QoS Aware Quorumcasting Over Optical Burst Switched Networks

A Thesis

Submitted for the Degree of

Doctor of Philosophy

in the Faculty of Engineering

by

B. G. Balagangadhar

Electrical and Communication Engineering

INDIAN INSTITUTE OF SCIENCE

BANGALORE – 560 012, INDIA

July 2008

TO

Lord Almighty,

My Parents,

& Family

Acknowledgements

I guess a page or two is not sufficient to mention about the persons who had helped, to reach this position of my life. First and foremost are my parents. In spite of all the hardships faced, they always motivated me to pursue Ph. D. They swallowed *poison* in the form of hardships only to give me *Amrutham*, the Nectar. To me they replicate **Lord Siva, Paravathi**. What more I can ask Lord Almighty, except thank Him for giving such a wonderful parents.

Whenever I am lost in darkness my *Mentor* acts as guiding light to the path of success. *Shri Gurubhyon Namaha*. It has been a great honor to work under Prof. T.Srinivas my research adviser. He has been helpful throughout my research and a highly motivating individual. I would like to thank Prof. Vinod M. Vokkarane, Assistant Professor, Department of Computer and Information science, *University of Massachusetts, Dartmouth*, for giving me an opportunity to work in National Science Foundation (NSF) project, which has greatly contributed to my thesis.

I would like to thank Prof. Vinod Sarma, Chairman ECE and Prof.Anurag Kumar, Divisional Chairman, Electrical Sciences for giving me a opportunity to pursue my research. I would like to thank Dean Engineering and Deputy Registrar (Paneer Selvam) for revoking my Ph.D registration. I sincerely thank Prof. Pappanareddy for his suggestions and constructive criticism. I also thank Prof. Selvarajan for reviewing my thesis.

I had great time working as Project Assistant, with Dr.Badrinarayan, Principal Research Scientist. I would also like to thank Dr.Shiva Leela, Senior Research Scientist and Mr.G.S.Hedge for their continuous assistance and support.

I would like to thank Prof. P.H.N.B.Sarma, who taught me Electrical Engineering for my under-graduate. I guess more than Electrical Engineering, I learned a way to lead the life as an respectable individual. He is my inspiration and role model.

I cannot forget my younger brother (Chinna), who has helped me a lot. He never let my parents feel, I was away. He took care of everything. My wife (Leela) has supported me in all aspects, I am lucky to have her as my life partner. Friends make every moment in life a cherishing one. Dinu, Kitty (Balakrishna) always have been helpful and motivating. I am really indebted to them. I enjoyed my stay at IISc with my friends Aditya, Pattu, Anand, Bharat. These guys made my life at IISc a memorable and ever-lasting. I would like to thank Rajesh, Graduate student in CIS Department, University of Massachusetts for helping me, with my transition from communications to computers. I thank Premkumar for useful discussions and suggestions. Akshay, Guru, and Sai are the best buddies I had in USA.

I would also like to thank my relatives, especially Babbai (Mr.Manikalaya Rao), Mamma (Mr.Subba Rao) and their family for encouraging me to pursue higher education. At last, I would like to thank everyone, who directly or in-directly helped me to reach this stage.

Vita

- **Education**
  - **M.Tech** in Opto-electronics and Laser Technology, June 2004, **International School of Photonics, Cochin University of Science and Technology**, Cochin, Kerala, India.

    Thesis title: *Topological Design Issues in Wide Area Wavelength Routed Optical Networks.*
  - **A.M.I.E.T.E (Electronics and Telecommunications)**, August 2001, **Institute of Electronics and Telecommunications**, New Delhi, India.
  - **B.Sc (Mathematics, Statistics, Computer Science)**, March 2001, **Osmania University**, Andhra Pradesh, India.
- **Research interests**
  - QoS based algorithms for WDM and OBS Networks.
  - Impairment-aware algorithms for Optical Networks.
  - Performance evaluation of Manycasting in Optical Burst Switched Networks.
- **Research Experience**
  - Visiting Scholar, (August 2007 – May 2008), CIS Department, University of Massachusetts Dartmouth. Project is "*Service Oriented Optical Networks (SOON)*" sponsored by National Science Foundation (NSF). Involved in analytical and simulation of QoS aware Manycasting in Optical-Burst Switched Networks.

Publications

- **Journals**

    1. **Balagangadhar. G. Bathula**, R. R. C. Bikram, V. M. Vokkarane, S. Talabattula, "Evaluation of Physical Impairment-Aware Manycasting Algorithms over Optical Burst-Switched (OBS) Networks," submitted to *IEEE/OSA Journal of Lightwave Technology.*

    2. **Balagangadhar. G. Bathula**, V. M. Vokkarane, S. Talabattula, "QoS based Manycasting Over Optical Burst Switched Networks (OBS)," submitted to *IEEE/ACM Transactions on Networking*

- **Conferences**

    1. **Balagangadhar. G. Bathula**, R. R. C. Bikram, V. M. Vokkarane, S. Talabattula, "Impairment-Aware Manycasting Algorithms Over Optical Burst Switched Networks," *to appear, Proceeding of IEEE International Conference on Computer Communication and Networks* (ICCCN)-2008, Virgin Islands, USA, Aug. 2008.

    2. **Balagangadhar. G. Bathula**, V. M. Vokkarane, R. R. C. Bikram, "Impairment-Aware Manycasting Over Optical Burst-Switched Networks", to appear, *IEEE International Conference on Communications* (ICC-2008), Beijing, China, May 2008, pp. 5234-5238.

    3. **B.G.Balagangadhar**, T.Srinivas and Rajappa Pappanareddy, "Constraint Based Routing for Wavelength Routed Optical Networks", *International Conference on Optics and Laser, ICOL 2005*, 12$^{\text{th}}$ Dec-15$^{\text{th}}$ Dec 2005, IRDE,

Dehradun, India.

4. **B.G.Balagangadhar**, T.Srinivas and Rajappa Pappanareddy, "Estimating Blocking Probability based on Network Layer and Physical Layer Constraints for High Speed Optical Networks", *International Conference on Photonics, 2006*, 13$^{\mathrm{th}}$ Dec-16$^{\mathrm{th}}$ Dec 2006, Hyderabad, India.

Abstract

Recently there is an emergence of many Internet applications such as multimedia, video conferencing, distributed interactive simulations (DIS), and high-performance scientific computations like Grid computing. These applications require huge amount of bandwidth and a viable communication paradigm to coordinate with multiple sources and destinations. Optical networks are the potential candidates for providing high bandwidth requirement. Existing communication paradigms include broadcast, and multicast. Hence supporting these paradigms over optical networks is necessary. Multicasting over optical networks has been well investigated in the literature. QoS policies implemented in IP does not apply for Wavelength division multiplexed (WDM) or optical burst switched (OBS) networks, as the optical counterpart for store-and-forward model does not exist. Hence there is a need to provision QoS over optical networks. These QoS requirements can include contention, optical signal quality, reliability and delay. To support these diverse requirements, optical networks must be able to manage the available resources effectively.

Destinations participating in the multicast session are fixed (or rather static). Due to the random contention in the network, if at least one or more destination(s) is not reachable, requested multicast session cannot be established. This results in loss of multicast request with high probability of blocking. Incorporating wavelength converters (WCs) at the core nodes can decrease the contention loss, however WCs require optical-electrical-optical (O/E/O) conversion. This increases the delay incurred by optical signal. On the other hand all-optical WCs are expensive and increase the cost of the network if deployed.

Goal of this thesis is, to provide hop-to-hop QoS on an existing all-optical network (AON) with no WC and optical regeneration capability. In order to minimize the request lost due to contention in AON, we propose a variation of multicasting called *Quorumcasting* or *Manycasting*. In Quorumcasting destinations can join (or leave) to (or from) the group depending on whether they are reachable or not. In other words destinations have to be determined rather than knowing them prior, as in the case of multicasting. Quorum pool is minimum number of destinations that are required to be participated in the session for successful accomplishment of the job (k be the size of quorum pool). Providing QoS for manycasting over OBS has not been addressed in the literature. Given the multicast group (with cardinality $m > k$) and the number of destinations required to be participated, the contribution of this work is based on providing necessary QoS.

In this thesis we study the behavior of manycasting over OBS networks. In OBS networks, packets from the upper-layer (such as IP, ATM, STM) are assembled and a burst is created at the edge router. By using O/E/O conversion at the edge nodes, these optical bursts are scheduled to the core node. Control header packet or burst header packet (BHP) is sent to prior to the transmission of burst. The BHP configures the core nodes and the burst is scheduled on the channel after certain offset time.

In the first part of the thesis, we explain the different distributed applications with primary focus on Grid over OBS (GoOBS). We study the loss scenario due contention and inadequate signal quality for an unicast case in OBS network. We further extend this to manycasting. We modify the BHP header fields to make the burst aware of not only contention on the next-hop link, but also bit-error rate (BER). By using recursive signal and noise power relations, we calculate the BER (or q-factor) of the link and schedule the burst only if the required BER threshold is met. Thus all the bursts that reach the next-hop node ensure that contention and BER constraint are met. This are called **"Impairment-Aware (IA) Scheduling"**. Burst loss in the network increases due to BER constraint. Hence we propose algorithms to decrease the burst loss and simultaneously providing the sufficient optical signal quality. We propose three algorithms called IA-shortest path tree (IA-SPT), IA-static over provisioning (IA-SOP), and IA-dynamic membership (IA-DM). In IA-SPT destination set is sorted in the non-decreasing order of the hop-distance from source. First k of them are selected and bursts are scheduled to these destinations along the shortest path. In IA-SOP we select additional $k'(\leq m-k)$ destinations where k' is the over provisioning factor. Over provisioning ensures that burst at least reach k of them, decreasing the contention blocking. However as the burst has to span more destinations, the fan-out of the multicast capable switch will be more and the BER could be high. In IA-DM destinations are dynamically added or removed, depending on contention and BER. Destination is removed and new destination is added based on the two constraints. Our simulation results shows that IA-DM out performs the other two algorithms in terms of request blocking. We show that IP-based manycasting has poor performance and hence there is a need for supporting manycasting over OBS networks. We verify our simulation results with the proposed analytical method.

In the next part, we focus on provisioning QoS in manycasting. QoS parameters considered for analysis include, signal quality i.e., optical signal to noise ratio (OSNR), reliability of the link and, propagation delay. In this work we consider application based QoS provisioning. In other words, given the threshold requirements of an application, our aim is to successfully schedule the burst to the quorum pool satisfying the threshold conditions. We use a de-centralized way of the scheduling the burst, using BHP. With the help of local-network state information, the burst is scheduled only if it satisfies multiple set of constraints. Corresponding reception of burst at the node ensures that all the QoS constraints are met and burst is forwarded to the next hop. QoS attributes are either multiplicative or additive. Noise factor of the optical signal and reliability factor are multiplicative constraints, where as propagation delay is additive. We define a path information vector, which provides the QoS information of the burst at every node. Using lattice theory we define an ordering, such that noise factor and propagation delay are minimum and reliability is maximum. Using path algebra we compute the overall QoS attributes. Due to multiple set of constraints, the request blocking could be high. We propose algorithms to minimize request blocking for Multiple Constrained Manycast Problem (MCMP). We propose two algorithms MCM-SPT and MCM-DM. We consider different set of service thresholds, such as real time and data service thresholds. Real time services impose restriction on signal quality and the propagation delay. On the other hand data services require high reliability and signal quality. Our simulation study shows that MCM-SPT performs better than MCM-DM for real-time services and the data services can be provisioned using MCM-DM.

Contents

List of Tables

List of Figures

List of Algorithms

Keywords

Optical Burst Switched Networks, Manycasting, Impairment-Aware algorithms, Blocking Probability, QoS constraints, Grid over OBS.

Abbreviations

WDM	Wavelength Division Multiplexing
DWDM	Dense Wavelength Division Multiplexing
OBS	Optical Burst Switched Networks
GoOBS	Grid over Optical Burst Switched Networks
PMD	Polarization Mode Dispersion
SPT	Shortest Path Tree
BHP	Burst Header packer
DM	Dynamic Membership
SOP	Static Over Provisioning
CBR	Constraint Based Routing
MPLS	Multi-Protocol Label Switching
MCMP	Multi-Constrained Manycast Problem
QoS	Quality of Service
BER	Bit Error Rate
SAN	Storage Area Network
CDN	Content Distribution Network

DIS	Distributed Interactive Simulation
OSNR	Optical Signal to Noise Ratio
AON	All Optical Network
WRON	Wavelength Routed Optical Network
EDFA	Erbium Doped Fiber Amplifier
ASE	Amplified Spontaneous Emission
JET	Just Enough Time
JIT	Just In Time
BCP	Burst Control Packet
WCC	Wavelength Continuity Constraint
MC-OXC	Multicast Capable - Optical Crossconnect Switch

Chapter 1

Optical Networks-An Overview

1.1 Introduction

In this age of information we need everything or anything to be at our fingertips. Sitting at one corner of the globe we want to access information located at other end of globe. In order to accomplish this, all we need is bandwidth. Bandwidth can be provided in any form, such as copper, optical fiber, and wireless. Traffic demands are increasing tremendously due to many emerging data-centric applications. This resulted in increase of data traffic compared to voice [1], [2]. These traffic demands can only be supported with optical fiber medium, which can ideally carry data rates to an order of Tb/s. Optical fiber communications gained importance due to its tremendous potential for supporting high data rate applications. Transmitting data in optical fiber over longest possible distance has been a challenge to many researchers. Though it seemed simple to send data over optical fiber, researchers had to overcome many obstacles. Evolution of optical fiber transmission systems started with light emitting diodes (LEDs) as the transmitters. Multimode fibers were used to carry the optical signal. This however suffered with many drawbacks, due to high optical signal loss. With the invention of lasers, researchers were able to couple the light over the single mode fiber, showing an significant improvement in the optical signal quality. Dispersion compensation fibers decrease the losses due to pulse dispersion. Due to the decrease in manufacturing cost of transmitters (lasers) and receivers (photo-detectors), wavelengths can be multiplexed into the fiber. This is called Wavelength Division Multiplexing (WDM). The network capacity has increased by using WDM technique and thus reducing the cost per bit transmitted per kilometer.

1.2 WDM Network

WDM is an approach to exploit huge-bandwidth, in which multiple users are multiplexed on the same fiber using different wavelengths. Each wavelength supports a single communication channel, operating at what ever rate one decides. With number of wavelengths coexisting on the single fiber one can tap huge bandwidth of the fiber. WDM network architectures, devices, components and sources are available in market place and are still emerging. Many WDM networks has been already been deployed mainly as the backbone networks. Routing of the wavelengths along the optical networks is carried through optical channels called *lightpaths* [3]. A lightpath may span multiple fiber links to provide "circuited-switched" inter-connection between the nodes, far away from the each other geographically. Each intermediate node provides an optical bypass facility. Fig. 1.1 shows the typical WDM network, with core nodes connected to edge routers. Each color represents a lightpath on which data is carried on the single wavelength. These wavelength routed optical networks (WRON) are called **transparent** optical networks or all-optical networks (AON). The major milestone in evolution of optical fiber transmission systems is development of erbium-doped fiber amplifiers (EDFA). A major advantage of EDFA is that, it is capable of amplifying the optical signal at all wavelengths simultaneously. This provided an way to increase the system capacity, i.e., the bit rate. EDFA's can be considered as the major catalyst for deployment of WDM transmission systems. Detailed description on EDFAs can be found in [4]. New developments in fiber Bragg gratings (FBGs) with 25 GHz or 50 GHz channel separation can be used as multiplexer and demultiplexer for dense WDM (DWDM) systems [5].

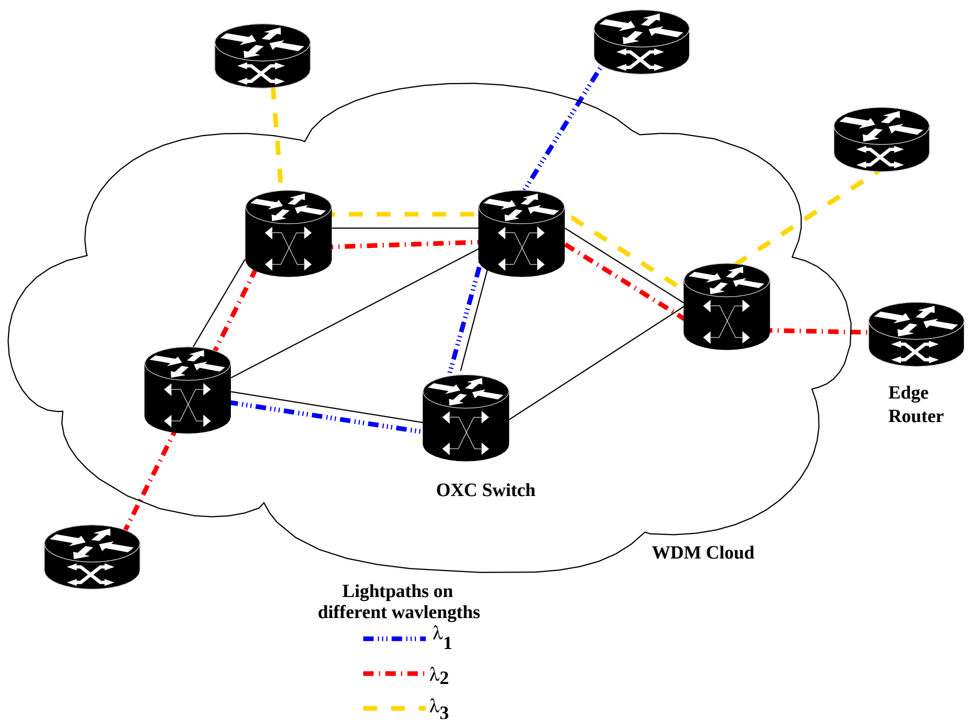


Figure 1.1: Schematic of Wavelength Division Multiplexed (WDM) Network.

1.3 Optical Burst Switched (OBS) Network

The OBS technology combines the advantages of optical circuit switching and optical packet switching [6]. An optical burst is defined as a number of continuous packets destined for a common egress point. The IP packets are combined to form a burst, with size varying from a single IP packet to large data. This allows for fine-grain multiplexing of data over single wavelength and therefore efficient use of optical bandwidth through sharing of resources (i.e., lightpaths) among number of users. In OBS technology, the data plane and optical plane are separated as shown in the Fig. 1.2. The burst aggregation and segregation is done at the ingress node by using electro-optical conversion. Prior to the burst transmission the burst control packet (BCP) or burst header packet (BHP) is created and sent towards the destination by the OBS ingress node. The BCP is sent on the out of band over separate wavelength channels for signaling and processed at the intermediate OBS routers. The architecture of OBS is explained in [7]. A comprehensive survey on optical burst switching can found in [8]. Signaling is an important aspect that can effect the performance of the network. For OBS network signaling plays an key role, since the core nodes in the OBS network are bufferless. In this work we focus on delayed reservation based signaling technique. In delayed reservation type of signaling the channel is reserved from actual arrival instant of the data burst at the node (or link). In order to implement this technique, BHP should carry the information of *offset time*. It is the time interval between BHP arrival and the data burst arrival. This type of signaling technique is called *Just Enough Time* (JET) signaling [6] [9]. On the other hand in immediate reservation signaling the channel is reserved from the instant the BHP reaches the node. *Just-In-Time* (JIT) signaling uses the immediate reservation [10] [11]. When multiple bursts contend for the same wavelength during signaling then data loss occurs and is called *contention loss*.

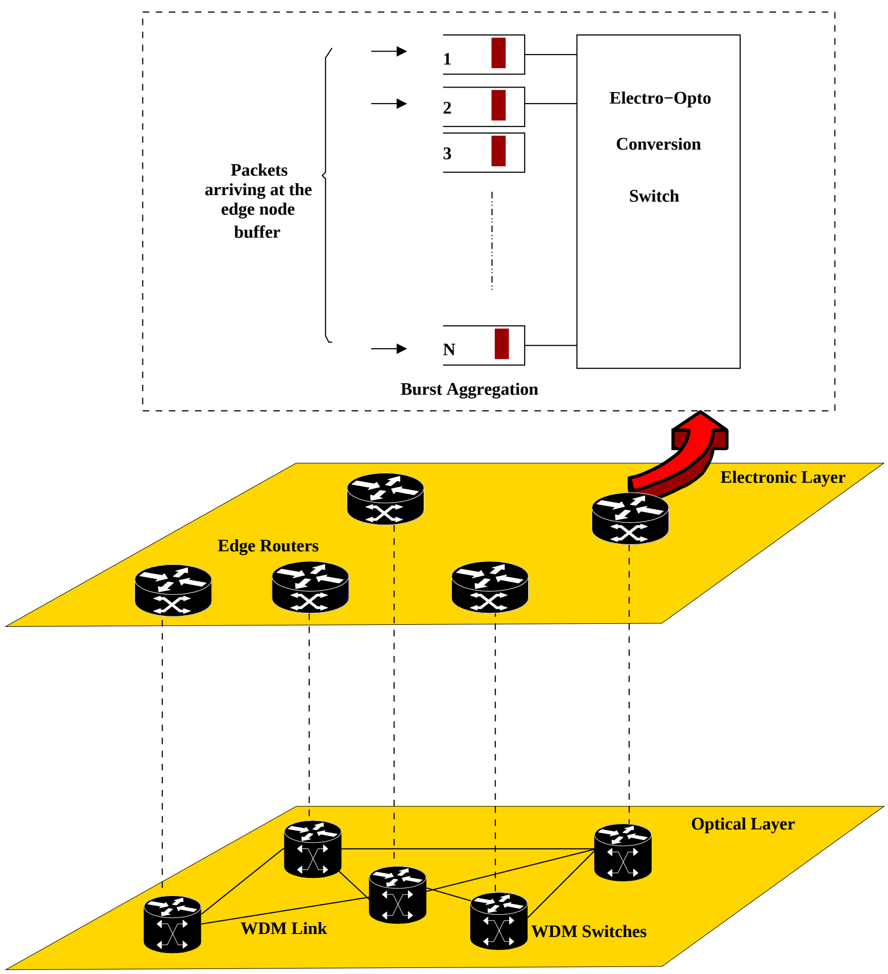


Figure 1.2: Schematic of typical OBS Network.

1.4 Different Communication Paradigms

In Fig. 1.3 different types of communication paradigms are shown. In multicasting, messages from source are forwarded to a group of destinations called *multicast group*. The cardinality of destinations in a multicast group is referred to as the *group size*. Multicast tree corresponds to a communication session established between source and fixed number of destinations which are already given. On the contrary manycasting or quorumcasting refers to a dynamic version of multicasting, where destinations have to determined instead of being given. In Fig. 1.3(b) there are five intended destinations (green & yellow), and out these five any three have to be selected. Thus a subset of destinations are selected from multicast group, which we call as *quorum pool*. Assuming the multicast group size to be m and quorum pool to be k, from combinatorics, number of possible subsets are $\binom{m}{k}$, where $m \geq k$. Depending upon the size of the multicast group, multicasting can be categorized as *small* or *large* group multicast [12].

Fig. 1.3(d) shows the variation in unicast (shown in Fig. 1.3(c)), where the any destination can be shown from the set of destinations. Performance analysis of anycasting to support Grid applications over OBS networks has been well investigated in the literature [13], [14].

1.5 Service Oriented Optical Networks (SOON)

Many emerging next-generation Internet applications, such as videoconferencing, multimedia distribution, and high-performance scientific computing, are characterized by high bandwidth requirements, multi-source and multi-destination communication paradigms, strict QoS requirements with respect to delay and loss, and high reliability and survivability requirements. In order to support these diverse requirements, emerging networks must be able to manage resources in a flexible manner.

This work investigates a service-oriented optical network architecture that is enabled by an underlying optical transport network capable of dynamic and agile resource allocation. An optical transport layer provides the basic mechanisms for reserving resources,

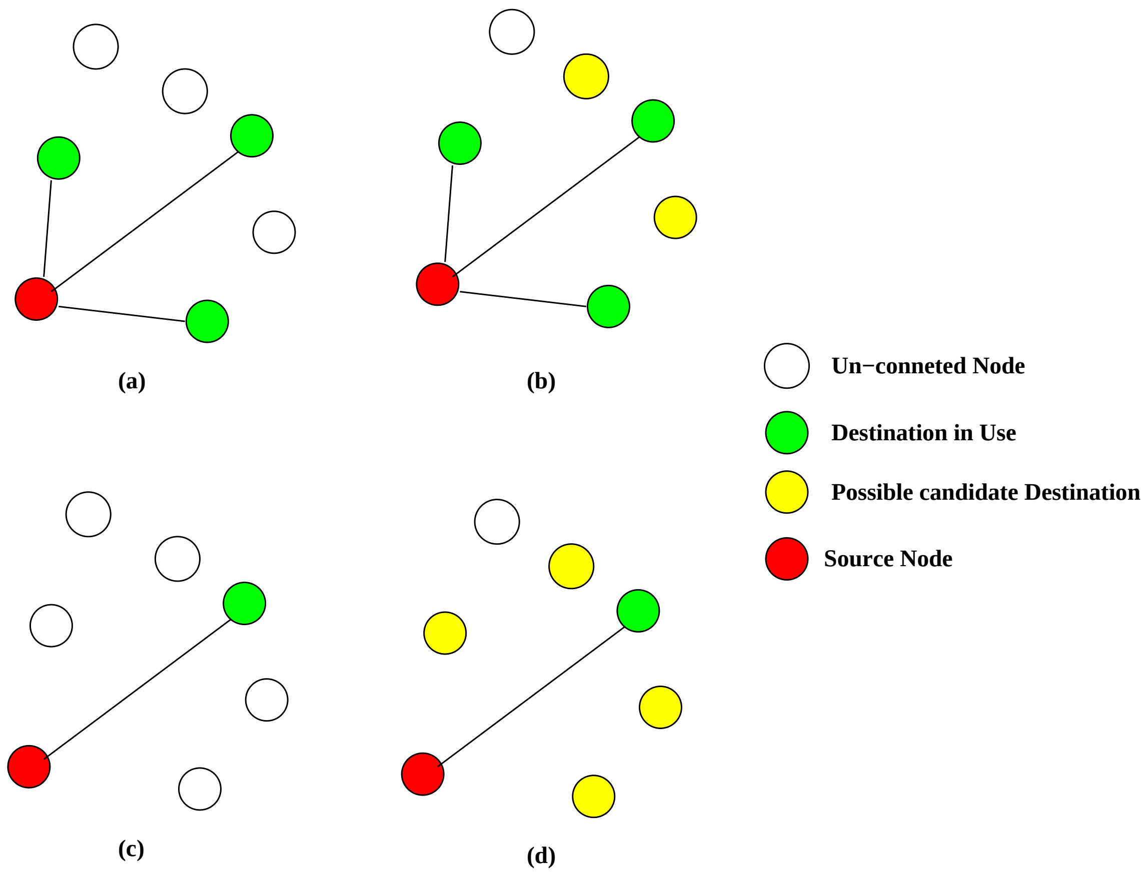


Figure 1.3: (a) Multicast (b) Manycast (c) Unicast (d) Anycast.

and an optical service layer builds a set of services over the basic optical transport layer. In the work, specific problems to be addressed include 1) investigating mechanisms for supporting manycast communication services in optical networks, 2) investigating mechanisms for providing impairment-awareness in such networks, and 3) developing methods for supporting reliable transport in optical networks. The proposed activity will result in a framework and an architecture for deploying a wide range of services in optical networks in order to support the requirements of the most demanding next-generation applications. The work will also result in the development of new algorithms for implementing manycasting and providing reliable services, and lead to the development of analytical models for evaluating these algorithms and protocols. Fig. 1.4 shows the proposed network architecture. Edge nodes convert the upper layer packets to Optical burst, before they are scheduled for transmission. Manycast requests are created at the edge node and core nodes route these bursts to the respective destinations. Bursts from different applications like Grid computing, storage area network or content distribution networks are provisioned with required QoS [15].

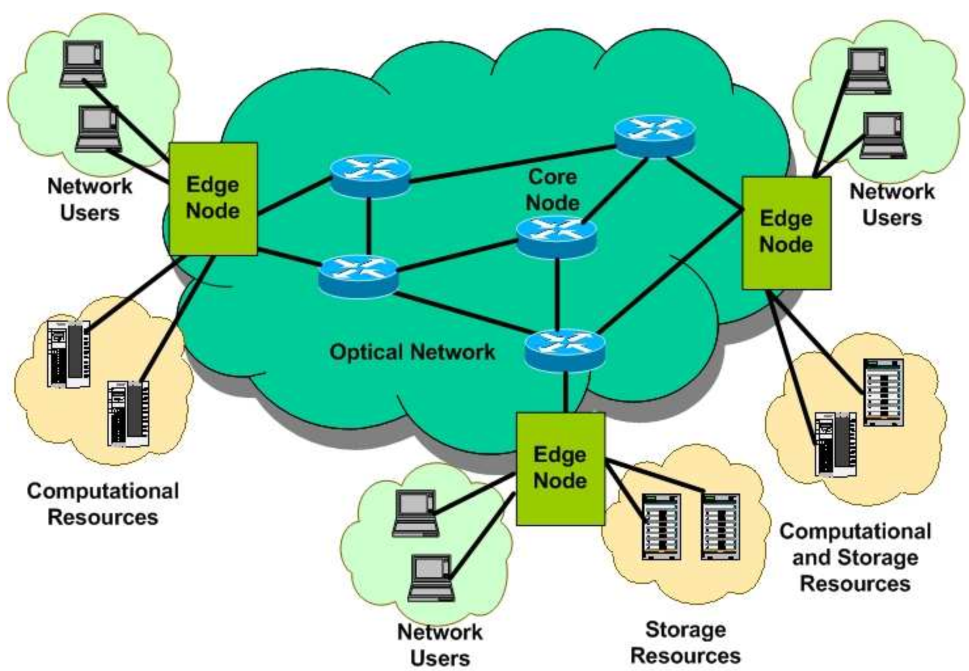


Figure 1.4: Service Oriented Optical Network (SOON).

