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

Mazen George Khair
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Electrical Engineering)
GRADE / DEGREE

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

Fault Localization for Next Generation Survivable All-Optical Networks

TITRE DE LA THÈSE / TITLE OF THESIS

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

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

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Martin Maier

Dorina Petriu

Amiya Nayak

Shervin Shirmohammadi

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

Fault Localization for Next Generation Survivable All-Optical Networks

Mazen George Khair

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

Ottawa-Carleton Institute of Electrical and Computer Engineering
School of Information Technology and Engineering
University Of Ottawa
Ottawa, Ontario, Canada

uOttawa

L'Université canadienne
Canada's university

December, 2008

Your file Votre référence
ISBN: 978-0-494-48657-3
Our file Notre référence
ISBN: 978-0-494-48657-3

"For God so loved the World that He gave us His only begotten Son, that whoever believes in Him should not perish but have ever lasting Life." John 3:16

Abstract

We propose a distributed fault localization protocol, called Parallel Limited Perimeter Vector Matching (P-LVM) protocol, for localizing multi-link failures in all-optical networks. The P-LVM protocol is based on the Limited Perimeter Vector Matching (LVM) protocol for localizing single-link failures. To handle multi-link failures, it tries to separate each failure in a small perimeter area after identifying each perimeter area with its corresponding failure and then localize the failures in parallel respectively in a distributed manner. Through simulation results, we show that the P-LVM protocol can effectively localize multi-link failures that occur simultaneously or at different times.

We also study the optimization problems in applying the LVM protocol in static all-optical networks. We consider two optimization problems: one is to optimize the traffic distribution so that the fault localization probability in terms of the number of localized links is maximized and the other is to optimize the traffic distribution so that the time for localizing a failed link is minimized. We formulate the two problems into an integer linear programming problem, respectively, and use the CPLEX optimization tool to solve the problems. By optimizing the traffic distribution the fault localization probability can be maximized and the fault localization time can be minimized. Moreover, a heuristic algorithm is proposed to evaluate the optimization results through simulation experiments.

Finally, we propose two different fault localization protocols for localizing single-link failures in multi-domain all-optical networks. These two protocols are based on LVM protocol mechanism, which restricts fault localization within a smaller perimeter area and can thus significantly reduce the time and space complexity for fault localization. The first protocol assumes the existence of power monitoring only at edge nodes, whereas the second protocol assumes no power monitors at the intermediate nodes or the edge nodes. The two protocols can localize both inter-domain and intra-domain link failures without exchanging any internal confidential domain-specific information between different domains. The analytical results show that both protocols can not only fast localize an inter-domain link failure between different domains but also localize an intra-domain link failure that affects inter-domain traffic faster than the Open-Shortest-Path-First (OSPF) protocol in a large network.

Acknowledgment

The following thesis, while an individual work, benefited from the insights and direction of several people. First, my Thesis Supervisor, professor Hussein T. Mouftah, exemplifies the high quality scholarship to which I aspire. In addition, professor Hussein T. Mouftah provided timely and instructive comments and evaluation at every stage of the thesis process, allowing me to complete this project on schedule.

Next, I would like to thank Dr. Jun Zheng for his insights that guided and challenged my thinking, substantially improving the finished work.

I also would like to thank my dearest friend Dr. Burak Kantarci, he was always there for me, he demonstrated a great example to me of how two different people from different background could achieve so much within a very short period of time. I am greatfull for the luck that brought us together.

In addition to the technical and instrumental assistance above, I received equally important assistance from family and friends. My parents provided on-going support throughout the thesis process, as well as technical assistance critical for completing the project in a timely manner. My parents instilled in me, from an early age, the desire and skills to obtain higher education. I am greatfull for their support.

I wish to thank my twin brother Muhannad, his wife Sahira and their kid Julian, I also want to thank my sister Sawsan, my brother in-law George and their kids Daniel, Marianne and Mathias for their unconditional support during the period of this work.

Finally, I would like to thank my friends Saed Hawwassh, Nizar Awwad, Botros Dalou and Wail Mardini for being always there when I needed them.

Table of Content

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Tables

Chapter 4

Acronyms

ALV: Affected-Link-Vector
AON: All-Optical Networks
B&B: Branch-and-Bound
BDI: Backward Defect Indicator
BER: Bit Error Rate
BGP: Border Gateway Protocol
CI: Confidence Interval
CP: Core Phase
CID: Channel Identifier
DP: Diagnosis Phase
DWDM: Dense Wavelength Division Multiplexing
E-BGP: External BGP
ER: Executive Route
ERQ: Executive Route Queue
ES: Executive Sink
FDI: Forward Defect Indicator
FF: First Fit
FLA: Fault Location Algorithm
HDFS: Heuristic Depth first Searching
I-ALV: Inter-domain Affected-Link-Vector
I-BGP: Internal BGP
I-ER: Inter-domain Executive Route
ILP: Integer Linear Programming
ISP: Internet Service Provider
ITU: International Telecommunication Union
LVP: Limited Perimeter Vector Matching Fault Localization Protocol
MEPP: Monitoring Equipment Placement Phase
OA: Optical Amplifiers
OCh: Optical Channel
OD: Optical Domain
OEO: Optical-Electronic-Optical
OPM: Optical Power Monitor
ORBGP: Optical Routing Border Gateway Protocol
OSA: Optical spectrum analyzer
OSC: Optical Supervisory Channel
OSNR: Optical Signal-to-Noise Ratio
OSPF: Open Shortest Path First
OXC: Optical crossconnects
PCP: Pre-Computational Phase
PES: Potential Executive Sink
P-LVM: Parallel Limited Perimeter Vector Matching protocol
PV: Path Vector
RTT: Round-Trip Time
RWA: Routing and Wavelength Assignment

STS-1: Synchronous Transport Signal 1
SONET: Synchronous Optical Networks
SPE: Synchronous Payload Envelope
SPEM: Shortest Path Eulerian Matching
SPN: Shared Perimeter Node
TFLA: Transparent Failure Location Algorithm
TI: Trace Identifier
vCCCP: Variant Constrained Cycle-Cover Problem
VoIP: Voice over Internet Protocol

Chapter 1 Introduction

1.1 Introduction

In the past decade, the explosive growth of the Internet and the emergence of various broadband applications, such as Internet telephony, video conferencing, video on demand, and interactive gaming, have imposed a tremendous demand for bandwidth capacity on the underlying telecommunications infrastructure. To meet the unprecedented demand for bandwidth capacity, telecommunications networks have experienced a bandwidth revolution by introducing fiber-optic technology. As the core of this revolution, optical fibers have proved to be an excellent physical transmission medium for providing huge bandwidth capacity. However, because of the limit of the electronic processing speed, it is impossible that all the bandwidth of an optical fiber will be exploited by using a single high-capacity optical channel or wavelength. The emergence of wavelength division multiplexing (WDM) technology has provided a practical solution to meeting this challenge. With WDM technology, multiple optical signals can be transmitted simultaneously and independently in different optical channels over a single fiber, each at a rate of a few gigabits per second, which significantly increases the usable bandwidth of an optical fiber [ZM04, Za03]

In an optical network, a fiber provides a number of optical channels or wavelengths to carry data traffic, each operating at a very high rate of several gigabits per second. A single network failure such as a fiber cut may cause the disruption of all the lightpaths that traverse the failed fiber link and can therefore lead to a large amount of data loss in the network. This would largely degrade and even disrupt the network services. To guarantee network services, a network must provide effective survivability capabilities to survive different types of network failures (e.g., a fiber cut or a node fault) [MH03, TS02].

In general, there are two basic survivability paradigms for an optical network: static protection and dynamic restoration. In static protection, spare network resources are

reserved at the time of network design or connection establishment for protection against network failures. In the event of a network failure, the disrupted network services are recovered by using the reserved network resources. Obviously, static protection is fast in service recovery as the reserved network resources are dedicated to network failures. However, it is inefficient in resource utilization because the reserved network resources cannot be used for other connections. In dynamic restoration, no spare resources are reserved in advance for restoration purposes. The network must dynamically discover spare network resources available in the network to recover the disrupted network services after network failures occur. Obviously, dynamic restoration is efficient in resource utilization but slow in service recovery [OM05, MG02].

1.2 Motivation

In the event of a network failure, an optical network must first detect and localize the failure before applying a static protection or dynamic restoration scheme. Traditionally, this is implemented by equipping optical power monitoring (OPM) at each node [KBB04], which allows the node to detect the optical signal loss and thus detect the failure of a lightpath. However, this would largely increase the cost of each network node. In Synchronous Optical Networks (SONETs), optical-electronic-optical (OEO) regenerators are used to detect the loss of light by monitoring high BER at an intermediate node.

On the other hand, most existing work [Yan02, RDB00] employs the same approaches used in traditional IP networks. This implies that each network element in all-optical networks, such as optical amplifiers and optical add–drop multiplexers (OADMs), is equipped with monitoring devices that are capable of reporting any signal degradation at any time. Fortunately, reporting signal degradation is very efficient when the network size is relatively small or since the signal is being repeated at every regeneration point. However, this does not work for all-optical networks. More efficient and less-costly approaches are needed especially when the size of the network grows significantly.

For all-optical networks, it is extremely important to reduce the network complexity and cost for the following reasons:

- The number of monitoring devices, such as optical power meter (OPM), optical spectrum analyzer (OSA), eye monitor, Bit Error Rate (BER) monitor, and Wavemeter which is designed to report any degradation in the signal continuously. The continuous flood of the huge amount of measured data from the monitoring devices results in a scalability issue as the network size grows. Furthermore, the tuning agility of future all-optical networks devices becomes a big challenge.
- Because of the fact that each monitoring device is testing specific component without considering its failure statistics, the continuous measure every unit time degrades the QoS specifications resulting in high cost and limiting the effectiveness of deploying all-optical networks. More importantly, the OPM not only requires proper placement but also it does not guarantee link localization [SS02]. Therefore, there should be a scheme that is used to localize the failure in the absence of OPM at the intermediate nodes [SM07]. Issues such as multi-fault localization, fast localization, increasing the number of localized links and inter-domain fault localization have to be addressed.

Protection and restoration in optical networks require that lightpaths are established and turn down in a dynamic way. Therefore, dynamic restoration is needed. The major advantage of dynamic restoration is that it has the ability to discover the available network resources and assigning them to affected traffic. This makes dynamic restoration protocols efficient in nature because they try to use network resources efficiently. Link restoration is the most efficient dynamic restoration mechanism because it tries to establish the shortest possible loop around the failed link. Link restoration is a single-hop point-to-point compared to path restoration, which is an end-to-end recovery scheme. Since link restoration tries to set-up the shortest possible loop around the failed link using a smaller number of links than path restoration. As a result, link restoration reduces the number of switches that need to be reconfigured. Furthermore, it requires less restoration time and waste less bandwidth. Furthermore, link restoration keeps the network running the way it is, preserving the network load balancing and hence avoiding any sudden chaos in the network. Therefore, link localization protocols are needed to feed up the protection scheme with the location of the failed link to achieve fast recovery.

LVM [Sic06] assumes that there is only one failure, which is not always the case. In many cases, multiple failures may occur in an all-optical network. However, multi-failure localization has not been widely studied for all-optical networks, which motivated us to propose a multi-failure localization protocol for all-optical networks. The protocol should be able to handle multi-link failures; also, it should try to separate each failure in a small perimeter area after identifying each perimeter area with its corresponding failure and then localize them in parallel in a distributed manner.

Successful fault localization for a given link depends on the traffic distribution. The Limited Perimeter Vector Matching (LVM) protocol was proposed for all-optical networks [Sic06], which can effectively localize a link failure if there are at least two lightpaths passing through the failed link, which belong to two different s-d pairs. However, if there are less than two lightpaths passing through the failed link, it is unable to localize the failure. We propose to study maximizing the number of localized links by providing a mathematical model to optimally distribute the traffic (or lightpaths) in the network so that in the event of a link failure the fault localization probability is maximized. We also plan to analyze the delay of the LVM protocol. We are interested in the problem to minimize the fault localization time when applying the LVM protocol to a static all-optical network. This can be achieved through a mathematical model that tries to optimize the traffic distribution in the network in order to minimize the fault localization time when applying the LVM.

To evaluate our optimization models, a heuristic approach should be used to evaluate our optimization models explained above. This heuristic approach attempts to satisfy the LVM protocol constraints by assigning the appropriate costs to the links to allow different SD pairs to utilize the same link. We show that greedy and non-greedy implementations of the heuristic approach provide link coverage bounded above by the optimization results. It also tries to distribute the traffic to minimize time localization at each link.

Fault localization can also be applied to fault localization in inter-domain all-optical networks. In this scenario, it becomes more difficult to localize a failure in a multi-domain scenario because the internal confidential domain-specific information (e.g., topology and traffic) in a domain is not allowed to be available to other domains. We will study fault localization in multi-domain all-optical networks with power monitoring available only at edge nodes. We are interested in localizing both inter-domain and intra-domain link failures

without exchanging any internal confidential domain-specific information between different domains. Furthermore, we will study fault localization for localizing single-link failures in multi-domain all-optical networks with no optical power monitoring devices available at neither intermediate nodes nor edge nods.

1.3 Thesis Objective

This thesis work studies the fault-localization issue in all optical networks. The objective of this thesis work includes the following aspects:

- To design a distributed fault-localization protocol, called Parallel Limited Perimeter Vector Matching (P-LVM), for localizing multi-link failures in all-optical networks based on the LVM protocol for localizing single-link failures in all-optical networks. The P-LVM protocol should be able to efficiently localize multi-link failures that occur simultaneously or at different times in all-optical networks.
- To formulate and solve the fault-localization optimization problem in applying the LVM protocol to static all-optical networks. Given the traffic demands, the fault-localization optimization problem is to optimize the traffic distribution so that the fault-localization probability in terms of the number of localized links can be maximized.
- To formulate and solve the fault-localization time optimization problem when applying the LVM protocol to static all-optical networks. Given the traffic demands, the fault-localization time optimization problem is to find a traffic distribution so that the time for localizing a failed link can be minimized.
- To present a heuristic algorithm to evaluate the above two optimization formulations, which performs wavelength assignment in a way that meets the requirements of the LVM protocol.
- To design a distributed fault-localization protocol for localizing single-link failures in multi-domain all-optical networks with power monitoring available only at edge nodes based on the LVM protocol.

- To design a distributed fault-localization protocol for localizing single-link failures in multi-domain all-optical networks with no power monitoring available at each edge node based on the LVM protocol.

1.4 List of Contributions

The contributions of this thesis work include the following six aspects following is the thesis contribution:

1. Proposed a distributed fault-localization protocol, called Parallel Limited Perimeter Vector Matching (P-LVM) protocol, for localizing multi-link failures in all-optical networks.
2. Studied the fault-localization optimization problem when applying the LVM protocol to static all-optical networks. This optimization problem is to optimize the traffic distribution so that the fault-localization probability in terms of the number of localized links can be maximized. We formulated the problem into an integer linear programming problem and use CPLEX to solve the problem.
3. Studied the fault-localization time optimization problem when applying the LVM protocol to static all-optical networks. This optimization problem is to find a traffic distribution so that the time for localizing a failed link can be minimized. We formulated the problem into an integer linear programming problem and use CPLEX to solve the problem.
4. Proposed a heuristic algorithm to evaluate the proposed optimization formulations, which performs wavelength assignment in a way that meets the requirements of the LVM protocol.
5. Designed a distributed fault-localization protocol for localizing single-link failures in multi-domain all-optical networks. This protocol is based on the LVM protocol. By assuming power monitoring available only at edge nodes, this protocol can localize both inter-domain and intra-domain link failures without exchanging any internal confidential domain-specific information between different domains.
6. Designed a distributed fault-localization protocol for localizing single-link failures in all-optical multi-domain networks. The protocol is based on the LVM protocol. By

performing a fault-localization process sequentially in the optical domains that an affected lightpath passes through, the protocol can localize both inter-domain and intra-domain link failures that affect inter-domain traffic without exchanging any internal confidential domain-specific information between different domains.

1.5 Thesis Outline

This thesis is organized in six chapters. In Chapter 2, we present a survey of all inter-domain and intra-domain localization techniques. In chapter 3, we describe all possible cases in which different combinations of multi-failures may occur in an all-optical network, introduce the proposed P-LVM protocol and its characteristics, and evaluate the performance of the protocol through simulation results. In chapter 4, we first give a brief introduction of the LVM protocol and review related work. Then we describe the integer linear programming formulations of the two optimization problems, respectively, and present the proposed heuristic algorithm. Finally, we evaluate the optimization results by comparing the results with those obtained using the proposed heuristic. In chapter 5, we present two different fault-localization protocols for localizing single-link failures in multi-domain all-optical networks and evaluate their performances through simulation results. In Chapter 6, we conclude this thesis and briefly provides highlights for future work.

Chapter 2 Inter and Intra domain localization protocols

2.1 Introduction

This chapter provides a survey about detection and localization protocols for survivable optical networks that guarantee service connectivity in the presence of a failure. We survey detection and localization technique in both intra- domain and inter-domain. This survey also investigates the existence of any Multi-fault localization protocols.

2.2 Protection Versus Restoration

Before we proceed with the description of localization protocols, we first introduce some basic concepts on protection and restoration, which are useful to the subsequent sections.
In general, there are two major groups for survivable techniques. The first one is known as *pre-planned protection* whereas the other is known as *dynamic restoration* [Muk06].

- Pre-planed protection in which protection resources are specified at the network planning stage and reserved to protect the network when a failure occurs [Do06]. Unfortunately, this approach does not make the best use of the available bandwidth as some of the resources and components are kept in a stand by stage and only used for protection when a failure occurred. On the other hand, pre-planed protection technique has a quick recovery time. Pre-planned protection can be further divided into two kinds:
    - The first one is known as *dedicated* in which a second copy of the data is transmitted on both protection and working path at the same time. This

resulted in reducing the network bandwidth to half. The nice thing about this protection type is that service recovery is 100% guaranteed [TP04].

- The second one is known as *shared protection* in which a second copy of the data is sent on the protection path once a failure has been detected. This allows better bandwidth utilization as the protection path can be used by a group of working paths to transfer low priority data on the absent of a failure. Once a failure is detected, this protection path will be evacuated which means that lots of data has to be buffered at the intermediate nodes. As it is well known, buffering means the need for O-E-O to convert the data to electronic domain to be able to store it as buffering in the optical domain does not exist yet. Therefore, this type of protection is not desired [IG00].

- Dynamic restoration comes as an alternative solution to overcome the weak points defined in pre-planned protection techniques. The major advantage of dynamic restoration is that it has the ability to discover the available network resources and assigning them to affected traffic. This makes dynamic restoration protocol efficient in nature as these protocol try to use network resources efficiently, unlike pre-planed technique which requires setting aside or adding protection resources [IMC03]. Unfortunately, dynamic protocols requires longer time and do not guarantee 100% service recovery as sufficient capacity may not be available at the time of the failure due to high network load [RS03, LS99]. Dynamic restoration mechanism has three different techniques:
  - *Disjoint path restoration:* This requires that the affected S_D pair tries to find an alternate route that does not share any link with the affected path and reroute the traffic on this newly established path [Bha98].
  - *Partially joint path restoration:* this technique reroutes the traffic through the shortest alternate path that usually has some common links with the affected path [CAQ02].
  - *Link restoration:* this simply tries to establish the shortest possible loop around the failed link.

As we have seen, link restoration is a single-hop point-to-point whereas path restoration is an end-to-end recovery scheme. Since link restoration tries to set-up the

shortest possible loop around the failed link using less number of links than path restoration. As a result, link restoration reduces the number of switches that should be reconfigured. Furthermore, it requires less restoration time. Furthermore, link restoration keeps the network running the way it is preserving the network load balancing avoiding any sudden chaos in the network.

Despite all these advantages, link restoration could congest the area around the failed link not to mention that it could set up considerably long restoration loops in highly loaded networks. It is important to mention that link restoration requires fault localization interval before the process of restoring that link starts [SM06].

On the other hand, path restoration does not require any fault localization; the restoration process simply starts after detecting the failed link. Destination node usually is responsible for activating an alarm to indicate the existence of a failure. Despite the fact that path restoration does not require a localization protocol to locate the failed link, it could be time consuming for diverse s-d pairs or when the network is heavily loaded. The reason for this is that for each affected lightpath an alternate route should be located [SC03].

Finally, link restoration only removes the capacity of the failed link whereas path restoration removes all affected lightpaths capacities from the network resources [LX05].
In this work, we are focusing on link restoration due to the highlighted advantages in the motivation section, and hence we propose different types of localization protocol for all-optical networks in which can be used with link restoration scheme.

2.3 Fault localization Protocols

In communication networks, the well known routing protocol Open Shortest Path First (OSPF) has introduced fault localization method [Moy et al 98]. Fault localization is being done through sending what is so called Hello packets, each router discover it's neighbors through sending these Hello packets on each interface. By default each router sends Hello packets every 10 sec and waits for the Hello packets from the neighbor router. If the router did not receive any Hello Packet from its neighbors within 40 sec, the router assumes that the

link to that neighbor is down. As a result, this information about this link will not be advertised in the routing information to other routers [RFC 2328].

Despite the fact that this is a practical way of localization a fault, it does not suit high bit rate networks such as optical networks as the failure localization takes a very long period of time. Hence, if a Hello interval is being short this will allow faster detection for faults; however, this will consume lots of bandwidth of the control channel which is undesirable thing. On the other hand, using longer Hello interval will defiantly slow down the detection of the failure. As a matter of fact, The Hello interval is chosen with respect to two criteria, the first one is the size of the network and the second one is the router dead interval which is usually four times the selected Hello interval [WB95].

Fault detection is also practiced in Voice over Internet Protocol (VoIP). One of these ways is that using short and low bandwidth control packets known as keep-alive messages. These messages are used to closely monitor the network status to help rapid switchover of calls [KK04].

Another protocol that is used for fault detection is known as ping-pong protocol. This protocol relies on centralized node known as dispatcher that acts as a master to control a set of nodes and their links in a parallel way. This method is also known as master-slave and pingger-pongger [SHS04]. The fact that this method suggests a centralized dispatcher makes it not efficient or scalable. Not to mention that, this protocol sometimes interpret returned low power signal as disconnections.

Another way to locate failure is to make use of stochastic. This notion of discovery has been described by such terms as knowledge extraction, database exploration, data pattern processing, information harvesting, or data dredging. This requires that the process of fault detection applied to large quantities of data, finding implicit knowledge hidden in the information. Data processing make use of many machine-intelligence and information techniques such as pattern analysis, induction decision trees, clustering, and time series analysis to locate a failure [TLH04]. These techniques are used to high disturbance systems such as power systems. As these methods relies on stochastic, they are some how time consuming and not always reliable. Again, such methods can not be used for fault localization in all-optical networks as these networks requires faster methods to met the restrictions introduced by service in which these networks supports.

All in all, despite the fact that these mentioned protocols has been put in service for a long time, and has been debugged and improved through out the years, none of them could be used for a high data bit rate like all-optical networks as these protocols are slow and inefficient. This suggests that a new localization and detection protocols are needed for all-optical networks.

2.3.1 Fault detection in SONET

SONET is builds of blocks, the basic block is known as synchronous transport signal 1 (STS-1). As shown in Figure 2. 1, STS-1 is build of two major sections, the first one is known as transport overhead which consist of 3 columns and 9 rows and carry the header information whereas the second one is the Synchronous Payload Envelope (SPE) which consist of 87 columns and 9 rows and carry the data.

The transport overhead is further divided into two other sections, the line and section overhead. Each of these sections has a set of bytes each has his own reason, we will only talk about those bytes that is being used for fault detection in SONET [Gor97].

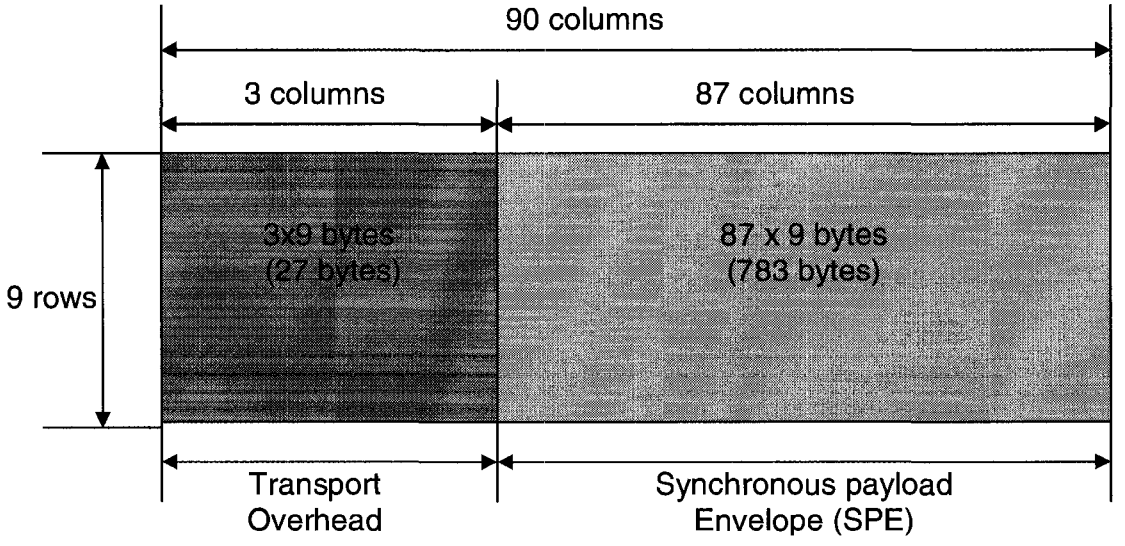

Figure 2. 1: The SONET STS-1 frame structure

- Trace Identifier (TI): A traffic data signal is commonly associated with a path trace identifier, or section trace identifier, which carries information identifying the data signal. The path trace, or section trace, information is normally used to verify network connectivity. Typically, this is achieved by comparing `expected` trace information with `received` trace information. If there is mismatch, then this is an indication of a misconnection. In an SDH/SONET network, the J1 or J2 byte of the traffic data structure normally serves as the path trace component, while the J0 byte serves as a section trace component [Bat01].
- Forward Defect Indicator (FDI): A specific example is of an optical network element that detects an abnormal condition and is required to assert a so-called "Forward Defect Indicator (FDI)". An FDI is a message generated by an Optical Channel (OCh) processing unit that is used to inform downstream equipment that an OCh is defective. Each FDI indicates the affected OCh and, optionally, also indicates the root cause of the defect. Also, an FDI is generated when an optical cross connect detects a cross point failure and is required to assert an OCh-FDI indicating the defect. In any of these situations it is clear that some optical channel overhead (OCh-OH) information will be required to be transported from network element locations at which the OCh Associated overhead is not available. Therefore, the only available path to transport the OCh-OH information is the Optical Supervisory Channel (OSC) which has the ability to read the overhead [Bat01].
- Backward Defect Indicator (BDI): destination node is adapted to pass failure information to the source node of the primary path, upon detecting a failure on the primary path, by inserting backward-defect-indicator (BDI) packets into a path that is at least partially disjoint from the primary path.

In fact, FDI and BDI signals are needed for more than one reason, these two signals can be used at different sub-layers to distinguish between fiber failures and channel failure. These are used to distinguish between the failure of a section of the link such as the failure between two amplifiers or indicating that the entire link is down. Furthermore, these signals are being used to allow specific nodes to act on respect to an alarm. Usually nodes adjacent to the failure will act to reroute the traffic through alternate path. So, when FDI received by a node

it will realize that the failure has been taken care by other nodes and no further actions is required [Gor97].

When a given link has experience certain type of failure, this failure will impact all downstream nodes by the interruption of data flow. This failure will be reported to the network management entity responsible within the domain. This failure will also be reported to neighbor nodes to inform these nodes not to take any actions. On response, the network management entity initiates a remote testing, fault correction, fault localization, isolation, and restoration procedure. The closest effected node to the failure performs the switchover and informs the management of the result. It is important to point out that in WDM transport networks some critical issues might rise up and it has to be taken care during the network design. Effects due to linear and nonlinear such as dispersion, losses, Cross-Phase-Modulation can cause lots of distortion and degrade the quality of the protected signal as the characteristics of the established new path could be different [GA03].

2.3.2 Mesh Optical Networks

Fault localization in All-optical mesh networks has not been deeply investigated and hence the number of research papers is very limited in this area. In opaque optical networks, fault detection can be achieved by electronically monitoring the loss of signal power at the receivers in an electro-optical switch [MT01].

In [ZPB06], proposes a fault localization technique in an all-optical overlaid-star TDM network. Two alternative techniques for the fault localization are proposed. The first one requires a control channel on each wavelength of a fiber link (one time slot in each TDM frame) to facilitate fault localization. Whereas the other requires that the source edge node transmits a small data block at the beginning of each non-allocated time slot. This small data blocks allows the destination edge node to immediately detect and localize a downstream link failure or a switch fabric fault, significantly reducing the detection time. More importantly, the second alternative technique does not require the reservation of a control channel on each wavelength of a fiber link, which is an advantage compared with alternative 1.

