2017

Design of Disaster-Resilient Datacenter WDM Networks

Umesh Shah
University of Windsor

Design of Disaster-Resilient Datacenter WDM Networks

By

Umesh Shah

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

2017

Design of DisasterResilient Datacenter WDM Networks

By

Umesh Shah

APPROVED BY:

_____

J Pathak

Odette School of Business

_____

D Wu

School of Computer Science

_____

A Jaekel, Co-Advisor

School of Computer Science

_____

S Bandhyopadhyay, Advisor

School of Computer Science

October 13, 2017

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 anyones 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

Survivability of data in datacenters, when a fault occurs, is turning into an upcoming challenge in planning cloud-based applications. At the point when such a disaster happens, a particular geological range is disrupted, and units of transmission systems (e.g., nodes and fibers) within the disrupted region end up faulty, leading to the loss of one or more demands. To deal with such a circumstance, a resilient communication code is required, so arrangements can be made to accommodate an alternative disaster-free path when a fault upsets the path utilized for data requests before the failure happens. In this work, we have shown a new approach to deal with this issue, on account of the static Route and Wavelength Assignment (RWA) in Wavelength Division Multiplexing (WDM) systems. In our approach, a set of communication demands can be handled only if it is feasible to i) Find the datacenter node ii) a primary lightpath that minimizes the effect of disasters that may disrupt lightpaths and iii) (for every disaster that upsets the primary lightpath), a backup lightpath that handles the disaster. We have presented, implemented and examined an efficient heuristic to solve this issue.

DEDICATION

In loving memory of my parents

Vipulkumar & Urmila Shah,

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Symbol	Definition
BP	Blocking Probability
DC	Datacenter
DCN	Datacenter Network
DEMUX	De-multiplexer
DWDM	Dense Wavelength Division Multiplexing
Gbps	Giga Bytes per second
IDE	Integrated Development Environment
ILP	Integer Linear Program
LAN	Local Area Network
MAN	Metropolitan Area Network
MUX	Multiplexer
NSFNET	National Science Network
QoS	Quality of Service
RWA	Routing and Wavelength Assignment
SRG	Shared Risk Group
Tbps	Tera Bytes per second
USANET	USA standard network topology
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing
WMD	Weapons of Mass Destruction

Chapter 1

INTRODUCTION

1.1 Overview

The increasing demand in communication networks such as e-mail, online interactive maps, social networks, web search and video storage has become an essential part of life. Datacenter (DC) network is a facility composed of networked computers and storage business of other organizations use to establish, develop, store and propagate large amounts of data, low latency, high availability, and low cost which can be delivered by optical networks [1]. Researchers show that datacenter networks based application are developing the network landscape, by replacing traditional hierarchical and connectivity-oriented Internet towards a more meshed and service-oriented infrastructure, offering applications the promise of scalability, availability, and responsiveness at very low costs [2, 3]. The risk of large-scale failures in DC is increasing rapidly, and the large area is disrupted because of natural disasters like earthquakes, hurricanes, tsunamis, atom bomb or deliberate attacks like Weapons of Mass Destruction (WMD). We assume that the set of disasters $\mathscr{D}$ in the network is known in advance. Each disaster $d \in \mathscr{D}$ disrupts one or more edges and nodes in the network. Disruption of an edge means no communication is possible using that edge. Disruption of an edge means that the affected node becomes inoperative so that we cannot use any edge to/from the node.

Cloud services delivered by datacenter networks pose new opportunities to provide protection against disasters. Survivability against disasters, both natural and man-made attacks, is becoming a major challenge. These events indicate that it is crucial to study and develop robust connection schemes to handle requests for communication in the case of a

disaster [2]. Such a robust communication scheme means that the network can handle requests for communication even when any disaster $d \in \mathscr{D}$. In this thesis, we have developed a heuristic based approach to the problem of designing robust Datacenter Network (DCN).

1.2   Wavelength Division Multiplexing Networks

The method of utilizing multiple optical signals on a single fiber is called Wavelength Division Multiplexing (WDM). WDM in the optical system has made it possible to design large communication system with high throughput. WDM has emerged a lot as the most popular method for the optical network, because of its elasticity and robustness. Since we can accommodate multiple wavelengths on a single fiber, a single fiber can carry great amounts of data. A single fiber can support up to 320 wavelengths. Each wavelength can have a capacity of 100 Gbps or more, ensuring in Tbps chunks of data flow through a fiber.

Lightpath: Signals travel through optical fibers as light. These signals are used to travel from a source to a destination node. A lightpath is an optical connection from source node to destination node. It begins from an end node, connects a number of fibers and router nodes, and ends in another end node [2]. Lightpaths are used to carry data in the form of optical signals. Several lightpaths can be transmitted on a single fiber using different carrier wavelengths. One of the challenges involved in designing wavelength routed networks is to develop efficient algorithms for establishing lightpaths in the optical network [4]. The algorithms must be able to choose routes and assign wavelengths to connections in a way which efficiently utilizes network resources (channels/wavelengths).

Routing and Wavelength Assignment (RWA): The problem of discovering a route for a lightpath and appointing a wavelength and resources to lightpaths in WDM networks is defined as the Routing and Wavelength Assignment. The RWA problems are considered as the NP-complete problems [5]. The main objective of the RWA problem is to establish as many lightpaths as possible, considering the resource limitations which minimizes the

network operation cost and increases the network performance [6].

In numerous applications, the real location of the server remains hidden from the user as it is not vital. In this case, it is possible to select the best source from the set of possible source to execute a job. The source which is used for fulfilling the request is called as the datacenter.

There are fundamentally two different demand allocations models for WDM optical networks. In static traffic model, the set of demands is fixed and known in advance. For dynamic traffic, the setup time and the duration of the demands are not known in advance, they are generated based on certain distributions.

1.3   Problem Statement & Solution Outline

The input to the problem is a set $\mathscr{R}$ of communication requests from users. A request $r \in \mathscr{R}$ is to communicate some file f_i to some node t_i in the datacenter network. The problem is to carry out the following tasks for each request $r \in \mathscr{R}$:

- find a suitable datacenter, say w_i, to store a copy of file f_i to handle case for fault-free communication,

- find a suitable datacenter, say w_j, to store a copy of file f_i to handle case for communication in the case of a disaster that affects the scheme used for fault-free communication,

- find an appropriate route and channel to define the primary lightpath for handling the request when the network is fault-free,

- find an appropriate route and channel to define the backup lightpath for handling the request when the network encounters some disaster $d \in \mathscr{D}$ that disrupts the primary lightpath. It is important to note that the same backup path will be used for all disasters that disrupt the primary lightpath.

3

Our objective is to solve the above problem such that the total number of copies of all the files is as small as possible.

While finding out the primary path and backup path for a set of requests, we must assure that the following conditions are fulfilled:

- The primary lightpath is from some datacenter w_i in the network to any user node t_i requesting the file.

- The backup lightpath to handle disaster will be from any datacenter node, say w_j, to the user node t_i. If the set of disasters $\mathscr{D}$ includes the source node w_i of the primary lightpath, w_j must be distinct from w_i.

- Each primary and backup lightpath needs to fulfill the wavelength continuity constraint [7] wavelength clash constraint [8] and the optical reach constraint [9]. This means that,

    - on all edges along the path used by a lightpath (primary or backup), the same carrier wavelength must be used.

    - on any edge that is shared by more than one lightpath, the same carrier wavelength cannot be used by more than one lightpath.

    - the length of the path used by any lightpath must be less than the maximum distance an optical signal can travel before the quality of the optical signal degrades below acceptable limits.

As we are considering static requests for our problem, the problem is to consider all the requests for communication request at the same time. This is a very complex task and solving it takes an inordinate amount of time. To make the problem tractable, a heuristic that relaxes some of the conditions needed to solve the problem optimally was proposed by Ms. Saja Al Mamori, Dr. Arunita Jaekel and Dr. Subir Bandyopadhyay. The heuristic

works within a reasonable time. The heuristic requires, as input, all the information about the network under consideration, the set $\mathscr{R}$ of requests to be handled and the set $\mathscr{D}$ of disasters to be handled. The heuristic uses an Integer Linear program (ILP).

In this thesis, we have studied the heuristic for solving the problem and have carried out extensive studies on the feasibility of this heuristic.

1.4 Thesis Organization

The rest of this thesis is organized as follows: Chapter 2 provides a review of some of the concepts and terminologies that are related to this work. It also includes a review of some of the closely related work of other researchers. In Chapter 3, we have defined the problem and presented the heuristic that we have studied. Chapter 4 discusses the simulation results of our experimentation and analysis of our obtained results. In Chapter 5, we have concluded the thesis by proposing some future work.

Chapter 2

REVIEW

2.1 Optical Networks

An optical network connects personal computers or any other devices which can generate or accumulate information in electronic form using optical fibers. An optical network comprises of numbers of nodes which are interdependent to each other to carry out communication across the network. Information is transmitted between sender and receiver node in the form of light pulses. Optical fibers are long, thin strands of glass having a diameter of a human hair. They are for the most part arranged in bundles, known as optical cables and are used to transmit signals over long areas [10]. Figure 2.1 shows the optical cable with the bundle of several optical fibers.
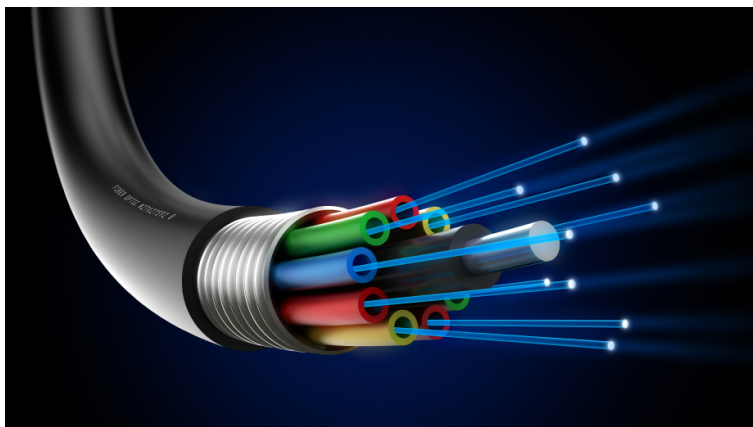
Figure 2.1: Optical Cable [11]

An optic fiber is comprised of three layers: core, cladding, and buffer. The cylindrical center is the innermost layer and is made of a very high-quality glass (silica) or plastic. The cladding is the external material covering the center and it is also made of glass. The
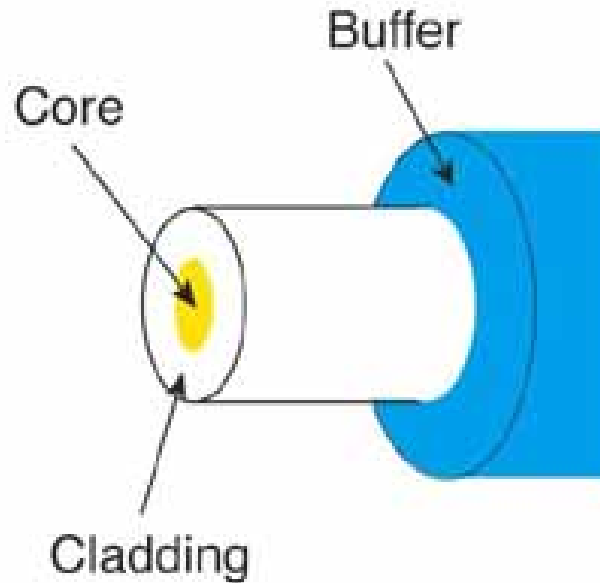
Figure 2.2: Layers in an optical fiber [12]
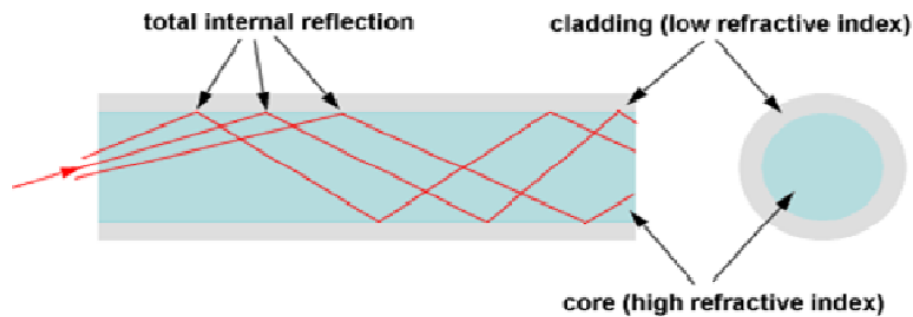
Figure 2.3: Total Internal Reflection inside optical fiber [13]

third layer i.e. buffer is known as the outer most layer of an optical fiber and is made up of plastic such as nylon or acrylic. A buffer protects both the center and cladding from any sort of physical harm. An optical signal goes through the center as light pulses and bounces into cladding which again reflects the light back to the core. This idea is known as total internal reflection and it results in lower light signal attenuation and less energy loss. Figure 2.2 and 2.3 demonstrates a symbolic optical fiber and its three layers.

