Dynamic service provisioning mechanisms for next-generation optical internet

Yu, Xiaojun

2014

Yu, X. (2014). Dynamic service provisioning mechanisms for next-generation optical internet. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/61868>

<https://doi.org/10.32657/10356/61868>

Nanyang Technological University

Dynamic Service Provisioning Mechanisms for Next-Generation Optical Internet

by

Yu Xiaojun

School of Electrical and Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy

November 2014

Declaration of Originality

I, **Yu Xiaojun**, declare that this thesis titled, “**Dynamic Service Provisioning Mechanisms for Next-Generation Optical Internet**” and the work presented in it are my own. I confirm that:

- This work was done wholly by myself while in candidature for a PhD degree at Nanyang Technological University.
- None of the work in this thesis has previously been submitted for a degree or any other qualification at NTU or any other institution.
- Any work that I have cited from the others, the source is always given. With the exception of such citations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.

Signed: Yu Xiaojun

Date: November 2014

Acknowledgements

Studying in NTU is one of the most precious experiences in my life. I owe my Ph.D to all the people around me who have offered me their encouragement, support and help. Without them, this dissertation would not have been completed.

First, I would like to express my utmost gratitude to my supervisor, Professor Cheng Tee Hiang, who gave me the chance to study in NTU under his supervision. During the past few years of my study, Prof. Cheng has provided me with the invaluable guidance, enthusiastic advice and generous support. Whenever I needed any help, Prof. Cheng has always been there offering it without any hesitation. Even when I was deeply frustrated with my little research progress within the beginning years, Prof. Cheng has never lost his faith in me; instead, he helped me regain my confidence with his unwavering support, encouragement and patience. I deem it my good fortune to be Prof. Cheng's student. I really do not know how to express my appreciation apart from saying thank you.

I am deeply appreciative of the help from Assoc. Prof. Xiao Gaoxi. Throughout my Ph.D study, Prof. Xiao has generously shared his vast research experiences and knowledge with me. He is always able to find some time for insightful discussions even when he was overloaded with work. His depth of knowledge, passion in research and dedication to excellence help me to survive from the difficult times. Apart from his academic support, Prof. Xiao has also taught me his strictness to details, positive attitude towards life and responsibility to family. All the discussions that I have had with Prof. Xiao will be the most invaluable treasure of my future life.

I would also like to thank all the excellent staff members of OPTIMUS- Photonics Centre of Excellence (previously NTRC) for providing me the outstanding family-like research environment and excellent laboratory facility. I will always be proud

of being a member of OPTIMUS and an NTRCer. My particular thanks go to Professor Shum Ping, who is enthusiastic, generous, and always ready there to help whenever anyone met with any difficulty; Asst. Prof. Luan Feng, who offered me my first job in NTU, and helped me out during the critical time; and Nanyang Asst. Prof. Liu Linbo for his generous help.

I am also grateful to all my wonderful colleagues and good friends Mr. Chong Lui Tat, Mr. Poh Khoo Yong, Lin Rongping, Han Yang, Wang Dawei, Huang Dezhaoh, Hu Wuhua, Wu Kan, Zhu Fan, Yu Shengrong(Timothy), Han Shiyong, Wang Zixiong, Liu Huanhuan, Xiong Fei, Li Yihui, Wong Jiahaur, Xiao Zhe, Wang Yubo, Ngor Pengty, Xie Linfu, Li Xiang, Zhu Senqiang, Luo Shan, Hu Yongxu, Mao Xiaohong, and all the other nice friends that I have met during my study and work in NTU. I benefited quite a lot from the many discussions with them. The wonderful time I have spent with them will always live in my memory. I wish them all the best in their future life.

I particularly wish to thank my parents and parents-in-law for their unconditional love and endless support. They are always supportive of the decisions I made. Their true love can always be felt and touched.

Last but not least, I would like to express my deepest love and gratitude towards my wife, Du Shujuan, for her unwavering love and boundless tolerance. Years before, she left her parents and friends in China to come to Singapore to support me. After the birth of our lovely baby, she gave up her job to return back to China to take care of our baby with great love and patience. Without the sacrifices she has made, it would be impossible for this dissertation to be completed.

Contents

Declaration of Originality	iii
Acknowledgements	v
Table of Contents	vii
Summary	xi
List of Figures	xiii
List of Tables	xv
Abbreviations	xvii
1 Introduction	1
1.1 Background and Motivations	1
1.2 Objectives	7
1.3 Original Contributions	9
1.4 Thesis Organization	11
2 Literature Review	15
2.1 Introduction	15
2.2 Emerging Technologies	16
2.2.1 Wavelength routed WDM networks	16
2.2.2 IP/MPLS over WDM	18
2.2.3 Lightpath	25
2.2.4 Light-tree	26
2.3 Traffic Grooming	31
2.3.1 Static and dynamic traffic grooming	33
2.3.2 Dynamic LSP routing for overlay IP/MPLS over WDM	35
2.3.3 Dynamic multicast grooming in WDM networks	38
2.4 Multicast Request Protection	41
2.5 Summary	45
3 Overlay Dynamic LSP Routing	47
3.1 Introduction	47
3.2 System Model and Definition	49

3.3	Proposed Algorithm	54
3.3.1	Graph generation and cost assignment	54
3.3.2	Data learning and cost expiration process	58
3.3.3	ELF algorithm	61
3.3.4	Algorithm complexity	62
3.4	Performance Evaluation	63
3.4.1	Performance in port-unlimited networks	65
3.4.2	Performance in port-limited networks	67
3.4.3	Port number influence	68
3.4.4	Historical data influence	70
3.5	Summary	71
4	Dynamic Multicast Traffic Grooming	73
4.1	Introduction	73
4.2	Lightpath VS. Light-Tree: A Brief Comparison	75
4.2.1	Physical layer node architectures	75
4.2.2	Network layer multicast traffic grooming methods	76
4.2.3	Possible further improvements	78
4.3	LPF Algorithm	79
4.3.1	Problem statement and main idea	79
4.3.2	Selection of fragmentation nodes	80
4.3.3	LPF dynamic multicast grooming algorithm	82
4.4	Results and Discussions	85
4.4.1	Performance metrics and simulation environment	85
4.4.2	Traffic models and loads	88
4.4.3	Destination number and unicast ratio	93
4.4.4	Utilization, delay and transceiver consumption	95
4.5	Summary	100
5	Overlay Multicast Traffic Grooming	103
5.1	Introduction	103
5.2	Chapter Preliminaries	106
5.2.1	Network Model	106
5.2.2	Node Architecture	107
5.2.3	Previous Work	108
5.2.4	Problem Statement	110
5.3	OMP for Multicast Traffic Grooming	111
5.3.1	IP/MPLS Layer DL Scheme	111
5.3.2	MTH Heuristic	113
5.3.3	LPF for WDM Layer Routing	114
5.3.4	Detailed OMP Mechanism	118
5.4	Performance Evaluation	119
5.4.1	Simulation Setup	120
5.4.2	OMP vs. SC in Different Cases	121

5.4.2.1	Different traffic loads	121
5.4.2.2	Different port resources	124
5.4.3	IP/MPLS Layer DL Influences	127
5.4.4	WDM Layer LPF Influences	128
5.4.5	WDM Layer Lightpath Cost Definition	131
5.4.6	WDM Layer Routing Strategies Influences	134
5.5	Summary	136
6	Dynamic Survivable Multicast Traffic Grooming	137
6.1	Introduction	137
6.2	Network Model and Problem Definition	139
6.2.1	Network Model	139
6.2.2	Problem Statement	140
6.2.3	Related Existing Work	141
6.3	LF-SSP Mechanism for SMTG	143
6.3.1	Segment Shared Protection (SSP) Method	143
6.3.2	Two-phase SSP Scheme	145
6.3.2.1	First-Phase Primary-Tree Route calculation	145
6.3.2.2	Graph generate for backup routing	147
6.3.2.3	LF method for segmentation	149
6.3.2.4	Second-phase backup calculation	151
6.3.3	The Proposed LF-SSP Mechanism	154
6.4	Performance Evaluation	156
6.4.1	Simulation Setup	156
6.4.2	LF-SSP Compared to Existing CL-SSP Method	158
6.4.2.1	Comparisons in a port-unlimited network	158
6.4.2.2	Comparisons in a port-limited network	158
6.4.3	LF-SSP Compared to SPT and LP Methods	160
6.4.3.1	Comparisons in a port-unlimited network	161
6.4.3.2	Comparisons in a port-limited network	162
6.4.4	Impacts of Port Resources on Network Performances	164
6.4.5	Impacts of Session Size on Network Performances	166
6.5	Summary	169
7	Conclusions and Future Work	171
7.1	Conclusions	171
7.2	Future Research Topics	174
7.2.1	Multi-user for overlay IP over WDM	175
7.2.2	Energy efficient dynamic multicast grooming	176
7.2.3	Request protection in traffic grooming process	178
	Publications	181
	Bibliography	183

Summary

Explosive Internet traffic growth over the past decades has enabled the rapid development and wide deployment of wavelength division multiplexing (WDM)-based networking technologies. At the same time, as IP/MPLS technology continues to gain popularity, it is widely believed that next-generation optical networks will adopt an IP/MPLS over WDM architecture with wavelength-routed WDM networks serving as the backbone. Since the capacity provided by a single wavelength channel in WDM networks is typically on the order of Gigabits per second (Gbps), while the bandwidth requirement of a request is typically much smaller than that, efficient schemes are desired to improve wavelength capacity utilization in WDM networks. This dissertation investigates several dynamic traffic provisioning problems with unicast and multicast traffic requests in the context of next-generation optical Internet. The main objective is to devise efficient mechanisms to optimize the overall network blocking performance.

We first study the dynamic label switched path routing problem in an overlay IP/MPLS over WDM network. In such a network, the limited information exchanges between the two layers of the network, if not properly devised, will seriously degrade the network performance. We propose to learn from the historical data of lightpath setup costs maintained by the IP/MPLS layer integrated service provider (ISP) when making routing decisions. A novel historical data-learning scheme, called *Existing Link First*, is proposed. Simulation results demonstrate that the proposed method significantly improves the network performance.

Next, we investigate the dynamic multicast traffic grooming problem in wavelength routed WDM networks. We conduct comprehensive comparisons between

the performances of two classes of existing schemes, namely the *lightpath*-based methods and *light-tree*-based methods. Our comparison results show that the lightpath-based schemes typically achieve better blocking performances than the light-tree-based schemes. We study the results carefully and explain the underlying cause of our findings. Inspired by such observations, we then proceed to propose a new lightpath based algorithm called the *lightpath fragmentation* (LPF) method. Simulation results show that the LPF method consistently outperforms all existing methods under different traffic loads.

By combining a data-learning scheme on the IP/MPLS layer to facilitate efficient routing and a lightpath fragmentation approach on the WDM layer to improve resource sharing, we propose in the third part of our work an *overlay multicast provisioning* (OMP) mechanism for dynamic multicast traffic grooming in overlay IP/MPLS over WDM networks. We show that OMP achieves much better blocking performances than all of the existing schemes. Further studies are carried out to determine the contributions of the data-learning scheme and lightpath fragmentation method, respectively, in various cases.

Finally, we consider the problem of protecting subwavelength level multicast requests through dynamic traffic grooming in WDM networks. A new mechanism called *lightpath fragmentation-based segment shared protection* (LF-SSP) is proposed for protecting multicast requests at the connection level. By improving the efficiency of resource sharing in multicast traffic grooming and its corresponding protection process, LF-SSP manages to outperform all of the existing methods by achieving better blocking performance when supporting both sub-wavelength- and wavelength-level dynamic multicast requests with shared protection.

List of Figures

1.1	Worldwide Internet traffic growth in the past decades	2
1.2	Bandwidth intensive applications growth in recent years	2
2.1	A typical wavelength routed WDM network architecture	16
2.2	A typical $N \times N$ optical-cross-connect structure.	17
2.3	A typical overlay IP/MPLS over WDM network	21
2.4	A typical LSR Architecture	22
2.5	A typical peer model IP/MPLS over WDM architecture.	23
2.6	An augmented IP/MPLS over WDM architecture.	24
2.7	Different multicast schemes in WDM networks	27
2.8	The first proposed MC-OXC and the SaD switch	29
2.9	Two node architectures for traffic grooming	32
2.10	A multicast request protected by different protection schemes . . .	42
2.11	An example with MTG-SSP method	45
3.1	The IP/MPLS over WDM network assumed	49
3.2	A typical node architecture utilized	52
3.3	Network topologies used in Chapter 3	65
3.4	ELF compared to existing algorithms in port-unlimited networks . .	66
3.5	ELF compared to existing algorithms in port-limited networks . . .	68
3.6	ELF compared to existing algorithms with add/drop ratio	69
3.7	Influence of the amount of historical data	71
4.1	An example of multicast traffic grooming	78
4.2	Network topologies utilized for simulations	87
4.3	Lightpath vs. light-tree: different traffic modelss	89
4.3	Lightpath vs. light-tree: different traffic modelss	90
4.4	Lightpath vs. light-tree: high traffic load	91
4.5	Lightpath vs. Light-tree: different traffic loads	92
4.6	Lightpath vs. Light-tree: different destination number	94
4.7	Lightpath vs. Light-tree: different unicast ratio	96
4.8	Lightpath vs. light-tree: capacity utilization	97
4.9	Average number of OEO conversions	98
4.10	Average number of transceiver pairs utilized	99
5.1	Serving a request with LPT and SC method	109
5.2	Network topologies utilized for simulations	120

5.3	OMP vs. SC methods under different traffic loads with $r = 1.0$. . .	122
5.4	OMP vs. SC methods under different traffic loads with $r = 0.6$. . .	123
5.5	OMP vs. SC versus add/drop ratio with $\rho = 300$	125
5.6	SC vs. OMP: average number of OEO conversions	126
5.7	Influence of DL scheme on OMP performance with $r = 0.6$	128
5.8	OMP_No_DL compared to OMP and SC schemes	129
5.9	Influence of LPF on OMP performance with $r = 0.6$	130
5.10	OMP_No_LPF is compared to OMP and SC schemes	131
5.11	Influence of reported link costs on OMP performance	132
5.12	The new lightpath fragmentation ratio	133
5.13	Influences of WDM layer routing strategies on OMP performance .	135
6.1	Path-shared vs. segment-shared scheme	144
6.2	LF method adopted for primary path protection.	151
6.3	An example for LF-SSP scheme	153
6.4	Two network topologies utilized for simulations	157
6.5	LF-SSP vs. CL-SSP for port-unlimited NSFnet	159
6.6	LF-SSP vs. CL-SSP for port-limited NSFnet	159
6.7	Average existing channel capacity utilization	161
6.8	LF-SSP vs. SPT and LP for port-unlimited NSFnet	162
6.9	LF-SSP vs. SPT and LP for port-limited NSFnet	163
6.10	LF-SSP vs. CL-SSP for different r	164
6.11	LF-SSP vs. SPT and LP for different r	165
6.12	LF-SSP vs. CL-SSP: influence of average session size	167
6.13	LF-SSP vs. SPT and LP: influence of average session size	168

List of Tables

3.1	Notations used in Chapter 3	53
-----	---------------------------------------	----

Abbreviations

ATM	Asynchronous Transfer Mode
BBR	Bandwidth Blocking Ratio
BGP	Border Gateway Protocol
CL-SSP	connection-level segment shared protection
ELF	Existing Link First
Gbps	Gigabytes per second
GC-OXC	Grooming-Capable OXC
GMPLS	Generalized Multi-Protocol Label Switching
HDTV	High-Definition Television
IETF	Internet Engineering Task Force
ILP	Integer Linear Programming
IP	Internet Protocol
ISP	Integrated Service Provider
KSP	K-loopless Shortest Path
LF-SSP	Lightpath-Fragmentation based Segment Shared Protection
LPF	LightPath Fragmentation
LSR	Label Switched Router
Mbps	Megabits per second
MC-OGSW	Multicast Capable Optical Grooming Switch
MC-OXC	Multicast Capable Optical Cross-Connect
MPH	Minimum-cost Path Heuristic
MPLS	Multi-Protocol Label Switching

MST	Minimum Spanning Tree
MTH	Multiple Tree Heuristic
NNI	Network-Network-Interface
NPF	Nearest Participant First
OADM	Optical Add/Drop Multiplexier
OBS	Optical Burst Switching
OCS	Optical Circuit Switching
OEO	Optical-Electrical-Optical Conversion
OIF	Optical Internetworking Forum
OMP	Overlay Multicast Provisioning
OPS	Optical Packet Switching
OSPF	Open Shortest Path First
OTN	optical transport network
OXC	Optical-cross-connect
PPH	Pruned Prim's Heuristic
PXC	Photonic-Cross-Connect
RWA	Routing and Wavelength Assignment
SMTG	Survivable Multicast Traffic Grooming
SONET	Synchronous Optical Networking
SSP	Segment Shared Protection
Tbps	Tera bit per second
UDP	User Datagram Protocol
UNI	User to Network Interface
WC-OXC	Wavelength-Convertible Optical Cross-Connect
WCC	Wavelength Continuity Constraint
WDM	Wavelength Division Multiplexing
WGXC	Wavelength-Grooming Cross-Connect

Chapter 1

Introduction

1.1 Background and Motivations

Over the past decades, the computer industry and telecommunications industry have experienced phenomenal growth, driven mainly by the strong market demand for computing power and network connectivity. Since late 1990's, the increase in the number of Internet hosts and users caused global network traffic to grow exponentially, which can be seen clearly from the worldwide Internet traffic growth shown in Figure 1.1 [1]. Statistics in [2] show that the Internet traffic almost double each year since 1997, and this growth is expected to continue into the foreseeable future.

In more recent years, with the inception of some new bandwidth-intensive network applications, such as on-line gaming, live-video-distribution and Internet TV, etc. as shown in Figure 1.2 [3], the Internet traffic growth is being further fuelled. Such tremendous growth in both worldwide Internet traffic and the various types of network services is exerting pressure on network operators to expand the capacity of their backbone networks rapidly to meet the demands.

FIGURE 1.1: Worldwide Internet traffic growth in the past decades.

FIGURE 1.2: Internet traffic growth of various bandwidth intensive applications in recent years.

In order to meet the exponentially growing Internet traffic, both Internet service providers and transport service providers urgently in need of a scalable networking technology to provide the desired capacity. Fortunately, as research in optical networking technology matures, the deployment of optical networking technologies comes into rescue [4, 5]. Among all the newly introduced optical technologies, wavelength division multiplexing (WDM) [6] based technology has been successfully deployed to cope with the traffic growth. With the WDM technology, the

achievable bandwidth of a single optical fiber could reach 1.6 Terabit/s (Tbps) in commercial networks [4]; while in the laboratory, recent research activity in designing 100 Tbps multicast-capable optical packet switching (OPS) router [7] as well as the 255 Tbps spatial division multiplexing based multicore fiber [8] further demonstrates the potential of WDM technology in meeting the ever-growing bandwidth demand in the future.

The initial deployment of WDM is limited to point-to-point transmission to increase the link capacity. Such is also known as first generation optical networking. With the rapid advancement and evolution of optical network technologies, optical networking has moved forward from point-to-point WDM transmission to all-optical reconfigurable backbone networks to take full advantage of the available bandwidth. Such reconfigurable WDM networks are known as the second-generation optical networks [9]. Through utilizing the Optical-cross-connects (OXCs) as well as the control plane technologies, such as Multi-Protocol Label Switching (MPLS) [10] and Generalized MPLS (GMPLS) [11, 12], this second generation optical networks support both format and protocol transparency in the control plane while simplifying hardware requirements in the data plane. It is also believed that the second generation WDM networks will be adopted by both integrated service providers (ISPs) and transport network service providers [13]. As research and development of new enabling technologies continue, some new networking paradigms with much finer granularities, such as optical packet switching (OPS) [14, 15] or optical burst switching (OBS) [16, 17], are predicted to be the key features of third generation optical networks.

Today's core network architecture consists of four layers: Internet Protocol(IP) and other content-bearing traffic, over Asynchronous Transfer Mode (ATM) network for traffic-engineering, over Synchronous Optical Networking (SONET) for transport, and over WDM for bandwidth capacity; hence, there are some functional overlap among these layers, and any bottleneck in a particular layer will

render the network unscalable to meet the ever-growing data traffic. In recent years, while IP has emerged as the primary network layer technology, advances in optical component technology, such as optical amplifiers, OXCs and Optical Add/Drop Multiplexiers (OADMs), are now transforming the optical layer from a static transmission facility into a dynamically reconfigurable network. The phenomenal growth in Internet traffic coupled with the enormous capacity that can be provided by WDM technology has bred the idea of an optical Internet [18], and a simplified, two-layered architecture that can leverage on the advantages of both IP and WDM networks is widely viewed as the right solution [18–20].

In this simplified two-layered optical Internet, the unnecessary intermediate layers are eliminated with “data directly over optics”. The WDM layer optical transport network (OTN) consists of multiple OXCs interconnected via WDM links in a mesh topology, while the IP layer networks consist of IP/MPLS label switched routers (LSRs) interconnected by virtual logical links, which are dynamically supported by the WDM layer wavelength switched optical channels that could span multiple OXCs. IP-based MPLS and GMPLS have been proposed as the control plane technology for IP and optical layer networks.

As for the interconnection models between the two layers’ control planes, several industrial organizations, including *Internet Engineering Task Force* (IETF) and the *Optical Internetworking Forum* (OIF), have already proposed some alternative architectural options [20, 21]. Detailed explanations of such architectures will be presented in the next Chapter. Among these proposed models, the simplest one is to treat the optical layer and the IP layer independently from each other. In such “overlay” network model, the optical layer only provides point-to-point *lightpath* [22] connection services to the IP layer networks, while the IP layer client routers can only make bandwidth connection requests to the optical network via a well defined signaling and routing interface, which is referred to as *User to Network Interface* (UNI) [23]. The two layers are controlled and managed independently,

and the only information exchanged between them is for service signaling and responses. Due to its simplicity, the overlay model is accepted as the most practical network architecture for near term implementation.

A critical issue for realizing an overlay IP over WDM interconnection model for next generation Internet is how to accommodate the dynamic incoming requests with efficient use of both IP and WDM layer network resources. Current research focuses on the use of MPLS switching scheme as a framework to provide real-time provisioning service and treat the two layers' resources separately, which thus makes inefficient resource utilizations of the WDM networks. To make efficient use of the networks resources with only limited information exchanges between the two layers is extremely important for overlay IP over WDM networks to be widely deployed.

In optical Internet, although each wavelength channel can provide up to 100 *giga-bytes per second* (Gbps) data transmission rate with the current optical networking technology [24], bandwidth demand of a single network application typically ranges from several to tens of *megabits per second* (Mbps). For the purpose of more efficient wavelength channel capacity utilization, a number of low bit-rate user requests are usually groomed together onto the high-speed wavelength channels for transmission. Such traffic bundling process is referred to as *traffic grooming* [25–27]. Traffic grooming not only improves wavelength utilizations, but also improves the utilizations of some other network resources, such as the transceivers and multiplexers in the network.

Many variations of the traffic grooming problem have been extensively investigated, and a number of provisioning mechanisms have been reported in the past years [25, 27]. Typically, such traffic grooming problems can be classified into two categories: the static ones and the dynamic ones. For the static problem, traffic demands are often given in advance or known to be static over time, and therefore,

the main objective of traffic grooming is to support all the traffic demands with minimum network cost. While for the dynamic traffic grooming problem, traffic demands usually arrive and leave the network dynamically without any information about the future arrivals. In such a case, traffic grooming then becomes a policy designing problem, and the main objective is usually to minimize the overall request blocking probability, or maximize the overall network throughput.

In recent years, with the wide deployment of more and more agile optical components and emerging networking technologies in modern optical networks, the traffic demands are now showing their dynamic nature, and it is expected that future optical networks should be able to respond dynamically to the traffic demands. Therefore, to efficiently utilize the network capacity, scalable and efficient online algorithms have to be devised. Furthermore, as network architectures move forward from ring based to mesh based, the future optical networks will be much more flexible in its configuration, and would be able to support both unicast and multicast traffic. In view of these important issues, we are motivated to investigate the dynamic traffic provisioning problem in future optical networks so that the unicast traffic, or multicast traffic, or both can be supported efficiently.

Optical networks are also prone to network node or fiber cut failures [28, 29]. In modern optical networks, since each wavelength channel can support up to 100G-bits per second of data, a single network failure can interrupt extensive number of on-going applications, and thereby, causes a huge amount of data and revenue losses to both network end-users and network operators. Therefore, network survivability becomes a critical issue in modern optical network design and operation process. Various restoration and protection based fault-recovery mechanisms have been proposed in literature [30, 31]. The initial work on network survivability is mainly focused on either wavelength/sub-wavelength granularity unicast request protection [32, 33] or wavelength level granularity multicast request protection [34–36]. In recent years, with the increasing popularity of some sub-wavelength

granulated bandwidth intensive multicast applications, such as on-line gaming, live-video-distribution and Internet TV, etc. as shown in Figure 1.2, protecting multicast requests in the traffic grooming process is becoming increasingly important. To the best of our knowledge, few schemes have been proposed for such problems, especially for the case with dynamic arriving and leaving multicast requests. Thus, we are motivated to investigate sub-wavelength granularity multicast request protection problem in the dynamic traffic grooming process.

1.2 Objectives

As discussed, the future optical networks will definitely support either unicast traffic, multicast traffic, or both. Specifically, since the network traffic is dynamic in nature while the optical networking technologies are becoming more and more sophisticated, it is widely believed that future optical networks will be able to respond to the user requests rapidly and allocate the required resource upon request dynamically. In a dynamic network environment, the request provisioning algorithms have an important impact on the network performances, and typically the main objective of the algorithm design is to minimize the network blocking performances. This dissertation aims to study the dynamic traffic routing/grooming problem in future optical networks. The main objectives are as follows:

- 1) *To investigate the dynamic unicast request provisioning problem and devise novel schemes to improve network blocking performances for overlay IP/MPLS over WDM networks* - Overlay IP/MPLS over WDM network is widely accepted to be the most practical network architecture for near-term deployment. In such a network, the main problem that limits its performance is the limited information exchanged between its two layers. Even though a number of

algorithms have been proposed for it, few methods have addressed such information exchange problem, and thus efforts are still needed to better understand such challenge. We will study such a problem in this dissertation, and propose novel provisioning mechanisms for the dynamic request, and also improve the network blocking performances..

- 2) *To compare the performance of lightpath and light-tree schemes in supporting sub-wavelength granularity multicast requests in WDM optical networks, and further improve its performance for supporting dynamic multicast traffic* -

Multicasting has become an essential capability of modern optical networks, and various lightpath and light-tree based schemes have been proposed for it. For sub-wavelength granularity multicast requests, however, no results have shown which scheme is a better choice in their dynamic grooming process. In order to facilitate multicast traffic provisioning while improving the network resource utilizations, we will investigate which scheme is a better choice for such a problem, and then propose new ways to further improve the network blocking performances.

- 3) *To study dynamic multicast request provisioning problem in overlay IP/MPLS over WDM network and to devise algorithms to improve the network performances* - Dynamic multicast request provisioning problem in overlay IP/MPLS over WDM networks is important but challenging for the limited information exchange issue that exists for the overlay networks, and it has not received much attention in the past years. Compared to the one to one unicast scenario, multicasting is more complex and deserves further study. Following our preliminary study on the unicasting problem, we extend our focus to the multicasting case, and also evaluate other problems that are related to such multicasting problems. We study the problem of how to improve resource sharing in the traffic grooming process.

- 4) *To study sub-wavelength granularity multicast requests survivability problem and propose new schemes to protect multicast requests against any single link failure in the dynamic traffic grooming process* - A single network failure may cause tremendous data loss to both network users and network operators, and it is especially true for networks supporting sub-wavelength granularity multicast requests. Since link failure is the most common problem in modern networks, it is worthwhile to protect multicast requests in the provisioning process. To the best of our knowledge, few mechanisms have been proposed addressing such a problem. Hence, we investigate how to protect sub-wavelength granularity multicast requests against any single link failure in the traffic grooming process; we also try to improve network resource sharing in both traffic grooming and request protection process so that the total resources allocated for a request protection are minimized.

1.3 Original Contributions

The main contributions of this thesis can be summarized as follows:

- 1) We compared various existing dynamic LSP routing schemes for overlay IP/MPLS over WDM networks, and proposed a novel historical data learning based routing and resource allocation scheme to facilitate IP/MPLS layer unicast request routing process - We conducted comprehensive literature review on dynamic LSP routing algorithms in overlay IP/MPLS over WDM networks, and found that all existing algorithms have ignored the fact that the IP/MPLS layer network may keep a record of the lightpath services that had been supported by the WDM layer network. Such records can help facilitate request provisioning in their routing process. To efficiently utilize such record,

we proposed to utilize the historical data record for logical link cost estimation. Through utilizing a novel historical data learning scheme, a dynamic LSP routing method, named Existing Link First (ELF) algorithm, is proposed. Extensive simulations are performed to compare the ELF algorithm with the other existing algorithms. Results show that the ELF algorithm significantly outperforms the existing ones under different traffic loads, with either limited or unlimited numbers of optical ports. We have also studied the influences of the number of candidate routes, add/drop ratio and the amount of historical data on network performances (Chapter 3).

- 2) We compared the blocking performances of various lightpath and light-tree based methods for dynamic multicast traffic grooming and propose a new lightpath fragmentation (LPF) based method to further improve the network blocking performance - Although both lightpath and light-tree schemes have been proposed for dynamic multicast traffic grooming, no previous work has compared their performances in minimizing network blocking probability. We compare these two schemes systematically, and find that the lightpath scheme has outperformed the light-tree scheme. Inspired by such comparison results, we propose a LightPath Fragmentation (LPF) based algorithm for dynamic multicast traffic grooming. We carry out extensive simulations to compare LPF to the other methods, and results show that LPF steadily outperforms all the other existing schemes with moderate cost of longer delay (Chapter 4).
- 3) We study the dynamic sub-wavelength granularity multicast traffic grooming problem in overlay IP/MPLS over WDM networks and propose an efficient overlay multicast provisioning (OMP) mechanism for such a problem - To facilitate request provisioning, the OMP mechanism jointly utilizes a historical data learning (DL) scheme at the IP/MPLS layer for logical link cost estimation, and a lightpath fragmentation (LPF) based scheme at the WDM layer for improving resource sharing in traffic grooming process. We conduct extensive

simulations to compare OMP to some existing algorithms in different cases; we also evaluate the respective influence of the DL and the LPF schemes on OMP performances through comparing OMP to some other mechanisms utilizing either the DL or the LPF scheme. The influences of different service cost definitions and different WDM layer routing strategies on OMP performances are also evaluated (Chapter 5).

- 4) We investigate the dynamic survivable multicast traffic grooming (SMTG) problem in meshed WDM optical networks and propose efficient algorithms to protect the sub-wavelength granularity multicast requests at the connection level in their dynamic multicast requests in traffic grooming process - Dynamic SMTG problem has not received much research attention in the past years, and few algorithms have been proposed. We study this problem in this dissertation and propose some efficient mechanisms. We first propose the connection-level segment shared protection (CL-SSP) method for such problem, and later improve it by incorporating a lightpath-fragmentation (LF) method to request protection process. We term this improved method as LF-SSP. LF-SSP tries to improve resource sharing in both traffic grooming and protection process while minimizing the total amount of resources allocated for a lightpath protection. We compare LF-SSP to some other existing methods either for subwavelength-level request protection or for wavelength-level protection in different cases, and evaluate the influences of some other factors on network performances (Chapter 6).

1.4 Thesis Organization

The remainder of this dissertation is organized as follows:

Chapter 2 briefly reviews some previous work that is related to this dissertation. We first introduce some emerging optical networking technologies for next-generation optical Internet and then introduce the traffic grooming problem, with a specific focus on dynamic traffic grooming; finally, we present the relevant work on multicast request protection problem.

Chapter 3 presents a historical data learning based dynamic LSP routing algorithm named existing link first (ELF) for overlay IP/MPLS over WDM networks. Simulation results with two different network topologies are also presented to study the performances of the ELF algorithm.

Chapter 4 systematically compares the blocking performances of lightpath and light-tree schemes in supporting dynamic multicast traffic grooming. A novel light-path fragmentation (LPF) based dynamic multicast traffic grooming algorithm is also proposed to improve the network blocking performances. Extensive simulations are conducted to compare the LPF method to the other algorithms in different cases.

Chapter 5 proposes an overlay multicast provisioning (OMP) mechanism for dynamic multicast traffic grooming algorithm in overlay IP/MPLS over WDM networks. Influences of the IP/MPLS layer data learning (DL) scheme, and the WDM layer lightpath-fragmentation (LPF) methods on OMP performance, as well as the WDM layer routing strategy are evaluated via extensive simulations.

Chapter 6 presents an algorithm, namely LF-SSP, for dynamic survivable multicast traffic grooming (SMTG). Performance of such method is compared to that of an existing method, namely, CL-SSP, for subwavelength-level multicast request protection in different cases via simulations. We also compare LF-SSP to some existing algorithms for wavelength-level multicast request protection in different cases. Influences of some other factors are also evaluated.

Chapter 7 summarizes the whole dissertation and discusses some possible future research directions.

Chapter 2

Literature Review

2.1 Introduction

This chapter gives a literature survey on recent advances in some areas that are related to this dissertation. It consists of four parts. The first part presents some emerging technologies for next-generation optical Internet, with a specific focus on the technologies that are largely related to research reported in the next few chapters of this dissertation. The second part describes the traffic grooming problem. We first introduce what the traffic grooming problem is about, and present some previous work on both static and dynamic traffic grooming problems. We explain why we focus on dynamic traffic grooming; and then summarize the previous efforts on dynamic LSP routing in overlay IP/MPLS network and dynamic multicast traffic grooming in WDM networks, respectively. The third part is on multicast request protection. We describe some existing algorithms on wavelength-level granulated multicast request protection, and then present two existing algorithms for sub-wavelength granulated multicast request protection. We also point out the problem with the assumption made in these two algorithms' using an example. The last section summarizes the chapter.

