End-To-End Signaling and Routing For Optical IP Networks

by

Lambros Pezoulas
B.A.Sc. in Eng., University of Waterloo, 1997

A Thesis Submitted to
The Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of

Master of Applied Science in Electrical Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
Department of Systems and Computer Engineering

Carleton, University
Ottawa, Ontario
January 27, 2003
© copyright
2003, Lambros Pezoulas

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-83476-X

Our file Notre référence

ISBN: 0-612-83476-X

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Canada

The undersigned hereby recommends to
The Faculty of Graduate Studies and Research

Acceptance of the Thesis

**End-To-End Signaling and Routing
for
Optical IP Networks**

submitted by

Lambros Pezoulas, B.A.Sc.

In partial fulfillment of the requirements

For the degree of Master of Applied Science in Electrical Engineering

Dr. I. Lambadaris, Thesis Supervisor

Dr. R. A. Goubran, Chair, Department of Systems and Computer Engineering

Carleton University

January 27, 2003

Abstract

This thesis outlines an approach (called OBGp) of extending the BGP routing protocol to support lightpath setup and management across an optical network. OBGp is a distributed approach, which gives more control to the edge customer. OBGp can be used in both networks with and without wavelength converters. In this thesis we also do a performance analysis of three different backward reservation algorithms that can be used by OBGp. We compare their performance in networks with sparse wavelength conversion. In sparse wavelength conversion networks only a few nodes support wavelength conversion. The optical paths in the network consist of a group of segments where each segment independently must meet the wavelength continuity constraint when setting up lightpaths across them. Because of the segmented nature of the lightpath traditional backward reservation algorithms can result in lower performance. We show how the First-Available algorithm bypasses this problem. Finally based on our performance measurements suggestions are made for the optimal placement of wavelength converters.

Acknowledgements

It is my great pleasure to express my deep sense of gratitude and appreciation to my supervisor Dr. Ioannis Lambadaris for his professional guidance and constant encouragement throughout the course of this investigation. I truly appreciate his generosity in sharing his valuable time and knowledge with me.

I would also like to extend my thanks to Dr. Changcheng Huang for his suggestions and helpful comments on my thesis and research work.

My special thanks goes to my beloved parents, Ioannis and Efstathia, my brother, Vasilios, and my sister, Georgia, who encouraged me throughout the course of my graduate studies. I would also like to mention my friend Konstantinos and thank him for his valuable suggestions.

Table of Contents

Acceptance Sheet	ii
Abstract	iii
Acknowledgements	iv
Table of Contents	v
List of Figures	viii
1 Introduction	1
1.1 Background	1
1.2 Previous Research	2
1.3 Thesis objectives	8
1.4 Thesis Organization	9
1.5 Thesis Contributions	9
2 Optical BGP Architecture	10
2.0 Introduction	10
2.1 OBGp Architecture	10
2.2 OBGp Requirements and Design	12
2.2.1 OBGp Message	15
2.2.2 OBGp Wavelength Table	18

2.3	OBGP Protocol	19
2.3.1	Discovery Phase	22
2.3.2	Reservation Phase	22
2.3.3	Setup Phase	23
2.3.4	Confirmation Phase	24
2.3.5	Teardown Phase	25
2.3.6	Error Phase	25
3	OBGP Signaling, A Backward Reservation Approach	28
3.0	Introduction	28
3.1	Wavelength Reservation Algorithm	29
3.2	First-Fit Algorithm	30
3.3	Random Algorithm	32
3.4	First Available Algorithm	33
4	OBGP Simulation Results	36
4.1	Linear Topology	36
4.2	Competing Traffic Streams	43
4.3	Multiple Competing Traffic Streams	60
5	Conclusions and Recommendations for Further Study	68
5.1	Conclusions	68

5.1.1	IP Control of Optical Networks with OBG	68
5.1.2	OBG Signaling	69
5.2	Recommendations for Further Study	70

REFERENCES	73
-------------------	-----------

List of Figures

Figure 1.1: An example optical network	1
Figure 2.1: Diagram of OBGP Router	14
Figure 2.2: OBGP Message Format	16
Figure 2.3: Lambda_RI Entry Format	17
Figure 2.4: OBGP Wavelength Table Fields	18
Figure 2.5: Four Phase vs. Two Phase OBGP	21
Figure 2.6: Error Scenarios	27
Figure 3.1: Two-Segment path	30
Figure 3.2: First-Fit Wavelength Reservation Algorithm	31
Figure 3.3: Random Wavelength Reservation Algorithm	32
Figure 3.4: First-Available Wavelength Reservation Algorithm	34
Figure 4.1: Linear Network Topology	36
Figure 4.2: Blocking Probability Graphs for Linear Topology	40
Figure 4.3: Nodal Discovery Blockage (Linear Topology)	43
Figure 4.4: Competing Traffic Streams	44
Figure 4.5: Blocking Probability (competition with no conversion)	45
Figure 4.6: Message Blockage Comparison (competition with no conversion)	46
Figure 4.7: Nodal Discovery Blockage (competition with no conversion)	49
Figure 4.8: Blocking Probability (competition with converter at node Y)	50

Figure 4.9: Message Blockage Comparison (competition with converter at node Y) . .	52
Figure 4.10: Nodal Discovery Blockage (competition with converter at node Y).	53
Figure 4.11: Blocking Probability (competition with converter at node X)	55
Figure 4.12: Message Blockage Comparison (competition with converter at node X). .	56
Figure 4.13: Blocking Probability (competition with converter at nodes X and Y) . . .	57
Figure 4.14: Message Blockage Comparison (competition with converter at X and Y) .	59
Figure 4.15: Network with multiple competing traffic streams	60
Figure 4.16: Nodal Discovery Blockage (multiple competition with no conversion) . .	61
Figure 4.17: Nodal Reservation Blockage (multiple competition with no conversion) .	63
Figure 4.18: Blocking Probability (multiple competition with no conversion).	64
Figure 4.19: Blocking Probability (multiple competition, varying converter location) .	65
Figure 4.20: Nodal Reservation Blockage (varying converter location)	67

To my Parents

Ioannis and Efstathia

Chapter 1

Introduction

1.1 Background

In recent years advances in wavelength-division multiplexing (WDM) technology, have enabled the deployment of systems that are capable of providing large amounts of bandwidth. These systems utilize Optical Cross Connects (OXC), which are capable of switching data entirely in the optical domain. Configuring these devices across a network enables one to establish all-optical connections, or lightpaths, between source nodes and destination nodes (figure 1.1). A lightpath is a contiguous optical path, which may be made up of one or more wavelengths connected across optical switches [17]. Data carried on these lightpaths avoid electronic conversion and processing at intermediate nodes, thereby alleviating the electronic bottleneck [14].

Figure 1.1: An example optical network

To compliment these high bandwidth optical networks though, intelligent methods of provisioning and control are required to enable greater flexibility and to reduce operation cost. A convergence is taking place between IP-based routers and multi-wavelength optical equipment to create an intelligent end-to-end integrated network [15]. Traditionally, as stated in [19], end-to-end circuit connections have been set up via network management systems (NMSs), which issue commands (usually under the control of a human operator) to the various network elements involved in the circuit via an equipment vendor's element management system (EMS). So far this has been the carrier's responsibility and the light path setup and configuration have been based on a centralized model. By moving this responsibility from the carrier to the customer and using a distributed approach, customers can better manage their optical wavelengths. This distributed approach gives more control to the edge customer.

The configuration of lightpaths in an All Optical Network can be divided into two tasks:

1. Determine the route on which the lightpath will be setup
2. Select the available wavelength(s) to be used across the route.

In general this is known as the Routing and Wavelength Assignment (RWA) problem.

1.2 Previous Research

Internet Protocol (IP) mechanisms are being extensively proposed for the management of optical resources [18, 20, 21]. Utilizing IP standards rapidly achieves vendor

interoperability and enables optical networks to continually track development in IP standards [18]. In [15, 17] a protocol called OBGp was proposed as a means for managing and setting up lightpaths across an All Optical Network. OBGp is an extension of the traditional Border Gateway Protocol (BGP) IP routing protocol. The research of [15, 17] show that there are advantages to integrate routing and signaling into one protocol. Both signaling and routing protocols are jointly used to carry information through the network. Signaling typically requires routing information, for example source routing. For failure protection and restoration, the interactions between signaling and routing are critical. Currently there exist various routing and signaling protocols that need to interwork in order to provide an end-to-end solution. In some cases this interworking of multiple protocols is complicated and results in poor scalability. It also requires that every router along the path implement the required routing and signaling protocols. Therefore, a better approach may be to tightly integrate the routing and signaling function into one protocol [15].