1.6 Grid over OBS (GoOBS)

Grid based applications form a new generation of distributed applications that combine scientific instruments, distributed data services, sensors and computing resources to solve complex scientific problems. Scientific problems can be bioinfomatics, particle physics, material modeling and engineering [16], [17], [18]. To meet the demands of Grid applications, research has been carried out to understand the role of networking. Resource utilization of the Grid depends on the available bandwidth on the link. Hence there is need to integrate Grid resources with emerging high-performance optical networks [19] [20]. Testbeds like Optical Metro Network Initiative (OMNI) [21], CA*net4 [22] and, TransLight [23] are aimed at developing high-performance networking blocks for the Grid.

Optical Burst switching (OBS) is a promising technology for the future where the bandwidth needs are very high. The OBS technology has the potential to bring several advantages for Grid networking [24],

1. Native mapping between bursts and Grid jobs: the bandwidth granularity offered by OBS networks allows efficient transmission of user job's with different traffic profiles.

2. Separation of Control and data planes: this allows all-optical data transmission with fast user/application initiated lightpath setup.

3. Electronic processing of the burst control packet (BCP) at each node, enables the network infrastructure to offer the Grid protocol layer functionality.

In this GoOBS architecture, one or more application *requests* or *jobs*, are assembled into a super-size packet called *data burst*, which is then transported over the optical core network and forwarded to their appropriate Grid resources. Each data burst has associated *control packet* containing information about the burst's duration, source node, the type of the Grid resources the burst requires, etc. Typically in OBS network architecture the control plane and the data plane are separated with an associated burst offset time. The Edge routers (ingress node) in the GoOBS architecture must have an intelligence of *G*rid Job classification apart from aggregating packets into burst, and de-aggregating. This job classification enables the network to provide the specialized services called *Grid Differentiated Services* (GridDiffServ).

Fig. 1.5 shows the Grid architecture and the layered GoOBS architecture [25]. The review of Grid architecture can be found in [26] and the functionality of each layer are explained in [27], [25]. Here we present the review of these architectures for the reader understanding.

1. *Fabric* provides the underlying base structure including the storage systems, computers, networks, system descriptors, etc.

2. *Connectivity* defines core communication and authentication protocols required for Grid-specific network transactions. Communication requirements include transport, routing and naming. Authentication protocols built on communication services provide cryptographically secure mechanisms for verifying the identity of users and resources.

3. *Resource* defines information and management protocols. Information protocols obtain information about the structure and the state of a resource and the management protocols are used to negotiate access to the shared resource, specifying resource requirement (Quality of Service).

4. *Collective* deals with interactions that are global in nature, such as resource discovery, brokering, monitoring against failures, data replication services, etc.

5. *Application* refers to many different commercial, scientific and engineering applications requiring one or more resources such as computing power, speed, data storage, etc.

Fig. 1.5(b) describes the OBS layer functionality as the networking layer of the Grid. The control plane functionality and protocols are separated from those of data plane [28]. This is because the control information is transmitted out-of-band in OBS networks.

1. *OBS Data Plane* transports incoming jobs from higher layers to appropriate Grid resources and ensures that the job has been successfully completed with the slack time.

   (a) *Job Aggregation and De-aggregation (JAD) Layer* aggregates the incoming jobs with similar properties, such as QoS or type of Grid resources required. JAD also de-aggregates the received bursts and are sent to the appropriate destination clients.

   (b) *Burst Framing Control (BFC)* encapsulates the burst from JAD to a proper frame structures. It also decodes incoming data burst frames structures.

   (c) *Medium Access Control (MAC)* sublayer in data plane maintains information about Grid and OBS resources. Protocols related to OBS resources include reservation of the wavelength, scheduling, offset time, contention resolution schemes and multicasting protocols.

2. *OBS Control Plane* is responsible for transmitting control packets (CPs), which contain the information necessary for switching and routing the data bursts (DBs). MAC sublayer acts the application layer of the control plane. Signaling of CPs is performed in domain (electrical) independent of data (optical).

   (a) *Burst Signaling Control (BSC)* layer receives the data burst properties including Grid resource type, destination address, Quality-of-Service (QoS), burst type etc., from MAC and determines the type of control packet to be transmitted to the next hop.

   (b) *Signaling Connection Control (SCC)* layer includes the routing algorithms for control packets in order to establish the physical path for outgoing bursts. Routing protocols include any communication paradigms discussed in Section-1.4.

   (c) *Signaling Frame Control (SFC)* layer receives bit streams containing the control packet type and its associated data burst properties.

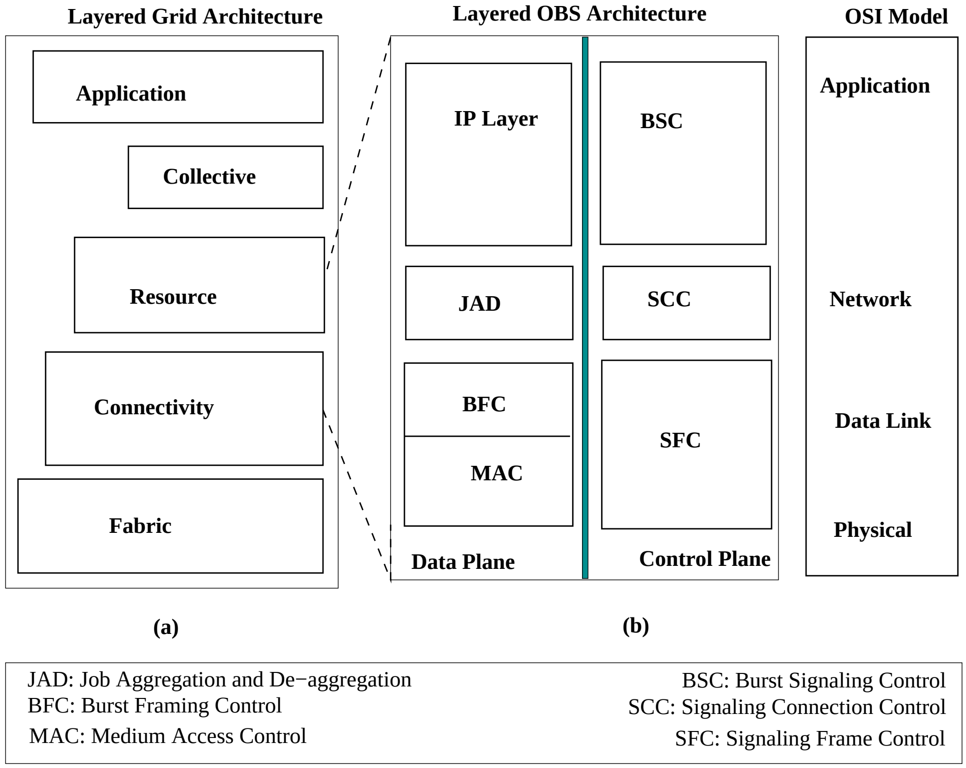


Figure 1.5: (a) A layered Grid Architecture (b) layered Grid-over-OBS architecture.

1.7 Contributions of this thesis

The reminder of thesis is organized into three chapters followed by a conclusion and future scope. In the *Chapter 2* we first explain the need for provisioning the bursts based on the impairments in the physical fiber link. Routing and Wavelength Assignment (RWA) algorithms are modified to take into account of loss encountered during the burst traversal along the fiber. We define QoS-Aware routing Algorithms (QARA) and explain the impact of the physical fiber link on the burst loss.

In *Chapter 3* we consider the manycasting scenario and study the effects of burst losses due to contention and BER. We propose the need for supporting manycasting over optical burst switched networks. Burst scheduling algorithms are proposed to make the BHP aware of impairments encountered in routing along the optical link. These are called *Impairment Aware* routing algorithms. In all the algorithms we use de-centralized (or distributed) way of scheduling the burst where each BHP maintains the information about the signal and noise powers. A burst is said to be successfully scheduled on the link, if and only if the link is free (i.e., no contention) and q-factor is more than the required threshold. We first propose Impairment-Aware Shortest-Path Tree (IA-SPT) which is based on computing the shortest path tree for a given manycast request. A variation of this is proposed as Impairment-Aware Static Over Provisioning (IA-SOP). In this scheme burst is sent to more than required destinations. This can decrease the contention loss, but due to increase in the fan-out of the switch, signal power decreases causing an increase in the BER. From simulation results, we observe that there is not much significant impact on the overall loss in IA-SOP. We propose Impairment-Aware Dynamic Membership (IA-DM), which is found to have significant decrease in loss due contention and BER. Finally with the help of analytical model, we verify our simulation results.

In *Chapter 4* for the first time, we propose algorithms to provide QoS for manycasting over OBS networks. In this chapter we use differentiated service provisioning methods. Based on the service threshold, bursts are scheduled. We propose an operation '$\circ$' to compute the path information vector which specifies the QoS attributes. The operation '$\circ$' performs multiplication on noise and reliability factors and addition on propagation delay. The proposed manycasting schemes are called Multi-Constrained manycasting algorithms (MCMP). We define threshold requirements based on data and real time service. Our proposed algorithms MCMP-Shortest path tree (MCMP-SPT) is found to perform better than MCMP-Dynamic membership (MCMP-DM) for real-time service. On the other hand the data services can be better provisioned using MCMP-DM. As the network topology is all-optical with no wavelength conversion and regeneration, cost is not considered as the constraint in this study.

Finally in *Chapter 5* we conclude this thesis, with possible future extensions in the work.

Chapter 2

Impairments in Wavelength Routed Optical Networks

2.1 Introduction

Wavelength Division Multiplexing (WDM) has been a promising technology for providing high bandwidth to the end users. Lightpaths (LPs) [3] are the basic communication channels between the two users on which an available free wavelength can be assigned. If there are no wavelength converters in the network, lightpath remains on same wavelength from source to destination. These networks are classified as All-Optical Networks (AON) also called Transparent optical networks. In other words the signal remains in optical domain from source to destination, without any optical-electronic-optical (O-E-O) conversion. This accounts for wavelength continuity constraint (WCC) i.e., call has be routed on the same wavelength from source to destination. If a call does not find a suitable wavelength it is dropped and hence lost. Because of the increasing traffic, these lightpaths are leased for varying durations. Hence a wavelength can be viewed as the circuit, making call blocking probability, an important metric for optical networks. Due to WCC *transparent optical network* suffers from high blocking probability. Transparent Optical Networks, are difficult to be deployed in large scale due to the following reasons,

- Cost of transmitters (lasers) or receivers (photodetectors) are expensive and the number of wavelengths being limited, some of the node pairs may not be connected directly through a single hop.

- Signal quality is effected as the lightpath traverse through different optical components [29].

- Wavelength Continuity Constraint (WCC).

One of the possible solution to minimize blocking probability is to have wavelength conversion incorporated in the nodes. These wavelength converters in the nodes are classified on the basics of wavelength conversion capability. In full wavelength conversion, any given wavelength from the set of wavelengths can be converted to any other wavelength. For example, let Λ be the set of wavelengths at the input to the wavelength converter, say $\lambda_k \in \Lambda$ can be converted to any $\lambda_p \in \Lambda$ with $k \neq p$. Networks which have these wavelength converters equipped on every node are called *Opaque Optical Networks*. Having full wavelength conversion capability makes the network analogous to the traditional circuit switched network. Hence the blocking probabilities are very less as compared to AON. However wavelength converter devices are still at the research stage and hence increase the cost of the network if deployed.

Many solutions and algorithms have been proposed, which optimize both blocking probability and number of converters used. Considerable research towards this aim has been carried which lead to the development of *Semi-Transparent* or *Translucent optical networks* [30], [31], [32]. It has been found that by placing these converters at optimal places, considerable improvement in blocking probability, can be achieved without much increase in the cost of the network. *Translucent* optical networks thus exhibit a trade off between cost and blocking probability.

Routing and Wavelength Assignment (RWA) algorithms deals with objective of routing the lightpaths and assigning a wavelength for the route. A comprehensive review of these algorithms can be found in [33], [34]. However these RWA algorithms have a potential drawback. They assume data is transmitted on the assigned wavelengths to be error free. But in practice this is not the case. Signal quality along the path (or route) is degraded due to the physical impairments. Thus in spite of assigning a free wavelength, data cannot be transmitted in that path due high bit error rate (BER). In order to mitigate this problem, RWA algorithms should be aware of impairments in physical layer. In [35] this problem has been discussed for the first time. Since then there had been many algorithms proposed in this direction [36], [37], [38], [39], [40]. As a result Quality of Service (QoS) of signals in the network is measured by the BER. Thus in AON, apart from meeting the QoS constraint, lightpath must satisfy *QoS constraint*. In addition to QoS awareness, RWA algorithm should also have desirable *fairness* [41].

In this dissertation, we study the transparent or all-optical networks and hence impact on signal quality along its traversal. All-optical switches or optical cross connects (OXCs) remove the bottleneck of electronic conversion step in the core nodes. All-optical switches are been subject to research and few are commercially available [42], [43].

Simulation of RWA is a time-consuming process and there is a need to provide analytical solutions for on-demand provisioning of lightpaths. Although the problem of analytically computing blocking probability in optical networks has been studied in the past [44], [45], [46], [47] using various models and assumptions, the physical layer issues were rarely accounted [48], [49].

2.2 What is QoS Aware Routing ?

Lightpaths carry data traffic via all-optical WDM channel. Setting up a lightpath for a connection request by using a routing and wavelength assignment technique is called *connection provisioning*. Intelligent provisioning of these connections is an important traffic engineering problem for better utilization of the resources. Making the routing algorithm aware of under lying physical layer is necessary for ensuring better quality of the signal. This lead to the development of algorithms which are based on the transmission impairments. These are called *QoS Aware Routing Algorithms (QARA)*. Without physical-layer impairment awareness, the network-layer RWA algorithm might provision a lightpath which does not meet the signal-quality requirement. Therefore a control plane should incorporate the characteristics of physical layer in setting up of lightpath for a new connection. Hence the lightpaths are configured only if the BER is less than the required BER threshold. We thus see that AONs are subject to WCC and *QoS constraint*.

All optical switches contains, functionality like: multiplexers, demultiplexers, switching fabric, wavelength converters and wavelength regenerators. Last two are optional in an OXC. In our work we done not consider wavelength converters or regenerators, but however one can extend this work incorporating these two functionalities in the switching architecture. All optical or transparent optical networks should be able to deliver the signal with low bit error rates (BER) over long distances. As the distance increases the optical signal is subject to physical impairments in the fiber. These physical impairments can be in the form of

- inter symbol interference (ISI) due to the chromatic dispersion in the fiber,
- non-linear effects like self-phase modulation (SPM), cross-phase modulation (XPM), polarization mode dispersion (PMD),
- noise accumulated due to spontaneous emission noise (ASE) in the Erbium doped fiber amplifiers (EDFA) and
- crosstalk which is more predominant in the Dense WDM systems.

Issues of routing in the optical plane with reconfigurable network elements and transmission impairments has been studied in [50]. Effect of PMD on the design of wavelength routed optical networks has been addressed in [51]. Impact of non-linear impairments like FWM for preserving the QoS are discussed in [52].

2.3 Network Description

Fig. 2.1 shows the path of the optical signal from source to destination. It traverses many optical cross-connects at the the intermediate nodes. As there is no wavelength conversion assumed, optical signal remains on the same wavelength all along it's path. Fig. 2.2 shows optical cross-connect switch.

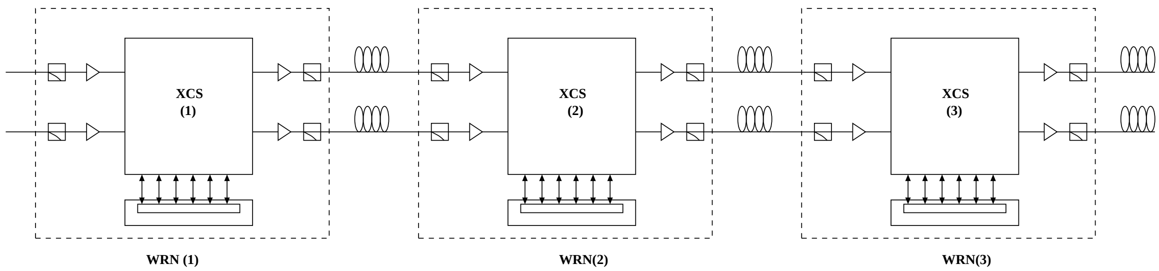


Figure 2.1: Optical nodes connected by WDM links with in-line amplifiers.

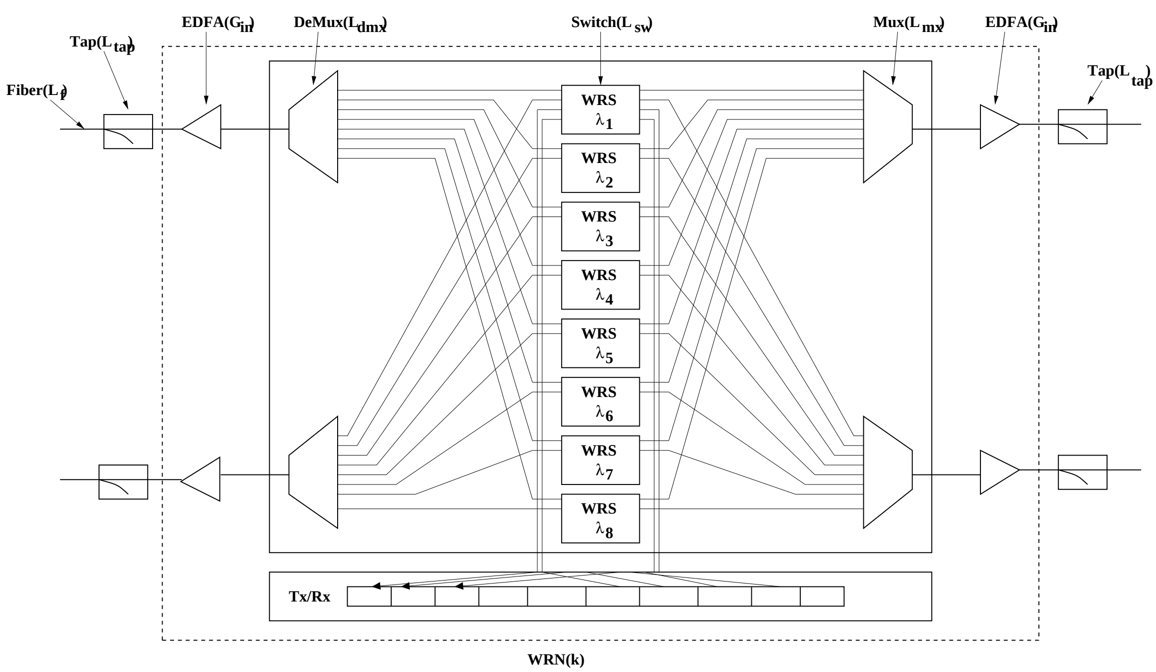


Figure 2.2: Optical Cross connect Switch.

Dynamic connection requests in WDM networks can be handled in a centralized or distributed way [53]. In the former, all paths to be established within a domain are computed and/or established by a single and centralized network element such as Network Management system (NMS) or path computation element (PCE). In the later, a distributed and intelligent optical control plane (OCP) embedded in each network element is responsible for route computation and lightpath establishment. In this work we follow a *distributed* connection establishment, where each cross-connect switch maintains the information about the BER of the optical signal. In this work we use distributed connection establishment (link-by-link) for the requests.

2.4 Simulation Results

Fig. 2.3 shows the National Science Foundation (NSF) network used for simulation. It consists of 14 nodes and 21 bi-directional links. The weights on the links indicate the physical fiber length between the nodes in kilometers. The architecture of the nodes in the Fig. 2.3 is shown in Fig. 2.2. We assume an in-line EDFA amplifiers placed in the link at a spacing of 70 kms. The network load is given by λ/μ, where λ is the burst arrival rate per second with Poisson process and μ being the service rate distributed exponentially. The source-destination pairs are uniformly distributed.

We consider the optical signal quality as the metric to qualify the link based on the BER. If the BER is greater than the required threshold the burst is dropped. Calculation of BER is based on the linear impairment losses like Amplified Spontaneous Noise (ASE) in the EDFA [4]. If we assume that the lightpath from the transmitter to the receiver goes through M optical amplifiers, with each introducing some noise power [4] and if each optical amplifier has same gain G, then we can obtain an upper bound on M, i.e., maximum number of spans, given by OSNR constraint as [50], [54],

$$M \leq \left\lfloor \frac{P_L}{2 \times OSNR_{min} \eta_{sp} h\nu (G-1) B_o} \right\rfloor \tag{2.1}$$

where P_L is the average optical power launched at the transmitter, η_{sp} is the excess

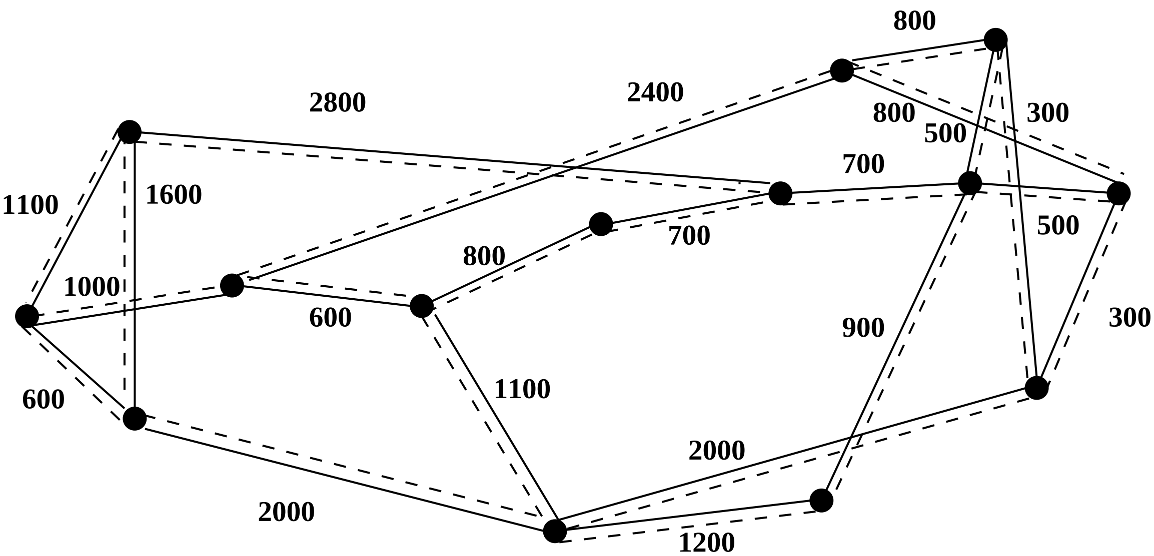


Figure 2.3: National Science Foundation network used for simulation.

noise factor, h is the Plank's constant, ν is the carrier frequency, and B_o is the optical bandwidth.

Signal and the ASE powers for all the source-destinations combinations for the NSF network are calculated. Fig. 2.4(a) shows the variation of signal power with the hop distance. At the initial node, signal transmitting power P_L is 1 mW. In the Fig. 2.4(a) we observe that the signal power increases with the increase in hop-distance for some of the source-destination combinations and decreases for other routes. Increase in the signal power is accounted for in-line EDFA used on these links. All the routes are based on the shortest-path distance between source and destination. In Fig. 2.4(b) we see that ASE noise linearly increases with the increase in the hop distance. Thus OSNR of the burst along the route decreases drastically as shown in the Fig. 2.5(a). For the routes having the hop distance of six, we see that OSNR of the burst at the destination is in the range of 2 to 10^{-1}. Fig. 2.5(b) shows the decrease in the q-factor of the optical signal with hop distance. If q-factor is less than 6 the BER would be greater than 10^{-9}. Thus we see that in spite of scheduling the burst on the link that is contention free, the burst can suffer with high BER at the destination, making it unacceptable. This indicates the need for making the burst scheduling schemes aware of the underlying physical layer.

Thus we see that burst can be lost due to lack of wavelength availability (contention)

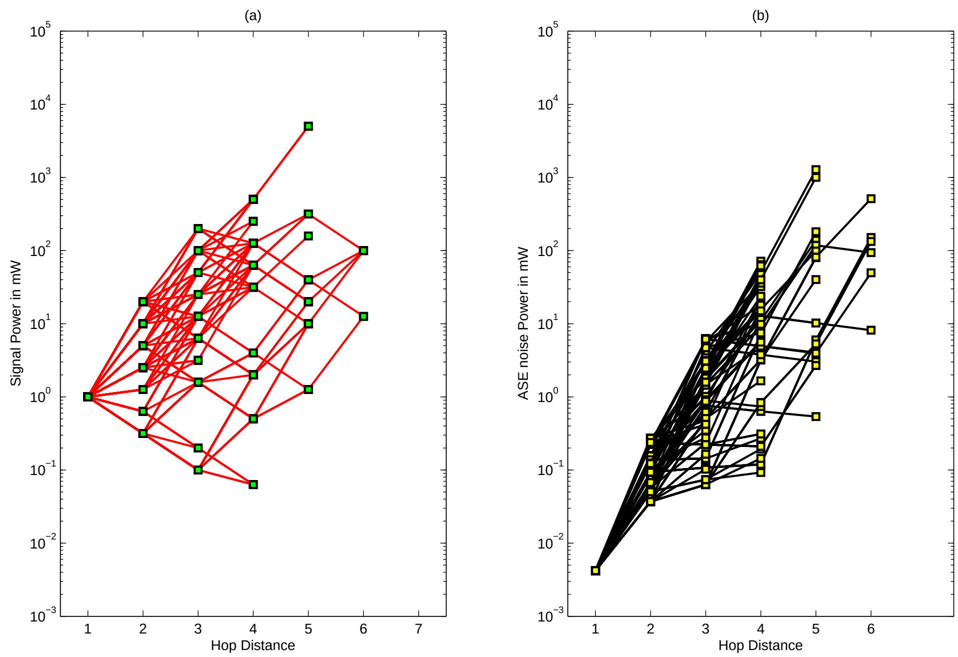


Figure 2.4: For each source and destination pair variation of (a) Signal Power, (b) ASE noise

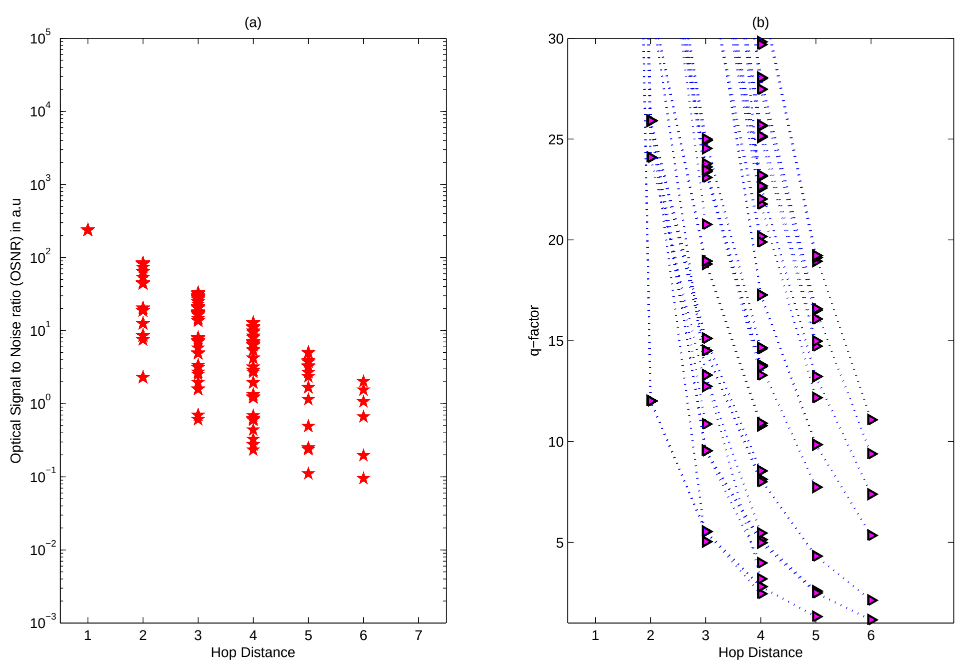


Figure 2.5: For each source and destination pair variation of (a) OSNR, and (b) q-factor in NSF network

or due to the insufficient signal quality (high BER). The performance of NSF network is studied based on network and physical layer constraints [54]. Performance of an any arbitrary network can be understood with blocking probability as the metric. Fig. 2.6(a) shows blocking probability versus network load for impairment aware and unaware conditions on a unicasting communication paradigm. Based on the OSNR constraint given in Eq. (2.1), blocking probability is calculated as show in Fig. 2.6(a). We observe from the Fig. 2.6(a) that taking physical layer to be ideal under estimates the blocking probability. We also observe that under low load conditions, the difference in the request blocking is very high compared to that under high load. This is due to the fact that at high load most of the burst are lost before due to contention, before they are really scheduled on the optical fiber link. Hence in the rest of this thesis, we consider the performance of the network at low loads. Fig. 2.7 shows the graph for request delay versus the network load. We see that the there is decrease in the average request delay for impairment-aware conditions than for impairment unaware conditions. When the burst is lost, then the delay of the burst is not taken into consideration, which causes an decrease in the average delay for impairment aware conditions.

