uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Hongxia Wang
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

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

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

P-Cycle in Multi-layer and Multi-failure Network Survivability

TITRE DE LA THÈSE / TITLE OF THESIS

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

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

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

C-H. Lung

A. Nayak

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

P-CYCLE IN MULTI-LAYER AND MULTI-FAILURE NETWORK SURVIVABILITY

By

Hongxia Wang

A thesis submitted to Graduate and Post-Doctoral Studies in partial fulfillment of the requirements for the degree of

Master of Applied Science

In

Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

School of Information Technology and Engineering

University of Ottawa

Ottawa, Ontario, Canada

Your file Votre référence
ISBN: 0-494-11445-2
Our file Notre référence
ISBN: 0-494-11445-2

Abstract:

Network survivability is a fascinating topic, in both research area and our daily life. P-cycles have gradually become a promising solution for mesh reliable networks, summed up as "ring like speed with mesh like efficiency".

In this thesis, we finely integrated p-cycle with multilayer technology and presented a new way for modern network: multilayer framework for p-cycle survivability. Based on this new framework, we also proposed a new mechanism which can survive multiple failures. This thesis aims to provide an efficient solution to adapt p-cycle to modern multilayer network architecture, and to survive multiple failures.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deep appreciation to my thesis supervisor, Prof. Hussein T. Mouftah, for his seasoned guidance, kindness and encouragement during my whole study period at the School of Information Technology and Engineering of the University of Ottawa, especially during the time of my pregnancy and nursery. It was him who gave me great support to help me to overcome the difficulties in my personal life and research work. It was him who introduced me to the intriguing kingdom of computer communication network.

Second, I can not be thankful enough to my dear family, my husband and my son, for their persistent love, affection and encouragement. Without their support, I would not be able to finish my program.

Next, I am deeply indebted to my parents who have taught me how to be a useful person to the society. They have tried their best to give me everything that I can never be able to pay it back.

Last but not least, it is my pleasure to acknowledge the group members in my lab for their valuable friendship, suggestions and inspiration during my study period.

ACRONYMS

AE	Apriori Efficiency
AP	Additional P-Cycle
APS	Automatic Protection Switching
ATM	Asynchronous Transfer Mode
BGP	Border Gateway Protocol
BLSR	Bidirectional Line Switched Rings
CoP	Class of Protection
CoS	Class of Service
DCPC	Distributed Cycle PreConfiguration Protocol
DWDM	Dense Wavelength Division Multiplexing
EON	European Optical Network
FDM	Frequency Division Multiplexing
GMPLS	Generalized MultiProtocol Label Switching
ILP	Integer Linear Programming
LCT	Link-Cycle Table
LSP	Label Switched Path
MFS	MultiFailure Survivability Scheme
MIP	Mixed Integer Programming

MPLS	MultiProtocol Label Switching
NP	Normal P-Cycle
NSF	National Science Foundation
OPCA	Off-Line P-Cycle Calculation Algorithm
OPPR	Optical Path Protected Ring
OSI	Open System Interconnection
OSPF	Open Shortest Path First
OSPR	Optical Shared Protection Ring
P-Cycle	Pre-Configured Protection Cycle
PMSM	P-Cycle Multilayer Survivability Mechanism
PP	Protection Pool
PPT	P-cycle Pool Table
QoS	Quality of Service
SBPP	Shared Backup Path Protection
SHR	Self-Healing Ring
SLA	Straddle Link Algorithm
SONET	Synchronous Optical Network
SRLG	Shared Risk Link Group
TDM	Time Division Multiplexing
TS	Topology Score
VC	Virtual Circuit
VP	Virtual Path
VWP	Virtual Wavelength Path

WDM	Wavelength Division Multiplexing
WP	Wavelength Path

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

λ	Wavelength
T($L_{j,i}$)	An SRLG resource topple T($L_{j,i}$), where j is link ID, showing that which layer this link is in, i is the identity of this link (link ID). For example, T($L_{3,2}$) means the 2nd link in layer3 (TDM link).
b	Bandwidth
$L_{j,i}$	The i_{th} link at layer j
$P_{j,k}$	The k_{th} P-Cycle at layer j
D_j	Layer j

Chapter 1 INTRODUCTION

1.1 Background

When our networks started to carry more traffic (gigabits per second range), the requirement of robustness of mature networks became critically important. Since the first appearance of optical technology, a single fiber began to carry significant amount of messages, especially after the advent of Wavelength Division Multiplexing (WDM) and Dense Wavelength Division Multiplexing (DWDM) technology. Any single failure can cause severe influence on thousands of customers. Network survivability has become an extremely crucial topic. In the new century of information era, nobody can live without talking on the phone, surfing through the internet, banking on line or taking distance courses. Different kinds of network services have already become a part of our life, including traditional data services and new real time multimedia services. A single failure in such networks can disrupt services for a large number of customers at once, causing a significant loss of revenues for service provides.

Extensive research work has been done in the area of network survivability. For almost 10 years, the work has been focused on two separate approaches: ring-based

and mesh-based network protection. The ring-based protection was first introduced in Synchronous Optical Network (SONET) and got well known for its fast reaction (around 50ms). However, its need for at least 100% spare capacity made investors hesitate. On the contrary, the mesh-based protection overcomes ring's low efficiency backside, realizing 50-70% redundancy, but in the meantime gives up ring's significant advantage, which is the fast restoration time. Mesh networks usually need more than 1 second to react [GROV00].

After a decade of this embarrassing dichotomy, in 1998 a totally new concept of protection, p-cycle, was proposed, which offered ring speed restoration in mesh networks, known as "ring-speed and mesh-efficiency". Since then, it has led to a revolution in the survivability research field.

After several years of hard work, the p-cycle concept has been studied in WDM and IP/MPLS networks. In some special cases, it can survive multiple failures. However, it is still a relatively new concept; tremendous research work is needed before it can be applied by the industry.

In this thesis, we introduced a framework to implement p-cycles in the Generalized MultiProtocol Label Switching (GMPLS) multi-layer environment, and also provide a mechanism for p-cycles to recover multiple failures. Our work has made p-cycles steps closer to real networks.

1.2 Motivation and Objective

The p-cycle has been assigned as "ring-speed and mesh-efficiency" which makes it an attractive solution for survivable networking. Up to now, it has been proposed to be suitable for WDM networks and IP networks; also it has the potential to be extended to GMPLS networks. However, no real research result has come out yet. As networks have begun to integrate different layers and different types of services together, the possibility of multiple failures becomes much higher. Because of the restriction of current algorithms to find proper p-cycles, the ability of p-cycle to survive multiple failures is very limited. It depends on the position where the second failure will occur.

This thesis aims to provide a complete simple solution to adapt p-cycles to modern GMPLS networks, and meanwhile to achieve the ability to survive multiple failures. Our objectives are as follows.

- To address the coordination problem in regular multi-layer networks.
- To introduce a new multi-layer framework to adapt p-cycles to multi-layer networking on the base of Shared Risk Link Group (SRLG) Tree structure.
- To provide a new scheme to solve failure information propagation and bandwidth information propagation problems in regular multi-layer networks.
- To propose a new mechanism and algorithm for p-cycles to survive multiple failures.

1.3 Thesis Contributions

This thesis has been committed to provide an efficient solution to integrate p-cycle with GMPLS and equip p-cycles with the ability to restore multiple failures. The contributions of this thesis are summarized as follows.

- The use of SRLG Tree to solve failure propagation and bandwidth propagation problems in regular multi-layer networks has been illustrated.
- Changes of some parameters from the original SRLG Tree have been made, in order to fit it to the p-cycle design.
- A new Link-Cycle Table has been proposed to let networks easily choose the appropriate p-cycle to recover failures.
- A novel mechanism has been provided for multi-failure recovery using p-cycle. This mechanism has two procedures, which are the off-line centralized calculation and the on-line distributed selection.
- In the off-line centralized calculation stage, an Off-line P-cycle Calculation Algorithm (OPCA algorithm) has been given to build up P-cycle Pool Tables (PPT) which will be used by the on-line distributed selection.
- A novel Protection Pool has been introduced to achieve multi-level protection and to provide Class of Protection (CoP).
- A versatile simulation program has been developed in Java language which verified our proposal.

1.4 Thesis Outline

This thesis is organized into six chapters. Chapter 1 gives an introduction of network survivability. The motivation, objectives and contributions of this thesis are also presented. Chapter 2 provides a comprehensive survey on the state-of-the-art of the network evolution, and different network protection/restoration methods. In Chapter 3, an SRLG Tree structure is presented and revised to fit the p-cycle environment. A new multi-layer framework is also addressed to adapt p-cycles into GMPLS networks by using the SRLG Tree. A two-stage scheme is proposed in Chapter 4 that extends the ability of p-cycles to survive multiple failures. In this chapter, a new algorithm OPCA is presented and the multi-level survivability issue is also addressed. Chapter 5 describes the simulation model. Performance analysis is also given based on simulation results. In the last chapter, we conclude our current research work and our future research target.

Chapter 2 STATE-OF-THE-ART OF NETWORK SURVIVABILITY

2.1 Introduction

The survivability of a network refers to a network's capability to provide continuous service in the presence of failures [ZHOU00]. After the optical fiber has become the dominant transport medium in communication networks because of its advantages in capacity, reliability, cost and scalability, different technologies have been proposed and used for survivable optical networks.

Before designing a survivable optical network, one must lay out the possible failures under which a network must survive. There are two familiar types of network failures: link failure and node failure. Link failure usually occurs after fiber cuts, while node failure is mostly due to equipment failure at network nodes. Man-made errors and uncontrollable natural phenomena, such as floods and earthquakes, can cause equipment and node failures as well. In an optical network, nodes are usually assumed to be robust, therefore in this thesis, we focus only on span/link failure.

Pre-designed protection [RAMA99a] and dynamic restoration [RAMA99b] are two general approaches to the design of survivable optical networks. Pre-designed protection refers to the fact that the recovery from network failures is based on preplanned schemes. Usually it relies on resources (fibers, wavelengths, switches, etc.) dedicated for protection purposes. In pre-designed protection, some resources are reserved for recovery from failures at either the connection setup or the network design time, and kept idle when there is no failure. From this aspect, the capacity is not efficiently used, but on the other hand, the level and speed of recovery can be guaranteed. Dynamic restoration implies that the spare capacity in a network is dynamically discovered and used to restore the affected services. It is typically more efficient than pre-designed protection in terms of the resource utilization. The drawback is, its restoration time is usually longer, and the 100 percent service recovery cannot be guaranteed if sufficient spare capacity is not available at the time of failure.

Most of today's communication networks use pre-designed protection mechanisms rather than dynamic restoration methods, because pre-designed protection can provide rapid and guaranteed recovery. Dynamic restoration still needs further improvement due to its long restoration time and uncertainty of recovery.

2.2 Network Survivability Before P-Cycle

Extensive research work has been done in the field of network survivability before the discovery of p-cycle technology. In [ZHOU00] [CHAL03] [RAMA99a] [RAMA99b], several basic protection techniques, such as Automatic Protection

Switching (APS), Dual Homing, Self-healing Ring (SHR), and Mesh Protection, are summarized by their characteristics.

Dual Homing is a relatively old technique and is not widely used in current networks. In Dual Homing, fibers are provided so that all data are sent simultaneously over two separate routes to and from the central network hub [ZHOU00]. This centralized structure makes the network vulnerable.

APS and SHR are the most common protection schemes used in non-WDM optical networks. APS is typically used to deal with link failures. It has three main architectures which have been used by mesh protection later: 1+1, 1:1 and 1:N. In general, m:n protection refers to a scheme in which m protection links/paths are shared among n working links/paths. SHR is more flexible than APS in that it can handle both link and node failures. Unidirectional SHR (USHR) and bidirectional SHR (BSHR) are two types of SHRs in the SONET system. SHR is a very successful technique, which has been widely used for many years, for survivable optical networks.

The significant success of SONET makes the ring-based survivable networking dominating the SONET era. DWDM optical rings have also been deployed. However, ring networks are considered as capacity-inefficient because of their dedicated resource protection structure. This is the main reason for the ring-to-mesh evolution.

The mesh network topology is widely employed in current telecommunication services due to its high efficiency and flexibility. Same as ring networks, mesh survivable networks use both pre-designed protection and dynamic restoration. From the network protection terminology, the survivable mesh approach can be divided into two basic paradigms: path protection/restoration and link protection/restoration. If we categorize it by the network resource sharing situation, it can also be classified as dedicated protection and shared protection. These different categories of survivable mesh networking techniques are shown in Figure 2-1.

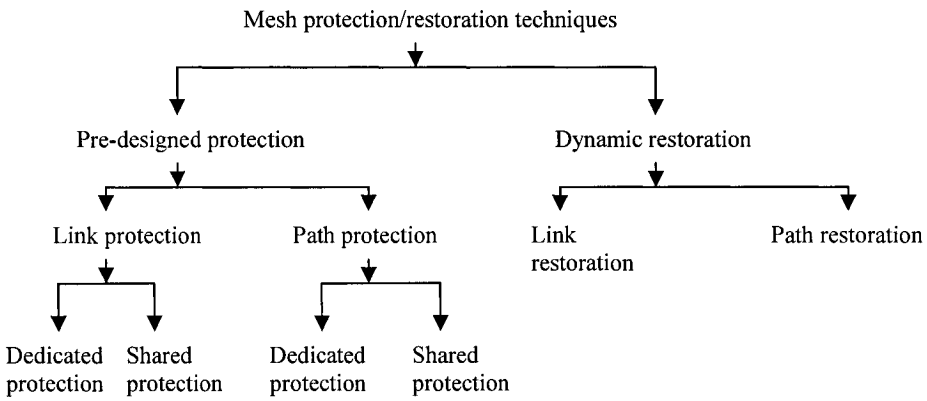


Figure 2-1 Mesh protection/restoration technique

In summary, the pre-designed shared link protection scheme is widely used at logical layers. At physical layers, in order to guarantee fast recovery, dedicated techniques are used. In this thesis, we will focus on the logical layer survivability.

2.3 Ring and Mesh Solitude

Currently, ring and mesh networks are two popularly used network topologies. SONET is the most typical and famous example for ring networks, while WDM networks mostly use the mesh topology.

The debate on which one is better in terms of survivability, ring or mesh, has been going on for nearly a decade. Each of them has some significant advantages and some serious drawbacks as well. These features are summarized in [GROV00] (see Table 2-1). Conventional Ring-based* survivability involves the use of BLSR (bidirectional line switched rings) and UPSR (unidirectional path-switched rings), or their optical versions OPPR and OSPR (optical path protected ring and optical shared protection ring). The important point is that rings use a simple switching mechanism which can achieve fast recovery in about 50-60 ms. However, its shortage of at least 100% redundancy, which means weak capacity efficiency, prohibits it from dominating the future network. Also the ring structure itself leads to inflexibility, and makes multi-ring network planning extremely complicated. These drawbacks determined that the trend of today's network is to migrate from ring to mesh topology.

Compared with rings, the mesh topology is more capacity-efficient. Inside a mesh network, each unit of the spare capacity is reusable in more ways. It permits a major

---

* Note that "ring" generically also includes 1+1, 1:1 APS architectures

reduction in the capacity that is required to serve the same set of demands. However, because of the general nature of solving a discrete capacitated multiple-path re-routing problem, the mesh restoration is not expected to be as fast as rings.

	Advantage	Disadvantage
RING	50msec restoration times	Planning of multi-ring network is too complex
		Hard to accommodate multiple service classes
		Inefficient and inflexible
		Ring-constraint routing
		Need at least 100% redundancy
MESH	Need only 50-70% redundancy	Up to 1.5sec restoration time
	Simple, exact capacity planning solutions	
	Easy and efficient design for multiple service classes	
	Efficient and flexible	
	Shortest-path routing	

Table 2-1 Comparison of ring and mesh networks

Table 2-1 shows a comparison of ring and mesh networks. From this table, one can easily get the desirable features for desirable networks. Exchange the drawback of mesh networks with the good point of ring networks, which means change mesh networks' slow restoration time (up to 1.5sec), to fast restoration time (around 50msec), and keep the rest of mesh advantages (low redundancy, simple capacity solutions, easy and efficient design for multiple service classes, efficient and flexible, and shortest path routing).

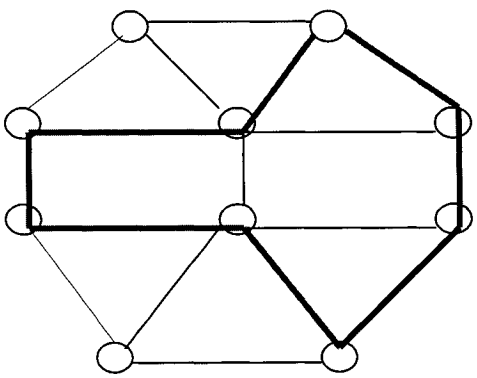
There has been effort to hybrid these two different technologies together in a ring-access/mesh-core division [GROV92] [BROW94], but in practice, ring and mesh components of such 'hybrids' continue to be designed and operated separately, along their respective principles. It has been a hard task for all network researchers to find a mechanism that can really benefit both ring and mesh techniques.

2.4 P-Cycles

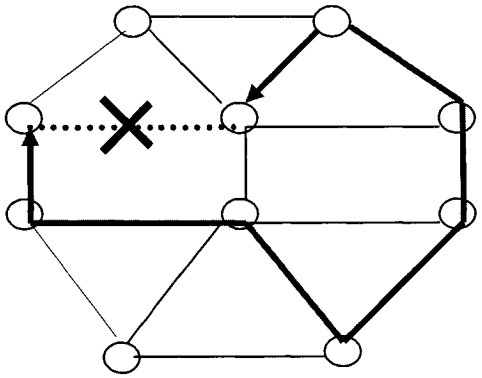
In 1998, a totally new strategy was introduced [GROV98]. This new concept, called "P-Cycle", is a fully pre-connected structure, which could merge these two different technologies (ring and mesh) together very well. It uses the same switching mechanism as a ring, but gives up virtually nothing in terms of required extra spare capacity in mesh networks. It combines the real-time switching simplicity and speed of rings, with the mesh-like efficiency, flexibility and freedom of a mesh in the routing of working paths. Since then p-cycles became well known as "ring-speed and mesh-efficiency".