A number of other mechanisms have also been proposed for the management and control of such “wavelength cloud” systems [22]. Most of these systems have been designed on variations of link state interior routing protocols such as OSPF, IS-IS [23] and PNNI [24] or complementary extensions of MPLS such as MPLmS [25]. These protocols allow for the optimized configuration and establishment of lightpaths across a single management domain. To date there has been little work done on developing protocols for the management of wavelengths between separate management domains. In addition, customer control and routing of wavelengths across a wavelength cloud is still

not possible. The research in [15, 17] tries to address these issues by using BGP which is an Inter-domain Routing protocol.

As it was mentioned in the previous section the RWA problem can be broken down into the routing and the signaling part. Using an existing IP routing protocol to determine the routes for the lightpaths the next problem to solve is the signaling. The signaling required for establishing a connection may be either centralized or distributed. Under centralized control, each request is processed by a central controller, which chooses a path and assigns an appropriate wavelength on each link along the path to establish the connection. There is a potential risk to having a single controller in the network. Since one controller maintains the state of the entire network, if it crashes it will result in bringing down the entire network. In large networks, centralized control is not feasible and therefore, it is essential to study distributed control protocols [2]. In contrast to the centralized approach, distributed control requires each node in the network to act as a separate controller and maintain its own local database. This makes this approach much more reliable in the event of a node failure in the network.

Previous research [1, 2, 3, 4] has studied and evaluated the performance of various distributed control protocols. The performance of these protocols was studied using two basic approaches. The first approach called *Path Multiplexing* (PM) requires the same wavelength to be used along the path [5]. This requirement is also known as the wavelength continuity constraint. The second approach called *Link Multiplexing* (LM), can establish a connection using possibly different wavelengths on different links along the path [5]. This requires each node in the network to have a wavelength converter. In

the case of PM no wavelength converters are required since the same wavelength is used end-to-end.

The distributed protocols studied in [1, 2, 3, 4] are classified into two categories, *forward wavelength reservation*, and *backward reservation* protocols. In *forward reservation* schemes wavelength resources are reserved along the forward path to the destination on a hop-by-hop basis [6]. In the distributed control approach a node only knows the status of its immediate links which makes the wavelength selection quite complicated since the source node doesn't know which wavelength will be available along the entire path. As the control message traverses the path in the forward direction, each node along the path attempts to reserve one or more wavelengths on the next link for the connection. If no wavelength can be found on the next link the request is blocked and the intermediate node sends a failure message back to the source, un-reserving any wavelengths that were temporarily reserved. Otherwise, at the destination node a wavelength is selected and a confirmation message is sent to the source node and at the same time un-reserves the other wavelength that were not chosen.

There are few limitations though with the forward reservation approach. First, in the case where only one wavelength is selected and reserved at each node there is no guarantee that the wavelength will be available on every link in the path. In the case where the wavelength is blocked, the source node can select another wavelength and reattempt to reserve it. This can result though in high setup times since several attempts might be required before a successful reservation. A second limitation of the forward reservation schemes is the over-reservation of resources in the case where more than one

wavelength are reserved along the forward path, with the expectation of finding at least one available wavelength end-to-end. During the time where the resources are temporarily reserved, they can't be used by other users even if they will never be used by the connection [6]. This may lead to blocking of subsequent requests [6].

In order to address the limitations of the forward reservation protocols, backward reservation schemes perform reservations after the control message has reached the destination [6]. As the control message traverses the path from source to destination it collects information as to which wavelengths are available first. Then, when the control message reaches the destination it has discovered a list of available end-to-end wavelengths. The destination node then selects one of these and sends a reservation message in the backward direction towards the source. By reserving wavelengths in the reverse direction wavelengths are idle for less time than if the wavelengths were reserved in the forward direction. Also. Only one wavelength is reserved, without having to temporarily block any other wavelengths, allowing them to be used for other connections. One possible drawback of a backward reservation scheme is that if multiple connections are set up simultaneously, it is possible that a wavelength available on a link in the forward direction will be taken by another connection request and no longer be available when the reservation message traverses the link in the reverse direction [6].

In general it has been found that backward reservation schemes outperform forward reservation schemes in the case where there are no wavelength converters [1]. Simulations performed in [1] confirmed that the lack of converters decreases the throughput of the network. Moreover, the simulations showed that a major portion of the

bandwidth loss due to the absence of wavelength conversion can be regained if a backward reservation protocol is used instead of the forward protocol [1]. In the case of a network with full wavelength conversion, different wavelengths can be used on the links across the path for which a lightpath is being setup. Thus, in this case there is no wavelength continuity constraint. This greatly simplifies the control overhead since only the availability of any wavelength is of concern and not the availability of a specific wavelength. In these cases a forward reservation scheme would be a better choice.

The studies in [1, 2; 3, 4] focus on networks either with full wavelength conversion where LM can be used, or networks with no wavelength converters at all where PM can be used. The benefits of using full wavelength conversion in a network have been studied in [12, 13] and compared to networks without wavelength conversion. These studies look at how various network parameters affect the performance under both cases. Obviously a network with full wavelength conversion, provide flexibility since the wavelength continuity constraint is not applicable. However quantitative results on the effectiveness of wavelength converters have been mixed [13]. Also, the high cost of such devices is a deferring factor. For this reason recent work has focused on networks with sparse wavelength conversion. It has been shown that sparse wavelength conversion is more realistic and efficient than full wavelength conversion [8, 9, 10]. The term sparse wavelength conversion means that a limited number of nodes in a network can convert all of the wavelengths in the network [11] as opposed to full wavelength conversion, which refers to a network where all nodes are equipped with wavelength converters. There has

not been any study on backward reservation algorithms in sparse wavelength conversion networks.

1.3 Thesis objectives

The aim of this thesis is to:

- i. Develop a protocol for managing optical resources, which belong to different autonomous domains through the IP layer by extending the traditional BGP protocol.
- ii. Demonstrate how routing and signaling can be successfully combined within the OBGp protocol.
- iii. Study the signaling part of OBGp and develop a backward reservation algorithm that works efficiently in sparse wavelength conversion networks.
- iv. Show how OBGp's backward reservation algorithm, address the main problem of backward reservation algorithms of multiple connections being set up simultaneously.
- v. Develop an OPNET model of the OBGp protocol, and compare, through OPNET simulations the performance of this algorithm to other backward reservation algorithms such as First-Fit and Random in various network topologies.
- vi. Observe nodal statistics such as the blocking at each node in various network topologies and find the optimal placement of the converters in the network for this wavelength reservation algorithm.

1.4 Thesis Organization

The remaining chapters of this thesis are organized as follows:

- Chapter 2:** Outlines the operation of the OBGp protocol and presents the OBGp Architecture.
- Chapter3:** Focuses on the signaling algorithm used by the OBGp protocol presented in Chapter 2. Describes three variations of the signaling algorithm and highlights the differences between these.
- Chaper4:** Analyzes and discusses the results achieved through the OPNET simulations. The backward reservation algorithm studied is compared to other basic algorithms such as First-Fit and Random algorithm in various network topologies.
- Chapter 5:** Conclusions and Recommendations for Further Study.

1.5 Thesis Contributions

As part of this research work the following publications have been made:

- L. Pezoulas, M. Francisco, I. Lambadaris, C. Huang, "Performance Analysis of a Backward Reservation Protocol In Networks with Sparse Wavelength Conversion" to be presented at ICC 2003.
- M. Francisco, L. Pezoulas, C. Huang, and I. Lambadaris, "End-To-End Signaling and Routing for Optical IP Networks," *Proceedings of ICC'02*, New York, April, 2002.

Chapter 2

Optical BGP Architecture

2.0 Introduction

The primary purpose of this chapter is to outline a method called Optical Border Gateway Protocol (OBGP), which can be used to provide an edge customer with a control method for establishing a lightpath across different autonomous domains through an optical network. OBGP is an extension of the traditional BGP IP routing protocol. This chapter will describe the development and implementation of OBGP by outlining the following areas: a) The design requirements for creating OBGP, b) The necessary extensions to BGP to create OBGP, c) The implementation of OBGP into a standard BGP framework, d) The benefits of using the OBGP approach as a way for setting up and managing lightpaths.

2.1 OBGP Architecture

The goals of an interdomain routing protocol such as BGP are very different from those of an intradomain protocol. The issue of scalability has much more significance for an interdomain gateway router, which must be able to forward any packet destined for anywhere in the Internet, versus a local router on a small subnet. In addition, the autonomous nature of network domains means that it is very difficult to calculate an

optimal path that crosses several domains. As a result, interdomain routing will only advertise network reachability.

There are several reasons BGP was chosen as the basis for this research towards a protocol for a light path setup and control. BGP is the current standard for interdomain routing in the Internet today. Utilizing the interdomain properties of BGP is fundamental to the design of OBGp, as many of the issues that external gateway protocols such as BGP were designed to deal with are similar to the management of multiple light paths in an optical network [3]. The widespread deployment of BGP is another important factor. It is easier to build something new based on a protocol that has been proved successful rather than creating a new one. It is very difficult to integrate a completely new protocol into the existing Internet. By minimizing the modifications to be made to the existing BGP, there is a better chance to integrate OBGp into the Internet.

