Traffic Grooming in Optical WDM Mesh Networks

Keyao Zhu
Hongyue Zhu
Biswanath Mukherjee

OPTICAL NETWORKS SERIES

Traffic Grooming in Optical WDM Mesh Networks

OPTICAL NETWORKS SERIES

Series Editor
Biswanath Mukherjee, *University of California, Davis*

Other books in the series:

SURVIVABLE OPTICAL WDM NETWORKS
Canhui (Sam) Ou and Biswanath Mukherjee, ISBN 0-387-24498-0
OPTICAL BURST SWITCHED NETWORKS
Jason P. Jue and Vinod M. Vokkarane, ISBN 0-387-23756-9

TRAFFIC GROOMING IN OPTICAL WDM MESH NETWORKS

KEYAO ZHU
Brion Technologies

HONGYUE ZHU
University of California, Davis

BISWANATH MUKHERJEE
University of California, Davis

Springer

Keyao Zhu
Brion Technologies, Inc.

Hongyue Zhu
University of California, Davis

Biswanath Mukherjee
University of California, Davis

TRAFFIC GROOMING IN OPTICAL WDM MESH NETWORKS

ISBN 0-387-25432-3 e-ISBN 0-387-27098-1 Printed on acid-free paper.
ISBN 978-0387-25432-6

Printed in the United States of America.

9 8 7 6 5 4 3 2 1 SPIN 11328056

springeronline.com

To our families and friends

Contents

Dedication v
List of Figures xiii
List of Tables xvii
Preface xix
Acknowledgments xxiii

1. OVERVIEW 1
   1.1 Background 1
   1.2 Traffic Grooming in SONET Ring Network 2
      1.2.1 Node Architecture 2
      1.2.2 Single-Hop Grooming in SONET/WDM Ring 4
      1.2.3 Multi-Hop Grooming in SONET/WDM Ring 5
      1.2.4 Dynamic Grooming in SONET/WDM Ring 6
      1.2.5 Grooming in Interconnected SONET/WDM Rings 8
   1.3 Traffic Grooming In Wavelength-Routed WDM Mesh Network 9
      1.3.1 Network Provisioning 10
      1.3.2 Network Design and Planner 12
      1.3.3 Grooming with Protection Requirement in WDM Mesh Network 13
      1.3.4 Grooming with Multicast in WDM Mesh Network 15
      1.3.5 Protocols and Algorithm Extensions for WDM Network Control 16

2. STATIC TRAFFIC GROOMING 17
   2.1 Introduction 17
   2.2 General Problem Statement 19
   2.3 Node Architecture 20
   2.4 Mathematical (ILP) Formulation 22

2.4.1 Multi-Hop Traffic Grooming 23
2.4.2 Single-Hop Traffic Grooming 28
2.4.3 Formulation Extension for Fixed-Transceiver Array 28
2.4.4 Computational Complexity 29
2.5 Illustrative Numerical Results From ILP Formulations 29
2.6 Heuristic Approach 33
2.6.1 Routing 33
2.6.2 Wavelength Assignment 34
2.6.3 Heuristics 35
2.6.4 Heuristic Results and Comparison 36
2.7 Mathematical Formulation Extension 39
2.7.1 Extension for Network Revenue Model 39
2.7.2 Illustrative Results 40
2.8 Conclusion 41
3. A GENERIC GRAPH MODEL 43
3.1 Introduction 43
3.1.1 Challenges of Traffic Grooming in a Heterogeneous WDM Mesh Network 44
3.1.2 Contributions of this Chapter 45
3.2 Construction of an Auxiliary Graph 46
3.3 Solving the Traffic-Grooming Problem Based on the Auxiliary Graph 50
3.3.1 The IGABAG Algorithm 51
3.3.2 The INGPROC Procedure and Traffic-Selection Schemes 51
3.3.3 An Illustrative Example 54
3.4 Grooming Policies and Weight Assignment 57
3.4.1 Grooming Policies 57
3.4.2 Weight Assignment 58
3.5 Numerical Examples 61
3.5.1 Comparison of Grooming Policies 61
3.5.2 Comparison of Traffic-Selection Schemes in a Relatively Small Network 64
3.5.3 Comparison in a Larger Representative Network 66
3.6 Conclusion 68
4. DYNAMIC TRAFFIC GROOMING 71
4.1 Introduction 71

4.1.1 Traffic Engineering In Optical WDM Networks Through Traffic Grooming 71
4.1.2 Optical WDM Network Heterogeneity 72
4.1.3 Organization 72
4.2 Node Architecture in a Heterogeneous WDM Backbone Network 73
4.3 Provisioning Connections in Heterogeneous WDM Network 75
4.3.1 Resource Discovery 76
4.3.2 Route Computation 78
4.3.3 Signaling 80
4.4 A Generic Provisioning Model 80
4.4.1 Graph Model 80
4.4.2 Engineering Network Traffic Using the Proposed Graph Model 83
4.4.3 Computational Complexity 84
4.5 Illustrative Numerical Examples 85
4.5.1 Comparison of Grooming Policies 85
4.5.2 Performance under Different Scenarios 87
4.6 Conclusion 92
5. GROOMING SWITCH ARCHITECTURES 93
5.1 Introduction 93
5.2 Grooming Switch Architectures and Grooming Schemes 94
5.2.1 Single-Hop Grooming OXC 94
5.2.2 Multi-Hop Partial-Grooming OXC 95
5.2.3 Multi-Hop Full-Grooming OXC 98
5.2.4 Light-tree-Based Source-Node Grooming OXC 99
5.2.5 Summary 100
5.3 Approaches and Algorithms 101
5.3.1 Single-Hop and Multi-Hop Grooming using an Auxiliary Graph Model 101
5.3.2 Source-Node Grooming Using Light-Tree Approach 103
5.4 Illustrative Numerical Results 105
5.4.1 Bandwidth Blocking Ratio (BBR) 106
5.4.2 Wavelength Utilization 110
5.4.3 Resource Efficiency Ratio (RER) 110
5.5 Conclusion 113

6. SPARSE GROOMING NETWORK 115
6.1 Problem Statement and Mathematical Formulation 116
6.1.1 Maximizing Network Throughput 118
6.1.2 Minimizing Network Cost 119
6.2 Heuristic Approaches 119
6.2.1 Grooming-Node-Selection Schemes 120
6.2.2 Traffic-Routing Schemes 120
6.3 Illustrative Numerical Examples 121
6.4 Conclusion 124

7. NETWORK DESIGN WITH OXCS OF DIFFERENT BANDWIDTH GRANULARITIES 125
7.1 Introduction 125
7.2 Problem Statement and Challenges 126
7.2.1 Problem Formulation 126
7.2.2 Challenges 127
7.2.3 Our Approach 130
7.3 Construction of an Auxiliary Graph 130
7.3.1 Node Representation 130
7.3.2 Circuits and Induced Topology 136
7.3.3 Auxiliary Graph for the Network 138
7.4 Framework for Network Design Based on the Auxiliary Graph 140
7.4.1 Algorithm for Routing a Connection Request 140
7.4.2 An Illustrative Example 142
7.4.3 Weight Assignment 146
7.4.4 Network Design Framework 148
7.5 Numerical Examples and Discussion 150
7.6 Conclusion 153

8. TRAFFIC GROOMING IN NEXT-GENERATION SONET/SDH 155
8.1 Virtual Concatenation 155
8.1.1 SONET Virtual Concatenation 156
8.1.2 Benefits of Virtual Concatenation: a Network Perspective 156
8.1.3 Illustrative Numerical Examples 158
8.2 Inverse Multiplexing 160
8.2.1 Problem Statement and Proposed Approaches 161
8.2.2 Illustrative Numerical Results 163
8.3 Conclusion 165

References 167

Index 173

List of Figures

1.1	Node architectures in a SONET/WDM ring network.	3
1.2	A SONET/WDM network with 4 nodes and 2 wavelengths.	5
1.3	Two possible configurations to support the traffic requests in Fig. 1.2.	5
1.4	SONET/WDM ring with/without a hub node.	6
1.5	Network design for 2-allowable traffic.	7
1.6	A sample interconnected-ring network topology and simplified architectures of the junction node.	9
1.7	An OXC with a two-level hierarchy and grooming capability.	11
1.8	Two different designs for a 4-node network [Cox and Sanchez, 2001].	13
1.9	A multi-layer protection example [Lardies et al., 2001].	14
1.10	Switch architecture for supporting multicast grooming [Sahasrabuddhe and Mukherjee, 1999].	16
2.1	Illustrative example of traffic grooming.	18
2.2	Node architecture 1: IP over WDM.	21
2.3	Node architecture 2: SONET over WDM.	22
2.4	Illustrative example of a fiber link, a lightpath, and a connection request.	24
2.5	(a) A 6-node network and (b) a 15-node network.	29
2.6	Network throughput vs. number of wavelengths for the network topology in Fig. 2.5(b) with 10 tunable transceivers at each node.	37
2.7	Network throughput vs. number of tunable transceivers for the network topology in Fig. 2.5(b) with 10 wavelengths on each fiber link.	38

2.8 Network throughput vs. number of wavelengths (size of fixed transceiver array) for the network topology in Fig. 2.5(b) with 12 tunable transceivers at each node. 38

3.1 (a) Physical topology of Network 1. (b) Virtual topology of Network 1. (c) Auxiliary graph of Network 1. 47

3.2 (a) Virtual topology of Network 1. (b) Corresponding auxiliary graph before routing the first traffic request T_1. (c) Corresponding auxiliary graph after routing the first traffic request T_1. 54

3.3 (a) Virtual topology of Network 1. (b) Corresponding auxiliary graph before routing the second traffic request T_2. (c) Corresponding auxiliary graph after routing the second traffic request T_2 using single-hop grooming. 55

3.4 (a) Virtual topology of Network 1. (b) Corresponding auxiliary graph before routing the second traffic request T_2. (b) Corresponding auxiliary graph after routing the second traffic request T_2 using multi-hop grooming. 56

3.5 Comparison of different grooming policies. (a) NSF network. (b) Comparison of different grooming policies using a non-blocking model. (c) Comparison of different grooming policies using a blocking model. 64

3.6 Comparison of traffic-selection schemes in a relatively small network. (a) Network 2: a 6-node network. (b) Average ratio of the amount of carried traffic by LCF to the amount of carried traffic by ILP. 65

3.7 Comparison of traffic-selection schemes in a larger representative network. (a) Network 3: a 19-node network. (b) Network throughput using different heuristics when each link has 8 wavelengths. (c) Network throughput using different heuristics when each link has 16 wavelengths. (d) Network throughput using heuristic LCF under different network configurations. 67

4.1 A multi-hop partial-grooming OXC. 75

4.2 Network state for a simple three-node network and the corresponding auxiliary graph. 81

4.3 Different grooming OXCs and their representations in the auxiliary graph. 84

4.4 Two alternative routes for a new connection request $(1, 2)$. 85

4.5 Percentage of blocked traffic when $Tx = 32$. 87

4.6 Percentage of blocked traffic when $Tx = 40$. 87

4.7	Performance of AGP when $Tx = 32$.	88
4.8	Performance of AGP when $Tx = 40$.	88
4.9	Sample network topology with 5 grooming nodes.	89
4.10	Traffic blocking ratio vs. offered load.	90
4.11	Normalized resource-efficiency ratio vs. offered load.	91
4.12	Connection blocking probability vs. offered load.	91
5.1	Examples of single-hop, multi-hop, and source-node grooming schemes.	96
5.2	Sample grooming OXC architectures: a multi-hop partial-grooming OXC and a source-node grooming OXC.	97
5.3	An overview of Time-Space-Time (TST) switch architecture.	99
5.4	A 24-node sample network topology.	107
5.5	Bandwidth blocking ratio (BBR) vs. load (in Erlangs) for different grooming OXCs under different bandwidth-granularity distributions.	108
5.6	Effect of different lightpath-establishment schemes and different number of grooming ports on the network performance of multi-hop partial-grooming OXCs.	109
5.7	Wavelength utilization (WU) vs. load (in Erlangs) for different grooming OXCs under different bandwidth-granularity distributions.	111
5.8	Normalized resource-efficiency ratio (RER) vs. load (in Erlangs) for different grooming OXCs under different bandwidth granularity distributions.	112
6.1	A sample network and two sparse-grooming network designs.	116
6.2	A sample sparse-grooming WDM network which carries two requests using four lightpaths.	117
6.3	Illustrative results from ILP formulation for the network in Fig. 6.1 assuming only one node has grooming capability.	121
6.4	Performance comparison between different G-Node selection schemes applied to the network in Fig. 5.4.	122
6.5	Network cost vs. network resources based on different cost ratio R.	123
7.1	State of the switches when routing traffic demand T_1 of bandwidth STS-1 from node 1 to node 4.	129
7.2	Network state after routing T_1 traffic demand.	129

7.3 A node with three different types of OXCs. (Each data path could be a multi-line, i.e., there may be multiple fibers in and out of the OC-192 OXC, multiple add and drop ports for each OXC, etc.) 131
7.4 Auxiliary graph for the node. 132
7.5 Initial auxiliary graph. 143
7.6 Corresponding auxiliary graph before routing the first traffic request T_1. 143
7.7 Corresponding auxiliary graph after routing the first traffic request T_1. 144
7.8 Corresponding auxiliary graph before routing the second traffic request T_2. 145
7.9 Corresponding auxiliary graph after routing the second traffic request T_2. 146
7.10 A 26-node WDM backbone network. 150
7.11 Comparison of total port cost in the four scenarios. 152
7.12 Comparison of number of transponders and wavelength-links used in the four scenarios. 152
7.13 Comparison of the lightpath utilization in the four scenarios. 153
8.1 An example of using VCAT to support different network services [Stanley, 2002]. 157
8.2 Illustrative results - Traffic pattern I. 159
8.3 Illustrative results - Traffic pattern II. 160
8.4 An illustrative example of inverse multiplexing in a SONET/SDH-based optical transport network. 162
8.5 Results for $K = 4$ paths. 164
8.6 . Performance results for different values of K based on MF algorithm. 164

List of Tables

2.1	Traffic matrix of OC-1 connection requests.	30
2.2	Traffic matrix of OC-3 connection requests.	30
2.3	Traffic matrix of OC-12 connection requests.	30
2.4	Throughput and number of lightpaths established (total traffic demand is OC-988).	31
2.5	Results: transceiver utilization (multi-hop case).	32
2.6	Results: wavelength utilization (multi-hop case).	32
2.7	Result: virtual topology and lightpath utilization (multi-hop case with T=5 and W=3).	33
2.8	Throughput results comparison between ILP and heuristic algorithms (total traffic demand is OC-988).	37
2.9	Results of comparison between revenue model and network throughput model.	40
3.1	Comparison of four operations.	58
3.2	The average traffic generated for the NSF network.	62
3.3	The weights of edges assigned in the experiments for the three grooming policies.	63
3.4	Performance comparison of ILP and different heuristics for routing static traffic demands.	66
3.5	The traffic generated for Network 3.	66
4.1	Average utilization of wavelength-links and transceivers when W=16 and L=300 Erlangs.	86
5.1	Summary of the characteristics of different optical grooming switches.	101
7.1	Comparison of three types of OXCs.	151
8.1	Traffic pattern II used in the study.	158

Preface

Optical networks based on wavelength-division multiplexing (WDM) technology offer the promise to satisfy the bandwidth requirements of the Internet infrastructure, and provide a scalable solution to support the bandwidth needs of future applications in the local and wide areas. In a wavelength-routed network, an optical channel, referred to as a *lightpath*, is set up between two network nodes for communication. Using WDM technology, an optical fiber link can support multiple non-overlapping wavelength channels, each of which can be operated at the data rate of 10 Gbps or 40 Gbps today. On the other hand, only a fraction of customers are expected to have a need for such a high bandwidth. Due to the large cost of the optical backbone infrastructure and enormous WDM channel capacity, connection requests with diverse low-speed bandwidth requirements need to be efficiently groomed onto high-capacity wavelength channels. This book investigates the optimized design, provisioning, and performance analysis of traffic-groomable WDM networks, and proposes and evaluates new WDM network architectures.

Organization of the Book

Significant amount of research effort has been devoted to traffic grooming in SONET/WDM ring networks since the current telecom networks are mainly deployed in the form of ring topologies or interconnected rings. As the long-haul backbone networks are evolving to irregular mesh topologies, traffic grooming in optical WDM mesh networks becomes an extremely important and practical research topic for both industry and academia. Chapter 1 gives an overview of traffic grooming in optical WDM network. The remaining chapters focus on traffic grooming in WDM mesh networks only.

In a wavelength-routed WDM network, instead of asking for the capacity of a full wavelength channel, a connection may only require a small fraction of the wavelength capacity. Chapter 2 investigates the problem of grooming static

traffic demands, i.e., a set of pre-known low-speed traffic streams, onto high-capacity lightpaths in a WDM-based optical mesh network. A mathematical formulation of this problem is presented and several connection-provisioning heuristics are also investigated.

To address the traffic-grooming problem, Chapter 3 presents a generic graph model, which captures the various capabilities and constraints of a network and which can be applied to both static and dynamic traffic-grooming problems. Based on the graph model, a grooming algorithm is proposed and different grooming policies are compared and evaluated. This graph model forms a principle method for solving traffic-grooming problems and is used and extended throughout the book.

In dynamic traffic grooming, connection requests with different bandwidth requirement come and go, and the future traffic is unknown. The graph model is used and extended to represent different grooming node architectures, and performance under various scenarios is compared in Chapter 4.

Grooming switch architectures have great impact on the performance of traffic grooming. In Chapter 5, the book explores different architectures and compares their capabilities and performance from different perspectives. Grooming algorithms are also proposed for different grooming architectures.

In a WDM mesh network, not all the nodes need to have grooming capabilities. A network with only a fractional of nodes having grooming functionalities may achieve comparable performance as the one in which all the nodes are grooming capable, but with much lower cost. How to design a network with sparse grooming is investigated and several heuristics are presented in Chapter 6.

To generalize the sparse-grooming problem, we consider the scenario where different nodes may employ different switching architectures. Some nodes do not have grooming capabilities, some nodes have full grooming capabilities, and the others have limited grooming capabilities which can only switch traffic at certain granularities. Design of a WDM network with switches of different bandwidth granularities to achieve cost-effectiveness is a practical and challenging problem, which is investigated in Chapter 7.

Next-generation SONET/SDH network can carry traffic in a finer granularity, and utilize link capacity more efficiently. Moreover, it enables a high-bandwidth connection to be carried by multiple diversely-routed, low-speed connections. This provides much more flexibilities to the network, and leads to new research problems in traffic grooming, which are explored in Chapter 8.

Intended Audience

This book is intended to be a reference book on traffic grooming in optical WDM mesh networks for industrial practitioners, researchers, and graduate students. The book explores various aspects of traffic-grooming problem and

includes state-of-the-art research results, and industrial practitioners and researchers should find this book to be of practical use.

KEYAO ZHU

HONGYUE ZHU

BISWANATH MUKHERJEE

Acknowledgments

Much of the book's material is based on research that we have conducted over the past several years with members of the Networks Research Laboratory at University of California, Davis. We would like to thank Dr. Hui Zang, now at Sprint Advanced Technology Laboratories, for her collaboration on Chapters 3, 4, 5, 6, 7, and 8, and Dr. Jing Zhang, now at Sun Microsystems, for her collaboration on Chapter 8. We would also like to acknowledge the following people of the Computer Science Department at UC Davis — Professor Dipak Ghosal, Professor Charles Martel, Amitabha Banerjee, Yurong (Grace) Huang, Dr. Glen Kramer (now at Teknovus), Dr. Canhui (Sam) Ou (now at SBC Services, Inc.), Smita Rai, Dr. Laxman H. Sahasrabuddhe (now at Park, Vaughan & Fleming LLP), Dr. Narendra Singhal (now at Microsoft Corp.), Dr. Jian Wang (now at Florida International University), Dr. Wushao Wen (now at McAfee), and Dr. Shun Yao (now at Park, Vaughan & Fleming LLP) — for their technical expertise and insightful discussion which have enabled us to better understand the subject matter.

A number of additional individuals whom we have the pleasure to collaborate with and whom we would like to acknowledge are the following: Dr. James Pan at Sprint Advanced Technology Laboratories, Dr. Mike O'Brien (formerly with Sprint Advanced Technology Laboratories), Dr. Takeo Hamada, and Dr. Ching-Fong Su, both at Fujitsu Laboratories of America.

This book would not have been possible without the support of our research on optical networks from several funding agencies as follows: US National Science Foundation (NSF) Grant Nos. ANI-9805285 and ANI-0207864; University of California Micro program; Alcatel Research & Innovation; and Sprint Advanced Technology Laboratories.

We gratefully acknowledge the people at Springer with whom we interacted — Alex Greene and Melissa Guasch — for their encouragement and assistance.

Finally, we wish to thank our family members for their constant support and encouragement.

Chapter 1

OVERVIEW

1.1 Background

Optical wavelength-division multiplexing (WDM) is a promising technology to accommodate the explosive growth of Internet and telecommunication traffic in wide-area, metro-area, and local-area networks. A single optical fiber strand has the potential bandwidth of 50 THz. Using WDM, this bandwidth can be divided into multiple non-overlapping frequency or wavelength channels. Each WDM channel may be operated at any speed, e.g., peak electronic speed of a few gigabits per second (Gbps) [Mukherjee, 1997, Ramaswami and Sivarajan, 1998]. Currently, commercially available optical fibers can support over a hundred wavelength channels, each of which can have a transmission speed of 10 Gbps and higher (e.g., OC-192 and OC-768).

While a single fiber strand has over a terabit-per-second bandwidth and a wavelength channel has over a gigabit-per-second transmission speed, the network may still be required to support traffic connections at rates that are lower than the full wavelength capacity. The capacity requirement of these low-rate traffic connections can vary in range from STS-1 (51.84 Mbps or lower) up to full wavelength capacity. In order to save network cost and to improve network performance, it is very important for the network operator to be able to "*groom*" multiple low-speed traffic connections onto high-capacity circuit pipes.

Different multiplexing techniques can be used for traffic grooming in different domains of optical WDM networks.

- *Space-division multiplexing (SDM)* partitions the physical space to increase transport bandwidth, e.g., bundling a set of fibers into a single cable, or using several cables within a network link [Barr and Patterson, 2001].
- *Frequency-division multiplexing (FDM)* partitions the available frequency spectrum into a set of independent channels. The use of FDM within an op-

tical network is termed (dense) wavelength-division multiplexing (DWDM or WDM) which enables a given fiber to carry traffic on many distinct wavelengths. WDM divides the optical spectrum into coarser units, called wavebands, which are further divided into wavelength channels [Barr and Patterson, 2001].

- *Time-division multiplexing (TDM)* divides the bandwidth's time domain into repeated time-slots of fixed length. Using TDM, multiple signals can share a given wavelength if they are non-overlapping in time [Barr and Patterson, 2001].

- *Dynamic statistical multiplexing or packet-division multiplexing (PDM)* provides "virtual circuit" service in an IP/MPLS over WDM network architecture. The bandwidth of a WDM channel is shared between multiple IP traffic streams (virtual circuits).

Although most research on traffic-grooming problems in the literature concentrate on efficiently grooming low-speed circuits onto high-capacity WDM channels using a TDM approach, the generic grooming idea can be applied to any optical network domain using the various multiplexing techniques mentioned above.

Traffic grooming is composed of a rich set of problems, including network planner, topology design (based on static traffic demand), and dynamic circuit provisioning (based on dynamic traffic demand). The traffic-grooming problem based on static traffic demands is essentially an optimization problem. It can be seen as a dual problem from different perspectives. One perspective is that, for a given traffic demand, satisfy all traffic requests as well as minimize the total network cost. The dual problem is that, for given resource limitation and traffic demands, maximize network throughput, i.e., the total amount of traffic that is successfully carried by the network.

In recent years, there has been an increasing amount of research activity on the traffic-grooming problem, both in academe and in industry. Researchers have been realizing that traffic grooming is a practical and important problem for WDM network design and implementation. In this chapter, we first review traffic grooming on SONET ring-based networks since ring topologies are employed extensively in telecom industry. Then, we introduce some research work on traffic grooming in irregular WDM mesh networks, which is the focus of this book, and various network architectures are presented and discussed.

1.2 Traffic Grooming in SONET Ring Network

1.2.1 Node Architecture

Synchronous Optical Network (SONET) and its equivalent Synchronous Digital Hierarchy (SDH) will be referred to as SONET throughout this book.

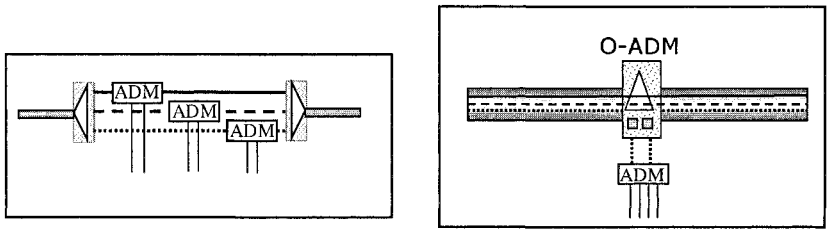


Figure 1.1. Node architectures in a SONET/WDM ring network.

SONET ring is the most widely used optical network infrastructure today. In a SONET ring network, WDM is mainly used as a point-to-point transmission technology. Each wavelength in such a SONET/WDM network is operated at OC-N line rate, e.g., $N = 192$. The SONET system's hierarchical TDM schemes allow a high-speed OC-N channel to carry multiple OC-M channels (where M is smaller than or equal to N). The ratio of N and the smallest value of M carried by the network is called "*grooming ratio*". Electronic add-drop multiplexers (ADMs) are used to add/drop traffic at intermediate nodes to/from the high-speed channels.

In a traditional SONET network, one ADM is needed for each wavelength at every node to perform traffic add/drop on that particular wavelength. With the progress of WDM, over a hundred wavelengths can now be supported simultaneously by a single fiber. It is, therefore, too costly to put the same amount of ADMs (each of which has a significant cost) at every network node since a lot of traffic is only bypassing an intermediate node. With the emerging optical components such as optical add-drop multiplexers (O-ADM) (also referred to as wavelength add-drop multiplexers (W-ADM)), it is possible for a node to bypass most of wavelength channels optically and only drop the wavelengths carrying the traffic destined to the node. An optical wavelength circuit between the electronic components at a node pair is called a "*lightpath*".

Compared with the wavelength channel resource, ADMs form the dominant cost in a SONET/WDM ring network. Hence, carefully arranging these optical bypasses can reduce a large amount of the network cost. Figure 1.1 shows different node architectures in a SONET/WDM ring network. It is clear that using O-ADMs can decrease the number of SONET ADMs used in the network and eventually bring down the network cost. Then the problems are, for a given low-speed set of traffic demands, which low-speed demands should be groomed together, which wavelengths should be used to carry the traffic, which wavelengths should be dropped at a local node, and how many ADMs are needed at a particular node?

1.2.2 Single-Hop Grooming in SONET/WDM Ring

A SONET/WDM ring network can have the node architecture shown in Fig. 1.1(b). OC-M low-speed connections are groomed on to OC-N wavelength channels. Assume that there is no wavelength converter at any network node. The traffic on a wavelength cannot be switched to other wavelengths. Based on this network model, for a given traffic matrix, satisfying all the traffic demands as well as minimizing the total number of ADMs is a network design/optimization problem and has been studied extensively in the literature.

Figures 1.2 and 1.3 show an example in which, by carefully grooming traffic in the SONET/WDM rings, some network cost saving can be achieved. Figure 1.2 shows a SONET/WDM ring network with 6 unidirectional connection requests. Each node is also equipped with an O-ADM (not shown in the figures). Assume that the SONET ring is also unidirectional (clockwise), the capacity of each wavelength is OC-N, and it can support two OC-M low-speed traffic requests in TDM fashion, i.e., $N = 2M$. In order to support all of the traffic requests, 8 ADMs are used in the network. Figure 1.3(a) shows a possible configuration. By interchanging the connections (1,3) and (2,3), wavelength 2 (shown as thick lines) can be optically bypassed at node 2, which results in one ADM savings at node 2. Figure 1.3(b) shows this configuration.

It has been proven in [Chiu and Modiano, 2000, Wan et al., 2000] that the general traffic-grooming problem is *NP*-Complete. The authors in [Wang et al., 2001] formulate the optimization problem as an integer linear program (ILP). When the network size is small, some commercial software can be used to solve the ILP equations to obtain an optimal solution. The formulation in [Wang et al., 2001] can be applied to both uniform and non-uniform traffic demands, as well as to unidirectional and bi-directional SONET/WDM ring networks. The limitation of the ILP approach is that the numbers of variables and equations increase explosively as the size of network increases. The computation complexity makes it hard to be useful on networks with practical size. By relaxing some of the constraints in the ILP formulation, it may be possible to get some results, which are close to the optimal solution for reasonable-size networks. The results from the ILP may give some insights and intuition for the development of good heuristic algorithms to handle the problem in a large network.

In [Wan et al., 2000, Zhang and Qiao, 2000, Simmons et al., 1999], some lower bound analysis is given for different traffic criteria (uniform and non-uniform) and network model (unidirectional ring and bi-directional ring). These lower bound results can be used to evaluate the performance of traffic-grooming heuristic algorithms. In most of the heuristic approaches, the traffic-grooming problem is divided into several sub-problems and solved separately. These heuristics can be found in [Chiu and Modiano, 2000, Wan et al., 2000, Wang et al., 2001, Zhang and Qiao, 2000, Simmons et al., 1999, Gerstel et al., 1998].

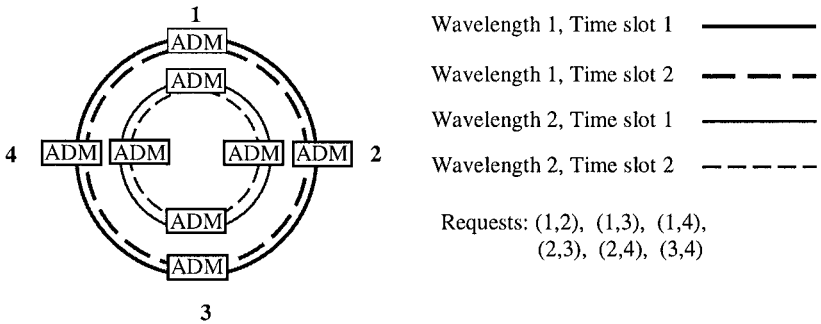


Figure 1.2. A SONET/WDM network with 4 nodes and 2 wavelengths.

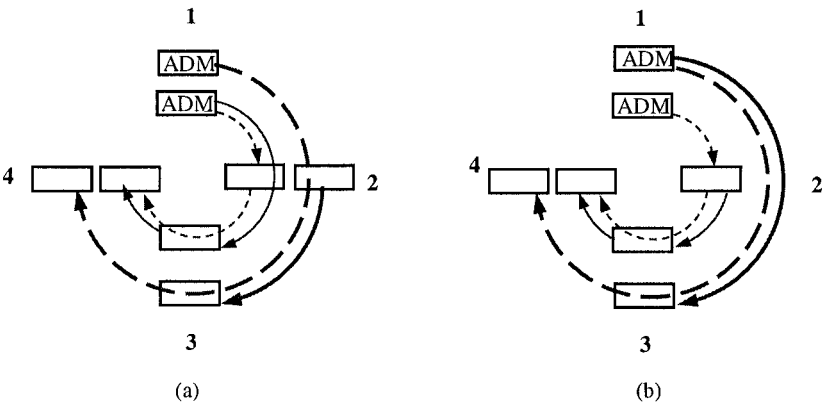
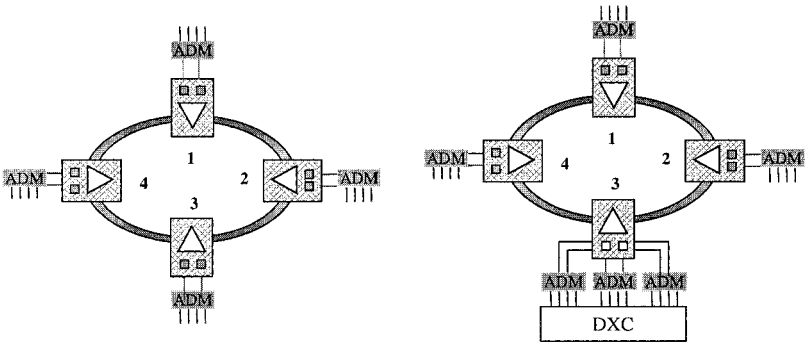


Figure 1.3. Two possible configurations to support the traffic requests in Fig. 1.2.

Greedy approach, approximation approach, and simulated annealing approach are used in these heuristic algorithms.

1.2.3 Multi-Hop Grooming in SONET/WDM Ring

In single-hop (a single-lightpath hop) grooming, traffic cannot be switched between different wavelengths. Figure 1.4(a) shows this kind of a network configuration. Another network architecture has been proposed in [Simmons et al., 1999, Gerstel et al., 2000], in which there are some nodes equipped with Digital Crossconnects (DXCs). In Fig. 1.4(b), node 3 has a DXC installed. This kind of node is called a hub node. Traffic from one wavelength/time-slot can be switched to any other wavelength/time-slot at the hub node. Because the traffic needs to be converted from optical to electronic at the hub node when



(a) Single-hop (without hub node) (b) Multi-hop (with a hub node)

Figure 1.4. SONET/WDM ring with/without a hub node.

wavelength/time-slot exchange occurs, this grooming approach is called multi-hop (multi-lightpath hops) grooming. Depending on the implementation, there can be a single hub node or multiple hub nodes in the network. A special case is that every node is a hub node, i.e., there is a DXC at every node. This kind of network is called point-to-point WDM ring network (PPWDM ring) [Gerstel et al., 2000].

The work in [Gerstel et al., 2000] provides some excellent theoretical analysis on comparing network cost of PPWDM ring, a SONET/WDM ring without hub node, a SONET/WDM ring with one or multiple hub nodes, etc. The authors of [Wang et al., 2001] have compared the single-hop grooming with multi-hop grooming (with one hub node) network performance using simulation. The results indicate that, when the grooming ratio is large, the multi-hop approach tends to use fewer ADMs, but when the grooming ratio is small, the single-hop approach tends to use fewer ADMs, and in general, the multi-hop approach uses more wavelengths than the single-hop approach.

1.2.4 Dynamic Grooming in SONET/WDM Ring

Instead of using a single static traffic matrix to characterize the traffic requirement, it is also possible to describe it by a set of traffic matrices. The traffic pattern may change within this matrix set over a period of time, say throughout a day or a month. The network needs to be reconfigured when the traffic pattern transits from one matrix to another matrix in the matrix set. The network design problem for supporting any traffic matrix in the matrix set (in a non-blocking manner) as well as minimizing the overall cost is known as a dynamic grooming problem in a SONET/WDM ring [Berry and Modiano, 2000].

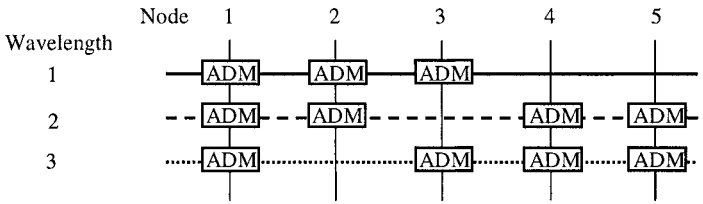


Figure 1.5. Network design for 2-allowable traffic.

Unlike the dynamic provisioning and grooming problem in a WDM mesh network, which will be introduced in Section 3, the dynamic-grooming problem proposed in [Berry and Modiano, 2000] is more likely a network design problem with reconfiguration consideration. The authors of [Berry and Modiano, 2000] have formulated the general dynamic-grooming problem in a SONET/WDM ring as a bipartite graph-matching problem and provided several methods to reduce the number of ADMs. A particular traffic matrix set is then considered and the lower bound on the number of ADMs is derived. They also provide the necessary and sufficient conditions so that a network can support such a traffic pattern. This kind of traffic matrix set is called a t-allowable traffic pattern. For a given traffic matrix, if each node can source at most t duplex circuits, we call this traffic matrix a t-allowable traffic matrix. The traffic matrix set, which only consists of t-allowable traffic matrices, is called a t-allowable matrix set or a t-allowable traffic pattern. We use an example from [Modiano and Lin, 2001] to illustrate dynamic traffic grooming for a t-allowable traffic pattern in a SONET/WDM ring.

Figure 1.5 shows a 5-node SONET/WDM ring network. Three wavelengths are supported in the network. Assume that each wavelength can support 2 low-speed circuits. The network configuration in Fig. 1.5 is a 2-allowable configuration, i.e., it can support any 2-allowable traffic matrix (set). For instance, consider a traffic matrix with request streams 1-2, 1-3, 2-3, 2-4, 3-4, 4-5, 4-5. The traffic matrix can be supported by assigning 1-3, 2-3 on wavelength 1, assigning 1-2, 2-4, 4-5, 4-5 on wavelength 2, and assigning 3-4 on wavelength 3. Note that, for a particular traffic matrix, there may be some redundant ADMs in the configuration. However, the configuration is able to support other potential t-allowable traffic matrices. Designing such configurations to support any t-allowable traffic matrix while minimizing the network cost is a very interesting problem with some practical utility. The authors in [Berry and Modiano, 2000] provide an excellent analysis on t-allowable traffic pattern. The study of dynamic-traffic grooming in a SONET/WDM ring with other generic traffic pattern can be potentially challenging research.

1.2.5 Grooming in Interconnected SONET/WDM Rings

Most traffic-grooming studies in SONET/WDM ring networks have assumed a single-ring network topology. The authors of [Wang and Mukherjee, 2002] have extended the problem to an interconnected-ring topology. Today's backbone networks are mainly constructed as a network of interconnected rings. Extending the traffic-grooming study from a single-ring topology to the interconnected-ring topology will be very useful for a network operator to design their network and to engineer the network traffic.

Figure 1.6(a) shows an interconnected SONET/WDM ring network with a single junction node. Multiple junction nodes may also exist in the interconnected-ring topology because of network survivability consideration. Various architectures can be used at the junction node to interconnect the two SONET rings. Figures 1.6(b)–(d) [Wang and Mukherjee, 2002] show some of the node architectures.

In Fig. 1.6(b), an O-ADM is used to drop some wavelengths at the junction node. ADMs and a DXC are used to switch the low-speed circuits between the interconnected rings. This node architecture has wavelength-conversion and time-slot interchange capability, i.e., a time-slot (low-speed circuit) on one wavelength can be switched to another time-slot on a different wavelength through this junction node.

Figure 1.6(c) uses an Optical Crossconnect (OXC) to interconnect the two rings. There are transparent and opaque technologies to build these OXCs. Transparent refers to all-optical switching, and opaque refers to switching with optical-electronic-optical (O-E-O) conversion. Depending on the implementation, the OXC may be equipped with or without wavelength-conversion capability. This node architecture can only switch the traffic at the wavelength granularity between the interconnected rings. Note that extra ADMs are needed to support local traffic originating from or terminating at the junction node.

Figure 1.6(d) shows a hierarchical node architecture with a switching capability on both wavelength and lower-speed circuit granularity.

The different node architectures at the junction nodes in the interconnected-ring network will add different constraints to the traffic-grooming problem. The work in [Wang and Mukherjee, 2002] presented the ILP formulation of the traffic-grooming problem in an interconnected-dual-ring topology, and proposed a heuristic algorithm to handle the problem for networks of practical size. Results are compared between the various junction nodes' interconnection strategies and grooming ratios. When the number of rings and the number of junction nodes increase in the interconnected-ring network, the network topology tends to become an irregular mesh topology.

Section 1.3 discusses some recent studies on traffic grooming on the WDM mesh topology. A comparison between the interconnected-ring approach and

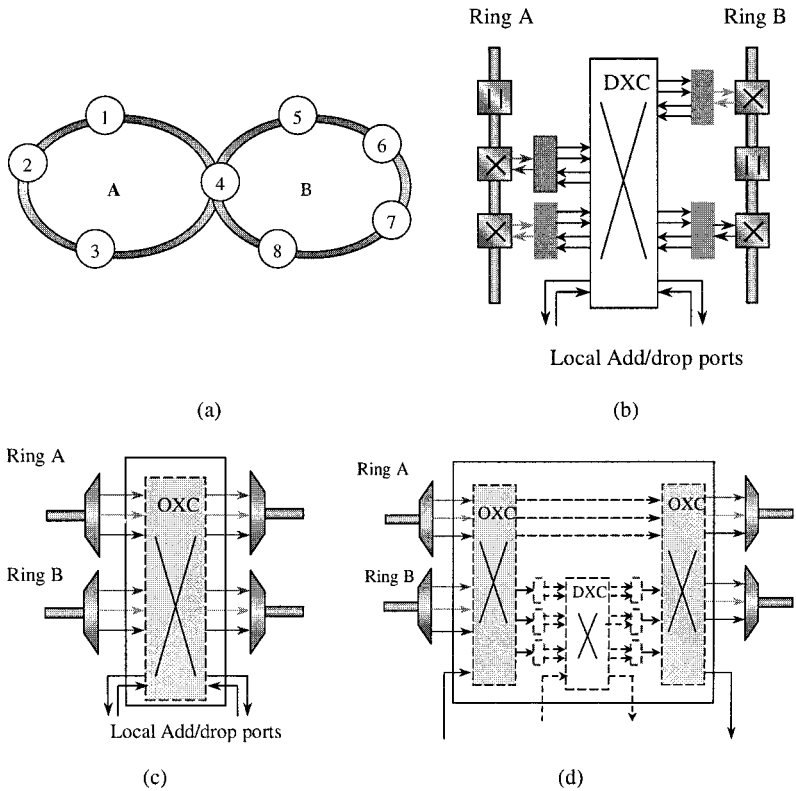


Figure 1.6. A sample interconnected-ring network topology and simplified architectures of the junction node.

the mesh approach will be a potential research challenge and hasn't been explored in the literature yet.

1.3 Traffic Grooming In Wavelength-Routed WDM Mesh Network

A significant amount of previous work on traffic grooming in the optical network literature is based on the ring network topology. Recently, traffic grooming in a WDM mesh network has started to get more attention. In this section, we review some recent work, which have been reported on this subject. We will also show some potential research challenges and directions.

1.3.1 Network Provisioning: Static and Dynamic Traffic Grooming

Although the SONET (interconnected) ring network has been used as the first generation of the optical network infrastructure, it has some limitations, which make it hard to accommodate the increasing Internet traffic. The next-generation optical network is expected to be an intelligent wavelength-routed WDM mesh network. This network will provide fast and convenient (point-and-click) automatic bandwidth provisioning and efficient protection mechanisms; and it will be based on an irregular mesh topology, which will make it much easier to scale.

When such a network is constructed, how to efficiently accommodate the incoming traffic requests is a network-provisioning problem. The traffic request can be static (measured by one or multiple fixed traffic matrices) or dynamic (measured by the arrival rate and the holding time statistics of a connection request). The work in [Zhu and Mukherjee, 2002b], based on static traffic demands, discusses the node architectures in a WDM mesh network, which has traffic-grooming capability. Grooming node architectures are discussed in detail in Chapter 5.

Figure 1.7 shows such an OXC architecture, which has hierarchical switching and multiplexing functionality. Instead of using a separate wavelength switching system and a grooming system (Fig. 1.6(d)), the OXC in Fig. 1.7 can directly support low-speed circuits and groom them onto wavelength channels through a grooming fabric (G-Fabric) and built-in transceiver arrays. This kind of OXC is called Wavelength-Grooming Crossconnect (WGXC) in [Thiagarajan and Somani, 2001b]. In a network equipped with a WGXC at every node, the grooming fabric and the size of the transceiver array provide another dimension of constraints on the network performance besides the wavelength-resource constraint. This is similar to the ADM constraint for traffic grooming in SONET/WDM ring networks.

The transceiver array used in the OXC can be either tunable or fixed. The authors in [Zhu and Mukherjee, 2002b] (see Chapter 2) consider a static traffic matrix set as the network traffic demands. Each traffic matrix in the matrix set represents a particular low-speed circuit request class. For given network resource constraints and traffic demands, the work studies how to maximize the network throughput under the network resource limitation. As stated in Section 1.1, minimizing cost and maximizing network throughput lead to two different perspectives on the same traffic-grooming problem. The authors formulate the problem as an ILP. A small network is used to show ILP results and two heuristic algorithms are proposed to study larger networks based on the observations from these results. Different network scenarios are considered and compared. They are single-hop grooming vs. multi-hop grooming,

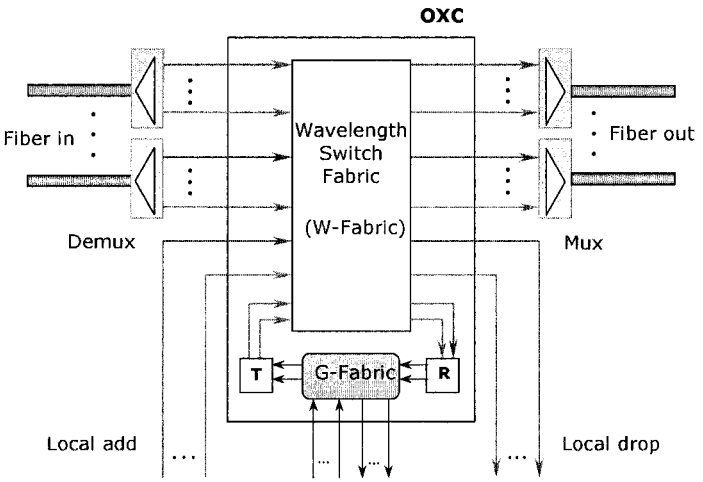


Figure 1.7. An OXC with a two-level hierarchy and grooming capability.

tunable transceivers vs. fixed transceivers, optimizing network throughput vs. optimizing network revenue, etc.

Unlike the work in [Zhu and Mukherjee, 2002b], the work in [Thiagarajan and Somani, 2001a, Thiagarajan and Somani, 2001b] considers a dynamic traffic pattern in a WDM mesh network. The work in [Thiagarajan and Somani, 2001b] has proposed a connection admission control (CAC) scheme to ensure that the network will treat every connection fairly. It has been observed in [Thiagarajan and Somani, 2001b] that, when most of the network nodes have grooming capability, the high-speed connection requests will have higher blocking probability than the low-speed connection requests in the absence of any fairness control. CAC is needed to guarantee that every class of connection requests will have similar blocking probability performance. The work in [Thiagarajan and Somani, 2001a] proposed a theoretical capacity correlation model to compute the blocking probability for WDM networks with constrained grooming capability.

The work in [Zhu and Mukherjee, 2002b] has assumed that every node is a WGXC node, and the grooming capability is constrained by the grooming fabric and transceiver array at every node. The work in [Thiagarajan and Somani, 2001b] has assumed that only a few of the network nodes are WGXC nodes and there is no constraint on these nodes. It will be a good extension to combine these assumptions and study the network performance as well as fairness in a static as well as a dynamic environment. This extension will be very practical and important to a service provider.

1.3.2 Network Design and Planner

Unlike the network-provisioning problem addressed in Section 1.3.1, the work in [Cox and Sanchez, 2001] studied how to plan and design such a WDM mesh network with certain forecast traffic demands. The problem is a network design and planner methodology. The problem description is as follows: given forecast traffic demand (static) and network node (locations), determine how to connect the nodes using fiber links and OXCs and route the traffic demands in order to satisfy all of the demands as well as minimize the network cost. The network cost is measured by the fiber cost, OXC or DXC port cost, and WDM system cost used in the network.

Figure 1.8 (from [Cox and Sanchez, 2001]) gives an example on this network design and planner problem considering traffic grooming. Figure 1.8(a) shows a four-node network and the traffic demands. Each link in Fig. 1.8(a) is a fiber conduit, which may carry multiple fiber links. Assume that the cost of a fiber going through one conduit is one unit and the capacity of a wavelength channel is OC-192. Five segments exist in Fig. 1.8(a): (A, B), (A, C), (A, D), (B, C), and (B, D). A segment is a sequence of fiber links that does not pass through an OXC [Cox and Sanchez, 2001]. There are two possible network design options to accommodate the traffic demands, which are shown in Figs. 1.8(b) and 1.8(c).

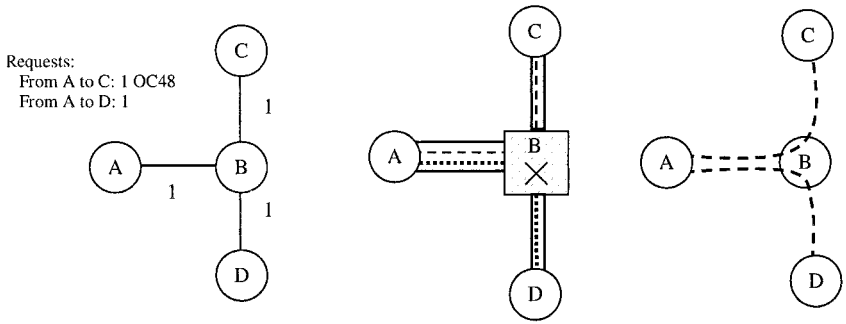
- Option 1 (Fig. 1.8(b)):
  - Place a fiber on segments (A, B), (B, C), and (C, D).
  - Install a WDM system on each fiber.
  - Place an OXC with 4 ports at node B to interconnect the wavelength channels.
  - There will be a total of 4 OXC ports used at nodes A, C, and D to add and drop traffic.

  Total cost for option 1 will be:

  $$Cost_{(option\ 1)} = 3 \cdot Cost_{fiber} + 3 \cdot Cost_{(WDM\ system)} + 8 \cdot Cost_{(OXC\ port)}$$

  The two demands will be carried by the dashed wavelength and dotted wavelength shown in Fig. 1.8(b).

- Option 2 (Fig. 1.8(c)):
  - Place a fiber on segments (A, C) and (A, D). These fibers will bypass node B.
  - Install a WDM system on each fiber.



(a) Node and fiber conduit location (b) Design option 1 (c) Design option 2

Figure 1.8. Two different designs for a 4-node network [Cox and Sanchez, 2001].

- There will be a total of 4 OXC ports used at nodes A, C, and D to add and drop traffic.

Total cost for option 2 will be:

$$Cost_{(option\ 2)} = 4 \cdot Cost_{fiber} + 2 \cdot Cost_{(WDM\ systems)} + 4 \cdot Cost_{(OXC\ ports)}$$

The two demands will be carried by the dashed wavelengths shown in Fig. 1.8(c).

From this example, we can see that each network element has its own cost function and the definitions of these cost functions will eventually determine how the network should be designed.

The authors in [Cox and Sanchez, 2001] have addressed this network design and planner problem. The problem is formulated as an ILP. Two heuristic algorithms are proposed for the mesh network design and the ring network design separately, i.e., design the network as an irregular mesh topology or an interconnected-ring topology. The authors compare the results between the mesh design and ring design. They find that (a) the mesh topology design has a compelling cost advantage for sufficiently large distance scales; (b) for ring technologies such as OC-192 BLSR, using WDM only results in cost savings when distances are sufficiently large; and (c) costs can be very insensitive to distance for ring technologies [Cox and Sanchez, 2001].

