Université d'Ottawa • University of Ottawa

Université d'Ottawa · University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

Mazen George KHAIR

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

M. A. Sc. (Electrical Engineering)

GRADE - DEGREE

School of Information Technology and Engineering

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

An Implementation Approach for an Inter-Domain Routing Protocol for DWDM

G. Bochmann

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

CO-DIRECTEUR DE LA THÈSE - THESIS CO-SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

T. Hall

C. Huang

J.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

An Implementation Approach for an Inter-Domain Routing Protocol for DWDM

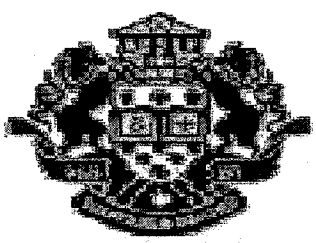
Thesis by

Mazen George Khair

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the Degree of Master of Applied Science, Electrical Engineering

March 15, 2004

Ottawa-Carleton Institute of Electrical and Computer Engineering
School of Information Technology and Engineering
University Of Ottawa
Ottawa, Ontario, Canada



Université d'Ottawa
University of Ottawa

Your file Votre référence
ISBN: 0-494-01509-8
Our file Notre référence
ISBN: 0-494-01509-8

Acknowledgements

First, I would like to thank my supervisors, Dr. Gregor von Bochmann and Dr. Jing Wu, whose comments and support have been a great help during the preparation of this work. I learned many things from them, even though, mainly implicitly rather than explicitly. I would like to thank my colleagues in the Optical Research Group for their cooperation and friendship. I think I should specially mention Abdel Maach for his valuable comments.

I would also like to express my gratitude to my family especially my parents, my twin brother and my sister, without whose help and support, this work would have been impossible.

"Trust in the LORD with all your heart and lean not on your own understanding; in all your ways acknowledge him, and he will make your paths straight." Proverbs 3:5, 6.

Abstract

Today's Internet applications are in need of additional bandwidth. Consequently, a large number of high capacity fiber optical networks have been deployed. The rapid increase in this need has resulted in more than one optical Autonomous Systems (AS) per geographical area. In order to make optimal use of these widely spread optical networks, an intelligent approach is needed to manage the resources of the networks.

The ultimate goal is to establish an end-to-end light path connection across multiple Autonomous Systems (ASs). The initiation of such a connection should be based on routing information that reflects the most recent topology of the network resources such as the available wavelengths, thus guaranteeing a lightpath set-up.

In this thesis, we propose a new routing protocol, called Optical Routing Border Gateway Protocol (ORBGP). ORBGP exploits most functionalities of the BGP routing protocol with some modifications to support the exchange of optical routing information. Furthermore, ORBGP introduces two new advertising policies to allow the edge nodes to advertise routing changes based on their needs. The performance of this protocol for different network

architectures has been investigated through simulation studies, and the results show promise.

Table of Contents

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Simple Network

ARPANET

Ring Architecture

European Network

Star Architecture

Appendix A

Appendix B

Appendix C

Appendix D

Appendix E

Appendix F

List of Tables

Chapter 3

Chapter 5

Chapter 6

Chapter 1 Introduction

1.1 Focus of the Thesis

As optical network deployment gains momentum, many countries have developed their own networks to support their national needs. Normally, each of these networks is partitioned into sub-areas to ease the management in the control plane. In fact, since these partitions belong to the same network administrations, the type and the amount of shared information among these partitions become an important issue, because shared information facilitates smart selection of the desired path with better Quality of Service (QoS) [BSO 02]. For example, exchanging summarized routing information among these areas about the paths that has the lowest cost and the lowest number of hops with a minimum delay will certainly promote a better QoS.

Fortunately, managing the control plane for a large network that belongs to a single network administration is relatively easy, because the network administrator has complete control over its network entities, in other words, he/she can have a complete knowledge of the topology and the resources available within the Autonomous System (AS). However, things become more challenging when different network administrators want to share each other's resources. Each AS is responsible for summarizing its routing information to advertise after careful analysis of the type and amount of information needed. This analysis is necessary due to privacy and security issues and the scope of shared information is defined by business agreements [BSO 02].

In optical networks, the major goal is to define a way for establishing a lightpath across multiple domains. This can be done by defining an interaction mechanism between intra-domain and inter-domain routing protocols. We investigate different protocols for intra-domain routing such as the so-called Intermediate Systems to Intermediate Systems (IS-IS) and the Open Shortest Path First (OSPF) protocols. The result of this study was to use OSPF as intra-domain routing protocol for optical network as a recommended protocol by the Internet Engineering Task Force (IETF). Besides, OSPF has already been extended to support optical routing [RFC 2370]. On the other hand, it was decided to extend the Border Gateway Protocol (BGP) to support optical routing because BGP is the only well known routing protocol that is used across ASs.

Our goal here is to enhance the BGP protocol to exchange optical routing information among different domains to support lightpath setup across different domains. The idea is to pass optical information, such as available wavelengths, obtained locally by OSPF to other domains by means of BGP. As was explained earlier, each network can be partitioned into sub-areas. Routers within each partitioned area exchange complete routing information among themselves; Routers that belong to the same partitioned area pass summarized routing information to other sub-areas through Area Border Routers (ABR). As a result, an ABR can extract the routing information that can be used to reach any sub-area within the same large optical network.

In fact, some of the ABRs can act as edge routers of the whole optical network. Therefore, those edge routers can summarize the routing information that is found in the sub-area level to find the useful routing information that can be used to reach each others. The useful routing information that can be used by edge routers to reach each other could be advertised to other large optical networks after careful analysis of the type of information to

be advertised. The advertised information could be passed by means of BGP update messages in the form of an optical attribute. Of course, this attribute would not include any specific detailed information about the AS that originates this update message. The only piece of information exchanged by this attribute is the available wavelengths that can allow other ASs to setup lightpaths that cross this specific AS for a specific destination.

1.2 Objective

The objective of this thesis is to develop a routing protocol for Inter-Domain routing for wavelength-routed optical networks. This protocol is based on the functionality of BGP. In other words, our routing protocol exploits BGP operation mechanisms and extends some of its features to include the exchange of optical attributes to support routing among optical ASs and modifying the advertising scheme of BGP to include information on the number of reserved or released wavelengths.

We are interested in investigating how successfully we can establish an end-to-end lightpath based on the available routing information exchanged by our proposed routing protocol. Besides, we are going to show the interaction between intra-domain and inter-domain protocols to support light path setup across multiple domains.

Unfortunately, the routing information provided by our proposed protocol through this interaction mechanism could not guarantee a light path setup for a number of reasons. For example, there might be a delay in reporting any routing information that has been changed due to the fact that each node may decide to advertise after a fixed amount of time. This delay makes the recent changes invisible to other edge nodes and hence these changes

become errors in the routing tables of other nodes and will lead to blocking if this wrong information is used in an attempt to establish a new lightpath.

Clearly, the frequency of these errors could be decreased depending on some parameters such as how frequent the routing information within each autonomous system (AS) and among ASs is refreshed.

Our goal is to study the performance of our proposed protocol depending on various parameters such as the rate of request, service time and refreshing period.

1.3 Thesis Organization

This thesis is organized as follows. Chapter 2 gives a quick overview of Dense Wavelength Division Multiplexing (DWDM) networks and IP over WDM architecture models. Chapter 3 describes the basic operation of most Internet routing protocols, and gives an overview of some intra-domain and inter-domain routing protocols such as RIP-v2, IS-IS, OSPF and BGP. It also discusses the advantages and disadvantages of these routing protocols in order to identify the best possible routing protocol that can be used for routing in optical networks. Chapter 4 discusses some related work. Chapter 5 discusses in detail the proposed protocol that performs routing among optical ASs. Chapter 6 discusses the selection process of the simulation tool and illustrates the simulation results. Finally, Chapter 7 presents a discussion of the results and draws some conclusions.

Chapter 2 Overview of DWDM Optical Networks

2.1 Generations of Digital Transport Networks

The simplest definition of the Internet is a collection of huge number of interconnected networks that are located in different geographical areas [Mou et al 03]. The first generation of digital carrier systems serving T1 and E1 links and was used for the first time in 1960 [Liu et al 02].

Due to the tremendous and continuous need for new applications that require huge bandwidth to meet the new style of human life, a second generation (2G) of digital carrier systems was deployed to achieve higher data bit rate such as the Synchronous Optical NETwork (SONET) [Tom et al 01]. Although, 2G had great features such as high data bit rate, nevertheless, it suffers form many weak points. First, it has low recovery time that can not be avoided unless a dedicated protection fiber was assigned to each channel [Tom et al 01]. Second, 2G networks have no routing control mechanism that works in a distributed fashion [Tom et al 01]. Finally, 2G creates compatibility problems among networks provided by different vendors.

To overcome these weak points defined in 2G networks, a third generation (3G) network architecture was proposed. 3G networks allows processing the data in the light domain without conversion of the data from the optical domain to the electrical domain and back to optical, as in 2G networks [Bla et al 02].

2.2 IP/WDM Architectural Modules

Internet Service Providers (ISP) provide different types of services such as Asynchronous Transfer Mode (ATM), Time Division Multiplexing (TDM) and IP. These services use different networks that work in different layers. Figure 2.1 shows the layering structure. The first approach was IP over ATM over Sonet over WDM layer.

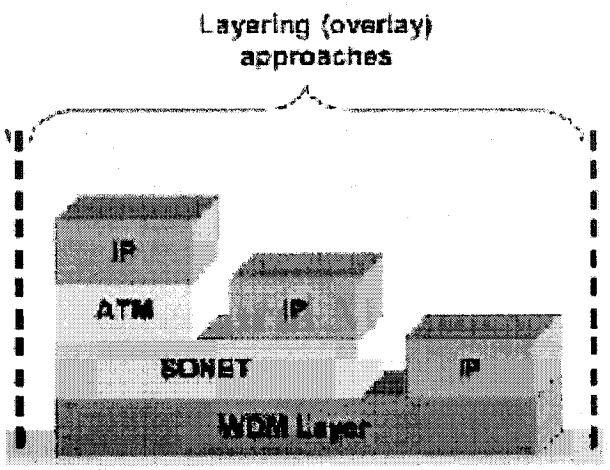


Figure 2. 1: layering structure

Much research has been done to find the best layering approach. For many reasons, IP over wave division multiplexing (WDM) approach has become the most preferable approach because it reduces the number of layers and hence decreases the cost. Furthermore, this approach reduces the overhead resulting from ATM or SONET [Mou et al 03]. Besides, the implementation of IP Gigabit routers makes it possible to avoid relying on SONET technology to produce high bit rates that can use the huge bandwidth available in WDM systems [Bla et al 02].

As a matter of fact, the new vision of the new Internet looks like that: each IP router should be capable of having the ability to establish or to tear down a light path to a certain

destination to reduce the pressure on its IP forwarding system [Wei 02]. This can be done by ensuring optical DWDM networks provide IP networks with enough information about the available wavelength that can be used to reach the desired destination through the optical domain[LLW 00], [SS 03], [CZ 96].

Chapter 3 Routing Protocols

3.1 Overview of Internet Routing Protocols

In this chapter, we will describe most of the important Internet routing protocols. We will consider the most common routing protocols such as Routing Information Protocol (RIP), IS-IS, OSPF and BGP.

Usually, routing protocols are compared on the basis of whether the routing protocol is an Internal Gateway Protocol (IGP) or External Gateway Protocol (EGP). Nevertheless, Internet routing protocols can be divided into two different categories in terms of operation mode [Ste et al 99]. The first category is called Distance Vector protocol (DV) while the other is known as Link State protocol (LS). Understanding these two modes of operation allows us to analyze the functionality, performance and scalability of all routing protocols.

3.1.1 Distance Vector Protocols (DV)

Figure 3.1 shows a small network topology that will be used to illustrate how DV protocols work. Normally, all routers pool their capabilities in achieving a distributed computation algorithm [Ste et al 99]. As a result, each router will know the best path to each destination. Usually, the routing algorithm tries to find the paths that have the minimum number of hops to the desired destination [Moy et al 98].

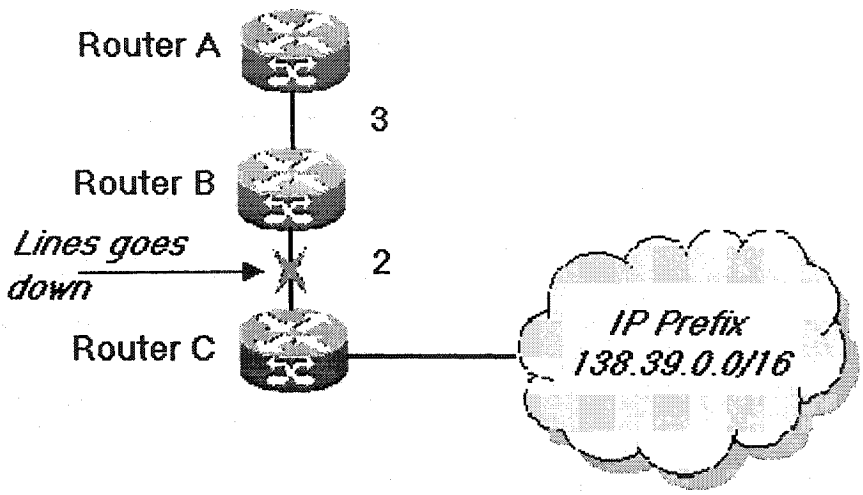


Figure 3. 1: Sample topology for DV

The purpose of the example illustrated in figure 3.1 is to show how Router A and B can forward packets to the destination 139.39.0.0/16. It is clear that Router A obtains its routing information from Router B which obtains its routing information through Router C [Ste et al 99]. Router C is connected directly to the host that has the following address 139.39.0.0/16 that is to be advertised. As a matter of fact, Router C has to inform Router B about the prefix 139.39.0.0/16 and Router B will pass this information on to Router A. The bold numbers between the routers represent the cost between routers. So, Router C sends a message to inform Router B that B can reach the prefix 139.39.0.0/16 through C with a cost equal to 2 [Ste et al 99].

Normally, a DV routing protocol running on any node sets up a table of prefixes to which it is connected. In our example, the initial routing table for Router C is shown in table 3-1.

Destination Prefix "Network"	Cost	Learned From
139.39.0.0/16	0	Self

Table 3. 1: Router C at Start-up

When Router B receives the advertisement from Router C, its routing table will look like Table 3-2.

Destination Prefix "Network"	Cost	Learned From
139.39.0.0/16	2	Router C

Table 3. 2: Router B after receiving routing information from C

Router B will pass this information to Router A after adding the cost between router A & router B which is equal to 3. As a result, Router B will have a routing table that looks like table 3-3 [Ste et al 99].

Destination Prefix "Network"	Cost	Learned From
139.39.0.0/16	5	Router B

Table 3. 3: Router A after receiving routing information from B

If Router A learns more information by any mean to new destinations, Router A will simply creates new entries for each destination that shows the cost and source of routing information for that destination [Ste et al 99].

DV protocols are relatively easy to be implemented and understood [Ste et al 99]. Nevertheless, DV protocols have many disadvantages. One of them is that since the Internet possesses a large number of prefixes, each routing table has to keep a large number of entries because each entry represents a way to reach a certain destination [Moy et al 98]. Consequently, exchanging these entries not only means a large number of routing messages will traverse the network but also it means that the entire routing tables is being exchanged [Moy et al 98].

One more disadvantage of the DV protocol is "counting to infinity". To understand this feature of DV protocols let us consider figure 1.1 that shows a failure between Router B and Router C. Certainly, Router B and C can detect this failure and both respond by removing the useless routing information [Ste et al 99]. Hence, Router B will remove the entry for the destination prefix 138.39.0.0/16 because it is no longer reachable and Router C removes the entry for the destination Router B. Router A, however, does not see that Router C is no longer connected to the network. Therefore, Router A will advertise this prefix 138.39.0.0/16 to Router B. Router B assumes that this information is a valid routing information to the prefix 138.39.0.0/16 and stores it into its table. When the time comes to refresh the routing table router B will send the prefix 138.39.0.0/16 to Router A after adding the cost 3. Again, Router A will store this invalid information. This behavior is called counting to infinity because this useless behavior keeps repeating until the cost of the unreachable destination reaches a maximum value at which it is considered to be unreachable [Ste et al 99].

3.1.2 Link State Protocols (LS)

LS protocols are completely different from DV [Ste et al 99]. In fact, unlike DV protocols LS routing algorithms use a distributed database approach [Moy et al 98]. Moreover, LS protocols have a relatively fast convergence time such that new routes are found quickly with minimum protocol overhead [Moy et al 98]. Usually, LS node generates a Link State Packet (LSP) when it has detected a new neighbor or the cost of the link to an existing neighbor has been changed [IDBer00].

Figure 3.2 shows an example for how LS protocols function [Ste et al 99]. Each router sends a Link State Packet (LSP) to all its neighbors. This LSP contains all the information about the links and the networks that are connected to the advertising node; the LSP also includes the cost for each link [Leo et al 01]. After each node receives a complete set of LSPs for the network each node can generate a topology for the whole network [Moy et al 98].

In our example, the topology shown in figure3.2 can be exchanged in LSPs. Table 3-4 has four columns each representing an LSP for router A, B, C and D respectively [Ste et al 99].

LSP for Router A		LSP for Router B		LSP for Router C		LSP for Router D	
Destinations	Cost	Destinations	Cost	Destinations	Cost	Destinations	Cost
138.39.0.4/30	0	138.39.0.8/30	0	138.39.0.12/30	0	138.39.0.16/30	0
Router B	4	Router B	4	Router A	3	Router A	10
Router C	3	Router C	1	Router B	1	Router C	3
Router D	10			Router D	3		

Table 3. 4: Shows the LSP for Router A, B, C and D

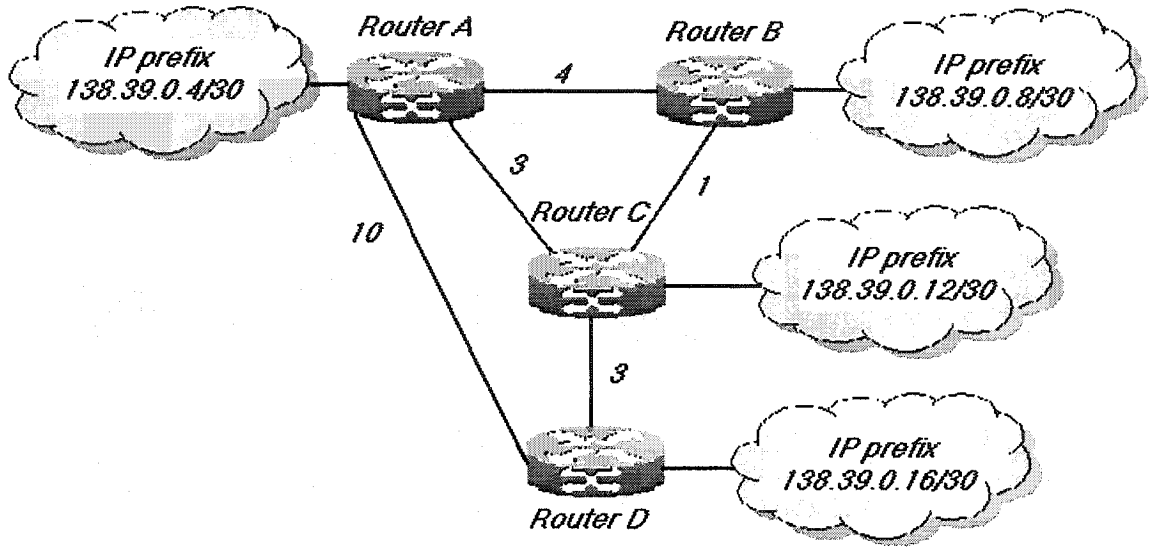


Figure 3. 2: Sample topology for LS

Normally, each router will have a copy of Table 3-4 that it can use in conjunction with Dijkstra's algorithm to find the best path to each destination [Moy et al 98]. As a result, each router will create a routing table that look like Table 3.3 for DV protocols.

Although LS protocols are powerful protocols nevertheless they are more complicated and more difficult to be implemented than DV protocols [Moy et al 98].

3.2 Intra-domain Routing Protocols

Today's Internet has many different routing protocols that are used to exchange information on the topology of the network that allows each individual router to make a relatively smart decision on how to forward its data packets to avoid congestion areas or to choose the path that has the best quality of service (QoS) [Moy et al 98]. In more complex networks, this is not an easy task to perform. Therefore, a more sophisticated routing protocol is necessary to achieve the optimal use of the network resources. These routing protocols should have the ability to exchange enough routing information such as the cost, delay, available bandwidth –

etc. Unfortunately, the larger the network the more information that has to be exchanged and hence scalability becomes a serious issue [Sta et al 00]. Recently, many routing protocols have been deployed in the Internet such as RIP, OSPF and IS-IS that will be reviewed in the following.

3.2.1 Routing Information Protocol (RIP v2)

RIP is a DV type routing protocol. RIP is based on the Bellman-Ford distance-vector routing algorithm [RFC2453]. This algorithm basically merges routing information provided by different routers into lookup tables. This algorithm has been used for routing computations in computer networks since the early days of the Advanced Research Projects Agency Network (ARPANET) [RFC2453]. RIP was designed to work as an Internal Gateway Protocol (IGP) in moderate-size Autonomous Systems (AS) [RFC2453]. Unfortunately, this protocol has many limitations. The most important one is that this protocol can not be scaled up to a large network because it does not support hierarchical routing [RFC2453]. Furthermore, as it is explained for the DV protocol, this protocol depends upon "counting to infinity" to resolve certain unusual situations [RFC2453]. Finally, unlike OSPF and IS-IS, RIP-v2 can not be used for networks supporting real-time applications because it uses a fixed "metric" to compare alternative routes instead of using real-time parameters such as measured delay, reliability, or load [Tho 01].

3.2.2 Intermediate System-Intermediate System (IS-IS)

IS-IS is an IGP for the Internet, used to exchange IP routing information among routers that belongs to the same AS. It was originally invented as a routing protocol for Open Systems

Interconnection (OSI[1]), however, it has been extended to include IP routing; the extended version is sometimes referred to as Integrated or Dual IS-IS [RFC1195].

Integrated IS-IS is a LS protocol type that is not widely deployed; on the other hand, it is used in a few relatively large networks that use Cisco Routers [Moy et al 98].

In a network that uses IS-IS as a routing protocols, each router makes use of its Protocol Data Unit (PDU) packet to originate a Link State PDU (LSP) that includes all the routing information about its neighbors. After the reception of all LSPs, each node performs Dijkstra algorithm processing to construct the routing table [RFC1195].

IS-IS domains support hierarchy routing by dividing the AS into IS-IS areas. Hierarchical routing is achieved by having two levels of routing [RFC1142] (see section 3.2.4).

The first level has nodes that deliver and receive LSPs from nodes that belong to the same area [RFC1142] whereas, level 2 routers act as area border nodes that perform routing among sub-domain that belong to the same AS.

3.2.3 Open Shortest Path First (OSPF)

OSPF is a LS type protocol. As in IS-IS, each router generates a packet known as a Link State Advertisement (LSA) [Moy et al 98]. Each LSA has complete information about the cost of the links to neighbours, besides it contains the information about the networks connected to that specific router. Figure 3.3 shows the structure of LSA [RFC 2328].

[1] OSI: This model is known as the seven layer model. It is defined by the International Organization for Standardization for network protocols. The seven layers are: 1- Physical, 2- Data link, 3- Internet, 4- Transport, 5- Session, 6- Presentation and finally 7- Application layer.

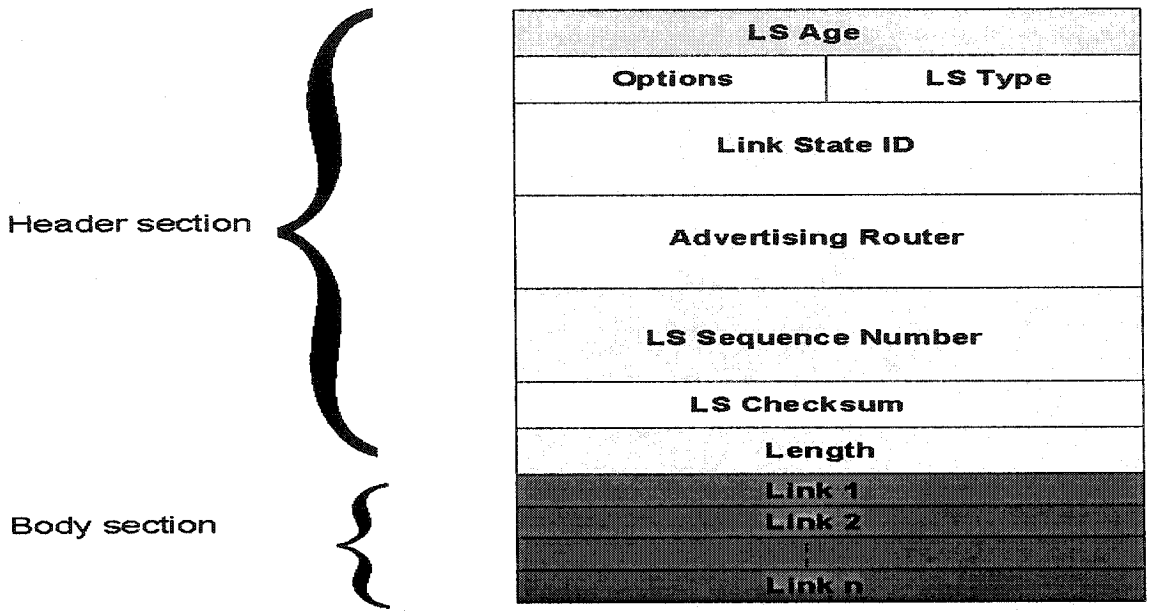


Figure 3. 3: LSA

After a node receives a LSA from all routers that belong to the same AS, each node runs Dijkstra's algorithm to find the best path for each destination.

All LSA have the same structure; we will explain briefly the task of each field in the header:

LS Age field: This field is set to zero once the node originates its LSA. This field is incremented by units of seconds at each hop to compensate for transmission delay. This field is also incremented at each node that stores this LSA. The LSA will be dropped if the LSA age field reaches its maximum value because it becomes old routing information [RFC 2328].

Options field: This field indicates how this LSA should be routed through the AS or how it could be handled [Moy et al 98]. This is due to the fact that each router has certain capabilities. For example, an edge router includes in the option field that it can handle external routing information; also a router may include in the option field that it can support

multicast routing. Therefore, this option field can help routers identify each other's capabilities.

LS type: This field specifies the type of routing information. For example, if the value of this parameter is 5 then the routing information is obtained from an external routing protocol like BGP whereas if the LS type field equals 9, 10 or 11 then the routing information is optical information.

Link State ID and Advertising Router fields: Each LSA has unique values for these two fields that identify the originator of LSA [RFC 2328].

LS Sequence Number: This field is incremented every time a router originates an LSA; this field helps to prevent any node from accepting old copies of LSA.

LS Checksum: This ensures that the received LSA has been received free of error.

Length: The length of the entire LSA.

Clearly, the length field is the last field in the header in figure 3.3 and what follows is the routing information for each link of the router originating this LSA.

3.2.4 Hierarchical Routing

Hierarchical routing is an existing feature in both IS-IS and OSPF. Both protocols allow each AS to be divided into more than one routing area. The objective is to reduce the number of LSA traversed back and forth. For example, each area exchanges complete information among its routers and exchanges routing information with other areas through Area Border Routers (see Figure 3.4).

For consistency, we believe that there should be three routing levels, the first level deals with routers located within each area, the second routing level is managed by Area Border Routers (ABRs) that summarize the routing information passed by the first level to

allow different ABRs that belongs to different routing areas to reach each other, whereas the third routing level obtains the useful routing information from the second level to allow edge routers of the same AS to identify the available resources among each others. The information stored in the third routing level is the useful information for inter-domain routing.

