Contents

List of Figures

Chapter 1

Reliability Issues in IP over Photonic Networks

Shin'ichi Arakawa and Masayuki Murata, Osaka University, Japan

1.1 Need for Improved Reliability in IP over Photonic Networks

The rapid growth in the number of users and in the number of multimedia applications on the Internet is dramatically increasing traffic volume on backbone networks. Very high speed networks are thus necessary. Moreover, the Internet protocol (IP) is emerging as a dominant technology, so the ability to carry the IP traffic efficiently is an important issue for the next–generation data–centric Internet. Recent advances in optical switches have led to optical technology with

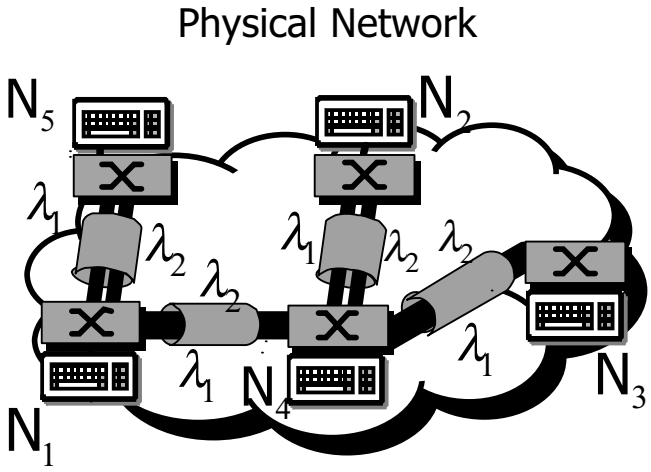


Figure 1.1: Physical WDM network: optical nodes are connected using optical fibers

networking capability. The so–call *photonic network* is a strong candidate for transporting IP traffic, and the integration of IP and optical networking technologies was the topic of a recent special issue [1].

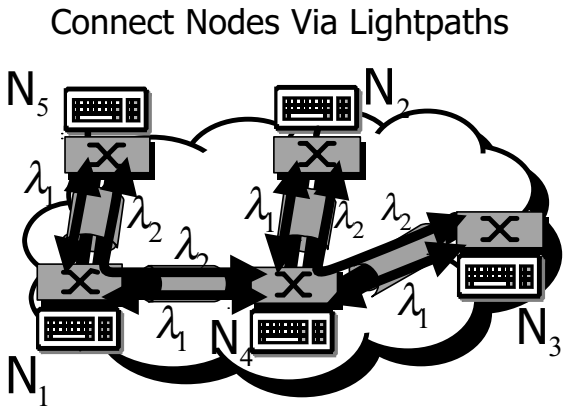
There are several candidate infrastructures for the photonic network. One is optical packet switching with optical code division multiplexing (OCDM) [2, 3]. Another is a wavelength division multiplexing (WDM), which allows multiple wavelengths to be carried on a single fiber. WDM–based photonic networks are a low–cost approach to handling the increased traffic volumes because of recent advances in WDM components, while OCDM technology is still immature in terms of both scalability and signal handling.

Currently, commercially available WDM transmission systems use only WDM technology on it fiber links (see Figure 1.1). Each wavelength in a fiber is treated as a physical link between network components (e.g., routers and switches). This

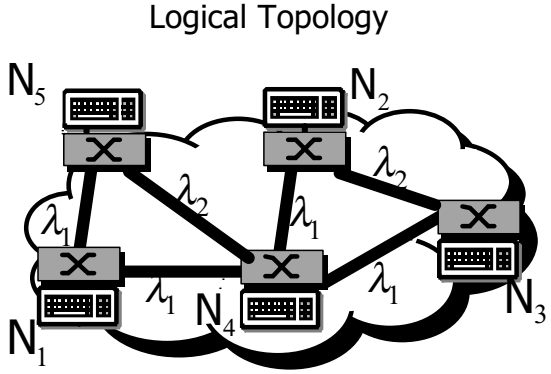
means that the conventional IP technique for handling multiple links can be used. Link capacity is increased by increasing the number of wavelengths on the fiber, which may resolve bandwidth bottlenecks in the link. However, simply resolving link bottlenecks in the face of exploding demands is not enough because it only shifts the bottlenecks to the electronic routers.

One way to alleviate 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 solving the bottlenecks at the electronic routers. The wavelength path (referred to as *lightpath*) provides a direct optical connection between two nodes by using multiple protocol lambda switching (MPλS) or generalized MPLS (GMPLS) technologies. A *logical topology* is constituted by wavelengths on the physical WDM network (see Ref. [4] and references therein). Here, the physical WDM network means an actual network consisting of optical nodes and optical–fiber links connecting nodes, as shown in figure 1.1. The lightpath is established on the physical network (see Figure 1.2(a))The actual traffic of the upper layer protocol is carried on the constructed logical topology (see figure 1.2).

The advances described above will lead to very high capacity networks, which will drive the need for a reliability mechanism embedded in the logical topology.



(a) Lightpath configuration



(b) Logical topology as seen from a upper layer protocol

Figure 1.2: Constructing logical topology by configuring lightpaths

A "reliability mechanism" is a functionality that enables recovery from unexpected failures of network components. Networks 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 the 1:1 protection scheme, and automatic protection switching [5] provides the reliability mechanism. Fiber allocation is sufficient for the 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, or ATM), may pass through successive network components. Thus, coordination of a reliability mechanism for each lightpath from 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). Other upper layer protocols, such as ATM and MPLS, also have a reliability mechanism, but the recovery time is significantly long. In contrast, a new route can be established within a few tens of milliseconds following a failure in WDM networks.

In general, there are two types of reliability mechanisms in WDM networks: *protection schemes* in which network resources are pre–determined and reserved for backup purpose, and *restoration schemes* in which network resources are dynamically computed and allocated only when a failure occurs. These two types of schemes are described in Section 1.3. In both, the network resources are wavelengths. Thus, the reliability mechanism in the optical domain is not always an obvious 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 opti-

cal domain with one in the electronic domain, we can obtain more reliable networks than the current Internet. In this chapter, we investigate methods to improve reliability in IP over photonic networks. Areas beging research include protection/restoration and multi-layer survivability schemes. Method for designing them are also being formulated.

1.2 Network Architectures for IP over Photonic Networks

We first describe the architecture model of a node in the network. As shown in Figure 1.3, a node has optical switches and electronic router. The switches consist of three main blocks: input section, non–blocking optical switches, and output section. In the input section, optical signals are demultiplexed into W fixed wavelengths, $\lambda_1, \ldots \lambda_w$. Then, each wavelength is transferred to an appropriate output port by a switch. Finally, in the output section, each wavelength is multiplexed again and sent to the next node. By configuring the switches along the path, a particular wavelength is carried from an input port to an output port without any electronic processing. If a lightpath terminates at this node, the IP packets on the lightpath are converted to electrical signals and forwarded to the electronic router. If a lightpath begins at this node, IP packets from the electronic router are transmitted over the lightpath after being converted to optical signals. Other structures for the node can be considered, but the above–mentioned node architecture is preferable since there is no need to modify the IP routing mechanism.

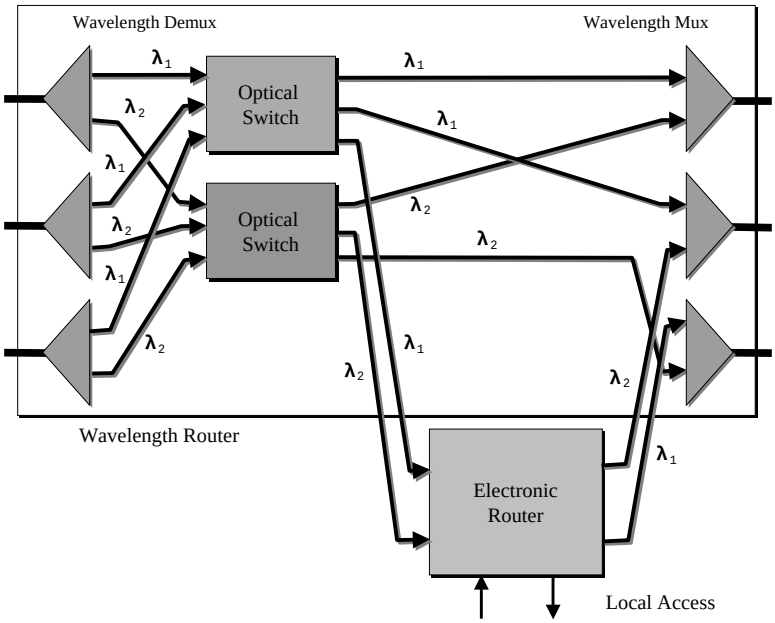


Figure 1.3: Architecture model of optical node

1.2.1 MPLS and GMPLS

In this section, we briefly introduce ATM–based MPLS. Then, after describing the concept of photonic MPLS, we present WDM–based MPLS as an example.