1.3.3 Grooming with Protection Requirement in WDM Mesh Network

The SONET/WDM ring network has been demonstrated to have reliable link-protection schemes. There is no need to consider the protection issue

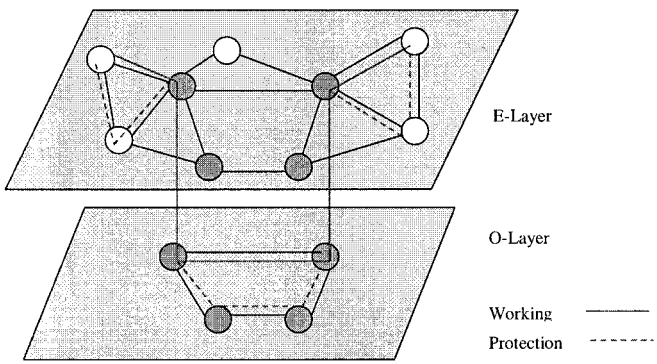


Figure 1.9. A multi-layer protection example [Lardies et al., 2001].

separately for groomed traffic in such a network. On the other hand, protection for groomed traffic should be studied in a WDM mesh network.

In a WDM mesh network, various protection schemes can be used depending on the network operator's preference and the customer's requirements. Either link-protection scheme or path-protection scheme may be applied on a WDM mesh network, and the protection resources can be dedicated or shared by the working circuits [Ramamurthy et al., 2003]. Although WDM protection schemes in mesh networks have been studied extensively, protection with traffic grooming is a new research area and has started to receive attention [Ou et al., 2003].

Different low-speed circuits may ask for different bandwidth requirements as well as protection service requirement. The low-speed circuits may be protected at either the electronic layer or at the optical layer. Figure 1.9 (from [Lardies et al., 2001]) shows an example of path protection in a network with electronic layer and optical layer. In Fig. 1.9, the shaded nodes are the nodes which are equipped with OXCs. Lightpaths can be established between these nodes, and low-speed connections can be groomed onto these lightpaths and transmitted in the optical domain.

Given a static traffic matrix and the protection requirement for every request (no protection, $1+1$ protection, $1:n$ protection, etc.), the authors in [Lardies et al., 2001] studied how to satisfy these connections' bandwidth and protection requirements while minimizing the network cost. Network cost is determined by the transmission cost and switching cost in a manner similar to that described in Section 1.3.2. The bandwidth requirement of a connection can be a fraction of a wavelength channel and some connections may be partially carried by the electronic layer. The authors of [Lardies et al., 2001] show how much benefit

there will be on network cost by grooming the traffic onto the optical domain instead of carrying them purely on the electronic layer. An ILP formulation is given and a simple heuristic is proposed. It should be possible to improve the heuristic and its performance presented in [Lardies et al., 2001]. The study of traffic grooming with protection requirement in a dynamic environment is a challenge and interesting topic.

1.3.4 Grooming with Multicast in WDM Mesh Network

Multicast applications such as video-on-demand and interactive games are becoming more and more popular. It is reasonable to estimate that there will be more such multicast applications which may require vast amount of bandwidth in the near future such as video conferencing, virtual reality entertainment, etc. Optical multicasting using "*light-tree*" [Sahasrabuddhe and Mukherjee, 1999] may be a good solution for these requirements. Since each wavelength will have capacity up to OC-192 (OC-768 in the future), multiple multicast sessions can be groomed to share the capacity on the same wavelength channel. In this case, the lightpaths or the light-trees can be established to accommodate multicast requests, which have lower capacity requirement than the bandwidth of a wavelength.

Figure 1.10 shows a simplified switch architecture, which can support multicast sessions with full wavelength capacity requirement or partial wavelength capacity requirement. With this architecture, the data on a wavelength channel from one incoming fiber or the local node can be switched to multiple outgoing fibers, and a full wavelength channel multicasting session can be maintained as much as possible in the optical domain. The DXC in Fig. 1.10 has multicast capability. This kind of electronic switch fabric is already commercially available. By combining this DXC with OE/EO conversion components (electronic mux/demux and transceiver), a low-speed multicast session can be groomed with other low-speed unicast/multicast sessions.

The work in [Singhal and Mukherjee, 2001] reports on some preliminary studies on multicast grooming in WDM mesh networks. The problem is defined as follows: given a set of multicast sessions with various capacity requirements, satisfy all of the multicast sessions, and at the same time, minimize the network cost. The network cost is measured by the wavelength-link cost used in the network. The authors show an ILP formulation for this problem and present some results based on some sample traffic matrices and network topologies. It is hard to scale the ILP approach to handle networks of practical size. Hence, simpler and efficient algorithms need to be explored to achieve near-optimal solutions. Multicast with grooming is a new research area and is expected to receive more attention in the optical networking literature.

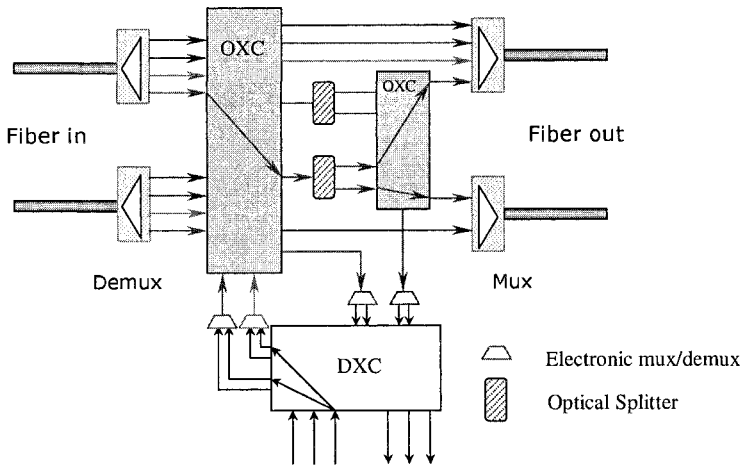


Figure 1.10. Switch architecture for supporting multicast grooming [Sahasrabuddhe and Mukherjee, 1999].

1.3.5 Protocols and Algorithm Extensions for WDM Network Control

Traffic grooming is a very important problem whose solution will enable us to fully develop an intelligent WDM optical transport network. The unified control plane of such a network is being standardized, and is known as Generalized Multi-protocol Label Switching (GMPLS) [Xu et al., 2000] in the Internet Engineering Task Force (IETF) forum. The purpose of this network control plane is to provide an intelligent automatic end-to-end circuit (virtual circuit) provisioning/signaling scheme throughout the different network domains. Different multiplexing techniques such as PDM, TDM, WDM, and SDM may be used for such an end-to-end circuit, and good grooming schemes are needed to efficiently allocate network resources.

There are three components in the control plane that need to be carefully designed to support traffic grooming, namely, resource-discovery protocol, signaling protocol, and path-computation algorithms. Several resource-discovery protocols based on traffic-engineering (TE) extensions of link-state protocols (OSPF, IS-IS) and link-management protocols have been proposed in IETF as the resource-discovery component in the control plane. The extensions of the MPLS signaling protocols are proposed as the signaling protocol in this control plane. An open issue is the design of efficient route-computation algorithms.

Chapter 2

STATIC TRAFFIC GROOMING

2.1 Introduction

Assigning network resources (e.g., wavelengths, transceivers) to successfully carry the connection requests (lightpaths) in an optical WDM mesh network is well known as the routing and wavelength assignment (RWA) problem [Mukherjee, 1997, Ramaswami and Sivarajan, 1998]. It is also known as a lightpath-provisioning problem. A lot of RWA studies have been reported in the optical networking literature, either based on static traffic demands [Mukherjee, 1997][Banerjee and Mukherjee, 1996] or based on dynamic traffic demands [Zhang and Qiao, 1998][Mokhtar and Azizoglu, 1998][Zang et al., 2000][Li and Somani, 1999][Jue and Xiao, 2000][Harai et al., 1997]. Most previous studies have assumed that a connection requests bandwidth for an entire lightpath channel. In this study, we assume the bandwidth of connection requests can be some fraction of the lightpath capacity, which makes the problem more practical.

We investigate the problem of how to "groom" low-speed connection requests to high-capacity lightpaths efficiently. The traffic-grooming problem has been studied on the SONET ring topology, as was discussed in Chapter 1. The objective function in these studies is to minimize the total network cost, measured in terms of the number of SONET add-drop multiplexers (ADMs). In this chapter, we consider irregular mesh WDM networks and assume that a connection requests bandwidth that is a fraction of the wavelength capacity.

Figure 2.1 shows an illustrative example of traffic grooming in a WDM mesh network. Fig. 2.1(a) shows a small six-node network. Each fiber has two wavelength channels. The capacity of each wavelength channel in this example is OC-48, i.e., approximately 2.5 Gbps. Note that the bandwidth of an OC-n channel is approximately $n \times 51.84$ Mbps. Each node is equipped

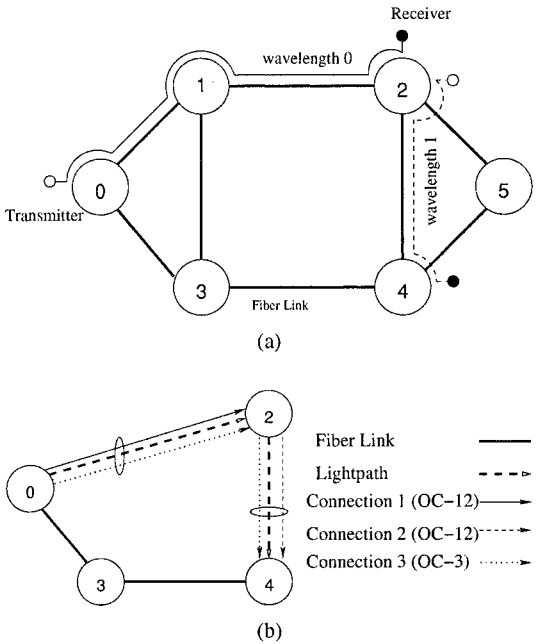


Figure 2.1. Illustrative example of traffic grooming.

with a tunable transmitter and a tunable receiver, both of which can be tuned to any wavelength. There are three connection requests: $(0, 2)$ with bandwidth requirement OC-12, $(2, 4)$ with bandwidth requirement OC-12, and $(0, 4)$ with bandwidth requirement OC-3. Two lightpaths have already been set up to carry these three connections, as shown in Fig. 2.1(a). Because of the resource limitations (transmitter in node 0 and receiver in node 4 are busy), we cannot set up a lightpath directly from node 0 to node 4; thus, connection 3 has to be carried by the spare capacity of the two existing lightpaths, as shown in Fig. 2.1(b). Different connection requests between the same node pair (s, d) can be either groomed on the same lightpath, which directly joins (s, d), using various multiplexing techniques, or routed separately through different virtual paths. A connection may traverse multiple lightpaths if no resources are available to set up a lightpath between the source and the destination directly.

We investigate the node architecture for the WDM optical network to support traffic-grooming capability. We study an optical wide-area WDM network which utilizes a grooming-capable optical node architecture, so that a group of lightpaths can be set up to optimally carry the low-speed connection requests.

We formulate the traffic-grooming problem in a mesh network as an optimization problem with the following objective function: for a given traffic

matrix set and network resources, maximize the (weighted) network throughput. The mathematical formulation is presented for static traffic demands. Several simple provisioning algorithms, i.e., heuristics, are also proposed and their performance is compared. Finally, we show how to extend the mathematical formulation to accommodate other network optimization criteria.

2.2 General Problem Statement

The problem of grooming low-speed traffic requests onto high-bandwidth wavelength channels on a given physical topology (fiber network) is formally stated below. We are given the following inputs to the problem.

1 A physical topology $G_p = (V, E_p)$ consisting of a weighted unidirectional graph, where V is the set of network nodes, and E_p is the set of physical links, which connect the nodes. Nodes correspond to network nodes and links correspond to the fibers between nodes. Though links are unidirectional, we assume that there are an equal number of fibers joining two nodes in different directions. Links are assigned weights, which may correspond to physical distance between nodes. In this study, we assume that all links have the same weight 1, which corresponds to the fiber hop distance. A network node i is assumed to be equipped with a $D_p(i) \times D_p(i)$ optical crossconnect ((OXC), also called wavelength-routing switch (WRS)), where $D_p(i)$ denotes the number of incoming fiber links to node i. For any node i, the number of incoming fiber links is equal to the number of outgoing fiber links.

2 Number of wavelength channels carried by each fiber is W. Capacity of a wavelength is C.

3 A set of $N \times N$ traffic matrices, where $N = |V|$. Each traffic matrix in the traffic-matrix set represents one particular group of low-speed connection requests between the nodes of the network. For example, if C is OC-48, there may exist four traffic matrices: an OC-1 traffic matrix, an OC-3 traffic matrix, an OC-12 traffic matrix, and an OC-48 traffic matrix.

4 The number of lasers (transmitters) (TR_i) and filters (receivers) (RR_i) at each node i. Note that the transceiver can be either wavelength-tunable or part of a fixed-tuned array.

Our goals are to determine the following:

1 A virtual topology $G_v = (V, E_v)$. The nodes of the virtual topology correspond to the nodes in the physical topology. A link between nodes i and j corresponds to an unidirectional lightpath set up between node pair (i, j).

2 Routing connection requests on the virtual topology to either minimize the total network cost or maximize total throughput. In this study, we consider maximizing total throughput.

2.3 Node Architecture

To carry connection requests in a WDM network, lightpath connections may be established between pairs of nodes. A connection request may traverse through one or more lightpaths before it reaches the destination. Two important functionalities must be supported by the WDM network nodes: one is wavelength routing, and the other is multiplexing and demultiplexing. An OXC provides the wavelength-routing capability to the WDM network nodes. Optical multiplexer/demultiplexer can multiplex/demultiplex several wavelengths to the same fiber link. Low-speed connection requests will be multiplexed on the same lightpath channel by using an electronic-domain TDM-based multiplexing technique. Figures 2.2 and 2.3 show two sample node architectures in a WDM optical network.

The node architecture is composed of two components: WRS and access station. The WRS performs wavelength routing and wavelength multiplexing/demultiplexing. The access station performs local traffic adding/dropping and low-speed traffic-grooming functionalities. WRS is composed of an Optical Crossconnect (OXC), Network Control and Management Unit (NC&M), and Optical Multiplexer/Demultiplexer. In the NC&M unit, the network-to-network interface (NNI) will configure the OXC and exchange control messages with peer nodes on a dedicated wavelength channel (shown as wavelength 0 in Figs. 2.2 and 2.3). The network-to-user interface (NUI) will communicate with the NNI and exchange control information with the user-to-network interface (UNI), the control component of the access station. The OXC provides wavelength-switching functionality. As shown in the example in Fig. 2.2, each fiber has three wavelengths. Wavelength 0 is used as a control channel for the NC&M to exchange control messages between network nodes. Other wavelengths are used to transmit data traffic.

In Fig. 2.2, each access station is equipped with some transmitters and receivers (transceivers). Traffic originating from an access station is sent out as an optical signal on one wavelength channel by a transmitter. Traffic destined to an access station is converted from an optical signal to electronic data by a receiver. Both tunable transceivers and fixed transceivers could be used in a WDM network. A tunable transceiver can be tuned between different wavelengths so that it can send out (or receive) an optical signal on any free wavelength in its tuning range. A fixed transceiver can only emit (or receive) an optical signal on one wavelength. To explore all of the wavelength channels on a fiber, a set of fixed transceivers, one per wavelength, can be grouped together to form a transceiver array. The size of a fixed transceiver array can be equal to or smaller than the

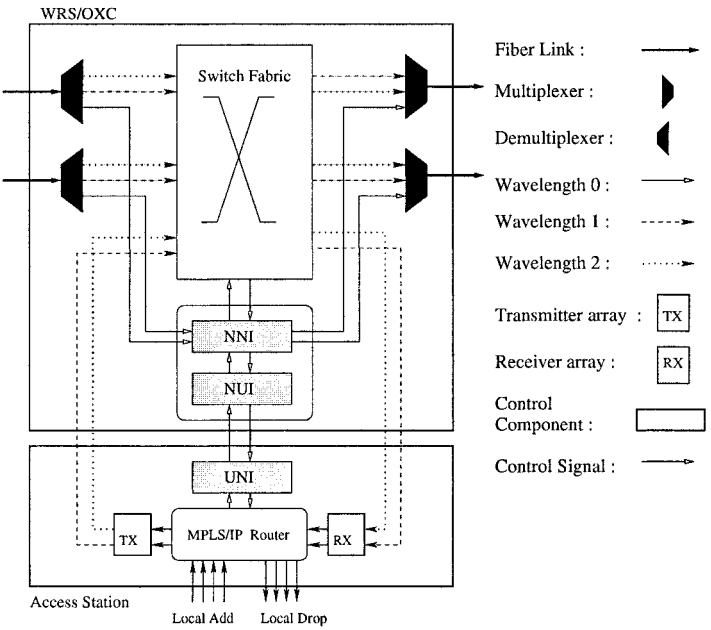


Figure 2.2. Node architecture 1: IP over WDM.

number of wavelengths on a fiber, and the number of transceiver arrays can be equal to or smaller than the number of fibers joining a node.

The access station in Fig. 2.2 provides a flexible, software-based, bandwidth-provisioning capability to the network. Multiplexing low-speed connections to high-capacity lightpaths is done by the MPLS/IP router using a software-based queueing scheme. The advantages of this model are that 1) it provides flexible bandwidth granularity for the traffic requests and 2) this MPLS/IP-over-WDM model has much less overhead than the SONET-over-WDM model, shown in Fig. 2.3. A potential disadvantage of this model is that the processing speed of the MPLS/IP router may not be fast enough compared with the vast amount of bandwidth provided by the optical fiber link.

In Fig. 2.3, each access station is equipped with several SONET Add/Drop Multiplexers (ADMs). Each SONET ADM has the ability to separate a high-rate SONET signal into lower-rate components [Chiu and Modiano, 2000]. In order for a node to transmit or receive traffic on a wavelength, the wavelength must be added or dropped at the node and a SONET ADM must be used. Generally, each SONET ADM is equipped with a fixed transceiver and operates only on one wavelength as shown in Fig. 2.3. The Digital Crossconnect (DXC) can interconnect the low-speed traffic streams between the access station and the

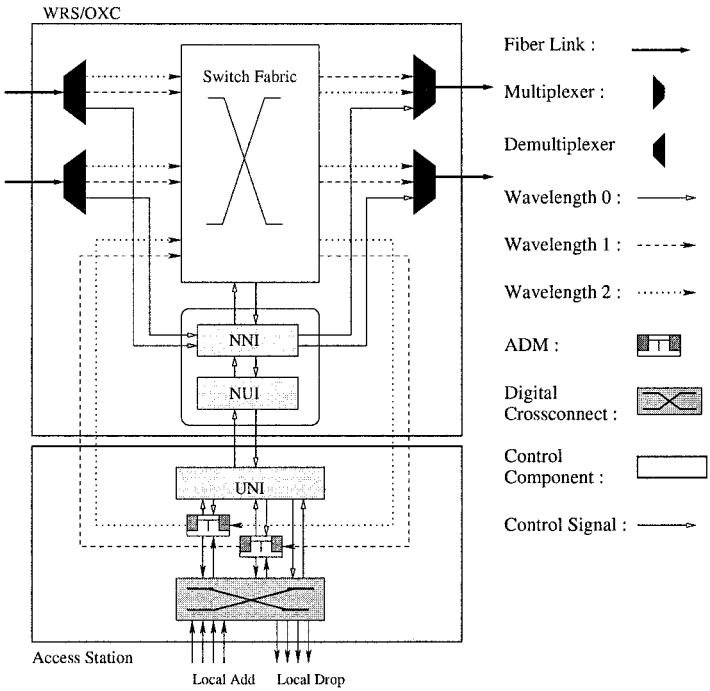


Figure 2.3. Node architecture 2: SONET over WDM.

ADMs. A low-speed traffic stream on one wavelength can be either dropped to the local client (IP router, ATM switch, etc.) or switched to another ADM and sent out on another wavelength. Figure 2.3 presents a SONET-over-WDM node architecture. SONET components (ADM, DXC, etc.) and SONET framing schemes can provide TDM-based fast multiplexing/demultiplexing capability, compared with the software-based scheme in Fig. 2.2. The disadvantage of this approach is the high cost of SONET components, such as ADM and DXC. In reality, both kinds of access stations may be used together to be connected with an OXC in order to have a multi-service provisioning platform (MSPP) for accessing an OXC in a carrier's network.

2.4 Mathematical (ILP) Formulation of the Traffic-Grooming Problem

The traffic-grooming problem in a mesh network with static traffic pattern turns out to be an Integer Linear Program (ILP). We make the following assumptions in our study:

1 The network is a single-fiber, irregular mesh network, i.e., there is at most one pair of (bidirectional) fiber link between each node pair.

2 The wavelength-routing switches in the network nodes do not have wavelength conversion capability. A lightpath connection must be set up on the same wavelength channel if it traverses through several fibers. Extension of this problem to include wavelength conversion is straightforward and it actually makes the problem simpler in terms of the number of variables and equations.

3 The transceivers in a network node are tunable to any wavelength on the fiber.

4 A connection request cannot be divided into several lower-speed connections and routed separately from the source to the destination. The data traffic on a connection request should always follow the same route. This constraint will be relaxed in Chapter 8 while exploiting SONET/SDH's virtual concatenation (VCAT) property.

5 Each node has unlimited multiplexing/demultiplexing capability and time-slot interchange capability. This means that the access station of a network node can multiplex/demultiplex as many low-speed traffic streams to a lightpath as needed, as long as the aggregated traffic does not exceed the lightpath capacity. This may only be true for the software-based provisioning scheme in Fig. 2.2, which may support virtual-circuit connections. The grooming capability of the node architecture in Fig. 2.3 is limited by the number of output ports of SONET ADM, and the size and the functionality of the DXC.

2.4.1 Multi-Hop Traffic Grooming

In this section, we assume that a connection can traverse multiple lightpaths before it reaches the destination. So, a connection may be groomed with different connections on different lightpaths. By extending the work in [Mukherjee, 1997, Banerjee and Mukherjee, 1997], which defines the collection of lightpaths in a WDM mesh network to form a virtual topology, we formulate the problem as an optimization problem. We will use the following notations in our mathematical formulation:

1 m and n denote endpoints of a physical fiber link that might occur in a lightpath.

2 i and j denote originating and terminating nodes for a lightpath. A lightpath may traverse single or multiple physical fiber links.

3 s and d denote source and destination of an end-to-end traffic request. The end-to-end traffic may traverse through a single lightpath or multiple light-

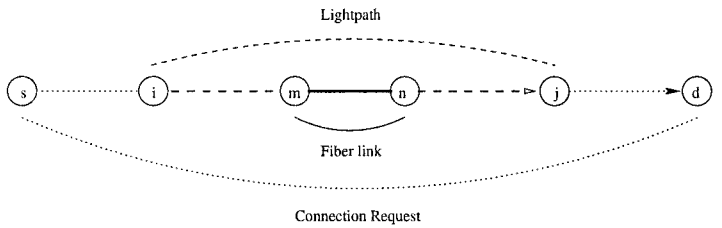


Figure 2.4. Illustrative example of a fiber link, a lightpath, and a connection request.

paths. Figure 2.4 shows how an end-to-end connection request may be carried.

4 y denotes the granularity of low-speed traffic requests. We assume $y \in \{1, 3, 12, 48\}$, which means that traffic demands between node pairs can be any of $OC-1$, $OC-3$, $OC-12$, and $OC-48$.

5 t denotes the index of $OC-y$ traffic request for any given node pair (s, d). For example, if there are ten $OC-1$ requests between node pair (s, d), then $t \in [1, 10]$.

- Given:
  - N: Number of nodes in the network.
  - W: Number of wavelengths per fiber. We assume all of the fibers in the network carry the same number of wavelengths.
  - P_{mn}: Number of fibers interconnecting node m and node n. $P_{mn} = 0$ for node pair which is not physically adjacent to each other. $P_{mn} = P_{nm} = 1$ if and only if there exists a direct physical fiber link between nodes m and n.
  - P_{mn}^{w}: Wavelength w on fiber P_{mn}. $P_{mn}^{w} = P_{mn}$.
  - TR_i: Number of transmitters at node i.
  - RR_i: Number of receivers at node i. Note that, in this set of ILP formulation, we assume all the nodes are equipped with tunable transceivers, which can be tuned to any of W wavelengths.
  - C: Capacity of each channel (wavelength).
  - Λ: Traffic matrix set. $\Lambda = \{\Lambda_y\}$, where y can be any allowed low-speed streams, 1, 3, 12, etc. In our study, $y \in \{1, 3, 12, 48\}$. $\Lambda_{y,sd}$ is the number of $OC-y$ connection requests between node pair (s, d).
- Variables:

- Virtual topology:
  - V_{ij}: Number of lightpaths from node i to node j in virtual topology. $V_{ij} = 0$ does not imply that $V_{ji} = 0$, i.e., lightpaths may be unidirectional.
  - V_{ij}^{w}: Number of lightpaths from node i to node j on wavelength w. Note that, if $V_{ij}^{w} > 1$, the lightpaths between node i and j on wavelength w may take different paths.
- Physical topology route:
  - $P_{mn}^{ij,w}$: Number of lightpaths between nodes (i, j) routed through fiber link (m, n) on wavelength w.
- Traffic route:
  - $\lambda_{ij,y}^{sd,t}$: The t^{th} $OC-y$ low-speed traffic request from node s to node d employing lightpath (i, j) as an intermediate virtual link.
  - $S_{sd}^{y,t}$: $S_{sd}^{y,t} = 1$ if the t^{th} $OC-y$ low-speed connection request from node s to node d has been successfully routed; otherwise, $S_{sd}^{y,t} = 0$.

- Optimize: Maximize the total successfully-routed low-speed traffic, i.e.,

$$Maximize : \sum_{y,s,d,t} y \cdot S_{sd}^{y,t} \tag{2.1}$$

- Constraints:
  - On virtual-topology connection variables:

$$\sum_{j} V_{ij} \quad \leq \quad TR_i \quad \forall i \tag{2.2}$$

$$\sum_{i} V_{ij} \quad \leq \quad RR_j \quad \forall j \tag{2.3}$$

$$\sum_{w} V_{ij}^{w} \quad = \quad V_{ij} \quad \forall i, j \tag{2.4}$$

$$int \quad V_{ij}, V_{ij}^{w} \tag{2.5}$$

  - On physical route variables:

$$\sum_{m} P_{mk}^{ij,w} \quad = \quad \sum_{n} P_{kn}^{ij,w} \quad if \quad k \neq i, j \quad \forall i, j, w, k \tag{2.6}$$

$$\sum_{m} P_{mi}^{ij,w} \quad = \quad 0 \qquad \forall i, j, w \tag{2.7}$$

$$\sum_{n} P_{jn}^{ij,w} \quad = \quad 0 \qquad \forall i, j, w \tag{2.8}$$

$$\sum_{n} P_{in}^{ij,w} = V_{ij}^{w} \quad \forall i,j,w \tag{2.9}$$

$$\sum_{m} P_{mj}^{ij,w} = V_{ij}^{w} \quad \forall i,j,w \tag{2.10}$$

$$\sum_{i,j} P_{mn}^{ij,w} \leq P_{mn}^{w} \quad \forall m,n,w \tag{2.11}$$

$$P_{mn}^{ij,w} \in \{0,1\} \tag{2.12}$$

- On virtual-topology traffic variables:

$$\sum_{i} \lambda_{id,y}^{sd,t} = S_{sd}^{y,t}$$

$$\forall s,d \quad y \in \{1,3,12,48\} \quad t \in [1,\Lambda_{y,sd}] \tag{2.13}$$

$$\sum_{j} \lambda_{sj,y}^{sd,t} = S_{sd}^{y,t}$$

$$\forall s,d,t \quad y \in \{1,3,12,48\} \quad t \in [1,\Lambda_{y,sd}] \tag{2.14}$$

$$\sum_{i} \lambda_{ik,y}^{sd,t} = \sum_{j} \lambda_{kj,n}^{sd,t} \quad if \quad k \neq s,d \quad \forall s,d,k,t \tag{2.15}$$

$$\sum_{i} \lambda_{is,y}^{sd,t} = 0$$

$$\forall s,d \quad y \in \{1,3,12,48\} \quad t \in [1,\Lambda_{y,sd}] \tag{2.16}$$

$$\sum_{j} \lambda_{dj,y}^{sd,t} = 0$$

$$\forall s,d \quad y \in \{1,3,12,48\} \quad t \in [1,\Lambda_{y,sd}] \tag{2.17}$$

$$\sum_{y,t} \sum_{s,d} y \times \lambda_{ij,y}^{sd,t} \leq V_{ij} \times C \quad \forall i,j \tag{2.18}$$

$$S_{sd}^{y,t} \in \{0,1\} \tag{2.19}$$

- Explanation of equations: The above equations are based on principles of conservation of flows and resources (transceivers, wavelengths, etc.).
  - Equation (2.1) shows the optimization objective function.

- Equations (2.2)–(2.3) ensure that the number of lightpaths between node pair (i, j) is less than or equal to the number of transmitters at node i and the number of receivers at node j.
- Equation (2.4) shows that the lightpaths between (i, j) are composed of lightpaths on different wavelengths between node pair (i, j). Note that the value of V_{ij}^{w} can be greater than 1. For example, in Fig. 2.1, two lightpaths on the same wavelength w can be set up between node 0 and node 5. One follows route $(0, 1, 2, 5)$, while the other follows route $(0, 3, 4, 5)$.
- Equations (2.6)–(2.10) are the multicommodity equations (flow conservation) that account for the routing of a lightpath from its origin to its termination. The flow-conservation equations have been formulated in two different ways [Ramaswami and Sivarajan, 1996]: (i) disaggregate formulation and (ii) aggregate formulation. In the disaggregate formulation, every i-j (or s-d) pair corresponds to a commodity, while in the aggregate formulation, all the traffic that is "sourced" from node i (or node s) corresponds to the same commodity, regardless of the traffic's destination. We employ the disaggregate formulation for the flow-conservation equations since it properly describes the traffic requests between different node pairs. Note that Equations (2.6)–(2.10) employ the wavelength-continuity constraint on the lightpath route.
  - Equation (2.6) ensures that, for an intermediate node k of lightpath (i, j) on wavelength w, the number of incoming lightpath streams is equal to the number of outgoing lightpath streams.
  - Equation (2.7) ensures that, for the origin node i of lightpath (i, j) on wavelength w, the number of incoming lightpath streams is 0.
  - Equation (2.8) ensures that, for the termination node j of lightpath (i, j) on wavelength w, the number of outgoing lightpath streams is 0.
  - Equation (2.9) ensures that, for the origin node i of lightpath (i, j) on wavelength w, the number of outgoing lightpath streams is equal to the total number of lightpaths between node pair (i, j) on wavelength w.
  - Equation (2.10) ensures that, for the termination node j of lightpath (i, j) on wavelength w, the number of incoming lightpath streams is equal to the total number of lightpaths between node pair (i, j) on wavelength w.
- Equations (2.11)–(2.12) ensure that wavelength w on one fiber link (m, n) can only be present in at most one lightpath in the virtual topology.

- Equations (2.13)–(2.19) 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 (channel) capacity.

2.4.2 Single-Hop Traffic Grooming

In this section, we assume that a connection can only traverse a single lightpath, i.e., only end-to-end traffic grooming is allowed. The formulation of the single-hop traffic-grooming problem is similar to the formulation of the multi-hop traffic-grooming problem, which was presented in the previous section, except for routing of connection requests on the virtual topology. We only present the different part as follows:

- On virtual-topology traffic variables:

$$\sum_{y,t} y \times S_{sd}^{y,t} \leq V_{sd} \times C \quad \forall s,d \tag{2.20}$$

$$S_{sd}^{y,t} \in \{0,1\} \tag{2.21}$$

2.4.3 Formulation Extension for Fixed-Transceiver Array

The mathematical formulations in the previous two sections are based on the assumption that the transceivers in a network node are tunable to any wavelength. If fixed-transceiver arrays are used at every network node and if M denotes the number of fixed-transceiver arrays used at each node, we can easily extend our formulation as follows:

- On virtual-topology connection variables:

$$\sum_{j} V_{ij}^{w} \leq M \quad \forall i,w \tag{2.22}$$

$$\sum_{i} V_{ij}^{w} \leq M \quad \forall j,w \tag{2.23}$$

$$\sum_{w} V_{ij}^{w} = V_{ij} \quad \forall i,j \tag{2.24}$$

$$int \quad V_{ij}, V_{ij}^{w} \tag{2.25}$$

The other parts of the formulations in the previous two sections are still the same. Equations (2.22)–(2.23) ensure that the number of lightpaths between node pair (i,j) on wavelength w is less than or equal to the number of transmitters at node i and the number of receivers at node j on wavelength w (every fixed-transceiver array only has one transceiver on each wavelength).

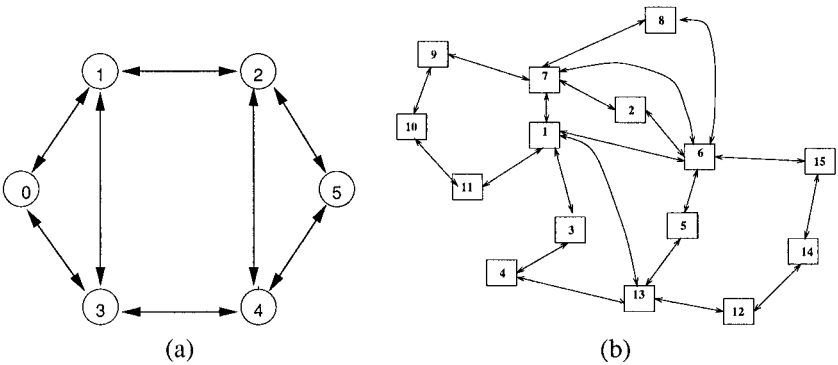


Figure 2.5. (a) A 6-node network and (b) a 15-node network.

2.4.4 Computational Complexity

It is well known that the RWA optimization problem is *NP*-complete [Mukherjee, 1997]. If we assume that each connection request requires the full capacity of a lightpath, then the traffic-grooming problem we are studying becomes the standard RWA optimization problem. It is easy to see that the traffic-grooming problem in a mesh network is also a *NP*-complete problem since the RWA optimization problem is *NP*-complete. As the number of variables and equations increases exponentially with the size of network and the number of wavelengths on each fiber, we use a small network topology as an example for obtaining ILP result. For large networks, we will use heuristic approaches.

2.5 Illustrative Numerical Results From ILP Formulations

This section presents numerical examples of the traffic-grooming problem using Fig. 2.5(a) as our physical topology. The traffic matrices are randomly generated. As an example, we allow the traffic demand to be any one of OC-1, OC-3, and OC-12. The traffic matrices are generated as follows: 1) the number of OC-1 connection requests between each node pair is generated as an uniformly-distributed random number between 0 to 16; 2) the number of OC-3 connection requests between each node pair is generated as an uniformly-distributed random number between 0 to 8; and 3) the number of OC-12 connection requests between each node pair is generated as an uniformly-distributed random number between 0 to 2. These three traffic matrices are shown in Tables 2.1 through 2.3, and the total traffic demand turns out to be the equivalent of OC-988. The capacity of each wavelength (channel) is OC-48.

Table 2.4 shows the corresponding result for the network throughput obtained from a commercial ILP solver, "CPLEX", based on different network

Table 2.1. Traffic matrix of OC-1 connection requests.

	Node 0	*Node 1*	*Node 2*	*Node 3*	*Node 4*	*Node 5*
Node 0	0	5	4	11	12	9
Node 1	0	0	8	5	16	6
Node 2	14	12	0	9	6	16
Node 3	4	11	15	0	1	5
Node 4	10	2	3	3	0	9
Node 5	2	1	8	15	13	0

Table 2.2. Traffic matrix of OC-3 connection requests.

	Node 0	*Node 1*	*Node 2*	*Node 3*	*Node 4*	*Node 5*
Node 0	0	6	2	1	5	4
Node 1	8	0	8	6	7	8
Node 2	1	3	0	0	2	7
Node 3	5	7	3	0	2	6
Node 4	6	4	5	0	0	2
Node 5	5	4	4	2	0	0

Table 2.3. Traffic matrix of OC-12 connection requests.

	Node 0	*Node 1*	*Node 2*	*Node 3*	*Node 4*	*Node 5*
Node 0	0	1	1	1	0	0
Node 1	1	0	1	1	0	2
Node 2	0	1	0	2	1	0
Node 3	2	0	2	0	2	0
Node 4	1	2	0	2	0	1
Node 5	1	1	2	2	2	0

resource parameters. In Table 2.4, T denotes the number of transceivers and W denotes the number of wavelengths. In the single-hop case, we only allow a connection to transverse a single lightpath, which means that only end-to-end traffic-grooming (multiplexing) is allowed. In the multi-hop case, a connection is allowed to traverse multiple lightpaths, i.e., a connection can be dropped at intermediate nodes and groomed with other low-speed connections on different lightpaths before it reaches its destination. Figure 2.1(b) showed a multi-hop grooming case, where connection 3 traversed two lightpaths: it was groomed with connection 1 on lightpath $(0, 2)$ and with connection 2 on lightpath $(2, 4)$.

Table 2.4. Throughput and number of lightpaths established (total traffic demand is OC-988).

	Multi-hop		*Single-hop*	
	Throughput	*Lightpath #*	*Throughput*	*Lightpath #*
T=3, W=3	74.7% (OC-738)	18	68.0% (OC-672)	18
T=4, W=3	93.8% (OC-927)	24	84.1% (OC-831)	24
T=5, W=3	97.9% (OC-967)	28	85.7% (OC-847)	24
T=7, W=3	97.9% (OC-967)	28	85.7% (OC-847)	24
T=3, W=4	74.7% (OC-738)	18	68.0% (OC-672)	18
T=4, W=4	94.4% (OC-933)	24	84.7% (OC-837)	24
T=5, W=4	100% (OC-988)	29	95.5% (OC-944)	28

As expected, the multi-hop case leads to higher throughput than the single-hop case.

We can see from Table 2.4 that, when the number of tunable transceivers at each node is increased from 3 to 5, the network throughput increases significantly, both in multi-hop case and in single-hop case. But when the number of tunable transceivers at each node increases from 5 to 7, the network throughput does not improve. This is because there are not enough wavelengths to setup more lightpaths to carry the blocked connection requests. Some illustrative results of transceiver and wavelength utilization for the multi-hop case are shown in Tables 2.5 and 2.6.

In the multi-hop case, when the transceiver is not a limited resource compared with wavelength, more short lightpaths may be set up to carry connections through multiple lightpaths, but this scenario is less likely in the single-hop case. This is shown in Table 2.4 where $T = 5$, $W = 3$ and $W = 4$. When $T = 5$, $W = 4$, if multi-hop grooming is allowed, the network throughput is 100%; otherwise, some connections get blocked. In multi-hop case, 29 lightpaths are established compared with 28 lightpaths in the single-hop case.

Tables 2.5–2.6 show some results for the node transceiver utilization and link wavelength utilization for the multi-hop case. When the number of transceivers is increased (from 3 to 5), the overall wavelength utilization is increased, as shown in Table 2.6. This is because more lightpaths have been established to carry the connection requests, shown in Table 2.4. As we mentioned before, when most of the links have fully utilized the available wavelengths, increasing the number of transceivers (from 5 to 7) will not help to improve the network throughput and will result in lower transceiver utilization, shown in Table 2.5 ($T = 7$ and $W = 3$).

Table 2.7 shows the virtual topology and the lightpath capacity utilization for the multi-hop case, when $T = 5$ and $W = 3$. As we can see, most of the lightpaths have high capacity utilization (above 90%). There are some node

Table 2.5. Results: transceiver utilization (multi-hop case).

		T=3, W=3	*T=5, W=3*	*T=7, W=3*
Node 0	Transmitter	100%	100%	71.4%
	Receiver	100%	100%	71.4%
Node 1	Transmitter	100%	100%	71.4%
	Receiver	100%	100%	71.4%
Node 2	Transmitter	100%	100%	71.4%
	Receiver	100%	100%	71.4%
Node 3	Transmitter	100%	100%	71.4%
	Receiver	100%	100%	71.4%
Node 4	Transmitter	100%	80%	57.4%
	Receiver	100%	80%	57.4%
Node 5	Transmitter	100%	80%	57.4%
	Receiver	100%	80%	57.4%

Table 2.6. Results: wavelength utilization (multi-hop case).

	T=3, W=3	*T=5, W=3*	*T=7, W=3*
Link (0,1)	33.3%	100%	100%
Link (0,3)	100%	100%	100%
Link (1,0)	100%	100%	100%
Link (1,2)	100%	100%	100%
Link (1,3)	33.3%	66.7%	66.7%
Link (2,1)	100%	100%	100%
Link (2,4)	66.7%	100%	100%
Link (2,5)	66.7%	100%	100%
Link (3,0)	33.3%	100%	100%
Link (3,1)	100%	66.7%	66.7%
Link (3,4)	66.7%	100%	100%
Link (4,2)	66.7%	100%	100%
Link (4,3)	66.7%	100%	100%
Line (4,5)	66.7%	66.7%	66.7%
Link (5,2)	66.7%	100%	100%
Link (5,4)	66.7%	66.7%	66.7%

pairs ($(0,1)$, $(1,3)$, etc.), which have multiple lightpaths set up, though the aggregate traffic between them can be carried by a single lightpath. The extra lightpaths are used to carry multi-hop connection traffic.

Table 2.7. Result: virtual topology and lightpath utilization (multi-hop case with T=5 and W=3).

	Node 0	*Node 1*	*Node 2*	*Node 3*	*Node 4*	*Node 5*
Node 0	0	2 (70%)	0 (100%)	1 (89%)	1 (100%)	1 (100%)
Node 1	1 (100%)	0	1 (100%)	2 (100%)	1 (100%)	0
Node 2	1 (100%)	1 (95%)	0	1 (100%)	1 (100%)	1 (70%)
Node 3	2 (100%)	1 (100%)	1 (100%)	0	0	1 (100%)
Node 4	1 (100%)	1 (100%)	0	0	0	1 (91%)
Node 5	0 (100%)	0	2 (98%)	1 (100%)	1 (100%)	0

In the ILP formulation, we treat the low-speed connection requests separately. The results from the ILP solutions show that, if there is a lightpath set up between (s, d), the low-speed connections between (s, d) tend to be packed on this lightpath channel directly. Based on this observation, we propose two simple heuristic algorithms for solving the traffic-grooming problem in large networks.

2.6 Heuristic Approach

The optimization problem of traffic-grooming is *NP*-complete. It can be partitioned into the following four subproblems, which are not necessarily independent:

1 Determine a virtual topology, i.e., determine the number of lightpaths between any node pair.

2 Route the lightpaths over the physical topology.

3 Assign wavelengths optimally to the lightpaths.

4 Route the low-speed connection requests on the virtual topology.

2.6.1 Routing

The routing and wavelength assignment problem (RWA) has received a lot of attention in the WDM networking literature. The current well-known routing approaches are fixed routing, fixed-alternate routing, and adaptive routing [Zang et al., 2000].

In fixed routing, the connections are always routed through a pre-defined fixed route for a given source-destination pair. One example of such an approach is fixed shortest-path routing. The shortest-path route for each source-destination pair is calculated off-line using standard shortest-path algorithms, such as Dijkstra's algorithm. If there are not enough resources to satisfy a connection request, the connection gets blocked.

In fixed-alternate routing, multiple fixed routes are considered when a connection request comes. In this approach, each node in the network is required to maintain a routing table that contains an ordered list of a number of fixed routes to each destination node. For example, these routes can be the first-shortest-path, the second-shortest-path, etc. When a connection request comes, the source node attempts to establish the connection on each of the routes from the routing table in sequence, until the connection is successfully established. Since fixed-alternate routing provides simplicity of control for setting up and tearing down connections, it is also widely used in the dynamic connection-provisioning case. It has been shown that, for certain networks, having as few as two alternate routes provides significantly lower blocking than having full wavelength conversion at each node with fixed routing [Ramamurthy and Mukherjee, 2002].

In adaptive routing, the route from a source node to a destination node is chosen dynamically, depending on the current network state. The current network state is determined by the set of all connections that are currently in progress [Zang et al., 2000]. For example, when a connection request arrives, the current shortest path between the source and the destination is calculated based on the available resources in the network; then the connection is established through the route. If a connection gets blocked in the adapting-routing approach, it will also be blocked in the fixed-alternate routing approach. Since each time a new connection request comes to a node, the route needs to be calculated based on the current network state, adaptive routing will require more computation and a longer setup time than fixed-alternate routing, but it is also more flexible than fixed-alternate routing.

In our heuristics, we will use adaptive routing.

2.6.2 Wavelength Assignment

Once the route has been chosen for each lightpath, the number of lightpaths traversing a physical fiber link defines the congestion on that particular link. With the wavelength-continuity constraint, we need to assign wavelengths to each lightpath such that any two lightpaths passing through the same physical link are assigned different wavelengths. We assume a single-fiber network system. There is only one fiber in each direction if two nodes are connected. Ten wavelength-assignment approaches have been compared in [Zang et al., 2000], and all of them were found to perform similarly. We will use one simple approach, First-Fit (FF). In FF, all wavelengths are numbered. When searching for an available wavelength, a lower-numbered wavelength is considered before a higher-numbered wavelength. The first available wavelength is then selected. The idea behind this simple scheme is that it tries to pack all of the in-use wavelengths towards the lower end of the wavelength space.

2.6.3 Heuristics

We propose two heuristic algorithms for the traffic-grooming problem. Let $T(s,d)$ denote the aggregate traffic between node pair s and d, which has not been successfully carried. Let $t(s,d)$ denote one connection request between s and d, which has not been successfully carried yet. Let C denote the wavelength capacity.

- *Maximizing Single-Hop Traffic (MST).* The basic idea of this heuristic is introduced in [Mukherjee, 1997] for the traditional virtual-topology design problem. This simple heuristic attempts to establish lightpaths between source-destination pairs with the highest $T(s,d)$ values, subject to constraints on the number of transceivers at the two end nodes, and the availability of a wavelength in the path connecting the two end nodes. The connection requests between s and d will be carried on the new established lightpath as much as possible. If there is enough capacity in the network, every connection will traverse a single lightpath hop. If there are not enough resources to establish enough lightpaths, the algorithm will try to carry the blocked connection requests using currently available spare capacity of the virtual topology. The pseudo-code for this heuristic is shown in Algorithm 2.1.

- *Maximizing Resource Utilization (MRU).* Let $H(s,d)$ denote the hop distance on physical topology between node pair (s,d). Define $T(s,d)/H(s,d)$ as the connection resource utilization value, which represents the average traffic per wavelength link. This quantity shows how efficiently the resources have been used to carry the traffic requests. This heuristic tries to establish the lightpaths between the node pairs with the maximum resource utilization values. When no lightpath can be set up, the remaining blocked traffic requests will be routed on the virtual topology based on their connection resource utilization value $t(s,d)/H^{'}(s,d)$, where $t(s,d)$ denotes a blocked connection request, and $H^{'}(s,d)$ denotes the hop distance between s and d on the virtual topology. The pseudo-code for this heuristic follows the same steps as the pseudo-code of MST, except that the node pairs and blocked connections are sorted based on their resource utilization values.

Both heuristic algorithms have two stages. Based on our observations from the ILP results, we find that packing different connections between the same node pair within the same existing lightpath, which directly joins the end points, is a very efficient grooming scheme. In the first stage, the algorithms try to establish lightpaths as much as possible to satisfy the aggregate end-to-end connection requests. If there are enough resources in the network, every connection request will be successfully routed through a single lightpath hop. This will minimize the traffic delay since the optical signals need not be converted into

Algorithm 2.1 MST

1 Construct virtual topology:

   (a) Sort all of the node pairs (s, d) according the sum of uncarried traffic request $T(s, d)$ between (s, d) and put them into a list L in descending order.

   (b) Try to setup a lightpath between the first node pair (s', d') in L using first-fit wavelength assignment and shortest-path routing, subject to the wavelength and transceiver constraints. If it fails, delete (s', d') from L; otherwise, let $T(s, d) = Max[T(s, d) - C, 0]$ and go to Step 1a until L is empty.

2 Route the low-speed connections on the virtual topology constructed in Step 1.

   (a) Satisfy all of the connection requests which can be carried through a single lightpath hop, and update the virtual topology network state.

   (b) Route the remaining connection requests based on the current virtual topology network state, in the descending order of the connections' bandwidth requirement.

electronic domain. In the second stage, the spare capacities of the currently established lightpath channels are used to carry the connection requests blocked in the first stage, and the algorithms give single-hop groomable connections higher priority to be satisfied.

2.6.4 Heuristic Results and Comparison

Table 2.8 shows the comparison between the results obtained from an ILP solver and the heuristic algorithms for the six-node network in Fig. 2.5(a). We can observe that the MST and MRU heuristic algorithms show reasonable performance when compared with the results obtained from the ILP solver. The heuristic approaches have much less computation complexity than the ILP approach. The two proposed algorithms are relatively simple and straightforward; by using other RWA algorithms instead of adaptive routing and first-fit wavelength assignment, it may be possible to develop other complex heuristic algorithms to achieve better performance.

Figures 2.6–2.8 show the results from the two heuristic algorithms, when applied to the larger network topology in Fig. 2.5(b). The traffic matrices follow the same distribution as we mentioned in Section 2.5.

In Fig. 2.6, we plot the network throughput versus the number of wavelengths on every fiber link when each node is equipped with 10 tunable transceivers. We observe that the MRU heuristic performs better than the MST algorithm with respect to network throughput. Since the number of tunable transceivers at each node is limited (10 in this case), when the number of wavelengths on each fiber link reaches a certain value (16 in this case), increasing the number of wavelengths does not help to increase the network throughput.

Table 2.8. Throughput results comparison between ILP and heuristic algorithms (total traffic demand is OC-988).

