OUKA
Osaka University Knowledge Archive

Title	On the Design Methods of Effective and Reliable IP over WDM Networks
Author(s)	Arakawa, Shin'ichi
Citation	大阪大学, 2003, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/1420
rights	
Note	

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

On the Design Methods of Effective and Reliable IP over WDM Networks

Shin’ichi Arakawa

Department of Informatics and Mathematical Science

Graduate School of Engineering Science

Osaka University

Preface

The popularity of the Internet and advancements in multimedia communication technologies have led to an exponential growth in Internet traffic. Moreover, IP (Internet Protocol) is emerging as a dominant technology, so the ability to carry IP traffic efficiently is an important issue for the next–generation data–centric Internet. WDM (Wavelength Division Multiplexing) technology provides multiple wavelengths each of the order of 10 Gbps. While link capacity can be increased by increasing the number of wavelengths multiplexed on the fiber, doing this does not eliminate the network bottlenecks caused by explosive traffic growth. It simply shifts the bottleneck to the electronic routers.

Recent advances in optical switches have led to WDM technology with networking capabilities. Using optical switches, a logical network consisting of wavelength channels (lightpaths) is built on a physical WDM network. The IP packets are carried on the logical network; the underlying WDM network provides only logical paths between the nodes. Networking in the optical domain has the potential to offer several network control functionalities such as functions of routing, congestion control, and reliability. While the TCP/IP protocol suite also has these functions, building the same functionality into more than one layer often gives an ill effect for the objective. An example of this is the TCP over ATM ABR service class, in which both the TCP and ABR service classes have their own congestion control mechanism. If the parameters of two layers are appropriately chosen, congestion control works very well. However, if they are not, the performance can degrade unexpectedly and actually become worse than with no congestion control mechanism. A more appropriate scenario for next–generation IP over WDM networks is limited use of the network control functionalities of WDM networks.

In this thesis, we propose methods for designing an effective and reliable IP over WDM network. When the WDM network is applied to an IP, the packet route is determined by the

routing protocol provided by the IP layer; therefore the end-to-end path provided by the logical topology of the WDM network is not suitable for an IP since an IP has its own metrics for route selection. Furthermore, the WDM network can also provide the IP layer with a reliability function. The routing protocol of an IP can find a detour and restore the traffic flow when a network component fails, but usually after a noticeable delay (typically 30 sec for updating the routing table). The reliability mechanism provided by the WDM network layer can offer much faster recovery, typically less than 50 msec.

We first describe a logical topology design algorithm in which a packet route is determined by the routing protocol provided by the IP layer. Thus, in designing the logical topology, the routes of the lightpaths are determined by considering the nature of the IP routing protocol. That is, we place lightpaths so that an IP packet experiences smaller delays on its end-to-end path. For this purpose, we try to reduce the number of electronic nodes, in addition to reducing the propagation delays between two end nodes. The routing stability of IP is another important issue in designing IP over WDM networks. Researches typically assume that the amount of traffic between nodes is given and fixed. In building IP networks, however, the issue of routing stability should also be considered. We compare the packet delays on the shortest and second-shortest end-to-end paths, and if the delays are much different, we conclude that the logical topology is robust against traffic fluctuation. We show through numerical examples that our algorithm is robust against routing instability. The simulation results show that our logical topology design algorithm increases the maximum throughput by 10–50% without sacrificing routing stability.

We next discuss the interaction between IP–layer reliability and optical–layer survivability, assuming that some lightpaths are protected by a WDM protection mechanism and the rest are restored by the IP–layer routing. To construct a reliable IP over WDM network, backup paths as well as primary paths should be embedded within the logical topology. The best approach to doing this depends on the protection schemes used and the types of failures that may occur. One approach is to recover from all types of failures in the optical layer, but this may require many wavelength resources and is thus less effective. In IP over WDM networks, IP routing has its own routing mechanism, so it may not be necessary to protect all the lightpaths by using the optical layer if doing so does not lead to cost savings even when a shared protection scheme is used. If we allow that several primary lightpaths cannot recover from some failure patterns and

that the resilience is left to the IP layer, we can expect more cost savings. We first formulate the design of a reliable IP over WDM network as an optimization problem and then propose a heuristic algorithm to relax the computational complexity. Using numerical examples, we show that our algorithm is best if the primary concern is traffic load at the IP routers after a failure. Based on our algorithm, we also discuss the effect of the interaction between the IP and WDM layers.

We next introduce QoR (Quality of Reliability), which is a concept related to QoS (Quality of Service) that reflects reliability in a WDM network. QoR can be used to guarantee a maximum recovery time set, based on the user's request, and guarantee that backup lightpaths will be available. In the conventional quality–based lightpath configuration methods, the failure-recovery quality is guaranteed only probabilistically. That is, these methods are aimed at improving the effective usage of network resources, but at the cost of a 100% guarantee of failure recovery. Thus, we introduce QoR as a new QoS metric aimed at providing highly reliable lightpaths. With QoR, both the time needed to recover from a single failure and 100% failure recovery are guaranteed, because building a highly reliable network is becoming increasingly more important than using network resources efficiently, especially as the number of wavelengths rises with advances in WDM technology. For this purpose, we propose heuristic algorithms; the results show that the number of wavelengths necessary to satisfy the QoR is reduced by 20 – 30.

In the above studies, we assume that traffic demand is known a priori. This assumption is, however, inappropriate when WDM technology is applied to the Internet. A more flexible network provisioning approach is necessary for the Internet. For capacity dimensioning of reliable IP over WDM networks, we propose a new approach, called "incremental capacity dimensioning", to designing the logical topology. There are three phases: an initial phase, an incremental phase, and a readjustment phase. With our approach, the logical topology can be adjusted according to incrementally changing traffic demand. During the incremental phase, primary paths are added as traffic increases. At the same time, the backup lightpaths are reconfigured since they do not affect the traffic carried on the operating primary paths. We describe a heuristic algorithm for selecting the set of backup lightpaths to be configured and formulate an optimization problem for reconfiguring them. The results show that the total traffic volume that the IP over WDM network can accommodate is improved by using our algorithm.

Acknowledgements

I would like to express my sincere appreciation to my adviser, Prof. Masayuki Murata of Osaka University, for his innumerable help, patience, encouragement, and continuous support. His technical and editorial advice was essential to the completion of this dissertation and has given me insights into the workings of academic research in general.

I am heartily grateful to the members of my dissertation committee, Prof. Hideo Miyahara, Prof. Makoto Imase, and Prof. Teruo Higashino, for reading my dissertation and providing many valuable comments.

I would like to acknowledge and thank Associate Prof. Naoki Wakamiya, Associate Prof. Hioryuki Ohsaki, and Associate Prof. Go Hasegawa of Osaka University for their helpful comments and feedbacks.

I would like to thank Prof. Ken–ichi Kitayama of Osaka University, and Dr. Hiroaki Harai of Communications Research Laboratory for enlightening discussions.

My thanks also go to my friends in the department for their inciting discussions, fellowship, and underpinning.

Last, but not least, I would like to express my deepest gratitude to my parents for their dedication and the many years of support they provided during my undergraduate and graduate studies, which provided the foundation for this work.

List of Papers

Book Chapter

1. S. Arakawa and M. Murata, *Reliability Issues in IP over Photonic Network*. Quality, Survivability and Reliability of Large Scale Systems – Case Studies: Olympic Games John Wiley & Son, Dec. 2002.

Journal papers

1. S. Arakawa, J. Katou, and M. Murata, "A design method for logical topologies with stable packet routing in IP over WDM networks," submitted to *IEICE Transactions on Communications*.

2. S. Arakawa, M. Murata, and H. Miyahara, "Functional partitioning for multi-layer survivability in IP over WDM networks," *IEICE Transactions on Communications*, vol. E83-B, pp. 2224–2233, Oct. 2000.

3. S. Arakawa, J. Katou, and M. Murata, "Design method of logical topologies with quality of reliability in WDM networks," to appear in *Photonic Network Communications*.

4. S. Arakawa and M. Murata, "Lightpath management of logical topology with incremental traffic changes for reliable IP over WDM networks," *Optical Network Magazine*, vol. 3, pp. 68–76, May 2002.

Refereed Conference papers

1. J. Katou, S. Arakawa, and M. Murata, “A design method of logical topology for IP over WDM networks with stable routing,” in *Proceedings of The Fifth Working Conference on Optical Network Design and Modeling (ONDM 2001)*, pp. 61–78, Feb. 2001.

2. S. Arakawa, M. Murata, and H. Miyahara, “Design methods of multi-layer survivability in IP over WDM networks,” in *Proceedings of OptiComm*, pp. 279–290, Oct. 2000.

3. J. Katou, S. Arakawa, and M. Murata, “Design method of logical topology in WDM network with quality of protection,” in *Proceedings of Workshop on Optical Networking: Technologies, Architectures and Management*, Nov. 2001.

4. S. Arakawa and M. Murata, “On incremental capacity dimensioning in reliable IP over WDM networks,” in *Proceedings of OptiComm*, pp. 153–163, Aug. 2001.

Non–Refereed Technical papers

1. S. Arakawa, J. Katou, and M. Murata, “A design method of logical topology with stable packet routing in IP over WDM network,” *IEICE General Conference*, Mar. 2001.

2. J. Katou, S. Arakawa, and M. Murata, “A design method of logical topology and its influence on IP routing in IP over WDM network,” *Technical Report of IEICE (*PNI2000-35*)*, pp. 32–38, Mar. 2001.

3. S. Arakawa, M. Murata, and H. Miyahara, “Design of lightpath networks with protections for IP over WDM networks,” *Technical Report of IEICE (*SSE99-111*)*, pp. 7–12, Nov. 1999.

4. S. Arakawa, M. Murata, and H. Miyahara, “Functional partitioning for multi-layered survivability in IP over WDM networks,” *Technical Report of IEICE (*PNI2000-3*)*, pp. 16–25, Oct. 2000.

5. J. Katou, S. Arakawa, and M. Murata, “Design method of logical topologies in WDM network with quality of protection,” *Technical Report of IEICE (*NS2001-212*)*, pp. 41–46, Feb. 2001.

6. S. Arakawa and M. Murata, “Incremental capacity dimensioning for reliable IP over WDM networks,” *Technical Report of IEICE (*PNI2000-34*)*, pp. 24–31, Mar. 2001.

7. S. Arakawa, S. Ishida, and M. Murata, “Management of logical topologies for dynamically changing traffic in reliable IP over WDM networks,” *Technical Report of IEICE (*DC2002-2*)*, pp. 7–14, Apr. 2002.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The rapid growth in the number of users and in the number of multimedia applications on the Internet is dramatically increasing traffic volumes on networks. WDM (Wavelength Division Multiplexing) technology, which provides multiple wavelengths each of the order of 10 Gbps on a fiber, is currently used by many ISPs (Internet Service Providers). While WDM technology offers a low-cost solution to the problem of growing traffic demands, the use of WDM for transporting IP traffic is still controversial. One aspect of the controversy is related to the protocol stack. ATM (Asynchronous Transfer Mode) technology can be used on SONET/SDH (Synchronous Optical NETwork / Synchronous Digital Hierarchy). If SONET, in turn, is built on WDM technology, the result is an *IP over ATM over SONET over WDM network*. An *IP over SONET over WDM network* is another solution since the ATM technology introduces only a protocol overhead (i.e., 5 byte cell header within a 53-byte cell). Another and more promising solution is the IP over WDM networks in which WDM is directly used by the IP (more exactly, there is a data link layer protocol such as PPP or HDLC between the IP and the WDM protocol stack).

There are several alternatives even for IP over WDM networks, depending on whether we utilize the capabilities of the WDM network or not. They include functions for routing, congestion control, and reliability. Currently, commercially available WDM transmission systems use WDM technology only on their fiber links (see Fig. 1.1). Each wavelength in a fiber is treated

as a physical link between network components (e.g., routers and switches). That is, each wavelength on the fiber is treated as a physical link between conventional IP routers, meaning that the conventional IP technique for handling multiple links can be used. Link capacity is increased by increasing the number of wavelengths on the fiber, which may eliminate bandwidth bottlenecks in the link. However, simply eliminating link bottlenecks in the face of exploding demand is not enough because it only shifts the bottlenecks to the electronic routers.

One way to eliminate router bottlenecks is to introduce optical switches. Suppose that each node has an optical switch directly connecting each input wavelength to an output wavelength, so that there is no electronic processing at the packet level. That is, no electronic routing is needed at the nodes. A wavelength path can be set up directly between two nodes via one or more optical switches (i.e., cross–connect switches). The intermediate nodes along the wavelength path are released from electronic routing, thereby eliminating the bottlenecks at the electronic routers.

If lightpaths are placed between every two nodes, then no electronic processing is necessary within the network. We can see such an example in a MPLS (Multi Protocol Label Switching) network [1]; the applicability of MPLS to the IP over WDM Network is now being discussed [2], which views the wavelength as a label. However, a great many wavelengths are necessary to establish such a network [3]. We thus need a compromise — establish a logical topology consisting of lightpaths by using the available wavelengths as much as possible. If a direct lightpath cannot be set up between two nodes, two or more lightpaths are used by packets to reach the destination. A example logical topology, or what we call *WDM path network*, is shown in Fig. 1.1. Packets sent from node N_1 to N_3 are forwarded on the direct logical path using wavelength λ_2. However, packets sent from node N_1 to N_4 takes two hops since there is no direct logical path. That is, the packets are first forwarded to node N_4 and then passed to node N_2. A logical view of the underlying network (i.e., the logical topology established by the WDM network) is shown in Fig. 1.3.

Many researchers have developed methods for designing the logical topology [4–6]. For example, the authors in [5] formulated a method for designing logical topology as an optimization problem and showed that the problem is NP-hard. In [6], the authors considered the logical topology design problem together with the packet routing problem so as to maximize the network throughput. Since the combined problem is computationally hard to solve, it is split

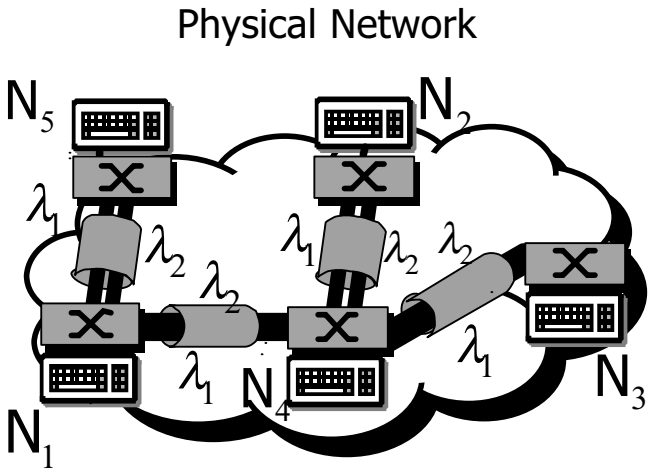


Figure 1.1: Physical WDM network: Optical nodes are connected using optical fibers

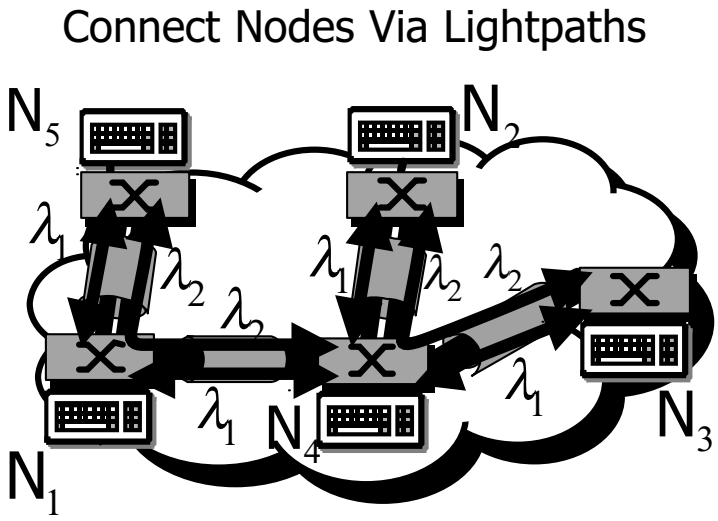


Figure 1.2: Constructing logical topology by configuring lightpaths

into two subproblems, which are solved independently. The routing problem is formulated as a linear-programming problem by imposing a delay constraint on each node pair. Several heuristics have also been proposed to relax the computational burden. However, when the WDM network is applied to IP, a packet route is determined by the routing protocol provided by the IP layer; therefore, the end-to-end path provided by the logical topology of the WDM network is not suitable for the IP, since the IP has its own metrics for route selection.

Advances in WDM technology lead to very high capacity networks, which will drive the need for a reliability mechanism embedded in the logical topology. “Reliability mechanism” is a functionality that enables recovery from unexpected failures of network components. Networks

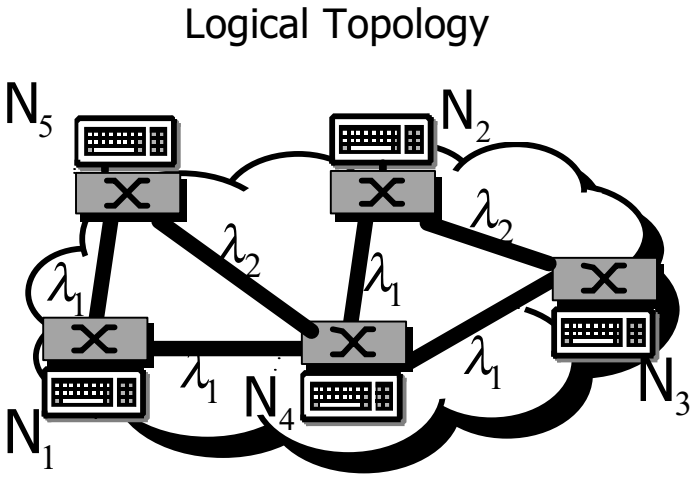


Figure 1.3: Logical topology as seen from upper layer protocol

will have to operate 99.999% of the time, meaning that downtime must be no more than five minutes per year. Without a reliability mechanism, the failure of a network component can lead to the loss of a large amount of data. In a traditional synchronous optical network/synchronous digital hierarchy (SONET/SDH) ring network, a backup fiber is allocated for each working fiber, in the case of a 1:1 protection scheme, and automatic protection switching [7] provides the reliability mechanism. Fiber allocation is sufficient for SONET/SDH networks because the optical signal is converted into an electronic signal at each node. However, in WDM networks, the optical signals, which are transparent to the upper layer protocol (e.g., IP, SONET/SDH, and ATM), may pass through successive network components. Thus, coordination of a reliability mechanism for each lightpath, end to end, is necessary for WDM networks.

Of course, in IP over WDM networks, IP itself has a reliability mechanism: link and/or node failures are avoided by finding a detour and then routing the IP traffic through it. However, the exchange interval of the routing metrics is long (e.g., 30 sec). In contrast, a new route can be established within a few tens of milliseconds following a failure in WDM networks. A reliability mechanism in the optical domain is not always an optimal solution because of the physical constraints on the number of wavelengths that can be carried in a fiber. By combining a reliability mechanism in the optical domain with one in the electronic domain, we can obtain more reliable networks than the current Internet.

1.2 Protection/Restoration Schemes

Reliability mechanisms in WDM networks can be roughly categorized into protection schemes and restoration schemes. Protection schemes allocate explicit resources for backup purposes, so they consume wavelengths. Restoration schemes do not allocate explicit resources in advance of a failure, so they do not consume wavelengths. When a failure occurs, backup paths are dynamically calculated and configured based on the current usage of network resources. The advantage of restoration schemes is that wavelength resources are not tied up for backup. However, they do not guarantee failure recovery. While protection schemes waste resources, they do guarantee recovery. Protection schemes can be further classified into *dedicated protection* schemes, in which a backup lightpath is dedicated to a primary lightpath, and *shared protection* schemes, in which several primary lightpaths can share the same wavelength as a backup lightpath, as long as their primary lightpaths are *failure independent*. Two primary lightpaths are failure independent if they do not share components that may fail.

1.2.1 Failures

We can consider three types of failure scenarios: *laser* failure, *link* failure, and *node* failure. A laser failure is a single–wavelength failure, caused by the failure of the transmitter or receiver designated for the wavelength. A link failure is caused by a fiber cut. If this happens in a WDM network, multiple lightpaths must be re–routed or switched onto backup paths. In the case of a node failure, a backup path must be set up for each lightpath passing through the failed node. Thus, a node failure is the most severe of the three scenarios. These failures can be detected by monitoring the optical signals passing through the network components. If a failure occurs, the node nearest to the failure components switches to a backup path if there is link protection. If there is path protection, the originating node of the corresponding lightpath switches to a backup lightpath. Before a network provider can replace a failed component, its location must be determined. Techniques for doing this are summarized elsewhere [8].

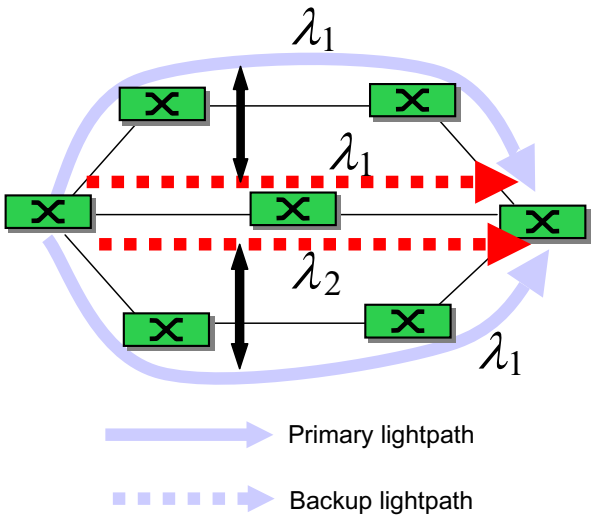


Figure 1.4: Dedicated protection

1.2.2 Dedicated Protection Schemes

As mentioned above, in dedicated protection schemes, a backup lightpath is dedicated to each primary lightpath, creating “1+1” or “1:1” protection (see Fig. 1.4). Each backup lightpath carries a copy of the signal carried by the corresponding primary lightpath in “1+1” protection. The receiver node thus receives two signals, one from the primary lightpath and one from the backup. It selects the better of the two signals. In the “1:1” scheme, a copy is not normally carried on the backup lightpath. The backup is used only when a failure occurs. The “1+1” scheme is thus worse in terms of bandwidth utilization because the bandwidths of the backup lightpaths are always being used for primary lightpath backup, while they can be used for low–priority IP traffic in the “1:1” scheme. However, since little coordination is needed to recover from failures, the recovery time is shorter in the “1+1” scheme.

1.2.3 Shared Protection Schemes

As mentioned above, in shared protection schemes, several primary lightpaths share one backup path. They can do this if the type of failures they cover are relaxed. A shared protection scheme must be carefully engineered so that any two primary lightpaths do not use a backup lightpath at the same time if a failure occurs. The backup resources are thus more effectively utilized, as shown in Fig. 1.5. Ramamurthy and Mukherjee, for example, showed that wavelength resources can be reduced by 20–44% using a shared path protection scheme [9].

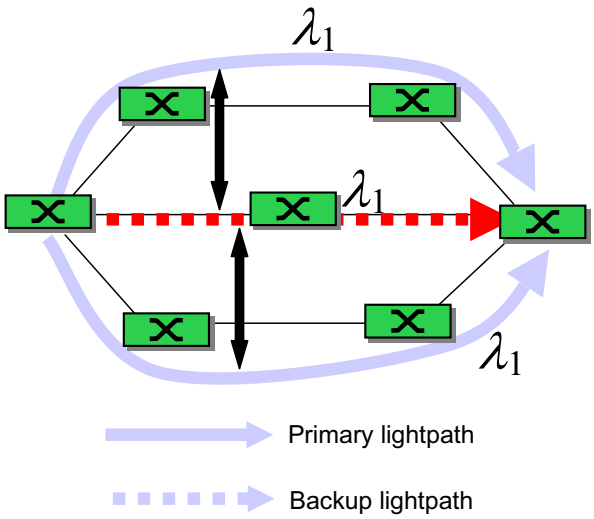


Figure 1.5: Shared protection

1.2.4 Restoration Schemes

Because a restoration scheme in the optical layer allocates a backup path only after failure occurs, wavelength resources can be more effectively used for transporting IP traffic. However, calculating alternate routes following a failure can take seconds or even minutes. Thus, a restoration scheme is usually combined with a protection scheme [10]. After failure recovery is completed by the protection scheme, a restoration scheme is used to provide either more efficient routes or additional protection against further failures before the first failure is fixed. A centralized management system can be used to calculate the alternate routes, and more sophisticated algorithms can be used to reduce the excess bandwidth required, so more complex mesh topologies can be supported.

1.3 Outline of Thesis

As discussed in Section 1.1, it is necessary to consider how to apply the networking capability of WDM to IP. A WDM network using optical switches has the potential to offer several network functionalities such as functions of routing, congestion control , and reliability. While the TCP/IP protocol suite also has these functions, building the same functionality into more than one layer often gives an ill effect for the objective. A more appropriate scenario for next–generation IP over WDM networks is limited use of the network control functionalities of WDM

networks.

