

On the Vision of Provisioning Multicast and Groupcast Traffic
in Next-Generation WDM-Based Optical Networks

by

Tanvir Rahman

A dissertation submitted to the Graduate Faculty in Engineering in partial
fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2009

©2009

TANVIR RAHMAN

All Rights Reserved

This manuscript has been read and accepted by the Graduate Faculty in Engineering in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date

Professor Mohamed A. Ali, Chair of Examining Committee

Date

Professor Muntaz K. Kassir, Executive Officer

Dr. Georgios Ellinas

[Adviser]

Professor Roger Dorsinville

[Supervision Committee]

Professor Neophytos Antoniadis

[Supervision Committee]

Professor Nicholas Madamopoulos

[Supervision Committee]

Abstract

ON THE VISION OF PROVISIONING MULTICAST AND GROUPCAST
TRAFFIC IN NEXT-GENERATION WDM-BASED OPTICAL NETWORKS

by

Tanvir Rahman

Adviser: Professor Georgios Ellinas

Wavelength-Division Multiplexed (WDM)-based optical transport networks, utilizing intelligent switches have emerged as the preferred architecture to accommodate the tremendous increase in traffic in today's networks. These networks need to provision high bandwidth sessions as well as a mixture of different types of sessions including unicast, multicast and groupcast.

Applications such as video on demand, multiparty conferencing, teaching via the web, grid computing, and other potential Internet applications will be the driving force behind the increase of multicast and groupcast traffic in the near future. For such applications to be viable, it is important that the mixture of unicast, multicast and groupcast traffic is routed efficiently through the optical network. To accommodate these trends, optical architectures with the capability to provision different types of services utilizing a number of intelligent optical cross-connect architectures have appeared in the forefront of optical networking technology.

The objective of this dissertation is to investigate problems associated with the provisioning of protected multicast sessions, as well as the network planning process and the provisioning of groupcast sessions. The primary contribution of this dissertation is a novel network protection technique for multicasting in optical networks, the proposal of planning processes for groupcast provisioning and novel heuristics

for groupcast routing and wavelength assignment. Both static and dynamic traffic cases were considered and evaluated. Using optical cross-connects that supports only lightpath establishment and exist commercially today, the thesis proposes novel approaches on how to establish groupcast connections (light-forests). Moreover, using drop-and-continue architectures, “linear light-trees” are constructed that also serve groupcast sessions. Assuming maturing multicast node architectures, this work proposes provisioning and protection solutions for multicast sessions as well as provisioning of groupcast sessions.

This dissertation fills an existing void in that area by developing a solution for provisioning mixed services for the general case of networks with arbitrary mesh topologies. It formulates models and algorithms that can be implemented by the network designers during the set-up of the network, to guarantee the operation of the network for different mixtures of applications.

ACKNOWLEDGMENTS

I like to express sincere appreciation to Professor Ellinas for his assistance in the preparation of this manuscript. Special thanks to Professor Mohamed Ali whose guidance and support during the doctoral program was a very enriching experience.

Many thanks go to my family: my wife Zineb, my daughter Boushra, my son Elias for being accommodating of my after hours studies and also for sacrificing their countless weekends. Also, I am grateful to my father, Mokbul Rahman, and my sister, Tushi, for their constant support during my unusual never-ending studies.

Finally, I am truly indebted to my mother, Rina Roy, for her dedication and determination in providing me with the best possible education. As her lofty ambitions for me has culminated to the completion of the doctoral degree, I would like to dedicate this thesis to my mother.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | Optical Multicast and Groupcast | 1 |
| 1.2 | Routing and Wavelength Assignment for Optical Groupcast | 6 |
| 1.3 | Thesis Objectives | 8 |
| 1.4 | Thesis Outline | 10 |
| 2 | Provisioning and Protection of Multicasts | 12 |
| 2.1 | Optical Multicast Node Architectures | 12 |
| 2.2 | Provisioning Multicast Sessions | 18 |
| 2.3 | Multicast Routing and Wavelength Assignment (MC-RWA) | 26 |
| 2.3.1 | Performance Evaluations | 28 |
| 2.3.2 | Provisioning Integrated Unicast and Multicast Traffic | 31 |
| 2.4 | Optical Multicast Protection | 33 |
| 2.4.1 | Multicast Protection Algorithms | 34 |
| 2.4.2 | Simulation Results | 37 |

| | | |
|----------|---|-----------|
| 3 | Lightpath-Based Optical Groupcast | 42 |
| 3.1 | Optical Cross-Connect Switch Architectures | 42 |
| 3.1.1 | Node Architectures Supporting Groupcast Sessions | 45 |
| 3.2 | Groupcast Routing and Wavelength Assignment (GC-RWA) | 47 |
| 3.2.1 | Groupcast Routing | 48 |
| 3.2.2 | Wavelength-Assignment | 49 |
| 3.2.3 | Example I: GC-RWA on OXC Network Supporting No Wave- length Conversion | 51 |
| 3.2.4 | Example II: GC-RWA on OXC Networks Supporting Wave- length Conversion | 54 |
| 3.2.5 | Optical Network Architecture | 56 |
| 3.2.6 | GC-RWA Performance | 61 |
| 3.3 | Static Groupcasts: Planning and Design of Sessions in OXC Networks | 66 |
| 3.3.1 | Lightpaths: Virtual Topology Design for Groupcast | 69 |
| 3.3.2 | VTD Formulation-I: OXC with Wavelength Converters | 70 |
| 3.3.3 | VTD Formulation-II: No Wavelength Conversion | 74 |
| 3.3.4 | The Network Design Problem: A Heuristic Approach | 77 |
| 3.3.5 | Heuristics and Illustrative Results | 80 |
| 3.4 | Dynamic Groupcasts: Performance Analysis in OXC Networks | 90 |
| 3.4.1 | GC-RWA performance in wavelength-continuous OXC-based net- work | 91 |

| | | |
|----------|--|-----------|
| 3.4.2 | GC-RWA Performance in OXC-Based Networks With Wavelength-Conversion Capabilities | 95 |
| 4 | Light-tree-Based Optical Groupcast | 98 |
| 4.1 | Drop-and-Continue (DaC) Node Architectures | 100 |
| 4.2 | Building Light-Forests from Light-Trees | 102 |
| 4.2.1 | Linear Light-trees | 103 |
| 4.2.2 | Linear Light-tree Performance Results | 105 |
| 4.3 | Static Groupcasts: Planning and Design of Sessions in MC-OXC Network | 108 |
| 4.3.1 | Light-Trees: Groupcast VTD Problem | 110 |
| 4.3.2 | Light-tree VTD I: Wavelength Conversion Case | 111 |
| 4.3.3 | Light-tree VTD II: Wavelength-Continuity Case | 115 |
| 4.3.4 | Network Design Problem utilizing MC-OXCs: Heuristic Approach | 119 |
| 4.4 | Static Groupcasts: Planning and Design of Sessions in DaC Networks | 122 |
| 4.4.1 | Data Plane: Groupcast Session Size | 122 |
| 4.5 | Dynamic Groupcast: Performance Analysis | 126 |
| 4.5.1 | Light-forest with mixed loads | 126 |
| 4.5.2 | Linear light-tree forest | 127 |
| 4.5.3 | Fairness Issue | 128 |
| 4.5.4 | Influence of Network Resources (Wavelengths) on Light-forest Performance | 129 |

5 Conclusions and Future Directions 132

5.1 Summary of Research 132

5.2 Future Directions 134

Bibliography 137

List of Figures

| | | |
|------|---|----|
| 1.1 | The optical communications evolution | 2 |
| 1.2 | MEF multipoint-to-multipoint service | 5 |
| 2.1 | Optical multicasting with multicast-incapable (MI) and multicast-capable (MC) nodes | 13 |
| 2.2 | Multicast node with M fibers and N wavelengths per fiber | 14 |
| 2.3 | (A) $N \times N$ SAD switch (B) SAD-based MC-OXC | 15 |
| 2.4 | All-optical Cross-connect Switch Architecture to support Multicast | 15 |
| 2.5 | Generic structure of $M^\lambda(N \times N)$ WDM MC-OXC architecture | 16 |
| 2.6 | A Banyan-style Multistage Interconnected Architecture (MIA) using SAD Switches | 17 |
| 2.7 | Multicasting using MOSPF | 23 |
| 2.8 | Overlay Model for Optical Network | 24 |
| 2.9 | Multicast trees from MCH, LTA and MCLT Algorithms on NSF Network | 28 |
| 2.10 | Blocking Performance vs. Multicast Group Size in NSF network | 29 |

| | |
|--|----|
| 2.11 Blocking Performance vs. Multicast Group Size in NSF and AT&T networks | 30 |
| 2.12 Average Multicast Session Cost vs. Multicast Group Size in NSF14 network | 31 |
| 2.13 Blocking Performance vs. Multicast Group Size in NSF network | 32 |
| 2.14 Example illustrating MCH and MC-CR algorithms | 36 |
| 2.15 Example illustrating CR2 algorithm | 39 |
| 2.16 Blocking performance vs. network load | 40 |
| 2.17 Blocking performance vs. multicast group size | 40 |
| 2.18 1+1 Multicast Protection Heuristic Performance: Blocking vs. Network Load | 41 |
| 3.1 OXC architectures supporting (A) Opaque and (B) Transparent switching | 43 |
| 3.2 Transparent optical cross-connect architecture | 45 |
| 3.3 Optical cross-connect architecture supporting groupcast | 46 |
| 3.4 NSF 14-node network | 49 |
| 3.5 A Network with 6 nodes and 9 links | 52 |
| 3.6 IP-over-DWDM network and IP-over-OTN network | 57 |
| 3.7 NSF 14 node network | 59 |
| 3.8 AT&T 24 node IP network | 60 |
| 3.9 Groupcast vs. multicast performance in wavelength-routed networks | 64 |

| | |
|--|----|
| 3.10 AT&T 24 node IP network | 64 |
| 3.11 Groupcast performance in a 14-node vs. a 24-node network | 66 |
| 3.12 Wavelengths required for groupcast in NSF14 vs. AT&T network | 67 |
| 3.13 Groupcast routing performance for minimum-cost-path vs. shortest-path-first in NSF14 | 68 |
| 3.14 Wavelengths per fiber required in NSF14 for minimum-cost-path and shortest-path-first algorithms | 69 |
| 3.15 Session size requirement (10% of total combinationd) for NSF14 | 78 |
| 3.16 Groupcast sessions in OXC-based network (NSF14) with and without wavelength conversion | 82 |
| 3.17 Wavelength usage per groupcast session in the wavelength-continuous OXC-based NSF14 network [Ref. Figure 3.4] | 86 |
| 3.18 Cost ratio per groupcast session in the wavelength-continuous OXC-based NSF14 network [Figure 3.4] | 87 |
| 3.19 Performance of groupcast routing and wavelength assignment (GC-RWA) for the dynamic traffic case | 93 |
| 3.20 Groupcast performance comparison between NSF and AT&T networks | 94 |
| 3.21 Wavelength assignment performance in the NSF network for groupcast sessions | 95 |
| 3.22 Performance of state-aware wavelength assignment algorithms for groupcast sessions in the NSF network | 96 |
| 3.23 State-aware wavelength assignment performance in NSF network for groupcast sessions | 97 |

| | | |
|------|---|-----|
| 4.1 | Approaches to support groupcast in wavelength-routed network . . . | 99 |
| 4.2 | Drop-and-continue optical cross-connect equipped with power splitters to support tap-and-continue architecture | 100 |
| 4.3 | Drop-and-continue node based on broadcast-and-select architecture . | 101 |
| 4.4 | Optical add/drop multiplexer (OADM) supporting drop-and-continue capabilities | 102 |
| 4.5 | Light-forest with 4 light-trees | 104 |
| 4.6 | Light-forest with 4 light-trees | 105 |
| 4.7 | Blocking probability vs. light-forest size for different light-forest algo- rithms | 107 |
| 4.8 | Performance of groupcast session for two different light-forest algo- rithms | 109 |
| 4.9 | Heuristic comparison for groupcast sessions in the NSF14 network [Fig. 3.4] | 122 |
| 4.10 | Optical-amplifiers cascaded in a chain | 125 |
| 4.11 | Performance of groupcast sessions under integrated traffic load (group- cast and unicast) | 127 |
| 4.12 | Performance of DaC-based groupcast sessions for different linear-tress routing approaches | 128 |
| 4.13 | Performance of MC-OXC-based groupcast sessions in NSF Figure 3.4 network with 64 Channel, 10,000 Calls, MCH Light-forest | 129 |
| 4.14 | Performance of MC-OXC-based groupcast sessions vs. group size . . | 130 |
| 4.15 | Performance of MC-OXC-based groupcast sessions vs. group size (net- work resource variation) | 131 |

List of Tables

| | | |
|-----|---|----|
| 2.1 | Multicast control plane | 22 |
| 2.2 | Multicast-tree Computation Algorithms | 27 |
| 2.3 | Dedicated 1+1 protection scheme using the MCH algorithm. | 34 |
| 2.4 | Dedicated 1+1 protection scheme using the MC-APE algorithm | 35 |
| 2.5 | Dedicated 1+1 protection scheme using the MC-CR algorithm. | 37 |
| 2.6 | Optimized Collapsed Ring (OCR) algorithm. | 38 |
| 2.7 | Collapsed Ring-2 (CR-2) algorithm. | 39 |
| 3.1 | Summary of optical switching technologies [1] | 44 |
| 3.2 | Wavelength selection algorithms for lightpath-based groupcast connections | 51 |
| 3.3 | Various light-forests with wavelength continuity constraint | 52 |
| 3.4 | Session, routes and wavelength assignment | 53 |
| 3.5 | Assignment of link numbers for Figure 3.5 | 53 |
| 3.6 | Session, routes and wavelength assignment for Example I | 54 |
| 3.7 | Wavelength assignment chart for Example I in Figure 3.5 | 54 |

| | | |
|------|--|----|
| 3.8 | Session, routes and wavelength assignment for Example II | 55 |
| 3.9 | Wavelength assignment chart for Example II in Figure 3.5 | 56 |
| 3.10 | The reference optical networks used in Chapters 3 and 4 | 60 |
| 3.11 | Lightpath-based groupcast RWA algorithm | 63 |
| 3.12 | Lightpath-based multicast RWA algorithm | 65 |
| 3.13 | Lightpath-based groupcast RWA algorithm #2 | 68 |
| 3.14 | Groupcast cost index (GCI) table for NSF14 network [Figure 3.4] . . | 77 |
| 3.15 | Sample of four groupcast sessions in NSF14 [Figure 3.4] | 80 |
| 3.16 | Lightpath-based groupcast with no wavelength conversion for mono- color light-trees | 81 |
| 3.17 | Lightpath-based groupcast with no wavelength conversion for multi- color light-trees | 83 |
| 3.18 | Lightpath-based groupcast with wavelength conversion for rainbow light-trees | 83 |
| 3.19 | Layered-graph based Groupcasting for Wavelength-continuous OXC- based Networks | 85 |
| 3.20 | List of 3 groupcast sessions for the NSF14 network of Figure 3.4 . . | 88 |
| 3.21 | The point-to-point lightpaths, wavelength assignment and lightpath cost resulting from three groupcast sessions in Table 3.20 | 90 |
| 3.22 | Continuous lightpath-based groupcast | 92 |
| 3.23 | Lightpath based groupcast in networks with wavelength conversion ca- pabilities | 96 |

| | | |
|-----|--|-----|
| 4.1 | Algorithms to build groupcast light-trees | 103 |
| 4.2 | Routing algorithms to build groupcast light-forest | 108 |
| 4.3 | Light-tree based groupcasts with no wavelength conversion | 120 |
| 4.4 | List of 20 groupcast sessions for the NSF14 network [Figure 3.4] . . . | 121 |

Chapter 1

Introduction

1.1 Optical Multicast and Groupcast

Wavelength-Division Multiplexed (WDM)-based optical transport networks, utilizing intelligent switches have emerged as the preferred architecture to accommodate the tremendous increase in traffic in today's networks. These networks need to provision high bandwidth sessions as well as a mixture of different types of sessions including unicast, multicast and groupcast.

The evolution of optical communication in the last two decades is shown in Figure 1.1. The first generation of optics served well to facilitate point-to-point large bandwidth communications which allowed long distance signal propagation between routers [2]. The second generation of optics dramatically improved on the bandwidth support by introducing wavelength division multiplexing (WDM); thus serving multiple communication channels on the same fiber. Also the second generation optics introduced the optical amplifiers, allowing the optical signal to travel thousands of kilometers (US coast-to-coast) without the need for regeneration. The third generation on optics gave birth to optical networking through the introduction of intelligent

optical cross-connects that were used to develop wavelength-routed WDM-based networks that can support point-and-click provisioning, as well as automatic reconfiguration and protection/restoration capabilities [3].

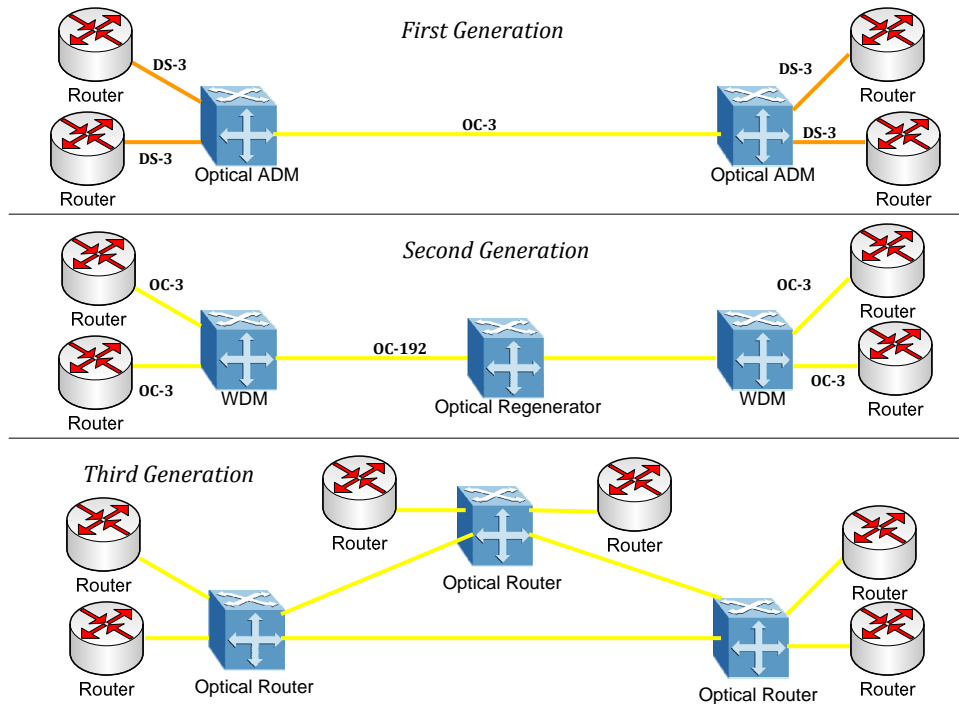


Figure 1.1: The optical communications evolution

The optical control plane was introduced with the second generation of optics as the optical network comprising of optical network elements had to be managed independent of the IP control plane. There were efforts to integrate the two control planes and the IP-over-WDM concept was introduced to achieve a joint control plane. Under a joint control plane the optical nodes support optical-electrical-optical (O/E/O) signal manipulation allowing for signal multiplexing and traffic grooming, as well as the support for multicasting relying on both the IP and optical layers. The alternative overlay model with segregated IP and optical control planes proved advantageous for all-optical communications (O/O/O), where IP network is treated as a client of

the optical transport network and multicasting is established in the optical layer by optical signal splitting.

With wide deployment of optical transport networks (OTN) and the emergence of bandwidth intensive applications, the idea of multicast and groupcast in optical networks stems from the ability to setup optical *point-to-multipoint* (multicast) and *multipoint-to-multipoint* (groupcast) sessions in wavelength-routed networks. Whereas in an optical network, a *lightpath* provides an end-to-end connection between a source and a destination node creating a unicast connection, a *light-tree* creates a multicast connection by providing an end-to-end connection between a source and multiple destination nodes [4]. A *light-forest* [5] on the other hand is a set of light-trees constructed to serve a groupcast or a collection of point-to-multipoint sessions.

There are several potential applications for optical multicast and groupcast. Applications such as video on demand, multiparty conferencing, teaching via the web, and other potential Internet applications will be the driving force behind the increase of multicast traffic in the near future. Scenarios where a group of participants can send and receive messages to and from all group members simultaneously are ideal candidates for groupcast sessions. Many applications that require groupcast are widely deployed, such as grid-computing [5], distributed interactive simulations, virtual private network (VPN) services and Ethernet LAN (E-LAN) services to name a few. As wavelength-routed networks are deployed in great numbers to meet scalability and bandwidth requirements in next-generation networks [6], groupcast will be more common in optical networks to serve multipoint-to-multipoint bandwidth intensive sessions.

Optical groupcasts are more natural and relevant for certain networks and applications. Some examples of networks that are ideal candidates of optical groupcast are described below:

- **Grid Networks:** The drive for of all-optical multicasting and groupcasting is strongly advocated by grid computing. The Global Lambda Integrated Facility [7] is paving the way for all-optical wavelength-routed global network to support grid computing. Grid networks are essential to support grid applications. Grid applications require distributed systems that may belong to different administrations to establish on-demand collaborative environment such that virtual organizations (VOs) are possible [8]. It is important to note that the definition of grid computing has expanded to include applied distributed computing technologies and environments that include everything from distributed high performance computing resources (traditional 'Grids') to horizontally scaled transactional systems supporting Service Oriented Architectures (SOA), across all scales and for all application domains [9].
- **Virtual Private Networks:** A virtual private network (VPN) is a communication network interconnecting two or more nodes or local area networks (LAN). Typically VPNs are setup over public Internet using IP security (IPSec) tunnels enabling virtual links between IP hosts. When N ($N \geq 3$) hosts partake in a single VPN, there are $N \times (N - 1)$ potential directed tunnels to enable seamless VPN transactions. Serving bandwidth intensive VPNs could be quite challenging for IP core routers, especially if diverse quality of service (QoS) is required [10]. Recent advances in WDM technology have unleashed huge amount of bandwidth in the order of terabits per second (Tbps), thus enabling the IP-over-WDM architecture to meet the bandwidth demand. For a given traffic demand, the problem of VPN design over WDM networks could be solved by establishing lightpaths in a wavelength-routed network [11, 12]. The establishment of multipoint-to-multipoint lightpaths for a group of nodes in a wavelength-routed mesh network is equivalent to the creation of a groupcast session.

- Ethernet LAN Services:** The Metro Ethernet Forum (MEF) has established standards and specifications for Ethernet services for carriers [13]. MEF standardization work will enable service providers to offer standardized services, which will be scalable and reliable, include quality of service definition for the services offered and supported by service management capabilities. MEF defines an E-Line service as a point-to-point Ethernet Virtual Connection (EVC), and an E-LAN service as any Ethernet service based on a multipoint-to-multipoint EVC as shown in Figure 1.2. The service specifications allow to serve all classes of service: from no performance assuring best-effort service to full performance assurances defined by Committed Information Rate (CIR), and associated guarantees regarding delay and loss, among other characteristics.

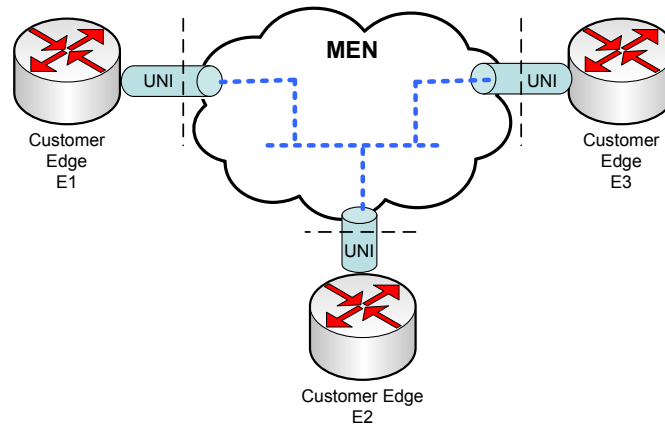


Figure 1.2: MEF multipoint-to-multipoint service

The MEF does not specify how to implement Ethernet services. However, the specifications point out what constitutes carrier-grade Ethernet services and what are the user-to-network interface (UNI) and network-to-network interface functions. As far as MEF is concerned, the standards apply to all flavors of Ethernet including, fiber, copper or microwave media irrespective of underlying infrastructures. The MEF develops specifications that define what constitutes carrier-grade Ethernet services, but not how providers should deliver Ethernet.

Traditionally, at the metro transport level, service providers have had essentially four options for how to deliver an Ethernet service: over SONET/SDH networks; natively switched and transported; over IP/MPLS networks; and using Wavelength Division Multiplexing (WDM). In addition, access methods for an Ethernet service includes copper, fiber, and even wireless.

For effective delivery of Carrier Ethernet, most service providers will be making use of fiber resources. Services such as SONET/SDH and native Ethernet can transport only a single channel over any given fiber. WDM technology helps service providers make better use of fiber infrastructures by enabling multiple wavelengths on a single fiber. However, WDM technology has one shortcoming for native Ethernet: poor manageability, leading to questionable reliability. Thus service providers will be looking at Optical Transport Network (OTN) to help support Ethernet-line (E-line) and Ethernet-LAN (E-LAN) services. The OTN (G.709) [14] is a global standard for a digital wrapper technology that enables multiple types of networks and services to be carried over a common fiber-optic infrastructure, including one that employs WDM and allows optical multicasting. Like SONET/SDH, OTN offers sophisticated OAM&P capabilities, including remote management and alarming.

1.2 Routing and Wavelength Assignment for Optical Groupcast

Groupcast, which is multipoint-to-multipoint communication, can be established using point-to-point or point-to-multipoint communications among all group members. Group communications are usually implemented by routing protocols such as IP multicast at the Layer 3 or by link-layer protocols such as Ethernet at Layer 2 of the OSI

reference model. With the advent of the optical transport network (OTN) optical groupcast at the physical layer (Layer 1 of the OSI reference model) has emerged as an important research topic.

The groupcast routing and wavelength assignment (GC-RWA) problem can be stated as follows: given a limited number of wavelengths and a set of groupcast calls, maximize the number of groupcast sessions admitted, or equivalently, minimize the call (or session) blocking probability. A groupcast is a set of multipoint-to-multipoint connections. Therefore multiple light-trees must be created to support a single groupcast session. A light-tree represents a single physical wavelength (optical channel) in all the fiber links that are part of the session-path. A network without waveband converter (WC) must obey the wavelength continuity constraint which degrades the network performance [15, 16]. Use of wavelength converters (WCs) can significantly improve network performance, however they are quite expensive and many researchers have proposed OXCs with limited WCs to increase network throughput [4].

The GC-RWA problem in essence has two basic sub-problems, namely, the routing problem and the wavelength assignment problem. The groupcast session paths may be chosen based on some criterion such as the wavelengths availability percentage in the OTN network. This becomes a coupled problem that is more suitable for dynamic traffic consideration. For the uncoupled case, session paths are first determined, followed next by the wavelength assignment phase. The uncoupled approach is suitable for OTN planning where the demand traffic matrix is given a priori.

For the dynamic GC-RWA problem, if the whole groupcast session cannot be admitted completely, it will be blocked, even though a large portion of the group members may be served. As many applications such as grid-computing and interactive multi-user applications have strict groupcast requirements, partial group communi-

cation establishment is not an acceptable outcome. Although the RWA problem is proven to be NP-hard for point-to-point lightpath [16, 15] and point-to-multipoint light-trees [4] establishment, many algorithms have been reported to address this problem. The GC-RWA is also an NP-complete problem as it requires the establishment of light-tree which is a well-known NP-complete problem. In this thesis we study many algorithms used for MC-RWA that are also extended to GC-RWA and investigate their network performance.

1.3 Thesis Objectives

The overall objective of this thesis was to introduce the concept of integrated provisioning of unicast, multicast and groupcast connections in WDM-based optical networks. We have well-established optical networks supporting point-to-point (unicast) sessions. Currently, the research community is actively addressing point-to-multipoint (multicast) sessions and multipoint-to-point sessions (example: EPON [17]) in all-optical networks. This thesis mainly addresses the last permutation option: multipoint-to-multipoint sessions in optical networks.

The first objective of the thesis was to develop a novel approach for provisioning of protected multicast sessions. Protection of services is of great importance in today's networks and is even more important when a large number of connections (as in the case of multicast sessions) can be potentially affected. This thesis proposes a novel algorithm for the protection of multicast connections in mesh optical networks. The proposed technique assumes any single link failure in the network and uses a "ring approach" to protect the multicast sessions in mesh networks.

The second objective of this thesis was to develop a network planning model to help design optical groupcast networks. The network planning process is typically the

design phase of the optical network, or any network at large [16]. The network design produces a network architecture identifying nodes and links, including determination of resources and their capacities, optimized to offer the planned network services under static traffic conditions.

The planning process for optical groupcast had several components. Firstly linear programs were designed with the objective of minimizing the network cost for a given set of optical groupcast sessions. Secondly, heuristics were designed to assess the performance of various Groupcast Routing and Wavelength Assignment (GC-RWA) algorithms based on a lightpath approach. Finally, a design methodology was generated based on static groupcast models to help with groupcast network planning.

The third major objective of this thesis was to develop GC-RWA algorithms that are based on multicast node architectures to support groupcast sessions in optical networks and provide a detailed assessment of these heuristics. A particular focus was placed on the performance of light-forests under dynamic traffic conditions. This work investigated the GC-RWA performance for two different networks, namely, multicast-capable optical switch based networks and finally optical drop-and-continue node based networks.

Three dynamic-traffic network simulators were built to investigate the three different types of optical networks: optical multicasting, lightpath-based groupcasting and light-tree based groupcasting. Additionally static-traffic network simulators were built to assess the performance of different multicast and groupcast RWA algorithms in each type of wavelength-routed network.

1.4 Thesis Outline

This dissertation investigates multicast and groupcast in optical wavelength-routed networks.

Chapter 2 deals exclusively with optical multicast. An overview of optical multicast session establishment using multicast capable and multicast incapable nodes using various routing and wavelength assignment algorithms is presented. This is followed by a novel approach on how to protect multicast sessions in mesh optical networks.

In Chapter 3 we consider typical OXC switch architectures to serve groupcast sessions. We define the concept of Groupcast Routing and Wavelength Assignment (GC-RWA) using lightpaths. We set up mixed integer linear programs (MILP) with the objective to minimize total cost for groupcast sessions in both wavelength conversion-capable and wavelength conversion-incapable OXC-based networks. In addition to the analytical model, we employ heuristics to investigate the GC-RWA performance for static traffic conditions. We develop a model for groupcast network design and capacity planning based on the proposed heuristics. Finally, we build a simulator to investigate the GC-RWA performance under dynamic traffic conditions.

In Chapter 4 we consider multicast-capable OXC nodes as well as drop-and-continue (DaC) node architectures to build light-trees to serve groupcast sessions. The problem of establishing efficient groupcast sessions through various routing and wavelength assignments (GC-RWA) is re-examined using light-trees. We consider static traffic between nodes and propose MILP models to analytically solve optical network design problems. We revise the groupcast network design model for light-trees and modify the lightpath-based heuristics accordingly. We consider the dynamic traffic conditions and light-tree based GC-RWA performance. We conclude the chap-

ter by comparing and contrasting the light-tree versus lightpath domain GC-RWA performance.

Chapter 5 summarizes the thesis and highlights the research contributions. Finally, open research topics are discussed that were not covered in this thesis.

Chapter 2

Provisioning and Protection of Optical Multicast Sessions

2.1 Optical Multicast Node Architectures

This section describes generic multicast node architectures that can be used to establish optical multicast and/or groupcast sessions.

Optical switching nodes can be classified into two broad categories: **multicast-capable (MC)** nodes and **multicast-incapable (MI)** nodes as shown in Figure 2.1. The generic functions of wavelength switching and wavelength conversion are the building blocks of most optical nodes (Figure 2.2). The multicast incapable nodes include typical optical cross-connect (OXC) and drop-and-continue (DaC) optical node architectures. Typical optical multicast-supporting nodes can be characterized by the following factors [18]:

- Signal splitting (S) the capability of signal splitting: either full, limited or no signal splitting support is provided at the node level.

- Splitting ratio (R) for the nodes supporting signal splitting, the signal power splitting ratio is either fixed or adjustable.
- Fan-out factor (F) represents the optical signal fan-out factor at the splitting nodes, which could be complete or limited and there could be networks with both type of optical splitters.
- Wavelength Conversion (C) represents the conversion capability of the network, which could be full, partial or no conversion capability depending on the optical node.
- Conversion Degree (D) represents whether the full range of conversion or limited range of conversion is supported.
- Signal amplification (A) indicates whether signal amplification occurs at the processing node. Typical signal amplification includes *retime*, *reshape* and *re-generation* of the optical pulses. The possible options are full, partial, or no amplification.

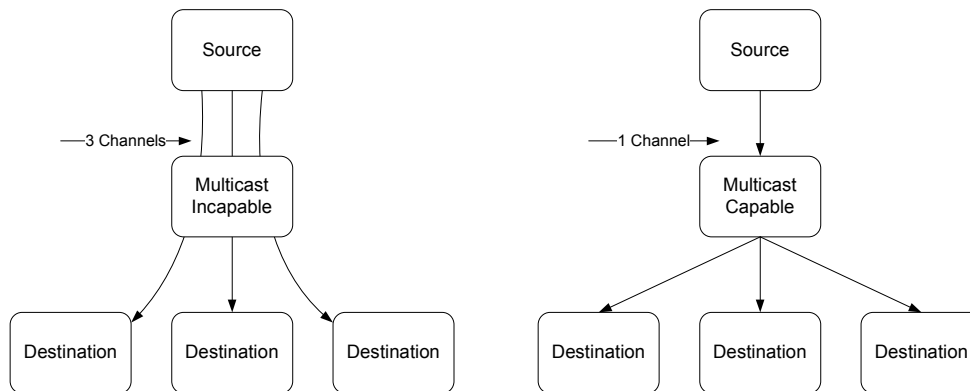


Figure 2.1: Optical multicasting with multicast-incapable (MI) and multicast-capable (MC) nodes

Many optical multicast-supporting nodes have been reported in the literature. Although "multicast-capable" broadly categorizes these systems, a diligent classification

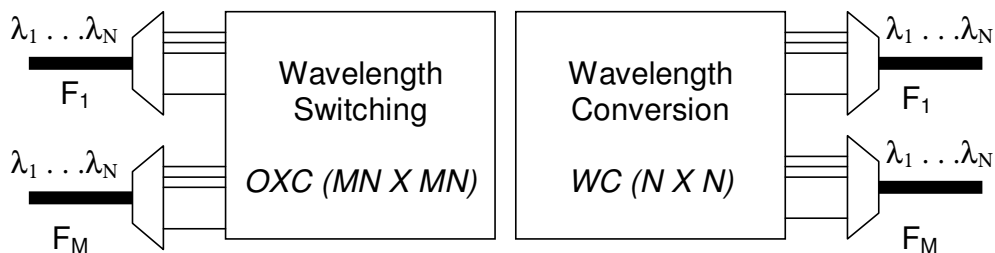


Figure 2.2: Multicast node with M fibers and N wavelengths per fiber

reveals much dissimilarities among the systems in reality. Therefore, it is important to note that many routing and wavelength assignment (RWA) algorithms are system specific and cannot be universally applied to all multicast-supporting nodes.

The key component of a multicasting-capable node is an optical splitter. An optical splitter is a passive optical device that can split an input signal without any knowledge about the optical characteristics of the input signal. By employing optical splitters in OXCs, optical layer multicast can be implemented. For example, the node shown in [19] is among the very first multicast-capable OXC (MC-OXC) architectures proposed in the literature. The splitter-and-delivery (SAD) as shown in Figure 2.3(A) can be characterized as a $S^{Full} : R^{Fixed} : F^{Complete}$ system. Many other complex switch architectures based on SAD architecture have been proposed in the literature [20], including the node in Figure 2.3(B) which is a $S^{Full} : R^{Fixed} : F^{Complete} : C^{Full}$ system also proposed in [19].

The transparent optical cross-connect presented in [21] is a multicast capable system characterized by $S^{Limited} : F^{Complete} : C^{Partial} : A^{Partial}$ node. The amplifier bank, converter bank and splitter bank shown in Figure 2.4 are limiting factors for the effective signal propagation range, the wavelength conversion range, and the number of signal copies respectively. The author points out that the splitter bank size should be augmented to reflect the nodal degree of its installation point, which is a great

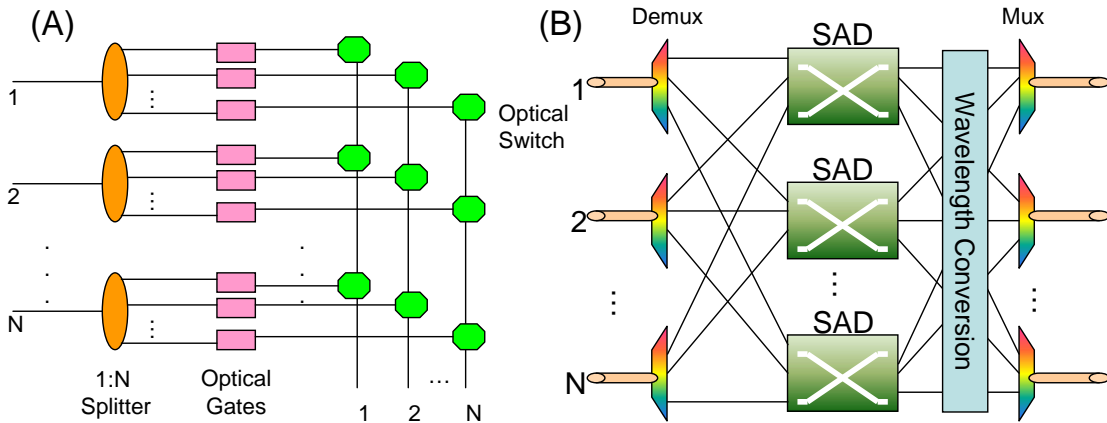


Figure 2.3: (A) $N \times N$ SAD switch (B) SAD-based MC-OXC

flexibility of this modular architecture. Also, we note that the wavelength conversion function is independent of the signal splitting function. However the signal amplification is a coupled problem with the signal spitting as an optical splitter splits an incoming signal n ways, therefore reducing the output signal power to $1/n$ of the incoming signal power.