2.2 Emerging Technologies for Next-Generation Optical Internet

2.2.1 Wavelength Routed WDM Networks

With the explosive increase in Internet traffic loads over the past few decades, wavelength routed WDM networks have emerged to be the dominant technology for backbone networks [9, 37]. A wavelength routed WDM network is comprised of a number of optical cross-connects (OXC) that are interconnected by some optical fiber links in a meshed topology; an optical link in the network is usually composed of two fibers in opposite directions with each carrying a number of wavelengths. Figure 2.1 illustrates a typical wavelength routed WDM network. In such a network, the transmission channels, referred to as lightpaths [22], are set up and torn down dynamically to provide the desired connectivity services to its client sub-networks. As can be seen in Figure 2.1, two lightpaths utilizing wavelengths λ_1 and λ_2 are set up. OXCs in the WDM network are utilized as wavelength routers to provide the wavelength level data routing and connection between different client sub-networks; the operations of the OXCs are controlled and managed by the OXC control units of the WDM network.

FIGURE 2.1: A typical wavelength routed WDM network architecture

A typical OXC architecture with N fiber links with each carrying W wavelengths is shown in Figure 2.2. As can be seen, the wavelengths within an optical fiber link are de-multiplexed first, and then sent to the corresponding switch fabrics; finally, they are switched to the desired output port and multiplexed again into a fiber link. For an OXC, it is equipped with some add and drop ports, which are utilized to transmit or terminate the traffic from the higher layer clients. Normally an add port comprise an electronic-to-optical converter, which is the transmitter or source of a wavelength channel while a drop port consists of an optical-to-electronic converter, which is the receiver or sink of a wavelength channel. An OXC could also be equipped with some wavelength converters so that it can switch the optical signal on an incoming wavelength of an input fiber to some other wavelengths on the output fiber. An OXC with wavelength converter is also known as wavelength-convertible optical cross-connect (WC-OXC). Wavelength converters are usually quite costly as they need to be implemented using expensive lasers and other optical devices.

FIGURE 2.2: A typical $N \times N$ optical-cross-connect structure.

In the network as shown in Figure 2.1, the client sub-networks are connected to each other via wavelength channels; in the IP/MPLS client sub-networks, the edge routers that are directly connected to the OXCs usually act as sources or destinations of the wavelength channels, and they implement the User-to-Network Interface (UNI) [23] for the WDM network and the client sub-networks. In the network shown in Figure 2.1, the wavelength channels may initiate or terminate at the edge routers, but the communication paths within the IP/MPLS sub-networks may continue in the electrical domain. The electrical domain communication channels are controlled and managed by the IP/MPLS protocol. With the convergence of IP-centric data network and the WDM optical transport network, it is widely believed that the IP/MPLS protocol will be the dominant protocol for the next-generation optical Internet.

2.2.2 IP/MPLS over WDM Convergence

The explosive growth of Internet traffic as well as the rapid advances in WDM networking technologies have made optical Internet desirable and possible [2, 18]. In addition, since multi-protocol label switching (MPLS) [10] and generalized MPLS (GMPLS) [11, 12] are converging as the new forwarding technology, a simplified two layered network architecture with IP/MPLS directly over WDM network is widely accepted as the most promising solution for next generation optical networks [19, 20]. In this two-layered architecture, the high-speed electronic routers equipped with MPLS functions are named label switched routers (LSRs) while those end-to-end paths that interconnect a pair of LSRs are called label switched paths (LSPs). An LSP may traverse more than one lightpaths; all LSPs together with the LSRs constitute the IP/MPLS layer logical topology of the IP/MPLS over WDM networks while the WDM layer of an IP/MPLS over WDM network consists of the OXCs interconnected by optical fiber links, each of which carries a number of wavelengths in a mesh topology.

As research in WDM networks and optical components advances, the development of IP/MPLS over WDM networking can be broadly classified into three generations [38]:

First generation IP/MPLS over point-to-point WDM networks: In such first generation networks, the WDM network at the physical layer is mainly utilized to provide point-to-point high-speed transmission connections between two adjacent routers. The Synchronous Optical Network (SONET) is typically used for framing and transport of the header information on WDM transmission channels, and IP packets are packed into SONET frames via the packet-over-SONET strategy. This model belongs to the IP over SONET over WDM three layered network architecture, and in such network, the IP routers are connected one another directly via multi-wavelength optical fiber links. Network topology under this architecture is fixed, network configuration is static and interaction between the IP and WDM layers are limited.

Second generation IP/MPLS over reconfigurable WDM networks: In the second generation IP/MPLS over WDM networks, the wavelength-level transmission channels are set up and torn down in the wavelength-routed WDM networks using the OXCs. Due to the configurability of the optical layer networks, protection and restoration can be implemented directly in the WDM layer network, hence, the SONET layer can be eliminated and a two-layered IP/MPLS over WDM architecture is achieved. In this network architecture, IP routers are connected to the OXCs ports through router interfaces, and the OXCs are connected to each other in a mesh topology via WDM fiber links. Through configuring OXCs appropriately, a router interface can be connected to any interface of another router. For the connections between the control planes of the IP layer and the WDM layer networks, several interconnection architectures have been proposed, namely overlay model, augmented model and peer model [21], which will be explained in detail later.

Third generation IP/MPLS over OPS/OBS Switched WDM Networks:

In this third generation network, data traffic is transported and switched in much finer granularity, e.g., optical packet switching (OPS) [14, 15] or optical burst switching (OBS) [16, 17] instead of optical circuit switching (OCS) [39], in optical switches in the WDM layer networks. Without the huge amount of OEO conversions, much faster data processing speed as well as much more efficient network resource utilization can be achieved. However, since there is no cost effective optical buffering solution as well as difficulties in implementing optical logical gates, this architecture is too futuristic and not realizable in the near future.

In our work, we focus on the IP/MPLS over reconfigurable WDM network, which seems to be most practical and realizable in the near future [38]. Specifically, for such networks, the IP/MPLS layer and the WDM layer network control planes can be combined in different fashions, namely overlay model, augmented model and peer model [20, 21]:

The Overlay Model: A typical overlay IP/MPLS over WDM network is a hierarchical interconnection model consisting of two layers, with the IP/MPLS layer residing over the WDM layer in a client-server fashion. In such networks, the IP/MPLS layer and the WDM layer networks are independent of each other, with IP/MPLS layer being the integrated service provider (ISP) and WDM layer being the bandwidth provider. The IP network has no knowledge of the optical network topology and resources.

With the overlay interconnection model, while the IP/MPLS layer network can only request lightpath services from the WDM layer, the optical network is responsible for setting up lightpaths between each IP/MPLS layer router pair. The only information exchanges between the two layers are the signaling messages for creating, modifying and deleting a switched lightpath through the UNI. Figure 2.3 shows a sample overlay IP/MPLS-over-WDM network architecture.

FIGURE 2.3: A typical overlay IP/MPLS over WDM network.

In overlay IP/MPLS over WDM networks, each network node is a photonic-cross-connect (PXC)¹ interconnected with zero, one or more LSRs through the UNI. When there is no LSR connected to it, a PXC is only responsible for switching the bypass traffic, from its input ports to its output ports transparently at the wavelength granularity [40, 41]. When a PXC is connected with one or more LSRs, it has additional functions. Figure 2.4 illustrates a typical PXC architecture with some LSRs connected to it. With several LSRs connected to it, the PXC can receive traffic terminated at the LSR to the local network or transmit traffic originated at the LSR from the local network. The LSRs are used to multiplex local traffic streams into a higher capacity request that PXC can support and also to generate/terminate traffic to/from a lightpath to local receiver.

The advantage of this model is that it is relatively simple to implement for static IP/MPLS over WDM networks, and it is particularly attractive to carriers, as

¹In this dissertation, we use PXC and OXC interchangeably as PXC is more frequently utilized in industry to indicate that an OXC is associated with some digital OXC functionalities.

FIGURE 2.4: A typical node architecture for overlay IP/MPLS over WDM networks. A PXC with multiple LSRs connected to it, where each LSR connects to the PXC through a fixed number of add/drop ports.

it aligns with common carrier practice of organizing their operational units into transport and switching units.

The Peer Model: In peer model networks, the IP router or LSRs and WDM layer OXCs are interconnected in a peer-to-peer fashion. The two layers are collapsed into a single integrated layer and are managed in a unified manner. In other words, the OXCs are treated in the same way as the routers or LSRs. With the adopted GMPLS function, all the IP routers and OXCs are controlled and managed by a single control plane, and each OXC is equipped with an IP address. In such a cases, the UNI, network-network-interface (NNI), and any other router-to-router interface are actually similar to each other from a routing and signaling perspective. In addition, the IP and optical networks run the same suite of routing protocol, e.g., open shortest path first (OSPF) [42] with suitable optical extension. While the topology and link state information kept by all the IP routers or LSRs and OXCs are identical to each other. Figure 2.5 illustrates a typical peer model IP over WDM network.

The Augmented Model: The augmented model allows the IP and optical networks to use separate routing protocols, while a certain kind of selected information from one routing protocol should be passed through the other. In other words, this model aims at making a compromise between the peer model and overlay model by allowing certain routing information to be exchanged between the two layers. For example, the reachable external IP addresses possessed by IP layer network could be passed through to the optical routing protocol, so that the reachability information could be accessed by its IP clients. For the information exchange between the IP domain and optical domain networks, the border gateway protocol (BGP) could be adopted in such network model. Figure 2.6 shows an augmented

FIGURE 2.5: A typical peer model IP/MPLS over WDM architecture.

network model for IP/MPLS over WDM networks.

The augmented network model tries to make the best of both IP and optical networks, and it also saves the complexity of the peer model networks. Yet the key design issue for this model is how routing information is exchanged at the IP-optical UNI. Some researchers suggested using BGP or OSPF for routing information exchanges between the two layers [43]; however, it remains to be seen if such a collapsed structure is feasible as there are distinct differences between the two layers. So far, no consensus has been reached on what kind of information should be shared between the two layers.

Among the three proposed interconnection models, the overlay interconnection model is the simplest in term of network control and management system implementation. It is widely believed that the overlay IP/MPLS over WDM network architecture will be a key component of next-generation optical Internet [19, 20, 38].

FIGURE 2.6: A typical augmented model for IP/MPLS over WDM networks.

2.2.3 Lightpath

In wavelength routed WDM networks, data is transmitted between different end node pairs by wavelength channels, which are known as lightpaths [22]. A lightpath is an end-to-end all-optical wavelength channel that connects the ingress and egress nodes of the network without experiencing any Optical-Electrical-Optical (OEO) conversions. Although an end-node pair that is connected by a lightpath may be geographically a number of links away from each other, they are only one-hop away from each other in the logical layer network. A lightpath provides the bandwidth capacity at the wavelength level granularity, and it can support a large number of sub-wavelength level requests simultaneously. Upon the client requests, lightpaths are set up and torn down dynamically. In the WDM networks, the set of lightpaths form the logical layer network. For example, in Figure 2.3, the WDM layer network is formed by four OXCs interconnected by four fiber links. In the WDM layer network, four lightpaths are constructed with each traversing a number of fiber links, while in the IP/MPLS layer network, these four lightpaths appear as four directional logical links.

For a given WDM network topology, which is usually represented by a graph $G(V, E)$ with some network resource utilization information, a typical network routing algorithm can be adopted for request provisioning. For example, for a unicast connection request from s to d , the Dijkstra's Shortest Path algorithm [44] can be used to find the minimum cost lightpath route to serve the connect request between s and d .

Compared to electrical switching, the biggest advantage of lightpath switching is its data transparency, which refers to the fact that data transmission is independent of the data bit rates, frame formats, protocol, etc.; the lightpath also decouples the optical layer network from its upper layer IP/MPLS network, and thus, makes network maintenance and updating easier.

In wavelength routed WDM networks, OXCs provide the routing and switching functions for lightpath based communication. If all OXCs in the network are not equipped with any wavelength converters, i.e., the OXCs are not capable of wavelength conversion, a new lightpath has to be associated with the same wavelength along its route, which is known as the wavelength continuity constraint (WCC); while if any of the OXCs are WC-OXC, such constraint can be relaxed.

2.2.4 Light-tree for Multicasting

Multicast is an efficient way to transmit information from one source to multiple destinations simultaneously. A multicast source sends a single copy of the data to all the members, and the data copy is needed only when at the branch of a tree, of which the source node is the root node and destinations are leaf nodes. Multicast saves a large amount of bandwidth resource when compared to the case of sending a separate copy of the data to each individual group member. Since IP has become the dominant network layer forwarding protocol while the demand for multicast service is growing steadily at the same time, extensive effort has been spent to design efficient IP multicast schemes [45].

IP multicast is based on an open service model. Any host or user can create a multicast group, send data to or receive data from a group and the multicast source only needs to know the group address. IP multicast utilizes the best-effort User Datagram Protocol (UDP) and it supports dynamic group management. Multicast members can dynamically join or leave a multicast group and there is no need for them to register, synchronize or negotiate with the center controller. Moreover, the IP multicasting does not require any special mechanisms for providing quality of service, security, or address allocation. The existing research on IP multicast has primarily focused on multicast group management, IP multicast protocol design,

FIGURE 2.7: Different multicast schemes in WDM networks: (a) Pure IP layer multicast; (b) IP multicast via multiple unicast; (c) WDM layer light-tree.

and IP multicast routing algorithm design. Detailed study on IP multicast can be found in [43, 45].

In recent years, as more and more bandwidth intensive multicast applications, such as multi-player gaming, video conferencing and interactive distance learning, etc., are becoming increasingly popular, a generalized lightpath concept, named light-tree, was introduced in [46] to support multicasting in WDM optical networks. A light-tree is an all-optical channel that supports data transmission transparently from a single source to multiple destinations simultaneously. Compared to IP multicast, light-tree typically relies on the optical splitting capability of its branching nodes for traffic copying. Compared to the one-to-one lightpath scheme, it is more bandwidth efficient in supporting wavelength level granulated multicast requests.

The efficiency of light-tree can be demonstrated with the example shown in Figure 2.7. Assume that a wavelength level granulated multicast request $R\{s, \{d_1, d_2\}\}$, where s and $\{d_1, d_2\}$ are the request source and destination set, respectively, arrives at the network. Figure 2.7 shows three typical schemes to support such a request. The first way is to use pure IP layer multicasting as shown in Figure 2.7(a). Each IP router on the multicast tree replicates the data packets, and then transmits the copy to down stream routers until the packets reach all the multicast destinations.

Apparently extensive OEO conversions are required at every router when using this way of multicasting. The huge amount of OEO conversions imposes heavy burden on the IP routers, such as the high-speed packet header processing units, the high power consumption and cooling equipment requirement, etc.; a long data processing delay is also introduced in the transmission process.

The second way is to utilize multiple one-to-one lightpath unicast as shown in Figure 2.7(b). Each multicast destination is connected to the multicast source via a lightpath in an end-to-end fashion. As can be seen, to support such a request, two separate lightpaths, which are s to d_1 using wavelength λ_1 and s to d_2 using λ_2 , are set up. Obviously, excessive bandwidth will be consumed. However, when the multicast destination size is large, bandwidth consumption will be extensive, and sometimes some requests will be blocked, which may further increase the blocking probability.

Finally, a light-tree strategy is utilized to realize optical multicasting at the WDM layer, taking advantage of the power-splitting capability of WDM switches as shown in Figure 2.7(c). Since the multiple destinations share the bandwidth of the same wavelength, this scheme provides more efficient utilization of bandwidth resource than the second scheme. Specifically, since the light-tree is an end-to-end connection channel from the request source to destinations, it does not require any OEO conversions in the transmission process.

To support light-tree communication, the multicast capable OXCs (MC-OXCs) are usually utilized to set up light-tree or light-forest in the WDM layer. Various kinds of MC-OXCs have been proposed in literature for optical layer multicast.

The first MC-OXC was proposed in [47], which is shown in Figure 2.8(a) and it is based on the so-called SaD switch as shown in Figure 2.8(b). An SaD switch consists of N light splitters, N^2 amplifiers and N^2 2×1 switches. All input signals to the SaD switch are using the same wavelength. Once the wavelength reaches

FIGURE 2.8: The first proposed MC-OXC and the SaD switch: (a) The first MC-OXC; (b) The SaD switch.

the input port, the light splitter splits the signal into N outputs, where the output signals are amplified by the optical amplifiers immediately. Through configuring the N^2 2×1 switches, each of the output signals can be switched to the appropriate set of output ports. In such SaD-based MC-OXC, each input optical signal can be directly switched to none, one, multiple, or all output ports. Therefore, the SaD-based architecture is strictly non-blocking.

Using this non-blocking SaD switch, the $N \times N$ MC-OXC can be built as shown in Figure 2.8(a). At each input port, a de-multiplexer is used to extract individual wavelengths, which are then directed to the appropriate SaD switch. The SaD switch performs the space switching and/or signal splitting for the wavelengths directed to it and also sends the wavelength to the desired output ports. Finally, the multiplexers at each output port are employed to combine the W signals on the individual wavelengths for transmission onto an outgoing fiber. This MC-OXC design can realize any permutation of optical multicast or unicast connections in a strictly non-blocking manner.

To reduce the power loss and the cost of SaD switch based MC-OXCs, some other power efficient and cost-effective switch architectures have been proposed

[48–51]. Largely speaking, the issues with designing cost-effective and power-efficient MC-OXC are related to the data-plane issue [52]. To conduct multicast communication, efficient light-tree set up schemes are also needed, which is related to the control plane issues.

The network layer multicasting problem is the well-known Steiner tree problem. Given a network topology $G(V, E)$, and a multicast request $R\{s, D\}$, where V , E , s and D are the network nodes, network edges, request source and request destinations, respectively, the problem is to find a minimum cost light-tree to connect s to all nodes in D . The Steiner tree problem is known as *NP*-complete [53] and is not scalable for networks with a large number of nodes. To approach such a problem in large size networks, approximation algorithms are usually adopted, and various heuristic algorithms have been proposed and reported in literature [54–56]. The following are three algorithms that are typically used for such a Steiner tree problem:

- a) Pruned Prim’s Heuristic (PPH) [54]: PPH begins by using the Prim’s algorithm [57] to find a minimum spanning tree (MST). When such minimum spanning tree is obtained, the PPH method then starts to prune the unnecessary tree branches until the final tree can reach only the desired destination nodes in D . The overall time complexity of PPH is $O(|V|^2 + |D||V|)$.
- b) KMB’s algorithm [55]: KMB’s algorithm is also based on the MST tree. The algorithm builds a complete directed distance graph H from G , and finds a MST tree U on H first; it then replaces each edge of U with the corresponding shortest path on G and again gets a MST tree U_s ; it finally prunes the leaves that are not in D from U_s to get the Steiner tree. The overall time complexity of PPH is $O(|D||V|^2)$.
- c) Nearest Participant First (NPF) algorithm [56]: NPF is a greedy-based algorithm. It starts from the node s , and tries to connect destination nodes in

D one by one to the partially built tree using the shortest paths. As NPF always connects the destination nodes using the minimum-cost path, it is also named the Minimum-cost Path Heuristic (MPH) algorithm. The overall time complexity of NPF is $O(|D||V|^2)$.

In our work within this dissertation, we will use the NPF algorithm to find the required tree for each multicast request.

2.3 Traffic Grooming

With the WDM technology, an optical fiber is able to support tens of wavelength channels with each providing capacity up to 100Gbps with the current transmission technology [58]. On the contrary, the bandwidth requirement of an individual user connection request is normally in the order of megabits per second (Mbps), which is much lower than the capacity provided by a single wavelength channel. Therefore, if a whole lightpath/light-tree is allocated to carry a sub-wavelength granulated user request, a large part of the bandwidth of such lightpath/light-tree will be wasted. Therefore, in order to efficiently utilize the wavelength channel capacity, multiple low-rate traffic demands are usually packed together onto the high-speed channels for transmission. Such a traffic bundling process is referred to as *traffic grooming* problem [59–61]. Initially the traffic grooming problem is studied in SONET/WDM ring networks, and the main objective is to minimize the overall network cost in terms of the number of optical add/drop multiplexers (OADMs). Later, such a problem is extended to meshed WDM networks to handle the static traffic, wherein the traffic grooming problem is also referred to as virtual topology design problem in some literature [25, 62].

To support traffic grooming, two typical kinds of switch architectures as shown in Figure 2.9 have been proposed in [63]. Figure 2.9(a) is a hierarchical switch

FIGURE 2.9: Node architectures for traffic grooming: (a) GC-OXC for lightpath based schemes; (b) MC-OGSW for light-tree based schemes.

architecture equipped with both wavelength switching and traffic grooming fabrics. With the switching fabrics, such a switch is able to support wavelength level data switching in the optical domain, while with the grooming fabric, it is able to realize the data grooming and switching in the electronic domain. In addition, such architecture is also able to realize traffic duplication with its grooming fabric in the electronic domain. This kind of switch architecture is called as grooming-capable OXC (GC-OXC) or Wavelength-Grooming Cross-connect (WGXC) [64], and is usually utilized to support the lightpath based unicast or multicast traffic grooming.

Figure 2.9(b) presents an architecture that is able to support multicasting services with either full wavelength capacity requirement or subwavelength level capacity requirement. For wavelength level granularity request, it is able to receive the optical signal from one input fiber, and send it to its splitter banks for signal duplicating, and finally switch them to the desired output ports all-optically. If the light-splitter banks are sufficient, the multicast requests can be transmitted as much as possible in the optical domain. While for the sub-wavelength granulated requests, the architecture is also equipped with a grooming fabric that is

responsible for traffic termination, aggregation and re-transmission in the electronic domain. Such architecture is usually called multicast capable optical grooming switch (MC-OGSW) [65], and is adopted for light-tree based multicast traffic grooming.

In recent years, an increasing amount of research efforts have been devoted to traffic grooming for both unicast and multicast traffic demands, and it is widely recognized that traffic grooming is a practical and critical technique for WDM network design and implementation [25, 63]. Below we give a brief review of the previous efforts on such problem.

2.3.1 Static and Dynamic Traffic Grooming

The traffic grooming problem can be classified into two categories: the static traffic grooming problem and the dynamic traffic grooming problem. For the static scenario, the traffic requests are known in advance or do not change over a period of time. From the algorithm design point of view, the main objective of this problem is generally to minimize the cost for supporting all the traffic demands, or to support as many request demands as possible with the given network resources. While for the dynamic scenario, the unicast/multicast traffic demands usually arrive and leave the network dynamically without any information about the future arrivals, and the main objective of algorithm design is usually to minimize the blocking probability or to maximize the network throughput.

The early-stage work on traffic grooming was mainly for tackling the static problems [25, 63], and conceptually such static problem can be divided into three sub-problems as below,

- a) The virtual topology design sub-problem: when given the traffic demands and the physical layer network topology, this problem is to determine the set of

lightpaths to be established for the traffic demands. The set of new lightpaths define the virtual network topology;

- b) The routing and wavelength assignment (RWA) sub-problem: after obtaining the virtual topology, it is essential to determine the physical layer route that corresponds to each virtual link and assign wavelength(s) to each route. In the RWA process, the distinct wavelength constraint, i.e., different lightpath passing through the same fiber link should be assigned different wavelength, and wavelength continuity constraint (WCC) [9], i.e., wavelength along a path route should be the same if no wavelength converter is utilized, must be taken into account. Extensive review on this sub-problem can be found in [66];
- c) The routing/grooming demands for the given traffic sub-problem: once the virtual network topology is finalized, the problem is how to meet the traffic demands efficiently using the lightpaths that have been set up. Since a request may transverse a number of existing lightpaths, and the different lightpath connections have to use OEO conversions, the main objective of this sub-problem is to minimize the OEO consumptions.

The static traffic grooming problem is well-known as an NP-complete problem. It is usually modeled as an integer linear programming (ILP) problem, and then is tackled with heuristic algorithms. Various approximation algorithms have been proposed for static unicast/multicast traffic grooming problem in literature. Comprehensive review on static traffic grooming problem can be found in [25, 67].

In recent years, with the developments of optical communication technologies, the traffic in optical networks now tends to show its dynamic nature; also, as the network traffic grooming locations are extending from core networks to metropolitan area networks, dynamic traffic grooming becomes increasingly popular, and extensive work has been done on it [26, 67]. Specifically, the authors in [26] classified the dynamic traffic grooming problems into two broad categories: *analysis*

problem and *design problem* according to different minimization objectives. While the analysis problem focuses on analyzing the network behaviors, e.g., network blocking performance, presuppose an existing policy and resource allocation under given traffic conditions, the design problem presuppose certain model that allows computation of certain objective under specific resource allocation and policy or to devise strategies to optimize certain objectives of interest. Among all issues that are related to both problems in the dynamic traffic grooming process, the network blocking performance is always the main concern, as it is the most important metric that measures the performance of a network.

For all work within this dissertation, we mainly focus on the design problem in dynamic traffic grooming process. Specifically, we aim to devise provisioning strategies to minimize the network blocking performance in different cases. Below we give a brief review to the previous work on dynamic unicast LSP routing in overlay IP/MPLS over WDM networks and multicast traffic grooming in WDM networks.

2.3.2 Dynamic LSP Routing in Overlay IP/MPLS over WDM Networks

In the past few years, dynamic unicast request provisioning problem has received extensive research interests in WDM networks [26, 68–70]. Specifically, the authors in [68, 69] have proposed different kinds of auxiliary graph models to represent the WDM networks; based on such graph models, various unicast request provisioning algorithms have also been presented. For the overlay IP/MPLS over WDM network, however, dynamic traffic grooming problem is relatively new, since the proposed graph models and existing algorithms cannot be directly extended to the overlay networks: in the overlay networks discussed in Section 2.2.2, the two layers are independent of each other, owned by two different owners, making the

complete information on the two different layers unavailable to any user. Due to such layered property of the overlay network, only a few algorithms have been proposed for the dynamic unicast traffic provisioning problem. We classify these existing algorithms into two categories, i.e., sequential routing and resource based routing, according to the routing decision strategies adopted by them.

Sequential routing algorithms: The Logical-Layer-First (LGF) and Optical-Layer-First (OPF) are two representative sequential routing algorithms. In the LGF algorithm, the following two steps are carried out upon the arrival of each transmission request: (1) try to route the request over the residual bandwidth on the existing logical links; (2) if Step 1 fails, then try to set up a new lightpath directly between the ingress and egress LSRs on the WDM-layer network. For the OPF algorithm, the above two steps are reversed. In both algorithms, if both steps fail, the request is blocked.

In [71], Ye *et al.* used the LGF sequential routing algorithm to set up the primary path in their integrated routing/protection strategy. Niu *et al.* presented both the LGF and OPF algorithms and made some comparisons between them when fixed-path routing is applied to set up new lightpaths [72]. Zhong *et al.* improved the OPF algorithm by adopting dynamic least congested shortest path routing in the WDM-layer network [73]. Although both Niu *et al.* and Zhong *et al.* have studied the influence of add/drop ratio on network performance, they did not take any measures to improve the optical port utilization when setting up new lightpaths. Therefore wavelength utilization tends to be low when lightpaths are long.

To improve both the port and wavelength resource utilizations, Ye *et al.* proposed an algorithm called Short Lightpath Establishment Approach (SLEA) [74]. Through dynamically assigning link costs in the auxiliary graph taking into consideration the optical hop constraint, i.e., assigning a high cost once the optical hop length reaches a certain threshold value and a low cost otherwise, SLEA tries

to eliminate the inefficient long lightpaths. Simulation results show that SLEA significantly improves the network blocking performances, and it even outperforms the integrated routing algorithm described in [75] if an appropriate hop constraint is found.

Resource based routing algorithms aim at efficiently utilizing the network resources. Two typical examples of such algorithms are the Existing Capacity First (ECF_OVLY) and Minimum Logical Hop (MLH_OVLY) methods proposed by Koo *et al.* [76]. The main idea of ECF_OVLY is to first try to use the residual bandwidths of the existing logical links to serve the arriving requests. If that fails, it then tries to set up some new lightpaths, not necessarily from source to destination, for the request. By encouraging setting up shorter lightpaths, ECF_OVLY lowers the chance that lightpaths are under-utilized for a long time.

MLH_OVLY aims to minimize the number of logical hops traversed by each incoming request. For each arriving request, MLH_OVLY first tries to serve it using a single-hop logical link, by either using an existing logical link or setting up a new lightpath between the source and destination LSRs; if that fails, it then tries to find a route with the minimum number of logical hops, where new lightpaths are set up when necessary. Simulation results in [76] show that MLH_OVLY outperforms ECF_OVLY.

In our earlier work described in [77], we proposed the use of historical data in LSP routing in overlay IP/MPLS over WDM networks. Through utilizing a novel but simple data learning scheme to estimate the IP/MPLS layer logical link costs and *K-loopless shortest path* (KSP) algorithm to find a number of IP/MPLS layer candidate routes, two simple LSP routing algorithms were proposed and satisfactory network blocking performances were obtained.

2.3.3 Dynamic Multicast Grooming in WDM Networks

In recent years, as more and more multicast applications such as multiparty conferencing, video distribution and HDTV etc. are becoming increasingly popular, multicast traffic is expected to constitute a large portion of the overall network traffic in the future. However, since the bandwidth needed for a multicast session is typically much lower compared to the capacity that can be provided by a single wavelength channel in today's WDM networks, traffic grooming is usually adopted to efficiently utilize the wavelength capacity.

The rapid advances and adoption of optical communication technologies make dynamic multicast traffic grooming increasingly important, and quite a few algorithms have been proposed for such a problem [65, 78–84]. Such algorithms typically utilize either lightpaths or light-trees, or both of them, to support dynamic multicast traffic. For convenience, we term those algorithms utilizing only the lightpaths for multicast traffic grooming as *lightpath based* approaches; and *light-tree based* approaches otherwise.

Lightpath based traffic grooming approaches: The first lightpath based dynamic multicast traffic grooming algorithm, named Maximizing Minimum Freeload (MMFL) method, was proposed in [78]. To simplify the calculations, MMFL limits the route selection for each multicast request to be on a single wavelength which, among all the wavelengths, has the largest overall residual capacity after accommodating the multicast request. The MMFL algorithm is simple, but the single wavelength constraint degrades the network resource utilizations.

To alleviate the unfavorable single-wavelength constraint, two other algorithms, which we called the logical-path-tree (LPT) method and saturated cut (SC) method respectively, were proposed in [79]. Both algorithms adopt the same idea of utilizing existing logical links to serve as many destinations as possible before setting up new lightpaths. Between them, SC achieves better performance by first

finding islands which include at least one of the sources/destinations and other nodes connected to them via existing links with sufficient residual capacities, and then connecting such islands by setting up new lightpaths. To the best of our knowledge, SC performs best among all the existing lightpath based methods.

Light-tree based traffic grooming approaches: For light-tree based dynamic traffic grooming, an established light-tree can either be fixed or changed in the grooming process. Depending on whether an established light-tree can be changed in grooming or not, we classify algorithms of this category into *fixed* light-tree methods and *adaptive* ones.

For the fixed light-tree methods, an established light-tree cannot be changed until the transmission going through it is finished; hence a fixed light-tree can only serve new requests whose destinations are supersets of the destinations of the established light-tree. The first set of four light-tree methods were proposed in [65], where an established light-tree can only be utilized to groom requests with the same destination set (though not necessarily the same source). Since the probability that two multicast sessions have the same destination set is low, the algorithms may lead to low network resource utilizations.

To release the “same destination set” constraint, algorithms are proposed to allow a new multicast session to utilize multiple existing light-trees if, and only if, these trees have disjoint destination sets which are all subsets of the destinations of the new session [80, 81]. Such algorithms include Multicast Tree Decompose (MTD) [80], light-tree division-destination branch node-based grooming scheme (LTD-DBNG) and light-tree division adjacent node component-based grooming scheme (LTD-ANCG) [81] etc. They adopt the same objective of utilizing the existing light-trees to serve as many destinations as possible. Compared to MTD, LTD-DBNG and LTD-ANCG have an additional step: they split a new light-tree into a few smaller ones such that the smaller trees have a better chance to be used

by future multicast sessions. Specifically, while LTD-DBNG splits a light-tree at an intermediate node only if this node is one of the destinations of the multicast session, LTD-ANCG allows more flexible splitting of a light-tree into a number of sub-trees according to a predefined priority list of different sub-tree topologies. Among the existing fixed light-tree methods, LTD-ANCG achieves the best blocking performance; hence it will be adopted in the performance comparisons.

For the adaptive light-tree methods, established light-trees can be changed dynamically for multicast traffic grooming [82, 83], though this may cause interruptions to ongoing traffic. Specifically, by modeling the network into a layered auxiliary graph, the multicast dynamic light-tree grooming algorithm (MDTGA) algorithm allows its established light-trees to be dropped, branched or extended to groom new requests, or to be contracted to release unused resources [82]; the efficient multicast traffic grooming algorithm (EMGA) method adopts a similar mechanism, but it dynamically changes its link costs to improve the overall grooming efficiency [83]. It has been shown in [83] that EMGA achieves better blocking performance than MDTGA in different cases. We include EMGA in performance comparisons in Chapter 4.

An interesting stop and go (S/G) light-tree mechanism was proposed in [84]. Such an S/G light-tree mechanism is based on a hybrid switch architecture, which is capable of both circuit- and packet-switching. Although such a mechanism is able to groom either unicast traffic or multicast traffic, or both, the switch architecture is quite different from what are being considered in current literature; it is also complicated and expensive for implementation.

2.4 Multicast Request Protection

Optical networks are vulnerable to various component failures, and a single failure of either software or a network element, e.g., fiber cut, OXC broken down, port on a client equipment failure, etc., can cause massive information loss and serious service interruptions to the end-users [30]. In networks supporting a large number of bandwidth-intensive multicast applications, such failure is even more serious since a single failure of a multicast session will affect several downstream users. The nearer a failure is to the session source, the more serious is its impact. Therefore, survivability mechanisms are essential to protect multicast sessions against network failures. Survivability methods can be classified into two categories: *restoration* [85] and *protection* [86]. While restoration is reactive and efficient in terms of resource utilization, protection is proactive and recovers fast response after a failure has occurred. In WDM backbone networks, protection is accepted to be the most preferred mechanism for survivability as it guarantees full recovery and fast recovery [87].