In this thesis, we first focus on the design of the end-to-end path provided by the logical topology of the WDM network to incorporate the IP route selection mechanism. We next focus on the reliability functions of IP and WDM. As described in Sections 1.2.2, 1.2.3, 1.2.4, the WDM network layer offers a reliability mechanism to the upper layer. The IP routing protocol can find a detour and restore traffic flow when a network component fails. By combining those two mechanisms appropriately, we should be able to construct networks much more reliable than the current Internet. In this thesis, we propose design methods to improve reliability in IP over WDM networks. A protection method is considered to provide a faster failure recovery.

Method for Designing Logical Topologies with Stable Packet Routing [11–14]

First, in Chapter 2, we describe a logical topology design algorithm in which a packet route is determined by the routing protocol provided by the IP layer. In IP over WDM networks, a packet route is determined by the routing protocol provided by the IP layer, and the underlying WDM network provides only (logical) paths between nodes. Thus, in designing the logical topology, the routes of the lightpaths are determined by considering the nature of the IP routing protocol. That is, we place lightpaths so that the IP packet experiences smaller delays on its end-to-end path. For this purpose, we try to reduce the number of electronic nodes, in addition to reducing the propagation delays between two end nodes. The routing stability of IP is another important issue in designing IP over WDM networks. Researchers typically assume that the amount of traffic between nodes is given and fixed. In building IP networks, however, the issue of routing stability should also be considered. In this thesis, we compare the packet delays on the shortest and second-shortest end-to-end paths, and if the delays are much different, we conclude that the logical topology is robust against traffic fluctuation. We will show through numerical examples that our proposed algorithm is robust against routing instability.

Functional Partitioning for Multi-layer Survivability in IP over WDM Networks [15–19]

To construct a reliable IP over WDM network, backup paths as well as primary paths should be embedded within the logical topology. The best approach to doing this depends on the protection schemes used and the types of failures that may occur. One approach is to recover from all types of failures in the optical layer, but this may require many wavelength resources and is thus less effectiveness. Among the several types of failures described in Section 1.2.1, we consider single–fiber failure. Multiple failures and node failures are assumed to be handled by the restoration functionality of the IP layer. We first discuss the reliable design for the single–layer case, i.e., recovery from a single–fiber failure is guaranteed in the WDM network.

We next describes multi–layer survivability, in which a subset of lightpaths are protected against failure. It may not be necessary to protect all the lightpaths by using the optical layer if doing so does not lead to cost savings even when a shared protection scheme (Section 1.2.3) is used. If we allow that several primary lightpaths cannot recover from some failure patterns and that the resilience is left to the IP layer, we can expect more cost savings. Consider the extreme case in which all wavelengths are used to establish the primary lightpaths, and no protection is established because failures are expected to seldom take place. Performance is maximized at the price of reliability. In this chapter, we discuss the interaction between IP–layer reliability and optical–layer survivability, assuming that lightpaths in a subset are protected by a WDM protection mechanism and that the rest are restored by the IP–layer routing function.

Methods for Designing Logical Topologies for Quality of Reliability [20–22]

Recent research has focused on providing QoS (Quality of Service) with respect to failure recovery in an optical WDM network [23–25]. Saradhi and Murthy introduced the concept of an *R–connection* [23]. They considered for use in dynamically establishing a reliable connection. The basic idea of the R–connection is that an application user specifies the level of reliability. The reliability levels of the connection are calculated based on a pre–specified reliability measurement for each network component. If the reliability requirement is not satisfied, the length of the primary lightpath covered by the partial backup lightpath is selected so as to enhance the reliability of the R–connection. Another way to provide QoP (Quality of Protection) is to use

the differentiated reliability (DiR) of a connection [24, 25]. This is the maximum probability that the connection will fail due to a single network component failing. With this approach, a continuous spectrum of reliability levels is provided. QoP was introduced to realize QoS in an optical network [24] through a probabilistic failure recovery model in which only a certain fraction of the traffic, specified by the user, is restored after failure. A approach different from previous ones [9, 24, 26] is to consider the possibility of two or more components failing at the same time (a multiple–failure assumption) and assume that each primary lightpath has its own reliability metric that can be determined from the failure probabilities of the network components [23]. Based on this approach, backup lightpaths are partially configured for the primary lightpath based on the specified probability.

However, in conventional QoP–based lightpath configuration methods, the failure-recovery quality is guaranteed only probabilistically. That is, these methods are aimed at improving the efficient usage of network resources, but at the cost of a 100% guarantee of failure recovery. We introduce QoR (Quality of Reliability) as a new QoS metric aimed at providing highly reliable lightpaths. With QoR, both the time needed to recover from a single–component failure and 100% failure recovery are guaranteed. Building a highly reliable network is becoming increasingly more important than using network resources efficiently, especially as the number of wavelengths rises with advances in WDM technology. Our approach is to build a logical topology by effectively using the available wavelengths in a way that guarantees the failure-recovery time and 100% failure recovery.

Incremental Lightpath Management for IP over WDM Networks [27–30]

Much of the previous research, including [9] and [18], assumed that traffic demand is known a priori, and from it, the optimal structure of the logical topology is obtained. Such an assumption is, however, inappropriate, especially when WDM technology is applied to the Internet. For traditional telephone networks, a network provisioning (or capacity dimensioning) method is well established. The target call blocking probability is first set, and then the number of telephone lines (or the capacity) needed to meet the requirement on the call blocking is determined. After installing the network, the traffic load is continuously measured, and as necessary, link capacity is increased to accommodate increased traffic. With this feedback loop, a telephone network

is well engineered to provide QoS in terms of call blocking probabilities. There are several rationales behind this successful approach.

- The call blocking probability is directly related to the user's perceived QoS.
- Capacity provisioning is easily based on stably growing traffic demands and the richness of historical statistics.
- There is a well-established fundamental theory, i.e., the Erlang loss formula.
- The network provider can directly measure a QoS parameter (the blocking probability) by monitoring the numbers of generated and blocked calls.

On the other hand, a network provisioning method suitable to the Internet has not yet been established, mainly due to three obstacles.

- The statistics obtained by traffic measurement are at the packet level, so the network provider cannot monitor or even predict the user's QoS.
- The explosion in traffic growth in the Internet makes it difficult to predict future traffic demand.
- There is no fundamental theory for the Internet like the Erlang loss formula in the telephone network.

Queueing theory has a long history and has been used as a fundamental theory for data networks (e.g., the Internet). However, queueing theory gives only the packet queueing delay and loss probability at the router, and router performance is only one component of the user's perceived QoS on the Internet. Furthermore, the packet behavior at the router is affected by the dynamic behavior of TCP, which is essentially window-based feedback congestion control [31].

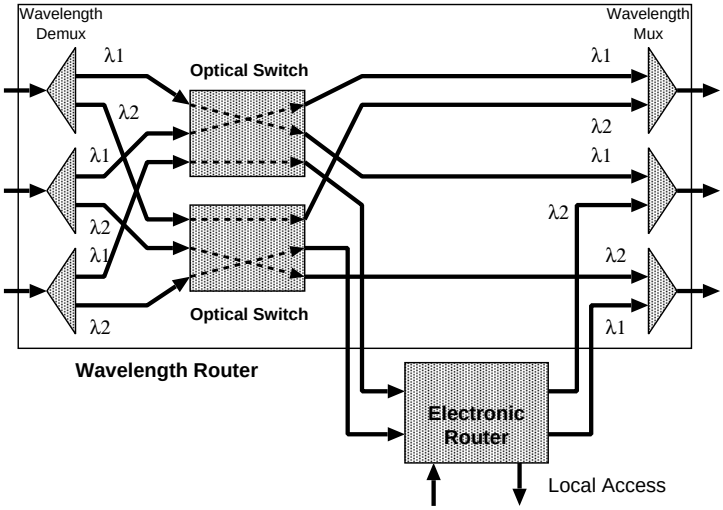
"Static" design in which the traffic load is assumed to be given a priori is thus completely inadequate. Instead, a more flexible network provisioning approach is necessary in the era of the Internet. Fortunately, the IP over WDM network can establish a feedback loop by using wavelength routing. If it is found through a traffic measurement that the user's perceived QoS is not satisfactory, new wavelength paths can be set up to increase the path bandwidth (i.e., the number of lightpaths).

For this purpose, we propose an "incremental logical topology management scheme", consisting of three phases for setting up primary and backup lightpaths; an *initial phase*, an *incremental phase*, and a *readjustment phase*. During the initial phase, a reliable IP over WDM network is built by setting up both primary and backup lightpaths. In this phase, even though we do not know the traffic demand, we have to establish the network. We do this by using statistics on traffic demands. In this scheme, the logical topology can easily be reconfigured , which is done in the incremental phase. During the incremental phase, the logical topology is reconfigured based on requests to set up new lightpath(s) due to changes in traffic demand or to mis–projections of traffic demand. We formulate the process of setting up lightpaths as an optimization problem. We also describe a heuristic algorithm, called a MRB (Minimum Reconfiguring for Backup lightpaths) algorithm, for selecting an appropriate wavelength. During the incremental phase, the backup lightpaths are reconfigured to achieve optimality. However, an incremental setup of the primary lightpaths may not lead to an optimal logical topology, and our logical topology might be under utilized compared to one designed using a static approach. Therefore, during the readjustment phase *both* primary and backup lightpaths are reconfigured. Note that the established lightpaths should be readjusted one-by-one so that service is not interrupted. In this thesis, however, we will mainly discuss the incremental phase; the issues related to the readjustment phase remain topics of future research.

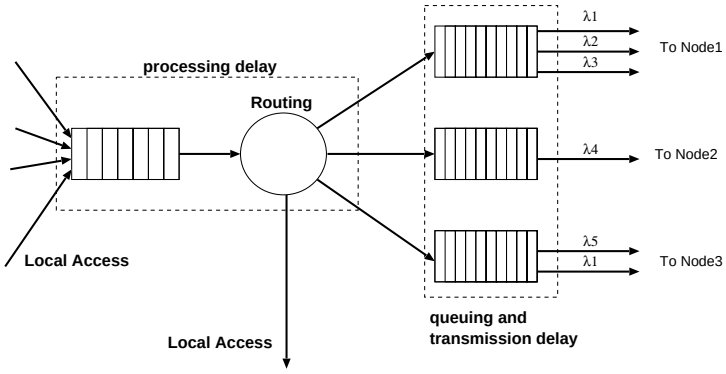
Chapter 2

Method for Designing Logical Topology with Stable Packet Routing in IP over WDM Networks

In this chapter, we describe a logical topology design algorithm in which a packet route is determined by the routing protocol provided by the IP layer. In designing the logical topology, the routes of the lightpaths should be determined by considering the nature of the IP routing protocol. That is, we place lightpaths such that the IP packet experiences smaller end-to-end delay. For this purpose, we try to reduce the number of electronic nodes, in addition to reducing the propagation delays between the end nodes. The routing stability of IP is another important issue in designing IP over WDM networks. Researchers typically assume that the amount of traffic between nodes is given and fixed. In building IP networks, however, the issue of routing stability should also be considered. We compare the packet delays on the shortest and second–shortest end-to-end paths, and if the delays are much different, we conclude that the logical topology is robust against traffic fluctuation. We will show through numerical examples that our proposed algorithm is robust against routing instability.



(a) Node architecture



(b) Model of electronic router

Figure 2.1: Node architectural model

2.1 Node Architectural Model

We first describe the architectural model of a node in the network. Figure 2.1 shows our architectural model of an optical node. Every optical node is equipped with optical switches and an electronic router. Each optical switch consists of three main components: input section, non-blocking switch, and output section. In the input section, optical signals are demultiplexed into W fixed wavelengths, λ_1,... λ_w. Each wavelength is switched into an appropriate output port of a non-blocking switch without changing its wavelength. The output section again multiplexes the output wavelengths of non-blocking switches into the fiber, and the optical signals goes to the next node. Note that a lightpath is set between two nodes by configuring non-blocking

switches along the path so that packets on a particular wavelength from the input port to the output port are forwarded with no electronic processing.

As described in Section 1.1, to reduce the number of wavelengths necessary, one-hop lightpaths are not always provided for all end-node pairs. If the lightpath is terminated at the node within the network, IP packets on that lightpath are converted to electronic signals and forwarded to the electronic router. The electronic router processes packet forwarding, just the same as conventional routers. If the packet should be further forwarded to other nodes, the electronic router puts it on the appropriate lightpath.

A model of electronic router is shown in Fig. 2.1(b). IP packets, which come from an optical switch or local access, are first buffered, and then these packets are processed on a FIFO (First In First Out) basis. When the packets are forwarded to the network, they are queued on the appropriate output-port buffer. In this chapter, we assume that multiple lightpaths between an adjacent node pair share the same buffer. We last note that the other structures of optical nodes can also be considered, but the above–mentioned node architecture is preferable since there is no need to modify the IP routing mechanism.

2.2 Design Algorithm for Logical Topology

A heuristic algorithm called MLDA (Minimum delay Logical topology Design Algorithm) was developed for establishing a logical topology [32]. MLDA works as follows. First, it places a lightpath between two nodes if there is a fiber directly connecting them. Then, it attempts to place lightpaths between nodes in the order of descending traffic demand. Finally, if there are any non–used wavelengths, it places as many lightpaths as possible randomly by using those wavelengths. Many conventional methods, including MLDA, focus on maximizing throughput, but they are not adequate for designing a logical topology suitable for carrying IP traffic since the IP routing protocol selects a route that has smaller delays on its end-to-end path.

We therefore developed a new logical topology design algorithm called SHLDA (Shortest-Hop Logical topology Design Algorithm). As described above, we assume that routing is performed only on the IP layer. Thus, the logical topology is designed by incorporating the nature of the route selection process used in the IP routing protocol. It is natural that the shortest path would be selected by the IP routing protocol for forwarding packets. By *short path* we mean

that the number of lightpaths between two nodes is small. Actually, queueing and propagation delays also affect route selection. Therefore, hop counts of lightpaths (i.e., the number of lightpaths the packet traverses) should be reduced as much as possible, and this is the primary objective of our algorithm.

Once a lightpath is allowed to be split between two end-node pairs, a series of lightpaths must be traversed to reach the destination, so the processing delays at the electronic routers must be considered. To incorporate these delays into the final determination of packet routes, we apply flow deviation method [33], which will be described in the next section.

MLDA uses traffic demand between node pairs to set up the next lightpath. In contrast, we use performance metric F_{ij} for node pair ij:

$$F_{ij} = \gamma_{ij} \times h_{ij}, \tag{2.1}$$

where γ_{ij} is the traffic demand from node i to j, and h_{ij} is the hop count of the minimum hop route for node pair ij on the physical topology. The hop count of a lightpath refers to the number of physical links that the lightpath traverses. Note that F_{ij} is equal to γ_{ij} in MLDA, i.e., MLDA does not consider the hop count of the lightpath, and uses only the propagation delay in determining the shortest route for the lightpath. In contrast, SHLDA uses the hop count as a metric in calculating the order of lightpath configuration. The propagation delay and the hop count are then taken into account in determining the route for each lightpath. In determining the route of the lightpath from node i to j, metric R_{ij} is given by the following equation,

$$R_{ij} = D_{ij} \times h_{ij}, \tag{2.2}$$

where D_{ij} is the total propagation delay of the route from node i to j. SHLDA selects the route with the smallest R_{ij} between nodes i and j. This enables a lightpath to be established that cuts through a large number of electronic routers. The SHLDA algorithm works as follows.

Step 1: Calculate metric F_{ij} for each node pair ij from traffic matrix $Q = q_{ij}$. In initially determining F_{ij}, h_{ij} is set simply as the hop count of the shortest physical path.

Step 2: Place a lightpath between two nodes if there is a fiber between them.

Step 3: Select node pair $i'j'$, where i' and j' are indices giving $\max_{ij} F_{ij}$. If $F_{i'j'} = 0$, go to Step 5; otherwise, go to Step 4.

Step 4: Find the shortest route for node pair $i'j'$ and check the availability of wavelengths in order to configure the lightpath. If more than one wavelength is available, use the wavelength with the lowest index to establish the lightpath. Then set $F_{i'j'} = 0$ and go back to Step 3. If there is no available wavelength, set $F_{i'j'} = 0$ and go back to Step 3.

Step 5: If non-used wavelengths remain, configure as many lightpaths as possible randomly using those wavelengths, as in MLDA.

2.3 Applying Flow Deviation Method

2.3.1 Description of Flow Deviation Method

In this subsection, we summarize the flow deviation method [33]. This method incrementally changes the flow assignment along a feasible and descent direction. Given objective function T, the method sets l_{ij} as a partial derivative with respect to λ_{ij}, where λ_{ij} is the flow rate of lightpath(s) between nodes i and j. The new flow assignment is then determined by using the shortest path algorithm in terms of l_{ij}. By incrementally changing from the old flow assignment to the new one, the optimal flow assignment is determined. The method works as follows.

Step 1: Prepare a feasible starting flow assignment, f^0. Let $n = 0$.

Step 2: Set $g \leftarrow f^n$. Assume that flow assignment f^n is represented as $\{x_{11}, \ldots, x_{pq}, \ldots, x_{NN}\}$.

Step 3: Calculate $l_{ij} = \frac{\partial T}{\partial \lambda_{ij}}$ and set new flow assignment $R(g)$ to $\{x'_{11}, \ldots, x'_{pq}, \ldots, x'_{NN}\}$ by solving the shortest path algorithm using metric l_{ij}.

Step 4: For each node pair ij, perform the following steps.

Step 4.1: Let v be the flow assignment by deviating the flow between nodes i and j from g toward $R(g)$. That is, the resulting flow assignment, v, is set to $\{x_{11}, \ldots, x'_{ij}, \ldots, x_{NN}\}$.

Step 4.2: Check whether v is feasible. In our case, feasible v means that the processing capability of IP routers and/or the capacity of a lightpath do not exceed its limits. If the v is not feasible, then the deviation at Step 4.1 is rejected, and go back to Step 4.

Step 4.3: Check whether v is decreasing. If $T(v) < T(g)$, g is allowed to be deviated toward v. Then, $g \leftarrow v$. And go back to Step 4. If $T(g) \leq T(v)$, the deviation from g toward $R(g)$ is rejected, and go back to Step 4.

Step 5: If $g = f^n$, stop iteration. Note that $g = f^n$ means there is no improvement of performance by deviating the flow. Otherwise, set $n \leftarrow n + 1$, and go back to Step 2.

2.3.2 Derivation of Metric l_{ij}

We next determine metric l_{ij} of the flow deviation. The following notations are used.

N: number of nodes in network

P_{ij}: propagation delay of lightpath ij

C: transmission capacity of each wavelength

μ: processing capability of an electronic router. Assumed to be identical among all routers for simplicity.

The following variables are also introduced.

a_{ij}^{sd}: when the packets are routed from node s to node d via the direct lightpath ij, the value is set to be 1. Otherwise, 0.

δ_i: the sum of all traffic switched by the IP electronic router at node i, except the traffic flow originating at node i.

Objective function T is given as the average T_{sd} (delay between node s and d), i.e.,

$$T = \frac{1}{N(N-1)} \sum_{s=1}^{N} \sum_{d=1}^{N} T_{sd} \tag{2.3}$$

As shown in Fig. 2.1, the delay incurred at a node consists of processing delay and transmission delay. Henceforth, the delay between nodes s and d consists of propagation delay, processing delay, and transmission delay. It thus follows that

$$T_{sd} = \Big[\sum_{ij} a_{ij}^{sd} P_{ij}\Big] + \Big[\sum_{ij} a_{ij}^{sd} Q_{ij}\Big] + \Big[\sum_{ij} (a_{ij}^{sd} R_i) + R_d\Big],$$

where Q_{ij} is the transmission delay of the packets on lightpath ij, and R_i is the processing delay in the electronic router for node i. In this chapter, Q_{ij} is determined by a $M/M/k_{ij}$ (where k_{ij} shows the number of lightpaths between node pair ij) queueing system, and R_i by a $M/M/1$ queueing system. A multiple number of lightpaths between the node pair is allowed , and those lightpaths share the same buffer (see Section 2). Q_{ij} and R_i are then determined as follows.

$$Q_{ij} = \frac{X_l}{l \cdot C - \lambda_{ij}} + \frac{1}{C} \tag{2.4}$$

$$R_i = \frac{1}{\mu - (\lambda_{ij} + \delta_i)} \tag{2.5}$$

where

$$X_l = \frac{p_0 \, (l\rho)^l}{(1-\rho)\, l!}$$

$$\rho = \frac{\lambda_{ij}}{k_{ij} \cdot C}$$

$$p_0 = \left\{ \sum_{x=0}^{k_{ij}-1} \frac{(k_{ij}\rho)^x}{x!} + \frac{(k_{ij}\rho)^{k_{ij}}}{k_{ij}!(k_{ij}-\rho)} \right\}^{-1}.$$

Three kinds of packets arrive at the electronic router of node i: packets destined for node i, packets arriving at node i from local access, and packets changing the lightpath at node i. Thus,

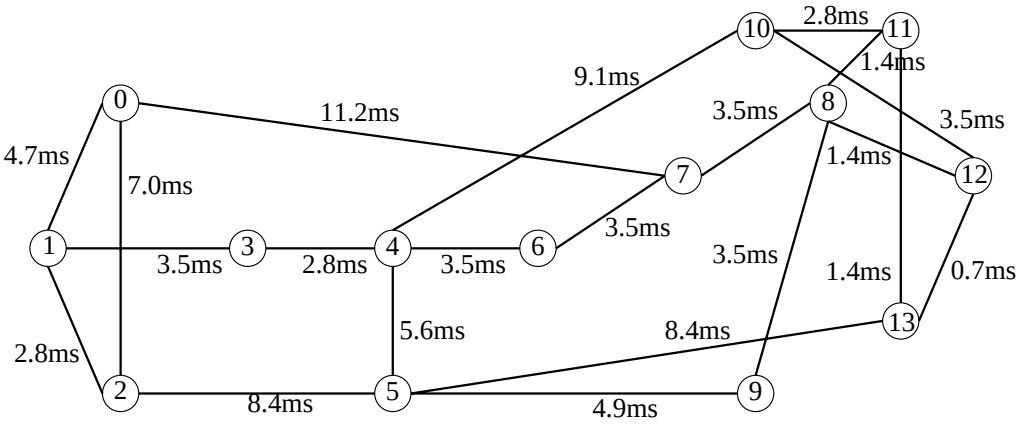


Figure 2.2: NSFNET

δ_i is given by the following equation.

$$\delta_i = \left[\sum_j \gamma_{ji} + \sum_j \gamma_{ij} + \sum_j a_{ij}^{sd} \gamma_{sd} \right] - \lambda_{ij}.$$

Note that λ_{ij} is the flow rate of lightpath(s) between nodes i and j. That is,

$$\lambda_{ij} = \sum_{sd} a_{ij}^{sd} \gamma_{sd}.$$

Eqs. (2.4) and (2.5) give l_{ij} as

$$l_{ij} = \frac{\partial T}{\partial \lambda_{ij}} = \frac{1}{N(N-1)} \sum_{s=1}^{N} \sum_{d=1}^{N} a_{ij}^{sd} \dot{x}_{sd},$$

where

$$x_{sd} = \frac{X_{k_{ij}}}{(l \cdot C - \lambda_{ij})^2} + \frac{1}{(\mu - (\lambda_{ij} + \delta_i))^2}.$$

2.4 Numerical Evaluation and Discussion

2.4.1 Network Model

As a network model, a 14–node NSFNET is considered (Fig. 2.2). A traffic matrix given in [6] is used in the numerical evaluation. Since the traffic matrix is given by a relative value, we introduce *traffic scale factor* α, and actual traffic demands between nodes are given by the

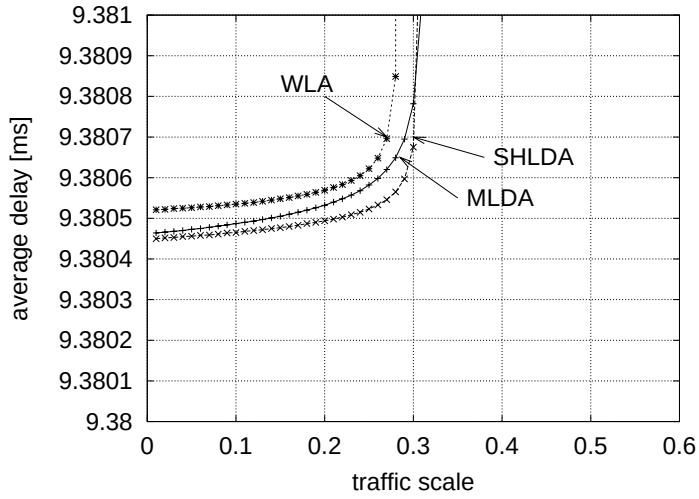


Figure 2.3: Average delay of end-to-end paths : W = 8, μ = 40 Mpps

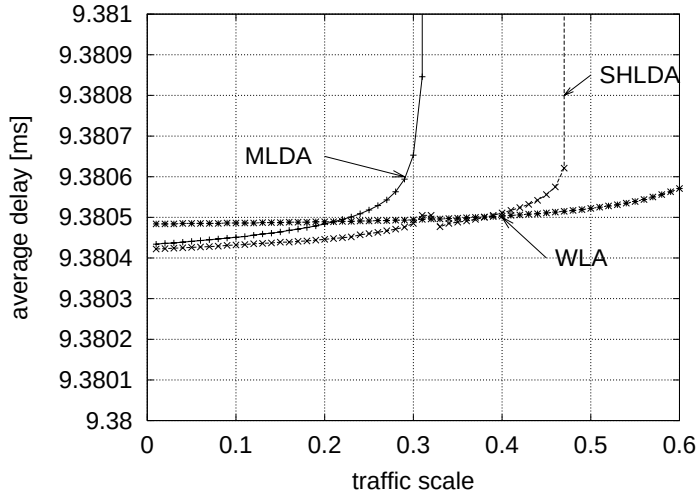


Figure 2.4: Average delay of end-to-end paths : W = 8, μ = 100 Mpps

traffic matrix multiplied by α. It is also assumed that the value of the given traffic matrix is represented in gigabits per second. And the transmission capacity of each wavelength is set to 10 Gbps. The packet processing capability of the electronic router, μ, is represented in pps (packet per second) under the assumption that the mean packet size is 1,000 bits long.