P-cycles were defined as "**P**re-configured **P**rotection **Cycles**" which are based on the formation of pre-configured cycles and formed out from the previously unconnected spare links of a mesh-restorable network [Grov98]. The unique use of "straddling link" makes p-cycles different from regular rings. Two important design methods were proposed in [GROV98], which are the Optimal Design of P-cycle Networks using LLP technology and the Distributed Cycle PreConfiguration (DCPC) protocol. The DCPC protocol is an adaptation of statelet processing rules of Self-healing Network protocols [GROV90][GROV91].

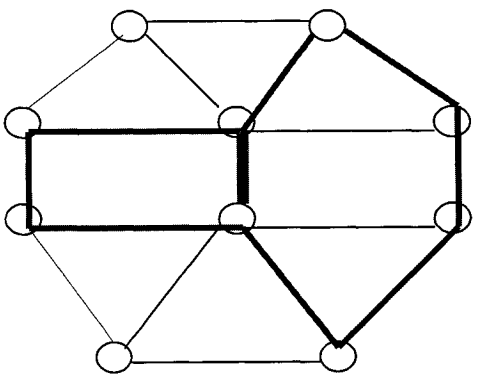
Figure2-2 illustrates the concept of p-cycle and its functions. Figure 2-2(a) is an example of a simple p-cycle in the network. In Figure 2-2(b), an on-cycle span breaks and the surviving arc of the cycle is used for restoration. Figure 2-2(c) shows an example of straddling span. When straddling span fails, two restoration paths are available to provide the restoration (as shown in Figure 2-2(d)).



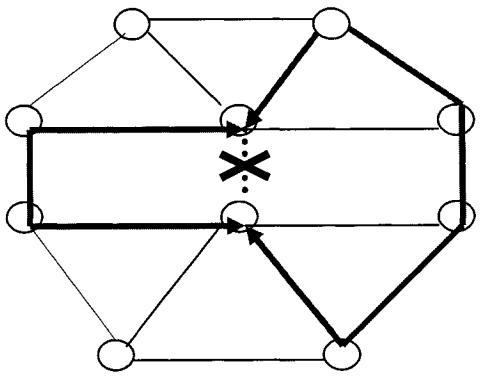
a) A simple p-cycle



b) In case of on-cycle span failure



b) Straddling span: an off cycle span having p-cycle nodes as endpoints



c) In case of straddling span failure

Figure 2-2 Use of p-cycles in span protection

An investigation of the deployment of p-cycles in WDM mesh networks with and without wavelength conversion was made in [SCHU02]. In virtual wavelength path (VWP) WDM networks, all nodes perform full wavelength conversion, while in wavelength path (WP) WDM networks nodes do not have any wavelength converter at all. Optimization models for both networks were provided and analyzed. VWP and WP networks are two extreme network conditions. It is not possible for all nodes to have full wavelength conversion capacity in a network (VWP case), because this will make the network too expensive. In WP networks, when no node has any wavelength converter, it becomes very hard to find paths to deliver the demanding traffic because of the wavelength continuity constraint.

To solve this problem, a new idea of "p-cycles in WDM networks with partial wavelength conversion" was brought forward [SCHU03c]. It focused on the relation between the cost associated with the number of required wavelength converters and the protection capacity-efficiency achieved. Mathematical models were formulated and the respective optimization problems for a pan-European network were solved. It was found out that the total number of converters required for a network can be greatly reduced, with only a small increase in the spare capacity by a strategy of associating wavelength converters with access points between a pure WP working layer and a set of pure WP p-cycle protection structures.

Although p-cycles offer a promising new approach to optical network survivability, the complexity of solving optimal p-cycle design problems is very high, because

each design needs to find out all the possible cycles in the whole network. The candidates can be thousands. New methods are needed to overcome this difficulty. Pre-selection and joint optimization issues were addressed in [GROV02b]. These two methods finely reduced the complexity of the p-cycle design problem and increased its capacity efficiency.

For the pre-selection, two metrics, topology score (TS) and apriori efficiency (AE), were used to rank all distinct cycles found from a certain network. TS and AE are based on insights about what makes for the most efficient p-cycles in the context of a give network design. The TS and AE of a cycle j is defined in the inset where S is the set of spans, x_{ij}=1 if span i is part of cycle j, x_{ij}=2 if span i straddles cycle j and x_{ij}=0 otherwise. C_i is the cost or distance of span i.

$$TS(j) = \sum_{i \in S} x_{ij} \qquad AE(j) = TS(j) / \sum_{(i \in S | x_{ij}=1)} C_i$$

After pre-selection, only a limited number of top-ranked cycles are represented in the optimal solution model as p-cycle candidates. Pre-selection significantly reduced the complexity of the p-cycle design.

Another method to lower the complexity is to optimize the choice of working routes in conjunction with the placement of the spare capacity. A 25% redundancy reduction can be gained by this joint design, when compared with optimal non-joint designs [GROV02b]. An upper bound has also been pointed out which can be used to cut the design complexity [STAM00].

Some new ways were also discovered to reduce the p-cycle chosen complexity [ZHAN02a][DOUC03a][MARD04]. A new algorithm called Straddling Link Algorithm (SLA) was proposed in [ZHAN02a] where Dijkstra's shortest path algorithm was used to find exactly one p-cycle for each link in the network. It guaranteed that only L (number of links) cycles will be counted and they are node-disjoint from each other. SLA dramatically reduced the complexity to O (NlogN) where N is the number of network nodes, but the price is that it loses the optimal efficiency because of the limited number of p-cycle candidates. Based on SLA, several algorithms were proposed in [DOUC03a]. These algorithms are SP-Add, SP-Expand and SP-Grow. They extend SLA cycles step by step until there is no further bigger cycles can be found, when using the same p-cycle merging technique introduced by [MARD04]. Finally a near-optimal solution can be found with no more than 10% difference from the Pure ILP and much less complexity.

The multi-failure survivability issue has also been discussed. [SCHU03a] and [SCHU03b] provided some conditions when p-cycle can manage dual failures.

The concept of p-cycle is originally conceived to be used in WDM and SONET transport networks. It also has been extended to the IP layer [STAM00]. P-cycles use virtual protection cycles for extremely fast restoration in IP networks. These virtual circuit-like p-cycles use MPLS Label Switched Paths (LSPs) or any other tunnelling and tag switching proposals. Because p-cycles are virtual circuits, they consume zero

capacity until the failure occurs and they are being used. Router failures can also be recovered by "Node-Encircling" p-cycles. An IP network design model using Mixed Integer Programming (MIP) formulation let "IP layer p-cycle recovery" become more feasible. In general, because IP networks are already restorable by routing protocols through dissemination of link-state and route advertisements, p-cycles provide an immediate real-time detour to prevent packet losses until the conventional global routing reconvergence occurs. Thus the p-cycle mechanism serves as a fast-acting but temporary protective measure, which secures network traffic while routing tables adapt globally to new network states [GROV00]. IP p-cycles are envisaged as the "fast" part of a "fast plus slow" overall recovery process.

The concept of "Path-Segment P-cycles" has also been introduced [SHEN03]. It effectively extends the p-cycle technique to cover not only on-cycle links or straddling links, but also paths or flow segments along a path, as well as working flows that transit a failed node. Another contribution of [SHEN03] is, it implied the potential of p-cycles to be used in a layered network structure. Soon after, the issue of bandwidth protection in MPLS networks were addressed [KANG03].

2.5 Summary

A variety of protection and restoration methods exists for communication networks, such as automatic protection switching (APS), ring protection, shared backup path protection (SBPP), mesh span/path restoration and p-cycles. These methods vary in two key parameters: the capacity efficiency and the restoration speed. They can also

be used in the same network to provide different qualities of protection service to different connections [BLOU03].

Among all these methods, the p-cycle has grasped more attention, because of its fascinating characteristics as "ring-speed and mesh-efficiency". Although all researchers are optimistic about its future, it is still relatively new concept and has not been employed by the industry yet.

In the following chapter, we will dedicate our effort on adapting p-cycles to real modern networks, which brings p-cycles at least one step closer to the industry.

Chapter 3 ADAPTING P-CYCLES TO GMPLS MULTI-LAYER ARCHITECTURE

3.1 Introduction

Modern telecommunication networks consist of different layers, such as IP-over-ATM, IP-over-MPLS, packet-over-optical networks, etc. No single layer can handle all different kinds of traffic, especially after the appearance of e-business, distance video-conferencing and many other real-time multimedia applications. The multi-layer network architecture offers a powerful solution to the increasing traffic pattern difference. It also invokes another important issue: how to manage all these different layers that have their own protocols and characteristics? The development of GMPLS provides a powerful instrument to seamlessly integrate different layers. With a common control panel, different layers start to work together smoothly under the GMPLS architecture.

In the previous chapters we introduced the p-cycle concept and summarized its advantages for the network protection. P-cycles have been studied extensively in the

literature. Now we will propose a new framework to adapt them into the GMPLS multi-layer architecture.

3.2 Motivation and Objective

In recent years, lots of research work has been carried out on the use of p-cycle protection and restoration in IP and WDM networks. However, all these attempts were carried out relatively individually. Nobody considered them intrinsically together, nor with any other kinds of networks. Actually p-cycles have a good potential to be used in different kinds of networks to fit the multi-layer environment. The reasons are as follows:

1) P-cycles apply in principle to any connection-oriented transport layer, including ATM using VP or VC constructs [GROV00]. This means p-cycles can be used in TDM, WDM, MPLS, FDM and many other networks. With the help of GMPLS, even previous connection-less IP packets can be encapsulated inside connection-oriented GMPLS packets so as to make the whole IP network transparent.
2) P-cycles can be set up with different sorts of technologies, such as LSP, virtual circuit, ligthpath, IP link, etc. In the GMPLS environment, this means "p-cycles can be set up with different hierarchical LSPs, such as Fiber LSP, λ LSP, TDM LSP and Packet LSP [GROV00]".
3) The p-cycle packet format matches hierarchical architectures inherently. Figure 3-1 shows the p-cycle packet format defined by [STAM00a]. One may observe that p-cycles can be extended to a hierarchical environment by encapsulating any incoming packet with a p-cycle header (which consists of

P-cycle ID, Destination Address, and Original Route Cost), and leaving the original packet as a transparent payload. For example, in an MPLS network, the p-cycle packet will look like in Figure 3-2.

p-cycle ID	Destination address	Original Route cost	**IP Packet**

Figure 3-1 p-cycle packet format at IP layer [STAM00a]

p-cycle ID	Destination address	Original Route cost	**MPLS Packet**	
			MPLS header	**IP Packet**

Figure 3-2 p-cycle packet format at MPLS layer

Subdividing a network into several layers and providing different layer protection using have several significant advantages:

1) ***Bandwidth saving***. When a single wavelength fails, we do not need to provide restoration for the whole span. Similarly when a single LSP fails, it is not wise to restore the whole λ, in order not to waste the bandwidth. However, networks do not have any other choice, but to restore the entire span or the total wavelength, if they are not in a multi-layer structure. In order to protect and consume the exact resource needed, not to waste bandwidth unnecessarily, we have to finely define the network and get the topology of each layer.

2) ***Much better granularity***. After a network has been subdivided into different layers, it can provide different functions at different layers and supply differentiated survivability.

3) ***Simplification of each layer's function***. Like the OSI 7 layer architecture, when an entity is subdivided into several smaller objects, each object becomes simpler. In multi-layer networks, each layer is simpler when compared with unlayered networks. Multi-layer networks are also more flexible and scalable.

4) ***Enhancement of network's power***. Networks become more powerful when their different layers work together, the network becomes more powerful. For example, transport networks alone can not protect against upper layer faults, such as router failures and network card failures, without upper layers' help. Sometimes the lower layer protection is required to offer faster and more cost-efficient coverage, which can not be achieved by upper layers.

3.3 SRLG Multi-Layer Architecture

There is an important problem of diversity integrity in all kinds of multilayered networks. "Related to the creation of physical layer, diversity is the need to be able to validate the details of physical structures that underlie logical protection or restoration routes to ensure integrity of the mapping from physical to logical diversity [GROV04]". This diversity problem also has to be solved before the p-cycle protection is employed.

In order to use p-cycles for multi-layer survivability, the network has to get the correct topology and resource information for each layer first, then deploys the p-cycle design at different layers and executes different layer p-cycle protection.

A new multi-layer differentiated protection architecture was proposed in [NASE04] that uses SRLG trees to aggregate link-state and bandwidth information for the communication between adjacent layers. Figure 3-3 illustrates all pertinent layers in today's transport networks.

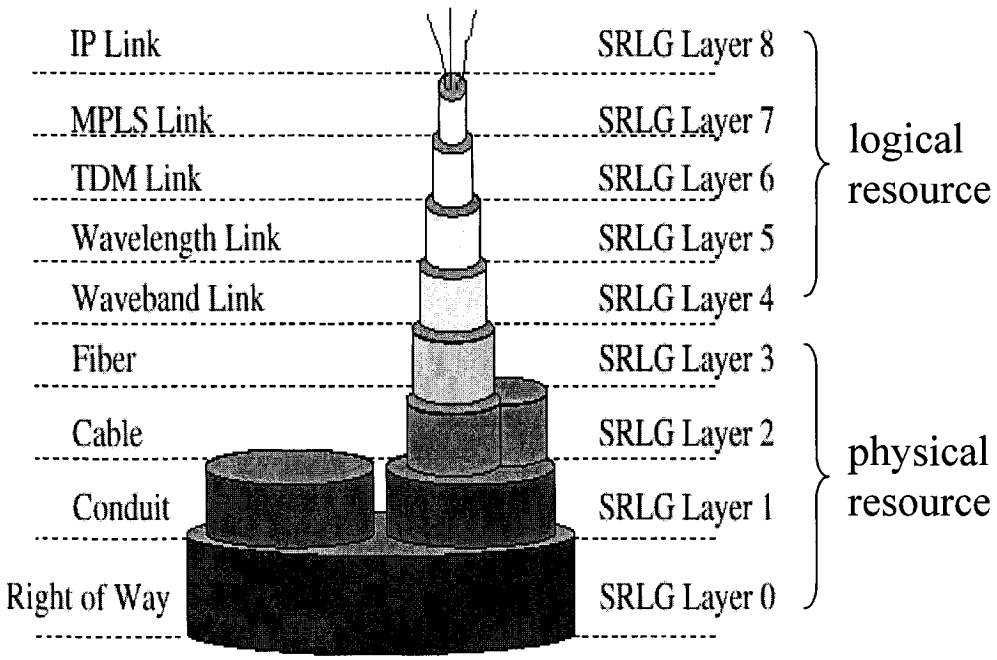


Figure 3-3 Multi-layer transport network architecture [NASE04]

The concept of SRLG (Share Risk Link Group) is used to present different layer links in GMPLS networks. As illustrated in Figure 3-3, networks are divided into 9 layers (IP layer, MPLS layer, TDM layer, Wavelength layer, Waveband layer, Fiber

layer, Cable layer, Conduit layer, Right of Way layer), with upper layer links tunnelled inside the adjacent lower layer links forming an SRLG group. A hierarchical SRLG tree is established to provide the failure propagation and the bandwidth propagation among different layers. Previous multi-layer networks suffer from two problems, resource usage redundancy and lack of guarantee for paths to be diverse at lower layers while they are diverse at an upper layer, due to the lack of coordination between different layers. The SRLG tree model successfully solved these two important issues by presenting the aggregate information between adjacent layers.

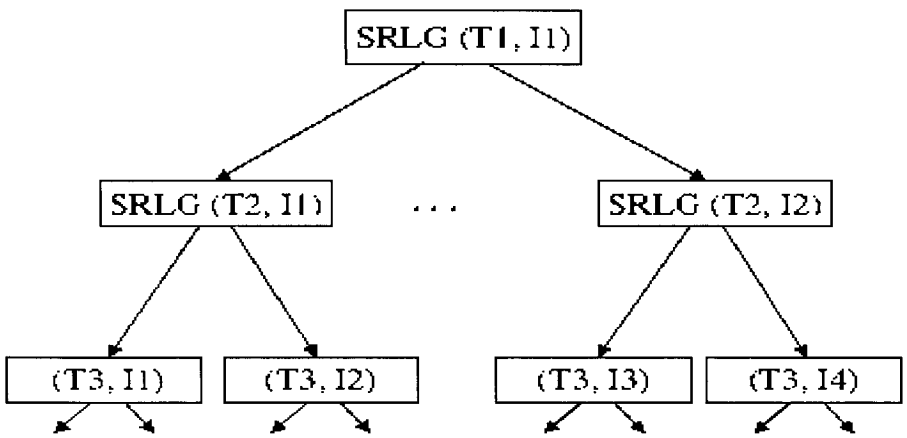


Figure 3-4 Topology database of SRLG Tree [NASE04]

Figure 3-4 represents the topology database of SRLG Trees. Through a top-down arrangement of SRLG resources in the multi-layer structure, each link at layer T_1 is associated with an SRLG resource topple (T_1, I_1), where I is the identity of the link. This resource can inherit a failure from one or more resources at layer T_2. In an

SRLG tree, the SRLG resource at layer T_1 is a parent (ancestor) of the SRLG resources at layer T_2 that inherits the failure. Similarly, every SRLG resource at layer T_2 is a parent of one or more SRLG resources at layer T_3, and so on, until all layers are reached. The SRLG resource at layer T_1 becomes the head of a tree whose leaves are SRLG resources at layers below T_1. The above procedure essentially creates a hanging SRLG tree from every link at layer T_1 [NASE04].

3.4 Adapting SRLG Tree to P-cycle Protection

In Figure 3-3, the four lower layers represent physical resources while the remaining upper layers represent logical resources [NASE04]. In a network, physical layers usually have their own protection/restoration requirement and procedure with more reliable and faster recovery scheme compared with logical layers. P-cycles are more suitable for logical resources. In this thesis, we consider only logical resource layers (the upper 5 layers) for the p-cycle protection, as shown in Figure 3-5. A revised SRLG Tree model is presented in Figure 3-6, using a bottom-up arrangement of SRLG resources, instead of the top-down sequence.

