ABSTRACT

Yufeng Xin. Topology Design of Large-Scale Optical Networks. (Under the direction of Professor George N. Rouskas and Professor Harry G. Perros).

Optical networks consisting of optical cross-connects(OXCs) arranged in some arbitrary topology are emerging as an integral part of the Internet infrastructure. The main functionality of these networks will be to provide reliable end-to-end lightpath connections to large numbers of electronic label switched routers (LSRs). We consider two problems that arise in building such networks. The first problem is related to the topology design of optical networks that can grow to Internet scales, while the second is related to the light-tree routing for the provision of optical multicast services.

In the first part of the thesis, we present a set of heuristic algorithms to address the combined problem of physical topology design (i.e., determine the number of OXCs required for a given traffic demand and the fiber links among them) and logical topology design (i.e., determine the routing and wavelength assignment for the lightpaths among the LSRs). We then extend our study to take a shared path-based protection scheme into consideration after presenting a detailed analysis and comparison of different protection strategies. In order to characterize the performance of our algorithms, we have developed lower bounds which can be computed efficiently. We present numerical results for up to 1000 LSRs and for a wide range of system parameters such as the number of wavelengths per fiber, the number of transceivers per LSR, and the number of ports per OXC.

In the second part of the thesis, we study the problem of constructing light-trees under optical layer power budget constraints, with a focus on algorithms which can guarantee a certain level of quality for the signals received by the destination nodes. We define a new constrained light-tree routing problem by introducing a set of constraints on the source-destination paths to account for the power losses at the optical layer. We investigate a number of variants of this problem, we characterize their complexity, and we develop a suite of corresponding routing algorithms. We find that, in order to guarantee an adequate signal quality and to scale to large destination sets, light-trees must be as balanced as possible. Our algorithms are designed to construct balanced trees which, in addition to having good performance in terms of signal quality, they also ensure a certain degree of fairness among destination nodes.

Topology Design of Large-Scale Optical Networks

by

Yufeng Xin

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial satisfaction of the
requirements for the Degree of
Doctor of Philosophy

Department of Computer Science

Raleigh

2002

Approved By:

Dr. Matt Stallmanne	Dr. Yahya Fathi
Dr. George N. Rouskas Chair of Advisory Committee	Dr. Harry G. Perros (Co-Chair)

To my parents . . .

Biography

Yufeng Xin was born in China in 1969. He received his M.E. and B.E. in 1993 and 1990 in China. He was a university lecturer and an electrical engineer in China before he began to work full time towards his Ph.D. degree at the Operations Research program with concentration on Computer Science at the NC State University at Raleigh, North Carolina, USA in 1998. He also obtained a degree of Master in Computer Science in 2001 from the Computer Science department at the NC State University.

Acknowledgements

I am very grateful to my advisors, Dr. Rouskas and Dr. Perros, for their invaluable support and guidance. They not only inspired many ideas in this thesis, but also taught me how to write concise and beautiful English sentences. Sincere appreciation is also extended to Dr. Fathi and Dr. Stallmann, members of my advisory committee, for their constructive criticism and suggestion.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The wide deployment of point-to-point wavelength division multiplexing (WDM) transmission systems in the Internet infrastructure has enhanced the need for faster switching at the core of the network. The corresponding massive increase in network bandwidth due to WDM has occurred in conjunction with a growing effort to modify the Internet Protocol to support different levels of Quality of Service (QoS). Label Switching Routers (LSRs) supporting Multi-Protocol Label Switching (MPLS) [19] are being deployed to address these two issues. On one hand, LSRs simplify the forwarding function, thereby making it possible to operate at higher data rates. On the other hand, MPLS enables the Internet architecture, built upon the connectionless Internet Protocol, to behave in a connection-oriented fashion that is more conductive to supporting QoS.

The rapid advancement and evolution of optical technologies makes it possible to move beyond point-to-point WDM transmission systems to an all-optical backbone network that can take full advantage of the available bandwidth by eliminating the need for per-hop packet forwarding. Such a network consists of a number of optical cross-connects (OXCs), arranged in some arbitrary topology, and provides interconnection to a number of client networks, e.g., IP subnetworks running Multi-Protocol Label Switching (MPLS) [64, 19].

Each OXC can switch the optical signal coming in on a wavelength of an input fiber link to the same wavelength in an output fiber link. An example OXC is showed in Figure 1.1. The OXC may also be equipped with converters that permit it to switch the optical signal on an incoming wavelength of an input fiber to any wavelength on an output

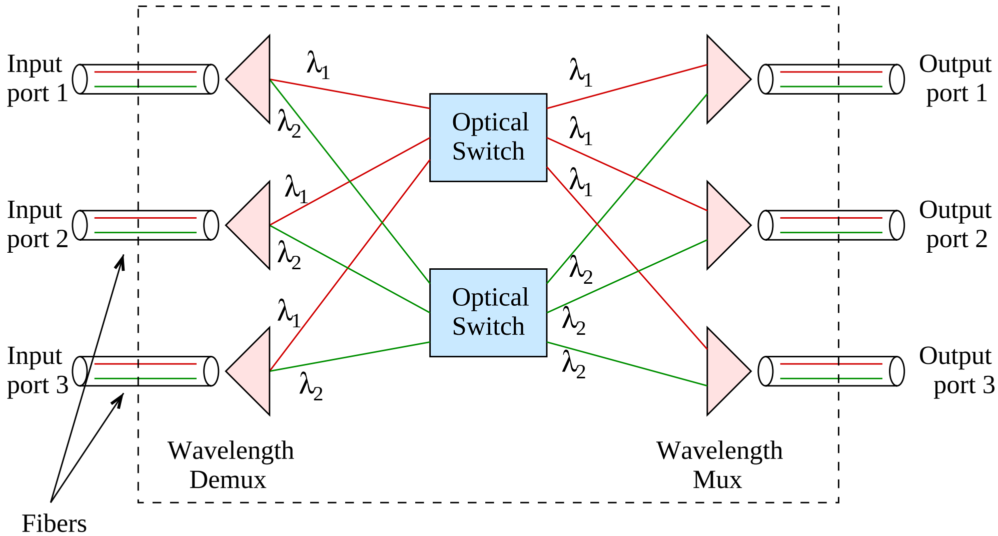


Figure 1.1: OXC

fiber link. The main mechanism of transport in such a network is the lightpath, which is a communication channel established between two OXCs or two LSRs and which may span a number of fiber links (physical hops). If no wavelength converters are used, a lightpath is associated with the same wavelength on each hop. This is the well-known wavelength continuity constraint. Using converters, a different wavelength on each hop may be used to create a lightpath. Thus, a lightpath is an end-to-end optical connection established between two LSRs.

Because of the wide deployment of ATM and SONET/SDH networks and equipment, a straightforward architecture of using WDM techniques to support IP subnetworks is IP/ATM/SONET/WDM. But this structure suffers from low scalability and low cost-efficiency due to the existence of multiple control planes. A novel IP-over-WDM architecture has been proposed to eliminate the intermediate ATM and SONET/SDH layers [6]. An extension to the MPLS protocol, Generalized MPLS (GMPLS) protocol, has been proposed for this architecture [31]. GMPLS supports multiple types of switching, including switching based on wavelengths usually referred to as Multi-Protocol Lambda Switching (MPλS). The new IP network with the MPLS extensions will take over the traffic engineering (TE) functions which used to be provided by the ATM layer, while the OXC-based WDM technology will take over the switching and protection functions previously provided by the

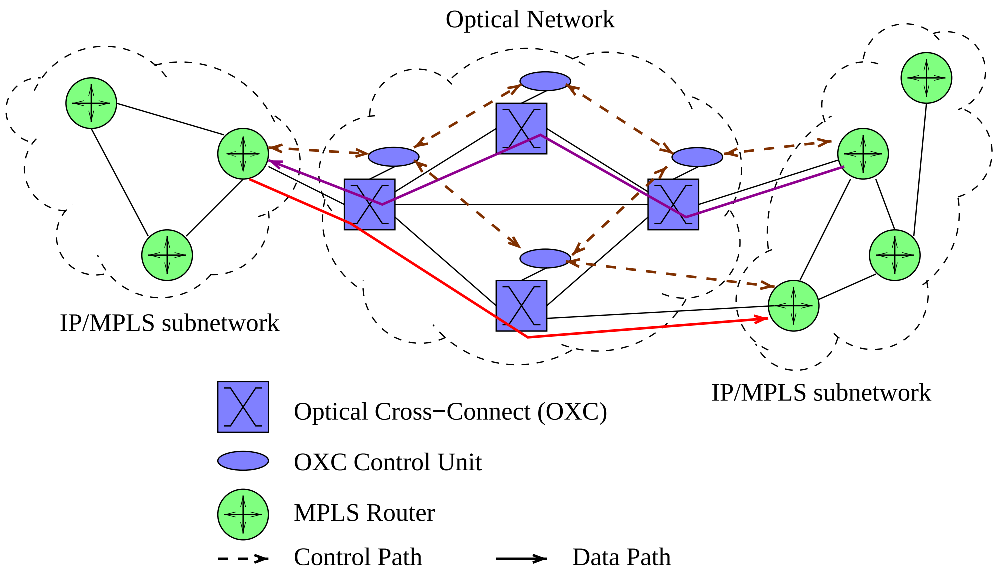


Figure 1.2: MPλS Optical Network

SONET/SDH layer. An IP router supporting MPLS is usually called the Label Switching Router (LSR). Such a network supports two types of label switched paths (LSP): LSPs similar to those in IP/MPLS networks (IP-LSP), and lightpaths (λ-LSPs). Therefore a link in a MPλS network might be a physical link (fiber hop) or a logical link (lightpath). A service requirement is satisfied by a service connection (path) passing one or several physical or virtue links. An example MPλS Optical Network is depicted in Figure 1.2.

Two models, the *overlay* model and the *peer* model have been proposed for MPλS networks. In the *overlay* model, the IP layer and the WDM layer use different control planes. The WDM layer manages a virtual topology made up of a set of lightpaths, while the IP/MPLS layer routes the traffic over the virtual topology and takes care of the TE requirements. In this model, the set of lightpaths (virtual topology) is usually static or semi-static. In the *peer* model, an integrated control plane will manage both IP/MPLS and WDM layers such that the physical links in the WDM layer and the logical links (lightpaths) in the IP layer will be managed in a unified way.

Currently, there is tremendous interest within both the industry and the research community in optical networks of OXCs. The Internet Engineering Task Force (IETF) is investigating the use of GMPLS [30]. Also, the Optical Domain Service Interconnection (ODSI) initiative and the Optical Internetworking Forum (OIF) are concerned with the interface between an LSR and the OXC to which it is attached as well as the interface between OXCs, and have several activities to address MPLS over WDM issues [75]. The majority of the research has been focused on the *overlay* model since it is easier to be implemented using current available techniques and existing networks.

1.1 Thesis Summary and Contributions

In this thesis we consider the problem of designing large-scale optical WDM networks of OXCs that provide reliable end-to-end lightpath services. The scale of the optical backbone network is characterized by the number of LSRs using its services and the number of lightpaths that it can support. We are interested in typical national or international networks, in which the number of LSRs can be in the hundreds or thousands, and the number of lightpaths that need to be established can be an order of magnitude greater than the number of LSRs. Given the number of LSRs, the number of wavelengths per fiber link, and a set of physical constraints (such as the number of transceivers at each LSR, the number of input/output ports at each OXC and the protection requirement), we address both the physical topology design problem (i.e., the number of OXCs required and their interconnectivity) and the routing and wavelength assignment. We assume no wavelength conversion in the network.

Since the problem is NP-complete, we present a set of heuristic algorithms to obtain a near-optimal solution in terms of the number of required OXCs, including a genetic algorithm to search the space of physical topologies. Due to the importance and complexity of considering the survivability in the topology designing of MPλS networks, we separately study the problem with and without the survivability requirement.

While some of the problems we consider have been studied earlier, our work differs from previous studies in several important ways. To the best of our knowledge, this is the first time that the problem of designing the physical *and* logical topology of a wavelength-routed network is fully formulated and solved. Also, whereas previously published algorithms have been applied to relatively small networks (e.g., 10- to 20-node topologies such as the NSFNet and the Arpanet) with few (less than one hundred) lightpaths, we consider large-scale networks of realistic size. We provide new insight into the design of WDM backbone networks by investigating the effect of various system parameters; again, we use realistic ranges for the values of these parameters, e.g., up to 128 wavelengths and up to 24 optical interfaces per LSR. Finally, we obtain lower bounds for the optimization problem and present results which indicate that our heuristics lead to good solutions.

The most important finding of study is that it is possible to build cost-effective optical backbone networks that provide rich connectivity among large numbers of LSRs with relatively few, but properly dimensioned, OXCs. In particular, the number of OXCs increases linearly with the number of attached LSRs, but the rate of the increase is rather slow. We also find that, in order to take advantage of an increasing number of wavelengths, the number of OXC ports must increase correspondingly. Otherwise, using additional wavelengths has little effect on the number and topology of OXCs required for a given number of LSRs. This result, coupled with the fact that the number of wavelengths in a fiber is expected to continue to increase in the foreseeable future, has implications on the OXC technology and design. For instance, our results indicate that 3-D MEMS switches which can scale to large port sizes would be more appropriate than 2-D MEMS switches which are limited to small port sizes [17]. Finally, we also find that it is possible to enhance the degree of connectivity among the LSRs by a large factor (through a corresponding increase of the number of optical interfaces at each LSR) with a relatively small incremental cost in terms of additional OXCs and fiber links among them.

For the survivability consideration, we first study the WDM protection strategy through comparing different protection/restoration schemes. A strategy based on the IP-restoration/WDM-protection and shared path-based WDM protection scheme turns out to be the final winner. We then study the effects of different implementation of this protection schemes to the network topology design. The performance of the joint and separate routing and wavelength assignment policy for the shared path-based protection is evaluated through numerical results. Finally, we integrate the above schemes into our topology design algorithm developed above. We compare the numerical results with those obtained when survivability is not required and those obtained when the dedicated path-based scheme is used.

Finally we study an important application problem for MPλS networks, the problem of light-tree routing in optical networks with light splitting capabilities. Since the effects of light splitting and power attenuation on optical signals are only partially mitigated by amplification, our focus is on algorithms which can guarantee a certain level of quality for the signals received by the destination nodes. The problem of signal quality does not arise in the context of multicast above the optical layer, and, to the best of our knowledge, it has not been directly addressed in the literature. We define a new constrained light-tree routing problem by introducing a set of constraints on the source-destination paths to account for the power losses at the optical layer. We investigate a number of variants of this problem, and we prove that they are all NP-complete. We also develop a suite of corresponding routing algorithms, one of which can be applied to networks with sparse light splitting and/or limited splitting fanout. One significant result of our study is that, in order to guarantee an adequate signal quality and to scale to large destination sets, light-trees must be balanced, or distance-weighted balanced (a term we define later). Numerical results demonstrate that existing algorithms tend to construct highly unbalanced trees, and are thus expected to perform poorly in an optical network setting. Our algorithms, on the other hand, are designed to construct balanced trees which, in addition to having good performance in terms of signal quality, they also ensure a certain degree of fairness among destination nodes.

The Thesis is organized as follows. We first present an extensive review of related work in Chapter 2. Section 2.1 is about the general topology design and the routing and wavelength assignment problem. We study and compare different protection/restoration schemes and figure out a survivability design strategy suitable for MPλS networks in Section 2.2. In Section 2.3, we review the work related to the problem of light-tree routing in the optical network.

The topology design problem of all-optical networks without considering protection is studied in Chapter 3. In Section 3.1 we describe the problem we study as well as the assumptions we make, and in Section 3.2 we present the formulation as an integer programming problem. In Section 3.3, we present the heuristic algorithms for generating a 2-connected graph, routing lightpaths, and assigning wavelengths. We also describe the genetic algorithm used for searching the space of physical topologies. We develop lower bounds for the optimization problem in Section 3.4. We present numerical results in Section 3.5, and we conclude the chapter in Section 3.6.

The protection problem is considered in Chapter 4. In Section 4.1, we present the heuristic algorithms for the protection schemes we study. We present numerical results in Section 4.2, and we conclude the chapter in Section 4.3.

The light-tree routing problem is studied in Chapter 5. In Section 5.1, we describe the multicast optical network under study, and we develop a model to account for optical signal losses in the network. In Section 5.2 we introduce the problem of constructing light-trees under constraints that ensure the quality of the optical signals received at destination nodes. We define three versions of the light-tree routing problem in Section 5.3, mainly differing on which type of power loss is the dominant factor for signal degradation. We characterize the complexity of all three versions of the problem, and we provide light-tree routing algorithms for each. We present numerical results in Section 5.4, and we conclude the chapter in Section 5.5.

Chapter 6 discusses future work and concludes the thesis.

Chapter 2

Related Work

In this chapter, we present an extensive review on research work related to our study. Section 2.1 discusses the general topology design problem in optical networks, as well as the routing and wavelength assignment problem. We study and compare different protection and restoration schemes and describe a survivability design strategy suitable for MPλS networks in Section 2.2. In Section 2.3, we review the work related to the problem of light-tree routing in optical networks.

2.1 Topology Design of All-Optical Networks

The problem of designing wavelength-routed networks of OXCs has received considerable attention in the last decade. In order to reduce the complexity of this problem, typically, it is broken down to two sub-problems: *network design* and *routing and wavelength assignment* (RWA). Network design involves *physical topology design* and *configuration design*.

The topology design involves the determination of the number of OXCs and their interconnectivity. The network configuration is concerned with the determination of the size of OXCs, the number of fibers and the set of lightpaths. Routing and wavelength assignment involves mapping lightpaths onto the physical topology and assigning wavelength to these lightpaths. The reader is referred to [44] for a general analysis of various formulations and solution approaches to the above problems.

Most of the work in the open literature focuses on the configuration design and the RWA problems under the assumption of a fixed fiber physical topology. That is, given a network traffic demand and a physical network topology, an optimal network configuration, a best virtual topology, and an optimal routing and wavelength assignment are obtained. In the case in which the set of lightpaths is also given, the problem is reduced to a pure RWA problem, which can be further decomposed into a routing sub-problem and a wavelength assignment sub-problem.

The design of virtual topologies has been studied extensively. Constrained by the limited number of available wavelengths and the available number of transceivers, it may not be possible to establish a lightpath between every pair of nodes. Consequently, only a selected set of nodes can be connected by lightpaths, leading to a virtual topology over the given physical network. In a virtual topology, the nodes correspond to actual physical network nodes, while the links correspond to lightpaths. A review of algorithms for virtual topology design can be found in [25].

Typically, the network design problem under a given physical topology can be formulated as an integer programming (IP) problem with the objective of optimizing a performance metric of interest. This IP problem has been shown to be NP-hard, and several heuristic algorithms have been proposed in the literature. These algorithms differ in their assumptions regarding the traffic demands, as well as in the performance metric used. Specifically, for the case of static traffic, i.e., when all the connections are known and they are static, the objective is typically to minimize network resource usage. In the case where the number of fibers is limited, the objective is usually to minimize the total number of wavelengths needed in the network. A linear programming relaxation technique to obtain the optimal solution for the RWA of a WDM ring network was proposed in [50], while a longest-lightpath-first heuristic was developed in [15] as an approximate solution to the RWA problem for a mesh network. In the case of a limited number of wavelengths per fiber, [1] provided an iterative scheme to minimize the costs associated with the working fibers for provisioning a static set of lightpaths on a given WDM network topology. A lightpath accommodation heuristic for a given set of lightpaths was developed to minimize the total number of OXC ports (another important network resource) in [57]. A general IP model for the virtual topology design was presented in [7], and a branch-and-bound algorithm to minimize the average lightpath length was described.

The design of a virtual topology was first formulated as an integer linear programming (ILP) problem in [77], where a heuristic algorithm combining simulation annealing and flow deviation was presented to minimize the delay and the maximum flow in a link. A genetic algorithm to calculate the optimal virtual topology and RWA to minimize the average signal delay was developed in [66]. An upper bound on the maximum carried traffic for any routing and wavelength assignment algorithm was derived in [61]. Most of the routing algorithms are adapted from Dijkstra's shortest path algorithm [66, 83, 44]. The wavelength assignment problem is equivalent to a graph coloring problem, and [85] adopted a known heuristic algorithm to solve it. The RWA problem for a multi-fiber network under dynamic traffic was solved in [83] using the layered-graph heuristic model.

The physical topology design problem has also received some attention. The relationship between the number of wavelengths needed and some topology parameters, such as connectivity, nodal degrees, and average hop distance, was studied in [9, 8] through simulation. A bound on the number of wavelengths given the connectivity requirements of the users and the number of switching states was derived in [11]. Simple approximate equations for the scaling properties of the number of wavelengths, the nodal degree, the total fiber lengths, and the maximum number of transit nodes of a lightpath were given in [41]. An analytical solution of the RWA problem for some regular topologies, such as shuffle and tori was carried out in [13, 53].

2.2 Protection in All-Optical Networks and Spare Capacity Allocation

One critical challenge to the design and management of optical networks is survivability. The survivability of a network is defined as its ability to recover the existing service connections that were broken upon failures. The need for highly survivable MPλS networks is obvious. We recall that a single fiber and an OXC have the capacity to transport and switch multi-terabit per second data streams, respectively. Therefore, even a single failure such as a fiber cut or an interface card malfunction with a short duration will result in huge data losses. Unfortunately, the probability of such kind of failures is not low. For instance, Hermes, a consortium of pan-European carriers, estimates an average of one cable cut every four days on their network. And 136 fiber cuts were reported by various United States carriers to the Federal Communications Commission in 1997 alone [32]. Though some authors have also addressed the dual failure scenario [21], we will only consider the possibility of a single failure during a particular time interval since the probability of dual failures is very low.

A communication network is often represented by a graph. For an MPλS network, fibers are represented by links and OXCs and LSRs are represented by nodes. By a network failure, we mean either a node failure or a link failure such that the connectivity of the network is decreased and some original service connections are broken. For example, in an MPλS network we described above, a link failure may be a fiber cut or a in-fiber amplifier failure, a node failure may be a failure of switching elements or interfaces in and between the LSR and OXC. The objective of the network survivability design is to make the probability of data loss as low as possible upon the link or node failure(s). In a survivable network, the original service path is referred to the primary path. Any node failure or link failure within the primary path will break down the service connections carried on this path. The fundamental mechanism of survivability design is to find available backup path(s) to substitute (restore) the broken primary path(s) upon failure(s) so that the affected service connections can be recovered. In order to make the data loss as low as possible in the restoration process, any survivability scheme must satisfy two requirements. First it must be able to provision enough spare capacity (usually in terms of bandwidth or redundant equipment). Second, the process must be as short as possible, i.e., it must have a short restoration time. For example, in the transport layer, the restoration time of $50ms$ provided by the SONET/SDH network is usually considered as the benchmark.

Upon a failure, a general restoration process will invoke the following four steps: fault detection, fault localization, fault notification, and fault restoration. The fault detection is performed individually at all layers through retrieving information from monitoring or testing. For example, at the physical layer, the available information is the optical power level and temperature; at the WDM layer, the available information is about the quality of the optical signal such as the signal-to-noise ratio (SNR) and crosstalk; and at the upper service layer, more detailed information on the quality of signal or data such as bit-error-rate (BER) can be obtained. Upon the detection of a fault, every layer will try to localize the fault. Consequently, a fault at a lower layer will trigger several alarms at the upper layers, which makes the fault localization a difficult task [54]. This is actually another major drawback of the multi-layered architecture of the IP/WDM network. Fortunately, there exist restoration schemes that do not require fault localization as we discuss later. After the fault is detected or localized, the next step is to notify the appropriate network element to find, configure, and provision the spare capacity of the backup path. The last step is, of course, to release the primary path configuration and resource and switch the service connection to the backup path.

Except fault detection which is decided by the physical layer techniques, the other three steps are decided by the survivability schemes, i.e., the way the backup paths are selected and the spare capacity is provisioned. Using different schemes, the fault localization may or may not be needed; the fault notification may be sent to different network elements; and the fault restoration function will choose different backup paths and use different capacity assignment policies. Therefore the network survivability problem can be decomposed into two subproblems: the routing problem (routing of primary and backup paths) and the capacity assignment problem. For WDM networks, the former is the lightpath routing problem and the latter is the wavelength assignment problem.

In the following, we compare and analyze several design strategies that different schemes may take and determine a survivability strategy suitable to the MPλS networks.

2.2.1 Survivability Design Strategies

- **Protection vs. Restoration.** Current spare capacity provisioning schemes upon link failures or node failures can be classified as pre-planned or protection mechanisms and adaptive restoration mechanisms [21] depending on whether the spare capacity is allocated before or after the failure, respectively. For protection mechanisms the backup paths are decided and spare capacity is reserved at the same time as the primary service paths are provisioned. Since a general requirement for protection is to recover every broken primary service connection upon a single failure, i.e., so-called 100must be assigned a backup path with assigned spare capacity. Therefore, the typical objective for a protection scheme is to minimize the total capacity usage, for instance, the total number of wavelengths required in optical networks. The implementation of a protection scheme depends on the way spare capacity is allocated. Spare capacity can be *jointly allocated* with primary service connection requirements, which results in minimal usage of total capacity. It can also be *independently allocated* after all the primary service requirements have been satisfied, which will provide better QoS performance to the primary requirements, but will consume more capacity. Consequently, we have the *joint provision scheme* and *separate provision scheme* in terms of allocation of the primary and backup capacity.

For restoration mechanisms the backup paths are found and provisioned dynamically after the actual occurrence of failures. Compared to protection mechanisms, the restoration schemes need more restoration time since they need to find out the available backup paths and allocate the spare capacity online. But they are more capacity-efficient since no spare capacity needs to be reserved prior. When network resource like the total number of wavelengths in an optical network is limited, spare capacity may not be available along the backup path for every broken primary service connection upon a single failure. In this case, those unrestored service connections will be blocked. Therefore, the typical objective of a restoration scheme is to minimize the total blocking probability upon a single failure. We note that spare capacity and primary capacity are always allocated separately, i.e., primary capacity is always allocated first.

- **Dedicated vs. Shared.** Protection mechanisms can be further classified as dedicated and shared protection. When backup paths are pre-connected, the service demand is transported along both the primary path and backup path simultaneously and the receiver picks up the best signal. This mechanism is referred to dedicated protection, or 1+1 protection. The spare capacity (bandwidth or wavelength) may also be shared among several service paths. The backup paths are pre-decided and the spare capacity is reserved but not used. When a particular service path is broken up, the service is switched to the reserved backup path. To survive from single node or link failures, the primary paths that share the same spare capacity must be mutually node- or link-disjoint and node- or link-disjoint with the backup path. This mechanism is usually referred to as the shared protection mechanism. There may be no sharing (1:1 protection) or sharing among up to N primary paths (1:N protection). Obviously, the *dedicated protection mechanism* requires more spare capacity than the *shared protection mechanism* since in the latter case the spare capacity is shared by several service paths and could be used to transport lower priority preemptive services when there is no fault. On the other hand, unlike the shared protection that needs to configure the backup path for a particular service path upon a failure, the dedicated protection scheme is faster since the receiver only needs to switch to the backup signal upon a failure.