BGP is a routing protocol while the setup and control of light paths are the functions of a signaling protocol. Both signaling and routing protocols are jointly used to carry information through the network. By combining routing with signaling in OBGp a lightweight solution is provided. Signaling functions can leverage some features of existing routing functions. For example, both routing and signaling messages can share a common transport connection (i.e. same TCP connection is used to exchange both routing and signaling messages). Thus, no extra overhead is required for setting up a separate communication channel for signaling messages. OBGp can also access directly the

database of BGP and therefore eliminate the synchronization requirement if two separate databases are used.

Another reason, which makes BGP ideal for signaling, is the AS_PATH attribute of a route. This attribute contains the ASs through which the announcement for the route has been passed. It gives an end-to-end view, which is necessary for setting up a light path. The AS_PATH also allows for better distribution of OBGp messages in the network. For example not all OBGp speakers in the network need to see this message if they are not in the specific AS_PATH thus reducing the amount of signaling traffic in the network. OBGp provides an interdomain solution that integrates heterogeneous domains into an end-to-end optical network. Different domains still can use their own intradomain solutions. Therefore OBGp can coexist with most of the existing solutions. Finally, as mentioned in the introduction the distributed nature of OBGp moves the responsibility of setting up a light path from carriers to edge customers. This allows the customer to set up a light path faster without having to wait for carriers to provision their connections which in most cases is a complex and time consuming task.

2.2 OBGp Requirements and Design

BGP does not contain any functionality or attributes for maintaining information on light paths across a network. However, BGP does contain *complete autonomous system (AS) path information* to reach a particular network, which makes it unique when compared to other routing protocols. This full path visibility enables routing policy decisions to be

made as well as preventing routing loops. Having full path visibility is a useful feature of BGP and can be leveraged for setting up a light path from one AS to another [3].

There are four types of messages specified for BGP [4]: OPEN, UPDATE, NOTIFICATION, and KEEP-ALIVE. An UPDATE message is used to provide routing updates within a network. These messages allow BGP routers to maintain consistent information regarding network reachability. UPDATE messages can add or withdraw routes from the Routing Information Base depending on the content of the message.

The basic high-level requirements for OBGp to work as a *light path setup* and *control* protocol are as follows:

- Provide a wavelength table to maintain light path information
- Carry light path reservation requests and responses between OBGp speaking devices
- Propagate up to date information on the status of resources in the network

Making minimum changes to BGP while meeting these requirements is a very important consideration. Leveraging current BGP properties is key to meeting these requirements. A typical OBGp Router consists of three components, illustrated in 2.1. The first component is a Border Gateway Router that can be found at the gateway of an AS. This router performs the BGP Peering and sits at the control plane for OBGp. The Switch Control Server (SCS) is the control interface between the router and the OXC. This interface maintains connection state for the OBGp process in the router. The

Optical Cross Connect (OXC) can be any type of cross connect either with wavelength conversion capability or not.

Figure 2.1: Diagram of OBGP Router

Initially the UPDATE message was considered for carrying the light path setup information in addition to normal reachability information [9]. This would require a few small extensions to be made to the normal UPDATE message. Given though that optical networks are continuously evolving, it was decided that it is more efficient to introduce a new message type to BGP. By using a separate message for the Optical part of OBGP it will be easier to make changes to the protocol. Furthermore, the changes required for supporting light path setup, are isolated without affecting the operation of the other messages, and allowing traditional BGP routers and OBGP capable routers to inter-work seamlessly. Thus, although the UPDATE message is sufficient for OBGP, by introducing a new message type more flexibility is provided to the OBGP approach.

The new message used, is called the OBGp message. The format of this message will be described in section 2.2.1. Another extension that must be made to provide traditional BGP with the functionality to setup light paths is the wavelength table at each OBGp router. This is required to store wavelength availability and setup information. The format of the wavelength table is described in section 2.2.2.

2.2.1 OBGp Message

The OBGp message is the primary message used for managing and configuring the optical cross connects and to update the wavelength tables at each OBGp speaking device. The OBGp message contains the standard fixed size header, which is common to all BGP messages. This header as described in [4] contains the Marker, Length, and Type fields. In order to distinguish OBGp messages from other BGP messages a new type code has been introduced making the set of type codes as follows:

- OPEN (1)
- UPDATE (2)
- NOTIFICATION (3)
- KEEPALIVE (4)
- OBGp (5)

In addition to the fixed-size header, the OBGp message contains the following fields:

4 bits	4 bits	4 bits	16 bits	Variable	Variable
Phase Id	Sub-code Id	Oper. Mode	Attributes Length	Attributes	Lambda RI

Figure 2.2: OBGp Message Format

Phase Id: This 4 bit field indicates the phase of the light path setup process. The possible phases which are described later on in the paper, are: i) Discovery, ii) Reservation, iii) Setup, iv) Confirmation, v) Teardown, vi) Error.

Sub-code Id: This 4 bit field is used by an OBGp message in the Confirmation or Error phase. For confirmation messages it allows to distinguish between Setup confirmation and Teardown confirmation messages. For error messages it allows to indicate in what phase the error occurred.

Operational Mode: This 4 bit field is used to indicate mode of operation of OBGp: two-phase or four-phase.

Attributes Length: This 16 bit unsigned integer indicates the total length of the Attributes field.

Attributes: The supported attributes for OBGp are the following:

AS_PATH: This is the AS Path from source to destination for which a light path is being setup. As an OBGp message travels from one router to the other along this path, this attribute does not change.

NEXT_HOP: The NEXT_HOP attribute defines the IP address of the border router that should be used as the next hop for the OBGp message.

CONN_ID: The CONN_ID attribute is used to identify light path connections. The connection identifier consists of a unique identifier plus the AS Path that corresponds to the specific light path.

Lambda RI: The Lambda Reachability Information (Lambda RI) contains a list of allowable wavelengths that can be used for the light path connection, each of which are formatted as shown in the following figure.

Lambda_num	Src_port_id	Dst_port_id	Seg_start	Seg_end
------------	-------------	-------------	-----------	---------

Figure 2.3: Lambda_RI Entry Format

The fields in the Lambda RI are described below :

Lambda_num: the wavelength number.

Src_port_id: the port on the crossconnect of the end point

Dst_port_id: the port on the crossconnect of the end point

Seg_start: the AS identifier of the start of a segment on which the wavelength is available

Seg_end: the AS identifier of the end of a segment on which the wavelength is available

2.2.2 OBGW Wavelength Table

Each OBGW router stores wavelength availability and setup information in its wavelength table. Using the wavelength table the OBGW process can easily determine which wavelengths are available to be used for a light path setup. Each wavelength entry in the table contains the following fields:

AS number	Lambda Id	Connection Id	Preference Level	Setup Complete
-----------	-----------	---------------	------------------	----------------

Figure 2.4: OBGW Wavelength Table Fields

AS number: the AS number of the neighbor border router to which the specific wavelength connects.

Lambda Id: This consists of the wavelength number, source port number and destination port number.

Connection Id: This consists of an identifier plus the AS Path that corresponds to the light path. If this field is null, then it means that the wavelength is available, otherwise it indicates that the wavelength is unavailable and cannot be used for another connection. The numeric identifier plus the AS path create a unique connection id. The numeric identifier within the connection id is not unique by itself across the network. It is only unique with respect to the same Source and same AS Path from that source.

Preference Level: This field is included for future use, to allow for assigning different preference levels to specific wavelengths.

Setup Complete: This field indicates if the connection for the specific wavelength has actually been setup on the OXC. This can be either True or False.

2.3 OBGP Protocol

In this section we consider the scenario where a source node S wants to setup a light path to a destination node D. Through regular BGP peering, node D has advertised its presence to the network and each node has AS path information and wavelength status information to reach node D. Thus, node S has the AS path required to reach node D. This is a path discovered through BGP and can be found in the BGP Routing Information Base. Each node in the path also has a wavelength table that stores the wavelength status information as described in the previous section.

Each OBGP node must be aware of its position within the AS Path in order to process correctly the OBGP messages. There are three types of OBGP nodes present in any optical network: Source, Intermediate, and Destination. A node's classification will change according to its relative position in the light path segment that is being reserved.

Light path Segment example: [Source_AS, Intermediate_AS1, Intermediate_AS2, ...
Destination_AS]

If a node is listed at the head of a light path reservation segment it is the Source, if it is at the tail, it is the Destination. Otherwise, the node is considered to be an intermediate node in the path segment such as AS1 in the above example.

The OBGp protocol can either operate in as a two-phase or a four-phase setup mode. In the four-phase setup mode, all resources have to be reserved first before the Setup of the actual OXCs can take place. This ensures that during the Setup phase, resources are available. On the other hand in the two-phase mode, Setup takes place immediately after having determined if there is a wavelength available to use. If there are many wavelengths available, then most likely there will be few setup failures due to competition scenarios. In this case a two-phase approach might be more efficient and faster approach. In the case though where there are many users trying to setup light path connections over the same path or sections of the same path, many competition failures may occur. In this case the four-phase approach can be more efficient since the competition will be detected in the reservation phase while trying to reserve the resources. Thus no time is wasted making the actual connection on the OXC, which can be costly. Also, no time has to be spent to destroy any connections on the OXCs.