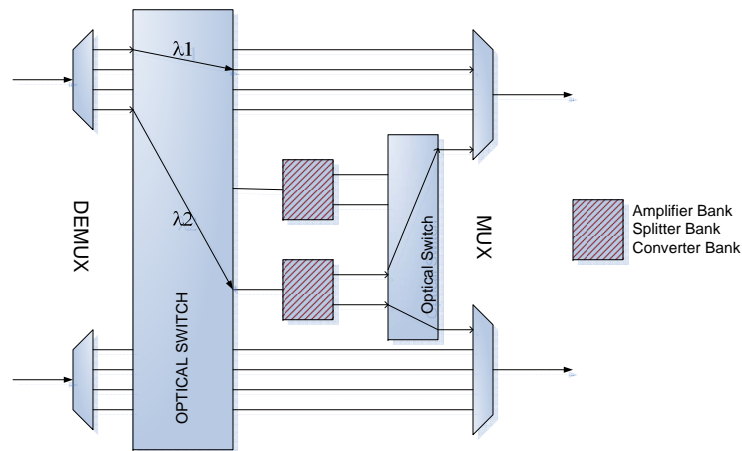


Figure 2.4: All-optical Cross-connect Switch Architecture to support Multicast

Finally the work in [22] considers a special switching element called *Multi-Wavelength Converters (MWC)* which is capable of replicating the input data from a wavelength to multiple outgoing wavelengths without using splitters. Several MWC technologies are available now [22] such as: four-wave mixing in Semiconductor Optical Amplifier (SOA), SOA delay interferometer, and supercontinuum generator in nonlinear fiber. The MWC-based WDM MC-OXC switch as shown in Figure 2.5 can be characterized by $S^{Full} : R^{Fixed} : F^{Complete} : C^{Full}$ node architecture. Work in [22] mainly focuses on the scalable architecture for MC-OXC and addresses the challenges resulting from the limitation in the number of wavelength conversion needed with traditional MC-OXC switches (Figure 2.4). The authors propose MWC to design strictly nonblocking WDM multicast switches and in doing so introduce a new class of multicast switches.

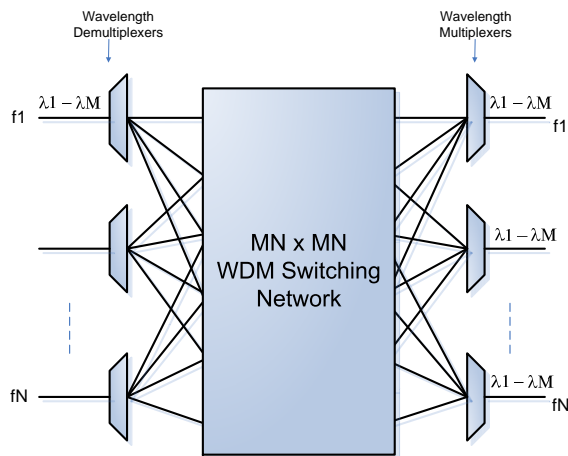


Figure 2.5: Generic structure of $M^\lambda(N \times N)$ WDM MC-OXC architecture

As we have discussed, the multicast capable optical switches vary in terms of switching design (architecture) and switching technology. The switch structure (independent of the switching technology) is the arrangements of switching systems which can be broadly classified into single-stage and multistage nodes. The characteristics of single-stage and multistage switches are described below:

- The architectures presented in Figures 2.3 and 2.4 are considered single-stage

switches, where links with the same wavelength from different input fibers are connected to one switching fabric. One major advantage of a single-stage interconnect design is its ability to scale for additional wavelengths. As it is apparent from Figure 2.2, additional switching is needed for every additional wavelength. A single-stage switching $M_{fiber}N_{wavelengths} \times M_{fiber}N_{wavelengths}$ realization on a single switching fabric is very challenging. Such a design requires $O(M^2N^2)$ switches, making it very expensive for typical large WDM networks. Thus the single-stage switch has limited connectivity pattern even if it is capable of full range wavelength conversion.

- Multistage switches are more economically realizable using large-scale interconnection of small switching units also known as the multistage interconnected architecture (MIA). The MIA interconnects a set of input fibers to a set of outputs fibers using multiple stages of small switching modules as shown in Figure 2.6. Any of the generic switch designs discussed above can be adapted to MIA to reduce the cost of switching.

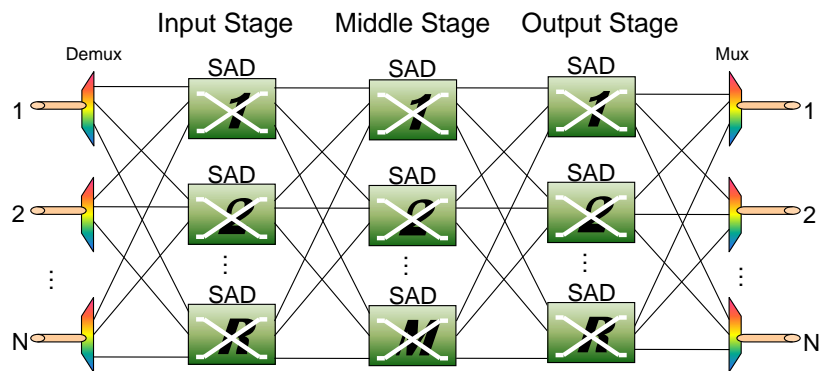


Figure 2.6: A Banyan-style Multistage Interconnected Architecture (MIA) using SAD Switches

2.2 Provisioning Multicast Sessions

The multicast session establishment in all-optical networks has been extensively studied in the literature which entails serving a multicast connection using wavelength-routed WDM networks. To support multicasting in an optical mesh network, a light-tree [12] or a Steiner Minimum tree (SMT) [12] must be build. Numerous heuristics exist to compute a multicast tree [15, 16], as finding the SMT has been proven to be an NP-complete problem.

The optical multicast routing and signaling problem is dependent on the underlying hardware building blocks in the network. The multicast traffic can be dispersed to its intended destinations in three fundamental ways:

- Relying only on IP or packet routing layer (I)
- Relying on the combination of both IP and optical layers (II)
- Relying only on the transmission media (optical layer) (III)

For option (I), packet duplication happens in the electronic layer and the optical transmission media is only used to transport the packets from source to destination. The advantage of IP layer multicast is that the routing decision of the IP routers is independent of the transmission layer and the IP multicast protocols would function on any physical layer transmission technology including wireless and optical media. The biggest disadvantage of IP layer multicast is the bottleneck created due to the electronic layer packet duplication and packet processing that is not scalable for bandwidth intensive multicast sessions as IP routers are designed to inspect and forward packets on hop-by-hop basis. For option (II), the routing decision for multicasting in IP-over-WDM networks is made using the joint control plane, while the bit stream duplication is accomplished in the optical layer using optical splitters. Finally, the

optical layer-based multicasting as described in option (III) only relies on the optical layer and there is no interaction with the electronic layer in the multicasting decision. The optical node is solely responsible for handling the multicast traffic. Multicast types (II) and (III) described above are identical to the control plane coordination between the WDM layer and IP layer for unicast traffic, namely the overlay and peer models. In the overlay approach, the job is totally handled by the WDM layer alone with partial or no coordination with the IP layer, while the peer model both IP and WDM layer coexist and cooperate. The thesis focuses on optical multicast and groupcast traffic and the remainder chapters will focus primarily on the overlay control plane model [23]. In an effort to provide a complete picture of optical multicasting, the peer model is briefly presented in the following section and elaborated wherever needed.

Multicasting has evolved from Type I to Types II and III as will be described below. Routing decisions in electronic layers are governed typically by the IP layer. Popular IP multicast protocols include:

- Internet Group Management Protocol (IGMP) [24]: It is a management communication protocol analogous to Internet Control Message Protocol (ICMP). IGMP is used by IP hosts and adjacent multicast routers to establish multicast group memberships. IP multicast sends messages from source to multicast group address and a host interested to participate in the IP multicast must use IGMP control messages to let the router know its intentions.
- Protocol-Independent Multicast (PIM): Supports both point-to-multipoint and multipoint-to-multipoint communication over IP. There are four variants of PIM:
 - PIM Sparse Mode (PIM-SM) explicitly builds unidirectional shared trees

- rooted at a Rendezvous Point (RP) per group, and optionally creates shortest-path trees per source [25].
- PIM Dense Mode (PIM-DM) implicitly builds shortest-path trees by flooding multicast traffic domain wide, and then pruning back branches of the tree where no receivers are present [26].
 - Bidirectional PIM explicitly builds shared bi-directional trees. It does not build a shortest path tree and it needs no source-specific state [27].
 - PIM Source Specific Multicast (PIM-SSM) builds trees that are rooted in just one source and appropriate for broadcasting of content. In SSM, an IP datagram is transmitted by a source S to an SSM destination address G, and receivers can receive this datagram by subscribing to channel (S,G) [28].
- Distance Vector Multicast Routing Protocol (DVMRP) [29]: Uses RIP protocol to share information between routers to transport IP Multicast packets among networks and to build the multicast routing table based on inter-router hop distance. The DVMRP protocol uses IGMP messages to exchange information with other routers.
 - Multicast Open Shortest Path First (MOSPF) [30]: Uses the open shortest path first (OSPF) protocol to support multicast routing and to share information about multicast sessions between OSPF areas and within a specific OSPF area.
 - Multicast BGP (MBGP) [31]: It is an extension to the BGP unicast routing protocol that allows parallel distribution of IP address families. The information about the topology of IP Multicast-capable routers is exchanged independently of the IPv4 topology allowing multicast topology to be different from the unicast routing topology. In addition to MBGP protocol for routing information, other

protocols such as the PIM family are needed to build trees and forward multicast traffic.

In IP-over-WDM networks, signaling is needed for lightpath establishment and graceful termination of a connection. Depending on whether we are using the peer or overlay model, the signaling algorithms will impact the network performance significantly [23, 32]. The overlay model in IP-over-WDM networks can apply any IP multicast routing to RSVP-TE [33] signaling protocol. The peer networks deploy OSPF-TE [34] for the routing and RSVP-TE protocol for unicast signaling and light-path provisioning [35]. As for multicasting, the MOSPF protocol can be superimposed on the peer model to support hybrid multicasting. The Traffic Engineering (TE) extensions to OSPF add dynamic properties to the route calculation algorithm. The properties are [34]:

- Maximum Reservable bandwidth
- Unreserved bandwidth
- Available bandwidth

The OSPF-TE fields are distributed between network nodes via the TLV fields of an opaque LSA. Each hybrid MOSPF router periodically collects information about multicast group membership via IGMP. This information, along with the above link-state information, is flooded to all other routers in the routing domain using OSPF LSAs. Each multicast node, since it understands the topology of the entire network, can then independently calculate a least-cost spanning tree with the multicast source as the root and the group members as leaves. This tree is the path that is used to route multicast traffic from the source to each of the group members. Note that all routers will calculate exactly the same tree, since they periodically share link-state information. The link state information is collected using the Link Management

| PROTOCOL | Function |
|--|---|
| LINK MANAGEMENT: Link Management Protocol | Neighbor discovery, Maintain control channel connectivity, Verify data link connectivity, Correlate link property information, Suppress downstream alarms, Localize link failures |
| ROUTING: Open Shortest Path First- Traffic Engineering (OSPF-TE) | Distribute TE link information, Advertise nodes in the network and create topology, Calculate constrained shortest path (CSPF) and Routing information for control and data plane |
| Multicast Open Shortest Path First (MOSPF) | Flood the multicast group membership information along with the link states, Build The shortest path multicast tree upon demand using Dijkstras algorithm. (All routers calculate the same source-based shortest-path delivery tree.) |
| SIGNALING: Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) | Signals setup/teardown/refresh of paths with QoS requirements (e.g., circuit size), Uses control channel to setup an optical LSP, Supports refresh reduction, Supports Explicit Route Object (ERO) and Record Route Object (RRO). |

Table 2.1: Multicast control plane

Protocol (LMP). LMP is an auxiliary protocol for GMPLS that runs between adjacent nodes. It has four functions: connectivity verification, property correlation (via Link Summary messages exchanged between neighbor nodes), fault isolation and control channel management. Among the services provided by LMP, there is a link ID service between nodes that can be used to identify physical resources. A control channel separated from data channel is used for LMP due to the special needs of the GMPLS environment and the variety of the transmission media involved. The control plane protocols for optical multicast are summarized in Table 2.1.

Note that in the peer model the optical subnet is treated as the peer of the IP subnet. The MOSPF will run both in the optical and the IP domains as the protocol is technology agnostic. Figure 2.7 illustrates MOSPF implementation in both the

optical and IP domains. Area 0 represents the optical core network while Area 1 and Area 2 represent the IP networks. Note that Multicast Area Border Router (MABR) routers must be able to communicate with both areas. The LMP and OSPF-TE with opaque-LSA run inside Area 0 only. The MOSPF membership LSAs run within Area 1 and Area 2.

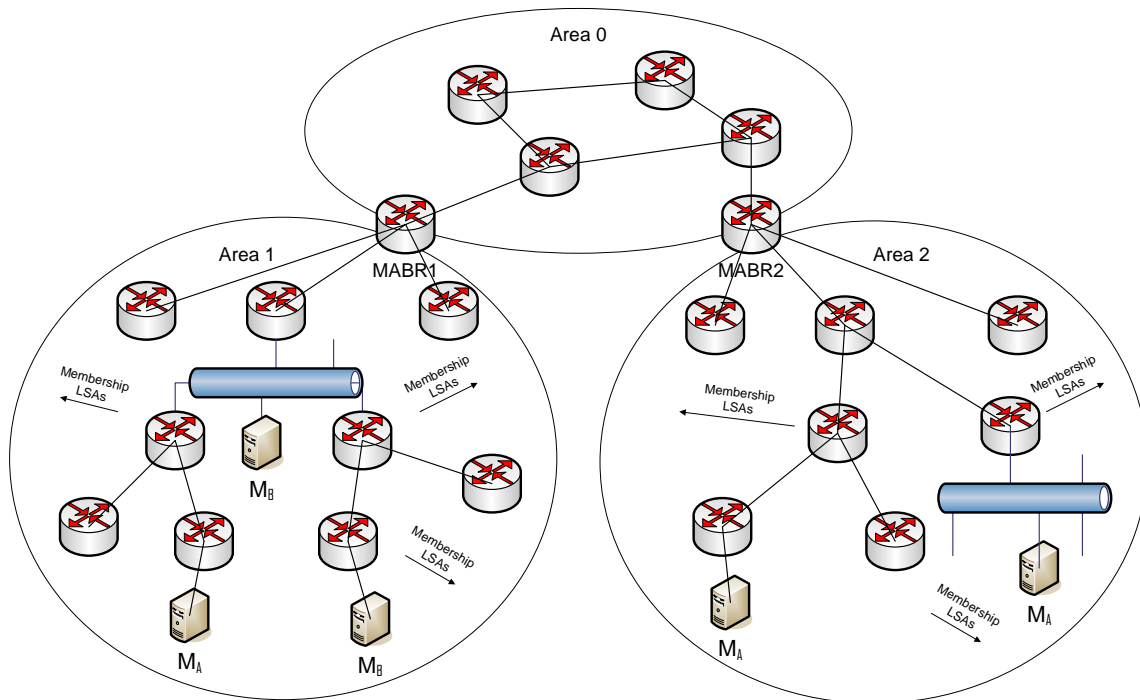


Figure 2.7: Multicasting using MOSPF

The routing and signaling are the two major building blocks for the optical control plane. The routing module helps to establish the end-to-end path for an optical session whereas the signaling module actually establishes the optical session through one or many optical nodes. In the optical context the number of wavelengths per fiber, the bandwidth per channel and the wavelength conversion capability are relevant information for the control plane. Additionally, link management information is needed for session protection and restoration mechanisms.

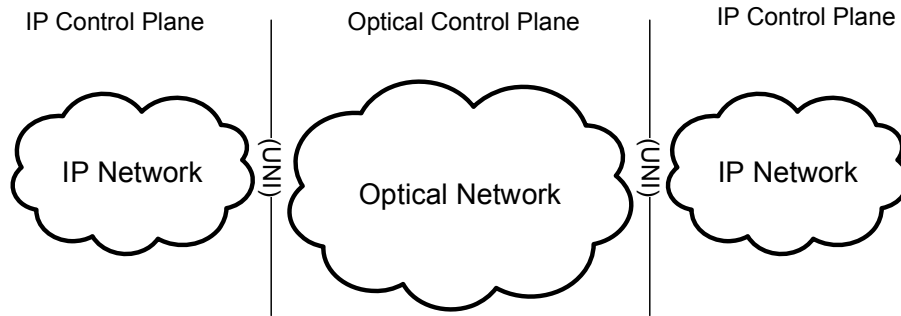


Figure 2.8: Overlay Model for Optical Network

For the optical multicast network, in this thesis we primarily consider the overlay model. The overlay model treats the IP network as a client of the optical network and hence defines a clear user network interface (UNI) as shown in Figure 2.8. From the overlay definition it is clear that the IP multicast protocols are incompatible in the optical domain. As optical multicasting is still evolving, the framework for optical multicast control plane is yet to be established. However we can implement an optical multicast control plane by leveraging the existing optical control plane protocols (such as OSPF-TE, RSVP-TE and LMP) and incorporate a few more functional blocks to handle the multicasting aspects. Below we highlight the functions of optical control plane protocols.

- **Link Management Protocol (LMP):** LMP is a link management protocol used for a variety of equipment including optical cross connects and photonic switches. In new optical equipment with many fibers and ports, each node is connected to many other nodes, so manual neighbor configuration is time-consuming. Neighbor discovery allows automatic inventory of links between nodes; automating identification between neighbors, and proper resource accounting. In addition, control channel connectivity is maintained through LMP Hello messages and

fault detection and localization is promoted as failure detected at node downstream of the flow is flagged upstream with a channel status message.

- **Open Shortest Path First Traffic Engineering (OSPF-TE):** OSPF-TE defines traffic engineering extensions to the OSPF routing protocol originally developed to distribute link information about data networks. It extends OSPF to allow the distribution of link information relating to the control channel as well as the information about the TE links in the transport network. OSPF-TE also sends information about the TE links and their switching capability (in the case of SONET for example). TE link information is made available using Opaque LSA (Type 10). The maximum bandwidth part of the Interface Switching Capability Descriptor sub-TLV is provided. In addition, the protection and restoration information is made available in Link Protection sub-TLV and SRLG sub-TLV where SRLG indicates shared link risked group information. The OSPF-TE asks Constrained Shortest Path First (CSPF) function for a route based on the following constraints: Minimum LSP Bandwidth, Encoding Type, and Switching Type. Once OSPF-TE receives from CSPF a route, it can indicate to RSVP what information to be used in the signaling message. An Explicit Route Object (ERO) is generated to indicate whether the signaling should be strict or loose. A strict ERO must be followed by the signaling message where a loose ERO is a combination of at least one loose hop and 0 or more strict hops. To reach a next loose hop, a node can request a new CSPF calculation.
- **ReSerVation Protocol-Traffic Engineering (RSVP-TE):** RSVP-TE is a signaling protocol extended from RSVP messages that are sent over the control channel to set up LSP (Label Switched Paths) in the transport/data layer. Paths are set up for a long period of time in the semi-permanent manner. Refresh is used

to keep the circuit state information synchronized and updated between nodes. Among the optical transport requirements are to provide paths with control channel failure survival and RSVP messages not traveling in the same as traffic bearing channel.

One way to support multicast is to construct multiple point-to-point lightpaths for the multicast session. Besides running LMP, OSPF-TE and RSVP-TE, the multicast source node must also calculate routes to the destination node set. This could be easily accomplished by looking-up the routing table in OSPF-TE, generated based out of constraint routing or Dijkstras algorithm, which will generate the shortest path to all destination in the optical network domain. A multicast function block (MFB) with the OXC will copy all the routes corresponding to the multicast destination set from OSPF routing table and trigger RSVP-TE to send signaling messages to the destination set. Once all lightpath establishment is completed, the MFB hands-off the call to the regular lightpath management control unit.

2.3 Multicast Routing and Wavelength Assignment (MC-RWA)

The multicast routing or light-tree computation requires full knowledge of the network connectivity. We assume that the multicast session request presented to the source node has the network topology information already stored in its database. The question of how to acquire and propagate link connectivity information is a separate issue which is typically addressed by routing protocols like OSPF. Once the multicast request arrives there are numerous algorithms available to compute the multicast tree as shown in Table 2.2.

| | |
|--|---|
| <i>Minimum Cost Heuristic (MCH)</i> | |
| <i>Step 1</i> | Given a network represented by the graph $G(V, E)$, with vertices and edges, and a source node s , compute the shortest path from the source node to all other nodes using the Dijkstra's shortest path algorithm. |
| <i>Step 2</i> | Given the destination set $D = \{d_1, d_2, \dots, d_n\}$, choose the corresponding paths as a set of links for each destination and combine links in a link set $L = \{l_1, l_2, \dots, l_k\}$. |
| <i>Step 3</i> | Delete the duplicated links from the link set ' L ' and the minimum cost multicast tree is constructed. |
| <i>Linear Tree Algorithm (LTA)</i> | |
| <i>Step 1</i> | Given a network represented by the graph $G(V, E)$, with vertices and edges, and a source node s and the destination set $D = \{d_1, d_2, \dots, d_n\}$, compute the shortest path from the source node to the destination d_1 using the Dijkstra's shortest path algorithm and augment the path in a link set $L = \{l_1, l_2, \dots, l_k\}$ |
| <i>Step 2</i> | For all the destination nodes that have been covered in the previous step remove them from the destination set D . Set the d_1 as the source node and go to <i>Step 1</i> . |
| <i>Step 3</i> | Repeat <i>Step 1</i> and <i>Step 2</i> until $D = \{\}$ and the resulting link vector ' L ' is the linear multicast tree. |
| <i>Minimum Cost Linear Tree Algorithm (MCLT)</i> | |
| <i>Step 1</i> | Given a network represented by the graph $G(V, E)$, with vertices and edges, and a source node s , compute the shortest path from the source node to all other nodes using the Dijkstra's shortest path algorithm. |
| <i>Step 2</i> | Given the destination set $D = \{d_1, d_2, \dots, d_n\}$, choose the destination d_k with the shortest path as a set of links for each destination and augment the links in a link set $L = \{l_1, l_2, \dots, l_k\}$. |
| <i>Step 3</i> | For all the destination nodes that have been covered in the previous step remove them from the destination set D . Set the d_k as the source node and go to <i>Step 1</i> . |
| <i>Step 3</i> | Repeat <i>Step 1</i> through <i>Step 3</i> until $D = \{\}$ and the resulting link vector ' L ' is the minimal cost linear multicast tree. |

Table 2.2: Multicast-tree Computation Algorithms

We present the following example in order to illustrate the working of the three algorithms in Table 2.2: MCH, LTA and MCLT. Figure 2.9 shows the NSF network with link costs, and a multicast session from TX^* to destination set $D = \{CO', PA', MD', MI'\}$. We apply the MCH and the effective path is as follows: $(TX^* \rightarrow CO')$, $(TX^* \rightarrow MD')$, $(TX^* \rightarrow GA \rightarrow PA' \rightarrow NY \rightarrow MI')$. Similarly, we apply LTA algorithms and the path we get is as follows: $(TX^* \rightarrow CO' \rightarrow NE \rightarrow IL \rightarrow PA' \rightarrow NJ \rightarrow MD' \rightarrow NY \rightarrow MI')$. Finally, we apply MCLT and obtain the following path: $(TX^* \rightarrow CO' \rightarrow NE \rightarrow IL \rightarrow PA' \rightarrow NY \rightarrow MI' \rightarrow NJ \rightarrow MD')$.

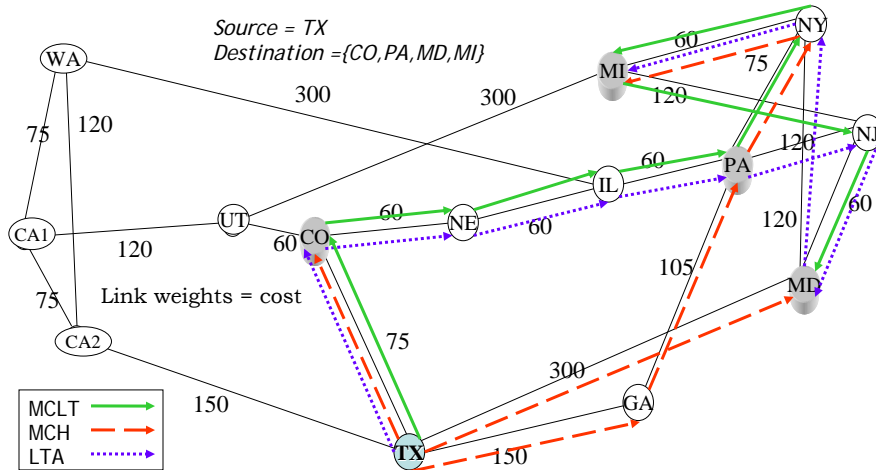


Figure 2.9: Multicast trees from MCH, LTA and MCLT Algorithms on NSF Network

2.3.1 Performance Evaluations

In this section, we present results for the performance evaluation of the various MC-RWA algorithms for 14-node, 21-link NSF network (Figure 3.4) and 24-node, 43-link AT&T network (Figure 3.10). Unless otherwise specified, the following assumptions and parameters are used throughout the simulation:

- Each node in the network is an MC-OXC with full splitting capability, but with no wavelength conversion capability.
- Each adjacent node pair is interconnected by one (bidirectional) fiber and each fiber is equipped with W wavelengths/fiber.
- There is no power loss due to splitting of the light signal and the signal is amplified to maintain the terminating nodes' minimum sensitivity level at the receiver.

Figure 2.10 shows the performance of MC-RWA using the MCH heuristic and the First-fit wavelength selection algorithm. We observe that as we increase the network resources in terms of number of wavelengths per fiber the blocking improves. We have used a load of 100 Erlangs to generate the graph in Figure 2.10 where all the traffic was comprised of multicast sessions only.

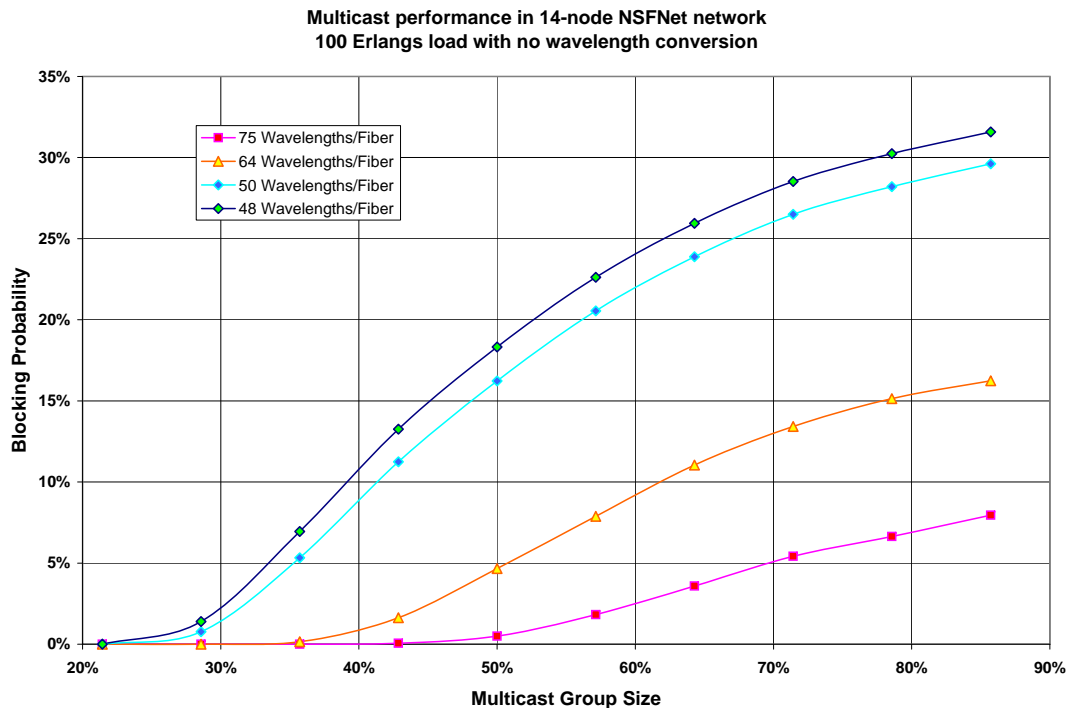


Figure 2.10: Blocking Performance vs. Multicast Group Size in NSF network

In the next experiment we compare the performance of MC-RWA in a large AT&T network (Figure 3.10) with a smaller NSF network (Figure 3.4) and the blocking performance results are presented in Figure 2.11. It appears that larger sessions perform worse in the larger network for the same number of wavelengths/fiber or network resource distribution.

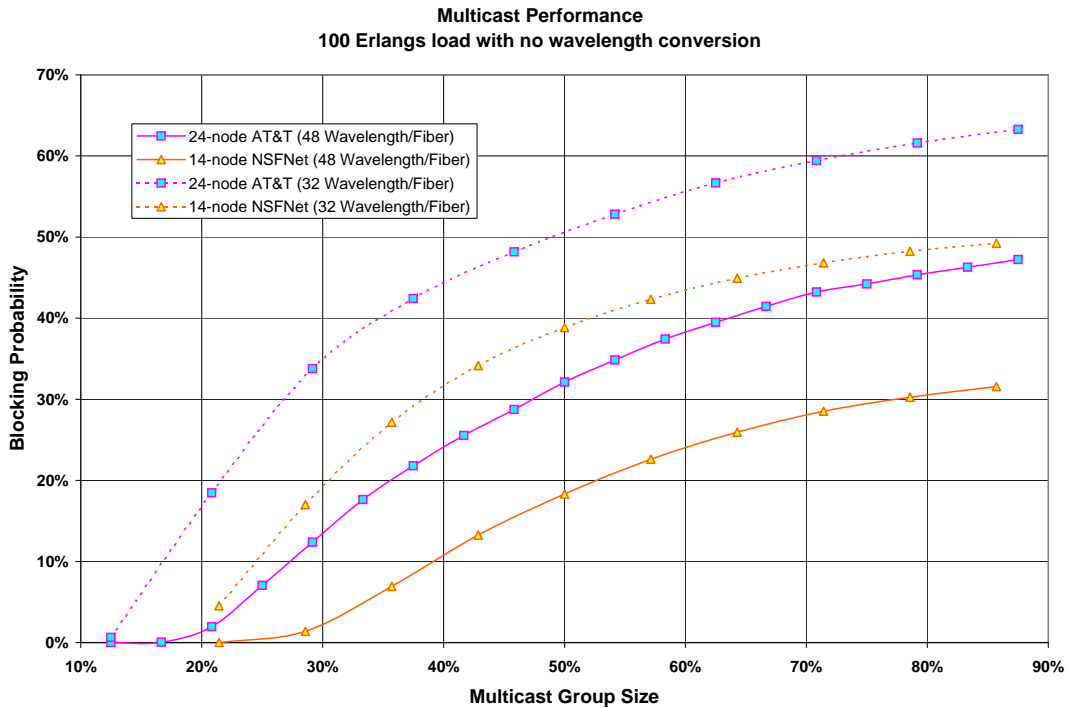


Figure 2.11: Blocking Performance vs. Multicast Group Size in NSF and AT&T networks

As MC-RWA includes both routing and wavelength assignment, we gain very little insight of how the routing algorithms are performing independent of the wavelength assignment algorithms. We decided to use the network cost as a metric to benchmark the performance of multicast routing algorithms described in Table 2.2. We conducted experiments on the NSF14 network (Figure 3.4) with 100,000 multicast sessions and computed the average session cost per multicast group size. Figure 2.12 summarizes the performance results of the multicast routing algorithms. In Figure 2.12 the ACTA clearly outperforms all other algorithms. We have introduced a variation for each of

the three algorithms described in Table 2.2 by substituting hop count as a cost metric instead of the actual link costs and for all the three algorithms (MCH, LTA and MCLT) their counter parts (MCH-hopcount, LTA-hopcount and MCLT-hopcount) show degraded performance in terms of network cost.

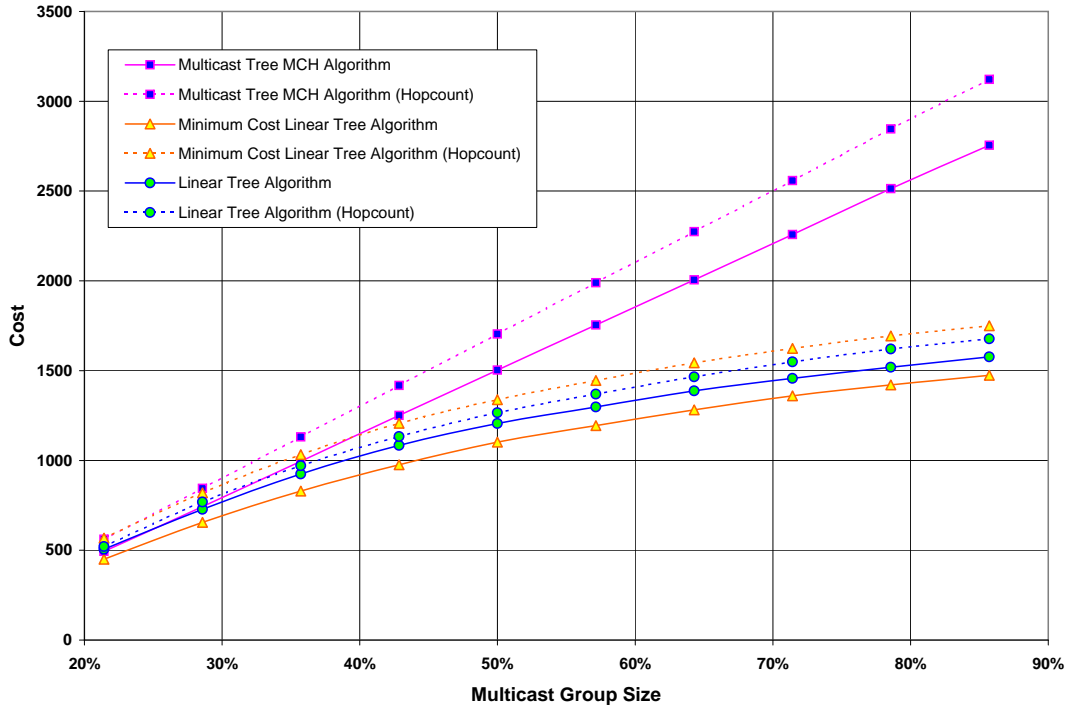


Figure 2.12: Average Multicast Session Cost vs. Multicast Group Size in NSF14 network

2.3.2 Provisioning Integrated Unicast and Multicast Traffic

In real networks, the multicast traffic will share network resources with unicast sessions and it is prudent to evaluate their performance in the presence of mixed network traffic load. We therefore conducted an experiment where we allowed on average half of the traffic to be general unicast traffic, from any node to any other node in the network. We then benchmarked the performance of each of the multicast heuristic algorithms described in Table 2.2 namely MCH, LTA and MCLT as well as the hop

count variation of the heuristics where the hopcount is used as a cost function instead of the link weights. Figure 2.13 is obtained by loading the NSF14 (Figure 3.4) with 100 Erlangs of traffic. The network links were designed with 32 wavelengths/fiber without any support for wavelength conversion at any multicast-capable node. Note that all nodes were multicast capable in the network. Using the first-fit algorithm for wavelength assignment, the network would block a calls if it was unable to route the call or no wavelength was available to connect the session. We observed that the MCH heuristic yields better performance than LTA and MCLT algorithms and MCH (Hopcount) outperforms all other heuristics.

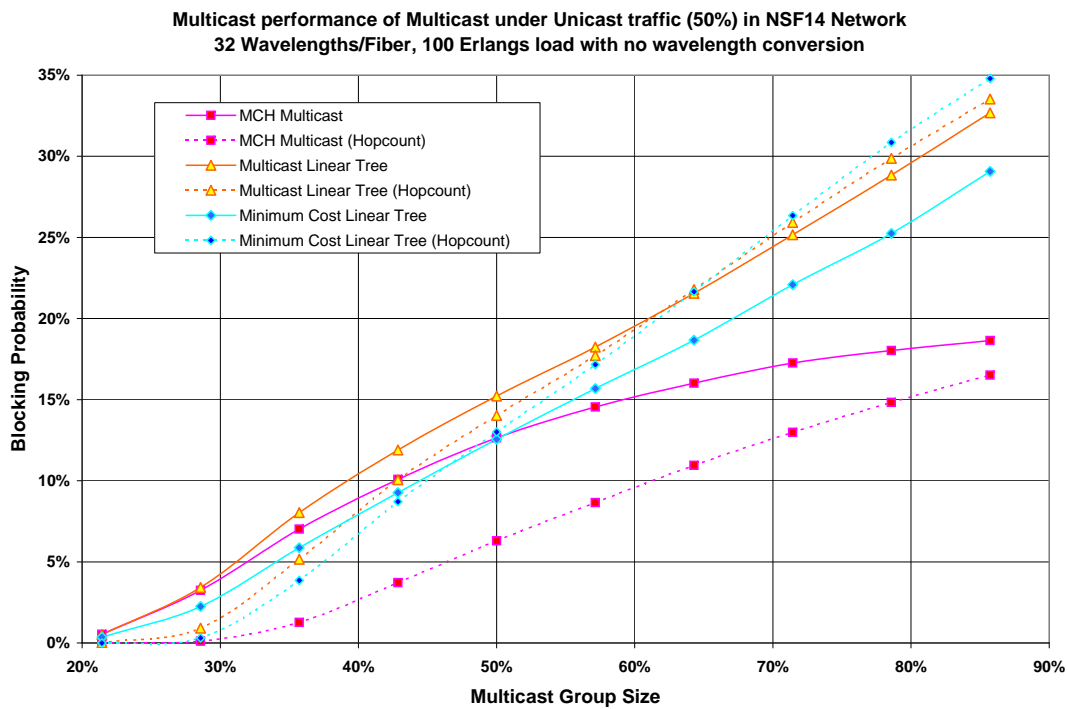


Figure 2.13: Blocking Performance vs. Multicast Group Size in NSF network

2.4 Optical Multicast Protection

For multicast applications to be viable, it is important that the multicast traffic is not only routed efficiently through the optical network but it is also protected against any possible failures in the network. This means that the multicast traffic has to be protected against any potential fiber (link) failure and that the recovery of the traffic has to be fast in order for the networks to meet their prescribed Service Level agreement (SLA) requirements [16].