To illustrate the functionality among the three routing levels, let us consider the following scenario, Figure 3.4 shows the first routing level of AS_2. We are assuming that AS_2 uses OSPF as its routing protocol. OSPF has the concept of hierarchal routing. The AS can be divided into more than one area.

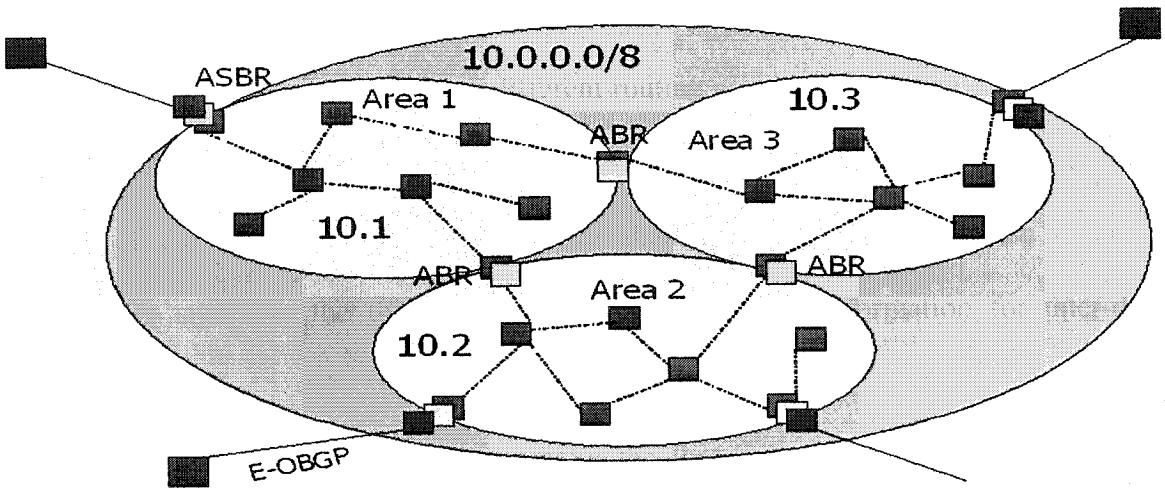


Figure 3. 4: AS-2 using OSPF as IGP

Clearly, AS_2 has three areas, routers within the same area exchange complete routing information. This information is used by each router within each area to construct a map for all destinations in this specific area. Each area exchange routing information with other areas through Area Border Routers (ABR). ABR can have access to routing information for both areas that are connected to this ABR.

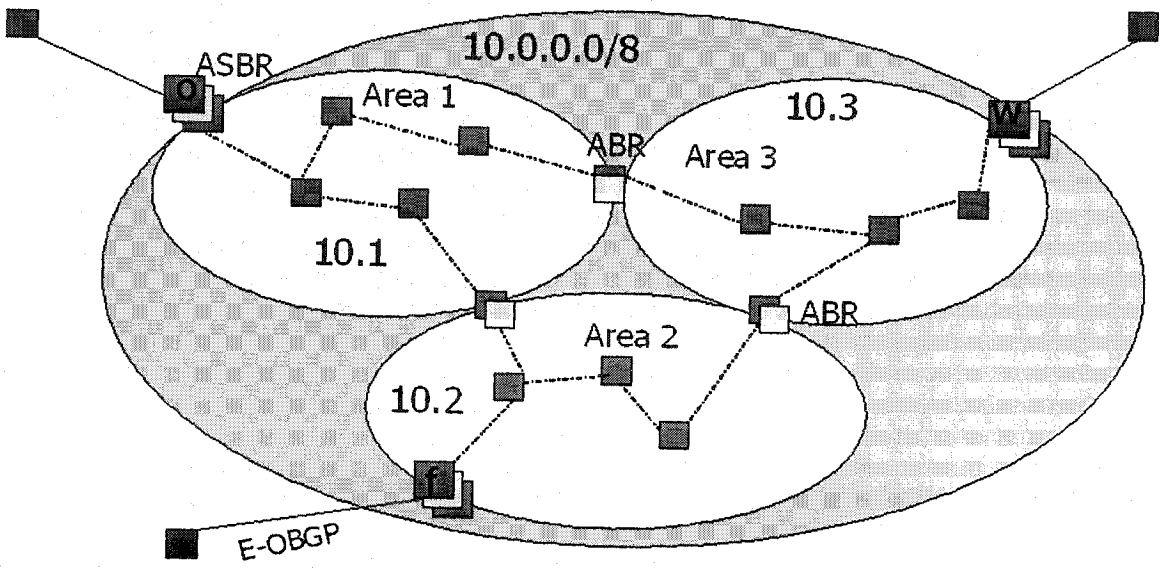


Figure 3. 5: AS_2 Area Level

ABR keeps information about how to reach other ABR located in different areas. Figure 3.5 shows the topology from the ABR point of view. For example, ABR between area 2 and 3 has the knowledge about the available wavelength that can be used to reach other ABRs located anywhere in AS_2.

However, if the network administrator of AS_2 finds that he/she has lots of resources to share with other ASs, this network administrator has to advertise an abstract version of the available resources to other AS. The question here is what should AS_2 advertise to other ASs? The answer may be deduced from Figure 3.6. Figure 3.6 shows the third routing level. This level of routing keeps track of the available resources among edge routers that belong to AS_2. For example, the edge router W keeps only the information about the resources that can reach other Edge Routers in AS_2, in our case router O and F.

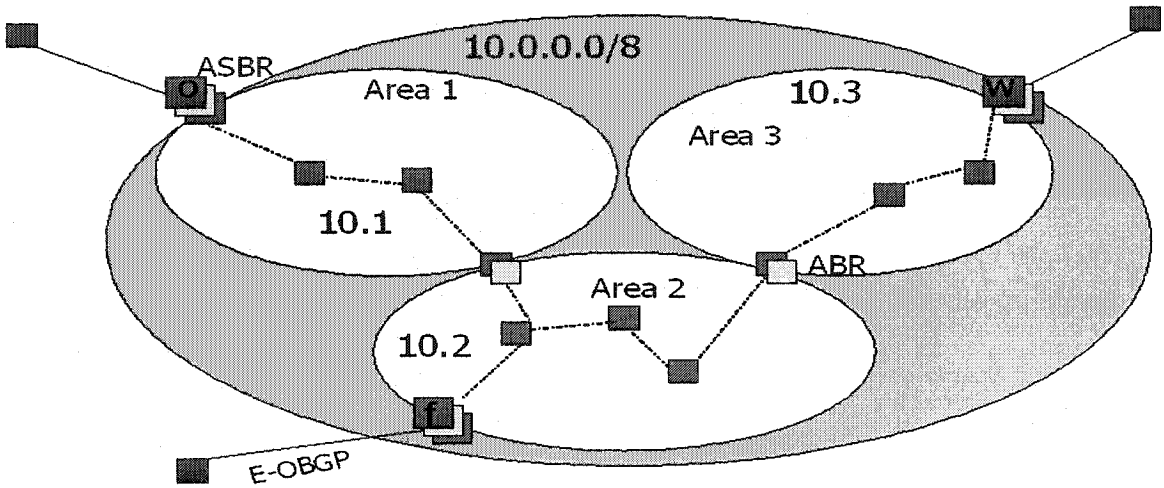


Figure 3. 6: AS_2 in the AS level

The information that is obtained in the third level could be advertised to other ASs. Of course, this information is a summary of the path that allows other AS to find paths that cross AS_2.

As a matter of fact, Edge Routers or Autonomous System Border Router (ASBR) should have extra space to store information about how to reach other ASBR in the same AS and other edge routers in different ASs. Besides, it stores information about how to reach ABR and finally it stores information about how to reach each router within the area that this ASRB belongs to. Figure 3.7 shows the required routing information in each routing level.

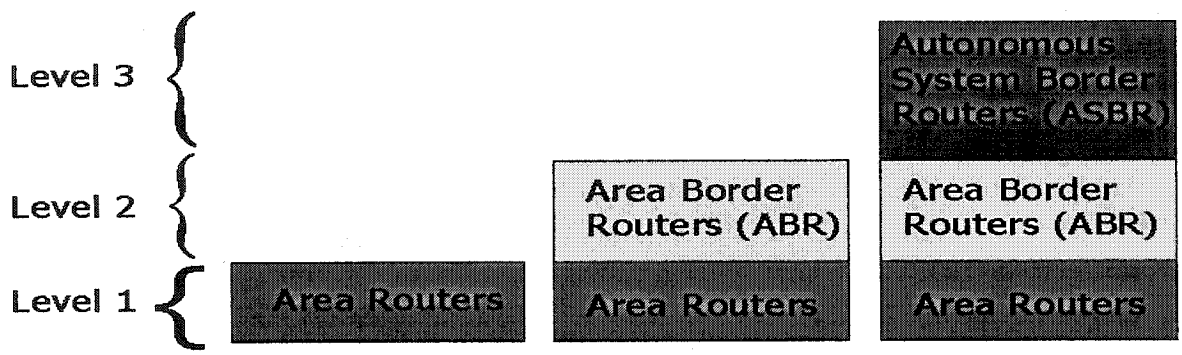


Figure 3. 7: Routing information in each level

3.2.5 Contrast among Intra-Domain routing protocols

In the previous sections, we investigated three intra-domain routing protocols to find out which one of them could be used for routing in optical networks. Based on our study, we made the following decisions:

First, we decided that RIP should not be used because of its weak points in terms of slow convergence due to certain network failures as was explained earlier in Section 3.1.1.

Second, we find that it is hard to choose between IS-IS and OSPF because they are very similar. In other words, there is no clear winner. However, we chose OSPF because the Internet Engineering Task Force (IETF) supported this protocol as an intra-domain routing protocols for optical networks [IDBer00]. Besides, OSPF is more efficient in using bandwidth. Unfortunately, this efficacy is counter balanced by the fact that OSPF is more complex [IDBer00]. Furthermore, OSPF can achieve better scalability because it can exchange more routing information than IS-IS [IDBer00]. The reason for this is that OSPF can have larger LSA packet sizes than IS-IS because IS-IS has a fixed routing packet size.

3.3 Inter-domain Routing

3.3.1 Border Gateway Protocol (BGP)

The *Border Gateway Protocol (BGP)* is an external routing protocol. BGP is a Path Vector (PV) type protocol. In PV, each border router advertises the destinations it can reach to its neighboring *Edge Routers* along with the information that describes various properties of the paths to these destinations. In other words, PV defines the route as a pairing between the destination and the attributes of the path to reach that destination. Thus the name path-vector

comes from the fact that each edge router receives from its neighboring edge router a vector that contains a set of routes [RFC 1322].

BGP is used to exchange routing information among different ASs located in different geographical regions [Moy et al 98]. The main task of BGP is to allow edge routers that belong to a certain AS to exchange network reachability information with other BGP systems located in different ASs [Ste et al 99]. This network reachability information includes information on the sequence of ASs that the reachability information passes through.

This information is used by edge routers to build a map of how ASs are connected to each other. Besides, each AS can make use of this routing information to detect and prevent routing loops. Loops can be detected because each AS can reject any routing information that has its AS name in it. Furthermore, BGP can give each AS the ability to apply a certain routing policy by allowing each AS to set the cost of the advertised link or prevent certain routing information to reach certain costumers [RFC1771].

Since BGP is a well-known protocol for inter-domain routing, we decided to adopt this protocol to achieve inter-domain routing among optical ASs and to make use of its routing features rather than to implement something from scratch.

BGP has been deployed in the Internet for a long time and it has the following features [RFC1771], [Moy et al 98], [Ste et al 99]:

- BGP functionality based on DV algorithms that tries to find the minimum number of hops to the desired destination [RFC1771].
- BGP relies on TCP as its carrier protocol with few modifications to achieve security [Tho 01].

- BGP has the ability to apply policy-based control; because BGP is manually configured, therefore each BGP router can decide which received addresses can be accepted. Besides, it can prevent some addresses propagateing to certain customers [Moy et al 98].
- Usually, edge routers that speak BGP advertise only one prefix to one of its BGP neighbours within an update message. Normally, the advertised address/prefix is associated with a number of attributes such as Cost, Next_hop, As_path –etc [Ste et al 99]. We will discuss these attributes later in this chapter.

Figure 3.8 shows a simple scenario for a network that has multiple ASs connected together. Clearly, there are five ASs, each has a number of edge routers. This scenario also shows the geographical working area for each routing protocol; intra-domain routing protocol like OSPF and inter-domain routing protocol like BGP.

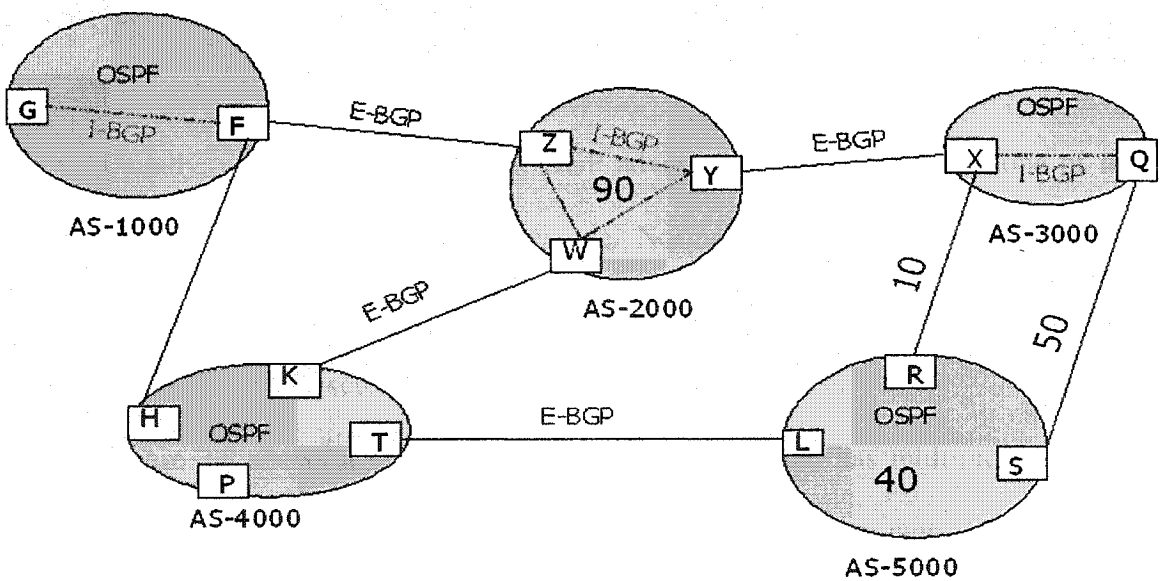


Figure 3. 8: A network scenario consisting of five ASs

During a BGP session between any two edge routers, the two participating peers exchange update messages. Each update message includes one prefix associated with a certain number of attributes such as AS_Path, Next_hop, cost --etc [RFC 1771] [Ste et al 99].

3.3.1.1 BGP Messages Types

BGP has four types of messages that share the same header structure. The four messages are: Open message, Notification message, Update message and finally Keep alive message [Tho 01].

Open message: is the first message that is used by any edge router after the TCP connection has been established. Usually, both participants send each other this message to identify each other [RFC1771].

Update message: this message is used to add or remove a prefix. For example, if a certain prefix no longer exists then the BGP update message will include this prefix in the withdraw section whereas if a new route "prefix" has been discovered then this prefix will be advertised as a new route [RFC1771].

Notification message: this message is used to report a fault during a BGP session [RFC1771].

Keep alive message: this message is sent during a BGP session between two BGP speakers to confirm that the BGP session is still alive [RFC1771]. The BGP session is kept alive until the two participants exchange their new routes.

3.3.1.2 BGP Path Attributes

BGP attributes are very important for any advertised route because these attributes allow the receiving BGP system to understand the reachability information for the received path [Tho 01]. Although, there are only seven path attributes defined by IETF in RFC1771, many new attributes are defined by other RFC documents to enhance BGP to support other functionalities.

In this section, we shall only discuss the seven attributes that are documented by the RFC1771. Usually, a certain number of these attributes are exchanged by a BGP update message to describe a route. The seven attributes are:

1- Origin Attribute: The Origin attribute can have three values: 0, 1 or 2. If the value of the Origin attribute is 0, then the Network Layer Reachability Information (NLRI) is received via an Interior routing protocol such as OSPF. If the Origin attribute has a value of 1, then the NLRI learned via an exterior routing protocol such as BGP [RFC1771]. Finally, if the Origin attribute has a value of 2, then the NLRI learned via a different means. It could be set for example by the network administrator [Tho 01].

2- AS_Path Attribute: The AS_Path attribute is essential to prevent loops that may occur among ASs [RFC1771]. To understand the benefit of this attribute, we will consider the example in figure 3.8. Assume that AS_3000 in figure 3.8 has a BGP session with AS_2000. This session will allow AS_2000 to construct a route to AS_3000. In turn, AS_2000 can have another BGP session with AS_4000 to pass on the route that was learnt from AS_3000. As a result, AS_4000 will get enough information to enable it to reach AS_3000 through AS_2000. Now AS_4000

performs the same task to pass the AS_Path information to AS_5000. If AS_5000 decided to advertise the route obtained by AS_4000 to AS_3000, AS_3000 would not accept the route because AS_3000 could easily detect that this route was originated by itself and hence, a routing loop is prevented.

3- NEXT_HOP Attribute: Usually, the value of this attribute is the *virtual interface*[2] or *loopback interface* if there is an internal session to exchange routing information among edge routers that belong to the same AS, whereas this attribute takes the IP address of the physical link of the source edge router within an External BGP session with the neighbor AS.

4- MULTI_ EXIT_DISC Attribute: This attribute is used when there is more than one connection between two adjacent ASs. This value can assist with choosing the optimal path across this AS [Ste et al 99]. For example, figure 3.8 shows that AS_5000 is connected with AS_3000 via two connections. AS_5000 can give different values for each connection when it advertises its routing information to AS_3000. Assume that AS_5000 will give a value equal to 10 for the connection between router S & Q, while it gives a value equal to 50 to the connection between router R & X. These values will guide AS_3000 to select between the two links. AS_3000 will choose the connection that has lower MULTI_ EXIT_DISC value. In this example, AS_3000 would choose the X-R connection because it has a value equal to 10, which is lower than the S-Q connection, which has a value of 50.

5- Local_Pref Attribute: there could be many possible ways to cross a given AS. This attribute can be used locally by each AS to control the path for crossing that AS. In

[2] Virtual interface is an address used for identifying a router, it has no relation with any physical or hardware interface. This address is very useful for performing Intra-Domain routing when there is no direct physical connection between edge routers that belong to the same AS.

other words, which edge routers that belong to the same AS will be used to transport the traffic [Ste et al 99]. For example, Figure 3.8 shows that Router Y in AS_2000 can allow Router X located in AS_3000 to reach Router F located in AS_1000 through two different paths. The first path could be through Y, Z whereas the other could be through Y, W and Z. In fact, the Local_Pref attribute allows edge routers Y, Z, and W that belong to AS_2000 to agree among themselves on the path that should be used to handle a certain path request.

6- Atomic_Aggregator Attribute: When two routes have overlapped at a certain edge router, this router can use this attribute to tell neighbor edge routers about this route overlapping [Tho 01].

7- Aggregator Attribute: It is necessary to aggregate more than one prefix into a single prefix for reason of scalability. This attribute indicates the AS and the router that perform the aggregation. For example, AS_2000 can aggregate the addresses of three ASs,(AS_3000, AS_4000, and AS_5000), into one address/prefix. This new aggregated address can be advertised to AS_1000.

As mentioned above, there are more than these seven attributes. Table 3-4 shows all the attributes that have been defined so far by IETF [Tho 01]. In fact, some of these attributes could be used in routing among optical ASs whereas others are redundant. Our work will focus on using some of these attributes to develop a routing protocol that will serve the needs of a WDM network. We shall discuss the benefits of only some of these attributes in optical networks in the following two chapters.

Attribute number	Attribute name
1	ORIGIN
2	AS_PATH
3	NEXT_HOP
4	MULTI_EXIT_DISC
5	LOCAL_PREF
6	ATOMIC_AGGREGATE
7	AGREGATOR
8	COMMUNITY
9	ORIGINATOR_ID
10	CLUSTER_LIST
14	MP-REACH-NLRI
15	MP-UNREACH-NLRI
16	EXTENDED_COMMUNITIES

Table 3. 5: BGP attributes

3.3.2 BGP operation "E-BGP and I-BGP"

When BGP is used between two different ASs, this mode of operation is referred to as External BGP (E-BGP) [Tho 01]. If an Internet Service Provider is using BGP to exchange routes that have been learned by an external BGP session or by any other means within its AS, then this mode of operation is referred to as Internal BGP (I-BGP) [Tho 01]. The most important fact about the operation of BGP as I-BGP is that each node has to peer with all other edge nodes located in the same AS through a logical connection [Ste et al 99]. The reason for these logical connections is to allow edge routers that belong to the same AS to exchange the routing information learned via different external sources. This mode of operation is known as "full-mesh I-BGP" [Moy et al 98]. In fact, since I-BGP sessions are logical sessions, there are no direct physical connections among the participants. Each AS configures these logical connections among its edge routers. For example in Figure 3.8, node Y in AS 2000 will have a logical peer with both edge routers Z and W. These logical paths are configured by the network administrator of AS_2000, the network administrator will specify the intermediate nodes that will be used to establish the I-BGP session between router Y and Z and the intermediate nodes that will be used to establish the I-BGP session between router Y and W. Unfortunately, these configured paths might not be configured properly causing some of the data packets to be lost somewhere along the path, or to be sent back and forth (sometimes called "oscillation") without reaching the required destination [GG 02], [BLRW 02]. A study on BGP miss-configurations found that up to 1200 prefixes in the Internet maybe suffering from miss-configuration every day [MWT 02].

3.4 Inter-Domain and Intra-Domain Routing protocols interaction

Today's Internet relies on different types of routing protocols to exchange reachability information, some of these routing protocols are used for Intra-domain routing like OSPF, IS-IS while others are used for Inter-domain routing such as, BGP, EGB [Tho 01]. For compatibility, routers should be able to make use of each route that has been learned from any of these different protocols to update its routing table if it finds that this route is better than the one stored in its routing table [Moy et al 98].

Normally, the processing of received routes should go through three phases at each router [Moy et al 98]. The first phase occurs when a route is received from either an internal or external routing protocol. This router should decide whether to accept this route or not [Moy et al 98].

The second phase is to compare this route with other possible routes that lead to the same destination in order to pick the best route to store in the routing table [Moy et al 98].

Finally, the third phase is to apply some decisions on each stored route to decide whether it could be nominated for advertising to neighbours or otherwise [Moy et al 98].

As a matter of fact, applying these three phases is a very hard task to perform because choosing amongst different routes learned by different routing protocols is a very difficult process [Tho 01]. For example, cost values may mean different things in different protocols and to interpret cost from OSPF to a cost in IS-IS will not reflect the exact value of the cost.

Usually, interpretation between routing protocols is needed when a route has to be found across several of ASs that use different IGP protocols [Moy et al 98]. For simplicity, there either should be a global agreement on how to interpret the attributes of a route learned

from another routing protocol, or assigning one protocol for Intra-domain and one for Inter-domain.

Chapter 4 Inter-Domain lightpath provisioning

4.1 Overview

Much research was done to find which routing protocol could be used for intra-domain routing for Optical networks. The result of this research showed that OSPF can be used as intra-domain routing protocol because it supports hierarchical routing and can be easily extended.

On the other hand, much effort has been directed to define a routing protocol that performs routing amongst different optical networks and many approaches proposed. Some of these approaches tried to perform lightpath provisioning among AS using BGP. Many approaches find that BGP could be extended for performing routing and signaling in the same time among different ASs to establish an end-to-end lightpath.

Since this work is concerned with lightpath provisioning amongst ASs using BGP, we will consider related works for enhancing BGP to support lightpath provisioning amongst ASs.

4.2 Related work

4.2.1 Optical BGP (OBGP): Inter-AS lightpath provisioning

The OBGP Internet draft [IDBLa 01] describes an approach to using BGP for lightpath provisioning among different ASs. Their proposed approach is based on the fact that different ASs are owners of their wavelengths and their OXCs. These ASs give their customers virtual control over their optical resources. In other words, customers may have the ability to set up or tear down a lightpath based on their need.

This approach suggests two phases of operation. The first phase is to exchange information about optical resources such as available wavelengths and reachability information such as As_Path, and cost amongst the ASs. The routing information is passed using multiprotocol BGP extensions [RFC 2858] and the BGP Extended Communities Attribute [RFC 1997].

The exchanged routing information will be used by each AS to construct a Routing Information Base (RIB). This RIB will be used to construct a map of the available wavelengths to each destination.

The second phase is the signalling phase. Each AS makes use of the RIB obtained from the first phase to setup a lightpath. BGP update messages are used for lightpath set-up or tear-down. The BGP update message is propagated across the desired ASs reserving the wavelengths and setting the OXCs along the path.

4.2.2 OBGP in Optical Networks

Another approach [JY 02] suggests two phases of operation exactly like the previous approach in Section 4.2.1. The first phase is to exchange routing information whereas the second phase is the signalling phase. Furthermore, this approach builds on the assumption that each AS has an OXC where it can share its optical resources with neighbor ASs. For example, in Figure 4.1, AS_4000 can make use of its combined unit[3] to setup a lightpath to any other ASs to reduce the pressure on its IP router. Furthermore, AS_4000 gives virtual control to AS_1000 to use its combined unit to setup a lightpath to any destination. This approach assumes a trusted relationship among different ASs (see Figure 4.1).

Moreover, each AS treats its combined unit as two separate routers. The first router is treated as a normal BGP speaker whereas the second router is treated as a virtual BGP router that represents the OXC. The virtual router advertises itself independently of the first router of the combined unit. Each virtual BGP router advertised its resources such as available wavelengths to all edge routers within its AS and to other edge routers located in different ASs. More importantly, each AS treats its combined units as completely separate ASs. In other words, any edge router outside of the AS that has the combined unit, assumes that the AS and the virtual router are two different ASs. For example, AS_1000 assumes that both AS_4000 and its combined unit are two different ASs.

Fortunately, treating each virtual router as a separate AS within each AS will eliminate the I-BGP mode of operation. As we explained earlier in chapter three, I-BGP is used to tunnel the external routing information to exchange reachability information among different ASs. Therefore, if each virtual router is an independent AS then there is no need to

[3] A combined unit has an IP router directly attached to a OXC as shown in Figure 4.1

tunnel any optical information. However, it will be difficult to run this protocol if each AS has more than one combined unit. This is because each combined unit will be treated as an individual AS within the same network that belongs to the same organization. The more combined units in one AS the more difficult it is to advertise or to coordinate between these combined units within the same AS. Furthermore, other ASs have to keep a large list of all combined units located in their domains and other domains.

