Resilience Provisioning Mechanisms for IP-Centric Optical Networks

By

Song Dong

SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Supervised by Dr. Chris Phillips

Department of Electronic Engineering

Queen Mary University of London

United Kingdom

November 2003

To Xumann and my forthcoming son

Abstract

To accommodate the increasing volume of Internet traffic brought about by the growing user community and new enterprise applications, optical networks are being deployed at unprecedented rates. This trend is changing the fundamental way in which optical transport networks are being designed and operated.

Traditional optical networks have been configured as static physical pipes to expand the transport capacity. In such a network, carrier-grade network resilience is provided by the protection and restoration facility in the SONET (Synchronous Optical Network) / SDH (Synchronous Digital Hierarchy) “transport” layer with the network topology mainly being ring-based. However, this solution has many limitations and cannot provide fast connection provisioning which is essential to the realisation of high-value broadband services. Thus there are efforts to develop a new generation of dynamic reconfigurable optical networks with an IP-centric control plane. This research investigates how optical resilience could be efficiently provisioned in such an infrastructure.

In addition, traditional optical networks have been deployed to meet the demands of predictable voice and private-line traffic. In this context, all the traffic is treated identically with full protection. Internet growth has diminished the predominance of voice traffic and private-line traffic relative to the much greater growth of data traffic, which presents a wider range of resilience requirements. For example, traffic generated by residential Internet access services typically requires a much lower grade of service than that of corporate financial transactions. Thus, a more cost-effective mechanism is needed to provide different resilience grades that better reflect the value of the traffic being carried.

The author addresses the above problems by proposing novel solutions for resilience provisioning mechanisms in the next generation optical network. Different resilience provisioning schemes are developed and comparatively assessed. Several schemes are proposed based on different application situations. These are then critically evaluated through analysis and simulation, and their strengths and weaknesses discussed. They are also compared against existing schemes.

In summary, the major achievements of this thesis are outlined as follows:

1. A flooding-based restoration scheme entitled Fast Restoration Scheme (FRS) is proposed. It uses flooding messages, instead of maintaining an up-to-date link state database, which is very expensive, to find the restoration path.

2. A novel restoration scheme named Adaptive Segment Path Restoration (ASPR) is proposed as an efficient resilience provisioning method for the real-time optical services. By dividing a long optical path into several segments and providing each with a separate backup path, it ensures a short restoration time at relatively low cost.

3. A Differentiated-Resilience Optical Services Model is proposed to provide multiple resilience levels to the optical services for a better network usage. The basic idea is to provide a range of resilience classes that better reflect the value of the traffic being carried. Simulation results demonstrate that it can provide a more cost-efficient resilient network than is possible with traditional protection mechanisms.

Acknowledgements

I would like to express my most sincere thanks to my supervisor, Dr. Chris Phillips, for his supervision and for all support and continuous encouragement he gave me in all moments of my study at Queen Mary University of London.

Thanks also go to my mentor, Dr. John Bigham, and my second supervisor, Professor Jonathan Pitts, for their technical help.

This research has been partially sponsored by Nortel Networks, Ltd. I am grateful to Nortel Networks for their sponsorship. I would also like to thank Mr. Robert Friskney, Mr. Nigel Baker and Dr. Xiang Lu at Harlow Labs, Nortel Networks for their help and technical support. I am very grateful to Robert for his extensive review of this thesis.

I would like to thank everybody in the Department of Electronic Engineering for creating and maintaining a nice and relaxing working environment. Thanks to Prof. Laurie Cuthbert, Mrs. Lynda Rolfe, Dr. Dong Chen, Andy Martin, Lin Du, Joy Feng, Gan, Petrit, Ho, Maria, and many others.

Special thanks go to Dr. Xingmin Meng and his family for their priceless help of many years.

I would also like to thank my friends and former colleagues at Huawei Technologies, namely Wang Shuzhong, Hao Xiuyan, Li Xin, Shuai Jinyu, Jiang Meikun, Wu Jun, Zhang Guorui and many others, for their help and encouragement.

I deeply thank my parents for their encouragement and support, and my wife Xumann for her endless love, encouragement, and for sharing with me the excitement and disappointment along the way.

Table of Content

List of Figures

List of Tables

Glossary

ADM	Add/Drop Multiplexer
APS	Automatic Protection Switching
AS	Autonomous System
ASPR	Adaptive Segment Path Restoration
ATM	Asynchronous Transport Mode
BLSR	Bidirectional Line Switched Ring
BSHR	Bidirectional Self-Healing Ring
BOD	Bandwidth on Demand
CR-LDP	Constraint-based Routing Label Distribution Protocol
CSPF	Constraint-based Shortest Path First
DCS	Digital Cross-connect System
DLE	Dynamic Lightpath Establishment
DROSM	Differentiated-Resilience Optical Services Model
DWDM	Dense Wavelength Division Multiplexing
DXC	Digital Cross-Connect
ECMF	Equal Cost Multi-path Forwarding
E/O	Electrical-to-Optical
FDM	Frequency Division Multiplexing
FRS	Fast Restoration Scheme
FSC	Fibre Switched Capable
GMPLS	Generalised Multi-Protocol Label Switching
GNS	Global Network State
IETF	Internet Engineering Task Force
ILP	Integer Linear Programming
IP	Internet Protocols
IP	Integer Programming
IS-IS	Intermediated System to Intermediate System
LDP	Label Distribution Protocol
LMP	Label Management Protocol
LP	Linear Programming
LMM	Label Mapping Message
LRM	Label Request Message
LSA	Link State Advertisement
LSC	Label Switched Capable

LSDB	Link State Database
LSP	Label Switching Path
LSR	Label Switching Router
MCF	Multi-Commodity Flow
MPLS	Multi-Protocol Label Switching
MPλS	Multi-Protocol Lambda Switching
ND	Node Degree
NDP	Neighbour Discovery Protocol
NMS	Network Management System
NP	Non-deterministic Polynomial
OADM	Optical Add/Drop Multiplexer
OAM&P	Operation, Administration, Maintenance and Provisioning
OC	Optical Carrier
OCh	Optical Channel
OMS	Optical Multiplex Section
OLT	Optical Line Terminal
OSPF	Open Shortest Path First
OSPF-TE	Open Shortest Path First Traffic Engineering Extension
OVPN	Optical Virtual Private Network
OXC	Optical Cross-Connect
O/E	Optical-to-Electrical
O/E/O	Optical-Electrical-Optical
PML	Path Merge Label Switching Router
PSC	Packet Switched Capable
PSL	Path Switch Label Switching Router
QoS	Quality of Service
RC	Resilience Class
RCL	Relative Capacity Loss
RFC	Request for Comment
RSVP	Resource reSerVation Protocol
RSVP-TE	Resource reservation Protocol Traffic Engineering Extension
RWA	Routing and Wavelength Assignment
SCA	Spare Capacity Allocation
SDH	Synchronous Digital Hierarchy
SHR	Self-Healing Ring
SLE	Static Lightpath Establishment
SONET	Synchronous Optical Network

SPE	Synchronous Payload Envelope
SPF	Shortest Path First
SRLG	Shared-Risk Link Group
STM	Synchronous Transport Module
STS	Synchronous Transport Signal
TDM	Time Division Multiplex (TDM) capable
TE	Traffic Engineering
TDM	Time Division Multiplexing
TLV	Type-Length-Value
TM	Terminal Multiplexer
UPSR	Unidirectional Path Switched Ring
USHR/L	Unidirectional Self-Healing Ring / Line
USHR/P	Unidirectional Self-Healing Ring / Path
VP	Virtual Path
VPI	Virtual Path Index (or Virtual Path Identifier)
VC	Virtual Circuit
VC	Virtual Container
VCI	Virtual Circuit Index (or Virtual Channel Identifier)
WDM	Wavelength Division Multiplexing
WIXC	Wavelength Interchanging Cross-Connect
WSXC	Wavelength Selective Cross-Connect

Chapter 1 Introduction

1.1 Problem Statement

To accommodate the rapid growth of the Internet, transport networks based on Wavelength Division Multiplexing (WDM) technology are increasingly being deployed in carrier networks. In such networks, a number of multiple data streams can be multiplexed into a single fibre, each operating at a few Gbit/s. Therefore, the aggregate throughput of this type of network is expected to be in the order of Tbit/s. Consequently, a single element (link or node) failure in the network could result in a large amount of data loss, which makes network resilience a key issue in the design of next generation optical networks.

At present, WDM is mostly deployed as point-to-point system and uses SONET (Synchronous Optical Networks) and SDH (Synchronous Digital Hierarchy) as the standard layer for interfacing to the higher layers of the protocol stack. SONET/SDH networks today are, for the most part, protected in the forms of rings. These rings are interconnected in order to provide overall network connectivity and resilience statically. This network infrastructure is well established and robust. However, it also has many limitations. Firstly, provisioning of new connections is usually achieved statically and manually, which takes months and may result in lost carrier opportunities. Secondly, more than half of the bandwidth needs to be reserved for protection against failures. Thirdly, upgrades and traffic growth are costly and difficult because increasing bandwidth on one link between nodes requires increasing it throughout the ring architecture, even where it is not needed.

The development of WDM transmission technology and more recently emerges of optical multiplexers and optical cross-connect (OXC) devices are moving optical networks towards a vision of all optical networks. In such a network, optical signals can be added and dropped to build connections without being converted into electrical domain, offering abundant and inexpensive bandwidth to the end users. To some, a key issue to realise such a vision will be that optical connections can be provisioned automatically to create bandwidth between end users, with timescales on the order of minutes or even seconds.

This requires a new generation of optical networking technology offering strong switch and router intelligence, along with a mesh network architecture. By using a network-wide control plane, mesh topologies make network configuration and traffic engineering much easier and flexible, enabling the all-important “point-and-click” dynamic provisioning.

Currently there is a consensus in the industry to extend IP protocols to serve as an IP-centric control plane for the optical network. This is built on the belief that signalling and routing mechanisms developed for IP traffic engineering applications could be re-used in optical networks.

At the same time, characteristics of traffic in the optical network have also changed. Traditional optical networks are dominated by voice traffic and private-line traffic. In such networks, all traffic is treated identically with full protection. Internet growth has diminished the predominance of voice traffic and private-line traffic relative to the much greater growth of data traffic, which presents a wider range of resilience requirements. For example, traffic generated by residential Internet access services typically requires a much lower grade of service than that of corporate financial transactions. In such a situation, providing all traffic with the same level of full protection proves to be very costly and wasteful.

As the fundamental infrastructure of optical networks changes, one also has to rethink the resilience provisioning in this new network environment. This research focuses on investigating resilience provisioning mechanisms in next generation IP-centric optical networks.

1.2 Contribution of this Research

Current advances in optical communication technology are moving optical networks towards a new generation, where optical connections can be added and dropped automatically and dynamically. There are also efforts to apply an IP-centric control plane to realise networking functions including neighbour topology discovery, automatic routing and connection establishment.

This research focuses on developing novel resilience provisioning mechanisms for this new generation of optical networks, which has as yet not attracted much attention from the research community.

The contribution of this research comprises three main elements:

1. A flooding-based reactive restoration scheme named Fast Restoration Scheme (FRS) is proposed.

   Flooding-based restoration uses the flooding messages to discover alternative paths after the failure occurs. It does not need the network node to maintain a global state of the network, thus it is easy to implement.

By maintaining a dynamically refreshing *Resource Table* in the *Receiver*, FRS precludes the possibility of link contentions and usually finishes the restoration connection with only one connection attempt. The mechanism of setting up restoration path from the *Selector* ensures loop-free restorations.

A patent by the author based on elements of this work has been filed by Nortel Networks [PAT1].

2. A novel resilience provisioning scheme entitled Adaptive Segment Path Restoration (ASPR) is proposed.

   In this approach, an LSP is divided into several segments. For each segment of the primary path, a separate backup path is provided. The segmentation of the primary path is adaptive to the topology of the network, allowing for more efficient resource usage whilst yielding restoration times comparable to link restoration. The implementation of the proposed scheme needs only some enhancement to the existing MPLS/GMPLS signalling protocols, which makes it simple and be able to work automatically. The comparative study and simulation results of the proposed scheme with others show that ASPR has the best restoration time performance, whilst remaining better than most other restoration schemes in terms of its spare capacity requirement.

3. A Differentiated-Resilience Optical Services Model (DROSM) for next generation optical network is proposed.

   In order to provide a range of resilience types that better reflect the value of the traffic being carried, this research proposes to provide differentiated levels of resilience for optical services.

   It considers classifying optical services according to their resilience requirements. Each resilience class is then provided with a different restoration strategy. The decision of restoration strategies is based on a novel analysis of optical restoration. In addition, a novel resource management mechanism is put forward to coordinate different resilience classes.

   A patent by the author based on elements of this work has been filed by Nortel Networks [PAT2].

1.3 Outline of the Thesis

Chapter 2 gives a brief overview of the evolution of optical networking technologies. Traditionally optical networks are static, in which services are manually provisioned and then exist for months or years. New automated means of managing the network resources, together with demands for greater flexibility from the customers, have led to the concept of dynamic optical networking. This chapter explains how this could happen and the main technologies being involved.

Chapter 3 gives a detailed introduction of resilience provisioning mechanisms in the optical network. In this chapter, a new classification framework of optical resilience provisioning mechanisms is presented to categorise all the existing resilience mechanisms. The purpose is to provide some insights into the inherent relations between different resilience provisioning mechanisms. The particular focus of the author's research is positioned within this framework allowing the contributions of this thesis to be placed in context.

Chapter 4 presents a novel flooding-based restoration scheme entitled Fast Restoration Scheme (FRS). A brief introduction of flooding-based restoration is first given, followed by details of the scheme. Then the scheme is evaluated via simulations with a realistic network.

In chapter 5, a new resilience provisioning scheme is proposed for mesh optical networks. This chapter starts with introduction of some basic terms used in MPLS / GMPLS restoration. Then the novel scheme is detailed together with a comparative simulation study.

Chapter 6 describes a novel Differentiated-Resilience Optical Services Model (DROSM), which proposes using different resilience provisioning for different optical services. The aim is to provide a more cost-efficient and flexible means for network carrier operators to exploit their network. In this chapter, the model is applied to wavelength-routed optical networks. Its performance is validated against different network topologies using different traffic patterns.

In chapter 7, the DROSM framework is further extended and applied to optical networks with wavelength conversion capabilities. Simulation results show that differentiated-resilience provisioning is also more cost-efficient than the single level resilience provisioning.

Chapter 8 concludes this research with the contributions being highlighted. It includes a general discussion and evaluation of the research. The integration of the novel schemes

proposed in this thesis is also discussed. Finally, a conclusion of this research is presented and areas for future work are also considered.

Chapter 2 Evolution of the Optical Network

2.1 Overview

Transport networks consist of functionalities necessary to provide cost-effective transport, multiplexing, routing, supervision, and survivability of service layer signals. Traditional optical networks are synchronous optical network (SONET) / synchronous digital hierarchy (SDH)-based time-division multiplexed (TDM) networks with wavelength division multiplexing (WDM) used strictly for fibre capacity. Provisioning an end-to-end optical service in such a network is extremely onerous and generally takes several months to accomplish. In addition, the topologies of traditional optical networks are fixed and the network configurations are all static. These characters make the network inflexible to accommodate the rapid increase of Internet traffic and unable to support more sophisticated broadband services, such as Bandwidth-on-Demand (BOD) and Optical Virtual Private Network (OVPN), etc [BEN01][VEE01].

In contrast, the emergence of a new generation optical networking technology, with the development of WDM transmission technology and more recently optical multiplexers and optical cross-connect (OXC) devices, is moving the optical network toward a pure optical network. In particular, these technologies provide the abilities to add, drop, and construct wavelength-routed networks, heralding a new era in which bandwidth is relatively abundant and inexpensive [SEN01]. In such a network, optical connections (lightpaths) can be provisioned automatically and dynamically to create bandwidth between end users. As a result, the new generation optical networks are more flexible, more cost-efficient, easier to manage, and able to support more sophisticated broadband services.

This chapter provides a background into the evolution of optical networking technologies with an emphasis on the new generation optical networking technologies.

2.2 Evolution of Optical Networks

2.2.1 TDM and WDM

There are currently two different multiplexing technologies in use in optical networks: Time Division Multiplexing (TDM) and Wavelength Division Multiplexing (WDM) [GOR01][LIUK02][BLA01] as shown in Figure 2-1.

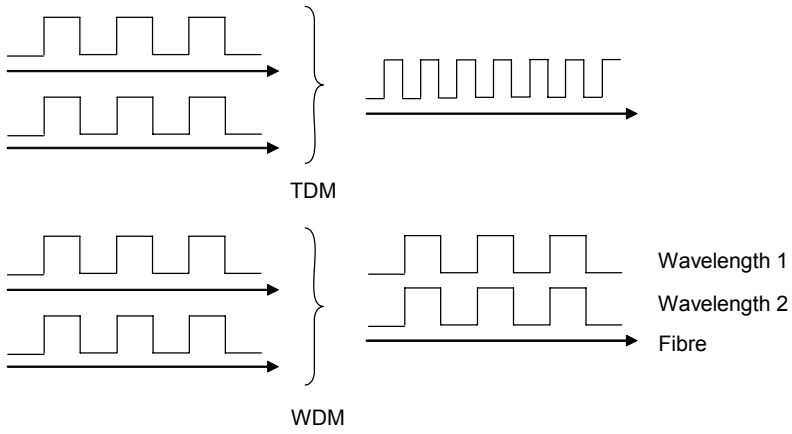


Figure 2-1: TDM and WDM

TDM is achieved through multiplexing many lower speed data streams into a higher speed stream at a higher bit rate by means of non-overlapping time slots allocated to the original data streams.

WDM is used to transmit data simultaneously at multiple carrier wavelengths through a single fibre. With this technology, the bandwidth of a channel is divided into multiple channels, and each channel occupies a part of the larger frequency spectrum. In WDM networks, each channel is called a **wavelength**.

TDM is widely used in the electrical domain to better utilise the carrier cable while WDM exploits the characteristic of light transmission in the fibre in the optical domain.

2.2.2 SONET/SDH Transport Network

Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH) are two TDM standards widely used by operators to transport and multiplex different tributary signals over optical links, thus creating a multiplexing structure called the SONET/SDH multiplex [BLA02].

The fundamental signal in SONET is STS-1, which operates at a rate of about 51Mb/s, while the fundamental signal of SDH is the STM-1, which operates at a rate of about 155 Mb/s. These two signals are made of contiguous frames that consist of a transport overhead (header) and a payload. To solve synchronisation issues, the actual data is transported in another internal frame that floats over two successive SONET/SDH payloads, and is named a Synchronous Payload Envelope (SPE) in SONET and a Virtual Container (VC) in SDH.

The transport networking functions of SONET/SDH are primarily performed by three broad classes of network element: terminal multiplexers (TMs), add/drop multiplexers (ADMs), and digital cross-connects (DXCs).

SONET/SDH networks can be configured in point-to-point, ring or mesh topologies, although most SONET/SDH networks are configured in a ring topology. Figure 2-2 shows how the network elements have been deployed to form a typical SONET/SDH transport network architecture. The speed of the links interconnecting ADMs usually starts at OC-3/ STM-1 (155 Mb/s) and can go up to OC-192 / STM-64 (10 Gb/s).

As a transport protocol using fibre optical links, SONET/SDH possesses a very rich set of network operation, administration, maintenance, and provisioning (OAM&P) capabilities. The SONET/SDH protocol also provides important protection and restoration capabilities, which will be discussed in the next chapter.

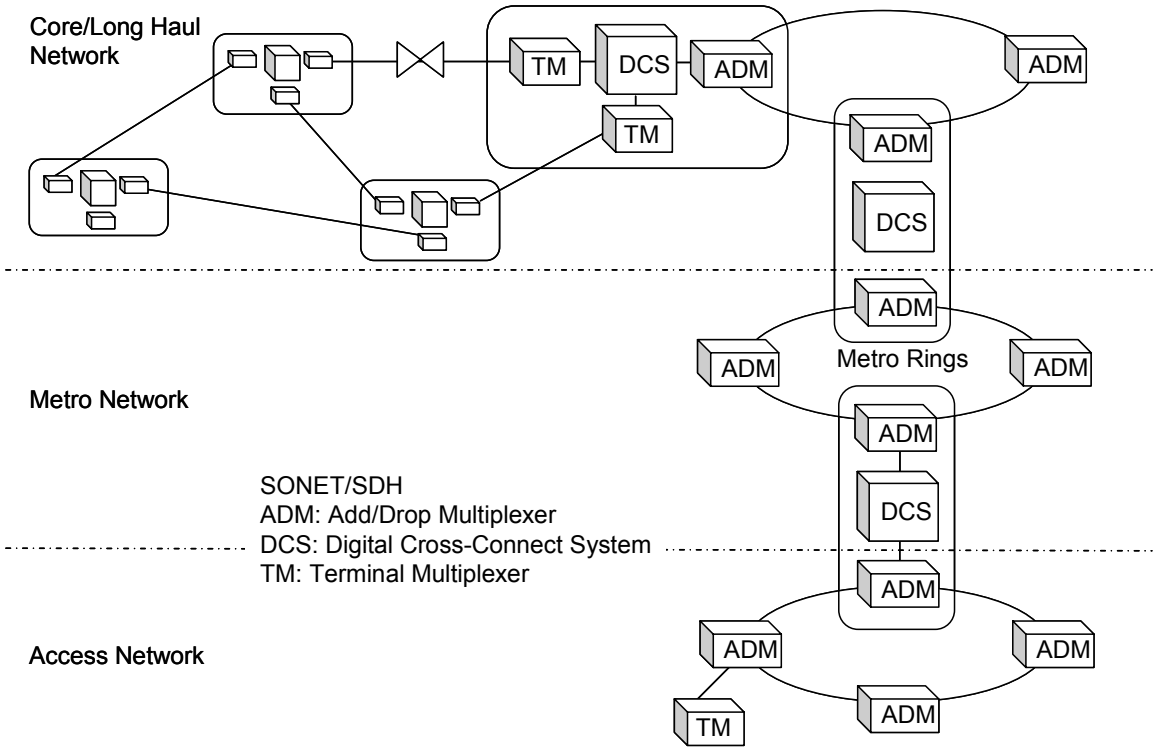


Figure 2-2: SONET/SDH Transport Network

2.2.3 Point-to-Point WDM Optical Network

WDM is built on a well-known concept called frequency division multiplexing (FDM) [LIUK02]. With this technology, the bandwidth of a channel (in its frequency domain) is divided into multiple channels, and each channel occupies a part of the larger frequency spectrum. In WDM networks, each channel is called a **wavelength**. This name is used because each channel operates at a different frequency and a different optical wavelength. In addition to the term wavelength, the term frequency slot, **lambda (λ),** and optical channel are also used to describe the optical WDM network channels.

2.2.3.1 WDM and DWDM

Essentially there is no difference between WDM and Dense WDM (DWDM) in optical networking [BLA02]. Both methods involve placing multiple wavelengths over a single strand of fibre optic cable. The only difference is the density of placement of the separate optical wavelengths. The most common spacing is referred to as a 100 GHz (0.8 nm) spacing. Others are emerging that pack the wavelengths closer together at spacing of 50 GHz (0.4 nm) and 25 GHz (0.2 nm).

A common set of wavelengths used today is in the 1550 nm region, which is referred to as the C band. Another frequency band is the L band, which operates above the C band in the 1574.37 nm to 1608.33 nm range.

DWDM systems allow the multiplexing of more than 160 wavelengths of 10 Gb/s (1.6 Tb/s per fibre with a 25 GHz spacing) by using both the C band and L band spectra. Some vendors are proposing a spacing of 12.5 GHz. Consequently, it will be possible to transmit 320 wavelengths of 10 Gb/s in a single fibre.

2.2.3.2 WDM Point-to-Point Networks

WDM was initially deployed as point-to-point systems to alleviate capacity exhaust in core transport networks. The traditional techniques for increasing capacity have included deployment of additional fibre and replacement of current capacity TDM transport systems with new higher-rate TDM systems. The former can be an expensive proposition while the latter generally requires replacement of transport systems and affords little in terms of equipment reuse. In contrast, point-to-point WDM systems offer a cost-effective capacity expansion for the transport network (Figure 2-3).

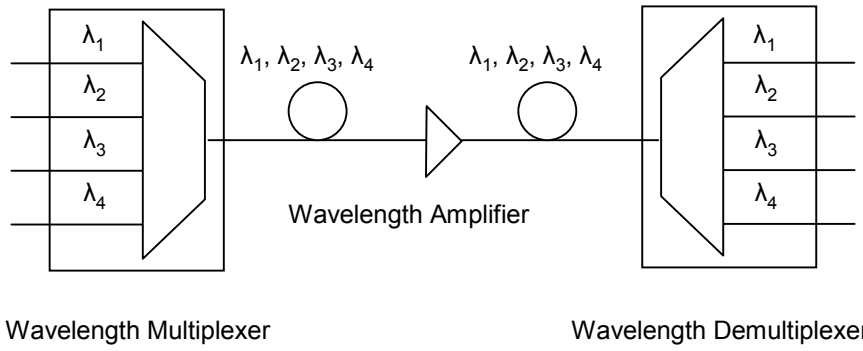


Figure 2-3: A Four-wavelength Point-to-Point WDM Transmission System with Amplifiers

2.2.4 WDM Optical Networking

The use of point-to-point WDM represents the first step toward optical networking, because it employs wavelength-based transport. However, in these initial point-to-point applications, most of the networking functionality remains the responsibility of the SONET/SDH-based TDM systems. The network resilience is also provisioned by the SONET/SDH-based transport layer. The optical layer only serves as statically deployed physical links to expand the transport capacity. When WDM networks are deployed as such,

every cross-connecting node performs optical-to-electrical (O/E) and electrical-to-optical (E/O) conversions, and switching is performed in the electrical domain [MAE98][CHA98].

As more and more point-to-point WDM systems are installed in the transport network, and more traffic is carried on WDM networks, it is desirable to reduce the number of O/E and E/O conversions in the network. The ultimate goal is to connect wavelengths on an end-to-end basis, where a wavelength goes through the network without O/E and E/O conversions. This process is known as optical networking, such connections are sometimes termed lightpaths, and such networks are known as transparent networks or all-optical networks.

As network traffic grows and optical channels increasingly become the medium for exchange in networks, carriers will need to manage capacity at the optical channel level. The use of WDM in the transport network will quickly evolve from point-to-point capacity expansion to scalable and robust optical transport networking applications catering to an expanding variety of client signals with equally varied service requirements. This is made possible by the mature of the key optical networking nodes.

2.2.4.1 Optical Add/Drop Multiplexer (OADM)

The main functionality of an OADM is to add, drop, or pass-through wavelength channels in a WDM enabled optical network. Figure 2-4 shows a possible structure of an OADM. In the figure, there are four input and output fibres, each of which supports n wavelengths. An incoming optical signal over the input fibre is demultiplexed through a wavelength demultiplexer. Each of the wavelength channels matches one fibre port. The demultiplexed signal can propagate directly through the fabric without changing wavelength or it can be dropped onto one of the fabric drop ports through a physical configuration of filters. Likewise, a wavelength can be added through an add fabric port and directed to a wavelength port by configuring corresponding filters. The outgoing wavelengths are multiplexed onto outgoing fibres and exit points.

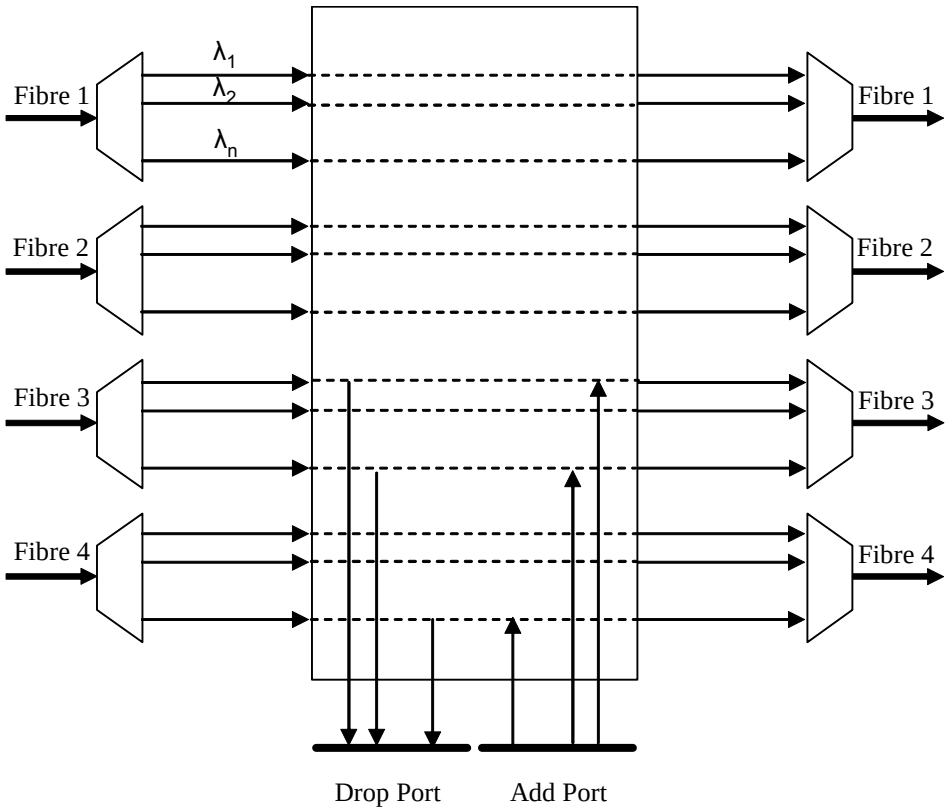


Figure 2-4: OADM

2.2.4.2 Optical Cross-connect (OXC)

An OADM isolates wavelengths to selectively access a wavelength channel. However, another useful function is to rearrange wavelengths from fibre to fibre within a WDM network. This is provided in an OXC. An OXC provides wavelength-level switching.

Figure 2-5 shows a possible structure of an OXC. In the figure, there are four input and output fibres, each of which has a number of wavelengths. Depending on the switch setting, a signal over a certain wavelength from one fibre can be connected to the same wavelength but on a different outgoing fibre. This can be accomplished without wavelength conversion. An OXC without wavelength conversion capability is also known as a Wavelength Selective Cross-Connect (WSXC) [CHA98][LIUK02].

In fact, it is likely that more than one signal will compete for a wavelength channel on one outgoing fibre, which causes outgoing fabric port contention. To ease this problem, wavelength conversion/interchange can be introduced to direct a wavelength to the fibre with a different optical frequency. An OXC that employs wavelength conversion is also known as a Wavelength Interchanging Cross-Connect (WIXC) [CHA98][LIUK02].

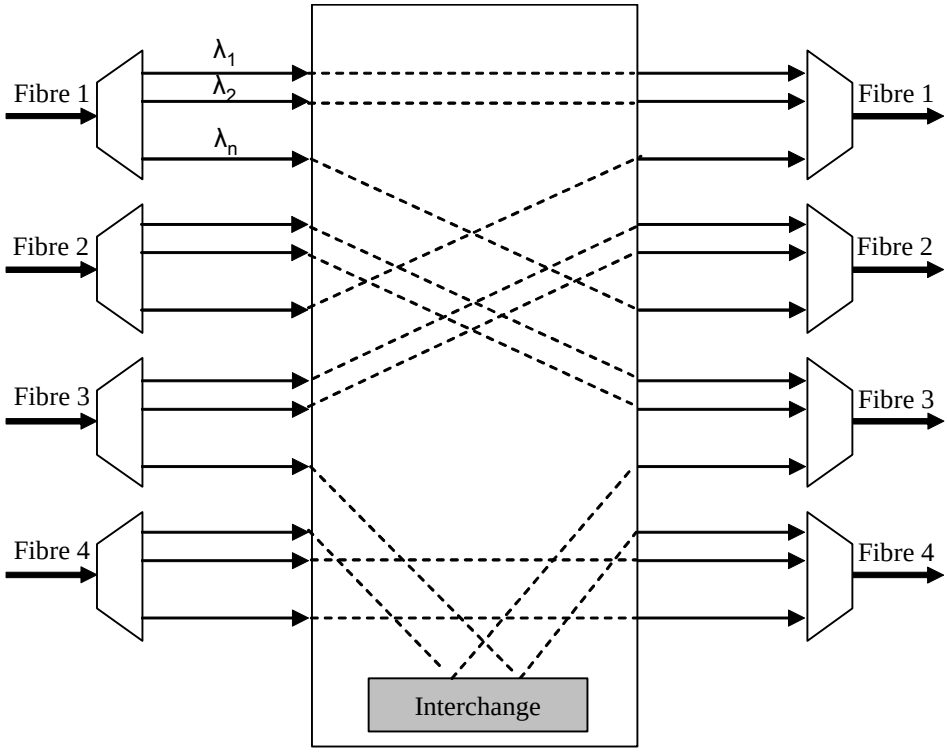


Figure 2-5: OXC

As the technique of optical wavelength conversion remains immature, wavelength conversion is usually performed through converting the optical signal to electricity, and then using the recovered bit stream to drive the modulation of a second wavelength. The electronic switching and processing costs of wavelength conversion at the each OXC can potentially be very high, leading to severe performance bottlenecks and limiting the delivery of optical bandwidth to the end users. Therefore, wavelength conversion should be used only when necessary.

2.2.4.3 Wavelength Continuity Constraint

As the cost is very high, wavelength conversion should be used sparingly. In the absence of wavelength conversion capability, a lightpath must occupy the same wavelength on all the fibre links through which it traverses. This property is called the wavelength-continuity constraint. An optical network that has no wavelength conversion capability is called a wavelength-routed optical network [ZAN01].

2.2.4.4 Routing and Wavelength Assignment (RWA)

The problem of finding a route for a lightpath and assigning a wavelength to the lightpath is known as the Routing and Wavelength Assignment (RWA) problem. The objective of the problem is to route lightpaths and assign wavelengths in a manner that

minimises the amount of network resources that are consumed, and at the same time ensures that no two lightpaths allocate the same wavelength on the same fibre link. In a wavelength-routed optical network, the RWA problem [ZAN01][ASS01] operates under the constraint that a lightpath must occupy the same wavelength along its route.

In a wavelength-routed optical network, the traffic can be either static or dynamic. The RWA problem can be considered under these two different traffic patterns. The static RWA problem applies to the case in which the set of connections required to be established is known in advance. The problem is then to set up lightpaths in a global fashion while minimising network resources such as the number of wavelengths or the number of fibres in the network. The static RWA problem is know as Static Lightpath Establishment (SLE) [ASS01][ZAN01] and can be formulated as an Integer Linear Program (ILP) [SCH98].

In the dynamic traffic pattern conditions, connections arrive to the network dynamically and remain for some finite amount of time before being dropped. This dynamic RWA problem is referred to as the Dynamic Lightpath Establishment (DLE) [ZAN01][ASS01] and can be further divided into two sub-problems: routing sub-problem and wavelength assignment sub-problem.

For the routing sub-problem, there are three basic approaches: fixed routing, fixed-alternate routing, and adaptive routing. For the wavelength assignment sub-problem, a number of schemes have been proposed. These schemes include: Random, First-Fit, Least-Used, Most-Used, Least-Loaded, Min-Product, MAX-SUM, and Relative Capacity Loss (RCL). Among the above schemes, RCL often achieves the best performance and the First-Fit technique gives nearly as good results. However, RCL requires global knowledge of the network status and the First-Fit assignment scheme requires only knowledge of the links along the route. First-Fit is also simple to implement [JUE].

2.3 IP-Centric Control Architecture

Due to the rapid growth of data traffic demand in recent years, the industry believes that optical networking is the key solution to keep up with the growth. As a result, considerable interest has been focused on optical networking. Key optical elements are being developed to increase network capacity and scalability. In order to automate the lightpath provisioning procedure, one of the key areas of focus is the optical control plane. The optical control plane is designed to provide simpler, faster and more flexible provisioning of optical connections in optical networks. Historically, a centralised connection management approach has been used to address this issue. The drawback of using a centralised approach is that it requires a

complex Network Management System (NMS). The process of integrating equipment from multiple vendors into a single NMS can be costly and lengthy. Therefore, a common control plane standard is important. Currently a consensus is emerging in the industry on utilising an IP-centric control plane within optical networks to support optical networking functionalities [SEN01][SAH03].

Multi-Protocol Label Switching (MPLS) [DAV00] is a control framework currently being developed as a standard to enable fast switching in IP networks. MPLS control mechanisms can be used to establish a label-switched path (LSP). The concept of MPLS and its constraint-based Traffic Engineering (TE) models can be extended to wavelength-routed optical networks as Multi-Protocol Lambda Switching (MPλS) [AWD01]. The Internet Engineering Task Force (IETF) is currently working on Generalised Multi-Protocol Label Switching (GMPLS) [BAN01][ASH01], a generalised control framework for establishing various types of connections, including lightpaths.

2.3.1 MPLS and Traffic Engineering

Packet-based MPLS uses labels to make forwarding decisions at the network nodes, in contrast to traditional destination-based hop-by-hop forwarding used in IP networks.

In connectionless network routing protocols, the packet forwarding decision is taken independently at each hop as the packet is sent from one hop to the next. In the traditional IP forwarding paradigm each router forwards the packet by using the IP destination address field in the packet header. Every router has a routing table that contains tuples of the form, <destination address, output interface>. The router reads the destination address from the header of an incoming packet and uses the routing table to forward it on the appropriate output interface.

MPLS is an advanced framework for fast label switching. In MPLS, a short fixed length value called a label is assigned to the packet, as the packet enters of the network. The packet is forwarded to its next hop together with this label. At subsequent hops, there is no further analysis of the packet's network layer header. When a packet reaches a core packet LSR, this LSR uses the label as an index into a forwarding table to determine the next hop and the corresponding outgoing label, writes the new label into the packet, and forwards the packet to the next hop.

An MPLS network consists of MPLS nodes called label switching routers (LSR) connected by circuits called label switching paths (LSP). Border LSRs in an MPLS domain act either as ingress or egress LSRs depending on the direction of the traffic being forwarded.

MPLS allows the establishment of LSPs between ingress and egress LSRs. Each LSP is associated with a forwarding equivalence class (FEC), which may be thought of as a set of packets that receive identical forwarding treatment at an LSR (e.g., the set of destination addresses lying in a given address range). To establish an LSP, a signalling protocol such as Label Distribution Protocol (LDP) / Constraint-based Routed LDP (CR-LDP) or Resource reSerVe Protocol Traffic Engineering extension (RSVP-TE), is required. Between two adjacent LSRs a short, fixed-length identifier called a label (significant only between the two LSRs) locally identifies an LSP. The signalling protocol is responsible for the inter-node communication that assigns and maintains these labels.

When a packet enters an MPLS-based packet network, it is classified according to its FEC and, possibly, additional rules that together determine the LSP along which the packet is sent. For this purpose, the ingress LSR attaches an appropriate label to the packet and forwards the packet to the next hop. The label may be attached to a packet either in the form of a header encapsulating the packet or it may be written in the circuit identifier field of the layer 2 encapsulation of the packet. In its generalized version, the label could be a value representing a time-slot, a wavelength, or even a fibre.

2.3.2 GMPLS Basics

The notion of an IP-centric control plane for optical networks was first described formally in an IETF Internet draft in November 1999. Note that this was after a number of vendors had already introduced the concept. This architecture was based on applying MPLS control concepts to optical networks. It was first called Multiprotocol Lambda Switching (MPλS) [AWD01], but later it was recognised that the same concepts could be generalised to control any circuit-switched network. Thus, the term generalised MPLS or GMPLS [ASH01][AWD01] is now used to describe the application of MPLS protocols to control other networks.

GMPLS is introduced to generalise the MPLS architecture to also consider non-packet-based bearer planes in addition to the conventional packet networks. The original MPLS mainly focuses on the data plane – the actual data traffic. On the other hand, GMPLS focuses on the control plane that performs connection management for the data plane for both Packet Switched Capable (PSC) interfaces and non-packet switched interfaces. These interfaces include:

- Packet Switched Capable (PSC),
- Time Division Multiplex (TDM) capable,
- Lambda Switched Capable (LSC),

- Fibre Switched Capable (FSC).

MPLS requires the LSP be set up between routers at both ends, while GMPLS extends the concept of LSP setup beyond routers. The LSP in GMPLS can be set up between any similar types of LSR at both ends. For example, the LSP can be setup between SONET/SDH ADMs to form a TDM LSP; the LSP can also be set up between two wavelength switching capable systems to form a LSC LSP; or the LSP can be set up between fibre switched capable optical cross-connect systems to form an FSC LSP (Figure 2-6).

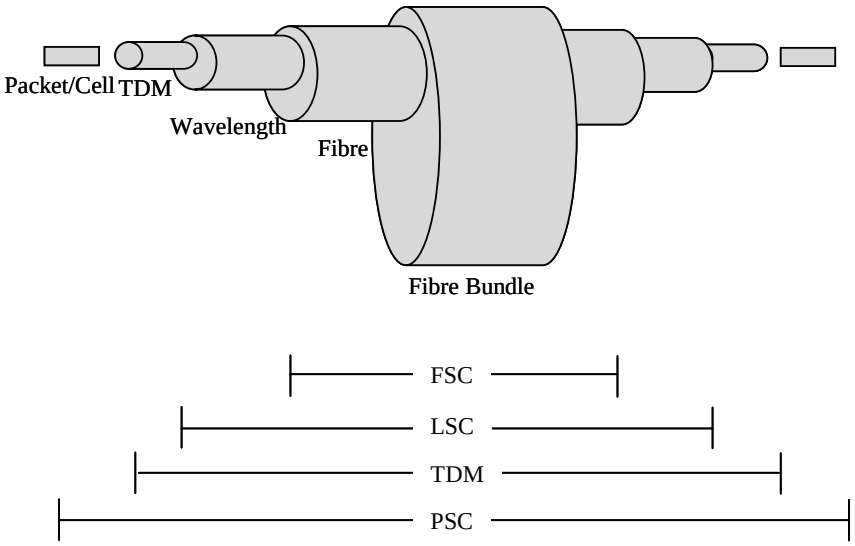


Figure 2-6: GMPLS interface hierarchy

2.3.3 IP-Centric Control Architecture for Optical Networks

One of the important uses of GMPLS is to address the control plane needs of optical networks. Such a control plane must includes several basic functions, such as network addressing, neighbour discovery, topology discovery, routing control and connection management, in order to support dynamic provisioning of lightpaths. To implement these functions, the development of GMPLS requires enhancements to existing IP signalling and routing protocols [BAN01].

2.3.3.1 Network Addressing

When IP protocols are extended to control optical networks, new constraints on the addressing and routing models are introduces since several hundreds of parallel physical links (e.g. wavelengths) can now connect two nodes. Most of the carriers already have today the

capability for several tens of wavelengths per fibre between two nodes. New generation of DWDM systems will allow several hundreds of wavelengths.

It becomes rather impractical to associate an IP address to each end of each physical link, to represent each link as a separate routing adjacency, and to advertise link states for each of these links. For that purpose, GMPLS enhances the MPLS routing and addressing model to increase their scalability.

It is assumed that each OXC in the optical network has a unique IP address which serves to identify the OXC and as a basis for creating an IP-centric control plane. A selector identifies further fine-grain information of relevance at an OXC. The selector can be formatted to have an adequate number of bits and a structure that expresses port, channel, and other identifications.

Within the WDM network, the establishment of lightpath segments between adjacent OXCs requires the identification of specific port, channel. In the framework of GMPLS, a label serves this function. The structure of the optical label is designed in such a way that it can encode all the required information (including WDM-specific information).

Another entity that must be identified is the Shared-Risk Link Group (SRLG). An SRLG is an identifier assigned to a group of optical links that share a physical resource. For instance, all optical channels routed over the same fibre could belong to the same SRLG. Similarly, all fibres routed over a conduit could belong to the same SRLG. The assignment of unique identifiers to these SRLGs within a WDM network is essential to ensure correct SRLG-disjoint path computation for protection and restoration.

Optical links between adjacent OXCs may be bundled for advertisement in a link state protocol. The component links within the bundle must be identifiable. In concert with SRLG identification, this information is necessary for correct path computation.

2.3.3.2 Neighbour Discovery

Routing within the WDM domain relies on knowledge of network topology and resource availability. This information may be gathered and / or used by a centralised system, or by distributed route computation entities. In either case, the first step towards network-wide link state determination is, for each OXC, to discover the status of local links to neighbours. In particular, each OXC must determine the up/down status of each optical link, the bandwidth and other parameters of the link, the identity of the remote end of the link, and the consistency of link parameters with the information available at the other end of the link.

The determination of these parameters could be based on a combination of manual configuration and an automated protocol running between adjacent OXCs. In general, this type of protocol can be referred to as a Neighbour Discovery Protocol (NDP). The Link Management Protocol (LMP) [LAN00] is an example of a NDP. It also contains other management functions such as link management and fault isolation.

LMP runs between a pair of nodes directly over IP with a distinct protocol ID. The core function set includes control channel management and link property correlation.

Control channel management constructs and maintains link connectivity between neighbouring nodes. This requires lightweight Hello messages that act as a fast keep-alive mechanism between the nodes. This message is encapsulated in an IP packet and sent to a designated IP multicast address. The content of the Hello message includes the IP address of the sending OXC, the port number of the link over which the packet is sent, and other parameters (e.g., SRLG information) whose consistency must be verified. This packet is received and processed by the neighbour, which repeats the received information along with the corresponding information from its side.

Link property correlation is used to exchange the local and remote property mapping. It is used to synchronise the link properties in the data plane, such as link multiplexing/demultiplexing capability and the encoding type of the data link, between the adjacent nodes. This is implemented using the LinkSummary message set including LinkSummary, LinkSummaryAck, and LinkSummaryNack.

Two optional procedures offered by LMP are link connectivity verification and fault localization. Link connectivity verification offers a testing procedure to verify the physical connectivity of the data-bearing links and identify any misconfigurations. Fault localization localizes failures in the WDM network.

2.3.3.3 Topology Discovery

Topology discovery enables each OXC in a network to build a database representing the network topology and resource availability. Topology discovery is accomplished by running extended versions of a distributed IP routing protocol such as Open Shortest Path First (OSPF) [MOY98] or Intermediate System to Intermediate System (IS-IS) in each OXC. Here only OSPF is discussed.

OSPF is a link-state routing protocol designed to run within a single area / Autonomous System (AS). Each node in the area describes its own link states by generating Link State

Advertisements (LSAs). These LSAs are distributed to all nodes in the network using a process called reliable flooding. This information is used to create a Link-State Database (LSDB), which describes the entire topology of the area. Once a network has converged to steady state, all nodes will have identical link-state databases. As a result, any node in the network can use its link-state database to calculate the best route to any other node in the network.

Standard OSPF is designed for routing IP datagrams. In OSPF Traffic Engineering extension (OSPF-TE) [KOM01], new features are introduced to support optical networks.

2.3.3.3.1 Link Bundling

In standard OSPF, each physical link between a pair of routers would result in a routing adjacency and being represented in a LSA. This means that routing protocol messages would be exchanged over each such link, and a LSA for each such link would be created and advertised to other nodes in the network. For optical networks, there would be a large number of parallel physical links between a pair of neighbours. Thus, the forming a routing adjacency and creating a separate LSA for each physical link will result in extreme traffic overhead. To address this, OSPF-TE [KAT01] treats all the parallel links between a pair of neighbours as a single routing adjacency. This mechanism is known as link bundling. With link bundling, all the parallel physical links between a pair of OXCs are coded in a single LSA, which is then *flooded*.

2.3.3.3.2 Resource Parameters

In standard OSPF, each link is assigned a cost. Such values are disseminated via LSA in the network for each node to calculate the route table. In an optical network, a logical routing adjacency could contain a large number of parallel physical links. Therefore, the extended OSPF should support carrying more information in LSAs. Such information consists of the representation of links and nodes in the network along with certain associated resource parameters (e.g. link cost, resource type and availability, SRLG information) that are critical to routing of lightpaths.

2.3.3.3.3 LSA Update

In standard OSPF, a LSA is re-originated when its link parameter changes. However in OSPF-TE, as link state advertisements carry more link parameters such as resource availability, care must be taken to ensure that this information is not generated too frequently with minor changes in resource states. A configurable threshold scheme needs to be

introduced whereby an OXC would generate a link state update only if a certain amount of link resource information has changed.

2.3.3.3.4 Source Routing Methodology

Standard OSPF is designed for routing IP datagrams. Hence, under standard OSPF each participating node would use an identical algorithm to compute a forwarding table that allows packets to be routed based on the destination address. Routing of an optical layer connection, on the other hand, requires that the entire path for the connection be computed at the source OXC and signalled to other OXCs in the path.

The new link representation and resource parameters are incorporated into OSPF through traffic engineering extensions. Extensions to OSPF for supporting GMPLS are described in [KOM01].

2.3.3.4 Signalling Protocols for Lightpath Establishment

The MPLS architecture for IP networks defines protocols for establishing Label Switched Paths (LSPs). These protocols are extended to provision of traffic engineering virtual circuits in an IP network. The TE-oriented characteristic enables those signalling protocols to be adapted for provisioning lightpaths in optical networks.

There are two choices for MPLS-based signalling protocols: Constraint-based Routed Label Distribution Protocol (CR-LDP) [RFC3036][RFC3037][RFC3214] or Resource Reservation Protocol with Traffic Engineering extensions (RSVP-TE)[RFC3209].

CR-LDP has its foundations in LDP, and is extended to incorporate the explicit route information. An explicit route is represented in a Label Request Message as a list of nodes along a constraint-based route. To establish the LSP, the ingress node sends out a Label Request Message to the downstream node along the explicit route. That downstream node then checks if it is the egress node (destination). If not, a Label Request Message carrying the refreshed explicit route is sent out further downstream until such a message reaches the egress node. If the check of resource availability is successful, the egress node will send a Label Mapping Message to the upstream node. Otherwise, an error Notification message is sent out. Each interim node only performs further action until a Label Mapping Message or a Notification message is received. A LSP is set up successfully after a Label Mapping Message reaches the ingress node. Unlike RSVP-RE, CR-LDP adopts a “hard state” mechanism, in which a established LSP needs no further refresh messages and will exist until a Label Release Message or a Label Withdraw Message is received.