In the four-phase mode of OBGp, the phases are: 1) Discovery, 2) Reservation, 3) Setup, and 4) Confirmation. In the two-phase mode the phases are: 1) Discovery and 2) Reservation-Setup. In the two-phase approach there is no need of an explicit confirmation message since the combined Reservation-Setup message that eventually reaches the Source serves as a confirmation. The two modes of operations are also illustrated in figure 2.5. These phases will also be described in more detail in the next sections. Note

that this chapter does not describe in detail the signaling algorithm used by OBGp, this will be done in chapter 3. This chapter focuses more on the general flow of messages and message types. Also, the last section of this chapter describes the various error scenarios.

(a) 4 Phase Setup

(b) 2 Phase Setup

Figure 2.5: Four Phase vs. Two Phase OBGp

2.3.1 Discovery Phase

A Discovery OBGp message is an OBGp message with its Phase Id field set to Discovery. When a source node S wants to setup a lightpath to a destination node D, it first determines the AS Path for this destination node from the local Routing Information Base provided by BGP. Then it initiates the light path setup request by sending a Discovery OBGp message to the next node in the AS Path. The source node S also fills the Lambda RI field of the Discovery OBGp message with all available wavelengths that connect to the next node in the AS Path. The AS Path and the connection id are stored in the attributes field of the message. The Discovery OBGp message traverses all nodes along the AS path probing for the available wavelengths that can potentially be used for establishing the lightpath connection. If the OBGp process determines that a specific wavelength is in use (by looking at the wavelength table, then it eliminates the wavelength from the list of available wavelengths. When the destination node receives the Discovery OBGp message, it contains all common wavelengths that can be used from source to destination for the light path setup. At this point the Discovery Phase has completed.

2.3.2 Reservation Phase

The Reservation phase is specific to the four-phase mode of operation of OBGp. In this phase resources are reserved in order to ensure their availability in the Setup phase. Once the wavelength(s) to be used has/have been determined, the Reservation OBGp message

is sent from Destination to Source. A Reservation OBGp message is an OBGp message with its Phase Id field set to Reservation. The AS Path and the connection id are stored in the attributes field of the message. When a node along the AS Path receives a reservation OBGp it finds this wavelength in its wavelength table and sets the Connection Id field to the Connection Id that was passed in the Attributes field of the Discovery OBGp message. When the Reservation message reaches the Source node and all nodes along the AS Path have updated their wavelength tables, then the Reservation phase has completed.

2.3.3 Setup Phase

In the four-phase mode the Setup phase takes place after resources have been reserved and the Reservation has completed. At this point all nodes along the AS Path have reserved a wavelength in their wavelength tables for the light path setup, but they have not actually made any hardware connections on the OXC. This is done in the Setup Phase. The source node S initiates the Setup Phase by sending a Setup OBGp message to the next node in the AS Path. A setup OBGp message is an OBGp message with its phase field set to Setup. In general when a node receives a Setup OBGp message, it makes the appropriate OXC connections and once this is complete it updates its wavelength table to indicate that Setup is complete. Eventually the Setup OBGp message is received at the destination node D and the Setup Phase is complete.

In the case of the two-phase mode the Setup phase takes place immediately after the Discovery phase. During the Setup phase in the two-phase mode, the OXC

connections are setup, and the wavelength table is updated at the same time to indicate that the specific wavelength is being used. In contrast to the four-mode operation the Setup phase proceeds from the Destination node to the Source node as shown in the diagram on the right in figure 2.5. Since the Setup phase is complete when the Setup OBGp message reaches the Source node, there is no need to send a separate confirmation message. The Setup message also serves as the confirmation of a successful setup.

2.3.4 Confirmation Phase

In the four-phase mode after the Setup Phase is complete, the source node is not aware of this. Until it receives some kind of notification, the source node can't start sending data over the optical light path. For this reason, the destination node D needs to send a Confirmation OBGp message back to the source along the AS path. A Confirmation OBGp message is an OBGp message with its phase field set to Confirmation. In this case since this is a confirmation for a Setup request, the sub-code field of the message is set to Setup. This is used in order to distinguish between Setup and Teardown Confirmation messages. Once the source node receives the Confirmation message this phase is complete and the source can start using the light path (in the case of a Setup Confirmation).

2.3.5 Teardown Phase

Teardown as opposed to setup is always a two-phase process since there is no competition in this case. Only the source node can request to teardown a light path. In order to achieve this the source node S, sends a Teardown OBGp message to the next node in the AS Path for the connection to be torn down. This information can be looked up in the wavelength table. When a node receives a Teardown OBGp message it first releases the hardware resources and then updates the wavelength table to indicate that the wavelength is available. Then it passes on the Teardown OBGp message to the next node in the AS Path. When the Teardown message is received at the destination node, the Teardown Phase is complete. The destination node then sends a Confirmation OBGp message with sub-code field set to Teardown back to the Source. This notifies the sources that the light path teardown was successful.

2.3.6 Error Phase

Whenever an error is detected, an OBGp Error message is sent back to the Source node. An Error OBGp message is an OBGp message with its phase set to Error and the sub-code field is set to the phase where the error was detected. The source node in a light path segment is considered to be the master and is the one that decides what action to be taken in the various error conditions. For example if the Source is notified that error occurred in the Setup phase, then it should immediately send a Teardown message to release the resources. Otherwise, the wavelength that was partially setup will appear that it is in use

and will not be available for other connections. 2.6 illustrates the flow of messages in the case of an error in different phases of the four-phase approach.

There are several sources of error. Some of the most important ones are the following:

- No available wavelength for setting up a light path. This will be detected in the Discovery phase if for example all common wavelengths are eliminated.
- Various competition cases where two light path setup requests both discover that a certain wavelength is available but then one of them reserves it first. The other one assuming that it is still available, attempts to reserve it and an error occurs. These errors are detected in the Reservation Phase for the four-phase mode and the Setup phase for the two-phase mode.
- Errors related to hardware failures on the OXCs. These types of errors will most likely be detected in the Setup Phase, when trying to make a connection on the OXC.

(a) Error in the Discovery Phase

(b) Error in the Reservation Phase

(c) Error in the Setup Phase

Figure 2.6: Error Scenarios

Chapter 3

OBGP Signaling, A Backward Reservation Approach

3.0 Introduction

In the previous chapter the necessary extensions to the traditional BGP protocol required to create OBGP were outlined as well as the benefits of using OBGP. Furthermore, it was shown that OBGP combines routing and signaling into one protocol. Given that routing is achieved by regular BGP, the OBGP extensions only need to focus on the signaling part. Thus by leveraging on BGP's routing capabilities, OBGP does not have to address the routing sub problem of the overall RWA problem, and can consider routing to be fixed. In fixed routing, a single fixed route is predetermined for each source-destination pair. When a connection request arrives, the network will attempt to establish a lightpath along the fixed route. Fixed routing schemes do not require the maintenance of global network state information [6]. A fixed routing approach is simple to implement; however it can be limited in term of routing options can lead to higher blocking [6]. Thus, an efficient signaling mechanism is critical in this case. This chapter will elaborate on the signaling part of the OBGP protocol, and present three variations of this signaling mechanism. The control algorithms described in this chapter can be used by other more general protocols as well and are not limited to OBGP use only.

3.1 Wavelength Reservation Algorithm

The Control Protocol described in this paper is a backward reservation protocol. The forward path is used for probing the available resources (wavelengths) and the actual reservation takes place in the backward path. This, wavelength assignment protocol supports optical networks with or without optical converters. In the case where there are no wavelength converters in the network the destination node acts as the main decision point, selecting the wavelength to be used for the entire path. In the case where there exist some optical converters, the end-to-end path consists of sub path segments. The endpoints of these segments are defined by the converter node locations. The wavelength continuity constraint is only applied to each individual segment independently. Two separate segments can use different wavelengths. Thus, each node with conversion capability makes a decision for one of the path segments connected to it. The destination is still a decision point for selecting the wavelength of the last segment in the path. As we will see further down the case of no converter is really a special case (i.e. single segment case) of the general algorithm that supports wavelength converters.

The performance of this control protocol depends on the algorithm used for making a wavelength selection at each of the decision points along the path. The algorithm studied in this paper is called First Available and we compare this to other basic algorithms such as Random selection and First-Fit selection. The following figure displays a scenario where the path consists of five nodes and there is one converter at node C. Thus, the path consists of two segments. The common wavelengths in segment 1 are λ_1 , λ_2 , λ_3 and in segment 2 λ_4 , λ_5 , λ_6 . We will use this diagram to describe our wavelength reservation/selection algorithm. We will first describe how First-Fit

backward reservation algorithm works and then describe First Available and highlight the differences of these two algorithms.