In mesh networks, fault localization can be done at different layers. At the physical layer, the power of a signal [Neu02] or the Optical Signal-to-Noise Ratio (OSNR) [WK02] can be monitored to detect and locate the failure. In [SS02], optical monitoring components like power meters are used to generate alarms whenever a signal loss is detected. This work also suggests that optimal placement of monitoring devices can limit the number of redundant alarms and provide sufficient information to localize a fault. This approach formulates the problem of receiving many redundant alarms and proposes an algorithm to solve it using an alarm matrix. Finding an optimal solution for this problem is basically trying to remove repeated detected signals from the network.

In [HE04], a network can be partitioned into sub-networks called "islands". Then a node or link failure is detected through the proposed protocol called island-by-island restoration protocol. It is worth mentioning that these islands are predefined and specified during the initial network design. These islands may be updated if the network topology has changed. This paper suggests three different partitioning methods, namely minimal island, shortest path island, and 2-stage Island. Despite all these partitions, the protocol lacks the capability of explicitly locating a failure.

In [SS08], this approach based on partitioning the network topology into many independent domains each one of them is run by a local fault-manager. The goal here is to reduce the alarm messages being exchanged back and forth crossing the entire network. So each independent domain will manage its alarm vector size in parallel with respect to other independent domains and limit the fault localization process within its independent domain. This monitoring hierarchy allows scalability as well as multiple fault-localization granularity levels.

In [ZH04, ZVS04], a network can be divided into a set of monitoring cycles, and each node and link is covered by a given cycle. Two algorithms, called Heuristic Depth first Searching (HDFS) and Shortest Path Eulerian Matching (SPEM), are proposed to find monitoring cycles. Each composed cycle is referred to as a monitoring cycle and there is a dedicated transceiver at a given node of that cycle. This will allow a set up of a loop-back dedicated supervisory channel for each cycle in which optical power, optical spectrum, optical SNR, high BER can be measured for each cycle. As a result, this will reduce the number of monitoring devices and thus reduce the overall cost of monitoring equipments.

Furthermore, this approach allows only the affected cycle to trigger an alarm, which reduces the number of triggered alarms, but at the cost of increased computational complexities.

In [ZH04, ZVS04], showed how a network can be divided into a set of monitoring cycles (m-cycles), and each node and link is covered by a given cycle in order to localize any failed link. In [ZHV06], an m-cycle construction for fault detection is formulated as a cycle cover problem with certain constraints. A heuristic spanning-tree based cycle construction algorithm is proposed and applied to different network architectures. The results of this approach showed nearly optimal performance. In [ZV07], the necessary conditions for achieving complete fault localizations are investigated in this approach. It has been shown that complete fault localization can be achieved if the m-cycles is constructed through a spanning-tree algorithm. A newly mathematical formulation problem known as variant constrained cycle-cover problem (vCCCP) is presented, and develops a branch-and-bound (B&B) algorithm for solving the newly defined vCCCP. The results showed that this new mathematical approach provides significantly improved results compared to the previous algorithm explained in [ZHV06]

In [WCZ05], aims to find effective methods that greatly reduce the hardware, cost and computational complexity for fault reporting and localization in future all-optical WDM networks. Instead of relying on network nodes that is assumed to be equipped with a performance-monitoring module that is active and reporting all the time. Instead, this paper assumes that optical signals can be sequentially sent along a set of lightpaths over an all-optical network to probe its state of health. Based on the results of these tests, a failure patterns for the network can be obtained. To minimize the required number of probes, the results of the previous tests are used to determine the successive probes to effectively localize the failure.

In [ZVSH04], it makes use of added pilot tones to wavelength channels which act as channels identifiers (CIDs) at input ports, this approach is knows as wavelength-routing fault detection scheme for concatenated optical crossconnects (OXCs) in all-optical networks (AONs). In this approach, routing errors at each OXCs can be simply localized by comparing CIDs at output ports with stored local routing information.

Fault Location Algorithm (FLA) is another fault localization scheme proposed in [MT00]. This scheme uses two operation phases to speed up the localization process. The

first phase, called Pre-Computational Phase (PCP), is performed off-line before any alarm is received. This phase results in generating a binary tree based on the established paths in the network. In the second phase, also called Diagnosis Phase (DP), a binary tree is traversed based on the received alarms from the active components to localize the failure. Despite the fact that this protocol can detect multiple failures, it is incapable of detecting failures associated with empty leaves. Moreover, the protocol model and component definitions are not applicable to transparent WDM networks.

Readers are referred to [KT96, AR97, MC98] for related work on fault detection at the physical layer using spectrum analyzer and photodiodes.

[LR97] proposes a finite state machine method, this approach becomes very complex in large-scale and dynamic networks make it hard for this approach to be deployed.
It is worth to mention that the International Telecommunication Union (ITU) is currently in the process of setting standard using GMPLS for fault detection on all-optical networks.

2.3.3 Limited-Perimeter Vector Matching Fault Localization Protocol

The LVM protocol is a fault localization protocol for localizing single-link failures in all-optical networks [Sic06, SM07]. This protocol assumes that no optical power monitoring is available at each intermediate node and only an edge node is able to detect the power loss or quality degradation of an optical signal. Therefore, it provides the maximum level of transparency by skipping any power monitoring or spectrum analysis at intermediate nodes on a lightpath. An advantage of the LVM protocol is that it limits fault localization within a smaller perimeter area and can thus significantly reduce the space and time complexities of fault localization. This protocol assumes that the control plane is reliable and de-coupled from the data plane.

To help understand this work, we first give a brief introduction of the fault localization process using the LVM protocol. The reader is referred to [SM07] for more details on the LVM protocol.

1) Once a link fails, all lightpaths passing through the failed link will be disrupted. As a result, the destination of each of the disrupted lightpaths will detect the signal loss of that lightpath and thus the failure.

2) Once a destination detects the failure, it will broadcast an ALARM message to all nodes in the network, which contains its node ID and the length of the shortest affected lightpath terminating at it.
3) Each sink node, also called potential executive sink (PES), will receive a copy of the ALARM messages sent by the other PESs. Each PES will compare the lightpath that terminates at it with the other lightpaths received by the alarms originated by other PESs. The PES that has the shortest lightpath will become the executive sink.
4) The executive sink will then create a vector of links on its shortest affected lightpath, called Affected-Link-Vector (ALV), defines a limited-perimeter, which consists of all the ALV links and their neighboring nodes, and then multicasts the ALV to all the nodes within the limited-perimeter.
5) Each sink receiving the ALV within the limited-perimeter will create its own link vector and compares the vector with the ALV. If its link vector has a common link with the ALV, the sink will send a binary vector back to the executive sink. The binary vector consists of a matching vector and the status of the corresponding lightpath. The matching vector contains the same links in the ALV. If the lightpath is working, its status is set to "1". Otherwise, the status is set to "0". For a working path, an element in the matching vector is set to "0" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "1". For a failed path, an element in the matching vector is set to "1" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "0".
6) After the executive sink collects the binary vectors from all the sinks within the limited-perimeter, it will perform a logical AND operation on all the collected binary vectors to determine the location of the failure. The link with "1" in the resulted binary vector will be determined as the failed link. If there are more than "1" in the resulted binary vector, the executive sink can extend the limited-perimeter by including the neighbors of the original perimeter nodes and run a second round of the above process.
7) Once the failed link is localized, the executive sink will broadcast a FAULT_LOCATION message to all nodes in the network, notifying them of the fault location.

Figure 2. 2 shows a network with a failed link b and the affected sinks, node 10, node 14, and node 17, which are also referred to as potential executive sinks.

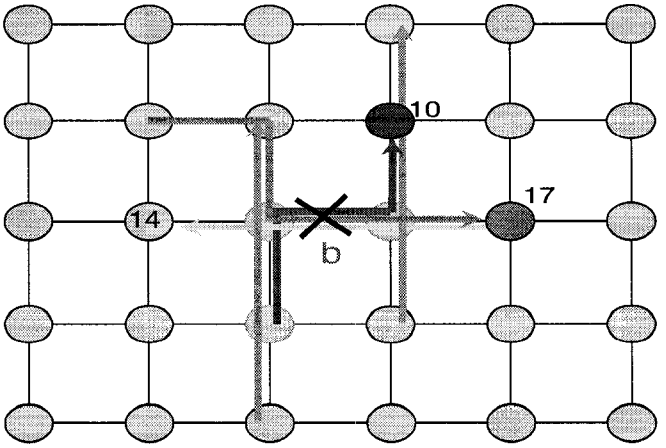

Figure 2. 2: A link failure with three affected sinks

Once the link fails, each potential executive sink will broadcast an alarm to all nodes in the network. Each alarm includes the node ID and the length of its shortest affected lightpath. Each potential executive sink will receive an alarm from every other affected node. For example, node 10 will have a copy of the alarm sent by each of the other two potential executive sinks, node 14 and node 17, respectively. After that, each potential executive sink will compare the length of its shortest affected lightpath with that of each shortest affected lightpath passed to it through the alarms. As a result, node 10 will become the executive sink and the associated lightpath of this node will become the executive route because node 10 has the shortest affected lightpath and the smallest node ID. Then the executive sink will create a vector called Affected-Link-Vector (ALV), which contains the links of the executive route, i.e., link a, link b, and link c, as shown in Figure 2. 3. The executive sink also defines a limited perimeter area, which consists of all the links and the nodes that are neighbors to the executive route. In Figure 2. 3, the limited perimeter area consists of nodes 4, 9, 11, 14, 17, 20, 22 and 27. After that, the executive sink will multicast the ALV vector to all nodes within the perimeter area. In response, each node within the perimeter area will create a binary vector and send it back to the executive sink if it has a lightpath terminating at it. This binary vector results from matching its lightpath with the ALV vector received from the executive sink, which is illustrated in Figure 2. 4.

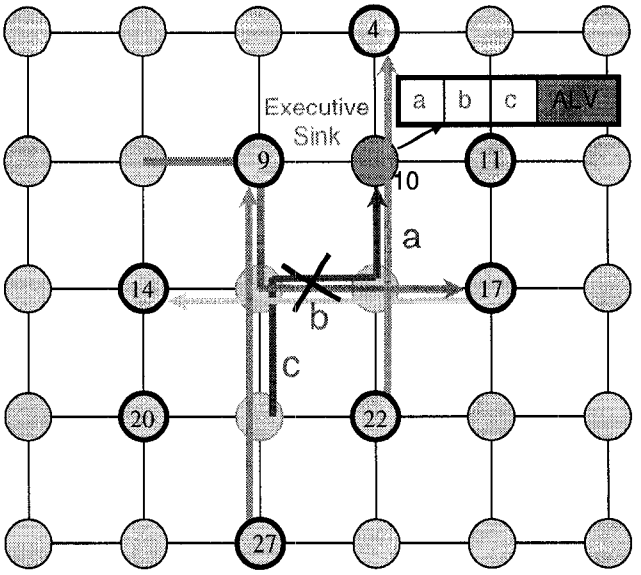

Figure 2. 3: A link failure with the exertive sink and the executive route

In Figure 2. 4, node 4 has a working lightpath that terminates at it. Thus it sets its status in the binary vector to "1". It also sets the cell for link a to "0" because link a exists in the ALV whereas link b and link c are set to "1". On the other hand, node 17 has an affected lightpath. Thus it sets its status in the binary vector to "0" and link b to "1" because link b does exist in the ALV. The other links are set to "0". Node 9 and node 14 will do the same and send their binary vectors to the executive sink.

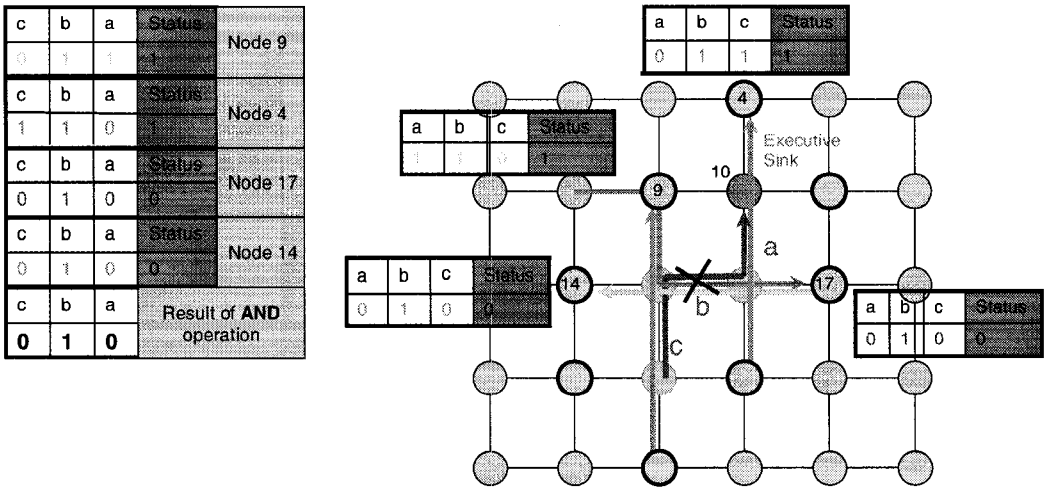

Figure 2. 4: Illustration of binary vector matching

After the executive sink collects all these binary vectors from nodes 4, 9, 14 and 17, it performs logical AND on the binary vectors to locate the failed link, as shown in Figure 2. 4. In the resulting vector, the failed link b has a value of "1" as shown on the left side of Figure 2. 4, which indicates that link b is a failed link, whereas the working links a and c have a value of "0". Once the failed link b is located, the executive sink will broadcast the information about this failed link to the entire network.

2.3.4 Multi-Failure Localization protocol

Multi-failure localization has not been widely studied in all-optical networks. In [MT00], Mas and Thiran proposed an efficient Fault Location Algorithm (FLA) for locating soft and hard failures in WDM networks. FLA uses two phases of operation to speed up the localization process. The first phase, called Pre-Computational Phase (PCP), is performed off-line before any alarm is received. In this phase, a binary tree is generated based on the established paths in the network. In the second phase, called Diagnosis Phase (DP), the binary tree is traversed based on the received alarms from active components. Although FLA can detect multiple failures, it is incapable of detecting failures associated with empty leaves. Therefore, it is not applicable to transparent WDM networks. In [MT05], Mas et al. proposed a Transparent Failure Location Algorithm (TFLA) based on FLA [MT00]. This algorithm assumes that network nodes are equipped with monitoring devices that can generate different types of alarms. These devices include optical power meter, optical spectrum analyzer (OSA), eye monitoring, BER monitoring, and Wavemeter. Each of these devices can detect a different type of failures and hence enable this algorithm to locate more than one type of failures at the same time. TFLA consists of three phases: a Pre-Computational Phase (PCP), a Core Phase (CP), and a third Monitoring Equipment Placement Phase (MEPP), which is a particular extension for optimal placement of new monitoring devices for fault localization. The disadvantage of this algorithm is that it assumes many monitoring devices, which increase the complexity and the cost of network nodes.

2.3.5 Inter-Domain Fault Localization Protocols

Fault localization has not been studied in the contest of inter-domain, there has been no material as far as this research is concern about inter-domain failure localization in all-optical networks. However, in IP networks, the only well known protocol for inter-domain routing is Border Gateway Protocol (BGP) that has a preliminary fault localization scheme. This protocol is also investigated to be used as a routing protocol to exchange wavelength reachability. [JY02], [FSHL01], [Kha04] have proposed an extensions to BGP routing protocol to exchange wavelength reachability among different Optical Domains (ODs), unfortunately, extending IP-network protocol to function in all-optical networks can not be used for fast fault localization as we will see in the following sections.

2.3.5.1 Border Gateway Protocol

The *Border Gateway Protocol (BGP)* is an external routing protocol. BGP is a Path Vector (PV) type protocol. In PV, each border router advertises the destinations it can reach to its neighboring *Edge Routers* along with the information that describes various properties of the paths to these destinations. In other words, PV defines the route as a pairing between the destination and the attributes of the path to reach that destination. Thus the name path-vector comes from the fact that each edge router receives from its neighboring edge router a vector that contains a set of routes [RFC1322].

BGP is used to exchange routing information among different ODs located in different geographical regions. The main task of BGP is to allow edge routers that belong to a certain OD to exchange network reachability information with other BGP systems located in different ODs [Ste99]. This network reachability information includes information on the sequence of ODs that the reachability information passes through.

This information is used by edge routers to build a map of how ODs are connected to each other. Besides, each OD can make use of this routing information to detect and prevent routing loops. Loops can be detected because each OD can reject any routing information that has its OD name in it. Furthermore, BGP can give each OD the ability to apply a certain routing policy by allowing each OD to set the cost of the advertised link or prevent certain routing information to reach certain costumers [RFC1771], [Tho01].

Figure 2. 5 shows a simple scenario for a network that has multiple ODs connected together. Clearly, there are five ODs, each has a number of edge routers. This scenario also shows the geographical working area for each routing protocol; intra-domain routing protocol like OSPF and inter-domain routing protocol like BGP.

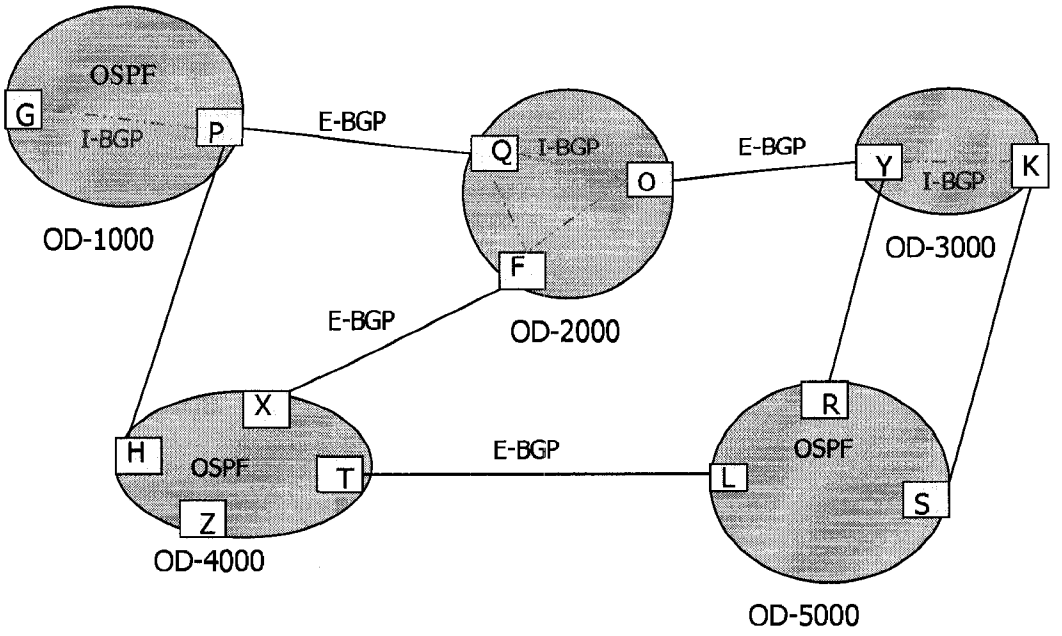

Figure 2. 5: A network Scenario consisting of five ODs

2.3.5.1.1 BGP Messages Type

BGP has four types of messages that share the same header structure. The four messages are: Open message, Notification message, Update message and finally Keep alive message [Tho01].

- Open message: is the first message that is used by any edge router after the TCP connection has been established. Usually, both participants send each other this message to identify each other [RFC1771].
- Update message: this message is used to add or remove a prefix. For example, if a certain prefix no longer exists then the BGP update message will include this prefix

in the withdraw section whereas if a new route "prefix" has been discovered then this prefix will be advertised as a new route [RFC1771].

- Notification message: this message is used to report a fault during a BGP session [RFC1771].
- Keep alive message: this message is sent during a BGP session between two BGP speakers to confirm that the BGP session is still alive.

The BGP session is kept alive until the two participants exchange their new routes [RFC1771].

2.3.5.1.2 BGP operation "I-BGP and E-BGP"

During a BGP session between any two edge routers, the two participating peers exchange open messages to identify each other and to negotiate a Hold Time. Each peer will send its new routes through update messages. Each update message includes one prefix associated with a certain number of attributes such as OD_Path, Next_hop, cost –etc. For more details about these attributes, readers can check the following [RFC1771] [Ste99]. Each peer can keep the session alive by sending Keepalive messages to the other peer, otherwise if the hold time expirees and no update or Keepalive message has been received then the other peer will be declared as unreachable (failure located) and the BGP session will be terminated. Since BGP works on top of TCP, the establishment of TCP session depends on the round-trip propagation time of packets, which is known as the round-trip time (RTT), is one of the most important parameters for determining TCP performance. And hence, it is not possible to use BGP for localization comparison. Not to mention that I-BGP does not run on every single intermediate node as I-BGP is pre-configured on selected routes.

When BGP is used between two different ODs, this mode of operation is referred to as External BGP (E-BGP). If an Internet Service Provider is using BGP to exchange routes that have been learned by an external BGP session or by any other means within its OD, then this mode of operation is referred to as Internal BGP (I-BGP) [Tho01]. The most important fact about the operation of BGP as I-BGP is that each node has to peer with all other edge nodes located in the same OD through a logical connection [Ste99]. The reason for these logical connections is to allow edge routers that belong to the same OD to exchange the

routing information learned via different external sources. This mode of operation is known as "full-mesh I-BGP". In fact, since I-BGP sessions are logical sessions, there are no direct physical connections among the participants. Each OD configures these logical connections among its edge routers. For example in Figure 2. 5, node Y in OD_2000 will have a logical peer with both edge routers Z and W. These logical paths are configured by the network administrator of OD_2000, the network administrator will specify the intermediate nodes that will be used to establish the I-BGP session between router Y and Z and the intermediate nodes that will be used to establish the I-BGP session between router Y and W. Unfortunately, these configured paths might not be configured properly causing some of the data packets to be lost somewhere along the path, or to be sent back and forth (sometimes called "oscillation") without reaching the required destination [GG02], [BLRW02]. A study on BGP miss-configurations found that up to 1200 prefixes in the Internet maybe suffering from miss-configuration every day [MWT02].

2.3.5.2 Optical Routing Border Gateway Protocol

Optical Routing Border Gateway Protocol (ORBGP) is built on the experience gained from BGP. In fact, much research is going on to improve the operation and scalability of BGP which suggests a promising future for this routing protocol [MF02].

ORBGP adopts most of BGP's features with some differences that allows wavelength information to be exchanged among edge routers to reflect the up-to-date bandwidth availability of the optical network. Furthermore, ORBGP introduces a new route advertising scheme which is triggered by the number of changes that took place in the link table of the advertising node [Kha04].

2.3.5.2.1 ORBGP Messages

As we explained earlier, ORBGP is built on the experience gained from BGP. In the previous section, we explained BGP in detail. Besides, we explained the message types and the purpose of these messages that are being used in BGP. ORBGP approach uses most of the BGP features.

The ORBGP session starts when the advertising node sends an open message to identify itself. Once the BGP session has been established, the two edge nodes exchange the optical routing information using the update message. The ORBGP make use of the update message to carry optical routing information between the communicating peers. Both nodes keep the ORBGP session alive by sending the keep alive message. Finally, both nodes report errors by simply sending a notification message indicating the failure type.

2.3.5.2.2 ORBGP Goals

There are many reasons that led to the definition of a new routing protocol. One reason is to have a routing protocol that serves the Internet Service Provider's (ISP's) basic needs. Of course, their basic need is to have routing information at their edge routers that guarantees any lightpath setup to any destination they wish to reach. Therefore, the aim of ORBGP is to guarantee a reliable mechanism to exchange optical routing information that reflects the most up-to-date topology of the global network.

[JY02] and [FSHL01] presented different existing approaches for performing lightpath provisioning across multiple optical domains. These two approaches suffer from the following weak points:

1. The assumption of a trusted relationship implicit in giving customers virtual control of their entities.
2. The lack of ability to have a complete control of the type of routing information advertised. For example, in Figure 2. 5, the network administrator of OD_2000 might decide to advertise the availability of 7 wavelengths out of 10 without specifying the color of the advertised wavelength that can be used to cross OD_2000. The reason for doing so is that OD_2000 does not want to reveal all its internal resources to the other ODs. Therefore, if this advertised information is used by other ODs to cross the OD_2000 domain, 30% of the lightpath requests coming from neighbors will be blocked. Normally, if a request is dropped by OD_2000, the neighbor who initiates that request will send another request asking for different wavelengths assuming that the first wavelength has been reserved by another customer.

3. The approaches extended BGP to perform both routing and signaling through two separate phases. The major problem with using BGP as a signaling protocol is that it is a configured protocol, therefore using this protocol to tunnel a lightpath request within certain OD is subject to oscillation due to the miss-configuration that may be introduced by the network administrator [GG02].

In order to avoid the weak point related to I-BGP miss-configuration, this approach assigned the I-BGP tasks to OSPF. Using OSPF can be more scalable if it also performs the I-BGP tunneling task along with the normal routing task in optical networks [RFC 2370], [Moy et al 98].

ORBGP is capable of advertising a specific amount of routing information and to give controlled privacy for the advertising OD.

Furthermore, ORBGP should avoid the advertisement policy defined in BGP. BGP does not allow edge nodes to advertise changes that took place in its routing table based on their needs. In fact, BGP performs automatic advertisement whenever a change takes place in its routing table. Of course, this makes no sense in the optical domain; an edge node should not advertise each time a wavelength has been released or reserved.

This approach suggest that fault localization will be done by two different protocols, for inter-domain, ORBGP will make use of BGP mechanism to locate a failure if a no keepalive or update message has received and hence, declare that link is down. On the other hand, since ORBGP information is tunneled within each OD through OSPF, hence intra-domain failure will be localized by OSPF as we explained earlier in this chapter. Unfortunately, depending on BGP or OSPF to perform localization requires lots of time and of course this does not meet all-optical networks requirements to meet desired services.

Chapter 3 Multi-Failure Localization in All-Optical Networks

3.1 Introduction

As we explained earlier, fault localization is a prerequisite for fault protection and restoration [AB08, AY08]. Multi-failure localization has not been widely studied in all-optical networks. In this chapter, we present a distributed fault localization protocol, called Parallel Limited Perimeter Vector Matching (P-LVM) protocol, for localizing multi-link failures in all-optical networks. P-LVM is based on the Limited Perimeter Vector Matching (LVM) protocol proposed for localizing single-link failures in all-optical networks [SM07]. To handle multi-link failures, it tries to separate each failure in a small perimeter area after identifying each perimeter area with its corresponding failure and then localize them in parallel in a distributed manner. Simulation results show that P-LVM can effectively localize multi-link failures with an average localization time below 40 sec, which is the localization time for OSPF (Open Shortest Path First) to localize a failure in traditional IP network.

The remainder of this chapter is organized as follows. In Section 3.2, we describe all possible cases in which different combinations of multi-failures may occur, present the proposed P-LVM protocol and its characteristics. In Section 3.3, we evaluate the performance of the protocol through simulation results. In Section 3.4, we conclude this chapter.

3.2 Distributed Multi-Failure Localization

In this section, we first identify different multi-failure cases and then present the P-LVM protocol and its characteristics.

3.2.1 *Multi-failure Localization*

To help understand the P-LVM protocol, it is necessary to identify different combinations of multiple failures. The P-LVM protocol is an extension of the LVM protocol. To avoid repetition we will only focus on the difference between LVM and the new requirements for P-LVM design. For ease of exposition, we will consider the scenarios with two link failures. However, the protocol can also be applied to the scenarios with more than two link failures. For each of the following cases, we detail how to locate the ES and the ER, for more details reader are encouraged to check section P-LVM or [SM07].

3.2.1.1 *De-coupled Multi-failure*

In this case, two failures occur simultaneously. They are apart from each other and there is no overlapping between the perimeters corresponding to the two failures.

Unlike LVM protocol that assumes all failures are due to one failed link, P-LVM has to be able to recognize the existence of two failures and apply separate localization process for each one of them. Figure 3. 1 shows two failures occurred at the same time, each failure affects two nodes. Nodes 12 and 22 have been affected by the left failure whereas nodes 17 and 27 have been affected by the right failure. Each affected node will broadcast an alarm that has the ID of the node and the affected lightpath. Hence, each Potential Executive Sink (PES) will have a copy of the four alarms. The square in the middle of the graph shows the information stored in each PES after receiving the four alarms from each other PES. The information shows that lightpaths that is terminated at node 12 and 22, has a shared link which is link b. The same thing for the other two lightpaths terminated at node 17 and 27 has a shared link which is link e. Therefore, each PES concludes that the failure affected node 12 and 22 did not affect node 17 and 27 because there are no shred links between them. Then

node 12 finds that it has shorter lightpath than node 22 so it start acting as first Executive Sink (ES 1), the same thing at node 17, it will realize that it has shorter lightpath than node 27 and it can start acting as ES 2. So, each ES will create its ALV vector and determined the limited-perimeter. Each ES will send the ALV vector within its perimeter area and wait for a respond from the sink nodes within its area, once the ES receives the respond from the perimeter nodes, the ES can perform a logical AND on the collected binary vectors to localize the fault. Once the failed link is determined the ES will broadcast this information to the entire network.