The problem of protection/restoration of unicast connections in mesh optical networks has also been extensively investigated in the last few years. The problem of effectively protecting multicast connections however has only recently started to receive some attention by the research community. This is precisely the focus of this section. This work investigates the problem of dedicated 1 + 1 protection of multicast connections in WDM mesh optical networks utilizing a number of different algorithms, and compares these algorithms among themselves and to the ones proposed in the literature.

Dedicated 1 + 1 protection of a multicast session requires establishing a working multicast tree and a protection multicast tree before assigning a wavelength to the session. Failing to find a working path, a protection path or a wavelength will result in blocking a multicast session. This problem is essentially equivalent to finding link-disjoint light-trees for a multicast session. Three algorithms for finding such working and protection light-trees are discussed and their simulation results are presented in the sections that follow.

| | |
|--------|---|
| Step 1 | For a working-path, create a light-tree using MCH. |
| Step 2 | Remove all edges along the working path. |
| Step 3 | Compute a protection-path based on the remaining graph using MCH. |

Table 2.3: Dedicated 1+1 protection scheme using the MCH algorithm.

2.4.1 Multicast Protection Algorithms

A minimum cost heuristic (MCH) is initially used in order to build a light-tree, in a two-step approach. First, for any multicast call, the minimum cost path (using weighted Dijkstras algorithm [16]) is computed from the source node to each of the destination nodes. Second, all the paths found in step one are combined to construct the multicast tree. Table 2.3 shows a very simple way of providing 1 + 1 protection for a multicast session using the minimum cost heuristic (MCH) to compute multicast trees.

The disadvantage of the above heuristic however, is that a multicast session cannot be protected when all of the links attached to a node are used to construct the primary tree for that session. This is clearly evident in networks with nodes with degree two (such as nodes s , and d_2 in the network example shown below in Figure 2.14), when both links attached to these nodes are used to construct the working tree. The second algorithm described in Table 2.4 addresses this issue by reserving some edges for future use in the protection tree. Specifically, one random link (edge) from the source and one random link (edge) from all the destination nodes are pruned and saved for protection tree calculation in the minimum cost with all pruned edges (MC-APE) algorithm. The algorithm in essence reduces the chance of creating an isolated node and consequently the blocking probability for multicast sessions by randomly removing an edge from source and all destination nodes before the working tree is computed.

| | |
|--------|---|
| Step 1 | For a working-path, randomly remove one (1) edge from the source and all destination nodes. |
| Step 2 | Create a light-tree based on the remaining graph using minimum cost heuristic (MCH). |
| Step 3 | Remove all edges along working path and restore the pruned edges from Step 1. |
| Step 4 | Compute a protection-path based on the resulting new graph from Step 3 using MCH. |

Table 2.4: Dedicated 1+1 protection scheme using the MC-APE algorithm

Computing link-disjoint multicast trees requires too many network links and finding a wavelength to serve a dedicated 1 + 1 multicast session becomes increasingly difficult with higher traffic loads. If one looks carefully at the definition of the 1 + 1 dedicated service, it is understood that at any given time all destination nodes will receive two simultaneous signals from the source and should survive any single link failure affecting either working or protection path. Unidirectional path-switched SONET rings (UPSR) [15] are well suited for multicasting and have the required 1 + 1 Automatic Protection Switching (APS) performance. Thus, much lower blocking probability for multicast session can be expected if a ring configuration could be established in a mesh network.

The minimum cost collapsed ring (MC-CR) algorithm described in Table 3 finds a ring path among the source and all destination nodes. Unlike SONET, it is not required to terminate the signal at the generating source node. Hence, two light-paths are established each starting in opposite directions from the source and terminating at the final node of the destination set after traversing the ring path. An illustrative example is shown in Figure 2.14. Note that dedicated protection using the MCH algorithm requires 9 links whereas the MC-CR algorithm uses only 6 links to protect the multicast session.

When inspected closely, it is apparent that the MC-CR algorithm is not cost opti-

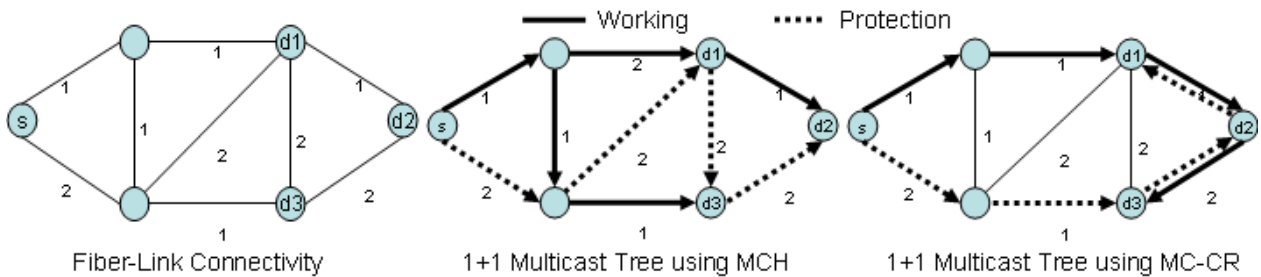


Figure 2.14: Example illustrating MCH and MC-CR algorithms

mized as in **Step 2** of Table 2.5 there are no selection criteria for choosing destination node sequences from the destination node set. Hence, the authors in [36] suggested to use Dijkstra's minimum cost algorithm to qualify the destination node selection criteria. The heuristic noted in [36] is called optimized collapsed ring (OCR) algorithm which is described in Table 2.6.

Although the OCR algorithm improves the performance compared to MC-CR algorithms as shown in [36], it has some limitations as illustrated in Figure 2.15. For example, in Figure 2.15 the OCR fails to serve d_1 and d_2 as the shortest path, always routing through node x . Also, the CR algorithm fails to serve this session, although clearly there exists a solution. So, we propose the collapsed ring-2 (CR2) algorithm described in Table 2.7 that would serve sessions like those in Figure 2.15. The main idea is to use hop count instead of the actual cost to find destination nodes that are close to the source. The CR2 algorithm outperforms the OCR heuristic consistently as shown in Figure 2.18 and the details of the simulation are presented in the next

| | |
|--------|--|
| Step 1 | Given a source = s and destination set = d_1, d_2, \dots, d_n , form a ring set of vertices v_1, v_2, \dots, v_n , where v_1 is the source node and $v_2 \dots v_n$ are destination nodes. |
| Step 2 | Find the minimum cost path from v_i to v_{i+1} where $1 \leq i \leq n - 1$ and remove all edges along the path. If a node v_x is part of the path, then remove the vertex v_x from the ring set. |
| Step 3 | Repeat Step 2 until all the vertices from the ring set are chosen and a working-path is created. |
| Step 4 | For the protection-path, find a minimum cost path from source to the last destination node in the working path and traverse the working path backward till the first destination node is reached. |

Table 2.5: Dedicated 1+1 protection scheme using the MC-CR algorithm.

performance section.

2.4.2 Simulation Results

We consider the NSFNET 14-node backbone network (Figure 3.4), where all nodes are interconnected with bi-directional fiber links and each link carries W wavelengths. We also assume that each node is equipped with a multicast-capable optical switch (MOSW) and no wavelength converters are used to establish a light-tree. We use a dynamic traffic model where multicast call requests arrive at each node according to a Poisson process and the holding time is negative exponentially distributed with a unit mean. In each simulation 50,000 multicast requests were generated, and the results were averaged over five simulation runs. We show the performance of the MC-RC algorithm under various network load conditions in Figure 2.16. To obtain the plot we fixed the multicast group size at 5 and varied the number of wavelengths ($W = 8, 16, 32, 64$) per fiber link.

We also considered 64 wavelengths per link in the first experiment to evaluate the blocking probability vs. session size for a network load of 100 Erlangs shown in

| | |
|--------|--|
| Step 1 | Given a source = s and destination set = d_1, d_2, \dots, d_n , form a ring set of vertices v_1, v_2, \dots, v_n , where v_1 is the source node and $v_2 \dots v_n$ are destination nodes. |
| Step 2 | For primary path, find the closest destination to the source using the shortest path algorithm, e.g. d , and remove all the links along the path. The new source is set to be d . |
| Step 3 | Repeat Step 2 until no more vertices are included in the ring set. This will form the primary path. |
| Step 4 | For backup path, find the shortest path from source to the last destination node in the backup path and traverse the primary path backward until the first destination is reached. |

Table 2.6: Optimized Collapsed Ring (OCR) algorithm.

Figure 2.17. MCH has the highest blocking probability as it allocates most of the links in building the primary tree. MC-APE is observed to perform better than MCH as less isolated nodes are created by pruning edges for the protection path. The MC-CR performs better than both algorithms for larger session sizes. Only for very small session sizes, where more links are wasted to establish the ring connectivity, the edge-pruning algorithm performs better than constructing a collapsed-ring.

For the optimized collapse ring (OCR) and the collapse ring-2 (CR2) heuristic performance comparison, we generated 100,000 fixed size multicast sessions. We used the same NSF14 network where each link was equipped with 64 wavelengths/fiber and no wavelength converters were used in the simulation. To benchmark the network, we also ran the MCH heuristic and all the blocking performance for multicast 1+1 protected sessions are shown in Figure 2.18. We note that the CR2 has consistently outperformed the OCR heuristic for all offered multicast group sizes and continuous network load of 100 Erlangs.

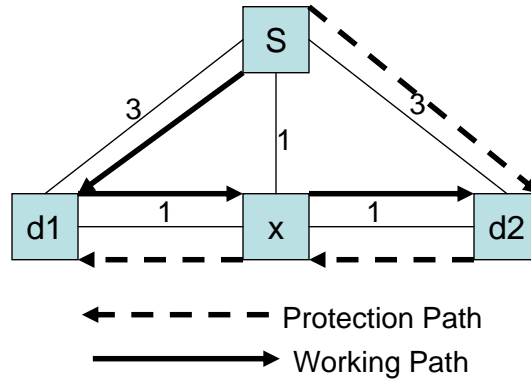


Figure 2.15: Example illustrating CR2 algorithm

| | |
|--------|--|
| Step 1 | Given a source = s and destination set = d_1, d_2, \dots, d_n , form a ring set of vertices v_1, v_2, \dots, v_n , where v_1 is the source node and $v_2 \dots v_n$ are destination nodes. |
| Step 2 | Replace the real cost of the network with a unit cost. For primary path, find the closest destination to the source using the shortest path algorithm, e.g. d , and remove all the links along the path. The new source is set to be d . |
| Step 3 | Repeat Step 2 until no more vertices are included in the ring set. This will form the primary path. |
| Step 4 | For backup path, find the shortest path from source to the last destination node in the backup path and traverse the primary path backward until the first destination is reached. |

Table 2.7: Collapsed Ring (CR-2) algorithm.

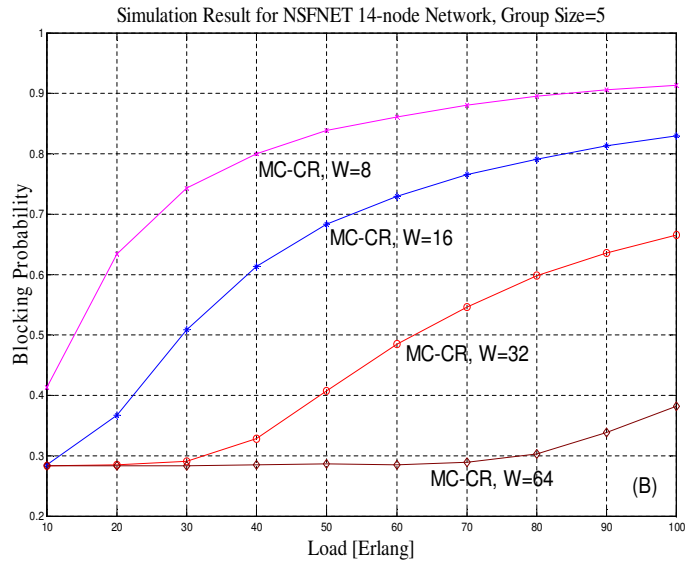


Figure 2.16: Blocking performance vs. network load

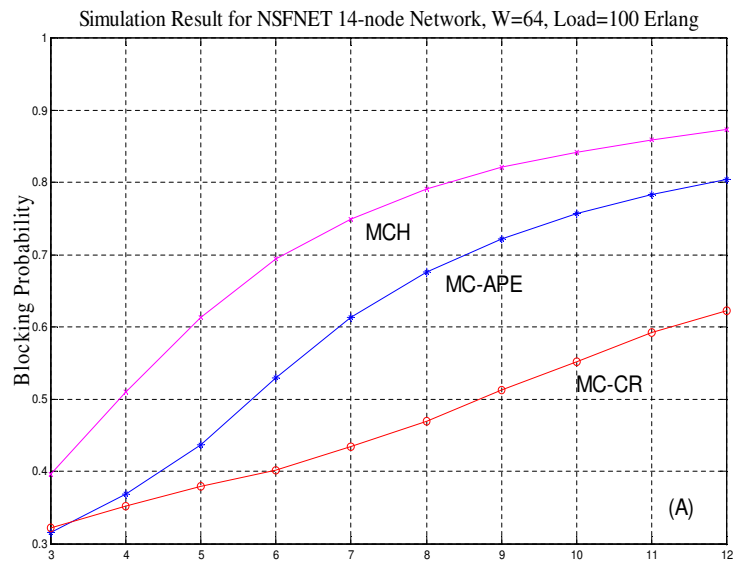


Figure 2.17: Blocking performance vs. multicast group size

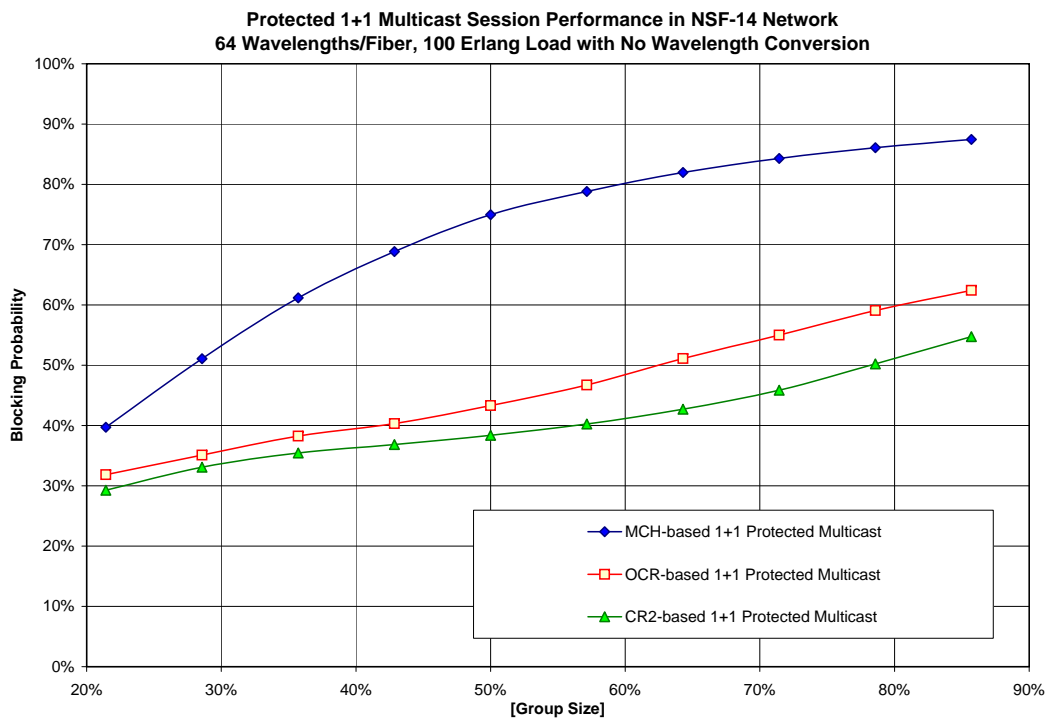


Figure 2.18: 1+1 Multicast Protection Heuristic Performance: Blocking vs. Network Load

Chapter 3

Lightpath-Based Optical Groupcast

In this thesis we investigate the performance of mixed traffic in WDM-based wavelength-routed optical networks. The mixed traffic can be further categorized into two major categories in particular: multicast traffic and groupcast traffic in the presence of unicast traffic. Chapter 2 described the provisioning of multicast traffic. This chapter deals specifically with the case of planning and provisioning of groupcast traffic. Performance of groupcast algorithms is evaluated in the presence of mixed (unicast and groupcast) traffic. The chapter starts with a description of node architectures that can be used for the provisioning of groupcast sessions that are lightpath-based.

3.1 Optical Cross-Connect Switch Architectures

Optical cross connects (OXC) are classified into two major categories: opaque and transparent, based on the characteristics of the optical signal flow through the switch. When an optical channel traversing the OXC transients through optical-electrical-

optical (O/E/O) conversion stages, it is referred to as an opaque switch. If the signal remains always in the optical domain when it passes through the OXC, then it is considered a transparent switch. The switching fabric inside an OXC may either be electronic or optical as shown in Figure 3.1. Opaque OXCs are either based on electrical switching technology or on optical switch fabrics surrounded by transponders/transceivers (providing optical to electrical to optical conversions), imposing the requirement of expensive optoelectronic interfaces. In OXCs using electrical switching, sub-wavelength switching granularities can be supported providing grooming capabilities for more efficient bandwidth utilization. The opaque OXCs also offer inherent regeneration, wavelength conversion, and bit-level monitoring. In transparent OXCs the incoming signals are routed through an optical switch fabric without the requirement of optoelectronic conversions offering transparency to a variety of bit rates and protocols. The switching granularity may vary and support switching at the fiber, the wavelength band, or the wavelength channel level. In this thesis we consider only transparent OXCs.

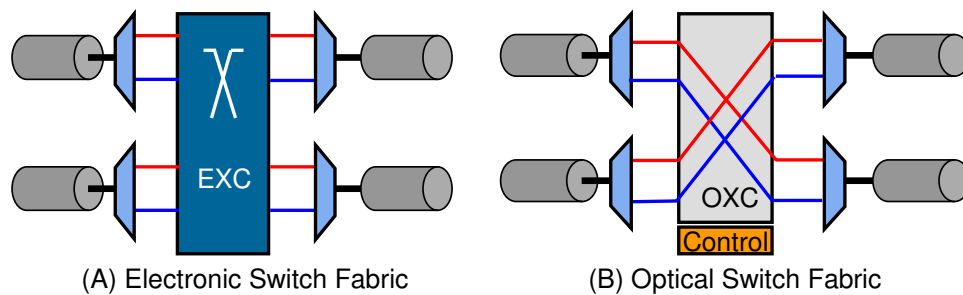


Figure 3.1: OXC architectures supporting (A) Opaque and (B) Transparent switching

A number of transparent switching technologies, some more mature than others, are available in the market today. Various physical layer factors such as crosstalk, amplifier noise, polarisation dependent loss (PDL), dispersion effects and filtering effects limit the OXC performance. A summary of major technologies and their performance characteristics is presented in Table 3.1.

There are many transparent OXC architectures proposed in the literatures [37, 38, 15]. From Table 3.1 it is obvious that most switching technologies have very limited port counts except for the 3-D MEMS technology. The number of ports per node is given by $[(\text{no. of fibers}) \times (\text{no. of wavelengths}) + (\text{no. of add/drop ports})]$. For example, an OXC with four fiber pairs each carrying 160 wavelengths and supporting 50% add/drop channels would require an 1920×1920 optical switch fabric that cannot be implemented by any single stage design. Most large OXC switches are put together using multistage crossbar-like designs such as the Clos design [1, 38, 15] using a switch fabric for each wavelength, as shown in Figure 3.2.

When the optical signal enters the switch on one wavelength and is switched out on another wavelength, the OXC is said to be capable of wavelength conversion. In general wavelength conversion has many benefits like low call-blocking probability, high call admission probability, protection and restoration support (for 1:N lightpaths) and greater flexibility in terms of lightpath provisioning. We shall investigate the effect of wavelength-conversion on groupcast sessions when evaluating the groupcast performance on OXC-based all-optical networks.

| Technology | Loss (dB) | Crosstalk (dB) | PDL (dB) | Switching Time | Max Switch Dimension | Maturity |
|------------------------|------------------|-----------------------|-----------------|-----------------------|-----------------------------|-----------------|
| 2-D MEMS | 5 | 55 | 0.2-0.5 | < 10ms | 32×32 | High |
| 3-D MEMS | 7 | 55 | 0.5 | < 10ms | $4,000 \times 4,000$ | Low |
| Bulk Mechanical | 3 | 55 | 0.2 | < 10ms | 16×16 | High |
| Liquid Crystal | 1 | 35 | 0.1 | < 10ms | 2×2 | High |
| Electro-optical | 8 | 35 | 0.5 | < 10ns | 16×16 | Low |
| Thermo-optical | 8 | 35 | 0.5 | < 10ms | 16×16 | High |
| Holographic | 1 | 35 | Low | < 10ns | 64×64 | Low |
| SOA | Gain | 35 | Low | < 10ns | 4×4 | Low |
| Bubble-based waveguide | 3 | 55 | 0.2 | < 10ms | 16×16 | High |

Table 3.1: Summary of optical switching technologies [1]

Optical cross connects (OXCs) are deployed at traffic terminating nodes or at

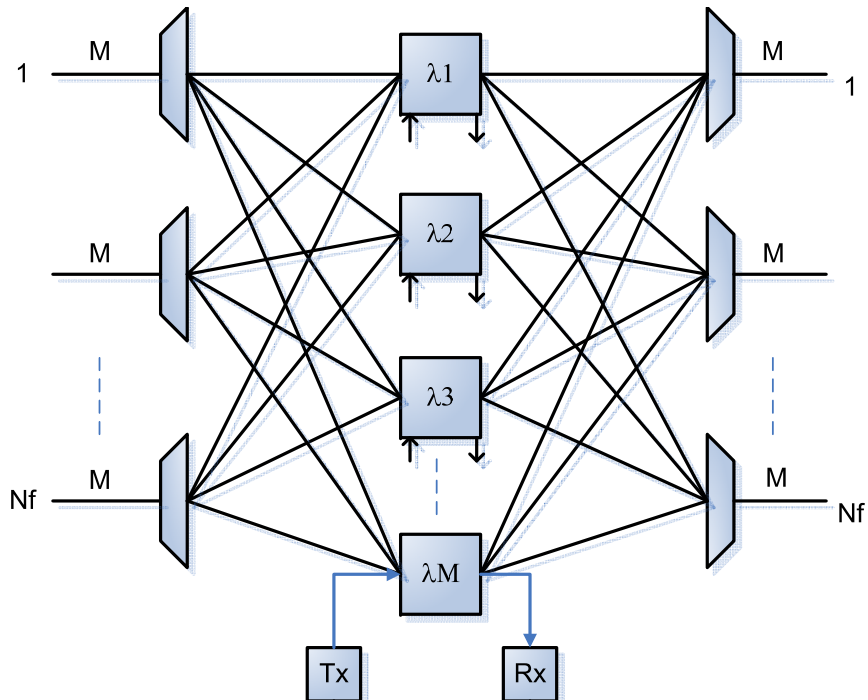


Figure 3.2: Transparent optical cross-connect architecture

junctions in the network for pass-through (switched) traffic. The OXCs cross-connect many fiber pairs and also support add and drop of local traffic providing the interface to the path-terminating equipment. The number of optical switches in the network is fixed and determined by the physical locations where the traffic has to be switched, originated or terminated. The number of wavelengths per fiber on the other hand can be modeled based on traffic demands and routing and wavelength-assignment algorithms.

3.1.1 Node Architectures Supporting Groupcast Sessions

Groupcast sessions tend to tie up a great number of ports as a result of supporting logical multipoint-to-multipoint sessions by physically routing point-to-point lightpaths. We propose a new node architecture that would partition the spectrum into two subgroups: subgroup A for optical channels supporting unicast traffic and sub-

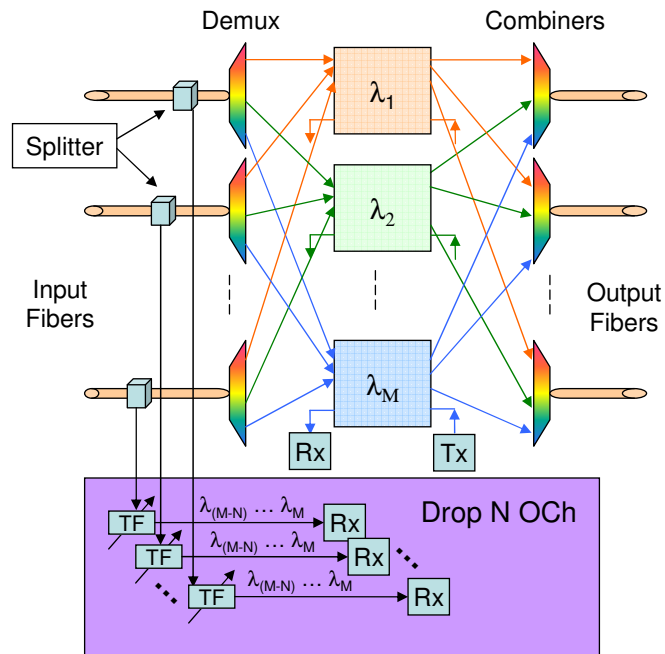


Figure 3.3: Optical cross-connect architecture supporting groupcast

group B for optical channels supporting groupcast traffic. This node architecture can be superimposed over an existing OXC by installing an augmented optical receiver module (and treating the outgoing optical signal through an optical amplifier). In Figure 3.3 the incoming fibers are tapped using optical splitters and a range of N fixed wavelengths are dropped using tunable filters thus allowing the subgroup B channels to support drop-and-continue signal path for groupcast sessions.

3.2 Groupcast Routing and Wavelength Assignment (GC-RWA)

Groupcast is a multipoint-to-multipoint communication that can be established using point-to-point or point-to-multipoint communications among all group members. In this chapter we investigate the performance of groupcast sessions in optical networks supporting unicast or point-to-point *lightpath* sessions only. Group communications are usually implemented by routing protocols such as IP multicast at the Layer 3 or by link-layer protocols such as Ethernet at the Layer 2 of the Open System Interconnection (OSI) reference model. With the advent of the wavelength-routed optical network, groupcasting at the Physical Layer (Layer 1 of the OSI reference model) has emerged as an important research topic.

Reference [39] includes similar work for unicast-based multicast trees. The USCH1 algorithm is introduced that uses n -wavelengths from source to n -destinations (group size n) to build a multicast session. Also, the MSCH2 Algorithm [40] initially computes the multicast light-trees for the multicast session using various shortest-path conventional multicast algorithms, and then it tries to assign a single wavelength to it. If no single wavelength is available along the computed tree, the tunnels, which are the branching points where the wavelength continuity constraint is violated are found [40]. At these nodes, the optical connection is terminated and a new optical connection is established over another wavelength.

The groupcast routing and wavelength assignment (GC-RWA) problem in the lightpath-based network can be decoupled into two separate subproblems, namely: groupcast routing subproblem and groupcast wavelength assignment subproblem.

3.2.1 Groupcast Routing

The objective of this subproblem is to find routable paths from source to destination for a set of nodes (group). Once the paths are found, we assign wavelengths to each path. The total cost of a groupcast session is the cumulative sum of all lightpath costs in the session. The routing is non-trivial for groupcast sessions. Given a group of k nodes $\{n_1, n_2, \dots, n_k\}$, there will be $n \times (n-1)$ routes, namely routes $\{r^1, r^2, \dots, r^{n \times (n-1)}\}$. The routing module investigates possible ways to establish the paths, which can be broadly classified into three types: *fixed routing* (FR), *alternate routing* (AR), and *exhaust routing* (ER) [41, 38].

- In the FR case, we compute the paths from source to all destinations using a specified algorithm (*shortest-path-first (SPF)* algorithm or *minimum-cost-path (MCP)* algorithm) and the same algorithm is used to compute all groupcast sessions. Note that the SPF and MCP produce the same result for hop-based routing where link costs are set to one. However, the performance would be different where link costs are non-uniform. We shall consider fixed routing for both static and dynamic traffic conditions.
- In the AR case, we compute k -alternate paths using the *k-shortest path* algorithm. For each groupcast session, we choose alternate routes and cycle through the route choices. For example, for $k = 2$ and a groupcast session size of 3, the route set is going to be $\{r_1^1, r_2^2, r_1^3, r_2^4, r_1^5, r_2^6\}$ where the route numbers are denoted by the superscript value and the chosen alternate route number is indicated by the subscript value. We shall consider adaptive routing under dynamic traffic conditions only.
- In the ER case, we consider the network state information which is updated dynamically as wavelength channels get assigned. The cost function is scaled

with the amount of wavelengths available per fiber. For example, assume that there are 16 wavelengths per fiber in the network shown in Figure 3.4. If the link between New York and Maryland has 2 wavelengths already assigned, then the link cost (120) would be incremented by $(14/16)$ to 225. For each session, the best route is chosen using a routing algorithm. We shall consider exhaustive routing under dynamic traffic conditions only.

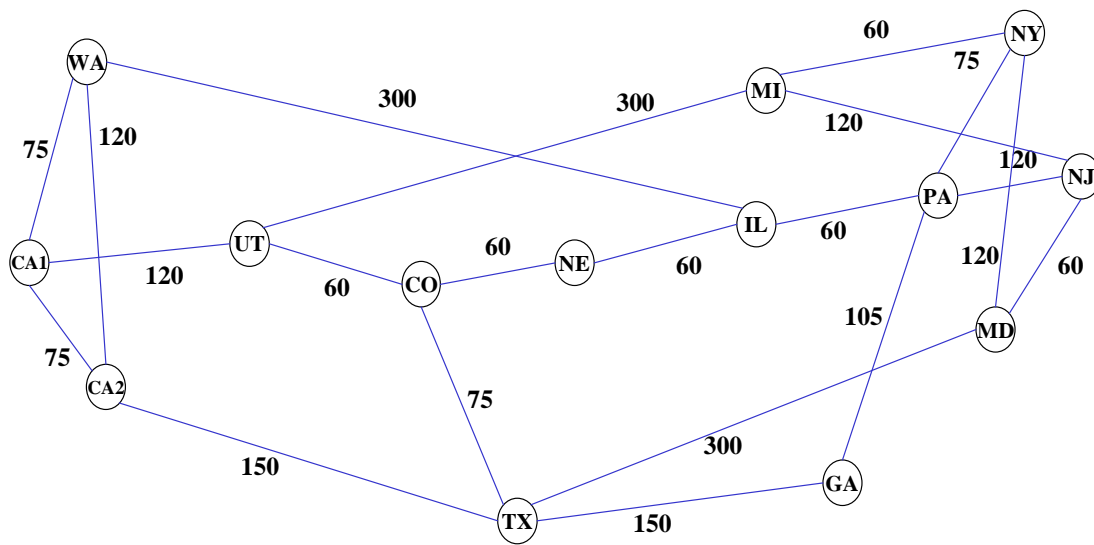


Figure 3.4: NSF 14-node network

In the static GC-RWA case, routing is relatively easier than the dynamic GC-RWA as the network state information is centrally kept and coordinated and no resource contention and signaling has to be dealt with.

3.2.2 Wavelength-Assignment

For the wavelength assignment subproblem, a number of heuristics have been proposed [42, 15, 38, 41]. The idea is to search for a wavelength that could be assigned to a specific route. If the network is not capable of wavelength conversion, then one specific wavelength must be assigned to the entire route. A light-forest is composed

of many light-trees. In a *homogeneous* light-forest, all the light-trees use the same wavelength. In a *heterogeneous* light-forest, different wavelengths are used to serve the groupcast session. If an individual light-tree uses a single wavelength then we call it a *monocolor light-tree* otherwise if each light-tree branch (lightpath from source to destination) uses separate wavelength then we call it a *multicolor light-tree*. If the wavelength continuity constraint is removed then the network can assign any wavelength at any link thus forming a *rainbow light-forest*. Table 3.3 summarizes the wavelength assignment schemes in a light-forest environment. The wavelength assignment algorithms are used to determine the optical channel for each link. We consider the following wavelength-assignment heuristics: *First-Fit* (FF), *Random Assignment* (RA), *Least-Used/SPREAD* and *Most-Used/PACK* as described in Table 3.2.

In the previous section (Section 3.1) we have discussed two basic OXC switch architectures (transparent and opaque) with and without wavelength conversion capabilities that could support various algorithms for route selection and wavelength assignment for optical groupcast sessions. Since we are only considering all-optical or transparent networks, groupcast routing and wavelength assignment (GC-RWA) will be investigated in following two contexts: (a) In the first context, we will determine the wavelength requirements to support groupcast of a certain group size in a given network; (b) In the second context, various GC-RWA algorithms will be considered with the objective of minimizing the blocking probability of the offered traffic load.

The following examples will illustrate the steps involved in establishing a groupcast session in a transparent wavelength-routed mesh networks.

| Wavelength Assignment Algorithms | |
|---|---|
| First-fit | Given a lightpath represented by the link set array $L = \{l_1, l_2, \dots, l_m\}$, ($m = L $), we find a free wavelength λ_j , ($1 \leq j \leq W$) to serve either lightpath, light-tree or light-forest, where W represents the total wavelengths per fiber. The wavelength channels are sequentially arranged in the ascending order prior to each wavelength search begins. |
| Random-fit | Given a lightpath $L = \{l_1, l_2, \dots, l_m\}$, ($m = L $), we find a free wavelength from W wavelengths to serve either lightpath, light-tree or light-forest. We search sequentially in the wavelength table, however the wavelength channels are randomly arranged to introduce disorder prior to each wavelength search. |
| SPREAD | Given a lightpath $L = \{l_1, l_2, \dots, l_m\}$, ($m = L $), we find a free wavelength from W wavelengths to serve either lightpath, light-tree or light-forest. The wavelength channels are arranged by their priority levels. The total priority is set to the total number of links in the network. When a wavelength is assigned to a link the priority level is decreased by one unit. The wavelengths are rearranged based on priority level in the highest to the lowest order prior to each wavelength search (load-balancing approach in terms of wavelength usage). |
| PACK | Given a lightpath represented by the link set array $L = \{l_1, l_2, \dots, l_m\}$, ($m = L $), we find a free wavelength W to serve either lightpath, light-tree or light-forest. The wavelength channels that are arranged by their priority levels. The total priority is set to the total number of links in the network. When a wavelength is assigned to a link the priority level is decreased by one unit. The wavelengths are rearranged based on priority level in the lowest to the highest order prior to each wavelength search (maximum reuse - pack as many connections as possible on the most used wavelength before moving the the next most-used wavelength, and so on). |

Table 3.2: Wavelength selection algorithms for lightpath-based groupcast connections

3.2.3 Example I: GC-RWA on OXC Network Supporting No Wavelength Conversion

Consider a network with 6 nodes, 9 links, and 4 wavelengths per link as shown in Figure 3.5. Assume that two groupcast session requests arrive for the node set $\{A, C, D\}$

| Light-Forest Model | Same Color Branch | Same Color Tree | Same Color Forest |
|------------------------|-------------------|-----------------|-------------------|
| Homogeneous | Yes | Yes | Yes |
| Heterogeneous | Yes | Yes | No |
| Monocolor light-trees | Yes | Yes | No |
| Multicolor light-trees | Yes | No | No |
| Rainbow | No | No | No |

Table 3.3: Various light-forests with wavelength continuity constraint

and $\{B,E,F\}$. In order to simplify the example, we assume that all nodes are managed by the same service provider and that the control plane of the OXCs are connected to a centralized management system.

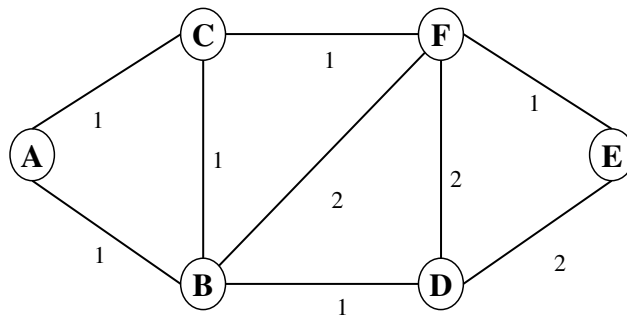


Figure 3.5: A Network with 6 nodes and 9 links

The following steps are taken in order to establish the groupcast session:

- **Step 1:** Determine the paths using shortest-path first algorithm¹ for connections (A,C) , (A,D) , (C,A) , (C,D) , (D,A) , and (D,C) in session 1 and the results are summarized in Table 3.4 .
- **Step 2:** Assign wavelengths to the routes found in **Step 1**. Note that we have a wavelength continuity constraint as wavelength conversion is not supported by the OXCs. The wavelengths assigned are shown in Table 3.4. Often, it is useful to list all the links as shown in Table 3.5 to facilitate wavelength assignments.