RSVP-TE adopts a “soft state” mechanism. Once the routing path is determined through a routing protocol or an external traffic engineering application, the RSVP daemon of ingress node start a session and sends a RSVP PATH message along the routing path to the destination. When the PATH message reaches the destination, the receiver initiates the reservation and sends a RSVP RESV message along the reserve routing path to the sender. During reservation, each node including the receiver and the sender is responsible for choosing its own level of reserved resources, a process known as admission control, to determine whether it can supply the desired Quality of Service (QoS). If the admission control succeeds, the corresponding parameters at the node’s packet classifier and scheduler are set and the reservation request message relays towards the data source. If the admission control fails, an error message is sent to the source. The established RSVP channels are soft state maintained, in which channel states are maintained at each node and applied with timers. These channels should be periodically refreshed through PATH and RESV messages. Otherwise, when the timer expires, the channel states will be deleted and the resource will be released.

Both these protocols allow hop-by-hop and explicitly routed signalling from a source to a destination node to establish unidirectional LSPs. New features must be introduced in these protocols to accommodate the peculiarities of lightpath provisioning in optical networks, including support for establishing bidirectional paths, support for establishing shared backup paths, and fault tolerance. Extensions for some of these requirements have already been proposed and are described for RSVP-TE in [BER02] and for CR-LDP in [ASH02].

2.4 Summary

This chapter provides a background into the evolution of optical networking technologies.

TDM-based SONET/SDH optical transports are widely used in traditional optical transport networks. Point-to-point WDM systems are introduced to expand the exhausted link capacity between two offices. O/E/O conversion is used for the traffic to transmit between those point-to-point systems. With more and more WDM systems being deployed, a straightforward approach is to eliminate the costly O/E/O conversion systems and let the optical signal pass through. It brings a vision of all-optical networks, in which lightpaths can be established dynamically and automatically. The development of some main optical components such as OADM and OXC drives us towards such a solution. However, new control plane technologies also need to be developed to realise the goal. One of the most

promising candidates is to utilise an IP-centric control plane, which is derived from MPLS, the newly developed control framework for IP networks. The derived IP-centric control plane technology is defined within the GMPLS framework: To be applied to lightpath provisioning, those protocols defined under the MPLS framework are extended to fit in the optical network environment; In addition, a neighbour discovery protocol is also developed to address the different characteristic of optical links.

Chapter 3 Resilience Provisioning Mechanisms in Optical Networks

3.1 Introduction

Network resilience (survivability) is a crucial issue that must be concerned when designing a network infrastructure in order to ensure the integrity of the different services that it supports. Network resilience refers to the ability of a network to recover services affected by failures that may be encountered during its operation [GEO99]. It may be considered as a component of QoS [MUR96][PAI97]. This issue has been attracting numerous studies for different networks.

The study of network resilience provisioning is to investigate how to provide continuous service in the presence of failures at a possible low cost. In this literature, a number of mechanisms have been proposed to be applied to different networks, for example, Equal Cost Multi-path Forwarding (ECMF) and dynamic routing in IP networks [MOY98]; VP / VC pre-planned protection and dynamic restoration in ATM networks [WU97][KAW99][MUR96]; APS, Self-Healing Rings, mesh-based protection, and dynamic restoration in SONET/SDH networks [WU92][WAS94]; optical APS and Self-Healing Rings in WDM optical networks [GER00][MOH00]. Although their names may be different when apply to different networks, the fundamental principles are quite similar.

Optical network resilience is even more important as the optical network acts as the foundation transport layer and carries huge amount of traffic on a single network element such as a fibre or a cross-connect. A break in a cable equipped with terabit/s optical transmission systems can disrupt the equivalent of 250 million telephone calls at once; especially in long-distance transport networks, when the probability of cable cuts is not negligible. For example, in a Pan-European network with 25,000 fibre-routed kilometres, a cable cut is statistically likely every four days [DEM99].

Therefore, resilience provisioning is an essential task when designing an optical network. Chapter 3 gives a detailed introduction of resilience provisioning mechanisms in the optical network. In this chapter, a new classification framework of optical resilience provisioning mechanisms is presented to put all the existing resilience mechanisms together. The purpose is to provide some insights into the inherent relationships between different resilience

provisioning mechanisms. The particular research area and the contributions of this thesis are also given within this framework.

Although the main discussion is focused on optical networks, the classification and study results may apply to other network technologies, such as ATM, IP, MPLS.

3.2 General Classification of Resilience Provisioning Mechanisms

In optical networks, spare resource is needed in order to protect service traffic against possible network element failures. The methods of how to provide resilience in a network are called resilience provisioning mechanisms, or resilient schemes. They are also called protection schemes and / or restoration schemes in some situations [MOH00][WU97]. Designing such resilient networks while minimizing spare capacity costs is, not surprisingly, a major concern of telecom industry and service providers.

Resilience provisioning mechanisms can be generally classified according to two different criteria: protection versus restoration and dedicated versus shared protection.

Figure 3-1 shows the general classification of resilience provisioning mechanisms.

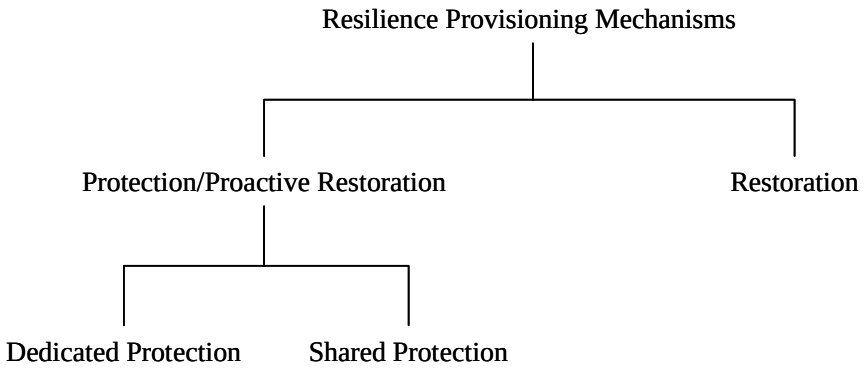


Figure 3-1: General Classification

3.2.1 Protection versus Restoration

The procedure of restoring affected traffic to a normal state is called restoration. Resilience provisioning mechanisms can be classified into two general categories: **protection** and **restoration**, depending on whether restoration route and resources are pre-assigned before the failure occurs or not. The technique that uses a pre-assigned alternative path and

capacity to ensure resilience is referred to as **protection**, and the technique that restore the affected traffic by finding an alternative path dynamically using available capacity is referred to as **restoration** [WU97][FUM00]. **Protection** and **restoration** are also referred to as **proactive/pre-planned restoration** and **reactive restoration** respectively in some studies [MOH00][COA91][MED99]. In this situation, restoration is referred to including both **protection** and **restoration**.

Protection pre-assigns **backup** paths and reserves resources at the time of establishing the **working** (or **primary**, both terms are used in this thesis) paths to protect traffic against possible failures, thus ensures a successful traffic restoration. Protection generally offers better traffic restoration speeds than reactive restoration, since it does not need the time-consuming path calculation/searching and connection reestablishment process. For example, in SONET/SDH, both Bidirectional Line Switched Ring (BLSR) and Unidirectional Path Switched Ring (UPSR) are protection and have a restoration specified to be less than 60 ms [ITUG872].

Restoration comes as the result of the introduction of mesh-based networks [GRO91][WU97]. As it identifies the restoration path only after the failure, restoration is much more flexible at choosing the alternative path, which results in better resource sharing. Restoration could be more cost-efficient and suitable for networks with rapid dynamic change of traffic demands, where pre-planned algorithms cannot provide a real-time solution [XIO99][RAM99]. The drawbacks of this approach are, firstly, that the amount of spare resource may not be adequate and thus cannot ensure a successful traffic restoration; secondly, that the restoration time can be several seconds or even longer, especially in heavily loaded networks [YE00]. In contrast to the 60 ms of SONET/SDH protection, a 2s restoration time goal [SOS94][WU94][GRO91] is commonly set for the dynamical distributed restoration using digital cross-connect (DCS) for ATM and SONET/SDH [GRO87][FUJ94] [YAN88][HAW95].

3.2.2 Dedicated versus Shared Protection

The **protection** performance can be improved by sharing the backup resources among different failure scenarios. Based on this issue, the resilience provisioning schemes of **protection** can be further classified into two categories: **dedicated protection** and **shared protection** [WU97][RAM99a][RAM99b].

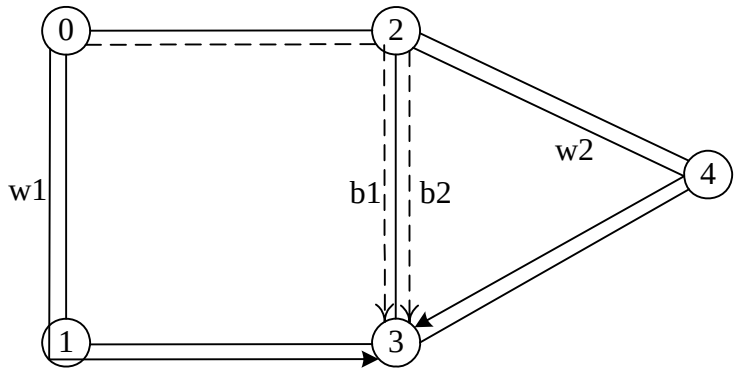


Figure 3-2: An illustration of dedicated / shared protection

In **dedicated protection** (also called 1+1 protection), when provisioning of the working path, a backup path is set up and dedicated to the connection. This method is illustrated in Figure 3-2. The figure shows two working paths, w1 and w2, and their respective backup paths, b1 and b2. At the time of a failure, **dedicated protection** only involves the switch actions at the two ends. In some cases, if the same signal is also transmitted in the backup path as in the working path, the traffic restoration only involves action at the merge point. Therefore, dedicated protection requires the least involvement of management system and has a very short restoration time [WU97][MOH00]. However, as the resource pre-allocated by each backup path only serves one particular working path, dedicated protection requires excessive resources for protection.

As the event of multiple failures is uncommon, one can assume there are always events of single failure in a network. Therefore two working paths that have no common network can share all or part of the resource used by their backup paths. In this case, it is called **shared protection** [WU97]. For example in Figure 3-2, because w1 and w2 have no common component, their backup paths, b1 and b2, could share the same channel along link 2-3. As it has better resource utilisation, **shared protection** is more cost-efficient when compared with **dedicated protection**. However, when applied to circuit-switched networks in which one incoming channel can only be connected to one outgoing channel at the same time [WON99][JUL94], shared protection requires more management involvement and its restoration time is relatively longer than that of dedicated protection. That is because signalling message is needed to inform the interim nodes to switch on the intended connection.

Reactive Restoration is proposed for better resource sharing among the backup paths by allocating resource only after the failure, thus is always **shared**.

3.2.3 Resilience Provisioning Mechanisms in Different Topologies

Network topology has a key effect on the resilience provisioning mechanisms. Figure 3-3 shows the classification of resilience provisioning mechanisms based on the network topology they are applied to (some terms are to be explained in Section 3.4.1). Usually only protection is used in the point-to-point / linear and ring topology networks, while both protection and restoration mechanism could be used in networks with a mesh topology.

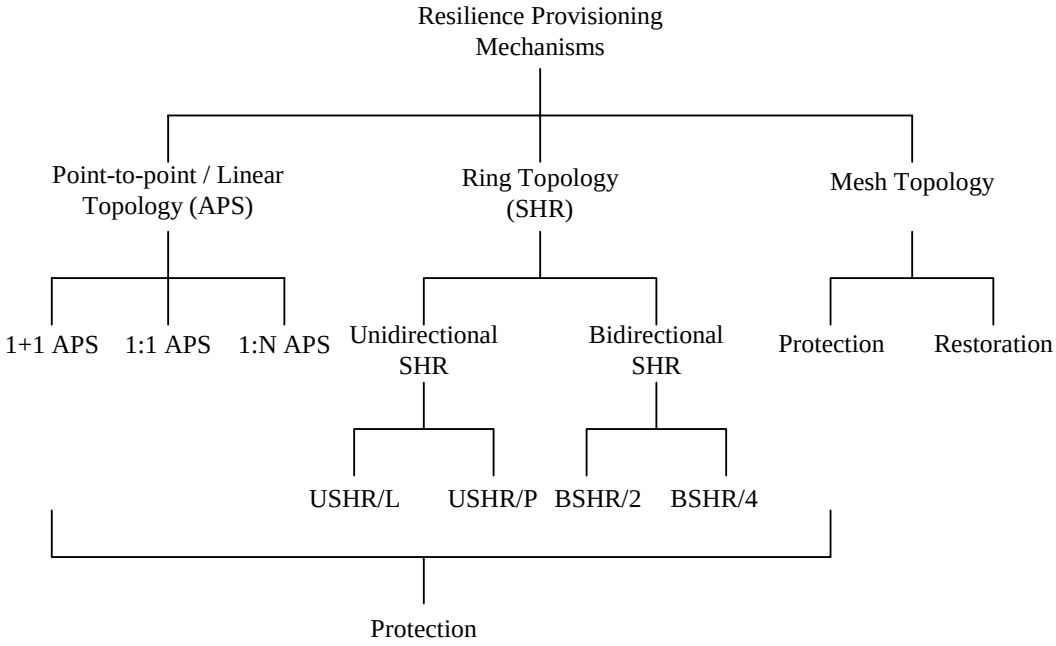


Figure 3-3: Resilience Provisioning Mechanisms

3.3 Resilience Provisioning Mechanisms in Linear Topology

3.3.1 Automatic Protection Switching

Automatic Protection Switching (APS) is typically used to handle link failures. It has three main architectures: 1+1, 1:1 and 1:N APS [WU97][ITUG872][FUM00]. The difference between the three architectures is the assignment of backup resources. In 1+1 APS (Figure 3-4(a)), a backup link is pre-assigned for every working link. The source node transmits the traffic on both the working and backup links. The receiver at the destination node compares the two signals and chooses the better one (e.g., the less noisy node). If one link fails, the destination node is still able to receive the signal on the operational link. In 1:1 APS (Figure 3-4(b)), every working link has a protection link, but the source and destination nodes switch to the backup link only when a failure on the working link is detected. Under normal

conditions, the backup link is either idle or used to carry low-priority traffic. Figure 3-4(c) shows how 1:*N* APS system works. In this scheme, *N* working links share a single backup link, thereby providing protection against the failure of any one of the *N* working links.

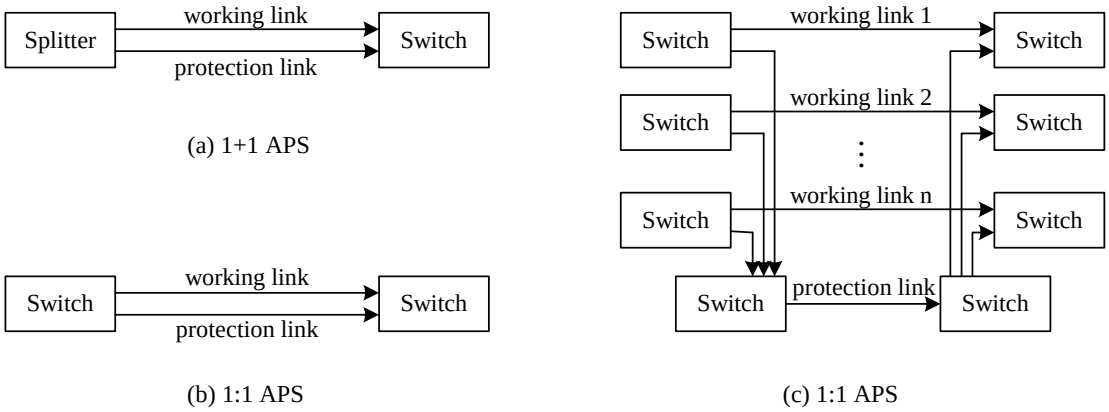


Figure 3-4: Automatic Protection Switching

3.4 Resilience Provisioning Mechanisms in Ring Topology

3.4.1 SONET/SDH Rings

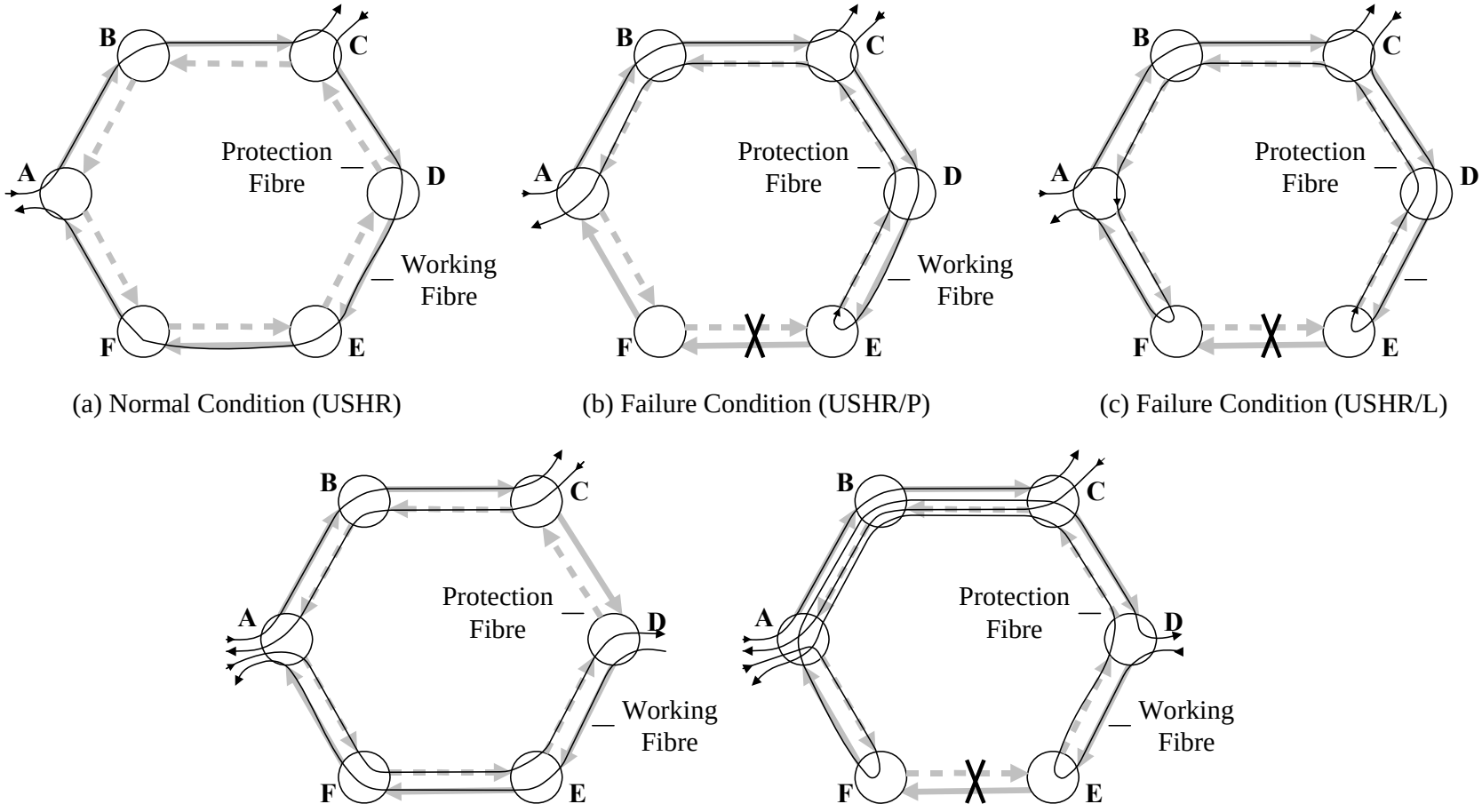
SONET SHR is more flexible than APS in that it can handle both link and node failures. SONET rings can be classified as **unidirectional** and **bidirectional** rings based on the routing principle during normal network conditions.

The self-healing protocol of a SONET rings can be implemented using overheads at the SONET line or path layer [ITUG872]. An SHR is a path switched self-healing ring if its protection switching is triggered by the SONET path layer signal. In contrast, an SHR is called a line switched SHR if its protection switching is triggered by the SONET line layer signal. Only unidirectional rings with path protection switching (called UPSR or USHR/P) and bidirectional rings with line protection switching (called BLSR or BSHR/L) are specified by ANSI and Bellcore requirements and commercially available [WU97].

In the unidirectional SHR, all traffic is routed in the same direction along one fibre called the working fibre. The second fibre is set aside as spare to protect the working fibre. Whenever a failure is detected, the system will automatically switch the affected traffic from the working fibre onto the protection one as illustrated in Figure 3-5 (a-c). In unidirectional SHR, every duplex connection will travel the whole circumference and the maximum link flow on the links of the ring will always be the sum of all connections passing through the

ring. Therefore the capacity required for both the working and protection fibres will be the sum of all connections carried by the ring. This type of ring is sometimes referred to as a dedicated protection SHR since for each demand there is a corresponding amount of space capacity set aside specifically to protect it [BLA02].

In the bidirectional SHR, every duplex connection travels through the same physical routing path as shown in Figure 3-5 (d-e). Bidirectional SHRs are further divided into two classes: 2-fibre BLSR (BLSR/2) and 4-fibre BLSR (BLSR/4), depending on spare capacity provisioning. For 4-fibre BLSR, two fibres serve as standby fibres that provide 1:1 protection. The 2-fibre BLSR uses only half the capacity of the fibre system for working traffic and reserves the other half as protection capacity. In the case of a failure, the affected traffic will be switched to the spare capacity as illustrated in Figure 3-5 (d-e).



(a) Normal Condition (USHR) (b) Failure Condition (USHR/P) (c) Failure Condition (USHR/L)

(d) Normal Condition (BSHR) (e) Failure Condition (BSHR)

Figure 3-5: Self Healing Rings

3.4.2 WDM Rings

SONET/SDH APS and Rings still work and restore the traffic in the electrical domain. With the development of WDM technology, the transport network is evolving towards an all-optical network. Thus the resilience provisioning purely in the optical domain is gaining more and more attention as an efficient method for protecting traffic.

Resilient schemes employed in WDM networks are often referred to as optical layer protection [GER00] in contrast to SONET protection, which is working in transport layer. Protection schemes similar to that in SONET point-to-point and ring-based network are adopted in the WDM point-to-point and ring-based network.

WDM protection operates either at the optical channel (OCh) section level or optical multiplex section (OMS) level. The main difference between OCh and OMS protection is represented by the granularity at which the layers operate. OCh protection works on individual lightpaths, thus allowing selective recovery of optical line terminal (OLT) failures. OMS protection works at the aggregate signal level, thus recovering all lightpaths present on the failed line concurrently. OCh and OMS are also referred to as Path layer and Line layer, respectively [FUM00][GER00][GER00a].

Classification of WDM protection schemes is shown in Table 3-1.

Network Topology	Line layer (OMS)		Path layer (OCh)	
	Dedicated	Shared	Dedicated	Shared
Point to point (Linear)	1+1 APS (OMS-DP)	1:1/1:N APS (OMS-SP)	1+1 APS (OCh-DP)	1:1/1:N APS (OCh-SP)
Ring	OULSR (OMS-DPRing)	OBLSR (OMS-SPRing)	OUPSR (OCh-DPRing)	OBPSR (OCh-SPRing)

Table 3-1: WDM APS and Rings

3.5 Resilience Provisioning Mechanisms in Mesh Topology

Optical networks are inherently mesh-based. As APS and SHR are well established and robust, resilience in traditional optical networks is usually provided by dimensioning networks into a ring-and-linear topology, using ring and linear-based protection schemes [BLA02] [GOR01]. The dimensioning of these networks into rings is a very complex task. In addition, it is not a cost-efficient solution to provide network resilience using ring-based

schemes since more than 100% backup resource is needed, in real networks sometimes more than 200% spare resource is needed [GRO00].

Mesh-based resilience provision schemes that operate on the entire network can be significantly more cost-efficient than ring-based schemes. For mesh-based resilience provisioning mechanism, the required spare to working resources ratio can typically be in the range of only 50%-70% for well-connected physical network graphs [GRO00][HER94] [IRA98], which make it very attractive to the network designer. At the same time, the rapid growth of Internet and e-commerce require the optical to be dynamically reconfigurable, which is made possible by the new development of WDM technology. The mesh network, because of its high connectivity, is well suitable for the dynamic resilience provisioning.

Resilience provisioning mechanisms in mesh networks could be classified according to different criteria:

- The use of pre-planned versus dynamic computed / discovered routes [WU97][DOV01];
- The use of link rerouting versus path rerouting [RAM99a][WU97];
- The use of centralized computation versus distributed computation / searching [WU97][GRO91][GEO99][MED99];
- The requirement of a database containing global network state (GNS) versus not required [CHE98][GEO99];
- The use of dedicated protection versus shared protection [MOH00][RAM99a] [WU97];
- The use of dimensioning the network into protection domains versus protecting the mesh network as a whole [GER00];
- The use of static provisioning versus dynamic provisioning [SEN01][ZAN01] [ASS01][YE01][RAMR01];

Figure 3-6 shows a classification framework of the resilience provisioning mechanisms in mesh networks.

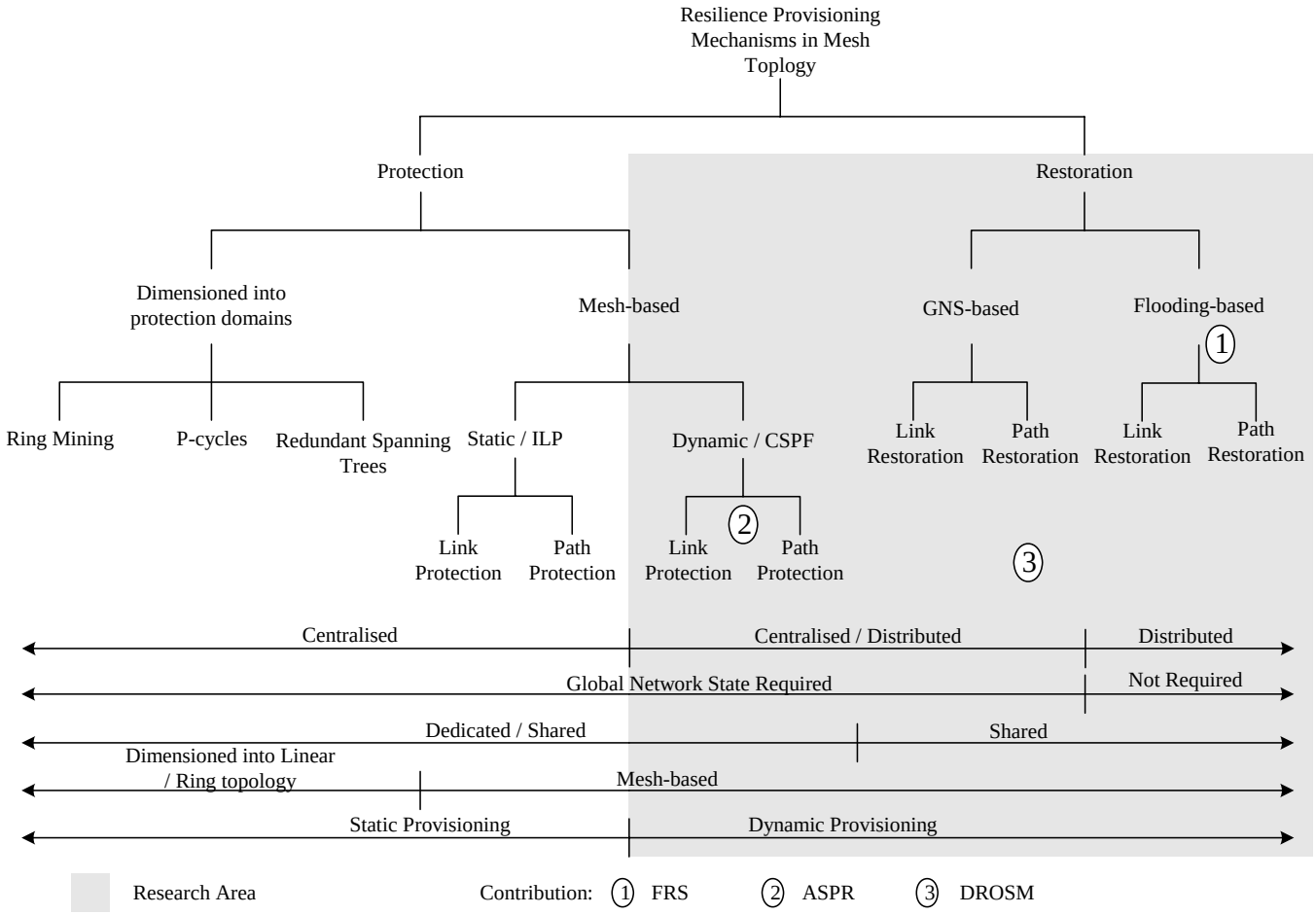


Figure 3-6: Resilience Provisioning Mechanisms in Mesh Topology

3.5.1 Basic Schemes

When provisioning resilience in mesh networks, some basic rules are followed. This section introduces these basic terms and rules.

3.5.1.1 Link-Disjoint versus Node-Disjoint Paths

In a mesh network work, the working path and backup path are either link-disjoint or node-disjoint [MOR00][RAM99a] [ANA00][DAC02].

Link-disjoint paths are those that do not share a single link along their routes. For example in Figure 3-7, path p1 and p2 are link-disjoint. As they do not share a common link and thus any single link failure will not affect both of them, p1 and p2 could serve as backup path of each other. However, since they are only link-disjoint, they may cross the same node in their routes. For example in Figure 3-7, both p1 and p2 travel through node D. Therefore, a node failure such as node D could let both p1 and p2 fail at the same time. A link-disjoint path can only protect against link failures. A link-disjoint path is also called edge-disjoint or arc-disjoint path in some researches [MAN02][GU96].

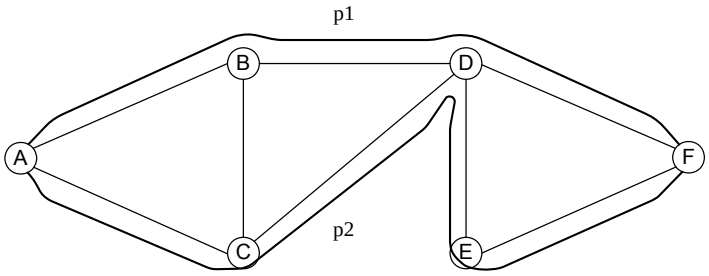


Figure 3-7: Link-Disjoint Paths

A more strict restriction is node-disjoint. The paths that do not share a single node except the source and destination node are called node-disjoint paths. Node-disjoint paths are also link-disjoint, as shown in Figure 3-8. Therefore, a node-disjoint path of a working path could serve to protect against both node and link failure. However, a node-disjoint path sometimes does not exist in a real network. In this case, link-disjoint path has to be used as the backup path. A node-disjoint path sometimes is also called a vertex-disjoint path [CHEC95][LAB92].

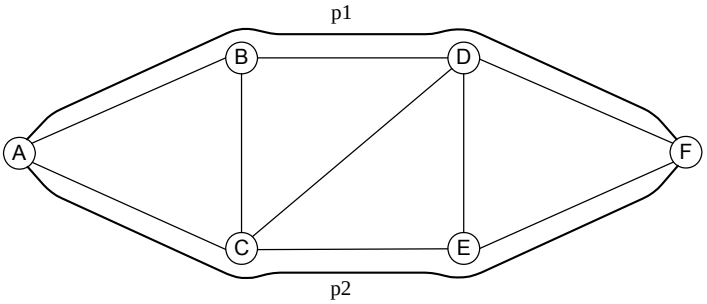


Figure 3-8: Node-Disjoint Paths

3.5.1.2 Link vs. Path Protection / Restoration

In a mesh network, protection and restoration could be either link-based or path-based.

Link protection / restoration employs local rerouting to cover a particular link. It reroutes traffic around the failed component. When a link fails, a new path is selected between the end nodes of the failed link, which is illustrated in Figure 3-9. Link restoration has an advantage of being able to restore traffic in a very short time since the rerouting of the traffic is close to the failure. However, it requires setting aside significant spare resources for the backup path [RAM99a][MOH00].

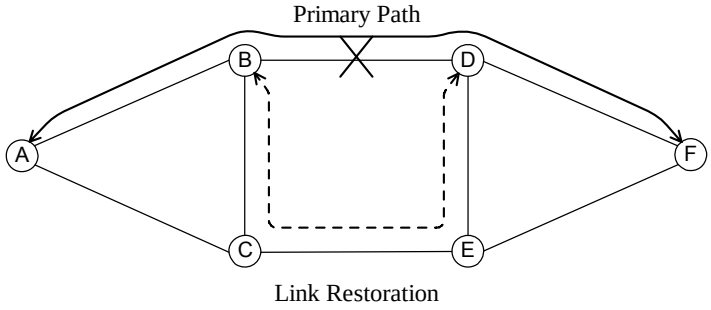


Figure 3-9: Link protection / restoration

Path protection / restoration uses end-to-end rerouting to cover the whole path. In path restoration, a backup path is established between the end nodes of the primary path. Path restoration has better performance on resource sharing thus requires less spare resource than link restoration. However, signalling is needed to notify the ingress OXC to switch over to the backup path, which results in a longer restoration time than link restoration [MOH00][RAM99a].

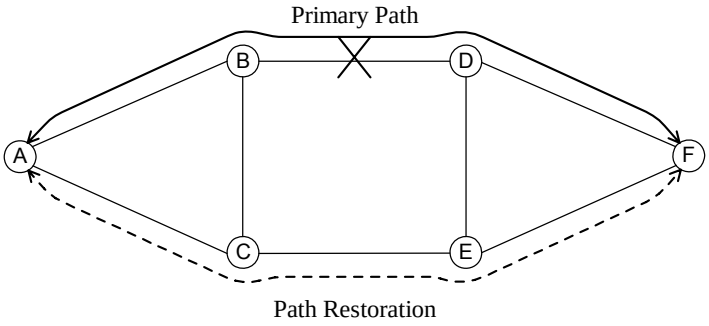


Figure 3-10: Path Protection / Restoration

3.5.2 Static Resilience Provisioning

Static resilience provisioning mechanisms assume that at the stage of network design traffic demands in the network are already known and will remain constant over time. Therefore, the design of how working and backup paths are deployed in the network could use optimisation algorithms to find a suitable solution. This procedure usually takes place at the planning stage of a network. The design of working and backup paths amounts to a multi-commodity flow (MCF) problem [LIU01], which is computationally complex, or non-deterministic polynomial (NP) hard. It can be solved by mathematical programming techniques.

3.5.2.1 Integer Linear Programming (ILP)

Linear and Integer programming are widely used mathematical techniques that are concerned with optimisation, which is with finding the best possible answer to a problem. The problem is usually formulated as a particular function to be maximized or minimized subject to several constraints. Linear Programming (LP) concerns optimising a linear function subject to linear side constraints. When in addition the variables are only allowed to take integer values, it is called Integer Programming (IP) or Integer Linear Programming (ILP) [SCH98]. Solving an ILP problem is NP-hard.

Resilience provisioning in the optical network is an optimisation problem, of which the objective is to minimise the network cost. ILP is usually used to formulate such a problem and to find a solution using some commercially available software tools, such as CPLEX. Unfortunately, the solving of ILP formulation is NP-hard and cannot produce a real-time solution [LIU01]. In addition, due to the rapid increase of the size of the path set, which forms the variables of the formulation, with the network size, the model will not scale for many

realistic situations. Therefore, the ILP formulations are practical only for small networks [RAM99a]. For larger networks, heuristic methods are needed to solve the problem [CAE98][SHY99][GRO99].

As it is time-consuming to find the solution, optimisation algorithms based on ILP formulation only apply to a one-time static design of resilience provisioning, which does not hold a particular time requirement. For dynamic provisioning, where the solutions are required to respond to dynamic changing of network topology and traffic patterns, solving ILP problem may not be practical.

Static resilience provisioning can be achieved either by dimensioning the network into protection domains with linear or ring topology or by protecting the mesh network as a whole (Figure 3-6). In both cases, ILP formulation is usually used to help produce a best possible result.

3.5.2.2 Ring Mining / Dimensioning

As the protection provided by APS and SHR is robust and easy to manage, mesh topologies could be dimensioned into linear or ring topologies and using these mechanisms. Furthermore given that SHR has many advantages over APS, a lot of studies [CAE98][SOR98][SEM94] [WAS94] are pursuing using SHR to cover mesh networks.

Planning and dimensioning self-healing rings in optical networks is a complex matter. The generic dimensioning of self-healing rings could be stated as follows:

Given a set of demand nodes N, a two-connected (meshed) network $G = (N, E)$ connecting these nodes via a set of edges (fibre, optical links) E, an O-D demand matrix $\boldsymbol{D} = (d_k)$, $k = (i, j)$ where d_k is the traffic demand between the origin $i = o(k)$ and destination $j = d(k)$ nodes of commodity k, with $i, j \in N$, one wants to determine a set of feasible SHRs that protects all demands at minimal cost.

This problem is apparently a very difficult one since whenever a set of nodes need to be connected to form a SHR, finding the best cycle that passes exactly one through each of them corresponds to solving an instance of the classical travelling salesman problem. If one temporarily ignores this matter of designing how to physically realise the different rings making up the network and only concentrates on the logical design, i.e., determining which nodes should be connected by which SHR and how each individual demand flowing on the network is to be protected, then the problem can be formulated as a large mixed ILP problem.

For example, in [SOR97][SOR98] the following formulation is used to model the design of a network using multiple interconnected unidirectional SHRs.

Letting $x_{ir} = 1$ if an ADM is placed at node i to connect it to ring r and 0 otherwise, $y_r = 1$ if ring r is used and 0 otherwise, v_{ir}^k (and v_{rj}^k) be the flow variables representing the amount of traffic of commodity k that accesses or leaves ring r at node i (from node j), and w_{rs}^{kl} the amount of inter-ring traffic of commodity k that passes from ring r to ring s at interconnection node l, one can write:

Min
$$\sum_{r\in R} c_r \sum_{i\in N} x_{ir} + f \sum_{r\in R} \sum_{s\in R, s\neq r} \sum_{k\in K} \sum_{l\in T} w_{rs}^{kl} \tag{1}$$

s.t.
$$\sum_{r\in R} v_{ir}^k = d^k \quad \forall k \in K, i = o(k) \tag{2}$$

$$\sum_{r\in R} v_{rj}^k = d^k \quad \forall k \in K, j = d(k) \tag{3}$$

$$v_{ir}^k + \sum_{s\in R, s\neq r} \sum_{l\in T\cap r\cap s} w_{sr}^{kl} = v_{rj}^k + \sum_{s\in R, s\neq r} \sum_{l\in T\cap r\cap s} w_{rs}^{kl} \quad \forall r \in R, \forall k \in K, i = o(k), j = d(k) \tag{4}$$

$$\sum_{k\in K} (v_{ir}^k + \sum_{s\in R, s\neq r} \sum_{l\in T\cap r\cap s} w_{sr}^{kl}) \leq u_r \quad \forall r \in R \tag{5}$$

$$\sum_{k\in K: o(k)=i} v_{ir}^k + \sum_{k\in K: d(k)=i} v_{ri}^k \leq \sum_{k\in K: o(k)=i \text{ or } d(k)=i} d^k x_{ir} \quad \forall i \in N, \forall r \in R \tag{6}$$

$$\sum_{i\in T} x_{ir} \geq 2y_r \quad \forall r \in R \tag{7}$$

$$\sum_{i\in N} x_{ir} \geq 16y_r \quad \forall r \in R \tag{8}$$

$$x_{ir}, y_r \in \{0,1\} \quad \forall r \in R \tag{9}$$

$$v_{ir}^k, v_{rj}^k, w_{rs}^{kl} \geq 0 \quad \forall i, j \in N, \forall r, s \in R (r \neq s), \forall k \in K, \text{and } \forall l \in T \cap r \cap s \tag{10}$$

where c_r represents the cost of an ADM for ring r, f the inter-ring traffic unit cost, d_k is the traffic demand for O-D pair $k = (i, j)$ as defined above, u_r is the capacity of ring r, and N,

T, *R*, *K* are respectively the sets of demand nodes, inter-ring transfer nodes, possible SHRs and commodities having to be protected. The objective (1) consists in minimising the sum of ADM and inter-ring traffic costs. Constraints (2-4) are demand satisfaction and flow conservation constraints, constraints (5) is the capacity constraint for the SHRs, and finally, constraints (7-9) are design constraints, ensuring respectively that no demand access or leave a given SHR at an origin or destination node if that node is not connected to that particular SHR, that each ring have at least two interconnection nodes, and that there be no more than a pre-determined number, here 16, of different nodes connected to any given ring.

This planning problem has been approached from various angles giving rise to quite different design problems depending on the simplifying assumptions considered and what characteristics the resulting network should have, for example the optimal placement of ADM and the how the rings are interconnected [SHI].

3.5.2.3 P-Cycles

A more efficient ring-like approach, termed P-cycles, has been proposed by Grover and Stamatelakis [GRO98][STA00], where the ring protects not only connections that are part of it, but also chords that run between ring nodes. An example of P-cycles is illustrated in Figure 3-11. Here, a protection cycle A-C-E-F-D-B-A is used to protect all the possible link failures. This approach was shown to be much more capacity efficient than ring dimension [GRO98]. However, the optimal design also results in an ILP problem, which makes it have the same problems as the ring-dimensioning method and only be suitable for static resilience provisioning. More information on P-cycle and its ILP formulations can be found in [GRO00][GRO98][SCHD02].

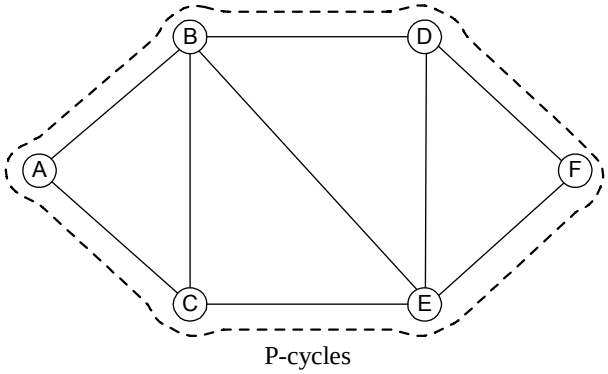


Figure 3-11: P-Cycles

3.5.2.4 Redundant Spanning Trees

Another application of dimensioning mesh network into protection topology is utilizing redundant spanning trees. Medard et al. [MED99][FIN97][GAL98] propose to use a pair of redundant spanning trees for pre-planned restoration in the presence of link or node failures. It involves topology construction in edge-redundant graphs and vertex-redundant graphs. For edge-redundant graphs, they propose an algorithm that constructs two directed trees rooted at the source vertex. One of them, the blue tree, is used as the working tree. The other, the red tree, is used for traffic protection. When a single link fails, every vertex in the graph can still be connected to the source vertex via either the red tree or the blue tree. For vertex-redundant graphs, they propose an algorithm that constructs two directed trees rooted at the source vertex. One of them, the blue tree, is used as the working tree. The other, the red tree, is used for traffic protection. When a single vertex (other than the source vertex) fails, every other vertex in the graph can still be connected to the source vertex via either the red tree or the blue tree.

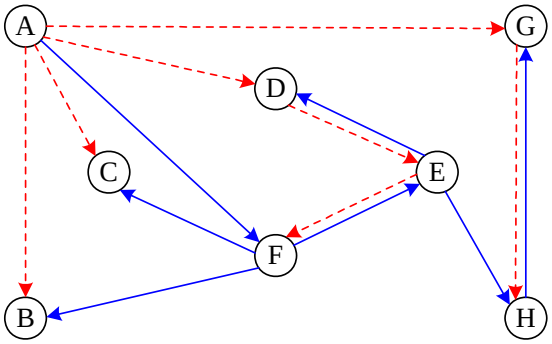


Figure 3-12: Redundant Spanning Trees

Here is an example given by [XUE02]. Figure 3-12 illustrates two directed trees rooted at the source vertex A, spanning all other vertices in the network. The tree with solid arcs is blue tree and the tree with dashed arcs is the red tree. The blue tree is the working tree and the red tree is the backup tree. If there is no link or vertex failure, each vertex in the network is connected to the source vertex via the blue tree.

When a single vertex failure happens, say at vertex G, the vertices F and H are no longer connected to A via the (broken) blue tree in Figure 3-12. However, they are connected to the source vertex A via the (also broken) red tree. For example, vertex F is connected to vertex A via (F, H) and (H, A).

Redundant spanning trees algorithms are particularly well suited to multicast networks and optical networks, where trees may be created by signal splitting [GAL98]. However, the

problem of finding minimum cost trees for a certain source and a set of destinations is the Steiner tree problem [WIN], which is NP-hard. Good surveys of the problem and heuristics can be found in [WIN92][MAK92][KAM02] and applications to networks can be found in [KAH92][MANI00][BAU97][BHAK83]. As the Steiner tree problem is NP-hard, the issue of cost minimization and capacity optimisation may be difficult.

3.5.2.5 Protect Mesh Network as a Whole

Static resilience provisioning in the mesh network could be taken by protecting the mesh network as a whole. These resilience-provisioning schemes are also called mesh-based resilient schemes. Mesh-based resilient schemes that operate on the entire network can be significantly more cost efficient than dimensioned ring-based schemes [IRA98][RAM99a]. This is built on the fact that backup paths have a wider range to share the spare resource in the network. Efficiency improvements ranging from 20% to 60% have been demonstrated in different application situations [GER00].

When protecting the mesh network as a whole using static resilience provisioning mechanism, the problem is to decide the optimal placement of the working and backup path for the traffic demands that are already known. This problem can also be expressed as spare capacity allocation (SCA) [LIU01][XIO99] and flow assignment. That is to determine where to place spare capacity in the network and how much spare capacity must be allocated to guarantee seamless communication against a set of failure scenarios.

SCA is an important part of the resilient network design. This problem has been investigated in different mesh-based networks such as ATM [XIO99], SONET/SDH [HER95] [IRA98][HERM97][ALR00][GRO99], WDM [RAM99a][CAE98], and IP/MPLS [OH00] network. In these researches, multi-commodity flow (MCF) models which results in an ILP formulation [LIU01] have been widely used to formulate the problem. In these models, the working and backup path of all traffic demand is pre-calculated (protection) to compose the search space for the design variables and the purpose is to minimise the total spare capacity required for the restoration from specific failure. The backup path resource is usually shared to reduce the spare capacity requirement.

The modelling of SCA results in an ILP formulation [LIU01]. The models can be further classified either for link protection, or for path protection (Figure 3-6). For example, in [XIO97], the following formulations are adopted to model network resilience provisioning using mesh-based protection for path protection and link protection, respectively.

Consider a network *G (N, L)* which has |*N*| nodes and |*L*| links, where |*N*| is the cardinality of set *N*. Each of its components (links and nodes) has two states: normal state 0 and failure state 1. The network state is completely described by a vector of these component states.

N: the set of nodes of the network;

A: the set of directed arcs of the network;

L: the set of links of the network, each link $l \in L$ is composed of two arcs (*a* and $a' \in A$) which have the same end nodes as *l* but with opposite directions;

$\prod$: the set of origin-destination node pairs (commodities);

S: a set of states of the network;

R_π^s: a set of candidate routes for commodity $\pi \in \prod$ when the network is in state $s \in S$;

γ_π^s : the traffic demand expressing minimal bandwidth requirement for commodity $\pi \in \prod$ when the network is in state $s \in S$;

$x_{r\pi}^s$: the bandwidth used by commodity π on route *r* when the network is in state s, $r \in R_\pi^s$;

δ_{rw}: the delta function which equals 1 when network component *w* is on route *r* and 0 otherwise;

F(*s*): a set of failed components when the network is in state $s \in S$;

c_a: capacity used on arc $a \in A$;

d_a: length of arc $a \in A$.

For path protection, the formulations could be for either global reconfiguration or failure-oriented reconfiguration.

For global reconfiguration, in case of failure, all traffic flows, affected and unaffected, are reconfigured so that the restoration ratio is guaranteed but the total network cost is

minimised. The resilient network design problem can thus be formulated as the following linear programming (LP) problem:

Min $\sum_{a \in A} d_a c_a$

s.t.

$$\sum_{r \in R_\pi^s} x_{r\pi}^s = \gamma_\pi^s, \quad \pi \in \prod, s \in S, \tag{11}$$

$$c_a \geq \sum_{\pi \in \prod} \sum_{r \in R_\pi^s} \delta_{ra} x_{r\pi}^s, \qquad a \in A, s \in S \tag{12}$$

where $x_{r\pi}^s \geq 0, r \in R_\pi^s, \pi \in \prod, s \in S$. In the above formulation, constraints (11) guarantee that the traffic demand γ_π^s between each node pair is satisfied in every possible network state. Constraints (12) ensure that the capacity assigned to arc *a* is large enough to accommodate the traffic flows on arc a for all possible network states.

In failure-oriented reconfiguration, only affected traffic flows are rerouted at failure events. Set s_0 as the normal operating state of the network. It can be formulated as the following LP problem:

Min $\sum_{a \in A} d_a (\sum_{\pi \in \prod} \sum_{r \in R_\pi^s} \delta_{ra} x_{r\pi}^{s_0} + c_a^{spare})$

s.t.

$$\sum_{r \in R_\pi^{s0}} x_{r\pi}^{s_0} = \gamma_\pi^{s_0}, \quad \pi \in \prod \tag{13}$$

$$c_a^{spare} \geq \sum_{\pi \in \prod} \sum_{r \in R_\pi^s} \delta_{ra} y_{r\pi}^s, \quad a \in A, s \in S - s_0 \tag{14}$$

$$\sum_{r \in R_\pi^s} y_{r\pi}^s - \sum_{r \in R_\pi^{s0}} U(\xi) x_{r\pi}^{s_0} = 0, \quad \pi \in \prod, s \in S - s_0 \tag{15}$$

where $x_{r\pi}^{s_0} \geq 0, y_{r\pi}^s \geq 0, \xi = \sum_{w \in F(s)} \delta_{rw}$ and $U(\xi) = 1$ if $\xi \geq 1$ and 0 otherwise. Here, $y_{r\pi}^s$ instead of $x_{r\pi}^s$ is used to denote a restoration flow on route *r* for the affected commodity π when the network is in state *s*, $r \in R_\pi^s, \pi \in \prod, s \in S - s_0$. The first term in the objective

function is the capacity required on arc *a* to carry traffic in normal network operating state s_0. The second term c_a^{spare} represents the additional (spare) capacity needed to restore affected traffic in case of failure. Constraints (15) ensure that for each commodity π, its affected traffic flows are completely restored in every possible failure scenario.