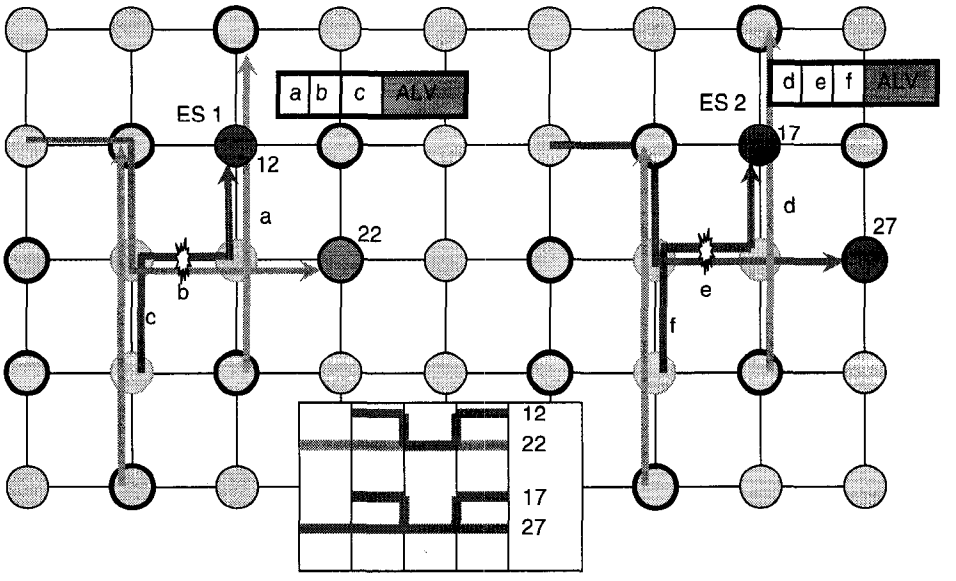

Figure 3. 1: De-coupled Multi-failure scenario

3.2.1.2 *De-coupled Multi-Failure Shared Lightpath*

This case is very similar to the previous Case. Two failures occur simultaneously. They are apart from each other and there is no overlapping between the perimeters corresponding to the two failures. However, there is a lightpath passing through the two failures.

Figure 3. 2 shows a network that has two failed links, it also shows that the affected lightpath that terminates at node 19 passes through the two failed links e and b. After exchanging alarm messages, each PES will have a copy of all alarms as shown in the square shown in the middle of Figure 3. 2 graph. Therefore, each PES will label this lightpath as un-valid information because it can not be used to localize a failure as it is not clear which link

caused the failure on this lightpath. Then the two ESs will be selected and each ES will run single fault localization as in LVM.

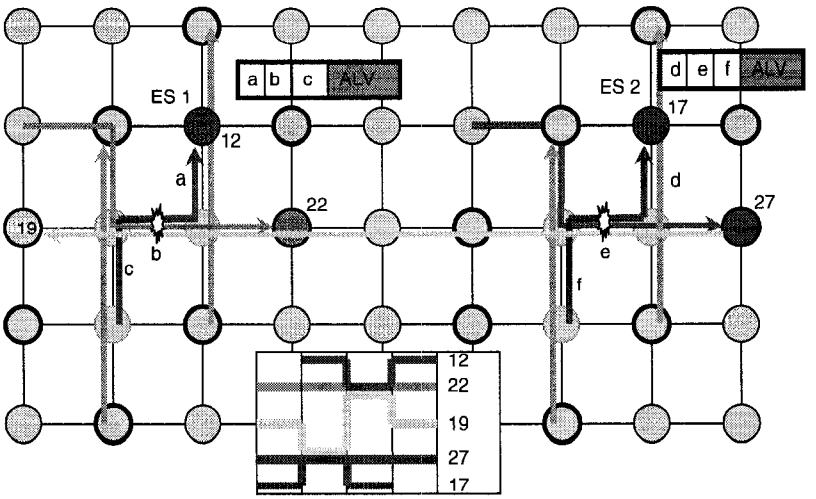

Figure 3. 2: De-coupled multi-failure shared lightpath

3.2.1.3 *Overlapping Perimeter Area with Shared Lightpath Scenario*

In this case, two failures occur simultaneously. They are close to each other and there is overlapping between the perimeters corresponding to the two failures. Moreover, a shared lightpath which terminates at node 17 passes through both perimeter areas and has shared links with both ER.

Following the same concept in the previous sections, each PES will get a copy of other PES alarms. As a result, each PES will have a copy of all affected lightpaths. Figure 3. **3** shows that each PES can locate that lightpath which terminate at node 11 and lightpath that terminates at node 29 have no shared links and they have the shortest ER. Therefore, node 11 will act as ES_1 and node 29 will act as ES_2. Each ES will run a single failure localization process to locate the responsible failed link that causes the detected failure. Furthermore, each ES can determine that the failed lightpath that terminates at node 17 has shared links with both Executive Routes (ERs) and hence it will not be used for fault localization.

Each ES then define its perimeter area and multicast its ALV, it is important to mention that few perimeter nodes can be included in both perimeter areas of the two ESs. Figure 3. 3 shows that node 20 located on both perimeter areas for the two different failures.

It also shows that node 20 has two lightpaths terminates at it. The first lightpath is originated from the upper perimeter area, whereas the other is originated from the lower perimeter area. We call this node as a Shared Perimeter Node (SPN). This SPN can distinguish between the two ES areas by checking the multicast ALV as each ES has unique ALV. Therefore, SPN will compare all lightpaths that terminates at it with the received ALVs. Any terminated lightpath at the SPN has shared links with ALV_1, the resulted binary vector will be sent back to ES 1. On the other hand, the resulted binary vector will be sent back to ES_2 if the lightpath has shared links with ALV_2. A binary vector for a working lightpath can be sent to both ESs if this working lightpath has shared links with both ALV1 and ALV 2.

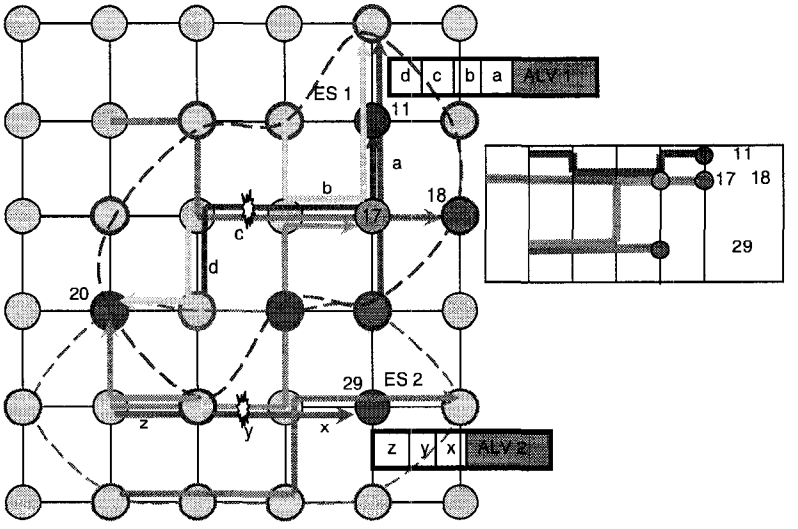

Figure 3. 3: Overlapping perimeter area with shared lightpath

It is also wroth to mention that, a node might act as an ES for both perimeter areas, so it will select two perimeter areas for each selected ER and multicast different ALV for each area. The dual ES will run binary vectors separately for each perimeter area as each area will send back messages that have different binary vectors.

3.2.1.4 *Sequential Multi-Failure Scenario*

In this case, two failures occur simultaneously. All failed lightpaths have shared links. Hence, once the first failed link is localized, the corresponding ES will find out that the

failed lightpath that terminates at node 14 does not pass through the failed link c, and hence node 14 will start another localization process.

Again, we assume that two failures occur almost at the same time. As a response, all PES exchange alarm messages, and as a result each PES will have a copy of all these alarms. Each PES will have the same information shown in the square located on the right side of Figure 3. 4, this square shows that all lightpaths has a common link, this will be interpreted by each PES as all lightpaths have failed due to single failure which might be the case.

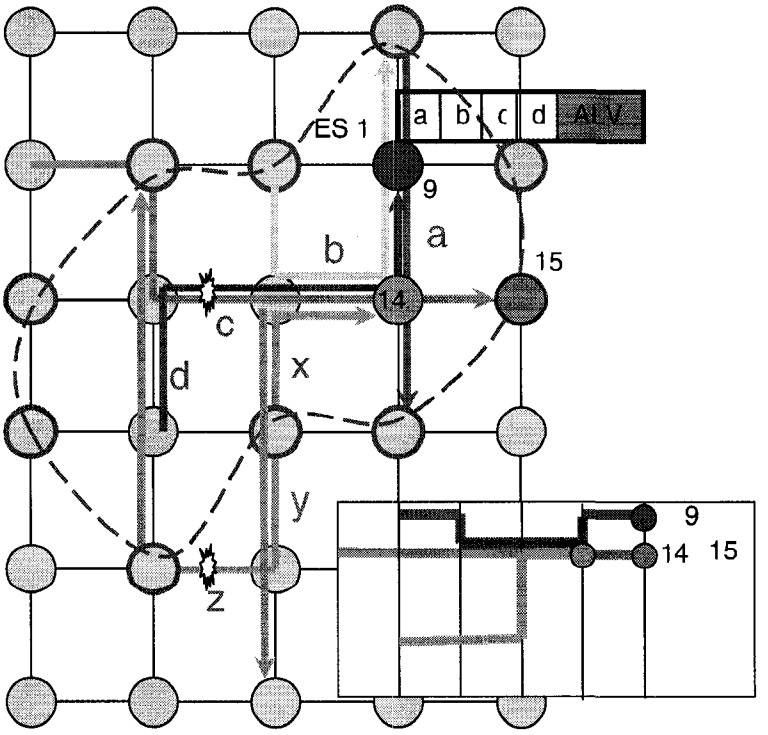

Figure 3. 4: Sequential Multi-failure recovery

Therefore, P-LVM starts the localization process by choosing node 9 as ES because it has the shortest ER, Node 9 will determine that link c is failed. However, before it broadcast the results about the failed link to the entire network, the ES will check whether all failed lightpath passes through link c. After checking link c, it can notice that the failed lightpath that terminates at node 14 does not go through the failed link c. Therefore, node 9 which acts as the current ES will trigger node 14 to act as new ES to locate the other failure and on the same time it broadcast the information about link c to the entire network. The new ES will

perform a fault localization process to locate the failed link z. Once it locates that link, it broadcast a message to the entire network indicating the location of the failed link z.

3.2.1.5 *Special Cases*

One special scenario is that two failures occur simultaneously and all failed lightpaths have shared links, as shown in Figure 3. 5. The executive sink node 11 will localize the failed link *b* and it will assume that no other failed links exist. In this case, one of the failed links will not be localized.

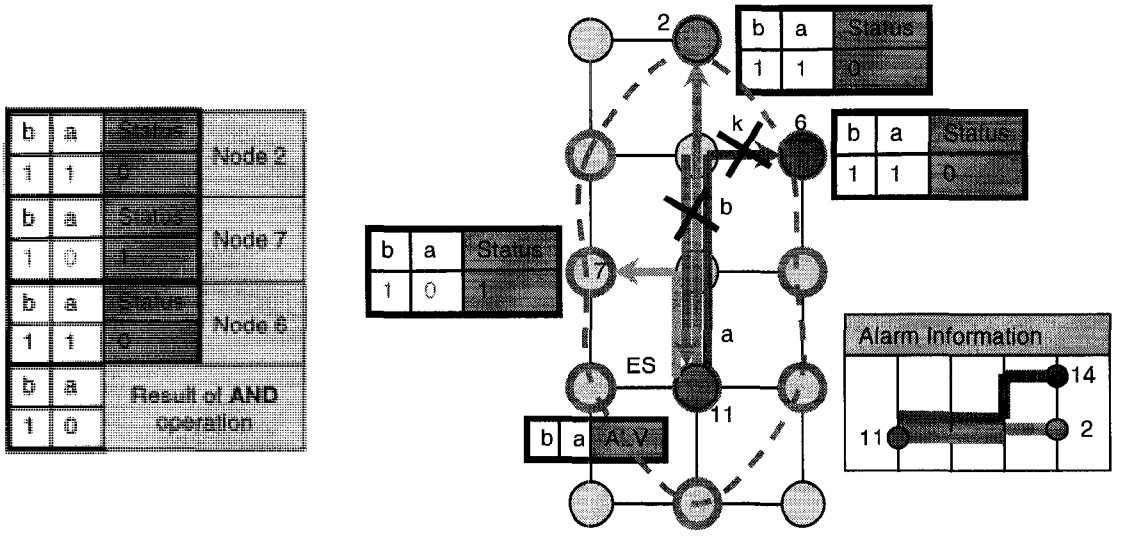

Figure 3. 5: Un-localized failed link

As it has been explained earlier, P-LVM starts the localization process by exchanging the alarms among all PES. Each PES will have a copy of all alarms after a certain period of time. Each PES realizes the alarm information and assumes that all failed lightpaths are due to single failure as all failed lightpaths have common link. So, node 11 will find that it has the shortest lightpath among all failed lightpaths and it start to act as ES. Then, ES define the perimeter area and multicast its ALV vector to all nodes within this area. As a response, node 2, 6 and 7 replies back with their binary vectors to the ES. The ES perform logic AND on the received binary vectors and find that link b is failed. ES also looks into link b and finds that all failed lightpaths are passing through link b. As a result, ES will assume that link b is the

only failed link and advertise this information to the entire network. It is clearly shown that the failed link k will not be localized unless enough lightpaths are available.

Another special scenario is that one failure occurs and all failed lightpaths have shared links, as shown in Figure 3. 6. In this case, all failed lightpaths share two or more links. In our example, all lightpaths shares link *c* and *b*. Therefore, the executive sink node 16 cannot tell exactly which link failed and hence it assumes that both links failed. The same scenario will be followed to localize the failure. After exchanging all alarms among all PES, PES number 7 will be selected as ES because it has the shortest lightpath and the smallest node ID. ES will determine the perimeter area and multicast its ALV to all nodes within that area. Nodes 3, 12 and 16 will respond back with their binary vectors to the ES as these nodes have lightpaths that terminates at them. Once the ES receive these binary vectors and perform logic AND on them, it will find out that both link b and c are failed. In fact, in this situation ES can not tell whether both links are down or only one of these links is down. Furthermore, if there is only one failed link ES can not tell which one of them is down.

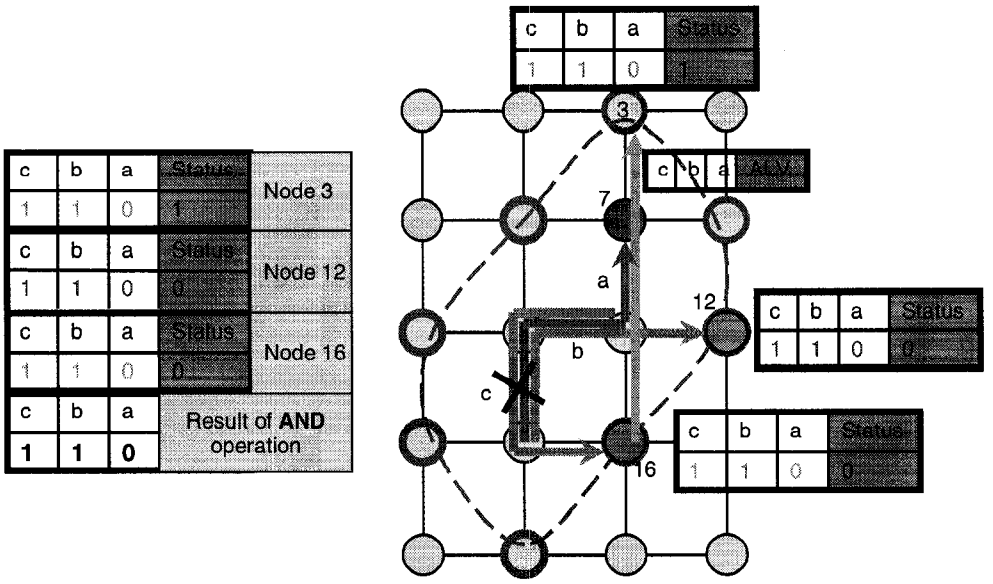

Figure 3. 6: Localization of Extra un-failed link

3.2.2 *P-LVM Protocol*

We now present the P-LVM protocol. We assume no optical power monitoring devices at any intermediate node. The only nodes that can detect the loss of light are the sink nodes of each lightpath. Each fiber has a number of lightpaths which is equal or less than the multiplexing degree supported at all intermediate nodes.

3.2.2.1 *Multi-failures Occurring at the Same Time*

The fault localization process for localizing such failures is described as follows.

- Once two link failures occur, all lightpaths that pass through the two failed links will be affected. In this case, each sink node that has an affected lightpath terminating at it will detect the failures. Note each sink node may have more than one affected lightpath terminating at it. At the beginning, each affected sink is considered as a potential executive sink (PES) but it is not clear that which failed link caused a particular PES node. Figure 3. 7 illustrates three affected lightpaths terminating at node 19, which are caused by two different failed links. To localize the failed links, node 19 has to advertise the information on all the three affected lightpaths in an alarm message to the entire network. If node 19 only advertises the information on the shortest affected lightpath passing through the failed link x, no information on the other two failed lightpaths (passing through link y) will be contained in the alarm message. In that case, no information on the failed link y could be exchanged in the network. As a result, the network will not be aware of this failure and hence will not be able to localize the failure. Therefore, node 19 has to advertise the information on all the three affected lightpaths terminating at it.
- All PES nodes pause for a pre-determined period. This pausing period will ensure all PES nodes have enough time to detect the failures. It is normally very short and is determined based on the propagation delay and the diameter of the network.

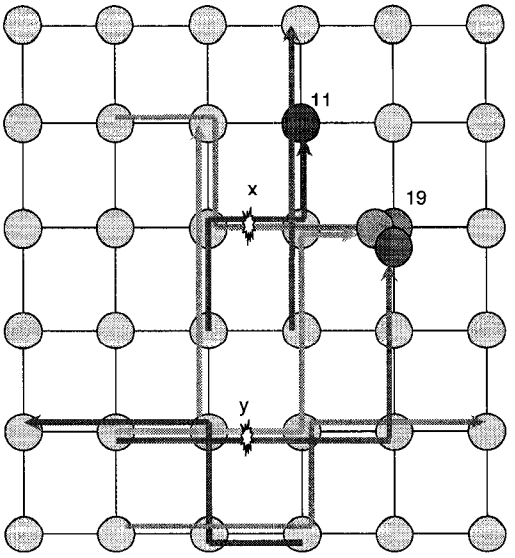

Figure 3. 7: Three affected lightpaths terminating at same PES

- During pausing, each PES node will generate an alarm message and broadcast the alarm to all other nodes in the network. Each alarm message will contain the ID of the PES node and all affected lightpaths terminating at the PES node. Therefore, each PES node will receive the alarm messages sent by the other PES nodes.
- Once a PES node receives the alarm messages from the other PES nodes, it will extract all the lightpaths and insert them into a list based on their lengths. If two lightpaths have the same length, the lightpath that originated from a node that has lower ID will be inserted first. Furthermore, if two lightpaths have the same length, originated from the same node, the lightpath that has lower connection ID will be inserted first in that list.
- Each PES node will compare the first lightpath in the list with all the other lightpaths in the list. If a PES node finds that there are two affected lightpaths that do not share any link. It will store each lightpath in a separate queue called Executive Route Queue (ERQ_i), where i=1, 2, 3, ... If the first lightpath has shared links with all the other lightpaths in the list, the second lightpath in the list will be considered. This process will continue until two separate lightpaths are found. If there are no separate lightpaths found, this case will be considered as a single-failure scenario, which can be handled similarly like the LVM protocol [SM07].

- Once two separate lightpaths are found, they will be compared against all the lightpaths remaining in the list.
  - If a lightpath remaining in the list has a common link with one of the two separate lightpaths, this lightpath will be added to the end of the *ERQ* storing that lightpath.
  - If a lightpath remaining in the list has a common link with both the two separate lightpaths, this lightpath will be ignored or removed because it could not provide any useful information for fault localization.
  - If a lightpath remaining in the list has no common link with any of the two separate lightpath, this lightpath will be added into a new ERQ because it represents a new failure, and the list will be compared against the new discovered lightpath. This process will continue until all the lightpaths remaining in the list has been compared.
- Since the affected lightpaths have been separated into different *ERQ*s in terms of the different failures that affect them, each *ERQ* can be used in a single-failure localization process, which can be performed in parallel with the other processes. Each parallel localization process is described as follows.
  - The first element in each *ERQ* is taken as the executive route (*ER*) because it has the shortest lightpath and the PES node of that lightpath will act as the executive sink (*ES*).
  - The *ES* will create an affected link vector (*ALV*) for its affected lightpath links and define a limited-perimeter area, which includes all the nodes and links on the *ER* as well as all neighbors of the nodes on the *ER*. After that, the ES will multicast the *ALV* to all nodes within the perimeter area.
  - Once a sink node within the perimeter area receives the *ALV*, it will create a binary vector by matching its lightpath with the ALV received from the *ES*. If the lightpath is unaffected, the lightpath status in the binary vector will be set to 1. For any link that exists on the executive route, the cell corresponding to that link in the binary vector will be set to 0. Otherwise, the corresponding cells will be set to 1. If the lightpath is affected, the lightpath status in the binary vector will be set to 0. For any link that exists on the executive route, the cell corresponding to that link will be set to 1. Otherwise, the corresponding cells will be set to 0.

- If the resulted binary vector has no shared link with the *ER*, the corresponding sink node will not send the binary vector back to the *ES*. However, if the resulted binary vector has shared links with the *ER*, the corresponding node will send the binary vector back to the *ES* which generated that corresponding ALV.
- After the *ES* collects all the binary vectors from the sink nodes within its perimeter area, it will perform logical AND on the binary vectors to locate the failed link. Once the failed link is located, the *ES* will broadcast the information on this failed link to the entire network.

3.2.2.2 *Multi-failures Occurring at the Different Time*

In this case, one failure occurs while the other failure is still in the process of being localized. P-LVM can still handle such cases.

- When a new failure occurs, the affected sinks will broadcast alarms to the entire network.
- If the ES of the first failure has not been identified, these new alarms will be included in the list of the affected lightpaths at each PES node and P-LVM will run the comparison again from the beginning to identify the new list of ES that can run fault localization for each failure. However, if a perimeter area has already been defined for the first failure, the ES can decide weather the localization process keeps going or stops as follows.
  - The ES will simply compare the ER with the lightpath route included in the newly received alarm. If there is a common link, the localization process should restart. This is due to the fact that some of the used lightpaths might have advertised the status as 1 in their binary vectors back to the ES but the failure will change the status of that lightpath from up to down or from 0 to 1. This will lead to unexpected results which make many fault locations unable to be localized properly.
  - Otherwise, if the ES finds no common links between the lightpath received in the alarm and the ER, it will simply ignore the alarm as the ES is sure that the alarm has no impact on the failure being localized.

3.2.3 *Protocol Characteristics*

The P-LVM protocol has the following characteristics:

1) It can localize multi-link failures that occur at the same time.
2) It can localize multi-link failures that occur at different times.
3) It can localize a failure in a perimeter area, which avoids flooding the network with extra control overheads.
4) It can localize a failure within different perimeter areas.

3.3 Performance evaluation

In this section, we evaluate the performance of the P-LVM protocol through simulation experiments. For evaluation, we consider two metrics: localization probability and localization delay. The localization probability is defined as the number of failures successfully localized over the overall number of failures. The localization delay is defined as the average time taken to localize a (dual) failure.

The simulation is performed on a platform developed in JAVA. In the simulation, we consider the National Science Foundation (NSF) network, as shown in Figure 3. 8. We assume that the connection requests arrive to the network randomly according to a Poisson process. The destination of each request is uniformly distributed over all nodes. The holding time of each connection is exponentially distributed. The routing algorithm is OSPF (Open Shortest Path First) and the wavelength assignment algorithm is First Fit (FF). In addition, we assume that the light speed inside a fiber is 2.14 x 108m/s, and the computational cycle (θ) for each node to determine the location of the failed link is 20ns. Since any lightpath between a pair of neighbor nodes can be used to localize the failure on the lightpath without using the P-LVM protocol, we exclude all possible lightpaths between neighbor nodes.
Finally, to ensure that our simulation results are accurate, we run each simulation for at least 40 times at each given rate of requests and calculated the average. We verified that running the simulation for 100 times is good enough to have all the values located within a confidence interval of 5% around the mean.

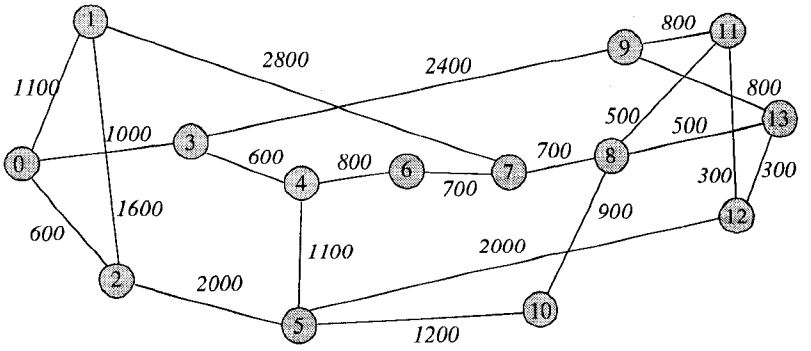

Figure 3. 8: NSF Network Topology

Figure 3. 9 shows the localization probability under different request rates. It is seen that when the request rate is small (i.e., less than 0.25), the localization probability increases remarkably with the increase of the request rate. This is because with the increase of the request rate more lightpaths are established in the network, which lead to more lightpaths available in a perimeter area and thus make fault localization easier. When the request rate is larger than a certain value, more lightpaths will not help more in fault localization and the localization probability thus tends to be stable. To be more specific, to localize a failure each link of the executive route has to have another lightpath that is originated from another source passing through it. If the status of this lightpath is "0" then this lightpath should not have any shared links with the other executive route as we are dealing with multi-failure here.

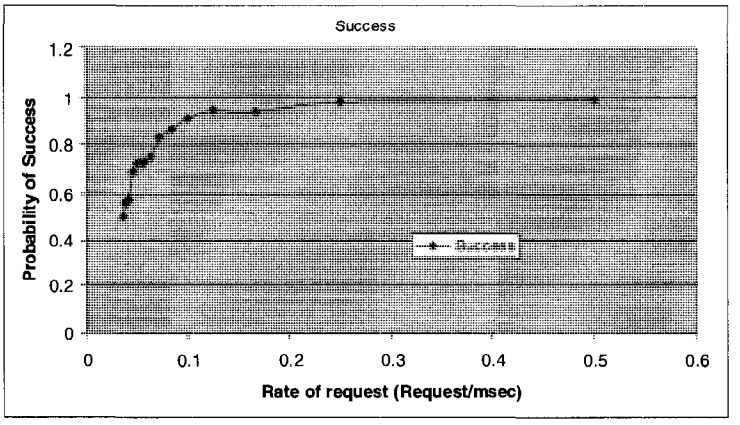

Figure 3. 9: Probability of successful localization

Figure 3. 10 shows the probability of unsuccessful localization. Obviously, with the increase of the request rate, this probability decreases remarkably when the request rate is small and tends to be stable when the rate becomes large. In addition, it shows the probability due to three different cases. In the first case, a failure occurs on a link does not have any lightpath passing through, we call this case "Unoccupied Link". In the second case, a failure occurs on a link which has lightpaths passing through, but during the localization process the perimeter area has no lightpath that terminates at it, we call this case as "Perimeter area". In the third case "Two ones", the AND operation by the executive sink results in two ones, indicating that either two failed links on the ER have been localized, which has a lower probability to occur, or there are not enough lightpaths that allow the ES to distinguish the failed links as explained in the second scenario of section 3.2.1.5.

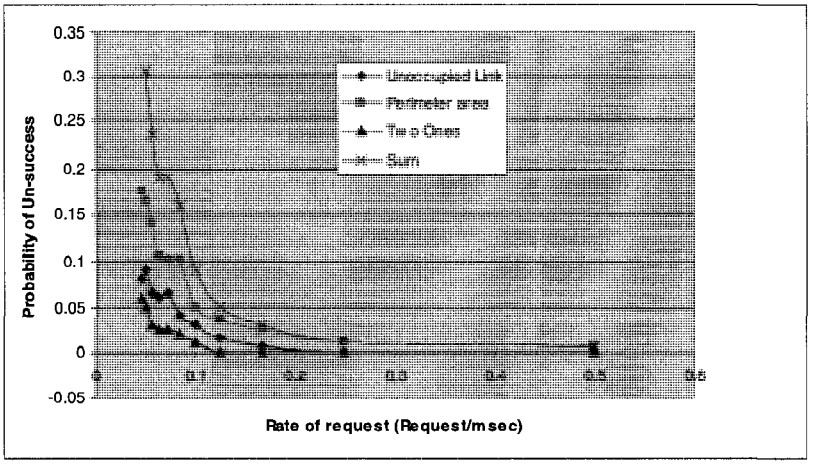

Figure 3. 10: Probability of unsuccessful localization

Figure 3. 11 shows the average localization delay as well as corresponding propagation delay and computational delay in seconds. It is seen that as the traffic increase the computational delay increases and hence the localization delay increases. This is because as the number of lightpaths increases the number of affected lightpaths will also increase, which increases the processing time to find two separate lightpaths. It is seen that the overall localization delay which is the sum of propagation and complexity delay is below 40 sec, which is the time needed to localize a failure using OSPF in traditional IP networks [Moy et al 98]. It is worth to mention that OSPF is also being used to exchange routing information in

opaque networks [RFC 2370] and GMPLS. Furthermore, this is the only well known protocol that is being used in both optical and IP networks that has localization capabilities.