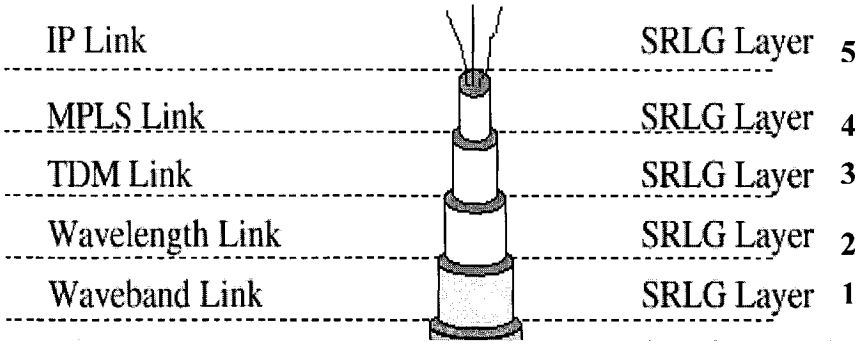


Figure 3-5 Simplified multi-layer network architecture

Different layers of the revised SRLG tree (see Figure 3-6) are associated with the revised multi-layer architecture (see Figure3-5). For example, the 1st layer of the SRLG tree is SRLG layer1 (Waveband link) in Figure3-5, the 2nd layer is SRLG Layer2 (Wavelength link), and etc. The root of this tree is Layer1 (Waveband) links which are tunnelled by one or more Layer2 (Wavelength) links. These layer2 wavelength links are children of that layer1 waveband link. Similarly, several Layer3 (TDM) links that are multiplexed inside one of those layer2 wavelength links become its children. Multiple MPLS links are the next generation of one of those TDM links. This information collection and calculation procedure continues until all layers get the necessary information. All the topology and resource information is stored inside a common database, which is organized in a hierarchical manner and shared by all layers.

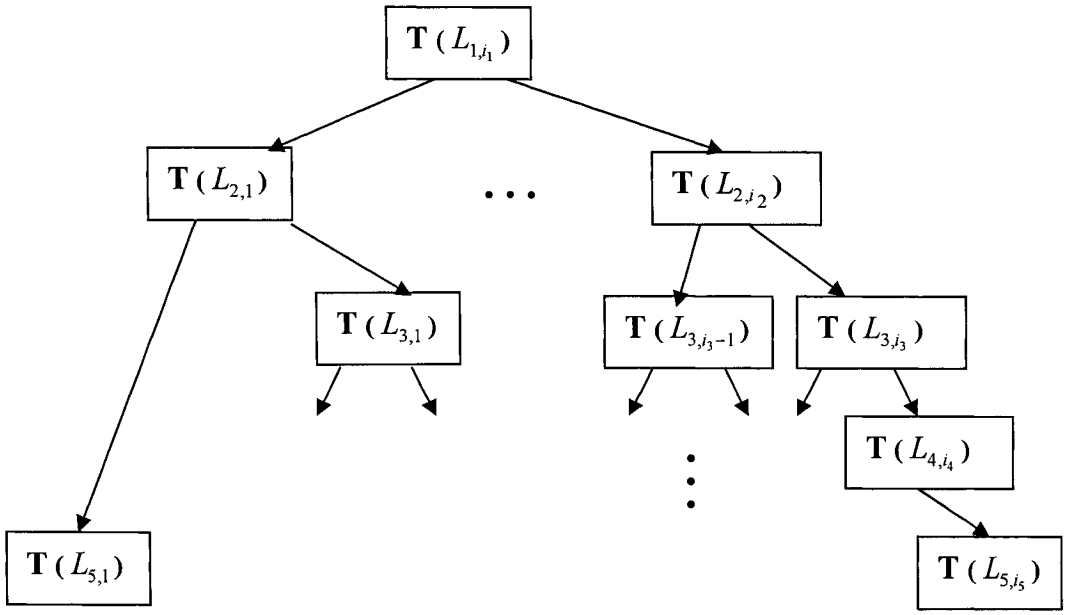


Figure 3-6 Revised Hierarchical SRLG tree

In the revised hierarchical SRLG tree, each leaf presents a certain link at a certain layer. A link is presented by an SRLG link topple T ($L_{j,i}$), where j is the layer ID, and i is the identity of this link (link ID). For example, T ($L_{3,2}$) represents the 2nd link at layer3 (TDM link).

There can be several special cases:

1) Some layers may be absent in a typical deployment of transport networks. For instance, inside a network, all MPLS links might be directly tunnelled inside several wavelength links, completely bypassing the TDM layer [NASE04]. The corresponding SRLG layer will be absent in the SRLG tree model.

2) A link at any given layer can be a direct ancestor of different generations. In Figure 3-6, T ($L_{2,1}$) is an ancestor of both link T ($L_{5,1}$) and link T ($L_{3,1}$).

In the revised SRLG tree, when a parent link fails, its entire children links will also fail. The failure information will propagate all the way to its descendants using top-down arrangement. This failure propagation relationship is elaborated in Figure 3-7. When the Layer2 link T (L_{2,i_2}) fails, the failure information is propagated to its children T (L_{3,i_3-1}) and T (L_{3,i_3}) through the direction of dotted arrows, furthermore to its grandchild T (L_{4,i_4}) and great-grandchild T (L_{5,i_5}). All these descendant links will fail after their ancestor T (L_{2,i_2}) fails.

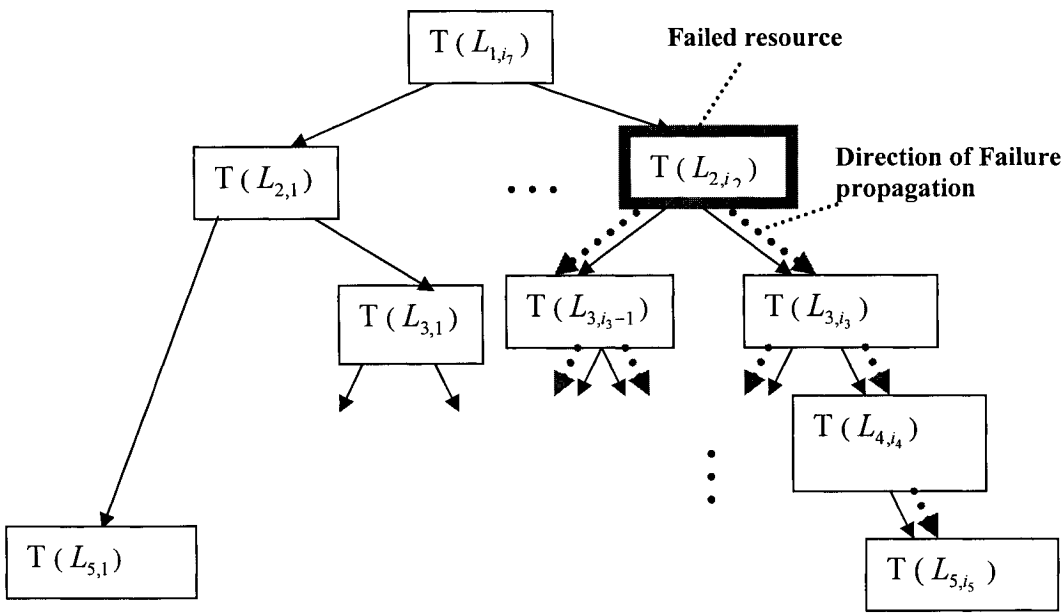


Figure 3-7 Failure propagation process in the revised SRLG tree

In the case of bandwidth allocation, the propagation process is managed in a bottom-up manner as illustrated in Figure 3-8. Suppose a new path setup demand with bandwidth *b* arrives at layer4 (MPLS layer). If this requirement is accepted and the new path passes through link T (L_{4,i_4}), bandwidth *b* will be reserved at T (L_{4,i_4}). In order to reserve bandwidth *b* at link T (L_{4,i_4}), all its lower layer ancestors have to reserve at least the same amount of bandwidth. If the lower layers do not have enough bandwidth, this call request can not be set up.

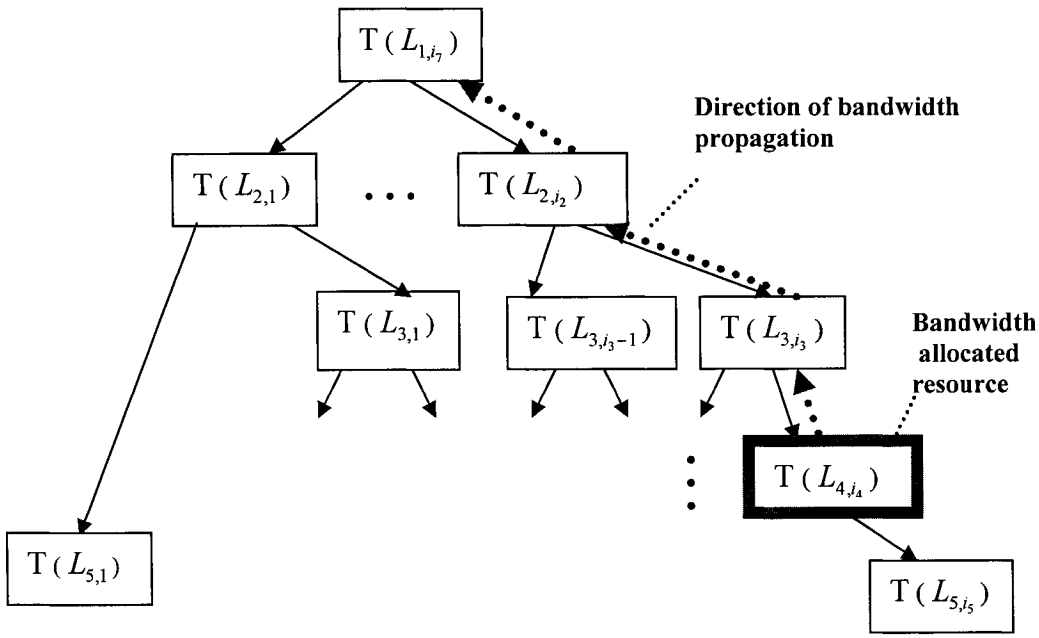


Figure 3-8 Bandwidth propagation in the revised SRLG tree

So T (L_{3,i_3}) in layer3 (TDM layer), T (L_{2,i_2}) in layer2 (wavelength layer) and

T (L_{1,j_1}) in layer1 (waveband layer) (if exist) will also claim at least bandwidth b. Accordingly, bandwidth b will be added to those links' working capacity and subtracted from their spare capacity records.

Through this revised SRLG tree model, each node is capable to get the most up-to-date network topology and resource information, including the residual capacity, the working capacity, and the spare capacity.

The revised SRLG Tree model has several advantages compared with the original SRLG model [NASE04], especially when used in the p-cycle design. The original SRLG model deployed a new architecture with better granularity and flexibility when compared with any unlayered structure, but the price it has paid is much higher complexity. In the revised model, some improvement has been done to simplify the model and reduce its complexity.

1) Changing the top-down structure to a bottom-up one makes the procedure to set up an SRLG tree easier. The root represents a waveband link. Wavelength links multiplexed in this waveband link become its first degree offspring. Several TDM links that are tunneled in those wavelength links turn into their coordinate descendants, the second degree offspring of that waveband link. Step by step using the same procedure, the whole tree will be set up.
2) When using the bottom-up model, nodes inside the tree will not have the Multiple Inheritance [NASE04] relationship in the network's topology database, which can simplify the structure and reduce the complexity.

3) The revised model only considers logical resource layers (IP, MPLS, TDM, Wavelength, Waveband), which further reduces the complexity and increases the scalability. This change also makes the model more suitable for the p-cycle protection. In a real transport network, physical layers usually need faster, more robust and more reliable protection/restoration methods, such as 1+1 dedicated or 1:1 shared mesh protection/restoration. P-cycles are more adequate to logical resource layers. When networks are in the stage of gathering network information, the propagation technology can still be applied to all network layers, in order to get benefits from SRLG model's convenience and accuracy.
4) Nodes inside the revised SRLG tree are defined as "links", not "resources", to simplify the problem. The bandwidth propagation mainly goes through links. In the case of node failures, usually one node failure can be converted to one or more link failures in the failure propagation. Links are the key resource in an information propagation procedure.

In the rest of this thesis, we use "SRLG Tree" to represent "Revised SRLG Tree" for simplicity.

3.5 P-cycle Multi-Layer Survivability Mechanism

Now we will introduce our novel P-cycle Multi-layer Survivability Mechanism (PMSM). After setting up the SRLG tree and gathering all the necessary information, the p-cycle design is executed at different layers using the centralized network management system [GROV00] [SCHU02a] [STAM00a] [SHEN03] [GROV02a]

[GROV02b] [GOUL03b] [ZHAN02a] which matches the GMPLS common control panel structure very well. SRLG trees secure that the system can get the appropriate up-to-date information of the link state, the working capacity and the spare capacity. Several algorithms have been deployed to find the optimal p-cycles [GROV98] [ZHAN02] [DOUC03a]. Our P-cycle Multi-layer Survivability Mechanism (PMSM) can use any of them. In other word, PMSM is algorithm-independent. After all layers have been reached and computed (See Figure 3-9), the final p-cycle information will be stored in the central database.

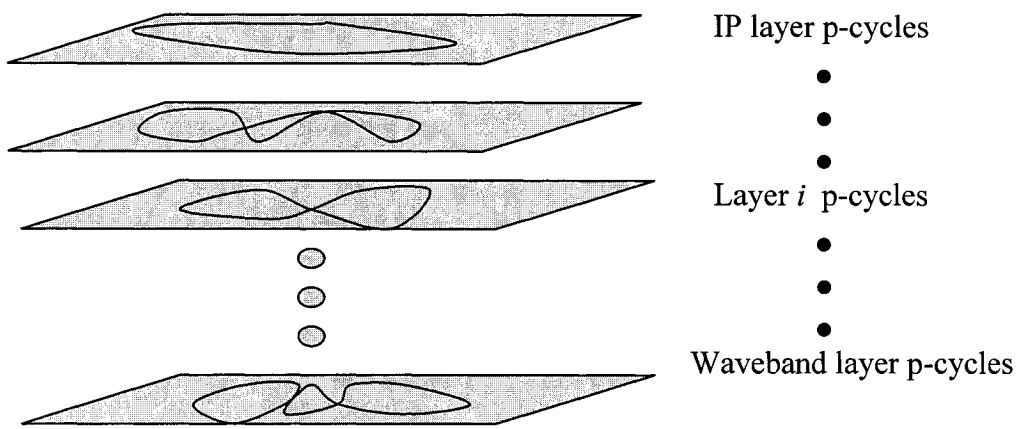


Figure 3-9 Multi-layer P-Cycles

Figure 3-9 shows the structure of multi-layer p-cycles. At each layer, p-cycles are designed independently, but work together when there are failures.

Three influential questions have to be answered before the p-cycle multi-layer survivability can be accomplished. They are summarized as follows.

1) How to identify different layer p-cycles and choose a proper one to use?

2) How can different layers coordinate with each other to prevent unexpected parallel restoration action due to their function duplication?

3) How can one layer inform other layers about the success or failure of the recovery so as to stop or trigger another layer to start a recovery process?

In order to solve the first problem, PMSM assumes that all nodes have the ability to maintain a Link-Cycle Table (see Table 3-1). Figure 3-10 shows a part of a sample network topology. From this figure, one can see it is a topology fragment of a network's layer1. We assume that $L_{1,1}$ and $L_{1,3}$ are protected by p-cycle $P_{1,1}$, where $L_{1,1}$ is an on-cycle link of $P_{1,1}$, $L_{1,3}$ is a straddling link. After all these relationships have been translated into a table, a new type of table, Link-Cycle Table (LCT), is established (as shown in Table 3-1).

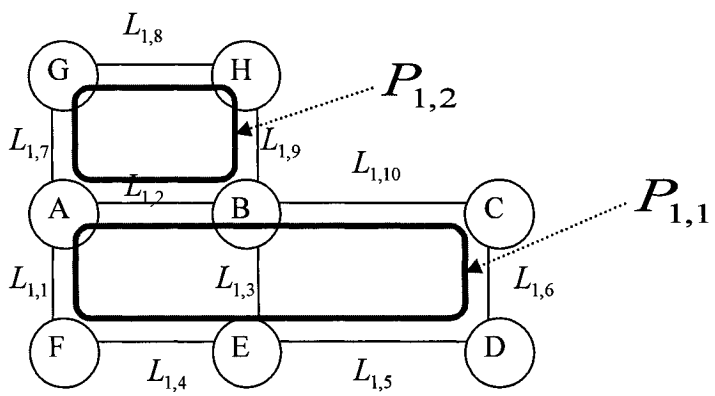


Figure 3-10 Example of a network part

Table 3-1 is formed by using the network topology shown in Figure 3-10.The first column is the Layer ID D_j, where j represents which layer the links and p-cycles are in. The second column is the Link ID $L_{j,i}$, where j is the Layer ID (same as the first

column), and i is the identity of this link (the same definition as SRLG tree). The third column is the P-cycle ID $P_{j,k}$, which is used to distinguish different p-cycles from each other with j representing the Layer ID, and k representing the p-cycle's identity. The fourth column shows all the links that the protection path will traverse when the p-cycle $P_{j,k}$ is used after the link $L_{j,i}$ fails.

D_j (Layer ID)	$L_{j,i}$ (Link ID)	$P_{j,k}$ (P-Cycle ID)	links protection path goes through when using $P_{j,k}$
D_1	$L_{1,1}$	$P_{1,1}$	$L_{1,2}, L_{1,4}, L_{1,5}, L_{1,6}, L_{1,10}$
D_1	$L_{1,2}$	$P_{1,2}$	$L_{1,7}, L_{1,8}, L_{1,9}$
D_1	$L_{1,3}$ *	$P_{1,1}$	$L_{1,1}, L_{1,2}, L_{1,4}$
D_1	$L_{1,3}$ *	$P_{1,1}$	$L_{1,5}, L_{1,6}, L_{1,10}$
...	...	...	...
D_5	$L_{5,i}$	$P_{5,k}$	...