For link protection, the two nodes directly connected by the failure link will be responsible for the restoration of traffic flows on that link. It can be formulated as the following LP problem.

Min $$\sum_{a \in A} d_a \left(\sum_{\pi \in \prod} \sum_{r \in R_\pi^{s0}} \delta_{ra} x_{r\pi}^{s_0} + c_a^{spare} \right)$$

s.t.

$$\sum_{r \in R_\pi^{s0}} x_{r\pi}^{s_0} = \gamma_\pi^{s_0}, \quad \pi \in \prod \tag{16}$$

$$c_a^{spare} \geq \sum_{b \in F(s)} \sum_{r \in R_b^s} \delta_{ra} y_{rb}^s, \quad a \in A, s \in S - s_0 \tag{17}$$

$$\sum_{r \in R_b^s} y_{rb}^s - \sum_{\pi \in \prod} \sum_{r \in R_\pi^{s0}} \delta_{rb} x_{r\pi}^{s_0} = 0, \quad b \in F(s), s \in S - s_0 \tag{18}$$

where $x_{r\pi}^{s_0} \geq 0, y_{rb}^s \geq 0$. Here the traffic flows on each arc $b \in A$ are regarded as one commodity. R_b^s is now the set of candidate routes for restoring traffic flows on arc *b* when it fails in network state $s \in S$, $F(s)$ is the set of failed arcs in network state *s*.

These formulations in [XIO99] are derived for ATM networks and also apply to WDM optical networks assuming that all the OXCs have wavelength conversion capabilities. In [RAM99a], ILP formulations are developed for the WDM optical network without wavelength conversion capabilities. These formulations include those for dedicated path protection, shared path protection, and shared link protection. These ILP solutions determine the routing and wavelength assignment of the working and backup path with an objective of minimising the total number of wavelengths used on all the links in the network.

A comparative study on the performance of different SCA schemes can be found in [LIU01].

3.5.2.6 Strength and Weakness of Static Provisioning

Static resilience provisioning is based on the assumption that traffic demands between each node in the network are already known and constant over time. Thus, based on this information, the network design and resilience provisioning could be tackled by formulating the problem into mathematical models, usually ILP models. As it has the knowledge of the entire set of traffic demands (as oppose to online algorithms within a dynamic environment that routes are chosen without awareness of future demands), the static resilience provisioning makes more efficient use of network capacity and projects a lower capacity requirement. As a result, this method is widely used in design and resilience provisioning of traditional optical networks.

However, optimisation algorithms based on ILP modelling of this problem prove to be NP-hard. It takes considerable time and efforts to calculate the result, which makes it unable to provide a real-time solution. It is impractical to use such a method to deal with a network with a fast changing state. Especially, this method only suits small networks. For larger networks, with the increase of network size and traffic matrix size, the above models are unable to scale to a realistic solution.

3.5.3 Dynamic Resilience Provisioning

Traditional static provisioning mechanisms cannot provide real-time solutions to a dynamic changing optical network as it assumes traffic demands are available at the network design stage and will remain constant in the network. In addition, the complicated nature of these optimisation algorithms, requiring considerable processing time, also limits their application in a dynamic environment.

In a dynamic environment, connection requests typically arrive one by one without knowledge of future demands. In addition, these connections may have a finite lifetime of varying duration that introduces yet more uncertainties.

Without the exact information of all traffic, one has to base the decision for resilience provisioning separately for each connection. The strategy is to find an optimal deployment for the newly coming connection request based on the up-to-date network situation. The lack of consideration for other connections, leads to results that cannot guarantee a globally optimal solution.

3.5.3.1 Shortest Path First (SPF) Algorithm

As a long period is required for the execution of complicated optimisation procedures, offline static resilience provisioning algorithms no longer apply to a dynamic environment. In this situation, traffic demands are no longer predictable and constant over time; connection requests arrive without knowledge of future demands and would remain a finite time in the network. Thus simplified algorithms are required to provide a real-time solution.

This simplification is usually made by considering a minimum cost solution for each single connection in a distributed manner. In this case, the Shortest Path First (SPF) algorithm and its constraint-based version (CSPF) are usually used. An efficient algorithm for finding the shortest path between two nodes in the network is Dijkstra's algorithm [BHA99].

Let $d(i)$ denote the distance of node i $(i \in V)$ from source node A; it is the sum of the cost of links in a possible path from node A to node i. Let $P(i)$ denote the predecessor of node i on the same path. Note that $d(A)=0$. The following steps result in the determination of the shortest path from A to Z:

Step 1. Start with $d(A)=0$,

$d(i) = L(Ai)$, if $i \in \Gamma_A$,

$= \infty$, otherwise;

Γ_i set of neighbor nodes of node i, $L(ij)$ = length of link from node i to node j.

Assign $S=V-\{A\}$, where V is the set of nodes in the given graph.

Assign $P(i)= A \ \forall i \in S$.

Step 2. a) Find $j \in S$ such that $d(j)=min\ d(i)$, $i \in S$.

b) Set $S=S-\{j\}$.

c) If $j=Z$ (the destination node), END;

otherwise, go to Step 3.

Step 3. $\forall i \in \Gamma_j$ and $i \in S$, if $d(j)+L(ji)<d(i)$, set $d(i)=d(j)+L(ji)$, $P(i)=j$.

Go to Step 2.

Dijkstra's algorithm only deals with non-negative weight edges. The Bellman-Ford algorithm [BHA99] solves the single-source shortest-paths problem in the more general case in which edge weight can be negative. Based on Bellman-Ford algorithm, constraint-based algorithms could be derived by transforming the network graph [BHA99].

Observe that Dijkstra's algorithm computes the shortest path with respect to link weight for a single connection at a time. The result can be very different to the paths that would be

selected when a batch of connections between a set of endpoints is requested for a given optimising objective [LIUK02]. Due to the complexity of some of the routing algorithms (high dimensionality, and integer programming problems, for example) and various criteria by which one may optimise the network, it may not be possible or efficient to run a full set of these versatile routing algorithms in a distributed fashion on every network node. Therefore, it is desirable to have such a basic form of path computation capability running on the network nodes for a dynamic situation.

As the resilience provisioning involves the decision of a pair of working and backup paths, dynamic resilience provisioning is a routing and wavelength assignment (RWA) problem with certain constraints [YE00][ZAN01]. It contains the routing and wavelength assignment for both the working and backup path. This pair of paths must be physically diverse without a common failure element. They should be either link-disjoint paths or node-disjoint paths.

The routing and wavelength assignment of both working and backup paths in a dynamic environment involves a path calculation using the Constraint-based Shortest Path First (CSPF) algorithm. The constraints include wavelength conversion capability of the OXC and physical diversity requirement of the pair of paths. The calculation of the backup path could occur either after the working path has been decided or at the same time as the working path calculation.

3.5.3.2 Centralised Scenario

The selection of working path and backup path at the time of resilience provisioning requires detailed knowledge of the global network state. These network state properties include network topology, bandwidth usage, available bandwidth of each link, and detailed deployment result of existing traffic if efficient backup path sharing is required. Collection of these properties and calculation of the working and backup path could be performed either in a centralised or a distributed manner.

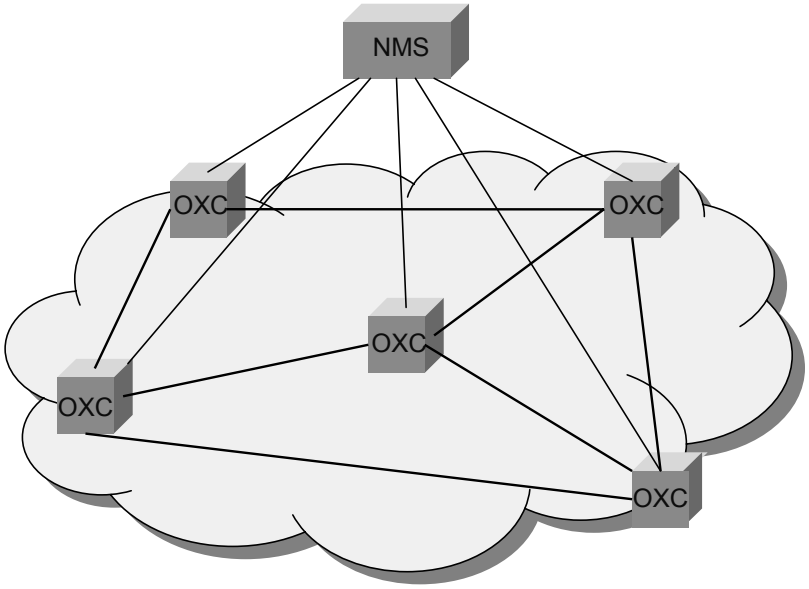


Figure 3-13: Centralised Scenario

In the centralised scenario, the detailed global network state information is collected and maintained by a network management system (NMS) (Figure 3-13). This is done through communication between the NMS and each node in the network. When there is a connection request, the NMS is responsible for the path selection and the connection setup.

The merit of the centralised mode is that nodes in the network do not require complicated processing capabilities because the maintenance of network state information and path calculation are performed in the NMS. The network structure with a centralised NMS suits for static resilience provisioning. It may also apply to a dynamic environment if in a small network. For static resilience provisioning, the NMS uses off-line optimisation algorithms (such as ILP) to calculate the working and backup path. For dynamic resilience provisioning, the NMS uses simplified algorithms (such as CSPF) to provide real-time solutions.

However, setting up and maintaining a NMS is costly. Each node in the network needs a connection with the centralised server. These connections need to be extremely reliable since they form the control plane of the network. Additional strategies are needed to protect the centralised server and its communication with each node in the network. Such communication also introduces certain latency. In addition, it requires the NMS have a high processing capability when connection requests are quite frequent. It is almost impossible when it is applied to a network with a large scope.

In contrast, it could be argued that a distributed solution should be followed for the dynamic resilience provisioning, especially in the IP-centric optical network infrastructure.

3.5.3.3 Distributed Scenario

The distributed solution eliminates the need for a centralised controller to manage reconfiguration. In the distributed scenario, every node in the network maintains a database about the network state. The path calculation usually takes place at the ingress OXC, using information contained in its local database. The local database is maintained and synchronised using a link state routing protocol, such as extended OSPF and IS-IS. Through the link state routing protocol, the information on each link is flooded to every node in the network. Therefore, each node will have the global network state. That the global network state is distributed to each node makes the whole system more reliable since it need not take into account of the possibility of collapse of the central server.

However, the distributed scenario does not provide a globally optimal solution to the problem. Another disadvantage of the distributed scenario is the database accuracy, decided by a trade-off between information detail and database re-convergence time. The database of each node needs to be as accurate as possible, thus prefers more information being flooded in the network. Nevertheless, more information being flooded means more time needed for the database to re-converge. Therefore, the database of each node cannot be updated in time, which means inaccuracy. The solution is to flood only aggregated information about each link, which makes the database of each node being updated in time by sacrificing some details of the link information. How to choose the aggregated information attracted a lot of research [SRI02][LI02].

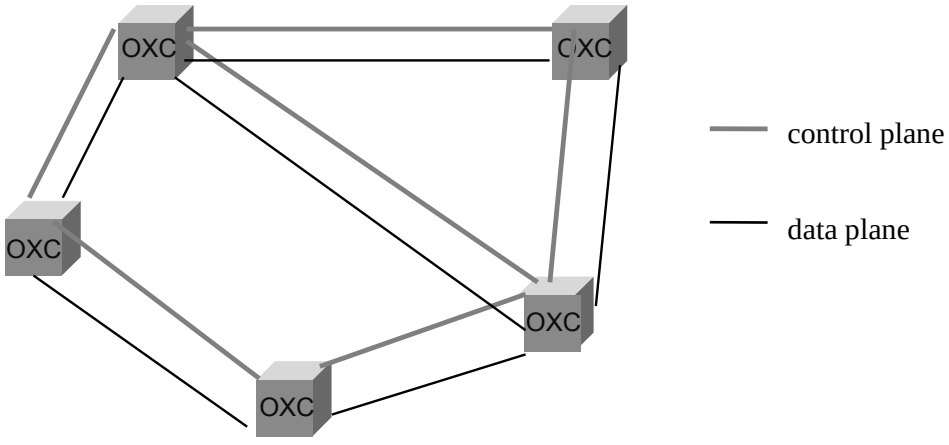


Figure 3-14: Distributed Scenario

For example, in the case of mesh-based shared protection, the detailed sharing information of each backup path is maintained locally and only aggregated information (e.g. total reserved/shared bandwidth) is flooded to other nodes in the network. Therefore, when path calculation occurs at the ingress OXC, it has no idea of the local sharing database of

other OXCs. It is therefore unable to determine backup sharability of each link for any given primary path. The backup path has to be calculated according to the aggregated information available in the database. Sharing of any link on the backup path is determined during signalling of the backup path. Information about the primary path is also carried in the signalling message. When it receives the signalling message, each OXC on the backup path decides whether the proposed backup path can reuse the resource already reserved by other backup paths.

3.5.3.4 GNS-Based versus Flooding-Based Restoration

As reactive restoration does not require pre-planned backup paths and it searches for alternative restoration paths only after the failure has occurred, it is more flexible in utilising the spare resource. Therefore it is more suitable for networks with dynamic changing traffic.

Search for restoration paths after a failure has occurred can be carried out using two different methods: in this thesis, one is called Global Network State (GNS)-based and the other is called flooding-based.

The first is by maintaining a database, which contains the global network state, either in a centralised or a distributed mode. For the centralised mode, a NMS is needed to collect the network state properties from each node in the network. For the distributed mode, each node of the network has a database maintaining aggregated information about the global network state. In this case, a link state routing protocol such as OSPF or IS-IS is needed to distribute the network state properties to each node.

By maintaining a database of the global network state, the NMS (centralise mode) or a network node (distributed mode) calculates restoration paths according to the current network state. The calculation algorithms that are based on SPF algorithm are conceptually simple and easy to implement.

However, it has several disadvantages. The cost of maintaining a network state database is very high. For the centralised mode, reliable connections are needed between the NMS and all nodes in the network. Its restoration is relatively slower. For the distributed mode, the database has only aggregated information of the link state of the network and cannot update in time immediately after the failure occurs. As a result, the path calculation algorithm may fail to find an alternative path while spare resource is still available.

In contrast, flooding-based restoration [GRO91][YAN88] does not require nodes in the network to maintain a global network state database; it does not need a NMS or a link state

protocol and thus is cheaper to implement. It finds restoration paths using message flooding. Flooding-based restoration may have a better restoration ratio than GNS-based restoration, since in the latter case, the database used by alternative path calculation cannot update in time immediately after the failure occurs.

The drawbacks of flooding-based restoration include: Firstly, the restoration time is longer than that of GNS-based restoration since the search for alternative paths is performed by flooding messages, which takes extra time in message propagation and processing. Secondly, the flooding messages make the communication overhead excessively high. In this case, a hop count is usually introduced to limit the flooding area of these messages. As a consequence, restoration paths can only be found in a confined area of the network, which may reduce the restoration ratio.

3.5.3.5 Strengths and Weaknesses of Dynamic Resilience Provisioning

The rapid evolution of Internet and e-commerce requires the new generation optical network to be reconfigurable to accommodate dynamic changing traffic. Moving from static to dynamic reconfigurable networks offers a number of advantages to both network providers and end-users. It enhances the efficiency of network usage through dynamic time-sharing of the resources, allowing end-to-end connections to be set up and released based on traffic demands. Automatic provisioning opens up the possibility of enabling users to pay for high-bandwidth connections based on usage.

In such a dynamic environment, connection setups are no longer rare. This optical network must be able to add and drop lightpaths automatically and dynamically. Accordingly, resilience provisioning must also be carried out dynamically in this new environment.

In order to provide real-time solutions, dynamic resilience provisioning mechanisms usually adopt simplified distributed algorithms. Therefore it is more suitable for traffic with fast changing patterns where future demands are not certain. These simplified algorithms are easy to implement and all the path calculation procedures can be performed automatically. Dynamic resilience provisioning also applies well to mesh network topologies.

However, distributed algorithms find an optimal solution for each single connection, without considering others and the global or even neighbour traffic situations, therefore degrading global network utilisation. Even more, over-time, the continual establishment and removal of these connections based on short-term expediency, can lead to an inefficient exploitation of the underlying network resources.

3.6 Research Focus and Contributions of this Thesis

As the new generation optical network is evolving to be reconfigurable and support dynamic changing traffic, new resilience provisioning mechanisms are needed to support this evolution. In addition, an IP-centric control plane is suggested to realise the dynamically reconfigurable optical network. Therefore, existing mechanisms should be examined to see if they are applicable in the new networking environment.

This research focuses on the study of dynamic resilience provisioning mechanisms for IP-centric optical networks. The main efforts include, firstly, investigating if there is any existing resilience provisioning scheme that could be adapted to the new networking environment; secondly, inventing new resilience provisioning schemes to cope with the mesh-based dynamic optical network environment. The main research area is shown in Figure 3-6.

The contributions of this research include three main parts (shown in Figure 3-6):

Firstly, a flooding-based reactive restoration scheme named Fast Restoration Scheme (FRS) is proposed.

Restoration is suitable for the rapid changing traffic in the network because it is more flexible if restoration paths are assigned only after the failure has occurred. Flooding-based restoration uses the flooding messages to discover alternative paths after the failure occurs. It does not need each node of the network or a NMS to maintain a global state of the network, thus it is easy to implement.

The FRS is a flooding-based restoration scheme applied in the IP-centric optical network environment. In addition, by maintaining a dynamically refreshing *Resource Table* in the *Receiver*, FRS precludes the possibility of link contentions and usually finishes the restoration connection with only one connection attempt. The mechanism of setting up restoration path from the *Selector* other than the *Receiver* ensures a loop-free connection. This scheme is introduced in Chapter 4. This work is published in [DON05] and a relevant patent has been filed by Nortel Networks [PAT1].

Secondly, a resilience-provisioning scheme named Adaptive Segment Path Restoration (ASPR) is proposed.

Dynamic resilience provisioning mechanisms in the mesh network topology can be classified as either link or path protection (or restoration) (Figure 3-6). Both methods have their own limitations.

A new resilience-provisioning scheme is proposed to compromise link and path protection. In this approach, a lightpath (or LSP) is divided into several segments. For each segment of the primary path, it is provided with a backup path. The segmentation of the primary path is adaptive to the topology of the network, allowing for more efficient resource usage whiles yielding restoration times comparable to link restoration. The implementation of the proposed scheme needs only some improvement to the existing MPLS/GMPLS signalling protocols, which makes it simple and be able to work automatically.

The comparative study and simulation results of the proposed scheme with others show that ASPR has the best restoration time performance, whilst it remains better than most other restoration schemes in terms of its spare capacity requirement. This approach could also be used to protect against multiple failure in the mesh network. The detailed work is presented in Chapter 5. Publications about this work are [DON2][DON4]. Especially, the latter has been awarded one of the six best papers in the ICT2002.

Finally, a Differentiated-Resilience Optical Services Model (DROSM) for next generation optical network is proposed.

In order to provide a range of resilience types that better reflect the value of the traffic being carried, this research proposes a novel model providing differentiated-resilience optical services.

In particular, optical services are classified according to their resilience requirements. Each resilience class is then provided with a different restoration strategy. The decision of restoration strategies is based on a novel analysis of optical restoration. In addition, a novel resource management mechanism is put forward to coordinate different resilience classes.

This model is applied to both optical networks with and without wavelength conversion capabilities, which are detailed in Chapter 7 and Chapter 6, respectively.

Publications presenting this work are [DON1][DON3]. A patent by the author relevant to this work has also been filed by Nortel Networks [PAT2].

3.7 Summary

In this chapter, the basic resilience provisioning mechanisms are introduced and a new classification framework of these mechanisms is presented. Firstly, the general classification of resilience provisioning schemes is introduced. Secondly, the resilience provisioning mechanisms are presented according to the topology they are applied to. Then, resilience-

provisioning mechanisms in mesh network topology are investigated in detail and requirements on resilience provisioning in the new generation optical network are illustrated. Finally, the scope of this research is defined and contributions of this research are also stated.

Chapter 4 Fast Restoration Scheme – A Flooding-Based Restoration for the Optical Network

Flooding-based restoration does not require each node having information about the whole network. Therefore it is easier and cheaper to implement. It also may have a better performance, as the database in each node used by a GNS-based restoration takes some time to update after the failure event. Some flooding-based restoration schemes have been studied in Digital Cross-connection Switches (DCS) and SONET networks

This chapter investigates the possible applications of flooding-based restoration in the IP-centric optical network. A new flooding based restoration scheme entitled Fast Restoration Scheme (FRS) is proposed. Different from the similar schemes in DCS and SONET networks, the proposed scheme has a shorter restoration time by utilising a *Resource Table* instead of signalling messages to decide the resource allocation of restoration paths. In addition, it avoids loops for restoration paths and therefore consumes less spare resource.

4.1 Overview

Resilience schemes can be classified into two general categories: **protection** and **restoration**, depending on whether resources are pre-allocated before the failure occurs or not. The technique that uses pre-assigned capacity to ensure survivability is referred to as protection, and the technique that reroutes the affected traffic after failure occurrence by using available capacity is referred to as restoration [WU97][FUM00]. Protection and restoration are also referred to as **proactive/pre-planned restoration** and **reactive restoration** in [MOH00][COA91][MED99], respectively.

Protection reserves resources, identifying backup paths at the time of establishing the working paths to protect traffic against possible failures. Since it does not need the time-consuming path calculation/search and connection reestablishment process, protection is capable of restoring traffic within a very short time. It is widely employed in the traditional optical networks.

Restoration comes as the result of the introduction of mesh-based networks. It has the merit of being more cost-efficient, since it does not reserve spare resources before the failure. The drawbacks of this approach are, firstly, that the amount of spare resource may not be adequate and thus cannot ensure a successful traffic restoration; secondly, that the restoration

latency can be several seconds or even longer, especially in heavily loaded networks [YE00]. However, since the requirement of traffic on resilience varies a lot and it can save a lot of spare resource, reactive restoration still has significance, especially for traffic with relatively lower resilience requirement.

The procedure of reactive restoration is a routing problem, which includes: propagation of the failure information to the relevant node, finding the restoration path based on the available resource, building up the restoration connection and switching over the traffic onto the alternative connection.

Accordingly, the finding of the restoration path could be performed by two means. One is using database maintained in NMS (centralised scenario) or in each OXCs (distributed scenario) to calculate restoration path, which is called GNS (Global Network Sate)-based. The other is using flooding messages in the network to discover the alternative path, which is called flooding-based.

GNS-based has many merits. It achieves its simplicity by transforming a distributed problem into a centralised one. By maintaining a database of the network state, the OXC calculates the alternative path according to the current network state. The calculation algorithms are conceptually simple and easy to implement. However, it has several disadvantages. The cost of maintaining a link state database is high, both for the centralised and for distributed scenario. For the centralised scenario, reliable connections are needed between the NMS server and all the OXCs in the network. Its restoration is relatively slower. For the distributed scenario, the database has only aggregated information of the link state of the network and cannot update in time immediately after the failure occurs. As a result, the path calculation algorithm may fail to find an alternative path while spare resource is still available.

In contrast, flooding-based restoration does not require the nodes in the network to maintain a global link state database of the network, which does not need the link state protocol and thus is cheaper to implement. The flooding-based restoration may have a better restoration ratio than GNS-based restoration, since in the latter case, the database used by alternative path calculation takes some time to update after the failure occurs. The drawbacks of flooding-based restoration include: Firstly, the restoration time is longer than that of GNS-based since the finding alternative path is performed by flooding messages, which takes extra time for message propagation and processing. Secondly, the flooding messages make the communication overhead excessively high. In this case, a hop count is usually introduced to

limit the flooding area of the messages. As a consequence, the possible alternative path can only be found in a confined area of the network, which may reduce the restoration ratio.

In this chapter, a novel flooding-based restoration scheme called Fast Restoration Scheme (FRS) is proposed. The scheme is fully distributed and depends only on the local state maintained at every individual node. It does not require each OXC to maintain a global link state database using a link state protocol. The simulation result shows that the restoration time of the proposed scheme is short and scales well.

4.2 Background and Related Work

Flooding-based restoration has received considerable attention for applying in different networks. It starts from the assumption that each node might cause problems of consistency among distributed databases when changes occur in the network. Therefore, the flooding-based restoration algorithms build up network information (topology and location of spare resources) required to restore the failed paths after the failure has occurred, guaranteeing an up to date view of the network. Grover et al. presented the first flooding-based restoration algorithm [GRO87][GRO91], later followed by several other researches [YAN88][KOM90] [CHO93][BIC93][CHO99]. All these algorithms are targeted at networks based on the SONET/SDH transmission standards. Researchers have also developed algorithms to integrate the flooding-based restoration with ATM networks.

Typically, a flooding-based restoration consists of two phases: *broadcasting phase* and *selection phase*. When there is a link failure, the node on one side is designated as the *Sender* and the other as the *Receiver* (or *Chooser*). The broadcasting phase is started when the *Sender* learns about the failure of the link. *Probe* messages are “flooded” from the *Sender*. Usually a *probe* message contains the following information: the ID of a flow that it intends to restore, the *Receiver* ID, the *node list* which is dynamically adjusted as it travels and a *hop count* to limit the broadcasting area. When nodes other than the *Sender* and *Receiver* receive the *probe* message, they will be only forward it on those interfaces with spare bandwidth. The *node list* is also used to prevent path loops. The *selection phase* begins after a *probe* message reaches the *Receiver*. A *connection request* message is then sent back along the path that is indicated by the *probe* message’s *node list*. Since the spare bandwidth could have been consumed by other restored flows, a *connection ack / failure* message is needed to be sure of the success of the connection. The *Receiver* attempts to restore the path based on the *node list* in the *probe* message. If a *connection failure* message is received, the connection procedure is started

again along a new path denoted by the next stored *probe* message. This continues until either the list of *probe* messages at the *Receiver* is exhausted or a *connection ack* has been received.

During the *broadcasting phase*, the purpose of the flooding *probe messages* is to find possible alternative paths. Thus, resources (bandwidths) cannot be reserved since the flooding message does not necessarily result in a successful restoration path. Therefore, a considerable number of connection attempts may fail due to resource contention during the *selection phase*. In this case, a *connection failure message* is used to notify the *Receiver* to start connection procedure again. Consequently, the restoration is delayed on account of the time consumed by failure notifications and multiple connection attempts.

The message flooding is usually confined in certain area around the failure so as to reduce the traffic burden caused by flooding messages, and for a faster restoration. Hence, link restoration rather than path restoration is normally adopted in flooding-based restoration algorithms. As a result, loops or route rewinds are usually formed in the restored connection, taking extra resource that may otherwise enable more failed connections to be restored. Figure 4-1 shows an example of a traffic loop caused by the simple *Sender* and *Receiver* flooding-based restoration algorithms.

Figure 4-1(a) shows the network state when the failure between nodes B and C occurs. The possible results of the flooding-based restoration are shown in Figure 4-1(b) and (c). Here, traffic loop consumes unnecessary resources that could be used for other connections. A better solution to these scenarios would be as shown in Figure 4-1(d) and (e), respectively.

The proposed flooding-based restoration algorithm, FRS, precludes possible resource contentions by building and maintaining a *Resource Table* in the *Receiver* during the traffic restoration time. The *Resource Table* records the network link state of the flooding area. Its content refreshes when a message arrives at the *Receiver* or a connection attempt is issued from *Receiver*. Thus, the *Resource Table* has the latest information about the network's link state. Before the *Receiver* originates a command for a connection attempt, it checks the *Resource Table* to see if there is enough resource along the proposed restoration path. If there is not enough, the connection attempt will simply not be sent out and the *Receiver* will consider a new restoration path denoted by another *probe* message. Therefore, most possible failures of the connection attempts can be precluded by means of checking the *Resource Table* while not using the *connection request* and *ack/failure message*. In this way, most restoration connections can be restored with only one connection attempt, which results in a faster restoration for the failed traffic.

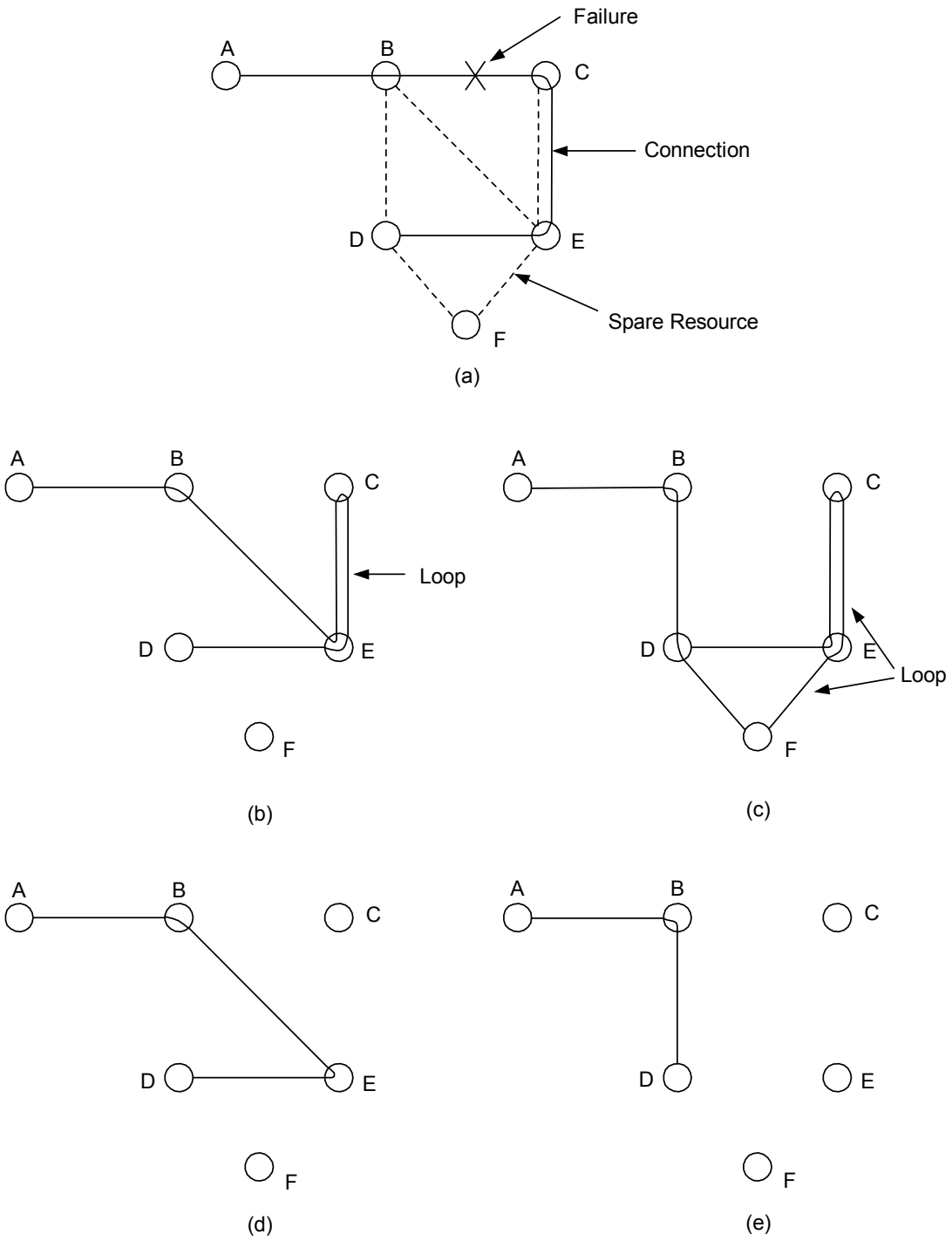


Figure 4-1: Example Connection Loop

FRS also provides a loop-free restoration, which enables more failed connection to be restored. This is achieved because that actual restoration paths for the failed connections are searched for between nodes other than just the *Receiver* and the *Sender*.

4.3 Fast Restoration Scheme

FRS provides a flooding-based restoration in the WDM optical network. It assumes that the optical network uses a GMPLS-based control plane. Therefore, each OXC is assigned an IP address as its LDP ID. Here, optical lightpaths, optical connections and LSPs have the same meaning.

FRS is basically composed of three phases including: a *broadcast phase*, a *selection phase* and a *connection phase*. During the *broadcast phase*, flooding of *probe messages* is used to search for possible restoration paths for the failed connections. The *selection phase* is used to decide the restoration paths and to find nodes that are assumed to initiate those restoration connections. The *connection phase* is to build up the restoration connections.

4.3.1 Broadcast Phase

The purpose of this phase is to search for all possible routes of the restoration paths. An OXC maintains information including a Label Mapping Table, Peer Nodes' LDP IDs, the number of wavelengths or optical channels on each interface and its own LDP ID.

In an optical network, the communication of two nodes requires bidirectional paths. It is suggested that these two bidirectional paths of one connection take the same route. Extensions have already been proposed and described for CR-LDP [BER02] and RSVP [ASH02] to establish bidirectional paths. It is also recommended that the bidirectional restoration paths for one connection take the same route as well. Therefore the search for restoration paths only need to focus on one unidirectional path and the other one takes the same route.

When a link failure is detected, of the two adjoining end nodes, the one with a smaller IP address is designated as a *Sender* and the other is designated a *Receiver*. The search for restoration paths focuses on those unidirectional LSPs that run through from the *Receiver* to the *Sender*. The nodes that a LSP travels through are called *on-path nodes* of that LSP. All the other nodes are called *tandem* node. Figure 4-2 shows an example of terms used for nodes in the network immediately after a failure.

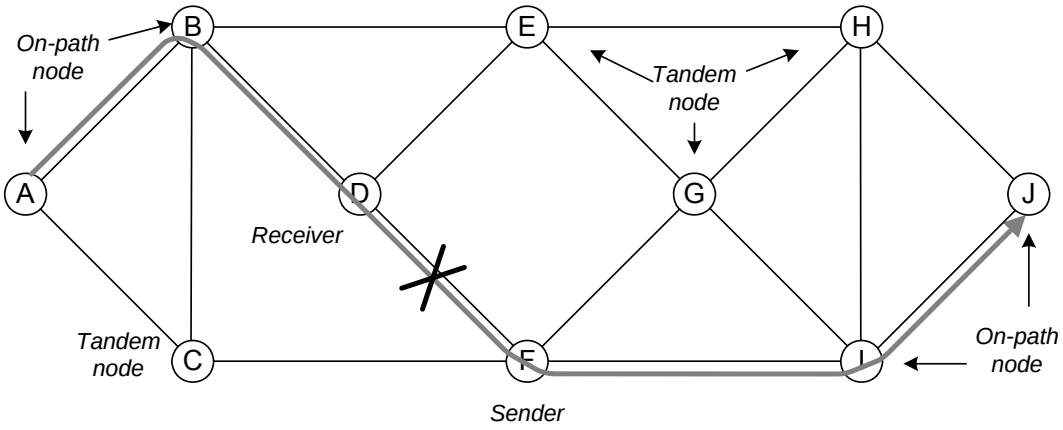


Figure 4-2: Node Terms

When a failure is detected, a *probe message* is created at the *Sender* for each affected LSP. A *probe message* consists of the following information:

- MT: Message type that denotes it as a *probe* message
- LSP_ID: The failed LSP ID
- RID: The *Receiver*'s LDP ID
- HC: Hop count, indicating the distance in hops from the *Sender*
- PV: Path Vector, recording the nodes that the probe message passed through
- SRV: Spare Resource Vector, recording the spare resource of the path

The *probe message* is supposed to find possible alternative paths for a failed LSP. It is only sent to those interfaces that have spare resource, except the failed interface. A *probe message* does not flood back to the interface it comes from. The *probe messages* for a failed LSP are flooded throughout the network, and terminate at an *on-path node* when they reach it from the interfaces other than that the failed LSP travels through.

On receiving a *probe message*, the *tandem* node updates fields in the *probe message* and then floods it further to those interfaces with spare resources. These updates include: The total spare resource on that interface is appended to SRV; the HC is incremented by one; the tandem node ID is appended in PV.

If a *probe* message for a failed LSP reaches an *on-path* node from interfaces that the LSP travel through, the *on-path* node updates the *probe* message and floods it further to those interfaces with spare resources. These updates are as the following: increase the HC by one; reconstruct a new SRV with only the number of spare wavelengths of the interface to which the refreshed probe message is to be sent; reconstruct a new PV with only current node LDP ID being inserted. Here the PV does not record the travelled path of the probe message along

the failed LSP. Its purpose is to avoid any possible loop in the restoration LSP downstream of the failure point.

It is essential to avoid the generation of redundant *probe messages* for fast and efficient restoration. In this case, a loop condition is avoided by checking the PV field. Furthermore, the *probe messages* of which the hop count (indicated by HC) exceeds a pre-determined limit are discarded. When a *probe message* for a failed LSP reaches an *on-path node* from the interfaces other than that the failed LSP travels through, the flooding process stops and the *probe message* is not further flooded.

An example of *probe message* flooding is depicted in Figure 4-3. Here, a LSP runs from A to J. When the link D-F fails, node F and D are designated as the *Sender* and the *Receiver*, respectively. The *Sender* creates a *probe message* for the failed LSP and sends a copy to each of the interfaces (to node C, node G and node I). When a *tandem* node receives a *probe message*, it first updates its content and then forwards it to other interfaces. When an *on-path* node receives a probe message from interfaces the failed LSP travels through (e.g. node I receives a probe message for the failed LSP from node F), it updates the probe message and floods it further to other interfaces. The flooding stops when *probe messages* arrive at an *on-path node* of the failed LSP from interfaces other than the failed LSP travels through. Here the probe message will not be forwarded again when it arrives at node A from node C, or at node B from node C or node E, or at node D from node E, or node F from node G or node C, or at node I from node G and node H, or at node J from node H.

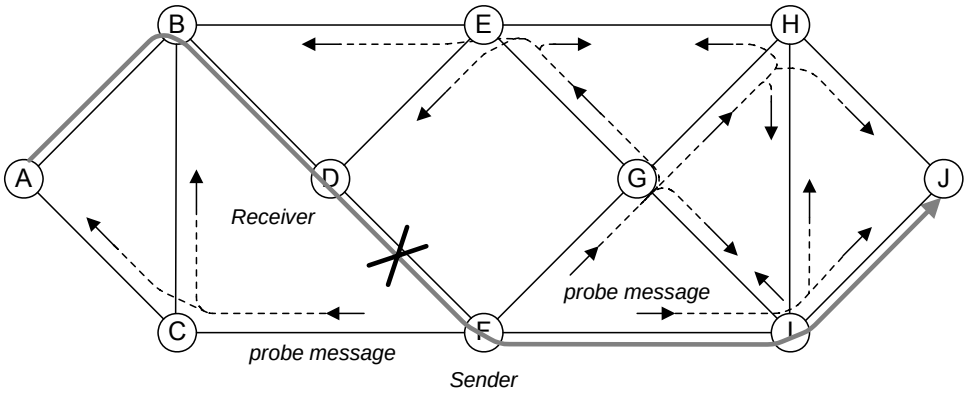


Figure 4-3: Flooding of Probe Messages

4.3.2 Selection Phase

Each *probe message* that reaches an *on-path* node of the failed LSP contains a suggestion of the restoration path for the failed LSP. As multiple *on-path* nodes, downstream and upstream of the failure, may receive a *probe message* for the same failed LSP, a mechanism is needed to decide which node is going to initiate the setting up of the restoration path.

The procedures of the *selection phase* shown in Figure 4-4 are used to decide the *Selector*, which is responsible for the initiation of the restoration path for a failed LSP. Setting up the restoration path is started from the *Selector*, while unlike other flooding-based restoration algorithms which start from the *Receiver*. This mechanism guarantees a loop-free restoration path upstream of the failure point.

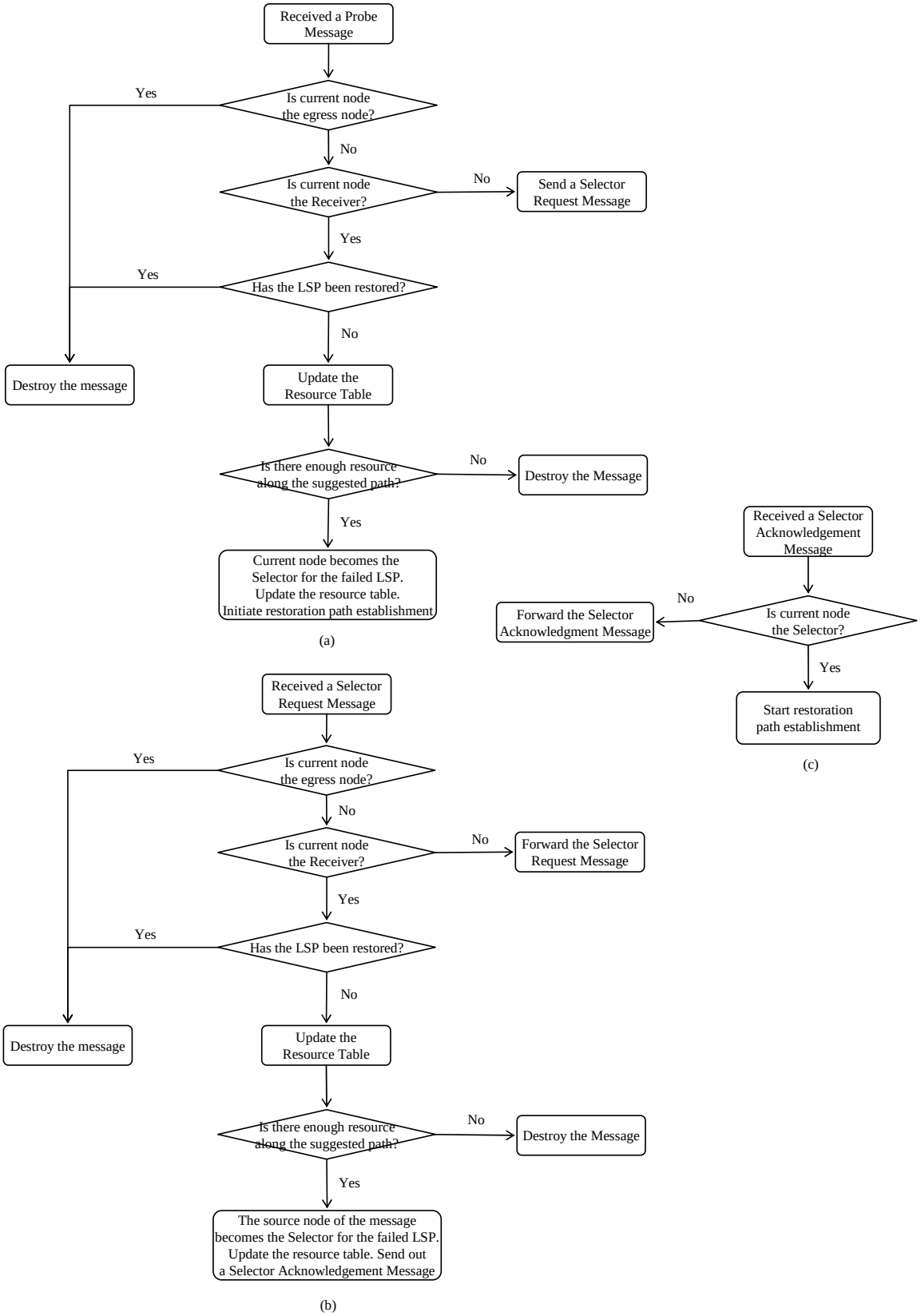


Figure 4-4: Selection Phase Procedures

During the *selection phase*, two types of messages, *Selector Request Messages (SReqM*) and *Selector Acknowledgement Messages (SAckM*), are used. They are the correspondence

between the *Receiver* and any *on-path node* that has received a *probe message* for a failed LSP. All the *on-path nodes* that have received a *probe message* are possible to initiate the setting up of the restoration path suggested by the *probe message*. However, in order to avoid the unnecessary resource contention, they need to consult the *Receiver* first. This is done via sending out a *Selector Request Message* to the *Receiver*. A *Selector Request Message* consists of the following information:

- MT: Message type that denotes it as a *SReqM*
- LSP_ID: the failed LSP ID
- RID: the *Receiver*'s LDP ID
- S_ID: the ID of the *on-path node* which creates this *SReqM*
- PV: Path Vector, same as that in the received *probe message*
- SRV: Spare Resource Vector, same as that in the received *probe message*

When an *on-path node* receives a *probe message* (this is done by detecting that the LSP_ID carried by the *probe message* is in the local Label Mapping Table), it creates a *Selector Request Message* and sends it downstream of the LSP. The node that receives a *Selector Request Message* forwards it downstream further until the message reaches the *Receiver* or the egress node of the failed LSP. An example of creating and sending *Selector Request Messages* is illustrated in Figure 4-5.

In this figure, when node A receives a *probe message* for the failed LSP, it creates a *Selector Request Message* and sends it to the downstream node B. Node B further forwards the message to the *Receiver*. Similarly Node B and Node I also create *Selector Request Messages* to request them to be designated as the *Selector* for the failed LSP.

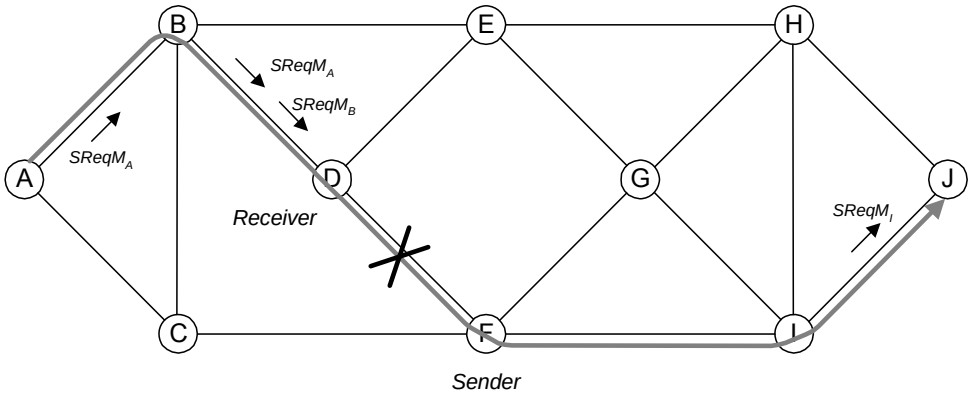


Figure 4-5: Sending out Selector Request Messages

The *Receiver* also has the possibility of getting a *probe message*, which accordingly contains a possible restoration path. It therefore also can be selected as the *Selector* for the failed LSP. For all these requests, the *Receiver* takes a strategy of *first come first serve*. The first request for taking charge of the restoration for a failed LSP will be designated as the *Selector* for it.

When the *Receiver* receives either a *probe message* or a *Selector Request Message* for a failed LSP, it checks whether the targeted LSP has been restored or not. If it has already been restored, the message is simply destroyed. If it has not been restored, the *Receiver* firstly uses the field of SWV in the message to update a *Resource Table* maintained in the node. The *Resource Table* has the following fields:

- S_ID: Source Node
- D_ID: Destination Node
- Resource

The *Resource Table* records the available resource in the flooding area. Its content is updated when a *probe message* or a *Selector Request Message* is received. Before a possible restoration path is approved, the content of the *Resource Table* is checked. If there is not enough spare resource along the suggested path, the message is simply dropped. If there is available resource for the suggested restoration, the *Receiver* either initiates the restoration path establishment if it has been designated as the *Selector* for the failed LSP, or sends back a *Selector Acknowledgement Message* if an *on-path* node has been approved as the *Selector*.

By this means, it can preclude resource contention, which otherwise takes several rounds of establishment attempts using signalling messages and therefore needs more time to finish the restoration.

Selector Acknowledgement Messages are used to notify the approved *Selector* that is to initiate the restoration of the failed LSP. A *Selector Acknowledgement Message* contains the following information:

- MT: Message type that denotes it as a *SAckM*
- LSP_ID: The failed LSP ID
- S_ID: *Selector* ID
- PV: Path Vector, copy of that in the *SReqM*

The *Selector Acknowledgement Message* is forwarded upstream along the failed LSP until it reaches its destination the *Selector*, which is defined by the S_ID field of the message.

During the propagation of the *Selector Acknowledgement Message* from the *Receiver* to the *Selector*, all the resource in this segment consumed by the failed LSP is released. This provides more resource for the restoration of other traffic. After it receives the confirmation, the *Selector* then starts setting up the restoration path suggested by the PV field contained in the message.

Another point that should be mentioned is that all the *on-path nodes* downstream of the *Sender* may also receive some *probe* messages. The restoration paths suggested by the PV field of these messages are actually not valid. Although the *on-path nodes* send out *Selector Request Messages* downstream along the failed LSP, these messages will eventually reach the egress node of the failed LSP and be destroyed. That they will never get to the *Receiver* ensures no invalid restoration path is produced.