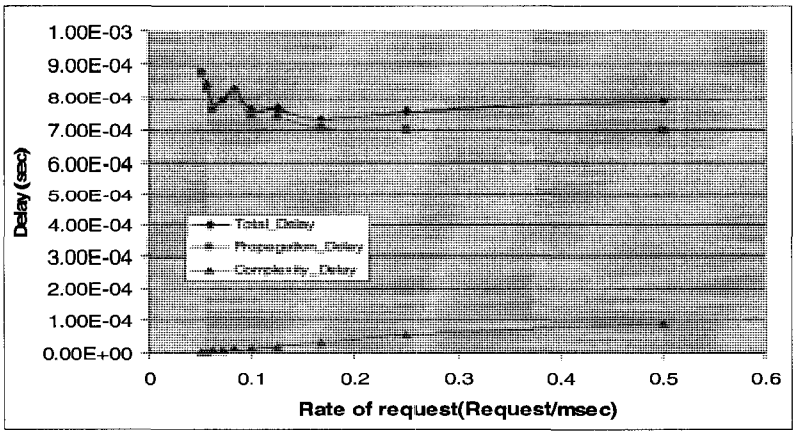

Figure 3. 11: Localization delay in sec

3.4 Summary

In this chapter, we proposed a distributed fault localization protocol, called Parallel Limited Perimeter Vector Matching (P-LVM) protocol, for localizing multi-link failures in all-optical networks. P-LVM is based on a Limited Perimeter Vector Matching (LVM) protocol for localizing single-link failures in all-optical networks. To handle multi-link failures, P-LVM tries to separate each failure in a small perimeter area after identifying each perimeter area with its corresponding failure and then localize the failures in parallel respectively in a distributed manner. The simulation results show that P-LVM can effectively localize multi-link failures that occur simultaneously or at different times in all-optical networks.

Chapter 4 Fault Localization-Constraint Connection Provisioning Optimization

4.1 Introduction

Network survivability is a big concern in all-optical networks. In this type of networks, a single network failure such as a fiber cut may cause a huge amount of data loss, which would largely degrade or even disrupt network services. To ensure network services, a network must incorporate efficient fault protection and restoration mechanisms to provide a high level of service survivability against different types of network failures [AS08, MH08, QH06].

Fault localization is a prerequisite for fault protection and restoration. LVM protocol was proposed for all-optical networks in [Sic06, SM07], which can effectively localize a link failure if there are at least two distinct lightpaths with different source-destination pairs passing through the failed link. However, if there are less than two lightpaths passing through the failed link, the protocol is unable to localize the failure. Therefore, an important problem when applying the LVM protocol in a static network, where the traffic (or lightpath) demand is known *a priori*, is to optimize the traffic distribution so that the fault localization probability in the network is maximized in the event of a link failure. On the other hand, traffic distribution may largely affect the time for localizing a failed link. Given a traffic demand, it is also important to optimize the traffic distribution so that the fault localization time in the network is minimized in the event of a link failure.

In this chapter, we study the optimization problems in applying the LVM protocol to static all-optical networks, and consider both the problem for maximizing the fault localization probability and the problem for minimizing the fault localization time. We formulate the two problems into an integer linear programming problem, respectively, and use the CPLEX optimization tool to solve the problems. We show that by optimizing the

traffic distribution the fault localization probability in terms of the number of localized links can be maximized and the fault localization time can be minimized. Moreover, we propose a heuristic algorithm to evaluate the optimization results through simulation experiments.

The rest of the chapter is organized as follows. In Section 4.2, we present the integer linear programming formulation of the two optimization problems. In Section 4.3, we present the proposed heuristic algorithm. In Section 4.4, we evaluate the optimization results. In Section 4.5, we conclude this chapter.

4.2 Problem Formulations

In this section, we present a formulation of the following two problems. One is to optimize the traffic distribution so that the fault localization probability in terms of the number of localized links is maximized and the other is to optimize the traffic distribution so that the time for localizing a failed link is minimized. This section provides the detailed description of the analytical models for both optimization problems.

4.2.1 Optimization for Fault Localization Probability

We consider a static network where a traffic demand matrix (or lightpath requests) is known *a priori*. Given the traffic matrix, the optimization problem considered is to optimally distribute the traffic (or lightpaths) in the network so that in the event of a link failure the fault localization probability is maximized, i.e. the number of links that can be localized by the LVM protocol is maximized. This optimization problem can be formulated into an integer linear programming problem. In the formalization, we assume that a traffic matrix does not contain any direct traffic demand (or lightpath request) between neighboring nodes. This is because if there is a direct lightpath between a pair of neighboring nodes, a failure on the link between the two nodes can be localized easily without using the LVM protocol. The notations used in the formulation are defined as follows.

- W: the number of wavelengths per link
- λ_{sd} : the number of lightpath requests from source node s to destination node d

- $LP_{ij\_r}^{sd\_w}$: A binary coefficient that indicates the reservation status of wavelength w on link (i,j) for a lightpath request r between source node s to destination node d. Since a wavelength on each link can only be assigned to one lightpath, the value of $LP_{ij\_r}^{sd\_w}$ can be either one or zero.
- U_{sd}^{ij}: the number of lightpaths established on the link from node i to node j between source node s to destination node d.
- $\overline{U}_{sd}^{ij}$: the number of lightpaths established on the unidirectional link between node i and node j between node pairs (s' - d') other than node pair (s - d).
- H_{sd}^{r}: the number of hops traversed by a lightpath between source node s to destination node d to accommodate lightpath request r
- D: the diameter of the network

Objective:

$$Max[\sum_{i}\sum_{j}\sum_{s}\sum_{d}(\overline{U}_{sd}^{ij} - U_{sd}^{ij})], \tag{1}$$

subject to

$$U_{sd}^{ij} = \sum_{r}\sum_{w} LP_{ij\_r}^{sd\_w}, \qquad \forall i,j,s,d \tag{2}$$

$$\overline{U}_{sd}^{ij} = \sum_{s'd'} U_{s'd'}^{ij}, \ s' \neq s \vee d' \neq d, \ \forall i,j,s,d \tag{3}$$

$$W - U_{sd}^{ij} - \overline{U}_{sd}^{ij} \geq 0, \quad \forall i,j,s,d \tag{4}$$

$$0 \leq U_{sd}^{ij} \leq \lambda_{sd}, \qquad \forall i,j,s,d \tag{5}$$

$$LP_{dk\_r}^{sd\_w} = 0 \tag{6}$$

$$LP_{ms\_r}^{sd\_w} = 0 \tag{7}$$

$$\sum_{w} LP_{ij\_r}^{sd\_w} \leq 1, \qquad \forall s,d,r,i,j \tag{8}$$

$$\sum_{j}\sum_{w} LP_{ij\_r}^{sd\_w} \le 1, \quad \forall i,s,d,r \quad i \ne s, i \ne d \tag{9}$$

$$\sum_{i}\sum_{w} LP_{ij\_r}^{sd\_w} \le 1, \quad \forall j,s,d,r \quad j \ne s, j \ne d \tag{10}$$

$$\sum_{w}\sum_{r} LP_{ij\_r}^{sd\_w} \le \lambda_{sd}, \quad \forall s,d,i,j \tag{11}$$

$$\sum_{s}\sum_{d}\sum_{r} LP_{ij\_r}^{sd\_w} \le 1, \quad \forall w,i,j \tag{12}$$

$$\sum_{w} LP_{ij\_r}^{sd\_w} + \sum_{w} LP_{ji\_r}^{sd\_w} \le 1, \quad \forall i,j,s,d,r \tag{13}$$

$$\sum_{W}\sum_{i} LP_{ik\_r}^{sdw} - \sum_{W}\sum_{j} LP_{kj\_r}^{sdw} = \begin{cases} -1 & if\ s = k \\ 1 & if\ d = k \\ 0 & otherwise \end{cases} \tag{14}$$

$$H_{sd}^{r} = \sum_{i}\sum_{j}\sum_{w} LP_{ij\_r}^{sd\_w}, \quad \forall s,d,r \tag{15}$$

$$H_{sd}^{r} \le D \tag{16}$$

$$\sum_{W}\left(\sum_{i}\sum_{k,k\ne s,d} LP_{ik\_r}^{sdw} - \sum_{k,k\ne s,d}\sum_{j} LP_{kj\_r}^{sdw}\right) = 0 \quad \forall s,d,r \tag{17}$$

The objective function in (1) is to maximize the defined coverage area. From the viewpoint of a particular *s-d* pair, a link is covered only if there are other *s-d* pairs that utilize it. Thus, $\left(\overline{U}_{sd}^{ij} - U_{sd}^{ij}\right)$ stands for the difference between the utilization of the link *(i,j)* by that particular *s-d* pair and that by other *s-d* pairs. Its value is supposed to be nonnegative to ensure that there are *s-d* pairs other than that particular *s-d* pair, which utilize the link. Therefore, the objective is to maximize the total of this difference. The link utilization is defined in (2), which provides the number of lightpaths that originate from node s and are destined for node d, and pass through link *(i,j)*. The complement of this utilization term is defined in (3) as $\left(\overline{U}_{sd}^{ij}\right)$, representing the number of lightpaths that originate from node *s* and are destined for node *d*, and pass through link *(i,j)*. The remaining part of the formulation is about the capacity and bound constraints. Constraint (4) guarantees that the utilization of a link is always less than the number of wavelengths on the corresponding link. Constraint (5)

is straightforward, i.e., the utilization of a *s-d* pair on the unidirectional link *(i,j)* is bounded to the number of demands between the *s-d* pair. Constraint (6) is for loop prevention by avoiding the destination node *d* to be treated as an intermediate node. Similarly, constraint (7) is for loop prevention by avoiding the source node s to be treated as an intermediate node. Constraint (8) ensures that at most one lightpath between a specific *s*-d pair passes through link *(i, j)* using wavelength *w*. Constraints (9) and (10) are also for loop prevention at intermediates nodes to guarantee that no lightpath pass through node *j* and *i* twice, respectively. Constraint (11) ensures that the number of established lightpaths between any s-d pair is not more than the actual traffic demands between these two nodes. Constraint (12) ensures that at most one lightpath passes through the link *(i, j)* using the same wavelength. Constraint (13) is to prevent looping on the same link back and forth. Constraint (14) guarantees flow conservation, where the number of lightpaths entering a node is equal to the number of lightpaths exiting that node on wavelength *w*. The hop count for a given lightpath between an *s-d* pair is defined in (15), and it is limited by the optical domain diameter in (16). Constraint (17) is the wavelength continuity constraint.

It is worth to mention that the throughput is not concern here, the reason for this is that when the traffic load is high LVM can efficiently localize any failed link, however, when the traffic is low throughput becomes less important and the number of links that can be localized when applying the LVM becomes the objective.

4.2.2 Optimization for Fault Localization Time

In this section, we first analyze the delay for localizing a failed link using the LVM protocol and then formulate the optimization problem for fault localization time.

4.2.2.1 Delay Analysis

We divide the fault localization process into six different phases and then analyze the delay in each phase, respectively. The overall delay is thus the sum of the delays in all phases. These phases include the detection phase, broadcasting phase, multicasting phase, matching phase, responding phase, and concluding phase.

During the detection phase, the average location of a failure is at the middle of the average length of an affected lightpath. Hence, the delay to detect a failure by a sink node

can be expressed as $T_1 = \gamma \times H_{sd}/2$, where γ is the expected value of propagation delay along a link and H_{sd} is the length of the executive route in terms of the number of hops between a given *s-d* pair. Note that in this phase a node that detects the failure will set a timer to time out after $2 \times \gamma \times D$, where D is the diameter of the network. The reason for setting this timer will be explained later in this section.

In the broadcasting phase, a sink node will broadcast an ALARM message to the entire network, the maximum delay for all nodes to receive this message is $T_2 = \gamma \times D$. To ensure that all the nodes receive the ALARM message, each sink node that detects the failure will wait for a period of time equal to twice the diameter of the network, i.e., $2 \times \gamma \times D$.

Up to this point, if we add all the delays from phase 1 to phase 2, it will be equal to the timer value each node has set. However, this time does not reflect the total actual waiting time because if a node located on one side of the network detects the failure, it will start its timer to count down. On the other hand, another affected sink may be located on the other side of the network and cannot start its timer unless it detects the failure after $\gamma \times D$. Therefore, we should consider the time delay difference between the two nodes and the delay up to this phase should equal to the sum of the time delays of the previous two phases plus an extra time, i.e., $T_{extra} = \gamma \times D$.

During the timer period, an affected sink that generated an ALARM message has also received an ALARM message from each of the other affected potential executive sinks. Each node compares the length of the affected lightpath contained in each received ALARM message with the length of its affected lightpath. In other words, at the end of the timer period, each potential executive sink can reach a conclusion on which potential executive sink should act as the executive sink. Therefore, it is not necessary to consider the computational delay of this period of time because this can be done in parallel while each potential executive sink is receiving the ALRAM messages.

In the multicasting phase, the executive sink will define a perimeter area and create its ALV vector, and then broadcast this vector to all nodes in the perimeter area. Since the executive route can be any route between a source and a sink, the time taken for the ALV to reach all nodes is $T_3 = \gamma \times (H_{sd} + 1)$.

In the responding phase, each perimeter node will send its binary vector to the executive sink, the delay in this phase is $T_5 = \gamma \times (H_{sd} + 1)$.

In the matching phase, each node within the perimeter area will compare the received ALV with the lightpath terminating at it. A perimeter node needs to compare two vectors that have lightpaths connecting two different *s-d* pairs. Therefore, the computational complexity of comparing the two vectors in this phase is $O[(H_{sd})^2]$. If we use θ to denote the computational cycle, the time delay in this phase is $\theta \times (H_{sd})^2$.

In the concluding phase, the executive sink requires one computational cycle θ to determine the location of the failed link, and hence the processing delay of this phase is θ.

By summing the delays in all phases, we have the total delay for localizing a failed link, which can be expressed as

$$\begin{aligned} T &= T_{extra} + \sum_{i=1}^{6} T_i = \gamma \times D + \gamma \times \frac{1}{2} H_{sd} + \gamma \times D + 2\gamma \times (H_{sd} + 1) + \theta \times (H_{sd})^2 + \theta \\ &= 2\gamma \times D + 2\gamma + \frac{5}{2}\gamma \times H_{sd} + \theta[(H_{sd})^2 + 1] \end{aligned} \tag{18}$$

As we can see, the only variable in the above formula is H_{sd}. If for any given *s-d* pair, this variable can be minimized to the smallest number of hops between the *s-d* pairs, the localization time is minimized.

4.2.2.2 Problem Formulation of Time Optimization

We consider a static network where traffic demands (or lightpath requests) are known *a priori* in the form of a traffic matrix. Given the traffic matrix, the optimization problem considered is to optimally distribute the traffic (or lightpaths) in the network so that in the event of a link failure the fault localization time is minimized, i.e., for any given s-d pair H_{sd} is minimized. For this purpose, we formulate the problem into an integer linear programming problem. Similarly, we assume that a traffic matrix does not contain any direct traffic demand between neighboring nodes. The notations used in the formulation are defined as follows.

- ρ : the total number of connection requests or the offered load

- $U_{\max}^{ij}$: the maximum number of lightpaths established on a link (i, j)
- H_{sd}^{wr} : the number of hops traversed by a lightpath between source node s to destination node d on wavelength w to accommodate the connection request r
- H_{Ave} : the average number of hops
- Len_{sd}^{r} : the length of a lightpath in terms of the number of hops between source node s to destination node d for a connection request r.

Objective:

$$Min[H_{avg} = \frac{1}{\rho}\sum_{s}\sum_{d}\sum_{w}\sum_{r} H_{sd}^{wr}] \tag{19}$$

subject to

$$H_{sd}^{wr} = \sum_{i}\sum_{j} LP_{ij\_r}^{sd\_w}, \qquad \forall s,d,r,w \tag{20}$$

$$\sum_{W}\left(\sum_{i}\sum_{k,k\neq s,d} LP_{ik\_r}^{sdw} - \sum_{k,k\neq s,d}\sum_{j} LP_{kj\_r}^{sdw}\right) = 0 \quad \forall s,d,r \tag{21}$$

$$\sum_{w} H_{sd}^{wr} = Len_{sd}^{r}, \quad \forall s,d,r \tag{22}$$

$$Len_{sd}^{r} > 1, \quad \forall s,d,r \tag{23}$$

$$U_{sd}^{ij} = \sum_{r}\sum_{w} LP_{ij\_r}^{sd\_w}, \quad \forall s,d,i,j \tag{24}$$

$$\sum_{s'd'} U_{s'd'}^{ij} > U_{sd}^{ij} \tag{25}$$

$$LP_{dk\_r}^{sd\_w} = 0 \tag{26}$$

$$LP_{ms\_r}^{sd\_w} = 0 \tag{27}$$

$$\sum_{s}\sum_{d}\sum_{r} LP_{ij\_r}^{sd\_w} \leq 1, \quad \forall w,i,j \tag{28}$$

$$\sum_{w} LP_{ij\_r}^{sd\_w} + \sum_{w} LP_{ji\_r}^{sd\_w} \leq 1, \quad \forall i,j \tag{29}$$

$$\sum_{sd} U_{sd}^{ij} \leq U_{\max}^{ij} \leq W \tag{30}$$

$$\sum_{W}\sum_{i} LP^{sdw}_{ik\_r} - \sum_{W}\sum_{j} LP^{sdw}_{kj\_r} = \begin{cases} -1 & if\ s = k \\ 1 & if\ d = k \\ 0 & otherwise \end{cases} \qquad (31)$$

$$\sum_{w}\sum_{r} LP^{sd\_w}_{ij\_r} <= \lambda_{sd}, \quad \forall s,d,i,j \qquad (32)$$

where constraint (20) provides the number of hops traversed between a given source node s to destination node d on wavelength w for a given connection request r. Constraint (21) is wavelength continuity constraint. Constraint (22) guarantees that no two lightpaths between a given *s-d* pair utilizes the same wavelength. Since the traffic matrix does not contain any traffic demand between neighboring nodes, the lower bound on the number of hops traversed by a lightpath is given in constraint (23). Constraint (24) represents the link utilization. Constraint (25) guarantees that at most one lightpath between two different *s-d* pairs passes through link *(i,j)*. This constraint is actually introduced by the protocol because each link should have more than one lightpath that belongs to different *s-d* pairs for a successful localization. Constraints (26) and (27) stand for preventing loops in the routes of the connections. A capacity constraint is given in (28) ensuring that at most one connection passes through link *(i,j)* using the same wavelength. Constraint (29) prevents oscillation of a connection's route over a link. Constraint (30) is a capacity constraint that guarantees the number of lightpaths established on the link between node i to node j is less than the maximum number of wavelengths available on that link. Constraint (31) presents flow conservation. By (32), the number of lightpaths established between a given *s-d* pair is set to be equal to the traffic demands between the node pair.

4.3 Proposed Heuristic

To evaluate our optimization formulations, we present a heuristic algorithm that provides wavelength assignments in a way that satisfies the requirements of the LVM protocol. As we explained earlier, each link has to have at least two lightpaths that belong to two different *s-d* pair. Furthermore, using the shortest hop will assure minimum localization time at each link. The following notations are used in the proposed heuristic algorithm:

- W: the number of wavelengths or channels per link
- λ_f^{ij}: the number of free channels on link (i,j)
- λ_o^{ij}: the number of occupied channels on link (i,j)
- I_{sd}: the flow ID for a lightpath established between source node s to destination node d.
- O_{sd}^{ij}: the number of lightpaths that start and terminate at the same s-d pair that pass through link (i,j)
- C_{ij}: the cost of link (i,j)
- β: a binary value that allows the heuristic algorithm to operate in a non-greedy or greedy mode
- ε: a negligible value that forces the channel to be selected
- L: the total number of links in the network

For each request, the cost of all links has to be set according to the following cost function. Once the new cost for each link has been assigned, each node will run *Dijkstra*'s algorithm to find the shortest path for each destination, the selected routes will be selected based on the updated cost matrix for the links:

$$C_{ij} = \begin{Bmatrix} \infty & , & \lambda_o^{ij} = W \\ 1 & , & (\lambda_o^{ij} < W) \wedge (O_{SD}^{ij} > 0) \wedge (\lambda_f^{ij} = (W - O_{SD}^{ij})) \\ \beta & , & \lambda_f^{ij} = W \\ \varepsilon & , & Otherwise \end{Bmatrix}$$

As shown in the cost equation C_{ij}, if the number of occupied channels on that link is equal to the maximum number of wavelengths, the link cost will be set to infinity. On the other hand, if the number of occupied channels is less than the maximum number of wavelengths on that link, and the number of similar s-d pairs that pass through this link is more than 0, and similar s-d pairs are the only lightpaths that pass through this link, the corresponding cost for that link will be set to 1. The third condition states that, if the number of free channels is equal to the maximum number of wavelengths on that link, the cost of that link will be set to a binary value depending on which mode of operation is desired. Otherwise, the cost of the link is given a negligible value.

As we explained, for each request the link cost will be re-assigned. The complexity of updating the cost matrix is $O(L \times W)$. Once the link costs are updated, *Dijkstra*'s algorithm will be invoked to find the path with the lowest cost between these two *s-d* pairs as shown in Figure 4. 1.

```
function ConnProv(int source, int dest, link [] Link, int NumOfLinks)
      begin
            Counter ← 0
            NextLink = Link[0]

            While (Counter < NumOfLinks) do
                  Begin
                        if(NextLink.OccupiedWaves = W)
                        begin
                              NextLink.Cost = INFINITY

                        Endif
                        Else if(NextLink.OccupiedWaves.SD()=(source,dest) )
                              Begin
                                    NextLink.Cost = 1
                              Endif
                       Else if(NextLink.OccupiedWaves = 0 )
                              Begin
                                    NextLink.Cost = β
                              Endif
                        Else
                                    NextLink.Cost = ϵ


                  Counter ← Counter + 1

                  EndWhile

            Dijkstra()
            ProvisionConnections()
            UpdateBlockingRatio()

      end
```

Figure 4. 1: Quasi-codes of the heuristic algorithm

This heuristic provides two operation modes: greedy and non-greedy. In the greedy mode, the value of β is set to zero. When the heuristic operates in greedy mode, it attempts to reserve wavelengths on the idle links. Therefore the cost of those links are assigned to zero. If it is not successful in assigning wavelengths on idle links, it goes with the links on which the wavelengths are utilized by other s-d pairs. Thus, the link costs are assigned to a negligible value (ε) close to zero.

When the heuristic operates in non-greedy mode, it does not distinguish the idle links and the links occupied by the same s-d pair of the incoming connection. Therefore it sets the

link costs to one in this situation. On the other hands, it gives the priority to the links that are occupied by other *s-d* pairs. In other words, the non-greedy mode prefers choosing occupied links. Hence, the non-greedy mode is good for reducing resource consumption. Note that we use greedy (non-greedy) approach and greedy (non-greedy) heuristic interchangeably. Figure 4 describes the procedures of the heuristic for provisioning a connection between a given *s-d* pair in quasi-codes.

The heuristic algorithm requires that the cost of all links has to be modified as follows:

- If there are no free wavelengths on that link, the cost for that link will be set to infinity.
- If the occupied wavelengths that pass through this link originated and terminated at the same *s-d* pair, the cost is set to 1.
- If all wavelengths are free, the cost will be set to a binary value β.
- Otherwise, the cost will be set to a negligible value ε.

Once the link costs are assigned, Dijkstra algorithm will be performed to find the path with the minimum cost. Finally, if there are available recourses, the connection will be provisioned. Otherwise a blocking is announced.

4.4 Numerical Results

In this section, we compare the optimization results for both the optimization problems with that for the proposed heuristic algorithm.

4.4.1 Optimization for Fault Localization Probability

In this section, we use CPLEX 9.0 optimization tool [CPLEX] running on P4-2GHz machine to solve the formulated optimization problem and analyze its optimization performance. To successfully localize a failed link using the LVM protocol, the link should have at least two lightpaths that belong to two different *s-d* pairs. If only one lightpath passes through a given link, we refer to this link as a fault-prone link because the status of this link cannot be determined by the perimeter nodes while localizing the failure. If no lightpath passes through a given link, the link is referred to as an idle link.

We define a traffic load of one lightpath demand between each pair of non-neighboring nodes as 100% load and applied different traffic loads to both NSFnet and 3x3 grid, including 25%, 36%, 50%, 75% and 100% load. Each traffic load (or matrix) is obtained by randomly removing lightpath demands from the 100% load matrix. Moreover, we use the fault localization probability as the performance metric. We assume that $L_{localized}$ is the total number of localized links in the event of link failures, L_{total} is the total number of links in the network, and L_{idle} is the number of idle links in the network. The fault localization probability P is defined as

$$P = \frac{L_{localized}}{L_{total} - L_{idle}} \qquad (33).$$

Figure 4. 2 shows the fault localization probability with NSFnet and 3x3 grid, respectively. It is seen that a smaller load results in a lower fault-localization probability. For each different load, we run the optimization 5 times. We plot the average of the five runs presenting 90% Confidence Interval (CI) for comparison of the coverage ratios with the two networks. The procedure of how to estimate the confidence interval is explained in Appendix A.

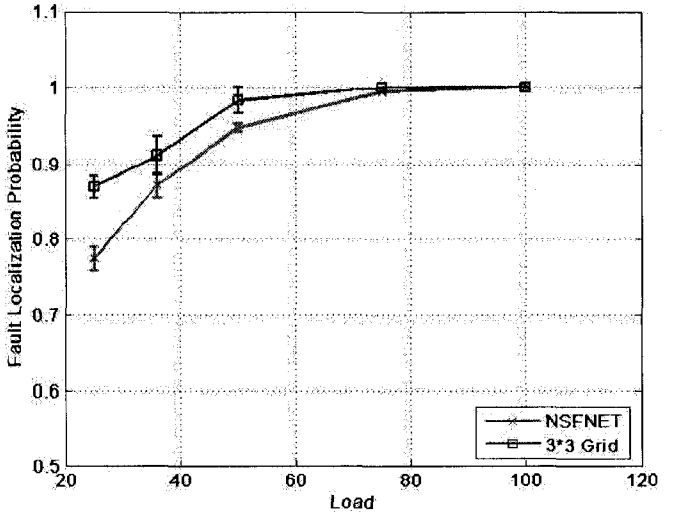

Figure 4. 2: Fault localization probability with 3x3 grid and NSFnet

4.4.1.1 3x3 Grid

We consider a grid network with $M \times N$ nodes, as shown in Figure 4. 3.

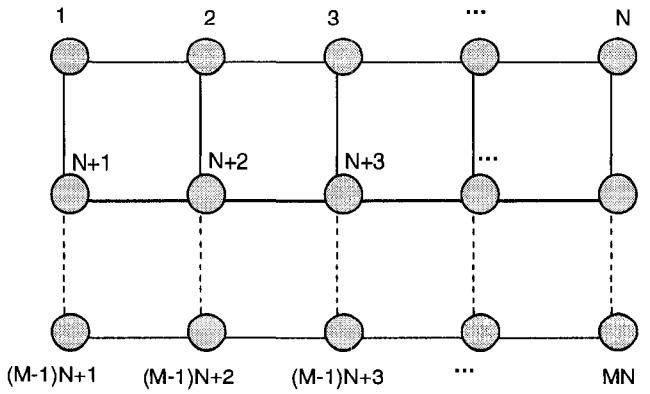

Figure 4. 3: Grid network

Figure 4. 4 shows the relationship among the network load, the number of fault-prone links, and the number of idle links. As we can see, the higher the load is, the lower the number of fault-prone and idle links.

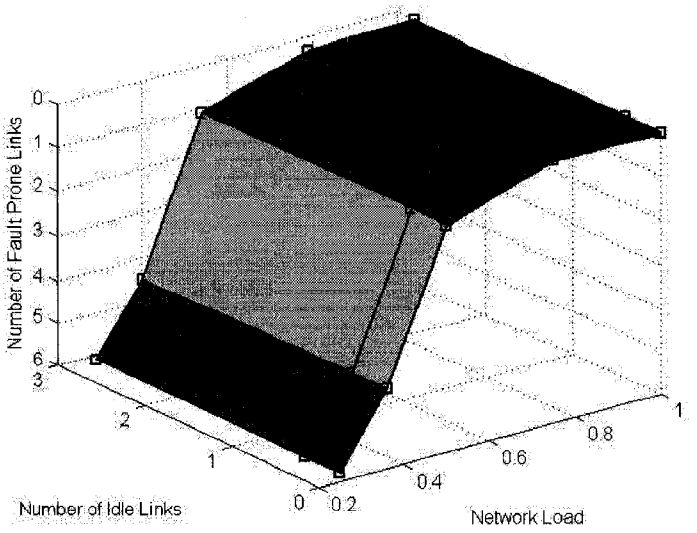

Figure 4. 4: Idle and fault-prone links versus network load with 3x3 grid

Figure 4. 5 shows the relationship among the network load, the maximum number of wavelengths, and the number of idle links. It is seen that as the load increases the number of idle links decreases because more lightpaths are established in the network. Moreover, as the load increases, the maximum number of wavelengths on each link increases as well.