The earliest motivation for MPLS was to simplify wide–area IP backbone networks by overlaying IP and the new emerging high–speed technology. During the mid–90's, the only solution was ATM, in which fixed–size packets (called cells) are switched in hardware at nodes. The main reason for this is that ATM can provide high–speed switching. In IP over ATM networks, ATM is used only for providing the link–level connectivity, although ATM itself had been developed

to offer its native networking capabilities. While there have been many excellent articles explaining ATM–based MPLS (e.g., [6, 7] and references therein), we start by introducing it because λ–MPLS can provide many similar functions in the optical domain. The primary concept of MPLS is to utilize the high speed packet–forwarding capability of the underlying network by using a label–swapping forwarding algorithm. A label is a short, fixed–length value carried in the packet header to identify the forwarding equivalence class (FEC). A label has only link–local significance; it corresponds to VPI/VCI of ATM. In MPLS, packet forwarding is performed as follows.

(1) At the ingress edge of the MPLS, i.e., at the ingress label switched router (LSR), the label–swapping forwarding algorithm maps the label of the destination address of an arriving IP packet to the initial label for injecting the packet onto the label–switched path (LSP). The ingress LSR performs a longest–prefix match routing table lookup to find an appropriate label, as in a conventional router.

(2) An LSP is set up between the ingress and egress LSRs by using a VP/VC connection in ATM. That is, an LSP is functionally equivalent to a virtual circuit.

(3) Within the network, the core LSR forwards the packet using the label–swapping forwarding algorithm. When a labeled packet arrives at the core LSR, it uses the label and the input port number to determine the next–hop output port number and the new label by an exact match search of the forwarding table.

This native function of ATM does not impose the processing burden of the longest–prefix matching of conventional IP routers.

(4) Finally, the egress LSR searches for the next link by performing a longest–prefix match table lookup, similar to that of conventional IP routers. Setting up the appropriate LSPs is another concern with MPLS. It is done by using the label distribution protocol (LDP), which supports two styles of label distribution: independent and ordered [6].

A key concept of MPLS is that multiple flows can be assigned to the same label (or LSP), and a stream can have granularity ranging from fine to coarse. The choice of granularity depends on the balance between the need to share the same label among many destinations and the need to maximize switching capability while husbanding resources. The granularity of LSP ranges from the IP prefix to application–level flow. The switched paths in the MPLS network take the form of a multipoint–to–point tree. The merging of the switched paths that occurs at a node when multiple upstream paths for a given stream are spliced to a single downstream switched path for the stream. In the case of ATM, however, merging is not always possible since most ATM switches are not capable of reassembling cells from multiple inbound VCs without the problem of cell interleaving. One solution to this problem is to use virtual paths rather than virtual channels to merge streams [8]. The merging of VPs creates a tree of VPs. Cell interleaving is prevented by assigning unique VCIs within each VP. MPLS performs explicit routing by combining a prespecified label to the LSP at the time the LSP is set up. This

makes it possible to introduce several features. One is traffic engineering, in which path selection is performed by taking into account network efficiency. Of course, there are reasons such a policy is not used for normal datagram networks, such as IP networks. One important reason is reliability, which is an active research area in MPLS. (See, for example, Ref. [6].)

While MPLS needs to establish a closed domain for using a new lower–layer technology, using photonic technology to build a very high speed Internet is useful. The recent advances in WDM technology that enable packet switching to be performed in the optical domain demonstrate the possibility of multi–protocol lambda switching (or λ–MPLS) [8, 9]. MPLS was recently extended to support various photonic networking technologies, including SONET/SDH and WDM. This generalized MPLS (GMPLS) [10] is now being standardized in the IETF. Hereafter, we focus on the emerging λ–MPLS photonic technology, which is a subset of GMPLS.

Among the several options of MPLS, explicit routing is the ability to explicitly determine the route a packet traverses. In such a network, a lightpath is established between end–node pairs (ingress/egress LSRs) based on traffic demand within the MPLS domain. The LSR in an electronic MPLS can generally perform various operations on packet labels, including label swapping, label merging, and label stacking [8]. However, it has been difficult to achieve those functions in the optical domain. One exception is that by viewing the wavelength as a label, label swapping can be performed by changing the incoming wavelength to a different wavelength at the optical cross–connect switch. However, high–speed wavelength

conversion is difficult to perform on a packet–by–packet basis with current technology, so the functionalities of core LSRs are very limited in λ–MPLS.

A key to achieving λ–MPLS is determining how to establish the logical topology offered to the upper layer protocol (IP in the current case). In the logical topology, wavelength paths are configured over the WDM physical network in order to carry IP packets over wavelength paths, so that no electronic processing is needed at intermediate nodes. Thus, λ–MPLS can potentially resolve router bottlenecks, but it still has several problems. The most difficult problem is capacity granularity: the unit of bandwidth between edge–node pairs in the MPLS domain is the wavelength capacity. It may sometimes be too large to accommodate traffic between node pairs. One approach to resolving the capacity granularity problem is wavelength merging [11], but the related technology is still immature. Thus, the lightpath should be set up in a circuit–switched fashion between the ingress/egress LSRs. For IP, it is natural to establish all–to–all connectivity among LSRs.

Once the logical topology is obtained, four functions are necessary in λ–MPLS. 1) Ingress LSR; maps an IP address to a wavelength label. 2) LSP (Label–Switched Path); the labeled wavelength, i.e., lightpath. 3) Core LSR (Core Label Switching Router); an optical cross–connect switching directly connecting input wavelength to output wavelength. Packet forwarding capability at the IP layer may be necessary if packets with different labels share the same lightpath. 4) LDP (Label Distribution Protocol); the logical topology design algorithm is utilized to implement the signaling protocol.

1.2.2 Logical Topology of Lightpaths and Its Design

There are three main network architectures for IP over WDM networks: *peering model*, *integrated model*, and *overlay model* [12]. The peering model separates the photonic network from the IP network, with different routing and signaling for each. The two networks interact with each other as peers using an exterior gateway protocol. GMPLS–based signaling and provisioning are generally used to provide a control plane for the photonic network. The integrated model integrates the IP router and optical cross–connect, and they are considered to be a single network element using a single network routing and signaling [13]. GMPLS technology is again generally used to provide a single integrated control plane. The overlay model allows intermediate electronic packet processing if a direct lightpath cannot be set up between two nodes. In this case, two or more lightpaths are used by packets to reach the destinations (*multi–hop approach* [14]). The packets on a lightpath terminating at a node are processed by the electronic IP router and forwarded to other nodes. In this network architecture, the traditional IP routing protocol is performed on the logical topology. The role of the WDM network is to provide optical connectivity to the upper layer protocol.

Many researchers have studied design methods for logical topology design, which entails a part of the routing and wavelength assignment (RWA) problem, which is NP–hard since the subset is already known to be a NP–hard [15]. Mukherjee et al. formulated this problem as a mixed–integer linear problem and solved it using a meta–heuristic approach [16]. However, its computation time is lengthy, so heuristic algorithms with various objective functions and constraints have been

studied (see Ref. [4] and references therein). The researchers assume that a protocol in the photonic network determines the actual route of the electronic packets, while the IP protocol also determines the route. The explicit routing functionality of MPLS may be used for the route determination. Katou et al. considered a logical topology design algorithm based on the nature of IP routing [17].

This previous research considered a single–fiber network. A multi–fiber network may have enough wavelengths to provide a fully meshed network. Furthermore, with multiple fibers, there is a *limited wavelength translation capability*. Thus, the single–fiber network is the worst–case scenario for evaluating performance. Xu et al. recently considered the multi–fiber case [18].

1.3 Protection/Restoration Schemes

As mentioned above, reliability mechanisms in WDM network can be roughly categorized into protection schemes and restoration schemes. Protection schemes allocate explicit resources for backup purposes, so they consume wavelengths. Restoration scheme does not allocate explicit resources, so they do not consume any 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 failure recovery. Protection scheme can be further classified schemes: *dedicated protection* schemes, in which a backup lightpath is

dedicated to its corresponding 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* of each other. Two primary lightpaths are failure independent of each other, if they do not share any supported failure components. The protection schemes have two types of backup–path coordination:*line* protection and *path* protection. With line protection, a loop path is set up around the supported failure components for backup. With path protection, a backup path is set up between each sender node and destination node.

