Control and management of multicast-capable IP over WDM networks

Chen, Sheng

2010

Chen, S. (2010). Control and management of multicast-capable IP over WDM networks.
Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/38587>

<https://doi.org/10.32657/10356/38587>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**CONTROL AND MANAGEMENT OF MULTICAST-
CAPABLE IP OVER WDM NETWORKS**

CHEN SHENG

SCHOOL OF ELECTRICAL AND ELECTRONIC ENGINEERING

2010

CONTROL AND MANAGEMENT OF MULTICAST- CAPABLE IP OVER WDM NETWORKS

CHEN SHENG

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

2010

Acknowledgment

This thesis concludes my study at Nanyang Technological University. I would not have gone so far without the support given by many individuals.

I wish to thank my first thesis supervisor, Prof. Poo Gee Swee, who retired from NTU in 2008. I thank him for his continuous encouragement and patient guidance during the early part of my PhD study. Those valuable discussions and conversations with him helped me grow and made me more confident to overcome the many obstacles in the path leading towards the completion of my study.

I also wish to thank Prof. Cheng Tee Hiang, who took over from Prof. Poo Gee Swee as my supervisor when Prof Poo left NTU. His enthusiasm, suggestion on how to consolidate and improve the research work I have done, and relentless effort in rendering assistance encouraged me to complete my PhD degree.

I would like to thank Prof. Ma Maode, Prof. Zhong Wende, Prof. Xiao Gaoxi, Prof. Shum Ping for their generous help and advices during the past several years. I wish to thank all of my research group colleagues, Dr. Zhou Yinzhu, Dr. Yin Yongning, Dr. Wang Haibo, Mr. He Xuzhou, Dr. Kelvin Yong and Dr. Sun Jianewei, for the many thoughtful discussions and collaborations. I also wish to thank Dr. Chen Bin, Dr. Zhang Feng for the help on my simulation works.

I gratefully acknowledge the financial support given by the Nanyang Technological University, which provided me a three-year research scholarship and opportunities to be the teaching assistant for a number of courses. I would like to thank all the professors, colleagues, and supporting technicians from Communication Lab II, InfoComm Research Lab and Network Technology Research Centre, for creating such a friendly working environment with enormous

ACKNOWLEDGMENT

resources.

I am greatly indebted to my beloved wife, Cai Chengjie, who has stood by me through thick and thin, and shared my happiness, frustration, successes and failures. Having her by my side makes this all meaningful.

Last but not least, I wish to thank my parents and parents-in-law. Their unwavering love and encouragement have been and will always be the source of my courage to undertake new challenges in future.

Summary

Wavelength-routed networks require control and management protocols to perform routing and wavelength assignment (RWA) functions, and exchange signaling information for resource reservations and other network operations. In this thesis, the problems in connection with the control and management of multicast-capable IP over WDM networks are studied. The focus is on investigating the performance of various protocols for dynamic lightpath provisioning, and developing efficient algorithms to improve the performance of multicast-capable IP over WDM networks, in which dynamic traffic grooming, optical multicasting, placement of wavelength converters and light splitters are supported for both unicast and multicast IP traffic.

Performances of various control and management protocols for dynamic lightpath provisioning have been studied. The results show that the centralized global information-based link state approach yields the lowest request blocking probability, low control message overhead and low connection setup time. However, when network size is large or the average event processing delay is long, the centralized approach will not be scalable.

Next, the signaling effect on dynamic traffic grooming in multicast-capable IP over WDM networks is investigated. Network performance is found to be poor when we incorporate the signaling requirement in a number of existing dynamic traffic grooming (DTG) algorithms. In order to mitigate the signaling effects, a new scheme called DTG-PRL is proposed to handle the DTG problem. The DTG-PRL is divided into two steps: (1) pre-reserve some lightpaths based on statistical data; and (2) dynamically groom traffic based on the established virtual topology. ILP

formulations and heuristic algorithms have been developed for DTG-PRL for both unicast and multicast IP traffic models. The simulation results show that DTG-PRL outperforms DTG algorithms in request blocking probability, network resource utilization, connection setup time, and control overhead efficiency.

Dynamic multicast routing and wavelength assignment (MC-RWA) is studied by considering the heterogeneity of network structures. A new generic graph model which is able to support various node architectures and heterogeneous network structures is proposed for the studying of the MC-RWA problem. Based on this graph model, dynamic MC-RWA problems for various light splitting and wavelength conversion scenarios are solved by means of an efficient multicast tree algorithm. This graph model provides a generic platform which can be used to study different aspects of the dynamic MC-RWA problem and can be adapted for studying other similar problems.

By extending the work on the generic graph model, placement of wavelength converters and light splitters is studied. In the study, both sparse and limited wavelength conversion and light splitting are considered in a holistic way. Four proposed algorithms, namely, WCF-I, WCF-D, LSF-I and LSF-D, for finding the optimal configuration of wavelength converters and light splitters are studied. From the results, it can be concluded that the presence of wavelength converters contributes more to improving the cost efficiency than the presence of light splitters when traffic load is not high; determining where to place wavelength converters and light splitters is far more important than determining the capability of individual wavelength converters and light splitters.

Contents

ACKNOWLEDGMENT	I
SUMMARY	III
CONTENTS	V
LIST OF FIGURES.....	VIII
LIST OF TABLES.....	XII
LIST OF ACRONYMS	XIII
CHAPTER 1 INTRODUCTION	1
1.1 Background	2
1.1.1 IP over WDM networks.....	2
1.1.2 Multicast in IP over WDM networks	4
1.2 Motivation.....	6
1.3 Objectives	9
1.4 Contributions.....	11
1.5 Organizations	13
CHAPTER 2 LITERATURE REVIEW.....	15
2.1 Network models and issues in IP over WDM networks	15
2.2 Routing and wavelength assignment	20
2.3 Lightpath establishment.....	24
2.4 Traffic grooming	28
2.5 Placement of wavelength converters and light splitters.....	30
2.6 Summary	33
CHAPTER 3 PERFORMANCE STUDY ON CONTROL PROTOCOLS FOR DYNAMIC LIGHTPATH PROVISIONING.....	35
3.1 Network model.....	37
3.2 Local information-based fixed and fixed-alternate link routing approach.....	38
3.3 Distributed global information-based link state approach	41
3.3.1 Parallel signaling mechanism.....	41
3.3.2 Backward-reservation sequential signaling mechanism.....	42
3.3.3 Forward-reservation sequential signaling mechanism	44
3.4 Centralized global information-based link state approach.....	46
3.5 Simulation and performance analysis	48
3.5.1 Topology and assumptions	48

3.5.2	Numerical results	49
3.5.3	More results on processing delay.....	56
3.6	Summary.....	61
CHAPTER 4 SIGNALING EFFECT ON DYNAMIC TRAFFIC GROOMING IN IP OVER WDM NETWORKS		62
4.1	Effect of signaling on DTG in IP over WDM networks.....	64
4.1.1	Signaling procedure in dynamic traffic grooming	64
4.1.2	Signaling algorithms for IP bandwidth allocation and lightpath provisioning.....	70
4.1.3	Problems of DTG due to the signaling	72
4.1.4	Signaling influence on DTG	74
4.2	Dynamic traffic grooming with pre-reserved lightpaths	77
4.3	ILP formulation for DTG-PRL.....	79
4.3.1	ILP Formulation for unicast IP traffic model.....	81
4.3.2	ILP Formulation for multicast IP traffic model	83
4.4	Heuristic approach for DTG-PRL	87
4.4.1	MWU algorithm.....	88
4.4.2	MWU-M algorithm.....	88
4.5	Simulation and numerical results	89
4.5.1	Simulation description	89
4.5.2	Topology and assumptions.....	90
4.5.3	Numerical results	93
4.6	Summary.....	109
CHAPTER 5 GENERIC GRAPH MODEL FOR DYNAMIC OPTICAL MULTICAST RWA.....		111
5.1	Definitions and construction of Generic Graph Model	115
5.2	Creation of wavelength conversion edges	119
5.2.1	Sparse wavelength conversion.....	120
5.2.2	Limited wavelength conversion.....	123
5.2.3	Hybrid wavelength conversion	124
5.3	Creation of light splitting edges	125
5.4	MC-RWA based on the Generic Graph Model.....	128
5.4.1	Cost model	128
5.4.2	MC-RWA algorithm.....	130
5.5	Numerical results.....	132
5.6	Summary.....	138

CHAPTER 6 PLACEMENT OF WAVELENGTH CONVERTERS AND LIGHT SPLITTERS USING THE GGM.....	140
6.1 Objective of placement problem.....	142
6.2 Algorithms for the placement of wavelength converter and light splitter	143
6.2.1 Greedy algorithm.....	143
6.2.2 Place wavelength converter first (WCF) strategies.....	144
6.2.3 Place light splitter first (LSF) strategies.....	149
6.3 Complexity analysis and adaptation of algorithms for other optimization criteria	150
6.3.1 Complexity analysis	150
6.3.2 Adaptation of algorithms for other optimization criteria	151
6.4 Simulation and numerical results.....	152
6.4.1 Topology and assumptions	152
6.4.2 Numerical results.....	153
6.5 Further discussions on cost efficiency	170
6.5.1 New definition of cost efficiency	170
6.5.2 Numerical results.....	171
6.6 Summary	182
CHAPTER 7 CONCLUSION AND FUTURE RECOMMENDATIONS.....	184
7.1 Research contributions.....	185
7.2 Recommendations for future research	187
AUTHOR'S PUBLICATIONS	190
APPENDIX NETWORK TOPOLOGIES	191
BIBLIOGRAPHY	194

List of Figures

Figure 2.1	IP over WDM network: Overlay model	16
Figure 2.2	IP over WDM network: Peer model	16
Figure 2.3	Classification of IP over WDM network issues	19
Figure 3.1	GMPLS optical network model.....	38
Figure 3.2	Local information-based fixed and fixed-alternate link routing approaches	40
Figure 3.3	Timing sequence of parallel signaling approach	41
Figure 3.4	Timing sequence of backward-reservation sequential signaling approach	43
Figure 3.5	Timing sequence of forward-reservation sequential signaling approach	44
Figure 3.6	The network model with a central manager	47
Figure 3.7	Time sequence of the centralized global information-based link state approach.....	48
Figure 3.8	RBP vs. TL (ETD)	51
Figure 3.9	RBP vs. TL (PTD)	51
Figure 3.10	CST vs. TL (ETD)	53
Figure 3.11	CST vs. TL (PTD)	53
Figure 3.12	CMO vs. TL (ETD)	55
Figure 3.13	CMO vs. TL (PTD).....	55
Figure 3.14	RBP vs. TL with different processing delay – <i>Fix</i> and <i>FixAlt</i>	57
Figure 3.15	RBP vs. TL with different processing delay – <i>Backward</i> and <i>Forward</i>	58
Figure 3.16	RBP vs. TL with different processing delay – <i>Parallel</i> and <i>Central</i>	58
Figure 4.1	Network model for IP over WDM network.....	65
Figure 4.2	Signaling procedure for DTG without available lightpath for new IP request.....	66
Figure 4.3	Signaling procedure for DTG with available lightpath for new IP request	67
Figure 4.4	Procedure algorithm of processing multicast IP request	68
Figure 4.5	Timing sequence of resource reservation signaling procedure	71
Figure 4.6	Timing sequence of resource release signaling procedure.....	72
Figure 4.7	IBBP vs. TL, Simple6 network, DTG-B with/without signaling.....	75

Figure 4.8	IBBP vs. TL, NSFNet network, DTG-B with/without signaling	75
Figure 4.9	NRU vs. TL, Simple6 network, DTG-B with/without signaling	76
Figure 4.10	NRU vs. TL, NSFNet network, DTG-B with/without signaling	76
Figure 4.11	Comparison between ILP-Unicast and MWU: IBBP vs. TL	94
Figure 4.12	Comparison between ILP-Unicast and MWU: NRU vs. TL	94
Figure 4.13	Comparison between ILP-Unicast and MWU: ICST vs. TL	95
Figure 4.14	Comparison between ILP-Unicast and MWU: CME vs. TL	95
Figure 4.15	Comparison between ILP-Multicast and MWU-M: MBBP vs. TL.....	96
Figure 4.16	Comparison between ILP-Multicast and MWU-M: MNRU vs. TL.....	96
Figure 4.17	Comparison between ILP-Multicast and MWU-M: MCST vs. TL.....	97
Figure 4.18	Comparison between ILP-Multicast and MWU-M: MCME vs. TL.....	97
Figure 4.19	Unicast performance comparison on Simple6: IBBP vs. TL	98
Figure 4.20	Unicast performance comparison on NSFNet: IBBP vs. TL	99
Figure 4.21	Unicast performance comparison on Simple6: NRU vs. TL	99
Figure 4.22	Unicast performance comparison on NSFNet: NRU vs. TL	100
Figure 4.23	Unicast performance comparison on Simple6: ICST vs. TL	100
Figure 4.24	Unicast performance comparison on NSFNet: ICST vs. TL	101
Figure 4.25	Unicast performance comparison on Simple6: CME vs. TL	101
Figure 4.26	Unicast performance comparison on NSFNet: CME vs. TL	102
Figure 4.27	Multicast performance comparison on Simple6: MBBP vs. TL.....	105
Figure 4.28	Multicast performance comparison on NSFNet: MBBP vs. TL	106
Figure 4.29	Multicast performance comparison on Simple6: MNRU vs. TL.....	106
Figure 4.30	Multicast performance comparison on NSFNet: MNRU vs. TL	107
Figure 4.31	Multicast performance comparison on Simple6: MCST vs. TL	107
Figure 4.32	Multicast performance comparison on NSFNet: MCST vs. TL	108
Figure 4.33	Multicast performance comparison on Simple6: MCME vs. TL.....	108
Figure 4.34	Multicast performance comparison on NSFNet: MCME vs. TL.....	109
Figure 5.1	(a) Physical topology of a given network $G(V_4, E_6, \Lambda_2)$; (b) Derived generic graph $G'(V'_{24}, E'_{36})$	117
Figure 5.2	(a) The multicast tree to be set up for the request $Req = \{A, \{C, D\}\}$ in G ; (b) The multicast tree to be set up for the request $Req = \{A, \{C, D\}\}$ in G'	119
Figure 5.3	Mode of (a) Convert-Switch; (b) Switch-Convert; (c)	

	Convert-Split; (d) Split-Convert.....	120
Figure 5.4	(a) The light tree found for request $r = \{A, \{C, D\}\}$ in G ; (b) Light tree setup in G' , Convert-Split mode, only 1 wavelength conversion edge is involved; (c) Light tree setup in G' , Split-Convert mode, 2 wavelength conversion edges are involved.	121
Figure 5.5	Create wavelength conversion edges in generic graph model.....	124
Figure 5.6	Create light splitting edges in generic graph model.	126
Figure 5.7	RBP for sparse splitting cases in different group sizes in NSFNet.....	133
Figure 5.8	RBP for sparse light splitting cases in different network topologies	134
Figure 5.9	RBP for limited light splitting cases in different network topologies	135
Figure 5.10	RBP for sparse conversion cases in different network topologies	136
Figure 5.11	RBP for limited conversion cases in NSFNet.....	137
Figure 5.12	RBP for Split-Convert and Convert-Split mode in NSFNet.....	138
Figure 6.1	Results of NSFNet ($T_r = 30Erl$) - variation of CE	158
Figure 6.2	Results of NSFNet ($T_r = 30Erl$) - variation of RBP.....	158
Figure 6.3	Results of NSFNet ($T_r = 30Erl$) - variation of AMTC	159
Figure 6.4	Cost Efficiency vs. Traffic Load – Mesh9.....	162
Figure 6.5	Cost Efficiency vs. Traffic Load – NSFNet.....	162
Figure 6.6	Request Blocking Probability vs. Traffic Load – Mesh9	163
Figure 6.7	Request Blocking Probability vs. Traffic Load – NSFNet	163
Figure 6.8	Average Multicast Tree Cost vs. Traffic Load – Mesh9.....	164
Figure 6.9	Average Multicast Tree Cost vs. Traffic Load – NSFNet.....	164
Figure 6.10	Comparison among optimal configurations – CE vs. Traffic Load.....	167
Figure 6.11	Comparison among optimal configurations – RBP vs. Traffic Load.....	168
Figure 6.12	Comparison among optimal configurations – AMTC vs. Traffic Load.....	169
Figure 6.13	Results of NSFNet ($T_r = 5Erl$) - variation of CE.....	175
Figure 6.14	Results of NSFNet ($T_r = 5Erl$) - variation of AMTC.....	175
Figure 6.15	Results of NSFNet ($T_r = 25Erl$) - variation of CE.....	176
Figure 6.16	Results of NSFNet ($T_r = 25Erl$) - variation of RBP	176

Figure 6.17	Results of NSFNet ($T_r = 25ErI$) - variation of AMTC	177
Figure 6.18	Cost Efficiency vs. Algorithms – Mesh9	179
Figure 6.19	Cost Efficiency vs. Algorithms – NSFNet.....	180
Figure 6.20	Request Blocking Probability vs. Algorithms – Mesh9.....	180
Figure 6.21	Request Blocking Probability vs. Algorithms – NSFNet.....	181
Figure 6.22	Average Multicast Tree Cost vs. Algorithms – Mesh9	181
Figure 6.23	Average Multicast Tree Cost vs. Algorithms – NSFNet	182
Figure A.1	Topology of Simple network with 6 nodes, 12 links. Each link in the network is composed of two fibers of opposite directions, and has the same physical lengths in 100 kilometers	191
Figure A.2	Topology of Mesh-torus network with 9 nodes, 18 links. Each link in the network is composed of two fibers of opposite directions, and has the same physical lengths in 100 kilometers.....	191
Figure A.3	Topology of New Jersey Local Access Transport Area network (NJ LATA) network with 11 nodes, 23 links. Each link in the network is composed of two fibers of opposite directions. The numbers next to the links denote the physical lengths in 1 kilometer. In simulation of Chapter 3, centralized global information-based link state approach, the central manager is located at node 4.	192
Figure A.4	Topology of COST239 network with 11 nodes, 26 links. Each link in the network is composed of two fibers of opposite directions.....	192
Figure A.5	Topology of NSFNet network with 14 nodes, 21 links. Each link in the network is composed of two fibers of opposite directions. The numbers next to the links denote the physical lengths in 100 kilometer.....	193

List of Tables

Table 4.1	ILP notations for unicast IP traffic model.....	81
Table 4.2	ILP formulation for unicast IP traffic model	82
Table 4.3	ILP notations for multicast IP traffic model	84
Table 4.4	ILP formulation for multicast IP traffic model	85
Table 4.5	Maximizing wavelength utilization algorithm	88
Table 4.6	Maximizing wavelength utilization for multicasting algorithm.....	89
Table 4.7	Common DTG algorithm for DTG-B and DTG-PRL.....	90
Table 5.1	States to manage switching / splitting edges	127
Table 5.2	MC-RWA algorithm.....	131
Table 6.1	WCF-I algorithm (Main algorithm).....	144
Table 6.2	Function Place_Device.....	145
Table 6.3	Function Adjust_Degree.....	146
Table 6.4	WCF-D algorithm (Main algorithm).....	148
Table 6.5	LSF-I algorithm (Main algorithm)	149
Table 6.6	LSF-D algorithm (Main algorithm).....	150
Table 6.7	Illustration of the steps and rounds, NSFNet14, Traffic load = 30Erl	154
Table 6.8	Configuration of wavelength converter and light splitter – Mesh9	160
Table 6.9	Configuration of wavelength converter and light splitter – NSFNet.....	160
Table 6.10	Illustration of the steps and rounds, NSFNet14, Traffic Load = 5Erl	172
Table 6.11	Illustration of the steps and rounds, NSFNet14, Traffic Load = 25Erl	173
Table 6.12	Configuration of wavelength converter and light splitter – Mesh9	178
Table 6.13	Configuration of wavelength converter and light splitter – NSFNet.....	178

List of Acronyms

ADM	Add-Drop Multiplexer
CST	Connection Setup Time
CME	Control Message Efficiency
CMO	Control Message Overhead
CR-LDP	Constraint-based Routing Label Distribution Protocol
DIR	Destination-Initiated Reservation
DLE	Dynamic Lightpath Establishment
DTG	Dynamic Traffic Grooming
DTG-B	Dynamic Traffic Grooming Basic
DTG-PRL	Dynamic Traffic Grooming with Pre-Reserved Lightpaths
DWDM	Dense Wavelength Division Multiplexer
DXC	Digital Cross-Connect
ETD	Exponential Traffic Distribution
FAXC	Forward After Cross-Connects
FBXC	Forward Before Cross-Connects
GGM	Generic Graph Model
GMPLS	Generalized Multiple Protocol Label Switching
HDTV	High Definition TV
IBBP	IP Bandwidth Blocking Probability
IETF	Internet Engineering Task Force
IIR	Intermediate-node Initiated Reservation
ILP	Integer Linear Programming
IP	Internet Protocol
LSA	Link State Advertisement
LSF-D	Light Splitter First with Decremental-placement
LSF-I	Light Splitter First with Incremental-placement
LSP	Label Switch Path
MBBP	Multicast Bandwidth Blocking Probability
MCST	Multicast Connection Setup Time
MCME	Multicast Control Message Efficiency
MC-OXC	Multicast Capable Optical Cross-Connect

LIST OF ACRONYMS

MC-RWA	Multicast Routing and Wavelength Assignment
MPLS	Multiple Protocol Label Switching
MPR	Multi-Path Routing
MNRU	Multicast Network Resource Utilization
NRU	Network Resource Utilization
OEO	Optical-Electronic-Optical
OLSP	Optical Label Switch Path
OSPF	Open Shortest Path First
OSPF-TE	Open Shortest Path First with Traffic Engineering
OXC	Optical Cross-Connect
PFBR	Probe Flooding-Based Routing
PTD	Pareto Traffic Distribution
RBP	Request Blocking Probability
RSVP	Resource Reservation Protocol
RSVP-TE	Resource Reservation Protocol with Traffic Engineering
RWA	Routing and Wavelength Assignment
SIR	Source-Initiated Reservation
SLE	Static Lightpath Establishment
SONET	Synchronous Optical Network
SaD	Splitter and Delivery
SPT	Shortest Path Tree
TE	Traffic Engineering
UNI	User Network Interface
VOD	Video-On-Demand
WCF-D	Wavelength Converter First with Decremental-placement
WCF-I	Wavelength Converter First with Incremental-placement
WDM	Wavelength Division Multiplexing

Chapter 1

Introduction

In the last decade, the pervasive use of Internet services has significantly increased the volume of Internet traffic. It is well recognized that optical Wavelength Division Multiplexing (WDM) networks will become the dominant transport infrastructure in future data networks. The inundation of such multimedia applications as video-streaming, video-conferencing and video-on-demand (VOD) in recent years makes network traffic more bandwidth intensive. Multicast techniques, which are able to transfer bandwidth intensive data streams from one source node to multiple destination nodes, are being developed and deployed to handle this type of traffic. There are many possible scenarios in the evolution of optical networks to support bandwidth intensive multicast applications. In the case of an IP over WDM network, one possible scenario is the provision of multicast support at the optical layer in addition to the existing IP multicast mechanism. In such a scenario, we envisage some of the optical layer equipment – such as wavelength routers and add/drop multiplexers – to be multicast capable and have wavelength conversion capability. The control and management of such a multicast-capable IP over WDM network will be very challenging and we aim to address such issues in this thesis. The rest of this chapter is organized as follows. Section 1.1 provides an overview of IP over WDM networks. Sections 1.2 to 1.5 cover motivations, objectives, major contributions, and organization of the thesis respectively.

1.1 Background

1.1.1 IP over WDM networks

A WDM network [1-3] is an optical network in which each optical fiber could support multiple data streams, each of which on a different wavelength. Technology advancements have made it possible for WDM networks to be designed in such a way that wavelengths can be switched and routed dynamically to set up end-to-end lightpaths on demand [4-6]. In the context of this thesis, a lightpath is an end-to-end optical circuit that does not go through any optical-electronic-optical (OEO) conversion [5]. In a wavelength-routed network, multi-gigabit-per-second up to terabit-per-second of data streams can be routed through optical switches and routers in the optical domain without going through OEO conversion. Wavelength-routed networks require a set of sophisticated control and management protocols to perform routing and wavelength assignment (RWA), reserve resources, and exchange of signaling information between optical network equipment.

A lightpath may span multiple fiber links, connected by optical routers and switches. To establish a lightpath, a route has to be selected and a wavelength on each span along the route has to be chosen. How the route is chosen and how the wavelength on each span of the route is assigned is collectively called the RWA problem [7-8]. If wavelength conversion is not facilitated by the network, RWA is subject to the wavelength continuity constraint, i.e., each lightpath must use the same wavelength on all the links traversed [6-8]. The RWA problem can be classified into the static lightpath establishment (SLE) problem and the dynamic lightpath establishment (DLE) problem. For SLE, connection requests are known in advance and one commonly used optimization objective is to set up lightpaths to carry all the connections such that the consumption of the network resource is

minimized. For DLE, a lightpath is established for each connection request when it arrives and released when the connection terminates. One commonly used optimization objective of DLE is to apply the RWA algorithm for each connection so as to minimize the request blocking probability (RBP).

Apart from RWA, signaling protocols are needed to set up and tear down connections, as well as to reserve the necessary resources. Two types of signaling protocols, namely forward and backward signaling protocols, have been investigated [9]. For the forward signaling protocol, the source starts to reserve the available wavelength channels along the route in the forward direction from the initiating node to the destination node by which an end-to-end lightpath is set up. When the reservation message reaches the destination node, the destination node chooses one wavelength channel among the wavelength channels reserved along the route and releases the rest. In the backward scheme, the source node sends a setup control message to the destination without reserving any resources. This setup control message will collect information about the availability of wavelength channels along the route. Upon receiving the setup control message, the destination node starts reserving wavelength channels along the route in the backward direction towards the source node.

There is an ongoing effort to develop the network control and management protocols for handling RWA and wavelength reservation. The Internet Engineering Task Force (IETF) is in the process of standardizing a set of distributed protocols within the generalized multiple protocol label switching (GMPLS) framework [10]. GMPLS is an extension of the multiprotocol label switching (MPLS) framework [11], which was developed as a standard to enable fast switching in IP networks. In particular, the Open Shortest Path First (OSPF) routing protocol, the

Constraint-based Routing Label Distribution Protocol (CR-LDP), and the Resource Reservation Protocol (RSVP) signaling protocol have been extended to support GMPLS. A summary of the proposed enhancements can be found in [12].

The data rate of each wavelength channel in today's commercial WDM systems is in the order of 2.5Gb/s to 10Gb/s (OC-48 to OC-192) and data rates at 40Gb/s (OC-768) is beginning to be commercially viable. In an IP over WDM network, typical bandwidth request of each IP connection is expected to be around OC-1, OC3 or OC-12, which is much lower than the wavelength capacity. In order to efficiently utilize the wavelength capacity, a number of independent lower-rate traffic streams can be multiplexed onto a single lightpath. This has given rise to the concept of traffic grooming [13], which is a technique for aggregating lower-data-rate traffic flows into a wavelength channel to optimize resource usage. Traffic grooming can be further classified into static traffic grooming and dynamic traffic grooming. The static traffic grooming problem is a special instance of the virtual topology design problem. Given the demand of low-rate circuit-like traffic between pairs of nodes, the problem is to assign traffic to wavelengths in such a way that minimizes the number of add-drop multiplexers (ADMs) used in the network. Due to the fast-changing dynamics of Internet traffic, static traffic grooming cannot efficiently accommodate changes in traffic, and this has led to the study of dynamic traffic grooming. Research in dynamic traffic grooming is mainly focusing on fast, on-line algorithms for reconfiguring the underlying optical network according to the IP traffic changes. Hence, it is very challenging to design RWA and signaling scheme for dynamic traffic grooming in IP over WDM networks.

1.1.2 Multicast in IP over WDM networks

Multicasting refers to point-to-multipoint communication in which there is one

source node and multiple destination nodes. There are many ways to implement multicasting. A simple but very inefficient way is to set up a unicast transport-layer connection between the source and each of the receivers and replicate the traffic over these unicast connections. This approach requires no explicit multicast support from the network layer since multicast is emulated using multiple point-to-point unicast connections. A second alternative is to provide explicit multicast support in the network. In this approach, a single IP packet is transmitted from the sending host, and then replicated at an IP router whenever it must be forwarded on multiple outgoing links in order to reach the receivers. Clearly, the second approach is more efficient in that only a single copy of an IP packet will ever traverse a link.

Multicasting has been extensively studied in traditional IP-centric electronic networks [14-16], and it has begun to attract the attentions of the optical networking community in recent years [17-19]. By utilizing optical devices such as light splitter, multicasting can be supported in the optical domain. A light splitter has the ability of splitting an incoming signal into n (≥ 2) outgoing signals, where n is known as the *fanout* of the light splitter.

Routing and wavelength assignment (RWA) needs to be expanded to multicast routing and wavelength assignment (MC-RWA) in order for wavelength-routed WDM networks to support multicasting [7]. The concept of light-tree, which is an extension of the concept of lightpath, has been proposed [17]. Similar to a lightpath, a light-tree is a light channel that originates from the source node and consists of a wavelength on each link that constitutes the light-tree. But unlike a lightpath, a light-tree spans multiple (instead of one) destination nodes. In short, a light-tree is a point-to-multipoint all-optical channel that may span multiple fiber links.

1.2 Motivation

In the control and management of multicast-capable IP over WDM optical networks, the establishment of lightpath/light-tree is one of the fundamental issues that needs to be resolved. When a set of connection setup requests are known in advance and lightpaths need to be set up to optimize certain performance objective, this is known as static lightpath establishment (SLE) problem. Generally the SLE problem can be formulated as an integer linear programming (ILP) problem [5-6], which in most cases is an *NP*-complete problem [5]. Due to the intensity of computation, ILP problems can only be solved computationally for small networks and sub-optimal heuristic methods [6-7] have to be used for larger networks. When the traffic pattern is dynamically changing, lightpaths have to be established and torn down dynamically and this is known as the dynamic lightpath establishment (DLE) problem [5, 7]. The DLE is also an *NP*-complete problem. Setting lightpaths in accordance with dynamic traffic changes not only involves RWA but also signaling for the reservation of the required network resources.

The RWA problem can be further divided into two separated sub-problems, i.e. routing problem and wavelength assignment problem. Various methods for solving the routing problem have been proposed [5, 7], which includes fixed routing, fixed-alternate routing and adaptive routing, etc. In the fixed routing scheme, a predetermined route for a given pair of source-destination nodes is always used. For fixed-alternate routing, a route from an ordered list of several predetermined routes will be used for a given pair of source-destination nodes. For adaptive routing, a route will be chosen dynamically for a given pair of source-destination nodes based on the current network traffic. The key issue in the wavelength assignment problem is how to select a wavelength for a given path when there are multiple wavelengths

satisfying the wavelength continuity constraint. Commonly used wavelength assignment methods are first-fit selection and random selection [7].

The part on signaling is concerned with network resource reservation and release. They are two commonly used methods for resource reservation, namely, the parallel method [20-21] and the sequential method [20, 22-23]. For the parallel method, reservation requests are sent out by the source node to all the other nodes along the selected route. For the sequential method, the reservation request passes through each node along the selected route in a hop-by-hop manner. The sequential method can be further divided into forward reservation and backward reservation.

With many varieties of algorithm and protocol available for RWA and signaling in dynamic lightpath provisioning, it is necessary to compare their performances in order to select the most appropriate ones. Motivated by this, we set out to compare the performances of various RWA-signaling solutions for dynamic lightpath provisioning, and aim to devise a better RWA-signaling solution with low request blocking probability, low signaling overhead, and short connection setup time.

Apart from lightpath provisioning, another key issue pertaining to the control and management of multicast-capable IP over WDM network that needs to be addressed is the distribution of IP traffic over the established lightpaths. Traffic grooming has been studied extensively for WDM/SONET ring networks [24-26] as well as mesh WDM networks based on static traffic demands [27]. Most recent studies on dynamic traffic grooming (DTG) [28-32] focus mainly on developing algorithms and strategies for IP traffic grooming. However signaling and its effect on grooming performance have not been considered. The common implicit assumption has been that the delay involved in signaling and resource information updating is negligible. This is not realistic as can be seen from the time it takes for

IP routers to exchange routing information and for route to converge. In addition, none of the studies considered multicast IP traffic in dynamic traffic grooming. All of these motivate us to find solutions for dynamic traffic grooming of both unicast and multicast traffic in the presence of additional constraints that exist in real-life networks.

Past studies on dynamic traffic grooming only focus on the problem of how to accommodate IP traffic on individual lightpaths established within the WDM optical layer. As a lightpath is a connection for one source and one destination, more than one lightpaths have to be employed to carry multicast traffic for a multicast session that involves one source and multiple destinations. Furthermore, inside an IP multicast-capable node where an IP packet can be replicated for different destinations, optical-electrical conversion will take place to convert the optical data stream into electrical IP packet flow. After the IP packets are replicated for different destinations, they will be converted back to optical data streams (electrical-optical conversion) and then be transmitted via various lightpaths to the intended destinations.

The availability of light splitters makes it possible for optical multicasting to remain in the optical layer. Due to the high cost of wavelength converters and light splitters [33-34], various strategies for implementing sparse and limited wavelength conversion and splitting have been proposed. Sparse conversion/splitting refers to the case where only a small subset of network nodes have the capability of wavelength conversion/light splitting. Limited wavelength conversion refers to the practical constraint that a wavelength converter can only convert an input wavelength to another wavelength within a finite wavelength band. Similarly, limited light splitting means that a light splitter has only a finite fan-out. If all the

network nodes are capable of unlimited wavelength conversion or infinite fan-out light splitting, the network is called a full wavelength conversion or full light-splitting network respectively.

