

University of Windsor

Scholarship at UWindor

Electronic Theses and Dissertations

Theses, Dissertations, and Major Papers

2018

Robust Data Center Network Design using Space Division Multiplexing

Ankita Biswas
University of Windsor

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

Recommended Citation

Biswas, Ankita, "Robust Data Center Network Design using Space Division Multiplexing" (2018). *Electronic Theses and Dissertations*. 7349.

<https://scholar.uwindsor.ca/etd/7349>

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.

Robust Data Center Network Design using Space Division Multiplexing

By

Ankita Biswas

A Thesis

Submitted to the Faculty of Graduate Studies
through the School of Computer Science
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

2018

© 2018 Ankita Biswas

Robust Data Center Network Design using Space Division Multiplexing

By

Ankita Biswas

APPROVED BY:

A. Sarker

Department of Mathematics and Statistics

S. Mavromoustak

School of Computer Science

S. Bandyopadhyay, Co-Advisor

School of Computer Science

A. Jaekel, Advisor

School of Computer Science

January 16, 2018

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

With the ever-increasing demand for data transmission in our generation where Internet and cloud concepts play a vital role, it has become very essential that we handle data in a most efficient way. A possible solution to overcome the capacity crunch problem which is so evident in future, is applications of Space Division Multiplexing, where we explore the remaining unused domain that is the spatial domain. Space Division Multiplexing using multi-core fibers (MCFs), and few-mode fibers (FMFs) has been studied in our work to enhance the data-carrying capacity of optical fibers, while minimizing the transmission cost per bit. The objective of our work is to develop a path protection scheme to handle communication requests in data center (DC) networks using elastic optical networking and space division multiplexing (SDM). Our approach to this problem is to 1) determine a dedicated primary and backup path, 2) possible allocation of spectrum using the flex-grid fixed-SDM model, 3) choose the best possible modulation format to minimize the number of sub carriers needed for data transfer, 4) measure the cost of the resources required to handle the new requests. We propose to evaluate the developed Integer Linear Programming (ILP) formulation based on this scheme, considering the possibility of disasters. We study the impact of the design on the cost of the solution, hence explore whether it promotes significant resource savings.

DEDICATION

*Dedicated to my parents Debasis and Gopa Biswas, and my
sister Suchita Biswas.*

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my advisor Dr. Subir Bandyopadhyay and Dr. Arunita Jaekel for their continuous support in my research, for their patience, motivation, and immense knowledge. Their guidance helped me throughout my time of research and writing of this thesis. I could not have imagined having better mentors for my study.

I would also like to thank my thesis committee members Dr. Animesh Sarker and Dr. Stephanos Mavromoustak for their valuable comments and suggestions for the completion of this thesis. I would like to sincerely thank Dr. Saja Al Mamoori, who has been the fundamental support throughout my work.

Also, I would like to thank my fellow lab mates for the stimulating discussions during the development of this work. Last but not the least, my deepest gratitude to my family for their unconditional love, support, and belief.

TABLE OF CONTENTS

DECLARATION OF ORIGINALITY	iii
ABSTRACT	iv
DEDICATION	v
ACKNOWLEDGMENT	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF SYMBOLS	xi
1 INTRODUCTION	1
1.1 Overview	1
1.2 Motivation	3
1.3 Problem Statement & Solution Outline	4
1.4 Scope of thesis	5
1.5 Structure of thesis	5
2 REVIEW OF RELATED TOPICS	7
2.1 Fundamentals of Optical Networks	7
2.1.1 Components of Optical fiber	8
2.1.2 Total Internal Reflection	9
2.2 Cloud Computing	9
2.3 Data Center	10
2.4 Disaster	11
2.4.1 Types of failures during a disaster	12
2.5 Faults in Optical Network	13
2.6 File Replication	13
2.7 Optical OFDM	15
2.8 Orthogonality Principle in OFDM	17
2.9 Routing and Spectrum Allocation	18
2.9.1 Spectrum Allocation in OFDM Networks	19
2.10 Space Division Multiplexing	20
2.10.1 Types of Fibers used in SDM	22
2.10.2 Ways of realizing SDM Transmission	24
2.11 Optical Reach	26
2.12 Modulation Format	27
2.13 Literature Review	29
3 OPTIMAL RSA FOR SDM OPTICAL NETWORK UNDER DYNAMIC TRAFFIC	34
3.1 Introduction to the problem	34
3.2 Assumptions made	36
3.3 The Objective of our algorithm	37
3.4 Idea of Virtual Node	38
3.5 Concept of Gaps	39
3.6 Notations used in the ILP	42
3.7 Formulation of ILP for optimal solution	45

3.7.1	Justification of ILP	52
4	SIMULATION EXPERIMENTS AND RESULTS	57
4.1	Experimental Setup	58
4.2	Performance Study	62
5	CONCLUSIONS AND FUTURE WORK	69
5.1	Conclusions	69
5.2	Future work	70
	BIBLIOGRAPHY	71
	VITA AUCTORIS	76

LIST OF TABLES

Table 4.11 Percentage increase in resource usage with increase in cores	65
---	----

LIST OF FIGURES

2.1	Optical Cable	8
2.2	Total Internal Reflection	9
2.3	Data Center Server [1]	11
2.4	Link Failure	12
2.5	Node Failure	12
2.6	Communication in fault free network	14
2.7	Communication in disaster affected network	15
2.8	Spectrum of WDM	16
2.9	Spectrum of OFDM	17
2.10	Orthogonality in OFDM	17
2.11	Six node network topology	19
2.12	Existing communication	20
2.13	Allocation of new resources	20
2.14	Six node network topology for SDM	21
2.15	Multiple core division in SDM	22
2.16	Single mode fiber	22
2.17	Multi core fiber	23
2.18	Fewer-mode and Multi-mode fiber	24
2.19	Models used in SDM	26
2.20	Six node network topology showing optical reach	27
2.21	Modulation format chart	28
3.1	Six node network topology with a virtual node at 0	38
3.2	A six node network topology with data centers	39
3.3	Available spectrum	40
3.4	Bandwidth allocation on available spectrum	40
3.5	Bandwidth allocation on cores of available spectrum	41
3.6	Gaps	41
4.1	Flow diagram for experimental setup	58
4.2	Topology 11-Node: 24 links	62
4.3	Topology 14-Node: 21 links	62
4.4	Resource usage 11-node, 14-node network	63
4.5	Blocking probability in 11-node, 14-node network	64
4.6	Resource usage vs number of requests with disaster with 4 cores	65
4.7	BP vs number of requests with 4 cores	66
4.8	Resource usage vs number of requests with 6 cores	66
4.9	BP vs number of requests with 6 cores	67
4.10	Resource usage vs number of cores with 50 demands	68
4.11	Percentage increase in resource usage with increase in cores	68

LIST OF SYMBOLS

Symbol	Definition
BP	Blocking Probability
BPSK	Binary Phase Shift Keying
COST-239	European Network Topology
DC	Datacenter
DCN	Datacenter Network
DEMUX	De-multiplexer
DWDM	Dense Wavelength Division Multiplexing
FMF	Fewer Mode Fiber
Gbps	Giga Bytes per second
GHz	Giga Hertz
ICT	Information and Communication Technology
ILP	Integer Linear Program
LAN	Local Area Network
MCF	Multi Core Fiber
MUX	Multiplexer
MCF	Multi Mode Fiber
NSFNET	National Science Network
MUX	Optical Carrier
OFDM	Orthogonal Frequency Division Multiplexing
QPSK	Quadrature Phase Shift Keying
QAM	Quadrature Amplitude Modulation
RWA	Routing and Spectrum Assignment
SDM	Space Division Multiplexing

SMF	Single Mode Fiber
SpRcs	Spectral and Spatial Resources
Tbps	Tera Bytes per second
TDM	Time Division Multiplexing
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing

Chapter 1

INTRODUCTION

1.1 Overview

The networks in which the dominant physical layer of technology for transport is optical fiber is known as optical networks. They can either be opaque or all-optical, or can be single-wavelength or based on dense wavelength division multiplexing (DWDM) [2]. In simple terms, it can be defined as a means of communication in which the electrical signal is converted into modulated optical signals for communicating data from specified source nodes to specified destination nodes in the network. The medium of communication is optical fiber cables. Optical networks are considered to be dominant in the information and communication (ICT) sector because of their all-optical approach for a wide-area network, where the required information can be transmitted in the optical domain between nodes and at a distance of thousands of kilometers, without any conversion to the electrical domain.

Optical fibers provide us with a higher rate of data communication, compared to copper cables. Also, it is cheap and is more resilient towards electromagnetic interference, and is capable of covering longer distances. We know optical fibers can provide a bandwidth capacity of the order of Gigabytes per second (Gbps) [3]. Furthermore, a capacity from 40 Gbps to 100 Gbps per channel is now commercially available in backbone networks in Wavelength Division Multiplexing (WDM) [3]. WDM is an optical technique where we can combine optical signals of different wavelengths onto a single strand of the optical fiber. This is achieved by adding multiplexer at the transmitters end and demultiplexer at the receivers end [4]. It combined two signals at a time on a single strand of the fiber earlier in 1978, when this technology was introduced, but now technology has enabled

a combination of up to 160 signals on a single strand of fiber at the same time. This advancement in technology has helped in increasing the capacity of the network without having to install more fibers.

The inflexible nature of WDM networks create limitations on network utilization. They require preallocated bandwidth request for a connection, even when the data rate claimed for the communication is not sufficient to fill the entire data carrying capacity of that bandwidth. Orthogonal Frequency-Division Multiplexing (OFDM) technology has recently come up as a promising technology for future high-speed optical transmission. The modulation technique used in OFDM achieves better spectral efficiency (i.e., how much data rate can be supported for a limited spectral bandwidth [5]) and impairment tolerance [6], it delivers the required capacity of bandwidth depending on the demand size. Thus, higher bandwidth capacity with an order of Terabits per second (Tbps) can be achieved with OFDM.

With the rapid advancement of technology, a lot of progress has been made in finding innovative ways to increase the data-carrying capacity of a single optical fiber, hence minimize the number of fibers used [7]. There are several possible methods for increasing transmission capacity over fixed bandwidth, but most of which are already in use [8]. Hence researchers wish to explore the one remaining unused dimension, which the spatial dimension. Thus, space division multiplexing (SDM) in optical networks seems to be coming up as a promising solution with the potential to overcome the possible capacity crunch problem in the backbone networks. In SDM, a single fiber is replaced with multiple fiber and used in parallel. This enables higher data transmission by using the same resources simultaneously, thus saving both time and resources. Researchers have been successful in exploring their attempt to optimize multiplexing in time, wavelength, phase, and polarization. Nowadays commercial systems utilize all the four dimensions to send more information through a single fiber [9].

1.2 Motivation

Optical networks form an integral part of current communications networks, carrying information anywhere starting from a few kilometers to miles on a transcontinental scale. Earlier in the 1980s, when fiber technology started to substitute copper wire, the first generation systems operated at bit rates of 40 Mb/s and required support from repeaters every 10 km. But today's technologically updated networks offer 100 Gb/s per wavelength [3]. However the increase in data traffic with the growth of smart devices, video-based applications and developments shows no signs of decreasing. The first step towards solving the capacity crunch problem is to explore the different modulation formats to obtain higher spectral efficiency. Going any further would in theory require the use of super-channels that include multiple sub-bands beyond the 50 GHz grid [10]. Even after so much advancement, the data transfer capacity remains limited, 400 Gb/s and 1 Tb/s running on legacy systems will be inadequate to bridge distances greater than 2,000 km in long links with high spectral efficiency [4]. These issues lead to a maximum achievable spectral efficiency for 2000 km for a single-mode fiber (fibers allowing a single light signal to propagate) of less than 7 bit/s/Hz [10] - the With traffic volumes increasing by around 40 percent each year, it hints that optical networks face a capacity crunch unless a step-up in technology can be achieved [4]. There are many ways to increase optical transmission system capacity over a fixed bandwidth, and most of them are already being used, including modulation using different amplitude levels, two orthogonal subcarriers (cosine and sine modulation) and polarization [11]. Space-division multiplexing has been proposed as a solution to enhance the capacity minimizing the cost per bit of fiber-optic transmission [11] [12].

1.3 Problem Statement & Solution Outline

Earlier Wavelength Division Multiplexing has been used for nearly a decade to come with communication schemes when one or more components failed in traditional networks when a disaster occurs. But with the advent of cloud computing, people now can access similar applications through the Internet irrespective of their physical locations. Cloud can be said to be a virtual data repository. It refers to accessing computer, information technology (IT), and software applications through a network connection, often by accessing data centers using wide area networking (WAN) or Internet connectivity [13]. Cloud computing increases productivity, helps increase cash flow and contributes to many more benefits such as flexibility, disaster recovery, security and likewise. Hence we can say technology has almost migrated to clouds now. With this it has become necessary to maintain all the data in the cloud somewhere, so that it is accessible to anyone anywhere as promised. Thus we came up with the concept of data center (DC) which can be defined as a centralized repository which possibly maybe physical or virtual for maintenance of data [14]. Approaches had to be modified from WDM when it started involving DCs as it takes into account multiple files stored in multiple locations. Hence, it has drifted to Elastic optical networking which allows each communication to dynamically adjust its resources (e.g., the optical bandwidth and modulation format), depending on bandwidth requirements and transmission characteristics for the communication [15]. Most of the existing work on DCs focuses on static lightpath allocation where the lightpaths are setup before the network starts operating. However presently it is expected that lightpaths should be created on user demand and resources should be restored after the communication is over, so that we can reuse the same resources for a different communication now. In our thesis we extend this approach to SDM where we allocate bandwidths for subcarriers or contiguous optical carriers (OCs), each using a certain modulation format and carrying a fraction of the aggregated traffic in

a fiber instead of lightpaths [15]. We consider a dynamic scenarios where, in response to each request for communication, the corresponding communication scheme is determined. Also, our approach takes data replication into account as a DC network must store multiple copies of each file in the network, so that at least one copy of each file is guaranteed to be available to handle any request for that file, even after a disaster happens. Not much progress has been made in this field earlier.

1.4 Scope of thesis

In the optical networking group, Dr. Al Mamoori has been investigating problems on SDM for DC networks. She has proposed an Integer linear programming (ILP) as well as a heuristic method to solve the resilient SDM data center network design problem for dynamic traffic [16]. In her thesis, she presented her experiments in for the heuristic approach but the performance of the optimal solution has not been studied yet. In this thesis, we would focus on the optimal results and also study the performance comparison with the heuristic approach.

1.5 Structure of thesis

The remaining thesis is organized as follows. In Chapter 2, we have reviewed basic concepts of optical networks based on SDM, the notion of disaster resilient techniques and data center networks. The algorithm for the data center network design is given in the form of an Integer Linear Program in the thesis of Dr. Al Mamoori. We have presented our proposed approach in Chapter 3. In order to make this thesis self contained, a detailed explanation of the ILP formulation is also given in Chapter 3. Chapter 4 is the primary contribution of this thesis and describes the implementation details of our approach, the simulation results, with critical comments. Chapter 5 provides the summary concluding

the thesis along with the directions of possible future work.

Chapter 2

REVIEW OF RELATED TOPICS

2.1 Fundamentals of Optical Networks

Over the past score of years, networks are known to have evolved from being relatively static, having fairly homogeneous traffic to being more configurable and carrying a heterogeneous number of services. Computer communication had started with copper wires as the medium that carry electrical signals, encoding the data to be transmitted from one device to another. Copper had many limitations as a medium of communication (e.g., most of the energy used to get wasted in the form of heat [12]). The tremendous increase in the need to communicate large volumes of data over the years increased the need of high speed networks with huge bandwidth. In the last two decades, enormous advancements have been made for using alternative media for communication. Hence, an optical network can be defined as a technological arrangement of signals that connects more than a single computer or any other devices which can generate or store data in electronic format using optical fibers [12]. It consists of a number of nodes which are interconnected using optical fibers and possesses the ability to carry out communication across the network using optical signals. It can be used for both local area networks as well as wide area networks. It can carry large amounts of data at high speed and over long distances. Not only does it provide a higher capacity but is also cost effective, hence can be efficiently used for new applications on the Internet, cloud computing, or other multimedia interaction [2].