2.5 Summary

In this chapter we have first introduced the need to study the impact of physical layer on routing and wavelength assignment problem. All-optical networks suffer with two constraints, (1) scarcity of number wavelengths and (2) decay in the optical signal quality. Assuming linear losses like ASE the blocking probability of the optical burst loss is calculated based on the two constraints mentioned above. In multicasting or manycasting the burst loss due to high BER could be more and hence there is need to develop schemes that reduces these losses. This aspect will be explored in the next chapter.

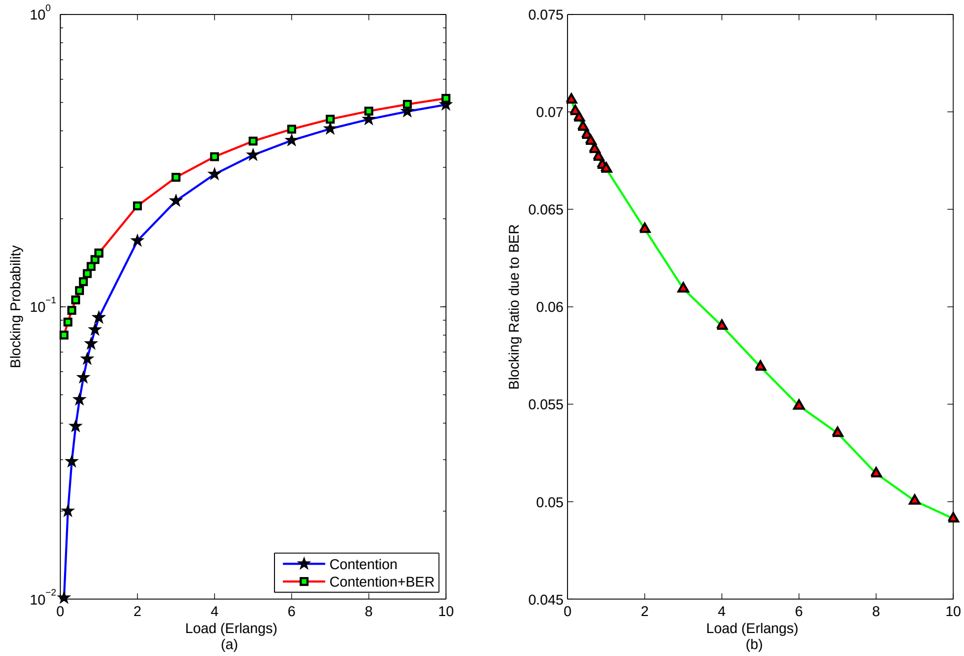


Figure 2.6: Performance of NSF network, for impairment aware and unaware conditions.

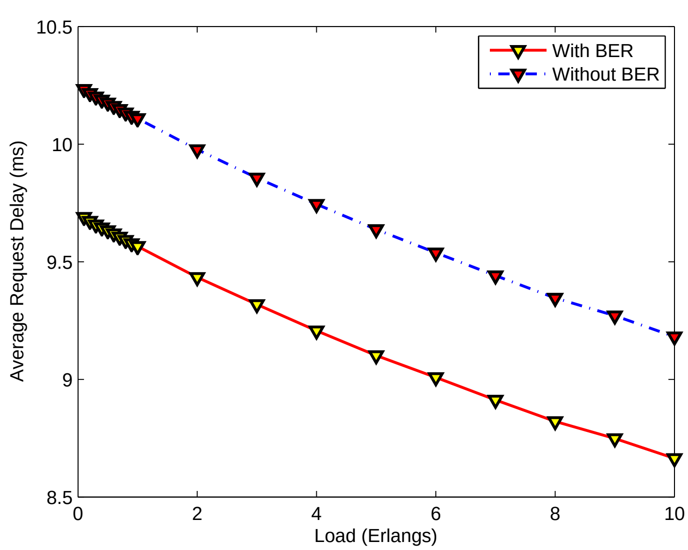


Figure 2.7: Delay performance of the shortest path (in ms)

Chapter 3

Impairment-Aware Manycasting Algorithms

In this chapter we discuss the effect of physical impairments on manycasting service over optical burst-switched (OBS) networks. Signal quality degradation in manycast networks is an important issue and it can occur due to fiber attenuation, splitter switch, and amplified spontaneous noise. These physical layer impairments causes the signal quality to be weak at the receiver and hence burst may not be detected or lost. Our objective is to select the best manycast destinations based on resource availability and quality of received signal. We propose three impairment-aware manycasting algorithms that take into account of the physical layer impairments. Using extensive simulation results, we compute average burst loss probability, both due to contentions and signal degradation. We develop analytical loss models for the proposed algorithms and verify them using simulation results. We have also compared our results with random destination selection (IP manycasting) using a Binomial model and observe that our algorithms outperform the random destination selection algorithm.

3.1 Introduction

Optical burst switched (OBS) network is a promising candidate to support high bandwidth Internet applications. These networks were proposed to overcome the technological constraints imposed by optical packet-switched networks. There has been recent emergence of many distributed applications that require high-bandwidth, such as grid computing, content distribution, and storage area networks. OBS networks have all the ingredients to support these applications. These applications require multiple destinations to be coordinated with a single source, and thus it seems multicasting is the way to implement these distributed applications. However in multicasting the destination set is fixed and the dynamic behavior of the network cannot be implemented. A variation in this is to dynamically vary the destinations depending in the status of the network. Hence in distributed applications, first step is to identify potential destination candidates and then select the required number. This is called *manycasting* and the problem is defined as follows: given a network $G(V, E)$, with V nodes and E edges, edge cost function is given by $g : E \rightarrow R^+$, an integer k, a source s, and the subset of candidate destinations $D_s \subseteq V$, $|D_s| = m \geq k$, where $|D_s|$ is the cardinality of the set D_s. If $k = 1$, one destination is chosen from the set D_s and this is called *anycasting* [55] [56].

Manycasting has caught attention of researches during the recent past, due to the emergence of many distributed applications [57]-[58]. Distributed applications, such as video conferencing, distributed interactive simulations (DIS), grid computing, storage area network (SAN), and distributed content distribution network (CDN) require large amount of bandwidths and an effective communication between single source and a set of destinations. Optical networks can provide huge bandwidth as required by these applications. Many of these distributed applications require user-controlled network infrastructure. Manycasting is also an attractive and viable communication paradigm for providing fault tolerance for the defense information infrastructures in battlefield [59]. Provisioning of connections based on QoS to these applications is an important issue [59].

In an OBS network, multiple packets to the same egress edge node are packed together in the form of single data burst at the ingress nodes. A control information for this data burst is transmitted ahead on separate channel and is called *burst header packet (BHP)*. BHPs are processed electronically at each intermediate node to reserve network resources before the data burst arrives at the node. After a certain offset time data burst is transmitted all-optically through the network.

Data loss in OBS network can occur either due to burst contentions or impairments in the fiber. *Burst contention* is a special issue in OBS networks, which occurs due to burstiness of IP traffic and the lack of optical buffering. Contention occurs when multiple bursts contend for the same outgoing port at the same time. Many schemes have been proposed to resolve burst contentions [60]. However all of these assume that the underlying physical fiber media is ideal. In other words, the burst that is allocated a wavelength is consider to be delivered error-free. But in practice this not the case. Bursts are transmitted all-optically in the fiber; they traverse through many optical components, such as fiber, multiplexer, demultiplexer, splitters and optical amplifiers. This causes the quality of the signal to degrade. Received signal have amplified spontaneous emission (ASE) noise due to optical amplifiers in the network [35]. The common metric to characterize the signal quality is optical-signal-to-noise ratio (OSNR), defined as the ratio of power of signal received to power of the ASE noise [61]. Multicast capable switches cause optical power to split depending on number of output ports. The power will be reduced as the signal propagates towards destination, thus decreasing OSNR. Bit error rate (BER) of the signal is related to OSNR. Decrease in OSNR causes an increase in BER. Thus a burst scheduled on a wavelength can be lost due to high BER of the signal. BER of the signal can be computed through q-factor [61]. If signal has low q, then BER of the signal is high and vice-verse. Thus a burst successfully scheduled on a wavelength, can be lost due to a low q. These impairments studies have been done extensively in past. Recent challenges are to develop impairment-aware routing algorithms before scheduling the data transmission [53]. As the first step toward implementing the impairment-aware manycasting, in this chapter we consider only the OSNR constraint. Therefore there is need to develop policies that implement manycasting considering both *burst contention* and *optical impairments*.

In this chapter we discuss the performance of different impairment-aware manycasting algorithms, using average blocking probability. Average blocking probability is computed with the help of discrete-event simulation model and later verified with help of analytical results. The rest of the chapter is organized as follows: Section 3.2 discusses issues of supporting manycasting over OBS networks. Section 3.3 and 3.4 discusses loss scenario in optical burst switched (OBS) network. In Section 3.5, we describe the proposed impairment-aware manycasting algorithms. Section 3.6 discusses the analytical model for the proposed manycasting algorithms. Results are presented in Section 3.7 where we compare our analytical results with simulation results. Section 3.8 summaries the chapter.

3.2 Manycasting Service

A manycast request is simply denoted by (s, D_s, k). We have to send the burst to k destinations out of m ($|D_s| = m$) possible candidate destinations. But due to the burst loss that occurs due to burst contention and/or signal degradation, there is no guarantee that exactly k destinations receive the burst. In general most multicasting solution approaches of are largely applicable to manycasting. Networks that can support optical multicast can also support optical manycasting. Thus, manycasting can be implemented by multicast-capable optical cross-connect (MC-OXC) switches as shown in Fig.3.2. Now when it comes to routing the burst, shortest-path tree (SPT) can be computed, as given below:

- *Step 1*: Find the shortest path from source s to all the destinations in D_s. Let $D_s = \{d_1, d_2, \ldots d_{|D_s|=m}\}$ and minimum hop distance from s to d_i, where $1 \leq i \leq m$ is $\mathbb{H}^{(s)} = \{h_1, h_2, \ldots h_m\}$.

- *Step 2*: All the destinations in D_s are sorted in the non-decreasing order according to the shortest distance from source s to the destinations. Let D'_m be the new set in this order given by $\{d'_1, d'_2, \ldots d'_m\}$.

- *Step 3*: Select the first k destinations from D'_m.

For a network of size n, each step requires the time-complexity of $O(n^2)$, $O(1)$, $O(n)$ respectively. If the shortest path distance to all the destinations are known, then the time-complexity of the SPT algorithm reduces to $O(n)$. We implement the SPT algorithm in distributed manner. *Step 1* is implemented by the unicast routing table. *Step 2* sorts the destinations at source node, in constant time. *Step 3* works as follows: First k destinations are selected from D'_m and BHP is sent to all next hop nodes (or child nodes). Let the child nodes be $\{c_1, c_2, \ldots, c_j\}$ where $1 \leq j \leq k$. Maximum number of child nodes can be k, if for each destination the next hop node is different. Upon receiving the BHP at the next-hop node, again the above mentioned three steps are under gone. The process ends when packet reaches all the k destinations or can discarded at the intermediate node (due to data loss). Even though the signal degradation along the shortest-path is less, but it is however not necessary that BER is within the threshold requirement. This indicates the need to bring the physical-layer awareness in manycasting algorithms, which are explained in the following sections. Bursts for the manycast are assembled in the same way as the unicast.When a burst is ready to transmit, a BHP will be sent out along the route for the manycast request [60]. The well-known OBS signaling protocols for unicast traffic, i.e., tell-and-wait (TAW), tell-and-go (TAG), just-in-time (JIT) and just-enough-time (JET) [62], can be used for manycasting with the modifications described in the above centralized or distributed version of the SPT algorithm. In this work manycasting is investigated with the help of just-enough-time signaling (JET).

Under JET source sends out a control packet which is followed by the burst after a certain period of time called *offset time*. This offset time is required because the control packet incurs processing delay at each switch while the burst does not. The offset time is at least Hd, where H is the number of hops between source and the destination, and d is the processing time incurred by the control packet at hop j, $1 \leq j \leq H$. Figure 3.1 shows the processing of control packet, before the burst arrives at the corresponding node. The bandwidth of the link j is reserved from the time the burst arrives (t) until the time it leaves ($t + L$), where L is the length of the burst.

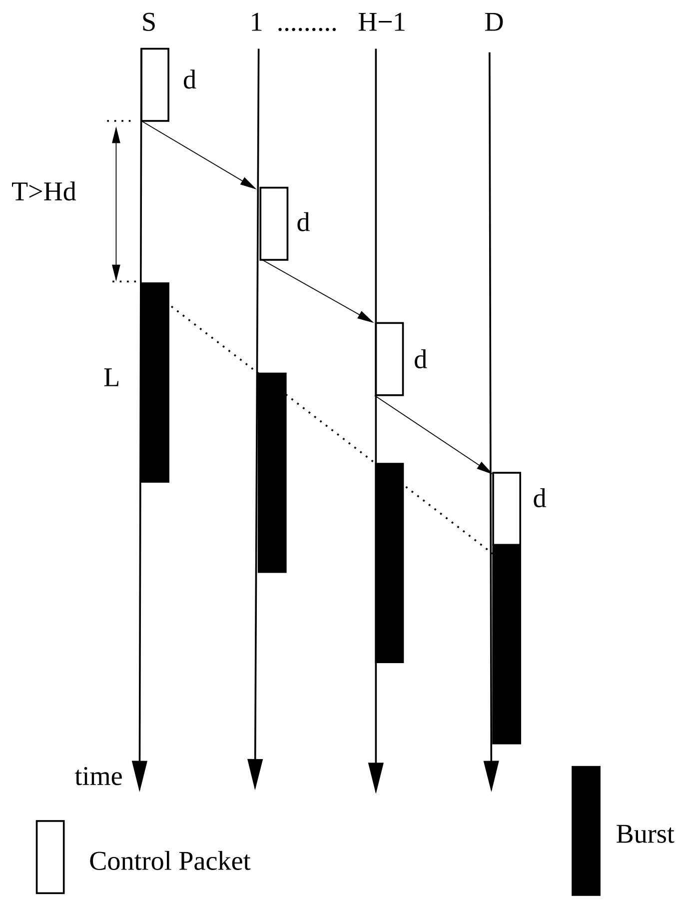


Figure 3.1: JET based signaling in OBS network

3.3 Contention Loss

A typical OBS network works as follows. Multiple packets to the same egress edge node are packed together into a *data burst* at ingress edge nodes. The control information for a data burst, contained in a *burst header packet (BHP)*, is transmitted on a separate control channel. BHPs are processed electronically at each intermediate node to reserve network resources before the data burst arrives at a node. Data bursts will then be routed all-optically on data channels through the network. This work has been presented in [60] and we explain it here for readers continuity.

Data loss due to *burst contention* is a special issue in OBS networks, because of the burstiness of IP traffic and the lack of effective optical buffering technique. Burst contention occurs when multiple bursts contend for the same outgoing channel on the same wavelength at the same time. There are many solutions to reduce the impact of burst contentions, such as in [63], [64], [65], [66]. In this work, we propose new schemes, like *shortest-path tree* (this is discussed in the previous section), *static over-provisioning* (SOP) and *dynamic membership* (DM), to alleviate the data loss problem in manycasting. The proposed schemes take into consideration the specific properties of manycast. The proposed schemes are not a replacement of existing contention resolution schemes but a complement to those schemes. That is, the proposed schemes could be used together with existing contention resolution schemes to further reduce data loss due to burst contentions.

Due to the lack of effective optical buffers in OBS networks and the one-way resource reservation mechanism, data loss resulting from burst contention cannot be avoided in typical OBS networks. The proposed schemes take into consideration the specific properties of manycast. The idea behind the new schemes is to improve data availability through controlled redundancy.

The first scheme, *static over-provisioning* (SOP) is motivated by the following observation. Because of the loss property of OBS networks, even at low load, we cannot guarantee that a burst will be received by k destinations, even if we could find the optimal solution (optimal k destinations out of D_c and the optimal route for the k destinations) for a manycast request (s, D_c, k). Instead of trying to avoid or resolve burst contentions such that a burst could reach the designated destination(s), we send the burst to a total of $k + k'$ destinations instead of only k destinations as indicated by the request, where $0 \leq k' \leq |D_c| - k$. As a special case, if $k' = 0$, there is no over-provisioning. If the bursts to some of the $k + k'$ destinations are lost, the total number of destinations which actually receive the burst may still be k or more with a high probability, such that the user requirement of a manycast request is satisfied. We then need to study two sub-problems: 1) how to decide the number k', given the request (s, D_c, k) and the network status; and 2) how to choose these additional k' destinations. In this work, we focus on the second sub-problem and evaluate the impact of a given k' on the network performance using SOP.

Here are the details of the SOP scheme. In SOP, the additional k' destinations will be decided before the burst (actually the BHP of the burst) is sent out from the source node. SOP could be used with either the centralized or the distributed version of the SPT routing algorithm, with a simple extension as follows: before the execution of *Step 3* in the SPT algorithm, we increase the value of k to the value of $k+k'$. By this extension, k' additional destinations, which have the shortest distances after k destinations have been selected, are included into the route tree. This extension is consistent with the idea of choosing the first k destinations which have the shortest distance from the source. Although we choose the additional k' destinations in this manner, alternatives are possible, such as choosing the additional k' destinations that have the least overlap with the route tree for the first k destinations.

The second scheme, *dynamic membership* (DM), takes a different perspective from SOP. In the SPT algorithm of Section 3.2 and the SOP scheme, the designated destinations and the route tree are decided at the source node, which is independent of network status. After that, the manycast request actually becomes a multicast request, which will be routed along the pre-calculated route tree. Since it is difficult for the source node to obtain the exact status information of nodes along the route tree, it may be a better choice to postpone such a decision until the burst actually arrives at intermediate nodes. In order to obtain such a flexibility, destination information should be included in the BHPs. Thus, the proposed DM scheme will work well with the distributed version of the

SPT algorithm.

Here are the details of the DM scheme. In DM, a designated set of k destinations is tentatively set up at the source node as before. Instead of discarding the remaining $(|D_c| - k)$ destinations, we evenly distribute the remaining destinations into all child branches at the source node. With the extra destinations, each burst that arrives at an intermediate node is still a manycast burst instead of a multicast burst. Then, if any designated destination is blocked at an intermediate node, we may send the burst to some of these extra destinations such that the total number of destinations which actually receive the burst is still no less than k. Therefore, in DM, the designated set of k destinations may change dynamically along the route tree according to the status of the network. In turn, the route tree itself may change accordingly. The algorithm for DM with the distributed version of SPT works as follows:

- [Input]: a manycast (u, D_u, k_u) arrives at Node u (the source or an intermediate node) with a candidate destination set D_u, among which k_u destinations are expected to be chosen as the actual destinations.

- [Output]: a list of (v_i, D_{v_i}, k_{v_i}) manycast requests to the next hop Node v_i, $i = 1, 2, \cdots, z$, where z is the number of child branches.

- *Step 1*: If $u \in D_u$, a copy of the burst will be dropped locally at Node u and update $D_u \leftarrow D_u - u$, $k_u \leftarrow k_u - 1$.

- *Step 2*: Sort the destinations in D_u in non-decreasing order according to the shortest distance from Node u to each destination.

- *Step 3*: Sequentially handle the destinations one by one from the ordered list until k_u destinations are successfully scheduled or all destinations are processed. For each destination $d_i \in D_u$, we find the next hop Node v_i to the destination from the unicast routing table. If Link $\langle u, v_i \rangle$ is freely available for the burst, $D_{v_i} \leftarrow D_{v_i} + d_i$, $k_{v_i} \leftarrow k_{v_i} + 1$. Otherwise, destination d_i is discarded.

- *Step 4*: For those untouched candidate destinations in *Step 3*, sequentially assign these nodes one by one d_i to the list of next hops Node v_i in a Round-Robin fashion, i.e., $D_{v_i} \leftarrow D_{v_i} + d_i$. Then, the algorithm terminates.

At the source node, the input to the above algorithm will be the manycast request itself (s, D_c, k). In *Step 4*, by distributing the untouched destinations evenly among the child branches, we may expect that each branch obtains some redundant protection from potential destination blocking. Note that DM is different from the well-known deflection routing scheme. In DM, if a destination is blocked, the destination is discarded and an alternative will be chosen to replace the blocked one. On the contrary, in deflection routing, if a destination is blocked on its primary route, an alternative route (if available) will be used to route the burst to the same destination.

With either SOP or DM, there is no guarantee that consecutive burst transmissions for the same manycast session will reach exactly the same set of destinations. We refer to this phenomena as *non-deterministic receiving* (NDR). NDR may be desirable or non-desirable, depending on the manycast application. For example, in quorum consensus, it is not required that the set of destinations which receive the data are exactly the same from transmission to transmission. Statistically, either SOP or DM can achieve some kind of load balancing. In this case, NDR is probably desirable. In another example, database protection via replication, a snapshot of the data should be sent through continual transmissions to the exact same set of sites among possible candidate sites. In this case, NDR will corrupt the integrity of the data. However, if the data can be transmitted with one transmission (such as placing it into one burst), NDR may again become desirable to achieve load balance.

3.4 Optical Impairment Loss

In this section we discuss the data loss due to physical impairments by computing q-factor. We first discuss the network architecture consisting of optical components that a signal traverses from source to destination in Section II-A. We then discuss the impairments, in such network, in Section II-B. We compute the quality factor of the signal on per hop basis in Section II-C. The parameters used for the computation of q- factor are tabulated in Table-I.

3.4.1 Network Architecture

Figure 3.2 shows the architecture for multicast-optical cross-connect (MC-OXC) using Splitter-and-Delivery (SaD) switch. As optical signal traverses from source to destination, it encounters losses due optical switches, multiplexer, demultiplexer, and fiber. Power loss can be compensated either by incorporating optical amplifiers or by increasing signal power at source. Fiber in-line amplification provided by the cascaded Erbium doped fiber amplifiers (EDFA), compensate the power loss due to attenuation in the fiber. However they increase the ASE noise in the channel, which in turn increases the BER. In this paper, we consider in-line amplification of signal, and hence the effect of ASE noise on the signal quality is used for the computation of BER. An $N \times N$ SaD switch proposed in [67] is used in the architecture for manycasting. It consists of N power splitters and N^2 2×1 optical gates which are used to reduce crosstalk and N^2 2×1 photonic switches as shown in Fig. 3.3. These switches are assumed to be configurable and hence can be instructed to split the incoming signal to any of $i = 1, \ldots N$ output ports [68], [69].

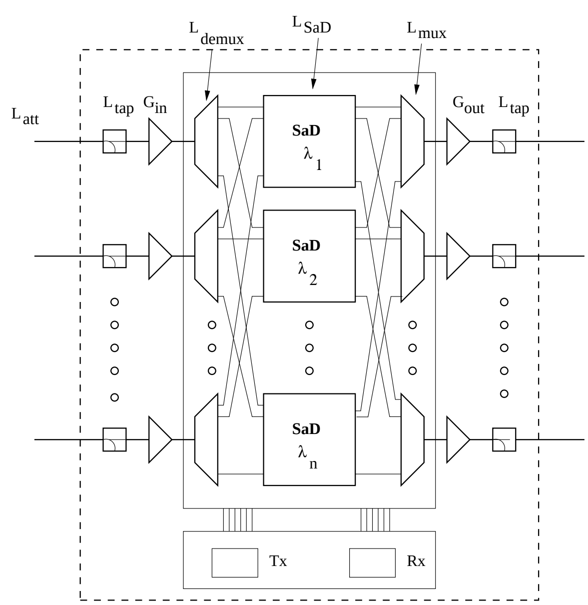


Figure 3.2: MC-OXC based on Splitter-and-Delivery Architecture.

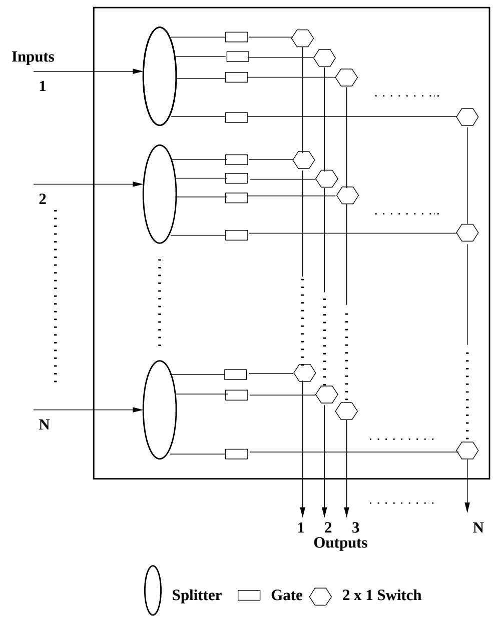


Figure 3.3: An $N \times N$ SaD Switch.

3.4.2 Calculation of q-factor on per-hop basis

- $L_{sp}(n) = 1/k(n)$ is loss due to the splitter at Node n, where $k(n)$ is the number of the output ports to which the signal is split, defined as *fan-out* of the splitter. If $k = 1$, then there is no splitting at the node and hence $L_{sp}(n) = 1$.

- L_n is physical distance between the nodes $\langle n, n+1 \rangle$, l is the distance between two amplifiers, then a_n, the number of amplifiers used between $\langle n, n+1 \rangle$ is given by,

$$a_n = \left\lceil \frac{L_n}{l} \right\rceil - 1. \tag{3.1}$$

We define l_n as the distance of fiber which is not been compensated by in-line amplification, given by:

$$l_n = L_n - a_n \times l. \tag{3.2}$$

- $L_{att}(n) = e^{-\alpha l_n}$ is loss due to the attenuation in the fiber, where α is the attenuation of the fiber.

- L_d, L_m, and L_t are defined as demultiplexer, multiplexer and tap losses, respectively.

- $L_{ins} = 2\log_2 NL_s + 4L_w$ is insertion loss of the SaD switch, where L_s is switch element insertion loss and L_w is waveguide or coupling loss and N is number of fibers, which is equal to number of input/output ports of the switch.

- G_{in} and G_{out}, are gains of the input and the output EDFA, respectively. Define $G_T = G_{in}G_{out}$ as the total gain provided by the amplifiers at the node.

- $\bar{G}$ is the saturated gain of the in-line EDFA. This gain is set to compensate the fiber loss between consecutive amplifiers given by $\bar{G} = e^{\alpha l}$.

- $P(n)$, $P_{ase}(n)$ are the signal and amplified spontaneous emission (ASE) noise power output at the n^{th} node, respectively.

- B_o and B_e are the optical and electrical bandwidths.

Recursive Power Relations: Here we derive recursive power relations similar to [35]. However the only difference, is that we consider in-line amplification and we use SaD switch instead of OXC. The output power at the Node n, $P(n)$ and is given by,

$$\begin{aligned} P(n) &= G_{in}G_{out}L_dL_mL_t^2L_{ins}L_{att}(n)L_{sp}(n-1)P(n-1), \\ &= G_TL_kL_{att}(n)L_{sp}(n-1)P(n-1), \\ &= G_TL_T(n-1)L_{sp}(n-1)P(n-1), \end{aligned} \tag{3.3}$$

where $L_k = L_dL_mL_t^2L_{ins}$, this loss is a constant for any node and $L_T(n-1) = L_kL_{att}(n)$.

$$\begin{aligned} P_{ase}(n) &= P_{ase}(n-1)L_T(n-1)G_T + G_{out}P_nL_T(n-1)\left[G_{in}-1\right]/L_t + \\ &\quad P_nL_t\left[G_{out}-1\right] + P_n[\bar{G}-1]a_n \end{aligned} \tag{3.4}$$

where $P_n = 2n_{sp}hf_cB_o$ with typical values given in Table-I. Due to in-line amplification of the signal using EDFA, there will be ASE noise along the route. Hence the last term in Eq. (3.4) represents the ASE noise along the fiber, and the first two terms represent the ASE noise due to EDFAs inside the node. We assume that this is as constant, when the wavelengths are centered around f_c. In the system of cascade amplifiers, the notion of sensitivity is not very useful when signal reaching the receiver has already added lot of noise [61]. In this case two parameters that are measured are, the average received signal power, $P(n)$ and received optical noise power $P_{ase}(n)$. The optical signal to noise ratio (OSNR) at node n is given by $OSNR(n) = P(n)/P_{ase}(n)$. By neglecting the receiver thermal noise and shot noise, the relationship between the q-$factor$ and OSNR is given by [61] [1],

$$q(n) = \frac{2\sqrt{\frac{B_o}{B_e}}\mathrm{OSNR}(n)}{1+\sqrt{1+4\mathrm{OSNR}(n)}}, \tag{3.5}$$

where $q(n)$ is defined as the quality factor of the link between nodes $\langle n, n+1\rangle$. Bit error rate of link n is given by,

$$BER(n) = 2 \text{ erfc}\left(\frac{q(n)}{\sqrt{2}}\right), \tag{3.6}$$

where, $erfc(x)$ is called complementary error function.

3.4.3 Assumptions

1. In the recursive power equations we have chosen the gain of the amplifiers (input/output) to be a constant, i.e., gain saturation effects of the amplifier are not considered.