2.4.2 Numerical Results and Discussions

We evaluate our SHLDA by comparing with MLDA. In addition to MLDA, we also consider WLA (WDM Link Approach), where a WDM technology is only utilized for point-to-point links between adjacent IP routers. Figure 2.3 compares the average delays obtained by the three

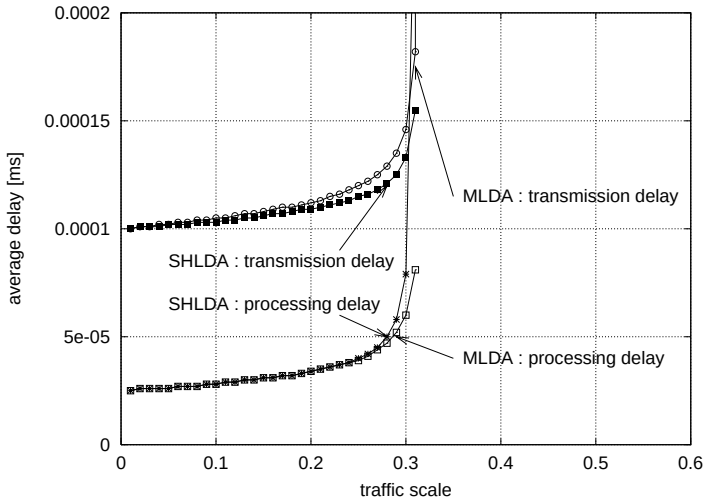


Figure 2.5: Average of packet processing and transmission delays at electronic routers : W = 8, μ = 40 Mpps

algorithms, SHLDA, MLDA and WLA. The horizontal axis shows the traffic scale factor α. The number of W is set to eight and the packet processing capacity of the IP router, μ, is set to 40 Mpps. In the figure, when α is small, no significant difference between the three algorithms can be seen. In the case of all three algorithms, the delays suddenly increase as α becomes large. It is notable that our SHLDA has the same performance as MLDA in terms of the maximum throughput, i.e., the saturation point of the delays.

Figure 2.4 shows the effect of increasing the packet forwarding capability of IP routers by changing μ from 40 Mpps to 100 Mpps. The other parameters are the same as those in Fig. 2.3. Comparing these two figures shows that the maximum throughput values by SHLDA is increased. On the other hand, an increased in the maximum throughput cannot be seen when we apply MLDA. To explain this result, the nodal delays were studied in more detail. Figs. 2.5 and 2.6 show the dependency of processing and transmission delays on α. As expected, the processing delay at the electronic router decreases both SHLDA and MLDA when the capability of the IP routers changes from 40 Mpps to 100 Mpps. Since the processing delay is reduced as the capacity of the IP router increases, the transmission delay becomes the bottleneck of the network. In that case, SHLDA becomes superior to MLDA.

Next, we set the number of wavelengths W to twelve and μ to 40 Mpps. The average delay is plotted in Fig. 2.7. By comparing Figs. 2.3 and 2.7, it is apparent that SHLDA exhibits the largest increase in maximum throughput. To see this more clearly, Fig. 2.8 presents components

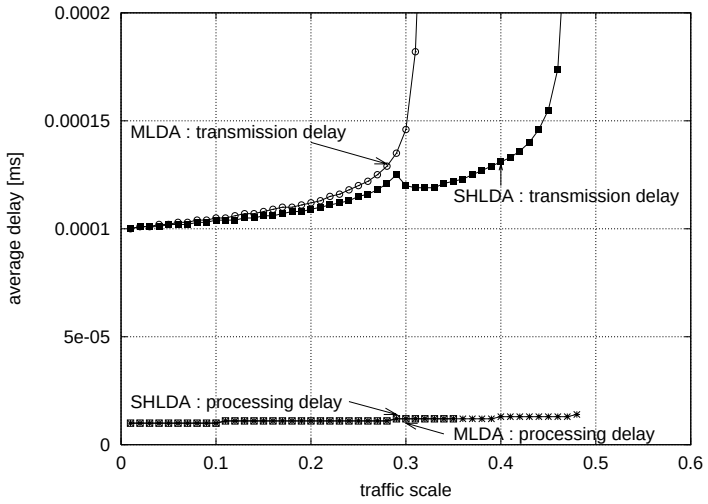


Figure 2.6: Average of packet processing and transmission delays at electronic routers: W = 8, μ = 100 Mpps

of delays.

According to Figs. 2.5 and 2.8, the transmission delay by MLDA is decreased more than that by SHLDA. Its reason can be explained as follows. SHLDA places lightpaths in a descending order of the product of the hop–count and traffic demand. As a result, a lightpath placed by SHLDA tends to utilize more links than the one by MLDA. Thus, MLDA can find more lightpaths than SHLDA as the number of available wavelengths increases. This leads to decreasing the transmission delay in the case of MLDA. Comparing the processing delay in Figs. 2.5 and 2.8 shows that, when the traffic scale factor is from 0.27 to 0.37, the processing delay at the IP router is decreased as the number of available wavelengths increases. Its effect is larger in the case of SHLDA. As mentioned before, the lightpaths placed by SHLDA tend to utilize more physical links. This results in more reduction of electric processing in SHLDA than that in MLDA.

The average delay determined by SHLDA by increasing the number of wavelengths is explained in the following. Figure 2.9, where W is 20 and μ is 40 Mpps, plots average delay against traffic scale factor. In this figure, SHLDA still attains a higher throughput than MLDA, but the difference is comparatively smaller than that in Fig. 2.7. The reason for this is that by increasing the number of wavelengths, the logical topologies obtained by SHLDA or MLDA become close to a fully meshed network. The advantage of SHLDA thus becomes small since it tries to reduce the traffic load on the IP router. We also show the case that μ is 100 Mpps. The

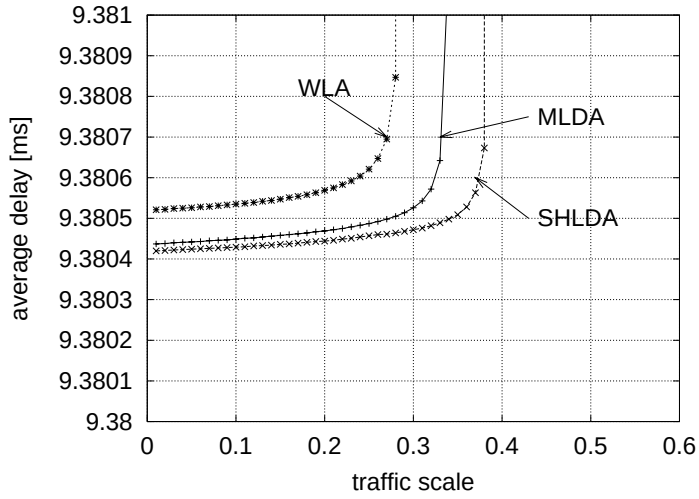


Figure 2.7: Average delay of end-to-end paths: W = 12, μ = 40 Mpps

result is plotted in Fig. 2.10 where we set $W = 12$. Comparing Figs. 2.7 and 2.10 also shows the effectiveness of SHLDA.

Lastly, we summarize the characteristics of WLA by observing Figs. 2.3, 2.4, 2.7 and 2.10. Figs. 2.3 and 2.7 indicate that the increase in the maximum throughput by WLA is very limited. This is because the processing delay at the electronic router is the primary bottleneck of the network; thus, the effect of increasing the number of wavelengths cannot be observed. As one can easily imagine, the results regarding for WLA are greatly improved as the capability of the IP router becomes large (compare Figs. 2.3, 2.4 and 2.10). Only in such a large capability, WLA is not a bad choice for IP over WDM networks.

2.4.3 Investigation on Routing Stability

We finally discuss the new logical topology design algorithm from the viewpoint of the stability of IP routing. In IP networks, it is necessary to avoid or at least to reduce unnecessary changes of the routes, which are caused by dynamically changing traffic demand. To evaluate routing stability, we show the packet delays of the first and second shortest end-to-end paths (lightpaths) determined by SHLDA . If these two values are close, the route of the IP packets may frequently change with traffic fluctuation.

Metric d_{sd}, which defines the difference of delays of the first and second shortest routes between node pair sd, is introduced here. From all possible combinations of source and destination

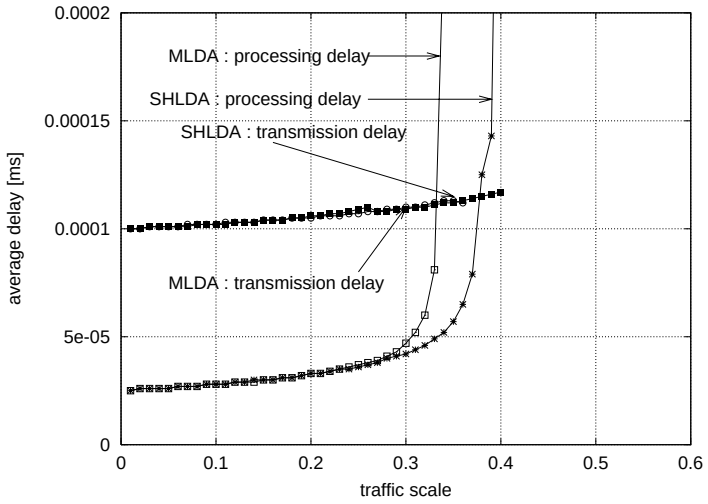


Figure 2.8: Average of packet processing and transmission delays at electronic routers: W = 12, μ = 40 Mpps

node pairs, the smallest one was chosen as d_{min}, i.e., $d_{min} = \min_{sd}\{d_{sd}\}$. We consider here that the design algorithm that provide the larger d_{min} gives a higher routing stability. Figs. 2.11 and 2.12 plot d_{min} obtained from SHLDA and MLDA as a function of α, where the number of wavelengths W is set to eight and twelve, respectively. The processing capacity of the IP router, μ, is identically set to 40 Mpps in both figures. The average value of d_{min} is also shown in the figures. It is clear that when W is 8 (Fig. 2.11), SHLDA is not very good especially when the traffic scale factor is large. However, it gives higher stability than MLDA when the number of wavelength is twelve (Fig. 2.12).

The problem with both of MLDA and SHLDA is that at several values of α, d_{min} takes very small values. This is mainly because SHLDA as well as MLDA is a "one–way algorithm". That is, there are no step-back operation in the algorithms; in other words, if the nodal delay is high, it is likely that the delay of the first shortest route becomes close to the delay of the second shortest one, since the nodal delay becomes dominat in such a region. We believe the situation can be avoided by reassembling the lightpaths to reduce the nodal delay, but this issue is one of our future research topics.

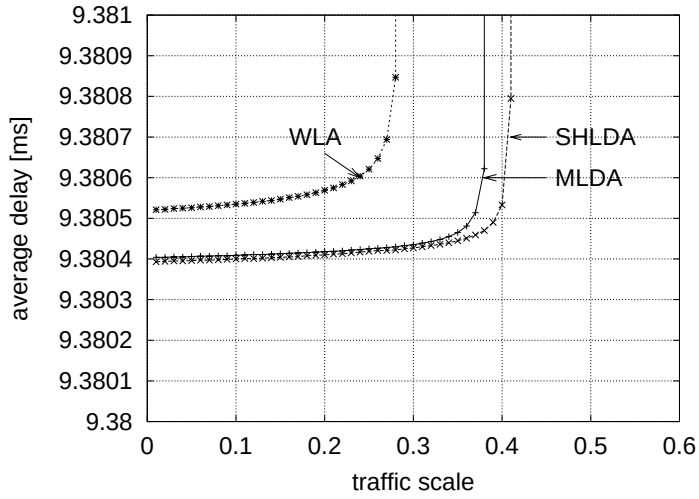


Figure 2.9: Average delay of end-to-end paths: W = 20, μ = 40 Mpps

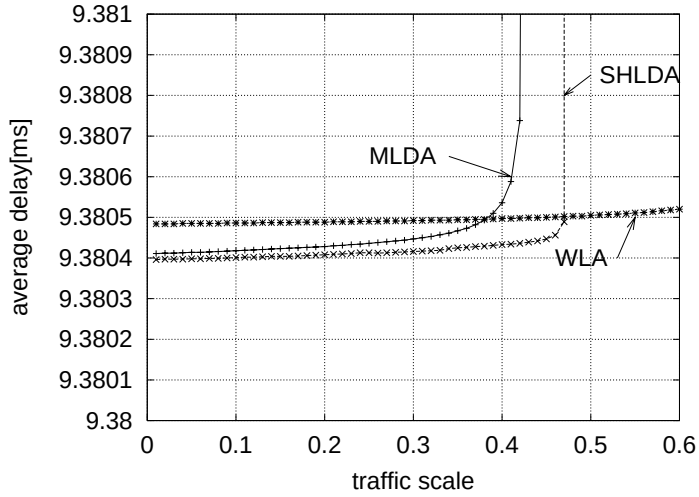


Figure 2.10: Average delay of end-to-end paths: W = 12, μ = 100 Mpps

2.5 Conclusion

We have proposed a new heuristic algorithm, SHLDA, for designing a logical topology by considering the delay between nodes as an objective metric. The proposed algorithm was compared with conventional methods in terms of the average packet delay and throughput. The results show that SHLDA becomes effective when the number of wavelengths is low and the processing capacity of a IP router is large. Furthermore, SHLDA was evaluated from a viewpoint of routing stability. It was found that SHLDA improves the maximum throughput, compared with a conventional algorithm, without sacrificing routing stability.

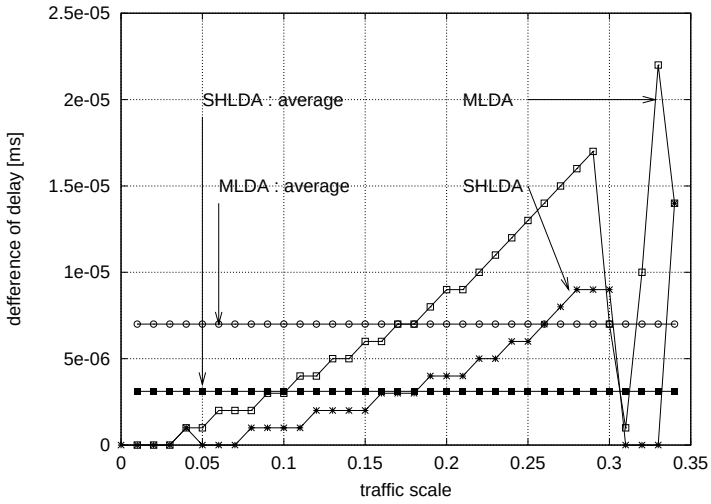


Figure 2.11: Route stability: Delay difference between the first and sencond shortest path (W = 8, μ = 40 Mpps)

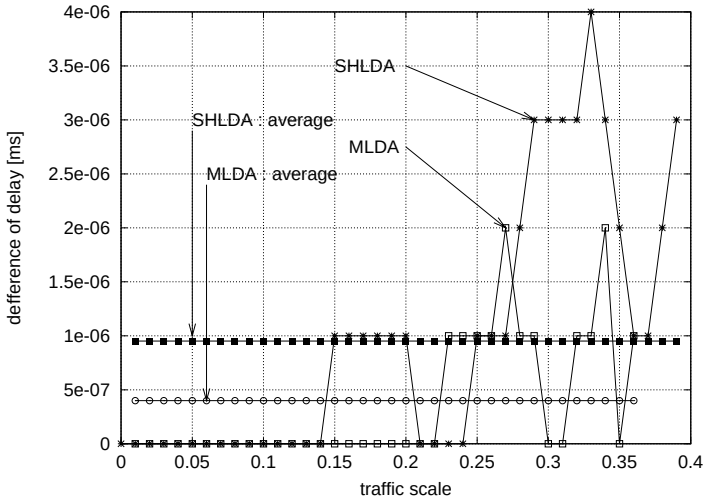


Figure 2.12: Route stability: Delay difference between the first and sencond shortest path (W = 12, μ = 40 Mpps)

Chapter 3

Functional Partitioning for Multi-layer Surivability in IP over WDM Networks

In this chapter, we discuss the multi–layer survivability in IP over WDM networks. In IP over WDM networks, IP routing has its own routing mechanism, so it may not be necessary to protect all the lightpaths by using the optical layer if doing so does not lead to cost savings even when a shared protection scheme is used. If we allow that several primary lightpaths cannot recover from some failure patterns and that the resilience is left to the IP layer, we can expect more cost savings. Multi–layer survivability, in which lightpaths in a subset are protected against failure, is investigated. Assuming the single–component failure in the network, we formulate the shared path protection mechanism as an optimization problem. It is formulated as MILP (Mixed Integer Linear Problem), and becomes computationally intensive as the network grows. Accordingly, we propose a heuristic algorithm and compare its results with the solution obtained by MILP. Through numerical examples, we compare the number of wavelengths required to make network reliable. We next consider the functional partitioning of IP routing and WDM protection. Based on using our algorithm, we also discuss the effect of interaction between the IP and WDM layers. We show that the largest-traffic-first approach is best if our primary concern is traffic load at the IP routers after a failure.

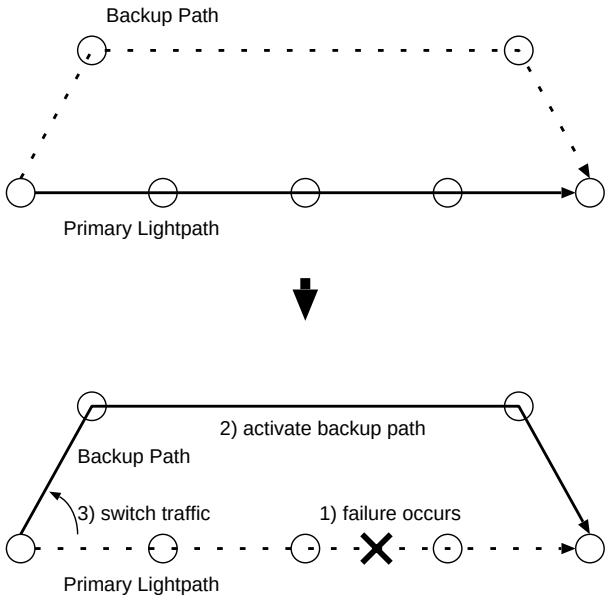


Figure 3.1: Path protection

3.1 Fault Tolerance Methods in WDM Networks

3.1.1 Protection Method

The protection method [9, 34] is a fast recovery method realized through mechanical switching in the optical domain. For each primary lightpath, backup lightpaths are determined and statically configured beforehand, and wavelengths for the backup lightpaths are reserved. There are two protection methods: path protection and link protection. In path protection, a backup lightpath is prepared between the source and destination nodes (Fig. 3.1). In contrast, in link protection a backup lightpath is prepared for each link of the primary lightpath (Fig. 3.2). In either case, when a network component fails along the primary lightpath, the corresponding backup lightpath is activated and traffic on the primary lightpath is switched to the backup lightpath. The protection method thus guarantees 100% reliability for primary lightpaths under the single–failure assumption. That is, whatever failure occurs, the lightpath can be restored and the lightpath bandwidth is not reduced due to a failure. However, since both protection methods reserve wavelengths for the backup lightpaths, the efficiency of wavelength usage is lower. There is a trade–off between fast recovery and efficient use of wavelength resources.

Accordingly, several methods aimed at using wavelengths more efficiently have been proposed [9, 23, 24, 26, 35, 36]. One promising method is *shared protection*, in which two or more

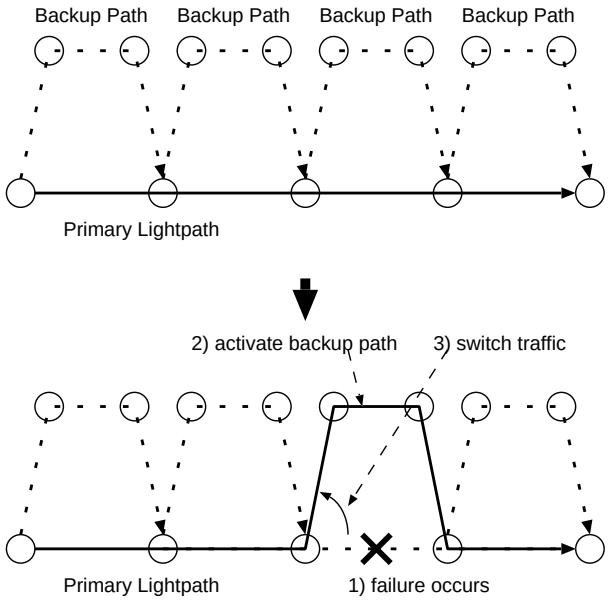


Figure 3.2: Link protection

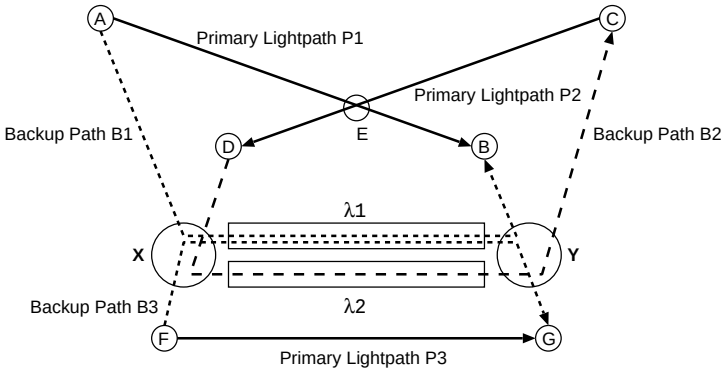


Figure 3.3: Shared protection method

primary lightpaths can share the same backup lightpath as long as the primary lightpaths are disjoint [9]. Figure 3.3 illustrates the idea of shared protection. Three primary lightpaths are denoted as $P1$, $P2$ and $P3$. $P1$ is placed between nodes A and B, and $P2$ and $P3$ connect node pairs CD and FG, respectively. Backup lightpaths $B1, B2$, and $B3$ protect primary lightpaths $P1$, $P2$, and $P3$, respectively. Primary lightpaths $P1$ and $P2$ both traverse intermediate node E. Furthermore, backup lightpaths $B1$, $B2$, and $B3$ are configured to use the link connecting node pair XY. Here, $B1$ and $B3$ share the same wavelength $\lambda 1$, whereas $B2$ uses $\lambda 2$. Note that $B1$ and $B2$ must use different wavelengths on the link since the corresponding primary lightpaths ($P1$ and $P2$) both use node E. If we assume that two or more components may fail at the same time, though, we cannot use the shared protection method. This is because the shared

protection method assumes that backup lightpaths whose primary lightpaths are disjoint will never be activated at the same time, hence the shared wavelength on the link will never create a conflict between the sharing backup lightpaths.

3.1.2 Restoration Method

Restoration is an alternative way to recover from failures in the optical layer. In a restoration method, a backup lightpath is dynamically determined when a failure occurs. Once a backup lightpath is found, the traffic on the primary lightpath affected by the failure is switched to the backup lightpath. Unlike the protection methods, a restoration method does not reserve any wavelength resources for the backup lightpaths before a failure. Therefore, the wavelengths are used more efficiently than with the protection methods. However, a restoration method cannot set up a backup lightpath if wavelength resources are not available. This means that a restoration method cannot provide a 100% guarantee of failure recovery. Moreover, since the backup lightpath is determined only after a failure occurs, a restoration method needs more time to restore a lightpath.

3.2 Single–Layer Case

Protection schemes for WDM networks have been widely studied [9,10,18,23,35,37–42]. Here, we consider the shared path protection scheme, which improves reliability against fiber failure, which is typically caused by the cutting of a fiber. The shared path protection mechanism is suitable for improving wavelength utilization if the WDM network is highly reliable and multiple failures seldom occur. Our objective is to minimize the number of wavelengths used on a link. The formulation in this subsection is based on that of Ramamurthy and Mukherjee [9].

3.2.1 *Problem Formulation*

We will use the following notation.

i, j: originating and terminating nodes of a logical link. The logical link between nodes i and j is lightpath ij.

m, n: end nodes of a physical link. The physical link connecting nodes m and n is physical link mn.

The following notations are used for characterizing the physical WDM network.