	Multi-hop (ILP)	*Single-hop (ILP)*	*Heuristic (MST)*	*Heuristic (MRU)*
T=3, W=3	74.7% (OC-738)	68.0% (OC-672)	71.0% (OC-701)	67.4% (OC-666)
T=4, W=3	93.8% (OC-927)	84.1% (OC-831)	89.4% (OC-883)	93.6% (OC-925)
T=5, W=3	97.9% (OC-967)	85.7% (OC-847)	94.4% (OC-933)	94.4% (OC-933)
T=7, W=3	97.9% (OC-967)	85.7% (OC-847)	94.4% (OC-933)	94.4% (OC-933)
T=3, W=4	74.7% (OC-738)	68.0% (OC-672)	71.0% (OC-701)	67.4% (OC-666)
T=4, W=4	94.4% (OC-933)	84.7% (OC-837)	93.1% (OC-920)	93.6% (OC-925)
T=5, W=4	100% (OC-988)	95.5% (OC-944)	100% (OC-988)	100% (OC-988)

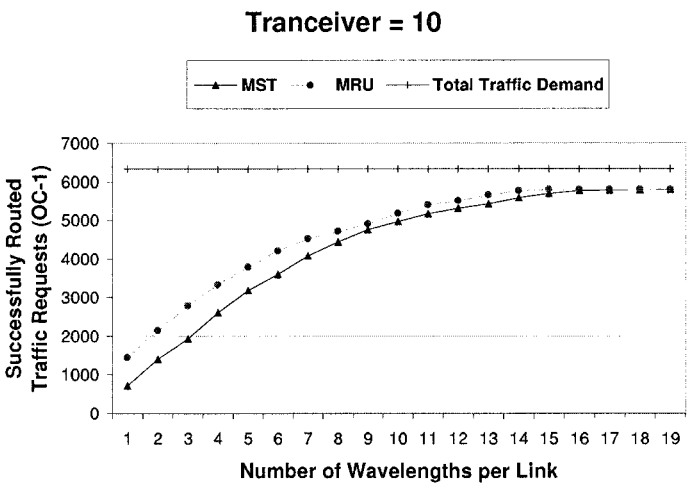


Figure 2.6. Network throughput vs. number of wavelengths for the network topology in Fig. 2.5(b) with 10 tunable transceivers at each node.

In Fig. 2.7, we plot the network throughput versus the number of transceivers at every node when each fiber link carries ten wavelengths. We compare the performance of the two heuristic algorithms. The results show that, when the number of transceivers is small (≤ 7 in this case), MST performs better than MRU. This is because MRU tries to utilize wavelengths efficiently. When the number of transceivers is small, wavelength is not the limiting resource in the network any more. So maximizing wavelength utilization will not help to improve the performance.

Figure 2.8 compares the performance using tunable transceiver and fixed transceiver in every network node. Each node is equipped with 12 transceivers

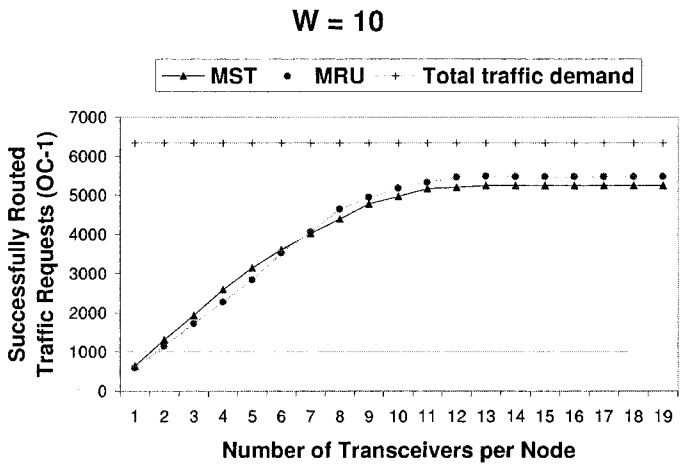


Figure 2.7. Network throughput vs. number of tunable transceivers for the network topology in Fig. 2.5(b) with 10 wavelengths on each fiber link.

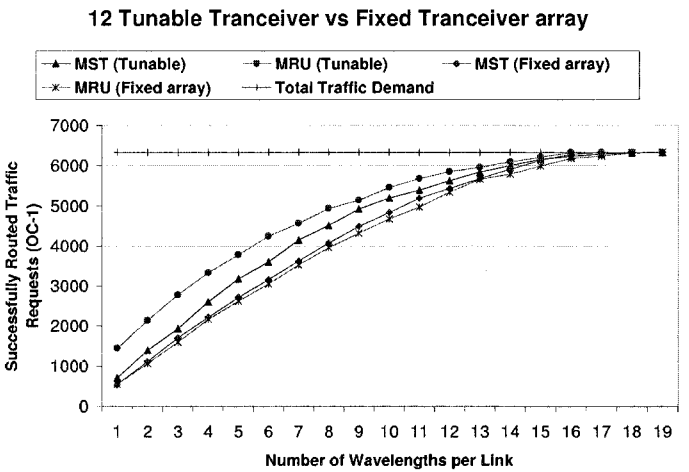


Figure 2.8. Network throughput vs. number of wavelengths (size of fixed transceiver array) for the network topology in Fig. 2.5(b) with 12 tunable transceivers at each node.

if we use tunable transceiver. Each node is equipped with one transceiver array if we use fixed transceivers and the size of the transceiver array is equal to the number of wavelengths supported by every fiber link. The results in Fig. 2.8 indicate

that, when nodes are equipped with the same number of transceivers, tunable-transceiver architecture has better performance. For the fixed-transceiver case, MST performs better than MRU.

2.7 Mathematical Formulation Extension for Other Optimization Criteria

In this section, we show how to extend our ILP formulations to handle different optimization criteria for the traffic-grooming problem.

2.7.1 Extension for Network Revenue Model

It has been shown earlier that the low-speed connection requests between the same node pair tend to be packed together on to the same lightpath channel. The connections which can be carried by a single lightpath channel are more likely to be satisfied than the connections which have to traverse multiple lightpaths, when they have the same bandwidth requirement and the optimization objective is to maximize network throughput. To make the problem more realistic, it is reasonable for us to assume that two connection requests may have different priority, even if they have the same bandwidth requirement. This is because different connection requests may have different end-node distance, quality-of-service requirement, etc. A connection's priority can be represented by a "weight" associated with it. In this section, we assume that the weight is determined by the bandwidth requirement and end-node distance of the connection request. For a given network topology and traffic demand, the objective is to maximize the weighted network throughput. Let W_i denote the weight of connection i, D_i denote the end-node distance of connection i, and C_i denote the bandwidth requirement of connection i. The connection's weight function is defined as:

$$W_i = D_i \times C_i^{\alpha} \tag{2.26}$$

where $0 \leq \alpha \leq 1$ and D_i is measured by the shortest-path distance of the connection's end nodes on the physical topology. Equation (2.26) is called the power-law cost function [Kleinrock, 1970]. It is used to study the actual tariffs demanded by communications services for high-speed telecommunication channels, and there is effectively a "quantity discount" (controlled by α) in that capacity cost (per unit of channel capacity) decreases as the capacity increases. Thus, the network weighted throughput becomes:

$$T = \sum_{i=1}^{K} D_i \cdot C_i^{\alpha} \cdot S_i \tag{2.27}$$

where $S_i = 1$ if connection request i has been satisfied; otherwise $S_i = 0$, and the total number of connection requests is K. T can also be called *"network*

Table 2.9. Results of comparison between revenue model and network throughput model.

	Optimize Revenue		*Optimize Throughput*
	Revenue	*Throughput*	*Throughput*
T=3, W=3	83.7%	72.4%	74.7%
T=5, W=3	98.5%	97.2%	97.9%
T=7, W=3	98.5%	97.2%	97.9%
T=3, W=4	83.7%	72.4%	74.7%
T=4, W=4	94.3%	91.7%	94.4%
T=5, W=4	100%	100%	100%

revenue". We can easily modify our ILP formulation to optimize T. The only part of the equations which should be modified is shown as follows:

- Optimize: Maximize network revenue.

$$Maximize : \sum_{y,s,d,t} D_{sd} \cdot y^{\alpha} \cdot S_{sd}^{y,t} \tag{2.28}$$

where D_{sd} denotes the distance between node pair (s, d).

2.7.2 Illustrative Results

In this section, we show some illustrative results to optimize network revenue using our ILP formulation extension. We use the same network topology and traffic matrix set as in Section 2.5. In Equation (2.28), D_{sd} is measured by the shortest-path hop distance between node s and d on the physical topology, and α is equal to 0.8.

Table 2.9 compares the results between the two optimization models. In Table 2.9, T denotes the number of tunable transceivers per node and W denotes the number of wavelengths per fiber link. Multi-hop grooming is allowed in both models. It is shown that, in the revenue model, when $T = 3$ and $W = 3$, the maximal achievable revenue is 83.7%, and 72.4% of traffic requests have been satisfied to achieve the revenue, while the maximal achievable traffic load the network can carry is 74.7%. In revenue model, we find that if there is a lightpath set up between (s, d), it may first be used to carry some long multi-hop connections (with higher weight) which will traverse this lightpath as an intermediate hop. Thus, some connections directly between (s, d) may be blocked. This means that packing different connections between the same node pair within the same existing lightpath, which directly joins the end points, is not a good grooming scheme any more. We find that, because of the quantity-discount parameter α in Equation (2.26), lower-speed connections are more

likely to be satisfied than higher-speed connection requests. It is obvious that different heuristics are needed based on the different optimization criteria.

2.8 Conclusion

This study was devoted to the traffic-grooming problem in a WDM mesh network. We studied the architecture of a node with grooming capability. We presented the ILP formulation for traffic-grooming in such a WDM mesh network. We compared the performance of the single-hop grooming approach and multi-hop grooming approach on a small six-node network with randomly generated traffic pattern. Results from ILP showed that the end-to-end aggregate traffic between the same node pair tends to be groomed on to the same lightpath channel, which directly joins the end points, if the optimization objective is to maximize the network throughput. Two heuristic approaches were also proposed for solving the traffic-grooming problem in large networks. We compared the performance of these two heuristic algorithms, MST and MRU, with different network resource parameters. The comparison results showed that MRU performs better if tunable transceivers are used and MST performs better if fixed transceivers are used. We extended the optimization problem to a network-revenue model and found a different grooming scheme, which can be used to design an efficient heuristic algorithm on the network-revenue model. We showed that, with proper extension, our ILP mathematical model can be used to develop good grooming schemes for different models. These schemes can be used to design efficient heuristic algorithms, which are practical for large and realistic networks.

Chapter 3

A GENERIC GRAPH MODEL

3.1 Introduction

As discussed in Chapter 2, traffic grooming is usually divided into four sub-problems, which are not necessarily independent: (1) determining the virtual topology that consists of lightpaths, (2) routing the lightpaths over the physical topology, (3) performing wavelength assignment to the lightpaths, and (4) routing the traffic on the virtual topology. The virtual-topology design problem [Chlamtac et al., 1992, Mukherjee et al., 1996, Ramaswami and Sivarajan, 1996, Ganz and Wang, 1994, Krishnaswamy and Sivarajan, 2001, Gerstel et al., 1999] is conjectured to be NP-hard [Mukherjee, 1997]. In addition, routing and wavelength assignment (RWA) [Zang et al., 2000] is also NP-hard [Chlamtac et al., 1993]. Therefore, traffic grooming in a mesh network is also a NP-hard problem [Zhu and Mukherjee, 2002b].

To solve the traffic-grooming problem, one approach is to deal with the four sub-problems separately. It first determines the virtual topology, then performs routing and wavelength assignment, and finally routes the traffic requests. There are considerable research results on each sub-problem already and they can be utilized to solve the traffic-grooming problem. Although this divide-and-conquer method makes traffic grooming easier to handle, it cannot achieve the optimal solution even if we can get the optimal solution for each sub-problem, since these four sub-problems are not necessarily independent and the solution to one sub-problem might affect how optimally another sub-problem can be solved. Sometimes, using the optimal solution for one sub-problem might not lead to the optimal solution to the whole problem. Moreover, this approach requires all the traffic requests to be known in advance, which cannot be satisfied in dynamic grooming.

Another approach is to solve the four sub-problems as a whole. Since it can take into account all the constraints regarding the four sub-problems simultaneously, this approach has a potential to achieve better performance. With static traffic, the traffic-grooming problem can be formulated as an integer linear program (ILP) (Chapter 2), and an optimal solution can be obtained for some relatively small networks. However, an ILP is not scalable and cannot be directly applied to large networks. One way to make the problem tractable is to develop heuristic algorithms and jointly solve the grooming problem for one connection request at a time. To the best of our knowledge, no integrated heuristic algorithm for solving the traffic-grooming problem has been developed for wavelength-routed networks in previous work.

The work in [Konda and Chow, 2001] formulates the static traffic-grooming problem as an ILP and proposes a heuristic to minimize the number of transceivers. In [Brunato and Battiti, 2002], several lower bounds for regular topologies are presented and greedy and iterative greedy schemes are developed. However, in both [Konda and Chow, 2001] and [Brunato and Battiti, 2002], the authors relax the physical-topology constraints, assuming all the virtual topologies are implementable on the given physical topology, i.e., they do not consider lightpath routing and wavelength assignment.

3.1.1 Challenges of Traffic Grooming in a Heterogeneous WDM Mesh Network

The WDM backbone network is expected to emerge as a multi-vendor, heterogeneous network. As WDM networks migrate from ring topologies to mesh topologies, it is very important to solve the traffic-grooming problem in a heterogeneous mesh network environment.

In terms of wavelength-conversion capability, heterogeneity means that some of the nodes in a network may have full wavelength-conversion capability (any incoming wavelength can be converted into any outgoing wavelength), some may have no wavelength-conversion capability (traffic must stay on the same wavelength when bypassing these nodes) [Subramaniam et al., 1996, Iness and Mukherjee, 1999], and some may have partial wavelength-conversion capability (some wavelengths can be converted into some other wavelengths) [Yates et al., 1996, Ramaswami and Sasaki, 1998, Venugopal et al., 1999, Tripathi and Sivarajan, 2000]. In previous work, however, it was assumed that all the nodes in a network either have wavelength-conversion capability or none has wavelength-conversion capability. In addition, if a node has this capability, it always has full wavelength-conversion capability. This all-or-nothing assumption may not be practical or valid in the future WDM network. It is necessary to address the partial and sparse wavelength-conversion scenarios.

In a WDM mesh network, each node must support two functionalities: wavelength routing, which can be accomplished by an optical crossconnect (OXC),

and optical multiplexing/demultiplexing, by which several wavelengths can be multiplexed to or demultiplexed from the same fiber-link. Besides, in order to groom low-speed connections onto a high-speed wavelength channel, a node will need to employ access stations, which can multiplex/demultiplex and switch low-speed connections using various multiplexing techniques, e.g., time-division multiplexing (TDM). A WDM mesh network may consist of systems from multiple vendors, and different vendors may employ different node architectures, which may have different grooming capabilities. Some architectures may have full grooming capabilities, while others may impose some constraints on the grooming capability, such as the number of transceivers used for originating and terminating groomable wavelength channels (also known as grooming ports). In addition, some nodes may have no grooming capability. These partial and sparse grooming-capability scenarios are very practical and should also be considered when solving the traffic-grooming problem.

To solve the traffic-grooming problem, the integrated approach is desirable not only because it has the potential to achieve better performance than the separated approach, but also because it can be used directly for dynamic traffic grooming, where the separated approach cannot be used. For a given connection request, the integrated approach should address the following issues: (1) Should this connection be routed on the current virtual topology if it is possible to do so? Sometimes, it may be better to set up a new lightpath even though the connection can be carried on the current virtual topology. (2) How to change the virtual topology to accommodate the connection? i.e., between which two nodes should we set up a new lightpath, if any? In some cases, we can set up a lightpath directly from the source of the traffic to the destination. In other cases, it is not necessary or possible to set up this lightpath and we may need to set up one or more lightpaths and route the connection onto these lightpaths and/or some existing lightpaths. Different decisions on these questions can result in different network performance. These decisions reflect the intentions of the network operator, and they are referred to as *grooming policies* [Zhu and Mukherjee, 2002a]. By using different grooming policies, a network operator can achieve various objectives, such as minimizing the number of wavelength-links, minimizing the number of lightpaths, minimizing the traffic hops on the virtual topology, etc. As the network state changes, the optimization objective may also need to change. Dynamically evolving the grooming policy according to the network state is also a challenge for traffic grooming. Dynamic traffic grooming in a WDM mesh network is addressed in Chapter 4.

3.1.2 Contributions of this Chapter

In this chapter, we propose a novel, generic graph model for traffic grooming in a heterogeneous WDM mesh network. In this model, various factors of heterogeneity of the network, such as the number of transceivers at each node, the

number of wavelengths on each fiber-link, as well as wavelength-conversion capabilities and grooming capabilities of each node, are represented by different edges in an auxiliary graph constructed by our model. Besides, this model can achieve various objectives using different grooming policies. Moreover, instead of designing a route-computation algorithm for each grooming policy, simple shortest-path route-computation algorithms can be used in this model to achieve various grooming policies by carefully choosing the weight functions for the edges in the auxiliary graph. Three different grooming policies are proposed and their performance is compared under blocking and non-blocking scenarios. Based on the auxiliary graph, we develop an integrated traffic-grooming algorithm which jointly solves the four traffic-grooming sub-problems. The integrated traffic-grooming algorithm can be applied to both static and dynamic traffic grooming. In static grooming, proper selection of the traffic requests is key to achieving a good network performance. We present several traffic-selection schemes based on this model and evaluate their performance for different network topologies.

The rest of the chapter is organized as follows. In Section 3.2, we demonstrate how to construct, according to the network state, an auxiliary graph, which is the basis of our graph model. Based on this, an integrated traffic-grooming algorithm and three selection schemes used by the algorithm for static traffic are proposed and an illustrative example is given in Section 3.3. In Section 3.4, the grooming policy is analyzed and three different grooming policies are proposed. Methods to choose the weight-assignment functions for the auxiliary graph to achieve these policies are also discussed. In Section 3.5, the performance of different grooming policies is shown for blocking and non-blocking scenarios. The performance of the three selection schemes used by the integrated traffic-grooming algorithm for static traffic is also compared under different network scenarios. Section 3.6 concludes the chapter.

3.2 Construction of an Auxiliary Graph

In order to solve the traffic-grooming problem, we first construct an auxiliary graph according to the given network configuration.

An illustrative example is shown in Fig. 3.1. In order to make the constructed auxiliary graph clear to see, we choose a very simple network topology. Network 1 (Fig. 3.1(a)) is a three-node network with four unidirectional fiber-links, each of which has two wavelengths. Node 0 has wavelength converters with full wavelength-conversion capability, node 1 has no wavelength converter, and node 2 has wavelength converters with limited wavelength-conversion capability in the sense that only wavelength λ_1 can be converted to λ_2. In the beginning, there is no lightpath in the network, so there is no edge in the virtual topology of Network 1, as shown in Fig. 3.1(b). An auxiliary graph is constructed as in Fig. 3.1(c).

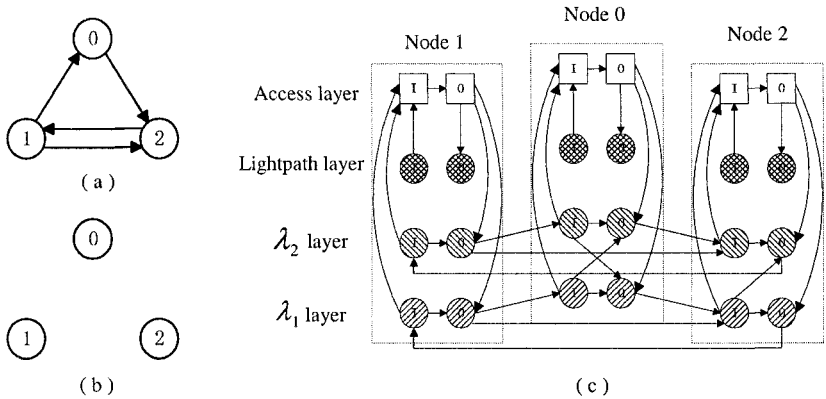


Figure 3.1. (a) Physical topology of Network 1. (b) Virtual topology of Network 1. (c) Auxiliary graph of Network 1.

In general, a network can be represented by a graph $G_0(V_0, E_0)$, where V_0 and E_0 are its node set and link set, respectively. Assuming that each link has W wavelengths, λ_1 through λ_W, we construct the corresponding auxiliary graph as follows.

Below, for clarity, we will use the terms *node* and *link* to represent a vertex and an edge, respectively, in the original network $G_0(V_0, E_0)$, and we will use the terms *vertex* and *edge* to represent a vertex and an edge in the auxiliary graph $G(V, E)$, respectively.

Auxiliary graph G is a layered graph with $(W + 2)$ layers (for the problem treatment in this chapter, although an alternate number of layers is quite possible for various generalization of the problem). Layers 1 through W denote the W *wavelength layers*, layer $(W+1)$ is called the *lightpath layer*, and layer $(W+2)$ is called the *access layer*, where a traffic flow starts and terminates. Each node has two ports on each layer, an input port and an output port. Let $N_i^{l,p}$ denote port p on layer l at node i; then $V = \{N_i^{l,p} \mid p \in \{0,1\},\ 1 \le l \le W+2,\ \forall i \in V_0\}$, where $N_i^{l,0}$ and $N_i^{l,1}$ denote the input port and the output port on layer l at network node i, respectively. Each edge in the auxiliary graph G has a property tuple $PT(c, w)$ associated with it, where c denotes the capacity of this edge and w denotes its weight. The edges are inserted in auxiliary graph G as follows.

- **Wavelength Bypass Edges (WBE).**
  There is an edge from the input port to the output port on each wavelength layer at node i:

$$\langle N_i^{l,0},\ N_i^{l,1} \rangle \in E \qquad \forall i \in V_0,\ 1 \le l \le W \tag{3.1}$$

We call the edge $\langle N_i^{l,0}, N_i^{l,1} \rangle$ *wavelength bypass edge* on layer l at node i and it is denoted as $WBE(i,l)$. The capacity of the edge is ∞.

- **Grooming Edges (GrmE).**
  There is an edge from the input port to the output port on access layer at node i if node i has grooming capability:

$$\langle N_i^{W+2,0}, N_i^{W+2,1} \rangle \in E \qquad \forall i \in V_0 \tag{3.2}$$

  We call the edge $\langle N_i^{W+2,0}, N_i^{W+2,1} \rangle$ *grooming edge* at node i and it is denoted as $GrmE(i)$. The capacity of the edge is ∞.

- **Mux Edges (MuxE).**
  There is an edge from the output port on the access layer to the output port on the lightpath layer at each node:

$$\langle N_i^{W+2,1}, N_i^{W+1,1} \rangle \in E \qquad \forall i \in V_0 \tag{3.3}$$

  We call the edge $\langle N_i^{W+2,1}, N_i^{W+1,1} \rangle$ *mux edge* at node i and it is denoted as $MuxE(i)$. The capacity of the edge is ∞.

- **Demux Edges (DmxE).**
  There is an edge from the input port on the lightpath layer to the input port on the access layer at each node:

$$\langle N_i^{W+1,0}, N_i^{W+2,0} \rangle \in E \qquad \forall i \in V_0 \tag{3.4}$$

  We call the edge $\langle N_i^{W+1,0}, N_i^{W+2,0} \rangle$ *demux edge* at node i and it is denoted as $DmxE(i)$. The capacity of the edge is ∞.

- **Transmitter Edges (TxE).**
  There is an edge from the output port on the access layer to the output port on wavelength layer l if there are transmitters available on wavelength λ_l at node i:

$$\langle N_i^{W+2,1}, N_i^{l,1} \rangle \in E \qquad \forall i \in V_0,\ 1 \le l \le W,\ TX_i^l > 0 \tag{3.5}$$

  where $TX_i^l (1 \le l \le W)$ denotes the number of transmitters that can operate at wavelength λ_l at node i.

  We call the edge $\langle N_i^{W+2,1}, N_i^{l,1} \rangle$ *transmitter edge* on layer l at node i and it is denoted as $TxE(i,l)$. The capacity of the edge is ∞.

- **Receiver Edges (RxE).**
  There is an edge from the input port on wavelength layer l to the input port

on the access layer if there are receivers available on wavelength λ_l at node i:

$$\langle N_i^{l,0},\ N_i^{W+2,0}\rangle \in E \qquad \forall i \in V_0,\ 1 \le l \le W,\ RX_i^l > 0 \tag{3.6}$$

where $RX_i^l (1 \le l \le W)$ denotes the number of receivers that can operate at wavelength λ_l at node i.

We call the edge $\langle N_i^{l,0},\ N_i^{W+2,0}\rangle$ *receiver edge* on layer l at node i and it is denoted as $RxE(i,l)$. The capacity of the edge is ∞.

- **Converter Edges (CvtE).**
  There is an edge from the input port on wavelength layer l_1 to the output port on wavelength layer l_2 at node i if wavelength λ_{l_1} can be converted to wavelength λ_{l_2} without using an access station at node i:

$$\langle N_i^{l_1,0},\ N_i^{l_2,1}\rangle \in E$$
$$\forall i \in V_0,\ \text{wavelength } l_1 \text{ is convertible to } l_2 \text{ at node } i \tag{3.7}$$

  We call the edge $\langle N_i^{l_1,0},\ N_i^{l_2,1}\rangle$ *converter edge* from layer l_1 to layer l_2 at node i and it is denoted as $CvtE(i,l_1,l_2)$. The capacity of the edge is ∞.

- **Wavelength-Link Edges (WLE).**
  There is an edge from the output port on wavelength layer l at node i to the input port on wavelength layer l at node j if there is a physical link from node i to node j and wavelength λ_l on this link is not used:

$$\langle N_i^{l,1},\ N_j^{l,0}\rangle \in E$$
$$(i,\ j) \in E_0\ ,\ \text{wavelength } \lambda_l \text{ on link } (i,\ j) \text{ is not used} \tag{3.8}$$

  We call the edge $\langle N_i^{l,1},\ N_j^{l,0}\rangle$ *wavelength-link edge* on layer l from node i to node j and it is denoted as $WLE(i,j,l)$. The capacity of this edge is the capacity of the corresponding wavelength on the link from node i to node j.

- **Lightpath Edges (LPE).**
  There is an edge from the output port on the lightpath layer at node i to the input port on the lightpath layer at node j if there is a lightpath from node i to node j:

$$\langle N_i^{W+1,1},\ N_j^{W+1,0}\rangle \in E$$
$$\exists \text{ a lightpath from node } i \text{ to node } j \tag{3.9}$$

  We call the edge $\langle N_i^{W+1,1},\ N_j^{W+1,0}\rangle$ *lightpath edge* from node i to node j and it is denoted as $LPE(i,j)$. The capacity of this edge is the residual capacity of the corresponding lightpath from node i to node j.

As a final step in constructing the auxiliary graph, we need to assign weights to each edge, i.e., determine w in the property tuple $PT(c, w)$ of each edge. The weights can reflect the cost of each network element (transceiver, wavelength-link, wavelength converter, etc.), and/or a certain grooming policy. By applying different weight settings, this graph model can be used to achieve different objectives. These weights can either be fixed, or they can be adjusted according to the current network state. We will discuss the weight function in Section 3.4.2.

Note that, for each edge, we can keep some other useful, edge-specific information in the property tuple also. For instance, for each lightpath edge, the routing and wavelength assignment information can be saved in the property tuple.

From the above procedure, it should be clear that the auxiliary graph reflects the current state of the network, which can be heterogeneous, with different nodes having different resources and capabilities.

3.3 Solving the Traffic-Grooming Problem Based on the Auxiliary Graph

Traffic grooming is usually divided into four sub-problems:

- determine the virtual topology of the network, i.e., which nodal transmitter should be directly connected to which nodal receivers,
- route the lightpaths over the physical topology,
- assign wavelengths to the lightpaths (this problem has been shown to be NP-hard in [Chlamtac et al., 1993] and there are various heuristics to solve it [Zang et al., 2000]), and
- route the traffic on the virtual topology.

In Chapter 2, these four sub-problems were usually solved separately. For instance, routing the traffic can only be done after the virtual topology has been selected. This approach has a drawback since these four sub-problems are not necessarily independent. It does not take into account the impact of routed traffic on the network state when determining the virtual topology. In addition, it cannot combine the knowledge about the physical-topology layer and virtual-topology layer to determine the route of the traffic. In our study, based on the auxiliary graph, we propose an integrated algorithm which jointly solves the four sub-problems. Since it can take advantage of all the updated information about the sub-problems, this approach has a potential to achieve very significant improvement in performance.

3.3.1 The IGABAG Algorithm

We first introduce the Integrated Grooming Algorithm Based on the Auxiliary Graph (IGABAG), which solves the four sub-problems for one traffic demand, and then we provide its complexity analysis.

Algorithm

The IGABAG algorithm needs initialization before being used. The initialization takes as a parameter the network configuration, which includes network topology, as well as node and link configuration; and according to the network configuration, it constructs the corresponding auxiliary graph G using the method discussed in Section 3.2.

The input of the IGABAG algorithm is a traffic demand, which is represented by $T(s, d, g, m)$, where s and d are the source and destination nodes, respectively; g is the granularity of the traffic demand, for instance, OC-3 or OC-48; and m is the amount of the traffic in units of g. The algorithm works as shown in Algorithm 3.1.

It can be observed that the IGABAG algorithm routes a given traffic request under the current network state and updates the network state after routing, making the auxiliary graph always reflect the current network state.

Complexity Analysis

Suppose there are N nodes in the network and each link has W wavelengths. In the corresponding auxiliary graph, there are $2 \times N \times (W + 2)$ vertices. Since the running time of shortest-path computation using Dijkstra algorithm is $O(V^2)$, where V is the number of the vertices in the graph, the running time of IGABAG algorithm is $O(N^2W^2)$. If each node in the network has full wavelength-conversion capability, all the wavelength layers can be collapsed into one wavelength layer since all the wavelengths are equivalent. In this special case, the running time of IGABAG algorithm is $O(N^2)$.

3.3.2 The INGPROC Procedure and Traffic-Selection Schemes

The IGABAG algorithm is used to accommodate one connection request, but in traffic-grooming problems, we need to route a set of requests. Based on the IGABAG algorithm, we propose the INtegrated Grooming PROCedure (INGPROC) to solve the traffic-grooming problem, shown in Algorithm 3.2. The input of INGPROC includes network configuration and a set of traffic requests.

Note that the INGPROC procedure can be applied to both static and dynamic grooming. For dynamic grooming, INGPROC just chooses the current traffic request in Step 2. For static grooming, where all the traffic demands are

Algorithm 3.1 IGABAG

Input: a traffic demand $T(s, d, g, m)$.

1 Delete the edges whose capacity is less than the bandwidth granularity of T, since they cannot accommodate T.

2 Find the shortest path p from the output port on the access layer of the source to the input port on the access layer of the destination of T on graph G. If not successful, restore the edges previously deleted in Step 1 and return -1.

3 If p contains wavelength-link edges, one or more lightpaths going through the corresponding wavelength-links needs to be set up. A lightpath starts whenever p travels through a transmitter edge, follows the subsequent wavelength-link edges, and terminates at the first receiver edge.

4 Route T along the pre-existing lightpaths in p and/or lightpaths set up according to p. If the capacity of the path, which is defined as the minimum capacity of the lightpaths along the path, is less than the entire amount of T, route the maximum amount possible, say n units, of the traffic granularity g.

5 Restore the edges previously deleted in Step 1.

6 Update graph G as follows:

   (a) For each lightpath newly set up, a lightpath edge from the output port of starting node of the lightpath to the input port of ending node is added on the lightpath layer.

   (b) The wavelength-link edges denoting the wavelength-links used by the lightpath are removed from the corresponding wavelength layers. Note that, if there are multiple fiber-links between the nodes, the wavelength-link edges are removed only when the corresponding wavelengths on all the fiber-links are used. So, this algorithm can also be used in the case where there are multiple fiber-links between the same node pair.

   (c) If there is no transmitter (receiver) available at node i on wavelength λ_l, the transmitter edge $TxE(i, l)$ (receiver edge $RxE(i, l)$) will be removed from G, i.e., this node cannot source/sink a lightpath on wavelength λ_l any more and can only be bypassed by a lightpath.

   (d) If there is no wavelength converter which can convert wavelength λ_{l_1} to wavelength λ_{l_2} available at node i, the converter edge $CvtE(i, l_1, l_2)$ will be removed from G.

   (e) Update the property tuple $PT(c, w)$ of the edges. For the lightpaths carrying the traffic T, the capacities of the lightpath edges denoting the lightpaths carrying the traffic T are decreased by the amount of the traffic routed. Updating the weights of the edges in the graph will change the grooming policies. We will discuss the grooming policies in Section 3.4.

7 If the entire traffic is accommodated, return 0. Otherwise, return $m - n$, which is the amount of the uncarried traffic in units of g.

known in advance, the order in which the requests are routed plays an important role in achieving good performance. We propose the following traffic-request-selection schemes for static traffic grooming.

Algorithm 3.2 INGPROC

Input: network configuration and a set of traffic requests.

1 Initialize IGABAG with network configuration.

2 Select a traffic demand $T(s, d, g, m)$ from the traffic-request set.

3 Apply IGABAG to T, and let k denote the return value.

4 If $k > 0$, insert $T(s, d, g, k)$ into the request set.

5 Go to Step 2 unless all the traffic has been routed, or no traffic can be routed with the remaining network resources.

- *Least Cost First (LCF).* LCF chooses the most cost-effective traffic request under the current network state and routes it. The cost of a traffic request is the weight of the shortest path for routing the traffic on the corresponding auxiliary graph divided by the amount of the traffic, which is computed as the granularity multiplied by the units of the traffic. Note that, after routing a connection, LCF needs to re-compute the cost of the unrouted connections under the updated network state. If there are N nodes in the network, W wavelengths on each link, and D traffic demands, the running time of INGPROC using LCF is $O(D^2N^2W^2)$, assuming no wavelength-conversion capability, or $O(D^2N^2)$, assuming full wavelength-conversion capability.

- *Maximum Utilization First (MUF).* MUF selects the connection with the highest utilization, which is defined as the total amount of the request divided by the number of hops from the source to the destination on the physical topology. The running time of INGPROC using MUF is $O(D \log D + DN^2W^2)$, assuming no wavelength-conversion capability, or $O(D \log D + DN^2)$, assuming full wavelength-conversion capability[1].

- *Maximum Amount First (MAF).* MAF selects the connection with the largest amount of demand and routes it. The running time of INGPROC using MAF is $O(D \log D + DN^2W^2)$, assuming no wavelength-conversion capability, or $O(D \log D + DN^2)$, assuming full wavelength-conversion capability.

We will compare the performance of these traffic-selection schemes in Section 3.5.

[1]Here, we use comparison sorts, such as heapsort and merge sort, to determine the order of the connections, whose running time is $O(D \log D)$. A linear-time sorting algorithm, such as counting sort, radix sort, and bucket sort, can also be applied to determine the order of the connections. Then, the running time of INGPROC is $O(DN^2W^2)$ and $O(DN^2)$, assuming no wavelength-conversion capability and full wavelength-conversion capability, respectively.

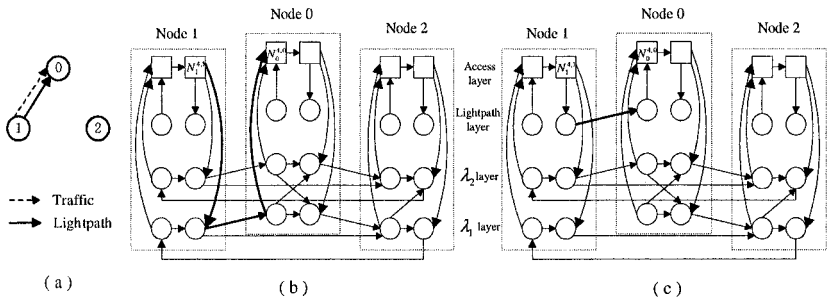


Figure 3.2. (a) Virtual topology of Network 1. (b) Corresponding auxiliary graph before routing the first traffic request T_1. (c) Corresponding auxiliary graph after routing the first traffic request T_1.

In dynamic grooming, connections arrive one at a time, hold for a certain time period, and terminate. When a connection terminates, the resource used for this connection must be released.

3.3.3 An Illustrative Example

To illustrate how the graph model and the IGABAG algorithm work, we consider an example based on the network in Fig. 3.1. Suppose the capacity of each wavelength is OC-48 and each node has grooming capability and two tunable transceivers.

The first connection request T_1 is $T(1, 0, \text{OC-12}, 2)$. To satisfy this request, we need to find in the auxiliary graph a path from $N_1^{4,1}$ to $N_0^{4,0}$. It is easy to see that there exists a path along the edges $TxE(1,1)$, $WLE(1,0,1)$, and $RxE(0,1)$, shown as bold lines in Fig. 3.2(b). Since this path contains a wavelength-link edge $WLE(1,0,1)$, which denotes a wavelength-link, we need to set up a lightpath L_1 using λ_1 on the fiber-link from node 1 to node 0. After setting up L_1, we need to add a lightpath edge $LPE(1,0)$ into the graph, which means that there is a lightpath from node 1 to node 0. Meanwhile, the wavelength-link edge $WLE(1,0,1)$ must be removed from the graph since this wavelength-link cannot be used to set up another lightpath later on. This connection T_1 then can be routed onto lightpath L_1 and the residual capacity of L_1 is 2 $\times$ OC-12. So the capacity of edge $LPE(1,0)$ is 24, which means that the capacity is equivalent to 24 OC-1s. The current virtual topology and the updated auxiliary graph are shown in Figs. 3.2(a) and 3.2(c), respectively.

Suppose the second connection request T_2 is $T(2, 0, \text{OC-12}, 1)$. Following the same procedure as above, we need to determine a path from $N_2^{4,1}$ to $N_0^{4,0}$. There exist several paths in the auxiliary graph.

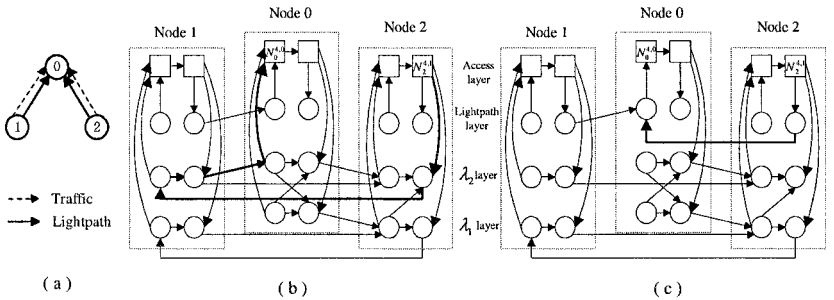


Figure 3.3. (a) Virtual topology of Network 1. (b) Corresponding auxiliary graph before routing the second traffic request T_2. (c) Corresponding auxiliary graph after routing the second traffic request T_2 using single-hop grooming.

- Case 1 (Single-hop grooming). One path is along the edges $TxE(2,2)$, $WLE(2,1,2)$, $WBE(1,2)$, $WLE(1,0,2)$, and $RxE(0,2)$, shown as bold lines in Fig. 3.3(b). This path contains edges $WLE(2,1,2)$ and $WLE(1,0,2)$, which denote wavelength λ_2 on the fiber-links from node 2 to node 1 and from node 1 to node 0, respectively. If this path is chosen, a lightpath L_2 consisting of these two wavelength-links needs to be set up. As a result, a lightpath edge $LPE(2,0)$ is added into the graph and the two wavelength-link edges $WLE(2,1,2)$ and $WLE(1,0,2)$ are removed from the graph. Since both receivers at node 0 are used, we remove all the receiver edges, i.e., edges $RxE(0,1)$ and $RxE(0,2)$, which means that node 0 cannot sink lightpaths any more. After the traffic is routed onto lightpath L_2, the capacity of lightpath edge $LPE(2,0)$ is 36 units. In this case, we set up one lightpath using two wavelength-links. Since the connection traverses a single lightpath, we call this approach *single-hop grooming*. Figures 3.3(a) and 3.3(c) show the current virtual topology and the updated auxiliary graph, respectively.

- Case 2 (Multi-hop grooming). Another path is along the edges $TxE(2,1)$, $WLE(2,1,1)$, $RxE(1,1)$, $GrmE(1)$, $MuxE(1)$, $LPE(1,0)$, and $DmxE(0)$, shown as bold lines in Fig. 3.4(b). This path contains edges $WLE(2,1,1)$ and $LPE(1,0)$, which denote wavelength λ_1 on the fiber-link from node 2 to node 1 and the lightpath from node 1 to node 0, respectively. If choosing this path, we need to set up a lightpath L_3 from node 2 to node 1 using wavelength λ_1 on the fiber-link from node 2 to node 1, and a lightpath edge $LPE(2,1)$ is added and wavelength-link edge $WLE(2,1,1)$ removed. Then, we route T_2 onto the newly setup lightpath L_3 and the pre-existing lightpath L_1. The capacities of lightpath edge $LPE(2,1)$ and $LPE(1,0)$ are 36 and 12, respectively. In this case, we have to route the connection onto

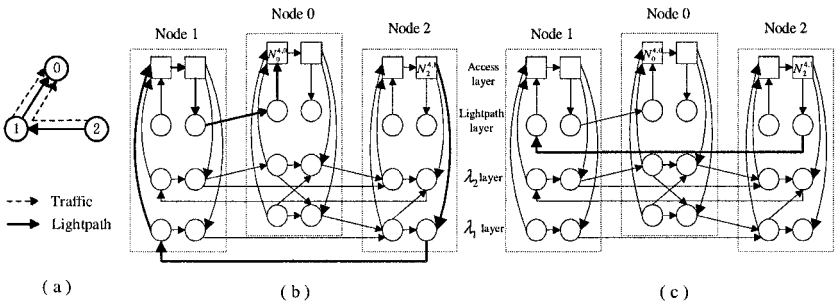


Figure 3.4. (a) Virtual topology of Network 1. (b) Corresponding auxiliary graph before routing the second traffic request T_2. (b) Corresponding auxiliary graph after routing the second traffic request T_2 using multi-hop grooming.

two lightpaths, but only one more wavelength-link is required for satisfying this traffic. Since the connection traverses multiple lightpaths, we call this approach *multi-hop grooming*. However, this kind of multi-hop grooming will add burden on the electrical devices, which are the bottleneck and major cost in a WDM network, at the intermediate node(s) (node 1 in this case). Figures 3.4(a) and 3.4(c) show the current virtual topology and the updated auxiliary graph, respectively.

Which path should be chosen depends on the grooming policy. Since the IGABAG algorithm chooses the shortest path, the grooming policy should be reflected in the weight-assignment function. We will discuss the grooming policies in Section 3.4.1.

Suppose the third traffic demand T_3 is $T(1, 0, \text{OC-}48, 1)$. If we use single-hop grooming for the second connection, we cannot find a path from $N_1^{4,1}$ to $N_0^{4,0}$ after removing all the lightpath edges since they cannot accommodate this traffic request and it will be blocked. However, if we use multi-hop grooming for the second connection, we can still find a path in the graph since there is a wavelength λ_2 available which can be used to set up a lightpath from node 1 to node 0 to carry the T_3 traffic.

From the above example, it can be seen that the IGABAG algorithm can deal with the grooming problem in the blocking manner under the wavelength and transceiver constraints using different grooming policies. It is a straightforward matter to verify that our algorithm can be used in non-blocking scenarios if enough network resources are given for the traffic demands.

One of the advantages of the graph model is that, if a path is obtained from the source to the destination in the auxiliary graph, all the four sub-problems of traffic grooming are solved simultaneously. Therefore, it can avoid the limitations introduced by having to solve the four sub-problems separately.

3.4 Grooming Policies and Weight Assignment

A grooming policy determines how to carry the traffic in a certain situation. It reflects the intentions of the network operator. In this section, we first analyze all the possible operations when routing a traffic request. The different ordering of the possible operations forms different grooming policies. Then, we discuss how to assign weights to edges in the auxiliary graph to achieve different grooming policies.

3.4.1 Grooming Policies

When solving the traffic-grooming problem, given a traffic demand $T(s, d, g, m)$, we need to determine how to route the traffic under the current network state.

In general, for a traffic demand $T(s, d, g, m)$ in a network, there are four possible operations that can be used to carry the traffic without altering the existing lightpaths. Note that we do not consider reconfiguring existing lightpaths, such as splitting or rerouting a lightpath, since then the traffic on the network would be interrupted.

- *Operation 1*: Route the traffic onto an existing lightpath directly connecting the source s and the destination d.
- *Operation 2*: Route the traffic through multiple existing lightpaths.
- *Operation 3*: Set up a new lightpath directly between the source s and the destination d and route the traffic onto this lightpath. Using this operation, we set up only one lightpath if the amount of the traffic is less than the capacity of the lightpath.
- *Operation 4*: Set up one or more lightpaths that do not directly connect the source s and the destination d, and route the traffic onto these lightpaths and/or some existing lightpaths. Using this operation, we need to set up at least one lightpath. However, since some existing lightpaths may be utilized, the number of wavelength-links used to set up the new lightpaths is probably less than that of wavelength-links needed to set up a lightpath directly connecting the source s and the destination d.

The characteristics of these four operations are summarized in Table 3.1.

Each operation must satisfy certain prerequisites before it can be applied. For instance, if there is no lightpath between the source and the destination of the traffic that can accommodate the traffic, then Operation 1 cannot be used. In some situations, all the operations are applicable, while in other situations, only some of them can be used. If none of them can be applied, the traffic must be blocked without reconfiguring the existing lightpaths.

Table 3.1. Comparison of four operations.

	Add new lightpath(s)	*Single-hop or multi-hop grooming*
Operation 1	No	Single-hop
Operation 2	No	Multi-hop
Operation 3	Yes	Single-hop
Operation 4	Yes	Multi-hop

In a situation where multiple operations can be applied, how to choose the operations is a matter of the grooming policy. By combining the various operations in different orders, we can achieve different grooming policies.

Here we present three different grooming policies. In each of them, for a given traffic demand $T(s, d, g, m)$, if there is a lightpath from s to d which can carry the traffic request, we always choose it since this is the best solution for the connection request, i.e., we always use Operation 1 when it is applicable.

- *Minimizing the Number of Traffic Hops (MinTH).*
  If Operation 1 fails, we always try to set up a lightpath from s to d and route the traffic onto this lightpath. Only when such a lightpath cannot be set up, we use multi-hop grooming. This policy is achieved by Operation 1 followed by Operation 3. After that, we will use Operations 2 and 4 and choose the one with fewer hops on the virtual topology.

- *Minimizing the Number of Lightpaths (MinLP).*
  This policy tries to set up the minimum number of *new* lightpaths to carry the traffic. If Operation 1 fails, we try to route the traffic using multiple existing lightpaths (Operation 2). If Operation 2 also fails, then we try to set up one or more lightpaths to accommodate the traffic using Operation 3 or Operation 4.

- *Minimizing the Number of Wavelength-Links (MinWL).*
  This policy tries to consume the minimum number of *extra* wavelength-links to carry the traffic. The difference between MinLP and MinWL is that, after Operations 1 and 2 fail, MinWL compares the number of wavelength-links used by Operations 3 and 4, and it chooses the one requiring fewer wavelength-links.

3.4.2 Weight Assignment

Policies MinTH, MinLP, and MinWL can be easily achieved by applying different weight-assignment functions to the graph model and using the IGABAG algorithm.

Since each grooming policy is achieved by combining the four operations in different ways, it is helpful to analyze the weight function of each operation first. Since IGABAG chooses the shortest path found in the auxiliary graph, the order in which the four operations are combined is determined by the relationship of the weight of each operation.

In the following discussion, the weight of each kind of edge is nonnegative and represented by the name of this kind of edge, e.g., the weight of a wavelength-link edge is represented by WLE. In addition, we assume that the same kind of edges has the same weight and there is no wavelength converter at each node, i.e., there is no converter edge in the auxiliary graph. It is straightforward to extend the discussion to more general cases.

- Operation 1 uses a single existing lightpath to route the traffic. Since each lightpath edge always has a mux and demux edge connected with it in the path, the weight of the path found in the auxiliary graph is:

$$MuxE + LPE + DmxE \tag{3.10}$$

- Operation 2 uses n ($n \geq 2$) existing lightpaths to carry the traffic. Since each lightpath edge always has a mux and demux edge connected with it and there is a grooming edge between two lightpaths, the weight of the path found in the auxiliary graph is:

$$n \times (MuxE + LPE + DmxE) + (n-1) \times GrmE \tag{3.11}$$

- Operation 3 sets up a lightpath between the source and the destination of the traffic and routes the traffic onto it. According to the IGABAG algorithm, the lightpath follows the path found in the graph, which consists of a transmitter edge, m ($m \geq 1$) wavelength-link edges, $m-1$ wavelength bypass edges, and a receiver edge. Therefore, the weight of the path found in the auxiliary graph is:

$$TxE + m \times WLE + (m-1) \times WBE + RxE \tag{3.12}$$

- Operation 4 sets up k ($k \geq 1$) lightpaths and routes the traffic onto them and k' ($k' \geq 0$) existing lightpaths. Supposing that each newly set up lightpath uses l_i ($l_i \geq 1, 1 \leq i \leq k$) wavelength-links, the weight of the path found in the auxiliary graph is:

$$\begin{aligned} &\sum_{i=1}^{k} (TxE + l_i \times WLE + (l_i - 1) \times WBE + RxE) \\ &\quad + k' \times (MuxE + LPE + DmxE) \\ &\quad + (k + k' - 1) \times GrmE \end{aligned} \tag{3.13}$$

To make Operation 1 the first choice among the four operations, we need to ensure that the value of Expression (3.10) is always the least among the four expressions.

Based on the analysis of the weight of each operation, we can easily manipulate the weights of the edges to satisfy the different grooming policies, as follows.

- MinTH policy tries to carry the current traffic request using the minimum number of lightpath hops on the virtual topology. It can be observed that, for each traffic, there is a grooming edge following each hop on the virtual topology except the last one. At the same time, if a grooming edge is encountered in the path, the traffic must experience a hop on the virtual topology. Therefore, minimizing the traffic hops is equivalent to minimizing the number of grooming edges in the path found by IGABAG. Hence, we need to assign a large weight to the grooming edges such that the weight of a path containing n $(n \geq 1)$ grooming edges is always greater than that of any path containing $n-1$ grooming edges. We call this kind of edges *dominant edges* in the graph.

- MinLP policy tries to set up the minimum number of new lightpaths for the current traffic request. For each newly set up lightpath, there must be a transmitter edge and a receiver edge in the path according to which the lightpath is set up. In addition, if there are n $(n \geq 0)$ transmitter edges and receiver edges in the path, n lightpaths must be set up according to IGABAG. Therefore, minimizing the number of lightpaths is equivalent to minimizing the number of transmitter edges and receiver edges. So, we should make transmitter edges and receiver edges dominant edges in the graph.

- MinWL policy tries to use as few unused wavelength-links as possible to accommodate the current traffic request. It is straightforward to see that we can achieve this policy by making wavelength-link edges dominant edges.

In grooming policies MinTH, MinLP, and MinWL, some operations are always applied before others, and these policies can be easily achieved by our graph model. However, this does not mean that the graph model can only apply the four operations in fixed orders. The order in which the four operations are performed by our model really depends on the weight assignment. If we appropriately assign weights to the edges in the auxiliary graph, the four operations can be applied in different orders at different times, which gives the network operator the maximum flexibility. For example, if we assign the weight to each edge according to the cost of the corresponding component, say, the weight of a TxE/RxE is the cost of a transmitter/receiver, the weight of a CvtE is the cost of a converter, the weight of a WLE is the cost of the corresponding wavelength-links, etc., the graph model will choose the most cost-effective

operation to route a connection, under the current network state. This intrinsic flexibility that the operations can be applied in different orders is one of the major advantages of our model.