1.3.1 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 designated transmitter or receiver 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 on 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 corresponding network components. If a failure occurs, a node nearest to the failure components switches into a backup path if there is link protection. If there is path protection, originating node of the corresponding lightpath switches into a backup lightpath. Before a network provider replace a failed component, the location of it must be

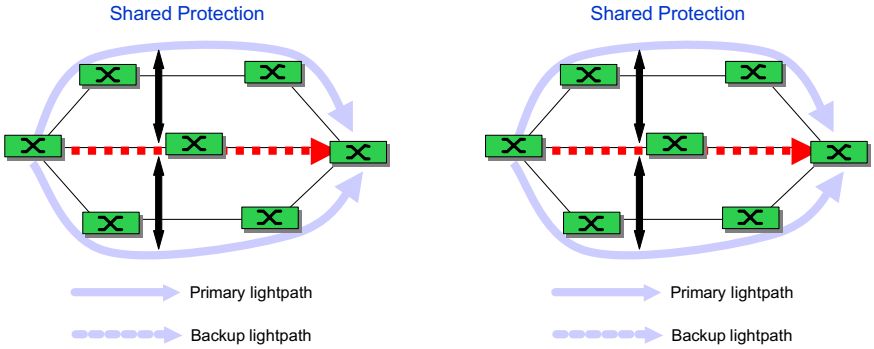
resolved. Techniques for doing this are summarized in Ref. [19].

1.3.2 Dedicated Protection Schemes

In dedicated protection schemes, a backup lightpath is dedicated each primary lightpath, so called “1+1” or “1:1” protection (see Figure 1.4(a)). The backup lightpath carries a copy of the signal carried by the primary lightpath in the “1+1” protection schemes. The receiver node thus receives two signals, one from the primary lightpath and one from the backup. The node 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 when there are no failures because the bandwidths of the backup lightpaths are always begin used, although the backup lightpaths can be used for low–priority IP traffic if there are “1:1” scheme. However, since little coordination is needed to recover from failures, the recovery time is shorter in “1+1” scheme.

1.3.3 Shared Protection Schemes

In the shared protection schemes, several primary lightpath share a backup path by relaxing the type of failures they concern. A shared protection scheme must be carefully engineered so that any two primary lightpaths can use a backup lightpath at the same time when a failure occurs. The backup resources are thus more effectively utilized, as shown in Figure 1.3.2. Ramamurthy and Mukherjee, for



(a) Dedicated protection (b) Shared protection

Figure 1.4: Illustrative example: use of wavelength resources in two protection schemes

example, showed that wavelength resources can be reduced by 20–44% using a shared path protection scheme [20]. However, it takes more time to recover from a fault because the backup path must be coordinated before using it.

1.3.4 Restoration Schemes

A restoration schemes in the optical layer calculates the backup path on the fly when a failure occurs. Since backup paths need not be allocated, wavelength resources can be more effectively utilized for transporting IP traffic. However, calculating alternate routes, following a failure can take second or even minutes. Thus, a restoration scheme is usually combined with a protection scheme [21]. After the failure recovery is completed by the protection, restoration 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.

1.4 Methods for Designing Reliability

To construct a reliable IP over WDM network, backup paths as well as primary paths should be embedded within a 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 a lot of wavelength resources and is less effectiveness. We consider a single fiber failure in this section. Multiple failures and node failures are assumed to be handled by the restoration functionality of the IP layer. Furthermore, it may not be necessary to protect all the lightpaths by using the optical layer if doing so does not lead to cost–saving even when a shared protection scheme is used. If we allows that several primary lightpaths cannot recover from some failure patterns and 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 will be maximized at the price of reliability.

In this section, we 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 function.

Subsection 1.4.1 describes the reliability design for single–layer case, and subsection 1.4.2 describes multi–layer survivability, in which a sub–set of lightpaths are protected against failure.

1.4.1 Single–Layer Case

Protection schemes for WDM networks have been widely studied [20–30]. Here, we consider the shared path protection scheme, which provides 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 the link. Our formulation in this subsection is based on that of Ramamurthy and Mukherjee [20].

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 is used for characterizing the physical WDM network.

N: number of nodes in a 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 is also 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 our 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:

$$w_{mn} = \sum_{w \in W} o_{mn}^{w}. \tag{1.1}$$

(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:

$$b_{mn} = \sum_{w \in W} m_{mn}^{w}. \tag{1.2}$$

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

$$o_{mn}^{k} + m_{mn}^{k} \leq P_{mn}. \tag{1.3}$$

(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:

$$c_{ij}^{k} = \sum_{w \in W} \sum_{it \in A_{ij}^{k}} g_{ij,pq,k}^{it,w}. \tag{1.4}$$

Note that it is unnecessary to use different wavelengths between 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):

$$g_{ij,pq,k}^{nt,w} = g_{ij,pq,k}^{tm,w} \quad \forall pq \in R_{ij}^{k}, \forall nt, tm \in A_{ij}^{k}. \tag{1.5}$$

(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:

$$g_{ij,p_1q_1,k}^{mn,w} = g_{ij,p_2q_2,k}^{mn,w} \quad \forall p_1q_1, p_2q_2 \in R_{ij}^{k}. \tag{1.6}$$

As is equation indicates, we assume that we allow to use the 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:

$$\sum_{ij} \sum_{k \in W : c_{ij}^k > 0 \wedge pq \in R_{ij}^k \wedge mn \in A_{ij}^k} g_{ij,pq,k}^{mn,w} \leq 1. \tag{1.7}$$

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

$$\varphi_{mn}\, m_{mn}^k \geq \sum_{w \in W} \sum_{(i,j):(c_{ij}^k > 0, mn \in A_{ij}^k)} \sum_{pq \in R_{ij}^k} g_{ij,pq,w}^{mn,k}. \tag{1.8}$$

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.

Heuristic Approaches

Formulations of the wavelength assignment problems 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 [31] can be used to solve it. However, an MILP can be solved only for 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. In the case of 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 the 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 to 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 value of 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 two 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 backup lightpath, check whether these backup lightpaths can share the wavelength. They can share if the corresponding primary lightpaths are fiber–disjoint, which means that they have no common link. If they can share the wavelength, delete link mn from B_{ij}^p and go to Step 3. Otherwise, the backup

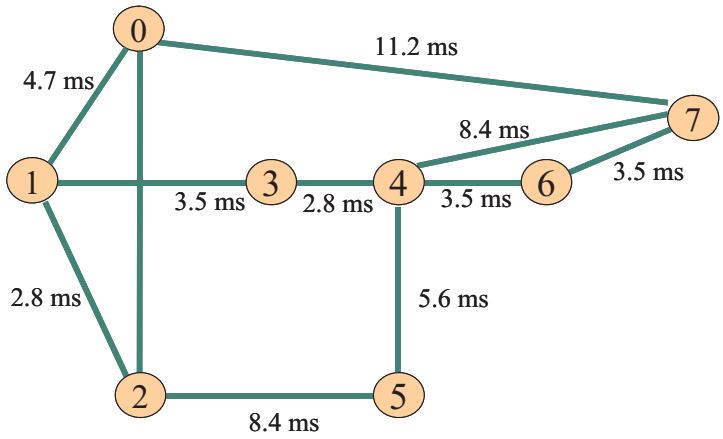


Figure 1.5: Physical topology of eight–node network

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 lightpath is 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.

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 dia-

grammed in Figure 1.5.

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 [32]. The MLDA algorithm works as follows. First, it set up a lightpath between nodes if there exists a fiber. Then, it attempts to set up lightpaths between nodes in the descending order of traffic rates. Finally, if some wavelengths are still unused, lightpaths are set up as much as possible using those wavelengths. The direct application 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 set up a lightpath even if the lightpath has already been set up, we do not set up multiple wavelengths between two nodes so that more wavelengths are left for possible use as backup lightpaths.