Table 3-1 Link-Cycle Table

Link $L_{1,3}$ was marked with a star in Table 3-1 to emphasize that it straddles $P_{1,1}$, so it gets two protection paths ($L_{1,2}$ - $L_{1,1}$ - $L_{1,4}$ and $L_{1,10}$ - $L_{1,6}$ - $L_{1,5}$) from the same p-cycle $P_{1,1}$. Other on-cycle links get only one protection path from $P_{1,1}$. Accordingly $L_{1,3}$ occupies two rows in the Link-Cycle Table when referring to the same p-cycle.

A network's Link-Cycle Table records its links at all layers, and their protection relationship with all p-cycles, including both the on-cycle and straddling relationship, with the help of the SRLG tree. It can provide the information about the inter-layer

and intra-layer relationship of any link-link, link-cycle (cycle here stands for p-cycle) and cycle-cycle.

The following part of this section is focused on how this Link-Cycle Table responds after a single failure. Suppose there is a Test Network. Figure 3-11 gives an example. After creating a SRLG tree for this network() , the multi-layer p-cycle design will be implemented by using any of the proposed p-cycle design methods, such as DCPC [GROV98], SLA [ZHAN02a], SP-Add [DOUC03a], Expand [DOUC03a], Grow [DOUC03a] or any other optimal p-cycle design [GROV98] [SCHU02a] [SCHU02b] [SCHU03c]. We assume that the corresponding Link-Cycle Table is obtained after the p-cycle design, as demonstrated in Table 3-2.

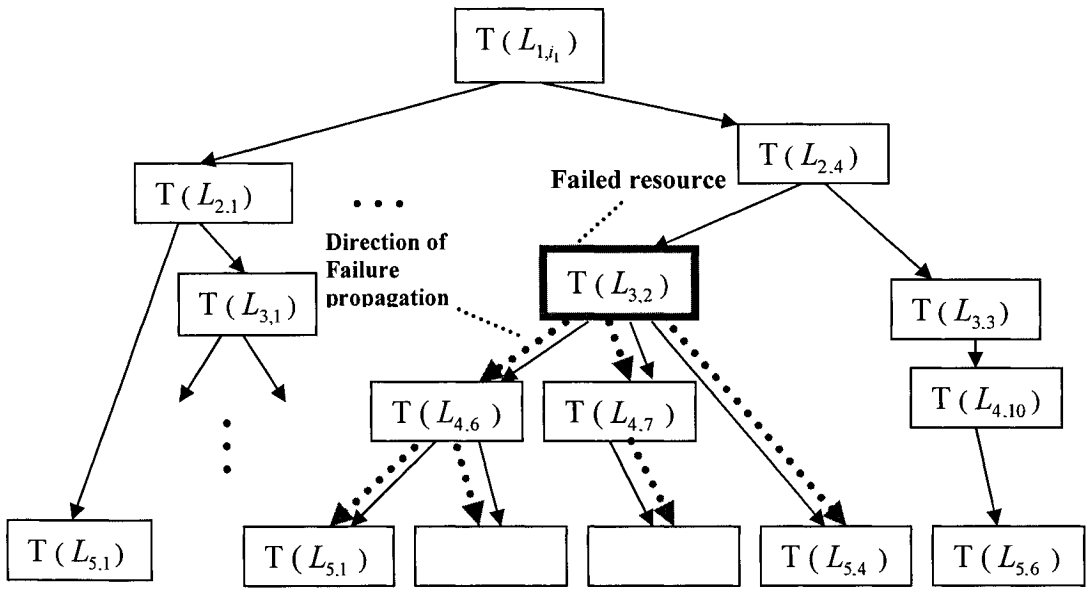


Figure 3-11 SRLG tree of Test Network

D_j (Layer ID)	$L_{j,i}$ (Link ID)	$P_{j,k}$ (P-Cycle ID)	links protection path goes through when using $P_{j,k}$
D_1	$L_{1,1}$	$P_{1,1}$	$L_{1,2}, L_{1,4}, L_{1,5}, L_{1,6}, L_{1,10}$
D_1	$L_{1,2}$	$P_{1,2}$	$L_{1,7}, L_{1,8}, L_{1,9}$
D_1	$L_{1,3}$ *	$P_{1,1}$	$L_{1,1}, L_{1,2}, L_{1,4}$
D_1	$L_{1,3}$ *	$P_{1,1}$	$L_{1,5}, L_{1,6}, L_{1,10}$
D_2	$L_{2,1}$	$P_{2,1}$	...
D_2	$L_{2,4}$	$P_{2,1}$	$L_{2,1}$,...
D_3	$L_{3,1}$	...	...
D_3	$L_{3,2}$	$P_{3,1}$	...
D_3	$L_{3,3}$	$P_{3,2}$	...
D_4	$L_{4,6}$	$P_{4,1}$	$L_{4,7}$,...
D_4	$L_{4,7}$	$P_{4,1}$	$L_{4,6}$,...
D_4	$L_{4,10}$	$P_{4,2}$	...
D_5	$L_{5,1}$	...	...
D_5	$L_{5,2}$	$P_{5,1}$	$L_{5,4}$,...
D_5	$L_{5,3}$ *	$P_{5,1}$	$L_{5,2}, L_{5,4}$,...
D_5	$L_{5,4}$	$P_{5,1}$	...
D_5	$L_{5,5}$	$P_{5,2}$	$L_{5,6}$,...
D_5	$L_{5,6}$	$P_{5,2}$	$L_{5,5}$,...
D_5	$L_{5,i}$	$P_{5,k}$	...

Table 3-2 Format of Link-Cycle Table for Test Network before failure

In the single failure case, it is very easy for the network to find the proper p-cycle after looking at its Link-Cycle Table. In this example, the network will find p-cycle $P_{2,1}$, which is in the same row as $L_{2,4}$. When there are more than one failure, the SRLG Tree and Link-Cycle Table will work together to get the solution. This multi-failure survivability issue will be discussed in Chapter4.

D_j (Layer ID)	$L_{j,i}$ (Link ID)	$P_{j,k}$ (P-Cycle ID)	links protection path goes through when using $P_{j,k}$
D_1	$L_{1,1}$	$P_{1,1}$	$L_{1,2}, L_{1,4}, L_{1,5}, L_{1,6}, L_{1,10}$
D_1	$L_{1,2}$	$P_{1,2}$	$L_{1,7}, L_{1,8}, L_{1,9}$
D_1	$L_{1,3}$ *	$P_{1,1}$	$L_{1,1}, L_{1,2}, L_{1,4}$
D_1	$L_{1,3}$ *	$P_{1,1}$	$L_{1,5}, L_{1,6}, L_{1,10}$
D_2	$L_{2,1}$	$P_{2,1}$	...
D_2	$L_{2,4}$ (down)	$P_{2,1}$	$L_{2,1}$,...
D_3	$L_{3,1}$	...	...
D_3	$L_{3,2}$ (down)	$P_{3,1}$	...
D_3	$L_{3,3}$ (down)	$P_{3,2}$	...
D_4	$L_{4,6}$ (down)	$P_{4,1}$	$L_{4,7}$ (down),...
D_4	$L_{4,7}$ (down)	$P_{4,1}$	$L_{4,6}$ (down),...
D_4	$L_{4,10}$ (down)	$P_{4,2}$	...
D_5	$L_{5,1}$ (down)	...	...
D_5	$L_{5,2}$ (down)	$P_{5,1}$	$L_{5,4}$ (down),...
D_5	$L_{5,3}$ (down)	$P_{5,1}$	$L_{5,2}$ (down), $L_{5,4}$ (down),...
D_5	$L_{5,4}$ (down)	$P_{5,1}$	...
D_5	$L_{5,5}$ (down)	$P_{5,2}$	$L_{5,6}$ (down),...
D_5	$L_{5,6}$ (down)	$P_{5,2}$	$L_{5,5}$ (down),...
D_5	$L_{5,i}$	$P_{5,k}$	...

Table 3-3 Link-Cycle Table for Test Network after failure

Beside the multi-failure survivability, another advantage of the Link-Cycle Table is the ability to provide sufficient information for different layers to coordinate with each other. The coordination problem between different layers in multi-layer networks has been a big challenge for a long time because of its impotency and difficulty. The Link-Cycle Table makes it easier by providing updated and detailed bandwidth/failure information to each node.

Technically networks can choose any layer as the protection layer. From the bandwidth consuming and recovery time points of view, networks still need to choose a proper layer to execute the recovery process. The lower the layer is, the more bandwidth it will consume to set up a backup path. On the contrary, the faster the recovery period will. There is always a trade-off between bandwidth consumed and recover time. Another interesting phenomenon is: when a single wavelength fails, hundreds of IP links may fail at the IP layer if this wavelength link is underneath them. If the IP layer is selected as the protection layer, this single wavelength link failure will be translated to multiple IP link failures. This increases the difficulty to recover the failure, or even cause the network to become unrecoverable. So the achievable availability of a protection connection depends highly on the proper identification and selection of a proper protection layer.

In PMSM, it is assumed that the system always starts recovery process at the layer that detects the failure. Usually nodes adjacent to the failure will detect the failure first. This means the layer that first detects the failure will be the layer where the failure is exactly located. If this failure can not be recovered at this layer, the adjacent upper layer will trigger its recovery process. There are also special cases that one needs to consider carefully. In a sparse network, maybe only several MPLS links are multiplexed inside a wavelength link. When this wavelength link is out of order, it is better to choose MPLS layer as the protection layer to recover those MPLS links, this way a lot of bandwidth can be saved. Because the SRLG Tree

offers the guarantee that working and backup paths of any requested connection are diverse against the failure at any desirable protection layer, the network can choose the best p-cycle without making any mistake [NASE04].

PMSM gives networks the flexibility to choose any layer protection with a default protection layer to be the layer that first detects the failure.

3.6 Summary

Before the development of GMPLS and IP/MPLS/WDM, networks used to work individually, even after the advent of IP over WDM, IP over SONET, and IP over MPLS, where different parts of networks still have their own protocols and protection mechanisms. GMPLS technology has made it possible to substantially merge all these networks together, and allow them to speak in the same language. By separating the control plane from the data plane, GMPLS provides an opportunity of seamless integration of multiple layers.

PMSM takes advantage of GMPLS's achievement for the p-cycle design. After setting up the SRLG tree for a desired network and storing the useful information inside a central database, the network is ready to provide multi-layer survivability using p-cycle. With any of the proposed p-cycle algorithms, a Link-Cycle Table can be established for each link. Once a failure occurs, the system will look at the Link-Cycle Tables and select a proper p-cycle at a proper layer to provide recovery.

There are two common network models to connect different layers together.

1) ***The peer model.*** Each node in the network has full capacity to gather the information of its own layer (link state, bandwidth, residual capacity, working capacity, spare capacity, etc...) and exchange the data with other nodes, including nodes at the same layer and nodes at different layers.
2) ***The overlay model.*** Only edge nodes and some important nodes (called as central nodes) have the ability to communicate with different layer nodes. All the other nodes have to get the required data from the central nodes. This model is more cost-effective with slower recovery speed.

PMSM can fit both models, because the system keeps all the necessary information inside a central database which can be shared by all nodes.

A new field is added into the regular p-cycle packet. This new field is the Layer ID, which is used to identify different layer p-cycles. The p-cycle packet format needs to change accordingly, as shown in Figure 3-12 and Figure 3-13 (refer to Figure 3-1 and Figure 3-2).

p-cycle ID	Layer ID	Destination address	Original Route cost	**IP Packet**

Figure 3-12 The p-cycle packet format at the IP layer

p-cycle ID	Layer ID	Destination address	Original Route cost	**MPLS Packet**	
				MPLS header	**IP Packet**

Figure 3-13 The p-cycle packet format at the MPLS layer

IP layer p-cycles have their unique characteristic. The IP layer is already restorable in the sense that routing protocols (such as OSPF and BGP) will indirectly provide protection after reconvergence [STAM00a]. P-cycles provide an immediate real-time detour to prevent the packet loss until conventional global routing reconvergence occurs. Thus the p-cycle protection will serve as an immediate real-time detour, in other word, a fast-acting but temporary protective measure which secures the network traffic. That is to say, IP p-cycles are envisaged as the "fast" part of a "fast plus slow" overall recovery process. This is also true in PMSM. Not only the IP layer, but also the MPLS layer had the same ability to recover itself. P-cycles act as a temporary fast recovery solution at both layers.

Chapter 4 MULTI-FAILURE SURVIVABILITY

4.1 Introduction

Before the blooming of internet, networks that can survive one single failure are sufficient for customers. Due to the exponential increase of service demands, the possibility of dual (even multiple) failures becomes inevitable. Customers start to pursue more robust and reliable services, which have pushed service providers and researchers to put more energy on searching for multiple failure solution.

So far, most of the work on p-cycle network design has focused on efficiently providing a guarantee for 100% restorability against any single span failure [SCHU04]. Some papers have discussed about dual-failure problems [SCHU03b] [SCHU03a], but they just made an analysis on the cases that multiple failure can be restored by p-cycles, left the situation that can not be survived by p-cycles unsolved. [SCHU04] recently produced a method to deal with multiple failures based on LLP design problem with a high complexity.

In this chapter, we will present a new method to provide multiple failure survivability, which is easier, simpler and more efficient than the solution given by [SCHU04].

4.2 Multi-Failure Survivability

Multiple failures can take place at one layer or different layers in a network. The latter problem is easier to solve. The reason is as follows. If two failure links (at different layers) have a relationship with each other (either ancestor or descendant of each other according to the SRLG tree), recovering the lower layer link will survive the higher layer link as well. If they have no relationship, which means one link is neither a parent nor a child of the other and they have no inheritance relationship, the network can execute both layer p-cycle protections simultaneously to survive both links at the same time. The former case (failures occur at the same layer) is more difficult to solve.

When multiple links fail at the same layer, they can automatically be recovered if these failure links belong to different p-cycles. In the view of any individual p-cycle, there is only one failure, so each p-cycle can provide recovery for each link separately. When multiple failures presented themselves inside the same p-cycle, this p-cycle will not be able to survive the second one, if there is no reconfiguration [SCHU03a] [SCHU03b]. In this chapter, we provide a solution to survive multiple failures which are at the same layer and the same p-cycle. It is called Multi-Failure Survivability Scheme (MFS).

There are two general ways to solve a multiple-failure problem.

1) Trying to provide the maximum multi-failure survivability ability while keeping the same network capacity.

2) Providing up to 100% multi-failure survivability by adding some additional capacity while keeping the additional capacity at a tolerant level or minimizing the extra capacity added.

MFS works with both models.

MFS has two procedures: the off-line calculation and the on-line selection. After the system completes the normal p-cycle calculation (any of the proposed algorithm in literature can be used, we call p-cycles found this way as "normal p-cycles"), it sets up Link-Cycle Tables (LCTs) [HWAN05] and becomes eligible for single failure. Based on the LCT and the network topology, a further execution is employed to find some backup p-cycles for each link. This procedure is similar to the one that to find the "Kth shortest path". These backup p-cycles are called as "additional p-cycles". Normal p-cycles and additional p-cycles together form a p-cycle pool for each link. Both the normal p-cycle algorithm and the backup p-cycle algorithm are complicated and time-consuming. They are done in an off-line manner. When there is failure (single failure or multiple failure), the system will choose the best p-cycle from a proper layer to react. This p-cycle selection action is an on-line process. It is very fast and needs much less CPU time than the off-line calculation.

The off-line and on-line multi-failure survivability scheme (MFS) makes use of both the precise and thorough calculation, and enables the system to behave within a short time. All the complicate calculation is done off-line and the on-line selection is very simple, so the recovery time will not be influenced by the complexity of the off-line calculation algorithm.

4.3 Off-Line Centralized Calculation

4.3.1 A Valuable Observation

Through the SRLG tree, each node in the network gets the most updated knowledge about the residual capacity, the working capacity and the spare capacity of all links. This information is also stored inside a central database for sharing purpose.

During study, we got an interesting observation: when we mention the failure recovery, we always suppose that there is a single failure, dual failures, or multiple failures. When there is a single failure, it is automatically assumed that all the other links and nodes, except this failed resource, are still working. When there are dual failures, only these two resources are out of service. The same claim can be applied to multi-failure cases. If all resources are down individually, and not caused by a common reason, the whole network will be out of service, no need to talk about survivability. This means, at a certain time, only one or several resources, more specifically in this thesis, only one or several links (spans) are down, while the other links (spans) are still working perfectly. This observation gives us an idea that, when there is a single failure, we can use the network's whole spare capacity to do our

protection job; when second failure comes out, the rest of the spare capacity can be used, which is the total spare capacity minus the part that has been used for the first failure; so as to the third failure, the fourth failure, …, and the nth failure.

4.3.2 OPCA algorithm

Based on the above valuable observation, we contrived our off-line p-cycle calculation algorithm (OPCA Algorithm).

1) Calculate normal p-cycles by using any regular p-cycle design method. Establish the Link-Cycle Table for the network. More than one p-cycle can exist for a certain link. In this case, there will be more than one p-cycle in the LCT for this link. After gathering information about normal p-cycles, the network is capable to survive any single link failure (or multiple failures that are inside different p-cycles).
2) Based on the normal p-cycle design, establish a protection topology for the network. A network's protection topology is built up by all the available spare capacity of all links inside the network. With the help of failure propagation and bandwidth propagation inside the SRLG tree, the network can get to know exactly how much bandwidth is left at each link. If one link fails, its capacity (both working capacity and spare capacity) will drop to zero. A link should have at least one unit spare capacity to be able to be counted; otherwise, its spare capacity will be marked as zero, and it will not appear in the network's protection topology.
3) For each link, calculate several shortest paths with lowest costs by using Dijkstra's Shortest Path Algorithm or any other shortest path algorithm to get

K paths in the protection topology. The source and the destination of these K paths are this link's two end nodes. K can be any number as one wishes. It is possible that the shortest paths found are less than K, when there are less than K disjoint paths between this link's two end nodes. We assume that the number of the final shortest paths found is k. Combining these k shortest paths together with the link itself, we get k backup cycles. We call these backup cycles "Additional P-cycles" (AP) as opposite to normal p-cycles. Figure 4-1 shows an example, indicating how to find additional p-cycle for link $L_{1,2}$. Figure 4-1(a) is the network's protection topology after procedure (1) (see Figure 3-10). The system finds the first shortest path using Dijkstra's Shortest Path Algorithm, called SP_1. SP_1 and $L_{1,2}$ together form a cycle $P_{1,2}$. Cycle $P_{1,2}$ becomes the first additional p-cycle for $L_{1,2}$. All the links, except link $L_{1,2}$ itself, which are in the same p-cycle as $L_{1,2}$ are deleted from the topology (see Figure 4-1(b)). At the rest topology, Dijkstra's Shortest Path Algorithm is deployed again to get another shortest path.