¹Any heuristic could have been used to determine the paths. However the shortest-path first is a good choice as OSPF uses the same algorithm.

| Session | (Source, Destination) | Route | Wavelength |
|---------|-----------------------|-----------------------------------|-------------|
| S1 | (A,C) | A \rightarrow C | λ_1 |
| | (A,D) | A \rightarrow B \rightarrow D | λ_1 |
| | (C,A) | C \rightarrow A | λ_1 |
| | (C,D) | C \rightarrow B \rightarrow D | λ_2 |
| | (D,A) | D \rightarrow B \rightarrow A | λ_1 |
| | (D,C) | D \rightarrow B \rightarrow C | λ_2 |
| S2 | (B,E) | B \rightarrow D \rightarrow E | λ_3 |
| | (B,F) | B \rightarrow F | λ_1 |
| | (E,B) | E \rightarrow D \rightarrow B | λ_3 |
| | (E,F) | E \rightarrow F | λ_1 |
| | (F,B) | F \rightarrow B | λ_1 |
| | (F,E) | F \rightarrow E | λ_1 |

Table 3.4: Session, routes and wavelength assignment

Table 3.4 is updated with route numbers and each route is represented as a set of links in Table 3.6. A two-dimensional table for link number and wavelength channel as shown in Table 3.7 is easier to maintain than complex details on path and wavelength assignment per connection. If a connection cannot be served by any available wavelength the session is blocked and it is marked as a failure.

- **Step 3:** Choose the following session and repeat **Step 1** and **Step 2** until we finish serving all the sessions. The routes and wavelength assignments for the second session {B,E,F} are shown in Table 3.5. Also, the session is recorded as a link-wavelength matrix in Table 3.7.

| | A | B | C | D | E | F |
|---|---|----|----|----|----|----|
| A | - | 1 | 2 | - | - | - |
| B | 3 | - | 4 | 5 | - | 6 |
| C | 7 | 8 | - | - | - | 9 |
| D | - | 10 | - | - | 11 | 12 |
| E | - | - | - | 13 | - | 14 |
| F | - | 15 | 16 | 17 | 18 | - |

Table 3.5: Assignment of link numbers for Figure 3.5

| Session | (Source, Destination) | Route {Link Set} | Wavelength |
|---------|-----------------------|------------------|-------------|
| S1 | (A,C) | {2} | λ_1 |
| | (A,D) | {1, 5} | λ_1 |
| | (C,A) | {7} | λ_1 |
| | (C,D) | {8, 5} | λ_2 |
| | (D,A) | {10, 3} | λ_1 |
| | (D,C) | {10, 4} | λ_2 |
| S2 | (B,E) | {5, 11} | λ_3 |
| | (B,F) | {6} | λ_1 |
| | (E,B) | {13, 10} | λ_3 |
| | (E,F) | {14} | λ_1 |
| | (F,B) | {15} | λ_1 |
| | (F,E) | {18} | λ_1 |

Table 3.6: Session, routes and wavelength assignment for Example I

Link No.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|-------------|----|----|----|----|----|----|----|----|---|----|----|----|----|----|----|----|----|----|
| λ_1 | S1 | S1 | S1 | | S1 | S2 | S1 | S1 | | S1 | | | | S2 | S2 | | | S2 |
| λ_2 | | | | S1 | S1 | | | S1 | | S1 | | | | | | | | |
| λ_3 | | | | | S2 | | | | | S2 | S2 | | S2 | | | | | |
| λ_4 | | | | | | | | | | | | | | | | | | |

Table 3.7: Wavelength assignment chart for Example I in Figure 3.5

3.2.4 Example II: GC-RWA on OXC Networks Supporting Wavelength Conversion

The major advantage of wavelength conversion is the dramatic increase in call admission probability. This example reconsiders the network shown in Figure 3.5 with 4 wavelengths per fiber. To illustrate the comparison, we take the same two groupcast sessions for the node set $\{A,C,D\}$ and $\{B,E,F\}$ as in Example I. The assumption is again that all nodes are managed by the same service provider and the control planes of the OXCs are connected to a centralized management system. The following steps are taken in order to establish the groupcast sessions when wavelength conversion is present:

- **Step 1:** Determine the paths using shortest-path first algorithm for connections (A,C), (A,D), (C,A), (C,D), (D,A), and (D,C) in session 1 and the results are summarized in Table 3.8.
- **Step 2:** Assign wavelengths to each path using the first available wavelength² and the assignments are shown in Table 3.8.
- **Step 3:** Choose the next session and repeat **Step 1** and **Step 2** until all the sessions are served. The routes and wavelength assignments for the second session {B,E,F} are shown in Table 3.8. Also, all the sessions are recorded as a link-wavelength matrix in Table 3.9 that better captures the wavelength utilization status. Note that wavelength one is more utilized in Example II than in Example I (see Table 3.7).

| Session | (Source, Destination) | Route {Link Set} | Wavelengths |
|---------|-----------------------|------------------|----------------------------|
| S1 | (A,C) | {2} | { λ_1 } |
| | (A,D) | {1, 5} | { λ_1, λ_1 } |
| | (C,A) | {7} | { λ_1 } |
| | (C,D) | {8, 5} | { λ_1, λ_2 } |
| | (D,A) | {10, 3} | { λ_1, λ_1 } |
| | (D,C) | {10, 4} | { λ_2, λ_1 } |
| S2 | (B,E) | {5, 11} | { λ_3, λ_1 } |
| | (B,F) | {6} | { λ_1 } |
| | (E,B) | {13, 10} | { λ_1, λ_3 } |
| | (E,F) | {14} | { λ_1 } |
| | (F,B) | {15} | { λ_1 } |
| | (F,E) | {18} | { λ_1 } |

Table 3.8: Session, routes and wavelength assignment for Example II

²There are many wavelength assignment (WA) algorithms like First-fit, Random-fit, etc. We are only considering the First-fit scheme in this example. See section 3.2.2 for detailed discussion on WA algorithms.

Link No.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|-------------|----|----|----|----|----|----|----|----|---|----|----|----|----|----|----|----|----|----|
| λ_1 | S1 | S1 | S1 | S1 | S1 | S2 | S1 | S1 | | S1 | S2 | | S2 | S2 | S2 | | | S2 |
| λ_2 | | | | | S1 | | | | | S1 | | | | | | | | |
| λ_3 | | | | | S2 | | | | | S2 | | | | | | | | |
| λ_4 | | | | | | | | | | | | | | | | | | |

Table 3.9: Wavelength assignment chart for Example II in Figure 3.5

3.2.5 Optical Network Architecture

Today's state of the art optical components can pack hundreds of wavelengths into a single fiber [43] and each channel can carry more than 40Gbps [37]. DWDM spectral grids standardized by ITU-T G.694.1 [44] specify channel spacing varying from 12.5 GHz to 100 GHz. With so much raw bandwidth capacity extractable from a single fiber strand, one might be tricked into thinking a DWDM point-to-point system is sufficient for wide area networks (WAN) for the foreseeable future. However, as it turns out this is not the case as the DWDM systems are quite expensive and do not scale well with IP core routers [6]. Although the IP-over-DWDM network architecture is widely deployed today (see Figure 3.6 left-hand side), the bandwidth bottleneck nevertheless lies in the electronic switch fabric. The traffic scaling problem can be readily identified by observing the sheer volume of IP traffic needed to be handled by the routers when it terminates hundreds of optical channels. Fortunately, optical communications has experienced tremendous advancement in recent years. The second generation optical communication systems enabled the multiplexing of multiple wavelengths, thus multiple optical channels could now share the same fiber. Finally, *wavelength-routing* introduced during the third generation of optical communication alleviated the electronic bottlenecks experienced by for point-to-point fiber-optic systems. With the introduction of optical cross-connects (OXC) in the optical transport network (OTN), IP-over-OTN architecture (see Figure 3.6, right-hand side) resulted

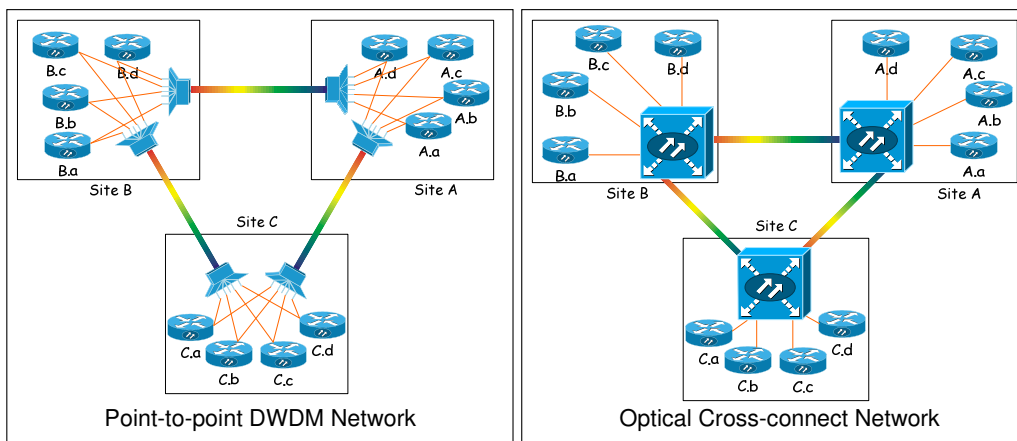


Figure 3.6: IP-over-DWDM network and IP-over-OTN network

in an attractive proposition for bandwidth scalability as well as economically competitive optical switching platforms. Interestingly, the IP-OXC per port cost (\$/port) is much lower than the IP-DWDM port cost excluding the system chassis cost [6].

To support optical groupcasting, we need a network that allows any-to-any node connectivity. A non-blocking N -node optical network that supports all possible traffic configurations requires $\sqrt{N/e}$ wavelengths to support all permutations of connections from input to output ports without multicasting [42]. Building a large non-blocking network can thus be very cost-prohibitive.

We assume that optical networks are built using optical switching elements (OTN architectures) as it is impossible to deploy fiber directly connecting every node to every other node. Given a wavelength-routed network with W wavelengths and N nodes, the minimum number of 2×2 optical cross-connect (OXC) switches required is given by the following equation [42]:

$$\log_2 OXC(N, W) \geq N \log_2 N - 2N \log_2 W - 1.44N \quad (3.1)$$

The number of OXC required per network is directly related to the number of concurrent sessions a network plans to support as suggested by Equation (3.1). Increasing the number of wavelengths per fiber helps reduce the number of OXCs required per network due to the *wavelength-routing* gain.

For a network with M transmitters, M receivers and W wavelengths, we can define an OXC reduction parameter \mathfrak{R} to be:

$$\mathfrak{R} = \frac{\log_2 OXC(M, W)}{\log_2 OXC(M, 1)} \quad (3.2)$$

To realize 50% of the \mathfrak{R} value or to reduce the OXC switches in the network by half we must have $M^{1/4}$ wavelengths.

An OXC-based all-optical network is designed to support lightpaths meeting a certain traffic demand matrix T . Traffic matrix T represents the connectivity requirements of the users to establish unidirectional sessions. So, for M users, each having one transmitter and one receiver ($M_t = M_r = M$) we will have a $T_{M \times M}$ traffic matrix. In a groupcast scenario, we have one transmitter establishing sessions with multiple receivers where $M_t \neq M_r$. We define a groupcast session (m, n) to be an ordered pair between a transmitter m and a receiver n . The traffic demand matrix T is a set of concurrent sessions or allowable simultaneous traffic states. The concept of *Permutation Traffic Set* is defined in [42] as the traffic set containing all one-to-one complete matching of transmitters and receivers. The *Permutation Traffic Set* for M users has $M!$ traffic elements. A subset of *Permutation Traffic Set* is the ρ -*Permutation Traffic Set* ($0 < \rho < 1$) defined in [42] to be the set of all one-to-one matchings of transmitters and receivers with at most ρM active sessions. Based on the definition, for ρ -permutation routing on M users we need at least $\sqrt{\rho M/e}$ wavelengths to avoid session blocking [42].

For a groupcast session of size N , $N(N - 1)$ lightpaths are required. A network supporting multiple groupcast sessions translates into a large M value for the traffic matrix T . Increasing the size of M directly impacts the number of ports in an OXC switch as it could be observed by substituting N (number of point-to-point nodes) with M (number of one-to-one sessions) in Equation 3.1. For optical DWDM networks, Equation 3.1 also shows that *wavelength-routing* and *wavelength-conversion* does not change the margin of switch size growth. However, Equation 3.2 shows that for large number of sessions, it is possible to reduce the number of switches by some factor if we use a large number of wavelengths. For example, for 10^8 sessions and 10^3 wavelengths, the number of switches could be reduced by a factor of 5 (Equation 3.2). Therefore, even with large groupcast sessions, wavelength-routing will have an impact on the OXC switch size by reducing the number of wavelengths required per fiber.

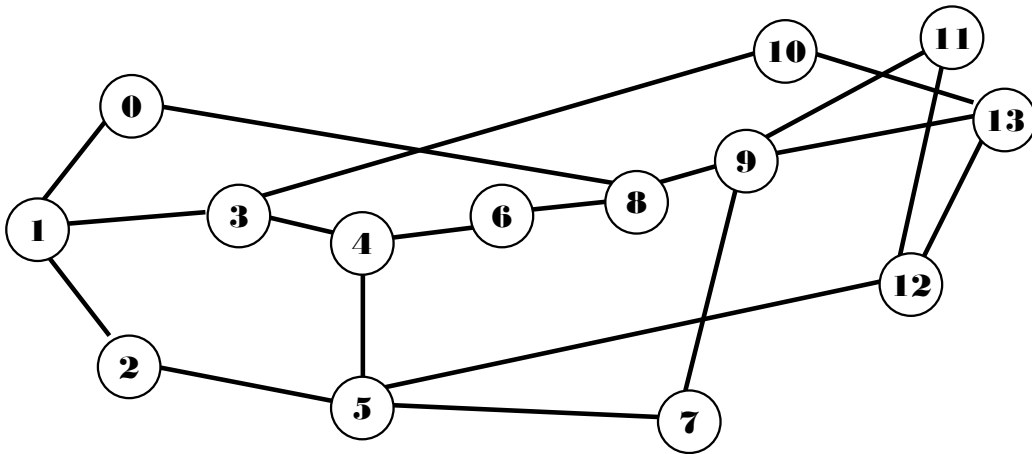


Figure 3.7: NSF 14 node network

For the evaluation of the algorithms developed during this thesis work (Chapters 3 and 4) and for comparison in network performance, we used some reference topologies for all-optical transport networks. After considering various optical network topologies [45, 46, 47, 48, 49], two reference topologies were chosen based on

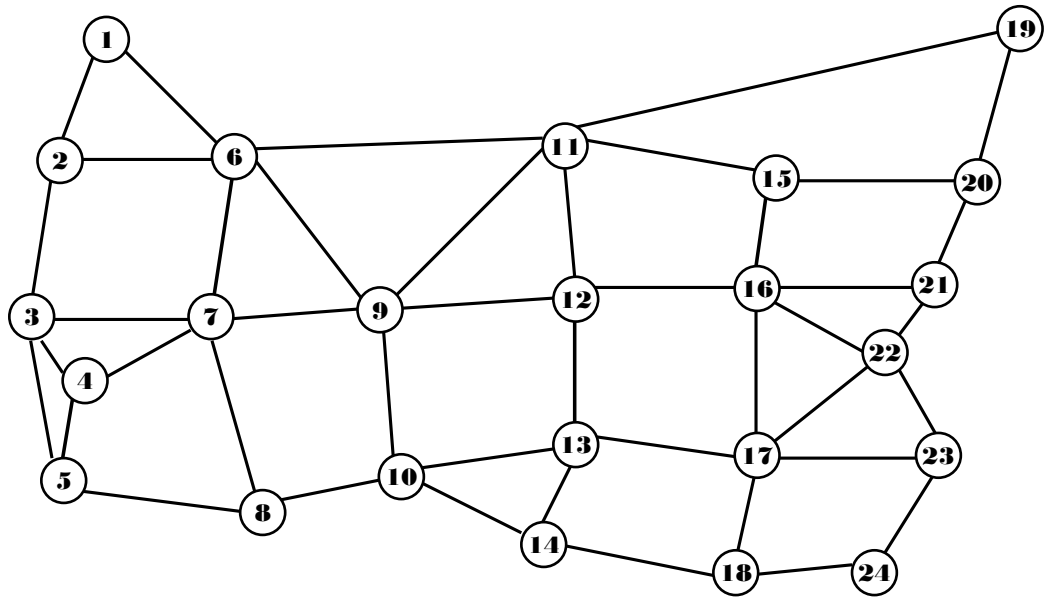


Figure 3.8: AT&T 24 node IP network

network size and complexities as shown in the Table 3.10 below. The NSF network topology is widely used in optical WDM network research papers as an illustrative wide-area backbone network topology as shown in Figure 3.7³. The second reference network consists of 24 nodes and 43 bidirectional links based on AT&T IP backbone topology [51]. This network is referenced in a number of optical multicast research literatures [52, 53] and is shown in Figure 3.8.

| Network Name | NSF14 | AT&T |
|----------------------|---------|---------|
| Total Nodes | 14 | 24 |
| Total Links | 21 | 43 |
| Average Nodal Degree | 3 | 3.5833 |
| Average Hop Distance | 2.14286 | 2.99275 |

Table 3.10: The reference optical networks used in Chapters 3 and 4

³In 1984, NSF began designing a high-speed successor to the ARPANET that would be open to all university research groups. It was the first WAN offering 56 kbps of bandwidth over leased lines. Later NSF14 was upgraded to T1 (1.5 Mbps) in 1991 and finally to T3 (45 Mbps) in 1992 before it was shutdown in April of 1995 [50].

3.2.6 GC-RWA Performance

In this section we investigate the performance of GC-RWA algorithms under static or pre-defined traffic conditions. Although we have described three routing algorithms and four wavelength assignment algorithms in Section 3.2, we selected a default routing and a default wavelength assignment algorithm for finding the required lightforests. The First-Fit (FF) algorithm is chosen as default wavelength assignment algorithm while the minimum-cost path is chosen as the default routing algorithm for analyzing the GC-RWA performance under static traffic planning scenarios. The remaining routing and wavelength assignment techniques described in Section 3.2 will be considered in Section 3.4 for the evaluation of the GC-RWA performance under dynamic traffic conditions.

In order to support groupcasting in an OXC-based network, first a group of nodes is selected. Then the network builds a set of point-to-point lightpaths from a source node from the group to the remaining group nodes' receivers. The process is repeated until all the nodes from the group are served as the source node. Once the routes are computed by the routing module, the wavelength assignment module secures a wavelength for each lightpath computed. The groupcast routing and wavelength assignment (GC-RWA) assumes the following inputs:

- The topology of the network depicted by a directed and weighted graph $G = (V, E)$, where V is the set of vertex representing the OXC nodes and E is the set of edges representing the fibers connecting the OXC nodes.
- Each link is assigned a weight to represent the cost (number of hops or equipment operating cost, or delay (in terms of the length of the link in kilometers)) of moving data from one node to the next.
- W is the number of optical channels carried by each fiber.

- The routing algorithm is specified as either *fixed routing (FR)* or *alternate routing (AR)* or *exhaust routing (ER)*. The FR is the default routing unless otherwise specified.
- The wavelength assignment algorithm is specified as either *first-fit (FF)* or *random-fit (RA)* or *least-used (SPREAD)* or *most-used (PACK)*. The FF is the default wavelength-assignment scheme unless otherwise stated.

The basic groupcast algorithm, described in Table 3.11, uses Dijkstra's algorithm to compute shortest paths from a source to all the nodes in the graph. Dijkstra's algorithm has time complexity on the order of $O(n^2)$, where $n = |V|$ is the number of nodes in the graph $G = (V, E)$. For a groupcast of size k , the process is repeated k times, thus increasing the time complexity to $O(kn^2)$.

For the wavelength assignment shown in Table 3.11 we consider *multicolor-light-trees* as described in Table 3.3. In the wavelength assignment sub-routine, W possible wavelengths are searched $|L|$ times where L is the lightpath length. For a given graph $G = (V, E)$, the length of the shortest path must be less than or equal to $|E| - |V|$. Therefore, the first-fit (FF) wavelength-assignment approach has time complexity on the order of $O(W \times (|E| - |V|))$. For the light-forest, the complexity of the FF algorithm increases to $O(W * (|E| - |V|)(k - 1)^2)$, where k is the group size. We know that for a set of n nodes, $n(n - 1)$ wavelengths are required to establish a groupcast session. We can therefore conclude that for large groupcast sessions the time complexity for the FF algorithm becomes $O(n^2(n - 1)^3)$ or $O(n^5)$.

We use weighted network topologies when evaluating the GC-RWA algorithms. The average number of links required to support multicast and groupcast sessions are compared in Figure 3.9 for the NSFNET node network [Figure 3.4]. A set of 64 random sessions were generated for a given group size for both multicast and

| Minimum Cost GC-RWA (MC_GCRWA) Algorithm | |
|---|--|
| Step 1 | Given a network $G = (V, E)$ and a group of nodes $\{v_1, v_2, \dots, v_n\}$, choose v_1 and compute the minimum cost path from v_1 to all other nodes $\{v_2, \dots, v_n\}$. |
| Step 2 | Store the routes found in Step 1 into $\{R_2^1, R_3^1, \dots, R_n^1\}$. |
| Step 3 | Repeat Step 1 and Step 2 for all the members of the group $\{v_1, v_2, \dots, v_n\}$ and construct the light-forest matrix: |
| | $F = \begin{pmatrix} R_2^1 & R_3^1 & \dots & R_n^1 \\ R_1^2 & R_3^2 & \dots & R_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ R_1^n & R_2^n & \dots & R_{n-1}^n \end{pmatrix}$ |
| | The light-forest F consists of n light-trees, where light-tree T^k is represented by the row-vector $T^k = [R_1^k R_2^k \dots R_n^k]_{(n-1) \times 1}$. Note that row-vector T^k has $(n-1)$ routes as no loop-back route is allowed. |
| Step 4 | Return light-forest F to wavelength assignment function. |
| Step 5 | Choose the route $F(1, 1)$ represented by the link set array $L_1 = \{l_1, l_2, \dots, l_m\}$, ($m = L_1 $) and find a free wavelength λ_j , ($1 \leq j \leq W$), where W represents the total wavelengths per fiber in the network. |
| Step 6 | Mark the wavelength λ_j chosen in Step 5 as unavailable in all the links from the link set L_i . |
| Step 7 | Repeat Step 5 and Step 6 for all the elements of F or light-paths represented by L_i where $(1 \leq i \leq (n-1)^2)$. |

Table 3.11: Lightpath-based groupcast RWA algorithm

groupcast connections. The MC_GCRWA algorithm [Table 3.11] is used for groupcast sessions while for multicast the MC_MCRWA algorithm as shown in Table 3.12 is used for multicast session establishment in Figure 3.9. As it could be easily noted from Figure 3.9, the network cost exponentially grows for groupcasting as the groupcast size increases.

The network cost is more reflective of the number of links utilized per session. To see the groupcast behavior in large networks we compare the groupcast performance in the 14-node NSF network [Figure 3.4] to the 24-node AT&T IP backbone network [Figure 3.10]. As the link costs are unique to each network, we use the average number

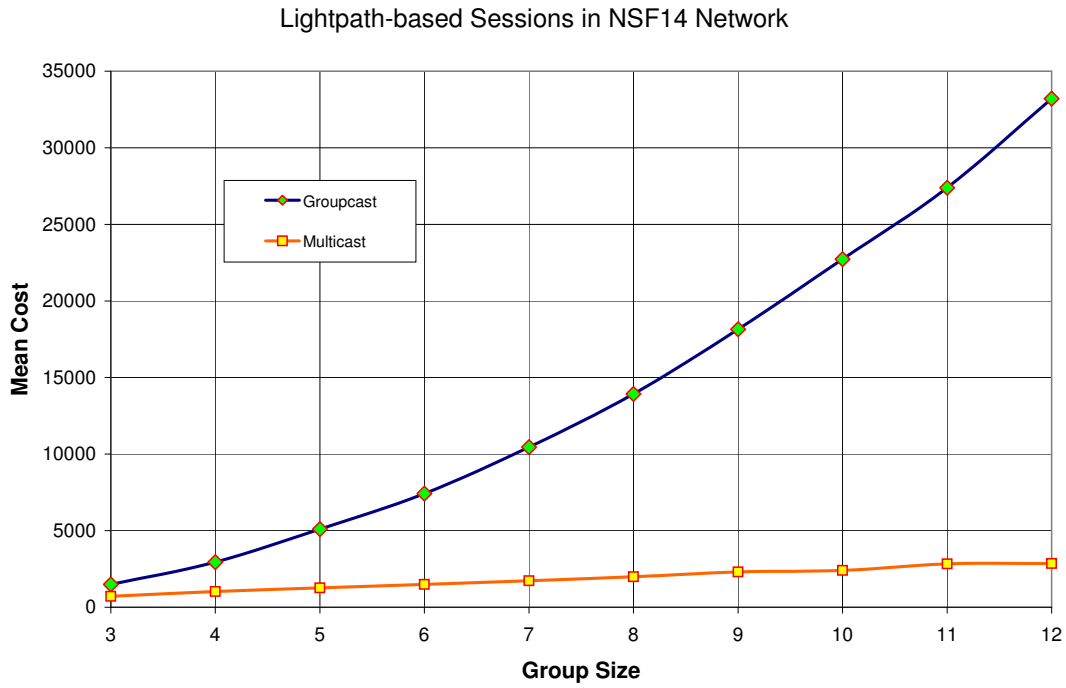


Figure 3.9: Groupcast vs. multicast performance in wavelength-routed networks

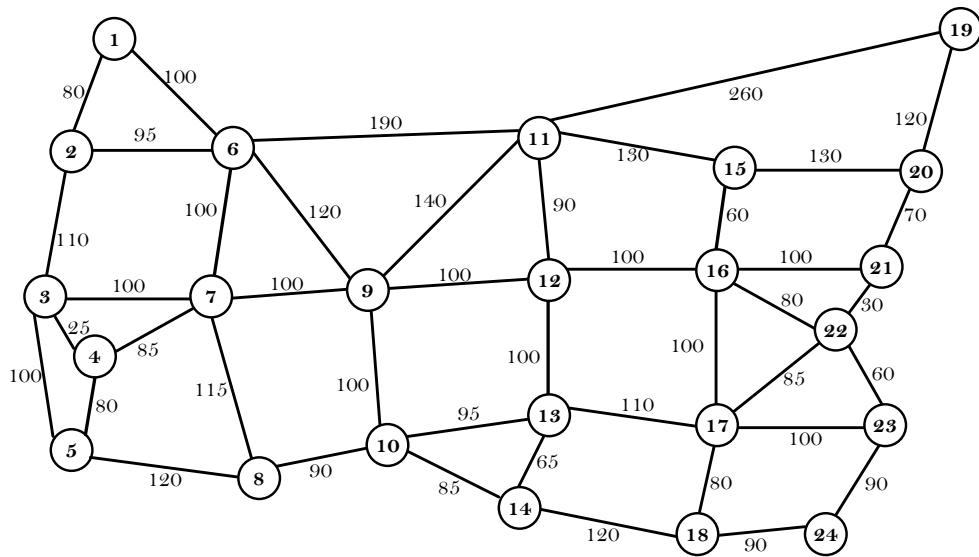


Figure 3.10: AT&T 24 node IP network

of links utilized per session for comparison as shown in Figure 3.11. Although each session requires at least one wavelength to be served, it is important to understand

| Minimum Cost Multicast RWA (MC_MCRWA) Algorithm | |
|--|---|
| Step 1 | Given a network $G = (V, E)$ and a group of nodes $\{v_1, v_2, \dots, v_n\}$, choose v_1 and compute minimum cost path from v_1 to all other nodes $\{v_2, \dots, v_n\}$. |
| Step 2 | Store the routes found in Step 1 into $\{R_2, R_3, \dots, R_n\}$. |
| Step 3 | Choose the route $R(1)$ represented by the link set array $L_1 = \{l_1, l_2, \dots, l_m\}$, ($m = L_1 $) and find a free wavelength λ_j , ($1 \leq j \leq W$), where W represents the total wavelengths per fiber in the network. |
| Step 4 | Mark the λ_j chosen in Step 3 as unavailable in all the links from the link set L_i . |
| Step 5 | Repeat Step 3 and Step 4 for all the elements of R or light-paths represented by L_i where ($1 \leq i \leq (n - 1)$). |

Table 3.12: Lightpath-based multicast RWA algorithm

the number of wavelengths required for a given size of groupcast session. Figure 3.12 shows the average number of wavelengths required for a given size of groupcast in the NSF14 and AT&T networks. We observe that the number of links required in large networks (24 node AT&T) is greater than that of small networks (14 node NSF14) as the session size increases. However, the number of wavelengths required in small networks (21 links NSF14) is greater than that of large networks (43 links AT&T). This is because the number of links required is dependent on the cost function and in large networks the maximum hop distance between nodes that are farthest apart is greater than the maximum hop distance between nodes that are farthest apart in small networks. As for the wavelength resource utilization, it is expected that the smaller nodal-degree network (less routing variation) will utilize the same links to establish a greater number of lightpath sessions irrespective of the network size when compared to a larger nodal-degree network (larger possibility of route variation). Note that the NSF14 14-node network has a nodal degree of 3, while the AT&T 34-node network has a nodal degree of 3.5833. In order to establish the comparison in Figure 3.11 and Figure 3.12, we generated 64 random groups of nodes for each group size and applied the GC-RWA algorithm as described in Table 3.11. For a group size of n , 64 n -tuple

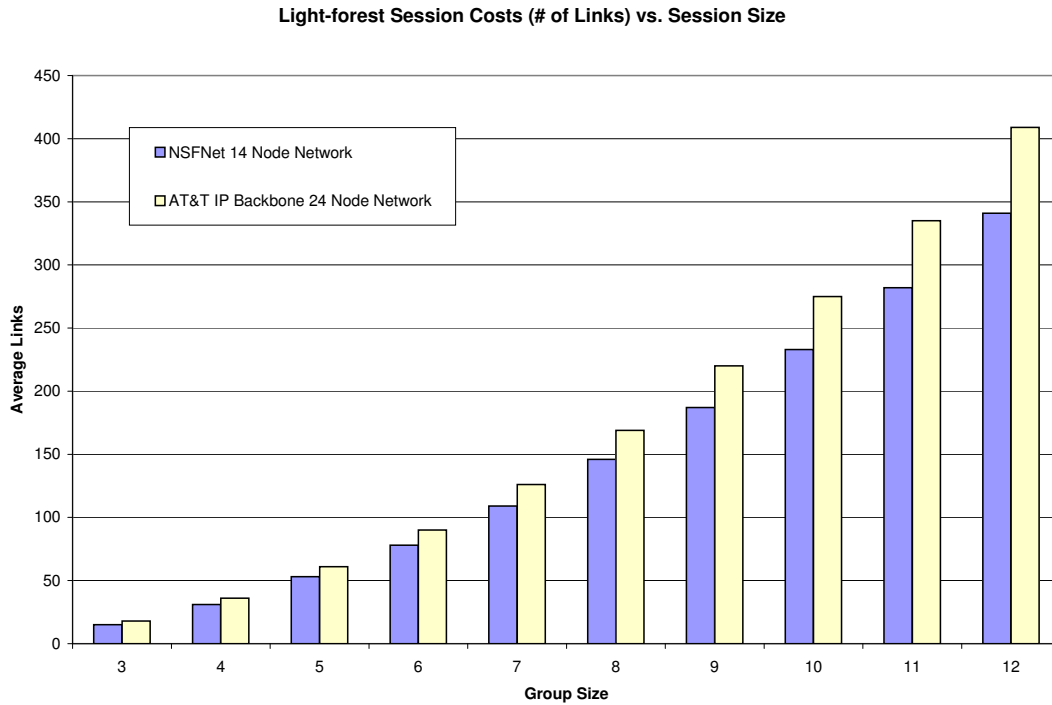


Figure 3.11: Groupcast performance in a 14-node vs. a 24-node network

random groupcast sessions were generated. The summation of all session costs is then divided by 64 to obtain an average session cost for the specified groupcast size.

3.3 Static Groupcasts: Planning and Design of Sessions in OXC Networks

The static groupcast stems from the idea of setting up all-to-all connectivities among a group of nodes for a given fixed or static traffic demand. The dynamic traffic demand has a time dimension attached to the call request that describes its duration. The dynamic groupcast is studied later in Section 3.4. A typical wavelength-routed WDM network can support a groupcast session by activating a number of lightpaths. It is important to understand the design limitations of provisioning lightpath-based static groupcast sessions.

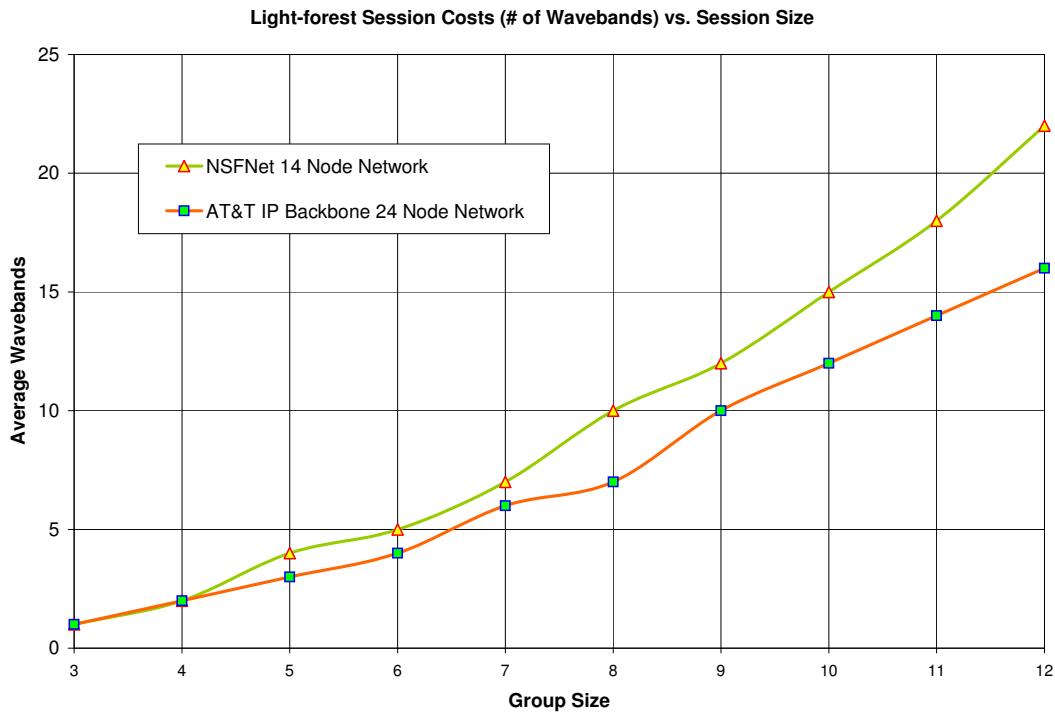


Figure 3.12: Wavelengths required for groupcast in NSF14 vs. AT&T network

The minimum cost routing described in Table 3.11 finds cost-optimized routes which do not take into account the wavelengths required to serve the session. Often, the number of wavelengths is a very expensive network resource and the cost function utilized to determine link weights does not take that into account. We obtain the graph in Figure 3.13 for GC-RWA performance by ignoring the link weights. We apply the shortest-path-first algorithm described in Table 3.13 that only considers the hop counts. It is obvious that the minimum-cost-path algorithm outperforms the shortest-path-first algorithm, but not by a great margin: there is only a 13.5% difference for the largest group size considered. However, when we look at the average number of wavelengths required to serve a given group size, the savings are significant using the shortest-path-first routing algorithm as shown in Figure 3.14. The savings on the number of wavelengths required per fiber are greater for larger group size (48% for the light-forest of size 12). When designing optical networks it is obvious that

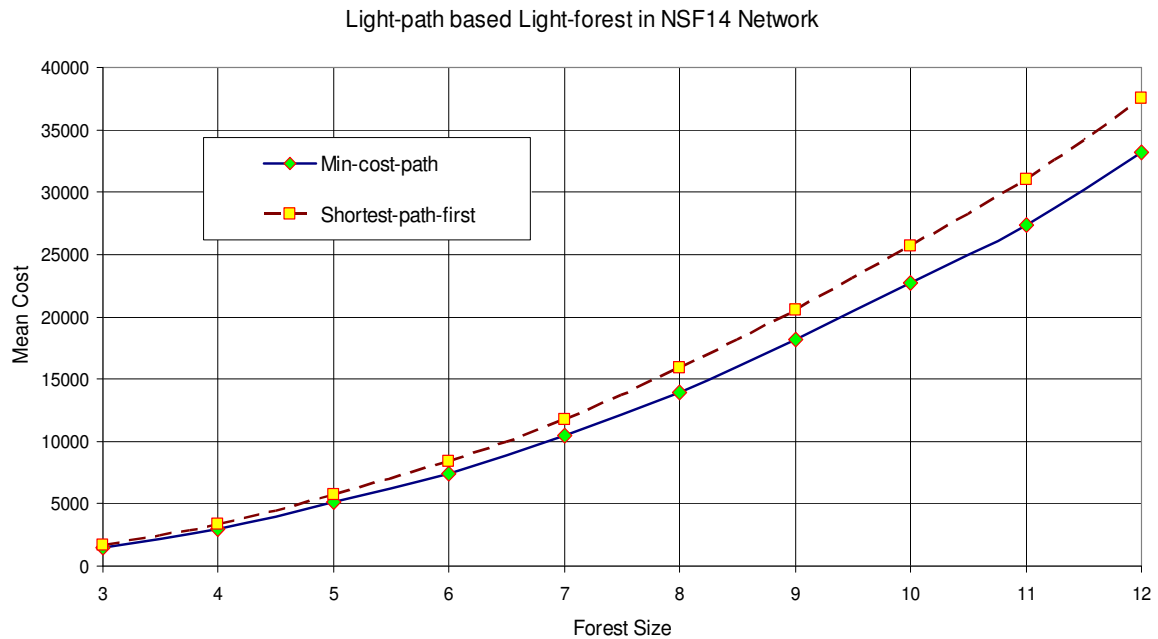


Figure 3.13: Groupcast routing performance for minimum-cost-path vs. shortest-path-first in NSF14

larger DWDM systems are cost-prohibitive and we can optimize the wavelengths-per-fiber requirement, given the traffic characteristics, by routing them on slightly more expensive routes.

| Shortest Path First GC-RWA (SPF_GCRWA) Algorithm | |
|---|--|
| Step 1 | Given a network $G = (V, E)$ and a group of nodes $\{v_1, v_2, \dots, v_n\}$, choose v_1 and compute shortest path from v_1 to all other nodes $\{v_2, \dots, v_n\}$ based on hop counts. |
| Step 2 to Step 7 | Follow the same instructions as stated in Step 2 to Step 7 from Table 3.11. |

Table 3.13: Lightpath-based groupcast RWA algorithm #2

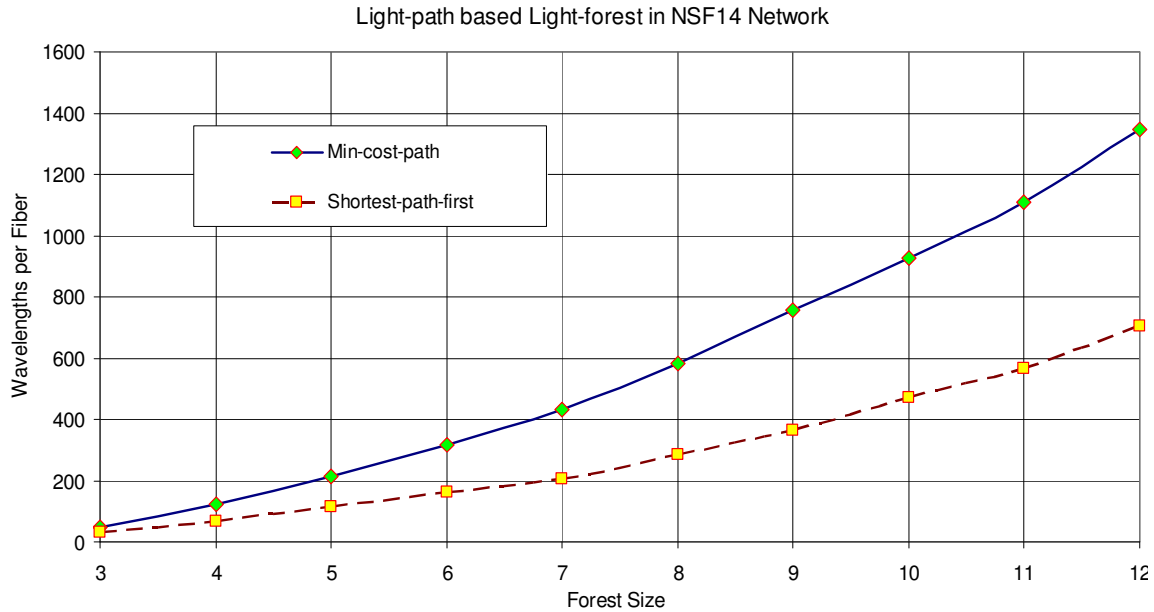


Figure 3.14: Wavelengths per fiber required in NSF14 for minimum-cost-path and shortest-path-first algorithms

3.3.1 Lightpaths: Virtual Topology Design for Groupcast

In this section we examine the static Virtual Topology Design (VTD) problem for groupcast traffic. The VTD problem of setting up a finite set of groupcast sessions at a minimum cost using lightpaths is considered using multicommodity flow and traffic flow on the virtual topology. In the first scenario, we assume that each OXC is equipped with wavelength conversion capabilities. In the subsequent scenario we add the wavelength-continuity constraint. The following parameters are assumed for the VTD problem:

- (i) $G(V, E)$: A physical graph representing the OTN network with a vertex set V and an edge set E . It contains N nodes and L bidirectional links. Each link has an assigned weight, which may correspond to the fiber distance between nodes implicitly representing propagation delay.