When various combinations of the capabilities and constraints pertaining to light splitting and wavelength conversion are taken into account, the MC-RWA problem will be much more complicated. It is unclear how MC-RWA can be handled in an efficient manner when the aforementioned constraints pertaining to wavelength conversion and light splitting need to be considered. Motivated by this, we aim to find a systematic and computational efficient solution to this class of problems. Once the MC-RWA is solved, placement of wavelength converters and light splitters is tackled. New algorithms are required to seek placement configuration of wavelength converters and light splitters such that the network cost and request blocking probability are minimized or kept within a reasonable limit. Apart from considering the WDM network per se, there is also a need to study the IP over WDM network as a whole; this motivates us to examine a number of multicast issues in IP over WDM network holistically.

1.3 Objectives

Optical device and WDM technologies have become more and more sophisticated but the control and management of IP over WDM networks are far from maturity. Although various models and algorithms have been proposed and studied in the context of RWA and traffic grooming, few have considered the signaling effect and delay caused by processing of management and control information. A study on the signaling effects on network performance is necessary. Also, within a multicast capable network with wavelength converters and light

splitters, existing MC-RWA algorithms only consider the existence of the wavelength converters and light splitters, while the capabilities (or limitations) of wavelength converters and light splitters have been ignored. A generic model is needed to take the practical constraints concerning wavelength conversions and light splitting into consideration. With this generic model, algorithms for placing wavelength converters and light splitters with different capabilities/limitations can be studied systematically.

In light of the aforementioned consideration, the following objectives for this thesis are set:

- *To study the performance of control and management schemes for dynamic lightpath provisioning.* Network performances for SLE and DLE schemes will be studied. An enhanced scheme for lightpath establishment is needed to minimize the request blocking probability, connection setup time, as well as the control message overhead.
- *To study the signaling effect in multicast-capable IP over WDM networks.* Effective algorithms need to be devised to deal with the problem of dynamic traffic grooming by taking the signaling effects into consideration. Both unicast and multicast IP traffic will be considered.
- *To study the dynamic MC-RWA problem pertaining to wavelength-routed WDM networks.* Novel models and algorithms need to be devised to tackle the MC-RWA problem, and heterogeneity of network structures will be considered.
- *To study the placement problem in association with wavelength converters and light splitters by taking into consideration the capabilities/limitations of wavelength convertors and light splitters.* New algorithms need to be proposed to find out the optimal configuration in terms of where wavelength converters

and light splitters are to be placed and what types of capabilities they should have to yield good cost efficiency and network performance.

1.4 Contributions

This thesis makes the following contributions to the body of knowledge on the control and management of multicast-capable IP over WDM networks.

- Various control protocols for dynamic lightpath provisioning have been compared, which include the local information-based fixed and fixed alternate link routing schemes, the distributed global information-based link state routing scheme with backward, forward, parallel reservation signaling, and the centralized global information-based link state routing scheme. The performance metrics used include the request blocking probability (RBP), the connection setup time (CST) and the control message overhead (CMO). The study enables us to draw important conclusions on how the schemes will perform for small as well as large WDM networks.
- A new technique called DTG-PRL has been proposed to solve the DTG problem with the signaling effects being considered. To the best of the author's knowledge, this is the first time that the signaling effects have been studied for the dynamic traffic grooming problem in WDM optical networks. DTG-PRL first pre-reserves some lightpaths based on statistical traffic observations, then dynamically grooms the IP traffic based on the established virtual topology. ILP formulations and heuristic algorithms are proposed to find the pre-reserved lightpaths for both unicast and multicast IP traffic. The study enables us to conclude that DTG-PRL is able to outperform previously proposed DTG algorithms in IP bandwidth blocking probability (IBBP), network resource

utilization (NRU), connection setup time (CST) and control message efficiency (CME).

- The MC-RWA problem for wavelength-routed WDM networks has been studied. A new Generic Graph Model (GGM) has been proposed. The GGM is based on a layered auxiliary graph which is generic and able to support various node architectures and heterogeneous network structures. Based on GGM, the dynamic MC-RWA problem can be solved by invoking an efficient multicast tree algorithm for various light splitting and wavelength conversion scenarios. Our proposed approach provides a universal platform to study different aspects of dynamic MC-RWA and related problems.
- By extending our previous work on GGM, we study the placement of wavelength converters and light splitters as well as how their capabilities will affect the cost efficiency and performance in a holistic manner. In our study, both the sparse and limited capability strategies are considered, so is the coexistence of both wavelength converters and light splitters in the same network. To the best of the author's knowledge, this is the first attempt to integrate both wavelength converters and light splitters with both sparse and limit strategies in one model for the study of placement problem. The proposed GGM can be effectively used to solve this complicate placement problem. The dynamic multicast traffic model is used for the simulation study. Four proposed algorithms, namely, WCF-I, WCF-D, LSF-I and LSF-D, for finding the optimal configuration of wavelength converters and light splitters are studied. The configuration includes information on the availability as well as the capability (and limitation) of the devices on each node in the network. From the results, it can be concluded that the presence of wavelength converters contributes more to improving the cost

efficiency than the presence of light splitters when traffic load is not high. The results also enable us to conclude that, comparing with the limited strategy, the sparse strategy on placing the wavelength converters and light splitters have more impact on the network performance and cost efficiency. In addition, our results lead us to believe that the configuration with the highest cost efficiency for a specific traffic load is not necessarily the optimal configuration that will yield the highest cost efficiency for other traffic loads.

1.5 Organizations

The rest of the thesis is organized as follows:

Chapter 2 contains a literature review on control and management of multicast-capable IP over WDM networks. Issues on routing and wavelength assignment, lightpath establishment, traffic grooming, as well as the placement of wavelength converters and light splitters are covered in this chapter.

Chapter 3 contains a performance study of control and management protocols for dynamic lightpath provisioning in IP over WDM networks. The signaling mechanism and RWA protocol are described in detail and simulations are made to compare their performances.

Chapter 4 contains an investigation of the effects of signaling on DTG. The newly proposed DTG-PRL technique, ILP formulations and heuristic algorithms for finding the pre-reserved lightpaths for unicast or multicast IP traffic are all presented in detail.

Chapter 5 contains a study on the problem of dynamic MC-RWA in wavelength-routed WDM networks. The proposed GGM is elaborated in this chapter.

Chapter 6 contains a study on the placement of wavelength converter and light splitter based on our proposed GGM in Chapter 5. Four new algorithms, namely, WCF-I, WCF-D, LSF-I and LSF-D, for finding the optimal configuration of wavelength converters and light splitters are introduced here.

Chapter 7 contains a summary of the key achievements and conclusions. Also included are discussions and recommendation for future research.

Chapter 2

Literature Review

In this chapter, we provide an overview of the control and management framework of multicast-capable IP over WDM networks. The organization of this chapter is as follows. In Section 2.1, network models for IP over WDM networks are reviewed and control and management issues are categorized. One of the most fundamental issues – routing and wavelength assignment problem – is reviewed in Section 2.2. Other important control and management issues such as lightpath establishment, traffic grooming, and placement of wavelength converters and light splitters are reviewed in Sections 2.3 to 2.5, respectively. Section 2.6 contains a summary of this chapter.

2.1 Network models and issues in IP over WDM networks

There are three possible architectures for IP over WDM networks; namely, the overlay model, the peer model, and the augmented model, respectively [12, 35-36].

For the overlay model shown in Figure 2.1, topological, routing and other information about the underlying optical network is hidden from the client layers. The optical layer regards the IP layer as a client to which it provides point-to-point lightpath via a user-network interface (UNI). There are two separate control planes, one operates within the IP/MPLS layer and the other operates within the optical layer, and they operate independently from each other.

Figure 2.1 IP over WDM network: Overlay model

Figure 2.2 IP over WDM network: Peer model

The peer model allows network equipment in the optical layer and IP/MPLS layer to regard each other as peer, and allows detailed topological, routing and other information to be exchanged, as shown in Figure 2.2. There is a common control plane for both layers and one instance of routing protocol operating in the two domains. The peer model allows fault handling, service provisioning, and end-to-end protection to be provided in a holistic manner across the IP/MPLS layer and WDM layer. One major advantage of this model is that routing and resource allocation are optimized for both layers as a whole, by which better optimization will be achieved compared to optimizing two layers separately. However, when the transport network and the client services are operated by different operators, peer model is impractical because each operator will not allow its internal network structure and other information to be accessible by other operators, and its network equipment to be easily interfered by other operators' equipment. A performance comparison between the peer model and the overlay model can be found in [37].

The augmented model has a separate control plan and routing protocol for the optical and IP layer. Though each domain runs its own routing protocol, some routing information is exchanged across the UNI. The provisioning and restoration functions of the optical network are independent from the IP clients.

The architecture for WDM networks that is widely expected to form the basis for a future all-optical infrastructure is built on the concept of *Wavelength Routing*. The services that a wavelength routed network offers to its IP client networks are in the form of logical connections implemented using *lightpath*, also called *Virtual Link*. Lightpaths are clear optical paths which may traverse a number of fiber links in the optical network. Information transmitted on a lightpath does not undergo any OEO conversions, and thus OXCs can be very simple because they do not need to do any

signal processing.

In [17, 38], the concept of a lightpath was generalized into that of a *light-tree*, which, like a lightpath, is a clear channel originating at a given source node and implemented with a single wavelength. But unlike a lightpath, a light-tree has multiple destination nodes; hence it is a point-to-multipoint channel. The physical links implementing a light-tree forms a tree rooted at the source node. Light-trees may be implemented by employing power splitters at the OXCs. An attractive feature of light-trees is the inherent capability for performing multicasting in the optical domain, by which wavelength routed light-trees are useful for transporting high bandwidth, real-time applications such as high-definition TV (HDTV). Therefore, OXC equipped with power splitters will be referred to as multicast capable OXC (MC-OXC).

By reviewing the architecture of multicast-capable IP over WDM networks, the issues in IP over WDM networks can be divided into two categories: data plane issues and control plane issues. A classification of the major issues in these two categories is shown in Figure 2.3.

Data plane issues are highly dependent on the transmission technologies used at the optical transmission level as well as the data transfer mode. Optical transmission technologies include, for example, transparency and wavelength-band switching [39-40] that allow better flexibility and scalability in (spectral) resource utilization. Optical packet switching technologies [41] also have a major impact on the data plane. The control plane performs network level coordination functions such as general network state information management, path selection, and signaling, as well as the control of local operation of specific network elements [42-43]. It can also be viewed as a complex distributed software component bundle divided into routing (e.g.

OSPF-TE, IS-IS) and signaling (e.g. RSVP-TE, CR-LDP) planes.

Figure 2.3 Classification of IP over WDM network issues

From the time scales that control operations are targeting at, the issues in the control plane may be classified into network design phase and network operation phase issues [44]. During the network design phase, control mechanisms such as capacity planning, placement of OXC, placement of wavelength converters and light splitters, static traffic grooming, static RWA, as well as the lightpath/light-tree protection are investigated. During the network operation phase, control mechanisms such as admission control [45], signaling, dynamic traffic grooming, dynamic RWA, as well as lightpath/light-tree restoration are examined.

In practice, it is very hard to clearly define the research boundary of these control

plane issues because the issues are highly coupled. For example, to solve problems in capacity planning and placement of network resources, strategies on routing of lightpaths and wavelength assignments will have a huge impact on the final solution. Similarly, to solve problems related to dynamic traffic grooming, dynamic RWA schemes and signaling protocols will affect the network performance significantly. In this literature review, we will focus on such issues as RWA, lightpath establishment, traffic grooming, and placement of wavelength converters and light splitters, which are highlighted in Figure 2.3.

2.2 Routing and wavelength assignment

Routing and wavelength assignment is the fundamental control problem of optical WDM networks. The performance of a network depends not only on its physical resources (e.g., OXCs, converters, fibers links, number of wavelengths per fiber, etc.) but also on how it is controlled. The objective of an RWA algorithm is to achieve the best possible performance within the underlying physical constraints. The RWA problem can be classified to static RWA where the traffic requirements are known in advance and dynamic RWA in which a sequence of lightpath requests arrive in some random fashion.

The static RWA problem needs to be tackled during the design and capacity planning phase of network architecting, and the dynamic RWA problem needs to be handled during the real-time network operation phase and involves the dynamic provisioning of lightpaths.

The RWA problem arises because in optical WDM networks, routing and wavelength assignment are intertwined and best to be handled together [7]. RWA is known to be an *NP*-complete problem because of the following two constraints:

- *Wavelength continuity constraint*: a lightpath must use the same wavelength on all the links along its path from the source node to the destination node.
- *Distinct wavelength constraint*: all lightpaths using the same link (fiber) must be allocated distinct wavelengths.

To find a solution for the RWA problem, a simple but normally less optimal approach is to decouple the problem into two separate sub-problems: the routing problem and the wavelength assignment problem.

Methods for solving the routing sub-problem include fixed routing, fixed-alternate routing and adaptive routing, which are briefly described below.

- Fixed routing always uses the same predetermined route for a given pair of source-destination nodes. Generally the fixed shortest-path route will be used and only wavelength assignment is to be carried out upon connection requests. Fixed routing is very simple; however, it may lead to high blocking probabilities when resources along the path are scarce.
- Fixed-alternate routing selects the route from an ordered list of predetermined routes for a given source-destination node pair. For instance, these routes may include the shortest-path route, the second-shortest-path route, and the third-shortest-path route, etc. Fixed-alternate routing can significantly reduce the connection blocking probability compared with fixed routing [7].
- Adaptive routing selects the route by dynamically searching all feasible routes between a given source-destination node pair based on current network state. Certain types of routes may be preferred over others, for example shorter paths may be given a higher priority. Adaptive routing has the advantage of lower blocking probability than fixed and fixed-alternate routing but it requires more complicated online computations and controls [7].

The wavelength assignment sub-problem mainly focuses on how to select a wavelength for a given path when there are multiple wavelengths satisfying the wavelength continuity constraint. The most popular wavelength assignment methods include first-fit selection and random selection, which are briefly described below.

- First-fit selection chooses the lowest-indexed feasible wavelengths from an indexed wavelength list.
- Random selection chooses wavelength randomly from the set of feasible wavelengths, usually with a uniform probability.

Both of the two wavelength assignment methods can be combined with the aforementioned routing methods into the RWA algorithms for the dynamic lightpath establishment.

The wavelength continuity constraint may be relaxed if the OXCs are equipped with wavelength converters [33, 46-48]. A wavelength converter is a single input/output device that can convert the wavelength of an optical signal arriving at its input port to a different wavelength as the signal departs from its output port. With wavelength converters, a lightpath is able to use different wavelengths along different physical links.

Different levels of wavelength conversion capability are possible. *Full wavelength conversion capability* implies that any input wavelength may be converted to any other wavelength. *Limited wavelength conversion* [47] means that each input wavelength may be converted to any of a specific set of wavelengths, which is not the complete set of available wavelengths. The advantage of full wavelength conversion is that it removes the wavelength continuity constraint, making it possible to establish a lightpath as long as each link along the path from source to destination has a free wavelength, which could be different for different

links. As a result, the RWA problem reduces to the classical routing problem, i.e. finding a suitable path for each connection in the network. On the other hand, with limited conversion, the RWA problem becomes more complex than with no conversion because the number of alternatives increases exponentially. Wavelength conversion (full or limited) makes more efficient use of wavelengths and results in better performance.

Since converter devices increase network cost, a possible middle ground is to use sparse conversion, which means to employ converters in some, but not all, OXCs in the network. In this case, a lightpath must follow the wavelength continuity constraint on those OXCs without wavelength conversion capability, but not on those OXCs with wavelength conversion capability. It has been shown that implementing full conversion in a small fraction of OXCs in the network is sufficient to achieve almost all the benefits of full wavelength conversion [48-49].

With the availability of MC-OXCs and the existence of multicast traffic demands, the problem of establishing light-trees to satisfy these demands arises. We will call this problem the multicast routing and wavelength assignment (MC-RWA) problem [50]. MC-RWA bears many similarities to the RWA problem discussed above. Specifically, the tight coupling between routing and wavelength assignment remains, and even becomes stronger. In the absence of wavelength conversion the same wavelength must be used by the multicast connection not just along the links of a single path but along all the links of the light-tree. Since the construction of optimal trees for routing multicast connections is by itself a hard problem [51], the combined MC-RWA problem becomes even harder. Some work on solving the MC-RWA problem by considering the wavelength conversion capability can be found in references [52-54].

2.3 Lightpath establishment

Lightpath establishment can be divided into static lightpath establishment (SLE) and dynamic lightpath establishment (DLE) problems. If the traffic pattern is known in advance, or it will remain unchanged over time, we call the problem SLE. Generally the solution of the SLE problem is computed in an off-line manner during network planning. The SLE problem has been well studied and many solutions have been proposed [7, 55-57]. If the traffic pattern is constantly changing, lightpaths have to be established and released dynamically in response. This constitutes the key challenge of the DLE problem.

Dynamic lightpath provisioning [8, 35, 58-60] needs a set of signaling and routing control protocols. These protocols are implemented at the control unit of each OXC. Currently most of the standardization work on defining control plane protocols in the optical network takes place under the auspices of IETF. Three components of the control plane are highly related to the provisioning of lightpaths in optical networks [44, 61]:

- *Topology and resource discovery.* The main purpose of the discovery mechanism is to disseminate network state information including resource usage, network connectivity, link capacity availability, and special constraints.
- *Route Computation.* This component employs RWA algorithms and traffic engineering functions to select an appropriate route for a requested lightpath
- *Lightpath Management.* Lightpath management is concerned with setup and tear-down of lightpaths, as well as coordination of protection switching in case of failure.

During real-time network operation, requests for setting up lightpaths will be submitted from the edge nodes when needed. Each time a request is received,

depending on the state of the network at the time of a request, an RWA algorithm must be executed in real time to determine whether it is feasible to accommodate the request. If so, the procedure of resource reservation will be started by invoking the signaling protocol. If the request cannot be accepted because of lack of resources, it is blocked. Due to the real-time nature of the problem, RWA algorithms in a dynamic traffic environment must be very simple.

The resource reservation function is carried out using such signaling protocol as Constraint-Based Label Distribution Protocol (CR-LDP) or Resource Reservation Protocol with Traffic Engineering (RSVP-TE). Resource reservation may be implemented in parallel or in sequential manner [62]. In the parallel case, the source node attempts to reserve the desired wavelength on each link in the route by sending a separate control message to each node in the route. Each node that receives the reservation request will send a positive acknowledgement (ACK) or a negative acknowledgement (NACK) back to the source after trying to reserve the specified wavelength. In a sequential case, reservation request is sent hop-by-hop along the route selected by the source. When the control message reaches the destination, it is processed and acknowledgement will be sent back toward the source node.

In the hop-by-hop mode, the actual reservation of link resources may be performed while either the control message is traveling in the forward direction toward the destination, or the control message is traveling in the reverse direction back toward the source.

- *Forward Reservation:* Wavelength resources are reserved along the forward path to the destination on a hop-by-hop basis. Wavelength assignment depends on what kind of information is available to the source node. If the source node is maintaining complete state information, it will be aware of which wavelength is

available and can make a reservation based on wavelength availability. In the case of source node only maintaining local information about the resource, an aggressive reservation scheme, which will over-reserve resources along the path, can be implemented. The drawback of this forward reservation scheme is that network resources may be over-reserved during the reservation process, which may lead to unnecessary blocking of certain connection requests and result in lower network utilization.

- *Backward Reservation:* To prevent over reservation of resources altogether, reservations may be made after the control message has reached the destination and is headed back to the source. In backward reservation, the source node sends control packets to the destination without reserving any resources. These control packets will collect information about wavelength usage along the path, and let the destination make the decision of which wavelength to be reserved in this connection. The drawback of backward reservation is that if multiple connections are set up simultaneously, it is possible that a wavelength available on a link in the forward direction will be taken by another connection request and no longer be available when the reservation message traverses the link in the reverse direction.

Dynamic lightpath provisioning attracts a lot of research interests in the past few years. Good overviews of the architectural and algorithmic issues involved in deploying an optical cross-connect mesh network, and methodologies and algorithms for dynamic lightpath computation can be found in [35, 61].

Control and management approaches, namely link state and distributed routing, are proposed and compared in [8]. Simulation results presented there show that the link state approach has lower blocking probability under low load, and has an

advantage in traffic engineering since explicit routing can be implemented if each node maintains global link state information. On the other hand, under high load, distributed routing has lower blocking probability and shorter connection setup delays.

Two distributed signaling protocols, namely probe flooding-based routing (PFBR) and multi-path routing based on adaptive routing (MPR), have been proposed and compared in [58]. The study shows that FBR yields better call acceptance rate but does not do so well in terms of connection setup time, average routing distance, as well as control message overhead.

A signaling scheme that forwards the message before the switch is configured, known as the forward before cross-connects (FBXC) scheme, has been proposed in [63]. Simulation results show that the FBXC scheme outperforms the forward after cross-connects (FAXC) scheme in terms of restoration time.

For supporting link state routing, OSPF has been extended with traffic engineering, i.e. OSPF-TE, which promises to provide more robust and powerful routing capabilities. Performance trade-offs between routing accuracy and protocol traffic overhead in OSPF-TE, in terms of different TE-LSA update mechanisms, has been studied in [64].

The dynamic routing problem with inaccurate link state information, has been discussed in [65]. Based on the threshold-triggered update scheme, a probabilistic method to model the uncertainty of link state parameters has been proposed. By defining cost metrics based on the uncertainty, two routing algorithms are proposed to optimize the specific cost metrics, so as to reduce the impact of inaccurate link state information.

Blocking probability for distributed lightpath establishment in WDM optical

networks has been analytically studied in [22]. Three basic types of connection blocking due to insufficient network capacity, outdated information and over-reservation for two kinds of reservation methods, i.e. destination-initiated reservation (DIR) and source-initiated reservation (SIR), have been discussed. The proposed analytical models show highly accurate results for both DIR and SIR.

To overcome the problem of wavelength contention conflict under highly dynamic traffic loads, an intermediate-node initiated reservation (IIR) signaling scheme has been proposed in [66]. Results show that the IIR scheme significantly lower the blocking probability caused by outdated information, and outperforms the classic DIR scheme at the cost of a modest increase in control overhead. In another study, accurate analytical models are proposed, and even extended for networks with no wavelength conversion and networks with sparse wavelength conversion [67].

2.4 Traffic grooming

With the existing optical communication technology, the data rate of each wavelength is in the order of 2.5-10Gbps, and channels operating at 40Gbps and beyond will be commercially available in the near future. These rates are too high for single source-to-destination connections even for backbone networks. To efficiently utilize this capacity, a number of independent lower-rate traffic streams must be multiplexed into a single lightpath. To inject traffic into the lightpath at the source node and to extract it from the lightpath at the destination node requires the use of digital logic, and thus the nodes must be equipped with digital cross-connect (DXC). With DXCs, it is possible to extract traffic from a lightpath, and re-injecting it into another lightpath. Traffic forwarding at the electronic level is possible, and such

forwarding is referred to as OEO switching. Due to the rate mismatch between the traffic demands and the lightpath capacity, all traffic components cannot be carried on a single lightpath from source to destination, and some electronic routing of traffic with OEO conversion becomes inevitable. However, a backbone node with OEO functionalities introduces delay, loss of throughput, and increased probability of errors. Hence there are good reasons to keep electronic routing to the minimum. Moreover, with the deployment of commercial WDM systems, the cost of the DXCs used for OEO switching can be one of the dominant cost components in building optical networks. Therefore the concept of grooming traffic into available wavelengths emerges to mitigate network cost and meet other design goals. Many variations of the grooming problems have been studied in the literature, for example, in References [13, 24-29, 31-32, 55, 68-76].

Early work on traffic grooming concentrated on static traffic grooming where all traffic demands are known to be reasonably stable over time, and embodied in a traffic matrix. In this case, the traffic grooming problem is closely related to the issues of virtual topology design and reconfiguration, which is well known to be *NP*-complete. In general, the virtual topology design problem can be formulated as an optimization problem, using principles from multi-commodity flow for physical routing of lightpaths and traffic flow on the virtual topology. Technique of integer linear programming (ILP) and mixed integer linear programming (MILP) are typically used for the formulation of the problem of virtual topology design. The objectives may aim to maximize the throughput [27], minimize the average packet hop distance [55], minimize congestion [56], or minimize the reconfiguration from an old virtual topology to a new virtual topology [55, 77]. Due to the computation complexity in solving ILP problems, heuristic algorithms are normally proposed to

solve the DTG problem for large networks.

As the explosive growth of the Internet traffic, applications such as peer to peer, video on demand, and video conferencing become popular. The traffic will become more bursty; therefore a more flexible traffic grooming mechanism is needed. For dynamic traffic, the algorithm for traffic grooming must be seen as a mechanism the network control plane can employ to make decisions in response to traffic change events [78]. Dynamic traffic grooming in IP over WDM network can be seen as a two-layer routing problem that needs to take into consideration the constraints in both IP and WDM layers, as well as the relationship between them. The traffic grooming policy can be divided into two categories depending on the IP over WDM network architecture: the overlay model and the peer model. In the peer model, the link usage information in both the IP and WDM layers can be used to perform route calculation. In the overlay model, the optical layer is invisible from the IP layer, and lightpath connection request has to be initiated by the IP layer via the UNI for new bandwidth demand. A novel graph model has been introduced for the peer model in [30], which considers a heterogeneous WDM mesh network environment and combines both IP and WDM layers into a single auxiliary graph. Two policies, the IP-Layer-First (ILF) and the Optical-Layer-First (OLF), have been studied for the overlay model in [45], and a so called One-Hop-First (OHF) policy has been studied in [58]. Some analytical models, which use the Markov process to study the performance of different traffic grooming strategies can be found in [74].

2.5 Placement of wavelength converters and light splitters

Wavelength converters and light splitters are two critical components in

realizing light-tree based optical layer multicasting. Wavelength converters can eliminate the wavelength continuity constraint and thus improve wavelength utilization. Light splitters can make optical multicasting more cost-efficient by minimizing the use of wavelengths. With these two types of devices, bandwidth request blocking probability can be reduced and network performance can be improved with well-designed network and node architecture, and RWA algorithms [46, 79-80].

However, for a light splitter with a fan-out of n (≥ 2), the power at the output of each splitter is only $1/n$ of that of the input power. In order to maintain a reasonably high output power level of the signal after splitting, a costly active amplification device is required to be installed in the splitter, which results in high total system cost [34]. In addition, wavelength converter is very expensive and is anticipated to remain expensive for many years to come [33]. Due to the aforementioned two factors, various strategies for sparse and limited wavelength conversion and splitting have been proposed. It has been shown in a number of studies that a relatively small number of wavelength converters and light splitters can achieve satisfactory performance [48, 81-82]. Nevertheless, the problem of how to properly place the wavelength converters and light splitters in a network is still very important.

Placement of wavelength converters and light splitters are more related to the issues of network design. Given a network topology, a certain number of wavelength converters and/or light splitters, and the traffic statistics between the nodes, the problem of interest is the optimal placement of converters in the network. Approaches to optimal placement depend on the optimality criterion and the traffic model [82]. Note that the placement problem contains the RWA problem as a sub-problem. The placement problem can be formulated as an ILP problem similar

to the RWA formulation. A number of studies on wavelength converter and light splitter placements can be found, but most of work consider the placements of wavelength converter and light splitter separately.

In [82], S. Subramaniam *et al.* propose a dynamic programming algorithm for placing a given number of converters in the network such that the network call blocking probability is minimized. B.Li *et al.* [83] propose two heuristic algorithms for wavelength converter placements for fix-alternate routing and least-loaded routing respectively, and the heuristic algorithms are shown to be optimal for the respective routing algorithms. X. Chu and Bo Li *et al.* [79] show that the wavelength converter placement and RWA algorithms are closely related in the sense that a well-designed wavelength converter placement mechanism for a particular RWA algorithm might not work well with a different RWA algorithm[79, 83]. S. Gao *et al.* [84] propose an optimization model for the converter placement problem as the minimization of a polynomial function of 0-1 variables under a linear constraint and describe a search algorithm that can efficiently find the optimal solution. M. E Houmaidi *et al.* [85] present a new analytical model that can capture link dependencies in all-optical WDM networks under uniform traffic.

N. K. Singhal *et al.* [86] investigate two switch architectures, an opaque (O-E-O) and a transparent (all-optical) one, to support multicasting in a WDM network. They also describe the mathematical formulations for RWAs that aim to optimize the global cost for several light-tree based multicast sessions for both switch architectures. In addition, the authors propose mixed integer linear program (MILP) formulations and heuristic algorithms for establishing a set of multicast sessions in a network with and without wavelength converters and with traffic grooming considerations. H. C. Lin *et al.* [87] study the splitter placement problem, of which

the goal is to place a given number of splitters in the network such that the average wavelength usage per link for a given set of multicast connections is minimized. The proposed k -maximum wavelength reduction method shows near-optimal average wavelength resource usage.

Some works have been done where the placement problems of wavelength converter and light splitter are considered jointly [19, 88-89]. R. Malli *et al.* [19] consider sparse light splitting and sparse wavelength conversion jointly, but the placement of the node with wavelength conversion or light splitting capability is more or less determined randomly. Y. Cao *et al.* [88] propose an ILP model to optimally solve the combined problem of placement of wavelength converters and light splitters. However the heuristic algorithm is not able to deal with large networks. Moreover, the objective of the proposed ILP model is to minimize the total number of wavelength channels required by static multicast traffic while the minimization of traffic blocking probability could be an important objective under dynamic multicast traffic. K. Yong *et al.* [89] proposes an analytical model for computing the approximate blocking probability in multicast optical virtual private network (M-OVPN), and a *Removal* algorithm has been proposed to solve the combined placement problem of wavelength converters and light splitters.

2.6 Summary

In this chapter, we provide a review on the issues related to control and management in multicast-capable IP over WDM networks. First, three network architectures for IP over WDM networks are reviewed. We categorize various control and management issues of IP over WDM network into data plane issues and control plane issues, and then concentrate on reviewing the control plane issues.

The control plane issues can be categorized into network design issues and network operation issues. We first review the problem of RWA, which is fundamental to both network design and network operation. Then we review the problem of lightpath establishment, which is also an important problem to be resolved for the control plane. After that, the traffic grooming problem is reviewed. Finally, we review the placement problem of wavelength converters and light splitters.

Chapter 3

Performance Study on Control Protocols for Dynamic Lightpath Provisioning

In this chapter, we study the performance of various control schemes for dynamic lightpath provisioning. Next generation Internet is moving towards IP/GMPLS/WDM networks with IP-based GMPLS control plane controlling intelligent OADM/OXC and optical devices within the underlying WDM networks. This network is required to provide fast responses to customers' demands for automatic setup and tear-down of lightpaths across the networks. Thus, appropriate control scheme is needed to provide dynamic lightpath provisioning in order to carry out RWA to determine the path and invoke the signaling function to reserve resources.

In a GMPLS network, the resource discovery function is typically carried out using the Open Shortest Path First (OSPF) routing protocol [90]. How path determination is carried out depends on whether a global-information based or a local-information based approach is adopted. For a global routing approach, a route for a connection request can be found based on the distributed global routing table maintained by each node. For a local routing approach with alternate routes, a hop-by-hop based routing scheme is used [91]. The resource reservation function is carried out using such signaling protocol as the Constraint-Based Label Distribution Protocol (CR-LDP) or the Resource Reservation Protocol with Traffic Engineering (RSVP-TE). Source reservation may be implemented in parallel [21] or in a

sequential manner [9, 62].

The global information-based link state approach has good performance in terms of request blocking probability (RBP), however excessive control messages need to be exchanged and this incurs heavy control overhead. The local information-based link state approach does not require many control messages to be exchanged. However, its RBP could be unacceptable when traffic load is high. Updated global link state information is the key requirement to achieve low RBP but high control message overhead (CMO) will be incurred to keep the info updated. There is a trade-off between the link state update frequency and the RBP. In addition, connection setup time (CST) is also closely related to the routing mechanism. CST will be short if the selected route has a low RBP and low hop count. Whether source reservation is executed in parallel or sequentially will have an impact on the connection setup time as well.

An alternative approach to the above is the centralized global information-based link state approach. In this approach, each network node sends its link state information to the central manager which collates all the information received and places it in a global link state information base. Lightpath establishment is carried out by the connection/lightpath manager which will execute the integrated RWA algorithm and the signaling algorithm upon the request of the source node. No resource broadcasting is required as all update messages to synchronize the global link state will be sent only to the manager. This helps to reduce the massive flow of update messages as required in the distributed approach. It could be a more effective control and management protocol for dynamic lightpath provisioning, as the virtues of the global information-based link state approach and the local information-based approach are combined together.

In this chapter, we carry out a performance study of various commonly adopted routing and signaling protocols for dynamic lightpath provisioning. These include the local information-based fixed and fixed alternate routing approaches; the distributed global information-based link state approaches with backward, forward, parallel reservation signaling; and the centralized global information-based link state routing protocol. The performance metrics used include RBP, CST and CMO.

The rest of this chapter is organized as follows. Section 3.1 introduces the network model. The local information-based fixed and fixed alternate link routing approaches, the distributed global information-based link state approaches with three different types of signaling, as well as the centralized global information-based link state approach are described in Sections 3.2, 3.3 and 3.4 respectively. The simulation model, results, and analysis are reported in Section 3.5. Finally in Section 3.6, we summarize the work reported in this chapter.