- **Path-based vs. Link-based.** The protection/restoration schemes can also be categorized as path-based and link/node-based schemes. Let us take the IP/WDM network as an example. For the path-based schemes, the fault notification is sent out by any one of the LSRs or OXCs (mostly, the destination) immediately after the failure to the source of the service path. The source will switch the service from the primary path to a backup path that is node-disjoint from the primary path if node failures are considered, or link-disjoint from the primary path if only link failures are considered. We can see that fault localization is not necessary in this case. In the link/node-based schemes, the fault notification is sent to the next router or OXC immediately upstream of the faulty link (or the link adjacent to the faulty node), and this router or OXC will be in charge of restoring the service, usually by finding an available backup path disjoint from the faulty link or node. In this scheme, fault localization is definitely needed. Also the path-based schemes require less spare capacity than the link-based schemes since the former only need to provision spare capacity for a whole path instead of every link along the path. Link-based schemes are usually faster since the path-based schemes need longer time to send the fault notification message to the source of the broken service, though fault localization is not necessary.

- **IP restoration vs. WDM protection.** As we have discussed above, when a failure occurs in a lower layer, all service paths in layers above this layer using the faulty link/node will also be broken. For example in the MPλS network, when a fiber is cut, all the λ-LSPs (logical links to the IP/MPLS layer) passing through this fiber will be broken down and all the IP-LSPs passing on those λ-LSPs will also be broken down. All the upper layers may have their own restoration schemes to restore the service paths in their particular layers. This network structure brings in a natural question, i.e., in which layer should we provision the protection or restoration? For an MPλS network, both the IP/MPLS layer and the WDM layer can fulfill this need. Recent research, however, indicates that schemes using restoration in the IP/MPLS layer and protection in the WDM layer are the best choice [33, 67]. The main reasons for this can be summarized as follows: (1) IP/MPLS operates at small traffic granularity and the traffic is highly dynamic, therefore it can afford a slow but dynamic restoration scheme, and (2) WDM operates at coarse traffic granularity and the virtual topology is semi-static, therefore a fast and spare capacity guaranteed protection scheme is more desirable.

While the restoration issues for IP/MPLS have been extensively studied, protection mechanisms for WDM layer are drawing increasing attention in recent years [23, 55, 60]. Conceptually, a survivable WDM network will use one of the protection schemes from Table 2.1. But the techniques implementing these schemes for the existing networks can not always be immediately applied to the WDM networks. The main reasons are: (1) the service connection in a WDM network is provisioned in units of wavelengths which is an integer, while in other networks, the service connection is provisioned in units of bandwidth that is not integer. Therefore the former is equivalent to an Integer Programming (IP) problem, while the later can be modelled as a Linear Programming (LP) problem. And we know that an IP problem is generally a NP-complete problem, but an LP problem often has polynomial solution. (2) A service connection is a lightpath in the WDM network, which has to meet the wavelength continuity constraint if no converters are used. (3) WDM operates at much coarser traffic granularity than other networks.

- **Link failures vs. Node failures.** When node failures are considered, the backup path must be node-disjoint from the primary path. For the path-based schemes, the backup path must be completely node-disjoint from the primary path, i.e., the two paths may not share the same node except for the source and destination nodes. For the link-based schemes, the backup path must be partially disjoint from the primary path, i.e., the backup path must bypass the faulty node. For link failures, the backup path only needs to be link-disjoint from the primary path, i.e., these two paths may not share the same link. We note that node-disjointness implies link-disjointness. Interestingly, most of the studies on the protection issues in the WDM layer do not take the node failures into account. The reasons can be explained as follows: (1) almost all the node equipment including the switching and transport equipment are redundant at the hardware level, which keeps the node failure probability very low, and (2) most of the failure modes such as router failure and interface failure, are not known to the WDM layer and the upper IP/MPLS layer has to take care of them.

Scheme	*Spare capacity usage*	*Restoration time*
Dedicated path protection	high	very fast
Shared path protection	low	slow
Dedicated link protection	very high	fast
Shared link protection	high	fast
Path restoration	very low	very slow
Link restoration	very low	slow

Table 2.1: Comparison of protection/restoration schemes

Therefore, we will take the same strategy in our study, i.e., node failures will not be considered.

2.2.2 A Survivability Strategy for MPλS Networks

From the above discussion, we clearly see that a survivability strategy based on the IP/MPLS restoration and WDM protection is a reasonable choice. As for the WDM protection, path-based and link-based shared protection schemes provide a tradeoff between the capacity usage and restoration time.

Combining the above different survivability mechanisms, we can obtain several different protection/restoration schemes. In Table 2.1, we tabulate these schemes and compare them qualitatively in terms of spare capacity usage and restoration time. Successful implementation and applications of these schemes can be found in many existing survivable networks. For instance, the retransmission mechanism in a TCP/IP network is a path restoration scheme, while the self-healing ring in the SONET/SDH network is a dedicated link protection scheme. Almost all of these schemes have been investigated in the GMPLS [5] and ATM networks [45].

2.2.3 Related Work for WDM Protection

Many protection and restoration schemes and algorithms for IP/MPLS, ATM, and SONET/SDH networks can be found in the literature. Generally, they are modeled as disjoint paths problems and multicommodity flow problems. For the WDM network, the routing problem is always coupled with the wavelength assignment problem. Therefore, the above methods can not be used directly. The routing problem can still be modeled as the disjoint paths problem, the wavelength continuity constraint and wavelength non-conflict constraint must be considered if the multicommdity flow model is to be used, in which case it becomes a NP-complete problem. Heuristics used for the protection/restoration problems for WDM networks have also been proposed by some researchers.

Several studies on the analysis and comparison of performance of different types of protection and restoration schemes for WDM networks can be found in [5], [23], [60], and [37]. [18] studied the interoperability between the WDM layer protection and upper service layer protection/restoration. The study on the interoperability is important because the disjoint logical links taken by the upper layer may contain a common physical link in the lower layer, which will make the upper layer restoration fail. [33] studied the optimal integration of IP restoration and WDM protection.

As the dominant optical transport network architecture, SONET/SDH provides a well-known robust survivability mechanism, the self-healing ring. It is actually a link-based dedicated loop-back recovery scheme. The building block of SONET networks are rings and every link is incorporated in a ring. When a link failure occurs, the service connection will be automatically loop-backed along the ring. However, a ring based WDM network architecture suffers from low cost-efficiency and scalability, so a mesh-based architecture is more promising [55]. Following the idea of self-healing ring, several ring-based protection schemes have been proposed for the mesh-based WDM network. Among them are the ring covers scheme [27], which finds the minimum number of rings to cover all links and every link is protected by its ring cover in the same way as the SONET/SDH self-healing ring, and the p-cycle scheme [38], which does not require every link to be covered by a ring. Links on the p-cycle are protected as in the ring cover, but links not covered by p-cycle will be protected by the loop-back paths along the p-cycle that they interconnected with. A generalized loop-back recovery scheme was proposed recently in [55]. The merits of the approach is that a primary digraph consisting of a set of fiber links and wavelengths is backed up by another backup digraph consisting of a set fiber links and wavelengths in the reverse direction of the primary digraph. This scheme enables the distributed operation for link or node failures and is proved simpler and more cost-efficient than the ring cover and p-cycle schemes. The main drawback of the above loopback methods is the high management overhead to maintain the rings or loopback digraphs. They also suffer from high time complexity.

A general link-based protection scheme preplans a disjoint path to connect the two ends of every link and assign wavelengths to this path. We note that the same wavelengths as those in the primary link have to be assigned to the backup path if there are no converters.

Even though the above link-based schemes have the advantage of fast recovery process, they suffer from the low bandwidth usage compared to path-based schemes. From previous studies, we have seen that dedicated path protection usually has the disadvantage of low bandwidth efficiency. Shared backup path protection (SBPP) has proven to be very efficient in spare capacity usage [60, 23]. [4] proves that static establishment outperforms dynamic establishment of protection paths. [21] studies the performance of the SBPP scheme upon dual link failure and optimal sharing. The effect of wavelength converters are considered in [49]. [22] considers the modularity of the capacity provision. As an extension to the shared backup path protection scheme, a shared backup tree can be created to protect a group of primary paths in [78]. [70] proposes a method to minimize the length difference between the protection and backup path using the model of NP-complete disjoint 2-path problem (N(E)D2PP).

In the above approaches, the problems are solved either by an Integer Programming (IP) model or by heuristics. We should keep in mind that the wavelength assignment problem is equivalent to the vertex coloring problem which is generally a NP-complete problem. Therefore the heuristic approach is more suitable for medium or large scale networks.

We also point out that the routing problem in the above general protection schemes is equivalent to link or node disjoint paths problems in graph theory. Sometimes it is called the diverse routing problem [14]. Some recent theoretical development on the disjoint path problem can be found in [26], [39], and [80]. We can also add some constraints on the routing of the primary and backup paths to reflect the QoS requirement. A common constraint is on the path length that can represent the delay constraint. Many of the related problems have been proved NP-complete in graph theory.

A few authors have also considered the effects of survivability schemes to the topology design of WDM networks. [23] studies the dependency on graph connectivity of different mesh protection and restoration schemes through simulation, while [86] studies topology design and upgrade of an optical network by bottleneck-cut identification.

2.3 Light-Tree Routing in All-Optical Networks

In [68, 29], the concept of a lightpath was generalized into that of a *light-tree*, which, like a lightpath, is a clear channel originating at a given source node and implemented with a single wavelength. But unlike a lightpath, a light-tree has multiple destination nodes, hence it is a point-to-multipoint channel. The physical links implementing a light-tree form a tree, rooted at the source node, rather than a path in the physical topology, hence the name. Light-trees may be implemented by employing optical devices known as *power splitters* [56] at the OXCs. A power splitter has the ability to split an incoming signal, arriving at some wavelength λ, into up to m outgoing signals, $m \geq 2$; m is referred to as the *fanout* of the power splitter. Each of these m signals is then independently switched to a different output port of the OXC. Due to the splitting operation and associated losses, the optical signals resulting from the splitting of the original incoming signal must be amplified before leaving the OXC. Also, to ensure the quality of each outgoing signal, the maximum fanout m of the power splitter may have to be limited to a small integer. If the OXC is also capable of wavelength conversion, each of the m outgoing signals may be shifted, independently of the others, to a wavelength different than the incoming wavelength λ. Otherwise, all m outgoing signals will be on the same wavelength λ. Note that, just like with wavelength converter devices, incorporating power splitters within an OXC is expected to increase the network cost because of the large amount of power amplification and the difficulty of fabrication.

Light-trees have several applications in optical networks, including:

- **Optical multicast.** An attractive feature of light-trees is the inherent capability for performing multicast in the optical domain (as opposed to performing multicast at a higher layer, e.g., the network layer, which requires electro-optic conversion). Therefore, light-trees can be useful for transporting high-bandwidth, real-time applications such as high-definition TV (HDTV). We note that TV signals are currently carried over distribution networks having a tree-like *physical* topology; creating a *logical* tree topology (light-tree) over an arbitrary physical topology for the distribution of similar applications would be a natural next step. Because of the multicast property, we will refer to OXCs equipped with power splitters as *multicast-capable* OXCs (MC-OXCs).

- **Enhanced virtual connectivity.** In opaque networks, the virtual degree of connectivity of each node is not tied to the number of its interfaces: electronic routing creates the illusion that a node can reach any other node in the network. In transparent networks, on the other hand, the degree of connectivity of each client node (e.g., IP/MPLS router) connected to the optical core is limited by its physical degree, i.e., the number of its optical transceivers [73]. A light-tree service would enable a client node to reach a large number of other client nodes independently of its physical degree, significantly enhancing the virtual connectivity of the network.

- **Traffic grooming.** Generalized MPLS (GMPLS) [30] makes it possible to tunnel a set of MPLS label-switched paths (LSPs) over a wavelength channel. Since switching at OXCs takes place at the granularity of a whole wavelength, a point-to-point lightpath allows the sharing of the wavelength bandwidth only between clients attached to the same ingress and egress OXCs. The light-tree concept offers a way to overcome this constraint, since it allows for the grooming and tunneling of a number of lower rate point-to-point LSPs to several destinations, regardless of the egress OXC to which these destinations attach.

2.3.1 The Steiner Tree Problem

To make efficient use of bandwidth in point-to-point networks, the typical approach for multicast communication is to build a multicast tree rooted at the source and spanning all the destinations in a given multicast group. Usually, a cost is assigned to each link of the network, and the objective is to determine the tree of minimum cost. This is the famous Steiner tree problem in graph theory [40], which is known to be NP-complete when the multicast group has more than two members [34]. Several heuristics and approximation schemes have been developed for the Steiner tree problem. These algorithms can be categorized roughly into the following three groups:

- **Shortest path-based heuristics (SPH).** The algorithm in [20] initializes the Steiner tree to the shortest path from the source to an arbitrary multicast member. It then repeatedly includes a new member by adding the shortest path between this member to the current partial tree, until all members have joined the tree. Many variants of this algorithm have been developed to improve the quality of the final tree, such as including the members in the order determined by their distance to the multicast tree instead of random inclusion [76], or growing the Steiner tree from the destinations instead of from the source [48].

- **Spanning tree-based heuristics (STH).** The algorithm in [47] first constructs a closure graph of the multicast nodes from the original graph using the cost of the shortest path between each pair of members. A minimum spanning tree of the closure graph is obtained (in polynomial time), and then the shortest paths in the original graph are used to replace the edges of this minimum spanning tree. Finally, the multicast tree is obtained by removing any cycles. This approach yields an approximation algorithm with a ratio of 2.

- **Metaheuristics**. Metaheuristics such as simulation annealing [24], genetic algorithms [28], and Tabu search [63] have been investigated to solve the Steiner tree problem and have been shown to perform well on average.

In practice, the nature of some multicast applications is such that the routing tree must satisfy certain constraints related to physical limitations (e.g., a limited fanout capability) or the desired quality of service (e.g., an upper bound on the end-to-end delay along any path of the tree). Constrained Steiner tree problems are at least as hard as the unconstrained one, and for certain constraints it has been shown that no polynomial-time approximation scheme exists. Several heuristics have been developed to compute constrained trees, most of which are based on the above heuristics for the unconstrained Steiner tree problem. For instance, the KPP algorithm [46] uses an approach similar to the one in [47] to compute an approximate Steiner tree in which the end-to-end delay along any path from the source to a destination node is bounded. The degree-constrained Steiner tree problem, in which it is assumed that some nodes may not support multicast (i.e., they cannot be used as branching points) or have a limited fanout capability, was studied in [12, 16], and appropriate heuristics were proposed. A constrained multicast tree problem in which the objective is to to bound both the end-to-end delay and the delay variation among all source-destination paths was studied in [65]; the total cost of the tree was not considered in that work, but it was shown that constructing such a constrained tree is an NP-complete problem.

2.3.2 Light-Tree Routing

With recent advances in MC-OXC technology [42], it is now possible to envision a future backbone network environment that provides a practical multicast service at the optical layer. Such a service will be implemented by using GMPLS-related protocols to establish light-trees on demand [29, 68]. While the problem of establishing a light-tree that spans a given source and a set of destination nodes bears some similarities to the Steiner tree problem, the nature of optical multicast introduces several new issues and complexities, as we discuss next.

Splitting an optical signal introduces losses, a problem not encountered in electronic packet- or circuit-switched networks, and thus, not addressed by existing routing tree algorithms. Even in the presence of optical amplifiers, this signal loss imposes a hard upper bound on the number of times a signal can be split, as well as on the number of hops that the signal can travel after every split operation. In the absence of wavelength conversion in the network (or even in networks with limited or sparse conversion capability), multicast routing is tightly coupled to wavelength allocation, an issue that does not arise in electronic networks. Also, optical networks may only have a sparse multicast switching capability, i.e., only a subset of the OXCs may be multicast capable. When only a few MC-OXCs are present in the network, a feasible multicast tree may not exist, and therefore the heuristics for degree-constrained multicast developed in [12] are not applicable at all. Finally, the problems of capacity planning of MC-OXCs and multicast routing strongly depend on one another.

Several recent research efforts have aimed to address some of the problems associated with optical multicast and the establishment of light-trees. Wavelength assignment in the presence of multicast has been studied in [69, 43, 52]. Multicast routing algorithms for networks with a sparse light splitting capability have been considered in [84, 79, 82]. To deal with the fact that a feasible multicast tree may not exist for a given source and destination set, the concept of a *light-forest* was developed in [84]. In general, all the multicast routing algorithms for optical networks assume unlimited fanout capacity at MC-OXCs, and each tree of a given light-forest must be assigned a different wavelength. The problem of optimally placing a small number of MC-OXCs in a WDM network has been studied in [2]. Finally, two designs for MC-OXCs have been proposed. The first is based on the splitter-and-delivery architecture [42], while the second is an enhancement of the former that results in better power efficiency [3].

Chapter 3

Topology Design of Large-Scale Optical Networks

We define the topology design problem as determining the minimum number of OXCs and the optimal topology among these OXCs. To model this problem, we first make some fundamental assumptions on the network structure. Then we formulate the problem as an integer programming problem. Since this optimization problem is NP-complete, we focus our study on the development of efficient heuristics. We design a set of heuristic algorithms for generating a biconnected graph, routing lightpaths, and assigning wavelengths. We also describe a genetic algorithm we use for searching the space of physical topologies. To evaluate the performance of our algorithms, we also develop lower bounds for this problem. Through extensive simulation for a large range of different parameters such as the number of wavelengths per fiber, the number of ports per OXC, and the number of transceiver per LSR, many valuable conclusions have been made, the most import one among which is that it is possible to build cost-effective optical backbone networks that provide rich connectivity among large numbers of LSRs with relatively few, but properly dimensioned, OXCs. The results also indicate that our algorithm lead to good solutions.

3.1 Problem Definition

We consider a number N of LSRs that are to be interconnected over an optical backbone network [10] which consists of OXC nodes supporting GMPLS. The service provided by the MPλS network of OXCs is the establishment of lightpaths among pairs of LSRs. We assume that each LSR has Δ optical transceivers, therefore, it may establish at most Δ incoming and at most Δ outgoing lightpaths at any given time. This constraint on the number of simultaneous lightpaths to/from an LSR is due both to optical hardware and cost limitations (reflected in the number of optical transceivers) and the traffic processing capacity of the LSR. We also assume that each fiber link in the network can support at most W wavelengths, and that each OXC has exactly P input/output ports. We let $\alpha, 0 \leq \alpha \leq 1$, be the desired degree of connectivity of the physical topology of OXCs, defined as:

$$\alpha = \frac{E}{M(M-1)/2} \tag{3.1}$$

where E represents the number of fiber links interconnecting the OXCs. Parameter α represents how dense the graph is. For an arbitrary graph, α ranges from 0 to 1, with 0 representing a graph that is totally disconnected and 1 representing a completely connected graph. We note that most of the existing backbone networks have an α value around 0.3.

The fundamental question we address in this chapter is:

> *What is the minimum number M of OXCs required to support the N LSRs, and what is the physical topology of the corresponding MPλS network?*

We believe that the answer to this question is of importance to service providers who need to deploy optical backbone networks in a cost-effective manner. We note that the cost of building an MPλS network will be mainly determined by: **(i)** the cost of the OXCs (including switch hardware and switch controller software), and **(ii)** the cost of (deploying or leasing) the fiber links between OXCs (including the cost of related equipment, such as optical amplifiers). While in our study we directly model only the OXC cost, we note that the fiber cost is indirectly taken into account through the parameter α: because of (3.1), for a given value of α, minimizing the number M of OXCs will also minimize the number of fiber links. Since α is an input parameter in our formulation, we believe that α in combination with the number of OXCs is representative of the overall cost of the MPλS network.

Clearly, in order to determine the number of OXCs in the backbone we need to take into account not only the number N of LSRs but also the traffic requirements (i.e., the number of lightpaths between pairs of LSRs), the survivability properties of the network, etc. In fact, different service providers may well have different and even conflicting requirements for their networks. Rather than trying to account for all possible design requirements, we are instead interested in providing a general framework that can help us answer the above question in a way that can provide practical guidelines for building MPλS networks. We therefore set the following requirements **R1**-**R5** that the MPλS network we design must satisfy. We believe that this list captures the salient features of the network and is sufficiently general to accommodate the requirements of a wide range of network providers.

R1. Each LSR accesses the backbone using two bidirectional fiber links, one to each of two different OXCs; both links are used for carrying traffic to and from each LSR.

R2. The physical topology of the OXCs is 2-connected.

R3. Each LSR maintains 2Δ simultaneous lightpaths to/from other LSRs.

R4. Two neighbor OXCs in the physical topology are interconnected by one bidirectional fiber link.

R5. The OXCs do not have any wavelength conversion capability.

The first two requirements (**R1** and **R2**) ensure that there are at least two edge-disjoint paths between any two LSRs, a necessary condition for a survivable network. **R3**-**R5** can be viewed as worst case requirements. **R3** ensures that the physical topology of the MPλS network can support the maximum number of simultaneous lightpaths (recall that no LSR can have more than Δ outgoing and Δ incoming lightpaths). In particular, we use $\mathcal{L}$ to denote the set of ΔN lightpaths that the network must support. Because of **R4**, the resulting network will use single-fiber links between pairs of adjacent OXCs. Finally, **R5** requires that a lightpath be assigned a single wavelength along all the physical links it traverses. **R4** and **R5** can be easily relaxed, but are included here because we believe that there will exist MPλS networks which will satisfy one or both of these requirements, at least during early deployment. Furthermore, we expect that relaxing either **R4** or **R5** will lead to a topology with a smaller number of OXCs compared to when both are in place, therefore, our results can be used as a worst-case scenario.

3.2 Integer Programming (IP) Formulation

Our objective is to determine the optimal physical topology of OXCs for establishing the given set $\mathcal{L}$ of lightpaths among the N LSRs, under the constraint on the number of wavelength W that can be supported in each fiber. This involves determining the minimum number of OXCs required as well as the links interconnecting the OXCs. Because of the difficulty of this problem, we choose an indirect approach to obtaining a near-optimal physical topology. Specifically, we first assume that the number M of OXCs in the physical topology is given and there is no constraint on the number of wavelengths. Consequently, we consider the problem of determining the links of the physical topology such that the number of wavelengths required to establish the set $\mathcal{L}$ of lightpaths is minimized. We formulate this problem as an integer programming (IP) problem. At the end of this section, we show how the solution to the IP problem can be used to obtain a physical topology with a near-optimal number of OXCs.

Let us start with the assumption that the number M of OXCs is given. Our objective then is to obtain:

1. the set of fiber links interconnecting the M OXC nodes (i.e., the physical topology), and

2. the routing and wavelength assignment for the lightpaths in the set $\mathcal{L}$ among the N LSRs

such that the required number of wavelengths per fiber link in the physical topology is minimized. The inputs to the problem are: the number of OXCs, M; the number of ports in each OXC, P; and the static traffic matrix $[v_{s,d}]$, where $v_{s,d}$ represents the number of lightpaths that have to be established between OXCs s and d, $s, d = 1, 2, \cdots, M$. This traffic matrix is derived directly from the lightpath set $\mathcal{L}$ as follows: if a lightpath needs to be set up from LSR A connected to OXC s and LSR B connected to OXC d, then we increment $v_{s,d}$. Let us define the following parameters:

- Objective function variable W_f. This is the total number of wavelengths needed to route all the lightpaths in the network (i.e., the maximum number of wavelengths used at any link of the physical topology).

- Number of wavelengths per fiber, W_c. This is *not* the same as W defined above. Instead, it is a variable used to set up the IP formulation. It is given a very high value (e.g., 1000) to ensure that the problem has a feasible solution.

- Wavelength usage variables y_w. $y_w = 1$ if wavelength $w, w = 1, \cdots, W_c$, is used in *some* link of the physical topology, zero otherwise.

- Physical topology variables $t_{i,j}$. $t_{i,j} = 1$ if there is a link between OXC i and j in the topology, and zero otherwise.

- Route variables $\lambda_{i,j}^{s,d}(k)$. $\lambda_{i,j}^{s,d}(k) = 1$ if the k-th lightpath from OXC node s to OXC node d uses physical link (i, j), and zero otherwise.

- Wavelength assignment variables $c_w^{s,d}(k)$. $c_w^{s,d}(k) = 1$ if the k-th lightpath from OXC node s to OXC node d uses wavelength w, and zero otherwise.

We can now set up the IP formulation as follows:

Objective function: Minimize $W_f = \sum_w y_w$

Subject to:

Traffic and lightpath constraints:

$$\sum_j \lambda_{i,j}^{s,d}(k) = \left\{ \begin{array}{ll} 1, & i = s \\ 2, & i \neq s \end{array} \right\} \quad \forall\ s, d, k \tag{3.2}$$

$$\sum_i \lambda_{i,j}^{s,d}(k) = \left\{ \begin{array}{ll} 1, & j = d \\ 2, & j \neq d \end{array} \right\} \quad \forall\ s, d, k \tag{3.3}$$

$$\sum_k \sum_j \lambda_{s,j}^{s,d}(k) = v_{s,d} \quad \forall\ s, d \tag{3.4}$$

$$\sum_k \sum_i \lambda_{i,d}^{s,d}(k) = v_{s,d} \quad \forall\ s, d \tag{3.5}$$

Physical topology constraints:

$$\sum_i t_{i,j} \leq p \quad \forall\ j \tag{3.6}$$

$$\sum_j t_{i,j} \leq p \quad \forall\ i \tag{3.7}$$

$$\sum_{i \in S, j \in \bar{S}} t_{i,j} \geq 2 \quad \forall\ S \subset V \tag{3.8}$$

$$t_{i,j} \geq \lambda_{i,j}^{s,d}(k) \quad \forall\, i, j, s, d, k \tag{3.9}$$

Wavelength capacity W_c per fiber constraint:

$$\sum_k \sum_{s,d} \lambda_{i,j}^{s,d}(k) \leq W_c \quad \forall\, i, j \tag{3.10}$$

Wavelength assignment constraints:

$$\sum_k \sum_w c_w^{s,d}(k) = v_{s,d} \quad \forall\, s, d \tag{3.11}$$

$$\sum_k \sum_{s,d} \lambda_{i,j}^{s,d}(k) c_w^{s,d}(k) \leq y_w \quad \forall\, i, j \tag{3.12}$$

Integer constraints:

$$t_{i,j} = 0, 1, \lambda_{i,j}^{s,d}(k) = 0, 1, c_w^{s,d}(k) = 0, 1, \forall\, i, j, s, d, w, k \tag{3.13}$$

Constraints (3.2) and (3.3) ensure that the traffic flowing out of an OXC node equals the traffic flowing into this node. Constraint (3.4) guarantees that the total number of wavelengths on all the links originating from node s for the lightpaths from s to d is equal to the number of lightpath requirements from node s to d. Constraint (3.5) is similar, and makes sure that the total number of wavelengths on all the links sinking to node d for the lightpaths from s to d is equal to the number of lightpaths from node s to d. Constraints (3.6) and (3.7) guarantee that the degree of a node cannot exceed p, the maximum available number of ports in an OXC used to interconnect to other OXCs. p plus the number of ports used for the LSRs attached to this OXC equals the total number of ports P of this particular OXC. Constraint (3.8) provides a necessary and sufficient condition for a 2-connected graph, a common requirement for a reliable communication network; note that V in this expression is the set of OXC nodes. Constraint (3.9) forces the lightpaths to be routed over existing physical links. Constraint (3.10) sets the maximum number of wavelengths per physical link, i.e., the total number of the lightpaths passing through a physical link can not exceed the maximum number of available wavelengths, W_c. Constraint (3.11) assigns a wavelength for every lightpath. Constraint (3.12) allows only one wavelength in a physical link to be used by a particular lightpath; note that this is a non-linear constraint. Finally, the last constraint ensures that all the variables take integer values.