2.2   Wavelength Division Multiplexing

The immense bandwidth requirements faced at present communication systems have stimulated the enormous deployment of optical backbone networks. Wavelength-Division-Multiplexing (WDM) has emerged as the most accepted technology for optical systems, because of its adaptability and robustness. In a WDM network, an end-to-end link is set up through a wavelength/channel, known as lightpath [5].

WDM system can exchange information on various channels by utilizing a single fiber. In WDM networks light from various laser sources, each with a different wavelength is joined into the single beam with the assistance of a multiplexer. At the receiving node, a De-multiplexer (DEMUX) is put that isolates the wavelengths from the beam into separate optical signals. Normally the transmitter comprises a laser and a modulator. The light source creates an optical carrier signal at either established or tunable wavelength. The receiver consists of a photodiode detector which changes an optical signal to into an electrical signal [14].
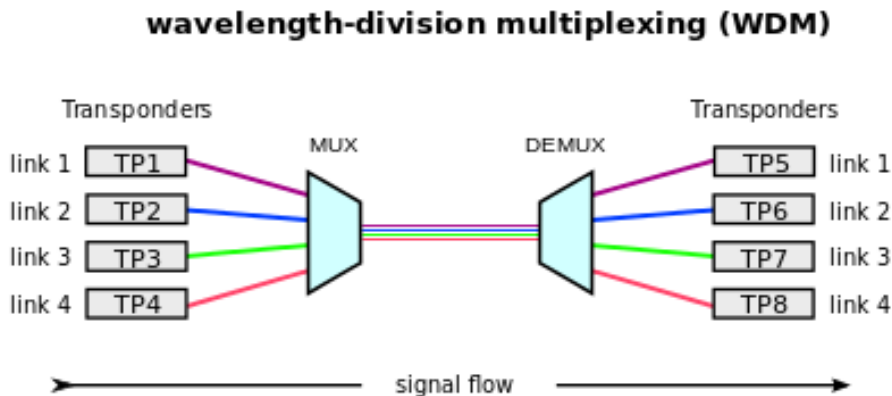
Figure 2.4: Wavelength Division Multiplexing System [15]

Figure 2.4 presents a WDM system with *n* channels (wavelengths). The sender node has *n* transmitters, each tuned to a vary wavelengths from λ_1 to λ_n.

2.3 Routing and Wavelength Assignment

Each lightpath must be allocated a route over the physical network, and a specific channel on each fiber it travels. For increasing the efficiency of wavelength-routed all-optical networks the issue of routing and wavelength assignment is crucial [16]. For a given physical system structure and the resource connection, the routing and wavelength assignment (RWA) problem is to find and a perfect route and wavelength out of the many possible choices for each connection so that no two paths sharing an edge have a similar wavelength [16].
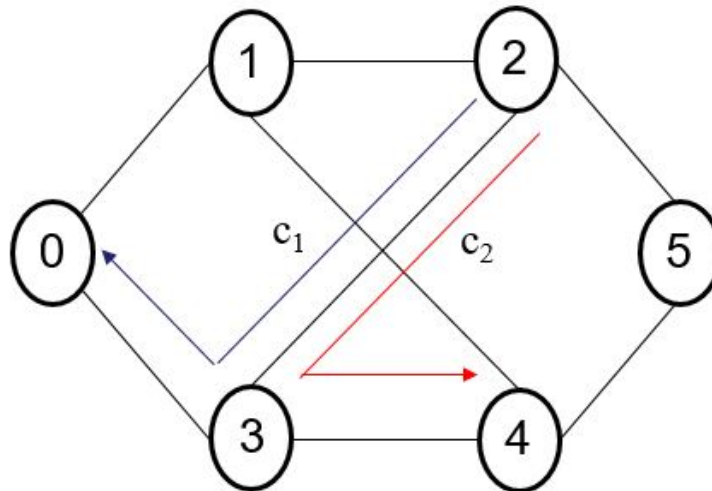
Figure 2.5: Routing Wavelength Assignment

The primary objective of the RWA problem is to establish as many lightpaths as possible, considering the resource limitations which minimizes the network operation cost and increases the network performance [6]. As can be found in Figure 2.5, a lightpath is accomplished by deciding a path of physical edges between the source and destination edge nodes, and reserving a specific wavelength on each of these edges for the lightpath.

There are three main constraints that should be fulfilled for a valid RWA:

I. Wavelength continuity constraint

II. Wavelength clash constraint

III. Optical reach

Wavelength continuity constraint states that a lightpath must use the same wavelength on all the links along its path from source to destination node [5]. Wavelength clash constraint states that the same wavelength cannot be assigned to more than one lightpath on the same link, at the same time. Optical reach [16] (also called as transmission reach [17]) is the maximum distance an optical signal can travel before signal regeneration is needed.

2.4 Lightpath

A lightpath is an optical connection from one end node to another. It starts from an end node, bridges several fibers and router nodes, and terminates in another end node [2]. A lightpath may or may not have multiple wavelengths from source to destination, which depends on the wavelength conversion capability of the network [2, 18]. A lightpath in a WDM network is a unidirectional optical connection between a source node and a destination node, which carries data in the form of encoded optical signals and may span multiple fiber links and use one or multiple wavelengths [5]. Different lightpaths in WDM networks can apply the same wavelength, as long as they do not share any common links. Because of the WDM technology, multiple lightpaths can be set up on the same fiber using different carrier wavelengths. A lightpath is established based on two criteria:

- Finding a route that utilizes resources on lightpath

- Assigning a unique wavelength to the lightpath on that route

In our work, we assume the wavelength continuity constraint and wavelength continuity clash constraint, which requires that the same wavelength to be maintained along the entire lightpath and not more than one wavelength can be assigned to lightpaths sharing a link.
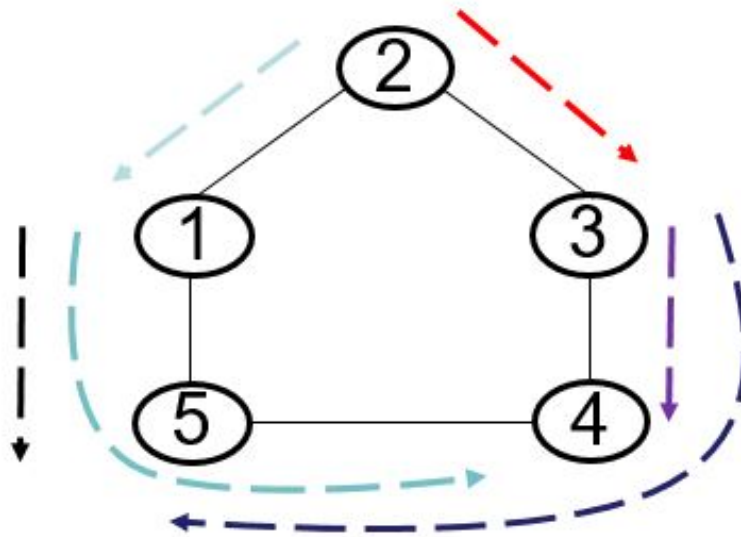
Figure 2.6: Lightpath setup in a 5-node network

Figure 2.6 shows the example of the physical topology with lightpaths. There are five nodes in the network and each node is connected to one or more nodes in the network by bi-directional links called edges. The lightpaths in the following examples are:

Lightpath 1: *node* $1 \rightarrow node$ 4

Lightpath 2: *node* $1 \rightarrow node$ 5

Lightpath 3: *node* $2 \rightarrow node$ 1

Lightpath 4: *node* $2 \rightarrow node$ 3

Lightpath 5: *node* $3 \rightarrow node$ 4

Lightpath 6: *node* $3 \rightarrow node$ 5

2.5   Datacenter

A datacenter can be pictured as a facility that utilized for storage, computing devices and delivering a large amount of information. See Figures 2.7 and 2.8. The resources and

information in a datacenter are served to clients through a network of datacenters, which is referred to as the cloud [19]. With the growth in demand for cloud services, immense amounts of digital content are being created and shared all the time over the system. The content and services of datacenters are replicated over multiple datacenters, with the aim that a client demand can be fulfilled by any datacenter that has the required content. This replication strategy also takes care of the issue of data accessibility in the event of a disaster (earthquake, tsunami, etc.), which may lead to failure of system components (failure if a node or link in the network) [2]. The durability of datacenters against disasters, both natu-
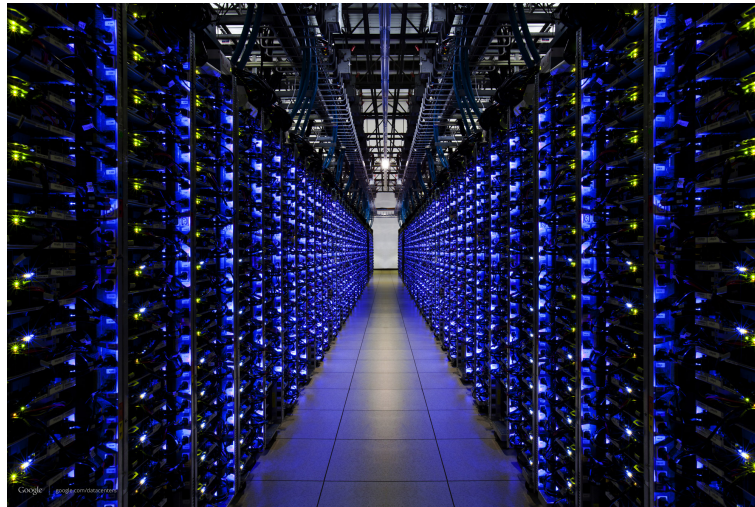
Figure 2.7: Datacenter [20]

ral and man-made attacks are turning into a major threat in designing cloud-based services, hence making cloud network design a significant issue. In our problem we consider that the communication demands are static in nature. Thus we utilize the idea of static lightpath allocation using path protection. When a communication demand for a specific file f, for the destination D is received, our goal is to search for resources to build up two lightpaths a primary lightpath and a backup lightpath.