A number of multicast protection mechanisms have been proposed to prevent the WDM networks from network failures in literature [34–36, 86, 88–91]. Generally, such protection mechanisms can be classified as *dedicated* [92] or *shared* [93] ones, according to whether the backup resources can be shared or not. Based on the way that the requests are protected, they can be further classified into four different categories: tree-based, ring-based, path-based and segment-based ones [88]. The tree-based scheme is a straightforward way to protect a multicast session: find a primary tree first, and then calculates a link-disjoint backup tree for each multicast session [94]. Two light-trees are deemed to be link-disjoint only if they do not share any common link along their routes. The two link-disjoint trees usually carry the same data to the request destinations. Once any link on the primary tree fails, the affected network nodes will detect such a failure and report to the network

FIGURE 2.10: A multicast request protected by different protection schemes. (a) Request $R\{s, \{C, D\}\}$ protected by link-disjoint trees. (b) Link-disjoint tree scheme not applicable for request $R\{s, \{E, F\}\}$. (c) A request $R\{s, \{D, E, F\}\}$ protected by OCR. (d) Request $R\{s, \{E, F\}\}$ protected by OPP-SDP scheme.

controller, which in turn will appropriately reconfigure the network nodes to use the backup tree for data transmission. Figure 2.10(a) shows an example for the link-disjoint tree-based protection scheme for multicast session $R\{s, \{C, D\}\}$. The tree-based scheme is fast in fault recovery and is usually utilized to provide $1 + 1$ dedicated protection. However, it needs excessive network resources and sometimes the desired link-disjoint trees may not be found in a mesh network, which thus degrades network resource utilizations and causes higher request blocking probability.

Ring or partial ring is set up to protect each multicast session with ring-based protection schemes [95, 96]. The authors in [95] proposed two algorithms, namely one ring for one multicast session (OFO) and one ring for all multicast sessions (OFA), for static multicast request protection. Both OFO and OFA connect the session source and destinations with a single ring. The difference is that OFO protect each session with a single ring without resource sharing among the other sessions, while OFA uses a Hamiltonian ring to connect all the network nodes and allows resource sharing among different multicast sessions. The authors in [96] proposed another ring-based method, namely optimized collapsed ring (OCR), for

multicast protection. OCR connects the session source and all destinations using a partial overlapped bidirectional ring, as the example shown in Figure 2.10(c). The main issue with all these ring based methods is their inefficiency in wavelength capacity utilization.

Many path-based algorithms have been proposed for multicast protections [34, 88]. The main idea of all path-based methods is to compute link-disjoint path pairs from a session source to all the destination members so that a multicast session will be able to survive from any single link failure of the primary and backup paths. Based on this idea, an algorithm called optimal path-pair-based shared disjoint paths (OPP-SDP) is proposed in [88]. It allows its backup route to share edges with its existing primary route and backup edges. Figure 2.10(d) shows an example where two OPP paths are used to protect a request. In [34], a similar path based protection algorithm, named multicast protection through spanning paths (MPSP), is proposed. Compared to OPP-SDP, MPSP considered the backup path capacity sharing among different multicast sessions.

Segment is a sequence of links from any destination or branch node to any other branch or destination node of a light-tree [88]. Different segment based protection schemes were proposed for multicast request protection [35, 36, 88, 97]. All these proposed protection methods always try to construct a primary light-tree first, and then divide such primary light-tree into segments; finally each segment is protected separately by different backup route to improve network survivability. It has been reported in [36] that the segmented based methods have better performances in terms of resource efficiency and blocking probability among all the proposed protection schemes; the restoration time for the segment-based method is also shorter when compared to the other protection schemes [35].

Recently a number of other new technologies, e.g., p -cycle [98] and network coding [99], have been introduced for solving the network survivability problem. Although

such new techniques have some very nice advantages, such as the high capacity utilization with network coding and the speed-recovery of p -cycle, they are still not good enough for multicast protection, especially for the dynamic traffic scenario. This is because p -cycle is not flexible and efficient for dynamic multicast sessions, and the network coding process introduces some additional computational complexity and OEO conversions. For the network coding scheme, the high-speed synchronization process is difficult to implement and the coding process can only be realized in the electronic domain.

It is worth noting that all the above multicast protection schemes have considered only full-wavelength granulated multicast requests. For sub-wavelength granulated multicast request protection, few algorithms have been proposed. Specifically, although the authors in [100] have proposed two algorithms, namely Multicast Traffic grooming with Segment Protection (MTG-SP) and Multicast Traffic grooming with Shared Segment Protection (MTG-SSP), for the survivable multicast traffic grooming (SMTG) problem, the assumptions made in the paper is not valid for current optical networks as pointed out by the authors in [35].

The example in Figure 2.11 illustrates the problem with the protection scheme proposed in [100]. For the topology shown in Figure 2.11, if there is a sub-wavelength granulated multicast request $R\{S, \{A, C, E\}\}$ that arrives at the network, a survivable route as shown in Figure 2.11(a) can be found for it with the MTG-SSP algorithm. When there is link failure happens at link $S \rightarrow A$ as shown in Figure 2.11(b), the backup route will be activated as shown in Figure 2.11(c). As can be seen, although C and E can be connected to S , service of A is still interrupted as there is no backup wavelength reserved along any route to node A .

Such a problem is caused by the assumption adopted in [100]. The authors stealthily assume that a wavelength channel that is reserved in a fiber can be utilized to

FIGURE 2.11: A multicast request $R\{S, \{A, C, E\}\}$ served with MTG-SSP method. (a) The working and backup paths found with MTG-SSP method. (b) A link failure occurred at link $S \rightarrow A$. (c) The service of A cannot be restored with MTG-SSP.

transmit light signals in two opposite directions with simple switch reconfigurations at the two end-nodes of the path. However, such function cannot be realized within the current optical communication infrastructure. In the current optical network, there are two fibers in opposite directions with each carries a number of wavelengths; the optical input and output ports on an end-node switch are utilized to connect the optical fibers incoming to and outgoing from this node. Since the switching ports at the end-nodes are fixed, the reserved wavelength channels in a fiber cannot be utilized in both directions. In addition, although there are some ways to realize the bi-directional switching function, e.g., placing a pair of optical circulators in between the two end-nodes of the fiber as described in [35], many changes need to be made to the existing network infrastructure, which is too expensive and not a practical solution for near term implementation.

2.5 Summary

This chapter gives a brief review of current research that is related to the topics to be explored in this dissertation. Specifically, we first conduct a survey of wavelength routed WDM networks and relevant optical networking technologies, which include the IP/MPLS over WDM networks, one-to-one lightpath for unicast communication, and one-to-many light-tree scheme for multicast communication; then

we review the traffic grooming problem; finally, the different multicast protection schemes are reviewed.

Chapter 3

Dynamic LSP Routing in Overlay IP/MPLS over WDM Networks

3.1 Introduction

With the explosive Internet traffic growth and the rapid advancements of optical networking technologies, it is widely believed that the IP/MPLS over WDM network will be a key component for building the next-generation Optical Internet [13, 18]. As described in Chapter 2, the control plans of the IP/MPLS layer and the WDM layer can be interconnected via several different models, namely overlay model, augmented model and peer model. With the two layers controlled and managed independently while only limited information exchanges between them, the overlay IP/MPLS over WDM network is widely accepted to be the most promising network architecture for near-term deployment among these three interconnection models [19, 20]. In addition, since the IP/MPLS layer and the WDM layer networks are usually owned by different operators, the ability to protect confidential transactional information also argues well for the overlay model.

In an overlay IP/MPLS over WDM network, the only information exchanges between the two IP/MPLS layer and the WDM layers through the User to Network Interface (UNI) are the service requests and responses. The simplicity of information exchanges makes the overlay model easy to be implemented and, hence, more extensive deployment [101]. The recently demonstrated service oriented optical networks [102, 103] is a practical application of such an overlay architecture. Various request routing algorithms have been proposed for dynamic LSP routing in overlay IP/MPLS over WDM networks [71–74, 76], and they can be classified as sequential routing and resource-based routing, as reviewed in Chapter 2, Section 2.3.2. We notice that none of these algorithms consider the fact that a logical-layer *integrated service provider* (ISP) typically has the historical records of its own service requests that have been supported by the WDM-layer network, which could be utilized by the ISP to make better routing decisions.

To utilize the historical record in LSP routing, originally, we proposed two novel dynamic LSP routing algorithms, namely *Existing Link First* (ELF) and *ENquire first* (ENF) methods respectively, for the first time to the best of our knowledge, for the overlay IP/MPLS over WDM networks in [104]. The main idea of the algorithms is to use lightpath set up cost that is reported by the WDM layer network to estimate the cost of a logical link, and to update the cost of such logical link, if the logical link has not been chosen by the logical layer routing algorithm for a certain period of time. The time interval is determined by the overall network traffic load and the number of requests that have arrived at the network.

Later we discovered that these schemes could be further improved. The improved algorithms lead to significantly better Bandwidth Blocking Ratio (BBR) performances with the revised ELF consistently perform better than the revised ENF in various different scenarios. The bandwidth blocking ratio (BBR), which is illustrated below, is defined as the total amount of blocked bandwidth to the total

amount of user requested bandwidth to be transmitted,

$$BBR = \frac{\sum \text{blocked request bandwidth}}{\sum \text{bandwidth of all requests}} \quad (3.1)$$

Hence, we choose to report only the revised ELF algorithm, and compare it with the best existing schemes.

In what follows, we first present the adopted overlay system model, and then describe the (revised) ELF algorithm in detail; finally, we compare the BBR performance of the ELF algorithm to the existing algorithms for different scenarios.

3.2 System Model and Definition

An overlay IP/MPLS-over-WDM network architecture is illustrated in Figure 3.1.

FIGURE 3.1: The IP/MPLS over WDM network assumed in this chapter.

The optical-cross-connects (OXCs) on the WDM layer are interconnected with optical fiber links each carrying a number of wavelengths, while routers on the

IP/MPLS layer are connected to OXCs through optical ports. The WDM layer network is a bandwidth provider, which provides a dynamic connectivity service to the IP/MPLS layer network in the form of lightpaths. The IP/MPLS layer network acts as an integrated service provider (ISP), and its topology changes dynamically following the arrivals and departures of service requests. The network control and management functions of the two layers are independent of each other.

With a centralized management system, each of the IP/MPLS and WDM network layers is controlled by a different network operator. The respective operators maintain all the information of their own network layers, and only share the necessary information with each other through the control channels. For each arriving request, the controllers on the two layers will work cooperatively in accordance with the agreed service contract. Specifically, when a request arrives, the IP-layer controller will first try to find a route in the logical layer. If this step fails, the request is transferred to the WDM layer through the UNIs. Note that whether a request needs to be transferred to the WDM layer is decided by the controller of the logical-layer. When the request has been transferred to the WDM layer, whether the request can be served will be decided by the controller of the WDM-layer. Also note that information exchanges between the two layers can be carried out via any UNI; it is not necessary for these information exchanges to be carried out via a direct connection between the controllers of the two layers (such a direct connection may not even exist). It is for the ease of illustration and discussion that a direct information exchange channel between the controllers in the two layers is shown in Figure 3.1.

In the overlay network architecture, each network node is a photonic-cross-connect (PXC) interconnected with zero, one or more LSRs through the UNI. With the LSRs connected, a PXC is responsible for switching the bypass traffic from its input ports to output ports transparently [40, 41], receiving traffic terminating at the LSR to the local network, and transmitting traffic originating at the LSR from the

local network. The LSRs are responsible for traffic grooming as well as generating and terminating connections in the electronic domain. Note that the number of add/drop ports of a PXC equals the number of transmitters/receivers the node has. Such input/output ports are typically expensive due to the high-speed electronic processing requirement. Therefore, to save on the network cost without sacrificing network performance, a favourable solution is to let each PXC to be equipped with a limited number of add/drop ports shared by all the wavelength channels going through it [105]. In this chapter, we call such a PXC architecture as *port-limited*. In contrast, in a *port-unlimited* PXC, each wavelength channel is assigned a dedicated pair of add/drop ports. Define *add/drop ratio* r [73] of a PXC as $r = N_p/N_W$ ($0 < r \leq 1$), where N_p denotes the number of add/drop port pairs and N_W is the number of incoming/outgoing wavelength channel pairs of the PXC. By definition, $r = 1$ for a port-unlimited PXC.

In this chapter, we consider the general case of an overlay IP/MPLS over WDM network with N network nodes and L bi-directional optical fiber links where each link carries W wavelengths. Without loss of generality, each network node is assumed to be a PXC interconnected with a single LSR [74], and all PXCs have no wavelength conversion capability. Note that the work can be easily extended to networks with full or partial wavelength conversion capability. The node architecture adopted in this chapter is shown in Figure 3.2.

For the network considered, it is assumed that there is a single ISP that manages the IP/MPLS network above the WDM network, and the ISP has full link-state information of the logical-layer network. Each time a new request arrives, the ISP will try to find an appropriate route for it and decide whether to use the existing logical links or to set up new lightpaths between LSRs. If new lightpaths need to be established, the ISP will send the request to the WDM-layer bandwidth provider, consulting on the costs of using (leasing) these lightpaths. Based on the service contract and network resource availability, the WDM-layer bandwidth

FIGURE 3.2: A typical node architecture with a PXC interconnected with a single LSR that is utilized in this chapter

provider either feedbacks with the lightpath setup costs or rejects the lightpath establishment request (e.g., if it runs out of wavelength channel and/or add/drop port resources). Such decisions made by the WDM layer operator are independent of those made by the logical-layer ISP, as it has no knowledge of either the network topology or the available resources on the logical layer; and vice versa. Once an incoming request is provisioned on the logical layer network using either the existing links or new lightpaths, or both, the route will be fixed by the ISP so that the end users' services will not be interrupted.

The information exchanges between the IP and WDM layers basically include only the cost enquiries and feedbacks for the candidate lightpaths to be set up. Since theoretically speaking the number of candidate lightpaths may increase exponentially when the network size (number of nodes) increases, the ISP has to smartly select a small subset of candidate lightpaths without any knowledge of WDM-layer topology and resources availability. To make good decisions, it makes sense for the ISP to make use of historical records of lightpath costs. In this chapter, we make

the assumption that the ISP can keep a record of the lightpath setup costs for certain past period of time as well as the time at which such costs were reported by the WDM-layer bandwidth provider. Such historical records can be used to estimate the cost for setting up each candidate lightpath and consequently decide on the candidate route(s). The cost estimation and candidate routes selection methods will be discussed in the next section.

Table 3.1 presents a summary of the notations used in this study.

TABLE 3.1: NOTATIONS USED IN THIS CHAPTER

Symbol	Means
K	Number of candidate routes for each incoming request
n	The ID number of each incoming connection request
$\bar{H}$	The average number of optical hops traversed by each lightpath
$L(i, j)$	Logical link between LSR i and j
$P(i, j)$	New lightpath between LSR i and j
ω_{ij}	Number of idle wavelengths along $P(i, j)$
C_{ij}	Cost of $P(i, j)$
L_{ij}	Cost of $L(i, j)$
H_{ij}	Minimum number of optical hops between OXC i and j
p_i	Number of idle optical ports available on LSR i
U	The maximum number of historical records of cost for each logical link kept by the IP-layer ISP
T^n	The arrival time of the n -th LSP request
C_{ij}^n	Cost of $L(i, j)$ reported by WDM layer operator at T^n
C_{ij}^{est}	Estimated cost for $L(i, j)$ after the expiration time T_{ij}
T_{ij}	Estimated expiration time for L_{ij}
T_{ij}^{cal}	Calculated expiration time for C_{ij}^{est}
r	Add/drop ratio
$\bar{R}_{ij}$	Estimated average changing rate of L_{ij} during the time before T^n

3.3 Proposed Algorithm

In what follows, we describe the proposed historical data learning based algorithm. Section 3.3.1 describes the graph generation and cost assignment process for making routing decisions. Section 3.3.2 discusses the historical data learning and cost updating strategies; Section 3.3.3 presents the complete ELF algorithm; and finally Section 3.3.4 analyzes the complexity of the algorithm.

3.3.1 Graph Generation and Cost Assignment

For each incoming request, the ELF algorithm runs the *K-shortest path* (KSP) algorithm [106] on top of a randomly generated graph to find a desired number of candidate routes. The randomly generated directed graph represents the IP/MPLS-layer network with its nodes being LSRs, and its edges being existing logical links with sufficient residual bandwidth for the incoming request or potential new lightpaths to be set up on the WDM layer. We call those edges corresponding to the existing logical links with sufficient residual bandwidths as *existing links*, and those corresponding to the potential new lightpaths as *candidate new lightpaths* (CNLs). After running the KSP algorithm on the randomly generated graph, each of the resulting candidate routes may consist of only existing links, only CNLs, or both.

For each CNL involved in the candidate routes, the logical-layer ISP will signal to the WDM-layer bandwidth provider to enquire its cost. If the CNL is finally chosen to serve the request, a new lightpath will be set up on the WDM layer to support it. However, since an IP-layer ISP does not have the link-state information of the WDM-layer network, CNLs on the candidate routes may turn out to be infeasible due to exhausted wavelength channel and/or input/output port resources. If all the candidate routes are infeasible, the request is blocked.

Upon receiving a lightpath establishment request from the IP/MPLS layer, theoretically speaking, the WDM-layer bandwidth provider can use any routing and wavelength assignment (RWA) strategy to decide whether and how to set up the required lightpaths. In this study, since the main focus is to study the use of historical data on the IP/MPLS layer, we adopt the shortest hop count path routing and first-fit wavelength assignment strategy on the WDM layer. Other more sophisticated RWA strategies certainly can also be used but we believe the insights provided would be similar. As shall be seen in Section 3.4, by using the simplest RWA strategy, the proposed algorithm already outperforms the best existing algorithms not using the historical data. In many cases, the performance improvements are one to two orders of magnitude.

For each lightpath enquired by the ISP, the WDM-layer bandwidth provider will respond with a market price (say, measured in dollar) and a virtual cost for setting it up. The virtual cost is what has been agreed in the service contract. The ISP will keep a record of the virtual cost and carefully utilize such records in deciding the candidate routes and candidate lightpaths.

The virtual cost shall reflect the resource consumption for setting up the enquired lightpath without revealing detailed WDM-layer information. Also it should discourage over-utilizing a certain link or PXC to avoid the emergence of hot spots. The virtual cost therefore should reflect resource consumption as well as resource redundancy/scarcity of the enquired lightpath. To give the ISP strong incentives to minimize the virtual cost for setting up a connection, the market price and the virtual cost have to have strong positive correlation (e.g., the market price may increase faster than being linearly proportional to the virtual cost). This is a reasonable assumption for setting up a required connection at a lower virtual cost is in the interest of the WDM bandwidth provider. The bandwidth provider therefore should be willing to reward the cooperative ISPs with a lower market price. In this study, we assume that the ISP will try to lower the virtual cost and

always select the candidate route with the minimum virtual cost. Cases in which the market price is not positively correlated to the virtual cost will not be studied in this thesis.

In a port-limited PXC, the limited number of optical ports plays an important role in governing network performance. To improve the performance of a network with limited numbers of wavelengths and optical ports, these two types of resources should be consumed in a balanced manner. Specifically, if both wavelengths and optical ports are abundant, the costs of consuming them should be low and not so different from each other; while if any one of them becomes scarce, the cost of consuming the scarce resource should become higher to impose a penalty for utilizing it. Therefore, we define the cost C_{ij} of a new lightpath $P(i, j)$ by taking into account the costs of consuming optical ports and wavelength resources as follows:

$$C_{ij} = \begin{cases} \left(\frac{\alpha}{p} - H_{ij} \ln \left(1 - \frac{1}{\omega_{ij}+1} \right) \right) \times amp & \text{if } \omega_{ij} > 0 \text{ and } p > 0 \\ \infty & \text{if } \omega_{ij} = 0 \text{ or } p = 0 \end{cases} \quad (3.2)$$

This cost function is intended for both port-limited and unlimited cases. Specifically, it consists of two parts. The first part reflects the cost of consuming a pair of optical ports at LSR i and LSR j : $p = \min(p_i, p_j)$ denotes the minimum number of optical ports available at the two end nodes of the candidate lightpath. The parameter $\alpha = \frac{(1-r)}{r(H+1)}$ regulates the relative weights of the costs of a wavelength and an optical port: a smaller add/drop ratio leads to a higher cost of consuming a pair of optical ports. Note that the cost for consuming optical ports reduces to zero in the port-unlimited case where $r = 1$. The second part calculates the cost of consuming a wavelength along each hop of the lightpath. The negative symbol is to ensure that the second part is a positive quantity, and $\omega_{ij} + 1$ is used to avoid generating an infinity value when $\omega_{ij} = 1$. Finally, amp is an amplification factor

regulating the ratio of the cost of using existing logical links and that of setting up new lightpaths. Such a definition of lightpath cost helps avoid selecting a route with too few idling optical ports or too few idling wavelengths, or too many optical hops. With the above-proposed cost function, the traffic load distribution on the WDM is expected to be more balanced.

We now discuss on the logical-layer link cost assignment. For simplicity, we classify the CNLs into *cost enquired* and *cost unknown ones*. If the cost of a link has been enquired before, it is a cost enquired link; otherwise, it is a cost unknown one. The costs of different types of links are defined as follows.

$$L_{ij} = \begin{cases} 1 & \text{an existing logical link} \\ M_{ij} & \text{a cost unknown virtual link} \\ C_{ij}^{est} & \text{a cost enquired virtual link} \\ 2M_{ij} & \text{a failed lightpath for } L(i, j) \end{cases} \quad (3.3)$$

where M_{ij} is the default value for the cost of $L(i, j)$. The default value can be suggested by the WDM-layer network operator to the IP-layer ISP, e.g., as using the average cost from past transactions, or calculated using some typical values of the relevant parameters. From our experiments, we found using the following simple method to calculate M_{ij} can achieve satisfactory performance.

For a given WDM-layer network with an average nodal degree δ , denote the average hop length of each lightpath when adopting the fixed minimum-hop routing method as $\bar{H}$, and the average number of idle wavelengths on each optical link for a particular network state as ω . Since the average number of add/drop ports on each node equals $W \times r \times \delta$ approximately, the average number of idle optical ports on each node statusfor a particular network state can be estimated as

$$p = W \times r \times \delta - \frac{1}{\bar{H} + 1} \times \delta \times (W - \omega) . \quad (3.4)$$

The second part on the right side of the above equation comes from the fact that each lightpath only uses add/drop ports on its two end nodes. To avoid having a zero or negative value for p under heavy traffic loads, we let

$$p = \max \left(W \times r \times \delta - \frac{1}{H+1} \times \delta \times (W - \omega), 1 \right) . \quad (3.5)$$

The default value of M_{ij} for $L(i, j)$ can be calculated as

$$M_{ij} = \left(\frac{\alpha}{p} - H_{ij} \ln \left(1 - \frac{1}{\omega + 1} \right) \right) \times amp , \quad (3.6)$$

where the value of ω can be anything between 0 and W , depending on the network state. Our experiences show that the performance of the proposed algorithm is not very sensitive to the value of ω . A convenient option with satisfactory performance is to let $\omega = \frac{W}{2}$, which is adopted in all the simulations reported in this chapter.

3.3.2 Data Learning and Cost Expiration Process

As described in Sec. 3.2, each time a new request arrives at the network, the WDM-layer bandwidth provider will report the costs of some CNLs to the IP-layer ISP upon request, and the IP-layer ISP will keep a record of such information. To avoid keeping excessive records of historical data, only a limited number of the latest records are kept for each logical link. As shall be seen later, keeping a large number of dated records may not help to improve network performance.

Utilizing the historical data records, the IP-layer ISP is able to estimate the cost of each logical link. However, as the WDM-layer network operations are independent of those on the IP/MPLS layer, the cost of building a new lightpath between two end nodes may change significantly over time. To avoid outdated information leading to bad routing decisions, we introduce a cost expiration strategy: a link

cost record which was updated long time ago (e.g., longer than a pre-defined threshold) is deemed as outdated and thus should be adjusted, e.g., towards a certain default value. The threshold time for a link-cost record to be adjusted is termed its *expiration time*. Specifically, the cost expiration process for a logical link works as follows: whenever the cost of a logical link is reported by the WDM layer, its logical-layer record is updated accordingly. Meanwhile its expiration time is calculated. Upon the arrival of a new transmission request at a certain time T^n , the expiration time of the logical link is compared to T^n . If the expiration time has not been reached yet, the link cost record is regarded as valid, and thus can be used directly in calculating the candidate routes. Whereas if the expiration time has already been reached, it means that the link cost information has been kept for too long and thus should be adjusted towards its default value. After the adjustment, a new expiration time is calculated if needed. The cost expiration process is repeated until the estimated link cost equals its default value, or it gets updated by the latest information reported from the WDM-layer network.

In this chapter, we propose to let the estimated link cost to be adjusted towards its default value in a few steps upon expiration. The corresponding expiration time is calculated by utilizing the average cost changing rate from the historical records. Specifically, upon receiving accurate information on the cost of a logical link $L(i, j)$, denoted as C_{ij}^n , at time T^n , let

$$d_{ij} = \text{sgn}(C_{ij}^n - M_{ij}) . \quad (3.7)$$

and define its average cost changing rate as

$$\bar{R}_{ij} = \begin{cases} \frac{1}{m} \sum_{t=1}^m \left| \frac{C_{ij}^t - C_{ij}^{t-1}}{T^t - T^{t-1}} \right| & m < U \\ \frac{1}{U} \sum_{t=m-U+1}^m \left| \frac{C_{ij}^t - C_{ij}^{t-1}}{T^t - T^{t-1}} \right| & m \geq U \end{cases} , \quad (3.8)$$

where U denotes the maximum number of historical data records kept for each link. The expiration time of the link cost can be calculated as

$$\delta t = \begin{cases} \infty & C_{ij}^n = M_{ij} \\ T^n - T^{n-1} & C_{ij}^n \neq M_{ij} \text{ and } \bar{R}_{ij} = 0 \\ \frac{|C_{ij}^n - M_{ij}|}{\bar{R}_{ij}} & C_{ij}^n \neq M_{ij} \text{ and } \bar{R}_{ij} \neq 0 \end{cases} \quad (3.9)$$

Then for any request arriving at a certain time T^m , the estimated cost of $L(i, j)$ can be calculated as follows,

$$C_{ij}^{est} = \begin{cases} C_{ij}^n - d_{ij} \Delta |C_{ij}^n - M_{ij}| \times \min\left(\frac{1}{\Delta}, \left\lfloor \frac{T^m - T^n}{\delta t} \right\rfloor\right) & \delta t \neq \infty \\ M_{ij} & \delta t = \infty \end{cases} \quad (3.10)$$

where parameter Δ ($0 < \Delta \leq 1$) is a constant for controlling how much the link cost should be adjusted towards its default value once it is deemed to be expired.

The next expiration time of the newly adjusted link cost estimation C_{ij}^{est} can be calculated as

$$T_{ij}^{cal} = \begin{cases} T^n + \left\lfloor 1 + \frac{T^m - T^n}{\delta t} \right\rfloor \delta t & C_{ij}^{est} \neq M_{ij} \\ \infty & C_{ij}^{est} = M_{ij} \end{cases} \quad (3.11)$$

Once the cost record of a logical link is deemed as expired, L_{ij} and T_{ij} are updated by C_{ij}^{est} and T_{ij}^{cal} respectively, and such updating process is repeated until C_{ij}^{est} equals to M_{ij} or until it is updated by the new cost reported from the WDM layer. Simulation results show that the BBR performance is not very sensitive to the value of Δ . In our simulations, we let $\Delta = 0.2$, which achieves slightly better performance than $\Delta = 1.0$ (where link cost is adjusted to be equal to its default value once it is deemed as expired).

Note that in such data expiration process, different logical links may have different

cost expiration time, which makes sense since different links may be under different traffic loads, leading to different rates of link-cost changes. Our experiences show that having different expiration time for different links leads to better performance than updating all link costs with the same interval. Furthermore, note that the data learning and cost expiration process does not introduce any significant additional computational complexity to the routing process.

With the graph generation method and the data learning and cost expiration strategies as described above, we now present the ELF algorithm.

3.3.3 Existing Link First (ELF) Algorithm

The algorithm begins by assigning a cost to each link in the logical-layer graph, and then running the KSP algorithm on it to find a desired number of candidate routes for each incoming request. As discussed in Sec. 3.3.1, some candidate routes may contain only the existing links while the others contain some CNLs. The algorithm gives priority to using existing links since such a strategy generally leads to better performance [68, 107]. Specifically, it checks all the candidate routes; if there exist routes with only existing links, the one with the minimum cost is selected to serve the request. If no such route exists, the IP-layer ISP will query the WDM layer for the costs of all the CNLs along the candidate routes. Based on the CNL costs provided by the WDM layer, the feasible route with the minimum overall cost, if any, is selected to serve the incoming request. Note that historical records of CNL costs are used in calculating the estimated link costs in the process of finding the candidate routes and such records are (partially) updated each time there is a feedback on link cost from the WDM-layer network.

Since the algorithm gives a higher priority to utilizing the existing logical links while enquiring WDM layer for CNL costs only when necessary, we call it the

Existing Link First (ELF) algorithm. Algorithm 3.1 shows the key steps of the proposed algorithm.

Algorithm 3.1: EXISTING LINK FIRST (ELF) ALGORITHM

input : Network $G(V, E)$, Request $R(s \rightarrow d, b)$, K

output: A path route for request $R(s \rightarrow d, b)$

```

1 Initialization. foreach PXC pair do find a minimum hop optical layer route;
  foreach Logical link do  $L_{ij} = M_{ij}$ ,  $T_{ij} = \infty$  ;
2 for Each arriving request do
3   if It is a connection request then go to Line 6;
4   else go to Line 20;
5 end
6 for n-th LSP request arriving at the network at time  $T^n$  do
7   update the estimated cost and expiration time for those links whose costs
   become expired, i.e., if  $T_{ij} < T^n$  then  $L_{ij} = C_{ij}^{est}$ ,  $T_{ij} = T_{ij}^{cal}$ ;
8   IP-layer graph generation and cost assignment;
9   Run the KSP algorithm to find  $K$  candidate routes;
10  Check the link property of all links along the  $K$  candidate routes;
11  if there exist routes containing only existing links, then
12    choose the route with the minimum cost;
13  else
14    Enquire  $C_{ij}^n$  for all CNLs along the  $K$  candidate routes;
15    Update cost and expiration time for all cost enquired CNLs:  $L_{ij} = C_{ij}^n$ ,
     $T_{ij} = T^n + \delta t$ ;
16    Choose the route with the minimum overall cost, if applicable;
17  end
18  Serve the connection request using the selected route if there exists at least
   one feasible route; otherwise, block the connection request;
19 end
20 Update both WDM and IP network status;

```

3.3.4 Algorithm Complexity

As reviewed in Chapter 2, since all the existing algorithms find their candidate routes for each arriving request using Dijkstra's shortest path algorithm, the computational complexity of these algorithms is $O(L + N \log N)$.

Compared with the existing algorithms, the differences in complexity for ELF mainly come from three aspects:

- (1) The K -loopless shortest path (KSP) algorithm used to find K logical-layer candidate routes for each incoming request, with complexity of $O(KN(L + N \log N))$;
- (2) The data learning and cost expiration process, which introduces hardly any additional computational complexity, as discussed;
- (3) The storage space required for keeping record of historical costs and their corresponding reporting time, with a complexity of $O(UN^2)$;

For typical optical networks with a moderate number of nodes, the increased computational complexity and storage space needed by the ELF algorithm should be acceptable considering the performance improvement.

3.4 Performance Evaluation

As in previous work [72–74, 76], a dynamic traffic model is used to study the blocking performance of the proposed algorithm. Without loss of generality, only unidirectional LSP requests are considered. Assume that the LSP requests arrive at the network independently following a Poisson process with a mean arrival rate of λ , and the LSP holding time is exponentially distributed with a unit mean, i.e., $\frac{1}{\mu} = 1$. The source and destination node pair of each LSP request is randomly chosen among all the network nodes. The bandwidth of each wavelength is divided into 16 units, and the number of bandwidth units requested by each LSP is an integer uniformly distributed between 1 and 16. Each LSP request has to be handled along a single route without splitting.

We evaluate the proposed algorithm mainly by measuring the *bandwidth blocking ratio* (BBR) [72, 73] of the network. Extensive simulations are carried out on two commonly used network topologies. As shown in Figure 3.3, they include the 14-node NSFnet and the 46-node USNET. The average number of optical hops traversed by each lightpath is $\overline{H} = 2.18$ for NSFnet and $\overline{H} = 4.4$ for USNET respectively. Since in real-life networks, the cost of setting up a new lightpath is generally much higher than that of using an existing logical link, we set the amplification factor *amp* to be 10 and 5 for NSFnet and USNET respectively, such that the cost of using a new lightpath is roughly 5 times [108] as much as that of using an existing logical link in a port-unlimited network, where on average, half of all the wavelength channels are still idle and the corresponding average number of idle optical ports can be calculated using Eq. 3.4. Some other assumptions adopted in the simulations include: 1) each fiber link carries $W = 8$ wavelengths; 2) for the SLEA algorithm, the optical hop constraint is set to 2 for both NSFnet and USNET; 3) the number of optical ports at a network node is set to $W \times r \times \delta_i$ with δ_i being the node's fanout degree. Results shown in each of the following figures are the average of more than 30 independent simulations, each of which simulating 10^5 connection requests. We observe that the simulation results turn out to be highly consistent, with variance smaller than 4% when there is a single logical-layer candidate route for each communication request (or in other words when $K = 1$), and smaller than 7% for $K = 2$ and $K = 3$.

In what follows we compare the BBR performances of the ELF algorithm against those of the existing ones in port-unlimited and port-limited networks respectively. We also evaluate the effects of the number of optical ports and the amount of recorded historical information. Since all conclusions hold for both topologies, unless otherwise specified, we present only the results on NSFnet for comparisons and discussions.

FIGURE 3.3: Two network topologies used for simulations. (a) 14-node NSFnet. (b) 46-node USNET.

3.4.1 Performance Comparison in Port-Unlimited Networks under Different Traffic Loads

We first consider the case with port-unlimited PXCs. Figure 3.4 shows the BBR performances against network traffic loads measured in Erlangs. As can be seen, ELF significantly outperforms the three existing algorithms, namely SLEA, MLH_OVLY and OPF respectively. Specifically, when $K = 1$, ELF outperforms MLH_OVLY and SLEA by about an order of magnitude, and even more over OPF. When $K = 3$, the improvements increase to more than two orders of magnitude.

Note that among the three existing methods, OPF performs the worst, which is not a surprise since it tries only a single candidate route at a single layer for each arriving request. MLH_OVLY outperforms OPF by 70% when traffic load is about 70 Erlangs and 40% when traffic load is around 100 Erlangs, as it tests multiple routes for each request. By imposing an optical hop constraint on new lightpaths, the effect of which is the same as that of using a limited number of wavelength converters on certain network nodes, SLEA outperforms MLH_OVLY by about 20% within the whole range of traffic loads.