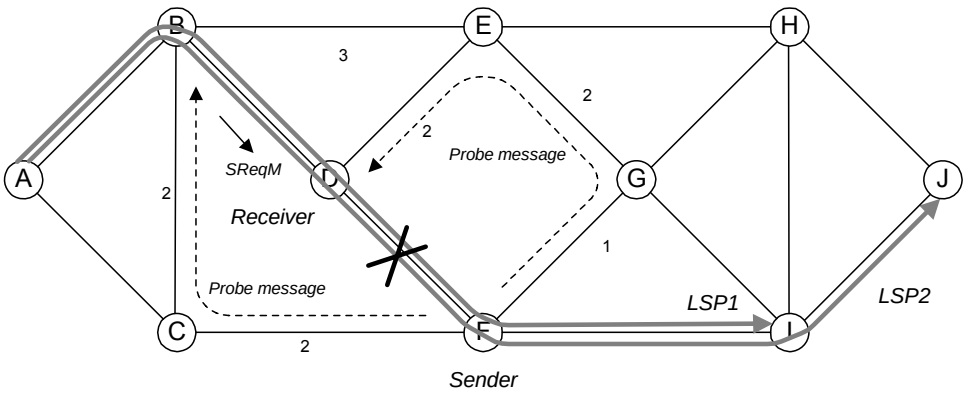


Figure 4-6: Choosing Selectors

Figure 4-6 illustrates how *Selectors* for different LSPs are chosen. Assume there are two LSPs affected by the failure of link DF. The numbers beside the links denote the spare resources (wavelength, optical channel) in each link. Suppose that a *probe* message for LSP1 arrives on the *Receiver* first. The *Receiver* will initialise the *Resource Table* as shown in Figure 4-7(a). By checking the *Resource Table*, the *Receiver* knows there is enough spare resource for the proposed restoration path DEGF. Therefore, the *Resource Table* is updated as Figure 4-7(b) before the *Receiver* starts the setting up of the restoration path for LSP1.

S_ID	D_ID	Resource
G	F	1
E	G	2
D	E	2

(a)

S_ID	D_ID	Resource
G	F	0
E	G	1
D	E	1

(b)

S_ID	D_ID	Resource
G	F	0
E	G	1
D	E	1
C	F	2
B	C	2

(c)

S_ID	D_ID	Resource
G	F	0
E	G	1
D	E	1
C	F	1
B	C	1

(d)

Figure 4-7: Updating *Resource Table*

The *probe message* for LSP2 may arrive at the *Receiver* from node E immediately after that for LSP1. The *Resource Table* needs to be synchronized according to the information contained in the new *probe message*. In the circumstance of inconsistent values of the same item, the smaller value is picked from the *Resource Table*. In this case, the table keeps intact as shown in Figure 4-7(b). Therefore, the *Receiver* will find there is no available spare resource for the proposed restoration path (DEGF) because no spare resource exists between node G and F.

Two *probe messages*, one for LSP1 and one for LSP2, may arrive on node B as shown in Figure 4-6. In this case, node B creates two *Selector Request Messages*, one for LSP1 and one for LSP2, and sends them to node D which is the *Receiver*. On receiving the *Selector Request Message* for LSP1, the *Receiver* gets that LSP1 has been restored. The message is simply destroyed. On receiving the *Selector Request Message* for LSP2, the *Resource Table* is updated again. In this case, it becomes as shown in Figure 4-7(c). Since there is enough resource along the proposed restoration path BCF, node B is assigned as the *Selector* for LSP2. The *Resource Table* is updated as shown in Figure 4-7(d) before a *Selector Acknowledge Message* is sent out.

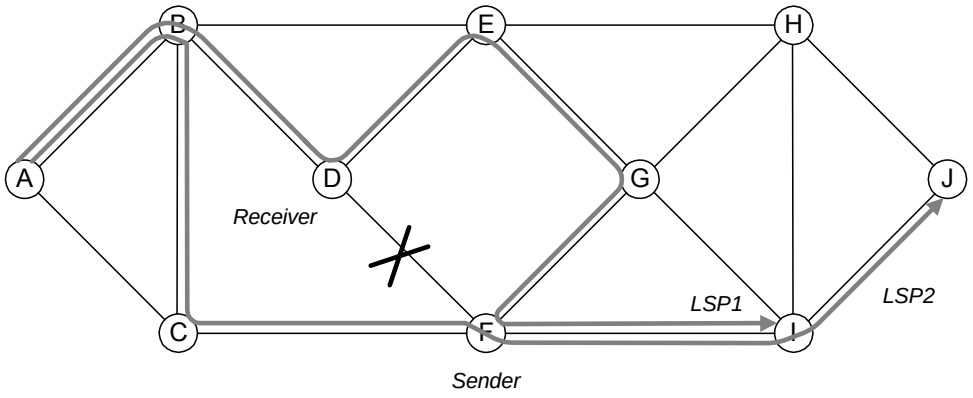


Figure 4-8: Restoration Results

4.3.3 Connection Phase

The *Selector* is responsible for starting the deployment of the restoration paths, which is done in the *connection phase*. The deployment of the restoration uses the standard signalling protocol CR-LDP or RSVP-TE.

As the bidirectional paths of a restoration need take the same route, the restoration for a LSP contains establishment of a pair of bidirectional paths at the same time. When the signalling message for the establishment of the restoration path reaches the merge node, the node checks whether it is the *Sender* or not. If it is not the *Sender*, a withdraw message is sent out the release the resource used by the failed LSP from the merge node to the *Sender*. This provides more resource for the restoration of other traffic.

As a result, the paths after restoration for the two failed LSPs are shown as in Figure 4-8.

4.4 Performance Evaluation

This section presents simulation results of the proposed Fast Restoration Scheme. An event-driven simulation tool OPNET^{TM} is used to model the scheme. The performance of implemented models has been verified and validated using simple network toplogoies.

The flooding-base restoration requires careful design of the amount of spare resource and its distribution in the network to ensure a full restoration of all traffic. The simulation uses a network model that is well known for evaluation of flooding-based restoration schemes.

The network model (New Jersey LATA network) used here for performance evaluation of the propose scheme, is shown in Figure 4-9(b). In the figure, each node is identified by a node number, and each link has two numbers that indicate the number of working and spare channels, respectively.

This model is also used by several other authors [YAN88][JOH93][BIC93][VAN96] to investigate their flooding-based restoration schemes in other networks, including DCS, SONET and ATM networks. In these flooding-based schemes, link restoration is usually used. Flooding messages are used to find restoration paths between the two end nodes of the failed link. As a result, restoration paths in these schemes are usually not loop-free, requiring more resource.

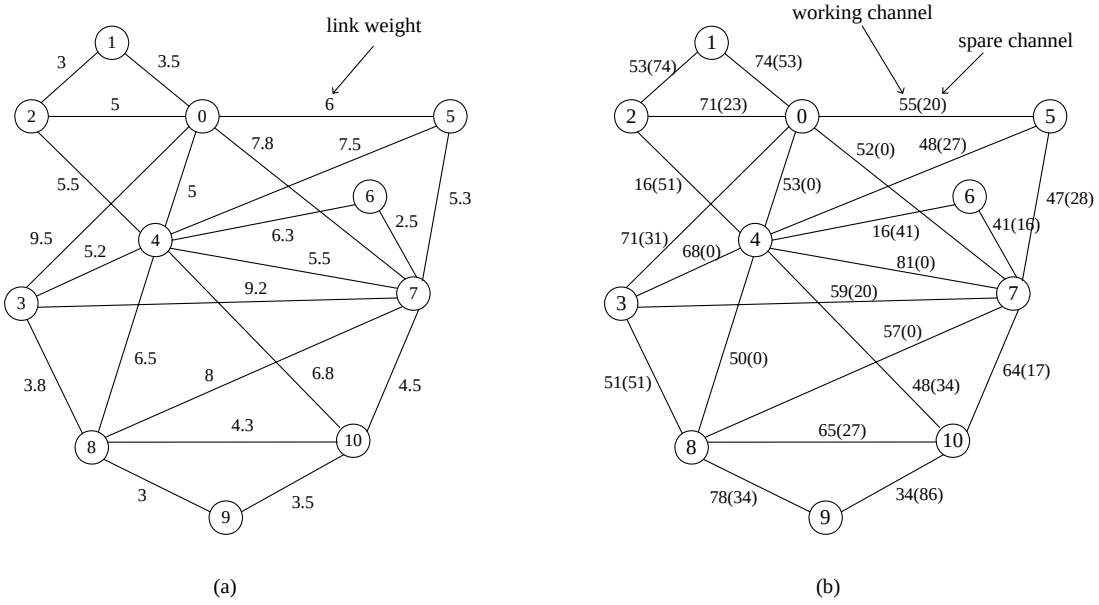


Figure 4-9: Performance Evaluation Network Model

In the proposed FRS, restoration paths do not simply run from the *Receiver* to the *Sender*. More path information is considered to prevent any loop for the restored traffic.

Therefore, the exact information of each LSP is required in the simulation to evaluate the scheme. In order to have the same deployment result as shown in Figure 4-9(b), a traffic matrix is assumed as shown in Table 4-1:

	0	1	2	3	4	5	6	7	8	9	10
0	-	28	41	59	38	17	8	8	15	2	8
1	-	-	44	12	7	8	8	8	2	2	8
2	-	-	-	1	1	30	1	1	1	1	1
3	-	-	-	-	41	18	8	59	17	16	18
4	-	-	-	-	-	30	7	80	18	13	47
5	-	-	-	-	-	-	8	15	8	8	8
6	-	-	-	-	-	-	-	8	5	2	2
7	-	-	-	-	-	-	-	-	44	10	14
8	-	-	-	-	-	-	-	-	-	48	47
9	-	-	-	-	-	-	-	-	-	-	10
10	-	-	-	-	-	-	-	-	-	-	-

Table 4-1: Traffic Matrix

Each unit of the matrix represents the number of *bidirectional* connections between two nodes, hence the triangular shape of the matrix. For example, there are 13 bidirectional LSPs between node 4 and node 9. The shortest path algorithm is used to deploy these connections. Figure 4-9(a) shows the weight of each link of the network used by path calculation.

The processing delay from the arrival of a message to the end of the processing depends on the processing capability of each OXC. In the simulation, the processing delay for each message is set at random between 1 and 5 ms. The average propagation delay of each link is set as 5 ms. The control channel between each two nodes is set to be 64 kb/s. The limit of HC is set to be 7. The time consumed by an OXC to switch on a connection is assumed as 10 ms. It is further assumed that the cross-connect requests resulting from all restoration messages received within each 10 ms interval are grouped into a single (batch) command and transmitted to the OXC at the end of the interval.

The simulation has been performed 10 times using different initial seeds. The result plotted in the graph is the average of these 10 simulations.

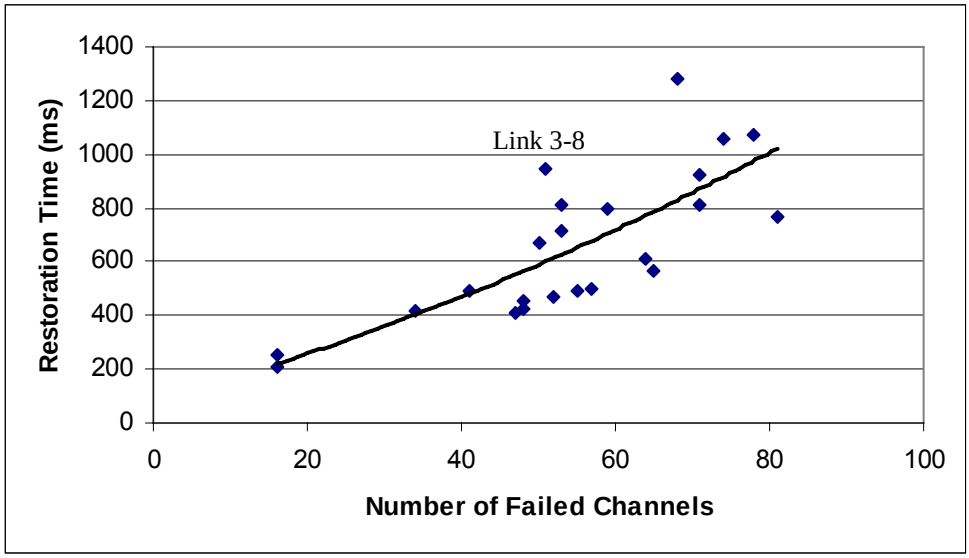


Figure 4-10: Restoration Time vs. Number of Failed Channels

Figure 4-10 and Figure 4-11 show the simulation results of restoration time after each of the 23 single link failures. As a result, 100% restoration can be achieved in each case. The result validates the performance of the proposed scheme FRS.

Figure 4-10 shows the average restoration time versus the number of failed wavelength channels. For example, 51 working paths are affected by the single link failure of Link 3-8 noted in the figure. It takes about 950 ms to restore all the failed connections.

The result shows that the restoration time varies a lot in cases of different link failures. For example, it takes only about 210 ms to fully restore all the traffic affected by the failure of link 4-6, whilst about 1280 ms is needed to restore all the failed connections after the failure of link 3-4.

The result also shows the distribution of spare resource affects the restoration time. Although the number of connections affected by the failure of link 3-4 is not the largest in all the single link failure cases, it takes the longest time to restore the failed traffic. This is due to the limited spare resource near the failure point. It requires the probe messages to be flooded to a wider area and thus consumes more time to restore the traffic.

Despite of this, there is generally a trend, which is illustrated by the trend line in the figure, of that a link failure with more working channels requires more time to restore the traffic. It is easy to understand this as more failed working channels triggers more probe messages that are flooded in the network and introduce more delay for the whole restoration process.

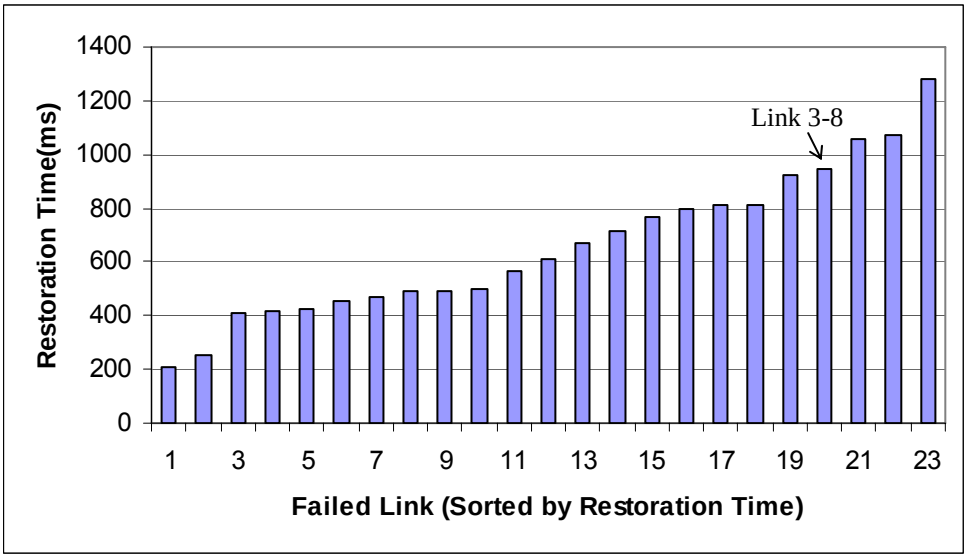


Figure 4-11: Restoration Time versus Failed Link (Sorted by Restoration Time)

The result also validates the analysis about flooding-based restoration presented in Chapter 3 of this thesis. Flooding-based restoration searches for alternative paths by flooding messages into the network. The restoration time mainly consists of the processing delay, the transmission delay and the propagation delay that are introduced by processing and transmitting the flooding messages. Due to the large number of flooding messages that are produced for the failed connections, the restoration normally takes several hundred ms to several second to finish.

However, this also gives a prospect that the restoration time could be cut shorter by improving the processing capability of each node and increasing the bandwidth of control links which are used by transmitting the flooding messages.

4.5 Summary

Flooding-based restoration has the advantage of being easy to implement. It uses flooding messages to search for the restoration paths. In this chapter, a novel flooding-based restoration entitled Fast Restoration Scheme (FRS) is proposed to provide dynamic restoration for the optical network. It has mainly two advantages compared with other flooding-based restoration schemes, which are proposed for the DCS, SONET and ATM networks.

First, traditional flooding-based restoration schemes search for restoration paths directly between the *Receiver* and *Sender* that usually result in traffic loops for the restoration paths.

Restoration paths searched by the FRS can originate and end at nodes other than these two nodes. This prevents any possible loop at both sides of the failed link. As a result, it consumes less resource to restore the traffic.

Second, unlike those flooding-based restoration schemes which utilise a try and acknowledgement signalling mechanism to solve link contentions, the FRS adopts a *Resource Table* mechanism to prevent possible link contentions and therefore requires less time.

The simulation results show that even with a low-speed signalling link (64 kb/s) the scheme can restore the traffic with a fairly short time.

Chapter 5 Adaptive Segment Path Restoration (ASPR)

5.1 Overview

Network resiliency schemes can be roughly classified into two categories as reactive restoration or proactive restoration/protection.

Reactive restoration commences only after a failure has taken place. A typical restoration action is to reallocate the unreserved network resources to fix the flows that were affected by the fault. The drawbacks of this approach are, firstly, that the amount of unreserved resources may not be adequate and some flows may have to be rejected and, secondly, that the recovery latency can be several seconds or even longer, especially in heavily loaded networks, since time is required to find and establish the alternative paths. This makes these schemes only suitable for best effort services. Proactive restoration (protection) reserves resources, identifying backup paths to protect traffic against possible faults at the time of establishing the primary paths. Since these schemes do not need the time-consuming connection reestablishment process, proactive restoration is capable of restoring traffic within a very short time. A successful application is within SONET/SDH protection rings. Here a restoration time of less than 50 milliseconds has become a benchmark within the industry. However, the drawback of proactive restoration is the high cost. There will generally be an investment of at least 100% in transmission capacity redundancy. For better resource utilization, resource sharing between backup paths can be employed. If two primary paths do not fail at the same time, their backup paths can be shared with each other, and thus the costs can be reduced.

In the mesh-based network, two basic schemes can be used for resilience provisioning. They are link-based protection / restoration and path-based protection / restoration.

Link protection / restoration employs local rerouting to cover a particular link. It reroutes traffic around the failed component. When a link fails, a new path is selected between the end nodes of the failed link. Link restoration has an advantage of being able to restore traffic in a very short time since the rerouting of the traffic is close to the failure. However, it requires setting aside significant spare resource for the backup path, which may not be affordable [RAM99a][MOH00].

Path protection / restoration uses end-to-end rerouting to cover the whole path. In path restoration, a backup path is established between two end nodes of the primary path. Path restoration has better performance on resource sharing and thus requires less spare resource than link restoration. However, signalling is needed to notify the ingress OXC to switch over to the backup path, which results in a longer restoration time than link restoration. So path restoration cannot satisfy the requirements of some real-time services [MOH00][RAM99a].

In order to achieve a restoration time that can satisfy the requirements of real-time services and consumes reasonable spare resource at the same time, this thesis proposes a novel resilience provisioning mechanism entitled Adaptive Segment Path Restoration (ASPR). The basic idea is to divide the whole path into several segments. For each segment, a backup path is provided for protection. The segmentation of the primary path is adaptive to the topology of the network, allowing for more efficient resource usage whilst yielding restoration times comparable to link restoration. The basic idea is applicable to not only optical networks but also other mesh-based networks.

As there are efforts to develop an IP-centric control plane, defined within the Generalised Multi-Protocol Label Switching (GMPLS) framework, to manage the next generation optical network, ASPR is designed to utilise its signalling protocols to realise automatic resilience provisioning. Since the GMPLS framework supports both Packet Switched Capable (PSC) interfaces and non-packet switched capable interfaces including TDM capable, Lambda Switched Capable (LSC), and Fibre Switched Capable (FSC) interfaces, ASPR is applicable to not only optical networks but also other circuit-switched and packet-switched networks including ATM networks, Frame Relay (FR) networks, and IP/MPLS networks, etc. Therefore, the following description and discussion of ASPR are not confined to optical networks. Its possible application in other types of networks is presented as well. In addition, in order to have a comprehensive investigation, the schemes that are used for comparative studies with ASPR include not only the two basic schemes (link restoration and path restoration) which are widely used in mesh-based networks, but also two newly MPLS restoration schemes proposed within Internet Engineering Task Force (IETF).

There are two reasons to compare ASPR with these MPLS restoration schemes. One is to evaluate the application of ASPR in MPLS networks. The other is that as MPLS may facilitate the convergence of network functionality on a common control and management plane (GMPLS) for different types of networks, the MPLS-based restoration schemes could possibly be applied to optical networks.

This chapter is organised as the following: Section 5.2 introduces some basic terms used in MPLS / GMPLS restoration and the two MPLS restoration schemes proposed within IETF. Section 5.3 presents the novel restoration scheme ASPR with illustrations of its applications in different types of networks. Section 5.4 describes the implementation of the simulation models. Section 5.5 gives the comparative study results. Some related work is discussed in Section 5.6. The summary of this work is in Section 5.7.

5.2 MPLS / GMPLS Restoration Context

5.2.1 MPLS / GMPLS Restoration Basics

In MPLS / GMPLS networks, a primary / working path is also called a primary / working LSP whilst a backup path is called a backup LSP.

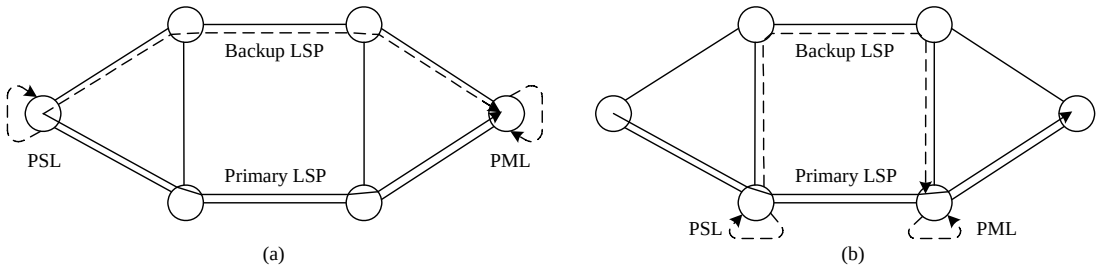


Figure 5-1: MPLS Restoration

The LSR that is responsible for switching the traffic from the primary / working path to the backup path is called the Path Switch LSR (PSL).

A Path Merge LSR (PML) is an LSR that is responsible for receiving the backup path traffic and either merges the traffic back onto the working path (Figure 5-1 (b), or, if it is itself the destination, passes the traffic on to the higher layer protocols (Figure 5-1 (a)).

Two mechanisms have recently been proposed for the restoration of Label Switched Paths (LSP) set up in MPLS networks, namely the RSVP Backup Detour [GAN01] and the Fast Rerouting scheme [FAR01][HAS00].

5.2.2 RSVP Backup Detour

Extensions to RSVP have been made to incorporate the concept of LSP tunnels into the RSVP flows. Together, these make it possible for routers using RSVP to create detours that

can route around downstream links and nodes. As a result, a LSP can quickly and automatically use an alternative by redirecting the user traffic to the pre-computed and pre-established detour routes in event of network link and node failures.

To protect from potential downstream link or node failures, a detour may be setup between the current node and one of the downstream nodes. The current node can be any node along a LSP except the LSP's egress, for the obvious reason that the egress node has no downstream link or node failure to speak of. Any downstream node of a LSP, which is more than one hop away from the current node, can be the detour merge point. For a penultimate node, only the immediate downstream link needs to be protected, so the egress node is the detour merge point.

In Figure 5-2, detour AC is created to protect against failures of link AB, link BC and node B. This detour path is computed and initiated at node A, and merges at node C. In this case, node B and D are not aware of the detour.

Likewise, detour BD is created by node B, to protect against failures of link BC, CD and node C. Detour DE is established to protect a local failure of link DE only.

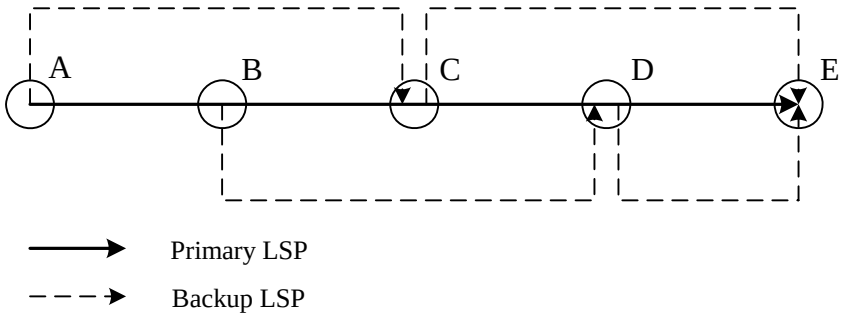


Figure 5-2: RSVP Backup Detour

This scheme provides fast and effective protection against all possible failures and is easy to implement. However, it requires much more resource redundancy than other schemes.

5.2.3 Fast Reroute

Fast Reroute approach is to reverse traffic at the point of the failure back to the ingress node of the protected LSP and redirect it via an alternative pre-configured LSP. This mechanism involves the setting up of two backup paths (separate from the working path). One of these backup paths, called the reverse path, runs in the opposite direction to the working path, from the penultimate node to the ingress node, via the same nodes that are along the

working path. The second backup path is established from the ingress node to the egress node via nodes that are path and link disjoint with the working path. When a failure arises, traffic is first redirected along the reverse path to the ingress node and from there it is forwarded along the alterative backup LSP.

For example in Figure 5-3, a revert backup LSP DCBA and an alternative LSP from A to E are established to protect against all possible link and node failures of LSP ABCDE.

This scheme is known for its low packet loss as it reroutes traffic near the fault without prior notification to the ingress node. However, this merit is obtained by introducing more traffic delay as the much longer route the restored traffic is to experience.

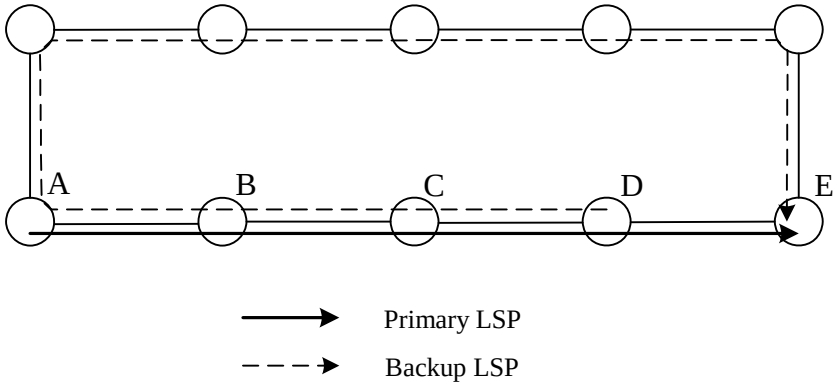


Figure 5-3: Fast Reroute

5.3 Adaptive Segment Path Restoration Scheme

In this chapter a novel resilience-provisioning scheme, entitled Adaptive Segment Path Restoration (ASPR) is proposed. It establishes backup LSPs for a given primary LSP using the standard MPLS/GMPLS signalling protocols. The basic idea is to divide a LSP into several segments. For each segment of the primary path, a backup path is provided. The segmentation of the primary path is adaptive to the topology of the network, allowing for more efficient resource usage whilst yielding restoration times comparable to link restoration.

ASPR is performed together with the deployment of primary / working LSPs; it consists of two phases. First, during the propagation of the forward signalling messages, the primary path is divided into several segments according to the topology of the network. For each segment a separate backup path is calculated. These actions are carried out along with the propagation of the forward signalling messages of the primary / working LSP. Then, along with the backward signalling messages, the segmentation of the primary LSP and backup

LSPs are further amended adaptively to the topology of the network. The backup LSPs are deployed only after the primary LSP is established.

Although this scheme does not have a particular requirement on the signalling protocol, and is capable of working with both CR-LDP and RSVP-TE, here only its application using CR-LDP is illustrated.

5.3.1 MPLS/GMPLS Traffic Restoration Cycle

MPLS / GMPLS traffic restoration occurs when there is a failure in one or more network components. A whole MPLS / GMPLS traffic restoration cycle includes **Failure Detection, Fail Notification** and **Traffic Restoration**. A more detailed discussion about the MPLS / GMPLS restoration cycle can be found in [SHA01]. The restoration time is defined as the interrupted period before the traffic is completely restored. It can be calculated as following:

$$T = t_d + t_n + t_s + t_r \qquad (1)$$

Here t_d is the time between the network link failure and when the failure is discovered by the MPLS / GMPLS restoration mechanism. This time may highly depend on lower layer protocols and usually is a constant for a given network.

t_n is the time taken by the notification message to travel from the LSR that detects the failure to the Path Switch LSR (PSL) which takes charge of the traffic switching to the backup LSP. In order to reduce the restoration time during its propagation, the notification message is usually assigned the premier priority during its transmission. After receiving a notification message, an intermediate LSR will forward it immediately without putting it in queue buffers. Thus t_n contains mainly the speed-of-light propagation delay and is proportional to the length of the primary LSP it protects.

$$t_n \propto L_{prim} \qquad (2)$$

t_s is the time consumed by the PSL which takes charge of the traffic switching from the primary LSP to the backup LSP. The operation of switching traffic to the backup LSP varies a lot in time and depends on the communication layer the MPLS / GMPLS signalling is associated with. In higher layers, the operation is to refresh the Label Mapping Table (IP layer) or Circuit Mapping (ATM and FR). In lower layers, it may consist of a cross-connect

action of an OXC (in the optical layer). However it varies, for different restoration schemes in the same network, t_s can be treated as a constant.

t_r represents the time taken by the diverted traffic travelling along the backup LSP till merging back into the Path Merge LSR (PML) after the failure point. This period is determined by the propagation delay, queuing delay, transmission delay and processing delay if MPLS restoration is performed in higher layers. In lower layer the queuing delay is replaced with the time taken by cross connect action of the physical switching unit, such as within an OXC in the optical layer. Compared to propagation delay and queuing delay, the processing delay is very small and can be omitted. Propagation delay is proportional to the length of the backup LSP. Thus,

$$t_r \propto L_{backup} \qquad (3)$$

In advance of the commencement of traffic flows, Service Level Agreements (SLAs) are decided to satisfy the QoS requirement of different services. The traffic flows are treated with different priorities in buffers, typically resulting in experiencing different delays. However, for the same service, the queuing delay can be treated approximately proportional to the number of hops of the backup LSP. Thus,

$$t_r \propto H_{backup} \qquad (4)$$

From (1), (2), (3) and (4), the following can be obtained

$$T \propto L_C, H_{backup} \quad where \quad L_C = L_{prim} + L_{backup} \qquad (5)$$

From (5), the following conclusions can be made:

1. The total restoration time is proportional to the sum of the length of primary and backup LSP, which we call **Restoration Length**.

2. The total restoration time is approximately proportional to the number of hops in the backup LSP.

3. With all the traffic parameters being the same in a given network, a restoration path deployment that has a shorter **Restoration Length**, L_C, and smaller backup path hops, H_{backup}, means a shorter restoration time.

4. Given two restoration path deployments, where the **Restoration Length** L_C is same, the one with smaller backup path hops H_{backup} provides a faster restoration.

Restoration can take place either at the source of a flow or close to the point of failure. The analysis confirms the common understanding that resolving the problem in the immediate vicinity to the fault provides faster restoration. That is why most restoration schemes try to reroute the traffic near the fault. However, the length of the backup path should also be considered when examining these schemes. According to our analysis, the actual restoration time is related to not only length from the fault to the traffic switching point, but also the length and hops of the backup path.

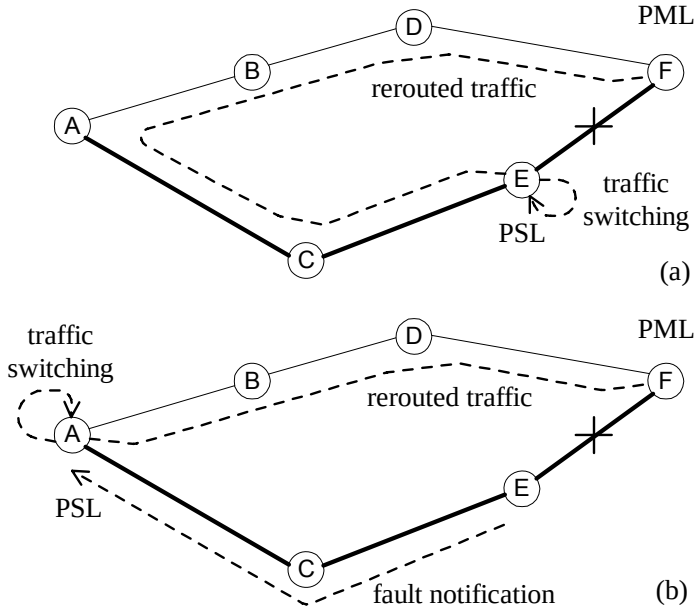


Figure 5-4: MPLS Traffic Restoration

Take for example in Figure 5-4. Suppose there is a primary LSP running along *ACEF*. Now the problem is that, given a link fault between *E* and *F* whether we should restore the traffic by switching the traffic at LSR *E* or at LSR *A*, which are illustrated in case (a) and (b). Most proposed MPLS restoration schemes (including Fast Reroute and RSVP backup Detour) adopt the first method, and reroute the traffic close to the fault. However, according to the analysis before, it is the downstream node that decides whether the traffic has been restored. Just switching the traffic to backup path does not necessarily mean that the restoration process has finished. If the rerouted traffic does not reach the PML *F* in time, the receiver of the flow can possibly drop connections. While for the second method in case (b), if the fault notification messages are given a higher priority than all the rest of the traffic, it will make sure that

$$t_{nb} + t_{rb} \leq t_{ra} \quad (6)$$

Thus,

$$T_b \leq T_a \quad (7)$$

That means case (b) has a better performance than case (a) in term of restoration time. In addition, it has the advantage of being more cost and management efficient because it does not need a spare backward path from *E* to *A* and a backup path *ABDF* can be shared to protect several fault scenarios which can arise at links: *AC*, *CE* or *EF*.

Another important issue is packet loss and delay. Case (a) has a better performance regarding packet loss. However, it is at the expense of increased traffic delay. In this case, the restored traffic has a longer path *ACECABDF* comparing with *AFBF* in case (b). It means that although case (a) reroutes the traffic quickly with less packet loss, it brings longer delay for the traffic.

Many customers do not want the longer delay. This is because most services with high reliability requirements can be placed into two categories. One example is mission critical services, such as financial data transfers. Another example is time critical services, such as IP-telephony and real-time video. These services are either sensitive to packet loss or to delay. Few are to both. Usually mission critical services are sensitive to packet loss while time critical services are sensitive to packet delay.

The benefit of case (b) is that once recovery is complete, the subsequent traffic delay will be typically much less than that in case (a), which is preferred by the time critical services. Although during the interim period packet loss may be greater, so long as the connection is maintained, higher layer protocols can be used to retransmit the missing information to satisfy the mission critical services. With reference to Figure 5-4, therefore the author proposes that traffic rerouting should take place at LSR *A*, which is adopted by the scheme of ASPR.

5.3.2 Adaptive Segment Path Restoration Algorithm

In order to restore the traffic within the time required by service whilst making more efficient use of the network's resources at the same time, ASPR divides the primary LSP into several segments. The end nodes of each segment are termed as **Segmentation Points**. The segmenting principle is that all the LSRs that have the same **Restoration Length** are put in the same segment. Then in each segment, a backup path is found to cover possible link failures within this segment. The purpose is to make the **Restoration Length** and backup

hops satisfy the QoS requirement of the different services being transported. The segmenting of the LSP is adaptive to the topology of the network and further to QoS requirement of each service.

Assume a MPLS network with a topology represented by the graph G (V, L), where V is the set of v nodes and L is the set of l links between the nodes. Furthermore, assume that graph G is two-edge redundant and therefore can be protected against any single link failure. A LSP P is a sequence $<v_1, l_{12}, v_2, l_{23}, \ldots, v_i, l_{i,i+1}, v_{i+1}, \ldots, v_{k+1}>$, where $l_{i,i+1}$ is a link with endpoints v_i and v_{i+1} (for $i = 1, \ldots, k$), v_1 is the ingress node and v_{k+1} is the egress node. The algorithm attempts to divide the primary LSP into several segments and deploys a backup LSP for each segment, respectively. Given a primary LSP P and the current node is denoted as v_1, Figure 5-5 shows how the next segmentation point is located.

```
begin SP(G, P)
    S_P = v2
    Lprim = l12
    Find backup LSP to v2:  Pb
    Lbackup = Length (Pb )
    LC= Lprim +Lbackup
    for i =2 to |P|-1 do
        Lprim + = li,i+1
        Find backup LSP to vi+1  Pb
        Lbackup = Length(Pb)
        if LC < Lprim +Lbackup
            break
        else
            S_P = vi+1
            LC = Lprim +Lbackup
end
```

Figure 5-5: Segmentation Point Location Procedure

In the algorithm, *S_P* represents the **Segmentation Point** and is initialised as next hop to the current node on the primary LSP. L_C is initialised as the **Restoration Length** of next hop. The algorithm is to find the farthest node to the current node on the primary LSP that has the same **Restoration Length** as that of the anterior nodes.

When the forward signalling message for path setup arrives the **Segmentation Point**, the procedure is performed again to find the next **Segment Point** until the path setting up message reaches the egress node. In this way, a LSP can be divided into segments automatically.

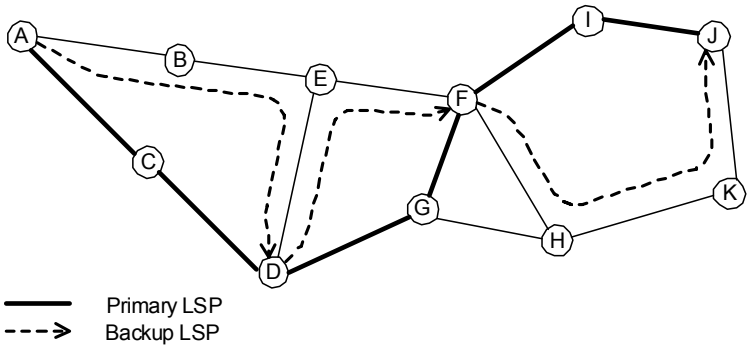


Figure 5-6: ASPR Path Segmentation Example

Figure 5-6 shows an example of how a **Segmentation Point** is found and the primary and backup LSPs are calculated. Given a primary LSP (*ACDGFIJ)*, the Segmentation Point Location Procedure is performed at the ingress LSR to find the next hop **Segmentation Point**. In this example, LSR *C* and LSR *D* have the same **Restoration Length** for the given LSP. Then LSR *D* is chosen as the **Segmentation Point**. When the next hop **Segmentation Point** has been defined, the backup LSP from current node, LSR *A*, to the **Segmentation Point**, LSR *D*, is calculated. In this case, it is *ABED*.

5.3.3 Setup of the Primary Path

When the ingress LSR receives a request to set up a primary LSP, the Segmentation Point Location Procedure is performed to find the next hop **Segmentation Point** (the first one). Then, the backup LSP of the first segment of the LSP (from the ingress node to the first **Segmentation Point**) is also calculated and stored with the pending Label Request Message in the ingress node. Deployment of the backup LSP of a segment of the primary path will initiate only after the downstream part of the primary path has been established.

5.3.3.1 Segmentation Point TLV

After the first **Segmentation Point** has been found, a **Segmentation Point TLV** is created as shown in Figure 5-7. This Type-Length-Value (TLV) is inserted as an optional parameter into the Label Request Message (LRM), which is sent out by the ingress node.

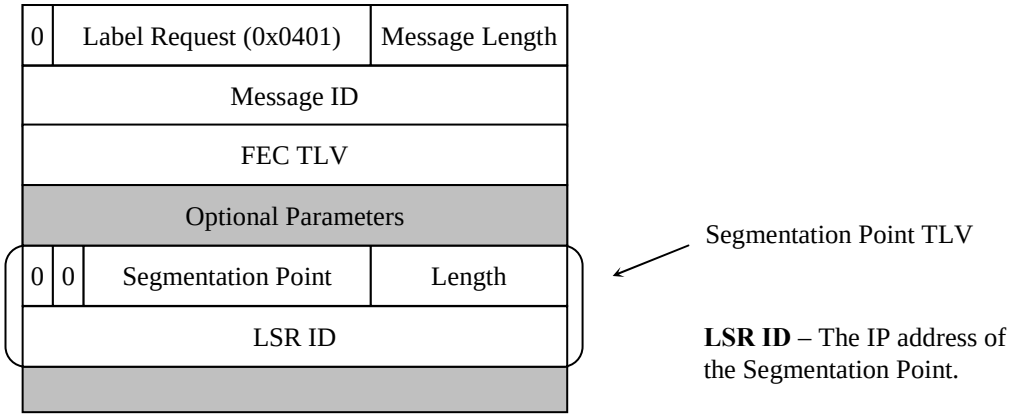


Figure 5-7: Segmentation Point TLV

5.3.3.2 Receiving a LRM Containing a Segmentation Point TLV

LRM procedures defined in [RFC3036] describe the routine actions to be taken when a LRM is received in a LSR. In order to implement the ASPR, updates to the LRM procedures for nodes except the egress node are illustrated in Figure 5-8.

When a downstream node except the egress node receives a LRM containing a **Segmentation Point TLV**, it checks if itself is the **Segmentation Point** denoted in the TLV. If not, the received **Segmentation Point TLV** is simply copied into the new LRM to be sent to the downstream node.

If the current node is the **Segmentation Point**, the Segmentation Point Location Procedure is executed to find the next hop **Segmentation Point**. Then, a backup LSP for the segment from current node to the next hop **Segmentation Point** of the primary path is calculated and the node list is saved with the pending Label Request Message. A new **Segmentation Point TLV** is created with the value set as the new **Segmentation Point**. The TLV is inserted into the new LRM that is then sent out to the downstream node.

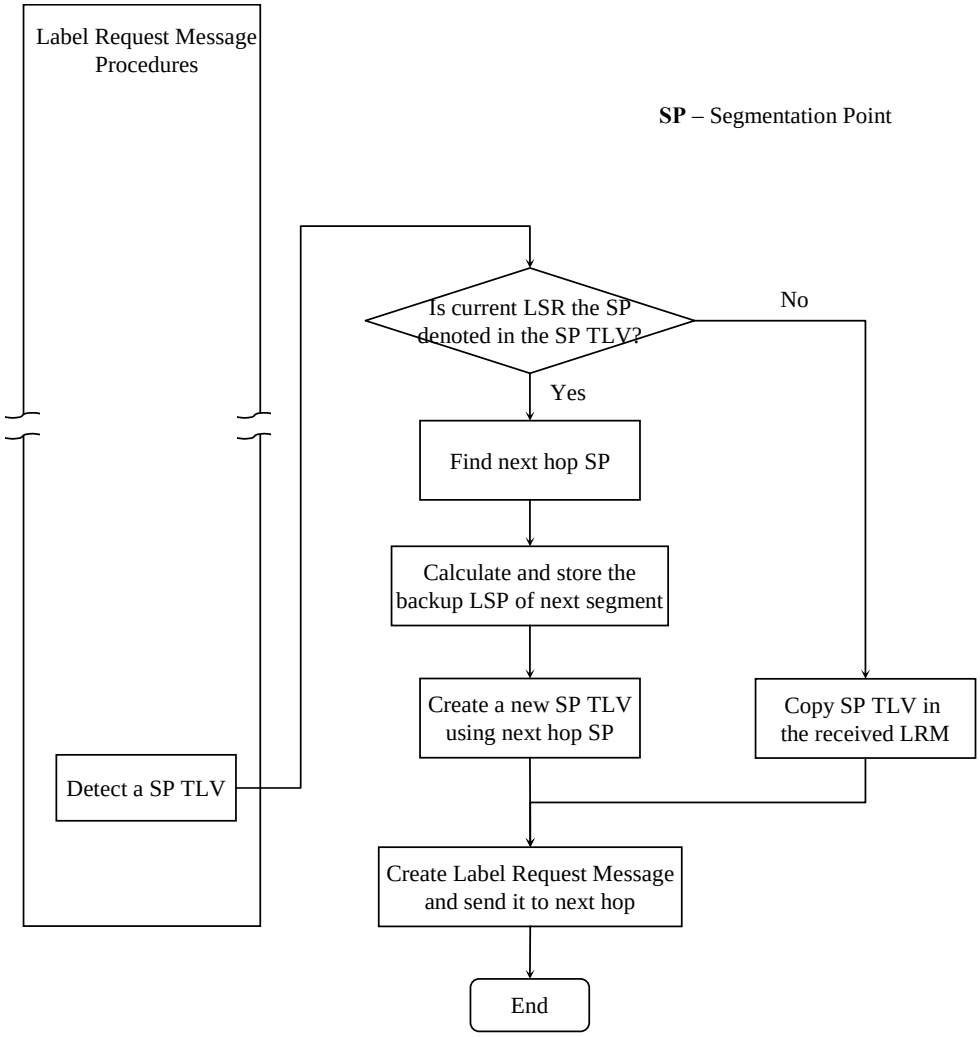


Figure 5-8: Receiving a LRM containing a Segmentation Point TLV (except the egress node)

The egress node is the last **Segmentation Point** of the primary LSP. If an egress node receives a LRM containing a **Segmentation Point TLV**, the value in this TLV must be the ID of the egress node. When a LRM reaches the egress node, the preliminary segmentation of the primary LSP is finished.

Following the LRM procedures defined in [RFC3036], a Label Mapping Message (LMM) is then created. In the ASPR, a new TLV termed **Primary Segment Path Vector TLV** is proposed to be put in the LMM to be sent to the upstream node.

5.3.3.3 Primary Segment Path Vector TLV

The **Primary Segment Path Vector TLV** (Figure 5-9) records the node list of current segment of the primary LSP. This information is to be used at the time of establishing the corresponding backup LSP, to decide whether the backup path can share resources that are used by other backup paths. The collection of this information is carried out along the

propagation of the LMM. Initially, the egress node ID is inserted into the **Primary Segment Path Vector**. The existence of a **Primary Segment Path Vector TLV** in a Label Mapping Message will also trigger the additional procedures for the ASPR. The specific operations of growing the **Path Segment Path Vector** are introduced in section 5.3.4.2.

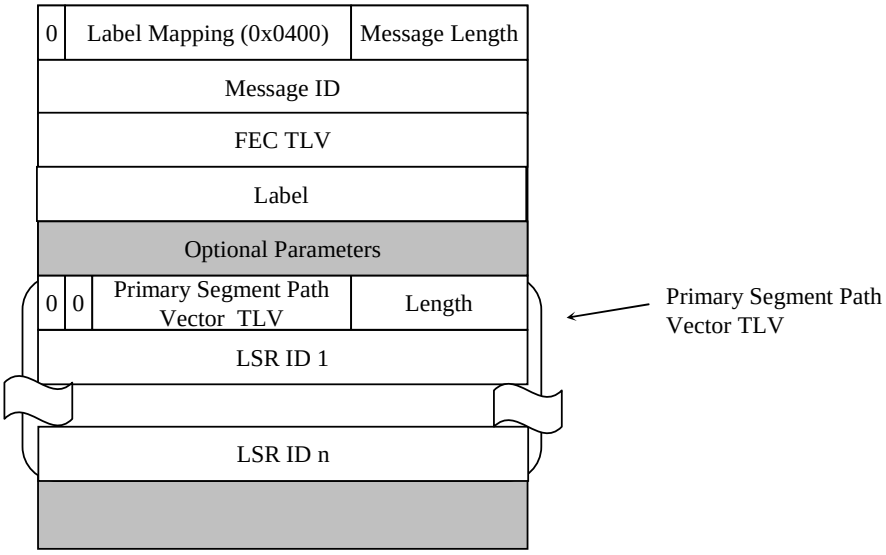


Figure 5-9: Primary Segment Path Vector TLV

5.3.4 Segment and Backup Path Refinement

In most circumstances, **route rewind** will take place for the backup LSPs, which is shown in Figure 5-10(a). Here, LSR *D* and *E* are set as the **Segmentation Points** and backup LSP origins for each segment path are calculated. As analyzed earlier, it is better to set LSR C and F instead of LSR D and E as the PSL for the second and third segments. Therefore, certain amendment is needed for the backup LSP to avoid **route rewind**. In this case, the saved backup LSPs at the **Segmentation Points** then may no longer be valid. To adjust the backup LSPs, a new optional parameter called **Backup Explicit Route TLV** is inserted in the Label Mapping Message.

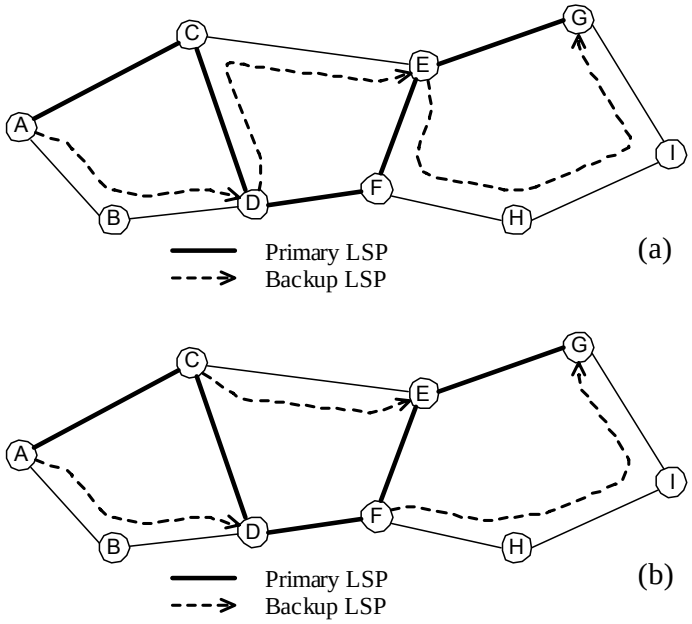


Figure 5-10: Segment and Backup Path Refinement

5.3.4.1 Backup Explicit Route Vector TLV

The **Backup Explicit Route Vector TLV** (defined in Figure 5-11) records the node list of the backup LSP for current segment of the primary LSP. It is initiated as the saved backup LSP at the Segmentation Point and is further adjusted along with the propagation of the Label Mapping Message if required. The specific operations are introduced in next section.