- (ii) Φ : A set of Φ groupcast sessions.
- (iii) G_i : The groupcast size for session i , where $3 \leq G \leq |V|$. The following observations are to be noted:
 - (a) A groupcast size of one is not possible. We need more than 1 node to form a group.
 - (b) A groupcast size of two is a special instance of point-to-point, bi-directional communication. However it is not a multipoint-to-multipoint communication session.
 - (c) A groupcast size of $|V|$ is a special case of groupcast called *allcast*, where each node establishes a broadcast session for the rest of the network nodes.
 - (d) In our study we do not consider any of the above cases: either (a) or (b) or (c). Thus, the smallest groupcast set has three members and the largest groupcast set has $|V| - 1$ members.

The objective is to set-up all Φ groupcast sessions on the given OTN while minimizing the cost. The cost of a groupcast session is defined as the sum of all the link weights traversed by the lightpaths used to establish the session.

3.3.2 VTD Formulation-I: OXC with Wavelength Converters

We formulate the VTD problem as an optimization problem, using principles from multicommodity flow for physical routing of lightpaths and traffic flow on the virtual topology using the following notation:

- (i) s^i and d^i denote *source node* and *destination node* for a lightpath in the groupcast session i , $S_i = s^i, d^i$.

- (ii) m and n denote *endpoints of a physical link* that might occur in a lightpath.
- (iii) i is used as an index for lightpath session number, where $i = 1, 2, \dots, k$. Note that to serve a groupcast session, we need at least a biconnected⁴ simple graph of lightpaths. A groupcast session for x nodes needs $x \times (x - 1)$ lightpaths. Hence $k = \sum_i^\Phi G_i * (G_i - 1)$.

We consider an OTN with all-optical (transparent) OXC nodes capable of full wavelength conversion.

- **Given:**

- N : Number of nodes in the network.
- W : Number of wavelengths per fiber. (This is a system-wide parameter.)
- P_{mn} : $P_{mn} = P_{nm} = 1$ if and only if a physical fiber exists between nodes m and n . Otherwise $P_{mn} = P_{nm} = 0$.
- w_{mn} : Weight or cost associated with a link. The cost could represent fiber distance, signal propagation delay, signal attenuation in a fiber, etc. Note that $w_{mn} = w_{nm}$, and $w_{mn} = \infty$ if $P_{mn} = 0$.

- **Variables:**

- V_{mn}^i : Virtual route equals to one, if there exists a lightpath from node m to n or $V_{mn}^i = 1$. $V_{mn}^i = 0$ otherwise.
- U_x^i : Unicast (lightpath) route, $U_x^i = 1$ if node x belongs to a lightpath, otherwise $U_x^i = 0$. A node can only be part of a session if it is either source, or destination or an intermediate OXC.
- F_{mn}^i : This is an integer commodity flow variable. Each destination node for a session needs one unit commodity.

⁴A graph with no articulation vertex is called a biconnected graph or non-separable graph [54]

- **Objective:** To optimize the total cost of all groupcast sessions:

$$\text{Minimize : } \sum_{i=1}^{i=k} \sum_{m,n} w_{mn} \cdot V_{mn}^i \quad (3.3)$$

Equation (3.3) is a linear objective function as it adds up the cost of individual lightpath sessions.

- **Constraints:**

- Point-to-point lightpath creation:

$$\forall i, \forall n \neq s^i : \sum_m V_{mn}^i = U_n^i \quad (3.4)$$

$$\forall i : \sum_m V_{ms^i}^i = 0 \quad (3.5)$$

$$\forall i : \sum_m V_{d^i m}^i = 0 \quad (3.6)$$

$$\forall i, \forall j = s^i, d^i : U_j^i = 1 \quad (3.7)$$

$$\forall m, n : \sum_i V_{mn}^i \leq P_{mn} \cdot W \quad (3.8)$$

- Commodity-flow constraints:

$$\forall i, \forall m \neq S^i : \sum_n F_{nm}^i = \sum_n F_{mn}^i \quad (3.9)$$

$$\forall i : \sum_n F_{s^i n}^i = 1 \quad (3.10)$$

$$\forall i : \sum_n F_{nd^i}^i = 1 \quad (3.11)$$

$$\forall i, m, n : V_{mn}^i = F_{mn}^i \quad (3.12)$$

- **Explanation of Equations:** The equations are written to create lightpaths to fully connect group nodes for every groupcast session. Equation (3.4) states

that every node belonging to a lightpath has an incoming edge except the source node. Equation (3.5) insures that the source node has no incoming edge and Equation (3.6) insures that the lightpath terminates at the destination node. Every source and destination node belongs to a lightpath (Equation 3.7). Equation (3.8) restricts the lightpaths between node m and n by the number of wavelengths (W) in either direction. The set of flow-conservation equations (3.9)-(3.12) ensures that all sessions leaving a source are transported via a lightpath to a destination node. Equation (3.9) ensures that all intermediate nodes in the lightpath have balanced incoming and outgoing flows. However there is only one outgoing flow for the source node (Equation 3.10) and there is only one incoming flow for the destination node (Equation 3.11). Finally Equation (3.12) ensures that every link occupied by a session has at least one flow.

- **Number of Variables and Constraints:** The groupcast model discussed above is an MILP formulation of the problem. The complexity of the problem is proportional to the number of variables in the system. For each session, we need an $N \times N$ matrix for V_{mn}^i and F_{mn}^i . We also need a flow vector of size N for U_x^i . So a total of $kN + 2kN^2$ variables are needed for k sessions. However, k depends on the groupcast size G and the total number of groupcast sessions Φ , thus $k = \sum_i^\Phi G_i * (G_i - 1)$. Therefore, the total number of variables grows linearly with respect to the number of groupcast sessions and quadratically with respect to the groupcast size or the total number of network nodes. The number of constraints in the above MILP is bounded by $O(kN^2)$. Just like the variables, the constraints grow linearly with respect to the number of sessions and quadratically with respect to the network size or the groupcast size. It is interesting to note that the total number of wavelengths W per fiber does not impact the MILP formulation size.

3.3.3 VTD Formulation-II: No Wavelength Conversion

If OXCs are not equipped with wavelength converters, the wavelength assigned to each lightpath must be the same throughout all the links traversed by the lightpath. This limitation is referred to as the *wavelength continuity constraint* which is unique to OTN. In this section we consider the case where each point-to-point lightpath has the same color wavelength for the entire length of the lightpath. The problem formulation remains the same as in the previous Section (3.3.2), but we now need to introduce some additional variables and constraints to capture the wavelength-continuity constraint.

We consider an OTN with all-optical OXC nodes that are not capable of wavelength conversion.

- **Given:**

- N : Number of nodes in the network.
- W : Number of wavelengths per fiber. (A system wide parameter.)
- P_{mn} : $P_{mn} = P_{nm} = 1$ if and only if a physical fiber exists between node m and node n . Otherwise $P_{mn} = P_{nm} = 0$.
- w_{mn} : Weight or cost of a link traversal. $w_{mn} = w_{nm}$, and $w_{mn} = \infty$ if $P_{mn} = 0$.

- **Variables:**

- V_{mn}^{ic} : Virtual route equals to one if there exists a lightpath from node m to n on wavelength color c ($V_{mn}^{ic} = 1$); otherwise $V_{mn}^{ic} = 0$.
- U_x^i : Unicast (lightpath) route, $U_x^i = 1$ if node x belongs to a lightpath, otherwise $U_x^i = 0$. A node can only be part of a session if it is either a source, a destination or an intermediate OXC.

- F_{mn}^i : This is an integer commodity flow variable. Each destination node for a session needs one unit commodity.
- C_c^i : Color of the wavelength, $C_c^i = 1$ if session i is on wavelength color c , otherwise $C_c^i = 0$.

- **Objective:** To optimize the total cost of all groupcast sessions:

$$\text{Minimize : } \sum_{i=1}^{i=k} \sum_{c=1}^{c=W} \sum_{m,n} w_{mn} \cdot V_{mn}^{ic} \quad (3.13)$$

Equation (3.13) is a linear objective function as it adds up the cost of individual lightpath sessions.

- **Constraints:**

- Point-to-point lightpath creation:

$$\forall i, \forall n \neq s^i : \sum_{m,c} V_{mn}^{ic} = U_n^i \quad (3.14)$$

$$\forall i : \sum_{m,c} V_{ms^i}^{ic} = 0 \quad (3.15)$$

$$\forall i : \sum_{m,c} V_{d^i m}^{ic} = 0 \quad (3.16)$$

$$\forall i, \forall j = s^i, d^i : U_j^i = 1 \quad (3.17)$$

$$\forall m, n : \sum_{i,c} V_{mn}^{ic} \leq P_{mn} \cdot W \quad (3.18)$$

$$\forall m, n, c : \sum_i V_{mn}^{ic} \leq P_{mn} \quad (3.19)$$

- Commodity-flow constraints:

$$\forall i, \forall m \neq S^i : \sum_n F_{nm}^i = \sum_n F_{mn}^i \quad (3.20)$$

$$\forall i : \sum_n F_{s^n i}^i = 1 \quad (3.21)$$

$$\forall i : \sum_n F_{nd^i}^i = 1 \quad (3.22)$$

$$\forall i, m, n, c : V_{mn}^{ic} = F_{mn}^i \quad (3.23)$$

– Wavelength-continuity constrains:

$$\forall i : \sum_c C_c^i = 1 \quad (3.24)$$

$$\forall m, n, i, c, (m \neq n) : V_{mn}^{ic} + V_{nm}^{ic} \leq C_c^i \quad (3.25)$$

- **Explanation of Equations:** The equations above are the same as in Section 3.3.2 except for the Equations 3.19, 3.24, and 3.25 that are newly introduced in this section. Equation (3.19) restricts the number of sessions between same node-pairs to the total number of wavelengths available. The set of wavelength-continuity equations (3.24) and (3.25) ensure that one session can choose only one lightpath and a session can occupy only the wavelength chosen by it (or all its links are on the same wavelength) respectively.
- **Number of Variables and Constraints:** The complexity of the problem has changed with the introduction of the wavelength colors. The total system variables is $kN + kN^2 + kW^2$ for k sessions in the above system. Knowing $k = \sum_i^\Phi G_i * (G_i - 1)$, the total number of variables grows linearly with respect to the number of groupcast sessions and the total number of wavelengths, while it still grows quadratically with respect to the groupcast size or the total number of network nodes. The number of constraints in the above continuous wavelength system is bounded by $O(kN^2W)$. Thus, the constraints grow linearly with respect to the number of sessions and the number of wavelengths, while they grow quadratically with respect to the network size or the groupcast size.

| Groupcast Size | Average Cost |
|----------------|--------------|
| 3 | 1496 |
| 4 | 2925 |
| 5 | 5118 |
| 6 | 7620 |
| 7 | 10554 |
| 8 | 14055 |
| 9 | 18218 |
| 10 | 22545 |
| 11 | 27400 |

Table 3.14: Groupcast cost index (GCI) table for NSF14 network [Figure 3.4]

It is very time consuming and sometimes sheer infeasible to solve the MILP equations stated in Sections (3.3.2, 3.3.3) for a large network (10+ nodes) with standard off-the-shelf commercial MILP softwares. Hence, we resort to heuristics and simulation to investigate the network design problem.

3.3.4 The Network Design Problem: A Heuristic Approach

A typical network design problem (NDP) seeks to minimize the network built-out cost assuming that the traffic demands are known *a priori*. It is not a greenfield approach where the network is to be built from scratch by laying fibers in the ground. The network topology is presented as an input to the design problem (NDP) with various limiting parameters such as fiber connectivities, number of wavelengths per fiber, and/or the node configurations (wavelength-conversion capability, multicast-capability, etc). The NDP problem objective is to minimizing the total network cost (assuming unlimited wavelengths per fiber).

In the early stages of the network design, the traffic demand matrix is often not available. However, we can still model the network and establish an average cost per group size. In order to estimate the network cost for any given combination of groupcast sessions, we establish a groupcast cost index (GCI) table as shown in

Table 3.14 based on the network in Figure 3.4. The table is based on random session generations and by computing their average costs. As only 10% of all possible group combinations is considered given the groupcast size, we use the distribution shown in Figure 3.15 to determine the number of sessions that must be considered per group size.

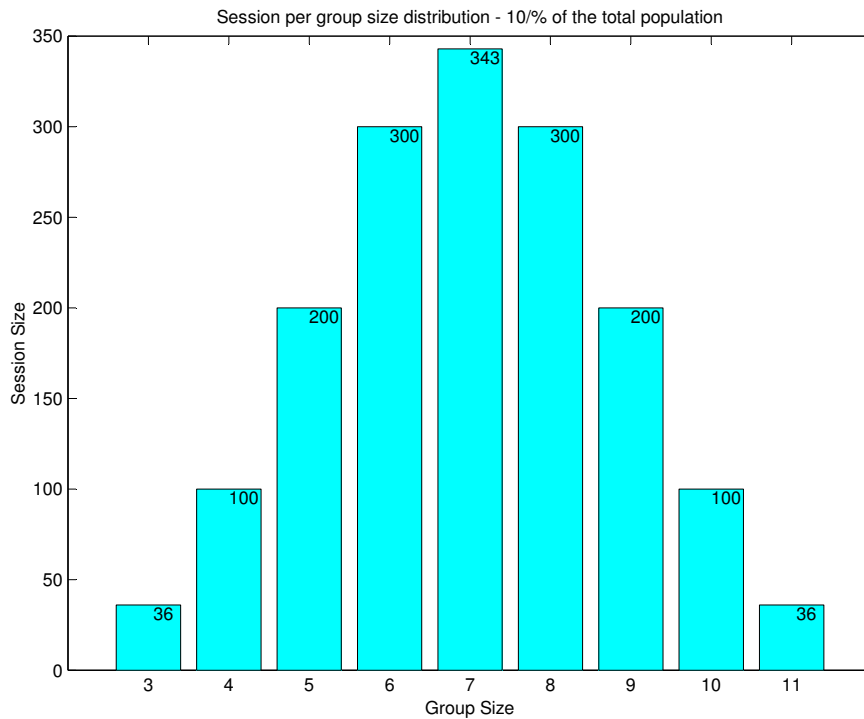


Figure 3.15: Session size requirement (10% of total combinationd) for NSF14

The following conditions are observed for the generation of the GCI table:

- Groupcast session participants are randomly selected.
- For each groupcast node the routing for point-to-multipoint, a non-cyclic multi-cast tree, is generated using Dijkstra's shortest path algorithm (minimum cost path algorithm). Note that any other routing algorithm will also work.
- There are $P(n, k) = \frac{n!}{k!}$ number of sessions for a given groupcast size (k), where n represents the total number of nodes in the network. Since the link costs

are symmetric and routing is done using shortest path, the sequence of nodes participating in the session is irrelevant. Therefore the session number could then be reduced to $C(n, k) = \frac{n!}{k!(n-k)!}$. We generated the Table 3.14 using 10% of all the possible sessions for a given groupcast size (Figure 3.15).

- All the network resources are available to a groupcast session.
- There is no contention for any resource among the sessions. For any given time, there is only one session in the network.
- A session is never blocked. (Neither routing nor wavelength assignment failure is a probable outcome while serving a session.)
- An infinite number (a large number for the model purpose) of wavelengths per fiber is used.
- To serve a groupcast session of size N , we establish $N \times (N - 1)$ point-to-point lightpaths.
- The wavelength assignment function may or may not bear any impact on the total cost of a session depending upon the design of the cost function design. (For example in exhaustive routing the cost of a session is dependent on both routing and wavelength assignment algorithms.)

Example I

We illustrate the application and relevance of the CGI table using the following example. For the groupcast sessions in Table 3.15, we use CGI values from Table 3.14 to compute the total cost of the sessions and compare it with the cost obtained using the heuristic algorithm from Table 3.11. We observe that there is an error margin of

4.71%, which is computed using the following equation:

$$\%error = \frac{|Cost_{Real} - Cost_{CGI}|}{Cost_{Real}} \quad (3.26)$$

| Group Nodes | Size | Real Cost | GCI Value |
|-----------------------|------|--------------|--------------|
| {0 6 5 12 8} | 5 | 5520 | 5118 |
| {5 9 1 10} | 4 | 3810 | 2925 |
| {0 4 6 7 9 11 8} | 7 | 8100 | 10554 |
| {8 9 7 2 13 6 11 1 5} | 9 | 17730 | 18218 |
| Total Cost | | 35160 | 36815 |

Table 3.15: Sample of four groupcast sessions in NSF14 [Figure 3.4]

For planning a network in terms of dimensioning, equipment cost and capital budgeting perspective an 4.71% margin of error is quite impressive given the fact only 10% of the population is used to generate the CGI table. To reduce the margin of error a larger sample from population (in terms of number of sessions) can be used.

3.3.5 Heuristics and Illustrative Results

The network is fundamentally divided into two categories: OXCs supporting wavelength conversion and OXCs with no wavelength conversion capability. For OXCs with no wavelength conversion we can have light-forests with monocolour light-trees or multicolour light-trees. In a monocolour environment the unicast based multicast tree is served by a single wavelength color. In a multicolour environment each unicast session for unicast-based multicast tree can be assigned its own wavelength color. The monocolour versus multicolour decision is primarily dependent on whether the optical control plane [physical layer] and the routing control plane [logical layer] are independent of each other or not. Optical control-plane-aware network routing algorithms like OSPF-TE [34] can effectively support multicolour light-trees. Table 3.16 shows the heuristic for monocolour light-tree forest generation and Table 3.17 shows

the heuristic for multicolor light-tree forest generation. Finally, Table 3.18 shows the heuristic for rainbow forest generation for OXCs with wavelength conversion capabilities. The monicolor light-tree has limited application for OXC-based networks. As all lightpaths originate from a given source node, the nodal degree of the source node becomes the limiting factor for multicast group size in the monicolor light-tree wavelength assignment algorithm. For example, the node NE in Figure 3.4 cannot have a multicast session with nodes CO, IL and GA using the same wavelength. It is obvious that it can have the same wavelength for any of the two destinations but not all three. Therefore, we shall not consider monicolor light-trees any further in this chapter. The monicolor light-tree concept shall be re-examined in the multicast capable optical cross-connect (MC-OXC) networks in Chapter 4. Note that monicolor light-trees are a subset of the multicolor light-tree set for the solution space of a given multicast group size.

| Optical Groupcast Heuristic for Monicolor Light-trees | |
|--|---|
| Step 1 | Given a network $G = (V, E)$ and a list of sessions S , choose the first session $S_1 = \{v_1, v_2, \dots, v_n\}$. |
| Step 2 | Begin with v_1 and compute shortest lightpaths from v_1 to all nodes ($v_2 \dots v_n$) based on hop counts. Construct light-tree T_1 with source node v_1 . |
| Step 3 | Find a wavelength ($\lambda_j, 1 \leq j \leq W$) from W wavelengths such that all the lightpaths from Step 2 can be served. Mark the wavelength λ_j as used along the lightpaths for light-tree T_1 . |
| Step 4 | Repeat Step 2 and Step 3 for all nodes in Session S_1 until all nodes are served (a total of n light-trees.) |
| Step 5 | Repeat Step 1 to Step 4 until all sessions in S are served (a total of $ S $ lightforests.) |

Table 3.16: Lightpath-based groupcast with no wavelength conversion for monicolor light-trees

The performance comparison for optical cross connects with and without wavelength conversion heuristics are shown in Figure 3.16 based on the heuristics in Table 3.18 and Table 3.17 respectively. In order to compare the performance of both

heuristics, 50 unique groupcast sessions ($n = 50$) were generated for a given groupcast session size. The average cost per session size is computed by collecting the sum of all session costs and dividing the total by the number of sessions. As it can be seen from Figure 3.16, the OXCs with wavelength conversion capability (OXC-WC) outperform the regular OXCs (without wavelength conversion capabilities). However, the primary difference between the two heuristics lies in the routing portion; which in the case of OXC-WC take advantage of the wavelength conversion capabilities and do not apply adaptive routing.

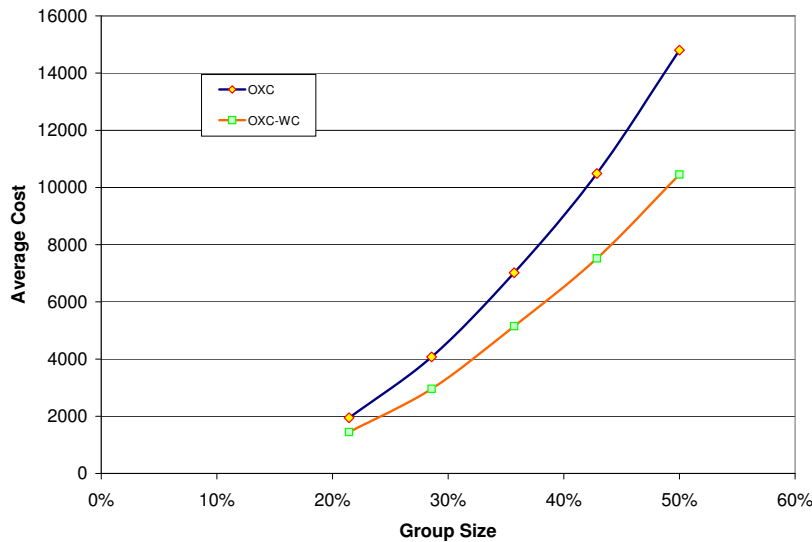


Figure 3.16: Groupcast sessions in OXC-based network (NSF14) with and without wavelength conversion

The adaptive or layered graph approach in Table 3.17 is not a good heuristic for cost optimization although Figure 3.16 does not actually show this. However, the session costs are exactly identical for both wavelength-conversion-capable and wavelength-continuous WDM mesh networks when the layered graph approach (**Step 2**) is ignored in the multicolor heuristic of Table 3.17. This is because the problem is then decoupled into two separate problems and routing becomes independent of wavelength assignment. Since the cost is directly associated with the routing of groupcast calls, the costs culminate to same figure for both wavelength conversion

| Optical Groupcast Heuristic for Multicolor Light-trees | |
|---|--|
| Step 1 | Given a set of groupcasts S_n , where n is the number of sessions, construct a list (Λ_k) by decomposing each session into multiple point-to-point source-destination pairs $\Lambda_k = \{s_k, d_k\}$. |
| Step 2 | Choose sequentially a wavelength $(\lambda_j, 1 \leq j \leq W)$ and construct graph $G = (V, E)$ to find the lightpath to serve the first session from Λ_k . |
| Step 3 | Mark the wavelength λ_j as used along the lightpath. If the lightpath is not found, choose the next wavelength in Step 2 and repeat Step 2 until the session is served. |
| Step 4 | Repeat Step 3 for all sessions until Λ_k is completed. |

Table 3.17: Lightpath-based groupcast with no wavelength conversion for multicolor light-trees

| Optical Groupcast Heuristic for Rainbow Light-forests | |
|--|---|
| Step 1 | Given a network $G = (V, E)$ and a list of sessions S_n , where n is the number of sessions, construct a list (Λ_k) by decomposing each session into multiple point-to-point source-destination sessions $\Lambda_k = \{s_k, d_k\}$. |
| Step 2 | Compute shortest lightpath for each session in Λ_k . |
| Step 3 | Make a link set $L_k = \{l_1, l_2, \dots, l_x\}$ required to construct the lightpath to serve Λ_k . Find a wavelength $(\lambda_j, 1 \leq j \leq W)$ from W wavelengths for each link in L_k . |
| Step 4 | Mark each wavelengths assigned $L_k = \{l_1^{\lambda_1}, \dots, l_x^{\lambda_j}\}$ for the link set L_k . |
| Step 5 | Repeat Step 2 to Step 4 until all sessions in Λ_k are served. |

Table 3.18: Lightpath-based groupcast with wavelength conversion for rainbow light-trees

capable and incapable networks. The advantage of wavelength conversion is better captured with dynamic traffic when the network performance is measured in terms of call blocking. Certainly, the wavelength conversion capabilities give considerable advantage as consistently demonstrated with lightpath performance [16] and light-tree performance [55]. As for static traffic, we shall look at other objectives beside cost optimization to holistically capture the network optimization in terms of both routing and wavelength assignment for groupcast sessions.

The objective to minimize cost assumes that we have unlimited wavelengths per fiber which may not be always applicable when designing networks. We can alternatively investigate the network design problem (NDP) in terms of minimizing the total number of wavelengths per fiber required to meet traffic demands. Note that we can consider routing cost when we make routing decisions, but we do not explicitly consider network cost as a system parameter for the wavelength minimization objective. We have observed earlier in Figure 3.13 that *min-cost-path* algorithm performs slightly worse than *shortest-path-first* in terms of wavelength requirements for unicast-based groupcast approach. We formulate the problem of designing an all-optical network with the objective to minimize the number of wavelengths required per fiber as follows:

- The network topology, represented by the graph $G = (V, E)$, is provided as an input to the problem.
- The link weight represents the cost associated with the traversal of the fiber route.
- The groupcast sessions are tabulated and provided as a input to the problem.
- The routing and wavelength assignment algorithm is specified as a heuristic.
- The order and selection criteria of a session is part of the heuristic.
- The output of the design process is the route specification and the wavelength assignment for the input sessions including a system-wide optimization requirement parameter: minimum number of wavelengths per fiber.

The heuristic to design wavelength optimized groupcast sessions is described in Table 3.19. Note that the sessions can be served in the ascending or descending order based on cost, which produces COST (ASC) and COST (DES) heuristics, or number

of links required in terms of network resource, which results in SIZE (ASC) and SIZE (DES) heuristics. If otherwise we serve the sessions in random order as they are provided as requirements for the network, the RANDOM heuristic can be used.

| Wavelength Optimized Optical Groupcast Heuristic | |
|---|--|
| Step 1 | Given a network $G = (V, E)$ and a list of sessions S_n , where n is the number of sessions, construct a list Λ by decomposing each session into multiple point-to-point source-destination sessions ($\Lambda_k = \{s_k, d_k\}$). |
| Step 2 | Compute shortest lightpath for each session in Λ_k and sort it based on cost or hop counts. It must be sorted in ascending or descending order. For random heuristic, ignore Step 2 . |
| Step 3 | Choose sequentially a wavelength ($\lambda_j, 1 \leq j \leq W$) from W wavelengths and construct a new graph $G_j = (V, E)$. Choose a session from Step 2 and compute the lightpath to serve it. Mark the wavelength λ_j as used along the lightpath. |
| Step 4 | Repeat Step 3 for all sessions until Λ_k is completed. |

Table 3.19: Layered-graph based Groupcasting for Wavelength-continuous OXC-based Networks

In order to evaluate the heuristics proposed in Table 3.19 we generated a number of groupcast sessions and applied the heuristics to design the network. The groupcast size varied uniformly between minimum groupcast size (3) and maximum groupcast size ($|V| - 1$). Note that including all nodes in the network in communication sessions is considered a broadcast session and a unicast session requires at least two nodes. The performance of the heuristics as a function of the average number of wavelengths per session is shown in Figure 3.17. The total number of wavelengths required to serve a session for a given heuristic is divided by the total number of sessions in order to obtain the average number of wavelengths required per session. The RANDOM algorithm consistently outperformed the rest of the heuristics for all session sizes except for session size of 10 where SIZE (ASC) tied with the RANDOM heuristic. For most sessions, we observe that sorting the unicast sessions based on cost or path size (in terms of hop counts) deteriorates the performance. For larger sessions (40

and 80), we observe that the performance of the heuristics can be summarized in the following order: $COST(ASC) \leq COST(DES) \leq SIZE(ASC) \leq SIZE(DES) \leq RANDOM$.

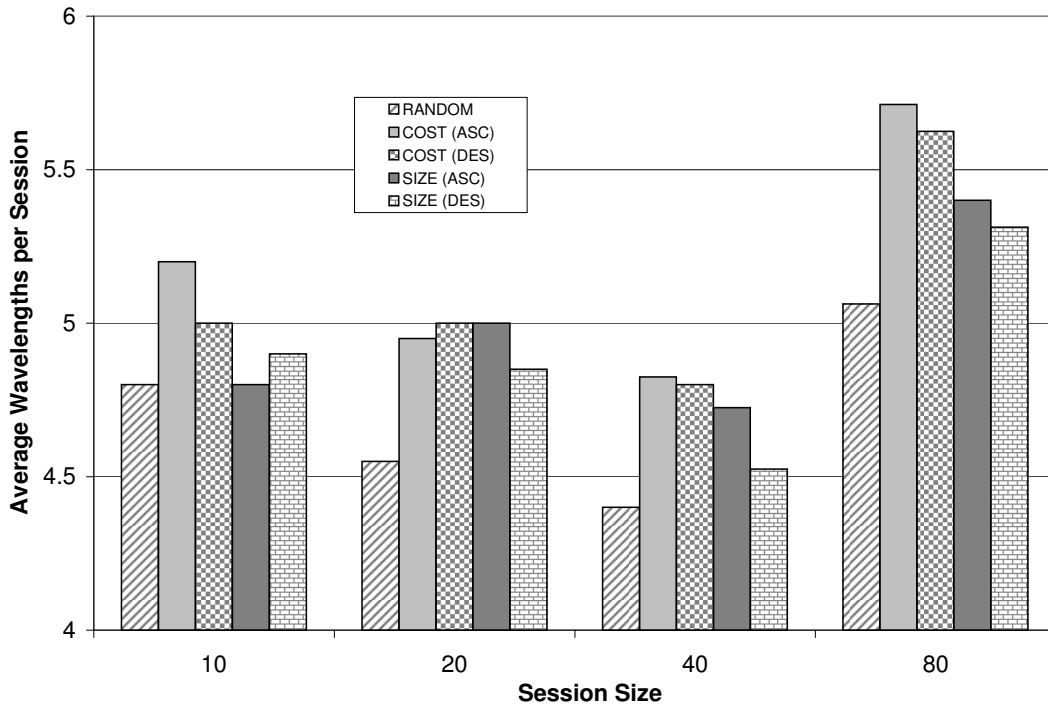


Figure 3.17: Wavelength usage per groupcast session in the wavelength-continuous OXC-based NSF14 network [Ref. Figure 3.4]

We also investigated the cost performance for the same heuristics and the results are shown in Figure 3.18. In order to compare different sessions we normalized the cost for groupcast sessions as follows: we took the average cost of links in the network and multiplied it by the average hop distance between two nodes in the network. The product is then multiplied by the total number of unicast sessions generated for each groupcast set to obtain the normalization value. The total cost for each groupcast set is divided by its normalization value to obtain the heuristic performance shown in Figure 3.18. Recall that each groupcast set has a uniformly distributed groupcast size between 3 to 13 nodes which translates into 6 to 156 unicast sessions per groupcast. In terms of cost, the RANDOM heuristic outperformed the rest of the heuristics for

larger sets of groupcast sessions ($n = 20$ and more groupcast sessions or $k = 1164$ and more of unicast sessions). Also, the COST (ASC) heuristic showed the worst performance for both small and large number of groupcast sets.

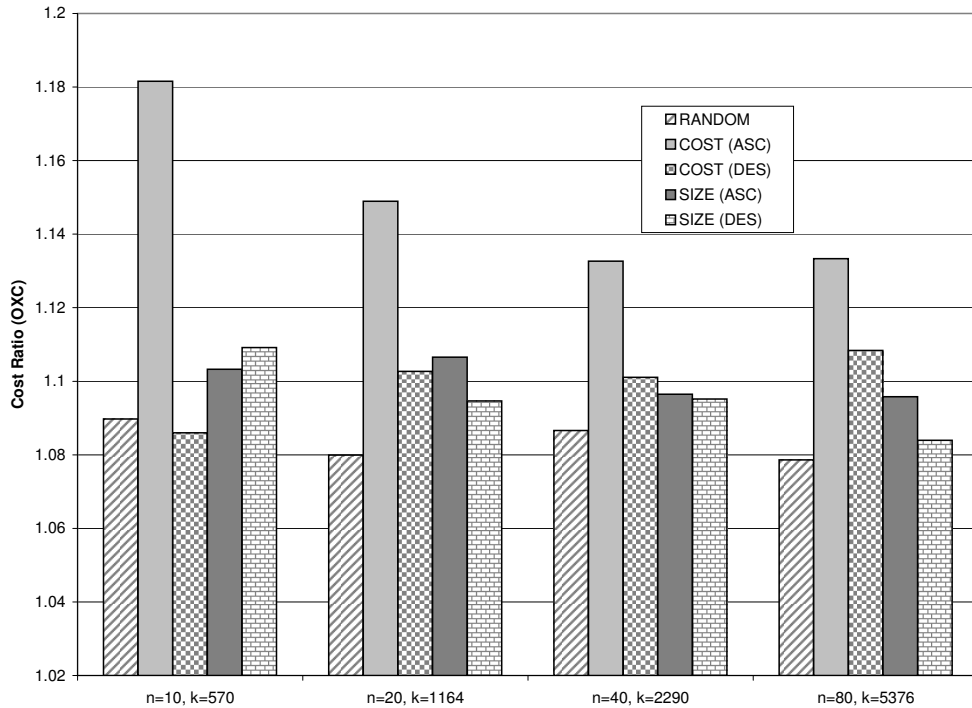


Figure 3.18: Cost ratio per groupcast session in the wavelength-continuous OXC-based NSF14 network [Figure 3.4]

We conclude this section by an illustrative example from the NSF network [Figure 3.4]. Assuming that we have an optical network with optical cross-connects deployed at each node supporting no wavelength conversion, we set-up three groupcast sessions as illustrated in Table 3.20. The group size and group members were randomly selected. The objective is to design lightpath assignments for establishing the groupcast connections while minimizing the wavelength requirements to serve all the sessions.