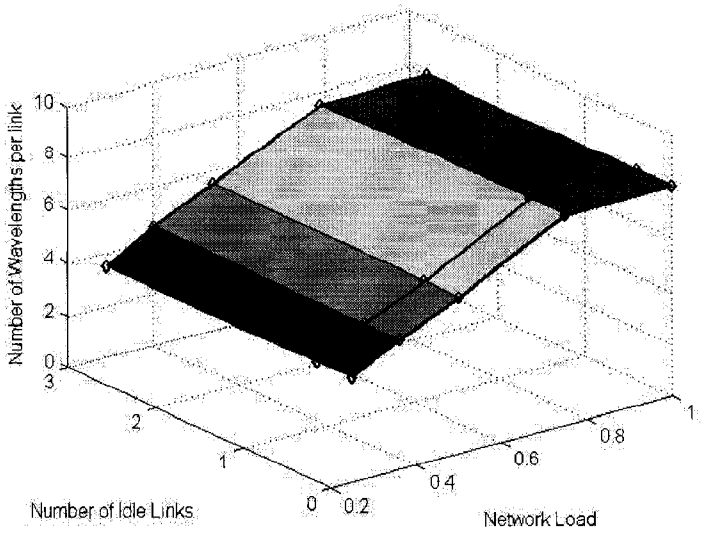

Figure 4. 5: Wavelengths and idle links versus network load with 3x3 grid

4.4.1.2 NSFnet

Figure 4. 6 and Figure 4. 7 show the results for NSFnet under 75%, 36%, and 25% load. In Figure 4. 6, it is seen that as the load increases the number of idle links and the number of fault-prone links decrease. In Figure 4. 7, it is seen that as the load increases the maximum number of wavelengths on each link increases.

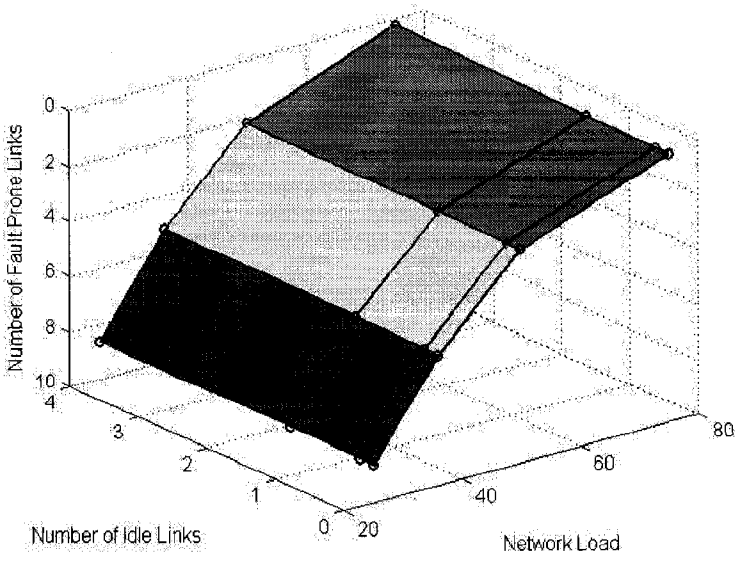

Figure 4. 6: Idle and fault-prone links versus network load with NSFnet

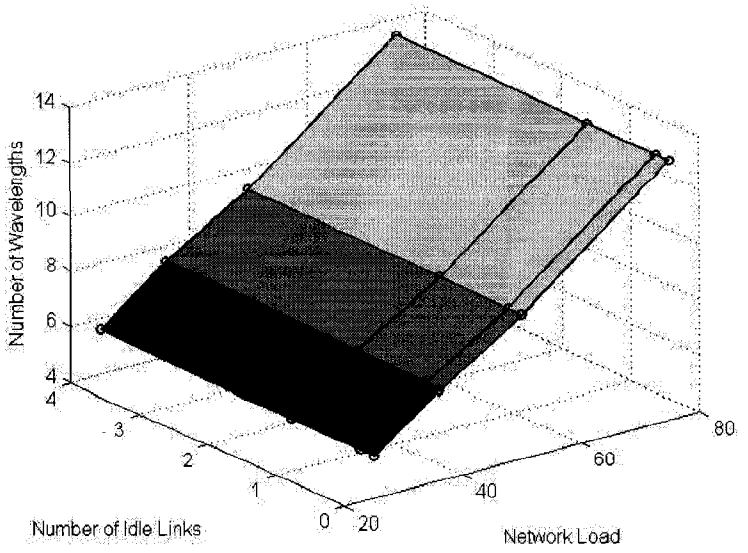

Figure 4. 7: Wavelengths and idle links versus network load with NSFnet

4.4.2 Optimization for Fault Localization Time

In this section, we use CPLEX 9.0 optimization tool [CPLEX] running on P4-2GHz machine to solve the formulated optimization problem and analyze its optimization performance. We

consider a grid network with 2, 4, 9, 16 nodes, respectively. For simplicity, we assume that the number of rows (N) is equal to the number of columns (M) in the network. In that case, the average localization time can be calculated as

$$T_{avg} = 4\gamma \times N - 2\gamma + \frac{5}{2}\gamma \times H_{avg} + \theta[(H_{avg})^2 + 1].$$

We also assume that the fiber length between nodes is 100km, the light speed in fiber is *(2.14 x 10^8 m/s)*, the computational cycle θ is 20ns, and the traffic matrix has no direct traffic between neighboring nodes. To obtain the numerical results, we apply a traffic matrix of two demands between any non-neighboring nodes.

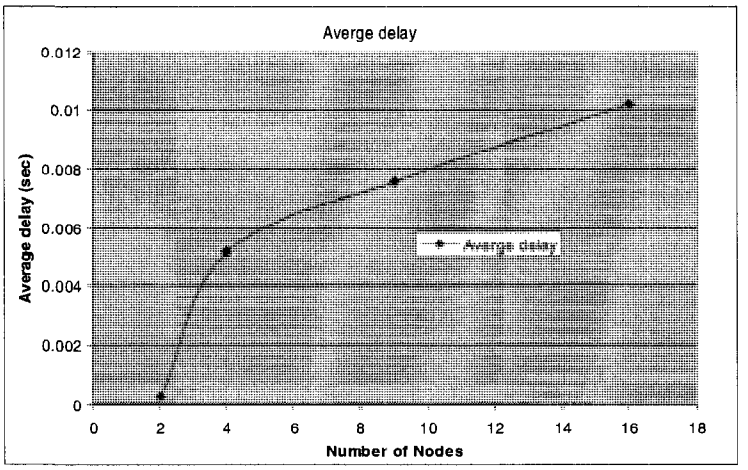

Figure 4. 8: Average localization time versus number of nodes

Figure 4. 8 shows the average localization time versus the number of nodes in the network. For a network with two nodes, there is no need to apply the LVM protocol because a failure can be detected and localized simply by a sink node. Thus, the time to localize such a failure is equal to the delay in the detection phase, i.e., $T_1 = \gamma \times H_{sd}/2$, and it is 0.234 ms. It is seen that with the number of nodes increasing the average localization time also increases. What is more interesting is that the average localization time is below 40 sec, which is the localization time needed to localize a failure when using OSPF in IP traditional networks [Moy et al 98].

Table 4. 1 shows the optimization results with different number of nodes, including the average hop count of a *s-d* pair H_{avg}, the number of wavelengths needed to make the

optimization, and the time taken to complete each run of optimization. It is seen that as the number of nodes increases more wavelengths are needed to make the optimization. Meanwhile, the average hop count of a *s-d* pair that minimizes the fault localization time also increases. In addition, it takes more time to complete the optimization.

Table 4. 1: Optimization results

Nodes	H_{avg}	Optimization Time (sec)	Number of wavelengths
2	1	0	1
4	2	0	4
9	2.50000008	48.23	10
16	3.083333373	85447.83	32
14 "NSF"	2.5214285916	4394.94	23

It is worth to mention that we have tried the optimization models for bigger networks. Unfortunately, the processing time was in the range of days, furthermore, larger memory were needed to handle the optimization models.

4.4.3 Performance Evaluation of the Heuristic in terms of Fault Localization Probability

In this section, we compare the optimization results for fault localization probability with the results using the heuristic for both the NSFnet and 3x3 grid networks.

Similar to the optimization environment, we applied a network load of 1 demand between non-neighbor nodes. We reduce the traffic load by a percentage and run the simulation five times under each load. At each load percentage, we generate different traffic distribution at each run. We have estimated the 90% CI for the five runs as explained in Appendix A, we have found that the 90% CI is small and hence can be removed from the graphs shown in this section.

Once we provision the load using First-Fit wavelength assignment, we search for all links that has no lightpaths passing through it.(i.e., idle links). Then we search for error-prone links that has one lightpath passing through it. After that, we search for the link that

has the maximum number of wavelengths. Finally, we generate a failure at each link and run the LVM to see if the failed link can be localized given the traffic distribution.

We run the simulation using the non-greedy and greedy heuristics, respectively. We assumed that the negligible value ε is 10^{-5}.

4.4.3.1 Performance Comparison for 3x3 Grid network

Figure 4. 9 shows the localization link ratio defined in equation 33 with the network load. As it is shown clearly, it shows that the optimization results and the results using the greedy and non-greedy heuristics, respectively. It is obvious that the results obtained using CPLEX is the best because the lightpaths are provisioned in such a way that ensures the maximum number of links that can be localized when a failure occurs.

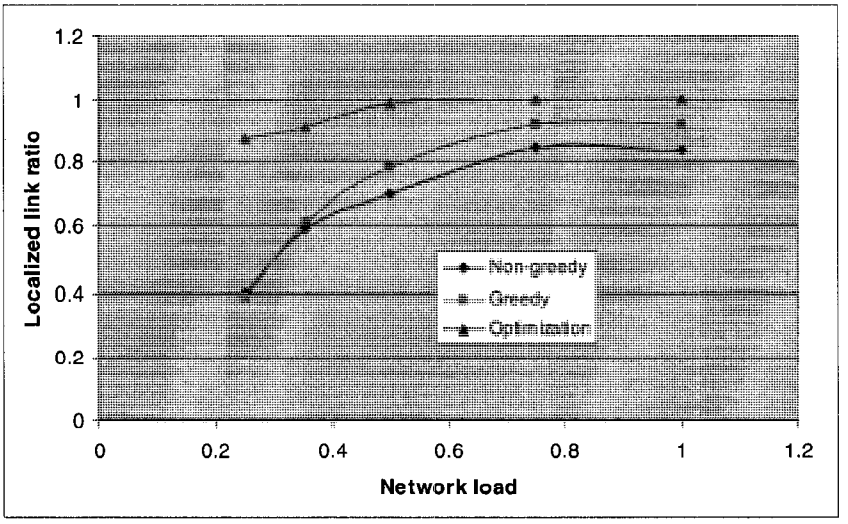

Figure 4. 9: Localized link ratio versus network load

Figure 4. 10 shows the acceptance rate under different network loads. The optimization results have the highest acceptance rate because all the traffic demands have been provisioned. The non-greedy heuristic also has a high acceptance, almost similar to the optimization results. However, we will see how efficient it will be in-terms of localizing the failed links. Moreover, the greedy heuristic has a lower rate of acceptance

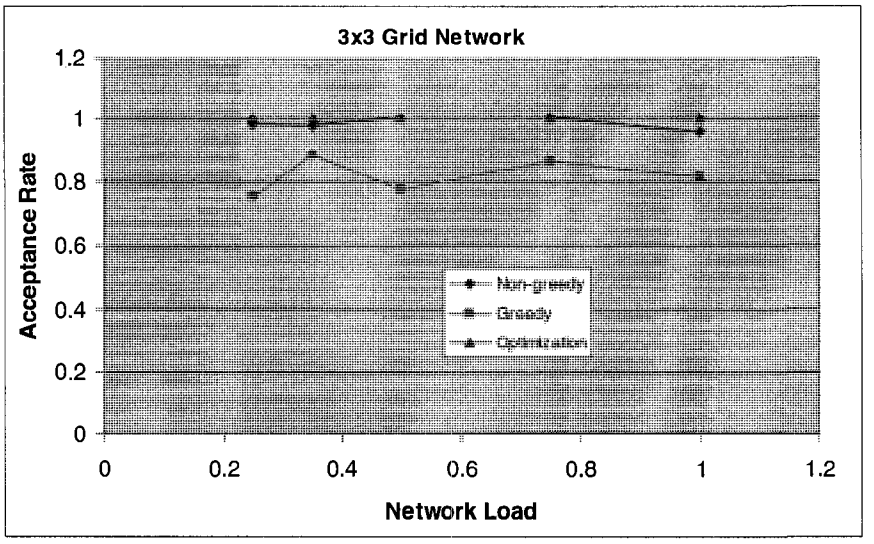

Figure 4. 10: Acceptance rate versus network load

Figure 4. 11 shows the number of idle links in the network. It is obvious that the non-greedy heuristic has the largest number of idle links. The reason is that the non-greedy heuristic tries to utilize already reserved links. Although we have seen that it has a high rate of acceptance, it leaves many links un-used and hence these idle links cannot be localized.

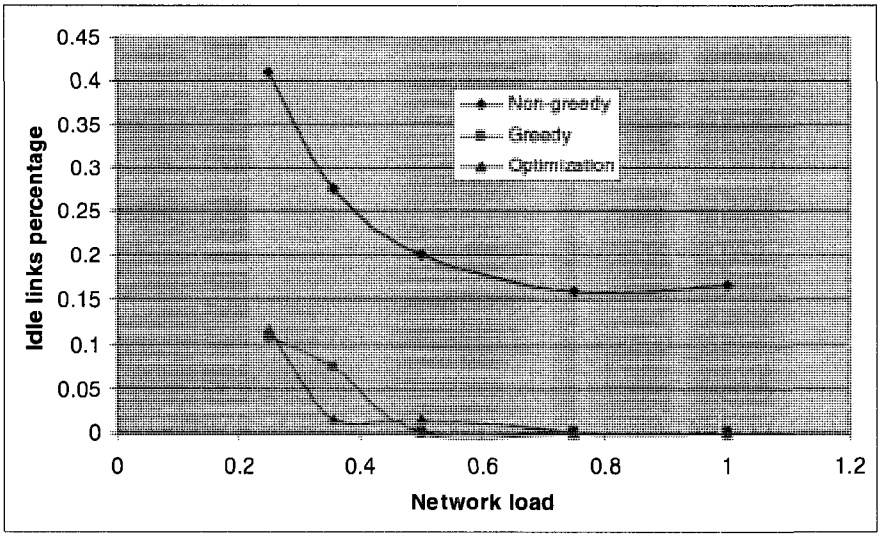

Figure 4. 11: Idle links versus network load

On the other hand, the greedy heuristic has a smaller number of idle links almost close to the optimization results due to the fact that the greedy heuristic assigns zero to idle

links. However, as we observed earlier, it has a low acceptance rate that prevents some of the connections from being provisioned.

Figure 4. 12 shows that the greedy heuristic has the highest percentage of error-prone links as the traffic decreases. On the other hand, the non-greedy heuristic has a smaller number of error-prone links and at the lowest traffic point it has the smallest number of error-prone links. This is due to the fact that the non-greedy heuristic leaves a huge number of links in the idle state and hence what is left will have most of the traffic concentrated on them, leaving a smaller number of error-prone links. It is seen that the optimization results obtained by CPLEX have the highest localized link ratio, highest acceptance rate, and as a result the minimum number of idle and error-prone links under all different loads. The non-greedy heuristic has a high acceptance rate but it leaves a large number of unoccupied links which can not be localized by LVM protocol. The greedy heuristic has a better localization ratio than the non-greedy one, but it does not allow all connections to be provisioned and still has higher idle links and error-prone links compared to the optimization results.

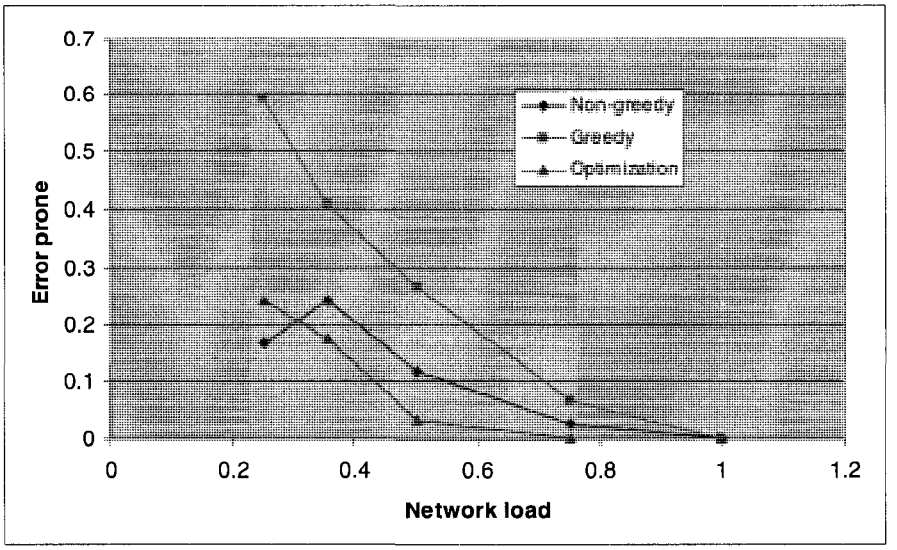

Figure 4. 12: Error prone versus network load

4.4.3.2 Performance Comparison for NSFnet

Figure 4. 13 shows the localized link ratio with NSFnet. It is seen that the localization ratio for the greedy heuristic has relatively better results than the optimization results at two points.

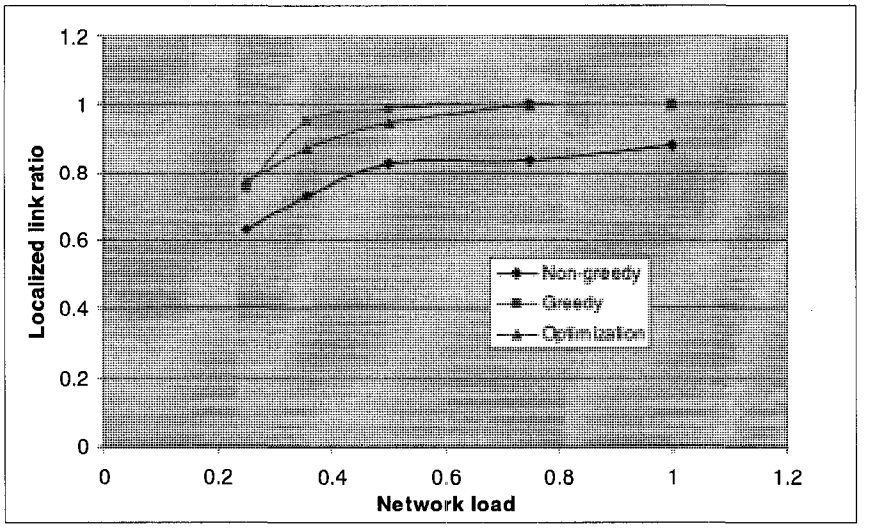

Figure 4. 13: Localization ratio versus network load

The greedy heuristic seems to perform better than the optimization. However, if we consider the fact that the optimization results were obtained based on the fact that NSFnet's diameter is 3 [BC92]; hence no lightpath that has more than three hops to be provisioned. Figure 4. 14 shows the average number of hops with the optimization, and the greedy and non-greedy heuristics. It is clearly shown that the optimization has the lowest number of hops as it has the lowest average value. This means that a smaller number of links are allowed to be used. And based on these constraints the average number of hops is 2.143 [NIST]. On the other hand, the limit is open for the number of hops that can be used to provision any connection in the heuristics. Therefore, the average number of hops is more than 2.5 for greedy and almost 3 for non-greedy.

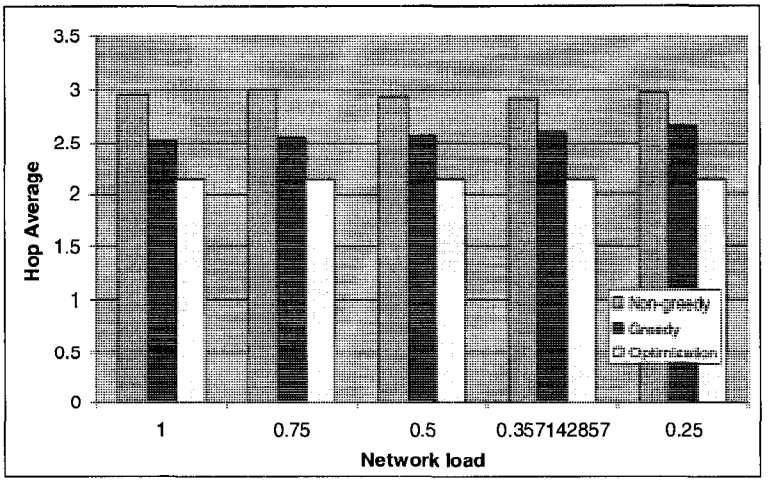

Figure 4. 14: Average number of hopes versus network load

Figure 4. 15 shows the acceptance rate of the optimization, and the greedy and non-greedy heuristics. It is obvious that the greedy heuristic has the lowest acceptance rate. It tries to use idle links. Hence, blocking is more likely to happen. However, the non-greedy one tries to avoid idle links but to use the already occupied links, leaving more resources available for incoming connections.

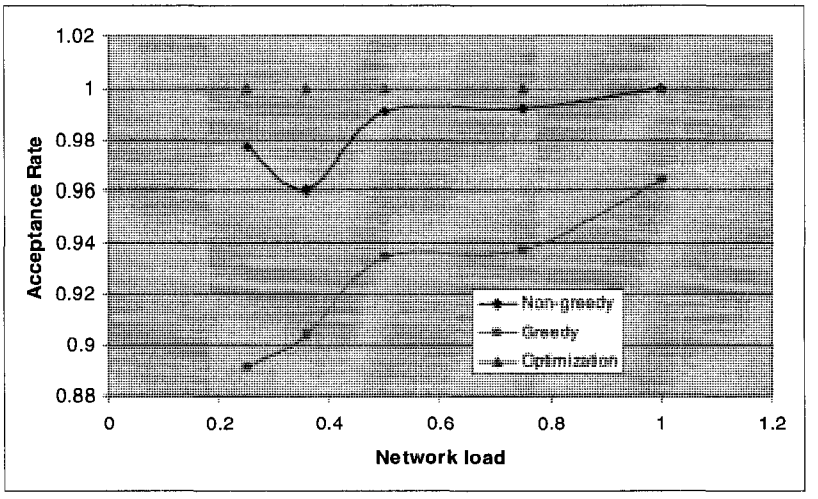

Figure 4. 15: Acceptance rate versus network load

Figure 4. 16 shows the idle links percentage. Because the non-greedy heuristic tries to utilize the already used links, it has the largest number of idle links. For the same reason,

the greedy heuristic seems to have better results because more links are being covered. However, this is not true because the optimization is only allowed to provision connections with maximum three hops and hence more idle links will exist compared to the greedy heuristic, which allows to provision connections with hops more than 3.

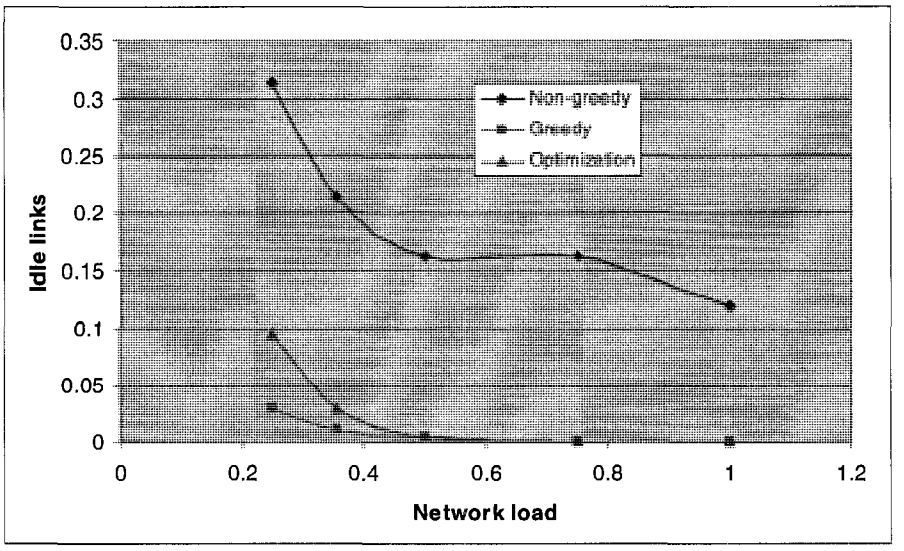

Figure 4. 16: Idle links percentage versus network load

Figure 4. 17 shows the error-prone links. It is seen that for the 3x3 grid, the non-greedy heuristic has the smallest number of error-prone links. The greedy heuristic tries to utilize the idle links. Hence, more error-prone links exist compared to the optimization.

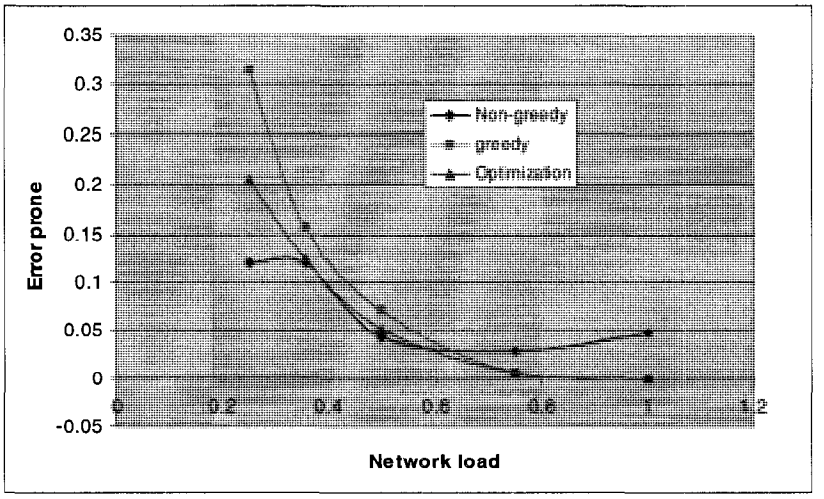

Figure 4. 17: Error prone links versus network load

To compare link utilization with the optimization, and the greedy and non-greedy heuristics, Table 4. 2 shows the average of the maximum number of wavelengths or link utilization needed at each link. It is seen that the non-greedy heuristic has the largest number of wavelengths because it tries to occupy the already occupied links. The greedy heuristic and the optimization have almost similar results.

Table 4. 2: Maximum link utilization

	3x3 Grid network				
Load percentage	100%	75%	50%	36%	25%
Optimization	8	7.5	5.6	5	3.8
Non-greedy	10	8.8	7.4	5.4	4.4
Greedy	6	5.8	5	3.6	3
	NSF				
Load percentage	100%	75%	50%	36%	25%
Optimization		13.4	9.2	7.4	5.6
Non-greedy	27	21.4	16.4	11.2	7.8
Greedy	15	12.4	8.4	7.4	4.8

4.4.4 Performance Evaluation of the Heuristic in terms of Fault Localization Time

In this section, we present the optimization results for minimizing the fault localization time using the LVM protocol. The objective is to find the best traffic distribution so that the localization time for any failed link is minimized while maximizing the acceptance of provisioned connections.

Similar to the optimization environment, we have applied a network load of 2 demands between non-neighboring nodes. Each point in the graphs corresponds to the average of five runs. We have also estimated the 90% CI for the five runs as explained in

Appendix A, we have found that the 90% CI is small and hence can be removed from the graphs shown in this section.

In each run, we provision the static load uniformly. Once the load is provisioned, we search for idle links. Then we search for error-prone links, and finally we search for the link that has the maximum number of wavelengths. We generate a failure at each link and run the LVM to see if the failed link can be localized given the traffic distribution. We run the simulation using the non-greedy and greedy heuristics, respective. We assumed the negligible value ε is 1e-5.

Figure 4. 18 shows the average hop count with the traffic distribution of the optimization, and the greedy and non-greedy heuristics, respectively. It is seen that the average hop count provides the minimum localization time at each link.

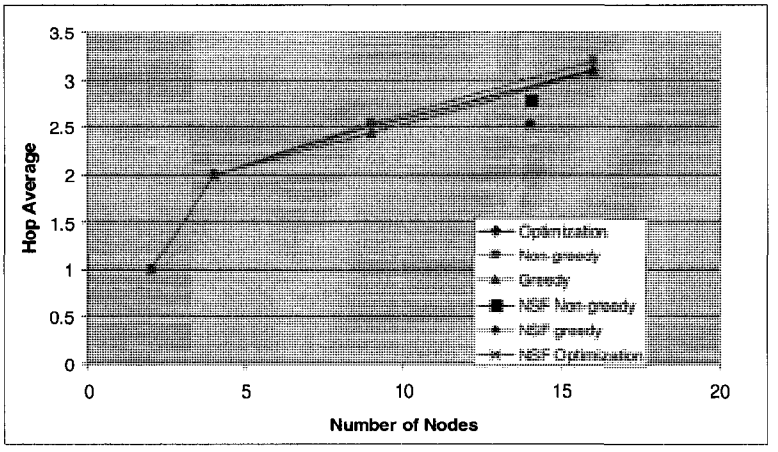

Figure 4. 18: Hop average versus number of network nodes

One might argue that there is not a significant difference between the optimization, and the greedy and non-greedy heuristics because all have similar results. However, both the greedy and non-greedy heuristics have a low acceptance ratio compared to the optimization. Figure 4. 19 shows that the acceptance rate is as low as 0.6 when the number of nodes is 9.