Note that our model can also be used for virtual-topology design. Suppose we are given the physical fiber topology, the maximal configuration of the nodes, and the traffic demands to be supported. Now, if we assign the weight to each edge according to the cost of the corresponding component, after routing all the traffic using the graph model, we can determine the virtual topology and the configuration of each node, such as the number of transceivers and converters at each node. Hence, our model can be used for network design while minimizing its total cost.

In dynamic grooming, the network state varies as connection requests come and go. To achieve good performance, the grooming policy should be evolved according to the current network state. For instance, if transceivers are becoming scarce resource, we should make full use of existing lightpaths to accommodate the new traffic and avoid setting up new lightpaths. The graph model can easily satisfy this requirement by adjusting the weights of edges according the current network state, i.e., the weight of an edge can be made a function of the network state. This capability of easily changing grooming policies makes the graph model very suitable for dynamic traffic grooming. We will explore this in detail in Chapter 4.

3.5 Numerical Examples

In this section, we first compare the performance of the three grooming policies MinWL, MinLP, and MinTH under the blocking and the non-blocking scenarios. Then, we compare the performance of the INGPROC procedure when using different traffic-selection schemes LCF, MAF, and MUF with the optimal solution obtained via an integer linear program (ILP) in a relatively small network. Finally, the performance of the three traffic-selection schemes is investigated for a larger, practical-sized network to demonstrate how the network throughput changes under different configurations.

3.5.1 Comparison of Grooming Policies

We compare the performance of these three grooming policies via simulation. The topology we used is the NSF network topology, which has 14 nodes and 21 links, as shown in Fig. 3.5(a). Each link is bidirectional and carries 32 wavelengths, and the capacity of each wavelength is OC-192. All nodes have grooming capability and there are 32 tunable transceivers and no wavelength converter at each node.

The traffic is randomly generated and uniformly distributed among all node pairs. For each node pair, there may exist several types of connections simul-

Table 3.2. The average traffic generated for the NSF network.

	# of connections	*Amount in terms of OC-1*
OC-3	54	2689
OC-12	55	5641
OC-48	54	11309
OC-192	8	2131

taneously, for instance, OC-3, OC-12, OC-48, and OC-192. The distributions of each type of connections are independent. For this example, the traffic is generated as follows.

- For each node pair (i, j), the probability that there is an OC-3 type of connection between them is 0.3. If there is an OC-3 connection between i and j, the amount m of the traffic $T(i, j, \text{OC-3}, m)$ is uniformly distributed between 1 and 32.
- For each node pair (i, j), the probability that there is an OC-12 type of connection between them is 0.3. If there is an OC-12 connection between i and j, the amount m of the traffic $T(i, j, \text{OC-12}, m)$ is uniformly distributed between 1 and 16.
- For each node pair (i, j), the probability that there is an OC-48 type of connection between them is 0.3. If there is an OC-48 connection between i and j, the amount m of the traffic $T(i, j, \text{OC-48}, m)$ is uniformly distributed between 1 and 8.
- For each node pair (i, j), the probability that there is an OC-192 type of connection between them is 0.05. If there is an OC-192 connection between i and j, the amount m of the traffic $T(i, j, \text{OC-192}, m)$ is uniformly distributed between 1 and 2.

In our simulation experiments, 10 different traffic matrices randomly generated according to the above distribution are used, and the average traffic distribution is shown in Table 3.2. On average, the total number of connection requests in a traffic matrix is 171 and the total traffic amount is equivalent to 21770 OC-1s. The network resources are enough to carry all the traffic demands, so this is a non-blocking model and the objective is to minimize the resources used to carry all of the traffic.

In Step 2 of INGPROC, we use the traffic-selection heuristic Least Cost First (LCF) to choose the traffic demand, which is discussed in Section 3.3.2. In the experiments, we assign weights to the edges such that all the requirements for the grooming policies are satisfied. The weights of different edges assigned in

Table 3.3. The weights of edges assigned in the experiments for the three grooming policies.

	MinTH	*MinLP*	*MinWL*
WLE	10	10	1000
GrmE	1000	20	0
TxE	20	200	20
RxE	20	200	20
LPE	1	1	1
MuxE	0	0	0
DmxE	0	0	0
WBE	0	0	0

the experiments for the three grooming policies are shown in Table 3.3. Note that the grooming policies will be achieved as long as the relationship of the weights of the different edges satisfies the grooming-policy requirements, no matter what value a specific edge is assigned.

The results based on 10 simulation experiments are shown in Fig. 3.5(b). It can be observed that, to carry all the traffic demands, the MinWL policy consumes the fewest wavelength-links, the MinLP policy sets up the minimum number of lightpaths, and the MinTH policy achieves the minimum number of average traffic hops on the virtual topology. This demonstrates that the weight-assignment functions of the three policies can really accomplish the corresponding grooming policies. In addition, the MinWL policy sets up the most number of lightpaths and the traffic experiences the largest number of hops on the virtual topology. This is because this policy prefers to use short lightpaths to carry connections. Since each lightpath will occupy one transmitter and one receiver at the source node and the destination node, respectively, the MinLP policy uses the fewest transceivers. The MinTH policy consumes the largest number of wavelength-links since it always tries to set up a lightpath from the source to the destination when the connection cannot be routed using a single existing lightpath.

We also study the performance of the three policies under a blocking scenario. The same NSF network topology and the same 10 traffic matrices are used. However, each link now has only 8 wavelengths, and each node has 12 tunable transceivers. Since the network resources are reduced and not all the requests can be satisfied, the objective in the blocking scenario is to maximize the carried traffic, i.e., the throughput of the network.

The results in Fig. 3.5(c) demonstrate that MinTH achieves the highest throughput among the three policies. This is because, in a blocking scenario, the network resources are limited. To improve the network throughput, we should use our limited resources efficiently, and single-hop grooming is usually more

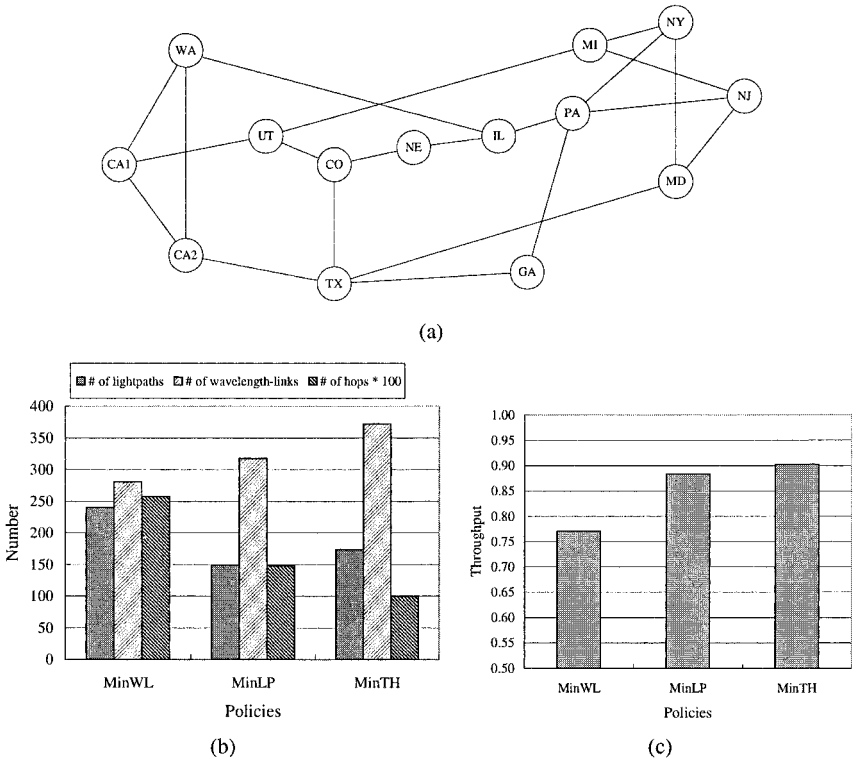


Figure 3.5. Comparison of different grooming policies. (a) NSF network. (b) Comparison of different grooming policies using a non-blocking model. (c) Comparison of different grooming policies using a blocking model.

efficient to use lightpath capacity to carry the traffic than multi-hop grooming. From the perspective of a traffic request, fewer hops mean that less resources (lightpaths) are used to accommodate the traffic. From the perspective of a lightpath from node s to node d, its efficiency is higher when using the same amount of capacity to carry the traffic whose source and destination are also s and d than to carry other traffic.

3.5.2 Comparison of Traffic-Selection Schemes in a Relatively Small Network

We compare the performance of our heuristics with the optimal solution obtained through an ILP (Chapter 2). Since the ILP can be solved only for small networks, for this comparison, we use a six-node network with bidirectional links shown in Fig. 3.6(a), and the traffic matrices are as follows. There are

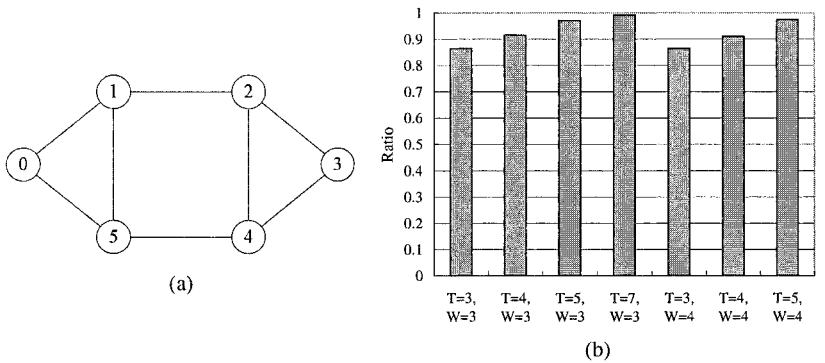


Figure 3.6. Comparison of traffic-selection schemes in a relatively small network. (a) Network 2: a 6-node network. (b) Average ratio of the amount of carried traffic by LCF to the amount of carried traffic by ILP.

three types of connections: OC-1, OC-3, and OC-12, and the amount of each connection type between each node pair is uniformly distributed between 0 to 16, 0 to 8, and 0 to 2, respectively. For our example, the total traffic amount becomes equivalent to 988 OC-1s. The capacity of each wavelength is OC-48. Each node has grooming capability with a limited number of transceivers and no wavelength converter. Since the network resources may not be enough to accommodate all the requests, our objective is to maximize the network throughput, i.e., the amount of successfully carried traffic. We use the MinTH policy in this experiment.

The results are shown in Table 3.4, where T denotes the number of transceivers at each node and W denotes the number of wavelengths per link. The numbers in the table are the percentage and the amount of the traffic routed using different traffic-selection schemes under different network configurations. We observe that the performance of LCF is better than those of the other two traffic-selection schemes in most cases and close or equal to the optimal solution.

To further compare the performance of our heuristics and that of the ILP, we tried 12 different traffic matrices with the same distribution as above. Figure 3.6(b) shows the average ratio of the amount of routed traffic by heuristic LCF to the amount of routed traffic by the ILP, under different network configurations. It can be observed that our heuristic can achieve a near-optimal solution, while using much less running time than the ILP. (The running time of LCF is less than 1 second on machine A[2] while it takes several minutes to more than 1 hour for the ILP to get the solutions on the same machine.)

[2]Machine A is a Windows PC with a 500-MHz Pentium III processor and 2-GB memory.

Table 3.4. Performance comparison of ILP and different heuristics for routing static traffic demands.

	ILP	*LCF*	*MUF*	*MAF*
T=3, W=3	76.7% (758)	67.9% (671)	65.9% (651)	65.0% (642)
T=4, W=3	95.7% (946)	90.9% (898)	85.3% (843)	86.6% (856)
T=5, W=3	96.9% (957)	96.5% (953)	89.5% (884)	94.0% (929)
T=7, W=3	96.9% (957)	96.8% (956)	96.9% (957)	96.9% (957)
T=3, W=4	76.7% (758)	67.9% (671)	65.9% (651)	65.0% (642)
T=4, W=4	96.4% (952)	95.5% (944)	85.3% (843)	85.8% (848)
T=5, W=4	100% (988)	100% (988)	100% (988)	100% (988)

Table 3.5. The traffic generated for Network 3.

	# of connections	*Amount in terms of OC-1*
OC-3	106	5175
OC-12	105	10272
OC-48	103	22368
OC-192	17	4877

3.5.3 Comparison of Traffic-Selection Schemes in a Larger Representative Network

We also examine the heuristics on a larger representative network (Network 3) shown in Fig. 3.7(a). This network has 19 nodes and 31 links. All the nodes have grooming capability and no wavelength converter. Each link is bidirectional, and each wavelength has a capacity of OC-192. Our experimental results are based on 10 different traffic matrices, which are randomly generated using the same method used for the NSF network. The average total number of connection requests is 330 and the average total traffic amount is equivalent to 42692 OC-1s. The distribution of the average generated traffic is shown in Table 3.5.

The blocking model and MinTH policy are used in this experiment. We vary the number of transceivers at each node and the number of wavelengths on each link to obtain the performance of the three traffic-selection schemes under different network configurations. Figures 3.7(b) and 3.7(c) show, when using the heuristics LCF, MUF, and MAF, how the network throughput changes as the number of transceivers at each node varies from 16 to 24, with the assumption that each fiber-link has 8 and 16 wavelengths, respectively. It can be observed that MUF performs better than MAF, and LCF performs best since LCF chooses

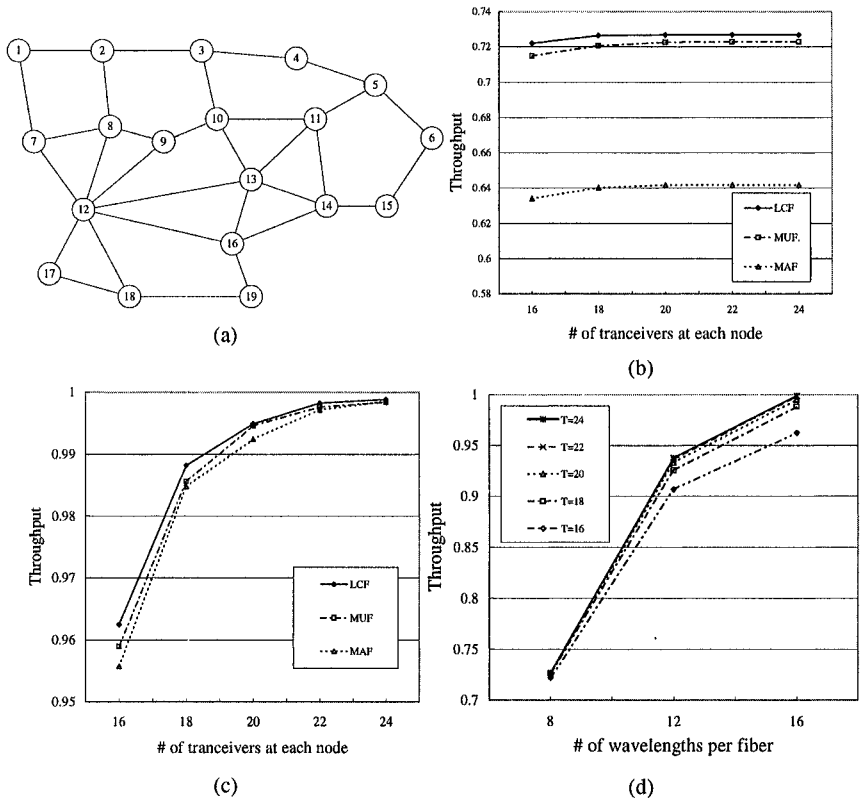


Figure 3.7. Comparison of traffic-selection schemes in a larger representative network. (a) Network 3: a 19-node network. (b) Network throughput using different heuristics when each link has 8 wavelengths. (c) Network throughput using different heuristics when each link has 16 wavelengths. (d) Network throughput using heuristic LCF under different network configurations.

the connection according to the current network state, while the other heuristics do not take this into account.

On the other hand, the time complexity of INGPROC when using selection scheme LCF is larger than when using MAF and MUF, as shown in Section 3.3.2. In this experiment, the running time of LCF is about 1 to 5 minutes on machine B[3] while the running time of MUF and MAF are both within several seconds on the same machine. This is because, after routing each connection, LCF needs to re-compute the cost of the remaining connection requests under

[3]Machine B is a Linux PC with a 1.7-GHz Pentium IV processor and 2-GB memory.

the updated network state, and computing the cost of each connection needs to determine the route of the connection on the auxiliary graph, whose time complexity is $O(N^2W^2)$, while MUF and MAF do not need this step. Note that the ILP cannot practically solve the problem of this size. It reported not enough memory available when we ran the ILP for this example on machine A or B. Even if each link has only 4 wavelengths and there are only 60 connection requests, the ILP will have 40,800 constraints and 1,420,597 variables and it could not obtain the solution within 3 days.

To further demonstrate the scalability of the IGABAG algorithm and the three traffic-selection heuristics, we conducted, on machine B, an experiment on a large, national-wide network with 277 nodes. We assume that each node has 20 tunable transceivers and no wavelength converter, and there are 20 wavelengths on each link. The running time of the IGABAG algorithm is less than 0.3 second. To route 3496 connections, MUF and MAF take about 17 minutes, while LCF is estimated to run about 20 days. It can be observed from these results that the IGABAG algorithm scales well with network size and can even satisfy the complexity requirement for on-line provisioning in dynamic traffic grooming. For static traffic, MUF and MAF can handle a large amount of connection requests in a reasonable time period, while LCF has a scalability problem as the number of connection requests increases.

Finally, Fig. 3.7(d) shows that, when the LCF heuristic is used, the network throughput increases as the number of wavelengths on each link increases. When each link has only 8 wavelengths, the throughput improves little as the number of transceivers is increased from 18 to 24. This is because the resource bottleneck is the number of wavelengths on each fiber-link rather than the number of transceivers at each node. For each node configuration, the network throughput increases as the number of wavelengths on each fiber increases.

3.6 Conclusion

In this chapter, we developed a novel and generic graph model for traffic grooming in a heterogeneous WDM optical mesh network. This model takes into account various constraints, such as transceivers, wavelengths, wavelength-conversion capabilities, and grooming capabilities, which means that it can be applied very generally to a heterogeneous WDM mesh network environment. With static grooming (where all connections to be set up are known a priori), it can achieve various objectives by using different grooming policies under blocking and non-blocking scenarios. Moreover, our ability to easily adjust the grooming policy according to the current network state makes this model very suitable for dynamic traffic grooming (where connections need to be set up one at a time). At the same time, the versatility of the model is accomplished by using a uniform mechanism for manipulating the edges in the auxiliary graph,

namely, by adjusting their weights and performing shortest-path computation, which is a very simple and novel feature of our proposed model.

Based on the auxiliary graph, we develop an integrated grooming algorithm, called IGABAG, which jointly solves the four traffic-grooming sub-problems for one traffic demand, and a grooming procedure, called INGPROC, which can accommodate both static and dynamic traffic grooming using the IGABAG algorithm. Among the three proposed grooming policies MinTH, MinLP, and MinWL, our study shows that MinWL consumes the minimum wavelength-links and MinLP uses the minimum transceivers under non-blocking scenarios, while the traffic travels using the minimum number of hops on the virtual topology in non-blocking scenarios when MinTH is used, and MinTH achieves the maximum throughput under blocking scenarios. For static traffic grooming, the LCF heuristic outperforms MUF and MAF when combined with the INGPROC procedure, while MUF and MAF scale better than LCF as the number of connection requests increases.

Chapter 4

DYNAMIC TRAFFIC GROOMING

4.1 Introduction

In recent years, the pervasive usage of Internet services has significantly increased the volume of Internet traffic. Due to the popularity of emerging Internet-based real-time applications and services, the network will become much more congested unless appropriate techniques are designed. Traffic engineering is an effective solution to control the network congestion and optimize network performance. As stated in [Awduche et al., 1999, Awduche et al., 2002], traffic engineering encompasses the application of technology and scientific principles to the measurement, modeling, characterization, and control of Internet traffic. The purpose of traffic engineering is to facilitate efficient and reliable network operations while simultaneously optimizing network resource utilization and traffic performance.

4.1.1 Traffic Engineering In Optical WDM Networks Through Traffic Grooming

It has been well recognized that optical wavelength-division multiplexing (WDM) networks will become the dominant transport infrastructure for Internet traffic. Using WDM technology, a fiber link can support multiple non-overlapping wavelength channels and provide over a Tbps link capacity. In such WDM networks, a lightpath provides a basic communication mechanism between two nodes. A lightpath is a wavelength circuit, which may span multiple fiber links and be routed by the intermediate optical switches between a given node pair. Although the bandwidth of a lightpath (i.e., a wavelength channel) in an optical WDM backbone network is quite high (10 Gbps (OC-192) today, and expected to grow to 40 Gbps (OC-768) soon), only a fraction of the customers are expected to have a need for such a high bandwidth. Many will be content

with a lower bandwidth, e.g., STS-1 (51.84 Mbps), OC-3, OC-12, OC-48, etc., for backbone applications. Since high-bandwidth wavelength channels will be filled up by many low-speed traffic streams, efficiently provisioning customer connections with such diverse bandwidth needs is a very important problem, which is the traffic-grooming problem. Due to the large cost of the optical backbone infrastructure and enormous WDM link capacity, provisioning inefficiency can cause significant capital overhead and a large amount of bandwidth waste. As discussed in previous chapters, traffic grooming enables a network operator to increase capacity utilization, minimize network cost, and optimize potential network revenue. Consequently, this chapter is devoted to the dynamic traffic-grooming problem in a heterogeneous, multi-granularity, optical WDM network environment, where connections with different bandwidth-granularity requirements arrive one at a time and each such connection needs to be properly routed through the network based on the current network state.

4.1.2 Optical WDM Network Heterogeneity

As the WDM technologies (transmission as well as switching technologies) keep maturing and optical network control and management protocols are being standardized, network operators are expected to be able to construct and operate a scalable, resource-efficient, mesh-based optical network, which can easily provide differentiated classes of service for different customers requirements. Thus, it is important for network operators as well as equipment vendors to understand how to efficiently provision connections of different bandwidth granularities in such WDM backbone mesh networks. A carrier's WDM backbone network is expected to be a multi-vendor, heterogeneous network since network equipment (NE) may need to be used from different vendors and new equipment has to co-exist with legacy equipment. For example: (1) network nodes may have optical crossconnects (OXCs) employing different architectures and technologies; (2) not all nodes may have wavelength conversion and traffic-grooming capabilities (i.e., sparse wavelength conversion and sparse grooming networks); (3) wavelength conversion and traffic grooming may only be available on certain wavelength channels (i.e., nodes may have partial wavelength conversion and partial grooming capability); and (4) different fiber links may support different numbers of wavelength channels, which may also operate at different speeds. This range of heterogeneity increases the challenge for the network operator to efficiently engineer the customers' traffic of different bandwidth granularities.

4.1.3 Organization

In this chapter, we investigate the problem of dynamically provisioning connections of different bandwidth granularities in mesh-based heterogeneous

WDM networks. The network is considered to have two levels of hierarchical switching capability: (1) wavelength switching and (2) SONET/SDH based low-speed circuit (time-slot) switching. The ideas of this study can also be applied to networks with other hierarchical switching capabilities, such as packet-based switching, waveband switching, fiber switching, etc.

This chapter is organized as follows: Section 4.2 presents different OXC architectures and their characteristics. Section 4.3 discusses different control issues on traffic grooming in a heterogeneous WDM network where different types of network components are inter-connected. Section 4.4 outlines a generic graph model (by building up on the work in Chapter 3) which can provide a framework to accommodate network heterogeneity and achieve different traffic-engineering objectives through different grooming policies. Section 4.5 presents illustrative numerical results. Section 4.6 concludes this study.

4.2 Node Architecture in a Heterogeneous WDM Backbone Network

Optical crossconnects (OXCs) are known to be the most important network elements in a carrier's WDM backbone network. Controlled by intelligent software in a centralized or distributed manner, OXCs can perform point-and-click or dial-up types of service to establish or tear down customers' connection requests in a matter of seconds (or perhaps milliseconds). There are transparent and opaque approaches to build these OXCs. The transparent approach refers to all-optical (O-O-O) switching, and the opaque approach refers to switching with optical-electronic-optical (O-E-O) conversion. Depending on their architectures and technologies employed, different OXCs may have different multiplexing and switching capabilities, which result in different capabilities for grooming low-speed traffic streams onto high-capacity wavelength channels. In this chapter, we consider OXCs of four categories. Below, we give a brief introduction on each type of OXC. A more detailed comprehensive study on their characteristics and performance will be presented in Chapter 5.

- *Non-grooming OXC*: This type of OXC can be built with either transparent or opaque approach. It has wavelength-switching capability. If the transparent approach is used, it is possible for this type of OXC to switch at higher bandwidth granularity, such as waveband (a group of wavelengths) or fiber. There is no low-data-rate port on a non-grooming OXC. Thus, extra aggregation/de-aggregation network elements are needed to work with this type of OXC if low-speed traffic streams need to be supported.

- *Single-hop grooming OXC*: Just like the non-grooming OXC, this type of OXC will only switch traffic at wavelength (or higher) granularity. On the other hand, it may have some lower-data-rate ports, which can directly support low-speed traffic streams. The traffic from these low-speed ports

can be multiplexed onto a wavelength channel using a TDM scheme, before the traffic enters the switch fabric. Since this type of OXC does not have the capability to switch low-speed streams at intermediate nodes, all of the low-speed streams on one wavelength channel at the source node will be switched to the same destination node.

- *Multi-hop partial-grooming OXC*: As shown in Fig. 4.1, the switch fabric of this type of OXC is composed of two parts: a wavelength-switch fabric (W-Fabric), which can be either all-optical or electronic, and an electronic-switch fabric which can switch low-speed traffic streams. The electronic-switch fabric is also called grooming fabric (G-Fabric). With this hierarchical switching and multiplexing architecture, this type of OXC can switch low-speed traffic streams from one wavelength channel to other wavelength channels and groom them with other low-speed streams without using any extra network element. Assuming that the wavelength capacity is OC-N and the lowest input port speed of the electronic switch fabric is OC-M ($N \geq M$), the ratio between N and M is called the *grooming ratio*. In this architecture, only a few of wavelength channels can be switched to the G-Fabric for switching at finer granularity. The number of ports, which connect the W-Fabric and G-Fabric, determines how much multi-hop grooming capability this OXC has.

- *Multi-hop full-grooming OXC*: This type of OXC can provide full-grooming functionality. Every OC-N wavelength channel arriving at the OXC will be de-multiplexed into its constituent OC-M streams before it enters the switch fabric. The switch fabric can switch these OC-M traffic streams in a non-blocking manner. Then, the switched streams will be multiplexed back onto different wavelength channels. An OXC with full-grooming functionality has to be built using the opaque approach. Note that the switching fabric of this type of OXC can be viewed as a large grooming fabric.

In a carrier's WDM backbone network, the interconnection of one or more types of OXCs may create different grooming scenarios, some of which have already been addressed in the literature [Zhu and Mukherjee, 2002b, Thiagarajan and Somani, 2001b, Ranganathan et al., 2002, Zhu et al., 2003a, Zhu and Mukherjee, 2002a, Xin et al., 2002]. The work in [Zhu and Mukherjee, 2002b] (Chapter 2) studied the network-design problem on a static traffic pattern using multi-hop partial-grooming OXCs at every network node. The work in [Ranganathan et al., 2002] quantitatively compared the network-cost tradeoffs for different network-design schemes using different types of OXCs for static traffic requests. The work in [Zhu and Mukherjee, 2002a][Xin et al., 2002][Thiagarajan and Somani, 2001b] explored the connection-provisioning problem for dynamic traffic requests with different bandwidth granularities in

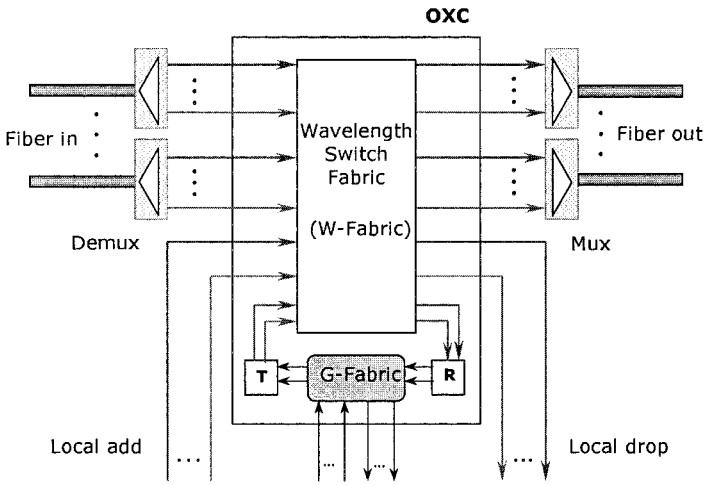


Figure 4.1. A multi-hop partial-grooming OXC.

WDM networks, which use multi-hop partial-grooming OXCs at every network node. The authors in [Zhu et al., 2003a] (Chapter 3) proposed a generic graph model to groom static traffic demands. The model assumes that, when a multi-hop partial-grooming OXC is used at a network node, all traffic requests will originate (or terminate) directly from (to) the G-Fabric of the OXC. This assumption may not hold in a heterogeneous WDM network since traffic streams can directly originate from (or terminate to) an OXC port (wavelength channel) without going through the G-Fabric. Moreover, the study in Chapter 3 did not explore how the network heterogeneity may affect the network performance and the network control plane. Hence, more studies are needed to efficiently engineer dynamic traffic in such a heterogeneous WDM backbone network. This is the topic of this chapter.

4.3 Provisioning Connections in Heterogeneous WDM Network

There are three important components in WDM network control, which determine how connections of different bandwidth granularities are provisioned. They are (1) resource-discovery protocol, (2) signaling protocol, and (3) route-computation algorithm. Resource-discovery protocols determine how the network resources are discovered, represented, and maintained in the OXCs' link-state databases (for distributed control) or by the network control and management system (for centralized control). Route-computation algorithms determine how the route of a low-speed connection request is computed and selected

according to the carrier's grooming policy [Zhu and Mukherjee, 2002a][Zhu et al., 2003a]. Grooming policy reflects the intention of a network operator on how to allocate network resources to a given request if multiple routes are available. It is a traffic-engineering decision. Signaling protocols determine how the connection is configured and how a network node allocates its local network resources to the connection, e.g., port mapping or label assignment for a connection. A unified control plane for intelligent WDM mesh-based networks is being standardized, and is known as Generalized Multi-Protocol Label Switching (GMPLS) [E. Mannie, et al., 2002] in the Internet Engineering Task Force (IETF) forum. The purpose of this network control plane is to provide an intelligent, automatic, end-to-end circuit (or virtual circuit) provisioning/signaling scheme throughout different network domains. As we mentioned above, the network may also be heterogeneous within one service domain. Thus, further extensions on the GMPLS control plane are needed to adapt it for the network heterogeneity caused by the inter-operation of OXCs from different vendors and the co-existence of new network equipment with legacy network equipment. These extensions will be elaborated in following sections. Note that, without loss of generality and for purposes of illustration, we will consider the WDM network to have full wavelength-conversion capability at every network node in our remaining sections.

4.3.1 Resource Discovery

As an abstract view, the network we are considering may have two types of network links: physical link (optical fiber) and virtual link (lightpath). A lightpath is also known as a traffic-engineering link (TE-link) in [E. Mannie, et al., 2002]. The switching granularity of the lightpath at any intermediate node should be full wavelength-channel granularity. A low-speed connection request may traverse one or multiple lightpaths. Because of the possibility of the interconnection of OXCs with different grooming capabilities, there may exist four types of lightpaths in a carrier's optical WDM network, as follows.

- *Multi-hop un-groomable lightpath*: A lightpath (i, j) is a multi-hop un-groomable lightpath if it is not connected with a finer-granularity switching element at its end nodes. This lightpath can only be used to carry the traffic directly between node pair (i, j). A lightpath, which is established between two non-grooming OXCs or single-hop grooming OXCs, is an example of this type of lightpath.

- *Source-groomable lightpath*: A lightpath (i, j) is a source-groomable lightpath if it is only connected with a finer-granularity switching element at its source node. All traffic on this lightpath has to terminate at node j, but the traffic may originate from any other network node as well. As an example, if a lightpath is established from a multi-hop full-grooming OXC to a non-

grooming OXC (or single-hop grooming OXC), it is a source-groomable lightpath.

- *Destination-groomable lightpath*: A lightpath (i, j) is a destination-groomable lightpath if it is only connected with a finer-granularity switching element at its destination node. All traffic on this lightpath has to originate from node i. At the lightpath destination node j, the traffic on lightpath (i, j) can either terminate at j or be groomed to other lightpaths and routed towards other nodes. As an example, if a lightpath is established from a non-grooming OXC (or single-hop grooming OXC) to a multi-hop full-grooming OXC, it is a source-groomable lightpath.
- *Full-groomable lightpath*: A lightpath (i, j) is a full-groomable lightpath if it connects to finer-granularity switching elements at both end nodes. This type of lightpath can be used to carry traffic between any node pair in the network. Note that source-groomable, destination-groomable, and full-groomable lightpaths can also be called multi-hop groomable lightpaths.

The states of both physical link (fiber link) and virtual link (lightpath) should be advertised by link-state advertisements (LSAs). Besides the neighbors connected by physical links, a node may also have other neighbors, which are connected to the node by virtual links and which are geographically far away in the network. As a simplified abstraction, we present the link state of each network link type as follows.

- *Fiber Link*: The representation of a fiber link (in a full wavelength-convertible network) can be denoted as $f(m, n, t, w, c)$, where m and n denote the end nodes of the fiber link, t denotes fiber index (for numbering multiple fibers between the same node pair), w denotes the available (free) wavelength channels on that fiber, and c denotes the administrative link cost. In a WDM network with wavelength-continuity constraint, more information is needed to identify the availability property of each individual wavelength channel. Note that this representation can be viewed as bundling multiple wavelength channels between the same pair of nodes as proposed in [E. Mannie, et al., 2002, K. Kompella, et al., 2002].

  If there are multiple fibers between the same node pair, they may be further bundled. The purpose of link bundling is to improve routing scalability by reducing the amount of information that has to be handled by the network control plane. This reduction is accomplished by performing information aggregation/abstraction [E. Mannie, et al., 2002]. After fiber links between the same node pair are bundled, they can be advertised as a logical link for route-computation algorithms.
- *Virtual Link*: The representation of a lightpath can be denoted as $l\ (i, j, v, t, m_1, m_2, c)$, where i and j denote the end nodes of the lightpath, v denotes the

lightpath type, t denotes the lightpath id, m_1 denotes the minimal reservable bandwidth on this lightpath, which is determined by the grooming ratio of the end nodes, m_2 denotes the maximal reservable bandwidth on this lightpath, which is bounded by the total available (free) capacity on the lightpath, and c denotes the administrative link cost. Multiple lightpaths (of the same type) between the same node pair can also be bundled and advertised as a logical link. Note that link bundling should also follow the network fault-management constraints, i.e., the links in the same bundle should belong to the same shared-risk link group (SRLG) [K. Kompella, et al., 2002].

Besides advertising the states of the network links, the state of a network node may also need to be advertised. This is because a multi-hop partial-grooming OXC can only perform multi-hop grooming on a limited number of lightpaths. So, if a network node employs this type of OXC, the available multi-hop grooming capability of the OXC should also be advertised when the node's state changes.

4.3.2 Route Computation

In an intelligent WDM network which we are considering, the route of a connection request will be computed either by the source node or by the network control and management system. Let $Req(s, d, r)$ denote a connection request, where s denotes the source node, d denotes the destination node, and r denotes the capacity requirement of the connection. There exist the following possibilities to route the connection request (as was discussed earlier in Section 3.4.1):

- *Operation 1*: Carry Req using an existing lightpath $l(s, d, v, t, m_1, m_2, c)$ between nodes s and d, if $m_1 \leq r \leq m_2$.
- *Operation 2*: Carry Req using multiple existing groomable lightpath.
- *Operation 3*: Carry Req by establishing a new lightpath (either groomable or non-groomable) between node pair (s, d) if enough resources exist.
- *Operation 4*: Carry Req using a combination of both existing groomable lightpaths and setting up new groomable lightpaths using available wavelength channels in fiber links and grooming resources in network nodes.

Since there are multiple ways to carry the connection request, multiple routes may be simultaneously available. The decision on how to choose a proper route from multiple candidate routes is a traffic-engineering decision and is known as the network operators' grooming policy. We have already studied different grooming policies in Chapter 3, under static traffic condition. In dynamic traffic environment, connections with various bandwidth requirements arrive in a network, stay for a certain period of time, and then leave the network. A grooming

policy may have different performance under dynamic traffic environment than static one. Moreover, dynamic traffic grooming might need to dynamically adjust the grooming policy according to the traffic pattern and current network state. Hence, grooming policies for dynamic traffic need to be investigated.

Below, we present four different grooming policies for dynamic traffic.

- *Minimize the Number of Traffic Hops on the Virtual Topology (MinTHV).* We first use Operation 1. If Operation 1 fails, we always try to set up a lightpath from s to d and route the traffic onto this lightpath (Operation 3). Only when such a direct lightpath cannot be set up, we use multi-hop grooming by either Operation 2 or Operation 4, and choose the one with fewer hops on the virtual topology (number of lightpaths). This policy chooses the route with the fewest lightpaths for a connection.

- *Minimize the Number of Traffic Hops on the Physical Topology (MinTHP).* We compare the number of wavelength-links used by all the four operations and choose the one with the fewest wavelength-links.

- *Minimize the Number of Lightpaths (MinLP).* This policy is similar to MinTHV but it tries to set up the minimal number of *new* lightpaths to carry the traffic. Operation 1 is attempted first. If it fails, we try to route the traffic using multiple existing lightpaths (Operation 2). If Operation 2 also fails, we try to set up one lightpath with the minimal number of wavelength-links either by Operation 3 or Operation 4. If such a lightpath is not feasible, we go with Operation 4 and set up two or more lightpaths.

- *Minimize the Number of Wavelength-Links (MinWL).* This policy is similar to MinTHP but it tries to consume the minimum number of *extra* wavelength-links, i.e., wavelength-links not being used by any lightpaths for now, to carry the traffic. The difference between MinLP and MinWL is that, if both Operations 1 and 2 fail, MinWL compares the number of wavelength-links used by Operations 3 and 4, and chooses the one requiring fewer wavelength-links; MinLP, on the other hand, compares the number of lightpaths used by Operations 3 and 4, chooses the one requiring fewer lightpaths, and uses the number of wavelength-links for tie-breaking.

In dynamic grooming, the network state changes as connection requests come and go. To achieve good performance, the grooming policy should be adjusted according to the current network state. For instance, if transceivers are becoming the more scarce resource, we should make full use of existing lightpaths to accommodate the new traffic and avoid setting up new lightpaths. This requirement can be easily satisfied by modifying the weights of edges according to the current network state. This capability of easily adjusting grooming policies makes the graph model very suitable for dynamic traffic grooming.

4.3.3 Signaling

After a route is successfully computed, every intermediate node along the route needs to be informed through appropriate signaling protocols. Two protocols — resource reservation protocol (RSVP) [Braden et al., 1997, Awduche et al., 2001] and constraint-based routing label distribution protocol (CR-LDP) [Andersson et al., 2001, B. Jamoussi, et al., 2002] with traffic-engineering extensions — have been proposed as the standard signaling protocols in the GMPLS control plane. In a heterogeneous WDM network, the route computed for a connection request will be composed of a sequence of intermediate node id as well as link bundle id. Since multiple candidate links may exist in a link bundle, an intermediate node needs to select one for the connection request when it needs to configure the OXC and establish the connection. If the link is a bundled lightpath link, then, based on the available capacity of each lightpath in the bundle and the bandwidth requirement of the request, different link-selection schemes can be used, e.g., random selection, first-fit selection, best-fit selection, etc.

4.4 A Generic Provisioning Model

4.4.1 Graph Model

The network heterogeneity increases the complexity of efficiently provisioning customers' connection requests. Hence, a generic bandwidth-provisioning model, which can incorporate various network elements (NEs) and accommodate different grooming policies, will enable network operators to manage their transport networks easily and efficiently, as well as reduce the overall cost (network cost and operation cost) significantly. This model should also be easy to implement and scale using a distributed control plane. Extending the work in Chapter 3, we propose such a simple, generic provisioning model for heterogeneous optical WDM networks.

Figure 4.2 shows an illustrative example of the provisioning model, which we use to explain how the model works. Figures 4.2(a) and 4.2(b) show the network state for a simple three-node network. The shaded node (node 0) is the node which employs a multi-hop partial-grooming OXC and the un-shaded nodes (nodes 1 and 2) are equipped with single-hop grooming OXCs. Each link in Fig. 4.2(a) represents a free wavelength channel between a node pair and each link in Fig. 4.2(b) represents an established lightpath. The lightpath $(0, 2)$ is a source-groomable lightpath, the lightpath $(1, 0)$ is a destination-groomable lightpath, and the lightpath $(2, 1)$ is a multi-hop un-groomable lightpath. A low-speed connection request from node 1 to node 2 can be carried by lightpaths $(1, 0)$ and $(0, 2)$. On the other hand, a request from node 2 to node 0 cannot traverse lightpaths $(2, 1)$ and $(1, 0)$ since node 1 does not have multi-hop grooming capability.

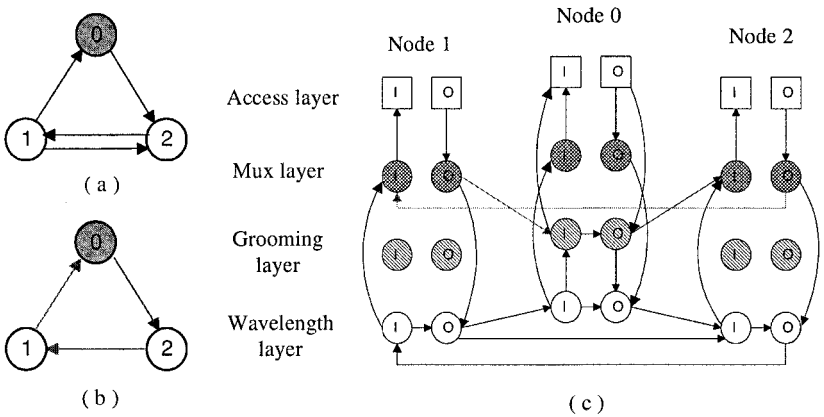


Figure 4.2. Network state for a simple three-node network and the corresponding auxiliary graph.

Figure 4.2(c) shows a graph representation of the network state in Fig. 4.2(a) and Fig. 4.2(b). The graph is divided into four layers, namely access layer, mux layer, grooming layer, and wavelength layer. The access layer represents the access point of a connection request, i.e., the point where a customer's connection starts and terminates. It can be an IP router, an ATM switch, or any other client equipment. The mux layer represents the OXC ports from which low-speed traffic streams are directly multiplexed (de-multiplexed) onto (from) wavelength channels without going through the grooming fabric. The grooming layer represents the grooming component of the network node. The wavelength layer represents the wavelength-switching capability and the link state of wavelength channels. A network node is divided into two vertices at each layer. These two vertices represent the input and output ports of the network node at that layer. The links in this graph model are named and work as follows.

- *Grooming switching link* connects the input port of the grooming layer to the output port of the grooming layer at a given node i, when node i has multi-hop traffic-grooming capability.

- *Wavelength switching link* connects the input port of the wavelength layer to the output port of the same layer at a given node i. It represents the wavelength-switching capability of the network node.

- *Mux link* connects the output port of the access layer to the output port of the mux layer at a given node i. It represents that the traffic starting from node

i can be packed to some wavelength channels and transmitted to another network node together without going through any grooming fabric.

- *Demux link* connects the input port of the mux layer to the input port of the access layer at a given node i. It represents that the traffic on a wavelength channel has been de-multiplexed and terminated at this node without going through any grooming fabric.

- *Mux to wavelength transmitting link* connects the output port of the mux layer to the output port of the wavelength layer at a given node i.

- *Wavelength to mux receiving link* is the link which connects the input port of the wavelength layer to the input port of the mux layer at a given node i.

- *Grooming link* connects the output port of the access layer to the output port of the grooming layer at a given node i, when node i has multi-hop grooming capability (i.e., free outgoing grooming ports). It represents that the traffic starting from node i can be groomed with other traffic streams (originating from node i or from other network nodes) to the same wavelength channel and transmitted to the next network node together.

- *De-grooming link* connects the input port of the grooming layer to the input port of the access layer at a given node i, when node i has multi-hop grooming capability (i.e., free incoming grooming ports). It represents that the traffic streams on a wavelength channel have been de-multiplexed, and then they may be either terminated at node i or switched to other lightpaths.

- *Grooming to wavelength transmitting link* connects the output port of the grooming layer to the output port of the wavelength layer at a given node i, when node i has multi-hop grooming capability (i.e., free incoming grooming ports). It denotes that a multi-hop groomable lightpath (i.e., either a source-groomable lightpath or a multi-hop full-groomable lightpath) can be originated at node i.

- *Wavelength to grooming receiving link* connects the input port of the wavelength layer to the input port of the grooming layer at a given node i, when node i has multi-hop grooming capability (i.e., free incoming grooming ports). It denotes that a multi-hop groomable lightpath (i.e., either a destination-groomable lightpath or a multi-hop full-groomable lightpath) can be terminated at node i.

- *Wavelength link* connects the output port of the wavelength layer at node i to the input port of the wavelength layer at node j. It denotes the availability of the wavelength channels between node pair (i, j).

- *Lightpath link* can start at the output port of the mux layer (grooming layer) at node i, and terminate at the input port of the mux layer (grooming layer) at node j. The four combinations of the end points represent the four possible lightpath types between node pair (i, j) discussed in Section 4.3.1.

Each link in Fig. 4.2(c) represents the availability of the corresponding network resource. A link is removed if the corresponding network resource is not available (e.g., deleting a wavelength link), and a link is added if the corresponding network resource becomes available from unavailable state (e.g., adding a lightpath link). In this graph representation, a customer's connection request will always originate from the output port of the access layer at the source node and terminate at the input port of the access layer at the destination node. After properly adjusting the administrative link cost, suitable routes can be found according to different grooming policies for a request by simply applying standard shortest-path route-computation algorithms. Through some straightforward extensions (by stretching the single wavelength layer to multiple layers, one for each wavelength), the network without full wavelength-conversion capability can also be properly modeled.

Figure 4.3 shows the corresponding graph representation of network nodes employing different traffic-groomable OXC architectures. They are extended from the original graph model proposed in Chapter 3. In the original graph model, there is a lightpath layer in the auxiliary graph, and all lightpath will originate/terminate at that lightpath layer. Hence, the auxiliary graph in Chapter 3 cannot model a network node to differentiate the four lightpath types and to simultaneously support all lightpath types we presented in Section 4.3.1 even if the node has this capability. By splitting the lightpath layer into the mux layer and the grooming layer, our proposed extended model is able to simultaneously support all types of lightpaths and incorporate all possible network states in a heterogeneous network environment.

4.4.2 Engineering Network Traffic Using the Proposed Graph Model

We can observe that, by constructing this auxiliary graph and using it for connection provisioning, the network control software can easily incorporate different network elements (i.e., OXCs with different characteristics) and accurately represent different network states. This strategy provides a platform for network operators to realize different grooming policies, and to eventually improve the provisioning flexibility and network resource efficiency.

Figure 4.4 shows an example on how to achieve different traffic-engineering objectives through different grooming policies by using our generic graph model. The network state and the graph representation are shown in Fig. 4.2. Assuming that there is a new traffic request from node 1 to node 2, Fig. 4.4

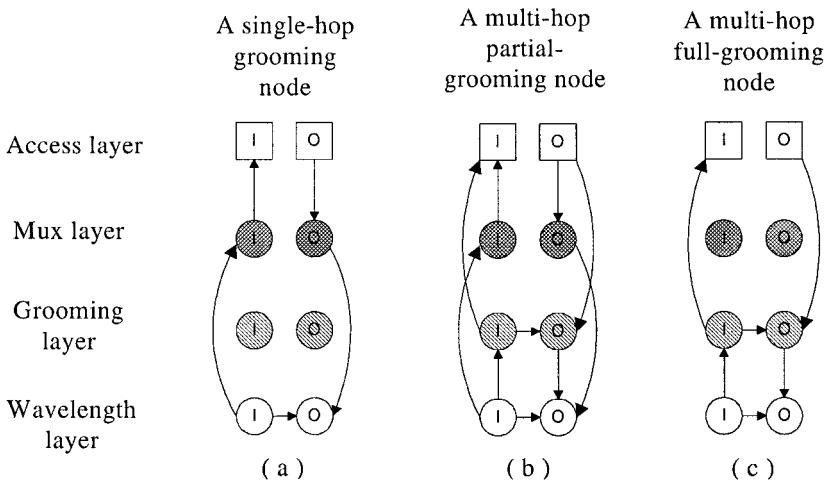


Figure 4.3. Different grooming OXCs and their representations in the auxiliary graph.

shows two possible routes (in thick links) for this connection request. The route shown in Fig. 4.4(a) traverses two existing lightpath links, while the route shown in Fig. 4.4(b) will employ two new wavelength channels. If the connection requires full wavelength-channel capacity, or if the overall bandwidth requirement of the future traffic demands between the node pair is estimated to be close to full wavelength-channel capacity, the route in Fig. 4.4(b) is preferred since the wavelength channels are fully utilized and no grooming is needed at node 0; otherwise, the route in Fig. 4.4(a) may be preferred if enough free capacity is available on the existing lightpaths. By assigning proper weights to each link in the graph in Fig. 4.2(c), the corresponding routes will be computed through standard shortest-path computation algorithms according to different grooming policies.

4.4.3 Computational Complexity

The computational complexity to construct such an auxiliary graph for a N-node WDM network is $O(2 \times N)$ for a full wavelength-convertible WDM network. This is because the auxiliary graph will consist of $N \times 4 \times 2$ nodes and at most $(N \times 4 \times 2)^2$ links. As we know, the computation complexity for a standard shortest-path algorithm in a N-node network is $O(N^2)$. Therefore, the computational complexity to provision a connection request using this model in a full wavelength-convertible WDM network is $O(N^2)$. If the WDM network does not have full wavelength-conversion capability, there will be $(2 \times (W+3) \times N)$ nodes and at most $(2 \times (W+3) \times N)^2$ links in the auxiliary graph, where W

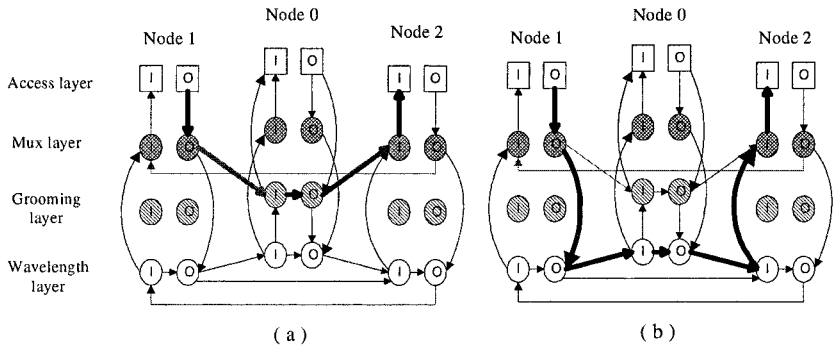


Figure 4.4. Two alternative routes for a new connection request $(1, 2)$.

is the number of wavelength channels a fiber supports in the network. Hence, the computational complexity to provision a connection request using this model will be $O(N^2 \times W^2)$ in such a wavelength-continuous WDM network.