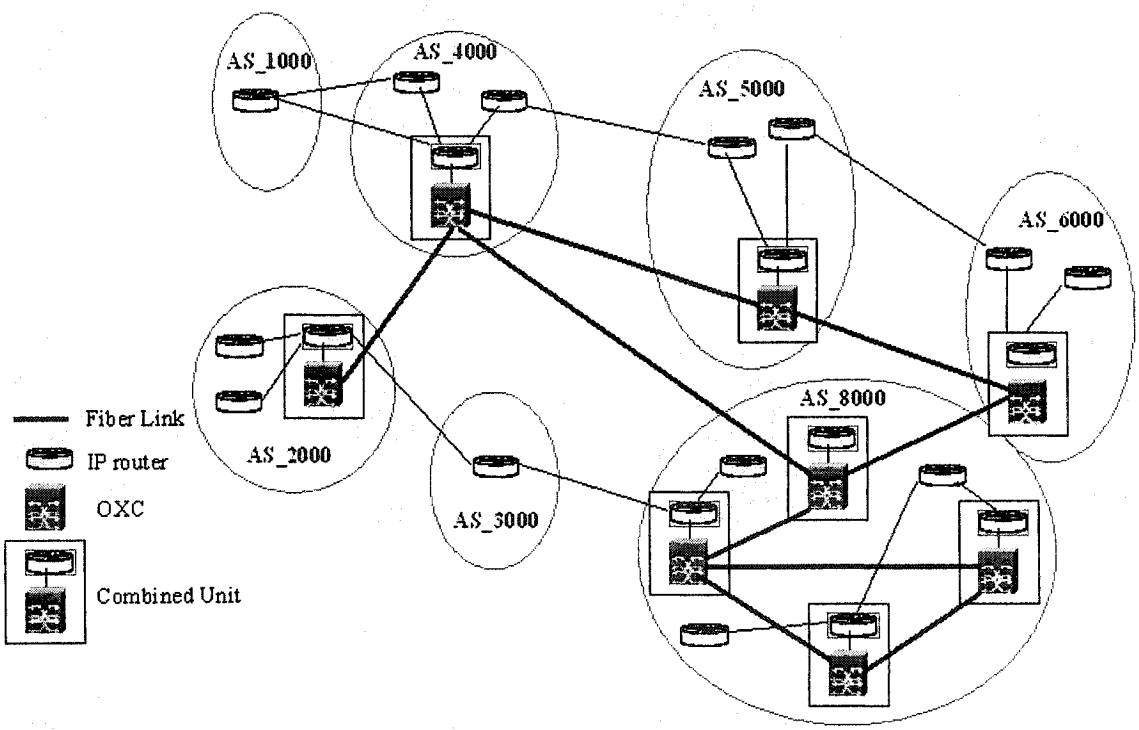


Figure 4. 1: IP load balancing using Combined Units

4.2.3 Inter-domain Signaling/Routing protocol in Optical Networks

Another approach [FSHL 01] has defined a new BGP message. This fifth message is used to set-up the requested wavelength.

This protocol has two modes of operation. The first mode is called the four-phase mode whereas the second is called the two-phase mode. In either mode, the node that initiates the lightpath request has no clear information about the available wavelength along the path. In the four-phase operation mode, the initiating node reserves a sub-set of the wavelengths at the source node hoping that one of these wavelengths will be available at the destination node. Figure 4.3 shows the four phase operation mode. Clearly, the four-phase mode shows that the initiating node discovers the available wavelengths between the source and destination along the path. Once the destination node receives the information about the available wavelengths along the path, it picks any of these available wavelengths and sends the reservation confirmation back to the source. Once the confirmation reaches the source node, the source node starts the third stage which is setting the OXC along the path and finally the destination node will send the confirmation back to indicate that the path is successfully established.

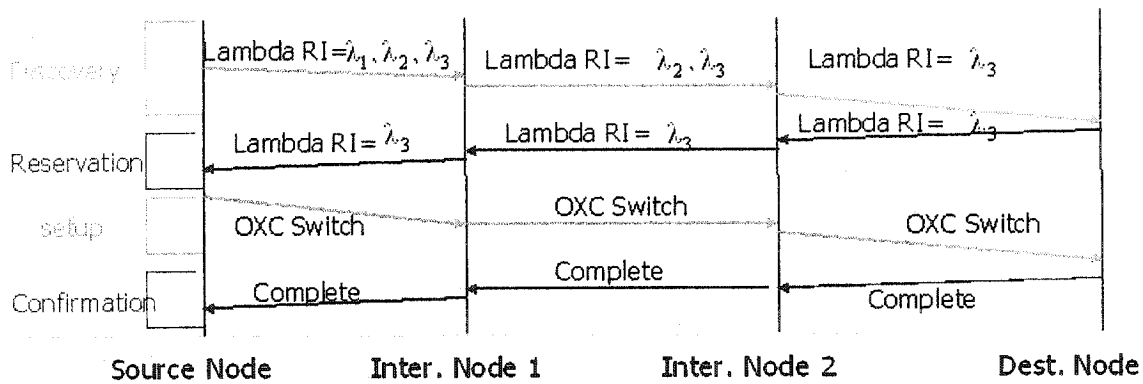


Figure 4. 2: Four phase operation mode

The two-phase mode is a simplified version of the four phase mode. In the two-phase mode, the discovery and reservation are combined in one phase whereas the set-up and confirmation phases are combined in another phase (Figure 4.3). The node that initiates the lightpath setup will pick randomly one available wavelength hoping that this wavelength is available along the path.

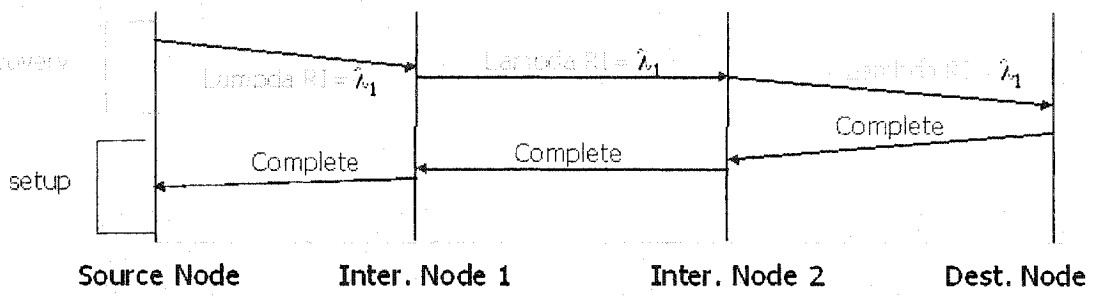


Figure 4. 3: Two phase operation mode

This approach has no clear mechanism for routing among ASs. In fact, it relies on the fifth signaling messages to update the manually pre-configured routing tables whenever a lightpath has been reserved. Unfortunately, these local changes are invisible to other edge routers. This leads to the problem that an other AS may initiate a lightpath using a wavelength that is already reserved.

4.3 Advantages and disadvantages of previous approaches

We investigated previous approaches that perform lightpath provisioning among ASs [IDBla01], [FSHL 01] & [JY 02].

The first approach is proposed by Canarie [IDBla01]. It is based on using BGP for light path provisioning. The proposed mechanism has two phases.

1. The first phase includes exchanging complete routing information about the topology of the ASs. This information is exchanged through the multi-protocol BGP extension and extended community and used by each OXC to create a lightpath Routing Information Base (RIB) to be used to establish lightpaths to other ASs
2. The second phase is the signaling phase. Each OXC makes use of the lightpath Routing Information Base (RIB) to be used to establish lightpaths for other ASs. Each edge node uses BGP update message to establish any requested lightpath.

The problem with this approach is that it assumes a complete trust relationship. Each AS gives complete control to its customers over its entities and other AS entities. Besides, this approach does not specify what type of intra-domain routing information should be exchanged/advertised. Finally, this approach uses BGP update messages as signaling messages to set-up the requested lightpath. This has the following disadvantages: first, since BGP is configured by a network operator, BGP does not provide automatic recovery, and more importantly, oscillation is more likely to occur due to a configuration mistakes (see section 3.3.2). Second, the oscillation problem will be worse if BGP is used to perform signaling.

The second approach also uses the same two phases of operation as the Canarie approach. Despite the similarities, the second approach proposes a new mechanism that would eliminate the I-BGP mode of operation. This can be done by assuming that each OXC is a completely separate unit. This unit is treated as an independent unit from its AS. In fact, eliminating the I-BGP mode of operation is a very good idea, because I-BGP is subjected to routing oscillation. However, the problem is that if each AS has more than one OXC then treating each of these OXC individually will cause a serious scalability problem. In other

words, each AS has to keep track of OXCs located in its own domain and other OXCs located in neighbor ASs.

The last approach uses BGP for signaling [FSHL 01]. They define a new message type to BGP message that works separately. This fifth message make uses of manually configured routing tables, this approach has many weak points. First, the routing tables are manually configured, i.e. the routing information is static. Second, defining a new fifth message in BGP is not a good idea because this will complicate the protocol implementation, In fact, a BGP update message can perform the same task. Finally, as we explained, BGP is not protected from oscillation when it is used within each AS.

Chapter 5 Optical Routing Border Gateway Protocol (ORBGP)

The purpose of this work is to define a routing protocol that can perform wavelength routing among ASs. In the previous chapter, we have explained three different approaches to achieving lightpath provisioning among optical ASs. In this chapter, we will show our approach for achieving inter-domain routing in optical networks. Our approach is to define a new protocol called Optical Routing Border Gateway Protocol (ORBGP). We will explain the features of our proposal and how it overcomes the weak points of the previous approaches such as assuming trusted relationships among different ASs. Besides, these approaches use the routing protocol BGP to perform the signaling task. BGP is already subjected to oscillation as we explained earlier in Section 3.3.2. Therefore, the same problem will arise in using BGP as a signaling Protocol. Therefore, we believe that for inter-domain signaling, ORBGP should not perform any signaling part leaving this task for GMPLS to avoid any routing loops that might took place due to miss-configuration.

5.1 ORBGP overview

Optical Routing Border Gateway Protocol (ORBGP) is built on the experience gained from BGP. In fact, much research is going on to improve the operation and scalability of BGP which suggests a promising future for this routing protocol [MF 02].

ORBGP adopts most of BGP's features with some differences that allows wavelength information to be exchanged among edge routers to reflect the up-to-date bandwidth availability of the optical network. Furthermore, ORBGP introduces a new route advertising scheme which is triggered by the number of changes that took place in the link table of the advertising node.

5.2 ORBGP goals

There are many reasons that led to the definition of a new routing protocol. One reason is to have a routing protocol that serves the ISP's basic needs. Of course, their basic need is to have routing information at their edge routers that guarantees any lightpath setup to any destination they wish to reach. Therefore, the aim of ORBGP is to guarantee a reliable mechanism to exchange optical routing information that reflects the most up-to-date topology of the global network.

Sections 4.2.1, 4.2.2 and 4.2.3 presented different existing approaches for performing lightpath provisioning across multiple optical domains. All these approaches suffer from some of the following weak points:

1. The assumption of a trusted relationship implicit in giving customers virtual control of their entities.
2. The lack of ability to have a complete control of the type of routing information advertised. For example, in Figure 3.8, the network administrator of AS_2000 might decide to advertise the availability of 7 wavelengths out of 10 without specifying the color of the advertised wavelength that can be used to cross AS_2000. The reason for doing so is that AS_2000 does not want to reveal all its internal resources to the

other ASs. Therefore, if this advertised information is used by other ASs to cross the AS_2000 domain, 30% of the lightpath requests coming from neighbors will be blocked. Normally, if a request is dropped by AS_2000, the neighbor who initiates that request will send another request asking for different wavelengths assuming that the first wavelength has been reserved by another customer.

3. The approaches extended BGP to perform both routing and signaling through two separate phases, as explained earlier in Chapter 4. The major problem with using BGP as a signaling protocol is that it is a configured protocol, therefore using this protocol to tunnel a lightpath request within certain AS is subject to oscillation due to the miss-configuration that may be introduced by the network administrator [GG 02] (see Section 3.3.2). In order to avoid the last weak point which is related to I-BGP, we assigned the I-BGP tasks to GMPLS (Figure 5.1). We will talk about GMPLS and how it can avoid oscillation later in this chapter. For the same reason, we believe that I-BGP should not perform any routing task within each AS. Therefore we assign the routing task within each AS to OSPF. In fact, using OSPF can be more scalable if it also performs the I-BGP task [Moy et al 98]

Our proposed routing protocol is capable of advertising a specific amount of routing information and to give controlled privacy for the advertising AS.

Furthermore, our routing protocol should avoid the advertisement policy defined in BGP. BGP does not allow edge nodes to advertise changes that took place in its routing table based on their needs. In fact, BGP performs automatic advertisement whenever a change takes place in its routing table. Of course, this makes no sense in the optical domain; an edge node should not advertise each time a wavelength has been released or reserved.

Figure 5.1 illustrates our point of view in terms of protocol interactions. As shown in the figure, ORBGP will perform only routing, whereas the signaling among the ASs will be performed by other means.

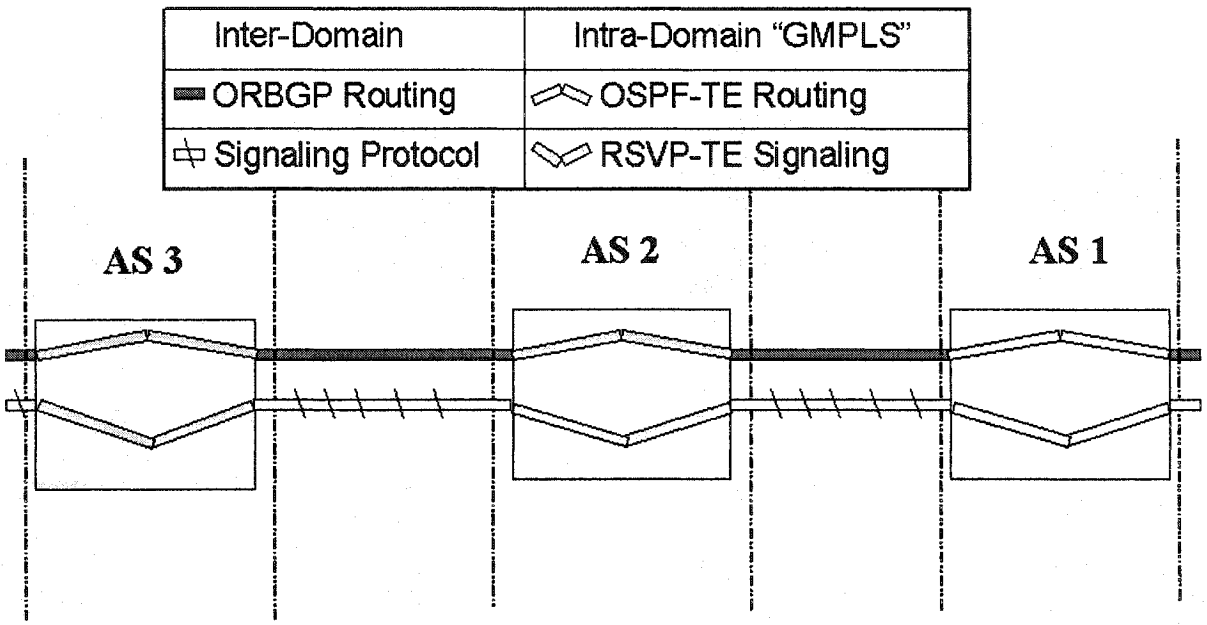


Figure 5. 1: Proposed protocol interactions

5.3 Routing information exchange: an example

Each Optical Cross Connector (OXC) should have a routing table that includes all the wavelengths to all destinations in the global network. Figure 5.2 shows a network example that will help us to illustrate the challenges that faces optical lightpath provisioning.

For example, router X in AS_3 should be able to reach router F in AS_1 on a certain wavelength; this information is a result of an interaction mechanism between intra-domain and inter-domain routing protocols.

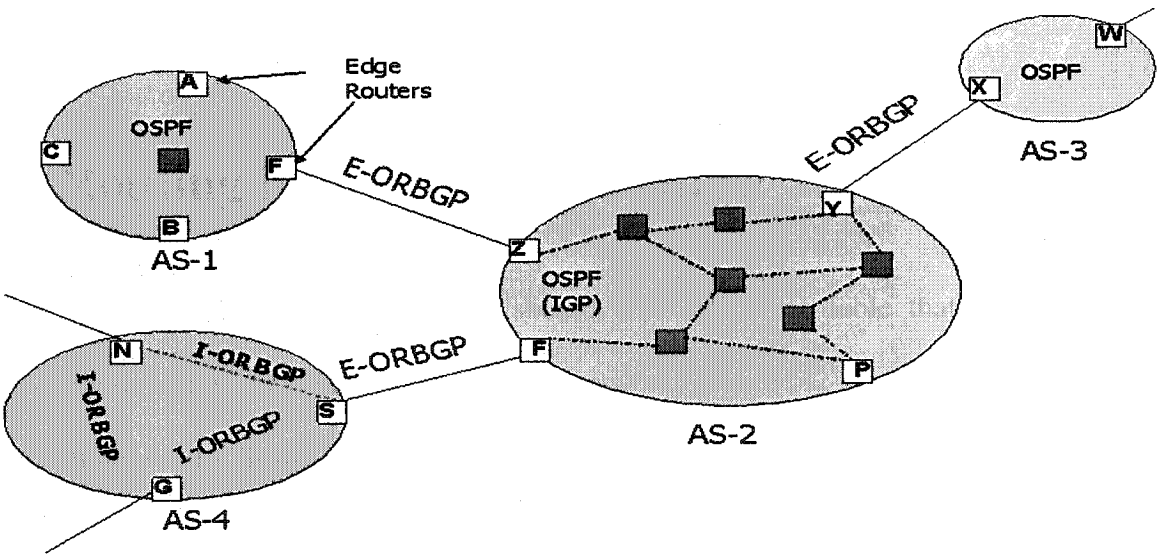


Figure 5. 2: Network example

Figure 5.3 shows a possible definition of an interaction mechanism. Clearly, routers Z and Y belong to the same AS. Furthermore, routers Z and Y have an extra space in their routing table to store information related to both intra-domain and inter-domain routing.

Normally, Z and Y should have the same information about the topology of their AS (AS_2) by the mean of OSPF. This means that both routers have enough information about the available wavelengths between them.

Moreover, since routers Z and Y are edge routers, they should also have the information about neighboring edge routers. Therefore, router Z knows which wavelengths are available to reach router F, located in (AS_1), and router Y knows which wavelengths are available to reach router X, located in (AS_3)

The path calculation algorithm across different domains is performed as follow:

- First, router Y will store the set of available wavelengths $\Lambda_{Y,X} = \{\lambda_1 \lambda_2 \lambda_3\}$, the set of valid wavelengths to reach AS_3.
- Second, router Y will intersect the available wavelengths that can be used to reach router X with the available wavelengths that can be used to reach router Z. The purpose of this intersection is to find the available wavelengths that allow router Z to reach router X.

  $\Lambda_{Y,Z}\{\lambda_1 \lambda_2\} \bigcap \Lambda_{Y,X}\{\lambda_1 \lambda_2 \lambda_3\} = \Lambda_{Z,X}\{\lambda_1 \lambda_2\}$.
- Third, router Y will pass the result of this intersection $\{\lambda_1 \lambda_2\}$ to router Z either by means of I-BGP as an update message or by the mean of OSPF as a LSA.
- Fourth, router Z will store the information obtained from router Y as a valid wavelengths to reach router X in AS_3.
- Fifth, router Z will intersect the information obtained from router Y with the available wavelengths that connected router Z to F. The purpose of this intersection is to find the available wavelengths that allow router F to reach router X in AS_3

  $\Lambda_{Z,F}\{\lambda_1 \lambda_3 \lambda_4\} \bigcap \Lambda_Y\{\lambda_1 \lambda_2\} = \Lambda_{F,X}\{\lambda_1\}$
- Sixth, the result of this intersection in router Z will be advertised to router F located in AS_1 by means of BGP (E-BGP).

- Finally, router F will store the information obtained from router Z as valid information to reach AS_3.

As a result, router F has enough information about how to reach AS_3. More importantly, router F does not know any details about AS_2 internal topology. The only piece of information that router F knows is that AS_3 is located after AS_2 and can be reached on λ1. The reason for hiding the internal topology is to prevent other ASs from using the available resources. Resources should be advertised only based on a business agreement otherwise it will be used for the benefit of other ASs.

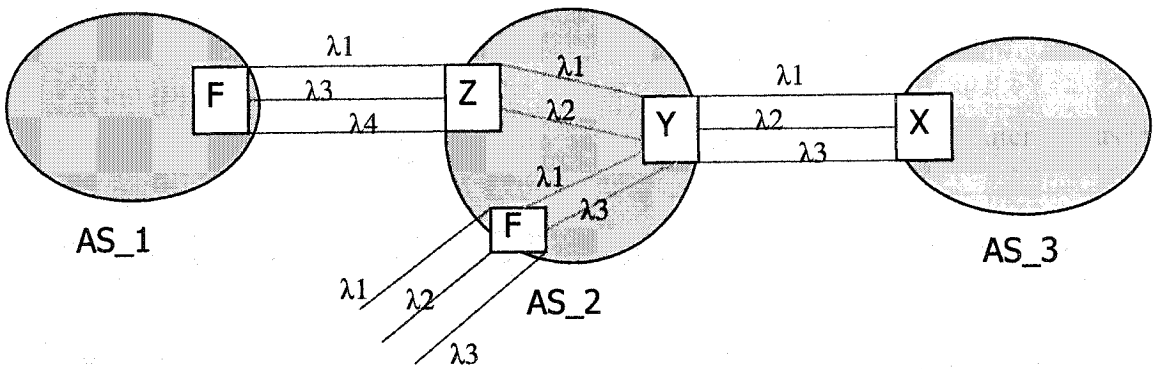


Figure 5. 3: Intra-Domain & Inter-Domain interaction mechanism

5.4 ORBGP structure

5.4.1 ORBGP messages

As we explained earlier, ORBGP is built on the experience gained from BGP. In Chapter 3, we explained BGP in detail. Besides, we explained the message types and the purpose of these messages that are being used in BGP [RFC1771]. In our approach, we are using most of the BGP features [Moy et al 98].

The ORBGP session starts when the advertising node sends an open message to identify itself. Once the BGP session has been established, the two edge nodes exchange the

optical routing information using the update message [RFC1771], [Moy et al 98], [Ste et al 99]. Both nodes keep the ORBGP session alive by sending the keep alive message. Finally, both nodes report errors by simply sending a notification message. In this section, we will only discuss the new information that has been added or replaced within the BGP update message to carry optical routing information.

Figure 5.4 shows the structure of an optical update message that is being exchanged. As explained earlier in Chapter 3, each update message advertises a route using a certain number of attributes. These attributes are:

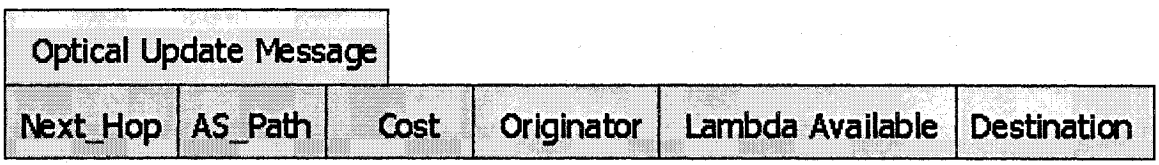


Figure 5. 4: Optical Update Message structure

- Next_Hop attribute field: this field indicates the port number/address of the source edge router that leads to the next hop that will eventually lead to the desired destination. The port number/address represents a unique number of the physical link if there is an external ORBGP session with the neighbor AS, whereas the next hop could be something like the *virtual interface* in BGP if there is an internal session to exchange routing information among edge routers that belong to the same AS.
- As_path attribute field: this field includes the list of the AS names that will lead to the desired destination.
- Cost attribute field: this is a new attribute that specifies the cost of the link being used.

- Originator attribute field: this field stores the name of the edge router originating this message. This value will be used by the receiving edge router to prevent sending the same routing information back to the sender.
- Available wavelengths: shows the status of the available wavelengths.
- Destination field: stores the destination prefix

5.4.2 Routing table structure

At each node there should be a routing table that shows the available wavelengths to all destinations. To simplify things, we considered the following network example shown in Figure 5.5. This small network has six ASs that represents a mesh network with a diameter equal to 2. Each AS has one OXC and each OXC is connected to the other OXC through a fiber link. We write $C_{i,j}$ for the cost of the link between the two OXCs i and j, as shown in the figure.

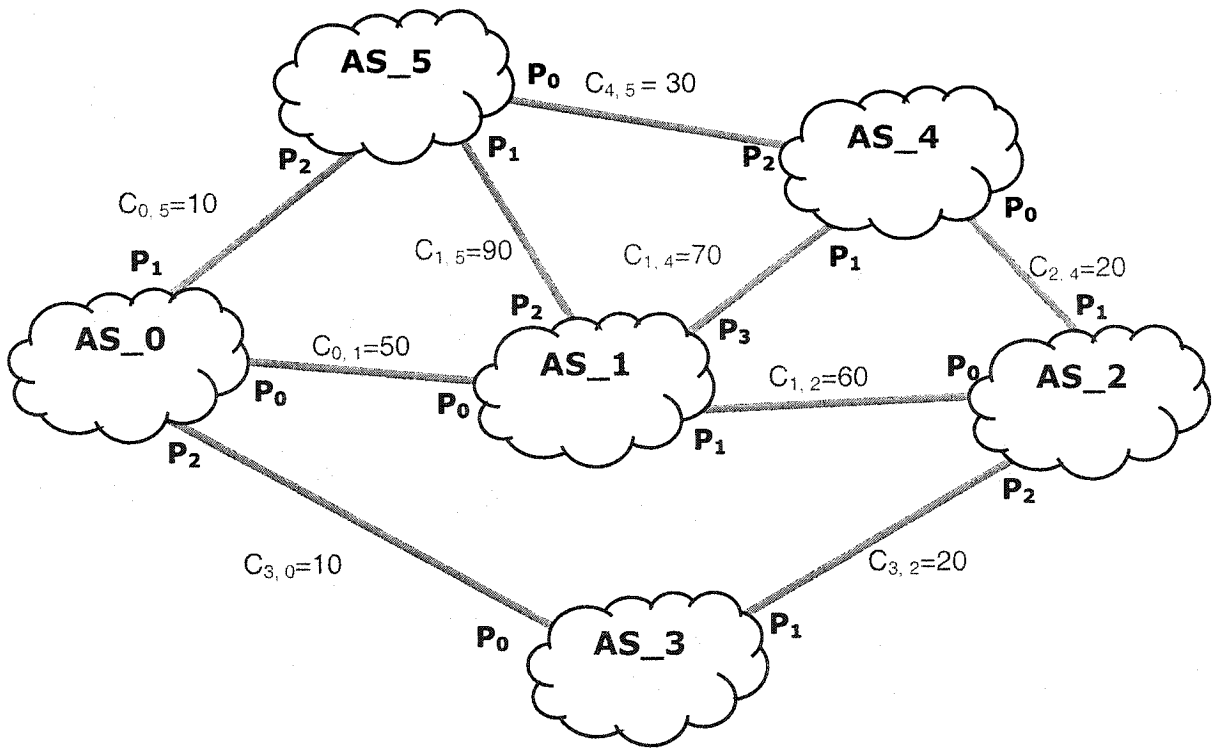


Figure 5. 5: Simple network example

Dest AS	NextHop	AS_Path	Cost	Current Wavelength										Originator
0	*****	0	0	T	T	T	T	T	T	T	T	T	T	0
1	Port_0	1	50	T	T	T	T	F	T	T	T	T	T	1
2	Port_2	3 2	30	T	T	T	T	T	T	T	T	T	T	3
3	Port_2	3	10	T	T	T	T	T	T	T	T	T	T	3
4	Port_1	5 4	40	T	T	F	T	F	T	T	T	T	T	5
5	Port_1	5	10	T	T	T	T	F	T	T	T	T	T	5

Table 5. 1: Routing Table for Node 0

Table 5-1 shows the routing table for AS_0. It includes all the necessary routing information that AS_0 should have to be able to establish a light path successfully to other ASs.

- The first column shows all the destinations that AS_0 can reach.
- The second column gives the port number/address of the link that should be used to reach the correspond destination in column number one. The third column shows the AS_Path attribute, the AS_path includes the list of the AS names the routing information traversed.
- The fourth column is the cost of the path; the cost could be a function of the fiber length. For example, the longer the fiber cable the higher the cost will be. Figure 5.5 shows that AS_0 can reach AS_4 through AS_5. Therefore, the cost between AS_0 and AS_4 is the summation of the cost along the path. In our example, the cost will be $C_{0,5} + C_{5,4} = 10 + 30 = 40$.
- The fifth column represents the available wavelengths to the corresponding destination in column 1.
- The sixth column represent a flag that indicates from were this route has been originated, for example AS_1 will set the originator attribute equal to 1 when AS_1

decides to advertise a route to AS_0. AS_0 will accept the advertised route if AS_0 finds that it is a useful route. Now, when AS_0 decide to advertise this route to other ASs, the originator attribute will prevent AS_0 from sending this route back to its origin, which is AS_1.

