Interdomain Routing In Optical Networks

Mark Joseph Francisco, Stephen Simpson, Lambros Pezoulas,

Changcheng Huang, Ioannis Lambadaris

Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6

Dept. Systems and Computer Engineering, Advanced Optical Networks Laboratory

Bill St.Arnaud

Canarie Inc., 110 O'Connor Street, Ottawa, ON K1P 1H1

ABSTRACT

This paper has outlined a method (called OBGP) of extending BGP to support lightpath setup and management across an optical network. The development of OBGP has been discussed by reviewing current BGP behavior and design requirements for OBGP. An implementation of OBGP using simulation tools has been presented, along with initial test results, which have shown that a seamless migration from BGP to OBGP is possible.

Keywords: OBGP, BGP, Wavelength Routing, Interdomain, WDM, Lightpath, Optical Cross Connect (OXC), NNI

1. INTRODUCTION

The demand for bandwidth in today's Internet infrastructure has led to the deployment of a large number of high capacity fiber optic networks. To compliment these high bandwidth optical networks, intelligent methods of provisioning and control are required to enable greater flexibility and to reduce operation cost. A convergence is taking place between IP-based routers and multi-wavelength optical equipment to create an intelligent end-to-end integrated network.

Traditionally, as stated in [1], end-to-end circuit connections have been set up via network management systems (NMSs), which issue commands (usually under the control of a human operator) to the various network elements involved in the circuit via an equipment vendor's element management system (EMS). So far this has been the carrier's responsibility and the lightpath setup and configuration have been based on a centralized model. By moving this responsibility from the carrier to the customer and using a distributed approach, we can simplify and easily allow customers to better manage their optical wavelengths. This distributed approach gives more control to edge customer.

Optical networks are *circuit switched networks* in nature due to their high traffic capacity and relatively slow switching speed. The Internet on the other hand is a connectionless packet switched network. Methods of integrating optical networks with the Internet are becoming the key issue of optical network architecture. While there are fewer issues concerning the data plane integration of optical networks and the Internet, there are significant debates on how the control planes should be integrated. The control planes in IP and optical networks can be loosely or tightly coupled. Three interconnection models are possible for IP over optical networks [2]: *overlay model, peer model and interdomain model*.

The *overlay model* is similar to classical IP over ATM model. Under this model, optical networks are treated as a transparent sublayer and the IP layer is independent of the corresponding protocols in optical networks. This approach has been widely studied and is being standardized in OIF (Optical Interworking Forum) and IETF (see [2] for example). While this approach may lead to early deployment of optical networks, it is still a loosely coupled solution. With the development of optical switches it is foreseeable that packet optical switches may be possible in future. Therefore it is important to develop an architecture that can more closely integrate optical networks with the Internet.

Under the *peer model*, optical cross-connects (OXCs) are treated just like any other router in terms of the control plane. This will allow a seamless interconnection of IP and optical networks. As pointed out in [2], the drawback of this approach is that it requires routing information specific to optical networks to be known to routers and may overburden router networks.

Different from the peer model the *interdomain model* allows separate instances of the IP and optical domains. Information from optical networks is selectively passed to the router networks to allow interworking through interdomain routing protocols. By choosing the information to be passed carefully we can avoid overburdening router networks. Therefore, the interdomain model combines the best of the peer and overlay models. Unfortunately we have not seen many discussions in this direction. This paper is our first effort towards the interdomain model.

In order to move in this direction we will show how the Border Gateway Protocol (BGP) can be extended to allow an edge customer to set up a lightpath to another site across multiple AS domains. This new flavor of the traditional BGP routing protocol will be called Optical Border Gateway Protocol (OBGP). *The goal of OBGP is to provide an edge network customer with a control method for establishing a lightpath through an optical network.*

This paper will describe the development and implementation of OBGP by outlining the following areas:

- The design requirements for creating OBGP
- The necessary extensions to BGP to create OBGP
- The implementation of OBGP into a standard BGP framework
- The benefits of using the OBGP approach, as a way for setting up lightpaths

Finally, conclusions and recommendations for future work are presented.