Research has been started introducing backup datacenters following the anycast principle to cut down bandwidth consumption [21, 22]. Data protection is a major issue in datacenter system, as the failure of a single datacenter should not cause the loss of any file from the
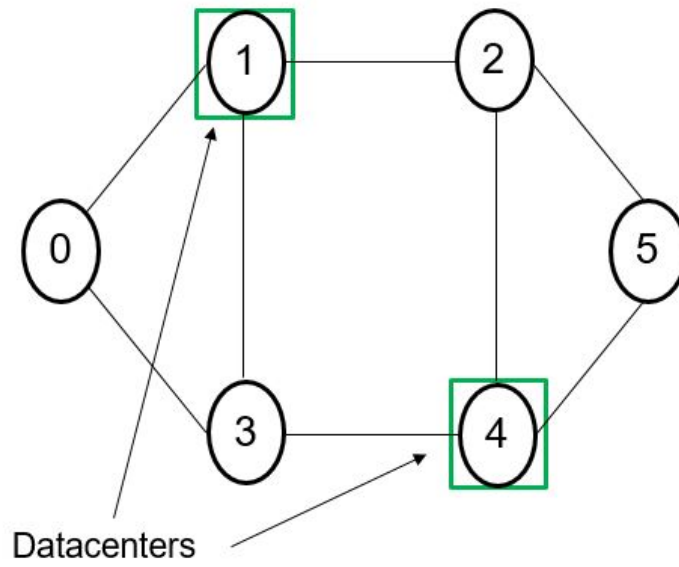
Figure 2.8: Nodes representing datacenters in a network

whole system. As per latest studies, it is crucial that the network supporting such services is volatile to data loss or service interruption, hence making cloud network design a critical issue. In [23] the importance of a certain replica is based on the popularity of data.

2.6   Replication

Cloud computing is a rising standard that gives computing communication and storage resources as a service over a system. Communication resources frequently turn into a bottleneck in service provisioning for many cloud applications. Thus, information replication which brings data (e.g., databases) nearer to information users (e.g., cloud applications) is seen as a promising result.

Keeping up replicas at multiple sites scales up the execution by reducing remote access delay and mitigating single point of failure. However, several infrastructures, such as storage devices and networking devices, are required to maintain data replicas [24]. Moreover, new replicas need to be integrated and any updates made at one of the sites need to be reflected at other locations.

In our work, we are trying to minimize the number of datacenters so that we can utilize the resources. This way, we can use as minimum replicas as possible and with respect to that minimum lightpaths.

2.7 Physical and Logical Topology
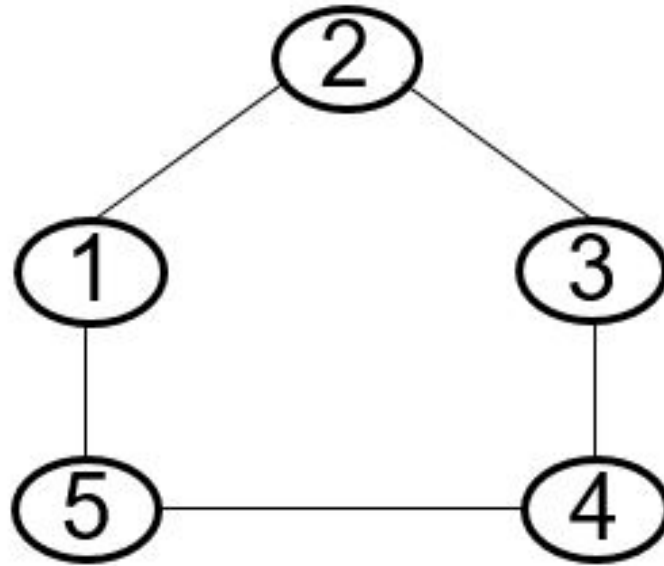
Figure 2.9: Physical Topology

A physical topology refers to the physical connectivity, utilizing optical fibers, between the nodes in the system. In Figure 2.9, we have shown a 5 node network, which shows a map like a structure of the system segments being used in a network. This structure demonstrates the relationship among different network items in which the circles shows the nodes of the network and stable lines demonstrates an actual unidirectional fiber that acts as the connection between two nodes. The physical topology of the network is demonstrated by graph G [N, E], where N is the set of nodes in the network and E is the edges of the network. Each edge of the system (i.e., the bi-directional connection between nodes) is built utilizing two unidirectional links.

The logical topology is the method that is taken after for setting up communication be-
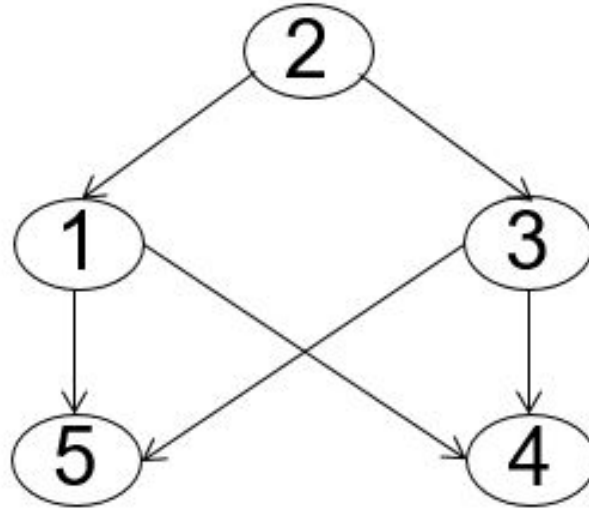
Figure 2.10: Logical Topology

tween a source and a destination node in a network. In this case, the lightpaths are consid-

ered as the edges in the network connecting to the nodes in the physical topology. These

nodes are the same ones in the physical topology. However the lightpath between the

sources and the destination nodes are the logical edges, and this representation of a net-

work is known as a logical or virtual topology [2]. Figure 2.10 shows the logical topology

that is established over the physical topology in Figure 2.9.

2.8 Different Lightpath Allocation Schemes

Network traffic or lightpath requests can be largely divided into two categories: static

lightpath demands and dynamic lightpath demands. The primary difference between them

is the lifetime of these demands. In static lightpath allocation all the demands are known

ahead of time. This is also referred to as permanent (or semi-permanent) lightpath alloca-

tion, because once the request is initiated that lightpath is likely to continue for a compar-

atively long time, weeks, months or years. After some time, if the traffic design changes,

another lightpath can be built up. The RWA relating to a set of static lightpath demands is

typically calculated offline.

However, in dynamic lightpath allocation the requests are not known in advance, they are taken care of as and when they happen. These demands have a specific lifetime i.e., a begin time when the lightpath is set up and an end time when the lightpath is brought down [25], which is generally of much shorter time compared to static lightpaths. These demands are brought down when communication is finished, and the resources allocated to the lightpath can be reused. The dynamic lightpath allocation is again divided into scheduled lightpath demands and ad-hoc lightpath requests. The demands for which the begin time and end time are known in advance (and are often periodic) are called Scheduled Lightpath Demands (SLD). The requests, for which we neither know the begin time nor the period of such requests in advance, are known as Ad-hoc Lightpath Demands (ALD) [26]. In our thesis we concentrate on static lightpath demands.

2.8.1 Static Lightpath Demands (SLD)

In static allocation all lightpaths are arranged ahead of time so either a particular light-path is pre-assigned for every possible source destination pair or the whole set of lightpath demands is known in advance and channel assignments are made for the requests as a whole [27].

In such a scheme, all the requests have to be considered when creating lightpath to support all the user requests. A static scheme guarantees that communication from a source to a destination will always be possible. The lightpaths are established in a way that all the requests fulfilled. In static schemes, none of the requests should be blocked because if it is blocked then we cannot have communication.

2.8.2 Dynamic Lightpath Demands (DLD)

In dynamic allocation, lightpaths are established on demand and are taken down when the communication is finished, and the WDM channels utilized for this communication are recovered for future use in some other communication [28, 29].

2.9 Literature Review

In this section, we discuss in detail, the papers that are directly related to our thesis.

In [1], the authors address the issue of path protection in datacenter systems that offer cloud services. They also focus on the problem of content placement, routing path in the network. The authors have proposed a new integrated Integer Linear Program (ILP) formulation to outline an optical datacenter system. The disaster protection scheme proposed in this paper uses anycast services and provides more security, yet uses less capacity than dedicated single-link failure protection. The authors developed their problem of allotting paths to high bandwidth connections, deciding content replica placement and providing shared path protection against single disaster failure (i.e., multiple disasters caused by a single disaster) for both paths and content. They also proposed the heuristic to derive a solution for a request (s, d) from the LP relaxation of the original ILP because the ILP does not scale well.

In [2], the authors resolve the issue of establishing the primary and backup path in datacenter network with shared backup paths. They focus on the problem of minimizing the average blocking probability (BP) for the new communication requests. For their problem, they are taking dynamic lightpath allocation with pre-defined replicas of the requested files. The authors developed a heuristic to ensure the survivability of a dynamic communication request in the presence of a disaster. Their approach was to find a survivable solution in the case of disaster to fulfill the user requests.

In [19], the authors show the solution for disaster-aware datacenter and content placement problem that intent to minimize the total risk of a network in terms of expected loss of file. They developed a dynamic content-management solution as an enhancement to the initial placement to make the system adaptable to changing situations and disasters. They have used a heuristic approach for solving the problem in datacenter network. They also

considered QoS (Quality of Service)/latency constraints and network resource usage.

In [21], the authors address the issue of fast and composed information backup in geographically distributed optical inter-datacenter networks with a specific goal to enhance the data-transfer efficiency of the regular datacenter backup. To avoid data loss in an optical inter-DC network, a cloud network usually uses different datacenters (DCs) for getting sufficient data redundancy.

In [30], the authors proposed a provisioning scheme to deal with demands for communication that consider the possibility of disasters that may affect multiple edges, nodes, and datacenters. They focussed to minimize the network resources, with respect to wavelength links, used to handle requests that use backup paths in the event of disasters. They used ILP method to share backup paths for communicating.

In [31], the authors considered (i) the issue of resource accommodation before disaster and (ii) re-allotment of resources after a fault. The research used a probabilistic mode for disaster ad proposed a proactive (i.e., before disaster) disaster-aware provisioning schemes with the goal of minimizing the loss in the case of a disaster. The creators also explored a reactive (i.e., after disaster) scheme for re-provisioning the network damaged by the system component failures resulting from the fault.

In [28], the authors provided a unique framework to recognize vulnerable point(s), given a WDM network. A one of a kind component of their framework is its capacity to adapt to a wide range of probabilistic attack and failure models. Their algorithm points in the plane that caused arbitrarily close to expected damage.

In [29], the authors represents joint design of datacenter network (DCN) placement. They tried to minimize the total network cost by formulating joint optimization. They minimized the cost by leveraging between the costs of DCNs and wavelengths.

2.10 Summary of Literature Review

Table 2.1: Literature Review Summary

Reference	Traffic Model	Solution Approach	Types of Failure addressed
Habib et al. [1]	Static	ILP	Single Link Failure
Ruchisree et al. [2]	Dynamic	Heuristic	Multi-Link Failure
Sifat et al. [19]	Dynamic	ILP/Heuristic	Single Link Failure
Mukherjee et al. [21]	Static	ILP	Single Link Failure
Saja et al. [30]	Dynamic	ILP	Multi-Link Failure
Tornatore et al. [31]	Static	ILP	Single Link Failure
Pankaj et al. [28]	Static	Probabilistic model	Multi-Link Failure
Jie Xiao et al. [29]	Static	ILP/Heuristic	Single Link Failure

Chapter 3

STATIC RWA FOR DATACENTER NETWORKS

3.1 Introduction

This chapter introduces the heuristic method we have used to solve the datacenter network design problem. The input to the heuristic is as follows:

1. The network topology defined by the graph $G = (N, E)$ where N denotes the set of nodes in the network including the datacenter nodes and other nodes from which the users can send requests for files and E denotes the bidirectional edges of the network, with each edge (i, j) representing a fiber link $i \leftrightarrow j$ between nodes i and j of the network.

2. Other details of the network including the length of each link $(i, j) \in E$, the set of channels K on each link in the network and the optical reach.

3. The set of requests, $\mathscr{R}$, where each request includes the file f_i requested and the node t requesting the file. We assume that the set of files specified in $\mathscr{R}$ consists of files $f_0, f_1, \ldots, f_m$.