4) This new path together with link $L_{1,2}$ form another cycle. This cycle will be the second additional p-cycle for link $L_{1,2}$. These additional p-cycles are named as AP_i, where "A" means additional, and "AP" stands for additional p-cycle. AP_i means the i_{th} additional p-cycle for a certain link. This procedure can keep going until there is no additional p-cycle can be found at the rest of the protection topology or the desired amount of additional p-cycles has already been found. This method assures that all additional p-cycles found

are disjoint of each other, because the links which are used by one cycle will not be used by any other cycles except the link being protected.

5) Rank additional p-cycles. All the additional p-cycles are ranked in the light of their topological score (TS) or apriori efficiency (AE) [GROV02b] or any other desirable criteria which are suitable for an individual network.

6) Establish P-Cycle Pool. After additional p-cycles are ranked, they are placed inside a new table with all normal p-cycles in front of them. This new table represents a P-Cycle Pool Table (PPT) for the link being protected (see Table 4-1).

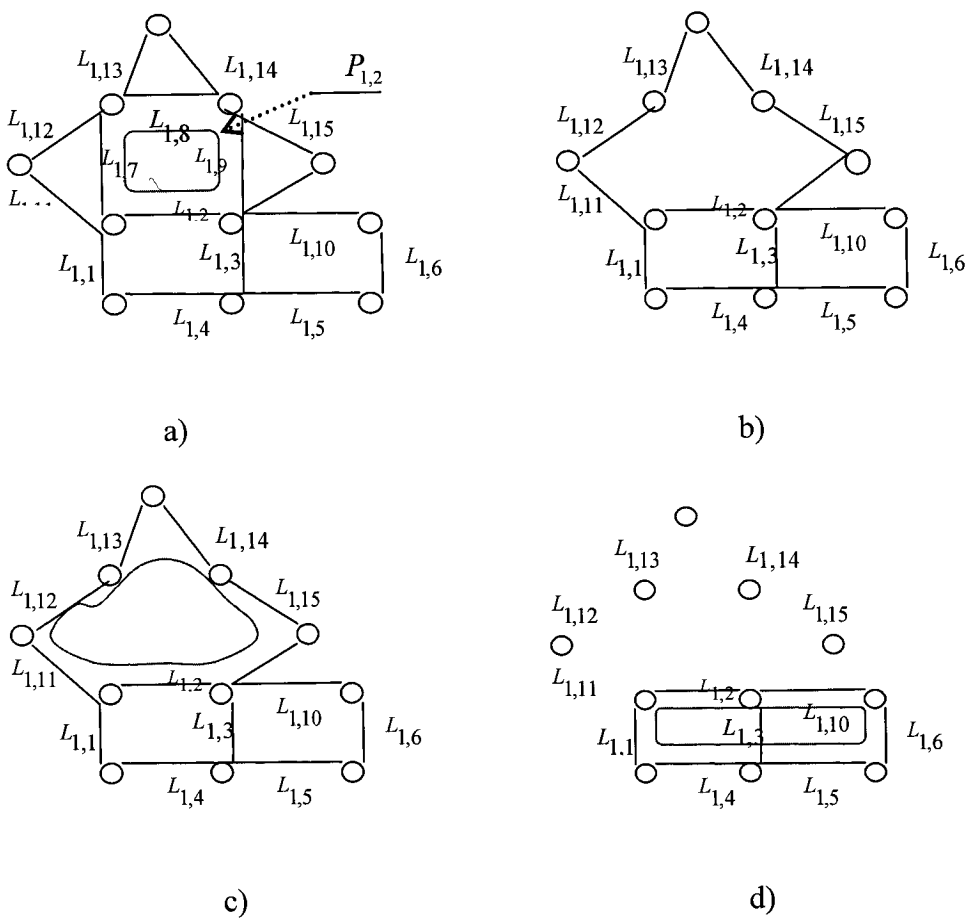


Figure 4-1 How to find Additional P-cycles

The p-cycle pool is formed by normal p-cycles together with all additional p-cycles. Any normal p-cycle (NP) is always placed on the top of any additional p-cycle (AP). In case if there is more than one normal p-cycle, they will be ranked by their TS or AE first, and be put at the proper position inside the p-cycle pool table. APs that have higher ranks will always get topper positions than APs with lower ranks. Whenever there is a failure, the system will choose a proper p-cycle from the p-cycle pool. Each node maintains all its adjacent links' p-cycle pools.

We would like to emphasize that the topology used here is the network's protection topology, not the original network topology. When a link is up, and it is not a part of any previous p-cycles, if its spare capacity has been used up, it still will not be able to provide any protection unit for any p-cycle. So this link will not appear at the protection topology. Also when a link fails, it will not have any available capacity any more. It will also be deleted from the protection topology. This guarantees that p-cycles found through the off-line calculation can really provide protection units. Each p-cycle offers exact one protection unit, the same as any other papers.

The advantage of our OPCA algorithm is obvious: additional p-cycles are diverse from each other (except the protected link itself). This characteristic enables the network to be able to survive multiple failures, which happen at the same time, inside the same p-cycle.

4.3.3 Some Improvements

In a dense network where the network capacity is almost used up, it is hard to find proper APs. In this condition, some changes can be made to the OPCA algorithm: try to find additional p-cycles as diverse as possible from each other, not necessarily to be 100% diverse. So the OPCA algorithm will have more chance to find APs for links.

An upper bound can also be defined: let the maximum amount of APs be n for each link. After finding n APs for a certain link, the system will stop searching and turn to the other link. If before the system gets n APs for a link, it can not find any proper AP any more, it will also turn to another link.

4.4 On-Line Distributed Selection

The on-line selection is a simple and fast behavior. It is executed after each link has got its p-cycle pool. Table 4-1 shows a sample P-cycle Pool Table (PPT) for $L_{1,2}$.

Inside a network, each node maintains its adjacent links' PPTs while the central database keeps record for the whole network, including different layers, different links, and their relationship.

As soon as nodes adjacent to the failure detect failure signals, they will look at the corresponding PPTs and choose proper p-cycles to use. This selection is always done with respect to p-cycles' rank, which means the system chooses p-cycles in a top-

down manner from the PPT. If the first one is not available, then the second; if the second one is still not available, then the third; till get one or fails to find any.

$L_{1,2}$:

$P_{1,2}$	$L_{1,7}, L_{1,8}, L_{1,9}$
AP_1	$L_{1,11}, L_{1,12}, L_{1,13}, L_{1,14}, L_{1,15}$
AP_2	$L_{1,1}, L_{1,4}, L_{1,5}, L_{1,6}, L_{1,10}$

Table 4-1 P-cycle Pool Table (PPT) for link $L_{1,2}$

If any link, through which the p-cycle traverse, fails or its bandwidth has been used up (e.g. already be used for other failure recovery), nodes will get this information, either through the SRLG tree or by its own protocol (when changes are inside the same layer), and revise their p-cycle pools' state. For example, if $L_{1,7}$ is down or does not have enough spare capacity, $L_{1,7}$ will be marked as "unavailable" inside the PPT. Whenever there is one link unavailable, the corresponding p-cycle will definitely become unavailable. Once the state of links changes, the status of p-cycles at each PPT will change right after. Table 4-2 shows an example of this change when $L_{1,7}$ is unavailable. The first available p-cycle inside the PPT, which is AP_1, will be chosen to recover $L_{1,2}$.

$L_{1,2}$:

$P_{1,2}$ (*unavailable*)	$L_{1,7}$ (*unavailable*), $L_{1,8}, L_{1,9}$
AP_1	$L_{1,11}, L_{1,12}, L_{1,13}, L_{1,14}, L_{1,15}$
AP_2	$L_{1,1}, L_{1,4}, L_{1,5}, L_{1,6}, L_{1,10}$

Table 4-2 P-cycle Pool Table for $L_{1,2}$ after $L_{1,7}$ becomes unavailable

The on-line selection follows an up-to-down order to make sure that the best available p-cycle will be selected first. P-cycles are marked "unavailable" at one time does not mean that they will be always unavailable. They can be available again, when the condition that makes it unavailable is removed. Each time the selection starts from the top of the PPT.

4.5 Class of Protection

An additional function that can be achieved by MFS is the ability to provide multilevel survivability, more specifically, the Class of Protection.

In the literature, several methods with different criteria have been introduced in the protection/restoration area. These include the dedicated 1+1 protection/ restoration, the shared 1:1, 1:n or m:n protection/restoration and the p-cycle protection. Each of them varies from each other in two key prospects: the capacity efficiency and the restoration speed. The dedicated 1+1 method is the fastest way for a survivable network, while it suffers from the high requirement of redundant capacity. The shared protection/restoration is capacity efficient, but it is relatively slow. The p-cycle is a newly proposed concept. It is the only technique that can achieve both mesh-efficiency and ring-speed.

In large networks, traffic differs from each other dramatically. Real-time multimedia services always have high priority, such as on-line banking services, tele-medicine services, and etc. There are also best-effort services that do not need any protection

capacity. Some other services are in the middle, which have lower priority than multimedia services, but higher priority than best-effort services. Different kinds of traffic with different priorities need to be treated differently in the protection/restoration domain. The concept of Class of Protection (CoP) has become one of the most important criteria to evaluate a network's quality.

MFS can be a platform for CoP. Each traffic gets a priority number that shows its required level of CoP upon its arrival. In this thesis, we define 3 level priorities P1, P2, P3, where P1 is the highest priority and P3 is the lowest. The first class protection is the "dedicated 1+1 protection". The second class is the "p-cycle" and the third one is the "shared protection" that can be either 1:1, 1:n or m:n. These 3 different levels are numbered by their different "restoration time" (see Table 4-3). A Protection Pool is formed by all protection paths that are calculated by these 3 different methods.

At the stage of off-line calculation, the system calculates several protection paths for every link, using these three different protection technologies (dedicated 1+1, shared 1:1 or 1:n or m:n, and p-cycle). Once a failure is detected, the on-line calculation will be operated, and the best protection path will be chosen to recover the failure.

Class of protection	Restoration time	M
Dedicated 1+1 protection	<60ms	1
p-cycle protection	<80ms	2
Shared protection	<200ms	3

Table 4-3 Class of protection (CoP) [BLOU03]

Table 4-4 demonstrates a sample protection pool for link $L_{1,2}$. The first row is M1 (Level 1 protection). The "D" in D1 presents "dedicated protection". The "S" in S1 means "shared protection". All protection paths are scheduled based on their CoP (Class of Protection). It is not necessary for the system to always calculate all these three level protection paths for all traffic. Some services do not need the dedicated protection. Sometimes the shared protection may be sufficient for a service. Best-effort services even do not need to consume any specific spare capacity at all. Their protection pool can be empty. In general, some traffic will have M2 and M3 protection paths, but no M1 level path while others may only have M3 level path. Once a link has a certain level protection path, the system will compute all its lower level protection paths. Whenever a failure takes place, nodes adjacent to the failure will look up the corresponding protection pool table and choose the first available protection path to perform. When there is more than one failure, and their first available protection paths need to traverse through the same spare capacity, their priority (P1, P2 or P3) will be compared. The pre-empty mechanism is used. The higher priority service will take this bandwidth and let the lower priority one to search for another protection path from its protection pool.

$L_{1,2}$:

M1	Protection path D1	All the links D1 traverse through
M2	$P_{1,2}$	$L_{1,7}, L_{1,8}, L_{1,9}$
M2	AP_1	$L_{1,11}, L_{1,12}, L_{1,13}, L_{1,14}, L_{1,15}$
M2	AP_2	$L_{1,1}, L_{1,4}, L_{1,5}, L_{1,6}, L_{1,10}$
M3	Backup path S1	All the links S1 traverse through

Table 4-4 Protection Pool [HWAN05]

4.6 Rearrangeable On-Line Selection

In a real service network, the ability to recover failures in order to make the network stable is more important than always providing the first degree service for clients, as long as the service level is in the range that can be accepted by clients. Based on this, we discovered an improved version of on-line selection algorithm, which is called "Rearrangeable On-line Selection" that can significantly increase network's ability to survive multiple failures. MFS using the rearrangeable on-line selection algorithm is called "Rearrangeable MFS" as opposite to the "Regular MFS", which uses the regular on-line selection algorithm.

The regular on-line selection algorithm always chooses the first available p-cycle from the failure link's PPT. Inside the PPT, p-cycles are ordered by their scores. P-cycles with higher scores always have higher chance to be selected, especially for the first failure. In case if this p-cycle traverses any link that is an on-cycle link of all p-cycles in the PPT for the second failure, the system will fail to survive the second one, although there exists a pair of p-cycles (one to survive the first failure and the other one for the second failure) that can recover both failures. Also because of p-cycle's own characteristics, the one that has longer circumstance tends to have higher rank, so the first p-cycle in the PPT usually traverses more links than any other p-cycles, and makes the possibility that the system fails to survive the second failure, very high. This badly affects the performance of our protection scheme.

A new version of on-line selection that is called rearrangeable on-line selection can solve this problem. For the first failure (suppose it is link1), the system will choose the first p-cycle from its PPT. When the second failure (suppose it is link2) happens, the system scans its p-cycle pool, if no p-cycle is available for the second failure, it will try the second p-cycle from link1's p-cycle pool, and scan link2's p-cycle pool again. This procedure continues until the system finds an available pair of p-cycles to survive both failures or it fails to find this pair.

The flowchart of the rearrangeable on-line selection algorithm is shown in Figure 4-2.

4.7 Summary

In this chapter, we initially present our motivation to provide multiple failure survivability. A novel Multi-Failure Survivability Scheme (MFS) is introduced, which consists of two stages: the off-line centralized calculation and the on-line distributed selection. In the off-line stage, an off-line p-cycle calculation algorithm (OPCA algorithm) is used to form a P-cycle Pool for each link in the network. Then a P-cycle Pool Table is built up for each link. Every node maintains the PPTs of its incoming and outgoing links. All PPTs are stored inside a central database for sharing purpose. If any node wants to check any information about other links' PPTs which are not maintained by itself, it will refer to the central database. When failures occur, the system triggers an on-line selection process and chooses proper p-cycles to recover the failures. A rearrangeable on-line selection algorithm is also proposed that can significantly increase the performance of MFS.

When the regular on-line selection algorithm is used, it is called "Regular MFS" while the name of "Rearrangeable MFS" is used for the rearrangeable on-line selection algorithm. Both of them can provide the Class of Protection (CoP) for different services.

MFS takes maximum advantage of the network resource to provide maximum survivability of multiple failures without adding any additional network capacity. If 100% dual failure survivability is needed, some additional capacity can be added at certain parts of the network.

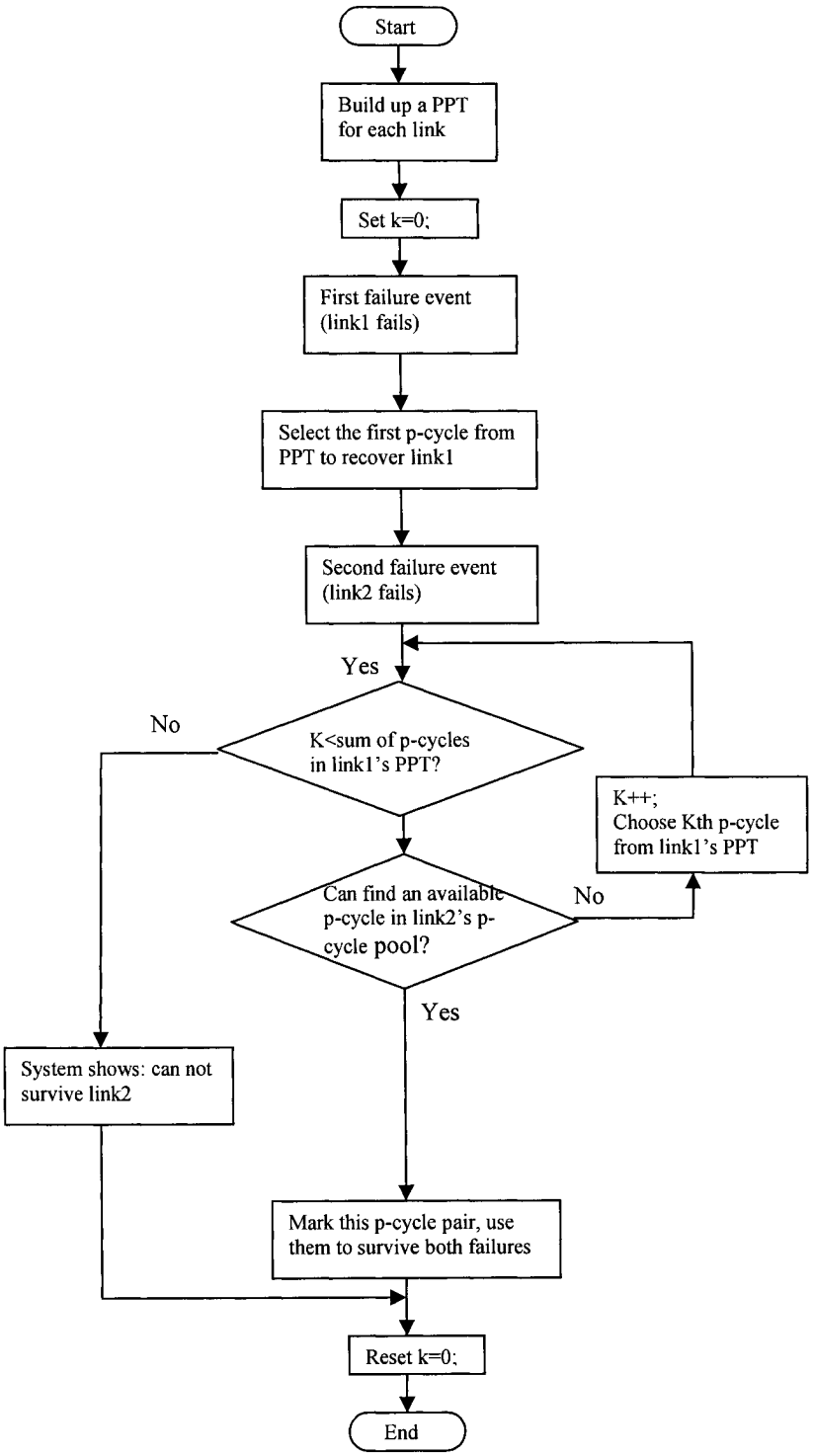


Figure 4-2 Flowchart of Rearrangeable On-line Calculation

Chapter 5 SIMULATION RESULTS AND PERFORMANCE ANALYSIS

5.1 Introduction

This chapter addresses the simulation issue. In this study, we performed intensive simulations to evaluate the performance of our 2-stage multi-failure survivability scheme (MFS) [HWAN05]. As we already discussed in the previous chapters, when multiple failures are at different p-cycles or the p-cycle reconfiguration is available, these multiple failures can be recovered by regular p-cycle designs. MFS is aimed to solve the most difficult situation, which may not be solved by the regular p-cycle solution: multiple failures occur simultaneously at the same p-cycle and the p-cycle reconfiguration is not possible. In this chapter, a simulator is developed. Simulation results are gathered and carefully analyzed to evaluate the performance of MFS.