2. OPTICAL BGP ARCHITECTURE

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

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

BGP is a routing protocol while the setup and control of lightpaths are the functions of a signaling protocol. We will show in this paper that there are advantages to integrate routing and signaling into one protocol. Both signaling and routing protocols are jointly used to carry information through the network. Signaling typically requires routing information, for example source routing. For failure protection and restoration, the interactions between signaling and routing are critical. Currently there exist various routing protocols and signaling protocols that need to interwork in order to provide an end-to-end solution. In some cases this interwork of multiple protocols is complicated and results in poor scalability. It also requires that every router along the path implement the required routing and signaling protocols. Therefore, it may be a better approach to integrate more closely the functions of routing and signaling.

By combining routing with signaling in OBGP we provide a lightweight solution. Signaling functions can leverage some features of existing routing functions. For example, both routing and signaling messages can share a common transport connection (i.e. same TCP connection is used to exchange both routing and signaling messages). Thus no extra overhead is required for setting up a separate communication channel for signaling messages. OBGP can also access directly the database of BGP and therefore eliminate the synchronization requirement if two separate databases are used.

Another reason, which makes BGP ideal for signaling, is the AS_PATH attribute of a route. The AS_PATH gives an end-to-end view, which is necessary for setting up a lightpath. The AS_PATH also allows for better distribution of OBGP messages in the network. For example not all OBGP speakers in the network need to see this message if they are not in the specific AS_PATH thus reducing the amount of signaling traffic in the network.

OBGP provides an interdomain solution that integrates heterogeneous domains into an end-to-end optical network. Different domains still can use their own intradomain solutions. Therefore OBGP can coexist with most of the existing solutions.

Finally, as mentioned in the introduction the distributed nature of OBGP moves the responsibility of setting up a lightpath from carriers to edge customers. This allows the customer to set up a lightpath faster without having to wait for carriers to provision their connections which in most cases is a complex and time consuming task.

2.1 OBGP Requirements and Design

BGP does not contain any functionality or attributes for maintaining information on lightpaths across a network. However, BGP does contain *complete autonomous system (AS) path information* to reach a particular network, which makes it unique when compared to other routing protocols. This full path visibility enables routing policy decisions to be made as well as preventing routing loops. Having full path visibility is a useful feature of BGP and can be leveraged for setting up a lightpath from one AS to another [3].

There are four types of messages specified for BGP [4]: OPEN, UPDATE, NOTIFICATION, and KEEP-ALIVE. An UPDATE message is used to provide routing updates within a network. These messages allow BGP routers to maintain consistent information regarding network reachability. UPDATE messages can add or withdraw routes from the Routing Information Base depending on the content of the message. UPDATE messages in BGP are the most relevant to the design of OBGP, since they will convey any new information from router to router. Specifically, UPDATE messages are comprised of a standard BGP header along with additional fields for path attributes.

The basic high-level requirements for OBGP to work as a *lightpath setup* and *control* protocol are as follows:

- Provide a set of attributes to maintain lightpath information
- Carry lightpath reservation requests and responses between OBGP speaking devices
- Propagate up to date information on the status of resources in the network

Making minimum changes to BGP while meeting these requirements is a very important consideration. Leveraging current BGP properties is key to meeting these requirements. A typical OBGP Router consists of three components, illustrated in Figure 1. The first component is a Border Gateway Router that can be found at the gateway of an AS. This router performs the BGP Peering and sits at the control plane for OBGP.

The Switch Control Server (SCS) is the control interface between the router and the OXC. This interface maintains connection state for the OBGP process in the router. The Optical Cross Connect (OXC) can be any type of cross connects, but for this paper, we assume this OXC has the capability of wavelength conversion. OXC without wavelength conversion will be discussed in the future.

Wavelength converters, as stated in [5], are devices, which take the data modulated on an input wavelength and transfer it to a different output wavelength. Wavelength converters improve network-blocking performance since it is not required to establish a common wavelength across all links in the route.