4. the set of disasters $\mathscr{D}$ that need to be considered. Each disaster $d \in \mathscr{D}$ is defined as the set of nodes and edges that will become inoperative if the disaster occurs.

The output produced by the heuristic is as follows:

1. For each file f_i specified in the set of requests $\mathscr{R}$, the set of datacenters which will contain a copy of file f_i.

2. For each request $r \in \mathscr{R}$, representing, say a request from node t for file f_i, the definition of a lightpath from a datacenter storing a copy of f_i to t. Such a lightpath will be defined by

    - a path $s_i \rightarrow \ldots \rightarrow t$,

    - a channel number $k \in K$.

    This lightpath will be called the *primary lightpath* and will be used for communication when the path $s_i \rightarrow \ldots \rightarrow t$ is not affected by any disaster.

3. For each request $r \in \mathscr{R}$, the definition of a lightpath from a datacenter storing a copy of f_i to t. Such a lightpath will be defined by

    - a path $s_j \rightarrow \ldots \rightarrow t$,

    - a channel number $k \in K$.

    This lightpath will be called the *backup lightpath* and will be used for communication when the network encounters a disaster that interrupts the primary lightpath.

Our objective is to reduce the total number of replications of files. In other words, we are minimizing the total number of copies of all the files specified in the set of requests $\mathscr{R}$. The primary and the backup lightpaths must satisfy the wavelength clash constraint, the wavelength continuity constraint and the optical reach constraint.

In an optimal algorithm, we must simultaneously satisfy all the required constraints and determine requested and obtain all the above solutions in an optimal solution. However, with all the constraints and requests, the problem becomes too much complex, and it takes

an enormous amount of time to get the solution. A heuristic[1] to solve the problem is given below.

3.2 The notion of dominant disasters

A disaster d is some event (e.g., earthquake, fire or sabotage) that is characterized by a set of resources $S_d = \{r_1, r_2, ..., r_p\}$, where each resource $r_i \in S_d$ is a component of the network (e.g. fibers, nodes) that will become inoperative when the disaster occurs. We assume that only one disaster can happen at any given point in time. WDM networks are deployed as wide-area networks (such as the USANET) where the distance between a pair of adjacent nodes is typically 100s of kilometers.

In most cases, disasters affect a local region (typically 1- 20 km). For instance, if nodes i and j are connected by a fiber, realizing an edge $i \rightarrow j$ of the network of length 100 km, and a disaster occurs that affects a circular area of radius 10 km, we can have an infinite number of disasters in radius of 10 km on the edge $i \rightarrow j$, each defining an area affected by a disaster. However, these disasters are inextinguishable, so far as their effects are concerned - the edge $i \rightarrow j$ is not usable after any of these disasters and therefore these disasters are equivalent, so far as network design is concerned.

We now consider two disasters d_1 and d_2. As before, d_1 corresponds to any of the disasters that only affects the edge $i \rightarrow j$, so that the set of affected network components for d_1 is $\{i \rightarrow j\}$. The site of disaster d_2 is such that node i itself is affected by the disaster. This means that node i becomes inoperative and we cannot use and edge to and from node i. The set of components affected by d_2 is the set of edges to/from i and the node i itself. Disaster d_2 is, in every sense, more catastrophic compared to d_1 and any scheme to handle disaster d_2 will be sufficient to handle disaster d_1.

---

[1]The heuristic was proposed by S. Al Mamoori, A. Jaekel, and S. Bandyopadhyay

Definition

If the set of network components affected by disaster d_i is a superset of the set of network components affected by disaster d_j, then d_i is said to *dominate d_j*.

In our example above, disaster d_2 dominates d_1. This definition leads to the concept of dominant disaster.

Definition

If a disaster d_k is such that no other disaster dominates d_k, then disaster d_k will be termed a *dominant* disaster.

For simplicity let us consider a wide-area network where

- fibers do not overlap or cross each other,

- two fibers are affected by a single disaster can happen only if the affected fibers are from/to a single node,

- all pairs of nodes are sufficiently far apart so that they are never affected by the same disaster.

In this scenario, it can be readily seen that the dominant disasters are the disasters that affect each node in the network. Our heuristic does not depend on how the dominant disasters are defined. The above assumption makes it easier to generate sample network scenarios of simulation and study.

3.3   A strategy for handling disasters

In our approach, we only deal with an approach to handle dominant disaster. The approach depends on which primary lightpath and which disaster we are considering. As mentioned before, corresponding to every request for communication our algorithm will generate a primary lightpath and a backup lightpath. We will use the idea of lightpath pro-

tection so that the primary lightpath will be used when the network is fault-free. When, due to a disaster, the primary lightpath is unusable, the backup lightpath will be used. In line with the philosophy of path protection both the primary and the backup lightpath will be provisioned at design time and all disasters requiring a backup lightpath will be handled by the same backup lightpath. We now discuss the following cases that we need to incorporate in our design.
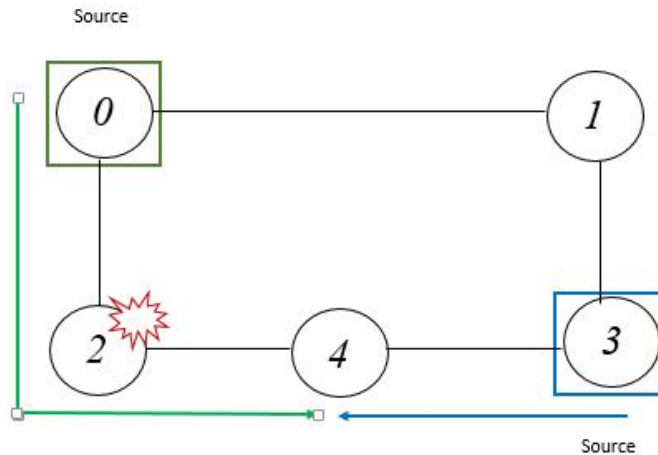
Figure 3.1: Establishment of backup lightpath

1. A disaster may affect the source of the primary lightpath. If this is a possible disaster, in order to handle it, the source of the backup lightpath must be different from that of the primary lightpath. For instance, let the primary lightpath uses the path $0 \to 2 \to 4$. So to handle a disaster at node 0, a backup lightpath path can be defined from $3 \to 4$ is feasible. As explained in Figure 3.1.

2. A disaster may affect the requesting node itself, i.e. a disaster at the destination node for the backup/primary lightpath. In such a case, the communication cannot proceed at all. In summary, we need not consider the case where a disaster affects the requesting node.

3. A disaster may affect an intermediate node in the path used by the primary lightpath. An example is a disaster at node 2. To handle such a disaster, the backup lightpath can be established from the same source node) but through different intermediate nodes (completely avoiding any in-ongoing or out-going link from the failed node in primary lightpath). For instance, if there can be no disaster involving node 0, a possible path for the backup lightpath could use the path $0 \rightarrow 1 \rightarrow 3 \rightarrow 4$. If however both node 0 and node 2 can be affected by (possibly different) disasters, we must use a backup path such as $3 \rightarrow 4$ which will avoid disasters affecting both nodes 0 and 2. This can be seen in Figure 3.2.
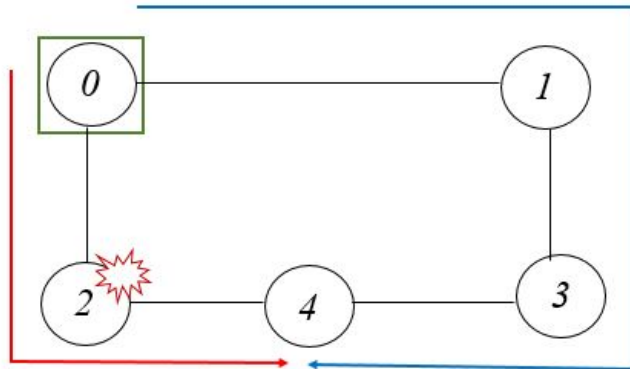
Figure 3.2: Establishment of backup lightpath to handle disaster on an intermediate node

3.4   The heuristic to be studied

In the heuristic we have assumed that the capacity of the datacenters are unlimited. This means that the replication strategy for any file f_i is independent of the replication strategy

for any other file $f_j, i \neq j$. We assume there is a database *DB* that will contains details of the channels already in use by each edge $(i, j) \in E$. At the start, this database is empty since the network is not initially supporting any lightpath. The steps of the heuristic are as follows:

Step 1) Partition the requests in $\mathcal{R}$ into subsets $\mathcal{R}_0, \mathcal{R}_1, \ldots, \mathcal{R}_m$ of requests, so that subset $\mathcal{R}_i$ contains all requests for file f_i.

Step 2) Repeat Steps $3 - 6$ for all $i, 0 \leq i \leq m$.

Step 3) Solve the RWA problem to find the RWA for the primary lightpaths to be used by all the requests in subset $\mathcal{R}_i$. This will also include solving the file placement problem for file f_i.

Step 4) Update the database *DB* with the channels and the paths corresponding to the primary lightpaths to handle requests in $\mathcal{R}_i$.

Step 5) Solve the RWA problem to find the RWA for the backup lightpaths to be used by all the requests in subset R_i. This will also include solving the file placement problem for file f_i, taking into account the existing locations of file f_i fixed in Step 3.

Step 6) Update the database *DB* with the channels and the paths corresponding to the backup lightpaths to handle requests in $\mathcal{R}_i$.

After Step 3 is over, request $r \in \mathcal{R}_i$ for communicating file f_i to node t will have a communication strategy to be used when the network is fault-free. This communication strategy will involve a path $s \rightarrow \ldots \rightarrow t$ for the primary lightpath and means that this step will decide to save one copy of file f_i in datacenter s. The intent here is to minimize the number of copies of file f_i, so that, each copy of file f_i will be used as the source of multiple primary lightpaths. Since Step 3 is solved many times, once for each file considered, in general, the

network is supporting a number of primary and backup lightpaths generated in previous iterations. When Step 3 is carried out for the ith iteration, the RWA mentioned in Step 3 must take into account all channels allotted for the requests in sets $\mathscr{R}_0, \mathscr{R}_1, \ldots, \mathscr{R}_{i-1}$ and avoid wavelength class constraints. The RWA for all requests in R_i must avoid wavelength clash constraints and the optical reach constraints mentioned earlier in Chapter 1. We do this using the database *DB*.

After Step 5 is over, request $r \in \mathscr{R}_i$ for communicating file f_i to node t will have a communication strategy to be used when the network encounters a disaster that disrupts the primary lightpath for request r. This communication strategy will involve a path $\hat{s} \rightarrow \ldots \rightarrow t$ for the backup lightpath and means that this step will decide to save one copy of file f_i in datacenter $\hat{s}$. This step takes into account the placement strategy for file f_i used in Step 3. Since we intend is to minimize the total number of copies of file f_i used for the primary and the backup lightpaths, each copy of file f_i will, in general, be used as the source of multiple primary and backup lightpaths. If possible, the source $\hat{s}$ for the backup lightpath for request r will be a node that is already hosting a copy of f_i. This is to allow us to reach our objective of minimizing the total number of copies of file f_i. The remaining comments for Step 5 are similar to those mentioned for Step 3 above.