Comparison between ELF and MLH_OVLY convincingly demonstrates the effectiveness of utilizing historical data to improve the BBR performance of overlay

FIGURE 3.4: BBR comparisons between the ELF algorithm and the existing ones under different traffic loads without optical port limitation ($r = 1.0$, $\Delta = 0.2$, $U = 5$).

IP-over-WDM networks: both methods find more than one candidate routes for each incoming request, yet ELF performs much better, thanks to its careful utilizations of historic data. Comparison between ELF and SLEA shows that proper utilization of historical information helps the IP-layer ISP to make better routing decisions, leading to better performance than that of using high-cost wavelength converters on certain network nodes.

Another interesting observation from Figure 3.4 is that increasing the number of candidate routes K for each incoming request improves the BBR performance of ELF. Specifically, when increasing the number of candidate routes from $K = 1$ to $K = 2$, the average BBR improvement is 80% under different traffic loads; further increasing from $K = 2$ to $K = 3$, the average further improvement is 71%. Such improvements mainly come from two aspects: (i) a larger value of K gives ELF a higher chance to find a more appropriate candidate route for each incoming

request, and thus lowers the blocking probability; and (ii) a larger K allows more CNLs to be involved in the candidate routes. Their costs are therefore updated more frequently, enabling more accurate selections of candidate routes.

3.4.2 Performance Comparisons in Port-Limited Networks under Different Traffic Loads

We now study the BBR performance in port-limited networks under different traffic loads. In this subsection, we let $r = 0.6$ for all LSRs, while the effects of different values of r will be evaluated in the next subsection. Figure 3.5 shows the BBR performance against network traffic loads. The results again demonstrate that ELF significantly outperforms the existing algorithms: on average, when $K = 1$, ELF performs about 77%, 57% and 37% better than OPF, MLH_OVLY and SLEA respectively; while when $K = 3$, the improvements over the three algorithms increase to about 89%, 79% and 69% respectively.

A noteworthy observation in Figure 3.5 is that when the number of candidate routes for each incoming request increases, the performance of ELF also improves, though the improvements are not as significant as those in port-unlimited networks. Results show that when increasing K from 1 to 2, the average improvement is about 46%. Further increasing K from 2 to 3, the average further improvement is at a much lower value of 11%. The improvements are mainly due to the fact that increasing K makes more CNLs to be involved in the candidate routes and thus increases information exchanges between the two layers. The effects of having more candidate routes to increase the chance of finding a feasible route, i.e., the contribution (i) discussed in the last subsection, become less significant in port-limited networks. This can be explained as follows: the limited port resources make longer routes going through more existing links (and consequently consuming more optical ports) less favorable.

FIGURE 3.5: BBR comparison between ELF and the existing ones under different traffic loads when considering optical port limitation ($r = 0.6$, $\Delta = 0.2$, $U = 5$).

Overall, the results shown in Figure 3.5 demonstrate the effectiveness of the ELF algorithm in port-limited networks.

3.4.3 Influence of the Limited Number of Optical Ports

Figure 3.6 compares the performance of the ELF algorithm with those of the three existing algorithms with different add/drop ratio r . The traffic load is fixed at 80 Erlangs. Results show that ELF significantly outperforms the existing algorithms within a wide range of add/drop ratio. For $K = 1$, ELF outperforms the existing algorithms once $r > 0.45$; when $K \geq 2$, it outperforms them once $r \geq 0.4$. Closer observations reveal that, when the add/drop ratio is large enough, e.g., $r > 0.55$, ELF outperforms any existing algorithms with a wide margin; whereas when r is of a small value, all the algorithms perform nearly the same. This can be explained

FIGURE 3.6: Performance comparison between ELF and the existing algorithms with different values of the add/drop ratio r . The traffic load is fixed at 80 Erlangs ($\Delta = 0.2$, $U = 5$).

as follows: when limited optical ports become the bottleneck resource dominating the network performance, different algorithms do not make significant differences. However, once the bottleneck constraint is relaxed to a certain extent, i.e., when r is large enough, the ELF algorithm, with its capability of more efficiently utilizing network resources, easily stands out.

Another interesting observation from Figure 3.6 is that the performance of ELF improves steadily with an increasing value of r , which is different from that of the existing algorithms of which the performances stay largely unchanged once r is larger than a certain threshold value. This is because when r is large enough, wavelength resources, instead of optical port resources, become the bottleneck. Once the existing algorithms such as OPF, MLH_OVLY and SLEA reach their respective best utilizations of wavelength resources, they will not benefit from the redundant optical port resources. On the contrary, the ELF algorithm, with its

logical-layer dynamic routing process, enjoys better flexibility in utilizing network resources. Specifically, by carefully utilizing historical records and the enquired information of CNL costs, ELF avoids those CNLs with limited wavelength resources, sometimes at the cost of using more optical ports. Better utilization of redundant optical port resources leads to better performance with an increasing value of r .

3.4.4 Influence of the Amount of Historical Data

As mentioned in Section 3.3.2, for each logical link, ELF only keeps record of a small number of latest link costs reported from the WDM layer. We now evaluate how the amount of historical link cost information affects the network performance.

As shown in Eq. (3.8), in the ELF algorithm, the amount of historical data kept for each logical link is decided by the parameter U . A larger value of U lets the average cost changing rate to be calculated over a relatively longer time duration, and vice versa. Figure 3.7 shows the BBR performance with different values of U . Simulation results show that a smaller value of U basically leads to better performance, yet when U is too small (e.g., $U = 1$ or 2), there exists bigger fluctuations in the BBR performance. Such observations can be explained as follows: the link cost changing rate may not be very stable, especially considering the fact that we are using the dynamic routing method on the IP/MPLS layer. . If we use a long-term average to estimate the link cost changing rate within a short period of time in future, over- or under-estimation may happen. That is why a larger value of U does not lead to better performance. When U is too small, however, the average changing rate is only affected by how link cost changes with the last one or two connection requests, which may easily cause fluctuations. In this chapter, we set $U = 5$ in our simulations as it leads to more stable BBR performance, convenient for comparison against those of the existing methods. If

only the BBR performance is concerned, however, setting $U = 1$ may be a better option: a small value of U helps achieve better BBR performance and save the space for historical data storage.

FIGURE 3.7: ELF performance with different number of recorded costs without optical port limitation ($r = 1.0$, $\Delta = 0.2$, $K = 1$).

3.5 Summary

In this Chapter, we present a study on the dynamic LSP routing problem in overlay IP/MPLS over WDM networks. To improve the overlay network performance, we proposed to learn from the historical information maintained by the IP/MPLS-layer ISP. By carefully utilizing a data learning and cost expiration scheme for logical link cost estimation, and adopting the KSP algorithm for logical-layer routing, a novel algorithm called Existing-Link-First (ELF) was proposed. Extensive simulation results show that the proposed algorithm significantly outperforms all

the existing ones under different traffic loads, with both limited or unlimited optical ports as long as such resources are not too restrictive. The very significant improvements in BBR performances come at a cost of a light increase in computational complexity and a small amount of historical data storage on the IP/MPLS layer.

Since the main focus of this study is on utilizing historical data for dynamic LSP routing, detailed discussions on information exchanges between the IP/MPLS and the WDM layers have been omitted. As IP/MPLS over WDM technologies are maturing quickly, appropriate protocols will be developed in the near future for such information exchanges.

The three algorithms we adopted for comparisons belong to two different classes. OPF and SLEA are sequential routing methods, which try to support a request on a single network layer. Information exchanges between different layers are not needed and thus, keeping historical data does not help to improve network performances. On the contrary, MLH_OVLY belongs to the resource based methods, which allow setting up of a new connection across two different layers to achieve more efficient utilization of network resources. Algorithms of this class may be revised to make use of the historical data to help to facilitate logical layer routing. It would be of our future research interest to investigate how much historical data learning could help to improve the performances of these algorithms. Another interesting topic for further research is to extend the ELF algorithm to support multicast communications.

Chapter 4

Dynamic Multicast Traffic Grooming in WDM Networks with Lightpath and Light-Tree

4.1 Introduction

As presented in Chapter 2, wavelength routed WDM networks have emerged to be the dominant infrastructure for backbone networks, of which lightpaths are widely utilized as communication channels [5]. In recent years, as more and more multicast applications are becoming increasingly popular, a generalized lightpath concept named light-tree has been introduced to support multicasting on the physical layer of WDM networks [46]. Under the assumptions that multicasting with lightpath is realized using multiple unicast lightpaths from the request source to each of the destinations, and a lightpath channel can only be tapped but not dropped at any intermediate destination node, previous results show that the light-tree scheme consumes less network resources than lightpath when serving multicast requests with wavelength level granularity [109].

However, the bandwidth requirement of a multicast session usually ranges from several to tens of megabits per second (Mbps). It is much lower compared to the 2.5 – 40 gigabits per second (Gbps) capacity that can be provided by a single wavelength channel in today’s WDM networks, hence, if a lightpath/light-tree is utilized to support a single multicast request, much of the channel capacity will be wasted. In order to improve the wavelength channel capacity utilizations, multicast traffic grooming is usually adopted to pack several multicast requests together onto high-speed wavelength channels for transmission [67]. Extensive previous studies on static multicast traffic grooming problems have been done in the past years [110–113].

With the rapid developments of optical communication technologies, dynamic multicast traffic grooming become an increasingly important research topic, and a number of algorithms have been proposed [65, 78–84, 114]. We classify such algorithms into lightpath-based and light-tree-based algorithms in Chapter 2 Section 2.3.3. Even though some simple comparisons between these two groups of approaches have been made for multicast transmission [114] and many-to-many transmission [78], to the best of our knowledge, there is no systematic comparison between their respective blocking performances for dynamic multicast traffic grooming.

In this chapter, we address the dynamic multicast traffic grooming problem in WDM networks, and present comprehensive comparisons between lightpath and light-tree schemes in different cases. Our main contributions are two-fold: first, we compare the performances of the existing lightpath and light-tree based grooming algorithms, and show that in most cases, the lightpath-based methods outperform the light-tree based ones. We discuss and explain such observations. Second, we propose a new lightpath-based algorithm, named LightPath Fragmentation (LPF) method, to further improve the network blocking performance. Numerous simulations show that the LPF method outperforms the existing algorithms in

different cases. The effects of the ratio of unicast traffic loads versus overall traffic loads and the average number of destinations for each multicast request are also studied.

Below we first present a brief comparison between the existing approaches and explain why the lightpath based approaches have a winning margin, and then describe the proposed LPF algorithm for network performance improvement; finally, we compare the blocking performances of all the algorithms in different cases.

4.2 Lightpath VS. Light-Tree: A Brief Comparison

4.2.1 Physical Layer Node Architectures

To support multicasting, some or all the network nodes need to be able to copy data, either in the electronic domain or optical domain or both, from a single input port to multiple output ports. To support traffic grooming, a node should be able to convert optical signals into electronic signals, perform appropriate traffic aggregation, and finally convert the signals back into optical domain. The devices that perform traffic grooming operations are called grooming fabrics [65].

Two typical kinds of switch architectures, which are described in Chapter 2 Section 2.3, have been proposed in literature for dynamic multicast traffic grooming; namely, grooming capable optical cross-connect (GC-OXC) [67] and multicast capable optical grooming switch (MC-OGSW) [65], and are shown in Figure 2.9(a) and Figure 2.9(b), respectively. For lightpath based dynamic multicast traffic grooming, the GC-OXC is usually utilized, while for light-tree based multicast traffic grooming, MC-OGSW is utilized. However, since the multicast capability of a MC-OGSW switch is constrained by the size of the splitter banks, the

MC-OGSW may cause some requests to be blocked if it is not equipped with sufficient light-splitter banks. In this chapter, since our main focus is to conduct fair comparisons between lightpath and light-tree based schemes, we assume that all MC-OGSWs are equipped with abundant light-splitter banks and hence there is no request blocking due to limited number of light-splitters.

For both architectures, the OXC is equipped with a number of add/drop ports, and the number of ports generally equals the number of transceivers on the node. Note that both the add/drop ports and the transmitters/receivers are of high costs due to their high-speed processing units. Hence, to save network cost without sacrificing network performance, each network node is usually equipped with a limited number of such port pairs shared by all wavelengths going through it. In this chapter, we define *add/drop ratio* r ($0 < r \leq 1$) as the ratio of the number of add/drop port pairs over the total number of wavelengths they support. For a node with $r < 1$, we call it a port-limited node; otherwise, it is a port-unlimited node.

Although previous results claimed that traffic duplication in the optical domain using passive light-splitters is less expensive than in the electronic domain [110], using light-splitters does not allow convenient traffic grooming. Moreover, optical splitting may result in higher power losses and degraded signal quality. It is worth noting that if both architectures are equipped with the same number of transceivers, an MC-OGSW switch may be more complex and expensive than a GC-OXC switch due to the power loss compensation units it may need to be equipped with.

4.2.2 Network Layer Multicast Traffic Grooming Methods

Existing results show that light-tree based schemes consume less network resource than lightpath based schemes when serving multicast requests with wavelength

level granularity [109]. It remains unclear which scheme is a better choice for requests with sub-wavelength granularities. We argue (and confirmed by extensive simulation results as reported later in this chapter) that the lightpath based methods may be more suitable for supporting dynamic traffic grooming, mainly because it allows more efficient traffic grooming. An illustrative example is shown in Figure 4.1.

In Figure 4.1, we assume that each node is equipped with only one pair of transceivers, and the fiber link is bidirectional carrying only a single wavelength in each direction. Suppose a multicast request $R1 : \{S, \{C, D\}, \frac{1}{4}\}$ arrives at the network, where $S, \{C, D\}$ and $\frac{1}{4}$ are the source, destination set and the bandwidth requested versus the channel capacity, respectively.

When the light-tree scheme is utilized, we can easily find a route for it as shown in Figure 4.1(a). However, though such a route is optimal for the current request, bandwidth may be wasted when a new request, e.g., $R2 : \{S, \{C\}, \frac{1}{4}\}$, arrives. Specifically, by accommodating it using the adaptive light-tree method, e.g., MDTGA or EMGA, $R2$ will be groomed into $R1$, and the bandwidth from B to D may be wasted; if a fixed light-tree algorithm is adopted, $R2$ will be blocked (Note that LTD-ANCG decomposes $R1$'s light-tree as shown in Figure 4.1(b)).

When a lightpath based scheme is adopted, however, the situation is different. For $R1$, a route as shown in Figure 4.1(c) can be found; while for $R2$, it can be served by the existing lightpath from S to C . Compared to the ordinary light-tree solution as shown in Figure 4.1(a), the lightpath solution achieves a better bandwidth blocking performance at the cost of consuming one more transmitter, but it also saves a light-splitter. When compared with LTD-ANCG as shown in Figure 4.1(b), it saves one receiver and one light-splitter.

From the above example, we observe that although lightpath may consume more transceivers for the current request in the traffic grooming process, it increases

FIGURE 4.1: An example of multicast traffic grooming: (a) ordinary light-tree solution; (b) light-tree decomposition with LTD-ANCG; (c) ordinary lightpath solution; (d) fragmentation of long lightpaths.

the chance that these transceivers may be conveniently utilized to groom future requests and consequently, helps improve network blocking performance. For networks with the given link capacity/transceiver resources, it is not easy to tell which scheme will perform better. As mentioned above, there are no detailed comparisons between the bandwidth blocking performances of these two schemes to the best of our knowledge.

4.2.3 Possible Further Improvements

The example above reveals that the main advantage of lightpath based schemes is that they allow more convenient “sharing” of resources between different multicast sessions. To make further improvements, the chance that different multicast sessions can be groomed should be further increased. One possible approach is to break a long lightpath into a few shorter ones if, and only if, such a move will *not* over-utilize transceiver resources, making them scarcer resources restricting the network in handling future multicast sessions.

An example is shown in Figure 4.1(c). Consider a new request $R3 : \{A, \{C, D\}, \frac{1}{4}\}$ that arrives after $R1$ and $R2$ have been successfully accommodated. Since there is no available wavelength along the route from A to C , this new request will be blocked, even if node A still has idle transmitters and the residual capacities along all the links are sufficient. However, if we break the long lightpath from S to C at the intermediate node A at the moment when it is set up, as shown in Figure 4.1(d), then $R3$ can be supported. Compared to the ordinary lightpath solution, the fragmentation process consumes an additional pair of transceivers for serving $R1$, yet it prevents $R3$ from being blocked, and may accommodate some future requests initiating or terminating at node A .

The above example shows that, by properly fragmenting long lightpaths into a few shorter ones while carefully keeping a balance between saving link capacity and transceivers, the network blocking performance may be improved.

4.3 Lightpath Fragmentation (LPF) Algorithm for Dynamic Multicast Traffic Grooming

4.3.1 Problem Statement and Main Idea

Let a network be represented as a graph $G(V, E)$, where V denotes the set of network nodes and E the fiber links. Assume that each link is composed of two fibers in opposite directions, each carrying W wavelengths. A multicast request is represented as $R\{s, D, b\}$, where s , D and b denote the source, the set of destination nodes, and the required bandwidth of the request, respectively. A request is served only when all its destination nodes can be served; otherwise, the request is blocked.

The dynamic multicast traffic grooming problem can be defined as follows. In a given network with dynamic arrivals of unicast/multicast connection requests,

based on the global information of link state and availability of transmitter/receiver resources on each node, a centralized algorithm is to be devised to support these connection requests with the objective of minimizing the network BBR. Note that we assume that any ongoing transmission cannot be interrupted. In other words, an existing lightpath cannot be fragmented or rerouted when there is ongoing transmission going through it.

In this section, we propose a new lightpath-based multicast traffic grooming algorithm which we call *LightPath Fragmentation* (LPF) algorithm. The main idea of the algorithm is to enhance resource sharing between different multicast sessions by adopting proper fragmentation of long lightpaths. We will show that, by keeping a balance between link capacity resources and transceiver resources, LPF algorithm significantly outperforms all the existing algorithms.

In what follows, we will first discuss how to choose the nodes along a lightpath for lightpath fragmentation and then present a detailed description of the LPF algorithm.

4.3.2 Selection of Fragmentation Nodes

Upon the arrival of a new multicast request, a tree route will be found for it. When lightpath based schemes are used, each tree may contain one or more lightpaths, each of which traverses one or multiple optical links. As discussed earlier, having long lightpaths may lower the chance of resources sharing in supporting future multicast sessions. We propose to fragment newly set-up long lightpaths if, and only if, such an action will not over-utilize network transceiver resources.

Suppose P is a new lightpath that is found for a multicast request, and n_i is an intermediate node of P , of which the fanout degree is d_i . Denote the number of idle transmitters and receivers on n_i as T_i and R_i respectively, and the numbers

of free wavelengths on the incoming and outgoing links which the new lightpath goes through as ω_{in} and ω_{out} respectively. To determine whether P should be fragmented at n_i , there is a need to figure out whether wavelength channels or transceivers are more limited resources (hereafter to be called as *bottleneck* resources) on this node. Lightpath fragmentation happens on node n_i if and only if wavelength channels turn out to be the bottleneck resources. Strictly speaking, which resources are the bottleneck resources will depend on the *traffic pattern* in particular, the traffic load between individual source-destination pairs. To avoid complicated computations based on the knowledge of traffic pattern, which anyway may not be easily available in many real-life systems, we consider the benchmark case where a lightpath is to be set up from node n_i to each of the other nodes in the network and see which resources become the bottleneck first. As shall be seen later, this simple method leads to satisfactory performance. Specifically, to figure out the bottleneck resources, two parameters are defined as follows:

$$\alpha_m = \min\left(\frac{T_i}{d_i \times \omega_{out}}, \frac{R_i}{d_i \times \omega_{in}}\right) \quad (4.1)$$

$$\alpha = \frac{1}{H_i} \quad (4.2)$$

where H_i is the average hop length of the shortest paths from n_i to all the other nodes. In the above equations, α_m denotes the smaller one between the add and drop ratios on n_i , and α is the add/drop ratio required for n_i to support lightpaths from itself to each of the other nodes. Note that for a certain network node, α is a constant once the network topology is given.

For each new lightpath that is found for supporting a multicast request, α_m is calculated on each intermediate node along the lightpath and then compared to α of the node. If $\alpha_m > \alpha$, transceivers are not regarded as bottleneck resources and hence the lightpath is fragmented at this node; whereas when $\alpha_m \leq \alpha$, it indicates

that the remaining transceiver resources on this node are limited and saving the transceiver resources may help to support the future requests. The lightpath is therefore not fragmented at this node. Adopting this simple strategy, we propose the LPF algorithm for dynamic multicast traffic grooming. Note that the algorithm merely aims at optimizing network blocking performance without worrying about any possible drawbacks of lightpath fragmentation, e.g., the possible longer delay. The algorithm however can be revised to take into account some other constraints, which will be briefly discussed in Section 4.4.4.

4.3.3 LPF Dynamic Multicast Grooming Algorithm

Since finding the optimal route for multicast traffic grooming is an NP-complete problem [67], we adopt the simple minimum cost path heuristic (MPH) [56] to find the tree route for each request. The main idea of MPH is to use the minimum cost paths to connect request destinations one by one to the tree for a request.

The main working steps of LPF are shown as follows.

With MPH, LPF utilizes existing logical links to serve as many destinations as possible in Step 1. Steps 2 – 12 then serve the remaining destinations, if any, by setting up new lightpaths and fragmenting the new lightpaths when applicable. Specifically, Step 4 serves a destination node by setting up a new lightpath; Step 5 fragments this new lightpath when applicable; and Steps 6 – 11 try to groom the fragmented new lightpaths into existing logical links. Finally, Steps 13 – 14 update the network status. Note that on both the logical and the optical layers, the “first-fit” wavelength assignment policy is adopted.

Different path length definition can be adopted in Step 4. The simplest way is to define the path cost as equal to its hop length. In our experiences, even with this simple definition, LPF manages to outperform all the existing methods, in many

Algorithm 4.1: LPF Multicast Traffic Grooming Algorithm

input : Network $G(V, E)$, Request $R\{s, D, b\}$

output: A set of lightpaths for serving $R\{s, D, b\}$

- 1 Grooming with existing lightpaths: Generate an auxiliary graph (AG) using existing lightpaths with enough residual bandwidth; Call **MPH** on AG to initiate a tree from s to connect as many members of D as possible. Remove served members from D , and if D is empty, go to Step 14; otherwise, save the partial tree T , and continue.// *logical-layer grooming*
 - 2 Add s and all the nodes on T to a set S ; // *optical-layer processing (2-12)*
 - 3 **while** $D \neq \Phi$ **do**
 - 4 Call **Optical-layer-routing-sub-algorithm**($G(V, E), S, D$), returns a new lightpath P ;
 - 5 Call **Lightpath-fragmentation-sub-algorithm**($G(V, E), S, P$), returns some new lightpaths;
 - 6 **for** each new lightpath P_i **do**
 - 7 Check existing lightpaths with enough residual bandwidth;
 - 8 **if** a certain existing lightpath P_e has the same source and destination as P_i **then**
 - 9 add P_e onto the partial tree T found in Step 1 and delete P_i from the new lightpath set.
 - 10 **end**
 - 11 **end**
 - 12 **end**
 - 13 Allocate transceivers and wavelength to each new lightpaths.
 - 14 Update residual capacity of all links of the logical-tree T .
-

Procedure 4.1: Optical-layer-routing-sub-algorithm

input : A network $G(V, E)$, two node sets S and D .

output: A new lightpath P to serve a request destination.

- 1 Generate an optical-layer AG; calculate all-to-all shortest paths between any node in S to any node in D by adopting a proper path cost definition (as later defined by Eq. (3)).
 - 2 Choose the shortest one among the shortest paths connecting a certain member in S to a certain member d in D . Denote the distance of the path as dis .
 - 3 **if** $dis < \infty$ **then**
 - 4 $S = S \cup \{d\}$; $D = D \setminus d$; Save the shortest path P ;
 - 5 **else**
 - 6 Block the request R , **break**;
 - 7 **end**
-

Procedure 4.2: Lightpath-fragmentation-sub-algorithm**input** : A network $G(V, E)$, a node set S and a lightpath P **output:** A set of new lightpaths.

```

1 while any node of  $P$  has not been checked do
2   for each intermediate node (if any)  $n_i$  along  $P$  do
3     Calculate  $\alpha_m$  for  $n_i$  ;
4     if  $\alpha_m > \alpha$  at  $n_i$  then
5       Fragment  $P$  at  $n_i$ , and get two new lightpaths  $P_a$  and  $P_b$ ;
6        $S = S \cup \{n_i\}$ ;  $P = P_b$ ;
7     end
8   end
9 end

```

cases by more than an order of magnitude. To further improve the performance, a better lightpath cost definition proposed in Chapter 3 Section 3.3.1, as shown below, is adopted

$$C_{ij} = \begin{cases} \frac{(1-r)}{p \times r \times (\overline{H}+1)} - H_{ij} \ln \left(1 - \frac{1}{\omega_{ij}+1} \right), & \text{if } \omega_{ij} > 0 \text{ and } p > 0 \\ \infty, & \text{if } \omega_{ij} = 0 \text{ or } p = 0 \end{cases} \quad (4.3)$$

where p is the smaller one among the number of transmitters at the source and the number of receivers at the destination of the lightpath; ω_{ij} is the number of available wavelengths along the lightpath route; $\overline{H}$ is the average path length of the network, and H_{ij} is the minimum hop length between two end-nodes of the lightpath. This function helps keep balanced consumptions of wavelength and transceiver resources. Specifically, if both resources are abundant, the costs of consuming them should be low and not so different from each other; while if any one of them becomes scarce, the cost of consuming the scarce resource will become higher to impose a penalty on its use. Extensive simulation results show that adopting this definition leads to the best performance among all the link-cost definitions we have tested. In this chapter, we use this definition in calculating the path cost.

4.4 Lightpath VS Light-Tree: Simulation Results and Discussions

Simulations are carried out to compare the lighptah and light-tree based grooming algorithms in different cases. Below we firstly present the performance metrics and simulation environment, and then show the simulation results in different cases.

4.4.1 Performance Metrics and Simulation Environment

Performances of all the algorithms are mainly measured by their bandwidth blocking ratio (BBR), which is defined by 3.1. To assess the grooming efficiency of lightpath and light-tree, the *average channel capacity utilization* (U_w) is defined as follows,

$$U_w \stackrel{def}{=} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{1}{W_u} \sum_{W_u} \frac{B(t)}{B} dt$$

where $B(t)/B$ measures the portion of channel capacity being utilized and W_u is the number of wavelengths being used. We see that U_w measures the capacity utilization of the channels being used for transmission.

To evaluate the required OEO conversions for each request which, as will be discussed in more details later, helps to reflect the average intermediate node processing delay experienced by each admitted session, the *average number of OEO conversions per session* is defined as

$$N_{\text{OEO}} = \frac{\sum_{NReq} \text{no. of OEO used for the request}}{NReq}$$

where $NReq$ is the number of admitted connection requests.

Two commonly used network topologies as illustrated in Figure 4.2 are adopted in our simulations. They are the 14-node, 21-link NSFnet, and the 11-node, 26-link

COST239 network. The results to be shown for each of these two networks are the average of at least five independent simulations, each of which with at least 10^5 connection requests.

The following are the key assumptions used in the simulations:

- 1) For lightpath based algorithms, all network nodes are equipped with GC-OXC, while for light-tree based ones, MC-OGSWs are utilized; for fair comparisons, both switches are equipped with the same number of transceivers. As aforementioned in Section 4.2.1, we assume that all MC-OGSWs have sufficient light-splitters. The fan-out degrees of all light-splitters are not limited, and the power of output signals from any light-splitter is high enough for detection.
- 2) Each link is composed of two fibers of opposite directions, each carrying $W = 32$ wavelengths; the capacity of each wavelength is $B = 16$ units.
- 3) Requests arrive at and leave the network according to a Poisson process with a rate of λ , and their holding time follows the negative exponential distribution with a mean value of $\mu = 1$; the bandwidth for supporting each request is an integer uniformly distributed in $[1, 16]$.
- 4) The number of transceivers on a node is set to be $W \times d_i \times r$, where d_i is the fanout degree of the node.
- 5) Signal power loss due to light-splitting or transmission attenuation is neglected.

Four request generation models are used in the simulations:

M1: a number of randomly pre-selected requests that arrive and leave the network independently. For NSFnet, such requests are $\{7, \{1, 5, 8\}\}$, $\{3, \{9, 12, 6\}\}$, $\{10, \{2, 4, 8\}\}$ and $\{11, \{0, 13, 6\}\}$; while for COST239, the requests are $\{2, \{10, 5, 7\}\}$, $\{1, \{0, 8, 5\}\}$, $\{4, \{6, 3, 10\}\}$ and $\{9, \{5, 7, 3\}\}$.

FIGURE 4.2: Two network topologies utilized for simulations: (a) 14-node NSFnet; (b) 11-node COST239 topology.

M2: the source and destination of a request are randomly chosen from two separate pre-selected sets. Similar to that in *M1*, we define them as $\{7, 3, 10, 11\}$ and $\{\{1, 5, 8\}, \{9, 12, 6\}, \{2, 4, 8\}, \{0, 13, 6\}\}$ for NSFnet, and $\{2, 1, 4, 9\}$ and $\{\{10, 5, 7\}, \{0, 8, 5\}, \{6, 3, 10\}, \{5, 7, 3\}\}$ for COST239.

M3: the request source and destinations are randomly chosen among all network nodes, with the destination number limited to a certain range. In this study, the destination number is uniformly distributed in $[2, 4]$ for both topologies.

M4: the source and destinations of the requests are randomly chosen; the destination number is distributed in $[2, N - 1]$ following the truncated geometric distribution with parameter q [81]. The average destination number is set to 3, and thus, for NSFnet, $q = 0.501$, while for COST239, $q = 0.504$.

By testing on the above models, we evaluate the performances of different algorithms in different cases with uniform and non-uniform traffic patterns and different numbers of destinations for each session.

Since all the conclusions hold for both topologies, unless otherwise specified, we present only the results on NSFnet for comparisons and discussions.

4.4.2 Lightpath vs. Light-tree: Effects of Different Traffic Models and Traffic Loads

Figure 4.3 compares the BBR performances of different algorithms under a fixed traffic load $\rho = 250$ Erlangs with different traffic patterns. From the simulation results, it is interesting to observe that: 1) for $M1$ and $M2$, there is no obvious winner among the *existing* lightpath and light-tree based algorithms, while for $M3$ and $M4$, the lightpath based methods have a significant winning margin over the light-tree based ones within the whole range of add/drop ratio; 2) the proposed LPF algorithm outperforms all the existing algorithms in different cases, especially when the network has a high add/drop ratio.

It is not difficult to understand why the performances of all the existing methods are similar for $M1$ and $M2$: when there are very limited options in selecting source/destinations for each session, connections are setup and torn down frequently only along a few tree routes. The large imbalanced distribution of traffic loads in the network degrades network resource utilizations and causes high BBR for all the algorithms. When multicast sessions are more evenly distributed in the network, e.g., as in $M3$ and $M4$, those algorithms that are able to achieve higher network resource utilizations gain a commanding winning margin.

Note that even for $M1$ and $M2$, the LPF method outperforms all the existing methods, esp. in networks with a high add/drop ratio, thanks to its enhanced resource sharing.

Due to space limit, hereafter we will use $M3$ as a representative case for performance comparisons, since the request generation in $M1$ and $M2$ has limited flexibility while $M4$ may lead to some over-sized light-trees (and rather bad performance) when the light-tree methods are used. Note that the following conclusions always hold: *i*) the differences between the performances of different algorithms

(a)

(b)

FIGURE 4.3: Lightpath vs. light-tree: BBR performances of the four different methods vs. add/drop ratio under traffic with four different patterns. Traffic load is set to be 250 Erlangs. (a) M1. (b) M2.

(c)

(d)

FIGURE 4.3: Lightpath vs. light-tree: BBR performances of the four different methods vs. add/drop ratio under traffic with four different patterns. Traffic load is set to be 250 Erlangs. (c) M3. (d) M4.

tend to be more significant under $M3$ and $M4$ compared to those under $M1$ and $M2$; and *ii*) LPF always outperforms all the existing algorithms.

FIGURE 4.4: Lightpath vs. light-tree: BBR performance versus add/drop ratio r under high traffic load $\rho = 450$ Erlangs for $M3$.

Figure 4.4 compares the performances of different algorithms under a higher traffic load of 450 Erlangs. As can be seen, LPF remains as the best-performing algorithm. In fact, it is the only algorithm which manages to drive BBR to lower than 1%. With a high add/drop ratio, it can drive BBR all the way to lower than 0.01%.

The observation that LPF outperforms even the best existing lightpath based method, namely the SC method, by more than an order of magnitude when add/drop ratio $r \geq 0.7$ demonstrates the benefits of lightpath fragmentation: the better utilization of the redundant transceiver resources helps enhance wavelength capacity sharing as well as releasing the wavelength continuity constraint, and consequently, significantly improve the network blocking performance.

(a)

(b)

FIGURE 4.5: Lightpath vs. Light-tree: BBR under different traffic loads. (a) Network with port-limited nodes where $r = 0.6$. (b) network with port-unlimited nodes where $r = 1.0$.

Figure 4.5 further compares the BBR performances of different algorithms with the given add/drop ratios but under varying traffic loads. Specifically, we consider two different cases where $r = 0.6$ and $r = 1.0$, respectively. As can be seen, though the differences between BBR of the LPF algorithm and those of the other algorithms tend to decrease under increasing traffic loads (which is not a surprise), even when under very heavy traffic loads where BBRs of all the other algorithms are close to or higher than 10%, the LPF algorithm still outperforms all the other algorithms by more than 50% where $r = 0.6$. For the case where $r = 1.0$, the difference is always higher than an order of magnitude; under moderate traffic loads, the differences are in three or four orders of magnitude.

4.4.3 Lightpath vs. Light-tree: Effects of the Number of Multicast Destinations and the Fraction of Unicast Traffic Loads

In this section, we evaluate how the BBR performances of all the algorithms change with two important factors of traffic pattern, namely the number of multicast destinations and the fraction of unicast traffic loads versus the overall traffic loads. Specifically, for the former evaluation, we adopt a request generating model that is similar to M3 but with the number of destinations uniformly distributed in $[2, 2(D_n - 1)]$, where D_n is the average number of destinations per multicast session; while for the latter evaluation, we define *unicast traffic ratio* as the ratio of unicast traffic to the overall network traffic loads, and let the destination number of each multicast request be uniformly distributed in $[2, 4]$.