In current networks, the possibility of dual failures is much higher than triple failure or more failures. So in this chapter, only dual failure cases are monitored.

The simulation process is made up of 3 parts:

(1) ***Get the proper p-cycle candidates***. The Grow Algorithm [DOUC03a] is used to find out all the proper cycles of the network. These cycles will be used as p-cycle candidates for further optimization. The Grow Algorithm is evolved from the Straddling Link Algorithm (SLA) [ZHAN02a] and has been considered as a simple and fast procedure to produce an initial subset of cycles with good efficiency, when compared with the optimal pure ILP solution, which is to find out all possible cycles of a network regardless their scores or ranks. We choose the Grow Algorithm because it has low complexity and good efficiency.

(2) ***Find p-cycles***. This is a p-cycle design procedure. We use AMPL to build up the Integer Linear Programming (ILP) model and use CPLEX as the solver to get the optimized solution. All cycles that have been chosen by CPLEX become p-cycles.

(3) ***Calculate the probability to survive dual failures***. Feed p-cycles that are found by CPLEX back to the third part of our simulation program to get the probability to recover dual failures. Two link failures that are inside a single p-cycle are randomly chosen. MFS is used to try to find a p-cycle pair to survive them. The probability of success is recorded.

Our simulator is implemented in Java language with object-oriented system design approach. The source code exceeds 5,800 lines.

5.2 Specification of Simulation

5.2.1 Test Networks

All simulations are performed over three different network topologies, i.e., the NSF backbone network, the US long haul network and the EON network. These networks are three widely used network topologies in both research and application fields, also they have different characteristics, such as the size and the degree.

Topologies of the NSF network, the US long haul network and the EON network are shown in Figure 5-1, Figure 5-2 and Figure 5-3, respectively.

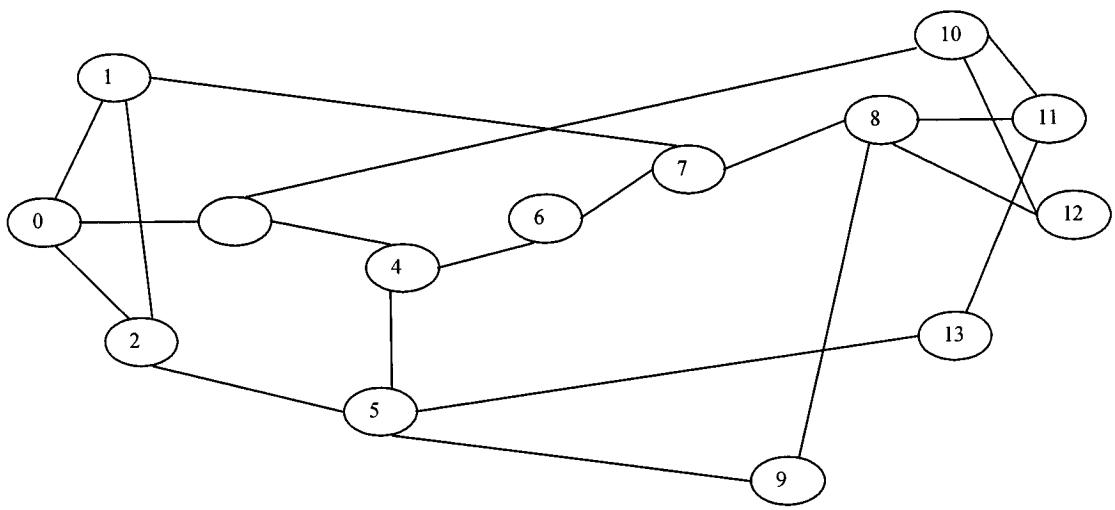


Figure 5-1 NSF backbone network topology (14 nodes, 21 links)

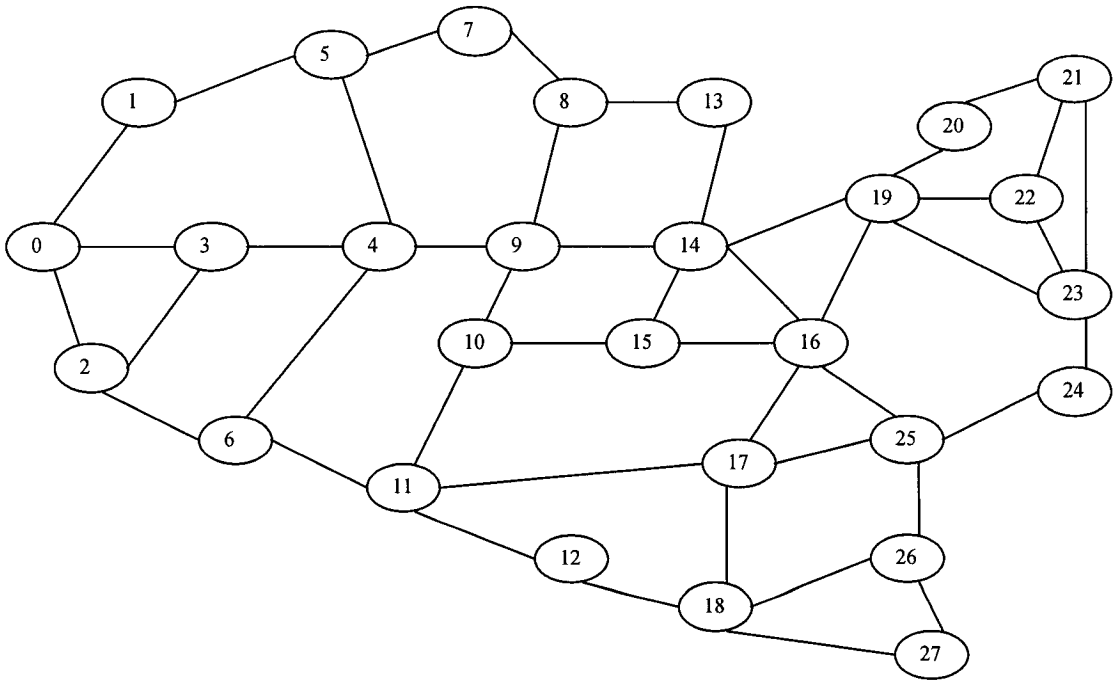


Figure 5-2 US long haul network topology (28 nodes, 45 links)

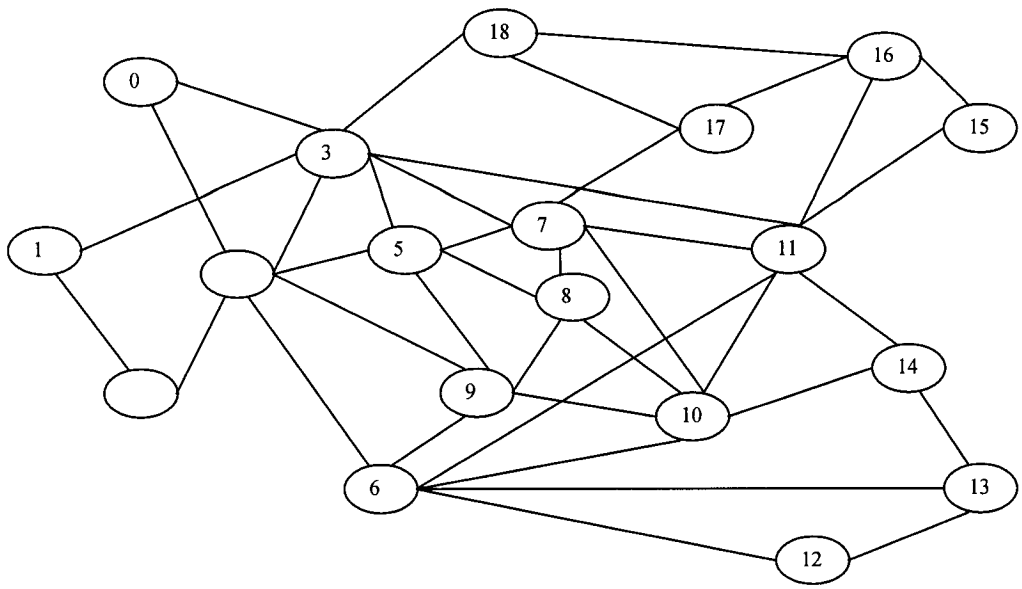


Figure 5-3 EON network topology (19 nodes, 39 links)

Table 5-1 shows some important characteristics of these three networks. Their differences in these three key aspects lead to different simulation result as we can see from the numerical results.

	NSF	US long haul	EON
Number of Nodes	14	28	19
Number of Links	21	45	39
Degree of network	3	3.21	4

Table 5-1 Three sample networks

5.2.2 Working Channel

In a real network, each span can have more than one working channel. At a certain time, some of these working channels are carrying the traffic, while the others are not. To investigate the performance of MFS scheme under different traffic patterns, we conduct simulations at different conditions. The number of working channels in each span is chosen to be 8, 16 or 32. In an optical network, this means each span has 8, 16 or 32 available wavelengths. The number of channels that carry the real traffic is randomly chosen from 0 to K (K is chosen to be 8, 16 or 32), while the others are idle. For example, in the case that each span has 8 working channels, if the number of busy channel is randomly chosen to be 3, this means there are 3 channels carrying the real traffic, and 5 other channels are idle. The p-cycle design will be executed based on 3 busy channels for this span. The other 5 idle channels can be used by p-cycles as a part of the network's spare capacity, or remain as the working capacity for further traffic demands.

The purpose to define different numbers of working channels (8, 16 or 32) is to find out whether the different number of working channels will have impact on the performance of MFS scheme.

5.2.3 Sample Time

In order to get more general results, the number of busy channel at each span is generated 10 times for each case (8, 16 or 32 working channel). Each time 1000 simulation runs are executed. Finally 10 results are gathered for each case and the average of the probability is calculated as final results.

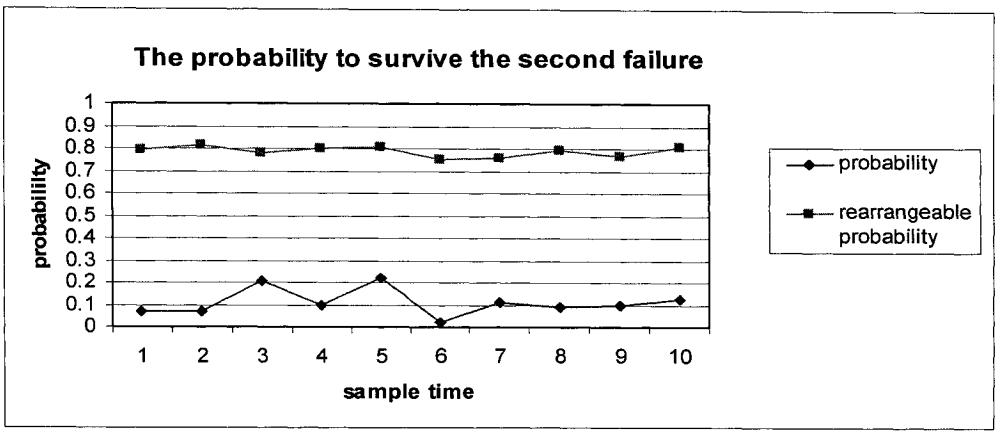
5.3 Results and Performance Analysis

In this section, we present various simulation results and conduct our analysis based on these results.

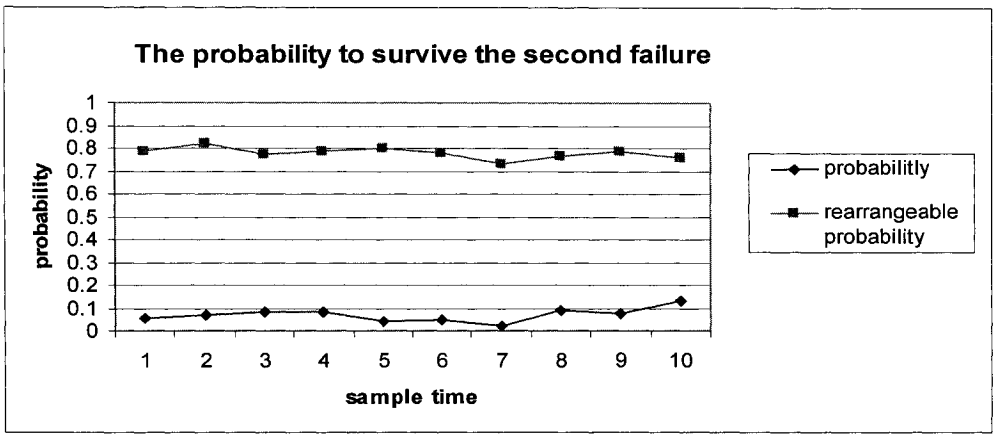
5.3.1 Numerical Results

5.3.1.1 Data Obtained From the NSF Network

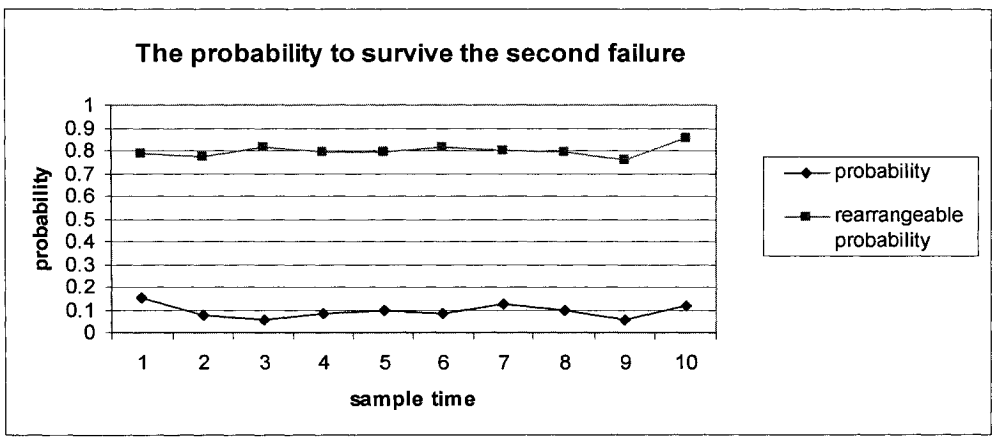
Numerical results of the probability to survive a second failure in the NSF network are shown in Figure 5-4(a)-(c). Figure 5-5 demonstrates the results when all the curves in Figure 5-4(a)-(c) are put inside one graph.



(a) NSF Network Probability when K=8



(b) NSF Network Probability when K=16



(c) NSF Network Probability when K=32

Figure 5-4 Probability for NSF Network to survive a second failure

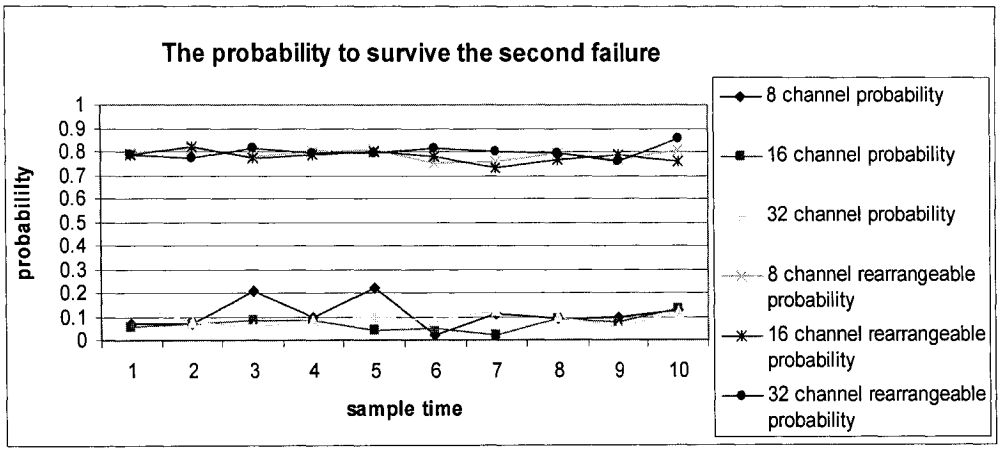
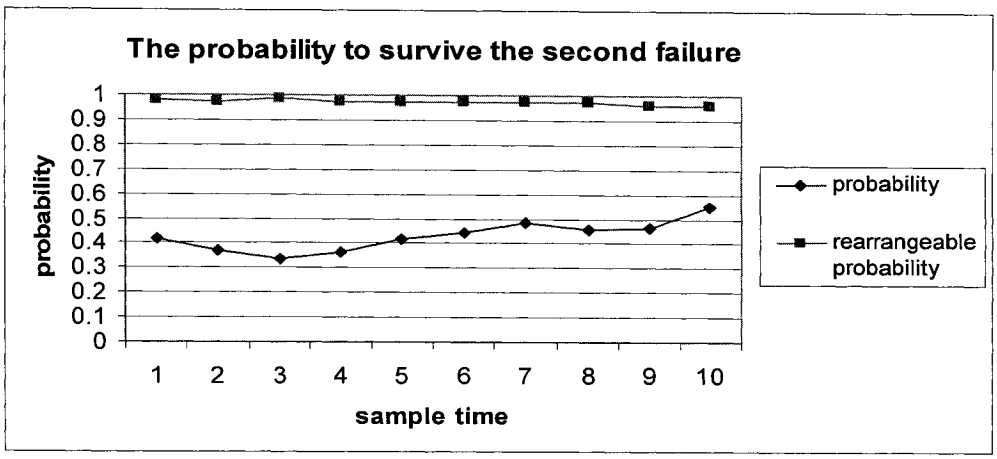


Figure 5-5 Probability of NSF network in all 3 cases

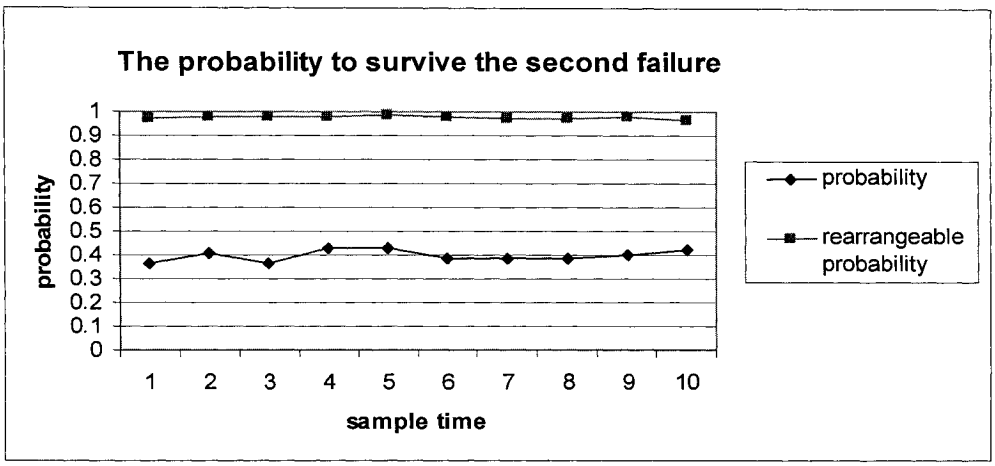
It is evident from the graphs that in the NSF network, the second failure can be survived by our MFS scheme with a probability of around 0.1 and 0.8, when using regular MFS and rearrangeable MFS, respectively, without adding up any extra capacity. All graphs show that the rearrangeable MFS has much better performance than the regular one. When the number of working channel changes from 8, 16 to 32, the probability of success remains almost the same.