The algorithm specified in Table 3.19 is used to generate the solution for the groupcast lightpath sessions and wavelength assignments captured in Table 3.21. As

| Session No. | Group Participants | Group Size |
|-------------|---------------------------------|------------|
| 1 | { GA, TX, MD, IL, CA1 } | 5 |
| 2 | { PA, WA, MI, TX } | 4 |
| 3 | { CO, GA, NE, PA, NY, IL, CA1 } | 7 |

Table 3.20: List of 3 groupcast sessions for the NSF14 network of Figure 3.4

expected, a total of 74 point-to-point lightpath sessions were generated for 3 groupcast sessions (size of 5, 4 and 7): particularly $P(5, 2) = \binom{5}{2}$ generated 20 sessions, $P(4, 2) = \binom{4}{2}$ generated 12 sessions and finally $P(7, 2) = \binom{7}{2}$ generated 42 sessions, thus a total of $(20 + 12 + 42)$ (74) sessions. The solution uses a total of 14 wavelengths and hence we need a minimum of 14×14 wavelength cross-connect (OXC) switching plane to serve all the sessions. Also we need at least 14 wavelengths per fiber to support all the groupcast sessions simultaneously.

| Lightpath Assignments for Groupcasts in Table 3.20 | | | | |
|--|-------------|---|------------|------|
| Source | Destination | Path | wavelength | Cost |
| {GA} | {TX} | {GA → TX, } | 1 | 150 |
| {GA} | {MD} | {GA → PA, PA → NJ, NJ → MD, } | 1 | 285 |
| {GA} | {IL} | {GA → PA, PA → IL, } | 2 | 165 |
| {GA} | {CA1} | {CA2 → CA1, TX → CA2, GA → TX, } | 2 | 375 |
| {TX} | {GA} | {TX → GA, } | 1 | 150 |
| {TX} | {MD} | {TX → MD, } | 1 | 300 |
| {TX} | {IL} | {CO → NE, TX → CO, NE → IL, } | 1 | 195 |
| {TX} | {CA1} | {CA2 → CA1, TX → CA2, } | 1 | 225 |
| {MD} | {GA} | {PA → GA, MD → NJ, NJ → PA, } | 1 | 285 |
| {MD} | {TX} | {MD → TX, } | 1 | 300 |
| {MD} | {IL} | {PA → IL, MD → NJ, NJ → PA, } | 3 | 240 |
| {MD} | {CA1} | {CA2 → CA1, TX → CA2, MD → TX, } | 3 | 525 |
| {IL} | {GA} | {IL → PA, PA → GA, } | 2 | 165 |
| {IL} | {TX} | {CO → TX, NE → CO, IL → NE, } | 1 | 195 |
| {IL} | {MD} | {IL → PA, PA → NJ, NJ → MD, } | 3 | 240 |
| {IL} | {CA1} | {UT → CA1, CO → UT, NE → CO, IL → NE, } | 2 | 300 |
| {CA1} | {GA} | {CA1 → CA2, CA2 → TX, TX → GA, } | 2 | 375 |
| {CA1} | {TX} | {CA1 → CA2, CA2 → TX, } | 1 | 225 |
| {CA1} | {MD} | {CA1 → CA2, CA2 → TX, TX → MD, } | 3 | 525 |

continued on next page

| Lightpath Assignments for Groupcasts in Table 3.20 (<i>Continued</i>) | | | | |
|---|-------------|---|------------|------|
| Source | Destination | Path | wavelength | Cost |
| {CA1} | {IL} | {CA1 → UT, UT → CO, CO → NE, NE → IL, } | 2 | 300 |
| {PA} | {WA} | {IL → WA, PA → IL, } | 1 | 360 |
| {PA} | {MI} | {PA → NY, NY → MI, } | 1 | 135 |
| {PA} | {TX} | {GA → TX, PA → GA, } | 3 | 255 |
| {WA} | {PA} | {WA → IL, IL → PA, } | 1 | 360 |
| {WA} | {MI} | {CA1 → UT, WA → CA1, UT → MI, } | 1 | 495 |
| {WA} | {TX} | {WA → CA2, CA2 → TX, } | 4 | 270 |
| {MI} | {PA} | {MI → NY, NY → PA, } | 1 | 135 |
| {MI} | {WA} | {IL → WA, PA → IL, MI → NY, NY → PA, } | 4 | 495 |
| {MI} | {TX} | {GA → TX, PA → GA, MI → NY, NY → PA, } | 5 | 390 |
| {TX} | {PA} | {TX → GA, GA → PA, } | 3 | 255 |
| {TX} | {WA} | {CA2 → WA, TX → CA2, } | 4 | 270 |
| {TX} | {MI} | {TX → GA, GA → PA, PA → NY, NY → MI, } | 4 | 390 |
| {CO} | {GA} | {CO → TX, TX → GA, } | 5 | 225 |
| {CO} | {NE} | {CO → NE, } | 3 | 60 |
| {CO} | {PA} | {CO → NE, NE → IL, IL → PA, } | 4 | 180 |
| {CO} | {NY} | {CO → NE, NE → IL, IL → PA, PA → NY, } | 5 | 255 |
| {CO} | {IL} | {CO → NE, NE → IL, } | 6 | 120 |
| {CO} | {CA1} | {UT → CA1, CO → UT, } | 1 | 180 |
| {GA} | {CO} | {TX → CO, GA → TX, } | 4 | 225 |
| {GA} | {NE} | {GA → PA, IL → NE, PA → IL, } | 5 | 225 |
| {GA} | {PA} | {GA → PA, } | 6 | 105 |
| {GA} | {NY} | {GA → PA, PA → NY, } | 7 | 180 |
| {GA} | {IL} | {GA → PA, PA → IL, } | 8 | 165 |
| {GA} | {CA1} | {CA2 → CA1, TX → CA2, GA → TX, } | 6 | 375 |
| {NE} | {CO} | {NE → CO, } | 3 | 60 |
| {NE} | {GA} | {NE → IL, IL → PA, PA → GA, } | 7 | 225 |
| {NE} | {PA} | {NE → IL, IL → PA, } | 8 | 120 |
| {NE} | {NY} | {NE → IL, IL → PA, PA → NY, } | 9 | 195 |
| {NE} | {IL} | {NE → IL, } | 3 | 60 |
| {NE} | {CA1} | {UT → CA1, CO → UT, NE → CO, } | 4 | 240 |
| {PA} | {CO} | {NE → CO, IL → NE, PA → IL, } | 6 | 180 |
| {PA} | {GA} | {PA → GA, } | 4 | 105 |
| {PA} | {NE} | {IL → NE, PA → IL, } | 7 | 120 |
| {PA} | {NY} | {PA → NY, } | 2 | 75 |
| {PA} | {IL} | {PA → IL, } | 9 | 60 |

continued on next page

| Lightpath Assignments for Groupcasts in Table 3.20 (Continued) | | | | |
|--|-------------|--|------------|--------------|
| Source | Destination | Path | wavelength | Cost |
| {PA} | {CA1} | { $UT \rightarrow CA1, CO \rightarrow UT, NE \rightarrow CO, IL \rightarrow NE, PA \rightarrow IL,$ } | 10 | 360 |
| {NY} | {CO} | { $NE \rightarrow CO, IL \rightarrow NE, PA \rightarrow IL, NY \rightarrow PA,$ } | 11 | 255 |
| {NY} | {GA} | { $PA \rightarrow GA, NY \rightarrow PA,$ } | 6 | 180 |
| {NY} | {NE} | { $IL \rightarrow NE, PA \rightarrow IL, NY \rightarrow PA,$ } | 12 | 195 |
| {NY} | {PA} | { $NY \rightarrow PA,$ } | 2 | 75 |
| {NY} | {IL} | { $PA \rightarrow IL, NY \rightarrow PA,$ } | 13 | 135 |
| {NY} | {CA1} | { $UT \rightarrow CA1, CO \rightarrow UT, NE \rightarrow CO, IL \rightarrow NE, PA \rightarrow IL, NY \rightarrow PA,$ } | 14 | 435 |
| {IL} | {CO} | { $NE \rightarrow CO, IL \rightarrow NE,$ } | 8 | 120 |
| {IL} | {GA} | { $IL \rightarrow PA, PA \rightarrow GA,$ } | 10 | 165 |
| {IL} | {NE} | { $IL \rightarrow NE,$ } | 3 | 60 |
| {IL} | {PA} | { $IL \rightarrow PA,$ } | 6 | 60 |
| {IL} | {NY} | { $IL \rightarrow PA, PA \rightarrow NY,$ } | 11 | 135 |
| {IL} | {CA1} | { $UT \rightarrow CA1, CO \rightarrow UT, NE \rightarrow CO, IL \rightarrow NE,$ } | 9 | 300 |
| {CA1} | {CO} | { $CA1 \rightarrow UT, UT \rightarrow CO,$ } | 3 | 180 |
| {CA1} | {GA} | { $CA1 \rightarrow CA2, CA2 \rightarrow TX, TX \rightarrow GA,$ } | 6 | 375 |
| {CA1} | {NE} | { $CA1 \rightarrow UT, UT \rightarrow CO, CO \rightarrow NE,$ } | 7 | 240 |
| {CA1} | {PA} | { $CA1 \rightarrow UT, UT \rightarrow CO, CO \rightarrow NE, NE \rightarrow IL, IL \rightarrow PA,$ } | 12 | 360 |
| {CA1} | {NY} | { $CA1 \rightarrow UT, UT \rightarrow CO, CO \rightarrow NE, NE \rightarrow IL, IL \rightarrow PA, PA \rightarrow NY,$ } | 13 | 435 |
| {CA1} | {IL} | { $CA1 \rightarrow UT, UT \rightarrow CO, CO \rightarrow NE, NE \rightarrow IL,$ } | 10 | 300 |
| Total | | | 14 | 17430 |

Table 3.21: The point-to-point lightpaths, wavelength assignment and lightpath cost resulting from three groupcast sessions in Table 3.20

3.4 Dynamic Groupcasts: Performance Analysis in OXC Networks

In a dynamic network the sessions are established as they arrive (on-the-fly) and teared-down after a requested period. We will present the performance of GC-RWA algorithms under dynamic traffic conditions in this section. In an OXC-based network, we can only establish point-to-point lightpaths as described in Section 3.2. Thus, for groupcast, we end up with $n \times (n - 1)$ lightpaths for a light-forest of size

n. The lightpath-based groupcast approach can be used to establish light-forests in both wavelength-continuous networks and networks with no wavelength-conversion capability.

We investigate dynamic groupcast performance for routing algorithms and wavelength assignment algorithms separately. Wavelength assignment can be constrained depending whether we use optical cross-connects (OXC) with and without wavelength conversion. The performance of groupcast routing and wavelength assignment (GC-RWA) in wavelength-continuous WDM-based mesh optical networks is presented in Section 3.4.1 while the performance in networks with no wavelength-conversion capability is presented in Section 3.4.2.

3.4.1 GC-RWA performance in wavelength-continuous OXC-based network

When the OXCs do not support wavelength conversion in the network, it is understood that the network is hardware-limited to continuous wavelength models. The wavelength assignment models for wavelength-continuous lightpaths are categorized based on the wavelength selection algorithms. In this section we study the performance of four popular wavelength assignment algorithms [16] [15] [41] for groupcasts namely: First-fit (FF), Random (RA), SPREAD and PACK as described previously in this chapter.

In order to compare the performance for various wavelength assignment heuristics, we maintain the groupcast routing module fixed. We use the minimum-cost heuristic routing (MCH) throughout this section unless otherwise stated. The general heuristic for lightpath-based GC-RWA is given in Table 3.22.

The GC-RWA heuristic performance is measured in terms of session blocking.

Table 3.22: Continuous lightpath-based groupcast

| | |
|---------------|--|
| Step 1 | Choose the first node from a groupcast node set $\{n_1, n_2, \dots, n_k\}$ where the group size is k . |
| Step 2 | Compute $(n - 1)$ lightpaths using the <i>shortest-path-first</i> or <i>min-cost</i> routing algorithm and save the routes in the route set $R^1 = \{r_1, r_2, \dots, r_{(n-1)}\}$ where R^1 indicates the route set for source node number one. |
| Step 3 | Repeat Step 1 and Step 2 until all the nodes from the node set are served and the route set $\{R^1, R^2, \dots, R^k\}$ is completed. |
| Step 4 | Choose the first route from the route set and find a wavelength using FF, RA, PACK or SPREAD wavelength-assignment algorithm. |
| Step 5 | If the wavelength x is chosen in Step 4 then the route set is updated with the wavelength number R_x^1 and wavelength x is booked in wavelength matrix $W_{L,N}$, where L is the total number of links in the network and N is the total number of wavelengths per link. |
| Step 6 | Repeat Step 4 and Step 5 until all the routes are served from the route set $\{R^1, R^2, \dots, R^k\}$. |

First we compare the performance of groupcast sessions in two different networks: 14-node NSF network (Figure 3.4) and 24-node AT&T network (Figure 3.10). Then we investigate the effect of wavelength assignment algorithms on GC-RWA performance.

The graphs obtained in this section are based on simulation results with the following assumptions: The simulator uses a dynamic traffic model where groupcast call requests arrive at each node according to a Poisson process and the holding time is negative exponentially distributed with a unit mean. In each simulation case 1,000 groupcast requests were generated, and the results were averaged over five simulation runs.

Figure 3.19 shows the performance of groupcasts in two different networks. We considered 128 wavelengths per link for each experiment in order to evaluate the blocking probability vs. traffic load [in Erlangs]. The group size in each case varied from 3 to $(|V| - 2)$, where $|V|$ is the total nodes in the network. It is apparent from the plots that the 14-node NSF network performs better than the 24-node AT&T network. Upon further investigation, we see a relationship between the nodal degree

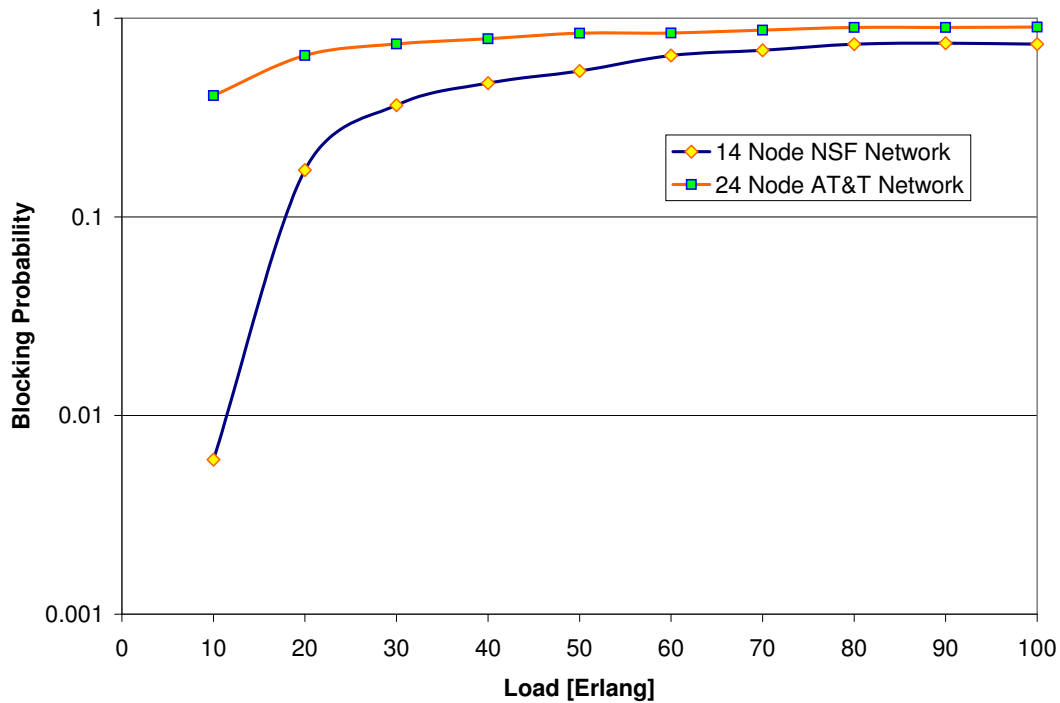


Figure 3.19: Performance of groupcast routing and wavelength assignment (GC-RWA) for the dynamic traffic case

of the network and the groupcast blocking performance. Figure 3.20 shows that as the network load increases, the performance of the blocking ratio for the AT&T network compared to the NSF network converges to the value of the nodal degree ratio for AT&T to the NSF network, which is 1.194.

Next, we investigate the performance of random wavelength assignment for GC-RWA. The *random-fit* algorithm as described in Table 3.22 differs from *first-fit* in terms of wavelength selection once the lightpaths are computed. For *random-fit* the wavelengths are randomly selected as oppose to sequential searching of free wavelengths. Figure 3.21 is generated based on GC-RWA simulations on the 14-node NSF network [Figure 3.4]. For the simulation we considered 64 wavelengths per fiber and we observed that the *random-fit* wavelength assignment yielded slightly poorer performance than *first-fit* wavelength assignment for optical groupcast sessions under the

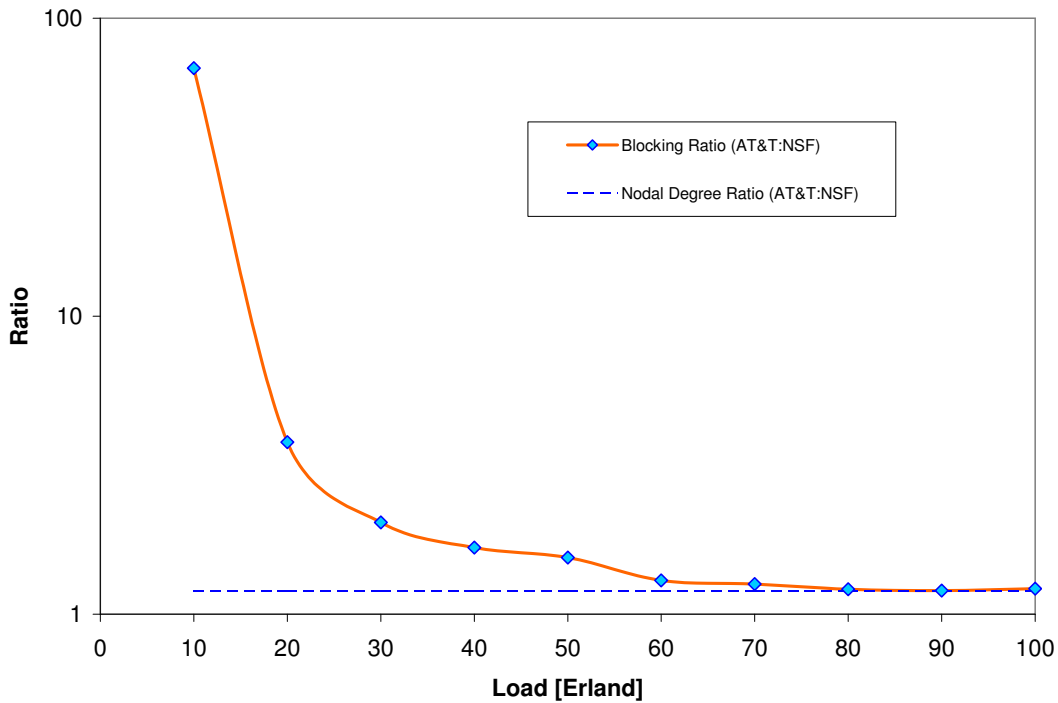


Figure 3.20: Groupcast performance comparison between NSF and AT&T networks majority of network traffic loads.

There are two wavelength assignment approaches: state-aware and state-unaware wavelength assignment algorithms. For the state-unaware case, the wavelength assignment algorithm does not inspect if the graph $G_\lambda = (V, E)$, indicating the physical graph at the specified wavelength, has the probability of accommodating the light-path request. On the other hand, the state-aware algorithms do exactly that, such as *PACK* and *SPREAD* as explained in Table 3.22. For the *PACK* algorithm we choose the wavelength that is serving most calls and for the *SPREAD* we choose the wavelength that has the least amount of calls. Figure 3.22 captures the performance of state-aware wavelength assignment algorithms. The simulation was conducted on the NSF network [Figure 3.4] using 64 wavelengths per fiber, and we observed that the *SPREAD* algorithm performs slightly better than the *PACK* algorithms as the network load increases.

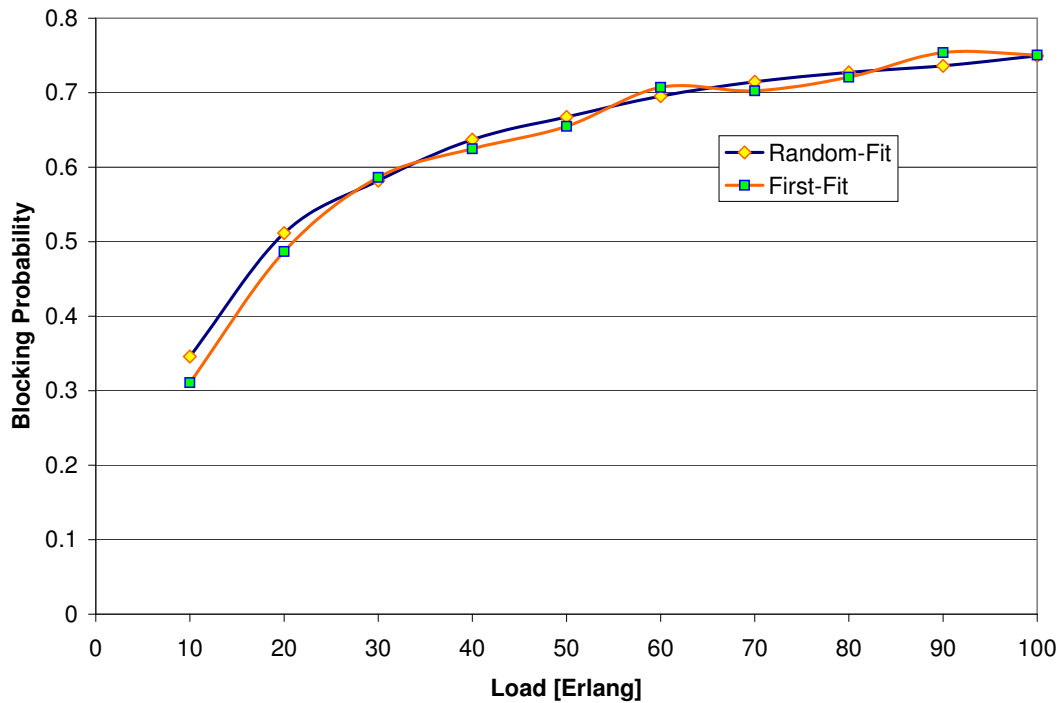


Figure 3.21: Wavelength assignment performance in the NSF network for groupcast sessions

3.4.2 GC-RWA Performance in OXC-Based Networks With Wavelength-Conversion Capabilities

In this section we study the performance of groupcast sessions when we discount the wavelength-continuity constraint. It is assumed that each optical cross-connect is equipped to convert signals from any incoming wavelength to any outgoing wavelength. It is important to point out that wavelength conversion facilitates the implementation of a distributed control plane. Under the distributed control plane model, the intermediate systems independently assign the wavelengths available in its outgoing links. The wavelength continuity model requires much more involved coordination among the intermediate systems and therefore a centralized control plane is preferable.

Figure 3.23 shows the performance of groupcast sessions with wavelength-conversion

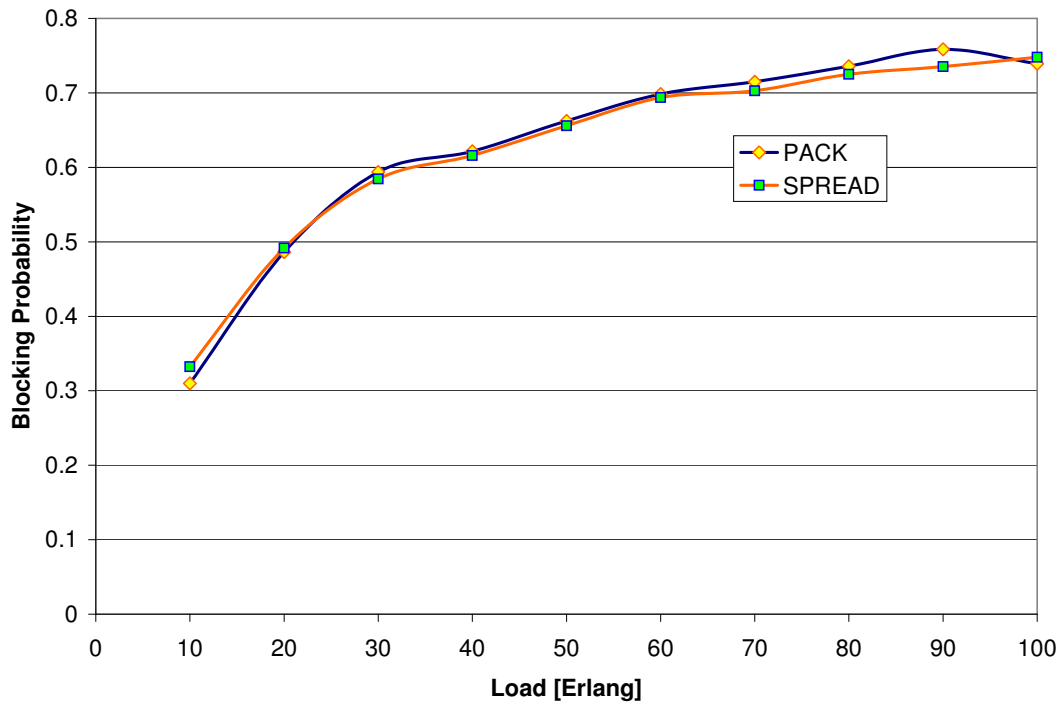


Figure 3.22: Performance of state-aware wavelength assignment algorithms for group-cast sessions in the NSF network

Table 3.23: Lightpath based groupcast in networks with wavelength conversion capabilities

| | |
|---------------|---|
| Step 1 | Choose the first node from a groupcast node set $\{n_1, n_2, \dots, n_k\}$ where the group size is k . |
| Step 2 | Compute $(n - 1)$ lightpaths using the <i>shortest-path-first</i> or <i>min-cost</i> routing algorithm and save the routes in the routes set $R^1 = \{r_1, r_2, \dots, r_{(n-1)}\}$ where R^1 indicates the route set for source node number one. |
| Step 3 | Repeat Step 1 and Step 2 until all the nodes from the node set are served and the route set $\{R^1, R^2, \dots, R^k\}$ is completed. |
| Step 4 | Choose the first route from the route set and find the first available wavelength. When the wavelength x is chosen the route is updated by the wavelength number R_x^1 . |
| Step 5 | Repeat Step 4 until all the routes are served from the route set $\{R^1, R^2, \dots, R^k\}$. |

versus wavelength-continuous groupcast sessions. We simulated 1,000 groupcast sessions with a uniformly distributed groupcast size and dynamic traffic loads. The traffic sessions requests were generated according to a Poisson process where the

holding time were negative exponentially distributed with a unit mean. We used the NSF 14-node network with 128 wavelengths per fiber to obtain the plots in Figure 3.23. For the wavelength-continuous case we used the *First-fit* algorithm for wavelength assignment. Similar to the wavelength-continuous case, the algorithm for the case of no wavelength conversion is described in Table 3.23. We observe that the groupcast performance for both wavelength-continuous networks and networks with wavelength-conversion capabilities are very comparable. The networks with wavelength conversion perform slightly better than the wavelength-continuous networks, however the gain is not significant.

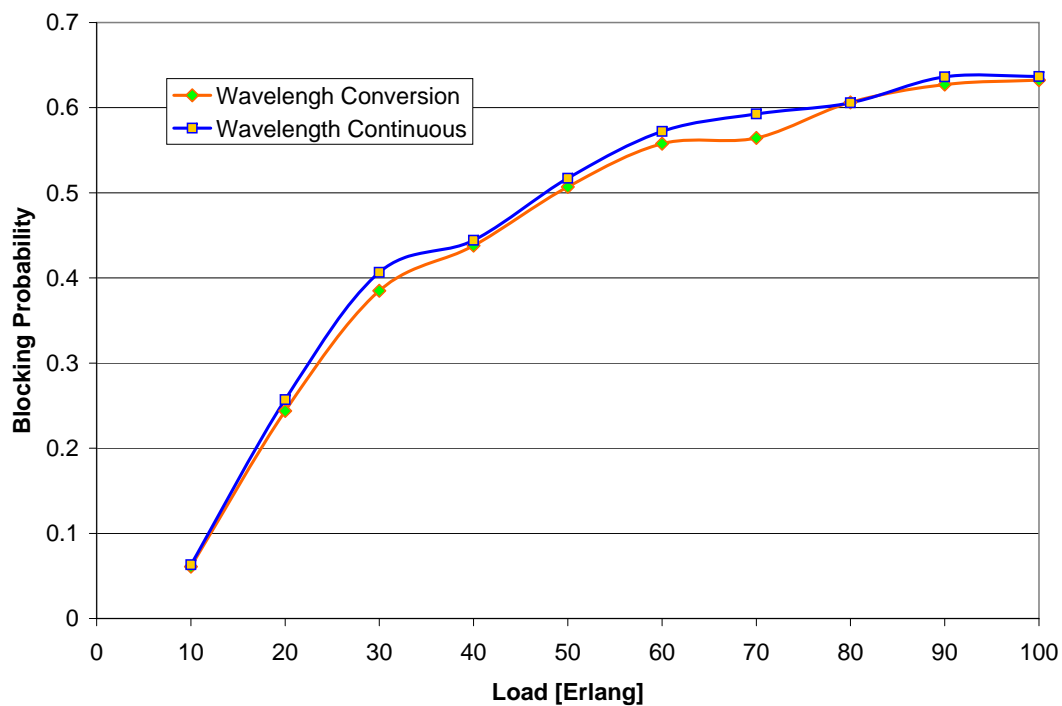


Figure 3.23: State-aware wavelength assignment performance in NSF network for groupcast sessions

Chapter 4

Light-tree-Based Optical Groupcast

As groupcast is a set of multipoint-to-multipoint connections, the unicast (lightpath) approach is not the only way to create light forests. Multiple light-trees can also be created to support a single groupcast session (light-tree based groupcast). A light-tree represents a single physical wavelength (optical channel) in all the fiber links that are part of the session-path. Constructing a light-forest among a group of N participants requires setting-up N conventional light-trees concurrently. Thus, the problem of routing and wavelength assignment for a light-forest (RWA Light-forest) is decomposed into two sub-problems: (1) Constructing a light-tree for each (source) node in the multicast group and (2) Finding a wavelength for each light-tree. Each step involves solving a set of problems that can influence overall RWA performance of the light-forest. To construct a cost-effective light-forest, we need to construct cost-effective light-trees. In general, finding efficient light-trees is an NP-complete problem and so heuristics are used instead to find the light-trees. Depending on the multi-cast optical node architecture used in the network, several approaches can be used

to build a light-forest. Two approaches are analyzed in this chapter and their performances are compared with each other. The first approach utilizes the conventional hybrid (OOO/OEO) multicast optical node architecture as described in Chapter 2 (see Figure 4.1(a)), whereas the second approach utilizes simple Drop-and-continue (DaC) node architectures as described in the section that follows. Figure 4.1(b) shows light-tree based groupcast session.

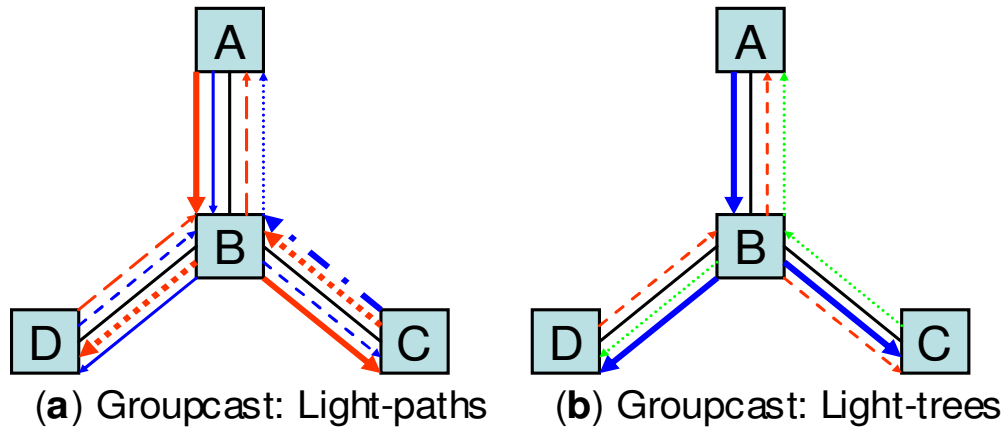


Figure 4.1: Approaches to support groupcast in wavelength-routed network

Optical multicast performance in the drop-and-continue context has been examined in [56, 20, 57, 19, 58, 59, 60, 18]. The hardware and OADM architecture performance is presented in [61, 62]. The optimal placement of splitters for power-efficient optical network design [57, 60, 63, 18, 58, 55] is also a common theme prevalent in the literature under the sparse spitting topic. Nevertheless, the usage of DaC and MC-OXC nodes to support groupcast connections has not been examined in the literature. This is precisely the contribution of this chapter.

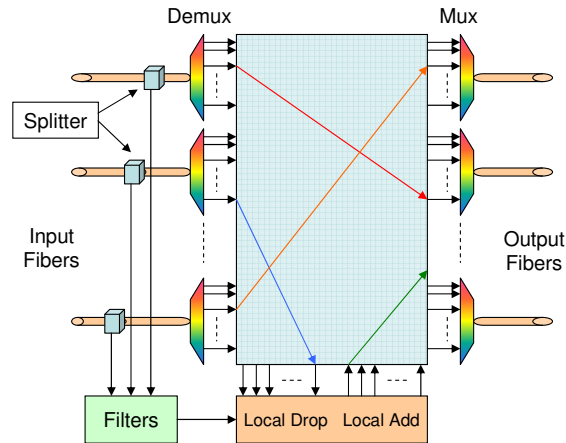


Figure 4.2: Drop-and-continue optical cross-connect equipped with power splitters to support tap-and-continue architecture

4.1 Drop-and-Continue (DaC) Node Architectures

A general architecture of a DaC utilizing optical cross-connects is shown in Figure 4.2. There are two other all-optical DaC architectures: broadcast-and-select using dynamic spectral equalizer (DSE) technology (Figure 4.3) and optical add/drop multiplexer (OADM) architecture (Figure 4.4). The OADM shown in the figure also deploys a redundant switching plane architecture for survivability purposes [64]. All incoming optical wavelength channels are demultiplexed, then switched in the switch fabric and remultiplexed. The switch allows wavelengths to pass through the node or to be added or dropped. For broadcast-and-select (BaS) technology, the splitters broadcast the multiwavelength output and single-channel filters select the appropriate wavelength to be dropped [65].

In optical multicasting the splitter is the fundamental optical device enabling an input optical signal to be split onto multiple outgoing links. For n -way splitter, the power of an incoming signal is reduced to $1/n$ for its outgoing signal power [16]. In optical groupcast drop-and-continue functionality can also be used to accommodate

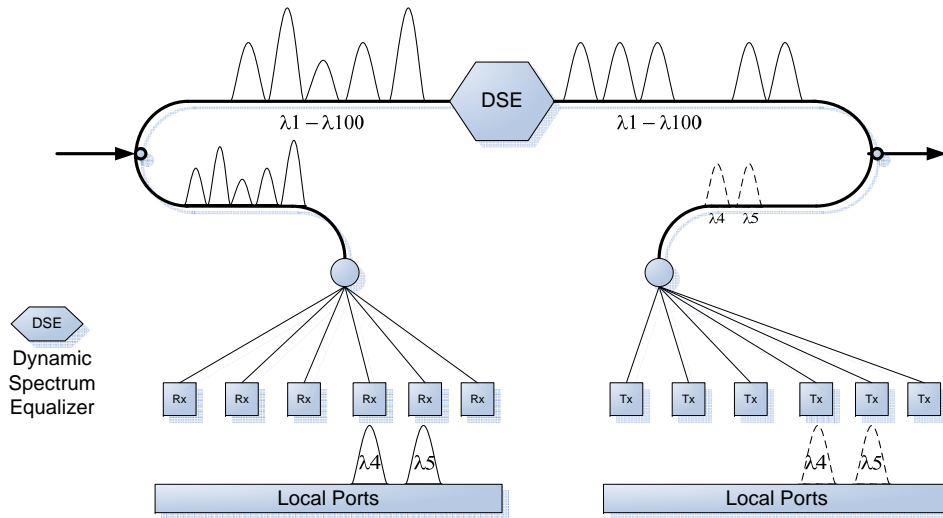


Figure 4.3: Drop-and-continue node based on broadcast-and-select architecture

the groupcast sessions.

In Chapter 3 we discussed the performance of optical groupcasting in networks with multicast-capable optical cross-connect (MC-OXC) nodes. The wavelength continuous and wavelength-conversion-capable light-forest sessions studied in Chapter 3 assumed that MC-OXC nodes are fully multicast capable. A full MC-OXC system supports directing the incoming signal to any of the outgoing links. All-optical (O-O-O) MC-OXC systems are not commercially available as of the writing of this thesis. It is expected however, that such commercial systems will be available in next few years [66]. The light-forests can also be constructed using sparse splitting capabilities of optical nodes, specifically using the Drop-and-Continue (DaC) node architectures [5]. The signal-splitting capability of an optical node in terms of the number of copies of an incoming signal that can be forwarded to outgoing links (the sparse splitting factor), dictates the multicast capability of an optical node. The DaC nodes are considered multicast incapable (MI) nodes [55]. In order to support multicasting in DaC-based networks, we need to construct a linear light-tree connecting all multicast destination nodes in a line. Intuitively it may appear that building linear

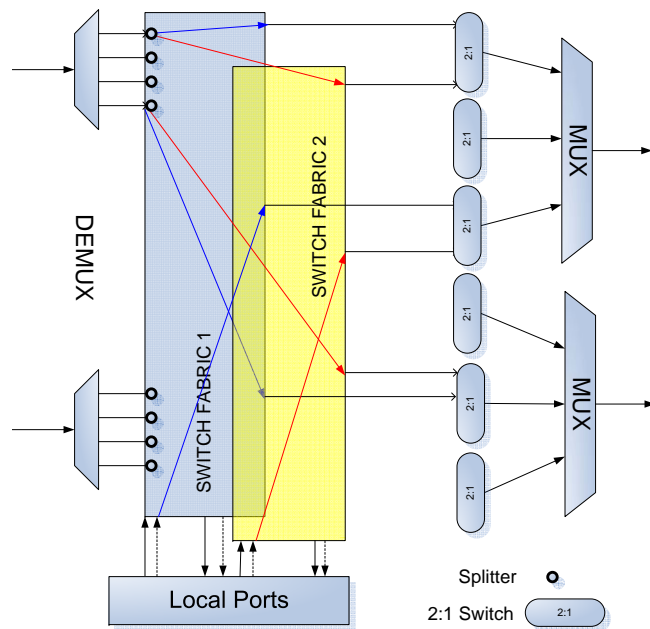


Figure 4.4: Optical add/drop multiplexer (OADM) supporting drop-and-continue capabilities

light-trees to support groupcasting will yield worse performance compared to regular light-trees, however it is quite the contrary as we will demonstrate later in this chapter [5]. Given that DaC technology is widely available as well as commercially deployed in many networks and linear light-trees are shown to exhibit improved network performance, it is important to research the advantages of using linear light-trees for optical groupcasting. In this chapter we shall investigate the groupcast performance using light-trees using both multicast-capable OXC nodes and all-optical DaC nodes.

