failures as well as node failures, it is not able to restore the end equipment failures, particularly transponders. Transponders can be protected in a (1: N) configuration by having a spare transponder for every N working transponders. One problem to overcome is that transponders today operate at a fixed wavelength. As a result, the set-a-side transponder may operate in different wavelength than the working transponder that does not function properly.

2.3.5 Self-Healing Optical Rings

WDM ring networks have 100% restoration capability. There are some challenges involved in WDM SHRs. In contrast to SONET SHRs, if the ring segments separating OADM are long, Optical Amplifiers (OAs) are required. These OAs are needed in both working and protection connections. When no fault is present, the protection fibers are configured in closed loops, so spontaneous emission of noise circulates in closed protection loops. WDM SHRs are either path switched or line switched. The path-switched rings are referred to as Unidirectional Path Protected Rings (UPPR) and the line-switched rings are known as Shared Protection Rings (SPRING). Another mechanism that is usually employed over WDM SHRs is *loop back*. In loop back two nodes adjacent to the failure are responsible for switching the affected traffic to the protection ring [Med1999]. Loop back mechanism is also applicable for recovery from node failures. However, sometime it leads to inefficient unnecessary examinations of all alternate loops.

2.3.6 Unidirectional WDM Path Protection Ring

In a UPPR the nodes are connected in a ring configuration using two fibers that propagate client signal in opposite directions around rings. One fiber is designed as a working fiber and the second one is for protection. The self-healing is achieved by using (1+1) protection. The signal is bridged at the transmitter by a splitter, resulting in two identical copies in two opposite directions. The receiver will switch to the protection ring if a failure happens. Signaling channel is not required since the receiver does not need to notify the sender of the fault condition. Because there is no duplicate electronics, a failure in transceivers results in data loss.

2.3.7 Bi-directional WDM Shared Protection Ring

There are two types of bi-directional SPRINGs, namely two-fiber and four-fiber SPRING. In four-fiber SPRING, two fibers are dedicated to working connections and the other two fibers are assigned for protection purpose. The working traffic travels around the working rings in both directions. The protection rings carry nothing unless a failure happens. To recover from a failure, the protection ring uses a ring switching. Then, when a segment of a ring is cut, the two end nodes close their switches and interconnect the working fiber to the protection fiber. As a result, the traffic is moved to the protection fibers away from the failure. An APS channel is required to coordinate the signaling at both ends. Only traffic that passes through a failed node can be restored. Traffic that originated or terminated at the failed node cannot be recovered.

In two-fiber SPRING, there are two rings, half of the capacity in each ring is used for working traffic and the other half of bandwidth is employed for protection purposes. In the event of a failure, similar to four-fiber SPRING, the working fiber is connected to the protection fiber through APS. It is notable that protection fiber does not carry any traffic prior to the failure occurrence. This scheme also needs APS signaling to be implemented. Like four-fiber SPRING, it cannot cope with transceiver failures.

2.4 Shared Protection Techniques for Mesh Networks

Suurballe introduces one of the first shared protection methods in early 80's [Sur84]. The algorithm is famous for its polynomial computational complexity for solving optimal disjoint paths in terms of the sum cost of the two paths in directed graph. Suurballe's Algorithm uses the link-state for deriving both a working and the corresponding protection paths. Thus, it is mostly useful for dedicated path protection. Another protection protocol is NETworkS Preplanned Automatic Restoration (NETSPAR) [Coa90]. NETSPAR is a hybrid scheme, as it employs a distributed fault detection protocol to determine failure, together with a centralized routing controller. The centralized routing controller allows re-routing policy to be much more deterministic than a fully distributed solution. The network can re-configure itself to a preplanned configuration to cover specific failures. This has the advantage that once the failure has been identified, the appropriate reconfiguration can take place without any additional time spent in generating a new configuration, as this will have already been computed in advance. The disadvantage is that if the failure type is not covered in the failure set, a compromise configuration has to be chosen which may not be optimal. This method works well assuming that the fiber system has extremely low errors and failures. For multi-failure, the algorithm may not perform adequately, however the chance of multiple failures combinations being in the failure set is much lower than for single failures.

Network Restoration using Neural Networks (NRNN) is also a hybrid approach but in contrast to NETSPAR, it does not require all nodes to store network information. Furthermore, it solves the restoration problem in a distributed way. As a result, the required memory units are less (1% of what in NETSPAR) and recovery is much faster than the counterpart [Wan94].

In recent mesh networks, recovery from failures is mostly implemented by link (line) protection. In this type of networks, under normal condition, switches connect working fibers to the OXCs. If a link fails, the protection switches at the end of the failed link move to their protection state. This automatically reroutes affected traffic around the failed link. The link protection uses *pre-configured protection cycles* (*p*-Cycle), which

could be computed off-line before the network is enfaced a failure or on-line after a link failure. The key to the successful protection operation of this protection technique is to find a set of protection cycles that cover the network links. Link protection is possible if a *p*-Cycle including the link is defined by the protection scheme [Gro02]. The dynamic (on-line) *p*-Cycles are formed by the spare capacity of a network. The set of link *p*-Cycles is chosen such that for every link the working connections are protected by *p*-Cycles of spare capacity [Sch02]. A study has revealed that for convertible WDM networks spare resource utilization for *p*-Cycles is about half of the working resources, while non-converting WDM networks reach a spare to working ratio of 71% [Sta00]. This scheme can be implemented either in centralized or distributed manner. The distributed approach is more scalable and faster but also more complex. Several variant of distributed schemes are introduced. In one type the failure information is flooded to all the network nodes. Each node then updates its table and also reconfigures its switches (if it is involved in failure) to recover the affected traffic [Ram99a]. Another possibility is to signal only the affected source-destination pairs to find the alternate routes in the first step and report the failure to other nodes in the next step [Cae98].

Another approach is to divide the network into multiple domains (in extreme case each domain is a ring) then protect each domain independently. As a result, when a failure occurs, it only disrupts a specific domain that can be reported to other domains [Zho00].

A new heuristic called the Efficiency Ratio of a unity-*p*-Cycle (ER-*p*-Cycle) was introduced in [Zha04]. The number of spare units of a unity-*p*-Cycle is equal to the number of spans on the cycle and the ER-*p*-Cycle is the ratio of the number of working units that are protected by the unity-*p*-Cycle to the number of spare units of the unity-*p*-Cycle. The problem, then, is to maximize ER-*p*-unities.

2.4.1 Linear Programming in Shared Protection Techniques

Researchers have noticed that before the best protection path is driven for a working path, a constraint in resource sharing must be imposed. This constraint is stated as *sharable spare capacity* for a protection path or the *dependency* of the protection path from its

working path. The dependency constraint imposes different design criteria such that most of the previously developed techniques were not able to address. One of the best-known solutions is the Two-Step-Approach [Tap01], which in general, divides the problem into two parts. In the first step, it finds a working path and in the next step it derives a protection path for the working path. The Two-Step-Approach is modified in different ways to introduce optimal solutions for protection purposes. In [Dat01], a scheme, called Pool-Reserved Backup Sharing (PRBS), is introduced that only optimizes the protection path problem. PRBS assumes that working paths are routed optimally. This method considers dependency and finds a protection path for a given working path from a backup path pool exclusive working paths.

To address the allocation of working paths, the study in [Xin02] inspects available shortest paths between source-destination (s-d) pairs one after another until the least-cost working and protection paths are realized. However, this method could be time-consuming for networks with numerous alternate paths between s-d pairs.

The study in [Bou02] proposes a method that assigns predefined costs to the sharable capacity links. The costs are defined by the probability of a link is used by a group of affected working paths in the event of a failure. Then, the algorithm attempts to maximize sharable capacities along with minimizing the costs. In [Xu-02], the metric for each link is defined by spare capacity and therefore, it minimizes the spare capacity.

In addition to the end-to-end protection mechanisms, there are numerous studies suggesting segmenting each working path and search for the protection segments to derive an alternate protection path. In [Sar02], a heuristic algorithm is proposed that considers the link-bandwidth and disregards the shared bandwidth. As a result, the scheme is not highly bandwidth-efficient.

A framework in [Kod01] has suggested that the faulty segment is detected and removed from the working path by its replacement protection segment. The performance of this method is segment-related and highly related to network topologies.

In [Ho-03], a protocol, called Short Leap Shared Protection (SLSP) along with a dynamic routing protocol named Cascade Diverse Routing (CDR), is proposed to perform segmented shared protection, in which investigating of k-shortest paths in each segment

is done. This scheme suffers from a large computational complexity but it still outperforms path-based shared protection mechanisms.

In [Ho-02], two heuristic techniques are proposed that improve the performance of Two-Step-Approach. The first scheme called Interactive Two-Step-Approach (ITSA) inspects k-shortest paths as the working path. It considers that with SRLG constraint, protection paths cannot share spare capacity if their working paths have a mutual link. Therefore, it solves the two paths one after another, where the working path is solved first and followed by its protection path on the residual network topology with the associated spare link-state in iterations. Spare link-state is a metric that is defined by considering the same SRLG.

Because of the computational complexity for on-line searching is high in ITSA, a second mechanism, called Maximum Likelihood Relaxation (MLR), which is a modified Dijkstra's algorithm is introduced. The MLR algorithm optimizes the allocation of a protection path while the working path is being calculated. The protocol aims to select a working path in such a way that the number of links containing enough shareable spare capacity is maximized. In other words, the working path is selected such that the number of links without enough sharable spare capacity and working link costs are jointly optimized. MLR is a much faster mechanism than ITSA. However, MLR cannot guarantee reaching to the optimal solution.

A work in [Kar03] assumes that protection paths are not sharable. This study introduces Max 2-Route Flow and 2-Critical Link metrics for formulating a Linear Program (IP) problem. The problem is solved to maximize Max 2-Route Flow and minimize 2-Critical Link. It is notable that this scheme is different from routing problems that maximize 1-flow or minimize 1-critical link.

2.4.2 Related issues

Bringing the above techniques characteristics into attention give emphasis to the following functions:

- Rapid signaling mechanisms need to propagate information related to the failure to other nodes promptly.
- Fast updating of protection routing tables is another requirement, which guarantees reliable restoration operations.
- Revertible protection switching is needed. Then, as soon as the failed link is fixed, traffic could return to the working paths and allow protection lines to be either allocated by low-priority classes or reserved as the protection bandwidth.
- Network service providers need to offer appropriate management tools to hide complexities and make protection and restoration schemes seem simple. For example, computing the best working and protection paths fast and automatic.

2.5 Dynamic Restoration

Pre-planning path protection schemes can provide 100% protection if they are supplied by 100% redundancy. To overcome this disadvantage, resource-sharing dynamic restoration mechanisms were introduced. Dynamic restoration mechanisms are based on two basic paradigms: *link restoration and path restoration*. Link restoration is achieved by the end nodes of the failed link via spinning around the failed link. *Loop-back* protocol [Med99] is the most known link restoration protocol. In loop-back two nodes adjacent to the failure are responsible for switching the affected traffic to another path. Loop-back mechanism is also applicable to recovery from node failures but sometimes it leads to inefficient unnecessary examinations of all alternate loops. For mesh networks, *Generalized Loop-back* [Med02] protocol has been introduced. This scheme finds a conjugate diagraph, called the secondary diagraph for the primary diagraph to flood the affected traffic with it. The flooding action with simultaneous failures can head to complications and ultimately to performance degradations. In path restoration, in contrast, an s-d pair dynamically discovers an alternate route on an end-to-end basis. However, the problem of finding two diversely routed paths (path restoration) in optical networks is much more difficult than the classical edge/node disjoint paths problem in graph theory [Hu-03]. This is because optical networks architecturally have two layers: the physical layer and the optical layer (virtual/logical layer). The physical layer consists

of fiber spans and nodes (representing locations where fiber spans terminated) and the optical layer consists of optical links (or lightpaths) and a subset of nodes contained in the physical layer. An optical link in the optical layer is a path connecting a pair of nodes via a set of fiber spans in the physical layer. The failures, including fiber cuts or failures in optical/electronic switching components, occur in the physical layer. However, the circuits are routed over optical links in the optical layer. Since an optical link is a path that may traverse several fiber spans and nodes, it can be affected by different failures in the physical layer. In addition, several optical links may traverse a single fiber span or node. Hence, a single failure in the physical layer can cause multiple optical link failures. The diverse routing problem is to find two paths between a pair of nodes in the optical layer such that no single failure in the physical layer may cause both paths to fail. In [Lee01], an example is provided to illustrate why the traditional edge/node-disjoint algorithms do not work for such a problem. Hu solved the open problem and proved that diverse routing is an NP-complete problem [Hu-03]. Several heuristic algorithms are proposed in the literature during the last two decades for path restoration. One of the earliest dynamic restoration algorithms is the Self-healing Algorithm (SA). The algorithm employs a two-way signaling method to find the shared diverse paths. It consists of two phases. In the first phase, the destination (sender) floods its signature to reach the source (chooser) by diverse routes. In the second phase, the chooser sends its signature back to the sender to identify the best path for the sender [Gro87]. The advantage of this restoration scheme is that it is very simple and completely database free, which can be useful if memory is limited. The speed of the Self-Healing network is influenced by the average propagating delay. To resolve the multi-failure problem, a technique in [Kom90] suggested adding a confirmation phase to the GSP protocol. Then, multi senders and choosers can be triggered instantly. Later, other sophisticated techniques such as FITNESS [Yan93], NETRATS [Chn93] with added phases and more computational complexities are proposed. A framework in [Nel03] proposes a localized re-routing approach called Failure Insensitive Routing (FIR) that provides short-term restoration for the transient failures. The idea behind the protocol is to reconfigure a backup route for the routers adjacent to the failure only, and if the failure persists for a longer duration, then

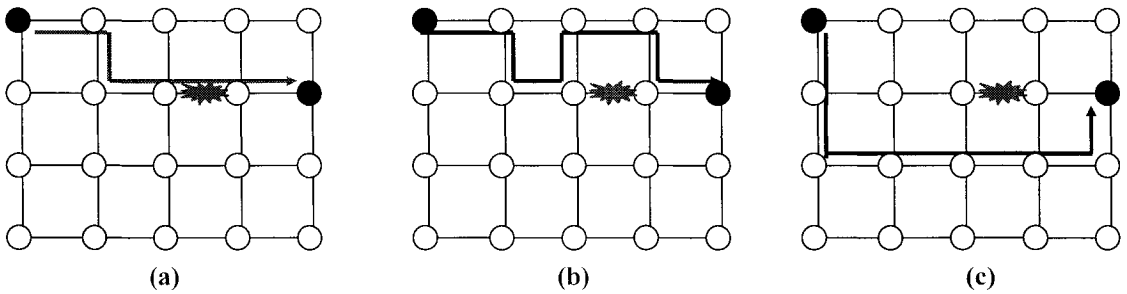

Figure 2.6: a) The failed path, b) link restoration, and c) path restoration.

the network routing protocol will be used to trigger a network wide restoration. The study shows that FIR significantly improves service availability for multiple failures scenarios. In [Jun02] a framework has been introduced that achieves dynamic path restoration by Backward Restoration Scheme (BRS). The proposed protocol searches for a replacement path between the destination-source by reserving all alternate paths. BRS finds a recovery path without employing any fault detection mechanism. However, this protocol suffers from the reservation delays of more than one examined path between the affected source-destination pair. In addition, the existence of several SRLGs that will fill the network with many in-progress reservations has not been considered. Figure 2.6 contrasts path and link restoration techniques. Both path and link restoration perform efficiently if they occupy the capacity of a shortest path/loop to the destination [Wu-95].

2.6 Multi-layer Protection in WDM Networks

Survivable functions, including fault detection, localization, alarm filtering, protection and restoration mechanisms are presented in different layers. This concept enables fault recovery procedures to either take place in a specific layer or take part in a hybrid of layers. For instance, recovery could be achieved via fault detection/localization in the physical layer followed by signaling in the optical layer and restoration in the network layer [Bis96]. However, splitting fault recovery processes among different layers requires coordination. In other words, coordination is necessary at multiple layers to assign a particular function to each layer and harmonize the layers in an effective way. In general, the set of rules describing the point of origination between of the fault recover process

and the interaction between layers is called an *escalation* or *interworking* strategy. Two escalation mechanisms have been proposed [Dem99] based on the layer at which the fault recovery process is initiated. In the bottom-up mechanism, recovery is started at the closest layer to the failure and escalated upward after a hold-off time associated to each layer. On the other hand, in top-down mechanism, recovery is always initiated at the uppermost layer and escalated downward. Hold-off time is not necessary in this type. The bottom-up approach is superior in terms of recovery time and cost, while the top-down approach is advantageous in terms of providing differentiated QoS.

There are several proposed protection mechanisms that try to link higher protection layers to the optical protection layer. One of the best one is Protection Interoperability WDM (PIW) that has been proposed to overcome a phenomenon called *failure propagation* [Arm97]. The risk of failure propagation exists wherever the links of a higher layer network are mapped into paths of a lower layer network. Failure propagation creates three problems: bottleneck problem, connectivity problem, and problem with grouping. PIW finds a solution to the mapping problem by considering the capacity and connectivity constraints. The PIW algorithm considers parameters that can be used to adapt the protection problem to different cases of higher layer protection schemes like SONET, IP or ATM networks [Cro00].

For IP over WDM networks, protection can be provided either by mapping the client layer into the optical layer, and relying mostly on the optical layer protection or it could be achieved by using Multi Protocol Label Switch (MPLS) over WDM [Dov01a]. There are also several challenges in applying MPLS to optical network restoration. First, an optical network connection is provisioned by cross connecting channels on individual OTSs along its path. This fact implies that zero-bandwidth paths cannot be established for later use. In contrast, in IP networks using MPLS, Label Switch Paths (LSPs) may be established such that no bandwidth is consumed [Dov01b]. Another challenging point about applying MPLS is raised by the bi-directional use of optical channels. This problem in bi-directional networks is usually created by the arrivals of staggered messages, where both ends of a link are able to assign the channel to arriving requests. There are some suggested solutions that use different *glare* techniques and also the Crank-back protocol

[Oka03]. The third challenge emerges when Resource reservation Protocol (RSVP) is employed as the background Connection Admission Control (CAC) protocol in MPLS. Applying RSVP would result in the cross-connection of channels in reverse directions, which will slow down the provisioning procedure. A practical way to solve this problem is to make use of Just Enough Time (JET) based scheme using PATH message in RSVP. A restoration mechanism, which uses PATH message and considers the glare has introduced as Robust Optical Layer End-to-End X-connection (ROLEX) in [Dov00].

2.7 Conclusion

In this chapter, we have surveyed optical network survivability techniques, including fault detection/localization, protection and restoration techniques in the literature. As it has been stated previously, most of the reviewed survivability techniques operate in opaque optical networks and rely on electronic signal monitoring/analysis. However, transparent all-optical networks, which are our focused attention networks, should skip any electronic monitoring and conversion to offer a considerably higher bandwidth. Unfortunately, all-optical technologies and devices are not mature yet. As a result, unlike opaque optical networks, transparent all-optical networks are unable to examine signal quality at the regenerators or measure BER (notice that analog signal monitoring at intermediate nodes may only allow them to sense signal loss). Although, novel strategies have been suggested which measure SNR by spectrum analyzers or via synchronous sampling [Mas03], they are neither economical nor scalable at the present time.

Furthermore, isolating failure in transparent all-optical networks is much harder than opaque optical networks, again, due to the lack of monitoring ability. Also, attack (intentional failure) detection and rectification is very complex. Therefore, previously defined optical networks terminologies, models, and techniques should be modified to consider the mentioned dissimilarities.

In the next chapter, we will introduce our deploying structural model and controlling mechanism for creating transparent all-optical networks. We give further details on some of the related reviewed mechanisms and present our proposed schemes for tailoring more reliable transparent all-optical networks in the coming chapters.

Chapter 3. Nested Routing Protocols

3.1 Introduction

TRANSPARENT All-optical networks are created by means of different structural models, control platforms, routing algorithms and switching techniques. In this section, first we introduce the employed components, the considered infrastructure, and the applied control mechanism which shape our network. Note that the network is one of the dominant all-optical networks in practice. Then, we review key routing mechanisms applicable to the network. We propose a novel signaling routing protocol that utilizes the unused bandwidth of establishing phases. We evaluate the protocol in terms of average bandwidth saving. The results reveal that the proposed protocol highly upgrades the network performance.

3.1.1 The Overlay Model

All-optical networks could be created based on different designs such as the *overlay*, *augmented*, *integrated*, *peer-to-peer* models [Rap01]. The all-optical network architecture considered throughout this thesis is the overlay model. In this structural design, optical switches interconnect data links and create the data network (also known as all-optical data channels), while the control units including E/O/E conversions and optical amplifiers interconnect control links and construct the control network (also called optical supervisory channels). The data channels are connected by optical switches and operate in a circuit switching fashion, while the control channels are constructed of electronic controllers and work in a packet switching mode. The traffic in the control network consists of small control packets resulting in much lighter traffic. Therefore, the control channel is usually implemented by one or more dedicated wavelengths in the same fiber link. When a connection request is arrived, a control packet in the control network traverses and configures the selected path switches to create a transparent optical path, namely a lightpath.

3.1.2 Controlling Mechanisms

Control mechanisms are primarily developed based on either centralized or distributed models. Despite this fact that centralized control mechanisms are relatively simple and work well for static traffic and small networks, they are considered to be infeasible for dynamic extended systems. In contrast, distributed control mechanisms are complex, but more scalable and reliable than centralized ones. Thus, distributed control model is preferably employed to manage dynamic traffic in scalable networks. In addition to these models, there is hierarchical model, which is the combination of centralized and distributed models. Hierarchical models mostly are applied to increasingly medium to large networks and dynamic information systems because of their ability to coordinate the network controlling messages. However, hierarchical management models cannot be economically and practically implemented for any network topology. Subsequently, researchers are still pursuing distributed rather than hierarchical management structures for all-optical mesh networks.

3.1.3 Routing and Wavelength Assignment

Given a set of lightpaths that need to be established on a network, and given a constraint on the number of wavelengths (the applied multiplexing degree), we need to determine the routes over which these lightpaths should be set up to maximize the number of paths. This problem is known as Routing and Wavelength Assignment (RWA). Conventionally, two types of RWA schemes have been defined, *Static* and *Dynamic* [Muk97]. In the static scheme also known as Static Lightpath Establishment (SLE), the lightpaths are set up in advance and corresponding wavelengths are fixed. In contrast, dynamic approach also called Dynamic Lightpath Establishment (DLS) establish lightpaths when requests have arrived and tear them down when the communication is over. Since the RWA problem is NP-Complete, it is often decoupled into routing and wavelength assignment sub-problems [Ram94]. It is extremely hard to obtain an exact solution to this problem and there are only heuristic solutions for it. Some of the proposed solutions are *random*

allocation, *most-used allocation*, *least-used allocation*, and *first-fit allocation*, [Ram95], [Ban02] and [Ram98]. Recently more sophisticated heuristic algorithms have been suggested that are originally devised for multi-fiber optical networks but are applicable on single fibers, such as *MAX-SUM* [Bar97], *relative capacity loss* [Zha98], [Spa00] and *distributed relative capacity loss* [Mukh98], [Zan00b], [Sto02].

3.1.3.1 Reservation Protocols

Reservation protocols create virtual channels (lightpaths) by selecting routes, assigning wavelengths, and configuring intermediate switches in networks. In general, reservation protocols employ two types of signaling protocols, namely *hop-by-hop*, and *parallel*. In hop-by-hop mechanism, a control signal is generated and sent one hop at a time, through the selected path toward the destination. This control signal is processed at each intermediate node and then forwarded to the next node. In the parallel scheme, a separate reservation signal is sent to each node along the path in parallel. Each node processes its signal and independently sends an acknowledgement back to the source.

In general, reservation protocols are classified into two major types based on the node at which the reservation is initiated [Sah00a]. In Forward Reservation Protocol (FRP), a source node decides on a route/wavelength and launches a reservation procedure. Accordingly, the source forwards a reservation signal toward the destination in the selected path. Once the reservation signal successfully reached the destination, it informs the source node of the successful reservation by sending an acknowledgement signal back to the source. In Backward Reservation Protocol (BRP), again the source node selects a route to the destination node but in contrast to forward reservation, the reservation is started from the destination [Som99]. Figure 3.1 illustrates the main differences between FRP and BRP. The shaded area represents wasted bandwidth of the restoration phases. It is easy to observe that BRP is more efficient than FRP.