5.3.1.2 Data Obtained From the US Long Haul Network

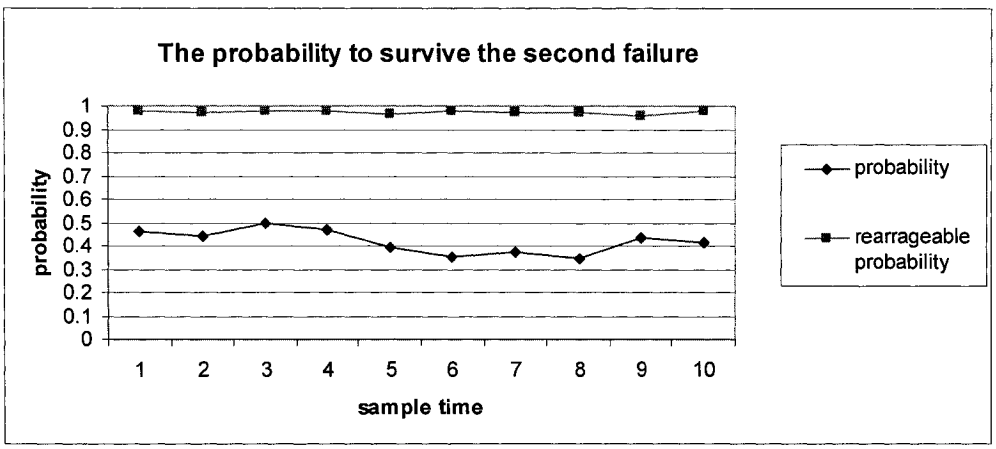
Numerical results of the probability to survive a second failure in the US Long Haul Network are shown in Figure 5-6(a)-(c). Figure 5-7 demonstrates the results when all the curves in Figure 5-6(a)-(c) are put inside one graph.



(a) US Long Haul Probability when K=8



(b) US Long Haul Probability when K=16



(c) US Long Haul Probability when K=32

Figure 5-6 Probability for US Long Haul to survive a second failure

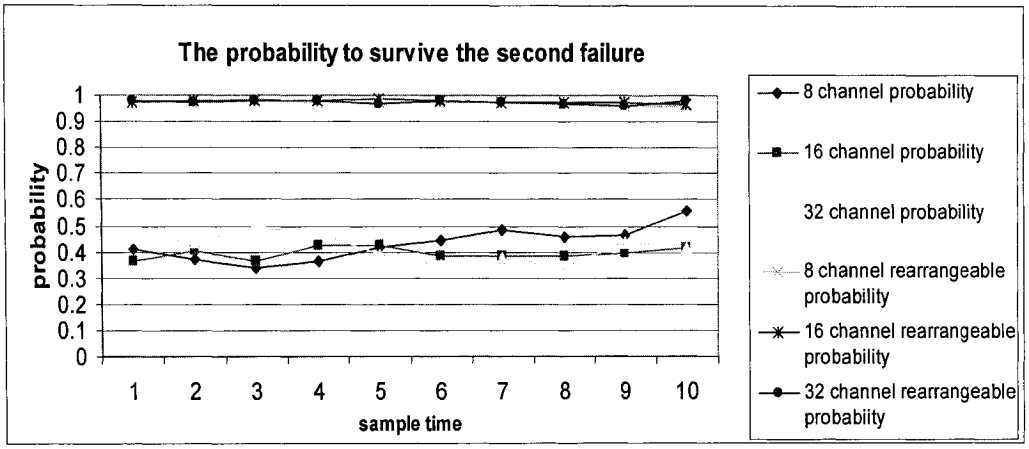
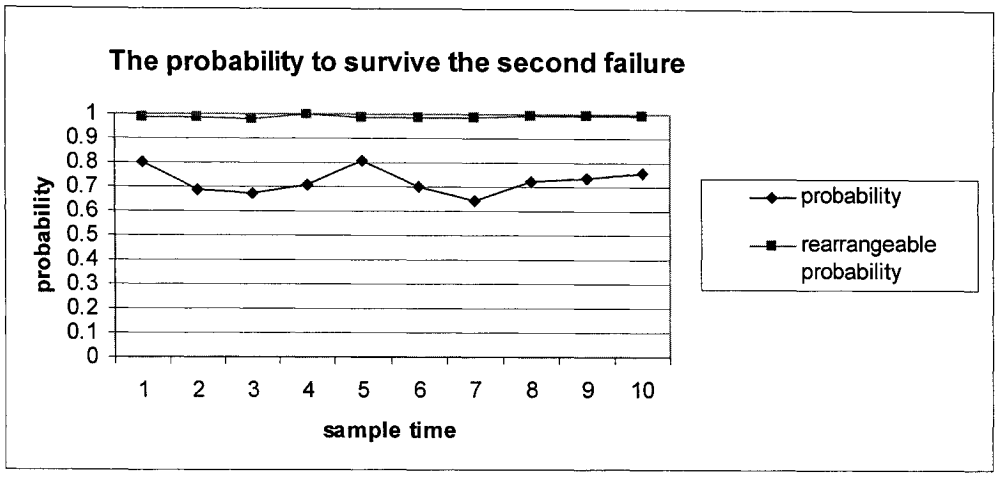


Figure 5-7 Probability for US Long Haul in all 3 cases

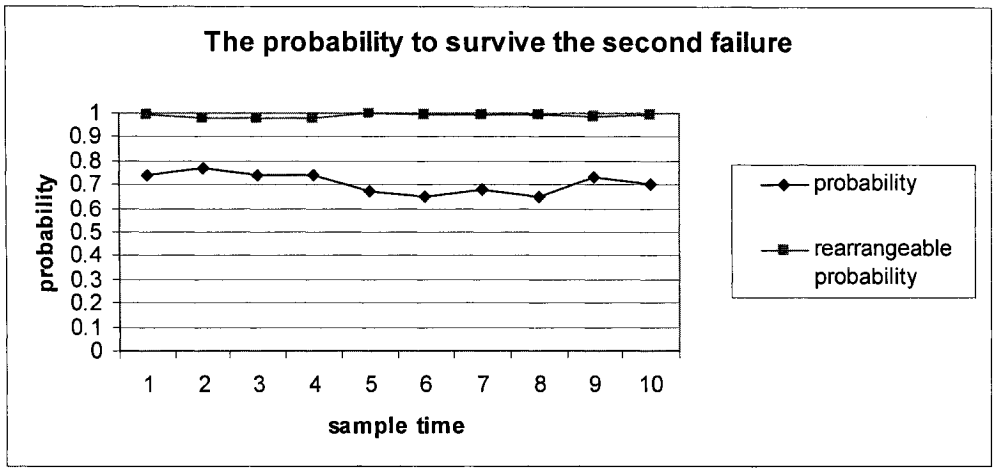
From the graphs we can see that in the US Long Hual network, the second failure can be survived by our MFS scheme with a probability of around 0.4 and 0.97, when using regular MFS and rearrangeable MFS, respectively, without adding up any extra capacity. All graphs show that the rearrangeable MFS has much better performance than the regular one. When the number of working channel changes from 8, 16 to 32, the probability of success remains almost the same. Compared with the NSF network, MFS has better performance in the US Long Hual Network.

5.3.1.3 Data Obtained From the EON Network

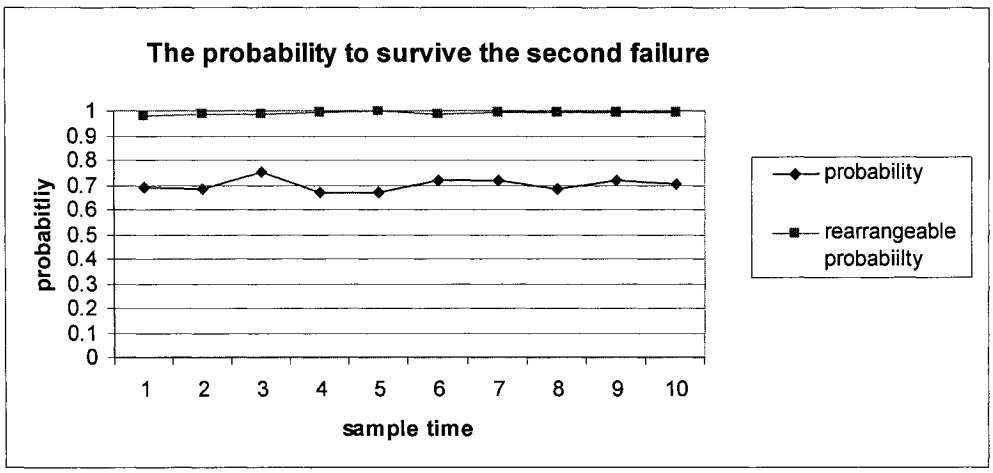
Numerical results of the probability to survive a second failure in the EON Network are shown in Figure 5-8(a)-(c). Figure 5-9 demonstrates the results when all the curves in Figure 5-8(a)-(c) are put inside one graph.



(a) EON Network Probability K=8



(b) EON Network Probability K=16



(c) EON Network Probability K=32

Figure 5-8 Probability of EON Network to survive a second failure

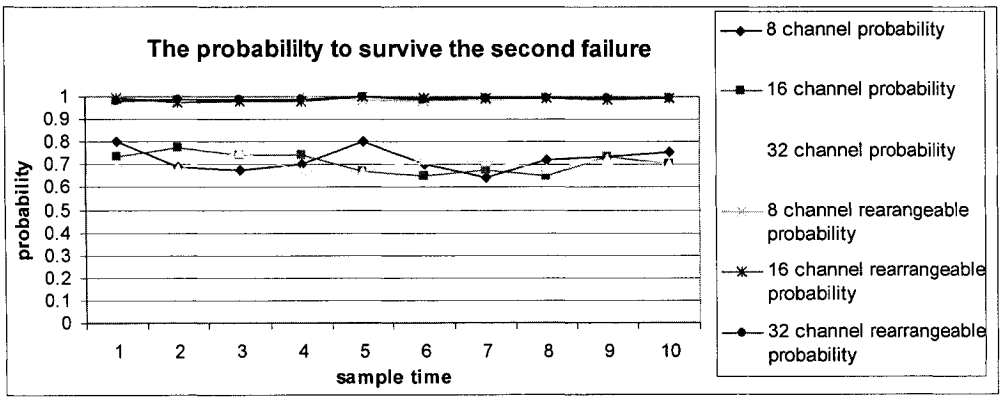


Figure 5-9 Probability for EON Network in all 3 cases

All the graphs illustrate that in the EON network, the second failure can be survived by our MFS scheme with a probability of around 0.75 and 0.98, when using regular MFS and rearrangeable MFS, respectively, without adding up any extra capacity. All graphs show that the rearrangeable MFS has much better performance than the regular one. When the number of working channel changes from 8, 16 to 32, the probability of success remains almost the same. When we compare the performance result with the previous two networks, we can see that, in the EON network, both the regular MFS and the rearrangeable MFS have higher probability to survive a second failure than in the NSF network or US long haul network. The reason of this performance difference is given in section 5.3.2.

5.3.2 Performance Analysis

5.3.2.1 The Regular MFS and the Rearrangeable MFS

From simulation results shown in Figure 5-4(a)-(c), Figure 5-6(a)-(c) and Figure 5-8(a)-(c), one may have the very first impression that the rearrangeable MFS has

accomplished a very good performance in all three test networks, while the efficiency of the regular MFS is relatively lower. The Rearrangeable MFS can achieve 0.76-0.852, 0.96-0.989, 0.975-0.996 probability to survive dual failures in the NSF network, the US Long Haul network and the EON network, respectively, while the regular MFS carries out 0.042-0.223,0.342-0.553,0.651-0.772 probability in these three networks, respectively.

From the regular MFS to the rearrangeable MFS, the probability to succeed is greatly improved. Both MFS schemes are executed at the situation that no extra spare capacity is added into the network after the p-cycle design. The simulation results have proved that our new MFS scheme (both the regular and the rearrangeable version) has acceptable performance to survive double failures, especially when using the rearrangeable one.

5.3.2.2 Different Number of Working Channels

From Figure 5-5, Figure 5-7 and Figure 5-9, we can conclude that in most networks, the difference of working channel number does not have obvious influence on the performance. In all three test networks, the range of success probability remains almost the same when the number of working channel is changing from 8, 16 to 32.

5.3.2.2 Different Network Topologies

All the figures demonstrate that MFS performs quite different when the network topology differs. Figure 5-11 shows the regular MFS probability to survive a second failure, while Figure 5-12 illustrates the rearrangeable MFS probability. From these

two figures, one can easily see the influence of different network topologies to MFS's execution efficiency.

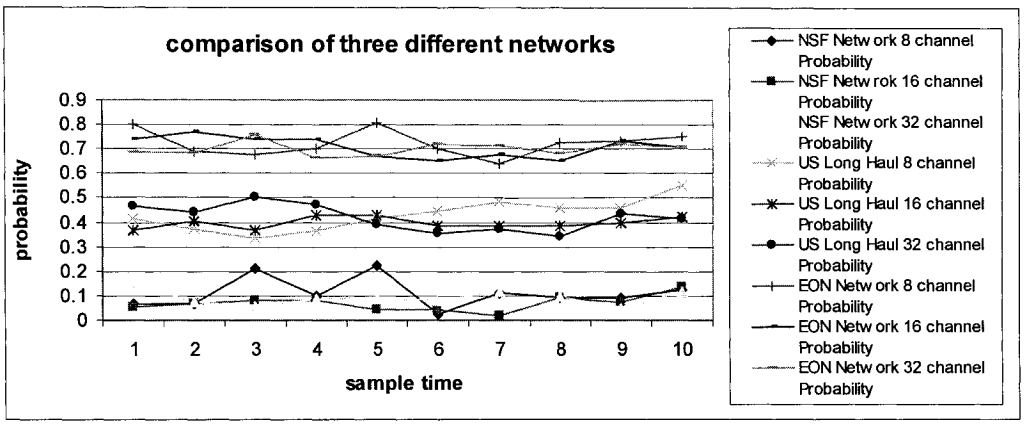


Figure 5-10 Regular MFS Probability Comparison for 3 networks

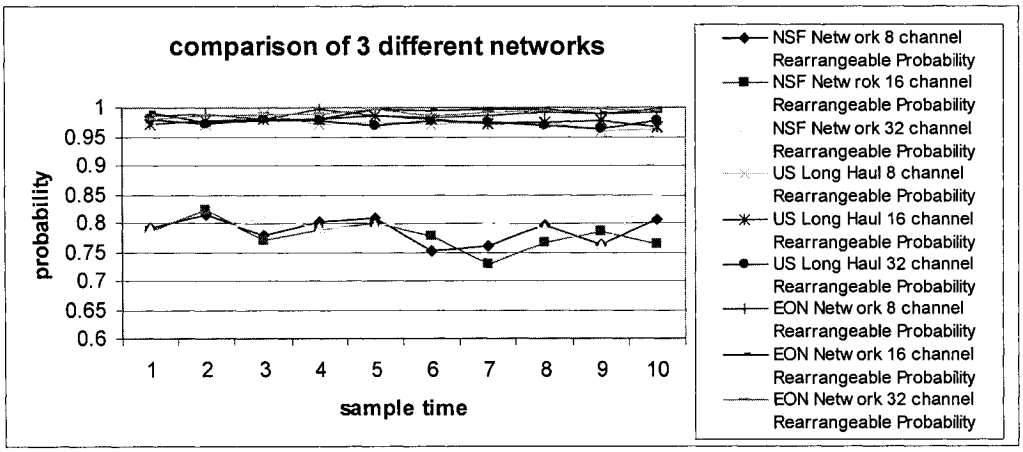


Figure 5-11 Rearrangeable MFS Probability Comparison for 3 networks

From the NSF network, the US Long Haul network to the EON network, the performance of both MFS schemes has a same trend: the probability of success is increasing. For the Regular MFS, it changes from an average of 0.09 in the NFS

network, to an average of 0.4 in the US Long Haul, till reaches 0.72 in the EON network, while the probability of the rearrangeable MFS is from an average of 0.79 in the NFS, to 0.97 in the US Long Haul, and 0.99 in the EON.