3.5   Concept of Virtual Node

One problem we encounter in this design is that the source of a lightpath is not specified in the request. The request simply specifies the file f_i to be retrieved. Our algorithm has to determine which node will be the source of the lightpath and store a copy of the requested file f_i at that node. Potentially any node in the network may be a source for the primary (or backup) lightpath for handling any request ($r \in \mathscr{R}$). This means that the well-known network flow algorithm for single commodity network flow [32] is not immediately
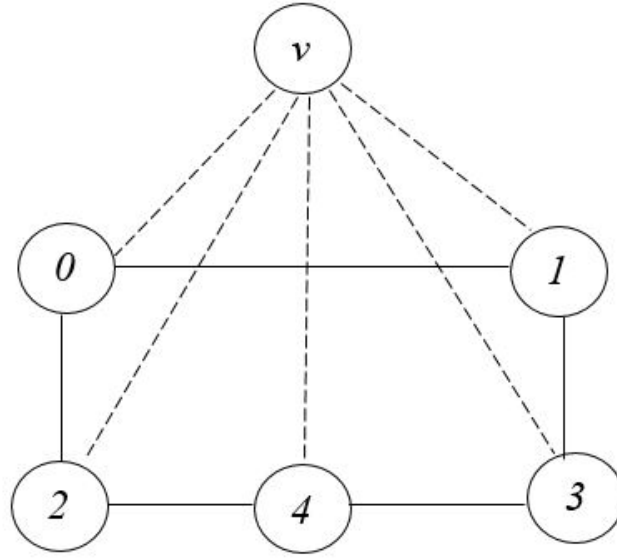
Figure 3.3: Concept of Virtual node

applicable.

The concept of virtual node is important in this context. In Step 3 of the ith iteration, the objective is to set up primary lightpaths for file f_i. At this point, any datacenter in the network is a potential candidate. We visualize a *virtual node v* which does not exist as a node in set N of nodes of the network. Node v is connected by a *virtual edge* of length 0 to *all* nodes in the network. When considering request $r \in \mathcal{R}$ to find a path $v \to s \to \ldots \to t$ from some node s that must store a copy of file f_i, we set up a path $s \to \ldots \to t$ from v to t. Since all edges from this virtual node has length 0, it does not affect the optical reach constraint. Simply by deleting the virtual edge $v \to s$, we can get path from s to t. This also makes it possible for us to specify our objective (Reduce the number of copies of file f_i)

For example, in Figure 3.3, there is a network of 5 physical (actually present in the network) nodes (0, 1, 2, 3, and 4) and a virtual node v. The virtual node is not part of the original network but is a concept to enable the use of a simple single commodity network flow algorithm. Figure 3.3 shows that the virtual node is connected to all the nodes in the network. One path from v to node 4 is $v \to 0 \to 2 \to 4$. If we select this path from virtual

node v, it corresponds to the actual path $0 \rightarrow 2 \rightarrow 4$.

3.6 Preamble

We now describe how we implemented our heuristic. The only steps which are complicated are steps 3 and 5. The remaining steps are trivial that we will not discuss. In the ith iteration, during Step 3, we set up, if possible, the primary lightpaths for set of requests $\mathcal{R}_i$ for file f_i. For this, we have developed an ILP formulation which we will call ILP-1.

During Step 5, we run the ILP formulation which we will call ILP-2 to set up the backup lightpaths. As explained in the heuristic above, ILP-2 takes advantage of the sites which already contain a copy of file f_i. From now on, for simplicity, we will drop the subscript i from $\mathcal{R}_i$ and f_i.

3.7 Integer Linear Programming (ILP)

formulations used in heuristic In this section we outline the RWA algorithm that has been proposed to find a feasible primary path in disaster free situation and feasible backup paths to handle disaster $d \in D$. We will use the following notation in formulations ILP-I and ILP-II:

N : the set of nodes (including the virtual node s).

E : the set of directed edges of the network. If i and j are nodes of the fiber network (including the datacenters), edge $(i \rightarrow j) \in E$ represents a fiber from node i to node j. Set E also includes the virtual edges from the virtual node s to each node in the network.

D : the predefined set of dominant disasters.

D^r : the set of dominant disasters that interrupts the primary lightpath for request r.

29

$(f,t):$ a request for file f, where t is the node requesting file f.

$R:$ the set of all requests for file f.

$E^r:$ the set of directed edges of the network which are not disrupted by disasters in D^r

$d_{max}:$ the optical reach.

$\ell_{ij}:$ the length of the fiber $(i,j) \in E$.

$K:$ set of channels per fiber.

$E^d:$ the set of edges disrupted due to disaster $d \in D$.

$c_{ij}^k:$ a constant for all channel $k \in K$ and edge $(i,j) \in E$ where

$$c_{ij}^k = \begin{cases} 1 & \text{if channel } k \text{ is used on edge } (i,j), \text{ either for} \\ & \text{a primary lightpath or for a backup lightpath} \\ 0 & \text{otherwise.} \end{cases}$$

$b_i:$ a constant for all node $i \in N$ where

$$b_i = \begin{cases} 1 & \text{if node } i \text{ is already used to save a copy of file } f \\ 0 & \text{otherwise.} \end{cases}$$

$x_{ij}^r:$ a binary variable for all edge $(i,j) \in E$ and all request $r \in R$ where

$$x_{ij}^r = \begin{cases} 1 & \text{if edge } (i,j) \text{ is used by the primary lightpath for request } r, \\ 0 & \text{otherwise.} \end{cases}$$

$y_{ij}^r:$ a binary variable for all edge $(i,j) \in E$ and all request $r \in R$ where

$$y_{ij}^r = \begin{cases} 1 & \text{if edge } (i,j) \text{ is used by the backup lightpath for request } r, \\ 0 & \text{otherwise.} \end{cases}$$

$w_i:$ a binary variable for all node $i \in N$ where

$$w_i = \begin{cases} 1 & \text{if edge } (0,i) \text{ is used by the primary lightpath, so that the first} \\ & \text{internal node in any path from the virtual node is } i \\ 0 & \text{otherwise.} \end{cases}$$

μ^{kr} : a binary variable for all channel $k \in K$ and all request $r \in R$ where

$$\mu^{kr} = \begin{cases} 1 & \text{if channel } k \text{ is used by the primary lightpath for request } r, \\ 0 & \text{otherwise.} \end{cases}$$

μ_{ij}^{kr} : a bounded variable for all channel $k \in K$, all request $r \in R$ and all edge $(i,j) \in E$

which is forced by the constraints in the ILP so that

$$\mu_{ij}^{kr} = \begin{cases} 1 & \text{if edge } (i,j) \text{ } and \text{ channel } k \text{ is used} \\ & \text{by the primary lightpath for request } r, \\ 0 & \text{otherwise.} \end{cases}$$

3.7.1 Formulation of ILP-1

Objective: Minimize

$$\sum_{i: i \in N} w_i \tag{3.1}$$

Subject to:

1. Enforce, for all node $i \in N$ and for all request $r \in R$, flow conservation on the path used by the primary lightpath to handle the request r is to communicate file f to node t.

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

2. Ensure that the length of the route used by each of the primary lightpaths is less than the optical reach d_{max}.

$$\sum_{j:(i,j) \in E} \ell_{ij} \cdot x_{ij}^r \leq d_{max} \quad \forall (i,j) \in E, r \in R \tag{3.3}$$

3. Compute the weight w_i for node i. In other words, determine whether node i has to be the source of any communication.

$$w_i \geq x_{si}^r \quad \forall (s,i) \in E, r \in R \tag{3.4}$$

31

$$w_i \leq \sum_{r \in R} x_{si}^r \qquad (3.5)$$

4. Ensure that exactly one channel k is used for the primary lightpath for request $r \in R$.

$$c_{ij}^k \cdot x_{ij}^r + \mu^{kr} \leq 1 \quad \forall (i,j) \in E, k \in K, r \in R \qquad (3.6)$$

$$\sum_k \mu^{kr} = 1 \quad \forall r \in R \qquad (3.7)$$

5. Determine the value of μ_{ij}^{kr},

$$\mu_{ij}^{kr} \leq \mu^{kr} \quad \forall k \in K, r \in R \qquad (3.8)$$

$$\mu_{ij}^{kr} \leq x_{ij}^r \quad \forall (i,j) \in E, r \in R \qquad (3.9)$$

$$\mu_{ij}^{kr} \geq \mu^{kr} + x_{ij}^r - 1 \quad \forall (i,j) \in E, k \in K, r \in R \qquad (3.10)$$

6. If the primary lightpaths for requests r_1 and r_2 share any edge $(i,j) \in E$, then the channel alloted to request r_1 must be different from the channel alloted to request r_2.

$$\mu_{ij}^{kr_1} + \mu_{ij}^{kr_2} \leq 1 \quad \forall (i,j) \in E, k \in K, r_1, r_2 \in R \qquad (3.11)$$

3.7.2   Justification of ILP-1

The objective is to minimize the total number of datacenters used for communicating the requests from the user. This is calculated by summation of the number of datacenters (w_i).

Constraint 3.2, corresponds to the flow balance conservation rule [32] for the primary path for the fault-free communication. x_{ij}^r is a binary variable (0/1) and its value becomes 1 if the edge $(i, j) \in E$ used in the path for the primary lightpath corresponding to the request; otherwise x_{ij}^r is 0.

Constraint 3.3, ensures that the length of the route used by each of the primary lightpaths is less than the optical reach d_{max}. This means that, summation of the lengths for all the edges used for a request $r \in R$ must not exceed the maximum distance optical signal can travel.

Constraints 3.4 & 3.5, compute the weight w_i for node i. That means we are measuring whether node i is the source of any communication in the network.It is also a binary variable. If our algorithm choose w_i as a datacenter than for some request r $x_{si}^r = 1$, so that constraint 3.4 means $w_i \geq 1$. In that case, the RHS for constraint 3.4 is at least 1, so that $w_i \leq$ some positive number. Since w_i is a binary variable, the only solution is $w_i = 1$. If no lightpath selects node i as the first node after s, constraints 3.4 & 3.5 become $w_i \geq 0$ and $w_i \leq 0$. The value of w_i in this case will be 0. In our heuristic we are trying to minimize the weights of nodes so that we will utilize the resources.

Constraints 3.6 & 3.7, make sure that exactly one channel k is used for the primary lightpath for a request $r \in R$. μ^{kr} is a binary variable (0/1) and its value becomes 1 if the channel k is used for the request $r \in R$. Its value should be exactly one. That means we are enforcing wavelength continuity constraint in our approach.

Constraints 3.8, 3.9 & 3.10, determine the value of μ_{ij}^{kr}. It simply states that if channel k is used for the primary lightpath for request r (i.e., $\mu^{kr} = 1$) and request r used edge $i \rightarrow j$, (i.e., $x_{ij}^r = 1$) then $\mu_{ij}^{kr} = 1$; otherwise $\mu_{ij}^{kr} = 0$. It is important to note that μ_{ij}^{kr} is bounded variable, not a binary variable.

Constraint 3.11, enforces the wavelength clash constraint with bounded variable μ_{ij}^{kr}. It states that if the primary lightpaths for requests r_1 and r_2 share any edge $(i, j) \in E$, then the

channel allocated to request r_1 must be different from the channel allocated to request r_2.