3.1 Network model

The network model is shown in Figure 3.1. It consists of a collection of Optical Cross-Connects (OXC) / Optical Add-Drop Multiplexers (OADM), each of which is controlled by a GMPLS control plane. The control plane assumes the same topology as the data plane. The control plane houses the routing and signaling algorithms as well as the management agent. Collectively, they control and manage the optical devices, configurations, routing, signaling, traffic engineering, etc. Inside each control plane, local or global information will be kept for routing purposes.

The following assumptions are made in this study:

- Switching between OXC/OADM is purely optical without OEO conversion.

We assume the necessity for wavelength continuity and no wavelength

conversion is available in the optical devices.

- Each lightpath forms an optical label switch path (OLSP) between IP routers.
- Routing information update is assumed to be completely reliable. No consideration is made on the possible error conditions.

Under the above assumptions, we consider three different approaches in the following sections.

Figure 3.1 GMPLS optical network model

3.2 Local information-based fixed and fixed-alternate link routing approach

In these protocols, each node keeps a static routing table, which consists of an ordered list of several fixed shortest routes to each destination node. It also maintains connection records and the status of each wavelength on each outgoing link of this node. When a connection request arrives, the ingress node establishes

the connection using each of the routes in sequence. The signaling steps are as follows:

- Upon receiving a request from the IP router, the ingress node sends a *RESERVE* message, with a set of available wavelengths to the egress node. The set of available wavelengths are not reserved during this probe stage, and the wavelengths can still be used for the routing of other connection requests.
- The intermediate node matches the set of wavelengths with its own set of wavelengths. If the result is null, an *NACK* message is sent back to the ingress node. Otherwise, it forwards the *RESERVE* message with this result. The wavelengths inside the set are still available for the matching of other connection requests. The process is repeated until the *RESERVE* message reaches the egress node.
- The egress node fixes a wavelength and sends an *ACK* message back to the preceding node along the route. The selected wavelength will be reserved by each intermediate node along the route. If the reservation is successful, the switch will be turned accordingly and this wavelength will not be used for routing other connection requests until it has been released when this connection is torn down. The process is repeated until the *ACK* message is returned to the ingress node.
- If the reservation is unsuccessful, *NACKs* will be sent to both the ingress node and those nodes that have reserved the selected wavelength so as to release the reservations.
- When an *NACK* is received in the ingress node, for the Fixed Link Routing approach, the ingress node will block the connection; for the Fixed-Alternate Link Routing approach, the ingress node will pick the second shortest route and

repeats the signaling process. If the second attempt fails, the connection is blocked.

The timing sequence is illustrated in Figure 3.2.

In these approaches, the eligible wavelengths in the set are still available for use by other connection requests. The wavelength may not be available during the reservation trip, if it has already been taken up by another connection request. In this case, the connection will be blocked. This blocking is the result of contention conflict.

Figure 3.2 Local information-based fixed and fixed-alternate link routing approaches

3.3 Distributed global information-based link state approach

In these protocols, each node maintains a global routing table that includes the network topology information and the wavelength usage on each link. The ingress node determines an optimal route to the egress node on one or more free wavelength according to a RWA scheme. Synchronized and identical link state information must be maintained at all nodes. Whenever a state change occurs, all the nodes must be informed. The updating can be done either periodically or triggered when the threshold is hit. To ease the comparison, triggering by threshold is assumed in this study. Signaling can be carried out either in parallel or in a sequential manner. We consider three different signaling mechanisms, namely parallel signaling mechanism, backward-reservation sequential signaling mechanism and forward-reservation sequential signaling mechanism.

3.3.1 Parallel signaling mechanism

Figure 3.3 Timing sequence of parallel signaling approach

Once a route and a fixed wavelength are selected, the ingress node will send separate reservation requests in parallel to each node in the route. Each request is routed on the shortest path. The signaling procedure is as follows:

- The ingress node selects a route and a wavelength according to a certain RWA scheme.
- It sends separate reservation requests in parallel to each node in the route. Each request is sent along the shortest path.
- If an intermediate node is able to reserve the wavelength, an *ACK* is sent back to the ingress node.
- If all the reservations are successful, the ingress node sends a *SETUP* message to each of the nodes. Each node then configures its switch accordingly and the connection is established.
- If one of the reservations is not successful, the call is blocked and the ingress node sends a *TAKEDOWN* message to each node except the one which sent the *NACK* to release the reserved resources.

The timing sequence is illustrated in Figure 3.3.

3.3.2 Backward-reservation sequential signaling mechanism

Once a route is selected according to the link state information, the source node sends the reservation message with a few probable wavelengths along the route. The reservation request will be forwarded sequentially between two neighboring nodes along the route. The egress node makes the decision to select a wavelength and then sends the *ACK* with the wavelength information back along the route. Each preceding node configures its switch when it receives the *ACK*. The signaling procedure is given below:

Figure 3.4 Timing sequence of backward-reservation sequential signaling approach

- Upon receiving a request from the IP router, the ingress node sends a *RESERVE* message, with a set of wavelengths available on the path selected by a certain RWA algorithm, to the egress node.
- The intermediate node matches the set of wavelengths with its own set of wavelengths. If the result is null, an *NACK* message is sent back to the ingress node. Otherwise, it forwards the *RESERVE* message with this result. The wavelengths inside the set are still available for the matching of other connection requests. The process is repeated until the *RESERVE* message reaches the egress node.
- The egress node selects a wavelength from the resulting set and configures its switch. It then sends an *ACK* message back to the preceding node. If the intermediate node successfully reserves the wavelength and configures its

switch, the *ACK* will be passed to its preceding node. The process is repeated until the *ACK* message is returned to the ingress node and the connection is established.

- If the intermediate node failed to reserve the wavelength, the *NACK* will be sent to the ingress node, as well as those nodes along the route that have reserved the wavelength so as to release the reservation.
- When an *NACK* is received, the ingress node will block the connection.

The timing sequence is illustrated in Figure 3.4.

In this approach, the contention request may be blocked due to contention conflict, as described in the local information-based fixed alternate link routing approach. In addition, in this approach, we let the node send the *ACK* to its preceding node after it sets the switch successfully. Thus, the time of setting the switches does not overlap. This will probably result in a long connection setup time.

3.3.3 Forward-reservation sequential signaling mechanism

Figure 3.5 Timing sequence of forward-reservation sequential signaling approach

In this mechanism, the source node selects a route and a wavelength and sends a reservation request. The request is forwarded sequentially between two neighboring nodes along the route. Each intermediate node reserves the wavelength, and forwards the reservation message before begins to set the switch for the connection. In this way, the setting of the switches will overlap in time, thus, reducing the total signaling time. This signaling procedure is similar to the FBXC scheme proposed in [63], which is expected to have better blocking performance and shorter connection setup time than the FAXC. The signaling procedure is given below:

- The source node selects a route and a wavelength according to a certain RWA scheme.
- The source node sends the *RESERVE* request to the next node along the route.
- If an intermediate node is able to reserve the wavelength on the link, it starts to set the switch for this connection and forwards the reservation request to its next hop along the route, an *ACK* message will be sent back to the source node as soon as the switch is ready.
- If all the reservations are successful, the source will receive all the *ACK* messages from the nodes along the route. The connection is successfully established.
- If any of the reservations fails, the call is blocked and all the agreed nodes will receive an *NACK* message from the failed node to release the reserved resources.

The timing sequence is illustrated in Figure 3.5.

In this protocol, if the global link state information of the ingress node is not up-to-date, it may choose an unavailable route, which will result in blocking. At the

same time, the node sends the *RESERVE* message to its next node before it begins to set the switch. In this way, the switch setting time can overlap, resulting in a lower connection setup time. However, if the reservation is aborted at some downstream node, those nodes that have begun to set the switch will have to stop and restore the switch. This will incur additional delay in releasing the connection.

3.4 Centralized global information-based link state approach

In this protocol, a central manager is invoked to maintain the global routing table and connection table. The first table provides the global routing information, network topology connectivity, and the link state information of all the nodes in the network. The second table gives the information of the connections established in this network. Each network node maintains only a local routing table. It sends the updates to the manager either periodically or when a threshold is reached. In this work, the threshold method is used and its value is set to 1. It means that an update is triggered whenever a change is made to the routing table. Figure 3.6 shows the network model with the central manager.

The signaling procedure is given as follows:

- Upon receiving a connection request from an ingress node, the manager checks the global routing table to determine the optimal route for this connection and at the same time selects a free wavelength for the route.
- The manager sends the *RESERVE* messages in parallel to each node along the route to reserve the wavelength.
- If an intermediate node is able to reserve the wavelength, it starts to configure its switch for this connection, updates the local connection table and local

routing table, and replies with an *ACK* message to the manager as soon as the switch is ready.

- After receiving all the *ACK* messages from the nodes along the route, the manager will update the connection table and global routing table, and inform the source IP router that the connection has been successfully established.
- If one of the intermediate nodes fails to make the reservation, an *NACK* message will be returned to the manager. The manager will send a *TAKEDOWN* message to each node except the one which sent the *NACK* to cancel the reservation. The connection is blocked.

The timing sequence is illustrated in Figure 3.7.

Figure 3.6 The network model with a central manager

Figure 3.7 Time sequence of the centralized global information-based link state approach

3.5 Simulation and performance analysis

3.5.1 Topology and assumptions

We adopt the New Jersey Local Access Transport Area network (NJ LATA) shown in Figure A.3 [92] for the simulation. This topology consists of 11 nodes interconnected by 2 unidirectional fibers. We assume each fiber carries 64 wavelengths.

For comparison, we adopt two traffic models in the simulations: 1) Exponential Traffic Distribution (ETD) and 2) Pareto Traffic Distribution (PTD). In ETD, call requests arrive at each node according to a Poisson distribution process with a network arrival rate λ , and the session holding time is assumed to be Exponential distribution with mean service time $1/\mu$. In PTD, the call request arrival rate and the session holding time both follow the Pareto distribution with $Shape = 1.2$ [93]. The mean session holding time in both PTD and ETD are assumed to be one second. The load (λ/μ) at each node is measured in Erlang. The average load per node is used

for plotting the performance diagrams. The processing time for a control message at each node is $10\mu s$, where $1\mu s = 10^{-6}s$. The processing time for a control message at the source node or the manager is $30\mu s$ because of the complexity of the connection request. The time to set up a switch connection is $500\mu s$. The propagation delay is $5\mu s/km$. In the simulation, each data point is obtained over a period of 100,000 connection requests.

We consider the following performance metrics:

- *Request Blocking Probability* (RBP): Probability that a connection cannot be established due to resource unavailability or contention along the desired route.
- *Connection Setup Time* (CST): Time to establish a connection once a connection request arrives.
- *Control Message Overhead* (CMO): The total traffic arising from all the control messages, including the routing messages for link state updating. We assume that the average size of one control message for all approaches is 30 bytes.

For convenience, we label the above protocols as *Fix*, *FixAlt*, *Backward*, *Forward*, *Parallel* and *Central*, respectively. *Fix* and *FixAlt* denote the local information-based fixed and fixed-alternate link routing approaches respectively. *Backward*, *Forward* and *Parallel* all represent the distributed global information-based link state approach in which *Backward* denotes sequential signaling with backward reservation; *Forward* denotes sequential signaling with forward-reservation; and *Parallel* denotes parallel signaling. *Central* denotes the centralized global information-based link state approach.

3.5.2 Numerical results

Figure 3.8, Figure 3.10 and Figure 3.12 show the results of ETD, and Figure

3.9, Figure 3.11 and Figure 3.13 show the results of PTD. Both of these two traffic models have very similar performances in terms of RBP, CST and CMO. Figure 3.8 and Figure 3.9 show the blocking probability versus traffic load for the above six protocols. As shown in the figures, *Fix* shows the worst performance and *Central* shows the best performance. In *Fix*, there is no link state updating and thus it has the highest blocking probability. In *Backward*, *Forward* and *Parallel*, since they all have link state updating, and may choose a route that is not likely to be blocked, their performances are better than *Fix*. *Forward* and *Parallel* have nearly the same blocking performance because they both have the same routing and wavelength assignment mechanism. The difference in signaling protocol between *Forward* and *Parallel* does not impact the RBP. *Backward* has a higher RBP compared with *Forward* and *Parallel*, although they have the same routing mechanism. The reason is due to the different signaling protocols. *Forward* and *Parallel* forward the messages before turning the switches, while *Backward* forwards the messages after turning the switches, which may cause outdated link state information and subsequently introduce more contention conflicts. However, *FixAlt* has a much better blocking performance compared with *Fix*, *Forward*, *Backward* as well as *Parallel*, when traffic load is lower than 3 Erlang for both traffic models. This is attributed to *FixAlt* has a second chance to try on the alternative route. As the traffic load increases to higher than 3 Erlang, even alternative routes and two chances of trying cannot mitigate blocking without dynamical link state updates. As a result, the RBP of *FixAlt* increases sharply and approaches that of *Fix*.

Figure 3.8 RBP vs. TL (ETD)

Figure 3.9 RBP vs. TL (PTD)

Nevertheless, by using centralized global routing and link state table, *Central* yields the best RBP performance among all the six protocols being compared, and even lower than *FixAlt* which has two chances in lightpath provisioning. Since the link state updates are only sent to the central manager which maintains the global routing table and link state table, contention conflicts can be reduced significantly during the RWA and signaling process, and this results in the lowest RBP.

Figure 3.10 and Figure 3.11 show the connection setup time versus traffic load. *Forward* yields the lowest connection setup time and *Parallel* yields the highest. *Backward*, *Fix*, *FixAlt* and *Central* yield moderate connection setup time with *Central* being the lowest among the four. When the RBP is not high, e.g. less than 1%, the shortest paths are often chosen for each approach. The difference in connection setup time can be attributed to the different signaling mechanisms used. *Parallel* requires 2 round-trip times to make a connection, which results in the highest setup time. *Forward* requires only 1 round-trip and the time durations taken to turn on different switches in the nodes are overlapping, which results in the lowest connection setup time incurred. *Central* incurs one more round-trip between the ingress node and the manager than *Forward*, thus, its connection setup time is slightly longer than *Forward*. It is nonetheless much lower than that of *Backward*, *Fix* and *FixAlt*, which have nearly identical average CST because of the similar signaling mechanisms used. As traffic load increases towards 4 Erlangs, the shortest paths increasingly become unavailable and longer paths have to be selected; hence, the CST will increase slightly. When traffic load is higher than 4 Erlang, *Fix* and *FixAlt* have a relatively high RBP (higher than 1%), which suggests that longer paths may be blocked with high probability, and only shorter paths can be set up successfully; hence the connection setup time decreases as the traffic increases.

Figure 3.10 CST vs. TL (ETD)

Figure 3.11 CST vs. TL (PTD)

Figure 3.12 and Figure 3.13 show the control message overhead versus traffic load. *Fix* and *FixAlt* have the lowest overhead because they do not generate link state update messages. The overhead traffic of *FixAlt* is slightly higher than that of *Fix* when traffic load is relatively high; this could be due to the retry overhead traffic of *FixAlt*. In the simulation, the update messages are generated as soon as changes occur, and all the update messages are broadcasted to all the nodes of the network, hence the routing overhead is $O(N \cdot (N-1))$. As *Parallel* uses parallel signaling, which generates individual reserve messages to each node in the route, its control message overhead is higher than that of the other two sequential cases, i.e. *Backward* and *Forward*. *Backward* and *Forward* have nearly the same CMO because they use the same updating strategy and have similar message overhead for signaling. *Central* also has relatively low overhead. The main reason is that the update messages are only confined between the manager and individual nodes, and no broadcasting is necessary. Hence, the routing overhead is only $O(N)$. This accounts for the low control message overhead in *Central*. In general, as the traffic load increases, the shortest paths become unavailable and longer paths have to be selected. The control message overhead increases slightly, except for *Fix* and *FixAlt*. When the RBP of *Fix* and *FixAlt* becomes very high, connections are more likely to be blocked and shorter paths have better chance to be accepted, which results in a slight decrease in the CMO.

Figure 3.12 CMO vs. TL (ETD)

Figure 3.13 CMO vs. TL (PTD)

In conclusion, the centralized global information-based link state approach yields the least RBP and low CMO for the set of assumptions made in the simulation. In a small-scale network, such as NJ LATA, the CST is comparable with the distributed global information-based link state approach with forward-reservation signaling, and much lower than the other distributed global information-based link state approaches.

3.5.3 More results on processing delay

In the previous study, we conclude that the centralized global information-based link state approach yields the least RBP, low CMO and comparable CST. However in reality, the centralized approach has such disadvantages as having single-point of failure and forming processing bottleneck. If network information, such as link state routing table, is highly centralized on one server, it will jeopardize the reliability of the whole network. In addition, if all the connection requests have to be processed by a single server, it will easily overload the server, and congest the signaling system. All of these factors need to be carefully considered in the design of real-life networks.

As such, one interesting question is under what conditions we still be able to take advantage of the centralized approach. Another follow-up question is, when the centralized approach is inappropriate, what kind of lightpath provisioning approach will be more appropriate to achieve good network performance.

We will now focus on the processing capability of the central manager, which can be easily observed by adjusting the event processing delay in the simulation study. We will extend our simulation study to make comparison among the few approaches being studied for different event processing delays.

In the new simulation study, we only use the ETD traffic model for comparison,

as the PTD model will yield similar results as the ETD model. All the other assumptions remain unchanged. We compare the RBP of all the above six lightpath provisioning approaches for average event processing delays in the order of 10^{-4} , 10^{-5} , 10^{-6} and 10^{-7} second. To be precise, the processing delays for general events and connection requests are assumed to be $[100\mu s, 10\mu s, 1\mu s, 0.1\mu s]$ and $[300\mu s, 30\mu s, 3\mu s, 0.3\mu s]$, respectively.

For clarity, we separate the results in three figures namely, Figure 3.14 to Figure 3.16. In the figures, D_p denotes the scale of the event processing delay.

Figure 3.14 RBP vs. TL with different processing delay – *Fix* and *FixAlt*

Figure 3.15 RBP vs. TL with different processing delay – *Backward* and *Forward*

Figure 3.16 RBP vs. TL with different processing delay – *Parallel* and *Central*

Figure 3.14 shows the comparison between *Fix* and *FixAlt*; Figure 3.15 shows the comparison between *Backward* and *Forward*; and Figure 3.16 shows the comparison between *Parallel* and *Central*.

In Figure 3.14, it can be found that for different time scales of event processing delay, the RBPs of *Fix* and *FixAlt* are quite close. The reason is that they have least control message overhead without link state updates. On the other hand, it can also be concluded that having shorter event processing delay will not help to reduce the RBP of *Fix* and *FixAlt*.

When the processing delay is in the order of 10^{-4} second, Figure 3.15 and Figure 3.16 show that the RBPs of *Backward*, *Forward*, *Parallel* and *Central* are significantly higher than those in other time scales of processing delay. This is especially true when traffic load is higher than 3 Erlang, where the RBPs of *Backward*, *Forward* and *Parallel* increase rapidly when the traffic increases. This could be due to the exponential increase in the number of link state update messages, which can not be processed in time and causes blocking as a result of outdated link state information. Similarly, for *Central*, because of its relatively low control message overhead, its RBP can still be maintained at an acceptable level until the traffic load approaches 5 Erlang. After that, the exponential increase in link state update messages has overwhelmed the central manager. Due to outdated link state information, the RBP increases steeply and exceeds those of *Backward*, *Forward* and *Parallel* around 7 Erlang, and finally approaches unity when traffic load is 10 Erlang.

When the processing delay is in the order of 10^{-5} , 10^{-6} , or 10^{-7} seconds, the RBPs of *Forward* and *Parallel* show obviously improvement as the processing delay becomes lower. The RBPs of *Central* in the time scales of 10^{-6} and 10^{-7}

approach 0 when the traffic load is less than 8 Erlang. When the traffic load is higher than 8 Erlang, RBP increases mainly due to lack of wavelength resource. However the RBP of *Backward* does not improve too much as processing delay becomes lower. The reason could be that the signaling procedure in *Backward* is sequential without time overlap for turning the switches. Therefore, the time saved from the processing delay has less impact compared with the delay incurred in turning all the switches sequentially, i.e. $500\mu s$ each time.

It should be noted that the actual processing delay is decided by the hardware or CPUs inside the control units of the routers and optical switches within the network. In addition, the performances of the few schemes being considered will be governed by both the processing time and service holding time used in the simulation. In our simulation, the average service holding time is set to 1 second. Though our simulation study may not be able to show the actual performance, it does show the relative performance between the schemes being compared.

In conclusion, when the processing delay is not large, e.g. in the order of 10^{-5} second or less in our simulation, *Central* is always the optimum approach for lightpath provisioning in term of RBP. However if the processing delay is large, as well as when the traffic load is high, *FixAlt* is more appropriate for lightpath provisioning because of its simplicity and low control message overhead. Among the other approaches, *Forward* is the second choice for lightpath provisioning in term of RBP. In addition, as a distributed global information-based link state approach, *Forward* is more robust in term of reliability compared with the centralized global information-based link state approach.

3.6 Summary

In this chapter, we present the performance comparison of various control protocols for dynamic lightpath provisioning, namely, the local information-based fixed and fixed-alternate link state routing approach, the distributed global information-based link state approach with backward, forward, parallel reservation signaling, and the centralized global information-based link state approach. We use RBP, CST and CMO as the performance metrics for the comparison. The centralized global information-based link state approach yields the lowest RBP and very low CMO when the processing delay is in the order of 10^{-5} second or lower. In a small-scale network, such as NJ LATA, its CST is comparable with the distributed global information-based link state approach with forward-reservation signaling, and much lower than the other distributed global information-based link state approaches.

When the network is large or when the average event processing delay is long, the centralized approach will not be scalable. Nevertheless, the distributed global information-based link state approach is also not scalable for large networks as the routing overhead will become excessive and paralyze the network. As such, local information-based fixed-alternate link routing approach is more appropriate for lightpath provisioning because of its simplicity and low control message overhead.

Chapter 4

Signaling Effect on Dynamic Traffic

Grooming in IP over WDM networks

As the WDM technology becomes more mature and widely used, the gap between the bandwidth requirement of a typical connection request (e.g., OC-1, OC-3, or OC-12) and the bandwidth of a single wavelength (capacity of a WDM channel, e.g., up to OC-192 or OC-768) becomes wider. In order to efficiently utilize the huge wavelength bandwidth resource, mechanisms for traffic grooming in IP over WDM network have been studied. By integrating low-speed traffic streams into high-capacity lightpaths, traffic grooming will enhance the wavelength bandwidth utilization, improve the network throughput and minimize the network cost [13, 72].

Most recent approaches on dynamic traffic grooming (DTG) [28-32] focused mainly on developing algorithms and strategies for IP traffic grooming. However traffic signaling and its effect on grooming performance have not been considered. In reality, the time consumed by signaling and information update at each node may have a strong impact on the network performance. In a practical network, signaling is needed for lightpath establishment and graceful disconnection. Typical signaling message processing time is about 0.1 ms ($1\text{ ms} = 10^{-3}\text{ s}$) for a simple message and $0.3 \sim 0.4\text{ ms}$ for a complex message, e.g., a connection request which includes route computation time. The timing figures quoted above were obtained from measurements in the AT&T prototype testbed [94-95]. Consequently, the IP router

in each node of the testbed [95] cannot process over 10000 events in one second. The excess events that are beyond the node processing capacity will be delayed. As time goes by, the delayed events accumulate and grow exponentially. They will either be discarded or waited in the queue, depending on the queuing strategy. Once the waiting time of IP connection request exceeds a certain threshold, the request has to be rejected or cancelled. On the other hand, the out-of-date link state information and the signaling delay may also lead to conflicts in resource allocation. From our study, we find that the blocking performance is greatly influenced by the signaling time and information update. New mechanisms are needed to include signaling and information update in DTG formulation in order to provide reasonable performance for traffic grooming in a real network. This has motivated the work described herein.

In this chapter, we investigate the effect of signaling on DTG by comparing the performance of DTG with and without considering the signaling effect. After the analysis of signaling impact on DTG in IP over WDM network, we propose a new technique to combat the adverse effect of signaling. We divide traffic grooming into two steps, 1) Pre-reserve some lightpaths based on statistical traffic observations; 2) Dynamically groom the traffic with the pre-established virtual topology, where some mature algorithms could be used. We call this scheme Dynamic Traffic Grooming with Pre-Reserved Lightpath (DTG-PRL). The traditional scheme that has no pre-reserved lightpaths is called Dynamic Traffic Grooming Basic (DTG-B). Details of the new technique will be presented in subsequent sections.

For a multicast-capable IP over WDM network, the traffic in the IP layer may contain multicast requests with various bandwidth requirements. In this chapter, we also study the scenarios with both unicast IP traffic and multicast IP traffic.

The rest of the chapter is organized as follows. The signaling consideration for DTG in IP over WDM network is introduced in Section 4.1. The proposed DTG-PRL technique is described in Section 4.2. For both unicast and multicast IP traffic models, the ILP formulations and heuristic approaches are presented in Section 4.3 and Section 4.4 respectively. Numerical simulations and results are reported in Section 4.5. Section 4.6 contains a summary of the work described in this chapter.

4.1 Effect of signaling on DTG in IP over WDM networks

Signaling is needed for any resource allocation and release. In this section, we shall discuss the effect of signaling on DTG in the following sequence: 1) General signaling procedure for DTG; 2) Signaling algorithms for IP bandwidth allocation and lightpath provisioning; 3) Problems of DTG due to signaling; and finally 4) Signaling effect as reflected in the simulation results.

4.1.1 Signaling procedure in dynamic traffic grooming

In an IP over WDM network, there are two separate signaling procedures: one at IP layer and another at WDM layer. This is illustrated in Figure 4.1 where a simple IP over WDM network is used. We follow the node architecture as described in [27]. Each node contains the functionality of both IP/MPLS router in the IP layer and OXC in the WDM layer. The communication between IP layer and WDM layer in each node is through the UNI. Two adjoining nodes are connected by a fiber link through the OXC component in each of them. Each fiber link may consist of several fibers, and each fiber may consist of several wavelength channels.

Figure 4.1 Network model for IP over WDM network

For illustration, we consider two scenarios with and without an available lightpath for the incoming IP connection request respectively. The IP connection request may be unicast request or multicast request. We will discuss them separately.

4.1.1.1 Signaling procedure for unicast IP request

Figure 4.2 and Figure 4.3 illustrate the signaling process for the new IP request with and without available lightpath, respectively.

Under the first scenario, when an IP connection request arrives, e.g. $IPReq: A \rightarrow B$ as shown in Figure 4.2, there is no available lightpath for this connection request. The IP/MPLS router in the IP Layer first needs to generate a lightpath request $LPReq: A \rightarrow B$ to the control plane of OXC in the WDM Layer through the UNI (Step 1). The WDM control plane in **Node A** will follow the lightpath provisioning procedure to establish a lightpath from **A** to **B** through the physical fiber links. This involves lightpath signaling to provide reservation and

acknowledgement (*Step 2* and *Step 3*). After the successful establishment of lightpath $LP: A \rightarrow B(w_0)$ from A to B on wavelength w_0 , the WDM control plane will send back a Lightpath Established Acknowledge (*Step 4*) to the IP/MPLS router. Thus a new virtual link $VL: A \rightarrow B$ is established in the virtual topology on **Node A**, and the bandwidth allocation procedure for the IP Connection Request ($IPReq: A \rightarrow B$) will proceed based on the updated virtual topology. Signaling on bandwidth reservation and acknowledgement (*Step 5* and *Step 6*) will be performed.

Figure 4.2 Signaling procedure for DTG without available lightpath for new IP request

Figure 4.3 Signaling procedure for DTG with available lightpath for new IP request

Under the second scenario, assume that a lightpath has already been established, i.e. $LP: A \rightarrow B \rightarrow C(w_1)$ as shown in Figure 4.3. Thus, there exists a virtual link $VL: A \rightarrow C$ in the virtual topology. Suppose $VL: A \rightarrow C$ is available for the incoming IP connection request $IPReq: A \rightarrow C$, one strategy to process this request is that the IP/MPLS router will first find the available routes based on the virtual topology. Then bandwidth allocation will be directly carried out along the $VL: A \rightarrow C$, using the signaling system to make reservation and acknowledgement (*Step 1* and *Step 2*). In case $VL: A \rightarrow C$ is not available for this connection or the bandwidth allocation procedure has failed, a new lightpath needs to be provisioned. This requires the signaling procedure described in the first scenario to be performed.

4.1.1.2 Signaling procedure for multicast IP request

Signaling procedure for multicast IP requests could be much more complicated

than that for unicast IP requests. For simplicity and easy comparison with unicast scheme, we implement the signaling procedure for multicast IP request by adapting the procedures for unicast IP request. The algorithm is illustrated in Figure 4.4.

Figure 4.4 Procedure algorithm of processing multicast IP request

When a multicast IP request is received by the source node, the *Shortest Path Tree* (SPT) will be found based on the status of global virtual topology located in the source node. SPT, derived from the definition of shortest path [51, 90], is defined as the multicast tree that spans the source node and all the destination nodes, and the sum of the distances of all the paths involved is the shortest. To find the SPT based on the virtual topology is the classic Steiner tree problem [51], which is well known to be *NP*-complete. We implement the simple and efficient algorithm, named SPT algorithm [51], to find the SPT. The SPT algorithm comes from the classic Dijkstra's shortest path algorithm. The basic idea of the SPT algorithm can be put in this way. It starts the tree from the source node. It continues to include the destination that is the nearest to the partially established tree until all the destinations are included in the tree.

If the SPT can be found, the multicast IP request will be decomposed into individual unicast IP requests, by which the bandwidth will be allocated separately on each virtual link involved. In case the SPT can not be found based on the current global virtual link states, the multicast IP request will simply be decomposed into individual unicast IP requests from the source node to each destination node belonging to the multicast group. When all the individual unicast IP requests have been successfully accommodated, the multicast IP request is deemed to be successful, otherwise it is deemed as failed. In this way, the multicast IP request is processed via reusing the signaling procedure of unicast IP request.

In the next section, we will discuss the signaling algorithm for IP bandwidth allocation and lightpath provisioning in details.

4.1.2 Signaling algorithms for IP bandwidth allocation and lightpath provisioning

Various routing and signaling protocols can be implemented in IP bandwidth allocation and lightpath provisioning. Different protocols and strategies may impact the DTG differently. Similarly, for both the peer model and overlay model, the signaling algorithms will have a significant impact on the network performance. For simplicity, we adopt the overlay model for IP over WDM networks [12, 35, 96], and a specific routing and signaling protocol in this study.

From the performance study in Chapter 3, we already know that the approach with distributed global information-based link state and forward-reservation sequential signaling has the lowest CST, relatively low CMO and comparable RBP among all the approaches with distributed global information-based link state. Therefore, we choose this approach for the routing and signaling here. As the implementation of these protocols in IP layer and WDM layer are quite similar, we categorically call the procedures *Resource Reservation* and *Resource Release*. The signaling algorithms are illustrated in Figure 4.4 and Figure 4.5.

In *Resource Reservation*, the source node selects a route and then sends a reservation request. For lightpath provisioning in the WDM layer, a fixed wavelength is also chosen. The selection of route as well as wavelength is based on the distributed global link state table on this node, where the link state information of each link in the network has been stored. For lightpath provisioning, link state simply means the status of each wavelength in each fiber link, whether it is “Occupied” or “Free”. For IP bandwidth allocation, the link state contains the status of each virtual link in the virtual topology, and the status indicates the value of the allocated bandwidth varying from 0 to the maximum capacity of the wavelength.

The reservation request is forwarded sequentially between two neighboring nodes along the selected route. Each intermediate node tries to reserve the resource, forwards the reservation message and then begins to set the devices for the connection. The timing sequence of signaling procedure is illustrated in Figure 4.4, and the signaling procedures are described in Section 3.3.3.

During the *Resource Release*, the source node sends out a release message along the activated route after receiving a disconnection request. The release message is forwarded sequentially from one node to another along the route. Each intermediate node releases the resource, forwards the release message and adjusts the device accordingly, and an *ACK* message will be sent back to the source node after the device adjustment. When all the *ACK* messages from the nodes along the route return to the source node, the connection is successfully taken down, and the global link state table in the source node will be updated. Figure 4.5 shows the timing sequence.

Figure 4.5 Timing sequence of resource reservation signaling procedure

Figure 4.6 Timing sequence of resource release signaling procedure

4.1.3 Problems of DTG due to the signaling

In practical networks, the function of DTG requires signaling support. Although the signaling procedure can be implemented in a variety of ways, some problems are inherent in the process irrespective of which strategy is adopted. These problems may produce adverse effect on the network performance. We shall discuss four of the most crucial factors below.

4.1.3.1 Contention conflict

On receiving a reservation request, the source chooses a suitable route based on the link state information at that time and sends the request. Nonetheless, signaling, though fast, takes time. During the signaling trip, the chosen resource may have already been occupied or become unavailable. The resource could be IP bandwidth or WDM wavelength. This situation is called *contention conflict*. In distributed network, contention conflict is common and unavoidable. However, previous works on dynamic traffic grooming did not consider signaling and ignore this problem.

They assume that resource allocation could be accomplished instantaneously. This is not realistic. In a practical network, especially when the traffic load is relatively high, contention conflict can cause substantial blocking which cannot be neglected.