Figure 3.1: Two-Segment path

3.2 First-Fit Algorithm

Figure 3.2 illustrates the First-Fit algorithm. In figure 3.2a the discovery message traverses the nodes probing for the available wavelengths. The format of the OBGp discovery message was described in the previous chapter. Along with the wavelengths the discovery message records the segment endpoints on which the wavelengths are available. This information is required in order to be able to support wavelength converters. Thus, discovery message 1 finds that $\lambda_1, \lambda_2, \lambda_3$ are available from node A to node B. When the discovery message traverses from node B to C (Disc.2) the endpoint information in the message is updated as shown in figure 3.2a since the discovery process has discovered that wavelengths $\lambda_1, \lambda_2, \lambda_3$ are available from A to C.

When a discovery message reaches a converter node it means that it has reached the end of the current segment and a new segment will be traversed. At this point the converter node needs to select which wavelength to use for the segment that has just been traversed before propagating the discovery message to the next node. In the First-Fit algorithm the first wavelength from the list of wavelengths for the specific segment is

selected. When node C sends the discovery message to node D it only propagates λ_1 indicating the segment on which it is available (A to C) and at the same time it discovers and sends the wavelengths that are available for the new segment ($\lambda_4, \lambda_5, \lambda_6$). When the discovery message reaches the destination, the destination node needs to select a wavelength to be used for the last segment, similarly to how node C selected a wavelength for segment A-C. Thus, λ_4 is selected at node E for segment C:E and this completes the discovery phase.

(a)

(b)

Figure 3.2: First-Fit Wavelength Reservation Algorithm

The reservation messages traverse the path in the backward direction as shown in figure 3.2b. The wavelengths selected for each segment in the discovery phase are reserved in each node's wavelength table, which was described in Chapter 2. When the reservation message reaches a converter node the wavelength for the previous segment is

popped off the list (since the reservation for that segment has completed). For example when C sends the reservation message to B it only passes λ_1 and it has completed reservation for segment C-E. Finally, the reservation message reaches node A and a wavelength has successfully been reserved on both segments that consist the path.

3.3 Random Algorithm

The Random selection algorithm is similar to the First-Fit algorithm that was just described with the only difference that wavelengths are selected *randomly* at converter nodes and at the destination node. Thus, referring to the previous example at node C $\lambda_1, \lambda_2, \lambda_3$ are available, and λ_2 is selected randomly instead of λ_1 which was selected with the First-Fit Algorithm. Similarly, node E selects λ_6 randomly. This is shown in figure 3.3. Apart from the selection mechanism the flow of messages is the same as in the First-Fit Wavelength Reservation Algorithm.

(a)

(b)

Figure 3.3: Random Wavelength Reservation Algorithm

3.4 First Available Algorithm

Having looked at First-Fit and Random wavelength selection algorithms we will present a new algorithm called First Available and compare its operation. The message flow for the First-Available algorithm is shown in figure 3.4. The first main difference with the other two algorithms is that at a converter node, the wavelengths for the current segment are all passed to the next segment. Thus, a wavelength selection is not made in the forward direction at converter nodes; rather all available wavelengths for each segment are propagated to the destination as shown in figure 3.4a. As a result though we end up with larger control messages than the case of First-Fit and Random. This is easily seen, by comparing the messages passed in figure 3.2 and figure 3.4. The worst case is a network with wavelength converters at every node. In this case the number of light-path segments is equal to the number of links in the path and using the First-Available algorithm would result in significant control overhead. The First-Available algorithm is intended for networks where a small fraction of the nodes have full wavelength conversion capabilities.

(a)

Figure 3.4: First-Available Wavelength Reservation Algorithm

The second difference is in the selection of a wavelength. In the previous algorithms once a wavelength was selected the rest of the wavelengths were discarded from the list. Then in the reservation phase the destination node or a converter node would attempt to reserve the specific wavelength in its wavelength table. In the case where there is competing traffic or multiple requests to the same node it is possible that between the time that discovery phase saw that the selected wavelength was available, and the time it actually tries to reserve the wavelength, that some other request has taken the wavelength and is no longer available. This would result in blocking the initial request.

In order to overcome such blocking First-Available algorithm does not discard the list of wavelengths after selecting a wavelength. It selects the first wavelength from the list and then tries to reserve it at the node where it was selected (destination node or converter node). If the reservation fails, the algorithm attempts to reserve the next wavelength from the list. Retries of wavelength reservations happens only at converter nodes or at the destination. This is repeated until a successful reservation is achieved or if

all the wavelengths on the list are attempted and failed. In the later case the request is considered blocked.

In figure 3.4b we see that the algorithm tried to reserve λ_4 and failed and then tried λ_5 . Since this was successfully reserved at the destination node this was the wavelength that was propagated to the next node in the path. Similarly, at the converter node C the algorithm tries to reserve λ_1 and fails. It then tries to reserve λ_2 that also fails. Finally it tries λ_3 and succeeds. This is the wavelength that is then propagated to the rest of the nodes in this segment.

Chapter 4

OBGP Simulation Results

In this chapter the performance of the three reservation algorithms described in Chapter 3 are compared and analyzed under different network scenarios. OBGP is used to implement and test each one of these algorithms in OPNET. Based on the statistics gathered from the OPNET simulations, recommendation is also made for to the optimal placement of wavelength converters in each case.

4.1 Linear Topology

The first scenario studied is a linear topology where a source node S makes multiple requests to a destination node D as shown in figure 4.1. Due to its simplicity this scenario is also used to verify the operation of the OBGP protocol by comparing to theoretical results. The requests at node S are generated according to a Poisson distribution with rate λ req/sec. Once a lightpath has been established, the holding time of the lightpath is exponentially distributed with average $1/\mu$ sec. None of the nodes in this scenario have wavelength conversion capability, thus the path consists of one segment for which the wavelength continuity constraint must be met.

Figure 4.1: Linear Network Topology

Before looking at the simulation results, a theoretical approximation for the blocking probability for each of the three algorithms is developed. We then compare these results to the simulation results.

1) First Available Algorithm: Erlang's First Formula can be used to estimate the blocking probability of the First Available Algorithm in a linear topology with no wavelength conversion. With the First Available algorithm the selected available wavelength will be available independently from the processing delay of the signaling messages at the nodes, and as long as there is at least one wavelength available there will be no blocking. Thus, the following formula allows us to approximate First Available Blocking in the linear topology:

$$P_{blk} = P[\text{lightpath request blocked}] = \frac{\left(\frac{\lambda}{\mu}\right)^s / s!}{\sum_{j=0}^s \left(\frac{\lambda}{\mu}\right)^j / j!}$$

s : is the number of wavelengths, which can be used end-to-end between the source-destination pair.

λ : the rate at which the source generates light-path requests to the specific destination.

μ : the mean holding time of the light-path after it has been established and setup.

2) First Fit Algorithm: In the case of the First Fit algorithm the blocking probability depends on the propagation delay. This is more evident when requests come very close together.

Let $P_{blk} = \frac{\left(\frac{\lambda}{\mu}\right)^s / s!}{\sum_{j=0}^s \left(\frac{\lambda}{\mu}\right)^j / j!}$ be the probability that there is no wavelength available.

Given that the round trip delay between the second last node and the last node in the source-destination path is d time units, then if two light-path requests come close together by a time difference of less than this propagation delay they will both see that a particular wavelength is available and attempt to reserve it. The first one will successfully reserve it but the second one will fail. Since requests are arriving according to a Poisson distribution with a rate of λ req/sec the probability that no light-path requests occur during a time interval of length d time units is given by:

$$P[\text{no lightpath requests during time interval } d] = e^{-\lambda d}$$

Thus, the Blocking probability for the First Fit Algorithm can be approximated as follows:

$$P_{first-fit-blk} = 1 - (1 - P_{blk})e^{-\lambda d}$$

3) Random Algorithm: In the case of Random algorithm the blocking probability depends on the propagation delay, as was the case with the First Fit algorithm.

The effect of time delays is smaller due to the random selection of the wavelength at the destination node.