Connectivity information in BGP is propagated through UPDATE messages. Both new routes and withdrawn routes are propagated through these messages to maintain up to date information on network reachability throughout the network. Lightpath setup must be performed by signaling between OBGP speaking devices. A message must be able to signal to the OBGP devices to setup, as well as propagate up to date information on the status of resources throughout the network. A new message format can be created and used to perform this function, but to fully leverage BGP properties, we explore the use of the UPDATE message in this paper. The BGP UPDATE messages can be leveraged for lightpath setup in two ways, first to propagate up to date information on the status of resources throughout the network and secondly, to carry lightpath setup requests and responses between OBGP speaking devices.

Using BGP to transmit lightpath setup requests and responses can be accomplished by modifying UPDATE messages to contain lightpath setup information in addition to normal reachability information. Propagating up to date lightpath reservation information to the entire network following a new lightpath setup can be accomplished by triggering UPDATE messages into the network once the setup is complete.
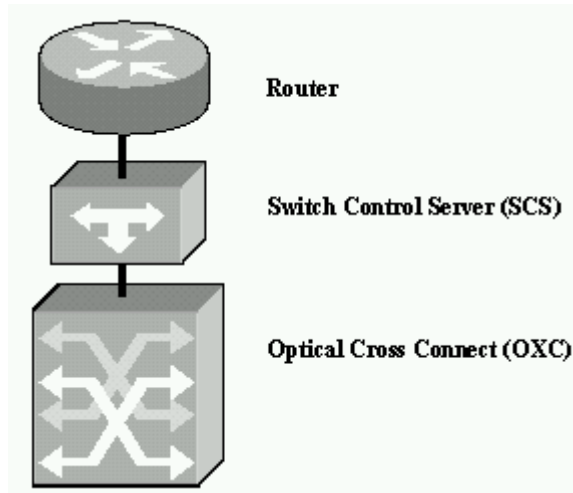
Figure 1. Diagram of OBGP Router

OBGP lightpath setup occurs when a source from an AS wants to get a direct path to a destination AS. Since BGP does not have this functionality extensions must be made. Proposed extensions to BGP include *wavelength availability attributes* and *lightpath setup phase tags,* which are carried by UPDATE messages.

The wavelength availability attributes accompany the AS Path Segment List to indicate the current wavelength availability status of a link between two neighboring domains. This attribute can have one of two values:

> **A**: One or more wavelengths are available for the link
> **U**: No wavelength is available for the link

Adaptive Routing can be supported using this wavelength availability attribute [6]. When an OBGP speaker receives an UPDATE message, it will check the wavelength availability attributes associated with the AS Path Segment List. If it finds no wavelength is available on a link, the OBGP decision process will reject the path. Otherwise, the OBGP decision process will decide whether to choose the path as a normal BGP decision process. When a source decides to setup a lightpath to a specific destination, it will find the AS Path from its OBGP database. This AS Path guarantees at least one wavelength is available on each link. This can significantly reduce the blocking probability although it may not be able to eliminate blocking entirely. Blocking can still happen if an optical network has wavelength continuity requirement or when competition occurs. Although broadcasting more information about a particular link (such as detailed information about each wavelength) can help an OBGP speaker to make a better decision on selecting a path and a channel, it is not a scalable solution due to the large amount of information to be flooded through the whole network. By using the wavelength availability attributes, updates to the rest of the network can be made only when the wavelengths of a link are fully utilized, or when a link has one or more wavelengths freed from a fully utilized state. This minimizes the information distributed, increases the convergence time and allows for a scalable solution in an interdomain network.

The lightpath setup phase tags indicate the current state of a lightpath setup process. This tag can have one of the following values:

> **N:** No reservation is occurring.
> **R:** Lightpath reservation request is in progress.
> **C:** Confirmation/Setup in progress.
> **W:** Withdrawal of a reservation in progress.

The lightpath setup phase tag can be encoded into the NLRI field of an UPDATE message with a new address family identifier called OBGP_Active thereafter [7]. In addition to this phase tag, the AS Path from the source to the destination will also be encoded into the OBGP_Active address to identify the ASs that are on the signaling path. Specifically, to request a lightpath through a BGP network, the source (or sender) of the request will need to propagate a new OBGP UPDATE

message with the OBGP_Active address to the destination and the destination will send another OBGP UPDATE message back as confirmation.