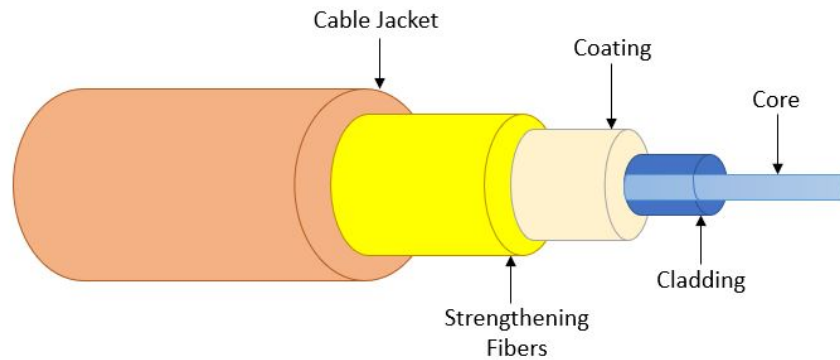


Figure 2.1: Optical Cable

2.1.1 Components of Optical fiber

An optical fiber is made of five parts.

- Core - The first layer in the fiber is the core. It is made up of silica. [12]
- Cladding - The second layer is the protective sheath, it is called cladding. It helps in prevention of data loss as it increases the cores total internal reflection [17].
- Plastic Coating - A plastic coat is wrapped around the fiber to reinforce the core and cladding for the third layer [17].
- Strengthening fibers - For added support, the fourth layer of the fiber cable is made of strengthening fibers.
- Cable jacket - Lastly the outer layer of the fiber is wrapped in a cable jacket to protect against elements [17]. The cladding has a lower refractive index, compared to the core, to keep the light signals inside the core [17]. Optical signals travel down a fiber-optic cable by bouncing repeatedly off the boundary between core and cladding. When an optical signal hits glass at an angle less than the critical angle [18], it reflects back in again due to total internal reflection. A very minimal loss during transmission allows Optical Fibers to

transmit light or data quickly over long distances.

2.1.2 Total Internal Reflection

Total internal reflection occurs when a propagating light wave strikes a surface or a medium at an angle which is larger than a particular critical angle with respect to the normal to the surface. The index of refraction of an optical material or glass is a measure of the speed of the light in the material, and any change in this index of refraction causes light to bend [18]. Refraction causes light to be reflected from the surface after a certain angle; hence this reflection is used by the optical fiber to trap light in the core of the fiber, which has a proper index of refraction.

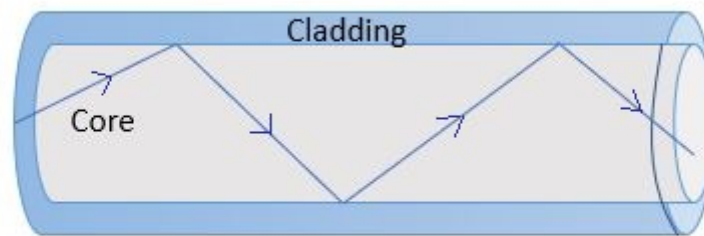


Figure 2.2: Total Internal Reflection

2.2 Cloud Computing

With significant advancements in Information and Communications Technology (ICT) over the last half-century, our generation seems to have migrated to cloud computing. Cloud computing is, in fact, being transformed into a model consisting of services that

are commoditized and delivered like traditional utilities such as water, electricity, gas, and telephony [18]. Clouds are nothing but remotely hosted servers on the Internet. We can access any resource or data irrespective of location, time or device if at all we are connected to the Internet [19]. It is a high capacity Internet-based computing which requires large data centers to store data. Cloud computing is a techno-scientifically distributed computing model; it differs from a traditional one as 1) it is largely scalable, 2) it can be encapsulated as an abstract entity that delivers various levels of services to customers outside the Cloud, 3) it is driven by markets of scale [44], and 4) the services can be dynamically configured (via virtualization or other approaches) and delivered on demand. Lately, industry leaders, research institutes, also Government are up to adopting Cloud Computing to solve their ever increasing computing and storage problems arising in the Internet Age [20].

2.3 Data Center

Presently, it is standard practice to reach content across the Internet autonomously without reference to the underlying hosting infrastructure of the Internet. This foundation consists of data centers that are observed, controlled and maintained around the clock by content providers. Providers such as IBM, Amazon, Microsoft, Google, Salesforce, have started to establish new data centers for hosting Cloud computing applications in several locations across the world to provide redundancy and secure reliability in case of disasters or site failures. A data center serves as a central computing/data storage resource to handle an organizations equipment, data resources, IT operations, storage and distribution of its data. Data centers house a networks most critical systems and are essential to the continuation of daily actions. Hence, the security and reliability of data centers and their information are of a prime priority to organizations [21].

The information we access over the Internet or say information we share over Internet, taking social networking sites like Facebook or Twitter as a common example, these data

are stored somewhere in the cloud. Organizations use data centers for the same purpose. Along with that, IT operations can be characterized as a crucial aspect in most organizational procedures around the world. Business continuity of a company relies on their respective information system department to run their operations [1]. To save a company's operations from catastrophic failure, it is necessary to provide a reliable IT infrastructure. Also, information security of a company is one their important concern, hence a data center also has to offer a safe environment which minimizes the chances of a security breach [1].



Figure 2.3: Data Center Server [1]

2.4 Disaster

Any accident or a natural catastrophe which causes major damage or interrupts the normal flow of activities is called a disaster. It may occur over a small area or can spread over a relatively larger area. In the context of optical networks, we might want to describe it as any natural or man-made phenomenon which disrupts the data flow in an information system. Examples of natural disasters may be an earthquake or tsunami, fiber cuts during

earth moving operations are disasters to be anticipated.

2.4.1 Types of failures during a disaster

- Link Failure- A single link goes down at a time (fiber cut)

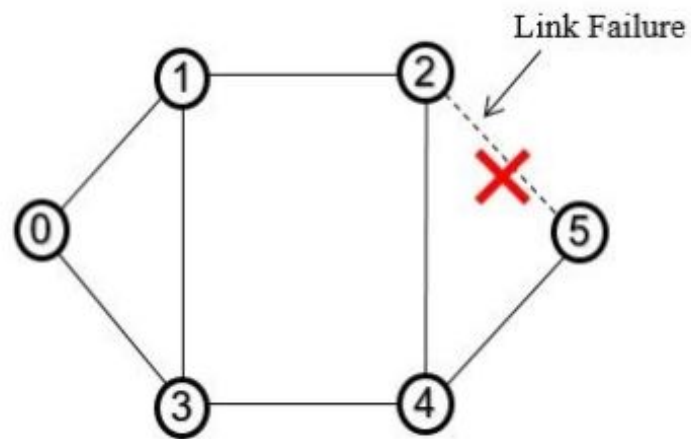


Figure 2.4: Link failure between node 2 and node 5

- Node Failure- All links to/from node go down simultaneously.

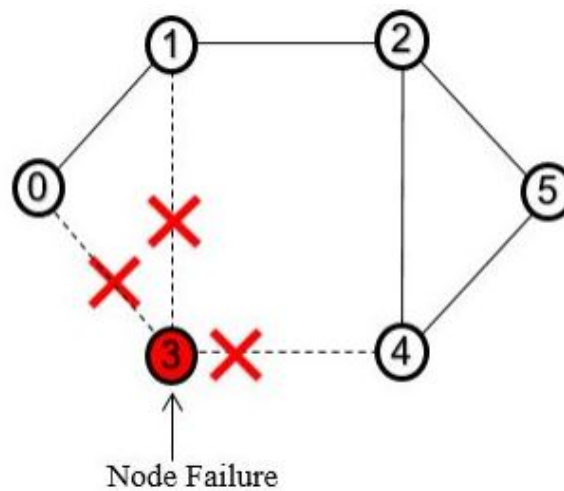


Figure 2.5: All the links connected to node 3 fails

2.5 Faults in Optical Network

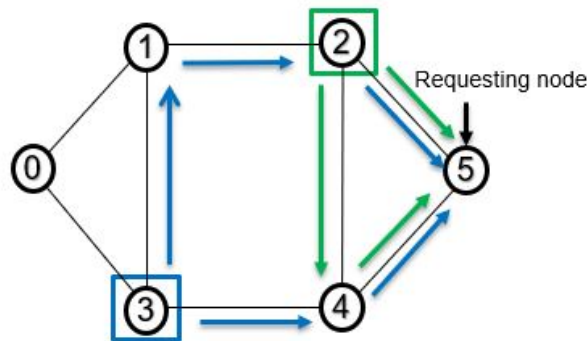
The Internet has become associated with all aspects of modern life, and the significance of a network interruption is considered a serious problem. It is observed that the Internet is not sufficiently resilient and survivable, and thus significant research is in progress to improve the situation. The objective of setting up a resilient network is to decrease the probability of a fault occurrence which might lead to failure in communicating data. This may be done by reducing the impact of an adverse event on network services. These defenses can be identified by developing and analyzing the threats. Therefore, the main techniques for designing disaster-resilient optical networks are to provide geographically diverse multiple paths for communication, so that the network can continue communication avoiding the parts of the network affected by a disaster. The next criterion for successfully designing a disaster-resilient optical network is to detect an adverse event or condition when it occurs. In this regard, the individual components such as routers can detect disasters and can determine when and where the defense mechanisms have failed. There are different ways to determine if there is a fault in the network. In the case of a disaster, an alternate path, called the backup path, can be used to overcome the effects of the disaster.

2.6 File Replication

The concept of saving multiple copies of files at various data centers in a cloud is called file replication. It is typically measured in Recovery Time Objective (RTO) and Recovery Point Objective (RPO). There might be various reasons why we might need to have replicas of the files in the data center. As example, there might occur a scenario where the shortest distance between two locations is greater than the maximum possible length of the cable. In

such scenarios we might want to fetch the data from the nearest data center. Other relevant reason can be the occurrence of a disaster (e.g., an earthquake or a landslide). To recover from a disaster, our primary responsibility is to restore the technological environment back to an operational level within a designated time frame [22]. This time frame is generally pre-determined by the requirements of the company or the institute [22]. Data loss depends on the time it takes to restore and reestablish the operational environment [22]. Earlier data replication was done by keeping a copy or backup of the data to remote sites. Whereas, data replication now refers more to almost immediately accessible data stored in multiple locations over the cloud. For large scale replication, it is mainly done for cloud based systems using data centers. [22].

- Replication strategy when a file f is requested in a fault free network:



fault-free data communication
 Consider DC at node 2 & 3 here

Figure 2.6: Communication in fault free network

There are multiple paths a route may select for successful data transmission. If we consider the above network topology in (Fig. 2.6), we see that there more than one paths are available, when data is requested from node 5 (just showing two paths from data center 3 and 2 in the above figure).

- Replication strategy when a file f is requested in a network affected by a disaster:

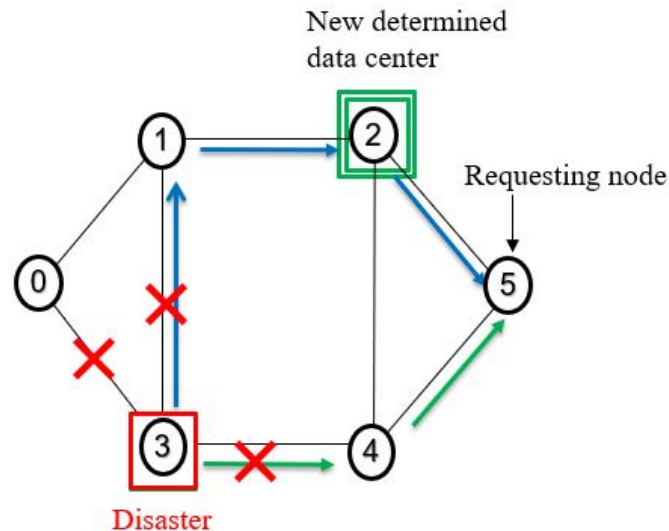


Figure 2.7: All the links connected to node 3 fails

If we consider the above network in (Fig. 2.7), when node 5 requests the file f , it was previously fetched from data center 3. But now if node 3 is affected by a disaster, a backup path from a new data center (which has a copy of file f) may be used.

2.7 Optical OFDM

The ever-increasing growth in data traffic requires more efficient and robust transmission methodology for data communication at more than 100 Gbps. Hence, it becomes necessary to reduce the total bandwidth requirement for data transmission. To achieve that investigators have investigated some technology other than WDM. Orthogonal Frequency-Division Multiplexing (OFDM) has appeared as a promising alternative to WDM. The main reason is that, with OFDM, the elastic nature of the technology means that each user will be al-

lotted exactly as much bandwidth he/she requires. OFDM-based spectrum-sliced elastic optical path network (SLICE) has higher spectrum efficiency compared to WDM because of the granularity of the sub-carrier frequencies. OFDM works for communication requests, which requires a larger bandwidth due to its elastic bandwidth allocation property and possess a better spectral efficiency and impairment tolerance [4]. In this technique, the spectrum on a fiber is split into a number of small bandwidths called the subcarriers. Each bandwidth is maintained by a modulated subcarrier. Each subcarrier carries a relatively low data rate signal that modulates the signal [23]. The number of subcarriers gets allocated according to the request. More than one contiguous subcarriers gets allocated when a connection demands a capacity larger than a single OFDM subcarrier. In this way, only the part of the available bandwidth on each fiber is used, as opposed to the WDM technique. Unlike in WDM technology, where fixed spacing is ensured between two carrier wavelengths to avoid interference within two signals, in OFDM technology, subcarriers overlap with one another. Due to the orthogonality principle of OFDM, the signals do not interfere with each other.

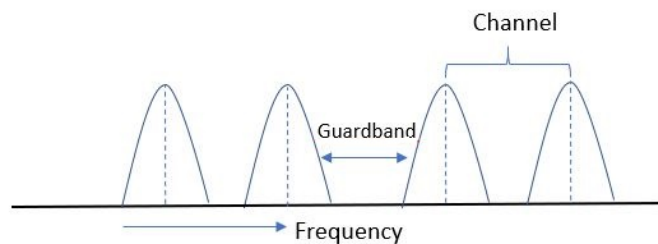


Figure 2.8: Spectrum of WDM

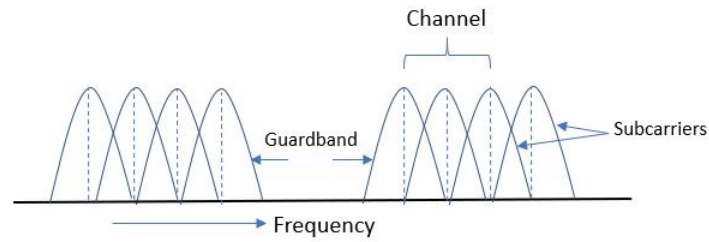


Figure 2.9: Spectrum of OFDM

2.8 Orthogonality Principle in OFDM

We see that in OFDM technology, subcarriers overlap with each other. The information on a specific subcarrier, however, does not get overlapped up with the other subcarrier, even though these subcarriers are overlapping. As shown in the figure 5 when a subcarrier is considered at its peak, other subcarriers have zero crossing at that particular point. Thus, information is only being read when the subcarrier is at its peak level. This certainly avoids the interference of two signals. The use of orthogonal subcarriers allows usage of more subcarriers per bandwidth which in turn increases the spectral efficiency.