The solution to the above problem gives an optimal topology for M OXC nodes, as well as the optimal lightpath routing and wavelength assignment that minimizes the number W_f of wavelengths used in any link. Let $W_f^\star$ be this optimal number of wavelengths. We note that $W_f^\star$ can be greater than, equal to, or less than the number of wavelengths W actually supported by the fiber links of the MPλS network. If $W_f^\star = W$, then the solution is not only optimal, but it is also feasible given the available number of wavelengths W. However, if $W_f^\star > W$, then this optimal solution is not feasible. In this case, we may have to increase the number M of OXCs that was given as input to the problem. By solving the same problem with a larger value of M, we will obtain a new optimal solution requiring a smaller number of wavelengths. On the other hand, if $W_f^\star < W$, the solution is feasible, but it may also be possible that another solution exist, one in which the physical topology consists of a smaller number of OXCs and which requires no more than W wavelengths. Thus, we can solve the same problem with a smaller value for M as input in the hope of finding such a solution.

The above observations naturally lead to a binary search approach to obtaining an optimal solution that requires no more than W wavelengths *and* minimizes the number of OXCs in the MPλS network. The binary search is illustrated in the overall algorithm shown in Figure 3.3.

3.3 The Heuristic Algorithm

The topology design problem presented in Section 3.1 is NP-complete since the incorporated wavelength assignment subproblem has been proved to be NP-complete. The IP model presented in Section 3.2 can only be solved for very small size networks in reasonable time because the number of variables and constraints increase much faster than the size of the network. The non-linear constraint (3.12) has made the IP model more difficulty. In this section, we present a set of heuristic algorithms for this problem that can be applied to medium and large size networks. We divide the problem into the following tasks: (i) generation of a feasible physical topology that satisfies all of the constraints (3.2)-(3.13), (ii) routing of lightpaths, and (iii) assignment of wavelengths to lightpaths. Each of these tasks is solved using a heuristic algorithm. We then use a genetic algorithm (GA) to generate additional feasible physical topologies, and we iterate in order to obtain a near-optimal solution with a minimum total number of wavelengths.

3.3.1 Generation of a Random Feasible Physical Topology

Recall that a feasible OXC network is at least 2-connected. In order to get a feasible topology, we first generate a random tree, then grow a 2-connected graph from it. If we number the leaves of the tree as $i, \cdots, H$, we can sequentially connect pairs of leaves with edges $(i, i+1), \cdots, (H-1, H)$ to obtain a 2-connected graph. Thus, the only question is how to generate a random tree of M nodes, where each node represents one of the OXCs.

We have adopted the method from [58] to generate a random tree of M nodes. Each tree of M nodes has an one-to-one relationship with a Prufer number that has $(M-2)$ digits. The digits are integers between 1 and M. Consider a tree $\mathcal{T}$ of M nodes numbered 1 to M in some manner. To obtain its Prufer number $P(\mathcal{T})$, we start with a null Prufer number (one with no digits) and we repeat the following steps to build $P(\mathcal{T})$ by appending one digit at a time to the right of the current Prufer number. Let i be the lowest numbered leaf in the tree, and let j be the parent of i. Then, j becomes the rightmost digit of $P(\mathcal{T})$. We remove i and edge (i, j) from $\mathcal{T}$. If i was the only child of j, then j becomes a leaf. If only two nodes remain in the tree, we stop; $P(\mathcal{T})$ has been formed. Otherwise, we repeat the above process with the new lowest numbered leaf.

The reverse process of obtaining a tree from a Prufer number $P(\mathcal{T})$ with $(M-2)$ digits in the range 1 to M consists of these steps. First, designate all nodes whose number does not appear in $P(\mathcal{T})$ as *eligible*. Let i be the lowest numbered eligible node, and let j be the leftmost digit of $P(\mathcal{T})$. Add the edge (i, j) to $\mathcal{T}$ and remove the leftmost digit j from $P(\mathcal{T})$. Designate i as no longer eligible. If j does not occur anywhere in what remains of $P(\mathcal{T})$, then designate j as eligible. Repeat the above procedure until no digits remain in $P(\mathcal{T})$, in which case there are exactly two nodes, say, i and j, still eligible. Add (i, j) to $\mathcal{T}$ and stop. Since there are exactly $(M-2)$ digits in the Prufer number and we remove a digit each time we add an edge to $\mathcal{T}$, the final graph $\mathcal{T}$ has exactly $M-1$ edges; it is also shown in [58] that $\mathcal{T}$ has no cycles, thus it is a tree.

As an example, suppose that $M = 8$, and that the Prufer number is 666585 (consisting of $M-2=6$ digits in the range 1-8). Nodes 1,2,3,4, and 7 do not appear in $P(\mathcal{T})$ and are designated as eligible for consideration. Node 1 is the lowest numbered eligible node, and digit 6 is the leftmost digit of the Prufer number. Consequently, we add the edge

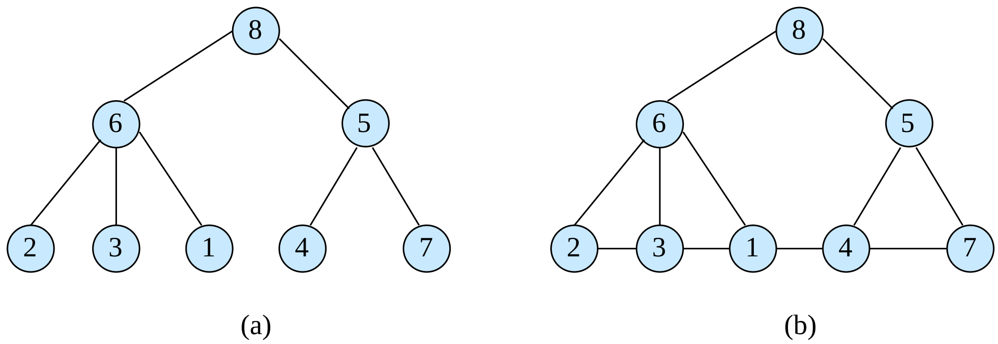


Figure 3.1: (a) Tree of 8 nodes corresponding to the Prufer number 666585, (b) the resulting 2-connected graph

(6,1) to the tree, we make node 1 ineligible, and we remove the leftmost digit 6 from the Prufer number. The remaining Prufer number is thus 66585, and since digit 6 appears in it, node 6 is not made eligible. In the second step, node 2 is the lowest numbered eligible node and digit 6 is the leftmost digit in the remaining Prufer number. We add edge (6,2) to the tree $\mathcal{T}$, we make node 2 ineligible, and we remove digit 6 from the Prufer number. The number which remains is 6585, so node 6 remains ineligible. In the third iteration, node 3 is the lowest numbered eligible node and the leftmost digit in the Prufer number is once again digit 6. Thus, we add edge (6,3) to the tree, we make node 3 ineligible, and we remove digit 6 to obtain the new number 585. Since the digit 6 does not appear in the new number, we make node 6 eligible. We continue in this manner, and we add edges (5,4), (8,6), and (5,7) to the tree, at which point the Prufer number becomes null (all its digits have been removed). At this stage there are only two nodes eligible, nodes 5 and 8. We next add edge (8,5) to obtain the tree shown in Figure 3.1(a).

Based on the above discussion, the following steps summarize the algorithm for generating a random feasible physical topology:

1. Given the number M of OXCs, randomly generate $(M-2)$ digits in the range of 1 to M to form $P(\mathcal{T})$.

2. Generate a tree $\mathcal{T}$ using the Prufer number, as above.

3. Construct a 2-connected graph by adding edges sequentially to connect the leaves of the tree $\mathcal{T}$.

After connecting the leaves of the tree in Figure 3.1(a), we get the 2-connected graph shown in Figure 3.1(b).

3.3.2 Routing and Wavelength Assignment

We now assume that we are given a 2-connected physical topology of M OXCs, as well as a set of lightpaths between pairs of OXCs that need to be established. We now present two algorithms, one to route each lightpath over a physical path of fiber links, and another to assign wavelengths to the lightpaths. Note that we treat the routing and wavelength assignment subproblems independently; this approach may require a larger number of wavelengths than a combined solution, but the latter is intractable while our approach can be applied directly to networks of realistic size. Furthermore, as we shall see, the routing algorithm takes into account the number of lightpaths using each link in order to minimize the number of wavelengths needed.

We use Dijkstra's shortest path algorithm to route the set of lightpaths over the given physical topology. In order to minimize the number of wavelengths used on a physical link, we use two heuristic approaches. First, the link weight used in Dijkstra's algorithm is dynamically adjusted to reflect the number of wavelengths already allocated on each link. Consider physical link ℓ and let C_ℓ be the actual link cost (C_ℓ is a constant) and w_ℓ be the number of lightpaths already using this link. Then, each time we run Dijkstra's algorithm to find a path for a certain lightpath, we use the quantity $L_\ell = C_\ell + Hw_\ell$ as the cost of link ℓ, where H is a tunable weight parameter. This cost function forces new lightpaths to be routed over less congested links in the physical topology, reducing the total number of wavelengths used in the network. For the next application of the algorithm, quantities w_ℓ are incremented for all links ℓ along the path of the just routed lightpath.

The second heuristic approach has to do with the order in which we consider the given lightpaths for routing. Specifically, we first sort the OXC nodes in an ascending order according to their degree (ties are broken arbitrarily). Starting with the first OXC node (the one with the smallest degree), we apply Dijkstra's algorithm to route all lightpaths that have this node as source or destination. We proceed in this manner by considering nodes with higher degrees. This method results to a considerably lower wavelength usage than when selecting the nodes randomly. This is because nodes with smaller degrees have fewer alternative links to route their lightpaths. In view of this, routing lightpaths originating or terminating at these nodes first will increase the wavelength use of their links. Because of the cost function described above, later lightpaths will tend to avoid the links around these nodes. On the other hand, if a node with a small degree was considered late in the process, its lightpaths would have to use one of its links regardless of how congested these links were, potentially increasing the overall number of wavelength required. We refer to this scheme as the smallest-degree-first-routing (SDFR) algorithm. The following steps summarize the heuristic algorithm for routing lightpaths:

1. Use the link weight function $L_\ell = C_\ell + Hw_\ell$.
2. Sort the nodes in an ascending order of their degree.
3. Consider each node in this order and use the Dijkstra's algorithm to build the shortest path for its lightpaths.

Once the physical links for each lightpath have been obtained, we need to assign wavelengths such that if two lightpaths share the same link then they are assigned a different wavelength. This wavelength assignment problem can be shown to be equivalent to the vertex coloring problem of an induced simple graph [85]. The induced graph is such that its vertices correspond to lightpaths in the original network, and vertices of the induced graph are linked by an edge only if the two corresponding lightpaths share the same physical link. A heuristic algorithm was developed in [85] to solve the vertex coloring problem. The algorithm uses a greedy approach to assign wavelengths (color) to the lightpath (vertex). We adopt this algorithm to perform wavelength assigned, since it has been shown to have good accuracy and to run in polynomial time. We also note that the upper bound of the number of distinct colors (wavelengths) used is equal to the maximum degree of the induced connection graph plus one.

3.3.3 The Genetic Algorithm (GA)

In the last decade, genetic algorithms (GAs) [36, 72] have proved to be a practical and robust optimization and search tool. These algorithms are based on the mechanisms of evolution and natural genetics that lead to the survival of the fittest by the process of natural search and selection. A GA generates a sequence of populations using a selection mechanism, and then applies crossover and mutation as search mechanisms. A GA is a global random search technique in the solution space of the problem and it usually avoids entrapping into a local optimization; for further details on the properties of GAs the reader is referred to [36, 72] and the references therein. The steps involved in a GA are as follows:

1. Design an efficient encoding scheme (chromosome) of the solution. Usually, this is a one-to-one bit string mapping to a solution.

2. Generate an initial set of feasible solutions. This set is referred to as a *generation*. The population size of the generation is determined by the variable G_s. This initial generation becomes the current generation.

3. If the stop criterion has been met, return the best solution in the current generation as the near-optimal solution and stop; otherwise, continue to Step 4 for another iteration. Usually the stop criterion is a predetermined number of iterations.

4. Randomly select solutions from the current generation based on their fitness values given by the value of the objective function. The selection probability for a solution is generally proportional to its fitness value. The selected solutions form the basis of the *offspring* generation.

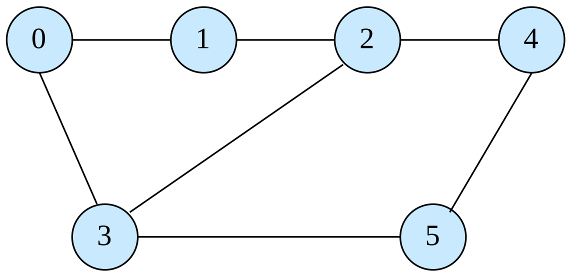
5. Perform crossover on the selected solutions. This step includes picking up pairs of strings at random, randomly choosing a crossover point, and switching the two strings after the point. This crossover is controlled by the *crossover rate* R_c. The algorithm invokes crossover only if a randomly generated number is less than R_c.

6. Perform mutation on the solutions obtained from Step 5, i.e., flip the bits of each string. Mutation is controlled by the *mutation rate* R_m, which is the probability that a bit will be flipped.

7. Calculate the fitness value for the new individuals if the new individual solutions are feasible. If an individual is not feasible, it is either made feasible or it is dropped.

8. Repeat Steps 4-6 until a new generation of the size $(G_s - 1)$ is generated. Set this as the current generation. Add the best solution from the last generation to the current generation. Go back to Step 3.

We use a GA algorithm to generate feasible physical topologies starting from the initial physical topology we obtained in Section 3.3.1. The objective is to search for physical topologies that will improve on the number of wavelengths required to establish



Adjacency Matrix

	0	1	2	3	4	5
0	0	1	0	1	0	0
1	1	0	1	0	0	0
2	0	1	0	1	1	0
3	1	0	1	0	0	1
4	0	0	1	0	0	1
5	0	0	0	1	1	0

Figure 3.2: (a) A physical topology with six nodes, (b) its adjacency matrix

the given set of lightpaths. Since we assume that the number M of OXCs is given, we define our solution space as the set of feasible physical topologies on M nodes that satisfy the constraints (3.2)-(3.13) given in Section 3.2. The fitness value of a feasible individual is the objective function W_f in our IP model (actually, it is the value of W_f obtained after we apply the heuristic routing and wavelength assignment algorithms of Section 3.3.2). We now proceed to describe the encoding of the solution, the calculation of the fitness value, the selection, crossover, and mutation strategies, and the handling of infeasible solutions.

Encoding. Consider a graph with M vertices. We number the vertices from 0 to $M-1$ and the edges of the graph from 0 to $\frac{M(M-1)}{2} - 1$. We read the string from left to right. Each edge (i, j) is numbered using an index k, which is defined according to the index of the two endpoints of the edge, i and j. Specifically, $k = \frac{M(M-1)}{2} - \frac{(M-i)(M-i-1)}{2} + j - i - 1,\ 0 \leq i < j < M$. The solution is encoded into a chromosome (i.e., a bit string of length $M(M-1)/2$) as follows. Each edge is represented by a bit in the bit string. If the edge exists, the bit is set to 1; otherwise it is set to 0. The position of the bit representing an edge is given by the edge index k.

As an example, let us consider the graph shown in Figure 3.2(a). The code of this graph is 10100 1000 110 01 1, which is simply the concatenation of the five rows above the diagonal of the adjacency matrix for this graph, shown in the shaded area of Figure 3.2(b).

Fitness value. This is the value of the objective function of the current solution (physical topology), i.e., the total number of wavelengths used in the network. This value is obtained by running the routing and wavelength assignment heuristic algorithms on the current physical topology.

GA parameters. The performance of a GA is determined by the right choice of control parameters: the crossover rate R_c, the mutation rate R_m, and the population size, G_s. In order to shorten the running time of the algorithm, we choose a relatively small population size and therefore a high level of string disruption. We set $R_c = 0.8$, $R_m = 0.1$, and $G_s = 25$. We use the following operations to generate an offspring from the parent generation until we get a new generation of the population size G_s. These operations work on the edge string encoding a graph.

- *Selection:* select two candidates from the parent generation using the roulette wheel selection mechanism. The individual with the highest fitness value (the total number of wavelengths) has the least probability to be selected.

- *Crossover:* If the random number we pick up is smaller than R_c, we crossover the two candidates obtained by the above selection operation. For example, let us assume that the two selected individuals from the parent generation are 1100001101 and 0110110011, and that the crossing point is the sixth digit. In this case, we obtain the following two new codes: 1100000011 and 0110111101.

- *Mutation:* We mutate bit-by-bit the two individuals obtained from the above crossover operation. For every bit of the string, we pick up a random number and flip this bit only if the random number is smaller than R_m. This results to two offspring. For example, after mutating the above two codes we may obtain: 1101000011 and 0111111101.

Infeasible individuals. It is possible that an offspring generated by the above process may correspond to an infeasible graph. That is, the graph may not be 2-connected or it may exceed the constraint on the connectivity α. Therefore, we run the following four tests for each generated offspring:

1. Check the nodal degree of each node to make sure that the constraint on the number of ports per OXC is satisfied.

2. Check if the total number of edges of the generated graph exceeds the given number of edges decided by the connectivity α.

3. Check the nodal degree to see if every node has a degree more than 2.

Complete Algorithm for Designing an MPλS Network

Input: The number N of LSRs, the set $\mathcal{L}$ of lightpaths, the number P of ports/OXC, the number W of wavelength/fiber, the connectivity α, and the number of generations G.

Output: A physical topology of M OXCs, and the RWA for the set $\mathcal{L}$ of lightpaths.

1. begin
2. $M_{max} \leftarrow N/2$, $M_{min} \leftarrow 1$ // Initialize range for number of OXCs
3. $num\_gen \leftarrow 0$ // Initialize the number of generations
4. $M \leftarrow (M_{max} + M_{min})/2$ // Binary search
5. $phy\_top \leftarrow$ initial physical topology generated using the algorithm in Section 3.3.1
6. For the physical topology phy_top:
7. Solve the RWA problem using the algorithms in Section 3.3.2
8. Calculate the objective function (i.e., fitness value) W_f
9. While $num\_gen < G$: // Genetic algorithm
10. $num\_gen \leftarrow num\_gen + 1$
11. Use the algorithm in Section 3.3.3 to create a new generation of physical topologies
12. For each individual phy_top in the new generation calculate the fitness value W_f using Steps 7 and 8
13. $W^\star \leftarrow$ smallest fitness value W_f from the last generation
14. $phy\_top^\star \leftarrow$ physical topology with smallest fitness value $W^\star$
15. If $W^\star > W$ then
16. $M_{min} \leftarrow M$; go back to Step 3
17. Else if $W^\star < W - 5$ then
18. $M_{max} \leftarrow M$; go back to Step 3
19. Else return M, $phy\_top^\star$
20. end of the algorithm

Figure 3.3: Algorithm for the design of MPλS networks

4. If the first three checks are successful, run a 2-connectivity check procedure that is based on the construction of the Depth-First-Search tree for the graph. Steps 3 and 4 combine to check if the network is 2-connected.

If tests 1 or 2 fail, then we drop the individual. If tests 3 or 4 fail, the graph is not 2-connected. We add edges between the disconnected OXCs until the graph becomes 2-connected. If the solution passes the above four tests, we run the RWA heuristics to obtain the minimum number of wavelengths needed in the network.

The complete algorithm is shown in Figure 3.3. For a particular value of the number M of OXCs, the complexity of the algorithm is dominated by Steps 6-18. In particular, the `while` loop in Steps 9-12 is executed G times, where G is the number of generations in the GA. For each of the G_s individuals of a generation, Steps 7 and 8 must be performed. Step 7 involves running Dijkstra's algorithm to find shortest paths for all $\mid \mathcal{L} \mid = \Delta N$ lightpaths, and running the wavelength assignment heuristic described in Section 3.3.2 on the resulting virtual topology; Step 8 takes constant time once Step 7 has completed. Let M be the number of OXCs for a given iteration of the binary search starting at Step 4. Dijkstra's algorithm takes time $O(\Delta N M^2)$, while the wavelength assignment heuristic we have implemented takes time $O(\Delta^2 N^2)$. Thus, Steps 7 and 8 for a single individual of a single generation when the number of OXCs is M takes time $O(\Delta N M^2 + \Delta^2 N^2)$. Accounting for all G generations and all G_s individuals of a generation, the time taken by the GA for a single value M of the number of OXCs is $O(GG_s(\Delta N M^2 + \Delta^2 N^2))$.

Since the wavelength assignment heuristic takes a considerable amount of time, in order to decrease the computation time, we have made the following modification to the algorithm in Figure 3.3. During the first $G - 1$ iterations of the GA algorithm, we use an upper bound in Step 7 of the algorithm to calculate the number of wavelength needed. This upper bound is easily calculated as it is the maximum degree of the induced connection graph plus one. In the last iteration we use The heuristic wavelength assignment algorithm of Section 3.3.2 is used only in the last iteration. This modified algorithm takes time only $O(G_s(G\Delta N M^2 + \Delta^2 N^2))$. Our experiments have shown that the significantly faster computation time of the modified algorithm has little effect on the accuracy of the results.

3.4 Lower Bounds on the Number of OXCs

We now present a lower bound on the number M of OXCs required given: the number N of LSRs, the number W of wavelengths, the degree of connectivity α of the optical network, the number P of ports of each OXC, and the number Δ of optical interfaces at each router.

We can obtain a lower bound simply by counting the number of OXC ports required in the network. Recall that E denotes the number of links in the optical network. Now, $2N$ OXC ports are needed to connect the N LSRs to the optical network (since each LSR is connected to two OXCs), while $2E$ ports are needed for the fiber links interconnecting the OXCs. Since the total number of OXC ports is MP, we have that:

$$MP \geq 2N + 2E = 2N + \alpha M(M-1) \tag{3.14}$$

Solving this equation for M, we obtain:

$$M \geq \left\lceil \frac{P + \alpha - \sqrt{(P+\alpha)^2 - 8\alpha N)}}{2\alpha} \right\rceil \tag{3.15}$$

where input parameters N, P, and α should be such that $(P+\alpha)^2 \geq 8\alpha N$.

The above result only considers the constraint on the number of ports per OXC and the degree of connectivity of the OXC network, and, it is proportional to the value of α, i.e., the larger the α is, the higher the lower bound is. In the algorithm we present in last section, we only set an upper bound for *alpha*. So this bound is not valid without a lower bound on α.

We now present another lower bound that takes the available number W of wavelengths into consideration and gets rid of α in the expression. Recall from requirement **R3** that the network must support $N\Delta$ lightpaths, where Δ is the number of optical interfaces at each of the N LSRs. Let D denote the average number of hops (fiber links traversed) over all lightpaths. Then, we have that:

$$N\Delta D \leq 2WE \tag{3.16}$$

where $2WE$ is the total number of link-wavelengths in the network, assuming that each fiber link between any pair of OXCs consists of two unidirectional fibers. Combining this result with equation (3.14), we obtain:

$$M \geq \frac{N\Delta D}{WP} + \frac{2N}{P} = \frac{N}{P}\left(\frac{\Delta D}{W} + 2\right) \tag{3.17}$$

From (3.17) we see that the lower bound on M depends on the lower bound on the average hop length D over all lightpaths. An immediate lower bound on D is 1, so we obtain a second lower bound on M as:

$$M \geq \frac{N}{P}\left(\frac{\Delta}{W} + 2\right) \tag{3.18}$$

To obtain a better lower bound on D, we note that, in a network with M nodes, at least $M/2$ nodes are at distance $\lfloor \log_d M/2 \rfloor$ or more from any given node [59], where d

is the maximum nodal degree. In our case, every OXC has P ports, among which $2N/M$ are used to interconnect the LSRs assigned to this OXC, so the maximum nodal degree in the core OXC network is $d = P - 2N/M$. Also the lower bound 1 is still valid for the left nodes. Therefore, we have $D \geq \frac{\log_d M/2+1}{2}$. If we use this value of D as the average hop length in equation (3.17), we have that:

$$W(MP - 2N) \geq N\Delta \frac{\log_d \frac{M}{2} + 1}{2}, \quad d = P - \frac{2N}{M} \tag{3.19}$$

Using iteration over M from 0 to N on both sides of the above inequality until it is satisfied, we can obtain the minimum feasible value of M.

We note that $M/2$ is generally greater than d for middle- or large-size networks, in which case $\frac{\log_d M/2+1}{2}$ is greater than 1. Therefore, the lower bound obtained from 3.19 is generally better than that obtained from 3.18. In the computation, we take the maximum value obtained from expressions 3.18 and 3.19.

3.5 Numerical Results

In this section we present results that illustrate how the different design parameters affect the number of OXCs required to interconnect a set of LSRs. We vary the design parameters as follows: the number N of LSRs is varied between 100-1000, the number W of wavelengths per link takes the values 32, 64, 128, the number P of ports per OXC varies from 16-64, and the number Δ of transceivers per LSR takes the values 4-24. We also set the upper bound on the degree of connectivity α of the MPλS network of OXCs to 0.4. We have assigned the N LSRs to the OXCs in a round-robin manner. That is, the first LSR is attached to the first and second OXC, the second LSR to the third and fourth OXC, and so on (recall that by requirement **R1** each LSR must attach to at least two OXCS). This assignment is made for convenience only, and is not inherent to our approach; in fact, the algorithm in Figure 3.3 can accommodate any arbitrary assignment of LSRs to OXCs. The set $\mathcal{L}$ of lightpaths is chosen so that each LSR has exactly Δ incoming and Δ outgoing lightpaths (see requirement **R3**) to a random set of other LSRs. Again, however, our algorithm will accommodate any set of lightpaths.

We should point out that the above ranges of the design parameters are based on realistic assumptions regarding the state of the technology and the size of the MPλS networks, and go far beyond the small networks to which previous virtual topology algorithms were limited. The results presented here illustrate that our algorithm can be applied to networks of size between one and two orders of magnitude greater than that of the networks studied previously. We also emphasize that the algorithm in Figure 3.3 computes not just the number of OXCs, but also the physical topology of the MPλS network (i.e., the physical links between the OXCs), as well as the routing and wavelength assignment for the set $\mathcal{L}$ of lightpaths between the LSRs. However, due to the large size of the resulting networks, it is not possible to draw the physical topology of fiber links or the logical topology of lightpaths here.

We first take a look at the performance of the GA algorithm we use to find the optimal physical topology under a certain number of wavelengths in Figure 3.4. It shows the results of the last iteration on the number of OXCs (the optimal ones). The number of LSRs is 600, 700, and 800 and the number of wavelengths is 32. We can see that the algorithm converges after 6 or 7 iterations to the wavelength constraint, which actually represents a current good physical topology of the network on the given number of OXCs.