2. We have assumed that q-factor is independent of the wavelength chosen. This assumption is valid when the wavelength spacing is less. Hence the carrier frequency f_c is chosen to be the central frequency of the wavelength band.

3. Signal degradation due to cross-talk and non-linearity in fiber have been ignored in the computation of q- factor.

4. Dispersion loss due to Polarization Mode Dispersion (PMD) is ignored.

[1]Derivation is given in Appendix A.1

3.4.4 Online Evaluation of q-factor using Burst Header Packet (BHP) Signaling

In a manycast scenario, we have the request in the form of (s, D_s, k), with $|D_s| = m$. In order to identify the best set of k destinations, we need to have the best possible tree, both in terms of load and quality (in other words high $q(n)$). Assuming the link to be free, we can route the optical signal. However the link may have a bad q value which in-turn results is high BER. If BER is greater than 10^{-9} then the signal cannot be recovered. Thus by keeping the threshold value for the BER we ensure that the signal received is acceptable. High BER corresponds to low q, so we say optical signal to be lost when q falls below the threshold value, q_{th}. Thus, the burst that was assumed to be transmitted by the network layer, cannot be recovered by the core node and is actually lost before reaching to *egress* node. The BHP used to reserve the channel for the OXC can also be used to make OXC aware of the q-factor. BHP can in-corporate the new field that stores q value. Initially, the q-field is set to a high value, and once BHP reaches the next node q value is updated using the recursive Eqs. (3.3, 3.4, 3.5). At every intermediate node, the BHP updates the q and checks the condition, $q > q_{th}$. If this is true the BHP proceeds further, else the burst is dropped. Burst loss due to signal impairment is defined as *optical-layer blocking*.

Successfully reception of the optical burst at the *egress* node is based on two issues, contention and link impairments. The manycasting algorithms proposed in [60] are modified to consider these two issues and are discussed in the next section.

3.5 Impairment-Aware Manycasting Algorithms

In this section we ensured the manycasting algorithms proposed in [60] to consider the signal degradation due to impairments in the fiber. In order to consider impairment-awareness during burst transmission, we modify the manycast request as $(u, D'_u, k_u, P(u), P_{ase}(u))$, where the last two tuple indicate signal and noise power respectively. u can be a source s or an intermediate node, with sorted destination set D'_u and intended number of destinations k_u. In all the algorithms considered, we have

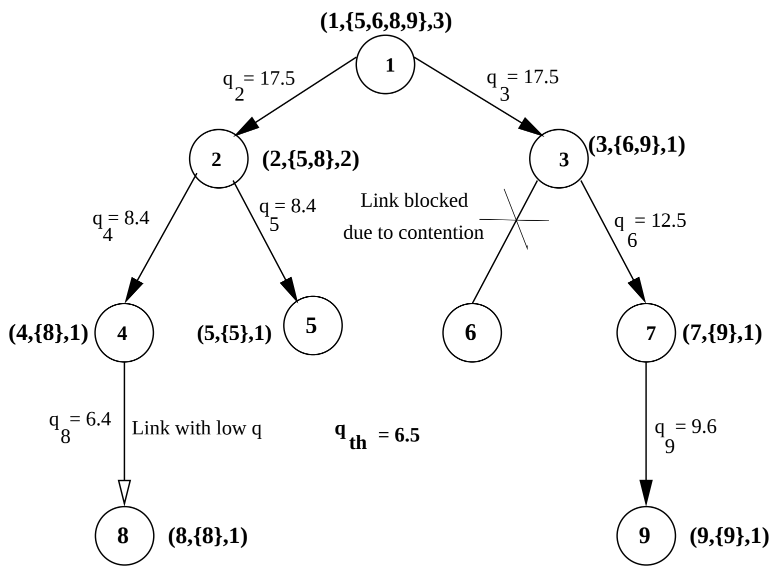
1. **Input:** The manycast request (u, D'_u, k_u) arrives at the source node with a candidate destination set D'_u, along with k intended destinations. The power inputs for this manycast request are $(P(u), P_{ase}(u))$. Hence we have $(u, D'_u, k_u, P(u), P_{ase}(u))$.

2. **Output:** Manycast request to the next-hop node after satisfying the BER constraint.

3. **Initialization:** At the source node, the manycast request is of the form $\big(s, D'_s, k_s, P(s), P_{ase}(s)\big)$. For every new burst entering the network, this manycast request is tagged to it. All other sets are initialized to null.

3.5.1 Impairment-Aware Shortest Path Tree (IA-SPT)

IA-SPT algorithm uses a precomputed shortest path tree. Based on the three steps mentioned in Section-II, the tree is constructed for each manycast request. Recursive power relations in Section 3.4 can be used to compute the OSNR of the optical signal along its path. If the link from the source node to one of the child nodes is free, then q is computed. If the q-factor is above the threshold value, q_{th}, then the channel is scheduled for burst transmission. Hence, the successful reception of the burst at the destination node guarantees that signal is error-free. This continues until k destinations are reached. If the burst reaches $< k$ destinations, then the manycast request is said to be blocked. As the IA-SPT is implemented on the pre-computed routing tree, it does not consider the dynamic nature of the network. This algorithm suffers from high burst loss, due to fixed routing along the shortest path tree and this is verified by simulation results. Other algorithms proposed, decrease the burst loss in the presence of optical layer impairments. In the pseudo-code lines 2-3 ensure that if the current node is the destination node then the destination set (D'_u) and intended number of destinations (k_u) are updated. These lines remain same for all the three algorithms used. Child nodes or the next-hop node set for Node u, are calculated using lines 5-8. For all child nodes the channel availability is checked using Line 10. Using recursive power relations described



Node	P(n)=Signal Power (mW)	P_{ase}(n)=ASE Power (mW)	Splits
1	1.00	0.0042	2
2	0.40	0.0110	2
3	0.40	0.0110	1
4	0.16	0.0160	1
5	0.16	0.0160	–
6	–	–	–
7	0.32	0.0160	1
8	0.12	0.020	–
9	0.25	0.020	–

Figure 3.4: Example explaining the three proposed algorithms works.

in Section-3.4, q-factor is computed and if the threshold condition is met, then we say that all the destinations corresponding to the child node n_j can be reached and this set is given by $S_D(n_j)$. $S_D(n_j) = k_u$ only when there is one child node for all the destination in D'_u. The new manycast request is thus formed at the child node n_j as given in Line 15. $\mathbb{D}$ is the set of all destinations that can be reached from node u. If $|\mathbb{D}| < k_u$, then the request is said to be blocked and probability of the request blocking is given by $1 - |\mathbb{D}|/k_u$.

Consider the example given in Fig. 3.4, in the case of IA-SPT, we select first $k_u = 3$ from $D'_c = \{5, 6, 8, 9\}$, i.e., $\{5, 6, 8\}$. As both the conditions in lines 10 and 14 are met, we have $S_D(2) = \{5, 8\}$ and $\mathbb{D} = \{5, 8\}$ and the new manycast request at $n_j = 2$ becomes $(2, \{5, 8\}, 2, P(2) = 0.4, P_{ase}(2) = 0.011)$. When $i = 2$, we have $n_j = 3$ and if the conditions are met then we have $S_D(3) = 6$ and $\mathbb{D} = \{5, 8\} \cup \{6\}$, that implies $|\mathbb{D}| = k_u$ and hence request $\big(1, \{5, 6, 8, 9\}, 3, P(1) = 1, P_{ase}(1) = 0.0042\big)$ is successful.

3.5.2 Impairment-Aware Static Over Provisioning (IA-SOP)

IA-SOP algorithm is similar to IA-SPT except that here we will not limit the number of destinations to k, but we send the burst to $k + k'$ destinations, where k' is such that $0 \leq k' \leq m - k$. With $k' = 0$, IA-SOP is similar to IA-SPT, i.e., no over-provisioning. In this algorithm, first $k + k'$, destinations are selected from the set D'_c. Sending the burst to more than k destinations ensures that it reaches at least k of them. However by doing over-provisioning the fan-out of the splitter increases, which increases BER. In spite of decrease in the contention loss, there is no significant improvement in the overall loss. From the simulation results we see that IA-SOP shows better performance than IA-SPT. The algorithm for IA-SOP is similar to that of IA-SPT, but with k_u replaced with $k_u + k'$. Thus the probability of request blocking is given by $1 - \min(|\mathbb{D}|, k_u)/k_u$. This is because if all the $k_u + k'$ are free then the burst is sent to more destinations than intended (i.e., k_u), but from the user perspective we have only k_u to be reached. If $|\mathbb{D}| > k_u$ implies $\min(|\mathbb{D}|, k_u) = k_u$, then the request blocking ratio is zero.

Consider the example shown in the Fig. 3.4, if we select $k'_u = 1$, then we have first $k_u + k'$ of D'_c as $\{5, 6, 8, 9\}$ and at two child nodes $2, 3$ the manycast requests are

Algorithm 1 Impairment Aware-Shortest Path Algorithm (IA-SPT)

Input: The manycast request (u, D'_u, k_u) arrives at the source node with a candidate destination set D'_u, along with the k intended. The power inputs for this manycast request are $\big(P(u), P_{ase}(u)\big)$. For clarity we denote the manycast request by $(u, D'_u, k_u, P(u), P_{ase}(u))$. where $P(u)$, $P_{ase}(u)$ are the signal and ASE powers at node u.

Output: Manycast request to the next hop node after satisfying the BER constraint.

Initialization: At the source node, the manycast request is of the form $\big(s, D'_s, k_s, P(s), P_{ase}(s)\big)$.
/* Update D'_u and k_u.*/
if $u \in D'_u$ **then**
    $D'_u \leftarrow D'_u \backslash \{u\}$.
    $k_u \leftarrow k_u - 1$.
    /* Destination set D'_u is the non-decreasing order of the hop distance*/.
else
    for $j \leftarrow 1$ to k_u **do**
        $n_j \leftarrow SPT[u, d'_j]$
        /* Next hop or child node is obtained from shortest path tree*/
        $N = N \cup \{n_j\}$
    end for
    for $i \leftarrow 1$ to $|N|$ **do**
        if $(\langle u, n_i \rangle = FREE)$ **then**
            $P(n_i) \leftarrow POW\_SIGNAL\big(P(u), |N|\big)$
            $P_{ase}(n_i) \leftarrow ASE\_SIGNAL\big(P_{ase}(u)\big)$
            $q(n_i) \leftarrow Q\_FACTOR\big(P(v_i), P_{ase}(v_i)\big)$
            if $\big(q(n_i) > q_{th}\big)$ **then**
                $D_{n_i} \leftarrow D_{n_i} \cup \{S_D(n_i)\}$
                /* $S_D(n_i)$ is the set of all destinations ($\subseteq \{d'_1, \ldots, d'_{k_u}\}$) that can be reached through child node n_i. $|S_D(n_i)| \leq k_u$ */
                $(n_i, D_{n_j}, |S_D(n_i)|, P(n_i), P_{ase}(n_i))$
                $\mathbb{D} \leftarrow \mathbb{D} \cup S_D(n_i)$
            else
                $DEST[n_j]$ are not reachable due to high BER.
            end if
        else
            $DEST[n_j]$ are not reachable due contention.
        end if
    end for
end if

Algorithm 2 Impairment-Aware Static Over Provisioning (IA-SOP)

Input: The manycast request (u, D'_u, k_u) arrives at the source node with a candidate destination set D'_u, along with the k intended. The power inputs for this manycast request are $\big(P(u), P_{ase}(u)\big)$. For clarity we denote the manycast request by $(u, D'_u, k_u, P(u), P_{ase}(u))$. where $P(u)$, $P_{ase}(u)$ are the signal and ASE powers at node u.

Output: Manycast request to the next hop node after satisfying the BER constraint.

Initialization: At the source node, the manycast request is of the form $\big(s, D'_s, k_s, P(s), P_{ase}(s)\big)$.
/* Update D'_u and k_u.*/
if $u \in D'_u$ **then**
  $D'_u \leftarrow D'_u \backslash \{u\}$.
  $k_u \leftarrow k_u - 1$.
  /* Destination set D'_u is the non-decreasing order of the hop distance*/.
else
  for $j \leftarrow 1$ to $k_u + k'_u$ **do**
    $n_j \leftarrow SPT[u, d'_j]$
    /* Next hop or child node is obtained from shortest path tree*/
    $N = N \cup \{n_j\}$
  end for
  for $i \leftarrow 1$ to $|N|$ **do**
    if $(\langle u, n_i \rangle = FREE)$ **then**
      $P(n_i) \leftarrow POW\_SIGNAL\big(P(u), |N|\big)$
      $P_{ase}(n_i) \leftarrow ASE\_SIGNAL\big(P_{ase}(u)\big)$
      $q(n_i) \leftarrow Q\_FACTOR\big(P(v_i), P_{ase}(v_i)\big)$
      if $\big(q(n_i) > q_{th}\big)$ **then**
        $D_{n_i} \leftarrow D_{n_i} \cup \{S_D(n_i)\}$
        /* $S_D(n_i)$ is the set of all destinations $(\subseteq \{d'_1, \ldots, d'_{k_u} \ldots, d'_{k_u+k'_u}\})$ that can be reached through child node n_i. $|S_D(n_i)| \leq k_u + k'_u$ */
        $(n_i, D_{n_j}, |S_D(n_i)|, P(n_i), P_{ase}(n_i))$
        $\mathbb{D} \leftarrow \mathbb{D} \cup S_D(n_i)$
      else
        $DEST[n_j]$ are not reachable due to high BER.
      end if
    else
      $DEST[n_j]$ are not reachable due contention.
    end if
  end for
end if

$\big(2, \{5, 8, 9\}, 2, P(2) = 0.4, P_{ase}(2) = 0.011\big)$, $\big(3, \{6, 9\}, 2, P(3) = 1, P_{ase}(3) = 0.011\big)$, respectively (assuming links $\langle 1, 2\rangle$, $\langle 1, 3\rangle$ are free and the q-factor is greater than the required threshold). Thus the request is successful.

3.5.3 Impairment-Aware Dynamic Membership (IA-DM)

IA-DM takes the dynamic network status into consideration. Instead of selecting the destinations before the burst is transmitted, we dynamically add members as possible destinations, depending on contention and quality of the link. IA-DM will work with a distributed version of SPT. The set of k-destinations is tentatively set up at the source node. We do not discard the remaining $m - k$ destinations, but instead keep them as child branches at the source node.

IA-DM algorithm is explained with an example shown in Fig. 3.4. Consider the manycast request $(1, \{5, 6, 8, 9\}, 3, P(1) = 1, P_{ase} = 0.0042)$ with signal and ASE powers as shown in Fig. 3.4. The table in the Fig. 3.4 shows the number of splits, input signal power and ASE power at each node. The output of IA-DM algorithm gives the manycast request at the next-hop node with signal and ASE values. These two values can be used to compute q-factor and thus qualify the outgoing link. $\mathbb{V}$ represents set of next-hop nodes (or child-nodes) for Node u, $\mathbb{Q}_{\mathbb{L}}$ represents set of nodes that have low q-factor, and $\mathbb{C}_{\mathbb{L}}$ represents set of nodes that are blocked due to contentions. These sets are initialized to *null* before the start of the algorithm. When the request arrives, and if $u \in D'_u$ then the burst is received locally and request is updated as shown in the lines 1-3. The set $\{5, 6, 8, 9\}$ is the sorted set of candidate destinations in the non-decreasing order of the hop-distance. Assuming Link $\langle 1, 2\rangle$ is free, $\mathbb{V}$ is updated, and the signal power, ASE power received at Node 2 are computed. Note that there is no split ($|\mathbb{V}| = 1$) and q-factor is computed as in lines 10-12. The condition for threshold is checked and thus the destination set at the next-hop node is updated. Lines 19-20 ensure that the number of destinations at all the child nodes does not exceed k_u, the number of destinations at the current node. The loop in line-5, is executed for all destinations. Hence the next destination in the order of non-decreasing hop-distance is Node 6. The child node for the current Node 1 is 3 and hence Link $\langle 1, 3 \rangle$ is checked for contention. If it is free then the split takes places at Node 1 and the power is divided equally among nodes 2 and 3 ($|\mathbb{V}| = 2$). Note that ASE power remains unchanged. Thus the new power and q values are computed using lines 10-12.

As the power of the signal is split at node 1, the new manycast request at the next-hop nodes 2 and 3 becomes, $(2, \{5, 8\}, 2, P(2) = 0.4, P_{ase}(2) = 0.011)$ and $(3, \{6, 9\}, 2, P(2) = 0.4, P_{ase}(2) = 0.011)$ respectively. At node 2, the next-hop nodes are $\{4, 5\}$ and the burst is scheduled assuming the links $\langle 2, 4 \rangle$ and $\langle 2, 5 \rangle$ are free. The new manycast request are updated accordingly. But as node 5 is the destination node lines 2-3 in the algorithm ensure that routing of the burst terminates at node 5. Along node 4 the burst continues to route towards node 8. As shown in the Fig. 3.4 we see that q-factor at node 8 is slightly lesser than the required threshold of $q_{th} = 6.5$. On the other side of the tree as link $\langle 3, 6 \rangle$ is blocked, so the burst has to be routed to other destination, which is at a longer distance then the node 6. Assuming all the links towards node 9 are free, we see that the q-factor is much lesser than it would have been at node 6 (if the link towards the node 6 was free). Thus we see that in IA-DM destinations can be added or removed dynamically and this decreases the request blocking in comparison with IA-SPT.

As we see here that the number of child nodes are not fixed and hence power relations need to be recomputed according to the split. Thus the processing delay will be more in case of IA-DM, when compared to the other two, i.e., IA-SPT and IA-SOP.

3.6 Analytical Model

In this section, we present the analytical model for the manycasting scenario. OBS signaling based on just-enough-time (JET) uses one-way reservation protocol in which data burst follows the control packet (BHP) without waiting for acknowledgment. As we use decentralized version of SPT which is based on node-to-node, we model the destination blocking rather than the session blocking. We use the Erlang-B model proposed by [70]. For a given output port of the switch the burst arrival process follows a Poisson process. Let w be the number of wavelengths used at each port. If we have only class of burst

Algorithm 3 Impairment Aware-Dynamic Membership Algorithm (IA-DM)

Input: The manycast request (u, D'_u, k_u) arrives at the source node with a candidate destination set D'_u, along with the k intended. The power inputs for this manycast request are $\big(P(u), P_{ase}(u)\big)$. For clarity we denote the manycast request by $(u, D'_u, k_u, P(u), P_{ase}(u))$. where $P(u)$, $P_{ase}(u)$ are the signal and ASE powers at node u.

Output: Manycast request to the next hop node after satisfying the BER constraint.

Initialization: At the source node, the manycast request is of the form $\big(s, D'_s, k_s, P(s), P_{ase}(s)\big)$.
/* Update D'_u and k_u.*/
if $u \in D'_u$ **then**
    $D'_u \leftarrow D'_u \backslash \{u\}$.
    $k_u \leftarrow k_u - 1$.
    /* Destination set D'_u is the non-decreasing order of the hop distance. */
else
    for $j \leftarrow 1$ to $|D'_u|$ **do**
        $n_j \leftarrow UNI\_CAST\big[u, d'_j\big]$
        if $(\langle u, n_j\rangle = FREE)$ **then**
            $\mathbb{V} \leftarrow \mathbb{V} \cup \{n_j\}$
            for $\mathbb{V} \leftarrow \mathbb{V} \cup \{n_j\}$ **do**
                $P(v_i) \leftarrow POW\_SIGNAL\big(P(u), |\mathbb{V}|\big)$
                $P_{ase}(v_i) \leftarrow ASE\_SIGNAL\big(P_{ase}(u)\big)$
                $q(v_i) \leftarrow Q\_FACTOR\big(P(u), P_{ase}(u)\big)$
                if $\big(q(v_i) > q_{th}\big)$ **then**
                    $D_{v_i} \leftarrow D_{v_i} \cup \{d(v_i)\}$.
                    /* $d(v_i)$ is the destination to be reached through child node v_i.*/
                else
                    $D_{v_j} \leftarrow D_{v_j} \backslash \{d(v_i)\}$
                    $\mathbb{Q}_L \leftarrow \mathbb{Q}_L \cup \{d(v_i)\}$
                end if
            end for
            while $\sum_{k=1}^{j} k_{n_k} < k_u$ **do**
                $k_{n_j} \leftarrow k_{n_j} + 1$
            end while
        else
            $\mathbb{C}_L \leftarrow \mathbb{C}_L \cup \{d_j\}$
        end if
    end for
end if

arrivals and the remaining offset time is equal for each burst at any switch , [70] the OBS system behaves exactly like an $M/G/w/w$ system for which exact blocking probability can be obtained using the Erlang B formula. Note that Erlang B is insensitive to service time distribution and hence we can use the $M/M/w/w$ queuing model for modeling blocking probability for the proposed impairment-aware manycasting. Let us define $\mu = 1/E[L]$, where $E[L]$ is the expected (or average) value of the burst length and λ be the arrival rate. Then the unicast load is given by $\rho = \lambda/\mu$. In manycasting, there are k intended destinations that have to be selected from m. We keep k to be the majority of the group, so we have $k \geq \lceil m/2 \rceil$. We define over-provisioning factor $\beta = k'/k$. β lies in the interval $0 < \beta < 1$ as long as we take $k \geq \lceil m/2 \rceil$ (Readers must note that this is not generalized inequality and hence does not hold for anycasting where $k = 1$). Hence the effective manycast load for IA-SPT and IA-SOP is given by,

$$\begin{aligned} \rho_m &= \lambda/(k\mu) \text{ for IA-SPT and} \\ &= \lambda\beta/(k\mu) \text{ for IA-SOP.} \end{aligned} \tag{3.7}$$

The manycast request blocking probability due to the contention B_C in the network is given by Erlang-B model [71] as

$$B_C = \frac{\rho_m^w/w!}{\sum_{i=0}^{w} \rho_m^i/i!}. \tag{3.8}$$

Bursts are scheduled once the links along its path to the destination are available. However, all the scheduled burst do not meet the BER requirement of the network. So some of the bursts will be dropped in spite of occupying a free channel. This blocking is referred to as *optical layer blocking* blocking and can be defined as,

$$\begin{aligned} B_Q &= \frac{\text{\# manycast requests dropped due to high BER}}{\text{\# manycast requests that find a free channel}} \\ &= \frac{\text{\# bursts dropped due to high BER } (q < q_{th})}{\text{\# bursts that find a free channel}}. \end{aligned} \tag{3.9}$$

Thus the overall blocking probability including *contention* B_C and *optical layer blocking* B_Q is given by,

$$\begin{aligned} B_{total} &= B_C + B_Q - B_C B_Q \\ &= B_C + (1 - B_C) B_Q. \end{aligned} \tag{3.10}$$

In the Eq. (3.10) we have considered blocking due to contention and insufficient BER are independent.

IA-DM adds and removes candidate destinations depending on whether or not the link is contention free. We use Poisson splitting to evaluate the blocking probability for IA-DM manycasting. Each burst carries the information about the destination set and the intended destinations k. If a particular child branch toward the destination in k of D'_c is blocked then, that destination is removed and a new destination is added from first $m - k$ thus maintaining total intended destination to be k (this is unlike the deflection routing where the burst is sent to the same destination, but through an alternative route). First k are selected from D'_c, i.e., $\{d'_1, \ldots, d'_k\}$. We define this destination set as *primary destinations* (D'_p). If any of the d'_i, $1 \leq i \leq k$ is blocked (with probability say $\widehat{q}$), then to satisfy the manycast request a destination is selected from the other $m - k$ destination set, i.e., $\{d'_{k+1}, \ldots, d'_m\}$ and we define this set as *secondary destinations* (D'_s). Secondary destinations are only used when at least one of the primary destination cannot be reached through its child nodes. Having partitioned D'_c into two disjoint sets, we model the arrival process using *Poisson splitting*. Let λ be the unicast arrival rate into the network. These arrivals are split into primary and secondary arrivals as independent arrivals based on the outcome of the Bernoulli trial with probability $\widehat{q}$, given by

$$\widehat{q} = \frac{(\lambda/\mu)^w / w!}{\displaystyle\sum_{i=0}^{w} (\lambda/\mu)^i / i!}. \tag{3.11}$$

This is similar to probability $\widehat{q}$ of failure in a Bernoulli trial, referred as *randomization* or *Poisson split* [71]. However note that split of arrival processes into two Poisson processes is valid only when each arrival is independent of assignment of other arrivals. We assume that we have, an estimate for contention blocking as $\widehat{q}$ and split the traffic based on the outcome of an experiment. Let us define X_r be a random variable which takes either 0 or 1. Thus we have,

$$X_r = \begin{cases} 1 & \text{if } d'_j \in D'_p,\ 1 \leq j \leq k, \text{ w.p. } (1-\widehat{q}) \\ 0 & \text{if } d'_j \in D'_s,\ k+1 \leq j \leq m, \text{ w.p. } \widehat{q}. \end{cases} \tag{3.12}$$

Hence we have two independent Poisson processes with arrival rates $\lambda_p = \lambda(1-\widehat{q})$ and $\lambda_s = \lambda\widehat{q}$ for primary and secondary destination sets, respectively. Thus the manycast load in case of IA-DM is given by,

$$\rho_m = \begin{cases} \lambda_p/k\mu & (\equiv \rho_m^{(p)}) \text{ for } D'_p \\ \lambda_s/(m-k)\mu & (\equiv \rho_m^{(s)}) \text{ for } D'_s. \end{cases} \tag{3.13}$$

In $\rho_m^{(s)}$, the denominator is the cardinality of D'_s i.e., $m-\ k$. Thus the overall manycast request blocking for IA-DM is given by,

$$B_{total}^{(IA-DM)} = B_{total}^{(p)} + (1 - B_{total}^{(p)})B_{total}^{(s)}, \tag{3.14}$$

where $B_{total}^{(p)}$ and $B_{total}^{(s)}$ are blocking probabilities of primary and secondary destinations, respectively, obtained from Eq.(3.10), with manycast loads as given by Eq.(3.13).

3.6.1 IP Manycasting

Selection of k destinations out of m by the IP layer is similar to the *random algorithm* in [57], we also present a simple analytical model for the manycasting with random selection of k destinations. Our results show that random selection of destinations has poor performance, hence supporting manycasting at the OBS layer is necessary. A manycast request is said to be blocked if the burst reaches less than k destinations. Hence given there are k destinations, probability that at least one of them is blocked is given by,

$$B_C^{(bino)} = \sum_{i=0}^{k-1} \binom{k}{i} (B_C)^i (1 - B_C)^{k-i}. \tag{3.15}$$

Hence the total blocking is,

$$B_{total}^{(bino)} = B_C^{(bino)} + (1 - B_C^{(bino)}) B_Q. \tag{3.16}$$

3.7 Numerical Results

In this section we present our simulation and analytical results. We consider *average request blocking* as performance metric. We define average request blocking ratio as given by [60]. Let f be the total number of requests used in the simulation. Consider a manycast request (s, D_s^f, k). Let $\mathbb{D}$ be the set of destinations which actually receive the data. Then *average request blocking* is given by,

$$B_{total}^{(Sim)} = \sum_f \big[1.0 - \min(|\mathbb{D}|, k)/k\big]/f. \tag{3.17}$$

We use notation m/k, which means $|D_s| = m$ and k intended destinations. As in [57] [59], we consider the candidate destinations set D_s at small, medium, and large sizes, and the intended destinations is a majority of the group. Three typical configurations, 3/2, 7/4, 11/6 were used. We use NSF network as shown in the Fig. 3.5 for our simulation studies. The numbers on the links indicate the distance between the nodes in kilometers. These links consists of in-line EDFAs spaced 70 kms apart (not shown in the figure for clarity). All the links in the network are bi-directional and have same transmission rate of 10 Gb/s. Burst arrivals follow Poisson process with an arrival rate of λ bursts per second. The length of the burst is exponentially distributed with expected service time of $1/\mu$ seconds. The unicast network load is then defined as λ/μ. The source and candidate destinations of a manycast request are evenly distributed among all the nodes. There are no optical buffers or wavelength converters in the network. We consider a single wavelength plane and hence $w = 1$ in Eq. (3.8),(3.11). The physical layer parameters

Table 3.1: Parameters used for computation of q-factor

Parameter	**Value**
Channel bit rate (B)	10 Gb/s
Optical Bandwidth (B_o)	70 GHz
Electrical Bandwidth (B_e)	$0.7 \times B$
Input power of the signal	1 mW (0 dBm)
Loss of Multiplexer/Demultiplexer	4 dB
Switch element insertion loss	1 dB
Waveguide fiber coupling loss	1 dB
Tap loss	1 dB
Fiber Attenuation Coefficient	0.3 dB/km
Gain of EDFA in MC-OXC (G_{in}, G_{out})	22 dB, 16 dB
ASE factor (n_{sp})	1.5
Planks Constant h	6.63×10^{-34} J-s
Carrier frequency f_c	193.55 THz
P_n in Eq. (3.4)	$2n_{sp}hf_cB_o$
Spacing between the amplifiers (l)	70 kms
q_{th}	6.5
Number of fibers/link (N)	2 (bi-directional)

used in the simulation model are shown the Table I. For all graphs, x-axis indicates unicast load.