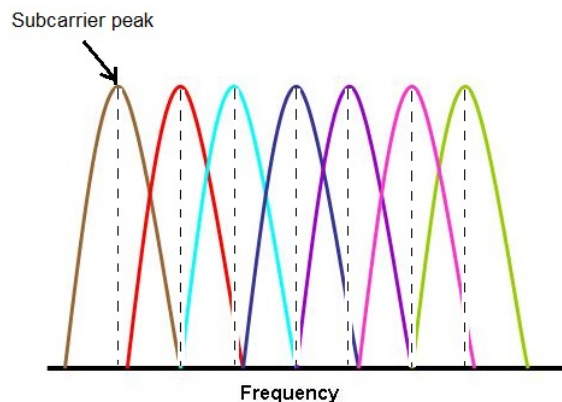


Figure 2.10: Orthogonality in OFDM

2.9 Routing and Spectrum Allocation

For a given network topology, a route and a spectrum has to be allocated to each communication request which include a source for the communication, the destination for the communication and, on each fiber in the path, an unallocated bandwidth sufficient to accommodate the required number of subcarriers for the communication. This process of determining the route and the spectrum for the communication request is known as the Routing and Spectrum allocation (RSA). The main purpose of RSA is to efficiently find an appropriate path and assign a spectrum for the communication request, such that no two requests are allotted the same bandwidth where one or more edges are being shared. Also, each communication request should maintain the same bandwidth. The objective of RSA is to maximize the number of connections on the network by minimizing the usage of frequency spectrum. RSA mainly consists of three constraints:

- **Spectrum Continuity Constraint:** According to this constraint, the spectrum assigned to each lightpath must be the same throughout the route. Thus the starting frequency/bandwidth should be same for each request from the source to the destination on all the fibers to be used in the lightpath.
- **Spectrum Clash Constraint:** This constraint indicates that any two lightpaths that share a common optical fiber must be allocated with non-overlapping spectrum or bandwidth that is separated by at least one guard band.
- **Spectrum Contiguity Constraint:** This constraint assures that the subcarriers allocated are contiguous in the spectrum [24].

2.9.1 Spectrum Allocation in OFDM Networks

- **Static Spectrum Allocation:** In this method of allocation the communication requests are known in advance. Once the connection is set up, it continues for a long time. It is not concerned about any request that is made after the network is deployed and spectral resources, once allotted, are never freed up.
- **Dynamic Spectrum Allocation:** In this method of allocation, new route selection and bandwidth allocation takes place in response to the user requests as they are received by the system. Once a communication is over, resources used for the communication are no more reserved, that is, they are freed up, so that it can be utilized for other requests.

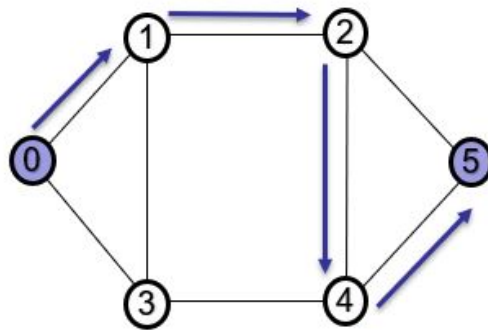


Figure 2.11: Six node network topology

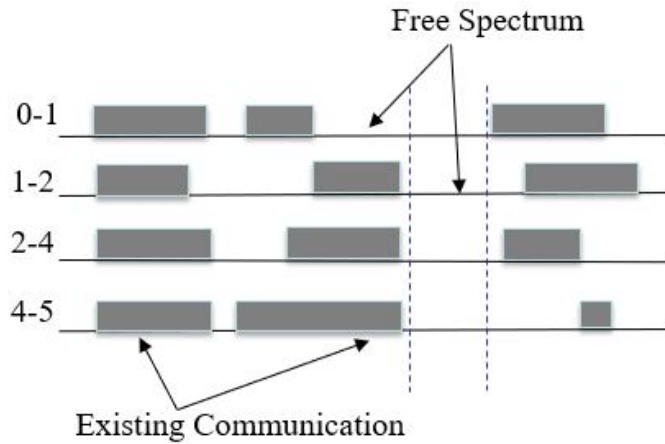


Figure 2.12: Existing communication

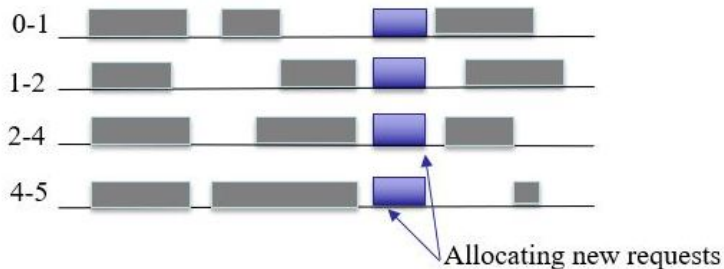


Figure 2.13: Allocation of new resources

2.10 Space Division Multiplexing

Common techniques used in optical communication are Wavelength-division multiplexing (WDM), Time-division multiplexing (TDM), Polarization-division multiplexing (PDM) but these have almost reached their scalability limits now. In order to satisfy the exponential growth of data traffic, optical networks have come up with the strategy of exploiting one remaining unused dimension, which is space. Increasing popularity and variety of bandwidth-demanding applications lead to incremental exhaustion of available spectral resources in currently networks. Space Division Multiplexing (SDM) architectures have

appeared as an up-and-coming technology for overcoming bandwidth limitations. Space Division Multiplexing (SDM) technology has attracted a lot of attention due to its superior advantages for optical communication. One of them being its greater data transmission capacity as it can transfer larger quantity of information through a single optical fiber making it both cost and time effective [25]. The difference from other ways of multiplexing and in SDM is that, a single fiber is replaced with multiple fibers and is used in parallel data transmission. As an example in the topology below, we have considered data to be flowing from node 4 to 5, earlier we considered bidirectional edges for data flow, but now in SDM multiple cores (four cores as shown in figure 2.15) are used to carry data from the source (node 4) to the destination (node 5). This enables larger data transmission using the same resource simultaneously.

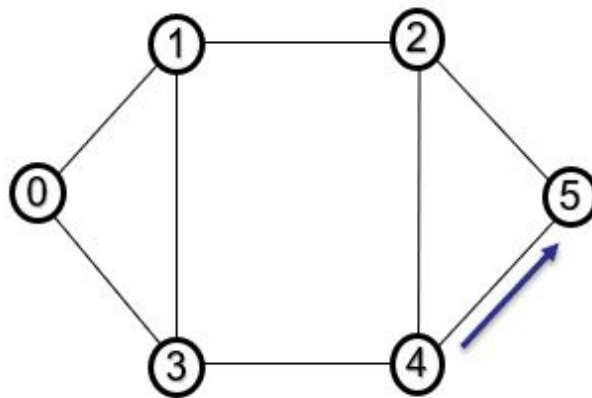


Figure 2.14: Six node network topology

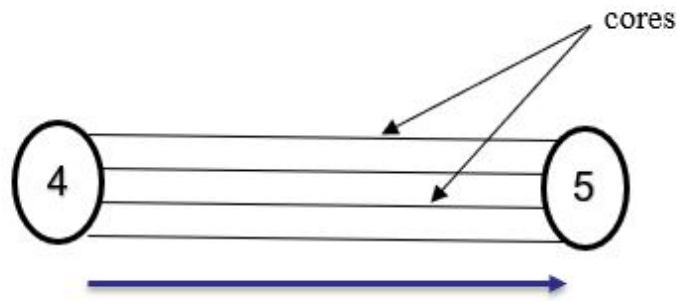


Figure 2.15: One super-channel id divided into multiple cores in SDM

2.10.1 Types of Fibers used in SDM

SDM supports parallel transmission of optical signals propagating simultaneously for a communication request. Numerous technologies have been introduced for SDM, as follows:

- **Single Mode Fiber:** Single Mode fiber optic cable the diametrical core is relatively small. It allows only a single mode of light to propagate. Hence, as the light passes through the core, the number of light reflections gets limited, lowering the attenuation and creating the ability for the signal to travel further. We can thus use SMF for long distance data transmission [26].

Single-mode Fiber



Figure 2.16: Single mode fiber

- **Multi Core Fiber:** This involves multiple single mode cores bundled in the same cladding. one-ring, linear-array, two-pitched, dual-ring, and hexagonal close-packed structure are examples of few types of core placement structures which has been proposed. One disadvantage of using MCF is that since multiple cores are bundled together, signal impairments might occur due to crosstalk between adjacent cores whenever signals are transmitted [9].

Multicore Fiber



Figure 2.17: Multi core fiber

- **Multi mode or Few-Mode Fiber (MMF/FMF):** We find that multi mode fiber optic cable has a relatively larger diametrical core which allows multiple modes of optical signals to propagate. Due to this, the number of light reflections created increases as the light passes through the core, thus allowing more data to pass through simultaneously. But the quality of the signal is reduced over long distances because of the high attenuation and dispersion rate with this type of fiber. These fibers can be used for short distance, data transmission or audio/video purposes in LANs [26].

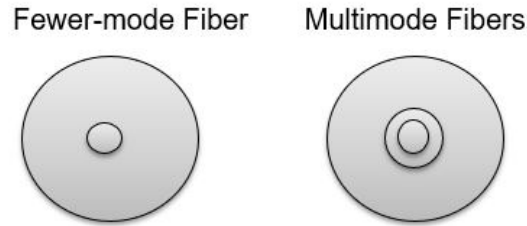


Figure 2.18: Fewer-mode and Multi-mode fiber

2.10.2 Ways of realizing SDM Transmission

SDM allows multi-carrier data transmission. Communication in SDM uses high-capacity super-channels consisting of multiple optical carriers on multiple fibers (in multi-fiber SDM) or multi-mode fibers. Each super-channel carries signals of a specific modulation format and carries a part of the data traffic. In SDM, the optical carriers can be transmitted using both spectral and spatial resources (SpRcs). Hence the super-channels can be formed in both frequency and spatial domain using the best resources. There are a few models which differ in the grade of spectral and spatial flexibility and are used in various scenarios of data transmission [9].

- Fixed-grid/Single - It is similar to WDM network with an expansion in the spatial domain. In this type of model, the spectrum is divided into fixed frequency grids, and single-carrier transmission is utilized.
- Flex-grid/Single - An additional flexibility is introduced in the spectral domain by means of a flexible grid. The optical carriers can now occupy flexibly allocated segments of spectrum which might vary to suit the user requests [9].
- Flex-grid/Fixed - In this model, the super-channels are extended in the spatial domain. Here, within the same spectrum segment, the OCs belonging to a super-channel can be transmitted using different SpRcs. The advantage of this

model is that the super-channels cannot overlap in the spectrum domain. The spectrum is split into fixed grids where the OCs can be flexibly allocated the spectrum requested by the user. The same bandwidth allocation throughout the channels should be maintained in this case. This model is mainly used in MMF/FMF systems [9].

- Flex-grid/Semi-flexible - In this model, the super-channels are further extended in the spatial domain by introduction of spatial groups. Here, SpRcs are divided into some groups and only one super-channel can be carried through each group within given spectrum window [9].
- Flex-grid/Flexible - In this model, full spectral and spatial flexibility in forming super-channels is considered. [9]. Although this scenario enables best resource utilization theoretically, it may lead to fragmentation of spectrum that is residual of unused gaps instead of fixed spacing which is generally not considered for allocation of new requests. Therefore, implementation of this model is definitely possible but also is challenging from the perspective of efficient demand allocation.

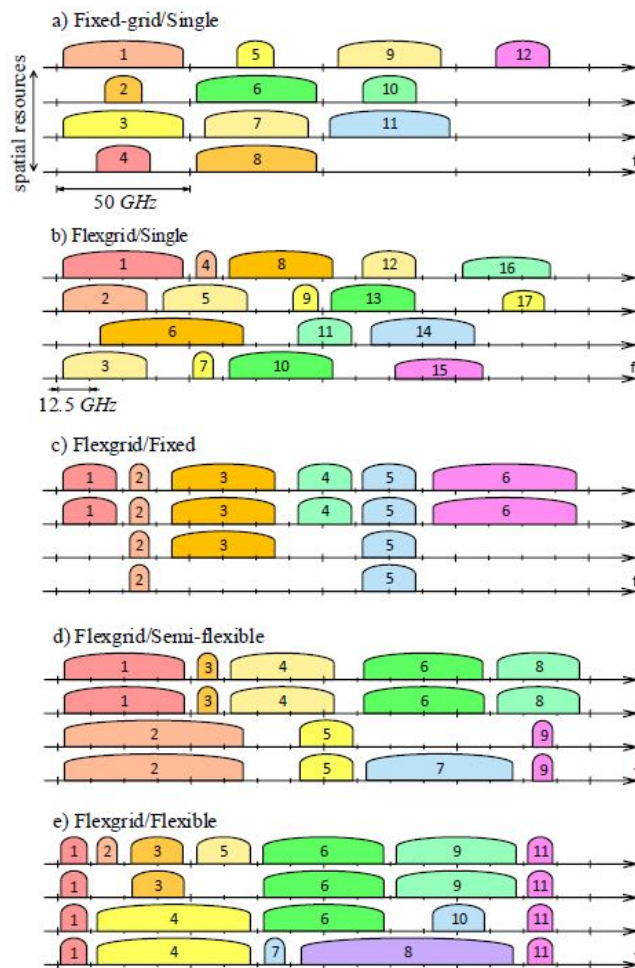


Figure 2.19: Models used in SDM [9]

2.11 Optical Reach

When an optical signal travels over long distances through optical fibers, the quality of the optical signal starts degrading after it has progressed a certain distance. Hence we can define optical reach as the distance an optical signal can travel before the quality of the light signal degrades to such a level that it becomes necessary to regenerate it. Various factors affect the optical reach in a network, like, the type of modulation format of the

signal, amplification of the signal, it's launched power, etc. The longer the optical reach of a signal, the less regeneration it requires over the network, hence serving the purpose of using less equipment thus lowering operating costs [27]. However, expensive equipment, such as transponders and amplifiers are required for achieving longer optical reach [27].

In (Fig. 2.20), if we consider a light path from node 0 to node 5, the length/optical reach of the light path is 1700 kms.

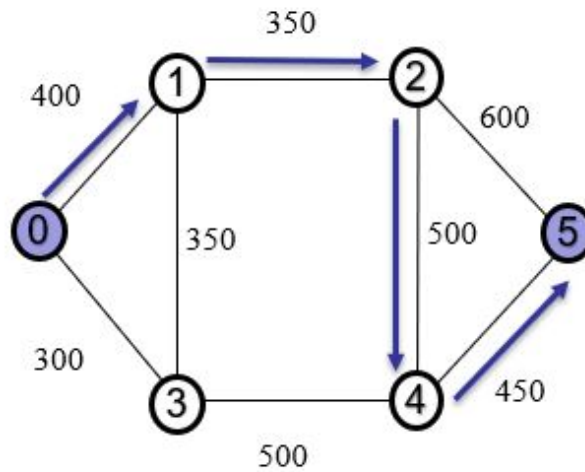


Figure 2.20:

2.12 Modulation Format

Optical fiber provides higher bit rate in long-distance transmission [28]. This can be further improved if we utilize the advanced modulation formats. Optical-fiber networks can transmit the Tb/s bits over few thousands of kilometers. It has been observed that an optical fiber networking system can have an attenuation coefficient 0.2db/km for many T Hz of bandwidth, exceeding the transmission distance than 10,000km and holding a capacity of more than 10Tb/s [29]. Advanced modulation formats also improve the channel utilization and capacity. Better performance of optical fiber can be achieved by using different

modulation formats. One of the reasons why using flex-grid is advantageous is its distance-adaptive nature. In flex-grid networks, the modulation format for each connection can be selected individually during the process of allocation. Distance-adaptive concepts add a new trade-off between optical reach and spectral efficiency. For example, modulations like BPSK do have a long optical reach, but at a cost of a low spectral efficiency. QPSK provides better spectral efficiency than BPSK. Again, other modulations like 16-QAM transmit the same bit rate occupying almost approximately one-fourth of the bandwidth of BPSK, but with a significant shorter optical reach.