(2) While the MLDA algorithm set 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 the one is not used in either algorithm considered in the algorithm, while the largest–traffic–first approach does need one needs it. We used the traffic matrix given in [32] for the reference purpose. 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 algorithms are compared in Table 1.1, which shows the number of wavelengths required to protect all the

Table 1.1: Number of wavelengths required to protect all lightpaths

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

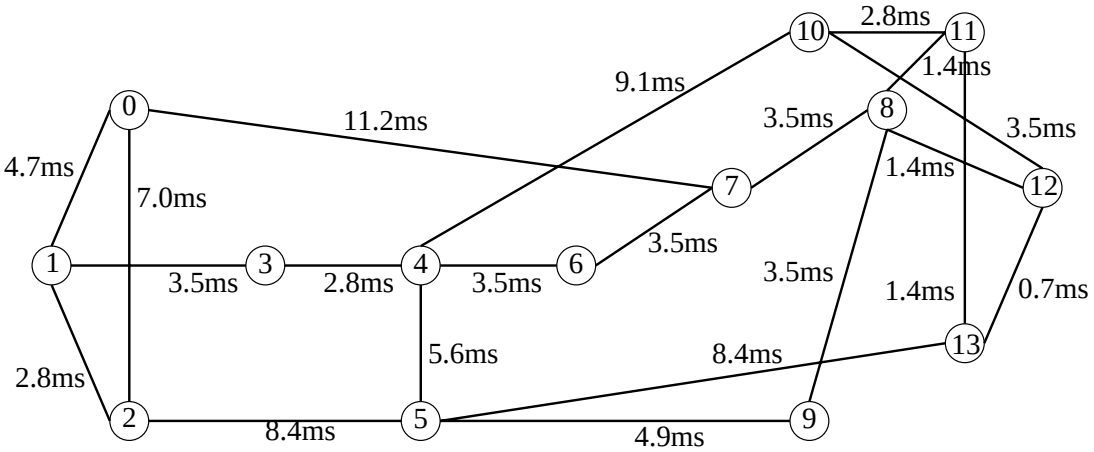


Figure 1.6: NSFNET backbone network model (14 nodes, 20 links)

lightpaths. Good results were obtained with both algorithms.

Results with Heuristic Approach

We next considered a 14–node NSFNET backbone network as the network model. The same traffic matrix [32] was used for reference purposes. Since the MLDA algorithm sets up lightpaths on the physical topology, we must identify the route which the IP packets pass. We modified Dijkstra's shortest path algorithms to consider the nodal processing delays. We assume that the nodal delays are derived from a M/M/1 queuing model and that the offered traffic rates are assumed to be $\sum_s \lambda^{sd}$.

Figure 1.7 compares the three approaches in terms of the number of wave-

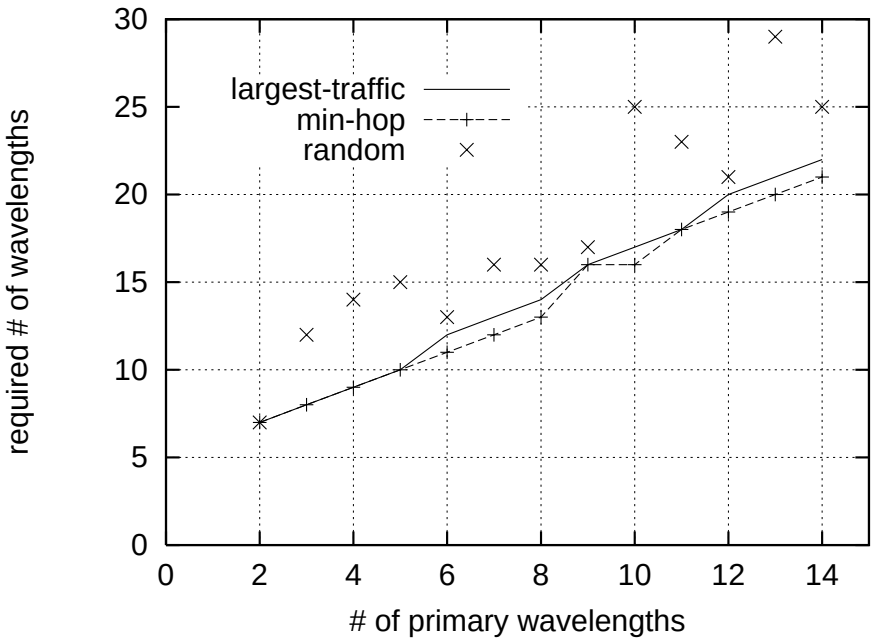


Figure 1.7: Number of wavelengths required to completely protect primary lightpaths

lengths 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 min–hop–first approach. The min–hop–first approach required the smallest number of wavelengths among the three approaches.

1.4.2 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 relationship 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 could not be utilized.

There is another problem. While the WDM protection mechanism can switch to the backup lightpath in the order of ten milliseconds, the 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 the table (typically in the order of ten seconds), it may find a route (which may consist of two ore more concatenated lightpaths) shorter than the backup lightpath allocated by the WDM protection mechanism.

The main cause of the above–mentioned problem is that we did not consider the possibility of 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 algorithms 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 $\{S\}$, identifying the elements of 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 exceed, 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 investigate a compromise between these objectives.

1.4.3 Numerical Examples and Discussion

We investigated the effect of IP/WDM interactions using the NSFNET backbone network model (see Figure 1.6).

As shown in Figure 1.8, the number of protected lightpaths depends on the number of wavelengths available in the fiber. To obtain this relationship, we use the MLDA algorithm [32] 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 algorithms, 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 wavelength, the min–hop–first approach protected more lightpaths than either the largest–traffic–first or random approach.

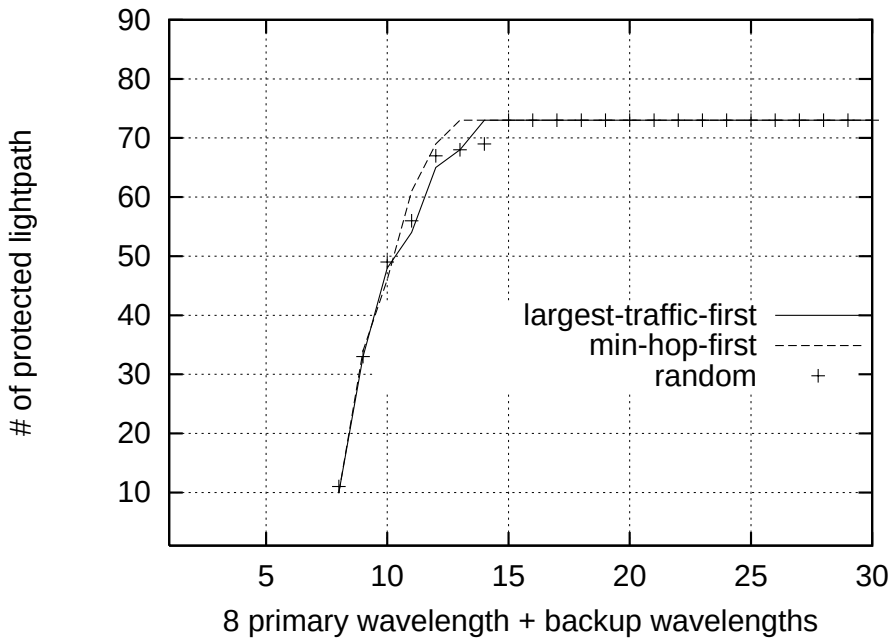


Figure 1.8: Number of protected lightpaths

We next fixed the total number of wavelengths, and changed the number of

wavelengths used for establishing primary lightpaths. Figure 1.9 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 increases. The min–hop–first approach protected the most lightpaths for any given number of backup wavelengths.

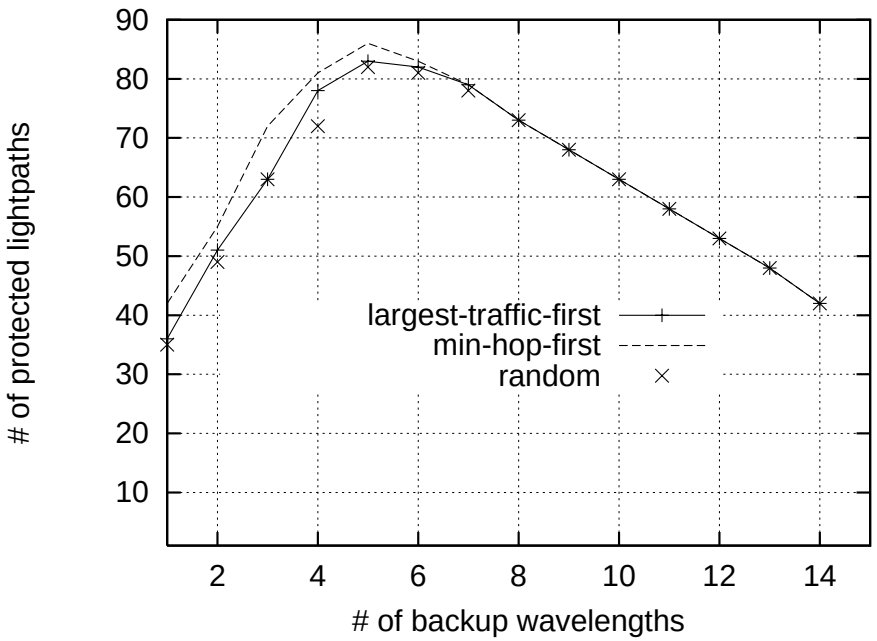


Figure 1.9: Number of protected lightpaths.