N: number of nodes in physical (and logical) network

W: number of wavelengths carried in a fiber

P_{mn}: physical topology defined by set $\{P_{mn}\}$. If a fiber connects nodes m and n, then $P_{mn} = 1$, otherwise $P_{mn} = 0$.

The following notation is used for representing the logical network.

V_{ij}: number of lightpaths between nodes i and j

R_{ij}^k: route of lightpath from node i to node j using wavelength k. It consists of a set of physical links: (i, m_1), (m_1, m_2), ..., (m_p, j).

A_{ij}^k: route of backup lightpath for primary lightpath from node i to node j using wavelength k. It consists of a set of physical links: (i, n_1), (n_1, n_2), ..., (n_q, j).

c_{ij}^k: If the primary lightpath uses wavelength k between originating node i and terminating node j, $c_{ij}^k = 1$, otherwise $c_{ij}^k = 0$. c_{ij}^k is determined from R_{ij}^k.

o_{mn}^k: If the primary lightpath uses wavelength k on physical link mn, $o_{mn}^k = 1$, otherwise $o_{mn}^k = 0$. o_{mn}^k can be determined from R_{ij}^k.

φ_{mn}: maximum number of backup lightpaths passing through physical link mn. It can be determined from A_{ij}^k.

The following variables are used to formulate the optimization problem.

w_{mn}: number of primary lightpaths on physical link between two directly connected nodes, m and n.

b_{mn}: number of backup lightpaths on physical link mn.

m_{mn}^w: If the backup lightpath uses wavelength w on physical link mn, $m_{mn}^w = 1$, otherwise $m_{mn}^w = 0$.

$g_{ij,pq,k}^{mn,w}$: If a lightpath originating at node i and terminating at node j uses wavelength k for the primary lightpath on physical link pq and also uses wavelength w between nodes m and n as a backup lightpath, $g_{ij,pq,k}^{mn,w} = 1$, otherwise $g_{ij,pq,k}^{mn,w} = 0$.

Using these notations, we next formulate the wavelength assignment problem for backup lightpaths as an optimization problem.

Objective function

Minimize number of wavelengths used:

$$\min \sum_{m,n} (w_{mn} + b_{mn}).$$

Constraints

(1) The number of primary lightpaths placed on physical link mn must equal the total number of primary lightpaths using wavelength w on that physical link:

$$w_{mn} = \sum_{w \in W} o_{mn}^{w}. \tag{3.1}$$

(2) Similarly, the number of backup lightpaths placed on physical link mn must equal the total number of wavelengths used on that link for the backup lightpaths:

$$b_{mn} = \sum_{w \in W} m_{mn}^{w}. \tag{3.2}$$

(3) Either one primary lightpath or one backup lightpath must use wavelength k on physical link mn if there is a fiber:

$$o_{mn}^{k} + m_{mn}^{k} \leq P_{mn}. \tag{3.3}$$

(4) The lightpath using wavelength k between node i and node j must be protected by a backup lightpath when physical link $pq \in R_{ij}^{k}$ fails:

$$c_{ij}^{k} = \sum_{w \in W} \sum_{it \in A_{ij}^{k}} g_{ij,pq,k}^{it,w}. \tag{3.4}$$

Note that it is unnecessary to use different wavelengths between the primary lightpath and the corresponding backup lightpath.

(5) The lightpath using wavelength k between node i and node j must use wavelength w on all links of the backup lightpath (i.e., the wavelength–continuity constraint should hold):

$$g_{ij,pq,k}^{nt,w} = g_{ij,pq,k}^{tm,w} \quad \forall pq \in R_{ij}^k, \forall nt, tm \in A_{ij}^k. \tag{3.5}$$

(6) For each fiber–failure scenario, a lightpath using wavelength k between node i and node j must use the same wavelength w on physical link $mn \in A_{ij}^k$ for the backup lightpath:

$$g_{ij,p_1q_1,k}^{mn,w} = g_{ij,p_2q_2,k}^{mn,w} \quad \forall p_1q_1, p_2q_2 \in R_{ij}^k. \tag{3.6}$$

As this equation indicates, we assume that we allow that different wavelengths can be used for the backup lightpath and corresponding primary path.

(7) When a failure occurs on physical link pq, at most one backup lightpath should use wavelength w on physical link mn if the corresponding primary lightpath traverses failure link pq:

$$\sum_{ij} \sum_{k \in W: c_{ij}^k > 0 \wedge pq \in R_{ij}^k \wedge mn \in A_{ij}^k} g_{ij,pq,k}^{mn,w} \leq 1. \tag{3.7}$$

(8) The number of backup lightpaths using wavelength k on physical link mn must be bounded:

$$\varphi_{mn}\, m_{mn}^k \geq \sum_{w \in W} \sum_{(i,j):(c_{ij}^k > 0, mn \in A_{ij}^k)} \sum_{pq \in R_{ij}^k} g_{ij,pq,w}^{mn,k}. \tag{3.8}$$

We do not distinguish two primary lightpaths having link disjoint routes in our formulation. In IP over WDM networks, paths having different routes are viewed by the IP layer as having different delays. Hence, IP selects the path providing the shortest delay, so it is not worthwhile to consider link disjoint routes. This is why we do not explicitly distinguish two primary lightpaths.

3.2.2 *Heuristic Approaches*

Formulation of the wavelength assignment problem for backup lightpaths using the shared path protection mechanism, as described above, results in a mixed integer linear problem (MILP), and a standard mathematical programming optimizer such as CPLEX [43] can be used to solve it. However, an MILP can be solved only when there is a small number of variables. In our case, the number of variables increases exponentially with the number of nodes and/or the number of wavelengths. We therefore need a heuristic approach applicable to large–scale networks.

Our basic idea is as follows. For shared path protection, several primary lightpaths are allowed to share a single wavelength as the backup lightpath. However, sharing of a backup lightpath is possible only when the corresponding primary lightpaths are fiber–disjoint. If the hop count of a primary lightpath is small, the possibility of conflicts with another lightpath is small. Here, the hop count of the lightpath refers to the number of physical links that the lightpath traverses. To enable more sharing while avoiding conflicts among lightpaths with large hop counts, we assign the backup lightpaths in ascending order based on the number of hop counts, which we call the *min–hop–first* approach. Assigning the wavelengths sequentially, starting with the smallest hop count lightpath, should reduce the number of wavelengths not assigned. After the lightpaths with the shorter hop counts are assigned as backup lightpaths, the lightpaths with larger hop counts can use wavelengths not yet assigned, since many wavelengths generally remain unused for those paths.

The following notation is used for explaining our min–hop–first approach.

h_{ij}^k: hop count of primary lightpath that uses wavelength k for node pair i and j.

A_{ij}^k: set of physical links used for backup lightpath for primary lightpath ij using wavelength k.

B_{ij}^k: set of links as yet unchecked as to whether a lightpath can be placed between nodes i and j using wavelength k. Initially, B_{ij}^k is set to A_{ij}^k.

Using this notation, we next describe our min–hop–first approach.

Step 1: Identify lightpath with smallest h_{ij}^k.

Step 2: For each wavelength p ($p = 1, 2, \cdots, W$), check whether the backup lightpath uses wavelength p between originating node i and terminating node j. More precisely, for each physical link connecting nodes m and n (i.e., link $mn \in B_{ij}^p$), do the following.

Step 2.1: If wavelength p on physical link mn is not used by another lightpath, delete link mn from B_{ij}^{p} and go to Step 3. If wavelength p is used by another lightpath, go to Step 2.2.

Step 2.2: If wavelength p on physical link mn is used by another primary lightpath, the backup lightpath cannot be set up using wavelength p. Return to Step 2 and examine the next wavelength. If wavelength p is used by a backup lightpath, check whether the two backup lightpaths can share the wavelength. They can share it if the corresponding primary lightpaths are fiber–disjoint, which means that they have no common links. If they can share the wavelength, delete link mn from B_{ij}^{p} and go to Step 3. Otherwise, the backup lightpath cannot be set up using wavelength p. Return to Step 2, and examine the next wavelength.

Step 3 If $B_{ij}^{p} = \phi$, assign wavelength p to link $mn \in A_{ij}^{p}$ and go back to Step 1. Otherwise, go back to Step 2.1 and examine the next link.

We also considered the *largest–traffic–first* approach, in which the lightpaths are selected in descending order based on the traffic load on the lightpaths. In the following subsections, we consider the *random* approach, in which the lightpath is selected randomly, for comparison purposes.

3.2.3 Numerical Examples

We first investigated the usefulness of IP over WDM networks with high reliability. CPLEX 6.5 was used to solve the optimization problem. Since it is hard to solve the problem for a large–scale network, we use a eight–node network diagrammed in Fig. 3.4.

We used our heuristic algorithms to examine its optimality, for which we needed its logical topology. For this purpose, we used the MLDA algorithm, a heuristic algorithm proposed by Ramaswami and Sivarajan [6]. The MLDA algorithm works as follows. First, it sets up a lightpath between nodes if there is a fiber between them. Then, it attempts to set up lightpaths between nodes in the descending order of traffic rates. Finally, if some wavelengths are still unused, as many lightpaths as possible are set up using those wavelengths. The direct applica-

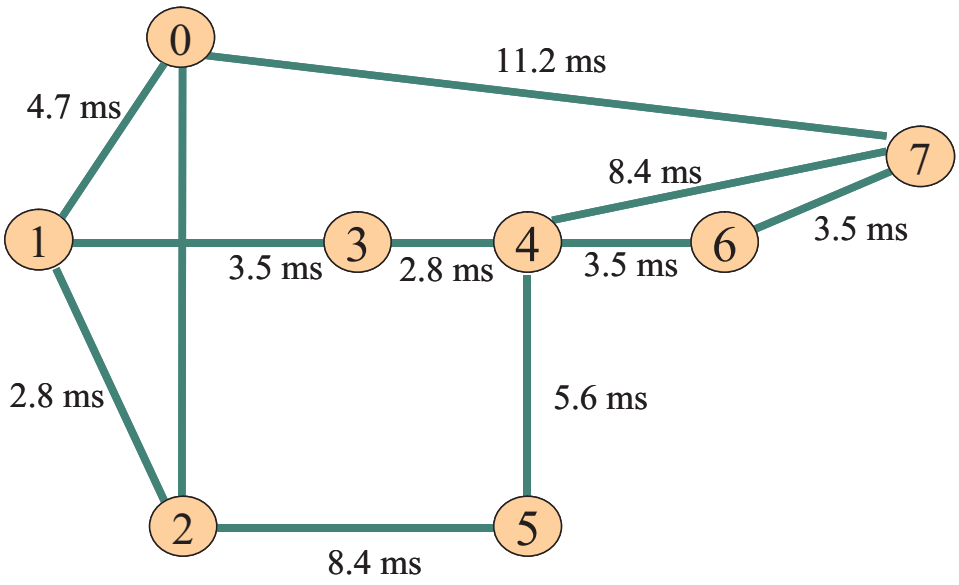


Figure 3.4: Physical topology of eight–node network

Table 3.1: Number of wavelengths required to protect all lightpaths

MILP	min–hop–first	largest–traffic–first
10	10	11

tion of the MLDA algorithm is not appropriate because it does not consider protection. We thus modified the algorithm as follows.

(1) While the MLDA algorithm sets up a lightpath even if the lightpath has already been set up, we do not set up multiple lightpaths between two nodes so that more wavelengths are left for possible use as backup lightpaths.

(2) While the MLDA algorithm sets up lightpaths randomly if any wavelengths remain unused, we do not assign them for the same reason as above.

The min–hop–first and random approaches do not require a traffic matrix since it is not used in either algorithm, while the largest–traffic–first approach does need one. We used the traffic matrix given in [6] for the reference purposes. We set the number of wavelengths used for primary lightpaths, that is, the wavelengths used by the MLDA algorithm, to five. The results of the optimization problem and our heuristic algorithm are compared in Table 3.1, which shows the number of wavelengths required to protect all lightpaths. Good results were obtained with both algorithms.

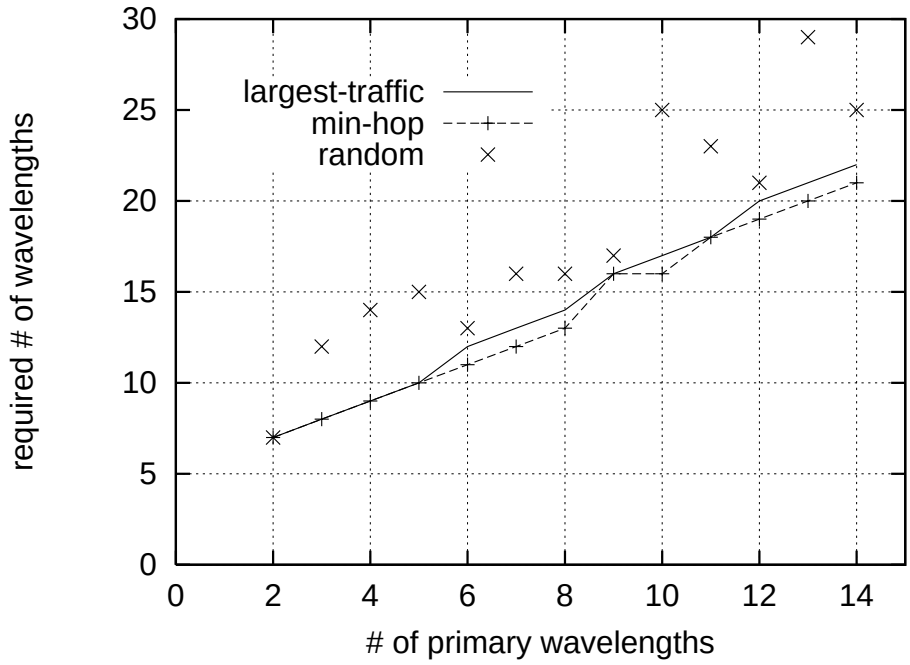


Figure 3.5: Number of wavelengths required to completely protect primary lightpaths

3.2.4 *Results with Heuristic Approach*

We next considered a 14–node NSFNET backbone network (Fig. 2.2) as the network model. The same traffic matrix [6] was used for reference purposes. Since the MLDA algorithm sets up lightpaths on the physical topology, we must identify the route of the IP packets. We modified Dijkstra's shortest path algorithm to consider the nodal processing delays. We assume that the delays are derived from a M/M/1 queueing model and that the offered traffic rates are assumed to be $\sum_s \lambda^{sd}$.

Figure 3.5 compares the three approaches in terms of the number of wavelengths required to protect all lightpaths. The horizontal axis shows the number of wavelengths used for the primary lightpaths. For example, if the primary lightpaths are established using ten wavelengths to establish the logical topology, an additional six wavelengths are needed to protect all lightpaths with the min–hop–first approach. The min–hop–first approach required the smallest number of wavelengths among the three approaches.

3.3 Multi Layer Survivability

Ideally, a WDM network would protect all lightpaths so that traffic on a primary lightpath could be switched to the backup lightpath within about ten milliseconds. However, we need to consider the tradeoff between the processing capability of the IP routers and the limitation on the number of wavelengths. Setting up more backup lightpaths protects more primary lightpaths, but because the number of wavelengths is limited, the number of primary lightpaths should be limited to increase the number of backup lightpaths. Reducing the number of primary lightpaths, however, increases the load on the IP routers, and bottlenecks at IP routers cannot be resolved. In contrast, increasing the number of wavelengths used for the primary lightpaths would enable more traffic to be carried by the primary lightpaths. However, in that case, the advantage of the protection mechanism of a WDM network cannot be used.

There is another problem. While the WDM protection mechanism can switch to the backup lightpath in the order of ten milliseconds, an IP router may change the route to a better one after the routing table is updated. Suppose that after a failure occurs, lightpath ij using wavelength k is switched to the backup lightpath. This naturally increases the propagation delay. After the router updates its table (typically in the order of ten seconds), it may find a route (which may consist of two or more concatenated lightpaths) shorter than the backup lightpath allocated by the WDM protection mechanism.

The main cause of this problem is that we did not consider the possibility of a route change in the design of the WDM protection mechanism described in section 1.4. To enable the wavelengths to be used more effectively, we changed our heuristic algorithm so that backup lightpaths that are not likely to be used by IP are not allocated. The changes to the min–hop–first approach are as follows.

(1) In Step. 1, after selecting lightpath h_{ij}^k, define set $\{S\}$; identify its elements, which are the node pairs, using h_{ij}^k.

(2) Calculate increased delay θ under the assumption that the backup lightpath is allocated.

(3) For every node pair sd in $\{S\}$, calculate the delay of primary lightpath d_{sd} and that of the second–shortest path, d_{sd}^a. Then, check whether the sum of d_{sd} and θ exceeds the delay of d_{sd}^a. If it does, check the next lightpath, $h_{i'j'}^{k'}$, without protecting the current lightpath, h_{ij}^k.

Determining how many wavelengths should be allocated for primary and backup lightpaths is difficult because it depends on the network capacity that must be provided by the primary lightpaths and on the network survivability that must be provided by the protection mechanism of the WDM network. We therefore used numerical examples to identify the best balance between these objectives.

3.3.1 Numerical Examples and Discussion

We investigated the effect of IP/WDM interactions using the NSFNET backbone network model (see Fig. 2.2).

As shown in Fig. 3.6, the number of protected lightpaths depends on the number of wavelengths available in the fiber. To obtain this relationship, we use the MLDA algorithm [6] to determine the logical topology. The number of wavelengths used for the primary lightpaths was fixed at eight, and the number of wavelengths for the backup lightpaths was increased from 0 to 22. Using the modified MLDA algorithm, we established 73 primary lightpaths. With seven backup wavelengths, these 73 lightpaths are completely protected with all three approaches (min–hop–first, largest–traffic–first, and random approaches)

Note that even without any backup wavelengths, the number of protected lightpaths is not 0 but 10. This is because, in the modified MLDA algorithm, wavelengths not allocated remain available to be used later for protection. Between 11 and 13 backup wavelengths, the min–hop–first approach protected more lightpaths than either the largest–traffic–first or random approach.

We next fixed the total number of wavelengths and changed the number of wavelengths used for establishing primary lightpaths. Figure 3.7 shows the results for 16 wavelengths. The horizontal axis shows the number of wavelengths used for backup lightpaths, and the vertical axis does the numbers of the lightpaths protected by WDM protection mechanisms. With all three approaches, the number of protected lightpaths first increased with the number of backup wavelengths, then decreased. This is because when the number of wavelengths reserved for backup is small, more lightpaths can be protected by increasing the number of wavelengths used for backup. However, as the number of wavelengths dedicated to backup increases, the number of primary lightpaths that can be generated decreases, and the number of wavelengths unused

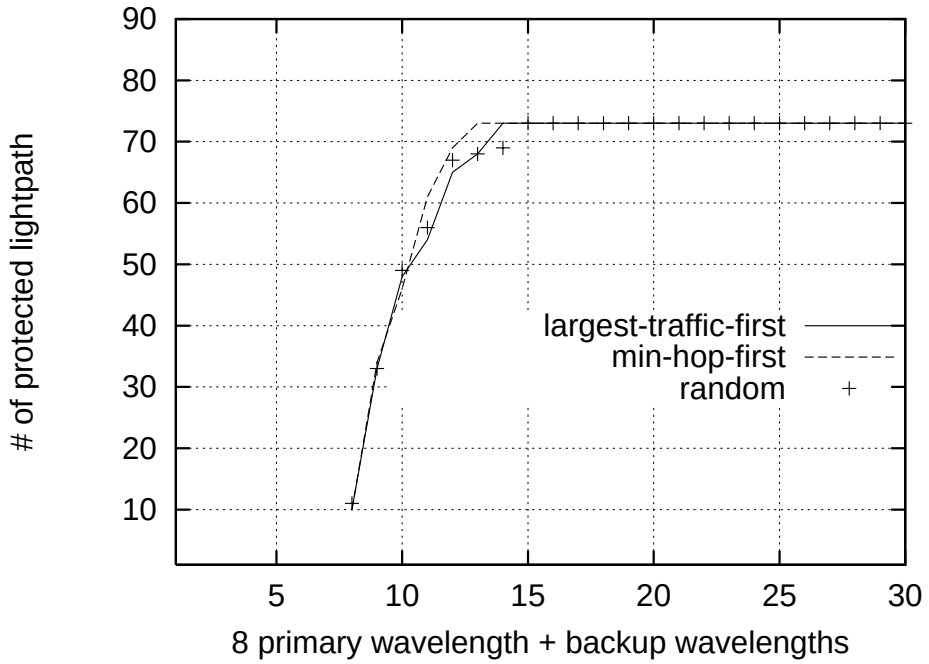


Figure 3.6: Number of protected lightpaths

increases. The min–hop–first approach protected the most lightpaths for any given number of backup wavelengths.

The increase in traffic volume at an IP router when a failure occurs is another important measure of the efficiency of the protection mechanism of an WDM network. To evaluate it, we again fixed the number of wavelengths at 16 and changed the number of wavelengths used for the primary lightpaths. For each number of wavelengths for primary lightpaths, we measured the increased loads at the routers after a single–fiber failure. By examining all cases of single–fiber failure, we identified the maximum load at each router. The increased traffic rates at each router, when 10, 12, and 14 wavelengths were used for the primary lightpath, are shown in Fig. 3.8, 3.9, and 3.10, respectively. The increased traffic rate was measured in terms of the packet rate [Mpps]. We assumed the packet length to be 1000 bits and the processing capability of the router to be 40 Mpps. The figures show that the maximum traffic rate at the routers gradually increased as the number of wavelengths used for primary lightpaths was increased. Conversely, the traffic rate at the routers increased as the number of backup lightpaths was reduced. With the min–hop–first approach, the loads were larger than with the largest–traffic–first approach. That is, the largest–traffic–first approach is a better choice for an IP over WDM network if the IP routers are a primary cause of bottlenecks within the network.

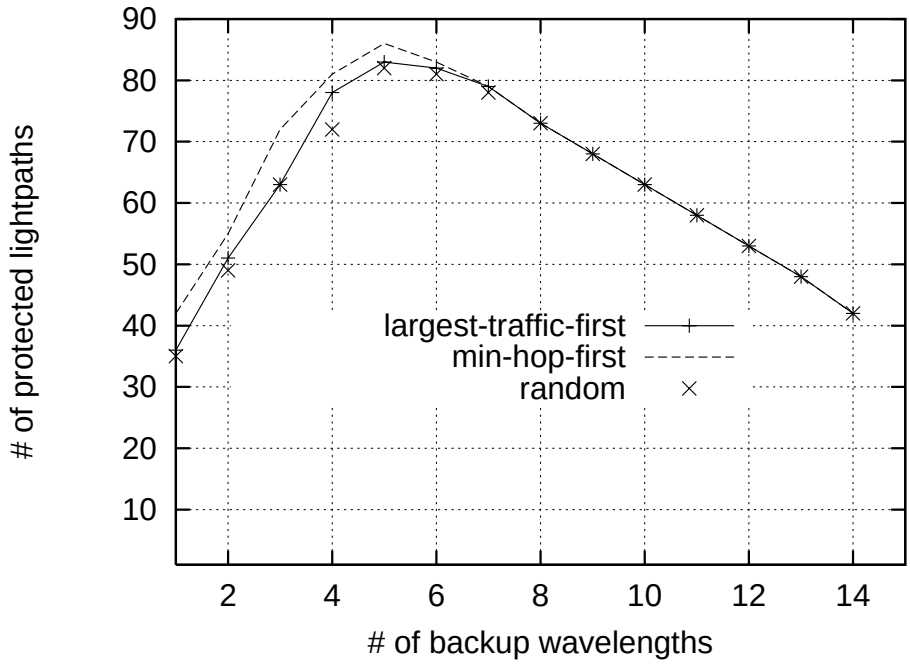


Figure 3.7: Number of protected lightpaths with fixed number of available wavelengths.

To clarify this difference, we next examined the three approaches in terms of traffic volume. As shown in Fig. 3.11, as the number of wavelengths used for the primary lightpaths was increased, the volume of traffic protected by the backup lightpaths first increased, then decreased, because the number of wavelengths available for backup got smaller. In contrast, the amount of traffic that can be restored by the IP routing protocol increases as the number of wavelengths used for the primary lightpaths is increased. The total volume of traffic not protected by the backup lightpaths is shown in Fig. 3.12. When the number of wavelengths in the fiber was below nine, the traffic was perfectly protected. However, when it exceeded nine, the volume of traffic not protected suddenly increased. Of course, it can be restored by IP routing after the routing table is updated, which we will discuss next.

First, however, from Figs. 3.11 and 3.12, we see that the largest–traffic–first approach protected more traffic than the min–hop–first approach. This is because it allocates the backup lightpaths based on traffic volume.

Finally, we discuss the traffic volume protected after the IP routing table is updated. Figure 3.13 shows the volume of traffic protected when the routing tables at the nodes were simultaneously updated. The difference from Fig. 3.11 is due to changes in several IP routes. Although IP does not select several backup lightpaths as routes, we must take this possibility into account.