Let $Pb_{[w-k]}$ be the probability that $w-k$ wavelengths are in use. Then the average number of wavelengths in use can be calculated as follows:

$$\text{Avg. number of wavelengths in use} = \sum_{k=0}^w (w-k) Pb_{[w-k]}$$

$$\text{Avg. number of available wavelengths} = Aw = 1 - \sum_{k=0}^w (w-k) Pb_{[w-k]}$$

Thus, each wavelength gets requests at a rate of $\lambda' = \lambda/Aw$. Since requests are arriving according to a Poisson distribution with a rate of λ' req/sec the probability that no light-path requests occur during a time interval of length d time units is given by:

$$P[\text{no lightpath requests during time interval } d] = e^{-\lambda'd}$$

Thus, the Blocking probability for the Random Algorithm can be approximated as follows:

$$P_{\text{random}} = 1 - (1 - P_{\text{blk}})e^{-\lambda'd}$$

Figure 4.2 shows the theoretical (dotted lines) and simulated (solid lines) blocking probability vs. light-path request arrival rate graphs for all three algorithms, for 10, 20 and 30 wavelengths. The figure also shows the 95% confidence intervals for the simulation graphs.

a) Linear Topology with 10 wavelengths

b) Linear Topology with 20 wavelengths

c) Linear Topology with 30 wavelengths

Figure 4.2: Blocking Probability Graphs for Linear Topology

From the blocking probability graphs in figure 4.2 it is shown that the simulation results are close to the theoretical approximations. The First Available algorithm has the lowest blocking probability of all three algorithms. Next is the Random algorithm and then is the First Fit algorithm with the highest blocking probability. Looking at the graphs in figure 4.2 one can also see that with larger number of end to end wavelengths, higher lightpath request rates can be achieved. In this case though the difference in performance between the First Fit and Random algorithm as well as the Random and First Available becomes bigger. This is expected since as it was shown in the theoretical analysis First Fit and Random algorithms depend on the probability of requests close together by a time less than the propagation delay of the last link. As the lightpath request generation rate gets higher this probability becomes larger giving a higher overall blocking probability.

In order to better understand where blocking occurs nodal statistics were also collected through OPNET. The bar graphs in figure 4.3 show the nodal blockage of discovery messages at each node for all three algorithms. For the First-Fit algorithm initially the majority of discovery messages are blocked at the destination. This is a result of lightpath requests coming very close to each other and thus generating discovery messages that see the same wavelengths to be available. When the first of these discovery messages reaches the destination node it reserves this wavelength and is no longer available for the other requests which have also seen the same wavelength to be available. As the lightpath request generation rate increases there starts to be more blockage at the source node. This is expected and is an indication of lack of resources. The discovery messages blocked at the intermediate nodes is very small compared to the source and destination nodes. The Random algorithm also has similar results although the

discovery messages blocked at the destination are dominant for a shorter time. In contrast to Random and First Fit, the First Available algorithm has the majority of discovery messages blocked at the source node. The discovery messages blocked at the destination is very small compared to the other two algorithms. This is because as explained earlier, in First Available as long as there are available wavelengths the request should not block.

a) Random (Linear Topology)

b) First-Fit (Linear Topology)

c) First Available (Linear Topology)

Figure 4.3: Nodal Discovery Blockage (Linear Topology)

Similar results were obtained for the case of 20, and 30 wavelengths. Also, all of the blockage in the network were due to discovery messages blocked. The reservation messages blocked were zero, and this is because there are no competing traffic streams. This case is studied in the next section.

4.2 Competing Traffic Streams

In this section we study a case of competing traffic streams. The network topology for this scenario is shown in figure 4.4. Node S makes lightpath requests to node D and node U makes lightpath requests to node W. Both S and U generate requests according to a Poisson distribution with a rate λ req/sec. The holding time of the lightpath is exponentially distributed with average $1/\mu$ sec. Initially none of the nodes is equipped with wavelength converters.

Figure 4.4: Competing Traffic Streams

From figure 4.4 it is clear that the two traffic streams have one common link (X-Y) which they have to share. Thus, as requests are being made by both nodes S and U, there is expected to be competition for the resources (available wavelengths) on link X-Y. The blocking probability graph for all three algorithms in this case is shown in figure 4.5. The figure shows that the Random algorithm in this case gives a lower blocking probability, then comes the First Available, and next is the First Fit algorithm with the highest blocking of all three.

Figure 4.5: Blocking Probability (competition with no conversion)

In order to explain these performance results the nodal statistics were also studied. In general in competition scenarios blocking can be a result either because of discovery message blockage or reservation message blockage. In the previous section we saw that only discovery messages were blocked since there were no competing traffic streams. In this case it was found that there was also blockage due to reservation messages being blocked. Figure 4.6 shows the percentage of discovery and reservation messages blocked in the network for the three algorithms.

a) Random (No Conversion)

b) First-Fit (No Conversion)

c) First Available (No Conversion)

Figure 4.6: Message Blockage Comparison (competition with no conversion)

The figure shows that as the lightpath request rate increases the effect of reservation blockage becomes smaller and discovery messages becomes the dominant reason for blocking. For Random and First Fit we see that at lower rates discovery and reservation messages blocked are close. For First Available though reservation message blockage has more impact on the overall blocking than the other two algorithms for lower rates. This is because less discovery messages are blocked at the destination nodes as was explained in the previous section. Then the destination nodes send more reservation messages towards the source, but once they reach node Y there is higher probability there that they get blocked due to competition if they choose the same wavelength. This was also verified by the simulation results which showed that all of the reservation blockage occurs at node Y.

We also observed at which nodes the discovery messages get blocked. This is shown in figure 4.7. For Random and First Fit initially there are discovery messages blocked at the destination due to the multiple requests problem explained in the linear topology scenario. As the lightpath request rate increases the majority of discovery blockage occurs at node X. This is an indication that at higher rates there is a lack of available wavelengths on the common link X-Y. With the First Available algorithm there is little blockage at the destination nodes and the majority of discovery blockage occurs at node X even at low lightpath request rates. Comparing the bar charts to the ones in figure 4.3 for the single traffic stream case we see similarities. In this case though since there are two competing streams of traffic the blocking does not happen at the actual source

nodes as was the case in the linear topology, rather it happens at the node where the two competing streams branch in to the same link. In our exapmle this is nodeX.

a) Random (No Conversion)

b) First-Fit (No Conversion)

c) First Available (No Conversion)

Figure 4.7: Nodal Discovery Blockage (competition with no conversion)

Having observed the case where no node has converter, we added converter capability to nodes in the network and observed how this affects the overall blocking of the three algorithms. From the non-converter scenario we found that node Y was the node where all the reservation messages were being blocked in the network. Thus, this was the first candidate for placing a converter. By adding wavelength conversion capability to node Y and running the simulations we got the blocking probability graphs for the three algorithms which are shown in figure 4.8.

Figure 4.8: Blocking Probability (competition with converter at node Y)

The graph in figure 4.8 also shows the blocking probabilities for the non-conversion scenario (dotted line graphs), in order to compare the performance in both cases. We see that the performance ranking of the three algorithms is different when using a converter a node Y. In this case First Available gives the lowest blocking probability and then comes the Random algorithm. The First Fit algorithm continues to have the highest blocking probability. This is because the First Available algorithm allows to try other wavelengths at converter nodes until one succeeds, or until it exhausts the list of available wavelengths in which case then the request is considered blocked. This can also be seen from the nodal blocking statistics in figure 4.9. Comparing this to the previous bar graphs in figure 4.6 we see that for the First Available algorithm the

percentage of reservation messages has been greatly reduced. Only discovery message blockage contributes to the overall blocking. The other two algorithms seem not to have benefited from adding the converter, and in fact there is a slight increase in the percentage of reservation messages being blocked. This is because with the Random and First-Fit algorithm a wavelength is selected at converter nodes but is not reserved until the discovery message reaches the destination which then has to send a reservation message back to this node. The longer the time it takes for the reservation to reach back to the converter node, there is higher probability that the selected wavelength is taken in the backward path from another request of the competing stream.

a) Random (Converter at Node Y)

b) First-Fit (Converter at Node Y)

c) First Available (Converter at Node Y)

Figure 4.9: Message Blockage Comparison (competition with converter at node Y)

In figure 4.10 we also look how the blocked discovery messages are blocked across the network as was done for the non-conversion scenario. For Random and First-Fit there is no change from the non-conversion scenario which was shown in figure 4.7. For the First Available algorithm though we see though now all of the discovery blocking is happening at node X and none at the other nodes.

a) Random (Converter at Node Y)

b) First-Fit (Converter at Node Y)

c) First Available (Converter at Node Y)

Figure 4.10: Nodal Discovery Blockage (competition with converter at node Y)

From this scenario we conclude that using a wavelength converter at a node where many reservations messages are blocked improves the overall blocking performance when using the First Available algorithm. If the Random or First Fit algorithms are used there is no improvement.

The next location considered for placing the converter was node X. In the non converter scenario we saw that node X was the main location where discovery messages were blocked. By adding a converter at node X instead of node Y we observed the performance of the three algorithms again. From figure 4.11 we see that the ranking is similar to the case of no converters in the network. Also, from the graphs we see that adding a converter at node X does not add performance improvement to any of the algorithms. Furthermore, by comparing with figure 4.8, we see that the Random and First-Fit algorithms give higher blocking probability from the case where the converter was placed at node Y. This is because the further away the converter is from the destination the longer it takes for the reservation message to reach this node and reserve the wavelength that was selected in the forward path during the discovery phase.