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 the 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 load at the router after a single fiber failure. By examining all cases of fiber single–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 Figures 1.10, 1.11, and 1.12, 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. Figures 1.10 through 1.12 show that the maximum traffic rate at the IP router gradually increased as the number of wavelength used for primary lightpaths was increased. That is, the traffic rate at the IP router 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 router is a primary cause of bottlenecks within the network.

To clarify this difference, we next examined the three approaches in terms of traffic volume. As shown in Figure 1.13, the number of wavelengths used for the primary lightpaths was increased, the volume of traffic protected by the backup lightpaths got larger, then decreased because the number of wavelengths available for backup got smaller. In contrast, the amount of traffic that can be restored by the

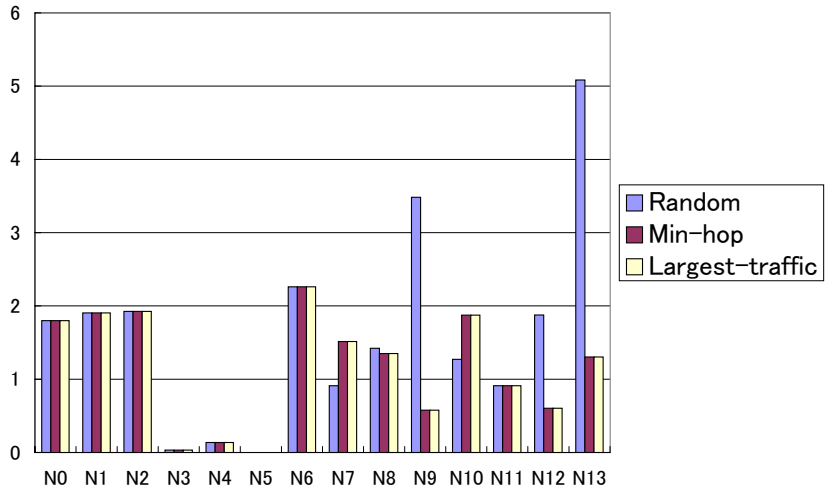


Figure 1.10: Maximum traffic load at IP router after single failure: number of wavelengths used for primary lightpaths is 10

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 Figure 1.14. When the number of wavelengths in the fiber was below nine, the traffic was perfectly protected. However, when it exceeded nine, the volume of the 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 Figures 1.13 and 1.14, 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 the traffic volume.

Finally, we discuss the traffic volume protected after the IP routing table is updated. Figure 1.15 shows the volume of traffic protected when the routing tables

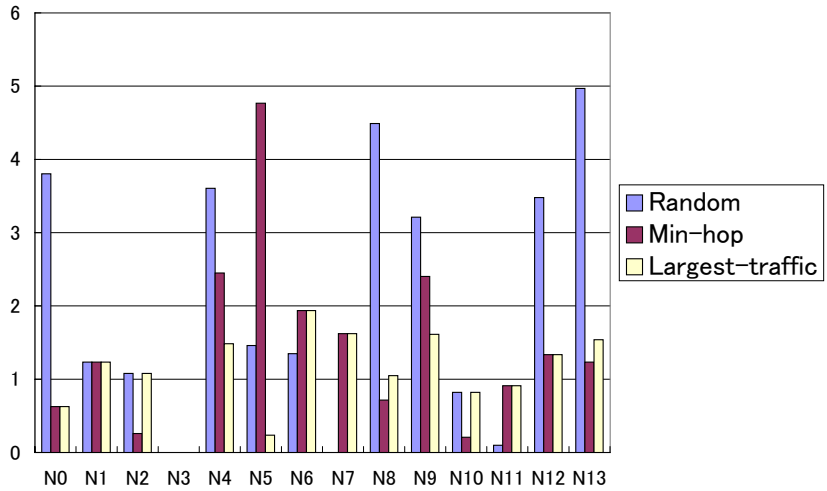


Figure 1.11: Maximum traffic load at IP router after single failure: number of wavelengths used for primary lightpaths is 12

at the nodes were simultaneously updated.

The difference from Figure 1.13 is due to changes in several IP routes. Although IP does not select several backup lightpaths as its routes, we must take this fact into account. It is one of our future research topics to build a set of perfect backup lightpaths such that IP chooses those lightpaths as its own routes.

Figure 1.16 is the complement to Figure 1.15; it shows the volume of traffic not protected after the routing tables were updated. These results clearly show that our proposed method 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–

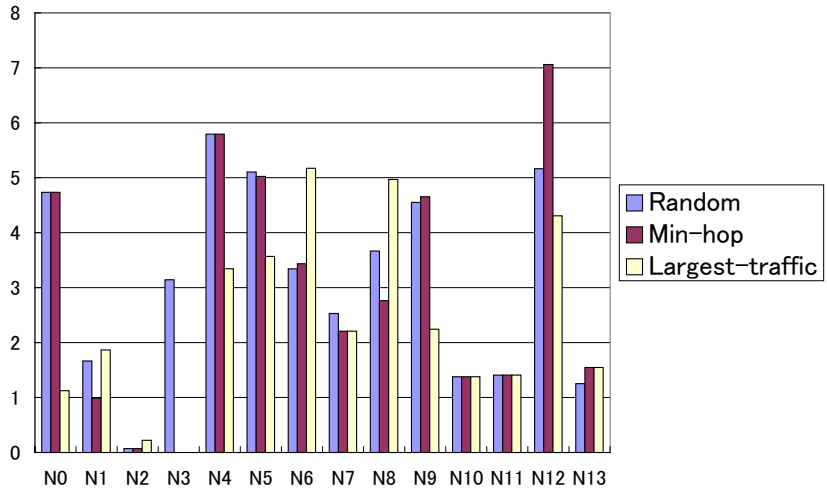


Figure 1.12: Maximum traffic load at IP router after single failure: number of wavelengths used for primary lightpaths is 14

traffic–first approach is better for reducing the traffic load at the IP router.

We also applied our heuristic algorithms to NTT's backbone networks, consisting of 49 nodes and 200 links. For the traffic matrix, we used publicly available traffic data [33]. We again found that the largest–traffic–first approach protects more traffic than the other approaches.

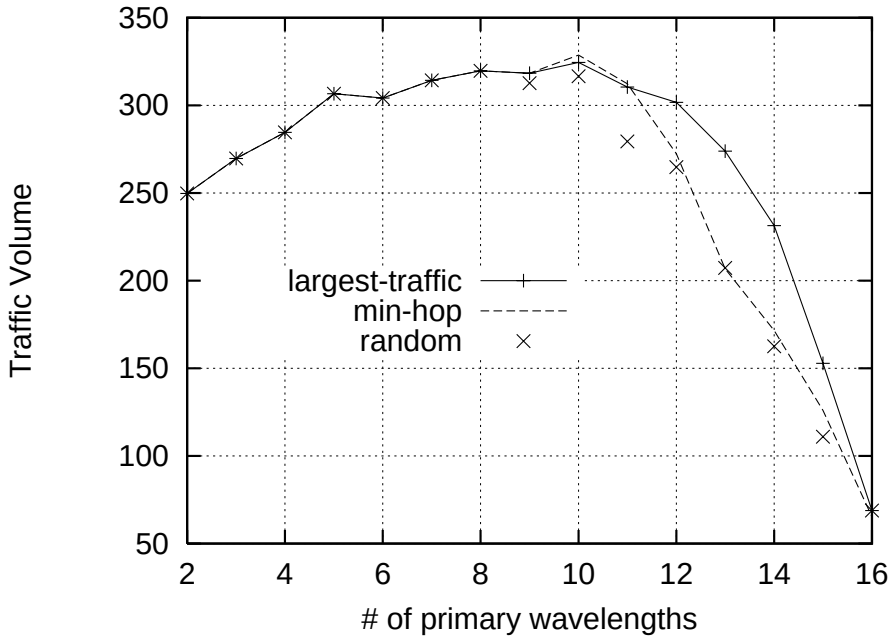


Figure 1.13: Total volume of traffic protected by backup lightpaths before IP routing table update