It is important to note that signaling in this paper is a two-phase process. When introducing competition and various failure scenarios, more phases may be needed to ensure safe creation of lightpaths.

2.2 OBGP Protocol

The new attribute and address family discussed in last section work in conjunction for lightpath setup. Figure 2 illustrates a basic network where node S wants to setup a lightpath to node D. Through regular BGP peering, node D has advertised its presence to the network and each node has AS Path information and wavelength status information to reach node D. Suppose that for node S to get to node D, traffic must travel through X and Y. This is a path discovered through BGP and can be found in the BGP Routing Information Base. Each node in the path also has wavelength availability status information, which is denoted by *A* (Available). To reserve the lightpath, node S initiates with an OBGP UPDATE message, as shown in Figure 2, requesting the reservation. The OBGP UPDATE message, shown in the box below the node information of S, contains AS Path information with the reservation request for the corresponding link. At each node, the availability is checked and the wavelength availability attribute is changed to *U* (Unavailable, if no wavelength is available), as illustrated in Figure 2 and Figure 3.
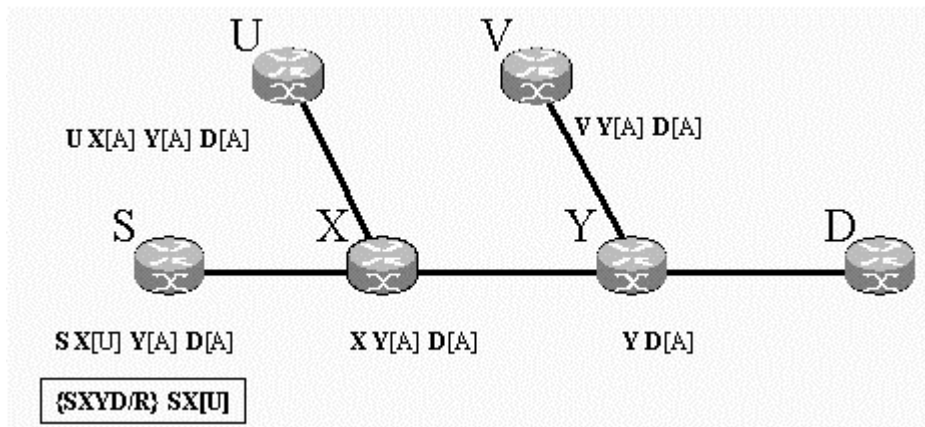

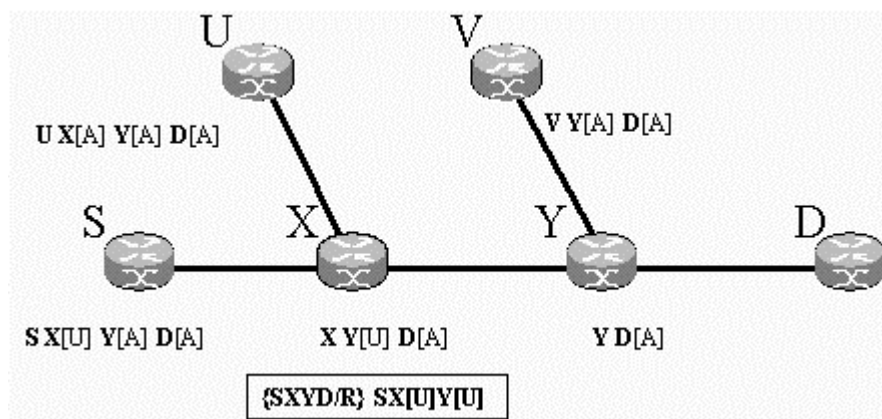**Figure 2. Network with Wavelength Status Attributes with Reservation initiated from S**