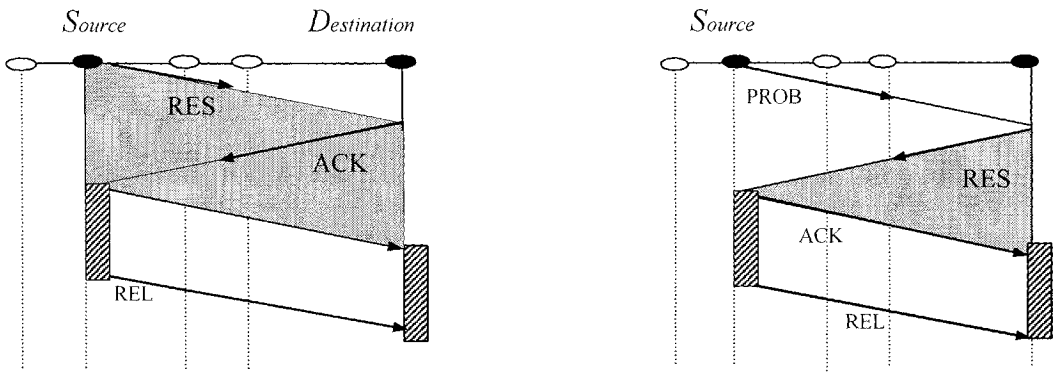

Figure 3.1: A comparison between the FRP (left) and BRP (right) reservation protocols.

3.1.4 Switching Techniques

Optical WDM networks employ three switching techniques: *photonic packet switching*; *optical burst switching;* and *optical circuit switching*. The performance of the photonic packet switching scheme is limited by the speed of electronics, since electronic buffering and address decoding are required at intermediate nodes. However, due to the high bandwidth, buffer overflow may occur that poses serious threats to the transmission of delay-sensitive traffic in this type of network. Although several solutions, such as employing optical delay lines, using deflection routing schemes, and applying wavelength conversions have been suggested, packet switching still suffers from packet loss, unpredictable delay, and call blocking. The second scheme, optical burst switching is an alternate technique to photonic packet switching and optical circuit switching. Examples of this type are Just-In-Time (JIT) and Just-Enough-time (JET) [Bal02]. This type speeds up the reservation process by estimating the time takes for a reservation to be processed in each node and sending the data right after that time without a need for any acknowledgement. However, their performance degrades dramatically for highly loaded networks [Qia99] since establishing a path may not be successful in primary attempts. Therefore, to fully capture the potential of optical bandwidth and remove the electronic bottleneck, optical signals should be transmitted in a pure circuit-switched fashion. No

buffering or decoding should be performed at intermediate nodes. In other words, optical signals should remain in the optical domain from the source to the destination.

3.2 Nested Reservation Protocol

We have introduced the idea of nesting data packets to utilize the unused bandwidth of reservation phases in [sic03a] as Nested Reservation Protocol (NRP). In general, NRP defines two types of reservation packets, namely independent and nested packets. Nested packets owe their existences to reservation packets. Therefore, conventional reservation packets are called independent reservation packets to contrast this attribute.

Unlike Backward Reservation Protocol (BRP) that locks a requested path between the destination and the source node via a reservation signal, NPR only attaches a time-stamp to each node along the path without locking the path by means of an independent reservation signal. The attached time-stamps automatically lock the corresponding links in coming pre-determined times. Each locking time is estimated based on the amount of round trip propagation delay between the current and the source node, which could be extracted from the node routing table. Then, the associated bandwidth stays unoccupied until the indicated time by the time-stamp is reached. This way, the reserved/unlocked bandwidth can then be assigned to a qualified request called a nested packet.

A nested reservation packet (the alternate reservation phase in NRP) examines the likelihood of transferring a nested data packet within the independent's reservation interval. To be more efficient, a nested path has to be a single-hop path and also a partial path of an independent path.

A nested reservation always assumes that an independent reservation is successful. Then, acting on this assumption, it locks its path similar to what BRP does. The locking process is necessary for a nested packet to avoid being disrupted by an independent failed reservation. The NRP protocol justifies the obligatory locking process of a nested path by indicating the following points:

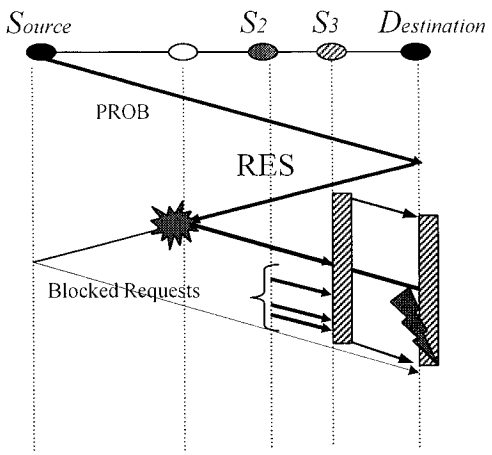

Figure 3.2: Upcoming requests have been blocked by the nested packet.

a) A nested packet is a short data packet. Therefore, the nested data can be completely transferred before new requests arrive; b) transferring the nested data can be accomplished during releasing the partially reserved lightpath of an unsuccessful independent restoration.

Thus, it is inferred that the transmission of a nested data packet does not postpone allocating the bandwidth to an upcoming real time request. This inference is realistic under a light load since reservations are usually successful and new arrivals are scarce. However, under a heavy load, the inference is not always correct; a nested packet can delay allocating the bandwidth to upcoming arrivals. This problem has been depicted in Figure 3.2. We illustrated a blocked independent reservation between an s-d pair. A nested packet is also shown between the source S_3 and the destination. All the arriving requests to the source S_2 have found the demanding bandwidth is occupied since the path has been locked by the nested reservation. This problem will ultimately saturate the network with a cluster of nested packets, since the protocol gives priority to single-hop nested requests over multi-hop independent requests. Giving priority to single-hop calls opposes the intention of reservation protocols that aim to bring fairness to networks.

Releasing the bandwidth right after failing an independent reservation cannot solve this problem, since this action destroys the nested data. In other words, it makes the protocol unreliable for nested data packets.

To remove the problem, the size of nested data packets should be adjusted (modified) based on the guaranteed free bandwidth, which could be defined by a successful or unsuccessful independent reservation. In the next section, we provide a detailed description of our novel protocol that efficiently solves the problem.

3.3 Signaling Nested Reservation Protocol

In this section, we introduce our novel signaling protocol called Signaling Nested Reservation Protocol (SNRP) that aims to assign the unused bandwidth of in progress reservations to non-real time requests. The proposed protocol neither hinders upcoming requests from grabbing the bandwidth nor interferes with nested data. The main implementation rules of SNRP are as follows: a) labeling rather than locking a requested path before transferring data to make use of the reservation bandwidth; b) signaling labeled nodes along the path to determine the allowed size of nested data. These two rules are put into practice for multi-hop request calls in the proposed protocol since single-hop requests are not involved in considerable unused bandwidth to be utilized by nested data. We describe SNRP by means of two scenarios. In the first scenario, reservations are mostly successful. Then, when a multi-hop request arrives, a backward reservation signal is disseminated from the destination toward the source. Note that this signal follows a forward probe signal that has been already received by the destination. So far the protocol operates like BRP. The main distinction is that the reservation signal in SNRP only labels links (nodes) along the path by "reserved", instead of locking them. Thus, SNRP acts similar to NRP in labeling. However, in contrast to NRP, no time-stamp is attached to define the locking time. When a node is successfully reserved (labeled), a permission signal is sent from the currently labeled node to the destination and all of the previously labeled nodes. Let us call this signal, *go signal*. Go signal enables each recipient of the signal to send a non-real time data packet to downstream neighbor. Go signal transfers the exact allowed packet size for nested data to the recipients. The permitted size of data packets is determined by the switching delay, the reservation holding time, and the round trip propagation delay between the currently labeled and the next node. Figure 3.3 shows how applying the SNRP protocol removes the unused bandwidth of an ongoing

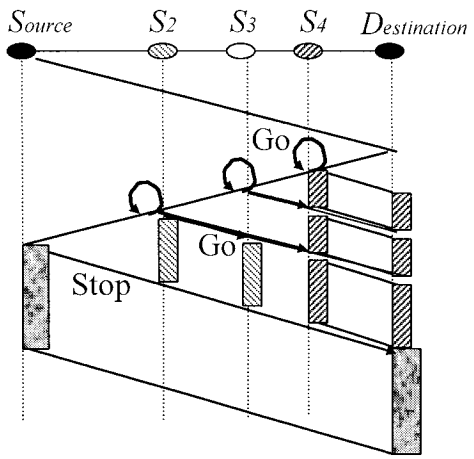

Figure 3.3: Go signals determine the size of the data packets.

reservation by assigning it to a set of nested non-real time data. Thick arrows represent go signals. For the sake of simplicity only the S_4 and S_2 have been depicted as activated sources, which have non-real time data for transmission. It is clear that other sources such as S_3 could be enabled once receiving the go signal.

Now, let us look at a different scenario in which a multi-hop reservation finds the path blocked. The path could be blocked due to allocating the bandwidth to an alternate simultaneous reservation which shares at least a link with the blocked restoration. Consequently, to inform the source of the failed reservation and trying a new path, the unavailable node forwards a *fail signal* to the source. In addition, it sends a prevention signal toward already permitted nodes to suppress them from sending more data packets. We call this prevention signal, *stop signal*. The stop signal can be piggybacked by "release signal" to remove the reserved labels from the partially reserved path. Since the stop signal does not encapsulate any information such as size or time, it could be summarized in one bit. As a result of receiving the stop signal, all the active sources are deactivated. This ensures that there will not be any nested packet after releasing a failed independent reservation to either delay assigning the bandwidth to upcoming requests or be disrupted by the released signal. Figure 3.4 shows how in SNRP the time and size of nested packets are determined to let the nested packets be transferred without playing a

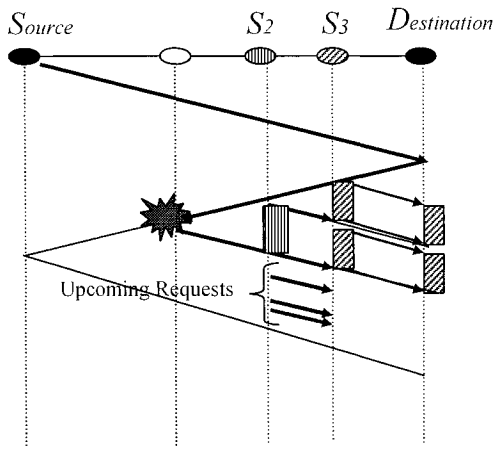

Figure 3.4: The nested data will be transferred before the FAIL signal returns to the destination and destroys the data.

negative role on upcoming requests or being affected by the unsuccessful independent reservation.

3.4 Comparison of NRP and SNRP

In this section we compare different features of NRP and SNRP. The proposed SNRP protocol has several advantages over NRP. The most important advantage is that it lessens the delay of non-real time data packets by sending nested data (note that it does not have any effect on real time data packets). The second benefit of this method is to shrink the unused bandwidth of reservation phases to an almost negligible amount. In contrast to NRP, which only utilizes the bandwidth of a single-hop destination-neighbor, every single-hop reserved pair is entitled to participate in the bandwidth usage. In addition, in NRP each reservation signal stores a time-stamp at each node all along the requested path. This function demands extending memory units at each node to store extra information. Conversely, the SNRP protocol only labels the path by reserved labels and does not require storing any time-stamp(s). Therefore, one bit can be used to exhibit the reserved state for each node. This implementation simplifies the protocol and reduces the memory units needed at each node. Furthermore, it will diminish the set-up time. Decreasing set-up time speeds up the reservation mechanism, which is highly appreciated in distributed systems.

Another benefit of the new protocol is its accuracy. By sending permission signals to every involved node along the reserved path, the length of data packets are restricted properly to the limited reservation phase bandwidth. Adjusting packet sizes ensures transferring data packets flawlessly. The accuracy of the protocol also guarantees allocating bandwidth without any delay to the arrival of real time calls. Figure 3.4 illustrates the successful transmission of nested data packets in case of a blocked reservation. The same figure also shows how the upcoming requests find enough bandwidth to launch their reservation signal without hesitation.

The last advantage of SNRP is related to its flexible locking process. Assume that a reservation signal is blocked. Two different methods are applicable: in the first method, the reservation is terminated immediately and then a retry mechanism is launched afterward; in contrast, the second type holds the line for a short time to examine the chance of releasing the wavelength during the holding time. In the latter case that the reservation time is extended, the go signal will be able to define a larger nested packet size. The exact size is computed based on different switching, propagation delay, and holding time. Furthermore, the bandwidth could be locked whenever the reservation is completed not based on rigid time.

Despite the above-mentioned advantages, SNRP has its own drawbacks. Signaling requires consuming bandwidth. Sending permission signals that encapsulate allowed packet sizes requires extending control packets. Extending control packets means utilizing more bandwidth from the control network, which may create congestion on control channels.

The following are several arguments, which lead us to disregard the negative effect of extending control packets:

- The number of ongoing reservations is limited in a system. Therefore, the number of the permission signals is relatively restricted.
- The size of data packets carried by permission signal can be factorized. This way, the size of a permission control packet could be reduced to a smaller number of bits.
- In general, the control network handles a light traffic load. Thus, it can tolerate the increased traffic without being congested and degraded.

In addition to the mentioned limitations, signaling constantly is a complicated procedure that prompts the reservation protocol prone to signal loss and also makes network management harder. However, in networks with a small number of nodes as it is in backbones, signaling can be smoothly and easily handled.

3.5 Performance evaluation

In this section we calculate the efficiency of the protocol to evaluate the improvement in bandwidth usage. Let us consider the following parameters for our evaluation.

N is the network diameter (maximum number of hops between nodes).

P is the probability of having a non-real time packet ready to be transmitted at each node.

x_i is the length of a non-real time packets. Index i, is the distance of a node from the destination. We assume that x_i has an exponential distribution with mean $1/\lambda$,

$$f(x_i) = \lambda e^{-\lambda x_i} U(x_i), \tag{3.1}$$

where $U(.)$ denotes the unit step function.

T *is* the propagation delay between any two successive neighboring nodes (we assumed an equal distance between nodes. $l_i = (i-1)T$ is the available free bandwidth when index i indicates the distance between the reserved and the destination node. Note that since (i-1) is a constant, $l_i = (i-1)T$ is the length of free bandwidth (time).

Saving bandwidth (SB) between two consecutive nodes in a reservation interval is either equal to zero in case no non-real time packet is available or is defined by the minimum of the pre-determined size of free bandwidth, l_i and the packet length x_i. Therefore, the saving bandwidth at node ith is equal to:

$$\mathrm{SB}_i = \begin{cases} \min(x_i, l_i) & \text{with probability, } P \\ 0 & \text{with probability, } (1\text{-}P) \end{cases} \tag{3.2}$$

As a result, the conditional mean of the saving bandwidth at each node equals:

$$E[\mathrm{SB}_i | x_i] = P \min(x_i, l_i). \tag{3.3}$$

Therefore, the average saving bandwidth at each node equals:

$$\begin{aligned} E[\mathrm{SB}_i] &= E[P\min(x_i, l_i)] \\ &= P\int_{-\infty}^{\infty}\min(x_i,(i-1)T)f(x_i)dx_i, \\ &= P\int_{0}^{(i-1)T} x_i\lambda e^{-\lambda x_i}dx_i + P\int_{(i-1)T}^{\infty}(i-1)T\lambda e^{-\lambda x_i}dx_i, \\ &= \frac{P}{\lambda}\left(1-e^{-\lambda(i-1)T}\right) \end{aligned} \tag{3.4}$$

Total average saving bandwidth, resulting from different nodes for an n-hop away s-d pair, would be as follows.

$$E[\mathrm{SB}\mid n] = \sum_{i=1}^{n} E[\mathrm{SB}_i], \tag{3.5}$$

where n stands for the distance between the source and the destination attempting to reserve the requested lightpath. For the sake of simplicity we assume, that the distance n as a random variable is uniformly distributed,[1] i.e., $P(n) = 1/N$ for $n = 1,2,\cdots,N$, we get.

$$\begin{aligned} E[\mathrm{SB}] &= E[E[\mathrm{SB}\mid n]] \\ &= \sum_{n=1}^{N}\frac{1}{N}E[\mathrm{SB}\mid n] \\ &= \frac{P}{\lambda N}\sum_{n=1}^{N}\sum_{i=1}^{n}\left(1-e^{-\lambda(i-1)T}\right). \end{aligned} \tag{3.6}$$

Simplifying the above equation using a simple geometric series, the total average saving bandwidth, i.e., the increased throughput, over a reservation period is given by:

$$E[\mathrm{SB}] = \frac{P}{\lambda}\left(\frac{N+1}{2} - \frac{N\left(1-e^{-\lambda T}\right) - e^{-\lambda T}\left(1-e^{-\lambda TN}\right)}{N\left(1-e^{-\lambda T}\right)^2}\right). \tag{3.7}$$

[1] A more reasonable assumption is to take the location of a source and a destination to be independent and uniformly distributed. However, the distance n would have a distribution with $P(n) = \frac{2(N+1-n)}{N(N+1)}$, for $n = 1,2,\cdots,N$. This will result in very complex expressions but almost similar value.

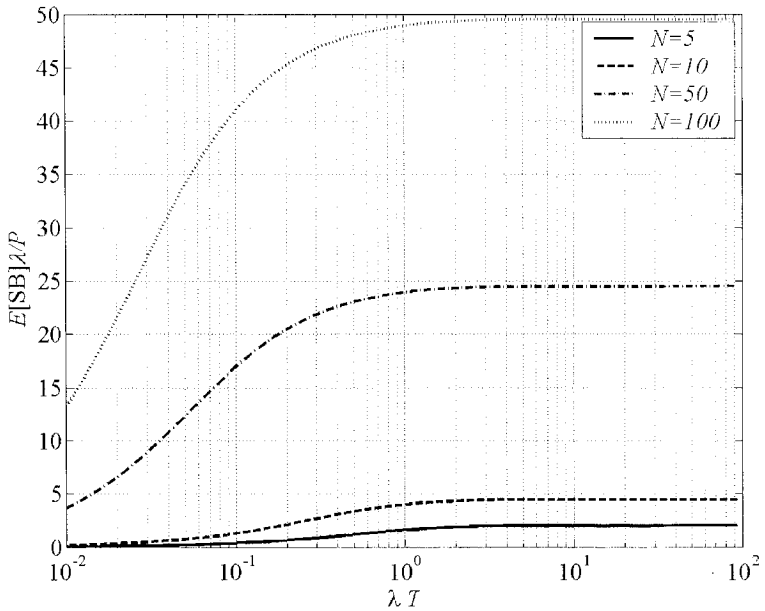

Figure 3.5: The ratio of the average saving bandwidth and the average of non-real time data packets (normalized over the probability), E [SB]/P, versus the ratio of the delay between nodes and the average of non-real time data packets λT.

Figure 3.5 shows $E[SB]\frac{\lambda}{P}$ versus λT. The following remarks are drawn from the above expression and justified by our intuitive expectations. Note that λT denotes the ratio of the delay between nodes and the average length of the non-real time data packets.

Remark 1: If λT is small enough, the above expression is approximated by $E[\mathrm{SB}] \approx P(N+1)/(2\lambda)$, i.e., the increased throughput is proportional to the average length of available non-real time packets, to the diameter of the network and also to the probability of having data available for transmission during reservation intervals. As it has been shown, increasing the diameter of network, N, increases the bandwidth usage.

Remark 2: The average saving bandwidth, E [SB], increases when the delay between nodes is increased. However, the improvement is saturated when the delay, T, is considerably bigger than the average length of non-real time data packets $1/\lambda$, i.e., $\lim_{T\to\infty} E[\mathrm{SB}] = P(N+1)/(2\lambda)$. The saving bandwidth is proportional to P, and increases with the increase of the average length of non-real time data packets.

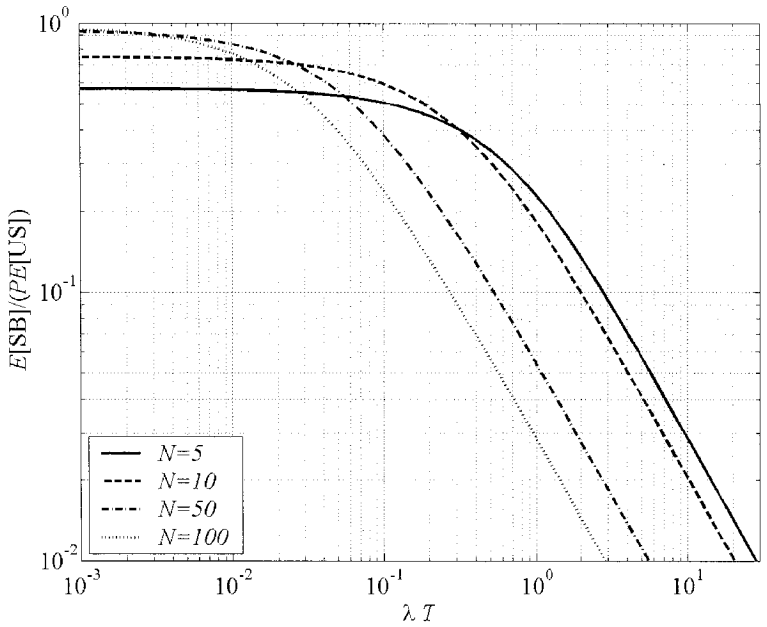

Figure 3.6: The ratio of the average saving bandwidth and the average of bandwidth, *E*[SB]/(*PE*[UB]), versus the ratio of the delay between nodes and the average of non-real time data packets λT.

If we assume that the average length of a real time data packet is L, the increase in throughput would be $E[\mathrm{SB}]$.

Thus, the improved average efficiency is,

$$\eta = \frac{L + E[\mathrm{SB}]}{L + E[\mathrm{UB}]}, \tag{3.8}$$

where E [UB] is the average unused bandwidth in the BRP protocol. We calculate the unused area,

$$E[\mathrm{UB}] = \sum_{n=1}^{N} \frac{1}{N} \sum_{i=1}^{n} T = \frac{1}{N} \sum_{n=1}^{N} n.\, nT = \frac{(2N+1)(N+1)T}{6}. \tag{3.9}$$

The improved average efficiency increases when the following ratio increases:

$$\frac{E[\mathrm{SB}]}{E[\mathrm{UB}]} = \frac{P}{\lambda T}\left(\frac{3}{2N+1} - \frac{6\left(N(1-e^{-\lambda T}) - e^{-\lambda T}(1-e^{-\lambda NT})\right)}{N(2N+1)(N+1)\left(1-e^{-\lambda T}\right)^2}\right). \tag{3.10}$$

The above expression represents the increased proportion of the bandwidth usage employing SNRP to the unused bandwidth applying BRP.

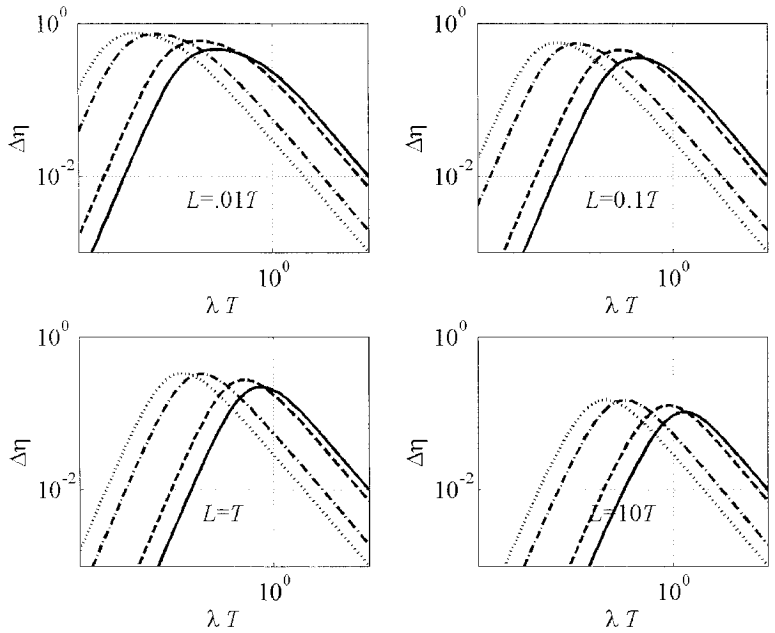

Figure 3.7: The improvement in network efficiency versus the ratio of the delay between nodes and the average of non-real time data packets λT, for different average real time data packet sizes, and when P=1.