Figure 4.6 shows the BBR performances of all the algorithms in networks with either port-limited or port-unlimited nodes. As can be seen, under a fixed traffic load, the BBR performances of all the algorithms degrade with an increasing average number of destinations. The main reason is obvious: requests with larger

(a)

(b)

FIGURE 4.6: Lightpath vs. Light-tree: effects of the number of multicast destinations. (a) Network with port-limited nodes where $r = 0.6$, traffic loads are 300 Erlangs. (b) Network with port-unlimited nodes where $r = 1.0$, traffic loads are 500 Erlangs.

number of destinations tend to consume more network resources. It is not a surprise to see that LPF has a larger winning margin in a port-unlimited network than in a port-limited network since it makes better use of the redundant transceiver resources.

Figure 4.7 compares BBR performances of different algorithms with an increasing unicast traffic ratio, still in networks with either port-limited or port-unlimited nodes. Simulation results show that in both networks, the BBR performances of all the algorithms improve with an increasing unicast traffic ratio, and again LPF has a bigger winning margin over the other algorithms in port-unlimited networks than in port-limited networks.

Figures 4.6 and 4.7 also show that the lightpath based algorithms outperform the light-tree based ones, and LPF outperforms its closest competitor, the SC algorithm, by at least an order of magnitude under most cases. The only exception is when the average number of multicast destinations is very large, where the exhausted link capacity resources drive all the algorithms to have a high BBR. In the most favorable case where the traffic loads are mostly unicast traffic, LPF outperforms its closest competitor by more than three orders of magnitude.

4.4.4 Lightpath Based Schemes: Capacity Utilization, Delay and Consumptions of Transceiver Resources

Figure 4.8 compares the average capacity utilization of non-idle wavelength channels of different algorithms. As can be seen, lightpath based schemes achieve much higher utilization than light-tree based ones within the whole range of add/-drop ratio. Such observation is not a surprise since, as mentioned earlier, lightpath based schemes help achieve easier grooming of traffic loads and consequently, much higher utilization of wavelength capacity resources.

(a)

(b)

FIGURE 4.7: Lightpath vs. Light-tree: effects of the unicast traffic ratio. (a) Network with port-limited nodes where $r = 0.6$, traffic loads are 300 Erlangs. (b) Network with port-unlimited nodes where $r = 1.0$, traffic loads are 500 Erlangs.

FIGURE 4.8: Lightpath vs. light-tree: average existing-channel capacity utilization.

While lightpath based algorithms lead to better BBR performance, a natural concern is that they may lead to more OEO conversions for each connection request and consequently lead to a longer delay. Figure 4.9 compares the average number of intermediate OEO conversions each connection has to go through, which provides an indirect yet good measurement of the expected delay experienced by each admitted request. As we can see, for $r \leq 0.4$, the numbers of OEO conversions required by SC and LPF decrease with an increasing value of r ; while when $r > 0.4$, such numbers increase with r for LPF, yet stay largely unchanged for SC. For EMGA and LTD-ANCG, such numbers remain largely unchanged throughout the whole range of r .

Such observation can be understood as follows: when transceiver resources are too limited, e.g., $r = 0.1$, the lightpath based methods can only set up a small number of lightpaths. Consequently, any admitted request tends to be groomed into existing lightpaths as far as possible, which causes a relatively large number

FIGURE 4.9: Average number of OEO conversions experienced by each admitted request. Traffic loads are 250 Erlangs.

of OEO conversions per request. When transceiver resources become redundant, more lightpaths can be set up, many of which may bypass some intermediate nodes along its route. The number of intermediate OEO conversions therefore decreases. When the number of transceivers further increases, e.g., $r > 0.4$, however, LPF starts to fragment more new lightpaths to improve link capacity utilizations, which pushes up the number of OEO conversions per request. For SC, on the other hand, it will not increase the transceiver consumptions per connection request even if such resources are redundant; thus its OEO conversions stay largely unchanged once it reaches its lowest value. For the light-tree based methods, since traffic grooming happens at a much lower probability, the number of OEO conversions per admitted request remains rather insensitive to the redundancy of the transceiver resources.

Figure 4.9 also reveals that while LPF may introduce a higher average number of OEO conversions per admitted request when $r > 0.65$, it significantly outperforms

the other methods by using roughly the same, or even fewer, OEO conversions per admitted request when $r < 0.65$. For example, putting Figure 4.3(c) and Figure 4.9 together, we see that when $r = 0.5$, LPF outperforms both SC and LTD-ANCG with fewer OEO conversions per admitted request.

FIGURE 4.10: Average number of transceiver pairs utilized all over the network, counted from the arrival of the 5,000-th connection request to the 10^5 -th connection request. Traffic loads are 250 Erlangs.

Figure 4.10 further shows that, compared to SC, LPF in fact does *not* consume more transceiver resources. Specifically, we compare the average number of transceiver pairs being utilized in the whole network within a period of time (Note that we do not compare the transceiver consumptions between lightpath based and light-tree based methods since a light-splitter in a light tree functionally acts as a one-to-many transceiver which, however, may request other resources such as more optical amplifiers.). As we can see, the total number of transceiver pairs consumed by LPF is actually slightly lower than that by SC within the whole range of r : the higher number of OEO conversions per admitted request does

not necessarily means a higher overall consumption of transceiver resources. The fragmentation of lightpath helps increase the efficiency of traffic grooming, and consequently keeps overall transceiver resource consumptions at a reasonably low level.

Note that in this study, we define the sole objective of LPF as improving the network BBR performance, without worrying about any possible drawbacks such as longer processing delay or higher transceiver consumptions etc. As a result, the algorithm tends to make full use of the transceiver resources when they are redundant. The algorithm can be easily revised to keep a balance between blocking performance and resource consumptions. For example, instead of measuring whether it helps to improve network blocking performance by fragmenting a lightpath on each of its intermediate nodes, we may revise the scheme to limit the number of fragmentations we could have for each lightpath (The upper bound of fragmentations we could have for a lightpath may depend on its hop length, the redundancy of its capacity resources, and/or something else.) by fragmenting only a few “most favorable” nodes along each lightpath. Such possible extensions, however, are out of the scope of this study and have to be discussed in our future work.

4.5 Summary

In this chapter, we studied dynamic multicast traffic grooming in meshed WDM networks with the main objective of minimizing the network bandwidth blocking ratio (BBR). We first compared the blocking performances of various existing lightpath and light-tree based approaches through extensive simulations, and showed that the lightpath based approaches generally outperform those light-tree based ones in different cases. Inspired by such observations, we then proposed the LightPath Fragmentation (LPF) method to further improve the network blocking

performances. Simulation results showed that LPF outperforms all the existing methods, mostly by at least an order of magnitude, with a modest cost of longer delay. The transceiver resource consumptions remain comparable with those of the best existing lightpath based methods. The effects of different factors including the average number of multicast destinations and the unicast traffic ratio were also evaluated. Results show that the LPF method significantly outperforms the existing methods with varying request destination numbers or unicast traffic ratio, and its superiority becomes even more obvious when the destination number is smaller or the unicast ratio is higher for its efficiency in resource sharing in grooming process.

It is worth noting that in this study, the lightpath and light-tree schemes are compared with respect to their performances in the dynamic multicast traffic grooming scenario. As for their BBR performances for serving wavelength-level multicast requests, further comparisons are needed as their performances strongly depend on how multicasting is implemented using lightpaths.

Chapter 5

Dynamic Multicast Traffic Grooming in Overlay IP/MPLS over WDM Networks

5.1 Introduction

As wavelength division multiplexing (WDM) network is playing the dominant role in the Internet backbone, it is widely believed that IP over WDM networks will be a key component of the next-generation Internet [18, 19]. The emerging networking technologies, such as Multi-Protocol Label Switching (MPLS), Generalized MPLS (GMPLS), User Network Interface (UNI), path computation element (PCE), etc., are also paving the way for such network revolution. For the interconnections between the IP/MPLS layer and the WDM layer networks, three architectural alternatives, as discussed in Chapter 2, have been proposed. Among these proposed architectures, the overlay model is deemed to be the most practical option for near-term deployment [38]. Extensive research has been done on

the overlay model, most of which has been focusing on handling dynamic unicast traffic [74, 76, 104].

In recent years, with the explosive growth of multicasting services, such as high-definition video distribution, online gaming, interactive distance learning, live-video conferencing, etc., it is widely believed that a large portion of the future Internet traffic will be multicast in nature. The bandwidth required for a typical multicast session is in the order of megabits per second (Mbps), which is much smaller compared to the 2.5 – 40 gigabits per second (Gbps) capacity that can be provided by a single wavelength channel in today's WDM networks. To efficiently utilize the wavelength capacity, multicast traffic grooming [67] is usually adopted to pack several multicast sessions onto wavelength channels for transmission. In recent years, dynamic multicast traffic grooming has become an important research issue, and various algorithms utilizing either lightpath, or light-tree, or both for dynamic multicast traffic grooming have been proposed [65, 78–83, 115].

Compared to peer model networks, dynamic multicast traffic grooming in overlay IP over WDM networks has not received much attention. This is because in such networks, the two layers of the network are managed by independent operators with very limited information exchanges between the networks, and thus the routing decisions made on one layer may lead to inefficient resource utilizations on the other layer. To the best of our knowledge, up till now only two methods have been proposed for tackling this problem. Both methods, which will be reviewed in Section 5.2, are easy to be implemented, yet not free from the inherent limits caused by limited information exchanges between the two layers.

Our previous study, which was presented in Chapter 3, on unicast traffic grooming in overlay networks shows that, by letting the two layers agree on a definition of the cost for setting up a new lightpath and allowing the IP/MPLS-layer operator to keep record of the recent service requests that have been supported by the WDM layer

network, the IP/MPLS-layer owner can make better routing decisions and improve network performance significantly [104]. To extend such results to multicast traffic grooming, however, requires nontrivial work. The issues to be studied include the definition of the cost for setting up new connections (not necessarily new lightpaths), the routing method, and more. Further, how to improve the efficiency of WDM-layer network resource sharing is also an important issue.

In this chapter, we address the dynamic multicast traffic grooming problem in overlay IP/MPLS over WDM networks. To help relax the constraint imposed by limited information exchanges in overlay networks while improving resource sharing in the traffic grooming process, an efficient overlay multicast provisioning (OMP) mechanism is proposed. By jointly utilizing a historical data learning (DL) scheme on the IP/MPLS layer for link cost estimation, and a lightpath fragmentation (LPF) based method on the WDM layer for improving resource sharing, OMP aims to minimize the bandwidth blocking ratio (BBR), which is defined by 3.1.

We perform extensive simulations to evaluate the performance of OMP mechanism in different cases, and compare OMP to the existing methods under different traffic loads, in networks with limited or unlimited optical port resources. To evaluate the respective influences of the DL scheme and the LPF method on OMP performance, we also devise some other provisioning mechanisms only utilizing the IP/MPLS layer DL scheme or the WDM layer LPF method, and compare them to OMP mechanism in different cases. Our results show that the IP-layer DL scheme contributes more to improve network performance than the WDM-layer LPF method. We also study the effects of other factors, including the definition of new connection cost and WDM-layer routing method, on OMP performances.

Below we first present the network model, the definition of the problem, and some most closely related existing results in Section 5.2, and then describe the proposed

OMP mechanism in Section 5.3; we evaluate the performance of OMP mechanism in Section 5.4, and conclude this chapter in Section 5.5.

5.2 Network Models, Previous Work and Problem Statement

5.2.1 Overlay IP/MPLS over WDM Network Model

An overlay IP/MPLS over WDM network is illustrated in Figure 3.1. With the overlay architecture, the operations and management of the two networks on different layers are independent of each other; the IP/MPLS layer is an integrated service provider (ISP) delivering the service requests between its end users while the WDM layer is the bandwidth provider providing the required connectivity services to its upper layer client(s).

In such an overlay network, the operator of each layer keeps all the information of its own layer, and sends its management commands to all its network elements via a centralized control system. Based on their service contracts, the two operators can also work cooperatively to provide the desired service fulfilling each arriving request. Specifically, upon the arrival of a multicast request, the IP-layer ISP first tries to find a route tree for it using only the existing logical links with sufficient residual bandwidth. If the effort fails, it will figure out the LSR pairs between which new lightpaths need to be set up, and query the WDM layer operator for the costs of setting up such lightpaths. After receiving the set up costs provided by the WDM layer, the IP-layer ISP will decide on the lightpaths to be purchased. Note that whether the WDM layer will be queried for the lightpath set up costs is decided by the IP layer operator while whether a new lightpath can indeed be set up or not is decided by the WDM layer operator. In the cost enquiry

process, the necessary information exchanges between the two layers are performed through well-defined network interfaces, i.e., UNIs, but not necessarily through the information exchange channels as shown in Figure 3.1.

In this study, we assume that there is only one ISP, and it has exact information of the IP/MPLS-layer network. We also assume that such IP-layer ISP can keep records of the lightpaths that have been supported by the WDM layer, their corresponding setup costs, as well as the time when such costs were reported. We extend the historical data learning (DL) scheme proposed in [104] from unicast to multicast case.

On the WDM layer network, we assume that the fixed minimum hop routing and first-fit wavelength assignment policy is adopted for lightpath routing. Note that a lightpath route could be very long and the long lightpaths may degrade resource sharing in traffic grooming process. As shown in our previous results that splitting long lightpaths into shorter ones helps improve resource sharing in the dynamic traffic grooming process [115], we assume that a lightpath fragmentation (LPF) based method is adopted in the lightpath routing process. With the LPF method, long lightpaths may be fragmented into shorter ones upon set up. We also assume that an established lightpath with ongoing transmission cannot be fragmented or rerouted.

Detailed description of the DL scheme and the LPF method will be presented in Section 5.3.

5.2.2 Node Architecture

A typical network node in an overlay IP/MPLS over WDM networks is an OXC interconnected with zero, one or more LSRs through an UNI [38]. By utilizing OXC, a node is able to transmit data traffic transparently from an input port to

an output port at the wavelength level granularity. Through the LSR(s), a node is also able to receive/transmit data traffic from/to the high-speed wavelength channels.

We assume that each network node in this study is an OXC interconnected with a single LSR, and all the OXCs have no wavelength conversion capability. The node architecture as shown in Figure 3.2 is assumed. For each node, the number of transmitters/receivers on it equals the number of add/drop ports on the OXC. Add/drop ports are typically expensive due to the need for high-speed electronics implementation. To save the network cost without sacrificing the network performance, an OXC is usually equipped with a limited number of add/drop ports that are to be shared by all the incoming/outgoing wavelengths. We use the *add/drop ratio* defined as below to represent the port resource on a network node:

$$r = N_P/N_W, (0 < r \leq 1) \quad (5.1)$$

where N_P is the number add/drop port pairs and N_W is the number of incoming/outgoing wavelengths the OXC has. If $r < 1$ for a node, we call it port-limited; else, it is port-unlimited.

5.2.3 Previous Work

To the best of our knowledge, only two methods have been proposed in literature for dynamic multicast traffic grooming in overlay IP/MPLS over WDM networks [79]. We term these methods as logical-path-tree (LPT) method and saturated cut (SC) method, respectively.

The main idea in these two methods is to utilize the IP layer's existing logical links to serve as many destinations as possible before setting up any new lightpaths. Specifically, LPT tries to find a route using existing logical links with enough

residual bandwidth for the request, and if such a step fails, it will then try to set up new lightpaths to connect the remaining destinations to the partial route found.

Compared to LPT, SC achieves better performance: it first identifies some islands, which contains either the source node s and nodes that can be reached from s , or at least a destination node d_i and those nodes which can reach d_i , using existing logical links with sufficient residual bandwidth, and then connect such islands using new lightpaths when necessary.

FIGURE 5.1: An example of serving multicast request $R\{s \rightarrow \{d_1, d_2, d_3\}\}$ with LPT and SC methods.

These methods are easy to implement, but has poor efficiency in resource utilization. An example is illustrated in Figure 5.1. Assume that at the time when multicast request $R\{s \rightarrow \{d_1, d_2, d_3\}\}$ arrives at the network, there are three logical links with enough residual bandwidth on the IP/MPLS layer, while on the WDM layer, all links have idle wavelengths except for the links BC and FE. When the LPT method is adopted, no route on the IP layer or the WDM layer can be found for this request. When the SC method is adopted, three islands as shown in Figure 5.1 can be found. However, although the two islands containing s and d_1

respectively can be connected via a new lightpath, the one with d_2 and d_3 cannot be connected to s ; thus the request has to be blocked.

By observing Figure 5.1 closely, it can be observed that if there is any chance to utilize the existing logical link between nodes n_1 and n_2 on the IP layer, the request in fact can be served. To handle such a situation, the OMP mechanism is proposed. Specifically, by estimating the cost of each IP layer logical link and then set up two new lightpaths from s to n_1 and n_2 to d_2 respectively, the OMP is able to fulfill the request.

5.2.4 Problem Statement

Denote the network as $G(V, E)$ with V and E being the sets of network nodes and fiber links respectively. Each WDM layer link consists of two fiber links in opposite directions, each of which carrying W wavelengths. A multicast request is represented as $R\{s, D, b\}$, where s , D and b are the request source, destination set and required bandwidth, respectively. A request is served only when all its destinations can be served; otherwise, it is blocked. Since our previous study showed that the lightpath scheme achieves better blocking performance over light-tree in dynamic traffic grooming process [115], we adopt the lightpath scheme to support multicast transmission in this study.

The dynamic traffic grooming problem in the overlay IP/MPLS over WDM networks can be defined as follows. Given an overlay IP over WDM network with certain network resources and dynamically arriving/leaving multicast requests, a request provisioning mechanism, which requires only limited information exchanges between the two different layers is to be devised to optimize the network BBR performance. For that purpose, network resource sharing must be enhanced on both layers as much as possible. As aforementioned, any established lightpath cannot be interrupted if it is still being used by any ongoing transmission.

5.3 OMP Mechanism for Dynamic Multicast Traffic Grooming

This section describes the proposed OMP mechanism. First, we present the historical data learning (DL) scheme for logical layer link cost estimation, followed by a description of the multiple tree heuristic (MTH) for finding a number of candidate route trees. Then we discuss the lightpath fragmentation (LPF) method for WDM layer routing. Finally, we present the OMP mechanism.

5.3.1 IP/MPLS layer historical data learning (DL) scheme

When requests arrive at the network, an auxiliary graph which represents the IP/MPLS layer network is generated. Edges of the graph consist of both existing logical links with sufficient residual bandwidth and potential new lightpaths to be set up. We call those existing lightpaths as *existing links*, and those new ones to be set up as *candidate new lightpaths* (CNLs).

Once the auxiliary graph is obtained, OMP assigns each graph edge an appropriate cost, and then finds a number of candidate routes. Hence, a candidate route for a request may consist of only existing links, or CNLs, or both. For simplicity, we further categorize those CNLs into *cost unknown* links and *cost enquired* links: if the cost of a link has never been reported by the WDM layer, it is a cost unknown link; otherwise, it is cost enquired. Below we briefly describe the link cost estimation process with the DL scheme.

As discussed in [104], for a cost unknown CNL between LSR i and j , it is reasonable for the WDM layer operator to provide the upper layer ISP a default link cost at the beginning of network operation. The default value calculated using the formula

below helps achieve satisfactory results:

$$M_{ij} = \left(\frac{1-r}{p \times r \times (\bar{H} + 1)} - H_{ij} \ln \left(1 - \frac{1}{\omega + 1} \right) \right) \times amp \quad (5.2)$$

where amp is an amplification factor; H_{ij} is the minimum number of optical hops between the two OXCs that are connected to LSRs i and j ; $\bar{H}$ is the average path length of the WDM network; $\omega = W/2$ is a representative value for the average number of idle wavelengths along a WDM link (As discussed in [104], network performance is not very sensitive to the value of ω); p is the average number of idle optical ports on a network node, and it can be calculated as

$$p = \max \left(W \times \delta \times r - \frac{1}{\bar{H} + 1} \times \delta \times (W - \omega), 1 \right) \quad (5.3)$$

where δ is the average nodal degree of the network.

For a cost enquired CNL, its cost can be estimated using the data learning scheme proposed in [104]. Specifically, for any request arriving at time T^m , the cost of a cost-enquired CNL can be estimated as follows,

$$C_{ij}^{est} = \begin{cases} C_{ij}^n - d_{ij} \Delta \left| C_{ij}^n - M_{ij} \right| \times \min \left(\frac{1}{\Delta}, \left\lfloor \frac{T^m - T^n}{\delta t} \right\rfloor \right), & \delta t \neq \infty \\ M_{ij}, & \delta t = \infty \end{cases} \quad (5.4)$$

and the expiration time of the above estimated cost can be calculated using the equation below,

$$T_{ij}^{cal} = \begin{cases} T_{ij} + \left\lfloor 1 + \frac{T^m - T^n}{\delta t} \right\rfloor \times \delta t, & C_{ij}^{est} \neq M_{ij} \\ \infty, & C_{ij}^{est} = M_{ij} \end{cases} \quad (5.5)$$

Each time a new request arrives at the network, the expiration time of all CNLs are compared to the request arriving time. If a link cost is regarded as being expired, its cost and expiration time will be updated by the new values, C_{ij}^{est} and T_{ij}^{cal} , respectively. Note that although some other schemes can also be devised for link cost estimation, we adopt this scheme for its simplicity. The results to be shown later in this chapter demonstrate that this simple scheme is able to achieve

satisfactory performance. Detailed derivation of the equation can be found in Chapter 3.

With the above described cost estimation process, the costs of the different types of auxiliary graph edges can be defined as follows,

$$L_{ij} = \begin{cases} 1 & \text{an existing logical link} \\ M_{ij} & \text{a cost unknown CNL} \\ C_{ij}^{est} & \text{a cost enquired CNL} \\ 2M_{ij} & \text{failed lightpath between LSR } i \text{ and } j \end{cases} \quad (5.6)$$

Once the costs of the auxiliary graph edges are known, a desired number of routes can be found using the appropriate multicast routing methods. Below we present the heuristic method we use in this study to find a desired number of logical trees for a multicast request.

5.3.2 Multiple tree heuristic (MTH) for IP/MPLS layer routing

Multicast traffic grooming is well-known to be an NP-complete problem, and heuristic methods, e.g., the *minimum cost path heuristic* (MPH) [56], are usually adopted for calculating multicast route. If MPH is directly adopted for multicast grooming, however, only a single tree can be found for a request. In an overlay network, the only tree found by MPH may not be good, or even feasible, for the request. We modify the MPH algorithm to find multiple candidate trees for an arriving request. We term the modified method as *multiple tree heuristic* (MTH).

The main idea of the MTH is to iteratively find one multicasting tree after another, until the required number of trees are found, or until no tree can be found. In

each iteration, for CNLs that are already included in multicasting trees found in earlier iteration(s), if any, we assign them with higher costs to discourage (but not strictly prevent) them from being used in the later multicasting trees again. Such an approach encourages MTH to include more CNLs in candidate trees and to enquire about their costs while avoiding the risk of missing some good candidate routes by strictly preventing CNLs from being included in multiple trees.

Procedure 5.1 summarizes the main steps of the MTH. We see that in Step 2, MTH adds all the existing logical links with sufficient residual bandwidth onto auxiliary graph; and then adds those CNLs that are involved in the previous trees and assign them costs in Steps 5 – 7; CNLs that are already in VL are assigned with higher costs. Step 8 assigns costs to the other edges of the auxiliary graph; Steps 9 – 13 find a logical tree for the current iteration. Note that MTH gives higher priority to using an existing logical tree: when a tree is found at the end of each iteration, MTH checks each of its links in Step 14. If the tree consists of only the existing logical links, it will be used to fulfill the request and the iterations stop; otherwise, the CNLs included in this tree will be recorded in a virtual link set. As aforementioned, these CNLs will be assigned with higher costs while being considered to be included in other trees in later iterations.

After the algorithm has been executed, MTH either returns a tree consisting of only the existing logical links, or a CNL set VL . The costs of CNLs included in VL are to be queried.

5.3.3 Lightpath fragmentation (LPF) for WDM layer routing

As discussed, if all the candidate IP/MPLS layer multicast trees found for a multicast request contain CNLs, the WDM layer operator needs to report the set up costs of such lightpaths to its IP layer counterpart based on their service contract.

Procedure 5.1: Multiple Tree Heuristic (MTH)**input** : The current network $G(V, E)$, request $R\{s, D, b\}$ and a number T .**output**: A CNL set, or a tree route for the request.

```

1 Clear the link set  $VL$  and auxiliary graph  $AG$ ;
2 Add all existing logical links with residual bandwidth  $\geq b$  onto  $AG$ ;
3 for tree number  $t_N = 1$  to  $t_N = T$  do
4   Node set  $S = \Phi$ ; add node  $s$  into  $S$ , let  $D$  be request's destination set;
5   for each candidate new lightpath  $VL_{ij}$  do
6     Add  $VL_{ij}$  onto  $AG$ ; if  $VL_{ij}$  is not in  $VL$ , set its link cost to be its estimated link cost;
7     otherwise, set its cost to be  $t_N$  times that of its estimated cost;
8   end
9   Assign costs to the other links on  $AG$  according to Eq. (5.6); compute all-to-all shortest
10  paths on  $AG$ ;
11  while  $D \neq \Phi$  do
12    Choose the minimum cost path  $P$  connecting a certain node in  $S$  to a certain node  $d$ 
13    in  $D$ ; add  $P$  onto tree route  $t$ ;
14    Check each link of  $P$ , if it is a CNL, add it into  $VL$ ;
15    Add all intermediate nodes along  $P$  into  $S$ ;  $D = D \setminus d$ ;
16  end
17  If set  $VL$  is empty, i.e., there exists a tree consisting of existing logical links only, save and
18  return the logical tree  $t$ ;
19 end
20 Return CNL set  $VL$ .

```

To fulfill such a purpose, a straightforward method is to adopt the lightpath cost definition used in [104], which is shown below:

$$C_{ij} = \begin{cases} \left(\frac{(1-r)}{p \times r \times (H+1)} - H_{ij} \ln \left(1 - \frac{1}{\omega_{ij}+1} \right) \right) \times amp, & \text{if } \omega_{ij} > 0 \text{ and } p > 0 \\ \infty, & \text{if } \omega_{ij} = 0 \text{ or } p = 0 \end{cases} \quad (5.7)$$

where all the parameters have the same meanings as those defined in Eq. (5.2) except for p and ω_{ij} . Here p is the smaller one among the number of transmitters at the source and the number of receivers at the destination of the lightpath; ω_{ij} is the number of idle wavelengths along the lightpath route. Such a definition tries to balance the consumptions of WDM layer wavelength and optical port resources: when both resources are abundant, the costs of consuming them should be low and not so different from each other; while if any of them becomes scarce, the cost of consuming it becomes higher.

After receiving the cost reported by the WDM layer, the IP layer ISP will recalculate the minimum-cost multicast tree and decide the lightpaths to be purchased. The WDM layer operator would then set up these lightpaths. Note that

some lightpaths may be long, which may degrade the utilization of WDM layer resources. To improve resource sharing in the grooming process, a lightpath fragmentation (LPF) method [115] is adopted in the WDM layer lightpath routing process. We will briefly describe the LPF method below.

Suppose n_i is an intermediate node along a new lightpath L that the IP layer operator wants to order. Denote the fan-out degree of n_i as d_i ; and the numbers of transmitters and receivers on n_i as T_i and R_i , respectively. To determine whether L should be fragmented at node n_i , two parameters are defined as follows,

$$\alpha_m = \min \left(\frac{T_i}{d_i \times W_{out}}, \frac{R_i}{d_i \times W_{in}} \right) \quad (5.8)$$

$$\alpha = \frac{1}{H_i} \quad (5.9)$$

where W_{out} and W_{in} are the numbers of idle wavelengths on the incoming and outgoing links that lightpath L goes through respectively, and H_i is the average number of optical hops from n_i to the other OXCs along the shortest paths on the WDM layer.

To determine whether a lightpath L should be fragmented at a node n_i , the main idea of LPF is to estimate whether the wavelength or the transceiver resources at n_i are more limited. A lightpath is fragmented at n_i only if the wavelength resources are regarded as more limited. Specifically, while α_m reflects the smaller one among the currently available add and drop ratios at n_i , α is the add/drop ratio required for n_i to support lightpaths from itself to each of the other nodes to be initiated from it. Thus, when $\alpha_m \geq \alpha$, we regard the transceiver resources as being more redundant and let the lightpath L be fragmented at n_i ; otherwise, we let L bypass n_i to avoid taxing on the limited transceiver resources. A more detailed discussion can be found in Chapter 4.

The key steps of the LPF method are as follows.

Procedure 5.2: Lightpath-fragmentation (LPF) method

input : A network $G(V, E)$, a lightpath L
output: A set of new lightpaths.

```

1 while any node of  $L$  has not been checked do
2   for each intermediate node (if any)  $n_i$  along  $L$  do
3     Calculate  $\alpha_m$  for  $n_i$  ;
4     if  $\alpha_m > \alpha$  at  $n_i$  then
5       Fragment  $L$  at  $n_i$ , and get two new lightpaths  $L_a$  and  $L_b$ ;
6        $L = L_b$ ;
7     end
8   end
9 end

```

When a new lightpath is ordered, the LPF method is used by the WDM layer to process the request, and a lightpath may be fragmented into several segments. However, note that Eq. (5.7) does not take into account the possibility of lightpath fragmentation when calculating the cost of a lightpath. This helps to simplify the calculation and keep the fragmentation operation, if any, to be transparent to the IP layer operator.

Equation (5.7) can be easily revised to take into account the effects of lightpath fragmentation in lightpath cost calculations. One possible way is to calculate the default link cost and the new lightpath set up cost as follows:

$$M_{ij} = \left(\sum_{seg} \left(\frac{1-r}{p \times r \times (\bar{H} + 1)} - H_{seg} \ln \left(1 - \frac{1}{\omega + 1} \right) \right) \right) \times amp \quad (5.10)$$

$$C_{ij} = \left(\sum_{seg} \left(\frac{1-r}{p_{seg} \times r \times (\bar{H} + 1)} - H_{seg} \ln \left(1 - \frac{1}{\omega_{seg} + 1} \right) \right) \right) \times amp \quad (5.11)$$

where $\omega = W/2$ which, as discussed in Section 5.3.1, is a representative value of the average number of idle wavelengths along a WDM link; p is the number of idle optical ports; p_{seg} is the smaller one among the number of transmitters at the source and the number of receivers at the destination of a segment after fragmentation; and H_{seg} and ω_{seg} are the hop length and the number of idle wavelengths along the new lightpath, respectively.

Equation (5.11) defines the cost of a new lightpath as the sum of all fragmented new lightpath segments. We call the CNL cost defined in Eq.(5.7) as a *rough* report, and the cost in Eq.(5.11) as an *accurate* report. The effect of using these two different definitions on the OMP performance will be evaluated in Section 5.4.

5.3.4 Overlay Multicast Provisioning (OMP) Mechanism

OMP utilizes the IP layer DL scheme for logical link cost estimation and the WDM layer LPF method for improving resource sharing. The key steps of the OMP method are presented below as Algorithm 5.1.

Algorithm 5.1: OMP for Dynamic Multicast Traffic Grooming

input : A network $G(V, E)$ and multicast request $R\{s, D, b\}$.

output: A tree route to serve $R\{s, D, b\}$.

- 1 Update the costs of all CNLs of which estimated costs are expired;
 - 2 Call Procedure 5.1; // *logical-layer grooming*
 - 3 If set VL is empty, go to Step 13; otherwise, continue;
 - 4 **for** each link in set VL **do**
 - 5 Enquire the WDM layer for the set up cost of the link;
 - 6 Update the IP/MPLS layer cost record for the link;
 - 7 **end**
 - 8 Based on the enquired link costs, run minimum cost path heuristic (MPH) algorithm again on the IP layer to find a logical tree t for the request; if any request destination cannot be connected, block the request, return;
 - 9 **for** each CNL on tree route t **do**
 - 10 Call Procedure 5.2, and return a set of new lightpath routes;
 - 11 For each new lightpath, allocate both wavelength and port resources;
 - 12 **end**
 - 13 Serve the request; update the IP/MPLS layer network status;
-

Steps 1 – 2 generate the logical layer auxiliary graph, and find a desired number of logical layer candidate trees for the request using the MTH heuristic; if there exists one logical tree consisting of existing logical links only, OMP uses this tree to serve the request in Step 3. If no such tree exists, however, the IP layer ISP will query the WDM layer operator for the costs of all CNLs in VL . Based on the lightpath setup costs reported by the WDM layer, OMP runs the MPH algorithm once again at the logical layer to find one logical tree for the request in Step 8;

Steps 9–12 fragment the new lightpaths to improve resource sharing, and establish them after processing. Finally, Step 13 updates the network status.

Note that once the set up cost of a CNL is reported by the WDM layer network, its upper layer cost record will be updated accordingly.

Finally, we have a brief discussion on the complexities of the heuristic algorithms. Both LPT and SC adopt the MPH algorithm to find the multicast tree for a request [79]. Their complexities can be calculated as $O(|D| |V|^2)$ and $O(|D(D+2)| |V|^2)$, where $|D|$ and $|V|$ denote the numbers of multicast destinations and network nodes, respectively. The OMP method also adopts the MPH algorithm to find the multicast trees. Since it firstly finds K candidate trees and then finds among them the one with the minimum cost, its complexity can be calculated as $O((K+1)|D| |V|^2)$. Note that OMP also requires storage space for recording the historical data, the complexity of which is $O(|V|^2)$. Overall, we see that the complexity of the OMP method remains at a reasonably low level.

5.4 Simulation Results and Discussions

Extensive simulations have been carried out to evaluate the performance of the OMP mechanism in different cases. Below we will present the simulation environment and performance metrics. Then we will compare the performance of OMP with *rough* report against that of an existing algorithm. We will also evaluate the influences of both the IP layer DL and WDM layer LPF methods on OMP performance, respectively. Finally, we will assess the influences of WDM layer lightpath cost report (*rough* vs. *accurate*) and WDM layer routing methods (fixed vs. dynamic shortest path) on OMP performance.