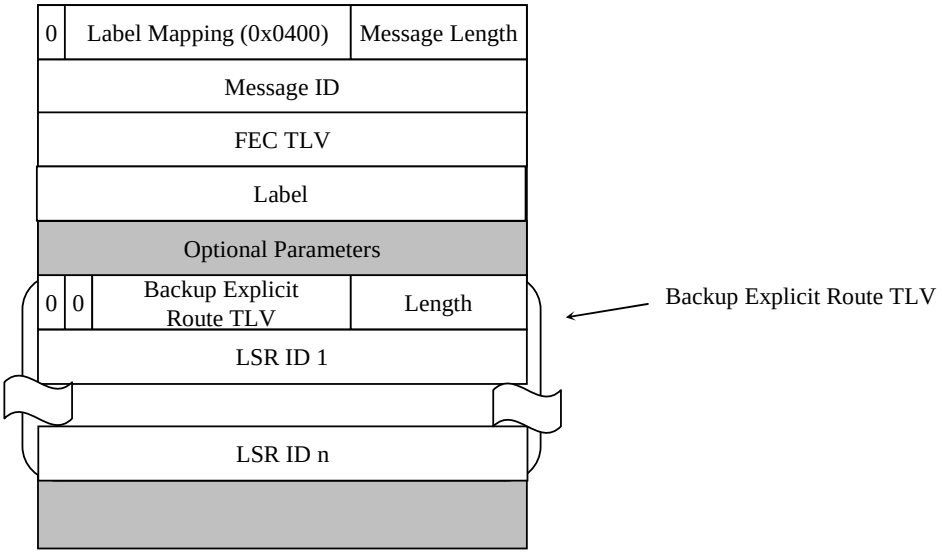
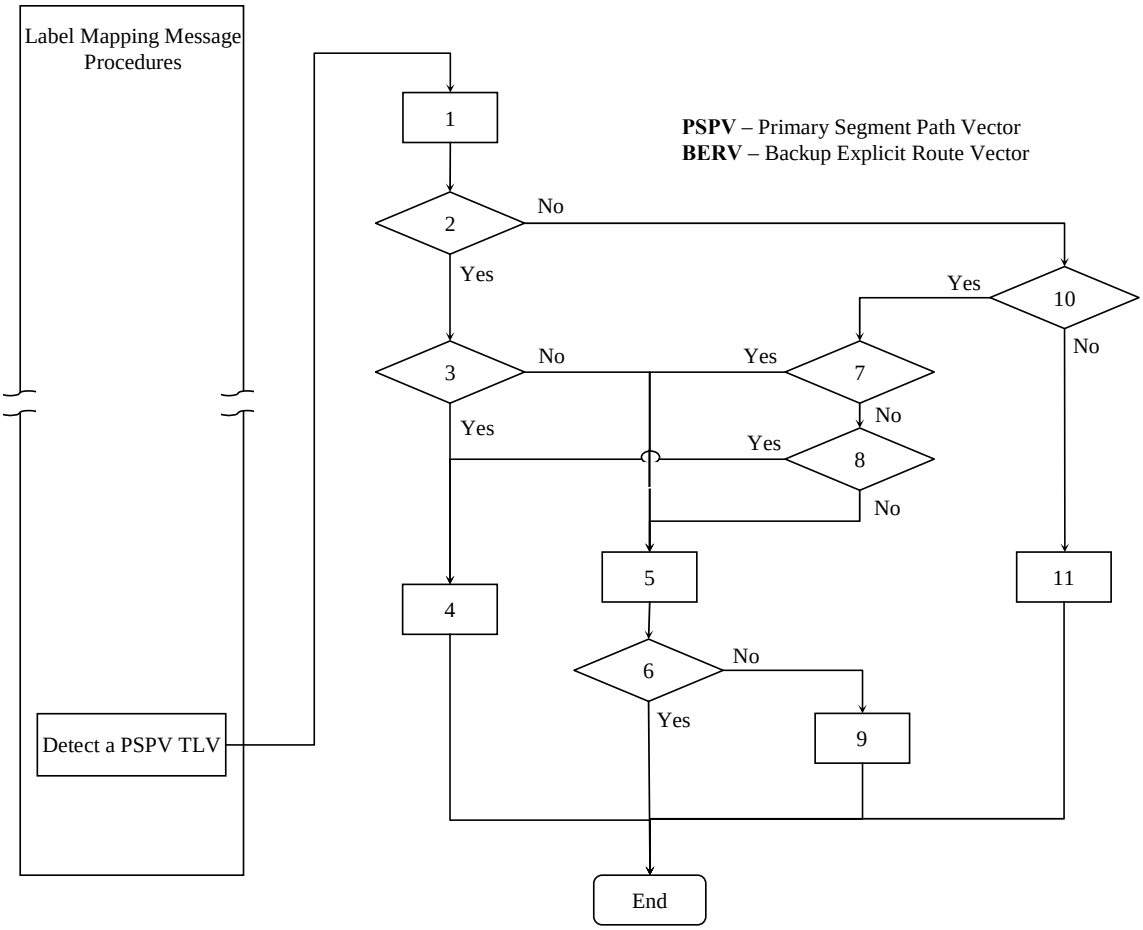


Figure 5-11: Backup Explicit Route Vector TLV

5.3.4.2 Refinement Procedures

When the Label Request Message for the primary LSP reaches the egress LSR, the egress LSR assigns a label for the LSP and creates a Label Mapping Message with a **Primary Segment Path Vector TLV**, which is then sent to the upstream node.



1 – Insert current LSR ID into the PSPV TLV node list.
2 – Is there a BERV TLV in the received Label Mapping Message?
3 – Does the next hop of Label Mapping Message have the same LSR ID as that of the next hop of the backup LSP (denoted by BERV TLV)?
4 – A new Label Mapping Message is created with the new PSPV TLV. Delete the first hop of backup LSP and create a BERV TLV and insert it in the new Label Mapping Message. Send the Label Mapping Message to the upstream node.
5 – The current LSR is the PSL for current segment. Initiate deployment of the backup LSP for current segment which is either denoted by BERV TLV or saved with the pending Label Request Message. The PSPV TLV is also inserted into the Label Request Message for the backup LSP to indicate the segment it intends to protect
6 – Is current node the ingress node?
7 – Is current node the ingress node?
8 – Does the next hop of Label Mapping Message have the same LSR ID as that of the next hop of the backup LSP?
9 – Create a new PSPV TLV with only current LSR ID in the node list and put it in the Label Mapping Message. Send the Label Mapping Message to the upstream node.
10 – Is there a saved backup LSP node list with the pending Label Request Message?
11 – Create a Label Mapping Message with the new PSPV TLV and send it to the upstream node.

Figure 5-12: Refinement Procedures

When a LSR receives a Label Mapping Message containing a **Primary Segment Path Vector TLV**, additional procedures to that defined in [RFC3036] are executed as shown in Figure 5-12.

Note here that the **Primary Segment Path Vector** records a segment of the primary path and the **Backup Explicit Route Vector** records its backup path. The refinement procedures are to make the segment of the primary path as long as possible and its backup path as short as possible within the same **Restoration Length**. In this sense, the refinement is like a backward segmentation.

Figure 5-10(b) shows the result of the refinement of the backup LSPs and the segmentation. Here, link *AC* is protected by backup LSP *ABD*. Link *CD* and *DF* are protected by backup LSP *CE*. Link *FE* and *EG* are protected by backup LSP *FHIG*.

5.3.5 Bandwidth Sharing and Setting Up the Backup Path

Resource sharing has been proven to be a very effective method to reduce the network costs [DOVC][DOV99]. It assumes that in a certain period, only singular separate failures are likely to arise. Thus, backup LSPs can share resources if they do not protect the same primary links.

During the setting up of the backup LSP, a record of the fault links, which it intends to protect, is carried by the **Primary Segment Path Vector TLV** in the Label Request Message. When a LSR is going to reserve bandwidth for the backup LSP, the algorithm shown in Figure 5-13 is performed to see if it can share resource with other backup LSPs.

```
begin LS(LST)
    for i = 1 to B_S do
        for j = 1 to Φ_P do
            share_f = false
            if L_i^P ∩ L_S ≡ φ
                L_i^P = L_i^P ∪ L_S
                share_f = true
    if share_f ≡ false
        Φ_P ++
        L_{Φ_P} = L_S
end
```

Figure 5-13: Resource Sharing Procedure

Here, a **Link Sharing Table (LST)** is created and stored for each port of each LSR. Let Φ_P be the total bandwidth reserved by the backup LSPs on port *p*. In **LST**, each bandwidth unit is denoted by a set of links L_i^P that it protects. When there is a request for the establishment a backup LSP (B_S denotes its bandwidth requirement), the primary links it intends to protect, denoted by L_S, is examined if the required bandwidth can be shared with other backup LSPs.

5.4 Simulation Models

Simulation is used in order to evaluate the performance of the proposed scheme ASPR against other schemes. The functions of ASPR described in former sections are implemented using OPNET™. In addition, other schemes such as Link Restoration, Path Restoration, Fast Reroute and RSVP Backup Detour are also implemented for comparative study.

5.4.1 Network Models

Figure 5-14 shows an example of the implemented network model. Each node in the network represents a GMPLS LSR that supports different data plane interfaces. To facilitate the performance evaluation, all nodes are assumed to have full wavelength conversion capability when the schemes are applied in the optical network. That is, a wavelength in one colour can be converted into any other colour if required.

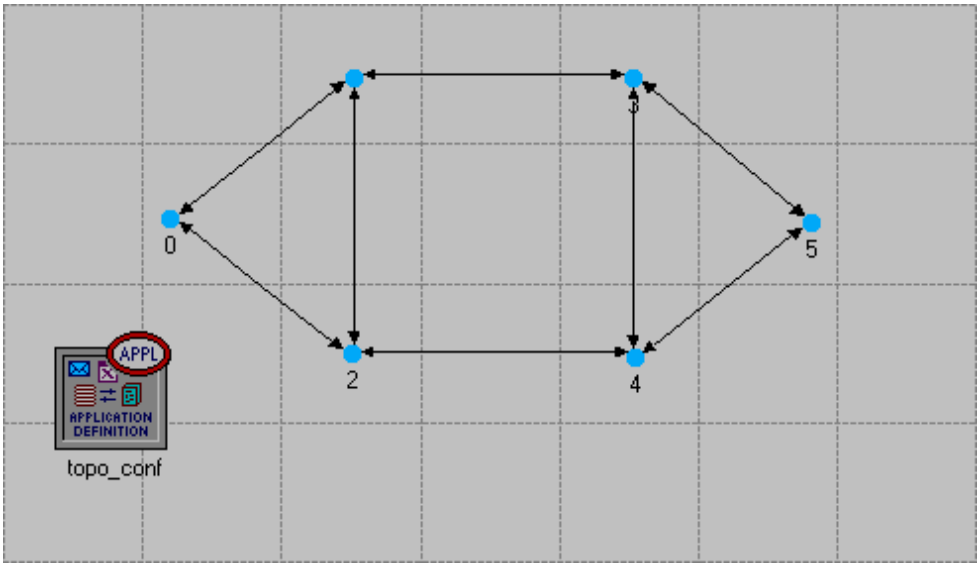


Figure 5-14: An Example of Simulation Network Model

These nodes are connected by physical links the bandwidth of which is set as infinite. The deployment results of these links are measured using a bandwidth unit, which varies in different applications. For example, a unit could be a wavelength channel when the scheme applies to optical networks.

5.4.2 Node Model

Each node (LSR) can initiate a connection request of which the destination node is randomly selected from the node set of the simulation network. It contains the basic functionalities of GMPLS CR-LDP and all other functionalities required to realise the resilience-provisioning schemes including ASPR, Link Restoration, Path Restoration, Fast Reroute and RSVP Backup Detour.

5.4.3 Verification and Validation

Having developed a simulation model, the node needs to be verified and validated. Verification determines whether the model does indeed perform as intended and validation shows whether the model is a true and accurate representation of the system modelled [PIT93]. This needs to be carried out at two levels, the first on a fine scale by looking at individual objects that make up the network and then at the whole network.

The verification and validation of the simulation models for ASPR is carried out from the following aspects.

First, the functionality performance of the simulation models are thoroughly tested using debugging tools provided by OPNET^{TM} during model implementation.

Second, an example network with a simple topology as shown in Figure 5-14 and certain traffic demands are used to validate the outcome of each scheme against the expected results.

Third, as Link Restoration and Path Restoration are extensively studied by other researchers with respect to other circuit-switched networks, there are already some well-accepted calculation and simulation results of these schemes applied to some well-known networks. In order to compare the results with that of other researchers, all network topologies used for performance evaluated are widely used by researchers. Therefore, part of the performance results (Link Restoration and Path Restoration) are verified and validated by comparing with some published results of other researchers using similar prerequisites.

5.4.4 Confidence Interval

System models that include stochastic behaviour have results that are dependent on the initial seeding of the random number generator. As a particular random seed selection can potentially result in an anomalous or non-representative behaviour, it is important for each model configuration to be exercised with several random number seeds, in order to be able to determine standard or typical behaviour. The basic principal applied here is that if a typical behaviour exists, and if many independent trials are performed, it is likely that a significant majority of these trials will fall within a close range of the standard.

Therefore, in the following section, each simulation is performed ten times using different initial seeds. The results are the average of the ten simulations. For example, the result of Spare Capacity Requirement of ASPR for the Toronto Metropolitan Network is shown in Table 5-1.

Initial Seed	Seed_128	Seed_97	Seed_34	Seed_8	Seed_71
SCR	1.24894	1.25222	1.27743	1.2639	1.26898
Initial Seed	Seed_25	Seed_210	Seed_11	Seed_7	Seed_92
SCR	*1.27311*	*1.27085*	*1.2699*	*1.25947*	*1.27078*

Table 5-1: Spare Capacity Requirement of ASPR for the Toronto Metropolitan Network

The 95% confidence interval is calculated as follows:

The mean is calculated as

$$\overline{X}(n) = \frac{\sum_{i=1}^{n} X_i}{n} = 1.26558$$

The sample variance is calculated as

$$\delta = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \overline{X}(n))^2}{n-1}} = 0.009296$$

The 95% confidence interval with a T-distribution is given by

$$\overline{X}(n) \pm t_{\alpha,N} \frac{\delta}{\sqrt{N}} = 1.26558 \pm 0.00665$$

As other parameters of the simulation results present similar small confidence intervals, only the average results of the ten simulations with different initial seeds are plotted in the result graphs for comparative study.

5.5 Performance Evaluation

Extensive experiments have been carried out to analyse the effectiveness of the proposed algorithm in a wide variety of network environments. Four real networks (Network 1: New Jersey LATA Network. Network 2: NSFNET. Network 3: U.S. Long-Distance Network. Network 4: Toronto Metropolitan Network) shown in Figure 6 are used in the comparison of different restoration schemes. Table 5-2 shows the parameters of the networks, in which ND (node degree) represents the network connectivity.

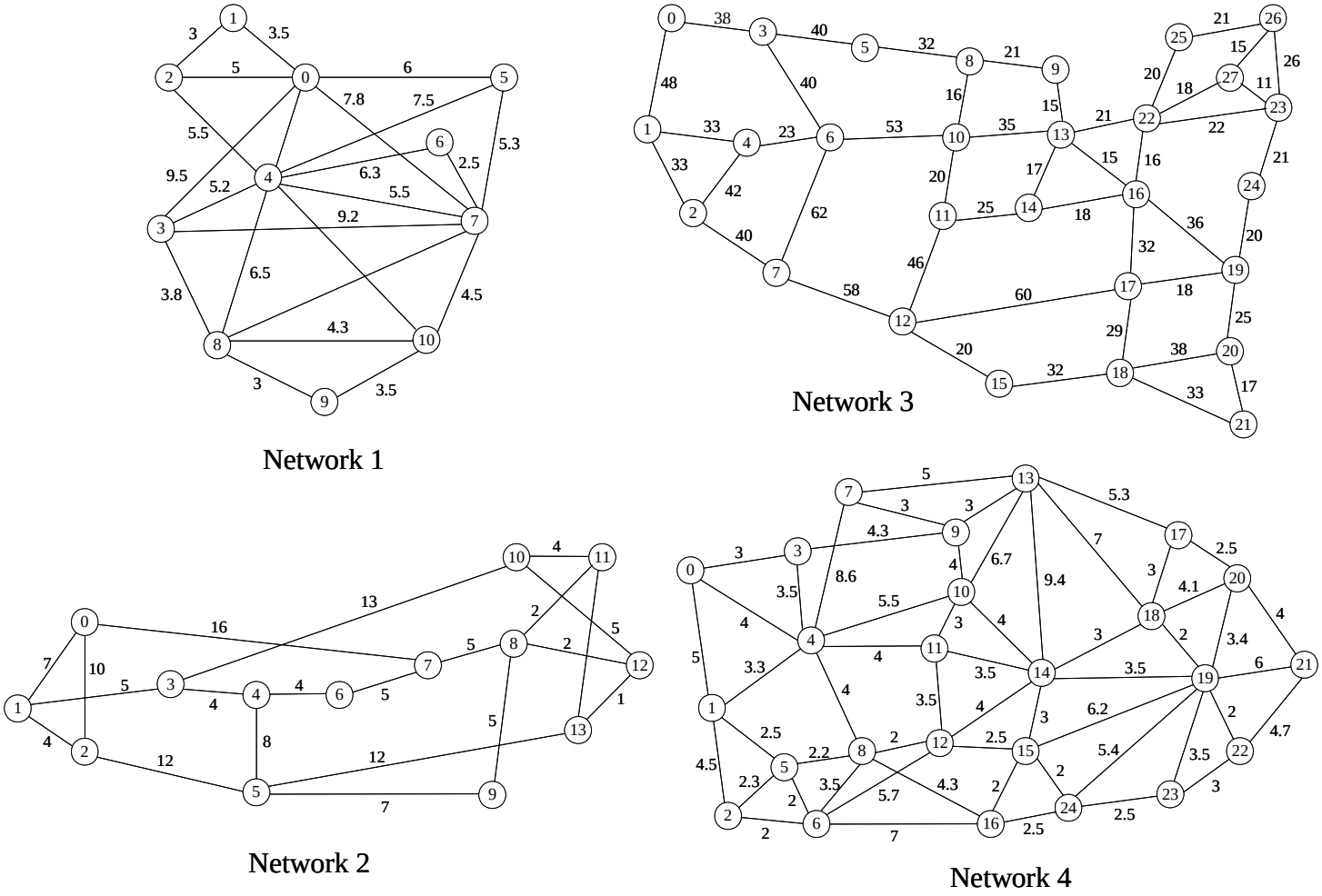


Figure 5-15: Examples of Network Topologies

Network	1	2	3	4
Node	11	14	28	25
Link	22	20	45	55
ND	4	2.86	3.21	4.4

Table 5-2: Network Parameters

The experiment implements the ASPR as well as other restoration schemes including: Link Restoration, Path Restoration, RSVP Backup Detour and Fast Reroute.

In all test networks, each node sets up 100 primary LSPs, of which the egress nodes are uniformly distributed over the set of network nodes. The bandwidth requirement of each flow is one unit. The shortest path algorithm is used to calculate the primary and backup LSPs. All the results are the average results of 10 simulations with different initial seeds.

5.5.1 Spare Capacity Requirement

The Spare Capacity Requirement is used to evaluate the cost efficiency of the restoration schemes. It is defined as the ratio of the total backup resource cost to that of the primary flows. The cost of each link is assumed to be proportional to its length as shown in Figure 5-15.

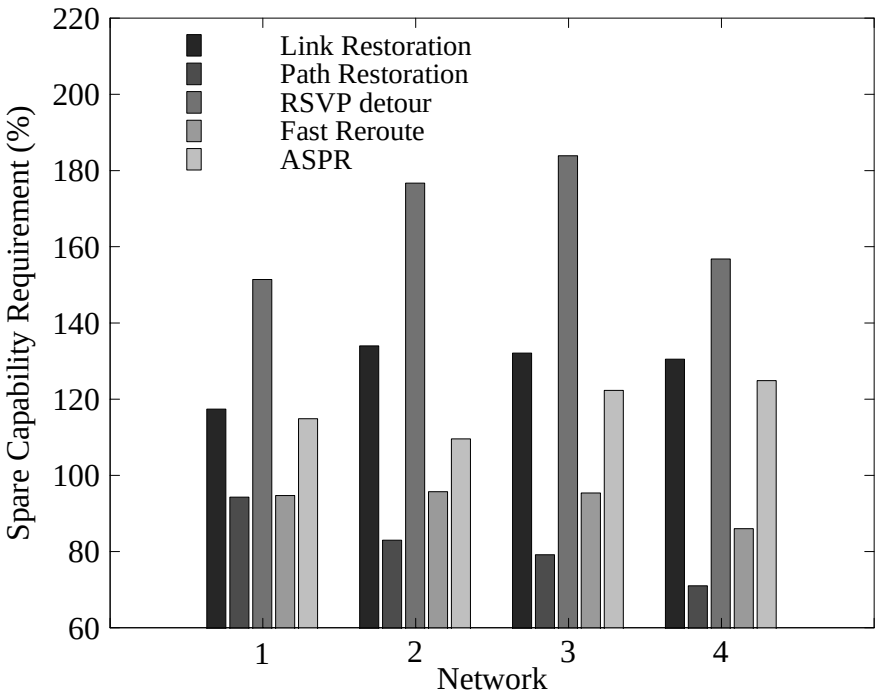


Figure 5-16: Spare Capacity Requirement

Figure 5-16 shows the results. Following can be observed,

(1) Path Restoration has the best performance in respect of the Spare Capacity Requirement.

(2) RSVP backup Detour has the worst performance and is even worse than that of Link Restoration. This is easy to understand. The RSVP backup Detour deploys a backup LSP at each LSR to the second next hop LSR, which makes the length of its backup LSPs nearly twice as that of the Link Restoration.

(3) ASPR has a better performance than Link Restoration. This is also expected since ASPR uses one backup LSP for all the possible links in the same segment, which results in better resource sharing.

(4) Fast Reroute has the second best Spare Capacity Requirement performance next to Path Restoration. Although it requires an additional reverse segment, it still provides good resource usage.

(5) Comparing to other schemes, ASPR performs better in a network with low connectivity (e.g. in Network 2). That is because in a low connectivity network, a segment

typically contains more links, which can share the same backup LSPs. While in a high connectivity network, the Segmentation Point Location Procedure may result in far fewer primary links per segment, which causes ASPR performance to be similar to link restoration.

5.5.2 Restoration Length

As explained before, the total restoration time is proportional to the sum of the length of primary and backup LSP, which we call the **Restoration Length**.

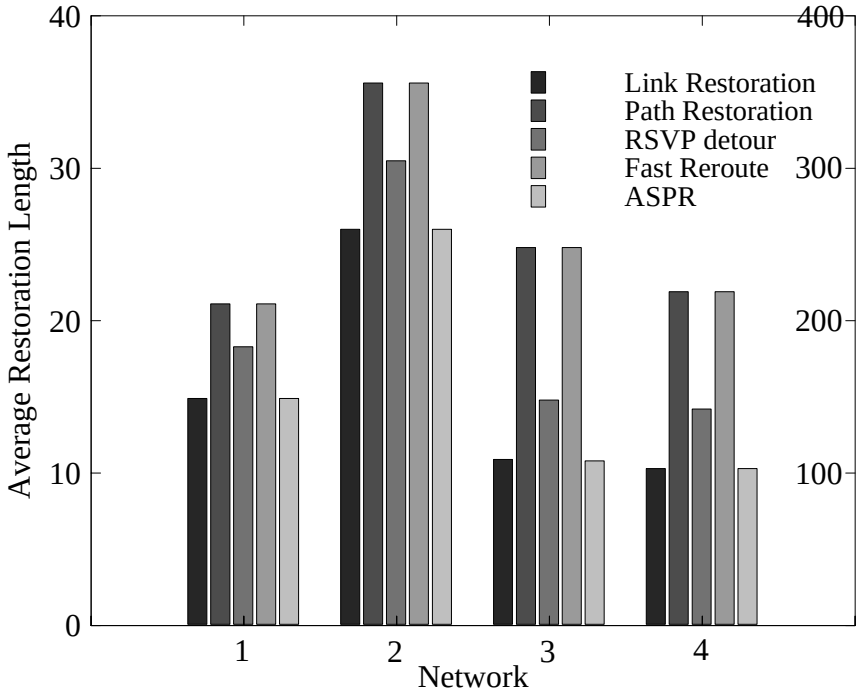


Figure 5-17: Average Restoration Length

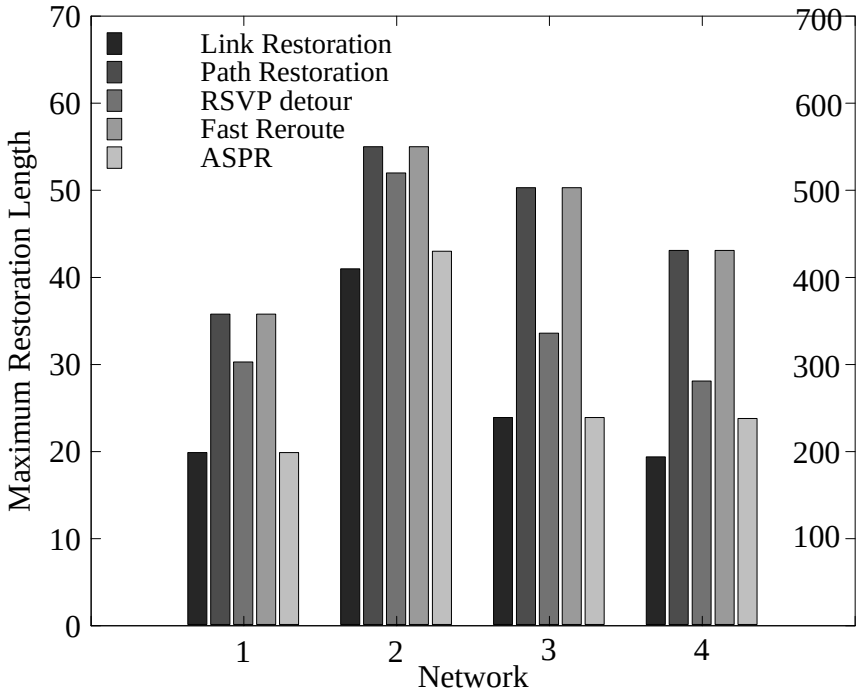


Figure 5-18: Maximum Restoration Length

Figure 5-17and Figure 5-18 show the average **Restoration Length** and the Maximum **Restoration Length** respectively. In these diagrams, the Y label for the networks 1, 2 and 4 is put on the left-hand side whilst that for network 3 is on the right-hand side. The following can be observed from these figures:

(1) ASPR and Link Restoration have the shortest average **Restoration Length**, nearly half that of Fast Reroute and Path Restoration. This means the restoration time of ASPR and Link Restoration is the shortest.

(2) Fast Reroute and Path Restoration have the longest average **Restoration Length**, which means their restoration times are correspondingly longer. The only difference between these two schemes is that Fast Reroute can restore the traffic quickly from the sender's point of view, thus the packet loss is lower. However, it is the receiver that decides whether the traffic has been restored or not. Therefore, from the receiver's point of view, there is not much difference in restoration time between Fast Reroute and Path Restoration. Fast Reroute's low packet loss is offset by the large propagation delays resulting from the typically longer restoration path lengths. This could be unacceptable to some delay-sensitive services.

(3) RSVP Backup Detours has a better performance than Fast Reroute and Path Restoration.

5.5.3 Backup LSP Hops

As analysed before, the restoration time is not only related to the **Restoration Length**, but also to the hop count of the backup LSP.

In higher layers (e.g. IP/MPLS), two or more incoming streams of a node can be mapped to one outgoing stream at the same time. Such a property enables virtual backup paths to be established before the failure to share the backup resource. Therefore, there is no need of signalling messages to notify the interim nodes of a backup path to set up the connection. However in these “soft” switched networks, more hops in a backup path mean more queuing delays the restored traffic expected to experience, which in turn will result in longer restoration time and traffic latency.

In the optical layer (as with other hard-circuit-switched layers), the sharing of resources between backup paths means that these backup paths cannot be established before the failure. Therefore, signalling message is required to notify the interim nodes (e.g. OXCs) to switch on the intended backup path. Such actions consume time, which makes up one of the major parts of the restoration time. Thus, more hops of backup paths mean more signalling latency and more OXC connection actions, which results in a longer restoration time and greater expense.

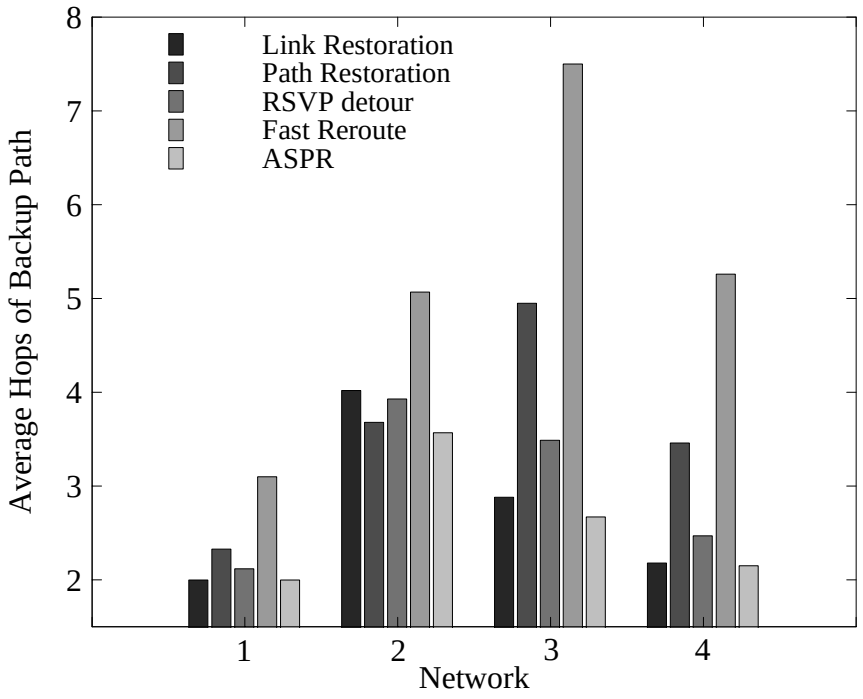


Figure 5-19: Average Hops of Backup Paths

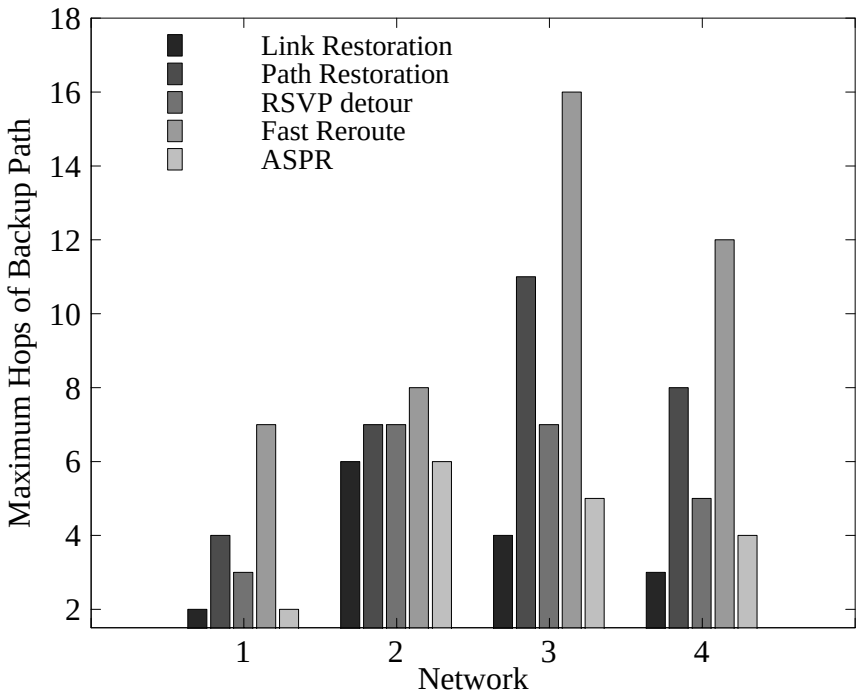


Figure 5-20: Maximum Hops of Backup Paths

Figure 5-19 and Figure 5-20 show the average backup LSP hop count and the maximum backup LSP hop count, respectively. The following can be observed,

(1) ASPR has the best average backup LSP hop count performance. For networks with high connectivity, such as network 1 and 4, the average hops of backup path of ASPR are only around 2 hops. For the worst case (network 2 with low connectivity), it is only 3.57 hops on average.

(2) Fast Reroute has the poorest performance. As it is based on the end-to-end restoration, the hops of backup path also depend on the size of the network. For a network with a big number of nodes and relatively low connectivity, such as network 3, its backup paths have 7.50 hops on average and 16 hops at maximum.

(3) RSVP Detour has a better performance than Fast Reroute and Path Restoration but a worse performance than ASPR.

5.6 Related Work

Related studies of dividing lightpaths into several protected segments can be found in [HO02][OU02].

In [HO02], a path is fragmented into nearly equal length parts to form a sequence of segments. Each segment is protected separately. At on extreme, when a path consists of only one segment, this approach is equivalent to path protection; at the other extreme, when each segment is a link, this approach is the same as link protection. However, the segmentation of a path, which is based on its length, does not consider the network topology to which it applies. The lack of awareness of network topologies often introduces route rewinds for the backup path. Moreover, this scheme requires manual interventions to help deploy the primary and backup paths.

In [OU02], a large mesh network is dimensioned into a set of non-overlapping areas, and a path traversing a sequence of such areas is divided into a sequence of segments, one per area. Each such segment is protected independently within its own area. Therefore, only diverting the traffic to the backup path in the area where a failure occurs can reduce the restoration time. However, this scheme requires offline algorithms to decide the dimensioning of protected areas. Another disadvantage of this scheme is that the segmentation of the whole path is not based on its length but on whether the path crosses the border of two areas or not. This often results in unnecessary segmentation for a short path that crosses the border by chance.

The proposed algorithm ASPR does not need any manual intervention and works automatically with the path establishment signalling protocol, which is important to the dynamic provisioning of resilience. In addition, the segmentation of the path is adaptive to the network topology, which avoids route rewinds for backup paths.

5.7 Summary

This chapter firstly presents a new method to investigate the performance of different MPLS/GMPLS restoration schemes. The **Restoration Length** is used as a measure of the performance of different schemes. In particular, the backup LSP hop count provides a measure of the relative efficiency of the alternatives. It also indicates the long run traffic latency.

Then, a novel MPLS/GMPLS restoration scheme called Adaptive Segment Path Restoration is proposed. The comparative study of ASPR with other schemes shows that, (1) the two suggested MPLS restoration schemes by IETF have their deficiencies. Fast Restoration has a better performance on cost efficiency and it has a merit of less packet loss. However its restoration time performance is the poorest. In addition, the backup path length and hop count are typically much larger, yielding latencies that could be unacceptable for the real-time services. RSVP Backup Detour has shorter Restoration Length and smaller hops of backup path than that of Fast Reroute, which means a faster restoration and less packet latency than that of Fast Reroute. However, its performance on spare capacity requirement is the poorest, and so poor (need average 170% redundancy) that it would be unacceptable to most service providers. (2) ASPR is shown to have the best restoration time performance with the least hop count on backup path, which has the least packet delay for the restored traffic. Both are favourable to the real-time services. In addition it is better than most of the other restoration schemes on Spare Capacity Requirement, and only next to Fast Reroute and Path Restoration.

Chapter 6 Differentiated Resilience Provisioning for Wavelength-Routed Optical Networks

This chapter presents a novel Differentiated Resilience Optical Services Model (DROSM), which suggests utilising differentiated resilience provisioning for the wavelength-routed optical network.

6.1 Overview

Despite recent events in the telecommunications industry, optical networks continue to grow rapidly to accommodate the rapid rise in data traffic brought on by new Internet and enterprise applications such as virtual private networks (VPNs) and e-commerce [BEN01]. At the same time, the introduction of optical networking with wavelength-division multiplexing (WDM) transmission technology, optical multiplexers and optical cross-connect (OXCs) devices, makes feasible the prospect of an "all-optical" Internet core. Configuring these devices enables one to establish all-optical connections, or lightpaths, between source and destination nodes. These all-optical lightpaths provide transparent data communication and offer potential for cost-savings by eliminating the electronic processing costs and bottlenecks at intermediate nodes.

In an all-optical WDM network, wavelength conversion capability adds to the cost and so must be used sparingly, so it is desirable for connections to use the same wavelength on all links along the route. This requirement is referred to as the wavelength continuity constraint. Lightpath-based WDM networks are generally referred to as wavelength-routed networks. If a common wavelength is not available on all links along the route, then the connection is blocked.

Traditional optical networks have been designed and deployed in the backbone as a service-independent layer to meet the demands of highly multiplexed and predictable voice and private-line traffic [BEN01]. In this context, the primary network requirement is high reliability. As a result, optical resilience is designed statically using offline algorithms and all traffic is treated identically and fully protected. However, the static network design and provisioning mechanism becomes more and more clumsy and inefficient with increasing traffic levels. Two issues are:

Firstly, the static optical resilience provisioning happens under the assumption that the operator has a demand forecast for a future time period, and decides how to add capacity to the network in an optimal manner to support the demand. This prediction becomes more and more difficult to make because of the explosion of the Internet traffic. Thus, dynamic optical provisioning becomes more and more important for the service provider to respond quickly and economically to customer demands.

Secondly, the Internet growth has diminished the predominance of voice traffic and private-line traffic relative to the much greater growth of data traffic, which presents a wider range of resilience requirements. For example, traffic generated by residential Internet access services typically requires a much lower grade of service than that of corporate financial transactions. Thus, a more cost-effective mechanism is needed to provide different resilience that better reflect the value of the traffic being carried.

Optical dynamic provisioning has been studied in a number of previous papers. [ZAN01][SEN01][ASS01] introduces the infrastructure of an IP-centric controlled optical work supporting optical dynamic provisioning. Some basic routing algorithms (fixed routing, semi-adaptive routing and adaptive routing) are also introduced with comparative simulation results being presented. Other researchers [BAR96][BIR96] focus on the analytical modelling and experimental analysis of blocking probability under the dynamic provisioning of optical connections. However, all the performance analysis in previous work assumed a single level of optical resilience provisioning. This research investigates the differentiated resilience provisioning and their coexistence in the optical network.

The rest of this chapter is organised as follows. Section 6.2 describes the novel Differentiated-Resilience Optical Services Model (DROSM) and all the associated mechanisms. Section 6.3 gives a simple description of the implementation of the simulation models. Section 6.4 presents the extensive simulation results of performance studies. The summary is given in Section 6.5

6.2 Differentiated-Resilience Optical Service Model

6.2.1 Optical Restoration Options

There is a problem in using GMPLS to implement optical network restoration. An important difference between GMPLS LSPs in the optical layer and MPLS LSPs in the IP layer is that zero-bandwidth paths cannot be established for later use in the former case. In the IP layer case, MPLS LSPs may be established whereby if no packets are switched into the

links along the path, no bandwidth is consumed. Switching packets onto these predefined paths is simple and rapid. While in the optical layer case, merging of multiple circuits into a single outgoing circuit at the same bit rate is generally not possible. This important difference between IP and optical networks becomes crucial in allocating restoration capacity [DOV01].

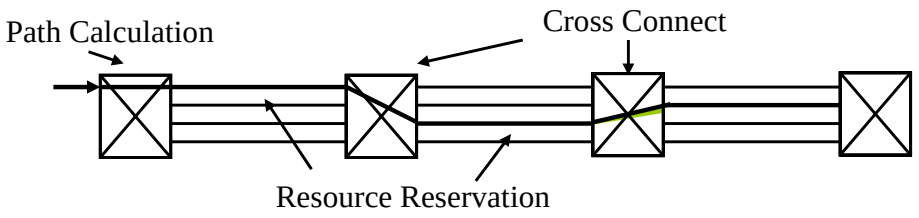


Figure 6-1: Provisioning an Optical Backup Path

Provisioning a GMPLS backup LSP connection consists of three actions (as shown in Figure 6-1): **path calculation**, **resource reservation** and **cross connect**. **Path calculation** involves the calculation of the link-disjoint / node-disjoint path. **Resource reservation** involves the reservation of wavelength channel for the backup LSP. **Cross connect** includes selection of the reserved resource and setting up the connection. The optical restoration options can be classified into four categories depending on whether these three actions for the backup path are performed before or after a failure event as illustrated in Table 6-1.

Category	Path Calculation	Resource Reservation	Cross Connect	Feasible
1	Before	Before	Before	Yes
2	Before	Before	After	Yes
3	Before	After	After	No
4	After	After	After	Yes

Table 6-1: Optical Restoration Options

For category 1, a separate connection is set up and set aside for restoration. In the event of a failure, restoration simply involves the cross-connect action at both ends of the connection, which makes this option the fastest and be able to achieve the restoration time benchmark (50ms) set by SONET. However, this option is also the most expensive since all the resource (time/space channel, wavelength) is allocated and the cross-connects are set before the failure, which cannot be shared by other traffic. This option is generally referred to as 1+1 or 1:1 dedicated protection.

For category 2, the backup paths are calculated and resources reserved before a failure event arises so it guarantees successful traffic restoration. As the cross-connects are set after the failure, the backup resource can be shared by other traffic prior to the failure event. However, signalling from the Path Switch LSR (PSL) is needed to notify the interim optical cross connects (OXCs) to select the reserved channel and build up the backup connection. Some constraints such as signalling propagation delay, cross-connect time and the large number of optical connections that need to be set up in response to a single failure make it hard to achieve the goal of 50ms restoration time. Experimental results [DOV99][LI01] show that this option is able to restore traffic within several hundreds ms.

For category 3, a backup path is pre-calculated and saved in the PSL. When there is a failure, the PSL sets up the backup path using either the CR-LDP or RSVP-TE signalling protocol. Since the resources are not reserved at the time of the original path calculation and they may be consumed later by other traffic, this option cannot ensure a successful restoration in a dynamically changing network, which makes this option impractical. This option only saves time by pre-calculating the path that would be requested if a failure were to arise. This could be trivial in an OXC with high processing performance.

For category 4, all the three actions are performed after failure. Although this option, referred to as fully dynamic restoration in [DOV01], is the most cost efficient, it requires a longer restoration time. This is caused by the inaccurate link state information for path calculation immediately after the failure and the large number of connections that typically need to be dealt with at the same time, in which case link contention often occurs. Nevertheless, the restoration time taken by this option is favourably comparable to that of the Distributed Network Restoration Scheme investigated in [GRO91][KOM90], since it uses link state information to calculate the alternative path instead of flooding messages to search one in the Distributed Network Restoration Scheme, which takes more time. In addition, the connection blocking caused by link contention in this option can be reduced using the scheme of crankback routing extensions to CR-LDP / RSVP-TE [IWA01]. Research in [GRO91] shows connections can be restored in under 2 seconds, thus, 2 second restoration times should be achievable for restoration option category 4.

6.2.2 Service Classification

According to the different optical restoration options, this research proposes a set of optical service resilience classes according to their resilience requirements as itemized in Table 6-2.

Service Class	RC1	RC2	RC3	RC4
Resilience Requirement	High	Medium	Low	Best Effort
Restoration Time	< 50 ms	< 500 ms	< 2 s	< 60 s
Resilience Strategy	Category 1	Category 2	Category 4	Category 4

Table 6-2: Service Classification and Resilience Strategies

Optical services of Resilience Class 1 (RC1) have the highest resilience requirements and require traffic restoration within 50 ms. Next comes optical services of Resilience Class 2 (RC2); these have medium resilience requirements of being restored within 500 ms. Following this, optical services of Resilience Class 3 (RC3) have relatively low resilience requirements with restoration times less than 2s if could be restored. Finally, optical services of Resilience Class 4 (RC4) are best-effort traffic and can be pre-empted by services of all other resilience classes. However, if there is spare resource, they could be restored after failure in around 60s.

6.2.3 Integration of Differentiated-Resilience with Optical Services

6.2.3.1 Link Status Classification

The link state routing protocol OSPF has been extended within the IETF to flood optical network information within an optical domain [KOM02][KOM01]. In OSPF, these optical router and link attributes are flooded as opaque Link State Advertisements (LSAs). These attributes are put in a hierarchical Type/Length/Value (TLV) triplet. Among the link attributes, there is a sub-TLV (Link Protection Type) dealing with network resiliency. The Link Protection Type represents the protection capability that exists for a link. Six protection capabilities are defined [KOM01][KOM2] so far: Extra Traffic, Unprotected, Shared, Dedicated 1:1, Dedicated 1+1 and Enhanced.

The definition of Link Protection Type is intended to provide information that can be used by path calculation algorithms to set up LSPs with the appropriate protection characteristics. However, this mechanism assumes that the resilience algorithm has already determined the protection type(s) of all the underlying links. The path calculation algorithm has no idea of how the resilience schemes have utilised the underlying resources and has no direct control over them. All that is available to the path calculation algorithm is information concerning the pre-calculated resilience state of each link. This lack of awareness between the

path calculation algorithm and the resilience enactment activity results in inefficient resource allocation.

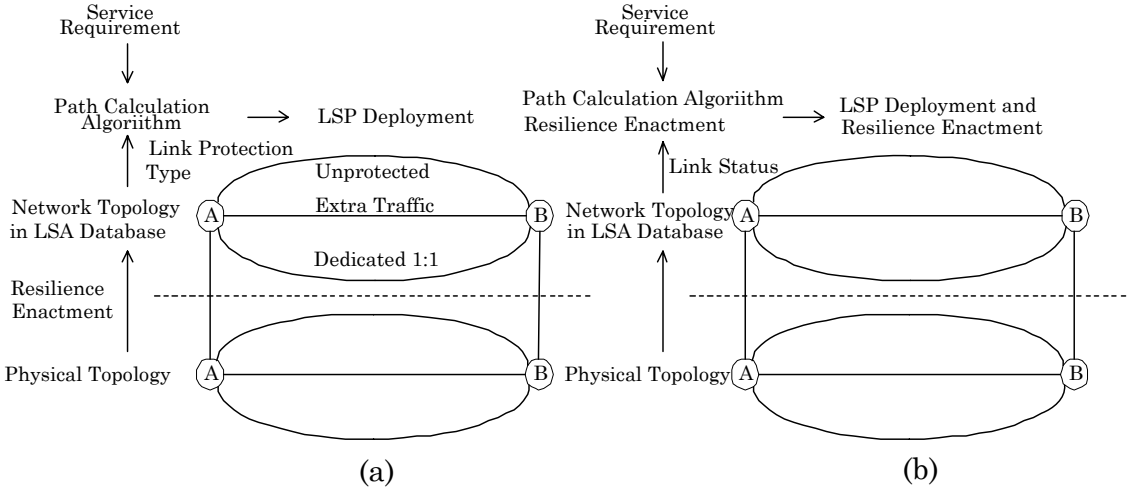


Figure 6-2: Inefficiency due to Lack of Link Status Awareness

For example in Figure 6-2 (a), assume that the resilience assignment of each link has been pre-established by resilience-provisioning algorithms such that there is one Unprotected, one Extra Traffic and one Dedicated 1:1 link. The path calculation mechanism can only select between these options even though it may have preferred an alternative assignment (i.e. three Unprotected links).

Instead, this research suggests a new link resource utilisation mechanism, shown in Figure 6-2 (b), which performs the resilience assignment and path calculation together, according to current service requirements. This assignment is dynamic, with the service requirements dictating how the links are configured, and so, better use is made of the available underlying resources.

In order to realise this mechanism, this research proposes a new optical link attribute, called **Link Status**. The Link Status attribute includes five types: *Unused*, *Used*, *Reserved*, *Shared* and *Held*.

Unused links are not used by any traffic nor reserved by backup LSPs. *Used* links are those occupied by non-preemptable working LSPs or backup LSPs for RC1 services. They cannot be shared with any other traffic. When the optical service is finished, *Used* links are released and the status transferred to *Unused*.

Reserved links are those used by backup LSPs of RC2 service. Additional parameters show which link or node failures they are protecting against. *Reserved* links for one backup

LSP can be shared with other backup paths that have node/link disjointed working paths, or used by preemptable traffic. In the former case, the link status remains as *Reserved* and the additional parameters are refreshed with new possible link and node failure events being added to the link state database. While in the latter case, the status becomes *Shared*.

Shared links are *Reserved* links used by preemptable (RC4) traffic. They cannot be shared by other preemptable (RC4) traffic, but the additional parameters are maintained. Once the preemptable service is terminated, the status becomes *Reserved*.

Finally, *Held* links are those being used by preemptable traffic (RC4). Figure 6-3 shows the Link Status Finite State Machine (FSM).

In a distributed routing mechanism [SEN01a][LI02], the detailed Link Status information is maintained in the local database of each OXC. Only the aggregated Link Status information is broadcast and available for path calculation in the ingress OXC.