It is interesting that this ratio is proportional to the probability P, and is a function of the number of nodes, and is inversely proportional to the delay and the inverse average length of non-real time packets, i.e., λT. Figure 3.6 shows this ratio (divided by P) versus the ratio of the delay between nodes and the average length of the non-real time packets.

Remark 3: From Figure 3.6 we conclude that the proposed method results in considerable improvement in bandwidth efficiency when $\lambda T \ll 1$, i.e., when the delay between nodes is smaller than the average length of non-real time packets.

We could calculate efficiency by replacing E [SB] and E [UB] in η. Therefore, E [SB] is related to the diameter N and traffic rate λ, if we assume that $p = 1$ (i.e., there is always non-real time packets to be delivered).

The results reveal that increasing non-real time data rate will increase efficiency. Furthermore, extending the network diameter has a positive effect on the network efficiency. Figure 3.7 depicts the improvement in efficiency of the network, $\Delta\eta=\eta-L/(L+E$ [UB]), where $P=1$. From this figure we conclude that the improvement is

substantial when the average size of the real time data packets and the reservation delays are comparable, e.g., $L<100TN$. The calculated efficiency is a lower bound, since the efficiency can be even higher when we consider allocating the unused bandwidth of failed reservations to nested data packets.

3.6 Conclusion

We proposed a novel bandwidth allocation technique based on the principles of Nested Restoration Protocol (NRP). The proposed protocol uses signaling to improve the performance of the background protocol and avoid the NRP drawbacks. Signaling Nested Reservation Protocol (SNRP) increases throughput by consuming the unused bandwidth of reservation phases. Two different types of reservation, namely independent and nested, are introduced. Independent reservations make use of two signals to control nested reservations. These signals are either permission (go) or prevention (stop) signals. Accordingly, the recipient nodes are either authorized to send restricted data packet or prevented from transferring any data packet. As a result, signaled nodes are able to utilize the bandwidth of the reservation interval efficiently and safely. This improvement comes at the cost of adding more traffic to the control network. The average bandwidth saving of a network is directly related to the average size of real time data packets, non-real time data packets, the diameter, and associated reservation delays of the network. Based on our results, we conclude that employing the proposed scheme substantially increases the network bandwidth usage and efficiency.

Chapter 4. Minimum Weighted-path Restoration Protocol

4.1 Introduction

ALTHOUGH optical networks provide wider bandwidth, more reliable service and faster speed to end users, they still suffer from service disruptions. In practice, one failure is reported every 30 minutes in average for a medium-sized network [Med99] caused by accidental damages and errors, or intentional attacks. Due to the huge bandwidth, a single failure in this type of networks leads to massive loss of data.

Failures occur in any part of optical networks, including nodes such as an OADM, or links, for instance an optical fiber. In addition to node and link failures, which are common scenarios in any communication network, channel failures are also common in all-optical networks. A channel failure is usually caused by failures in transmitting and/or receiving equipment(s) operating on the wavelength, or an attack on the involved amplifiers such as an in-band jamming. However, link failures are the most common because of their widespread expansion of fibers.

Another related issue is the probability of having multiple failed links or channels simultaneously in a network. The probability of having a multi-failure case would be less than experiencing a single failure if we consider equal and small failure probabilities for links (channels). For instance, if link failure rate is p, the probability of having a single failed link in an n-link network driven by $np(1-p)^{n-1}$ is always higher than having a two-link failure calculated by $0.5n(n-1)p^2(1-p)^{n-2}$ for $p < 2(n+1)^{-1}$. Therefore, we narrow down our discussion to more common scenarios, i.e., single link failures.

4.2 Models, Components, and Layers

Optical networks are designed based on different models such as the overlay, augmented, integrated peer-to-peer models. We consider the overlay model for creating our all-

optical network (details of this model can be found in Chapter 3). We distinguish between optical components that are capable of monitoring/analyzing and those that do not have such abilities in our network. Realistically, all-optical components such as Optical Amplifiers (OA) have limited or no electronic monitoring and analysis abilities. As a result, they may be able to sense loss of signal, but cannot realize high BER or launch any restoration procedure. In contrast, there are components capable of detecting failures and taking proper actions in response to service disruptions such as Optical Cross-connects (OXCs).

Although fault detection and recovery can be achieved in different layers, for example in the Physical (Ph) layer through electronic processes using spectrum analyzer and via Automatic Protection Switches (APSs), based on our management technique point of view, only the end-point of a lightpath (a sink) can report a failure by generating informative alarms in the Optical Channel (OCh) layer. This is mainly due to our intention to do restoration within the optical layer and consequently, to preserve the high bandwidth of all-optical communications. We assumed that critical parameters for having an effective restoration in OCh layer have been properly taken care of. These critical factors can be summarized as follows: influences due to nonlinearities, dispersion, and wavelength-to-wavelength interactions is not falsely reported as an interruption of service; interoperability, i.e., transporting the optical signals from one provider domain to another and monitoring signals parameters is achieved with the same accuracy.

4.3 Dynamic Restoration

Survivable techniques can be categorized into two major groups: pre-planned protection and dynamic restoration techniques. Pre-planned protection techniques refer to the fact that protection resources, which are specified at the planning stage, protect a system from failures. Since some of the resources and components are only used for protection in this technique, the network bandwidth is dramatically reduced. However, recovery from failures is accomplished with low delays and high probabilities. Two kinds of protection are possible, namely dedicated and shared protection. In the dedicated case, a second copy of data is simultaneously transmitted on both the working and the coupled

protection paths. Accordingly, the network bandwidth is narrowed down to half in this case. However, 100% service recovery is guaranteed in the dedicated protection. In the shared case, a second copy of data is sent through the shared protection path only after detecting a failure. To be more bandwidth-efficient, the capacity of the protection path is assigned to low priority traffic before detecting a failure and will be pre-emptied after detecting the failure. However, performing pre-empting procedures for optical networks requires buffering which is still the bottleneck of optical networks. To overcome the drawbacks associated with pre-planned protection techniques, dynamic restoration protocols are introduced. Dynamic restoration implies discovering the unoccupied network capacities and assigning them to affected traffic. Dynamic restoration protocols are more bandwidth-efficient than their static counterpart since they do not require setting aside protection resources. However, restoration time is usually longer. Furthermore, 100% service recovery cannot be guaranteed, as sufficient capacities may not be available at the time of failures.

4.4 Link and Path Restoration

Network restoration is achievable through two main techniques, namely *link* and *path restoration*. Figure 4.1 illustrates link and path restoration approaches. Link restoration restores the failed connection by finding an alternate route between the two ends of the failed link. The key to a successful restoration operation in this type is to find a set of protection cycles (loop) that cover the network links. This scheme can be implemented either in a centralized or a distributed manner. The distributed approach is more scalable,

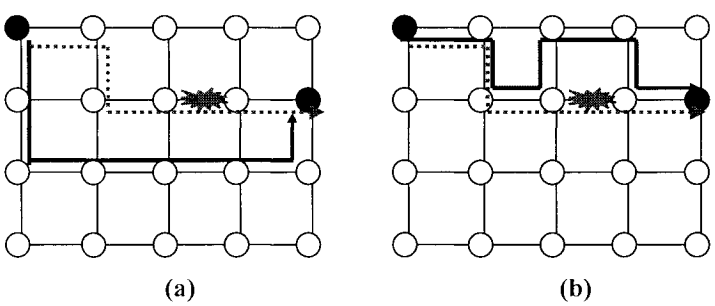

Figure 4.1: a) path restoration and b) link restoration.

but more complex. Several types of distributed schemes are introduced in the literature. Most of the introduced techniques, broadcast the failure information to all network nodes after fault localization [Zan00]. Another possible approach is to signal only the affected source-destination pairs to find the bypassing routes in the first step and report the failure to other nodes in the next step [Med99]. Another method is implemented by dividing the network into multiple domains (in an extreme case each domain is a ring) then protects each domain independently. When a failure occurs, it only disrupts a specific domain that can be reported to other domains [Med02]. Link restoration is usually faster than path restoration since it reroutes the affected traffic to a shorter recovery route (a loop around the failed link). Consequently, setting up the restoration paths, by link restoration, takes less configuration and switching times. However, link restoration requires fault localization in advance.

In contrast, path restoration finds an alternate path between the source and destination nodes. Accordingly, path restoration is an end-to-end technique rather than a point-to-point one. In the path restoration there is no need for fault localization due to this fact that the restoration path is a complete disjoint path. As a result, the restoration process is simple but normally more time-consuming than link restoration. This scheme could be also developed in centralized and distributed manner with its own advantages and disadvantages. Figure 4.2 compares efficient and inefficient path and link restorations.

On the other hand, these techniques can be implemented in a pre-planned protection or a dynamic restoration. Although pre-planned protection techniques guarantee high protection rate, they require expanding resources or setting aside capacities in their dedicated style, or performing pre-empting procedures in their shared protection form. Therefore, we propose a dynamic restoration protocol that resourcefully establishes recovery paths in the following section.

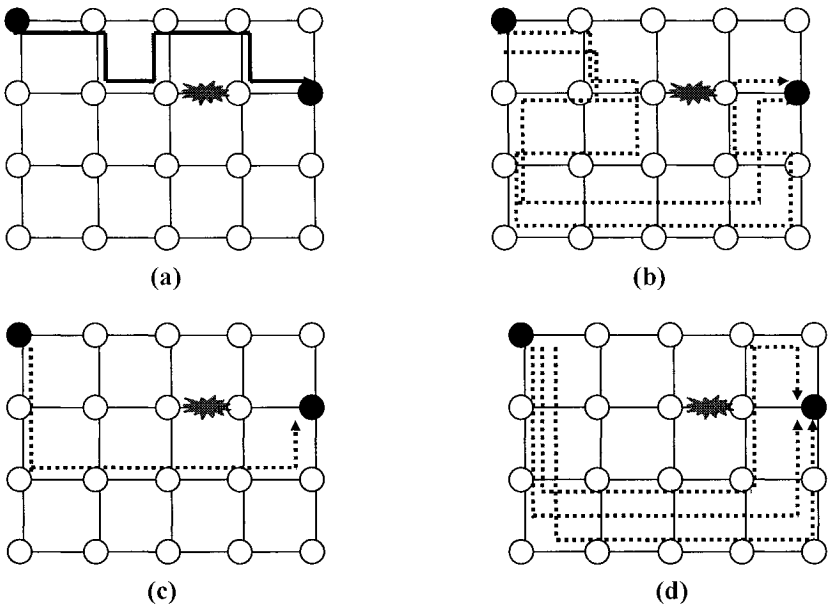

Figure 4.2: a) an efficient link restoration, b) inefficient path restorations, c) an efficient path restoration, and d) inefficient path restorations.

4.5 Minimum Weighted-path Restoration Protocol

Path restoration basically removes the capacity of affected paths from the network by disregarding the working links of affected paths. This action results creates a sudden shock to the load balancing of systems since a set of affected lightpaths need to be simultaneously set up over disjoint paths. The more lightpaths traverse through the failed link, the more severe the shock and chaos would be. In addition, path restoration could be very slow for distant nodes mostly due to associated switching configuration delays. On the other hand, link restoration could cause congestion around the failed link (a hot spot) since the capacity of shortest loops, around the failed link, is limited. To deal with these drawbacks, we propose a novel restoration protocol. *Minimum Weighted-path Restoration Protocol* (MWRP) is a tradeoff between link and path restoration with added rules and phases. This protocol is activated when a sink does not receive expected data. Protocol implementation steps can be itemized as follows:

- A fast fault localization protocol is launched to pinpoint the failed link. The location of the severed link then is reported to other nodes for table modification.
- The shortest paths between the affected s-d pair excluding the failed link are listed. Let us call this list, *S-List*. S-list can be simply extracted from the *k*th shortest paths information stored in network switches.
- Each alternate path $p_k = \{i_{k1}, i_{k2}, \cdots, i_{k\,n}\}$ is assigned a weight L_p derived from the number of involved links in it and a related factor called length factor, $L_p = \delta l_p$ where δ is defined as the length factor and l_p is the number of links in p.
- Each alternate path p is also assigned a weight S_p based on the number of to-be-configured switches (added switches) and an associated factor called switching factor, $S_p = \phi u_p$, where ϕ is the switching factor and u_p is the number of added (extra) switches. Adding these factors is the key point of the protocol since they control the significance of their coefficient and guarantee the convergence of the protocol to a joint solution.
- Alternate paths are sorted based on the cumulative assigned weights, a restoration path with the smallest weight is chosen, i.e.,

$$p = \arg\min_{p} \{S_p + L_p\}. \tag{4.1}$$

Accordingly, the proposed protocol discovers one of the *k*th shortest paths with limited number of added switches. Figure 4.3 demonstrates the affected working path and also shows the restored path by the MWRP scheme.

As we explained, a balance between minimizing the total number of links and switches is

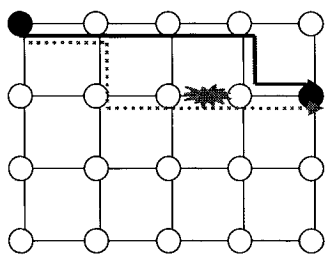

Figure 4.3: Minimum weighted-path restoration protocol (MWPR) set up a segmented restoration path.

reached by defining the length and switching factors. Network Service Providers (NSPs), based on requirements and network circumstances, specify these factors. For instance, if a network is a backbone and switching delays are comparable with propagation delays, NSPs could assign almost the same weight to both factors. On the other hand, for LAN networks whose switching delays are much higher than propagation delays, they could increase the switching factor to stress on minimizing the related delay.

It is interesting to see that the impact of other parameters can be similarly entered in the protocol. For instance, numbers of links can be replaced by traffic loads. Then, the least loaded paths are listed in S-Lists and consequently, the MWRP protocol finds the path with minimum load and number of to-be-configured switches.

4.5.1 Advantages and Disadvantages

Since the proposed protocol discovers a restoration path with the minimized number of added switches, some of the working links are shared with the restoration path. Having a partially joint path offers several advantages: restoration is partly handled by the affected path not by a completely disjoint path, which helps the network load to stay in a better balance after failure; the average restoration delay is reduced. More paths are going to be restored. However, creating weighted lists and vectors for different paths and comparing them positively increases the computational complexity.

4.6 Performance Evaluation

To evaluate the proposed protocol, we calculate its availability for mesh networks and compare the results with those of link and path restoration.

4.6.1 Reliability and Availability Theory

The theory of reliability and availability analyzes the structure of a system based on a set of distinct subsystems connected to obtain an intended function. Reliability R is defined as the probability that a system will perform its intended function for a specified period of time under assumed circumstances. On other hand, availability A is the probability that

a system is available for use at a given time. Availability may be viewed as the fraction of time that a system is in the operational state independent of how many times it has previously failed and was restored. So, if we assume that systems, components and subsystems are not repairable, we will refer to reliability, while the availability is a typical feature of restorable systems. Reliability of a system is defined as the probability that a system operates without failure for a length of time greater than t, i.e., if the random variable T is the time to failure, then, $R(t) = P(T > t) = 1 - F_T(t)$. For repairable system, Availability is defined by

$$A = \frac{E[T]}{E[T_R] + E[T_B]}. \tag{4.2}$$

Where $E[T]$, $E[T_R]$ and $E[T_B]$ are the Mean Time To Failure (MTTF), the Mean Time To Repair (MTTR) and the Mean Time Between successive Failures (MTBF). Generally a system is made up of functional elements characterized by MTBF and MTTR which are assigned as there manufacturing characteristics. In order to obtain the availability of a system, the following steps must be taken: system decomposition in functional elements; development of the mathematical model, which considers the relation among the subsystems; availability evaluation of every subsystem using the reliability values of each element; and availability evaluation of the whole system. In this work, the system under study is a mesh network, where the subsystems are the optical links and nodes routed on this topology. We assume that nodes are fully reliable and therefore, we take into account only link failures. Our notation allows us to analyze system reliability and availability as follows.

If we denote H_i as the event that the ith subsystem is operational up to time t, $\{T_i > t\}$ and H as the event that the system is operating at time t, $\{T > t\}$, $R_i(t) = P(H_i)$ represents the ith element reliability and $R(t) = P(H)$ is the system reliability. Otherwise, if H_i is defined as the event that the ith element is operating at time t independently of what has previously occurred, $A_i(t) = P(H_i)$ is the ith element availability and $A(t) = P(H)$ represents the system availability.

Let us assume that there are n alternate restoration paths, $\{p_1, p_2, \cdots, p_n\}$ for each working path, where each path is shown with its links, $p_k = (i_{1,k}, i_{2,k}, \cdots, i_{l_{p_k},k})$. Therefore, a path is available if at least one of the alternate paths is available and a given path is available if all the associated links are available.

$$P(H) = P\left(\bigcup_{k=1}^{n} \bigcap_{m=1}^{l_{p_k}} H_{i_{m,k}} \right), \tag{4.3}$$

Assuming that the collection of link failure events, $H_{i_{m,k}}$ are independent, we obtain:

$$A(t) = P(H) = 1 - \prod_{k=1}^{n} \left(1 - \prod_{m=1}^{l_{p_k}} P(H_{i_m}) \right), \tag{4.4}$$

$$A(t) = 1 - \prod_{k=1}^{n} \left(1 - \prod_{m=1}^{l_{p_k}} A_{i_m}(t) \right), \tag{4.5}$$

To compare and identify the most available technique, we need to evaluate the set of alternate paths $\{p_1, p_2, \cdots, p_n\}$ for each approach as well as the link availability $A_{i_m}(t)$. For the sake of simplicity, it is reasonable to assume that the availabilities are equal for various links, i.e., $A_{i_m}(t) = B(t)$ and for all links and techniques. Therefore, we get

$$A(t) = 1 - \prod_{k=1}^{n} \left(1 - (B(t))^{l_{p_k}} \right), \tag{4.6}$$

Based on (4.6), the availability depends on the number of alternate routes n, and the length of each route l_{p_k}. Then, we need to evaluate n and l_{p_k} to calculate the related availability $A(t)$ of each protocol. Since the number of paths and links depend on the network topology, we use the graph theory to investigate the protocols differences. We first prove that the number of alternate restoration paths for a working path decreases as we remove more links from the network.

Lemma: For a given network graph, $G^{(0)} = (V, E)$ we assume that nodes are completely reliable/fully available. Then, a failure occurs only as a result of an edge purged. Also,

we assume that there are no leaves in the graph (i.e., nodes are at least 2-edge connected) to perform dynamic shared restoration successfully. We denote:

$G^{(0)}$ is at least 2 - edge connected

$$G^{(1)} = (V^{(1)}, E^{(1)}), \quad \text{where} \begin{cases} E^{(1)} = E - F^{(1)} \\ F^{(1)} = \{e_1\}, \quad e_1 \text{ is a single edge} \end{cases}$$

$$G^{(n)} = (V^{(n)}, E^{(n)}), \quad \text{where} \begin{cases} E^{(n)} = E - \bigcup_{k=1}^{n} F^{(k)} \\ F^{(n)} = \{e_n\}, \quad e_n \text{ is a single edge} \end{cases}$$

$S^{(n)}$ is the set of all cycles in $G^{(n)}$.

Claim: the number of cycles in $G^{(n)}$ is decreasing with *n*.

Proof: We have

$S^{(1)} = S^{(0)} - C^{(1)}$

$S^{(n)} = S^{(0)} - C^{(n)}$,

where $C^{(1)} = \{\text{cycles with at least one edge in } F^{(1)}\}$ and

$C^{(n)} = \{\text{cycles with at least one edge in } \bigcup_{k=1}^{n} F^{(k)}\}$.

$C^{(n)} = C^{(n-1)} \cup \{\text{cycles with no edges in } F^{(1)} \text{ and any edge in } F^{(n)}\}$

Since $|C^{(n)}| \geq |C^{(n-1)}|$, we conclude:

$|S^{(n)}| \leq |S^{(n-1)}|$. ■

This proof illustrates that the algorithm, which eliminates more links from the network graph ends up with a smaller number of alternate paths. Therefore, path restoration has the smallest *n* (alternate paths) since it removes all of the working links from the network. In contrast, MWRP and link restoration only remove a single edge (the failed link) from the network graph. Then, they possess a higher *n* (number of alternate paths).

Next, let us evaluate the availability of a given path p_k (i.e., $A_{pk} = \prod_{i \in p_k} A_i(t)$) for our different restoration protocols. For simplicity, we assume that all links maintain equal availability prior to any failure occurrence. $A_i(t) = B(t)$. Thus, the path availability equals $(B(t))^m$ for an *m*-hop path.

.

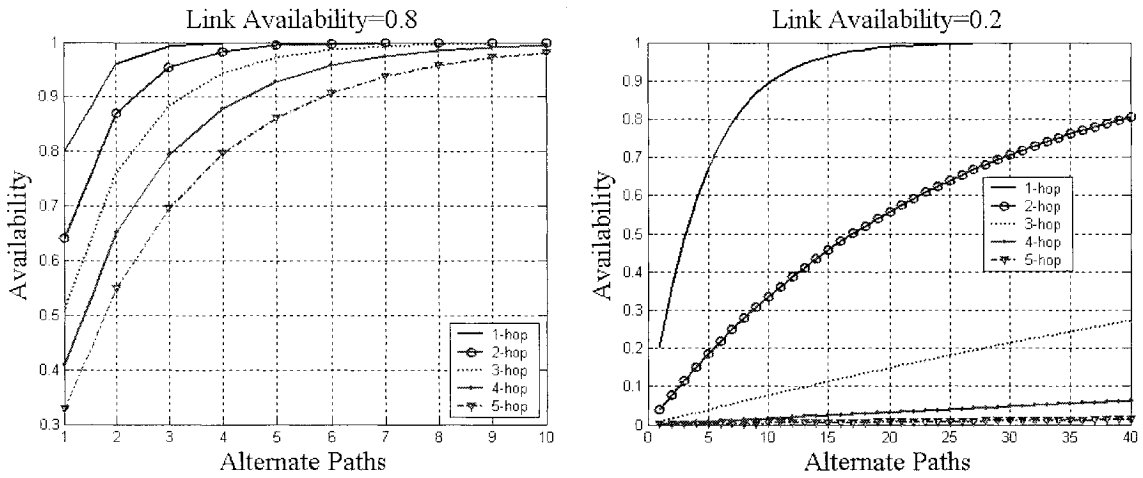

Figure 4.4: The network availability improves by increasing the number of alternate paths or decreasing the number of hops in paths.

Since the MWRP protocol minimizes the number of links and to-be-configured switches in a restoration path, the length of the path could not be longer than those set up by path or link restorations. Two extreme cases are as follows.

If the MWRP backup path comprises all the working links, it will be identical to the path defined by link restoration. On the other hand, if none of the working links are used in the restoration path, the MWRP path will be the same as the one figured out by path restoration. As a result, parameter l_{P_k} in (4.6) is the smallest possible for the MWRP protocol. Thus, A_{Pk} or the path availability of the proposed protocol would be larger than those of the counterpart protocols.

From the above, we conclude that the probability of a path not being available is reduced exponentially with increasing the number of alternate paths. Also, the availability degrades exponentially with increasing the number of involved links. Figure 4.4 depicts these points by related curves based on the calculated results. Therefore, refereeing to the outcomes, we conclude that the MWRP protocol provides the highest availability among the compared protocols.

4.7 Conclusion

Equipping communication networks with survivable components and techniques is a necessary task of Network Service Providers (NSPs) since from the client point of view, the higher the level of service continuity is, the more reliable and desirable a network would be. Accordingly, optical networks require superior survivable services to preserve their very high bandwidth. Although dynamic restoration and pre-planned protection techniques are both deployed over optical networks, restoration mechanisms are more preferable due to high bandwidth-efficiency. Accordingly, in this chapter we proposed a dynamic restoration protocol. Minimum Weighted-path Restoration Protocol (MWRP) scheme discovers a restoration path that comprises the minimum number of links (hops) and added switches between an s-d pair. The MWRP protocol uses two weighting factors, namely length and switching factors, for emphasizing on related criteria and reaching a joint solution.

We evaluated the protocol in terms of availability against path and link restoration. The results revealed that applying the MWRP protocol increases the network availability exponentially. This result was projected since the MWRP restoration processes are carried out through the most resourceful alternate paths. However, MWRP definitely has a higher average restoration delay than link restoration protocol due to the supplementary phases and related computational delays.