In Figure 3.5 we plot the number M of OXCs in the MPλS network against the number N of LSRs. Three plots are given for three different values of the number of wavelengths per link, $W = 32, 64, 128$. For these results, we have let $\Delta = 12$ and $P = 64$. We make two important observations. First, the number M of OXCs increases almost linearly with the number N of LSRs, but the slope of the curves is moderate. In particular, an increase by a factor of ten in the number of LSRs results in an increase in the number of OXCs by a factor between four (for $W = 32$) and seven (for $W = 128$). Since we have kept the degree of connectivity at around 0.4 for all physical topologies, the corresponding increase in the number of links in the topology, as N increases, is similar. This result implies that OXC networks to interconnect very large number of LSRs can be built cost-effectively. The second observation is that the larger the number of wavelengths available at each fiber link, the smaller the number of OXCs required for a given number N of LSRs. This result is expected, however, we note that a two-fold increase in the number of wavelengths (from 32 to 64) reduces the number of OXCs by less than 1/2 (between 44% for $N = 100$ and 41% for $N = 1000$). We also see the effect of diminishing returns, since a second two-fold increase in the number of wavelengths (from 64 to 128) results in a smaller reduction in the number of OXCs (between 40% for $N = 100$ and 13% for $N = 1000$). Note that, for a given value of N, the number of lightpaths in the set $\mathcal{L}$ remains constant at ΔN across the three curves in Figure 3.5, so one would expect a larger decrease in the number of OXCs as the number of wavelengths increases. However, recall that $2N$ OXC ports are needed to attach the LSRs to the OXC network. Since the number of ports per OXC remains constant at $P = 64$, as W increases, the number of OXCs needed is constrained by the number of ports required rather than the number of lightpaths that need to be established. In other words, in order to take full advantage of the larger number of wavelengths in the fiber, OXCs with a larger number of ports must be employed.

In Figure 3.6 we fix the number of wavelengths to $W = 64$ and the number of ports per OXC to $P = 64$, and we plot the number of OXCs against the number N of LSRs. Three curves are shown, one for a different value of the number of transceivers per LSR, $\Delta = 4, 8, 12$. Note that the curve for $\Delta = 12$ is identical to the middle curve of Figure 3.5 for $W = 64$, although the scales in the two figures are different. Again, we see that the number of OXCs increases linearly with the number of LSRs. Recall that the number ΔN of lightpaths that must be established for a given value of N increases linearly with Δ. However, the curves in Figure 3.6 show that the number of OXCs for a given N value needed to support the larger number of lightpaths increases much more slowly than Δ. For instance, for $N = 100$, 6 OXCs are needed for $\Delta = 4$, 8 OXCs for $\Delta = 8$, and 10 OXCs for $\Delta = 12$. For $N = 1000$, the corresponding number of OXCs are 36, 42, and 46, respectively. These results indicate that a relatively small incremental cost (in terms of additional OXCS and fiber links to interconnect them) can provide a significantly richer connectivity among the LSRs.

In Figure 3.7 we plot the number of OXCs against the number of wavelengths when the number of LSRs is constant at $N = 300$ and the number of ports per OXC is $P = 64$. Three curves for different numbers of transceivers per LSR are shown, $\Delta = 4, 8, 12$. The results are as expected. Specifically, the number of OXCs needed decreases as the number W of wavelengths per fiber increases, but the curves flatten out once $W > 80$. As we mentioned above, this reflects the fact that a larger number of ports per OXC is needed to take full advantage of the large number of wavelengths. Also, more OXCs are required as Δ increases, but the results are consistent with the previous figure in that the increase in the number of OXCs is significantly slower than the increase in Δ (and the corresponding increase in the number of lightpaths to be established). Similar observations can be made from Figure 3.8, where we let $N = 300$ and $P = 64$, and we plot the number of OXCs against the number Δ of transceivers per LSR. Note that Δ (and, consequently,

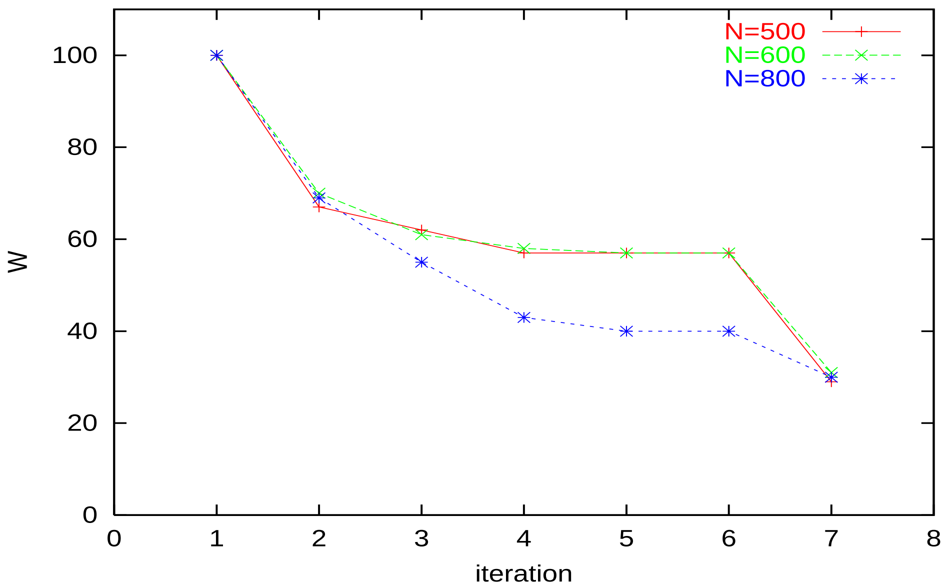


Figure 3.4: The performance of the GA

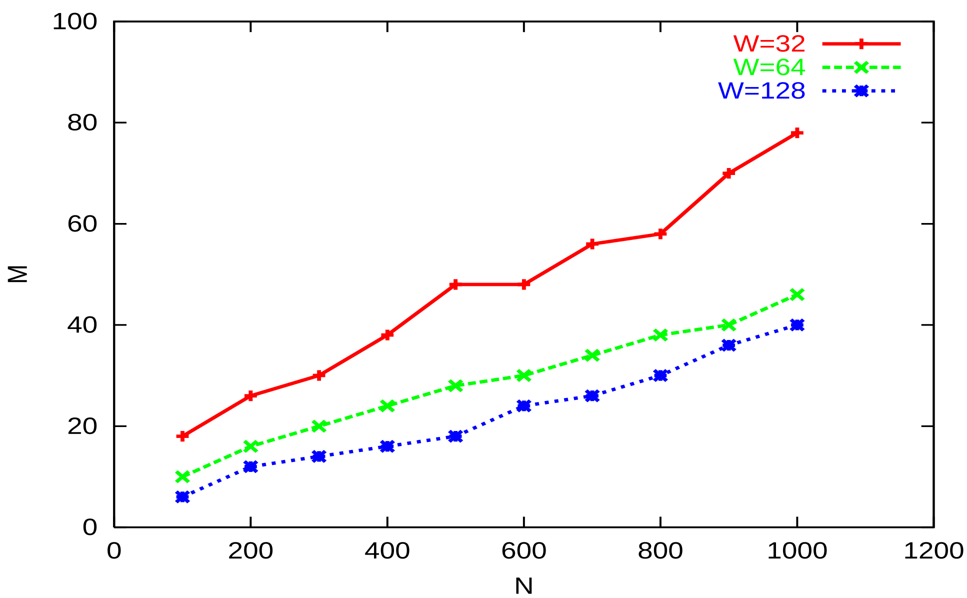


Figure 3.5: No. OXCs in the physical topology ($\Delta = 12$, $P = 64$)

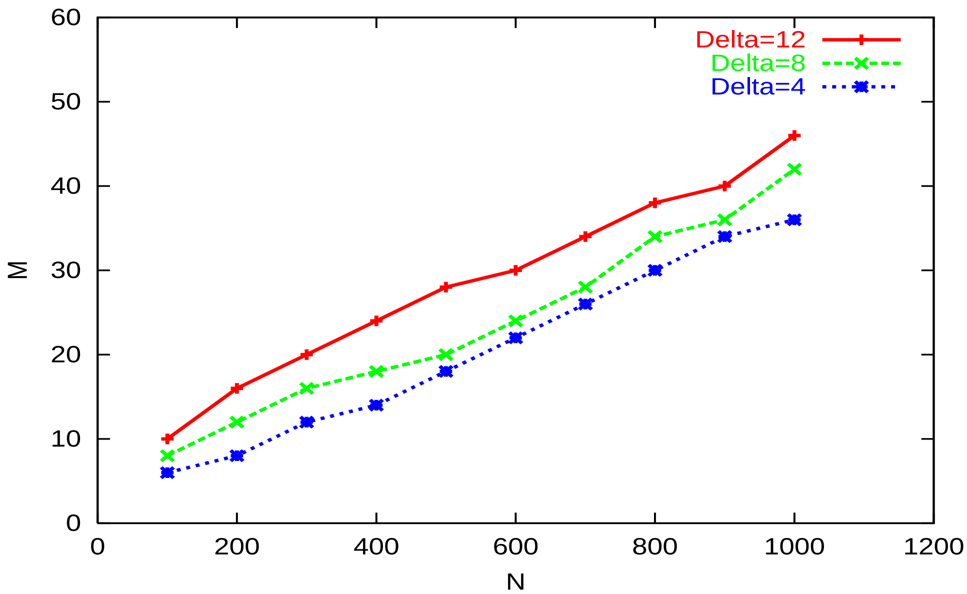


Figure 3.6: No. OXCs in the physical topology ($W = 64$, $P = 64$)

the number of lightpaths) increases by a factor of six from 4 to 24, the number of OXCs required increases much slower, from 20 to 48 when $W = 32$. In fact, the increase in the number of OXCs is even slower for larger number of wavelengths, from 10 to only 20 for $W = 128$.

In Figure 3.9 we let $N = 300$ and $W = 32$ and we plot the number of OXCs against the number P of ports, for three different values of Δ, $\Delta = 4, 8, 12$. There is a sharp drop in the number of OXCs initially, as P increases from 16 to 24, but the curves level off after that. Also, we again see that the value of Δ does not significantly affect the number of OXCs. These results indicate that by employing OXCs of medium size (in terms of P) can have a dramatic effect in the number of OXCs required. While OXCs with many ports are expected to be more expensive than those with few ports, the dramatic drop in the curves of Figure 3.9 indicates that it may be cost-effective to employ the former; also, for the same degree of connectivity α, fewer OXCs implies significant savings in fiber links.

Finally, in Figures 3.10 and 3.11 we compare the results of our heuristic algorithm to the lower bound we presented in Section 3.4. Figure 3.10 plots the middle curve of Figure 3.5 and the corresponding lower bound; thus, these plots correspond to the following values of the input parameters: $\Delta = 12$, $P = 64$, and $W = 64$. As we can see, the

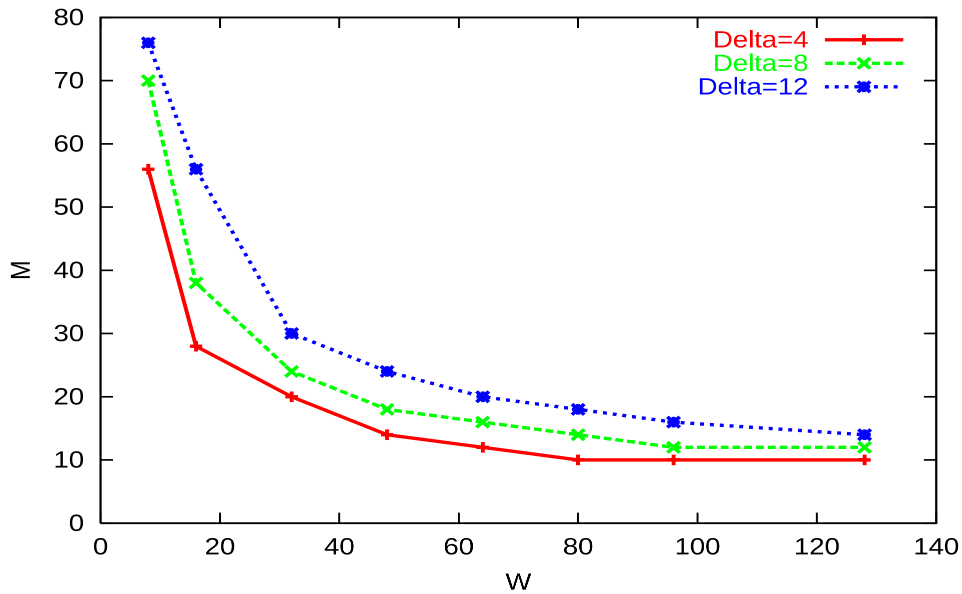


Figure 3.7: No. OXCs in the physical topology ($N = 300$, $P = 64$)

two curves have very similar behavior, and the results from our heuristic are not far away from the lower bound. We emphasize that the lower bound is obtained simply by counting the number of network resources that are absolutely necessary to support the given set of lightpaths. In other words, there is no guarantee that there exists a physical topology with a number of OXCs equal to the lower bound, such that it is possible to route and assign wavelengths to the given set of lightpaths. Therefore, the optimal number of OXCs lies somewhere between the two curves in Figure 3.10.

Figure 3.11 is similar to Figure 3.10, but it plots the bottom curve of Figure 3.6 and the corresponding lower bound, for $W = 64$, $P = 64$, and $\Delta = 4$. Again, we see our heuristic returns a number of OXCs that is close to the lower bound, and this result is consistent across the range of the number N of LSRs shown in the figure. Very similar results regarding the relative performance of our heuristic and the lower bound have been obtained for a wide range of the input parameters.

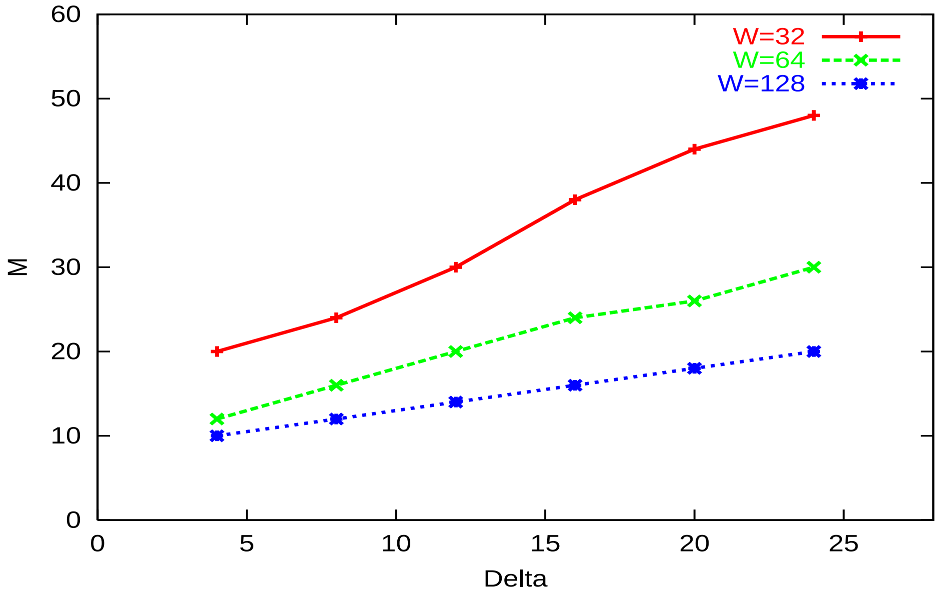


Figure 3.8: No. OXCs in the physical topology ($N = 300$, $P = 64$)

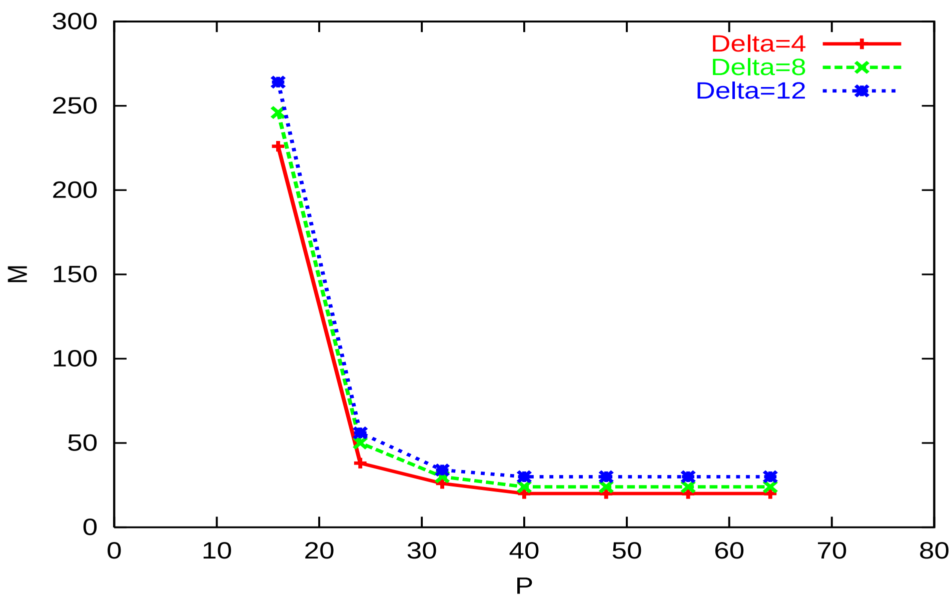


Figure 3.9: No. OXCs in the physical topology ($N = 300$, $W = 32$)

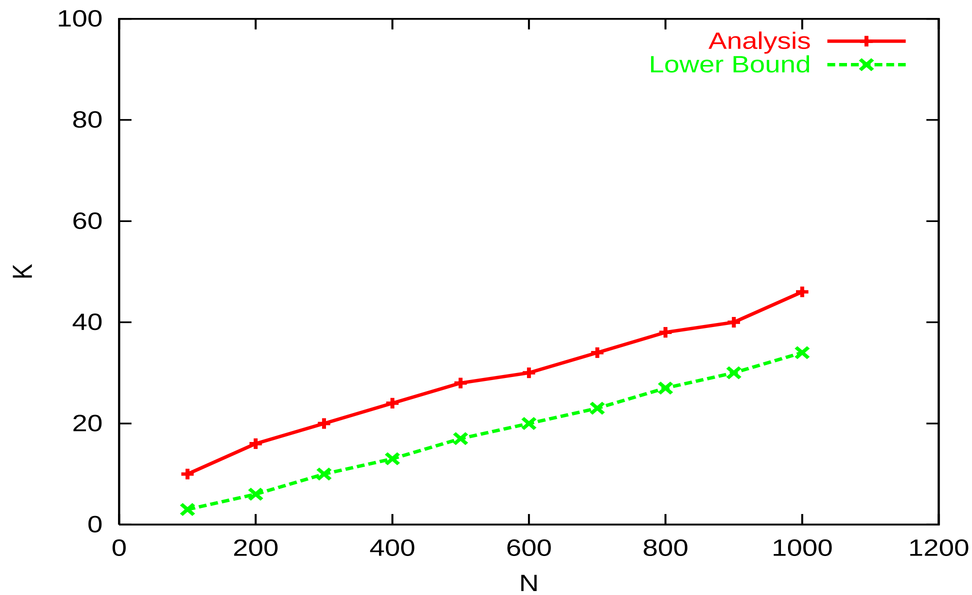


Figure 3.10: Heuristic vs. lower bound ($\Delta = 12$, $P = 64$, $W = 64$)

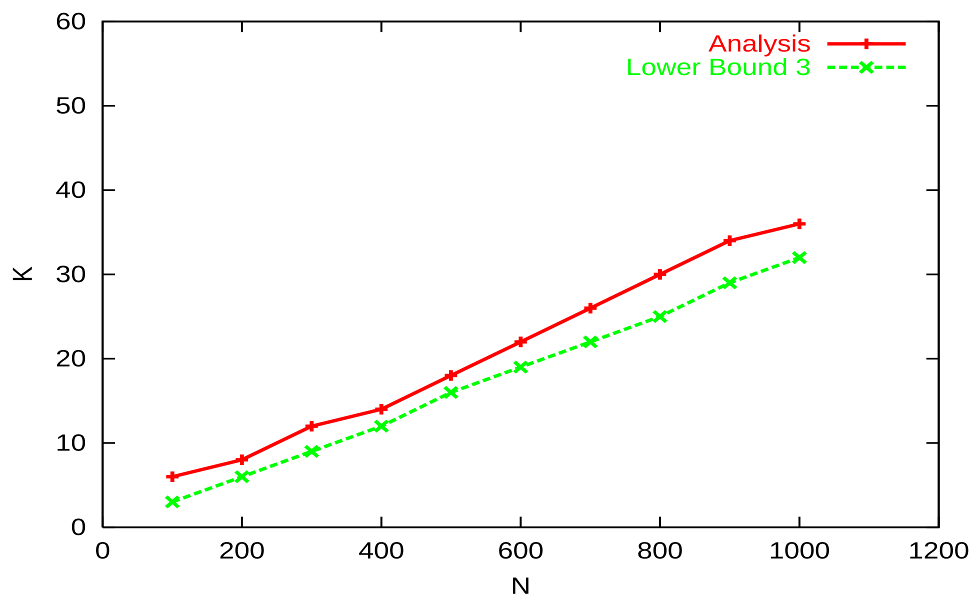


Figure 3.11: Heuristic vs. lower bound ($W = 64$, $P = 64$, $\Delta = 4$)

3.6 Concluding Remarks

We have described a set of heuristic algorithms for the physical and logical topology design of large-scale optical networks of OXCs. Our objective has been to minimize the number of OXCs given a constraint on the number of wavelengths per fiber link and certain constraints (i.e., biconnectivity) on the physical topology. We presented routing and wavelength assignment heuristics, as well as a genetic algorithm to iterate over the set of physical topologies. We have applied our algorithms to design networks that can accommodate hundreds of LSRs and several thousands of lightpaths. Our results are close to optimal, and they have shed new light into the design of MPλS networks. The most important finding is that it is possible to build cost-effective networks that provide rich connectivity among the LSRs with relatively few, but properly dimensioned, OXCs.

Chapter 4

Protection and Spare Capacity Allocation

Survivability is a fundamental requirement to the design of modern high-speed networks. In this chapter, we demonstrate how to integrate this requirement into the algorithms we developed in last chapter. As we have pointed out in Section 2.2, a shared path-based protection scheme is the most suitable one for the optical network in terms of spare capacity consumption and restoration time. We focus on heuristic algorithms to implement this scheme since it is a also NP-complete problem. As in the previous chapter, we decompose the problem into the primary/backup paths routing subproblem and the wavelength assignment subproblem and we develope heuristics to solve each subproblem. We also conduct numerical analysis using these algorithms and we compare them with the results from Chapter 3 in which protection was not considered. The results show that considering protection in the topology design will require more resources in terms of number of OXCs and number of fibers. However, the resource increase is moderate, confirming the conclusion we made in the previous chapter that cost-effective optical backbone networks that provide rich connectivity among large numbers of LSRs can be built with relatively few, but properly dimensioned, OXCs.

4.1 Protection Schemes: Algorithms and Implementation

In this section we focus on the shared path-based protection schemes. Since the related routing and wavelength assignment problem is NP-complete, we develop two heuristic algorithms that are polynomial in running time.

4.1.1 Routing of Link-Disjoint Paths

The routing of a pair of primary/backup paths is equivalent to the disjoint path problem since they should not share a common physical link or node to prevent the lightpath from one-link or one-node failure.

As we have discussed in Section 2.2, we will only consider single link failures, so a pair of link-disjoint paths is sufficient for the routing of the primary/backup paths. The link-disjoint problem can be defined as follow.

Definition 4.1.1 *Given a network $G = (V, E)$, a source node $s \in V$, and a sink (destination) node $t \in V$, find two paths from s to t that have no links in common.*

We note that the networks we generate in 3 are biconnected, which guantantees the existence of a pair of disjont paths between any pair of nodes in the network. A straightforward way to find a pair of link-disjoint paths for a connection (s, t) (s is the source and t is the destination) is a two-step algorithm that works as follows: Find the shortest path first and mark it as the primary path p_1. Then remove all links of p_1 from E and find the shortest path again from s to t and mark it as the backup path b_1. Obviously, p_1 and b_1 are link-disjoint. The advantage of this method is that we can always find a shortest path for the primary path, which is desirable for the QoS requirement of the connection service. Plus its implementation is very simple. The drawback of this method is that sometimes it may fail to find a pair of link-disjoint paths though there actually exists one [14]. In this case the pair of link-disjoint paths does not contain the shortest path between the source and the destination. In case that this method fails, we adopt the algorithm developed in [14], which can guarantee a pair of link-disjoint paths to be found when there exists one.

4.1.2 Wavelength Assignment for Shared Path-Based Protection

We consider two approaches, namely: separate routing and joint routing, to solve the routing and wavelength assignment problem for primary and backup paths.

1. **Separate routing, joint assignment**

   The primary and backup paths for each connection are routed one by one as a pair of link-disjoint paths. Then these paths are converted into an induced graph Q and a vertex coloring algorithm is used to assign the wavelength. Each vertex in this graph represents a lightpath and two vertices are connected by a link only if they share a common fiber link in the original graph under certain conditions. There are five types of relationship between any two pairs of primary/backup paths and accordingly their relationship in the induced graph is decided as follows. Suppose that $p1$ and $p2$ are the primary paths for two connections, $b1$ and $b2$ are the corresponding link-disjoint backup paths. We then have:

   - $p1$ and $b1$ are disjoint: $p1$ and $b1$ can use the same color, so there is no link between them in Q; $p2$ and $b2$ are disjoint: $p2$ and $b2$ can use the same color, so again there is no link between them in Q;
   - $p1$ and $b2$ overlap: $p1$ and $b2$ must use different colors, so $p1$ and $b2$ are connected in Q; $p2$ and $b1$ overlap: $p2$ and $b1$ must use different colors, so $p2$ and $b1$ are connected in Q;
   - $p1$ and $p2$ disjoint: $p1$ and $p2$ can use the same color, $b1$ and $b2$ may also use the same color. Therefore, there is no link between $p1$ and $p2$, or between $b1$ and $b2$ in Q.
   - $p1$ and $p2$ overlap, $b1$ and $b2$ also overlap: $p1$ and $p2$ must use different colors, $b1$ and $b2$ must also use different colors. Therefore, $p1$ and $p2$, $b1$ and $b2$ are connected in Q.
   - $p1$ and $p2$ overlap, $b1$ and $b2$ disjoint: $p1$ and $p2$ must use different colors, $b1$ and $b2$ can use the same color. So $p1$ and $p2$ are connected, but there is no link between $b1$ and $b2$ in Q.