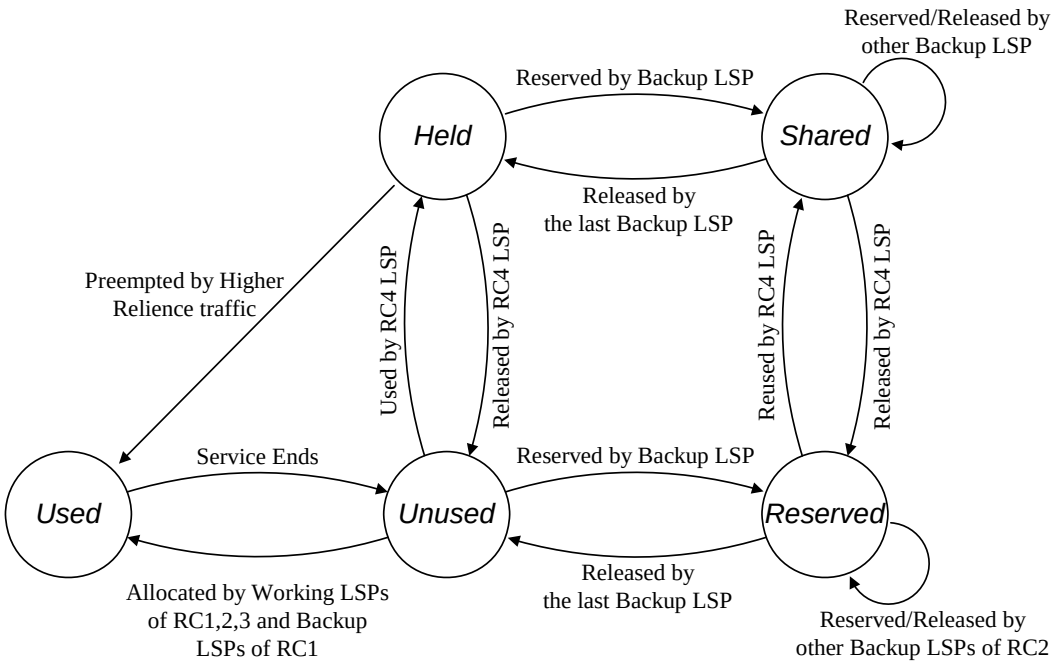


Figure 6-3: Link Status Finite State Machine

6.2.3.2 Resilience Strategies

In the proposed model, traffic is generally protected or restored using end-to-end (path) restoration within an individual optical domain. Path restoration has the advantage of being more cost-efficient than link restoration. However, in a large optical domain, the time taken by path restoration for a long end-to-end optical path may be unable to satisfy service with

high resilience requirement (RC1) and restore traffic within 50 ms. In this case the Adaptive Segment Path Restoration (ASPR) scheme proposed in Chapter 5 of this thesis can be used to divide the long optical path into several segments, with a backup LSP deployed for each segment.

This strategy is not needed for services with a low resilience requirement (Resilience Class 2, 3 and 4) since the extra time taken by end-to-end restoration is a relatively small part of the total restoration time.

For RC1 traffic, optical restoration category 1 is used. For RC1 service, the working LSP can use links whose status is *Unused* or *Held* when resources are limited and pre-emption is permitted at deploying time. Similarly, the backup LSP of RC1 service can use *Unused* links or *Held* links (if pre-emption is permitted).

For RC2 traffic, optical restoration category 2 is used. The pre-calculated and allocated backup LSPs ensure the traffic can be restored successfully. Since the cross-connects are set only after the failure occurs, the resource taken by the backup LSP can be reused. At the time of LSP path calculation and deployment, the working LSP uses *Unused* links, or *Held* links if resource is limited and pre-emption is permitted. Then, the status of these links becomes *Used*. The backup LSP can use *Unused* links, or *Reserved* links, or *Held* links, or *Shared* links. Accordingly, the link status of these links becomes *Reserved*, *Reserved, Shared* and *Shared*, respectively.

For RC3 traffic, optical restoration category 4 is used (Note that category 3 is impractical). The node-disjoint / link-disjoint restoration LSPs are calculated and deployed only after a failure. RC3 service cannot be pre-empted. So, at the time of the alternative LSP provisioning in response to a failure, only those links with a status of *Unused* or *Held* (if resource is limited and pre-emption is permitted) can be used.

For RC4 traffic, optical restoration category 4 is used. The restoration of RC4 services starts at the time when OSPF has re-converged (within 60s). At this time, each node has accurate (or more up-to-date) information of the revised network state after the failure. Also, as the links previously used by the failed services of type RC1, RC2 and RC3 have been reclaimed, more resource for restoration will be available. Unlike RC3, the restoration LSP of RC4 service can use *Unused* and *Reserved* links. However, the restoration of RC4 could fail due to there being no appropriate spare resource.

Both RC3 and RC4 are unprotected services and could experience unsuccessful restoration if the resource is limited. However, there is still a necessity to distinguish between

these two classes. RC3 cannot be pre-empted and can pre-empt RC4. Their restoration times are also different.

6.2.4 Functional Model

6.2.4.1 Resilience Provisioning

The resilience-provisioning algorithm is performed at the ingress OXC, which serves as the Path Switching LSR (PSL). Different resilience strategies are provided for the different optical services. For RC1 and RC2 services, a pair of link-disjoint / node-disjoint working and backup paths is calculated according to the aggregated link information.

For RC2, sharing of any link on the backup path is determined during signalling of the proposed backup path. Information about the working path is also carried in the signalling message to establish the backup path. When it receives the signalling message, each OXC on the backup path decides whether the proposed backup path can reuse the resource already reserved by other backup paths. The decision is based on the detailed Link Status information, which is maintained in each OXC's local database. For RC3 and RC4 services, only the working path is calculated and deployed.

6.2.4.2 Restoration Procedure

Almost all the processes are originated from the PSL. The Path Merge LSR (PML) performs active actions only in RC1 restoration.

For RC1, when the PML detects a loss of light in the working path, the signal in the protecting path will be used. For other resilience classes, a notification message is needed to inform the PSL. When the PSL is notified of the failure, it starts the main restoration process. The main restoration process retrieves information about the failed services and produces two child processes immediately (i.e. the RC2 and RC3 restoration processes). It also schedules the RC4 restoration process for subsequent action once OSPF re-convergence has taken place, as shown in Figure 6-4.

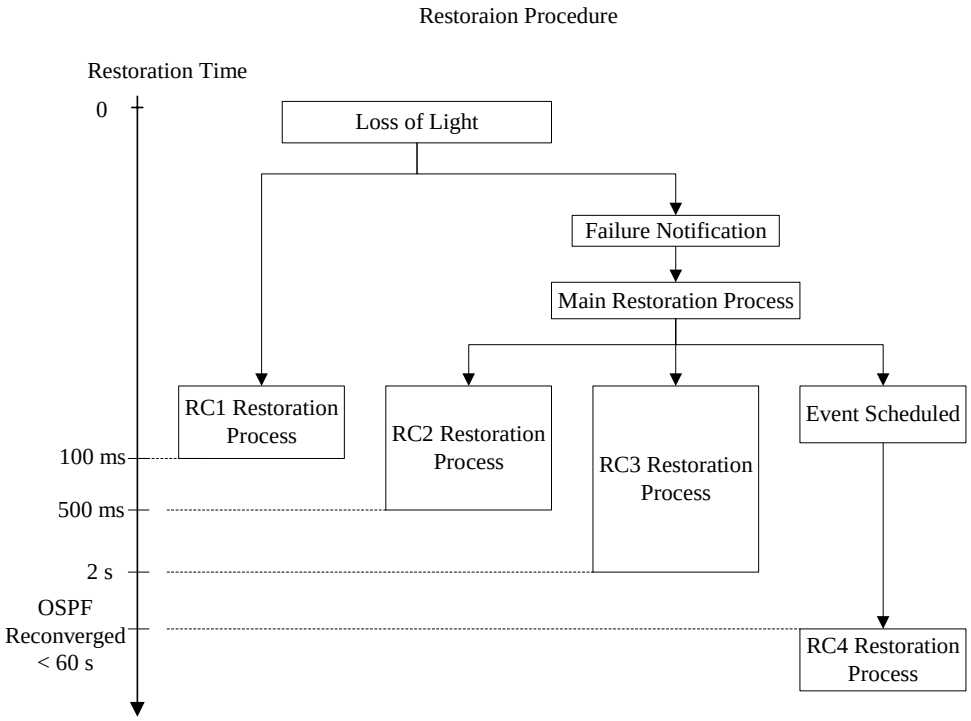


Figure 6-4: Restoration Procedure Time-Line

6.2.4.3 Failure Recovery Procedure

A robust and practical resilience scheme requires an easy means of transferring the restored traffic back once the network failure has been repaired. This is essential if the network is to be automatically maintained. This procedure is referred to as **Failure recovery** and the new working LSP as the **recovery LSP**.

As for the restoration procedure, the **failure recovery** procedure is mainly performed by the PSL. When the PSL learns of the revival of the previously failed component(s), it recovers the RC1, RC2 and RC3 services one by one. For a RC1 or RC2 service, the **recovery LSP** tries to use the same path as the original failed LSP. If this is not possible (because the link is already being used by other traffic), the PSL must calculate a new pair of working and backup LSPs.

For a RC3 service, the PSL calculates a new LSP directly without referring to the failed one. All traffic is switched to the recovery LSP only after the pipes have been established, allowing the service outage to be less than 100ms and therefore satisfies the service resilience requirement.

The recovery of RC4 services may not be necessary; however, this can take place once OSPF has again reconverged.

6.3 Simulation Models

Simulation is used to evaluate the proposed differentiated-resilience-provisioning scheme with other single level resilience schemes. This section gives a description of the implemented models, which are used in the simulation.

In the simulation network, each Optical Cross-Connect (OXC) is modelled as an OPNET node model. Each node could initiate a call and become the ingress node of a lightpath to any other nodes. Figure 6-5 shows its function modules. It contains four function blocks including: Service Generator, Connection Manager, GMPLS/CR-LDP and OSPF link state database (LSDB).

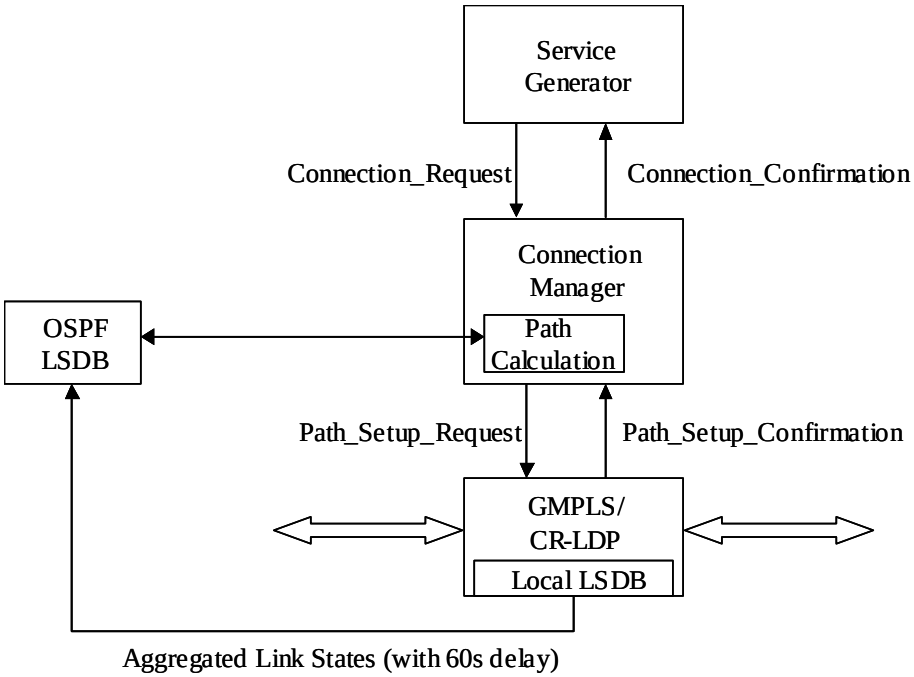


Figure 6-5: Simulation Model

6.3.1 Service Generator

The Service Generator automatically produces requests to set up connections with specified resilience requirement. The resilience requirement could be RC1 (Dedicated), or RC2 (Mesh-shared), or RC3 (Unprotected) or RC4 (Best-effort). In order to evaluate the network performance at different traffic loads, two parameters of the lightpath connection request are defined. These two parameters are connection rate (defined as the frequency of the connection request or the reciprocal of connection arrival interval) and holding time (defined as the life time of each connection). The traffic load is measured in Erlangs, which can be calculated by multiplying the connection rate with the average connection holding time. In the

simulations, the connection arrival interval is assumed with a Poisson distribution while holding time follows a negative exponential distribution.

Service Generator also receives confirmation of connection requests and records the connection statistics.

6.3.2 Connection Manager

This module takes charge of the path calculation and monitoring the path setup of a connection request. When there is a lightpath connection request, it retrieves information about the current network state from the OSPF link state database (LSDB) and calculates the route of the primary and backup path according to the request's resilience requirement. The route calculation and connection deployment of the backup path initiate only after those of the primary path have been finished successfully. For a lightpath with a higher resilience requirement, it is established successfully only if both its primary path and backup path have been set up successfully.

6.3.3 GMPLS/CR-LDP

This module acts as the signalling process which realizes most of the major functions of CR-LDP with GMPLS extensions, defined in [ASH02][RFC3036]. In addition, it also maintains a local database that records the detailed sharing information between backup LSPs.

6.3.4 OSPF LSDB

In order to provide a routing topology for each OXC to calculate the route of a lightpath, a module that provides simplified OSPF functions is implemented. This module maintains a database, which records the link state of the simulation network.

6.3.5 Verification and Validation

Having developed a simulation model, the node needs to be verified and validated. Verification determines whether the model does indeed perform as intended and validation shows whether the model is a true and accurate representation of the system modelled [PIT93]. This needs to be carried out at two levels, the first on a fine scale by looking at individual objects that make up the network and then at the whole network.

The verification of validation of the simulation models for ASPR is carried out from the following aspects.

First, the functionality performance of the simulation models are thoroughly tested using debugging tools provided by OPNET™ during model implementation.

Second, an example network with a simple topology and certain traffic demands are used to validate the outcome against the expected results.

Third, the performance results of single level resilience-provisioning schemes (dedicated and mesh-shared protection) are also validated by comparing with some published results [SEN01][LI02][RAM01] of other researchers using similar prerequisites.

6.3.6 Confidence Interval

System models that include stochastic behaviour have results that are dependent on the initial seeding of the random number generator. As a particular random seed selection can potentially result in an anomalous or non-representative behaviour, it is important for each model configuration to be exercised with several random number seeds, in order to be able to determine standard or typical behaviour. The basic principal applied here is that if a typical behaviour exists, and if many independent trials are performed, it is likely that a significant majority of these trials will fall within a close range of the standard.

Therefore, in the following section, most simulations are performed ten times using different initial seeds. The results are the average of that of the ten simulations. The confidence interval will not be shown in the figure if it is minute compared with the average sample value.

6.4 Performance Results

Extensive simulations have been performed to evaluate the proposed Differentiated-resilience Optical Services Model for the wavelength-routed optical network. For comparative studies, the results are compared with two schemes (dedicated protection and mesh-shared protection) which only provide single level of resilience.

6.4.1 Simulation and Network Parameters

Figure 6-6 shows the network topologies that are used in the simulation. Table 6-3 shows the parameters of the networks, in which ND (average nodal degree) represents the network connectivity.

Network	1	2	3	4
Node	11	14	28	25
Link	22	20	45	55
ND	4	2.86	3.21	4.4

Table 6-3: Network Parameters

In these networks, each node (OXC) is assumed with no wavelength conversion capability. Each link has 40 wavelength channels and each channel has a default cost of 1 unit. The number along each link is the link weight that is only used by routing protocol OSPF to calculate the route for a lightpath.

In order to evaluate the performance of dynamic provisioning of lightpaths in an optical network, traffic can be usually modelled as one of the two types: incremental traffic and dynamic traffic [ASS01].

For incremental traffic, the connection requests arrive one by one. A lightpath is established for each connection, and once being established, the lightpath remains in the network indefinitely.

For dynamic traffic, a lightpath is set up for each connection request as it arrives, and the lightpath is released after some finite amount of time, which is called the holding time.

Both types of traffic are used to evaluate the performance of the proposed scheme.

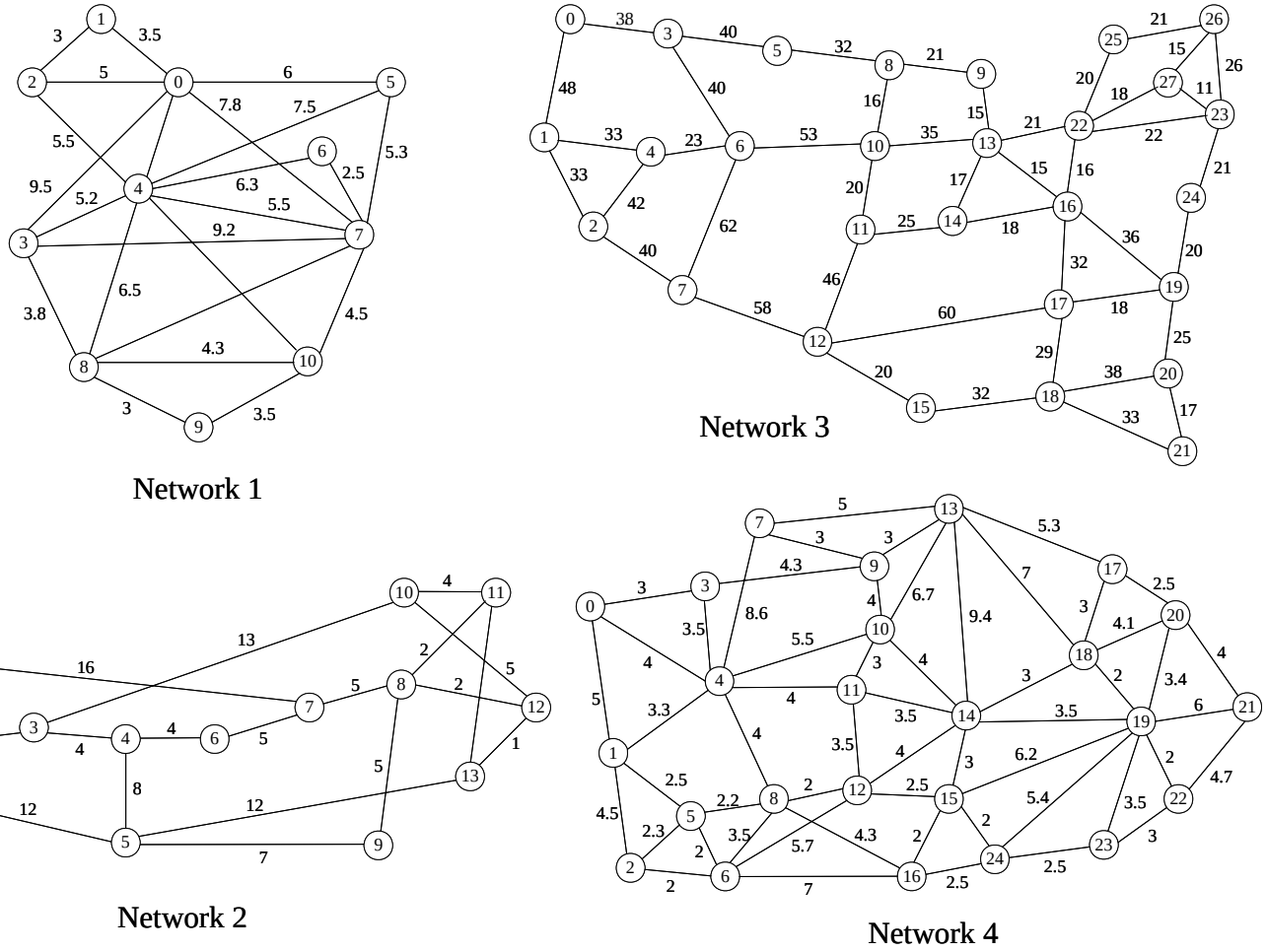


Figure 6-6: Examples of Network Topologies

6.4.2 Dynamic Traffic Scenario

Firstly, dynamic traffic is used to examine the performance of these resilience-provisioning schemes. In this set of experiments, the connection request arrives with a Poisson distribution and each lightpath has an average holding time of 100s with an exponential distribution. The traffic load is calculated by multiplexing the connection arrival rate with the average holding time. In order to increase the connection probability, three connection attempts are made for each lightpath request. It assumes the traffic pre-emption is not permitted at the path establishment. In the differentiated-resilience scenario, the ratio of connection requests for the four resilience classes is 1:1:1:1.

Each experiment is performed 10 times with each time a different initial seed. Those presented in the figures are the average results of the ten simulations.

6.4.2.1 Blocking Probability

The ratio of the number of blocked connections to that of the total connection requests is termed as blocking probability. This parameter is often used to evaluate the performance of a network with dynamic traffic.

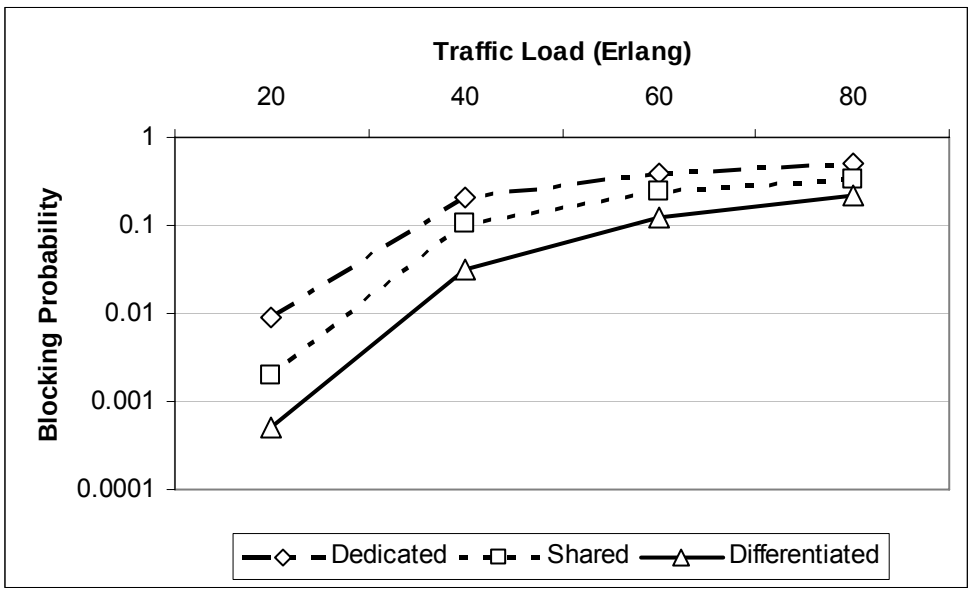


Figure 6-7: Blocking Probability of the LATA Network

Figure 6-7 shows the blocking probability as a function of the traffic load per node for the LATA network.

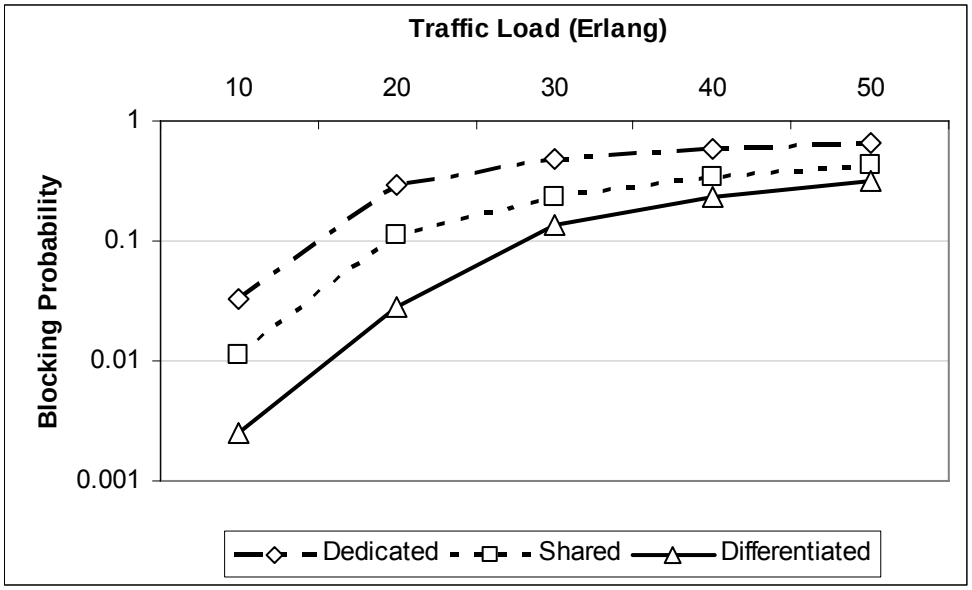


Figure 6-8: Blocking Probability of the NSF Network

Figure 6-8 shows the blocking probability as a function of the traffic load per node for the NSF network.

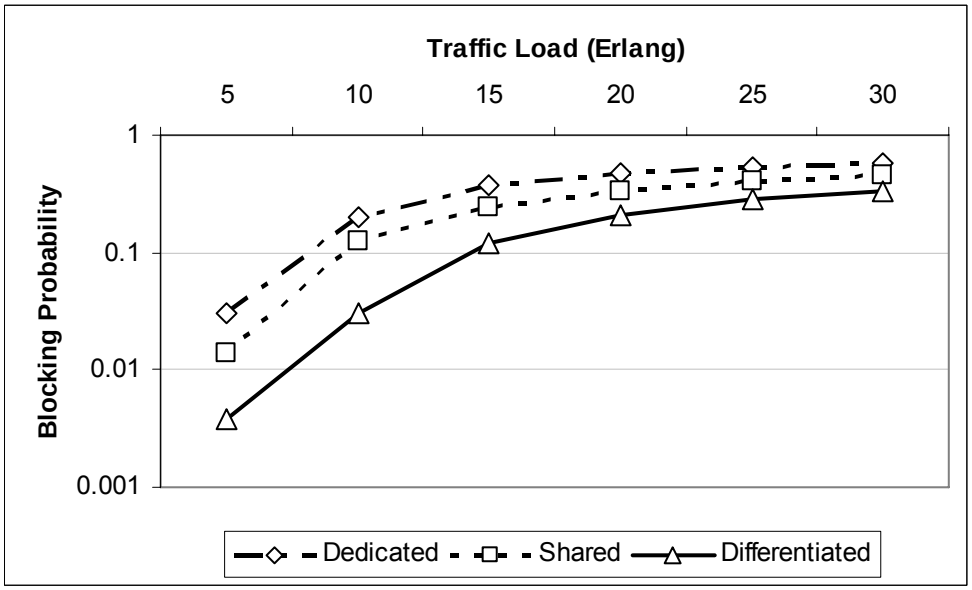


Figure 6-9: Blocking Probability of the USA Long Haul Network

Figure 6-9 shows the blocking probability as a function of the traffic load per node for the USA long haul network.

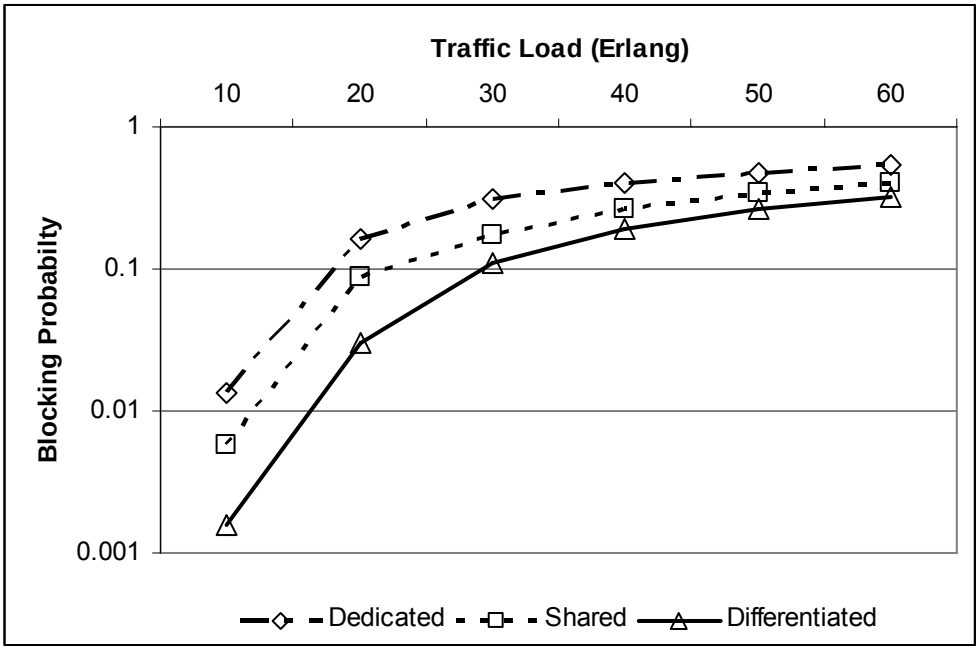


Figure 6-10: Blocking Probability of the Toronto Metropolitan Network

Figure 6-10 shows the blocking probability as a function of the traffic load per node for the Toronto metropolitan network.

As expected, in all networks differentiated-resilience scenario has a better performance than that of dedicated and shared protection.

6.4.2.2 Total Deployed Connections

In this set of experiments, the average number of successfully established connections for each scheme are collected and shown as a function of the traffic load.

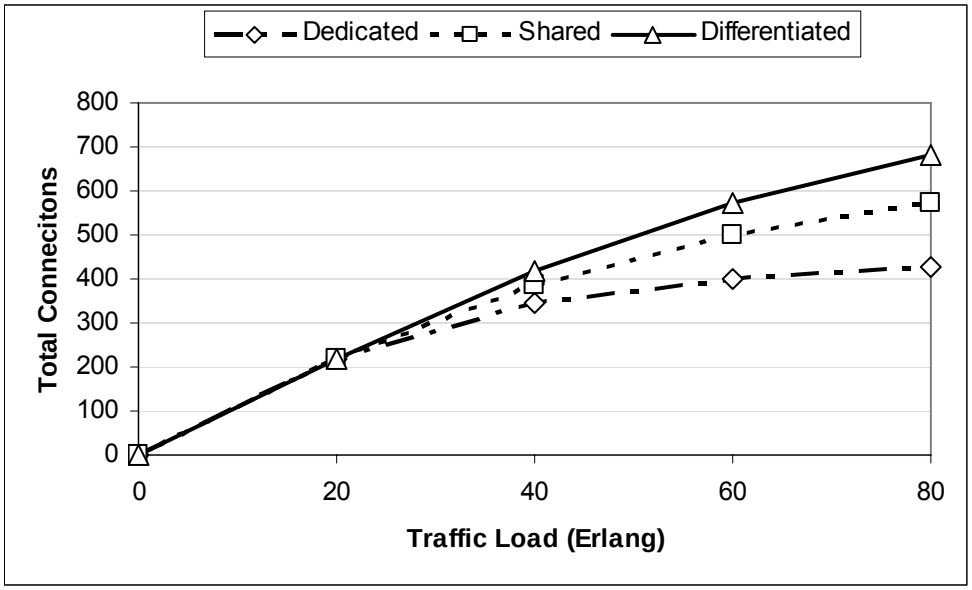


Figure 6-11: Total Deployed Connections in the LATA Network

Figure 6-11 shows the average number of deployed connection as a function of the traffic load per node in the LATA network.

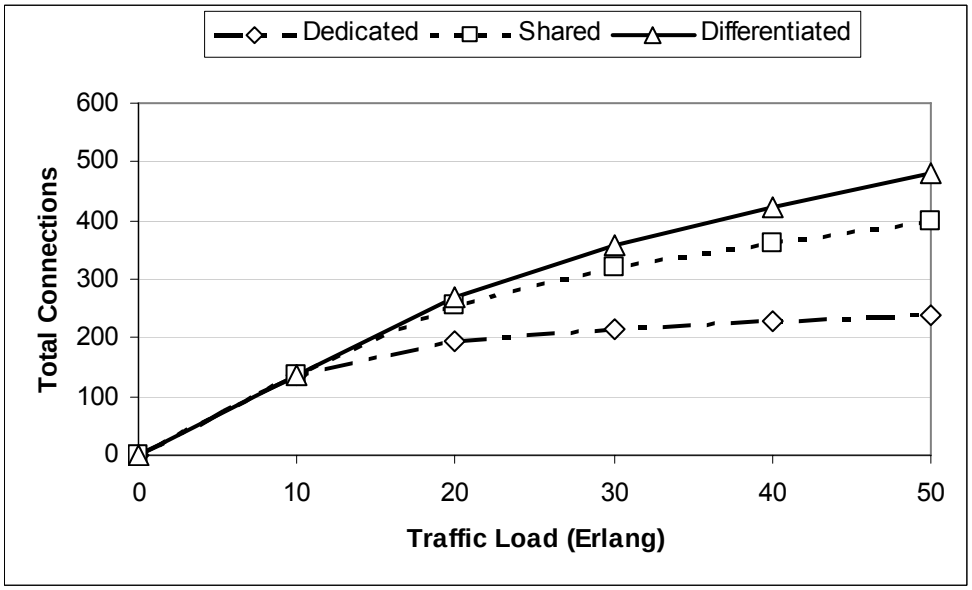


Figure 6-12: Total Deployed Connections in the NSF Network

Figure 6-12 shows the average number of deployed connections as a function of the traffic load per node in the NSF network.

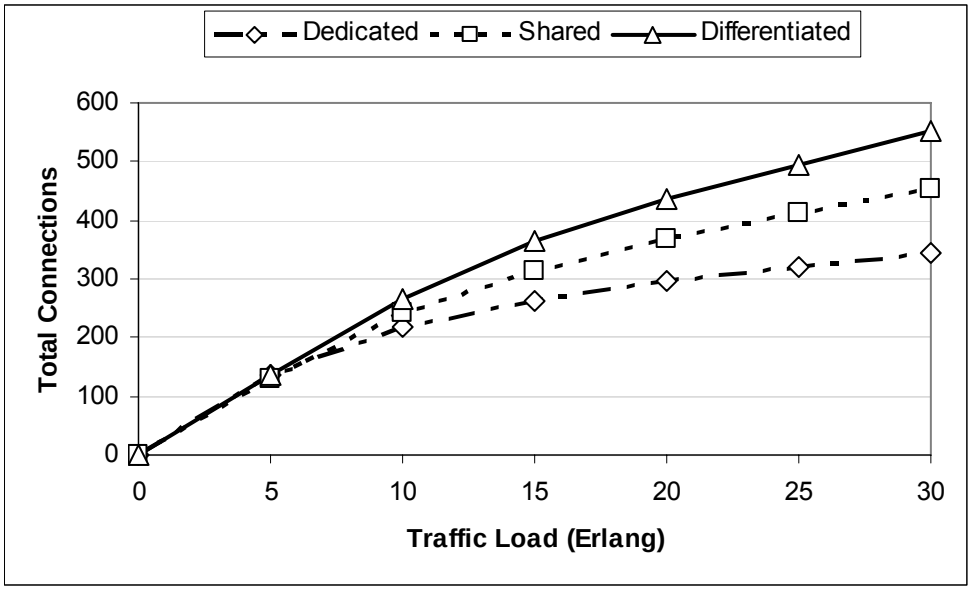


Figure 6-13: Total Deployed Connections in the US Long Haul Network

Figure 6-13 shows the average number of deployed connections as a function of the traffic load per node in the US long haul network.

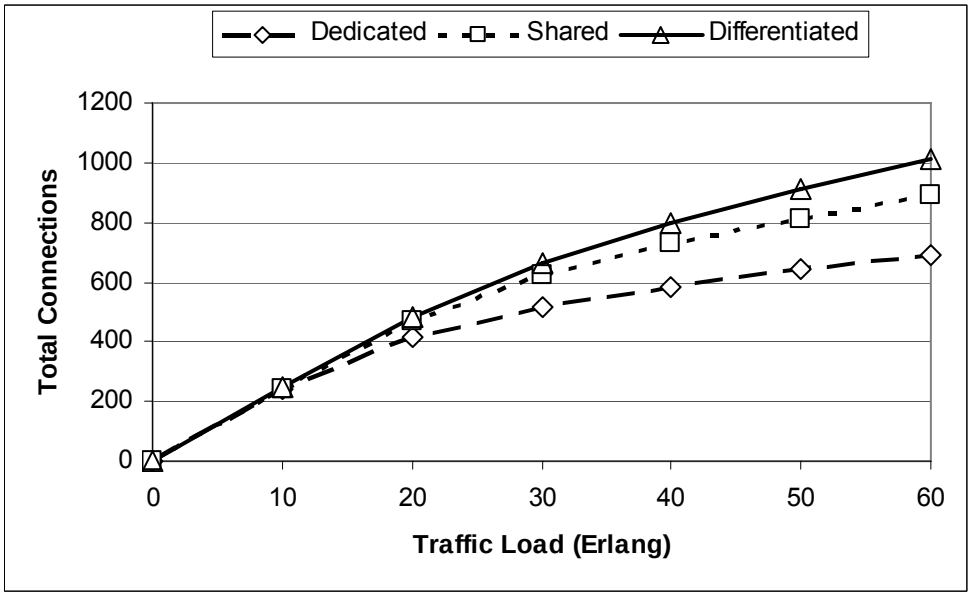


Figure 6-14: Total Deployed Connections in the Toronto Metropolitan Network

Figure 6-14 shows the average number of deployed connections as a function of the traffic load per node in the Toronto metropolitan network.

Figures listed above show that, for dedicated protection, the average number of connections being deployed is the smallest. The average number of connections with the differentiated-resilience approach is greater that of dedicated and shared protection. The result

confirms that differentiated resilience provisioning has the advantage that more connections can be deployed compared with traditional single level fully protected resilience. For example in Figure 6-14, when each node has a traffic load of 60 Erlangs, the network accommodates an average of 1013 connections. This compares favourably against 689 connections in dedicated protection scenario. The differentiated resilience-provisioning scheme provides nearly 50% more connections. Shared protection provides 892 connections; the differentiated resilience-provisioning scheme is still able to provide about 14% more connections.

6.4.2.3 Resource Allocation

In this set of experiments, the average amount of resource that is used at different traffic loads is observed.

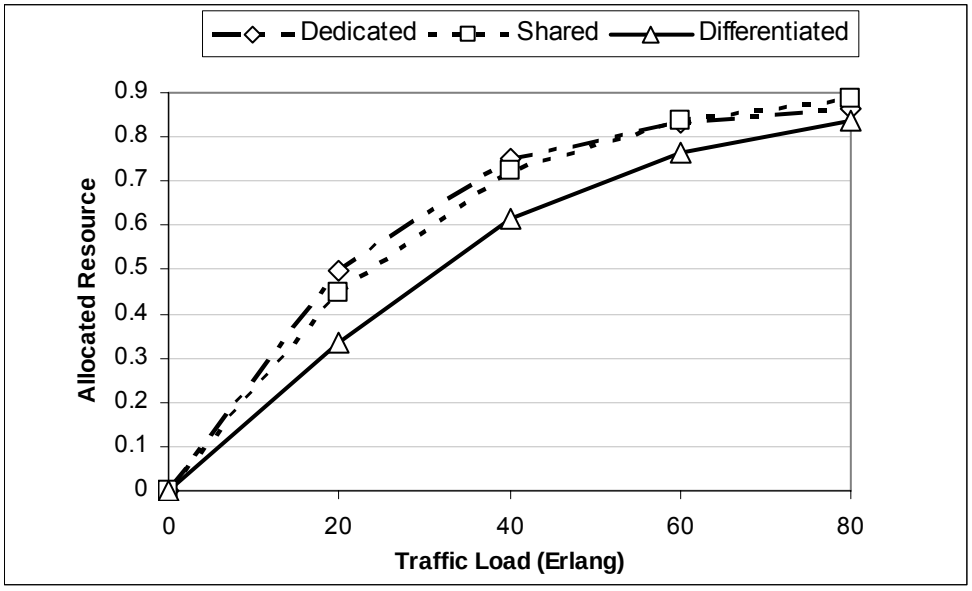


Figure 6-15: Resource Allocation in the LATA Network

Figure 6-15 shows the average allocated resource as a function of the traffic load per node in the LATA network.

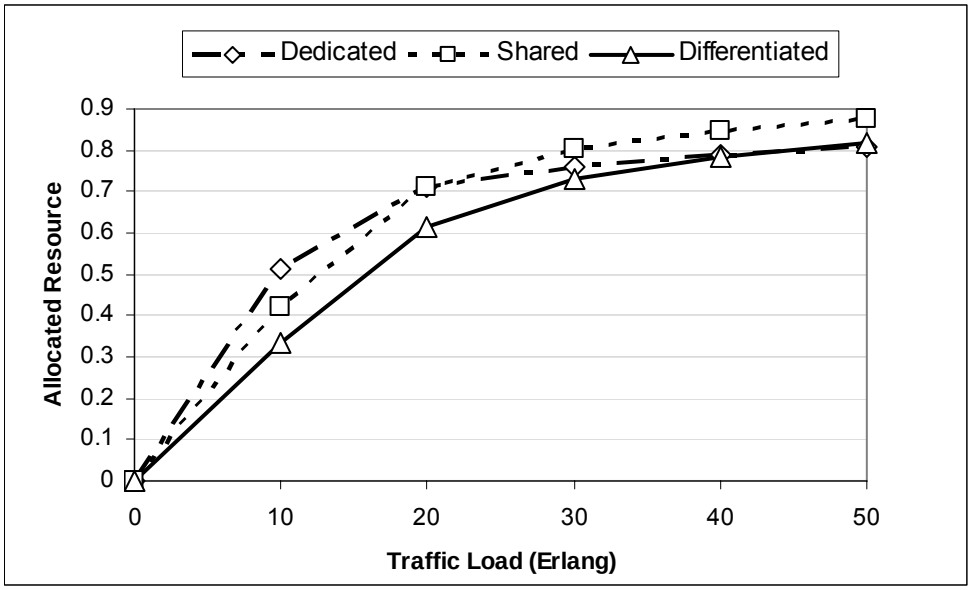


Figure 6-16: Resource Allocation in the NSF Network

Figure 6-16 shows the average allocated resource as a function of the traffic load per node in the NSF network.

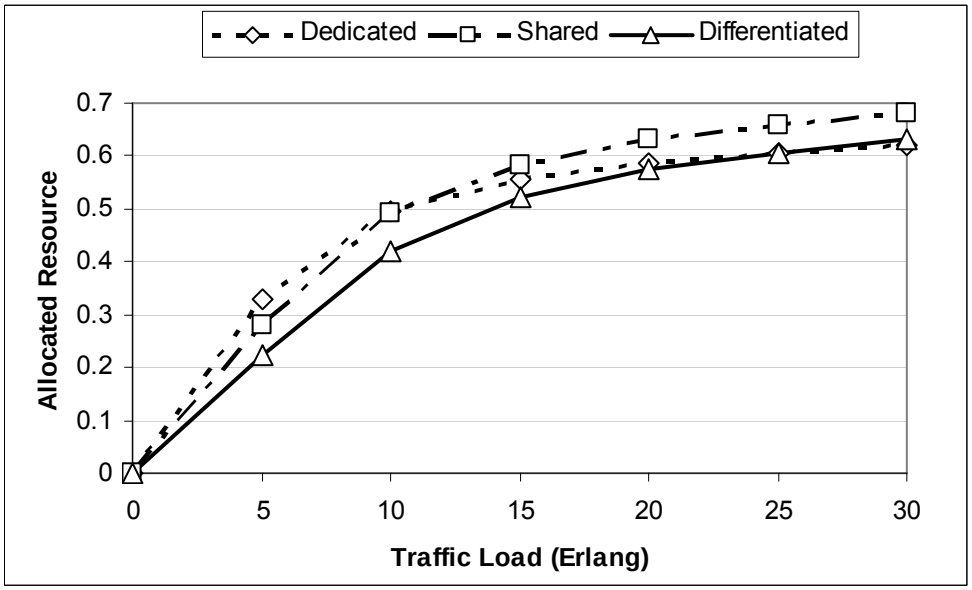


Figure 6-17: Resource Allocation in the US Long Haul Network

Figure 6-17 shows the average allocated resource as a function of the traffic load per node in the US long haul network.

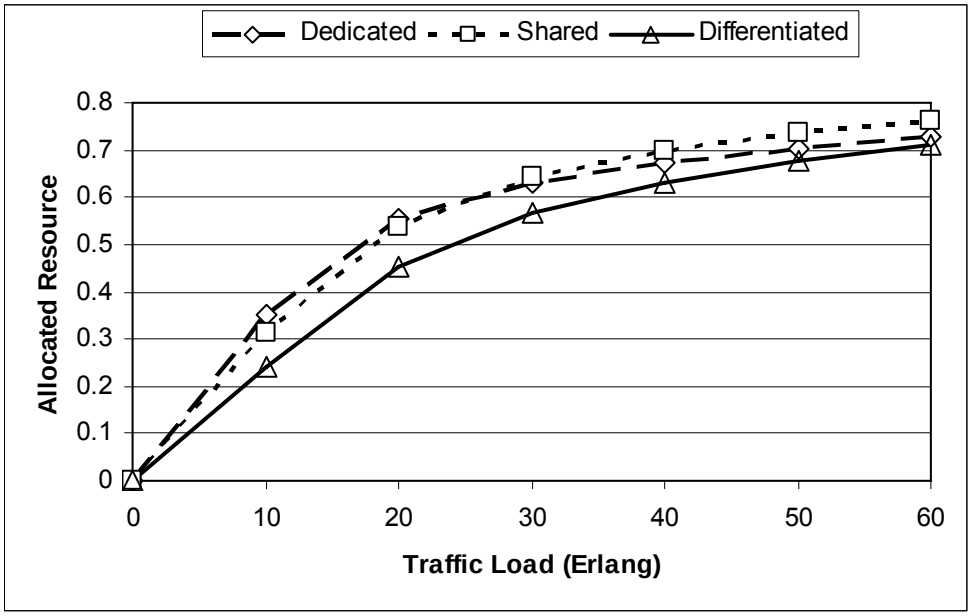


Figure 6-18: Resource Allocation in the Toronto Metropolitan Network

Figure 6-18 show the average allocated resource as a function of the traffic load per node in the Toronto metropolitan network.

These figures show that differentiated resilience provisioning usually consumes less network resource than the single level resilience provisioning schemes including dedicated protection and shared protection. In contrast to expectations, in some cases shared protection allocates more resource than dedicated protection. This is because, given the same connection request rate, shared protection has a higher connection acceptance and thus allows more traffic to be deployed in the network which then consume more resource. As the traffic load increases, much more connections will be deployed for differentiated-resilience provisioning. It is therefore expected that the average allocated resource for differentiated-resilience provisioning will eventually take over that of shared protection and dedicated protection. This will be validated in the following experiments in the incremental traffic scenario.

Together with the perform in blocking probability, the conclusion can be drawn that differentiated resilience provisioning requires less network resources and has lower blocking probability than that of either dedicated or shared protection.

6.4.3 Incremental Traffic Scenario

This section uses incremental traffic to assess the performance of the differentiated resilience optical services model and two single level resilience-provisioning mechanisms: dedicated protection and shared protection.

In the incremental traffic case, connection requests arrive sequentially, a lightpath is established for each connection, and the lightpath remains in the network indefinitely. For each node, the connection requests have a Poisson distribution.

As results of the four networks are similar, only those of Network 4, the Toronto Metropolitan Network, are presented here.

6.4.3.1 Capacity Performance

Capacity performance is used to evaluate how many connections can be deployed in a given network for a particular resilience provisioning mechanism. A better capacity performance means more connections can be accommodated in a network. This attribute is of great interest to the network carriers as more connections mean more revenue could be possibly generated.

To find out the maximum number of connections a network can hold, incremental traffic is used to deliberately exhaust the network resource until no more connections can be established in the network.

Figure 6-19 to Figure 6-24 show the deployment results of the two single level resilience-provisioning mechanisms (dedicated protection and shared protection) and the proposed differentiated-resilience optical services model.

For each of the two single level resilience-provisioning mechanisms, ten simulations are carried out.

In the dedicated protection scenario, only 727 ± 8 (95% confidence interval) connections are deployed in the network.

In the shared protection scenario, there are average 1237 ± 11 (95% confidence interval) connections being established in the network.

In the differentiated resilience-provisioning scenario, the deployment results could vary a lot according to different connection request ratios of the four Resilience Class types of service. The ratio of different types of traffic really depends on what network carriers want their networks to be. For comparative study, extensive experiments of different connection request ratios have been performed to evaluate the proposed schemes. However, here, without losing its inherent characteristics, only several typical ratios are presented to describe the performance and what they mean to the network carriers.

In one example shown in Figure 6-19, the approximate request ratio of these four types of traffic is 2:2:1:1. In order to provide a similar amount of higher resilience traffic to that of the dedicated protection scenario, the maximum call attempts for RC1 services is set to 10 instead of the default value 3.

The result shows that the network accommodates 1565 connections in total, of which 375 are RC1 connections, 398 are RC2 connections, 307 are RC3 and 485 are RC4 connections. Such a deployment result shows that differentiated-resilience provisioning offers the ability to carry the same amount of "premium grade" (RC1 & RC2) guaranteed fast protection traffic as with purely dedicated protection, whilst offering approximately the same capacity over again in lower resilience quality connections, which have been gaining in popularity recently.

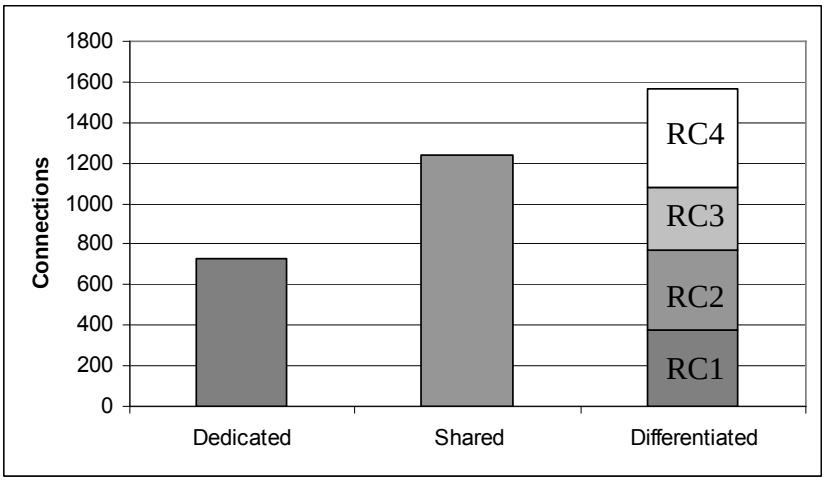


Figure 6-19: Deployment Result A (Request Ratio 2:2:1:1)

The result of other cases with different ratio of connection requests of different traffic are shown as the following:

In Figure 6-20, the approximate request ratio of these four types of traffic is 1:1:1:1, the reattempts of all connection requests are set to a default value of 3. For the differentiated resilience provisioning scenario, the network accommodates totally 1778 connections, namely 251 RC1 connections, 439 RC2 connections, 504 RC3 and 584 RC4 connections.