Chapter 5. Open Link Restoration Protocol

5.1 Introduction

EQUIPPING communication networks with survivable components and mechanisms, which recover them from failures, is a mandatory task of service providers. However, protecting networks via redundancy and by expanding physical and financial resources is not always possible. Therefore, service providers have focused on employing dynamic restoration mechanisms that make use of networks spare capacities for recovery purposes. Accordingly, in this chapter we propose a novel dynamic restoration mechanism that efficiently discovers unoccupied alternate resources for recovery of the affected traffic at the time of fault detection. The proposed protocol is an adaptation of link restoration algorithm with supplementary rules and phases. As a result, its performance in extreme cases is relatively coupled with that of link restoration algorithm. We analyze the proposed protocol in terms of system reliability. The analytical results reveal that the proposed protocol increases the network reliability considerably. We also evaluate the performance of the proposed protocol in terms of throughput via simulation. The simulation outcomes demonstrate that the proposed protocol significantly enhances the network throughput. However, the protocol has a higher computational complexity and average restoration delay than the link restoration algorithm.

5.2 Single or Multiple Layer Restorations

Survivable functions could either take place in a specific layer or takes part in a hybrid of layers. For instance, recovery could be achieved via fault detection/localization in the physical layer followed by signaling in the optical layer and restoration in the network layer as it is in IP over WDM networks or might be completed in a single layer similar to Self Healing Rings (SHRs). In general, the set of rules that describe the origin of the fault recovery process and the interaction between layers is called an *escalation* or *interworking* strategy. Two escalation mechanisms have been defined in the layer in

which the fault recovery process is initiated. In the *bottom-up* mechanism, recovery is started at the closest layer to the failure and escalated upward after a hold-off time for each layer. On the contrary, in the *top-down* mechanism, recovery is always initiated at the uppermost layer and escalated downward. Hold-off time is not necessary in this type. The bottom-up approach is superior in terms of recovery time and cost while the top-down approach is advantageous in terms of providing differentiated QoS. In addition to the escalation strategies, there are other restoration mechanisms that attempt to link higher layers to lower layers. However, splitting fault recovery processes between different layers and then linking them requires coordination. Coordination and incorporation between layers makes those protocols complex and increases restoration delays. Thus, it is simple and efficient to apply single layer restoration protocols for recovery from failures.

5.3 Open Link Restoration Protocol

In an attempt to upgrade the performance of dynamic restoration protocols, we propose a novel restoration method called *open link restoration protocol*. Open link restoration protocol is an adaptation of the link restoration protocol; with this difference that it is allowed (open) to consider and employ the working links of an affected path (a primary path) in the restoration loop. The protocol is activated when an affected node announces service disruption by disseminating corresponding alarms. We assume that critical factors for having an effective restoration in the optical layer have been properly taken care of. These critical factors are: influences due to wavelength-to-wavelength interactions is not wrongly reported as an interruption of service; transporting the optical signals from one provider domain to another and monitoring different signal parameters is carried out with the same accuracy (interoperability is entirely achieved).

The implementation steps of the proposed protocol could be summarized as following:

- The failure location is identified. Like the link restoration protocol, the open link restoration protocol requires to pinpoint the location of a failure. Fault localization is mandatory since the protocol should exclude the failed link from the restoration loop.

For opaque optical networks with monitoring and analysis abilities at intermediate nodes this step could be overlooked.

- Subsequent to fault localization, the failure information is flooded to the network nodes. This action ensures that routing tables are valid. Note that validating routing tables will accelerate the process of recovery for in progress restorations and secure upcoming request calls from examining the failed link.
- The shortest loop (the one that has the less number of to-be-configured switches) between the end-points of the failed link is discovered. A modified form of the generalized loop-back protocol is launched to figure out the most efficient restoration loop between the disconnected nodes. The modification allows the links of the failed path (the primary path) to be included in the secondary diagraph. In other words, no restriction is imposed on the links which establish the restoration loop. In addition, it is assumed that the common links between the primary and the secondary diagraphs can accommodate enough bandwidth for setting up the restoration loops, i.e., it is assumed that the common switches between two paths have been already configured. We call the modified generalized loop-back protocol, *open loop-back protocol*. It is clear then that the primary and secondary diagraphs in the open loop-back protocol are more extensive than those in the generalized loop-back protocol.
- The information of the shortest open loop-back is sent to the source node.
- Once the source node received the open loop-back signal, it learns about the restoration loop route. The source node then checks the loop links against the failed path (the primary path). If there is any replicated link, the source node removes them from both the restoration loop and the affected path. The capacities of the replicated links are released by a corresponding signal disseminated toward the destination by the source node.
- The affected traffic is rerouted through the verified/substantiated path.

Figure 5.1 shows how the restoration path is established in two simple steps. In the first step the restoration loop is drawn between the two-end of the failed link by the open loop-back protocol. Then, in the second step, once the restoration signal reached the source node, the links of the restoration loop are compared with those of the primary path

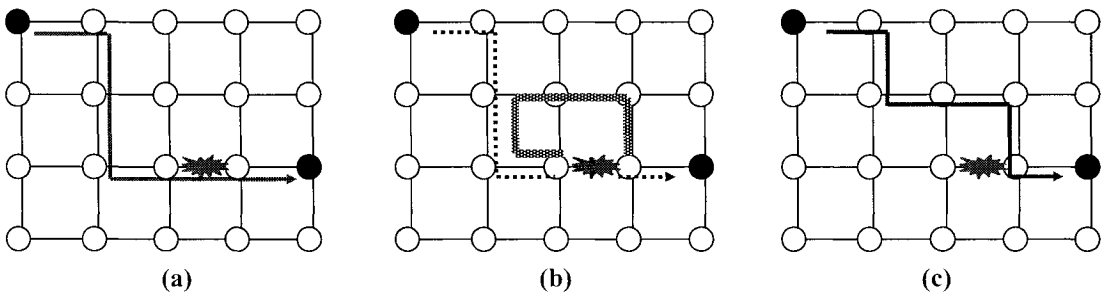

Figure 5.1: a) the working path, b) the restoration loop, and c) the verified restoration path.

and finally, the restoration path is shaped by releasing the mutual links between the restoration loop and the affected path. In subsection 5.4 we evaluate the performance of the proposed protocol and compare it with link and path restoration protocols. At this stage, we could intuitively project that the open link restoration protocol is the most resourceful restoration protocol among them since its performance would be similar to the link or path restoration protocol only in extreme cases: if the restoration loop defined by open link restoration comprises all the working links of the failed path, the resultant will be identical to one set up by link restoration; on the other hand, if the restoration loop excludes all the working links of the primary path, the resulting restoration path will be equivalent to the one discovered by path restoration.

5.3.1 Functioning Layers

Path restoration only takes place in the OCh layer since failed lightpaths are independently swapped to alternate lightpaths. In contrast to path restoration, link and open link restorations are possible in other sublayers of the optical layer such as the OCh or the OMS layer. In the OCh layer, link restoration reroutes the affected traffic of each lightpath through a loop around the failed link. Each restoration loop then could be completely disjoint from other restoration loops. Consequently, each restoration path might vary from others. For open link restoration in the OCh layer, restoration paths would be even more distinct since they might also be entirely diverse in terms of covering the working links of the failed lightpaths. In view of the fact that restoration

paths in the case of open link restoration are more diverse, the burden of affected traffic is more evenly distributed in the network. Therefore, although open link restoration is more complex, it increases the chance of successful restorations.

In the OMS layer, link restoration restores the affected lightpaths by means of identical restoration loops. As a result, although restoration paths might be different in regards to their primary links, they are similar in terms of their restoration links. For open link restoration, although restoration loops are identical at the beginning, they are modified in integration with primary paths at the end. Therefore, restoration paths would be more distinct than those of link restoration. Although it is obvious that link and open link restorations could function faster in the OMS layer than in the OCh layer, finding a single restoration loop, which could accommodate the required bandwidth of affected traffic, would be difficult or impossible for highly loaded networks. Consequently, link and open link restorations could not provide the same number of successful restorations in the OMS layer as they do in the OCh layer. Thus, open link restoration should be adapted and performed in the appropriate layer based on networks circumstances and requirements to offer higher speed or efficiency alternatively.

5.4 Network Reliability

For calculating network reliability, we assume that in average there are n disjoint paths between s-d pairs. We also assume that the time to failure for these paths are independent and exponentially distributed, i.e. the probability that a given path is in working state at time t is $R(t) = e^{-\lambda t}$, where λ is the path failure rate. This implies that path failures are memoryless. We use the following lemma to explain the relationship between the path length and its reliability assuming that link failures are independent.

Lemma 1: If multiple independent subsystems (multiple links/hop of a path) with exponential reliabilities, $R_i(t) = \exp(-\lambda_i t)$ the overall reliability of the system (the path) remains exponential $R_{\text{path}}(t) = \exp\left(-\sum_i \lambda_i t\right)$, with the sum of link failure rates as the failure rate of the overall system $\lambda = \sum_i \lambda_i$.

The above lemma also states that λ in $R(t) = e^{-\lambda t}$ for a given path is proportional to the number of links.

We presume that there are W wavelengths available on each path, such that every wavelength is independently occupied with a given probability $p \in [0,1]$. In other words, the probability p could represent the traffic load on each path, i.e., $p = 0$ means that there is no traffic and $p = 1$ indicates full traffic load (the path is fully utilized).

The network reliability is defined as the probability of a network being in the working state at time t. Therefore, using the Bayes's formula, the network reliability is given by

$$R(t) = \sum_{k=0}^{n-1} P_{rob}(\mathbf{x} = k) R(t|\mathbf{x} = k), \tag{5.1}$$

where $\mathbf{x}$ is the random variable denoting the number of failed paths, $P(\mathbf{x} = k)$ is the probability of having k failed paths out of n paths and $R(t|\mathbf{x} = k)$ is the conditional probability that the k paths could be restored by alternate paths, i.e., the probability that n-k working paths could carry the whole traffic that was originally loaded on all n paths. Since the failures of alternate paths are independent, the total number failures k has a binominal distribution. Accordingly, we have,

$$P_{rob}(\mathbf{x} = k) = \binom{n}{k} \left(1 - \exp(-\lambda t)\right)^k \exp\left(-\lambda t(n-k)\right), \tag{5.2}$$

The total number of wavelengths is equal to nW, since each path provides W wavelengths. The event of being "in use" for these wavelengths is jointly independent and is assumed to be binary random variables. Therefore, the distribution of the total number of "in use" wavelengths, $\mathbf{y}$, is as follows:

$$P(\mathbf{y} = m) = \binom{nW}{m} p^m (1-p)^{nW-m}, \tag{5.3}$$

The probability $R(t|\mathbf{x} = k)$ is equal to the probability that the affected traffic in k failed paths could be accommodated by n-k alternative paths; this is possible if and only if the total number of wavelengths in use, $\mathbf{y}$, is smaller than or equal to the number of

existing alternate wavelengths (n-k) W, i.e. using (5.3) the conditional reliability given k failures is given by,

$$R(t|\mathbf{x}=k)=P_{rob}(\mathbf{y}\le(n-k)W)=\sum_{m=0}^{(n-k)W}\binom{nW}{m}p^m(1-p)^{nW-m}, \qquad (5.4)$$

Substituting (5.2) and (5.4) in (5.1), we get

$$R(t)=\sum_{k=0}^{n-1}\sum_{m=0}^{(n-k)W}\binom{n}{k}\binom{nW}{m}(1-\exp(-\lambda t))^k\exp(-\lambda t(n-k))p^m(1-p)^{nW-m}, \qquad (5.5)$$

Now, let us investigate which protocol provides us with the highest network reliability. To do so, we draw equation (5.5) curves by assigning different values to the variables. Figure 5.2 shows the results of changing failure rate, network load, and multiplexing degree on network reliability for different path lengths (hop basis). We analyze these curves as follows.

Figure 5.2.a has been drawn when failure $\lambda t = 0.2$, network load 0.2, and multiplexing degree W=4. The curves reveal that the network reliability increases as paths get shorter (the number of hops in paths decreases). The reason is linked to the failure probability: longer paths have higher failure chance; plus, shorter paths have higher recovery chance. Also, we observe that the reliability of a path increases as the number of alternate paths increases. This performance is due to increasing the chance of establishing alternate paths.

Figure 5.2.b shows the result of increasing the failure rate when other parameters are fixed (network load 0.2, multiplexing degree W=4). We observe that for lengthy paths, the network reliability decreases as the number of paths increases. This is because the chance of having failure for lengthy paths is higher in this case. As a result, network reliability could noticeably decrease for relatively long paths.

Figure 5.2.c depicts the effect of increasing multiplexing degree when other parameters are fixed (failure $\lambda t = 0.2$, network load 0.2). With increasing the multiplying degree W, the network reliability increases. The reason behind this outcome is simple: the network load is fixed, while the number of virtual channels is increased. Consequently, increasing the number of alternate paths has improved the chance of successful restorations.

Figure 5.3.d reveals that for heavily loaded networks, reliability could be decreasing rather than increasing (as it is in lower loads, Figure 5.3.a-c). This is because, in high loads, any failure disrupts numerous lightpaths. Also, as the network gets bigger, more failed cases are emerged. Consequently, restoring large heavily loaded networks becomes harder than small lightly loaded ones. Furthermore, discovering and setting up shorter alternate paths are more possible than longer paths.

Based on our results, the shortest paths always are more reliable than the longer ones (for the worst scenarios 5.3.c-d, the reliability of shorter paths still is higher even in case of higher numbers of alternate paths). Accordingly, we conclude that the open link restoration protocol, with the smallest restoration paths, obtains the highest reliability.

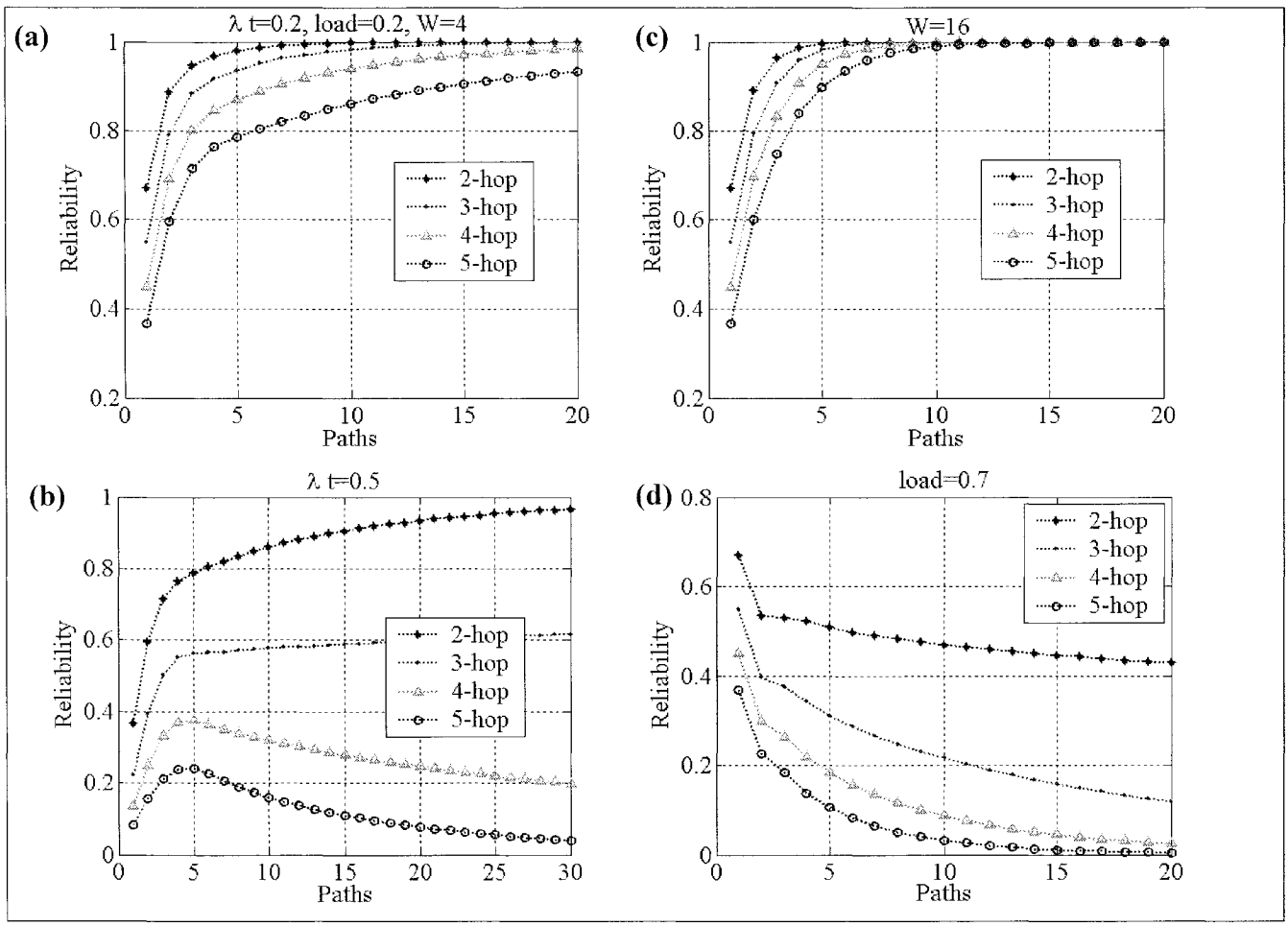

Figure 5.2: The effect of modifying multiplexing degree, failure rate, and network load on the network reliability.

It is easy to show that the number of hops defined by open link restoration protocol always is smaller than or equal to the counterpart protocols.

Network reliability is also related to Mean-Time-To-Repair (MTTR). Restoration delays including rerouting and configuration delays are determining factors to obtain MTTR. Accordingly, MTTR depends on the number of to-be-established switches between the end points of the failed link/path. Therefore, the associated MTTR of open link restoration is always smaller than or equal to the path or link restoration protocol. As a result, networks equipped by open link restoration mechanisms always provide the highest level of system reliability.

5.5 Simulation Assumptions and Results

We also evaluate the discussed protocols via simulation. We compare the related throughput for a variety of mesh networks shaped by matrices of $M+1$ by $N+1$ nodes. Mesh networks are chosen since they have remarkable potentials for restoration. For path restoration, Backward Restoration Scheme (BRS) is selected. This protocol searches backwardly for a restoration path between the d-s pair by examining alternate paths consecutively. For link restoration, we assume that a fast fault localization technique detects and localizes the failure by a pre-determined time. Then, a generalized loop-back restoration protocol is performed.

The parameters in this section are set based on similar simulation works in the literature. We assume that the location of a failed link is a random variable with a uniform distribution. In addition, we set fault localization time to be 0.05 sec. The configuration time of OXCs is considered fixed and assumed 0.125 sec. The average propagation delay is set to 0.025 sec for each link. The average computational delay for a decision making command is assumed to be 0.0001 sec (the amount of this parameter is set by considering the current computer clocks). We assume that each rerouting decision in average takes 0.001 sec according to our topology (note that for smaller or larger routing tables this amount needs to be modified and adjusted). Without loss of generality, we assume that each node performs full wavelength conversion, i.e., a lightpath could comprise any free wavelength on its way. It is assumed that requests are generated at each node according

to Poisson distribution. In our simulation, this parameter is set in a way that shows the throughput in an ample range. Data packets are randomly generated. To avoid having a zero data size, we add a small fixed amount equal to 0.125sec to data packets. We assume a zero holding time. This means that if unsuccessful, lightpath reservations are dropped immediately. (It is proven that the holding time does not have a significant effect on throughput [Qia96]. S-d pairs are uniformly distributed. For calculating the throughputs, we obtain the ratio of the successful transmissions to the whole number of generated calls (200 calls).

The effects of the following parameters are investigated:

- *Multiplexing degree*: The number of virtual channels (wavelengths) supported by each link is changed between 2 and 16.
- *Network size*: The simulation is run over matrix-formed mesh networks of 3x3 to 9x9 in order to examine the effect of increasing the network size.

Figure 5.3 displays the simulation results for different network sizes confronting 10 discrete link failures. The curves of the open link restoration protocol has been drawn with 90% confidence interval as they have been marked by "*".

In general, the results indicate that the open link restoration protocol always outperforms the counterpart protocols. Also, they reveal the following facts about the protocols:

- The link restoration protocol has the lowest throughput.
- The value of path restoration throughput is always less than open link restoration throughput and more than link restoration throughput.
- The open link restoration throughput shows at least 10% increase over the path restoration protocol for the multiplexing degree larger than 4.
- The path restoration throughput demonstrates at least 3% increase over the link restoration for the multiplexing degree larger than 4.

Acting similar to segmented protection protocols is the main reason behind the highest throughput in the case of open link restoration (note that the open link restoration protocol is a restoration technique not a pre-defined protection method). This means that the proposed protocol dynamically discovers the best segments for recovering affected paths. In other words, in contrast to segmented path protection protocols that pre-actively

switch to protection segments, open link restoration technique reactively defines restoration segments.

Accordingly, the latter technique utilizes the bandwidth more efficiently. However, the open link restoration protocol has a higher computational complexity and average restoration delay because of the modification procedure.

In general, the speed of the restoration protocols depends on several factors, including the computational cycle, the pace of fault detection and localization, the switching delay, the average propagation delay, the network size as well as the traffic. Thus, restoration times are really case-related and can be noticeably varied from a network to another.

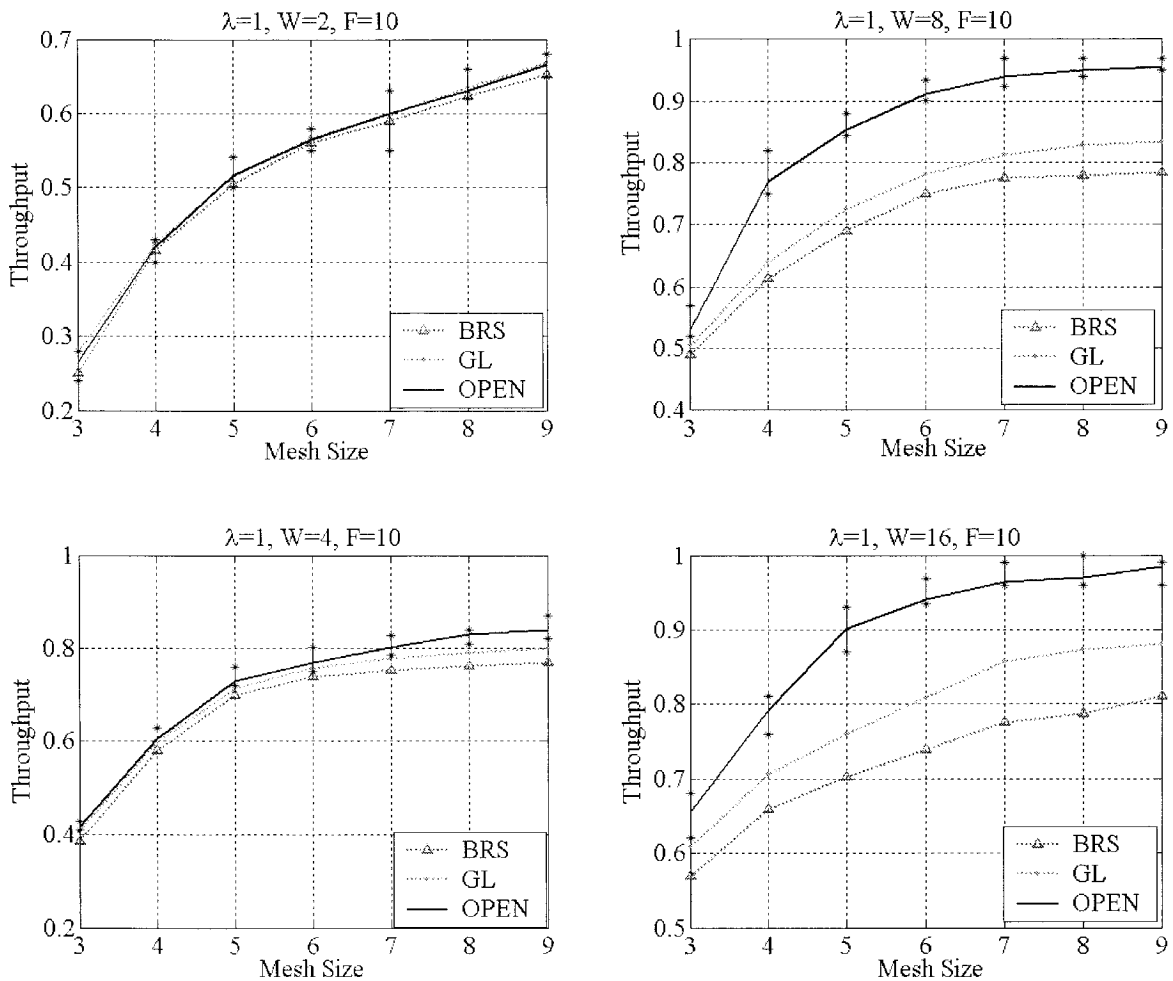

Figure 5.3: Comparisons between the restoration protocols in terms of throughput for various network sizes and different multiplexing degrees.