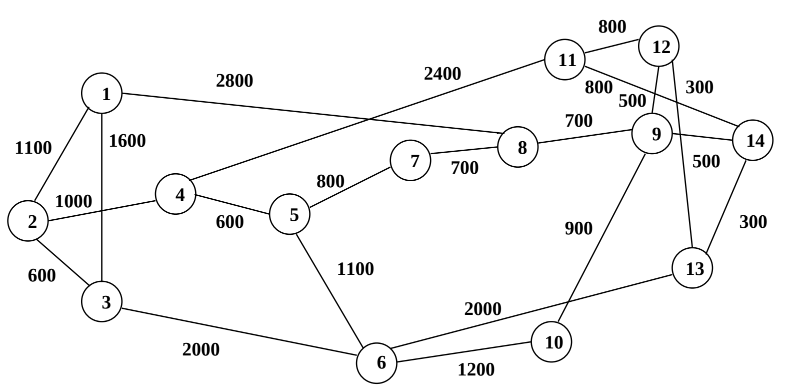


Figure 3.5: The NSF network consisting of 14 nodes and 21 bi-directional links.

With the help of discrete-event simulation model the average request blocking for contention (without impairment aware) is computed using the Eq.(3.17). This contention blocking is verified with the analytical model proposed in Section 3.6. In this case we assume underlying physical layer to be ideal. Figs. 3.6, 3.7, show the Comparison of

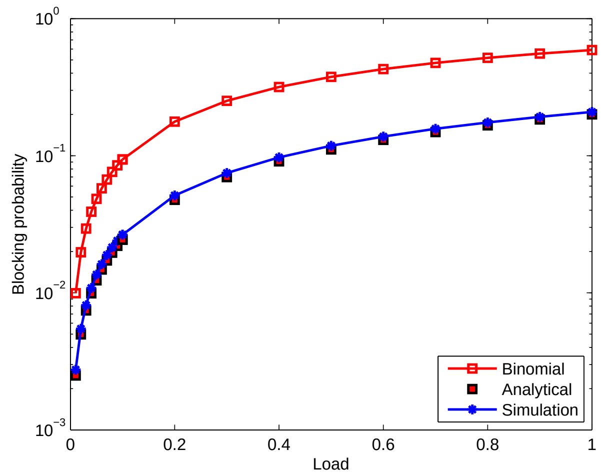


Figure 3.6: Comparison of Blocking Probability for Low and Medium loads for SPT with 7/4 manycast configuration

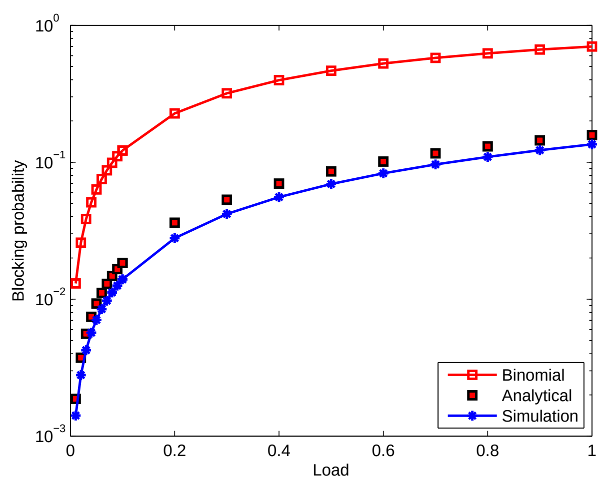


Figure 3.7: Comparison of Blocking Probability for Low and Medium loads for SOP with 7/4 manycast configuration

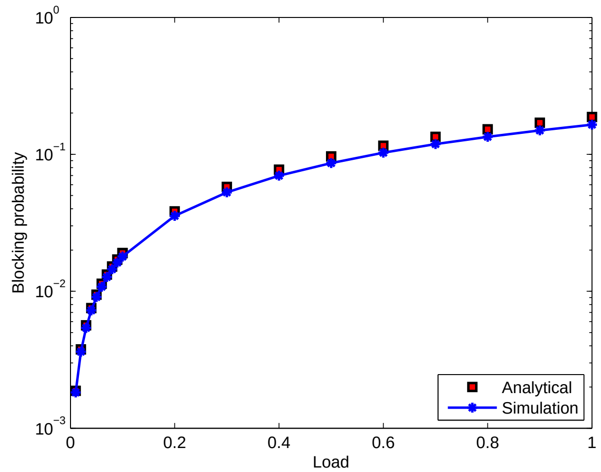


Figure 3.8: Comparison of Blocking Probability for Low and Medium loads for DM with 7/4 manycast configuration

simulation with analytical results for SPT and SOP respectively. Also we see that random selection (Binomial Model) of destinations by the IP layer has high probability of blocking and hence suffers from poor performance. Fig. 3.8 shows the Comparison of simulation and analytical results for DM. In these results the low and medium loads were considered because at high loads all the schemes perform almost similar.

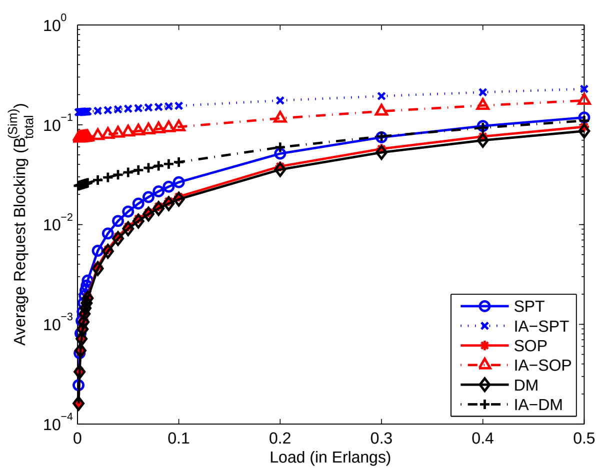


Figure 3.9: Comparison of algorithms with and without impairment awareness.

Using discrete-event simulations we compute $B_{total}^{(Sim)}$ using Eq. (3.17) and compare our results for without impairment-awareness, as given in [60]. Fig. 3.9 show the comparison of impairment-aware average request blocking to regular algorithms. From these graphs we observe there is significant difference in $B^{(Sim)}_{total}$ under low load conditions. This is because under low load conditions, contention blocking will be less and hence regular algorithms used in [60] does not provide the correct estimate of blocking. From the Fig. 3.9 we also observe that IA-DM has lower blocking than IA-SPT and IA-SOP and thus, impairment-aware manycasting over OBS, can be improved by using IA-DM. From the Fig. 3.9 we observe that without impairments all the three algorithms perform almost similar. However in the presence of impairments there is is significant reduction in the burst loss, when IA-DM is used. Our simulation results show that even under high loads IA-DM is better than the other two as shown in Fig. 3.10.

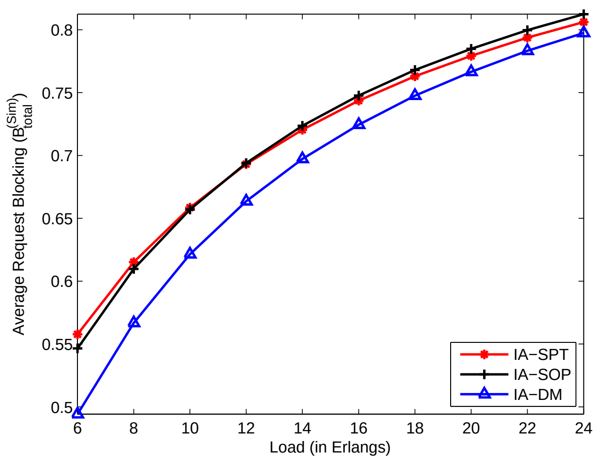


Figure 3.10: The blocking performance comparison between IA-SPT, IA-SOP and IA-DM for manycast configuration 7/4 under High load.

We validate our simulation results with the analytical model explained in Section 3.6. Fig. 3.11 shows that our model is accurate for IA-SPT. This graph also indicates that random selection of k destinations from D_c (IP-Manycasting) has poor performance compared IA-SPT. Significant reduction in the blocking can be achieved by using IA-SPT.

From Fig. 3.12 we observe that our analytical model over-estimates the blocking probability of IA-SOP at low loads. This is due to the size of intended destinations. In our case we have $k' = 3$, which is equivalent to multicasting. However at high loads these results converge.

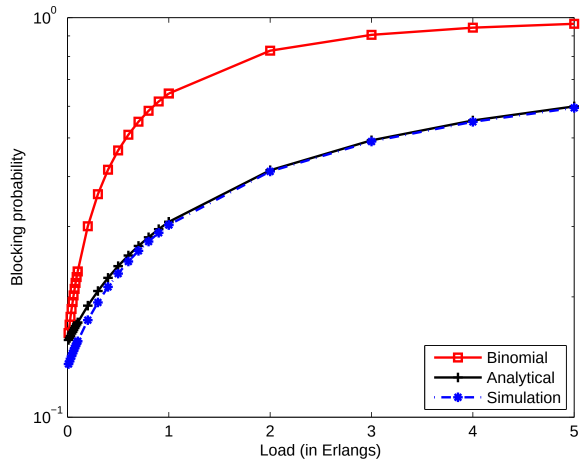


Figure 3.11: Comparison of Binomial, Analytical and Simulation results for overall blocking probability for IA-SPT under low load with 7/4 manycast configuration.

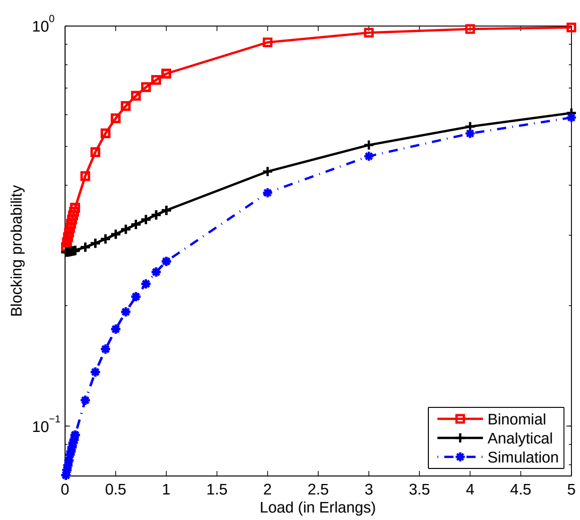


Figure 3.12: Comparison of Binomial, Analytical and Simulation results for overall blocking probability for IA-SOP with $k' = 3$ under low load with 7/4 manycast configuration.

Finally we validate our simulation results for IA-DM using Poisson-splitting. From Fig. 3.13 we observe that Poisson split model slightly over-estimates the blocking probability than simulation. This is because of the Eq. (3.17) does not distinguish between primary and secondary destinations as in Poisson split. However the difference being very small, it provides a good estimate for the impairment-aware manycasting. Also by using Poisson-splitting we maintain the arrival process to secondary destinations as Poisson distribution and this makes analysis computationally efficient. In the Fig. 3.13 we also compare our results without split, which clearly validate our simulation results.

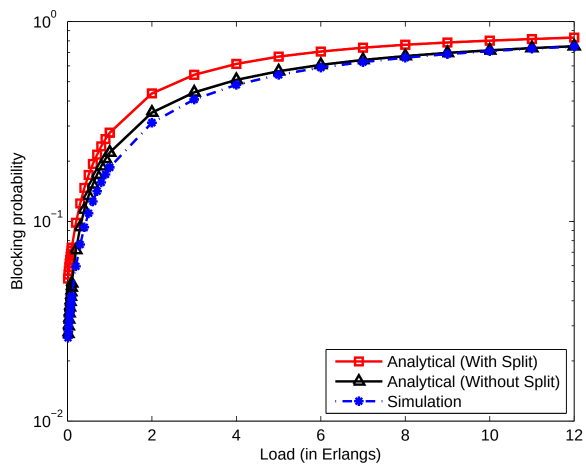


Figure 3.13: Comparison of Analytical (with and without Poisson split) and Simulation results for overall blocking probability for IA-DM under low load with 7/4 configuration.

We compare the performance of other manycast scenarios like 3/2 and 11/6. Fig. 3.14 shows the average request blocking for 3/2 manycast configuration. We observe that there has been a significant reduction is the burst loss using IA-DM when compared with other two algorithms IA-SPT and IA-SOP. This reduction is attributed to the decrease in the contention loss. Static behavior of the IA-SPT and IA-SOP causes more requests to be blocked due to contention. In the Fig. 3.14 we also see that the comparison of these algorithms, having no impairments. However the improvement in the blocking probability for DM is small.

Comparison of these algorithms for 11/6 manycast configuration is given in Fig. 3.15. In the case of 11/6 the algorithms SOP and DM have almost same loss. However in the

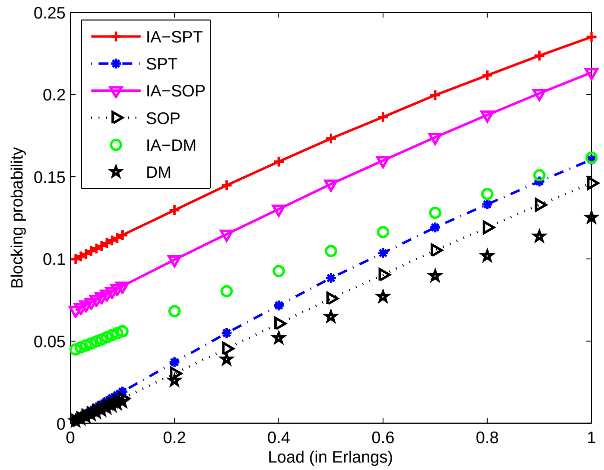


Figure 3.14: Comparison of algorithms with and without impairment aware for manycast configuration of 3/2.

presence of impairments, there is significant decrease in the burst loss for IA-DM when compared with IA-SOP and IA-SPT for the same given load. Thus we see that for all manycast configurations IA-DM performs better in terms of request blocking.

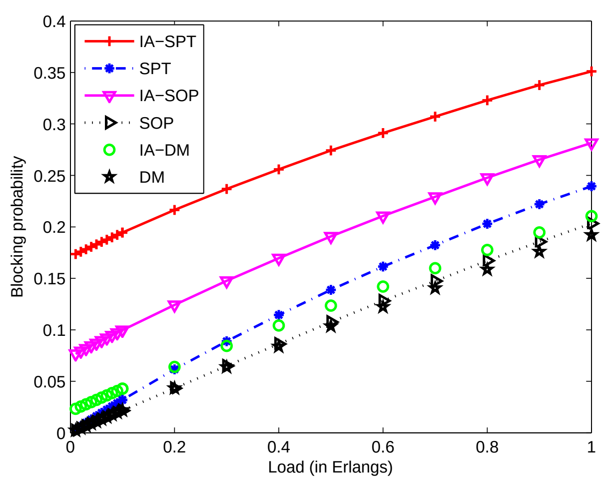


Figure 3.15: Comparison of algorithms with and without impairment aware for manycast configuration of 11/6.

3.8 Summary

In this chapter, we discuss issues related to burst loss in optical burst switched networks (OBS). We indicate the need for supporting manycasting by the OBS network, to improve the performance in terms of burst loss. We propose algorithms to decrease the burst loss both for contention and impairment loss. As these algorithms use a de-centralized way of scheduling the burst, computation of q-factor can be done by per hop basis rather than end-to-end. We indicate that BER based signaling using BHP has significant impact in calculating data loss in OBS networks. We propose three impairment-aware algorithms IA-SPT, IA-SOP and IA-DM. Through extensive simulation and numerical analysis, we show that IA-DM has better performance for supporting manycasting over OBS. Analytical models for the algorithms has been proposed. We finally verify our discrete-event simulation models with analytical results.

Chapter 4

QoS Based Manycasting for OBS Network

Many distributed applications require a, group of destinations to be co-ordinated with a single source. Multicasting is a communication paradigm to implement these distributed applications. However in multicasting, if at least one of the members in the group cannot satisfy the service requirement of the application, the whole multicast request is said to be blocked. On the contrary in manycast (or quorumcast) destinations join or leave the group, depending on whether it satisfies the service requirement or not. This dynamic membership based destination group decreases the request blocking. In this work we study the behavior of manycasting over optical burst switched networks (OBS) based on multiple quality of service (QoS) constraints. These multiple constraints can be in the form of physical-layer impairments, transmission delay, and reliability of the link. Each application requires its own QoS threshold attributes. Destinations qualify only if they satisfy the required QoS constraints set-up by the application. We propose a decentralized way of routing the burst towards its destination. With the help of local-network state information, available at each node the burst is scheduled only if it satisfies multiple set of constraints. Correspondingly reception of the burst at the node ensures that all the QoS constraints are met and burst is forwarded to the next-hop. Due to multiple constraints, burst blocking could be high. We propose algorithms to minimize request blocking for

Multiple Constrained Manycast Problem (MCMP). With the help of simulations we have calculated the average request blocking for the proposed algorithms. Our simulation results show that MCM-shortest path tree (SPT) performs better than MCM-dynamic membership (DM) for delay constrained services and real-time service, where as data services can be provisioned using MCM-DM.

4.1 Introduction

In Chapter 3 we have seen a need to bring the BER awareness in the optical-control plane. However OBS network needs an more effective routing algorithms to support Quality of Service (QoS) for the distributed applications. Three approaches for the routing include,

- Route the control packets on hop-by-hop basis, as in IP network, using the fast look-up algorithm to determine the next-hop
- In Multi-Protocol Label Switched (MPLS) network, a packet is marked with a label, which is used to route the packet through the network.
- Constrained-routing version of MPLS can be used to explicitly setup routes. This explicit routing is very useful in a constrained based routed OBS network, where the traffic routes have to meet certain QoS metrics such as delay, hop-count, BER or bandwidth.

In this work we use the constrained based version of routing for providing necessary QoS to manycast communication paradigm. Later in this chapter we show that the performance of the routing algorithm can be improved with the use of *adaptive* routing. For instance the traffic is usually forwarded along the shortest-path, however an equal and higher hop count are also identified and are used in a Multi-Path Routing (MPR) strategy. MPR uses alternative routes when the shortest path is congested. Due to MPR techniques, the burst tend to choose the longer paths and hence the QoS parameters such as, BER, delay and reliability of the path can be degraded. Hence there is need to obtain an trade-off between the least congested path and QoS.

In this chapter we propose algorithms that provide QoS in manycasting over OBS networks. We also propose a mathematical framework for destination selection policies based on QoS constraints as required by certain applications. Our approach can incorporate multiple constraints related to different services. The proposed methods are service-centric and completely decentralized, as they use only local-network state information. The rest of the chapter is organized as follows: We first discuss the related work in this topic in Section 4.1.1. In Section 4.2 mathematical formulation for ordering destinations based on service constraints is discussed. In Section 4.3 we explain the proposed algorithm with the help of an illustrative example. Section 4.4, discusses the performance evaluation of algorithm with the blocking probability as the metric. Finally, Section 4.5 summaries this chapter.

4.1.1 Related Work

Manycasting work was first reported independently by [57] and [72] as quorumcast problem and κ-Steiner tree problem. It is defined as an edge cost function $g : E \rightarrow R^{+}$, an integer κ, a Source s and the subset of candidate destinations $D_s \subseteq V$, $|D_s| = m \geq \kappa$, find a minimum spanning κ destinations in D_s. Cost of the tree is the sum of the cost of edges on the tree. Manycast request can be denoted by (s, D_s, κ). The manycast problem is found to be NP-hard in [72]. As IP layer is above the WDM/OBS network, the selection of κ destinations by the IP layer is similar to the random algorithm in [57]. In [73] this random algorithm has been verified using Binomial model and found to provide poor performance. Thus supporting manycasting in OBS networks is necessary for bandwidth-efficient manycasting [60]. Apart from constructing minimum cost tree that spans from Source s, to manycast group members, the need for QoS routing has been discussed in [59]. This paper discuss the quality of a tree in terms of source-destination delay constraints imposed by applications that use the tree. As delay-constrained quorumcast routing problem is NP-complete, an efficient heuristic QoS routing algorithm has been proposed in [59] with cost of the quorumcast tree close to that of optimal routing tree.

Apart from supporting manycasting over optical networks, we also need to provision QoS in OBS networks. This is because, QoS provisioning methods in IP will not apply to optical counterpart, as there is no store-and forward model [74]. Such mechanisms for QoS provisioning in IP over OBS networks must consider the physical characteristics and limitations of optical domain. Physical characteristics of the optical domain include optical-signal degradation which is an important concern in transparent optical networks, propagation delay incurred from source to destination, especially in OBS networks and reliability of the link from catastrophic effects. When the optical signal traverses in the transparent optical network, where there is an absence of electrical regenerators there will be significant loss of power due to many impairments. These impairments can be amplified-spontaneous emission (ASE) noise due to EDFA amplifiers, attenuation loss, multiplexer/demultiplexer loss, optical-cross connect switch loss (OXC), and split loss (for mutlicast capable switches) [75]. Incorporating electrical regenerator can significantly increase the cost of the network and end-to-end propagation delay can be high because of O/E/O conversion. Challenges and requirements for introducing impairment-awareness into the management and the control planes in WDM networks has been discussed in [53]. Decrease in optical signal to noise ratio (OSNR) increases bit error rate (BER) of the signal and hence the signal can be said to lost if BER is more than the required threshold [35], [37]. Manycasting (or multicasting) requires the OXC to split the signal. Multicasting over optical network can be done by the OXC switch incorporating the splitter-and-deliver (SaD) switch [67]. Depending on the fan-out of the switch the input power significantly decreases compared to unicast, thus decreasing OSNR. Multicasting under the optical layer constraint has been discussed in [68]. Power-efficient multicasting for optimizing BER has been studied in [69]. For the first time, impairment-awareness for implementing manycasting over OBS networks has been addressed in [75]. This paper discusses the importance of physical layer awareness and computes the burst loss due to contention and high BER. Further in [73] performance of different algorithms has been discussed and an analytical model has been proposed for calculating burst loss probability.

Reliability is an important issue in designing storage-area networks (SAN). SAN are supported over fiber-channel (FC), and hence threat to failure can occur due to cable cuts, physical attacks, catastrophic effect. Reliability factor is thus necessary for end-to-end path. The work proposed in [76] discusses about reliability for SAN. Analytical models are developed for calculation of long-term failures, service availability, and link failures. Reliability factor as the multiplicative constraint has been discussed in [77], [78]. Optical burst switched networks meet the requirements of computationally intensive Grid based applications known as Grid OBS (GOBS) [24]. Success of the grid depends on the quality of service the network can provide to ensure successful completion of the job [79]. So it is necessary to provision services such as average end-to-end delay, and BER in GOBS networks. Performance analysis of end-to-end propagation delay and blocking probability for OBS based grids using anycasting has been presented in [25]. Different types of anycasting algorithms has been compared in [55] with the shortest-path unicast routing, where the destinations has the specific address. Manycasting over OBS networks based on multiple resources has been addressed in [80].

4.2 Mathematical Formulation

4.2.1 Notations

(s, D_s, κ) is a manycast request where s is the source node, D_s is the destination set, elements of which are probable candidates for the particular service request, κ is the minimum number of destinations that are required to participate in the manycast session for the job to be successfully completed. Manycast session is also denoted by m/κ. Manycasting can be understood as the dynamic version of the multicast, where in the members can leave and new members can join, so that κ of them will always participate in the session [81]. Number of ways κ of them can be selected from $|D_s| = m$ is $\binom{m}{\kappa}$. We define a set $\wp_\kappa(m)$ called the power set which contains all the $\binom{m}{\kappa}$ combinations. Our work focuses on selecting the best possible set $\in \wp_\kappa(m)$ which can meet the service demands effectively. Most of the distributed applications, for example Grid computing requires BER to be low, high reliability of the path, and low propagation delay. Destinations chosen must be able to provide these quality of service attributes. In this section we present a mathematical framework to quantify a destination based on these three service attributes. A destination is said to qualify as the member of quorum pool, if it satisfies the service requirements of the application. Notation used for describing lattices can be found in [82].

4.2.2 Lattices

The lattice structure is described here, for the reader to understand few sections of this chapter. It is best explained by the special relation $\preccurlyeq$ on an arbitrary set say A, called *ordering*, which has the following properties,

$$\Omega_1 \preccurlyeq \Omega_1 \;\forall\; \Omega_1 \in A \qquad \text{reflexivity} \tag{4.1}$$

$$\Omega_1 \preccurlyeq \Omega_2 \wedge \Omega_2 \preccurlyeq \Omega_1 \implies \Omega_1 = \Omega_2 \;\forall\; \Omega_1, \Omega_2 \in A \qquad \text{anti-symmetry} \tag{4.2}$$

$$\Omega_1 \preccurlyeq \Omega_2 \wedge \Omega_2 \preccurlyeq \Omega_3 \implies \Omega_1 \preccurlyeq \Omega_3 \;\forall\; \Omega_1, \Omega_2, \Omega_3 \in A \qquad \text{transitivity} \tag{4.3}$$

If all the elements in the set A are comparable, which is equivalent to $\Omega_1 \preccurlyeq \Omega_2$ or $\Omega_2 \preccurlyeq \Omega_1$ being defined, the ordering is total and A is the totally ordered set or a chain. If all of the comparisons are not defined, A is called *partially ordered* set or *poset*. For such kind of ordering a *lower* and *upper* bound of each subset $B \subseteq A$ exist. Let those elements be $\perp$ and $\top$ for which,

$$\forall\, \Omega \in B,\ \perp \preccurlyeq \Omega \tag{4.4}$$

and

$$\forall\, \Omega \in B,\ \Omega \preccurlyeq \top \tag{4.5}$$

If the least or greatest elements of such bounds exist, they are called the *greatest lower bound (g.l.b)* or *infimum* and *least upper bound (l.u.b) supremum.* Lattice is an ordered set

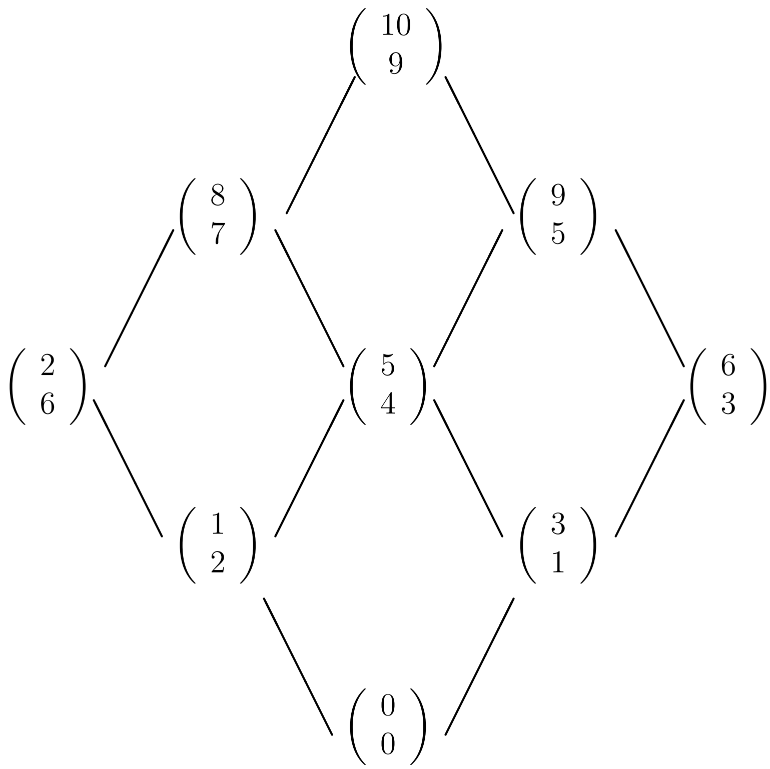


Figure 4.1: Example showing the Hasse Diagram

set in which every pair of elements has a least upper bound and a greatest lower bound. Ordering can be illustrated with the example shown in the Fig.4.2.2. The example shows the lattice on s set of two-dimensional vectors where,

$$\Omega_1 \preccurlyeq \Omega_2 \Leftrightarrow \begin{pmatrix} \omega_{11} \\ \omega_{21} \end{pmatrix} \preccurlyeq \begin{pmatrix} \omega_{12} \\ \omega_{22} \end{pmatrix} \Leftrightarrow (\omega_{11} \leq \omega_{12}) \wedge (\omega_{21} \leq \omega_{22}) \tag{4.6}$$

4.2.3 Service Attributes

We define η_j, γ_j, and τ_j as noise factor, reliability factor, and end-to-end propagation delay for the Link j, respectively. Noise factor is defined as ratio of input optical signal to noise ratio ($OSNR_{i/p} \equiv OSNR_i$) and output optical signal to noise ratio ($OSNR_{o/p} \equiv OSNR_{i+1}$), thus we have

$$\eta_j = \frac{OSNR_{i/p}}{OSNR_{o/p}}, \tag{4.7}$$

where $OSNR$ is defined as the ratio of the average signal power received at a node to the average ASE noise power at that node. The OSNR of the link and q-factor are related as,

$$q = \frac{2\sqrt{\frac{B_o}{B_e}}\mathrm{OSNR}}{1+\sqrt{1+4\mathrm{OSNR}}}, \tag{4.8}$$

where B_o and B_e are optical and electrical bandwidths, respectively [61]. Bit-error rate is related to the q-factor as follows,

$$BER = 2\ \mathrm{erfc}\left(\frac{q}{\sqrt{2}}\right). \tag{4.9}$$

From Eqs.(4.7,4.8,4.9), we see that q-factor is a function of η and hence BER can be quantified on the basis of η. If the BER is not to exceed certain threshold (say 10^{-9}), then there exists a corresponding noise factor threshold, say η_{th}. The signal is said to be lost due to high BER, if the end-to-end η is greater than η_{th} and thus cannot be recovered at the destination. The overall noise factor of a burst that has, traversed H hops is given by,

$$\eta_H = \prod_{k=1}^{H} \eta_k, \tag{4.10}$$

where in the above equation, the product is performed H times, starting from initial the link.