3.7.3 The formulation of ILP-2

Objective: Minimize

$$\sum_{i:i\in N} w_i \cdot (1 - b_i) \tag{3.12}$$

Subject to:

1. Enforce, for all node $i \in N$ and for all request $r \in R$, flow conservation on the path used by the backup lightpath to handle the request r is to communicate file f to node t.

$$\sum_{j:(i,j)\in E^r} y_{ij}^r - \sum_{j:(j,i)\in E^r} y_{ji}^r = \begin{cases} 1 & \text{if } i = s, \\ -1 & \text{if } i = t, \\ 0 & \text{otherwise.} \end{cases} \tag{3.13}$$

2. Ensure that the length of the route used by each of the backup lightpaths is less than the optical reach d_{max}.

$$\sum_{j:(i,j)\in E^r} \ell_{ij} \cdot y_{ij}^r \leq d_{max} \quad \forall (i,j) \in E^r, r \in R \tag{3.14}$$

3. Compute the weight w_i for node i. In other words, determine whether node i has to be the source of any communication.

$$w_i \geq y_{si}^r \quad \forall (s,i) \in E^r, r \in R \tag{3.15}$$

$$w_i \leq \sum_{r\in R} y_{si}^r \tag{3.16}$$

4. Ensure that exactly one channel k is used for the backup lightpath for request $r \in R$.

$$c_{ij}^k \cdot y_{ij}^r + \mu^{kr} \leq 1 \quad \forall (i,j) \in E^r, k \in K, r \in R \tag{3.17}$$

34

$$\sum_k \mu^{kr} = 1 \quad \forall r \in R \tag{3.18}$$

5. Determine the value of μ_{ij}^{kr},

$$\mu_{ij}^{kr} \leq \mu^{kr} \quad \forall k \in K, r \in R \tag{3.19}$$

$$\mu_{ij}^{kr} \leq y_{ij}^{r} \quad \forall (i,j) \in E^{r}, r \in R \tag{3.20}$$

$$\mu_{ij}^{kr} \geq \mu^{kr} + y_{ij}^{r} - 1 \quad \forall (i,j) \in E^{r}, k \in K, r \in R \tag{3.21}$$

6. If the backup lightpaths for requests r_1 and r_2 share any edge $(i,j) \in E^r$, then the channel alloted to request r_1 must be different from the channel alloted to request r_2.

$$\mu_{ij}^{kr_1} + \mu_{ij}^{kr_2} \leq 1 \quad \forall (i,j) \in E^{r}, k \in K, r_1, r_2 \in R \tag{3.22}$$

3.7.4 Justification of ILP-2

The objective is to minimize the total number of datacenters used for communicating the requests from the user for the backup lightpaths. This is calculated by summation of the number of datacenters (w_i). If node i already used by the primary lightpath, $b_i = 1$. For such nodes, the cost of using the node as a source for a backup lightpath is ignored by the objective function.

Constraint 3.13, corresponds to the flow balance conservation rule for the backup path for the fault-free communication. y_{ij}^{r} is a binary variable (0/1) and its value becomes 1 if the edge $(i,j) \in E^{r}$ used in the backup path for the accommodating request otherwise y_{ij}^{r} is 0. In this constraint, the topology to be used is E^{r}, the topology that survives after removing all edges affected by all the disasters in D^{r}.

The explanations for all the other constraints are omitted since these constraints are identical to those in ILP-I.

3.7.5 Analysis of the ILP

There are four binary variables in the ILP as follows:

There is one variable $x_{ij}^r(y_{ij}^r)$ for each edge $(i,j) \in E$ (E^r) and for a request $r \in R$ for primary(backup) lightpath. There is one variable w_i for all nodes in the network where $i \in N$. One more variable is μ^{kr} for each combination of where $k \in K$ and $r \in R$. Hence, the formulation has $(2|R||E| + |N| + |R||K|)$ binary variables.

There is only one continuous variable in the ILP. That variable is μ_{ij}^{kr} for each combination of where each edge $(i,j) \in E$, channel $k \in K$ and request $r \in R$. Thus the formulation has $(|R||K||E|)$ continuous variables.

Here notice that, the continuous variable is the crucial factor determining the complexity of the algorithm.

Table 3.1: Analysis of binary variables on various topologies

Comparison of Binary and Continuous Variable			
Nodes	Edges	Binary Variables	Continuous Variables
11	63	4027	15120
14	56	3614	13440
24	110	6864	25890

Chapter 4

EXPERIMENTATION AND RESULT

In this section, network standards, simulation and outcomes achieved using the proposed Heuristic method will be explained. Simulation results for the proposed methodology have been shown under the various scenarios in the network system. Results noted in the tables are the average of 3 different runs for every topology. The algorithm is capable of generating result for practical sized networks. All the experiments have been conducted on IBM ILOG CPLEX [33], using Intel Core i7-4510U CPU 2 GHz processor.

4.1 Simulation Setup

4.1.1 Network Setup
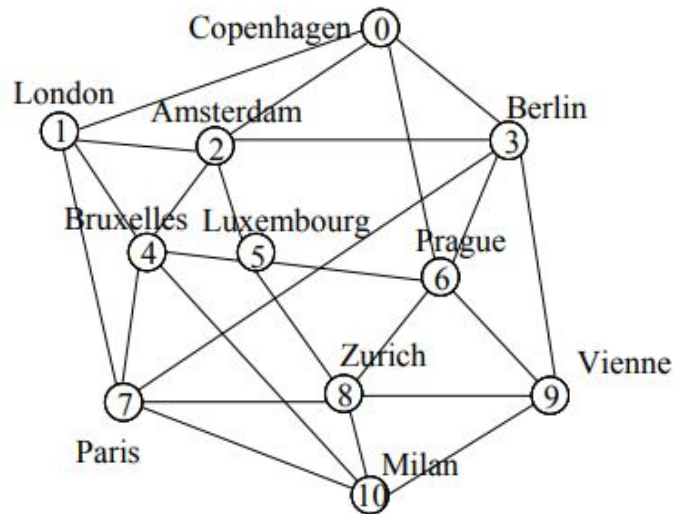

Figure 4.1: COST-239 network (11-node topology)

Several well-known network topologies have been considered, ranging from 11 to 24

nodes.

The topologies used for the experiment are as followed:

1. 11-node COST-239 network (Fig. 4.1)
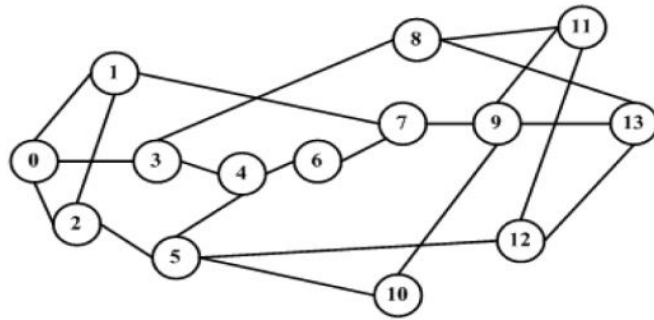
2. 14-node NSFNET (Fig. 4.2)

3. 24-node USANET (Fig. 4.3)
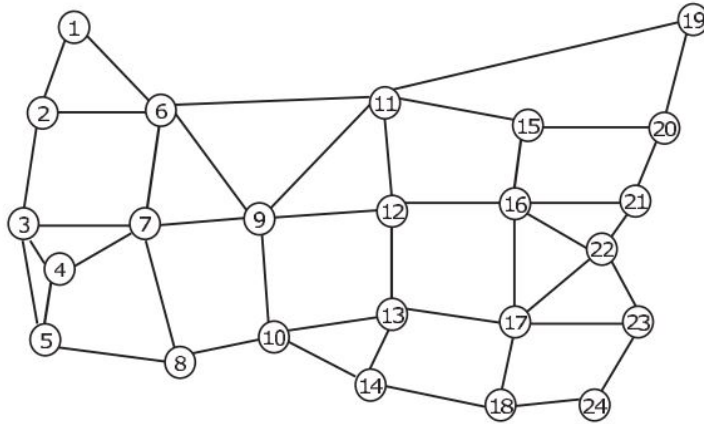

Figure 4.2: NSFNET network (14-node topology)

Figure 4.3: USANET network (24-node topology)

4.1.2 Algorithm inputs

The following parameters were given as input to our RWA calculation:

1. **Network topology:** The network topology comprises of a set of nodes (N) and fiber links (E) associating the nodes of the network.

2. **Request-file:** The request-files are comprised of many communication requests from a source to the destination node. The source is a document (file f_i) located at some datacenter node. On the other hand, a destination is a node that requests a replica of files f_i. Every request was created arbitrarily from the arrangement of 3 files (f_i).

3. **Disaster nodes:** We consider all specified disasters at the same time. All the edges entering or leaving from the specified disaster nodes are not accessible for processing communication requests.

4.2 Performance evaluation of proposed approach

For the simulations discussed in this Section, the network topology, set of requests, number of channels and disaster scenarios have been provided as an input. For each topology, we have considered 8 and 16 channels per edge, and two disasters occurring at node 1 and node 9. We have taken different sets of requests ranging in size from 30 requests to 90 requests, for 3 different files. Each value reported in the Tables is achieved by taking the average of 3 experimental runs.

Tables 4.1 and 4.2 show the simulation results for 11-node COST-239 topology (consisting of 52 one-directional edges as well as 11 virtual edges) with 8 and 16 channels per fiber respectively. The results for phase-1 indicate that the number of datacenters required increase with the number of requests. This happens because our algorithm enforces the

Table 4.1: Comparison of resource usage in 11-Node network (8 channels)

Network Topology	Disaster	Files	No. of request	Primary		Backup	
				Datacenter	Wavelength-link	Datacenter	Wavelength-link
11-node (Channel 8)	[1,9]	File1	10	[5,9]	27	[5,7]	20
		File2	10	[4,9]	15	[5,7]	12
		File3	10	[6,8]	22	[5,11]	15
		File1	15	[6,9]	36	[6,10]	18
		File2	15	[6,9]	20	[8,10]	25
		File3	15	[4,6]	25	[4,8]	14
		File1	20	[4,6,9]	49	[5,7,8]	48
		File2	20	[2,8,9]	34	[3,5,10]	38
		File3	20	[3,5,9]	22	[5,8,10]	15

optical reach constraints for lightpath allocation. As more destination nodes are included more datacenters are needed, since existing ones may not be within the optical reach new destination nodes. We also note that it is possible to accommodate the same number of requests with fewer datacenters, if we have more available channels per fiber. The number of wavelength-links needed is higher with 16 channels, because having fewer datacenters typically lead to longer paths for servicing each communication request.

Phase-2 is only run when some disaster is affecting one or more of the primary lightpaths. In phase-2, we use the output of the phase-1 (wavelength-links database) and repeat for the request sets of all the files. So some of the channels are already utilized by the phase-1. We have reported that on average phase-2 also needs two to three additional datacenters.

Table 4.2: Comparison of resource usage in 11-Node network (16 channels)

Network Topology	Disaster	Files	No. of request	Primary		Backup	
				Datacenter	Wavelength-link	Datacenter	Wavelength-link
11-node (Channel 16)	[1,9]	File1	15	[9]	37	[2,10]	31
		File2	15	[9]	21	[4,5]	37
		File3	15	[9]	24	[5,10]	25
		File1	20	[6,9]	49	[5,10]	47
		File2	20	[5,9]	26	[5,11]	25
		File3	20	[7,9]	24	[5,11]	34
		File1	30	[2,9]	69	[5,11]	68
		File2	30	[5,9]	48	[5,11]	35
		File3	30	[2,9]	40	[4,8]	60

Figure 4.4 shows how the number of datacenter nodes (DCNs) needed varies with the request set size for the 11-not topology. The first two bars in each group indicates the
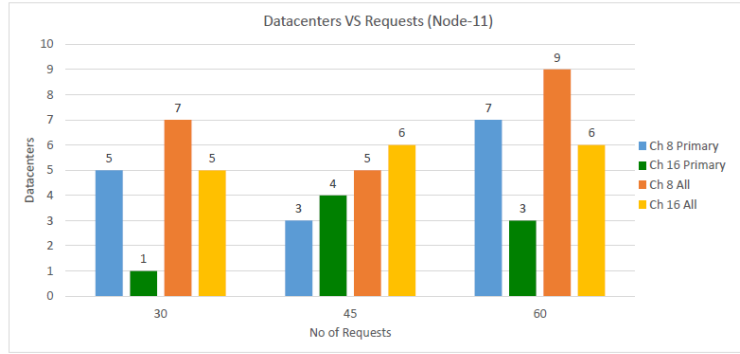
Figure 4.4: Number of datacenters needed vs size of request set for 11-node network
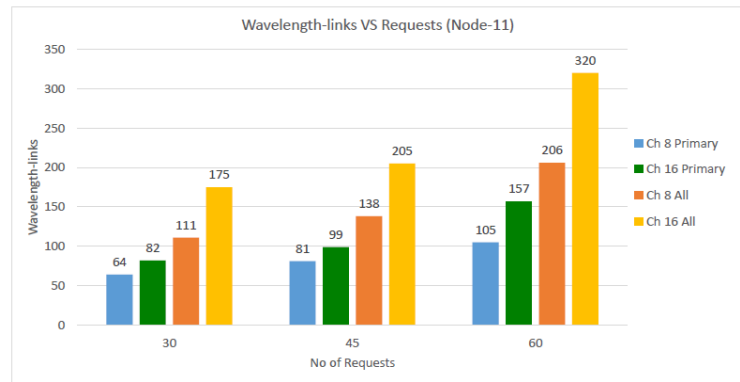
Figure 4.5: Number of wavelength-links needed vs size of request set for 11-node network

number of DCNs needed for primary paths only using 8 and 16 channels. The last two bars indicate the corresponding values when considering both primary and backup lightpaths. Although with 16-channels there was usually a reduction in the DCNs, in some cases we observed a slight increase. This likely due to the distribution of the destination nodes. The number of wavelength-links required increased with the number of requests, as shown in Figure 4.5.