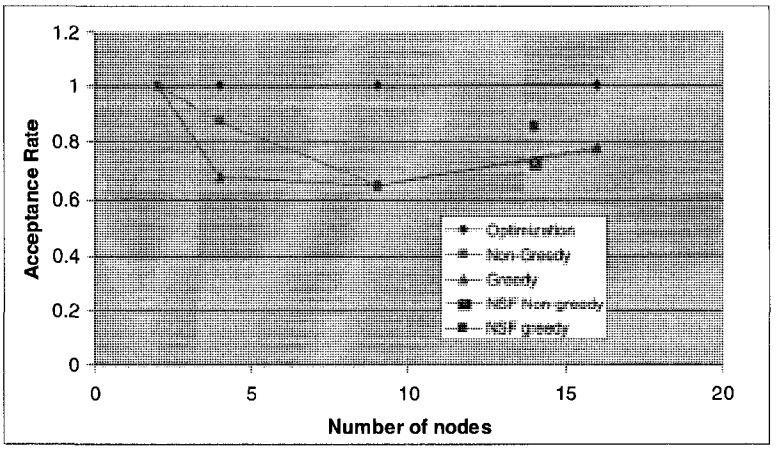

Figure 4. 19: Acceptance rate versus number of network nodes

There are still other issues that we have not considered yet. Both the greedy and non-greedy heuristics do not guarantee successful localization for all links despite the large number of traffic demands. Table 4. **3** shows the localization ratio with greedy and non-greedy.

Table 4. 3: Localization ratio with greedy and non-greedy

	Greedy			
Nodes	Idle Links	Error prone	Localized	Un-Localized
1	0	0	1	0
4	0	2.8	1	0
9	0	0.6	0.96667	0.033333333
14	0	0	1	0
16	0	0	1	0
	Non-greedy			
Nodes	Idle Links	Error prone	Localized	Un-Localized
1	0	0	1	0
4	0.25	1	0.65	0.1
9	0.04167	0.8	0.89167	0.066666667

14	0.01905	0	0.98095	0
16	0.0375	0.8	0.95417	0.008333333

It is worth to mention that larger network sizes can be considered using the heuristic algorithm. Both greedy and non-greedy mode can be used as both are scaleable. However, due to comparison we only used the same network sizes that have been used in the optimization models.

4.5 Summary

In this chapter, we studied the optimization problems in applying the LVM protocol for localizing single-link failures in static all-optical networks. Two optimization problems were considered: the problem for maximizing the fault localization probability and the problem for minimizing the fault localization time. Given the traffic demands, the former problem is to optimize the traffic distribution so that the fault localization probability in terms of the number of localized links is maximized while the latter is to optimize the traffic distribution so that the fault localization time for each failed link is minimized. We formulized the problems into an integer linear programming problem, respectively, and use CPLEX to solve the problems. Through numerical results, we showed that by optimizing the traffic distribution the fault localization probability can be maximized and the solution to the problem can also provide the maximum number of wavelengths needed on each link to obtain the maximum localization probability. Moreover, the fault localization time can be effectively minimized and is below the OSPF localization time (40sec) in traditional IP networks. Finally, we proposed a heuristic algorithm that can be used to evaluate the optimization results. This heuristic algorithm tries to satisfy the LVM protocol's conditions through assigning appropriate costs to the network links to allow different *s-d* pairs to utilize the same link. It has two operation modes: greedy and non-greedy. Both the greedy and non-greedy heuristics provide link coverage but not as well as the optimization results. The heuristics also try to distribute the traffic to minimize the localization time at each link

For future work, we will extend the proposed optimization and heuristics to a dynamic traffic environment. We will also consider partial and full wavelength-conversion issues for fault localization-constrained connection provisioning.

Chapter 5 Inter–domain Fault Localization

5.1 Introduction

Fault localization is a prerequisite for fault protection and restoration [K07, JS06, HZ06, M05]. Even more challenging, it is more difficult to localize a failure in a multi-domain scenario because in such a scenario the internal confidential domain-specific information (e.g., topology and traffic) on a domain is not allowed to be available to other domains.

In this chapter, we propose a fault localization protocol for localizing single-link failures in multi-domain all-optical networks. The protocol is based on a limited-perimeter vector matching (LVM) mechanism, which restricts fault localization within a smaller perimeter area and can thus significantly reduce the time and space complexity of fault localization. [SM07]. By assuming power monitoring available only at edge nodes, this protocol can localize both inter-domain and intra-domain link failures without exchanging any internal confidential domain-specific information between different domains. We show through analytical results that it can not only fast localize an inter-domain link failure between different domains but also localize an intra-domain link failure that affects inter-domain traffic faster than the Open-Shortest-Path-First (OSPF) protocol in a large network.

Furthermore, we also propose a fault localization protocol for localizing single-link failures in multi-domain all-optical networks with no optical power monitoring available at each node. The protocol is based on a limited-perimeter vector matching (LVM) mechanism, which restricts fault localization within a smaller perimeter area and can thus significantly reduce the time and space complexity of fault localization [SM07]. By performing a fault localization process sequentially in the optical domains that an affected lightpath passes through, the protocol can localize both inter-domain and intra-domain link failures that affect inter-domain traffic without exchanging any internal confidential domain-specific

information between different domains. We show through analytical results that it can not only fast localize an inter-domain link failure between different domains but also localize an intra-domain link failure that affects inter-domain traffic faster than the Open-Shortest-Path-First (OSPF) protocol in a large network.

5.2 Distributed Fault Localization for Multi-Domain All-Optical Networks with Partial Power Monitoring

In this section, we present a fault localization protocol for localizing single-link failures in multi-domain all-optical networks.

5.2.1 Protocol Description

We consider a multi-domain all-optical network with no power monitoring available at each intra-domain nodes and only edge nodes are equipped with power monitors. Each fiber has multiple wavelengths and control messages are delivered out-of-band.

The fault localization protocol can be divided into two phases: inter-domain localization phase and intra-domain localization phase. The inter-domain localization phase is intended to localize a faulty domain or an inter-domain link failure, while the intra-domain localization phase is intended to localize a link failure within a faulty domain. The procedures in each phase are described as follows.

5.2.1.1 Case A: Inter-domain Localization Phase

1) Once a link fails, all lightpaths passing through the failed link will be disrupted. As a result, the destination of each of the disrupted lightpaths will detect the signal loss of that lightpath and thus the failure. Meanwhile, all edge nodes after the failed link on a disrupted lightpath will also detect the signal loss of that lightpath.
2) Once a destination detects the failure, it will broadcast an INTRA_ALARM message to all nodes within its domain, which contains its node ID and the length of the shortest affected lightpath terminating at it. At the same time, it will start a timer set

to a pre-determined period of time (2D) and then wait until it receives a SUPPRESS message or the timer times out.

3) Once an ingress edge node detects the failure, it will send an INTER_ALARM message to its upstream edge node. The INTER_ALARM message contains the domain ID, the node ID, and the IDs of all disrupted lightpaths coming from the upstream edge node. At the same time, if it is within the same domain as the destination of a disrupted lightpath, it will send a SUPPRESS message to the destination of that disrupted lightpath within the domain.
4) Once an egress edge node detects the failure, it will broadcast an INTER_ALARM message to all nodes within its domain, including its upstream edge node of each disrupted lightpath passing through it. The INTER_ALARM message contains the domain ID, the node ID, and the IDs of all disrupted lightpaths coming from the upstream edge node.
5) If an unaffected edge node receives an INTER_ALARM message from its downstream edge node with a different domain ID, it can conclude that the failure is on the inter-domain link from itself to the downstream edge node. In this case, the unaffected edge node will broadcast an INTER_FAULT_LOCATION message to all edge nodes in the multi-domain network notifying them of the failed inter-domain link.
6) If an unaffected edge node receives an INTER_ALARM message from its downstream edge node with the same domain ID, it can conclude that the failure is within its own domain. In this case, it will send a GO_AHEAD message back to the downstream edge node which sent the INTER-ALARM message. At the same time, it will broadcast an INTER_FAULT_LOCATION message to all edge nodes in the network.
7) If an affected sink receives a SUPPRESS message, it knows that the failure is not within its own domain and will stop any reaction for localizing the failure.
8) If an affected sink receives an INTRA_ALARM message before its timer times out, it will simply store the message.
9) If an affected sink receives an INTER_ALARM message before its timer times out, it will extend the timer for additional 2D and then continue to wait until the timer

times out.

10) If the timer of an affected sink times out with no INTER_ALARM message received, it means that the failure is within the domain of the affected sink and there is no inter-domain lightpath affected. In this case, the localization process goes into the intra-domain localization phase with no inter-domain lightpaths affected.

11) If an egress edge node receives a GO_AHEAD message after sending an INTER_ALARM message, the localization process goes into the intra-domain localization phase with inter-domain lightpaths affected.

Note that it is possible that there are multiple affected edge nodes and unaffected edge nodes within the faulty domain. In this case, an unaffected edge node can be viewed as a virtual source and an affected edge node can be viewed as a virtual destination or sink.

5.2.1.2 Case B: Intra-domain Localization Phase with Inter-Domain Lightpaths Affected

1) If an affected egress edge node receives a GO_AHEAD message after sending an INTER_ALARM message, it will broadcast an INTRA_ALARM message to all other nodes within the domain, which contains its node ID and the length of the shortest affected lightpath passing through it. At the same time, it will start a timer set to a pre-determined period of time (D). If the source of a lightpath is not within the same domain, only the partial path between the virtual source and the affected sink of the lightpath or the partial path between the virtual source and the virtual sink of the lightpath is considered when computing the path length. As a result, each affected or virtual sink will receive the INTRA_ALARM messages sent by the other affected and virtual sinks within the faulty domain.
2) Once the timer of an affected or virtual sink times out, the sink will compare the length of its own shortest path with the lengths of the shortest paths of the other affected and virtual sinks contained in the received INTRA_ALARM messages. The affected or virtual sink with the shortest affected lightpath will become the executive sink.
3) The executive sink will then create a vector of links on its shortest affected lightpath, called Affected-Link-Vector (ALV), defines a limited-perimeter, which consists of all the ALV links and their neighboring nodes, and then multicasts the ALV to all the nodes

within the limited-perimeter.

4) Each sink or virtual sink receiving the ALV within the limited-perimeter will create its own link vector and compares the vector with the ALV. If its link vector has a common link with the ALV, the sink will send a binary vector back to the executive sink. The binary vector consists of a matching vector and the status of the corresponding lightpath. The matching vector contains the same links in the ALV. If the lightpath is working, its status is set to "1". Otherwise, the status is set to "0". For a working path, an element in the matching vector is set to "0" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "1". For a failed path, an element in the matching vector is set to "1" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "0".
5) After the executive sink collects the binary vectors from all the sinks within the limited-perimeter, it will perform a logical AND operation on all the collected binary vectors to determine the location of the failure. The link with "1" in the resulted binary vector will be determined as the failed link. If there are more than "1" in the resulted binary vector, the executive sink can extend the limited-perimeter by including the neighbours of the original perimeter nodes and run a second round of the above process.
6) Once the failed link is localized, the executive sink will broadcast an INTRA_FAULT_LOCATION message to all nodes within its domain, notifying them of the fault location.

5.2.1.3 Case C: Intra-domain Localization Phase with no Inter-Domain Lightpaths Affected

1) Once the timer of an affected sink times out with no INTER_ALARM message received, the sink will compare the length of its own shortest path with the lengths of the shortest paths of the other affected sinks contained in the received INTRA_ALARM messages. The affected sink with the shortest affected lightpath will become the executive sink.
2) The executive sink will then create a vector of links on its shortest affected lightpath, called Affected-Link-Vector (ALV), defines a limited-perimeter, which consists of all the ALV links and their neighboring nodes, and then multicasts the ALV to all the nodes

within the limited-perimeter.

3) Each sink receiving the ALV within the limited-perimeter will create its own link vector and compares the vector with the ALV. If its link vector has a common link with the ALV, the sink will send a binary vector back to the executive sink. The binary vector consists of a matching vector and the status of the corresponding lightpath. The matching vector contains the same links in the ALV. If the lightpath is working, its status is set to "1". Otherwise, the status is set to "0". For a working path, an element in the matching vector is set to "0" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "1". For a failed path, an element in the matching vector is set to "1" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "0".
4) After the executive sink collects the binary vectors from all the sinks within the limited-perimeter, it will perform a logical AND operation on all the collected binary vectors to determine the location of the failure. The link with "1" in the resulted binary vector will be determined as the failed link. If there are more than "1" in the resulted binary vector, the executive sink can extend the limited-perimeter by including the neighbours of the original perimeter nodes and run a second round of the above process.
5) Once the failed link is localized, the executive sink will broadcast an INTRA_FAULT_LOCATION message to all nodes within its domain, notifying them of the fault location.

5.2.2 Illustrative Example

Figure 5. 1 illustrates the main operation procedures of the proposed fault localization protocol. In this example, we consider two link failures: One is in optical domain 3 (OD-3) whereas the other is between edge nodes A and B. Since the protocol can only handle a single failure at a time, we consider the failure on link A-B first. Once this failure occurs, the sink node in OD-2, OD-4 and OD-5 will detect it. In addition, edge nodes B, C, D, G, H, E, and F will also detect the failure because they are edge nodes with power monitoring devices. The subsequent procedures are as follows.

- The first nodes to respond are the edge nodes close to the sinks. In this example, edge nodes D, H, and F will broadcast a SUPPRESS message to all the nodes belongs to their domains, respectively, because they realize that the failure is not within their domains.
- At the same time, each edge node that detects the failure will send an INTER_Domain alarm message to the upstream edge node inquiring whether the upstream node detects the failure. For example, edge node D will send a message to edge node C and at the same time H to G, G to B, C to B, F to E, E to B, and B to A.
- In response, the upstream nodes will reply to the downstream indicating whether it also detects the failure. In this example, edge node A will be the only one that will reply with a message indicating no power loss. This information can be used at edge node B to conclude that link A-B is down. Then edge node B will advertise the fault location "Link A-B" to all other domains.

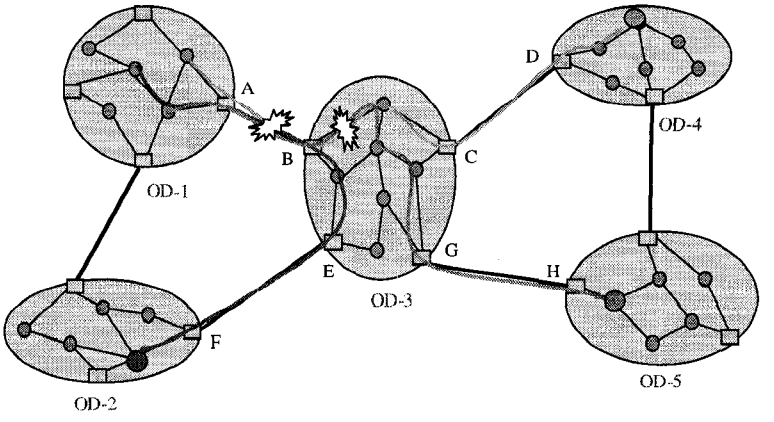

Figure 5. 1: Inter-domain failure example

Now let us consider the link failure within OD-3. Once it occurs, the sink nodes in OD-4 and OD-5 will detect the failure. On the other hand, edge nodes C, D, G, and H will also detect the failure. In this case, edge nodes D and H will broadcast a SUPPRESS message to all nodes in their domains preventing them from taking further action. At the same time, each edge node will send an INTER_domain alarm message to the upstream node, i.e., D to C, H to G, G to B, and C to B. As a result, both G and C will receive a message from edge node B indicating no power loss. Thus, these two nodes will realize that the failure is within OD-3, and then advertise this failure information to all other optical

domains, indicating that OD-3 is the faulty domain. In the mean time, OD-3 will continue the fault localization process to locate the failed link.

5.2.3 Performance Analysis

In this section, we analyze the performance of the proposed fault localization protocol in terms of the maximum delay and average delay for localizing a single link failure in the three different scenarios.

5.2.3.1 Case A: Inter-domain link failure scenario

i) Maximum delay

To calculate the maximum delay, we assume that a failure occurs at the upstream end of an inter-domain link. In this case, the downstream node takes $d_{inter\_links} \times \gamma$ to detect the failure. At the same time, it needs to send an INTER_ALARM message to the upstream edge node to indicate that it has detected a failure. Therefore, the maximum delay taken to locate the failure is

$$T_{Max\_Inter\_domain} = 2 \times d_{inter\_links} \times \gamma$$

where, d_{inter_links} is the maximum length among all inter-domain links and γ is the expected value of propagation delay along the inter-domain link.

ii) Average delay

To calculate the average delay, we assume that a failure occurs at the middle of an inter-domain link. This takes $\frac{1}{2} \times E[d_{inter\_links}] \times \gamma$ time for the downstream edge node to detect the failure. At the same time, the downstream node needs to send an INTER_ALARM message to the upstream edge node to indicate that it has detected a failure. Therefore, the average delay taken to locate the failure can be calculated as

$$T_{Ave\_Inter\_domain} = \left(\frac{1}{2} \times E[d_{inter\_links}] \times \gamma\right) + \left(1 \times E[d_{inter\_links}] \times \gamma\right) = \frac{3}{2} \times E[d_{inter\_links}] \times \gamma$$

where $E[d_{inter\_links}]$ is the average length of all inter-domain links.

5.2.3.2 Case B: Inter-domain failure caused by intra-domain link failure

i) Maximum delay

To calculate the maximum delay, we divide the fault localization protocol into seven different phases and then calculate the delay in each phase, respectively. The overall delay is thus the sum of the delays in all phases. These phases include (1) detection phase, (2) broadcasting phase, (3) distinguishing phase, (4) multicasting phase, (5) matching phase, (6) responding phase, and (7) concluding phase.

In the detection phase, the maximum delay for an egress edge node to detect a failure is the maximum length of an optical domain. To calculate the maximum delay, we assume that a failure occurs right next to the farthest upstream edge node. Therefore, the maximum delay in this phase is $T_1 = \gamma \times D_{OD}$, where, D_{OD} is the longest diameter among all the domains. It is worth mentioning that in this phase an intra-domain node will also detect the failure and will set a timer to time out after $4 \times \gamma \times D_{OD}$. The reader is referred to Case B in Section 5.2.1 for details.

In the broadcasting phase, the downstream node will send an INTER_ALARM message to the upstream edge node, the maximum delay for the upstream edge node to receive this message is the same as in phase 1, i.e., $T_2 = \gamma \times D_{OD}$. Note that each intra-domain node that detects the failure will also broadcast an INTRA_ALARM message to all nodes within its domain.

In the distinguishing phase, the upstream node has known that there is a failure within its domain. Thus, it sends a GO_AHEAD message to all the nodes within the domain to start the localization process. The maximum time taken for the GO_AHEAD message to be

received by all nodes can be calculated as $T_3 = \gamma \times D_{OD}$. To ensure that all the nodes receive the GO_AHEAD message, the last affected node that receives the GO_AHEAD message could be next to the downstream node because this node does not know its location with respect to the downstream edge node. It assumes that it is the first to detect this GO_AHEAD message and has to wait for an additional $\gamma \times D_{OD}$ time to ensure that all the nodes within its domain have received the GO_AHEAD message.

If we add up the delays from phase 1 up to phase 3, the total delay is equal to the timer value $4 \times \gamma \times D_{OD}$ that has been set at each node once it detects a failure. For analysis purpose, however, this time does not reflect the total actual waiting time because if the node next to the upstream edge node detects a failure, it will start its timer to count down. On the other hand, another affected node may be located next to the downstream node and cannot start its timer unless it detects the failure after $\gamma \times D_{OD}$. Therefore, we should consider the time difference between the two nodes. And hence the delay up to this phase should equal to the sum of the time delay of the previous three phases plus T_{Max_Extra}, where $T_{Max\_Extra} = 2 \times \gamma \times D_{OD}$.

During the timer period, an affected node that generated an ALARM message has also received an ALARM message from each of the other affected potential executive sinks. Each node compares the length of the affected lightpath contained in each received ALARM message with the length of its affected lightpath. In other words, at the end of the timer period, each potential executive sink has reached a conclusion on which potential executive sink should act as the executive sink. Therefore, it is not necessary to consider the computational complexity of this period of time because this can be done in parallel while each potential executive sink is receiving the ALRAM messages.

In the multicasting phase, the executive sink will define a perimeter area and create its ALV vector, and then broadcast this vector to all nodes in the perimeter area. Since the executive route can be a route between a virtual source and a virtual sink of that domain, the time taken for the ALV to reach all nodes is $T_4 = \gamma \times D_{OD}$.

In the matching phase, each node within the perimeter area will compare the received ALV with the lightpath terminated at it. In the worst case, a perimeter node needs to compare two vectors that have lightpaths traversing the entire domain. Therefore, the computational

complexity of comparing two vectors in this phase is $O(D_{OD1}, D_{OD1})$, which implies that the maximum complexity is $O(D_{OD}^2)$. If we use θ to denote the computational cycle, the time delay in this phase is $T_5 = \theta \times D_{OD}^2$.

In the responding phase, each perimeter node will send its binary vector to the executive sink, the maximum delay in this phase is $T_6 = \gamma \times D_{OD}$.

In the concluding phase, the executive sink requires one computational cycle θ to determine the location of the failed link, and hence the time complexity of this phase is $T_7 = \theta$.

Therefore, the overall delay for localizing an intra-domain link failure, $T_{Max\_Inter\_due\_Intra}$, can be calculated as

$$\begin{aligned} T_{Max\_Inter\_due\_Intra} &= T_{Max\_Extra} + \sum_1^7 T_i \\ &= (7 \times \gamma \times D_{OD}) + \theta \times D_{OD}^2 + \theta \end{aligned}$$

For a grid network with $M \times N$ nodes, $D_{OD} = M + N - 2$. Hence, we have

$$T_{Max\_Inter\_due\_Intra} = 7\gamma(N + M - 2) + \theta((N + M - 2)^2 + 1).$$

For simplicity, if we assume that $M = N$, the overall delay can be calculated as

$$T_{Max\_Inter\_due\_Intra} = 14\gamma(N - 1) + \theta((2(N - 1))^2 + 1)$$

ii) Average delay

To calculate the average delay, we can use the method similar to the above analysis for the maximum delay. Hence, the overall average delay is

$$\begin{aligned} T_{Ave\_Inter\_due\_Intra} &= (T_{Ave\_Extra}) + \sum_{i=1}^{7} T_i \\ &= (\gamma D_{Ave\_OD} + \gamma E[d_{sd}]) + \gamma \frac{1}{2} E[d_{sd}] + 2\gamma D_{Ave\_OD} + 2\gamma(E[d_{sd}] + 1) + \theta E[d_{sd}]^2 + \theta \\ &= 3\gamma D_{Ave\_OD} + \gamma \frac{7}{2} E[d_{sd}] + 2\gamma + \theta(E[d_{sd}]^2 + 1) \end{aligned}$$

where $E[d_{sd}]$ is the average hop distance and D_{Ave_OD} is the average diameter among all domains.

For a grid network, $E[d_{sd}]=(N+M)/3$ and $D_{Ave\_OD}=N+M-2$ [NIST]. Hence, we have

$$T_{Ave\_Inter\_due\_Intra} = 3\gamma(N+M-2)+\frac{7}{2}\gamma\frac{N+M}{3}+2\gamma+\theta\left(\left(\frac{N+M}{3}\right)^2+1\right)$$

$$T_{Ave\_Inter\_due\_Intra} = \gamma\left(\frac{25N+25M-24}{6}\right)+\theta\left(\left(\frac{N+M}{3}\right)^2+1\right)$$

For simplicity, if we assume that $M = N$, the overall delay can be calculated as

$$T_{Ave\_Inter\_due\_Intra} = \gamma\left(\frac{25N-12}{3}\right)+\theta\left(\frac{4N^2+9}{9}\right)$$

5.2.3.3 Case C: Intra-domain fault localization scenario

i) Maximum delay

To calculate the maximum delay, we follow the same analysis for Case B. The total maximum delay can be calculated as

$$\begin{aligned} T_{Max\_Intra} &= \left(T_{Max\_Extra}\right)+\sum_{i=1}^{7}T_i \\ &= \left(2\times\gamma\times D_{OD}\right)+4\times\gamma\times D_{OD}+\theta\times D_{OD}^2+\theta \end{aligned}$$

or in terms of N and M,

$$T_{Max\_Intra} = \left(6\times\gamma\times(N+M-2)\right)+\theta\times(N+M-2)^2+\theta$$

If we assume $M = N$, we have

$$T_{Max\_Intra} = \gamma 12(N-1)+\theta(4N^2-8N+5)$$

ii) Average delay

Similarly, following the same analysis for case B, the total average delay can be calculated as

$$T_{Ave\_Intra} = (T_{Ave\_Extra}) + \sum_{i=1}^{7} T_i$$

$$= (\gamma D_{Ave\_OD} + \gamma E[d_{sd}]) + \gamma \frac{1}{2} E[d_{sd}] + \gamma D_{Ave\_OD} + 2\gamma(E[d_{sd}]+1) + \theta \times E[d_{sd}]^2 + \theta$$

$$= \gamma \frac{7}{2} E[d_{sd}] + 2\gamma D_{Ave\_OD} + 2\gamma + \theta(E[d_{sd}]^2 + 1)$$

or in terms of N and M,

$$T_{Ave\_Intra} = \gamma \frac{7}{2} \frac{N+M}{3} + 2\gamma(N+M-2) + 2\gamma + \theta\left(\left(\frac{N+M}{3}\right)^2 + 1\right)$$

If we assume $M = N$, we have

$$T_{Ave\_Intra} = \gamma \frac{19N}{3} - 2\gamma + \theta\left(\frac{4N^2+9}{9}\right)$$

5.2.4 Analytical Results

To evaluate the performance of the proposed fault localization protocol, we give the numerical results of the localization delay with the three different cases. For case B and case C, we also compare the localization delay with that of OSPF. For comparison, we consider a multi-domain all-optical networks and each domain is a grid network with $M \times N$ nodes, as shown in **Figure 5. 2**. We assume that the light speed inside a fiber is 2.14 x 10^8m/s, the length of each link within a domain is 10km, the length of an inter-domain link is 100km, and the computational cycle θ is 20ns.

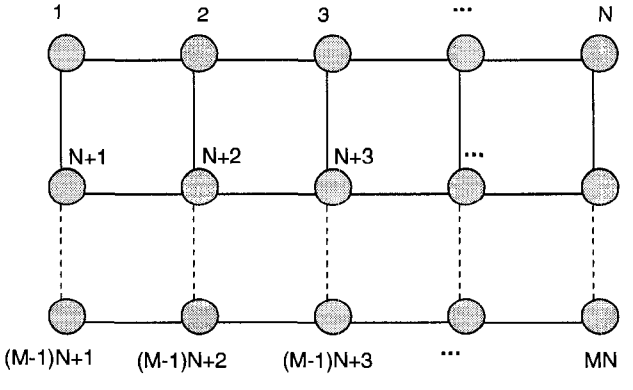

Figure 5. 2: Grid Network

Under these assumptions, the fault localization delay in case A is very small with the maximum and average delay values being 9.34E-4 and 7.00E-4, respectively.

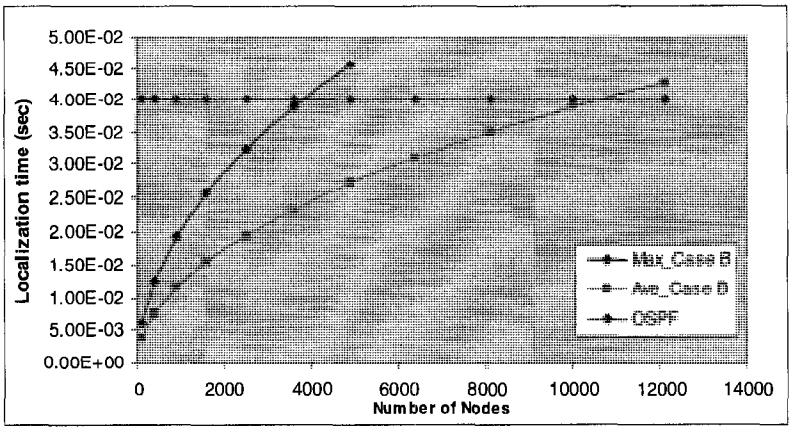

Figure 5. 3: Localization delay: case B vs OSPF

Figure 5. 3 and Figure 5. 4 show the fault localization delay with case B, case C, and OSPF, respectively. It is seen that given the same number of network nodes the localization delay with case C is smaller than that with case B in terms of both the maximum delay and the average delay. In the worst case, the maximum localization delay with case B is smaller than that with OSPF (less than 40s) if the number of nodes is not larger than 3600.