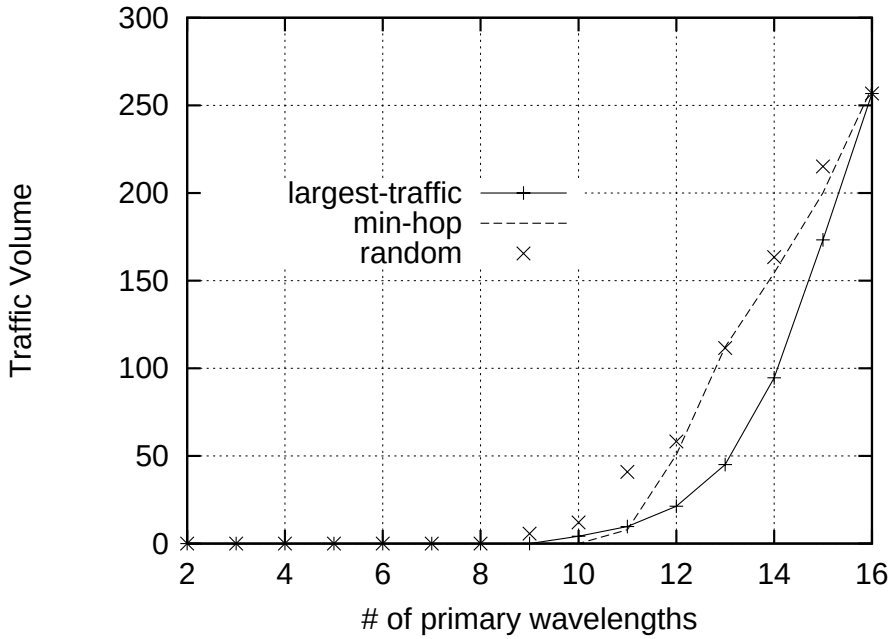


Figure 1.14: Total volume of traffic not protected by backup lightpaths before IP routing table update

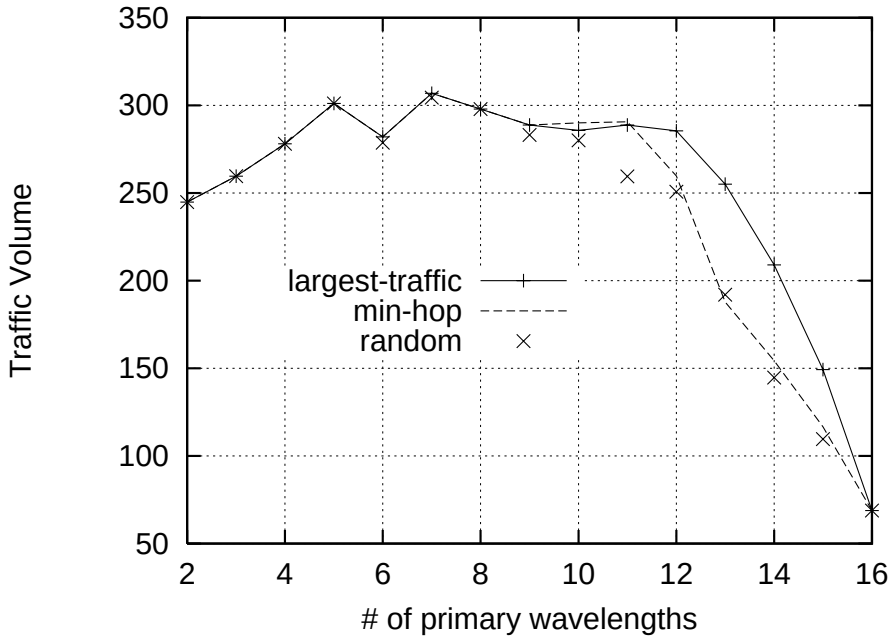


Figure 1.15: Total volume of traffic protected by backup lightpaths after IP routing table update

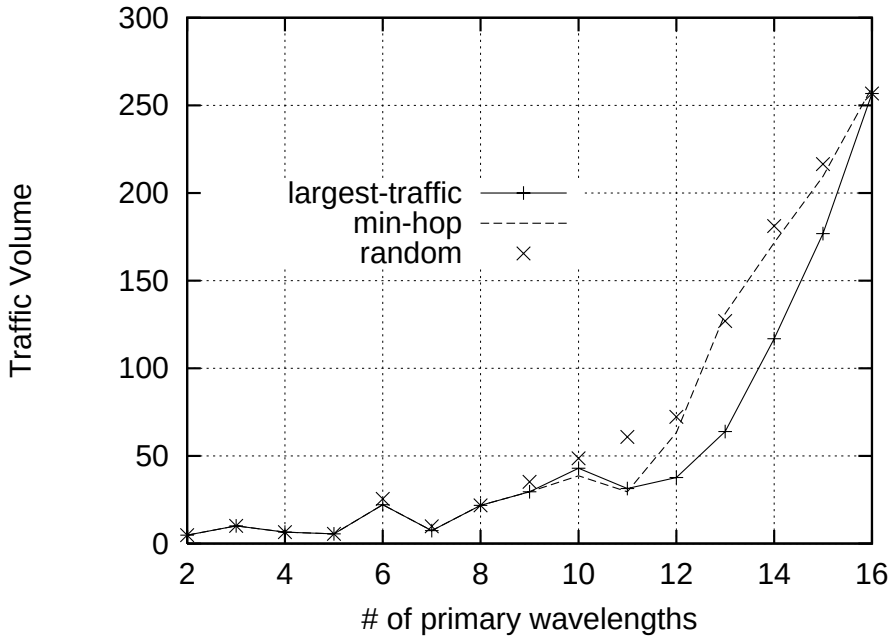


Figure 1.16: Total volume of traffic not protected by backup lightpaths after IP routing table update

1.5 Implementation Issues

1.5.1 Reconfigurability Issue

A lot of past researches including [20] and [30] assume that traffic demand is known a priori. Then, an optimal structure of the logical topology is obtained. Such an assumption is, however, apparently inappropriate especially when the WDM technology is applied to the Internet. In the traditional telephone network, a network provisioning (or capacity dimensioning) method has already been well established. The target call blocking probability is first set, and the number of telephone lines (or the capacity) is determined to meet the requirement on the call blocking. After installing the network, the traffic load is continuously measured, and if necessary, the link capacity is increased to accommodate the increased traffic. By this feedback loop, the telephone network is well engineered to provide QoS (Quality of Service) in terms of call blocking probabilities. Rationales behind this successful positive feedback loop include: (1) the call blocking probability is directly related to the user's perceived QoS in the telephone network, (2) capacity provisioning is easily based on stably growing traffic demands and the rich experiences on past statistics, (3) we have well-established fundamental theory, i.e., Erlang loss formula, and (4) the network provider can directly measure a QoS parameter (i.e., 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. By contrast with the telephone network, there are several

obstacles. (1) The statistics obtained by traffic measurement is packet level and henceforth the network provider cannot monitor or even predict the user's QoS, (2) an explosion of the traffic growth in the Internet makes it difficult to predict a future traffic demand, (3) there is no fundamental theory in the Internet like the Erlang loss formula in the telephone network. A queueing theory has a long history and has been used as a fundamental theory in the data network (i.e., the Internet). However, the queueing theory only reveals the packet queueing delay and loss probability at the router. The router performance is only a component of the user's perceived QoS in the Internet. Furthermore, the packet behavior at the router is reflected by the dynamic behavior of TCP, which is essentially the window-based feedback congestion control [34].

According to the above discussions, the "static" design that the traffic load is assumed to be given a priori is completely inadequate. Instead, a more flexible network provisioning approach is necessary in the era of the Internet. Fortunately, the IP over WDM network has a capability of establishing the above–mentioned feedback loop by utilizing wavelength routing. If it is found through the traffic measurement that the user's perceived QoS is not satisfactory, then new wavelength paths are set up to increase the path bandwidth (i.e., the number of lightpaths).

In this section, we explain our incremental approach to capacity dimensioning of reliable IP over WDM networks [35]. 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

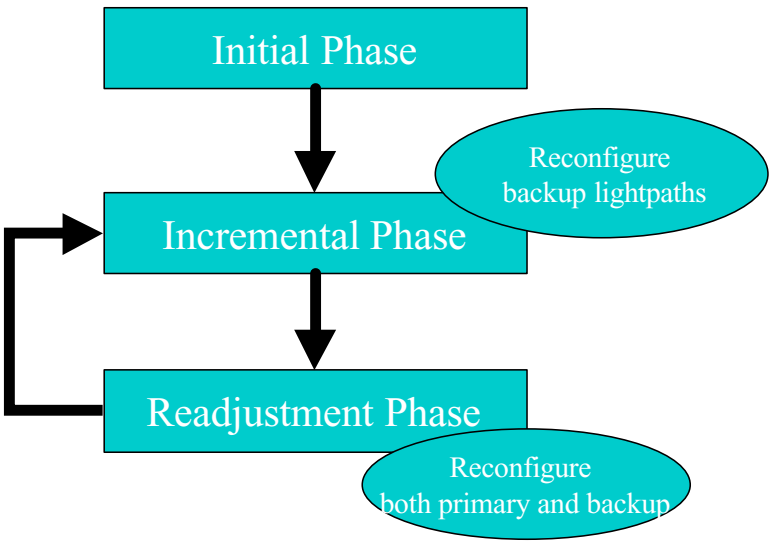


Figure 1.17: Three–step approach to reconfigure logical topology of reliable IP over WDM networks

wavelengths, alert signals are generated so that the network provider can increase the number of fibers to meet the increasing traffic demand.