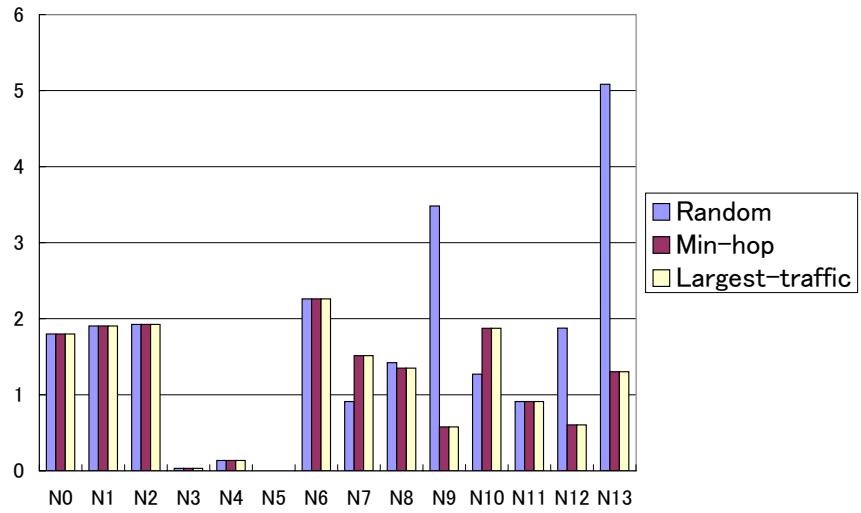


Figure 3.8: Maximum traffic load at IP routers after single–fiber failure; number of wavelengths used for primary lightpaths is 10

One of our future research topics to build a set of perfectly backed–up lightpaths such that IP chooses those lightpaths as its own routes. Figure 3.14 is the complement to Fig. 3.13; it shows the volume of traffic not protected after the routing tables were updated.

These results clearly show that our proposed algorithm can be used to estimate the number of wavelengths required for primary and backup lightpaths to achieve a good compromise between high performance (by establishing a WDM logical topology) and high reliability (by protecting a larger number of primary lightpaths). Using it, we found that the min–hop–first approach is better for improving network reliability, while the largest–traffic–first approach is better for reducing the traffic loads at the IP routers.

We also applied our heuristic algorithms to NTT's backbone networks, which have 49 nodes and 200 links. For the traffic matrix, we used publicly available traffic data [44]. We again found that the largest–traffic–first approach protects more traffic than the other approaches.

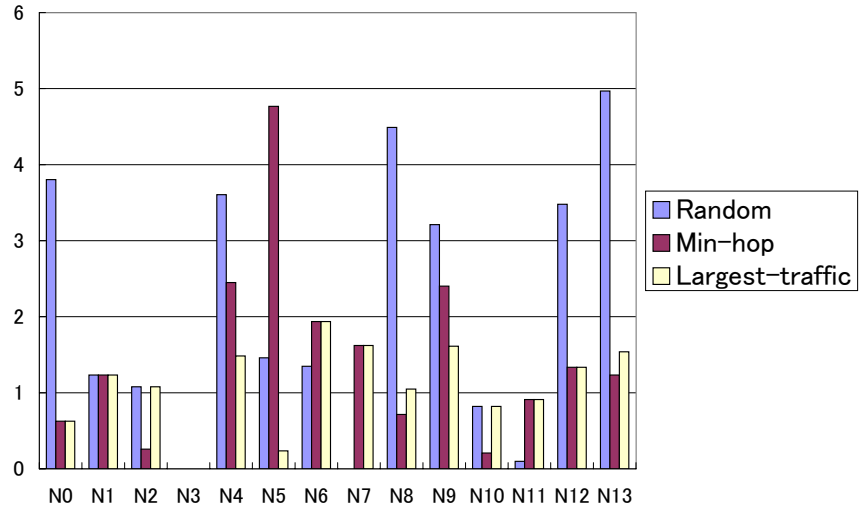


Figure 3.9: Maximum traffic load at IP routers after single–fiber failure:; number of wavelengths used for primary lightpaths is 12

3.4 Conclusion

In this chapter, we discussed the multi–layer survivability in IP over WDM networks. In Section 3.2, we considered the reliability mechanism in the IP over WDM network. Assuming a single–fiber failure in the network, we formulated the shared path protection mechanism as an optimization problem. It is formulated as MILP, and computationally intensive as the network size grows. Accordingly, we proposed heuristic algorithms and compared the results with the solution obtained by MILP. Through numerical examples, we compared the number of wavelengths required for the reliable network. We next considered the functional partitioning of IP routing and WDM protection to improve reliability. Based on using our heuristic algorithm, we discussed the effect of interaction between the IP and WDM layers. We showed that the largest-traffic-first approach is best if our primary concern is traffic load at the IP routers after a failure.

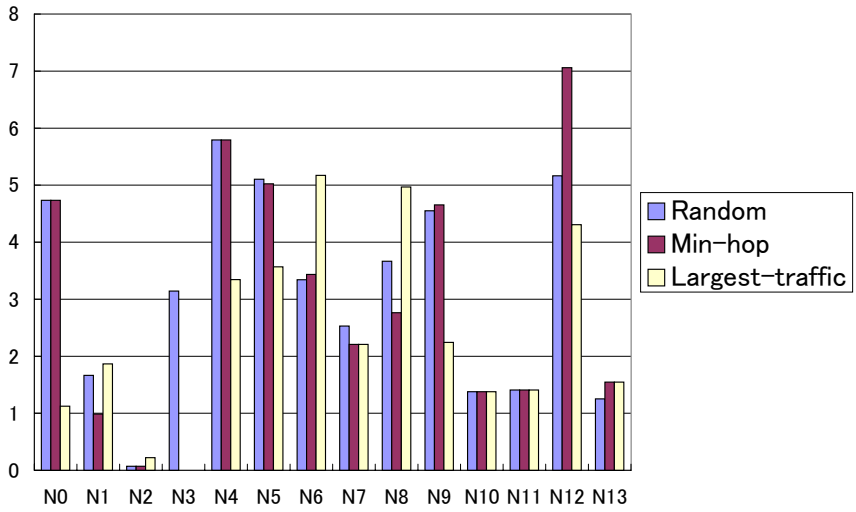


Figure 3.10: Maximum traffic load at IP routers after single–fiber failure: number of wavelengths used for primary lightpaths is 14

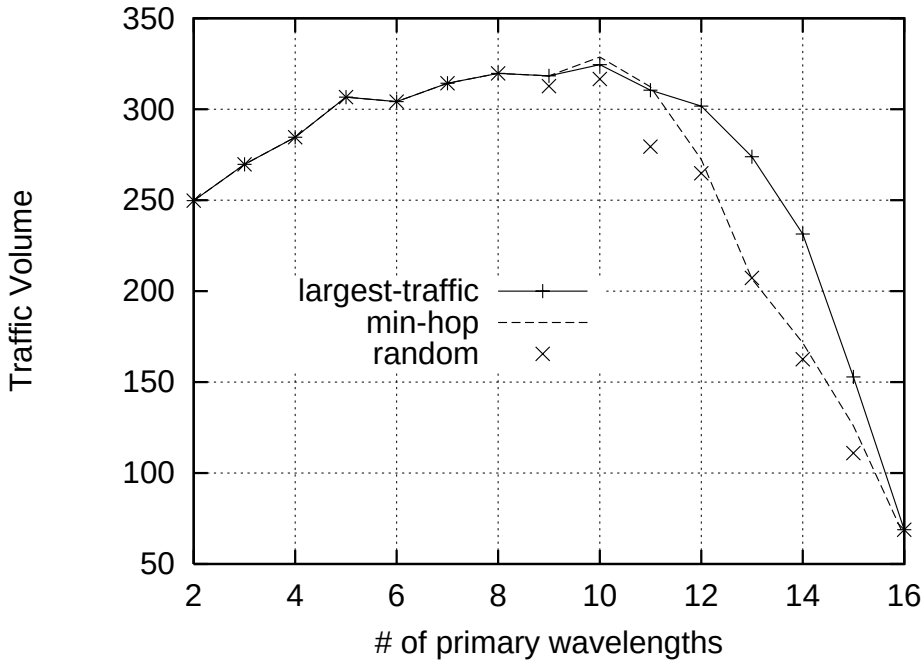


Figure 3.11: Total volume of traffic protected by backup lightpaths before IP routing table update

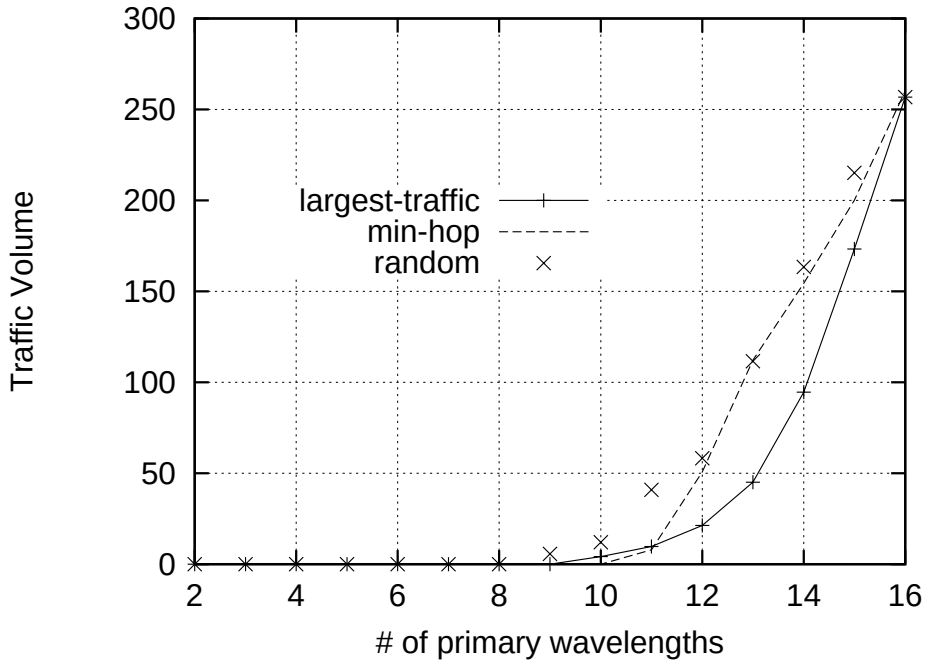


Figure 3.12: Total volume of traffic not protected by backup lightpaths before IP routing table update

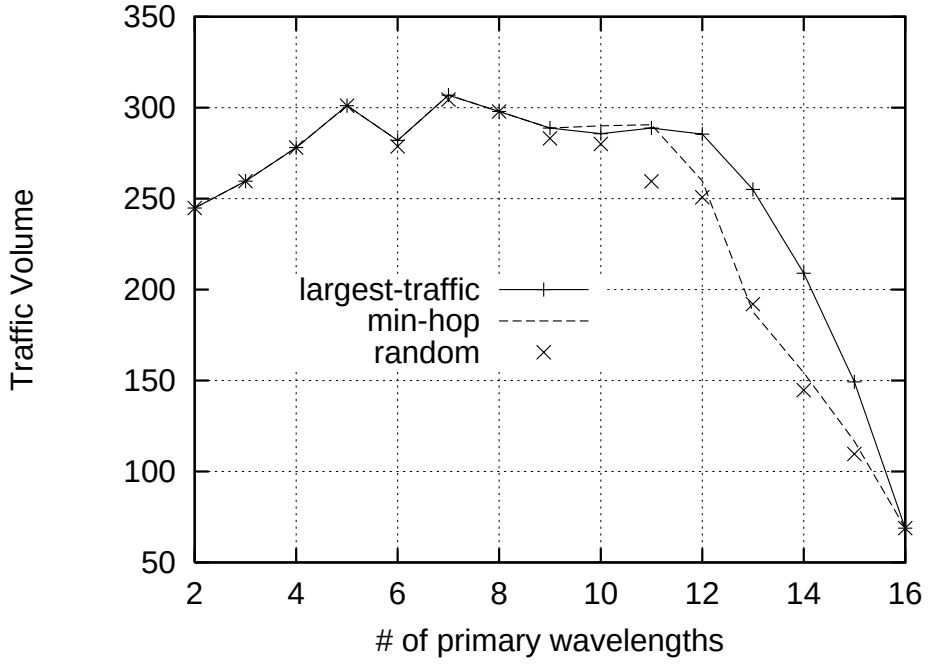


Figure 3.13: Total volume of traffic protected by backup lightpaths after IP routing table update

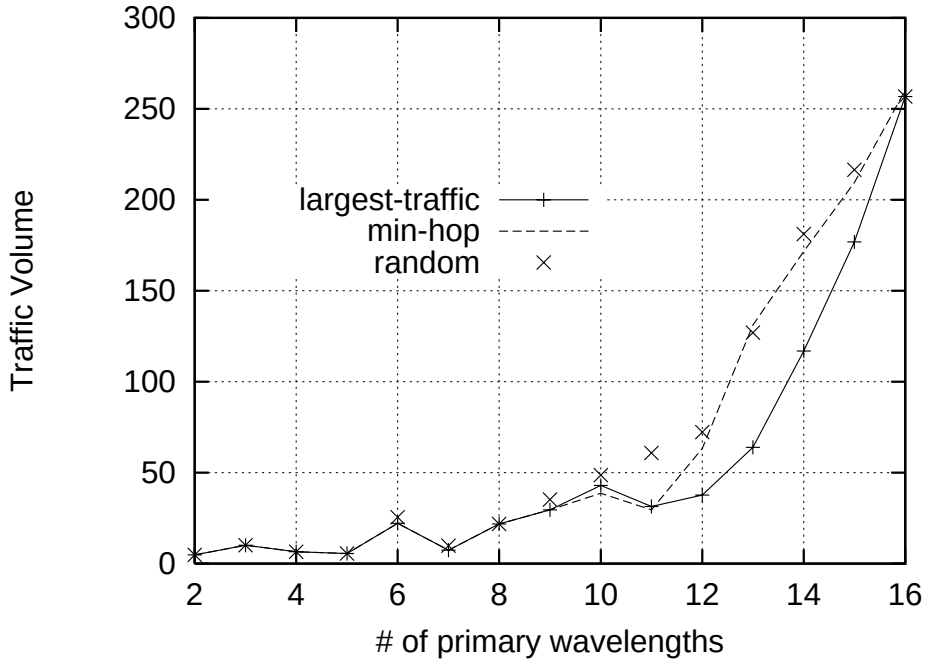


Figure 3.14: Total volume of traffic not protected by backup lightpaths after IP routing table update

Chapter 4

Methods for Designing Logical Topologies for Quality of Reliability

Building a highly reliable network is becoming more important as the number of wavelengths increase with advances in WDM technology. In this chapter, we introduce QoR (Quality of Reliability), a concept related to QoS that reflects reliability in a WDM network. QoR can be used to guarantee a maximum recovery time, based on the user's request, and guarantee that backup lightpaths will be available. In the conventional quality–based lightpath configuration methods, the failure-recovery quality is guaranteed only probabilistically. That is, these methods are aimed at improving the efficient usage of network resources, but at the cost of a 100% guarantee of failure recovery. Thus, we introduce QoR as a new QoS metric aimed at providing highly reliable lightpaths.

4.1 Quality Metrics in Existing Fault Tolerance Methods

Several researchers have discussed methods to design logical topologies with protection [9, 23, 24]. Most of the existing protection methods try to minimize the number of wavelengths when designing the logical topology or to maximize the total throughput within the network. The shared protection method is an effective way to further reduce the number of wavelengths needed under the single–failure assumption.

Wavelength resources can also be used more effectively if we introduce several classes of guarantee with respect to the probability of failure recovery [24]. The conventional protection

method only guarantees complete failure recovery (i.e., a single class with a 100% guarantee). Likewise, another guarantee class with a smaller probability of failure recovery can be offered [23]. That is, backup lightpaths are provided for only the connections requesting a higher class of protection, and thus a higher probability of failure recovery.

Recent research has focused on providing QoS with respect to failure recovery in an optical WDM network [23, 24]. QoP (Quality of Protection) was then introduced to realize QoS in an optical network [24], through a probabilistic failure recovery model where only a certain fraction of traffic, which can be specified by the user, is restored after failure. A different approach from [9, 24, 26]. is to consider the possibility of two or more components failing at the same time (a multiple–failure assumption) and assume that each primary lightpath has its own reliability metric that can be determined from the failure probabilities of the network components [23]. Based on this approach, backup lightpaths are partially configured for the primary lightpath according to the specified probability. However, in these QoP–based lightpath configuration methods, the failure-recovery quality is guaranteed only probabilistically. That is, these methods are aimed at improving the effective usage of network resources, but at the cost of a 100% guarantee of failure recovery.

In this chapter, we introduce a new metric to define QoS with respect to the reliability provided by the optical layer. This metric, which is based on the maximum recovery time defined as the maximum time between failure occurrence and the time at which traffic is switched to the backup lightpath, is QoR. The QoR can be used to guarantee the maximum recovery time according to user requests and provide a 100% guarantee that a backup lightpath will be available.

4.2 QoR and Recovery Time Modeling

4.2.1 QoS Classification based on Maximum Failure Recovery Time

In the QoR definition, class is associated with the maximum recovery time. By specifying a QoR class, we can guarantee a corresponding maximum recovery time upon failure for each connection. In the QoR system we propose, QoR_1 (the highest class) guarantees the minimum failure recovery time. QoR_∞ provides no lightpath protection, and the actual failure recovery

Table 4.1: QoR (Quality of Reliability)

QoR$_1$	failure recovery within D_{min}
QoR$_2$	failure recovery within $(D_{min} + D_{scale})$
QoR$_3$	failure recovery within $(D_{min} + 2D_{scale})$
$\vdots$	$\vdots$
QoR$_n$	failure recovery within $(D_{min} + (n-1)D_{scale})$
$\vdots$	$\vdots$
QoR$_\infty$	no lightpath protection provided

is left to the upper-layer protocol (e.g., IP). More specifically, QoR_n guarantees the maximum recovery time associated with class n, denoted as $RT(QoR_n)$. One of its simplest forms is

$$RT(QoR_n) = a + b * f(n), \tag{4.1}$$

where a, b, and $f(n)$ are determined by the network operator based on the network environment. By configuring $f(n)$, a QoR class can be represented in an arithmetic or geometric progression, or any other form. In the numerical evaluation of Section 4.4, $f(n)$ is set simply as

$$f(n) = n - 1, \tag{4.2}$$

and $a = D_{min}$ is the minimum recovery time, which includes the time needed to switch from the primary lightpath to the backup lightpath. The step–width of the recovery time is $b = D_{scale}$, which includes the processing time to propagate the failure information and to reserve wavelengths at each node of the backup lightpath. The function $RT(QoR_n)$ should be appropriately determined for a given network environment, but specification of only a class–dependent recovery time is not sufficient. We must consider a more precise definition of the recovery time. The node–pair dependent recovery time is discussed in Section 3.2.

4.2.2 QoR Specification for Each Node Pair

There may be no route that can be used to configure backup lightpaths in a way that guarantees the maximum recovery time specified for the QoR class. Figure 4.1 shows an example of such a case. In the figure, there are two routes from node A to node F. One is $[A \rightarrow B \rightarrow C \rightarrow D \rightarrow$

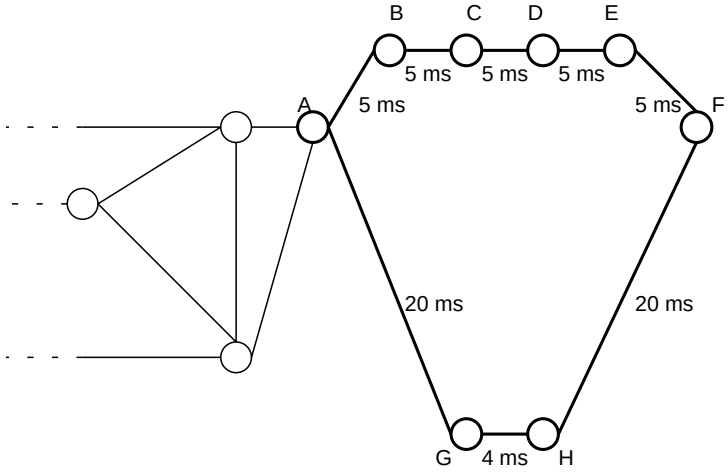


Figure 4.1: Example topology

Table 4.2: QoR dependent on node pair

QoR	Maximum recovery time	QoR_{12}		QoR_{ij}	
QoR_1	D_{min}	—		—	
QoR_2	$D_{min} + 1 * D_{scale}$	—		$QoR_{ij}(1)$	
QoR_3	$D_{min} + 2 * D_{scale}$	$QoR_{12}(1)$		$QoR_{ij}(2)$	
QoR_4	$D_{min} + 3 * D_{scale}$	$QoR_{12}(2)$	$\cdots$	$QoR_{ij}(3)$	$\cdots$
QoR_5	$D_{min} + 4 * D_{scale}$	$QoR_{12}(3)$		$QoR_{ij}(4)$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
QoR_∞	No protection lightpaths	$QoR_{12}(\infty)$		$QoR_{ij}(\infty)$	

$E \rightarrow F]$, and the other is $[A \rightarrow G \rightarrow H \rightarrow F]$. The propagation delay of the first route is 25 ms in total, while that of the second is 44 ms. In this situation, if node pair AF requires a QoR class with a maximum recovery time of 20 ms, no lightpath route would provide the required recovery time. The recovery time includes the time needed to propagate the failure notification, and this takes more than 20 ms regardless of the route assigned to the primary lightpath.

Thus, the QoR concept should be extended to allow the network operator to specify the QoR class for each node pair ij. This means the network operator will begin by examining the smallest possible recovery time for node pair ij, determined by including the propagation delay between nodes i and j, the node delay for lightpath switching, and so on. This minimum time is set as the recovery time for the highest class for node pair 12, which is represented as $QoR_{12}(1)$. The recovery times of the lower classes, $QoR_{12}(2)$, $QoR_{12}(3)$, $\ldots$ are then determined in the same way. In the example shown in Table 4.2, the original QoR classes are defined by Eq. (4.1). First, $QoR_{12}(1)$ for node pair 12 is mapped to QoR_3. Then, the network operator maps

$QoR_{12}(n)$ to QoR_{n+2}. The network operator makes this decision for each node pair. The mapped QoR_{ij} are provided to end users, and an end user using node pair ij can choose the preferred class from $QoR_{ij}(\cdot)$.

4.2.3 Modeling Recovery Times

In this section, we describe the behavior of the protection method and explain how the recovery time is determined. As shown in Fig. 4.2, primary lightpath L is protected by several backup lightpaths P_x $(1 \leq x \leq B)$. Here, B is the number of backup lightpaths for primary lightpath L and is at most equal to the number of intermediate nodes that the primary lightpath traverses. We also define *segment* x as a part of the primary lightpath between the source and destination nodes of P_x (denoted as S_x and D_x, respectively). Using this notation, we will describe the protection method and show how the recovery time is modeled.

To provide QoR, we need to set up several backup lightpaths in such a way that the maximum recovery time of each segment provided by each backup lightpath does not exceed a threshold value. For this purpose, we modify the SLSP (Short Leap Shared Protection) method [34]. In the original SLSP, several backup lightpaths are configured for each primary lightpath, so that any two neighboring backup lightpaths overlap (Fig. 4.3). Unlike the shared protection methods, SLSP enables recovery from a node failure. For example, if a failure occurs at node D, node C switches the traffic to the backup lightpath directly connected to node H.

The quality metric is estimated by specifying the maximum length of the backup lightpath such that its length will be shorter than the threshold [34]. However, when SLSP is used, only the length of the backup lightpath is specified. In contrast, we want to allow users to specify the maximum recovery time for primary lightpath L. Such a QoR can be realized by allocating backup lightpaths in a way that ensures the maximum recovery time of each segment is smaller than that segment's threshold. Positioning two neighboring segments so that they overlap also enables recovery from a single-node failure.