From the above topology in (Fig. 2.20) where the length of the path is 1700 km, we see from the modulation format table that 8-QAM is the best suitable format which can be applied to the network.

Modulation Format	Optical Reach (km)	Spectral Efficiency(bp s/Hz)(Mp)
BPSK	9600	1
QPSK	4800	2
8-QAM	2400	3
16-QAM	1200	4

Figure 2.21: Modulation formats [30]

2.13 Literature Review

Krzysztof Walkowiak et al. in [9] discusses integer linear programming(ILP) optimization models for SDM optical networks. The proposed models exhibit different ways of data transmission through SDM, each characterized by different flexibility in the use of spatial and spectral resources. They have also analyzed the details of the ILP models describing the technological aspects of SDM. Their work is concerned with generic ILP formulations that can be used to model various versions of SDM which can lead to the discovery of different ways of using the spatial domain available in SDM networks. The author describes the types of fibers we can use in SDM networks for parallel transmission of optical signals and briefly discusses the concept of super-channels. When a high-capacity super-channel is transmitted over the network, it might consist of some optical carriers (OCs), which get generated or terminated using devices called transceiver, using a particular modulation format and carrying a portion of the flowing traffic. The OCs in SDM can be transmitted using both spectral and spatial resources (SpRcs). Authors have focused on channel based modeling in this paper and described a few models through which SDM transmission can be realized, namely Fixed-grid single, Flex-grid single, Flex-grid fixed, Flex-grid semi flexible, Flex-grid/flexible. They further formulate three ILP models and apply it to three SDM scenarios, the first being Flex grid/Flexible. The main objective was to minimize the overall spectrum usage of the network to provision the demands, considering all the links and SpRcs. The next considered model uses a Flex grid/Single scenario when each request is designated to at most one SpRc on each link. Although the solution estimated by solving this model satisfied all the constraints, it failed to address one major issue introduced by SDM, i.e., Light path-to-Spatial-Resource (LtoSR) assignment. The last model discusses Flex grid/Fixed scenario when all the SpRcs (fibers, cores, modes) are considered as one

single entity while having the freedom of switching spectral slices freely. They compared the optimal solution given by the three ILP models and concluded that the first two did not seem to be scalable when solved directly using exact methods, so they might as well come up with a heuristic solution of it in future.

The paper [30] concentrates on a distance-adaptive scenario when the same connection request can be transmitted with different modulation formats and is associated with different spectral efficiencies and optical reaches. As spectrum fragmentation degrades the performance (in terms of blocking probability), in flex-grid networks the modulation format can be decided independently for each connection during the resource allocation process. For example, modulations like BPSK, have longer optical reach, but have low spectral efficiency, whereas 16-QAM transmits same data using one-fourth of the bandwidth. The paper focuses on the allocation of dynamic connection requests when they are received, deciding its route, modulation, and the band to be allotted through Routing and Spectrum Assignment (RSA). Zaragoza et al. reviews existing RSA proposals in the paper which apply to heterogeneous flex-grid networks, evaluating their blocking performance averaged among services, their fairness in balancing the blocking probabilities observed by different services. Simulation study has been performed to analyze the performance of two different RSA algorithms namely blocking model and incremental model. The authors also propose Partial-Sharing-Partitioning (PSP), that finds a compromise between network capacity and fairness. The Partial-Sharing-Partitioning (PSP) scheme manages the spectrum in heterogeneous flex-grid network. When a request is sent, the PSP controller allocates it into the dedicated partition for service. If infeasible, it tries to allocate it into the shared partition. The request gets blocked if it fails to do so. In the paper, the authors compare the network capacity and fairness performances for a set of previously proposed algorithms, revealing their merits in both dimensions. Results show that PSP algorithm performs better for optimizing the network capacity, enforcing fairness among services.

In the paper [31] Muhammad et al. discusses that for encompassing the capacity limitation of the networks based on single-core optical fiber (SCF) and for attaining far higher transmission throughput and spectral efficiency, research community researchers have come up with the idea of exploiting the only remaining unused dimension, i.e., space. For assigning of spectrum slices by the traffic demands in flex-grid networks, the well-known routing and wavelength assignment (RWA) problem is modified to the routing and spectrum allocation (RSA) problem. The flex-grid networks use multi-core fibers for spectrum and core allocation during traffic demands. Due to the spectrum non-overlapping constraint in flex-grid different traffic demands with common spectrum slices cannot traverse through the same network link. Although, demands can be routed through the same link but different cores if they share some common spectrum slices [31]. The paper also investigates the routing, spectrum and core allocation (RSCA) problem for flex-grid optical networks formulates RSCA network planning problem using ILP. For a RSCA problem, all the traffic demands in the network are known before i.e the requests are not dynamic. The authors give an optimal solution for provisioning the demands through proper allocation of spectrum and core, while efficiently utilizing all the spectrum resources. The optimal solution evaluates the number of spectrum slices required to serve a given traffic demand, and also fulfill all the other constraints, such as inter-core crosstalk and spectrum overlapping simultaneously [31]. The paper gives a brief sketch of the crosstalk issue for networks with MCF. To incorporate the inter-core crosstalk in the ILP formulation two different approaches were proposed. First, slices with same spectrum cannot be assigned to different demands that transmit through neighboring cores unless the crosstalk level at the receiver end is beneath a given threshold. The other approach is to pre-compute the crosstalk values for all the paths and choose a route with the value of crosstalk below a given threshold [31]. A scalable and effective heuristic is proposed for the same problem. After observing the solutions, they concluded that their proposed ILP model is flexible for small networks and low traffic

but bigger topologies with high traffic demands efficient solutions within optimal time can be obtained if heuristic algorithms are used. The applied heuristic strategy distributes the traffic demands according to the required spectrum and path. To establish a new request, the spectrum resources are gradually increased after every iteration until the demands are met in the network. The results conclude that the proposed heuristic can generate a solution which might be close to optimal solution in polynomial time [31].

The paper [15] gives an overview of the latest developments and possible approaches concerning flexible optical networking and the observed advantages that spatially flexible networking approaches can contribute to optical networks. It refers to the capacity of a network to dynamically adjust its resources, such as the modulation format, or optical bandwidth according to the required bandwidth and transmission characteristics of each requested connection. Klouidis et al. discusses the different channel allocation options over the available fiber dimensions which are its wavelength, bandwidth or space approved by the latest relevant technology advances and research efforts. Next, the possible flexible networking approaches considering both spectral and spatial network flexibility are commented on and discussed, identifying their benefits, limitations, and synergies [15]. The author says that the pure SDM approach of placing a spatial demultiplexer is prior to a spectral demultiplexer to process and extract the spatially multiplexed data can also be applied to the spectrally flexible networking systems. The last part focuses on network design, planning, operation, and control issues, addressing the latest advancements in flexible optical networking, while also highlighting the primary research directions and required modifications for the introduction of flexibility in the space dimension [15].

Wang et al. in [32] discusses that the Orthogonal Frequency Division Multiplexing (OFDM) technology enables the elastic and flexible bandwidth allocation in the SLICE network [33]. When a user request requires multiple sub-carriers, consecutive sub-carriers in the spectrum domain are allocated and overlapped without using the spectral gap or the

guard-band frequencies [32]. Similar to RWA in WDM networks, the SLICE network deploys the routing and spectrum allocation (RSA) process to serve the traffic demands [32]. The objective of the paper [32] is to study an optimal solution and also propose new approaches for the analysis of the number of subcarriers used in a general mesh network. The authors present Integer Linear Programming (ILP) formulations for optimal RSA with two optimizations goals. The first being, minimizing the maximum subcarrier index among all the fibers, and the second being the minimization of the total allocated sub-carriers over all the fibers, keeping five constraints in mind, namely, the traffic demand constraint and Sub-carrier capacity constraint, Spectrum Continuity Constraint, Guard-Carrier Constraint, Sub-carrier Consecutiveness Constraint [32]. They also analyze the upper and lower bounds for the maximum subcarrier index in a SLICE network. These methods improve the bounds for the case with predetermined routing knowledge. Moreover, the simulations presented in [33] review two other efficient heuristic algorithms, namely the BLSA or the balanced load spectrum allocation and SPSR or the shortest path with maximum spectrum reuse under different optimization goals [32].

Chapter 3

OPTIMAL RSA FOR SDM OPTICAL NETWORK UNDER DYNAMIC TRAFFIC

This chapter gives a brief overview of the optimum algorithm for handling a request for communication in a data center network using SDM. As explained in the introduction, the algorithm presented here appears in [16] and is given here for completeness.

- Introduction to the problem.
- Aim of our problem.
- Assumptions made.
- Idea of Virtual Node.
- Concept of Gaps.
- An ILP Formulation to solve the problem.
- Justification of the ILP.

3.1 Introduction to the problem

A network consists of nodes (representing source and destinations of data, typically computers or data centers) and edges (representing fibers in the network). A cloud is a network of data centers which handles customer requests irrespective of the location of the user. The problem is to process a request for a file determine a strategy for robust data communication. The data communication strategy must ensure that the requested file may be

communicated to the node requesting the file even if a disaster (such as an earthquake or fire) occurs. For maintaining a robust DC network, it must be ensured that the network stores multiple copies of each file in different nodes of the network, so that at least one copy of each file is available even after a disaster happens, ensuring no interruption of services. In our investigation we are considering the dynamic scenario where a specific communication scheme is determined in response to each request for communication. If we consider a network of n nodes and e bidirectional edges, and set of disasters D , our replication strategy must ensure that, for all the disaster scenarios where $d \in D$, and for all the communication requests possible, it is possible to determine a fault-free path to enable communication. The replication strategy in [16] determines how many copies of a file, (say m copies) is required as well how many data centers ($S_j^1, S_i^2, \dots, S_i^m$) should have a copy of the file f_i [16]. We determine the route for communication both primary communication as well as backup communication when a primary path is affected by any disaster using an integer linear program. In other words, the problem we are considering is to handle a request for a file f_i from a node t . We assume that, before the network is deployed, the replication strategy that file f_i is saved at nodes, say, S_i^j , $1 \leq j \leq m$ has been already determined. The problem is as follows:

For a fault free network :

The communication strategy has to ensure that, for file f_i , it is possible to have a possible path, also called the *primary path* from some node S_i^j to t which may be used for communication using SDM. As we also intend to minimize the resources used, here we consider three modulation formats along with their respective spectral efficiencies. We also take into account the fact that if a modulation format has a higher efficiency, the optical reach is lower. Our objective is to find the shortest viable path for communication, such that the length of the path is less than the optical reach for BPSK format. BPSK format is known to have the longest optical reach (9600 km). We have other modulation formats as

QPSK (4800 km), 16QAM (2400 km) which can handle communication with lesser optical reach [15] in a more efficient manner.

For a network affected by any disaster $d \in D$:

It must be possible to have a fault-free viable path, appropriate for communication using SDM, called a *backup path* which avoids any disaster d that disrupts the primary path, from some node S_i^k to t . Our scheme for the backup path ensures that when a primary path is disrupted by disaster d , irrespective of the disaster, it must be possible to use the same backup path for communication to node t . Since we intend to deal with a large set of communication requests via a cloud connecting to DCs, spectrally and spatially flexible optical networks are likely to be ideal candidates for ultra high speed communication needed for data center networks. We use flex-grid fixed-SDM model for this purpose, since this is technologically viable and yet offers great scope in terms of the choices for modulation format, bandwidth used for communication and space multiplexing. Our proposed scheme uses the idea of dedicated path protection to determine path for backup communication to handle disasters.

3.2 Assumptions made

- We have already determined the replication strategy for our system so that the locations for all the files $f_1, f_2..f_n$ are known. In general, each file will be replicated at several data centers, so that, if there are m copies of file f_i , we know that copies of file f_i are saved at data centers $S_i^1, S_i^2..S_i^m$.
- The network uses OFDM for data communication.
- The network uses flex-grid fixed model for SDM networks.
- The network is supposed to already support a number of on-going communication when a new request for transmitting file f_i to node t is received. The

details of each existing communication are known to us, so that, for each ongoing communication, we have all information about the scheme for fault-free communication and the scheme to handle disaster $d \forall d \in D$.

3.3 The Objective of our algorithm

For our algorithm, we consider the set $\{S_i^1, S_i^2, \dots, S_i^m\}$ as the set of data centers having copies of file f_i , and the set D for disasters that might occur, where each disaster is defined by the set of edges which are adversely affected by the disaster. Our objective is to determine paths for the primary and the backup communication to handle the new request for communication of file f_i to some node t satisfying the following conditions:

- The primary communication will be from some node $S_i^j, 1 \leq j \leq m$, to node t [16].
- The backup communication will be from some node $S_i^l, 1 \leq l \leq m$. Here $j \neq l$, since a disaster can disrupt the source S_i^j used in the primary communication. This scheme for backup communication cannot use any of the edges affected by all disaster $d \in D$ that disrupts the primary communication [16].
- The length of the path used by the primary and backup communication will determine the optimum modulation format to be used by the primary and backup communication.
- Minimize the number of subcarriers needed to carry out the new communication.
- Measure the cost of the resources needed to handle the new request by the sum of the spectrum bandwidths needed for the new communication.

3.4 Idea of Virtual Node

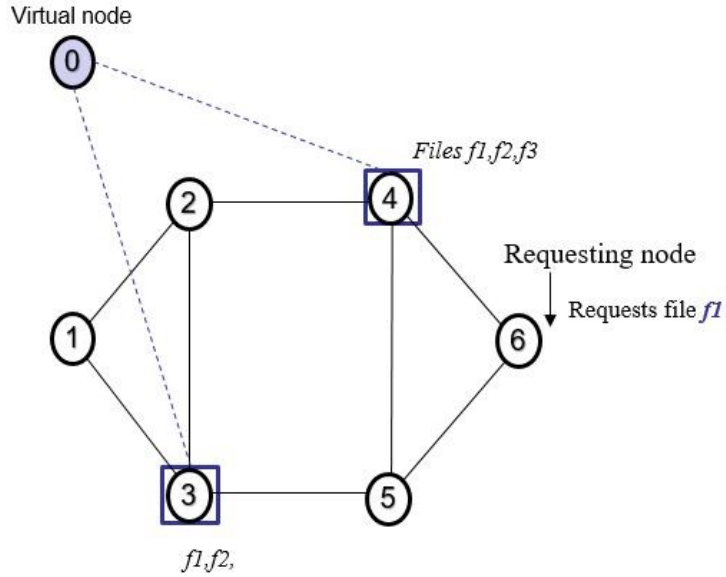


Figure 3.1: Six node network topology with a virtual node at 0

When a file f_i is requested for communication, the source of the communication is not specified. In other words, only file f_i to be retrieved is specified. Our algorithm needs to determine which node will act as the source of the lightpath. This must be one of the nodes $S_i^j; 1 \leq j \leq m$ that stores a copy of the requested file f_i . To help determine which data center node S_i^j (S_i^l will be the source, for the primary (or backup) path, it is convenient to visualize a new *virtual node* s and some new *virtual edges* from s as follows. For each data center $S_i^j; 1 \leq j \leq m$, we visualize a single virtual edge from virtual node s to data center S_i^j of length 0. This virtual node v does not exist as a node in the network. Since all edges from this virtual node has length 0, it does not affect the optical reach constraint. We can now use standard network flow algorithms to find a primary (backup) path from the virtual node s to node t . Simply by deleting the virtual edge $s \rightarrow S_i^j$ ($s \rightarrow S_i^l$), we can get the primary (backup) path from S_i^j (S_i^l) to t .