4.1.3.2 *Blocking due to outdated link state*

In a network based on distributed link state, such as OSPF, the source node may choose an unavailable route if the global link state is outdated. This will result in blocking. This kind of blocking is highly dependable on the updating strategies of link state. The purpose of link state update is to keep the global link state table up-to-date and identical in network nodes. In reality, updating takes time to cover the entire network. Therefore, blocking due to outdated link state is unavoidable. How to effectively update and spread the information are key issues to be addressed. Frequent update will keep information afresh but the control overheads of control messages will be high. Infrequent update will incur less overheads but the information is likely to be outdated. Different strategies will produce different effects on DTG. This problem is rather complex and out of the scope of this thesis. S. Shen has conducted a detailed study on this problem [97].

4.1.3.3 *Blocking due to node processing delay*

The node processing delay has an effect on blocking. If the node capacity is low, the excess events that are beyond the node processing capacity will either be discarded or queued, depending on the strategy adopted. Once the waiting time of IP connection request exceeds a certain threshold that the user can endure, the request has to be cancelled. As the traffic load increases, the blocking will become worse. This is despite the fact that some bandwidth resources may still be available. Ways to alleviate the problem include optimizing the queuing policy, decreasing control message overheads, enhancing the processing speed which is hardware

dependent, and so on.

4.1.3.4 *Double-layer signaling*

For DTG, the signaling procedures are more complex due to double-layer signaling, i.e. IP layer and WDM layer signaling. Signaling in WDM layer is responsible for lightpath provisioning. Signaling in IP layer is needed for bandwidth allocation. As shown in Figure 4.1, the signaling procedures in these two layers are independent. This means independent signaling protocols are used in each layer. Nonetheless, as the IP bandwidth allocation depends on the virtual topology, the blocking probability of the IP connection requests is influenced by the variation of the virtual topology, i.e. the establishment and release of lightpaths. One key strategy of achieving a high IP bandwidth utilization is to reduce the blocking probability of lightpath connections and maintain a relatively stable virtual topology.

4.1.4 **Signaling influence on DTG**

In order to find out the effect of signaling on DTG, we carry out a performance comparison between DTG-B without signaling and DTG-B with signaling for both unicast and multicast IP traffic models.

We evaluate the performance by IP Bandwidth Blocking Probability (IBBP) and Network Resource Utilization (NRU). For a fair comparison, a so-called basic strategy is applied, i.e. no retry is attempted. We assume that the event queue at each node is infinite. The information update threshold is set to 1, which means as soon as the link state information changes, an update message will be spread to the entire network. Detailed information about the simulation is provided in Section 4.5, and the topologies used are Simple6 shown in Figure A.1 and NSFNet shown in Figure A.5.

Figure 4.7 IBBP vs. TL, Simple6 network, DTG-B with/without signaling

Figure 4.8 IBBP vs. TL, NSFNet network, DTG-B with/without signaling

Figure 4.9 NRU vs. TL, Simple6 network, DTG-B with/without signaling

Figure 4.10 NRU vs. TL, NSFNet network, DTG-B with/without signaling

Figures 4.7 ~ 4.10 show the comparison results. It can be seen that the DTG-B without signaling has lower IBBP and higher NRU for both Simple6 and NSFNet networks. As the simulations are conducted in exactly the same environment, the difference can only be attributed to the signaling effect. For DTG-B without signaling, real-time link state updating (threshold = 1) is assumed to be perfect, that will contribute to good performance. However, for DTG-B with signaling, real-time link state updating takes time, which aggravates blocking performance. Even at a low load, the IBBP increases sharply and the NRU reaches its upper bound because of the large processing delay caused by a large number of updating messages. This demonstrates that the signaling and information update are key factors affecting the DTG performance and should be studied properly.

In summary, we have explained the signaling procedure and protocols, as well as the effect of signaling on DTG in this section. The simulation results show that the performance of DTG is greatly affected by signaling. As signaling is required in a practical network, new mechanisms are needed to combat the negative impact of signaling. In the next section, we shall present the proposed mechanism, DTG-PRL, to solve the problem.

4.2 Dynamic traffic grooming with pre-reserved lightpaths

Although contention conflict, outdated link state and node processing delay may occur in both IP and WDM layers, in real-life network with traffic grooming, the IP traffic will be much more dynamic than the WDM traffic which is a consolidation of the IP traffic. In other words, the WDM traffic will fluctuate significantly. If there is a way to find out the optimal lightpath distribution and

pre-reserve those lightpaths, contention conflict, outdated link state and node processing delay in the WDM layer will reduce. In this section, we propose a new DTG-PRL mechanism to reduce the negative effect of signaling. We make use of a pre-reserved virtual topology to carry out dynamic traffic grooming. Two steps are involved: 1) pre-reserve some lightpaths based on statistical traffic observations and 2) dynamically groom the traffic with the pre-established virtual topology.

The concept of pre-reserving resources has been applied in lightpath protection and restoration [98]. The idea of an envelope has been developed in [98], which is based on reserving wavelength links. We extend this concept and apply it to the problem of dynamic traffic grooming with signaling. In our approach, a lightpath is used as the minimum unit of the resources that will be pre-reserved. This differs from the previous envelope concept which uses separate wavelength links as the minimum unit of resource. The reason for adopting this approach lies in the following two aspects.

First, step 1) will reduce the conflicts brought in by the process of lightpath establishment, as the pre-reserved lightpaths are carefully selected. Once reserved, the lightpaths will not be released unless the statistical properties of the traffic are changed. However in DTG-B, the virtual topology has to be established dynamically. Under low IP traffic load, lightpaths are set up and released frequently, which tends to increase the blocking probability of lightpath provisioning. Under high IP traffic load, dynamic lightpath establishment faces resource shortage and the blocking probability becomes high. We expect a well pre-planned virtual topology will eliminate these drawbacks.

Second, if the virtual topology is well established, the control messages for signaling in WDM layer and link state updating will be reduced. Lower blocking

probability and shorter connection setup time are expected in the DTG-PRL mechanism.

In this way, the dynamic traffic grooming problem can be divided into two sub-problems: 1) How to select the pre-reserved lightpaths, i.e. pre-establish the virtual topology; 2) How to dynamically groom the traffic on the established virtual topology.

Sub-problem 1) is a virtual topology design problem and some algorithms have been developed previously (e.g. [27]) for static traffic grooming. For dynamic traffic grooming, applying the existing static algorithms is inappropriate. The challenging problems are how to generate a well-balanced traffic matrix that matches the distribution of dynamic traffic load, and how to obtain the optimal virtual topology suitable for dynamic traffic grooming.

We solve the sub-problem 1) by proposing an ILP formulation, as well as a heuristic algorithm for large networks. As the pre-reserved lightpaths will not consume all the resources in the whole network, the unused wavelength links are still available for a DTG-B scheme; and some suitable algorithms that have been proposed in the literature can be adopted to solve sub-problem 2). We will concentrate on solving sub-problem 1) here, and compare the performance of DTG-PRL scheme with DTG-B scheme by means of simulation. The ILP formulation and the heuristic algorithm will be described in the following sub-sections. A comparison between the ILP formulation and heuristic algorithm will be made by means of numerical analysis in Section 4.3.

4.3 ILP formulation for DTG-PRL

How to select pre-reserved lightpaths for DTG-PRL can be formulated as an

ILP optimization problem. In our formulation, both throughput and wavelength cost are considered. The throughput is defined as the total traffic of successfully routed IP connection requests with different bandwidth demands. The requested bandwidth of low-speed traffic (IP connections) varies and can be OC-1, OC-3, OC-6, OC-12, OC-24 or OC-48. The wavelength cost is defined as the number of occupied wavelengths on each fiber link in the entire network. We also make the following assumptions in our study:

- There is one fiber link in each direction between neighboring nodes. Each fiber consists of 64 wavelengths;
- No wavelength conversion is considered;
- The transceivers in a network node can be tuned to any wavelength on the fiber;
- Each node has sufficient transceivers for input and output lightpaths;
- The event queue in each node is large; i.e., no signaling events will be discarded.

The objective of the optimization process is to find out how the established lightpaths are distributed based on a given traffic matrix and network topology. The optimal lightpath distribution is defined as a situation where the established lightpaths may provide maximum traffic throughput and occupy as few wavelength links as possible. In order to achieve both objectives, we adopt a two-round ILP calculation. As achieving the maximum traffic throughput is the main objective, it will be calculated in the first round. The second round is meant to calculate the minimum number of wavelength link; the result of the first round will be counted as an additional constraint. We discuss the formulations for unicast IP traffic model and multicast IP traffic model separately.

4.3.1 ILP Formulation for unicast IP traffic model

The notations used are listed in Table 4.1.

Table 4.1 ILP notations for unicast IP traffic model

m, n	denote the endpoints of a physical fiber link that might occur in a lightpath;
i, j	denote the originating and terminating nodes for a lightpath. A lightpath may traverse single or multiple physical fiber links;
s, d	denote the source and destination of the end-to-end traffic request. The end-to-end traffic may traverse through a single or multiple lightpaths;
x	denotes the granularity of low-speed traffic requests. We assume $x \in \{1, 3, 6, 12, 24, 48\}$, which means that traffic demands between node pairs can be any of OC-1, OC-3, OC-6, OC-12, OC-24 and OC-48;
r	denotes the index of OC- x traffic request for any given node pair (s, d) . For instance, if there are 5 OC- x requests between node pair (s, d) , then $r \in [1, 5]$;
N	denotes the number of nodes in the network;
W	denotes the number of wavelengths per fiber. We assume all of the fibers in the network carry the same number of wavelengths;
C	denotes the maximum capacity of each wavelength. We use OC-1 as the unit;
F_{mn}	denotes the number of fiber interconnecting node m and node n . For our study in this thesis, $F_{mn} \in \{0, 1\}$. $F_{mn} = F_{nm} = 0$ indicates no physical connection between node m and node n . $F_{mn} = F_{nm} = 1$ indicates there exists one unidirectional fiber link between these two nodes in both $m \rightarrow n$ and $n \rightarrow m$ direction;
Λ^x	denotes the traffic matrix set of OC- x connection requests. $\Lambda^x = \{\Lambda_{sd}^x\}$, where Λ_{sd}^x is the number of OC- x connection requests between node pair (s, d) ;
P_{ij}^w	denotes the number of lightpaths from node i to node j on wavelength w . $P_{ij}^w > 1$ means the lightpaths between node i to node j on wavelength w may take different paths;
$L_{mn}^{ij,w}$	denotes the number of lightpath from node i to node j on wavelength

w , employing fiber (m, n) . $L_{mn}^{ij,w} \in \{0,1\}$;

Q_{mn}^w denotes whether the wavelength w is occupied on fiber (m, n) or not.
 $0 \leq Q_{mn}^w \leq F_{mn}$. For our study here, $Q_{mn}^w = 1$ means w is occupied,
otherwise $Q_{mn}^w = 0$;

$\lambda_{ij,x}^{sd,r}$ $\lambda_{ij,x}^{sd,r} = 1$ means the r^{th} OC- x low-speed traffic request from node s to
node d employs lightpath (i, j) as an intermediate virtual link;
otherwise $\lambda_{ij,x}^{sd,r} = 0$;

$S_{sd}^{x,r}$ $S_{sd}^{x,r} = 1$ means the r^{th} OC- x low-speed traffic request from node s to
node d has been successfully routed; otherwise $S_{sd}^{x,r} = 0$.

The ILP is formulated as in Table 4.2.

Table 4.2 ILP formulation for unicast IP traffic model

Round 1 – Maximizing the throughput

Objective function:

$$\text{Maximize: } \sum_{s,d,x,r} x \cdot S_{sd}^{x,r} \quad (4.1)$$

Constraints - on physical route variables:

$$\sum_m L_{mk}^{ij,w} = \sum_n L_{kn}^{ij,w}, \quad k \neq i, j \quad \forall i, j, w, k \quad (4.2)$$

$$\sum_m L_{mi}^{ij,w} = \sum_n L_{jn}^{ij,w} = 0, \quad \forall i, j, w \quad (4.3)$$

$$\sum_n L_{in}^{ij,w} = \sum_m L_{mj}^{ij,w} = P_{ij}^w, \quad \forall i, j, w \quad (4.4)$$

$$\sum_{i,j} L_{mn}^{ij,w} = Q_{mn}^w, \quad \forall m, n, w \quad (4.5)$$

$$Q_{mn}^w \leq F_{mn}, \quad \forall m, n, w \quad (4.6)$$

$$L_{mn}^{ij,w} \in \{0,1\} \quad (4.7)$$

Constraints - on virtual-topology traffic variables:

$$\sum_i \lambda_{ik,x}^{sd,r} = \sum_j \lambda_{kj,x}^{sd,r}, \quad k \neq s, d \quad \forall s, d \quad r \in [1, \Lambda_{sd}^x] \quad (4.8)$$

$$\sum_i \lambda_{is,x}^{sd,r} = \sum_j \lambda_{dj,x}^{sd,r} = 0, \quad \forall s, d \quad r \in [1, \Lambda_{sd}^x] \quad (4.9)$$

$$\sum_j \lambda_{sj,x}^{sd,r} = \sum_i \lambda_{id,x}^{sd,r} = S_{sd}^{x,r}, \quad \forall s, d \quad r \in [1, \Lambda_{sd}^x] \quad (4.10)$$

$$\sum_{s,d,x,r} x \cdot \lambda_{ij,x}^{sd,r} \leq C \cdot \sum_w P_{ij}^w, \quad \forall i, j \quad (4.11)$$

$$S_{sd}^{x,r} \in \{0,1\} \quad (4.12)$$

Round 2 – Minimizing the wavelength links

Objective function:

$$\text{Minimize:} \quad \sum_w \sum_{m,n} Q_{mn}^w \quad (4.13)$$

Additional Constraint for 2nd round:

Suppose the result of the object value in the first round is $T_{\max}$, then

$$\sum_{s,d,x,r} x \cdot S_{sd}^{x,r} = T_{\max} \quad \forall s,d,x,r \quad (4.14)$$

Equation (4.1) shows the optimization objective function for maximizing the traffic throughput. Equations (4.2) to (4.4) are the multi-commodity equations (flow conservation) that account for the routing of a lightpath from its origin to its termination. Equations (4.5) to (4.7) ensure that wavelength w on one fiber link (m,n) can only be present in at most one lightpath in the virtual topology. Equations (4.8) to (4.12) are responsible for the routing of low-speed traffic requests on the virtual topology, and they take into account the fact that the aggregate traffic flowing through lightpaths cannot exceed the overall wavelength capacity. Equation (4.13) shows the optimization objective function for minimizing the wavelength links based on all the constraints from (4.2) to (4.12), and equation (4.14) to dictate that the total traffic throughput should be equal to the object value produced in the first round.

4.3.2 ILP Formulation for multicast IP traffic model

A multicast IP traffic request can be represented by

$$Req = \{s, D\}, \quad s \in V_N, \quad D = \{d_i \mid d_i \in V_N, d_i \neq s, 0 \leq i \leq z-1\} \quad (4.15)$$

Here V_N denotes the set of nodes in the network, s is the source node, D is a set of destination nodes, and z is the number of destinations of Req . In our discussion, the multicast traffic request will have only one source node, hence the

multicast group size of Req is $1+z$. The ILP notations and formulations for multicast IP traffic model presented in Table 4.3 and Table 4.4 are extended from the ILP formulation for the unicast IP traffic model,.

Table 4.3 ILP notations for multicast IP traffic model

m, n	denote the endpoints of a physical fiber link that might occur in a lightpath;
i, j	denote the originating and terminating nodes for a lightpath. A lightpath may traverse single or multiple physical fiber links;
s, D	denote the group of IP multicast request, s denotes the source node, D denotes the group of destination nodes. The IP multicast traffic must traverse through multiple lightpaths when $\ D\ > 1$;
d	denotes one single destination node, where $d \in D$;
z	denotes the number of destinations of Req , $z = \ D\ > 1$;
x	denotes the granularity of low-speed traffic requests. We assume $x \in \{1, 3, 6, 12, 24, 48\}$, which means that traffic demands of one IP multicast traffic request can be any of OC-1, OC-3, OC-6, OC-12, OC-24 and OC-48;
r	denotes the index of OC- x traffic request for any given multicast group (s, D) . For instance, if there are 5 OC- x requests for multicast group (s, D) , then $r \in [1, 5]$.
N	denotes the number of nodes in the network;
W	denotes the number of wavelengths per fiber. We assume all of the fibers in the network carry the same number of wavelengths;
C	denotes the maximum capacity of each wavelength. We use OC-1 as the unit;
F_{mn}	denotes the number of fiber interconnecting node m and node n . For our study in this paper, $F_{mn} \in \{0, 1\}$. $F_{mn} = F_{nm} = 0$ indicate no physical connection between node m and node n . $F_{mn} = F_{nm} = 1$ indicate there exists one unidirectional fiber link between these two nodes in both $m \rightarrow n$ and $n \rightarrow m$ direction;
Λ^x	denotes the traffic matrix set of OC- x connection requests. $\Lambda^x = \{\Lambda_{sD}^x\}$, where Λ_{sD}^x is the number of OC- x connection requests for multicast group

(s, D) ;

P_{ij}^w denotes the number of lightpaths from node i to node j on wavelength w . $P_{ij}^w > 1$ means the lightpaths between node i and node j on wavelength w may take different paths;

$L_{mn}^{ij,w}$ denotes the number of lightpath from node i to node j on wavelength w , employing fiber (m, n) . $L_{mn}^{ij,w} \in \{0, 1\}$;

Q_{mn}^w denotes whether the wavelength w is occupied on fiber (m, n) or not. $0 \leq Q_{mn}^w \leq F_{mn}$. For our study here, $Q_{mn}^w = 1$ means w is occupied, otherwise $Q_{mn}^w = 0$;

$\lambda_{ij,x}^{sD,r}$ $\lambda_{ij,x}^{sD,r} = 1$ means the r^{th} OC- x low-speed multicast request from node s to destination group D employs lightpath (i, j) as an intermediate virtual link; otherwise $\lambda_{ij,x}^{sD,r} = 0$;

$S_{d,x}^{sD,r}$ $S_{d,x}^{sD,r} = 1$ means the r^{th} OC- x low-speed multicast request from node s to destination group D has been successfully routed for destination d , where $d \in D$; otherwise $S_{d,x}^{sD,r} = 0$;

$S_{sD}^{x,r}$ $S_{sD}^{x,r} = z$ means the r^{th} OC- x low-speed traffic request from node s to destination group D has been successfully routed for all the destinations in D ; otherwise $S_{sD}^{x,r} = 0$.

Table 4.4 ILP formulation for multicast IP traffic model

Round 1 – Maximizing the throughput

Objective function:

$$\text{Maximize: } \sum_{s,D,x,r} x \cdot S_{sD}^{x,r} \quad (4.16)$$

Constraints - on physical route variables:

$$\sum_m L_{mk}^{ij,w} = \sum_n L_{kn}^{ij,w}, \quad k \neq i, j \quad \forall i, j, w, k \quad (4.17)$$

$$\sum_m L_{mi}^{ij,w} = \sum_n L_{jn}^{ij,w} = 0, \quad \forall i, j, w \quad (4.18)$$

$$\sum_n L_{in}^{ij,w} = \sum_m L_{mj}^{ij,w} = P_{ij}^w, \quad \forall i, j, w \quad (4.19)$$

$$\sum_{i,j} L_{mn}^{ij,w} = Q_{mn}^w, \quad \forall m, n, w \quad (4.20)$$

$$Q_{mn}^w \leq F_{mn}, \quad \forall m, n, w \quad (4.21)$$

$$L_{mn}^{ij,w} \in \{0, 1\} \quad (4.22)$$

Constraints - on virtual-topology traffic variables:

$$\sum_i \lambda_{ik,x}^{sD,r} = \sum_j \lambda_{kj,x}^{sD,r}, \quad k \neq s, k \notin D, \forall (s, D) \quad r \in [1, \Lambda_{sD}^x] \quad (4.23)$$

$$0 \leq \sum_j \lambda_{kj,x}^{sD,r} \leq z, \quad k = s \text{ or } k \in D, \forall (s, D) \quad r \in [1, \Lambda_{sD}^x] \quad (4.24)$$

$$\sum_i \lambda_{is,x}^{sD,r} = 0, \quad \forall (s, D) \quad r \in [1, \Lambda_{sD}^x] \quad (4.25)$$

$$\sum_i \lambda_{id,x}^{sD,r} = S_{d,x}^{sD,r}, \quad d \in D, \forall (s, D) \quad r \in [1, \Lambda_{sD}^x] \quad (4.26)$$

$$\sum_{d \in D} S_{d,x}^{sD,r} = S_{sD}^{x,r}, \quad \forall (s, D) \quad r \in [1, \Lambda_{sD}^x] \quad (4.27)$$

$$\sum_{s,D,x,r} x \cdot \lambda_{ij,x}^{sD,r} \leq C \cdot \sum_w P_{ij}^w, \quad \forall i, j \quad (4.28)$$

$$S_{d,x}^{sD,r} \in \{0, 1\} \quad (4.29)$$

$$S_{sD}^{x,r} \in \{0, z\} \quad (4.30)$$

Round 2 – Minimizing the wavelength links

Objective function:

$$\text{Minimize:} \quad \sum_w \sum_{m,n} Q_{mn}^w \quad (4.31)$$

Additional Constraint for 2nd round:

Suppose the result of the object value in the first round is $T_{\max}$, then

$$\sum_{s,D,x,r} x \cdot S_{sD}^{x,r} = T_{\max} \quad \forall (s, D), \forall x, \forall r \quad (4.32)$$

For a network with N nodes, given a multicast group size $1+z$, the total number of set D can be denoted by C_{N-1}^z , and the total number of multicast group (s, D) can be given by:

$$N \cdot C_{N-1}^z = \frac{N!}{z! \cdot (N-1-z)!} \quad (4.33)$$

Equation (4.16) shows the optimization objective function for maximizing the traffic throughput. Equations (4.17) to (4.22) are the multi-commodity equations (flow conservation) that account for the routing of a lightpath from its origin to its termination, which are the same as Equations (4.3) to (4.8) in ILP formulations for

unicast IP traffic model. Equations (4.23) to (4.30) are responsible for the routing of low-speed traffic requests on the virtual topology, and they take into account the fact that the aggregate traffic flowing through a set of lightpaths cannot exceed the overall wavelength capacity. Compared with the ILP formulations for unicast IP traffic model, Equation (4.24) serves to restrict outbound traffic from any source node or destination node from exceeding the multicast destination group size, which avoids loops to be created when the multicast tree is established over the virtual topology. Equation (4.25) serves to ensure that the inbound traffic to the source node is zero, and Equation (4.26) is meant for evaluating the routing status of an individual destination. Equation (4.27) is meant for evaluating the routing status of all the destinations, which also shows the final routing status of the multicast request. Equation (4.31) shows the optimization objective function for minimizing the wavelength links based on all the constraints from (4.17) to (4.30), as well as an additional Equation (4.31) which reflects that the total traffic throughput is equal to the object value produced in the first round.

4.4 Heuristic approach for DTG-PRL

It is well known that the traffic grooming problem for a mesh network is an *NP*-complete problem [27]. The number of variables and equations increases exponentially as the numbers of nodes and wavelengths on each fiber increase. Consequently, we have to resort to a heuristic approach for DTG-PRL for large networks. For both unicast and multicast IP traffic models, we develop two heuristic algorithms, namely Maximizing Wavelength Utilization (MWU) and MWU for multicasting (MWU-M), respectively. The proposed heuristic algorithms are aimed at both maximizing the traffic throughput and minimizing the wavelength cost as in

the ILP formulations.

4.4.1 MWU algorithm

The MWU algorithm follows a simple rule that the higher the wavelength utilization for the traffic between a node pair, the higher is its priority for resource allocation. In this way, MWU may find an optimal distribution of established lightpaths, i.e. large traffic with low wavelength cost is likely to be accepted. MWU is described in Table 4.5.

Table 4.5 Maximizing wavelength utilization algorithm

-
1. For each node pair (s, d), calculate the sum of unallocated traffic request $T(s, d)$ and the shortest hop distance $H(s, d)$ on the physical topology between (s, d); 2. For each node pair (s, d), calculate the wavelength utilization: $WU(s, d) = T(s, d) / H(s, d)$; 3. Find out the maximum $WU_{\max} = T(s', d') / H(s', d')$; 4. Try to setup a lightpath between (s', d') using First-Fit wavelength assignment and shortest-path routing, subject to the wavelength continuity constraint; 5. If the lightpath setup fails, that means no lightpath can accommodate any traffic request between (s', d'), hence set $T(s', d') = 0$, the traffic is blocked; otherwise, set $T(s', d') = \text{Max}[T(s', d') - C, 0]$, where C denotes the capacity of the wavelength. 6. Goto step 1 until $T(s, d) = 0 \quad \forall s, d$.
-

4.4.2 MWU-M algorithm

The MWU-M algorithm will follow the rule that: the higher the wavelength utilization of the IP multicast request group (s, D) , the higher its priority for resource allocation. The difference between MWU and MWU-M is that the former only needs to find the shortest lightpath from the source to the destination while the latter needs to find the SPT which consists of multiple lightpaths (virtual links), spanning the source node and all the destination nodes in the multicast group (s, D) . Thus, MWU-W is able to find an optimal distribution of established lightpaths over the physical topology of a network such that large traffic with low

wavelength cost has a higher chance to be accepted. MWU-M is described in Table 4.6.

Table 4.6 Maximizing wavelength utilization for multicasting algorithm

-
1. For each IP multicast request group (s, D) , calculate the sum of unallocated traffic request $T(s, D)$;
 2. For each IP multicast request group (s, D) , find the multicast tree with least involved hop number $H(s, D)$ on the physical topology;
 3. For each IP multicast request group (s, D) , calculate the wavelength utilization: $WU(s, D) = T(s, D) / H(s, D)$;
 4. Find out the maximum $WU_{\max} = T(s', D') / H(s', D')$;
 5. Try to setup multiple lightpaths to satisfy the multicast tree for (s', D') using First-Fit wavelength assignment and shortest-path routing, subject to the wavelength continuity constraint for individual lightpath;
 6. If any one of the lightpath is failed to setup, that means no multicast tree can accommodate any multicast traffic request for group (s', D') , hence set $T(s', D') = 0$, the multicast traffic is blocked; otherwise, set $T(s', D') = \text{Max}[T(s', D') - C, 0]$, where C denotes the capacity of the wavelength.
 7. Goto step 1 until $T(s, D) = 0 \quad \forall s \quad \forall D$.
-

4.5 Simulation and numerical results

4.5.1 Simulation description

In the simulation we use the *First Come First Serve* (FCFS) strategy for processing each IP connection request, and *First-Fit* (FF) strategy for choosing the available lightpaths and wavelengths. The common procedure for DTG, which is used to solve sub-problem 2), can make use of any algorithm proposed in the literatures [28-32]. Without loss of generality, we adopt a common DTG algorithm as described in Table 4.7 for both DTG-B and DTG-PRL schemes.

Table 4.7 Common DTG algorithm for DTG-B and DTG-PRL

1.	Get an IP connection request;
2.	Find the available route based on the virtual topology, if found, goto step 3; otherwise goto step 4;
3.	Proceed with the signaling to allocate bandwidth resources for this connection; if success, the connection will be set up, goto step 1; otherwise goto step 4;
4.	Request for a new lightpath and proceed with the signaling for lightpath provisioning; if the new lightpath is successfully established, goto step 3; otherwise the connection is blocked.

There are some differences between DTG-PRL and DTG-B when the abovementioned algorithm is executed. For DTG-B, the virtual topology is empty at the beginning, and will have to be dynamically changed as the lightpath is established and released during the simulation. However, for DTG-PRL, it has a well established virtual topology at the beginning, and pre-reserved lightpaths will not be released during the simulation. Note that in DTG-PRL, pre-reserved lightpaths do not occupy all the wavelength resources; the spare capacity is available for dynamic lightpath set up just like DTG-B. Consequently, the virtual topology in DTG-PRL is contributed by both the pre-reserved lightpaths and the dynamically established lightpaths.

4.5.2 Topology and assumptions

We adopt the Simple6 shown in Figure A.1, and the NSFNet shown in Figure A.5 for simulation. The following assumptions are made for both of them:

- Each fiber contains 64 wavelengths. Wavelength capacity is OC-192 (10Gbps).
- The Pareto distribution [93, 99] is adopted for IP traffic in order to make the comparison easier and the signaling effects to be seen easily. The traffic over WDM layer is dominated by the demands from the IP layer, according to the

specific routing and signaling strategies. Both the IP request arrival interval $1/\lambda$ and the IP service holding time $1/\mu$ follow the Pareto distribution with $Shape = 1.2$ (typical value for Ethernet traffic). Mean service holding time is 1.0 second, and traffic load in Erlang for each node-pair is calculated by means of λ/μ . The processing time for a control message at each node is $10\mu s$, where $1\mu s = 10^{-6}s$. The processing time for a control message at the source node or the manager is $30\mu s$ because of the complexity of the connection request. The time to set up a switch connection is $500\mu s$. The propagation delay is $5\mu s/km$.

- The given traffic matrix $\Lambda^x = \{\Lambda_{sd}^x\}$ defined in Table 4.1 for unicast and $\Lambda^x = \{\Lambda_{sD}^x\}$ defined in Table 4.3 for multicast are both randomly generated according to the traffic model described above. Λ_{sd}^x and Λ_{sD}^x denote the average number of OC- x connection requests for node pair (s, d) and for multicast group (s, D) , respectively, that have been normalized by the mean service holding time.
- We set an observation period of over 100,000 IP connection requests on each node. The request bandwidth follows the ratio of OC-1 : OC-3 : OC-6 : OC-12 : OC-24 : OC-48 = 1 : 1 : 1 : 1 : 1 : 1.
- Four performance metrics are employed, the definitions of which are given below:
 - 1) *IP Bandwidth Blocking Probability (IBBP)*: This is the ratio of the requested IP bandwidth which is not successfully accommodated by the total requested IP bandwidth of all IP connections:

$$IBBP = 1 - \frac{\sum_{(Successful\ IP\ Connection\ Request)} (Requested\ BW)}{\sum_{(Executed\ IP\ Connection\ Request)} (Requested\ BW)} \quad (4.34)$$

- 2) *Network Resource Utilization* (NRU): This is the ratio of effectively utilized IP bandwidth to the total usable link capacity in the entire network during the simulation:

$$NRU = \frac{\sum_{(Successful\ IP\ Connection\ Request)} (Requested\ BW \times Service\ Holding\ Time)}{(Simulation\ Period) \times \sum_{(Links)} \sum_{(Wavelengths)} (Wavelength\ Capacity)} \quad (4.35)$$

- 3) *IP Connection Setup Time* (ICST): The average time period from the arrival of the IP connection request to the time when the IP connection is successfully set up. Note that, only successful IP connections are counted:

$$ICST = \frac{\sum_{(Successful\ IP\ Connection\ Request)} (t_{successfully\ set\ up} - t_{request\ arrival})}{\sum_{(Successful\ IP\ Connection\ Request)}} \quad (4.36)$$

- 4) *Control Message Efficiency* (CME): This is the ratio of IP connection throughput in bits to control message overhead in bits. CME denotes how much effective traffic is successfully accommodated by a certain length of control messages overhead. Higher CME shows better performance efficiency in term of control message overhead:

$$CME = \frac{\sum_{(Successful\ IP\ Connection\ Request)} (Requested\ BW \times Service\ Holding\ Time) \text{ (bits)}}{\sum_{(Control\ Messages)} (Length\ of\ Message) \text{ (bits)}} \quad (4.37)$$

- Different strategies on information message updating and signaling retry policies in IP or WDM layer may impact the network performance strongly. Comparisons among different strategies are rather complex, and need further study. For a clear and fair comparison between DTG-B and DTG-PRL, we

adopt the following strategy. The information updating threshold is set to 1 in the WDM layer and 1, 5, 10 in the IP layer; no retry is allowed during signaling in both IP and WDM layers. As the IP connection requests are more dynamic than the lightpath connection requests, the control overhead in the IP layer is heavier than that in the WDM layer. Thus, we set 1 as the updating threshold in the WDM layer, and several different values in the IP layer, in order to simultaneously compare the influence brought in by information updating.

4.5.3 Numerical results

4.5.3.1 *Comparison between ILP formulation and heuristic approach*

Figures 4.11 ~ 4.14 show the comparison results of ILP formulation and heuristic algorithm MWU on the Simple6 network. The simulations are based on the scenarios of DTG-PRL with signaling. We focus only on the graphs of ILP and MWU under various conditions. It can be seen that, the graphs of MWU have a good match with those of ILP. The average absolute errors of MWU compared with ILP in IBBP are only about 2.1%, 0.81% and 0.72% for $U=1, 5, 10$ respectively. And in NRU, the average absolute errors are only about 0.55%, 0.20% and 0.17% for $U=1, 5, 10$ respectively. Thus, the proposed heuristic algorithm MWU provides comparable results as the ILP formulation. Subsequently, when we employ MWU for the simulation of a larger network, NSFNet, the results should be reliable. Similar results for comparing the performances of multicast ILP formulation and heuristic algorithm MWU-M on Simple6 network are shown in Figures 4.15 ~ 4.18.