   We associate a vector L_j for each of the $|E|$ links. L_j consists of all the paths that use this link. We use this data structure to decide if two primary paths are overlapping

Algorithm for shared path-based protection using separate routing and joint assignment
Input: The physical topology of OXC network $G(V, E)$, the set of lightpaths, $\mathcal{L}$
Output: Routing and wavelength assignment for the primary/backup paths for $\mathcal{L}$ based on the path protection

1. **begin**
2. **while** ($\mathcal{L}$ is not empty)
3. Remove a ligthpath requirement l_i from $\mathcal{L}$
4. Find a pair of link-disjoint paths for l_i
5. Let the shorter one as the primary path p_i, the longer one as the backup path b_i
6. **end while**
7. Convert all the primary/backup paths into the induced graph Q according to the five relationships defined in this section
8. Assign the colors for the graph Q
9. end algorithm

Figure 4.1: Algorithm for path protection using separate routing and joint assignment

or disjoint. The complete algorithm is depicted in Figure 4.1. We call this algorithm Separate-Routed Shared Path-based Protection (SSPP).

For a particular number N of LSRs and number M of OXCs, the `while` loop in Steps 2-6 takes $O(2\Delta NM^2)$ time since we use Dijkstra's algorithm for the lightpath routing, Step 7 takes $O(2\Delta^2N^2)$ for the construction of the induced graph Q, and Step 8 takes $O(4\Delta^2N^2)$ for the graph coloring since we totally have $2\Delta N$ lightpahts.Therefore, the time taken by the SSPP algorithm is $O(2\Delta NM^2 + 6\Delta^2N^2)$.

2. **Joint routing, separate assignment**
We use the heuristics developed in Chapter 3 to route the set of primary paths $\mathcal{L}_p$ first and assign the minimum number of wavelengths for them by coloring a induced graph Q_p. In Q_p, every node represents a primary lightpath and two nodes in Q_p will be interconnected by a link if the corresponding lightpaths have at least a physical link in common in the original network. Secondly, we route a disjoint backup path for each primary path in the network using the two-step method we described in Section 4.1.1 and assign an available wavelength for it. We note that the wavelength assigned to the backup lightpaths should not conflict with those wavelengths having been assigned to the set of primary lightpaths. Therefore, the wavelength assigment

for the set of backup lightpaths $\mathcal{L}_b$ can be converted to a constrained vertex coloring problem. After we obtain the routes of the set of backup lightpaths, we first convert it into a color constrained induced graph Q_b, in which a node represents a backup lightpath and the interconnection relationship among these nodes is decided by the following conditions:

- $p1$ and $p2$ overlap, $b1$ and $b2$ also overlap: $b1$ and $b2$ must use different wavelength (color). Therefore, $b1$ and $b2$ are connected in Q_b.
- $p1$ and $p2$ overlap, $b1$ and $b2$ disjoint: $b1$ and $b2$ can use the same wavelength (color). Thus, no link exists between $b1$ and $b2$ in Q_b.
- $p1$ and $p2$ disjoint: $b1$ and $b2$ can use the same wavelength (color), and no link exists between $b1$ and $b2$ in Q_b.

Every node in the Q_b is associated with a color constraint set. The color constraint set C_v for a particular node (backup lightpath) v in Q_b is decided by the wavelength assignment for the set of the primary lightpaths. A backup lightpath v may not use any wavelength (color) already used by a primary lightpath passing any physical link in its route.

The constrained vertex coloring problem can be defined as follows:

Definition 4.1.2 *Let $Q_b = (V, E)$ be a simple graph. Each node $v \in V$ has a set of colors C_v that may not be used. What is the minimum number of colors to cover all the nodes in V such that any two nodes u and v are not assigned the same color if $e(u, v) \in E$ and the color assigned to a node v, $c_v \notin C_v$.*

In order to solve the constrained vertex coloring problem, we propose an efficient greedy algorithm that combines the construction of the induced graph Q_b with wavelength assignment. The basic idea is as follows: We begin with an empty graph Q_b, the set of backup lightpaths L_b, and the set of primary lightpath L_p with assigned colors. At each step, we first remove a backup lightpath v from L_b and assign it with a constrained vector C_v. C_v is determined by checking every physical link along the

Algorithm for shared path-based protection using joint routing and separate assignment
Input: The physical topology of OXC network $G(V, E)$, the set of lightpaths, $\mathcal{L}$
Output: Routing and wavelength assignment for the primary/backup paths for $\mathcal{L}$ based on the path protection

1. **begin**
2. **while** ($\mathcal{L}$ is not empty)
3. Remove a ligthpath requirement l_i from $\mathcal{L}$
4. Find the shortest path for l_i as the primary path p_i
5. **end while**
6. Convert all the primary paths into the induced graph Q_p according to their link jointness relationship
7. Assign the colors for the graph Q_p
8. **while** ($\mathcal{L}$ is not empty)
9. Remove a ligthpath requirement l_i from $\mathcal{L}$
10. Find a path link-disjoint from p_i, let it be the backup path b_i for l_i
11. Decide constrained colors to the vector C_i for b_i by checking colored Q_p
12. Add constrained colors to C_i for b_i by checking the colored Q_b
13. Color b_i by the smallest indexed color not in C_i
14. Add b_i to the Q_b
15. **end while**
16. **end algorithm**

Figure 4.2: Algorithm for path protection using joint routing and separate assignment

route of v and adding all the colors used by primary lightpahts into C_v. We then compare v with every other backup lightpath u that have been put into Q_b with assigned colors. If the primary lightpaths of u and v share a common link and u and v themselves also share a common link, v can not use the color of u, c_u, and we add c_u into C_v. Finally we pick up the smallest color not belonging to C_v as c_v and add v into Q_b. This process continues until all L_b is empty and we obtain the constrained induced graph Q_b with assigned colors.

The complete algorithm is depicted in Figure 4.2. We call this algorithm Joint-Routed Shared Path-Based Protection (JSPP).

For a particular number N of LSRs and number M of OXCs, Steps 2-7 for the routing and wavelength assignment of primary lightpaths take $O(\Delta N M^2 + \frac{3}{2}\Delta^2 N^2)$ since we use Dijkstra's algorithm for the lightpath routing and we have ΔN primary lightpaths, the `while` loop in Steps 8-15 takes $O(\Delta N M^2 + \Delta^2 N^2 + \frac{\Delta^2 N^2}{2})$ for the routing and constrained wavelength (color) assignment of backup lightpaths. Therefore, the time taken by the JSPP algorithm is $O(2\Delta N M^2 + 3\Delta^2 N^2)$.

The above complexity analysis for the SSPP and JSPP algorithms is based on worst case analysis. We can see that these algorithms have similar time complexity, but the JSPP algorithm has a smaller constant coefficient than the SSPP algorithm.

To evaluate the effects of different protection schemes on the topology design of optical Networks, we modify the algorithm we developed in Chapter 3 to take protection into consideration in the routing and wavelength assignment of lightpaths. We substitute Step 7 in Figure 3.3 with the two algorithms we developed in Section 4.1.2. We expect that, when lightpath protection is taken into account, a larger number of wavelength will be needed to satisfy a given set of lightpath requests for a given physical topology. Alternatively, if the number of wavelengths is fixed, protection will require a larger number of OXCs, and therefore, a more expensive physical topology. The numerical results we present in the next section confirm our expectations, but indicate that the increase in network resources is moderate.

4.2 Numerical Results

We use the same network model and assumption we used in Chapter 3. We present two groups of numerical analysis in this section. The first group is used to investigate the relative efficiency of the two algorithms we developed in Section 4.1.2 in terms of wavelength usage and time complexity . The second group of results studies the effect of important network parameters such as the available number of wavelengths, the number of OXC ports, etc., on the number of OXCs and the physical topology when different protection policies are adopted. In addition to the two shared path-based protection algorithms we developed, we also include the results for a dedicated path-based scheme (DPP) and the results from Chapter 3 (which are shown under the label "Primary") as benchmarks. The DPP scheme is actually a 1+1 path protection scheme which we implement as follows: we route every lightpath requirement as a pair of primary and backup paths one by one, then assign the wavelength for all of these lightpaths in a way that any two lightpaths, primary or backup, do not share a same wavelength as long as they have a link in common. In this scheme, there is no wavelength sharing between backup paths, and therefore it always requires a larger number of wavelengths for a given topology. As a result, it also requires a larger number of OXCs in our topology design problem.

4.2.1 Algorithm Comparison

Figure 4.3 and Figure 4.4 depict the performance of the path-based protection algorithms. We vary the number N of LSRs from 100-1000, the number M of OXCs is set to 1/20 of N, and the number Δ of transceivers per LSR is fixed to 4. From Figure 4.3, we can clearly see that shared path-based protection is significantly more efficient in terms of wavelength usage compared to the DPP scheme. The gap between the shared and dedicated schemes becomes larger as the number N of LSRs increases. When N equals 1000, the number of wavelength required by the DPP scheme is almost twice that required by the shared protection schemes. At the same time, SSPP and JSPP require only a moderately larger number of wavelengths than the primary-only scheme that does not provide any protection. Between the two shared protection schemes we study, the one based on separate routing of primary/backup paths (SSPP) outperformed the one based on the joint routing of all connection requirements (JSPP). As a trade-off, Figure 4.3 shows that JSPP is much more time efficient than SSPP and is close to the primary-only scheme. In General, we have found that SSPP requires two or three times more computation time than JSPP. This result validates the complexity analysis results we presented earlier. The DPP scheme has the worst running time performance, and always requires two or three times more computation time than SSPP.

4.2.2 The Effects of System Parameters

In Figure 4.5, we plot the number M of OXCs in the MPλS network against the number N of LSRs. For these results, we have let $\Delta = 4$, $W = 64$ and $P = 64$. We make two important observations. First, for all of the four algorithms, the number M of OXCs increases almost linearly with the number N of LSRs and the slope is moderate. Second, JSPP and SSPP require a similar number M of OXCs, which is usually 1/3 more than that required by the primary-only case. Also, DPP is the worst scheme since it requires

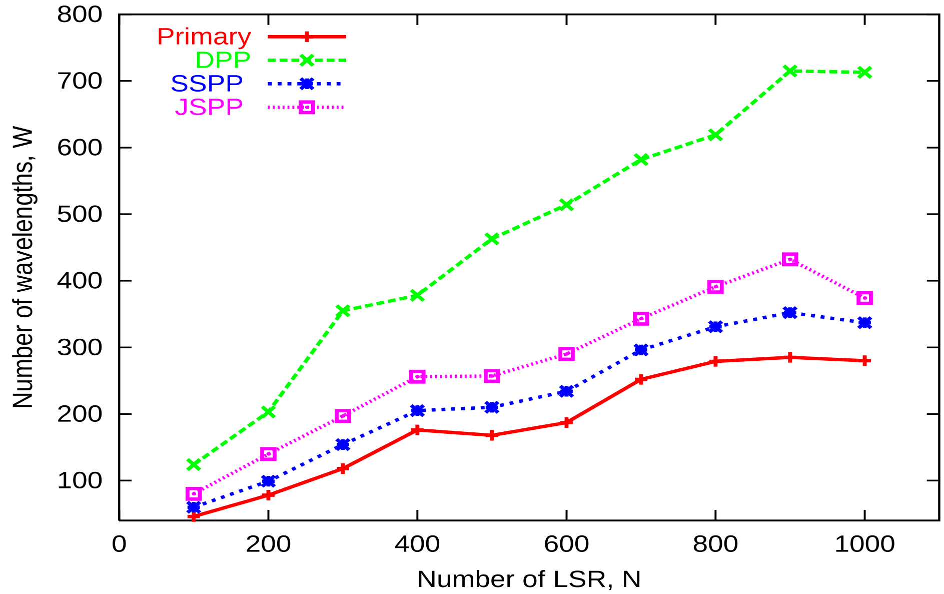


Figure 4.3: No. wavelengths for shared path-based protection

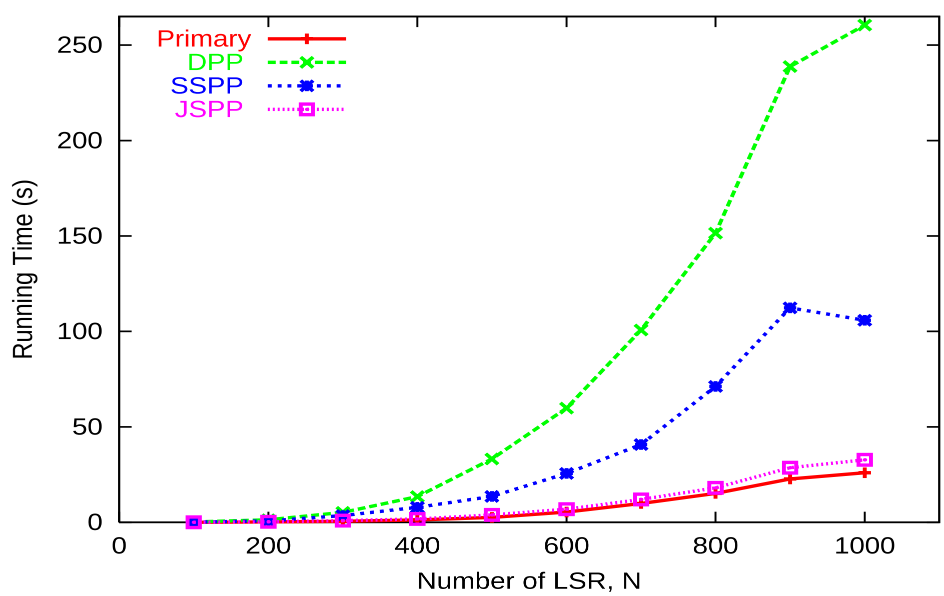


Figure 4.4: Running time for shared path-based protection

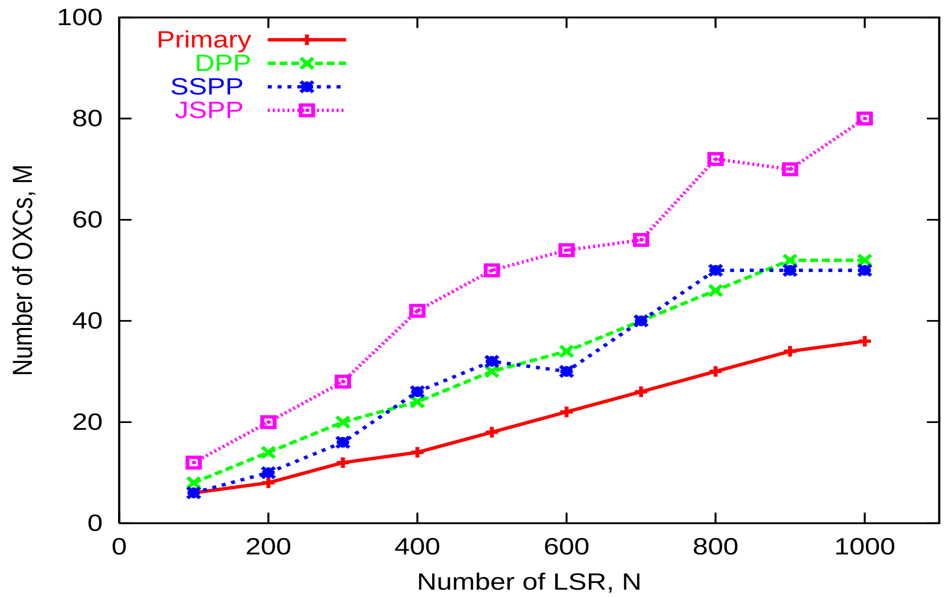


Figure 4.5: No. OXCs in the physical topology ($\Delta = 4$, $P = 64$, $W = 64$)

twice as many OXCs as the primary-only case. These results imply that OXC networks to interconnect very large number of LSRs can be built cost-effectively when protection is required and the shared path-based protection scheme is a good choice.

In Figure 4.6, we plot the number of OXCs against the number W of wavelengths when the number of LSRs is constant at $N = 300$, the number of transceivers per LSR is $\Delta = 8$, and the number of ports per OXC is $P = 64$. The results are as expected. The number of OXCs needed decreases as the number W of wavelength per fiber increases. The curves flatten out once $W > 80$, which reflects the fact that a larger number of port per OXC is needed to take full advantage of a large number of wavelengths. Also, the difference among the numbers of OXCs required for the different protection schemes tends to diminish as W increases.

A larger number of OXCs is required as Δ increases as shown in Figure 4.7, where we let $N = 300$, $P = 64$, and $W = 64$. While the increase in the number of OXCs is significantly slower than the increase in Δ for the primary-only case, the increase in the number of OXCs when considering protection is much faster. However, the increase in M is still slower than the increase in Δ for shared protection schemes JSPP and SSPP, but it is almost the same for DPP. Note that as Δ (and, consequently, the number of lightpaths)

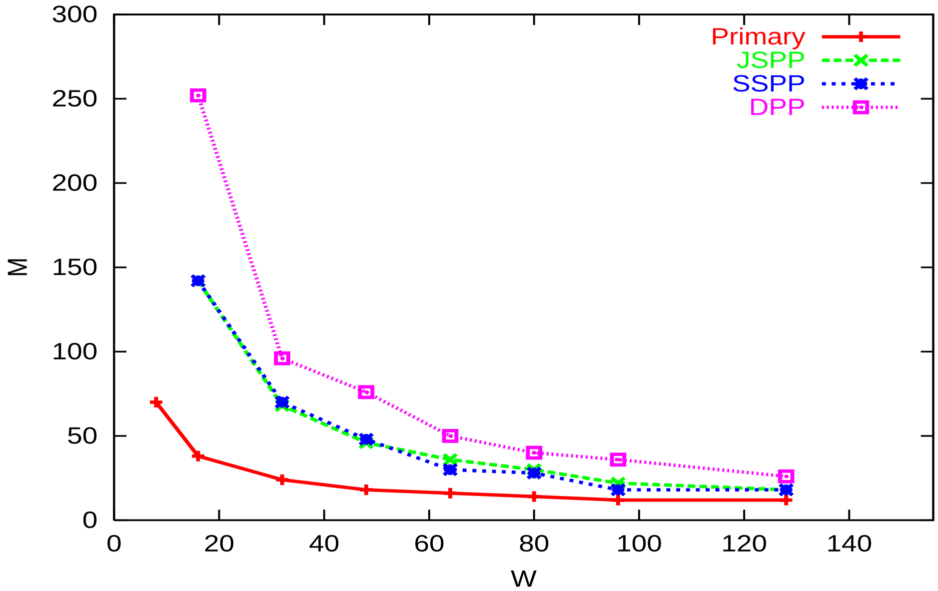


Figure 4.6: No. OXCs in the physical topology ($N = 300$, $P = 64$, $\Delta = 8$)

increases by a factor of six from 4 to 24, the number of OXCs required increases slower for JSPP, from 18 to 86, and for SSPP, from 18 to 90.

In Figure 4.8, we let $N = 300$, $W = 32$, and $\Delta = 8$, and we plot the number M of OXCs against the number P of ports per OXC. When P equals 16, a feasible solution could not be found even for $M = 300$ for any of three protection schemes. But a solution with a reasonable number M of OXCs is possible for $P \geq 24$ (except for DPP, when P=24). We also observe that increasing P beyond 30 does not have a significant effect on the number of OXCs (due to the limitation on the number W of wavelengths). Note that, in general OXCs with many ports are more expensive than those with few ports. Therefore, our results indicate that optical OXC networks can be built cost-effectively with OXCs with ports of medium size even when protection is required.

4.3 Concluding Remarks

We have described two heuristic algorithms for shared path-based protection in optical networks. We use these algorithms along with the ones we developed in the pre-

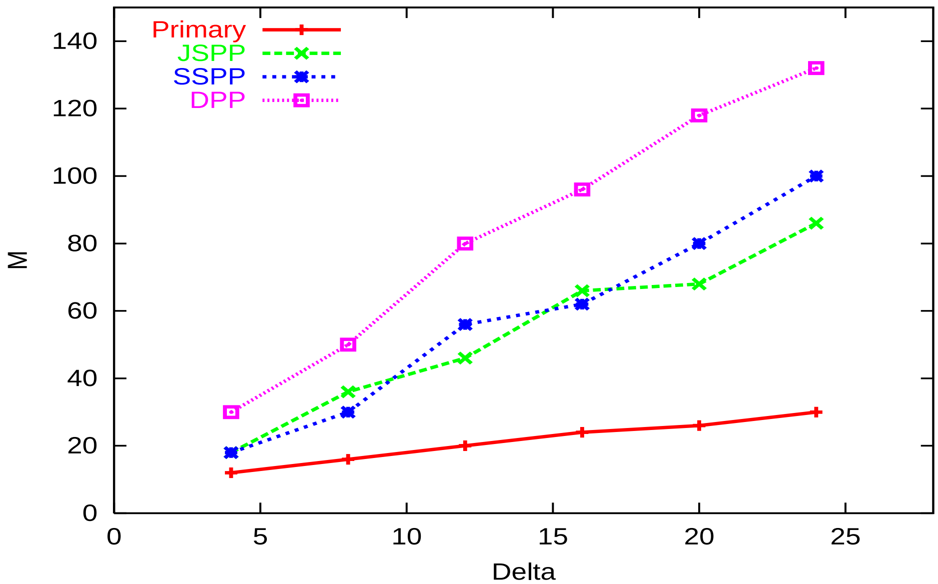


Figure 4.7: No. OXCs in the physical topology ($N = 300$, $P = 64$, $W = 64$)

vious chapter, to compute the number of OXCs and corresponding physical topology to interconnect a given number of LSRs to satisfy their lightpath requirement. We have also applied the algorithms to design networks that can accommodate up to 1000 LSRs and several thousands of lightpaths. The results indicate that survivability requirements increase the cost of the network in terms of number of OXCs or wavelengths, but only modestly. The results also imply that shared path-based protection outperforms the dedicated path-based protection significantly. Among the two heuristics we developed, the one based on the joint routing and separate wavelength assignment (JSPP) has similar performance in terms of the solution quality with the one based on the separate routing, joint wavelength assignment(SSPP), but the former runs much faster than the latter.

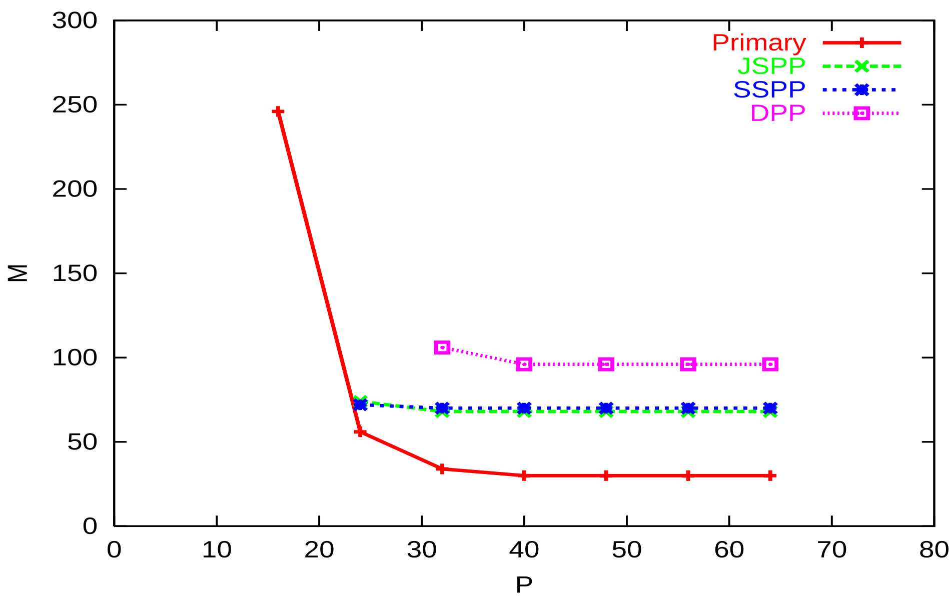


Figure 4.8: No. OXCs in the physical topology ($N = 300$, $W = 32$, $\Delta = 8$)

Chapter 5

Constrained Ligth-Tree Routing

Multicast communication in legacy networks has been studied extensively. It is usually accomplished through a mulitcast tree routed at the source node. When QoS is required, a constrained multicast tree can be constructed. The concept of a light-tree is the extension of a multicast tree to optical networks, and it can support multicasting and enhance the virtual connectivity and traffic grooming in the optical layer. A specific QoS problem for light-tree routing in the optical layer is motivated by the effects of light splitting and power attenuation of optical signals which are only partially mitigated by amplification. The light-tree constructed must guarantee a certain level of quality for the signals received by the destination nodes. The problem of signal quality does not arise in the context of multicast above the optical layer. Therefore, we define a new constrained light-tree routing problem to account for the power losses. We investigate a number of variants of this problem that we prove to be NP-complete. We also develop a suite of corresponding routing algorithms, one of which can be applied to networks with sparse light splitting and/or limited splitting fanout. One significant result of our study is that, in order to guarantee an adequate signal quality and to scale to large destination sets, light-trees must be balanced, or distance-weighted balanced (a term we define later). Numerical results demonstrate the advantage of our algorithms over existing algorithms in terms of the tree balancing and the fairness among the destinations.

5.1 The Multicast Optical Network

We consider an optical WDM network with N nodes interconnected by fiber links. Each of the links is capable of carrying W wavelengths, and each of the nodes is equipped with an OXC with P input ports and P output ports. The OXC at (some of) the nodes is multicast-capable (MC-OXC). A $P \times P$ MC-OXC consists of a set of W $P \times P$ splitter-and-delivery (SaD) switches, one for each wavelength; Figure 5.1 shows a 3×3 MC-OXC for $W = 2$ wavelengths. In addition to the W SaD switches, P demultiplexers (respectively, multiplexers) are used to extract (respectively, combine) individual wavelengths. The SaD switch design was first proposed in [42] and was later modified in [3] in order to reduce cost and improve power efficiency. A $P \times P$ SaD switch, as it was proposed in [42], is shown in Figure 5.2. It consists of P power splitters, P^2 optical gates (to reduce the excessive crosstalk), and P^2 2×1 photonic switches. As in [3], we assume that the splitters are configurable, in that they can be instructed to split the incoming signal into m output signals, $m = 1, \cdots, P$; note that $m = 1$ corresponds to no power splitting, i.e., no multicast, while $m = P$ corresponds to a broadcast operation. By appropriately configuring the corresponding m 2×1 photonic switches, each of the p signals resulting from the splitting operation can be switched to the desired output ports.

In a transparent network, optical signals suffer losses as they travel from source to destination node. We distinguish two types of losses:

1. *Signal attenuation.* This is due to the propagation of light along the fibers between the source and destination nodes. Optical amplifiers (EDFAs) are used along the optical paths to boost the power of the information-carrying signals in order to compensate for the signal attenuation. However, optical power amplification is not perfect, and there is a limit on the number of times a signal may be amplified. Thus, it has been suggested in [74] that power attenuation (along with other physical layer impairments, such as dispersion) be taken into account when routing lightpaths in a transparent optical network.