Figure 3. Network with a Lightpath Reservation in Progress from S to D

If the entire lightpath is granted, OBGP will signal a new UPDATE message back to the Source, with the lightpath setup phase tag set to Confirmation (**C**). In Figure 4, the lightpath reservation has completed and D sends a new UPDATE message to confirm the lightpath, which is shown in the box. As this message propagates through the network, the corresponding

nodes update their wavelength availability status attributes, as shown in Figure 5. Note that nodes not on the AS path also get updated.
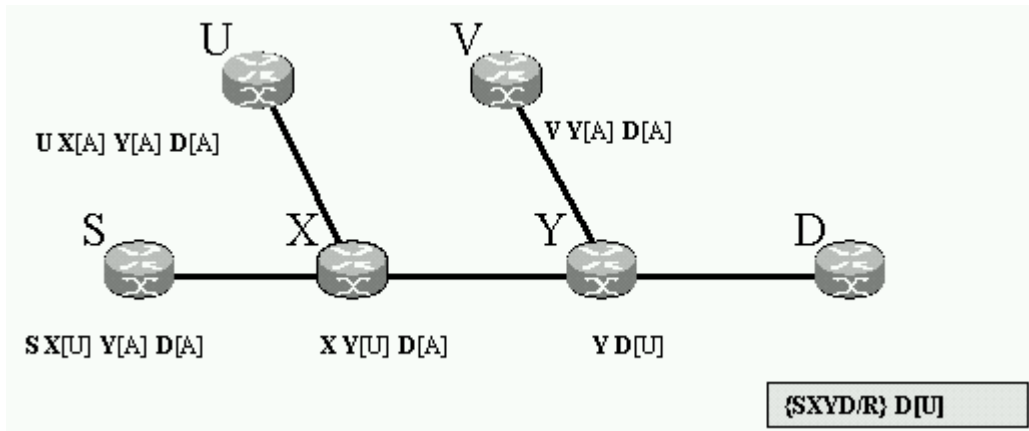

Figure 4. Network with a Completed Lightpath Reservation and a Confirmation Initiation
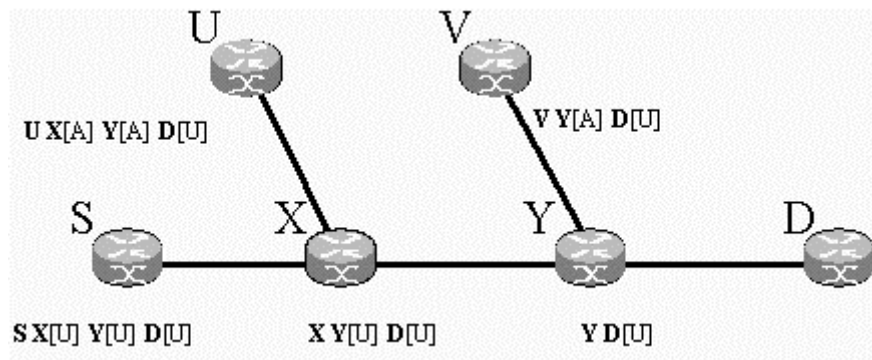

Figure 5. Network with a Complete Confirmed Lightpath from S to D

If the reservation is denied at any stage, a new UPDATE message is sent back to the source with the wavelength availability attributes set to A (Available) and phase tag set to W (Withdraw). The UPDATE is sent back through the same path to the source deconstructing the current reservation.

Competition amongst lightpath reservations must be addressed. Assume both node S and node V send reservation request to node D in Figure 2. In this example, node V is physically closer to node D, and acquires the reservation to node D first. Node S's reservation is propagated through node X and then to node Y, where node Y cannot grant the next reservation since it is being used by the reserved lightpath from node V to node D. Node S's reservation must be withdrawn at this point. If there is a failure while reserving a lightpath, a similar process occurs as the competition case. OBGP will begin to withdraw the current lightpath. Future research must be completed to fully explore this scenario.

In both the reservation and withdrawal scenarios mentioned so far, each OBGP node must be aware of its position within the AS Path in order to process correctly the OBGP UPDATE messages. There are three types of OBGP nodes present in any optical network: Source, Intermediate, and Destination. A node's classification will change according to its relative position in the lightpath segment that is being reserved.