The recovery time is modeled as shown in Fig. 4.2. When a failure occurs in segment x, the nodes next to the failed component send information to the nodes that precede it. When the failure information arrives at node S_x, it reserves wavelengths on the prepared backup lightpath, P_x, by sending a reservation signal to D_x through nodes $k, k+1, \ldots k+H_x$. Here, H_x is the hop

count of backup lightpath P_x. When the activation is completed, node S_x switches the traffic on the primary lightpath onto P_x. The recovery time when a failure occurs in segment x thus consists of three factors;

- Delay needed to propagate the failure information to node S_x
- Configuration time needed to reserve wavelengths at each node of backup lightpath P_x
- Switching time needed to move the traffic from the failed primary lightpath onto backup lightpath P_x

Thus, the maximum recovery time when a failure occurs in segment x (denoted as RT_x) is

$$RT_x = \sum_{k=S_x}^{h_x} d_{k(k+1)} + D_{node} \times (H_x + 1) + D_{conf}, \tag{4.3}$$

where D_{node} is the wavelength reservation time needed at each node along P_x, and D_{conf} is the switching time at node S_x. In Eq. (4.3), d_{ij} is the propagation delay between nodes i and j. h_x is the maximum hop count that the failure information has to traverse in segment x:

$$h_x = \begin{cases} D_x - 1, & D_x \leq S_{x+1}, \\ S_{x+1} - 1, & S_x < S_{x+1} < D_x. \end{cases} \tag{4.4}$$

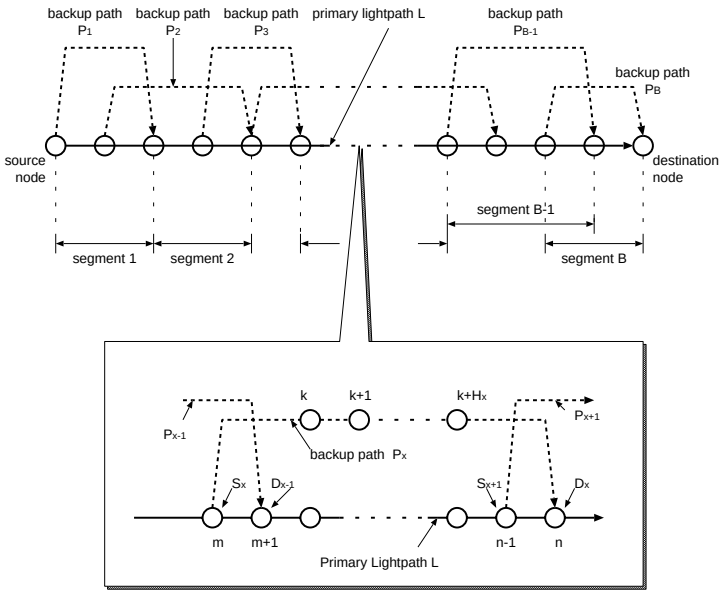


Figure 4.2: Primary lightpath protected by several backup lightpaths P_x $(1 \leq x \leq B)$

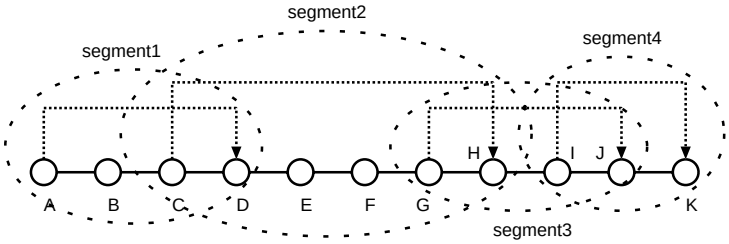


Figure 4.3: Illustrative example of SLSP

The maximum recovery time for primary lightpath L, $RT_{max}(L)$, is the maximum of RT_x for each segment x, and thus,

$$RT_{max}(L) = \max_{1 \le x \le B} RT_x. \tag{4.5}$$

4.3 Logical Topology Design Algorithms for Satisfying QoR Requirements

In this section, we describe three heuristic algorithms for designing logical topologies that satisfy the QoR requirements. The objective in designing the logical topology is to minimize the number of wavelengths when the traffic volume and QoR requirements for each node pair are given. In essence, all three algorithms work as follows.

Step 1: For each node pair ij, set metric β_{ij} based on $QoR_{ij}(\cdot)$, which is used to determine the order of node pairs assigned to lightpaths.

Step 2: In descending order of metric β_{ij}, assign the route and the wavelengths.

The route of a backup lightpath is assumed to be configured on the shortest hop route between source node S_x and destination node D_x, and the route is disjoint with the links or nodes of its primary lightpath, L, except nodes S_x and D_x. The backup lightpath is set up based on the hop count because, as shown in Eq. (4.3), the failure recovery time is highly dependent on the number of hops in the recovery model.

Before explaining how the wavelength is allocated to the backup lightpaths, we should mention that wavelength conversion is not taken into consideration here, so the same wavelength must be used for each lightpath (i.e., a wavelength continuity constraint). When a backup lightpath is set up to protect one segment of L, the same wavelength on L must be assigned to the

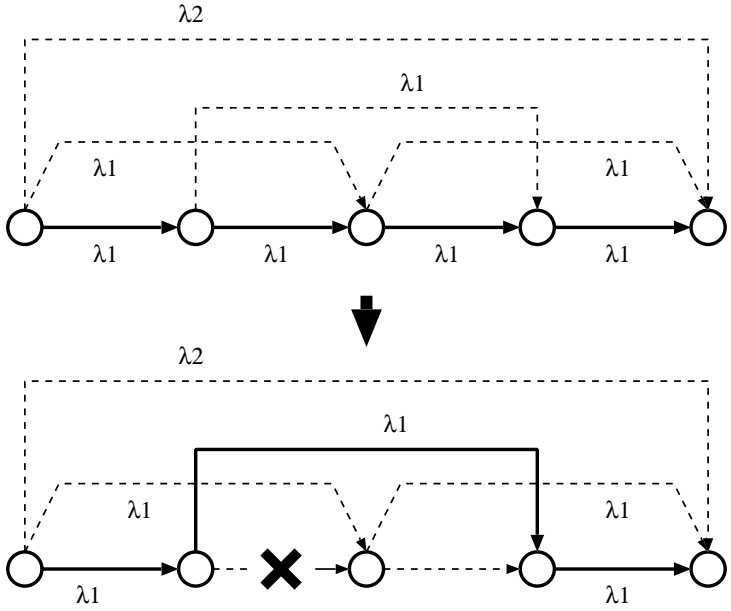


Figure 4.4: Wavelength continuity

backup lightpath since the backup lightpath will become part of the primary lightpath after a failure (Fig. 4.4). However, when the source and destination nodes of the backup lightpath are identical to those of L, the wavelength of the backup lightpath does not have to be the same as that assigned to the primary lightpath because in this case the backup lightpath does not share any links with the primary lightpath.

In what follows, we will introduce two algorithms for wavelength assignment of primary and backup lightpaths: Max–Shared algorithm and Layered Graph algorithm.

4.3.1 First–Fit Algorithm

The First–Fit algorithm first determines the routes of the primary and backup lightpaths. This is a combinational optimization problem to determine routes for the best set of a primary lightpath and backup lightpaths. To simplify the algorithm, the primary lightpath is routed by selecting the route with the smallest propagation delay between nodes, while the backup lightpath is set on the route that has the minimum hop count on the link/node disjoint path.

After the routes of all the primary and backup lightpaths are determined, a wavelength is assigned to each lightpath based on the First–Fit (FF) policy [45]. The FF policy works as follows. If the algorithm discovers that several wavelengths $\{\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_n}; i_1 < i_2 < \ldots <$

i_n } are available for the lightpath, it selects the one with the lowest index (i.e., λ_{i_1} is selected). Note that the assignment depends on whether the source and destination nodes of the backup lightpath are the same as those of the primary lightpath. That is,

- If the nodes are identical, different wavelengths can be assigned to the primary lightpath and the corresponding backup lightpath. Therefore, the algorithm first searches for an available wavelength for the primary lightpath. The wavelength for the backup lightpath is then determined independently of the wavelength assignment for the primary lightpath.

- If a backup lightpath only partially protects the primary lightpath, the primary lightpath and the set of backup lightpaths must be assigned the same wavelength to satisfy the wavelength continuity constraint.

4.3.2 Max–Shared Algorithm

In the Max–Shared algorithm, the routes of the primary lightpath and a set of backup lightpaths are determined and then wavelengths are assigned to these lightpaths. The routing algorithm for primary and backup lightpaths is the same as for the First–Fit algorithm, i.e., finding the minimum propagation delay for the primary lightpath and the minimum hop count for the backup lightpaths. The difference from the First–Fit algorithm is in the wavelength assignments. In the Max–Shared algorithm, all available wavelengths are examined for possible assignment to both the primary and backup lightpaths, and the best one is chosen. During the evaluation of each wavelength, the number of links newly used for the backup lightpath is counted, and the count is set as the cost of the wavelength. Only if the source and destination nodes of a backup lightpath are the same as those of the primary lightpath, the wavelength for the backup lightpath is assigned independently of the primary lightpath. Note that we select the wavelength with minimum cost if several wavelengths are available for the backup lightpath.

The Max–Shared algorithm enables more efficient use of wavelength resources compared to the First–Fit algorithm. This is because the it assigns a wavelength to each set of primary and backup lightpaths selected from all possible wavelengths to maximize the number of wavelengths that are shared with other lightpaths, while the First–Fit algorithm does not try all available wavelengths.

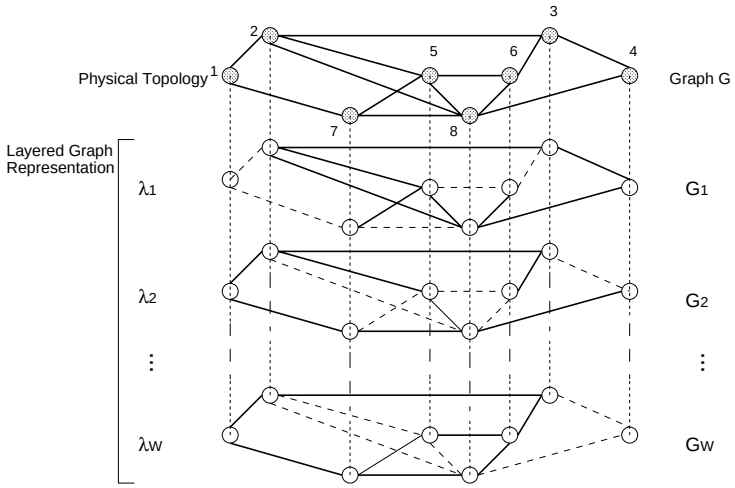


Figure 4.5: Example of a layered graph (number of wavelengths = W)

4.3.3 Logical Topology Design Algorithm based on a Layered Graph

We next propose a new algorithm which is based on a layered graph. The layered graph consists of a set of wavelength graphs $G_n(1 \leq n \leq W)$, each of which corresponds to the graph for wavelength λ_n [46]. Wavelength graphs are independent of each other if wavelength conversion is not allowed. The layered graph enables us to determine both the route and the wavelength of the lightpath at the same time by calculating the shortest route for each wavelength. Figure 4.5 shows an example of a layered graph in which the number of wavelengths is set to W. The solid lines in each wavelength graph, G_n, indicate that wavelength λ_n is free on that link, whereas dotted lines indicate that the wavelength is already being used for a primary or backup lightpath. The metric for each edge of G_n is the propagation delay of the corresponding link. To determine the wavelength to be assigned to each set of primary and backup lightpaths, we introduce cost C^n for each wavelength λ_n; it denotes the number of links where wavelength λ_n is newly used by the set of primary and backup lightpaths. The proposed algorithm works as follows.

Step 0: Set w, representing the number of wavelengths needed to construct the logical topology, to 0.

Step 1: For each possible lightpath between nodes i and j, perform Steps 2 through 4.

Step 2: Update w by calculating the number of wavelengths already used by some links.

Step 3: From λ_1 to λ_{w+1}, perform the following steps. (Assume that λ_n is currently chosen in the following steps.)

Step 3.1: Check whether a route consisting of only unreserved wavelengths exists between node pair ij on graph G_n. If such a route does not exist, the primary lightpath cannot be set up. If so, go back to Step 3 and check the next wavelength on G_{n+1}. Otherwise, the primary lightpath, denoted by L_{ij}, is set up on the route using λ_n, and update the metric of edges on G_n. That is, delete the corresponding links on L_{ij} from G_n and set the cost of primary lightpath C_p^n to the number of deleted links.

Step 3.2: Based on SLSP, a set of backup lightpaths $\{P_1, P_2, \ldots, P_k\}$, each of which should satisfy the QoR_{ij} requirements, can be derived. For this purpose, the route of the backup lightpaths are determined such that the backup lightpaths are disjoint to the primary lightpath, L_{ij}, and the hop count of the route is minimal. To satisfy these two conditions, calculate C_r^n, the cost for assigning wavelength λ_n to backup lightpath P_r ($1 \leq r \leq k$) and determine the set of backup lightpaths for L_{ij}.

Step 3.2.1: When the source node and destination node of P_r are identical to those of L_{ij}, P_r can be tentatively assigned to each wavelength λ_i ($1 \leq i \leq w+1$). If the backup lightpaths are partially configured at L_{ij}, perform Step 3.2.2 only for graph G_n because the backup lightpath partially protecting the primary lightpath must be assigned the same wavelength as the primary lightpath.

Step 3.2.2: If backup lightpath P_r can be set up on wavelength graph G_e, count the number of links that are newly used on G_e, and set the cost, C_e, of P_r to the number. After checking all wavelengths (i.e., G_1 through G_{w+1}), select the e' for which cost $C_{e'}$ of the corresponding $G_{e'}$ is minimal. Then, set $C_{e'}$ to C_r^n.

Step 3.3: Set C^n to $C_p^n + \sum_{r=1}^{k} C_r^n$. Here, C^n is the cost of wavelength λ_n for setting up both the primary and backup lightpaths between nodes i and j. Go back to Step 3.

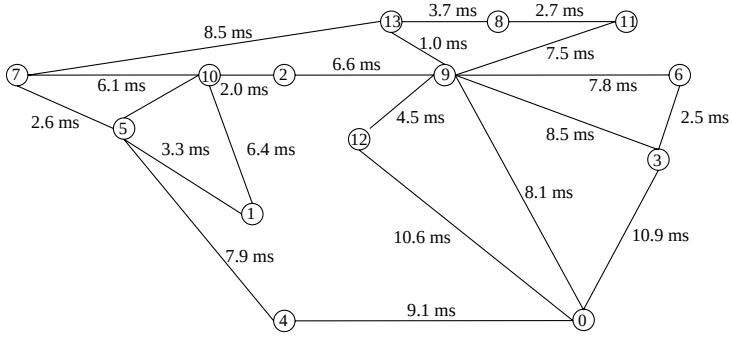


Figure 4.6: 14–node random network

Step 4: Select a such that C^a is the minimum value of $\{C^1, C^2, \ldots, C^{w+1}\}$ and assign wavelength λ_a to P_a and P_r (which is partially protecting P_a). Then, assign $\lambda_{e'}$, which is pre–calculated in Step 3.2.2, to the path protection backup lightpath.

The algorithm calculates the cost of assigning the primary and backup lightpaths for each wavelength in Steps 3.1 and 3.2, respectively. In Step 3.3, cost C_r^n is calculated for each backup lightpath r on λ_n, where cost means the number of newly used wavelength resources. Step 3 determines the actually used wavelength that minimizes the cost of assigning both the primary and backup lightpaths and sets up the lightpaths using λ_a. Note that the above algorithm counts the number of wavelengths needed, w. However, when the number of wavelengths is set to W, Steps 3.1 through 3.4 are performed from λ_1 to λ_W.

4.4 Numerical Evaluation and Discussion

4.4.1 Network Models

We used a 14–node NSFNET model (Fig. 2.2) and a traffic matrix (Table 4.3) to evaluate the three algorithms. The traffic matrix contains relative values of the amount of traced traffic on NSFNET in 1992. We introduced traffic scale factor α and used the traffic matrix multiplied by α as the actual traffic demand. We assumed Gbps to be the unit for the traffic matrix.

The bandwidth of each wavelength was set to 10 Gbps, and a connection whose requested bandwidth exceeded 10 Gbps was assigned multiple lightpaths to carry the traffic. When two or more lightpaths were assigned to a connection, we set them on the same route.

We also used a randomly generated network with 21 links placed randomly within the 14–node network. Note that the numbers of links and nodes were the same as for NSFNET. The propagation delay for a link was also given randomly and ranged from 0.7ms to 11.2ms, which are the shortest and longest propagation delays of links in the original NSFNET. A traffic matrix for the network was randomly selected between 0.0004 and 21.030, the minimum and maximum values in Table 4.3.

In the following subsections, $D_{min} = 10\text{ms}$, $D_{scale} = 2\text{ms}$, $D_{node} = 1\text{ms}$, and $D_{conf} = 0$.

4.4.2 Evaluation Results and Discussion

First, we will look at the number of wavelengths needed with each algorithm when every node pair ij requests the same QoR_{ij}. More specifically, in the current example, the network operator prepares QoR_{ij} classes that are dependent on node pair ij. For example, consider node pair $3, 4$ in NSFNET (Fig. 2.2). If the primary lightpath is set to route $[3 \rightarrow 4]$ and a backup lightpath is set to $[3 \rightarrow 1 \rightarrow 2 \rightarrow 5 \rightarrow 4]$, the maximum recovery time is 6.8 ms. If lightpaths are set on different routes, the maximum recovery time will be more than 6.8 ms. Therefore, 6.8 ms is the minimum of the maximum times that can be guaranteed for node pair $3, 4$. Here, if D_{min} and D_{scale} are set to 5 ms and 1 ms, respectively, the maximum recovery time guaranteed in QoR_2 is 6 ms and that in QoR_3 is 7 ms. Accordingly, $QoR_{34}(1)$ is set to QoR_2.

In the current example, however, all node pairs are assumed to request the same class to simply show the relationship between QoR_{ij} and the number of wavelengths needed with each algorithm. The horizontal axis in Fig. 4.7 shows the class number that all node pairs request. The vertical axis shows the number of wavelengths needed to set up all the primary and backup lightpaths to fulfill the requests. To obtain this figure, we used the NSFNET network model (Fig. 2.2) and set traffic scale factor α to 1.

The proposed algorithm, based on the layered graph, enabled the wavelength resources to be used more efficiently than with the other algorithms, especially when QoR_{ij} was high (e.g., $QoR_{ij} = 1$ or 2). When QoR_{ij} is high, more backup lightpaths must be configured throughout the network to achieve the required recovery times. In this situation, the layered graph algorithm can determine routes for each primary and backup lightpath in a way that requires fewer additional wavelength resources. Note that a solid line without points represents the result when no

Table 4.3: Traffic matrix for NSFNET

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.000	0.109	0.206	0.014	0.045	0.004	0.043	0.145	0.051	0.010	0.007	0.008	0.000	0.033
1	1.171	0.000	0.856	0.062	1.112	0.777	0.362	1.579	0.366	1.661	0.203	3.781	0.483	1.319
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.031	0.341	1.364	0.000	0.190	0.060	0.070	0.288	0.200	0.326	0.307	0.669	0.008	0.401
4	0.028	6.751	1.902	0.343	0.000	0.403	1.077	6.222	2.402	1.792	0.045	7.903	0.997	0.529
5	0.000	0.581	0.342	0.552	0.340	0.000	0.261	0.268	0.087	0.387	0.004	0.084	0.006	0.248
6	0.175	2.202	10.231	0.447	2.203	0.790	0.000	11.410	1.982	2.195	0.078	7.140	0.033	3.284
7	0.239	6.384	21.030	0.852	2.821	0.266	9.708	0.000	4.395	3.300	1.137	4.863	0.553	1.385
8	0.645	1.893	3.735	0.600	2.499	0.681	2.506	6.102	0.000	3.962	1.452	12.750	2.334	0.076
9	0.005	3.529	1.026	0.373	2.234	0.948	0.498	5.708	0.684	0.000	0.630	1.764	0.591	0.076
10	0.010	0.102	0.313	0.169	0.024	0.006	0.081	0.145	0.058	0.712	0.000	0.084	0.006	0.050
11	0.128	2.615	0.100	0.594	2.486	0.132	0.549	4.057	2.953	2.237	1.050	0.000	0.101	0.054
12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
13	0.073	2.909	1.363	0.989	3.561	1.207	0.644	2.879	0.467	0.000	0.399	0.000	1.075	0.000

Table 4.4: Traffic matrix for random network

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.000	6.014	16.019	16.596	7.874	7.979	9.556	20.655	17.433	16.887	16.318	3.662	8.016	21.928
1	10.809	0.000	8.875	16.940	17.114	4.149	10.389	1.429	12.390	9.286	14.597	0.614	9.435	23.283
2	6.535	18.131	0.000	1.331	6.372	10.558	21.717	12.767	17.530	5.591	20.742	17.462	2.246	4.555
3	10.349	18.561	22.590	0.000	8.741	16.489	9.399	17.612	23.805	2.514	12.137	10.195	18.315	0.528
4	19.477	8.912	1.138	4.912	0.000	8.195	22.045	13.420	23.898	18.793	14.354	21.615	7.561	22.260
5	3.207	18.679	15.722	19.825	13.611	0.000	2.072	14.386	12.201	1.189	21.251	11.976	9.178	21.057
6	4.866	21.311	21.628	23.178	12.215	17.105	0.000	8.090	3.729	12.394	6.662	1.775	16.190	20.936
7	10.944	6.544	18.552	8.881	4.804	12.135	3.561	0.000	20.522	7.960	7.548	12.970	12.723	19.745
8	14.156	0.354	22.097	23.330	11.787	2.964	11.021	9.415	0.000	2.142	23.233	16.897	0.608	2.962
9	5.291	21.642	19.109	21.477	18.579	20.430	18.397	3.511	5.311	0.000	13.577	15.642	23.244	10.099
10	13.978	6.792	13.446	17.077	16.913	17.978	17.428	15.011	7.688	5.215	0.000	17.971	18.705	5.007
11	20.109	8.318	21.900	11.093	1.657	3.191	8.736	20.762	15.044	3.315	7.572	0.000	23.817	6.822
12	12.880	13.394	12.840	2.504	23.489	17.194	9.293	3.315	10.272	2.206	21.289	18.076	0.000	7.593
13	4.977	13.667	1.564	14.059	18.670	12.049	22.373	16.570	23.139	0.030	10.137	22.251	11.169	0.000

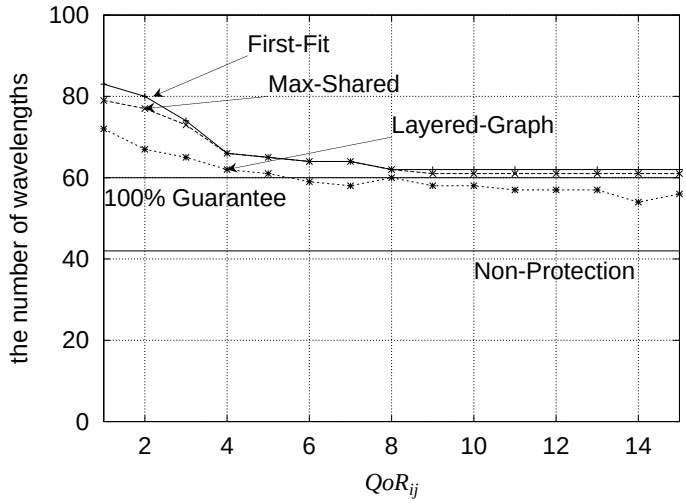


Figure 4.7: QoR_{ij} vs. number of wavelengths in NSFNET ($\alpha = 1$)

backup lightpath is prepared for a primary lightpath (labeled as Non–Protection in each figure), or the result when only one backup lightpath is configured for each primary lightpath based on the layered graph, which guarantees 100% reliability (labeled as 100% Guarantee in each figure). The number of wavelengths needed with our QoR was at most 100% more than what was needed with no protection. Moreover, the number of wavelengths needed with the three algorithms was at most 50% more than the result for a 100% guarantee.

In Figs. 4.7, 4.8, 4.9, and 4.10, the number of wavelengths needed for a 100% guarantee exceeded the number needed with the layered graph algorithm at lower QoR_{ij}. This is because even for the lower QoR_{ij}, the number of backup lightpaths configured by the layered graph algorithm slightly exceeded the number needed for a 100% guarantee. As a result, the layered graph algorithm required fewer wavelengths than in the 100% guarantee case. This tendency was also observed when the algorithms were applied to a randomly generated network (Fig. 4.8).

When the traffic volume was increased ($\alpha = 2$ in Fig. 4.9 and $\alpha = 5$ in Fig. 4.10), the layered graph algorithm still enabled the most efficient use of wavelength resources. We then limited the number of wavelengths, W, to 20 and configured QoR_{ij} as in the previous evaluations. The number of blocked connections due to a lack of available wavelength resources and the total traffic volume at the blocked connections are shown in Figs. 4.11 and 4.12, respectively, for the NSFNET model with $\alpha = 1$. There was no significant difference among the three algorithms when W = 20. However, when the number of wavelengths was set to 50, the

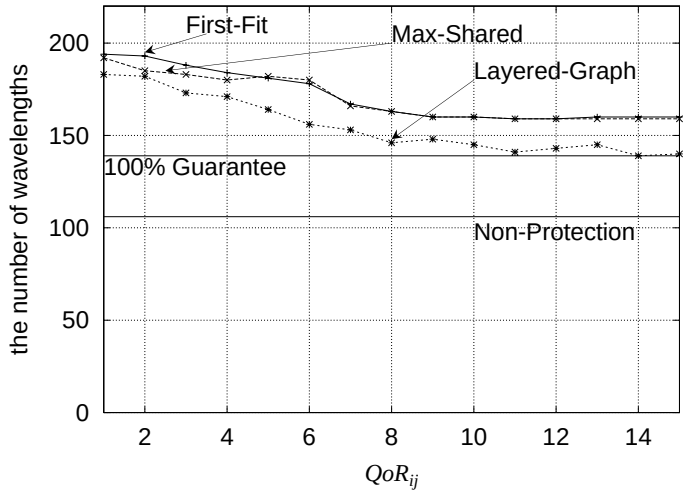


Figure 4.8: QoR_{ij} vs. number of wavelengths in a Random Network ($\alpha = 1$)

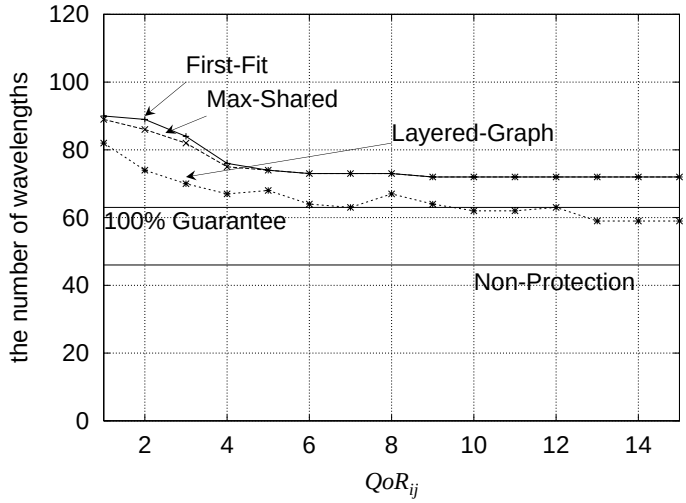


Figure 4.9: QoR_{ij} vs. number of wavelengths in NSFNET ($\alpha = 2$)

advantage of the layered graph algorithm became significant in terms of the number of blocked connections and the amount of blocked traffic, as shown in Figs. 4.13 and 4.14, respectively. This was because the greater number of available wavelengths made it easier to find available wavelength resources for the backup lightpaths that could be shared with other backup lightpaths. In other words, more wavelengths enables more wavelength sharing, and the advantage of the layered graph algorithm becomes increasingly significant as the number of wavelengths rises. Note that there was no blocking in the case of no available backup lightpaths when the number of wavelengths was set to 50.