5.4.3 Link table

Each link at each edge node has a table that keeps track of the available wavelengths with the neighbor edge router located in the neighbor AS. Table 5.2 shows the configuration of the link that connects AS_0 to node AS_1.

Dest AS	NextHop	Free Wavelengths									
1	Port_0	T	T	T	T	F	T	T	T	T	T

Table 5. 2: Link Configuration between node 0 and node 1

This table will keep track of the available wavelengths on this link, in other words, this table will be updated when a wavelength request is received by the edge node. So, the status of the requested wavelength will be changed from True to False if the wavelength status was True. Whereas the status is changed from False to True if the received message is a tear down message.

5.4.4 Routing policies and Information processing among ASs

Routing policy could be defined as a set of conditions specified by the network administrator. For example, a network administrator may decide to accept a path that has the smallest number of hops as a desired path, whereas others may decide to choose the path that

has the lowest cost. The routing policy could be different from one network administrator to another; it depends on business agreements between the ASs [Moy et al 98].

Normally, when a node receives a message, it has to check the type of the message. If this message is a routing message, the node makes use of the configured policy to make decisions on whether to accept the advertised path or not. The following example illustrates an algorithm that prefers a path that has lowest number of hops over the cost. The example starts when the node receives the routing message; the node performs the following:

- It checks the AS_Path attribute. If this node finds that its AS name is in the received AS_Path attribute, it will reject this message [RFC 1771], otherwise
- It will check its routing table for other routes to the same destination with a smaller number of hops. If there is one, then this new route is rejected, otherwise
- If there are two different routes that have the same number of hops, this node will check for the cost. If the cost of the received message is higher than the one stored locally, then this route is rejected, otherwise
- The received route overwrites the old route because it has the most recent status of the available wavelengths.

5.5 Advertising Policies

At this point, we are familiar with the needs of optical routing protocol, we discussed earlier in this chapter the structure of our proposed protocol ORBGP. However, we did not explain anything about when are the routing messages exchanged, or what are the conditions that force an edge node to advertise?

In fact, we consider the following two schemes for advertisement. *The first scheme* is based on the expiration of a time-out/refreshing period[4] whereas *the second scheme* is based on number of changes[5] that took place in a given link table.

5.5.1 First scheme for advertisement

Any node that uses the first scheme will wait until the expiration of the timer that triggers the advertisement. Once the timer expired, that node will check the number of changes for each link table, if the number of changes in any link table is larger than a specified threshold, which is set by the network administrator, then an advertisement will be sent to all neighbours in adjacent domains, and the timer is set again.

As a matter of fact, triggering the refreshment continuously based on fixed period of time does not make sense because if the link table might experience a relatively high number of changes and the timer of the refreshing period is not expired yet. Therefore, these changes would not be advertised and hence introducing errors at other routing tables.

5.5.2 Second scheme for advertisement

The second scheme is to advertise independent of time. This can be done by introducing a counter at each link table, which counts the number of status changes individual wavelengths (being reserved or being released). When the value of the counter becomes higher than a threshold set by the network administrator, then an advertisement will be done.

[4] Time-out/refreshing period: it means how long the edge node will wait to advertise or report changes that have taken place in its link table.
[5] number of changes: This represents the number of reserved and released wavelengths in each link table.

This second scheme might not be suitable for low request rates because some nodes might not need to advertise since there were very few changes. Therefore, it would be better to use the two schemes together.

In fact, no matter which scheme is used, if a node decides to advertise because its timer has expired or a certain link table experienced relatively high number of changes, hence that node will perform the following advertisement algorithm:

- For each row of column 5 in Table 5.1, the node intersects the available wavelength with the available wavelengths in the link table.
- The advertising node inserts the result of wavelengths intersection along with the corresponding route in an update message.
- the node modifies the originator, the cost and the As_path attributes and finally,
- The node sends the update message that includes the changes to its neighbors.

5.6 Inter-Domain and Intra-Domain signaling

5.6.1 Intra-Domain Signaling

Although signaling is not the aim of this work, nevertheless we still believe that I-BGP should not perform any signaling task within each AS. Oscillation will be very serious and a difficult problem to be solved within the AS. Therefore, we assigned this task to GMPLS [BDLT 01]. GMPLS is concerned with the control plane that performs complete management for all types of connections for both packet switching and non-packet switching, such as Time Division Multiplexing (TDM), wavelength switching, and finally fiber switching [RFC 3471]. GMPLS uses either the Constraint-Based Routing Label Distribution Protocol (CR-LDP) or the Resource Reservation Protocol (Traffic Engineering)

(RSVP-TE) for signaling purposes [RFC 3472], [RFC 3473]. In fact, there is no clear winner between these two signaling protocols; however, the trend is towards RSVE-TE, since this protocol has been deployed for a long time [BF 00]. RSVP-TE has a very efficient way to detect loops and hence avoids oscillations. This works as follows. Each node that originates a new connection keeps track of all the intermediate nodes by defining an object known as RECORD_ROUTE object [RFC3209]. This object is carried in the connect request packet and keeps a record of all intermediate nodes encountered [RFC3209].

5.6.2 Inter-Domain Signaling

Up to this point, all we were talking about was the exchange of routing information among different ASs; we defined the routing message types, mechanisms and specification of the proposed routing protocol.

Despite the fact that this work is about performing routing among different ASs, nevertheless we decided to have a brief talk about a simple signaling protocol that will help us in evaluating our routing protocol in the context of our simulation studies. We used three types of messages for simulation purposes:

1. Request message
2. Confirm message
3. Tear down message

These three messages are used by each edge router to establish and tear down lightpaths.

The common structure of the signaling message is in Figure 5.6.

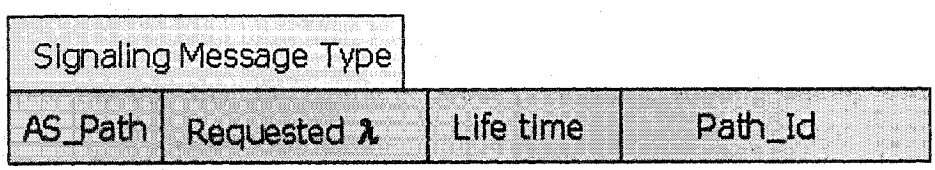


Figure 5. 6: Signaling message structure

AS_Path: This attribute contains the list of the ASs that the signaling message has to traverse to reach the destination.

Requested λ: This field specifies the requested wavelength that will be used to establish the lightpath between the source and the destination.

Life Time: this specifies the duration of the requested lightpath.

Path_Id: this is a unique number to identify the lightpath.

5.6.3 Processing signaling messages among ASs

Normally, the Request Message will be triggered by a costumer request and will be sent after consulting the routing table for available resources. If there are resources, then the initiating node will change the status of the requested wavelength in its link table and send a Request Message to the neighbor node that leads to the destination after increasing the counter that keeps track of the number of changes that occurred over this link. In turn, the neighbor node will check the destination field of the Request Message to find out if this request is to one of its customers. If it does not recognize the destination field, the neighbor node will check the available resources for the requested wavelength. If the requested wavelength is available, it will change the status of this wavelength and increase the counter at the link table and finally forward the Request Message to the next neighbor node that will lead eventually to the destination node. Once the destination node receives the Request Message that has its address, it will send back a confirmation message back to the source node to confirm the reservation.

Once the light path has been successfully established, the destination node will set a timer that will last for the duration of the established lightpath. Normally, once the duration of the requested lightpath is expired, the destination node will send back a tear-down message to the source node.

Unfortunately, sometimes the requested wavelength is not available along the path. Therefore, if an intermediate node finds that the requested wavelength is no longer available it will send back a tear down message to the source node.

And finally, if any node receives a tear-down message, the node will change the status of the wavelength in the corresponding link table and increment the counter that keeps track of the number of changes that occurred on the link.

5.7 Performance of the proposed routing protocol

5.7.1 Information accuracy versus routing overhead

The ultimate goal of the routing protocol is to keep the routing information as fresh as possible at each node. Therefore, each node should refresh its routing table as often as possible. Of course, the need for refreshing is directly related to the rate of requests, in other words, the higher the rate of request is the faster each node has to advertise these changes.

One way of doing the advertisement is to advertise after a fix period of time. One could assume that the shorter the time period the better, but this is not always true because of the following two reasons: First, the request rate could be very low and the need of frequent advertisement is not necessary. Second, when the refreshing period becomes very small, the

large number of routing messages will flood the network and make it difficult for the nodes to keep track of them.

Another solution for reducing the protocol overhead, while keeping the routing information relatively accurate, is to advertise based on need. For example, each node would advertise when a certain percent of changes have occurred in its link tables.

5.7.2 Blocking types

Unfortunately, it may happen that the wavelength reservation could not be achieved when a connection request is sent because the chosen wavelength is not available on some link of the chosen path. This is what we called blocking. For the purpose of this discussion, we assume that the information about all the network resources is correctly stored in a global table, i.e. if any wavelength has been released or reserved at a certain node the global table is updated simultaneously with that node.

In fact, there are three type of blocking that we are interested in:

- Justified Refusal. This blocking occurs when, both the global table and the routing table of the initiating node show that the requested wavelength for the desired destination is not available.

  P_{JR}= Number of Justified Refusal/ Total number of requests

- Unjustified Acceptance. In this type of blocking, the global table shows that the requested wavelength is not available whereas the routing table of the edge node shows that the requested wavelength is available. The reason for this difference is that each node may wait for a certain period of time to report about new reserved resources to other nodes. Consequently, if the corresponding wavelengths at other nodes is being used, it will lead to blocking somewhere along the path and hence causing Unjustified Acceptance blocking.

  P_{UA}= Number of Unjustified Acceptance / Total number of requests

- Unjustified Refusal. In this type of blocking, the global table shows that the requested wavelength is available whereas the routing table of the edge node shows that the requested wavelength is not available. Again, the reason for this difference is that

each node may wait for a certain period of time to report about new released resources to other nodes. Hence, other nodes assume that these resources are still reserved and hence never use them if a customer asks for them causing Unjustified Refusal blocking to occur.

P_{UR}= Number of Unjustified Refusal / Total number of requests

Clearly, the total blocking probability is equal to the sum of three types of blocking.

$P_T = P_{JR} + P_{UA} + P_{UR}$.

On the other hand, we did not mention anything about the probability of a lightpath being successfully established. We call the probability of a lightpath being successfully established as Justified Acceptance (P_{JA}). In this case, both the global table and the routing table of the edge node show that the requested wavelength is available for the desired destination.

The ideal case is to have both P_{UA} and P_{UR} equal to zero, this means that the distributed routing information is 100% correct and reflects accurately the topology and resource allocation in the global network. The above three mentioned blocking types are our major concern because they are directly related to the performance of our proposed routing protocol.

5.7.3 Objectives for our simulation studies

Clearly, the aim of the simulations is to investigate the different blocking probabilities. Besides, we are interested in finding the best advertising policy. Is it better to have regular refresh periods or refreshing based on the number of changes or both of them as we explained earlier in Section 5.5.

The results of our work should show how we can increase the resource utilization of the network and reduce the blocking probability without having too high overhead due to the routing protocol.

And finally, we will show the effect of the network architecture on the three types of blocking probabilities.

Chapter 6 Simulation Results, Performance Analysis and Evaluation

6.1 Simulation tools

We tried many approaches to implement our routing protocol, we decided that this protocol could be implemented using OPNET. Unfortunately, we faced many technical problems using the OPNET simulation tool because the available simulation package for simulating the BGP protocol has many detailed features that we are not interested in. Besides, this BGP implementation has many details that are very hard to follow and hard to extend. Therefore, we decided to use Java as new implementation tool. In fact, JAVA gave us the flexibility that we needed.

In section 6.1, we will discuss in details the types of difficulties we faced with OPNET, and we will show the advantages of our simulation model written in JAVA.

6.1.1 OPNET Simulation tool

OPNET stands for OPtimum NETwork performance simulation tool [OPNET]. This simulation tool is a powerful tool in terms of protocol implementation; most of the Internet protocols have been implemented within the OPNET framework based on the IETF specifications. The implementations of these protocols were done in a way to allow other

users to study and investigate the operation of these protocols by changing the network configuration and certain other parameters.

Furthermore, OPNET has some scenarios that are already implemented to evaluate the protocol performance in a very realistic environment. These performance parameters include propagation delay, jitter, queuing and delay. Figure 6.1 shows a network scenario that is implemented by OPNET to evaluate the BGP performance.

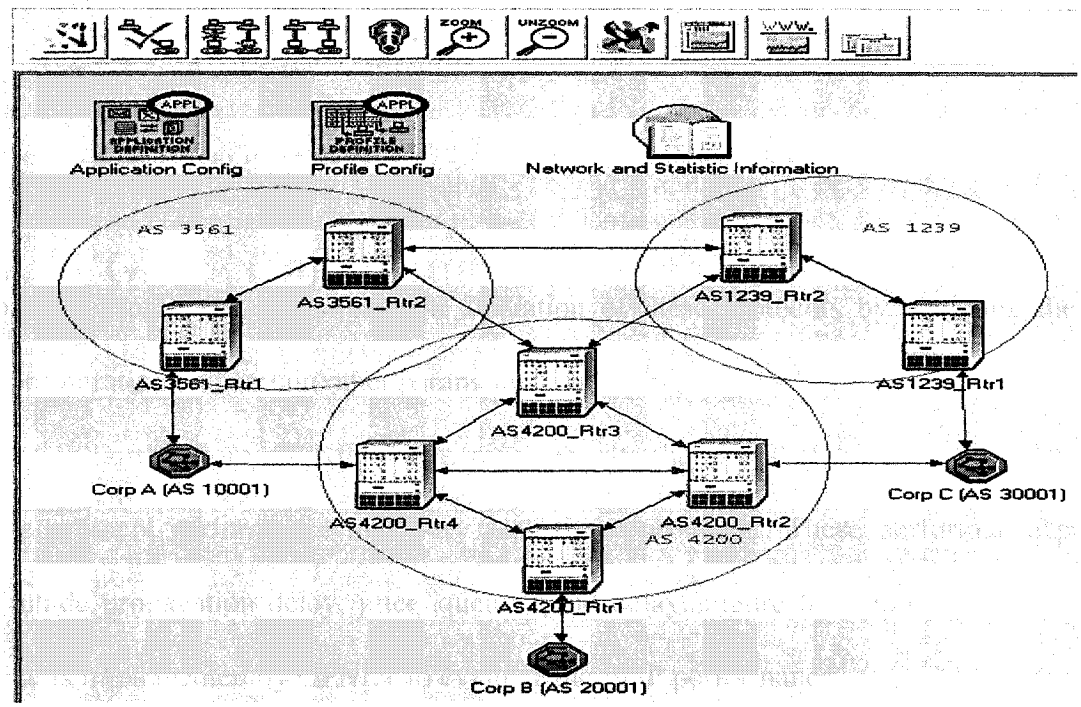


Figure 6. 1: BGP simple configuration scenario

Clearly, Figure 6.1 shows six ASs connected to each other. AS 10001, AS 20001 and AS 30001 are subnets, each has only one edge router. AS1239, AS3561 and AS 4200 are normal ASs, each of them has more than one edge router. Running this model will allow network reachability information to be exchanged by means of BGP update messages to construct a map of AS connectivity at each edge router.

This model was chosen because it has implemented almost all the detailed features of BGP. The idea here is to consider this scenario as starting point and to work on enhancing this scenario to support optical routing.

After careful study of the OPNET BGP model to find a starting point for our work, the idea of enhancing this model could be summarized as follows:

Each router should have optical configuration tables for all its interfaces. For example, router AS1239_Rtr2 should have two configuration tables: the first table has the information about the external connections, while the other table contains information about the internal connections. Table 6.1 shows the structure of the configuration table for external connections for the AS1239_Rtr2 router. The external interface data represent the optical links with AS 3561 and AS 4200

	My AS Number	Neighbor IP Address	Neighbor AS	Fiber Number	λ 1	λ 2	λ 3	λ 4	λ 5	λ 6	λ 7	λ 8	λ 9	λ 10	total number of λ
external Interface	1239	132.0.14.1	3561	1	1	1	1	1	1	1	1	1	1	1	10
external Interface	1239	129.0.30.1	4200	1	1	1	1	1	1	1	1	1	1	1	10

Table 6.1: External Configuration data for AS1239_Rtr2 router

On the other hand, Table 6.2 shows the structure of the configuration table for the internal connections for AS1239_Rtr2. The internal interface represents the available resources among the edge routers that belong to the same AS. Internal interface information is obtained locally by means of the OSPF protocol running within the domain.

Each optical configuration table is read when the corresponding router has to start a BGP session with its neighbor router. For example, router AS1239_Rtr2 will read Table 6.1 if AS1239_Rtr2 has to start an external BGP session with router AS3561_Rtr2 located in AS 3561 or router AS4200_Rtr3 located in AS 4200.

	My AS Number	Neighbor AS	Fiber Number	λ 1	λ 2	λ 3	λ 4	λ 5	λ 6	λ 7	λ 8	λ 9	λ 10	total number of λ
Internal Interface	1239	1239	1	1	1	1	1	1	1	1	1	1	1	10

Table 6.2: Structure of OSPF file for AS 1239.

Table 6.2 is read if and only if any of the edge routers located in AS1239 has to perform an internal BGP session. For example, Router AS1239_Rtr2 will read this table if it has to perform an internal BGP session with Router AS1239_Rtr1.

Unfortunately, we encountered many difficulties when we further pursued our research using the OPNET tool due to the following reasons:

1. It is very hard to implement the ORBGP protocol using OPNET. The problem is not only the implementation; it is how to evaluate this routing protocol. ORBGP is a routing protocol, in other words, it is only concerned with exchanging routing parameters and it does not give any evidence that this routing information is true or not. As a matter of fact, this requires a signaling protocol that has to perform the reservation and this requires enormous work in OPNET.
2. Many features have to be designed to meet optical routing needs. The existing BGP model implementation restricted us from doing so. For example, BGP triggers an advertisement once a single change takes place in its routing table. In fact, this is not acceptable in optical networks because a node should not advertise every time a wavelength has been reserved or released.
3. The BGP code is hard to follow because there is no clear documentation to show the operation of the BGP model in OPNET. Besides, changing any parameter in the configuration tables provided for the model can lead to a huge change in the protocol behavior which is hard to follow and/or to predict.

4. Finally, we are not interested in certain features of BGP, such as keeping the I-BGP operation or certain attributes, such as the Atomic_Aggregator, Aggregator and Origin attributes that we discussed earlier in section 3.3.1.2.

6.1.2 A JAVA Simulation model

To avoid the above difficulties, we decided to implement a new simple simulation platform in JAVA for investigating routing in optical networks. In fact, this platform was build based on the experience gained with OPNET. The new implementation in JAVA allowed us to be more flexible in our design, and we can investigate several features that were hard to obtain in OPNET. For example, OPNET's BGP model advertises whenever a single change has took place in the network, whereas the new protocol advertise when a certain number of changes took place in a given link table.

In our simulation model, we defined an AS object and a link object. Each AS object represents an AS that is run by a single administrator. Each node object includes the routing information stored in a table (see Section 5.4.2). On the other hand, the link object is used to connect these nodes together. The link object has the information about the status of its wavelengths (see Section 5.4.3).

Furthermore, we defined a global routing table; this table is used to keep track of the reservation status of all resources in the simulated network model. This table has accurate information that can be used to verify the accuracy of the distributed routing information located in each node.

The global table and the information in the link nodes are updated simultaneously each time a new connection is established or teared down. On the other hand, the routing

tables of the node are only updated based on either one of the two advertising scheme explained in Section 5.5.

To evaluate the operation of the proposed protocol, we implemented several network architectures, namely the simple architecture of Figure 5.5, the ARPANet architecture of Figure 6.4.1, the ring architecture of Figure 6.5.1, and finally a European network shown in Figure 6.6.1. Finally, to ensure that our simulation results are accurate, we run each simulation for at least 40 times at each given rate of requests and calculated the average. We verified that running the simulation for 40 times is good enough to have all the values located within a confidence interval of 5% around the mean. Every time the simulation moves to a higher rate, we excluded the first run out of the 40 in order to avoid the effect of the initialization bias.

6.2 Simulation parameters

In our simulation, the requests are coming randomly and uniformly distributed over the nodes of the network. The requests arrive at the source node following Poisson distribution, and include the number of the requested wavelength, the duration of the requested connection and finally, the destination node. The selection of the destination node is also uniformly distributed over the total number of the network nodes.

The arrival of the rate of requests follows a Poisson distribution. We chose a Poisson distribution because this distribution reflects a realistic environment for any network environment. Poisson distribution has three conditions to be met [BCB et al 00]:

1. Requests come one at a time.

2. Between two consecutive events, the inter-arriving time is exponentially distributed , and
3. The coming request is completely independent from the previous requests.

There are many parameters for which we should investigate their effect on the performance of the routing protocol. These parameters are:

- The refreshing period. This period will decide how often we should send routing advertisements to guarantee up-to-date routing information at each edge node (see Section 5.5.1).
- Number of changes. This parameter represents the number wavelengths status changes will trigger a routing advertisement. (see Section 5.5.2).
- The number of wavelength per link.
- The request rate for new lightpaths.
- The average duration of the requested lightpaths
- And finally, the internal blocking probability. This parameter represents the internal blocking probability for each AS (see Section 5.2).

6.3 Simulation results for a simple network

All the simulations in this section are performed on the network structure shown in Figure 5.5. The duration of the requested lightpaths is uniformly distributed between 0 and 800msec. And finally, the number of wavelengths per fiber is 10 and the lightpath request rates takes values between 0 and 1 request/msec, it is assumed that the refreshing period is relatively small compared to the lifetime of the requested wavelength.

6.3.1 Effect of the Refreshing Period

This section investigates the effect of the refreshing period on the three blocking probabilities P_{JR}, P_{UA} and P_{UR}.

Figure 6.3.1.1 compares P_{JR} at different refreshing periods. The lower curve represents the value of P_{JR} obtained when the refreshing period equal to 250msec while the upper curve represents the ideal curve. The ideal curve represents the value of P_{JR} obtained when the global table is being used. As we explained earlier in Section 5.7.2, the global table has accurate knowledge about the global network resources. Hence, using the global table will eliminate both P_{UA}, P_{UR}. As a result, more lightpaths will be established and hence better network utilization.

On the other hand, when we do not use the global table, each node relies on the refreshing period to update its routing table. This will lead to non-zero values for P_{UA} and P_{UR} because waiting for the expiration of the refreshing period will delay the report of changes that took place at the node and hence the corresponding at other nodes becomes erroneous. The presence of P_{UA}, P_{UR} blocking will prevent a certain number of lightpaths to be established leaving more lightpaths available for costumers and hence reducing P_{JR}. This explains the difference between the ideal curve and the other curves at refreshing period 20, 40, 80, 130, 200 and 250msec.

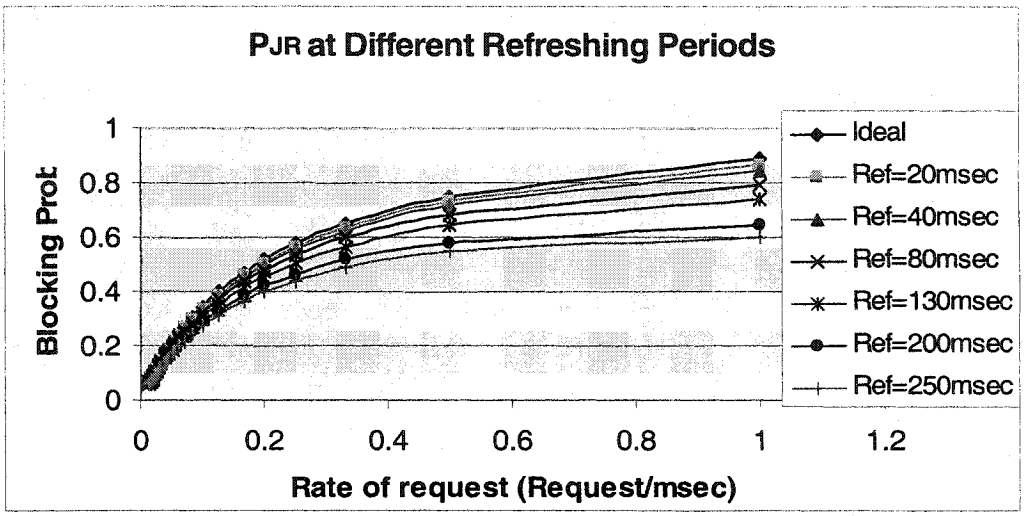


Figure 6.3.1.1: Probability of Justified Refusal at different refreshing Periods

Figure 6.3.1.2 shows the effect of the refreshing period on P_{UA}. It shows a clear difference between blocking for the ideal case and blocking for refreshing period = 250msec. As a matter of fact, P_{UA} can be reduced by decreasing the refreshing period because fast refreshing will keep the routing table relatively up-to-date and hence P_{UA} will be smaller.

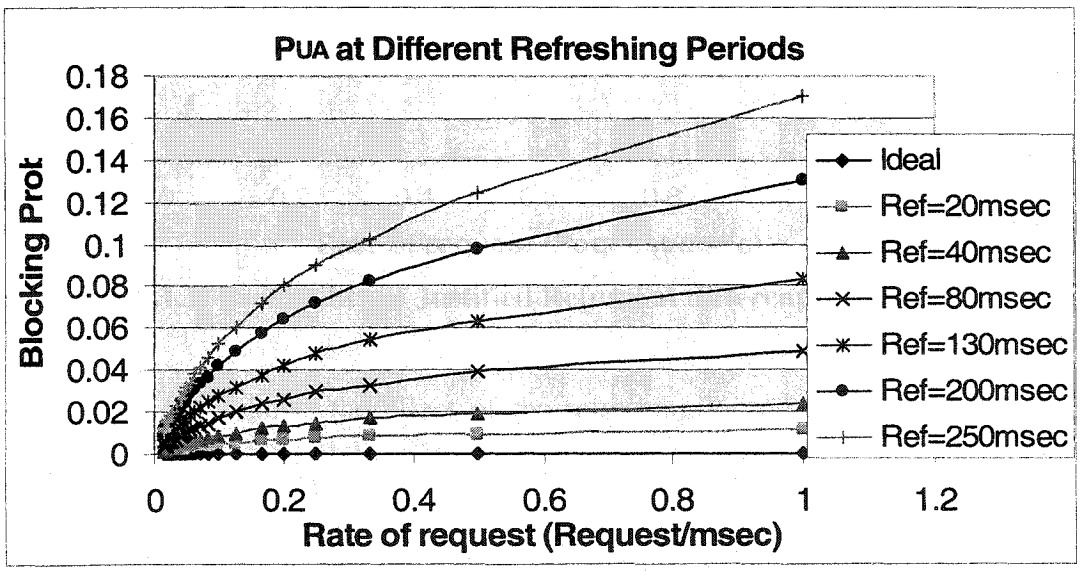


Figure 6.3.1.2: Probability of Unjustified Acceptance at different refreshing periods

Figure 6.3.1.3 shows the effect of the refreshing period on P_{UR}; it is clearly shown that the smaller the refreshing period the better the situation.