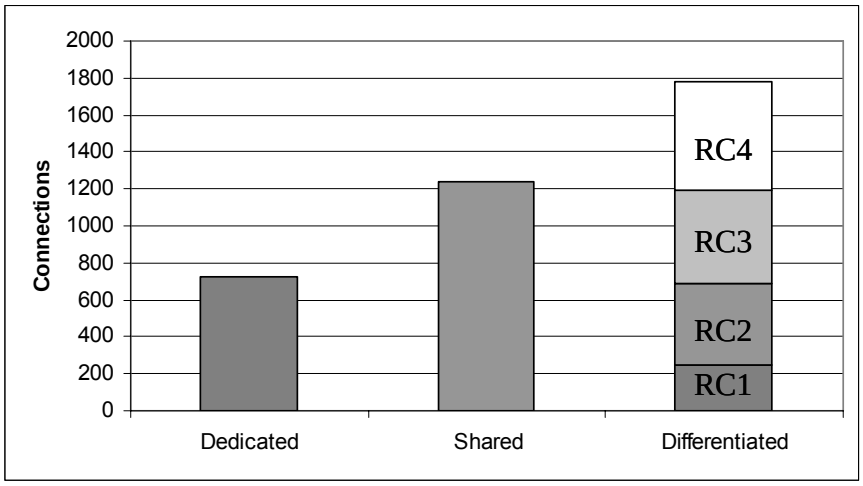


Figure 6-20: Deployment Result B (Request Ratio 1:1:1:1)

In Figure 6-21, the approximate request ratio of these four types of traffic is 1:2:3:4. The deployment result for the differentiated-resilience provisioning is that the network accommodates a total of 1869 connections, of which 131 are RC1 connections, 378 are RC2 connections, 563 are RC3 and 797 are RC4 connections.

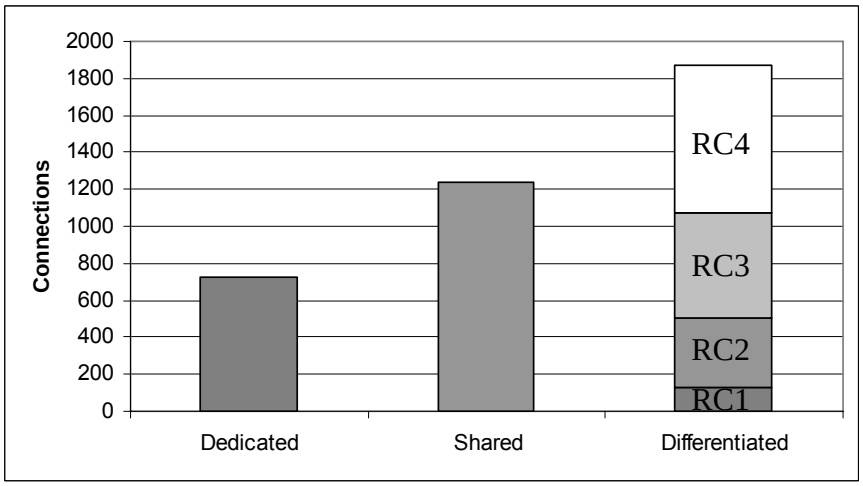


Figure 6-21: Deployment Result C (Request Ratio 1:2:3:4)

In Figure 6-22, the approximate request ratio of these four types of traffic is 1:1:3:5. In total 1900 connections have been established in the network in the differentiated-resilience provisioning scenario, of which 114 are RC1 connections, 205 are RC2 connections, 605 are RC3 connections and 976 are RC4 connections.

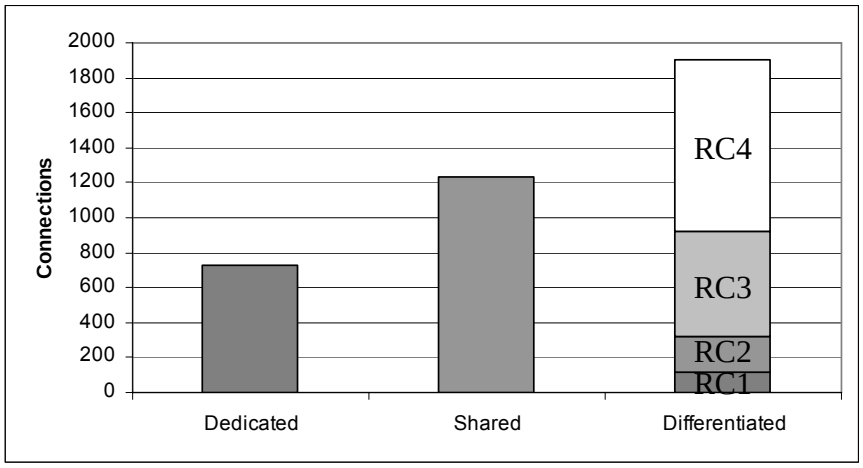


Figure 6-22: Deployment Result D (Request Ratio 1:1:3:5)

In Figure 6-23, the approximate request ratio of these four types of traffic is 1:1:2:4. The network accommodates a total of 1845 connections in the differentiated-resilience provisioning scenario, of which 151 are RC1 connections, 264 are RC2 connections, 474 are RC3 and 956 are RC4 connections.

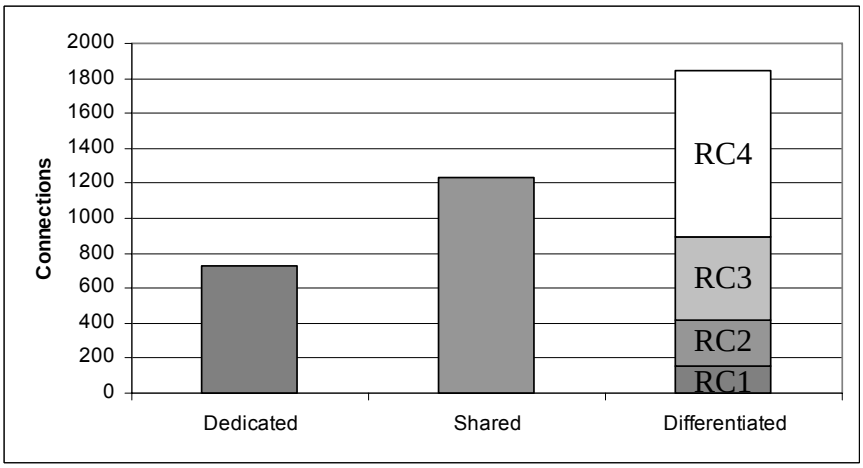


Figure 6-23: Deployment Result E (Request Ratio 1:1:2:4)

In Figure 6-24, the approximate request ratio of these four types of traffic is 1:1:3:6. The network accommodates a total of 1882 connections in the differentiated-resilience provisioning scenario, of which 116 are RC1 connections, 185 are RC2 connections, 561 are RC3 and 1020 are RC4 connections.

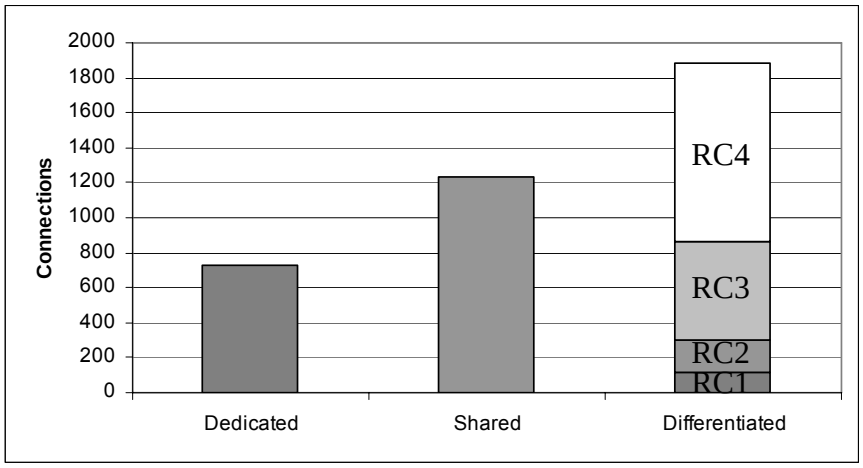


Figure 6-24: Deployment Result F (Request Ratio 1:1:3:6)

The figures listed above present varied deployment results under different ratios of the four types of traffic. However, in each case, the differentiated resilience provisioning scenario provides the capability to accommodate much more traffic, more than twice of what can be supported in the dedicated protection scenario, which is usually used in traditional optical networks.

Shared protection provides fairly good capacity performance. However, it can only ensure that the failed traffic is restored within several hundred milliseconds. Although huge amounts of traffic in today's Internet have a relatively lower grade of resilience requirement, there are some services which actually have a higher resilience requirement and thus require traffic to be restored within 50 ms. Such a high resilience requirement is not achievable using shared protection scheme. In contrast, the differentiated resilience provisioning provides more choice: service with higher resilience requirements can be met with a faster restoration strategy.

As shown in the above figures, the different ratios of the four types of traffic offered present different deployment results. The ratio could be decided according to the actual applications. In other words, the resilience provisioning policies could be adapted according to different levels of network resources availabilities, to make best use of the network: when network resource is abundant, the higher resilience service might be provided for more of the traffic; when the spare network resources are at a lower level, relatively lower resilience might be employed. This solution can lead to a better trade off between network utilization and resilience services. Thus, the ratio of the different resilience classes would depend on both the service requirement and the available network resources.

6.4.3.2 Resource Utilisation

The differentiated resilience provisioning mechanism has another advantage in that the network resource can be more efficiently used as shown in Figure 6-25. In the dedicated protection scenario, due to the wavelength continuity constraint and dedicated protection requirement, only 87.6% of the total resource can be used. About 12.4% of the network resources are unable to be used by any more dedicated protected connections. In the shared protection scenario, 97% of the network resources are used either by primary paths or by backup paths whilst leaving about 3% network resource to be un-exploited. For the differentiated resilience provisioning scenario, due to its flexibility, all network resource can be utilised due to its flexibility.

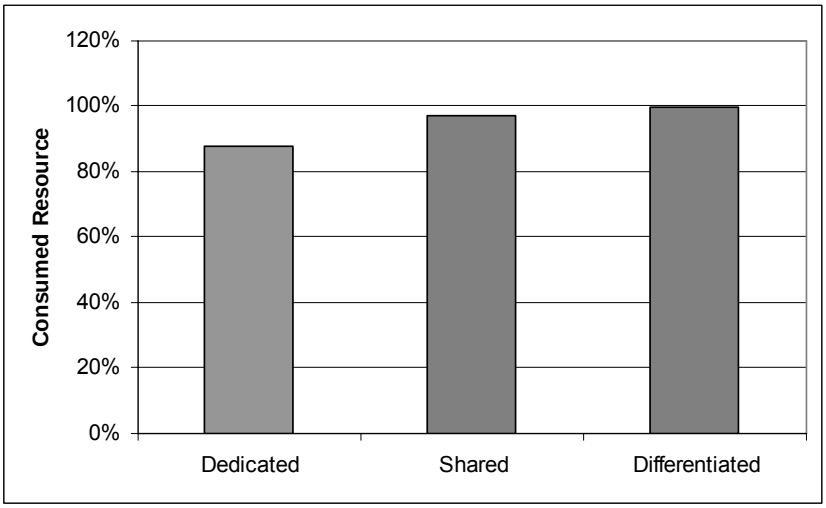


Figure 6-25: Useable Resource

6.4.3.3 Restoration Ratio of Single Link Failures

This section is to evaluate the performance of the proposed differentiated-resilience optical services model when a single link failure occurs in the network.

In this set of experiments, the network state is set in accordance with the deployment results of incremental traffic in the differentiated-resilience provisioning scenario listed in Section 6.4.3.1. That is to say, no spare resource is left in the network. The expected results would be that the RC2 and RC3 traffic can only be restored by pre-empting the RC4 traffic.

Without loss of generality, four separate cases of single link failure are presented here. Of the four links, two are located at the edge (Link0-3 and Link6-16) and two are in the core (Link11-14 and Link 14-19) of the Toronto Metropolitan Network.

Figure 6-26 to Figure 6-31 show the restoration results of the four types of traffic in different situations.

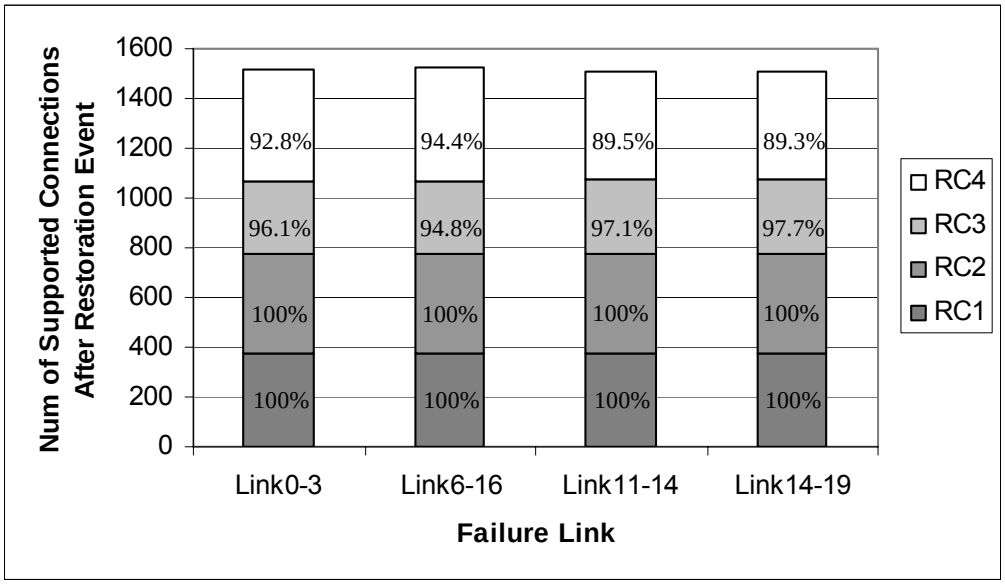


Figure 6-26: Restoration Ratio after Single Link Failure – A (2:2:1:1)

Figure 6-26 shows the number of supported connections after a restoration event for the deployment result shown in Figure 6-19.

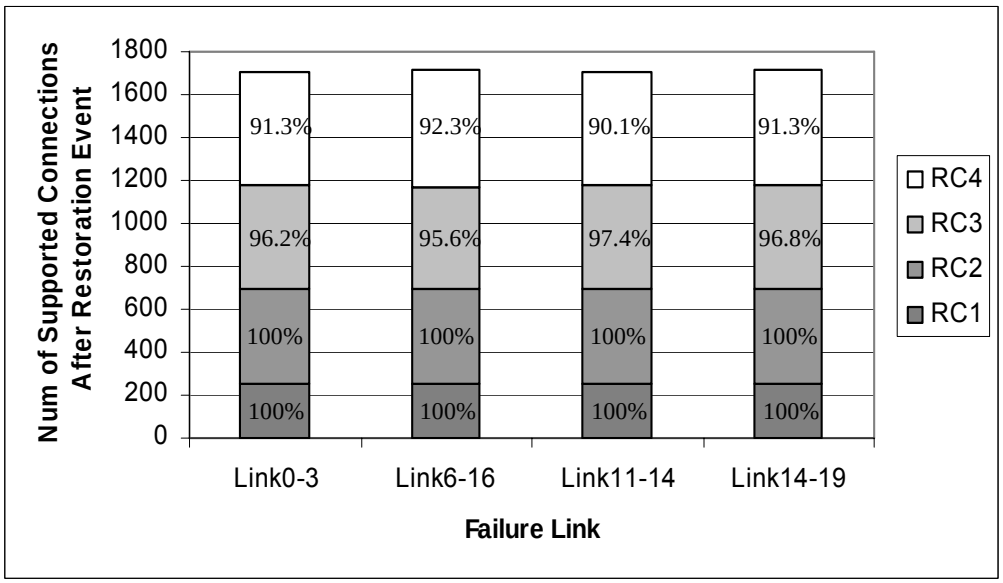


Figure 6-27: Restoration Ratio after Single Link Failure – B (1:1:1:1)

Figure 6-27 shows the number of supported connections after a restoration event for the deployment result shown in Figure 6-20

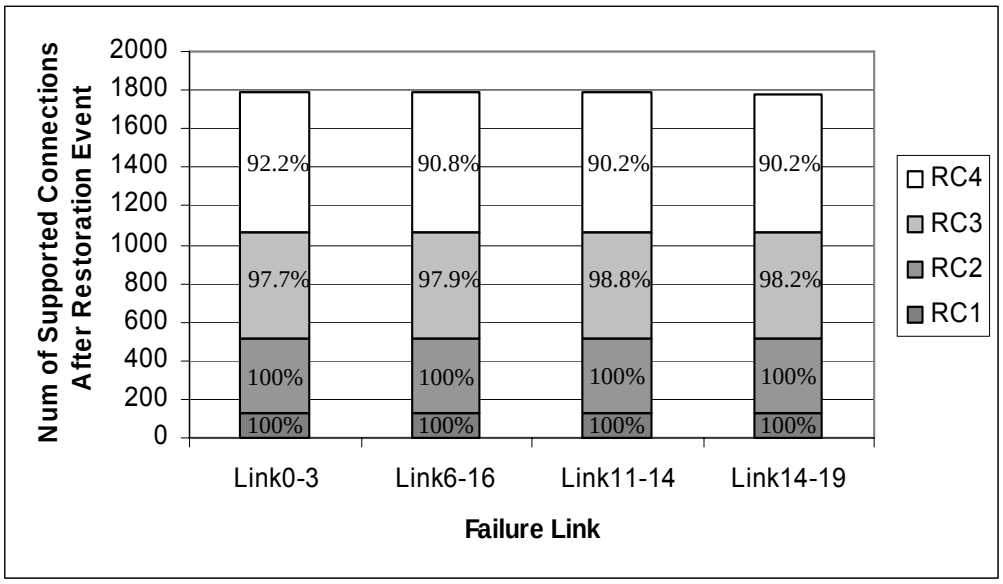


Figure 6-28: Restoration Ratio after Single Link Failure – C (1:2:3:4)

Figure 6-28 shows the number of supported connections after a restoration event for the deployment result shown in Figure 6-21.

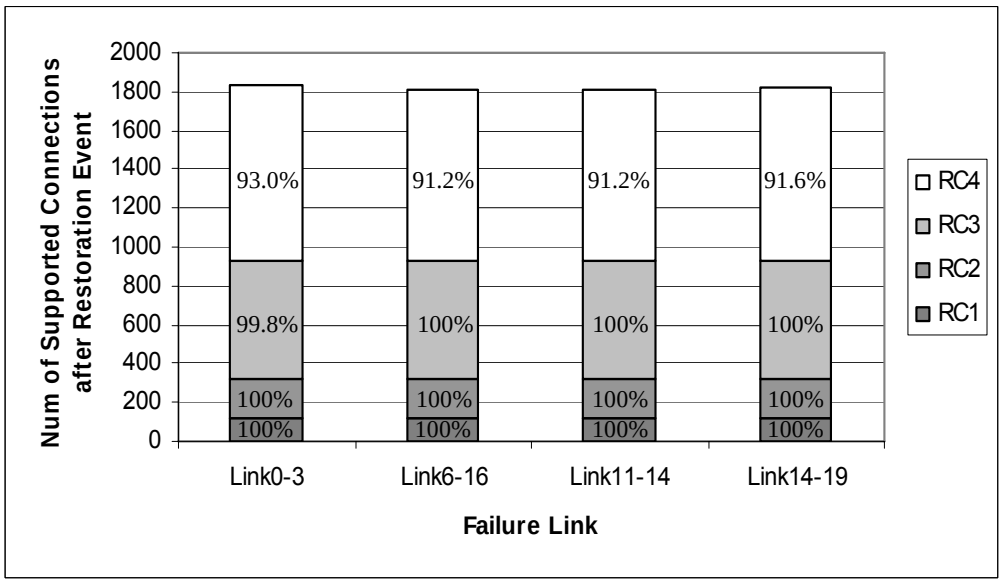


Figure 6-29: Restoration Ratio after Single Link Failure – D (1:1:3:5)

Figure 6-29 shows the number of supported connections after a restoration event for the deployment result shown in Figure 6-22.

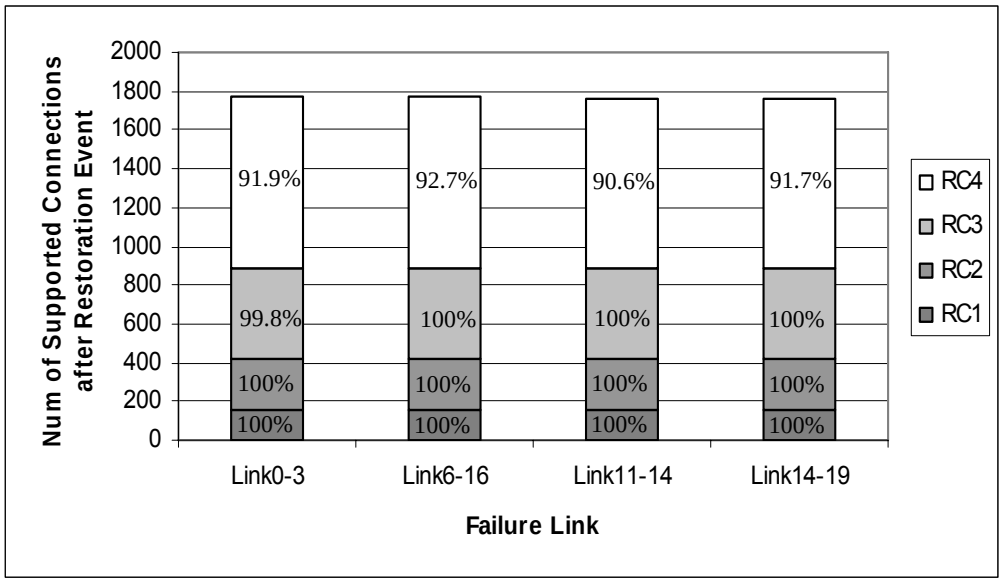


Figure 6-30: Restoration Ratio after Single Link Failure – E (1:1:2:4)

Figure 6-30 shows the number of supported connections after a restoration event for the deployment result shown in Figure 6-23.

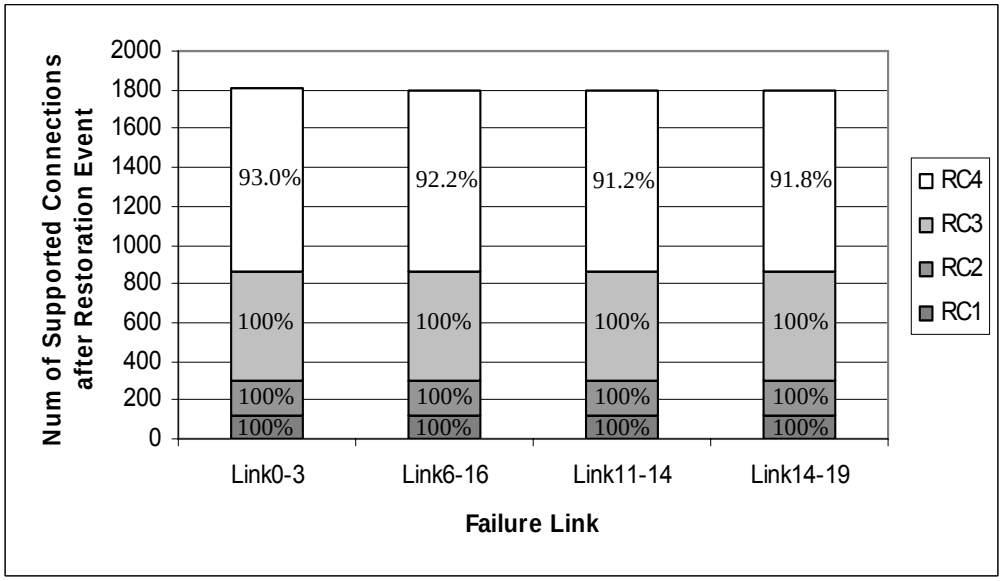


Figure 6-31: Restoration Ratio after Single Link Failure – F (1:1:3:6)

Figure 6-31 shows the number of supported connections after a restoration event for the deployment result shown in Figure 6-24.

In all the cases, as expected, all affected RC1 and RC2 traffic are restored. As there is no spare resource left in the network, the restoration of RC2 traffic is carried out by pre-empting RC4 traffic.

In all the cases, the affected RC4 traffic cannot be restored. This is because there is no spare resource in the network. In addition, the restoration of RC2 and RC3 traffic are performed by pre-empting RC4 traffic, which introduces further loss for RC4 traffic. As a result, certain connections of RC4 traffic (figures show it against the total RC4 connections supported in the network) are lost. This is not a surprise as RC4 traffic is proposed as best-effort traffic. It could be restored when resource is available.

What needs full exploration is the restoration results of RC3 traffic in the different study cases. While more than 95% of RC3 traffic are either maintained or restored, the losses of RC3 traffic due to the single link failure are different in the study cases. In some cases (Case A, B and C) some connections of RC3 traffic are lost whilst in others (Case D, E and F) all the affected RC3 traffic is restored. The result poses an interesting question: How to ensure a full RC3 traffic restoration under a single link failure? This is what network carriers care most because, if the RC3 traffic could be sure to be restored after a failure event, it is much easier for them to sell this service to their customers.

Since in all the study cases, the initial network states are set as that no spare resource is available, the restoration of RC2 and RC3 are achieved only by pre-empting RC4 traffic. Here, the ratios of actual deployed connections of RC2 and RC3 traffic to that of RC4 traffic are plotted in Figure 6-32.

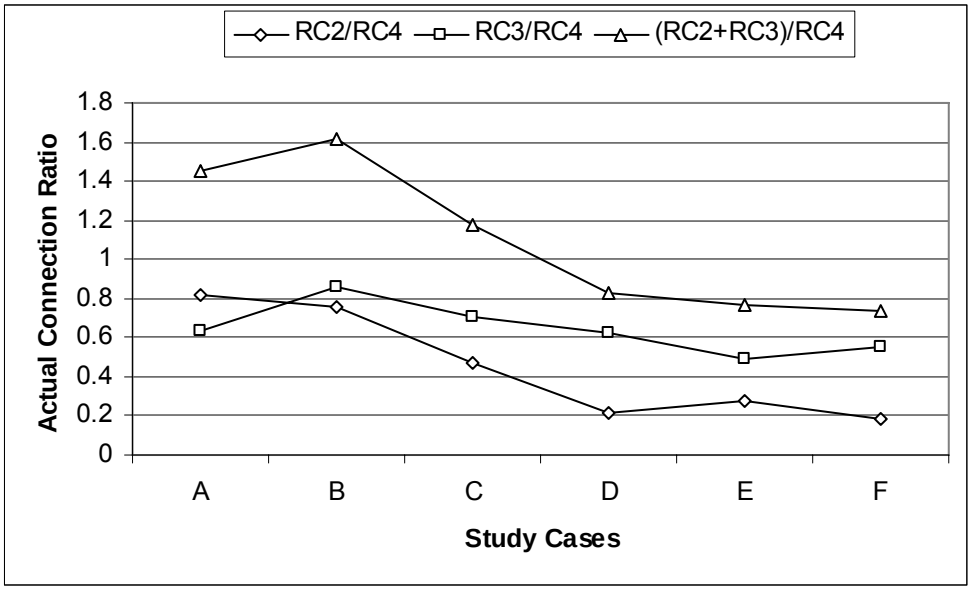


Figure 6-32: Actual Connection Ratio of Different Traffic in the Study Cases

Figure 6-32 shows that the loss of RC3 traffic due to a single link failure is roughly related to (RC2+RC3) / RC4, although the exact relationship is worth further studying. When the ratio of the total amount of RC2 and RC3 traffic to the amount of RC4 traffic is less than 0.8, there is a greater chance that all RC3 traffic affected by a single link failure can be restored. Notice that in Case D and F, in the event of Link0-3 failure, a very small portion (actually just one failed connection in each case) cannot be restored. Remember that in this irregular network topology, each link has been assumed with an equal number (40) of wavelength channels. If some optimisation is introduced at the network planning stage, the ratio of the sum of RC2 and RC3 traffic to RC4 traffic is expected to be higher to ensure full restoration for RC3 traffic. In other words, more resource could be used for traffic with a more stringent resilience requirement instead of best-effort traffic. Such an investigation is worth pursuing further.

Here, the property that all nodes (OXCs) in the network have no wavelength conversion capability further reduces the possibility of RC3 traffic being fully restored. In the wavelength-routed optical network, a lightpath must occupy the same wavelength on all the links along its path. If full or even partial wavelength conversion capability is available for each OXC, the ratio of the sum of RC2 and RC3 traffic to RC4 traffic is expected to be even larger with more resource that could be allocated to traffic with higher resilience requirement.

If the deployment results of Case A and F are put together as shown in Figure 6-33, an interesting point arises. Case A provides the same amount of guaranteed traffic as with dedicated protection whilst offering roughly the same amount of traffic over again in lower

resilience grades. In Case F, as RC3 traffic can be fully restored, it offers more guaranteed traffic and even more traffic over again in lower resilience grade. Now a question is whether a network carrier should adopt the deployment strategy of Case A or Case F. The author thinks the policy really depends on the application situations. If the guaranteed service requires a shorter restoration time, Case A could be adopted. If such a requirement is relatively loose, Case F could be followed. In each case, far more connections can be achieved by differentiated-resilience provisioning.

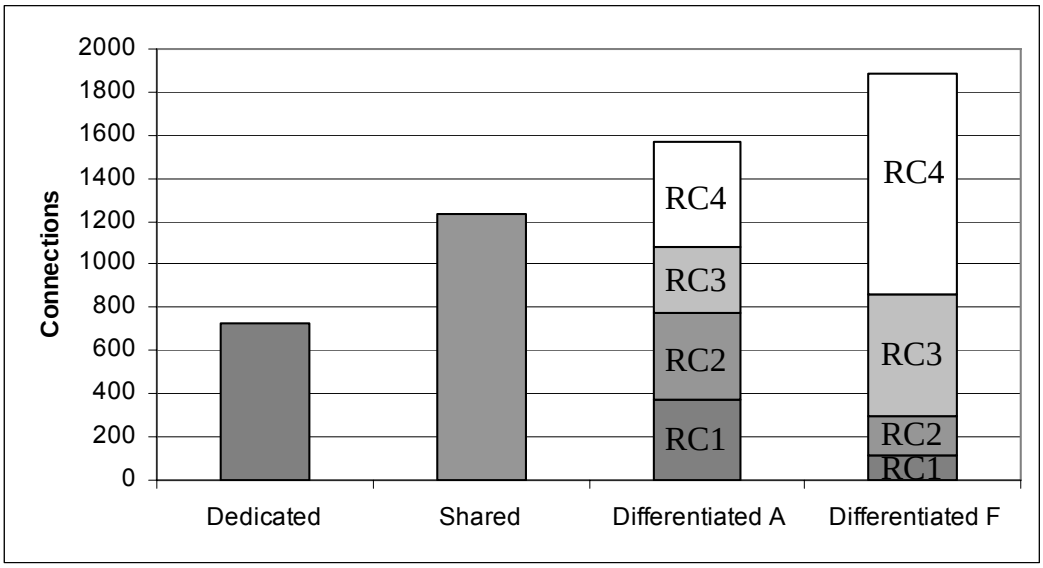


Figure 6-33: Choices of Different Network Deployment Pattern

In other ways, the resilience provisioning policies might also be adapted according to different states of network resources, such as resource usage, to make even better use of existing network resources. When network resource is abundant, higher resilience service might be provided for the traffic; when network resources are at a lower level, relatively lower resilience might be adopted. This solution can lead to a better trade off between network utilization and resilience services. Thus, the ratio of the different resilience would depend on both the service requirement and the states of the network. The differentiated resilience provisioning policy is a topic for further work.

6.5 Summary

Static resilience provisioning using full protection is usually adopted in traditional optical networks. It is very costly to use to cope with the dynamic data traffic usually presents varied resilience requirements.

In this chapter, a differentiated-resilience optical services model is proposed to provide optical resilience in the dynamic environment. This chapter argues that it is more cost-effective and more flexible to provide different resilience that better reflects the value of the traffic being carried.

The rationale and mechanisms of this model are introduced in detail. In order to evaluate its performance, extensive simulations have been carried out to compare the proposed scheme with the traditional single level resilience provisioning schemes. The merits of the dynamic differentiated-resilience provisioning and resource-sharing model are confirmed through simulation, and demonstrate that, by better matching the service provided to that required, a significant network resource saving can be made – or conversely, more traffic can be accommodated in the network.

Chapter 7 Differentiated Resilience Provisioning for Optical Networks with Wavelength Conversion Capabilities

This chapter extends and investigates the idea of differentiated resilience provisioning presented in Chapter 6 for the optical network with wavelength conversion capabilities.

7.1 Overview

Differentiated-resilience provisioning has been proved in Chapter 6 to be more cost-efficient than the single level resilience provisioning for the wavelength-routed optical network. It provides a more flexible means for network carriers to exploit their network by offering resilience grades that better reflect the value of the traffic being carried.

It is expected these merits apply to optical networks with wavelength conversion capabilities as well. However, with the presence of wavelength conversion capabilities in OXCs, a lightpath can occupy different colours of wavelength along its route, providing greater possibilities for resource sharing between different backup paths. Therefore, the work presented in this chapter is not to simply apply the model described in Chapter 6 to the new network environment. Instead, some improvements have been pursued to utilise the new property (wavelength conversion capabilities) in the OXC thoroughly.

This chapter is organised as the following: Section 7.2 details the improvements and the relevant mechanisms. Section 7.3 presents the simulation test and results. Section 7.4 serves as a summary.

7.2 Improvements to the Differentiated-Resilience Optical Services Model

Without the wavelength continuity constraint, a lightpath can take any colour of wavelength on the links along its route. Therefore it is more likely that resources could be shared between different backup paths. The basic rationale behind the extension is that, in the wavelength-routed optical network backup paths of RC1 traffic are not shared, whilst in networks with wavelength conversion capabilities efforts are put to reuse those as well.

7.2.1 Service Classification

The same set of optical resilience classes as that in Chapter 6 are proposed according to their respective resilience requirement as shown in Table 7-1.

Service Class	RC1	RC2	RC3	RC4
Resilience Requirement	High	Medium	Low	Best Effort
Restoration Time	< 50 ms	< 500 ms	< 2 s	< 60 s
Resilience Strategy	Category 1	Category 2	Category 4	Category 4

Table 7-1: Resilience Classes

7.2.2 Link Management

The same link management strategy as that in Chapter 6 is adopted to utilise the underlying resource. That is, a five states optical link attribute entitled *Link Status* is used for path calculation and resilience deployment.

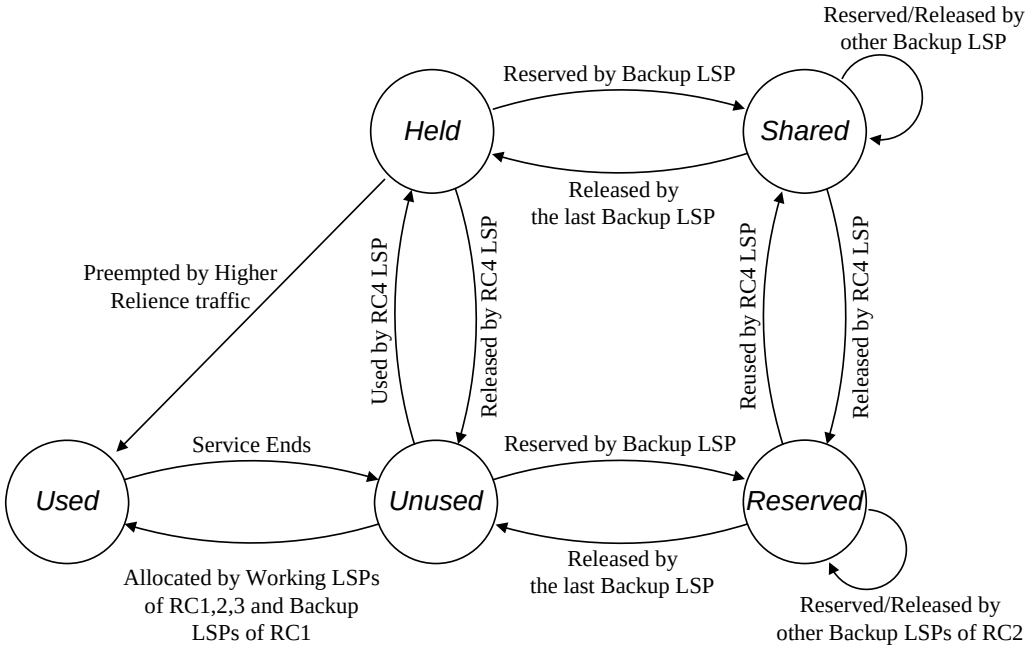


Figure 7-1: Link Status Finite State Machine

7.2.3 Resilience Strategies

7.2.3.1 General Strategy

As discussed as that in Chapter 6, traffic is generally protected or restored using end-to-end (path) restoration within an individual optical domain. Path restoration has the advantage of being more cost-efficient than link restoration. However, in a large optical domain, the time taken by path restoration for a long end-to-end optical path may be unable to satisfy the requirements of services with high resilience requirements (Resilience Class 1). In this case, long optical paths can be segmented into several pieces, with backup LSPs deployed for each piece using the ASPR algorithm proposed in [DON02], detailed in Chapter 5. This strategy is not needed for services with a low resilience requirement (Resilience Class 2, 3 and 4) since the extra time taken by end-to-end restoration is a relatively small part of the total restoration time.

7.2.3.2 Resilience Class 1

Optical services of Resilience Class 1 (RC1) have the highest resilience requirements and require traffic restoration within 50 ms.

For RC1 traffic, optical resilience category 1 is used. For example in **Figure 7-2**, a backup LSP AEIJ is set up for the working LSP AFJL. Since resilience category 1 performs all the actions (path calculation, path assignment, cross-connect) before the failure, it can ensure traffic restoration within 50ms. However, because all of the resources used by the backup LSP are dedicated, it is very expensive to deploy.

In order to reuse the dedicated resource, this chapter proposes a new optical link type for the optical opaque LSA, called the *Virtual Optical Link*. The whole backup LSP is treated as a *Virtual Optical Link*, which means that its resource can be reused as a whole. For example, this *Virtual Optical Link* AEIL can be used as backup LSP for LSP ADHL and/or as part of backup LSP for LSP BCGK, or used by pre-emptable traffic. In the former cases, LSP ADHL could be RC1 traffic and LSP BCGK could be RC2 traffic.

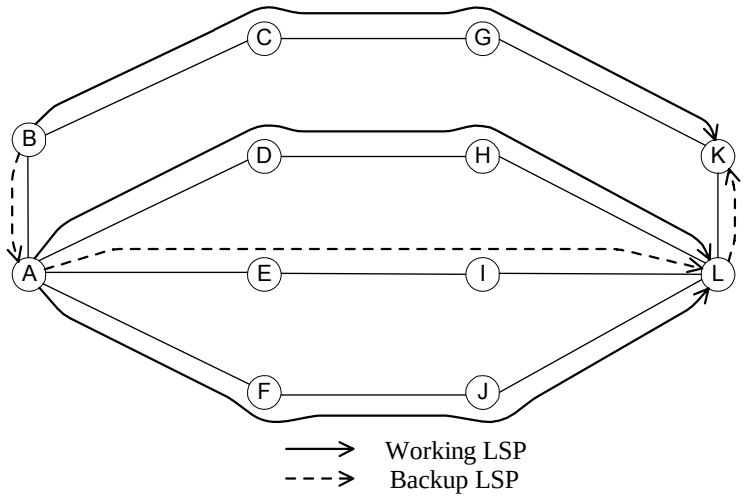


Figure 7-2: Resilience Strategy for RC1

The status of each of the links used by LSP AEIL, transfers from *Unused* to *Used*, which means each individual link cannot be used separately. A *Virtual Optical Link* is produced with its status set to *Reserved* and with additional parameters identifying which network components (i.e. nodes / links) it protects (i.e. LSP AFJL in this case). This *Virtual Optical Link* can be reused as a whole. If it is shared by further protecting LSP ADHL and/or LSP BCGK (in addition to LSP AFJL), the status remains *Reserved* with refreshed additional parameters. If it is used by preemptable traffic, the status becomes as *Shared*.

For RC1 services, the working LSP can use links whose status is *Unused* or *Held* when resources are limited and traffic preemption is permitted. Backup LSPs can use *Unused* and *Reserved* links.

Note here the improvement to the model in Chapter 6 is to reuse the resource occupied by backup paths of RC1 traffic. Compared to that in wavelength-routed network detailed in Chapter 6, here with wavelength conversion capabilities, chances of reusing such resource are much greater. Thus a mechanism (*Virtual Optical Link*) is proposed to realise this function.

After this improvement, the RC1 traffic employs 1:1 or N:M dedicated protection instead of 1+1 dedicated protection which is adopted by the model in Chapter 6.

7.2.3.3 Resilience Class 2

Optical services of Resilience Class 2 (RC2) have medium resilience requirements of being restored within 500 ms.

For RC2 traffic, optical resilience category 2 is used. The pre-calculated and allocated backup LSPs ensure the traffic can be restored. Since the cross-connects are set only after failure, the backup LSPs can be reused in pieces, as shown in Figure 7-3.

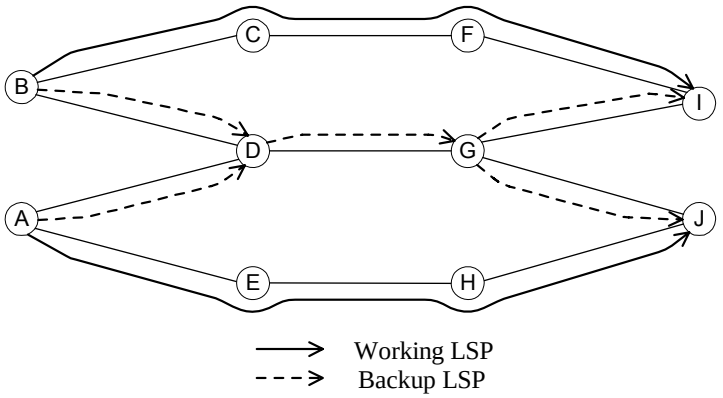


Figure 7-3: Resilience Strategy for RC2

At the time of LSP path calculation and deployment, the working LSP uses *Unused* links, or *Held* links if resource is limited and pre-emption is permitted. Then, the status of these links becomes *Used*. The backup LSP can use *Unused* links, or *Reserved* links, or *Held* links, or *Shared* links. Accordingly, these links become *Reserved*, *Reserved*, *Shared* and *Shared*, respectively.

7.2.3.4 Resilience Class 3

Optical services of Resilience Class 3 (RC3) have relatively low resilience requirements with restoration times less than 2 s.

For RC3 traffic, optical resilience category 4 is used, as category 3 is impractical. The node-disjoint / link-disjoint restoration LSPs are calculated and deployed only after a failure.

RC3 services cannot be pre-empted. So, at the time of the alternative LSP provisioning in response to a failure, only those links with a status of *Unused* or *Held* (if resource is limited and pre-emption is permitted) can be used.

7.2.3.5 Resilience Class 4

Optical services of Resilience Class 4 (RC4) have very low resilience requirements with restoration times around 60s. They are not guaranteed and can be pre-empted by services of all other resilience classes.

For RC4 traffic, optical resilience category 4 is used. The restoration of RC4 services starts at the time when OSPF has re-converged (within 60 s). At this time, each node has accurate information of the revised network states after the failure. Also, the links previously used by the failed services of type RC1, RC2 and RC3, having been reclaimed, will provide more resource for restoration. Unlike RC3, the restoration LSP of RC4 services can use *Unused* and *Reserved* links. However, the restoration of RC4 could fail due to there being no spare resource.

7.2.4 Functional Model

7.2.4.1 Resilience Provisioning

As described in Chapter 6, the resilience-provisioning algorithm is performed at the ingress OXC, which serves as the PSL. Different resilience strategies are provided for the different optical services. For RC1 and RC2 services, a pair of link-disjoint / node-disjoint working and backup paths is calculated according to the aggregated link information. Sharing of any link on the backup path is determined during signalling of the proposed backup path. Information about the working path is also carried in the signalling message. When it receives the signalling message, each OXC on the backup path decides whether the proposed backup path can reuse the resource already reserved for other backup paths. The decision is based on the detailed *Link Status* information, which is maintained in OXC's local database. For RC3 and RC4 services, only a working path is calculated and deployed.

7.2.4.2 Restoration Procedure

As backup paths of RC1 traffic are reused instead of being dedicated for one protection such as in Chapter 6, the restoration procedure therefore changes as the following.

Almost all the processes are originated and performed in the PSL. The PML performs active actions only for RC1 restoration after a failure notification is received.

When the PSL is notified of the failure, it starts the main restoration process. The main restoration process retrieves information about the failed services and produces three child processes, immediately (i.e. the RC1, RC2 and RC3 restoration processes). It also schedules the RC4 restoration process for subsequent action once reconvergence has taken place, as shown in Figure 7-4.

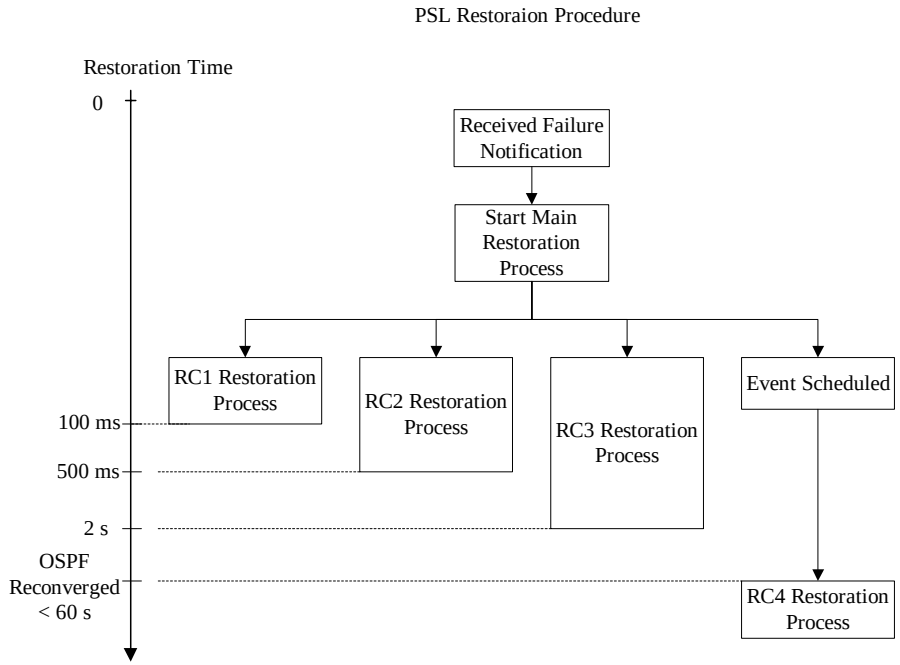


Figure 7-4: PSL Restoration Procedure

The RC1 restoration process takes charge of restoring the RC1 services. It involves the PSL switching the traffic onto the pre-calculated and deployed backup pipes.

The RC2 restoration process is responsible for the restoration of RC2 services. It involves the PSL sending out signalling messages to connect the pre-calculated and reserved link pieces. It does not require the PML to originate any action.

The RC3 restoration process restores RC3 services. It involves the PSL calculating the node/link-disjoint paths and deploying the calculated LSP, according to the link states. Only the *Unused* links and *Held* links (if resource is limited) are considered by the path calculation algorithm. Link contention may occur in this restoration and the connection request may fail. In this case, a new restoration path will be calculated and deployed by the PSL. The restoration does not require the PML to originate any action.

When the previously scheduled RC4 restoration event arises, the RC4 restoration process is activated. It is responsible for the restoration of RC4 services. At the time RC4 restoration process is activated, OSPF has already reconverged with the resource previously occupied by the failed LSPs having being reclaimed. The PSL will then calculate and deploy the restoration LSPs for RC4 services. Unlike RC3, *Unused* and *Reserved* links can be used by the restoration LSPs. The restoration does not require the PML to originate any action.

7.3 Performance Evaluation

In order to evaluate the proposed model, simulations are performed using the OPNET™ environment. The Toronto metropolitan network (Figure 7-5), with 25 nodes and 55 links, is used for a case study.

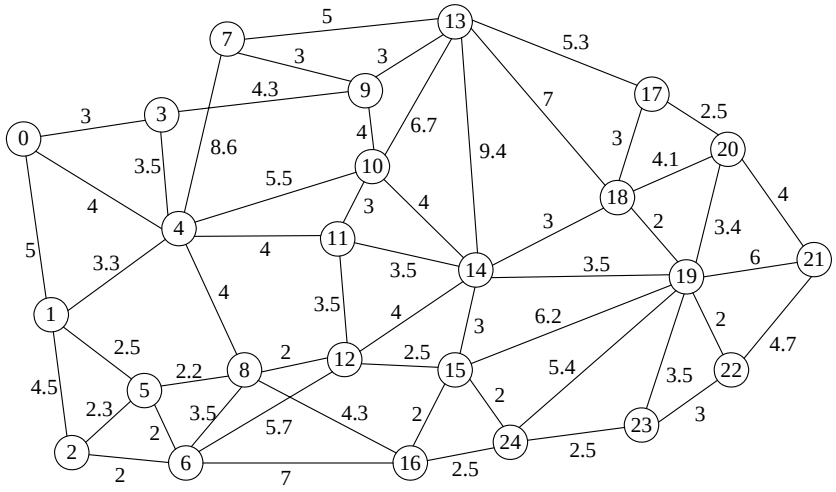


Figure 7-5: Example Network Topology

The traffic demand is assumed as a full-mesh set of 6000 (=24*10*25) lightpaths. Each connection request has a bandwidth of 1 unit. A ratio of 1:2:3:4 is used for RC1:RC2:RC3:RC4 traffic, respectively. The source and destination nodes are selected randomly for each lightpath. Dijkstra's shortest path first algorithm is used for the path calculation.