Figure 4.11 Comparison between ILP-Unicast and MWU: IBBP vs. TL

Figure 4.12 Comparison between ILP-Unicast and MWU: NRU vs. TL

Figure 4.13 Comparison between ILP-Unicast and MWU: ICST vs. TL

Figure 4.14 Comparison between ILP-Unicast and MWU: CME vs. TL

Figure 4.15 Comparison between ILP-Multicast and MWU-M: MBBP vs. TL

Figure 4.16 Comparison between ILP-Multicast and MWU-M: MNRU vs. TL

Figure 4.17 Comparison between ILP-Multicast and MWU-M: MCST vs. TL

Figure 4.18 Comparison between ILP-Multicast and MWU-M: MCME vs. TL

4.5.3.2 Comparison between DTG-B and DTG-PRL with unicast IP traffic model

Figures 4.19 ~ 4.26 show the performance comparisons using the Simple6 and NSFNet networks for unicast IP traffic model. To keep the figures simple, for Simple6, we show only the ILP results whereas for NSFNet, we show the heuristic results. We make comparisons between scenarios of DTG-PRL and DTG-B with and without signaling in order to find the differences. For those strategies with signaling, the information updating threshold in IP layer, U , is set to 1, 5 and 10. For those strategies without signaling, the ICST and the CME are meaningless, so only strategies with signaling are shown in the figures of ICST and CME. In the figure, WS denotes “with signaling”, NS denotes “without signaling”, U denotes “Information updating threshold in the IP layer”.

Figure 4.19 Unicast performance comparison on Simple6: IBBP vs. TL

Figure 4.20 Unicast performance comparison on NSFNet: IBBP vs. TL

Figure 4.21 Unicast performance comparison on Simple6: NRU vs. TL

Figure 4.22 Unicast performance comparison on NSFNet: NRU vs. TL

Figure 4.23 Unicast performance comparison on Simple6: ICST vs. TL

Figure 4.24 Unicast performance comparison on NSFNet: ICST vs. TL

Figure 4.25 Unicast performance comparison on Simple6: CME vs. TL

Figure 4.26 Unicast performance comparison on NSFNet: CME vs. TL

It can be seen that in both Simple6 and NSFNet, DTG-PRL schemes outperform DTG-B schemes under the same strategy in all four metrics. This means DTG-PRL schemes have lower IP bandwidth blocking probability, higher network resource utilization, lower IP connection setup time and higher control message efficiency. The graphs of Simple6 and NSFNet show similar trends in each metric.

Figures 4.19 ~ 4.20 show the IBBP and Figures 4.21 ~ 4.22 show the NRU. Under the same strategy, the difference between DTG-B and DTG-PRL is only due to the pre-reserved lightpaths. The performance improvement due to pre-reserved lightpaths is obvious. For instance, when the IBBP of DTG-PRL $U=1, 5, 10$ is around 10%, the IBBP of DTG-B $U=1, 5, 10$ is up to 20% ~ 30% in Simple6 and 20% ~ 25% in NSFNet. Besides, in NSFNet for example, the highest NRU of DTG-B $U=1, 5, 10$ are around 5%, 11%, 15% at 15, 30 and 40 Erlang respectively,

whereas the highest NRU of DTG-PRL $U=1, 5, 10$ are around 7%, 13%, 17% at 20, 40 and 50 Erlang respectively. This illustrates that the gain from DTG-PRL is up to 40%, 20% and 13% for strategies of $U=1, 5, 10$ in comparison with the DTG-B in NSFNet.

Figures 4.23 ~ 4.24 show the ICST and Figures 4.25 ~ 4.26 show the CME respectively. As the load increases, the ICST of DTG-B decreases first and increases later, and the CME of DTG-B increases first and decreases later. Nevertheless, the ICST and CME of DTG-PRL maintain consistently increasing and decreasing trend respectively. This can be explained as follows. In DTG-PRL, some carefully selected lightpaths are pre-reserved and will not be released, whereas in DTG-B, the lightpaths are established or released as soon as they are requested or freed. At low load, lightpaths vary frequently in DTG-B, which increases the ICST. As the load increases, established lightpaths are not easily released because they are likely to be occupied for a longer period. The lightpath variation decreases, as well as the ICST. As the load increases further, the ICST increases again due to the processing delay. For DTG-PRL, ICST is barely influenced by the lightpath variation, so that it will increase consistently. For CME, the reason is similar. High lightpath variation induces more control message overhead, which lowers the control message efficiency. However, as load increases, the effective throughput decreases, so does the CME. It is also clear that the difference of ICST between DTG-B and DTG-PRL is bigger in NSFNet. This may be concluded that in a more scalable network, the average time saving by Pre-Reserved Lightpaths during the IP connection setup could be more.

In addition, it is observed that IP layer information updating threshold has a distinct impact on network performance. For either DTG-B or DTG-PRL, the

scenario without signaling has the best performance because of its unrealistic assumption that the time taken for information updating and signaling is negligible. Among the scenarios with signaling, the $U=10$ strategy has the best performance and $U=1$ strategy has the worst performance. The reason can be explained as follows. Generally, a lower updating threshold means a more frequent update, thus the virtual link (lightpath) information in the IP layer is expected to be up-to-date. On the contrary, the frequent updating of virtual link information brings a higher control message overhead, which tends to congest the signaling procedure of connection establishment, and eventually results in high blocking. The appropriate setting of the updating threshold is a complex issue. A good strategy is needed here.

4.5.3.3 *Comparison between DTG-B and DTG-PRL with multicast IP traffic*

Figures 4.27 ~ 4.34 show the performances of DTG-B and DTG-PRL for the Simple6 and NSFNet networks for multicast IP traffic. It can be seen that DTG-PRL outperforms DTG-B in all the performance metrics.

Without the signaling effect, the results of DTG-B and DTG-PRL for multicast IP traffic show a higher MBBP and lower MNRU compared to the results for unicast IP traffic model. This can be seen by comparing Figures 4.19 ~ 4.22 with Figures 4.27 ~ 4.30. The reason is that both DTG-B and DTG-PRL need multiple IP connection accommodation procedures for one multicast IP request. The increased complexity of accessing multicast IP traffic increases the blocking probability of the multicast IP requests. Moreover, the situation becomes even worse when the signaling effects are taken into account, which can be seen in Figures 4.27 ~ 4.30. The MBBP has a steep increase when the load is around 15 Erlang for Simple6 network and 1.5 ~ 2.0 Erlang for NSFNet network, where the load is still

considered low for unicast IP traffic. The peak of MNRU is also much lower compared with those of unicast IP traffic model.

Figures 4.31 ~ 4.32 show the MCST of Simple6 and NSFNet networks. Although DTG-PRL still has a lower connection setup time than DTG-B, the differences between them shrink compared with those of unicast IP traffic model. This can be explained as follows. In the DTG-PRL for multicast IP traffic model, lightpaths but not light-trees will be pre-reserved. It means that DTG-PRL also needs multiple resource reservation procedures to accommodate multicast IP requests for individual destinations, which is similar to DTG-B. For DTG-B, when the virtual topology has become stable, the total time for multicast connection setup will have little difference compared with that of DTG-PRL.

Figure 4.27 Multicast performance comparison on Simple6: MBBP vs. TL

Figure 4.28 Multicast performance comparison on NSFNet: MBBP vs. TL

Figure 4.29 Multicast performance comparison on Simple6: MNRU vs. TL

Figure 4.30 Multicast performance comparison on NSFNet: MNRU vs. TL

Figure 4.31 Multicast performance comparison on Simple6: MCST vs. TL

Figure 4.32 Multicast performance comparison on NSFNet: MCST vs. TL

Figure 4.33 Multicast performance comparison on Simple6: MCME vs. TL

Figure 4.34 Multicast performance comparison on NSFNet: MCME vs. TL

4.6 Summary

Previous research work on dynamic traffic grooming did not consider the effect of signaling and related information update. The performance is poor when we incorporate the signaling in previous DTG algorithms. This shows that the effect of signaling cannot be ignored in practical networks. In this chapter, we discuss the effect of signaling on DTG and propose a new technique called DTG-PRL to solve the DTG problem in practical network. DTG-PRL consists of two steps, 1) Pre-reserve some lightpaths based on statistical traffic observations; 2) Dynamically groom the traffic based on the established virtual topology. For both unicast and multicast IP traffic model, we have developed the ILP formulations and the MWU/MWU-M heuristic algorithms for the purpose. It has been shown that the results of MWU/MWU-W have a good match with those of ILP. This establishes

the reliability of the MWU/MWU-W algorithm when applied to scalable networks. Subsequently, we carry out the comparisons between DTG-PRL and DTG-B. The results show that the DTG-PRL outperforms DTG-B when comparing the metrics of *IP bandwidth blocking probability* (IBBP), *network resource utilization* (NRU), *IP connection setup time* (ICST), and *control message efficiency* (CME). Information updating strategy is identified to be one of the key factors affecting the performance of dynamic traffic grooming. This deserves further study in future.

Chapter 5

Generic Graph Model for Dynamic Optical Multicast RWA

We have discussed the DTG with multicast IP traffic load in Chapter 4, and found that an IP multicast request that has more than one destinations has to pass through more than one lightpaths (virtual links). Between two joint lightpaths, traffic needs to go through OEO conversion at the switch node where they join. OEO may limit the rate of traffic flow and introduce more processing delay. Unfortunately, multicasting on the IP layer will result in frequent OEO conversion, which is not efficient.

A strong mandate has been provided to develop multicast capability in optical WDM network in order to support bandwidth-intensive multimedia applications such as high-definition television (HDTV), video-conferencing, video-on-demand (VOD), and so on [18, 50]. Further more, optical multicast also offer the following advantages [38]:

- Protection: efficient optical channel and line dedicated protection;
- Performance: improved performance (no store and forward) compared to packet-oriented multicast technology;
- Cost Reduction: reduction of the total number of transceivers deployed in optical networks;
- Resource utilization: overall network throughput improvement by reducing the number of wavelengths used per fiber link (i.e. minimizing the overall

bandwidth usage per fiber link).

Hence, multicasting in the optical layer attracts more research interests recently.

In order to transfer multicast data in WDM network, an optical multicast tree is normally required to be set up at the optical layer. Such tree-structured optical connections are referred to as light-trees [17]. Hardware and software implementations are required to realize light-tree-based optical layer multicasting.

Switch nodes in the network need to possess light splitting and wavelength conversion capabilities to enable a wavelength from any input port to be switched/routed to multiple output ports at the same or a different wavelength. The light splitting operation can be realized with a splitter-and-delivery (SaD) switch [100] or a configurable light splitter [101]. Besides light splitting, wavelength conversion is another important operation performed by wavelength converters which are able to convert an input wavelength into a different output wavelength, via OEO conversion or in an all-optical manner [102].

Various node architectures and network structures can be constructed and supported for different light splitting and wavelength conversion capabilities. Basically, the various capabilities can be classified into full, sparse, and limited cases. The term “full” means that light splitters or wavelength converters exist in every network node and there is no limitation in the wavelength that can convert or the number of outputs and input signal can be split. In general, such cases are easy to analyze but not practical nor cost-efficient as large number of components are required, and may also impose fabrication complexity. To overcome this, both network level (sparse cases) and node level (limited cases) solutions are preferred. The term “sparse” refers to only a subset of the nodes in the network have light splitting or wavelength conversion capabilities [48]. Note that wavelength

converters and light splitters need not co-locate within the same network node. For the limited cases, a node's capability of light splitting or wavelength conversion is limited by some constraints. For limited light splitting [103], the input light signal can only be split into a limited number of output ports due to power loss constraint; For limited wavelength conversion [47], an input wavelength can only be converted into certain wavelengths but not all wavelengths due to the constraints of limited wavelength converters.

Besides all the above hardware related issues, efficient control protocols and algorithms are needed to support optical layer multicasting. To establish a light-tree, one needs to find a route from the source node to all the destination nodes and then assign an available wavelength to each link along the route. This process is called multicast routing and wavelength assignment (MC-RWA). The MC-RWA problem has been studied for various node architectures and network structures. However, owing to the variety of hardware considerations, different combinations of light splitting and wavelength conversion make MC-RWA a difficult problem to solve. There are many studies [104-105] on MC-RWA problem for various light splitting scenarios without considering wavelength conversions. Similarly, there are also many studies on various wavelength conversion scenarios without light splitting being considered [52, 106-107]. Reference [19] is a rare article that considers both sparse light splitting and sparse wavelength conversion jointly but only for static MC-RWA, in which a set of requests are known beforehand and these requests need to be met with the objective of maximizing network throughput or minimizing network resource usage.

It is unlikely for real-life networks to have homogeneous configurations for all their nodes like those studies mentioned above. More likely a network will consist

of network equipment supplied by various vendors [32]. Replacement and upgrading will see new equipment operating besides old ones. As such, it may be desirable to study the problem based on the assumption that network nodes are heterogeneous in nature. Therefore, in order to efficiently solve the MC-RWA problem for heterogeneous networks, the model for studying the MC-RWA problem should be: 1) generic so that all combinations of wavelength conversion and light splitting capabilities in different network nodes can be covered; 2) simple enough to meet with on-line computing demand for dynamic traffic; 3) flexible and extendable for potential network upgrade.

In this chapter, a generic graph model is proposed based on layered auxiliary graph for MC-RWA problem in relation to wavelength-routed WDM networks. Based on the graph model, an efficient shortest path tree (SPT) algorithm is employed to solve the dynamic MC-RWA problem for various light splitting and wavelength conversion scenarios. Furthermore, the proposed model can be used as a common platform to study and evaluate other related problems and algorithms, such as splitter and converter placement, multicast traffic grooming and so on.

The rest of the chapter is organized as follows. In Section 5.1, the proposed graph model is introduced. In Section 5.2, the creation of wavelength conversion edges is discussed. In Section 5.3, the creation of light splitting edges is discussed. The MC-RWA algorithm is derived and presented in Section 5.4. In Section 5.5, numerical results are obtained and presented. Finally, we summarize this chapter in Section 5.6.

5.1 Definitions and construction of Generic Graph Model

A WDM network can be modeled as a directed graph $G(V_N, E_L, \Lambda_W)$, where $V_N = \{v_0, v_1, \dots, v_{N-1}\}$ denotes the set of N switch nodes, $E_L = \{e_0, e_1, \dots, e_{L-1}\}$ denotes the set of L directed wavelength links, and $\Lambda_W = \{\lambda_0, \lambda_1, \dots, \lambda_{W-1}\}$ denotes the set of W wavelengths supported in each wavelength link. For each node $v \in V_N$, the degree of node v is denoted as d_v . In this thesis, the degree of a node is defined as the number of nodes that it is connected to via direct links. Note that each directly connected node pair is assumed to be connected by two directional wavelength links in opposite directions. Hence, the total number of directed wavelength links L in G is given by:

$$L = \sum_{v \in V_N} d_v, \text{ where } 1 \leq d_v \leq N-1 \quad (5.1)$$

From the given directed graph $G(V_N, E_L, \Lambda_W)$, a generic graph $G'(V_{N'}, E_{L'})$ can be generated, where $V_{N'}$ denotes the set of N' vertices and $E_{L'}$ the set of L' directed edges in the generic graph G' . To avoid confusion, the terms ‘node’ and ‘link’ are used for the physical network G while the terms ‘vertex’ and ‘edge’ are used for the derived generic graph G' . G' consists of W layers and each layer is mapped to the network state of one of the W supported wavelengths. On each layer of G' , each node $v \in V_N$ is mapped to d_v input vertices and d_v output vertices. Each of the input (output) vertex represents the port from which the traffic enters(leaves) the node. The total number of the vertices in G' is given by:

$$N' = \left(2 \sum_{v \in V_N} d_v \right) \cdot W = 2 \cdot L \cdot W \quad (5.2)$$

On each layer of G' , each directed wavelength link $e \in E_L$ is mapped to one

unique wavelength edge. A wavelength edge is a directed edge originating from an output vertex to an input vertex between two different switching nodes on the same wavelength layer. Assume that for an arbitrary switching node $v \in V_N$, the traffic from any input port of v can be switched to any output port of v by an OXC. Node v will be represented as d_v input vertices and d_v output vertices with $d_v \cdot d_v$ optical switching edges interconnecting these input and output vertices in a full mesh structure on each layer of G' . Here, an optical switching edge refers to a directed edge originating from an input vertex to an output vertex inside the same switching node on the same wavelength layer. The total number of the edges in G' is given by:

$$L' = \left(\sum_{v \in V_N} (d_v)^2 + L \right) \cdot W \quad (5.3)$$

So far, the above discussion does not consider wavelength converters and light splitters. If a switching node has the capability of wavelength conversion or light splitting, wavelength conversion or light splitting edges will also have to be created in G' . Here, a wavelength conversion edge refers to a directed edge originating from an input/output vertex in a switching node in one wavelength layer to another input/output vertex in the same switching node in a different wavelength layer. A light splitting edge is a directed edge originating from an input vertex to an output vertex inside the same switching node on the same wavelength layer. A light splitting edge is very similar to an optical switching edge; the minor difference being that only one switching edge originating from one input vertex can be occupied but multiple splitting edges originating from one input vertex may be occupied at the same time. Based on the type or limitation of the wavelength converter and light splitter, the wavelength conversion edges and light splitting

edges are created accordingly. We will discuss the construction of wavelength edges and splitting edges in details later.

Figure 5.1 (a) Physical topology of a given network $G(V_4, E_6, \Lambda_2)$;
(b) Derived generic graph $G'(V'_{24}, E'_{36})$

Figure 5.1 shows a simple example based on a given network $G(V_4, E_6, \Lambda_2)$, where $N = 4$, $L = 6$ and $W = 2$. Two wavelength layers, λ_0 and λ_1 , are mapped onto the generic graph G' . Suppose node **B** has full wavelength conversion and full light splitting capability. Since $d_B = 3$, 6 directed (or 3 bi-directed, as shown in Figure 1(b)) wavelength conversion edges are created to connect the corresponding output vertices of **B** between layer λ_0 and λ_1 . On each

wavelength layer, any input vertex of $\mathbf{B}$ will have 3 directed light splitting edges connecting to the 3 output vertices respectively. An optical multicast request Req in a WDM network can be denoted as

$$Req = \{p, Q\}, \text{ where } p \in V_N, Q = \{q_i \mid q_i \in V_N, q_i \neq p, 0 \leq i \leq m-1\} \quad (5.4)$$

Here p is the source node, Q is a set of destination nodes, and m is the number of destinations of Req . Denote the multicast group size of Req as z_{Req} ; z_{Req} is given by:

$$z_{Req} = 1 + m \quad (5.5)$$

Following the example in Figure 5.1, if a multicast request $Req = \{A, \{C, D\}\}$ is given, a multicast tree will be set up to satisfy the request Req . Assume that the chosen multicast tree is as shown in Figure 5.2(a). Figure 5.2(b) illustrates the setting up of the multicast tree in the generic graph. The tree starts from the output vertex of source node A on wavelength layer λ_0 . After passing a wavelength edge directed to B , it arrives at the input vertex of B where the signal is split via two splitter edges into two copies -- one goes to the output vertex connecting node C while the other goes to the output vertex connecting node D . The copy that goes to node C arrives at its destination by directly traversing the wavelength edge from B to C on layer λ_0 . The other copy that goes to node D reaches its destination through a wavelength conversion edge from layer λ_0 to layer λ_1 and the wavelength edge from B to D on layer λ_1 .

Though this example only illustrates full wavelength conversion and full light splitting in node B , it can be easily seen that the proposed generic graph model can model any combination of wavelength converters and light splitters in a node.

Figure 5.2 (a) The multicast tree to be set up for the request $Req = \{A, \{C, D\}\}$ in G ;
 (b) The multicast tree to be set up for the request $Req = \{A, \{C, D\}\}$ in G' .

5.2 Creation of wavelength conversion edges

Strategies involving the use of wavelength converters in a WDM network can be classified into the following categories: Sparse wavelength conversion and limited wavelength conversion. These two strategies are not mutually exclusive and can be implemented in a hybrid manner. However, due to the complexity of implementing a hybrid strategy, there is virtually no study on it. By converting the physical topology graph into a generic graph model, it becomes straightforward to study various hybrid strategies involving both sparse and limited wavelength

conversions.

5.2.1 Sparse wavelength conversion

If only a subset of the switching nodes in a network have the capability of wavelength conversion while other nodes do not have, this is known as sparse wavelength conversion. The rule for incorporating sparse wavelength conversion into the generic graph model is simple: Create wavelength conversion edges among the wavelength layers for the node that has the capability of wavelength conversion.

The wavelength conversion edges can be created on input vertices or output vertices depending on how it is implemented within the switching node. The difference is illustrated in Figure 5.3.

Figure 5.3 Mode of (a) Convert-Switch; (b) Switch-Convert; (c) Convert-Split; (d) Split-Convert.

Figure 5.4 (a) The light tree found for request $r = \{A, \{C, D\}\}$ in G ; (b) Light tree setup in G' , Convert-Split mode, only 1 wavelength conversion edge is involved; (c) Light tree setup in G' , Split-Convert mode, 2 wavelength conversion edges are involved.

Figure 5.3(a) and 5.3(b) show the Convert-Switch and Switch-Convert modes respectively for a node without light splitting capability while Figure 5.3(c) and 5.3(d) show the Convert-Split and Split-Convert modes respectively for a node with light splitting capability. For the case of a switching node in which no light splitter is used, two wavelength conversion modules are needed for both the Convert-Switch mode and the Switch-Convert mode regardless of whether the wavelength conversion modules are placed at the inputs or the outputs of the switch. However for the case of a switching node with light splitting capability, there is a difference between placing wavelength conversion modules at the inputs or the outputs of the switching node. For the Convert-Split mode, if all the output ports always use the same wavelength, only one wavelength conversion module is needed at the input. However, for the Split-Convert mode, two individual wavelength conversion modules are needed for the two output ports.

Based on the example described in Section 5.1.1, Figure 5.4 illustrates the difference between the models for Convert-Split mode and Split-Convert mode. For a multicast tree request $Req = \{A, \{C, D\}\}$, if the light-tree is found as shown in Figure 5.4(a), the light-tree setup in the generic graph model will depend on where wavelength converters are placed in node **B**. Compared with the Split-Convert mode in Figure 5.4(c), the Convert-Split mode in Figure 5.4(b) may save one wavelength conversion edge, and hence the routing cost. The cost issue will be discussed in Section 5.1.4.

5.2.2 Limited wavelength conversion

When an input wavelength can only be converted into certain wavelength but not the full range of wavelength within a switching node, this is known as limited wavelength conversion. The threshold model proposed in [47] is used in this paper. Let g_c denote the conversion degree of a limited wavelength converter. In the threshold model, a converter with conversion degree g_c can convert an input wavelength to g_c adjacent wavelengths on either side of the input wavelength in addition to the input wavelength itself. Therefore, for a given number of supported wavelengths W , we have:

$$0 \leq g_c \leq \lfloor W/2 \rfloor \quad (5.6)$$

Note that $g_c = 0$ is equivalent to no wavelength conversion, and $g_c = \lfloor W/2 \rfloor$ implies full wavelength conversion.

In order to establish the generic graph model, all the possible wavelength conversion edges need to be created. Note that only one conversion edge can be occupied for each wavelength at any time instance. From the wavelength layer λ_i , $2g_c$ wavelength conversion edge will be created and linked to each layer from $\lambda_{(i-g_c) \pmod W}$ to $\lambda_{(i+g_c) \pmod W}$. Figure 5.5 illustrates an example for layer λ_i and its adjacent layers in Switch/Splitter-Convert mode. Combined with Figure 5.1(a), the total number of wavelength conversion edges $L'_{c,v}$ in node v is given by:

$$L'_{c,v} = W \cdot 2g_c \cdot d_v \quad (5.7)$$

Figure 5.5 Create wavelength conversion edges in generic graph model

5.2.3 Hybrid wavelength conversion

If both sparse and limiting wavelength conversions are used in a WDM network, this is known as hybrid wavelength conversion. Assume that the conversion degree for node v is denoted as $g_{c,v}$. Note that for a node without wavelength conversion capability, $g_{c,v} = 0$. The total number of the wavelength conversion edges L'_c in the generic graph model G' with hybrid wavelength conversion is given by:

$$L'_c = \sum_{v \in V_N} L'_{c,v} = 2W \cdot \sum_{v \in V_N} (g_{c,v} \cdot d_v), \text{ where } 1 \leq d_v \leq N-1, 0 \leq g_{c,v} \leq \lfloor W/2 \rfloor \quad (5.8)$$

Suppose all the nodes have the same $g_{c,v} = g_c$, which is a case of homogeneous limiting wavelength conversion network, then L'_c can be simplified to:

$$L'_c = 2Wg_c \cdot \sum_{v \in V_N} d_v = 2WL \cdot g_c = N' \cdot g_c \quad (5.9)$$

Suppose all the nodes are capable of full wavelength conversion without any limitation, then L'_c can be simplified further to:

$$L'_c = 2WL \cdot \lfloor W/2 \rfloor = W(W-1)L \quad (5.10)$$

5.3 Creation of light splitting edges

The creation of light splitting edges also can be classified into sparse, limited and hybrid. For a sparse splitter network, light splitting edges need to be created for each node that has light splitters when we model it using the generic graph model. Let $g_{s,v}$ denote the splitter degree of node v in a limited splitter network. Note that $g_{s,v}$ is in fact the number of output ports that a splitter in node v can support; thus,

$$1 \leq g_{s,v} \leq d_v \quad (5.11)$$

Figure 5.6 Create light splitting edges in generic graph model.

It is assumed that, by default, each node has a non-blocking OXC that is able to realize all the input-output cross-connection combinations without blocking. From the functionality point of view, light splitting has a similar function, that is, to transfer data from any input port to any output port. The subtle difference is that, for switching edges, all the edges originating from the same input vertex are shared whereas for splitting edges, they are dedicated. In the hybrid splitter network, for a node without light splitting capability, $g_{s,v} = 1$; and for a node with full light

splitting capability, $g_{s,v} = d_v$.

Figure 5.6 shows an example of how light splitting edges are created using the generic graph model. Assume that for an arbitrary node $v \notin \{v_0, v_1, v_2, v_3\}$, the node degree is $d_v = 4$. Figure 5.6(a) shows all the possible edges used by the switch or light splitter. Consider a traffic flow from node v_1 to v_2 arriving at the input vertex of v . If node v has no splitter, i.e. $g_{s,v} = 1$, the traffic will pass through the edge $v_1 \rightarrow v_2$, and all other three edges ($v_1 \rightarrow v_0$, $v_1 \rightarrow v_1$, $v_1 \rightarrow v_3$) will become invalid for future routing, as shown in Figure 5.6(b). On the other hand, if $g_{s,v} = 2$, then all the other three edges will still be valid for future routing, as shown in Figure 5.6(c), until the maximum number of supported ports is reached. The last case in which the maximum number of supported ports is reached is shown in Figure 5.6(d), in which two splitting edges ($v_1 \rightarrow v_2$, $v_1 \rightarrow v_3$) are occupied while the other two edges ($v_1 \rightarrow v_0$, $v_1 \rightarrow v_1$) become invalid.

Since the switching edges and splitting edges have the same possible connections, in order not to make the graph too complex, we do not distinguish between these two types of edges but make use of a two-dimension state to show their states, as shown in Table 5.1.

Table 5.1 States to manage switching / splitting edges

For routing Edge status	Valid	Invalid
Busy	-	<i>State I</i>
Idle	<i>State II</i>	<i>State III</i>

A currently busy edge, i.e. one that carries a traffic flow, is always invalid for routing and is deemed to be in **State I**. A currently idle edge will be valid for routing (**State II**) if the number of other busy splitting edges that have the same input vertex is less than $g_{s,v}$; otherwise it is invalid for routing (**State III**).

The total number of possible light splitting edges in the generic graph model G' is the same as the number of possible switching edges, which is given by:

$$L'_s = W \cdot \sum_{v \in V'_N} (d_v)^2 \quad (5.12)$$

Compared with existing graph models for traffic grooming problem [32, 108], the proposed graph model makes two important contributions. Firstly, this model is a more general one, which directly considers the MC-RWA problem. Generally, the solutions to routing and wavelength assignment problem form the basis for solving many other related problems, such as traffic grooming, virtual topology design and network survivability. Secondly and more importantly, the proposed model is the first to consider the heterogeneity in network nodes with different light splitting and wavelength conversion capabilities, which are the key components for the optical layer multicast.

5.4 MC-RWA based on the Generic Graph Model

5.4.1 Cost model

By assigning a cost to each edge in the generic graph $G'(V'_N, E'_L)$, the RWA problem can be converted to a classic Steiner tree problem [51], which is well known to be *NP*-complete. The SPT algorithm that has been proposed in [51] can be used to solve the multicast RWA problem. A brief introduction of the SPT algorithm can be found in section 4.1.1.2. The algorithm has the complexity of $O(m \cdot N^2)$

and an approximate upper bound of $O(2 - 2/m)$ compared to the optimal solution [51], where N is the number of vertices in the graph and m is the number of the destinations. Due to its simplicity and efficiency, we apply the SPT algorithm in the Generic Graph Model to meet the real time computing demand for dynamic traffic.

Let the cost of wavelength link edge $e_{v_i v_j}$ in G' be denoted as $C_w(e_{v_i v_j})$, where $C_w(e_{v_i v_j})$ is a function of the wavelength link distance between two nodes v_i and v_j connected by $e_{v_i v_j}$. Assume that the cost of a wavelength link in a unit distance is fixed at α , and the distance between v_i and v_j is D_{ij} , we have:

$$C_w(e_{v_i v_j}) = D_{ij} \cdot \alpha = D_{ji} \cdot \alpha, \quad v_i, v_j \in V_N \quad (5.13)$$

Suppose all the wavelength conversion edges, light splitting edges, and optical switching edges have the same cost of C_c , C_{sp} , and C_{sw} respectively, and suppose the multicast tree found via the SPT algorithm for $Req = \{p, Q\}$ is denoted as $T_r = T(V_r, E_r)$, the cost of T_r is given by:

$$C(T_r) = \sum_{e_w \in E_r} C_w(e_w) + \sum_{e_c \in E_r} C_c + \sum_{e_{sp} \in E_r} C_{sp} + \sum_{e_{sw} \in E_r} C_{sw} \quad (5.14)$$

Here, e_w , e_c , e_{sp} and e_{sw} denote wavelength link edge, wavelength conversion edge, light splitting edge and optical switching edge occupied in T_r respectively. Thus, given a multicast request $Req = \{p, Q\}$, the task is to find a multicast tree T_r such that $C(T_r)$ is minimum.

In a real-world network, C_c , C_{sp} , and C_{sw} are hard to be determined accurately. Somani A. K. *et al.* proposed a parametric cost model in [109] to compare the cost of a multi-fiber network to a single-fiber wavelength-convertible network. Without loss of generality, by following the concept of the parametric cost

model, we propose the following approximation to compare the cost of wavelength link to the cost of wavelength conversion edge and light splitting edge [53, 109]. Let $C_c = \kappa_c \cdot \alpha$, $C_{sp} = \kappa_{sp} \cdot \alpha$, and $C_{sw} = \kappa_{sw} \cdot \alpha$, equation (5.14) can be changed to:

$$C(T_r) = \alpha \cdot \left(\sum_{e_w \in E_r} D_w + L_{e_c} \cdot \kappa_c + L_{e_{sp}} \cdot \kappa_{sp} + L_{e_{sw}} \cdot \kappa_{sw} \right) \quad (5.15)$$

Here, L_{e_c} , $L_{e_{sp}}$, and $L_{e_{sw}}$ denote the number of occupied wavelength conversion edges, light splitting edges and optical switching edges in T_r . The ratio κ_c , κ_{sp} and κ_{sw} can be approximated using the real hardware cost of wavelength converters, light splitters and OXCs, including costs of components, configuration, and maintenance, etc. From equation (5.15), it can be easily found that different values of κ_c , κ_{sp} and κ_{sw} will impact the cost of T_r as well as the way different edges are chosen. During the execution of the SPT algorithm, longer wavelength link edges and edges with a higher κ value will have a lower possibility of being chosen and vice versa, to minimize the total cost of the selected multicasting tree.

By means of the derived generic graphic model, the cost of wavelength converters, light splitters, OXCs, as well as wavelength links will all be taken into consideration during the execution of the multicast RWA. It makes the study on the effects of wavelength converters and light splitters straightforward and easy to compare.

5.4.2 MC-RWA algorithm

Note that, for each node in the physical network G , there are corresponding input and output vertices on each wavelength layer in the graph G' . We add a pseudo node $\hat{v}$ for each of the switching node $v \in V_N$, and then connect each

pseudo node to its corresponding input and output vertices on each wavelength layer by zero-weight edges. In this way, the MC-RWA problem is reduced to finding a shortest path tree in the graph, which is rooted at pseudo source node $\hat{v}_s$ and spanning each of destination nodes $\hat{v}_d$. Therefore, for a new multicast request $Req = \{s, D\}$ in the generic graph $G'(V'_N, E'_L)$, where s is the source, D is the set of destination nodes, the details of MC-RWA algorithm based on SPT algorithm can be described in Table 5.2.