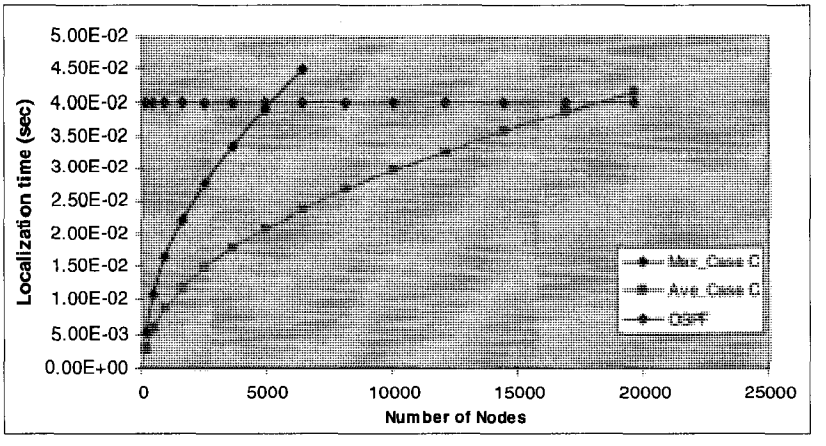

Figure 5. 4: Localization delay: case C vs OSPF

5.3 Distributed Fault Localization for Multi-Domain All-Optical Networks without Power Monitoring

5.3.1 Inter-domain Fault Localization Protocol Description

In this section, we present a distributed fault localization protocol for localizing single-link failures in multi-domain all-optical networks. We consider a network with no optical power monitoring available at either intra-domain nodes or edge nodes. Each fiber has multiple wavelengths and control messages are delivered out-of-band.

The fault localization protocol is consists of two phases: intra-domain localization phase and inter-domain localization phase. The intra-domain phase is intended to identify whether a link failure is within a particular domain and at the same time localize the failure if it is within that domain, while the inter-domain phase is intended to localize the failed inter-domain link if the intra-domain phase identifies that a failure is on an inter-domain link. In the description of the protocol, we consider three failure cases: A, B, and C, which will be explained in the description. For ease of exposition, we consider an example shown in Figure 5. 5.

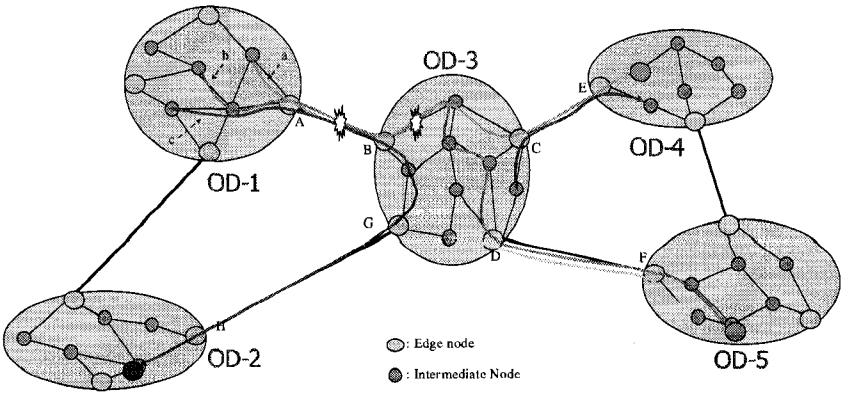

Figure 5. 5: Inter-domain failure localization

5.3.1.1 Case A: Inter-domain failure cased by Intra-domain failure

In this case, a link failure occurs within a domain, e.g., in OD-3, and affects inter-domain traffic, e.g., the lightpaths a and b which terminate in OD-4 and OD-5, respectively, as shown in Figure 5. 5.

When the failure occurs in OD-3, the sink nodes of lightpaths a and b terminating in OD-4 and OD-5 will detect the failure. Once detecting the failure, they will run a modified LVM process, respectively. The ingress edge nodes E and F in OD-4 and OD-5 will act as the virtual sources of the affected lightpaths a and b. The procedures to localize the failure within OD-4 and OD-5 are described as follows.

1) The sink node of an affected lightpath, also called potential executive sink, will broadcast an INTRA_DOMAIN_ALARM to its entire domain once it detects a failure.
2) The sink node will wait for a period of time equal to twice the diameter of its domain. The reason is that each node does not know its location in its domain with respect to the edge nodes and it will assume that it is the first to detect the failure. Hence, it has to allow enough time for other possible sink nodes located on the other side of the optical domain to detect the failure, and also enough time to receive the alarms originated from those possible sinks.
3) The ingress edge node of the affected lightpath will also set a timer equal to twice the diameter of its domain once it receives the first alarm and will act as a virtual source of the affected lightpath.
4) After receiving all alarms, the potential executive sink with the shortest lightpath will act as the executive sink. The ingress edge node which the affected lightpath passes through will act as a virtual source. This is because no internal confidential information about the topology of this domain is allowed to be sent out to other domains.
5) The executive sink will then create a vector of links on its shortest affected lightpath, called Affected-Link-Vector (ALV), define a limited-perimeter, which consists of all the ALV links and their neighboring nodes, and then multicast the ALV to all the nodes within the limited-perimeter.
6) Each sink receiving the ALV within the limited-perimeter will create its own link vector and compares the vector with the ALV. If its link vector has a common link with the ALV, the sink will send a binary vector back to the executive sink. The binary vector consists of a matching vector and the status of the corresponding lightpath. The matching

vector contains the same links in the ALV. If the lightpath is working, its status is set to "1". Otherwise, the status is set to "0". For a working path, an element in the matching vector is set to "0" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "1". For a failed path, an element in the matching vector is set to "1" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "0".

7) After the executive sink collects the binary vectors from all the sinks within the limited-perimeter, it will perform a logical AND operation on all the collected binary vectors to determine the location of the failure. The link with "1" in the resulted binary vector will be determined as the failed link. If there are more than "1" in the resulted binary vector, the executive sink can extend the limited-perimeter by including the neighbors of the original perimeter nodes and run a second round of the above process starting from step 5. However, if the AND operation results in no "1" in the resulted binary vector, it means that no failed link exists within this domain.
8) The executive sink will broadcast a FALSE_ALARM message to the entire domain.
9) Once the ingress node receives the FALSE_ALARM message, it will send an INTER_DOMAIN_ALARM to its upstream edge node which the failed lightpath passes through, informing it of the failed lightpath ID and the localization result. For the example in Figure 5. 5, the failure is not within either OD-4 or OD-5. This result will be broadcast to the entire domain by the corresponding executive sink. Once the ingress edge node E of OD-4 and ingress edge node F of OD-5 receive the result, they will forward the result to their upstream egress edge nodes, i.e., edge node C and edge node D of OD-3, indicating no failure localized within their domains.
10) The egress edge node of the upstream optical domain initiates another localization process within its domain. The egress edge node will act as a virtual sink for the affected lightpath and the corresponding ingress edge node will act as a virtual source. Note that there may be other egress edge nodes that have received INTER_DOMAIN_ALARM and will initiate another localization processes. Each of them starts an independent localization process starting from the multicasting phase (i.e., step 5) because the failed lightpath is known. For the example in Figure 5. 5, both egress edge nodes C and D of OD-3 will initiate another localization process, respectively, which starts from the multicasting phase. Note that the two processes will start independently at different

times.

11) The first egress edge node that localizes the failure will send an INTER_DOMAIN_LOCALIZATION message to other domains, indicating the failed link. For the example in Figure 5. 5, executive sink C will be the first to advertise the failed link because it has a shorter affected lightpath compared to egress edge node D and thus detects the lightpath failure first. Moreover, OD-4 is closer to the failed link compared to OD-5. Therefore, edge nodes B, D, and G will also advertise the result about the faulty domain to other domains.

If the localization process results in no failed link, the virtual executive sink, also called egress edge node, will send a FALSE_ALARM message to the entire domain to indicate that no failed link is localized. In response to the alarm message, the ingress edge node which the affected lightpath passes through will start over again from step 9 to track the source of the failure. This process continues until the source domain of the affected lightpath is reached and the failed link is localized.

5.3.1.2 Case B: Inter-domain failure caused by Inter-domain failure

In this case, a failure occurs on an inter-domain link, e.g., on link A-B, as shown in Figure 5. 5. Once the failure occurs, the sink nodes in OD-4, OD-5, and OD-2 will detect the failure and then run the modified LVM process. The running results will be sent to their upstream edge nodes C, D, and G. These edge nodes will then run multiple LVM processes in OD-3, independently. Each independent session starts from the multicasting phase because the affected lightpaths are known. The first edge node which finishes the LVM process will send the result to the entire domain. For the example in Figure 5. 5, there is no failed link within OD-3. Egress edge node C will be the first to advertise the result to the entire domain. Hence the ingress edge node B of OD-3 will forward the result to the egress edge node A of OD-1. Edge node A will then act as a virtual sink and run another LVM process in OD-1 and find out that there is no failed link within its domain. Since edge node A knows that OD-1 is the source domain of the affected lightpaths, it will act as an inter-domain executive sink and initiate the inter-domain localization phase. Edge node A considers the shortest affected lightpath that terminates in OD-4 as the Inter-domain Executive Route (I-ER).

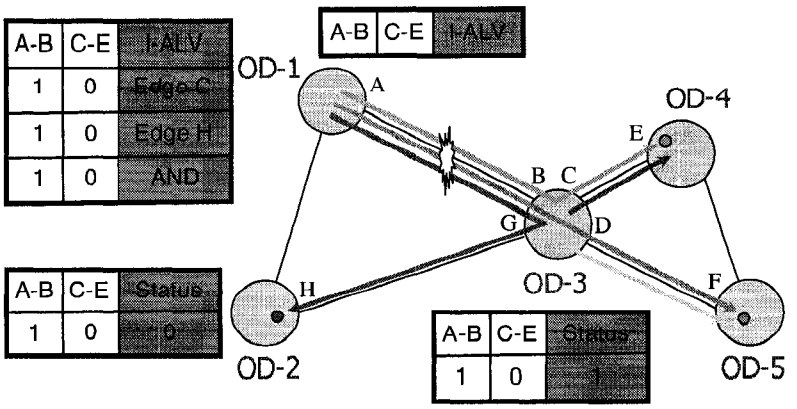

Figure 5. 6: different optical domains

Once the first phase ends and no failed link is localized, the second phase is started. The procedures are described as follows.

1) The egress edge node of the source domain will act as the inter-domain executive sink. In the second phase, each domain will be considered as one single edge node.
2) The IES will then create a vector of links on its inter-domain affected lightpath, called Inter-domain Affected-Link-Vector (I-ALV), define an inter-domain limited-perimeter, which consists of all the I-ALV links and their neighboring ODs, and then multicast the I-ALV to all ODs within the inter-domain limited-perimeter. The inter-domain executive sink makes use of the BGP routing table to construct the I-ALV and the inter-domain limited-perimeter. For the example in Figure 5. 5, node A is the inter-domain executive sink. Figure 5. 6 illustrates the procedures described above.
3) Each OD that acts as a sink receiving the I-ALV within the limited-perimeter will create its own link vector and compare the vector with the I-ALV. If its link vector has a common link with the ALV, the sink will send a binary vector back to the inter-domain executive sink. The binary vector consists of a matching vector and the status of the corresponding lightpath. The matching vector contains the same links in the ALV. If the lightpath is working, its status is set to "1". Otherwise, the status is set to "0". For a working path, an element in the matching vector is set to "0" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "1". For a failed path, an element in the matching vector is set to "1" if the corresponding link is found in the sink's link vector. Otherwise, it is set to "0". For the example in Figure 5. 5, we assume the first to

reply is edge node C in OD-3, indicating that it has a working lightpath that passes through link C-E and terminates in OD-4. This information will be sent back in the binary vector to edge node A. Similarly, this will be done by edge node H in OD-2, which will send back a binary vector indicating that it has a failed lightpath that passes through link A-B.

After the inter-domain executive sink collects the binary vectors from all the OD sinks within the inter-domain limited-perimeter, it will perform a logical AND operation on all the collected binary vectors to determine the location of the failure. The link with "1" in the resulted binary vector will be determined as the failed link. If there are more than "1" in the resulted binary vector, the executive sink can extend the limited-perimeter by including the neighbors of the original perimeter OD and run a second round of the above process. Therefore, once the binary vectors are received by node A, it will perform logic AND on them and determine that link A-B is down, as shown in Figure 5. 6. Then it will advertise this result to all other ODs

5.3.1.3 Case C: Intra-domain failure cased by Intra-domain failure

In this case, a link failure occurs within a domain and only affects intra-domain traffic. Once the failure occurs, the sink nodes within its domain will detect the failure. In response, they will run the LVM protocol to localize the failure. The localization process is similar to that described in case A from step 1 to step 7, which will not be repeated here. Once the failed link is localized, the executive sink will broadcast an INTRA_FAULT_LOCATION message to all nodes within its domain, notifying them of the fault location.

5.3.2 Performance Analysis

In this section, we analyze the performance of the proposed fault localization protocol in terms of the maximum delay for localizing a single-link failure in the three different cases.

5.3.2.1 Case A: Inter-domain failure cased by Intra-domain failure

To analyze the maximum delay, we divide the fault localization process into six different phases and then calculate the delay in each phase, respectively. The overall delay is thus the

sum of the delays in all phases. These phases include (1) detection phase, (2) broadcasting phase, (3) multicasting phase, (4) matching phase, (5) responding phase, and (6) concluding phase.

To calculate the maximum delay in the detection phase, we assume that a failure occurs right next to the farthest downstream edge node. Therefore, the maximum delay can be calculated as

$$T_1 = H(\gamma \times D_{OD}) + (H-1)(\gamma_{inter\_link}),$$

where D_{OD} is the diameter length of an optical domain, γ_{inter_link} is the propagation delay of the longest inter-domain link, γ is the propagation delay of the longest intra-domain link, and H is the number of optical domains that an affected lightpath passes through. $H-1$ is the number of inter-domain links that an affected lightpath passes through.

Note that in this phase an intra-domain node will also detect the failure and will set a timer to time out after $2 \times \gamma \times D_{OD}$. We will explain the reason for setting this timer later in this section.

In the broadcasting phase, a sink node will broadcast an INTRA_ALARM message to its entire domain, the maximum delay for all nodes within its domain to receive this message is $T_2 = \gamma \times D_{OD}$. To ensure that all the nodes receive the INTRA_ALARM message, each sink node that detects the failure will wait for a period of time equal to twice the diameter of its domain $2 \times \gamma \times D_{OD}$.

Up to this point, if we add all the delay from phase 1 to phase 2, it will be equal to the timer value each node has set, excluding the delay from other domains in phase 1.However, this time does not reflect the total actual waiting time because if the node next to the upstream edge node detects the failure, it will start its timer to count down. On the other hand, another affected sink may be located next to the downstream node and cannot start its timer unless it detects the failure after $\gamma \times D_{OD}$. Therefore, we should consider the time delay difference between the two nodes and the delay up to this phase should equal to the sum of the time delays of the previous two phases plus T_{Max_Extra}, where $T_{Max\_Extra} = \gamma \times D_{OD}$.

During the timer period, an affected sink that generated an ALARM message has also received an ALARM message from each of the other affected potential executive sinks. Each node compares the length of the affected lightpath contained in each received ALARM

message with the length of its affected lightpath. In other words, at the end of the timer period, each potential executive sink can reach a conclusion on which potential executive sink should act as the executive sink. Therefore, it is not necessary to consider the computational delay of this period of time because this can be done in parallel while each potential executive sink is receiving the ALRAM messages.

In the multicasting phase, the executive sink will define a perimeter area and create its ALV vector, and then broadcast this vector to all nodes in the perimeter area. Since the executive route can be a route between a virtual source and a virtual sink of that domain, the time taken for the ALV to reach all nodes is $T_3 = \gamma \times D_{OD}$.

In the matching phase, each node within the perimeter area will compare the received ALV with the lightpath terminating at it. In the worst case, a perimeter node needs to compare two vectors that have lightpaths traversing the entire domain. Therefore, the computational complexity of comparing two vectors in this phase is $O(D_{OD1}, D_{OD1})$, which implies that the maximum complexity is $O(D_{OD}^2)$. If we use θ to denote the computational cycle, the time delay in this phase is $T_4 = \theta \times D_{OD}^2$.

In the responding phase, each perimeter node will send its binary vector to the executive sink, the maximum delay in this phase is $T_5 = \gamma \times D_{OD}$.

In the concluding phase, the executive sink requires one computational cycle θ to determine the location of the failed link, and hence the processing delay of this phase is $T_6 = \theta$.

Therefore, the maximum delay for localizing an intra-domain link failure can be expressed as

$$\begin{aligned} T_{Inter\_Intra}^{\max} &= T_{\max\_extra} + \sum_{i=1}^{6} T_i \\ &= \gamma \times D_{OD} + H(\gamma \times D_{OD}) + (H-1)(\gamma_{inter\_link}) \\ &\quad + \gamma \times D_{OD} + \gamma \times D_{OD} + \theta \times D_{OD}^2 + \gamma \times D_{OD} + \theta \\ &= \gamma \times D_{OD}(H+4) + (H-1)(\gamma_{inter\_link}) + \theta \times D_{OD}^2 + \theta \end{aligned}$$

For a grid network with $M \times N$ nodes, $D_{OD} = M + N - 2$. Hence, we have

$$T_{Inter\_Intra}^{\max} = \gamma(N+M-2)(H+4) + (H-1)(\gamma_{inter\_link}) + \theta((N+M-2)^2+1)$$

For simplicity, if we assume that $M = N$, the overall delay can be calculated as

$$T_{Inter-Intra}^{\max} = 2\gamma(N-1)(H+4) + (H-1)(\gamma_{inter\_link}) + \theta((2(N-1))^2+1)$$

However, this is only the localization delay for the first optical domain, the failed link could be in another domain and hence the maximum delay for other domains has to be also included. The total delay has to include the localization time of all domains that perform a localization process on an affected lightpath. If we assume that an affected lightpath passes through H domains and $H-1$ inter-domain link. The total delay can be expressed as

$$\begin{aligned} T_{Inter\_Intra}^{total\_\max} &= T_{Inter\_Intra}^{\max} + (H-1)\sum_{i=3}^{6} T_i + (H-1)(\gamma_{inter\_link}) \\ &= T_{Inter\_Intra}^{Max} + (H-1)(2\gamma D_{OD} + \theta \times D_{OD}^2 + \theta) + (H-1)(\gamma_{inter\_link}) \\ &= 2\gamma(N-1)(3H+2) + H\theta((2(N-1))^2+1) + 2(H-1)(\gamma_{inter\_link}) \end{aligned}$$

As we can see, the delay in other domains includes only the delay in the multicasting phase because the affected lightpath is known and there is no need to detect the failure and broadcast the alarms. Furthermore, the delay due to inter-domain links should be also included.

5.3.2.2 Case B: Inter-domain failure cased by Inter-domain link failure

In this case, the second localization phase starts after the first localization phase is done. Therefore, the total delay should include the delay in the first and second phases. Since the affected lightpath is known, the selected inter-domain executive sink will start from the multicasting phase. Therefore, the total maximum delay can be calculated as

$$T_{Inter\_Inter}^{Total\_Max} = T_{Inter\_Intra}^{Total\_Max} + \sum_{i=3}^{6} IntT_i,$$

where $\sum_{i=3}^{6} IntT_i$ is the inter domain delay in the multicasting phase, matching phase, responding phase, and concluding phase, respectively. The total delay in the second phase is

$$T_{Inter\_Inter}^{total\_\max} = T_{Inter\_Intra}^{total\_\max} + 2\left(\gamma((H-2)\times D_{OD}) + \left((H-1)\gamma_{inter\_link} + \gamma_{inter\_link}\right)\right) + \theta\left((H)^2\right) + \theta$$
$$= T_{Inter\_Intra}^{total\_\max} + 2\left(\gamma((H-2)\times D_{OD}) + \left(H\gamma_{inter\_link}\right)\right) + \theta\left((H)^2\right) + \theta$$

where $2\left(\gamma((H-2)\times D_{OD}) + \left(H\gamma_{inter_link}\right)\right)$is for the multicasting and replying phases within the perimeter area, $\theta\left((H)^2\right) + \theta$ is the delay in the matching and concluding phase. As we can see, the maximum number of links that could be compared is the maximum number of links that could be found in the inter-domain perimeter area.

If we substitute $D_{OD} = M + N - 2$ and for M=N, the total delay can be expressed as

$$T_{Inter\_Inter}^{total\_\max} = T_{Inter\_Intra}^{total\_\max} + 2\gamma((H-2)\times(2N-2)) + 2\left(H\times\gamma_{inter\_link}\right) + \theta\left((H)^2\right) + \theta$$

5.3.2.3 Case C: Intra-domain failure cased by Intra-domain link failure

To calculate the maximum delay in this case, we follow the same analysis for Case B. The total maximum delay can be calculated as

$$T_{Intra\_Intra}^{\max} = \left(T_{\max\_extra}\right) + \sum_{i=1}^{7} T_i$$
$$= (\gamma\times D_{OD}) + 4\times\gamma\times D_{OD} + \theta\times D_{OD}^2 + \theta$$

or in terms of N and M,

$$T_{Intra\_Intra}^{\max} = (5\times\gamma\times(N+M-2)) + \theta\times(N+M-2)^2 + \theta.$$

If we assume $M = N$, we have

$$T_{Intra\_Intra}^{\max} = \gamma 10(N-1) + \theta\left(4N^2 - 8N + 5\right).$$

5.3.3 Analytical Results

To evaluate the performance of the proposed fault localization protocol, we show the localization delays with the three different cases. We also compare their localization delays with that of OSPF, i.e., 40 s. For comparison, we consider a multi-domain all-optical

network, in which each domain is a grid network with $M \times N$ nodes, as shown in **Figure 5. 2.** We assume that the light speed inside a fiber is 2.14 x 10^8m/s, the length of each link within a domain is 10km, the length of an inter-domain link is 100km, and the computational cycle θ is 20ns.

Figure 5. 7 shows the localization delays with the three cases, where the number of domains is 4 while the number of nodes in each domain is variable. It is seen that the localization delay with case B is below 40sec as long as the number of nodes in each domain is below 350 nodes. This means that in this case the proposed protocol outperforms OSPF if the number of nodes in each domain does not exceed than 350.

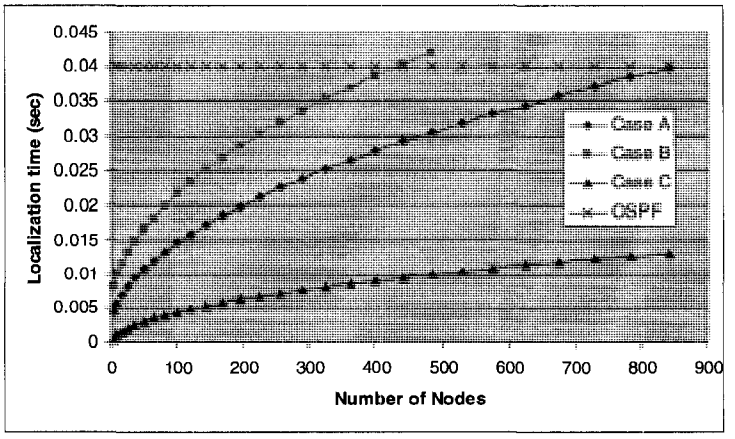

Figure 5. 7: Localization delay vs number of nodes for 4 domains

Figure 5. 8 shows the localization delays with the three cases, where the number of domains is increased to 6. We can see that compared to Figure 4 the maximum delay increases. This is due to the fact that an affected lightpath could traverse 6 optical domains and hence more time is needed to localize the failure. For case B, the delay reaches 40sec if the domain size reaches approximately 150 nodes. This means that in this case the proposed protocol outperforms OSPF if the number of nodes in each domain is less than 150.

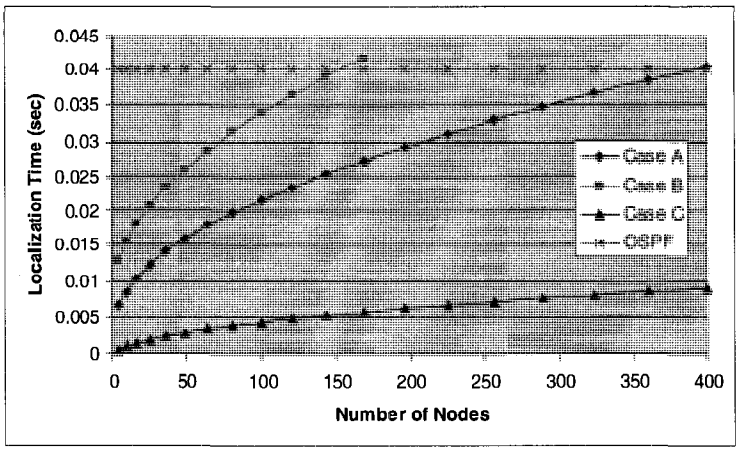

Figure 5. 8: Localization time vs number of nodes for 6 domains

Figure 5. 9 shows the localization delays, where the number of nodes in each domain is 16 while the number of domains is variable. It is seen that for case B the delay reaches 40 sec if the number of domains becomes 13. Note that the delay with case C is unaffected because it is independent of the number of domains.

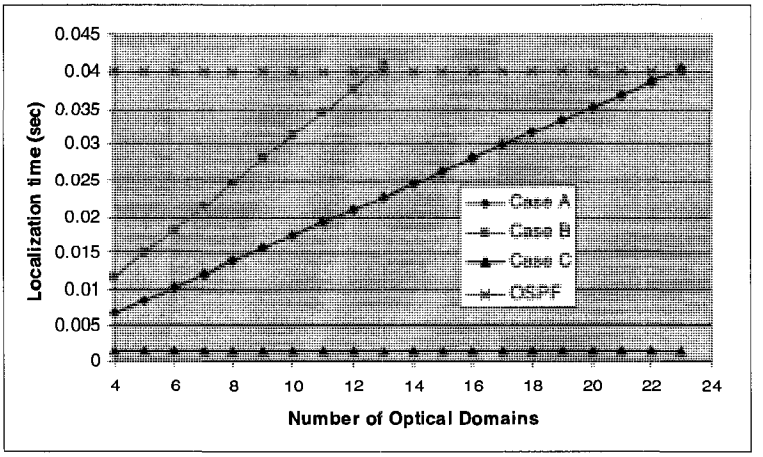

Figure 5. 9: Localization time vs number of optical domains for 16 nodes

Figure 5. 10 shows the localization delays, where the number of nodes in each domain is increased to 25. For case B, the delay reaches OSPF localization time if the number of domains reaches approximately 11.

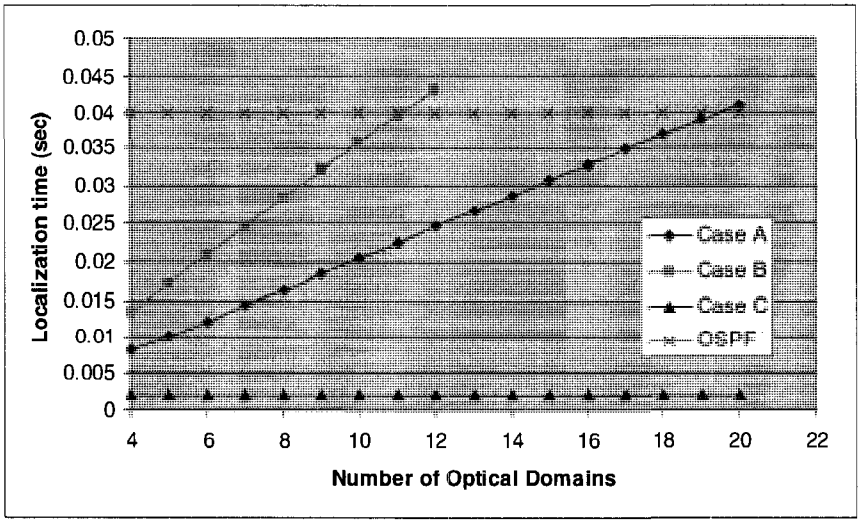

Figure 5. 10: Localization time vs number of optical domains for 25 nodes

5.4 Summary

In this chapter, we proposed a fault localization protocol for localizing single-link failures in multi-domain all-optical networks. The proposed protocol is based on the limited-perimeter vector matching (LVM) mechanism and assumes power monitoring available only at edge nodes, It can not only localize inter-domain link failures between different domains, but also localize intra-domain link failures within a single domain. In particular, it does not require exchanging any internal confidential domain-specific information between different domains. The analytical results have shown that the proposed fault localization protocol can not only fast localize an inter-domain link failure, but also localize an intra-domain link failure that affects inter-domain traffic faster than the Open-Shortest-Path-First (OSPF) protocol for a large network.

Furthermore, we propose a distributed fault localization protocol for localizing single-link failures in all-optical multi-domain networks. The protocol is also based on a limited-perimeter vector matching (LVM) mechanism and assumes no power monitoring at either edge nodes or in intermediate nodes. By performing a fault localization process sequentially in the optical domains that an affected lightpath passes through, the protocol can localize both inter-domain and intra-domain link failures that affect inter-domain traffic

without exchanging any internal confidential domain-specific information between different domains. We show through analytical results that it can not only fast localize an inter-domain link failure between different domains but also localize an intra-domain link failure that affects inter-domain traffic faster than the Open-Shortest-Path-First (OSPF) protocol in a large network.

Chapter 6 Conclusion and Future Work

6.1 Summary and Concluding Remarks

This thesis studies the major issues on fault localization in both intra-domain and inter-domain all optical networks.