2. *Splitting loss.* An m-way splitter (similar to those shown in Figure 5.2) is an optical device which splits an input signal among m outputs. For an ideal device, the power of each output is $(1/m)$-th of that of the original signal; in practice, the splitting operation introduces additional losses and the power of each output is lower than

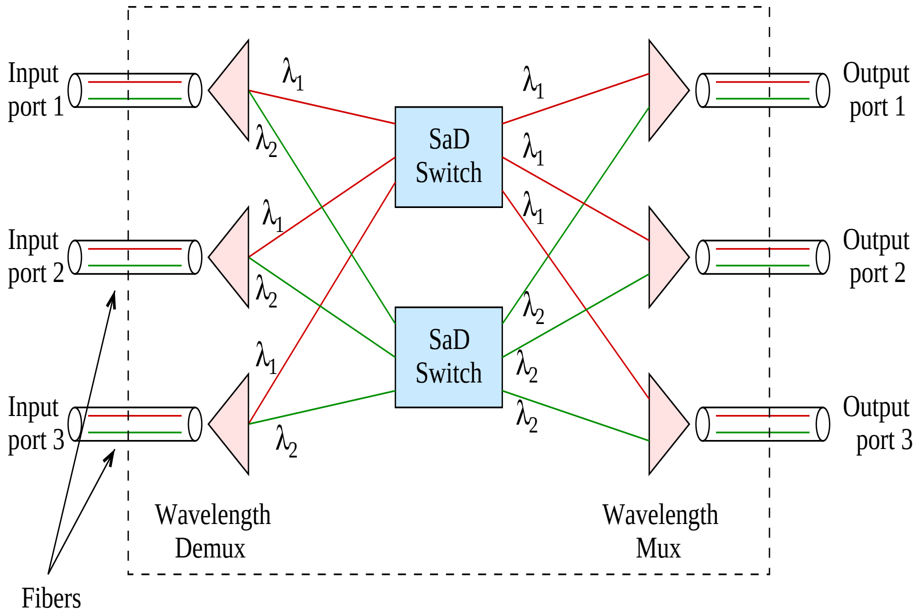


Figure 5.1: A $3X3$ MC-OXC based on the SaD switch architecture, $W = 2$

that of the ideal case. Splitting losses occur within MC-OXCs at the branch points of light-trees carrying point-to-multipoint signals. While amplification may partially compensate for the power loss due to light splitting, it is clear that this type of loss must be taken into account for light-tree routing.

The next two subsections discuss the two types of losses in more detail.

5.1.1 Power Attenuation Along a Fiber Link

The output power P_{out} at the end of a fiber of length L is related to the input power P_{in} by

$$P_{out} = P_{in}e^{-\alpha L} \tag{5.1}$$

where α is the fiber attenuation ratio [73]; near 1550 nm, we have that $4.34\alpha = 0.2 = \alpha_{dB}$. In general, distributed feedback (DFB) lasers put out about 50 mW (17 dB) of power after the output signal is boosted by an amplifier, while the sensitivity of avalanche photodiode (APD) receivers at 2.5 Gb/s is -34 dB [62]. Therefore, from (5.1), we obtain the maximum transmission distance in a fiber as $L_{max} = 255$ km. For any fiber link whose length is greater than L_{max}, a number of EDFA amplifiers must be added to compensate for the

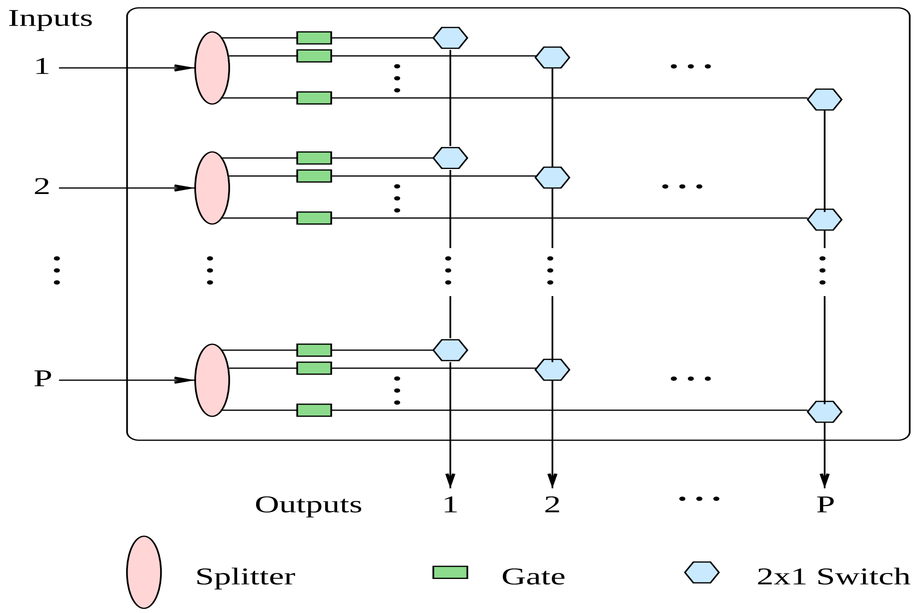


Figure 5.2: A $P \times P$ SaD switch

power attenuation, so that the receiving power at the end of the fiber is no less than -34 dB. When using optical amplifiers, other constraints must be considered, including the maximum permissible power on a fiber, the effects of fiber nonlinearities, and the receiver sensitivity. Consequently, in current practice, amplifier spacings range from 20 km to 100 km.

Suppose that the span of length between two consecutive amplifiers (EDFAs) in the optical network is S, and that the gain of each amplifier is denoted by G (S and G are assumed to be parameters which are fixed for a particular fiber system). Then, the power received at the end of a fiber link of length L is related to the input power as follows:

$$P_{out} = P_{in} \left(Ge^{-\alpha S}\right)^{\frac{L}{S}} = P_{in} \left(G^{\frac{1}{S}} e^{-\alpha}\right)^{L} = P_{in} Q^{L} \tag{5.2}$$

where $Q = G^{\frac{1}{S}} e^{-\alpha} < 1$ is a constant determined by the fiber system. Expression (5.2) describes the signal attenuation within a fiber link equipped with optical amplifiers, as a function of the link length L.

5.1.2 Power Loss due to Light Splitting at the MC-OXC

Let us now consider a signal that arrives at some input port of an MC-OXC such as the one shown in Figure 5.1. This signal is split into m output signals at the SaD switch corresponding to the input's wavelength. The m output signals are then switched to the appropriate output ports of the MC-OXC. We assume that the power splitters at the SaD switch (refer to Figure 5.2) are configurable, such that a multicast optical signal does not always need to be split P times, where P is the number of input/output ports of the SaD switch. Instead, the multicast signal is split into exactly m signals, $m = 1, \cdots, P$, where m is the out-degree of the node in the corresponding light-tree. Configurability is made possible by new devices such as the compact multimode interference couplers with tunable power splitting ratios that were reported recently in [51]. We also assume that the tunable power splitting ratio can be controlled by the multicast signaling protocol, making it possible to realize MC-OXCs with any desirable fanout $m, m = 1, \cdots, P$.

Given these assumptions, the power loss (in dB) at an MC-OXC for an input signal that is split into m output signals is given by [3]:

$$Loss_{SaD} = 10 \log_{10} m + \beta(P) \tag{5.3}$$

In the above expression, the term $\beta(P)$ captures losses due to the multiplexing and demultiplexing of signals, as well as the insertion and coupling losses at the 2×1 switching elements (refer to Figure 5.2). Since the number of switching elements in the signal path is equal to the number P of input/output ports of the SaD switch, then this term is a function of P.

From (5.3), we can now derive the output power of each of the m output signals as a function of the input power as follows:

$$P_{out} = \frac{10^{-\frac{\beta(P)}{10}}}{m} P_{in} \leq \frac{P_{in}}{m} \tag{5.4}$$

Expression (5.4) assumes that signals are not amplified as they leave the MC-OXC. To compensate for the power loss due to light splitting, optical amplifiers may be placed at the output ports of the MC-OXC. Let G denote the gain of an amplifier, and define $R = 10^{-\frac{\beta(P)}{10}} G$. R is a constant for a given SaD switch, and is determined by the number of ports P of the switch, the losses incurred at the various elements of the switch, and the amplifier gain. Then, the output power of a signal that has undergone m-way splitting is given by:

$$P_{out} = \frac{R P_{in}}{m} \leq P_{in} \tag{5.5}$$

5.2 The Light-Tree Routing Problem

We represent a network of MC-OXCs by a simple graph $G = (V, A)$. V denotes the set of nodes (i.e., MC-OXCs), and A, the set of arcs, corresponds to the set of (unidirectional) fiber links connecting the nodes. We will also use $N =\mid V \mid$ to refer to the number of nodes in the network. We define a *distance function* $\mathcal{D} : A \rightarrow \mathcal{R}^+$ which assigns a non-negative weight to each fiber link in the network. More specifically, the value $\mathcal{D}(\ell)$ associated with link $\ell = (u, v) \in A, u, v \in V$, is the geographical distance that the optical signal travels along the link ℓ from node u to node v.

Under the light-tree routing scenario we are considering, an optical signal originating at some *source* node $s \in V$ in the network must be delivered to a set $M \subseteq V - \{s\}$ of destination nodes. In general, several point-to-multipoint sessions may proceed concurrently within the network, each characterized by a source node and a destination set. We assume that communication in the network is connection-oriented, and that point-to-multipoint connections are established by issuing a *connect request*; similarly, at the conclusion of a session a *disconnect request* is issued. In response to a connect request, and prior to any optical signal been transmitted from the source to the destinations, a connection establishment process is initiated. Central to the connection establishment is the determination of a light-tree, i.e., a set of paths between the source and the destinations, over which the optical signal will be carried for the duration of the point-to-multipoint session.

Let s and M be the source and destination set, respectively, of a certain point-to-multipoint session. We let $T = (V_T, A_T)$ denote the light-tree, rooted at s, for this session. The light-tree is a subgraph of G (i.e., $V_T \subseteq V$ and $A_T \subseteq A$) spanning s and the nodes in M (that is, $M \cup \{s\} \subseteq V_T$). In addition, V_T may contain *relay* nodes, that is, nodes intermediate to the path from the source to a destination. Relay nodes do not terminate the optical signal transmitted by the source node s; rather, they simply split and/or switch the signal towards the downstream links of the light-tree. We let $H_T(s, v)$ denote the unique path from source s to destination $v \in M$ in the light-tree T. We define $P_{in}(s)$ as the power of the optical signal injected into the network by the source node s, and $P_{out}(s, v)$ as the power of the optical signal received by destination $v \in M$. The output power $P_{out}(s, v)$ at destination v is related to the input power at the source s through the following expression:

$$P_{out}(s, v) = P_{in}(s) \times L^{(atten)}(s, v) \times L^{(split)}(s, v) \tag{5.6}$$

In the above expression, parameter $L^{(atten)}(s,v)$ (respectively, $L^{(split)}(s,v)$) accounts for the power loss due to attenuation (respectively, light splitting) along the path from s to v in the light-tree T; we assume that both parameters include the effects of amplification.

Recall that expression (5.2) relates the input and output signal power for a single fiber link. The expression can be generalized to a path from a source s to a destination v in a straightforward manner, allowing us to express $L^{(atten)}(s,v)$ as follows:

$$L^{(atten)}(s,v) = \prod_{\ell \in H_T(s,v)} Q^{\mathcal{D}(\ell)} = Q^{\sum_{\ell \in H_T(s,v)} \mathcal{D}(\ell)} < 1 \tag{5.7}$$

Similarly, we can obtain an expression for $L^{(split)}$ by considering all MC-OXCs in the path from s to v in the light-tree T, and applying expression (5.5). Let us define $F_T(u)$ as the fanout of the MC-OXC at node u of the light-tree T, with respect to the optical signal carried on this light-tree [1]. The fanout $F_T(u)$ corresponds to the quantity m in expression (5.5). As a result, we obtain:

$$L^{(split)}(s,v) = \prod_{u \in H_T(s,v)} \frac{R}{F_T(u)} < 1 \tag{5.8}$$

We note that, as we explained in the previous section, quantities Q and R in expressions (5.7) and (5.8) are constants for a given optical network.

5.2.1 Path Constraints to Ensure Optical Signal Quality

We now introduce two parameters that can be used to characterize the quality of the light-tree as perceived by the application making use of the point-to-multipoint optical communication. These parameters relate the end-to-end power loss along individual source-destination paths to the desired level of signal power at the receivers, as follows.

- *Source-destination loss tolerance*, Δ. Parameter Δ represents an upper bound on the acceptable end-to-end power loss along any path from the source to a destination node. This parameter reflects the fact that if the optical signal power falls below the receiver sensitivity, then the information carried by the signal cannot be recovered.

- *Inter-destination loss variation tolerance*, δ. Parameter δ is the maximum difference between the end-to-end losses along the paths from the source to any two destination

[1] Note that node v may be part of a different light-tree T', with a different source and destination set; its fanout with respect to T' may be different than its fanout with respect to T.

nodes that can be tolerated by the application. This parameter can be thought of as a measure of *fairness* among the destination nodes of the light-tree.

By supplying values for parameters Δ and δ, the application in effect imposes a set of constraints on the optical signal power at the receivers of the light-tree:

$$P_{out}(s,v) \geq \Delta P_{in}(s) \qquad \forall\, v \in M,\ \Delta \leq 1 \tag{5.9}$$

$$\frac{1}{\delta} \leq \frac{P_{out}(s,v)}{P_{out}(s,u)} \leq \delta \qquad \forall\, v,u \in M,\ \delta \geq 1 \tag{5.10}$$

We will refer to (5.9) as the *source-destination loss constraint*, while (5.10) will be called the *inter-destination loss variation constraint*. We will also say that tree T is a *feasible* light-tree for a point-to-multipoint session with source s and destination set M, if and only if T satisfies both (5.9) and (5.10). Note that, in order for the application to proceed, it is necessary and sufficient that a *single* feasible light-tree be constructed, since *any* feasible tree can meet the quality of service requirements as expressed by parameters Δ and δ.

5.3 Optical Signal Power Constrained Light-Trees

Let Δ and δ be the loss and loss variation tolerances, respectively, as specified by a client application that wishes to initiate a point-to-multipoint session. Our objective is to determine a light-tree such that the power losses along all source-destination paths in the tree are within the two tolerances. This problem, which we will call the *Power Constrained Light-Tree* (PCLT) problem, can be formally expressed as follows.

Problem 5.3.1 (PCLT) *Given a network $G = (V, A)$, a source node $s \in V$, a destination set $M \subseteq V - \{s\}$, a distance function $\mathcal{D} : A \rightarrow \mathcal{R}^+$, a loss tolerance Δ, and a loss variation tolerance δ, does there exist a light-tree $T = (V_T, A_T)$ spanning s and the nodes in M, that satisfies both constraints (5.9) and (5.10)?*

In the next three subsections we study three variants of the PCLT problem. The variants mainly differ in the assumptions made regarding the degree to which each of the two types of power loss (i.e., loss due to attenuation or light splitting) affects the quality of the received signal. As we explain, the assumptions depend on the geographical span of the light-tree and the size of the destination set, and it is possible that different variants of the PCLT problem apply to different light-trees within the *same* optical network. Therefore, we characterize the complexity of, and provide light-tree algorithms for, all three variants of the PCLT problem.

5.3.1 The PCLT Problem Under Power Attenuation Only

Let us first consider the PCLT problem under the assumption that power attenuation is the dominant factor in determining the signal quality at the receivers of the light-tree. In other words, we assume that $L^{(split)}(s,v) \approx 1$ in expression (5.6), for all destinations v. This is a reasonable assumption when **(i)** the source of the point-to-multipoint session and the destination nodes are separated by large geographical distances, and/or **(ii)** there is a small number of destination nodes, thus, the optical signal only needs to undergo a small number of splitting operations. In this case, we can use (5.7) to rewrite the source-destination constraint (5.9) and the inter-destination loss variation constraint (5.10) as follows.

$$\begin{aligned} L^{(atten)}(s,v) \geq \Delta \quad &\Rightarrow \quad Q^{\sum_{\ell \in H_T(s,v)} \mathcal{D}(\ell)} \geq \Delta \\ &\Rightarrow \sum_{\ell \in H_T(s,v)} \mathcal{D}(\ell) \leq \log_Q \Delta \qquad \forall\, v \in M \end{aligned} \tag{5.11}$$

$$\begin{aligned} \frac{1}{\delta} \leq \frac{L^{(atten)}(s,v)}{L^{(atten)}(s,u)} \leq \delta &\Rightarrow Q^{\left|\sum_{\ell \in H_T(s,v)} \mathcal{D}(\ell) - \sum_{\ell \in H_T(s,u)} \mathcal{D}(\ell)\right|} \leq \delta \\ &\Rightarrow \left| \sum_{\ell \in H_T(s,v)} \mathcal{D}(\ell) - \sum_{\ell \in H_T(s,u)} \mathcal{D}(\ell) \right| \leq \log_Q \delta \qquad \forall v,u \in M \end{aligned} \tag{5.12}$$

Note that the last step in (5.11) is due to the fact that constants Q and Δ are such that $0 < \Delta, Q < 1$.

An interesting observation regarding constraints (5.11) and (5.12) is that they represent two conflicting objectives. Indeed, the loss constraint (5.11) dictates that short paths be used. But choosing the shortest paths may lead to a violation of the loss variation constraint (5.12) among nodes that are close to the source and nodes that are far away from it. Consequently, it may be necessary to select longer paths for some nodes in order to satisfy the latter constraint. Then, the problem of finding a feasible light-tree becomes one of selecting paths in a way that strikes a balance between these two objectives.

The PCLT problem with constraints (5.11) and (5.12) is equivalent to the *delay- and delay variation-bounded multicast tree* (DVBMT) problem studied in [65, 71]. Specifically, the loss constraint (5.11) is equivalent to the delay constraint of DVBMT, while the loss variation constraint (5.12) is equivalent to the delay variation constraint of DVBMT. We proved in [65] that the DVBMT problem is NP-complete whenever the size of the destination set $| M | \geq 2$. Consequently, if we ignore the power loss due to the splitting of the optical signal at the branch nodes of the light-tree, the PCLT problem is also NP-complete. In this case, the heuristics developed in [65, 71] can be applied directly to construct a light-tree that satisfies both constraints (5.11) and (5.12).

5.3.2 The PCLT Problem Under Splitting Losses Only

Let us now turn our attention to the case when signal attenuation is negligible (i.e., $L^{(atten)} \approx 1$ in expression (5.6)), and power loss due to light splitting is the dominant factor affecting signal quality at the receivers. This situation may arise when **(i)** the destination set includes a large number of nodes, and/or **(ii)** the source and destination nodes are located in close proximity to each other. We can then use expression (5.8) to rewrite constraints (5.9) and (5.10) as follows (recall that $F_T(w)$ is the fanout of node w with respect to light-tree T, in other words, it denotes the number of times the optical signal traveling along light-tree T is split at node w).

$$\prod_{w \in H_T(s,v)} \frac{R}{F_T(w)} \geq \Delta \qquad \forall\, v \in M \tag{5.13}$$

$$\frac{1}{\delta} \leq \frac{\prod_{w \in H_T(s,v)} \frac{R}{F_T(w)}}{\prod_{w \in H_T(s,u)} \frac{R}{F_T(w)}} \leq \delta \qquad \forall\, u, v \in M \tag{5.14}$$

Let us interpret constraints (5.13) and (5.14). Without loss of generality, let us assume that $R = 1$, i.e., that the power of each of the $F_T(w)$ output signals at node w is $(1/F_T(w))$-th of that of the input signal; our conclusions are valid even when $R > 1$. When $R = 1$, the denominator of the left hand side of (5.13) corresponds to the product $\prod_{w \in H_T(s,v)} F_T(w)$ along the path from the source s to destination v. We will call this product the *split ratio of node* v, and its inverse corresponds to the residual power of the optical signal received at node v after all the splits along the path. We can see that constraint (5.13) imposes an upper bound on the split ratio on the path to each destination node in set M.

Let us now turn our attention to constraint (5.14). When $R = 1$, it states that the split ratios of any two paths from the source to two destination nodes v and u should be within a tight range from each other, where the tightness of the range is determined by parameter δ. Therefore, this constraint suggests that light-trees must be as balanced as possible. To see why, suppose that a light-tree is constructed for a set of K destinations such that one destination node, say v, is directly connected to the root (source) while the remaining $K - 1$ nodes are all in a different subtree connected to the root. It is clear that, even after amplification (i.e., $R > 1$), node v will receive a signal of better quality than the other $K - 1$ destinations: the signal arriving at node v is of the same quality as the one traveling towards the other subtree, but the latter signal will have to be split several times (and thus, it will degrade further) before it reaches each of the $K - 1$ destinations in the subtree. Such an unbalanced tree has two important disadvantages. First, it introduces unfairness, since receivers at small depth in the (logical) tree receive a signal of better quality than receivers at large depth, *independently* of their geographical distance to the source. Second, it is not scalable, since it may introduce excessive losses that make it impossible to deliver a signal to a given number of destinations. To see this, consider a worst case scenario where the tree is a binary one and is recursively constructed such that the left subtree consists of exactly one receiver, while the right subtree contains all remaining receivers and consists of left and right subtrees in a similar way. It is easy to see that the receiver at depth one (in the left subtree of the whole tree) receives a signal that has undergone one split and its power is one-half of that of the original signal. On the other hand, the receiver at depth K (the rightmost leaf of the tree) receives a signal that is the result of K splits, and its power is $(1/2^K)$-th of that of the original signal. While extreme, this scenario illustrates the pitfalls of unbalanced trees for the multicast of optical signals.

The requirement that the light-tree be as balanced as possible is a direct consequence of the fact that when an optical signal undergoes m-way splitting, its power is equally divided among the m output signals. Thus, this requirement is unique to optical layer multicast. To the best of our knowledge, the problem of constructing balanced multicast trees has not been studied in the literature, since it does not arise in the context of multicast above the optical layer. We now prove the problem of constructing balanced multicast trees to be NP-complete. In the next subsection we present a suite of heuristics to obtain balanced light-trees that satisfy constraints (5.13) and (5.14).

Our proof is by reduction from the Exact Cover by Three-Sets (X3C) problem [35], a well-known NP-complete problem defined as:

Definition 5.3.1 (X3C) *Given a set $S = \{S_i\}$ with $3k$ elements for some natural number k and a collection $Y = \{Y_j\}$ of subsets of the set, each of which contains exactly three elements, do there exist in the collection Y k subsets that together cover the set S?*

Theorem 5.3.1 *The PCLT problem under constraints (5.13) and (5.14) is NP-complete.*

Proof. Clearly, PCLT belongs in the class NP, since a solution to the PCLT problem can be verified in polynomial time. We now transform the NP-complete X3C problem to PCLT. Consider an arbitrary instance of the X3C problem consisting of **(i)** a set $S = \{S_i\}$ of elements, where $|S| = 3k$ for some natural number k, and **(ii)** a collection $Y = \{Y_j\}$ of subsets of S, each subset containing exactly three elements of S. Let $m = |S|, n = |Y|$. We construct a corresponding instance of PCLT as follows. The graph $G = (V, A)$ has $n + m + 1$ nodes, with $V = \{s, Y_1, Y_2, \cdots, Y_n, S_1, S_2, \cdots, S_m\}$, where s is the source node and $M = S = \{S_i\}$ is the destination set of the light-tree. The set A of links is:

$$\begin{aligned} A &= \{(s, Y_1), (s, Y_2), \cdots, (s, Y_n)\} \\ &\quad \cup \{(Y_j, S_i) | Y_j \in Y \wedge S_i \in Y_j\} \end{aligned} \tag{5.15}$$

In other words, there is a link from s to every node Y_i, and a link from every node Y_i to every node S_j which is a member of Y_i (see Figure 5.3). The distance function is defined as $\mathcal{D}(\ell) = 1, \ \forall \ \ell \in A$ (in fact, the distance function can be arbitrary; since this variant of PCLT neglects power attenuation, the constraints (5.13) and (5.14) do not depend on the link weights). Finally, the loss and loss variation tolerances are $\Delta = \frac{1}{3k}$ and $\delta = 1$, respectively.

It is obvious that this transformation can be performed in polynomial time. We now show that a feasible light-tree for the PCLT problem exists if and only if set S has an exact cover. If S has a cover $X = \{Y_{\pi_1}, Y_{\pi_2}, \cdots, Y_{\pi_k}\}$, the tree containing the source s, the set of nodes $X = \{Y_{\pi_1}, Y_{\pi_2}, \cdots, Y_{\pi_k}\}$, and the set of nodes $S = \{S_i\}$ is a feasible solution for PCLT. This is because the split ratio of each destination node S_i is equal to $\frac{1}{3k}$, and the tree satisfies both constraints (5.13) and (5.14). Conversely, let T be a feasible light-tree for PCLT. Then, T must contain the source node s, all destination nodes S_i,

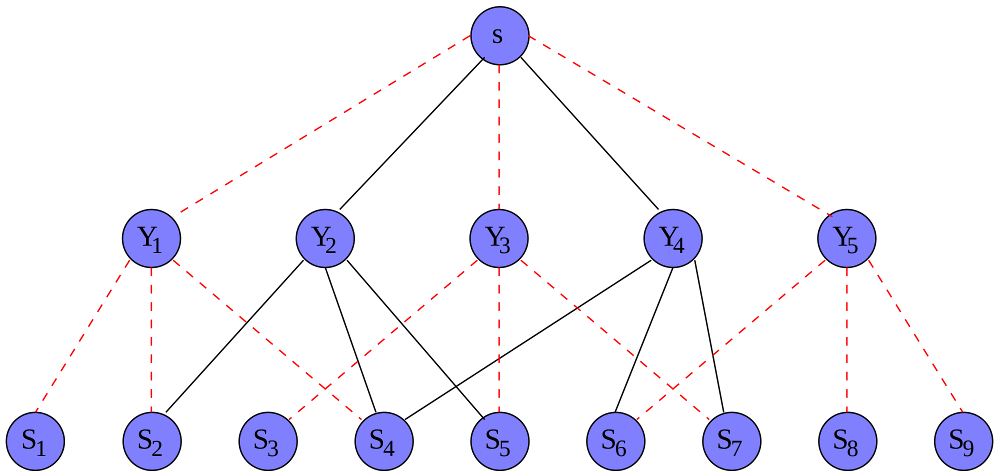


Figure 5.3: Instance of PCLT corresponding to an instance of X3C with $k = 3, S = \{S_1, \cdots, S_9\}, Y_1 = \{S_1, S_2, S_4\}, Y_2 = \{S_2, S_4, S_5\}, Y_3 = \{S_3, S_5, S_7\}, Y_4 = \{S_4, S_6, S_7\}, Y_5 = \{S_6, S_8, S_9\}$, and exact cover $\{Y_1, Y_3, Y_5\}$; the light-tree T is denoted by dashed lines

and a subset X of $Y = \{Y_j\}$. Since $\delta = 1$, all destination nodes S_i have the same split ratio. By construction of the PCLT instance, each destination node S_i must have exactly one parent in the light-tree T: if some node S_i had more than one parents a loop would exist (contradicting the hypothesis that T is a tree), and if it had no parent, it would not be connected to the tree T (again contradicting the hypothesis that T is a solution to the PCLT problem, i.e., it spans all destination nodes). Therefore, the nodes in the subset X of Y contained in the light-tree T exactly cover the set S, implying that X is a solution to the instance of the X3C problem.

Balanced Light-Tree Algorithm

We now present an algorithm for the version of the PCLT problem discussed above. The objective of any such algorithm would be to construct a feasible light-tree, i.e., one that satisfies both constraints (5.13) and (5.14). Note, however, that since the PCLT problem is NP-complete, any polynomial-time algorithm may fail to construct a feasible light-tree for a given problem instance, even if one exists. The algorithm we present can be used to search through the space of *candidate* trees (i.e., trees spanning s and the nodes in M) for a feasible solution to the PCLT problem. Our algorithm either returns a feasible tree, or, having failed to discover such a tree, it returns one for which **(i)** the maximum split ratio of any node in M, and **(ii)** the maximum difference between the split ratios of any pair of nodes in M, are minimum *over all trees considered by the algorithm*.