Table 5.2 MC-RWA algorithm

-
1. Initially, the route tree for the request includes only the pseudo source node $\hat{v}_s$. For each pseudo destination $\hat{v}_d$ of D , repeat *Step 2* to *Step 3*.
 2. Search the shortest paths on the layered auxiliary graph from the partially established route tree to all remaining pseudo destination nodes. Choose the destination which is nearest to the partially established route tree. If no such destination exists, go to *Step 4*; otherwise, continue to *Step 3*.
 3. Iterate through vertices and edges on the shortest path to the chosen destination to augment the route tree. Set all the traverse wavelength link edges along the path to be busy. Set the traverse wavelength conversion edges and light splitting edges to states in Table 5.1 according to the sparse or limit strategies. Then go to *Step 5*.
 4. Block such request (for request blocking policy) or block such destination (for destination blocking case) and go to *Step 5*.
 5. End of algorithm.
-

In general, the complexity of the above MC-RWA algorithm results mainly from the SPT algorithm. In the established generic graph $G'(V'_N, E'_L)$, the running complexity of the SPT algorithm is $O(m \cdot (N')^2)$, where N' is the number of vertices in the graph and m is the number of destinations [51]. From equations (5.1) and (5.2), the complexity of the SPT algorithm, can be derived to:

$$O\left(m \cdot (N')^2\right) = O\left(m \cdot 4 \cdot W^2 \cdot \left(\sum_{v \in V_N} d_v\right)^2\right) = O\left(4m \cdot (N \cdot W \cdot d_{avg})^2\right),$$

$$\text{where } d_{avg} = \frac{1}{N} \sum_{v \in V_N} d_v \quad (5.16)$$

5.5 Numerical results

In this section, the proposed MC-RWA algorithm based on the graph model is evaluated using European COST239 and NSFNet network as shown in Figure A.4 and Figure A.5. We assume the number of wavelengths in each fiber to be $W = 8$. Assume that the multicast group size has a default value of $z = 6$, the multicast request is generated randomly by choosing z nodes from all the network nodes, following a uniform distribution. One node is assigned as the source node and the rest destination nodes. The arrival multicast requests follow a Poisson distribution with the mean arrival rate of λ . For each admitted multicast request, the holding time follows an exponential distribution with a mean holding time of μ . Without loss of generality and for the sake of simplicity, assume that the weights of wavelength link edge and light splitting edge are all one unit and each wavelength-conversion edge has 50 units. This is because the cost of wavelength conversion is much higher than the cost of wavelength as well as light splitter [109]. For each data point in the figure showing simulation results, 100,000 multicast requests are evaluated. For each multicast request, if any destination cannot be reached, the request will be blocked. The evaluation metric used is termed *request blocking probability* (RBP).

Figure 5.7 RBP for sparse splitting cases in different group sizes in NSFNet

First, we focus on light splitting related issues by assuming that the network does not support wavelength conversion. Figure 5.7 shows the results for the sparse splitting case by varying the number of light splitting nodes in the NSFNet network. $L = 0$ is the case in which no node in the network has the capability of light splitting; $L = 14$ is the case of full light splitting, and $L = 7$ is the case in which half of the nodes in the network has light splitting capability. Nodes with light splitting capability are randomly assigned. Once assigned, these nodes are fixed throughout the simulation. In general, when the number of light splitting capable nodes increases, the RBP decreases. It can be seen that the blocking performance for the case of $L = 7$ approaches that of full light splitting ($L = 14$). This suggests that it is not necessary for the network to be equipped with full light splitting capability as it does not improve the blocking performance significantly compared

to if 50% of the nodes have light splitting capability. Using fewer light splitting capable nodes will result in significant cost saving. By and large, the trends of the RBP graphs for different group sizes are similar. The larger the multicast group size is, the higher the RBP.

Figure 5.8 RBP for sparse light splitting cases in different network topologies

In Figure 5.8, the performances of the 3 different light splitting cases for NSFNet and COST239 networks are compared. Here, “No” means no node in the network has the capability of light splitting; “Full” refers to the case of full light splitting, and “Half” means that half of the nodes in the network possess light splitting capability. In the figure, the RBPs for COST239 are much smaller than those for NSFNet. The reason is that the average node degree for COST239 is larger than that of NSFNet – 2.36 ($= 26/11$) for COST239 and 1.5 ($= 21/14$) for NSFNet – which makes the incoming requests in COST239 to have a higher chance to be

routed from one node to the others.

Figure 5.9 RBP for limited light splitting cases in different network topologies

Figure 5.9 shows the blocking performance for the limited light splitting case, i.e., all the nodes are assumed to have splitting capabilities but such capabilities are limited by the splitting degree, g_s . When g_s increases from 1 to 3, the RBP decreases. The performance corresponds to $g_s = 2$ is close to that of $g_s = 3$. This indicates that by using limited-splitting-degree-capable splitters in the network, good blocking performance can still be obtained. The differences between $g_s = 2$ and $g_s = 3$ in COST239 network are larger than the corresponding differences in NSFNet. This is because the average node degree of NSFNet is 1.5, which means that the splitter with $g_s = 2$ can satisfy most of the splitting operation requirement. Thus, it can be concluded that when the splitting degree is close to the average node

degree of the network, good blocking performances can still be obtained by using limited light splitters. For limited light splitters with the same splitting degree, the RBP for limited splitting case will be closer to that of full splitting case for networks with a smaller average node degree.

Figure 5.10 RBP for sparse conversion cases in different network topologies

In Figures 5.10 ~ 5.12, the blocking performances are evaluated for various wavelength conversion cases. The network is assumed to have full light splitting capability for fair comparisons. Figure 5.10 shows the RBP for the sparse wavelength conversion case by varying the number of nodes with wavelength conversion capability. In the figure, “No” means no node in the network has the capability of wavelength conversion; “Full” refers to the case of full wavelength conversion, and “Half” means about half of the nodes in the network possess wavelength conversion capability. Nodes with wavelength conversion capability are

randomly assigned and then fixed throughout the simulation. In general, for different topologies, the more nodes with wavelength conversion capabilities, the lower the RBP will be.

Figure 5.11 RBP for limited conversion cases in NSFNet

Figure 5.11 shows the RBP for the limited wavelength conversion case by varying the conversion degree, g_c . $g_c = 1$ means that the converter can convert an input wavelength to one adjacent wavelength on either side of the input wavelength in addition to the input wavelength itself. By using limited wavelength conversion, reasonably good blocking performance can still be achieved compared with the full conversion case. Finally, Figure 5.12 shows the RBP for converters under Split-Convert and Convert-Split modes. The network is assumed to have full light splitting capability and half of the network nodes are assumed to have wavelength conversion capability. It is important to note that different converter construction

approaches may lead to different blocking performances. In general, the Convert-Split mode requires fewer wavelength converters than the Split-Convert mode. The Convert-Split mode trades off blocking performance with fewer wavelength converters used.

Figure 5.12 RBP for Split-Convert and Convert-Split mode in NSFNet

5.6 Summary

In this chapter, a new graph model has been developed for studying dynamic MC-RWA in wavelength-routed WDM networks. Based on the graph model, an efficient shortest path tree algorithm is employed to solve the dynamic MC-RWA for different light splitting and wavelength conversion scenarios. This provides a way to handle heterogeneity in network node structures. The proposed generic graph model enables the dynamic MC-RWA problem for sparse/limited splitting and/or sparse/limited conversion to be jointly studied to provide clearer and deeper

insights of optical layer multicasting. Based on this generic graph model, the blocking performance is investigated for Convert-Split mode and Split-Convert mode (Figure 5.12). In general, this graph model provides a useful platform for studying dynamic MC-RWA problem for different light splitting and wavelength conversion scenarios. To the best of the authors' knowledge, this is the first attempt in applying the generic graph model to address the dynamic MC-RWA problem. In the next chapter, the model will be applied to study the placement problem of splitters and converters in the network.

Chapter 6

Placement of Wavelength Converters and Light Splitters using the GGM

As described in Chapter 5, light splitter and wavelength converter are two critical components in realizing a light-tree based optical layer multicasting [17]. However, due to the high cost of wavelength converter and light splitter, strategies for sparse and limited wavelength conversion and splitting have to be considered. Meanwhile, it has been shown in a number of studies that a relatively small number of wavelength converters and light splitters can achieve satisfactory performance [48, 81-82]. Nevertheless, how to properly place the wavelength converters and light splitters in a network is still an ongoing research problem.

After conducting a comprehensive literature review of prior works on placements of wavelength converters and light splitters in Section 2.5, we made the following observations. Firstly, placement problems in relation to wavelength converters and light splitters have always been studied separately. To the best of our knowledge, few prior work has ever considered placements of wavelength converters and light splitters holistically within the same network. Secondly, the traffic model used in most studies of wavelength converter placements is based on the assumption of pure unicast communication, and the impact of multicast connection sessions are completely ignored. Though the existence of multicast connection sessions is considered in the study of light splitter placements, a set of static multicast connection sessions which is known a priori is often assumed.

Thirdly, only the absence or presence of wavelength converters and light splitters in a node is considered but not such details as their capabilities and complexities. For example, assume that two nodes N_1 and N_2 with node degrees of $d_{N_1} = 2$ and $d_{N_2} = 3$ respectively have full wavelength conversion capability. In many previous studies on wavelength converter placement problem, only one wavelength conversion module is counted for each node, which does not give a good indication of the number of wavelength converters needed in each node and the entire network. Since a node with a higher node degree has more input-output ports for connecting to the neighboring nodes, it should contain more wavelength converters for all the input-output ports so as to support full wavelength conversion in this node. Hence the amount of hardware required to implement wavelength conversion will differ from node to node. This observation is also applicable to previous studies on limited wavelength conversion and placement of light splitters. The above observations suggest that a different approach is needed to assess the amount of wavelength conversions needed and the corresponding cost. This motivates us to find better models and algorithms to solve the wavelength converter and light splitter placement problem for dynamic traffic.

In this chapter, by extending our earlier work on the Generic Graph Model [110], we propose four algorithms to find the optimal placements of wavelength converters and light splitters in terms of such performance metrics as *Cost Efficiency* (CE), *Request Blocking Probability* (RBP), and *Average Cost of Multicast Tree* (ACMT). Simulation studies carried out for different topologies will be reported and discussed. Since the generic graph model has been described in Chapter 5 in details, in this chapter, we focus on the algorithms for the placement problem by using the generic graph model.

The rest of this chapter is organized as follows. In Section 6.1, the objective of the placement problem is described. In Section 6.2, the algorithms for the placements of wavelength converters and light splitters are described in details. In Section 6.3, the complexity of the proposed algorithms is analyzed. The adaptation of the algorithm is discussed as well. Numerical results are presented and discussed in Section 6.4, and finally conclusions are drawn in Section 6.5.

6.1 Objective of placement problem

In previous studies, the objective of the placement problem is to find an optimal placement of the given number of wavelength converters and light splitters such that the network performance is optimized. However, most of the studies make the assumptions that network nodes with wavelength converters or light splitters have full wavelength conversion or full light splitting capability. This assumption completely ignores the cost of wavelength conversion and light splitting for multicast RWA. Also, as most of these studies do not model the node structure and use of wavelength converters and splitters within the node, it is hard to compare the exact usage of wavelength conversion and light splitting links. By means of the proposed generic graph model, we reinvestigate the placement problem by taking into consideration the usages and costs of wavelength converters and light splitters.

The placement problem in our study can be stated as: *For the given dynamic optical multicast traffic model, find the appropriate number of wavelength converters and light splitting links for each node, such that the cost efficiency is optimal for the entire network.*

Here, cost efficiency is defined as the ratio of request success probability over the average multicast tree cost, i.e.,

$$\eta = (1 - P_b) / C(T) \quad (6.1)$$

where P_b denotes the blocking probability of the request and $C(T)$ denotes the average multicast tree cost. Definition of $C(T)$ can be found from Equations (5.14) and (5.15) in section 5.2.1.

6.2 Algorithms for the placement of wavelength converter and light splitter

6.2.1 Greedy algorithm

The greedy algorithm is an algorithm which tries to go through all the possible combinations to find the optimal solution. The concept is straightforward but such a brute force approach is not practical when the number of possible combinations is enormous. If the wavelength degree and splitter degree can be decided separately for each node, from Equations (5.6) and (5.11) in Chapter 5, it can be found that the possible choices of wavelength converters in node v is $(\lfloor W/2 \rfloor + 1)$, and light splitters in node v is d_v . Here W denotes the number of wavelength per fiber link, d_v denotes the node degree in node v . Hence, the number of all the possible combinations is given by:

$$(\lfloor W/2 \rfloor + 1)^N \cdot \prod_{v_i \in V_N} d_{v_i} \quad (6.2)$$

For a complex network topology with a large N , it is very computationally intensive to find the best solution. For example, for the NSFNet network shown in Figure A.5, the number of possible combinations yielded by (6.2) is 2.3×10^{16} when $W = 8$, making it impossible to implement the greedy algorithm and faster algorithms need to be devised.

Since the placement of wavelength converters and light splitters can be

independent, two possible strategies can be implemented; namely, place wavelength converter first and place light splitter first.

6.2.2 Place wavelength converter first (WCF) strategies

The WCF strategy will first place the wavelength converters in the network, and then place light splitters. We propose and study two algorithms, namely, WCF with Incremental Placement (WCF-I) and WCF with Decremental Placement (WCF-D).

6.2.2.1 WCF-I algorithm

Define $V_C = \{v \mid v \in V_N\}$ to record those nodes that will have wavelength converters placed and $V_S = \{v \mid v \in V_N\}$ for those nodes that will have light splitters placed. Define $\bar{V}_C = \{v \mid v \in V_N, v \notin V_C\}$ and $\bar{V}_S = \{v \mid v \in V_N, v \notin V_S\}$ as the complement set of V_C and V_S , respectively.

The WCF-I algorithm can be described as follows. It contains three parts: Main algorithm (see Table 6.1), Function Place_Device (see Table 6.2) and Function Adjust_Degree (see Table 6.3).

Table 6.1 WCF-I algorithm (Main algorithm)

```
// WCF-I Algorithm (Main Algorithm)

BEGIN
  // Step 0: Initialization, all nodes have no WC and LS
   $V_C = V_S = \emptyset$ ;
  Establish generic graph  $G'$  based on  $V_C$  and  $V_S$ ; // which are  $\emptyset$ 
  Proceed with MC-RWA on  $G'$ , find  $\eta_{\emptyset}$ ;
   $\eta_{MAX} = \eta_{\emptyset}$ ;

  // Step 1: Place wavelength converters with full WC capability
   $[V_C, V_S, \eta_{MAX}] = \text{Place\_Device}(\text{FWC}, V_C, V_S, \eta_{MAX})$ ;

  // Step 2: Place light Splitters with full LS capability
   $[V_C, V_S, \eta_{MAX}] = \text{Place\_Device}(\text{FLS}, V_C, V_S, \eta_{MAX})$ ;

  // Step 3: Adjust Degree of wavelength conversion  $g_C$ 
```

```

[  $V_c$ ,  $g_c$ ,  $\eta_{MAX}$  ] = Adjust_Degree(FWC,  $V_c$ ,  $g_c$ ,  $\eta_{MAX}$ );

// Step 4: Adjust degree of light splitting  $g_s$ 
[  $V_s$ ,  $g_s$ ,  $\eta_{MAX}$  ] = Adjust_Degree(FLS,  $V_s$ ,  $g_s$ ,  $\eta_{MAX}$ );

// Find the highest cost efficiency
Return [  $V_c$ ,  $g_c$ ,  $V_s$ ,  $g_s$ ,  $\eta_{MAX}$  ];
END

```

Table 6.2 Function Place_Device

```

// Function Place_Device
Function Place_Device(INPUT:Device_Type,  $V_c$ ,  $V_s$ ,  $\eta_{MAX}$ ; OUTPUT:  $V_c$ ,  $V_s$ ,
 $\eta_{MAX}$ )
BEGIN
     $V = \emptyset$ ;
    IF Device_Type == FWC THEN,
         $V = V_c$ ;
    ELSEIF Device_Type == FLS THEN,
         $V = V_s$ ;
    END IF

    WHILE  $V \neq V_N$  DO
        FOR i=0 TO N-1,
            IF  $v_i \in \bar{V}$  THEN,
                Add  $v_i$  to  $V$ ;
                Establish generic graph  $G'$  based on updated  $V_c$  and  $V_s$ ;
                Proceed MC-RWA on  $G'$ , find  $\eta_i$ ;
                Remove  $v_i$  from  $V$ ;
            END IF
            i = i + 1;
        END FOR

        m = -1; // m is the index of the node to be added to  $V_c$ 
        FOR i=0 TO N-1,
            IF  $\eta_i > \eta_{MAX}$  THEN,
                 $\eta_{MAX} = \eta_i$ ;
                m = i;
            END IF
        END FOR

        IF m > -1 THEN
            Add  $v_m$  to  $V$ ; // Find higher cost efficiency
        ELSE
            Break WHILE Loop; // Can not find higher cost efficiency
        END IF
    END WHILE

```

```

Return [  $V_c$  ,  $V_s$  ,  $\eta_{MAX}$  ];
END

```

Table 6.3 Function Adjust_Degree

```

// Function Adjust_Degree
Function Adjust_Degree(INPUT:Device_Type,  $V$  ,  $g$  ,  $\eta_{MAX}$  ; OUTPUT:  $V$  ,  $g$  ,
 $\eta_{MAX}$  )
BEGIN
  FOR i=0 TO N-1,
    IF  $v_i \in V$  THEN,
       $g_{MIN} = g_{MAX} = 1$  ;
      IF Device_Type == FWC THEN,
         $g_{MIN} = 1$  ;  $g_{MAX} = \lfloor W/2 \rfloor$  ;
      ELSEIF Device_Type == FLS THEN,
         $g_{MIN} = 2$  ;  $g_{MAX} = d_{v_i}$  ;
      END IF

      FOR  $m = g_{MAX}$  TO  $g_{MIN}$  ,
        Establish generic graph  $G'$  based on updated  $V$  ;
        Proceed MC-RWA on  $G'$  , find  $\eta_i$  ;
        IF  $\eta_i > \eta_{MAX}$  THEN,
           $\eta_{MAX} = \eta_i$  ;  $g_i = m$  ;
        END IF
         $m = m - 1$  ;
      END FOR
    END IF

     $i = i + 1$  ;
  END FOR

  Return [  $V$  ,  $g$  ,  $\eta_{MAX}$  ];
END

```

The basic concept of WCF with incremental placement algorithm is to place wavelength converters or light splitters one by one until the highest cost efficiency is found. The algorithm has five steps. *Step 0* is the initialization step, which serves to find out the cost efficiency with no wavelength converter and light splitter in the whole network. *Step 1* and *Step 2* sparsely place wavelength converters first followed by light splitters, which fulfill the sparse placement strategy. In these

two steps, the wavelength converters and light splitters to be added to the network are assumed to have full capability of wavelength conversion and light splitting. After that, *Step 3* and *Step 4* restrict the degree of wavelength converters and light splitters that are added to the network. These two steps fulfill the limit wavelength conversion and light splitting strategy. After each step, the highest cost efficiency will be recorded and passed to the next step, and V_C , V_S , g_C , g_S will be changed accordingly. Since the underlying algorithms of sparse placement strategies for wavelength converters and light splitters are the same, we combine them into one function **Place_Device**. For the same reason, the algorithms for the limit strategy are also combined into function **Adjust_Degree**. We use **Device_Type** to differentiate a wavelength converter from a light splitter.

Inside function **Place_Device**, the devices will be added to the network in several rounds. In each round, the device will be temporally added to each candidate node contained in $\bar{V}_C$ or $\bar{V}_S$ individually and the cost efficiency is calculated. Only the scenario with the highest cost efficiency will be chosen, i.e. a wavelength converter or light splitter will be placed in the candidate node.

Inside function **Adjust_Degree**, the degree adjustment is only carried out in sequence for those nodes that have been placed with devices. The adjustment range for the degree is defined by $[g_{\min}, g_{\max}]$. Since each node that has been placed with a wavelength converter or a splitter has full capability of wavelength conversion or light splitting respectively to start with, the node degree is initialized to $g_{\max}$, and is decremented by 1 in each round until it reaches $g_{\min}$. $g_{\min}$ can be set to 0, in which case, it indicates that the selected device is not available.

After all the five steps have been executed, V_C , V_S , g_C , g_S will record the

placement information and degree limitation of each node in the network, together with the highest cost efficiency that has been found.

6.2.2.2 WCF-D algorithm

Contrast to the WCF-I algorithm, the WCF-D algorithm finds the highest cost efficiency in a reverse way, i.e. remove wavelength converters and light splitters one by one until the highest cost efficiency is found. WCF-D also contains five steps, the same as those in WCF-I. During the initialization, WCF-D begins with all nodes having full capability of wavelength conversion and light splitting. In *Step 1* and *Step 2* when wavelength converters and light splitters are placed, only one node will be chosen and its device will be removed such that the network with the rest of devices will have the highest cost efficiency.

Compared with WCF-I algorithm, WCF-D algorithm is almost the same except the changes from V_C to $\bar{V}_C$, and V_S to $\bar{V}_S$ in *Step 1* and *Step 2*. The main algorithm of WCF-D is described in Table 6.4.

Table 6.4 WCF-D algorithm (Main algorithm)

```
// WCF-D Algorithm (Main algorithm)
BEGIN
  // Step 0: Initialization, all nodes have full WC and LS
   $\mathbf{V}_C = \mathbf{V}_S = \mathbf{V}_N$ ; // i.e.  $\bar{\mathbf{V}}_C = \bar{\mathbf{V}}_S = \emptyset$ 
  Establish generic graph  $\mathbf{G}'$  based on  $\mathbf{V}_C$  and  $\mathbf{V}_S$ ;
  Proceed MC-RWA on  $\mathbf{G}'$ , find  $\eta_{\text{FULL}}$ ;
   $\eta_{\text{MAX}} = \eta_{\text{FULL}}$ ;

  // Step 1: Place wavelength converter with full WC capability
   $[\bar{\mathbf{V}}_C, \bar{\mathbf{V}}_S, \eta_{\text{MAX}}] = \text{Place\_Device}(\text{FWC}, \bar{\mathbf{V}}_C, \bar{\mathbf{V}}_S, \eta_{\text{MAX}})$ ;

  // Step 2: Place light splitter with full LS capability
   $[\bar{\mathbf{V}}_C, \bar{\mathbf{V}}_S, \eta_{\text{MAX}}] = \text{Place\_Device}(\text{FLS}, \bar{\mathbf{V}}_C, \bar{\mathbf{V}}_S, \eta_{\text{MAX}})$ ;

  // Step 3: Adjust degree of wavelength conversion  $\mathbf{g}_C$ 
   $[\mathbf{V}_C, \mathbf{g}_C, \eta_{\text{MAX}}] = \text{Adjust\_Degree}(\text{FWC}, \mathbf{V}_C, \mathbf{g}_C, \eta_{\text{MAX}})$ ;
```

```

// Step 4: Adjust degree of light splitting  $\mathbf{g}_s$ 
[  $\mathbf{V}_s$ ,  $\mathbf{g}_s$ ,  $\eta_{\text{MAX}}$  ] = Adjust_Degree(FLS,  $\mathbf{V}_s$ ,  $\mathbf{g}_s$ ,  $\eta_{\text{MAX}}$ );

// Find highest cost efficiency
Return [  $\mathbf{V}_c$ ,  $\mathbf{g}_c$ ,  $\mathbf{V}_s$ ,  $\mathbf{g}_s$ ,  $\eta_{\text{MAX}}$  ];
END

```

6.2.3 Place light splitter first (LSF) strategies

LSF strategies will place wavelength converters after light splitters. We propose and study two algorithms, namely, LSF with Incremental placement (LSF-I) and LSF with Decremental placement (LSF-D), similar to WCF strategies. Since the basic steps are essentially the same as those in WCF-I and WCF-D, we will show the main algorithms below without giving further explanation.

6.2.3.1 LSF-I algorithm

Table 6.5 LSF-I algorithm (Main algorithm)

```

// LSF-I Algorithm (Main algorithm)
BEGIN
// Step 0: Initialization, all nodes have no WC and LS
 $\mathbf{V}_c = \mathbf{V}_s = \emptyset$ ;
Establish generic graph  $\mathbf{G}'$  based on  $\mathbf{V}_c$  and  $\mathbf{V}_s$ ; // which are  $\emptyset$ 
Proceed MC-RWA on  $\mathbf{G}'$ , find  $\eta_{\emptyset}$ ;
 $\eta_{\text{MAX}} = \eta_{\emptyset}$ ;

// Step 1: Place light splitters with full LS capability
[  $\mathbf{V}_c$ ,  $\mathbf{V}_s$ ,  $\eta_{\text{MAX}}$  ] = Place_Device(FLS,  $\mathbf{V}_c$ ,  $\mathbf{V}_s$ ,  $\eta_{\text{MAX}}$ );

// Step 2: Place wavelength converters with full WC capability
[  $\mathbf{V}_c$ ,  $\mathbf{V}_s$ ,  $\eta_{\text{MAX}}$  ] = Place_Device(FWC,  $\mathbf{V}_c$ ,  $\mathbf{V}_s$ ,  $\eta_{\text{MAX}}$ );

// Step 3: Adjust degree of light splitting  $\mathbf{g}_s$ 
[  $\mathbf{V}_s$ ,  $\mathbf{g}_s$ ,  $\eta_{\text{MAX}}$  ] = Adjust_Degree(FLS,  $\mathbf{V}_s$ ,  $\mathbf{g}_s$ ,  $\eta_{\text{MAX}}$ );

// Step 4: Adjust degree of wavelength conversion  $\mathbf{g}_c$ 
[  $\mathbf{V}_c$ ,  $\mathbf{g}_c$ ,  $\eta_{\text{MAX}}$  ] = Adjust_Degree(FWC,  $\mathbf{V}_c$ ,  $\mathbf{g}_c$ ,  $\eta_{\text{MAX}}$ );

// Find highest cost efficiency

```

```

Return [  $V_c$ ,  $g_c$ ,  $V_s$ ,  $g_s$ ,  $\eta_{MAX}$  ];
END

```

6.2.3.2 LSF-D algorithm

Table 6.6 LSF-D algorithm (Main algorithm)

```

// LSF-D Algorithm (Main algorithm)
BEGIN
  // Step 0: Initialization, all nodes have full WC and LS
   $V_c = V_s = V_N$ ; // i.e.  $\bar{V}_c = \bar{V}_s = \emptyset$ 
  Establish generic graph  $G'$  based on  $V_c$  and  $V_s$ ;
  Proceed MC-RWA on  $G'$ , find  $\eta_{FULL}$ ;
   $\eta_{MAX} = \eta_{FULL}$ ;

  // Step 1: Place light splitter with full LS capability
  [ $\bar{V}_c$ ,  $\bar{V}_s$ ,  $\eta_{MAX}$ ] = Place_Device(FLS,  $\bar{V}_c$ ,  $\bar{V}_s$ ,  $\eta_{MAX}$ );

  // Step 2: Place wavelength converters with full WC capability
  [ $\bar{V}_c$ ,  $\bar{V}_s$ ,  $\eta_{MAX}$ ] = Place_Device(FWC,  $\bar{V}_c$ ,  $\bar{V}_s$ ,  $\eta_{MAX}$ );

  // Step 3: Adjust degree of light splitting  $g_s$ 
  [ $V_s$ ,  $g_s$ ,  $\eta_{MAX}$ ] = Adjust_Degree(FLS,  $V_s$ ,  $g_s$ ,  $\eta_{MAX}$ );

  // Step 4: Adjust degree of wavelength conversion  $g_c$ 
  [ $V_c$ ,  $g_c$ ,  $\eta_{MAX}$ ] = Adjust_Degree(FWC,  $V_c$ ,  $g_c$ ,  $\eta_{MAX}$ );

  // Find highest cost efficiency
  Return [  $V_c$ ,  $g_c$ ,  $V_s$ ,  $g_s$ ,  $\eta_{MAX}$  ];
END

```

6.3 Complexity analysis and adaptation of algorithms for other optimization criteria

6.3.1 Complexity analysis

The kernel of the proposed algorithm is to establish the generic graph G' and execute MC-RWA on G' . The complexity of the algorithm can be analyzed in two parts – the executing times of the kernel and the complexity of the kernel itself.

▪ *Executing times of the kernel*

Since it is hard to know when the highest cost efficiency can be found, we will use the maximum times of possible executions to estimate the complexity. For WCF-I, during the initialization step, the kernel will be executed once. The place device steps will be executed at most N rounds, in each round the number of executions will be decreased by 1 from N to 1. Therefore, the number of times the kernel will be executed in each of Step 1 and Step 2 is $N(N+1)/2$. During the adjust degree steps, the number of executions are decided by the number of nodes that contain the devices, and the range of the device degree $[g_{\min}, g_{\max}]$. Thus, the number of times the degree of wavelength converters will be adjusted is $\sum_{v \in V_C} \lfloor W/2 \rfloor$, and light splitter is $\sum_{v \in V_S} (d_{v_i} - 1)$. Hence, the total number of times the kernel will be executed for all the steps is:

$$1 + 2 \cdot N(N+1)/2 + \sum_{v \in V_C} \lfloor W/2 \rfloor + \sum_{v \in V_S} (d_{v_i} - 1) \quad (6.3)$$

The above expression can be simplified as:

$$O(N^2) + O(N \cdot W) + O(N \cdot d_{avg}), \quad \text{where } d_{avg} = \frac{1}{N} \sum_{v \in V_N} d_v \quad (6.4)$$

▪ *Complexity of the kernel*

The complexity of the kernel can be found in section 5.2.2, Equation (5.16).

6.3.2 Adaptation of algorithms for other optimization criteria

The proposed four algorithms can be used to find the optimal placement for wavelength converters and light splitters. In many previous studies on the placement problem, the number of converters or splitters is given beforehand. By adding a constraint on the maximum number of devices in the function

Place_Device, our algorithms can also be used for finding the optimal placement for a given number of converters or splitters.

When the maximum number of devices is reached after many rounds of adding devices (for WCF-I and LSF-I) or removing devices (for WCF-D and LSF-D), the loop will be stopped. Therefore, the final placement will be the best solution that can be found by the algorithms for the given number of wavelength converters or splitters. To make the discussion simple, we do not place a constraint on the number of devices in this paper so that the optimal solution for a given network topology will be found.

The proposed four algorithms can also be used for criteria other than the cost efficiency. By simply replacing the comparison condition for cost efficiency in function **Place_Device** and function **Adjust_Degree** with a different measure, the algorithms can be used to find the optimal configuration based on different criteria. In this thesis, only the cost efficiency criteria is considered and studied.

6.4 Simulation and numerical results

6.4.1 Topology and assumptions

A simulation study has been carried out for our proposed algorithms. The topologies used in our simulation are the 9-node mesh-torus network and 14-node NSFNet network shown in Figure A.2 and Figure A.5 respectively.

The assumptions made for our simulation study are as follows:

- Each link represents two unidirectional fibers, and each fiber consists of 8 wavelengths, i.e., $W = 8$.
- The group size of the dynamic multicast request is $z_r = 4$, i.e., for one

multicast, there exists one source node and three destination nodes. We set an observation period of over 10,000 multicast requests from each node.

- The request arrive interval ($1/\lambda$) follows Poisson distribution and the service holding time ($1/\mu$) follows the exponential distribution. The mean service holding time is 1.0 second and the traffic load is calculated in Erlang (λ/μ).
- The default unit cost of a wavelength link edge is $\alpha = 10$ per 100km, the cost of wavelength conversion edge is $\kappa_c \cdot \alpha = 20$, the cost of light splitting edge is $\kappa_{sp} \cdot \alpha = 10$, and the cost of OXC edge is $\kappa_{sw} \cdot \alpha = 1$.
- For simplicity, only the Convert-Switch mode and Convert-Split mode shown in Figure 5.3(a) and Figure 5.3(c) are assumed.
- Three performance metrics are used, and their definitions are as follows:
 - 1) *Request Blocking Probability (RBP)* - P_b : This is the ratio of failed multicast connection requests to total multicast connection requests:

$$P_b = 1 - \sum r_{Success} / \sum r \quad (6.5)$$

- 2) *Average Multicast Tree Cost (AMTC)* - $C(T)$: This is the average cost value of all the successfully established multicast trees:

$$C(T) = \sum C(T_{r_{Success}}) / \sum r_{success} \quad (6.6)$$

- 3) *Cost Efficiency (CE)* - η : This is the ratio of multicast requests success probability to average multicast tree cost, the formula for computing this is given in equation (6.1).

6.4.2 Numerical results

6.4.2.1 Execution of the four proposed algorithms

Table 6.7 shows the steps and rounds for the 4 proposed algorithms WCF-I,

LSF-I, WCF-D and LSF-D for the NSFNet network when traffic load is 30 Erlang. In *Step 1* and *Step 2* for WCF-I and LSF-I, the wavelength converters and light splitters are added onto the nodes one by one, in increasing CE for each round. Similarly in *Step 1* and *Step 2* for WCF-D and LSF-D, the wavelength converters and light splitters are removed from nodes one by one, in increasing CE for each round. Some of the steps or rounds are skipped because of unsatisfied criteria, e.g., a higher CE cannot be found any more in the current step. The skipped steps and rounds are denoted by $\Downarrow$. In *Step 3* and *Step 4*, the degree of every available wavelength converter and light splitter is determined based on the CE. Finally, the ultimate highest CE is found among the four highest CEs found by the four respective algorithms. For the case shown in Table 6.7, the final CE found by WCF-D is the highest among the four proposed algorithms.