Figure 4.11: Blocking Probability (competition with converter at node X)

This can also be seen from figure 4.12, which shows the percentage of discovery and reservation messages blocked in the network when node X is equipped with conversion capability. Comparing these bar graphs to the ones in figure 4.6, which were for the non-conversion scenario, we see that the percentage of reservation blockage is higher for the Random and First-Fit algorithms. For the First-Available algorithm though it is about the same.

a) Random (Converter at Node X)

b) First-Fit (Converter at Node X)

c) First Available (Converter at Node X)

Figure 4.12: Message Blockage Comparison (competition with converter at node X)

The final case studied for this network scenario was the case of having converters at both nodes X and Y. The blocking probability for this case is shown in figure 4.13. In this graph we compare the performance to the case of having a converter at node Y. We see that the ranking of the three algorithms are the same. But there is no additional improvement in performance, which is a result of the converter at node X, which causes more reservation blockage to occur for the Random and First-Fit algorithm.

Figure 4.13: Blocking Probability (competition with converter at nodes X and Y)

Figure 4.14 shows the percentage of discovery and reservation messages blocked in the network when node X and node Y is equipped with conversion capability.

a) Random (Converter at Node X and Y)

b) First-Fit (Converter at Node X and Y)

c) First Available (Converter at Node X and Y)

Figure 4.14: Message Blockage Comparison (competition with converter at X and Y)

After observing all of the above scenarios we can see that the Random and First-Fit algorithms do not offer improvement when using a wavelength converter and their performance is close to the performance achieved without converters. On the other hand using the First-Available algorithm with a converter can improve the blocking performance. The placement though of the converter in this case is important. It was observed that the optimal placement for the converter in this case was the node where there was high reservation message blockage. By placing a converter at a node where there is discovery message blockage does not offer any additional improvement. In the next section we will look at scenarios where there are more than two competing traffic streams.

4.3 Multiple Competing Traffic Streams

In the previous section we saw that a node at which there is large reservation message blockage is a good candidate for placing a wavelength converter. These nodes are the nodes where there is high outgoing traffic branching off on to multiple links. The previous scenario only had one such node in the network. In this section we study the case where there are more than one node in the network where we have high reservation message blockage. The network for this scenario is shown in figure 4.15. In this scenario node S makes light-path requests to node D, node U makes light-path requests to node W, and node V makes light-path requests to node Q. We have simulated this network with 10 wavelengths on each link. Lightpath requests are generated according to a Poisson distribution with a rate of λ req/sec. The holding time of the light-paths is exponentially distributed with a mean of $1/\mu$ sec.

Figure 4.15: Network with multiple competing traffic streams.

As in the previous case we initially look at the case where there are no wavelength converters in the network. Figure 4.16 shows the percentage of discovery messages blocked at each node as the lightpath request rate increases.

a) First-Fit (No Conversion)

b) Random (No Conversion)

c) First Available (No Conversion)

Figure 4.16: Nodal Discovery Blockage (multiple competition with no conversion)

From figure 4.16 we can see that X and Y are the nodes where the majority of discovery message blockage occurs. We also observe that node Y has higher discovery blockage than X. This is because at node Y as shown in figure 4.15 there is more branching in traffic streams. For Random and First Fit we see also some blockage at the destination nodes for small lightpath request rates. The reason this happens was clearly explained in the previous two sections. In figure 4.17 we also graph the percentage of reservation messages blocked at each node in the network. It was found that all reservation blocking happens at nodes Y and Z for all three algorithms. From the graphs we see that all three algorithms have about the same reservation blockage pattern. We also see that there is about a 10% higher reservation blockage at node Y than node Z. This is because there is more outgoing traffic from node Y than from node Z.

a) Random (No Conversion)

b) First-Fit (No Conversion)

c) First Available (No Conversion)

Figure 4.17: Nodal Reservation Blockage (multiple competition with no conversion)

The blocking probability graph is shown if figure 4.18 for all three algorithms. Again we see that in the absence of wavelength conversion the Random algorithm gives a better performance.

Figure 4.18: Blocking Probability (multiple competition with no conversion)

After collecting the various statistics for the non-conversion scenario we observed how the converter placement affects the blocking performance in the network. Since as it was discussed earlier, Random and First Fit do not offer any improvement when using wavelength converters, in this section we only compare the performance of the First Available algorithm when varying the converter placement in the network. We do compare though with the performance of the Random algorithm in the case of no conversion since it was found that in the non-conversion scenario Random performs better than First Available. In figure 4.19 we plot the blocking probability for the First Available algorithm with the converter placed at different nodes in the network.

Figure 4.19: Blocking Probability (multiple competition, varying converter location)

From the figure we see that by placing the wavelength converter at node Y lowers the blocking probability of the First Available algorithm compared to the non-conversion scenario but is still slightly above the Random algorithm in a network with no conversion. On the other hand if the converter is placed at node Z we see a better improvement than placing the converter at node Y, and also the blocking probability is lower than the Random algorithm in the non-conversion scenario. The reason why placing the converter at node Z results in better performance than placing it at node Y, is because the closer the converter is to the destination nodes there is higher probability that the retries at the converter node will result eventually in a successful reservation. The further away the converter is from the destination node there is a probability that the list

of available wavelengths seen in the forward direction is different on the reverse path during the reservation phase. Finally we see from the same figure that the maximum performance improvement is achieved by placing two converters in the network one at node Y and one at node Z.

In figure 4.20 we also look at the percentage of reservation messages blocked at each node varies as we change the converter placement.

a) First Available (Converter at Node Y)

b) First Available (Converter at Node Z)

c) First Available (Converter at Node Y and Node Z)

Figure 4.20: Nodal Reservation Blockage (varying converter location)

When the converter is placed at node Y, we see that the reservation blockage at node Y is reduced compared to the non-conversion case bar graph shown in figure 4.17. As the lightpath request rate increases though the reservation blocking at node Y starts to get higher. At the same time though the overall percentage of reservation messages blocked in the network gets smaller compared to the discovery messages blocked, which means the main reason for blocking is the lack of resources and not so much the reservation blockage due to competition. Similarly from the same figure we see that when the converter is placed at node Z the reservation blocking at node Z is reduced compared to the non-conversion case. In the case where there are converters at both node Y and node Z the percentage of reservation messages blocked at these nodes is very close and resembles the bar graph we got for the non-conversion scenario.

Chapter 5

Conclusions and Recommendations for Further Study

5.1 Conclusions

5.1.1 IP Control of Optical Networks with OBG

The explosion of the Internet will demand new and innovative solutions for the IP/Optical convergence. This thesis has presented a method called OBG of extending BGP to support light path setup and management across an optical network. The development of OBG has demonstrated an example of the convergence of the IP and Optical domain. The implementation of OBG has been discussed by reviewing current BGP behavior and design requirements for OBG. An implementation of the OBG proposal using simulation tools has been presented, along with various test results.

The main goal of developing and simulating OBG was to extend the BGP protocol to provide added functionality for light path setup. This goal has been met by demonstrating how OBG can be used to set up light paths in a simulation network. These initial demonstrations prove that the proposed scheme for OBG is a viable mechanism and is worthy of future study. The modifications made so far to BGP include the storage of wavelength information in the wavelength table at each node, and the new OBG message type. Some behavior changes have been made to BGP. However OBG is still compatible with BGP. The overall changes are small and acceptable. A seamless migration from BGP to OBG should be possible.

Apart from the simplicity and small changes required for implementing OBGp there are other attractive features of this protocol. BGP is an inter-domain routing protocol that is used in the Internet and, more generally in inter-networks based on IP. By using the underlying routing capabilities of BGP to implement the OBGp protocol allows to configure, setup and manage wavelengths between many different autonomous domains. Using OBGp enterprises at the edge can manage their own wavelength configuration across a carrier's network. The conventional solution to date is for a carrier to offer a managed lightpath service to the customers at the edge. OBGp moves this responsibility from the Carrier to the Edge customers, which now have more control over managing their wavelength space.

5.1.2 OBGp Signaling

After presenting the OBGp protocol this thesis focused on the signaling mechanisms used by OBGp. OBGp was used as a means to implement and compare the performance of various possible signaling algorithms. OBGp uses a backward reservation approach. Backward reservation algorithms have been studied by many researchers and have been shown to perform well in networks without wavelength converters. In the case though of networks which have wavelength converters, the use of backward reservation algorithms such as Random and First-Fit do not offer any additional performance improvement as shown in this thesis. In this paper we have introduced a new backward reservation algorithm called First Available, which makes better use of the wavelength converters capabilities by lowering the average blocking probability. This is accomplished by allowing the nodes that are considered to be the main decision points for selecting

wavelengths to try another wavelength if the first choice does not succeed. Furthermore, First-Available addresses the problem encountered in backward reservation algorithms when multiple connections are set up simultaneously. In this case it is possible that a wavelength available on a link in the forward direction will be taken by a request that comes earlier by a very small time difference. By allowing multiple attempts on different wavelengths, a lower blocking probability is achieved.