Tables 4.3 and 4.4 show the results for the 14-node NSFNET topology with 8 and 16 channels per fiber respectively. For this topology, we considered the optical reach to be 3000 *km* as some of the links in the topology are more than 2000 *km* long. We note that for some of the simulation cases, no additional resources were required to service backup lightpaths. This is because in those cases none of the primary paths were affected by the

Table 4.3: Comparison of resource usage in 14-Node network (8 channels)

Network Topology	Disaster	Files	No. of request	Primary		Backup	
				Datacenter	Wavelength-link	Datacenter	Wavelength-link
14-node (Channel 8)	[1,9]	File1	10	[6,7]	25	-	-
		File2	10	[5,8]	25	[6,8]	22
		File3	10	[3,10]	23	[5,12]	17
		File1	15	[2,5,12]	39	-	-
		File2	15	[4,5,8]	44	[6,7,11]	45
		File3	15	[3,5,9]	46	[4,5,8,11]	23
		File1	20	[1,2,5,14]	81	[2,5,11,14]	55
		File2	20	[3,6,7,10]	75	[3,4,7,12]	41
		File3	20	[3,5,8,12]	32	[3,4,5,6,7,8,14]	17

specified disasters. It is possible that under a different set of disasters backup resources would be needed for these communication requests. In other respects, the results followed a similar pattern to the 11-node topology.

Table 4.4: Comparison of resource usage in 14-Node network (16 channels)

Network Topology	Disaster	Files	No. of request	Primary		Backup	
				Datacenter	Wavelength-link	Datacenter	Wavelength-link
14-node (Channel 16)	[1,9]	File1	15	[6]	38	-	-
		File2	15	[6]	23	-	-
		File3	15	[6,14]	20	-	-
		File1	20	[2,8]	52	[8,10]	44
		File2	20	[5,6]	50	[4,8]	36
		File3	20	[5,7]	36	[6,10]	41
		File1	30	[6,8]	64	-	-
		File2	30	[6,10]	57	-	-
		File3	30	[5,10]	62	-	-

Figures 4.6 and 4.7 show the resource usage for the 14-node topology. For this topology, there is a consistent reduction in the number of DCNs needed with 16 channels, compared to 8 channels. The number of DCNs is expected to increase when backup lightpaths are also considered. But we observe that in some cases, there was no such increase. This is because, for these cases, none of the primary paths were affected by the specified disasters. Hence, no backup lightpaths were needed. The number of wavelength links typically increased, with a reduction in DCNs, for the same number of requests. However, there were a few exceptions. This is because, with 8 channels, some of the shorter paths did not have any available wavelengths.

Figure 4.6: Number of datacenters needed VS size of request set for 14-node network

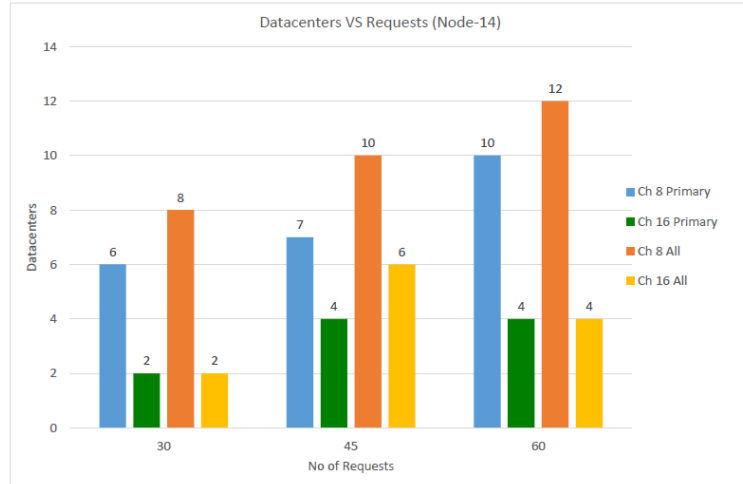Table 4.5: Comparison of resource usage in 24-Node network (8 channels)

Network Topology	Disaster	Files	No. of request	Primary		Backup	
				Datacenter	Wavelength-link	Datacenter	Wavelength-link
24-node (Channel 8)	[1,9]	File1	15	[7,16,17]	37	-	-
		File2	15	[2,16,23]	24	-	-
		File3	15	[7,13,21]	37	-	-
		File1	20	[3,13,21]	51	-	-
		File2	20	[7,16,21]	46	-	-
		File3	20	[7,12,17]	53	-	-
		File1	30	[7,9,18,21]	80	[2,11,13,21]	103
		File2	30	[3,13,14,21]	11	-	-
		File3	30	[7,9,16,24]	58	[4,8,16,18]	54

Table 4.5 and 4.6 illustrate the simulation results for the 24-node topology with 8 channels and 16 channels respectively. The results for this topology follow the same patter as for the 11-node and 14-node topologies.

Figures 4.8 and 4.9 show the DCN and wavelength-link usage for the 24-node topology. For 45 requests we observed no differences in terms of the DCN, when increasing the number of channels from 8 to 16. Also, none of the primary lightpaths were affected by the specified disasters, so no additional resources were needed for backup lightpaths.
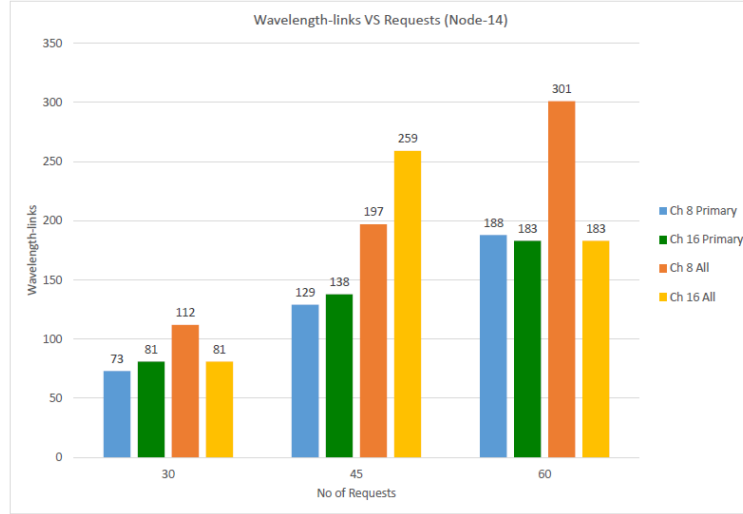
Figure 4.7: Number of wavelength-links needed VS size of request set for 14-node network

Table 4.6: Comparison of resource usage in 24-Node network (16 channels)

Network Topology	Disaster	Files	No. of request	Primary		Backup	
				Datacenter	Wavelength-link	Datacenter	Wavelength-link
24-node (Channel 16)	[1,9]	File1	15	[7,16,17]	35	-	-
		File2	15	[6,16]	28	-	-
		File3	15	[2,10,21]	35	-	-
		File1	20	[7,12,21]	49	-	-
		File2	20	[7,11,21]	47	-	-
		File3	20	[7,12,21]	39	-	-
		File1	30	[2,9,11]	77	[2,13,21]	52
		File2	30	[5,13,21]	64	-	-
		File3	30	[7,9,22]	63	[7,13,17]	71

4.3 Comparison with previous work

We compared the results of our proposed approach to the disaster-zone failure (DZF) ILP presented in [1], adapted for dedicated single link failure. We used the 11-node COST-239 topology with 32 channels per fiber link and different request sets consisting of 20, 25 and 30 requests. Figures 4.10 and 4.11 show the results in terms of the number of wavelength-links used and the number of DCNs needed respectively. We see that with 20 requests, the DZF requires 3 DCNs and 55 wavelength-links, while our proposed approach uses more wavelength-links, but fewer DCNs. Simulation results demonstrate proposed approach can achieve significant reduction in number of datacenters, compared to DZF
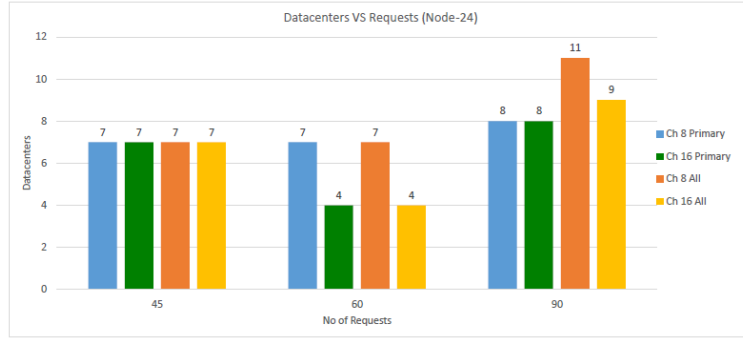
Figure 4.8: Number of datacenters needed vs size of request set for 24-node network
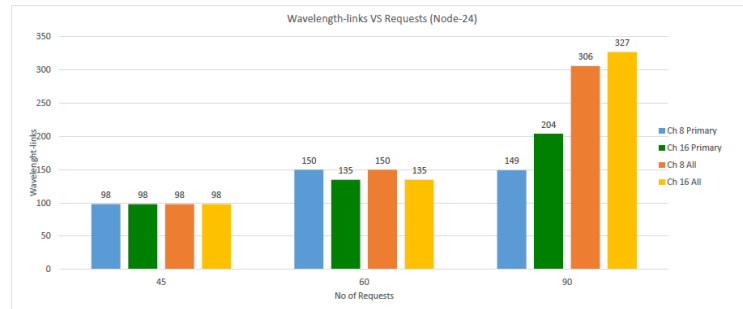

Figure 4.9: Number of wavelength-links needed vs size of request set for 24-node network