5.4.1 Simulation Environment and Performance Metrics

Two commonly used network topologies are used in our simulations. As shown in Figure 5.2, they are 14-node, 21-link NSFnet and 24-node, 43-link USnet topologies, respectively. NSFnet has an average nodal degree of 3 and an average shortest path length of 2.18. For USnet, the two parameters are 3.58 and 2.99, respectively.

FIGURE 5.2: The network topologies utilized for simulations. (a) The 14-node 21-link NSFnet. (b) The 24-node 43-link USnet topology.

The following are some assumptions adopted in the simulations:

- 1) Each fiber link carries $W = 32$ wavelengths, the capacity of each wavelength channel is $B = 16$ units;
- 2) Requests arrive/leave the network dynamically as a Poisson process with a mean rate λ ; their holding time follows a negative exponential distribution with mean $\mu = 1$; bandwidth requirement of each request is an integer uniformly distributed in $[1, 16]$;
- 3) Source and destination nodes of all requests are randomly chosen among all network nodes; the number of destination nodes of each request is an integer uniformly distributed in $[2, 4]$ for NSFnet, and in $[2, 7]$ for USnet;

- 4) The cost of utilizing a new lightpath is about 5 times [108] that of using an existing logical link; thus, the parameter *amp* is set to be 40 and 25 for NSFnet and USnet, respectively;
- (5) For the IP/MPLS layer historical data learning scheme, the parameters are the same as those adopted in [104].

The BBR Performance of OMP is compared to that of the existing saturated cut (SC) method proposed in [79]. Results shown in each figure are averaged from at least five independent simulation experiments, each of which running 10^5 requests or more. Since all the conclusions hold for both topologies, unless otherwise stated, we present only the results on NSFnet for comparisons and discussions.

5.4.2 Performance Comparisons between OMP and SC Method in Different Cases

5.4.2.1 Comparisons under different traffic loads

We compare OMP and SC methods in networks with either limited or unlimited optical ports under different traffic loads. As can be seen in Figure 5.3, OMP outperforms SC within the whole range of traffic loads in port-unlimited NSFnet. Specifically, when under low traffic loads, e.g., around 450 Erlangs, OMP outperforms SC by more than two orders of magnitude; while under higher traffic loads, e.g., around 600 Erlangs, OMP still outperforms SC by about 50%.

Figure 5.4 compares OMP with SC in port-limited NSFnet topology where $r = 0.6$ for all OXCs, under different traffic loads. As we can see, OMP again significantly outperforms SC under different traffic loads: when under low traffic loads, e.g., $\rho = 310$ Erlangs, OMP outperforms SC by more than an order of magnitude, while

FIGURE 5.3: Comparison between OMP and SC methods under different traffic loads in port-unlimited NSFnet ($r = 1.0$).

when under heavier traffic loads, e.g., $\rho = 360$ Erlangs, OMP outperforms SC by more than 50%.

Together, Figures 5.3 and 5.4 convincingly demonstrate satisfactory BBR performance of OMP in overlay IP/MPLS over WDM networks. Such performance is due to the combined contribution of the IP/MPLS layer DL scheme and the WDM layer LPF method. Individual contributions from each of them will be further evaluated in Sections 5.4.3 and 5.4.4, respectively.

Another interesting observation in Figures 5.3 and 5.4 is that having a larger number of logical layer candidate routes does not always lead to significant improvement in the BBR performance. Such an observation is different from that

FIGURE 5.4: Comparison between OMP and SC in port-limited NSFnet network under different traffic loads ($r = 0.6$).

for dynamic LSP routing as reported in Chapter 3, wherein the network performance improves with an increasing number of candidate routes. Many reasons contribute to this observation, and the main one among them is that the logical layer link cost estimation process is a reasonably good choice for route selection for the connection request, even when we try to find only one single candidate route. Specifically, our simulation results show that the first candidate route found by MTH has a high chance ($\geq 95\%$) to be chosen as the final route for the request.

In what follows, we will evaluate the impact of optical port resource availability on OMP performance.

5.4.2.2 Comparisons in networks with different port resources

Figure 5.5 compares OMP versus SC with different add/drop ratios. The traffic load is fixed at $\rho = 300$ Erlangs. As can be seen, when the add/drop port resource is too limited, e.g., $r < 0.4$, there is no obvious winner between the two methods; once the add/drop ratio is large enough, e.g., $r \geq 0.5$, OMP significantly outperforms SC. Specifically, when $r \geq 0.6$, OMP outperforms SC by more than an order of magnitude. Such observation can be explained as follows: when the port resource is too limited, different algorithms would not make a significant difference; once the port resource has some slack, the scheme that is capable of utilizing network resources more efficiently will shine. Note that, when there is slack in the port resource, the BBR performance of SC does not improve further while the performance of OMP steadily improves with an increase in the add/drop ratio.

It is worth noting that, in Figure 5.5, the OMP does not make significant improvements with an increase in the number of IP/MPLS layer candidate routes, due to the same reason as explained earlier.

To figure out whether OMP leads to too many intermediate OEO conversions for each connection request, which is not desirable since extensive OEO conversions may lower the transmission speed while increasing transmission cost, Figure 5.6 compares the average number of intermediate OEO conversions experienced by each multicast request for both the SC and OMP methods. As can be seen, when $r < 0.6$, the average number of OEO conversions experienced by each request decreases with an increasing value of r in the SC method, while when $r > 0.6$, this number stays almost unchanged. The observations however are very different for the OMP method: when $r < 0.2$, the average number decreases with an increasing value of r ; for $r > 0.2$, the average number increases with r . Specifically, for $r < 0.45$, a request served by OMP usually experiences a smaller number of OEO conversions, while for $r > 0.5$, OMP has a higher number of OEO conversions for

FIGURE 5.5: Comparison between OMP and SC versus add/drop ratio in NSFnet network under traffic load $\rho = 300$ Erlangs.

each connection request. The highest value of about 2.6, however, appears to be acceptable for most applications.

Such observations can be explained as follows: when the port resource is limited, e.g., $r < 0.2$, only a few short lightpaths can be set up between each LSR pair, and most of the admitted requests are served by these lightpaths, which leads to a larger average number of intermediate OEO conversions for both methods. With more add/drop port resource, more end-to-end lightpaths can be set up between each LSR pair, intermediate OEO conversions hence will reduce for both algorithms. For SC, once the number of intermediate OEO conversions reaches its lowest value, it will not be changed further. For OMP, however, since the algorithm

FIGURE 5.6: Average number of OEO conversions experienced by each multi-cast request served with SC and OMP ($\rho = 300$ Erlangs).

is designed to make the best use of the excess resource to improve the network BBR performance as much as possible, some new lightpaths will be fragmented, which leads to an increased number of intermediate OEO conversions.

Putting Figure 5.5 and Figure 5.6 together, we see that for a moderate add/drop ratio of $r = 0.6$, the OMP methods, by increasing the average number of intermediate OEO conversions for about 13% (from 1.81 to an acceptable value of 2.04), improves the BBR performance to be more than an order of magnitude better than that of the SC method.

Since increasing the number of candidate routes seldom leads to any significant improvements in BBR performance, hereafter we will only present the results

obtained with a single logical layer candidate route for each connection request.

5.4.3 Influences of IP/MPLS Layer Data-Learning (DL) Scheme

In this section, we evaluate the impact of IP/MPLS layer historical data learning (DL) scheme on OMP performance. For comparison purpose, we devise an “OMP without data learning” (OMP_No_DL) method. Specifically, the method is nearly the same as the OMP method except that for the IP layer auxiliary graph edge cost assignment, instead of using the DL scheme, it assigns a cost of 1 to using existing logical links and a cost of 5 to using CNLs.

Figure 5.7 compares OMP_No_DL against SC and OMP in port-limited NSFnet under different traffic loads. Results show that without the IP layer DL scheme, OMP_No_DL performs the worst for the full range of traffic loads: SC outperforms OMP_No_DL by more than 60% on average; while OMP is more than an order of magnitude better under light traffic loads, e.g., when $\rho = 310$ Erlangs, and about 80% better under heavy traffic loads, e.g., when $\rho = 370$ Erlangs.

To further demonstrate the significant effects of the DL scheme, Figure 5.8 compares OMP_No_DL against SC and OMP in NSFnet with different port resource. Results show that when the port resource is too limited, e.g., $r \leq 0.3$, there is no obvious winner among the three methods; when $r > 0.3$, however, OMP_No_DL again performs the worst. Specifically, OMP outperforms OMP_No_DL by more than one order when $r > 0.55$, while SC outperforms OMP_No_DL once $r > 0.4$.

The above comparisons clearly illustrate the major impacts of the IP layer DL scheme on network BBR performance: by estimating the cost of each logical link

FIGURE 5.7: Performance of OMP_No_DL compared to OMP and SC methods under different traffic loads in port-limited NSFnet ($r = 0.6$).

using historical data, the DL scheme helps choose the right route for each incoming request, which improves the BBR performance by one, to several orders of magnitude.

5.4.4 Influences of WDM Layer Lightpath Fragmentation (LPF) Method

To evaluate the effect of the WDM layer LPF method on OMP performance, we devise an “OMP without LPF method” (OMP_No_LPF), which is nearly the same as OMP but not using the LPF method on the optical layer.

FIGURE 5.8: Performance of OMP_No_DL compared to OMP and SC in NSFnet with different optical port resources ($\rho = 300$ Erlangs).

Figure 5.9 compares OMP_No_LPF against OMP and SC under different traffic loads in port-limited NSFnet where $r = 0.6$ for all OXCs. As can be seen, OMP_No_LPF outperforms SC within the full range of traffic loads. But it performs nearly the same as OMP under moderate and high traffic loads; it is only outperformed by OMP when light traffic loads. Such an observation can be explained as follows: under light traffic loads, most connection requests can be served using the existing lightpaths. Lightpath fragmentation, which enhances wavelength resource sharing, leads to better performance. Under heavy traffic loads, more lightpaths need to be set up. The limited port resource soon gets exhausted, mainly because they have been used to support the new end-to-end lightpaths.

FIGURE 5.9: Performance of OMP_No.LPF is compared to OMP and SC under different traffic loads in port-limited NSFnet ($r = 0.6$)

Figure 5.10 further compares OMP_No.LPF against SC and OMP in NSFnet with different port resources. As can be seen, when the port resources are limited, e.g., $r < 0.4$, the three methods perform nearly the same; while when the port resources are abundant, OMP_No.LPF and OMP outperform SC. More port resources also lead to bigger differences between the performances of OMP and OMP_No.LPF. The results again demonstrate that LPF helps to improve network performance, especially when port resources are abundant.

FIGURE 5.10: Performance of OMP_No_LPF is compared to OMP and SC schemes in NSFnet with different add/drop ports ($\rho = 300$ Erlangs).

5.4.5 Influences of WDM Layer Lightpath Set up Cost

Definition

In the earlier subsections, Eq. (5.7) was used to define the cost for setting up a new lightpath. As discussed, such a definition does not take into account the lightpath fragmentation effect. It would be of research interest to figure out the impacts on BBR performance when Eqs. (5.2) and (5.7) are replaced by Eqs. (5.10) and (5.11) respectively in order to reflect the lightpath segmentation on WDM layer.

Figure 5.11 compares OMP (with rough and accurate reported lightpath costs) against SC in NSFnet network with different optical port resources. As discussed earlier, when port resources are abundant, e.g., when $r > 0.4$, OMP with either

rough or accurate reported link cost performs much better than SC. As to the performances of OMP with two definitions of lightpath cost, we observe they are quite similar to each other. Specifically, OMP with accurate reports only marginally outperforms OMP with rough report when $r > 0.5$. Such an observation is not difficult to be understood: when port resources are too limited, few lightpaths are fragmented; hence rough and accurate reports report roughly the same value; while as port resources increase, more lightpaths are fragmented, accurate reports thus give more accurate cost information. However, since even under such a case the fragmented lightpaths count for only a small fraction of all lightpaths, the performance differences is insignificant.

FIGURE 5.11: Performance of OMP (with rough and accurate reported link costs) compared to SC in NSFnet with different port resources ($\rho = 300$ Erlangs).

To verify the above observations, Figure 5.12 shows the fragmentation ratio, which is defined as the number of fragmented lightpaths versus the total number of lightpaths. As can be seen, when add/drop ratio $r < 0.4$, virtually no lightpath is fragmented. Therefore, the performances of OMP with rough and accurate lightpath costs are nearly the same. When the add/drop ratio $r > 0.4$, though some new lightpaths can be fragmented, i.e., $\alpha_m \geq \alpha$ on a certain node along a new lightpath, the fragmentation ratio is quite low; hence the differences in BBR performances based on two different cost definitions remain to be insignificant.

FIGURE 5.12: The ratio of new lightpath that are fragmented to the total number of new lightpaths in NSFnet with different port resources ($\rho = 300$ Erlangs).

Similar observations hold for the USnet topology: The fragmentation ratios are 0.3% and 1.6% for add/drop ratio $r = 0.5$ and $r = 1.0$ respectively under a traffic

load of $\rho = 200$ Erlangs. The two different cost definitions therefore do not lead to significant differences in BBR performance.

Note that the above results are obtained when the fixed minimum hop routing and first-fit wavelength assignment policies are adopted on the WDM layer. We have also tested the case of adopting dynamic minimum cost path routing and first-fit wavelength assignment, and found that the above conclusions hold.

5.4.6 Influences of the WDM Layer Routing Strategies

In this section, we evaluate the influences of WDM layer RWA policies on OMP performance.

For comparison purpose, we devise a new scheme which adopts the same IP layer routing method as that of OMP, and yet uses the dynamic minimum-cost path routing and first-fit wavelength assignment policies on the WDM layer. We call it the OMP (dynamic) method. To further assess the influences of LPF on OMP performance, OMP (dynamic) without LPF method is also included in the comparison. Note that for OMP (dynamic), the queried cost of a CNL is the cost of the dynamic shortest path calculated in the WDM layer network. Also note that we use Eq. (5.7), i.e., rough report of lightpath cost, to define the lightpath set up cost since, the two different definitions of lightpath cost lead to nearly the same performance.

Figure 5.13 compares OMP with different routing strategies against SC in NSFnet with different port resources. As can be seen, with an increasing add/drop ratio, OMP with either dynamic or static WDM layer routing method outperforms SC within the full range of add/drop ratio; while for OMP, the results show that the performances for different routing methods are nearly the same when $r < 0.5$; when $r > 0.5$, OMP with dynamic RWA starts to perform better. The results

FIGURE 5.13: Performances of OMP with different WDM layer routing strategies compared to SC in NSFnet with different port resources ($\rho = 300$ Erlangs).

are reasonable and can be explained as follows: when the port resources are the resource bottleneck, OMP with either dynamic or static RWA scheme, though different in their capabilities of exploring wavelength resources, would not lead to significantly different performances. With more abundant port resources, OMP with dynamic WDM layer RWA scheme is able to find more appropriate lightpaths for a request, and consequently leads to better performance.

Figure 5.13 also shows that LPF on WDM layer leads to more significant improvements when the dynamic RWA policy is adopted and port resources are abundant ($r > 0.6$).

5.5 Summary

In this chapter, we studied the dynamic multicast traffic grooming problem in overlay IP/MPLS over WDM networks. An efficient overlay multicast provisioning (OMP) mechanism which jointly utilizes an IP/MPLS layer historical data learning (DL) scheme and a WDM layer lightpath fragmentation (LPF) based method was proposed. Simulation results demonstrated that OMP significantly outperforms the existing methods under different traffic loads, in networks with limited or unlimited optical port resources. We assessed the respective influences of DL and LPF methods on OMP performances, and showed that both DL and LPF method help improve the OMP performance, and contributions by the DL scheme are much more significant. Influences of the different definitions of WDM layer lightpath cost and different WDM layer routing strategies on OMP performance have also been evaluated. Results show that when the port resources are limited, neither different lightpath definitions nor WDM layer routing strategies lead to significantly different performances; whereas when port resources are abundant, both factors strongly affect the network BBR performance. Between them, the improvements achieved by adopting a more efficient RWA scheme appear to be more significant than those achieved by having accurate lightpath cost information.

Chapter 6

Multicast Request Protection in Dynamic Traffic Grooming

6.1 Introduction

As bandwidth intensive multicast applications, such as high-definition-television (HDTV), interactive distant learning, live-video conferencing, etc., become increasingly popular, it is widely believed that a large portion of future Internet traffic will be multicast in nature. However, because optical networks are vulnerable to various component failures, a single failure may cause massive information loss and serious service interruptions for the end users. Hence, survivability mechanisms are essential to protect multicast sessions against network failures. As reviewed in Chapter 2, various mechanisms have been proposed for wavelength-level multicast request protection in WDM networks [34–36, 88–91].

Compared to the extensive research efforts directed toward wavelength-level multicast request protection, subwavelength level multicast request protection, which is also known as survivable multicast traffic grooming (SMTG), has received relatively little attention in recent years. Specifically, although two methods of addressing

such problems, which will be reviewed in Section 6.2.3, have been proposed in [100], some assumptions adopted therein may not necessarily be valid in modern optical communication networks (as discussed in Chapter 2). Therefore, we do not adopt these assumptions in our proposed scheme.

In this chapter, we address the problem of protecting subwavelength level multicast requests through dynamic traffic grooming. A novel mechanism, named the lightpath-fragmentation (LF)-based segment shared protection (LF-SSP) scheme, is proposed to protect multicast requests at the connection level. The primary objective of the design of the algorithm is to protect requests against any single link failure while minimizing the network's bandwidth-blocking ratio (BBR), which is defined by 3.1. By adopting the LF method to fragment new lightpaths into shorter segments to improve resource sharing for both grooming and request protection, the LF-SSP scheme attempts to minimize the overall network resources allocated to protect each request. We perform extensive simulations to evaluate the performance of LF-SSP. We first compare LF-SSP to an existing method for sub-wavelength level multicast request protection, and then extend the comparison to some existing methods for wavelength-level request protection. The results indicate that LF-SSP outperforms the existing methods in certain cases when the network resources are not overly limited. In addition, the effects of some other factors, including the add/drop port resources and the average number of multicast request destinations, are also studied.

The remainder of this chapter is organized as follows. Section 6.2 presents the network model, related work and the problem to be addressed. Section 6.3 describes the proposed LF-SSP scheme for dynamic SMTG. The simulation results are presented and discussed in Section 6.4. Section 6.5 concludes the chapter.

6.2 Network Model and Problem Definition

6.2.1 Network Model

We consider multicast traffic grooming in wavelength-routed WDM networks. The network can be represented by a directed graph $G = (V, E)$, where V and E are the sets of nodes and fiber links, respectively. Specifically, we assume that the physical-layer topology of the network is a set of nodes interconnected by fiber links. Each fiber link is composed of two fibers travelling in opposite directions, each of which carries W wavelengths. The capacity of each wavelength is B units. Each network node is equipped with a grooming-capable optical cross-connect (GC-OXC) [67] as shown in Figure 2.9(a). The GC-OXC is equipped with some number of add/drop ports, which is generally equal to the number of transceiver pairs on the node. Because both add/drop ports and transmitters/receivers are high-cost components, a network node is usually equipped with a limited number of ports that are shared by all wavelengths passing through the node. In this chapter, we define the *add/drop ratio* r ($0 < r \leq 1$) as the ratio of the number of add/drop port pairs over the total number of wavelengths passing through the OXC. We refer to a network node with $r < 1$ as a port-limited node; otherwise, it is a port-unlimited node.

To support multicasting services, a lightpath-based scheme is adopted in this chapter: as demonstrated in our previous study, lightpath-based approaches consistently outperform the light-tree based ones in achieving better bandwidth-blocking performance for dynamic multicast traffic grooming [115]. A lightpath occupies a wavelength along its route, a transmitter at its source node, and a receiver at its destination node, whereas a multicast request may traverse several lightpaths along its route and consumes a portion of the bandwidth provided by each lightpath through which it passes.

6.2.2 Problem Statement

The dynamic SMTG problem in WDM networks can be formulated as follows. Upon the arrival of each multicast request $R\{s, D, b\}$, where s , D and b denote the request source, specified destination and required bandwidth respectively, the central network controller must identify both primary and backup routes for the request using accurate global information regarding the network state so that the request will be able to survive any single link failure. The primary objective of the design of the algorithm is to minimize the network BBR by improving resource sharing in both the traffic-grooming and request-protection processes.

We assume that all multicast requests arrive/leave the network dynamically with no prior information regarding future requests and that a request is supported only when all its destinations can be served; otherwise, the request will be blocked. All requests are protected at the connection level, i.e., the connection to each request destination shall pass through a certain number of *survivable* lightpaths. A lightpath is “survivable” if it is protected by a link-disjoint backup path, which may pass through one or multiple lightpaths with sufficient bandwidth reserved for backup purposes. The backup paths for different survivable lightpaths can share capacities when their corresponding primary paths are link-disjoint to each other. It has been shown in [116] that grooming primary and backup paths separately helps to improve network performance. In this chapter, we refer to lightpaths that are utilized on working paths as *working lightpaths* and those that are used solely for request protection as *backup lightpaths*.

To be more precise, in this study, we assume that each fiber link comprises of two fibers travelling in opposite directions, and that a single network link failure will sever the fibers in both of the two opposite directions. We also assume that an ongoing transmission cannot be interrupted until it is complete.

6.2.3 Related Existing Work

The problem of wavelength-level multicast request protection has been extensively studied, and various link-, path- and segment-based protection schemes have been proposed [34–36, 88–91]. Among them, the segment-based protection schemes, as discussed in Section 2.4, are believed to be most efficient for dynamic multicast traffic protection; hence, a segment-based protection scheme is considered in this chapter.

The first segment-based protection method was proposed in [88]. By protecting each segment on the primary tree using a link-disjoint path from the segment source to its end node, the proposed segment-based method is able to protect a multicast session against any single link failure. By adopting a similar idea, the authors of [35] have proposed a different segment-based mechanism known as the Segment-based Protection Tree (SPT). In SPT approach, each primary tree is divided into tree-segments, each of which is then protected using a link-disjoint tree, which connects the session source and all its destinations. The minimum-cost survivable topology, which accounts for the costs for both the primary and backup trees, is chosen to fulfill the connection request. It has been demonstrated that the SPT mechanism outperforms the best existing path-based method in certain cases.

More recently, another new segment-based protection method known as level protection (LP) has been proposed in [36]. Once a primary tree has been found, the LP scheme attempts to protect the session destinations one by one in ascending order of their distances from the request source, with the objective of efficiently sharing resources on both the primary and backup trees. The LP scheme has been demonstrated to achieve far superior performance with respect to the algorithms proposed in [34, 89, 90].

To date, the problem of subwavelength-level dynamic SMTG has received little attention. Specifically, the authors of [100] have proposed the first two algorithms to attempt to address this problem, namely, Multicast Traffic Grooming with Segment Protection (MTG-SP) and Multicast Traffic Grooming with Shared Segment Protection (MTG-SSP). As discussed in Section 2.4, because the assumptions adopted by these algorithms are not necessarily valid in modern optical-network infrastructures and yet, the implementations of these methods strongly rely on these assumptions, we do not consider these methods for comparison in this chapter.

In [117], we proposed a scheme known as connection-level segment shared protection (CL-SSP) for SMTG. For each multicast request, CL-SSP attempts to protect the session against any single link failure at the connection level. To improve resource sharing, CL-SSP adopts the simple approach of splitting new primary/backup lightpaths into shorter segments on intermediate nodes that have redundant transceiver resources. Our results indicated that such a simple approach, although very effective for supporting multicast transmissions without protection [115], may not easily achieve satisfactory performance for supporting survivable dynamic traffic grooming. In fact, among the several methods we studied, the best performance was achieved by fragmenting only the backup lightpaths. A more careful approach to lightpath fragmentation is unquestionably needed.

In this study, we propose a new LF-SSP method in which the lightpath fragmentation of working and backup lightpaths is considered simultaneously to achieve better BBR performance. The BBR performance of the new method is compared against that of the best existing wavelength-level schemes, including the SPT and LP methods, and that of the subwavelength-level CL-SSP method in which only backup lightpaths are split. For simplicity, we henceforth refer to CL-SSP with only backup-path splitting simply as the CL-SSP method.

6.3 LF-SSP Mechanism for Survivable Dynamic Multicast Traffic Grooming

This section presents the proposed LF-SSP mechanism in detail. We first briefly describe the main underlying concept of the segment shared protection (SSP) method for survivable dynamic multicast traffic grooming, and then discuss the major steps of the proposed two-phase LF-SSP method. Finally, the detailed LF-SSP mechanism is presented.

6.3.1 Segment Shared Protection (SSP) Method

Segment shared protection is an efficient mechanism that has been utilized to support survivable unicast [116, 118, 119] and multicast transmissions [35, 36, 88]. For unicast services, the basic concept of SSP is to split the working path into a few segments and protect each of them separately. It has been demonstrated in the literature that segment protection offers better blocking performance than path- or link- protection methods [116, 118, 119]. The example presented in Fig. 6.1 illustrates the efficiency of the SSP scheme over the others in resource sharing. If path-protection scheme is adopted, as shown in Fig. 6.1(a), the backup-path capacity for the lightpath from s to d cannot be shared, whereas if a segment-protection scheme is adopted, resource sharing of the backup path is permitted as shown in Fig. 6.1(b).

To support survivable multicast requests, SSP also attempts to fragment a multicast tree into segments, e.g., a number of paths or smaller trees, and then protects each segment separately using a link-disjoint topology. Meanwhile, to improve resource utilization, the backup resources for different segments can be shared. Typically, there are two approaches to backup-resource sharing : *self-sharing* and *cross-sharing* [120]. Self-sharing refers to the sharing of resources among backup

FIGURE 6.1: A comparison between the path-shared protection and SSP schemes against single link failure. (a) Path shared protection. (b) SSP scheme for traffic grooming.

routes within the same session, and cross-sharing refers to the sharing of backup resources among different sessions.

In this study, both self- and cross-sharing are considered in the proposed SSP scheme for survivable multicast traffic grooming. Specifically, to identify the sharing potential of the backup paths, each backup lightpath is associated with a link set $\mathbf{A}$. Consider a backup lightpath e : its associated link set can be described as $\mathbf{A} = \{a | a \in E, 0 < b_e^a \leq B\}$, where b_e^a represents the amount of capacity that would be rerouted to the backup lightpath e if its protected link a fails. Hence, the total amount of capacity that has been reserved for protection on e is $b_t = \max_{a \in \mathbf{A}} b_e^a$, and its residual capacity is $r_m = B - b_t$; and the capacity along e that can be shared by the backup path of any new primary lightpath l , which is link-disjoint to e , is $b_f^l(L) = b_t - \max_{a \in (L \cap \mathbf{A})} b_e^a$, where L denotes the set of links that l goes through. Note that $b_f^l(L)$ is always positive.

The proposed SSP mechanism consists of two main phases: the first phase is to establish a primary tree route from the request source to all its destinations using either only some existing survivable lightpaths, or some new lightpaths as well, if and only if setting new lightpaths is necessary; and the second phase

is to protect each new primary lightpath, if any, against single link failure via segment protection. Specifically, because we have previously demonstrated in [117] that carefully fragmenting new primary/backup lightpaths into shorter ones helps to improve the network BBR for segment protection, a lightpath-fragmentation (LF) method is adopted in the proposed SSP method to conservatively split new lightpaths into shorter ones. In this manner, the SSP mechanism attempts to minimize the network resources allocated for request protection.

In what follows, we present the detailed two-phase SSP scheme for survivable dynamic multicast traffic grooming.

6.3.2 Two-Phase SSP Scheme for Survivable Dynamic Multicast Traffic Grooming

6.3.2.1 First-phase Primary-Tree Route Calculation

When a lightpath-based approach is adopted, the primary tree for the multicast transmission is composed of a number of lightpaths. Because establishing new lightpaths consumes additional wavelength and transceiver resources, we use existing survivable lightpaths to serve as many request destinations as possible, and new lightpaths are established to serve the remaining destinations only when it is necessary to do so.

It is well known that the identification of the minimal-cost primary tree route is an NP-complete problem. We adopt the simple yet efficient minimum-cost path heuristic (MPH) [56] to connect the request destinations one by one to the primary tree using minimum cost paths. Procedure 6.1 presents the process for calculating the primary route T for a multicast request.

Procedure 6.1: Primary_Route_Calculation_sub_Algorithm**input** : A network $G(V, E)$ and a multicast request $R\{s, D, b\}$.**output:** A primary tree route T for serving $R\{s, D, b\}$.

- 1 Generate an auxiliary graph (AG) using all existing survivable lightpaths; call **MPH** on AG to initiate a tree from s to serve as many members of D as possible. Remove those served members from D . If D is empty, go to Step 12; otherwise, save the partial tree T and continue.
// *Grooming with existing survivable lightpaths*
- 2 Add s and all the other nodes of T to a node set S . // *optical layer primary tree routing (2-11)*
- 3 **while** $D \neq \Phi$ **do**
- 4 Generate an optical-layer AG; calculate all-to-all shortest paths between each node in S and each node in D using the lightpath-cost definition presented in Eq. 6.1;
- 5 Choose the single shortest path P among all paths that connect any member of S to a member d of D . Denote the distance of P by dis .
- 6 **if** $dis < \infty$ **then**
- 7 $S = S \cup \{d\}$, $D = D \setminus d$; save the path P onto the primary tree T ;
- 8 **else**
- 9 Block the request $R\{s, D, b\}$ and break;
- 10 **end**
- 11 **end**
- 12 Return the primary tree route T .

Step 1 attempts to utilize the existing survivable lightpaths to serve as many request destinations as possible. If there are sufficient existing survivable lightpaths, the request will be served directly; otherwise, the remaining destinations will be connected to the primary tree one by one using new lightpaths in Steps 2 – 11.

Note that various path-cost definitions can be adopted in Step 4 to define the distance between two nodes. To improve the network's BBR performance while balancing the resource consumption, the cost definition shown below is adopted [104]:

$$C_{ij} = \begin{cases} \frac{1-r}{p \times r \times (\overline{H}+1)} - H_{ij} \ln \left(1 - \frac{1}{\omega_{ij}+1}\right) & \text{if } \omega_{ij} > 0 \text{ and } p > 0 \\ \infty & \text{if } \omega_{ij} = 0 \text{ or } p = 0 \end{cases} \quad (6.1)$$

where p is the smaller of the transmitter number at the source node and receiver number at the destination node of the lightpath, ω_{ij} is the number of wavelengths available along the lightpath route, $\overline{H}$ is the average path length of the network, and H_{ij} is the minimum hop length between the two end nodes of the lightpath. The intent of this function is to balance the consumption of wavelength and transceiver resources. Specifically, if both resources are abundant, the costs of consuming them will be low and similar to each other, whereas if either becomes

scarce, the cost of consuming the scarce resource will increase. Our previous studies have demonstrated that such a definition leads to satisfactory performance [104]

Once the primary tree route is identified for a request in the first phase, the proposed method enter its second phase to identify a backup route for each new primary lightpath, wherein auxiliary graphs are generated for routing purposes.

6.3.2.2 Graph Generation for Backup-Route Calculation

For each new lightpath added to the primary tree route T , the SSP scheme is adopted to protect it at the connection level. Specifically, the lightpath is fragmented into segments, which we refer to as *active segments* (ASs) following [119], and each segment is then protected using one or more link-disjoint *backup segments* (BSs). For example, the new primary lightpath shown in Fig. 6.1(b) is fragmented into three ASs, i.e., $S \rightarrow n_1$, $n_1 \rightarrow n_3$, and $n_3 \rightarrow d$, and each AS is protected by a separate link-disjoint BS. The nodes n_1 and n_3 , each of which acts as the source node of a segment, are called switching nodes.

To identify backup routes for a multicast session $R\{s, D, b\}$, an auxiliary graph with appropriate edge costs is generated. On such a graph, an edge that represents a backup route along which a wavelength has been reserved for protection is called an *existing* backup lightpath, whereas an edge that denotes a route along which a new wavelength must be reserved for protection is called a *new* backup lightpath. Suppose that P is a new primary lightpath to be protected, AS is an active segment of P , and SG is the set of links that AS traverses. The link-cost assignment policy

below is adopted for the calculation of AS's backup route:

$$L_{ij} = \begin{cases} \infty & \text{an existing backup lightpath } e \text{ that is not link - disjoint} \\ & \text{with AS or its capacity } r_m + b_f^{AS}(SG) < b; \\ \varepsilon + b_{ad} & \text{an existing backup lightpath } e \text{ that is link - disjoint with} \\ & \text{AS with capacity } r_m + b_f^{AS}(SG) \geq b; \\ b \times C(e) & \text{a new backup lightpath to be set up and that is link -} \\ & \text{disjoint with AS;} \end{cases} \quad (6.2)$$

where ε is a small positive value (10^{-2} used in this study) and b_{ad} is the amount of additional capacity that should be reserved on an existing backup lightpath for AS protection. It can be calculated as $b_{ad} = \max(0, b - b_f^{AS}(SG))$.

As indicated in Eq. (6.2), the costs assigned to the edges on the auxiliary graph are different from each other: the costs for existing backup lightpaths that do not have sufficient residual bandwidth or are not link-disjoint with AS are infinitely large, whereas the costs for those existing lightpaths that do have sufficient residual capacity are proportional to the additional capacity that must be reserved for request protection. For new backup lightpaths to be established, their assigned costs must equal the bandwidth required by the request multiplied by the cost of the new lightpath $C(e)$, which is defined as follows:

$$C(e) = \begin{cases} \left(\frac{(1-r)}{p \times r(H+1)} - H_e \ln \left(1 - \frac{1}{\omega_e + 1} \right) \right) \times amp & \text{if } \omega_e > 0 \text{ and } p > 0 \\ \infty & \text{if } \omega_e = 0 \text{ or } p = 0 \end{cases} \quad (6.3)$$

where all symbols have the same meanings as those defined in Eq. (6.1), and amp is an amplification factor. This definition ensures that the cost of establishing a new backup lightpath, i.e., $C(e)$, is usually approximately 5 times [108] that of using an existing lightpath when the network is in its typical operation status. (In this chapter, we assume that $\omega_e = W/2$ for the typical network status. For a detailed discussion please refer to [104].)

Once the auxiliary graph is generated, any shortest-path algorithm can then be applied to identify a backup route for each primary path. However, because long

lightpaths may impose a strict wavelength-continuity constraint on lightpath routing, thereby degrading resource sharing in the grooming process, an LF method is adopted in the proposed SSP method for both AS determination and BS fragmentation.