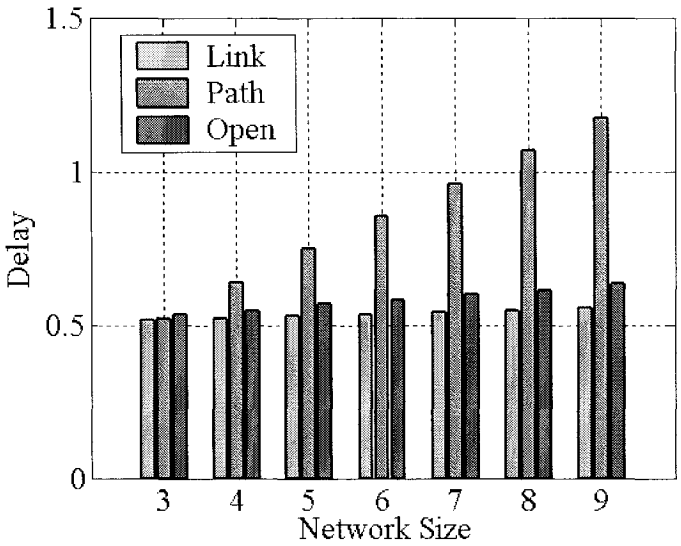

Figure 5.4: A comparison between the delays of the examined Protocols.

We assume that the associated distance between an s-d pair is proportional to the network dimension [Sic04a] and defined by the expected value of lightpath lengths based on M and N, $E[d_{sd}] = \frac{M+N+4}{3}$.

We also assume that network is under low traffic load. As a result, restoration routes would be the shortest alternate paths in the network. Figure 5.4 shows the associated restoration delays of the compared protocols based on our previous assigned values of parameters.

Note that link restoration could be always the fastest if the following conditions are met:

- Fault detection and localization delays are insignificant with respect to restoration delays. This condition requires deploying fast failure localization schemes which are typically executed in the optical layer.
- Switching delays are much higher than the propagation delays. Note that the introduction of high-speed switches has reduced the gap between switching and propagation delays in backbone optical networks.
- Restoration is achieved in the link layer (OMS). This way, the entire group of affected paths is recovered by an identical restoration loop and analogous processes are avoided.

- Restoration loops are the shortest loop comprising limited number of links. This condition is usually met if the network is under low traffic load.

5.6 Conclusion

This chapter focuses on the design and performance evaluation of a novel dynamic restoration mechanism for optical telecommunication networks. Accordingly, it proposes an efficient restoration protocol called open link restoration protocol. The open link restoration protocol employs a modified form of the generalized loop-back algorithm, namely open loop-back protocol, for detecting a restoration loop. Since open loop-back is allowed to consider all links of a failed path for forming a restoration loop, the resulting restoration path may comprise a set of redundant links. In such a case, the restoration path requires some modifications and verification. The modification process is accomplished at the source node by comparing the restoration loop with the failed path. Consequently, the mutual links between the failed path and the restoration loop are released via a corresponding signal sent by the source toward the destination node. Due to the modification phase, the restoration process of the open link restoration protocol becomes slightly more complex and lengthy than that of the link restoration protocol. The open link restoration protocol always discovers the shorter restoration paths and consequently provides a higher reliability than the examined link restoration protocol (Generalized loop-back). However, since open link restoration protocol is an adaptation to the link restoration protocol, in extreme circumstances, its defined restoration path could be identical to that of the link restoration protocol. In addition, the simulation results prove that the open link restoration protocol outperforms the examined path restoration protocol (Backward restoration scheme) by increasing the network throughput over 10% and maintaining a lower average restoration delay at the same time.

Chapter 6. Signaling Fault Localization Protocols

6.1 Introduction

PROVIDING resilience against failures is an important requirement for many high-speed networks. Similarly, equipping very high-bandwidth transparent optical networks with protection and restoration components is an imperative necessity. Although failures cannot be avoided, quick detection, identification, and restoration make networks look more robust and reliable. Therefore, it is essential to develop fast and dependable fault localization and restoration mechanisms. To accelerate the speed of recovery, fault localization and restoration schemes should be performed in the optical layer. Generally speaking, these techniques are feasible through different layers, but it is preferable to recover all-optical networks in the optical layer to preserve the full transparency and the high bandwidth. Accordingly, in this chapter we propose two novel fault localization schemes which function in the optical layer (the path sublayer). The proposed techniques use signaling in the control network through either the control (supervisory) or established data channels to locate a failure. We evaluate their performance in terms of restoration delay.

6.2 Fault localization

Network survivability or the network ability to tolerate failures is a vital feature of every reliable network. Protecting and restoring a network from service disruption becomes even more important as the effective bandwidth increases. To recover from fault occurrences, dynamic restoration protocols are preferred and recommended over the pre-planed dedicated ones due to their scalable capability and because of their efficient resource utilization.

Fast recovery is achieved by means of fast fault localization and restoration. There are several fault localization proposals to monitor channel continuity and quality using transceivers power levels. However, monitoring equipment is expensive and employing

them in all-optical networks removes full transparency of pure optical communications. In other words, when a failure occurs in all-optical networks, the network edge routers may be able to detect the existence of a fault, but the exact location of the fault cannot be indicated. No advanced optical technique is introduced at this time [Gha00]. In addition, SONET/SDH detection techniques, which are significantly developed and widely used, with current features, are not applicable to transparent WDM networks. Therefore, it is essential to devise compatible fault localization mechanisms. Lacking the ability to localize failures introduces significant problems as follows.

- The location of a failure is not reported to other nodes in order to modify their routing tables. Maintaining the accuracy of routing tables is necessary to prevent upcoming circuit requests to be examined over faulty links. Notice that the location of failed links can be flooded through the network very fast even with the distributed management mechanisms by means of OSCs.
- Every disconnected s-d pair is enforced to communicate through a completely disjoint path instead of only skipping the failed part. This disadvantage removes the entire affected path capacity from the network, which is not a resourceful approach.
- The interruption of service could be a harsh shock especially for highly loaded affected lightpaths. Assigning the disrupted load to completely disjoint alternate lightpaths can spread the shock around and creates chaos.
- On the other hand, fault detection pinpoints the affected link, which results in rerouting the traffic more informatively. In addition, it helps maintaining routing tables correct. Furthermore, fast fault localization in the optical layer prevents activating higher layers protection mechanisms and therefore, it would be more resource, time, and cost effective.

6.3 Fault Detection Protocol in SONET

Fault detection in SONET is accomplished using overhead bytes and also by electronic monitoring through Digital Cross Connects (DXCs) to detect loss of data or high Bit Error Rate (BER) [Dov99]. The role of this overhead is to ensure that the connection is well-managed before or after any service disruption. However, in the absence of alarm

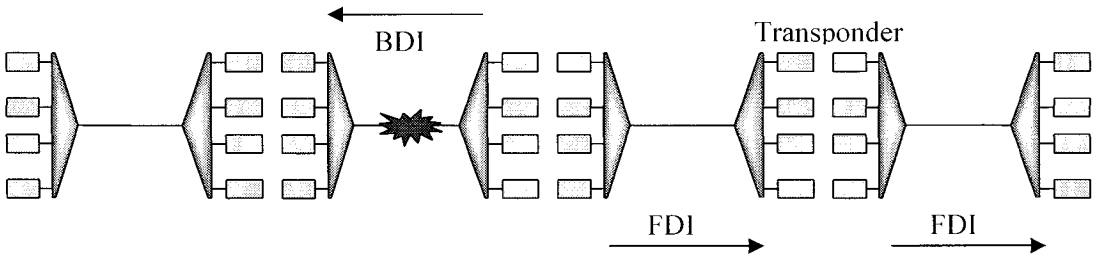

Figure 6.1: Fault detection in SONET using electronic monitoring and signaling.

management to report a failure, whenever a single failure happens, it may cause multiple alarms to be generated and consequently redundant/improper reactions to be taken by several nodes in response to the reported failure. Therefore, it is important to detect the root-cause alarm and suppress the remaining alarms. This can be accomplished by using two special signals. Forward Defect Indicator (FDI) that conveys the failure information to the downstream nodes. The signal is used to suppress false alarms generated by downstream nodes. The FDI signal is sometimes referred to as the Alarm Indication Signal (AIS). On the other hand, Backward Defect Indicator (BDI) is a response by the end terminal acknowledging the notification of a fault condition in the network. BDI is always used in conjunction with FDI. Similar to FDI, it traverses through the control network (OSCs). Figure 6.1 shows fault detection signaling in SONET.

A major reason to use the fault indicator signals is to stop triggering other nodes from protection switching [Ram98]. For example, the node adjacent to a failure detects the failure and may trigger a protection switch to reroute traffic around the failure. At the same time, other nodes downstream may decide to reroute traffic through other alternate paths. Receiving FDI ensures that the task of rerouting traffic is taking care of by only one node. This way other nodes will not invoke their protection switching mechanisms. Implementing the above-mentioned SONET fault detection/localization technique is not possible in wavelength-routed networks. This is due to employing OXCs, which in contrast to DXCs are incapable of electronic monitoring and measuring BER. However, optical sensors could economically and easily detect signal loss. In the next section, we introduce our fault localization technique, which uses optical sensors to pinpoint the location of failures.

6.4 Fault Localization Signaling Protocol

In this subsection we introduce a novel protocol called *Fault Localization Signaling Protocol (FLSP)* that can identify the location of a failure on a failed path in a distributed way. The main improvement of the proposed protocol is that it operates within the optical layer.

We presume that failures happen during the data transfer mode due to the short interval time of lightpath establishing phases. This implies that the destination node is aware of the source node and the setup route to the source before an interruption of service occurs. The idea behind the FLSP protocol is simple and is based on the following principles:

- Signaling nodes along the failed path.
- Receiving acknowledgement signals back from the signaled nodes.
- Discovering the first lost acknowledgment signal.

The protocol is activated once a connection is disrupted. This happens when the connected destination does not receive expected information from the source. Then, the destination initiates the FLSP protocol, which operates as follows:

- Fault localization signals are disseminated in parallel to all nodes along the disconnected path by the destination node (sink). These signals are short control packets traversing in the control network. There are equal numbers of localizing signals to the number of hops between the s-d pair. Fault localization signals could be numbered based on their distance to the destination (e.g., for an n-hop path, the closest link to the destination is numbered link $\#1$ and the farthest would be link $\#n$). However, labeling is not necessary because they only trigger associated nodes to reply by sending acknowledgement signals back to the destination node. Notice that attaching link labels would be useful for fault detection/localization mechanisms in higher layers in case the optical layer fails to indicate the location of the occurred failure. Fault localization signals traverse through the control network (OSCs), which is assumed secure. However, it is notable that even the severed control network does not impair the FLSP protocol, as it will be clarified further.

- Each recipient of a fault localization signal is triggered to send a small acknowledgment signal back to the destination through the faulty lightpath. These acknowledgment signals are short data packets that could contain the corresponding link numbers. For the sake of simplicity, the fault localization signals can be also reflected back to the destination. This procedure can be achieved by an optical sensor, which is invoked to send a signal back to the destination in response to the localizing signal or via mirroring the fault localization signals by the associated node in the OSC. Note that the returning signals should traverse through the data network.
- Acknowledgement signals are received by the destination. Each acknowledgement signal requires specific amount of time to reach the destination node based on its propagation time. The arrival of an acknowledgement signal guarantees an unbroken link. Either a timer or a counter could do the task of identifying which signal has been properly returned to the destination. A timer can be deployed to verify the propagation delays. A counter can complete the identification too. The counter counts the number of acknowledgment signals. Notice that in the latter case, there might be more than one absent checking signal. If only one signal is missing, then, the link adjacent to the source node has been disconnected (link #*n*). Otherwise, if there were more than one missing acknowledgment signals, the next number after the counted number would be the number of the broken link. In the case of having a disconnection in the control channels (a fiber cut case), the protocol operates in the same way. This is because, in this case, no fault localization (checking) signal could advance and consequently no acknowledgement signal is received from this point further.
- The location of the failure is broadcasted to other nodes.
- A fast link restoration protocol is launched.

6.5 Two Implementation models

Figure 6.2 a) shows how the FLSP protocol operates. As it has been depicted, six fault localization signals have been sent toward the source node consecutively. In response, each recipient independently disseminates an acknowledgement signal back to the destination. It can easily be seen that due to the channel cut only signals #1, #2, and #3

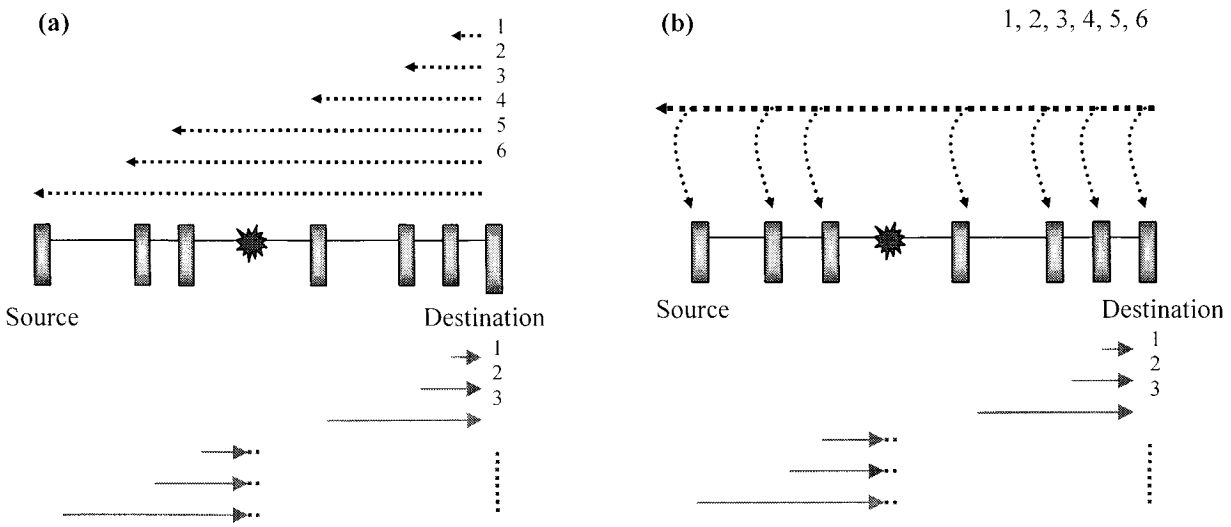

Figure 6.2: Two implementations of the FLSP protocol.

are received on time or counted to "3" totally. Therefore, link #4 is detected as the failed link. As it has been illustrated in 6.2 b), signaling the nodes has been structured in sequences. Another design, which improves the FLSP protocol performance, is to broadcast a single localization signal. Instead of sending several localizing signals (equal to the number of hops), only a single localizing signal is sent to the source. The broadcast signal will trigger each node on his way. Each triggered node, then, sends a responding (acknowledgment) signal back to the destination node.

Obviously, sending a single localizing signal decreases traffic in the control network, which results in saving control bandwidth. Also, broadcasting is more robust than sequential signaling due to this fact that sequential signaling is susceptible to signal loss. More details on FLSP performance evaluation and comparisons can be viewed in [Sic04b].

6.6 Rolling Back Signaling Protocol

In this section we introduce another proposed protocol called *Rolling Back Signaling Protocol (RBSP)* [Sic04a] that is able to identify the location of a failure more efficiently than FLSP. The main advantage of the proposed protocol is that it operates in the optical layer. The new protocol detects the failed link and then could recover the affected data by

link or path restoration. We presume that a failure happens during data transfers since the duration of lightpath establishments are relatively short and negligible. The presumption implies that the destination node is aware of the source node and the setup route, before any interruption of service occurs. It is also assumed that no intermediate monitoring or analysis is possible which means intermediate nodes are unable to report high BER. However, supplementary equipments, including optical sensors, could be deployed at the output ports. The protocol will be activated once a connection is disrupted, when the connected destination does not receive expected data or the sink detects high BER. Then, the destination initiates the RBSP protocol which operates as follows.

- An alarming signal is backwardly disseminated toward the upstream neighbor by the destination node (sink). The alarming signal is a short control packet which only warns the upstream node of data disruption. We call this alarm a Rolling Back Alarm (RBA) since it is backwardly transmitted. This alarm traverses through the control network (OSCs). It is notable that even the severed channel could not impair the RBSP protocol in view of the fact that the disconnection will stop the progress of the RBA alarm to upstream nodes.
- If the link has not been cut, the recipient of RBA (or the output sensor of the port) is triggered to send an acknowledgment signal (ACK) to the downstream node(s). This action is necessary to suppress any restoration process at upstream node(s). The node also sends the RBA alarm to its upstream neighbor along the path (rolls back the RBA alarm).
- If the link has been disconnected, the recipient is not triggered and consequently, does not acknowledge receiving RBA. The sender after a short waiting time, comprehend that there is a disconnection. Then, a NACK signal is send to downstream nodes. The sender launches a link restoration protocol. Generalized Loop-back restoration scheme (GL) could be applied for a fast link restoration.
- Or, once the NACK signal is received by the destination, it will activate a path restoration protocol such as Backward Restoration Scheme (BRS) to reroute the affected traffic through an alternate disjoint path.

- The identified location of the failed connection is flooded through the network. This action is necessary to validate routing tables and also to prevent upcoming calls from blocking as a result of the disconnection. As an advantage, receiving the failure information by other nodes could accelerate their restoration processes.

Figure 6.3 illustrates how the RBSP protocol functions. As it has been shown, the alarming signal (RBA) is rolled back toward the upstream nodes until it reaches the disconnection. Then, the time-out node (waiting for an ACK signal) sends a NACK signal to downstream nodes. In response to the NACK signal, BRS or GL is activated to bypass the failed part and restore the connection through the shortest path between two disrupted ends or two failed points.

6.7 Performance Evaluation

To evaluate the performance of the proposed protocol, we calculate the restoration time for a regular matrix-formed mesh network. Mesh networks are the most popular kind of networks, which have considerable potential for restoration purposes since they are at least two-edge connected. For the sake of simplicity, we assume that the restoration path can be established without either experiencing failure or going through more than n extra links. In the first step, we calculate the required restoration time of RBSP for an example path. Then, we compare the calculated results with that of BRS. The BRS scheme figures out a restoration path between the d-s pair by probing alternate paths in parallel [Zeh02].

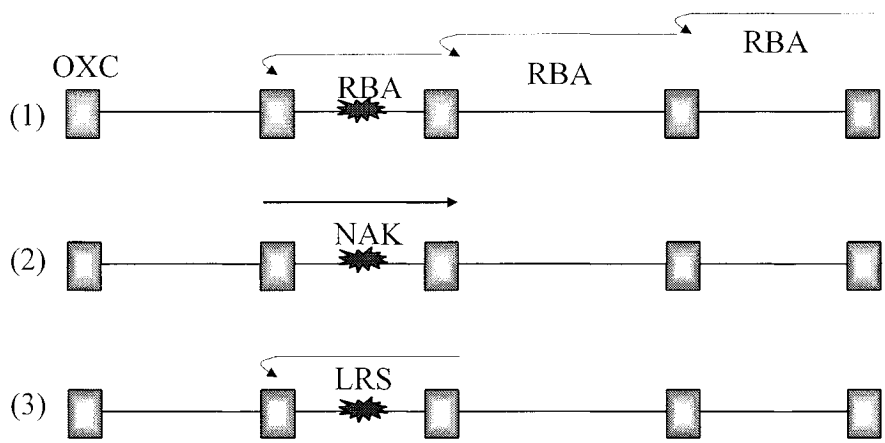

Figure 6.3: The illustration of the required steps to detect and restore a failure in RBSP.

In the next step, we generalize the calculations and compare the results. We define the connection restoration time for a broken connection as a time taken to set up the restoration path. We also define the following notation to calculate restoration times.

T_{BRS} Time taken to restore a path by the BRS scheme.

T_{RBSP} Time taken to restore a path by the RBSP protocol.

T_d Time taken for the destination to detect a failure.

T_r Time taken for a node to compute a new route (routing delay).

T_s Time taken for a node to configure a switch (switching delay).

T_c Time taken for a node to check signal loss condition.

T_p Propagation delay between two adjacent nodes considered equal for all links.

N_{sd} Number of hops between a source-destination pair.

N_d Number of hops from a severed connection to the destination.

In the following example, we assumed a 6-hop working path of a mesh network has been disconnected at the 4th hop. We also showed few extra links and nodes around the failed link since they will be used for setting up the restoration loop and path.