(disaster-zone failure) based algorithm approach by approximately $20 - 30\%$. The same pattern is observed for the other request sets as well. We note that one of the factors resulting in higher wavelength usage in our approach is that we enforce the *wavelength continuity constraint*. This leads to more wavelengths being needed compared to networks where this constraint is not enforced. Even though we use more wavelength-links, our algorithm achieves lower cost because wavelength converters are not needed at every node. Furthermore, we also consider the optical reach, so lightpaths longer than the optical reach are not allowed. This eliminates the need for expensive optical regenerators.
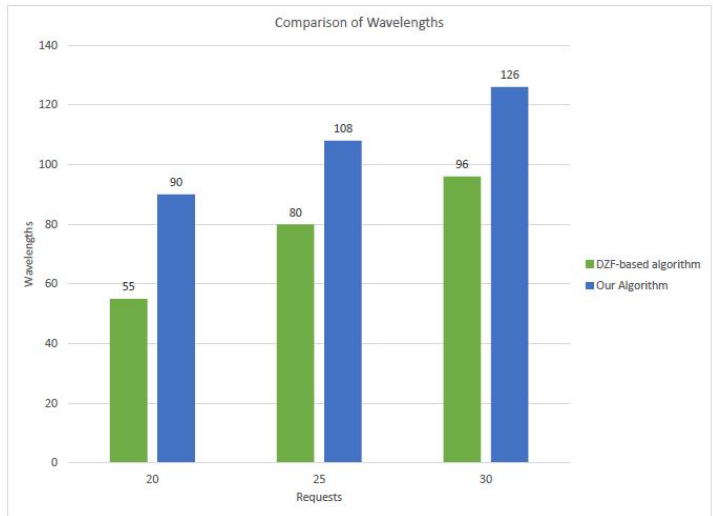
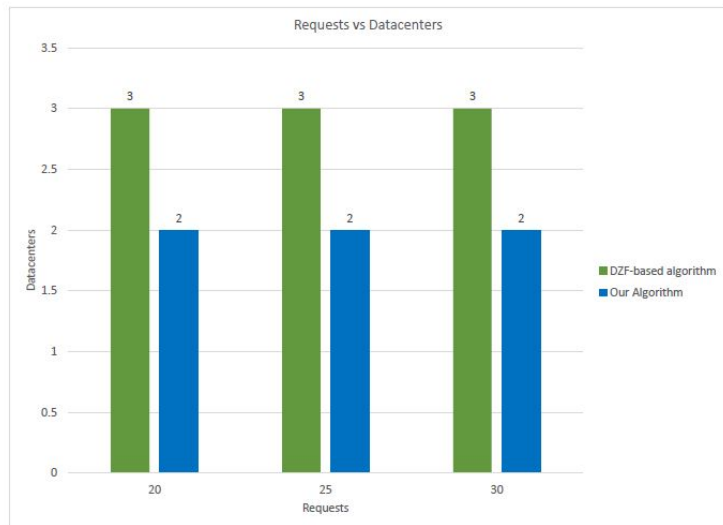Figure 4.10: Wavelengths VS Requests (COST-239 ch-32)

Figure 4.11: Datacenters VS Requests (COST-239 ch-32)

Chapter 5

CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

Storage of a large amount of important and sensitive information is heavily relied upon datacenters (DC) and this data is then transmitted to end users through various communication channels. Catastrophic events and man-made attacks cause substantial harm to these communication networks. During such disasters, various geographical areas get affected, which leads to failure of communication networks either through a node failure or broken optics cable. There is an ever-increasing need to have a communication network that is more resilient to such disasters or responsive in such situations. Due to the failure of specific sections of the network, more and more requests might need to be transferred/communicated, which results in increased traffic on channels. In our work here, we have tried to apply a heuristic solution to the problem, which to the best of our knowledge is the first attempt of this kind. Our aim is to minimize the total number of replicas of files in a network when the network uses static lightpath allocation and has pre-defined optical reach. In our scheme, backup lightpaths were only to be used in case of disaster occurs on the primary lightpaths. Our simulations with different sets of communication requests ran on various network topologies. The final report gives needed network performance under different scenarios.

5.2   Future work

With frequent occurrences of natural disasters and increased directed attacks, network vulnerability to multiple cascading and correlated failures has become a major concern. Large portions of communication networks may be damaged by any disaster. With everything from software to infrastructure being offered as a cloud service, we have cutting edge opportunity to provide resilient datacenter based communication systems in a cost-effective manner. Through our work, we have presented an efficient heuristic to ensure the survivability of a static set of communication requests in the presence of set of disasters. Our objectives are to establish minimum no. of datacenters, minimum lightpaths, with wavelength continuity constraint, wavelength clash constraint, and optical reach constraint. In a fairly well-connected network, our approach is guaranteed to identify survivable solutions. A possible direction to our future work might include, handling dynamic communication requests with respect to optical reach and optimal solution for handling the static as well as dynamic requests.

BIBLIOGRAPHY

[1] M. F. Habib, M. Tornatore, M. De Leenheer, F. Dikbiyik, and B. Mukherjee, "Design of disaster-resilient optical datacenter networks," *Journal of Lightwave Technology*, vol. 30, no. 16, pp. 2563–2573, 2012.

[2] R. Das, S. Bandyopadhyay, S. Al Mamoori, and A. Jaekel, "Dynamic provisioning of fault tolerant optical networks for data centers," in *Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on*. IEEE, 2017, pp. 1–4.

[3] M. F. Bari, R. Boutaba, R. Esteves, L. Z. Granville, M. Podlesny, M. G. Rabbani, Q. Zhang, and M. F. Zhani, "Data center network virtualization: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 909–928, 2013.

[4] H. Zang, J. P. Jue, B. Mukherjee *et al.*, "A review of routing and wavelength assignment approaches for wavelength-routed optical wdm networks," *Optical networks magazine*, vol. 1, no. 1, pp. 47–60, 2000.

[5] J. Zheng and H. T. Mouftah, *Optical WDM networks: concepts and design principles*. John Wiley & Sons, 2004.

[6] S. Varanasi, S. Bandyopadhyay, and A. Jaekel, "Impairment-aware dynamic routing and wavelength assignment in translucent optical wdm networks." in *ICDCN*, 2014, pp. 363–377.

[7] D. Banerjee and B. Mukherjee, "A practical approach for routing and wavelength assignment in large wavelength-routed optical networks," *IEEE Journal on selected areas in communications*, vol. 14, no. 5, pp. 903–908, 1996.

[8] N. Skorin-Kapov, "Routing and wavelength assignment in optical networks using bin packing based algorithms," *European Journal of Operational Research*, vol. 177, no. 2, pp. 1167–1179, 2007.

[9] 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.

[10] S. Bandyopadhyay, *Dissemination of Information in Optical Networks:: From Technology to Algorithms.* Springer Science & Business Media, 2007.

[11] "Datatalk," 2017, [Online; accessed 19-September-2017]. [Online]. Available: http://datatalk.co/communications/fibre-optic-cable-solutions/

[12] "The fiber optic association, inc.,," 2017, [Online; accessed 19-September-2017]. [Online]. Available: http://www.thefoa.org/Lennie/jargon.html

[13] "Fiber optic for sale co." 2017, [Online; accessed 19-September-2017]. [Online]. Available: https://www.fiberoptics4sale.com/blogs/archive-posts/95146054-optical-fiber-tutorial-optic-fiber-communication-fiber

[14] P. K. Agarwal, A. Efrat, S. K. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, "The resilience of wdm networks to probabilistic geographical failures," *IEEE/ACM Transactions on Networking (TON)*, vol. 21, no. 5, pp. 1525–1538, 2013.

[15] Wikipedia, "Wavelength-division multiplexing — wikipedia, the free encyclopedia," 2017, [Online; accessed 19-September-2017]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Wavelength-division_multiplexing&oldid=788337161

[16] A. E. Ozdaglar and D. P. Bertsekas, "Routing and wavelength assignment in optical networks," *IEEE/ACM Transactions On Networking (ton)*, vol. 11, no. 2, pp. 259–272, 2003.

[17] H. Song, B.-W. Kim, and B. Mukherjee, "Long-reach optical access networks: A survey of research challenges, demonstrations, and bandwidth assignment mechanisms," *IEEE communications surveys & tutorials*, vol. 12, no. 1, 2010.

[18] R. Ramaswami and K. N. Sivarajan, "Design of logical topologies for wavelength-routed optical networks," *IEEE Journal on Selected areas in communications*, vol. 14, no. 5, pp. 840–851, 1996.

[19] S. Ferdousi, F. Dikbiyik, M. F. Habib, M. Tornatore, and B. Mukherjee, "Disaster-aware datacenter placement and dynamic content management in cloud networks," *Journal of Optical Communications and Networking*, vol. 7, no. 7, pp. 681–694, 2015.

[20] "Globaldots," 2017, [Online; accessed 19-September-2017]. [Online]. Available: http://www.globaldots.com/verizon-auctions-data-centers-google-cloud-attacks-amazon-weeks-news/

[21] S. S. Savas, F. Dikbiyik, M. F. Habib, M. Tornatore, and B. Mukherjee, "Disaster-aware service provisioning with manycasting in cloud networks," *Photonic Network Communications*, vol. 28, no. 2, pp. 123–134, 2014.

[22] F. L. Presti, C. Petrioli, and C. Vicari, "Dynamic replica placement in content delivery networks," in *Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, 2005. 13th IEEE International Symposium on*. IEEE, 2005, pp. 351–360.

[23] T. A. Neves, L. M. Drummond, L. S. Ochi, C. Albuquerque, and E. Uchoa, "Solving replica placement and request distribution in content distribution networks," *Electronic Notes in Discrete Mathematics*, vol. 36, pp. 89–96, 2010.

[24] D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Y. Zomaya, "Energy-efficient data replication in cloud computing datacenters," *Cluster computing*, vol. 18, no. 1,

pp. 385–402, 2015.

[25] Q. Rahman, S. Bandyopadhyay, and Y. Aneja, "Optimal regenerator placement in translucent optical networks," *Optical Switching and Networking*, vol. 15, pp. 134–147, 2015.

[26] S. Bandyopadhyay, Q. Rahman, S. Banerjee, S. Murthy, and A. Sen, "Dynamic lightpath allocation in translucent wdm optical networks," in *Communications, 2009. ICC'09. IEEE International Conference on*. IEEE, 2009, pp. 1–6.

[27] I. Chlamtac, A. Ganz, and G. Karmi, "Lightnets: Topologies for high-speed optical networks," *Journal of Lightwave Technology*, vol. 11, no. 5, pp. 951–961, 1993.

[28] O. Gerstel and S. Kutten, *Dynamic wavelength allocation in WDM ring networks*. IBM TJ Watson Research Center, 1996.

[29] O. Gerstel, G. Sasaki, and R. Ramaswami, "Dynamic channel assignment for wdm optical networks with little or no wavelength conversion," in *Proceedings of the Annual Allerton Conference on Communication Control and Computing*, vol. 34. University of Illinois, 1996, pp. 32–43.

[30] S. Al Mamoori, A. Jaekel, and S. Bandyopadhyay, "Designing resilient wdm data center networks for dynamic lightpath demands," in *Computers and Communications (ISCC), 2017 IEEE Symposium on*. IEEE, 2017, pp. 724–729.

[31] D. Andrei, H.-H. Yen, M. Tornatore, C. U. Martel, and B. Mukherjee, "Integrated provisioning of sliding scheduled services over wdm optical networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 1, no. 2, 2009.

[32] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, "Network flows: theory, algorithms, and applications," 1993.

[33] "Ibm ilog cplex," 2017, [Online; accessed 19-September-2017]. [Online]. Available: https://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/

[34] C. V. Saradhi and S. Subramaniam, "Physical layer impairment aware routing (pliar) in wdm optical networks: issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 4, 2009.

[35] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the vulnerability of the fiber infrastructure to disasters," *IEEE/ACM Transactions on Networking (TON)*, vol. 19, no. 6, pp. 1610–1623, 2011.

[36] G. Ellinas, A. G. Hailemariam, and T. E. Stern, "Protection cycles in mesh wdm networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 1924–1937, 2000.

[37] N.-H. Bao, L.-M. Li, H.-B. Luo, H.-F. Yu, and Z.-Z. Zhang, "Impairment aware sharing constraint relaxed path protection in translucent optical networks," *Optical Engineering*, vol. 51, no. 4, pp. 045 002–1, 2012.

[38] M. Hashimoto and K. Miura, "Layer-wise topology design for cost effective ip-optical networks," *Communications and Network*, vol. 4, no. 01, p. 88, 2012.

VITA AUCTORIS

NAME: Umesh Shah

PLACE OF BIRTH: Ahmedabad, India

YEAR OF BIRTH: 1991

EDUCATION: Shri Damubhai Shukla Shaikshanik Sankul,

Ahmedabad, India, 2010

Gujarat Technological University,

India, 2014

University of Windsor, M.Sc,

Windsor, ON, 2017