The *balanced light-tree* (BLT) algorithm, described in detail in Figure 5.4, takes as input an initial tree T_0 spanning the source s and destination nodes in M; the issue of constructing this initial tree is addressed shortly. In general, tree T_0 may be infeasible, i.e., it may violate (5.13) and/or (5.14). The key part of the BLT algorithm is the tree balancing procedure that is implemented by the **while** loop in Steps 4-18 of Figure 5.4. Consider an intermediate light-tree T, and let u (respectively, v) denote the leaf node with maximum (respectively, minimum) split ratio. The idea behind the BLT algorithm is to delete node u from T, and add it back to the tree by connecting it to some node y in the path from source s to v. Doing so reduces the split ratio of node u, but it also increases the split ratio of all nodes below node y in the tree; therefore, this pair of delete/add operations is performed only if it does not increase the split ratio of any node beyond that of node u (refer also to the **if** statement in Steps 14-17 of Figure 5.4). Thus, after each iteration of the algorithm, the split ratio of the node with the maximum value is decreased, in an attempt to satisfy constraint (5.13). While the split ratio of some other node(s) is increased, it does not increase beyond the previous maximum value. As a result, the difference between the maximum and minimum split ratio values also decreases with each iteration, as required by constraint (5.14). The algorithm terminates after a certain number of iterations, or if two successive iterations fail to reduce the maximum split ratio; the latter condition is not shown in Figure 5.4 in order to keep the pseudocode description simple.

In order to completely specify the BLT algorithm, we now explain how to select the node y in the path from s to v (the node with the minimum split ratio) to connect node u (the one with the maximum split ratio). Let Y denote the number of nodes in the path from source s to node v. We consider three different criteria for selecting a node $y \in Y$ to which to connect node u, resulting in three variants of the BLT algorithm.

1. **Shortest path (BLT-SP).** In this variant, we select node y such that the path from y to u is shortest among the paths from any node in Y to u.

2. **Minimum split ratio (BLT-MSR).** In this case, node y is one with the smallest split ratio among all nodes in Y.

3. **Degree constraint (BLT-D).** This is similar to BLT-MSR, except that the node y selected must be such that its fanout is no more than a maximum value F. F

Balanced Light-Tree (BLT) Algorithm

Input: A graph $G = (V, A)$ representing the network of MC-OXCs, a source node $s \in V$, a destination set $M \subseteq V$, a loss tolerance Δ, a loss variation tolerance δ, and an initial light-tree T_0 spanning the set $\{s \cup M\}$

Output: A light-tree T_f spanning the set $\{s \cup M\}$, and such that either **(i)** T_f is feasible, or **(ii)** the difference between the maximum and minimum split ratio of any two nodes in M is minimum

begin
  $T \leftarrow T_0$ // Initialize the light-tree
  $h \leftarrow 1$ // Number of iterations
  while ($h \leq$ MAX_ITER)
    Use depth-first search to calculate the split ratio of all nodes in M
    if (light-tree T is feasible) **then** return T
    $u \leftarrow$ node with maximum split ratio
    $v \leftarrow$ node with minimum split ratio
    $w \leftarrow$ the first node in the path from u to s in tree T
      s.t. $w \in M$ or w has a fanout > 1
    $Y \leftarrow$ set of nodes in the path from v to s in tree T
    In graph G, compute the shortest path from u to every node in Y
    $y \leftarrow$ a node in Y selected based on a predefined criterion (see Section 5.3.2)
    if (maximum split ratio of T does not increase) **then**
      Delete the path from w to u in tree T
        // Delete the node with the maximum split ratio
      Add the shortest path from y to u to T
        // Add the node back to T on a different path
    end if
  end while
  return T
end algorithm

Figure 5.4: General balanced light-tree (BLT) algorithm

may correspond to the maximum fanout capacity of the SaD switches at each MC-OXC. With this selection criterion, the resulting light-tree will have a bounded degree (fanout). Note that, if we use a different value of F for each node in the network, then the algorithm can be used in optical networks with sparse light splitting, since multicast-incapable OXCs can be accounted for by letting $F = 1$ for these nodes.

Finally, we use the SPH algorithm [76] to construct an initial tree T_0 that spans the source node s and the destination set M. The SPH algorithm is a fast algorithm which has been used successfully as a starting point for several constrained Steiner tree problems [12]. The algorithm starts with a partial tree consisting of the shortest path from the source s to some destination node. It then repeatedly extends the partial tree to another destination node u, until all destination nodes have been included. A new destination node u is connected to the partial tree by including the shortest path from some node y of the tree to u. Therefore, the issue arises of selecting the node y of the partial tree to which to connect node u. For each variant of the BLT algorithm, we use the corresponding selection criterion to select node y of the partial tree.

Regarding the complexity of the BLT algorithm shown in Figure 5.4, it is straightforward to verify that the worst-case running time is $O(N^2 I)$, where N is the number of nodes in the network and I, an input parameter, is the number of iterations of the **while** loop in Steps 4-18. We note that the worst-case complexity is the same for all three variants of the BLT algorithm.

5.3.3 The General PCLT Problem

We now consider the most general version of the PCLT, which arises when both signal attenuation and light splitting contribute to the degradation of the quality of the signal as it travels through the optical network. In this case, the signal power received at each destination node is related to the signal power emitted by the source node through expressions (5.6)-(5.8), and the light-tree must be constructed such that constraints (5.9) and (5.10) be satisfied. Clearly, this version of the PCLT problem is also NP-complete, since it includes as special cases the two versions studied in Sections 5.3.1 and 5.3.2, both of which are NP-complete.

An interesting observation regarding this general version of the PCLT problem is that there is a tradeoff between the number of times a signal may be split and the distance that the signal can travel. Signals that have been split multiple times may not be able to travel over large distances, even after amplification, and vice versa. This tradeoff, which is unique to optical networks, is not taken into account by existing multicast routing algorithms. In this case, it would be desirable to have receivers which are far away (in terms of distance traveled by the optical signal) from the source, be closer to the source in the (logical) light-tree. This way, the signal arriving to these receivers will have undergone a smaller number of splits. In this case, the resulting light-tree will not necessarily be balanced (in the traditional definition of the term), but rather it must be balanced in a manner that accounts for the geographical locations of the various receivers relative to the source. In other words, the number of signal splits for each receiver must be appropriately weighted by the distance to the receiver.

Based on the above observations, we modify the BLT algorithm shown in Figure 5.4 to construct *distance-weighted* balanced light-trees; we will call this algorithm *weighted* BLT (WBLT). The main idea is to consider the tree node with the largest *total* loss and attempt to reduce its splitting loss by moving it closer to the source in the logical light-tree. Doing so may increase the attenuation loss (since the node may be added to the tree on a longer path), but it will also decrease its splitting loss, possibly resulting in a smaller total loss. This weighted balancing procedure can be accomplished by making the following small changes in the algorithm of Figure 5.4: in Step 7 (respectively, Step 8), select the node with the maximum (respectively, minimum) total loss, and in the **if** statement in Step 14, check whether the maximum total loss at any node of the tree increases. Otherwise, the algorithm remains unchanged. Note that, since there are three variants of the BLT algorithm, we also have three variants of WBLT, namely, WBLT-SP, WBLT-MSR, and WBLT-D.

5.4 Numerical Results

We have used simulation to evaluate the average case performance of the light-tree routing algorithms on randomly generated graphs. The graphs were generated using the method described in [81]. The nodes of the graphs were placed in a grid of dimensions 5000×5000 km, an area roughly the size of the continental United States. The weight of each link was set to the Euclidean distance between the pair of nodes connected by the link. To test the performance of our algorithms, we randomly generated graphs with a number of nodes ranging from 50 to 110, and we varied the size of the destination set from 5-15% of the number of nodes in the graph. In all the results shown in this section, each point plotted represents the average over 300 graphs for the stated number of nodes. We have also computed 95% confidence intervals which are not shown, since they are very narrow and including them would affect the clarity of the figures. For algorithm BLT-D, we set the degree constraint as 4, a reasonable value for the maximum fanout of an MC-OXC.

5.4.1 The BLT Algorithms

We first study the performance of the three variants of the BLT algorithm (namely, BLT-SP, BLT-MSR, and BLT-D) for the PCLT problem under splitting losses only. We consider three performance measures:

1. *maximum split ratio*, which captures the quality of the signal at the destination node where it is worst,
2. *maximum-to-minimum split ratio*, which reflects the difference between the best and worst signal quality, and is a measure of inter-destination fairness, and
3. *number of links of the light-tree*, which captures the amount of resources (e.g., wavelengths) consumed by the point-to-multipoint session.

In Figures 5.5-5.7 we plot the behavior of the algorithms in terms of the three metrics as a function of the number N of nodes in the network, for light-trees with a number of destinations equal to 15% of the number of nodes; Each figure shows three pairs of plots, each pair corresponding to one of the variants of the BLT algorithm, BLT-SP, BLT-MSR, and BLT-D. The two plots within each pair correspond to two light-trees: the initial light-tree T_0, provided as input to the BLT algorithm [2], and the final light-tree returned by the algorithm after the tree balancing procedure (the **while** loop in Steps 4-18 of Figure 5.4).

Figure 5.5 shows the maximum split ratio for the three algorithms, before and after the tree balancing procedure. Let us first concentrate on the initial trees. As we can see, the initial tree for BLT-SP has the worst average performance, while the maximum split ratios of the initial trees for BLT-MSR and BLT-D are much smaller (especially for large networks), with BLT-D being slightly better than BLT-MSR. In particular, the maximum split ratio of the initial tree constructed by BLT-SP is significantly larger than the size of the destination set; for instance, for $N = 100$, the destination set has 15 nodes, but the maximum split ratio is around 48; in other words, without amplification, the corresponding destination node would have received (1/48)-th of the power of the signal transmitted by the source. Even after amplification, this signal will have undergone severe degradation due to splits. Note

[2] Note that, while the SPH algorithm [76] is used to construct the initial light-tree T_0, a different criterion is used by each BLT variant to determine how a new destination node is connected to the partial tree, as we explained in Section 5.3.2. Therefore, the initial light-tree is different for each BLT variant.

that BLT-SP corresponds to the pure SPH algorithm [76], which has been used extensively in the literature for the Steiner tree problem. Naturally, the SPH algorithm does not take into account optical layer power constraints, and thus, it may produce very unbalanced trees. This result indicates that algorithms not specifically designed with these constraints in mind would have very poor performance in the context of optical layer multicast. On the other hand, BLT-MSR and BLT-D are variants of SPH that take the split ratio into account when building the initial tree. As we can see, such customization results in significant improvements in performance with respect to this metric.

Let us now turn our attention to the final trees produced by the three algorithms. We immediately see that the tree balancing procedure is successful in reducing significantly the maximum split ratio from that of the initial tree, for all three algorithms. Specifically, the improvement (decrease) in the maximum split ratio ranges from about 50% (for the BLT-MSR and the BLT-D algorithms) to 70% (for the BLT-SP algorithm). In other words, the signal quality at the destination where it is worst, is 50-70% better, depending on the algorithm, in the final, balanced tree compared to the initial tree. Furthermore, the maximum split ratio of the final trees increases more slowly with the number of nodes than that of the initial trees. We also observe that the BLT-SP algorithm shows the best improvement after the balancing operation, and its final trees SP have a maximum split ratio smaller than that of the corresponding final trees constructed by BLT-MSR and BLT-D. This result is due to the fact that the BLT-SP algorithm does not impose any constraints on the final tree (e.g., compared to the BLT-D algorithm), and thus, it is able to find better trees. Overall, the results of Figure 5.5 suggest that the suite of BLT algorithms can be used to construct light-trees with good performance in terms of signal power degradation. Consequently, light-trees can scale to large destination sets and networks sizes. Such scalability may not be possible with currently available algorithms, since the resulting light-trees (refer to the initial tree for BLT-SP in Figure 5.5) have a high maximum split ratio which also increases quickly with the number of network nodes.

Figure 5.6 plots the maximum-to-minimum split ratio for the initial and final trees of all three algorithms. This is a measure of the worst to best signal power at the destinations, i.e., a measure of fairness. As we can see, BLT-SP has the worst performance (both for the initial and final trees), while the performance of the initial trees constructed by BLT-MSR and BLT-D is better. More importantly, the final trees of BLT-MSR and BLT-D have a very low value (around 2.5), suggesting fair treatment of the destination

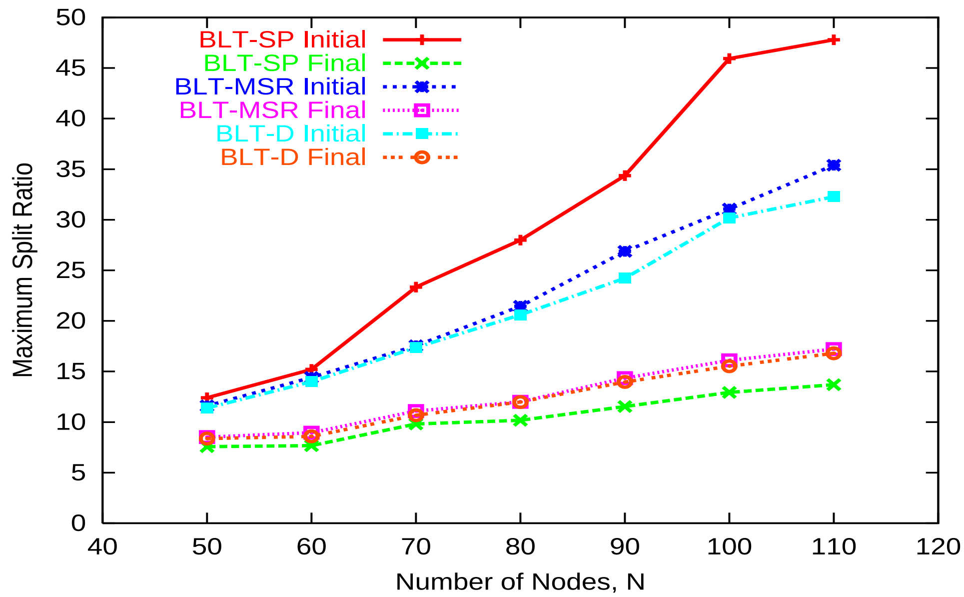


Figure 5.5: Maximum split ratio, destination set size = $.15N$

nodes. Furthermore, this low value of the maximum-to-minimum split ratio remains almost constant across the range of network sizes considered, again indicating that the fairness property scales to networks and destination sets of realistic size.

Figure 5.7 plots the number of edges of the initial and final trees for the three algorithms. The trees constructed by BLT-SP have fewer edges than those by BLT-MSR and BLT-D. Also, performing tree balancing increases the number of edges of the final tree, regardless of the algorithm employed. This result illustrates the penalty involved in balancing the tree to reduce the maximum split ratio and improve the signal quality at the destinations where it is worst. In order to balance the light-tree, destinations with high split ratios are added closer to the source by extending the tree and using additional relay nodes and edges. Consequently, balanced trees use additional network resources, including relay nodes, links, and wavelengths. Thus, there is a tradeoff between using resources efficiently and balancing the light-tree to accommodate optical layer power constraints.

Very similar results have been obtained when the number of destination nodes is equal to 5% or 10% of the number of nodes. The results are presented in Figures 5.8-5.13.

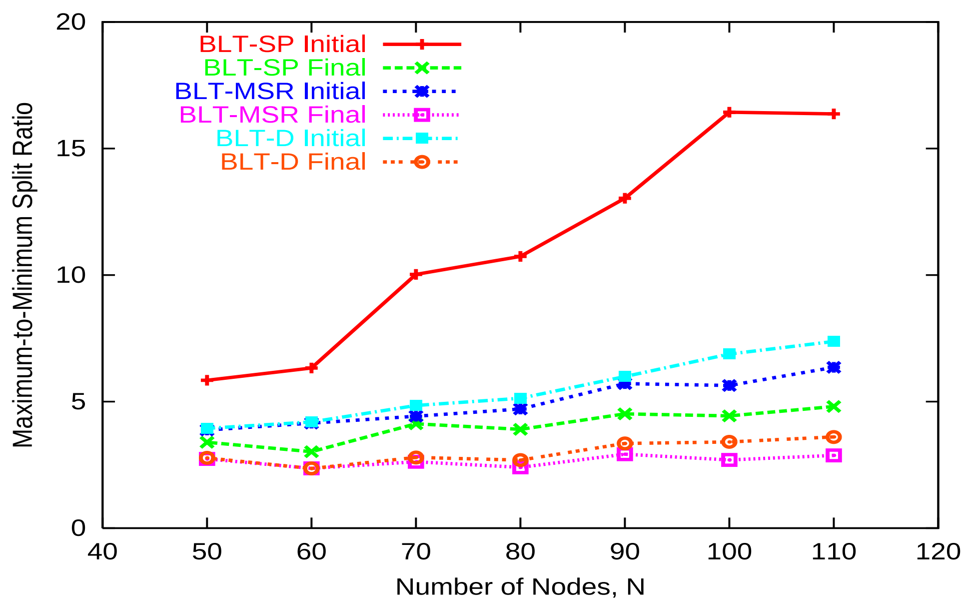


Figure 5.6: Maximum-to-minimum splitting ratio, destination set size = $.15N$

5.4.2 The WBLT Algorithm

We now demonstrate the operation of the WBLT algorithm which constructs distance-weighted light-trees by taking into account losses due to both attenuation and light-splitting. We define parameter $S, S > 0$, to capture the relative importance of loss due to attenuation and loss due to power splitting: when $S > 1$, loss due to attenuation is the dominant component of total loss, while when $S < 1$, splitting loss dominates. As we discussed above, the value of S (i.e., whether it is greater than or less than one) depends on several network parameters including the diameter of the network, the destination set size, the distance between amplifiers, and the technology of amplifiers, SaD switches, and power splitters. By varying the value of S we are able to investigate a wide range of relative values for the power splitting and attenuation losses.

Figure 5.14 plots the maximum loss (in dB), against parameter S. Due to space constraints, we only show results for the WBLT-D algorithm. The initial tree is constructed using Dijkstra's algorithm, and consists of the shortest paths from the source to all destinations; thus, this tree minimizes loss due to attenuation. The figure shows three pairs of plots: one for the total loss, one for the loss due to attenuation, and one for loss due to

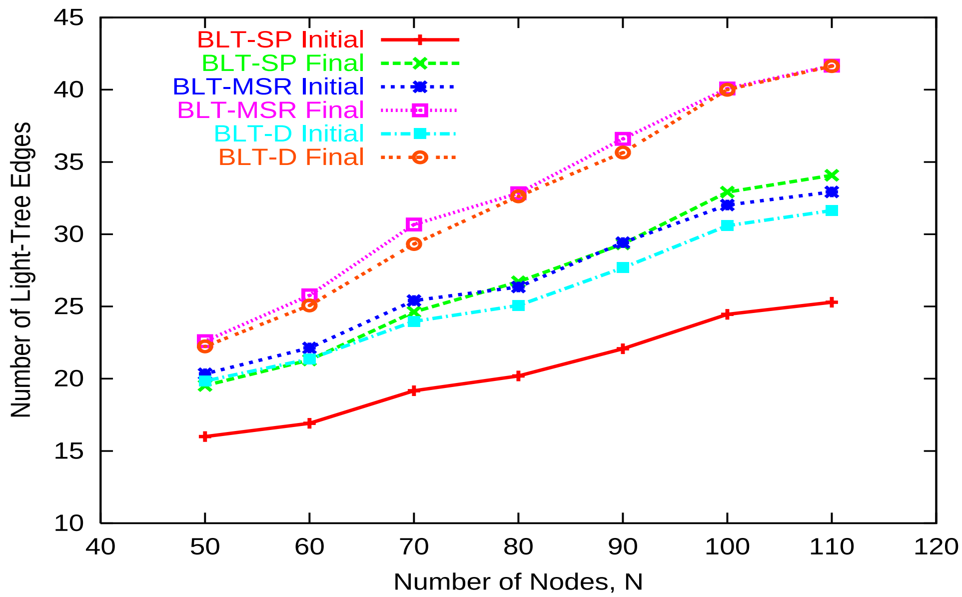


Figure 5.7: Cost, destination set size = $.15N$

power splitting. Each pair consists of one plot corresponding to the initial tree, and one corresponding to the final tree after applying WBLT-D to the initial tree.

As we can see from Figure 5.14, the total loss tracks the dominant loss component (attenuation or power splitting). The total loss is smaller for the final tree, especially when loss due to power splitting dominates. The decrease in total loss can be more than 50% at low values of S (note that both axes are shown in log scale). The plots corresponding to loss due to attenuation and power splitting explain how the distance-weighted balancing operation of WBLT is successful in reducing the total loss. Specifically, WBLT moves nodes that are far away from the source (in geographical distance) closer to the source in the light-tree. Doing so increases the loss due to attenuation (compare the corresponding plots for the initial and final tree), but reduces the loss due to power splitting (again, compare the corresponding plots). This operation is particularly successful when loss due to power splitting is dominant or even roughly equivalent to loss due to attenuation (i.e., for values of S up to 3 in the figure). When loss due to attenuation is dominant (e.g., for $S = 10$), the WBLT algorithm has little effect on total loss. This result is expected, of course, since the initial tree is optimal with respect to attenuation, and any reduction in loss due to power splitting would have negligible effect on total loss.

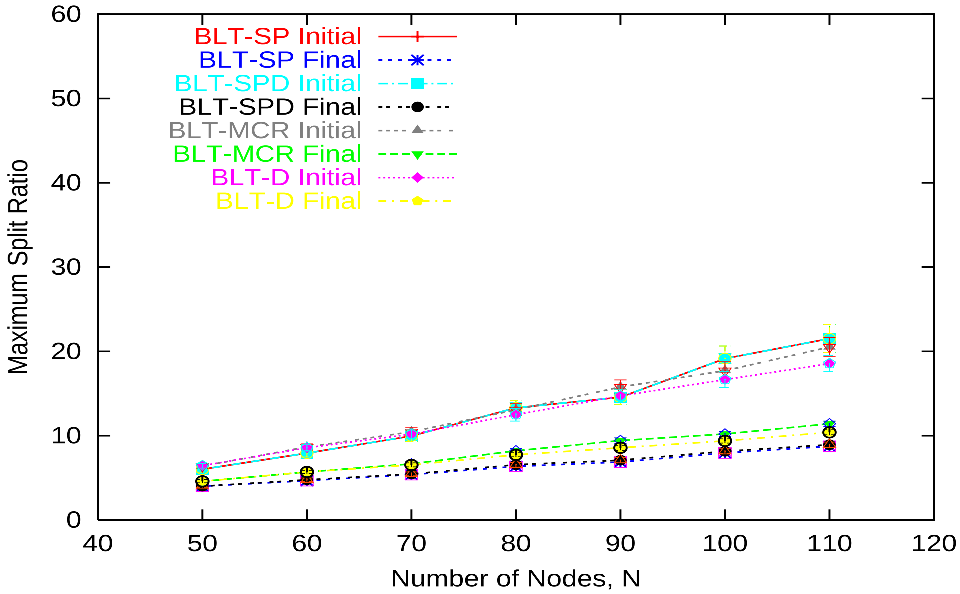


Figure 5.8: Maximum split ratio, destination set size = $.10N$

5.5 Concluding Remarks

We have studied the light-tree routing problem under optical layer power budget constraints. We considered both attenuation and splitting loss as factors affecting the quality of signals delivered to the destination nodes. We introduced a set of constraints on the end-to-end paths in order to guarantee an adequate signal quality and to ensure a measure of fairness among the destination nodes. These constraints require the light-tree to be balanced or distance-weighted balanced. We proved that constructing such a light-tree spanning a given source and destination node set is an NP-complete problem. We developed a number of algorithms for building balanced trees, and we investigated their performance through extensive simulation experiments on a large number of randomly generated network topologies.