For example, (Fig. 3.1) below, represents a topology with 6 physical nodes numbered 1, 2, 3, 4, 5 and 6 and a virtual node is assigned the number 0. The virtual node is not a part of the original network but is a concept to enable the use of a simple single commodity network flow algorithm. If nodes 3 and 4 contain a copy of the file to be communicated, Fig. 3.1 shows that virtual node 0 is connected to all the potential data center nodes (node 3 and 4 in Fig. 3.1) in the network which are potential sources for the communication. Our algorithm determines the optimal path from the virtual node to the requesting node based on factors such as the distance from the requesting node, available spectrum on each fiber in the path. We might then get the actual path for communication simply by deleting the virtual edge. For instance if 3 is the first node in the selected path $0 \rightarrow 3 \rightarrow 5 \rightarrow 6$ from node 0 to the requesting node 6, the actual path for communication is that obtained by removing the virtual edge $0 \rightarrow 3$ in Fig. 3.1) and give the path for communication $3 \rightarrow 5 \rightarrow 6$ in Fig. 3.1)

3.5 Concept of Gaps

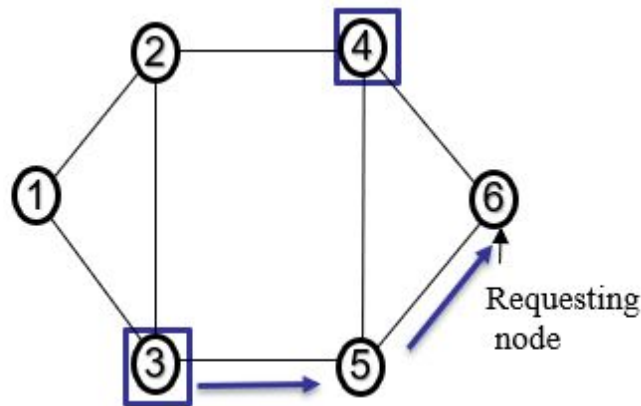


Figure 3.2:

Fig. 3.2 shows a six node network with data centers at node 3 and node 4 storing file f_i , and node 6 is requesting file f_i . The edges $3 \rightarrow 5$ and $5 \rightarrow 6$ indicate that the path

for primary communication is $3 \rightarrow 5 \rightarrow 6$. We are using flex-grid fixed model for SDM networks for allocating bandwidth to the new requested communication.



Figure 3.3: Available spectrum

Fig. 3.3 shows the available spectrum on the node $3 \rightarrow 5$ and node $5 \rightarrow 6$ where the resource allocation should take place for setting up the new communication request.



Figure 3.4: Bandwidth allocation on available spectrum

The grey rectangular boxes in Fig. 3.4 depicts the bandwidth allocation on available spectrum on node $3 \rightarrow 5$ and node $5 \rightarrow 6$.

In SDM, a single link from the source node to destination node represents multiple cores, and each core should have equal bandwidth allocation over the spectrum. Fig. 3.5 shows that the node $3 \rightarrow 5$ further gets divided into multiple cores (4 here) and are now responsible for carrying out the communication. Similarly, node $5 \rightarrow 6$ gets subdivided into four cores for the new request. Equal bandwidth allocation of spectrum takes place in all the cores because of the flex-grid fixed model.

3.6 shows a bundle of fibers, representing an edge $i \rightarrow j$ (edge $3 \rightarrow 5$ in Fig. 3.6). Selected portions of each fiber in this packet are already assigned to existing ongoing communication, which means we cannot use them for the new request, rather we will have to

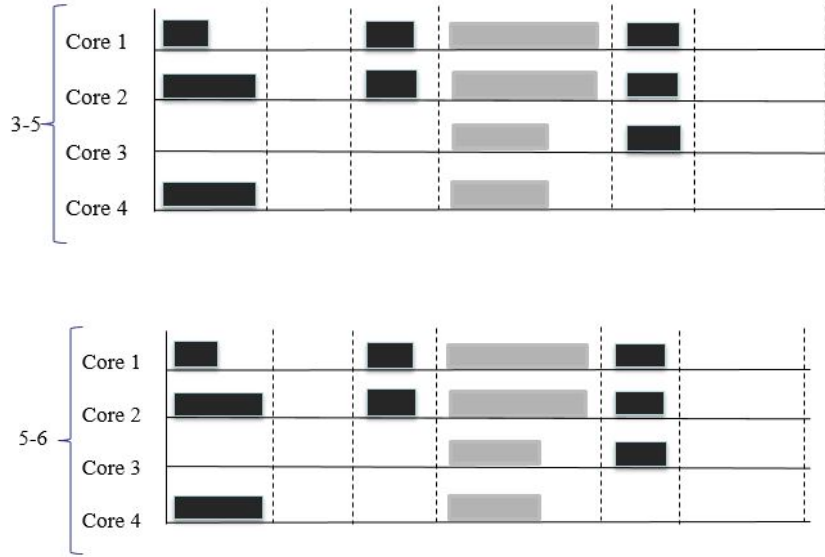


Figure 3.5: Bandwidth allocation on cores of available spectrum

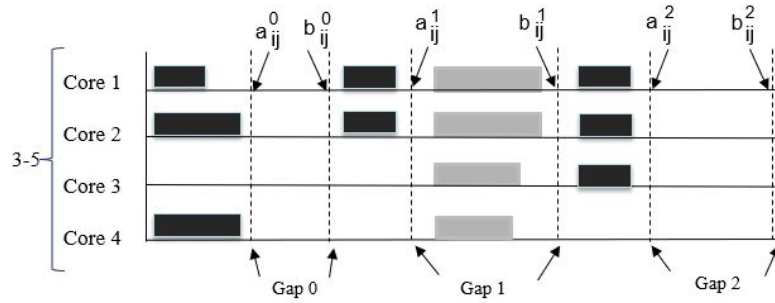


Figure 3.6: Gaps

assign the unused spectrum to the new communication requests. Fig. 3.6 shows a possible allocation of spectrum using the flex-grid fixed-SDM model. The black boxes in the Fig. 3.6 represent spectrum which has already been assigned to some other communication. If a new request for communication uses edge $i \rightarrow j$, the spectrum allotted to the new communication must be within the bandwidths shown in Fig. 3.6. We call these permissible bandwidths the gaps on edge $(i \rightarrow j)$. In the situation described in Fig. 3.6, there are 3 gaps, where gap 0 has a starting subcarrier a_{ij}^0 and ending subcarrier b_{ij}^0 . Similarly gap 1 has a starting subcarrier as a_{ij}^1 and ending subcarrier b_{ij}^1 , gap 2 has a starting subcarrier as a_{ij}^2 and ending subcarrier b_{ij}^2 . When a new communication is requested, a bandwidth which

lies within one of these 3 gaps should be allocated. In Fig. 3.6 we select gap 2 according to the bandwidth requirement of the requested communication.

3.6 Notations used in the ILP

In this section we outline the notations used to formulate the algorithm proposed in [16].

$N(E)$: set of end nodes (bidirectional edges) in the physical network topology.

E^d : set of bidirectional edges (i.e., links) in the network that are disrupted by disaster d .

G_{ij} : set of all gaps on edge $(i, j) \in E$.

C : set of core per fiber.

$s(t)$: source (destination) of the new request for communication. In this problem s is the virtual node as discussed above.

B : a constant denoting the bandwidth requested, measured by the total number of subcarriers needed for communication.

k_0 : a constant denoting the packing density for 16-QAM modulation format.

k_1 : a constant denoting the packing density for 8-QAM modulation format.

k_2 : a constant denoting the packing density for QPSK modulation format.

d_0 : a constant denoting the optical reach for 16-QAM modulation format.

d_1 : a constant denoting the optical reach for 8-QAM modulation format.

d_2 : a constant denoting the optical reach for QPSK modulation format.

M : a large number.

$a_{ij}^g(b_{ij}^g)$: a constant for all gap $g \in G_{ij}$ and edge $(i, j) \in E$ that represents the starting (ending) subcarrier wavelength of the g^{th} spectrum gap on edge (i, j) .

ℓ_{ij} : a constant denoting the length of the edge $(i, j) \in E$.

n^P : number of subcarriers needed for the primary communication.

n^B : number of subcarriers needed for the backup communication.

ℓ^P : a variable denoting the length of the path used by the primary communication scheme.

ℓ^B : a variable denoting the length of the path used by the backup communication scheme.

$\theta(\omega)$: an integer variable representing the starting subcarrier number of the new request for the primary (backup) communication scheme.

$\phi(\psi)$: an integer variable representing the ending subcarrier number of the new request for the primary (backup) communication scheme.

F_0^P : a binary variable with a value 1 if the primary communication uses 16-QAM format; 0 otherwise.

F_1^P : a binary variable with a value 1 if the primary communication uses 8-QAM format; 0 otherwise.

F_2^P : a binary variable with a value 1 if the primary communication uses QPSK format; 0 otherwise.

G_1^P : a binary variable with a value 1 if the primary communication may use 8-QAM format; 0 otherwise.

G_2^P : a binary variable with a value 1 if the primary communication may use QPSK format; 0 otherwise.

F_0^B : a binary variable with a value 1 if the backup communication uses 16-QAM format; 0 otherwise.

F_1^B : a binary variable with a value 1 if the backup communication uses 8-QAM format; 0 otherwise.

F_2^B : a binary variable with a value 1 if the backup communication uses QPSK format; 0 otherwise.

G_1^B : a binary variable with a value 1 if the backup communication uses 8-QAM format; 0 otherwise. This includes the possibility that $F_0^B = 1$

G_2^B : a binary variable with a value 1 if the backup communication uses QPSK format; 0 otherwise. This includes the possibility that $F_0^B = 1$ or $F_1^B = 1$

q^d : a binary variable with a value 1 if disaster d affects the primary communication; 0 otherwise.

w_{ij}^d : a binary variable for all edge $(i, j) \in E$ where $w_{ij}^d = 1$ if edge (i, j) appears in the path used to handle disaster d ; 0 otherwise.

$x_{ij}(y_{ij})$: a binary variable for all edge $(i, j) \in E$ where $x_{ij} = 1$ ($y_{ij} = 1$) if the path used by primary (backup) communication scheme for the new request uses edge (i, j) ; 0 otherwise.

$x_{ij}^g(y_{ij}^g)$: a binary variable for all gap $g \in G$ and edge $(i, j) \in E$ edge where $x_{ij}^g = 1$ if gap g on edge (i, j) is used by the primary (backup) communication scheme; 0 otherwise.

3.7 Formulation of ILP for optimal solution

The ILP given below is proposed in [16] determines the optimal primary and backup communication to communicate file f_i to node t . The ILP determines the path for the primary (backup) communication scheme, such that the total spectrum needed is as low as possible. In the description below set of constraints, the constraint number with sub notation alphabetical a denotes the constraint for primary path, whereas the one's with alphabetical b denotes the constraints for the backup path.

Objective Function

$$\text{minimize}(\phi - \theta) + (\psi - \omega) \quad (3.1)$$

Subject to

1. Flow balance equations must be satisfied by the primary path used for the primary communication (backup path used when disaster d happens, $\forall d \in D$)

$$\sum_{j:(i,j) \in E} x_{ij} - \sum_{j:(j,i) \in E} x_{ji} = \begin{cases} 1, & \text{if } i = s, \\ -1, & \text{if } i = t, \quad \forall i \in N \\ 0, & \text{otherwise.} \end{cases} \quad (3.2)$$

$$\sum_{j:(i,j) \notin E^d} w_{ij}^d - \sum_{j:(j,i) \notin E^d} w_{ji}^d = \begin{cases} q^d, & \text{if } i = s, \\ -q^d, & \text{if } i = t, \quad \forall i \in N \\ 0, & \text{otherwise.} \end{cases} \quad (3.3)$$

2. Set $q^d = 1$ if disaster d disrupts the primary communication. Otherwise q^d must be set to 0.

$$q^d \geq x_{ij} \quad \forall (i, j) \in E^d, \quad d \in D \quad (3.4)$$

$$q^d \leq \sum_{i,j:(i,j) \in E^d} x_{ij} \quad \forall d \in D \quad (3.5)$$

3. Determine the backup path to be used when any disaster interrupts the primary path.

$$y_{ij} \leq (1 - q^d).M + w_{ij}^d \quad \forall (i, j) \in E, \quad d \in D \quad (3.6)$$

$$y_{ij} \geq w_{ij}^d \quad \forall (i, j) \in E, \quad d \in D \quad (3.7)$$

$$y_{ij} \leq \sum_d w_{ij}^d \quad \forall (i, j), \quad d \in D \quad (3.8)$$

4. Find the length of the primary (backup) path.

$$\ell^P = \sum_{i,j:(i,j) \in E} x_{ij} \cdot \ell_{ij} \quad (3.9a)$$

$$\ell^B = \sum_{i,j:(i,j) \in E} y_{ij} \cdot \ell_{ij} \quad (3.9b)$$

5. Determine whether 16-QAM is the best modulation format to be used in primary (backup) communication.

$$M.F_0^P \geq (d_0 - \ell^P) \quad (3.10a)$$

$$M.F_0^B \geq (d_0 - \ell^B) \quad (3.10b)$$

$$M.(1 - F_0^P) \geq (\ell^P - d_0) \quad (3.11a)$$

$$M.(1 - F_0^B) \geq (\ell^B - d_0) \quad (3.11b)$$

6. Find the number of subcarriers needed for primary (backup) communication if 16-QAM is the best modulation format

$$n^P \geq F_0^P \cdot B/k_0 \quad (3.12a)$$

$$n^B \geq F_0^B \cdot B/k_0 \quad (3.12b)$$

$$n^P \leq M(1 - F_0^P) + B/k_0 \quad (3.13a)$$

$$n^B \leq M(1 - F_0^B) + B/k_0 \quad (3.13b)$$

7. Determine whether 8-QAM may be used for primary (backup) communication.

$$M \cdot G_1^P \geq (d_1 - \ell^P) \quad (3.14a)$$

$$M \cdot G_1^B \geq (d_1 - \ell^B) \quad (3.14b)$$

$$M \cdot (1 - G_1^P) \geq (\ell^P - d_1) \quad (3.15a)$$

$$M \cdot (1 - G_1^B) \geq (\ell^B - d_1) \quad (3.15b)$$

8. Determine whether 8-QAM is the best modulation format to be used for primary (backup) communication.

$$F_1^P \leq G_1^P \quad (3.16a)$$

$$F_1^B \leq G_1^B \quad (3.16b)$$

$$F_1^P \leq 1 - F_0^P \quad (3.17a)$$

$$F_1^B \leq 1 - F_0^B \quad (3.17b)$$

$$F_1^P \geq G_1^P - F_0^P \quad (3.18a)$$

$$F_1^B \geq G_1^B - F_0^B \quad (3.18b)$$

9. Find the number of subcarriers needed for primary (backup) communication if 8-QAM is the best modulation format

$$n^P \geq F_1^P \cdot B/k_1 \quad (3.19a)$$

$$n^B \geq F_1^B \cdot B/k_1 \quad (3.19b)$$

$$n^P \leq M(1 - F_1^P) + B/k_1 \quad (3.20a)$$

$$n^B \leq M(1 - F_1^B) + B/k_1 \quad (3.20b)$$

10. Determine whether QPSK may be used for primary (backup) communication.

$$M.G_2^P \geq (d_2 - \ell^P) \quad (3.21a)$$

$$M.G_2^B \geq (d_2 - \ell^B) \quad (3.21b)$$

$$M.(1 - G_2^P) \geq (\ell^P - d_2) \quad (3.22a)$$

$$M.(1 - G_2^B) \geq (\ell^B - d_2) \quad (3.22b)$$

11. Determine whether QPSK is the best modulation format to be used for primary (backup) communication.

$$F_2^P \leq G_2^P \quad (3.23a)$$

$$F_2^B \leq G_2^B \quad (3.23b)$$

$$F_2^P \leq 1 - F_0^P \quad (3.24a)$$

$$F_2^B \leq 1 - F_0^B \quad (3.24b)$$

$$F_2^P \leq 1 - F_1^P \quad (3.25a)$$

$$F_2^B \leq 1 - F_1^B \quad (3.25b)$$

$$F_2^P \geq G_2^P - F_0^P - F_1^P \quad (3.26a)$$

$$F_2^B \geq G_2^B - F_0^B - F_1^B \quad (3.26b)$$

12. Only one modulation format is assigned for one request.

$$F_0^P + F_1^P + F_2^P = 1 \quad (3.27a)$$

$$F_0^B + F_1^B + F_2^B = 1 \quad (3.27b)$$

13. Find the number of subcarriers needed for primary (backup) communication if QPSK is the best modulation format

$$n^P \geq F_2^P \cdot B/k_2 \quad (3.28a)$$

$$n^B \geq F_2^B \cdot B/k_2 \quad (3.28b)$$

$$n^P \leq M(1 - F_2^P) + B/k_2 \quad (3.29a)$$

$$n^B \leq M(1 - F_2^B) + B/k_2 \quad (3.29b)$$

14. Exactly one spectrum gap of one core must be used on each fiber on the path used for the primary (backup) communication of the new request.