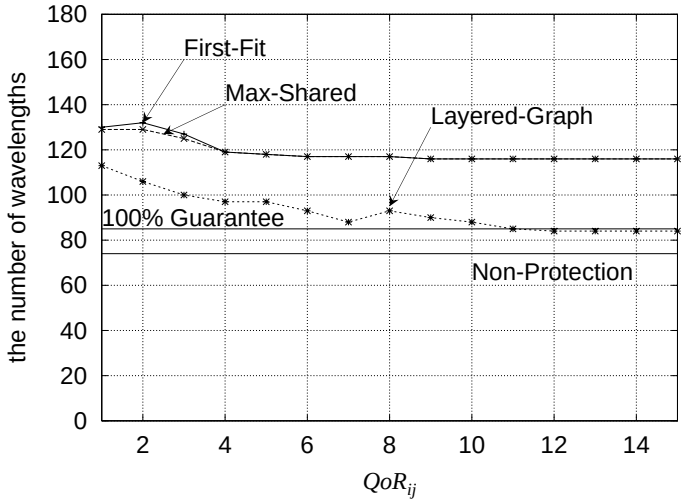


Figure 4.10: QoR_{ij} vs. number of wavelengths in NSFNET ($\alpha = 5$)

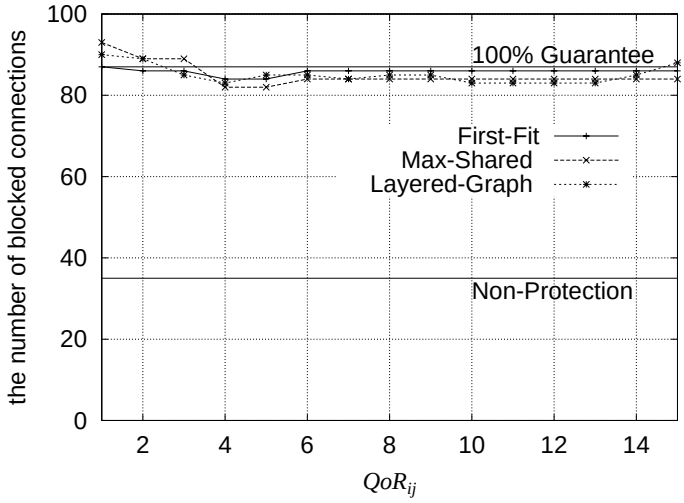


Figure 4.11: Number of blocked connections in NSFNET (W = 20)

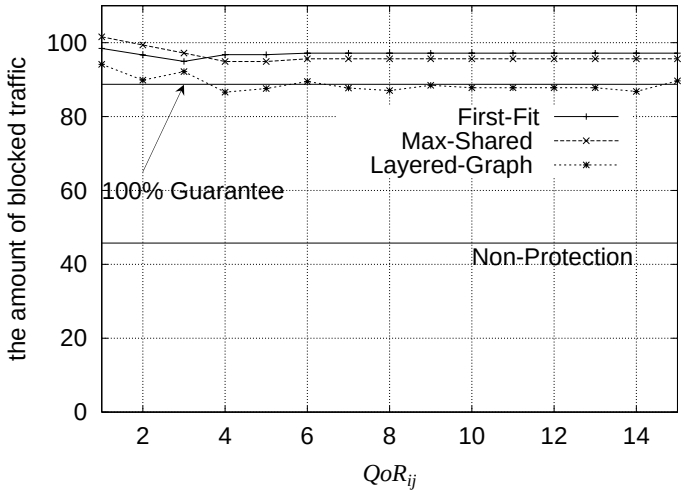


Figure 4.12: Amount of blocked traffic in NSFNET (W = 20)

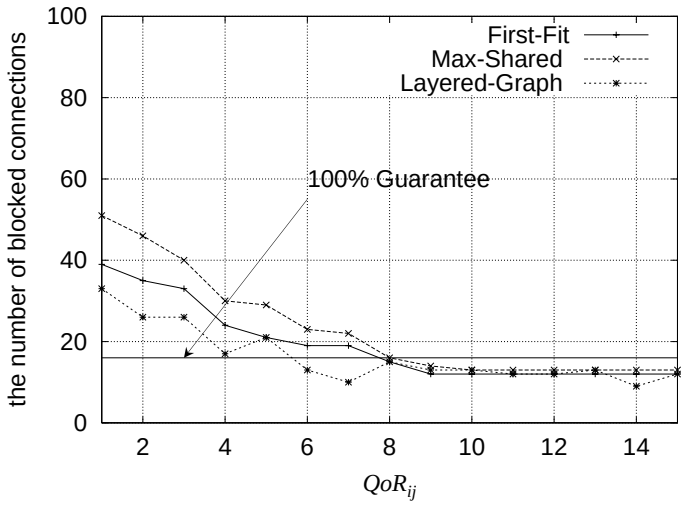


Figure 4.13: Number of blocked connections in NSFNET (W = 50)

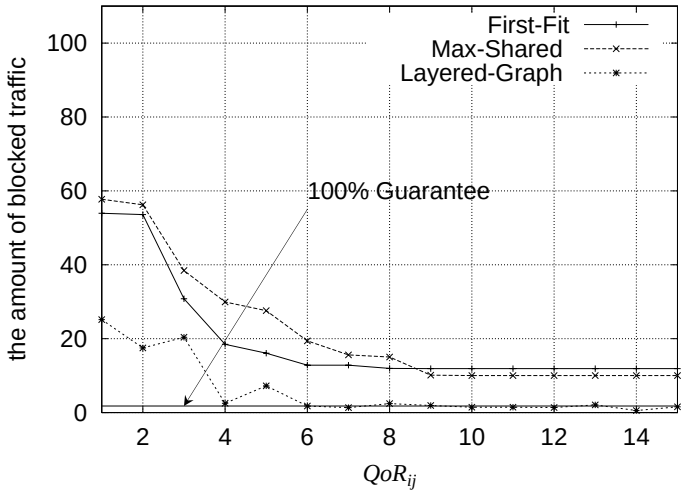


Figure 4.14: Amount of blocked traffic in NSFNET (W = 50)

4.5 Conclusion

In this chapter, we introduced QoR, a concept related to QoS that reflects reliability in a WDM network. QoR can be used to guarantee a maximum recovery time set according to the user's request and provide a 100% guarantee that backup lightpaths will be available. By extending QoR, we can specify a QoR for each node pair ij as QoR_{ij}. We introduced two heuristic algorithms based on QoR_{ij} that can be used to design a logical topology with a protection method that satisfies QoR_{ij} requirements. The objective of the algorithms is to minimize the number of wavelengths needed to carry the overall traffic and provide fault tolerance within QoR requirements. Numerical results showed that the algorithm, which is based on a layered graph, enables more efficient use of wavelength resources than is possible with the other algorithms, especially as the requested traffic volume grows. The algorithm also allows more connections to be carried when using a limited number of wavelengths.

Chapter 5

Incremental Lightpath Management for IP over WDM Networks

In the studies described in the previous chapters, we assumed that traffic demand is known a priori. Such an assumption is, however, inappropriate when WDM technology is applied to the Internet. A more flexible network provisioning approach is necessary for the Internet. In this chapter, we propose a new approach called "incremental capacity dimensioning" for dimensioning the capacity of reliable IP over WDM networks. Our incremental approach consists of three steps for building the logical topology: an initial phase, an incremental phase, and a readjustment phase. With our approach, the logical topology can be adjusted according to incrementally changing traffic demand. During the incremental phase, primary paths are added as traffic increases. At the same time, the backup lightpaths are reconfigured since they do not affect the traffic carried on the operating primary paths. Our algorithm, called MRB (Minimum Reconfiguring for Backup lightpath), assigns the wavelength route in such a way that the number of backup lightpaths to be reconfigured is minimized. Our results show that the total traffic volume which the IP over WDM network can accommodate is increased by using our MRB algorithm. We also introduce another QoR implementation within our three–step approach and explain how our optimization formulation supports the QoR.

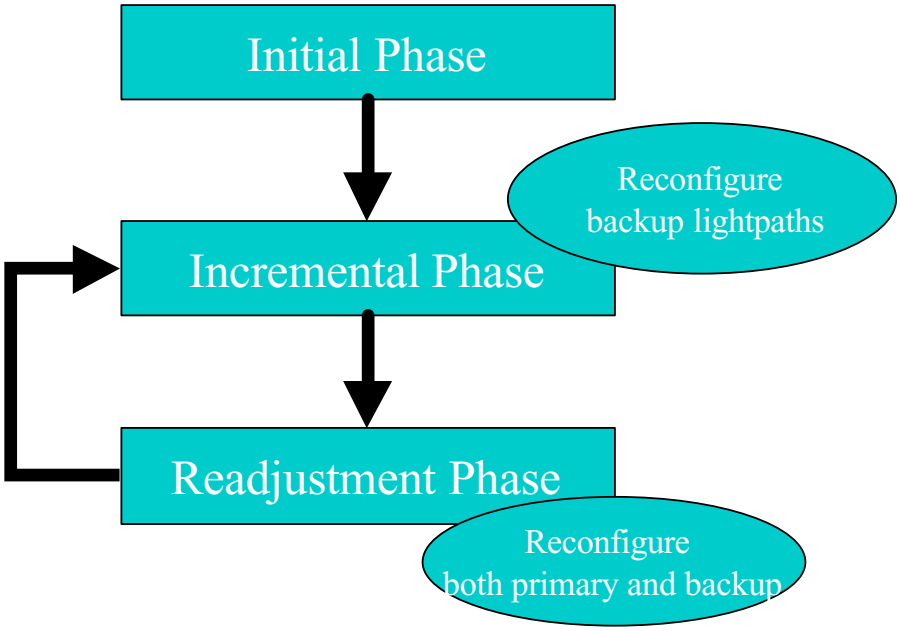


Figure 5.1: Three–step approach to reconfiguring logical topology of reliable IP over WDM network

5.1 Managing Logical Topology for Reliable IP over WDM Networks

In this section, we explain our incremental approach to capacity dimensioning of reliable IP over WDM networks [29]. It consists of initial, incremental, and readjustment phases, which will be described in the following subsections in turn. In each phase, if a sufficient number of lightpaths cannot be set up due to a lack of wavelengths, alert signals are generated so that the network provider can increase the number of fibers to meet the increasing traffic demand.

5.1.1 Initial Phase

In the initial phase, primary and backup lightpaths are set up for given traffic demands. Our approach allows for the likelihood that the projected traffic demands are incorrect. The lightpaths are adjusted in the incremental phase.

Existing methods for designing the logical topology can be used in this phase. They include the method for designing the logical topology for primary lightpaths described in [6], and the heuristic algorithm for setting up backup lightpaths for IP over WDM networks. In this

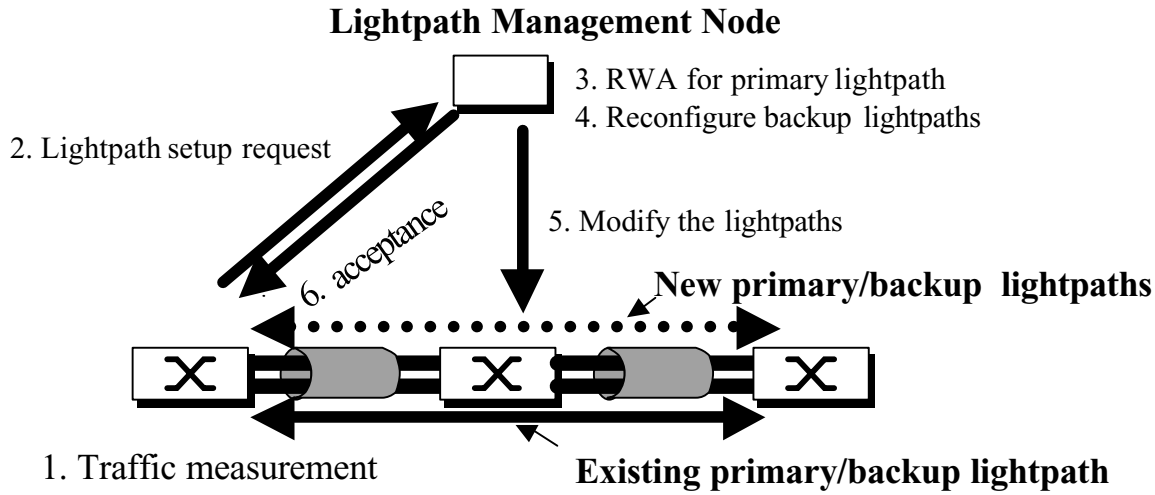


Figure 5.2: Logical topology management model used in incremental phase

phase, the number of wavelengths used for setting up the lightpaths should be minimized so that wavelengths remain for handling increased traffic volume in the incremental phase.

5.1.2 Incremental Phase

The logical topology established in the initial phase must be changed as the patterns of traffic change. This is done in the incremental phase. Our logical topology management model is illustrated in Fig. 5.2. In this model, traffic measurement is mandatory. One way to measure it is to monitor lightpath utilization at the originating node. If it exceeds some threshold C_{th} $(0 < C_{th} < 1)$, the node requests the lightpath management node (LMN), a special node for managing the logical topology of a WDM network, to set up a new lightpath. This is a simplest form of a measurement–based approach. However, this approach is insufficient for a data network; we need an active measurement approach to meet the user–oriented QoS requirement.

In our model, we assume that the LMN eventually knows the actual traffic demand through traffic measurement. It then solves the routing and wavelength assignment problem for both the primary and backup lightpaths. A message to set up a new lightpath is returned to the originating node, and the result is reflected in the WDM network.

As lightpath setup requests are generated, the number of wavelengths available decreases, eventually leading to blocking. To minimize the possibility of blocking, the backup lightpaths

are reconfigured for more effective use of the wavelengths. Only the backup lightpaths are reconfigured because they do not carry traffic unless a failure occurs. The primary lightpaths are not changed in this phase so the active traffic flows are not affected by the lightpath reconfiguration. In this phase, an algorithm is needed for assigning a routing and wavelengths for the new primary lightpaths and one for reconfiguring the backup lightpaths. They will be described in Sections 5.2.1 and 5.2.2 in detail.

5.1.3 Readjustment Phase

In the readjustment phase, inefficient usage of wavelengths, which is caused by the dynamic and incremental wavelength assignment in the incremental phase, is resolved. To improve wavelength usage, all the lightpaths, including the primary ones, are reconfigured. A static design method can be used to do this. Unlike in the initial phase, however, the primary lightpaths are already transporting traffic. The effect of reconfiguration on service interruption should thus be minimized, even if the resulting logical topology is a semi–optimal solution. This is because a global optimal solution will likely require rearranging most of the lightpaths within the network. Thus, the new logical topology should be configured step by step from the old one. One promising method for doing this is the branch–exchange method proposed by Labourdette et al. [47].

Another important issue in this readjustment phase is *when to reconfigure* the logical topology. A straightforward approach is to do it when an alert signal is generated. (An alert signal means a lightpath cannot be set up due to the lack of wavelengths.) The logical topology is reconfigured so as to minimize the number of wavelengths used enabling the lightpath to be accommodated. References [48, 49] gave a reconfiguration policy for this issue, but they only address the primary lightpaths. Further study is needed to include the rearrangement of the backup lightpaths.

5.2 Incremental Capacity Dimensioning

As we described in Section 5.1, the LMN solves the routing and wavelength assignment problem for each new primary lightpath and an optimization problem for reconfiguring the set of

backup lightpaths. We will now describe these in more detail.

5.2.1 Routing and Wavelength Assignment for Primary Lightpath

For each new lightpath setup request, the LMN first solves the routing and wavelength assignment problem for the primary lightpath. The primary lightpath is selected from among the free wavelengths and the wavelengths being used for backup.

If there is a lightpath having the same source–destination pair as the new lightpath, the new lightpath is set up along the same route as the existing lightpath. This is because in IP over WDM networks, the IP layer recognizes that paths on different routes are viewed as having different delays. Hence, the IP layer selects the path with the lower delay, and there is no effect of having multiple lightpaths between source–destination pairs. In some cases, route fluctuation may occur between multiple routes. If no existing lightpath has the same source–destination pair, the new lightpath is set up along the shortest route.

With our minimum reconfiguring for backup lightpath (MRB) algorithm, wavelengths are selected such that the number of backup lightpaths to be reconfigured is minimized. By minimizing the number of backup lightpaths to be reconfigured, we minimize the amount of change to the optimal logical topology obtained in the initial or readjustment phase. Note that actual wavelength assignment is done only after the backup lightpaths are successfully reconfigured (see algorithm below). If there is no available wavelength, an alert signal is generated. More specifically, our algorithm is works as follows.

MRB algorithm

Step 1 For each wavelength k, set $\phi_k = \{\ \}$.

Step 2 Determine the number of backup lightpaths along the route of the requested primary lightpath, P_{new}, that must be reconfigured. For each wavelength k, do Step 3.

Step 3 For each link pq along the route of P_{new}, check whether wavelength k is currently being used. If it is being used by a primary lightpath, set $\phi_k \leftarrow \infty$ and return to Step 2. If it is being used by a backup lightpath (P_{old}), set $\phi_k = \phi \cup P_{old}$. After all the wavelengths have been checked, return to Step 2 and examine the next wavelength. Otherwise, go to Step 4.

Step 4 Select wavelength k' such that the number of elements of $\phi_{k'}$ is minimal.

When multiple lightpaths are necessary between the source–destination pair, lightpaths cannot be set up along different routes. Multiple lightpaths with different routes are prohibited because the IP routing may not choose those paths. That is, IP routing puts all packets onto the primary lightpath with the shortest delay. Multiple lightpaths with different routes can be avoided by using an explicit routing in MPLS [50], and the traffic between the source–destination pair can be divided between the multiple primary lightpaths by explicitly determining the lightpath to use via labels [51]. In this case, our algorithm can be extended so that if there is no available wavelength along the shortest path, the next shortest route is checked for possible wavelength assignment.

5.2.2 Optimization Formulation for Reconfiguring Backup Lightpaths

If a wavelength currently allocated for backup is selected for a new primary wavelength, the backup lightpaths must be reconfigured within the logical topology. Here we describe an optimization formulation that minimizes the number of wavelengths used for backup lightpaths. By doing this, the possibility of the next arriving lightpath setup requests being blocked is minimized. A shared protection scheme is used to improve the use of wavelengths [9]. Before formulating the optimization problem, we summarize the notations used to characterize the physical WDM network.

N: number of nodes in physical WDM network

W: number of wavelengths per fiber

P_{mn}: physical topology defined by set $\{P_{mn}\}$. If there is a fiber connecting nodes m and n, $P_{mn} = 1$, otherwise $P_{mn} = 0$.

C_{mn}: cost between node m and n. Here, we use the propagation delay.

We next introduce the parameters used to represent the route and wavelengths of primary lightpaths where its backup lightpaths are to be reconfigured.

P_{ij}^k: If a backup lightpath for a primary lightpath between node i and node j using wavelength k must be reconfigured, $P_{ij}^k = 1$, otherwise $P_{ij}^k = 0$. P_{ij}^k is determined using the our MRB algorithm.

R_{ij}^k: route of lightpath from node i to node j using wavelength k. It consists of a set of physical links: (i, m_1), (m_1, m_2), ..., (m_p, j).

o_{nm}^w: If the primary lightpath uses wavelength k on physical link mn, $o_{mn}^k = 1$, otherwise $o_{mn}^k = 0$. o_{mn}^k is determined from R_{ij}^k.

A_{ij}^k: set of routes of backup lightpaths for primary lightpath from node i to node j using wavelength k. It consists of a set of physical links: (i, n_1), (n_1, n_2), ..., (n_q, j).

φ_{nm}: maximum number of backup lightpaths on physical link mn. It is determined from A_{ij}^k.

We use the following variables to formulate the optimization problem.

b_{nm}: number of backup lightpaths placed on physical link mn.

m_{nm}^w: If the backup lightpath uses wavelength w on physical link mn, $m_{mn}^w = 1$, otherwise $m_{mn}^w = 0$.

$g_{ij,pq,k}^{mn,w,r}$: If the lightpath originating at node i and terminating at node j uses wavelength k for the primary lightpath on physical link pq and wavelength w between nodes m and n as a backup lightpath on the $r-$th alternate route, $g_{ij,pq,k}^{mn,w,r} = 1$, otherwise $g_{ij,pq,k}^{mn,w,r} = 0$.

We can now formulate the optimization problem.

Objective function

Minimize number of wavelengths used for backup lightpaths:

$$\min \sum_{mn} b_{mn}. \tag{5.1}$$

Constraints

1. The number of backup lightpaths placed on physical link mn must equal the sum of the number of wavelengths used on that link for the backup lightpaths:

$$b_{mn} = \sum_{w \in W} m_{mn}^{w}. \tag{5.2}$$

2. Either a primary lightpath or a backup lightpath must use wavelength k on physical link mn if there is a fiber.

$$o_{mn}^{k} + m_{mn}^{k} \leq P_{mn} \tag{5.3}$$

3. The lightpath using wavelength k between nodes i and j must be protected by a backup lightpath when physical link $pq \in R_{ij}^{k}$ fails. That is, if $P_{ij}^{k} = 1$,

$$\sum_{w \in W} \sum_{r \in A_{ij}^{k}} \sum_{it \in r} g_{ij,pq,k}^{it,w,r} = 1. \tag{5.4}$$

   Note that it is unnecessary to use the same wavelength for the primary and corresponding backup lightpaths.

4. The lightpath using wavelength k between nodes i and j must use wavelength w on all links of the backup lightpath ($r \in A_{ij}^{k}$) when a link between node p and node q fails. Namely, if $P_{ij}^{k} = 1$,

$$g_{ij,pq,k}^{nt,w,r} = g_{ij,pq,k}^{tm,w,r}, \qquad \forall pq \in R_{ij}^{k}, \forall nt, tm \in r, \forall r \in A_{ij}^{k}. \tag{5.5}$$

   This is called the "wavelength continuity constraint".

5. The lightpath using wavelength k between nodes i and j must use wavelength w for the backup lightpath. This means, for each fiber–failure scenario along the lightpath using wavelength k between nodes i and j, the same wavelength, w, is utilized. That is, if $P_{ij}^{k} = 1$,

$$g_{ij,p_1q_1,k}^{pq,w,r} = g_{ij,p_2q_2,k}^{pq,w,r}, \qquad \forall p_1q_1, p_2q_2 \in R_{ij}^{k}. \tag{5.6}$$

   As this equation indicates, we allow the use of different wavelengths for the backup path

against the failure of the corresponding primary path.

6. When physical link pq fails, at most one backup lightpath can use wavelength w on physical link mn if the corresponding primary lightpath traverses failed link pq.

$$\sum_{ij} \sum_{k \in W: pq \in R_{ij}^k} \sum_{r \in A_{ij}^k: mn \in r} \sum_{mn \in r} g_{ij,pq,k}^{mn,w,r} \leq 1 \tag{5.7}$$

7. The number of backup lightpaths using wavelength k on physical link mn must be bounded.

$$\varphi_{mn} \times m_{mn}^w \geq \sum_{k \in W} \sum_{ij} \sum_{r \in A_{ij}^k: mn \in r} \sum_{pq \in R_{ij}^k} g_{ij,pq,w}^{mn,k,r} \tag{5.8}$$

8. For two primary lightpaths between nodes i and j using wavelengths k and k', the cost of the corresponding backup lightpaths must be the same along routes $r(\in A_{ij}^k)$ and $r'(\in A_{ij}^{k'})$. That is, if $P_{ij}^k = 1 \wedge P_{ij}^{k'} = 1 \wedge r \equiv r'$,

$$\sum_{w} \sum_{mn \in r} C_{mn} \times g_{ij,pq,k}^{mn,w,r} = \sum_{w'} \sum_{m'n' \in r'} C_{m'n'} \times g_{ij,pq,k'}^{m'n',w',r'}. \tag{5.9}$$

Note that in Eqs. (5.7) and (5.8), we do not impose the condition $P_{ij}^k = 1$. This is because wavelength sharing is allowed only if the corresponding primary lightpaths are link–disjoint.