Initial Phase

In the initial phase, primary and backup lightpaths are set up for given traffic demands. As described above, our approach allows for the likelihood that the projected traffic demands are incorrect. Lightpaths are adjusted in the incremental phase.

Existing design methods for the logical topology can be used in this phase. They include the method for designing the logical topology for primary lightpaths described in [32], and the heuristic algorithm for setting up backup lightpaths for

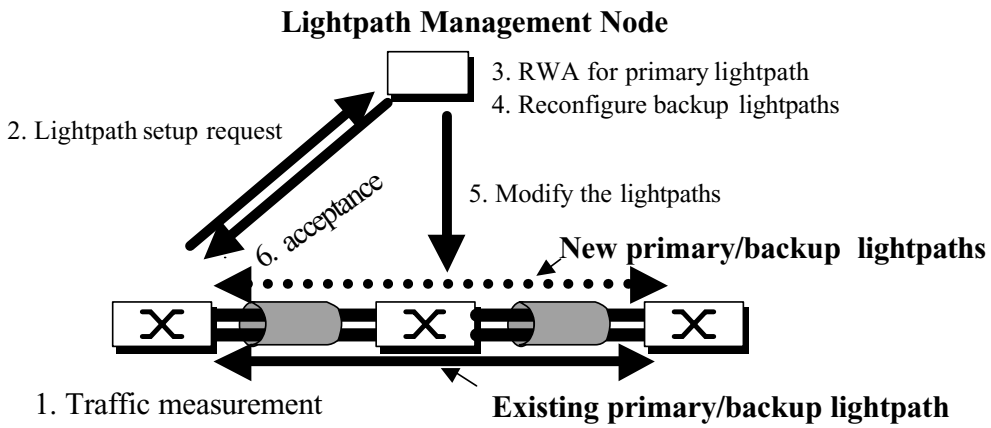


Figure 1.18: Logical topology management model used in the incremental phase

the IP over WDM network as described in section 1.4. In this phase, the number of wavelengths used for setting up the lightpaths should be minimized so that wavelengths remain for handling the increased traffic volume in the incremental phase.

Incremental Phase

The logical topology established in the initial phase must be changed as the patterns of traffic changes. This is performed in the incremental phase. Our logical topology management model is illustrated in Figure 1.18. 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 α $(0 < \alpha < 1)$, 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 by the traffic measurement. It then solves the routing and wavelength assignment problem for both 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 decrease, eventually leading to blocking. To minimize the possibility of blocking, we reconfigure the backup lightpaths for more effective use of the wavelengths. Only the backup lightpaths is reconfigured because the backup lightpaths do not carry traffic unless a failure occurs. We do not change the primary lightpaths in this phase so that the active traffic flows are not affected by the lightpath reconfiguration. In this phase, we need an algorithm for assigning a routing and wavelengths for the new primary lightpaths and one for reconfiguring the backup lightpaths. They will be described in subsection 1.5.2 in detail.

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 recofiguration 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. [36].

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, and consequently the lightpath will be accommodated. References [37, 38] give a reconfiguration policy for this issue, but they only address the primary lightpaths, and a further study is necessary to include the rearrangement of the backup lightpaths.

1.5.2 Incremental Capacity Dimensioning

As we described in Subsection 1.5.1, LMN solves the routing and wavelength assignment problem for new primary lightpath and an optimization problem for reconfiguring the set of backup lightpaths. We will now describe these in more detail.

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 viewd as having different delays. Hence, the IP layer selects the path with the lower delay, and there is no effect of having multiple lightpaths among 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.

We propose a minimum reconfiguring for backup lightpath (MRB) algorithm for assigning the wavelengths of primary lightpath, 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 the algorithm below). If there is no available wavelength, an alert signal is generated. More specifically, our al-

gorithm 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 begin 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 of 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. We prohibit multiple lightpaths with different routes because the IP routing may not choose those paths. That is, IP routing puts all packets onto the primary lightpath with the shorter delay. Multiple lightpaths with different routes can be avoided by using an explicit routing in MPLS [39], 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 [9]. 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.

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, we can expect that possibility of blocking the next arriving lightpath setup requests is minimized. We use a shared protection scheme to improve the use of wavelengths [20]. Before formulating the optimization problem, we summarize the notations we use to characterize the physical WDM network.

N: number of nodes in physical WDM network

W: number of wavelengths in a 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 we use to represent the logical topology after the route and wavelength of the new primary lightpath is determined using the MRB algorithm.

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 a 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 our optimization problem.

b_{nm}: number of backup lightpaths placed on the 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 our optimization problem.

Objective function

Minimize number of wavelengths used for backup lightpaths:

$$\min \sum_{mn} b_{mn}. \tag{1.9}$$

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:

$$b_{mn} = \sum_{w \in W} m_{mn}^{w}. \tag{1.10}$$

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

$$o_{mn}^{k} + m_{mn}^{k} \leq P_{mn} \tag{1.11}$$

3. 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. That is, if $P_{ij}^{k} = 1$,

$$\sum_{w \in W} \sum_{r \in A_{ij}^{k}} \sum_{it \in r} g_{ij,pq,k}^{it,w,r} = 1. \tag{1.12}$$

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 $P_{ij}^k = 1$,

$$g_{ij,pq,k}^{nt,w,r} = g_{ij,pq,k}^{tm,w,r}, \qquad \forall pq \in R_{ij}^k, \forall nt, tm \in r, \forall r \in A_{ij}^k. \tag{1.13}$$

   This is called the "wavelength continuity constraints".

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

$$g_{ij,p_1q_1,k}^{pq,w,r} = g_{ij,p_2q_2,k}^{pq,w,r}, \qquad \forall p_1q_1, p_2q_2 \in R_{ij}^k. \tag{1.14}$$

   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.

$$\sum_{ij} \sum_{k \in W: pq \in R_{ij}^k} \sum_{r \in A_{ij}^k: mn \in r} \sum_{mn \in r} g_{ij,pq,k}^{mn,w,r} \leq 1 \tag{1.15}$$

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

$$\varphi_{mn} \times m_{mn}^w \geq \sum_{k \in W} \sum_{ij} \sum_{r \in A_{ij}^k: mn \in r} \sum_{pq \in R_{ij}^k} g_{ij,pq,w}^{mn,k,r} \tag{1.16}$$

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 $P_{ij}^{k} = 1 \wedge P_{ij}^{k'} = 1 \wedge r \equiv r'$,

$$\sum_{w} \sum_{mn \in r} C_{mn} \times g_{ij,pq,k}^{mn,w,r} = \sum_{w'} \sum_{m'n' \in r'} C_{m'n'} \times g_{ij,pq,k'}^{m'n',w',r'}. \qquad (1.17)$$

Note that in Eqs. (1.15) and (1.16), 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 re set up along the same route. Eq. (1.17) defines this constraint. As described above, the option of explicit routing in MPLS can be used. If it is, the above constraint can be eliminated.

Evaluation

To evaluate our proposed algorithm, we simulated the incremental phase. We used a network consisting of 14 nodes and 21 links as the physical topology (see Figure 1.6). The number of wavelengths in each fiber, W, was 50. As an initial condition, we allocated one primary lightpath for each node–pair. This emulated the initial phase of our approach. The traffic rate given in [32] was used for reference purpose. 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 the primary lightpaths were determined based on the first–fit policy [23]. The

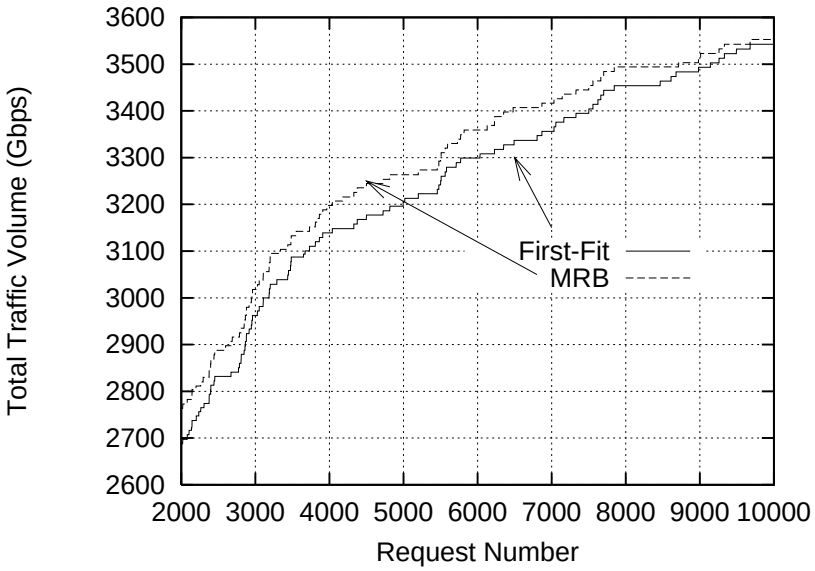


Figure 1.19: Total traffic volume with first–fit and MRB algorithms

wavelengths of 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 measure 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, a requests to set up a new lightpath 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 problem described in subsection 1.5.2. We used the CPLEX opti-

mizer to solve the problem. 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. If the utilization exceeded the threshold, 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 the new lightpath. In the first–fit approach, the wavelength for the new primary lightpath is always checked from λ_1 to λ_W. 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 with a lightpath setup request was blocked due to the lack of available wavelengths. As shown in Figure 1.19, we can see that the MRB algorithm is slightly better than 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 1.20 shows that using our MRB algorithm reduces the value of γ_2 and improve the usage of the wavelengths.