The reason of this trend can be explained by looking at Table 5-1. The NSF network has 14 nodes and 21 links with a network degree of 3. The US Long Haul network has 28 nodes and 45 links with a network degree of 3.25. The EON network has 19 nodes and 39 links with the highest degree among those networks, which is 4. The higher a network's degree is, the better MFS's performance will be.

The reason is as follows. When a network has a higher degree, there will be more links connected to each node, and more possible paths between any two nodes, which means more cycles can be found for the links between these two nodes. This determines that there will be more available cycles being found by the OPCA Algorithm and there will be more cycles gathered at each link's PPT, so more options can be found when the system looks up the PPTs. At the end, MFS will have a higher possibility to survive the second failure.

When the network degree increases only a little, the performance of MFS jumps a lot. For example, when the network's degree increases from 3 to 3.25 (compare the NSF network with the US Long Haul network), the probability to recover the second failure jumps from 0.09 to 0.4 when using the regular MFS, and jumps from 0.79 to 0.97 with the rearrangeable MFS. The degree of the EON network is 4. Its

probability to survive dual failures by the regular MFS can already achieve 0.72, and the rearrangeable MFS can reach almost 100% success rate already.

5.3.2.3 Different Number of Links and Nodes in a Network

Another interesting observation is that the difference of the number of nodes and links inside a network does not affect the performance result. The US Long Haul network has more nodes and links than the EON network, but because it has a lower degree, MFS outperforms in the EON than in the US Long Haul.

5.3 Conclusion

In this chapter, we provide detailed simulation results to evaluate the performance of MFS, including the regular MFS and the rearrangeable MFS. Simulation results show that MFS works well with different network topologies. The rearrageable version of MFS has better performance than the regular one. Both versions will have better chance to succeed when the network has higher degree.

Chapter 6 CONCLUSIONS AND FUTURE RESEARCH

6.1 Summary and Concluding Remarks

This thesis provided a complete solution for the protection and restoration using the p-cycle technology in the GMPLS multi-layer architecture, which is the P-cycle Multi-layer Survivability Mechanism (PMSM). A novel MFS multi-failure survivability scheme is also proposed by which the network can recover multiple failures.

To adapt the p-cycle into a multi-layer environment, an SRLG tree structure is used. The SRLG tree concept is first presented by [NASE04]. In this thesis, we use a revised version of SRLG tree, which successfully simplified the SRLG model and reduced its complexity. The Link-Cycle Table (LCT) concept is employed to record all the necessary network resource information.

Compared to any unlayered architecture, this multi-layer structure has many advantages, such as bandwidth saving, better granularity, single layer function's simplification and the ability to form more powerful networks.

Based on this multi-layer network structure, we proposed the Multi-Failure Survivability Scheme (MFS) to provide multiple failure survivability. MFS is composed of two stages: (1) off-line calculation procedure, and (2) on-line selection procedure. At the off-line calculation stage, a new algorithm OPCA is used to form a P-Cycle Pool Table (PPT) for every link. Cycles in the PPT are chosen in the top-down sequence when the link fails. The Rearrangeable MFS is also provided, which can extraordinarily improve the performance of MFS.

MFS has been proved to have very good efficiency by our simulation results. In a sparse network, such as the NFS network, the rearrangeable MFS can survive more than 70% of dual failures. Once the degree of the network increases, the probability of success also increases. In the US Long Haul network (with a degree of 3.25), the success probability of regular MFS and rearrangeable MFS becomes 0.4 and 0.97 respectively. In the EON network (with a degree of 4), 72% dual failures can be recovered by the regular MFS and 98% by the rearrangeable MFS. This means in a dense network, most of the second failure can be recovered by the rearrangeable MFS.

MFS also offers networks a potential to provide the Class of Protection, which gives networks the power to provide better service.

MFS is efficient, because it is executed with no extra capacity cost. It is implemented after the regular p-cycle design without adding any extra spare capacity. It can also be used to survive the third or the fourth failure, or even more, by using the same algorithm.

6.2 Future Research

In this thesis, we provided a mechanism named PMSM to adapt the p-cycle concept to GMPLS networks. A new scheme called MFS (Multi-Failure Survivability Scheme) is also proposed to survive the network from multiple failures. As an extension to the work we have carried out in this thesis, the future research work can be summarized as follows:

- In a multilayered network, the coordination between different layers is very important. How can different layers properly communicate with each other? A new multi-layer protocol is needed to solve this problem.
- In this thesis, we provide algorithms and simulation results with the object of "maximize the probability to survive the second failure without adding up any extra network capacity". Algorithms and simulations can also be employed by using another object: "minimize the extra network capacity while realizing 100% dual failure survivability". New algorithms are needed to solve the problem about where and how to add the extra network capacity.

- In this thesis, we implement simulations to calculate the probability for MFS to survive dual failures. Implementation can also be carried out to simulate MFS's performance when there are three, four or even more failures at the same time.
- PMSM and MFS can be executed in any kind of networks. In this thesis, we only conducted algorithms and simulations for general network designs, without any specification on network type. If they are used inside an optical network, the assumption is that all the nodes have full wavelength conversion ability. If one wants to use them specifically in an optical network that not all nodes can do wavelength conversion, the wavelength continuity constraint has to be taken into consideration. The OPCA algorithm has to be revised to add up the wavelength continuity constraint. We left this interesting topic to our future research work.

References:

[BANE01a] A. Banerjee, J. Drake, J. P. Lang, B. Turner, K. Kompella and Y. Rekhter, "Generalized multiprotocol label switching: an overview of routing and management enhancements," IEEE Communications Magazine, vol. 39, no. 1, pp. 144 – 150, Jan. 2001.

[BANE01b] A. Banerjee, J. Drake, J. P. Lang, B. Turner, D. Awduche, L. Berger, K. Kompella and Y. Rekhter, "Generalized multiprotocol label switching: an overview of signaling enhancements and recovery techniques," IEEE Communications Magazine, vol. 39, no. 7, pp. 144 – 151, July 2001.

[BLOU03] J. Blouin, S. Anthony, W. D. Grover and N. Hadi, "Benefits of p-cycles in a mixed protection and restoration approach," Proceedings of DRCN 2003, Banff, Canada, pp. 203-211, Oct. 2003.

[BROW94] G. N. Brown, W. D. Grover, J. B. Slevinsky and M. H. MacGregor, "Mesh.Arc networking: an architecture for efficient survivable self-healing networks," Proceedings of ICC 1994, pp. 471 - 477, Mainz, Germany, May 1994.

[CHAL03] S. Chalasani and V. Rajaravivarma, "Survivability in Optical Networks," Proceedings of the 35th Southeastern Symposium on System Theory 2003, pp. 6 – 10, March 2003.

[CHAN05] G. K. Chang and J. Yu, "Multirate payload switching using a swappable optical carrier suppressed label in a packet-switched DWDM optical

network," IEEE Journal of Lightwave Technology, vol. 23, no. 1, pp. 196 – 202, Jan. 2005.

[DAVI03] W. G. David, "IETF Work on Protection and Restoration for Optical Networks," Optical Networks Magazine, July/August 2003.

[DOUC01] J. Doucette and W. D. Grover, "Comparison of Mesh Protection and Restoration Schemes and the Dependency on Graph Connectivity," Proceedings of DRCN 2001, Budapest, Hungary, pp. 121 - 128, Oct. 2001.

[DOUC03a] J. Doucette, D. He, W. D. Grover and O. Yang, "Algorithmic approaches for efficient enumeration of candidate p-cycles and capacitated p-cycle network design", Proceedings of DRCN 2003, Banff, Canada, pp. 212 – 220, Oct. 2003.

[DOUC03b] J. Doucette and W .D. Grover, "Node-Inclusive Span Survivability in an Optical Mesh Transport Network," Proceedings of NFOEC 2003, Orlando, USA, pp. 634 - 643, Sept. 2003.

[GROV90] W. D. Grover, "Method and apparatus for self-healing and self provisioning networks," U.S. Patent No. 4,956,835, 1990.

[GROV91] W. D. Grover and B. D. Venables, "Development and performance verification of a distribured asynchronous protocol for real-time network restoration," IEEE Journal on Selected Areas in Communications 1991, pp. 112 - 125, vol. 9, no. 1, Jan. 1991.

[GROV92] W. D. Grover, "Case studies of survivable ring, mesh, and mesh-arc hybrid networks," Proceedings of IEEE Globecom 1992, Orlando,

Florida, pp. 633 - 638, Dec.1992.

[GROV98] W. D. Grover and D. Stamatelakis, "Cycle-oriented distributed preconfiguration: ring-like speed with mesh-like capacity for self-planning network restoration," Proceedings of ICC 1998, Atlanta, USA, pp. 537 – 543, vol. 1, June 1998.

[GROV00] W. D. Grover and D. Stamatelakis, "Bridging the ring-mesh dichotomy with p-cycles," Proceedings of DRCN 2000, Munich, Germany, pp. 92 – 104, April 2000.

[GROV02a] W. D. Grover, J. Doucette, M. Clouqueur and D. Stamatelakis, "New options and Insight for survivable transport networks," IEEE Communications Magazine, vol. 40, no. 1, pp. 34 – 41, Jan. 2002.

[GROV02b] W .D. Grover and J. Doucette, "Advances in optical network design with p-cycles: Joint optimization and pre-selection of candidate p-cycles," IEEE/LEOS Summer Topi 2002, pp. WA2-49 - WA2-50, July 2002.

[GUMA05] A. Gumaste and S. Q. Zheng, "Protection and restoration scheme for light-trail WDM ring networks," Proceedings of ONDM 2005, Milan, Italy, pp. 311 – 320, Feb. 2005.

[GUN03] K. D. Gun, K. C. Hee and K. K. Young , "Multiple Hierarchical Protection Schemes for Differentiated Services in GMPLS networks," Proceedings of ConTEL 2003, Zagreb, Croatia, vol.2, pp. 661 - 664, June 2003.

[GXUE03a] G. Xue and R. Gottapu, "Finding protection cycles in DWDM

networks," Proceedings of HPSR 2003, Torino, Italy, pp. 305 – 309, June 2003.

[GXUE03b] G. Xue and R. Gottapu, "Efficient construction of virtual p-cycles protecting all cycle-protectable working links," Proceedings of HPSR 2003, Torino, Italy, pp. 305 – 309, June 2003.

[HERZ94] M. Herzberg and S. J. Bye, "An Optimal Spare-Capacity Assignment Model for Survivable Network with Hop Limits," Proceedings of GLOBECOM 1994, San Francisco, USA, vol. 3, pp. 1601 – 1606, Dec.1994.

[HWAN05] H. Wang and H. T. Mouftah, "P-cycles in Multi-layer and Multi-failure Network Survivabiligy," Proceedings of ICTON 2005, Barcelona, Spain, July 2005.

[JAIS05] A. Jaiszczyk, "Automatically switched optical networks: benefits and requirements," IEEE Communications Magazine, vol. 43, no. 2, pp. S10 - S15, Feb. 2005.

[KANG03] J. Kang and J. R. Martin, "Bandwidth Protection in MPLS Networks Using p-Cycles Structure," Proceedings of DRCN 2003, Banff, Canada, Oct. 2003.

[KERI05] H. Kerivin, D. Nace and T. Pham, "Design of Capacitated Survivable Networks With a Single Facility," IEEE/ACM Transactions on Networking, vol. 13, no. 2, pp. 248–261, April 2005.

[MAED04] W. Mardini, O. Yang and G. Q. Wang, "Mapping p-cycles from planar graph to cycle graphs," Proceedings of CISS 2004, Princeton, New

Jersey, pp. 807 - 812, Mar. 2004.

[MURA97] K. Murakami and H. S. Kim, "Comparative study on restoration schemes of survivable ATM networks," Proceedings of INFOCOM 1997, Kobe, Japan, vol. 1, pp. 345 - 352, April 1997.

[NASE04] H. Naser and H. T. Mouftah, "A multilayer differentiated protection services architecture," IEEE Journal on Selected Areas in Communications, vol. 22, no. 8, pp. 1539 – 1547, Oct. 2004.

[OKI05] E. Oki, K. Shiomoto, D. Shimazaki, N. Yamanaka, W. Imajuku and Y. Takigawa, "Dynamic multilayer routing schemes in GMPLS- based IP+optical networks," IEEE Communications Magazine, vol. 43, no. 1, pp. 108 - 114, Jan. 2005

[PERE05] M. Perenyi, J. Breuer, T. Cinkler and C. Gaspar, "Grooming node placement in switched multilayer networks," Proceedings of ONDM 2005, Milan, Italy, pp. 413 – 419, Feb. 2005.

[RAMA99a] S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks. Part I-Protection," Proceedings of INFOCOM 1999, New York, USA, vol. 2, pp. 744 – 751, March 1999.

[RAMA99b] S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks. II. Restoration," Proceedings of ICC 1999, Vancouver, Canada, vol. 3, pp. 2023 – 2030, June 1999.

[RAMA03] S. Ramamurthy, L. Sahasrabuddhe and B. Mukherjee, "Survivable WDM mesh networks," Journal of Lightwave Technology, vol. 21, no. 4, pp. 870 – 883, April 2003.

[ROBE03] S. Roberto, S. Marina, O. Gianpaolo, R. Federica and F. Fabio, "A Multilayer Solution for Path Provisioning in New-Geneartion Optical/MPLS Networks," Journal of Lightwave Technology, vol. 21, no. 5, May 2003.

[SCHU02a] D. A. Schupke, C. G. Gruber and A. Autenrieth, "Optimal configuration of p-cycles in WDM networks," Proceedings of ICC 2002, New York, USA, vol. 5, pp. 2761 – 2765, April/May 2002.

[SCHU02b] D. A. Schupke, "Fast and efficient WDM network protection using p-cycles," 2002 IEEE/LEOS Summer Topi , pp. WA1-47 – WA1-48, July 2002.

[SCHU03a] D. A. Schupke, "Multiple failure survivability in WDM networks with p-cycles," Proceedings of ISCAS 2003, Bangkok, Thailand, vol. 3, pp. III-866–III-869, May 2003.

[SCHU03b] D. A. Schupke, "The tradeoff between the number of deployed p-cycles and the survivability to dual fiber duct failures," Proceedings of ICC 2003, Ottawa, Canada, 11-15 vol. 2, pp. 1428 – 1432, May 2003.

[SCHU03c] D. A. Schupke and M. C. Scheffel, "Configuration of p-cycles in WDM Networks with Partial Wavelength Conversion," Photonic Network 2003, vol. 6, no. 3, pp. 239 - 252, Nov. 2003.

[SCHU04] D. A. Schupke, W. D. Grover and M. Clouqueur, "Strategies for Enhanced Dual Failure Restorability with Static or Reconfigurable p-Cycle Networks," Proceedings of ICC 2004, Madison, USA, vol. 3, pp. 1628 – 1633, June 2004.

[SCHU05] D. A. Schupke, "On Hamiltonian cycles as optimal p-cycles", IEEE Communications Letters, vol. 9, no. 4, pp. 360 - 362, April 2005.

[SHEN03] G. Shen and W. D. Grover, "Extending the p-cycle concept to path segment protection for span and node failure recovery," IEEE Journal on Selected Areas in Communications, vol. 21, no. 8, pp. 1306 – 1319, Oct. 2003.

[STAM00a] D. Stamatelakis and W. D. Grover, "IP layer restoration and network planning based on virtual protection cycles," IEEE Journal on Selected Areas in Communications, vol. 18, no. 10, pp. 1938 - 1949, Oct. 2000.

[STAM00b] D. Stamatelakis and W. D. Grover, "Theoretical underpinnings for the efficiency of restorable networks using preconfigured cycles ("p-cycles"),"IEEE Transactions on Communications, vol. 48, no.8, pp. 1262 – 1265, Aug. 2000.

[SUKY02] L. SuKyong, K. Chul and G. David, "Hierarchical Restoration Scheme for Multiple Failures in GMPLS Networks," Proceedings of ICPPW 2002, Vancouver, Canada, pp. 177 – 182, Aug. 2002.

[VARG05] J. C. Varghese, A. Dutta, A. Cheng, D. Chee, M. Elaoud, T. McAuley, I. Sebuktekin, B. Kim and K. D. Wong, "Integrated networking technologies for a survivable network," Proceedings of WCNC 2005, New Orleans, USA, vol. 4, pp. 2424–2429, March 2005.

[WASH05] A. N. Washington, C. C. Hsu, H. Perros and M. Devetsikiotis, "Approximation Techniques for the Analysis of Large Traffic-Groomed Tandem Optical Networks," Proceedings of Simulation Symposium

2005, San Diego, USA, pp. 15 - 22, April 2005.

[YANG05] X. Yang, L. Shen and B. Ramamurthy, "Survivable Lightpath Provisioning in WDM Mesh Networks Under Shared Path Protection and Signal Quality Constraints," Journal of Lightwave Technology, vol. 23, no. 4, pp. 1556 – 1567, April 2005.

[ZHAN02a] H. Zhang and O. Yang, "Finding protection cycles in DWDM networks," Proceedings of ICC 2002, New York, USA, vol. 5, pp. 2756 – 2760, April/May 2002.

[ZHAN02b] H. Zhang and A. Durresi, "Differentiated multi-layer survivability in IP/WDM networks," Proceedings of NOMS 2002, Florence, Italy, pp. 681 – 694, April 2002.

[ZHAO03] J. Zhao, L. Lei, Y. Ji and D. Xu, "Integrated Multilayer Survivability Strategy with Inter-Layer Signaling," Proceedings of ICCT 2003, Beijing, China, vol. 1, pp. 612 - 616, April 2003.

[ZHOU00] D. Y. Zhou and S. Suresh, "Survivability in Optical Networks", IEEE Network, vol. 14, no. 6, pp.16 – 23, Nov/Dec. 2000.