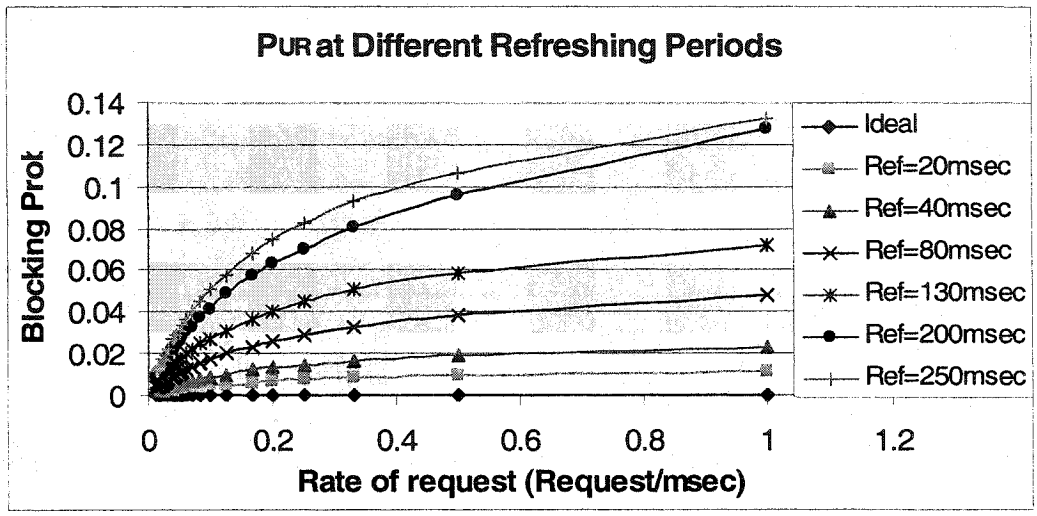


Figure 6.3.1.3: Probability of Unjustified Refusal at different refreshing periods

As a result, the refreshing period that makes P_{JR} higher is the refreshing period, which minimizes the other two types of blockings, P_{UA} and P_{UR}. Figures 6.3.1.2 and 6.3.1.3 show that P_{UA} and P_{UR} are relatively similar in value. For example, $P_{UA} = 0.081$ whereas $P_{UR} = 0.076$ at a refreshing period equal to 130msec. This can be understood by the fact that statistically to have an Unjustified Acceptance blocking P_{UA} is almost the same as having Unjustified Refusal blocking P_{UR} at variable rate of request because both of them result from the slow report of the routing protocol. The reason for this small difference is due to the contention blocking when two customers request the same wavelengths at the same time, hence one of those two customers will be blocked causing P_{UA} to increase.

Figure 6.3.1.4 shows the probability of a call being refused. The probability that a call is refused is the sum of the three types of blockings P_{JR}, P_{UA}, and P_{UR}. Clearly, the sum of the three types of blocking has approximately an equal values at all refreshing periods. The reason for this is that P_{JR} has a high value at small refreshing period and low value at high refreshing period. On the other hand, P_{UA} and P_{UR} have small values at short refreshing period and high vales at longer refreshing period. Therefore, at any refreshing period the sum

of the three values will be almost the same. In other words, at fast refreshing period the resources of the global network is being used more efficiently since there are no errors in the routing tables causing P_{JR} to rise whereas if the refreshing period is slow, errors will be introduced at the routing tables causing both P_{UA} and P_{UR} to rise. The summation of the value that both P_{UA} and P_{UR} gained is approximately equal to the value that P_{JR} decrease.

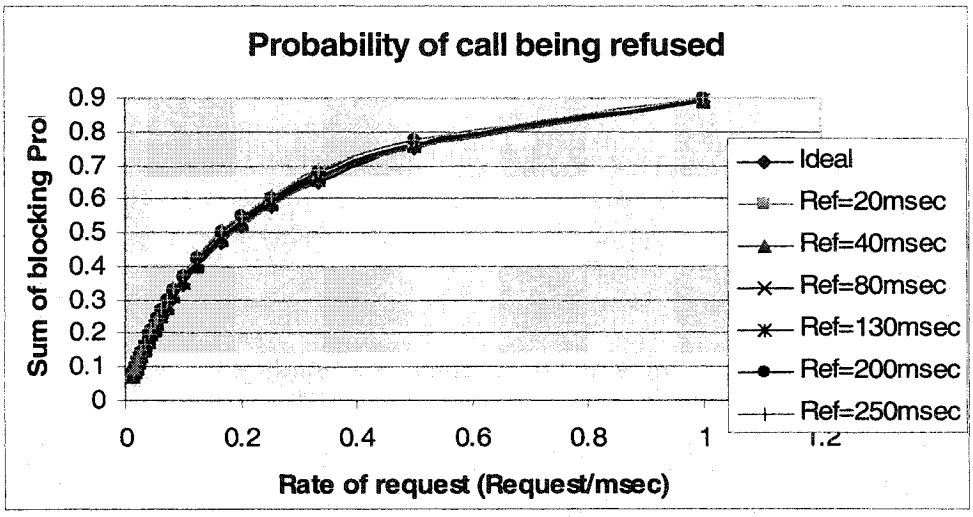


Figure 6.3.1.4: Probability of a call being refused

Figure 6.3.1.5 shows that the link utilization increases for faster refreshing periods, whereas slow refreshing shows lower link utilization. Slow refreshing cause more errors in the routing tables and hence reduces the number of lightpaths that are successfully established.

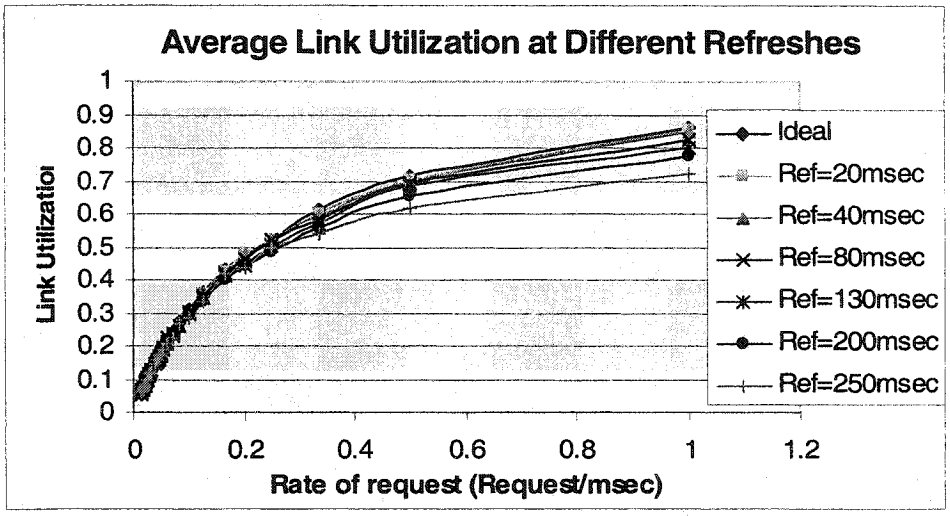


Figure 6.3.1.5: link utilization at different refreshing periods

6.3.2 The effect of Intra-Domain blocking

To simulate the effect of intra-domain blocking, we assume that each OXC is a complete AS that has its intra-domain routing protocol and its own internal blocking probability due to internal requests for the advertised resources. We also assume that each network administrator has the right to advertise a certain amount of its resources to neighbor ASs. In general, the type of information being passed from one AS to the other is based on their business agreement. This information can vary from full information about the internal topology to zero information [IDXu02] (see Section 5.2). Figure 6.3.2.1 shows that the intra-domain blocking probability is 0% for each AS. This means that all requests will be passed because each AS advertises all its available resources and there are no internal requests that might occupy these advertised resources.

In this section, we will investigate the effect of intra-domain blocking assuming that the blocking probability has the values of 0%, 1%, 5%, 10%, 20% and 30%. Besides, we fixed the refreshing period at 20msec because we are not interested in studying the effect of

refreshing periods here. Finally, we set the number of wavelengths to be 100 and the average life time of the requested lightpath is uniformly distributed with a mean equal to 4000msec.

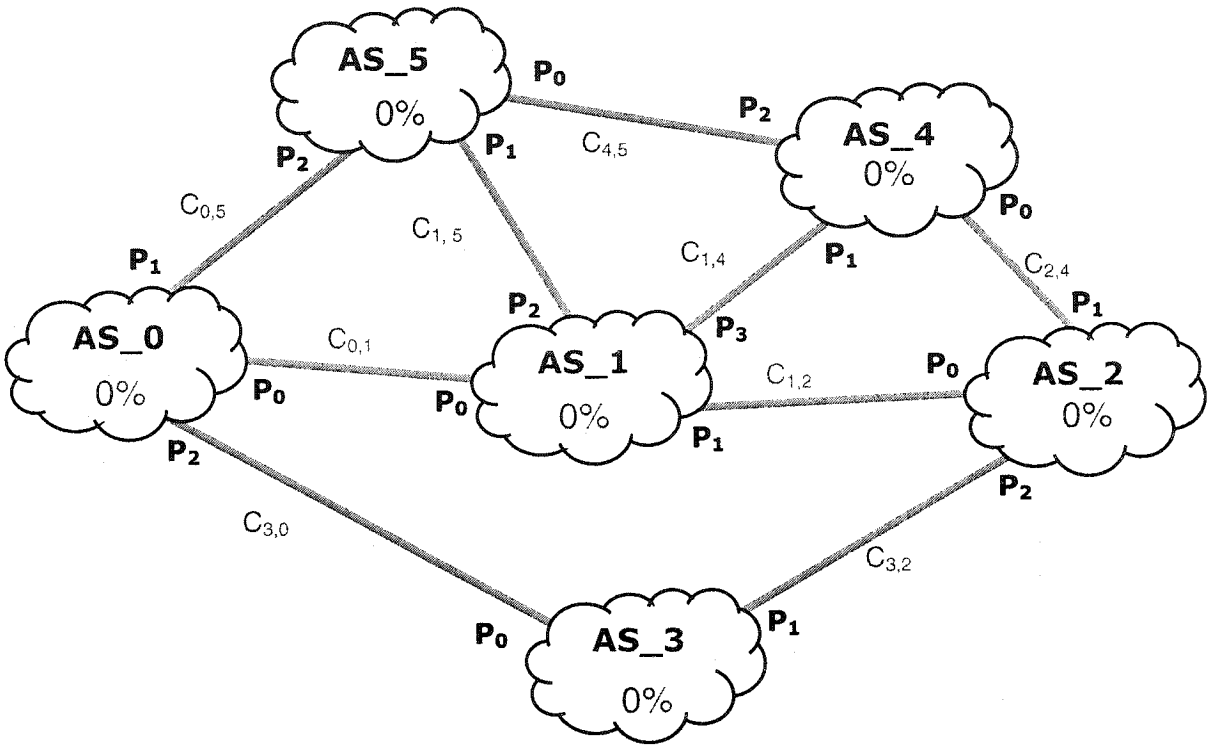


Figure 6.3.2.1: shows the Intra-Domain Blocking probability

Figure 6.3.2.2 illustrates the effect of the intra-domain blocking on P_{JR}. Clearly, the intra-domain blocking has effect on P_{JR}. Increasing the intra-domain blocking will prevent some resources from being successfully established leaving more available resources at the edge routers that will reduce P_{JR}.

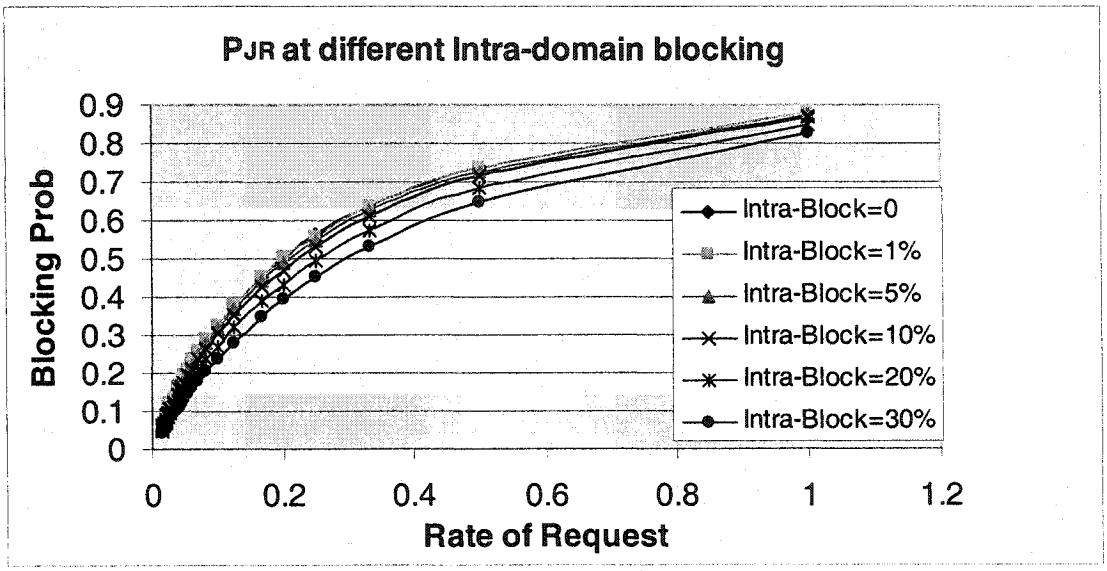


Figure 6.3.2.2: P_{JR} at 0, 1, 5, 10, 20 & 30% Intra-domain blocking

Figure 6.3.2.3 shows the effect of intra-domain blocking on P_{UA}. In this graph, P_{UA} represents the value that resulted from both, the blocking that took place due to errors introduced in the routing tables and the blocking that took place due to intra-domain blocking. It is clear that P_{UA} is significantly increased due to intra-domain blocking.

For example, at intra-domain blocking equal to 30%, P_{UA} is equal to 0.35. One might wonder why the P_{UA} decreases as the rate of request increases. The reason for this is that the definition of P_{UA} is the number of initiated lightpath that has been blocked by the network over the total number of requests. Of course, as the number of total requests increases P_{UA} will decrease.

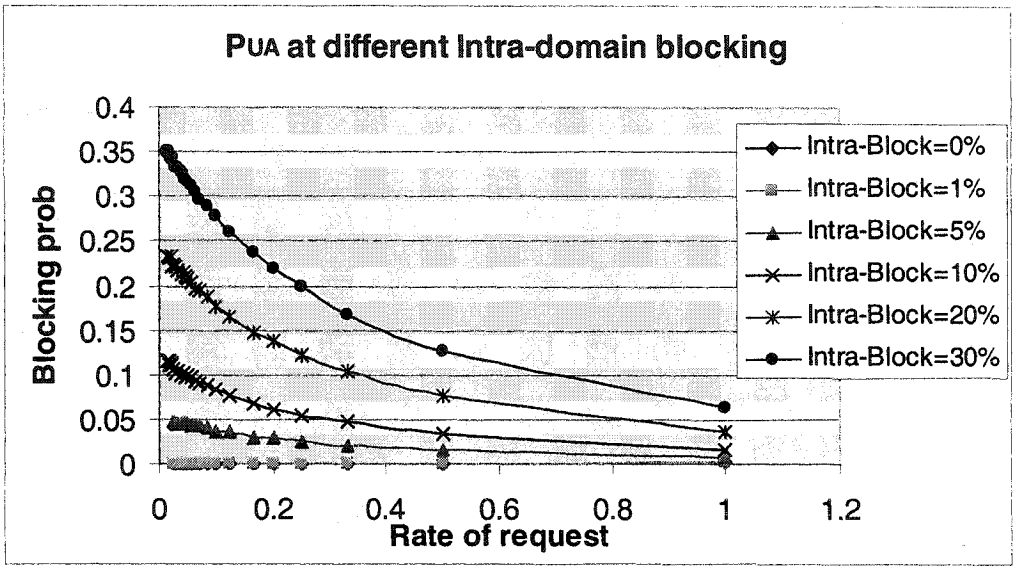


Figure 6.3.2.3: P_{UA} at 0, 1, 5, 10, 20 & 30% Intra-domain blocking

Figure 6.3.2.4 shows that P_{UR} decreases as the intra-domain blocking value increases. P_{UR} follow the trend of P_{JR}.

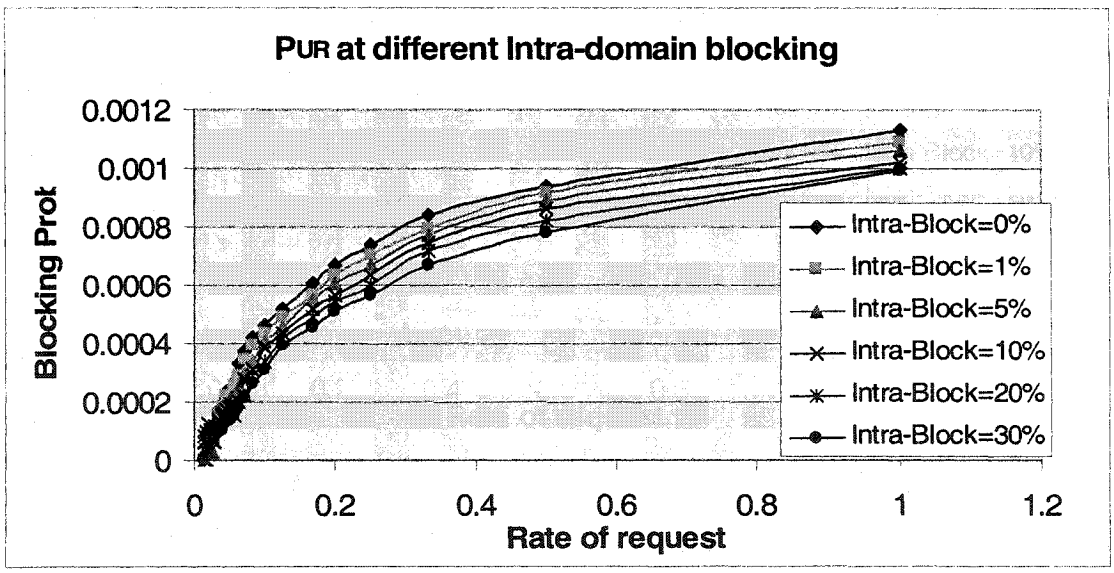


Figure 6.3.2.4: P_{UR} at 0, 1, 5, 10, 20 & 30% Intra-domain blocking

Finally, we investigate the effect of the intra-domain blocking on the average link utilization. Figure 6.3.2.5 shows the average link utilization for 0%, 1%, 5%, 10%, 20% and 30% intra-domain blocking.

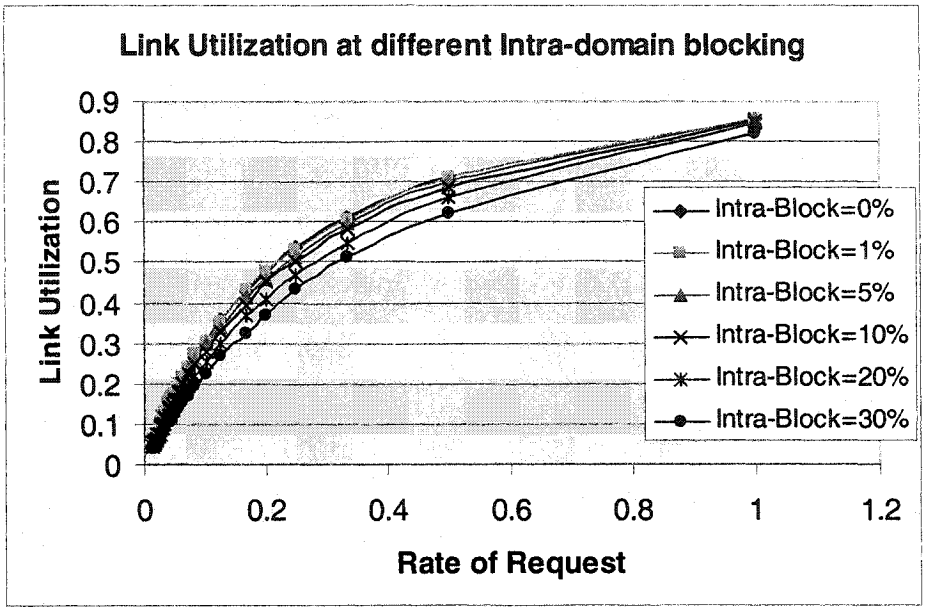


Figure 6.3.2.5: Link utilization at 0, 1, 5, 10, 20 & 30% Intra-block

Clearly, this graph shows that intra-domain blocking has effect on the average link utilization. This is because of the fact that intra-domain blocking prevents some resources to be successfully established that will reduce the average link utilization.

6.3.3 Threshold change percent

In Section 5.5, we talked about two schemes for advertisement. The first scheme is to trigger the advertisement based on the expiration of the refreshing timer, whereas the second scheme considers the number of changes that took place in the link table to trigger an advertisement.

We find that the second scheme is better because if the number of changes exceeded a certain threshold then an automatic advertisement will take place without waiting for an expiration of the timer. This means that the advertisement is done based on the amount of information that has changed since the last up-date.

In this section, we will investigate the triggering of the advertisement based on the following percentage of change in the link table: 0, 5, 10, 20, 30, & 40 %.

Figure 6.3.2.1 shows the simulated network. We assume that each link has 64 wavelengths. We increase the number of wavelengths for this part of the simulations for granularity purposes. This is because if we have only 10 wavelengths, then each reserved wavelength represents 10 percent of change. However, if each link has 64 wavelengths then every 6 wavelengths represent 10%.

Figure 6.3.3.2.1 shows P_{JR} at different thresholds. clearly Figure 6.3.3.2.1 shows that waiting for the counter to indicate three changes out of 64, i.e. "5% of change", does not have such a bad effect on P_{JR}, however, waiting for 26 changes out of 64, i.e. "40% of change", has a bad effect on P_{JR}. Again, if P_{JR} is reduced, this is not necessary a good sign. It means that more errors were introduced into the routing table due to the slow reporting and hence less lightpaths will be established and consequently more wavelengths will be available for costumers, thus reducing P_{JR}.

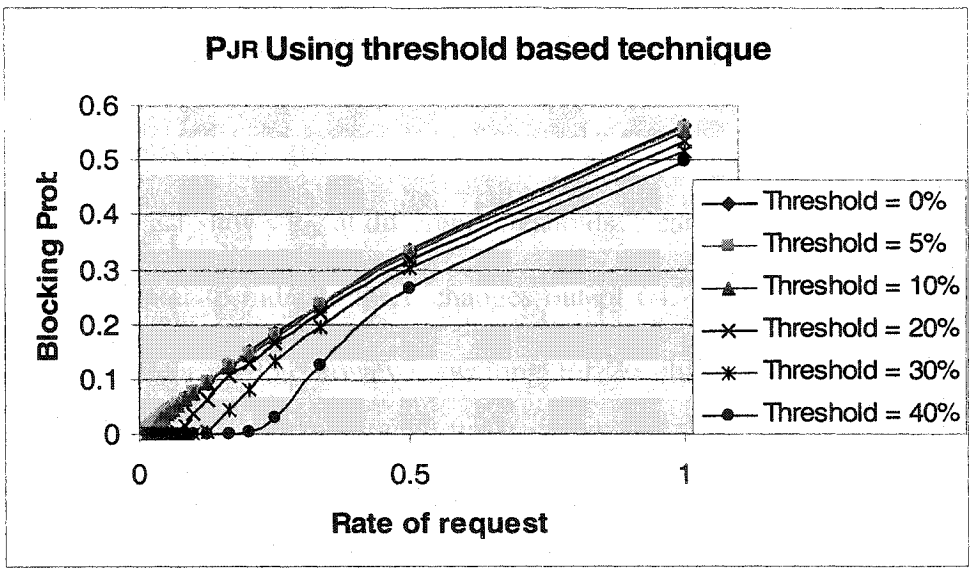


Figure 6.3.3.2.1: P_{JR} at Different Threshold

Figure 6.3.3.2.2 shows P_{UA} at different thresholds. The higher the threshold is the higher P_{UA} will be. P_{UA} is strongly affected by the threshold value. As we defined P_{UA} earlier, P_{UA} occurs due to slow reporting of the routing protocol. As a matter of fact, if the threshold value increases, this will cause lots of errors in the routing tables because each edge node will report changes slowly causing more errors in the routing tables and hence the requested lightpaths will not be successfully established due to the fact that the initiation of requests based on wrong routing information will definitely lead to higher values of P_{UA}.

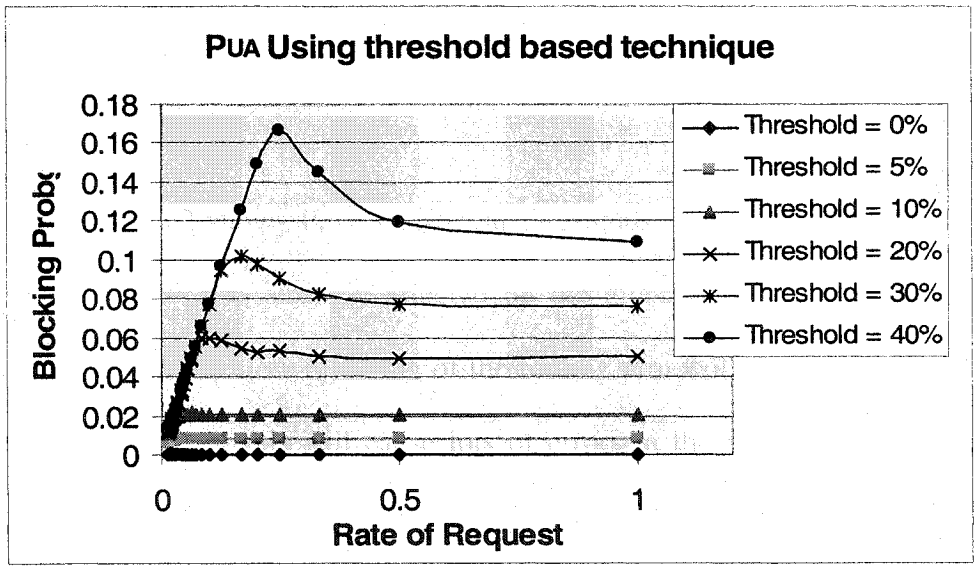


Figure 6.3.3.2.2: P_{UA} at different threshold

All curves in figure 6.3.3.2.2 shows that the value of P_{UA} starts to increase as the request rate increases, however, for the lower curves P_{UA} does not increase that much, because changes are propagated more frequently keeping the routing tables up-to-date, whereas when the refreshing period is very large, as in the upper curves, P_{UA} increased dramatically.

The upper curve in Figure 6.3.3.2.2 shows that P_{UA} is increasing as the request rate increased to a certain point, then it decreases to a relatively lower value. The reason for this behavior is that as the request rate increases, more lightpaths will be established. These

changes are not reported fast enough to the other edge nodes and hence the value of P_{UA} will increase. However, once the slow advertisement starts to report these changes, P_{UA} will decrease to a certain level because the edge routers will start updating their routing tables about the status of the reserved wavelengths. Eventually, the reserved wavelengths will be shown at each edge router and hence they can not be used by any other customer till they are released. Increasing the number of reserved wavelengths will reduce the number of signaling messages that might contribute to P_{UA} and this explains the strange behavior of the upper curve.

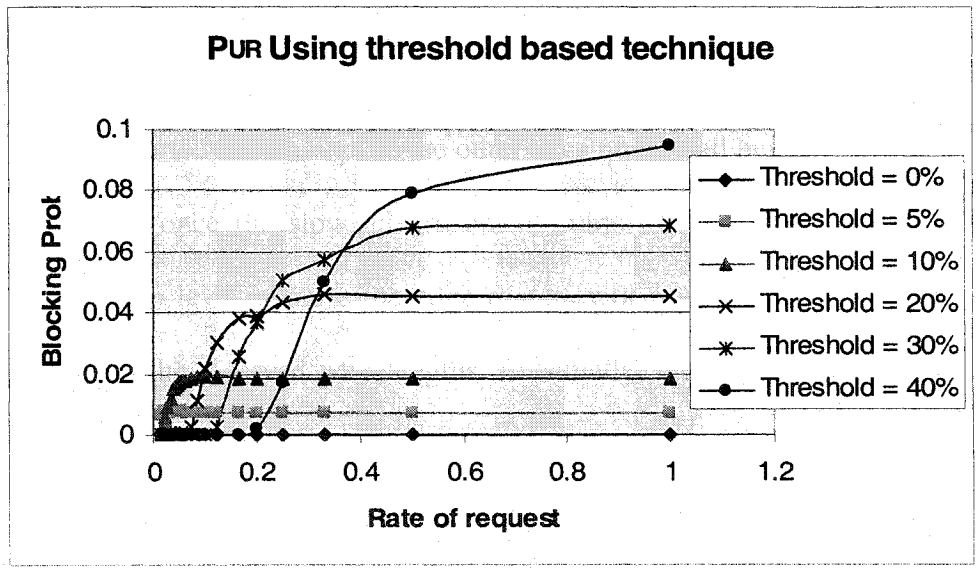


Figure 6.3.3.2.3: P_{UR} at different thresholds

Figure 6.3.3.2.3 shows that also P_{UR} is strongly affected by the threshold value. It shows that P_{UR} can reach a value approximately equal to 0.1 at request rate equal to 1 when the threshold equals 40%.