$$\sum_{g \in G} x_{ij}^g = x_{ij} \quad \forall (i, j) \in E \quad (3.30a)$$

$$\sum_{g \in G} y_{ij}^g = y_{ij} \quad \forall (i, j) \in E \quad (3.30b)$$

15. The starting subcarrier of the primary (backup) communication for the new request must be greater than or equal to the starting subcarrier of the g^{th} gap on edge (i, j) .

$$\theta \geq a_{ij}^g \cdot x_{ij}^g \quad \forall (i, j) \in E, \quad g \in G \quad (3.31a)$$

$$\omega \geq a_{ij}^g \cdot y_{ij}^g \quad \forall (i, j) \in E, \quad g \in G \quad (3.31b)$$

16. The ending subcarrier of the primary (backup) communication for the new request must be less than or equal to the ending subcarrier of the g^{th} gap on edge (i, j) .

$$\phi \leq b_{ij}^g + M \cdot (1 - x_{ij}^g) \quad \forall (i, j) \in E, \quad g \in G \quad (3.32a)$$

$$\psi \leq b_{ij}^g + M \cdot (1 - y_{ij}^g) \quad \forall (i, j) \in E, \quad g \in G \quad (3.32b)$$

17. The total number of subcarriers on all gaps and cores used by the primary (backup) communication must be greater than or equal to the required bandwidth.

$$(\phi - \theta + 1) \geq n^P / |C| \quad (3.33a)$$

$$(\psi - \omega + 1) \geq n^B / |C| - M(1 - q) \quad (3.33b)$$

$$q \geq q^d \quad \forall d \quad (3.33c)$$

$$q \leq \sum_d q^d \quad (3.33d)$$

$$psi \leq M.q \quad (3.33e)$$

18. The value of the ending subcarrier ϕ (ψ) should be greater than that of the value of the starting subcarrier θ (ω) of the primary (backup) communication for the new request of the g^{th} gap on edge (i, j) .

$$\phi \geq \theta \quad (3.34a)$$

$$\psi \geq \omega \quad (3.34b)$$

3.7.1 Justification of ILP

Along with establishing a route for primary as well as backup communication during disasters for the new communication request, the objective function of our algorithm is to minimize the resource usage for the new communication request by minimizing the sum of the spectrums used for the primary and backup communication. This is calculated by $\phi - \theta$ for the primary communication and $\psi - \omega$ for the backup communication.

Constraint 3.2 corresponds to the flow conservation equation for the primary path for the new communication request where x_{ij} is a binary variable (0/1). It's value becomes 1 if the link $(i \rightarrow j) \in E$ is used in the primary path for the new communication request; otherwise x_{ij} is 0.

When a disaster affects any node or edge in the network used for the primary communication, a backup path needs to be established, to ensure the normal data flow. Constraint

3.6 to 3.7 calculates the route for backup communication. When disaster d happens, q^d is set to 1; otherwise it is 0. If $q^d = 0$, constraint 3.3 ensures that $w_{ij}^d = 0$. The value of w_{ij}^d is later required in 3.6 and 3.7 to determine a backup path.

In constraint 3.4 and 3.5, q^d , $d \in D$, is set to 1 when a disaster affects the primary communication; 0 otherwise.

If $q^d = 0$, w_{ij}^d in constraint 3.3 becomes equal to 0 which means constraints in 3.6 and 3.7 become $y_{ij} \leq M$ and $y_{ij} \geq 0$, which are both trivial constraints. Again, if $q^d = 1$, constraint 3.6 and 3.7 become $y_{ij} \leq w_{ij}^d$ and $y_{ij} \geq w_{ij}^d$, so that $y_{ij} = w_{ij}^d$. In other words, the path for the backup communication is defined by y_{ij} , which avoids all disasters that affect the path for the primary communication and thus sets up a new path for backup communication.

Constraint 3.9a and 3.9b determines the length of the primary and backup path.

Constraints 3.10a (3.10b), and 3.11a (3.11b) ensure that F_0^P (F_0^B) = 1, if the length of the path for the primary (backup) communication does not exceed the optical reach d_0 for 16-QAM modulation format, otherwise F_0^P (F_0^B) = 0. If F_0^P (F_0^B) = 1, constraints 3.12a (3.12b) and 3.13a (3.13b) ensure that the number of subcarriers or n^P (n^B) = B/k_0 ; otherwise, constraints 3.12a (3.12b) and 3.13a (3.13b) become trivial. Thus, if 16-QAM can be used, the number of subcarriers n^P is determined by the packing density for 16-QAM.

Constraints 3.14a (3.14b), and 3.15a (3.15b) ensure that G_1^P (G_1^B) = 1, if the length of the path for the primary (backup) communication does not exceed the optical reach d_1 for 8-QAM modulation format, otherwise G_1^P (G_1^B) = 0. If the length of the path for the primary (backup) communication does not exceed the optical reach d_0 for 16-QAM modulation format, clearly it does not exceed the optical reach d_1 for 8-QAM modulation format. Therefore, if $F_0^P = 1$ ($F_0^B = 1$) $G_1^P = 1$ ($G_1^B = 1$). Constraints 3.16a, 3.17a, 3.18a ensure that if $F_0^P = 1$ (i.e., 16-QAM may be used) $F_1^P = 0$; Otherwise, if $G_1^P = 0$, $F_1^P = 0$; otherwise $F_1^P = 1$. This may be achieved using a non-linear constraint $F_1^P = (G_1^P = 1) \wedge (F_0^P = 0)$. Constraints 3.16a, 3.17a, 3.18a is the linear equivalent of the non-linear constraint. A

similar argument is applicable for the backup path.

Constraints 3.19a (3.19b) and 3.20a (3.20b) determines the number of subcarriers needed if the primary (backup) communication uses 8-QAM as the modulation format to carry out the new communication request.

Similarly, constraints 3.21a (3.21b), and 3.22a (3.22b) ensure that $G_2^P (G_2^B) = 1$, if the length of the path for the primary (backup) communication does not exceed the optical reach d_2 for QPSK modulation format, otherwise $G_2^P (G_2^B) = 0$.

A more detailed explanations about the choice of modulation format for the new requested communication would be, constraints 3.14a (3.14b) and 3.15a (3.15b) are similar to constraints 3.10a (3.10b) and 3.11a (3.11b) and ensure that $G_1^P (G_1^B) = 1$, if the length of the path for the primary (backup) communication does not exceed the optical reach d_1 for 8-QAM; otherwise $G_1^P (G_1^B) = 0$. In other words, if $G_1^P (G_1^B) = 1$, 8-QAM may be used for the primary (backup) communication. If the length of the path for the primary (backup) communication does not exceed the optical reach d_0 for 16-QAM, it clearly does not exceed the optical reach d_1 for 8-QAM. Therefore, if $F_0^P (F_0^B) = 1$, $G_1^P (G_1^B)$ will always be 1. Our objective is to make sure that we should use 16-QAM, if possible; otherwise, we should use 8-QAM. In other words, $F_1^P (F_1^B) = 1$ only when $F_0^P (F_0^B) = 0$ and $G_1^P (G_1^B) = 1$; otherwise it is 0. It may be readily verified that constraints 3.16a (3.16b), 3.17a (3.17b) and 3.18a (3.18a) ensures that this condition is satisfied. Constraints 3.21a (3.21b) and 3.22a (3.22b) are like constraints 3.14a (3.14b) and 3.15a (3.15b) and ensure that $G_2^P (G_2^B) = 1$, if the length of the path for the primary (backup) communication does not exceed the optical reach d_2 for QPSK; otherwise $G_2^P (G_2^B) = 0$. To determine $F_2^P (F_2^B)$ we use a technique similar to that for $F_1^P (F_1^B)$. Here we need to ensure that $F_2^P (F_2^B) = 1$ only when $G_2^P (G_2^B) = 1$ (i.e., QPSK may be used) and $F_1^P (F_1^B) = 0$ (i.e., 8-QAM cannot be used) and $F_0^P (F_0^B) = 1$ (i.e., 16-QAM cannot be used). It may be verified that constraints 3.23a (3.23b) to 3.26a (3.26b) ensures that this condition is satisfied.

3.27a (3.27b) ensures that only one modulation format is assigned for one request operating that time. Assignment of modulation format happens on determination of the distance of the route requested and the optical reach of the respective format. Only one among $F_0^P/F_1^P/F_2^P$ can be 1 for a certain request.

3.28a (3.28b) and 3.29a (3.29b) determines the number of subcarriers needed if the primary (backup) communication uses QPSK as the modulation format to carry out the new request.

Constraint 3.30a (3.30b) ensures that exactly one gap on the spectrum is used for the primary (backup) path in communication.

Constraints 3.31a (3.31b) states that the starting subcarrier frequency of the new primary (backup) communication should be always equal to or greater than the starting subcarrier frequency of the gap $g \in G$ on fiber $(i \rightarrow j), \forall (i \rightarrow j) \in E$.

Constraint 3.32a (3.32b) states that the ending subcarrier frequency of the new primary (backup) communication should always be less than or equal to the ending subcarrier frequency of the gap $g \in G$ on fiber $(i \rightarrow j), \forall (i \rightarrow j) \in E$. If the new primary (backup) communication uses the g^{th} gap on the fiber $(i \rightarrow j), \forall (i \rightarrow j) \in E$ then the value of $x_{ij}^g(y_{ij}^g)$ becomes 1. Thus the right hand side of the constraint 3.16a (3.16b) equals the ending frequency of the gap (i.e., b_{ij}^g) and hence the ending frequency of the communication is less than or equal to the ending frequency of the gap g . But if the new primary (backup) communication does not use the g^{th} gap on the fiber $(i \rightarrow j), \forall (i \rightarrow j) \in E$ then the value of $x_{ij}^g(y_{ij}^g)$ becomes 0. Since M is a large constant, the right hand side of the constraint 3.16a (3.16b) becomes a very large quantity. The left hand side of constraint will always be less than the right side of the constraint.

Constraints 3.33a (3.33b) determines that the total number of subcarriers on all gaps and cores used by the primary (backup) communication must be greater than or equal to the required bandwidth of the new communication request.

Constraints 3.33c, 3.33d, refeq31e makes sure that if primary path does not come across any node where disaster has happened, that is q_d is 0 where $d \in$ all the nodes affected by disaster, then no paths for backup will be created, hence evaluation of ending values for gap for backup path is also not necessary. But if a disaster happens, assigning a value of 1 to q_d , q becomes 1 as well. As psi is forced to be less than M which is a very large number, we determine the actual value of psi from constraint 3.32b.

Constraints 3.34a (3.34b) ensure that the value of the ending subcarrier denoted by ϕ (ψ) should be greater than that of the value of the starting subcarrier θ (ω) of the primary (backup) communication for the new request of the g^{th} gap on the edge (i, j) .

Chapter 4

SIMULATION EXPERIMENTS AND RESULTS

One of the efficient techniques in computer networking that are used to study and analyze the performance of an algorithm developed for a specific problem is simulation. By using a network simulator, one does not need to deploy a network physically. The ILP formulation evaluated in this research generates an optimal solution, which can be used as a benchmark for the other heuristics. The primary objective of this simulation study was to evaluate the proposed ILP formulation.

The important metric for our studies is the *blocking probability*, defined as the ratio of the number of requests for communication which could not be handled by the network (due to lack of spectral resources) to the total number of requests for communication.

We have carried out several sets of experiments to study our formulation. For the experiments we have used three networks - a 8-node network, an 11-node network (the COST-239 network) and 14-node network (the NSFNET). In our experiments, if two nodes x and y in a network has an edge, $x \rightarrow y$, it means that these two nodes are connected using bi-directional optical fibers with multiple cores. For each such network link, available spectral resources are divided into fixed slices or slots, where each slice may be used by a subcarrier having the corresponding carrier wavelength. We have carried out experiments considering 50, 70 and 90 slices per fiber and compared the resource utilization and the blocking probability in all the networks with multiple demand sets. We have also carried out experiments on the 8-node, the 11-node and the 14-node network having 4, 6 and 7 cores in each fiber and compared the resource utilization and blocking probability. For processing a communication request, we need input data such as the file requested, the destination node (i.e., the node requesting the file), information about the replication strategy (i.e., which

data centers contain copies of file f_i for all files f_i), a definition of the disasters that may affect the network.