Below, we present the main idea of the LF method.

6.3.2.3 LF Method for AS Determination and BS Fragmentation

Suppose that L is a primary/backup route along which a new wavelength should be reserved; n is an intermediate node of L with a fan-out degree of d_n . The numbers of idle transmitters and receivers on n are denoted as T_n and R_n , respectively, and the numbers of free wavelengths along the incoming and outgoing links of L are denoted as ω_{in} and ω_{out} , respectively. The three parameters defined below are used to determine whether the path route L can be fragmented at n :

$$\alpha_m = \min\left(\frac{T_n}{d_n \times \omega_{out}}, \frac{R_n}{d_n \times \omega_{in}}\right) \quad (6.4)$$

$$\alpha_n^p = \frac{1}{H_n^p} \quad (6.5)$$

$$\alpha_n^b = \frac{1}{H_n^b} \quad (6.6)$$

where H_n^p and H_n^b are the average shortest primary and backup path lengths from n to all other network nodes, respectively.

In the above equations, α_m denotes the smaller of the add and the drop ratios on node n for the current network status; α_n^p and α_n^b are inversely proportional to the average hop length from node n to all the other network nodes, indicating the availability of the port resources for lightpath fragmentation. When $\alpha_m > \alpha_n^p$

($\alpha_m > \alpha_n^b$), it indicates that the transceiver resources at n are relatively redundant and it is beneficial to segment L at n so that there is no need to establish too long lightpaths, and thereby alleviates the wavelength continuity constraint on L 's establishment; otherwise, the transceiver resources are regarded as relatively limited at n , and it is better for L to bypass n to conserve transceiver resources.

We refer to a node n as *fragment node* if $\alpha_m > \alpha_n^p$ for a primary lightpath or $\alpha_m > \alpha_n^b$ for a backup path that passes through n . Procedure 6.2 presents the major steps of LF method.

Procedure 6.2: Lightpath_fragmentation_scheme(LF)

input : A network $G(V, E)$ and a lightpath L

output: A set of new lightpaths.

```

1 while any node of  $L$  has not been checked do
2   for each intermediate node (if any)  $n$  along  $L$  do
3     Calculate  $\alpha_m$  for  $n$  ;
4     if  $\alpha_m > \alpha_n^p$  (or  $\alpha_m > \alpha_n^b$ ) at  $n$  then
5       Fragment  $L$  at  $n$  and obtain two new lightpaths  $L_a$  and  $L_b$ ;
6        $L = L_b$ ;
7     end
8   end
9 end
```

The LF method is adopted to segment both the primary and backup paths of each survivable lightpath that will be set up. For example, using the LF method, the survivable lightpath presented in Fig. 6.1(b) could be fragmented as illustrated in Fig. 6.2, wherein n_1 , n_3 , 1, and 2 are fragment nodes.

Even though such a lightpath-fragmentation process helps alleviate the wavelength-continuity constraint, our previous results in [117] demonstrated that overly greedy path fragmentation still degrades the network BBR. This degradation occurs because protecting a large number of ASs requires a large amount of protection resources, which ultimately affects the network performance.

Inspired by this observation, we propose the conservative application of the LF method for the protection of new lightpaths using the proposed SSP scheme. Specifically, we adopt a similar two-phase process to that applied in the PROMISE

FIGURE 6.2: LF method adopted for primary path protection.

method proposed in [118], and employ the LF method for both AS determination and BS fragmentation in the second-phase of the proposed SSP mechanism. The primary objective is to minimize the network resources allocated for lightpath protection.

6.3.2.4 Second-Phase LF-Based Backup-Route Calculation

After the first-phase calculation, we have obtained a primary tree that may contain some new primary lightpaths. Let us denote a new lightpath to be protected as P ; the number of optical hops that it traverses as n , and all nodes along P are labeled from 0 to n . Assume that m ($0 < m < n$) is an intermediate node along P , AS_m is an active segment from node m to node n along P , C_m is the shortest link-disjoint backup path that has been identified for AS_m protection, and $C_{m,i}$ is the shortest backup path that has been identified for protecting the active segment between node m and node i ($m < i < n$). The costs of the latter two paths are $C(m)$ and $C(m,i)$, respectively; they are the sums of the costs of each backup lightpath that C_m and $C_{m,i}$ traverse. Procedure 6.3 presents the major steps in the second-phase LF-based backup-route calculation.

The second phase of the proposed SSP mechanism begins by testing each intermediate node m along the primary lightpath P in Step 2. A fragment node is chosen in Step 3. Step 4 tentatively selects a few links along P to be active segment AS_m ; in Steps 5 and 6, the algorithm attempts to identify the shortest link-disjoint path to protect the active segment AS_m from one end to the other.

Procedure 6.3: LF_based_Backup_Path_Provisioning**input** : A network $G(V, E)$ and new primary lightpath P for $R\{s, D, b\}$.**output**: A set of backup lightpaths to protect P .

```

1 Clear the set of fragment nodes  $S_F$ ;
2 for each intermediate node  $m = n - 1$  to 0 along  $P$  do
3   if  $m$  is a fragment node with  $\alpha_m > \alpha_m^p$  then
4     Set the links from  $m$  to  $n$  along  $P$  to be  $AS_m$  ;
5     Generate an auxiliary graph (AG) using both existing backup lightpaths and new
6     backup lightpaths to be established, and assign each edge a cost according to Eq.
7     (6.2); // Graph generation for backup routing
8     Calculate the shortest path that is link-disjoint with  $AS_m$  on AG for its protection and
9     record this path as  $C_m$ ; //  $C_m$  initialization
10    for each node  $i$  in  $S_F$  do
11      Set AS to be the links from  $m$  to  $i$  along  $P$  ;
12      Generate an auxiliary graph (AG') using both existing backup lightpaths and new
13      backup lightpaths to be established, and assign each edge a cost according to Eq.
14      (6.2); // graph generation for AS protection
15      Calculate the shortest path that is link-disjoint with AS from  $m$  to  $i$  on AG' to
16      protect AS and denote this path by BS; // backup routing
17      Call Procedure 6.2 (Lightpath Fragment scheme) to process each new backup
18      lightpath along BS and return a set of new lightpaths, calculate the cost  $C(m, i)$ 
19      using Eq. (6.1), and then multiply the sum cost of these new lightpaths by  $amp$ ;
20      record the processed BS route as  $C_{m,i}$ ; // LF for backup route processing;
21       $C_m = \min(C_m, C_{m,i} + C_i)$ ; // choose the back route for  $AS_m$ 
22    end
23     $S_F = S_F \cup \{m\}$ 
24  end
25 end
26 Return  $C_0$  for protecting the new lightpath  $P$ .

```

Steps 7–13 recursively test the possible fragmentation of the active segment AS_m , and ultimately choose a minimum- cost backup path for the protection of AS_m . In Step 14, the tested fragment node is recorded; this node might be used later as a switching node for some other segment. Note that a backup path BS may contain both existing backup lightpaths and new lightpaths to be established, and a new backup lightpath will be segmented as long as it contains any fragment nodes. The above process is repeated for all fragment nodes along P until the backup path with the minimum cost is identified to protect P . Once the backup path is determined for P , the segmentation of P itself and the fragmentation of all BSs are also determined.

Figure 6.3 presents an illustrative example of the second-phase backup-path provisioning process. Assume that a primary lightpath P connecting nodes 0 and 4

FIGURE 6.3: An illustrative example of LF-based backup path provisioning. (a) The primary lightpath P to be protected; (b) a tentative segment AS to be protected; (c) the minimum cost backup path for AS protection; (d) extended AS and its possible backup routes; (e) a minimum cost path for AS protection; (f) possible backup paths for tentative AS protection that include node 1; (g) minimum cost backup path for AS protection; (h) the tentative segment P and all its possible backup routes; (i) the final backup path for P protection.

is to be protected; nodes 1, 2 and 3 are fragment nodes. At the beginning, for the active segment between nodes 3 and 4, the minimum-cost link-disjoint backup path shown in Fig. 6.3(b) is identified for protection, as shown in Fig. 6.3(c). In Fig. 6.3(d), the active segment AS_m is extended to include links between fragment nodes 2 and 3. To protect this updated AS_m , there are two options for backup route selection, as shown in Fig. 6.3(d): one is the shortest link-disjoint path directly from node 2 to node 4, and the other one is a combination of two backup paths between nodes 2 and 3 and the nodes 3 to 4, respectively. After the LF-based path segmentation and cost calculation, the combined backup path shown in Fig. 6.3(e) is chosen for AS_m protection. This procedure is repeated for the active segment AS_m to incorporate fragment nodes along P one by one, and ultimately, the minimum-cost path shown in Fig. 6.3(i) is chosen to protect P .

For a new lightpath with n intermediate fragment nodes, the second-phase backup-path calculation sub-algorithm will check all possible routes for its protection and will ultimately choose the one that requires the minimum amount of backup resources. The BSs can be identified using any shortest-path algorithm, e.g., Dijkstras algorithm, which has a complexity of $|E| + |V| \log(|V|)$, where $|E|$ and $|V|$ are the numbers of network edges and network nodes, respectively; the total numbers of BSs calculations is limited to $n(n + 1)/2$.

Using the primary- and backup-route provisioning sub-algorithms described above, we now present the LF-SSP scheme for survivable dynamic multicast traffic grooming.

6.3.3 The Proposed LF-SSP Mechanism

Algorithm 6.1 presents the major steps of the LF-SSP scheme for survivable dynamic multicast traffic grooming.

To encourage traffic grooming when fulfilling a request, LF-SSP gives higher priority to the use of existing survivable lightpaths. Hence, Step 1 attempts to identify a primary route T_P for the request, and if a tree route that comprises only existing survivable lightpaths exists, it will be adopted to directly fulfill the request; otherwise, new lightpaths must be established for the request. Steps 2 – 5 attempts to protect each new primary lightpath on T_P using the LF-based backup-path calculation sub-algorithm to minimize the amount of backup resources allocated for each lightpath; if a backup route cannot be found for any primary lightpath, the request will be rejected. Steps 6 – 20 allocate the required network resources for a survivable lightpath, and finally Step 21 updates the network status. Note that for the traffic grooming in Step 1 and the lightpath wavelength assignment in Steps 2 – 20, the first-fit wavelength-assignment policy is adopted if there is more than one candidate route.

Algorithm 6.1: LF-SSP for dynamic multicast traffic grooming

input : A network $G(V, E)$ and a multicast request $R\{s, D, b\}$.
output: A survivable multicast tree route for $R\{s, D, b\}$.

- 1 Call Primary_Route_Calculation_sub_Algorithm, which returns a primary route T_P ; if T_P consists only of existing survivable lightpaths, go to Step 18 and otherwise, continue;
- 2 **for** each new lightpath P on T_P **do**
- 3 Call LF_based_Backup_Path_Provisioning for P protection;
- 4 If backup paths (BSs) can be found for P , assign the BSs to P and denote P as survivable on T_P , and if BSs cannot be found, block the request;
- 5 **end**
- 6 **for** each new primary lightpath P on T_P **do**
- 7 Fragment P according to its BSs; allocate transceiver and wavelength resources to each fragmented new AS of P ; if any resource is unavailable, block the request;
- 8 **for** each new backup lightpath along the BSs **do**
- 9 Call Procedure 6.2 to fragment the lightpath, if possible;
- 10 **for** each new lightpath L obtained from fragmentation **do**
- 11 Reserve a wavelength along the route L ;
- 12 **if** source node of L is neither s nor a fragment node on P **then**
- 13 Allocate a transmitter at this node;
- 14 **end**
- 15 **if** end node of L does not belong to D or is not a fragment node on P **then**
- 16 Allocate a receiver at this end node;
- 17 **end**
- 18 **end**
- 19 **end**
- 20 **end**
- 21 Update the residual capacities of all survivable lightpaths along route T_P .

Finally, we compare the complexity of LF-SSP to those of the existing wavelength- and subwavelength-level methods. For an arriving request with destination size D , all these methods firstly find a minimum spanning primary tree (MST) with a complexity of $O(D|V|^2)$, where $|V|$ is the number of network nodes. SPT method then protects each segment of such MST using a link-disjoint tree, its complexity hence is $O(3D|V|^3)$; LP protects the request destinations on the MST one by one in an ascending order of their distances from the source, with complexity of $O(0.5D^2|V|^2)$. CL-SSP protects each new lightpath on the MST using a link-disjoint path its complexity therefore is $O(D|V|^2 + D(|E| + |V|\log(|V|)))$ with $|E|$ being the number of network links. For the LF-SSP method, since it adopts Procedure 6.3 to minimize the resources for both request grooming and protection, its complexity is $O(D|V|^2) + 0.5D^2(|E| + |V|\log(|V|))$ in the worst case. In a backbone network with a moderate number of network nodes, the complexity of LF-SSP is acceptable.

6.4 Performance Evaluation

In this section, we conduct a number of experiments to evaluate the performance of the LF-SSP method in various cases. We first present the simulation setup, and then compare the LF-SSP to existing methods for either subwavelength- or wavelength-level multicast request protection; finally, we assess the influences of several factors on LF-SSP performance.

6.4.1 Simulation Setup

We simulate a dynamic network environment to evaluate LF-SSP performance in various cases. We assume that all multicast requests arrive/leave the network dynamically according to Poisson distribution with mean rate λ and that their holding time is negative-exponentially distributed with a mean of $\mu = 1$; the request source and destinations are randomly chosen from among all network nodes, and the number of destinations of each request, denoted by N , is assumed to follow a truncated geometric distribution [81] defined by a parameter $q(0 < q < 1)$. Specifically, Eq. (6.7) and Eq. (6.8) presents the probability that a request has k destinations and the average number of request destinations, respectively:

$$P(N = k) = \frac{(1 - q)q^{k-1}}{q - q^{|V|}}, 2 \leq k \leq |V| - 1 \quad (6.7)$$

$$\begin{aligned} E(N) &= \sum_{k=2}^{|V|-1} k \times P(N = k) \\ &= \frac{2q - q^2 - |V|q^{|V|-1} + (|V| - 1)q^{|V|}}{(1 - q)(q - q^{|V|-1})} \end{aligned} \quad (6.8)$$

where $|V|$ is the number of network nodes.

FIGURE 6.4: Two network topologies utilized for simulations. (a) 14-node 21-link NSFnet topology and (b) 24-node 43-link USnet topology.

The two commonly used network topologies depicted in Fig. 6.4, 14-node 21-link NSFnet and 24-node 43-link USnet, are utilized for the simulations. We assume the followings:

1. Each network link consists of two fibers travelling in opposite directions, each carrying $W = 32$ wavelengths; the full capacity of each wavelength is $B = 16$ units.
2. The averaged number of destinations of each session is 3. For subwavelength requests, the required bandwidth is an integer that is uniformly distributed in the interval $[1, 16]$, and for wavelength-level multicast requests, the required bandwidth is 16.
3. The number of transceivers on a network node is $W \times d_i \times r$, where d_i is the fan-out degree of the node.
4. Loss of signal power loss because of transmission attenuation or light splitting is not considered.

In the following figures, the results are averages of at least five independent simulations, each was produced based on 10^5 connection requests. Because the two topologies generate relatively consistent performance results, unless otherwise

specified, we present only the results obtained using the NSFnet topology for comparisons and discussions.

6.4.2 Comparison of LF-SSP to the Existing Subwavelength-Level CL-SSP Method

6.4.2.1 Comparisons in a port-unlimited network

Figure 6.5 compares the LF-SSP method to the existing CL-SSP method in a port-unlimited network under various traffic loads for subwavelength-level multicast request protection. It is evident that LF-SSP outperforms CL-SSP throughout the entire range of traffic loads: under low traffic loads, e.g., $\rho = 65$ Erlangs, LF-SSP is superior to CL-SSP by more than three orders of magnitude, whereas under higher traffic loads, e.g., $\rho = 110$ Erlangs, LF-SSP remains superior to CL-SSP by approximately 49.1% in terms of network BBR.

6.4.2.2 Comparisons in a port-limited network

Figure 6.6 presents a comparison between LF-SSP and CL-SSP in a port-limited network, where $r = 0.6$ for all network nodes, under various traffic loads. The results illustrate that LF-SSP again consistently outperforms CL-SSP throughout the entire range of traffic loads in the request-grooming process. Specifically, when the traffic load is low, e.g., $\rho = 60$ Erlangs, LF-SSP is superior by more than three orders of magnitude, whereas under higher traffic loads, e.g., $\rho = 100$ Erlangs, it outperforms CL-SSP by approximately 65.2%.

The results presented above in both Fig. 6.5 and Fig. 6.6 convincingly demonstrate that LF-SSP can achieve satisfactory results in networks with either limited or unlimited port resources. Compared to CL-SSP, LF-SSP utilizes the LF method

FIGURE 6.5: LF-SSP compared to CL-SSP in the port-unlimited NSFnet topology under various traffic loads ($r = 1.0$).

FIGURE 6.6: LF-SSP compared to CL-SSP in the port-limited NSFnet topology under various traffic loads ($r = 0.6$).

more conservatively when splitting new primary/backup lightpaths to minimize the total resources allocated for the protection of a new request. This additional step allows the LF-SSP method to more efficiently utilize the network resources.

To evaluate the resource utilization of LF-SSP, we define a new metric, the average utilization of existing channel capacity, which is defined as the ratio of the total traffic load carried by network to the total capacity of all existing channels, for performance measurement. Figure 6.7 presents the values of this metric for both methods in the port-limited NSFnet topology under various traffic loads. It is evident that because CL-SSP splits new backup lightpaths as long as such lightpaths can be fragmented at any of their intermediate nodes, its wavelength utilization is lower. By contrast, through conservatively utilizing the LF method to minimize the resources allocated for request protection, the LF-SSP method improves the utilization of the channel capacity by approximately 15% on average. This improvement clearly demonstrates that it is beneficial to adopt the additional step of conservatively utilizing the LF method to minimize the resources allocated for request protection.

6.4.3 Comparison of LF-SSP to Existing Wavelength-Level SPT and LP Methods

To further assess the performance of the LF-SSP method, we extend its application to the protection of wavelength-level multicast requests and compare its performance to that of the two best existing methods for this task in various cases.

FIGURE 6.7: The average utilization of existing channel capacity versus the traffic loads in the port-limited NSFnet topology ($r = 0.6$).

6.4.3.1 Comparisons in a port-unlimited network

Figure 6.8 compares LF-SSP to SPT and LP in the port-unlimited NSFnet topology under various traffic loads. It is evident that LF-SSP reliably outperforms both SPT and LP methods throughout the entire range of traffic loads. Specifically, when under low traffic loads, i.e., $\rho < 45$ Erlangs, LF-SSP consistently outperforms LP by approximately one order of magnitude, and outperforms SPT by about 83.6%; whereas when under higher traffic loads, e.g., $\rho = 70$ Erlangs, the performances of LF-SSP is still slightly better than those two methods. Such results can be understood: when the traffic load is low, the network resources are abundant, and thereby, the method that is able to make more efficient the network resources can easily stand out; when the traffic load is high, however, the wavelength resource becomes limited, and therefore, the performance differences between the different methods are not so obvious.

FIGURE 6.8: LF-SSP compared to the SPT and LP methods in the port-unlimited NSFnet topology under various traffic loads ($r = 1.0$).

6.4.3.2 Comparisons in a port-limited network

Figure 6.9 compares LF-SSP to both the SPT and LP methods in the port-limited NSFnet topology under various traffic loads. Again we can see that, LF-SSP evidently achieves better performance than both algorithms throughout the entire range of traffic loads: when under low traffic loads, e.g. $\rho = 45$ Erlangs, it outperforms LP method by more than one order in of magnitude, and outperforms SPT by more than 80%; whereas when under higher traffic loads, e.g. $\rho = 70$ Erlangs, performance of LF-SSP is still slightly better than those of the SPT and LP methods. Such observation is due to the fact that when the traffic load is low, the network resources are relatively abundant, and thus LF-SSP is able to fragment the new lightpaths to improve resource sharing in protection process, whereas when the traffic load is high, both primary and backup lightpaths are less fragmented, which drags down the resource sharing efficiency, and thus performance superiority

of LF-SSP method is not obvious.

FIGURE 6.9: LF-SSP compared to the SPT and LP methods in the port-limited NSFnet topology under various traffic loads ($r = 0.6$).

The results presented in Fig. 6.8 and Fig. 6.9 illustrate that LF-SSP is able to achieve satisfactory performance if it is utilized to protect wavelength-level multicast requests. It is also interesting to note that a comparison between Figures 6.8 and 6.9 indicates that LF-SSP performs better when the port resources are abundant. This is attributed to the fact that LF-SSP is designed to make the best possible use of transceiver resources to minimize the network BBR.

Below, we compare LF-SSP to the existing methods in various cases to further assess the influence of transceiver resources on the LF-SSP performance.

6.4.4 Effects of the Add/drop Port Resources on Sub-wavelength and Wavelength Level Protection

Figure 6.10 offers a comparison between LF-SSP and CL-SSP in the NSFnet network with various port resources for subwavelength-level multicast request protection. It is clear that when the port resources are limited, i.e., $r < 0.2$, the performance difference between the two methods is not so big, whereas when the port resources increase, the performance of LF-SSP improves rapidly and exceeds that of CL-SSP by more than an order of magnitude once $r > 0.3$. When the port resources are relatively abundant, e.g., $r > 0.7$, LF-SSP outperforms CL-SSP by nearly two orders of magnitude.

FIGURE 6.10: Comparison between LF-SSP and CL-SSP in terms of the add/drop ratio in the NSFnet topology for subwavelength-level request protection (traffic load $\rho = 70$ Erlangs).

Figure 6.11 compares LF-SSP to the SPT and LP methods for the protection of wavelength-level multicast requests in the NSFnet network. It is again clear that

when the port resources are limited, i.e., $r < 0.35$, the BBR performances of these methods are not very different from one another. When the port resources are not so bottlenecked, i.e., $r > 0.35$, the BBR performance of LF-SSP improves and supersedes those of SPT and LP; when $r > 0.5$, LF-SSP is superior to LP by nearly two orders of magnitude, and outperforms SPT by approximately one order of magnitude.

FIGURE 6.11: LF-SSP compared to the SPT and LP methods in terms of the add/drop ratio in the NSFnet topology for wavelength-level multicast request protection (traffic load $\rho = 35$ Erlangs).

The above results in both cases can be understood as follows: when the port resources are bottlenecked, the use of different methods cannot lead to any large differences in the results because of the limited resources; when the port resources are no longer bottlenecked, however, the method that can make efficient use of the port resources stands out. As LF-SSP is designed to make the best possible use of the port resources, its BBR performance improves once the port resources are

reasonably sufficient. The conservative lightpath fragmentation helps improve the resource sharing in request protection process.

Such comparisons clearly demonstrate that by utilizing the LF method to minimize the total resources allocated for request protection, LF-SSP is able to achieve satisfactory performance when protecting either wavelength- or subwavelength-level multicast requests in networks with various amounts of port resources.

6.4.5 Effects of the Average Number of Destinations of Each Request

We also study the influences of the average numbers of request destinations on the performances of LF-SSP for supporting subwavelength- and wavelength-level multicast requests in NSFnet topology with limited and unlimited port resources. Because the conclusions hold for networks with either limited or unlimited port resources, we present only the results obtained for port-limited networks.

Figure 6.12 compares LF-SSP and CL-SSP in the port-limited NSFnet topology for the protection of subwavelength-level multicast requests. Simulation results show that LF-SSP evidently outperforms CL-SSP: when the average size of the multicast sessions is small, e.g., $E(N) < 4.5$, LF-SSP is superior to CL-SSP by more than one order of magnitude, whereas when the size of the multicast sessions increases, the performance difference between the two methods decreases. This observation can be understood as follows: if the size of the average multicast-session is small, the network resources are relatively abundant and the performance superiority of LF-SSP is more obvious because this method uses network resources more efficiently. When the average session size is large, the network resources are relatively scarce, LF-SSP cannot cause an equally significant difference as that in the former case because of the bottlenecked resources, the performance difference between the two methods is therefore reduced.

FIGURE 6.12: LF-SSP compared to CL-SSP for sub-wavelength level multicast request protection in the NSFnet topology ($r = 0.6, \rho = 60$ Erlangs).

Figure 6.13 compares LF-SSP to the SPT and LP methods for the protection of wavelength-level multicast requests in the port-limited NSFnet topology. It is evident that LF-SSP achieves better BBR performance than the SPT and LP methods in the simulated cases. As can be seen, LF-SSP achieves much better BBR performances. Specifically, when the multicast-session size is small, i.e., $E(N) < 4$, LF-SSP outperforms LP by about an order in magnitude. Even when the average multicast-session size is larger, i.e., $E(N) > 4.5$, it still outperforms LP by 54% when averaged over all the simulated cases. When compared to the SPT method, LF-SSP also achieves much better performances in different cases: when $E(N) < 4$, it outperforms SPT by more than 60%, while when $E(N) > 4.5$, its performance is still slightly better over that of SPT. Such results can be understood: when the average multicast-session size is small, the network resources are relatively abundant, and thus, the LF-SSP method is able to achieve better BBR

performance through utilizing the lightpath fragmentation method to improve resource sharing in request protection process; when the traffic load is high, however, the network resources become limited. In this case, resource sharing with lightpath fragmentation becomes less significant, and thus, performance superiority of LF-SSP becomes smaller.

FIGURE 6.13: LF-SSP compared to SPT and LP for wavelength level multicast request protection in the NSFnet topology ($r = 0.6, \rho = 30$ Erlangs).

The above results again demonstrate that LF-SSP steadily outperforms the existing methods in networks with either limited or unlimited port resources. It is worth noting that the proposed LF-SSP scheme is designed for subwavelength-level multicast request protection. It is not a surprise that for wavelength-level request protection, there may still be nontrivial space for the performance of the algorithm to be further improved. For example, the three parameters defined in (6.4)-(6.6) can be optimized to further improve the network resource sharing for wavelength-level request provisioning. We compare LF-SSP versus SPT and LP without making such changes since our main focus is not to design the best possible

algorithm for wavelength-level protection. Instead, we present such comparisons simply because there exists too few existing algorithms for subwavelength-level request protection. Our results show that, even without making any changes to LF-SSP for wavelength-level request protection, the proposed algorithm nevertheless achieves better performance than LP and SPT in different cases. When the network resources are abundant, and the session size is not too large, it steadily outperforms the best existing method by about 60% under low traffic loads.

6.5 Summary

In this chapter, we addressed the problem of protecting subwavelength-level multicast requests against single link failure in dynamic traffic grooming. An efficient mechanism, lightpath-fragmentation-based segment shared protection (LF-SSP), was proposed. LF-SSP attempts to minimize network blocking performance by adopting the LF method to conservatively split the new primary/backup lightpaths into shorter ones to improve resource sharing while minimizing the total resources allocated for request protection. We performed extensive simulations to evaluate the performance of LF-SSP in various cases. The results demonstrated that LF-SSP is capable of achieving satisfactory performance for either subwavelength- or wavelength-level multicast request protection in networks with limited or unlimited resources under various traffic loads; the performance superiority of LF-SSP compared to the existing methods is more evident when the network resources are relatively abundant. The influences of variations in the number of add/drop ports and the average multicast-session size on the LF-SSP performance were also evaluated.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

To satisfy the ever-increasing bandwidth demand of Internet users, future optical networks should be capable of supporting a variety of network applications and provide extremely high capacity. For this reason, WDM-based optical networking technologies are currently under rapid development and have been widely deployed in both access and backbone networks. Furthermore, as IP/MPLS technology continues to gain popularity, it is widely accepted that IP/MPLS over WDM will be a key component for next-generation optical Internet. However, because the capacity provided by a single-wavelength channel is much higher than the bandwidth requirement of a user request, dynamic request-provisioning mechanisms are needed to efficiently utilize the capacity provided by modern optical networks. This dissertation is devoted to an investigation of the problems related to dynamic request-provisioning mechanisms in next-generation optical Internet applications. The primary objective is to devise efficient provisioning mechanisms for subwavelength-granularity unicast or multicast requests in their dynamic traffic-grooming processes.

We began this dissertation with a brief review of some previous work regarding wavelength-routed WDM networks in Chapter 2. First, we introduced some of the basic concepts of WDM networks and some emerging technologies that have been proposed for the next-generation optical Internet infrastructure. We also briefly discussed the advantages and shortcomings of these technologies for near-term deployment. Then, we presented some existing work concerning the traffic-grooming problem in WDM networks, in particular, the dynamic grooming problem for either unicast or multicast traffic; finally, we introduced some basic elements of network survivability problems. We briefly described some schemes that have been proposed in the literature for wavelength-level multicast request protection in WDM networks and also discussed the problems from which these existing protection schemes suffer.

As overlay IP/MPLS over WDM is widely accepted to be the most practical network architecture for near-term deployment, it is essential for such a network to be capable of efficiently supporting unicasting services. We addressed the dynamic LSP routing problem in overlay IP/MPLS over WDM networks in Chapter 3. Through an extensive literature review, we found that the primary factor that limits the blocking performance of such a network is the limited information exchange between its two layer networks. Although various routing mechanisms have been proposed for such networks, none has considered the fact that the IP/MPLS-layer operator may have a record of all lightpath services that have been supported by the WDM-layer network. Such a record, if properly utilized, can facilitate the IP/MPLS-layer request-routing process. Based on this understanding, we proposed a historical data-utilization scheme to make use of the data record kept by an IP/MPLS-layer network and devised an Existing Link First (ELF) algorithm to improve the network blocking performance. Simulations were performed to compare the performance of the ELF algorithm to those of the best existing methods in various cases, and the results demonstrated that the proposed ELF algorithm

significantly outperforms the existing methods in various cases.

In Chapter 4, we studied the dynamic multicast traffic-grooming problem in wavelength-routed WDM networks. In the literature, both lightpath and light-tree schemes have been proposed to support dynamic multicast traffic grooming in WDM networks. However, we noted that although some very simple comparisons have been performed between two potential schemes for traffic grooming in certain cases, no results have firmly established which scheme is the better choice for dynamic multicast traffic grooming. We compared the blocking performances of two such schemes in dynamic multicast traffic grooming. We found that the lightpath-based methods are always superior to light-tree-based methods because lightpaths are more easily shared in the traffic-grooming process. Inspired by these results, we then proposed a lightpath-fragmentation (LPF)-based method to further improve the network blocking performance. We also performed extensive simulations to compare the LPF method to other existing methods in various cases to evaluate its performance. The influences of various parameters were also studied via simulation.

Because the future optical Internet infrastructure will need to support not only unicast requests but also multicast traffic, Chapter 5 investigated the dynamic multicast traffic-grooming problem in overlay IP/MPLS over WDM networks, and a dynamic multicast traffic-grooming algorithm, the overlay multicast provisioning (OMP) mechanism, was proposed. The OMP mechanism utilizes a historical data-learning (DL) scheme at the IP/MPLS layer for logical link-cost estimation and a lightpath-fragmentation (LPF) -based method at the WDM layer for resource-sharing improvement. The performance of the OMP scheme was compared to those of other existing schemes in various cases. The results clearly demonstrated the effectiveness of OMP in various cases. We also evaluated the respective influences of both the DL and LPF schemes on the OMP performance by comparing OMP to some other provisioning mechanisms. The results demonstrated

that the improvement in the OMP performance achieved using the DL scheme is more significant than that achieved using the LPF scheme when the fixed routing strategy is adopted at the WDM layer. We also assessed the influences of various lightpath-cost definitions and various WDM-layer routing policies on the OMP performance.

As the capacity provided by a single fiber continues to increase, network resiliency will prove to be essential in WDM networks because a single network failure may lead to tremendous amount of data loss. Chapter 6 investigates the problem of subwavelength-granularity multicast request protection in the dynamic traffic-grooming process. To improve resource sharing in both the traffic-grooming and protection processes, we proposed a method called the lightpath-fragmentation (LF)-based segment shared protection (LF-SSP) scheme for SMTG. Through utilizing the LF method to conservatively split new primary/backup lightpaths into shorter ones, LF-SSP tries to minimize the overall resources allocated for request protection. LF-SSP was compared to some existing methods for either subwavelength-level dynamic SMTG or wavelength-level multicast request protection in various cases. We also evaluated the influence of some other factors, such as the number of add/drop ports and the average number of request destinations, on the LF-SSP blocking performance.

7.2 Future Research Topics

The primary focus of this dissertation is the problem of dynamic request provisioning in future optical Internet applications, and a number of algorithms for either unicast or multicast traffic grooming have been proposed. Building on the preliminary framework that has been described in this dissertation, many other issues still deserve further investigation in the future. Below, we briefly present some possible topics for future study based on the groundwork laid by this dissertation.

7.2.1 Multi-user Service Provisioning in Overlay IP/M-PLS over WDM

For the approach to overlay IP/MPLS over WDM networks adopted in Chapter 3 and Chapter 5, we assumed that there is only one integrated service provider on the WDM-layer network that provides desired services to its end users. We also assumed that this service provider maintains complete cost records for all lightpath/light-tree services supported by the WDM-layer bandwidth provider. Although this assumption may be true in some cases, in practical network operations, there is highly likely to be more than one integrated service providers on the IP/MPLS layer sharing the same optical WDM-layer network. In such a scenario, each integrated service provider on the IP/MPLS layer can have only its own partial record of the lightpath/light-tree services supported by the WDM-layer network. Therefore, when provisioning unicast traffic, multicast traffic, or both, a single integrated service provider can only predict the network status of the WDM network using its own partial cost record. As it was demonstrated in Chapter 3 that the limited information exchange between the IP/MPLS layer and the WDM layer has already led to considerable difficulties in IP/MPLS-layer link-cost estimation, this availability of only partial cost records would make the problem even more severe and complex.

For the benefit of both IP-layer integrated service providers and WDM-layer bandwidth providers, efficient provisioning mechanisms are needed to improve the utilization of network resources and also to reduce the probability of request blocking. Specifically, for IP/MPLS-layer ISPs, more sophisticated mechanisms that can efficiently utilize partial historical data records to facilitate IP/MPLS-layer request routing are needed. Using such a method, an IP-layer ISP will be able to make more efficient routing decisions for its arriving requests, which will assist the ISP

in balancing its traffic distribution and also in improving its network services, e.g., lowering the occurrence of network blocking, and ultimately increasing its profits.