As it has been depicted, the restoration time by the RBSP protocol is,

$$T_{RBSP} = T_d + 5T_c + 14T_p + 4T_s.$$

Similarly, the corresponding restoration time for the BRS scheme would be,

$$T_{BRS} = T_d + 8T_r + 16T_p + 4T_s.$$

To generalize the problem, we assume that a failed link is successfully restored by a loop of n links. Therefore, the connection restoration time with RBSP is estimated as follows:

$$T_{RBSP} = T_d + (N_d + 1)T_p + (N_d + 1)T_c + nT_p + (n+1)T_s + (N_d + n - 1)T_p. \quad (6.1)$$

On the other hand, the required restoring time by the BRS scheme is equal to,

$$T_{BRS} = T_d + (N_{sd} + n - 1)T_p + (N_{sd} + n - 1)T_r + (n+1)T_s + (N_{sd} + n - 1)T_p. \quad (6.2)$$

As a result, if T_c and T_r are assumed equivalent, the difference in the restoration times is,

$$T_{BRS} - T_{RBSP} = 2(N_{sd} - N_d - 1)T_p + (N_{sd} - N_d + n - 2)T_r. \quad (6.3)$$

It is reasonable to assume that the location of a failure is uniformly distributed along the path. From this assumption we have,

$$E[N_d] = \frac{N_{sd}+1}{2}. \tag{6.4}$$

Therefore,

$$T_{BRS} - T_{RBSP} = (N_{sd}-3)T_p + \frac{(N_{sd}+2n-5)}{2}T_r. \tag{6.5}$$

The distance between two nodes *s, d* would be,

$$d_{sd} = |x_s - x_d| + |y_s - y_d|.$$

Where (x_s, y_s) and (x_d, y_d) are the *x-y* coordinates of the source and destination, respectively. We assume that the coordinates of source and destination are independent and also x and y coordinates of each are independent uniform random variables distributed. Thus, for a mesh matrix-formed network of *(M+1) (N+1)* nodes, we have:

$$P(x_s = m) = P(x_d = n) = \frac{1}{M+1}, \qquad \forall m, n = 0, 1, \cdots, M,$$

$$P(y_s = p) = P(y_d = q) = \frac{1}{N+1}, \qquad \forall p, q = 0, 1, \cdots, N,$$

Then, the expected value of the distance is given by,

$$E[d_{sd}] = E[|x_s - x_d|] + E[|y_s - y_d|]$$

In order to calculate the average distance we use the distribution of $|x_s - x_d|$ and $|y_s - y_d|$ as follows, respectively:

$$P(|x_s - x_d| = i) = \sum_k P(x_s = k)P(x_d = i-k),$$

$$P(|y_s - y_d| = j) = \sum_k P(y_s = k)P(y_d = j-k),$$

where k in the above summations take all possible values. We easily get:

$$P(|x_s - x_d| = i) = \frac{2(M+1-i)}{M(M+1)}, \qquad \forall i = 1,2,\cdots,M,$$

$$P(|y_s - y_d| = j) = \frac{2(N+1-i)}{N(N+1)}, \qquad \forall j = 1,2,\cdots,N,$$

Therefore, we can get the following results,

$$E[|x_s - x_d|] = \sum_{i=0}^{M} i\left(\frac{2(M+1-i)}{M(M+1)}\right) = \frac{M+2}{3}. \tag{6.6}$$

Similarly, we have

$$E[|y_s - y_d|] = \frac{N+2}{3}.$$

As a result, we can define the expected value of the distance based on M and N as follows:

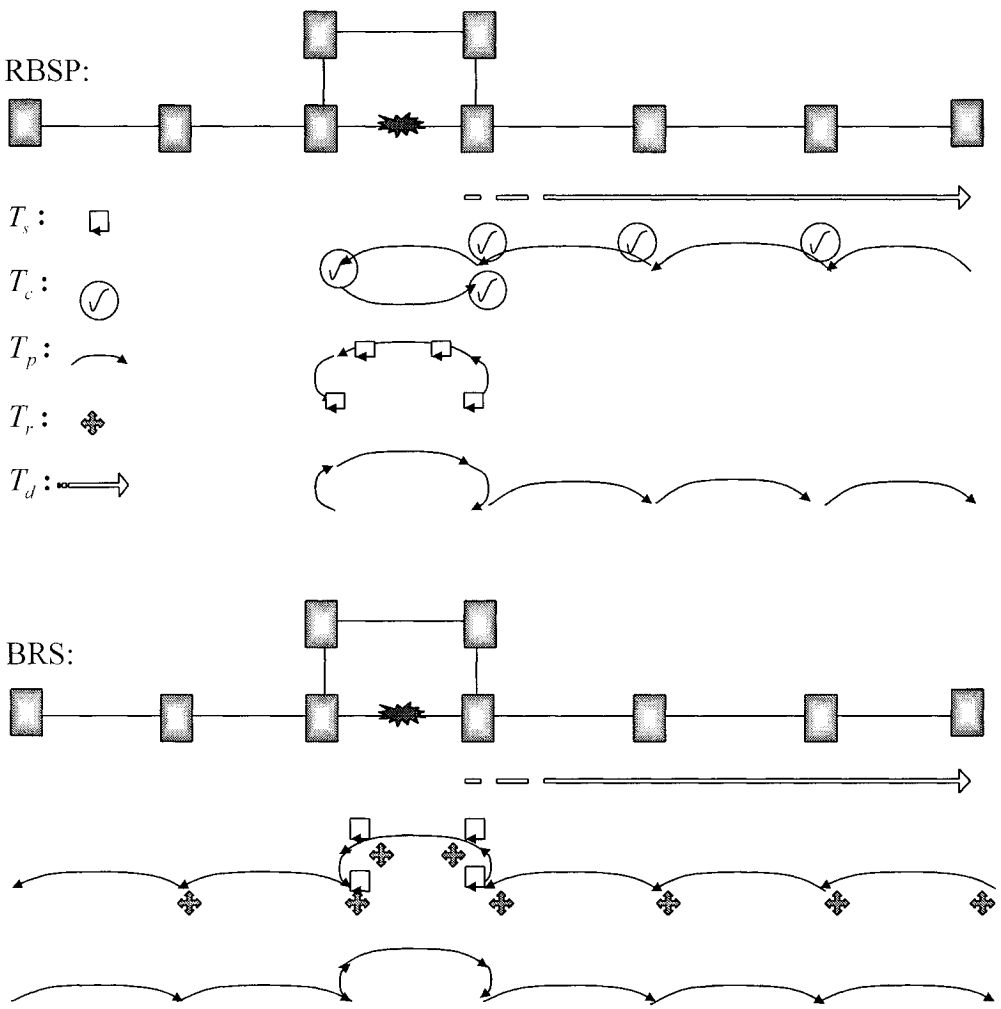

Figure 6.4: Illustration of different time delays for implementing RBSP and BRS.

$$E[d_{sd}] = \frac{M+N+4}{3}. \tag{6.7}$$

$$\text{or } E[N_{sd}] = \frac{M+N+4}{3}.$$

Hence, taking into consideration that $E(d_{i,j}) = N_{sd}$, equation (6.5) is generalized to,

$$T_{BRS}|_n - T_{RBSP}|_n = \frac{(M+N-5)}{3}T_p + \frac{(M+N+6n-11)}{6}T_r. \tag{6.8}$$

Although, the first part in equation (6.8) would be negative for $M+ N<5$, in view of the fact that the second part is always positive since $n \geq 2$ and $T_p << T_r$, we could conclude that the resulting equation is always positive.

Another important point is that for a fixed n, the result of equation (6.8) increases linearly by the network size. As a result, RBSP considerably outperforms BRS in large networks.

6.8 Conclusion

In this chapter we proposed two novel fault localization protocols called Fault Localization Signaling Protocol (FLSP) and Rolling Back Signaling Protocol (RBSP), which both function in the optical layer. The FLSP fault localization protocol is similar to Pingger-Pongger protocol [Lu-04] with this difference that the proposed protocol is a distributed fault localization technique. The FLSP protocol could be implemented in sequential and broadcast ways. Similar to FLSP, RBSP disseminates controlling signals in the OSCs to detect the location of a failure. Once a destination (sink) detects the failure, it sends an alarming signal (RBA) to its upstream neighbor backwardly along the lightpath. In response, the recipient sends an ACK signal back to the sender and rolls back the RBA signal to its upstream node. If, otherwise, the recipient of RBA does not see the RBS, no signal is sent back to the sender. Then, after a short waiting period, the sender sends a NACK signal to the downstream nodes and also launches a link restoration process (e.g., the generalized loop-back restoration mechanism). The proposed protocol comprises several advantages: due to locating the faulty connection, the restoration could be achieved by link restoration which is generally faster than path

restoration; the protocol takes place in the optical layer and maintains a high-speed; the RBSP protocol validates routing tables and secures upcoming calls from trapping into the failed link. We calculated the restoration time of RBSP and compared it with that of Backward Restoration Scheme (BRS). The results reveal that RBSP has a much shorter restoration delay. This achievement comes at the cost of increasing traffic in the control network (OSCs) and adding optical sensors to output ports.

Chapter 7. Limited-perimeter Vector Matching Fault Localization Protocol

7.1 Introduction

THIS chapter deals with the issue of fault localization in the all-optical domain. In order to provide the maximum level of transparency, the protocol skips any optical power monitoring or spectrum analysis at intermediate nodes. Accordingly, the protocol is activated when a set of alarming sinks emerges in the system. The proposed fault localization protocol is implemented by means of five distinct phases: pausing, flooding, multicasting, matching, and concluding phases. The key point, which distinguishes the proposed protocol from link-state protocols, is that the fault localization is dynamically restricted to a relatively small area called "limited-perimeter". Owing to restricting the fault localization area, the associated time and space complexities are dramatically narrowed down. In addition to decreasing the computational complexity, the restriction enables the protocol to track down simultaneous multi-failure. To evaluate the performance of the proposed scheme, the time and space complexities are calculated and the results are compared with those of a boundless link-state protocol.

7.2 Path or link Restoration

Debates over the pros and cons of dynamic restoration techniques, including path and link restoration protocols continue as the introduced versions are improved and new versions are proposed. However, the following comparisons highlight some of their merits and drawbacks and justify our focus to develop a fast and capable fault localization technique to be employed by link restoration.

Link restoration is a single-hop point-to-point, while path restoration is an end-to-end recovery technique. Link restoration restores the affected path with generally less number of links than path restoration. Consequently, link restoration decreases the number of to-

be-reconfigured switches and the associated restoration delays. In addition, since link restoration keeps a number of working links (segments of the working path) in place, it preserves the network load balancing. In other words, link restoration does not create a sudden chaos in the system. However, link restoration could cause congestion around the failed link. In addition, this technique may set up considerably long restoration loops. Although, this phenomenon could severely alter the protocol functionality, typically it only happens in highly loaded networks. Link restoration is always followed by a fault localization interval. The fault localization time for large network could be long and undesirable. However, at the expense of the fault localization delay, routing tables are mended and restoration processes are completed faster.

In contrast, path restoration implementation does not require any fault localization and is started immediately after a fault alarm is activated. However, path restoration could be time-consuming for distant s-d pairs and also when networks are heavily loaded. This is mostly due to performing reconfiguration procedures and often repeating these procedures for alternate paths before successfully establishing restoration paths.

Furthermore, link restoration only removes the failed link capacity by pinpointing the failure location, while path restoration practically removes the affected path capacity from the network resources by excluding the entire set of affected paths.

Thus, we focus on link restoration due to the reviewed merits and in particular its high restoration speed. Consequently, in this work we propose a fast fault localization scheme applicable to link restoration.

7.3 Limited-perimeter Vector Matching Fault Localization Protocol

In this section we explain in detail our proposed protocol called Limited-perimeter Vector Matching fault localization protocol (LVM). We assume that each fiber is multiplexed in the applied network. Then, the number of lightpaths traversing a link could be as large as the multiplexing degree. In general, once such a link fails, the following scenario will be shaped:

A group of affected sinks detects signal loss or high BER by dropping the data. Since we assume that the occurred failure is only detectible by end terminals, for instance at the connected OXCs or OADMs, the closest sink to the failed link would be the first one that discovers the failure. The LVM protocol is launched in reaction to the alarming sink as follows:

- The alarming sink is recognized as a *potential executive-sink*. If the linked OXC drops the data of more than one lightpaths passing through the failed link, then the shortest lightpath among them is chosen as the associated path to the potential executive-sink.
- The potential executive-sink pauses for a pre-determined period of time to allow the entire network nodes to reach their steady states. The pausing interval is very short and defined based on the network characteristics, including the diameter and the average propagation delay.
- Every potential executive-sink reports its status to the network nodes by flooding a corresponding alarm. The alarm encapsulates the node ID and also the path length. Accordingly, every potential executive-sink receives other potential executive-sinks information. Figure 7.1 shows a group of potential executive-sinks along with their ID and associated path lengths, in an example network with 25 nodes having the link 12-15 failed.

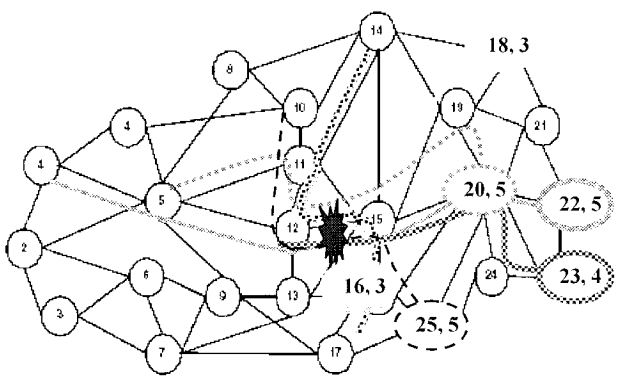

Figure 7.1: Potential executive-sinks compete for conducting the LVM protocol.

- Each recipient then compares its path with those of other recipients. If its path length is the smallest among others, it recognizes that it can act as the *executive-sink* and conduct the fault localization. In case there are several equal path lengths, the recipients also compare IDs. In such a case, a sink with the smallest ID is selected as the executive-sink. Other criteria for selecting the executive-sink in such a case are also possible.
- In the next step, the executive-sink creates a vector of its affected lightpath links. Let us call this vector Affected-Link-Vector (ALV). Meanwhile, the executive-sink also defines the limited-perimeter, which comprises all the ALV neighboring nodes and links. Figure 7.2 illustrates the ALV path and the coupled limited-perimeter, in our example.
- The executive-sink multicasts a copy of the ALV vector to the nodes within the limited-perimeter. Receiving the ALV vector suppresses the recipients from taking further steps as potential executive-sinks. Note that the suppressing action is performed only on the nodes within the defined limited-perimeter. Consequently, those potential executive-sinks, which have not received ALV, have a chance to run a second LVM. To avoid having either coincident LVMs for a single failure or inefficiently fault localization, a second LVM has to be performed only after checking the result of the first LVM.
- Each ALV recipient creates its own link-vector and compares its elements with the

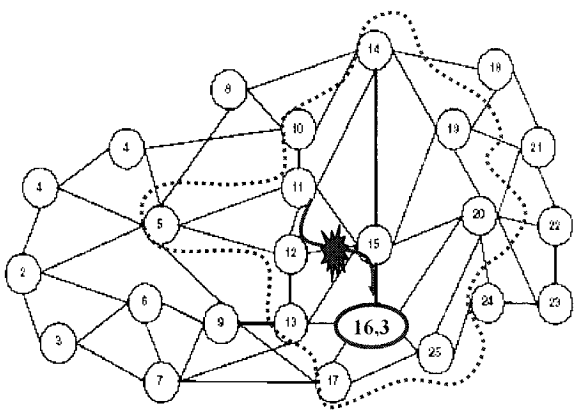

Figure 7.2: The limited perimeter defined by sink 16.

ALV vector. The outcome of elements matching is demonstrated by means of a binary vector correspondence to ALV. For an affected lightpath, the corresponding cell of the binary vector is set to 1 if the same link is found in the recipient link-vector. In contrast, the corresponding cell of the binary vector is set to 0 if the examined link is not matched with any elements of the recipient link-vector. For a working lightpath, reverse order of bits show analogous conditions. The matching vector along with the recipient status ("1" for working and "0" for failed) is reported back to the executive-sink.

- The executive-sink collects the binary vectors from the nodes within the limited-perimeter. It performs logical AND on the binary vectors to define the disconnection. If the executive-sink could not define the disconnection in this stage (if there are more than one bits set to "1" in the binary vector), it could launch the second round of search by extending the limited-perimeter to a further edge. For instance, the limited-perimeter could expand over the neighbors of the considered nodes.
- The executive-sink broadcasts the location of the failure. Then, network nodes update their routing tables to facilitate ongoing restoration procedures and guarantee upcoming connections, not to try over the failed link.
- In the last step, a link restoration protocol such as the generalized loop-back protocol [Med02] is exploited to restore the affected paths.

7.4 A Descriptive Example

In this section we describe the procedures of the LVM protocol by means of an example mesh network. For the sake of simplicity we assumed that only few lightpaths have been established in the network. It is obvious that increasing the number of lightpaths will accelerate the matching process since additional binary vectors are supplemented. We demonstrate each link by its two end-nodes. Let us assume that link {12- 15} is disconnected while the established lightpaths in the network are: P_1= {5- 12- 15- 19- 20}, P_2= {10- 11- 12- 15- 16- 25}, P_3= {3- 6- 9- 13- 16}, P_4={11- 12- 15- 16}, P_5= {14- 11- 12- 13}, P_6= {14- 11- 12- 15- 16- 17}, P_7= {4- 5- 12- 15- 20- 22}, P_8={5- 12- 15- 20- 24- 23}, P_9={17- 25- 24- 23}, and P_{10}= {12- 15- 19- 18}. Then, the disconnection affects

several paths, including P_1, P_2, P_4, P_6, P_7, P_8, and P_{10}. Since path P_4 has the shortest path and also the smallest ID, sink 16 wins the competition among all potential executive-sinks and becomes the executive-sink. Accordingly, the related limited-perimeter, comprising all the neighboring nodes of the executive-sink path, is specified as it has been illustrated in Figure 7.2. Hereafter, only paths within the defined limited-perimeter are involved in the fault localization procedure.

As the first step in the matching procedure, the ALV vector of executive-sink is formed by sink 16 and disseminated to the limited-perimeter nodes. As a result, all the sinks receive a copy of ALV as follows:

11- 12	12- 15	15- 16	ALV

If we assume that the ALV recipients respond based on their distance from the executive-sink, then, sink 17 would be the first recipient which makes its own link-vector and matches up its elements to the ALV vector.

14- 11	11- 12	12- 15	15- 16	16- 17

The related matching vector of this sink is tailored by setting the corresponding cells of an equivalent binary vector to ALV. The path status is attached and the resulting binary vector is sent back to the executive-sink. Note that in practice, only a binary vector of 4 bits is sent back to the executive-sink.

11-12	12-15	15-16	Status
1	1	1	0

Next, sink 13 generates its link-vector and compares it with ALV. Although this path has not been affected by the failure, it should reply to the executive-sink since it has been extended within the limited-perimeter.

14- 11	11- 12	12- 13

Since this path is in the working state, a reverse order of bits is presented in the associated binary vector.

11-12	12-15	15-16	Status
0	1	1	1

Then, after receiving this vector, the executive concludes that link {11-12} has not failed. Later, sink 25 makes its link-vector and compares it with the ALV vector.

10- 11	11- 12	12- 15	15- 16	16- 25

As a result, its mutual links and status is reported back to the executive-sink as follows,

11-12	12-15	15-16	Status
1	1	1	0

Then, sink 20 generates its link-vector to discover the common terms with ALV.

5- 12	12- 15	15- 19	19- 20

The following matching vector is disseminated to the executive-sink.

11-12	12-15	15-16	Status
0	1	0	0

Although executive-sink is still receiving binary vectors from other sinks (P_7, P_8, P_{10}), it is able to discover the failed link {12-15} at this stage.

111-12	[illegible]	15-16
0	1	0

Finally, the executive-sink broadcasts the failed link information to the network nodes.

7.5 Performance Evaluation

In this section we evaluate the proposed protocol in terms of time and space complexities. To simplify the calculations, we divide the fault localization procedure to the following discrete phases:

I. The pausing phase
II. The flooding phase
III. The multicasting phase
IV. The matching phase
V. The concluding phase

In Phase I, the network nodes pause to reach their steady-state conditions. The related pausing period is determined by the average distance between s-d pairs, i.e. $E[d_{sd}]$ and the expected value of propagation delays between neighboring nodes, i.e. α .

$$T_1 = E[d_{sd}]\alpha , \tag{7.1}$$

During the second phase, potential executive-sinks flood their information through the network. In this phase, the sink with the shortest path and ID wins the competition and

conducts the LVM fault localization procedures. The associated flooding time interval is given by,

$$T_2 = D\alpha\,, \tag{7.2}$$

Where D stands for the network diameter based on the number of hops.

Delay in multicasting phase is also related to $E[d_{sd}]$ since the maximum path lengths within the limited-perimeter comprise at most two additional hops to the shortest path. Therefore, the associated delay equals:

$$T_3 = (E[d_{sd}]+2)\alpha\,, \tag{7.3}$$

The order of computational complexity of matching two vectors with lengths L_1 and L_2 *is* $O(L_1L_2)$. Therefore, the computational complexity of comparing each two vectors in the matching phase is on the order of $O(d_{s_1d_2}d_{s_2d_2})$. Thus, the computational complexity in average is on the order of $O(E[d_{sd}]^2)$. If we assume that all recipients perform this process concurrently, the related delay for phase IV is proportional to the computational complexity as follows:

$$T_4 = E[d_{sd}]^2\ \beta\,, \tag{7.4}$$

Where β represents the computational cycle.

Finally, we assume that the last phase could be completed in a single time cycle, β .

Therefore, the overall delay for the defined phases is calculated as follows:

$$T_{LVM} = \sum_{i=1}^{5} T_i = E[d_{sd}]\alpha + D\alpha + (E[d_{sd}]+2)\alpha + E[d_{sd}]^2\beta + \beta\,, \tag{7.5}$$

Let us now calculate the time complexity based on a matrix-formed mesh network with $M+1$ by $N+1$ nodes. The distance between any s-d pairs is calculated as follows. The detail of calculation can be found in Section 6.5.

$$E[d_{sd}] = \frac{M+N+4}{3}\,, \tag{7.6}$$

Now that we calculated the expected value of the distance between s-d pairs, we are able to present the time complexity based on M and N.

$$T_{LVM} = \sum_{i=1}^{5} T_i = (M+N)\alpha + \left(\frac{M+N+7}{3}\right)2\alpha + \left(\frac{M+N+4}{3}\right)^2 \beta + \beta, \qquad (7.7)$$

If we assume that $\beta \geq \alpha$, the calculated delay is mainly related to the quadratic term. However, α for some networks could be much larger than β. Then, we investigate the effect of varying these parameters on T_{LVM} to find the dominant term. To reduce the number of variables, let us assume that $M = N$ and $\beta = 1$. We vary α based on N, i.e., $\alpha = N, 2N, \cdots$. Figure 7.3 illustrates the results of our comparison. The plotted curves reveal that when $\alpha = 2N = M + N$ both terms have the equal effect on the delay (note that for an example matrix of $M = N = 5$, α is found to be 10 when β is set to 1). However, as α increases, its related delay becomes dominant.

On the other hand, the space complexity can be calculated by defining the following phases:

I. The multicasting phase

II. The matching phase

In the first phase, nodes within the defined limited-perimeter received a copy of ALV. Assuming that *m* bits represent each link, we have

$$S_1 = mE[d_{sd}], \qquad (7.8)$$

In the matching phase, the executive-sink receives matching vectors with the ALV length from *n* recipients. Then,

$$S_2 = nE[d_{sd}], \qquad (7.9)$$

Therefore, the space complexity would be equal to,

$$S_{LVM} = (m+n)E[d_{sd}]$$

$$S_{LVM} = (m+n)\frac{M+N+4}{3}, \qquad (7.10)$$

For comparison, let us consider a distributed link-state protocol that does not impose any restriction on the searching area. Similar to VLM processes, we calculate a lower bound on the time complexity as follows:

$$T_X = (M+N)3\alpha + \left(\frac{M+N+4}{3}\right)^2 \beta + \beta, \tag{7.11}$$

Correspondingly, the space complexity would be,

$$S_X = (m+M+N)\left(\frac{M+N+4}{3}\right), \tag{7.12}$$

Comparing these results with those of the LVM protocol demonstrates the advantages of the proposed protocol in accelerating the fault localization process and saving memory.

$$T_X - T_{LVM} = \left(\frac{4(M+N)-14}{3}\right)\alpha, \tag{7.13}$$

$$S_X - S_{LVM} = (M+N-n)\left(\frac{M+N+4}{3}\right), \tag{7.14}$$

In addition to the proven enhancements, LVM would be able to localize multi-failure. Multi-failure detection and localization is possible if the defined limited-perimeters do not overlap (no link or nodes is shared between limited-perimeters). Note that multiple failures cannot be localized simultaneously since only a single executive-sink is recognized after the flooding phase. Consequently, a second executive-sink can be

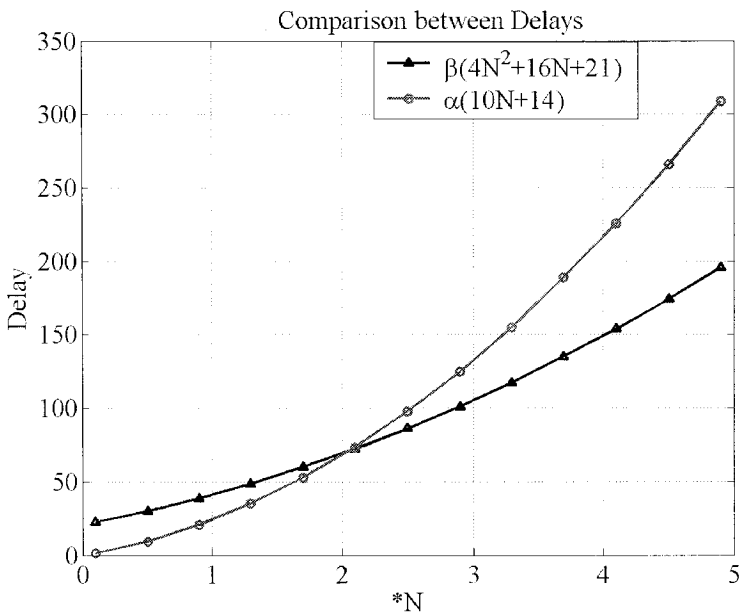

Figure 7.3: A comparison between the propagation and computational delays.

discovered right after the first executive-sink multicasting phase. However, this may produce multi-running LVMs for a single failure and cause inefficient processing. To eliminate this problem, a second LVM has to be performed only after checking the result of the first LVM. Thus, multiple failures could be consecutively identified by a short interval between them. Since the time difference between two tandem failures pragmatically is not zero, we infer that LVM can fast and effectively respond to multi-failure cases.

7.6 Conclusion

In this chapter we proposed a novel fault localization protocol applicable to all-optical networks. The proposed protocol namely, Limited-perimeter Vector Matching fault localization protocol (LVM) is highly efficient due to operating in the optical layer. The protocol implementation could be described in five steps. First, the system is paused to reach the steady-state condition. Second, the connected sink to the shortest affected path is identified to conduct the fault localization protocol (the executive-sink is chosen). Third, based on the executive-sink path links, a relatively small region called "limited-perimeter" is defined. Then, the executive-sink path information (Affected-Link-Vector) is disseminated to the nodes within the limited-perimeter. Forth, each recipient compares its own link-vector with ALV and sends the matching results via an ALV correspondence binary vector back to the executive-sink. In the last step, the executive-sink, through performing logical AND on the binary vectors, discovers the failed link. We evaluated LVM in terms of time and space complexities against a similar boundless centralized protocol. The analytical results reveal that the time and space complexities are less than those of the counterpart protocol. The difference in performance grows as the network is expanded to a larger one via increasing the nodes (the diameter) or/and the distance (the amount of average propagation delay). These outcomes were logically expected since LVM could dynamically define and limit the fault localization area. In addition to providing lower complexities, the protocol is able to pin down and locate multiple failures if they occur in non-overlapped limited-perimeters. Thus, we conclude that LVM

is a powerful fault localization method with remarkable advantages for survivable transparent all-optical networks.