Table 6.7 Illustration of the steps and rounds, NSFNet14, Traffic load = 30Erl

<i>Algorithms</i> [Step, Round]	WCF-I	LSF-I	WCF-D	LSF-D
[Step 0, Round 00]	∅	∅	Full WC & LS	Full WC & LS
[Step 1, Round 00]	+ C06	+ S08	⇓	- S11
[Step 1, Round 01]	+ C08	+ S00		- S08
[Step 1, Round 02]	+ C10	⇓		- S13
[Step 1, Round 03]	+ C07			- S02
[Step 1, Round 04]	+ C03			- S01
[Step 1, Round 05]	+ C13			- S12
[Step 1, Round 06]	+ C04			- S09
[Step 1, Round 07]	+ C01			- S00
[Step 1, Round 08]	+ C12			⇓
[Step 1, Round 09]	+ C11			
[Step 1, Round 10]	+ C05			
[Step 1, Round 11]	+ C00			
[Step 1, Round 12]	⇓			
[Step 1, Round 13]				
[Step 2, Round 00]	+ S10	+ C06	- S11	⇓
[Step 2, Round 01]	+ S07	+ C10	- S08	
[Step 2, Round 02]	+ S04	+ C04	- S13	
[Step 2, Round 03]	⇓	+ C07	- S02	
[Step 2, Round 04]		+ C03	- S01	
[Step 2, Round 05]		+ C13	- S12	
[Step 2, Round 06]		+ C12	- S09	
[Step 2, Round 07]		+ C08	- S00	

[Step 2, Round 08]		+ C11		
[Step 2, Round 09]		+ C02		
[Step 2, Round 10]		+ C01		
[Step 2, Round 11]		+ C05	↓	
[Step 2, Round 12]		+ C09		
[Step 2, Round 13]		↓		
[Step 3, Round 00]	C00 / d=4	S00 / d=0	C00 / d=3	S03 / d=2
[Step 3, Round 01]	C01 / d=4	S08 / d=1	C01 / d=4	S04 / d=3
[Step 3, Round 02]	C03 / d=4		C02 / d=2	S05 / d=2
[Step 3, Round 03]	C04 / d=4		C03 / d=4	S06 / d=4
[Step 3, Round 04]	C05 / d=4		C04 / d=4	S07 / d=2
[Step 3, Round 05]	C06 / d=3		C05 / d=2	S10 / d=4
[Step 3, Round 06]	C07 / d=4		C06 / d=4	
[Step 3, Round 07]	C08 / d=4	↓	C07 / d=4	
[Step 3, Round 08]	C10 / d=3		C08 / d=4	
[Step 3, Round 09]	C11 / d=4		C09 / d=4	↓
[Step 3, Round 10]	C12 / d=4		C10 / d=4	
[Step 3, Round 11]	C13 / d=4		C11 / d=4	
[Step 3, Round 12]	↓		C12 / d=4	
[Step 3, Round 13]			C13 / d=4	
[Step 4, Round 00]	S04 / d=3	C01 / d=4	S03 / d=3	C00 / d=4
[Step 4, Round 01]	S07 / d=3	C02 / d=4	S04 / d=3	C01 / d=4
[Step 4, Round 02]	S10 / d=4	C03 / d=4	S05 / d=0	C02 / d=4
[Step 4, Round 03]		C04 / d=4	S06 / d=4	C03 / d=4
[Step 4, Round 04]		C05 / d=4	S07 / d=3	C04 / d=4
[Step 4, Round 05]		C06 / d=4	S10 / d=4	C05 / d=2
[Step 4, Round 06]		C07 / d=4		C06 / d=4
[Step 4, Round 07]		C08 / d=4		C07 / d=4
[Step 4, Round 08]	↓	C09 / d=3		C08 / d=4
[Step 4, Round 09]		C10 / d=4	↓	C09 / d=4
[Step 4, Round 10]		C11 / d=4		C10 / d=4
[Step 4, Round 11]		C12 / d=4		C11 / d=4
[Step 4, Round 12]		C13 / d=4		C12 / d=4
[Step 4, Round 13]		↓		C13 / d=4
η	1.753×10^{-3}	1.739×10^{-3}	1.760×10^{-3}	1.759×10^{-3}
P_b	1.7241%	2.3327%	1.5213%	1.5358%
$C(T)$	560.50	561.61	559.65	559.68

+/-: Add/Remove; C/S: Converter/Splitter; d: Degree; Numbers: Node Index

Figures 6.1 ~ 6.3 show the stepwise variation of CE, RBP and AMTC in different rounds in search of the highest CE. The round number in these figures excludes all the skipped steps and rounds.

From the results shown in Table 6.7, it can be seen that the incremental placement algorithms (WCF-I and LSF-I) take more rounds to place wavelength converters while decremental placement algorithms (WCF-D and LSF-D) take more

rounds to remove light splitters. This means in the final solution, the number of nodes that have wavelength converters is higher than the number of nodes that have light splitters. Therefore, it can be easily seen that all the four algorithms take more rounds to set the degrees of wavelength converters than the degrees of light splitters. This suggests that under the observed traffic load (i.e. 30 Erlang) for NSFNet network, the usage of wavelength converters contributes more to improve the CE than the usage of light splitters. The reason is as follows. Although light splitters may help to reduce the usage of wavelength links by splitting the light at an intermediate node inside a light-tree, it has to follow the wavelength continuous constraint, i.e., the light-tree is constraint to use a single wavelength for all the wavelength links concerned if there is no wavelength converter in the network. As the traffic load increases, it could be more difficult to reserve a single wavelength for the light-tree consisting of more wavelength links. On the other hand, the wavelength converter is meant for breaking the wavelength continuous constraint. With the help of wavelength converters, it is possible to establish a light-tree with wavelength links over different wavelength layers. But a separate lightpath from the source node to each of the destination node has to be established if there is no light splitter in the network. Nonetheless, this is probably not a major issue when traffic load is not high, as the number of wavelength links would be sufficient to accommodate the traffic. For example, at 30Erl for NSFNet, the RBP is low (< 0.09) even without wavelength converter and light splitter (see Figure 6.2, Round 0). In addition, the multicast tree cost could be reduced by carefully selecting wavelength conversion edges. Therefore, incorporating wavelength converters in the network will be more cost efficient than incorporating light splitters for low load multicast traffic.

From the results shown in Figures 6.1 ~ 6.3, it can be seen that incremental placement algorithms (WCF-I and LSF-I) start from a lower CE, higher RBP and higher AMTC. This is due to the different initialization stages of these two kinds of strategies, i.e. no capability of wavelength conversion and light splitting for WCF-I and LSF-I but full capability for WCF-D and LSF-D. As the execution of the algorithms proceeds, the CE of WCF-I and LSF-I increase sharply to approach those of WCF-D and LSF-D, as shown in Figure 6.1. From the initialization stage to round number 15, the CE of WCF-I increases 25% from 1.401×10^{-3} to 1.751×10^{-3} , LSF-I increases 23% from 1.401×10^{-3} to 1.725×10^{-3} , WCF-D increases 2.3% from 1.720×10^{-3} to 1.760×10^{-3} , and LSF-D only increases 2.15% from 1.720×10^{-3} to 1.757×10^{-3} . In addition, the rate of increase in CE for WCF-I is faster than that of LSF-I, e.g. the CE for WCF-I in round 1 is nearly the same as that of LSF-I in round 3. This is due to LSF-I trying to add splitters first within the first two rounds but the splitters contribute less to improving the CE, as explained above. A similar behavior is observed in the RBP and AMTC, as shown in Figures 6.2 and Figure 6.3 respectively.

Figure 6.1 Results of NSFNet ($T_r = 30Erl$) - variation of CE

Figure 6.2 Results of NSFNet ($T_r = 30Erl$) - variation of RBP

Figure 6.3 Results of NSFNet ($T_r = 30ErI$) - variation of AMTC

For WCF-I and LSF-I after round 15, and for WCF-D and LSF-D after round 8, the algorithms start the steps to find the degrees for the available wavelength converters and light splitters, after executing the steps to place them. From then on, the values of all the three metrics begin to stabilize, with the exception of a slight change for LSF-I in round 26. It can be seen that, in *Step 1* and *Step 2*, all the placed devices are assumed to have full capability of wavelength conversion and light splitting. Even after the degrees have been adjusted in subsequent steps, most of them remain close to full capability of wavelength conversion and light splitting. This phenomenon shows that in the process of finding the optimal configuration (in terms of CE) for the wavelength converters and light splitters, where to place a device is more crucial than that the device's capability (degree).

6.4.2.2 *Optimal configurations (in terms of CE) for different traffic loads*

We make the comparison based on two network topologies, i.e. Mesh9 and NSFNet shown in Figure A.2 and Figure A.5, under the same assumptions. Table 6.8 and Table 6.9 show the optimal configurations for Mesh9 and NSFNet respectively, in terms of CE yielded by the various algorithms for different traffic loads. The algorithm that yielded the highest CE overall is indicated in the last column. The configurations are shown in the *Node-Degree* format, where *Node* is the node identifier and *Degree* indicates the capability of wavelength conversion or fan-out of the light splitters.

Table 6.8 *Configuration of wavelength converter and light splitter – Mesh9*

Traffic Load	Converter (Node-Degree)	Splitter (Node-Degree)	Algorithm
30 Erl	0-4, 1-4, 4-4, 5-4, 6-4, 7-4, 8-4	∅	LSF-D
35 Erl	0-4, 1-4, 3-4, 4-4, 5-4, 6-4, 7-4, 8-4	∅	LSF-D
40 Erl	∅	∅	WCF-I
45 Erl	∅	∅	WCF-I
50 Erl	∅	∅	WCF-I
55 Erl	∅	∅	WCF-I
60 Erl	∅	∅	WCF-I

Table 6.9 *Configuration of wavelength converter and light splitter – NSFNet*

Traffic Load	Converter (Node-Degree)	Splitter (Node-Degree)	Algorithm
10 Erl	10-4, 4-4, 7-4, 3-4, 6-2, 1-2, 12-2, 5-4, 13-4, 9-4, 11-4	∅	LSF-I
15 Erl	3-4, 10-4, 4-4, 7-4, 6-4, 12-4, 1-4, 9-4, 5-4, 13-4, 8-4, 2-4	∅	WCF-I
20 Erl	1-1, 2-4, 3-4, 4-4, 5-4, 6-4, 7-4, 8-4,	∅	LSF-D

	9-4, 10-4, 11-3, 12-4, 13-4		
25 Erl	0-4, 1-4, 2-2, 3-4, 4-4, 5-4, 6-4, 7-4, 9-4, 10-4, 11-4, 12-4, 13-4	10-2	WCF-D
30 Erl	0-3, 1-4, 2-2, 3-4, 4-4, 5-2, 6-4, 7-4, 8-4, 9-4, 10-4, 11-4, 12-4, 13-4	3-3, 4-3, 6-4, 7-3, 10-4	WCF-D
35 Erl	0-2, 1-4, 2-4, 3-4, 4-4, 5-4, 6-4, 7-4, 8-4, 9-4, 10-4, 11-4, 12-4, 13-4	1-3, 2-3, 3-3, 4-3, 5-2, 6-4, 7-3, 9-2, 10-4	WCF-D
40 Erl	3-4, 6-4, 10-4, 12-4, 13-4, 8-4, 1-4, 4-4, 7-4, 2-4, 9-4, 5-4, 11-4	10-4, 7-3, 4-3, 6-4, 3-3, 9-2	WCF-I
45 Erl	0-4, 1-4, 2-4, 3-4, 4-4, 5-4, 6-4, 7-4, 8-4, 9-4, 10-2, 11-4, 12-4, 13-4	0-3, 1-3, 2-3, 3-3, 4-3, 5-2, 6-4, 7-3, 9-2, 10-4, 11-3, 12-3	LSF-D
50 Erl	0-4, 1-4, 2-4, 3-4, 4-4, 5-4, 6-4, 7-4, 8-4, 9-4, 10-4, 11-4, 12-3, 13-4	0-3, 1-3, 3-3, 4-3, 5-2, 6-4, 7-3, 9-2, 10-4, 11-3, 12-3, 13-3	LSF-D

Figure 6.4 ~ 6.9 show the results in graphical form, comparing 6 different cases. Note that in these figures, “Full” denotes the case of full wavelength conversion and splitting and “Non” denotes the case of no wavelength conversion and light splitting.

Figure 6.4 Cost Efficiency vs. Traffic Load – Mesh9

Figure 6.5 Cost Efficiency vs. Traffic Load – NSFNet

Figure 6.6 Request Blocking Probability vs. Traffic Load – Mesh9

Figure 6.7 Request Blocking Probability vs. Traffic Load – NSFNet

Figure 6.8 Average Multicast Tree Cost vs. Traffic Load – Mesh9

Figure 6.9 Average Multicast Tree Cost vs. Traffic Load – NSFNet

Mesh9 network is a small symmetrical network with $d_v = 4$ for all the nodes. From Table 6.8, it can be seen that the best configurations for different traffic loads do not contain any light splitters. In addition, for traffic load higher than 35Erl, the best configuration is the case in which no wavelength conversion and light splitting is used. From Figure 6.4, Figure 6.6 and Figure 6.8, it can be seen that although the case of full wavelength conversion and splitting has the lowest RBP, its CE is significantly lower than other cases due to its extreme high AMTC. The reason for this is that for a small network with a large d_v , such as Mesh9, when the traffic load is high, request blocking is mainly caused by the lack of wavelength links rather than the lack or limitation of wavelength converters or light splitters. The improvement in RBP resulting from introducing wavelength converters and light splitters is insufficient to overcome the additional cost introduced by these devices.

It can be seen from Figure 6.4 that, for traffic load higher than 45Erl, the cost efficiency of WCF-D and LSF-D decrease faster than the cost efficiency of WCF-I and LSF-I because decremental algorithms have the tendency of keeping more wavelength converters and light splitters so as to lower the total cost efficiency.

Note that the average node degree of NSFNet network is $\bar{d}_v = 3$, which is lower than the multicast group size. The positive effect of introducing light splitters when traffic load is high can be seen from Table 6.9. When traffic load is high, light splitter is cost efficient; however, when traffic load is low, e.g. less than 25Erl, no light splitter is chosen in the optimal configurations. On the contrary, wavelength converters are always chosen in the optimal configurations. Results show that most of the nodes need full capability of wavelength conversion at various traffic loads. From Figure 6.5, Figure 6.7 and Figure 6.9, it can be seen that all four cases of our proposed algorithms are neck in neck in performance based on the three chosen

metrics. The performances of WCF-I, LSF-I, WCF-D and LSF-D are very close to those of Full WC Full LS for RBP, much higher in CE when load is low (e.g. lower than 30Erl), and slightly higher when load is high (e.g. higher than 30Erl). It can be concluded that for the NSFNet network, when traffic load is low, neither wavelength conversion nor light splitting is cost efficient; when traffic load is high, the cost efficiency for the case with full wavelength conversion and light splitting capability for all nodes approaches that of the optimal configuration. It is also clear from this study that different traffic loads have different optimal configurations in terms of CE.

6.4.2.3 *Robustness of configuration optimized for a specific traffic load*

It is interesting to know whether the optimal configuration (in terms of CE) for wavelength converters and light splitters for a specific traffic load for a network is still reasonably optimal for other different traffic loads. The results shown in Figure 6.10, Figure 6.11 and Figure 6.12 shed light on this question.

Based on the optimal placements of 30Erl on NSFNet obtained from the different algorithms, we make a comparison of the cost efficiency, request blocking probability and average multicast tree cost for different traffic loads and scenarios, which are, namely, 1) Full wavelength conversion and full light splitting; 2) Full wavelength conversion and null light splitting; 3) Null wavelength conversion and full light splitting; 4) Null wavelength conversion and null light splitting; 5)-8) Proposed 4 algorithms, WCF-I, WCF-D, LSF-I and LSF-D based on the optimal configurations for 30 Erl, which are shown in Table 6.9; 9) Results based on the optimal configurations for respective traffic loads, named as optimal.

Figure 6.10 Comparison among optimal configurations – CE vs. Traffic Load

It can be seen that the two cases of null wavelength conversion are the worst, i.e. low CE, high RBP and high AMTC. This clearly shows that without wavelength converters in the network, light splitters are not effective in improving the CE, RBP and AMTC, even though multicast traffic is expected to benefit from light splitting.

It is clear from the graphs that the optimal case always has the highest CE for different traffic loads though it is not the best based on RBP and AMTC. For the other 6 cases, the differences in performance are not so significant. At low traffic load (lower than 30Erl), the case of LSF-I (30Erl) and Full WC Null LS outperform WCF-I (30Erl), WCF-I (30Erl) outperforms WCF-D (30Erl), and WCF-D (30Erl) outperforms LSF-D (30Erl) and Full WC Full LS. At high traffic load (higher than

30Erl), the performances of these scheme are the exact reverse. It can be observed that although the RBP of Full WC Full LS is always lower than that of Full WC Null LS (see Figure 6.11), the CE of Full WC Null LS could be higher than Full WC Full LS, especially when the traffic load is low. Hence, we can conclude that light splitters are not cost effective when the traffic load is low.

Figure 6.11 Comparison among optimal configurations – RBP vs. Traffic Load

Figure 6.12 Comparison among optimal configurations – AMTC vs. Traffic Load

From this study, it is clear that the best configuration (in terms of CE) of wavelength converters and light splitters obtained at specific traffic load is not necessary the optimal for other scenarios and traffic loads but generally quite close to being optimal in terms of CE.

To summarize, the following conclusions can be drawn from this study:

- 1) The usage of wavelength converters contributes more to improve the cost efficiency than the usage of light splitters when traffic load is not high.
- 2) The decision on whether to place a wavelength converter and light splitter in a node is more important than the decision on the devices' capabilities. In other words, the sparse wavelength conversion and light splitting strategy is more crucial than the limited wavelength conversion and light splitting strategy.

- 3) For a small network with a relatively large average node degree, such as Mesh9, introducing wavelength converters and light splitters may not be cost effective. However, for a large network with a relatively small average node degree (e.g. less than the multicast group size), like NSFNet, introducing wavelength converters and light splitters is generally cost effective, especially when the traffic load is high.
- 4) The optimal configuration for wavelength converters and light splitters obtained for a specific traffic load and pattern will not always be optimal for other traffic loads and patterns. The optimal configuration is traffic dependant.

6.5 Further discussions on cost efficiency

6.5.1 New definition of cost efficiency

From Figure 6.2, we realize that the blocking probability for the final configurations found by WCF-I and LSF-I are much higher than those found by WCF-D and LSF-D. The difference between highest η_H (found by WCF-D) and lowest η_L (found by LSF-I) is about $\Delta_\eta = (\eta_H - \eta_L) / \eta_L = 1.21\%$, however the difference between highest P_{bH} (found by LSF-I) and lowest P_{bL} (found by WCF-D) is up to $\Delta_{P_b} = (P_{bH} - P_{bL}) / P_{bL} = 53.3\%$. This obvious difference between Δ_η and Δ_{P_b} shows that the cost efficiency defined in Equation 6.1 is not sensitive to the blocking probability. In a real network, lower blocking probability to a certain extent is always a basic requirement to avoid losing customers, and such a definition of cost efficiency makes it virtually impossible to design a network having a very low blocking probability for all four proposed algorithms. For example, to lower the blocking probability from 1% to 0.1%, the numerator of the

objective function only slightly increases from 0.99 to 0.999, meanwhile the denominator, which reflects the cost of the network, may go up significantly. The cost efficiency would then be decreased.

To make the metric of cost efficiency more sensitive to the changes in blocking probability, we redefine the cost efficiency as follows:

$$\eta = -\log_{10} (P_b / P_{MAX} + P_{MIN}) / C(T), \text{ where } 0 < P_{MIN} \ll P_{MAX} < 1 \quad (6.7)$$

Here, P_b and $C(T)$ still denote the blocking probability of the request and the average multicast tree cost, respectively. P_{MIN} and P_{MAX} are constant, denoting the acceptable minimum and maximum blocking probability, respectively. P_{MIN} is a positive value that is close to 0, which is used to avoid the numerator becoming infinity when $P_b = 0$. When $P_b \rightarrow P_{MAX}$, the numerator will approach 0, which makes $\eta \rightarrow 0$. By introducing the logarithm scale in the numerator, the order of improvement in blocking probability can be reflected appropriately. In addition, by using P_{MAX} as an upper bound for the blocking probability, the cost efficiency will be set to 0 or even negative values when P_b is high. By means of this new definition of cost efficiency, we expect the final configurations found by the four proposed algorithms to have high cost efficiency and fairly low blocking probability.

6.5.2 Numerical results

A simulation study has been carried out based on the enhanced definition of cost efficiency with $P_{MAX} = 10^{-1}$ and $P_{MIN} = 10^{-7}$.

6.5.2.1 Execution of the four proposed algorithms

Table 6.10 and Table 6.11 show the steps and rounds for the 4 proposed algorithms WCF-I, LSF-I, WCF-D and LSF-D for the NSFNet network when traffic load is 5 Erlang and 25 Erlang. When traffic load is 5 Erlang, the blocking probability is always 0, therefore the cost efficiency is only decided by $C(T)$.

Table 6.10 Illustration of the steps and rounds, NSFNet14, Traffic Load = 5Erl

<i>Algorithms</i> [Step, Round]	WCF-I	LSF-I	WCF-D	LSF-D
[Step 0, Round 00]	∅	∅	Full WC & LS	Full WC & LS
[Step 1, Round 00]	+ C10	+ S00	⇓	- S10
[Step 1, Round 01]	+ C04	⇓		- S07
[Step 1, Round 02]	+ C07			- S04
[Step 1, Round 03]	+ C03			- S05
[Step 1, Round 04]	+ C06			- S03
[Step 1, Round 05]	+ C01			- S12
[Step 1, Round 06]	+ C12			- S01
[Step 1, Round 07]	+ C13			- S06
[Step 1, Round 08]	+ C05			- S11
[Step 1, Round 09]	+ C09			- S09
[Step 1, Round 10]	+ C11			- S13
[Step 1, Round 11]	+ C02			- S02
[Step 1, Round 12]	+ C00			- S08
[Step 1, Round 13]	+ C08			- S00
[Step 2, Round 00]	⇓	+ C10	- S10	⇓
[Step 2, Round 01]		+ C04	- S07	
[Step 2, Round 02]		+ C07	- S04	
[Step 2, Round 03]		+ C03	- S05	
[Step 2, Round 04]		+ C06	- S03	
[Step 2, Round 05]		+ C01	- S12	
[Step 2, Round 06]		+ C12	- S01	
[Step 2, Round 07]		+ C05	- S06	
[Step 2, Round 08]		+ C13	- S11	
[Step 2, Round 09]		+ C09	- S09	
[Step 2, Round 10]		+ C11	- S13	
[Step 2, Round 11]		+ C02	- S02	
[Step 2, Round 12]		⇓	- S08	
[Step 2, Round 13]			- S00	
[Step 3, Round 00]	C00 / d=4	S00 / d=0	C00 / d=4	⇓
[Step 3, Round 01]	C01 / d=4	⇓	C01 / d=4	
[Step 3, Round 02]	C02 / d=1		C02 / d=1	
[Step 3, Round 03]	C03 / d=4		C03 / d=4	
[Step 3, Round 04]	C04 / d=4		C04 / d=4	
[Step 3, Round 05]	C05 / d=4		C05 / d=4	
[Step 3, Round 06]	C06 / d=3		C06 / d=4	
[Step 3, Round 07]	C07 / d=4		C07 / d=4	
[Step 3, Round 08]	C08 / d=4		C08 / d=4	

[Step 3, Round 09]	C09 / d=1		C09 / d=1	
[Step 3, Round 10]	C10 / d=4		C10 / d=4	
[Step 3, Round 11]	C11 / d=4		C11 / d=4	
[Step 3, Round 12]	C12 / d=4		C12 / d=4	
[Step 3, Round 13]	C13 / d=4		C13 / d=4	
[Step 4, Round 00]		C01 / d=4		C00 / d=4
[Step 4, Round 01]		C02 / d=3		C01 / d=4
[Step 4, Round 02]		C03 / d=3		C02 / d=1
[Step 4, Round 03]		C04 / d=4		C03 / d=4
[Step 4, Round 04]		C05 / d=4		C04 / d=4
[Step 4, Round 05]		C06 / d=4		C05 / d=2
[Step 4, Round 06]	⇓	C07 / d=4	⇓	C06 / d=4
[Step 4, Round 07]		C09 / d=1		C07 / d=4
[Step 4, Round 08]		C10 / d=4		C08 / d=4
[Step 4, Round 09]		C11 / d=4		C09 / d=1
[Step 4, Round 10]		C12 / d=4		C10 / d=4
[Step 4, Round 11]		C13 / d=4		C11 / d=4
[Step 4, Round 12]		⇓		C12 / d=4
[Step 4, Round 13]				C13 / d=4
η	1.475×10^{-2}	1.475×10^{-2}	1.475×10^{-2}	1.475×10^{-2}
P_b	0%	0%	0%	0%
$C(T)$	474.45	474.49	474.45	474.45

+/-: Add/Remove; C/S: Converter/Splitter; d: Degree; Numbers: Node Index

Table 6.11 Illustration of the steps and rounds, NSFNet14, Traffic Load = 25Erl

<i>Algorithms</i> [Step, Round]	WCF-I	LSF-I	WCF-D	LSF-D
[Step 0, Round 00]	∅	∅	Full WC & LS	Full WC & LS
[Step 1, Round 00]	+ C06	+ S07	- C05	- S08
[Step 1, Round 01]	+ C10	+ S03	- C08	- S11
[Step 1, Round 02]	+ C07	+ S06		- S00
[Step 1, Round 03]	+ C03	+ S10		- S01
[Step 1, Round 04]	+ C04	+ S01		- S06
[Step 1, Round 05]	+ C00	+ S13		
[Step 1, Round 06]	+ C12	+ S04		
[Step 1, Round 07]	+ C01	+ S05	⇓	
[Step 1, Round 08]	+ C08	+ S09		
[Step 1, Round 09]	+ C13			⇓
[Step 1, Round 10]	+ C11	⇓		
[Step 1, Round 11]	+ C09			
[Step 1, Round 12]	+ C05			
[Step 1, Round 13]	+ C02			
[Step 2, Round 00]	+ S07	+ C06	⇓	- C02
[Step 2, Round 01]	+ S05	+ C07		⇓
[Step 2, Round 02]	+ S04	+ C10		
[Step 2, Round 03]	+ S01	+ C04		
[Step 2, Round 04]	+ S03	+ C01		
[Step 2, Round 05]	⇓	+ C12		
[Step 2, Round 06]		+ C03		

[Step 2, Round 07]		+ C09		
[Step 2, Round 08]		+ C00		
[Step 2, Round 09]		+ C08		
[Step 2, Round 10]		⇓		
[Step 2, Round 11]				
[Step 2, Round 12]				
[Step 2, Round 13]				
[Step 3, Round 00]	C00 / d=4	S01 / d=3	C00 / d=4	S02 / d=3
[Step 3, Round 01]	C01 / d=4	S03 / d=3	C01 / d=4	S03 / d=3
[Step 3, Round 02]	C02 / d=4	S04 / d=4	C02 / d=4	S04 / d=3
[Step 3, Round 03]	C03 / d=4	S05 / d=2	C03 / d=4	S05 / d=2
[Step 3, Round 04]	C04 / d=4	S06 / d=4	C04 / d=4	S07 / d=3
[Step 3, Round 05]	C05 / d=4	S07 / d=3	C06 / d=4	S09 / d=2
[Step 3, Round 06]	C06 / d=4	S09 / d=2	C07 / d=1	S10 / d=4
[Step 3, Round 07]	C07 / d=4	S10 / d=2	C09 / d=4	S12 / d=3
[Step 3, Round 08]	C08 / d=4	S13 / d=2	C10 / d=4	S13 / d=3
[Step 3, Round 09]	C09 / d=4	⇓	C11 / d=2	⇓
[Step 3, Round 10]	C10 / d=4		C12 / d=4	
[Step 3, Round 11]	C11 / d=4		C13 / d=4	
[Step 3, Round 12]	C12 / d=4		⇓	
[Step 3, Round 13]	C13 / d=4			
[Step 4, Round 00]	S01 / d=3	C00 / d=4	S00 / d=3	C00 / d=4
[Step 4, Round 01]	S03 / d=3	C01 / d=4	S01 / d=3	C01 / d=4
[Step 4, Round 02]	S04 / d=3	C03 / d=2	S02 / d=3	C03 / d=4
[Step 4, Round 03]	S05 / d=2	C04 / d=4	S03 / d=3	C04 / d=1
[Step 4, Round 04]	S07 / d=3	C06 / d=4	S04 / d=3	C05 / d=4
[Step 4, Round 05]	⇓	C07 / d=4	S05 / d=2	C06 / d=4
[Step 4, Round 06]		C08 / d=4	S06 / d=3	C07 / d=4
[Step 4, Round 07]		C09 / d=1	S07 / d=3	C08 / d=4
[Step 4, Round 08]		C10 / d=4	S08 / d=3	C09 / d=4
[Step 4, Round 09]		C12 / d=4	S09 / d=2	C10 / d=4
[Step 4, Round 10]		⇓	S10 / d=4	C11 / d=4
[Step 4, Round 11]			S11 / d=3	C12 / d=4
[Step 4, Round 12]			S12 / d=3	C13 / d=4
[Step 4, Round 13]			S13 / d=3	
η	2.721×10^{-3}	2.667×10^{-3}	2.590×10^{-3}	2.714×10^{-3}
P_b	0.3517%	0.3527%	0.3747%	0.3396%
$C(T)$	534.36	544.75	550.64	541.17

+/-: Add/Remove; C/S: Converter/Splitter; d: Degree; Numbers: Node Index

Figures 6.13 ~ 6.14 show the stepwise variation of CE and AMTC in different rounds when traffic load is 5Erl; Figures 6.15 ~ 6.17 show the stepwise variation of CE, RBP and AMTC when traffic load is 25Erl.

Figure 6.13 Results of NSFNet ($T_r = 5ErI$) - variation of CE

Figure 6.14 Results of NSFNet ($T_r = 5ErI$) - variation of AMTC

Figure 6.15 Results of NSFNet ($T_r = 25Erl$) - variation of CE

Figure 6.16 Results of NSFNet ($T_r = 25Erl$) - variation of RBP

Figure 6.17 Results of NSFNet ($T_r = 25Erl$) - variation of AMTC

From the results, it can be observed that when traffic load is low (e.g. 5Erl), the usage of wavelength converters contributes more to improve the cost efficiency than the usage of light splitters. In addition, determining where to place wavelength converters and light splitters is far more important than determining the capability of individual wavelength converters and light splitters. The difference between highest η_H (found by WCF-I) and lowest η_L (found by WCF-D) is about $\Delta_\eta = (\eta_H - \eta_L) / \eta_L = 5.06\%$, and the difference between highest P_{bH} (found by WCF-D) and lowest P_{bL} (found by LSF-D) is decreased to $\Delta_{P_b} = (P_{bH} - P_{bL}) / P_{bL} = 10.3\%$. The difference between Δ_η and Δ_{P_b} has been reduced significantly and the new definition of cost efficiency shows adequate sensitivity to the blocking probability.

6.5.2.2 *Optimal configurations (in terms of CE) for different traffic loads*

Table 6.12 and Table 6.13 show the optimal configurations for Mesh9 and NSFNet respectively, in terms of the new CE yielded by the various algorithms for different traffic loads. The algorithm that yielded the highest CE overall is indicated in the last column. Compared with Table 6.8 and Table 6.9, which are results based on previous definition of CE, the algorithm tends to keep more wavelength converters and light splitters in order to lower the blocking probability.

Table 6.12 *Configuration of wavelength converter and light splitter – Mesh9*

Traffic Load	Converter (Node-Degree)	Splitter (Node-Degree)	Algorithm
20 Erl	5-4, 3-2	∅	WCF-I
25 Erl	1-3	∅	WCF-I
30 Erl	0-4, 2-1, 3-4, 4-4, 5-4, 7-4, 8-4	0-4, 1-4, 2-4, 3-4, 4-4, 5-4, 6-4, 7-4, 8-4	WCF-D
35 Erl	1-3, 8-4, 3-4, 0-4, 5-4, 2-4, 7-1, 4-4, 6-4	∅	WCF-I
40 Erl	8-4, 6-4, 1-4, 5-4, 3-4, 4-3, 2-4, 0-4, 7-4	∅	WCF-I

Table 6.13 *Configuration of wavelength converter and light splitter – NSFNet*

Traffic Load	Converter (Node-Degree)	Splitter (Node-Degree)	Algorithm
5 Erl	0-4, 1-4, 2-1, 3-4, 4-4, 5-4, 6-4, 7-4, 8-4, 9-1, 10-4, 11-4, 12-4, 13-4	∅	LSF-D
10 Erl	0-4, 1-4, 2-3, 3-4, 4-4, 5-4, 6-4, 7-4, 8-4, 9-4, 10-4, 11-4, 12-4, 13-4	∅	LSF-D
15 Erl	0-4, 1-4, 2-4, 3-4, 4-4, 5-4, 6-2, 7-4, 8-4, 9-4, 10-4, 11-4, 12-4, 13-4	0-3, 1-3, 2-3, 4-3, 5-2, 7-3, 9-2, 10-4, 11-3, 13-3	LSF-D
20 Erl	0-4, 1-4, 2-4, 3-4, 4-4, 5-4, 6-4, 7-4, 8-4, 9-4, 10-4, 11-4, 12-4, 13-4	1-3, 3-3, 4-3, 5-2, 6-4, 7-3, 10-4, 11-3, 12-3, 13-3	LSF-D
25 Erl	6-4, 10-4, 7-4, 3-4, 4-4, 0-4, 12-4, 1-4, 8-4, 13-4, 11-4, 9-4, 5-4, 2-4	7-3, 5-2, 4-3, 1-3, 3-3	WCF-I

Figure 6.18 ~ 6.23 show the results in graphical form, comparing the six different cases, as in Figures 6.4 to 6.9 in Section 6.4.2.2.

For the Mesh9 network, the RBPs found by WCF-I, LSF-I, WCF-D and LSF-D are close under the observed traffic loads. The AMTCs has more significant impact to the CE. Obviously the incremental algorithms (WCF-I and LSF-I) tend to have a low AMTC because they try to minimize the usage of wavelength converters and light splitters. This can explain why WCF-I and LSF-I will find a configuration with a higher CE in general, except for one instance when the traffic load is 30 Erlang, where WCF-D finds a configuration with the highest CE (see Figure 6.18). The reason can be easily found in Figure 6.20 that this configuration has zero RBP. As zero RBP can lead to a very high CE from the new definition of CE, it shows that the new definition of CE is sensitive to RBP.