4.1 Experimental Setup

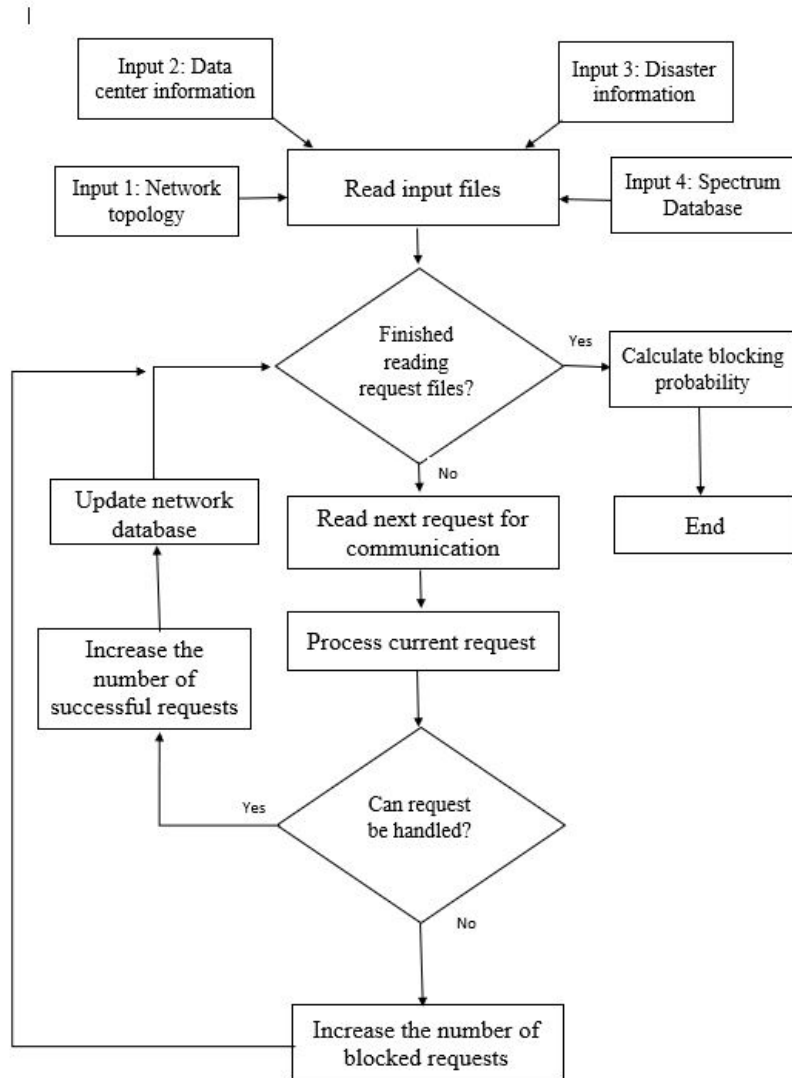


Figure 4.1: Flow diagram for experimental setup

The objective of the ILP described in Chapter III is to find an optimal scheme for handling a request for communication. This means that, for the new request, the ILP determines

- a primary path that is used for communication when the network is fault-free,
- a backup path, that avoids the disasters that affects the primary path, which will be used when the primary path is not usable, and
- a scheme to determine the subcarrier wavelengths and the cores that will be used for communication.

The objective for the scheme is that the total amount of spectrum used is as small as possible.

We have used a simulator, written in C, that uses the IBM ILOG CPLEX to determine the blocking probability under different scenarios. CPLEX is a tool developed by IBM [34], which can be used to solve linear optimization problems, commonly known as Linear Programming (LP) problems, including Integer linear programming (ILP). CPLEX Optimizer provides flexible, high-performance mathematical programming solvers for linear programming, quadratic programming (QP) problem, quadratically constrained programming (QCP) and mixed integer programming (MIP) problems [34]. A simplified block diagram for the simulator is shown in Figure 4.1.

As shown in Figure 4.1, the input to the simulator is the following:

- a file containing a definition of the network topology, consisting of the list of nodes in the network and the list of edges in the network, each edge representing a bidirectional fiber link.
- a file defining the data centers. This includes which nodes are data centers and which data centers contain file.

- a file defining the disasters to be considered. Each disaster is defined as a set of edges which become inoperative when that disaster takes place.
- a file containing the spectrum and other information giving the number of slots per fiber and the number of cores in each fiber.
- a file containing many requests. Each request consists of a) the file requested, b) the node requesting the file, and c) the required data rate for communication, specified by the number of subcarriers needed. In our discussions below this will be called the request file.

The simulator maintains an internal database that includes the information in the first four input files mentioned above. The internal database also contains the current network state - defined by the spectrum used on each core of each fiber in the network. When the simulator starts, the network is “empty” in the sense that no spectrum is initially allotted to any communication. When a new request for communication is processed and resources are available to handle the request, the network state is updated to include the spectrum resources allotted to set up the spectrums that must be reserved to handle the new request.

The first step (shown in the first block) of our simulator reads the first four input files. Then the simulator enters into a loop that iterates until all the requests in the request file have been considered (condition for continuing the loop is shown in the second block). In a given iteration of this loop, the simulator

1. reads the next request for the request file.
2. generates constraints for our ILP to solve in order to handle the request currently being considered. These constraints take into account the network state, the input files read, and the current request and are saved in a file of type lp.
3. invokes the CPLEX ILP solver. The solver reads the .lp file and attempts to find, if possible, an optimal solution for the problem.

4. If the solver is able to find a solution, the output generated by the solver is a file that gives
 - a primary path that is used for communication when the network is fault-free,
 - a backup path, that avoids the disasters that affects the primary path, which will be used when the primary path is not usable, and
 - a scheme to determine the subcarrier wavelengths and the cores that will be used for communication. The assignment of subcarriers is determined by the spectral efficiency of the modulation formats (16-QAM, 8-QAM, BPSK) and the optical reach for each of these formats.

The simulator reads the file generated by the CPLEX solver, updates the network state and increments the number of successful requests.

5. If the solver is unable to find a solution, the simulator increments the number of blocked requests.

When all the requests have been processed, the simulator exits the iterative process described above. The simulator calculates the blocking probability by computing the ratio between the number of blocked requests to the total number of requests provided for the simulation.

4.2 Performance Study

The topologies used for the experiments are:

1. a 8-node network,
2. the COST-239 network,

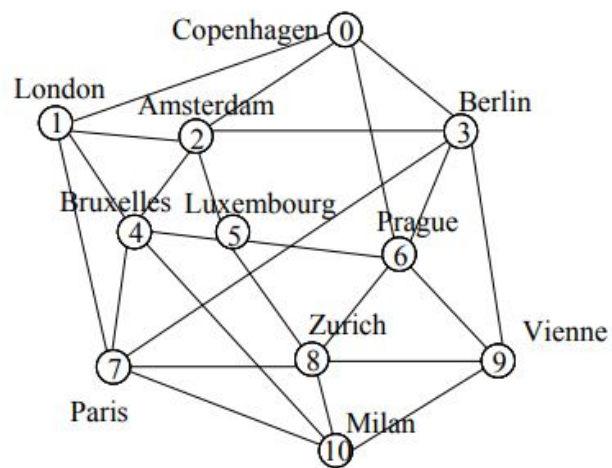


Figure 4.2: COST-239 network [35]

3. the NSFNET.

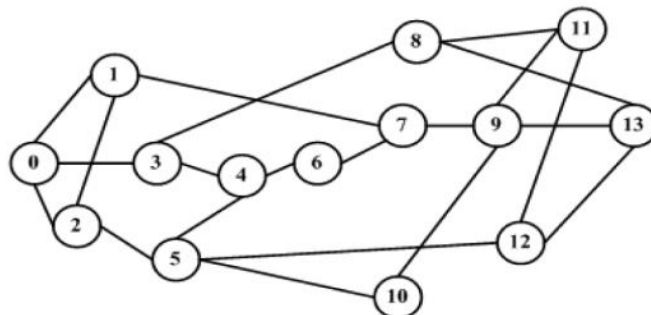


Figure 4.3: NSFNET [36]

We present the experimental results in this section. We have studied how the efficiency of the algorithm was affected by various parameters such as an increase in the number of disasters, an increase in the number of data centers, an increase in the number of cores/fiber. The simulation was carried out on a 8-node network, the COST-239 network (an European network with having 11 nodes), NSFNET (a 14-node network in the US). The simulation was done for 30, 50, 70 requests for a specified network topology. Each value in the graphs shown below are based on the average values which we got after running a particular size of request file 3 times, each time with different request files. We analyzed the blocking probability and the resource usage as the network size increased, the number of cores per fiber link increased, the number of disasters in each network increased, and the number of data centers to handle those disasters increased.

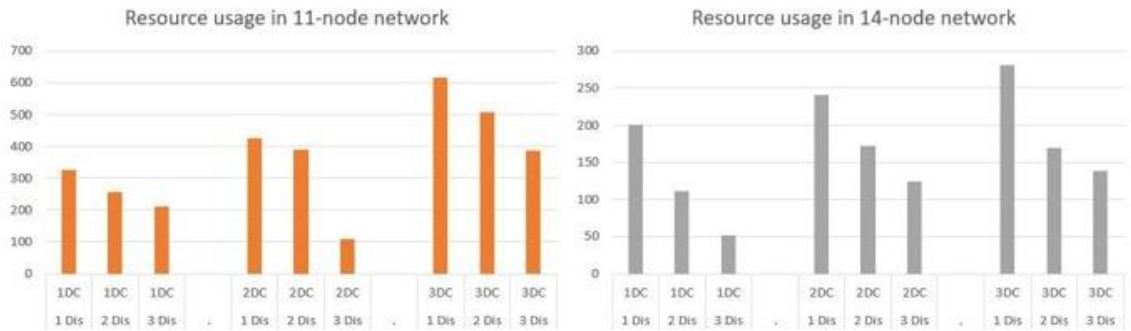


Figure 4.4: Resource usage 11-node, 14-node network for 50 requests.

1. When a disaster occurs in the node which is requesting a communication itself, there exists no destination where the file could be sent. Such disasters clearly cannot be handled and are excluded from the study.
2. After comparing the results of 8-node, 11-node, 14-node network for 50 requests, from Fig 4.4 we observed that the resource usage of a network decreases with the increase in the number of disasters for their respective networks. As the number of disasters increase, the number of blocked requests increases.

Since the number of requests that are successful, decreases, resource consumption decreases.

3. If we increase the number of data centers in each network we observe that the resource usage increases significantly. Data centers are located at a certain distance from each other. Thus if a path from a data center gets blocked because of a disaster, the file can be fetched from an alternate DC of the network. Thus we can handle more requests which increases the resource consumption by the network.

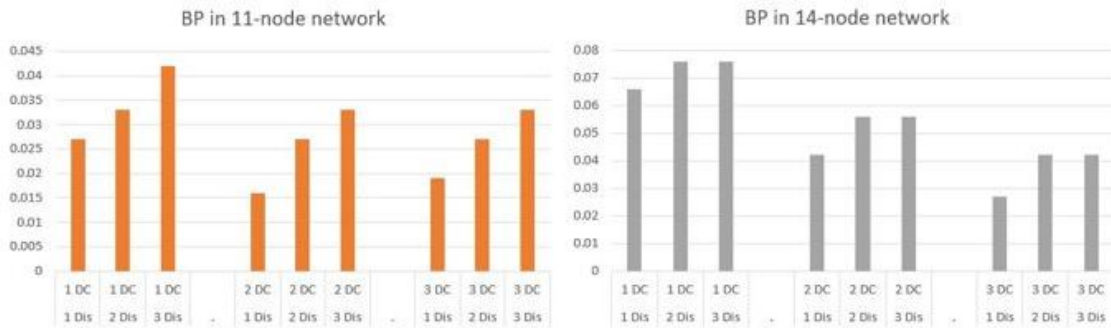


Figure 4.5: Blocking probability in 11-node, 14-node network for 50 requests

4. After comparing the results of the 8-node, 11-node, 14-node network for 50 demands from Fig 4.5, it is observed that the blocking probability increases with the number of disasters in each network. But again, the blocking probability decreases as the number of data centers increases in each network, providing more feasible routes for handling more requests.

5. We perform several sets of comparison on the NSF network as follows:

Fig 4.6 shows the resource usage of NSFNET with 2, 3, and 4 data centers, 2 disasters and 4 cores per fiber link. One of the objective of our algorithm is to minimize the resource usage which we can clearly see from the graph.

- We observe that for 30, 50, 70 requests, the resource consumption

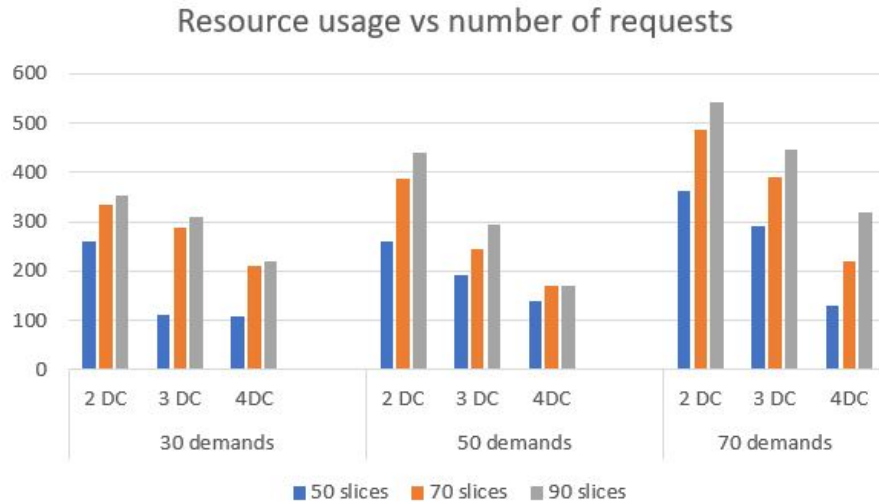


Figure 4.6: Resource usage vs number of requests with 2 disasters and 4 cores

decreases with the increase in the number of data centers in the network. Establishing more data centers help in finding a feasible scheme for communicating files to requested nodes from locations which are closer and thus requiring less resources.

- When we compare resource usage for 50, 70 and 90 slices, we find that the requests that are blocked due to lack of free resources using, say, 50 slices, can easily be handled when we increase the number of slices to a higher value (e.g., 90). Hence, the total resource usage increases with the increase of the number of slices.
- Resource usage increases with the increase in number of requests from 30 to 70. As expected, when the number of requests is increased, the consumption of resources in order to satisfy them also increases simultaneously.

6. Fig 4.7 shows the blocking probability (BP) of NSFNET with 2, 3, 4 data centers, 2 disasters and 4 cores per fiber link. Another objective of our problem is to reduce the blocking probability of the network so that we can process more

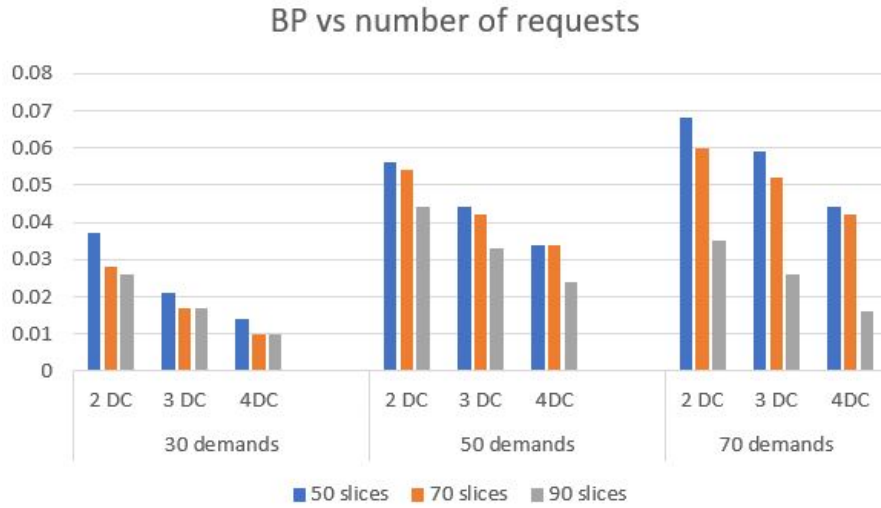


Figure 4.7: BP vs number of requests with 2 disasters and 4 cores

requests. Thus from Fig 4.7 we see that the BP of a network decreases with the increase in number of data centers in each network. Also, as expected, the BP can be reduced when we increase the number of slices from 50 to 90, thus enabling availability of more resources to successfully handle more requests.