Chapter 8. Conclusion

8.1 Concluding Remarks

NETWORK survivability or the network ability to withstand failures is an important feature of every reliable network. Accordingly, creating resilience in high capacity transparent WDM networks is a basic requirement for designing the next generation of optical networks. When financial and physical resources are limited or where scalability is needed, dynamic restoration protocols are preferred over protection mechanisms for recovery purposes. However, most of the introduced and developed restoration techniques for opaque optical networks are not applicable to transparent all-optical networks. Furthermore, the restoration techniques exclusively devised for transparent optical networks reroute affected traffic to completely disjoint paths to skip fault localization. Rerouting traffic to alternate disjoint paths, in the absence of fault localization, introduces several problems, as detailed in Section 6.2.

To upgrade recovery techniques, compatible/competent fault localization mechanisms need to be deployed in all-optical networks. Although, several fault localization techniques have been previously introduced, most of them are not well-matched to optical networks. In addition, SONET/SDH fault detection techniques, which are significantly developed and widely used, with their current features are not applicable to all-optical networks. Therefore, it is essential to devise a specific kind of fault localization to be used as a perquisite for connection restoration mechanisms and also preserve the high speed of all-optical networks. This thesis takes the first necessary steps in developing high performance fault localization and rerouting techniques by introducing novel signaling algorithms. These protocols use signaling to either detect the location of a failure or to set up the most efficient restoration path. We have introduced two signaling methods for fault localization namely, Fault Localization Signaling Protocol (FLSP) and Rolling Back Signaling Protocol (RBSP). Applying these fault

localization protocols increases the speed of fault localization, maintains the network load in a better balance, and also prevents blocking caused by invalid routing decisions. Consequently, it enhances bandwidth utilization and network performance. We have also proposed Limited-perimeter Vector Matching (LVM) protocol that highly enhances the fault localization performance in large transparent optical networks. We have analytically calculated the average restoration delays for those fault localization protocols and shown the resulting improvements. However, the improvements come at the price of increasing traffic in the Optical Supervisory Channels (OSCs).

In addition, we have introduced a novel bandwidth allocation protocol called Signaling Nested Reservation Protocol (SNRP). This protocol could noticeably increase throughput by reallocating reservations bandwidth to non real time data. We have statistically modeled and calculated the increase in the network throughput. We also proposed two novel dynamic restoration scheme called Minimum Weighted-path Restoration Protocol (MWRP), and Open Link Restoration Protocol. These protocols open a new window to high-efficiency dynamic restoration by partially path restoration. We have evaluated their performance through extensive simulation. Also, we have numerically proved that the proposed signaling protocols offer exciting enhancements in terms of system availability and reliability to all-optical WDM networks. However, due to sophisticated nature of these protocols, we observed increases in the computational complexities.

8.2 Future Research

As an extension to this work, the following list of open problems is suggested as potential future research areas:

The proposed fault localization mechanisms only locate faulty links/channels. Further improvements and provisioning could enable them to additionally detect faulty nodes.

In the Nested Signaling Reservation Protocol (NSRP), we have considered networks survivable. The failure of any node or link in the control network may bring the entire network down. Furthermore, the effect of failures in data channels can halt ongoing data transfers and decrease the efficiency. We have assumed that those types of failures have been foreseen and prevented. One solution to this problem would be to provide a dual

control network and also alternate paths for virtual connections through the network. This requires adding more resources, which ultimately increases network costs. Another related work can bring the probability of failures into the analysis. However, the analysis will become much more complicated.

This work has assumed that wavelength conversions are possible and prompt. Investigating the impact of deploying real wavelength conversions with current features over various topologies for the proposed restoration and fault localization protocols is another problem to be solved.

References

[Ago93] N. Agoulmine, D. S. Neuman, and M. Oliveira, "Distribution of Management Over Multi-Domains Network Management Systems," *Proceedings of IEEE Global Telecommunications Conference (GLOMCOM'93)*, pp. 1217-1221, Nov. 1993.

[Ala04] W. Alanqar and A. Jukan, "Extending End-to-end Optical Service Provisioning and Restoration in Carrier Networks: Opportunities, Issues, and Challenges," *IEEE Communications Magazine*, Vol. 42, No. 1, pp. 52-60, Jan. 2004.

[Ali01] M. Ali, D. Elie-Dit-Cosaque, and L. Tancevski, "Enhancements to Multi-Protocol Lambda Switching (MPLS) to Accommodate Transmission Impairments," *Proceedings of IEEE Global Telecommunications Conference (GLOBCOM'01)*, pp. 70-75, Nov. 2001.

[Amr97] A. Amrani and J. Roldan "Optical Monitoring System for Scalable All-optical networks," *Proceedings of IEEE Lasers and Electro-Optics Society (LEOS'97)*, pp. 270-271, Nov. 1997.

[App04] R. Appelman and Z. Zalevsky, "All-optical Switching Technologies for Protection Applications," *IEEE Optical Communications Magazine*, Vol. 42, No. 11, pp. 35-40, Nov. 2004.

[Arc03] D. Arci, D. Petecchi, G. Maier, A. Pattavina, and M. Tornatore, "Availability Models for Protection Techniques in WDM Networks," *Proceedings of IEEE Design of Reliable Communication Networks (DRCN'03)*, pp. 158-169, Oct. 2003.

[Arm92] J. Armitage, O. Corchat, and J. Y. Le Boudec, "Design of a Survivable WDM Photonic Network," *Proceedings of IEEE Global Telecommunications Conference (GLOBCOM'92)*, pp. 633-638, Dec. 1992.

[Bal02] I. Baldine, G. N. Rouskas, H. G. Perros, and D. Stevenson, "Jumpstart: A Just-in-Time Signaling Architecture for WDM Burst-Switched Networks," *IEEE Communications Magazine*, Vol. 40, No. 2, pp. 82-96, Feb. 2002.

[Bar97] R. B. Barry and S. Subramaniam, "The MUX-SUM Wavelength Assignment Algorithm for WDM Ring Networks," *Proceedings of IEEE Optical Fiber Communication Conference (OFC'97)*, pp. 121-122, Feb. 1997.

[Bis96] M. Bischoff and M. N. Huber, "Survivability Concept for All-optical Transport Networks," *Proceedings of International Conference on Communications Technology (ICCT'96)*, pp. 884-887, May 1996.

[Bon00] P. Bonefant and A. Rodriguez, "Optical Data Networking," *IEEE Communications Magazine*, Vol. 38, No. 3, pp. 63-70, Mar. 2000.

[Bou02] E. Bouillet, J. F. Labourdette, G. Ellina, and R. Ramamurthy, "Stochastic Approaches to Compute Shared Mesh Restored Lightpaths in Optical Network Architectures," *Proceedings of IEEE INFOCOM'* 02, pp. 801-807, June 2002.

[Cae98] B. V. Caenegem, W. V. Pary, F. D. Turck, and P. M. Demeester, "Dimensioning of Survivable WDM Networks," *IEEE Journal on Selected Areas of Communications*, Vol. 16, No. 7, pp. 1146-1157, Sep. 1998.

[Chi003] A. L. Chiu and J. Strand, "An Agile Optical Layer Restoration Method for Router Failures," *IEEE Network,* Vol. 17, No. 2, pp. 38-42, Mar. 2003.

[Chn93] R. S. K. Chng, M. S. Sinclair, C. P. Botham and M. Mahony, "Multi-layer Real-time Restoration for Survivable Networks," *IEE Colloquium on Safety Critical Distributed Systems (CSCDS'93)*, Vol. 5, No. 189, pp. 1-5, 1993.

[Coa91] B. A. Coan, "Using Distributed Topology Update and Prep-planned Configurations to Achieve Trunk Network Survivability," *IEEE Transactions on Reliability*, Vol. 40, No. 4, pp. 404-416, 1991.

[Cro00] O. Crochat, J. Y. Le Boudec, and O. Gerstel, "Protection Interoperability for WDM Optical Networks," *IEEE Transactions on Networking*, Vol. 8, No. 3, pp. 384-395, 2000.

[Cro98] O. Crochat and J-Y. Le Boudec, "Design Protection for WDM Optical Networks," *IEEE Journal on Selected Areas of Communications*, Vol. 16, No. 7, pp. 1158-1165, Sept. 1998.

[Dat01] S. Datta, S. Sengupta, and S. Biswa, "Efficient channel Reservation for Backup Paths in Optical Networks," *Proceedings of IEEE Global Telecommunications Conference (GLOBCOM'01)*, pp. Nov. 2001.

[Del70] O. E. Delange, "Wideband Optical Communication systems part II-Frequency Division Multiplexing," *IEEE Proceeding*, Vol. 58, No. 10, pp. 1683- 1690, Oct. 1970.

[Dem99] P. Demeester, M. Gryseels, and A. Autenrith, "Resilience in Multi-layer Networks," *IEEE Communications Magazine*, Vol. 37, No. 8, pp. 70-76, Aug. 1999.

[Dip05] G. Dipnarayan, H. C. Doan, O. S. Yang, and K. J. Seng, "A Centralized Server Based on Signaling and Resource Reservation Protocol Design for Fast Protection and Restoration in Optical Ring Networks," *Proceedings of IEEE International Performance, Computing, and Communications Conference (IPCCC'05),* pp. 631-632, Apr. 2005.

[Dob02] P. D. Dobbelaere and K. Falta, "Digital MEMS for Optical Switching," *IEEE Communications Magazine*, Vol. 40, No. 3, pp. 88-95, Mar. 2002.

[Dos99] B. D. Doshi, "Optical Network Design and Restoration," *Bell Labs Technical Journal*, Vol. 4, No. 1, pp. 58-84, Mar. 1999.

[Don05] W. Dongmei, L. Guangzhi, and A. Chin, "On Wavelength Management for Restoration in WDM Ring Networks," *Proceedings of Optical Fiber Communication Conference/ National Fiber Optic Engineers Conference (OFC/NFOEC'05)*, Vol. 3, pp. 6-11, Mar. 2005.

[Dov01a] R. Doverspike and J. Yates, "Challenges for MPLS IN Optical Network Restoration," *IEEE Communications Magazine*, pp. 89-96, Feb. 2001.

[Dov01b] R. Doverspike, G. Sahin, J. L. Strand, and R. W. Tkach, "Fast Restoration in a Mesh Network of Optical Cross-connects," *Proceedings of Optical Fiber Communication Conference (*OFC'99), Vol. 1, pp. 170-172, Mar. 1999.

[Fum00] A. Fumagalli and L. Valcarenghi, "IP restoration vs. WDM Protection: Is There an Optimal Choice?" *IEEE Network*, Vol. 14, No. 6, pp. 34-41, Dec. 2000.

[Ger97] O. Gerstel and R. Ramaswami, "Multiwavelength Optical Network Architectures and Protection Schemes," *Photonic Networks, Advances in Optical Communications,* pp. 42-51, 1997.

[Ger98] O. Gerstel, R. Ramaswami, and G. H. Sasaki, "Fault Tolerant Multi-wavelength Optical Rings With Limited Wavelength Conversion," *IEEE Journal on Selected Areas of Communications*, Vol. 6, No. 7, pp. 1166-1178, Sep. 1998.

[Ger00] O. Gerstel and R. Ramaswami, "Optical Layer Survivability: a Service perspective," *IEEE Communications Magazine*, Vol. 38, No. 3, pp. 104-113, Mar. 2000.

[Gha00] N. Ghani, S. Dixit, and T. Wang, "On IP-over-WDM integration," *IEEE Communications Magazine*, Vol. 38, No. 3, pp. 72-84, Mar. 2000.

[Gio05] A. Giorgetti, L. Valcarenghi, and P. Castoldi, "Single-Layer versus Multilayer Preplanned Lightpath Restoration," *IEEE/ACM Journal of Lightwave Technology*, Vol. 23, No. 10, pp. 3206-3212, Oct. 2005.

[Gof02] D. R. Goff, "A Brief History of Fiber Optic Technology," *Focal Press*, Woburn, Massachusetts, 2002.

[Gro87] W. D. Grover, "The Self-healing Network: A Fast Distributed Restoration Technique for Networks Using Digital Cross-connect Machines," *Proceedings of IEEE Global Telecommunication Conference (GLOBCOM'87)*, pp. 1090-1095, 1987.

[Gro02] W. D. Grover and J. Doucette, "Advances in Optical Network Design with *P*-cycles: Joint Optimization and Pre-selection of Candidate *P*-cycles," *Proceedings of IEEE/LEOS Summer Topicals*, pp. WA49-WA50, 2002.

[Hai03] S. Hairong, J. J. Han, and H. Levendel, "Availability Requirement for a Fault-Management Server in High-availability Communication Systems," *IEEE Transactions on Reliability*, Vol. 52, No. 2, pp. 238-244, June 2003.

[Hai04] A. G. Hailemariam, G. Ellinas, and T. Stern, "Localized Failure Restoration in Mesh Optical Networks," *Proceedings of IEEE Optical Fiber Communication Conference (OFC'04)*, pp. 23-27, Feb. 2004.

[Hec99] J. Hecht, "City of Light: The Story of Fiber Optics," *Oxford University Press*, New York, 1999.

[Ho-02a] P. -H. Ho, J. Topolcai, and H. T. Mouftah, "On Achieving Survivable Routing for Shared Protection in Survivable Next-Generation Internet," *Optical Networks Magazine*, Vol. 4, No. 6, pp. 34-46, 2002.

[Ho-02b] P. -H. Ho and H. T. Mouftah, "A Framework of Service Guaranteed Shared Protection for Optical Networks," *IEEE Communications Magazine*, pp. 97-103, Feb. 2002.

[Ho-03] P. -H. Ho and H. T. Mouftah, "Allocation of Protection Domains in Dynamic WDM Mesh Networks," *Proceedings of IEEE International Conference on Communications (ICC'03)*, pp. 29-36, 2003.

[Hu-03] J. Q. Hu, "Diverse routing in optical mesh networks," *IEEE Transactions on Communications*, Vol. 51, No. 3, pp. 489-494, Mar. 2003.

[Ira96] R. R. Iraschko, M. H. McGregor, and W. D. Grover, "Optimal Capacity Placement for Path Restoration in Mesh Survivable Networks," *Proceedings of IEEE International Conference on Communications (ICC'96)*, pp. 1568-1574, June 1996.

[Kar03] K. Kar, M. Kodialam, and T. V. Lakshman, "Routing Restorable Bandwidth Guaranteed Connection Using Maximum 2-Route Flows," *IEEE Transactions on Networking*, Vol. 11, No. 5, pp. 772-780, Oct. 2003.

[Kar04] M. Karol, P. Krishnan and J. Jenny Li, "Rapid Fault Detection and Recovery for IP Telephony," *Proceedings of IEEE International Conference on Communications (ICC'04)*, pp. 1484-1489, June 2004.

[Kob96a] Y. Kobayashi, Y. Tada, S. Matsuoka, N. Hirayama, and K. Hagimoto, "Supervisory Systems for All-optical Network Transmission Systems," *Proceedings of IEEE Global Telecommunications Conference (GLOBCOM'96)*, pp. 933-937, Nov. 1996.

[Kod01] M. Kodialam and T. V. Lakeshman, "Dynamic Routing of Locally Restorable Bandwidth Guaranteed Tunnels Using Aggregated Link Usage Information," *Proceedings of IEEE INFOCOM'01*, pp. 376-385, 2001.

[Kom90] H. Komine, T. Chujo, T. Ogura, K. Miyazaki and T. Soejima, "A Distributed Restoration Algorithm for Multiple-link and Node Failures of Transport Networks," *Proceedings of IEEE Global Telecommunications Conference (GLOBCOM'90)*, pp. 459-463, Nov. 1990.

[Laz92] A. A. Lazar, W. Wang, and R. H. Deng, "Models and Algorithms for Network Failure Detection and Identification: A Review," *Proceedings of Singapore International Conference on Communications/International Symposium on Information Theory and Its Applications SICC/ITA*, pp. 999-1003, Nov. 1992.

[Lee01] K. Lee and K. Siu, "An Algorithmic Framework for Protection Switching WDM Networks," *Proceedings of Optical Fiber Communication Conference/ National Fiber Optic Engineers Conference (OFC/NFOEC'01)*, pp. 402-410, July 2001.

[Lee05] C. M. Lee, CX. K. Chan, and L. K. Chen, "A Single-Fiber Bidirectional Self-Healing WDM Multi-hop Mesh Network With All-Optical Cyclic Deflection Routing for Link Restoration," *IEEE Photonics Technology Letters*, Vol. 1, No. 11, pp. 2463 – 2465, Nov. 2005.

[Li-97] C. -S. Li and R. Ramaswami, "Automatic Fault Detection, Isolation, and Recovery in Transparent All-optical Networks," *IEEE Journal of Light wave Technology,* Vol. 15, No. 10, pp. 1784-1793, Oct. 1997.

[Li-05] L. Li, M. M. Buddhikot, C. Chekuri, and K. Guo, "Routing Bandwidth Guaranteed Paths with Local Restoration in Label Switched Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 23, No. 2, pp. 437-449, Feb. 2005.

[Lim05] K. W. Lim, E. S. Son, K. H. Han, and Y. C. Chung, "Fault Localization in WDM Passive Optical Network by Reusing Downstream Light Sources," *IEEE Photonics Technology Letters*, Vol. 99, pp. 1-3, June 2005.

[Lu-04] S. Lu, H. Wu, S. Yan, and Y.Tasi, "Testing and Diagnosis Techniques for LUT-based FPGA's," *Proceedings of IEEE 13th Asian Test Symposium (ATS'04)*, pp. 414-419, 2004.

[Mac00] J. C. MacEchen and R. L. Chesser, "Vulnerabilities in the Open Shortest Path First Interior Gateway Protocol," *Proceedings of IEEE Military Communications Conference (MILCOM'00)*, pp. 1224-1228, Oct. 2000.

[Man04] J. Manchester, D. Saha, and S. K. Tripathi, "Protection, Restoration, and Recovery," *IEEE Network*, Vol. 18, No. 2, pp. 3-4, 2004.

[Mar93] M. A. Marsan, "Topologies for Wavelength Routing All-Optical Networks," *IEEE/ACM Transactions on Networking*, Vol. 1, No. 5, pp. 534-546, Oct. 1993.

[Mas00] C. Mas and P. Thiran, "An Efficient Algorithm for Locating Soft and Hard Failures in WDM Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 10, pp. 1900-1911, Oct. 2000.

[Mas03] C. Mas and I. Tomkos, "Failure Detection for Secure Optical Networks," *Proceedings of International Conference on Transparent Optical Networks (ICTON'03)*, pp. 70-75, May 2003.

[Med97] M. Medard, S. R. Chinn, and P. Saengudomlert, "Security Issues in All-optical Networks," *IEEE Network*, pp. 42-48, May 1997.

[Med98] M. Medard, S. R. Chinn, and P. Saengudomlert "Attack Detection in All-optical Networks," *Proceedings of Optical Fiber Communication Conference (OFC'98)*, pp. 272-273, Feb. 1998.

[Med99] M. Medard, S. G. Finn, and R. A. Barry, "WDM Loop-back Recovery in Mesh Networks," *Proceedings of IEEE INFOCOM'99*, Vol. 2, pp. 752-759, 1999.

[Med02] M. Medard, S. G. Finn, R. A. Barry, W. He, and S. S. Lumetta, "Generalized Loop-Back Recovery in Optical Mesh Networks," *IEEE/ACM Transactions on Networking*," Vol. 10, No. 1, pp.153-164, Feb. 2002.

[Mei00] Y. Mei and C. Qiao, "Distributed Control Schemes for Dynamic Lightpath Establishment in WDM Optical Networks," *Proceedings of Optical Network Workshop (ONW'00)*, pp. 28-40, Feb. 2000.

[Muk96] B. Mukherjee, B. Banerjee, D. Ramamurty, and S. Mukherjee, "Some Principals for Designing a Wide-Area WDM Network," *IEEE/ACM Transactions on Networking*, Vol. 4, No. 5, pp. 684-695, Oct. 1996.

[Muk97] B. Mukherjee, *Optical Communication Networks,* London: McGraw-Hill, 1997.

[MUK00] B. Mukherjee, "WDM Optical Communication Networks: Progress and Challenges," *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 10, pp. 1810-1824, Oct. 2000.

[Nel03] S. Nelakuditi, S. Lee, Y. Yu, and A. Zhang, "Failure Insensitive Routing for Ensuring Service Availability," *Proceedings International Workshop on Quality of Service (IWQoS'03)*, pp. 287-304, June 2003.

[Oka03] T. Okamawari, S. Inglima and A. McGuire, "Design of Reliable Communication Networks," *Proceedings of Design of Reliable Communication Networks (DRCN'03)*, pp. 122-130, Oct. 2003.

[Pat01] J. K. Patel, S. U. Kim, and S. Subramaniam, "A Framework for Managing Faults and Attacks in All-optical Transparent Networks," *Proceedings of Defense Advanced Research Projects Agency/Information Survivability Conference and Exposition (DARPA/DISCEX'01)*, pp. 137-145, June 2001.

[Qia96] C. Qiao and Y. Mei, "Wavelength Reservation under Distributed Control," *Digest of IEEE/LEOS Summer Topic Meeting on Broadband Optical Networks: embedding technologies and applications*, pp. 45-46, Aug. 1996.

[Qia99] C. Qiao and M. Yoo, "Optical Burst Switching (OBS) a New Paradigm for an Optical Internet," *Journal on High Speed Networks*, Vol. 8, No. 1, pp. 69-84, 1999.

[Ram94] R. Ramaswami and K. N. Sivarajan, "Optimal Routing and Wavelength Assignment in All-Optical Networks," *Proceedings of IEEE INFOCOM'94*, Vol. 2, pp. 970-979, June 1994.

[Ram98] B. Ramamurthy and B. Mukherjee, "Wavelength Conversion in WDM Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 7, pp. 1061-1073, Sep. 1998.

[Ram99a] B. Ramamurthy, D. Datta, and H. Feng, "Transparent vs. Opaque vs. Translucent Wavelength-routed Optical Networks," *Proceedings of Optical Fiber Communication Conference (OFC' 99)*, pp. 59-61, Feb. 1999.

[Ram99b] B. Ramamurthy and B. mukherjee, "Survivable WDM Mesh, part I-Restoration Networks," *Proceedings of IEEE INFOCOM'99*, pp. 744-751, 1999.

[Ram99c] B. Ramamurthy and B. Mukherjee, "Survivable WDM Mesh, part II-Restoration Networks," *Proceedings of IEEE International Conference on Communications (ICC'99)*, pp. 2023-2030, 1999.

[Ram01] B. Ramamurthy, S. Yaragoral, and X. Yang, "Translucent Optical WDM Networks for the Next-Generation Networks," *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM'01),* Vol. 1, pp. 60-64, Nov. 2001.

[Ram02] B. Ramamurthy and A. Ramakrishnan, "Design of Virtual Private Networks (VPNs) over Optical Wavelength Division Multiplexed (WDM) Networks," *SPIE Optical Networks Magazine*, Vol. 3, No. 1, pp. 59-63, Jan./Feb. 2002.

[Rama97] R. Ramaswami and A. Segall, "Distributed Network Control for Optical Networks," *IEEE/ACM Transactions on Networking*, Vol. 5, No. 6, pp. 936-943, Dec. 1997.

[Rama98] R. Ramaswami and K. N. Sivarajan, "Optical Networks: a Practical Perspective," *Morgan Kaufman*, 1998.

[Rama99] R. Ramaswami and A. Segall, "Distributed Network Control for Wavelength Routed Optical Networks," *Proceedings of IEEE International Conference on Communications (ICC'99)*, Vol. 3, pp. 2023-2030, 1999.

[Ramp01] L. Rampros, G. Chatzilias, and A. Manzalini, "Design and Experiments of an Automatic Switched Optical Network (ASON)," *Proceedings of European Conference on Optical Communication (ECOC'01)*, pp. 256-257, 2001.

[Sah00a] D. Saha, "Forward Reservation Protocol with Intermediate Unlock (FRP-IU) for Dynamic Path Establishment in All-Optical Networks (AONs)," *Proceedings of SPIE Terra-bit Optical Networking (TON'00)*, pp. 4213-4225, Nov. 2000.