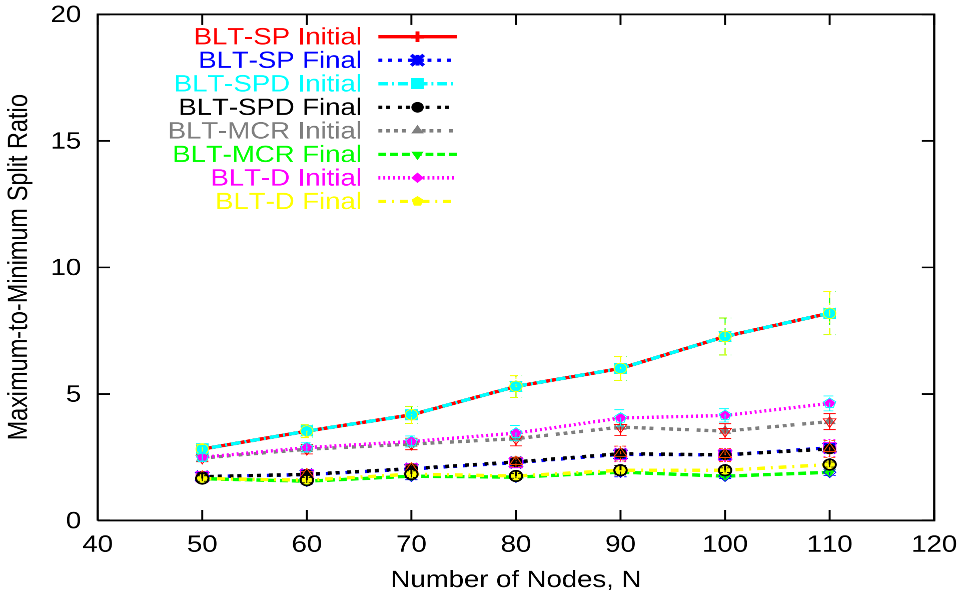


Figure 5.9: Maximum-to-minimum splitting ratio, destination set size = $.10N$

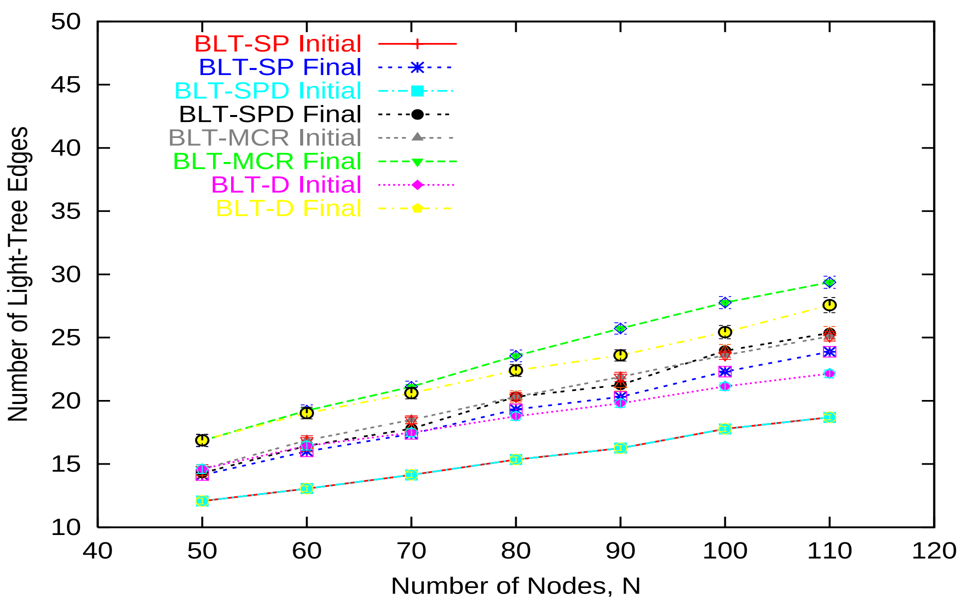


Figure 5.10: Cost, destination set size = $.10N$

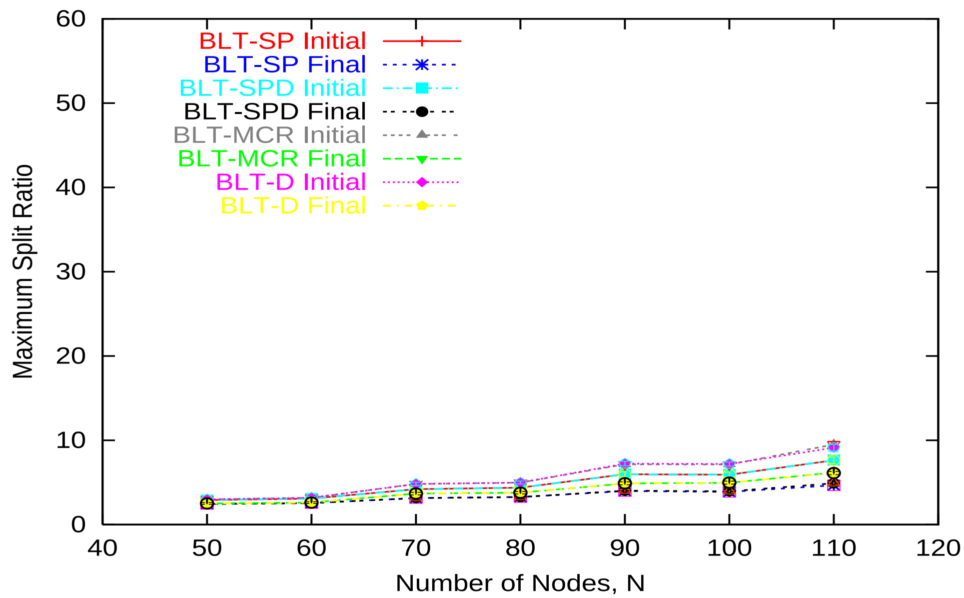


Figure 5.11: Maximum split ratio, destination set size = $.5N$

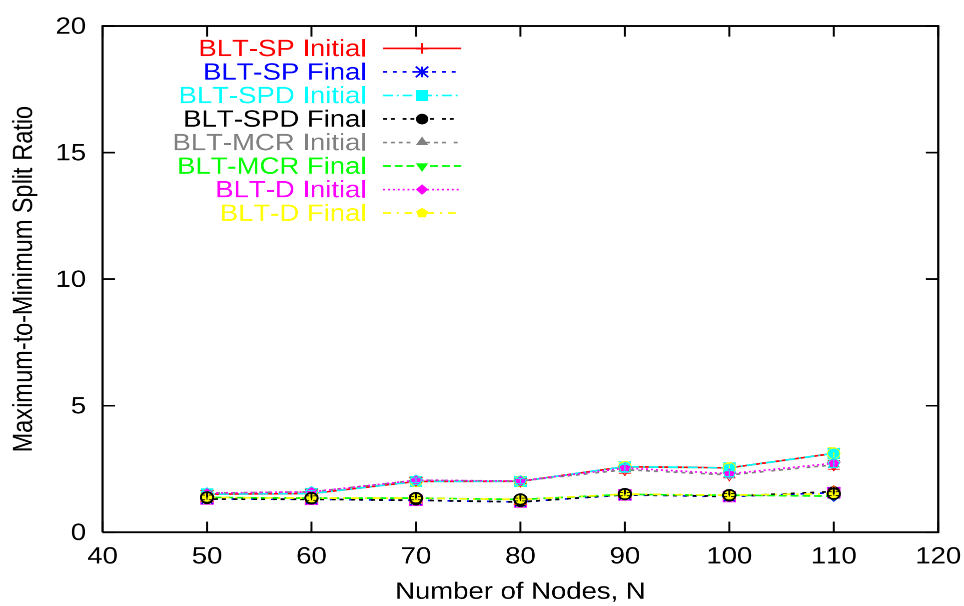


Figure 5.12: Maximum-to-minimum splitting ratio, destination set size = $.5N$

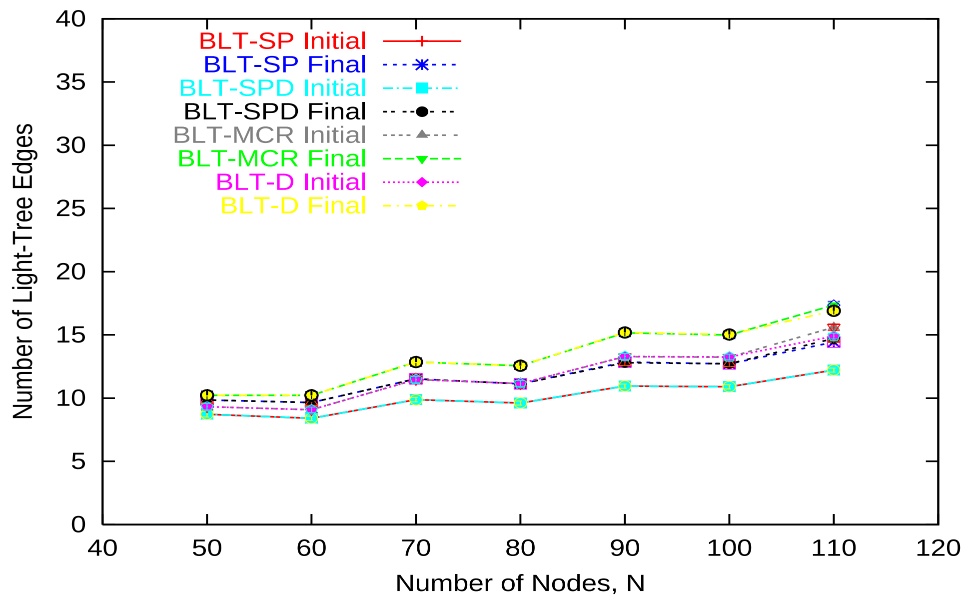


Figure 5.13: Cost, destination set size = $.5N$

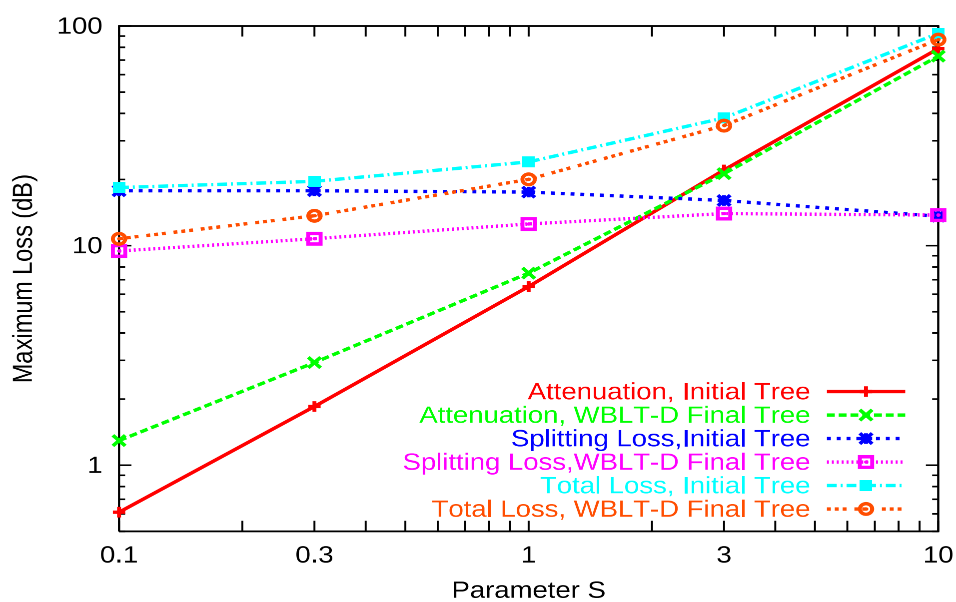


Figure 5.14: Maximum loss for WBLT-D, destination set size = $.15N$

Chapter 6

Conclusions and Future Work

We have examined the problem of integrated physical topology and virtual topology design for large scale optical networks. Our emphasis has been on providing efficient heuristic algorithms to find near optimal solutions for optical networks that are required to support up to several thousands of LSRs, each of which may be equiped with tens of transceivers. We extended our study to include the survivability requirement into the network design. The results demonstrate that it is possible to design a network that scales well with the size of the LSRs and the traffic requirements. We have also examined an important application for optical networks, the light-tree routing problem, which is mainly used to support multicasting at the optical layer. Our algorithm takes the power loss into consideration and guarantees a certain level of fairness for the power distribution among members in a multicasting group through the construction of a balanced light-tree.

6.1 Future Work

Our work on the physical and logical topology design and optical layer multicast can be extended in several directions. As we mentioned earlier, at the optical layer, the offered traffic in the form of lightpaths (logical topology) appears static at certain time scales. Furthermore, most topology design studies (including ours) assume that the majority of the traffic is point-to-point. However, over the long term, traffic patterns may change, and the offered traffic may increase and require additional resources. Therefore, the problem of migrating the virtual topology to satisfy the new requirements is an important one. Furthermore, a different objective in physical topology design would be to implement a physical topology that would be able to satisfy the maximum increase in offered traffic without additional resources (wavelength or OXCs). Also, since multicast services are becoming popular in the Internet, and the amount of multicast traffic is expected to increase significantly, then multicast demands must be taken into account in the design of optical networks. In this case, we have to investigate the routing and wavelength assignment problem for multicast connections.

Another attractive field of study is the design of a survivable network. In this work, we only considered the protection requirement in the optical layer, but we did not consider the interoperation between the optical layer and upper service layer(s) (e.g., the IP or MPLS layer). We need to conduct additional research on efficient protection/restoration schemes in such a network.

Although link-based protection is considered highly capacity-consuming, its fast restoration time is very attractive to network service providers and worthy of further investigation. In an optical network, one or more primary lightpaths with different wavelengths may use a particular link. When this link fails, we have to find one or more backup paths with the same number of wavelengths to recover the broken lightpaths. If wavelength converters are used, finding backup paths for a particular faulty link is equivalent to the integer flow multicommodity problem. We only need to assign the number of wavelengths to the backup paths, but the backup wavelength need not be the same as the primary wavelength. The problem becomes more complicated when converters are not used since we have to find the same group of wavelengths in the backup paths for the recovery of primary ligthpaths, which may be not always available. The wavelengths we need may have been occupied by other primary or backup lightpaths. Therefore we may not be able to recover all the ligthpaths based on the link protection. In order to demonstrate the performance of the link-based protection scheme, we may define a recovery ratio, γ, to represent the weighted average percentage of recovered primary lightpaths over the total primary ligthpaths for each link. Efficient algorithms need to be developed to minimize γ.

An extension to the light-tree routing problem can be made to consider networks with sparse light splitting. In our current network model, we assume that every OXC is multicast-capable. However, multicast-capable OXCs are very expensive and not technically ready for the commercial deployment. A more practical assumption would be that only a fraction of the OXCs are multicast-capable. the optimal allocation of the multicast-capable OXCs and the balanced light-tree routing problem under such a sparse light splitting scenario are both valuable problems worthy of further investigation. A possible approach is to combine the concept of a light-forest with our balancing algorithm to obtain a balanced light-forest for a given multicast group, which will minimize both the number of wavelengths and the power loss.

Bibliography

[1] M. Alanyali and E. Ayanoglu. Provisioning algorithms for WDM optical networks. *IEEE/ACM Transactions on Networking*, 7(5):767–778, October 1999.

[2] M. Ali and J. Deogun. Allocation of splitting nodes in wavelength-routed networks. *Photonic Network Communications*, 2(3):245–263, August 2000.

[3] M. Ali and J. Deogun. Power-efficient design of multicast wavelength-routed networks. *IEEE Journal on Selected Areas in Communications*, 18(10):1852–1862, 2000.

[4] Vishal Anand and Chunming Qiao. Static versus dynamic establishment of protection paths in wdm networks. *Journal of High Speed Network*, 10(4):317–327, 2001.

[5] A. Banerjee, J. Drake, J. P. Lang, B. Turner, D. Awduche, L. Berger, K. Kompella, and Y. Rekhter. Generalized multiprotocol label switching: An overview of signalling enhancement and recovery techniques. *IEEE Communication Magazine*, pages 144–151, July 2001.

[6] A. Banerjee, J. Drake, J. P. Lang, B. Turner, K. Kompella, and Y. Rekhter. Generalized multiprotocol label switching: An overview of routing and management enhancements. *IEEE Communication Magazine*, pages 144–150, January 2001.

[7] D. Banerjee and B. Mukherjee. Wavelength-routed optical networks: Linear formulation, resource budgeting, and a reconfiguration study. *IEEE/ACM Transactions on Networking*, 8(5):598–607, October 2000.

[8] S. Baroni and P. Bayvel. Key topological parameters for the wavelength-routed optical networks. In *Proceeding of ECOC 1996*, 1996.

[9] S. Baroni and P. Bayvel. Wavelength requirements in arbitrarily connected wavelength-routed optical networks. *IEEE/OSA Journal of Lightwave Technology*, 15(2):242–251, February 1997.

[10] S. Baroni, J.O. Eaves, M. Kumar, and M.A. Qureshi. Analysis and design of backbone architecture alternatives for IP optical networks. *IEEE Journal on Selected Areas in Communications*, 18(10):1980–1994, October 2000.

[11] R. A. Barry and P. A. Humblet. On the number of wavelengths and switches in all-optical networks. *IEEE Transactions on Communications*, 42(2/3/4):583–591, Feb-Apr 1996.

[12] F. Bauer and A. Varma. Degree-constrained multicasting in point-to-point networks. In *Proceedings of INFOCOM '95*, pages 369–376. IEEE, April 1995.

[13] B. Beauquier. All-to-all communication for some wavelength routed all-optical networks. *Networks*, 30, 1999.

[14] Ramesh Bhandari. *Survivable networks: algorithms for diverse routing*. Kluwer, USA, 1999.

[15] I. Chlamtac, A. Ganz, and G. Karmi. Lightpath communications: An approach to high bandwidth optical WANS. *IEEE Transactions on Communications*, 40(7):1171–1182, July 1992.

[16] C. Chow. On multicast path finding algorithms. In *Proceedings of INFOCOM '91*, pages 1311–1317. IEEE, 1991.

[17] P. B. Chu, S-S. Lee, and S. Park. MEMS: The path to large optical crossconnects. *IEEE Communications*, 40(3):80–87, March 2002.

[18] O. Crochat, J. L. Boudec, and O. Gerstel. Protection interoperability for wdm optical networks. *IEEE/ACM Trans. On Networks*, 8(3):384–395, June 2000.

[19] B. Davie and Y. Rekhter. *MPLS Technology and Applications*. Morgan Kaufmann Publishers, San Diego, California, 2000.

[20] M. Doar and I. Leslie. How bad is naive multicast routing. In *Proceedings of INFOCOM '93*, pages 82–89. IEEE, April 1993.

[21] John Doucette, Matthieu Clouqueur, and Wayne D. Gover. On the service availability and respective capacity requirements of shared backup path protected mesh networks. *Submission to the Optical Network Magzine*, March 2002.

[22] John Doucette and Wayne D. Grover. Influence of modularity and economy-of-scale effects on design of mesh-restorable dwdm networks. *IEEE Journal on Selected Areas in Communications*, 18(10):1912–1923, October 2000.

[23] John Doucette and Wayne D. Grover. Comparison of mesh protection and restoration schemes and the dependency on graph connectivity. In *DRCN'2001*, pages 1–9, 2001.

[24] K. A. Dowsland. Hill-climbing simulated annealing and the steiner problem in graphs. *Eng. Opt.*, 17:91–107, 1991.

[25] R. Dutta and G. N. Rouskas. A survey of virtual topology design algorithms for wavelength routed optical networks. *Optical Networks*, 1(1):73–89, January 2000.

[26] Tali Eilam-Tzoreff. The disjoint shortest paths problem. *Discrete applied methematics*, 85:113–138, October 1998.

[27] G. Ellinas, R. G. Hailemarian, and T. E. Stern. Protection cycles in mesh wdm networks. *IEEE Journal on Selected Areas in Communications*, 18:1924–1937, October 2000.

[28] H. Esbensen. Computing near-optimal solutions to the steiner problem in a graph using a genetic algorithm. *Networks*, 26:173–185, 1995.

[29] D. Papadimitriou *et al.* Optical multicast in wavelength switched networks – architectural framework. IETF Draft <draft-poj-optical-multicast-01.txt>, July 2001. Work in progress.

[30] P. Ashwood-Smith *et al.* Generalized MPLS – signaling functional description. IETF Draft <draft-ietf-mpls-generalized-signaling-06.txt>, April 2001. Work in progress.

[31] E. Mannie etc. Generalized multi-protocol label switching (gmpls) architecture. *IETF Internet Draft*, draft-ietf-ccamp-gmpls-architecture-01.txt, November 2001.

[32] P. F. Fonseca. Pan-european multi-wavelength transport network network design, architecture, survivability and sdh networking. In *DRCN'98*, page P3, 1998.

[33] A. Fumagalli and L. Valcarenghi. Ip restoration vs. wdm protection: Is there an optimal choice. *IEEE Network Magazine*, pages 34–41, Nov./Dec. 2000.

[34] M. R. Garey, R. L. Graham, and D. S. Johnson. The complexity of computing steiner minimal trees. *SIAM Journal of Applied Mathematics*, 32(4):835–859, June 1977.

[35] M. R. Garey and D. S. Johnson. *Computers and Intractability*. W. H. Freeman and Co., New York, 1979.

[36] M. Gen and R. Cheng. *Genetic Algorithms and Engineering Optimization*. Wiley, New York, 2000.

[37] Ornan Gerstel and Rajiv Ramaswami. Optical layer survivability-an implementation perspectivce. *IEEE Journal on Selected Areas in Communications*, 18(10):1885–1899, October 2000.

[38] Wayne D. Grover and D. Stamatelakis. Cycle-oriented distributed preconfiguration: Ring-like speed with mesh-like capacity for self-planning network reconfiguration. In *ICC'1998*, pages 537–543. IEEE, 1998.

[39] F. Guerriero and R. Musmanno. Label correctin methods to solve multicriteria shortest path problems. *J. of Optimizition Theory and Application*, 111(3):589–613, Dec. 2001.

[40] S. L. Hakimi. Steiner's problem in graphs and its implications. *Networks*, 1:113–133, 1971.

[41] D.R. Hjelme, A. Royset, and B.J. Slagsvold. How many wavelengths does it take to build a wavelength routed optical network. In *Proceeding of ECOC 1996*, 1996.

[42] W. S. Hu and Q. J. Zeng. Multicasting optical cross connects employing splitter-and-delivery switch. *IEEE Photonics Technology Letters*, 10:970–972, July 1998.

[43] X.-H. Jia, D.-Z. Du, X.-D. Hu, M.-K. Lee, and J. Gu. Optimization of wavelength assignment for QoS multicast in WDM networks. *IEEE Transactions on Communications*, 49(2):341–350, Feb. 2001.

[44] E. Karasan and E. Ayanoglu. Effects of wavelength routing and selection algorithms on wavelength conversion gain in WDM optical networks. *IEEE Journal on Selected Areas in Communications*, 16(7):1081–1096, September 1998.

[45] R. Kawamura. Architectures for atm network survivability. *IEEE Communication Survey*, 1(1), 1998.

[46] V. P. Kompella, J. C. Pasquale, and G. C. Polyzos. Multicast routing for multimedia communication. *IEEE/ACM Transactions on Networking*, 1(3):286–292, June 1993.

[47] L. Kou, G. Markowsky, and L. Berman. A fast algorithm for steiner trees. *Acta Informatica*, 15:141–145, 1981.

[48] J. Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. Amer. Math. Soc.*, 7:48–50, . 1956.

[49] Myungmoon Lee, Jintae Yu, Yongbum Kim, and Jinwoo Park. A restoration method independent of failure location in all-optical networks. *Computer Communication*, 25(10):915–921, June 2002.

[50] T. Lee, K. Lee, and S. Park. Optimal routing and wavelength assignment in WDM ring networks. *IEEE Journal on Selected Areas in Communications*, 18(10):2146–2154, October 2000.

[51] Juerg Leuthold and Charles H. Joyner. Multimode interference couplers with tunable power splitting ratios. *IEEE/OSA Journal of Lightwave Technology*, 19(5):700–706, May 2001.

[52] D. Li, X. Du, X. Hu, L. Ruan, and X. Jia. Minimizing number of wavelengths in multicast routing trees in WDM networks. *Networks*, 35(4):260–265, 2000.

[53] M.A. Marsan, A. Bianco, and E. Leonardi. Topologies for wavelength-routing all-optical networks. *IEEE/ACM Transactions on Networking*, 1(5):534–546, October 1993.

[54] Carmen Mas and Patrick Triran. An efficient algorithm for locating soft and hard failure in wdm networks. *IEEE Journal on Selected Areas in Communications*, 18(10):1900–1911, October 2000.

[55] M. Medard, R. A. Barry, S. G. Finn, W. He, and S. S. Lumetta. Generalized loop-back recovery in optical mesh networks. *IEEE/ACM Trans. On Networks*, 10(1):153–164, February 2002.

[56] B. Mukherjee. *Optical Communication Networking.* McGraw-Hill, 1997.

[57] N. Nagatsu, S. Okamoto, and K. Sato. Optical path cross-connect system scale evaluation using path accommodation design for restricted wavelength multiplexing. *IEEE Journal Selected Areas in Communications*, 14(5):893–902, June 1996.

[58] C.C. Palmer and A. Kershenbaum. An approach to a problem in network design using genetic algorithms. *Networks*, 26:151–163, 1995.

[59] R.K. Pankaj and R.G. Gallager. Wavelength requirements of all-optical networks. *IEEE/ACM Transactions on Networking*, 3(3):269–280, June 1995.

[60] S. Ramamurthy and B. Mukherjee. Survivable wdm mesh network, part i-protection. In *INFOCOM-99*, pages 744–751. IEEE, 1999.

[61] R. Ramaswami and K. Sivarajan. Routing and wavelength assignment in all-optical networks. *IEEE/ACM Transactions on Networking*, 3(5):489–500, October 1995.

[62] R. Ramaswami and K. N. Sivarajan. *Optical Networks*. Morgan Kaufmann Publishers, San Francisco, California, 1998.

[63] C. C. Ribeiro and M. C. De Souza. Tabu search for the steiner problem in graphs. *Networks*, 36(2):138–146, September 2000.

[64] E. Rosen, A. Viswanathan, and R. Callon. Multiprotocol label switching architecture. RFC 3031, January 2001.

[65] G. N. Rouskas and I. Baldine. Multicast routing with end-to-end delay and delay variation constraints. *IEEE Journal on Selected Areas in Communications*, 15(3):346–356, April 1997.

[66] D. Saha, M.D. Purkayastha, and B. Mukherjee. An approach to wide area WDM optical network design using genetic algorithms. *Computer Communications*, 22:156–172, 1999.

[67] L. Sahasrabuddhe, S. Ramamurthy, and B. Mukherjee. Fault management in ip-over-wdm networks: Wdm protection vs. ip restoration. *IEEE Journal on Selected Areas in Communications*, 20(1):21–33, Janurary 2002.

[68] L. H. Sahasrabuddhe and B. Mukherjee. Light-trees: Optical multicasting for improved performance in wavelength-routed networks. *IEEE Communications*, 37(2):67–73, February 1999.

[69] G. Sahin and M. Azizoglu. Multicast routing and wavelength assignment in wide-area networks. In *Proceedings of SPIE*, volume 3531, pages 196–208, November 1998.

[70] Arunabha Sen, Baohong Shen, and Subir Bandyopadhyay. Survivability of lightpath networks-path length in wdm protection scheme. *Journal of High Speed Network*, 10(4), 2001.

[71] P-R. Sheu and S-T. Chen. A fast and efficient heuristic algorithm for the delay- and delay variation-bounded multicast tree problem. *Computer Communications*, 2001.

[72] M. Srinivas and M. Patnaik. Genetic algorithms: A survey. *IEEE Computer*, 27(6):17–26, June 1994.

[73] T. E. Stern and K. Bala. *Multiwavelength Optical Networks*. Prentice Hall, Upper Saddle River, New Jersey, 2000.

[74] J. Strand, A. L. Chiu, and R. Tkach. Issues for routing in the optical layer. *IEEE Communications*, pages 81–96, February 2001.

[75] D. H. Su and D. W. Griffith. Standards activities for MPLS over WDM networks. *Optical Networks*, 1(3), July 2000.

[76] H. Takahashi and A. Matsuyama. An approximate solution for the steiner problem in graphs. *Math. Japonica*, 24(6):573–577, 1980.

[77] B. Mukherjee *et al.* Some principles for designing a wide-area WDM optical network. *IEEE/ACM Transactions on Networking*, 4(5):684–696, October 1996.

[78] L. Tran, K. Steenhaut, and Ann Nowe. Efficient usage of capacity resources in survivable mpλs networks. In *Submission to Net'2002*, 2002.

[79] W.-Y. Tseng and S.-Y. Kuo. All-optical multicasting on wavelength-routed WDM networks with partial replication. In *ICIN-2001*, pages 813–818. IEEE, 2001.

[80] Jens Vygen. Np-completeness of some edge-disjoint paths problems. *Discrete applied methematics*, 61:83–90, 1995.

[81] B. W. Waxman. Routing of multipoint connections. *IEEE Journal on Selected Areas in Communications*, 6(9):1617–1622, December 1988.

[82] K.-D. Wu, J.-C. Wu, and C.-S. Yang. Multicast routing with power consideration in sparse splitting WDM networks. In *ICC-2001*, pages 513–517. IEEE, 2001.

[83] 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*, 18(10):2130–2137, October 2000.

[84] X. Zhang, J. Y. Wei, and C. Qiao. Constrained multicast routing in WDM networks with sparse light splitting. *Journal of Lightwave Technology*, 18(12):1917–1927, December 2000.

[85] Z. Zhang and A. Acampora. A heuristic wavelength assignment algorithm for multihop WDM networks with wavelength routing and wavelength reuse. *IEEE/ACM Transactions on Networking*, 3(3):281–288, June 1995.

[86] Keyao Zhu, Laxman Sahasrabuddhe, and Biswanath Mukherjee. Topology design and upgrade of an optical network by bottlenect-cut identification. *Journal of High Speed Network*, 10(4):293–301, 2001.