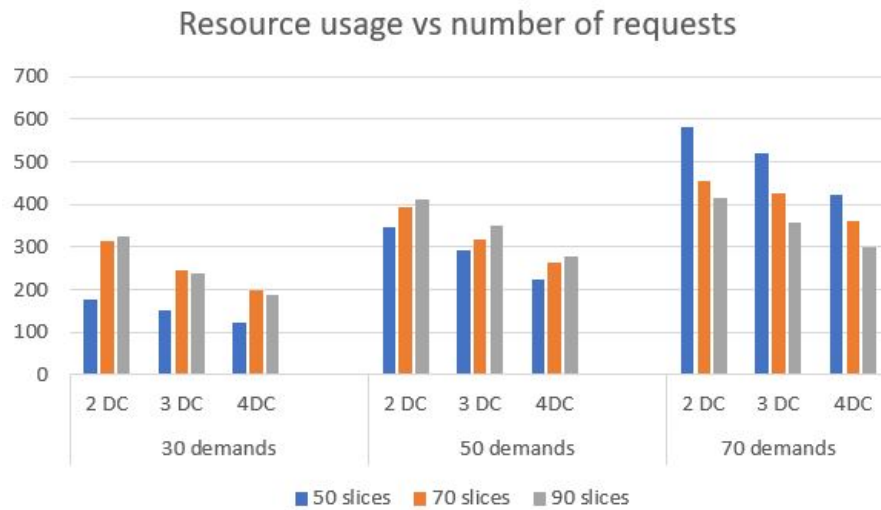


Figure 4.8: Resource usage vs number of requests with 2 disasters and 6 cores

7. Fig 4.6 depicts the resource usage of NSFNET with 2, 3, and 4 data centers, 2 disasters and 6 cores per fiber link. If we compare Fig 4.6 with Fig 4.8 shown

above, we observe that the resource usage decreases with the increase in the number of cores for the same number of requests, data centers, and number of disasters. For the same bandwidth, the resources required with each fiber having 6 cores will definitely be lesser than that when each fiber has 4 cores. Introducing the concept of cores in SDM thus helps us save resources. The total bandwidth required to handle the same set of requests is less when the number of cores is more. As expected, resource usage decreases with an increase in the number of data centers in the network, decreases with increase in the number of subcarriers/slices and increases with increase in the number of requests.

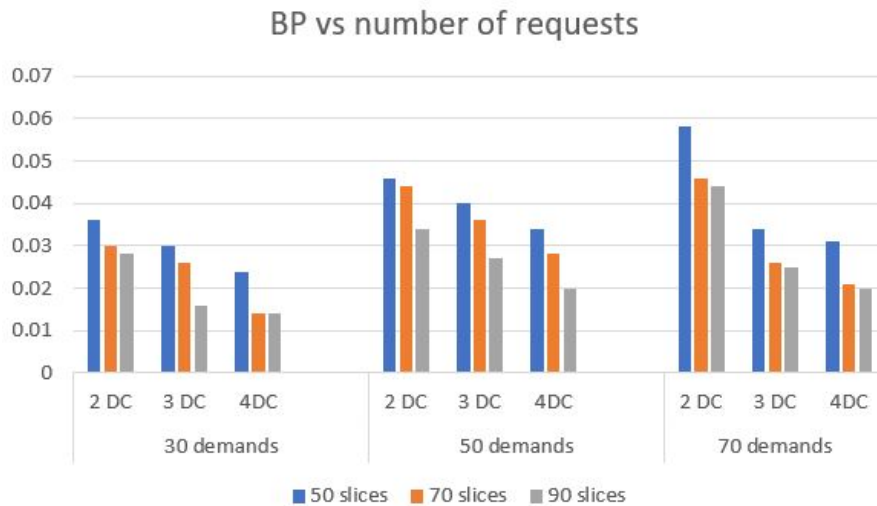


Figure 4.9: BP vs number of requests with 2 disasters and 6 cores

8. Fig 4.9 shows the BP of NSFNET with 2, 3, 4 data centers, 2 disasters and 6 cores per fiber link. If we compare Fig 4.7 with Fig 4.9 shown above, we observe that the BP of a network decreases with an increase in the number of cores, as more requests can be accommodated now with fewer subcarriers. Also, the BP decreases with increase in number of data centers in the network and increases with the number of subcarriers per core.
9. Fig 4.10 shows the resource usage of NSFNET with 2, 3, 4 data centers, band-

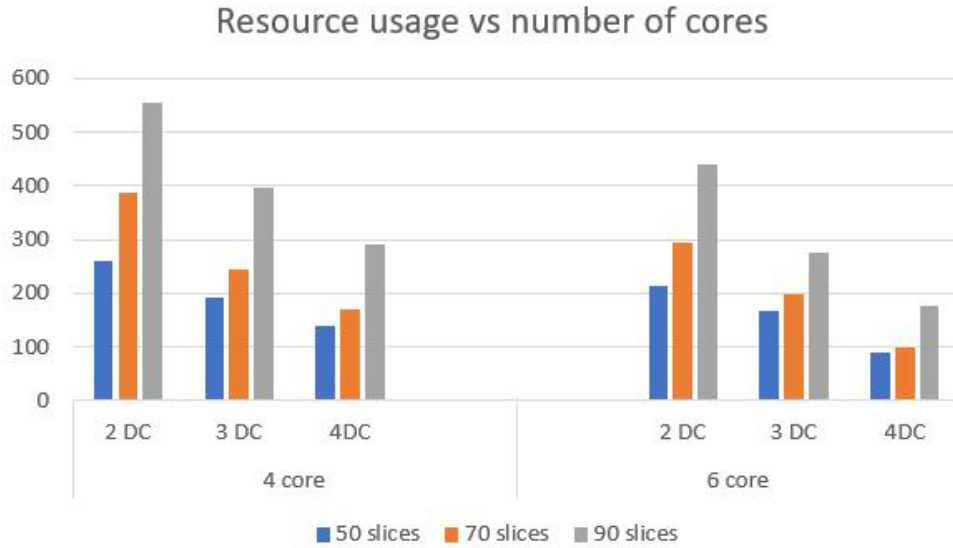


Figure 4.10: Resource usage vs number of cores

width as 48, 2 disasters and 50 requests. This graph shows better comparison of the resource usage of a network with 4 and 6 cores. For every demand set executed, we found out that the resource needed was less when we increased the number of cores from 4 to 6.

Resource usage	4 core			7 core			Percentage increase in resource consumption		
	2 DC	3 DC	4DC	2 DC	3 DC	4DC	2 DC	3 DC	4DC
50 slices	286	190	140	184	168	161	8.16	7.84	7.28
70 slices	318	245	170	190	296	269	7.31	7.2	7.77
90 slices	346	395	290	506	470	346	8.88	8.5	7.33

Figure 4.11: percentage increase in resource usage with increase in cores

10. Fig 4.11 depicts the percentage increase in the NSFNET for 50 requests, 96 subcarriers, 2 disasters and 2, 3, 4 data centers with 4 and 7 cores. From table 4.11 we see that there is approximately 7-8% decrease in resources used when we increased the number of cores from 4 to 7, thus minimizing bandwidth required.

Chapter 5

CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

In our work we have studied an optimal algorithm for fault-tolerant communication in data center networks using space division multiplexing. The concept of keeping multiple copies of files in multiple data centers has been used to make our network more robust. Under any circumstances of disaster or network breakage, our backup communication does not let any disruption affect the normal communication. We use dedicated path protection scheme to find an alternate route during the same. Also, we have applied SDM technology enabling us to use multiple cores for data transmission. The application of modulation format facilitates improvement in efficiency in determining the subcarriers required to carry out a requested communication. After comparison of results of multiple networks, we can conclude that, the BP of a network can be reduced by increasing :

- (i) the number of DCs,
- (ii) the number of slots/slices per core, or
- (iii) the number of cores.

Similarly resource utilization in SDM networks can be decreased by increasing:

- (i) the number of DCs.
- (ii) the number of slots/slices per core.
- (iii) the number of cores.

5.2 Future work

Frequent occurrences of man-made or natural disasters render networks vulnerable to multiple cascading and correlated failures. This has become a significant concern along with the rise in data traffic and development of new technology. We have used the flex-grid fixed model of SDM technology for designing a robust data center network. In the future, we can incorporate the flex-grid flex model which is known to provide the best flexibility among the existing models in SDM regarding allocation of spectral-spatial resources. We have used dedicated path protection scheme in our work and possible improvements for future includes using the shared path protection scheme (SPP) technique instead of dedicated path protection (DPP) as this might lead to increased resource savings.

BIBLIOGRAPHY

- [1] D. Technologies:, “Data center solutions.” [Online]. Available: <http://dynamictechnologies.ca/portfolio/twitter-bootstrap-3/>
- [2] B. Wen, N. M. Bhide, R. K. Shenai, K. M. Sivalingam *et al.*, “Optical wavelength division multiplexing (wdm) network simulator (owns): architecture and performance studies,” *SPIE Optical Networks Magazine*, vol. 2, no. 5, pp. 16–26, 2001.
- [3] K. Christodoulopoulos, I. Tomkos, and E. A. Varvarigos, “Elastic bandwidth allocation in flexible ofdm-based optical networks,” *Journal of Lightwave Technology*, vol. 29, no. 9, pp. 1354–1366, 2011.
- [4] A. Kodakanchi, “Energy aware routing in optical grid networks,” 2015.
- [5] M. Jinno, B. Kozicki, H. Takara, A. Watanabe, Y. Sone, T. Tanaka, and A. Hirano, “Distance-adaptive spectrum resource allocation in spectrum-sliced elastic optical path network [topics in optical communications],” *IEEE Communications Magazine*, vol. 48, no. 8, 2010.
- [6] K. Christodoulopoulos, I. Tomkos, and E. A. Varvarigos, “Routing and spectrum allocation in ofdm-based optical networks with elastic bandwidth allocation,” in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*. IEEE, 2010, pp. 1–6.
- [7] D. Richardson, J. Fini, and L. Nelson, “Space-division multiplexing in optical fibres,” *Nature Photonics*, vol. 7, no. 5, pp. 354–362, 2013.
- [8] Coriant:, “Space division multiplexing a new milestone in the evolution of fiber optic communication.” [Online]. Available: http://modegap.eu/wp-content/uploads/2013/12/coriant_whitepaper_sdm_121413.pdf

- [9] K. Walkowiak, P. Lechowicz, M. Klinkowski, and A. Sen, "Ilp modeling of flexgrid sdm optical networks," in *Telecommunications Network Strategy and Planning Symposium (Networks), 2016 17th International*. IEEE, 2016, pp. 121–126.
- [10] www.nokiasiemensnetworks.com, "Space division multiplexing: A new milestone in the evolution of fiber optic communication." [Online]. Available: http://modegap.eu/?publication=space-division-multiplexing&wppa_open=yes
- [11] O. CNX, "Reflection, refraction, and dispersion." [Online]. Available: <https://courses.lumenlearning.com/boundless-physics/chapter/reflection-refraction-and-dispersion/>
- [12] S. Bandyopadhyay, *Dissemination of Information in Optical Networks:: From Technology to Algorithms*. Springer Science & Business Media, 2007.
- [13] SDxCentral:, "What is the cloud." [Online]. Available: <https://www.sdxcentral.com/cloud/definitions/what-is-cloud/>
- [14] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," in *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4. ACM, 2008, pp. 63–74.
- [15] D. Klonidis, F. Cugini, O. Gerstel, M. Jinno, V. Lopez, E. Palkopoulou, M. Sekiya, D. Siracusa, G. Thouénon, and C. Betoule, "Spectrally and spatially flexible optical network planning and operations," *IEEE Communications Magazine*, vol. 53, no. 2, pp. 69–78, 2015.
- [16] S. B. Saja Al Mamoori, Arunita Jaekel, "Robust data center network design based on space division multiplexing," in *Optical and Grid Computing, Accepted for publication 2018, International Conference on Computing, Network and Communications*. ICNC, 2014, pp. 1–4.
- [17] G. Keiser, *Optical fiber communications*. Wiley Online Library, 2003.

- [18] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility," *Future Generation computer systems*, vol. 25, no. 6, pp. 599–616, 2009.
- [19] D. Agrawal, S. Das, and A. El Abbadi, "Big data and cloud computing: current state and future opportunities," in *Proceedings of the 14th International Conference on Extending Database Technology*. ACM, 2011, pp. 530–533.
- [20] I. Foster, Y. Zhao, I. Raicu, and S. Lu, "Cloud computing and grid computing 360-degree compared," in *Grid Computing Environments Workshop, 2008. GCE'08*. Ieee, 2008, pp. 1–10.
- [21] Cyberpedia, "What is data center: An overview of data centers." [Online]. Available: <https://www.paloaltonetworks.com/cyberpedia/what-is-a-data-center>
- [22] TechTarget, "Disaster recovery." [Online]. Available: <http://searchdisasterrecovery.techtarget.com/>
- [23] K. Technologies:, "Concepts of orthogonal frequency division multiplexing (ofdm) and 802.11 wlan." [Online]. Available: http://rfmw.em.keysight.com/wireless/helpfiles/89600b/webhelp/subsystems/wlan-ofdm/Content/ofdm_basicprinciplesoverview.htm
- [24] L. Velasco, A. Castro, M. Ruiz, and G. Junyent, "Solving routing and spectrum allocation related optimization problems: From off-line to in-operation flexgrid network planning," *Journal of Lightwave Technology*, vol. 32, no. 16, pp. 2780–2795, 2014.
- [25] C. Carboni and G. Li, "Novel applications of space-division multiplexing," *Frontiers of Optoelectronics*, vol. 9, no. 2, pp. 270–276, 2016.
- [26] Multicom:, "Single mode vs. multi-mode fiber optic cable." [Online]. Available: <https://www.multicominc.com/training/technical-resources/>

single-mode-vs-multi-mode-fiber-optic-cable/

- [27] J. M. Simmons, "On determining the optimal optical reach for a long-haul network," *Journal of Lightwave Technology*, vol. 23, no. 3, p. 1039, 2005.
- [28] S. Jawla and R. Singh, "Different modulation formats used in optical communication system," *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, vol. 8, no. 4, 2013.
- [29] J. Kahn, "Modulation and detection techniques for optical communication systems," in *Coherent Optical Technologies and Applications*. Optical Society of America, 2006, p. CThC1.
- [30] J.-L. Izquierdo-Zaragoza, P. Pavon-Marino, and M.-V. Bueno-Delgado, "Distance-adaptive online rsa algorithms for heterogeneous flex-grid networks," in *Optical Network Design and Modeling, 2014 International Conference on*. IEEE, 2014, pp. 204–209.
- [31] A. Muhammad, G. Zervas, D. Simeonidou, and R. Forchheimer, "Routing, spectrum and core allocation in flexgrid sdm networks with multi-core fibers," in *Optical Network Design and Modeling, 2014 International Conference on*. IEEE, 2014, pp. 192–197.
- [32] Y. Wang, X. Cao, and Q. Hu, "Routing and spectrum allocation in spectrum-sliced elastic optical path networks," in *Communications (ICC), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–5.
- [33] M. Jinno, H. Takara, and B. Kozicki, "Dynamic optical mesh networks: drivers, challenges and solutions for the future," in *Optical Communication, 2009. ECOC'09. 35th European Conference on*. IEEE, 2009, pp. 1–4.

- [34] I. Analytics:, “Cplex optimizer.” [Online]. Available: <https://www.ibm.com/analytics/data-science/prescriptive-analytics/cplex-optimizer>
- [35] W. He, J. Fang, and A. K. Somani, “A p-cycle based survivable design for dynamic traffic in wdm networks,” in *Global Telecommunications Conference, 2005. GLOBECOM'05. IEEE*, vol. 4. IEEE, pp. 5–pp.
- [36] S. Shirazipourazad, C. Zhou, Z. Derakhshandeh, and A. Sen, “On routing and spectrum allocation in spectrum-sliced optical networks,” in *INFOCOM, 2013 Proceedings IEEE*. IEEE, 2013, pp. 385–389.

VITA AUCTORIS

NAME: Ankita Biswas

PLACE OF BIRTH: West Bengal, India

YEAR OF BIRTH: 1990

EDUCATION: Assembly of God Church School
West Bengal, India, 2008
West Bengal University of Technology
West Bengal, India, 2012
University of Windsor, M.Sc
Windsor, ON, 2018