Figure 6.18 Cost Efficiency vs. Algorithms – Mesh9

Figure 6.19 Cost Efficiency vs. Algorithms – NSFNet

Figure 6.20 Request Blocking Probability vs. Algorithms – Mesh9

Figure 6.21 Request Blocking Probability vs. Algorithms – NSFNet

Figure 6.22 Average Multicast Tree Cost vs. Algorithms – Mesh9

Figure 6.23 Average Multicast Tree Cost vs. Algorithms – NSFNet

For the NSFNet network, all the four proposed algorithms can find configurations with higher CE, lower RBP and lower AMTC compared with Non and Full cases under different traffic loads, but the differences among the four algorithms are very small.

6.6 Summary

In this chapter, we study the placement problem of wavelength converters and light splitters using the Generic Graph Model. Four algorithms, namely, WCF-I, WCF-D, LSF-I and LSF-D, have been proposed to find the optimal configuration of wavelength converters and light splitters in a network. Simulation studies have been carried out and three metrics: Cost Efficiency, Request Blocking Probability and Average Multicast Tree Cost are used to compare the different configurations and

algorithms. All of the proposed algorithms are able to find a placement configuration with a higher CE and comparable RBP, compared with the case of full wavelength conversion and full light splitting. The study led us to make the following conclusions. Firstly, wavelength converters contribute more to improve the cost efficiency than light splitters when traffic load is not very high. Secondly, determining where to place wavelength converters and light splitters is far more important than determining the capability of individual wavelength converters and light splitters. The former should be part of critical design consideration where the latter could be used for fine tuning. Thirdly, introducing wavelength converters and light splitters is generally cost efficient for a large network with a relatively small average node degree when traffic load is high. The optimal configuration that is obtained for a specific traffic load will not always be optimal for other traffic loads.

Chapter 7

Conclusion and Future Recommendations

Over the last decade there is a massive increase in network bandwidth due to a wide-spread deployment of WDM transmission technology. It has heightened the need for faster switching in the core network. The rapid advancement and evolution of optical communication technologies makes it possible to move beyond point-to-point WDM transmission systems to an all-optical backbone network that can take full advantage of the available bandwidth by eliminating the need for per-hop packet forwarding. As IP over WDM network architecture moves toward a dynamically reconfigurable switched architecture, the increased flexibility and agility require more intelligence to be built into the IP over WDM network elements to facilitate network control and management.

In this thesis, we study the control and management of the multicast-capable IP over WDM networks. Our focus is on investigating the performance of various protocols for dynamic lightpath provisioning, and developing efficient algorithms to improve the network performance in multicast-capable IP over WDM networks, where dynamic traffic grooming, optical multicasting, placement of wavelength converters and light splitters are supported for both unicast and multicast IP traffic loads. This chapter summarizes our main contributions and recommends some potential directions for future research.

7.1 Research contributions

In Chapter 3, performances of various control protocols for dynamic lightpath provisioning are studied; they are namely, the local information-based fixed and fixed alternate link routing approaches, the distributed global information-based link state approaches with backward, forward, parallel reservation signaling and the centralized global information-based link state approach for dynamic lightpath provisioning. Three performance metrics, namely, RBP, CST and CMO are used for the comparison. The centralized global information-based link state approach gives the least RBP and very low CMO when the processing delay is in the order of 10^{-5} second or lower. In a small-scale network, such as NJ LATA, its CST is comparable with the distributed global information-based link state approach with forward-reservation signaling, and much lower than the other distributed global information-based link state approaches. As the network size increases, or the average event processing delay becomes larger, the centralized approach may not be scalable. Nevertheless, the distributed global information-based link state approach is also not scalable for large networks as the routing overhead will become excessive, rendering the networks to be practically useless. As such, local information-based fixed-alternate link routing approach is more appropriate for lightpath provisioning because of its simplicity and low control message overhead.

In Chapter 4 the signaling effects on dynamic traffic grooming for both unicast and multicast IP traffic models are studied. A new technique called DTG-PRL is proposed to solve the DTG problem. DTG-PRL consists of two steps, 1) Pre-reserve some lightpaths based on statistical traffic observations; 2) Dynamically grooms the traffic based on the established virtual topology. For both unicast and multicast IP traffic model, the ILP formulations and the heuristic algorithms (MWU/MWU-M)

are proposed. The MWU/MWU-W is found to have similar performance compared with ILP formulations. A performance comparison between DTG-PRL and a DTG algorithm called DTG-B is carried out. The results show that the proposed DTG-PRL outperforms DTG-B.

By using light splitters, data transmission from one point to multiple points can be realized in the optical layer. It eliminates O-E-O conversions between IP layer and WDM layer, and removes the bottlenecks for large bandwidth transmission. However, with wavelength converters and light splitters, multicast routing and wavelength assignment (MC-RWA) becomes more complex. In Chapter 5, a new graph model named Generic Graph Model is developed for dynamic MC-RWA for wavelength-routed WDM networks. Based on the GGM, an efficient shortest path tree algorithm is employed to solve dynamic MC-RWA for a number of light splitting and wavelength conversion scenarios. This provides a way to handle heterogeneous network structures. Based on this GGM, the blocking performance is investigated for Convert-Split mode and Split-Convert mode. The study shows that the GGM could provide a generic platform for studying the dynamic MC-RWA problem for different light splitting and wavelength conversion scenarios.

The placement problem for wavelength converters and light splitters is studied in Chapter 6 based on the GGM developed in chapter 5. Four new algorithms, namely, WCF-I, WCF-D, LSF-I and LSF-D, are proposed to find the optimal configuration of wavelength converters and light splitters. Three metrics: namely, Cost Efficiency, Request Blocking Probability and Average Multicast Tree Cost are used to compare the performances of different configurations and algorithms. All of the proposed algorithms are able to find a placement configuration with a higher CE and comparable RBP, compared with the case of full wavelength conversion and

full light splitting. The study led us to draw the following conclusions. Firstly, wavelength converters contribute more to improve the cost efficiency than light splitters when traffic load is not very high. Secondly, determining where to place wavelength converters and light splitters is far more important than determining the capability of individual wavelength converters and light splitters. The former should form a key part of critical design considerations where the latter could be used for fine tuning. Thirdly, introducing wavelength converters and light splitters is generally cost effective for a large network with a relatively small average node degree when traffic load is high. The optimal configuration that has the highest cost efficiency obtained for a specific traffic load will not always be optimal in cost efficiency for other traffic loads.

7.2 Recommendations for future research

Although research on control and management of multicast-capable IP over WDM networks has been ongoing for a few years, it is far from complete and many problems need to be carefully studied. For future research, the following directions may be pursued:

- **Strategy of information updating for dynamic traffic grooming in real-life networks** – In our study of signaling effect on dynamic traffic grooming in chapter 4, the strategy for information update is found to be one of the key factors affecting the performance of dynamic traffic grooming. Apart from WDM link state update on the optical layer [97], information update for dynamic traffic grooming includes the updating of virtual link status, which is essential for real-time IP traffic to be groomed. New and better strategies need

to be devised for implementation on IP and/or WDM layer such that the network could have the best trade-off between request blocking probability and control message overhead, high network resource utilization, and level of congestion.

- **Dynamic light-tree provisioning, protection and restoration** – Many previous research works only focus on lightpath provisioning, protection and restoration for unicast communication. As the concept of light-trees is gaining acceptance, more research effort is directed towards developing algorithms and other enabling technologies for optical multicasting. Existing routing and signaling protocols for lightpath provisioning have to be extended in order to support light-tree provisioning. Further more, since a single fiber cut can disrupt the transmission of information to many destinations, it is crucial to protect and have suitable mechanisms for fast restoration of multicast sessions. Therefore, dynamic light-tree provisioning, protection and restoration problem needs detailed studies.

- **Dynamic traffic grooming for light-tree based optical multicasting with wavelength conversion capability** – Efficiently grooming low-speed connections into high-capacity lightpaths will improve the network throughput and reduce the network cost. In Chapter 4, we present some work on dynamic traffic grooming for IP-based multicasting. However, in our study, data transmission in the WDM layer is confined to directed lightpaths from one source node to one destination node. Further investigations and novel algorithms are required to optimize the network resources efficiently for

light-tree based optical multicasting. X. Huang et al. has conducted a study on dynamic traffic grooming with sparse grooming capability for light-tree based optical multicasting [108]. However, compared to our proposed GGM, the graph model presented in [108] does not consider the wavelength conversion capability. There are more practical issues and constraints that need to be taken considerations for the study of dynamic traffic grooming for optical multicasting with both wavelength conversion and light splitting capability.

- **Optimization algorithm for placement of wavelength converters and light splitters based on generic graph model** – Our proposed generic graph model provides a generic platform to study dynamic MC-RWA problem for different light splitting and wavelength conversion scenarios. However, for a network with a large N , L , and W , the computational time of the proposed algorithms will increase dramatically. Novel optimization algorithms that are computationally efficient are required. Moreover, since the assumed value of κ_c , κ_{sp} and κ_{sw} for the simulation study will significantly impact the final results, it is interesting to further investigate the optimal configurations for different values of κ_c , κ_{sp} and κ_{sw} .

Author's Publications

Journal Papers

1. S. Chen and G.-S. Poo, "Study of signaling effects on Dynamic Traffic Grooming in IP/MPLS over WDM network", *Computer Communications*, vol. 30, no. 18, 2007, pp. 3586-3597.
2. Y. Zhou, G.-S. Poo, S. Chen, P. Shum, and L. Zhang, "Dynamic multicast routing and wavelength assignment using generic graph model for wavelength-division-multiplexing networks", *IET Communications*, vol. 2, no. 7, 2008, pp. 951-959.
3. S. Chen, T. H. Cheng, and G.-S. Poo, "Placement of Wavelength Converter and Light Splitter in WDM Network by Generic Graph Model", *Computer Communications*, vol. 33, no. 7, 2010, pp. 868-883.

Conference Papers

1. S. Chen and G.-S. Poo, "Dynamic generation of activity plan for policy-based management", in *Proc. of 12th IEEE International Conference on Networks*, Singapore, 2004, pp. 114-118.
2. S. Chen and G.-S. Poo, "Comparison of protocols for dynamic lightpath provisioning in GMPLS optical network", in *Proc. of NTRC Research Forum*, Singapore, 2004, pp. 26-34.
3. S. Chen and G.-S. Poo, "Performance Study of Control and Management Protocols for Dynamic Lightpath Provisioning", in *Proc. of 14th IEEE International Conference on Networks*, vol. 2, Singapore, 2006, pp. 1-6.

Appendix

Network Topologies

Figure A.1 Topology of Simple network with 6 nodes, 12 links. Each link in the network is composed of two fibers of opposite directions, and has the same physical lengths in 100 kilometers

Figure A.2 Topology of Mesh-torus network with 9 nodes, 18 links. Each link in the network is composed of two fibers of opposite directions, and has the same physical lengths in 100 kilometers

Figure A.3 Topology of New Jersey Local Access Transport Area network (NJ LATA) network with 11 nodes, 23 links. Each link in the network is composed of two fibers of opposite directions. The numbers next to the links denote the physical lengths in 1 kilometer. In simulation of Chapter 3, centralized global information-based link state approach, the central manager is located at node 4.

Figure A.4 Topology of COST239 network with 11 nodes, 26 links. Each link in the network is composed of two fibers of opposite directions.

Figure A.5 Topology of NSFNet network with 14 nodes, 21 links. Each link in the network is composed of two fibers of opposite directions. The numbers next to the links denote the physical lengths in 100 kilometer.

Bibliography

- [1] B. Mukherjee, *Optical WDM networks*: Springer, 2006.
- [2] R. Ramaswami and K. N. Sivarajan, *Optical networks : A practical perspective*, 2nd ed.: Morgan Kaufmann Publishers, 2002.
- [3] T. E. Stern and K. Bala, *Multiwavelength optical networks: A layered approach*: Addison-Wesley, 1999.
- [4] P. Green, "Progress in optical networking," *IEEE Communication Magazine*, vol. 39, no. 1, 2001, pp. 54-61.
- [5] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: an approach to high bandwidth optical WAN's," *IEEE Transactions on Communications*, vol. 40, no. 7, 1992, pp. 1171-1182.
- [6] R. Ramaswami and K. N. Sivarajan, "Routing and wavelength assignment in all-optical networks," *IEEE/ACM Transactions on Networking*, vol. 3, no. 5, 1995, pp. 489-500.
- [7] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," *Optical Networks Magazine*, vol. 1, no. 1, 2000, pp. 47-60.
- [8] H. Zang, J. P. Jue, L. Sahasrabudhe *et al.*, "Dynamic lightpath establishment in wavelength-routed WDM networks," *IEEE Communication Magazine*, vol. 39, no. 9, 2001, pp. 100-108.
- [9] X. Yuan, R. Melhem, R. Gupta *et al.*, "Distributed control protocols for wavelength reservation and their performance evaluation," *Photonic Network Communications*, vol. 1, no. 3, 1999, pp. 207-218.
- [10] E. Mannie, *Generalized multi-protocol label switching (GMPLS) architecture*, IETF RFC 3945, 2004.
- [11] E. Rosen, A. Viswanathan, and R. Callon, *Multiprotocol label switching architecture*, IETF RFC 3031, 2001.
- [12] A. Banerjee, J. Drake, J. P. Lang *et al.*, "Generalized multiprotocol label switching: An overview of routing and management enhancements," *IEEE Communication Magazine*, vol. 39, no. 1, 2001, pp. 144-150.
- [13] R. Dutta and G. N. Rouskas, "Traffic grooming in WDM networks: Past and

- future,” *IEEE Network*, vol. 16, no. 6, 2002, pp. 46-56.
- [14] “Special issue on network support for multipoint communication,” *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 3, 1997, pp. 273-588.
- [15] C. K. Miller, *Multicast networking and applications*: Addison-Wesley, 1999.
- [16] S. Paul, *Multicasting on the Internet and its applications*: Kluwer Academic, 1998.
- [17] L. H. Sahasrabuddhe and B. Mukherjee, “Light-trees: Optical multicasting for improved performance in wavelength-routed networks,” *IEEE Communication Magazine*, vol. 37, no. 2, 1999, pp. 67-73.
- [18] G. N. Rouskas, “Optical layer multicast: Rationale, Building blocks, and Challenges,” *IEEE Network*, vol. 17, no. 1, 2003, pp. 60-65.
- [19] R. Malli, X. Zhang, and C. Qiao, “Benefit of multicasting in all-optical networks,” in *Proc. of SPIE All-Optical Networking: Architecture, Control, and Management Issues*, 1998, pp. 209-220.
- [20] Y. Mei and C. Qiao, “Efficient distributed control protocols for WDM all-optical networks,” in *Proc. of IEEE ICCCN 6th*, 1997, pp. 150-153.
- [21] R. Ramaswami and A. Segall, “Distributed network control for optical networks,” *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, 1997, pp. 936-943.
- [22] K. Lu, G. Xiao, and I. Chlamtac, “Analysis of blocking probability for distributed lightpath establishment in WDM optical networks,” *IEEE/ACM Transactions on Networking*, vol. 13, no. 1, 2005, pp. 187-197.
- [23] C. Xin, Y. Ye, S. Dixit *et al.*, “An emulation-based comparative study of centralized and distributed control schemes for optical networks,” *Optical Networks Magazine*, vol. 3, no. 5, 2002, pp. 44-51.
- [24] A. L. Chiu and E. H. Modiano, “Traffic grooming algorithms for reducing electronic multiplexing costs in WDM ring networks,” *Journal of Lightwave Technology*, vol. 18, no. 1, 2000, pp. 2-12.
- [25] P.-J. Wan, G. Calinescu, L. Liu *et al.*, “Grooming of arbitrary traffic in SONET/WDM BLSRs,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, 2000, pp. 1995-2003.
- [26] X. Zhang and C. Qiao, “On scheduling all-to-all personalized connections

- and cost-effective designs in WDM rings,” *IEEE/ACM Transactions on Networking*, vol. 7, no. 3, 1999, pp. 435-445.
- [27] K. Zhu and B. Mukherjee, “Traffic grooming in an optical WDM mesh network,” *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, 2002, pp. 122-133.
 - [28] B. Chen, W.-D. Zhong, and S. K. Bose, “A path inflation control strategy for dynamic traffic grooming in IP/MPLS over WDM network,” *IEEE Communication Letters*, vol. 8, no. 11, 2004, pp. 680-682.
 - [29] S. Huang, M. Bo, J. Zhang *et al.*, “Dynamic traffic grooming with adaptive routing in optical WDM mesh networks,” in *Proc. of SPIE Network Architectures, Management, and Applications III*, 2005, pp. 60222Y-(1-8).
 - [30] H. Zhu, H. Zang, K. Zhu *et al.*, “A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks,” *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, 2003, pp. 285-299.
 - [31] K. Zhu, H. Zang, and B. Mukherjee, “A comprehensive study on next-generation optical grooming switches,” *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 7, 2003, pp. 1173-1186.
 - [32] K. Zhu, H. Zhu, and B. Mukherjee, “Traffic engineering in multigranularity heterogeneous optical WDM mesh networks through dynamic traffic grooming,” *IEEE Network*, vol. 17, no. 2, 2003, pp. 8-15.
 - [33] J. M. H. Elmirghani and H. T. Mouftah, “All-optical wavelength conversion: Technologies and applications in DWDM networks,” *IEEE Communication Magazine*, vol. 38, no. 3, 2000, pp. 86-92.
 - [34] C. S. R. Murthy and M. Gurusamy, *WDM optical networks : Concepts, design, and algorithms*: Prentice Hall, 2002.
 - [35] C. Assi, A. A. Shami, M. A. Ali *et al.*, “Optical networking and real-time provisioning: An integrated vision for the next-generation Internet,” *IEEE Network*, vol. 15, no. 4, 2001, pp. 36-45.
 - [36] J. Y. Wei, “Advances in the management and control of optical Internet,” *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 4, 2002, pp. 768-785.
 - [37] W.-D. Zhong, X. Niu, B. Chen *et al.*, “Performance comparison of overlay and peer models in IP/MPLS over optical networks,” *Photonic Network Communications*, vol. 9, no. 1, 2005, pp. 121-131.

-
- [38] D. Papadimitriou, D. Ooms, J. Jones *et al.*, *Optical multicast in wavelength switched networks: Architectural Framework*, draft-poj-optical-multicast-02.txt, IETF IPO Working Group, 2001.
- [39] X. Cao, V. Anand, and C. Qiao, "Waveband switching in optical networks," *IEEE Communication Magazine*, vol. 41, no. 4, 2003, pp. 105-112.
- [40] E. Dotaro, D. Papadimitriou, L. Noirie *et al.*, *Optical multi-granularity & architectural framework*, draft-dotaro-ipo-multi-granularity-00.txt, IETF IPO working group, 2001.
- [41] T. S. El-Bawab and J.-D. Shin, "Optical packet switching in core networks: Between vision and reality," *IEEE Communication Magazine*, vol. 40, no. 9, 2002, pp. 60-65.
- [42] C. Xin, Y. Ye, T.-S. Wang *et al.*, "On an IP-centric optical control plane," *IEEE Communication Magazine*, vol. 39, no. 9, 2001, pp. 88-93.
- [43] G. M. Bernstein, J. Yates, and D. Saha, "IP-centric control and management of optical transport networks," *IEEE Communication Magazine*, vol. 38, no. 10, 2000, pp. 2-8.
- [44] S. Dixit, *IP over WDM : building the next-generation optical Internet*: Wiley-Interscience, 2003.
- [45] X. Niu, W.-D. Zhong, G. Shen *et al.*, "Connection establishment of label switched paths in IP/MPLS over optical networks," *Photonic Network Communications*, vol. 6, no. 1, 2003, pp. 33-41.
- [46] B. Ramamurthy and B. Mukherjee, "Wavelength conversion in WDM networking," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 7, 1998, pp. 1061-1073.
- [47] R. Ramaswami and G. Sasaki, "Multiwavelength optical networks with limited wavelength conversion," *IEEE/ACM Transactions on Networking*, vol. 6, no. 6, 1998, pp. 744-754.
- [48] S. Subramaniam, M. Azizoglu, and A. K. Somani, "All-optical networks with sparse wavelength conversion," *IEEE/ACM Transactions on Networking*, vol. 4, no. 4, 1996, pp. 544-557.
- [49] Y. Zhu, G. N. Rouskas, and H. G. Perros, "A path decomposition approach for computing blocking probabilities in wavelength-routing networks," *IEEE/ACM Transactions on Networking*, vol. 8, no. 6, 2000, pp. 747-762.
- [50] Y. Zhou and G.-S. Poo, "Optical multicast over wavelength-routed WDM

- networks: A survey,” *Elsevier Journal on Optical Switching and Networking*, vol. 2, no. 3, 2005, pp. 176-197.
- [51] F. K. Hwang, D. S. Richards, and P. Winter, *The Steiner Tree Problem*: Elsevier Science, 1992.
 - [52] Y. Zhou and G.-S. Poo, “Multicast wavelength assignment for sparse wavelength conversion in WDM networks,” in *Proc. of IEEE INFOCOM 25th*, Barcelona, Spain, 2006, pp. 1-10.
 - [53] Y. Zhou and G.-S. Poo, “Multi-wavelength multicast wavelength assignment algorithm for limited wavelength conversion in wavelength-division multiplexing networks,” *IET Communications*, vol. 1, no. 4, 2007, pp. 776-783.
 - [54] Y. Zhou, “Wavelength assignment algorithms for optical multicast in WDM networks,” Thesis of Ph.D, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 2007.
 - [55] D. Banerjee and B. Mukherjee, “Wavelength-routed optical networks: Linear formulation, resource budgeting tradeoffs, and a reconfiguration study,” *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, 2000, pp. 598-607.
 - [56] R. M. Krishnaswamy and K. N. Sivarajan, “Design of logical topologies: A linear formulation for wavelength-routed optical networks with no wavelength changers,” *IEEE/ACM Transactions on Networking*, vol. 9, no. 2, 2001, pp. 186-201.
 - [57] A. Sridharan and K. N. Sivarajan, “Blocking in all-optical networks,” *IEEE/ACM Transactions on Networking*, vol. 12, no. 2, 2004, pp. 384-397.
 - [58] C. Assi, A. A. Shami, M. A. Ali *et al.*, “Integrated routing algorithms for provisioning “sub-wavelength” connections in IP over WDM networks,” *Photonic Network Communications*, vol. 4, no. 3, 2002, pp. 377-390.
 - [59] A. A. Shami, C. Assi, M. A. Ali *et al.*, “Integrated routing/signaling protocols in GMPLS-based optical networks,” *Optical Networks Magazine*, vol. 3, no. 4, 2002, pp. 16-23.
 - [60] A. A. Shami, C. Assi, I. Habib *et al.*, “Connection management protocols for dynamic lightpath provisioning in future WDM networks,” *Photonic Network Communications*, vol. 6, no. 1, 2003, pp. 25-32.
 - [61] S. Sengupta and R. Ramamurthy, “From network design to dynamic provisioning and restoration in optical cross-connect mesh networks: an architectural and algorithmic overview,” *IEEE Network*, vol. 15, no. 4, 2001,

pp. 46-54.

- [62] H. Zang, L. Sahasrabudhe, J. P. Jue *et al.*, "Connection management for wavelength-routed WDM networks," in *Proc. of IEEE Global Telecommunications*, 1999, pp. 1428-1432.
- [63] M. Goyal, G. Li, and J. Yates, "Shared mesh restoration: a simulation study," in *Proc. of Optical Fiber Communication (OFC)*, 2002, pp. 489-490.
- [64] H. M. Alnuweiri, L.-Y. K. Wong, and T. Al-Khasib, "Performance of new link state advertisement mechanisms in routing protocols with traffic engineering extensions," *IEEE Communication Magazine*, vol. 42, no. 5, 2004, pp. 151-162.
- [65] J. Li, G. Mohan, E. C. Tien *et al.*, "Dynamic routing with inaccurate link state information in integrated IP-over-WDM networks," *Computer Networks*, vol. 46, 2004, pp. 829-851.
- [66] K. Lu, J. P. Jue, G. Xiao *et al.*, "Intermediate-node initiated reservation (IIR): a new signaling scheme for wavelength-routed networks," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 8, 2003, pp. 1285-1294.
- [67] G. Xiao, K. Lu, and I. Chlamtac, "An evaluation of distributed wavelength provisioning in WDM optical networks with sparse wavelength conversion," *Journal of Lightwave Technology*, vol. 22, no. 7, 2004, pp. 1668-1678.
- [68] B. Chen, S. K. Bose, and W.-D. Zhong, "Priority enabled dynamic traffic grooming," *IEEE Communication Letters*, vol. 9, no. 4, 2005, pp. 366-368.
- [69] B. Chen, W.-D. Zhong, and S. K. Bose, "Applying saturated cut method for dynamic traffic grooming in IP/MPLS over WDM networks," *IEEE Communication Letters*, vol. 10, no. 2, 2006, pp. 117-119.
- [70] F. Farahmand, X. Huang, and J. P. Jue, "Efficient online traffic grooming algorithms in WDM mesh networks with drop-and-continue node architecture," in *Proc. of International Conference on Broadband Networks 1st*, 2004, pp. 180-189.
- [71] J. Huang, Q. Zeng, J. Liu *et al.*, "Dynamic traffic grooming with multi-granularity traffic in WDM optical mesh networks," in *Proc. of SPIE Network Architectures, Management, and Applications*, 2003, pp. 988-993.
- [72] E. Modiano and P. J. Lin, "Traffic grooming in WDM networks," *IEEE Communication Magazine*, vol. 39, no. 7, 2001, pp. 124-139.
- [73] J. Wang, W. Cho, V. R. Vemuri *et al.*, "Improved approaches for

- cost-effective traffic grooming in WDM ring networks: ILP formulations and single-hop and multihop connections,” *Journal of Lightwave Technology*, vol. 19, no. 11, 2001, pp. 1645-1652.
- [74] C. Xin, C. Qiao, and S. Dixit, “Traffic grooming in mesh WDM optical networks - performance analysis,” *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, 2004, pp. 1658-1669.
- [75] T. Ye, Q. Zeng, G. Zhu *et al.*, “Dynamic traffic grooming using hop constraint lightpath establishment approach,” in *Proc. of SPIE Network Architectures, Management, and Applications II*, 2005, pp. 1325-1332.
- [76] Y. Zhu and R. Lin, “Dynamic traffic grooming in survivable WDM networks,” in *Proc. of SPIE Optical Transmission, Switching, and Subsystems III*, 2005, pp. 60212C-(1-9).
- [77] A. Gencata and B. Mukherjee, “Virtual-topology adaptation for WDM mesh networks under dynamic traffic,” *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, 2003, pp. 236-247.
- [78] B. Mukherjee, R. Dutta, A. E. Kamal *et al.*, “Traffic Grooming for Optical Networks Foundations, Techniques, and Frontiers,” *Optical Networks*, Springer Science+Business Media, LLC, 2008.
- [79] X. Chu, B. Li, and I. Chlamtac, “Wavelength converter placement under different RWA algorithms in wavelength-routed all-optical networks,” *IEEE Transactions on Communications*, vol. 51, no. 4, 2003, pp. 607-617.
- [80] S. Sankaranarayanan and S. Subramaniam, “Comprehensive performance modeling and analysis of multicasting in optical networks,” *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 9, 2003, pp. 1399-1413.
- [81] M. Ali, “Optimization of splitting node placement in wavelength-routed optical networks,” *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, 2002, pp. 1571-1579.
- [82] S. Subramaniam, M. Azizoglu, and A. K. Somani, “On optimal converter placement in wavelength-routed networks,” *IEEE/ACM Transactions on Networking*, vol. 7, no. 5, 1999, pp. 754-766.
- [83] B. Li, X. Chu, and K. Sohraby, “Routing and wavelength assignment vs. wavelength converter placement in all-optical networks,” *IEEE Communication Magazine*, vol. 41, no. 8, 2003, pp. S22-S28.
- [84] S. Gao, X. Jia, C. Huang *et al.*, “An optimization model for placement of wavelength converters to minimize blocking probability in WDM

- networks,” *Journal of Lightwave Technology*, vol. 21, no. 3, 2003, pp. 684-693.
- [85] M. E. Houmaidi and M. A. Bassiouni, “Dependency-based analytical model for computing connection blocking rates and its application in the sparse placement of optical converters,” *IEEE Transactions on Communications*, vol. 54, no. 1, 2006, pp. 159-168.
- [86] N. K. Singhal, L. H. Sahasrabudde, and B. Mukherjee, “Optimal multicasting of multiple light-trees of different bandwidth granularities in a WDM mesh network with sparse splitting capabilities,” *IEEE/ACM Transactions on Networking*, vol. 14, no. 5, 2006, pp. 1104-1117.
- [87] H.-C. Lin and S.-W. Wang, “Splitter placement in all-optical WDM networks,” in *Proc. of IEEE Global Telecommunications*, 2005, pp. 306-310.
- [88] Y. Cao and O. Yu, “Optimal placement of light splitters and wavelength converters for multicast in WDM networks,” in *Proc. of International Conference on Communications, Circuits and Systems*, 2005, pp. 585-589.
- [89] K.-M. Yong, G.-S. Poo, and T. H. Cheng, “Optimal placement of multicast and wavelength converting nodes in multicast optical virtual private network,” *Computer Communications*, vol. 29, no. 12, 2006, pp. 2169-2180.
- [90] J. Moy, *OSPF Version 2*, IETF RFC 2328, 1998.
- [91] J. P. Jue and G. Xiao, “An adaptive routing algorithm for wavelength-routed optical networks with a distributed control scheme,” in *Proc. of IEEE ICCCN 9th*, 2000, pp. 192-197.
- [92] M. Medard, S. Lumetta, and L. Li, “A network management architecture for robust packet routing in mesh optical access networks,” *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 4, 2002, pp. 822-833.
- [93] V. Paxson and S. Floyd, “Wide area traffic: the failure of poisson modeling,” *IEEE/ACM Transactions on Networking*, vol. 3, no. 3, 1995, pp. 226-244.
- [94] D. Wang, G. Li, J. Yates *et al.*, *Optical NNI inter-working demonstration*, IEEE/LEOS Summer Topical Meeting, July 2002.
- [95] M. Goyal, J. Yates, G. Li *et al.*, “Benefits of restoration signaling message aggregation,” in *Proc. of Optical Fiber Communication (OFC)*, 2003, pp. 203-204.
- [96] D. O. Awduche, “MPLS and traffic engineering in IP networks,” *IEEE Communication Magazine*, vol. 37, no. 12, 1999, pp. 42-47.

- [97] S. Shen, "Link-state update in wavelength-routed networks," Thesis of Ph.D, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 2005.
- [98] W. D. Grover, "The protected working capacity envelop concept: An alternate paradigm for automated service provisioning," *IEEE Communication Magazine*, vol. 42, no. 1, 2004, pp. 62-69.
- [99] W. Willinger, M. S. Taqqu, R. Sherman *et al.*, "Self similarity through high-variability: Statistical analysis of Ethernet LAN traffic at the source level," *IEEE/ACM Transactions on Networking*, vol. 5, no. 1, 1997, pp. 71-86.
- [100] W. S. Hu and Q. J. Zeng, "Multicasting optical cross connects employing splitter-and-delivery switch," *IEEE Photonics Technology Letters*, vol. 10, no. 7, 1998, pp. 970-972.
- [101] J. Leuthold and C. H. Joyner, "Multimode interference couplers with tunable power splitting ratios," *Journal of Lightwave Technology*, vol. 19, no. 5, 2001, pp. 700-707.
- [102] J. M. Yates, M. P. Rumsewicz, and J. P. R. Lacey, "Wavelength converters in dynamically-reconfigurable WDM networks," *IEEE Communications Surveys & Tutorials*, vol. 2, no. 2, 1999, pp. 2-15.
- [103] M. Ali and J. S. Deogun, "Power-efficient design of multicast wavelength-routed networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, 2000, pp. 1852-1862.
- [104] M.-T. Chen, S.-S. Tseng, and B. M. T. Lin, "Dynamic multicast routing under delay constraints in WDM networks with heterogeneous light splitting capabilities," *Computer Communications*, vol. 29, no. 5, 2006, pp. 1492-1503.
- [105] X. Zhang, J. Y. Wei, and C. Qiao, "Constrained multicast routing in WDM networks with sparse light splitting," *Journal of Lightwave Technology*, vol. 18, no. 12, 2000, pp. 1917-1927.
- [106] B. Chen and J. Wang, "Efficient routing and wavelength assignment for multicast in WDM networks," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, 2002, pp. 97-109.
- [107] G.-S. Poo and Y. Zhou, "A new multicast wavelength assignment algorithm in wavelength-routed WDM networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 4, 2006, pp. 2-12.
- [108] X. Huang, F. Farahmand, and J. P. Jue, "Multicast traffic grooming in

- wavelength-routed WDM mesh networks using dynamically changing light-trees,” *Journal of Lightwave Technology*, vol. 23, no. 10, 2005, pp. 3178-3187.
- [109] A. K. Somani, M. Mina, and L. Li, “On trading wavelengths with fibers: A cost-performance based study,” *IEEE/ACM Transactions on Networking*, vol. 12, no. 5, 2004, pp. 944-951.
- [110] Y. Zhou, G.-S. Poo, S. Chen *et al.*, “Dynamic multicast routing and wavelength assignment using generic graph model for wavelength-division-multiplexing networks,” *IET Communications*, vol. 2, no. 7, 2008, pp. 951-959.