Figure 6.3.3.2.4 shows the average link utilization at different thresholds; clearly the higher the threshold is the less utilization we have.

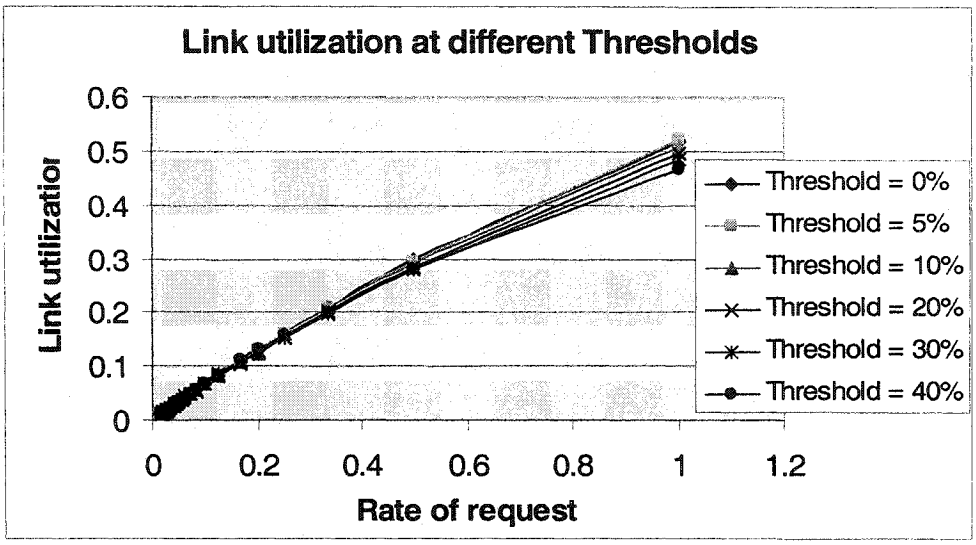


Figure 6.3.3.2.4: Link Utilization at different thresholds

6.3.4 Link utilization

Figure 6.3.4.4 shows the utilization of the individual link in the simple network; it is clearly shown that the link utilization increases when the request rate increases.

Furthermore, we can see that the link utilization for all fibers varies within a certain margin. For example, at a rate of 0.35 request/msec, the link utilization varies between a maximum value equal to 0.7 and a minimum value equal to 0.4, depending on the individual link.

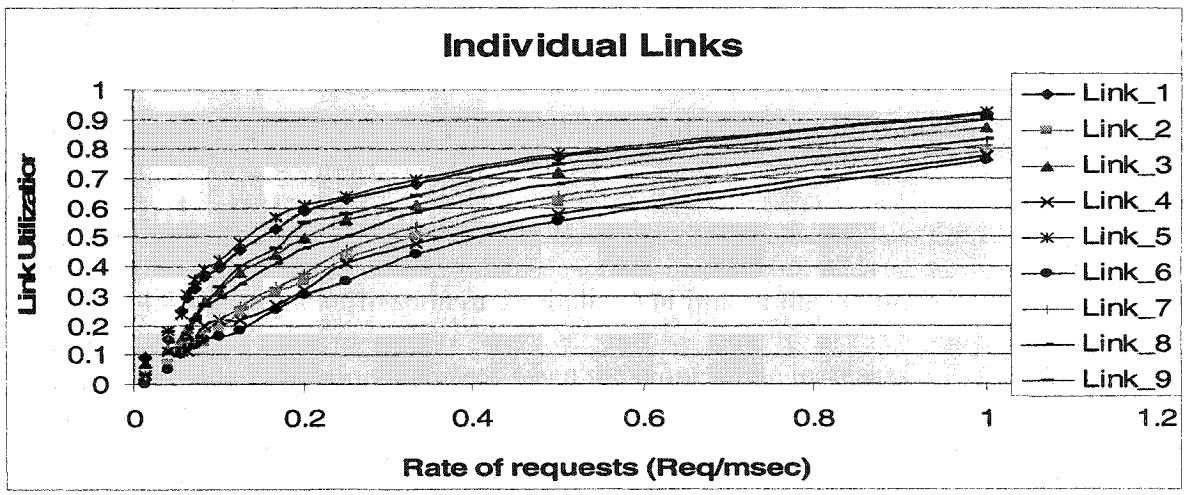


Figure 6.3.4.1: Individual link utilization at different rate

6.4 Advanced Research Projects Agency Network

This section investigates the three types of blocking within the ARPANET architecture shown in Figure 6.4.1. This network has a diameter of three hops.

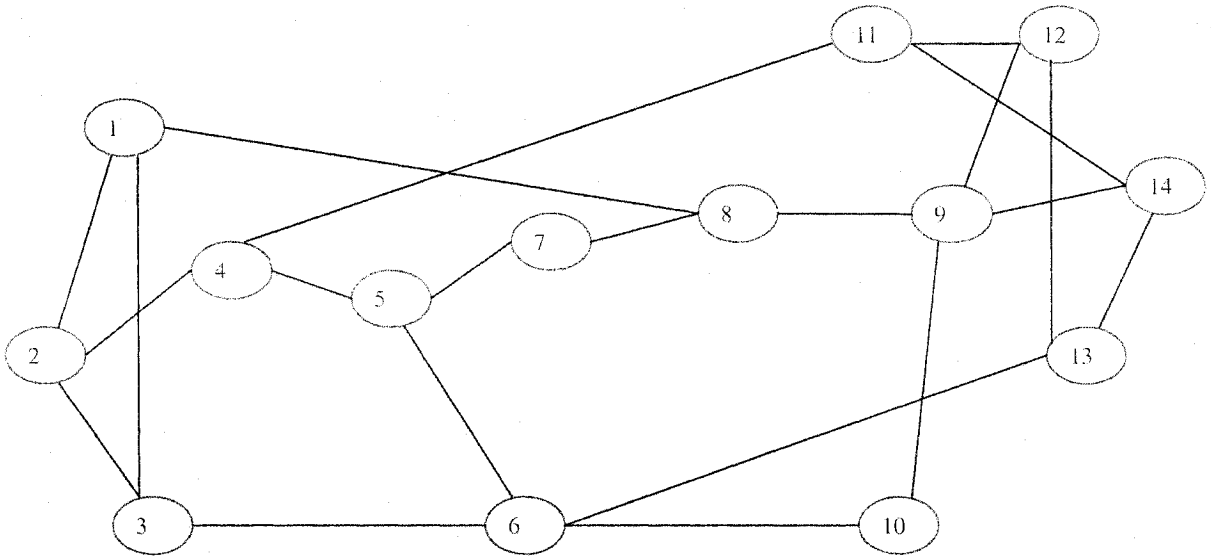


Figure 6.4.1: ARPANet.

6.4.1 Effect of the Refreshing Period

In this section, we are interested in finding the effect of the refreshing period over the blocking probabilities and the link utilization. We chose the number of wavelengths to be 10 and the average life time of the requested wavelengths to be 400msec. The rate of request was changed from 0 to 1 request/msec, and it is assumed that the refreshing period is very small compared to the lifetime of the requested wavelength.

All the results are similar to the one we obtained for the simple network. The effect of the refreshing period for the ARPANET is shown in Appendix A.

6.4.2 Intra-Domain blocking effect

In this section, we investigate the effect of the intra-domain blocking on the three types of blocking and the link utilization for intra-domain blocking probabilities equal to 0, 5, 10, 20, 30%. Again, the results are similar to those for the simple network. They are shown in Appendix B.

6.4.3 Advertising based on the number of changes

In this section, we only investigate the advertising based on the number of changes.

We are interested in investigating the effect of increasing the threshold on the three types of blocking and the link utilization. Without any surprise, the results follow the same trend as the results obtained for the simple network architecture, as shown in Appendix C.

6.4.4 The effect of the network diameter on P_{JR}, P_{UA} and P_{UR}

Logically, if the number of hops increases then the probability of establishing a lightpath successfully across the path will be reduced. In this section, we investigate the blocking that will take place at each hop. In other words, we want to investigate the effect of increasing the diameter of the network on the three types of blocking. We perform this simulation with the ARPANET architectures, which has a diameter of 3 (Figure 6.4.1). We set the number of wavelength to 64 and we fixed the refreshing period to be 20msec. Normally, if the refreshing period is increased, P_{JR} will decrease whereas P_{UA}, P_{UR} will increase.

Figure 6.4.4.1, Figure 6.4.4.2 and Figure 6.4.4.3 show the three types of blocking probability (P_{JR}, P_{UA}, P_{UR}) for three different hop counts. Each of these graphs shows the blocking rate for single, double and triple hop paths.

The value of P_{JR} for single hop paths is calculated by dividing the number of blocked requests destined for a single hop path by the total number of requests that are destined for a single hop path; and similarly for double and triple hop paths.

The blocking rates for the other types of blocking (P_{UA}, P_{UR}) are calculated similarly. The three figures show the same behavior; clearly, the blocking rate is increased as the number of hops increased. This is due to the fact that the longer the path is, the more likely some blocking will take place. Furthermore, the longer the path is, the more likely P_{UA} and P_{UR} blocking will occur due to the longer time needed for the new routing information to be propagated.

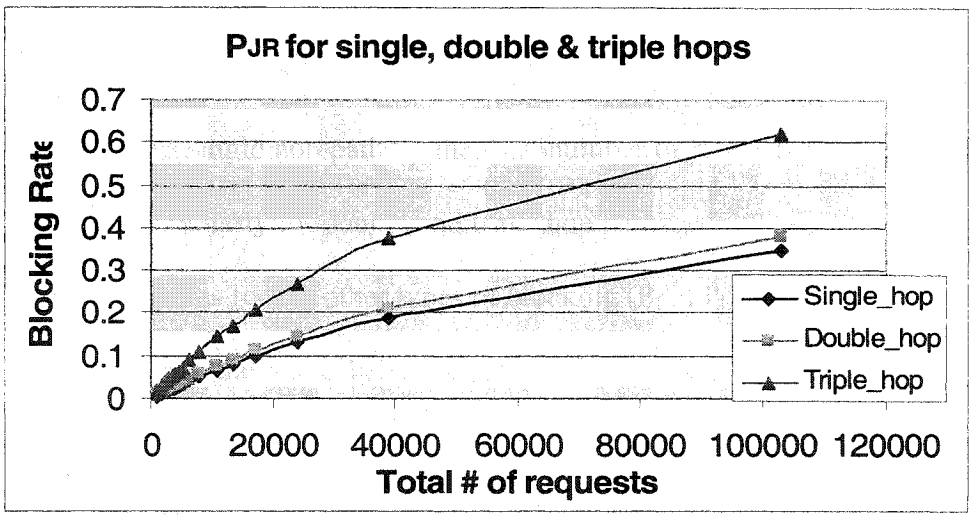


Figure 6.4.4. 1: P_{JR} at single, double and triple hops

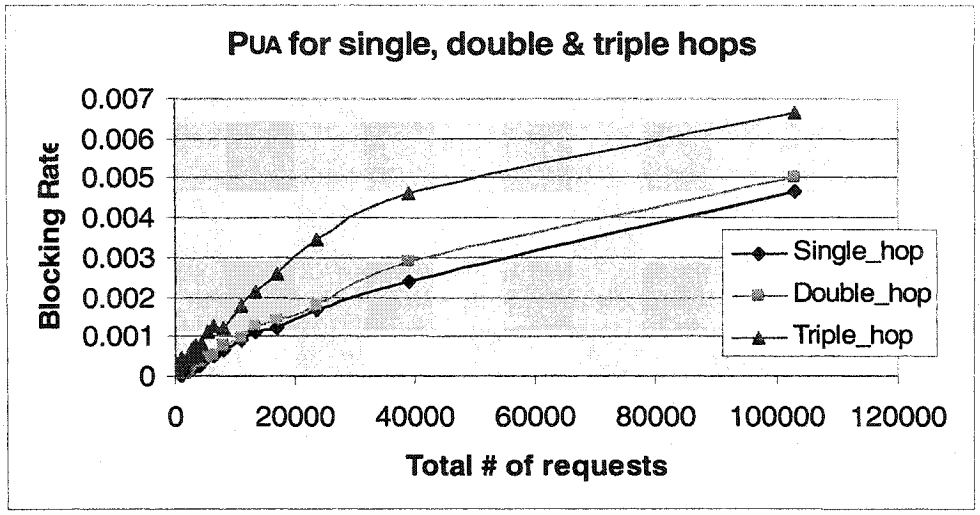


Figure 6.4.4. 2: P_{UA} at single, double and triple hops

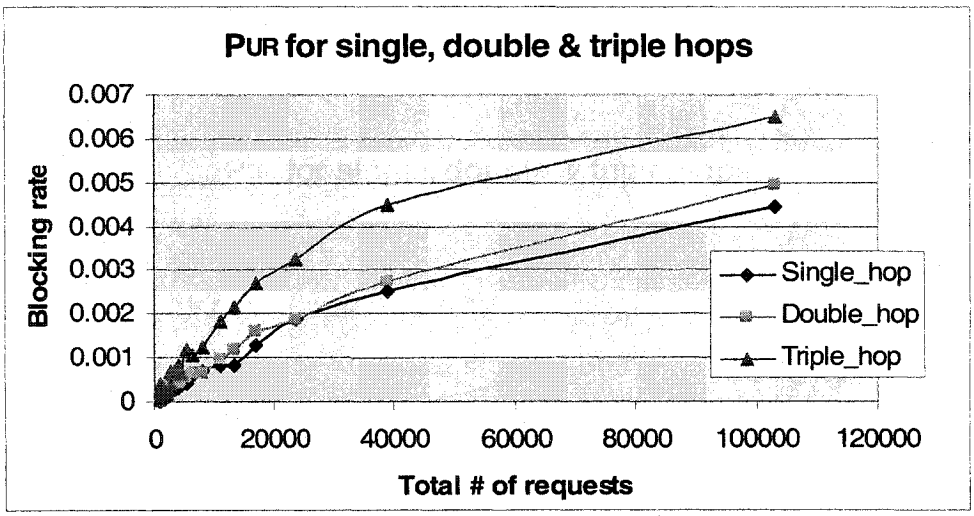


Figure 6.4.4. 3: P_{UR} at single, double and triple hops

We may wonder why there is a blocking for single hop. Logically, there should not be. The reason for this is that the routing table in a given node is not updated simultaneously with the link tables within our simulation model, i.e. when a request comes we update the status of the wavelengths in the link table without changing anything in the routing table.

Therefore, a certain wavelength might appear available in the routing table while it is reserved on the link table, which may cause blocking for a single hop.

6.5 Ring Network

This section investigates the blocking probability for a ring network, as shown in Figure 6.5.1. Normally, a ring architecture has the lowest number of connections. We are interested to see the effect of reducing the number of links on the three types of blocking probabilities and the link utilization.

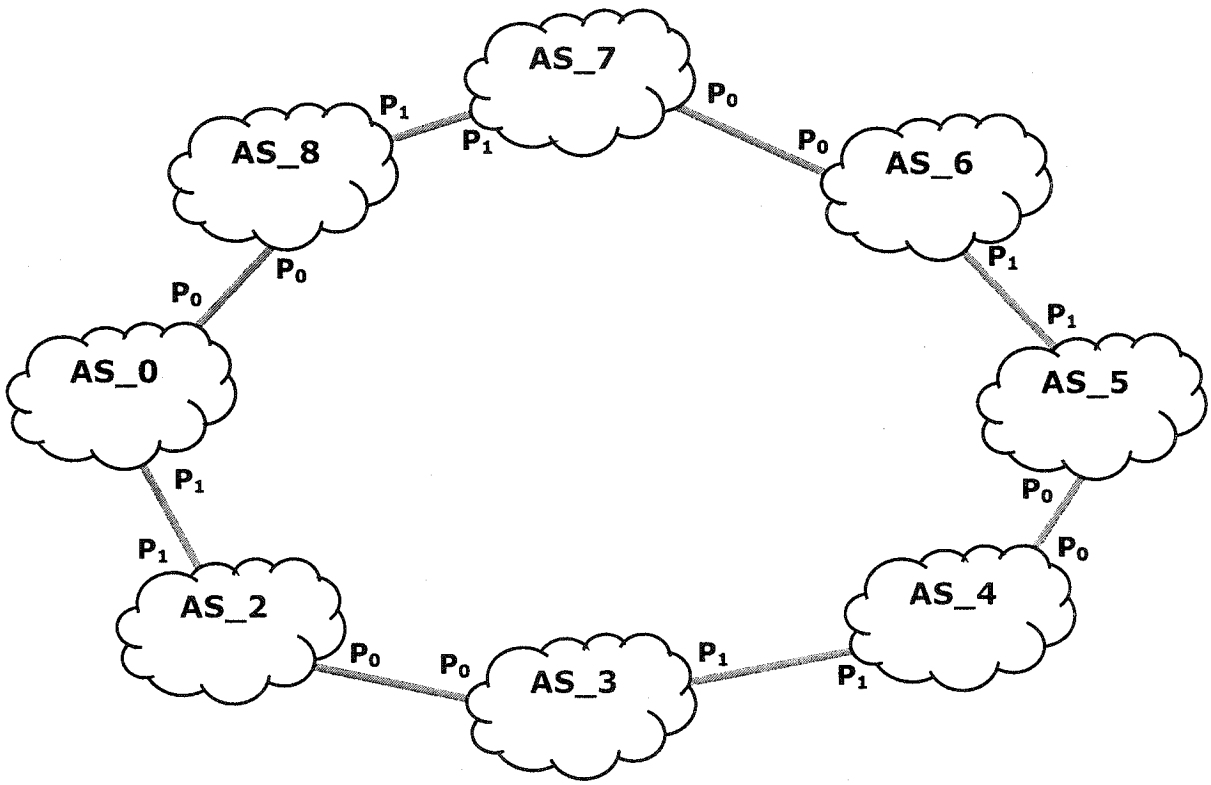


Figure 6.5.1: Ring Architecture

Appendix C shows the results for this network. Clearly, P_{JR} for the ring architecture is relatively high. This is due to the fact that the ring architecture has poor connectivity among the nodes; this will increase the blocking for lightpaths requests destined for triple or four hop paths. Furthermore, this architecture increases the propagation time that is required to advertise the routing changes.

6.6 European network

This section investigates another network with relatively high connectivity, as shown in Figure 6.6.1[CLL 03]. Appendix D shows the results for this network. The results show the same trend obtained for the ARPANET, however, this network has a slightly lower P_{JR}. This is due to the fact that this network has more connections that can absorb a higher number of requests.

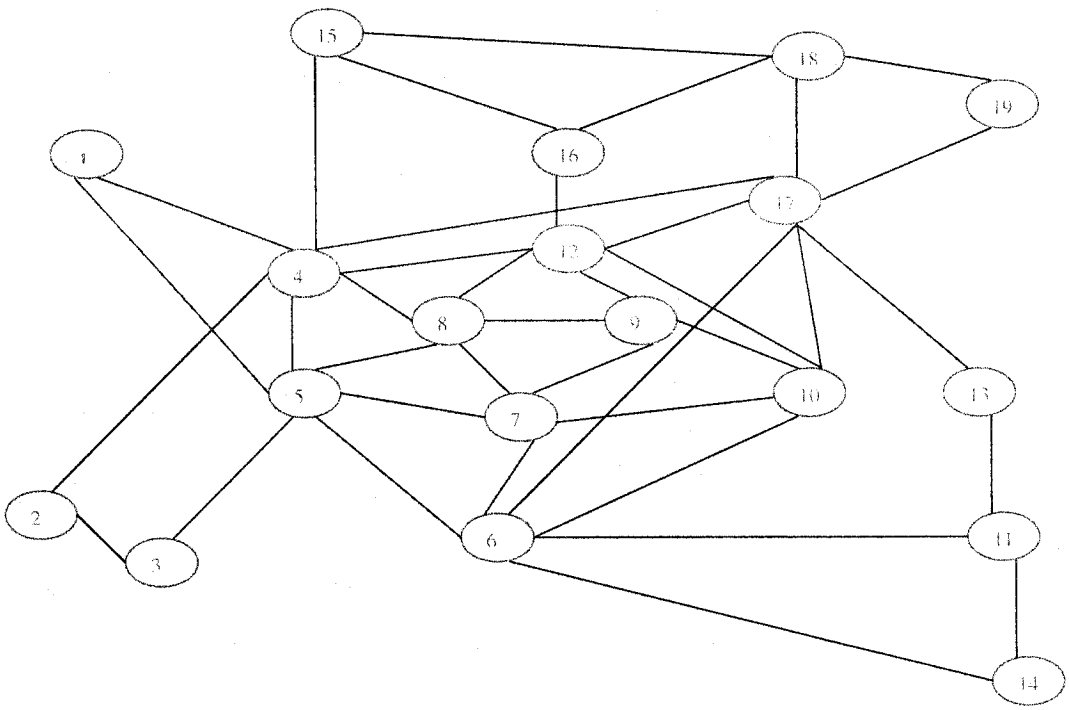


Figure 6.6.1: Network

6.7 Star Architecture

The considered star architecture has a special topology that is different from the previous architectures, as shown in Figure 6.7.1. This architecture is used to connect different AS through a core switch. The special feature obtained from this star architecture is that the number of hops is always one. This implies faster propagation of the routing information and hence more lightpaths will be successfully established. Each AS will inform other ASs about the available wavelengths through the core switch. The core switch has no incoming nor outgoing traffic; it is only used to connect the ASs at the optical level. Appendix F shows the results for this architecture.

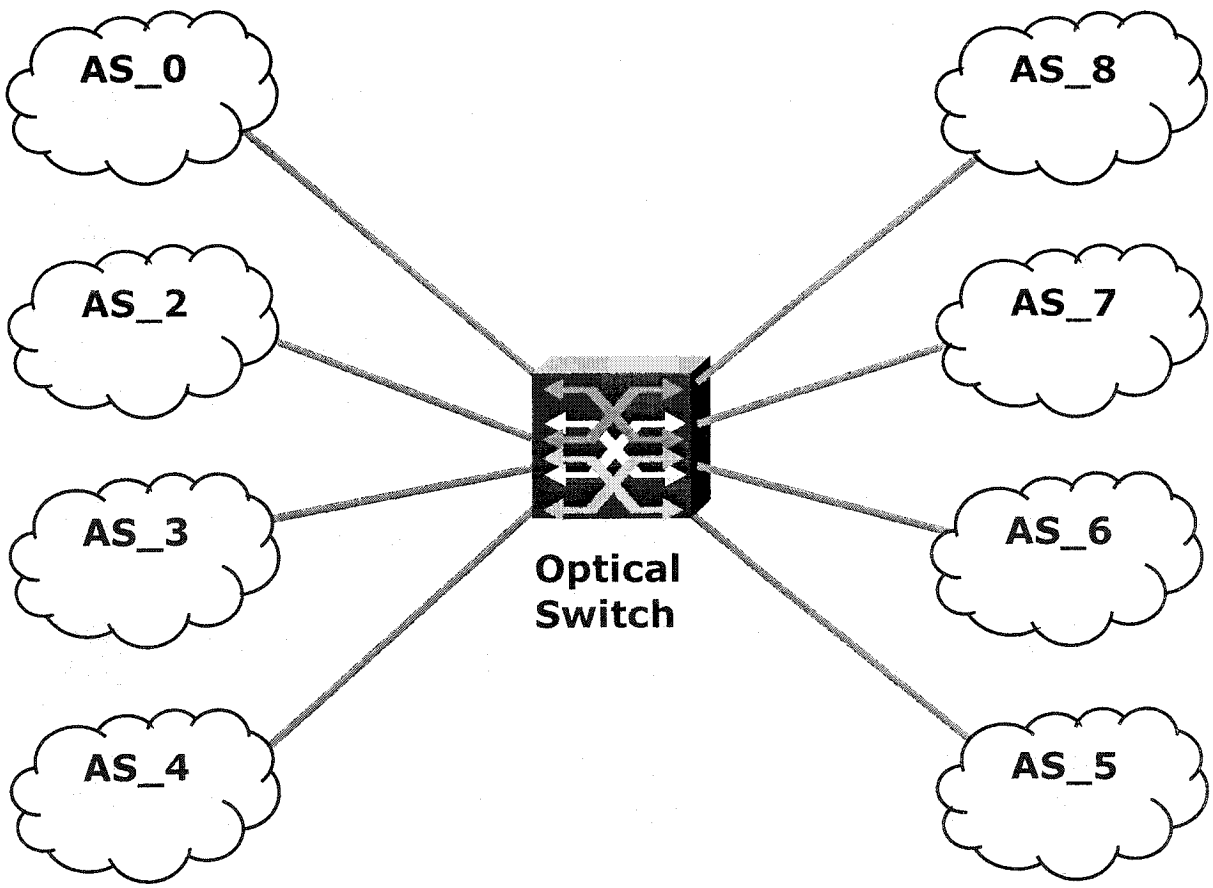


Figure 6.7.1: Star Architecture

Chapter 7 Discussion & Conclusion

7.1 List of contributions

Our major goal was to define a routing protocol that performs optical routing among ASs. We proposed a new routing protocol called Optical Routing Border Gateway Protocol (ORBGP). It exploits most of BGP functionality with some modifications to support the exchange of optical routing information.

The major contributions of our work are the following:

- We extended BGP to exchange the optical routing information by defining a new attribute called "optical attribute".
- We defined two new advertising schemes, *the first scheme* based on the expiration of a time-out/refreshing period, and *the second scheme* is based on the number of changes that took place in a given link table (see Section 5.5).
- We investigate the performance of ORBGP on different network architectures; we showed that the diameter and the connectivity of the network introduce a challenge that faces the performance of ORBGP. In other words, the higher the network diameter and the smaller the connectivity, the more difficult it is to establish a lightpath.
- We proposed a new distribution of the routing and signaling functions among different protocols in the context of IP over optical networks (see Figure 5.1). We proposed the following for future optical control plane:
  - OSPF can perform the I-BGP task.

- GMPLS performs signaling within each AS
- ORBGP perform routing among AS
- New signaling protocol has to be defined for signaling among ASs

We believe that this environment of protocol interaction is very useful since it eliminates the weak points of I-BGP and allows each AS to have a complete privacy on the type of the advertised routing information.

In conclusion, our work shows great features compared to the previous approaches. Our approach provides the following:

- Unlike any previous approach, ORBGP enables edge nodes to advertise the routing changes that occur in any of its link tables to update the other nodes' routing tables. We showed that each node can advertise when a certain number of changes have occured in any of its link tables. The number of changes required to trigger an advertisement is set by the network administrator (see Section 5.5).
- ORBGP provides complete privacy for the advertising node by giving this node the ability to specify the type and the amount of advertised routing information. For example, we showed in Section 5.3 that the only advertised information is the wavelength that could be used to cross the adverting AS without giving any detailed information about the internal topology.
- Our approach aims to assure the privacy of each AS. In the other approaches, explained in Sections 4.2.1, 4.2.2 and 4.2.3, each AS gives neighbor ASs virtual control on its OXCs to set-up or tear-down a lightpath.
- We assigned the I-BGP task to OSPF to eliminate the risk of having routing loops. We have discussed the weak points of the I-BGP Operation mode in Section 3.3.2.

- ORBGP performs only a routing task, it does not perform any signaling task. We have noted that the other approaches explained in Sections 4.2.1, 4.2.2 and 4.2.3 use BGP update messages to perform signaling. Keeping in mind that the I-BGP operation mode is manually configured and subjected to miss-configuration, loops are more likely to occur if I-BGP operation mode is used during signaling.
- ORBGP allows edge routers to treat each destination AS as one prefix. ORBGP provides each edge node the wavelengths that could be used to reach that AS. This approach is very scalable compared to other approaches. For example, we have explained in Section 4.2.2 that each OXC is treated as an independent unit. In other words, each independent unit is treated as an independent AS despite the fact that all these independent units belong to a single domain that is managed by a single network administrator.