Meanwhile, for WDM-layer networks, such efficient mechanisms are also required for the bandwidth providers to improve their revenue. For example, a WDM-network operator may also maintain its own service records regarding its upper-layer clients . Through the effective use of such service records, the WDM-layer network operator would be able to optimize its allocation of resources to its various upper-layer clients such that it could utilize the minimum possible network resources to support the largest amount of upper-layer client services and thus achieve the highest possible profit. This topic is, to some extent, related to game theory problem, and to the best of our knowledge, no prior work has been done on this front.

7.2.2 Energy-Efficient Dynamic Multicast Grooming

In wavelength-routed WDM networks, both lightpath and light-tree schemes have been utilized to support dynamic multicast communications. Chapter 4 compared the blocking performances of two such schemes in dynamic multicast grooming. The results indicated that lightpath schemes reliably outperform light-tree schemes in achieving better blocking performances in various cases; however, a lightpath scheme requires a larger number of OEO conversions for each admitted multicast request. In WDM networks, because the data-transmission rate is very high, high-speed data-processing units are usually of very high cost, and the OEO conversions also introduce considerable intermediate-node processing delay . Therefore, multicast traffic grooming utilizing only a lightpath scheme could incur high costs and long delays. By contrast, multicast transmission using a light-tree scheme may lead to a higher probability of request blocking but requires fewer OEO conversions, i.e., lower network operation costs, in the traffic-grooming process. Therefore, a

trade-off exists between the network blocking performance and the practical network operation costs, and the ability to maintain an appropriate balance between the two is critical for network operators.

In Chapter 4, we were concerned only with the network blocking performance. In practice, however, a bandwidth provider typically attempts to provide the best possible service to its clients while simultaneously reaping the highest possible benefits on its own behalf by reducing the amount of high-cost equipment in use. In such a case, it would be beneficial to utilize a combination of lightpath and light-tree schemes in WDM networks to support the Internet traffic carried by these networks. Currently, although some mechanisms that utilize either lightpath, or light-tree, or both schemes for multicast traffic grooming problem have been proposed in the literature, none of these schemes has considered the question of how to improve network performances while simultaneously minimizing network costs. By combining the benefits of both lightpath and light-tree schemes, it should be possible to improve network blocking performances while minimizing overall network costs.

It is worth mentioning that throughout this dissertation, we assumed that all network nodes have no wavelength-conversion capabilities because of the high cost of wavelength converters. In recent years, however, as more and more physical-layer optical networking technologies have been rapidly advancing, high-speed wavelength converters have begun to emerge and may be deployed soon in future optical networks. Therefore, further research will permit the study of the influence of wavelength converters on networks that support dynamic multicast traffic grooming. It will also become possible to consider full, limited or fixed network-node wavelength-conversion capabilities.

For wavelength-routed WDM networks that support dynamic multicast services, it will also become necessary to consider some other constraints, such as sparse

network-node wavelength-conversion capabilities, sparse light splitting or limited numbers of add/drop ports, as these constraints are usually major concerns in practical networks.

7.2.3 Request Protection in the Traffic Grooming Process

Survivability is an important issue for practical modern optical networks. Chapter 6 investigated the protection of subwavelength-granularity multicast requests in dynamic traffic grooming. This preliminary work can be immediately extended to the following topics of future study:

1. By considering the energy efficiency, the mechanism could be extended to utilize both lightpath and light-tree schemes to protect multicast requests. In Chapter 6, a lightpath-based scheme was utilized to support dynamic subwavelength-granularity multicast requests because of the superior blocking performance of this type of scheme. As discussed above, however, for the purposes of energy consumption, it is not efficient to utilize lightpaths alone. Instead, the joint utilization of both lightpath and light-tree schemes would be more energy efficient. Thus, the problem of energy-efficient dynamic survivable multicast traffic grooming is worth further investigation. For example, there is a need for sophisticated algorithms to decide when and how to utilize lightpaths to improve network-resource sharing in the traffic-grooming process while simultaneously utilizing light trees to reduce the energy consumption in the network caused by OEO conversion. Although the design of such a mechanism is highly dependent upon the networks traffic pattern, which cannot be straightforwardly determined in many real-life network systems, typical scenarios can be considered to provide benchmark estimations.
2. It is desirable to extend the current protection algorithm from single link failure to cases of multiple link failures, node failures, or both. Chapter 6 considered

the protection of multicast requests from single link failure only. In practical network operation, however, occasional serious accidents, such as water leakage into undersea fiber links, shark bites, large earthquakes, or volcanic explosions, may cause multiple types of network failures simultaneously. In this case, efficient and resilient algorithms are needed to recover the affected network as quickly as possible. The multiple-failure scenario is much more complex than single link failure, especially for subwavelength-granularity multicast request traffic grooming. To the best of our knowledge, no previous efforts have been devoted to the protection of subwavelength-granularity multicast requests against multiple network failures in dynamic traffic grooming.

3. The simultaneous protection of dynamic unicast and multicast requests in WDM networks is also a topic of our future interest. The current literature regarding request protection has focused either on pure unicast or pure multicast traffic, and few efforts have focused on the simultaneous protection of these two types of traffic in the traffic-grooming process. In modern communication networks, however, both unicast and multicast traffic will coexist. As previously discussed, both lightpath and light-tree schemes for traffic transmission should be utilized simultaneously to conserve network costs without sacrificing network performance. Specifically, because an existing lightpath can be utilized to support unicast traffic, multicast traffic, or both, whereas an existing light tree can only be shared by certain multicast requests because of issues of resource efficiency and service security, there is a need for sophisticated protection mechanisms that account for self- and cross-sharing as well as intra-request resource sharing.

Publications

Journal Papers:

1. **X. Yu**, G. Xiao, and T.-H. Cheng, “Historical data learning based dynamic LSP routing for overlay IP over WDM networks,” *Optical Fiber Technology*, vol. 19, no. 4, pp. 308–319, Aug. 2013.
2. **X. Yu**, G. Xiao, and T.-H. Cheng, “Dynamic multicast traffic grooming in optical WDM mesh networks: Lightpath vs light-tree,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 5, no. 8, pp. 870–880, Aug. 2013.
3. **X. Yu**, G. Xiao, and T.-H. Cheng, “An Efficient Mechanism for Dynamic Multicast Traffic Grooming in Overlay IP/MPLS over WDM Networks,” *Optical Fiber Technology*, vol. 20, no. 4, pp. 341–352, Aug. 2014.
4. **X. Yu**, G. Xiao, and T.-H. Cheng, “A Segment-Shared Protection Mechanism for Dynamic Survivable Multicast Traffic Grooming in WDM Networks,” Submitted to *Optical Fiber Technology*. (Under Review).

Conference Papers:

1. **X. Yu**, G. Xiao, and T.-H. Cheng, “Historical data learning based dynamic LSP routing for overlay IP over WDM networks,” in *Proc. IEEE ICICS '2009*, Dec. 2009, pp. 1–5.
2. **X. Yu**, G. Xiao, and T.-H. Cheng, “Performance Comparisons between Light-path and Light-tree Schemes in Dynamic Multicast Traffic Grooming Process,” in *Proc. IEEE Photonics Global Conference (PGC) '2012*, Dec. 2012, pp. 1–5.
3. **X. Yu**, G. Xiao, and T.-H. Cheng, “Connection Level Segment Shared Protection for Dynamic Multicast Traffic Grooming,” in *Proc. IEEE ICICS '2013*, Dec. 2013.
4. **X. Yu**, G. Xiao, T.-H. Cheng and D. Huang, “Impacts of link-cost and routing methods on overlay network performance for dynamic traffic grooming” in *Proc. IEEE ICOCN '2014*, Nov. 2014.

Bibliography

- [1] (2012, Jun.) Cisco visual networking index (vni). [Online]. Available: http://www.cisco.com/en/US/netsol/ns827/networking_solutions_sub_solution.html
- [2] A. M. Odlyzko, “Internet traffic growth: Sources and implications,” in *Proc. SPIE*, vol. 5247, Sep. 2003, pp. 1–15.
- [3] M. Dankberg, “Transforming satellite broadband,” *SatMagazine*, pp. 54–56, Jun. 2008. [Online]. Available: <http://www.satmagazine.com/2008/SatMagJune08.pdf>
- [4] P. E. Green, “Progress in optical networking,” *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 54–61, Jan. 2001.
- [5] R. Ramaswami, “Optical networking technologies: What worked and what didn’t,” *IEEE Commun. Mag.*, vol. 44, no. 9, pp. 132–139, Sep. 2006.
- [6] R. Ramaswami, “Multiwavelength lightwave networks for computer communication,” *IEEE Commun. Mag.*, vol. 31, no. 2, pp. 78–88, Feb. 1993.
- [7] T.-H. Cheng, Y.-K. Yeo, Y. Wang, G. Xiao, L. Zhou, D. Wang, Z. Xu, and J. Liu, “Design of a 100tb/s multicast-capable optical packet router (invited),” in *Proc. of 2010 International Conference on Communications and Mobile Computing ICCMC2010*, vol. 2, Shenzhen, China, Apr. 2010, pp. 18–22.
- [8] R. G. H. van Uden, R. A. Correa, E. A. Lopez, F. M. Huijskens, C. Xia, G. Li, A. Schlzgen, H. de Waardt, A. M. J. Koonen, and C. M. Okonkwo, “Ultra-high-density spatial division multiplexing with a few-mode multicore fibre,” *Nature Photonics*, vol. 8, no. 11, p. 865–870, Nov. 2014.
- [9] G. N. Rouskas and H. G. Perros, “A tutorial on optical networks,” *NETWORKING 2002 Tutorials, LNCS 2497*, pp. 155–193, May 2002.
- [10] E. Rosen, A. Viswanathan, and R. Callon, “Multiprotocol label switching architecture,” *IETF RFC 3031*, Jan. 2001.
- [11] G. Swallow, J. Drake, Boeing, H. Ishimatsu, and Y. Rekhter. (2005, Oct.) Generalized multiprotocol label switching (GMPLS) use.

- [12] D. Papadimitriou, M. Vigoureux, K. Shiomoto, D. Brungard, and J. L. Roux. (2010, Oct.) Generalized MPLS (GMPLS) protocol extensions for multi-layer and multi-region networks (MLN/MRN).
- [13] M. J. O'Mahony, C. Politi, D. Klonidis, R. Nejabati, and D. Simeonidou, "Future optical networks," *IEEE J. Lightwave Technol.*, vol. 24, no. 12, pp. 4684–4696, Dec. 2006.
- [14] S. Yao, B. Mukherjee, and S. Dixit, "Advances in photonic packet switching: An overview," *IEEE Commun. Mag.*, vol. 38, no. 2, pp. 84–94, Feb. 2000.
- [15] S. J. B. Yoo, "Optical packet and burst switching technologies for the future photonic internet," *IEEE J. Lightwave Technol.*, vol. 24, no. 12, pp. 4468–4492, Dec. 2006.
- [16] C. Qiao and M. Yoo, "Optical burst switching (OBS) - a new paradigm for an optical internet," *Journal of High Speed Networks*, vol. 8, no. 1, pp. 69–84, Jan. 1999.
- [17] Y. Chen, C. Qiao, and X. Yu, "Optical burst switching: A new area in optical networking research," *IEEE Network*, vol. 18, no. 3, pp. 16–23, May/June 2004.
- [18] M. A. Ali, A. A. Shami, C. M. Assi, Y. Ye, and R. Kurtz, "Architectural options for the next-generation networking paradigm: Is optical internet the answer?" *Photonic Network Communications*, vol. 3, no. 1/2, pp. 7–21, Jan. 2001.
- [19] N. Ghani, S. Dixit, and T.-S. Wang, "On IP-over-WDM integration," *IEEE Commun. Mag.*, vol. 38, no. 3, pp. 72–84, Mar. 2000.
- [20] B. Rajagopalan, D. Pendarakis, D. Saha, R. S. Ramamoorthy, and K. Bala, "IP over optical networks: Architectural aspects," *IEEE Commun. Mag.*, vol. 38, no. 9, pp. 94–102, Sep. 2000.
- [21] B. Rajagopalan, J. Luciani, and D. Awduche. (2004, Mar.) IP over Optical Networks: a Framework. [Online]. Available: <http://www.rfc-archive.org/getrfc.php?rfc=3717>
- [22] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: an approach to high bandwidth optical WAN's," *IEEE Trans. Commun.*, vol. 40, no. 7, pp. 1171–1182, Jul. 1992.
- [23] (2004, Feb.) User network interface (UNI) 1.0 signaling specification, release 2. [Online]. Available: www.oiforum.com/public/documents/OIF-UNI-01.0-R2-RSVP.pdf
- [24] L. Mehedy, M. Bakaul, A. Nirmalathas, and E. Skafidas, "OFDM versus single carrier towards spectrally efficient 100 gb/s transmission with direct detection," *IEEE/OSA J. OPT. COMMUN. NETW.*, vol. 4, no. 10, pp. 779–789, Oct. 2012.

- [25] R. Dutta and G. N. Rouskas, "Traffic grooming in WDM networks: Past and future," *IEEE Network*, vol. 16, no. 6, pp. 46–56, Nov/Dec 2002.
- [26] S. Huang and R. Dutta, "Dynamic traffic grooming: The changing role of traffic grooming," *IEEE Communications Surveys & Tutorials*, vol. 9, no. 1, pp. 32–50, Jan. 2007.
- [27] R. Dutta, A. E. Kamal, and G. N. Rouskas, *Traffic grooming for optical networks: foundations, techniques, and frontiers*. Springer, 2008.
- [28] G. Maier, A. Pattavina, S. D. Patre, and M. Martinelli, "Optical network survivability: protection techniques in the WDM layer," *Photonic Network Communications*, vol. 4, no. 3, pp. 251–269, 2002.
- [29] J.-P. Vasseur, M. Pickavet, and P. Demeester, *Network recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS*. USA: Elsevier, 2004.
- [30] C. S. Ou and B. Mukherjee, *Survivable Optical WDM Networks*. Springer, 2005.
- [31] D. Zhou and S. Subramaniam, "Survivability in optical networks," *IEEE Network*, vol. 14, no. 6, pp. 16–23, November/December 2000.
- [32] S. Ramamurthy, L. Sahasrabudhe, and B. Mukherjee, "Survivable WDM mesh networks," *IEEE/OSA J. Lightw. Technol.*, vol. 21, no. 4, pp. 870–883, Apr. 2003.
- [33] W. Yao and B. Ramamurthy, "Survivable traffic grooming with path protection at the connection level in WDM mesh networks," *IEEE/OSA J. Lightw. Technol.*, vol. 23, no. 10, pp. 2846–2853, Oct. 2005.
- [34] H. Luo, L. Li, H. Yu, and S. Wang, "Achieving shared protection for dynamic multicast sessions in survivable mesh WDM networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 9, pp. 83–95, Dec. 2007.
- [35] L. Long and A. E. Kamal, "Tree-based protection of multicast services in WDM mesh networks," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*. IEEE, Nov. 2009, pp. 1–6.
- [36] T. Panayiotou, G. Ellinas, and N. Antoniadis, "Segment-based protection of multicast connections in metropolitan area optical networks with quality-of-transmission considerations," *IEEE/OSA J. OPT. COMMUN. NETW.*, vol. 4, no. 9, pp. 692–702, Sep. 2012.
- [37] B. Mukherjee, *Optical WDM Networks*. New York, USA: Springer Science+Business Media, Inc., 2006.
- [38] J. Y. Wei, "Advances in the management and control of optical internet," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 4, pp. 768–785, May 2002.

- [39] K. J. Barker, A. Benner, R. Hoare, A. Hoisie, A. K. Jones, D. K. Kerbyson, D. Li, R. Melhem, R. Rajamony, E. Schenfeld, S. Shao, C. Stunkel, and P. Walker, "On the feasibility of optical circuit switching for high performance computing systems," in *IEEE Computer Society 2005 ACM/IEEE conference on Supercomputing*. ACM/IEEE, 2005, p. 16.
- [40] S. Melle, D. Perkins, and C. Villamizar, "Network cost savings from router bypass in IP over WDM core networks," in *Proc.OFC/NFOEC '2008*, Feb. 2008.
- [41] X. Cao, V. Anand, Y. Xiong, and C. Qiao, "A study of waveband switching with multilayer multigranular optical cross-connects," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 7, pp. 1081–1095, Sep. 2003.
- [42] J. T. Moy, *OSPF: Anatomy Of An Internet Routing Protocol*. Addison-Wesley, 1998.
- [43] R. Wittmann, *Multicast Communication: Protocols and Applications*. Morgan Kaufmann, 2001.
- [44] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.
- [45] L. H.Sahasrabuddhe and B. Mukherjee, "Multicast routing algorithms and protocols: a tutorial," *IEEE Network*, vol. 14, no. 1, pp. 90–102, January/February 2000.
- [46] L. H.Sahasrabuddhe and B. Mukherjee, "Light-trees: Optical multicasting for improved performance in wavelength-routed networks," *IEEE Commun. Mag.*, vol. 37, no. 2, pp. 67–73, Feb. 1999.
- [47] W. Hu and Q. Zeng, "Multicasting optical cross connects employing splitter-and-delivery switch," *IEEE Photonics Technol. Lett.*, vol. 10, no. 7, pp. 970–972, Jul. 1998.
- [48] M. Ali and J. S. Deogun, "Power-efficient design of multicast wavelength-routed networks," *IEEE J. Sel. Areas. Commun.*, vol. 18, no. 10, pp. 1852–1862, Oct. 2000.
- [49] R. Varrazza, I. B. Djordjevic, and S. Yu, "Active vertical-coupler-based optical crosspoint switch matrix for optical packet-switching applications," *IEEE/OSA J. Lightwave. Technol.*, vol. 22, no. 9, pp. 2034–2042, Sep. 2004.
- [50] M. Ali and J. S. Deogun, "Cost-effective implementation of multicasting in wavelength-routed networks," *IEEE J. Lightwave Technol.*, vol. 18, no. 12, pp. 1628–1638, Dec. 2000.
- [51] X. Qin and Y. Yang, "A cost-effective construction for WDM multicast switching networks," in *Proc. ICC '2002*, vol. 5, Apr. 2002, pp. 2902–2906.
- [52] G. N. Rouskas, "Optical layer multicast: Rationale, building blocks, and challenges," *IEEE Network*, vol. 17, no. 1, pp. 60–65, January/February 2003.

- [53] S. L. Hakimi, "Steiner's problem in graphs and its implications," *Networks*, vol. 1, no. 2, pp. 113–133, 1971.
- [54] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2001.
- [55] L. Kou, G. Markowsky, and L. Berman, "A fast algorithm for steiner trees," *Acta informatica*, vol. 15, no. 2, pp. 141–145, 1981.
- [56] H. Takahashi and A. Matsuyama, "An approximate solution for the steiner problem in graphs," *Math. Japonica*, vol. 24, no. 6, pp. 573–577, 1980.
- [57] R. C. Prim, "Shortest connection networks and some generalizations," *Bell system technical journal*, vol. 36, no. 6, pp. 1389–1401, 1957.
- [58] M. Arikawa, T. Okamoto, M. Muraki, D. Ogasahara, E. L. T. de Gabory, T. Ito, and K. Fukuchi, "Transmission of a 127 Gb/s PM-QPSK signal over a 3350 km SMF-only line with chromatic dispersion compensation using real-time DSP [invited]," *IEEE/OSA J. OPT. COMMUN. NETW.*, vol. 4, no. 11, pp. B161–B167, Nov. 2012.
- [59] A. L. Chiu and E. H. Modiano, "Traffic grooming algorithms for reducing electronic multiplexing costs in WDM ring networks," *IEEE/OSA J. Lightwave Technol.*, vol. 18, no. 1, pp. 2–12, Jan. 2000.
- [60] P.-J. Wan, G. Calinescu, and O. Frieder, "Grooming of arbitrary traffic in SONET/WDM BLSRs," *IEEE J. Sel. Areas. Commun.*, vol. 18, no. 10, pp. 1995–2003, Oct. 2000.
- [61] X. Zhang and C. Qiao, "An effective and comprehensive approach for traffic grooming and wavelength assignment in SONET/WDM rings," *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 608–617, Oct. 2000.
- [62] B. Mukherjee, D. Banerjee, S. Ramamurthy, and A. Mukherjee, "Some principles for designing a wide-area WDM optical network," *IEEE/ACM Transactions on Networking*, vol. 4, no. 5, pp. 684–696, Oct. 1996.
- [63] K. Zhu and B. Mukherjee, "A review of traffic grooming in WDM optical networks: Architectures and challenges," *Optical Networks Magazine*, vol. 4, no. 2, pp. 55–64, 2003.
- [64] S. Thiagarajan and A. K. Somani, "Capacity fairness of WDM networks with grooming capabilities," *Optical Networks Magazine*, vol. 2, no. 3, pp. 24–31, 2001.
- [65] A. Khalil, A. Hadjiantonis, C. M. Assi, A. Shami, G. Ellinas, and M. Ali, "Dynamic provisioning of low-speed unicast/multicast traffic demands in mesh-based WDM optical networks," *IEEE/OSA J. Lightw. Technol.*, vol. 24, no. 2, pp. 681–692, Feb. 2006.

- [66] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," *Optical Networks*, vol. 1, no. 1, pp. 47–60, Jan. 2000.
- [67] A. E. Kamal, "Algorithms for multicast traffic grooming in WDM mesh networks," *IEEE Commun. Mag.*, vol. 44, no. 11, pp. 97–105, Nov. 2006.
- [68] H. Zhu, H. Zang, K. Zhu, and B. Mukherjee, "A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks," *IEEE/ACM Trans. Netw.*, vol. 11, no. 2, pp. 285–299, Apr. 2003.
- [69] Y. Wang and B. Ramamurthy, "A link bundled auxiliary graph model for constrained dynamic traffic grooming in WDM mesh networks," *IEEE J. Sel. Areas. Commun.*, vol. 23, no. 8, pp. 1542–1555, Aug. 2005.
- [70] B. Chen, G. N. Rouskas, and R. Dutta, "On hierarchical traffic grooming in WDM networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 16, no. 5, pp. 1226–1238, Oct. 2008.
- [71] Y. Ye, C. Assi, S. Dixit, and M. A. Ali, "A simple dynamic integrated provisioning/protection scheme in IP over WDM networks," *IEEE Commun. Mag.*, vol. 39, no. 11, pp. 174–182, Nov. 2001.
- [72] X. Niu, W. D. Zhong, G. Shen, and T. H. Cheng, "Connection establishment of label switched paths in IP/MPLS over optical networks," *Photonic Network Communication*, vol. 6, no. 1, pp. 33–41, Jul. 2003.
- [73] W. Zhong, X. Niu, B. Chen, S. K. Bose, and Y. Jin, "Performance comparison of overlay and peer models in IP/MPLS over optical networks," *Photonic Network Communication*, vol. 9, no. 1, pp. 121–131, Jan. 2005.
- [74] T. Ye, Y. Su, K. Zhu, L. Leng, Q. Zeng, and Y. Jin, "SLEA: A novel scheme for routing in overlay IP/WDM networks," *IEEE/OSA J. Lightwave Technol.*, vol. 23, no. 10, pp. 2934–2944, Oct. 2005.
- [75] Q. Zheng and G. Mohan, "An efficient dynamic protection scheme in integrated IP/WDM networks," in *Proc. IEEE ICC'2003*, vol. 2, May 2003, pp. 1494–1498.
- [76] S. Koo, G. Sahin, and S. Subramaniam, "Dynamic LSP routing in IP/MPLS over WDM networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 12, pp. 45–55, Dec. 2006.
- [77] X. Yu, G. Xiao, and T.-H. Cheng, "Historical data learning based dynamic LSP routing for overlay IP over WDM networks," in *Proc. IEEE ICICS '2009*, Dec. 2009, pp. 1–5.
- [78] G. V. Chowdhary, C. Siva, and R. Murthy, "Dynamic multicast traffic engineering in WDM groomed mesh networks," in *Proc. IEEE Broadnets '2004*, 2004.

- [79] B. Chen, S. K. Bose, and W.-D. Zhong, "Applying saturated cut method in dynamic multicast traffic grooming in IP/MPLS over WDM mesh networks," in *2005 Fifth International Conference on Information, Communications and Signal Processing*, Dec. 2005, pp. 1207–1211.
- [80] C. Lu, X. Nie, S. Wang, and L. Li, "Efficient dynamic multicast traffic grooming algorithm on WDM networks," in *Proc. SPIE Int Soc Opt Eng*, vol. 6022, Nov. 2005, pp. 602 230–1–602 230–10.
- [81] R. Lin, W.-D. Zhong, S. K. Bose, and M. Zukerman, "Light-tree configuration for multicast traffic grooming in WDM mesh networks," *Photonic Netw. Commun.*, vol. 20, no. 2, pp. 151–164, Oct. 2010.
- [82] X. Huang, F. Farahmand, and J. P. Jue, "Multicast traffic grooming in wavelength-routed WDM mesh networks using dynamically changing light-trees," *IEEE J. Lightwave Technol.*, vol. 23, no. 10, pp. 3178–3187, Oct. 2005.
- [83] L. Liao, H. Wen, L. Li, and S. Wang, "New dynamic grooming approaches for multicast traffic in meshed wavelength-division-multiplexing networks," *SPIE Opt. Eng.*, vol. 46, no. 2, pp. 25 004–1–25 004–7, Feb. 2007.
- [84] J. E. Sierra, L. F. Caro, F. Solano, J. L. Marzo, R. Fabregat, and Y. Donoso, "Dynamic unicast/multicast traffic grooming using S/G light-tree in WDM networks," in *Proc. SPECTS 2008*, Edinburgh, UK, Jun. 2008, pp. 263–268.
- [85] S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks. II - restoration," in *Proc. IEEE International Conference on Communications '1999*, vol. 3, Jun. 1999, pp. 2023 – 2030.
- [86] S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks. part I-protection," in *Proc. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies INFOCOM '99*, vol. 2, Mar. 1999, pp. 744 – 751.
- [87] H. Drid, B. Cousin, M. Molnar, and S. Lahoud, "A survey of survivability in multi-domain optical networks," *Computer Communications*, vol. 33, no. 8, pp. 1005–1012, May 2010.
- [88] N. K. Singhal, L. H. Sahasrabudde, and B. Mukherjee, "Provisioning of survivable multicast sessions against single link failures in optical WDM mesh networks," *IEEE/OSA J. Lightw. Technol.*, vol. 21, no. 11, pp. 2587–2594, Nov. 2003.
- [89] H. Luo, L. Li, and H. Yu, "Algorithm for protecting light-trees in survivable mesh wavelength-division-multiplexing networks," *Journal of Optical Networking*, vol. 5, no. 12, pp. 1071–1083, Dec. 2006.
- [90] X. Wang, L. Guo, L. Pang, J. Du, and F. Jin, "Segment protection algorithm with load balancing for multicasting WDM mesh networks," in *10th International Conference on Advanced Communication Technology 2008. (ICACT 2008.)*, vol. 3. IEEE, 2008, pp. 2013–2016.

- [91] F. Zhang and W.-D. Zhong, "Performance evaluation of optical multicast protection approaches for combined node and link failure recovery," *IEEE/OSA J. Lightw. Technol.*, vol. 27, no. 18, pp. 4017–4025, Sep. 2009.
- [92] C. S. Ou, K. Zhu, J. Zhang, H. Zhu, B. Mukherjee, H. Zang, and L. Sahasrabudde, "Traffic grooming for survivable WDM networks: dedicated protection [Invited]," *IEEE/OSA Journal of Optical Networking*, vol. 3, no. 1, pp. 50–74, Jan. 2004.
- [93] C. Ou, K. Zhu, H. Zang, L. H. Sahasrabudde, and B. Mukherjee, "Traffic grooming for survivable WDM networks-shared protection," *IEEE J. Sel. Areas. Commun.*, vol. 21, no. 9, pp. 1367–1383, Nov. 2003.
- [94] N. K. Singhal and B. Mukherjee, "Protecting multicast sessions in WDM optical mesh networks," *IEEE/OSA J. Lightw. Technol.*, vol. 21, no. 4, pp. 884–892, Apr. 2003.
- [95] P. Leelarusmee, C. Boworntummarat, and L. Wuttisittikulkij, "Design and analysis of five protection schemes for preplanned recovery in multicast WDM networks," in *IEEE/Sarnoff Symposium on Advances in Wired and Wireless Communication 2004*, 2004, pp. 167–170.
- [96] A. Khalil, A. Hadjiantonis, G. Ellinas, and M. Ali, "Dynamic provisioning of survivable heterogeneous multicast and unicast traffic in WDM networks," in *IEEE International Conference on Communications, 2006. ICC'06.*, vol. 6. IEEE, 2006, pp. 2465–2470.
- [97] L. Liao, L. Li, and S. Wang, "Multicast protection scheme in survivable WDM optical networks," *Journal of Network and Computer Applications*, vol. 31, no. 3, pp. 303–316, Aug. 2008.
- [98] F. Zhang, W.-D. Zhong, and Y. Jin, "Optimizations of p -cycle-based protection of optical multicast sessions," *IEEE/OSA J. Lightwave. Technol.*, vol. 26, no. 19, pp. 3298–3306, Oct. 2008.
- [99] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.
- [100] C. Lu and L. Li, "Dynamic multicast traffic grooming for survivable WDM mesh networks," in *Proc. International Conference on Communications, Circuits and Systems 2008*, May 2008, pp. 567–571.
- [101] J. Vasseur, J. Le Roux, and Y. Ikejiri, "A set of monitoring tools for path computation element (PCE)-based architecture," RFC 5886, June, Tech. Rep., 2010.
- [102] B. Martini, F. Baroncelli, P. Castoldi, A. Muchanga, and L. Wosinska, "The service oriented optical network (SOON) project," in *Proc. 2005 7th International Conference on Transparent Optical Networks, 2005, IEEE ICTON '2005*, Jul. 2005, pp. 429–432.

- [103] G. S. Z. amd Valerio Martini, Y. Qin, E. Escalona, R. Nejabati, D. Simeonidou, F. Baroncelli, B. Martini, K. Torkmen, and P. Castoldi, "Service-oriented multigranular optical network architecture for clouds," *J. Opt. Commun. Netw.*, vol. 2, no. 10, pp. 883–891, Oct. 2010.
- [104] X. Yu, G. Xiao, and T.-H. Cheng, "Historical data learning based dynamic LSP routing for overlay IP over WDM networks," *Optical Fiber Technology*, vol. 19, no. 4, pp. 308–319, Aug. 2013.
- [105] G. Shen and W. D. Grover, "Dynamic path-protected service provisioning in optical transport networks with a limited number of add/drop ports and transmitter tunability," *IEEE J. Sel. Areas. Commun.*, vol. 25, no. 6, pp. 121–134, Aug. 2007.
- [106] E. Q. V. Martins and M. M. B. Pascoal. (2000) A new implementation of Yen's ranking loopless paths algorithm. [Online]. Available: <http://citeseer.ist.psu.edu/martins00new.html>
- [107] A. Giorgetti, N. Sambo, I. Cerutti, N. Andriolli, and P. Castoldi, "Label preference schemes for lightpath provisioning and restoration in distributed GMPLS networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 27, no. 6, pp. 688–697, Mar. 2009.
- [108] E. A. Doumith and M. Gagnaire, "Impact of traffic predictability on WDM EXC/OXC Network performance," *IEEE J. Sel. Areas. Commun.*, vol. 25, no. 5, pp. 895–904, Jun. 2007.
- [109] R. Malli, X. Zhang, and C. Qiao, "Benefit of multicasting in all-optical networks," in *SPIE Proc. Conf. All-optical Networking*, Rio de Janeiro, Brazil, Nov. 1998, pp. 209–220.
- [110] H. V. Madhyastha, G. V. Chowdhary, N. Srinivas, and C. S. R. Murthy, "Grooming of multicast sessions in metropolitan WDM ring networks," *Computer Networks*, vol. 49, no. 4, pp. 561–579, Nov. 2005.
- [111] N. K. Singhal, L. H. Sahasrabuddhe, and B. Mukherjee, "Optimal multicasting of multiple light-trees of different bandwidth granularities in a WDM mesh network with sparse splitting capabilities," *IEEE/ACM Trans. Netw.*, vol. 14, no. 5, pp. 1104–1117, Oct. 2006.
- [112] R. Ul-Mustafa and A. E. Kamal, "Design and provisioning of WDM networks with multicast traffic grooming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 4, pp. 37–53, Apr. 2006.
- [113] Y. Zhu, Y. Jin, W. Son, W. Guo, W. Hu, W.-D. Zhong, and M.-Y. Wu, "Multicast flow aggregation in IP over optical networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 5, pp. 1011–1021, Jun. 2007.
- [114] S. S. W. Lee, H.-H. Yen, and A. Chen, "Traffic grooming for IP multicast over WDM networks using light-path and light-tree schemes," in *Proc. of the Ninth International Conference on Networks ICN 2010*, Menuires, France, Apr. 2010, pp. 291–293.

- [115] X. Yu, G. Xiao, and T.-H. Cheng, "Dynamic multicast traffic grooming in optical wdm mesh networks: Lightpath vs light-tree," *IEEE/OSA J. OPT. COMMUN. NETW.*, vol. 5, no. 8, pp. 870–880, Aug. 2013.
- [116] C. Ou, K. Zhu, H. Zang, L. H. Sahasrabuddhe, and B. Mukherjee, "Traffic grooming for survivable WDM networks-shared protection," *IEEE J. Sel. Areas. Commun.*, vol. 21, no. 9, pp. 1367–1383, Nov. 2003.
- [117] X. Yu, G. Xiao, and T.-H. Cheng, "Connection-level segment shared protection for dynamic multicast traffic grooming," in *Proc. IEEE ICICS '2013*, Dec. 2013.
- [118] D. Xu, Y. Xiong, and C. Qiao, "Novel algorithms for shared segment protection," *IEEE J. Sel. Areas. Commun.*, vol. 21, no. 8, pp. 1320–1331, Oct. 2003.
- [119] P.-H. Ho, J. Tapolcai, and T. Cinkler, "Segment shared protection in mesh communications networks with bandwidth guaranteed tunnels," *IEEE/ACM Transactions on Networking*, vol. 12, no. 6, pp. 1105–1118, Dec. 2004.
- [120] N. K. Singhal, C. Ou, and B. Mukherjee, "Cross-sharing vs. self-sharing trees for protecting multicast sessions in mesh networks," *Computer Networks*, vol. 50, no. 2, pp. 200–206, 2006.