Reliability is the other factor considered for providing services. We define the reliability factor γ of a link, such that $0 \leq \gamma \leq 1$, indicates the percentage of reliability for a particular link. The reliability prediction method involves the calculation of down times contributed to all building blocks required to establish end-to-end network path [76]. In this work we assign a number generated from a uniformly distributed random variable $\sim U[0.6, 1]$ for each link in the network. The end-to-end reliability for the path traversing H-hops is calculated as,

$$\gamma_H = \prod_{k=1}^{H} \gamma_k. \tag{4.11}$$

As our aim is show how the service provisioning can be done, calculation of γ's is beyond the scope of this work.

The last attribute that we consider as an important service parameter for distributed applications is propagation delay. If τ is the propagation delay of a link, then end-to-end delay for H hops, is given by,

$$\tau_H = \sum_{k=1}^{H} \tau_k. \tag{4.12}$$

4.2.4 Path Information Vector

The service attributes can be used to maintain the local network information and by properly comparing these vectors, destinations can be chosen. Comparison of multi-dimension metrics can be done using the notion of lattices [81]. Lattices are explained using the ordering denoted by $\preccurlyeq$, which has the properties of reflexivity, antisymmetry, and transitivity. We denote the information vector at Link j as,

$$\Omega_j = \begin{pmatrix} \eta_j \\ \gamma_j \\ \tau_j \end{pmatrix}. \tag{4.13}$$

Definition 1. *Let Ω_j and Ω_k be the two information vectors for the links j and k, respectively. We define $\Omega_j \preccurlyeq \Omega_k$ and comparable if and only if*

$$(\eta_j \leq \eta_k) \wedge (\gamma_j \geq \gamma_k) \wedge (\tau_j \leq \tau_k). \tag{4.14}$$

Definition 2. *Ω_j and Ω_k are not comparable if and only if any one or two the inequalities in (4.14) are false. In other words if $(\eta_j > \eta_k)$ or/and $(\gamma_j < \gamma_k)$ or/and $(\tau_j > \tau_k)$. We denote them by $\Omega_j \parallel \Omega_k$.*

From the Eqs.(4.10,4.11,4.12), we see that the service attributes are either multiplicative (product) or additive (sum). The ordering condition in Eq. (4.14) is chosen such that, noise factor and propagation delay are minimum, and reliability is maximum. Each

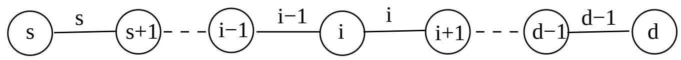


Figure 4.2: This explains the notation used in the Eq. (4.18).

information vector is a 3-tuple and hence it is a 3-dimensional vector space over real field $\mathbb{R}$, which is denoted by $\mathbb{R}^3$. The operation over multi-dimensional vectors is given by,

$$\circ : \Omega_j \in \mathbb{R}^3,\, \Omega_k \in \mathbb{R}^3 \rightarrow \Omega_j \circ \Omega_k \in \mathbb{R}^3. \tag{4.15}$$

where the operation $\circ$ on two vectors Ω_j and Ω_k is given by,

$$\Omega_j \circ \Omega_k = \begin{pmatrix} \eta_j \eta_k \\ \gamma_j \gamma_k \\ \tau_j + \tau_k \end{pmatrix} \tag{4.16}$$

Definition 3. *The path information vector from Source s to Destination d, is denoted by $\Omega_{\langle s,d \rangle}$ and is given by,*

$$\Omega_{\langle s,d \rangle} = \Omega_{\langle s,s+1 \rangle} \circ \ldots \circ \Omega_{\langle j,j+1 \rangle} \circ \ldots \Omega_{\langle d-1,d \rangle}, \tag{4.17}$$

where $\Omega_{\langle j,j+1 \rangle}$ is the information vector for the link between the nodes $\langle j, j+1 \rangle$ as shown in the Fig.4.2.

Thus using Eqs. (4.10,4.11,4.12) above equation becomes,

$$\Omega_{\langle s,d \rangle} = \begin{pmatrix} \displaystyle\prod_{k=s}^{d-1} \eta_k \\ \displaystyle\prod_{k=s}^{d-1} \gamma_k \\ \displaystyle\sum_{k=s}^{d-1} \tau_k \end{pmatrix}. \tag{4.18}$$

We use the notation $\Omega_{\langle s,d \rangle}$ for path information and Ω_j for the link information vector. However if the path consists of a single link, then from the Eq. (4.17), we get

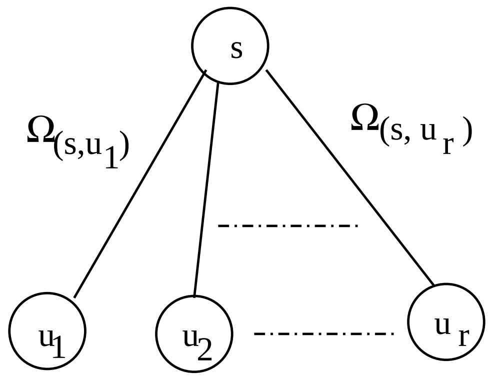


Figure 4.3: The child nodes or the next-hop nodes are the intermediate nodes towards the destination.

$\Omega_{\langle s,d\rangle} = \Omega_{\langle j,j+1\rangle} \equiv \Omega_j$.

Definition 4. *Consider a manycast request of the form* (s, κ, D_s). *Let* $D_s = \{d_1, d_2, \ldots, d_m\}$. *We define the next-hop (or the child nodes) corresponding to* s *as* $\{u_1, u_2, \ldots, u_r\}$ *as shown in Fig. 4.3, where* $r \in \mathbb{Z}^+$. *From Definition 3, there exists an information vector* $\Omega_{\langle s,u_i\rangle}$ $\forall$ $1 \leq i \leq r$.

If u_i is any intermediate node, then the overall information vector from source s to the destination is computed using Eq. (4.18) with upper limit replaced by $u_i - 1$

Definition 5. *We define differentiated service set as* $\Theta = \{\theta_1, \theta_2, \ldots, \theta_S\}$. *For each service* $\theta_p \in \Theta$ *there exists a threshold parameter (or constraint) that is defined as* $\top^{(\theta_p)}$ *and is given by*

$$\top^{(\theta_p)} = \begin{pmatrix} \eta_{max}^{(\theta_p)} \\ \gamma_{min}^{(\theta_p)} \\ \tau_{max}^{(\theta_p)} \end{pmatrix}. \tag{4.19}$$

For the successful establishment of QoS-based manycast session, the chosen destinations must satisfy the service requirements as defined in Eq. (4.19).

Theorem 4.2.1. *If* $\Omega_{\langle s,d\rangle} \preccurlyeq \top^{(\theta_p)}$, *then all the link information vectors* $\Omega_{\langle j,j+1\rangle}$*'s* $\forall$ $j \in$

$\{s, s+1, \ldots, d-1\}$, along the path from $\langle s, d\rangle$, are comparable to $\top^{(\theta_p)}$, i.e., $\Omega_{\langle j,j+1\rangle} \preccurlyeq \top^{(\theta_p)}$.

Proof. Given $\Omega_{\langle s,d\rangle} \preccurlyeq \top^{(\theta_p)}$, then by Definition 3 we have

$$\begin{aligned}
\Omega_{\langle s,s+1\rangle} \circ \ldots \circ \Omega_{\langle j,j+1\rangle} \circ \ldots \Omega_{\langle d-1,d\rangle} &\preccurlyeq \top^{(\theta_p)} \\
\Rightarrow \Omega_{\langle j,j+1\rangle} \circ \begin{pmatrix} \prod\limits_{\substack{k=s \\ k\neq j}}^{d} \eta_k \\ \prod\limits_{\substack{k=s \\ k\neq j}}^{d} \gamma_k \\ \sum\limits_{\substack{k=s \\ k\neq j}}^{d} \tau_k \end{pmatrix} &\preccurlyeq \top^{(\theta_p)}.
\end{aligned} \tag{4.20}$$

From the Eqs. (4.19) and (4.20) we get,

$$\eta_j \prod_{\substack{k=s \\ k\neq j}}^{d} \eta_k \leq \eta_{max}^{(\theta_p)}$$

$$\Rightarrow \eta_j \leq \frac{\eta_{max}^{(\theta_p)}}{\prod\limits_{\substack{k=s \\ k\neq j}}^{d} \eta_k} \leq \eta_{max}^{(\theta_p)}$$

Last inequality follows from the fact that $\eta_k > 1$ and hence we have $\eta_j \leq \eta_{max}^{(\theta_p)}$. Similarly for other two service attributes we have,

$$\gamma_j \geq \frac{\gamma_{max}^{(\theta_p)}}{\prod\limits_{\substack{k=s \\ k\neq j}}^{d} \gamma_k} \geq \gamma_{min}^{(\theta_p)}$$

$$\Rightarrow \gamma_j \geq \gamma_{min}^{(\theta_p)} \; (\because 0 \leq \gamma_k \leq 1)$$

$$\tau_j \leq \left(\tau_{max}^{(\theta_p)} - \sum_{\substack{k=s \\ k\neq j}}^{d} \tau_k \right) \leq \tau_{max}^{(\theta_p)}$$

$$\Rightarrow \tau_j \leq \tau_{max}^{(\theta_p)} \; (\because \tau_k \in \mathbb{R}^+)$$

Thus $\forall\ j \in \{s, s+1, \ldots, d-1\}$ we have $\Omega_{\langle j,j+1\rangle} \preccurlyeq \top^{(\theta_p)}$. Hence proved. $\square$

Definition 6. *A Path $\langle s, u_i \rangle$, where u_i can be child node or any intermediate node towards the destination is said to be feasible for service $\top^{(\theta_p)}$, if and only if $\Omega_{\langle s,u_i\rangle} \preccurlyeq \top^{(\theta_p)}$.*

In the case of u_i being a child node (or next-hop node), information vector is given by Eq. (4.13) and for intermediate node it is given by Eq. (4.18) with d replaced with u_i.

Theorem 4.2.2. *The path $\langle s, d\rangle$ is a feasible path if and only if $\top^{(\theta_p)}$ is the upper bound $\forall\ \Omega_{\langle s,j\rangle}$, where $j \in \{s+1, \ldots, d-1, d\}$.*

Proof. 1. *If Condition:* Let $\Omega_{\langle s,d\rangle}$ is the feasible path. Then from the Definition 6 we have,

$$\Omega_{\langle s,d\rangle} \preccurlyeq \top^{(\theta_p)}. \tag{4.21}$$

Consider $\Omega_{\langle s,s+1\rangle}$ and $\Omega_{\langle s,s+2\rangle}$, then we have

$$\begin{pmatrix} \eta_s \\ \gamma_s \\ \tau_s \end{pmatrix} \preccurlyeq \begin{pmatrix} \eta_s\eta_{s+1} \\ \gamma_s\gamma_{s+1} \\ \tau_s + \tau_{s+1} \end{pmatrix}$$

Above inequality follows from the fact that $\eta > 1$, $0 < \gamma < 1$, and $\tau \in \mathbb{R}^+$. Thus we have

$$\Omega_{\langle s,s+1\rangle} \preccurlyeq \Omega_{\langle s,s+2\rangle} \ldots \preccurlyeq \Omega_{\langle s,d\rangle} \tag{4.22}$$

From the Eqs. (4.21) and (4.22) we see that $\top^{(\theta_p)}$ is the upper bound for $\Omega_{\langle s,j\rangle}$, where $j \in \{s+1, \ldots, d-1, d\}$.

2. *Only if:* Given $\top^{(\theta_p)}$ is the upper bound for $\Omega_{\langle s,j\rangle}$, where $j \in \{s+1,\ldots,d-1,d\}$, then we have

$$\Omega_{\langle s,d\rangle} \preccurlyeq \top^{(\theta_p)}$$

and hence the path $\langle s,d\rangle$ is feasible.

□

Theorem 4.2.3. *If $\Omega_{\langle s,u_i\rangle}$, the information vector corresponding to the path from $\langle s,u_i\rangle$ is not feasible then all the destinations using u_i as the intermediate node does not qualify as the members in quorum pool.*

Proof. From the Definition 6 we see that, if $\Omega_{\langle s,u_i\rangle}$ is not feasible path then either $\top^{(\theta_p)} \preccurlyeq \Omega_{\langle s,u_i\rangle}$ or $\Omega_{\langle s,u_i\rangle} \parallel \top^{(\theta_p)}$. If the former condition were true then,

$$\begin{pmatrix} \eta_{max}^{(\theta_p)} \\ \gamma_{min}^{(\theta_p)} \\ \tau_{max}^{(\theta_p)} \end{pmatrix} \preccurlyeq \begin{pmatrix} \prod_{k=s}^{u_i-1} \eta_k \\ \prod_{k=s}^{u_i-1} \gamma_k \\ \sum_{k=s}^{u_i-1} \tau_k \end{pmatrix} \tag{4.23}$$

From the Eq. (4.23) we see that as all the constraints for the service θ_p are not met and thus the destinations using u_i as intermediate node disqualify to be in the quorum pool. In similar way at least one of the constraint is not meet, then we have $\Omega_{\langle s,u_i\rangle} \parallel \top^{(\theta_p)}$ and thus all the destinations using u_i as intermediate node does not satisfy the service requirements of θ_p. We define this set as NRD, non-reachable destinations due to insufficient QoS. □

Lemma 4.2.1. *A manycast request (s, D_s, κ) is said to lost (or blocked) due to insufficient QoS, if and only if cardinality of the set that does not satisfy QoS is greater than κ.*

Proof. We require any of the κ members in the group for successful completion of the session. But from the Theorem 4.2.3 if the number of destinations that cannot be reached through all u_i's is greater that κ, then $|NRD| > \kappa$. As $\kappa \geq \lceil m/2 \rceil$ number of remaining destinations $(m - |NRD|) < \kappa$ and hence the request is blocked. □

4.3 Multi Constrained Manycast Problem (MCMP)

In this section we explain the proposed Multi-Constrained Manycast algorithms with the help of an example. We propose two algorithms, MCM-Shortest Path Tree (SPT) and MCM- Dynamic Membership (DM) for evaluating the performance of the manycasting with quality of service (QoS) constraints. These proposed algorithms are distributed wherein, each node individually maintains the network state information and executes the algorithm. Algorithms implemented in the centralized way, may fail due to a single failure and resulting in poor performance. Our proposed algorithms have the following functionality:

1. Handling multiple constraints with help of link state information available locally.
2. Service differentiated provisioning of manycast sessions.
3. Finding the best possible destinations in terms of service requirements for the manycast sessions.

We first discuss the steps to implement the distributed version of the shortest path tree (SPT) which is given in [73], [60], [75].

- *Step 1*: Find the shortest path from Source s to all the destinations in D_s. Let $D_s = \{d_1, d_2, \ldots d_{|D_s|=m}\}$ and minimum-hop distance from s to d_i, where $1 \leq i \leq m$ is $\mathbb{H}^{(s)} = \{h_1, h_2, \ldots, h_m\}$.
- *Step 2*: All the destinations in D_s are sorted in the non-decreasing order according to the shortest distance from Source s to the destinations. Let D'_s be the new set in this order given by $\{d'_1, d'_2, \ldots, d'_m\}$.
- *Step 3*: Select the first κ destinations from D'_s.

Step 1 is implemented using unicast routing table, with the time-complexity of this step being $O(n^2)$ for a network with n nodes. *Step 2* sorts the destinations in constant time $O(1)$. *Step 3* we select the first κ of them from D'_s, with $O(n)$. Once the first κ of them are selected, the burst header packet (BHP) is sent to corresponding child nodes

$\{u_1, u_2, \ldots, u_r\}$ or the next-hop nodes where $1 \leq r \leq \kappa$ from the Source s in the manycast request (s, D_s, κ). Construction of routing tree starts from the Source s. Once the BHP is received at the corresponding child nodes, the data burst is scheduled along the channel. OBS is based on one-way reservation protocols, such as Just-Enough-Time (JET) and Tell-And-Go (TAG) [83], [6], in which data burst is scheduled after certain *offset time* without waiting for the acknowledgment. In this work we use JET signaling protocol for the manycasting. Upon receiving the data burst at the corresponding child nodes based on the QoS constraints, the manycast request is updated as $(u_i, D_{u_i}, \kappa_{u_i})$ where u_i is a child node for Source s in the previous iteration, D_{u_i} are all the destinations in D_s that can be reached through u_i, and clearly $D_{u_i} \subseteq D_s$. κ_{u_i} is updated accordingly if any of the u_i is a destination and we make sure that $\sum_{i=1}^{p} \kappa_{u_i} \leq \kappa$. With u_i as the Source node all the above three steps are performed and this iteration proceeds until minimum κ of them are reached. We thus see that SPT works in the distributed way and each node executes the algorithms based on the local network state information.

In the case of Dynamic Membership (DM), the above mentioned three steps differ slightly. In DM instead of discarding the rest of $m - \kappa$ destinations, we keep them and are used if any of the first κ are blocked due to the contention or in-sufficient QoS on the link. Former blocking is referred to as *contention blocking* and the later is referred to as *QoS blocking*. Detailed description of the algorithms *MCM-SPT* and *MCM-DM* is explained the following sections. Before we begin with the description of the routing algorithms, it is important to know about the OBS Control plane which can implement these functionality effectively.

4.3.1 OBS Control Plane

In the OBS layered architecture we find two important planes: *data plane* and *control plane*. *Control plane* allows scheduling and reservation protocols to be performed in a domain (electrical) different from data plane (optical). Detailed description of OBS data and control planes can be found in [84], [25]. Examples of control packets are burst header packet (BHP), network control packet (NCP), and burst confirmation packet (BCP). In

Table 4.1: Control Packets Frame Fields

BHP Frame Field	Description
Id	Burst Identification number used for sequencing
Source (u)	Initial or starting node of the burst
Quorum members (D_u)	These are the probable destinations to which burst can be reached.
κ_u	Number of members in manycast session
$\top^{(\theta_p)}$	Threshold information vector for service θ_p.
$\Omega_{\langle u-1,u\rangle}$	Link information vector corresponding to the link between $\langle u-1, u\rangle$.
Ingress Channel	Wavelength used for the data burst
Duration	Duration of the data burst in seconds
Offset	Time offset between the control packet the data packet

Manycast ID	Burst ID	Source Node (u)	Destination Set (D_u)	Qurom Pool (κ_u)	$\top^{(\theta_p)}$	$\Omega_{(u-1,u)}$

Figure 4.4: Burst header packet fields considered for the analysis.

this work we use BHPs as the control packets and we propose the new BHP field which provides information about the QoS. In previous works [75, 73] the BHP was modified to accommodate q-factor (i.e., BER) and burst were scheduled based on the BER threshold. Table 4.1 lists possible fields associated with QoS based scheduling of bursts.

We consider six fields of Table 4.1 in BHP as shown in the Fig. 4.4.

4.3.2 Multi-Constraint Manycast-SPT (MCM-SPT)

In this section we explain MCMP with help of shortest-path tree. The pseudo-code for this is given in Algorithm 4. When the new burst arrives in the network, it is assigned a unique burst ID, id. A BHP is created to this burst, with all the fields shown in the Fig. 4.4, where $u = s$, destination set D_s, quorum pool κ_s, threshold parameters $\top^{(\theta_p)}$ for the service θ_p and the initial information vector $\Omega_{initial} = [1, 1, 0]^T$. This is indicated as Line 1, in the algorithm. Destinations in D_s are sorted along the shortest path using

$SORT.SP[D_s]$. Next-hop nodes from s to $d'_j \in D_s$ are calculated and added to the set N. Loop in Lines 7-10 selects the first κ_u destinations from D'_s and next-hop nodes are added to N for each destination d'_j using Line 9. Every link to the next-hop node is checked for contention using Line 12. If the link is found free, the path information vector is calculated using path algebra explained in Section 4.2.4. If $\Omega_{\langle s,n_j\rangle} \preccurlyeq \top^{(\theta_p)}$, then the link $\langle s, n_j\rangle$, qualifies the QoS threshold attributes for the service θ_p. All the destinations which use n_j as the intermediate node is given by the set $DEST[n_j]$ and hence the new destination set is given by D_{n_j} as shown in Line 15. BHP at node n_j is updated with the new values as given by Line 17. One must note here that the burst ID remains same, until it reaches the κ_s destinations. If condition given in Line 14 were false, i.e., $\Omega_{\langle u,n_j\rangle} \parallel \top^{(\theta_p)}$ or $\top^{(\theta_p)} \preccurlyeq \Omega_{\langle u,n_j\rangle}$, then according to Theorem 4.2.3 and Lemma 4.2.1 we have manycast request to be not meet, as the minimum number of members in the pool are less then the required κ_u. We refer to this blocking as *QoS Blocking*. Burst is removed from the network due to in-sufficient QoS parameters. *Contention blocking* occurs when an arriving burst finds the channel occupied. Burst is removed from the network if the condition in Line 12 and Line 15 are not met. This algorithm repeats until the κ_s destinations are covered for a burst.

Consider a manycast request of the form $(s = 2, D_s = \{6, 7, 11\}, \kappa = 2)$, this can be represented by 3/2. Fig. 4.5 shows the shortest-path tree for the given manycast request of the NSF network in Fig. 3.5, with links shown in dotted lines. Let the service threshold be $\top^{(\theta_p)} = [\eta_{th} = 6, \gamma_{th} = 0.6, \tau_{th} = 20 \text{ ms}]^T$. In order to guarantee QoS for the service θ_p, our aim is to identify destinations that have overall noise factor $\eta \leq 6$, reliability $\gamma \geq 0.6$ and propagation delay $\tau \leq 20$ ms. Burst enters network at Source $s = 2$, burst header fields in Fig. 4.4 are updated with the values and path information vector is initialized to $\Omega_{initial} = [1, 1, 0]^T$ as given in the step-1 of the Algorithm 4. Using $SORT.SP$ the destination set D_2 is sorted in the non-decreasing order of the distance and the new set is given by $D'_2 = \{7, 6, 11\}$. In MCM-SPT we select the first $\kappa_2 = 2$ from D'_2, we have $\{7, 6\}$. Next hop nodes for these destinations are given by $N = \{4\}$. Assuming the Link $\langle 2, 4\rangle$ is free (no contention), we compute the path information vector $\Omega_{\langle 2,4\rangle} \leftarrow \Omega_{\langle 2,4\rangle} \circ \Omega_{initial}$

given in Line-13. Computation of noise factor is done using the parameter values given in [73], [75]. We assume the input power at Node 2 as $P(2) = 1$ mW, with ASE noise as $P_{ase}(2) = 0.0042$ mW. Taking ratio of these two powers we get OSNR at Node 2 as $OSNR(2) = 238$ and based on all the losses mentioned in [73], and [75] we have the OSNR at Node 4 as $OSNR(4) = 56.53$. Using Eq. (4.7) we get $\eta_{\langle 2,4\rangle} = 4.21$. Propagation delay of the burst along the link (ms) is calculated as distance (km) to the velocity of light (250 km/ms). The information vector is given by,

$$\begin{aligned}\Omega_{\langle 2,4\rangle} &= \Omega_{\langle 2,4\rangle} \circ \Omega_{initial}, \\ &= \begin{pmatrix} 4.21 \\ 0.98 \\ 4.0 \end{pmatrix} \circ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \\ &= \begin{pmatrix} 4.21 \\ 0.98 \\ 4.0 \end{pmatrix}.\end{aligned}$$

We thus see that $\Omega_{\langle 2,4\rangle}$ is with in the threshold requirement of the service θ_p, satisfying condition in Line-14. A new BHP is created at node 4, with same burst ID as, $(id, 4, D_4 = \{6,7\}, \kappa_4 = 2, \top^{(\theta_p)}, \Omega_{\langle 2,4\rangle})$. Algorithm exits in Line-18. The same algorithm is repeated, however the Step-1 is skipped as this is the old burst. Lines 2-4 are used when one of the intermediate node is a destination. Assuming the Link $\langle 4,5\rangle$ is free we have,

$$\begin{aligned}\Omega_{\langle 4,5\rangle} &= \Omega_{\langle 2,4\rangle} \circ \Omega_{\langle 4,5\rangle}, \\ &= \begin{pmatrix} 4.21 \\ 0.98 \\ 4.0 \end{pmatrix} \circ \begin{pmatrix} 1.0702 \\ 0.99 \\ 2.4 \end{pmatrix}, \\ &= \begin{pmatrix} 4.5056 \\ 0.972 \\ 6.4 \end{pmatrix}.\end{aligned}$$

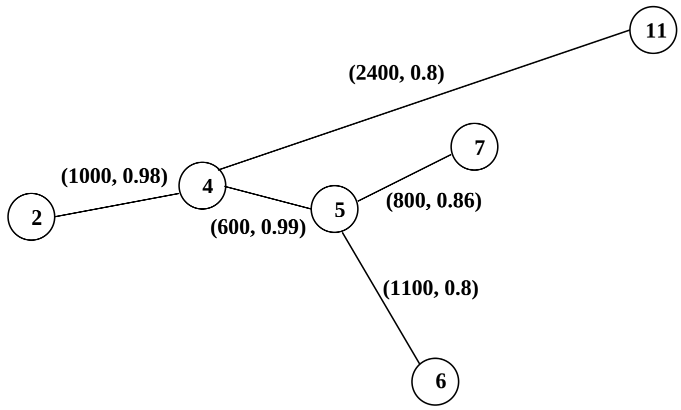


Figure 4.5: Example used for explaining proposed algorithms. The weights indicate distance and reliability factor of the links. Other nodes and links of the NSF network are not shown for clarity. Each link consists of in-line amplifiers placed 70 km apart.

Assuming links from Node 5 to Node 7 and Node 8 to be free, we have $\Omega_{\langle 5,7\rangle} = [5.44, 0.83, 10.8]^T$ and $\Omega_{\langle 5,6\rangle} = [4.839, 0.776, 9.6]^T$. We thus see that the QoS threshold conditions for the service θ_p are met. The manycast session is successful for a given service. Same manycast request can be blocked for different service threshold conditions like $\top^{(\theta_q)} = [5, 0.8, 10]^T$. If at least one of the destination is not reachable through the next-hop node, due to contention or insufficient QoS, entire manycast request is said to be blocked. This is executed by the Lines 19 and Line 23.

4.3.3 Multi-Constraint Manycast-DM

MCM-DM is given in Algorithm-5. It contains two procedures (1) for calculating the QoS parameters and updating BHP with the new values, defined as $Procedure.QoS()$ and (2) for calculating the number of destinations that can be reached from the next-hop Node n is greater than κ_u, defined as $Procedure.Block()$. Details of these procedures are given in Appendix B.2 and Appendix B.1 respectively. Instead of discarding $|D_u| - \kappa_u$ destinations as in MCM-SPT, we keep these destinations as secondary destinations and use them if any of the first κ_u are blocked. Intuitively one can understand that request blocking could be reduced in the case of MCM-DM, as members in the quorum pool are added or removed dynamically. While adding the destinations into the quorum pool the burst traversal can be along a longer path, deteriorating QoS parameters. Resulting QoS

Algorithm 4 Multi-Constraint Manycast-Shortest Path Tree (MCM-SPT)

Input: The manycast request $(id, u, D_u, \kappa_u, \top^{(\theta_p)}, \Omega_{\langle u-1,u\rangle})$ arrives at the Source node u with a candidate destination set D_u, along with the κ intended.

Output: Manycast request to the next hop (or child) node after satisfying QoS parameters for the service θ_p.

Initialization: When the burst first enters the network with the request $(s, D_s, \kappa_s, \top^{(\theta_p)})$, we tag the request with a burst ID and $\Omega_{initial}$, where $\Omega_{initial} = [1, 1, 0]^T$. We therefore have the request as $(id, s, D_s, \kappa_s, \top^{(\theta_p)}, \Omega_{initial})$.

if $u \in D_u$ **then**
  $D_u \leftarrow D_u \backslash \{u\}$;
  $\kappa_u \leftarrow \kappa_u - 1$;
else
  $D'_u \leftarrow SORT.SP[D_u]$;
  for $j \leftarrow 1$ to κ_u **do**
    $n_j \leftarrow NEXT.HOP.NODE.SP[u, d'_j]$;
    $N = N \cup \{n_j\}$;
  end for
  for $j \leftarrow 1$ to $|N|$ **do**
    if $LINK\langle u, n_j\rangle = FREE$ **then**
      $\Omega_{\langle u,n_j\rangle} \leftarrow \Omega_{\langle u-1,u\rangle} \circ \Omega_{\langle u,n_j\rangle}$;
      /*`QoS parameters computed using the path algebra`*/
      if $\Omega_{\langle u,n_j\rangle} \preccurlyeq \top^{(\theta_p)}$ **then**
        $D_{n_j} \leftarrow DEST[n_j]$;
        $UPDATE.BHP[n_j]$;
        $(id, n_j, D_{n_j}, \kappa_{n_j}, \top^{(\theta_p)}, \Omega_{\langle u,n_j\rangle})$;
      else
        $DELETE.BURST[id]$;
        `exit;`
        /*`QoS Blocking`*/
      end if
    else
      $DELETE.BURST[id]$;
      `exit;`
      /*`Contention Blocking`*/
    end if
  end for
end if

blocking could be high when compared to MCM-SPT. This algorithm is explained using the same example $(2, \{6, 7, 11\}, 2)$ for which the manycast tree is shown in Fig.4.5. Let the threshold conditions for the service θ_p be $[\eta_{th} = 6, \gamma_{th} = 0.6, \tau_{th} = 20 \text{ ms}]^T$. BHP is created for this burst as the part of initialization as given in the Line 1 of the algorithm.