Chapter 3 proposed a distributed fault localization protocol, called Parallel Limited Perimeter Vector Matching (P-LVM) protocol, for localizing multi-link failures in all-optical networks. P-LVM is based on a Limited Perimeter Vector Matching (LVM) protocol for localizing single-link failures in all-optical networks. To handle multi-link failures, P-LVM tries to separate each failure in a small perimeter area after identifying each perimeter area with its corresponding failure and then localize the failures in parallel respectively in a distributed manner. The simulation results show that P-LVM can effectively localize multi-link failures that occur simultaneously or at different times in all-optical networks.

Chapter 4 provides a comprehensive study for the optimization problems in applying the LVM protocol for localizing single-link failures in static all-optical networks. Two optimization problems were considered: the problem for maximizing the fault localization probability and the problem for minimizing the fault localization time. Given the traffic demands, the former problem is to optimize the traffic distribution so that the fault localization probability in terms of the number of localized links is maximized while the latter is to optimize the traffic distribution so that the fault localization time for each failed link is minimized. We formulized the problems into an integer linear programming problem, respectively, and use CPLEX to solve the problems. Through numerical results, we showed that by optimizing the traffic distribution the fault localization probability can be maximized and the solution to the problem can also provide the maximum number of wavelengths needed on each link to obtain the maximum localization probability. Moreover, the fault

localization time can be effectively minimized and is below the OSPF localization time (40sec) in traditional IP networks. Finally, we proposed a heuristic algorithm that can be used to evaluate the optimization results. This heuristic algorithm tries to satisfy the LVM protocol's conditions through assigning appropriate costs to the network links to allow different *s-d* pairs to utilize the same link. It has two operation modes: greedy and non-greedy. Both the greedy and non-greedy heuristics provide link coverage but not as well as the optimization results. The heuristics also try to distribute the traffic to minimize the localization time at each link

Chapter 5 proposes two different fault localization protocols for localizing single-link failures in multi-domain all-optical networks. Both protocols are based on a limited-perimeter vector matching (LVM) mechanism, which restricts fault localization within a smaller perimeter area and can thus significantly reduce the time and space complexity for fault localization. The first inter-domain fault localization protocol assumes power monitoring available only at edge nodes, whereas the second protocol assumes no power monitors at the intermediate nodes nor the edge nodes. The two protocols can localize both inter-domain and intra-domain link failures without exchanging any internal confidential domain-specific information between different domains. The analytical results show that both protocols can not only fast localize an inter-domain link failure between different domains but also localize an intra-domain link failure that affects inter-domain traffic faster than the Open-Shortest-Path-First (OSPF) protocol in a large network.

6.2 Future Work

For future work, we suggest the following work to be done:

- Extending the proposed optimization and heuristics models to a dynamic traffic environment. The optimization model should be enhanced to have the ability to deal with dynamic traffic, or fast heuristic approaches should be introduced to provision the dynamic traffic fast to achieve fast localization or maximize the number of links that can be localized while still satisfying maximum number of requests.
- Considering partial and full wavelength-conversion issues [GOMCJ08, JWTF08] for fault localization-constrained connection provisioning. Introducing wavelength

conversation will allow provisioning the lightpaths in totally different way and hence affecting localization time or the number of failed links that can be localized when a failure occurs. Furthermore, issues like proper placement for wavelength conversion also should be addressed.

- Applying the multi-failure localization protocol for inter-domain all optical networks. Since inter-domain means that many ODs, the probability of having more than one failure increases. Therefore, applying a multi-failure protocol becomes a tempting solution. Multi-failure protocol could be used with either power monitors at the edge nodes or simply without power monitors anywhere.
- Providing clear interaction between intra-domain and inter-domain localization protocols, we have seen many papers in the literature discuss the interaction between intra-domain and inter-domain routing protocols [BYCK06, NJAL07]. The same thing should be applied among localization protocols.
- Providing clear interaction between intra-domain and intra domain routing protocols with intra-domain and inter-domain localization protocols:
    - For intra-domain, we have seen that localization protocols such as LVM or P-LVM can localize the failed link faster than OSPF. Therefore, localization protocols should be able to trigger the routing protocol such as OSPF to exchange the information about the failed link to the entire network and also to update the routing table to exclude this link from their current and future routes. Furthermore, during the localization process, localization protocols need to access the routing table to get the connectivity of the nodes when the perimeter area is being defined by the ES.
    - For inter-domain, edge nodes should be able to make use of intra-domain routing information to identify any lightpath that is passing through or terminates within its domain. Furthermore, edge nodes should be able to access inter-domain routing information to realize the connectivity among ODs when the edge node is constructing the inter-domain perimeter area or forming the inter-domain affected link vector (I-ALV).

- Optimizing intra-domain traffic distribution based on different criteria such as fiber length, available bandwidth to achieve fast localization and or increase number of localized links.
- Implementing an optimization model for traffic distribution to achieve fast localization or increase number of localized links in inter-domain all optical networks. This optimization model should consider inter-domain configuration policies, fiber length, available bandwidth and other inter-domain parameters.
- Implementing an intra-domain availability-awareness [KGC08, KICTON08, KaMO08] optimization model for fault localization in all-optical networks based on the LVM protocol. The same thing can be applied for inter-domain all-optical networks.
- Fault localization can also be extended to be applied in GMPLS and multi granular optical networks [NM06, CAQ07]. Traffic distribution should consider the different layers introduced in GMPLS for successful localization at each single layer.

References

[AB08] M. S. Ab-Rahman, "Optical Network Restoration and Migration Using OXADM", 10th International Conference on Advanced Communication Technology (ICACT 2008), Volume 3, Page(s): 2155 – 2160, February 2008.

[AR97] A. Amrani, J. Roldan and G. Junyent, "Optical monitoring system for scalable all-optical networks", Proceedings of IEEE Lasers and Electro-Optics Society (LEOS 1997), Volume 2, Page(s): 270 - 271, November 1997.

[AS08] R. Asthana, Y. N. Singh, "Distributed Protocol for Removal of Loop Backs and Optimum Allocation of p-Cycles to Minimize the Restored Path Lengths", Journal of Lightwave Technology, Volume 26, Issue 5, Page(s):616 -627, March 2008.

[AY08] A. Askarian, Yuxiang Zhai, S. Subramaniam, Y. Pointurier, M. Brandt-Pearce, "Protection and Restoration from Link Failures in DWDM Networks: A Cross-Layer Study", IEEE International Conference on Communications (ICC 2008), Page(s):5448 - 5452, May 2008.

[Bat01] R. J. Bates, "Optical switching and networking handbook", McGraw-Hill, 2001, ISBN: 007137356X.

[BC92] H. Braun, B. A. Chinoy, "The National Science Foundation Network," The San Diego Supercomputer Centre (SDSC) Applied Network Research group, GA-21029, September 1992.

[BCB et al 00] J. Banks, J. S. Carson II, B. L. Nelson, "Discrete-Event System Simulation", Prentice Hall PTR, Inc., 2000, ISBN: 0-13-088702-1.

[BH06] Y. Ben-Ezra, M. Haridim, B. I. Lembrikov, "All-Optical AGC of EDFA Based on SOA", Proceedings of IEEE Journal of Quantum Electronics, Volume 42, Issue 12, Page(s): 1209 – 1214, December 2006.

[Bha98] B. Bhandari, "Survivable Networks: Algorithms for Diverse Routing", Kluwer Academic Publishers, 1998, ISBN: 9780792383819.

[BLRW02] A. Basu, C. L. Ong, A. Rasala, G. Wilfong, "Route Oscillation in I-BGP with Route Reflection", Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications (ACM SIGCOMM 2002), Page(s): 235 -247, August 2002.

[BYCK06] D. Bauer, M. Yuksel, C. Carothers, S. Kalyanaraman, "A Case Study in Understanding OSPF and BGP Interactions Using Efficient Experiment Design", Proceedings of the 20th ACM/IEEE/SCS Workshop on Principles of Advanced and Distributed Simulation (PADS 2006), Page(s):158 – 165, May 2006.

[CAQ02] S. Chauban, V. Anand, C. Qiao, "Sub-Path Protection: A new framework for Optical Layer Survivability and Its Quantitative Evaluation", Buffalo NY, Technical report (Accessed February 2002).

[CAQ07] X. Cao, V. Anand, C. Qiao, "Waveband switching for dynamic traffic demands in multigranular optical networks", IEEE/ACM Transactions on Networking, Volume 15, Issue 4, Page(s): 957-968, August 2007.

[CPLEX] CPLEX , http://www.ilog.com.

[Do06] J. Doucette, "Joint working and spare capacity design of node-inclusive spanrestorable optical networks", Proceedings of Optical Fiber Communication Conference, 2006 and the 2006 National Fiber Optic Engineers Conference (OFC 2006), Page(s): 10 pp., March 2006.

[FSHL01] M. Francisco, S. Simpson, C. Huang, I. Lambadaris, B. St-Arnaud, "Inter-Domain Routing in Optical Networks", Proceedings of Optical Networking and Communications Conference (Opticomm 2001), Volume 4599, Page(s): 120-129, August 2001.

[GA03] A. Gumaste and T. Anthony, "DWDM network designs and engineering solutions", Indianapolis, IN: Cisco Press, 2003, ISBN: 1587050749.

[GG02] T. G. Griffin, G. Wilfong, "On the Correctness of IBGP configuration", Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications (ACM SIGCOMM 2002), Page(s): 17 - 29, August 2002.

[GOMCJ08] M. Galili, L.K. Oxenlowe, H.C.H. Mulvad, A.T. Clausen, P. Jeppesen, "Optical Wavelength Conversion by Cross-Phase Modulation of Data Signals up to 640 Gb/s", IEEE Journal of Selected Topics in Quantum Electronics, Volume 14, Issue 3, Page(s):573 – 579, May/June 2008

[Gor97] W. Goralski, "SONET : a guide to Synchronous Optical Network", McGraw-Hill, c1997, ISBN: 0070245630.

[Grov03] W. D. Grover, "Mesh-based Survivable Networks", Prentice Hall, 2003, ISBN: 013494576x.

[HE04] A. G. Hailemariam, G. Ellinas, T. Stern, "Localized failure restoration in mesh optical networks", Proceedings of IEEE Optical Fiber Communication Conference (OFC2004), Volume 1, Page(s): 23-27, February 2004.

[HZ06] M. Held, L. Zhou, "Redundancy, Restorability and Path Availability in Optical Mesh Networks", Proceedings of International Conference on Transparent Optical Networks (ICTON 2006), Volume 3, Page(s):120 - 125, June 2006.

[IG00] R. R. Iraschko, W. D. Grover, "A highly efficient path-restoration protocol for management of optical network transport integrity", Proceedings of IEEE Journal on Selected Areas in Communications, Volume 18, Issue 5, Page(s): 779 – 794, May 2000.

[IMC03] M. Ilyas, H. T. Mouftah, CRC Press, "The Handbook of Optical Communication Networks", CRC Press, Inc. Boca Raton, FL, USA, 2003, ISBN: 0849313333.

[JS06] D. R. Jeske, A. Sampath, "Restoration Strategies In Mesh Optical Networks: Cost Vs. Service Availability", 12th Pacific Rim International Symposium on Dependable Computing (PRDC '06), Page(s):147 - 153, December 2006.

[JWTF08] C. Jiajia, L. Wosinska, M. Tacca, A. Fumagalli, "Dynamic routing based on information summary-LSA in WDM networks with wavelength conversion", Proceedings of 10th Anniversary International Conference on Transparent Optical Networks (ICTON 2008), Volume 3, Page(s):50 – 54, June 2008.

[JY02] S. Jeong, C. Youn, "Instability Analysis for OBGP Routing Convergence in Optical Networks", Proceedings of the International Conference on Information Networking, Wireless Communications Technologies and Network Applications-Part I (ICOIN 2002), Page(s): 346-358, January 2002.

[K07] S.V. Kartalopoulos, "Channel Protection with Real-Time and In-Service Performance Monitoring for Next Generation Secure WDM Networks", IEEE International Conference on Communications (ICC 2007), Page(s):1235 – 1240, June 2007.

[KaMO08] B. Kantarci, H. T. Mouftah, S. Oktug, "Arranging Shareability Dynamically for the Availability-Constrained Design of Optical Transport Networks", IEEE Symposium on Computers and Communications (ISCC 2008),Volume 1, July 2008.

[KBB04] D. C. Kilper, R. Bach, D. J. Blumenthal, D. Einstein, T. Landolsi, L. Ostar, M. Preiss, and A. E.Willner, "Optical performance monitoring", Journal of Lightwave Technology, Volume 22, Issue 1, Page(s): 294–304, January 2004.

[KGC08] B. Kantarci, H. T. Mouftah, S. Oktug, "Differentiated Availability-Aware Connection Provisioning in Optical Transport Networks", In 2008 Proceedings of IEEE Global Communications Conference (GLOBECOM 2008), New Orleans, LA, USA, December 2008.

[Kha04] M. G. Khair, "An Implementation Approach for an Inter-Domain Routing Protocol for DWDM", Masters Thesis, University of Ottawa, March 2004.

[KICTON08] B. Kantarci, H. T. Mouftah, S. Oktug, "Connection Provisioning with Feasible Shareability Determination for Availability-Aware Design of Optical Networks", Proceedings of the 10th International Conference on Transparent Optical Networks (ICTON 2008), Volume 3, Page(s): 19-22, June 2008.

[KK04] M. Karol, P. Krishnan, J. J. Li, "Rapid fault detection and recovery for IP telephony", Proceedings of IEEE International Conference on Communication (ICC 2004), Volume 3, Page(s): 1484 – 1489, June 2004.

[KT96] Y. Kobayashi, Y. Tada, S. Matsuoka, N. Hirayama, K. Hagimoto, "Supervisory systems for all-optical network transmission systems", Proceedings of IEEE Global Telecommunications Conference (GLOBECOM 1996), Volume 2, Page(s): 933 – 937, November 1996.

[LR97] C. Li, R. Ramaswami, "Automatic fault detection, isolation, and recovery in transparent all-optical networks", Journal of Lightwave Technology, Volume 15, Issue 10, Page(s): 1784 - 1793, October 1997.

[LS99] E. Limal, K. E. Stubkjaer, "An algorithm for link restoration of wavelength routing optical networks", Proceedings of IEEE International Conference on Communications (ICC 1999). Volume 3, Page(s): 2055 – 2061, June 1999.

[LX05] C. Lee, X. Sun, C. Chan, L. Chen, "A single-fiber bidirectional self-healing WDM multihop mesh network with all-optical cyclic deflection routing for link restoration", Proceedings of IEEE Photonics Technology Letters, Volume 17, Issue 11, Page(s): 2463 – 2465, November 2005.

[M05] H. T. Mouftah, "Future trends in the design of survivable optical networks", Proceedings of the 7th International Conference on Transparent Optical Networks (ICTON 2005), Volume 1, Page(s): 52, July 2005.

[MC98] M. Medard, S. R. Chinn, P. Saengudomlert, "Attack detection in all-optical networks", Proceedings of Optical Fiber Communication Conference and Exhibit (OFC 1998), Page(s): 272 - 273, February 1998.

[MF02] O. Maennel, A. Feldmann, "Realistic BGP Traffic for Test Labs", Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications (ACM SIGCOMM 2002), Page(s): 235 - 247, August 2002.

[MG02] C. Siva, R. Murthy, M. Gurusamy "WDM Optical Networks Concepts, Design, and Algorithms", Prentice Hall, 2002, ISBN: 0130606375.

[MH03] H. T. Mouftah, P. Ho, "OPTICAL NETWORKS Architecture and survivability", Kluwer Academic Publishers, Inc., 2003, ISBN: 1-4020-7196-5.

[MH08] A. Mykkeltveit, B. E. Helvik, "On provision of availability guarantees using shared protection", International Conference on Optical Network Design and Modeling (ONDM 2008), Page(s):1 – 6, March 2008.

[Moy et al 98] J. T. Moy, "OSPF: Anatomy of an Internet Routing Protocol", Addison Wesley Longman, Inc., 1998, ISBN: 0-201-63472-4.

[MT00] C. Mas, P. Thiran, "An efficient algorithm for locating soft and hard failures in WDM networks", Proceedings of IEEE Journal on Selected Areas in Communications, Volume 18, Issue 10, Page(s):1900 – 1911, October 2000.

[MT01] C. Mas, P. Thiran, "A review on fault location methods and their application to optical networks," SPIE Optical Networks Magazine, Volume 2, Issue 4, Page(s):73-87, July/August 2001.

[MT05] C. Mas, I. Tomkos, O. K. Tonguz, "Failure Location Algorithm for Transparent Optical Networks", Proceedings of IEEE Journal on Selected Areas in Communications, Volume 23, Issue 8, Page(s):1508 – 1519, August 2005.

[Muk06] B. Mukherjee, "Optical WDM Networks", Springer, January 2006, ISBN: 978-0-387-29055-3.

[MWT02] R. Mahajan, D. Wetherall and T. Anderson, "Understanding BGP configuration", Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications (ACM SIGCOMM 2002), Page(s):3 - 16, August 2002.

[Neu02] R. E. Neuhauser, "Optical performance monitoring (OPM) in next-generation optical networks", Proceedings of SPIE Network design and management, Volume 4909, Page(s):34-42, August 2002.

[NIST] National Institute of Standards and Technology, "Catalog of various network connectivity models", http://w3.antd.nist.gov/wctg/netanal/netanal_netmodels.html.

[NJAL07] K.-K. Nguyen, B. Jaumard, C. Assi, M. Lanoue, "Toward a Distributed Control Plane Architecture for Next Generation Routers", Proceedings of 4th European Conference on Universal Multiservice Networks (ECUMN 2007), Page(s): 173 – 182, February 2007.

[NM06] N. Naas, H. T. Mouftah, "Optimum Planning of GMPLS Transport Networks", International Conference on Transparent Optical Networks (ICTON2006), Volume 3, Page(s):70-73, June 2006.

[OM05] C. Ou, B. Mukherjee, "Survivable Optical WDM Networks", Springer 2005, ISBN: 0387237569.

[PK07] S. Pachnicke, P. M. Krummrich, E. Voges, E. Gottwald, "Transient gain dynamics in long-haul transmission systems with electronic EDFA gain control", Journal of Optical Networking, Volume 6, Issue 9, Page(s):1129-1137, September 2007.

[QH06] Yang Qin, K. C. How, "Comparison studies on recovery schemes of combining protection and restoration at different layers in IP over WDM networks", 10th IEEE Singapore International Conference on Communication systems (ICCS 2006), Page(s):1 – 5, October 2006.

[RDB00] G. Rossi, T. E. Dimmick, D. J. Blumenthal, "Optical performance monitoring in reconfigurable WDM optical networks using subcarrier multiplexing," Journal of Lightwave Technology, Volume 18, Issue 12, Page(s):1639–1648, December 2000.

[RFC1322] D. Estrin, Y. Rekhter, S. Hotz "A Unified Approach to Inter-Domain Routing", RFC 1322, May 1992.

[RFC1771] Y. Rekhter, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, March 1995.

[RFC 2328] J. T. Moy, "OSPF Version 2", STD 54, RFC 2328, April 1998.

[RFC 2370] R. Coltun, "The OSPF Opaque LSA Option", RFC 2370, July 1998.

[RS02] R. Ramaswami, K. N. Sivarajan, "Optical networks: a practical perspective", San Francisco : Morgan Kaufmann Publishers, c2002, ISBN: 1558606556.

[RS03] S. Ramamurthy, L. Sahasrabuddhe, B. Mukherjee, "Survivable WDM mesh networks", Journal of Lightwave Technology, Volume 21, Issue 4, Page(s):870 – 883, April 2003.

[SC03] J. Strand, A. Chiu, "Realizing the advantages of optical re-configurability and restoration with integrated optical cross-connects", Journal of Lightwave Technology, Volume 21, Issue 11, Page(s): 2871 – 2882, April 2003.

[SHS04] L. Shyue-Kung, W. Hung-Chin, Y. Shoei-Jia, T. Yu-Cheng, "Testing and diagnosis techniques for LUT-based FPGA's", Proceedings of IEEE 13th Asian Test Symposium (ATS2004), Page(s):414 – 419, November 2004.

[Sic06] A. V. Sichani, "Signaling Protocols for Survivable All-optical Networks", Ph.D Thesis, University of Ottawa, January 2006.

[SM06] A.V. Sichani, H.T. Mouftah, "Performance Evaluation of Dynamic Restoration Techniques for Survivable Optical Networks", Proceedings 11th IEEE Symposium on Computers and Communications (ISCC 2006), Page(s):961 – 966, June 2006.

[SM07] A.V. Sichani, H. T. Mouftah, "Limited-perimeter vector matching fault-localization protocol for transparent all-optical communication networks", IEEE Journal on Selected Areas in Communications, Volume 1, Issue 3, Page(s): 472 – 478, June 2007.

[SS02] S. Stanic, S. Subramaniam, H. Choi, G. Sahin, H. Choi, "On monitoring transparent optical networks", Proceedings of International Conference on Parallel Processing Workshops (ICPP 2002), Page(s):217 – 223, August 2002.

[SS08] S. Stanic, S. Subramaniam, "Distributed Hierarchical Monitoring and Alarm Management in Transparent Optical Networks", Proceedings of IEEE International Conference on Communications (ICC '08), Page(s):5281-5285, May 2008.

[Ste99] J. W. Stewart, "BGP-4: Inter-Domain Routing in the Internet", Addison Wesley Longman, Inc., 1999, ISBN: 0-201-37951-1.

[Tho01] S. A. Thomas, "IP Switching and Routing Essentials: Understanding RIP, OSPF, BGP, MPLS, CR-LDP, and RSVP-TE", WILEY, Inc.,2001, ISBN: 0-471-03466-5.

[TLH04] S. K. Tso, J. K. Lin, H. K. Ho, C. M. Mak, K. M. Yung, Y. K. Ho, "Data mining for detection of sensitive buses and influential buses in a power system subjected to disturbances", Proceedings of IEEE Transactions on Power Systems, Volume 19, Issue 1, Page(s): 563 – 568, February 2004.

[TP04] S. W. Tak, P. Prathombutr, J. Jang, E. K. Park, "A restoration software framework for survivable WDM optical networks", Proceedings of 13th International Conference on Computer Communications and Networks (ICCCN 2004), Page(s):11 – 16, October 2004.

[TS02] P. Tomsu, C. Schmutzer, "Next Generation Optical Networks, The Convergence of IP Intelligence and Optical Technologies", Prentice Hall, 2002, ISBN: 013028226x.

[WB95] W. V. Wollman, Y. Barsoum, "Overview of open shortest path first, version 2 (OSPF V2) routing in the tactical environment", Proceedings of IEEE Military Communications Conference (MILCOM 1995), Volume 3, Page(s):925 – 930, November 1995.

[WCZ05] Y. Wen, V. M. S. Chan, L. Zhang, "Efficient fault diagnosis algorithms for all-optical WDM networks with probabilistic link failures", IEEE/OSA Journal of Lightwave Technology, Volume 23, Issue 10, Page(s): 3358–3371, October 2005.

[WK02] W. Weingartner, D. C. Kilper, "OSNR Monitoring for Fault Management in High Speed Networks", Proceedings of the 28th European Conference on Optical Communication (ECOC 2002),Volume 3, Page(s):1- 2, September 2002.

[Yan02] W. G. Yang, "Sensitivity issues of optical performance monitoring, "IEEE Photonics Technology Letters, Volume 14, Issue 1, Page(s):107–109, January 2002.

[Za03] H. Zang, "WDM Mesh Networks, Management and Survivability", Kluwer Academic Publishers, Inc., 2003, ISBN: 1 402073550.

[ZH04] H. Zeng, C. Huang, A. Vukovic, "Monitoring cycles for fault detection in meshed all-optical networks", Proceedings of International Conference on Parallel Processing/International Workshop in Optical Networks and Management (ICPP 2004), Page(s):434 – 439, August 2004.

[ZHV06] H. Zeng, C. Huang, A. Vukovic, "A Novel Fault Detection and Localization Scheme for Mesh All-Optical Networks Based on Monitoring-Cycles", Journal of Photonic Network Communications, Springer US, Volume 11, Issue 3, Page(s): 277-286, May 2006.

[ZM04] J. Zheng, H. T. Mouftah, "Optical WDM Networks, Concepts and Design principles", Wiley-Interscience, 2004, ISBN: 0471671703.

[ZPB06] J. Zheng, C. Peng, G. v. Bochmann, "Fault Detection and Localization Scheme for All-Optical Overlaid-Star TDM Networks", First International Conference on Communications and Networking in China (ChinaCom 2006), Page(s):1 - 9, October 2006.

[ZV07] H. Zeng, A. Vukovic, "The Variant Cycle-Cover Problem in Fault Detection and Localization for Mesh All-Optical Networks", Photonic Network Communications, Springer US, Volume 14, Issue 2, Page(s): 111-122, 2007.

[ZVS04] H. Zeng, C. Huang, A. Vukovic, M. Savoie, "Fault Detection and Path Performance Monitoring in Meshed All-Optical Networks", Proceedings of IEEE Global Telecommunications Conference 2004 (GLOBECOM 2004), Volume 3, Page(s):2014-2018, November - December 2004.

[ZVSH04] H. Zeng, A. Vukovic, M. Savoie, C. Huang, "Wavelength-Routing Fault Detection in an AON Testbed Utilizing Concatenated Pilot Tones", Proceedings of SPIE Photonics North 2004 (SPIE PN 2004), Volume 5579, Page(s): 620-627, September 2004.

Biography

Mazen Khair graduated from Al-Mustansiriya University where he received his B.Sc. from the Electrical and Computer Engineering Department in 1997, and his M.Sc. in 2004 from the School of Information Technology and Engineering at University of Ottawa. He was an intern at the Communication Research Center (CRC) of Canada between March 2002 and March 2004. He enrolled in the Ph.D. program of the same institute in 2004, and is a research and teaching assistant at the School of Information Technology and Engineering of University of Ottawa since 2002. He is a student member of IEEE.

During his work on Ph.D. thesis, he published the following papers:

Journal papers:

- A. Hafid, A. Maach, M. Khair, and J. Drissi , "Advance lightpath provisioning in inter-domain optical networks", Journal of Optical Networking, Volume 4, Issue 11, Page(s):714-736, August 2005.

- M. Khair, B. Kantarci, J. Zheng, and H. T. Mouftah, "Optimization for Fault Localization in All-Optical Networks", Journal of Lightwave technology. (Under review)

International conference papers:

- M. Khair, B. Kantarci, J. Zheng, and H. T. Mouftah, "Performance Optimization for Fault Localization in All-Optical Networks", The fifth international conference on broadband communications, networks and systems (BroadNets08), September 2008, London, UK.

- M. G. Khair, J. Zheng, and H. T. Mouftah, "Distributed Fault Localization for Multi-Domain All-Optical Networks with Partial Power Monitoring", IEEE Symposium on Computers and Communications (ISCC'08), July 2008, Marrakech, Morocco.

- M. G. Khair, B. Kantarci, J. Zheng, and H. T. Mouftah, "Optimization for Minimizing Fault Localization Time in All-Optical Networks", 10th Anniversary International Conference on Transparent Optical Networks (ICTON 2008), Volume 3, Page(s): 63-66, June 2008, Athens, Greece.

- M. Khair, J. Zheng, and H. T. Mouftah, "Distributed Fault Localization for Multi-Domain All-Optical Networks without Power Monitoring", 13th International Telecommunications Network Strategy and Planning Symposium (NETWORKS 2008), September-October 2008, Budapest, Hungary.

- M. G. Khair, B. Kantarci, J. Zheng, and H. T. Mouftah, "Connection Provisioning Constrained to Fault Localization in All-optical Networks", 23rd International Symposium on Computer and Information Sciences (ISCIS 2008), October, Istanbul, Turkey.

- M. G. Khair, T. J. Hall, "Just Enough Routing and wavelength Assignment (RWA) scheme for metro-ring architecture", The 9th International Conference on Advanced Communication Technology (ICACT2007), Volume 2, Page(s):1397 – 1401, February 2007, Phoenix Park, Republic of Korea.

- R. Radziwilowicz, M. G. Khair, S. A. Paredes, T. J. Hall, "Design issues for edge nodes in agile all-photonic networks", 2006 Workshop on High Performance Switching and Routing (HPSR 2006), Page(s):63-69, June 2006, Poznan, Poland.

- A. Hafid, M. Abdelilah, K. G. Mazen, J. Drissi, "Optical routing border gateway protocol-based advance lightpath setup", Proceedings of the 2005 Systems Communications (ICW '05) , Page(s):223 – 228, August 2005, Montreal, Canada.

- M. Khair, J. Zheng, and H. T. Mouftah, "Distributed Multi-Failure Localization Protocol for All-Optical Networks", The 13th International Conference on Optical Networking Design and Modeling (ONDM 2009), February 2008, Braunschweig, Germany. "Under review".

Appendix A

Confidence Interval Estimation

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

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

Therefore, we have, $P\left[L = \mu - 1.96\,\delta/\sqrt{N} \le \overline{X} \le \mu + 1.96\,\delta/\sqrt{N} = U\right] = 0.95$.

For instance, if we consider a group of data for 12 independent runs per simulation experiment: $X_1, X_2, \cdots, X_n$, = 20, 35, 29, 25, 35, 39, 22, 27, 37, 40, 36, 21, then, we calculate our parameters including $\overline{X}$ = 30.5, δ = 7.342281, and $1.96\,\delta/\sqrt{N}$ = 4.17124285. As a result, confidence limits would be $[L = 26.32875715,\ U = 34.67124]$. In our simulation results, 90% confidence intervals are depicted by "-".