[Sah00b] D. Saha, "An Efficient Wavelength Reservation Protocol for Lightpath Establishment in All-Optical Networks (AONs)," *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM'00)*, Vol. 2, pp. 1264-1268, Nov. 2000.

[Sar02] C. V. Saradhi and C. Silva Ram Murthy, "Dynamic Establishment of Segmented Protection Paths in Single and Multi-fiber WDM Mesh Networks," *Proceedings of SPIE Conference on Optical Communications (COPTICOMM'02)*, pp. 77-98, Aug. 2002.

[Sch02] D. A. Schupke, "Fast and Efficient WDM Network Protection Using p-Cycles," *IEEE Network*, pp. 47-48, 2002.

[Sen00] D. Sengupta, D. Saha, and D. Datta, "Performance Analysis of Access Protocols for Multi-level Optical WDM MANs Using Wavelength Routing Switch," *Proceedings of IEEE International Conference on Communications (ICC'00)*, Vol. 3, pp. 1633-1637, 2000.

[Shy04] S-K Lu, H-C Wu, and S-J Yan, "Tsai Testing and Diagnosis Techniques for LUT-Based FPGA's," *Proceedings of Test Symposium (TS'04)*, pp. 414-419, Nov. 2004.

[Sic03a] A. V. Sichani and H. T. Mouftah, "A new signaling nested protocol for bandwidth allocation in WDM all-optical network" *Proceedings of International Symposium on Telecommunications (IST'03)*, pp. 366-370, Oct. 2003.

[Sic03b] A. V. Sichani and H. T. Mouftah, "A Novel Distributed Progressive Reservation Protocol for WDM All-Optical Networks," *Proceedings of IEEE*

International Conference on Communications (ICC'03), Vol. 26, No.1 pp. 1463-1467, May. 2003.

[Sic03c] A. V. Sichani and H. T. Mouftah, "A Nested Path Reservation Protocol for Multiplexed All-Optical Interconnection Networks," *Proceedings of Optical Networks Design and Management (ONDM'03),* pp. 401-411, Feb. 2003.

[Sic04a] A. V. Sichani and H. T. Mouftah, "Rolling Back Signaling Protocol- A Novel Fault Localization Protocol for WDM Mesh Networks," *CIC China Communications Magazine*, Vol. 1, No. 1, pp. 101-105, Dec. 2004.

[Sic04b] A. V. Sichani and H. T. Mouftah, "A Novel Broadcasting Fault Detection Protocol in WDM Networks," *Proceedings of Queens Biennial Symposium on Communications (QBSC'04),* pp. 222-224, May 2004.

[Sic05a] A. V. Sichani and H. T. Mouftah, "Signaling Nested Reservation Protocol a Novel Reservation Protocol for Creating All-optical Network," *IEEE Journal on Selected Areas in Communications*, Vol. 23, No. 2, pp. 277-282, Feb. 2005.

[Sic05b] A. V. Sichani, S. Gazor, and H. T. Mouftah, "Minimum Weighted-path Restoration Protocol for Survivable All-optical Networks," *Mediterranean Journal on Computer and Networks,* to appear.

[Sic05c] A. V. Sichani and H. T. Mouftah, "Performance Evaluation of Dynamic Restoration Techniques for Survivable WDM Networks," submitted to *IEEE International Symposium on Computers and Communications (ISCC'06).*

[Sic05d] A. V. Sichani and H. T. Mouftah, "Maximum Mutual Links-*k*th Shortest Path Protocol, A Survivable Routing Scheme for WDM Mesh Networks," *Proceedings of International Symposium on Telecommunications 2005 (IST'05),* pp. 109-116, Sept. 2005.

[Sic05e] A. V. Sichani and H. T. Mouftah, "Open Link Restoration Protocol for Survivable Optical Networks" *IEEE/OSA Journal of Lightwave Technology,* submitted.

[Sic05f] A. V. Sichani and H. T. Mouftah, "Limited-perimeter Vector Matching Fault Localization protocol for All-optical Communication Networks," in patenting

process/to be submitted to *Elsevier Optical Switching and Networking Journal.*

[Sin05] R. K. Sinha, Y. Fang, and R. Doverspike, "Performance Study of Mesh Restoration," *Proceedings of Optical Fiber Communication Conference/ National Fiber Optic Engineers Conference (OFC/NFOEC'05)*, Vol. 2, pp. 3-10, Mar. 2005.

[Som99] A. K. Somani and L. Li, "Dynamic Wavelength Routing Using Congestion and Neighborhood Information," *IEEE/ACM Transactions on Networking*, Vol. 7, No. 5, pp. 779-786, Oct. 1999.

[Spa00] J. Spath, "Resource Allocation for Dynamic Routing in WDM Networks," *Computer Networks*, Vol. 32, pp. 519-538, May 2000.

[Sta02] S. Stanic, S. Subramaniam, H. Choi, G. Sahin, and H-A. Choi, "On Monitoring Transparent Optical Networks," *Proceedings of International Conference on Parallel Processing Workshop 2002 (ICPPW'02)*, pp. 218-223, Aug. 2002.

[Sto02] G. Stoica and D. Sengupta, "On Dynamic Wavelength Assignment for Wavelength-Routed All-Optical Networks," *SPIE Optical Networks Magazine*, Vol. 3, No. 1, pp. 68-76, Jan. 2002.

[Str01] J. Strand, A. Chiu, and R. Tkach, "Issues for Routing in Optical Layer," *IEEE Communications Magazine*, Vol. 39, No. 2, pp. 81-87, Feb. 2001.

[Sur84] J. W. Surrballe and M. M. B. Pascoal, "A Quick Method for Finding Shortest Pairs of Disjoint Paths," *IEEE Network*, Vol. 14, No. 2, pp. 325-336, 1984.

[Tap01] J. Tapolcai, P. Laborczi, P. H. Ho, T. Cinkler, A. Recski, and H. T. Mouftah, "Algorithm for Asymmetrically Weighted Pair of Disjointed Path in Survivable Networks," *Proceedings of Design of Reliable Communication Networks (DRCN'01)*, pp. 239-249, Oct. 2001.

[Tso04] S. K. Tso, J. K. Lin, H. K. Ho, C. M. Mak, K. M. Yung, and Y. K. Ho, "Data Mining for Detection of Sensitive Buses and Influential Buses in a Power System Subjected to Disturbances," *IEEE Transactions on Power Systems,* Vol. 19, No. 1, pp. 563-568, Feb. 2004.

[Wan94] C. J. Wang, H. Y. Zhou, and C. H. Chow, "Automatic Network Restoration Using a Two-level Associative Memory," *Proceedings of IEEE International Conference on Neural Networks (ICNN'94),* pp. 3565-3570, June 1994.

[Wol95] W. V. Wollman and Y. Barsoum, "Overview of Open Shortest Path First, Version 2 Routing in the Tactical Environment," *Proceedings of IEEE Military Communications Conference (MILCOM'95)*, pp. 925-930, Nov. 1995.

[Wu-95] T. H. Wu, "Emerging Technologies for Fiber Network Survivability," *IEEE Communications Magazine*, Vol. 33, No. 6, pp. 58-74. Feb. 1995.

[Xin02] C. Xin, Y. Ye, S. Dixit, and C. Qiao, "A Joint Lightpath Routing Approach in Survivable Optical Networks," *Optical Networks Magazine*, pp. 23-32, May 2002.

[Xu-02] D. Xu, C. Qiao, and Y. Xiong, "An Ultra-fast Shared Path Protection Scheme-Distributed Partial Information Management, Part II," *Proceedings of IEEE International Conference on Network Protocols (ICNP'02)*, pp. 344-353, Nov. 2002.

[Yan93] C. Yang and S. Hasegava, "A Failure Immunization Technology for Network Service Survivability," *Proceedings of IEEE Global Telecommunications Conference (GLOBCOM'93),* pp. 1549-1554, 1993.

[Yeo01] T. Yeow, K. L. Eddie Law, and A. Goldenberg, "MEMS Optical Switches," *IEEE Communications Magazine*, Vol. 39, No. 11, pp. 158-163, Nov. 2001.

[Yon05] T. Yongning, E. S. Al-Shaer, and R. Boutaba, "Active Integrated Fault Localization in Communication Networks," *Proceedings of IFIP/IEEE International Symposium on Integrated Network Management (ISINM'05),* pp. 543-556, May 2005.

[Yua99] X. Yuan, R. Melhem, R. Gupta, Y. Mei, and C. Qiao, "Distributed Control Protocols for Wavelength Reservation and their Performance Evaluation," *Photonic Network and Communication*, Vol. 1, No. 3, pp. 207-218, 1999.

[Zan00a] H. Zang, J. P. Jue, and B. Mukherjee, "A Review of Routing and Wavelength Assignment Approaches for Wavelength-Routed Optical WDM Networks," *SPIE Optical Networks Magazine*, Vol. 1, No. 3, pp. 47-60, Jan. 2000.

[Zan00b] H. Zang, L. Sahasrabuddhe, J. P. Jue, S. Ramamurthy, and B. Mukherjee, "Connection Management for Wavelength-Routed WDM Networks," *Proceedings of IEEE Global Telecommunications Conference (GLOBCOM'00)*, pp. 1428-1432, May 2000.

[Zen04] H. Zeng, C. Huang, and A. Vukovic, "Monitoring Cycles for Fault Detection in Meshed All-optical Networks," *Proceedings of International Conference on Parallel Processing/International Workshop in Optical Networks and Management (ICPP/ONCM'04)*, pp. 434-439, Aug. 2004.

[Zha98] X. Zhang and C. Qiao, "Wavelength Assignment for Dynamic Traffic in Multi-fiber WDM Networks," *Proceedings of IEEE International Conference on Computer Communications and Networks (ICCCN'98)*, pp. 479-485, Oct. 1998.

[Zha01] Z. Zhang, J. Fu, D. Guo, and L. Zhang, "Lightpath routing for intelligent optical networks," *IEEE Network*, Vol. 15, No. 4, pp. 28-35, Aug. 2001.

[Zha04] Z. Zhang, W. D. Zhong, and B. Mukherjee, "A Heuristic Method for Design of Survivable WDM Networks with *p*-Cycles," *IEEE Communications Letters*, Vol. 8, No. 7, pp. 467-469, July 2004.

[Zhe02a] J. Zheng and H. T. Mouftah, "Destination-Initiating Path Restoration Protocol for Wavelength-routed WDM Networks," *IEE Proceedings on Communications*, Vol. 149, No. 1, pp. 18-22, Feb. 2002.

[Zhe02b] J. Zheng and H. T. Mouftah, "Distributed Lightpath Control Based on Destination Routing for Wavelength-Routed WDM Networks," *SPIE Optical Networks Magazine*, Vol. 3, No. 4, pp. 38-46, July 2002.

[Zho00] D.Zhou and S. Subramaniam, "Survivability in Optical Networks," *IEEE Network*, Vol. 14, No. 6, pp. 16-23, Nov. 2000.

Appendix A

Simulation Program

We have used MATLAB 6.5 for simulating the performance of the considered restoration protocols. Since this simulation has involved both real time and discrete-events, it has been challenging and lengthy. Therefore, the main program and the linked functions have been written in modules to be easily debugged. The simulation model consists of a main program and several key functions which establish the network architecture and circumstances. These functions create the network topology, generate calls, produce failures, and finally evaluate the restoration protocols. In the following, we describe some key functions.

The first function shapes a matrix, namely switching matrix which develop the network structure. Switching ports (input and output ports coupled to each wavelength) are simulated such that each row corresponds to a certain node in which a particular wavelength is supported (locked/unlocked). Accordingly, the number of rows in the switching matrix is defined by multiplying the number of nodes by the number of wavelengths. A row of the switching matrix which is "in use" is filled with the user (caller) ID, the blocked time, the released time (based on the caller data size), the status, and other related information respectively. A "1" in the status cell of each row indicates that the related port (wavelength) is locked by a request. Routinely, the row is cleared at the release time of the call.

The second function generates request calls by creating a matrix, namely calling matrix. Each row of the calling matrix is filled by the information of a call, including the source, the destination, the data length, the arrival time, and ID.

s-d pairs in this matrix are randomly selected in the way that no pair obtains identical nodes. Time intervals between calls and the data sizes in the calling matrix follow exponential distribution functions with parameters $\lambda_{request}$ and λ_{Data} respectively. The

average of the arrival rate ($\frac{1}{\lambda_{request}}$) is set to be 1 sec and the average of the data rate ($\frac{1}{\lambda_{data}}$) is 10 sec. A constant value of 0.125 sec is added to the data column of the calling matrix to avoid having a negligible data size. In our simulation, the calling matrix has been created by 200 calls. Therefore, it comprises 200 rows.

To investigate the impact of failures, a failing matrix similar to the calling matrix has been created. Failure locations are random variables with Uniform distribution. Failure occurrence times are randomly picked within the first and last arrival time range. Failure durations are also random variables and could last up to 10 sec. For the purpose of our simulation, we have generated 10 failures.

We set the average fault localization time to be 0.05 sec. The configuration time of OXCs is considered fixed and equal to 0.125 sec for all calls. The average propagation delay is set to 0.025 sec on each link. The computational cycles are assumed to be

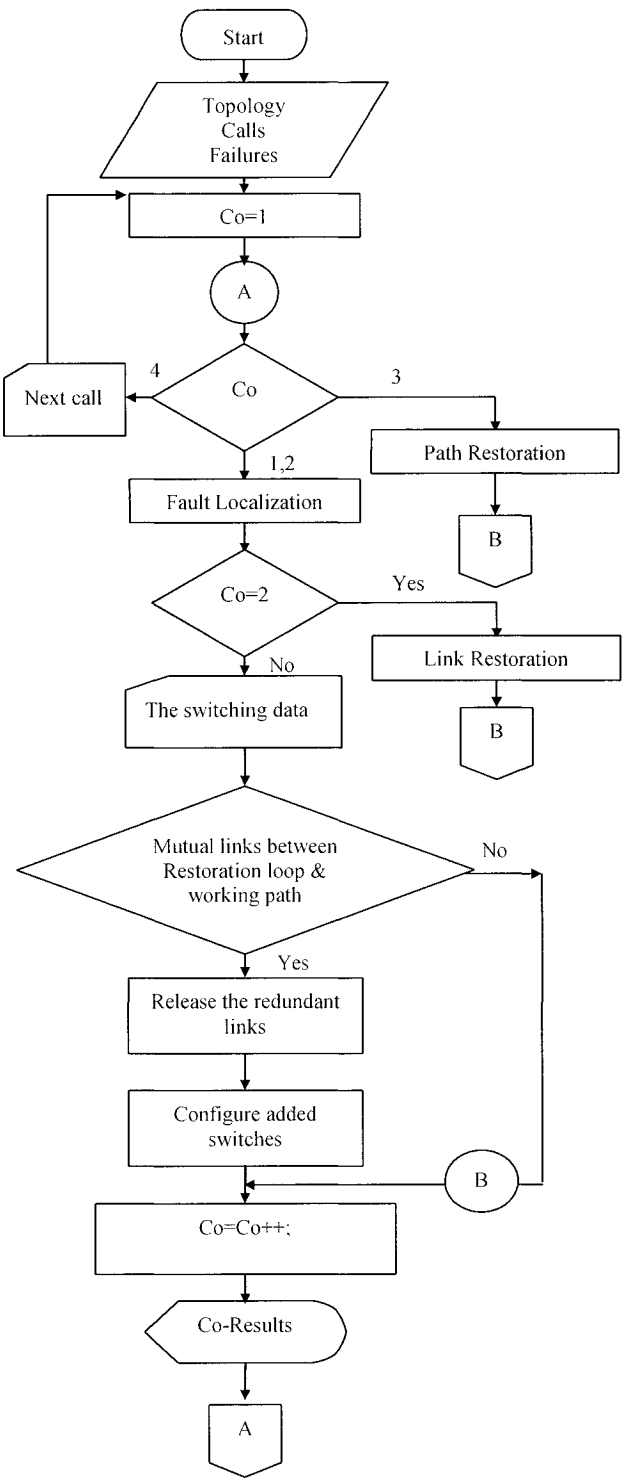

Figure A.1 Main flowchart of the compared restoration protocols.

0.00001 sec. We assume that each re-routing decision takes 0.01 sec. Without loss of generality, we assume that each node performs full wavelength conversion, i.e., a lightpath could comprise any free wavelength on its way. We assume a zero holding time. This means that if unsuccessful, a lightpath reservation is dropped immediately. (Research has proven that the holding time does not have a significant effect on throughput). We relax the initial retry-list, which restricts the set of wavelengths under consideration to the multiplexing degree.

Eventually after generating all the network matrices, the switching, calling and failing matrices have been sorted on the arrival/occurrence time columns to arrange and define the order of events. Note that the release time of the existing calls and the repaired time of the created failures have been included and sorted as complementary events. To include these events, copies of the calling and failing matrices, with an exchange on their arrival and release time columns, are added to the data and the entire data is sorted. Finally, the main program calls three performance evaluation functions, namely path, joint path, and link restoration functions. These functions examine the restoration ability of the applied protocols and illustrate the associated throughput via the related curves. Figure A.1 depicts the associated flowchart of the main program.

The effects of the following parameters on the performance have been observed:

- *Multiplexing degree*: The number of virtual channels (wavelengths) supported by each link changed between 2 and 16.
- *Network size*: The simulation is run over symmetrical mesh networks of size 3x3 to 9x9.

Appendix B

Confidence Interval Estimation

Let X be an unknown parameter to be estimated based on a random sample $X_1, X_2, \cdots, X_n$, from a distribution with that parameter X. We want an interval [L, U], where L and U are two statistics calculated from $X_1, X_2, \cdots, X_n$, such that the interval includes X with a pre-assigned probability, $1-\alpha$. Specifically, we want to have $P\{L \le X \le U\} \ge 1-\alpha$ regardless of the true value of X. Note that L and U are random variables, being functions of $X_1, X_2, \cdots, X_n$, so [L, U] is a random interval. It is called a $100(1-\alpha)\%$ Confidence Interval (CI), $1-\alpha$ is called confidence level, and L and U are confidence limits.

Consider a random sample $X_1, X_2, \cdots, X_n$ with $N(\mu, \delta^2)$ distribution, where δ^2 is assumed to be known and μ is the sample mean $\overline{X}$. From the properties of the normal distribution, it follows that there is 95% probability that $\overline{X}$ falls within a distance of $1.96\,\delta/\sqrt{N}$ from μ : $P\left[\mu - 1.96\,\delta/\sqrt{N} \le \overline{X} \le \mu + 1.96\,\delta/\sqrt{N}\right]$. $\overline{X}$ is within a distance of $1.96\,\delta/\sqrt{N}$ from μ if and only if μ is within a distance of $1.96\,\delta/\sqrt{N}$ from $\overline{X}$. Therefore, we have, $P\left[L = \overline{X} - 1.96\,\delta/\sqrt{N} \le \mu \le \overline{X} + 1.96\,\delta/\sqrt{N} = U\right] = 0.95$.

For instance if we consider a group of data for 10 independent runs per simulation experiment: $X_1, X_2, \cdots, X_n$, = 85, 89, 98, 83, 88, 77, 94, 92, 80, 89; then, we can calculate our parameters including $\overline{X}$ = 87.5, 6.4261, and $1.96\,\delta/\sqrt{N}$ = 3.9768. As a result, confidence limits would be [*L=83.5232, U=91.4768*]. In our simulation results, 90% confidence intervals are depicted by "*".

Resume

1400 Fisher Crescent
Kingston, ON
K7M 8T1
Canada
Phone (613) 384-8497
E-mail
sichani@site.uottawa.ca

Atousa Vali Sichani

Education

2002 - 2005 **University of Ottawa** Ottawa, Canada
Ph. D. student, Electrical and Computer Engineering

2000 - 2002 **Queen's University** Kingston, Canada
M. Sc. Electrical and Computer Engineering

1993 - 1994 **Sorbonne University** Paris, France
French language

1998 - 1993 **Isfahan University of Technology** Isfahan, Iran
B. Sc. Electrical and Computer Engineering

Academic Awards

[1] Ontario Graduate Scholarship 2005- 2006 (**OGS**)
[2] Ottawa Excellence Award 2005- 2006 (**OEA**)
[3] Ontario Graduate Scholarship 2004- 2005 (**OGS**)
[4] Ottawa Excellence Award 2004- 2005 (**OEA**)
[5] Ontario Graduate Scholarship 2003- 2004 (**OGS**)
[6] Ottawa Excellence Award 2003- 2004 (**OEA**)
[7] Queen's Graduate Award 2001- 2002 (**QGA**)
[8] Queen's Graduate Award 2000- 2001 (**QGA**)

Academic Work Experience

2002- 2005 **Research Assistant**	University Of Ottawa	Ottawa, Canada
2002- 2003 **Teaching Assistant**	University Of Ottawa	Ottawa, Canada
2000- 2002 **Research Assistant**	Queen's University	Kingston, Canada
2000- 2002 **Teaching Assistant**	Queen's University	Kingston, Canada
1993- 1993 **Teaching Assistant**	Isfahan University of Technology	Isfahan, Iran

Publications

[1] **A. V. Sichani** and H. T. Mouftah, "Signaling Nested Reservation Protocol a Novel Reservation Protocol for Creating All-optical Network," *IEEE Journal on Selected Areas in Communications,* Vol. 23, No. 2, pp. 277-282, Feb. 2005.

[2] **A. V. Sichani** and H. T. Mouftah, "Rolling Back Signaling Protocol- A Novel Fault Localization Protocol for WDM Mesh Networks," *CIC China Communications Magazine,"* Vol. 1, No. 1, pp. 101-105, Dec. 2004.

[3] **A. V. Sichani,** S. Gazor, and H. T. Mouftah, "Minimum Weighted-Path Restoration Protocol for Survivable All-optical Networks," accepted in *Mediterranean Journal on Computer and Networks*, 2006.

[4] **A. V. Sichani** and H. T. Mouftah, "A Novel Broadcasting Fault Detection Protocol in WDM Networks," *Proceedings of QBSC'04,* pp. 222-224, May 2004.

[5] **A. V. Sichani** and H. T. Mouftah, "A New Signaling Nested Reservation Protocol for Bandwidth Allocation in WDM All-Optical Networks" *Proceedings of IST'03,* pp. 366-370, Oct.

2003.

[6] **A. V. Sichani** and H. T. Mouftah, "A Novel Distributed Progressive Reservation Protocol for WDM All-Optical Networks," *Proceedings of IEEE ICC'03,* Vol. 26, No.1 pp. 1463-1467, May. 2003.

[7] **A. V. Sichani** and H. T. Mouftah, "A Nested Path Reservation Protocol for Multiplexed All-Optical Interconnection Networks," *Proceedings of ONDM'03*, pp. 401-411, Feb. 2003.

[8] **A. V. Sichani** and H. T. Mouftah, "Maximum Mutual Links-*k*th Shortest Path Protocol, A Survivable Routing Scheme for WDM Mesh Networks," *Proceedings of IST'05,* pp. 109-116, Sep. 2005.

[9] **A. V. Sichani** and H. T. Mouftah, "Link Versus Path Restoration Techniques for Survivable WDM Networks," accepted to *IEEE ISCC'06.*

[10] **A. V. Sichani** and H. T. Mouftah, "Open Link Restoration Protocol for Survivable Optical Networks" submitted to *IEEE/OSA Journal of Lightwave Technology.*

[11] **A. V. Sichani** and H. T. Mouftah, "Limited-perimeter Vector Matching Fault Localization protocol for All-optical Communication Networks," submitted to *IEE Transactions on communication.*

Posters and Presentation

[1] Nested Reservation Protocol (NRP), CITO Innovators Show, 2004.

[2] Forward-Backward Reservation, University of Ottawa, 2003.

[3] Network Optimization Techniques, Queen's University, 2002.

[4] Redundant Arrays of Inexpensive Disks (RAID), Queen's University, 2002.

[5] Storage Area Networks (SAN), Queen's University, 2001.

Patent

A. V. Sichani, H. T. Mouftah and University of Ottawa, "Limited-perimeter Vector Matching Fault Localization protocol," In Progress.

Services and Activities

IEEE JSAC Reviewer, 2005

IEEE ICC Reviewer, 2004 - 2005

IEEE LCN Reviewer, 2005

Social/Family-Coordinator of Iranian Student Association of Queens University (ISAQU), 2002- 2004.

Member of Iranian/ Afghan Women's Rights.