7.3.1 Network Capacity Requirement

The link capacities are initially set to infinity. For comparative study, two single level resilience-provisioning schemes, shared protection and 1+1 dedicated protection, are also implemented. Shared protection uses optical restoration category 2 and 1+1 dedicated protection uses optical restoration category 1. Both of them provide only a single resilience level. For comparison, the traffic demand in these two scenarios is 3600 (=24*6*25) lightpaths, which is equal to the total demand of RC1, RC2 and RC3 traffic in the differentiated resilience-provisioning scenario. The total bandwidth of each scheme is shown in Figure 7-6.

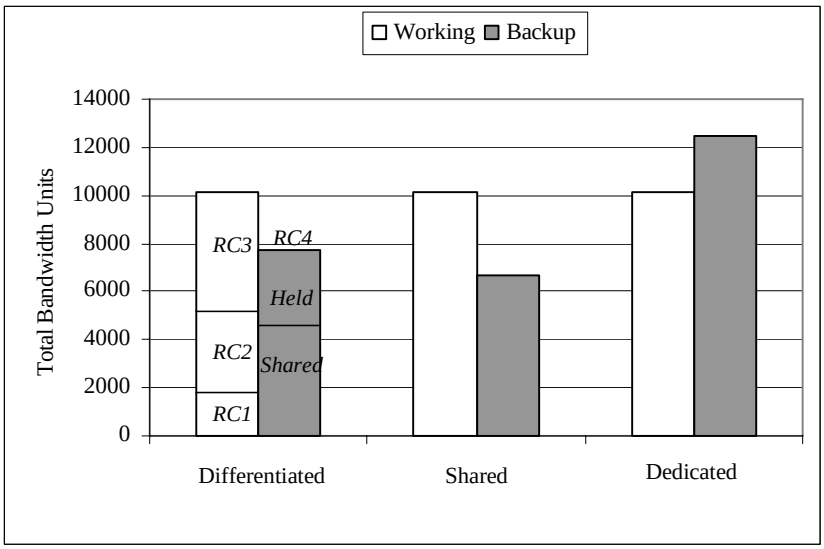


Figure 7-6: Network Capacity Requirement

The result shows that, the ratio of backup capacity to working capacity in the shared protection scenario is 65.6% and that of the dedicated protection scenario is 123.8%. In the differentiated-resilience provisioning scenario, the ratio of capacity consumed by RC4 traffic to that by RC1, RC2 and RC3 traffic is 76.1%, of which 56.6% is with the status of *Shared* and 43.3% is with the status of *Held*. Although the differentiated resilience provisioning scheme requires 6.3% more total capacity than that of shared protection, it provides 66.7% more services of a lower resilience.

7.3.2 Restoration Ratio of Single Link Failures

This section investigates the performance of the proposed scheme when a single link failure occurs. In this set of experiments the link capacities and network state are set to the traffic deployment results of the 6000 lightpaths of the differentiated-resilience provisioning scenario of Section 7.3.1.

Without loss of generality, failure events of four links (two edge links and two core links) are considered. The performance results are shown in Figure 7-7.

The figure shows that all the RC1, RC2 and RC3 traffic in these four scenarios are restored. However, different ratios of RC4 traffic are lost due to failure events and traffic preemptions. As there is no spare resource, all RC1, RC2 and RC3 traffic are restored, probably by preempting RC4 traffic.

All RC3 traffic is also fully restored after each of the four single failures. It is not a surprise as with the wavelength conversion capability being present in each node, there is a greater chance that RC3 traffic can be fully restored.

As the network state is set as the deployment result discussed in Section 7.3.1, there is no spare resource (*Link Status* as *Unused*) in the network. Therefore the affected connections of RC4 traffic cannot be restored. In addition, the restoration of affected RC1, RC2 and RC3 traffic is carried out by pre-empting RC4 traffic, which introduces further RC4 traffic loss. This is clearly exhibited in Figure 7-7.

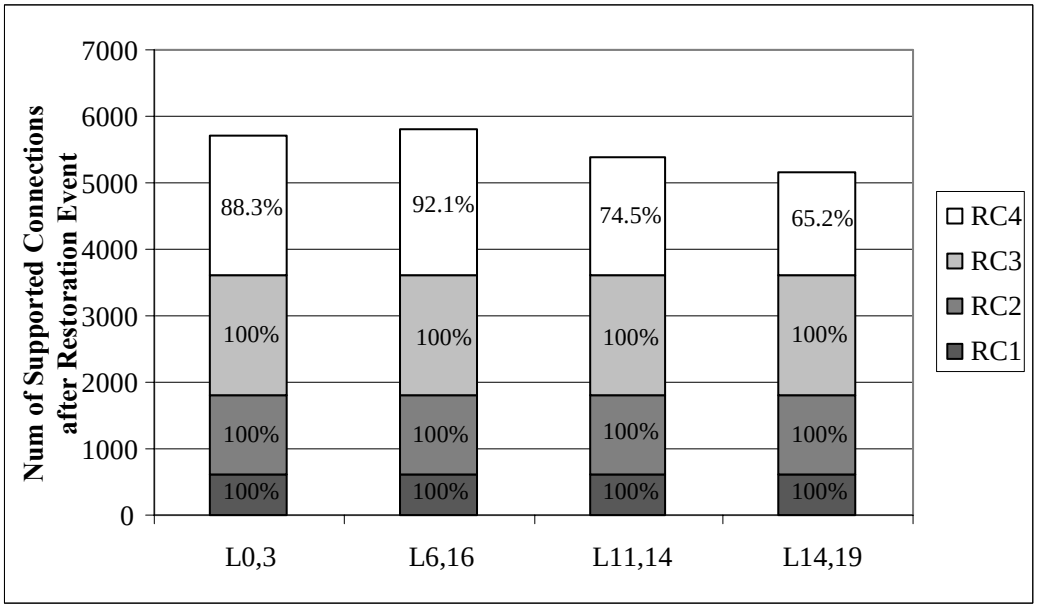


Figure 7-7: Restoration Ratio of Single Link Failure

Note here that the loss of RC4 traffic is greater than that in experiments with wavelength-routed networks presented in Chapter 6. This is because here backup paths of RC1 traffic could be shared by RC4 traffic. Therefore the restoration of RC1 traffic leaves more RC4 traffic being interrupted. Also, more of the RC4 traffic is potential restoration resource for any given RC3 traffic to restore itself with.

The loss of RC4 traffic due to different link failures, unlike that presented in Chapter 6, presents larger variations, with less connections being dropped due to edge link failures and more due to core link failures. For example, after the failure of Link 6-16, about 92.1% RC4 traffic is kept whilst after the failure of Link 14-19, about 34.8% of RC4 traffic fails. This is because here the capacity of each link is initially set as infinite. The establishment of the 6000 connections means that the core links are heavily loaded whilst the edge links are lightly loaded. Therefore, the core link failures have more traffic being affected and as a result more

RC4 traffic being preempted. In Chapter 6, in order to model the wavelength-routed optical network, each link is assumed to have the same number of wavelengths. Thus, there is no big difference between the traffic affected by core link failures and that by edge link failures.

Despite the difference, the result shows that differentiated-resilience provisioning also provides a more cost-efficient solution for the optical network with wavelength conversion capabilities. It also offers much more useful RC3 service, but an arguable much more lower value RC4 service.

7.4 Summary

Differentiated resilience provisioning provides a more cost-efficient solution by providing different resilience grades that better reflect the value of the traffic being carried. This has proven to be true in wavelength-routed optical networks by research presented in Chapter 6. In this chapter, this idea is further extended and applied to the optical network with wavelength conversion capabilities. Some improvements have been made on the model proposed in Chapter 6 to reuse the backup resource used by RC1 traffic. The simulation results show that the advantages of differentiated resilience provisioning hold also in optical networks with wavelength conversion capabilities.

Chapter 8 Discussion and Conclusion

8.1 Discussion

Current advances in optical communication technology are leading to flexible, highly configurable optical networks. The near future should see a migration from the current static wavelength-based control and operation to more dynamic IP-centric routing and resource management schemes. Future optical networking will most likely be based on fast circuit switching, in which end-to-end optical pipes are dynamically created and removed by means of signalling protocols and fast provisioning algorithms.

New resilience provisioning mechanisms are needed to support this evolution as resilience is usually provisioned statically in the traditional optical network. This research has focused on the study of dynamic resilience provisioning mechanisms for IP-centric optical networks. Different resilience mechanisms including static and dynamic schemes have been studied and compared thoroughly to find their inherent relations. Based on this, efforts have been made to develop novel resilience provisioning mechanisms in a dynamic environment.

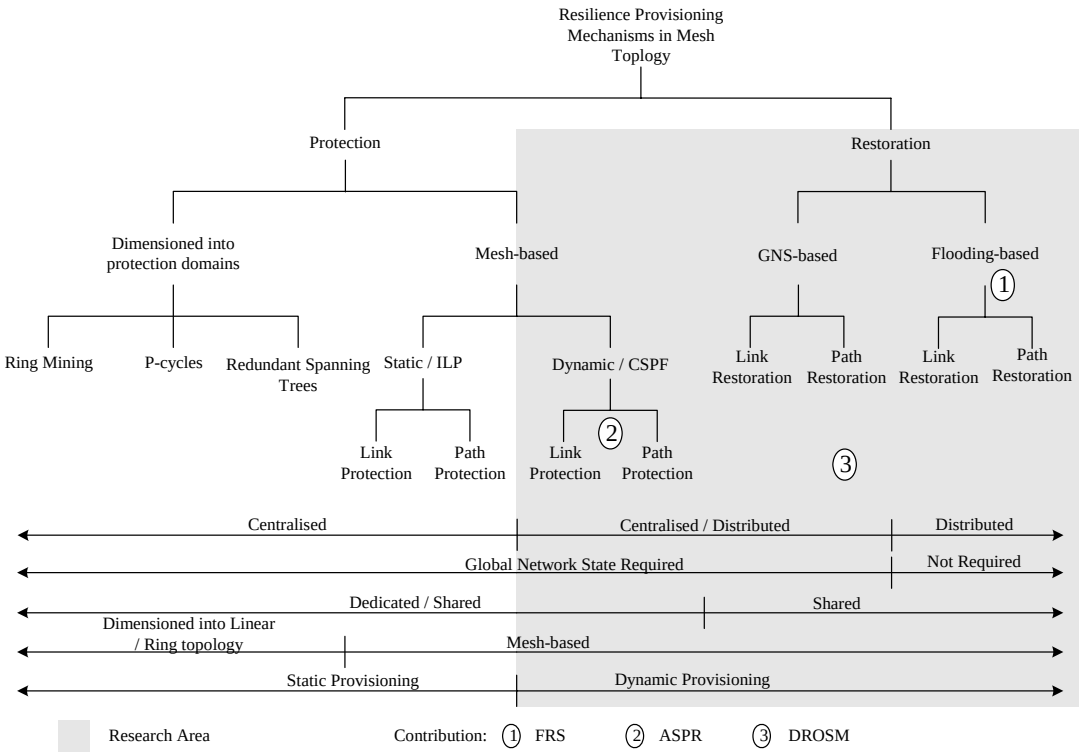


Figure 8-1: Research Focus and Contribution Summary

As shown in Figure 8-1, the following have been observed for optical resilience provisioning in a dynamic environment.

Dynamic resilience provisioning applies to networks with a mesh topology. In particular, dynamic resilience provisioning provides optical resilience based on the entire mesh network. In a static environment, resilience could be provided by dimensioning the mesh network into protection domains with simpler ring or linear topologies. However, the topology is built bearing in mind the anticipated traffic demands – which is difficult to determine in a dynamically provisioned environment. Although it is easy to operate and manage networks with a simple topology separately, design and management of a network comprised of these separate domains has proved to be very complicated [GRO02]. As a result, provisioning an optical connection and its resilience across domains in such a network is extremely manual and takes several months to accomplish. Therefore, in order to deploy optical resilience dynamically, it is essential that schemes operate on the entire network.

Dynamic resilience provisioning can employ schemes that either adopt dedicated protection or share backup resource between connections without a common failure component. In particular, protection could be either dedicated or shared. Dedicated protection possesses the shortest restoration time. However, because the resource pre-allocated by each backup path only serves one particular working path, dedicated protection requires excessive resources for protection. Shared protection is more cost efficient as working paths with no common failure component can share backup resource between each other. However, its restoration time is relatively longer than dedicated protection because signalling messages are needed to inform the interim nodes to set up the connection. Because the backup resource is pre-allocated before failure events, both dedicated and shared protection ensure a successful restoration.

Unlike protection, (reactive) restoration decides and allocates restoration paths only after the failure has occurred. Restoration has a better resource sharing as it is much more flexible at choosing alternative paths. In addition, this high flexibility makes restoration be much suitable for networks with a dynamic traffic demand. However, its restoration time is much longer than protection because of the time consumed by searching and establishing the alternative paths. In addition, as the alternative paths are not defined and the required resource is not allocated before the failure, restoration cannot guarantee successful restoration for all the affected traffic. Conversely, in order to achieve full restoration for all traffic, considerable

network resource redundancy is required. It is very complicated to determine the quantity and distribution of the redundant resource, especially in a dynamic environment.

The decision of restoration paths either before (for protection) or after (for restoration) requires detailed information of the network state. These network state properties include network topology, bandwidth usage, available bandwidth of each link, and detailed deployment result of existing traffic if efficient backup path sharing is required. Collection of these properties and provision of the backup / restoration path could be performed either in a centralised or a distributed mode. Static resilience provisioning usually adopts the centralised mode as it assumes all the traffic demands are already known and will remain constant, therefore, the central control point could utilised offline algorithms to achieve a globally optimal solution. However, in a dynamic environment, connection requests arrive one by one without knowledge of future demands. Traditional solutions that use offline algorithms to achieve a globally optimal solution no long apply. Even in a centralised mode, provisioning decisions have to be based on each single connection request, similar to the distributed approach. In addition, the communication for collecting network state and issuing command between the central control point and each node of the network introduces significant latency. The only advantage of a centralised mode in the dynamic environment is that it does not require every node in the network have high computing capabilities. Although the centralised mode might apply to very small networks in a dynamic environment, dynamic resilience provisioning should normally adopt a distributed mode.

Reactive restoration after a failure event can be carried out using two different mechanisms: Global Network State (GNS) based or flooding-based restoration. The first is by maintaining a database, which contains the global network state, either in a centralised or a distributed mode. For the centralised mode, a NMS is needed to collect the network state properties from each node in the network. For the distributed mode, each node of the network maintains a database containing aggregated information about the global network state. In this case, a link state routing protocol such as OSPF or IS-IS is needed to distribute and synchronise the network state properties.

By maintaining a database of the global network state, the NMS (centralised mode) or a network node (distributed mode) calculates restoration paths according to the current network state. The calculation algorithms that are based on SPF algorithm are conceptually simple and

easy to implement. However, the maintaining a global network state database requires extra expense. For the centralised mode, reliable connections are needed between the NMS and all nodes in the network. Its restoration is relatively slower. For the distributed mode, the database has only aggregated information of the link state of the network and cannot update in time immediately after the failure occurs. As a result, the path calculation algorithm may fail to find an alternative path while spare resource is available.

In contrast, flooding-based restoration does not require nodes in the network to maintain a global network state database; it does not need a NMS or a link state protocol and thus is cheaper to implement. It finds restoration paths using message flooding. The drawbacks of flooding-based restoration include: firstly, the restoration time is longer than that of GNS-based restoration since the search for alternative paths is performed by flooding messages, which take extra time in message propagation and processing. Secondly, the flooding messages make the communication overhead excessively high. In this case, a hop count is usually introduced to limit the flooding area of these messages. As a consequence, restoration paths can only be found in a confined area of the network, which may reduce the restoration ratio.

In brief, the development of WDM transmission technology and more recently emergence of optical multiplexers and optical cross-connect (OXC) devices are moving optical networks towards a vision of all optical networks. In order to offer abundant and inexpensive bandwidth to the end users, optical connections and resilience must be able to be deployed automatically and dynamically to achieve high flexibility and cost-efficiency. In a dynamic environment, connection requests arrive one by one without knowledge of future demands which makes traditional static resilience provisioning mechanisms inappropriate.

This research focuses on optical resilience provisioning in a dynamic environment. Particularly, the research is performed in a manner that is sympathetic to current efforts in the industry to utilise an IP-centric control plane to provide networking functionalities. The research has been carried out from different aspects based on observations of the new characteristics of resilience provisioning in a dynamic environment (shown in Figure 8-1).

The first contribution exploits the observation of that restoration is much more flexible and well suitable for a dynamic environment. Although the complex networking functionalities of the IP-centric control plane make the GNS-based restoration quite

straightforward, the flooding-based approach may still have value if complex functionalities are expensive to implement in all nodes in the network.

Flooding-based restoration uses flooding messages to discover alternative paths after the failure has occurred. It does not need each node of the network or a centralised NMS to maintain the global network state of the network, thus it is easy to implement.

Flooding-based restoration is initially designed for DCS and SONET networks, applying in the electrical domain. In this research, a new flooding-based reactive restoration scheme named Fast Restoration Scheme (FRS) is proposed to apply in the optical domain.

In addition, novel mechanisms have been included to enable it to finish more quickly and require less spare resource by achieving loop-free restorations. By maintaining a dynamically refreshing *Resource Table* in the *Receiver*, FRS precludes the possibility of link contention and finishes the restoration process with only one connection attempt. The capability of setting up restoration path between nodes other than just the *Receiver* and the *Sender* ensures loop-free restorations.

Simulations have validated its performance. The simulation results also show that the restoration time mainly consists of processing delay, transmission delay and propagation delay that are taken by processing and transmitting the flooding messages. A failure on a link with more working channels makes FRS take more time to restore the traffic. The restoration time can be reduced by improving the processing capability of each node and increasing the control channel capacity. However, because of the large number of flooding messages produced for the connections affected by a failure, the restoration normally takes several hundred milliseconds to several seconds to finish.

The second contribution is built on the observation that dynamic resilience provisioning needs to operate across the entire mesh network. Resilience provisioning mechanisms operating on the entire mesh network can be classified as either link or path protection (or restoration). These two schemes are the fundamental strategies to provision resilience in a dynamic environment.

However, both methods have their own limitations. Link protection (or restoration) could be too costly to implement whilst path protection (or restoration) could take a long time.

A new resilience-provisioning scheme named Adaptive Segment Path Restoration (ASPR) is proposed to offer a new means of providing resilience in the mesh network. In this

approach, a lightpath (or LSP) is divided into several segments. For each segment of the primary path, a backup path is established. The segmentation of the primary path is adaptive to the topology of the network, allowing for more efficient resource usage whilst yielding restoration times comparable to link restoration. The implementation of the proposed scheme needs a slight extension to the existing MPLS/GMPLS signalling protocols, which makes it simple to implement and be able to work automatically. The comparative study and simulation results of the proposed scheme with others show that ASPR has the best restoration time performance, while it remains better than most other restoration schemes in terms of the spare capacity requirement.

The significance of this scheme lies in that, in a mesh network, to provide link protection for an optical connection may be too expensive whilst to offer end-to-end path restoration may result in a long restoration time, which does not suit services with high resilience requirements; instead, by dividing a long optical connection into several segments with each segment being provided a backup path, the cost can be drastically reduced and, at the same time, fast restoration is able to be achieved. In addition, the introduction of a new concept of restoration length and its corresponding application in the scheme enable ASPR to achieve a fast restoration which is comparable to the link restoration.

This approach could also be used to protect against multiple failures in the mesh network. In a real network that covers a large area, multiple failures are not uncommon. Traditional static resilience provisioning mechanisms solve the problem by dimensioning the large network into smaller protection domains with special topologies. In each protection domain, failures can be assumed to arise separately. In this way, resilience provisioning schemes for only single failures can be applied. In a dynamic environment, it is more economical to protect a network as a whole. In this case, a long path may encounter multiple failures while in a shorter segment multiple failures can be regard as unlikely. However, evaluation of the application of ASPR in this area requires further research and validation.

The last contribution is based on observation of that the optical service presents more varied resilience requirements than before that can be utilised to increase cost-efficiency of the network.

As analysed in a former part of this thesis, static resilience provisioning can employ centralised offline algorithms to achieve a deployment result close to optimum. This is built on the assumption that all the traffic demands are known in advance and will remain constant

in the network. However, in a dynamic environment it is unrealistic to employ a centralised control system to provide real-time solutions. In addition, the dynamic characteristics of traffic make it impossible to achieve an overall optimal solution. Therefore, resilience for each connection is provisioned separately in a distributed pattern. This could result in inefficient exploitation of the network resources.

In order to improve efficiency of resilience provisioning in the dynamic environment, this research proposes to provide differentiated levels of resilience for optical services. This is built on observations that the once dominant voice traffic has been surpassed by data traffic which presents a more varied nature; the majority of data traffic has a relatively low resilience requirement. However, traditional resilience provisioning mechanisms provide full protection and treat all traffic equally. This has proved to be very costly and wasteful.

In contrast, this research suggests optical resilience should be provided selectively to better reflect the value of the traffic being carried. A differentiated resilience optical services model (DROSM) has been proposed where optical services are classified according to their resilience requirements. Each resilience class is then provided with a different restoration strategy. In addition, a novel resource management mechanism is put forward to coordinate different resilience classes.

The model is first applied to optical networks without wavelength conversion capabilities (wavelength-routed optical networks). Then it is further extended to apply in optical networks with wavelength conversion capabilities.

Simulations in both situations have proved that differentiated-resilience provisioning provides a more cost-efficient and flexible solution than single level resilience provisioning schemes. This has been witnessed in both dynamic and incremental traffic scenarios. Particularly in the incremental traffic scenario, for both networks with and without wavelength conversion capabilities, differentiated-resilience provisioning offers the ability to carry the same amount of “premium grade” (RC1 & RC2) guaranteed fast protection traffic as with purely dedicated protection, whilst offering approximately the same capacity over again in lower resilience quality connections, which have been gaining in popularity recently.

The simulation also shows that RC3 traffic can be fully restored if resource is available. This gives network carriers more choices to manipulate their networks. If the guaranteed service has a stringent requirement, more resource can be assigned to RC1 and RC2 traffic whilst RC3 and RC4 serve for lower resilience quality connections. If such a requirement is relatively loose, even RC3 could be employed for guaranteed services.

Alternatively, the resilience provisioning policies might also be adapted according to different states of network resources, such as resource usage, to make even better use of existing network resources. When network resources are abundant, higher resilience service might be provided for the traffic; when network resources are at a lower level, relatively lower resilience might be adopted. This solution can lead to a better trade off between network utilisation and resilience services. Thus, the ratio of the different resilience would depend on both the service requirement and the states of the network. The mechanism for dynamically tuned differentiated-resilience provisioning is for future work.

To sum up, this research has focused on resilience provisioning mechanisms in a dynamic environment. Three novel schemes have been proposed. However, these three schemes are not isolated from each other. They can function together within an integrated system.

8.1.1 Integration of FRS and DROSM

The basic idea of DROSM is to provide different levels of resilience that better reflect the value of traffic being carried. In the proposed model, all of the different restoration strategies assume each node in the network maintains a database containing information about the Global Network State (GNS). That resilience provisioning for different service adopts similar schemes (both backup and restoration paths are calculated using the GNS maintained in the database) enables them easier to coordinate with each other. Functional procedures are also more likely to be reused.

However, the synchronisation of these distributed databases requires signalling supports (e.g. OSPF-TE) and high processing capabilities at each node. When these supports and capabilities are not available or are limited, FRS - which requires less signalling support and processing capability - can be adopted.

In this situation, traffic with higher resilience requirements can utilise pre-planned dedicated protection or shared protection while that with a lower resilience requirement can use flooding-based restoration FRS. In particular, each node must be aware of which connections are with higher resilience and which connections use FRS. Furthermore, at the time of a failure, probe messages are only flooded for the latter.

8.1.2 Integration of ASPR and DROSM

DROSM gives the framework and strategies to provision differentiated-resilience for optical services.

In the proposed model, traffic is generally protected or restored using end-to-end (path) restoration within an individual optical domain. Path restoration has the advantage of being more cost-efficient than link restoration. However, in a large optical domain, the time taken by path restoration for a long end-to-end optical path may be unable to satisfy services with a high resilience requirement (RC1) and to restore traffic within 50 ms. In this case the Adaptive Segment Path Restoration (ASPR) scheme can be used to divide the long optical path into several segments, with a backup LSP deployed for each segment.

This strategy is not needed for services with a low resilience requirement (Resilience Class 2, 3 and 4) since the extra time taken by end-to-end restoration is a relatively small part of the whole restoration time.

8.2 Conclusion

This research focuses on dynamic resilience provisioning for the IP-centric optical network which has as yet not attracted much attention from the research community. The main efforts include, firstly, investigating the applications of existing resilience provisioning mechanisms in the new network environment, secondly, inventing new resilience provisioning schemes to cope with the mesh-based dynamic optical network environment.

The research starts with a thorough investigation of existing resilience provisioning mechanisms. A new classification framework as shown in Figure 8-1 is proposed to put all these schemes together. The purpose is to provide some insights into the inherent relations between these different resilience provisioning mechanisms, which might serve as clues to find new solutions.

The key contributions of this research are as follows:

Firstly, a novel flooding-based reactive restoration scheme named Fast Restoration Scheme (FRS) is proposed. FRS is a flooding-based restoration scheme applying to the optical layer. In addition, novel mechanisms make it be able to finish more quickly with only one attempt needed for each connection, and require less spare resource by achieving loop-free restorations.

Secondly, a resilience-provisioning scheme entitled Adaptive Segment Path Restoration (ASPR) is proposed to offer a new option to provide resilience in the mesh network. By dividing a long optical connection into several segments with each segment being provided a backup path, the cost can be drastically reduced, and at the same time fast restoration is able to be achieved. In addition, the introduction of a new concept of restoration length and its corresponding application in the scheme enable ASPR to achieve a fast restoration which is comparable to the link restoration.

Finally, a Differentiated-Resilience Optical Services Model (DROSM) for next generation optical networks is proposed. It suggests optical resilience be provided to better reflect the value of traffic being carried. In particular, optical services are classified according to their resilience requirements. Each resilience class is then provided with a different restoration strategy. In addition, a novel resource management mechanism is put forward to coordinate different resilience classes. The model is applied to both optical networks with and without wavelength conversion capabilities.

These schemes are not simply isolated and work alone. They could be integrated at appropriate conditions to provide a more comprehensive solution.

Although much effort has been put on the study of resilience provisioning in the new network environment, this is not to say the research is complete and exhaustive. Particularly, several immediate topics shown as follows are worth further pursuing.

8.3 Future Work

As stated before, one possible area for future work is to investigate the application of Adaptive Path Segment Restoration (APRS) for multiple failures. Dividing a working path into several segments and providing each a separate backup path can afford traffic protection against multiple failures with a certain spatial distribution. The merits of this method could be evaluated by analysis and simulation results.

Another future work topic is to examine the policy for differentiated resilience provisioning. Differentiated resilience provisioning has proven to be more cost-efficient and flexible than single resilience provisioning mechanisms. However, how to choose different ratios of the different resilience classes and what is its influence is on network carriers remains worth further pursuit. In addition, selecting different ratios for the resilience classes

based on changing circumstances could yield even better resource efficiencies. The solution could lead to a better trade off between network utilisation and resilience services.

Finally, a further topic of research is motivated by the following observation: Dynamic resilience provisioning utilises a distributed management system. The deployment of each optical connection and its resilience enactment are decided separately without taking into account other connections in the network. The resulting placement can differ significantly from an optimal solution for the whole network.

Therefore, future work could introduce some coordination between these separated nodes in a network. By communicating with others, a node may be able to make a more considerate decision for a connection's resilience provisioning. The communication strategy between these nodes could utilise Agent or other Artificial Intelligence (AI) technologies. The goal would be to see a more efficient and better deployment of finite resources across the whole network.

Author's Publications and Patents

[DON1] Dong Song, Phillips Chris and Friskney Robert, "Differentiated-Resilience Provisioning in the Wavelength-Routed Optical Network", 18th International Teletraffic Conference (ITC), Berlin, September 2003.

[DON2] Dong Song and Phillips Chris, "Adaptive Segment Path Restoration (ASPR) in MPLS Networks", IFIP & IEEE Net-Con' 2002, Paris, October 2002.

[DON3] Dong Song and Phillips Chris, "A Differentiated-Resilience Optical Services Model for Next Generation Optical Networks", ICWLHN 2002 & ICN 2002, Atlanta, August 2002.

[DON4] Dong Song and Phillips Chris, "A New Service Restoration Scheme in MPLS Networks", International Conference of Telecommunication 2002, Beijing, June 2002. Best paper award.

[DON5] Dong Song, Phillips Chris, and Lu Xiang, "Multi-Protocol Lambda Switched Fast Restoration for Meshed Optical Network", 17th UK Teletraffic Symposium, Dublin, May 2001.

[PAT1] Lu Xiang and Dong Song, "Fast Restoration in Optical Mesh Network," US patent filing number: 10/323,409, 2001

[PAT2] Friskney Robert, Dong Song and Phillips Chris, "Differentiated-Resilience Provisioning in Optical Networks," US patent filing number: 09/846,096, 2002

[PAT3] Friskney Robert, Baker Nigel, Davis Fiona, Dong Song and Phillips Chris, "Methods and Apparatus for Determining a Path in a Communication Network," Patent filed 30/9/2003, US filing number not yet allocated. Nortel ref "16087ID".

References

[ALR00] Al-Rumaih A., et al., “Spare capacity planning for survivable mesh networks,” in Proceedings IFIP-TC6 networking 2000, Paris, May 2000

[ANA00] V. Anand and C. Qiao, "Dynamic Establishment of Protection Paths in WDM Netowrks, Part I", in 9th International Conference on Computer Communications and Networks, October 2000.

[ASH01] Ashwood-Smith P., et al., “Generalized MPLS – Signalling Functional Description,” IETF RFC 3471, January 2003

[ASH02] Ashwood-Smith P., et al., "Generalized MPLS Signalling - CR-LDP Extensions", IETF RFC 3472, January 2003

[ASS01] Assi C., et al., “Optical Networking and Real-time Provisioning: An Integrated Vision for the Next-Generation Internet,” IEEE Network, pp. 36-45, July/August 2001

[AWD01] Awduche D and Rekhter Y., “Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control with Optical Crossconnects,” IEEE Communications Magazine, March 2001

[BAN01] Banerjee A., et al., “Generalised Multiprotocol Label Switching: An Overview of Routing and Management Enhancements,” IEEE Communications Magazine, Vol. 39, No. 1, January 2001, pp. 144-150

[BAR96] Barry R.A. and Humblet P.A., "Models of Blocking Probability in All-Optical with and without Wavelength Changers", *IEEE JOURNAL of Selected Area in Communications*, Vol. 14, No. 5, pp. 858-867, 1996.

[BAU97] Bauer F. and Varma A., "ARIES: A Rearrangeable Edge-based Online Steiner Algorithm", *IEEE Journal Selected Areas on Communication*, Vol. 15, pp. 382-397, April 1997

[BEN01] Benjamin D., et al., “Optical Services over the Intelligent Optical Network,” IEEE Communications Magazine, September 2001, pp. 73-78

[BER02] Berger L., et al., "Generalized MPLS Signaling - RSVP-TE Extensions", IETF RFC 3473, January 2003

[BHA99] Bhandari R., “Survivable Network- Algorithms for Diverse Routing,” Kluwer Academic Publications, 1999

[BHAK83] Bharath-Kumar K. and Jaffe J.M., "Routing to Multiple Destinations in Computer Networks", *IEEE Transaction on Communications*, Vol. COM-31, pp. 343-351, Mar. 1983

[BIC93] Bicknell J., Chow C.E., and Syed S., "Performance Analysis of Fast Distributed Network Restoration Algorithms", in *IEEE GLOBECOM '93*, 1993.

[BIR96] Birman A., "Computing Approximate Blocking Probabilities for a Class of All-Optical Networks", *IEEE Journal of Selected Area in Communications*, Vol. 14, No. 5, pp. 852-857, 1996.

[BLA02] Black U., “Optical Networks: Third Generation Transport Systems”, New Jersey: Prentice Hall, 2002

[CAE98] Caenegem B.B., et al., “Dimensioning of Survivable WDM Networks,” IEEE Journal on Selected Areas of Communications, Vol. 16, No. 7, pp. 1146-1157, September 1998

[CHA98] Chan V.W.S., et al., “Architecture and Technologies for High-Speed Optical Data Networks,” Journal of Lightwave Technology, Vol. 16, No. 12, December 1998, pp. 2146-2168

[CHE98] Chen S. and Nahrstedt K., “An Overview of Quality-of-Service Routing for the Next Generation High-Speed Networks: Problems and Solutions,” IEEE Network, Special Issue on Transmission and Distribution of Digital Video, November/December 1998

[CHEC95] Chen C.-C. and Chen J., "Vertex-disjoint Routings in Star Graphs", in *IEEE* First Conference on Algorithms and Architectures for Parallel Processing, April 1995

[CHO93] Chow C.E., Bicknell J., and McCaughey S., "A Fast Distributed Network Restoration Algorithm", in *12th Phoenix Conference on Computer and Communication* March, 1993.

[CHO99] Chow C.E. and Hansmats A., "Design and Analysis of One Prong Network Restoration Algorithms", in *Performance, Computing and Communications Conference, 1999 IEEE International*

[COA91] Coan B.A., et al, “Using Distributed Topology Update and Preplanned Configurations to Achieve Trunk Network Survivability,” IEEE Transactions on Reliability, Vol. 40, No. 4, pp. 404-416, October 1991

[DAC02] Dacomo A., et al., "Design of Static Resilient WDM Mesh Networks with Multiple Heuristic Crieria", in 21st Annual Joint Conference of the IEEE Computer and Communications Societies, June 2002

[DAV00] Davie B. and Rekhter Y., “MPLS Technology and Applications,” Academic Press, 2000

[DEM99] Demeester P., et al, “Resilience in Multiplayer Networks”, IEEE Communications Magazine, August 1999, pp. 70-75

[DOV99] R.D. Doverspike, et al., "Fast Restoration in a Mesh Network of Optical Cross-connects", in *OFC-99*. San Diego, Feburary 1999.

[DOV01] Doverspike Robert and Yates J., “Challenges for MPLS in Optical Network Restoration,” IEEE Communication Magazine, pp. 89-96, February 2001

[DOVC] Dovrolis C. and Ramanathan P., "Resource Aggregation for Fault Tolerance in Integrated Service Networks", Computer Communication Review, April 1998.

[FAR01] Farrel A. and Miller B., "Surviving Failures in MPLS Networks", White Paper, Data Connection Limited, http://www.dataconnection.com, Feburary 2001.

[FIN97] Finn S.G., Medard M.M., and Barry R.A., "A Novel Approach to Automatic Protection Switching Using Trees", *IEEE International Conference on Communication (ICC) 1997*, Montreal, June 1997

[FUJ94] Fujii H. and Yoshikai N., “Double Search Self-healing Algorithm and its Characteristics,” *Electron. Commun. Jpn.*, pt. 1, Vol. 77, No. 3, 1994

[FUM00] Fumagalli A. and Valcarenghi L., "IP Restoration vs. WDM Protection: Is There an Optimal Choice?" IEEE Network, November/December 2000, pp. 34-41

[GAL98] Gallager R.G., et al., "Multicast Automatic Protection Switching in Arbitrary Redundant Graphs", *IEEE International Conference on Communications (ICC) 1998*, June 1998

[GAN01] Gan D.-H., et al., "A Method for MPLS LSP Fast-Reroute Using RSVP Detours", Internet Draft draft-gan-fast-reroute-00.txt, work in progress, April 2001.

[GEO99] Georgatsos P., T'Joens Y., et al, "Towards Resilient Network and Services," ACTS NIG-G5.

[GER00] Gerstel O. and Ramaswami R., "Optical Layer Survivability – An Implementation Perspective", IEEE Journal on Selected Areas in Communications, Vol. 18, No. 10, October 2000, pp. 1885-1899

[GER00a] Gerstel O. and Ramaswami R., "Optical Layer Survivability: A Services Perspective", *IEEE Communications Magazine*, pp. 104-113, March 2000.

[GOR01] Goralski W., "Optical Networking & WDM," Osborne: McGraw-Hill, 2001

[GRO87] Grover W.D., "The Selfhealing Network: A Fast Distributed Restoration Technique for Networks Using Digital Crossconnect Machines," in Proceedings of GLOBECOM'87, Vol. 2, pp. 1090-1095, 1987

[GRO91] Grover W.D., et al, "Development and Performance Assessment of a Distributed Asynchronous Protocol for Real-time Network Restoration," IEEE Journal on Selected Areas in Communications, Vol. 9, No. 1, pp. 112-125, January 1991

[GRO98] Grover W.D. and Stamatelakis D., "Cycle-oriented distributed preconfigration: ring-like speed with mesh-like capacity for self-planning network restoration", *ICC '98*, 1998.

[GRO99] Grover W.D., Iraschko R.R., and Zheng Y., "Comparative Methods and Issues in Design of Mesh-Restorable STM and ATM Networks,"

Telecommunication Network Planning, pp. 169-200, Kluwer Academic Publishers, 1999

[GRO00] Grover W.D. and Stamatelakis D., "Bridging the Ring-mesh Dichotomy with P-cycles", in *Proceedings of Design of Reliable Communication Networks (DRCN 2000)*. Munich Germany, April 2000

[GRO02] Grover W.D., et al., “New Options and Insights for Survivable Transport Networks,” IEEE Communication Magazine, January 2002

[GU96] Gu Q.-P. and Tamaki H., "Routing a Permutation in the Hypercube by Two Sets of Edge-disjoint Paths", in *10th International Parallel Processing Symposium*, April 1996.

[HAS00] Haskin D. and Krishnan R., "A Method for Setting an Alternative Label Switched Paths to Handle Fast Route", Internet Draft draft-haskin-mpls-fast-reroute-05.txt, work in progress, November 2000.

[HAW95] Hawker I., Johnson D. and Chng R., “Distributed Restoration in Telecommunications Networks,” Fifth IEE Conference on Telecommunications, pp. 26-29, Mar 1995

[HER94] Herzberg M. and Bye S., “An Optical Spare-Capacity Assignment Model for Survivable Networks with Hop Limits,” IEEE GLOBECOM’94, pp. 1601-1607, 1994

[HERM97] M. Heraberg, D. Wells, and A. Herschtal, “Optimal resource allocation for path restoration in mesh-type self-healing networks,” in Proceedings of 15th International Teletraffic Congress, Washington, D.C., June 1997, vol. 15

[HO02] Ho P.-H. and Mouftah H.T., "A Framework for Service-Guaranteed Shared Protection in WDM Mesh Networks", *IEEE Communications Magazine*, pp. 97-103, February 2002.

[IRA98] Iraschko R., MacGregor M., Grover W.D., “Optimal Capacity Placement for Path Restoration in STM or ATM Mesh Survivable Networks,” IEEE/ACM Transactions on Networking, Vol. 6, No. 3, June 1998

[ITUG872] ITU-T Rec. G.872, “Architecture of Optical Transport Networks”, February 1999

[IWA01] Iwata A., et al., "Crankback Routing Extensions for MPLS Signaling", Internet Draft draft-iwata-mpls-crankback-01.txt, work in progress, July 2001.

[JUE] Jue J.P., “Lightpath Establishment in Wavelength-Routed WDM Optical Networks,” http://www.utdallas.edu/~jjue/cs6v81/articles/dle.pdf

[JUL94] Julka V., Cassandras C.G., “Optimal Call Admission in Circuit-switched Networks,” in Proceedings of the 33rd IEEE Conference on Decision and Control, Vol. 3, pp. 14-16, December 1994

[KAH92] Kahng A.B. and Robins G., "A New Class of Iterative Steiner Tree Heuristics with Good Performance," *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 11, Issue. 7, pp. 893-902, July 1992

[KAM02] Kamei S. and Kakugawa H., "A Self-stabilizing Algorithm for the Steiner Tree Problem", in *21st IEEE Symposium on Reliable Distributed Systems*, October 2002.

[KAT01] Katz D., et al., “Traffic Engineering Extensions to OSPF,” Internet draft

[KAW99] Kawamura R. and Ohta H., “Architectures for ATM Network Survivability and Their Field Deployment”, IEEE Communications Magazine, August 1999, pp. 88-94

[KOM90] Komine H., et al., "A Distributed Restoration Algorithm for Multiple-link and Node Failures of Transport Networks", in *IEEE GLOBECOM '90*, 1990.

[KOM01] Kompella K., et al., "Routing Extensions in Support of Generalized MPLS", Internet Draft draft-ietf-ccamp-gmpls-routing-02.txt, work in progress, February 2001

[KOM02] Kompella K., et al., "OSFP Extension in Support of Generalized MPLS", Internet Draft draft-ietf-ccamp-osf-gmpls-extensions-04.txt, work in progress, February 2002

[LAB92] Labourdette J.P., Acampora A.S., and Hart G.W., "Reconfiguration Algorithms for Rearrageable Lightwave Networks", in INFOCOM'92, May 1992

[LAN00] Lang J.P., et al., “Link Management Protocol (LMP),” Internet draft, July 2000

[LI01] Li G., et al., "Experiment in Fast Restoration using GMPLS in Optical / Electronic Mesh Networks", in *OFC-2001*. Anaheim, CA, March 2001

[LI02] Li G., et al., "Efficient Distributed Path Selection for Shared Restoration Connections", *INFOCOM’02*. New York, June 2002.

[LIU01] Liu Y., Tipper D. and Siripongwutikorn P., “Approximating Optimal Spare Capacity Allocation by Successive Survivable Routing,” IEEE INFOCOM, 2001

[LIUK02] Liu Kevin H., “IP over WDM,” John Wiley & Sons, 2002

[MAE98] Maeda M. W., “Management and Control of Transparent Optical Networks,” IEEE Journal on Selected Areas in Communications, Vol. 16, No. 7, September 1998, pp. 1008-1023

[MAK92] Makki K. and Pissinou N., "The Steiner Tree Problem with Minimum Number of Vertices in Graphs", in *Proceedings of the Second Great Lakes Symposium on VLSI*, Feb 1992.

[MAN02] Manohar P., Manjunath D., and Shevgaonkar R.K., "Routing and Wavelength Assignment in Optical netowrks from Edge-disjoint Path Algorithms", *IEEE Communications Letters*, Vol. 6, Issue 5, pp. 211-213, May 2002.

[MANI00] Mandoiu I.I., Vazirani V.V., and Ganley J.L., "A New Heuristic for Rectilinear Steiner Trees", *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 19, Issue. 10, pp. 1129-1139, Oct. 2000

[MED99] Medard M., et al., "Redundant Trees for Preplanned Recovery in Arbitrary Vertex-Redundant or Edge-Redundant Graph", *IEEE/ACM Transactions on Networking*, Vol. 7, No. 5, pp. 641-652, October 1999.

[MOH00] Mohan G. and Murthy C.S.R., “Lightpath Restoration in WDM Optical Networks”, IEEE Network, November/December 2000, pp. 24-32

[MOY98] Moy J., “OSPF Version 2”, IETF RFC 2328, April 1998

[MUR96] Murakami K. and Kim H. S., “Virtual Path Routing for Survivable ATM Networks”, IEEE/ACM Transactions on Networking, Vol. 4, No. 1, February 1996, pp. 22-39

[OH00] Oh T.H., Chen T.M., and Kennington J.L., “Fault restoration and spare capacity allocation with QoS constraints for MPLS networks,” IEEE GLOBECOM., Nov. 2000

[OU02] OU C., Mukherjee B., and Zang H., "Sub-Path Protection for Scalability and Fast Recovery in WDM Mesh Networks", in *OFC2002*March 2002.

[PAI97] Pai D.J. and Owen H.L., “An Algorithm for Bandwidth Management with Survivability Constraints in ATM Networks,” in Proceedings of ICC’97, Vol. 1, pp. 261-266

[RAM99a] Ramamurthy S. and Mukherjee B., “Survivable WDM Mesh Networks, Part I – Protection,” IEEE INFORCOM’99, pp. 744-751, 1999

[RAM99b] Ramamurthy S. and Mukherjee B., “Survivable WDM Mesh Networks, Part II – Restoration,” IEEE ICC’99, Vol. 3, pp. 2023-2030, June 1999

[RAMR01] Ramamurthy R., et al., “Capacity Performance of Dynamic Provisioning in Optical Networks,” Journal of Lightwave Technology, Vol. 19, No. 1, January 2001

[RFC3037] Thomas B. and Gray E., "LDP Applicability", RFC 3037, January 2001.

[RFC3036] Andersson L., et al., "LDP Specification", RFC 3036, January 2001.

[RFC3209] Awduche D., et al., "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC3214] Ash J., et al., "LSP Modification Using CR-LDP", RFC 3214, January 2002

[SAH03] Saha D., Rajagopalan B. and Bernstein G., “The Optical Network Control Plane: State of the Standards and Deployment,” IEEE Optical Communications, August 2003, pp. 529-534

[SCH98] Schrijver A., “Theory of Linear and Integer Programming,” John Wiley & Sons, 1998

[SCHD02] Schupke D.A., Gruber C.G., and Autenrieth A., "Optimal Configuration of P-cycle in WDM Networks", in *ICC 2002*, May 2002

[SEM94] Semal P. and Wirl K., “Optimal Clustering and Ring Creation in the Network Planning System PHANET,” 6th International Network Planning Symposium – Networks’94, Budapest, Hungary, September 1994, pp. 303-308

[SEN01] Sengupta S., Ramamurthy R., “From Network Design to Dynamic Provisioning and Restoration in Optical Cross-Connect Mesh Networks: An Architectural and Algorithmic Overview,” IEEE Network, pp. 46-54, July/August 2001

[SEN01a] Sengupta S. and Ramamurthy R., "Capacity Efficient Distributed Routing of Mesh-Restoraed Lightpaths in Optical Networks", in *GLOBECOM'01*2001

[SHA01] Sharma V., et al., "Framework for MPLS-based Recovery", Internet Draft draft-ietf-mpls-recovery-frmwrk-03.txt, work in progress, July 2001.

[SHI] Shi J.X., Fonseka J.P., “Dimensioning of Self-Healing Rings and Their Interconnections,”

[SHY99] Shyur C.-C., Lu T.-C., and Wen U.-P., “Applying Tabu Search to Spare Capacity Planning for Network Restoration,” Computers & Operations Research, Vol. 26, No. 10, pp. 1175-1194, October 1999

[SOR97] Soriano P., et al., "Multi-SHR Design for Zonal Networks", in *EURO / INFORMS Joint Meeting*. Barcelona, Spain, July 1997.

[SOR98] Soriano P., et al., "Designing Survivable Networks with Multiple Self-Healing Rings", in *INFORMS Meeting*. Montreal, Canada, April 1998.

[SOS94] Sosnosky J., “Service Application for SONET DCS Distributed Restoration,” IEEE Journal on Selected Areas in Communications, Vol. 12, pp. 59-68, January 1994

[SRI02] Srisuresh P. and Joseph P., "TE LSAs to extend OSPF for Traffic Engineering", Internet Draft draft-srisuresh-ospf-te-02.txt, work in progress, January 2002.

[STA00] D. Stamatelakis and W.D. Grover, "Theoretical Underpinnings for the Efficiency of Restorable Networks Using Preconfigured Cycles ("p-cycles")",

IEEE Transaction on Communications, Vol. 48, No. 8, pp. 1262-1265, August 2000.

[VAN96] Vanderstraeten H., et al., “Integration of Distributed Restoration Procedures in the Control Architecture of ATM Cross-Connects,”

[VEE01] Veeraraghavan M., et al., “Architecture and Protocols that Enable New Applications on Optical Networks,” IEEE Communications Magazine, March 2001, pp. 118-127

[WAS94] Wasen O.J., Wu T.-H., and Cardwell R.H., “Survivable SONET Networks – Design Methodology”, IEEE Journal of Lightwave Technology, Vol. 12, No. 1, January 1994, pp. 205-212

[WIN] Winter P., "Steiner Tree Problem in Networks: A Survey", *Networks*, pp. 129-167

[WIN92] Winter P. and Smith J.M., "Path-distance Heuristics for the Steiner Problem in Undirected Networks", *Algorithmica*, Vol. 7, No. 2-3, pp. 309-327, 1992

[WON99] Wong E.W.M, Chan A.K.M., Yum T.-S.P., “A Taxonomy of Rerouting in Circuit-switched Networks,” IEEE Communications Magazine, Vol. 37, No. 11, pp. 116-122, November 1999

[WU92] Wu T.-H., “Fiber Network Service Survivability”, Norwood, MA: Artech House, 1992

[WU97] Wu T-H. and Yoshikai N., “ATM transport and network integrity”, San Diego, London: Academic Press, 1997

[XIO97] Xiong Y. and Mason L.G., “Restoration Strategies and Spare Capacity Requirements in Self-healing ATM Networks,” INFOCOM’97, Vol. 1, April 1997

[XIO99] Xiong Y. and Mason L.G., “Restoration Strategies and Spare Capacity Requirements in Self-healing ATM Networks,” IEEE/ACM Transactions on Networking, Vol. 7, No. 1, pp. 98-110, February 1999

[XUE02] Xue G., Chen L., and Thulasiraman K., "QoS Issues in Redundant Trees for Protection in Vertex-Redundant or Edge-Redundant Graphs", ICC’02, 2002

[YAN88] Yang C.H. and Hasegawa S., “FITNESS: Failure Immunization Techonlogy for Network Service Survivability,” in Proceedings of GLOBECOM’88, Vol. 3, pp. 4731-4736

[YE00] Ye Y., Dixit S., and Ali M., "On joint protection/restoration in IP-centric DWDM based optical transport networks", *IEEE Communications Magazine*, vol. 38, Issue: 6, pp. 174-183, June 2000

[ZAN01] Zang H., et al, “Dynamic Lightpath Establishment in Wavelength-Routed WDM Networks,” IEEE Communications Magazine, pp. 100-108, September 2001