4.5 Illustrative Numerical Examples

4.5.1 Comparison of Grooming Policies

We compare the performance of different grooming policies on the network topology shown in Fig. 3.7(a), which has 19 nodes and 31 links. All the nodes have grooming capability but no wavelength-conversion capability. Each link is bidirectional with $W = 16$ wavelengths in each direction, and each wavelength has a capacity of OC-192. The traffic arrival is a Poisson process and the connection holding time is exponentially distributed (whose average value is normalized to unity in our studies reported here). The traffic is uniformly distributed among all node pairs. There are four types of connection requests: OC-3, OC-12, OC-48, and OC-192, and the proportion of the number of these connections is 6:6:6:1. For a connection request $T(s, d, g, m)$, m, the amount of the traffic in unit of g, is uniformly distributed between 1 and 32, 1 and 16, 1 and 8, and 1 and 2 for OC-3, OC-12, OC-48, and OC-192 types of connections, respectively. We simulate 100,000 connection requests to obtain the network performance under a certain scenario and a grooming policy. We ran our simulation experiments on a Linux PC with a 1.5-GHz Pentium IV processor and 512-MB memory. Each data point reported in the illustrations in this section took between 6-9 minutes of running time on this computer.

Table 4.1 shows the average utilization of wavelength-links (U_W) and the average utilization of transceivers (U_{Tx}) when the network load L is 300 Erlangs. When each node has only 16 transceivers, the utilization of transceivers is very high since they are the more constrained resources. When there are 32

Table 4.1. Average utilization of wavelength-links and transceivers when W=16 and L=300 Erlangs.

	Tx=16		*Tx=32*		*Tx=40*	
	U_W	U_{Tx}	U_W	U_{Tx}	U_W	U_{Tx}
MinTHV	0.7819	0.9858	0.8878	0.7264	0.8905	0.5884
MinTHP	0.5674	0.9901	0.7354	0.8165	0.7361	0.6910
MinLP	0.7403	0.9807	0.8890	0.8007	0.8918	0.6651
MinWL	0.6201	0.9859	0.8133	0.8683	0.8120	0.7825

transceivers at each node, the utilizations of both transceivers and wavelengths are quite balanced and high as well. If there are 40 transceivers at each node, we have relatively more transceivers; hence, wavelength-links become the more constrained resources, so the utilization of the transceivers is relatively lower.

Figure 4.5 shows the network performance under different grooming policies with $Tx = 32$ transceivers at each node. We observe that, as the network load increases, the percentage of blocked traffic also increases, but different grooming policies have different blocking probabilities. Grooming policy MinTHV performs best, followed by MinTHP, MinLP, and MinWL in sequence.

When we alter the network configuration by adding more transceivers at each node, the performance of each of the policies also changes, as shown in Fig. 4.6. In this scenario, each node has 40 transceivers instead of 32. We observe that, now, MinTHP outperforms MinTHV and achieves the best results, and MinLP becomes the poorest-performing policy. This is because, in this network configuration, there are relatively more transceivers in the network so that wavelength-links become the more constrained resources. Recall that MinTHP utilizes wavelength-links more efficiently than other policies; hence, it performs the best in this case.

Another observation from both Figs. 4.5 and 4.6 is that MinTHV and MinTHP always perform better than MinLP and MinWL in terms of percentage of blocked traffic. This is because MinTHV and MinTHP examine the *overall* resource requirement of a given connection, while MinLP and MinWL only consider the new lightpaths to be set up or the extra wavelength-links used by these new lightpaths while setting up the connection. Therefore, MinTHV and MinTHP are more resource-efficient.

From the above results, we can observe that different grooming policies have different performance under various network configurations, which suggests that a grooming policy should be adjusted according to the current network state.

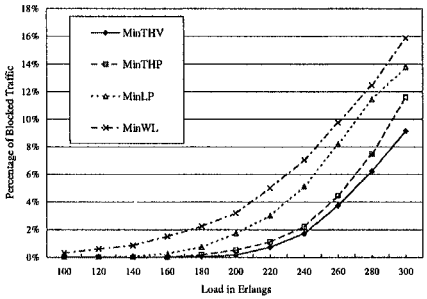


Figure 4.5. Percentage of blocked traffic when $Tx = 32$.

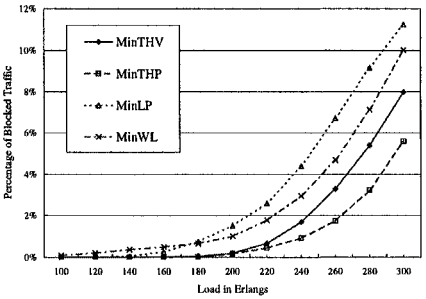


Figure 4.6. Percentage of blocked traffic when $Tx = 40$.

Adaptive Grooming Policy (AGP)

Since MinTHV performs best when transceivers are the more constrained resources and MinTHP gives the best results when wavelength-links become more scarce resources, we try to utilize the advantages of these two grooming policies by combining them together. Here, we present an Adaptive Grooming Policy (AGP) which, for each connection request, switches between MinTHV and MinTHP according to the current network state.

We use the ratio of the number of unused wavelength-links in the network to the total number of available transceivers at all nodes as an indicator of the network state. If the ratio is larger than Δ_1, MinTHV will be used to avoid setting up lightpaths since transceivers are more scarce resources at this time; if the ratio is less than Δ_2, MinTHP will be employed to try to save wavelength-links as much as possible; if the ratio is in between, the policy will not be changed.

For our numerical examples, we report results for $\Delta_1 = 1.2$ and $\Delta_2 = 1.0$. (We experimented with other combinations of these two parameters, and found these choices of values to perform the best for the network topology in Fig. 3.7(a).) Our results are shown in Figs. 4.7 and 4.8. We observe that the Adaptive Grooming Policy (AGP) achieves the best results under different network configurations.

4.5.2 Performance under Different Scenarios

In this subsection, we show illustrative numerical examples on the network performance by using our proposed model to provision different bandwidth-granularity connection requests. The network topology used in our simulation experiments is shown in Fig. 4.9. It represents a typical operator's optical backbone network topology, which has 24 nodes and 43 bi-directional links. We simulated a dynamic traffic environment, where the traffic arrival process is

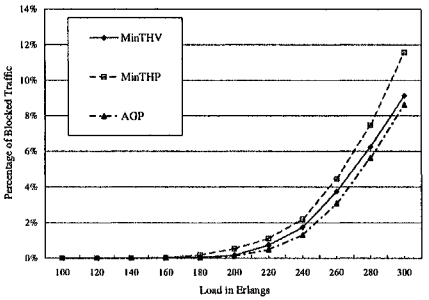


Figure 4.7. Performance of AGP when $Tx = 32$.

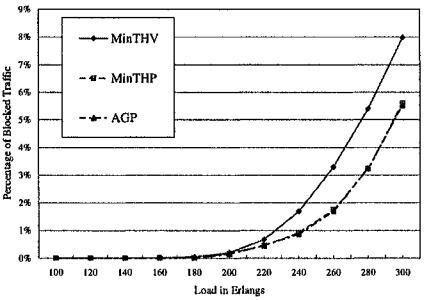


Figure 4.8. Performance of AGP when $Tx = 40$.

Poisson and the connection-holding time follows a negative exponential distribution. The capacity of each wavelength channel is OC-192. For the illustrative results shown here, the bandwidth requirements of the connection requests follow an uniform distribution between OC-3, OC-12, OC-48, OC-192 (i.e., OC-3 : OC-12 : OC-48 : OC-192 = 1 : 1 : 1 : 1); connection requests are uniformly distributed among all node pairs; average connection-holding time is normalized to unity; the cost of a fiber link is modeled as unity; and load (in Erlang) is defined as the connection-arrival rate times the average connection-holding time, which denotes the average number of connections the network is carrying at any instant. Note that, from the connections' bandwidth distribution and Erlang load, it is straightforward to calculate the network offered load in the unit of OC-1.

We employ three metrics to evaluate the network performance, namely Traffic Blocking Ratio (TBR), Connection Blocking Probability (CBP), and Resource Efficiency Ratio (RER).

- *Traffic Blocking Ratio* represents the percentage of the amount of blocked traffic over the amount of bandwidth requirement of all traffic requests during the entire simulation period.

- *Connection Blocking Probability* represents the percentage of the total number of blocked connection requests over the number of all traffic requests during the entire simulation period. The connections with different bandwidth requirements may experience different connection blocking probabilities.

- *Resource Efficiency Ratio* represents how efficiently connections are routed and groomed. It can be computed as the average of network carried traffic (in terms of OC-1 unit) divided by the total allocated network capacity (i.e., total number of allocated wavelength links times 192) over the entire simulation period.

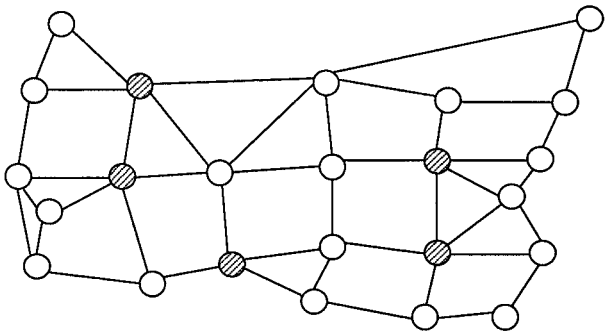


Figure 4.9. Sample network topology with 5 grooming nodes.

Let α denote the RER. α can be computed as follows:

$$\alpha = \frac{\sum_i \rho_i \times t_i}{\sum_i \gamma_i \times t_i},$$

where t_i is the time period between the i^{th} event (connection arrival or departure) and $(i+1)^{th}$ event, ρ_i is the network carried load during the time period t_i, and γ_i is the total number of wavelength links used during t_i. (Please note that ρ_i and γ_i do not change during time period t_i as there is no other event during that period.) If every connection required full wavelength-channel capacity (i.e., no capacity waste) and all connections were routed through the shortest path (in hop distance), the resource efficiency ratio will equal $\frac{1}{H}$, where H is the average hop distance of the network topology. This gives us an upper bound on the resource efficiency ratio. The average hop distance of the network topology in Fig. 4.9 is 3; hence, the upper bound for resource efficiency ratio for this network is approximately 33.3%.

A connection is provisioned based on the following policies.

- The connection is provisioned through the least-cost route. The cost of a lightpath link is equal to the overall cost of the concatenated fiber links it traverses. Every fiber link has unit cost.
- If there are multiple least-cost routes and the connection does not require full wavelength-channel capacity, select the route employing the minimal number of free wavelength links.
- If there are multiple least-cost routes and the bandwidth requirement of the connection requires full wavelength-channel capacity, select the route traversing the minimal numbers of electronic grooming fabrics.

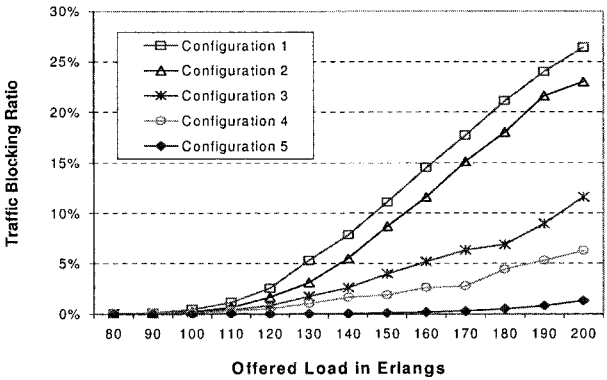


Figure 4.10. Traffic blocking ratio vs. offered load.

Note that it is also possible to apply other grooming policies by using the proposed provisioning model. Additional grooming policies can be found in Chapter 3.

Figures 4.10–4.12 show illustrative numerical examples based on different network configurations. In Configuration 1, all network nodes are only equipped with single-hop grooming OXCs. In Configurations 2, 3, and 4, the shaded nodes in Fig. 4.9 are equipped with multi-hop partial-grooming OXCs. The numbers of grooming ports in multi-hop partial-grooming OXCs are 4, 8, and 16 in Configurations 2, 3, and 4, respectively. In Configuration 5, all shaded nodes in Fig. 4.9 are equipped with multi-hop full-grooming OXCs[1]. Each bi-directional link in Fig. 4.9 contains two uni-directional fibers and each fiber supports eight wavelength channels.

In Fig. 4.10, we have plotted the traffic blocking ratio versus network offered load in Erlangs. We observe that, as the load increases, the traffic blocking ratio increases. When the network has more grooming capability, the traffic blocking ratio can be improved significantly. Figure 4.11 shows the normalized (to theoretical upper bound) resource efficiency ratio of different network configurations. It is straightforward to see that multi-hop grooming capability can help to increase the resource efficiency ratio, which will, in turn, reduce the traffic blocking ratio (shown in Fig. 4.10). If there is no multi-hop groomable node in the network, at least one lightpath is needed between every node pair whenever there is any low-speed request. Hence, it may lead to a full-mesh logical con-

[1]Note that the problem of how to select the multi-hop grooming nodes is an interesting research topic on network design and optimization. It is addressed in Chapter 6. The work in [Zhu et al., 2002, Lardies et al., 2001] also presented some related study. In this work, the multi-hop grooming nodes are selected from among the nodes with large nodal degree (4 or 5).

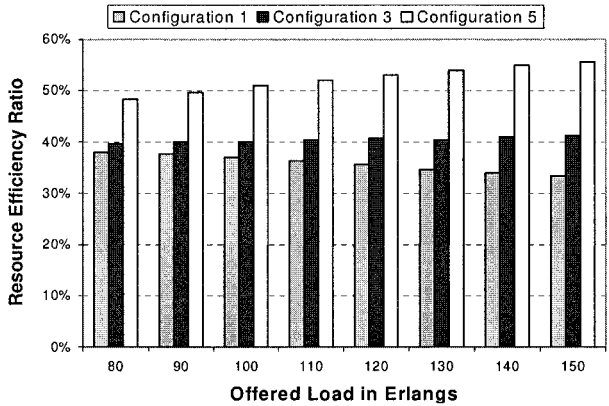


Figure 4.11. Normalized resource-efficiency ratio vs. offered load.

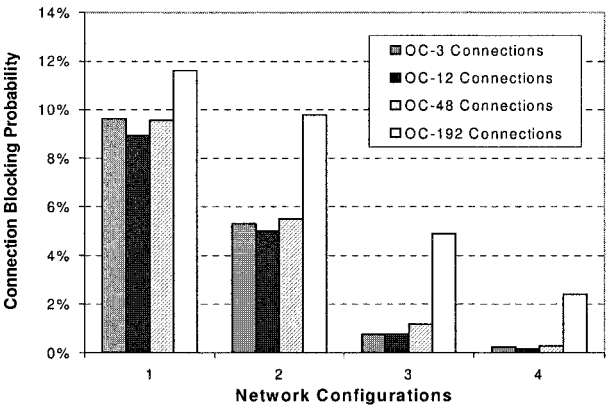


Figure 4.12. Connection blocking probability vs. offered load.

nectivity. This tends to increase the probability of connection blocking (due to lack of a free wavelength) and low wavelength-capacity utilization (due to under-utilization of the established lightpaths).

Figure 4.12 shows the blocking probabilities of connections with different bandwidth granularities. We can observe that connections requiring a full wavelength-channel capacity experience more chances to be blocked. This leads to a fairness concern. As Fig. 4.12 shows, the unfairness can become quite severe (relatively) when the network employs more grooming capability. This is because, when low-speed connections are carried by the network, the resources tend to be fragmented. Therefore, high-speed connections tend to

be blocked more often since they cannot fit into the available trunk capacities. The authors in [Thiagarajan and Somani, 2001b] have addressed this unfairness issue and proposed a call-admission-control (CAC) algorithm to achieve the capacity fairness. This algorithm tries to keep the traffic blocking ratio of each connection bandwidth-granularity class at the same level by using statistical information to perform admission control. Another possible solution to resolve this fairness issue is trunk reservation or capacity reservation, i.e., pre-reserve a certain amount of resources for connections with high-granularity bandwidth requirement.

Different network topologies, different resource parameters (i.e., different wavelength number, different multi-hop grooming node locations, etc.) and different connection bandwidth-granularity distributions were also experimented with, and we observed similar results.

4.6 Conclusion

We discussed the challenges to provisioning connections of different bandwidth granularities in heterogeneous WDM backbone networks. This heterogeneity may arise in a large, multi-vendor, inter-operational nation-wide backbone network. We investigated how to properly extend the existing standards to incorporate this heterogeneity and proposed a simple, graph-based, provisioning model. Our proposed model is scalable and easy to implement in the existing control plane. Using this model, different traffic-engineering optimization objectives can be easily achieved through different grooming policies. Our results illustrated the effectiveness of the proposed model and the relationship between the nodal grooming capability and the network throughput, connection blocking probability, and resource efficiency. We observed that employing more grooming capacity can help a network operator to improve the network performance and the utilization of the link capacity. However, this may also increase the network cost since more electronic processing is needed. We verified the blocking performance unfairness between connections of different bandwidth-granularity classes due to resource-usage characteristics. We also observed that the unfairness problem can become more severe when a network has more grooming capability. This may lead to an interesting research topic in the future.

Chapter 5

GROOMING SWITCH ARCHITECTURES

5.1 Introduction

As WDM switching technology keeps maturing, optical WDM networks are expected to evolve to arbitrary-mesh topologies, and network provisioning will migrate from an on-site manual interconnection process to a point-and-click or on-demand automatic switching and connection process. Such an intelligent optical WDM network is emerging under the joint efforts of optical switch (or OXC) development, optical network control plane standardization, and extensive optical WDM network research and experimental activities in industry and academe. Among different optical WDM switching technologies (i.e., circuit switching, burst switching, and packet switching), WDM circuit switching is believed to be the most practical approach to enable the next-generation optical network. Consequently, this study concentrates on an intelligent optical circuit-switched WDM transport network.

In this chapter, we systematically investigate and evaluate the characteristics of different optical grooming switches, i.e., the optical crossconnects (OXCs) with traffic-grooming capability. According to various OXC architectures, we explore and propose different traffic-grooming schemes, and compare the performance of the grooming OXCs and grooming schemes under a dynamic traffic environment. To the best of our knowledge, this is one of the first studies which comprehensively examines the characteristics and performance of different types of optical grooming OXCs for dynamic traffic under different connection bandwidth-granularity distributions. Our investigation should assist network operators to cost-effectively design and operate an optical groomable WDM backbone network, and it will be also helpful for equipment vendors to develop high-performance grooming OXCs.

The rest of the chapter is organized as follows. In Section 5.2, we introduce the architectures of different grooming OXCs and explore their corresponding grooming schemes. Section 5.3 proposes the detailed approaches and algorithms for the different grooming schemes using different grooming OXCs. Illustrative numerical results are presented and analyzed in Section 5.4. Section 5.5 concludes the study.

5.2 Different Grooming Switch Architectures and Corresponding Grooming Schemes

In an optical WDM network, a lightpath provides a basic communication mechanism between two network nodes. From traffic-grooming perspective, a lightpath is a circuit with full wavelength capacity. It may span one or multiple fiber links and be switched by intermediate nodes. Low-speed traffic streams can be packed onto a lightpath at its end nodes by grooming OXCs. There are transparent (all-optical) or opaque switching technologies to implement those OXCs. Transparent (all-optical) technology refers to the switching without optical-to-electronic (OE) conversion. Opaque technology refers to the switching with OE conversion. Different technologies and architectures may lead to different grooming OXCs, which may be capable of different grooming schemes. Specifically, there are three grooming schemes, namely *single-hop grooming, multi-hop grooming, and light-tree-based source-node grooming*. Each grooming OXC may support one or more grooming schemes. Different types of grooming OXCs with their corresponding grooming schemes are reviewed in the following subsections.

5.2.1 Single-Hop Grooming OXC

An OXC is called a single-hop grooming OXC if (1) it can only switch at wavelength granularity, and (2) it has low-data-rate interfaces (ports) which can directly support low-speed traffic streams from client network elements. Note that, co-operating with a separated network aggregation equipment (e.g., an electrical multiplexer), an OXC with only wavelength-granularity input/output ports and only switching at wavelength granularity is equivalent to and can be viewed as a single-hop grooming OXC. Using this type of OXCs, low-speed traffic from client equipment can be multiplexed onto a wavelength channel using a TDM scheme. Since this type of OXC does not have the capability of switching low-granularity traffic at intermediate nodes, all of the low-speed streams on one wavelength channel at the source node will be switched to the same destination node, i.e., a low-speed connection can only traverse a single lightpath hop. Therefore, this end-to-end grooming scheme is called single-hop grooming scheme.

Figure 5.1(a) shows how a low-speed connection ($C1$) is carried by a lightpath ($L4$) from node 1 to node 5 using the single-hop grooming scheme. Note that, in Fig. 5.1(a), nodes 1 and 5 are equipped with single-hop grooming OXCs, which can only switch at wavelength granularity.

5.2.2 Multi-Hop Partial-Grooming OXC

A multi-hop partial-grooming OXC consists of two switch fabrics, a wavelength-switch fabric (W-Fabric) which can be implemented using either transparent (all-optical) or opaque (electronic) technology, and an electronic-switch fabric which can switch low-speed traffic streams. The electronic-switch fabric is also called grooming fabric (G-Fabric). With this hierarchical switching and multiplexing architecture, this type of OXC can switch low-granularity traffic streams from one wavelength channel to other wavelength channels and groom them with other low-speed streams without using any extra network element. Assuming that the wavelength capacity is OC-N and the lowest input port speed of the electronic switch fabric is OC-M ($N \geq M$), the ratio between N and M is called the "grooming ratio". In this architecture, only a few of the wavelength channels (lightpaths) can be switched to the G-Fabric for switching at finer granularity. The number of ports, which connect the W-Fabric and G-Fabric, determines how much multi-hop grooming capability the OXC has. Figure 5.2(a) shows a simplified multi-hop partial-grooming OXC architecture. A multi-hop partial-grooming OXC can support both single-hop grooming and multi-hop grooming schemes.

Figure 5.1(a) also shows how a low-speed connection ($C2$) can be carried by a sequence of multiple lightpaths ($L1$, $L2$, and $L3$) from node 1 to node 5. Note that nodes 2 and 3 are equipped with multi-hop partial-grooming OXCs, and only their G-Fabrics are shown in the figure.

Figure 5.1(a) also shows that there may exist four types of lightpaths in a WDM network which employs multi-hop partial-grooming OXCs. Assuming that, in Fig. 5.1(a), all network nodes are equipped with multi-hop partial-grooming OXCs, and only the G-Fabrics of nodes 2 and 3 and W-Fabrics of nodes 1, 4, and 5 are shown, lightpaths $L1$, $L2$, $L3$, and $L4$ represent these four lightpath types.

- *Multi-hop-ungroomable lightpath* ($L4$): A lightpath (i, j) is a multi-hop-ungroomable lightpath if it is not connected with a finer-granularity switching element at either of its end nodes. This lightpath can only be used to carry the traffic directly between node pair (i, j). Lightpath $L4$ in Fig. 5.1(a) is a multi-hop-ungroomable lightpath. For simplicity, this type of lightpath may also be called an ungroomable lightpath.

- *Source-groomable lightpath* ($L3$): A lightpath (i, j) is a source-groomable lightpath if it is only connected with a finer-granularity switching element

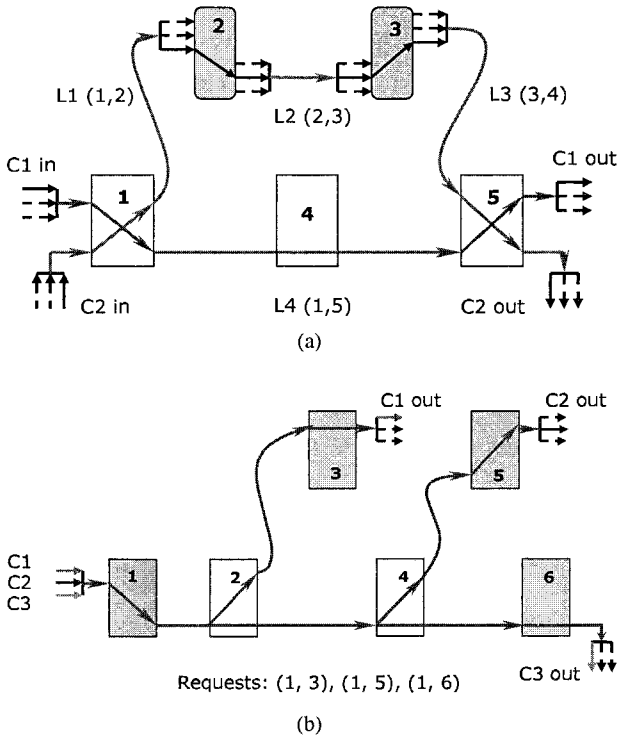
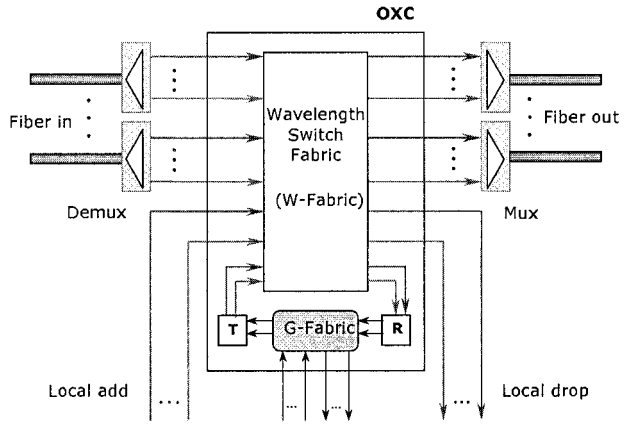


Figure 5.1. Examples of single-hop, multi-hop, and source-node grooming schemes.

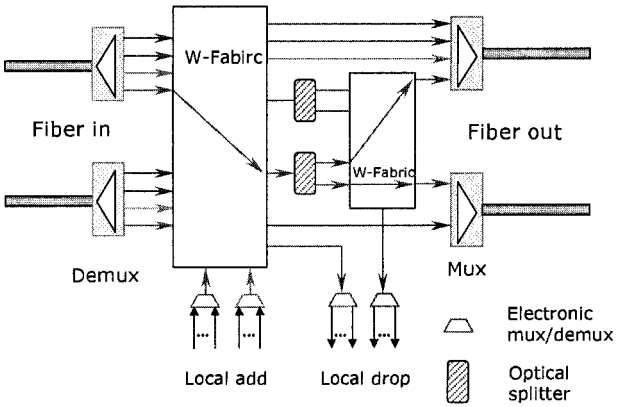
at its source node. All traffic on this lightpath has to terminate at node j, but the traffic may originate from any other network node as well. Lightpath $L3$ in Fig. 5.1(a) is a source-groomable lightpath.

- *Destination-groomable lightpath* ($L1$): A lightpath (i, j) is a destination-groomable lightpath if it is only connected with a finer-granularity switching element at its destination node. All traffic on this lightpath has to originate from node i. At the lightpath destination node j, the traffic on lightpath (i, j) can either terminate at j or be groomed to other lightpaths and switched towards other nodes. Lightpath $L1$ in Fig. 5.1(a) is a destination-groomable lightpath.

- *Full-groomable lightpath* ($L2$): A lightpath (i, j) is a full-groomable lightpath if it connects to finer-granularity switching elements at both end nodes. This lightpath can be used to carry traffic between any node pair in the network. Lightpath $L2$ in Fig. 5.1(a) is a full-groomable lightpath. Please note



(a)



(b)

Figure 5.2. Sample grooming OXC architectures: a multi-hop partial-grooming OXC and a source-node grooming OXC.

that the source-groomable lightpath, destination-groomable lightpath, and full-groomable lightpath can be also called multi-hop-groomable lightpath.

In an optical WDM network employing multi-hop partial-grooming OXCs, the lightpaths can either be established dynamically according to current connection requests, or be pre-planned based on forecasted traffic demands.

5.2.3 Multi-Hop Full-Grooming OXC

A multi-hop full-grooming OXC can provide full-grooming functionality, i.e., every OC-N wavelength channel arriving at the OXC will be demultiplexed into its constituent OC-M streams before it enters the switch fabric. The switch fabric can switch these OC-M traffic streams in a non-blocking manner. Then, the switched streams will be multiplexed back onto different wavelength channels. An OXC with full-grooming functionality has to be built using the opaque approach.

When a WDM network employs multi-hop full-grooming OXC at every network node, every wavelength channel on each fiber link forms a full-groomable lightpath between adjacent node pairs. Therefore, the virtual topology (i.e., lightpath topology) is pre-determined and exactly the same as the physical topology (fiber topology). A low-speed traffic stream can be easily switched from one time slot of a wavelength channel to another time slot of a possibly-different wavelength channel at every intermediate node which it traverses. A multi-hop full-grooming OXC can support both single-hop and multi-hop grooming schemes.

The time-space-time (TST) switching technology can be used to implement the multi-hop full-grooming OXC switch fabric. Figure 5.3 shows an overview of such a multi-stage switching architecture. Each input port of the fabric shown in Fig. 5.3 consists of n time slots and the fabric can switch N time slots in a non-blocking manner. Hence, the total number of input ports (as well as output ports) is N/n. Such a fabric consists of 3 stages. The first stage performs time-slot interchange (TSI) functionality, through which each of the n time slots from an input port can be directed towards any of the k time slots towards the second stage. The second stage performs space-switching functionality, through which N/n traffic streams (one time slot from each input port) can be switched towards the third stage. Note that the second stage may only require a single space-switch component since, at any instance, there are at most N/n traffic streams to be switched towards the third stage (one time slot from each input port). This architecture is called a time-shared-space switch. Again, the third stage performs time-slot interchange functionality. Based on the "Clos" switch architecture, in order to support non-blocking switching at lower granularity level (say, STS-1), the number of time slots supported by the internal connections (between stage 1 and 2, and between stage 2 and 3) should be nearly two times the number of time slots carried by the input/output ports (i.e., $k = 2n-1$). In other words, the internal speed of such a TST-based switch should nearly double the date rate of its input/output ports, which may lead to a significant technical difficulty to develop a large-scale multi-hop full-grooming OXC operating at high speed (e.g., OC-192, approx. 10 Gbps and OC-768, approx. 40 Gbps).

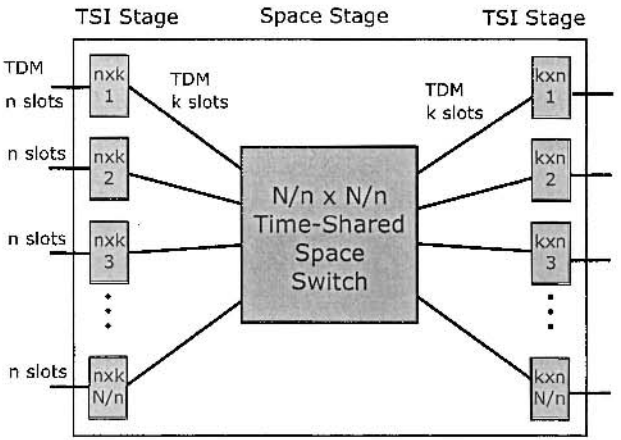


Figure 5.3. An overview of Time-Space-Time (TST) switch architecture.

5.2.4 Light-tree-Based Source-Node Grooming OXC

Optical "light-tree' has been proposed to support multicast applications in optical WDM networks [Sahasrabuddhe and Mukherjee, 1999][Singhal and Mukherjee, 2001]. A light-tree is a wavelength tree which connects one source node to multiple destination nodes. Through a light-tree, traffic from the source node will be delivered to all destination nodes of the tree. In a light-tree, the node which originates the traffic is called the "root" node, and the nodes which terminate the traffic are called the "leaf" nodes. Note that a leaf node can also serve as an intermediate multicast node since it can both receive traffic and forward the traffic to other nodes using its multicast capability, i.e., such a node may have the "Drop-and-Continue" property. In order to support multicast, an OXC needs to duplicate the traffic from one input port to multiple output ports. For an OXC using transparent switching technology, this duplication can be realized in the optical domain using an optical splitter by splitting the power of an optical signal from one input port to multiple output ports. For an OXC using opaque technology, the traffic duplication can be easily accomplished by copying the electronic bit stream from one input port to multiple output ports. Figure 5.2(b) shows a simplified architecture of a multicast-capable OXC using the transparent technology.

Figure 5.1(b) shows how the OXCs' multicast capability can be used to perform traffic grooming. There are three low-speed traffic steams from the same source node 1 to different destination nodes 3, 5, and 6, where aggregated bandwidth requirement is lower than the capacity of a wavelength channel in this example. By setting up a light-tree, these three traffic streams can be packed

onto the same wavelength channel, and delivered to all destination nodes (i.e., light-tree leaf nodes). At each destination node, only the appropriate traffic stream is picked up and relayed to the client equipment. In this way, the low-speed traffic from the same source node can be groomed to the same wavelength channel and be sent to different destination nodes. We call this grooming scheme light-tree-based source-node grooming scheme. Please note that, if a connection between a node pair requires full wavelength-channel capacity, a light-tree becomes a lightpath. From traffic-grooming perspective, the multicast-capable OXC can be called a light-tree-based source-node grooming OXC. Such an OXC can support the light-tree-based source-node grooming scheme as well as the single-hop grooming scheme.

5.2.5 Summary of Different Grooming Schemes and Grooming OXC Architectures

As a summary, the single-hop grooming scheme can only groom traffic from the same source node to the same destination node; the light-tree-based source-node grooming scheme can groom traffic from the same source node to different destination nodes; and the multi-hop grooming schemes may groom traffic from different source nodes and to different destination nodes.

Table 5.1 summarizes the characteristics of different optical grooming OXCs. The multi-hop full-grooming OXC has the best grooming capability and provisioning flexibility, but it can only be implemented using the opaque technology. Therefore, it requires a significant amount of electronic processing, which potentially leads to poor scalability and high cost (normalized per-bit switching cost). Since it has built-in wavelength-conversion capability and full-grooming capability, the network control of this OXC encounters less physical constraints and will be relatively easy to develop.

The single-hop grooming OXC, on the other hand, has poor grooming capability and does not have too much flexibility to provision connections of different bandwidth granularities, since only the single-hop grooming scheme is supported. Both transparent (less mature) and opaque (more mature) technologies can be used to develop this type of OXC. As the OXC always switches traffic at high granularity, it may have the largest switching capacity and the lowest cost (normalized per-bit switching cost). The single-hop grooming OXC with transparent technology can have good scalability (for wavelength-band switching or fiber switching). Depending on the implementation, the OXC may employ the wavelength-continuity constraint (if it is built with transparent technology and has no wavelength-conversion capability or only has partial wavelength-conversion capability). Hence, certain intelligent control algorithms are required. Provisioning connections in a WDM network with the wavelength-continuity constraint is known as a standard routing and wavelength assignment (RWA) problem and has been well addressed in the literature.

Type \ Charac.	Grooming capability	Provisioning flexibility	Switching capacity	Cost	Scalability	Optical bypassing	Technology maturity (hardware)	Technology maturity (control software)
Single-hop grooming OXC	poor	poor	largest	low	good	can	medium	medium
Source-node grooming OXC	good	good	largest	meduim	good	can	medium	low
Multi-hop partial-grooming OXC	better	better	large	medium	good	can	medium	low
Multi-hop full-grooming OXC	best	best	small	high	poor	cannot	high	high

Table 5.1. Summary of the characteristics of different optical grooming switches.

Most characteristics of the source-node grooming OXC and the multi-hop partial-grooming OXC are between those of the single-hop grooming OXC and the multi-hop full-grooming OXC. Intelligent algorithms are needed for WDM networks which employ light-tree-based source-node grooming OXCs (or multi-hop partial-grooming OXCs) to efficiently set up light-trees (or multi-hop-groomable lightpaths). Compared with the RWA problem, there are relatively few references in the literature on this subject. More efforts are needed to investigate the traffic-grooming problem and to develop the novel algorithms, which can fully utilize the grooming capability of OXCs and optimize the network's resource efficiency.

5.3 Approaches and Algorithms

In this section, we present two approaches and algorithms to efficiently achieve the proposed grooming schemes in an optical WDM network, one for single-hop and multi-hop grooming schemes and the other for the light-tree-based source-node grooming scheme.

5.3.1 Single-Hop and Multi-Hop Grooming using an Auxiliary Graph Model

Grooming Policies and an Auxiliary Graph Model

In a traffic-groomable WDM network, there may be multiple ways to carry a low-speed connection request, i.e., there may exist multiple routes from a given source node to a given destination node, each of which may use different amount of network resources, e.g., wavelength channels, grooming ports, etc. The decision on how to choose a proper route among multiple candidate routes is known as the network operator's "grooming policy", as was discussed in Chapter 3. Different grooming policies reflect the network operator's intention

on how to engineer their traffic using the available network resources. For example, a low-speed connection can be carried through existing lightpaths, by setting up a new lightpath between the given source-destination node pair, or by a combination of both existing lightpaths and newly established lightpaths. The effect of different grooming polices on network performance was addressed in Chapters 2 and 3.

We extend the generic graph model, which was proposed in Chapter 3, to handle the single-hop grooming and the multi-hop grooming schemes. The extended model can uniformly incorporate different grooming OXC architectures (single-hop grooming OXC, multi-hop partial-grooming OXC, and multi-hop full-grooming OXC) and can easily realize different grooming policies. In this model, an auxiliary graph is constructed for a given network state. The route of a connection request is computed based on the auxiliary graph. We use a graph $G(V, E)$ to represent a given network state, where V denotes the network node set (i.e., the OXCs) and E denotes the network link set (i.e., the fiber links and the lightpath links). We use $G'(V', E')$ to denote the corresponding auxiliary graph, where V' denotes the vertex set and E' denotes the edge set. From now on, for clarity, we will use the terms *node* and *link* to represent a vertex and an edge in the original network state graph $G(V, E)$, respectively, and the terms *vertex* and *edge* to represent a vertex and an edge in the auxiliary graph $G'(V', E')$, respectively.

The auxiliary graph $G'(V', E')$ can be divided into four layers, namely the access layer, the mux layer, the grooming layer, and the wavelength layer. The access layer represents the access points of a connection request, i.e., the points where a customer's connection starts and terminates. It can be an IP router, an ATM switch, or any other client equipment. The mux layer represents the OXC ports from which low-speed traffic streams are directly multiplexed (demultiplexed) onto (from) wavelength channels and switched by the W-Fabric without going through any G-Fabric. It can be an electronic multiplexer/demultiplexer, a SONET ADM, etc. The grooming layer represents the grooming component of the network node, e.g., the grooming fabric. The wavelength layer represents the wavelength-switching capability, i.e., the W-Fabric, and the link state of wavelength channels. A network node is divided into two vertices at each layer. These two vertices represent the input port and output port of the network node at that layer.

The detailed construction of the auxiliary graph can be found in Section 4.4. Figures 4.3–4.4 illustrate examples of the auxiliary graph model.

Grooming Using Auxiliary Graph (GUAG) Algorithm

This auxiliary graph can be used to provision customers' connection requests. By assigning proper weights (i.e., administrative costs) to each edge in the auxiliary graph, suitable routes will be computed through standard shortest-path

route-computation algorithms according to different grooming policies. More detailed study on how to use this graph model to achieve different grooming policies and how these grooming policies may affect network performance can be found in Chapter 3. Based on this model, we design an algorithm, called Grooming Using Auxiliary Graph (GUAG) (see Algorithm 5.1), which can be used to provision connections of different bandwidth granularities. This algorithm can be used in a WDM network which employs single-hop grooming OXCs, multi-hop partial-grooming OXCs, or multi-hop full-grooming OXCs.

Computational Complexity of GUAG

In GUAG, the time complexity to construct an auxiliary graph for a N-node, full-wavelength-convertible WDM network is $O(N^2)$, because the auxiliary graph of such a network consists of $2 \times 4 \times N$ vertices and at most $(2 \times 4 \times N)^2$ edges. Consequently, the computational complexity of Step 1 and Step 2 in Algorithm GUAG is $O(N^2)$. Step 3 computes a least-cost route between two given vertices based on the auxiliary graph using standard shortest-path algorithms. Since the auxiliary graph has $2 \times 4 \times N$ nodes, the computational complexity of Step 3 is also $O(N^2)$. The computational complexity of the remaining steps in GUAG is $O(N)$. Therefore, the overall computational complexity of GUAG is $O(N^2)$ for a full-wavelength-convertible WDM network. Using the same analogy, we know that the computational complexity of GUAG will be $O(N^2W^2)$ in a wavelength-continuous WDM network, where W is the number of wavelength channels supported by each fiber link.

5.3.2 Source-Node Grooming Using Light-Tree Approach

The proposed graph model and the GUAG algorithm cannot be used to handle the light-tree-based source-node grooming scheme. Therefore, we design another algorithm to perform efficient traffic grooming in a WDM network employing multicast-capable OXCs. Note that, in a WDM network using the light-tree-based source-grooming scheme, the network state can be represented as a graph $G(V, E, T)$, where V denotes the network node set, E denotes the network link set, and T denotes the established light-tree set.

Grooming Using Light-Tree (GULT) Algorithm

The proposed algorithm for the light-tree-based source-grooming scheme is described in Algorithm 5.2 (GULT).

Computation Complexity of GULT

In Step 2 of GULT, the computational complexity to find the closest node i to the destination node d in a given light-tree T_s is $O(N^2)$. This is because node i can be found by constructing a shortest-path tree T_d rooted at node d. In T_d, we

Algorithm 5.1 Grooming Using auxiliary Graph (GUAG)

Input: Network state graph $G(V, E)$, and a connection request $Req(s, d, c)$ where s and d denote the source and destination nodes of the request ($s, d \in V$), and c denotes the connection's capacity requirement.

Output: A route R between nodes s and d, which satisfies the connection's capacity requirement and a new network state $G_{new}(V_{new}, E_{new})$ after provisioning the connection.

1 Construct the auxiliary graph $G'(V', E')$ according to network resource availability and the bandwidth requirement of the request, i.e., (a) if there is no free wavelength on a fiber connecting a node pair (i, j), there is no edge connecting the vertices of the wavelength layer between node pair (i, j); (b) if there is no lightpath link between a node pair (i, j), or lightpaths between (i, j) do not have enough free capacity for Req, there is no corresponding edge in the auxiliary graph; and (c) if the multi-hop partial-grooming OXC at a given node i has used up all grooming ports, there is no edge between the vertices of the grooming layer and the wavelength layer in node i.

2 Assign proper weights to each edge in G', according to the grooming policy. The grooming policy we use in this study is described as follows.

   (a) If there is any multi-hop-ungroomable lightpath between (s, d) with enough free capacity, carry Req using this lightpath.

   (b) Otherwise, a connection is provisioned through the least-cost route. In this study, the cost of a fiber link is assumed to be unity. The cost of a lightpath link is equal to the overall cost of the concatenated fiber links it traverses.

   (c) If there are multiple least-cost routes and the connection does not require full wavelength-channel capacity, choose the route which employs the minimal number of free wavelength links.

   (d) If there are multiple least-cost routes and the bandwidth requirement of the connection is full wavelength-channel capacity, choose the route which traverses the minimal number of G-Fabrics.

   Please refer to Chapter 3 on how to assign proper edge weights to the auxiliary graph according to a grooming policy.

3 Compute a route R' based on the auxiliary graph G'. If no path exists, return NULL.

4 Map R' in G' back to a route R in the original network state graph G.

5 Allocate resource and update the network state according to the route R. It may include the operations of (a) updating the number of available wavelength on fiber links along the route R, if necessary; (b) creating new lightpath links, if necessary; (c) updating the number of grooming ports in a partial-grooming OXC traversed by R, if necessary; and (d) updating the free capacity of the lightpaths involved in R, if necessary. Note that, if there are multiple applicable lightpaths between a node pair (i, j) along the route R, we will randomly choose one for Req.

6 Return R and the updated network state $G_{new}(V_{new}, E_{new})$.

Algorithm 5.2 Grooming Using Light-Tree (GULT)

Input: Network state graph $G(V, E, T)$, and a connection request $Req(s, d, c)$ where s and d denote the source and destination nodes of the request (i.e. $s, d \in V$), and c denotes the connection's capacity requirement.

Output: A light-tree T_s' which is rooted at node s and covers node d as a leaf node, and a new network state $G_{new}(V_{new}, E_{new}, T_{new})$ after provisioning the connection.

1 Let T_s denote a given established light-tree rooted at node s. If node d is one of leaf nodes of T_s, and T_s has enough free capacity for Req, let $T_s^{'}$ be equal to T_s, and go to Step 5. Note that, if multiple such trees exist, one will be randomly picked to be T_s'.

2 For any T_s which has enough free capacity for Req, compute a least-cost route from every tree node j in T_s to node d, subject to available wavelength-link constraints, and available splitter constraints at node j (if the transparent technology is used to build the OXC at node j, as shown in Fig. 5.2(b)). Let i denote the node which has the least-cost route to d among all the tree nodes. If no such tree exists, go to Step 4.

3 Find the $T_s^{'}$ and the corresponding node i' so that node i' is the closest one to node d among all tree nodes of all candidate trees. If node i' is the root node of $T_s^{'}$, go to Step 4; otherwise, expand the tree to include a new branch from node i' to node d. Therefore, node i' needs to duplicate the traffic originating from node s and switch it to node d. After that, go to Step 5.

4 Set up a light-tree from node s to node d following the least-cost route based on the current network state. Note that this special light-tree instance only has one leaf (destination) node. Hence, it is equivalent to a lightpath. Let this light-tree be $T_s^{'}$.

5 Allocate free capacity of $T_s^{'}$ to Req. If $T_s^{'}$ is a new tree or if $T_s^{'}$ is a established tree but a new branch has been added, update the corresponding wavelength-link availability information.

6 Return $T_s^{'}$ and the new network state $G_{new}(V_{new}, E_{new}, T_{new})$.

can easily find the closest node (to node d) which is also in tree T_s. Assuming that there are K candidate light-trees which are rooted at node s and can be expanded to support the connection request, the computational complexities for Steps 2 and 3 will be $O(KN^2)$. The computational complexity of the other steps in GULT is $O(N)$. Hence, the overall computational complexity of GULT is $O(KN^2)$ in a full wavelength-convertible network. In a wavelength-continuous network, GULT's computational complexity is $O(KWN^2)$, where W is the number of wavelengths supported by a fiber link.

5.4 Illustrative Numerical Results

We simulate a dynamic network environment to evaluate the performance of different optical grooming OXCs and their corresponding grooming schemes, using the proposed algorithms GUAG and GULT. The following parameters are used in our example study reported here.

- Every network node is equipped with the same type of grooming OXCs.

- The connection-arrival process is Poisson and the connection-holding time follows a negative exponential distribution.
- The capacity of each wavelength is OC-192.
- The network has full wavelength-conversion capability and each fiber link can support eight wavelength channels.
- A multi-hop partial-grooming OXC has six (incoming and outgoing) grooming ports.
- A light-tree-based source-node grooming OXC is built using the opaque technology and hence has unlimited multicast capability.
- A connection request can have any bandwidth granularity of $OC-1$, $OC-3$, $OC-12$, $OC-48$ and $OC-192$. Three bandwidth-granularity distributions for the number of connection requests ($OC-1 : OC-3 : OC-12 : OC-48 : OC-192$) are examined, namely, (a) $3:3:3:3:1$, (b) $1:1:1:1:1$, and (c) $1:1:1:1:3$.
- Connections are uniformly distributed among all node pairs.
- Average connection-holding time is normalized to unity; network offered load (in Erlang) is defined as the connection-arrival rate times the average holding time times a connection's average bandwidth and it is normalized to the unit of OC-192.
- The example network topology used in our simulation experiments is shown in Fig. 5.4.

5.4.1 Bandwidth Blocking Ratio (BBR)

Bandwidth blocking ratio (BBR) represents the percentage of the amount of blocked traffic over the amount of bandwidth requirement of all traffic requests during the entire simulation period. Note that pure blocking probability cannot reflect the effectiveness of the algorithm, as connections have different bandwidth requirements.

Figure 5.5 compares the network performance (BBR vs. load) by using different optical grooming OXCs under different connection bandwidth-granularity distributions. In Fig. 5.5, the multi-hop full-grooming OXC shows the best network performance, and multi-hop non-grooming OXC shows the worst performance. We make the following observations. (a) Without any low-granularity multiplexing and switching functionality, the light-tree-based source-node grooming OXCs can significantly improve the network performance compared to multi-hop non-grooming OXCs. (b) Connection bandwidth-granularity distribution plays an important role on the network performance. When there are

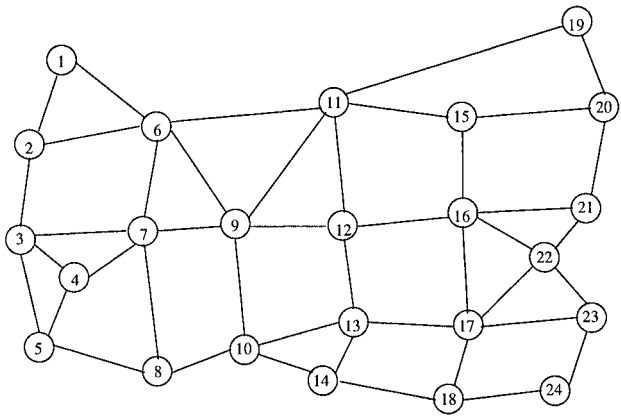
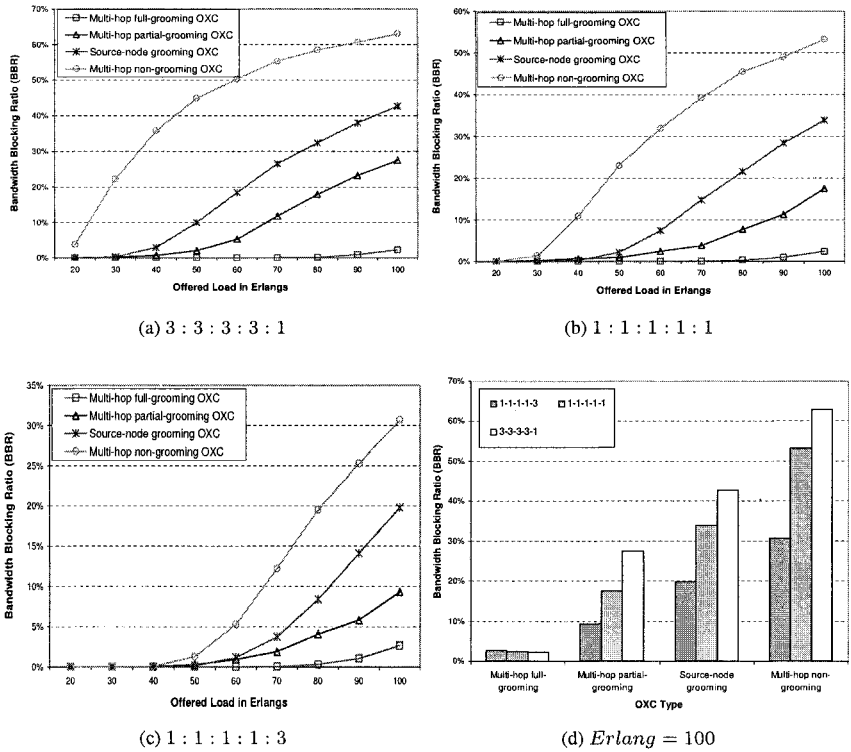


Figure 5.4. A 24-node sample network topology.

a lot of low-speed connection requests, multi-hop full-grooming OXCs outperform multi-hop partial-grooming OXCs as shown in Fig. 5.5(a). As the number of high-speed connections increases, the performance gap between the multi-hop full-grooming OXC and the partial-grooming OXC is significantly reduced. Although not shown here, our experimental results verify that, when all connections require full wavelength-channel capacity, all OXCs have the same network performance. Note that the high BBR region shown in Fig. 5.5 may not be realistic in a practical WDM backbone network. They are shown for illustration purposes to fairly compare the grooming OXCs' network performance under the same network load. (c) As shown in Fig. 5.5(d), the multi-hop full-grooming OXC performs almost the same under different bandwidth-granularity distributions. On the other hand, under the same network load, if there are more low-speed connections, other types of grooming OXCs lead to worse network performance. This is because low-speed connections may potentially cause a large bandwidth wastage and under-utilize the wavelength capacity if a network node does not have multi-hop full-grooming capability.