4.2 Building Light-Forests from Light-Trees

To construct a cost-effective light-forest, we need to construct cost-effective light-trees [67]. In general, finding efficient light-trees is an NP-complete problem and so heuristics are used instead to find the light-trees. One simple heuristic to find such a

| Step | (A) MCH light-tree | (B) Linear light-tree |
|------|---|---|
| 1 | Given a group of N nodes, set of multicast vertices v_1, v_2, \dots, v_n , set v_1 as source node and $v_2 \dots v_n$ as destination nodes. | Given a group of N nodes, set of multicast vertices v_1, v_2, \dots, v_n , set v_1 as source node and $v_2 \dots v_n$ as destination nodes. |
| 2 | Find the minimum cost path from v_1 to v_i where $2 = i = n$ and list all link numbers that form the light path. (For new paths, append link numbers at the end of the list and eliminate the repeated link numbers.) | Find the minimum cost path from v_i to all destination nodes. Choose and omit the shortest next destination from the set and remove all links from its path. If a node v_y is included, then remove v_y from the destination set. |
| 3 | Repeat Step 2 until all the vertices from the set are chosen and a light-tree is created. | Repeat Step 2 until all the vertices from the set are chosen and a light-tree is created. |

Table 4.1: Algorithms to build groupcast light-trees

light-tree is the minimum cost heuristic (MCH). MCH computes the shortest distance paths from the source node to all multicast destination nodes and then collapses the common intermediate links to one lightpath (Table 4.1(A)).

For example, Figure 4.5 shows a light-forest of size 4 constructed using MCH light-trees. Next, a linear light-tree heuristic that utilizes a simple Drop-and-continue (DaC) node architecture is proposed.

4.2.1 Linear Light-trees

Light-forest sessions are difficult to accommodate as they are very resource-intensive. A Drop-and-continue (DaC) node architecture allows for the establishment of light-trees in a linear manner (“linear light-trees”), as a copy of the transiting message can be dropped at each DaC node. Because of the wavelength continuity constraint for each light-tree, linear light-trees offer considerable resource savings when they are

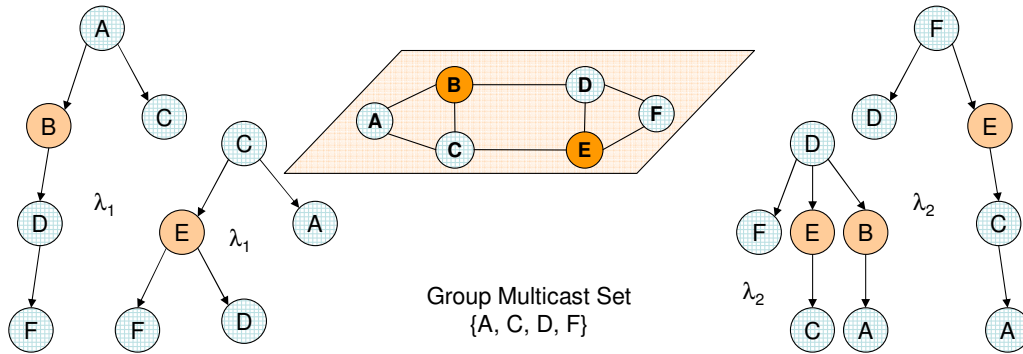


Figure 4.5: Light-forest with 4 light-trees

used to build light-forests, compared to the usage of MCH light-trees as shown below in Figure 4.5.

For group multicasting, once a linear light-tree is found, the same tree is used with different wavelength combinations for the establishment of multicast sessions for each group member. Figure 4.6 below illustrates how the linear tree is useful in saving network resources. In this example wavelength λ_1 is used by node A and D , for multicast trees $(A \rightarrow C \rightarrow E \rightarrow F \rightarrow D)$ and $(D \rightarrow F \rightarrow E \rightarrow C \rightarrow A)$ respectively. In the case of intermediate multicast nodes C and F , two linear light-trees must be constructed to serve multicast sessions $(C \rightarrow A)$ and $(C \rightarrow E \rightarrow F \rightarrow D)$ for node C and $(F \rightarrow D)$ and $(F \rightarrow E \rightarrow C \rightarrow A)$ for node F . For the six-node network shown in Figure 4.5, to build a light-forest for group A, C, F, D using the MCH light-tree approach we will utilize 17 links, whereas when we use the linear light-tree approach we will utilize 16 links (as shown in Figure 4.6).

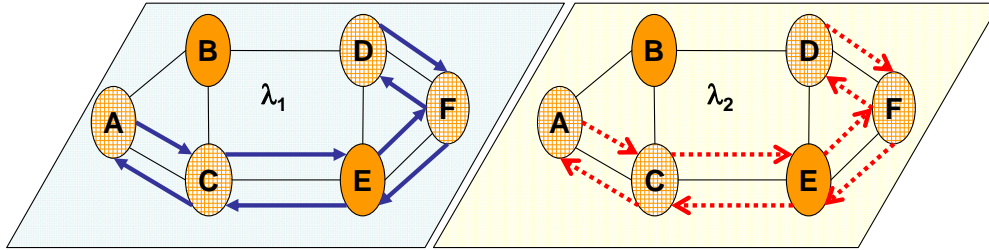


Figure 4.6: Light-forest with 4 light-trees

4.2.2 Linear Light-tree Performance Results

A group multicast is a highly distributive process, where each participating node is assumed to have a global connectivity information database. Each participating group member is capable of calculating a multicast tree taking itself as the root using one of the algorithms described in Table 4.1(A). Once the tree is calculated, the wavelength assignments for the tree are requested and published to other group members. To arbitrate wavelength contentions among light-trees, a random node sequence is generated for the multicast group. The first node generates its tree and requests optical channels for its tree, publishing them to the second node. The second node generates its tree and considers optical channels available excluding the requested channels from the first node; the third considers channels excluding the first and second trees requests. For the N th tree, wavelength channels available are only those left after $N - 1$ trees reserved their channels. If a node fails to find a tree, it publishes a null-tree which may lead to session failure. Once all channels are found to serve the light-forest, the call is admitted to the network. This distributive process holds true for the light-trees generated using the MCH approach. For linear light-tree generation,

each node generates and advertises its linear tree first without wavelength requests. When N linear trees are made available to each node, the least cost tree is chosen as the light-tree model. Wavelength requests are generated once the tree is selected and they follow the same random node sequence for request arbitration as in the MCH light-forest procedure.

To evaluate the RWA algorithms for the establishment of light-forests we simulated 50,000 calls for each light-forest size. We set 25% of the arriving calls to be group multicast of size N and the remaining were unicast calls. Call arrivals followed the Poisson distribution while the holding times for the calls were exponentially distributed with a mean of 1 unit time. The network load was set at 100 Erlangs. The unicast calls were routed using the shortest path first (SPF) algorithm and the wavelength assignments were done using the first-fit approach. For light-forests, the multicast routing followed the heuristics described in Table 4.1(A), while the wavelength assignments were also done using the first-fit approach.

In Figure 4.7, the graph shows the blocking probability for different light-forest sizes for a 14-node mesh network (shown in Figure 3.4) with 64 wavelengths per fiber direction. Note that the linear light-tree approach consistently performs better than the approach based on MCH light-trees for light-forest sizes larger than 30%.

Additional simulation results were obtained for GC-RWA heuristics for the case of MCH-based techniques used to route groupcast connections in networks with multicast capable OXCs. Table 4.2(B) demonstrates this algorithm.

The LTCG algorithm supports multicast using the optical splitters in the MC-OXCs. Wavelength conversion is not considered at the MC-OXC nodes and hence an entire light-tree must be served by a single contiguous wavelength. Using this heuristic we can find each multicast tree in $O(nk - n)$ time for a k -size groupcast and we can reserve wavelengths for the tree in $O(n)$ time. Therefore, the total time

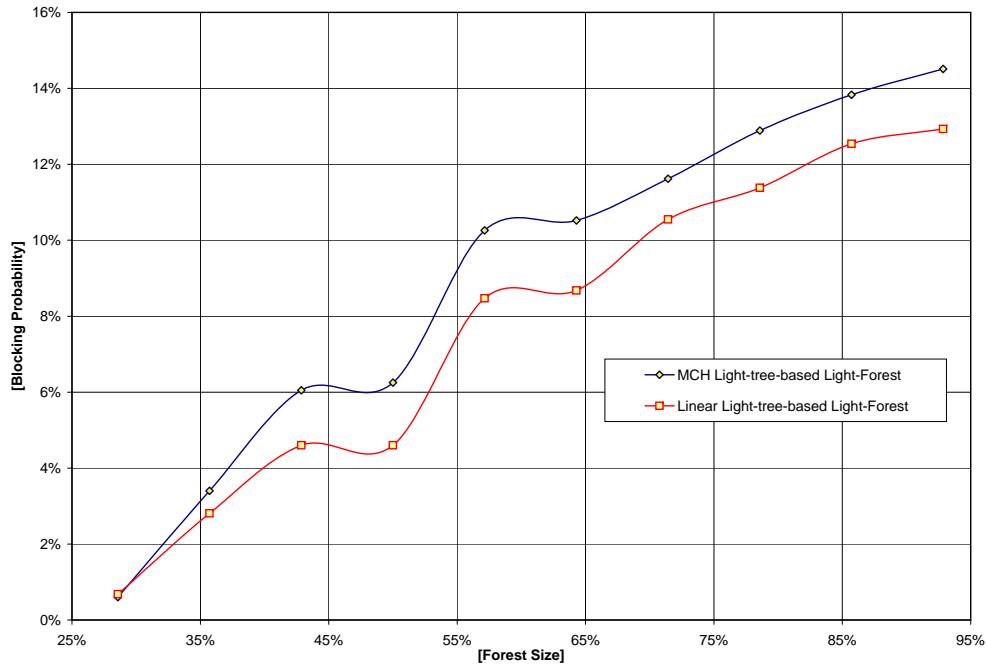


Figure 4.7: Blocking probability vs. light-forest size for different light-forest algorithms

complexity to set up each light-forest is $O(nk^2)$. It is important to note that although the time complexity for the lightpath and light-tree based heuristics using the MCH approach is similar, their performance is not. Due to heavy wavelength utilization per call for the lightpath approach, we expect it to perform poorer than the light-tree approach.

A groupcast session is a highly control plane dependent process, where each participating node must have full global connectivity information. To evaluate the MCH-based lightpath and light-tree GC-RWA algorithms for the establishment of light-forests we simulated 2,500 calls for a group size of 5. Call arrivals followed the Poisson distribution while the holding times for the calls were exponentially distributed with a mean of 1 unit time. The network load was varied from 10 to 100 Erlangs. For light-forests, the groupcast routing followed the heuristics described in Tables 4.2(A) and Tables 4.2(B), while the wavelength assignments were done using the first-fit [Section 3.2.2] approach. In Figure 4.8, the graph shows the blocking probability

| Step | (A) Lightpath groupcast (LPGC) Algorithm | (B) Light-tree groupcast (LTGC) Algorithm |
|------|---|---|
| 1 | Calculate all pairs of minimum cost paths (MCH) and store the paths in a table. | Calculate all pairs of minimum cost paths (MCH) and store the paths in a table. |
| 2 | For a groupcast call, find the shortest paths from one group participant to all other group members (set of lightpaths). Repeat the process until all the group members are served (make light-forest). | For a groupcast call, find and combine the shortest paths from one group participant to all other group members (make light-tree). Repeat the process until all the group members are served (make light-forest). |
| 3 | Find wavelengths for each lightpath till the complete light-forest is illuminated. | Find wavelengths for each light-tree till the complete light-forest is illuminated. |

Table 4.2: Routing algorithms to build groupcast light-forest

for different light-forest sizes for a 14-node mesh network 3.4 with 100 wavelengths per fiber direction. Note that the light-tree groupcast approach consistently performs better than the lightpath approach.

4.3 Static Groupcasts: Planning and Design of Sessions in MC-OXC Network

When building a light-forest for a groupcast request, the solution must be cost effective. Point-to-point lightpath-based groupcasting as established in Chapter 3 for OXC-based networks is costly and network resource intensive. Three cost variables are of interest when light-forests are considered [55]:

- The number of wavelengths needed to establish the session. Wavelengths are the most expensive resource in the network. When branches of light-trees in the forest overlap on a link, different wavelengths have to be used for each tree.

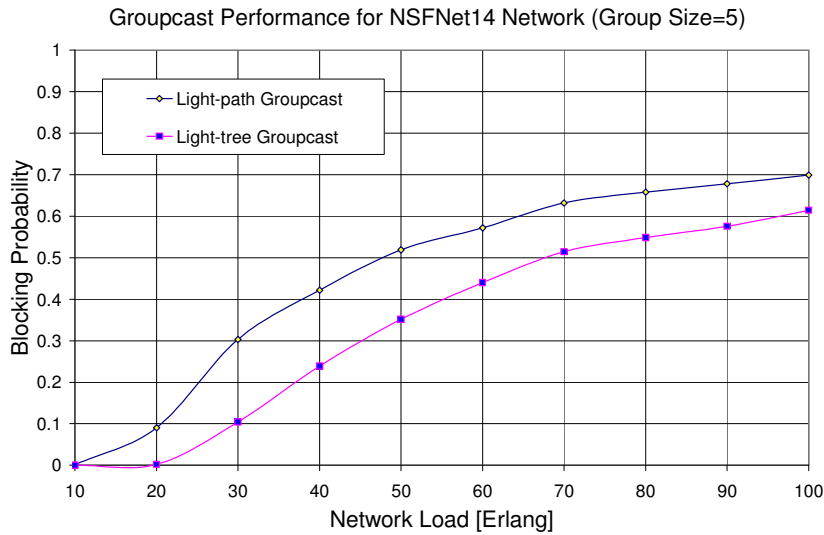


Figure 4.8: Performance of groupcast session for two different light-forest algorithms

Therefore, the number of wavelengths needed by the light-forest is simply the maximum number of overlaps of the links.

- The total number of links or branches needed to establish the session. Since the total number of branches represent total bandwidth consumed in a light-tree, the same metric could be extended to light-forests. Thus the total number of links represents the total bandwidth consumed to establish the light-forest.
- The average number of hops, or diameter of the light-forest connectivity graph. The average hops between two nodes in the groupcast session can be used for optical link budget and signal attenuation performance.

Optical groupcasting is very resource intensive as creation of light-paths soar quadratically with the linear increment of groupcast size. Similar scaling issue with multicast routing (point-to-multipoint communication) has been improved by deploying multicast-capable wavelength router (MCWR) [12]. A MCWR is essentially a multicast capable OXC (MC-OXC) switch that allows formation of light-trees by splitting the incoming light into $(1 : N)$ ways. A light-tree is a point-to-multipoint

generalization of a lightpath [4]. Thus using light-trees we expect to reduce the number of transceivers required and the average wavelength consumption per fiber-link to drop for groupcast sessions.

4.3.1 Light-Trees: Groupcast VTD Problem

In this section we study the static virtual topology design (VTD) in a network with MC-OXCs supporting groupcast traffic. The objective of this VTD problem is same as in Section (3.3.2): to set-up a finite set of groupcast sessions at a minimum cost, but using light-trees instead of lightpaths. In the MILP formulations below we consider two basic models: MC-OXCs with and without wavelength conversion capabilities.

The following inputs are assumed to the VTD problem:

- (i) $G(V, E)$: A physical graph representing the OTN network with a vertex set V and an edge set E . It contains N nodes and L bidirectional links. Each link has an assigned weight, which may correspond to the fiber distance between nodes implicitly representing propagation delay.
- (ii) W : Total number of wavelength channels carried by each fiber.
- (iii) Φ : A set of Φ groupcast sessions.
- (iv) G_n : The groupcast size for the n^{th} session¹; where $3 \leq G_n \leq |V|$.

The objective is to set up all Φ groupcast sessions on the given OTN while minimizing the cost. The cost of a groupcast session is the sum of all the link weights used to establish the groupcast.

¹Groups must have at least three nodes, but cannot include all the nodes. See section (3.3.1) for detail discussion.

4.3.2 Light-tree VTD I: Wavelength Conversion Case

Below we formulate the cost optimization problem using principles from multicommodity flow for physical routing of light-trees and traffic flows on the virtual topology:

- (i) s^i and d^i denote the *source* and *destination* nodes respectively in the i^{th} multicast session.
- (ii) m and n denote *endpoints of a physical link* that might occur in a light-tree.
- (iii) i is used as an index for light-tree session number, where $i = 1, 2, \dots, k$. To serve a groupcast session of size x , we need x light-trees. We therefore refer to a groupcast session of size x as a **light-forest** [5] with x light-trees. To establish Φ light-forests, we need to set-up k light-trees, where $k = \sum_{n=1}^{n=\Phi} G_n$.

We consider an OTN with all optical (MC-OXC) switches equipped with wavelength converters.

- **Given:**

- N : Number of nodes in the network.
- W : Number of wavelengths per fiber.
- P_{mn} : Number of fibers connecting node m to node n . If there is no fiber link, then $P_{mn} = 0$, $P_{mn} \in \mathbb{Z}$.
- w_{mn} : Weight or cost associated with a link². As the graph $G(V, E)$ is undirected, $w_{mn} = w_{nm}$, and if $P_{mn} = 0$ then $w_{mn} = \infty$.
- Φ : Total number of light-forest sessions.

²Weights render same meanings as in Section (3.3.1)

- S_i : Session representing a light-tree for $i = 1, 2, \dots, k$. For a given light-forest j , there are G_j nodes forming $|G_j|$ light-trees. Each light-tree session (S_i) has a source node and a set of destination nodes denoted by $\{s^i, d^1, d^2, \dots, d^y\}$ where $y = |G_j| - 1$. The cardinality of the light-tree is represented by G_i which is essentially the forest size.

- **Variables:**

- V_{mn}^i : This is a boolean variable which is equal to one if the optical link between nodes m and n is in-use by a lightpath that is part of the light-tree (multicast) session i , zero otherwise.
- O_x^i : This is a boolean variable which is equal to one when a node x belongs to an optical session i (node x is the source, the destination or an intermediate MC-OXC node serving a light-tree for the session S_i).
- F_{mn}^i : This is an integer commodity flow variable. Each destination node for a session needs one unit commodity. To serve a G_i sized light-tree session, $(G_i - 1)$ units of commodity flow are needed. F_{mn}^i also represents the number of destination nodes in session S_i downstream from the link between nodes m and n .

- **Objective:** To optimize the total cost of all groupcast sessions:

$$\text{Minimize} : \sum_{i=1}^{i=k} \sum_{m,n} w_{mn} \cdot V_{mn}^i \quad (4.1)$$

Equation (4.1) is a linear objective function as it adds up the cost of individual light-trees to serve all light-forest sessions.

- **Constraints:**

- Point-to-Multipoint connectivity constraints:

$$\forall i, \forall j \in S_i : O_j^i = 1 \quad (4.2)$$

$$\forall i, \forall n \neq s^i : \sum_m V_{mn}^i = O_n^i \quad (4.3)$$

$$\forall i : \sum_m O_{ms^i}^i = 0 \quad (4.4)$$

$$\forall i, m : \sum_n V_{mn}^i \leq D_p(m) \cdot O_m^i \quad (4.5)$$

$$\forall m, n : \sum_i V_{mn}^i \leq P_{mn} \cdot W \quad (4.6)$$

$$\forall i, \forall m \neq d^i : \sum_n V_{mn}^i \geq O_m^i \quad (4.7)$$

– Commodity-flow constraints:

$$\forall i, \forall m \neq S^i : \sum_n F_{nm}^i = \sum_n F_{mn}^i \quad (4.8)$$

$$\forall i, \forall m = s^i : \sum_n F_{ns^i}^i = 0 \quad (4.9)$$

$$\forall i, \forall m = d^i : \sum_n F_{nm}^i - \sum_n F_{mn}^i = 1 \quad (4.10)$$

$$\forall i, \forall m = s^i : \sum_n F_{s^i n}^i = (G_i - 1) \quad (4.11)$$

$$\forall i, \forall m, \forall n : F_{mn}^i \leq V_{mn}^i \cdot (G_i - 1) \quad (4.12)$$

$$\forall i, \forall m, \forall n : V_{mn}^i \leq F_{mn}^i \quad (4.13)$$

– Light-forest constraint:

$$\forall i, \forall m, \forall n : F_{mn}^i < G_i \quad (4.14)$$

- **Explanation of Equations:** The light-tree creation constraints are more involved than the lightpath constraints shown in MILP formulations from Section 3.3.2. However, the essence of the equations are the same: establish source

to destination routes and then let the commodities flow from source to destinations. Equation (4.2) states that every source and destination nodes in a session are part of a light-tree. Equation (4.3) ensures that every node that is part of a light-tree has an incoming node except the source. To isolate sources from receiving their own signals, equation (4.4) eliminates lightpaths back into the sources. Any intermediate MC-OXC should not split the incoming signal into copies greater than its nodal degree and equation (4.5) captures that constraint. Equation (4.6) restricts the number of lightpaths between nodes m and n to atmost the maximum number of wavelengths per fiber for either direction of the link. Finally all participating nodes in a light-tree must have at least one outgoing link except the destination nodes, which is captured by equation (4.7). The commodity-flow constraints are also straightforward. Equation (4.8) states that the sum of incoming flows to a transient multicast node must be equal to the sum of exiting flows except for the session set which contains source and destination nodes. Equation (4.9) clarifies that there are no flows going into the source nodes. Equation (4.10) ensures that there should be exactly one flow terminating at each destination node. As for the originating flows, equation (4.11) states that each source node must output $(G_i - 1)$ flows. Finally Equations (4.12) and (4.13) state that every link not occupied by a session has no flow on it and every session setup for a light-tree must have positive flows, respectively. Light-trees in a light-forest have an additional constraint (Equation 4.14): the number of flows for a light-tree session handled by a link must be less than the light-forest size.

- **Number of Variables and Constraints:** We need $kN + 2kN^2$ variables, where $k = \sum_{n=1}^{n=\Phi} G_n$. The equation grows linearly with the number of light-forest sessions (Φ) and the average forest size (\bar{G}), and it grows quadratically with the number of nodes in the network. As for the constraints, they are

bounded by $O(\Phi \cdot \bar{G} \cdot N^2)$ showing once again a second order relationship with the number of network nodes. For the VTD problem, the light-tree based MILP formulation is smaller than the lightpath based MILP formulation (Sec. 3.3.1). The number of variables increase by a $(\bar{G} - 1)$ factor which could be quite significant for large number of light-forests.

4.3.3 Light-tree VTD II: Wavelength-Continuity Case

If MC-OXCs are not capable of wavelength conversion, then each light-tree must use a single color wavelength. To ensure that the wavelength-continuity constraint is met, additional variables and constraints are introduced.

Assuming all the same parameters as described in Section (4.3.2), we consider an OTN with (MC-OXC) switches that have no wavelength conversion capabilities.

- **Given:**

- N : Number of nodes in the network.
- W : Number of wavelengths per fiber.
- P_{mn} : Number of fibers connecting node m and node n . If there is no fiber link, then $P_{mn} = 0$, $P_{mn} \in \mathbb{Z}$.
- w_{mn} : Weight or cost associated with a link³. As the graph $G(V, E)$ is undirected, $w_{mn} = w_{nm}$, and if $P_{mn} = 0$ then $w_{mn} = \infty$.
- Φ : Total number of light-forest sessions.
- S_i : Session representing a light-tree for $i = 1, 2, \dots, k$. For a given light-forest j , there are G_j nodes forming $|G_j|$ light-trees. Each light-tree session (S_i) has a source node and a set of destination nodes denoted by

³Weights render same meanings as in Section (3.3.1)

$\{s^i, d^1, d^2, \dots, d^y\}$ where $y = |G_j| - 1$. The cardinality of the light-tree is represented by G_i which is essentially the forest size.

• **Variables:**

- V_{mn}^{ic} : This is a boolean variable which is equal to one, if the optical link between nodes m and n is in use by a lightpath of color c for the light-tree (multicast) session i , zero otherwise.
- O_x^i : This is a boolean variable which is equal to one when a node x belongs to an optical session i (node x is the source or the destination or the intermediate MC-OXC node serving a light-tree for the session S_i), 0 otherwise.
- F_{mn}^i : This is an integer commodity flow variable. Each destination node for a session needs one unit commodity. To serve a G_i sized light-tree session, $(G_i - 1)$ units of commodity flow are needed. F_{mn}^i also represents the number of destination nodes in session S_i downstream from the link between nodes m and n .
- C_c^i : Color of the wavelength. $C_c^i = 1$ if session i is on wavelength color c , otherwise $C_c^i = 0$.

• **Objective:** To optimize the total cost of all groupcast sessions:

$$\text{Minimize : } \sum_{i=1}^{i=k} \sum_{c=1}^{c=W} \sum_{m,n} w_{mn} \cdot V_{mn}^{ic} \quad (4.15)$$

Equation (4.15) is a linear objective function that adds up the cost of individual colored light-trees to serve all light-forest sessions.

• **Constraints:**

– Point-to-Multipoint connectivity constraints:

$$\forall i, \forall j \in S_i : O_j^i = 1 \quad (4.16)$$

$$\forall i, \forall n \neq s^i : \sum_{m,c} V_{mn}^{ic} = O_n^i \quad (4.17)$$

$$\forall i : \sum_m O_{ms^i}^i = 0 \quad (4.18)$$

$$\forall i, m : \sum_{n,c} V_{mn}^{ic} \leq D_p(m) \cdot O_m^i \quad (4.19)$$

$$\forall m, n : \sum_{i,c} V_{mn}^{ic} \leq P_{mn} \cdot W \quad (4.20)$$

$$\forall i, \forall m \neq d^i : \sum_{n,c} V_{mn}^{ic} \geq O_m^i \quad (4.21)$$

– Commodity-flow constraints:

$$\forall i, \forall m \neq S^i : \sum_n F_{nm}^i = \sum_n F_{mn}^i \quad (4.22)$$

$$\forall i, \forall m = s^i : \sum_n F_{ns^i}^i = 0 \quad (4.23)$$

$$\forall i, \forall m = d^i : \sum_n F_{nm}^i - \sum_n F_{mn}^i = 1 \quad (4.24)$$

$$\forall i, \forall m = s^i : \sum_n F_{s^i n}^i = (G_i - 1) \quad (4.25)$$

$$\forall i, \forall m, \forall n : F_{mn}^i \leq \sum_c V_{mn}^{ic} \cdot (G_i - 1) \quad (4.26)$$

$$\forall i, \forall m, \forall n : \sum_c V_{mn}^{ic} \leq F_{mn}^i \quad (4.27)$$

– Wavelength-continuity constraint:

$$\forall i : \sum_c C_c^i = 1 \quad (4.28)$$

$$\forall i, \forall c, \forall m, n (n > m) : V_{mn}^{ic} + V_{mn}^{ic} < C_c^i \quad (4.29)$$

- **Explanation of Equations:** The monocolor light-tree concept introduces new wavelength-continuity constraints to the previously stated set of equations from Section 3.3.2. When routing calls, we need to consider wavelength color and therefore introduced the color variable to Equations (4.17), (4.19), (4.20) and (4.21). The commodity-flow constraints are mostly the same (Equations (4.22-4.25)). However Equations (4.26) and (4.27) need to be wavelength selective when identifying every link not occupied by a session to remove flow from and ensuring that every session setup for a light-tree has positive flows, respectively. Finally, wavelength-continuity constraints have to be respected. The litmus test for the color, Equation (4.28) enforces the constraint that each light-tree is of single color. Equation (4.29) further ensures that the wavelength not chosen during a session is not occupied and the links of a lightpath are on the same wavelength.
- **Number of Variables and Constraints:** We need $2kN + kN^2 + kWN^2$ variables, where $k = \sum_{n=1}^{\Phi} G_n$. The equation grows linearly with the number of forest sessions (Φ) and the average forest size (\bar{G}), and the number of wavelengths selected. However, the number of variables grows quadratically with the number of nodes in the network. As for the constraints, they are bounded by $O(\Phi \cdot \bar{G} \cdot N^2)$ showing once again second order relationship with the number of network nodes. For the VTD II problem, the light-tree based MILP formulation is smaller than the lightpath based MILP formulation (Section 3.3.3). The number of variables increase by a factor $(\bar{G} - 1)$ which could be quite significant for large number of light-forests.

4.3.4 Network Design Problem utilizing MC-OXCs: Heuristic Approach

In this section we study the network design problem (NDP) for multicast-capable optical cross-connect (MC-OXC) networks. The problem is defined as follows: Given the physical network topology, a specified number of wavelengths on each fiber, and the traffic matrix (groupcast or mixed unicast and groupcast connections), find the routes and the wavelength assignment of the incoming sessions such that the total traffic demand cost is optimized.

In a network with no wavelength conversion, the wavelength continuity constraint has to be respected when routing point-to-multipoint traffic. A groupcast session of size k will generate k -multicast requests and each session is served by a single continuous wavelength. For routing purpose, each wavelength layer is presented as a graph, thus forming W -graphs for W -wavelengths per fiber network. For groupcast session size k , k -multicast trees are generated. When we have n -groupcast sessions to plan for, we expand each groupcast session to multiple multicast sessions and form a new session list S . Each multicast session is then routed and assigned a wavelength using the heuristic described in Table 4.3.

As indicated in **Step 4** of Table 4.3, we can optimize the session list based on the destination group size or the multicast-tree cost. Additionally, the list can be arranged in an ascending or descending order. Thus, the five heuristics used are: (1) no sorting of the sessions (RANDOM) (2) sorting the sessions based on cost in ascending order (COST (ASC)) (3) sorting the sessions based on cost in descending order (COST (DES)) (4) sorting the sessions based on the destination group size in ascending order (SIZE (ASC)) sorting the sessions based on the destination group size in descending order (SIZE (DES)). Figure 4.9 summarizes the results obtained using the cost optimization heuristics. Each set of results is grouped based on the

| Cost Optimization Heuristic for Groupcast in MC-OXC Network | |
|---|--|
| Step 1 | Given a groupcast session list with n -session, select the first groupcast session. |
| Step 2 | For the groupcast size k , generate k -multicast-trees using Steiner-tree minimum cost heuristic. |
| Step 3 | Index each tree found in Step 2 in terms of source (s), destinations (d), cost (c) and wavelength (w) in the session list $S_i = \{s_i, d_i^1, d_i^2, \dots, d_i^k, c_i, w_i\}$. |
| Step 4 | Repeat Step 2 and Step 3 until all of the n -groupcast sessions are served. Sort the sessions in S based on destination group size or session cost (if specified otherwise). Go to the next step. |
| Step 5 | Select the first session from the list S . |
| Step 6 | Choose the first wavelength (W_i) from the W -wavelength list and construct a graph G' based on the wavelength availability per link. If the wavelength is unavailable the link is removed from the original physical graph G . |
| Step 7 | Route the S_i -session on the graph G' from Step 6 using Steiner-tree minimum cost heuristic. For successful routing, assign the wavelengths to the links selected during multicast-tree construction and update the cost c_i in S_i list. If the multicast routing fails, increment the wavelength (W_{i+1}) and go to Step 6 . |
| Step 8 | Repeat Step 6 and Step 7 until all the sessions from list S_i are served. |
| Step 9 | Compute the total cost of the sessions using $S_{cost} = \sum_{i=1}^{i= S } c_i$. |

Table 4.3: Light-tree based groupcasts with no wavelength conversion

total number of groupcast sessions (n) which in turn generates k -multicast sessions. In order to establish the cost ratio of the sessions to the network we divide the total session cost by the total network cost, which is the sum of all link costs in the network (multiplied by two to capture the fact that the links are bi-directional) and then multiplied by the total number of groupcast sessions. For example, the total network cost is 6330 for the NSF network [Figure 3.4].

As it can be seen from the plots in Figure 4.9, the sorting of sessions in ascending order based on multicast group size (SIZE(DES)) gave the best performance results for the large groupcast sessions ($n \geq 30$). However, the sorting of sessions in descending

order based on multicast group size exhibited the worst performance in most cases (for $n = 30$, COST(DES) was the worst performer). Although sorting the list based on cost sizes seemed to be an interesting proposition, however this heuristic never yielded good results except for one groupcast session (for $n = 10$, COST(DES) was the best performer). The RANDOM heuristic exhibited random behavior pattern.

| Session No. | Group Participants | Group Size |
|-------------|--|------------|
| 1 | { CO, CA1, WA, PA } | 4 |
| 2 | { GA, TX, MD, IL, CA1 } | 5 |
| 3 | { PA, WA, MI, TX } | 4 |
| 4 | { CO, GA, NE, PA, NY, IL, CA1 } | 7 |
| 5 | { PA, NE, CA2, NJ, GA, NY, WA, TX, IL } | 9 |
| 6 | { NE, UT, IL, PA, TX, CA1 } | 6 |
| 7 | { CO, IL, MD, MI, GA, CA2, NY, CA1, UT, TX, NE, WA } | 12 |
| 8 | { WA, MI, GA } | 3 |
| 9 | { NE, MD, CA1, MI, NY, GA, UT, WA, CA2, PA } | 10 |
| 10 | { MD, TX, PA, WA, CA2, IL, CO, NE, CA1, UT, NJ } | 11 |
| 11 | { NY, CA2, IL, NJ, MD, CO, NE, MI } | 8 |
| 12 | { CA1, MI, NY } | 3 |
| 13 | { UT, MD, WA, NY, NJ, PA, CA1, CO, GA } | 9 |
| 14 | { PA, NY, GA, TX, UT, CO, IL, NE, CA1, MD, CA2 } | 11 |
| 15 | { MI, GA, PA, CA2, CA1, UT } | 6 |
| 16 | { PA, NJ, CA1, NE, IL, CO, CA2, NY, WA, GA, MD } | 11 |
| 17 | { NE, CO, GA, WA, MI, TX, CA1, UT, NY, NJ, PA, IL } | 12 |
| 18 | { CO, NJ, MI, NE, TX, CA1, CA2, IL, GA, NY } | 10 |
| 19 | { CA1, WA, NE, CA2 } | 4 |
| 20 | { NE, CA2, TX, IL, CO } | 5 |

Table 4.4: List of 20 groupcast sessions for the NSF14 network [Figure 3.4]

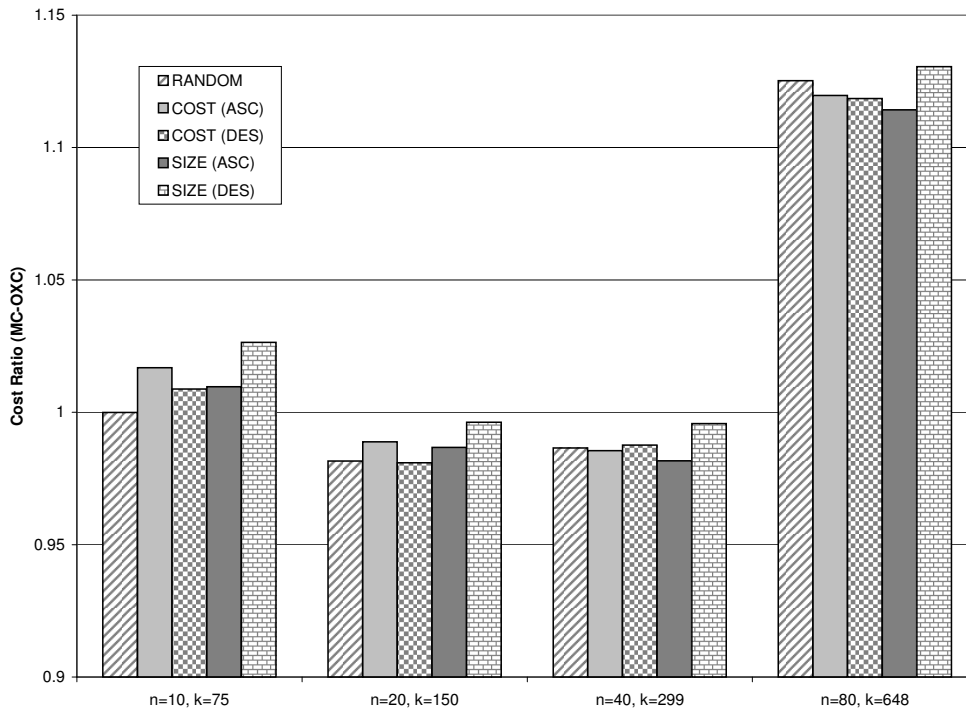


Figure 4.9: Heuristic comparison for groupcast sessions in the NSF14 network [Fig. 3.4]

4.4 Static Groupcasts: Planning and Design of Sessions in DaC Networks

4.4.1 Data Plane: Groupcast Session Size

We investigate the design challenges regarding the groupcast session size (group size) when DaCs are used to establish groupcast sessions. The linear multicast light-tree in terms of the number of drop ports may be optical-power limited. The problem of building light-trees under optical power constraints has been studied in [68], [18] and [69]. The solutions proposed place an upper bound on the number of times the optical signal may be split. In [69] the authors emphasize the importance of establishing balanced light-trees as algorithms creating highly unbalanced light-trees result in degrading performance.

It is well accepted that cross-layer optimization (e.g., routing with physical layer constraints) such as power-constraint routing and wavelength assignment will yield superior network performance. However, this approach is impractical, as most optical networks are build with some design constraints, such as node locations and optical link (power) budgets. The routing and wavelength assignment and higher layer protocols evolve with time, however physical layer components (optical fibers and optical nodes) are not continuously replaced in order to accommodate improved RWA algorithms or newer Layer 2 or Layer 3 protocols. Therefore a more pragmatic way to approach the design and planning of linear light-tree groupcast networks, is to ascertain layer-agnostic and layer-independent network designs.