7.2 Future work

- For further research, we are interested in investigating the performance and the scalability of OSPF if it is used to perform the I-BGP task. This is because I-BGP has many disadvantages, e.g. it is manually configured and subject to oscillation.
- We are also interested in route aggregation. In other words, how could we aggregate a certain number of ASs into one prefix. We are interested in the performance of ORBGP across many aggregated networks.
- Finally, we are interested in introducing new policy parameters to increase the link protection and network utilization, for instance:

- Protection parameter. Each AS can advertise a number to indicate the protection availability for the requested lightpath. The higher this number the higher the link protection is.
- Diversity parameter and shared risk. This parameter allows a certain node to indicate a divergent route to the same destination. For example, if node A has two established lightpaths to node B. The Diversity parameter allows node A to determine whether those two lightpaths have no single point of failure along the whole path. And hence this will allow node A to provide lightpaths to its costumers with high robustness because, if one path failed, node A can use the other lightpath.

Bibliography

[RFC .]: Reference to a Request For Comments (RFC) document issued by the IETF (Internet Engineering Task Force).

[ID]: Reference to an Internet Draft (ID) document issued by the IETF (Internet Engineering Task Force).

Books and Papers

[BCB et al 00] Jerry Banks, John S. Carson II, Barry L. Nelson, "Discrete-Event System Simulation". ISBN 0-13-088702-1, Prentice Hall PTR, Inc., 2000.

[BDLT 01] Ayan Banerjee, John Drake, Jonathan Lang, "Generalized Multi-protocol Label Switching: An Overview of Signaling Enhancements and Recovery Techniques", Communications Magazine, IEEE, pp. 144 -151, Jan. 2003.

[Bla et al 02] Uyless D. Black, "Optical Networks: Third generation Transport Systems". ISBN 0-13-060726-6, Prentice Hall PTR, Inc., 2002.

[BLRW 02] Anindya Basu, Chih-Hao Luke Ong, April Rasala, Gordon Wilfong, "Route Oscillation in I-BGP with Route Reflection", In Proceedings of ACM SIGCOMM, pp. 235 - 247, August 2002.

[BSO 02] Greg M. Bernstein, Vishal Sharma, Lyndon Ong, "Inter-domain optical routing", Journal of Optical Networking '02 , Vol. 1, No. 2, pp. 80 - 92, Feb. 2002.

[CLL 03] Xiaowen Chu, Bo Li, Jiangchuan Liu, "wavelength Converter Placement under a Dynamic RWA Algorithm in Wavelength-Routed All-Optical Network", IEEE Transactions on Communications, Vol. 51, No. 4, pp. 607 - 617, April 2003.

[CZ 96]Imrich Chlamtac, Tao Zhang, "Lightpath (wavelength) Routing in Large WDM Networks", IEEE Journal on selected areas in communication, Vol. 14, No. 5, pp. 909 - 913, June 1996.

[FSHL 01] Mark Francisco, Stephen Simpson, Changcheng Huang, Ioannis Lambadaris, Bill St-Arnaud, "Inter-Domain Routing in Optical Networks", Opticomm 2001, Denver, Colorado, August 2001.

[GG 02] Timothy G. Griffin, Gordon Wilfong, "On the Correctness of IBGP configuration", In Proceedings of ACM SIGCOMM, pp. 17 - 29, August 2002.

[JY 02] Sangjin Jeong, Chan-Hyun Youn, "Instability Analysis for OBGP Routing Convergence in Optical Networks", In Proceedings of ICOIN-16, Korea, Jan. 2002

[Leo et al 01] Alberto Leon-Garcia, Indra Widjaja, "Communication Networks: Fundamental Concepts and Key Architectures". ISBN 0-07-250353-X, McGraw-Hill, Inc., 2001.

[Liu et al 02] Kevin H. Liu, "IP over WDM". ISBN 0-470-84417-5, John Wiley & Sons Ltd, 2002.

[LLW 00] Kevin H. Liu, Changdong Lui, John Y. Wei, "Overlay vs. Traffic Engineering for IP/WDM Networks", In GLOBECOM '00, IEEE , Vol. 2 , pp. 1293 -1297, Nov. 2000.

[MF 02] Olaf Maennel, Anja Feldmann, "Realistic BGP Traffic for Test Labs", In Proceedings of ACM SIGCOMM, pp. 235 - 247, August 2002.

[Mou et al 03] Hussein T. Mouftah, Pin-Han Ho, "OPTICAL NETWORKS Architecture and survivability". ISBN 1-4020-7196-5, Kluwer Academic Publishers, Inc., 2003.

[Moy et al 98] John T. Moy, "OSPF: Anatomy of an Internet Routing Protocol". ISBN 0-201-63472-4, Addison Wesley Longman, Inc., 1998.

[MWT 02] Ratul Mahajan, David Wetherall, Tom Anderson, "Understanding BGP configuration", In Proceedings of ACM SIGCOMM, pp. 3 - 16, August 2002.

[SS 03] Sasaki, G.H. Ching-Fong Su, "The interface between IP and WDM and its effect on the cost of survivability", Communications Magazine, IEEE, Vol. 41, pp. 74 - 79, Jan. 2003.

[Sta et al 00] William Stallings, "Data & computer communication". ISBN 0-13-084370-9, Prentice Hall PTR, Inc., 2000.

[Ste et al 99] John W. Stewart, "BGP-4: Inter-Domain Routing in the Internet". ISBN 0-201-37951-1, Addison Wesley Longman, Inc., 1999.

[Tho 01] Stephen A. Thomas, "IP Switching and Routing Essentials: Understanding RIP, OSPF, BGP, MPLS, CR-LDP, and RSVP-TE". ISBN 0-471-03466-5, WILEY, Inc.,2001.

[Tom et al 01] Peter Tomus, Christian Schmutzer, "Next Generation Optical Networks". ISBN 0-13-028226-x, Prentice Hall PTR, Inc., 2001.

[Wei 02] John Y. Wei, "Advances in the Management and Control of Optical Internet", In IEEE J. Select. Area Commun., vol. 20, no.4, pp. 768 - 785, May 2002.

RFC

[RFC1771] Y. Rekhter, "A Border Gateway Protocol 4 (BGP-4)", March 1995.

[RFC2453] G. Malkin, "RIP Version 2", November 1998.

[RFC1195] R. Callon, "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments", December 1990

[RFC1142] D. Oran, "OSI IS-IS Intra-domain Routing Protocol", February 1990

[RFC 2328] John T. Moy, "OSPF Version 2", STD 54, April 1998.

[RFC 2370] R. Coltun, "The OSPF Opaque LSA Option", July 1998.

[RFC 2858] T. Bates, Chandra, Katz, and Y. Rekhter, "Multiprotocol Extensions for BGP4", June 2000, RFC 2858

[RFC 1997] R. Chandra, P. Traina, "BGP Communities Attribute", August 1996

[RFC 3209] D. Awduche, L. Berger, D. Gan, T. Li, and V. Srinivasan, "*RSVP-TE: Extensions to RSVP for LSP Tunnels*", RFC 3209, December 2001.

[RFC 3471] L. Berger, "*Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description*", RFC 3471, January 2003.

[RFC 3472] P. Ashwood-Smith, L. Berger, January, "*Generalized Multi-Protocol Label Switching (GMPLS) Signaling Constraint-based Routed Label Distribution Protocol (CR-LDP) Extensions*, RFC 3472, 2003.

[RFC 3473] L. Berger "*Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions*", RFC 3473 January 2003.

[RFC 1322] D. Estrin, Y. Rekhter, S. Hotz "*A Unified Approach to Inter-Domain Routing*", RFC 3473 May 1992.

Internet Draft

[IDBer00] G. Bernstein, "Optical Inter Domain Routing Considerations", draft-ipo-optical-inter-domain-00.txt, Internet Draft, Work in Progress, November 2001

[IDBla01] Marc Blanchet, " Optical BGP (OBGP): InterAS lightpath Provisioning", draft-parent-obgp-01.txt, Internet Draft, Work in Progress, August 2001

[IDXu02] Yangguang Xu, "A BGP/GMPLS Solution for Inter-Domain Optical Networking ", draft-xu-bgp-gmpls-02.txt, Internet Draft, Expiration date 2002, June 2002

Links

[BF 00] Paul Brittain, Adrian Farrel, MPLS Traffic engineering a Choice of signaling Protocols, Analysis of the similarities and differences between the two primary, white paper, Data Connection, January 2000.

[OPNET] OPNET Technology Inc., http://www.opnet.com

Appendices:

A – Effect of the refreshing period for the ARPANET

Figure A.1 shows the effect of the refreshing period on P_{JR} blocking. There are seven curves in this graph. These curve shows that the refreshing period does have effect on P_{JR}. We can tell that the longer the refreshing period, is the lower the P_{JR}. Clearly, Figure A.1 shows that at refreshing period equal to 10msec causes $P_{JR} = 0.8$ whereas at refreshing period equal to 250msec causes $P_{JR} = 0.7$.

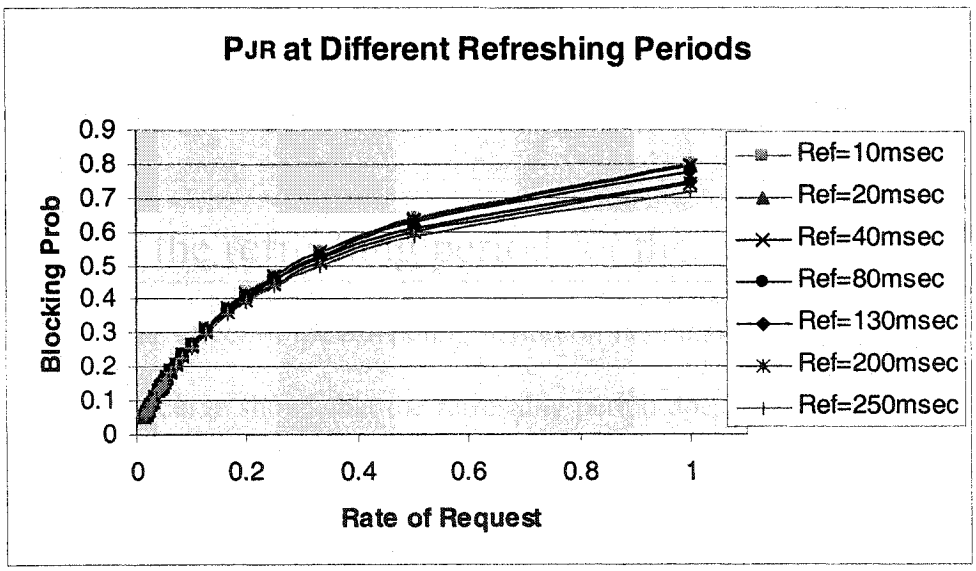


Figure A. 1: P_{JR} at Different refreshing periods

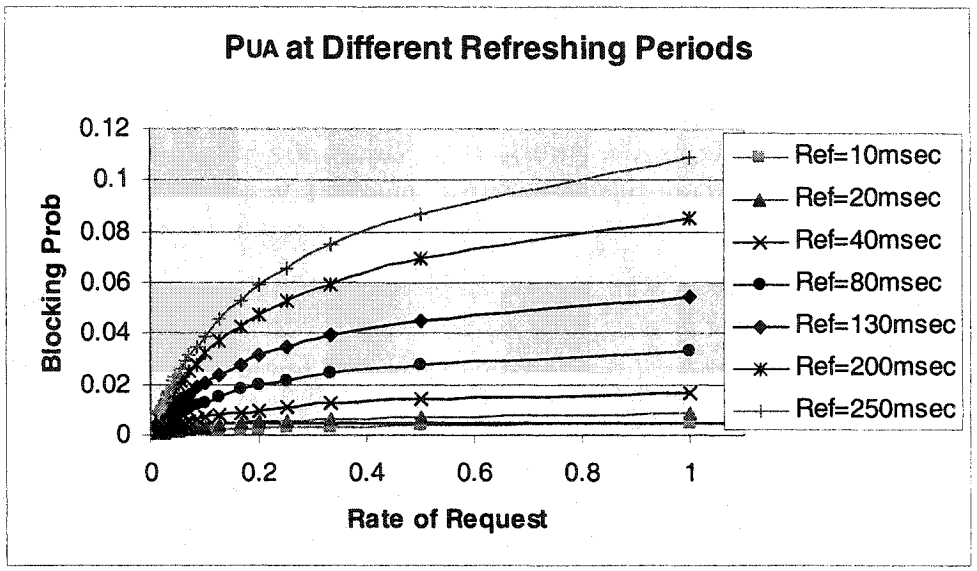


Figure A. 2: P_{UA} at different refreshing periods

Figure A.2 and Figure A.3 show the effect of the refreshing period on both P_{UR} and P_{UA}. Clearly, slow refreshing periods cause higher P_{UR} and P_{UA} because more errors will be introduced at routing tables due to the slow report of routing changes.

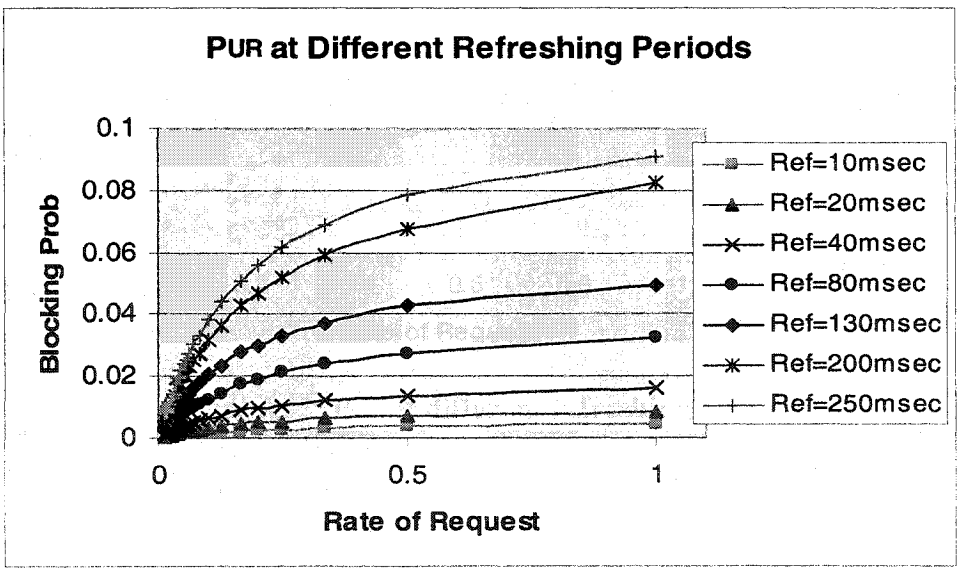


Figure A. 3: P_{UR} at different refreshing periods

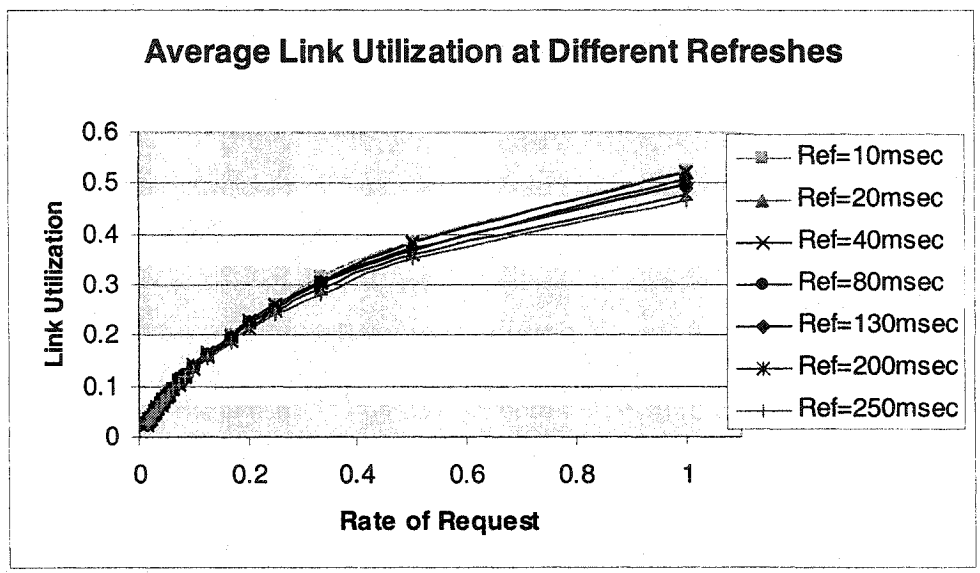


Figure A. 4: Link Utilization at different refreshing periods

Figure A.4 shows that the shorter the refreshing period is, the higher the link utilization.

In fact, the ARPANET architecture showed relatively lower average link utilization compared to the simple network architecture. The bigger the network is, the more likely the requested light path will not be successfully established because the light path has to cross more ASs.

B – Effect of the intra-domain blocking for the ARPANET

Figure B.1, Figure B.2, Figure B.3 and Figure B.4 show the effect of Intra domain blocking probability on P_{JR}, P_{UA}, P_{UR} and link utilization, respectively.

The number of wavelengths used in this simulation is 100 for granularity purposes, and the refresh period is fixed at 20msec whereas the mean of the average life time of the requested wavelength is 4000msec.

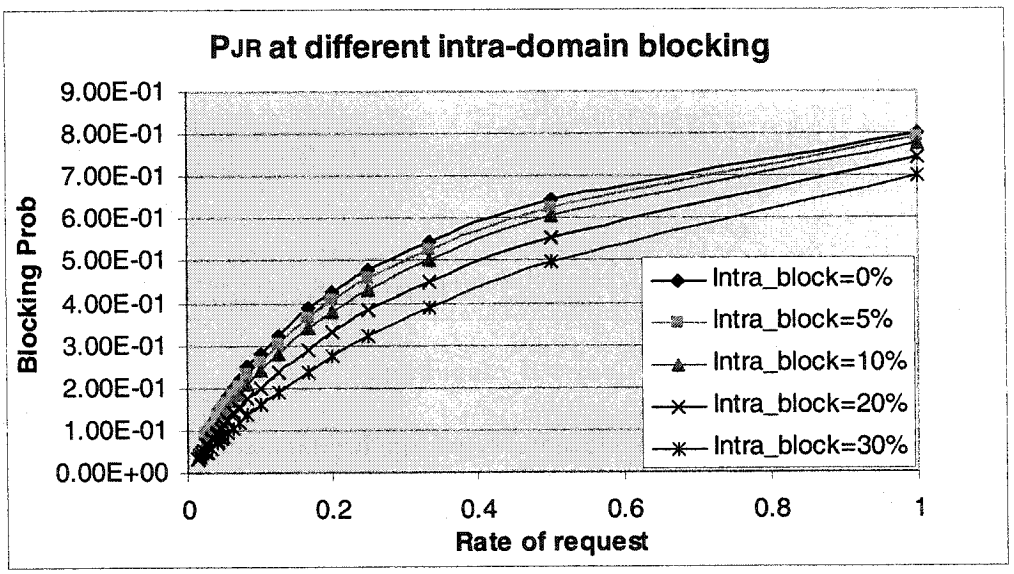


Figure B. 1: P_{JR} at 0, 5, 10, 20, 30% Intra domain blocking probability

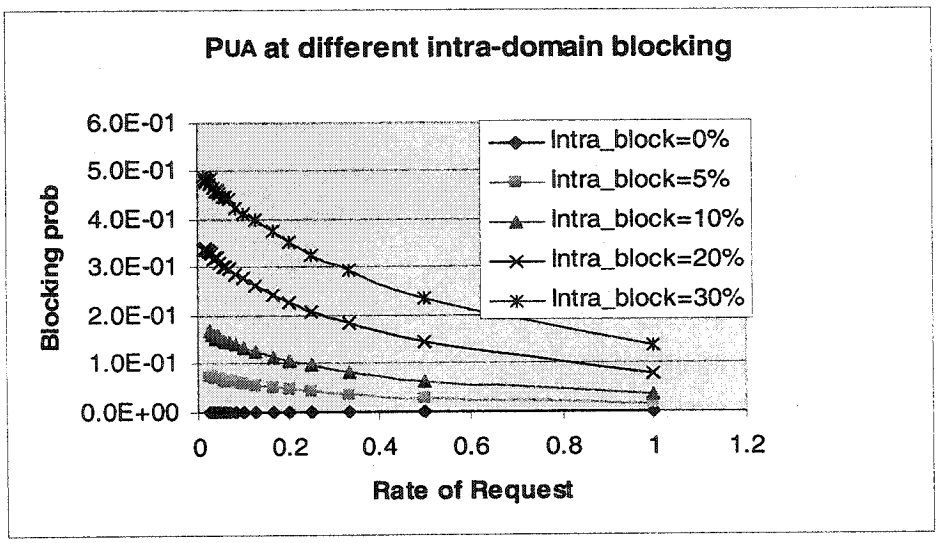


Figure B. 2: P_{UA} at 0, 5, 10, 20, 30% Intra domain blocking probability

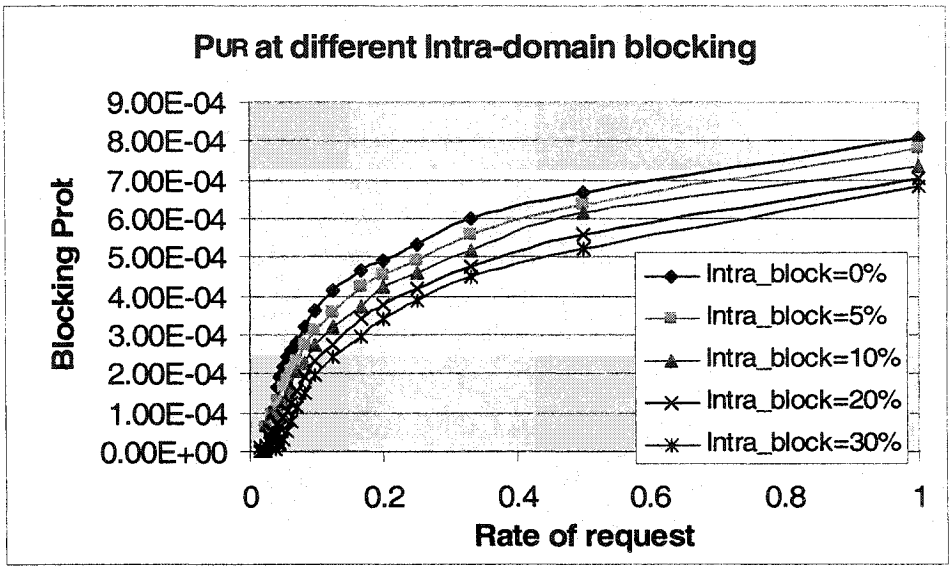


Figure B. 3: P_{UR} at 0, 5, 10, 20, 30% Intra domain blocking probability

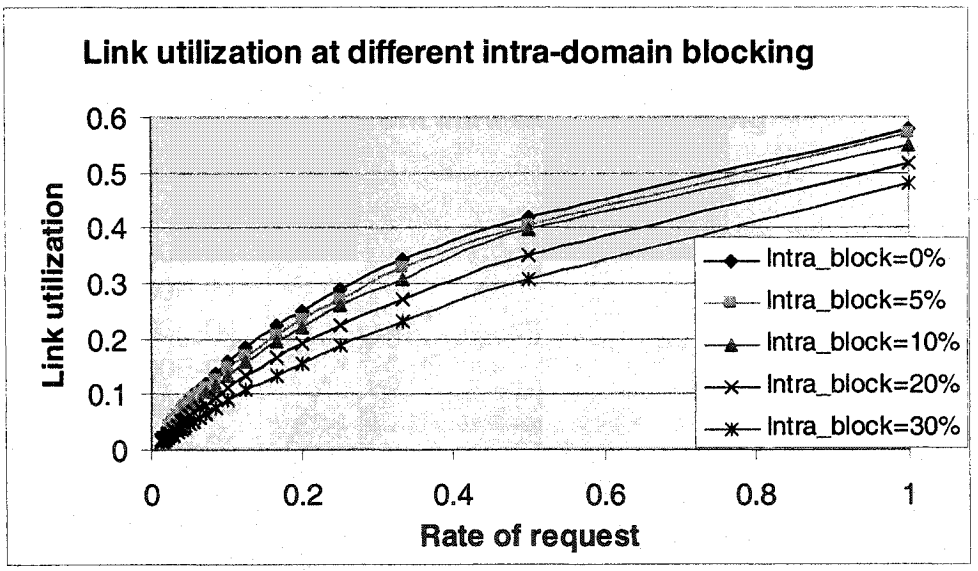


Figure B. 4: Link utilization at 0, 5, 10, 20, 30% Intra domain blocking probability

C – Threshold change percent for the ARPANET

Figure C.1, Figure C.2, Figure C.3 and Figure C.4 show the effect of different threshold values on P_{JR}, P_{UA}, P_{UR} and average link utilization, respectively. In this simulation, we choose the number of wavelengths to be 64 and the average life time of the requested wavelengths to be 400msec. The rate of request was changed from 0 to 1 request/msec, and it is assumed that the refreshing period is very small compared to the lifetime of the requested wavelength.