Lightpath Segment example: [Source_AS   Intermediate_AS1   Intermediate_AS2 … Destination_AS]

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

Four scenarios have been identified when a BGP node receives an OBGP UPDATE message. According to the lightpath setup phase tag value, the BGP process will behave differently. The different scenarios are described in Table 1 below.

Table 1 displays the behavior necessary for OBGP to function properly. From this table, an algorithm can be developed to allow OBGP to meet its lightpath setup and management requirements. The algorithm must be distributed across all BGP nodes wishing to participate in lightpath setup. It is important to note that OBGP will only use BGP as a basis for implementation. OBGP behavior will exist on a concurrent plane with traditional BGP and will be distinguished by the value of the OBGP_ACTIVE tag.

	Scenario Conditions	Action Taken
1	Intermediate node receives an OBGP UPDATE message with phase tag set to "R"	If this node is not in the path segment discard this message. Otherwise, change the lightpath setup attributes accordingly and send the OBGP UPDATE message to the next hop.
2	Destination receives an OBGP UPDATE message with phase tag set to "R"	Generate a new OBGP UPDATE message with phase tag set to "C". The AS Path is the same as the received OBGP UPDATE message. Propagate the new OBGP UPDATE message back to the peer that originated the request.
3	Intermediate node receives an OBGP UPDATE message with phase tag set to "C"	Change the lightpath setup attributes accordingly and advertise the OBGP UPDATE message to its peers.
4	Source receives an OBGP UPDATE message with phase tag set to "C"	Lightpath setup completed. Ready to send information.

Table 1: OBGP Scenario Behavior

3. TESTING OBGP

3.1 OBGP Test Network

OBGP has been implemented using OPNET network simulation software. There are several reasons for using a network simulation package such as OPNET rather than implementing the protocol in hardware. The main reason is that OPNET provides a flexible development environment. In addition, protocols such as BGP are clearly defined in OPNET using state based methods

The test network was designed to cover the major points of the OBGP algorithm. There are two networks in the test network: the electrical network and the optical network. The electrical network is used for OBGP as the control plane while the optical network is treated as the actual data plane of the overall optical network The electrical network performs its regular functions of packet switching and forwarding, but primarily, exchanging information about BGP and discovering the optimal paths within a network. The components necessary for implementing an OBGP test network included:

- 5 BGP-enabled IP routers with standard interfaces
- 2 BGP-enabled IP router with an SCS (switch control server) interface into the router and external interface to the OXC
- 2 OXCs (optical cross connect)

The OPNET test network is depicted in Fig 6. The electrical network consists of seven gateway routers. These routers are BGP and OBGP enabled. The network through AS 3561, AS 20000, AS 1239 and AS 911 is the core of the OBGP reservation. Periphery nodes at AS 20000, AS 1239 and AS 911 have been added to test the OBGP algorithm.

At AS 20000 and AS 1239 the control for the OXCs has been enabled within their respective domain. The optical interfaces from AS 3561 and AS 911 generate random traffic from one to another

To simulate a lightpath reservation request from AS3561 to the destination AS911, a request was simulated. The lightpath reservation was made to AS911, which spanned two intermediate nodes, AS20000 and AS1239. The sequence of a normal lightpath reservation is illustrated in Figure 6, with the description below:

0 AS 3561 sends the OBGP UPDATE Message to AS 20000.
1 AS 20000 receives the OBGP UPDATE Message, saves it in its Local Routing Information Base and sends it to the next AS in the AS Path List.
2 AS 1239 receives the OBGP UPDATE Message, saves it in its Local Routing Information Base and sends it to the next hop.
3 AS 911 receives the OBGP UPDATE Message, saves it and grants the wavelength reservation. An OBGP UPDATE is sent back to the advertising port (AS 1239).
4 AS 1239 receives the message and advertises it to its peer.
5 AS 20000 receives the message and advertises it to its peer.
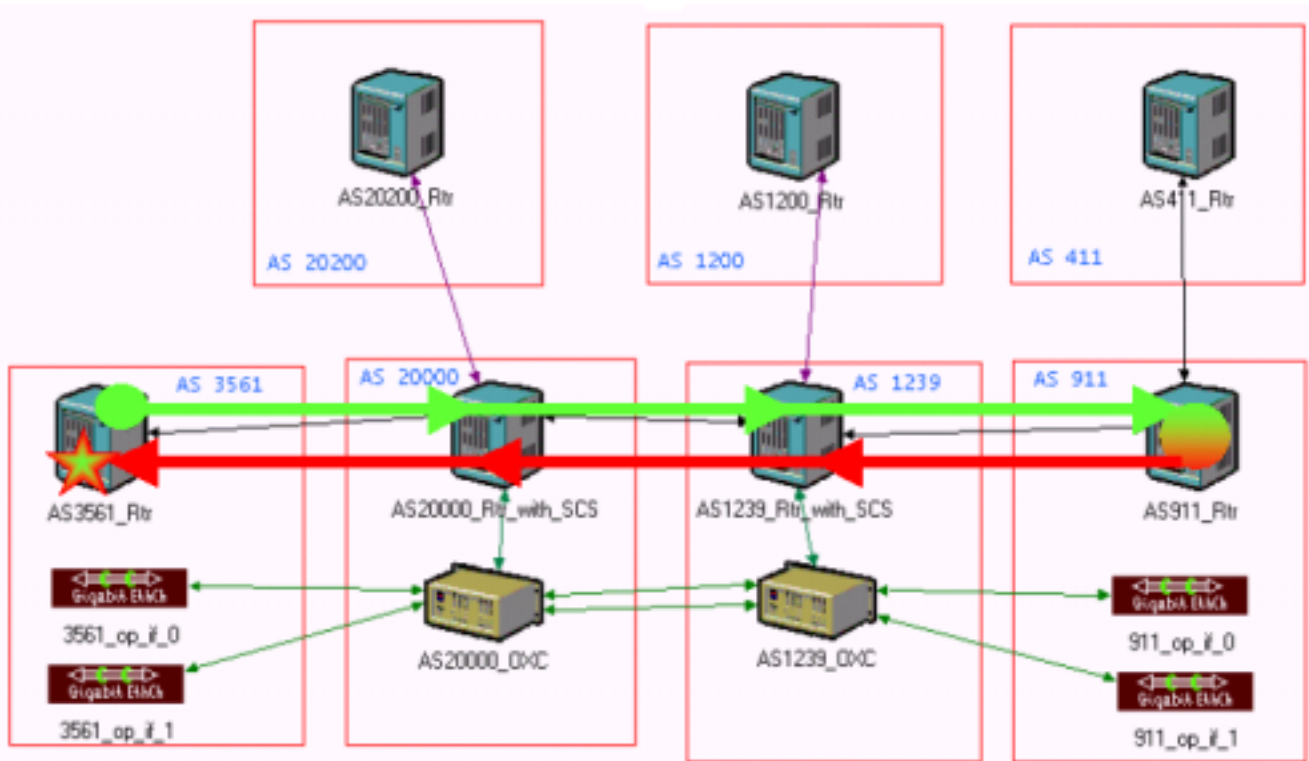6 AS 3561 receives the message indicating a fully reserved and configured lightpath.

Figure 5: Test Network and sequence of events for OBGP Lightpath Reservation.

3.2 OBGP Test Results

Round Trip Time (RTT) was measured in the OBGP test network to determine the amount of time to reserve a lightpath using OBGP. The RTT specification is the sum of propagation delay, queuing delay and processing delay at each node. The round trip time is a measure of a lightpath setup time. Round Trip Times have been collected from reservations made throughout the network to AS911, illustrated in Table 2.

Source of Reservation	Round Trip Distance (km)	Round Trip Time (seconds)
AS3561_Rtr	2200	0.013
AS20000_Rtr_with_SCS	1400	0.008
AS1239	800	0.005

Table 2: Round Trip Time of Lightpath Reservation for Source to AS911.