As we can see from Fig. 5.5, compared to the multi-hop full-grooming OXC, the multi-hop partial-grooming OXC shows reasonably good network performance, while using significantly less amount of low-speed electronic processing. Besides the grooming policy and the corresponding grooming algorithm used, the performance of multi-hop partial-grooming OXCs is determined by two factors:

- How much grooming capacity a multi-hop partial-grooming OXC may have?



(a) 3 : 3 : 3 : 3 : 1

(b) 1 : 1 : 1 : 1 : 1

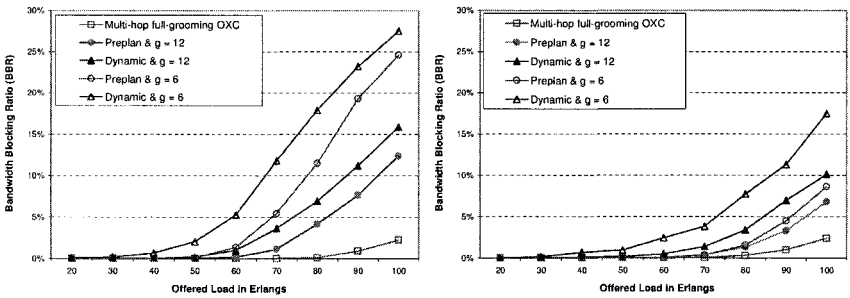
(c) 1 : 1 : 1 : 1 : 3

(d) $Erlang = 100$

Figure 5.5. Bandwidth blocking ratio (BBR) vs. load (in Erlangs) for different grooming OXCs under different bandwidth-granularity distributions.

- How to cost-effectively establish the multi-hop-groomable lightpath (i.e., source-groomable lightpath, destination groomable-lightpath, and full-groomable lightpath) to perform multi-hop grooming?

One approach to improve the performance of the multi-hop partial-grooming OXC is to increase its grooming capacity. Recall that the grooming capacity of a multi-hop partial-grooming OXC is determined by the size of its G-Fabric and the number of grooming ports (G-ports) connecting the W-Fabric and the G-Fabric. Another approach is to optimize the establishment of multi-hop-groomable lightpaths. These lightpaths can either be dynamically set up (on-demand) or be statically pre-planned. In the dynamic groomable-lightpath establishment approach, it is assumed that future traffic demands are unknown. Hence, instead of considering long-term global optimization, groomable lightpaths are set up according to the requirement of the current connection request. On the other hand, the pre-planning approach tries to pre-establish a



(a) 3 : 3 : 3 : 3 : 1 (b) 1 : 1 : 1 : 1 : 1

Figure 5.6. Effect of different lightpath-establishment schemes and different number of grooming ports on the network performance of multi-hop partial-grooming OXCs.

certain number of groomable lightpaths based on the future traffic projection. Low-speed connection requests are then dynamically provisioned using these pre-planned resources. When all pre-planned groomable lightpaths have been saturated, new groomable lightpaths can be dynamically established for new requests subject to network resource constraints.

Figure 5.6 shows how the number of grooming ports and different groomable-lightpath establishment approaches may affect the network performance of multi-hop partial-grooming OXCs under different connection bandwidth-granularity distributions. In Fig. 5.6, g denotes the number of grooming ports which a multi-hop partial-grooming OXC has. We observe that, as g increases, the performance of the multi-hop partial-grooming OXC improves. We have also simulated a very simple lightpath pre-planning scheme, called Embedded on Physical Topology (EPT). In EPT, full-groomable lightpaths are pre-established between every adjacent node pair. These lightpaths form a grooming layer, which has exactly the same topology as the physical fiber topology. A low-speed connection will be carried by jointly utilizing the resources on this grooming layer as well as on the physical topology through the grooming policies described in GUAG. Unlike a dynamically-established lightpath which will be torn down if it does not carry any traffic, a pre-planned-groomable lightpath will never be torn down. If one wavelength channel on each fiber link is used for EPT, we called it a one-degree EPT (1-EPT), which is simulated in our experiments. Figure 5.6 shows that the 1-EPT pre-planning scheme does improve the performance of multi-hop partial-grooming OXC compared to the dynamic groomable-lightpath establishment scheme. This is because the dynamic scheme may greedily establish groomable lightpaths without considering the possible future traffic demands. Although the GUAG algorithm attempts to perform local (or short-term) resource optimization, the grooming layer, which is formed by dynamically-established lightpaths, may not be optimal and effi-

cient to carry the future requests. If every multi-hop partial-grooming OXC has enough grooming capability (i.e., enough grooming ports) and W-EPT is used (where W is the number of wavelength channels supported by every fiber link), a network employing multi-hop partial-grooming OXCs will be equivalent to a network with multi-hop full-grooming OXCs everywhere. Hence, they will have the same network performance.

Note that, besides the grooming policy used in GUAG, applying other grooming policies may further improve the network performance. Please see Chapters 2 and 3 for other possible grooming policies and their effect on network performance. Similarly, besides 1-EPT, other pre-planning schemes can also be used and it is possible for them to improve the network performance. For example, the Integer Linear Program (ILP) formulation proposed in Chapter 2 can be used to pre-plan groomable lightpaths, and may achieve better network performance. A potential drawback of the pre-planning approach is that it may not be scalable. This is because network operators may have to re-do the pre-plan procedure when the network needs to be scaled or when the traffic pattern fluctuates. The effect of other preplanning schemes (2-EPT, ILP approach, etc.) should be interesting for future research

5.4.2 Wavelength Utilization

Wavelength utilization (WU) represents the average number of used wavelength links over the total number of wavelength links supported by the network during the entire simulation period.

In Fig. 5.7, we plot the average (weighted by time) WU versus the network offered load for different grooming OXCs under different connection bandwidth-granularity distributions. It is straightforward to see that, under the same network offered load, single-hop grooming OXCs will exhaust wavelength links more quickly than other OXC types. We can also observe from Fig. 5.7(d) that, under the same network offered load, the more low-speed connections the network supports, the more wavelength links tend to be used. For the same bandwidth-blocking performance, a lower wavelength utilization is desirable since fewer wavelength links are consumed to carry the same load.

5.4.3 Resource Efficiency Ratio (RER)

Wavelength utilization shown in Fig. 5.7 may not be the best metric to measure the resource usage of different grooming schemes and grooming OXCs. From Fig. 5.7, one cannot tell how efficiently the allocated wavelength channels are utilized. Resource efficiency ratio (RER) is a more suitable metric to measure the grooming performance of different OXCs and different grooming schemes. RER represents how efficiently connections are routed and groomed. It can be computed as the average (weighted by time) network carried traffic

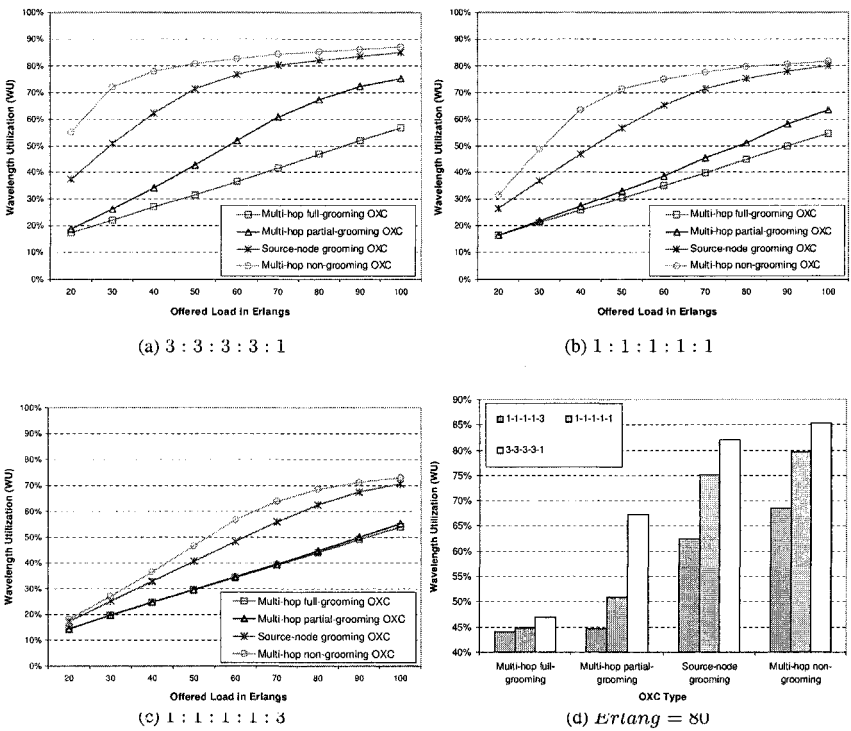


Figure 5.7. Wavelength utilization (WU) vs. load (in Erlangs) for different grooming OXCs under different bandwidth-granularity distributions.

(in terms of OC-1 units) divided by the total allocated network capacity (i.e., total number of allocated wavelength links times 192) over the entire simulation period. If we consider "minimal hops" as our objective for a route-computation algorithm, then the inverse of the average hop distance is the RER upper bound. This upper bound is achieved only when every connection requires OC-192 bandwidth and follows the shortest path. Since there are limited resources (as in our case), not every connection can follow the shortest path and the upper bound may not be achievable. Let α denote the RER, which can be computed as follows:

$$\alpha = \frac{\sum_i \rho_i \times t_i}{\sum_i \gamma_i \times t_i},$$

where t_i is the time period between the i^{th} event (connection arrival or departure) and $(i+1)^{th}$ event, ρ_i is the network carried load during the time period t_i, and γ_i is the total number of wavelength links used during t_i. (Please note that

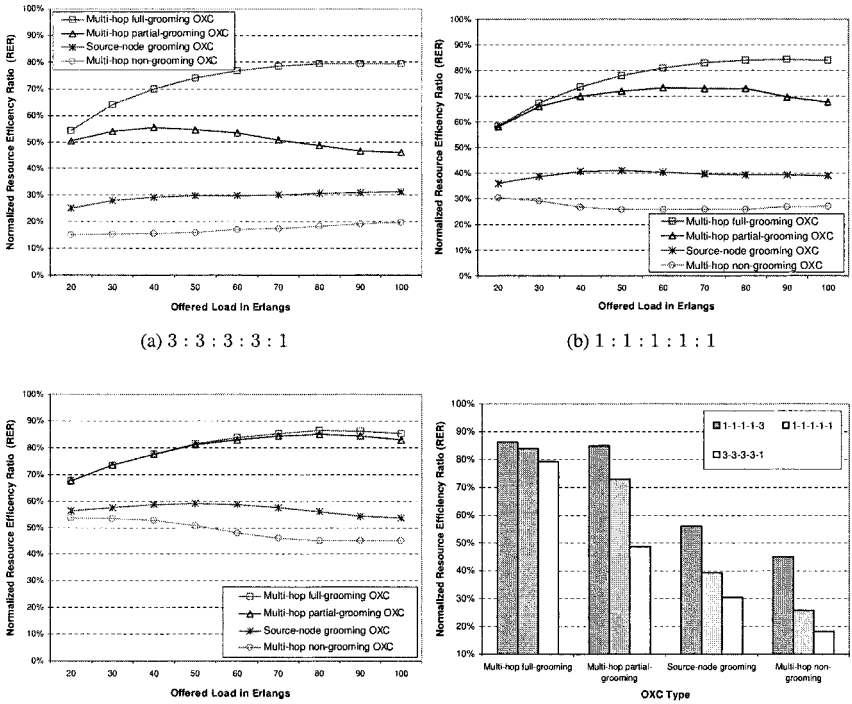


Figure 5.8. Normalized resource-efficiency ratio (RER) vs. load (in Erlangs) for different grooming OXCs under different bandwidth granularity distributions.

ρ_i and γ_i do not change during time period t_i as there is no other event during that period.)

Figure 5.8 shows the normalized (to upper bound) resource efficiency ratio (RER) versus network offered load for different grooming OXCs under different bandwidth-granularity distributions. The higher RER means that a network can route and groom traffic requests more efficiently. Thus, a network with high RER can have low bandwidth blocking ratio (BBR), which is shown in Fig. 5.5. This explains why, under the same network offered load, the multi-hop full-grooming OXC uses the least number of wavelength channels (i.e., the lowest wavelength utilization) but carries the most amount of traffic (i.e., the lowest bandwidth blocking ratio). We can also observe from Fig. 5.8 that, under the same network offered load, when the number of low-speed connections increases, the resource efficiency ratio will decrease. This is because the network

resources tend to be fragmented while the number of low-speed connections increases, and the difficulty for a network to fully utilize allocated wavelength channels also increases.

5.5 Conclusion

In this chapter, we presented four optical grooming OXC architectures and compared their characteristics. Three grooming schemes – single-hop grooming, multi-hop grooming, and light-tree-based source-node grooming – for these OXCs were investigated. We proposed two algorithms, GUAG and GULT, to efficiently provision connections of different bandwidth granularities. The performance of different grooming OXCs and grooming schemes was compared using the proposed algorithms under a dynamic traffic environment. We also investigated the effect of different bandwidth-granularity distributions on the network performance of different grooming OXCs. We observed that the light-tree-based source-node grooming using the OXCs' multicast capability can significantly improve the network performance of the single-hop grooming OXC without employing any low-granularity electronic processing. The multi-hop full-grooming OXC showed the best network performance in terms of network bandwidth blocking ratio, wavelength utilization, and resource efficiency ratio. However, it may encounter a scalability problem since a large amount of electronic processing is required at low-speed granularity. With a few nodes having low-granularity switching capability, the multi-hop partial-grooming OXC showed good performance compared to other grooming OXCs, which makes it as a good alternative when the multi-hop full-grooming is not needed at every network node. In order to fairly compare their performance, we assumed that a network uniformly employs one type of grooming OXCs. This assumption may not always be practical and can be relaxed in next-generation optical backbone networks, where different OXCs with different grooming capabilities may be inter-connected and may need to co-exist. This problem, which was initially studied in [Zhu et al., 2003b], will be discussed in Chapter 7. The proposed algorithms, GUAG and GULT, employed one traffic-grooming policy and light-tree establishment approach. Other algorithms with different traffic-grooming policies and light-tree establishment approaches will be interesting topics for future research.

Chapter 6

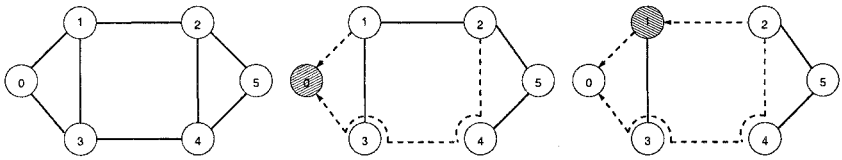
SPARSE GROOMING NETWORK

In most previous research on traffic grooming in WDM mesh networks, it is assumed that every network node has traffic-grooming capability, which may not be practical or cost-effective in a nationwide WDM backbone network.

Figure 5.2(a) shows a sample G-OXC architecture. There are two switching fabrics in this OXC, a wavelength-switching fabric (W-Fabric) and a grooming fabric (G-Fabric). Because a grooming OXC may be more costly than an OXC without grooming capability (i.e., the OXCs which only have the W-Fabric), and the G-OXCs from different vendors may have different grooming capability, in a multi-vendor inter-operational WDM mesh network, only a few network nodes may have traffic-grooming capability. We call this type of network a *"sparse-grooming network"*, and we call a node which has traffic-grooming capability to be a *grooming node* (G-Node).

Hence, the problem of designing a sparse-grooming WDM mesh network is a very important and practical problem. In this chapter, we investigate the problem of efficiently designing a sparse-grooming WDM mesh network for a given set of traffic requests via theoretical formulation as well as simulation experiments. The results from our research indicate that, through careful network design, a sparse-grooming WDM network can achieve similar network performance as a full-grooming network, while significantly reducing the network cost.

Figure 6.1 shows an example of designing a sparse-grooming WDM mesh network. Figure 6.1(a) shows a six-node network, where each edge represents a pair of unidirectional fiber links. For simplicity of illustration, let us assume that each fiber supports one wavelength channel, which can carry two low-speed connection requests. And only one G-OXC is allowed to be used in the network. Assume that there are four low-speed requests, among which three are between node pair $(2, 0)$ and one between node pair $(1, 0)$. Two network designs are



(a) Network 1: A 6-node network (b) Design I (c) Design II

Figure 6.1. A sample network and two sparse-grooming network designs.

shown in Fig. 6.1. The shaded nodes represent the G-Nodes, and the dashed lines represent established lightpaths.

In Design I (Fig. 6.1(b)), node 0 is the G-Node. Two lightpaths — $(1,0)$ and $(2,0)$ — are established. Three requests can be satisfied by these lightpaths: two between $(2,0)$, and one between $(1,0)$. In Design II (Fig. 6.1(c)), three lightpaths are set up: $(2,0)$, $(1,0)$, and $(2,1)$. Note that, in Design II, all requests can be satisfied. Thus, this simple example illustrates that improved performance can be achieved by carefully choosing the G-Node and engineering the traffic in the network.

In this chapter, we study the problem of how to efficiently design such a sparse-grooming WDM mesh network. Section 6.1 gives the formal problem statement. Then, mathematical formulations for two objective functions are presented. Fast heuristic algorithms, which can be used to handle large network topologies, are proposed in Section 6.2. Then, illustrative numerical results from mathematical formulations as well as heuristics are shown and analyzed in Section 6.3. Finally, Section 6.4 concludes the study.

6.1 Problem Statement and Mathematical Formulation

We formally define the problem as follows: given a network topology $G(V,E)$, where V represents the node set and E represents the link set of the network; given a traffic matrix Λ, where each element represents the number of low-speed requests between a node pair; assume that each wavelength channel can support g low-speed requests (where g is known as *grooming ratio*); design the network such that either one of following two objectives can be optimized:

1 For a given amount of network resources (number of G-OXCs N_g, and number of wavelengths W on each fiber), maximize the network throughput.

2 Carry all traffic requests, while simultaneously minimizing the overall network cost, which is determined by the number of wavelength channels and G-OXCs used in the network.

Figure 6.2 shows a WDM network in which two low-speed requests ($C1$ and $C2$) are being carried. Both requests are from node 1 to node 5. In Fig. 6.2,

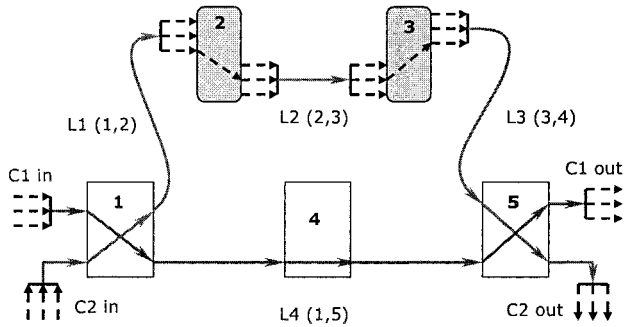


Figure 6.2. A sample sparse-grooming WDM network which carries two requests using four lightpaths.

nodes 2 and 3 are G-Nodes, and only their G-Fabrics are shown. Nodes 1, 4, and 5 are equipped with OXCs without grooming capability. Figure 6.2 shows that, in a sparse-grooming mesh network, a low-speed request can be carried either by a single lightpath ($C1$) or by traversing multiple lightpaths and G-Nodes ($C2$).

We formulate the problem mathematically and it turns out to be an integer linear program (ILP). To conserve space, only the most important equations of the formulation are presented here. The following notations and variables are used:

- (m, n), (i, j), and (s, d) denote the end nodes of a fiber, a lightpath, and a low-speed request, respectively.
- There are four types of lightpaths, represented by (i, j), (i', j), (i, j'), and (i', j'). (i, j) denotes a lightpath which does not connect with any G-Fabric at its end nodes, e.g., $L4(1, 5)$ in Fig. 6.2. (i', j), (i, j'), and (i', j') denote the lightpaths which are connected to G-Fabrics at the source node, the destination node, and both nodes, respectively, e.g., $L3$, $L1$, and $L2$ in Fig. 6.2(b). We use V_{ij}, $V_{i'j}$, $V_{ij'}$, and $V_{i'j'}$ to denote the numbers of each type of lightpaths.
- P_{mn}^{ij}, $P_{mn}^{i'j}$, $P_{mn}^{ij'}$, and $P_{mn}^{i'j'}$ denote the number of wavelengths which have been used to support each type of lightpaths, on fiber (m, n). P_{mn} denotes the number of fiber links between node pair (m, n).
- T_{ij}^{sd}, $T_{i'j}^{sd}$, $T_{ij'}^{sd}$, and $T_{i'j'}^{sd}$ denote the total amount of traffic between node pair (s, d), which are being carried by the lightpaths (i, j), (i', j), (i, j'), and (i', j'), respectively. T_{sd} denotes the successfully carried traffic between (s, d). Λ_{sd} denotes the total offered traffic between (s, d). Note that $T_{sd} \leq \Lambda_{sd}$.

- $M_i = 1$ if node i is a G-Node; otherwise, $M_i = 0$.

6.1.1 Maximizing Network Throughput

Given W wavelengths per fiber link and N_g as the number of G-OXCs which can be used in the network, the problem can be formulated as follows:

- Objective Function:

$$Maximize : \sum_{s,d} T_{sd} \tag{6.1}$$

- Constraints:

$$\sum_{m} P_{mk}^{ij} = \sum_{n} P_{kn}^{ij} \tag{6.2}$$

$$\sum_{i,j} (P_{mn}^{ij} + P_{mn}^{i'j} + P_{mn}^{ij'} + P_{mn}^{i'j'}) \leq W \times P_{mn} \tag{6.3}$$

$$T_{sk'}^{sd} + \sum_{i} T_{i'k'}^{sd} = T_{k'd}^{sd} + \sum_{j} T_{k'j'}^{sd} \tag{6.4}$$

$$\sum_{s,d} T_{i'j'}^{sd} \leq g \times V_{i'j'} \tag{6.5}$$

$$T_{sd}^{sd} + T_{s'd}^{sd} + T_{sj'}^{sd} + T_{s'j'}^{sd} = T_{sd} \tag{6.6}$$

$$T_{sd}^{sd} + T_{sd'}^{sd} + T_{i'd}^{sd} + T_{i'd'}^{sd} = T_{sd} \tag{6.7}$$

$$\sum_{j} (V_{i'j} + V_{i'j'} + V_{ji'} + V_{j'i'}) \leq M_i \times D \tag{6.8}$$

$$\sum_{i} M_i \leq N_g \tag{6.9}$$

Equation (6.2) is the flow-conservation equation at node k for establishing lightpaths (i, j), which may use node k as an intermediate node. There are similar equations for lightpaths (i', j), (i', j), and (i', j'). Equation (6.3) is the resource-constraint equation for a fiber link (m, n), i.e., the total number of lightpaths (four types) carried by fiber (m, n) cannot exceed the total number of wavelength channels in fiber (m, n). Equation (6.4) is the flow-conservation equation at any intermediate node k for routing the requests between (s, d), which may use node k as an intermediate G-Node. Equation (6.5)

is the resource-constraint equation for lightpaths (i', j'), i.e., the total number of low-speed connections carried by the lightpath (i', j') cannot exceed the overall capacity of lightpath (i', j'). There are similar equations for lightpaths (i, j), (i', j), and (i, j'). Equation (6.6) guarantees that, for a given node pair (s, d), the amount of traffic successfully flowing out from the source node should be equal to the amount of traffic that can be successfully carried between (s, d). Equation (6.7) captures the similar constraint at the destination side. Equation (6.8) ensures that node i is a G-Node if there is a lightpath connected to its G-Fabric (i.e., node i initiates or terminates a lightpath at its G-Fabric). D is a large constant, which can be the upper bound on the maximum number of lightpaths that can originate from any node. Equation (6.9) ensures that there are at most N_g G-Nodes in the network.

6.1.2 Minimizing Network Cost

Let C_w denote the cost of supporting one wavelength channel in the network, and C_g denote the extra cost of employing a G-OXC instead of an OXC without grooming capability. The second network-design objective can be achieved using the following formulation. Most equations are the same as those in Section 6.1.1, and we only present the different ones below:

- Objective Function:

$$Minimize : W \times C_w + N_g \times C_g \tag{6.10}$$

- Constraints:

$$T_{sd} = \Lambda_{sd} \tag{6.11}$$

6.2 Heuristic Approaches

The computational complexity makes the ILP formulations only suitable for designing small or moderate-sized WDM networks. When the network size is large, i.e., there are several tens of network nodes (e.g., Fig. 5.4), efficient heuristics are needed to solve the problem.

For a given network topology and given values of N_g and W, we design a heuristic $\delta(N_g, W)$ which chooses G-Nodes and routes the traffic requests to maximize the network throughput (for the first objective function). The second objective function can be achieved by starting with small values of N_g and W, and gradually increasing N_g and W until all the traffic requests are satisfied. In this section, we concentrate on the development of such a heuristic $\delta(N_g, W)$ for the first objective function.

Given a WDM network topology $G(V, E)$ and a list of traffic demands Λ, the heuristic $\delta(N_g, W)$ is composed of the following two steps.

1 Choose N_g nodes as G-Nodes from V, based on certain cost function $F_c(v)$ for a given node v (to be elaborated below).

2 Route the traffic requests on the network, subject to the network resource constraints. Let $F_r(\delta)$ denote the algorithm to be used to route the traffic requests. When the traffic is static (known a priori), the algorithm may contain some backtracking procedure; when the traffic is dynamic or incremental, no backtracking is allowed, and the requests will be carried on a first-come first-served basis. In this study, we consider static traffic.

6.2.1 Grooming-Node-Selection Schemes

We propose and study the characteristics of three different cost functions for selecting G-Nodes, namely nodal-degree selection, bypass-traffic selection, and random selection. Note that some ideas on these node-selection schemes are borrowed from sparse-wavelength-converter-placement studies [Iness and Mukherjee, 1999, Arora and Subramaniam, 2000].

- *Nodal-Degree Selection*: In this scheme, the first N_g nodes which have the maximum nodal degree are picked to be G-Nodes. If several nodes have same nodal degree and only some of them can be chosen, random selection is used to break any ties.

- *Bypass-Traffic Selection*: For a given node v, $F_c(v)$ is computed as the total amount of traffic which may bypass the node, assuming each traffic request is routed on the physical network topology using a standard shortest-path routing algorithm. The N_g nodes which have the maximum amount of bypass traffic can be selected as the G-Nodes. Instead of routing the traffic requests between a node pair (s, d) using a single shortest-path route, it may be also possible to compute K $(K \geq 2)$ alternate paths between (s, d) and bifurcating the traffic among these K alternate paths.

- *Random Selection*: In this scheme, N_g nodes are randomly picked to be G-Nodes.

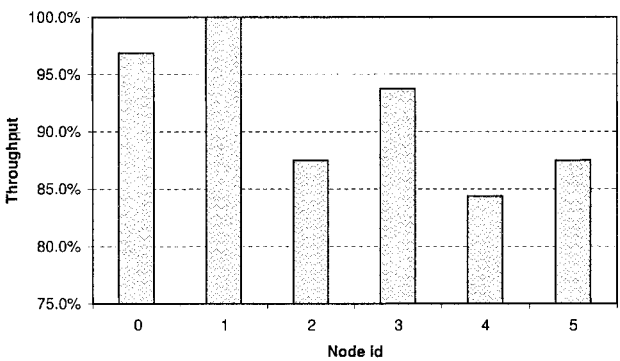
One can also design other schemes to select the G-Nodes. Similar heuristics exist in the study of sparse wavelength conversion in WDM networks [Iness and Mukherjee, 1999, Arora and Subramaniam, 2000].

6.2.2 Traffic-Routing Schemes

A simple algorithm is used in our study to perform traffic routing after the network resources have been determined and allocated. Given a set of traffic requests, a request list is generated based on a random permutation of all the requests. The requests are then served sequentially following their order in the

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

(a) Traffic Matrix



(b) Network Throughput vs. Grooming Node Id

Figure 6.3. Illustrative results from ILP formulation for the network in Fig. 6.1 assuming only one node has grooming capability.

list. L such random lists can be generated and tried. The best one among these L results will be taken as the final result. In our study, we choose L as 50. If L is equal to 1, the traffic pattern is equivalent to incremental traffic where connection requests arrive one at a time.

6.3 Illustrative Numerical Examples

Figure 6.3 shows an illustrative numerical example based on the sample network shown in Fig. 6.1(a). For simplicity of exposition, let each fiber in this example support two wavelengths ($W = 2$), each wavelength can carry two low-speed connection requests ($g = 2$), and there is one G-OXC which needs to be placed in the network ($N_g = 1$). Figure 6.3(a) shows a randomly-generated traffic matrix, in which each element represents the number of low-speed connection requests between a node pair. The total number of low-speed requests is 32 in this example. We use "CPLEX", a commercial optimization

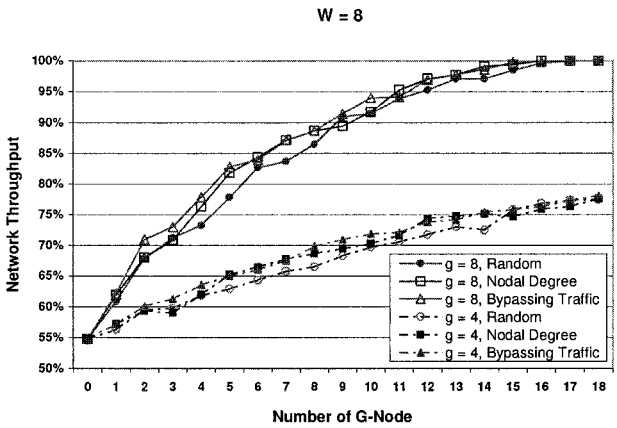


Figure 6.4. Performance comparison between different G-Node selection schemes applied to the network in Fig. 5.4.

software package, to solve the ILP formulation. For this example, our study shows that node 1 is the best candidate node to be the G-Node. Then, we force our formulation to artificially select each of the other nodes separately to be the G-Node and compare the network throughput with that of the best design (i.e., selecting node 1). The performance comparison is shown in Fig. 6.3(b). The horizontal axis in Fig. 6.3(b) represents the node chosen as the G-Node. The vertical axis in Fig. 6.3(b) shows the optimal network throughput (obtained via ILP formulation) by choosing the corresponding node as the G-Node. It is observed that 100% network throughput can be achieved if the G-Node is node 1; at the other performance extreme, if node 4 is chosen as the G-Node, the network throughput is below 85%. The results indicate that a network operator can increase the network throughput as well as reduce the network cost (using less grooming equipment in the network) by carefully designing a sparse-grooming WDM mesh network.

When a network has several tens of nodes, the ILP approach becomes computationally intractable; hence, heuristic will need to be used. Figures 6.4 and 6.5 show the results obtained on the network topology of Fig. 5.4, using the heuristic algorithms proposed in Section 6.2. The network contains 24 nodes and 43 bidirectional fiber links. Traffic demands are randomly generated (integer values) between each node pair with uniform distribution between $(0, 3)$. Each fiber link can support eight wavelength channels ($W = 8$). Different values of the grooming ratio g will be investigated below.

Figure 6.4 shows the performance comparison between different G-Node selection schemes. It can be observed that random selection does not perform

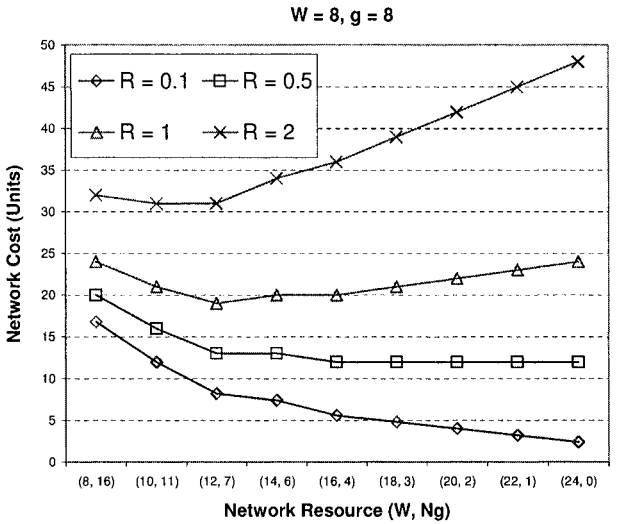


Figure 6.5. Network cost vs. network resources based on different cost ratio R.

as well as the others. Selecting the G-Node by bypassing traffic achieves better performance than selecting the G-Node by nodal degree in most cases. This is because the Bypass-Traffic-Selection Scheme considers the network topology as well as the traffic intensity at each node. We can also observe that, if the grooming ratio is equal to eight ($g = 8$), when the number of G-Nodes exceed a certain bound ($N_g = 16$ in this case), no additional performance gain can be achieved by having more grooming nodes in the network. We can also observe that, when the grooming ratio is large (e.g., $g = 8$), having more grooming nodes will achieve better network performance than the case when the grooming ratio is small (e.g., $g = 4$).

Figure 6.5 shows different network costs based on different values of the cost ratio R, which is defined as follows:

$$R = C_w / C_g \tag{6.12}$$

where C_w denotes the cost (equipment cost, operational cost, etc.) to support one wavelength channel in the network, and C_w denotes the cost to employ grooming capability in a network node. The network cost C_n can be represented as:

$$C_n = N_g \times C_g + W \times C_w = (N_g + R \times W) \times C_g \tag{6.13}$$

The horizontal axis (W, N_g) in Fig. 6.5 represents, for a given number of wavelength channels on each fiber link (W), how many G-Nodes (N_g) are needed to

satisfy all the requests. By normalizing the cost C_g to be one unit, the vertical axis represents the overall network cost computed using Eqn. (6.13).

We can observe in Fig. 6.5 that, for a given cost ratio R, an optimal design with minimal overall network cost can be achieved. This optimal configuration reflects the cost trade-off between a grooming node and a wavelength channel. When the cost of a wavelength channel is cheaper compared to the cost of a grooming node (for $R = 0.1$, and $R = 0.5$), more wavelength channels and less grooming nodes should be used, and vice versa (for $R = 2$). Note that, although the cost of a wavelength channel will decrease as WDM technology keeps maturing, supporting more wavelength channels in a nationwide WDM backbone network may still be expensive because of the large geographic distance between the network nodes, the number of fiber links a backbone network may have, and the cost for network monitoring, maintenance, and management.

6.4 Conclusion

This study was devoted to the problem of designing a WDM mesh backbone network with sparse traffic-grooming capability. The mathematical formulations (ILPs) for two design objective functions were presented. Due to the large computational complexity of ILPs, three heuristic algorithms were also proposed to solve large instances of the problem. Our results from both the mathematical formulations and heuristics show that, by employing a limited number of grooming nodes, the network capacity can be used more efficiently and the network performance can be improved significantly. We also showed that it is possible to find a balance between the number of wavelength channels and the number of grooming nodes used in the network. This balance will eventually reduce the network cost. Further study is needed on the effect of sparse grooming on dynamic traffic, and more intelligent heuristic algorithms could also be developed.

Chapter 7

NETWORK DESIGN WITH OXCS OF DIFFERENT BANDWIDTH GRANULARITIES

7.1 Introduction

As WDM technology advances, the capacity of a wavelength channel continues to increase (from OC-192 to OC-768 and possibly beyond). However, the bandwidth requirement of a typical connection request (referred to as a traffic demand) is versatile (e.g., STS-1, OC-3, OC-12, OC-48, and OC-192), and usually a small fraction of the bandwidth of a WDM channel. To efficiently use the bandwidth, grooming switches have been developed which can pack/unpack low-speed connections onto/from high-speed WDM channels and switch at sub-wavelength granularities. Different grooming switches may have different grooming granularities. For instance, some grooming switches can groom at STS-1 level, i.e., they are capable of unpacking a wavelength channel down to STS-1 timeslots, switching those STS-1 timeslots independently, and packing them back onto wavelength channels. Some other grooming switches may do grooming only at OC-48 level (assuming the capacity of a wavelength channel is greater than OC-48, say OC-192). Although this kind of grooming switches provides less flexibility in grooming, the port cost may be less than that of STS-1 grooming switches. These two kinds of grooming switches are both opto-electro-opto (OEO) switches, and need to be surrounded by transponders, which perform O/E and E/O conversion. On the other hand, all-optical OXCs do not need transponders for OXC ports, which is a significant savings in optical transport networks, but at the price of no grooming capability and wasting channel capacities due to under-utilized wavelength channels. Since the WDM backbone network topologies are usually irregular and traffic requests are of different bandwidth granularities, it is not necessary to deploy the same kind of OXC at all the nodes, especially if the OXCs are interoperable, which seems to be a major goal of current standards activities, e.g., those of the Optical In-

ternetworking Forum (OIF). If the granularities of all the OXCs in a network are the same, we call this network *granularity-homogeneous network*; otherwise, we call it *granularity-heterogeneous network*. When designing a WDM backbone network, it is desirable to exploit the benefits of all types of OXCs to accommodate the traffic — which is typically non-uniform across all node pairs — while reducing the network capital expenditures.

In this chapter, we focus on the design of a WDM mesh backbone network with OXCs of different bandwidth granularities to minimize the network-wide OXC port cost. Specifically, we determine the type of OXC at each node, compute the route of each traffic request, and calculate the total OXC port cost. To the best of our knowledge, this is the first research report which considers the design of a network using OXCs with different switching granularities. As related work, in [Ranganathan et al., 2002], the authors compare the network cost when using different node architectures, but they assume all the grooming nodes to have the same STS-1 grooming capability.

The chapter is organized as follows. Section 7.2 presents the problem formulation and the challenges faced when designing a granularity-heterogeneous network. In Section 7.3, we construct the auxiliary graph for the network with different node architectures. Based on the auxiliary graph, a traffic-provisioning algorithm is proposed in Section 7.4, which is employed by the framework for network design. Section 7.5 demonstrates an example of a network design using the framework. Section 7.6 concludes the chapter.

7.2 Problem Statement and Challenges

7.2.1 Problem Formulation

The network design problem addressed here can be formulated as follows.

- Given:
  - the physical topology of a network,
  - a traffic matrix which contains various bandwidth requirement between different nodes,
  - the types of OXCs which can be deployed in the network, and
  - the port cost of each type of OXC.
- We need to determine:
  - the type of OXC at each node, and
  - the route for each traffic request.
- Objective:
  - minimize the total OXC port cost of the network while accommodating all the traffic demands.

A traffic demand is represented by the notation $T(s, d, g, m)$, where s and d are the source and destination nodes, respectively; g is the granularity of the traffic demand, for instance, OC-3 or OC-48; and m is the amount of the traffic in units of g. Without loss of generality, we assume the finest granularity of the bandwidth of a traffic demand is STS-1 for the rest of the chapter.

Note that another version of the problem formulation can be specified where each node in the network can have more than one type of OXC. Certainly, this is a more general case. However, only one OXC at each node is more practical today due to ease of maintenance and cost; hence, we proceed with this version in the rest of this study while noting that our approach can easily be adapted to the general node architecture with multiple types of OXCs.

7.2.2 Challenges in Designing a Granularity-Heterogeneous Network

When designing a network with OXCs of different bandwidth granularities, the first question we need to answer is: How to choose an OXC for each node? When determining which type of OXCs should be used at a node, there are several factors we need to consider: traffic originating from or terminating at this node, bypassing traffic, and the types of OXCs at other nodes. For instance, if most of the traffic from and to this node and bypassing traffic are of granularity of lightpaths, it is reasonable to put an OXC with no grooming capability at this node. Hence, the route of the traffic should be considered when determining the node architecture. On the other hand, the route of the traffic demand also depends on the type of OXCs at each node. If the type of OXC at a node is changed, some traffic in the network may need to or have to change the route to adapt to the change. Therefore, the type of OXCs at each node and the route of traffic interplay with each other, causing the network-design problem to be a difficult one.

We also need to ensure that the designed network can fully accommodate the traffic demands, and furthermore, we need to determine how to route the traffic in the network. However, even in a granularity-heterogeneous network, in which the type of OXCs at each node is already known, routing a connection request and representing the network state are still significantly complex problems, compared with those in a network where only OXCs with STS-1 grooming granularity or OXCs with no grooming capability exist.

Consider a network where some nodes have STS-1 grooming capability, while the others have no grooming capability. Although this network is granularity-heterogeneous, the complexity of computing a route for a connection and representing the network state is similar to that in a network where only OXCs with STS-1 grooming granularity or OXCs with no grooming capability exist. Note that, in this special case, an OXC either has finest grooming granularity or no grooming capability. In general, however, the routing and

representation problem is more complicated when there are OXCs with other granularities. In this chapter, we address the general case unless stated otherwise.

In a network consisting of only non-grooming OXCs, the free capacity of a lightpath can only be used by the connections which have the same source and destination node as the lightpath. In an STS-1 grooming granularity network, if there is a lightpath between two nodes, only the timeslots[1] carrying traffic are switched to outgoing lightpaths or dropped at the end node, and all the other free timeslots are not switched, which means the free capacities on the lightpath can always be accessible by the end nodes of the lightpath. However, in a multi-granularity network, OXCs switch traffic at different granularities. For a given traffic demand requiring certain bandwidth, certain amount of timeslots on the lightpaths along the route of the traffic are allocated to the connection. At a STS-1 grooming switch, only the timeslots occupied by the traffic are switched from the incoming OXC port to the outgoing OXC port. However, at an OXC with a switching granularity coarser than the bandwidth granularity of the traffic request, some free timeslots may also be switched together with those timeslots taken by the traffic, causing these free timeslots to bypass this node and become unavailable to this node. The fundamental observation is that the timeslots within the switching granularity of an OXC are transparent to the OXC and these timeslots can only be operated as a whole.

Here, we give a small example. To carry a traffic demand $T_1(1, 4, \text{STS-1}, 2)$, we setup lightpaths L_1 (from node 1 to node 2), L_2 (from node 2 to node 3), and L_3 (from node 3 to node 4), and then route T_1 onto these lightpaths. (Note that the terms STS-1 and OC-1 are used interchangeably.) The switching granularities of the nodes 1, 2, 3, and 4 are STS-1, OC-3, STS-1, and STS-1, respectively. The capacity of a wavelength channel is OC-12 for this illustration. Figure 7.1 shows the switching state of the OXCs.

Since the capacity of a lightpath is OC-12, we can view a lightpath as a communication channel with 12 timeslots (TS_1 to TS_{12}). Node 1 puts traffic T_1 into the first two timeslots, TS_1 and TS_2, of lightpath L_1. Since the switching granularity of node 2 is OC-3, it can only unpack lightpath L_1 into 4 segments (each of capacity OC-3), with 3 timeslots in each segment, and node 2 can switch these segments as individual entities. In order to route traffic T_1 onto lightpath L_2, node 2 switches the first segment in lightpath L_1 to a segment, say the third segment, in lightpath L_2. Now the traffic, if any, in timeslots TS_1, TS_2, and TS_3 of lightpath L_1 will be in timeslots TS_7, TS_8, and TS_9 of lightpath L_2, respectively. Note that, although the order of the timeslots where the traffic is put may change when the traffic is switched at node 2, the order of

[1] Without loss of generality, the time domain is assumed here for traffic multiplexing within a wavelength channel.

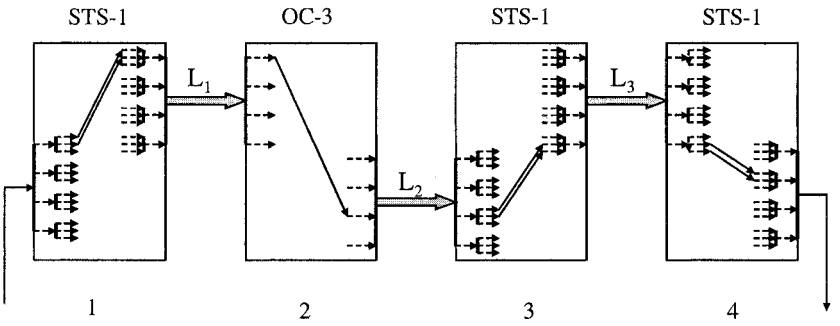


Figure 7.1. State of the switches when routing traffic demand T_1 of bandwidth STS-1 from node 1 to node 4.

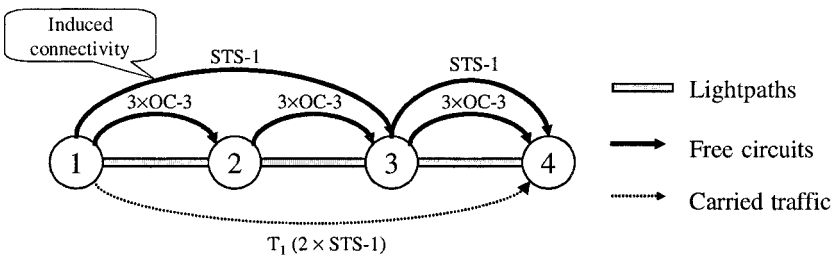


Figure 7.2. Network state after routing T_1 traffic demand.

the timeslots within a segment cannot be changed because a segment is treated as an integral entity. At node 3, lightpath L_2 is terminated and unpacked into timeslot level since the switching granularity of node 3 is STS-1, i.e., at the timeslot level. Here, only the timeslots used by traffic T_1, i.e., timeslots TS_7 and TS_8, are switched onto lightpath L_3; timeslot TS_9 will not be switched at this instant, and it can be switched to any free timeslot in any outgoing lightpath or dropped at this node later for future traffic. Finally, traffic T_1 reaches node 4 via lightpath L_3 and is dropped at node 4.

Because the switching granularity of node 2 (OC-3) is coarser than the bandwidth granularity of the traffic demand (2 × STS-1), there is a free STS-1 timeslot (timeslot TS_3 in L_1) switched onto L_2 (timeslot TS_9 in L_2) by the OXC at node 2. Although this timeslot goes through node 2, it cannot be accessed by node 2 due to the switching configuration for traffic T_1. Any traffic carried by this timeslot will bypass node 2 and directly reach node 3, where it can be switched to any free outgoing lightpath or be dropped at that node. This is equivalent to having an STS-1 circuit directly connecting node 1 and

node 3, and this circuit is called an *induced connectivity*. If there is an STS-1 traffic demand from node 1 to node 3 later, it can be directly put into timeslot TS_3 in lightpath L_1 at node 1 and dropped from timeslot TS_9 in lightpath L_2 at node 3, without any change in switching configuration at node 2. Figure 7.2 shows the network state (virtual connectivity) after routing T_1. These circuits form another topology above the virtual topology, in which each edge is a lightpath, and traffic demands should be routed on this topology instead of on the virtual topology.

7.2.3 Our Approach

To accommodate the diverse characteristics of multi-granularity networks, we enhance the graph model proposed in Chapter 3. The original graph model can route a traffic demand according to the current network state, and update the network state after carrying the traffic, but only for the network in which all the grooming OXCs have the same (STS-1) grooming capability. The graph model has been extended by us now so that it can be applied to a network with OXCs of different grooming granularities. The extended graph model can also represent the situation where several types of OXCs co-exist at the same node, and intelligently choose the appropriate type of OXC to carry the traffic demand. Based on this enhanced and versatile model, we investigate a framework for designing a granularity-heterogeneous network.

7.3 Construction of an Auxiliary Graph

In the graph model, we construct an auxiliary graph according to the network state, compute a route in the auxiliary graph for a given traffic request, set up the connection according to the route, and then update the auxiliary graph to reflect the changes in the network state. We first demonstrate how to construct the auxiliary model for a node according to the node architecture, and then show how to build the auxiliary graph for the entire network.

7.3.1 Node Representation

The representation of a node in the graph model is determined by the node architecture. Different node architectures have different graph representations. With respect to node architecture, we focus on how OXCs with different switching granularity are interconnected, if there are multiple OXCs at the same node, and where the traffic requests can be originated and terminated.

Figure 7.3 depicts an example of a generalized node architecture. In this node architecture, there are three OXCs: OXC O_1 with STS-1 grooming granularity, OXC O_2 with OC-48 grooming granularity, and OXC O_3 with no grooming capability which can only switch at lightpath granularity, say OC-192 in this example. O_1 and O_2 are connected to O_3 which is directly connected to in-

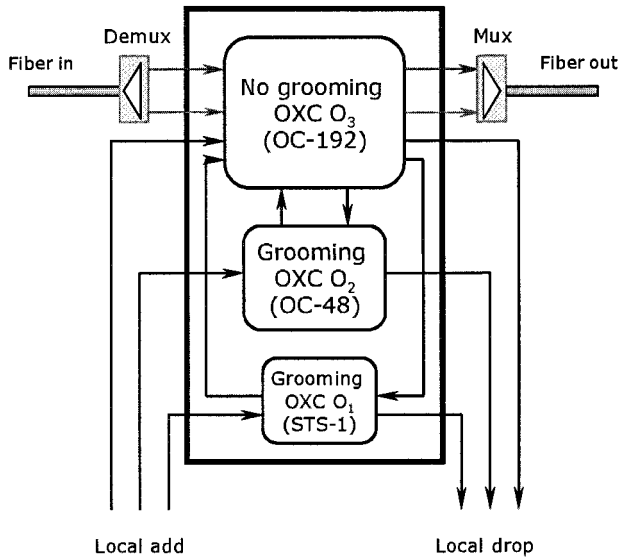


Figure 7.3. A node with three different types of OXCs. (Each data path could be a multi-line, i.e., there may be multiple fibers in and out of the OC-192 OXC, multiple add and drop ports for each OXC, etc.)

coming and outgoing fibers. Local traffic can be added to or dropped from all three OXCs. Assume that there are two wavelengths in this network (λ_1 and λ_2). Also, for purpose of illustration, assume that this node can convert λ_1 to λ_2 without using grooming OXCs, but not vice versa. The corresponding auxiliary graph for this node architecture is constructed as in Fig. 7.4.