Optical-Power Constraints

A linear light-tree size (number of branches or drop-offs) in all-optical DaC-based networks is determined by the received optical-power (assuming that no signal regeneration is taking place at intermediate nodes). For the linear light-tree size to be viable, the last destination node must detect sufficient power, satisfying the following equation:

$$P_{Source} - P_{Attenuation} - P_{DaC} \geq P_{Receiver} \quad (4.30)$$

In Equation 4.30 the source power is represented by P_{Source} and the receiver sensitivity represented by $P_{Receiver}$ is the minimum signal power required to detect the signal. Although ideally optical amplifiers can be utilized to compensate losses, we must nevertheless consider networks with no optical amplifiers or regenerators, where Equation 4.30 places an upper bound on the number of drops including the final lightpath terminating point. Thus, if we wish to drop the signal at n -DaC nodes (not including the final destination) we need to solve the following equation:

$$P_{Source} - n \times (P_{Attenuation} + P_{DaC}) \geq P_{Receiver} \quad (4.31)$$

where P_{DaC} represents the power attenuation when the signal passes through a DaC-based node.

Typical attenuation figures in Equation 4.31 should include coupling losses (1dB per connector) [70] splice losses (0.1dB/splice) and fiber attenuation. The fiber attenuation varies from 0.25dB/km [70] to 2dB/km [15] depending on wavelength and fiber type. The following example will illustrate how to establish an upper bound on the number of drops for a session. For example, we assume the following parameters:

- Maximum power available from transmitter source equals 1mW [70] ($P_{Source} = 1mW$).
- 10 wavelengths/fiber are available and the bandwidth is 10 Gbps for each wavelength ($P_{Source} = -10dBm$).
- Minimum power required for signal detection at the receiver is -30dBm [70] ($P_{Receiver} = -30dBm$).
- Fiber attenuation is 0.25dB/km [70] and the average distance between two nodes is 4 km.
- In the DaC architecture shown in Figure 4.2, assume that 5% of the input power is tapped at the input ports. The 5% tap at the DaC nodes incurs a coupling loss of 0.1 dB [20].

Using Equation 4.31 we have $-10dBm - n \times (1 + 0.1 + 0.2)dB \geq -30dBm$, which leads to $n \leq 15.3$. With this simple example we demonstrate that we can have up to

15 DaC nodes, on average 4 km apart, before the final destination of a linear light-tree multicast session.

It is important to mention that the optical signal-to-noise-ratio (OSNR) is a very critical parameter for certain transmission systems that could limit the optical multicast size. It is assumed that the OSNR value at the optical line terminating equipment (LTE) is well within the required bit error rate (BER) for high speed optical communication systems.⁴

In fact, for thermal-noise-limited systems, the noise figure F is given in terms of the ratio of input optical amplifier signal-to-noise ratio to output amplifier's signal-to-noise ratio [70] as shown in the equation below.⁵

$$F = \frac{(S/N)_{in}}{(S/N)_{out}} \quad (4.32)$$

The EDFA systems deployed in cascaded chain as shown in Figure 4.10 experience loss of signal power between successive amplifiers. The fractional transmission loss α_k lowers the signal-to-noise ratio between amplifiers. The final noise figure for a cascaded amplifier chain is given by Equation 4.33.

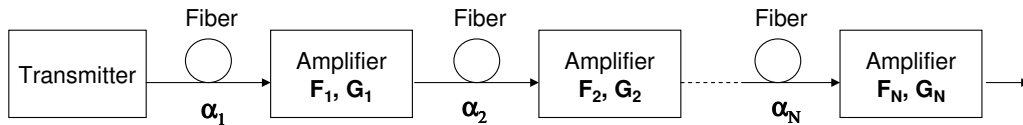


Figure 4.10: Optical-amplifiers cascaded in a chain

⁴Typically Erbium-Doped Fiber Amplifiers (EDFA) are used at each optical cross-connect (including DaCs) to amplify the optical signals [15]. The EDFA amplifies the optical signal as well as the noise due to amplified spontaneous emission and hence imposes a lower limit on the power received at the photodetector. In general, EDFAs can reduce the SNR power by 3dB [70] and given typical relative intensity noise (RIN) value of -130dB/Hz, a 10GHz channel will need OSNR_≤30dB.

⁵For nonthermal noise limited systems, the effect of amplifier noise on the SNR must be computed individually. Shot-noise-limited systems using Avalanche Photodiodes (APDs) or heterodyne receivers do not suffer from the amplifier's SNR effect. [70]

$$F = \frac{F_1}{\alpha_1} + \frac{F_2}{\alpha_1 G_1 \alpha_2} + \dots + \frac{F_N}{\alpha_N \prod_{i=1}^{N-1} (\alpha_i G_i)} \quad (4.33)$$

The online fiber amplification benefits multicast signals especially for DaC-based networks. However, for unicast traffic it is better to have pre- and post-amplification at the node level to readjust the signal strength and to minimize the EDFA amplified spontaneous emission. The authors in [18] investigate the optimal placement of optical amplifiers for given optical multicast sessions. The optimum placement of amplifiers for mixed-type traffic is an open research topic.

Considering backbone networks such as NSF14 (Figure 3.4) where the average distance between nodes is 1,221 km, the optical amplifiers must be employed at a minimum distance of 600 km and a maximum distance of 3,000 km, in order to be able to compensate for such large fiber attenuation. It is clear that optical amplifier-free networks may be considered for Metropolitan Area Networks (MANs), however amplifier free wide area networks (WANs) are impractical and will not be considered for groupcasting in DaC-based all-optical groupcast network design.

4.5 Dynamic Groupcast: Performance Analysis

4.5.1 Light-forest with mixed loads

First we investigate the performance of GC-RWA under increased load. We observe the consistent behaviour of light-forest sessions as we increase the sessions from 25% with 75% unicast traffic initially to 100% groupcast traffic only (final). The plot in Figure 4.11 shows performance of four varied traffic load profiles. Note that the minimum-cost heuristic was used for light-tree setup.

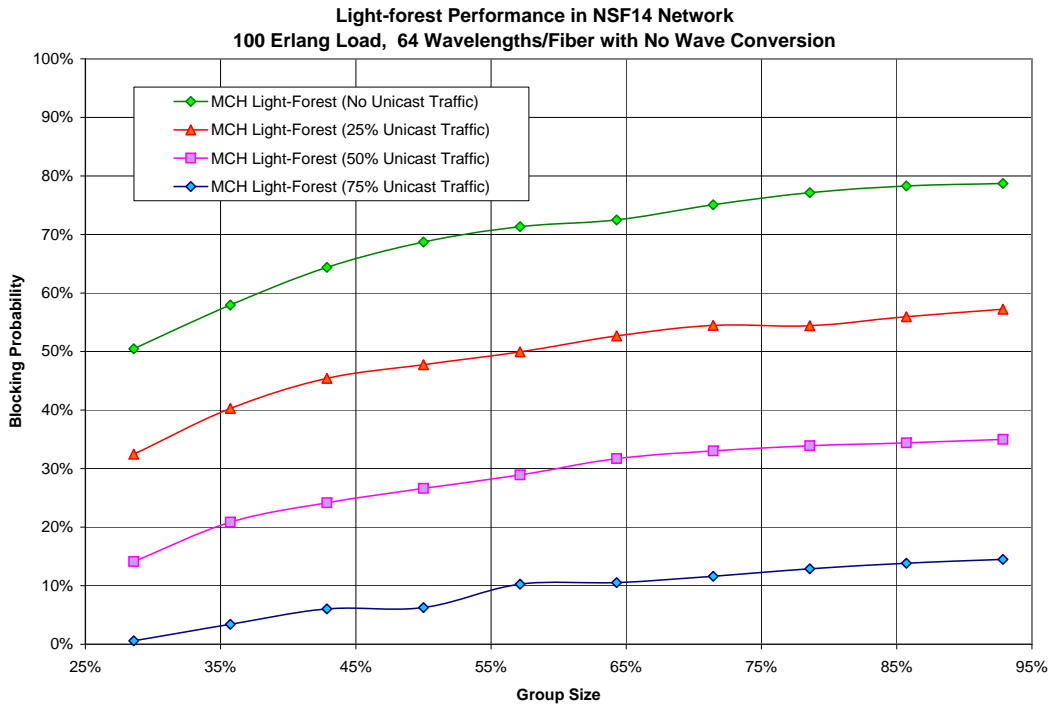


Figure 4.11: Performance of groupcast sessions under integrated traffic load (groupcast and unicast)

4.5.2 Linear light-tree forest

For DaC-based linear-tree algorithms, next we investigate the routing influence on the overall groupcast performance. We used four approaches to calculate the routing namely: (i) shortest-path from one destination to the next in random order (random); (ii) sequenced manner to traverse the nodes based on their IP addresses; (iii) choose the closest node first (iv) closest node first based on the hop distance (not the link cost). The shortest hop algorithm outperformed the rest of the three routing algorithms, except for sequenced algorithm, which showed good results for larger groupcast sessions. Figure 4.12 shows the DaC-based groupcast performance for various linear-tree routing algorithms.

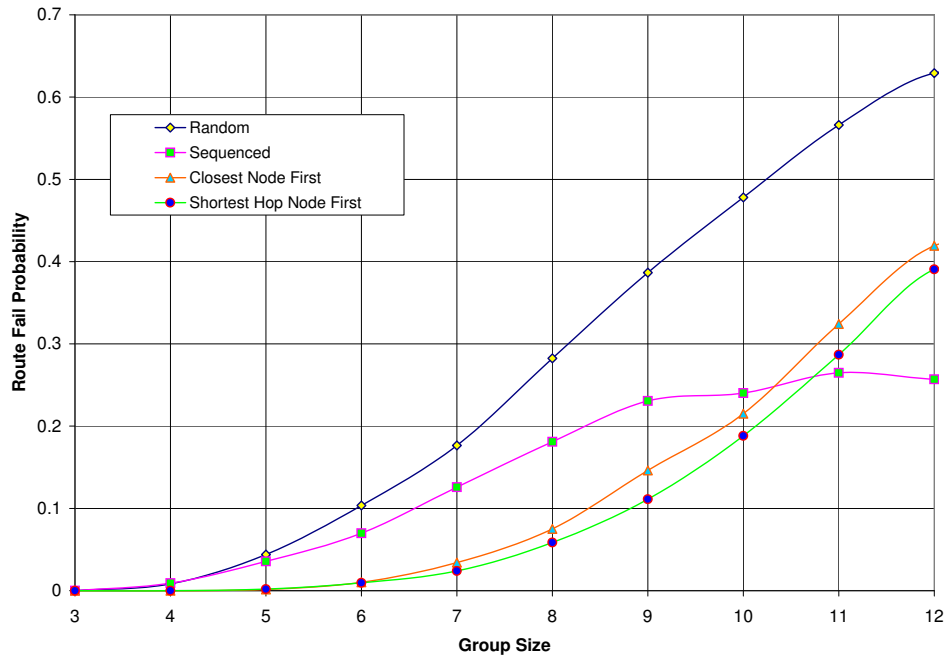


Figure 4.12: Performance of DaC-based groupcast sessions for different linear-tress routing approaches

4.5.3 Fairness Issue

As we admit more groupcast calls into the network, we observe that the groupcast calls are blocked discriminantly because of its heavy network resource requirements compared to unicast calls. For example in Figure 4.13 the average blocking of the network conceals the fact that the groupcast blocking is quite high compared to unicast traffic in the same network. Blocking has to be fairly distributed for both unicast and multicast traffic. Thus network planners should put in reserved resources to support planned groupcast session during heavy traffic condition, otherwise may suffer significant degradation of groupcast performance.

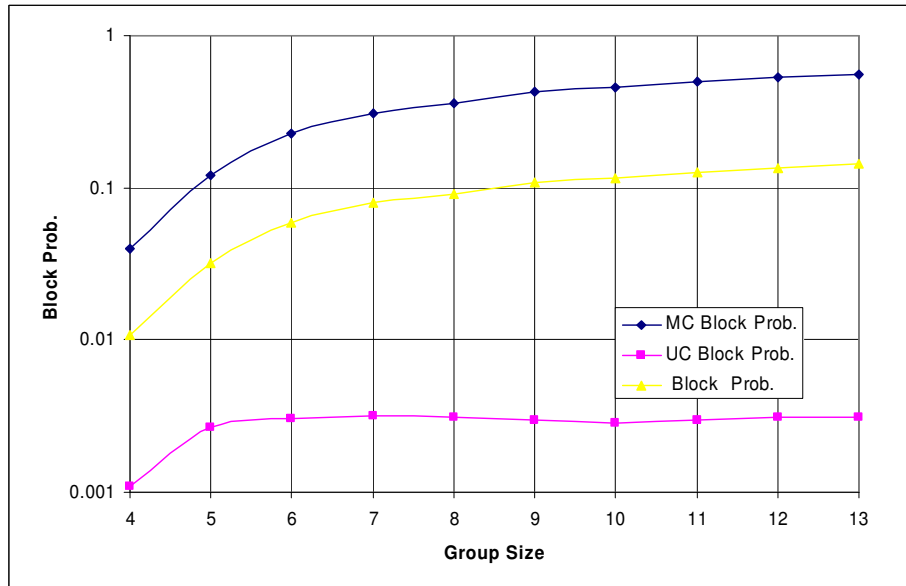


Figure 4.13: Performance of MC-OXC-based groupcast sessions in NSF Figure 3.4 network with 64 Channel, 10,000 Calls, MCH Light-forest

4.5.4 Influence of Network Resources (Wavelengths) on Light-forest Performance

In this section we investigated the relationship between the network resource allocation and the groupcast performance. In the first experiment, we let the number of wavelengths to be proportionate to the number of nodes in the network and increased the the network resources by a constant number. The plot in Figure 4.14 shows the performance of the groupcast sessions in the NSF network. We observe the improvement for increased network allocation of wavelengths/fiber.

Next we varied the network allocation ratio with respect to the number of links in the network and plotted the results in Figure 4.15. We observed no direct correlation between the network resources allocated for the groupcast session and the size of its

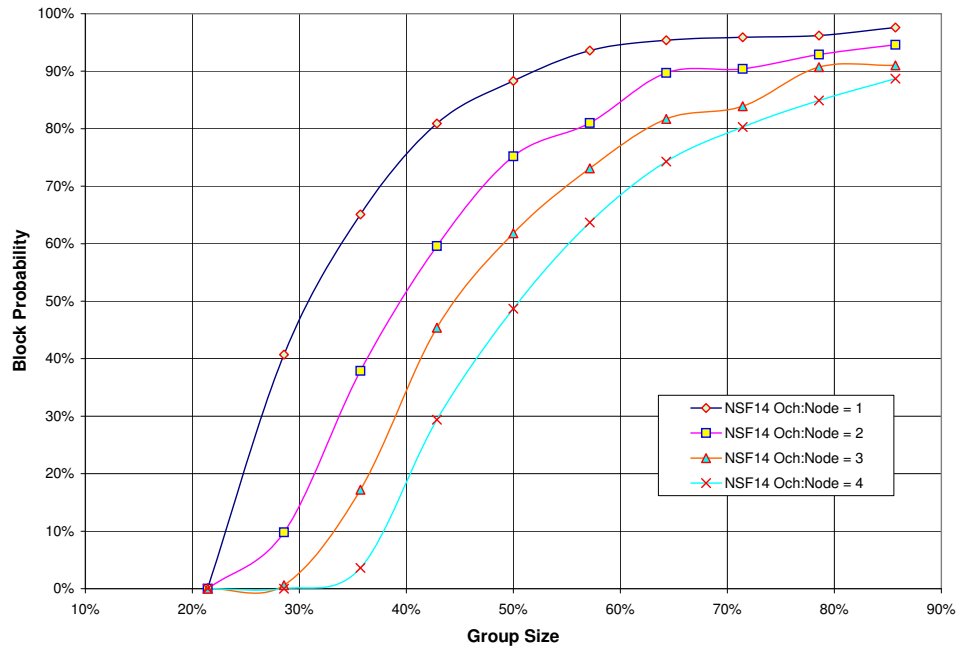


Figure 4.14: Performance of MC-OXC-based groupcast sessions vs. group size network (in terms of the number of links or nodes).

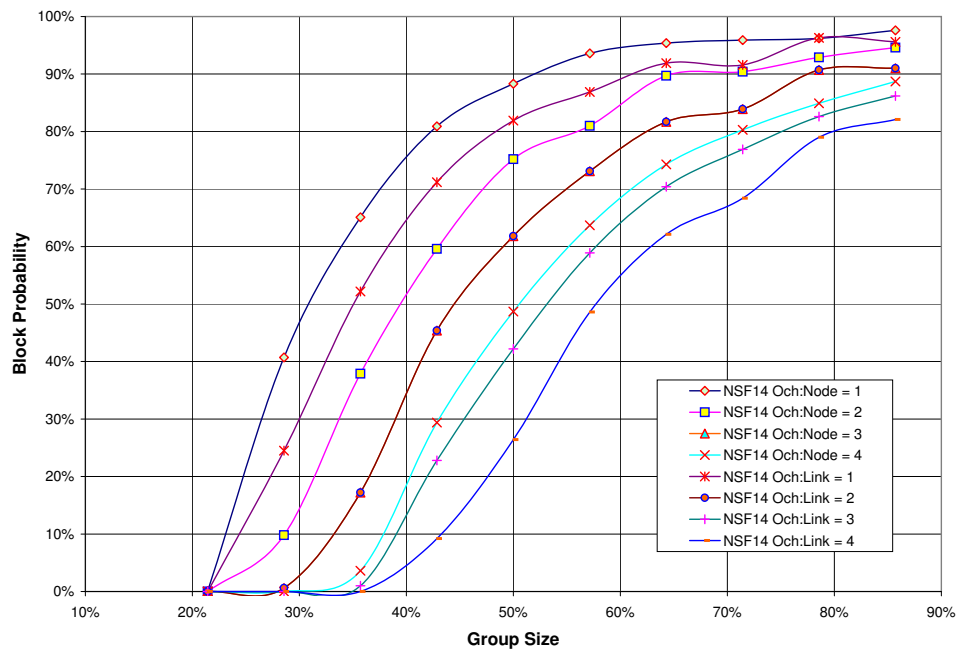


Figure 4.15: Performance of MC-OXC-based groupcast sessions vs. group size (network resource variation)

Chapter 5

Conclusions and Future Directions

5.1 Summary of Research

The main focus of this dissertation was to derive provisioning and network design techniques as well as evaluate the performance of these techniques for WDM-based wavelength-routed optical networks with mixed-type of traffic (including unicast, multicast and groupcast traffic). The mixed traffic can be categorized into two major categories in particular: multicast traffic and groupcast traffic in the presence of unicast traffic.

The basic motivation behind the proposed work was discussed in Chapter 1 and included such reasons as the emergence of several applications in today's networks (such as video on demand, multiparty conferencing, teaching via the web, grid computing, etc) that will be the driving force behind the increase of multicast and groupcast traffic in the near future.

All possible multicast node architectures that can be used to establish multicast and/or groupcast sessions are investigated in Chapter 2 and a short literature review

is provided on multicast routing and signaling as well. A novel protection approach for protection of multicast sessions is included in the discussion of Chapter 2. This approach, namely the minimum cost collapsed ring (MC-CR) algorithm, finds a ring path among the source and all destination nodes. Unlike SONET, it is not required to terminate the signal at the generating source node. Hence, two lightpaths are established, each starting in opposite directions from the source and terminating at the final node of the destination set after traversing the ring path. The MC-CR technique is shown to require less resources than the dedicated protection approach using the MCH algorithm to protect the multicast sessions.

Chapter 3 dealt specifically with the case of planning and provisioning of groupcast traffic using lightpaths. Performance of groupcast algorithms was evaluated in the presence of mixed (unicast and groupcast) traffic. The chapter gave a description of node architectures that can be used for the provisioning of groupcast sessions that are lightpath-based and then investigated groupcast routing and wavelength assignment in two contexts: (a) Determined the wavelength requirements to support groupcast of a certain group size in a given network; (b) Considered various GC-RWA algorithms with the objective of minimizing the blocking probability of the offered traffic load. Chapter 3 also examined the static Virtual Topology Design (VTD) problem for groupcast traffic and set-up MILP formulations for setting up a finite set of groupcast sessions at a minimum cost using lightpaths for networks with and without wavelength conversion.

Constructing light-forests utilizing light-trees instead of lightpaths was another novel technique described in this thesis. Chapter 4 describes how multiple light-trees can also be created to support a single groupcast session. Two approaches were analyzed in this chapter and their performances were compared with each other. The first approach utilized the conventional multicast-capable optical cross-connect

node architecture as described in Chapter 2, whereas the second approach utilized simple Drop-and-continue (DaC) node architectures. Algorithms for constructing light-trees based on MC-OXCs using the MCH approach, as well as algorithms for developing linear light-trees using the DaC node designs were developed and analyzed. Additional considerations on the physical layer constraints imposed for the linear light-tree design were discussed. Finally, the static virtual topology design (VTD) was studied in a network with MC-OXCs supporting groupcast traffic. The objective of the VTD problem was to set-up a finite set of groupcast sessions, using light-trees, at a minimum cost. MILP formulations were presented that considered two basic models: MC-OXCs with and without wavelength conversion capabilities.

The thesis achieved its intended objectives. The research took advantage of the wide body of results available for provisioning of unicast and multicast connections and developed new planning and provisioning methodologies for solving the groupcast provisioning problem in WDM-based wavelength-routed optical networks. The research developed simple and novel models and algorithms that can be implemented before the network becomes operational (static heuristics and network planning techniques), as well as dynamic approaches that are used to provisioning groupcast connections on-the-fly.

5.2 Future Directions

During the course of this research, several challenging problems have presented themselves. Since the thesis cannot address the large number of questions that arose, they are noted as topics for future exploration. This section identifies some of these problems and tries to provide preliminary answers and directions for solving them.

There are several important directions in which the research on groupcast pro-

visioning can be extended. One of the immediate directions for further research is fault restoration for groupcast sessions. Once light-forests are in session, they must survive single link (and possibly node) failures. An approach similar to multicast 1+1 protection as shown in Chapter 2 can be provided by building rings for all light-trees created in the mesh topology [71] [16]. Different classes of groupcast services should be accommodated simultaneously and they must be bounded by SLAs. Protection and restoration of multiple failures is another open problem, albeit a very difficult one as it will be rather complicated to protect against simultaneous failures. One possible solution to this problem is to provision partial groupcast or multicast sessions, without providing service to a small percentage of the destination nodes.

Another related issue to be addressed is the problem of traffic grooming in groupcast sessions. The network should be able to address the provisioning of connections with several different types of granularities.

Another interesting question is the fairness issues that arise when we try to provision mixed services such as unicast and groupcast connections. The larger connections have the tendency to get blocked more often than the smaller connections and thus an approach that employs a fairness scheme has to be developed, so that all connections are treated the same way during the provisioning process. In the opposite end of the spectrum there is much work to be done in the area of admission control for these networks. If a network has limited resources and a large groupcast session arrives and uses a large portion of the network resources, the smaller connections cannot be admitted in the network. A possible solution to this problem is a network state-aware indicator that governs the admission control for multicast and groupcast traffic.

Another significant open question is the development of a control plane for the provisioning of mixed-type connections, especially when these connections cross over several administrative domains. Routing, and signaling extensions that incorporate

groupcast and mixed-type applications have to be developed and standardized, and groupcast session management has to be developed especially for protection and restoration of these services.

Bibliography

- [1] A. Tzanakaki, I. Zacharopoulos, and I. Tomkos, "Broadband Building Blocks," *IEEE Circuits and Devices Magazine*, vol. 20, pp. 32–37, Mar.-Apr. 2004.
- [2] D. Cavendish, "Evolution of Optical Transport Technologies: From SONET/SDH to WDM," *IEEE Communications Magazine*, pp. 164–172, June 2000.
- [3] A. Rodriguez-Moral, P. Bonenfant, S. Baroni, and R. Wu, "Optical Data Networking: Protocols, Technologies, and Architectures for Next Generation Optical Transport Networks and Optical Internetworks," *Journal of Lightwave Technology*, vol. 18, no. 12, pp. 1855–1870, 2000.
- [4] L. H. Sahasrabudde and B. Mukherjee, "Multicast Routing Algorithms and Protocols: A Tutorial," *IEEE Network*, pp. 90–102, Jan-Feb 2000.
- [5] T. Rahman, M. A. Ali, and G. Ellinas, "Building Light-Forest to Support Group Multicast Sessions in Mesh-based Optical Grid Networks," in *OFC/NFOEC Technical Digest*, Anaheim, CA, Mar. 2006, paper OWQ2.
- [6] S. Sengupta, V. Kumar, and D. Saha, "Switched Optical Backbone for Cost-Effective Scalable Core IP Networks," *IEEE Communications Magazine*, pp. 60–70, June 2003.
- [7] "Global Lambda Integrated Facility," [Online], <http://www.glif.is/>.

- [8] P. Szegedi, Z. Lakatos, and J. Spath, "Signaling architectures and recovery time scaling for grid applications in IST project MUPBED," *IEEE Communications Magazine*, vol. 44, no. 3, March 2006.
- [9] "Open Grid Forum," [Online], <http://www.ggf.org/>.
- [10] Y. Qin, K. Sivalingam, and B. Li, "QoS for Virtual Private Networks (VPN) over Optical WDM Networks," in *OPTICOMM 2000*, 2000.
- [11] B. Ramamurthy and A. Ramakrishnan, "Design of Virtual Private Networks (VPNs) over Optical Wavelength Division Multiplexed (WDM) Networks," in *OPTICOMM 2000*, Oct. 2000, pp. 76–86.
- [12] L. H. Sahasrabuddhe and B. Mukherjee, "Light-Trees: Optical Multicasting for improved Performance in Wavelength-Routed Networks," *IEEE Communication Magazine*, pp. 67–73, February 1999.
- [13] "The Metro Ethernet Forum," <http://metroethernetforum.org/>.
- [14] D. Papadimitriou, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control," RFC 4328 (Proposed Standard), Jan. 2006. [Online]. Available: <http://www.ietf.org/rfc/rfc4328.txt>
- [15] R. Ramaswami and K. N. Sivarajan, *Optical Networks: A Practical Perspective*, 2nd ed. San Francisco, CA: Morgan Kaufmann Publishers, 2002.
- [16] T. E. Stern and K. Bala, *Multiwavelength Optical Networks: A Layered Approach*. Upper Saddle River, NJ: Prentice Hall PTR, 2000.
- [17] A. Hadjiantonis, T. Rahman, S. Sherif, A. Khalil, G. Ellinas, M. Arend, M. A. Ali, and A. Shami, "A Novel Decentralized Ethernet-Based PON Architecture," in *IEEE ICC2004*, Paris, France, Jun. 2004, paper ON06-8.

- [18] A. M. Hamad and A. E. Kamal, "Optimal power-aware design of all-optical multicasting in wavelength routed networks," in *IEEE International Conference on Communications*, no. 1, June 2004, pp. 1796–1800.
- [19] W. S. Hu and Q. J. Zeng, "Multicasting optical cross connects employing splitter-and-delivery switch," *IEEE Photonics Technology Letter*, vol. 10, pp. 970–972, July 1998.
- [20] M. Ali and J. Deogun, "Power-efficient Design of Multicast Wavelength-routed Networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 1852–1862, October 2000.
- [21] n. K. Singhal and b. Mukherjee, "Architectures and algorithm for multicasting in WDM optical mesh networks using opaque and transparent optical cross-connects," in *Optical Fiber Communication Conference and Exhibit*, vol. 2, 2001, pp. TuG8–1 – TuG8–3.
- [22] H. S. Hamza and J. S. Deogun, "A new class of strictly nonblocking WDM multicast switching networks," in *Broadband Communications, Networks and Systems*, Sept. 2007, pp. 477 – 485.
- [23] C. Assi, A. Shami, and M. Ali, "Optical networking and real-time provisioning: an integrated vision for the next-generation Internet," *IEEE Network*, vol. 15, pp. 36–45, 2001.
- [24] B. Cain, S. Deering, I. Kouvelas, B. Fenner, and A. Thyagarajan, "Internet Group Management Protocol, Version 3," RFC 3376 (Proposed Standard), Oct. 2002, updated by RFC 4604. [Online]. Available: <http://www.ietf.org/rfc/rfc3376.txt>
- [25] B. Fenner, M. Handley, H. Holbrook, and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)," RFC

- 4601 (Proposed Standard), Aug. 2006, updated by RFC 5059. [Online]. Available: <http://www.ietf.org/rfc/rfc4601.txt>
- [26] A. Adams, J. Nicholas, and W. Siadak, "Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)," RFC 3973 (Experimental), Jan. 2005. [Online]. Available: <http://www.ietf.org/rfc/rfc3973.txt>
- [27] M. Handley, I. Kouvelas, T. Speakman, and L. Vicisano, "Bidirectional Protocol Independent Multicast (BIDIR-PIM)," RFC 5015 (Proposed Standard), Oct. 2007. [Online]. Available: <http://www.ietf.org/rfc/rfc5015.txt>
- [28] S. Bhattacharyya, "An Overview of Source-Specific Multicast (SSM)," RFC 3569 (Informational), Jul. 2003. [Online]. Available: <http://www.ietf.org/rfc/rfc3569.txt>
- [29] D. Waitzman, C. Partridge, and S. Deering, "Distance Vector Multicast Routing Protocol," RFC 1075 (Experimental), Nov. 1988. [Online]. Available: <http://www.ietf.org/rfc/rfc1075.txt>
- [30] J. Moy, "Multicast Extensions to OSPF," RFC 1584 (Historic), Mar. 1994. [Online]. Available: <http://www.ietf.org/rfc/rfc1584.txt>
- [31] T. Bates, Y. Rekhter, R. Chandra, and D. Katz, "Multiprotocol Extensions for BGP-4," RFC 2858 (Proposed Standard), Jun. 2000, obsoleted by RFC 4760. [Online]. Available: <http://www.ietf.org/rfc/rfc2858.txt>
- [32] A. Shami, C. Assi, I. Habib, and M. Ali, "Performance evaluation of two GMPLS-based distributed control and management protocols for dynamic lightpath provisioning in future IP networks," in *IEEE ICC02*, 2002.
- [33] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels," RFC 3209 (Proposed

- Standard), Dec. 2001, updated by RFCs 3936, 4420, 4874, 5151. [Online]. Available: <http://www.ietf.org/rfc/rfc3209.txt>
- [34] D. Katz, K. Kompella, and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2," RFC 3630 (Proposed Standard), Sep. 2003, updated by RFC 4203. [Online]. Available: <http://www.ietf.org/rfc/rfc3630.txt>
- [35] S. Salsano, A. Botta, P. Iovanna, M. Intermite, and A. Polidoro, "Traffic engineering with OSPF-TE and RSVP-TE: Flooding reduction techniques and evaluation of processing cost," *Elsevier Computer Communications*, vol. 29, p. 20342045, Jan. 2006.
- [36] A. Khalil, A. Hadjiantonis, G. Ellinas, and M. Ali, "Dynamic Provisioning of Survivable Heterogeneous Multicast and Unicast Traffic in WDM Networks," in *IEEE International Conference on Communications*, vol. 6, June 2006, pp. 2465–2470.
- [37] S. V. Kartalopoulos, *Introduction to DWDM Technology Data in a Rainbow*. Piscataway, NJ: IEEE Press, 2000.
- [38] K. Bala, "Routing In Linear Lightwave Networks," Ph.D. dissertation, Columbia University, 1993.
- [39] J. He, S. G. Chan, and D. Tsang, "Routing and Wavelength Assignment for WDM Multicast Networks," in *IEEE Global Telecommunications Conference*, vol. 3, no. 1, Nov. 2001, pp. 1536–1540.
- [40] Y. Sun, J. Gu, and D. Tsang, "Multicast Routing in All-Optical Wavelength-Routed Networks," *Optical Network Magazines*, vol. 2, no. 4, pp. 101–109, Jul/Aug 2001.

- [41] C. S. R. Murthy and M. Gurusamy, *WDM Optical Networks: Concepts, Design, and Algorithms*. Prentice Hall PTR, 2002.
- [42] R. A. Barry and P. A. Humblet, “On the Number of Wavelengths and Switches in All-Optical Networks,” *IEEE Transactions on Communications*, vol. 42, no. 2/3/4, pp. 583–591, February/March/April 1994.
- [43] Han Y. Ryu et al., “Stable Single-Frequency Fiber Ring Laser for 25-GHz ITU-T Grids Utilizing Saturable Absorber Filter,” *IEEE Photonics Technology Letters*, vol. 17, pp. 1824 – 1826, September 2005.
- [44] *Spectral grids for WDM applications: DWDM frequency grid*, ITU-T G.692.1 Standard, May 2002.
- [45] “Level 3 Network,” [Online], http://www.level3.com/about_us.
- [46] “Global Crossing,” [Online], http://www.globalcrossing.com/network/network_interactive_map.aspx.
- [47] “360 Network,” [Online], <http://www.360networks.com/>.
- [48] “TeleGeography,” [Online], <http://www.telegeography.com/maps/>.
- [49] “KPN International,” [Online], <http://www.kpn.com/kpn/show/id=1561743>.
- [50] “The NSFNET Backbone Service,” [Online], <http://www.merit.edu/networkresearch/projecthistory/nsfnet/>.
- [51] K. Zhu, “Network Topologies,” [Online], <http://networks.cs.ucdavis.edu/~zhuk/topologies.html>.
- [52] K. Zhu, H. Zang, and B. Mukherjee, “A Comprehensive Study on Next-generation Optical Grooming Switches,” *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 1173–1186, Sep. 2003.

- [53] N. K. Singhal, L. H. Sahasrabudde, and B. Mukherjee, "Provisioning of Survivable Multicast Sessions Against Single Link Failures in Optical WDM Mesh Networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 21, pp. 2587–2594, Sep. 2003.
- [54] G. Chartrand and P. Zhang, *Introduction to Graph Theory*. McGraw Hill Higher Education, 2005.
- [55] J. He, S. H. G. Chan, and D. H. K. Tsang, "Multicasting in WDM Networks," *IEEE Communications Surveys & Tutorials*, vol. 4, no. 1, pp. 2–20, 2002.
- [56] F. Farahmand, X. Huang, and J. P. Jue, "Efficient Online Traffic Grooming Algorithms in WDM Mesh Networks with Drop-and-Continue Node Architecture," in *Proceedings of the First International Conference on Broadband Networks, BROADNETS'04*. IEEE Computer Society, 2004.
- [57] M. Ali, "Optimization of Splitting Node Placement in Wavelength-routed Optical Networks," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 1571–1579, October 2002.
- [58] E. Miller, R. Libeskind-Hadas, D. Barnard, W. Chang, K. Dresner, W. M. Turner, and J. R. Hartline, "On the complexity of virtual topology design for multicasting in WDM trees with tap-and-continue and multicast-capable switches," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1601–1612, Nov. 2004.
- [59] J. R. K. Hartline, R. Libeskind-Hadas, K. M. Dresner, E. W. Drucker, and K. J. Ray, "Optimal virtual topologies for one-to-many communication in WDM paths and rings," *IEEE/ACM Transactions on Networking*, vol. 12, no. 2, pp. 375–383, April 2004.

- [60] G. N. Rouskas, "Optical layer multicast: Rationale, building blocks, and challenges," *IEEE Network*, vol. 17, no. 1, pp. 60–65, January/February 2003.
- [61] S. Frisken, H. Zhou, D. Abakoumov, G. Baxter, and S. Poole, "High performance 'Drop and Continue' functionality in a Wavelength Selective Switch," in *OFC/NFOEC Technical Digest*, Anaheim, CA, Mar. 2006, post Deadline Paper PDP14.
- [62] H. Hinderthr and L. Friedrich, "A Novel, Multi-Service Add-Drop-Multiplexer Architecture for Access and Backhauling Applications With 4.3Gbit/s Line Rate," in *OFC/NFOEC Technical Digest*, Mar. 2007, paper JWA91.
- [63] O. Yu and Y. Cao, "Placement of Light Splitters and Wavelength Converters for Efficient Multicast in All-Optical WDM Networks," *IEICE Transaction. INF. & SYST*, vol. E89-D, no. 2, pp. 709–718, February 2006.
- [64] *GR-253-CORE Synchronous Optical Network (SONET) Transport Systems: Common Generic Criteria*, 4th ed., Telcordia, Dec. 2005.
- [65] J. Bayne and M. Sharma, "Broadcast-and-select OADM Enables low-cost Transparency," *Lightwave Special Reports: Annual Technology Forecast*, Dec. 2001, http://lw.pennnet.com/display_article/128774/13/ARTCL/none/none/1/Broadcast-and-select-OADM-enables-low-cost-transparency.
- [66] N. K. Singhal, L. H. Sahasrabudde, and B. Mukherjee, "Optimal Multicasting of Multiple Light-Trees of Different Bandwidth Granularities in a WDM Mesh Network With Sparse Splitting Capabilities," *IEEE/ACM Transactions on Networking*, vol. 15, no. 5, pp. 1104–1117, October 2006.
- [67] T. Rahman, M. A. Ali, and G. Ellinas, "Performance Evaluation of Light-Forests to Serve Groupcast Sessions in WDM Mesh Networks," in *OFC/NFOEC Technical Digest*, San Diego, CA, Mar. 2009, JWA88.

- [68] Y. Xin and G. N. Rouskas, “Multicast Routing Under Optical Layer Constraints,” in *IEEE Infocom*, vol. 4, 2004, pp. 2731–2742.
- [69] —, “Optical layer multicast: Rationale, building blocks, and challenges,” *OSA Journal of Optical Networking*, vol. 3, no. 5, pp. 282–302, May 2004.
- [70] J. C. Palais, *Fiber Optic Communications*, 5th ed. Pearson Prentice Hall, 2005.
- [71] T. Rahman and G. Ellinas, “Protection of Multicast Sessions in WDM Mesh Optical Networks,” in *OFC/NFOEC Technical Digest*, Anaheim, CA, Mar. 2005, paper OTuK5.