When we set up multiple backup lightpaths between originating node i and terminating node j, we should set them up along the same route for the same reason multiple primary lightpaths are set up along the same route. Equation (5.9) defines this constraint. If the option of explicit routing in MPLS [50] is used, the constraint can be eliminated.

5.2.3 Evaluation

To evaluate our proposed algorithm, we simulated the incremental phase. using a network consisting of 14 nodes and 21 links as the physical topology (see Fig. 2.2). The number of wavelengths in each fiber, W, was 50. As an initial condition, one primary lightpath was allocated for each node–pair, which simulated the initial phase of our approach. The traffic rate given in [6] was used for reference purposes. The primary lightpaths were set up on the shortest route, i.e., the path along which the propagation delay was the smallest. The wavelengths of

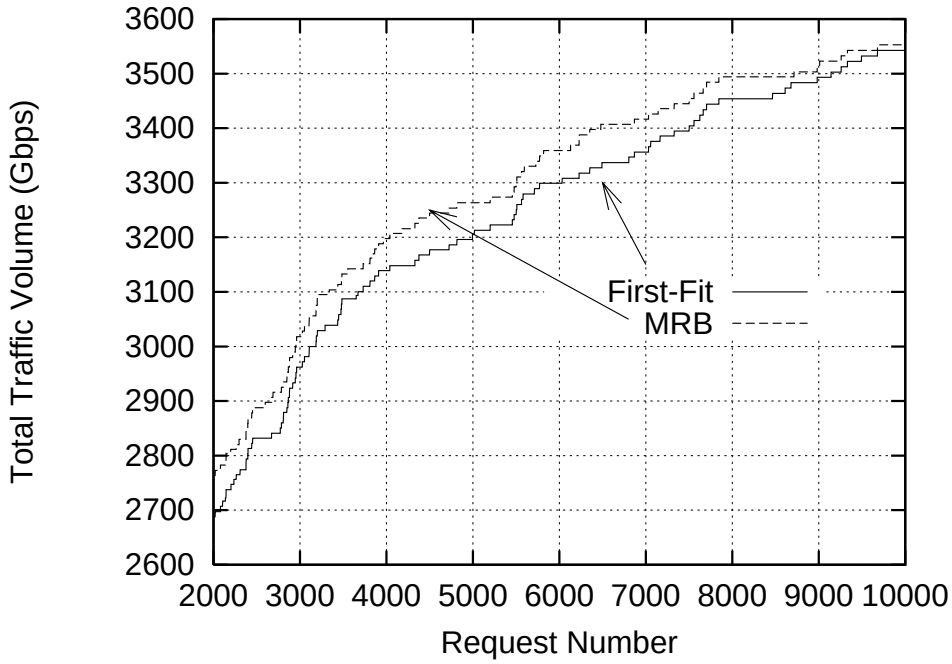


Figure 5.3: Total traffic volume with first–fit and MRB algorithms

the primary lightpaths were determined based on the first–fit policy [35]. The wavelengths of the backup lightpaths were determined by using the min–hop–first algorithm, which assigns the wavelengths in descending order of the hop–count of the primary lightpaths.

In our proposed framework, each node measures the traffic volume, and if the utilization of the primary lightpath exceeds the threshold value, a lightpath setup request is generated. However, in our simulation, we did not consider such a scenario. Instead, we simply considered that during the incremental phase, requests to set up new lightpaths arrived randomly at the node pairs. The volume of traffic demand was randomly set between 0 and C (Gbps), where C represents the wavelength capacity. In our simulation, C was 10 Gbps.

For each lightpath setup request, we used the MRB algorithm and solved the optimization formulation described in Section 5.2.2 using the CPLEX optimizer. We generated $10,000$ lightpath setup requests, and for each request, the node checked whether the utilization of the primary lightpath exceeded 80% of the lightpath capacity (i.e., C_{th} = 0.8). If it did, the node generated a lightpath setup request. The wavelength of the new primary lightpath was determined using our MRB algorithm, and the optimization problem was solved to reconfigure the backup lightpaths if necessary. We counted the number of blocked requests as a performance measure. For comparison purposes, we also considered the first–fit approach for establishing

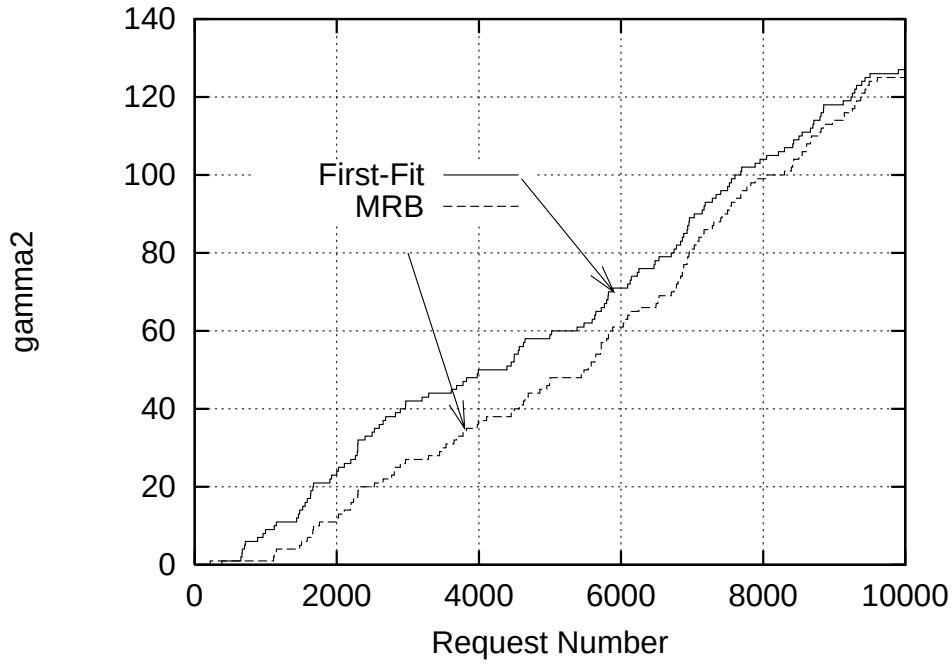


Figure 5.4: Number of lightpath setup requests rejected because backup lightpaths could not be reconfigured

the new lightpath. In the first–fit approach, the wavelengths from λ_1 to λ_W are checked in a sequential for the new primary lightpath. If an available wavelength is found (say, λ_m), then the new primary lightpath is set up using λ_m.

We compared the total traffic volume with the number of requests. The volume did not increase when a lightpath setup request was blocked due to the lack of available wavelengths. As shown in Fig. 5.3, the results of MRB algorithm is slightly better than that of the first–fit approach.

We also compared the number of lightpath setup requests rejected because backup lightpaths could not be reconfigured. We denote the number rejected by γ_2. Recall that the primary lightpath setup request is rejected (1) if the primary lightpath cannot be set up due to the lack of a wavelength (γ_1) or (2) if the backup lightpath cannot be reconfigured (i.e., γ_2). A lower value of γ_2 means more requests for primary lightpaths can be accepted by reconfiguring backup lightpaths. Figure 5.4 shows that using the MRB algorithm reduces γ_2 and improve the usage of the wavelengths.

5.2.4 Distributed Approaches

So far, we have considered a centralized approach to establishing the logical topology. In general, the centralized approach has a scalability problem, especially when the number of wavelengths and/or the network is large. Our main purpose is to propose a framework for the incremental use of wavelengths in IP over WDM networks. We can thus replace the centralized approach with a distributed approach in our framework.

Anand and Qiao proposed a heuristic algorithm for setting up primary and backup lightpaths on demand [35]: routes and wavelengths are assigned for each lightpath setup request. Backup lightpaths can be reconfigured to meet future lightpaths setup requests, so wavelengths are used more effectively. However, only dedicated protection is considered, so more wavelengths are needed. As described in Chapter 4, a shared protection scheme is more appropriate in IP over WDM networks since IP routing can also protect against failure. A distributed algorithm for shared protection scheme is considered by Yuan [41].

Mohan et al. considered a restoration method [52]. They call a connection request with a reliability requirement a *D–connection* (dependable connection). They divided methods for establishing connections into reactive and pro–active. In the reactive methods, if an existing lightpath fails, a search is initiated to find a lightpath that does not use the failed components. In the pro–active methods, backup lightpaths are identified and resources are reserved along the backup lightpaths. The backup lightpaths are set up when primary lightpath is established.

5.2.5 Quality of Reliability Issue

Quality of reliability, or Quality of protection (QoP), is one aspect of quality of service (QoS) that is suitable for reliable IP over WDM networks. The implementation of QoP has been considered by several research groups [23–25, 53]. One suggested way to provide QoP is to split each primary lightpath into several segments [23, 53]. Doing this enables quick handling of the failure signals sent to the originating node on the primary lightpath. Saradhi and Murthy introduced the concept of an *R–connection* [23] for dynamical establishment of a reliable connection. The basic idea of the R–connection is that an application user specifies the level of reliability for a connection and the reliability levels of the connection are calculated based on a pre–specified reliability measurement for each network component. If the reliability require-

ment is not satisfied, the length of the primary lightpath covered by a partial backup lightpath is selected so as to increase the reliability of the R–connection. Another way to provide QoP is to use the differentiated reliability (DiR) of a connection [24, 25]: the maximum probability that the connection will fail due to a single network component failing. With this approach, a continuous spectrum of reliability levels is provided.

Here, we describe another QoP implementation within our three–step approach and explain how our optimization formulation differs to support QoP. We introduce three QoS classes with respect to reliability.

Class 1. Provide both primary and backup lightpaths in the incremental phase if wavelengths are available.

Class 2. Provide a backup path, but it can be taken by a primary lightpath , which belongs to QoS class 1, if a wavelength is not available.

Class 3. Provide only primary lightpaths; no protection mechanism is provided.

These QoS classes can easily be provided by modifying the logical topology design algorithm. We introduce the following notation.

QoP_{ij}: If backup lightpaths must be provided between nodes i and j in the incremental phase, $QoP_{ij} = 1$, otherwise $QoP_{ij} = 0$.

In the incremental phase, QoP classes 2 and 3 are treated the same. Thus, we simply set QoP_{ij} to 0 for both classes. To provide both primary and backup lightpaths in the incremental phase, we change Eq. (5.4):

$$QoP_{ij} = \sum_{w \in W} \sum_{r \in A_{ij}^k} \sum_{it \in r} g_{ij,pq,k}^{it,w,r}. \tag{5.10}$$

If $QoP_{ij} = 0$, $g_{ij,pq,k}^{it,w,r}$ is also set to 0, and backup lightpaths for QoP classes 1 and 2 can be provided.

5.3 Conclusion

In this chapter, we proposed an incremental use of the wavelengths , called “incremental capacity dimensioning”, in reliable IP over WDM networks. It provides a network structure flexible against traffic change. Three phases (initial, incremental, and readjustment) are introduced for this purpose. In the incremental phase, only the backup lightpaths are reconfigured for an effective use of wavelengths. In the readjustment phase, on the other hand, both the primary and backup lightpaths are reconfigured, since the incremental setup of the primary lightpaths tends to utilize the wavelengths ineffectively. In the readjustment phase, a one-by-one readjustment of the established lightpaths toward a new logical topology is performed so that we can achieve service continuity of the IP over WDM network. The branch–exchange method may be used to do this. Improving the algorithm to minimize the number of the one-by-one readjustment operations is left for future work.

Chapter 6

Conclusion

In this thesis, we have proposed methods for designing IP over WDM network to integrate the routing function of IP and the networking capability of WDM. When the WDM technology is applied to IP, a packet route is determined by the routing protocol provided by the IP layer, and thus the end-to-end path provided by the logical topology of the WDM network is not suitable for IP since IP has its own metrics for route selection. WDM network can also provide a reliability function to the IP layer.

We first focused on route selection on IP. In Chapter 2, we proposed a new heuristic algorithm, SHLDA, for designing a logical topology based on the delays between nodes as an objective metric. Comparison of the proposed algorithm with conventional ones in terms of the average packet delay and throughput showed that it is effective when the number of wavelengths is low and the processing capacity of the IP routers is large. Evaluation of SHLDA from the viewpoint of routing stability showed that SHLDA improves the maximum throughput, compared with a conventional algorithm, without sacrificing routing stability.

In Chapter 3, we first formulated the shared path protection mechanisms as an MILP optimization problem, assuming a single-fiber failure in the network. Because it becomes computationally intensive as the network grows, we proposed heuristic algorithms and compared its solutions with that obtained by MILP. Through numerical examples, we compared the number of wavelengths necessary for reliable network. We next considered the functional partitioning of IP routing and WDM protection. Based on our heuristic algorithm, we also discussed the effect of interaction between the IP and WDM layers. We showed that the largest-traffic-first

approach is best if our primary concern is traffic load at the IP routers after a failure.

In Chapter 4, we introduced QoR, a concept related to QoS that reflects reliability in a WDM network. QoR can be used to guarantee a maximum recovery time set, based on the user's request and guarantee that backup lightpaths will be available. By extending QoR, we can specify a QoR for each node pair ij as QoR_{ij}. We described a heuristic algorithm based on QoR_{ij} that can be used to design a logical topology with a protection method that satisfies QoR_{ij} requirements. The objective of this algorithm is to minimize the number of wavelengths needed to carry the overall traffic and provide fault tolerance within QoR requirements. Numerical results showed that the algorithm, which is based on a layered graph, enables more efficient use of wavelength resources than is possible with other algorithms, especially as the requested traffic volume grows. The algorithm also allows more connections to be carried when using a limited number of wavelengths.

We next proposed a framework for an incremental use of the wavelengths in reliable IP over WDM networks in Chapter 5. Our incremental approach consists of three steps for building the logical topology: an initial phase, an incremental phase, and a readjustment phase. With our approach, the logical topology can be adjusted according to incrementally changing traffic demand. During the incremental phase, primary paths are added as traffic increases. At the same time, the backup lightpaths are reconfigured since they do not affect the traffic carried on the operating primary paths. Our proposed algorithm, called MRB (Minimum Reconfiguring for Backup lightpath), assigns the wavelength route in such a way that the number of backup lightpaths to be reconfigured is minimized. Our results showed that the total traffic volume the IP over WDM network can accommodate is increased by using this algorithm. We also describe another QoR implementation within our three–step approach and explained how our optimization formulation supports the QoR.

In the readjustment phase, a one–by–one readjustment of the established lightpaths toward a new logical topology should be performed so that service is not interrupted. we can achieve a service continuity of the IP over WDM networks. The branch–exchange method can be used for this purpose. However, the algorithm must be concerned about the backup lightpaths. This issue is left for future work. As we mentioned in this thesis, networking in the optical domain has the potential to offer several network control functionalities such as functions of routing, congestion control, and reliability. In this thesis, we investigated the functional partitioning of

routing and reliability between the IP and WDM. Our future research topic is to consider how to apply the networking capability of WDM to the function of congestion control.

Bibliography

[1] "MPLS IETF homepage," *http://www.ietf.org/html.charters/mpls-charter.html*.

[2] D. Awduche and Y. Rekhter, "Multi-protocol lambda switching: Combining MPLS traffic engineering control with optical crossconnects," *IEEE Communications Magazine*, vol. 39, pp. 111–116, Mar. 2001.

[3] R. Ramaswami and K. N. Sivarajan, "Routing and wavelength assignment in all–optical networks," *IEEE/ACM Transactions on Networking*, vol. 3, pp. 489–500, Oct. 1995.

[4] R. Dutta and G. N. Rouskas, "A survey of virtual topology design algorithms for wavelength routed optical networks," *Optical Network Magazine*, vol. 1, pp. 73–89, Jan. 2000.

[5] B. Mukherjee, D. Banerjee, S. Ramamurthy, and A. Mukherjee, "Some principles for designing a wide-area WDM optical network," *IEEE/ACM Transactions on Networking*, vol. 4, pp. 684–695, Oct. 1996.

[6] R. Ramaswami and K. N. Sivarajan, "Design of logical topologies for wavelength-routed optical networks," *IEEE Journal on Selected Areas in Communications*, vol. 14, pp. 840–851, June 1996.

[7] G. Ellinas and T. Stern, "Automatic protection switching for link failures in optical networks with bidirectional links," in *Proceedings of GLOBECOM*, 1996.

[8] C. Mas and P. Thiran, "A review on fault location methods and their application to optical networks," *Optical Network Magazine*, vol. 2, pp. 73–87, July/Augest 2001.

[9] S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks part *i* – protection," in *Proceedings of Infocom'99*, pp. 744–751, March 1999.

[10] O. Gerstel and R. Ramaswami, “Optical layer survivability: A services perspective,” *IEEE Communications Magazine*, vol. 15, no. 4, pp. 104–113, 2000.

[11] S. Arakawa, J. Katou, and M. Murata, “A design method of logical topology with stable packet routing in IP over WDM network,” *IEICE General Conference*, Mar. 2001.

[12] J. Katou, S. Arakawa, and M. Murata, “A design method of logical topology and its influence on IP routing in IP over WDM network,” *Technical Report of IEICE (*PNI2000-35*)*, pp. 32–38, Mar. 2001.

[13] J. Katou, S. Arakawa, and M. Murata, “A design method of logical topology for IP over WDM networks with stable routing,” in *Proceedings of The Fifth Working Conference on Optical Network Design and Modeling (ONDM 2001)*, pp. 61–78, Feb. 2001.

[14] S. Arakawa, J. Katou, and M. Murata, “A design method for logical topologies with stable packet routing in IP over WDM networks,” submitted to *IEICE Transactions on Communications*.

[15] S. Arakawa, M. Murata, and H. Miyahara, “Design of lightpath networks with protections for IP over WDM networks,” *Technical Report of IEICE (*SSE99-111*)*, pp. 7–12, Nov. 1999.

[16] S. Arakawa, M. Murata, and H. Miyahara, “Functional partitioning for multi-layered survivability in IP over WDM networks,” *Technical Report of IEICE (*PNI2000-3*)*, pp. 16–25, May 2000.

[17] S. Arakawa, M. Murata, and H. Miyahara, “Design methods of multi-layer survivability in IP over WDM networks,” in *Proceedings of OptiComm*, pp. 279–290, Oct. 2000.

[18] S. Arakawa, M. Murata, and H. Miyahara, “Functional partitioning for multi-layer survivability in IP over WDM networks,” *IEICE Transactions on Communications*, vol. E83-B, pp. 2224–2233, Oct. 2000.

[19] S. Arakawa and M. Murata, *Reliability Issues in IP over Photonic Networks*. Quality, Survivability and Reliability of Large Scale Systems – Case Studies: Olympic Games, John Wiley & Son, Dec. 2002. (Chapter in Book).

[20] J. Katou, S. Arakawa, and M. Murata, “Design method of logical topologies in WDM network with quality of protection,” *Technical Report of IEICE (*NS2001-212*)*, pp. 41–46, Feb. 2001.

[21] J. Katou, S. Arakawa, and M. Murata, “Design method of logical topology in WDM network with quality of protection,” in *Proceedings of Workshop on Optical Networking: Technologies, Architectures and Management*, Nov. 2001.

[22] S. Arakawa, J. Katou, and M. Murata, “Design method of logical topologies with quality of reliability in WDM networks,” to appear in *Photonic Network Communications*.

[23] C. V. Saradhi and C. S. R. Murthy, “Routing differentiated reliable connections in single and multi–fiber WDM optical networks,” in *Proceedings of Opticomm*, pp. 24–35, Aug. 2001.

[24] O. Gerstel and G. Sasaki, “Quality of protection (QoP): A quantitative unifying paradigm to protection service grades,” in *Proceedings of Opticomm*, pp. 12–23, Aug. 2001.

[25] A. Fumagalli and M. Taaca, “Differentiated reliability (DiR) in WDM rings without wavelength converters,” in *Proceedings of ICC*, June 2001.

[26] G. Mohan and A. K. Somani, “Routing dependable connections with specified failure restoration guarantees in WDM networks,” in *Proceedings of IEEE INFOCOM 2000*, pp. 1761–1770, March 2000.

[27] S. Arakawa and M. Murata, “Incremental capacity dimensioning for reliable IP over WDM networks,” *Technical Report of IEICE (*PNI2000-34*)*, pp. 24–31, Mar. 2001.

[28] S. Arakawa, S. Ishida, and M. Murata, “Management of logical topologies for dynamically changing traffic in reliable IP over WDM networks,” *Technical Report of IEICE (*DC2002-2*)*, pp. 7–14, Apr. 2002.

[29] S. Arakawa and M. Murata, “On incremental capacity dimensioning in reliable IP over WDM networks,” in *Proceedings of OPTICOMM*, pp. 153–163, Aug. 2001.

[30] S. Arakawa and M. Murata, “Lightpath management of logical topology with incremental traffic changes for reliable IP over WDM networks,” *Optical Network Magazine*, vol. 3, pp. 68–76, May 2002.

[31] M. Murata, “Challenges for the next–generation Internet and the role of IP over photonic networks,” *IEICE Transactions on Communications*, vol. E83-B, pp. 2153–2165, Oct. 2000.

[32] R. Ramaswami and K. N. Sivarajan, “Design of logical topologies for wavelength-routed all-optical networks,” in *Proceedings of IEEE INFOCOM ’95*, pp. 1316–1325, Apr. 1995.

[33] L. Fratta, M. Gerla, and L. Kleinrock, “The flow deviation method: An approach to store-and-forward communication network design,” *Networks*, vol. 3, pp. 97–133, 1973.

[34] P.-H. Ho and H. T. Mouftah, “A framework of a survivable optical Internet using short leap shared protection (SLSP),” in *Proceedings of 2001 IEEE Workshop on High Performance Switching and Routing*, pp. 21–25, May 2001.

[35] V. Anand and C. Qiao, “Dynamic establishment of protection paths in WDM networks, part i,” in *Proceedings of the 9th IEEE International Conference on Computer Communications and Networks (IC3N 2000)*, Oct. 2000.

[36] M. Kodialam and T. V. Lakshman, “Dynamic routing of locally restorable bandwidth guaranteed tunnels using aggregated link usage information,” in *Proceedings of IEEE INFOCOM 2000*, pp. 902–911, March 2000.

[37] H. Zang and B. Mukherjee, “Connection management for survivable wavelength–routed WDM mesh networks,” *Optical Network Magazine*, vol. 2, pp. 17–28, July/Augest 2001.

[38] E. Modiano and A. Narula, “Survivable routing of logical topologies in WDM networks,” in *Proceedings of IEEE INFOCOM*, Apr. 2001.

[39] M. Kodialam and T. Lakshman, “Dynamic routing of bandwidth guaranteed tunnels with restoration,” in *Proceedings of IEEE INFOCOM*, Apr. 2000.

[40] B. Doshi *et al.*, “Optical network design and restoration,” *Bell Labs Technical Journal*, vol. 4, pp. 58–84, January–March 1999.

[41] S. Yuan, "A heuristic routing algorithm for shared protection in connection–oriented networks," in *Proceedings of Opticomm*, pp. 142–152, Aug. 2001.

[42] M. Sridharan and A. K. Somani, "Revenue maximization in survivable WDM networks," in *Proceedings of Opticomm*, pp. 291–302, Oct. 2000.

[43] "CPLEX homepage," *http://www.cplex.com*.

[44] "NTT Information Web Station," available at `http://www.ntt-east.co.jp/info-st/network/traffic/index.html`. (in Japanese).

[45] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," *Optical Network Magazine*, pp. 47–60, January 2000.

[46] V. Anand and C. Qiao, "Static versus dynamic establishment of protection paths in WDM networks," in *Proceedings of IEEE INFOCOM 2001*, April 2001.

[47] J-F. P. Labourdette, F. W. Hart, and A. S. Acampora, "Branch-exchange sequences for reconfiguration of lightwave networks," *IEEE Transactions on Communications*, vol. 42, pp. 2822–2832, Oct. 1994.

[48] I. Baldine and G. N. Rouskas, "Traffic adaptive WDM networks: A study of reconfiguration issues," *IEEE Journal of Lightwave Technology*, vol. 19, no. 4, pp. 433–455, 2001.

[49] I. Baldine and G. N. Rouskas, "Dynamic reconfiguration policies in multihop WDM networks," *Journal of High Speed Networks*, vol. 4, no. 3, pp. 221–238, 1995.

[50] D. O. Awduche, "MPLS and traffic engineering in IP networks," *IEEE Communications*, pp. 42–47, Dec. 1999.

[51] D. O. Awduche, Y. Rekhter, J. Drake, and R. Coltun, "Multi-protocol lambda switching: Combining MPLS traffic engineering control with optical crossconnects," *IETF Internet Draft*. draft-awduche-mpls-te-optical-02.txt.

[52] G. Mohan, C. Murthy, and A. Somani, "Efficient algorithms for routing dependable connections in WDM optical networks," *IEEE/ACM Transactions on Networking*, vol. 9, Oct. 2001.

[53] P.-H. Ho and H. Mouftah, “A framework of a survivable optical internet using short leap shared protection (SLSP),” in *Proceedings of IEEE Workshop on High Performance Switching and Routing*, May 2001.