In general, the auxiliary graph for node i can be represented by $G_i(V_i, E_i)$, where V_i and E_i are its vertex set and edge set, respectively. In general, if there are W wavelengths on each link in the network, λ_1 through λ_W, and C possible grooming granularities at node i, F_1 through F_C, where F_1 is the finest grooming granularity and F_C is the coarsest grooming granularity, the auxiliary graph for node i is constructed as a layered graph with $(W + C + 3)$ layers. If the capacity of a lightpath is OC-192, the typical grooming granularities can be STS-1, OC-3, OC-12, and OC-48, as shown in Fig. 7.4.

- Layers 1 through W denote the W *wavelength layers*,
- Layer $(W + 1)$ is called the *transponder layer*,
- Layer $(W + 2)$ is called the *lightpath layer*,
- Layers $(W + 3)$ through $(W + C + 2)$ denote the C *grooming layers*, i.e., from F_C grooming layer to F_1 grooming layer, and

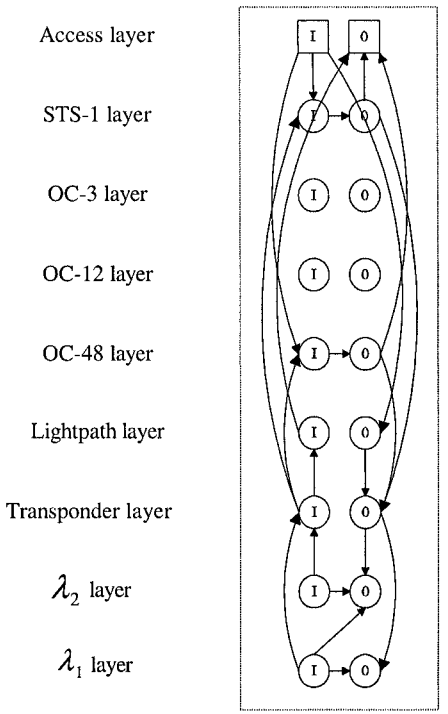


Figure 7.4. Auxiliary graph for the node.

- Layer $(W+C+3)$ is called the *access layer*, where a traffic flow starts and terminates.

There are two ports on each layer, an input port and an output port, shown as a vertex marked with "I" and a vertex marked with "O", respectively, in Fig. 7.4. (The numbers 0 and 1 will be used to refer to the input port and the output port in the discussion below.) Let $N_i^{l,p}$ denote port p on layer l at node i, then $V_i = \{N_i^{l,p} \mid p \in \{0,1\},\ 1 \le l \le (W+C+3)\}$, where $N_i^{l,0}$ and $N_i^{l,1}$ denote the input port and the output port on layer l at node i, respectively. According to the node architecture, the edges in the auxiliary graph for node i are inserted as follows.

1 **Wavelength Bypass Edges (WBE).**
There is an edge from the input port to the output port on each wavelength layer at node i if node i can bypass traffic without using grooming switches:

$$\langle N_i^{l,0},\ N_i^{l,1} \rangle \in E_i, \qquad 1 \le l \le W. \tag{7.1}$$

We call the edge $\langle N_i^{l,0},\ N_i^{l,1}\rangle$ a *wavelength bypass edge* on layer l at node i and it is denoted as $WBE(i,l)$. If a path[2] contains $WBE(i,l)$, that means a lightpath newly set up to carry the connection will bypass node i on λ_l.

2 **Wavelength Converter Edges (WCE).**
There is an edge from the input port on wavelength layer l_1 to the output port on wavelength layer l_2 at node i if wavelength λ_{l_1} can be converted to wavelength λ_{l_2} without using a grooming OXC at node i:

$$\begin{aligned}&\langle N_i^{l_1,0},\ N_i^{l_2,1}\rangle \in E,\\ &\text{wavelength } \lambda_{l_1} \text{ is convertible to } \lambda_{l_2} \text{ at node } i.\end{aligned} \tag{7.2}$$

We call the edge $\langle N_i^{l_1,0},\ N_i^{l_2,1}\rangle$ a *wavelength converter edge* from layer l_1 to layer l_2 at node i and it is denoted as $WCE(i,l_1,l_2)$. For example, there is a wavelength converter edge $WCE(i,1,2)$ in Fig. 7.4, which is from the input port on λ_1 layer to the output port on λ_2 layer, because this node can convert λ_1 to λ_2. There is no wavelength converter edge $WCE(i,2,1)$ in Fig. 7.4 since λ_2 cannot be converted to λ_1 at this node. If a path contains $WCE(i,l_1,l_2)$, that means a lightpath newly set up to carry the connection will be converted from λ_{l_1} to λ_{l_2} at node i without using a grooming OXC.

3 **Grooming Fabric Edges (GFE).**
There is an edge from the input port to the output port on F_j grooming layer at node i if node i has grooming capability at F_j granularity:

$$\langle N_i^{W+2+j,0},\ N_i^{W+2+j,1}\rangle \in E_i, \qquad 1 \le j \le C. \tag{7.3}$$

We call the edge $\langle N_i^{W+2+j,0},\ N_i^{W+2+j,1}\rangle$ a F_j *grooming fabric edge* at node i and it is denoted as $GFE(i,F_j)$. If a path contains $GFE(i,F_j)$, that means the connection will be switched by an OXC with F_j granularity to an outgoing lightpath or the access station at node i.

4 **Wavelength Add Edges (WAE).**
There is an edge from the output port on the transponder layer to the output port on wavelength layer l at node i:

$$\langle N_i^{W+1,1},\ N_i^{l,1}\rangle \in E_i, \qquad 1 \le l \le W. \tag{7.4}$$

We call the edge $\langle N_i^{W+1,1},\ N_i^{l,1}\rangle$ a *wavelength add edge* at node i and it is denoted as $WAE(i,l)$. If a path contains $WAE(i,l)$, that means we need to set up a new lightpath at node i using λ_l to carry the connection.

[2]For conciseness, a path means the path computed for a connection in the auxiliary graph.

5 **Wavelength Drop Edges (WDE).**
There is an edge from the input port on wavelength layer l to the input port on the transponder layer at node i:

$$\langle N_i^{l,0},\ N_i^{W+1,0}\rangle \in E_i \qquad 1 \le l \le W. \tag{7.5}$$

We call the edge $\langle N_i^{l,0},\ N_i^{W+1,0}\rangle$ a *wavelength drop edge* at node i and it is denoted as $WDE(i,\ l)$. If a path contains $WDE(i,\ l)$, that means a lightpath newly set up to carry the connection will be terminated from λ_l at node i.

6 **Grooming-Transponder Edges (GTE).**
There is an edge from the output port on the F_j grooming layer to the output port on the transponder layer at node i if node i has an OXC with an available port which can perform grooming at F_j granularity and is connected to the transponders directly or via an OXC without grooming capability:

$$\langle N_i^{W+2+j,1},\ N_i^{W+1,1}\rangle \in E_i, \qquad 1 \le j \le C. \tag{7.6}$$

We call the edge $\langle N_i^{W+2+j,1},\ N_i^{W+1,1}\rangle$ a *grooming-transponder edge* at node i and it is denoted as $GTE(i,\ F_j)$. If a path contains $GTE(i,\ F_j)$, that means we need to set up a new lightpath originating from an OXC with F_j granularity at node i to carry the connection.

7 **Transponder-Grooming Edges (TGE).**
There is an edge from the input port on the transponder layer to the input port on the F_j grooming layer at node i if node i has an OXC with an available port which can perform grooming at F_j granularity and is connected to the transponders directly or via an OXC without grooming capability:

$$\langle N_i^{W+1,0},\ N_i^{W+2+j,0}\rangle \in E_i, \qquad 1 \le j \le C. \tag{7.7}$$

We call the edge $\langle N_i^{W+1,0},\ N_i^{W+2+j,0}\rangle$ a *transponder-grooming edge* at node i and it is denoted as $TGE(i,\ F_j)$. If a path contains $TGE(i,\ F_j)$, that means a lightpath newly set up to carry the connection will be terminated by an OXC with F_j granularity at node i.

8 **Lightpath-Transponder Edges (LTE).**
There is an edge from the output port on the lightpath layer to the output port on the transponder layer at node i:

$$\langle N_i^{W+2,1},\ N_i^{W+1,1}\rangle \in E_i. \tag{7.8}$$

We call the edge $\langle N_i^{W+2,1},\ N_i^{W+1,1}\rangle$ a *lightpath-transponder edge* at node i and it is denoted as $LTE(i)$. If a path contains $LTE(i,\ F_j)$, that means

we need to set up a new lightpath originating from a non-grooming OXC at node i to carry the connection. This lightpath can only be used to carry the traffic starting from node i.

9 **Transponder-Lightpath Edges (TLE).**
There is an edge from the input port on the transponder layer to the input port on the lightpath layer at node i:

$$\langle N_i^{W+1,0},\ N_i^{W+2,0} \rangle \in E_i. \tag{7.9}$$

We call the edge $\langle N_i^{W+1,0},\ N_i^{W+2,0} \rangle$ a *transponder-lightpath edge* at node i and it is denoted as $TLE(i)$. If a path contains $TLE(i,\ F_j)$, that means a lightpath newly set up to carry the connection will be terminated by a non-grooming OXC at node i. This lightpath can only be used to carry the traffic terminating at node i.

10 **Grooming Add Edges (GAE).**
There is an edge from the input port on the access layer to the input port on the F_j grooming layer at node i if node i can add traffic into an OXC with F_j granularity grooming capability using an unused OXC port:

$$\langle N_i^{W+C+3,0},\ N_i^{W+2+j,0} \rangle \in E_i, \qquad 1 \le j \le C. \tag{7.10}$$

We call the edge $\langle N_i^{W+C+3,0},\ N_i^{W+2+j,0} \rangle$ a *grooming add edge* at node i and it is denoted as $GAE(i,\ F_j)$. If a path contains $GAE(i,\ F_j)$, that means the connection will be added to the network at its source node i using an unused port of an OXC with F_j granularity.

11 **Grooming Drop Edges (GDE).**
There is an edge from the output port on the F_j grooming layer to the output port on the access layer at node i if node i can drop traffic from an OXC with F_j granularity grooming capability using an unused OXC port:

$$\langle N_i^{W+2+j,1},\ N_i^{W+C+3,1} \rangle \in E_i, \qquad 1 \le j \le C. \tag{7.11}$$

We call the edge $\langle N_i^{W+2+j,1},\ N_i^{W+C+3,1} \rangle$ a *grooming drop edge* at node i and it is denoted as $GDE(i,\ F_j)$. If a path contains $GDE(i,\ F_j)$, that means the connection will be dropped from the network at its destination node i using an unused port of an OXC with F_j granularity.

12 **Lightpath Add Edges (LAE).**
There is an edge from the input port on the access layer to the output port on the lightpath layer at node i if node i can add traffic into an OXC with no grooming capability using an unused OXC port:

$$\langle N_i^{W+C+3,0},\ N_i^{W+2,1} \rangle \in E_i. \tag{7.12}$$

We call the edge $\langle N_i^{W+C+3,0}, N_i^{W+2,1} \rangle$ a *lightpath add edge* at node i and it is denoted as $LAE(i)$. If a path contains $LAE(i)$, that means the connection will be added to the network at its source node i using an unused port of a non-grooming OXC.

13 **Lightpath Drop Edges (LDE).**
There is an edge from the input port on the lightpath layer to the output port on the access layer at node i if node i can drop traffic from an OXC with no grooming capability using an unused OXC port:

$$\langle N_i^{W+2,0}, N_i^{W+C+3,1} \rangle \in E_i. \tag{7.13}$$

We call the edge $\langle N_i^{W+2,0}, N_i^{W+C+3,1} \rangle$ a *lightpath drop edge* at node i and it is denoted as $LDE(i)$. If a path contains $LDE(i)$, that means the connection will be dropped from the network at its destination node i using an unused port of a non-grooming OXC.

14 **Grooming Cascade Edges (GCE).**
There is an edge from the output port on the F_j grooming layer to the input port on the F_k grooming layer at node i if traffic can flow from an unused OXC output port with F_j granularity grooming capability to an unused OXC input port with F_k granularity grooming capability at node i and granularities F_k and F_j are different:

$$\langle N_i^{W+2+j,1}, N_i^{W+2+k,0} \rangle \in E_i, \quad 1 \leq j, k \leq C, j \neq k. \tag{7.14}$$

We call the edge $\langle N_i^{W+2+j,1}, N_i^{W+2+k,0} \rangle$ a *grooming cascade edge* at node i and it is denoted as $GCE(i, F_j, F_k)$. If a path contains $GCE(i, F_j, F_k)$, that means the connection will be groomed by an OXC with F_k granularity after having been groomed by an OXC with F_j granularity.

Each edge has a property tuple $PT(c, w)$ associated with it, where c denotes the capacity of this edge and w denotes its weight. The capacity of each edge in the auxiliary graph G_i is assigned ∞. We will discuss weight assignment in Section 7.4.3.

7.3.2 Circuits and Induced Topology in Granularity-Heterogeneous Network

As we know, a lightpath may traverse one or more wavelength-links, and a collection of lightpaths may form a virtual topology in a wavelength-routed network. Traffic requests are carried by these lightpaths. A traffic demand can traverse multiple lightpaths before it reaches its destination node. At intermediate nodes, grooming OXCs are used to switch the traffic from an incoming

lightpath to an outgoing lightpath. If the grooming granularity of the OXC is coarser than the bandwidth granularity of the traffic request, some free timeslots in the incoming lightpath may also be switched together with those timeslots taken by the traffic to the outgoing lightpath, causing these free timeslots to bypass this node and become inaccessible to the node. Some or all of these free timeslots, possibly along with other free timeslots, may be further switched to the next lightpath by the next intermediate node, until they are not switched any longer or reach the destination. We refer to these switched free timeslots by the term *circuit*. In addition, the free timeslots not switched at the OXCs are also called circuits, and these circuits traverse only one lightpath. Therefore, an unused lightpath can also be viewed as a circuit. One of the most important properties of a circuit is that traffic will come out at the end of a circuit *if and only if* it is put into the timeslots at the starting point of the circuit. A circuit can go through multiple lightpaths, and a lightpath can be traversed by multiple circuits. Hence, circuits form another topology, which we call *induced topology*, above the virtual topology. The relationship between circuits and lightpaths is analogous to that between lightpaths and fiber-links.

In granularity-heterogeneous network, a traffic demand cannot be routed according to the virtual topology since lightpath connectivity may not provide accurate information for routing the traffic. For instance, if there are some free timeslots in the lightpath from node 1 to node 2 and there is a traffic request from node 1 to node 2 asking for one timeslot, this lightpath may not be able to deliver the traffic from node 1 to node 2 since the situation may exist where all these free timeslots are switched further to another node 3 and traffic in these timeslots cannot be dropped at node 2. However, circuits provide exact information on whether or not traffic from one node can reach another node. Therefore, traffic demands should be routed on the induced topology instead of on the virtual topology. The relationship between induced topology and virtual topology is analogous to that between virtual topology and physical topology.

For a lightpath, both of its ends are OXCs, which may have different switching granularities. Since a circuit is a fraction of a lightpath or a concatenation of lightpaths in terms of bandwidth, it is also started from an OXC and terminated at another OXC, and the switching granularity of the starting OXC and the terminating OXC may be different. Meanwhile, a circuit itself has its own bandwidth granularity. Therefore, a circuit can be represented by the notation $CT(s, g_s, t, g_t, g_c, m)$, where s and t are starting and terminating nodes of the circuit, respectively; g_s and g_t are the switching granularities of the starting OXC and terminating OXC, respectively; g_c is the bandwidth granularity of the circuit; and m is the amount of bandwidth of the circuit in unit of g_c. Moreover, we extend the concept of circuits by regarding also as circuits the connections between OXCs and their access stations, such as SONET add-drop multiplexer, ATM switches, and IP routers, and between different grooming OXCs inside a

node. So, the end of a circuit can be either an OXC or an access station, and the granularity of the client can be various and is assumed as the finest granularity, STS-1, in the following development below. One key observation is that the switching granularity of the starting OXC g_s and that of the terminating OXC g_t are both no coarser than the bandwidth granularity of the circuit g_c, and the switching granularities of all intermediate nodes traversed by the circuit are all coarser than g_c.

After carrying a traffic request, a circuit may be decomposed into one or more smaller circuits if the bandwidth requirement of the traffic is less than the bandwidth of the circuit. For example, a circuit $CT(i, \text{OC-3}, j, \text{STS-1}, \text{OC-48}, 3)$ will be decomposed into $CT_1(i, \text{OC-3}, j, \text{STS-1}, \text{OC-48}, 2)$, $CT(i, \text{OC-3}, j, \text{STS-1}, \text{OC-12}, 3)$, and $CT_3(i, \text{OC-3}, j, \text{STS-1}, \text{OC-3}, 1)$ after accommodating a traffic request $T(i, j, \text{OC-3}, 3)$. This means that circuits are more dynamic than lightpaths, and they outnumber lightpaths in the network; hence, more intelligence is required to keep track of circuits than lightpaths.

7.3.3 Auxiliary Graph for the Network

Based on the auxiliary graph of each node, we can construct the auxiliary graph for the entire network.

In general, the physical topology of the network can be represented by a graph $G'(V', E')$, where V' and E' are its node set and link set, respectively. Assuming that each link has W wavelengths, λ_1 through λ_W, and there are C possible grooming granularities at nodes in the network, F_1 through F_C, where F_1 is the finest grooming granularity and F_C is the coarsest grooming granularity, we construct the corresponding auxiliary graph $G(V, E)$ for the network G' as follows, where V and E are the vertex set and edge set of graph G, respectively.

The auxiliary graph G is a layered graph with $(W + C + 3)$ layers. First, we construct an auxiliary graph $G_i(V_i, E_i)$ for each node i in the network G' according to the node representation method described above. These nodal auxiliary graphs are components of the network auxiliary graph $G(V, E)$. The vertex set V is the union of the vertex set of the auxiliary graph for each node i in the network G', i.e.,

$$\begin{aligned} V &= \bigcup V_i, \qquad \forall i \in V'. \\ &= \{N_i^{l,p} \mid p \in \{0,1\},\ 1 \le l \le W+C+3,\ \forall i \in V'\}. \end{aligned} \tag{7.15}$$

All the edges in the auxiliary graph of each node i in the network G' are in the edge set E, so

$$E_i \subset E, \qquad \forall i \in V'. \tag{7.16}$$

Then, we insert some additional edges according to the network state, as outlined below.

- **Wavelength-Link Edges (WLE).**
  There is an edge from the output port on wavelength layer l at node i to the input port on wavelength layer l at node j if there is a physical link from node i to node j and wavelength λ_l on this link is not used:

$$\langle N_i^{l,1},\ N_j^{l,0}\rangle \in E, \qquad (i,\ j) \in E';$$
$$\text{wavelength } \lambda_l \text{ on link } (i,\ j) \text{ is not used.} \tag{7.17}$$

  We call the edge $\langle N_i^{l,1},\ N_j^{l,0}\rangle$ a *wavelength-link edge* on layer l from node i to node j and it is denoted as $WLE(i, j, l)$. The capacity of this edge is the capacity of the corresponding wavelength on the link from node i to node j. If a path contains $WLE(i, j, l)$, that means a lightpath traversing link from node i to node j on wavelength λ_l needs to be set up to carry the connection.

- **Circuit Edges (CE).**
  Based on the positions of the ends of the corresponding circuits, circuit edges can be classified as internode circuit edges and intranode circuit edges.

  - **Internode Circuit Edges.**
    There is an edge from the output port on the F_{k_1} grooming layer or input port on the access layer at node i to the input port on the F_{k_2} grooming layer or output port on the access layer at node j if there is a circuit from an OXC with k_1 grooming granularity or a client at node i to an OXC with k_2 grooming granularity or a client at node j:

$$\begin{aligned}
&\langle N_i^{W+2+k_1,1},\ N_j^{W+2+k_2,0}\rangle \in E\\
&\langle N_i^{W+C+3,0},\ N_j^{W+2+k_2,0}\rangle \in E\\
&\langle N_i^{W+2+k_1,1},\ N_j^{W+C+3,1}\rangle \in E\\
&\langle N_i^{W+C+3,0},\ N_j^{W+C+3,1}\rangle \in E,\ 1 \le k_1, k_2 \le C
\end{aligned}$$

$$\exists \text{ a corresponding circuit from node } i \text{ to node } j \text{ in the network.} \tag{7.18}$$

  - **Intranode Circuit Edges.**
    The intranode circuit edges can be parallel with any of the following edges: grooming add edges, grooming drop edges, and grooming cascade edges. We call those edges *grooming-port-associated edges*. The difference between these edges and intranode circuit edges is that these edges exist when there are unused OXC ports in the node, while intranode circuit edges denote the circuits inside a node, which are *induced* by using the OXC port to carry traffic requests.

Circuit edges are denoted as $CE(i, l_i, j, l_j, g_c, m)$, where i and j are nodes in the network, which can be the same; and l_i and l_j are the starting layer and ending layer of the circuit edge in the auxiliary graph, respectively, which can be from $(W + 3)$ to $(W + C + 3)$; g_c is the bandwidth granularity of the corresponding circuit; and m is the amount of bandwidth of the corresponding circuit in unit of g_c. The capacity of a circuit edge is the bandwidth of the corresponding circuit. If a path contains $CE(i, l_i, j, l_j, g_c, m)$, that means the connection will be carried by using the corresponding circuit from node i to node j (if $i \neq j$), or will be added from or dropped to the access station or flow between different types of OXCs within a node (if $i = j$) by sharing a used OXC port with existing traffic.

Although we assume the numbers of wavelengths are the same on all fiber-links in the network, it is straightforward to construct the auxiliary graph for the network where different links may have different number of wavelengths.

Note that granularity-homogeneous network is a special case of granularity-heterogeneous network. The method for constructing the auxiliary graph described above can also be applied to granularity-homogeneous network.

From the construction of the auxiliary graph, it should be clear that the auxiliary graph reflects induced topology and the current state of the network.

7.4 Framework for Network Design Based on the Auxiliary Graph

7.4.1 Algorithm for Routing a Connection Request

To carry a traffic request in a network, several questions need to be answered: (1) Should a new lightpath or lightpaths be set up to accommodate the traffic? Sometimes, it may be better to add a new lightpath or lightpaths even though the connection can be carried on the current virtual topology. (2) How to add the new lightpath(s)? In some cases, we can set up a lightpath directly from the source of the traffic to the destination. In other cases, it is not necessary or possible to set up this lightpath and we may need to set up one or more other lightpaths. (3) How to route the traffic on the virtual topology? As we mentioned before, the virtual topology is not able to provide sufficient information about the connectivity between nodes in granularity-heterogeneous network, and the induced topology should be used to route the traffic.

Based on the auxiliary graph, we propose the Algorithm for Routing a Connection request (ARC) in a given network, which could be granularity-heterogeneous.

The ARC algorithm needs initialization before being used. The initialization takes as a parameter the network configuration, which includes the network topology, as well as the node and the link configurations; and according to the

network configuration, it constructs the corresponding auxiliary graph G using the method discussed in Section 7.3.

The ARC algorithm takes a traffic demand as the input and works as shown in Algorithm 7.1.

Algorithm 7.1 ARC

Input: a traffic demand $T(s, d, g, m)$.

1 Delete the edges whose capacity is less than the bandwidth granularity of T, since they cannot accommodate T.

2 Find the shortest path p from the input port on the access layer of the source to the output port on the access layer of the destination of T on graph G. If not successful, restore the edges previously deleted in Step 1 and return -1.

3 If p contains wavelength-link edges, one or more lightpaths going through the corresponding wavelength-links need to be set up. A lightpath starts whenever p travels through a wavelength add edge, follows the subsequent wavelength-link edges, and terminates at the first wavelength drop edge. Note that a lightpath is also a circuit. If p contains grooming-port-associated edges, the corresponding intranode circuits need to be set up.

4 Route T along the pre-existing circuits in p and/or circuits set up according to p. If the capacity of the path, which is defined as the minimum capacity of the circuits along the path, is less than the entire amount of T, route the maximum amount possible, say n units, of the traffic granularity g.

5 Restore the edges previously deleted in Step 1.

6 Update graph G as follows:

   (a) For each circuit newly set up, a corresponding circuit edge is added.

   (b) The wavelength-link edges denoting the wavelength-links used by the lightpath are removed from the corresponding wavelength layers.

   (c) Starting from the origin of p, for each circuit used by T, remove the corresponding circuit edge. If the circuit is decomposed into one or more circuits, which have different bandwidth granularities, add the corresponding circuit edges into the auxiliary graph. Each of these circuits starts from the same OXC as the original circuits, possibly extends to the following circuits along p, and terminates at the first OXC along p that has a switching granularity not coarser than the bandwidth granularity of this circuit or at the destination of the traffic.

   (d) For each node along p, check all the edges in the auxiliary graph of that node. Remove the edges whose existence conditions are not valid any longer. For instance, all the wavelength converter edges (WCE) in a node auxiliary graph should be removed if there is no wavelength converter available at this node; If there is no free port with a grooming granularity available at a node any more, all the grooming add/drop edge (GAE and GDE), grooming-transponder edge (GTE), transponder-grooming edge (TGE), and grooming cascade edge (GCE) connected to the layer representing that grooming granularity will be removed from the auxiliary graph of that node.

7 If the entire traffic is accommodated, return 0. Otherwise, return $m - n$, which is the amount of the uncarried traffic in units of g.

Complexity Analysis

The complexity of the ARC algorithm is determined by the running time of shortest-path computation in Step 2. Suppose there are N nodes in the network; each node has C OXCs with different grooming granularities; each OXC has P ports; and each link has W wavelengths. In the corresponding auxiliary graph, there are $2N(W + C + 3)$ vertices. As there may be multiple parallel circuit edges between two vertices, we need to count the number of edges explicitly. Within the auxiliary graph of a node, the number of wavelength converter edges (WCE) is $O(W^2)$; the number of grooming cascade edges (GCEs) is $O(C^2)$. Hence, the total number of edges in the auxiliary graph of a node except intranode circuit edges is $O(W^2 + C^2)$. The number of wavelength-link edges (WLE) in the auxiliary graph of the whole network is $O(WN^2)$. Since each node has $P \cdot C$ OXC ports, it can set up $O(PC)$ lightpaths. Each lightpath can be decomposed into C circuits, then the number of circuit edges in the auxiliary graph of the whole network is $O(NPC^2)$. Therefore, the total number of edges in the auxiliary graph of the whole network is $O(W^2 + C^2 + WN^2 + NPC^2)$. Since the running time of shortest-path computation using Dijkstra algorithm is $O(V^2 + E)$, where V and E are the number of the vertices and edges in the graph, respectively, the running time of the ARC algorithm is $O(N^2(W + C)^2 + W^2 + C^2 + WN^2 + NPC^2)$, i.e., $O(N^2(W + C)^2 + NPC^2)$. Usually, the auxiliary graph is not dense, so the first part $O(N^2(W + C)^2)$ is dominant. If each node in the network has full wavelength-conversion capability, for instance only OEO switches are used, all the wavelength layers can be collapsed into one super wavelength layer since all the wavelengths are equivalent. In this special case, the running time of the ARC algorithm is $O(N^2C^2 + NPC^2)$.

7.4.2 An Illustrative Example

To illustrate how the auxiliary graph and the ARC algorithm work, we use the same example mentioned in Section 7.2. In the example network, there is one link between node 1 and node 2, between node 2 and node 3, between node 3 and node 4. Note that there may be other nodes and links in the network also, but for purpose of simplicity, we only focus on these four nodes in this example. Suppose the switching granularities of nodes 1, 2, 3, and 4 are STS-1, OC-3, STS-1, and STS-1, respectively. Each link has two wavelengths (λ_1 and λ_2), and the capacity of a wavelength channel is OC-12 for this illustration. Initially, there is no lightpath in the network and the auxiliary graph of this part of the network is shown in Fig. 7.5.

When a traffic demand $T_1(1, 4, \text{STS-1}, 2)$ arrives, we need to find in the auxiliary graph a path from the input port on the access layer at node 1 to the output port on the access layer at node 4, shown as shaded ports in Fig. 7.6. It is easy to

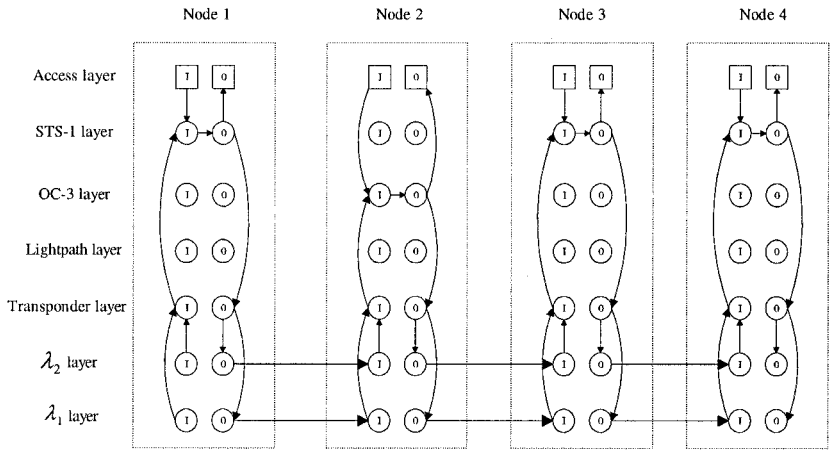


Figure 7.5. Initial auxiliary graph.

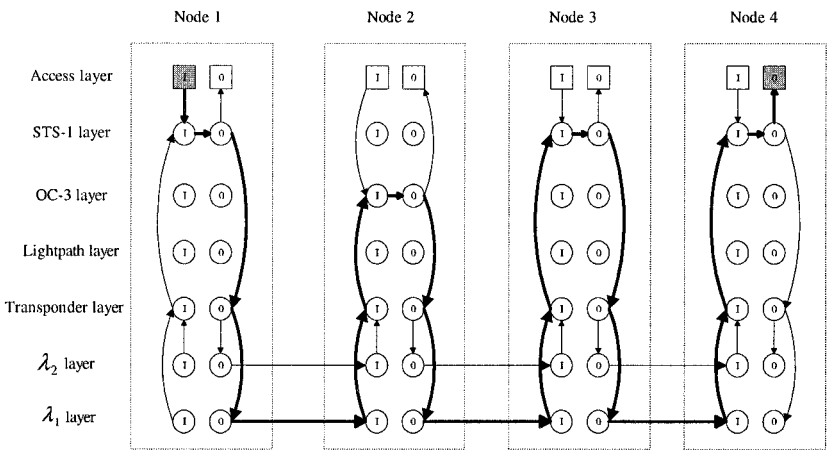


Figure 7.6. Corresponding auxiliary graph before routing the first traffic request T_1.

see that there exists a path p along the edges $GAE(1, \text{STS-1})$, $GFE(1, \text{STS-1})$, $GTE(1, \text{STS-1})$, $WAE(1, 1)$, $WLE(1, 2, 1)$, $WDE(2, 1)$, $TGE(2, \text{OC-3})$, $GFE(2, \text{OC-3})$, $GTE(2, \text{OC-3})$, $WAE(2, 1)$, $WLE(2, 3, 1)$, $WDE(3, 1)$, $TGE(3, \text{STS-1})$, $GFE(3, \text{STS-1})$, $GTE(3, \text{STS-1})$, $WAE(3, 1)$, $WLE(3, 4, 1)$, $WDE(4, 1)$, $TGE(4, \text{STS-1})$, $GFE(4, \text{STS-1})$, and $GDE(4, \text{STS-1})$, shown as bold lines in Fig. 7.6. Since this path contains wavelength-link edge $WLE(1, 2, 1)$, $WLE(2, 3, 1)$, and $WLE(3, 4, 1)$, which denote wavelength-

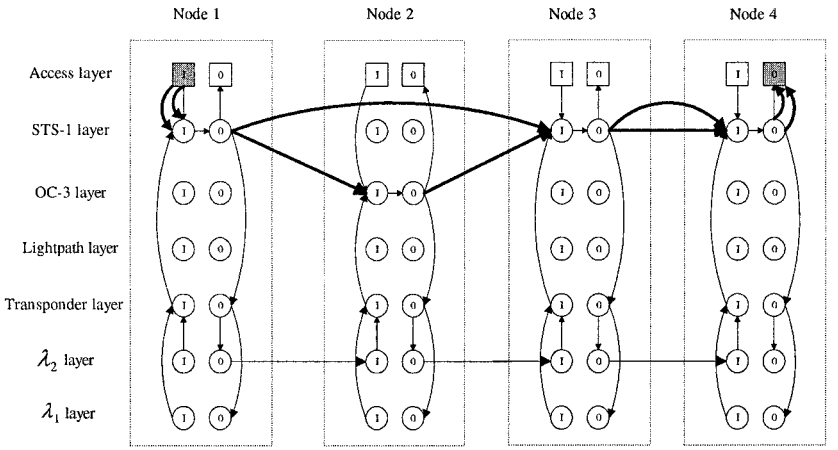


Figure 7.7. Corresponding auxiliary graph after routing the first traffic request T_1.

links, we need to set up three lightpaths (circuits) L_1= CT(1, STS-1, 2, OC-3, OC-12, 1), L_2= CT(2, OC-3, 3, STS-1, OC-12, 1), and L_3= CT(3, STS-1, 4, STS-1, OC-12, 1) using λ_1 on the fiber-link from node 1 to node 2, from node 2 to node 3, and from node 3 to node 4, respectively. Since this path also contains grooming-port-associated edges GAE(1, STS-1) and GDE(4, STS-1), two intranode circuits CT_1= CT(1′, STS-1, 1, STS-1, OC-12, 1) and CT_2= CT(4, STS-1, 4′, STS-1, OC-12, 1) are also set up. Here, i' denotes the client at node i. Then, traffic demand T_1 is routed along these newly set up circuits. The capacity of the path (OC-12) is greater than the bandwidth requirement of T_1 (2 · STS-1), and T_1 is successfully accommodated. After routing T_1, the auxiliary graph needs to be updated. Five circuit edges, i.e., CE_1= CE(1, 7, 1, 6, OC-12, 1), CE_2= CE(1, 6, 2, 5, OC-12, 1), CE_3= CE(2, 5, 3, 6, OC-12, 1), CE_4= CE(3, 6, 4, 6, OC-12, 1), and CE_5= CE(4, 6, 4, 7, OC-12, 1), are added into the auxiliary graph. Three wavelength-link edges, i.e., WLE(1, 2, 1), WLE(2, 3, 1), and WLE(3, 4, 1), are removed since they have been used and are not available any more. After routing T_1, circuit CT_1 is decomposed into two circuits: CT_3= CT(1′, STS-1, 1, STS-1, OC-3, 3) and CT_4= CT(1′, STS-1, 1, STS-1, STS-1, 1); circuit edge CE_1 is replaced by two circuit edges CE(1, 7, 1, 6, OC-3, 3) and CE(1, 7, 1, 6, STS-1, 1). After routing T_1, circuit L_1 is decomposed into two circuits: CT_5= CT(1, STS-1, 2, OC-3, OC-3, 3) and CT_6= CT(1, STS-1, 2, OC-3, STS-1, 1). As the granularity of the OXC at node 2 (OC-3) is coarser than that of CT_6, CT_6 traverses node 2 and extends further to circuit L_2 until it arrives at node 3 where the granularity of the OXC (STS-1) is not coarser than that of CT_6 (STS-1). Now CT_6 becomes CT(1,

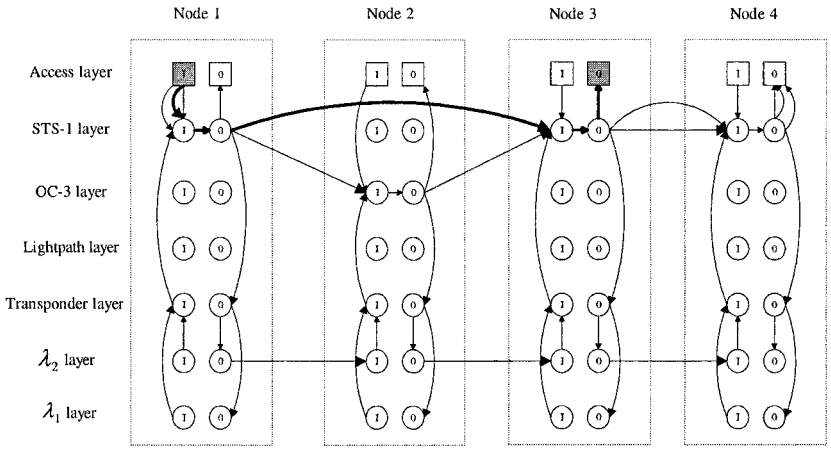


Figure 7.8. Corresponding auxiliary graph before routing the second traffic request T_2.

STS-1, 3, STS-1, STS-1, 1); and circuit edge CE_2 is replaced by two circuit edges CE(1, 6, 2, 5, OC-3, 3) and CE(1, 6, 3, 6, STS-1, 1). After routing T_1, circuit L_2 is decomposed into two circuits: CT_7= CT(2, OC-3, 3, STS-1, OC-3, 3) and CT_8= CT(2, OC-3, 3, STS-1, STS-1, 1). As CT_8 becomes a part of CT_6, circuit edge CE_3 is replaced by one circuit edge CE(2, 5, 3, 6, OC-3, 3). Similarly, circuit edge CE_4 is replaced by two circuit edges CE(3, 6, 4, 6, OC-3, 3) and CE(3, 6, 4, 6, STS-1, 1); circuit edge CE_5 is replaced by two circuit edges CE(4, 6, 4, 7, OC-3, 3) and CE(4, 6, 4, 7, STS-1, 1). To carry traffic demand T_1, each of nodes 1, 2, 3, and 4 uses one unused input port and one unused output port. Figure 7.7 shows the updated auxiliary graph after carrying traffic demand T_1, which represents the network state shown in Fig. 7.2.

Suppose another traffic demand $T_2(1, 3, \text{STS-1}, 1)$ comes. There is a path in the auxiliary graph from the input port on the access layer at node 1 to the output port on the access layer at node 3, shown in Fig. 7.8: CE(1, 7, 1, 6, STS-1, 1), GFE(1, STS-1), CE(1, 6, 3, 6, STS-1, 1) GFE(1, STS-1), and GDE(3, STS-1). Since there is no wavelength-link edge in the path, we do not need to set up any new lightpath; an intranode circuit CT_9= CT(3, STS-1, 3′, STS-1, OC-12, 1) will be set up because the path contains a grooming-port-associated edge GDE(3, STS-1). Hence, a new circuit edge $CE_6 = CE$(3, 6, 3, 7, OC-12, 1) is added into the auxiliary graph. T_2 is carried by existing circuits CT_4= CT(1, STS-1, 1′, STS-1, STS-1, 1), CT_6= CT(1, STS-1, 3, STS-1, OC-1, 1), and newly set up circuit CT_9= CT(3, STS-1, 3′, STS-1, OC-12, 1). After carrying T_2, CT_4 and CT_6 have no free bandwidth left and

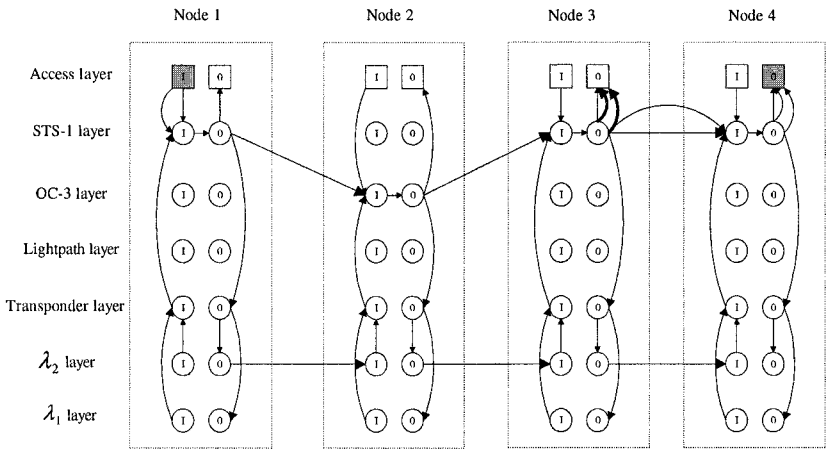


Figure 7.9. Corresponding auxiliary graph after routing the second traffic request T_2.

the corresponding circuit edges $CE(1, 7, 1, 6, \text{STS-1}, 1)$ and $CE(1, 6, 3, 6, \text{STS-1}, 1)$ are removed from the auxiliary graph; CT_9 is decomposed into two circuits: $CT(3, \text{STS-1}, 3', \text{STS-1}, \text{OC-3}, 3)$ and $CT(3, \text{STS-1}, 3', \text{STS-1}, \text{STS-1}, 2)$. Hence, circuit edge CE_6 is replaced by two circuit edges $CE(3, 6, 3, 7, \text{OC-3}, 3)$ and $CE(3, 6, 3, 7, \text{STS-1}, 2)$. To carry T_2, nodes 1 and 2 consume no new OXC port; T_2 shares OXC ports with existing traffic (T_1) at these two nodes. Node 3 uses one unused OXC output port to drop T_2 and does not consume new OXC input port. Figure 7.9 shows the updated auxiliary graph after carrying the traffic demand T_1 and T_2.

7.4.3 Weight Assignment

Since the ARC algorithm applies the shortest-path algorithm to compute the route for a connection request, the route depends on the weight of each edge in the auxiliary graph. In order to choose a proper route for a connection, the weight function should be carefully designed. If we assign the weight to each edge according to the cost of the corresponding component, say the weights of GFE are proportional to the cost of the corresponding grooming fabric, the weight of a WCE is the cost of a converter, the weight of a WLE is proportional to the cost of the corresponding wavelength-links, etc., the ARC algorithm will choose the most cost-effective operation to route a connection, under the current network state. After routing all the traffic using the graph model, we can determine the virtual topology and the induced topology, as well as the configuration of each node, such as the number of OXC ports needed at each node.

In this network design problem, for a given traffic request, how to choose the type of OXCs appropriately at each node along the route is a very important issue if there are multiple different types of OXCs at each node. A fixed weight function, which does not change the weights of the edges, cannot solve the problem. For example, if there is a connection request $T_1(i, j, \text{OC-48}, 1)$ and there are three type of OXCs—OXCs with switching granularities STS-1, OC-48, and OC-192—available at an intermediate node along the path computed for the T_1, it may be more desirable to use the OXC with OC-48 grooming granularity. However, if the bandwidth requirement of T_1 is 2 $\times$ STS-1, the OXC with STS-1 grooming granularity may become the best choice. If the weights of the edges are fixed, the route for the connections will be the same. Therefore, in order to intelligently choose the type of OXCs, the weight of an edge may be different for different traffic requests, i.e., the weight function should be dynamic and traffic requests should be one of the parameter for determining the weight function.

If the port cost of each type of OXC is known, the cost of switching a single timeslot within each type of OXC can be calculated by dividing the port cost with port rate (in terms of the number of timeslots it can support.) To accommodate a connection, the number of timeslots required by the connection have to be switched at an OXC. The cost of switching these timeslots needs to be included in the total cost. Moreover, if the switching granularity of the OXC is coarser than that of the connection, some additional timeslots may also have to be switched. Although these timeslots are not taken by the connection and can be used to carry other connections, they lose some value because they can only carry the connections which go to the same output OXC port as the current connection since the switching fabric is already configured. To discourage this from happening, some penalty may apply in this situation.

Here is an illustrative example. Suppose the port costs of the OXCs with switching granularities STS-1, OC-48, and OC-192 are 5, 4, and 1, respectively, and the port rates are all OC-192. Then, the cost of switching a single timeslot is $5/192$, $1/48$, and $1/192$. To carry a connection $T(i, j, \text{OC-48}, 1)$, node i needs to determine to which OXC it should add the connection. By applying the method described above, we can get three cost values:

$$Cost_{STS-1}(T) = 48 \times 5/192 = 1.25 \tag{7.19}$$

$$Cost_{OC-48}(T) = 48 \times 1/48 = 1 \tag{7.20}$$

$$\begin{aligned} Cost_{OC-192}(T) &= 48 \times 1/192 + \delta \cdot (192 - 48) \times 1/192 \\ &= 0.25 + 0.75\delta \end{aligned} \tag{7.21}$$

When using STS-1 and OC-48 grooming OXCs, there is no penalty because no free timeslot is switched after carrying the connection. However, if an OXC with OC-192 switching granularity is used, 144 (192 − 48) free timeslots are switched, causing some penalty. In Equation (7.21), δ is the *penalty ratio*, which is defined as the ratio of penalty for wasting a timeslot to normal cost for a single timeslot. It is clear that OC-48 grooming OXC is the best choice for this connection. In order to make $Cost_{OC-48}$ the least, δ must be greater than 1. On the other hand, if penalty ratio δ is given, we can determine which OXC has the least cost for the connection.

In general, for a connection $T(s, d, g, m)$, the switching cost of T using an OXC O is:

$$Cost_O(T) = \frac{C_O \cdot g \cdot m}{R_O} + \frac{C_O \cdot \delta}{R_O}\left(\left\lceil \frac{g \cdot m}{g_O} \right\rceil \cdot g_O - g \cdot m\right) \tag{7.22}$$

where C_O is per-port cost of OXC O, R_O is the port rate of OXC O, and g_O is the switching granularity of OXC O. Note that the values of R_O, g_O, and g are in terms of the number of the smallest timeslots the network supports.

It is obvious that the switching cost not only depends on the port cost, but also depends on the bandwidth requirement of the connection, switching granularity of the OXC, and penalty ratio. If we assign the weights of wavelength bypass edges and grooming fabric edges using the corresponding switching cost mentioned above, the weights of these edges will dynamically change accordingly. If the penalty ratio is properly set, the ARC algorithm can intelligently choose a suitable OXC for connections.

When determining the weights of the edges in the auxiliary graph, we can also take into account load balancing. For example, if the load of a specific link is above some threshold, the weight for the wavelength-link edges representing the corresponding wavelength links can be increased to discourage connections from using them. Also, per-port cost can change for each individual OXC as the number of ports used varies. This reflects the facts that the per-port cost may be a function of the size of an OXC because different sizes of OXCs may employ different architectures.

7.4.4 Network Design Framework

The ARC algorithm computes a route for a give traffic demand, based on the auxiliary graph, which has the capability to represent the case where there are multiple different types of OXCs at a single node. By using the dynamic weight assignment for the edges in the auxiliary graph, the ARC algorithm can choose suitable OXCs along the route for a traffic demand. Based on the ARC algorithm, we propose the network design procedure, as shown in Algorithm 7.2.

Algorithm 7.2 Network Design Procedure

1 Assume every node in the network has all types of OXCs.

2 Compute a route for each traffic demand using the ARC algorithm, until all the traffic has been carried.

3 Determine the type of OXC at each node.

(a) Count the number of used ports of each type of OXCs at each node. Let Q_i^j be the number of used ports of OXC with j switching granularity at node i.

(b) For each type of OXC at each node, calculate the number of OXC ports needed if all the traffic going through this node is carried only by this type of OXC. To calculate this number, we use *Port Conversion Ratio* $PCR^{j \to k}$, which is defined as the number of ports of OXC with switching granularity k needed to replace one port of OXC with switching granularity j. For example, to replace one port of OXCs with OC-192 or OC-48 switching granularity, one port of STS-1 grooming OXC is enough as long as they operate at the same port rate. On the other hand, however, more than one OC-48 grooming port may be needed to replace one STS-1 grooming port since an OC-48 grooming port is not flexible as an STS-1 grooming port and may not groom the traffic as an STS-1 grooming port does. We can estimate the value of Port Conversion Ratio $PCR^{j \to k}$ as follows. Assuming that there is only one type of OXCs (OXCs with switching granularity j or k) in the network, route all the traffic demands and count the number of the OXC ports needed in each case. The ratio of these numbers is a good estimation of Port Conversion Ratio. Port Conversion Ratio is a parameter to the design procedure and can be tuned to get better performance.

Let $\bar{Q}_i^k$ be the calculated number of ports of OXC with switching granularity k at node i, then

$$\bar{Q}_i^k = \sum_j Q_i^j \cdot PCR^{j \to k} \tag{7.23}$$

(c) Suppose C_k is the per-port cost of OXC with switching granularity k, then $\bar{Q}_i^k \cdot C_k$ is the port cost at node i using an OXC with switching granularity k. We choose at node i the type of OXC with switching granularity k such that $\bar{Q}_i^k \cdot C_k$ is the least among all the types of OXCs at node i.

4 Reroute all the traffic demands in the determined network configuration and calculate the network-wide OXC port cost.

In Step 2, traffic demands are routed one by one. The order in which the requests are routed will affect the results. There are several traffic-request-selection schemes proposed in Chapter 3, which can be employed here. One of the schemes is Maximum Utilization First (MUF), which selects the connection with the highest utilization. Here, utilization is defined as the total amount of the request divided by the number of hops from the source to the destination on the physical topology. We choose MUF in Step 2 because it has been shown in Chapter 3 that MUF has good performance and scalability.

The running time of network design procedure is dominated by Step 2. Suppose there are D traffic demands, and the running time of the ARC algorithm is

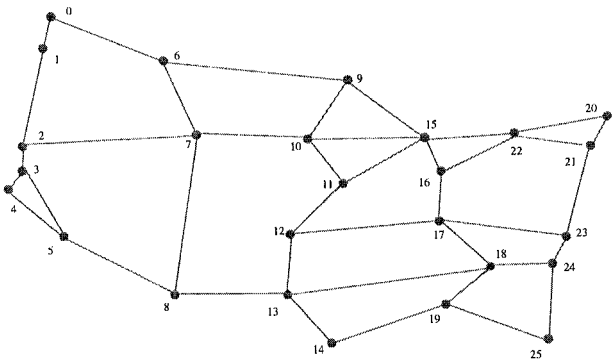


Figure 7.10. A 26-node WDM backbone network.

$O(A)$, then running time of design procedure is $O(D \log D + DA)$. Please see Section 7.4.1 for the detailed information about the running time of the ARC algorithm.

Note that, in Step 2, for each traffic demand $T(s, d, g, m)$, we apply the ARC algorithm. Let x denote the value returned by the ARC algorithm. If $x > 0$, we need to compute another route for $T(s, d, g, x)$ to satisfy the bandwidth requirement. If $x = -1$, it means the traffic cannot be accommodated in the network. In this case, the operator may need to reconfigure the network to add more resources, and then restart the network design procedure.