The reservation from AS3561 to AS911 takes a total time of 13ms for a network that spans 2200 kilometers. The RTT may appear to be a small number, but is acceptable considering the size and simplicity of the test network. This means that the processing time is small compared to the propagation time. Through the use of the OBGP Active tag, BGP can process OBGP messages much faster than regular BGP messages. What is not considered in this RTT is the switch time of each OXC, as it was not factored into the simulation.

4. FUTURE RESEARCH

OBGP has been shown as a valid way of reserving and creating end-to-end lightpaths from source to destination in an interdomain optical network. Yet, there is still further research that must be explored to ensure that OBGP will work with all scenarios.

Firstly, the two-phase reservation process may not be sufficient or robust enough to handle all scenarios that may happen. Results shown in this paper confirm that the reservation with two phases is sufficient for the simple linear network. Yet, further research must be considered to explore the implications of two-phase reservation versus more phases.

It is assumed that each domain has the capability to convert wavelengths. When some domains do not have this capability, the *wavelength continuity constraint* [5] is introduced. This constraint requires that the same wavelength must be used along a lightpath. To support this scenario, some new capabilities must be incorporated into the approach discussed in this paper.

The solution proposed in this paper is to try to minimize changes to BGP protocol. The signaling capabilities have been encoded into the BGP UPDATE message (specifically the NLRI field). In future we may introduce more capabilities to support various scenarios. Therefore it may be useful to explore the usage of a new OBGP message type. This message can be easily introduced using the BGP Capabilities feature [8]. Through this new message, less impact would be made to existing BGP functionality when more features are introduced into OBGP.

Network reliability is a very important issue, which OBGP must consider. By creating several diversely routed paths from source to destination, the end-to-end connection from source to destination can be protected under failure of a link. The current approach of OBGP uses the AS Path list to determine the destination. Since BGP stores only one path to a destination, further research must be conducted to allow OBGP to set up alternate paths.

5. CONCLUSIONS

This paper has outlined a method of extending BGP to support lightpath setup and management across an optical network. The development of OBGP has been discussed by reviewing current BGP behavior and design requirements for OBGP. An implementation of OBGP using simulation tools has been presented, along with initial test results.

The main goal of developing and simulating OBGP was to extend the BGP protocol to provide added functionality for lightpath setup. This goal has been met by demonstrating how OBGP can be used to set up lightpaths in a simulation network. These initial demonstrations prove that OBGP is a viable mechanism and is worthy of future study. The modifications made so far to BGP include two new tags and one new attributes. Some behavior changes have been made to BGP. However OBGP is still compatible with BGP. We feel that overall changes are small and acceptable. A seamless migration from BGP to OBGP is possible.

The explosion of the Internet will demand new and innovative solutions for the IP/Optical convergence. The development and implementation of OBGP has demonstrated an example of this convergence. Continued research into OBGP is expected, as further studies will revise the details of the protocol.

REFERENCES

1. G. Bernstein, et al., Some Comments on GMPLS and Optical Technologies, Network Working Group, November 2000.
2. B. Rajagopalan, et al., IP over Optical Networks: Architectural Aspects; IEEE Communications Magazine, Sept. 2000.
3. B. St.Arnaud, et al., Optical BGP Networks Discussion Paper, March 20, 2000, Revised Draft July 16, 2000.
4. Y. Rekhter, et al., A Border Gateway Protocol 4 (BGP-4), IETF Request for Comments 1771, March 1995.
5. J. Yates, M. Rumswewicz, J.P.R. Lacey, Wavelength Converters in Dynamically-Reconfigurable WDM Networks, IEEE Communication Surveys, 1999.
6. H. Zang, et al., A Review of Routing and Wavelength Assignment Approaches for Wavelength-Routed Optical WDM Networks, SPIE Optical Networks Magazine, January 2000.
7. T. Bates, et al., Multiprotocol Extensions for BGP-4, IETF Request for Comments 2858, June 2000.
8. R. Chandra, et al., Capabilities Advertisement with BGP-4, IETF Request for Comments 2842, May 2000.