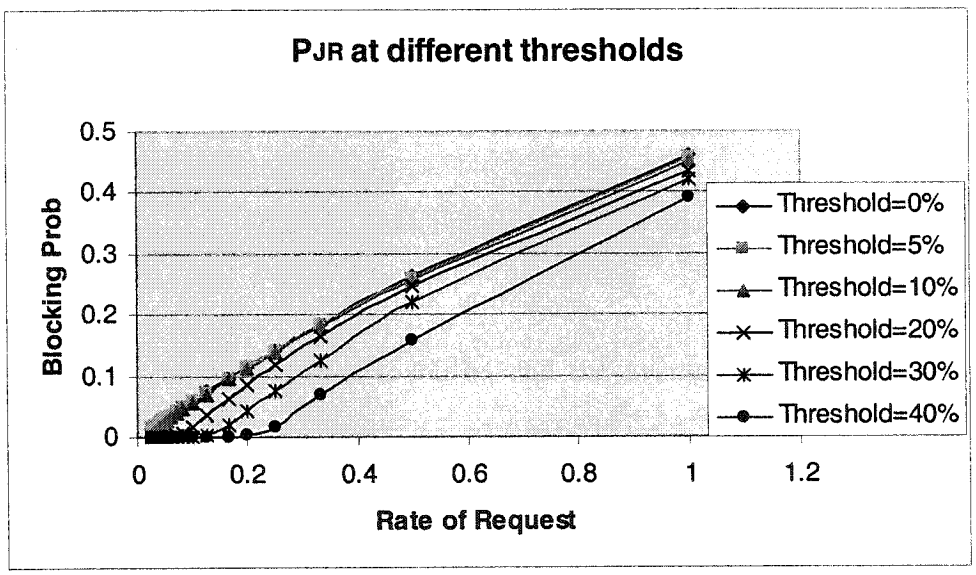


Figure C. 1: The effect of different threshold values on P_{JR}

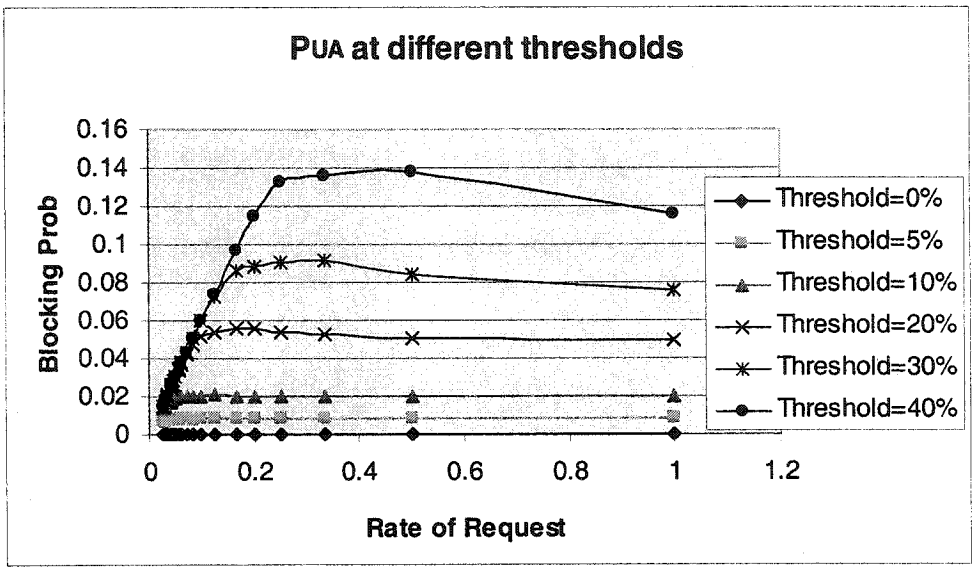


Figure C. 2: The effect of different threshold values on P_{UA}

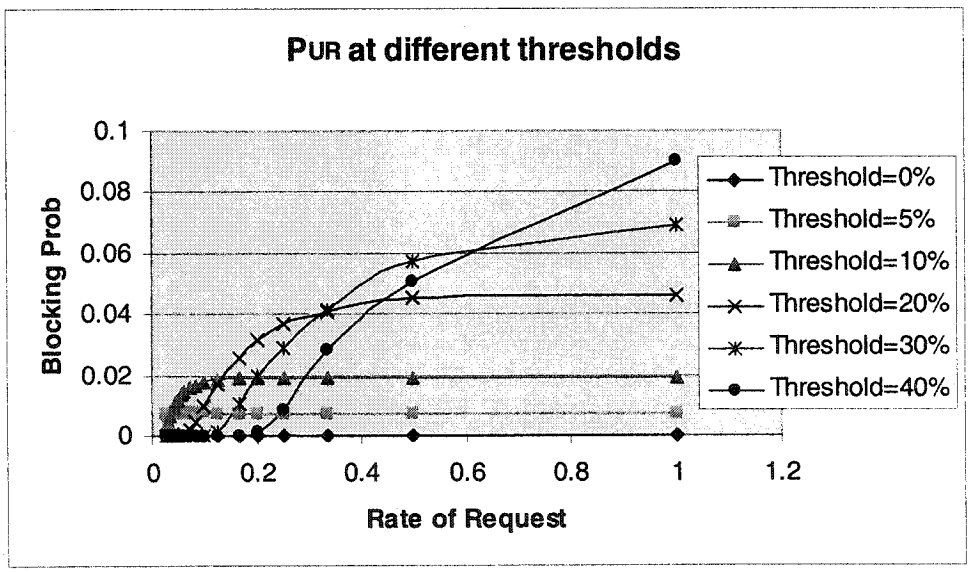


Figure C. 3: The effect of different threshold values on P_{UR}

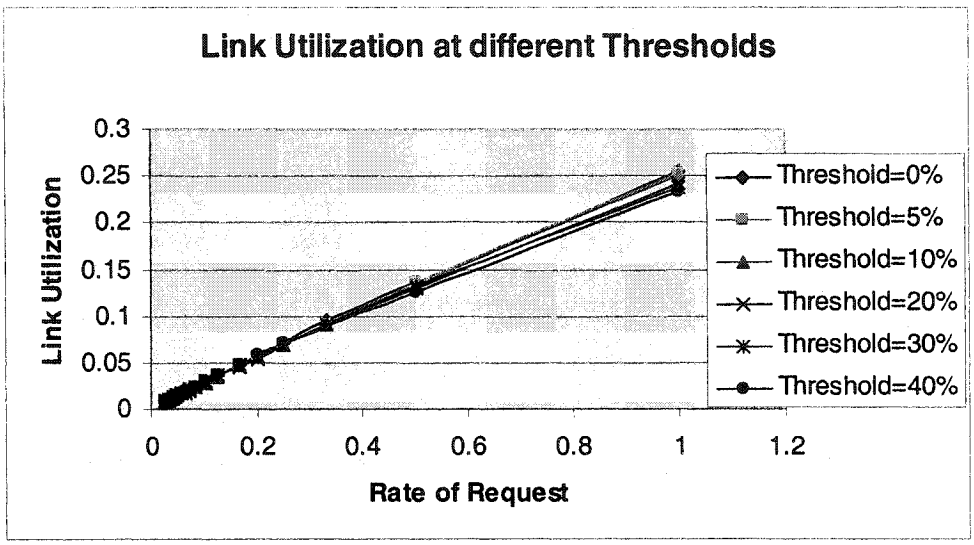


Figure C. 4: Link utilization at different thresholds

D – Threshold change percent for the ring network

Figure D.1, Figure D.2, Figure D.3 and Figure D.4 show the effect of different threshold values on P_{JR}, P_{UA}, P_{UR} and link utilization, respectively.

The simulation on the ring network is performed using 64 wavelengths.

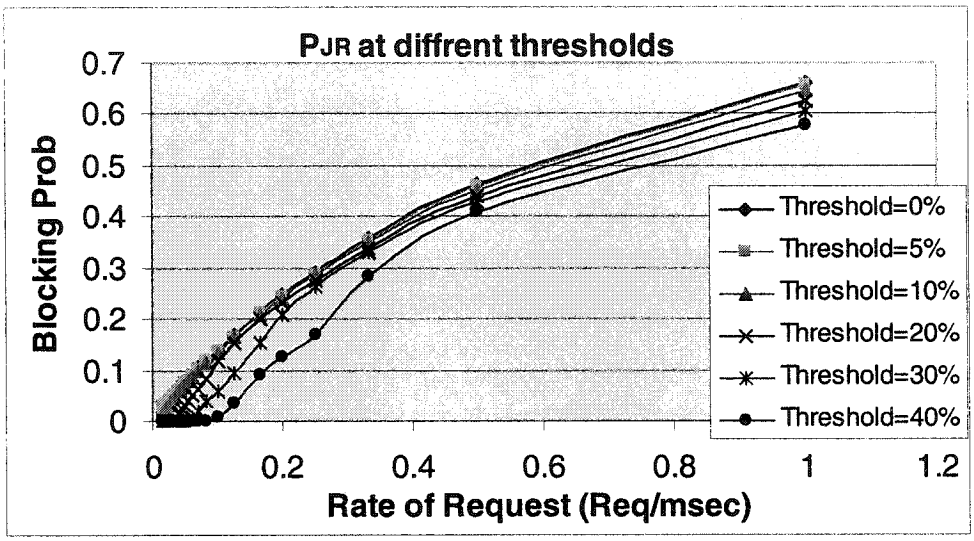


Figure D. 1: The effect of different threshold values on P_{JR}

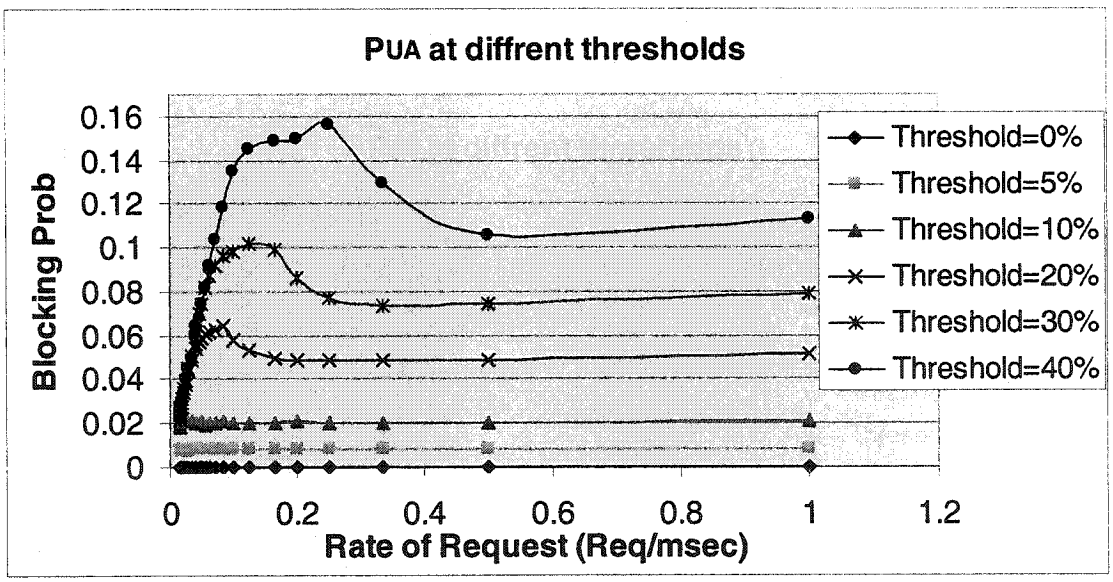


Figure D. 2: The effect of different threshold values on P_{UA}

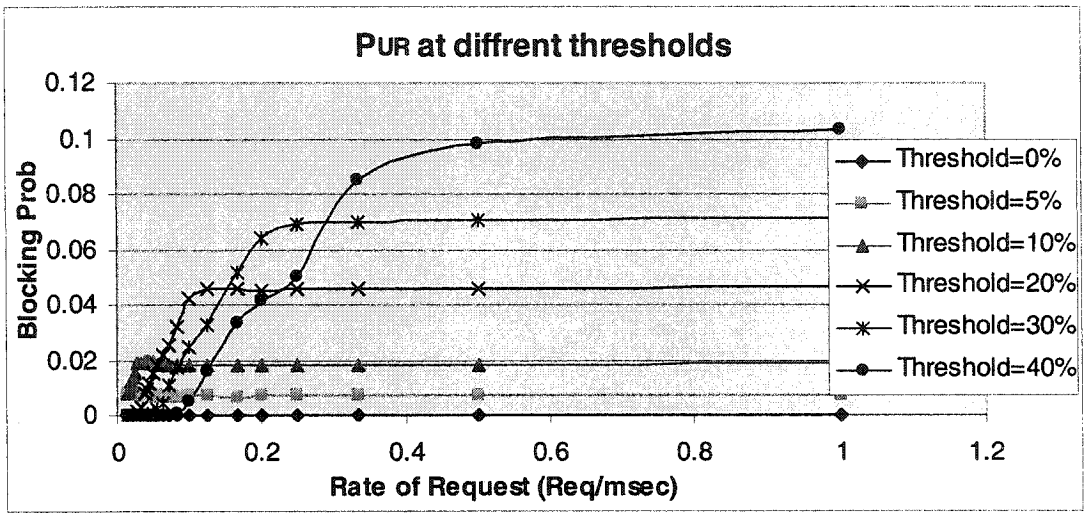


Figure D. 3: The effect of different threshold values on P_{UR}

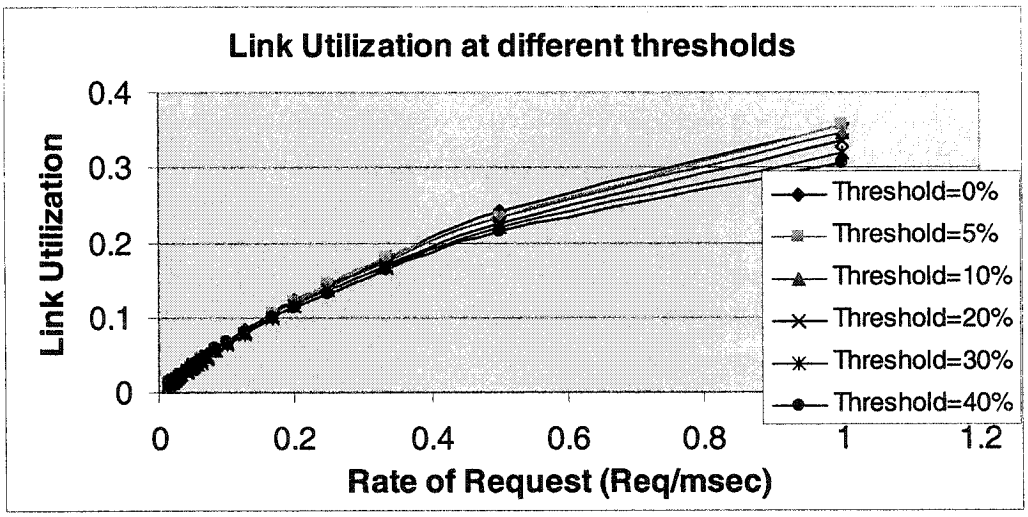


Figure D. 4: Link utilization at different thresholds

E – Threshold change percent for the European network

Figure E.1, Figure E.2, Figure E.3 and Figure E.4 show the effect of different threshold values on P_{JR}, P_{UA}, P_{UR}, and link utilization, respectively. We used the threshold based technique for advertising for the same reasons we explained earlier.

The number of the wavelengths used here is 64.

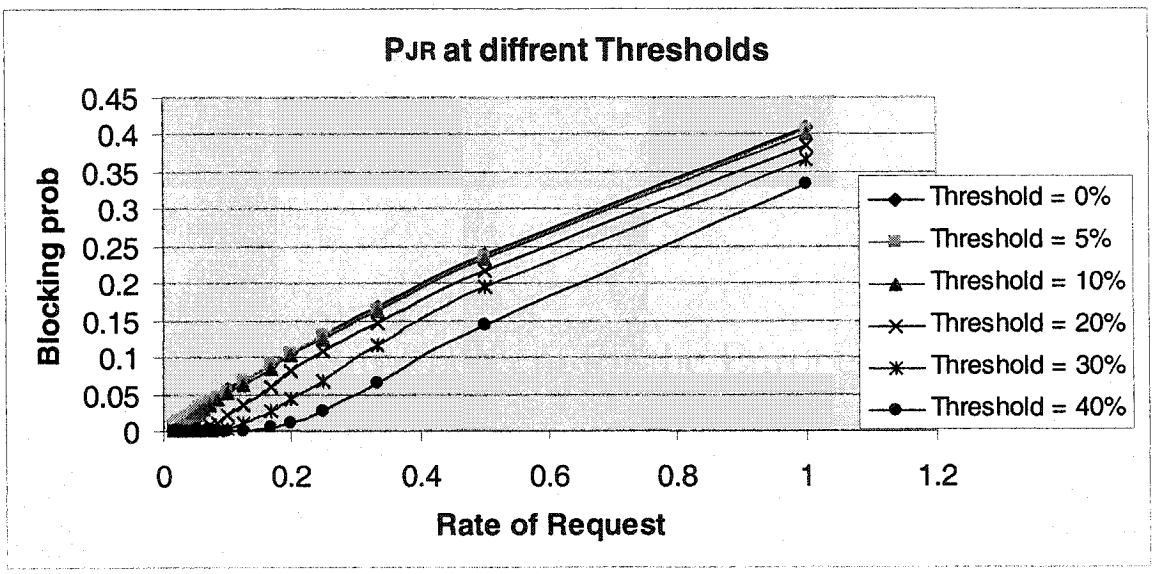


Figure E. 1: P_{JR} at different thresholds

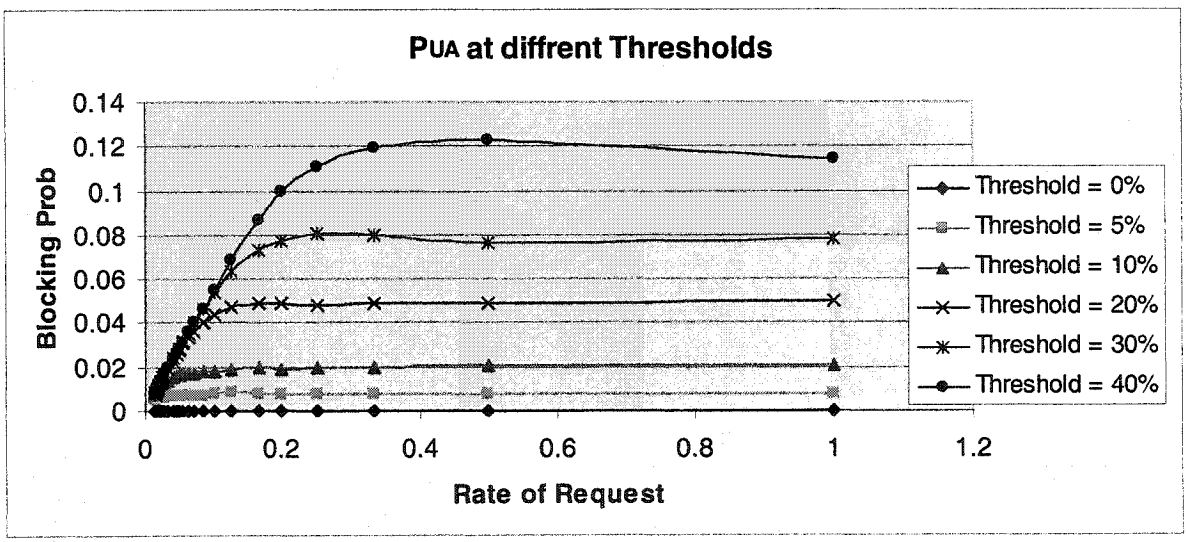


Figure E. 2: P_{UA} at different thresholds

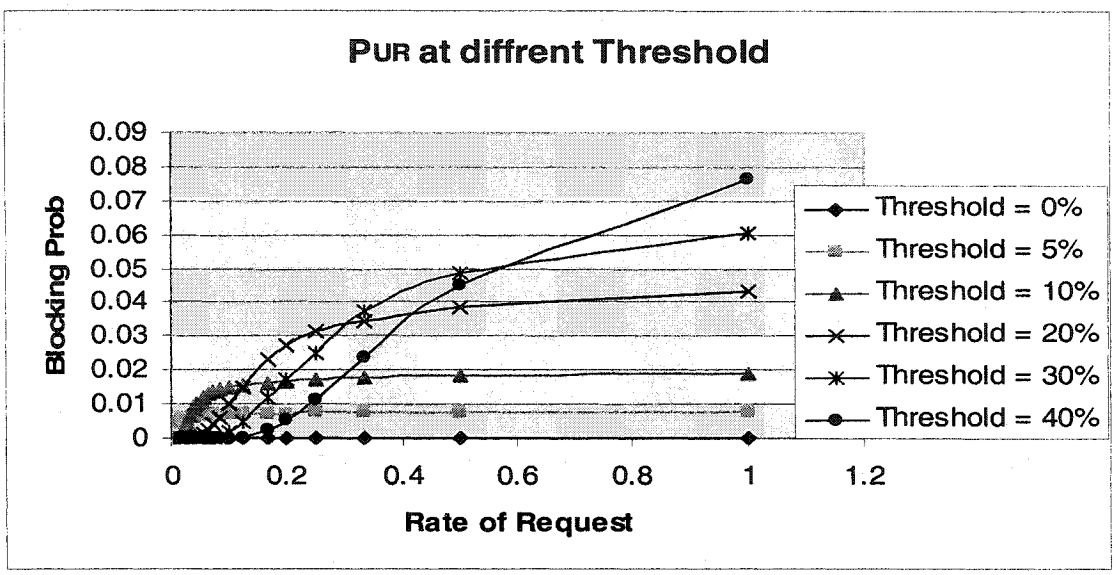


Figure E. 3: P_{UR} at different thresholds

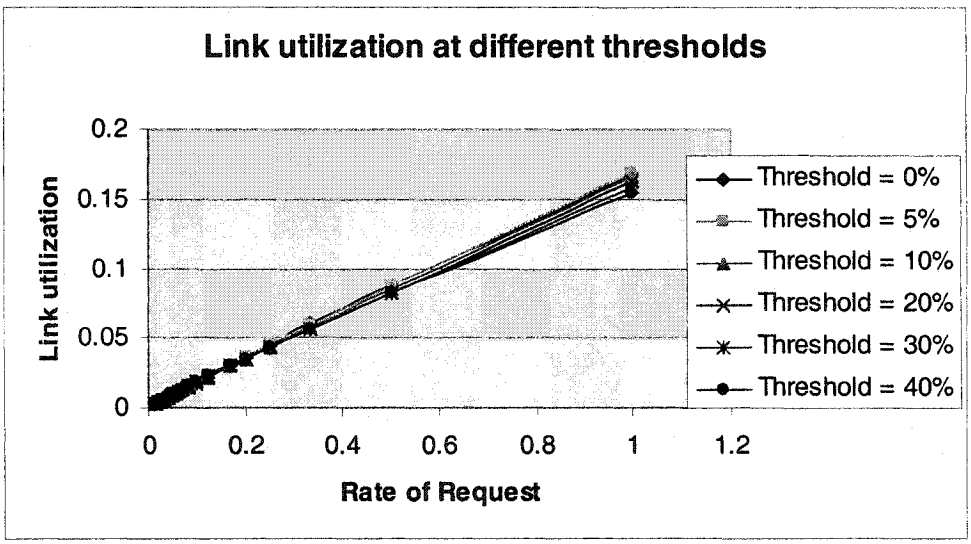


Figure E. 4: Link utilization at different thresholds

F – Threshold change percent for the star network

Figure F.1, Figure F.2 and Figure F.3 show the effect of different threshold values on P_{JR}, P_{UA}, and P_{UR}, respectively.

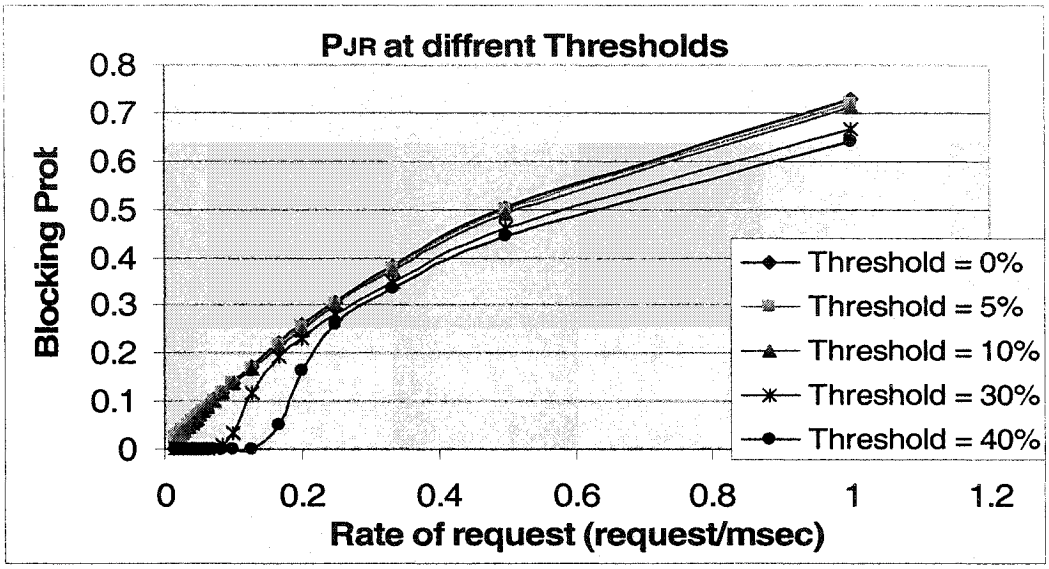


Figure F. 1: P_{JR} at different thresholds

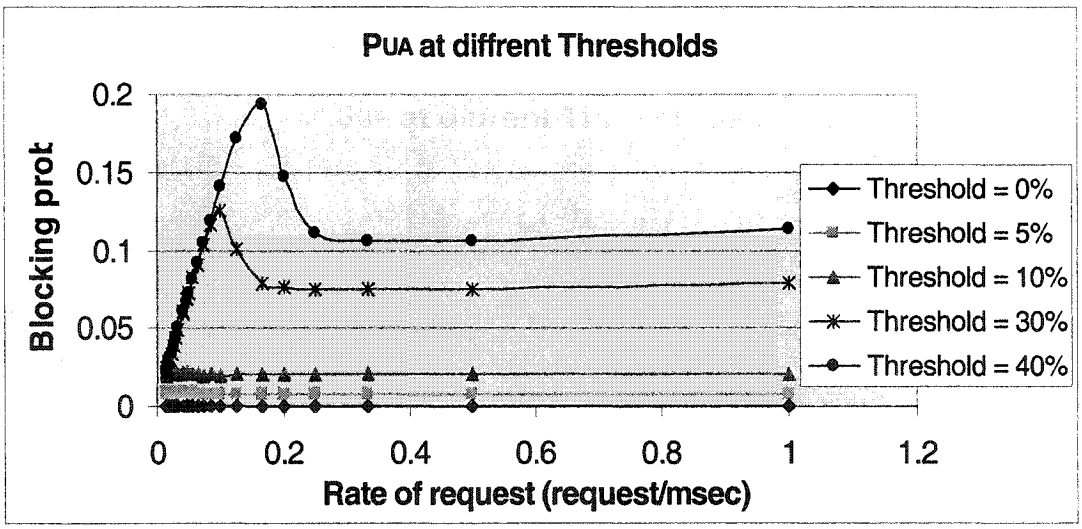


Figure F. 2: P_{UA} at different thresholds

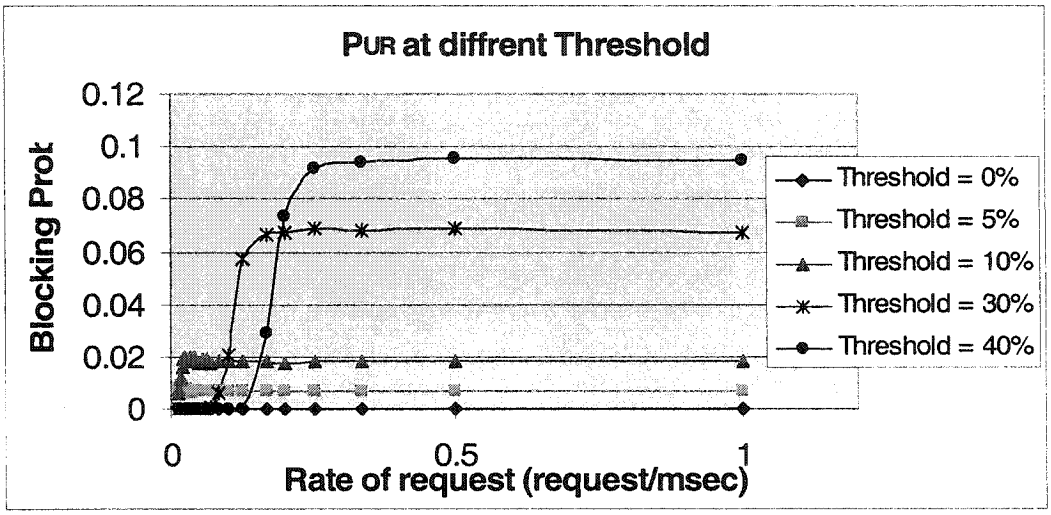


Figure F. 3: P_{UR} at different thresholds

G – Acronyms

ABR: Area Border Router
ASBR: Autonomous System Border Router
AS: Autonomous Systems
ATM: Asynchronous Transfer Mode
ARPANET : Advanced Research Projects Agency Network
BGP: Border Gateway Protocol
CLNP: Connection-less Network Protocol
CR-LDP: Constraint-based Routing Label Distribution Protocol
DV: Distance Vector protocol
DWDM: Dense Wavelength Division Multiplexing
EGP: External Gateway Protocol
GMPLS: Generalized Multi Protocol Label Switching
IGP: Internal Gateway Protocol
IS-IS: Intermediate System- Intermediate System
LS: Link State protocol
LSP: Link State Packets
LSA: Link State Advertisement
NLRI: Network Layer Reachability Information
OSPF: Open Shortest Path First
OSI: Open Systems Interconnection
OPNET: OPtimum NETwork performance
PDU: Protocol Data Unit
RIP: Routing Information Protocol
RSVP-TE: Resource reservation Protocol- Traffic Engineering
SONET: Synchronous Optical NETwork
SDH: Synchronous Digital Hierarchy
TCP/IP: Transmission Connection Protocol /Internet Protocols
TDM: Time division Multiplexing
QoS: Quality of Service
P_{JR}: Probability of Justified Refusal
P_{UA}: Probability of Unjustified Acceptance
P_{UR}: Probability of Unjustified Refusal
P_{JA}: Probability of Justified Acceptance
P: Ideal blocking
PV: Path Vector