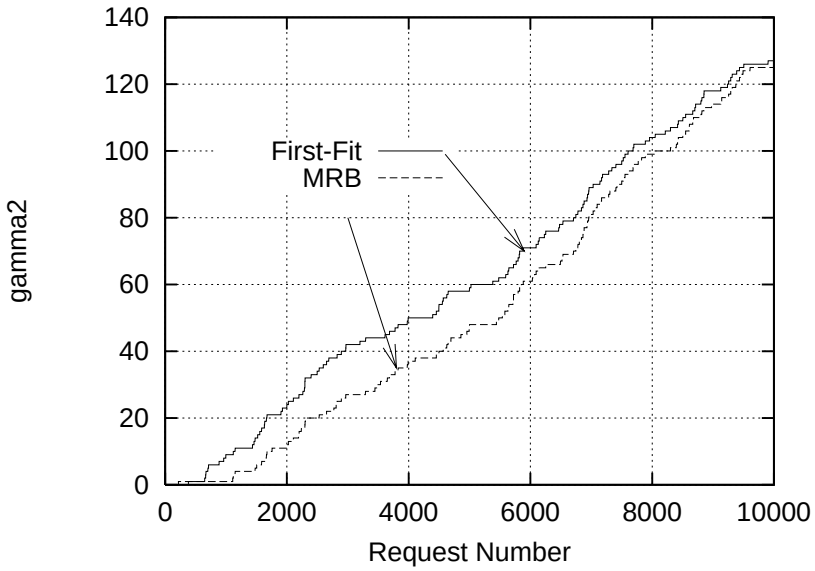


Figure 1.20: Number of lightpath setup request rejected because backup lightpaths could not be reconfigured

1.5.3 Distributed Approaches

So far, we consider 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 size 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 [23], Routes and wavelengths are assigned for each lightpath setup request. Backup lightpaths can be reconfigured to meet future lightpaths setup requests, so that wavelengths are used more effectively. However,

only dedicated protection is considered, so more wavelengths are needed.

As described above, a shared protection scheme is more appropriate in IP over WDM networks since IP routing can also protect against failure. A distributed algorithms for shared protection scheme is introduced in [28].

Mohan et al. considered a restoration method [40]. 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 calculated at the time of establishing primary lightpath.

1.5.4 Quality of Reliability Issue

Quality of reliability or Quality of protection (QoP) is one aspect of quality of service (QoS) that is suitable for the reliable IP over WDM networks. The implementation of QoP has been considered by several research groups [26, 41–43]. One suggested way to provide QoP is to split the primary lightpath into several segments [26, 41]. 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* [26]. They considered the problem of 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 con-

nection are calculated based on a pre–specified reliability measurement of 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 is to use the differentiated reliability (DiR) of a connection [42, 43]. 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.

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 with the above QoS class 1 if a wavelength is not available.

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

This QoP mechanism can easily be implemented 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 in the same. We simply set QoP_{ij} to 0 for both classes. To provide both primary and backup lightpaths in the incremental phase, we change Eq. (1.12):

$$QoP_{ij} = \sum_{w \in W} \sum_{r \in A_{ij}^k} \sum_{it \in r} g_{ij,pq,k}^{it,w,r}. \tag{1.18}$$

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

1.6 Summary and Future Research Topics

In this chapter, we discussed the reliability issues in IP over WDM networks. We first described the multi–layer survivability in IP over WDM networks. Assuming a single–failure within a network, we formulated a shared link protection mechanism as an optimization problem. It is formulated as an MILP and becomes computationally intensive as the network grows in size. We thus proposed two heuristic approaches and compared them with the solution obtained by the formulation. Through numerical examples, we compared the number of wavelengths required for network reliability. We next considered the functional partitioning of IP routing and WDM protection for improving reliability. Based on our heuristic algorithm, we discussed the effect of interaction between IP and WDM layers. Simulation results showed that the largest-traffic-first approach is best if our primary concern is traffic load at the IP router after a failure.

We next proposed a framework for the incremental use of wavelengths in IP

over WDM networks with protection. Our framework provides a flexible network structure against changes in traffic volume. Three phases (initial, incremental, and readjustment) were introduced for this purpose. In the incremental phase, only the backup lightpaths are reconfigured to improve the use of wavelengths. In the readjustment phase, both primary and backup lightpaths are reconfigured, since an 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 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 research.

Bibliography

[1] IEEE NETWORK, *a special issue on IP–Optical Integration*, vol. 15. July/Augest 2001.

[2] M. Murata and K. Kitayama, "A perspective on photonic multiprotocol label switching," *IEEE NETWORK*, vol. 15, pp. 56–63, July/Augest 2001.

[3] K. Kitayama, N. Wada, and H. Sotobayashi, "Architectural considerations for photonic IP router based upon optical code correlation," *IEEE Journal of Lightwave Technology*, vol. 18, pp. 1834–1844, Dec. 2000.

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

[6] B. Davie, P. Doolan, and Y. Rekhter, *Switching in IP Networks - IP Switching, Tag Switching, and Related Technologies*. Morgan Kaufmann, 1998.

[7] G. Arnutage, "MPLS: The magic behind the myths," *IEEE Communications Magazine*, pp. 124–131, Jan. 2000.

[8] N. Ghani, S. Dixit, and T.-S. Wang, "On IP-over-WDM integration," *IEEE Communications Magazine*, pp. 72–84, Mar. 2000.

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

[10] C. Huang, V. Sharma, S. Makam, and K. Owens, "A path protection/restoration mechanism for MPLS networks," *IETF Internet Draft*, 2000.

[11] J. Bannister, J. Touch, A. Willner, and S. Suryaputra, "How many wavelength do we really need? a study of the performance limits of packet over wavelength," *Optical Networks Magazine*, vol. 1, pp. 11–28, Apr. 2000.

[12] J. Wei, C. Liu, S. Park, K. Liu, R. Ramamurthy, H. Kim, and M. Maeda, "Network control and management for the next generation Internet," *IEICE Transaction on Communications*, vol. E83-B, pp. 2191–2209, Oct. 2000.

[13] M. Kodialam and T. Lakshman, "Integrated dynamic IP and wavelength routing in IP over WDM networks," *Proceedings of INFOCOM*, 2001.

[14] B. Mukherjee, "WDM–based local lightwave networks part ii, multihop systems," *IEEE Network Magazine*, July 1992.

[15] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: An approach to high bandwidth optical WAN's," *IEEE Transactions on Communications.*, vol. 40, pp. 1171–1182, July 1992.

[16] 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.

[17] J. Katou, S. Arakawa, and M. Murata, "Design method of logical topology with stable routing for IP over WDM networks," *Proceedings of Optical Network Design and Modeling*, Feb. 2001.

[18] S. Xu, L. Li, and S. Wang, "Dynamic routing and assignment of wavelength algorithms in multifiber wavelength division multiplexing networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 2130–2137, Oct. 2000.

[19] 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.

[20] S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks, part i - protection," *Proceedings of Infocom*, pp. 744–751, Mar. 1999.

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

[22] 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.

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

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

[25] M. Kodialam and T. Lakshman, "Dynamic routing of bandwidth guaranteed tunnels with restoration," *Proceedings of IEEE INFOCOM*, Apr. 2000.

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

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

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

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

[30] 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.

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

[32] 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.

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

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

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

[36] 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.

[37] 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.

[38] 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.

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

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

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

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

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