Destinations are sorted along the shortest-path and $D'_2 = \{7, 6, 11\}$. The next-hop nodes is given by the Lines 7-10. In this case we have $N = \{4\}$. At Node 4, the destination node set is $D_4 = \{6, 7\}$. As we select only the first κ_2 of them, destination Node 11 is left out. All destinations $|D_u - \kappa_u|$ are added in round-robin to the destination set at the child nodes (or next-hop node). Here as there is only one child node, we have $D_4 = \{6, 7\} \cup \{11\}$.[1] Loop in Lines 13-22 selects the primary destinations, in this case $\{6, 7\}$. Next-hop Node for node 6 is Node 4, assuming the link $\langle 2, 4\rangle$ to be free and as the condition in Lines 15-17 is meet (since $\kappa_4 = 0$) QoS parameters are calculated using $Procedure.QoS$. BHP at Node 4 is updated as $(id, 4, D_4 = \{6, 7, 11\}, k_4 = 1, \top^{(\theta_p)}, \Omega_{\langle 2,4\rangle})$, with $\Omega_{\langle 2,4\rangle} = [4.21, 0.98, 4]^T$. For the next iteration i.e., for Destination 7, the next-hop node being same, we have $k_4 = 2$, which is updated in the BHP at Node 4. Finally burst is scheduled and the BHP at node 4 is now $(id, 4, D_4 = \{6, 7, 11\}, k_4 = 2, \top^{(\theta_p)}, [4.21, 0.98, 4]^T)$. In the similar way assuming Link $\langle 4, 5\rangle$ to be free the BHP at node 5 is $(id, 5, D_5 = \{6, 7, 11\}, k_4 = 2, \top^{(\theta_p)}, [4.5056, 0.972, 6.4]^T)$. If the Link $\langle 5, 7\rangle$ is not free, $k = 1$ and the BHP at Node 6 becomes $(id, 6, D_6 = \{6\}, 1, \top^{(\theta_p)}, [4.839, 0.776, 9.6]^T)$. As $k = 1$ the loop in Line-24 is executed. First condition in Line-26 ($k \leq |D_5| - 2$), ensures that number of blocked destinations due to contention in not greater than the secondary destinations. Next-hop node to $d'_{2+1} = 11$ from Node $u = 5$ is Node 4. Thus the burst node Node 5 is updated as $(id, 5, D_5 = 11, 1, \top^{(\theta_p)}, \Omega_{\langle 4,5\rangle})$. Burst at Node 5 is routed to Node 11 along Node 4. We

[1] The procedure for round-robin is described in Appendix B.3.

have

$$\begin{aligned}\Omega_{\langle 5,4\rangle} &= \Omega_{\langle 4,5\rangle} \circ \Omega_{\langle 5,4\rangle} \\ &= \begin{pmatrix} 4.5056 \\ 0.972 \\ 6.4 \end{pmatrix} \circ \begin{pmatrix} 1.0104 \\ 0.99 \\ 2.4 \end{pmatrix} = \begin{pmatrix} 4.55 \\ 0.96228 \\ 8.8 \end{pmatrix},\end{aligned}$$

and the BHP at Node 4 is $(id, 4, \{11\}, 1, \top^{(\theta_p)}, [4.55, 0.96228, 8.8]^T)$. Finally if the link $\langle 4, 11\rangle$ is free, BHP is updated with $(id, 11, \{11\}, 1, \top^{(\theta_p)}, [5, 0.76, 18.4]^T)$. We thus see that the manycast request which was earlier blocked in MCM-SPT if one of the destination is blocked is now satisfied. As MCM-DM adds destinations on the longer path, it is necessary to see whether the route to the destination is with in the QoS threshold requirements of the service.

4.4 Performance Evaluation

In this section we present our simulation results. We consider *average request blocking* as performance metric. We define average request blocking ratio as given by [60]. Let f be the total number of manycast requests used in the simulation. Consider a manycast request (s, D_s^f, κ). Let $\mathbb{D}$ be the set of destinations which actually receive the data. Then *average request blocking* is given by,

$$B_{avg} = \sum_f \left[1.0 - \min(|\mathbb{D}|, \kappa)/\kappa\right]/f. \tag{4.24}$$

NSF network shown in the Fig. 4.6 is used for our simulation. All the links are bi-directional and have same transmission rate of 10 Gb/s. Burst arrivals follow Poisson process with an arrival rate of λ bursts per second. The length of the burst is exponentially distributed with expected service time of $1/\mu$ seconds.

Algorithm 5 Multi-Constraint Manycast-Dynamic Membership (MCM-DM)

Input: The manycast request $(id, u, D_u, \kappa_u, \top^{(\theta_p)}, \Omega_{\langle u-1,u\rangle})$ arrives at the Source node u with a candidate destination set D_u, along with the κ intended.

Output: Manycast request to the next hop (or child) node after satisfying QoS parameters for the service θ_p.

Initialization: $(id, s, D_s, \kappa_s, \top^{(\theta_p)}, \Omega_{initial})$.
if $u \in D_u$ **then**
    $D_u \leftarrow D_u \backslash \{u\}$;
    $\kappa_u \leftarrow \kappa_u - 1$;
else
    $D'_u \leftarrow SORT.SP[D_u]$;
    for $j \leftarrow 1$ to κ_u **do**
        $n_j \leftarrow NEXT.HOP.NODE.SP[u, d'_j]$;
        $N = N \cup \{n_j\}$;
    end for
    $ROUND.ROBIN.DEST[|D_u| - \kappa_u]$;
    $k \leftarrow 0$;
    for $j \leftarrow 1$ to κ_u **do**
        if $LINK\langle u, n_j\rangle = FREE$ **then**
            while $\sum_{k=1}^{|N|} \kappa_{n_k} < \kappa_u$ **do**
                $\kappa_{n_j} \leftarrow \kappa_{n_j} + 1$;
            end while
            Procedure.QoS( );
        else
            $k \leftarrow k + 1$;
        end if
    end for
end if
for $i \leftarrow 1$ to k **do**
    $n_k \leftarrow NEXT.HOP.NODE.SP[u, d'_{k+\kappa_u}]$;
    if $(k \leq |D_u| - \kappa_u) \& (LINK\langle u, n_k\rangle = FREE)$ **then**
        Procedure.QoS( );
    else
        Procedure.Block( );
    end if
end for

4.4.1 Assumptions

1. Only one wavelength is considered for analysis. Hence the dependency of q-factor on the wavelength is ignored.

2. Wavelength converters are not used in the network.

3. Calculation of noise factor is based on, losses due to attenuation, mux/demux, tap and split loss. Only amplified spontaneous emission (ASE) noise can be considered for OSNR. Shot noise and beat noise are ignored.

4. Effects of offset time are ignored.

5. In line amplifiers along the links are placed, with spacing of 70 km between the amplifiers.

6. There are no optical buffers or wavelength converters in the network.

7. Reliability factor is same along both directions of the fiber.

As already mentioned, we have intended destinations, i.e., quorum pool to be majority of the group $\kappa \geq \lceil m/2 \rceil$. Candidate destination group (or quorum group) can be small, medium or large. Three typical configurations 3/2, 7/4 and 11/6 are considered for simulations. First we present simulation results for 7/4 manycast configuration. We differentiate among service requirements, i.e., different services put different constraints. Differentiated services considered for simulation are $\top^{(\theta_1)} = [5.7, 0.6, 20]^T$, $\top^{(\theta_2)} = [5.7, 0.6, 10]^T$, $\top^{(\theta_3)} = [4.25, 0.9, 10]^T$ and $\top^{(\theta_4)} = [4.25, 0.8, 10]^T$. We consider $\top^{(\theta_2)}$ as the real-time service, since it has more stringent delay requirement. Service $\top^{(\theta_1)}$ can be for data service as it has less relaxed delay requirements. Other two services have high threshold requirements.

Figure 4.7 shows the performance of the MCMP-SPT for different set of services. More the requirements of the service, more is the blocking. As MCM-SPT uses shortest-path routing, one can expect to have a lower QoS blocking, but however due to the random contention along the links, if any one of the destination is not reachable, entire manycast

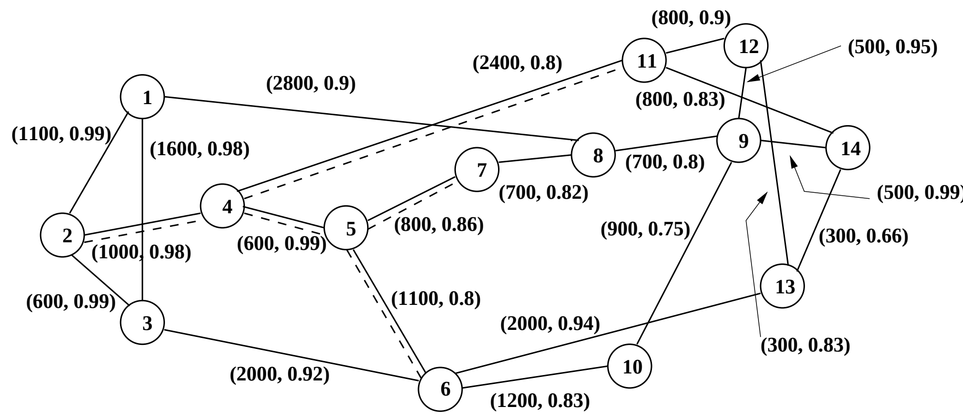


Figure 4.6: NSF network with 14 nodes and 21 bi-directional links. The weights represent distance in km and the corresponding reliability factor of the links respectively.

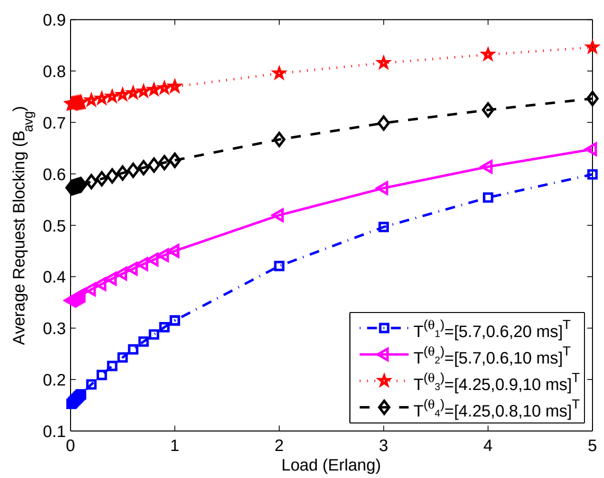


Figure 4.7: Blocking Probability performance of SPT for different service thresholds.

request would be blocked. On the contrary, MCM-DM adds or removes destinations based on the contention in the network. However destinations which are added to the quorum pool can be at a longer distance than the destination which is not reachable. As the result, QoS of this destination can be decreased. In spite of decrease in values, if the path-information vector is with in the threshold condition of the service, the request can be satisfied. Fig. 4.8 shows average request blocking for MCM-DM under different service thresholds. At high loads, most of the blocking would be contention blocking and hence the effect of QoS will not be understood much. As our aim is show the effects of QoS, all the results are simulated under medium network load conditions.

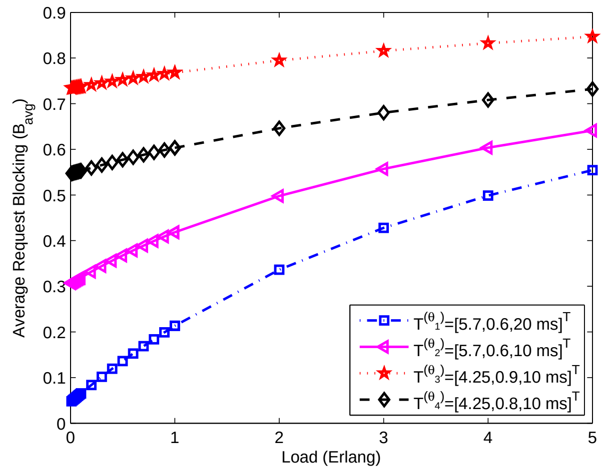


Figure 4.8: Blocking Probability performance of DM for different service thresholds.

The Fig.4.9 shows the comparison of average request blocking for the two proposed algorithms. We see that for the data service requirement $\top^{(\theta_1)}$, there is the significant reduction in the request blocking for the network loads between $(0, 1]$. As the network load increase the performance of two algorithms converges and for loads greater than 5, the request blocking is same. Under real-time service requirements like service θ_2, we observe from Fig. 4.10 that performance of MCMP-DM is reduced and difference in the request loss between MCM-SPT and MCM-DM has been decreased. This can be accounted to the fact that, while adding secondary destinations, the longer paths have to be traversed and hence the delay increases, causing a destination to disqualify based on the delay constraint. We can thus observe that MCM-DM can be chosen for data service application, where there is not specific upper bound on the propagation delay of the burst. Data service based distributed applications like SAN and CDN have more priority on η and γ rather than τ. We have also simulated the performance of the algorithms for more stringent QoS requirements, like service θ_3. We observe here that both the algorithms for this service requirements behave same. By relaxing the constraint on reliability, we observe a significant decrease in the request loss in Fig. 4.12.

Same set of services were simulated for the other two configurations, i.e., 3/2 and 11/6 at low loads. At higher loads contention in the network could be large and hence the

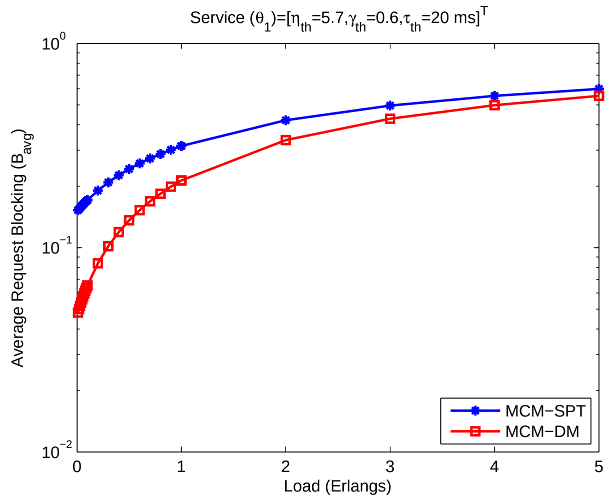


Figure 4.9: Blocking Probability performance of MCM-SPT and MCM-DM for service threshold of $\top^{(\theta_1)} = [\eta_{th} = 5.7, \gamma_{th} = 0.6, \tau_{th} = 20 \text{ ms}]^T$.

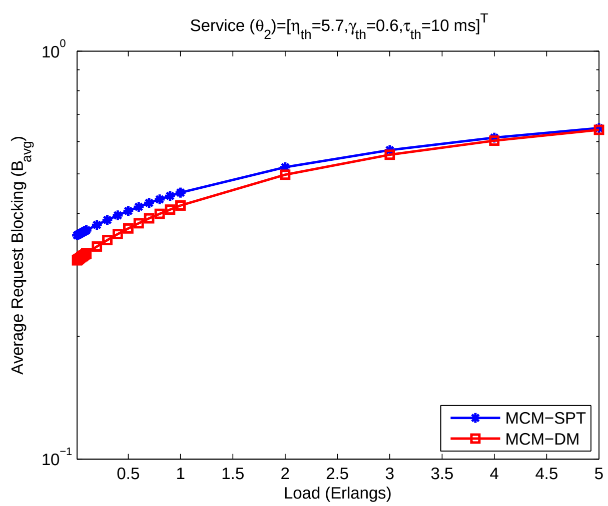


Figure 4.10: Blocking Probability performance of MCM-SPT and MCM-DM for service threshold of $\top^{(\theta_2)} = [\eta_{th} = 5.7, \gamma_{th} = 0.6, \tau_{th} = 10 \text{ ms}]^T$.

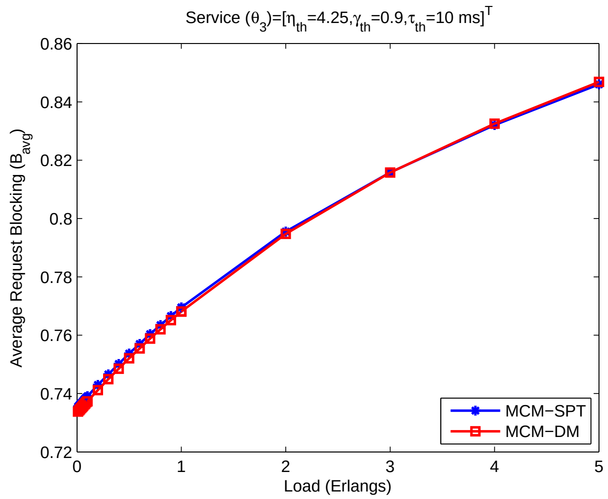


Figure 4.11: Blocking Probability performance of MCM-SPT and MCM-DM for service threshold of $\top^{(\theta_3)} = [\eta_{th} = 4.25, \gamma_{th} = 0.9, \tau_{th} = 10 \text{ ms}]^T$.

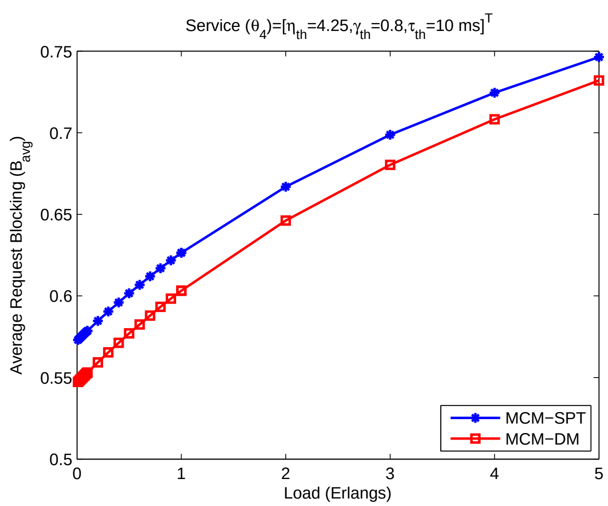


Figure 4.12: Blocking Probability performance of MCM-SPT and MCM-DM for service threshold of $\top^{(\theta_4)} = [\eta_{th} = 4.25, \gamma_{th} = 0.8, \tau_{th} = 10 \text{ ms}]^T$.

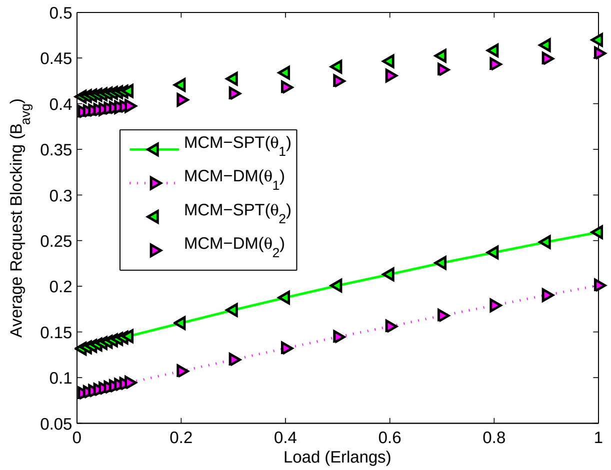


Figure 4.13: Blocking Probability performance of MCM-SPT and MCM-DM for 3/2 manycast configuration for services θ_1 and θ_2.

effect of the QoS may not be significant. Hence we restricted our simulation study only at low loads. Fig. 4.13 shows the performance of 3/2 manycast configuration for services θ_1 and θ_2. As we know that θ_1 has more relaxed threshold parameters, hence in case of θ_1 we can improve the blocking marginally using MCM-DM. But in the case of θ_2, where the delay requirement is only 10 ms, we observe that both algorithms offer same performance in terms of request blocking.

We observe an interesting result in Fig.4.14. In the case of θ_3 for 3/2 manycast configuration, we find MCM-SPT to offer lower request blocking than MCM-DM. This is because service θ_3 has high QoS requirement (real-time service). In the view to decrease the request blocking MCM-DM schedules the burst on the longer paths, which causes service attributes to exceed beyond the threshold requirements. Once again we see that MCM-DM can be only used when there are much relaxed QoS parameters (data services).

Finally we also simulate 11/6 manycast configuration for the four services. Fig. 4.15 and Fig. 4.16 show much similar performance to that of 7/4 and 3/2.

We simulate the impact of each service attribute on the network for 7/4 manycasting. In the other words, only one service attribute (i.e., either delay, BER or reliability) is considered, with others two service attributes neglected. Then the problem becomes a single constrained problem. Let service threshold for the delay constrained (DC) $(\top^{(\theta_d)})$

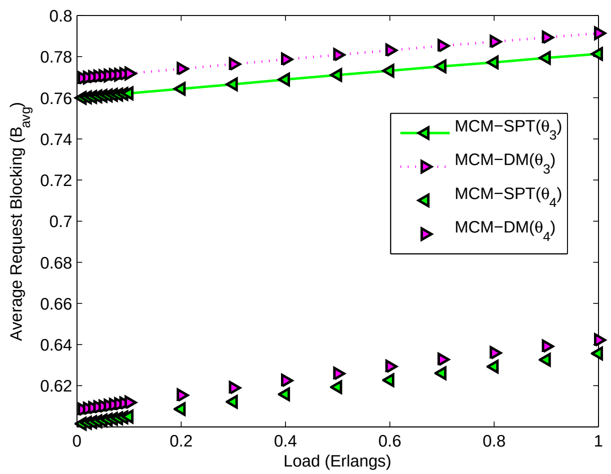


Figure 4.14: Blocking Probability performance of MCM-SPT and MCM-DM for 3/2 manycast configuration for services θ_3 and θ_4.

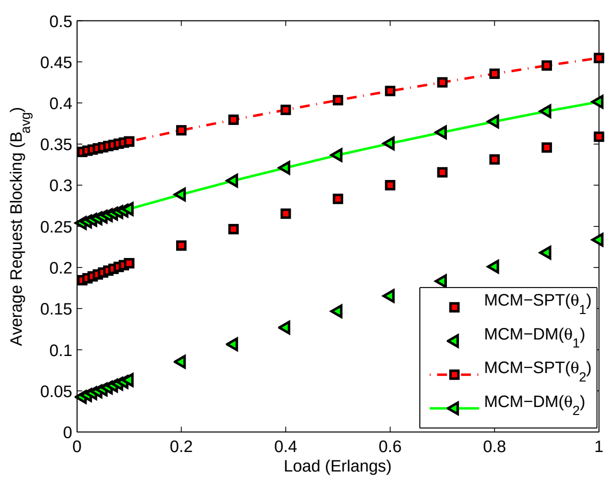


Figure 4.15: Blocking Probability performance of MCM-SPT and MCM-DM for 11/6 manycast configuration for services θ_1 and θ_2.

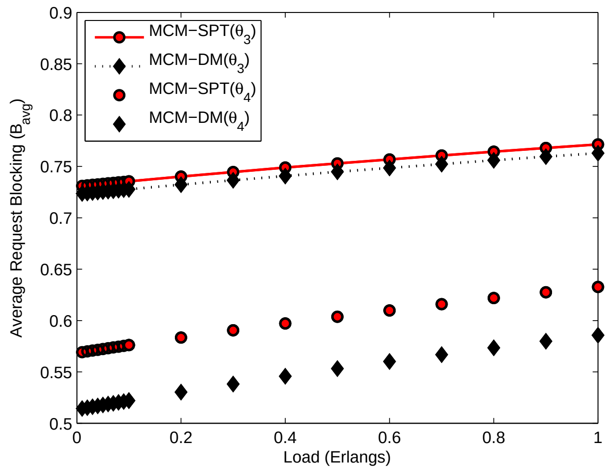


Figure 4.16: Blocking Probability performance of MCM-SPT and MCM-DM for 11/6 manycast configuration for services θ_3 and θ_4.

be $[\infty, 0, 10]^T$. In this case we removed the threshold requirements on noise factor and reliability by keeping $\eta_{max}^{(\theta_d)} = \infty$ and $\gamma_{min}^{(\theta_d)} = 0$. Similarly we consider BER constrained (BC) and Reliability constrained (RC) with service thresholds given by $(\top^{(\theta_b)})$ be $[4.25, 0, \infty]^T$ and θ_r be $[\infty, 0.8, \infty]^T$ respectively. Fig. 4.17 shows the performance of 7/4 manycast configuration for MCM-SPT and MCM-DM. We observe a significant decrease in the request blocking for MCM-DM compared to MCM-SPT in the case of BER constrained (BC). Services θ_b and θ_r performance is almost similar for both MCM-SPT and MCM-DM.

4.5 Summary

In this chapter we have evaluated the performance of manycasting over optical networks for providing QoS. Using lattice theory we were able to calculate multiplicative and additive QoS attributes. By using a distributed scheduling bursts are routed to the destinations based on the contention and QoS conditions. Two algorithms were proposed in a view to decrease the average request loss for manycasting. Performance of these algorithms has been studied under differentiated services. From the simulation results it has been observed that MCM-DM performs better in the case of data services and MCM-SPT for real-time services. This work proposes the necessity of providing QoS to

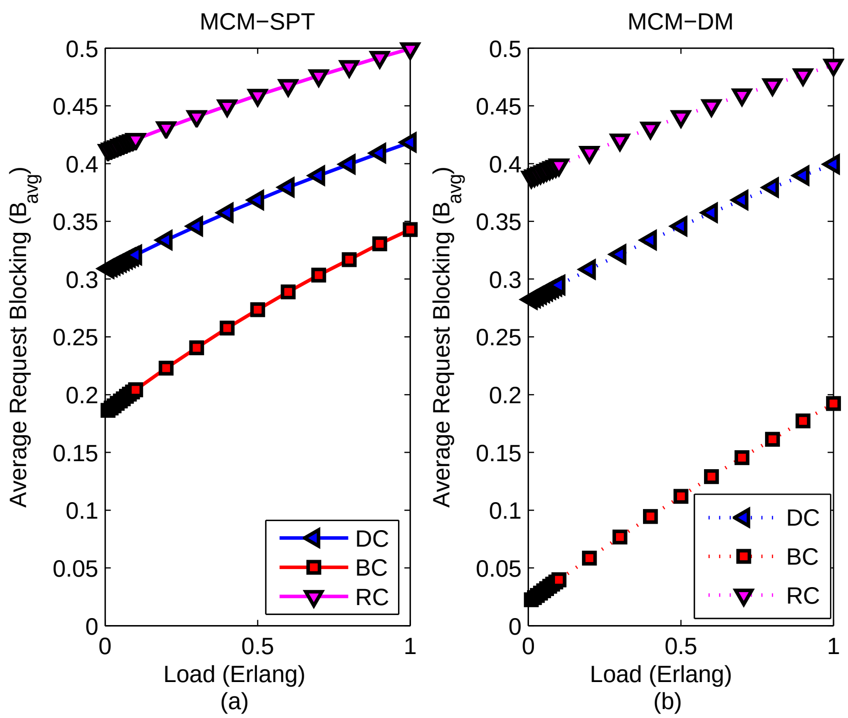


Figure 4.17: Performance of attribute specific 7/4 manycasting for (a) MCM-SPT and (b) MCM-DM.

the manycasting over OBS networks.

Chapter 5

Conclusion and Future Scope

5.1 Conclusion

With the advent of many distributed Internet applications, demand for the bandwidth has been increased tremendously. Optical networks are the potential source for supporting these new emerging Internet applications. Hence there a need to develop intelligent optical network control plane protocols, called *optical control plane*. This work focused on supporting Grid applications over optical networks. In order to support Grid applications, the control plane signaling should be capable of handling of dynamic connection requests. We have primarily focused on developing intelligent distributed control plane rather than the centralized. In a Grid network, there is need for revisiting the RWA algorithms for the path computation, in regard to physical layer impairments in AON networks.

In order to accomplish the above mentioned needs/demands we have proposed *Manycasting*. Manycasting is found to be a viable communication for Grid and many distributed Internet applications. Manycasting can implement, user controlled network infrastructure for supporting dynamic and interactive service of the Grid. In this work we have indicated the need for supporting manycasting by optical burst switched networks for improving performance of the network.

All optical network architecture has been presented. Linear optical impairments like ASE were considered in this work. Burst loss based on contention and signal quality is evaluated for unicasting. Later, this work was extended to manycasting problem. A distributed version of shortest path tree was proposed. Using multicast capable switch architectures, the signal impairment loss is calculated. Signal and ASE powers were evaluated recursively for each link and are updated at the nodes. Thus network elements (NEs) have the intelligence about the signal quality and bursts are scheduled only if the desired signal quality is maintained. Burst header packets are used for control plane signaling. BHPs which were earlier used to carry the information about the wavelength availability, are now incorporated with additional functionality of q-factor. To reduce the blocking probability due to contention and optical signal quality algorithms have been proposed in *Chapter 3*.

- Impairment Aware Shortest Path Tree (IA-SPT) selects the first k destinations out of m in-terms of the shortest distance. Then the bursts are routed to these destinations. However due to random contention in the network, burst may not reach k of them and hence the manycast request can be lost. This results in the poor performance of IA-SPT.

- Impairment Aware Static Over Provisioning (IA-SOP) selects first $k+k'$ destinations where $0 \leq k' \leq m-k$. This kind of over provisioning can help burst to reach more destinations than k and thus satisfying the manycast request. However by over-provisioning, the fan-out of the multicast capable-optical crossconnect (MC-OXC) increases and hence the power loss due to the optical signal split could be high. This causes an increase in burst loss due to OSNR.