7.5 Numerical Examples and Discussion

We conducted simulation experiments of the above design principles on a typical nation-wide backbone network. The topology is shown in Fig. 7.10. It has 26 nodes and 40 bi-directional links. Each link has 50 wavelengths and the capacity of a wavelength channel is OC-192. The bandwidth granularity of a traffic demand can be STS-1, OC-3, OC-12, OC-48, and OC-192, and the total traffic bandwidth requirement distribution of these 5 granularities is $\alpha_1 : \alpha_2 : \alpha_3 : \alpha_4 : \alpha_5$, respectively. The traffic is uniformly distributed between all the nodes. There are 3 types of OXCs, whose characteristics are shown in Table 7.1; these OXC types are chosen for this study as representatives of the diverse characteristics in OXC technologies. The per-port cost ratio of Type I, Type II, and Type III OXCs is $\beta_1 : \beta_2 : \beta_3$.

We compare the port cost in four scenarios. In Scenario 1, there is only a Type I OXC at each node; in Scenario 2, only Type II OXCs are deployed in the network; and in Scenario 3, each node only has a Type III OXC. The networks in these three scenarios are granularity-homogeneous. In Scenario 4, each node can only employ one type of OXC, but different nodes may have

Table 7.1. Comparison of three types of OXCs.

OXC	*Switching technology*	*Grooming capability?*	*Capacity of the OXC port*	*Grooming granularity*	*Need transponders for bypass traffic?*
Type I	OOO	No	OC-192	N/A	No
Type II	OEO	Yes	OC-192	OC-48	Yes
Type III	OEO	Yes	OC-192	STS-1	Yes

different types of OXCs, and the network will be granularity-heterogeneous network. We use the network design framework described in 7.4.4 to determine the type of OXC at each node, and all three types of OXCs can co-exist in the network. In the experiments reported here, the ratio of $\alpha_1 : \alpha_2 : \alpha_3 : \alpha_4 : \alpha_5$ is $5 : 1 : 1 : 3 : 3$, which is based on the projected traffic distribution of a typical nation-wide WDM backbone network, and the per-port cost ratio $\beta_1 : \beta_2 : \beta_3$ is $1 : 3 : 4$. Note that these ratios are sample inputs to our network design procedure, and more-accurate data, when available, can be plugged into our model. In addition, the weights of wavelength-link edges and circuit edges representing lightpaths are 10 and 1, respectively; the weight of the circuit edges representing the derived circuits is the summation of the weight of the circuits it derived from; the weight of grooming fabric edges and wavelength bypass edges dynamically change using the method described in Section 7.4.3; the penalty ratio (δ) is 10 and port conversion ratios $PCR^{\text{STS-1}\rightarrow\text{OC-192}}$ and $PCR^{\text{STS-1}\rightarrow\text{OC-48}}$ are 5.3 and 1.6[3], respectively. Port conversion ratio $PCR^{\text{OC-48}\rightarrow\text{OC-192}}$ can be derived as $5.3/1.6 = 3.3$. Note that port conversion ratios $PCR^{\text{OC-192}\rightarrow\text{OC-48}}$, $PCR^{\text{OC-192}\rightarrow\text{STS-1}}$, and $PCR^{\text{OC-48}\rightarrow\text{STS-1}}$ are set to 1 because the port rates of different OXCs are the same and one OXC port with finer granularity can replace only one OXC port with coarser granularity.

Figure 7.11 shows the total port cost, which is normalized by the per-port cost of all-optical OXCs, in the four scenarios; Fig. 7.12 shows the number of transponders used and wavelength-links used in the network; and Fig. 7.13 shows the lightpath utilization in the four scenarios. For the given traffic distribution and port cost ratio, the total port cost of the network in Scenario 1 is the highest, followed by the cost in Scenarios 2 and 3, and Scenario 4 achieves the lowest port cost. In Scenario 1, since the OXCs do not have grooming capability, the lightpath utilization is very low (6.5%) and 3822 OXC ports are used, resulting in highest total port cost despite the lowest per-port cost. In addition, this scenario uses the largest amount of wavelength-links to carry all the traffic.

[3] We experimented with other combinations of these parameters, and found these choices of values to perform the best for this example.

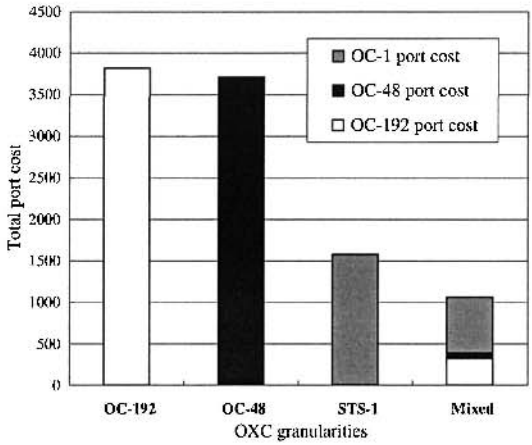


Figure 7.11. Comparison of total port cost in the four scenarios.

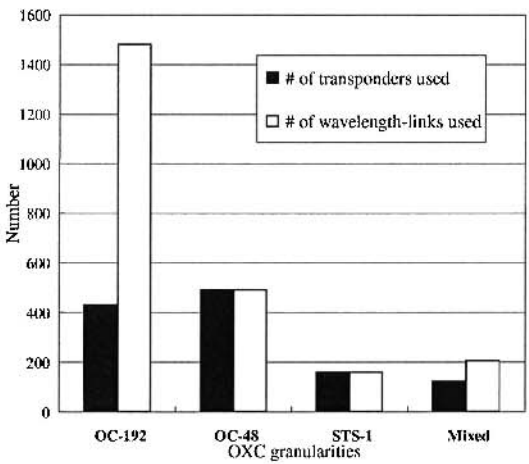


Figure 7.12. Comparison of number of transponders and wavelength-links used in the four scenarios.

In Scenario 3, although the per-port cost of the type of OXCs used is the highest, the total port cost is less than that in Scenarios 1 and 2. This is because Type III OXCs can efficiently pack low-speed connections onto high-speed wavelength channels, making the lightpath utilization relatively high (86%). Hence,

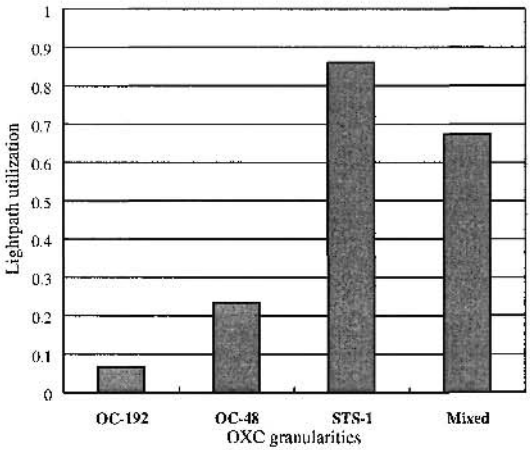


Figure 7.13. Comparison of the lightpath utilization in the four scenarios.

the total number of OXC ports (394[4]), WDM transponders used (160), and wavelength-links used (160) are lower than those in Scenarios 1 and 2.

However, there is still room for improvement. For instance, not all of the nodes need such high flexibility in grooming fabric; some nodes may achieve similar performance with coarser grooming granularity or even no grooming capability, with the coordination of other nodes, thus further reducing the cost. This can be observed in Scenario 4. In this scenario, we choose an appropriate type of OXC for each node. Compared with Scenario 3, although more OXC ports may be used, the total port cost and the number of transponders used in the network are reduced about 33% and 23%, respectively, at the price of using more wavelength-links and lower lightpath utilization. This is because some Type III OXCs at some nodes are replaced with Type I and Type II OXCs, which have lower per-port cost than Type III, and Type I OXCs do not need transponders for bypassing traffic.

7.6 Conclusion

In this chapter, we investigated the problem of designing a WDM backbone network which may consist of OXCs with different switching granularities. Our objective was to minimize the total network-wide OXC port cost. We found that, in a granularity-heterogeneous network, routing traffic on the virtual topology can result in additional connectivities between nodes, which form an induced

[4]Total port cost for Scenario 3 (STS-1 OXC) = 1576; per-port cost (β_3) = 4; so number of ports = 1576/4 = 394.

topology above the virtual topology and which make traffic provisioning much more complex. To accommodate the characteristics of such a network, we enhanced our previous graph model to be able to represent different node architectures. Based on the extended graph model, we proposed a framework for WDM backbone network design to better utilize the benefits of different type of OXCs, which have different bandwidth granularities. Our results demonstrate that using different type of OXCs will yield better network performance, and a design using our framework can reduce the network-wide OXC port cost.

Chapter 8

TRAFFIC GROOMING IN NEXT-GENERATION SONET/SDH

8.1 Virtual Concatenation

Optical SONET/SDH networks are the dominant infrastructure to support data and voice traffic in backbone and metro-area networks. With the maturity of WDM switching technology, we can now deploy a multi-service optical WDM network employing intelligent optical crossconnects (OXCs). But SONET/SDH will still be very important as the framing layer to support efficient and intelligent operation, administration, maintenance and provisioning (OAM&P) functionalities in backbone and metro networks.

As data traffic continues to increase substantially, the inefficiency of transporting variable-length packet through fixed-length SONET frames emerges as a major concern when network operators try to optimize the usage of their bandwidth to support various services, such as IP, frame relay, Ethernet, etc. In traditional SONET/SDH multiplexing hierarchy, frames of multiple low-speed traffic streams (say, STS-1 frame, approx. 51.84 Mbps) are combined to form the frame of a high-speed stream. In order to support high-speed traffic from single client source, e.g., an ATM switch, N "contiguous" lower-order SONET containers are merged into one of greater capacity. This is called SONET/SDH concatenation technique.

Usually, SONET/SDH concatenation is implemented at certain speeds, such as STS-3c, STS-12c, etc, which provide a tiered bandwidth-allocation mechanism for different client services. Unfortunately, although "contiguous" and "tiered" concatenation is simple for implementation, it is not very flexible or efficient, especially in a multi-service network.

From network node perspective, traffic streams from different client network equipment are to be discretely mapped into different tiers of SONET bandwidth trunks (data containers), which may result in huge capacity waste. For example,

carrying a Gigabit Ethernet connection using a concatenated OC-48 pipe (approx. 2.5 Gbps) leads to 60% bandwidth wastage. From a network perspective, time-slot contiguous requirement imposes a constraint for traffic provisioning and may degrade network performance in a dynamic traffic environment where resources are easy to be fragmented. This constraint also makes it more difficult for operators to perform efficient traffic grooming, i.e., packing different low-speed traffic streams onto high-capacity wavelength channels.

8.1.1 SONET Virtual Concatenation

Virtual concatenation (VCAT) helps a SONET/SDH-based network to carry traffic in a finer granularity, and utilize link capacity more efficiently. The basic principle of VCAT [ANSI T1X1.5 2001-062, 2001, ITU G.707, 2002] is that a number of smaller containers, which are not necessarily contiguous, are concatenated, to create a bigger container. Depending on a network's switching granularity, virtual concatenation is possible for small container size from VC-1.5 up to STS-3c.

Figure 8.1 shows an example of how to support multiple services using a single OC-48 channel through VCAT. In Fig. 8.1, an OC-48 channel is used to carry two Gigabit Ethernet, one 200 Mbps Fibre Channel, and two STS-1 TDM voice streams. Through a STS-1 switch, traffic can be switched onto different OC-12 pipes, and these OC-12 pipes can be sent through the network over various routes. Figure 8.1 also illustrates the potential load-balancing benefit brought by VCAT. Also, note that, when traffic from one client is sent over different routes, VCAT mappers at the destination node need to compensate for differential delay between bifurcated streams when they are reconstructed. Currently, a commercially available device may support up to 50 ms (+/-25 ms) differential delay with external RAM, which is equivalent to a 10,000 km difference in route length [Hills, 2002].

In general, a virtually-concatenated SONET channel made up of $N\times$ STS-1 is transported as individual STS-1s across the network; at the receiver, the individual STS-1s are re-aligned and sorted to recreate the original payload. VCAT can be supported at the edge OXC (in port cards) or in a separate traffic-aggregation elements connecting client equipment and the OXC. In this study, we explore, from a network perspective, benefits of SONET virtual concatenation in an optical WDM network under dynamic traffic.

8.1.2 Benefits of Virtual Concatenation: a Network Perspective

From a network perspective, VCAT can benefit optical networks in the following aspects:

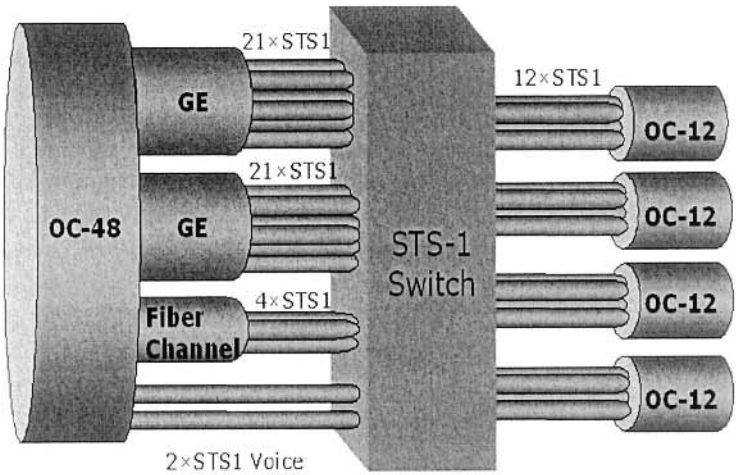


Figure 8.1. An example of using VCAT to support different network services [Stanley, 2002].

- *Relax time-slot alignment and contiguity constraints.* Instead of aligning to particular time-slots and consisting of N contiguous STS-1 time-slots in a wavelength channel, a high-speed STS-N channel can (theoretically) be made up by any N STS-1 time-slots and carried by different wavelength channels.

- *More efficiently utilize channel capacity to support multiple types of data and voice services.* Instead of mapping data traffic (packet/cell/frame) into SONET frames in a discrete tiered manner, optical networks can carry data traffic in a more resource-efficient way. Traffic granularity can be increased in the unit of 1.6 Mbps (VT1.5) in metro-area networks, and 51 Mbps (STS-1) or 150 Mbps (STS-3c) in backbone networks.

- *Bifurcate traffic streams to balance network load.* With VCAT, it is possible to split a high-speed traffic stream into multiple low-speed streams and route them separately through the network. This enables traffic to be distributed in the network more evenly, and hence improves network performance.

Our investigation is based on discrete-event simulations. Connections with different bandwidth granularities come and leave network, one at a time, following a Poisson arrival process and negative exponential-distribution holding time.

Note that, depending on implementation, VCAT mappers at the receiver node may only be able to handle certain number of routes (denoted by t) for a single

Table 8.1. Traffic pattern II used in the study.

Class	*Rate*	*Without VCAT*	*With VCAT*
1	50 Mbps	STS-1	STS-1
2	100 Mbps	STS-3c	STS-2
3	150 Mbps	STS-3c	STS-3
4	200 Mbps	STS-12c	STS-4
5	400 Mbps	STS-12c	STS-8
6	600 Mbps	STS-12c	STS-12
7	1 Gbps	STS-48c	STS-21
8	2.5 Gbps	STS-48c	STS-48
9	5 Gbps	STS-192c	STS-96
10	10 Gbps	STS-192c	STS-192

connection, in which case the connection will be blocked after t routes have been examined, and the connection has not yet been fully provisioned.

8.1.3 Illustrative Numerical Examples

The topology shown in Fig. 5.4 is used in our simulation.

Two traffic patterns are studied (Pattern I and II), one consisting of five service classes and the other has ten service classes. Capacity of each wavelength is OC-192. In Pattern I, data rates for each class are approximately 51 Mbps, 153 Mbps, 622 Mbps, 2.5 Gbps, and 10 Gbps, which can be perfectly mapped into the tiered SONET container, i.e., OC-1, OC-3c, OC-12c, OC-48c, and OC-192c. Table 8.1 shows Pattern II's service classes, service rates, and corresponding SONET containers with or without VCAT.

When traffic bifurcation is needed, a simple route computation and traffic bifurcation heuristic is applied to a connection request. The heuristic works as shown in Algorithm 8.1.

Algorithm 8.1 A simple greedy traffic-bifurcation algorithm

1 A shortest path is computed according to link cost for the request.

2 Bandwidth of the route is calculated. Bandwidth of the route is constrained by the link along the route with minimal free capacity. Then, update the available capacity of the links along the route.

3 Remove the link without free capacity and repeat Steps 1 and 2 until connection can be carried by the set of routes computed or no more routes exist for the connection.

Figures 8.2 and 8.3 illustrate network performance in terms of bandwidth blocking percentage (BBP) as a function of offered load in Erlangs. BBP is used as the performance measurement metric since connections from different classes

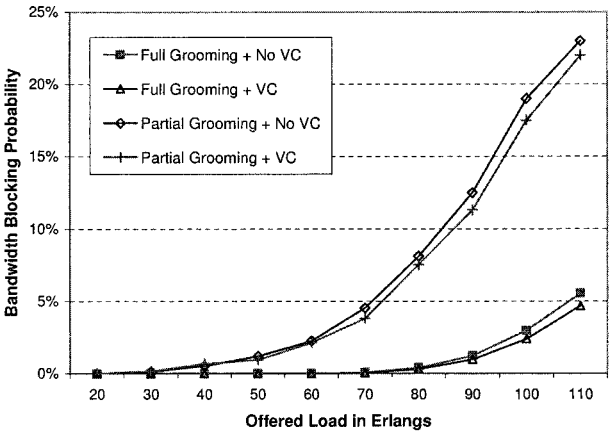


Figure 8.2. Illustrative results - Traffic pattern I.

have different bandwidth requirements. Network offered load is normalized to the unit of OC-192. Each fiber is assumed to support 8 wavelengths.

Figure 8.2 shows the network performance for Pattern I with or without VCAT. Two network configurations are examined, i.e., all nodes are either equipped with STS-1 full-grooming switches or partial-grooming switches. Note that, in a partial-grooming switch, only a limited number of wavelengths (6 in our simulations) can be switched to a separate grooming switch (or grooming fabric within an OXC) to perform traffic grooming. Observe that there is 5-10 percent performance gain by using VCAT. In Pattern I, every class can be perfectly mapped into one of tiered SONET containers, and no bifurcation is assumed in this example. Therefore, performance improvement shown in Fig. 8.2 comes solely from eliminating time-slot alignment and contiguity constraints provided by VCAT. This observation is similar to the effect of wavelength conversion in a wavelength-routed WDM network, where each connection requires full wavelength.

Figure 8.3 shows how VCAT can improve performance when the network needs to support data-oriented services with different bandwidth requirements, assuming full-grooming OXCs everywhere. Now, BBP is significantly reduced by employing VCAT. Meanwhile, more improvement can be achieved by allowing a simple traffic bifurcation scheme. In our study, a connection will be bifurcated only if no single route with enough capacity exists. The results for different values of t have been examined and some are shown in Fig. 8.3, i.e., $t = 4$, 8, and unlimited. It is expected that more advanced load-balancing and traffic-bifurcation approaches can further improve network throughput. Additional results based on different network configurations (e.g., different number

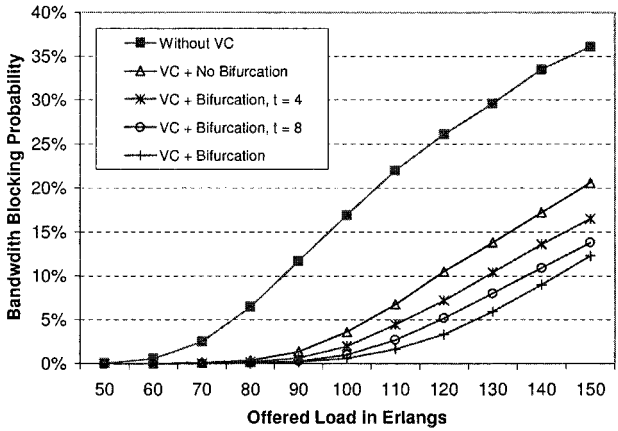


Figure 8.3. Illustrative results - Traffic pattern II.

of wavelengths, different network topologies, different OXC switching capabilities) were also experimented with, and similar results were observed.

Besides the three benefits we have quantitively demonstrated, SONET/SDH VCAT can also benefit an optical network on network compatibility, resiliency, and control and management. VCAT works across legacy networks. Only end nodes of the network are aware of containers being virtually concatenated. For resiliency, since individual members of a virtually-concatenated channel may be carried through different routes, a network failure may only affect partial bandwidth of a connection. Hence, a connection may still get service under reduced bandwidth before failures are fixed (best-effort services) or protection/restoration schemes are activated (priority services). With VCAT, link-state information can be maintained at an aggregated level since time-slot alignment and container-continuity constraints are handled by end nodes. Moreover, built on virtual concatenation, the link-capacity adjustment scheme (LCAS, approved by ITU-T as G.7042) allows network operators to adjust pipe capacity while in use (on the fly). This increases the possibility for on-demand traffic provisioning and on-line traffic grooming/re-grooming and makes SONET/SDH-based optical WDM network more data friendly. However, more intelligent algorithms and mechanisms need to be explored in order to fully utilize the benefits provided by this technique.

8.2 Inverse Multiplexing

The emerging new-generation SONET/SDH techniques, namely virtual concatenation (VCAT) and link-capacity adjustment scheme (LCAS), possess some attractive features which can not only eliminate different limitations of tradi-

tional SONET/SDH networks but also deliver some new benefits. VCAT technology can break the stringent SONET/SDH digital hierarchy by virtually concatenating multiple lower-speed time-slots (VT1.5, STS-1, STS-3c), to form a higher-speed bandwidth pipe (STS-12c, STS-48c, STS-192c). Moreover, with the support of intelligent network control protocols, each independent member of a VCAT group can be supported through a different route between the source and destination node pair of the request. This scheme is also known as the inverse-multiplexing mechanism, which is illustrated in Fig. 8.4.

8.2.1 Problem Statement and Proposed Approaches

The benefits of VCAT technique to SONET/SDH optical networks, such as resource efficiency, service resiliency, flexible bandwidth provisioning, etc., have been widely recognized in both industry and academe [Zhu et al., 2003c, Cavendish et al., 2002, Valencia, 2002]. In order to fully exploit its benefits, the key problem for a network management system is to optimally compute multiple low-capacity paths to provide enough aggregated bandwidth for a high-speed request. Note that, although a standard max-flow algorithm can compute the maximal network capacity between any given node pair in the network in polynomial time, it cannot control the number of distinct paths to be returned. Therefore, it is possible that, in a STS-1 switched SONET/SDH network, a STS-48c request may be carried by 48 distinct (but not necessarily disjoint) paths in the worst case. Using too many distinct paths to support an inverse-multiplexed high-speed connection has some serious disadvantages:

- It may increase the complexity of the network management system and signaling protocol, especially in a distributed network-control environment. The link-state database in the source node of the connection has to keep track of each individual path members for the inverse-multiplexed service. Furthermore, the signaling messages have to be sent along each path to establish or tear down the connection. This may introduce significant management complexity and overhead.

- Combined with LCAS, VCAT-based inverse-multiplexing approach may improve the service resiliency since each network failure may only affect part of the bandwidth of a service. Therefore, instead of losing the entire service, a *degraded service* can still be offered to the user before the failure is fixed or the protection/restoration schemes are activated. However, having too many path members for a connection may also increase the risk of service interruption. It is straightforward to see that more path members means higher service disruption rate when a network failure occurs.

Because of these issues, network operators may want to limit the number of diversely-routed VCAT members for an inverse-multiplexed connection in

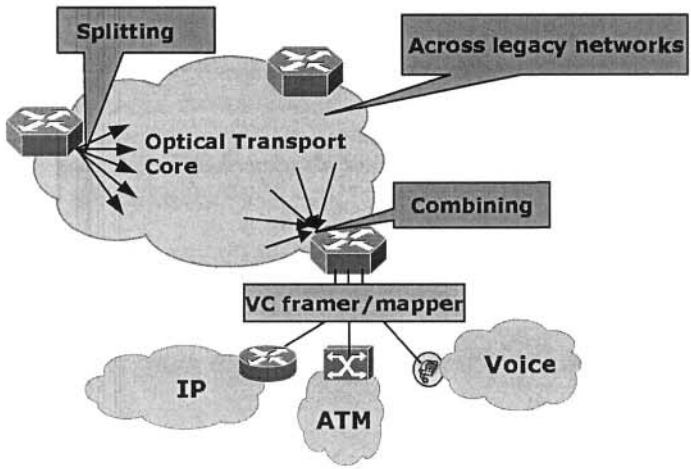


Figure 8.4. An illustrative example of inverse multiplexing in a SONET/SDH-based optical transport network.

their optical networks. This can be achieved by setting a control parameter for maximal allowed path number in the inverse-multiplexing route-computation algorithm of the provisioning system. We formally state the problem as follows:

- Given:
  - An optical network, modeled by a graph $G\ (V,\ E,\ C\ (e,\ w))$, where V is the node set and E is the link set. Each link consist of multiple wavelength channels. $C\ (e,\ w)$ denotes the available capacity on link e and wavelength channel w. The bandwidth of each channel can be shared by different users using SONET/SDH-based TDM scheme.
  - The basic time-slot component of each channel is g. g can be equal to VT1.5, STS-1, etc. g is also known as the basic data-rate granularity supported by the network.
  - Each network node has switching capacity in granularity of g. Traffic can be switched between different wavelength channels and different fiber links. Depending on the implement and node architecture, a network node can be either a full-grooming switch (in the granularity of g) or a partial-grooming switch Chapter 5.
  - A network-wide inverse-multiplexing control parameter K for maximal allowed path number.
  - A high-speed request R with bandwidth requirement B.

Algorithm 8.2 Shortest-Path-First Inverse-multiplexing Algorithm (SPF)

1 A shortest path is computed according to link cost for request R.

2 Bandwidth of the route is calculated. Bandwidth of the route is constrained by the link along the route with minimal free capacity. Then, update the available capacity of the links along the route.

3 Remove the links without free capacity and repeat Steps 1 and 2 until request R can be carried by the set of routes computed or K routes have been examined.

Algorithm 8.3 Widest-Path-First Inverse-multiplexing Algorithm (WPF)

1 Compute the widest path, i.e., the one which has maximum available capacity. Such a widest-path computation algorithm can be developed by modifying Dijkstra's shortest-path algorithm.

2 Update the available capacity of the links along the route.

3 Remove the links without free capacity and repeat Steps 1 and 2 until request R can be carried by the set of routes computed or K routes have been examined.

Algorithm 8.4 Max-Flow Inverse-multiplexing Algorithm (MF)

1 Compute the max-flow from the source node to the destination node of the connection request R using a variation of Ford-Fulkerson's maximum-flow algorithm [Cormen et al., 2001]. Note that, in each iteration of the algorithm, we compute the widest path between the node pair to be the augmenting path. Let $G'(V', E', C'(e, w))$ denote the graph constructed by the computed max-flow, where V' denotes the node set which the max-flow covers, E' denotes the link set which the max-flow traverses, and $C'(e, w))$ denotes the amount of capacity needed on link e and wavelength channel w to support the max-flow.

2 Apply Algorithm 8.3, i.e., WPF, on the graph $G'(V', E', C'(e, w))$.

- Find: k distinct paths, where $1 \leq k \leq K$, such that the aggregated capacity they offer $> B$.

As we have mentioned, a standard max-flow algorithm cannot be directly applied since it is not able to control the number of distinct paths used to carry the request in polynomial time. In fact, we proved that such a problem is NP-complete. Therefore, we develop several heuristic algorithms, which are presented in Algorithms 8.2, 8.3, and 8.4.

8.2.2 Illustrative Numerical Results

To study the performance of the proposed heuristic algorithms, we have simulated a dynamic traffic environment on a 24-node optical WDM mesh network shown in Fig. 5.4. Our investigation is based on discrete-event simulations. Connections with different bandwidth granularities come and leave, one at a time, following a Poisson arrival process and negative-exponentially-distributed

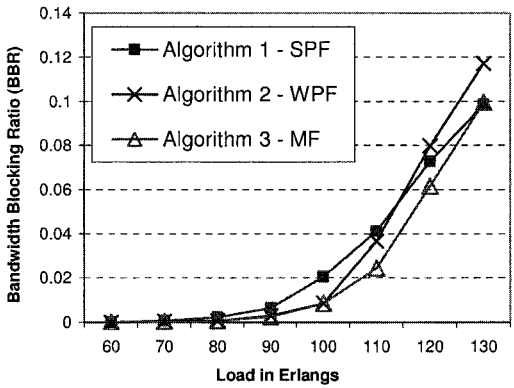


Figure 8.5. Results for $K = 4$ paths.

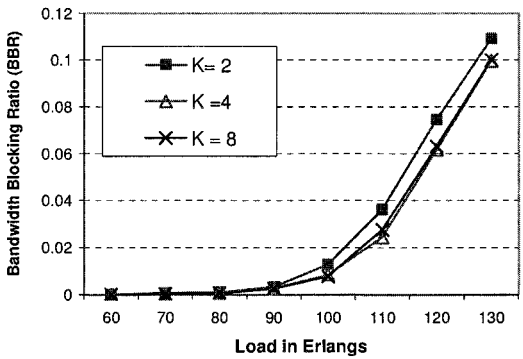


Figure 8.6. Performance results for different values of K based on MF algorithm.

holding time. The network supports ten service classes, with bandwidth requirements 51 Mbps, 100 Mbps, 150 Mbps, 200 Mbps, 400 Mbps, 600 Mbps, 1 Gbps, 2.5 Gbps, 5Gbps, or 10 Gbps. Note that, without VCAT, some service class cannot be perfectly mapped to a SONET/SDH framing time-slot. Capacity of each wavelength is OC-192 and there are 8 wavelength channels on each fiber link. Each network node is assumed to be equipped with a STS-1 full grooming switch.

Figures 8.5 and 8.6 show the performance of three proposed heuristic algorithms based on different values of K. Since requests may have different bandwidth requirements, the network load is normalized to 10 Gbps and shown in Erlangs load on the horizontal axis. The vertical axis shows the bandwidth

blocking ratio (BBR), which is computed as the total amount of blocked bandwidth over the total bandwidth requirements. In general, we can observe from Fig. 8.5 that WPF and MF have better performance than SPF in terms of BBR under reasonable network load. When the network load increases, SPF starts to outperform WPF and has very close performance compared to MF. This is because SPF always tries to use the least amount of resources (in terms of wavelength links) to carry a request. Therefore, when the network load is high enough, it may have a similar or even better performance compared to the other heuristics. We have also observed that, in most cases, MF shows the best network performance. In Fig. 8.6, we show the effects of different values of K using the MF algorithm. We can observe that the network performance can be improved by increasing K from 2 to 4. But increasing K from 4 to 8 provides not much benefit as shown in Fig. 8.6. Higher values of K sometimes may lead to worse network performance, especially under high network load (shown in Fig. 8.6, when load $>$ 110 Erlangs). This is because longer paths tend to be picked when the more path members have to be used to carry a connection. This may lead to higher bandwidth consumption and worse network performance. Please note that more advanced heuristics may achieve better performance with the cost of higher implementation complexity or longer computational time.

8.3 Conclusion

In this chapter, we investigated the traffic-grooming problem in next-generation SONET/SDH networks. Virtual concatenation (VCAT) provides more flexibilities to the optical transport networks. We demonstrated the benefits of VCAT in terms of network performance, and proposed several heuristics for inverse multiplexing. Our results show that Max-Flow performs best among the proposed heuristics in most cases.

References

[Andersson et al., 2001] Andersson, L., Doolan, P., Feldman, N., Fredette, A., and Thomas, B. (2001). LDP specification. *RFC 3036*.

[ANSI T1X1.5 2001-062, 2001] ANSI T1X1.5 2001-062 (2001). Synchronous optical networks (SONET).

[Arora and Subramaniam, 2000] Arora, A. S. and Subramaniam, S. (2000). Converter placement in wavelength routing mesh topoligies. *Proc. IEEE ICC*, pages 1282–1288.

[Awduche et al., 1999] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and McManus, J. (1999). Requirements for traffic engineering over MPLS. *RFC 2702*.

[Awduche et al., 2001] Awduche et al. (2001). RSVP-TE: Extensions to rsvp for LSP tunnels. *RFC 3209*.

[Awduche et al., 2002] Awduche, D., Chiu, A., Elwalid, A., Widjaja, I., and Xiao, X. (2002). Overview and principle of internet traffic engineering. *RFC 3272*.

[B. Jamoussi, et al., 2002] B. Jamoussi, et al. (2002). Constrained-based LSP setup using LDP. *RFC 3212*.

[Banerjee and Mukherjee, 1996] Banerjee, D. and Mukherjee, B. (1996). A practical approach for routing and wavelength assignment in large wavelength-routed optical networks. *IEEE Journal on Selected Areas in Communications*, 14(5):903–908.

[Banerjee and Mukherjee, 1997] Banerjee, D. and Mukherjee, B. (1997). Wavelength-routed design of logical topologies for wavelength-routed optical networks. *Proc. IEEE INFOCOM*, 1:269–276.

[Barr and Patterson, 2001] Barr, R. S. and Patterson, R. A. (2001). Grooming telecommunication networks. *SPIE Optical Networks Magazine*, 2(3):20–23.

[Berry and Modiano, 2000] Berry, R. and Modiano, E. (2000). Reducing electronic multiplexing costs in SONET/WDM rings with dynamically changing traffic. *IEEE Journal on Selected Areas in Communications*, 18(10):1961–1971.

[Braden et al., 1997] Braden, R., Zhang, L., Berson, S., Herzog, S., and Jamin, S. (1997). Resource reservation protocol (RSVP) - version 1 functional specification. *RFC 2205*.

[Brunato and Battiti, 2002] Brunato, M. and Battiti, R. (2002). A multistart randomized greedy algorithm for traffic grooming on mesh logical topologies. In *ONDM 2002 Working Conference*, Torino, Italy.

[Cavendish et al., 2002] Cavendish, D., Murakami, K., Yun, S., Matsuda, O., and Nishihar, M. (2002). New transport services for next-generation SONET/SDH systems. *IEEE Communications Magazine*, 40(5):72–79.

[Chiu and Modiano, 2000] Chiu, L. and Modiano, H. (2000). Traffic grooming algorithms for reducing electronic multiplexing costs in WDM ring networks. *IEEE/OSA Journal of Lightwave Technology*, 18(1):2–12.

[Chlamtac et al., 1992] Chlamtac, I., Ganz, A., and Karmi, G. (1992). Lightpath communications: an approach to high bandwidth optical WAN's. *IEEE Transactions on Communications*, 40(7):1171–1182.

[Chlamtac et al., 1993] Chlamtac, I., Ganz, A., and Karmi, G. (1993). Lightnets: Topologies for high speed optical networks. *IEEE/OSA J. Lightwave Technol.*, 11:951–961.

[Cormen et al., 2001] Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. (2001). *Introduction to Algorithms*. MIT Press.

[Cox and Sanchez, 2001] Cox, L. A. and Sanchez, J. (2001). Cost saving from optimized packing and grooming of optical circuits: mesh versus ring comparisons. *SPIE Optical Networks Magazine*, 2(3):72–90.

[E. Mannie, et al., 2002] E. Mannie, et al. (2002). Generalized multi-protocol label switching (GMPLS) architecture. *Internet Draft, Work in Progress, draft-ietf-ccamp-gmpls-architecture-02.txt.*

[Ganz and Wang, 1994] Ganz, A. and Wang, X. (1994). Efficient algorithm for virtual topology design in multihop lightwave networks. *IEEE/ACM Trans. Networking*, 2(3):217–225.

[Gerstel et al., 1998] Gerstel, O., Lin, P., and Sasaki, G. (1998). Wavelength assignment in a WDM ring to minimize cost of embedded SONET rings. In *Proc., IEEE INFOCOM '98*, pages 94–101, San Francisco, CA.

[Gerstel et al., 1999] Gerstel, O., Lin, P., and Sasaki, G. H. (1999). Combined WDM and SONET network design. In *Proc., IEEE INFOCOM '99*, volume 2, pages 734–743, New York, NY.

[Gerstel et al., 2000] Gerstel, O., Ramaswami, R., and Sasaki, G. H. (2000). Cost-effective traffic grooming in WDM rings. *IEEE/ACM Trans. Networking*, 8(5):618–630.

[Harai et al., 1997] Harai, H., Murata, M., and Miyahara, H. (1997). Performance of alternate routing methods in all-optical switching networks. *Proc. IEEE INFOCOM*, 2:516–524.

[Hills, 2002] Hills, T. (2002). Next-generation SONET. *Lightreading Report.*

[Iness and Mukherjee, 1999] Iness, J. and Mukherjee, B. (1999). Sparse wavelength conversion in wavelength-routed WDM networks. *Photonic Network Communications Journal*, 1(3):183–205.

[ITU G.707, 2002] ITU G.707 (2002). Network node interface for the synchronous digital hierarchy (SDH).

[Jue and Xiao, 2000] Jue, J. P. and Xiao, G. (2000). An adaptive routing algorithm with a distributed control scheme for wavelength-routed optical networks. *Proc. Ninth International Conference on Computer Communications and Networks, (Las Vegas, NV).*

[K. Kompella, et al., 2002] K. Kompella, et al. (2002). Link bundling in mpls traffic engineering. *Internet Draft, Work in Progress, draft-ietf-mpls-bundle-03.txt.*

[Kleinrock, 1970] Kleinrock, L. (1970). Analytic and simulation methods in computer network design. *AFIPS Conference Proceedings*, 42:569–579.

[Konda and Chow, 2001] Konda, V. R. and Chow, T. Y. (2001). Algorithm for traffic grooming in optical networks to minimize the number of transceivers. In *2001 IEEE Workshop on High Performance Switching and Routing*, pages 218–221, Dallas, TX.

[Krishnaswamy and Sivarajan, 2001] Krishnaswamy, R. M. and Sivarajan, K. N. (2001). Design of logical topologies: A linear formulation for wavelength-routed optical networks with no wavelength changers. *IEEE/ACM Trans. Networking*, 9(2):186–198.

[Lardies et al., 2001] Lardies, A., Gupta, R., and Patterson, R. A. (2001). Traffic grooming in a multi-layer network. *SPIE Optical Networks Magazine*, 2(3):91–99.

[Li and Somani, 1999] Li, L. and Somani, A. K. (1999). Dynamic wavelength routing using congestion and neighbor-hood information. *IEEE/ACM Transactions on Networking*, 7:779–786.

[Modiano and Lin, 2001] Modiano, E. and Lin, P. (2001). Traffic grooming in WDM networks. *IEEE Commun. Mag.*, 39(7):124–129.

[Mokhtar and Azizoglu, 1998] Mokhtar, A. and Azizoglu, M. (1998). Adaptive wavelength routing in all-optical networks. *IEEE/ACM Transactions on Networking*, 6:197–206.

[Mukherjee, 1997] Mukherjee, B. (1997). *Optical Communiation Networks*. McGraw-Hill, New York.

[Mukherjee et al., 1996] Mukherjee, B., Ramamurthy, S., Banerjee, D., and Mukherjee, A. (1996). Some principles for designing a wide-area optical network. *IEEE/ACM Trans. Networking*, 4(5):684–696.

[Ou et al., 2003] Ou, C., Zhu, K., Zang, H., Sahasrabuddhe, L., and Mukherjee, B. (2003). Traffic groomig for survivable WDM networks — shared protection. *IEEE J. Select. Areas Commun.*, 21(9):1367–1383.

[Ramamurthy and Mukherjee, 2002] Ramamurthy, S. and Mukherjee, B. (2002). Fixed-alternate routing and wavelength conversion in wavelength-routed optical networks. *IEEE/ACM Transactions on Networking*, 10(3):351–367.

[Ramamurthy et al., 2003] Ramamurthy, S., Sahasrabuddhe, L., and Mukherjee, B. (2003). Survivable WDM mesh networks. *IEEE/OSA J. Lightwave Technol.*, 21(4):870–883.

[Ramaswami and Sasaki, 1998] Ramaswami, R. and Sasaki, G. (1998). Multiwavelength optical networks with limited wavelength conversion. *IEEE/ACM Trans. Networking*, 6(6):744–754.

[Ramaswami and Sivarajan, 1996] Ramaswami, R. and Sivarajan, K. N. (1996). Design of logical topologies for wavelength-routed optical networks. *IEEE J. Select. Areas Commun.*, 14(5):840–851.

[Ramaswami and Sivarajan, 1998] Ramaswami, R. and Sivarajan, K. N. (1998). *Optical Networks: A Practical Perspective*. San Francisco: Morgan Kaufmann Publishers.

[Ranganathan et al., 2002] Ranganathan, R., Blair, L., and Berthold, J. (2002). Architectural implications of core grooming in a 46-node USA optical network. In *Proc., OFC '02*, pages 498–499, Anaheim, CA.

[Sahasrabuddhe and Mukherjee, 1999] Sahasrabuddhe, L. H. and Mukherjee, B. (1999). Lightrees: optical multicasting for improved performance in wavelength-routed networks. *IEEE Communications Magazine*, 37(2):67–73.

[Simmons et al., 1999] Simmons, J., Goldstein, E., and Saleh, A. (1999). Quantifying the benefit of wavelength add-drop in WDM rings with independent and dependent traffic. *IEEE/OSA Journal of Lightwave Technology*, 17(1):48–57.

[Singhal and Mukherjee, 2001] Singhal, N. and Mukherjee, B. (2001). Architectures and algorithms for multicasting in WDM optical mesh networks using opaque and transparent optical crossconnects. In *Proc. OFC 2001*, volume 2, pages TuG8-1–TuG8-3.

[Stanley, 2002] Stanley, S. (2002). Next-gen SONET silicon. *Lightreading Report*.

[Subramaniam et al., 1996] Subramaniam, S., Azizoglu, M., and Somani, A. K. (1996). All-optical networks with sparse wavelength conversion. *IEEE/ACM Trans. Networking*, 4(4):544–557.

[Thiagarajan and Somani, 2001a] Thiagarajan, S. and Somani, A. K. (2001a). A capacity correlation model for WDM networks with constrained grooming capabilities. In *Proc., IEEE ICC '01*, volume 5, pages 1592–1596, Helsinki, Finland.

[Thiagarajan and Somani, 2001b] Thiagarajan, S. and Somani, A. K. (2001b). Capacity fairness of WDM networks with grooming capabilities. *SPIE Optical Networks Mag.*, 2(3):24–31.

[Tripathi and Sivarajan, 2000] Tripathi, T. and Sivarajan, K. N. (2000). Computing approximate blocking probabilities in wavelength routed all-optical networks with limited-range wavelength conversion. *IEEE J. Select. Areas Commun.*, 18(10):2123–2129.

[Valencia, 2002] Valencia, E. H. (2002). Hybrid transport solutions for TDM/data networking services. *IEEE Communications Magazine*, 40(5):104–112.

[Venugopal et al., 1999] Venugopal, K. R., Shivakumar, M., and Kumar, P. S. (1999). A heuristic for placement of limited range wavelength converters in all-optical networks. In *Proc., IEEE INFOCOM '99*, volume 2, pages 908–915, New York, NY.

[Wan et al., 2000] Wan, P. J., Calinescu, G., Liu, L., and Frieder, O. (2000). Grooming of arbitrary traffic in SONET/WDM BLSRs. *IEEE J. Select. Areas Commun.*, 18(10):1995–2003.

[Wang et al., 2001] Wang, J., Cho, W., Vemuri, V. R., and Mukherjee, B. (2001). Improved approaches for cost-effective traffic grooming in WDM ring networks: ILP formulations

and single-hop and multihop connections. *IEEE/OSA Journal of Lightwave Technology*, 19(11):1645–1653.

[Wang and Mukherjee, 2002] Wang, J. and Mukherjee, B. (2002). Interconnected WDM ring networks: Strategies for interconnection and traffic grooming. *SPIE Optical Networks Mag.*, 3(5):10–20.

[Xin et al., 2002] Xin, C., Ye, Y., Dixit, S., and Qiao, C. (2002). An integrated lightpath provisioning approach in mesh optical networks. *Proc. OFC*, pages 547–549.

[Xu et al., 2000] Xu et al. (2000). A framework for generalized multi-protocol label switching (GMPLS). *http://search.ietf.org/internet-drafts/draft-many-ccamp-gmpls-framework-00.txt.*

[Yates et al., 1996] Yates, J., Lacey, J., Everitt, D., and Summerfield, M. (1996). Limited-range wavelength translation in all-optical networks. In *Proc., IEEE INFOCOM '96*, volume 3, pages 954–961, San Francisco, CA.

[Zang et al., 2000] Zang, H., Jue, J. P., and Mukherjee, B. (2000). A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks. *SPIE Optical Networks Mag.*, 1(1):47–60.

[Zhang and Qiao, 1998] Zhang, X. and Qiao, C. (1998). Wavelength assignment for dynamic traffic in multi-fiber WDM networks. *Proc., 7th International Conference on Computer Communications and Networks*, pages 479–485.

[Zhang and Qiao, 2000] Zhang, X. and Qiao, C. (2000). An effective and comprehensive approach for traffic grooming and wavelength assignment in SONET/WDM rings. *IEEE/ACM Trans. Networking*, 8(5):608–617.

[Zhu et al., 2002] Zhu, H., Zang, H., Zhu, K., and Mukherjee, B. (2002). Dynamic traffic grooming in WDM mesh networks using a novel graph model. *Proc. IEEE Globecom*, 3:2696–2700.

[Zhu et al., 2003a] Zhu, H., Zang, H., Zhu, K., and Mukherjee, B. (2003a). A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks. *IEEE/ACM Trans. Networking*, 11(2):285–299.

[Zhu et al., 2003b] Zhu, H., Zhu, K., Zang, H., and Mukherjee, B. (2003b). Cost-effective WDM backbone network design with OXCs of different bandwidth granularities. *IEEE Journal on Selected Areas in Communications*, 21(9):1452–1466.

[Zhu and Mukherjee, 2002a] Zhu, K. and Mukherjee, B. (2002a). On-line approaches for provisioning connections of different bandwidth granularities in WDM mesh networks. In *Proc., OFC '02*, pages 549–551, Anaheim, CA.

[Zhu and Mukherjee, 2002b] Zhu, K. and Mukherjee, B. (2002b). Traffic grooming in an optical WDM mesh network. *IEEE J. Select. Areas Commun.*, 20(1):122–133.

[Zhu et al., 2003c] Zhu, K., Zang, H., and Mukherjee, B. (2003c). Exploiting the benefit of virtual concatenation technique to the optical transport networks. *Proc. OFC 2003.*

Index

ADM, 3
    O-ADM, 3
    SONET ADM, 3
    W-ADM, 3
ARC (Algorithm for Routing a Connection), 140
auxiliary graph, 46
    access layer, 47
    Converter Edges (CvtE), 49
    Demux Edges (DmxE), 48
    Grooming Edges (GrmE), 48
    Lightpath Edges (LPE), 49
    lightpath layer, 47
    Mux Edges (MuxE), 48
    property tuple, 50
    Receiver Edges (RxE), 48
    Transmitter Edges (TxE), 48
    Wavelength Bypass Edges (WBE), 47
    wavelength layer, 47
    Wavelength-Link Edges (WLE), 49

bandwidth blocking ratio (BBR), 106
bipartite graph, 7

connection admission control (CAC), 11
Connection Blocking Probability (CBP), 88
CPLEX, 29

dominant edge, 60
DXC, 5
dynamic grooming
    SONET/WDM ring, 6

Embedded on Physical Topology (EPT), 109
enhanced graph model, 130–140
    access layer, 132
    Circuit Edges (CE), 139
        Internode Circuit Edges, 139
        Intranode Circuit Edges, 139
    Grooming Add Edges (GAE), 135
    Grooming Cascade Edges (GCE), 136
    Grooming Drop Edges (GDE), 135
    Grooming Fabric Edges (GFE), 133
    grooming layer, 131
    Grooming-Transponder Edges (GTE), 134
    Lightpath Add Edges (LAE), 135
    Lightpath Drop Edges (LDE), 136
    lightpath layer, 131
    Lightpath-Transponder Edges (LTE), 134
    transponder layer, 131
    Transponder-Grooming Edges (TGE), 134
    Transponder-Lightpath Edges (TLE), 135
    Wavelength Add Edges (WAE), 133
    Wavelength Bypass Edges (WBE), 132
    Wavelength Converter Edges (WCE), 133
    Wavelength Drop Edges (WDE), 134
    wavelength layer, 131
    Wavelength-Link Edges (WLE), 139
extended graph model, 80

FDM, 1

G-Fabric, 10, 95
GMPLS, 16
granularity-heterogeneous network, 126
granularity-homogeneous network, 126
graph model, 45
grooming policy, 45, 57–58, 79
    MinLP, 58, 79
    MinTH, 58
    MinTHP, 79
    MinTHV, 79
    MinWL, 58, 79
grooming ratio, 3
grooming-node-selection scheme
    bypass-traffic selection, 120
    nodal-degree selection, 120
    random selection, 120
GUAG (Grooming Using Auxiliary Graph), 102
GULT (Grooming Using Light-Tree), 103

heterogeneity, 44, 72

IGABAG, 51
ILP, 4, 22–29, 117–119
induced connectivity, 130
induced topology, 137
INGPROC, 51–54
interconnected ring, 8
inverse multiplexing, 160
inverse-multiplexing algorithm
    Max-Flow (MF), 163
    Shortest-Path-First (SPF), 163
    Widest-Path-First (WPF), 163

light-tree, 15, 99
lightpath, 3
    destination-groomable, 77, 96
    full-groomable, 77, 96
    multi-hop un-groomable, 76, 95
    source-groomable, 76, 95
link
    fiber, 77
    virtual, 77
link-capacity adjustment scheme (LCAS), 160

MPLS, 21
MRU, 35
MSPP, 22
MST, 35
multi-hop grooming, 56
    SONET/WDM ring, 5
multicast, 15

NC&M, 20
network design framework, 148
network revenue, 40
NNI, 20
node architecture
    SONET/WDM ring, 2
NP-Complete, 4, 29

O-E-O conversion, 8
OXC, 8, 73
    multi-hop full-grooming, 74, 98
    multi-hop partial-grooming, 74, 95–97
    non-grooming, 73
    single-hop grooming, 73, 94–95
    source-node grooming, 99

PDM, 2
penalty ratio, 148
physical topology, 19
Port Conversion Ratio (PCR), 149
PPWDM ring, 6

quantity discount, 39

Resource Efficiency Ratio (RER), 88, 110
routing, 33
    adaptive, 34
    fixed, 33
    fixed-alternate, 34
RWA, 17

SDM, 1
single-hop grooming, 55
    SONET/WDM ring, 4
SONET, 2
SONET/WDM, 3
sparse grooming, 115

t-allowable, 7
TDM, 2
time-space-time (TST) switching, 98
Traffic Blocking Ratio (TBR), 88
traffic engineering, 16, 71
traffic grooming, 2
    integrated approach, 45
traffic request selection
    Least Cost First (LCF), 53
    Maximum Amount First (MAF), 53
    Maximum Utilization First (MUF), 53

UNI, 20

virtual concatenation (VCAT), 156
virtual topology, 19
virtual-topology design, 61

W-Fabric, 95
Wavelength Assignment, 34
    First-Fit, 34
wavelength conversion, 44
    full, 44
    partial, 44
    sparse, 44
WDM, 1
weight assignment, 58, 146
WGXC, 10
WRS, 19