It was shown in Chapter 3 that the First-Available algorithm introduces extra control overhead compared to the other algorithms. This overhead increases with the number of converters, thus the networks studied in this thesis were networks with sparse wavelength conversion. In such networks the placement of the few converters affects the overall performance of the network. The right choice of converter placement is critical. By gathering nodal statistics from the network we found that the nodes that have the highest reservation blockage in the network are the candidate locations for the converters. We further concluded that the nodes with highest reservation blockage are the ones with multiple traffic streams branching out on to multiple links.

5.2 Recommendations for Further Study

The OBGp proposal has been shown as a valid way of reserving and creating end-to-end lightpaths from source to destination in an interdomain optical network. Although as shown in this thesis there is value in using the OBGp protocol as a means of setting up lightpaths, further research must still be done, in order to explore and ensure that OBGp will work efficiently and reliably with all scenarios.

Further research has to be done in identifying and ensuring that all possible error conditions can be managed and resolved properly without leaving the system in an instable state. The first phase of OBGp supports detecting message during the signaling part of setting up the lightpaths as described in chapter 2. The protocol takes care of tearing down the connection and releasing any reserved wavelengths. The case though where an error occurs after the lightpath has been set up and is in use, has not been addressed in this thesis. Furthermore, we have assumed a reliable IP control plane in the first phase of OBGp. In reality though it is possible, that a problem can occur in the control plane and be lost after a lightpath has been set up in the optical domain. In this case a way has to be found to control the optical path. This is not a simple problem and needs further study.

In our first approach the OBGp protocol uses the AS path provided by traditional BGP routing, when setting up a lightpath in the Optical layer. This path might be the optimal path between two nodes in the IP domain, but the same might not be true in the Optical domain. For this reason, an alternate routing approach in OBGp might be worth investigating. Currently BGP learns routes from its peers, and then uses a selection process to decide which ones to use. Only one route is stored in the Local Routing Information Base of a BGP router. For OBGp purposes this could be changed to allow for storage of alternate routes as well. Simulations would have to be performed to see how this change would affect the performance of the protocol.

The other main contribution of this thesis was the implementation and performance analysis of the First-Available signaling algorithm used in OBGp. More simulations could be performed with different traffic characteristics. In our simulations

this far we generate lightpath setup requests according to a Poisson process and the holding times of the lightpaths are exponentially distributed. Different topologies can also be simulated.

Finally, it would also be useful to compare the OBGp performance using the First-Available algorithm, to other approaches such as GMPLS. This comparison has not been performed as part of this thesis. Additional work would be required to implement the necessary models in OPNET to test both approaches under similar conditions.

REFERENCES

- [1] X. Yuan et al., "Distributed Control Protocols For Wavelength Reservation And Their Performance Evaluation", *Photonic Networks and Communications*, Vol. 1, No. 3, 1999, pp. 207–218.
- [2] Y. Mei and C. Qiao., "Efficient Distributed Control Protocols for WDM All-Optical Networks", In *Proceedings of International Conference on Computer Communication and Networks (IC3N)*, Las Vegas, NV, Sept. 1997 pp. 150-153.
- [3] C. Qiao and Y. Mei., "Wavelength Reservation Under Distributed Control", In *Digest of IEEE/LEOS Summer Topic Meetings on Broadband Optical Networks: enabling technologies and applications*, August 1996, pp. 45-46.
- [4] X. Yuan, R. Melhem and R. Gupta, "Distributed path reservation algorithms for multiplexed all-optical interconnection networks", Extended version in *IEEE Transactions on Computers*, Vol. 48, No. 12, pages 1355-1363, December 1999.
- [5] C. Qiao and Y. Mei, "On the Multiplexing Degree Required to Embed Permutations in a Class of Networks with Direct Interconnects", in *Proc. IEEE Symposium High-Performance Computer Architecture*, Feb. 1996, pp. 118-129.

- [6] H. Zang et al., "Dynamic Lightpath Establishment in Wavelength-Routed WDM Networks", *IEEE Communications Magazine*, September 2001.
- [7] H. Zang, J.P. Jue, and B. Mukherjee, "A Review of Routing and Wavelength Assignment Approaches for Wavelength-Routed Optical WDM Networks", *Optical Networks Magazine*, vol. 1, no. 1, pp. 47-60, Jan. 2000.
- [8] R. Ramaswami and G. H. Sasaki, "Multiwavelength Optical Networks with Limited Wavelength Conversion" in *Proc. INFOCOM '97*, pp. 489-498.
- [9] O. Gerstel et al., "Fault Tolerant Multiwavelength Optical Rings with Limited Wavelength Conversion", *IEEE Journal on Sel. Areas in Comm.*, Vol. 16, No. 7, Sept. '98, pp. 1166-1178.
- [10] O. Gerstel, R. Ramaswami and G. Sasaki, "Benefits of Limited Wavelength Conversion in WDM Ring Networks", *Optical Fiber Commun. Conf.*, pp. 119120, Feb. 1997.
- [11] S. Subramaniam, and A. K. Somani, "All-Optical Networks with Sparse Wavelength Conversion", *IEEE/ACM Transaction on Networking*, Vol. 4, No. 4, August 1996.

- [12] M. Kovacevic and A. S. Acampora, "Benefits of Wavelength Translation in All-Optical Clear Channel Networks", IEEE J. Select. Areas Commun., vol. 14, pp. 868-880, June 1996.
- [13] R. Barry and P. Humblet, "Models of Blocking Probability in All-Optical Networks with and Without Wavelength Changers", IEEE J. Sel. Areas. Commun. 14, pp. 858-865, (1996).
- [14] D. Saha, "An Efficient Wavelength Reservation Protocol for Lightpath Establishment in All-Optical Networks (AONs)", 2000.
- [15] M. J. Francisco, L. Pezoulas, C. Huang, I. Lambadaris, "End-To-End Signaling and Routing for Optical Networks" in Optical Networks, Opticomm 2001, Denver, Colorado, August 2001.
- [16] L. Pezoulas, M. Francisco, I. Lambadaris, C. Huang, "Performance Analysis of a Backward Reservation Protocol In Networks with Sparse Wavelength Conversion" to be presented at ICC 2003.
- [17] B. St.Arnaud, et al., "Optical BGP Networks" Discussion Paper, March 20, 2000, Revised Draft July 16, 2000.

- [18] Jennifer Yates, Graham Smith, Panagiotis Sebos, Cristina Cannon, Pablo Arias, Jordan Rice and Albert Greenberg, “IP Control of Optical Networks: Design and Experimentation”, Optical Fiber Commun. Conf., March 2001.
- [19] G. Bernstein and V. Sharma, “Some Comments on GMPLS and Optical Technologies”, Internet Draft, Work in Progress, draft-bernstein-gmpls-optical-00.txt, November 2000.
- [20] A. Greenberg, G. Hjalmtysson, and J. Yates, “Smart Routers - Simple Optics: A network Architecture for IP over WDM” Optical Fiber Conference, 2000.
- [21] D. Awduche et al., “Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control with Optical Crossconnects”, Internet Draft, Work in progress, draft-awduche-mpls-te-optical-03.txt, April 2001.
- [22] J. Luciani, B. Rajagopalan, D. Awduche, B. Cain, B. Jamoussi, “IP over Optical Networks – A Framework”, draft-ip-optical-framework-00.txt, September 10, 2000, <http://www.ietf.org/ietf/1id-abstrsacts.txt>.

- [23] Kompella, K., Rekhter, Y., Awduche, D., et al, “Extensions to IS-IS/OSPF and RSVP in support of MPLambdaS”, draft-kompella-mpls-optical-00.txt, <http://www.ietf.org/ietf/1id-abstrsacts.txt>

- [24] ATM Forum, “Private Network-Network Interface Specification: Version 1.0”, March 1996, <http://www.atmforum.org>.

- [25] D. Awduche, Y. Rekhter, J. Darke, R. Coltun, “Multi-protocol Lambda Switching: Combining MPLS Traffic Engineering Control with Optical Crossconnects”, draft-awduche-mple-te-optical-01.txt, <http://www.ietf.org/ietf/1id-abstrsacts.txt>.

- [26] J. W. Stewart, “BGP 4 Inter-Domain Routing in the Internet”, Addison-Wesley, 1999.

- [27] T. Stern, K. Bala, “Multiwavelength Optical Networks A Layered Approach”, New Jersey, Prentice Hall, Inc, 2000.

- [28] J. Banks, J. Carson II, B. Nelson, D. Nicol, “Discrete-Event System Simulation”, Prentice Hall, Inc, 1999, pp 398-417.

- [29] J. Warland, P. Varaiya, “High Performance Communication Networks”, Second Edition, Morgan Kaufman Publishers, 2000.
- [30] T. N. Saadawi, M. H. Ammar, A. El Hakeem, “Fundamentals of Telecommunication Networks”, John Wiley & Sons, Inc., 1994.