- Impairment Aware Dynamic Membership (IA-DM) has been proposed to overcome these shortcomings. In this algorithm destinations are classified as primary and secondary destinations. Due to the contention if, one of the primary destination is blocked, then the burst is routed to a destination selected from secondary. This selected destination can be on the longer distance and if the burst still meets the required BER threshold, then manycast request is said to be satisfied.

The performance evaluation of these algorithms has been carried using discrete-event simulation model and are verified with the proposed analytical model.

Finally the impact of QoS provisioning for manycasting over OBS networks has been discussed. In this part of the work we have focused on QoS attributes like noise factor, reliability and propagation delay. These service attributes were calculated based on link-by-link using the operation defined by $\circ$. BHPs were modified in accordance to maintain the path information vector of the QoS attributes. The QoS provisioning problem has been addressed as Multi-Constraint Manycast Problem (MCMP). In this part we have considered two algorithms,

- MCMP-Shortest path Tree (SPT) based on the shortest path tree computation.
- MCMP-Dynamic Membership (DM) in which destinations join or leave the quorum pool based on the QoS requirements.

These proposed algorithms are distributed in nature. The performance of these algorithms were evaluated based on different service requirements. We have differentiated services as data or real time service. Results were presented for different set of service thresholds. With the help of simulations we conclude that MCM-SPT performs better than MCM-DM for real-time services and the data services can be provisioned using MCM-DM.

5.2 Future Scope

In the present work we have considered AON and hence network cost was not the part of the performance evaluation. However this work can be further extended to translucent optical networks, where the wavelength converters (WCs) and regenerators (WRs) are sparsely located. By optimizing the placement of these WCs and WRs cost optimization can be achieved.

RWA algorithms are time consuming and hence there is a need to develop efficient analytical models. These analytical models should be able to incorporate the physical layer impairments.

In this entire study we have considered impact of linear impairments. Modeling of non-linear impairments for the manycasting scenario would be necessary for the data rates greater than 10 Gbps and beyond. We are also investigating multi constraint manycast problem based on Genetic algorithms (GA).

Appendix A

Relationships between *q*-factor, OSNR, and noise factor

A.1 Derivation of Eq. (3.5)

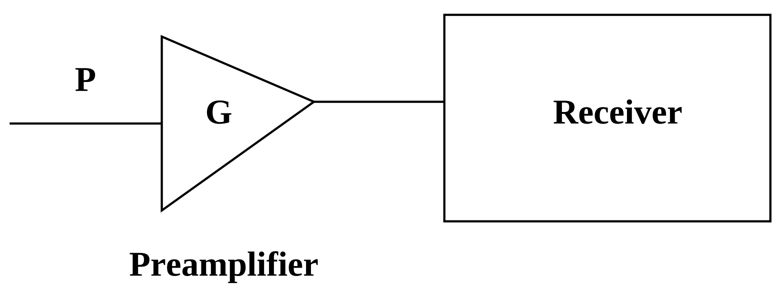


Figure A.1: Receiver with optical preamplifier

The amplifier spontaneous emission (ASE) noise power at the output of the amplifier for each polarization mode is given by,

$$P_N = n_{sp} h f_c (G-1) B_o, \tag{A.1}$$

where n_{sp} is a constant called the spontaneous emission factor, G is the amplifier gain and B_o is the optical bandwidth. Two fundamental modes are present in the fiber and hence the total noise power at the output of the amplifier is $2P_N$. n_{sp} depends on the population inversion within the amplifier. Typically it is around 2-5 for most of the amplifiers. We define,

$$P_n = n_{sp} h f_c.$$

The photo-detector produces a current that is proportional to the incident power. The signal current is given by,

$$I = \mathcal{R} G P, \tag{A.2}$$

where P is the received optical signal power as shown in the Fig. A.1, $\mathcal{R}$ is the responsivity of the photo-detector. In addition to shot and thermal noise, there are signal-spontaneous noise, and spontaneous-spontaneous noise currents at the receiver given by,

$$\sigma^2_{sig-spont} = 4\mathcal{R}^2 G P P_n (G-1) B_e, \tag{A.3}$$

$$\sigma^2_{spont-spont} = 2\mathcal{R}^2 [P_n(G-1)]^2 (2B_o - B_e) B_e, \tag{A.4}$$

where B_e is the electrical bandwidth. Derivation of these variances are given in [61]. In the case of systems with cascades of optical amplifiers, two parameters that are measured are the average received signal power $\bar{P}_{rec}$ and received optical signal noise power P_{ase}. In the case of optically preamplified receiver, $P_{ase} = 2P_n(G-1)B_o$. The optical signal to noise ratio (OSNR) is defined as $\bar{P}_{rec}/P_{ase}$. q-factor is given by,

$$q = \frac{I_1 - I_0}{\sigma_0 + \sigma_1} \tag{A.5}$$

Assuming $I_0 = 0$ we have $\bar{P}_{rec} = P_1/2$ and the Eq. (A.5) becomes,

$$q = \frac{\mathcal{R} G P_1}{\sigma_0} \left(1 + \frac{\sigma_1}{\sigma_0}\right)^{-1}$$

$$q = \frac{\mathcal{R}(2\bar{P}_{rec})}{\sqrt{2\mathcal{R}^2[P_n(G-1)]^2(2B_o - B_e)B_e}} \left(1 + \sqrt{1 + \frac{4\mathcal{R}^2 G P_1 P_n (G-1) B_e}{2\mathcal{R}^2[P_n(G-1)]^2(2B_o - B_e)B_e}}\right)^{-1}$$

$$\begin{aligned} q &= \frac{4 \times OSNR}{\sqrt{\frac{2B_e(2B_o - B_e)}{B_o^2}}} \left(1 + \sqrt{1 + 8 \times OSNR \times \frac{B_o}{2B_o - B_e}}\right)^{-1} \\ &= \frac{2\sqrt{\frac{B_o}{B_e}} \times OSNR}{1 + \sqrt{1 + 4 \times OSNR}} \text{ if } 2B_o \gg B_e \qquad (A.6) \\ &= \frac{2\sqrt{2} \times OSNR}{1 + \sqrt{1 + 8 \times OSNR}} \text{ if } B_o = B_e. \qquad (A.7) \end{aligned}$$

In this thesis calculation of q-factor is based on Eq. (A.6) even for $B_o = B_e$ case. Error that occurs due to this, is however negligible.

A.2 Calculation of Noise factor threshold

We calculate the noise factor threshold η_{th} using Eqs.(4.7), (4.8), (4.9). From Eq.(4.7), the noise factor of the link j is given by,

$$\eta_j = \left(\frac{P(j)}{P_{ase}(j)}\right) \left(\frac{P_{ase}(j+1)}{P(j+1)}\right)$$

For the path from Source s to the Destination d, the overall noise factor is given by,

$$\eta_{\langle s,d \rangle} = \left(\frac{P(s)}{P_{ase}(s)}\right) \left(\frac{P_{ase}(d)}{P(d)}\right)$$

We assume the transmitting power of the receiver is $P(s) = 1$ mW. The ASE noise power at Source node s is given by,

$$\begin{aligned} P_{ase}(s) &= P_{ini}L_dL_mL_tL_{ins}(G_{in}-1)G_{out} \\ &\quad +P_{ini}L_t[G_{out}-1] \end{aligned} \tag{A.8}$$

$P_{ini} = 2n_{sp}hf_cB_o$ where n_{sp} spontaneous-emission factor, h is the Plank's Constant, and f_c is the central frequency of the optical signal. L_d, L_m, L_t, and L_{ins} are demultiplexer, multiplexer, tap, and insertion loss of the optical cross-connect switch respectively. G_{in} and G_{out} are the input and output gains of the erbium doped fiber amplifier (EDFA) in switch. Parameter values and the switch architecture can be found in [75], [73], [35]. By using the Eq. (A.8) we get the $P_{ase}(s) = 0.0042$ mW. Thus the OSNR at Source s will be $OSNR(s) = 238$ a.u. For BER of 10^{-12} we need $q \approx 7$, for which the required $OSNR(d)$ is obtained by solving the Eq. (4.8),

$$7 = \frac{2\sqrt{\frac{B_o}{B_e}}OSNR(d)}{1+\sqrt{1+4OSNR(d)}} \tag{A.9}$$

For a system operating $B = 10$ Gb/s with $B_o = 70$ GHz and $B_e = 0.7B$, $OSNR(d) = 56$ which is obtained by solving Eq. A.9. Hence if the $OSNR(d) < 56 (= OSNR_{min})$ then the BER will increase beyond 10^{-12}. Thus the noise factor threshold $\eta_{max} = 4.25$ corresponding to $q = 7$. Similarly for $q = 6$, $\eta_{max} \approx 6$. Thus we see that as long as the noise-factor of the burst is $\leq \eta_{max}$, burst can be scheduled for transmission. We derive the relation for noise-factor threshold (η_{max}) and q-factor threshold (q_{th}).

$$\eta_{max} = \frac{OSNR(s)}{OSNR_{min}}. \tag{A.10}$$

In order for the BER to be lesser than the given threshold, the OSNR at the destination should be greater than the $OSNR_{min}$. Thus the Eq. (4.8) at the threshold conditions is given by,

$$q_{th} = \frac{2\sqrt{\frac{B_o}{B_e}}OSNR_{min}}{1+\sqrt{1+4OSNR_{min}}} \tag{A.11}$$

Solving this equation for $OSNR_{min}$ we get,

$$ONSR_{min} = q_{th}\left(q_{th} + \sqrt{\frac{B_o}{B_e}}\right)\frac{B_o}{B_e} \tag{A.12}$$

Substituting Eq. (A.12) in Eq. (A.10) we get,

$$\eta_{max} = \frac{OSNR(s)}{q_{th}\left(\frac{B_o}{B_e}\right)\left(q_{th} + \sqrt{\frac{B_o}{B_e}}\right)}. \tag{A.13}$$

Appendix B

Procedures used in MCM-DM algorithm

B.1 Pseudo Code for Procedure.Block()

Procedure.Block( ) {

$NOT.REACH.DEST[n] \leftarrow DEST[n]$;

if $|NOT.REACH.DEST[n]| > \kappa_u$ **then**

  $DELETE.BURST[id]$;

else

  `continue`;

end if} /* `End of Procedure.Block( )`*/

B.2 Pseudo Code for Procedure.QoS()

Procedure.QoS( ) {

$\Omega_{\langle u,n\rangle} \leftarrow \Omega_{\langle u-1,u\rangle} \circ \Omega_{\langle u,n\rangle}$;

if $\Omega_{\langle u,n\rangle} \preccurlyeq \top^{(\theta_p)}$ **then**

  $UPDATE.BHP[n]$;

  $(id, n, D_n, \kappa_n, \top^{(\theta_p)}, \Omega_{\langle u,n\rangle})$;

else

Procedure.Block( );

end if} /* `End of Procedure.QoS( )`*/

B.3 Round Robin Procedure

Details of the round-robin procedure used in the Algorithm 5 is given below,

$N_u \leftarrow NEXT.HOP.NODES\{k_u\}$;

$i \leftarrow 1$;

for $l \leftarrow 1$ to $|D_u| - k_u$ **do**

$D_{N_i} \leftarrow D_{N_i} \cup \{d_{k_u+l}\}$;

if $i < |N_u|$ **then**

$i \leftarrow i + 1$;

else

$i = 1$;

end if

end for

Bibliography

[1] B. Mukherjee, "WDM optical communications networks: Progress and challenges," *IEEE Journal of Selected Areas in Communications (JSAC)*, vol. 18, no. 10, pp. 1810–1824, 2000.

[2] Y. Lee and B. Mukherjee, "Traffic engineering in next-generation optical networks," *IEEE Communictions Surveys & Tutorials*, vol. 6, no. 3, pp. 16–33, 2004.

[3] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: An approach to high bandwidth optical WANs," *IEEE Transcations on Communications*, vol. 40, no. 07, pp. 1171–1182, 1992.

[4] E. Desurive, *Erbium-Doped Fiber Amplifiers: Principles and Applications.* New York, USA: John Wiley, 1994.

[5] Lightreading. [Online]. Available: http://www.lightreading.com

[6] C. Qio and M. Yoo, "Optical burst switching OBS-a new paradigm for an optical internet," *Journal of High Speed Networks*, vol. 8, no. 1, pp. 69–84, 1999.

[7] T. Battestilli and H. Perros, "An introduction to optical burst switching," *IEEE Communication Magazine*, vol. 41, no. 8, pp. S10–S15, 2003.

[8] V. M. Vokkarane, "Design and analysis of architectures and protocols for optical burst-switched networks," Ph.D. dissertation, University of Texas at Dallas, Texas, USA, 2004.

[9] C. Qiao and M. Yoo, "Choices, features and issues in optical burst switching," *Optical Networking Magazine*, vol. 1, no. 2, pp. 36–44, 1999.

[10] I. Baldine, G. N. Rouskas, H. G. Perros, and D. Stevenson, "JumpStart: A just-in-time signaling architecture for WDM burst-switched networks," *IEEE Communication Magazine*, vol. 40, no. 2, pp. 82–89, 2002.

[11] J. Y. Wei and R. I. McFarland, "Just-in-time signaling for WDM optical burst switching network," *IEEE/OSA Journal of Lightwave Technology*, vol. 18, no. 12, pp. 2019–2037, 2000.

[12] X. Huang, Q. She, T. Zhang, K. Lu, and J. P. Jue, "Modeling and performance analysis of small group multicast with deflection routing in optical burst switched networks," *IEEE Journal on Selected Area in Communications*, vol. 26, no. 3, pp. 74–86, 2008.

[13] L. Guanghong and L. Zhaohong, "Performance analysis of communication paradigm based on anycast in the grid over OBS," in *Proc. IEEE International Conference on Convergence Information Technology (ICCIT 2007)*, Gyeongju, Republic of Korea, Nov. 2007, pp. 1874–1879.

[14] M. D. Leenheer, F. Farahmand, K. Lu, P. Thysebaert, B. Volckaert, F. D. Turck, B. Dhoedt, P. Demeester, and J. P. Jue, "Anycast routing in optical burst switched grid networks," in *Proc. IEEE International European Conference on Optical Communications, (ECOC-'05)*, Glasgow, Scotland, Sept. 2005.

[15] Service Oriented Optical Networks (SOON). [Online]. Available: http://www.cis.umassd.edu/~vvokkarane/soon

[16] M. Russell, G. Allen, G. Daues, I. Foster, E. Seidel, J. Novotny, J. Shalf, and G. Laszewski, "The astrophysics simulation collaboratory: A science portal enabling community software development," *Cluster Computing, Springer Netherlands*, vol. 5, no. 3, pp. 297–304, 2002.

[17] M. Russell, G. Allen, G. Daues, I. Kelley, M. Russell, E. Seidel, J. Shalf, and M. Tobias, “The astrophysics simulation collaboratory portal: A framework for effective distributed research,” *Cluster Computing, Springer Netherlands*, vol. 21, no. 2, pp. 259–270, 2005.

[18] myGrid. [Online]. Available: http://www.mygrid.org.uk

[19] J. Mambretti, J. Weinberger, J. Chen, E. Bacon, F. Yeh, D. Lillethun, B. Grossman, Y. Gu, and M. Mazzuca, “The photonics TeraStream:enabling next generation applications through intelligent optical networking at iGrid 2002,” *Journal of Future Computer Systems, Elsevier Press*, vol. 19, no. 6, pp. 897–908, 2003.

[20] R. Grossman, Y. Gu, D. Hamleburg, D. Hanley, X. Hong, J. Levera, M. Mazzuca, D. L. J. Mambretti, and J. Weinberger, “Experimental studies using photonic data services at iGrid 2002,” *Journal of Future Computer Systems, Elsevier Press*, vol. 19, no. 6, pp. 945–956, 2003.

[21] Optical Metro Network initiative (OMNI). [Online]. Available: http://www.icair.org/omninet

[22] M. Blanchet, F. Parent, and B. S. Arnaud. (March 2001) Optical (OBGP): InterAS lightpath provisioning, IETF network working group report. [Online]. Available: http://www.search.ietf.org/internet-drafts/draft-parent-obgp-01.txt

[23] Translight. [Online]. Available: http://www.startap.net/translight

[24] R. Nejabati. (2006) Grid optical burst switched networks (GOBS). [Online]. Available: www.ogf.org/Public_Comment_Docs/Documents/Jan-2007/OGF_GHPN_GOBS_final.pdf

[25] F. Farahmand, M. D. Leenheer, P. Thysebaert, B. Volckaert, F. D. Turck, B. Dhoedt, P. Demeester, and J. P. Jue, “A multi-layered approach to optical burst-switched based grids,” in *Proc. IEEE International Conference on Networking and Services (ICNS-’06)*, Silicon Valley, CA, USA, July 2006, pp. 63–69.

[26] I. Foster, "The grid: A new infrastructure for 21st century science," *Physics Today*, vol. 55, no. 2, pp. 42–47, 2002.

[27] I. Foster, C. Kesselman, and S. Tuecke, "The anatomy of grid," *International Journal of High Performance Computing*, vol. 15, no. 3, pp. 200–222, 2001.

[28] F. Farahmand, V. M. Vokkarane, J. P. Jue, J. J. P. C. Rodrigues, and M. M. Freire, "Optical burst switching network: A multi-layered approach," *Journal of High Speed Networks (JHSN)*, vol. 16, no. 2, pp. 105–122, 2007.

[29] B. Ramamurthy, H. Feng, D. Datta, J.P.Heritage, and B. Mukherjee, "Transparent vs. opaque vs. translucent wavelength-routed optical networks," in *Optical Fiber Communication (OFC '99) Technical Digest*, CA, USA, Feb. 1999, pp. 59–61.

[30] X. Yang and B. Ramamurthy, "Dynamic routing in translucent WDM optical networks," in *Proc. IEEE ICC-2002*, New York, USA, Apr. 2002, pp. 2796–2802.

[31] B. Ramamurthy, S. Yaragoral, and X. Yang, "Translucent optical WDM networks for next-generation backbone networks," in *Proc. IEEE Globecom-2001*, Texas, USA, Dec. 2001, pp. 60–64.

[32] X. Yang and B. Ramamurthy, "Dynamic routing in translucent WDM optical networks: The intradomain case," *IEEE/OSA Journal of Lightwave Technology*, vol. 23, no. 3, pp. 955–971, 2005.

[33] B. Mukherjee, *Optical Communication Networks*. New York, USA: McGraw-Hill, 1997.

[34] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength routed optical WDM networks," *Optical Networks Magazine*, vol. 1, no. 1, pp. 47–60, 2000.

[35] B. Ramamurthy, D. Datta, H. Feng, J. P. Heritage, and B. Mukherjee, "Impact of transmission impairments on the teletraffic performance of wavelength-routed optical

networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 17, no. 10, pp. 1713–1723, 1999.

[36] B. Ramamurthy, *Design of Optical WDM Networks - LAN, MAN and WAN.* Norwell, MA, USA: Kluwer Academic Publishers, 2000.

[37] Y. Huang, J. P. Heritage, and B. Mukherjee, "Connection provisioning with transmission impairment consideration in optical WDM networks with high-speed channels," *IEEE/LEOS Journal of Lightwave Technology*, vol. 23, no. 3, pp. 982–993, 2005.

[38] T. Deng and S. Subramaniam, "QoS-friendly wavelength assignment in dynamic wavelength routed optical networks," *Photonics Network Communications*, vol. 10, no. 1, pp. 5–22, 2005.

[39] M. A. C. Lima, A. C. César, and A. F. R. Araújo, "Optical network optimization with transmission impairments based on genetic algorithm," in *Proc. SMBO/IEEE International Microwave and Optoelectronics Conference (IMOC)*, Iguaza Falls, Brazil, Sept. 2003, pp. 361–365.

[40] Y. Pointurier, M. Brandt-Pearce, T. Deng, and S. Subramaniam, "Fair QoS-aware adaptive routing and wavelength assignment in all-optical networks," in *Proc. IEEE International Conference on Communications ICC*, Istanbul, Turkey, June 2006, pp. 2433–2438.

[41] Y. Pointurier, "Cross-layer design of all-optical networks incorporating crosstalk effects," Ph.D. dissertation, University of Virgina, USA, 2006.

[42] Lambda optical systems networks. [Online]. Available: http://www.lambdaopticalsystems.com

[43] Movaz networks. [Online]. Available: http://www.movaz.com

[44] A. Biram, "Computing approximate blocking probabilites for the class of all-optical networks," *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, pp. 852–857, 1996.

[45] M. Duser and P. Bayvel, "Analysis of dynamically wavelength-routed optical burst switched network architecture," *IEEE/OSA Journal of Lightwave Technology*, vol. 20, no. 4, pp. 574–585, 2002.

[46] K. Lu, G. Xiao, and I. Chamtac, "Analysis of blocking probability for distributed lightpath establishment in WDM optical networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 1, pp. 187–197, 2005.

[47] A. Sridharan and K. N. Sivarajan, "Blocking in all-optical networks," *IEEE/ACM Transactions on Networking*, vol. 12, no. 2, pp. 384–397, 2004.

[48] I. Karamitsos, "Analytical model for optical burst switching (OBS) networks with linear transmission impariments," in *Proc. IEEE Second International Conference on Wireless and Optical Communications Networks (WOCN) 2005*, Dubai, United Arab Emirates (UAE), Mar. 2005, pp. 192–195.

[49] I. Karamitsos, A. Tzanakaki, and I. Tomkos, "Study of physical effects as a performance parameter in optical burst switched networks," in *Proc. IEEE 7th International Conference on Transparent Optical Networks ICTON 2005*, Barcelona, Spain, July 2005, pp. 228–231.

[50] J. Strand and A. L. Chiu, "Issues for routing in the optical layer," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 81–87, 2001.

[51] M. Ali and L. Tancevski, "Impact of polarization mode dispersion on the design of wavelength routed networks," *IEEE Photonics Teechnology Letters*, vol. 14, no. 5, pp. 720–722, 2002.

[52] I. E. Fonseca, M. R. N. Riberiro, and H. Waldman, "Preserving global optical QoS in FWM impaired dynamic networks," *IEEE Electronic Letters*, vol. 40, no. 3, pp. 191–192, 2004.

[53] R. Martinez, F. Cugini, N. Andriolli, L. Wosinska, and J. Comellas, "Challenges and requirements for introducing impairment-awareness into management and control

planes of ASON/GMPLS WDM networks," *IEEE Communication Magazine*, vol. 44, no. 12, pp. 76–85, 2007.

[54] B. G. Balagangadhar, T. Srinivas, and P. Rajappa, "Constraint based routing for wavelength routed optical networks," in *Proceedings of International Conference on Optics and Lasers (ICOL)-2005*, Dehradun, India, Dec. 2005.

[55] M. D. Leenheer, F. Farahmand, K. Lu, T. Zhang, P. Thysebaert, B. Volckaert, F. D. Turck, B. Dhoedt, P. Demeester, and J. P. Jue, "Anycast algorithms supporting optical burst switched grid networks," in *Proc. IEEE International Conference on Networking and Service, (ICNS-'06)*, Silicon Valley, CA, USA, July 2006, pp. 63–69.

[56] B. G. Bathula and J. M. H. Elmirghani, "Providing QoS for anycasting over optical burst switched grid networks," in *Proc. of Workshop on Service Aware Optical Grid Networks*, Beijing, China, Oct. 2008.

[57] S. Y. Cheung and A. Kumar, "Efficient quorumcast routing algorithms," in *Proc. IEEE INFOCOM-'94*, Toronto, Ontario, Canada, June 1994, pp. 840–847.

[58] N. Garg, "Saving an epsilon: a 2-approximation for the k-mst problem in graphs," in *Proc. on 33rd Annual ACM Symposium on Theory of Computing-'01*, Baltimore, MD, USA, July 2005, pp. 396–402.

[59] B. Wang and J. C. Hou, "An efficient QoS routing algorithm for quorumcast communication," in *Proc. IEEE Network Protocols-'01*, Riverside, California, Nov. 2001, pp. 110–118.

[60] X. Huang, Q. She, V. M. Vokkarane, and J. P. Jue, "Manycasting over optical burst-switched networks," in *Proc. IEEE International Conference on Communication (ICC-'07)*, Glasgow, Scotland, May 2007, pp. 2353–2358.

[61] R. Ramaswami and K. N. Sivarajan, *Optical Networks*. San Francisco, CA, USA: Morgan Kaufmann Publishers, 2004.

[62] Y. Chen, C. Qiao, and X. Yu, "Optical burst switching: A new area in optical networking research," *IEEE Network*, vol. 18, no. 3, pp. 16–23, 2004.

[63] B. Ramamurthy and B. Mukherjee, "Wavelength conversion in WDM networking," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 7, pp. 1061–1073, 1998.

[64] C. Gauger, "Dimensioning of FDL buffers for optical burst switching nodes," in *Proceedings, Optical Network Design and Modeling (ONDM 2002)*, Torino, Feb. 2002.

[65] V. M. Vokkarane and J. P. Jue, "Segmentation-based non-preemptive channel scheduling algorithms for optical burst-switched networks," *IEEE/OSA Journal of Lightwave Technology (JLT), Special Issue on Optical Networks*, vol. 23, no. 10, pp. 3125–3137, 2005.

[66] C. Hsu, T. Liu, and N. Huang, "Performance analysis of deflection routing in optical burst switched networks," in *Proc. IEEE INFOCOM 2002*, New York, USA, June 2002, pp. 66–73.

[67] W. S. Hu and Q. J. Zeng, "Multicasting optical cross connects employing splitter-and-delivery switch," *IEEE Photonics Technology Letters*, vol. 10, no. 7, pp. 970–972, 1998.

[68] Y. Xin and G. N. Rouskas, "Multicast routing under optical constraint," in *Proc. IEEE International Conference of Computer and Communications Societies (INFOCOM-'04)*, Hnng Kong, China, Mar. 2004, pp. 2731–2742.

[69] M. Ali and J. Deogun, "Power-efficient design of multicast wavelength routed networks," *IEEE Journal on Selected Area in Communications*, vol. 18, no. 10, pp. 1852–1862, 2000.

[70] H. L. Vu and M. Zukerman, "Blocking probability for priority classes in optical burst switching networks," *IEEE Communications Letters*, vol. 6, no. 5, pp. 214–216, 2002.

[71] D. Bertsekas and R. Gallager, *Data Networks.* Upper Saddle River, NJ: Prentice-Hall, 1992.

[72] R. Ravi, R. Sundaram, M. V. Marathe, D. J. Rosenkrantz, and S. S. Ravi, "Spanning trees short or small," in *Proceedings of the fifth annual ACM-SIAM symposium on Discrete algorithms*, Arlington, Virgina, United States, Jan. 1994, pp. 546–555.

[73] B. G. Bathula, V. M. Vokkarane, R. R. C. Bikram, and T. Srinivas, "Impairment-aware manycast algorithms over optical burst switched networks," in *to appear, Proc. IEEE International Conference on Computer Communications and Networks (ICCCN)-'08*, St.Thomas U.S. Virgin Islands, USA, Aug. 2008.

[74] A. Kaheel, T. Khattab, A. Mohamed, and H. Alnuweiri, "Quality-of-service mechanisms in IP-over-WDM netowrks," *IEEE Communications Magazine*, vol. 40, no. 12, pp. 38–43, 2002.

[75] B. G. Bathula, V. M. Vokkarane, and R. R. C. Bikram, "Impairment-aware manycasting over optical burst switched networks," in *to appear, Proc. IEEE International Conference on Communication (ICC)-'08*, Bejing, China, May 2008, pp. 5234–5238.

[76] X. Qiu, R. Telikepalli, T. Drwiega, and J. Yan, "Reliability and availability assessment of storage area network extension solutions," *IEEE Communication Magazine*, vol. 43, no. 3, pp. 80–85, 2005.

[77] A. Jukan and G. Franzl, "Constraint-based path selection methods for on-demand provisioning in WDM networks," in *Proc. IEEE International Conference of Computer and Communications Societies (INFOCOM-'02)*, New York, NY, USA, June 2002, pp. 827–836.

[78] ——, "Path selection methods with multiple constraints in service-guaranteed WDM networks," *IEEE/ACM Transactions on Networking*, vol. 12, no. 1, pp. 59–72, 2004.

[79] A. Jukan, "Optical control plane for the grid community," *IEEE Communications Surveys & Tutorials*, vol. 9, no. 3, pp. 30–44, 2007.

[80] Q. She, X. Huang, N. Kannasoot, Q. Zhang, and J. P. Jue, "Multi-resource manycast over optical burst switched networks," in *Proc. IEEE International Conference on Computer Communications and Networks (ICCCN-'07)*, Honolulu, Hawaii, USA, Aug. 2007, pp. 222–227.

[81] A. B. Przygienda, "Link state routing with QoS in ATM LANs," Ph.D. dissertation, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, 1995.

[82] B. A. Davey and H. A. Priestley, *Introduction to Lattices and Order.* Cambridge, United Kingdom: Cambridge University Press, 2002.

[83] M. Yoo and C. Qiao, "Just-enough-time (JET): A high speed protocol for burst traffic in optical networks," in *Proc. IEEE/LEOS Summer Topical Meetings Digest for Conf. Technologies Global Information Infrastructure*, Montreal, Canada, Aug. 1997, pp. 26–27.

[84] J. P. Jue and V. M. Vokkarane, *Optical burst-switched networks.* Germany: Springer, Optical Networks Series, 2005.