A Stable Path Selection Algorithm for Protecting Optical Networks using OBGP Backup Routing

Wen-Fong Wang
Dept. Computer and Communications Engineering
National Yunlin University of Science & Technology
Douliu city, Yunlin, Taiwan
wwf@yuntech.edu.tw

Lih-Chyau Wuu
Institute of Information Engineering
National Yunlin University of Science & Technology
Douliu city, Yunlin, Taiwan
wuulc@yuntech.edu.tw

Abstract—BGP is an inter-domain routing protocol that allows an autonomous system (AS) to apply local policies for selecting the best route and decide whether this route is propagated to other ASes or not. When a network failure occurs, BGP withdraws the failed path and selects immediately an alternate path for backup routing. In this investigation, we study the Optical Border Gateway Protocol (OBGP), which can give edge network users in optical networks an automatic control means to establish a lightpath through optical AS domains. Actually, OBGP inherits the features from BGP for optical networks. However, in previous research, it indicates that BGP cannot guarantee the system stability of backup routing under failures; this is due to the inconsistency of local policies between two neighboring ASes. For the stable and safe backup routing of OBGP, we propose a stable path selection scheme for optical AS domains and draw properties for conducting local policies. To verify the scheme, a prototype implementation of our approach and a test environment are fulfilled for functional testing. From the testing, it shows the basic functions of our scheme are correct and effective.

Keywords—OBGP; BGP; Backup Routing; Convergence

I. INTRODUCTION

Optical networks have become more accessible to users on the edge of communications networks after fibre optic cables were laid in and among many communities by carriers. With WDM (Wavelength Division Multiplexing) technology, the users can create high bandwidth connections to their peer groups by employing the leased links and wavelengths of the optical networks. To light up dim fibers, carriers are willing to operate their own "wavelength cloud" to offer such lightpath service to edge network users. In this situation, establishing connectivity of lightpaths requires manual provision and management.

There are few ways for managing and configuring wavelengths between network domains, which allow edge network users to manage their own lightpaths across several wavelength clouds. By shifting the responsibility of administrating lighpaths to users, carriers allow the users to manage their own optical wavelengths better and avoid some expensive services such as lightpath management provided by the carriers. In [1], the authors show that the board gate protocol (BGP) can be extended to allow an edge user to set up a lightpath to peers across AS domains. This approach is called Optical BGP or OBGP. It is a distributed approach, which gives control to edge users, and allows them to manage their leased objects better. OBGP can provide an inter-domain routing and signaling capability that integrates heterogeneous domains into an end-to-end optical network and can coexist with most of the existing intra-domain solutions.

In the related research, M. Francisco, et al. presented new attributes and tags carried by UPDATE messages to reserve optical wavelengths for a lightpath setup [2]. Another work in [3] proposed a new message type for OBGP, called "OBGP message" to achieve end-to-end signaling and routing for optical networks. The authors created a wavelength table for each OBGP router to store wavelength availability and setup information. In [1], [4], the authors described the applications and functional requirements of OBGP and investigated the lightpath provisioning for inter-domain routing. To extend the BGP protocol for optical networks, a few new optional attributes have been considered and created in the protocol data units of BGP so that wavelength information can be encoded into the routing information base (RIB) of BGP. In [5], the authors discussed a broad range of issues related to the requirements for general inter-domain and inter-area routing in optical networks. They reviewed the applicability of existing routing protocols in the Internet and telecommunications for various optical routing. In our investigation, we follow the results in [1], [4], which seem to be more promising and realizable.

One common weakness in most optical networks is that any link or router failure among ASes would cause the significant loss of transmitted data. In the Internet, there are thousands of ASes connected, whereas an AS is a collection of routers and links operated by a single institution. To increase the reliability of networks under link or router failures, backup routing schemes could be used to withdraw a failed route and select an alternate path to recover the communication service. For this alternate path we can call it a backup path. Nevertheless, the backup path is not easy to select and must be constrained by some commercial relationships between ASes. In some failure scenarios, the backup route would introduce a BGP convergence problem [6], which results in protocol divergence. The work in [7] presents a general model for backup routing while allowing each AS to apply local routing policies that are consistent with the commercial relationships it has with its neighbors. The authors proved their model is inherently safe in the sense that the global system remains safe

This work was supported by the National Science Council, Taiwan, R.O.C., under contracts NSC96-2221-E-224-006.

under any combination of link and router failures. The safe characteristic means that the sets of routing policies would never lead to BGP divergence.

In this study, a stable path selection scheme for OBGP safe backup routing is considered. In our approach, several properties are suggested for an AS to follow in setting its routing policies, and an algorithm is proposed for OBGP to find best safety backup paths. With our approach, OBGP routers can select a best and safe backup path to restore transmission quickly and attain minimum data loss, while a link or router fails. Throughout the paper, the two words, path and route, are used interchangeably.

This paper is organized as follows. Section II describes the OBGP protocol and its incorporated architecture with OXCs. Section III specifies a safety model for backup routing [8], [9], [10], which can avoid the convergent problem of BGP as local policies apply in case a failure occurs in AS domains. In Section IV, our scheme for safe backup routing in optical AS domains is presented. We sketch and formalize the new properties of OBGP to be guidelines, which govern the setting of AS local routing policies. Also, an algorithm is devised to find the best and safe backup route for OBGP entities. In Section V, the implementation of our scheme and a functional testing environment are described. Finally, our work is concluded in Section VI.

II. OBGP System Architecture

BGP can reach different AS domains by using dedicated AS paths based on path vector routing. The usage of AS paths enables routing decisions to prevent routing loops. Having full path visibility is rather useful to enhance BGP for setting up a lightpath from one AS to others. To allow BGP carrying lightpath information, the BGP's OPEN and UPDATE messages can be used to include lightpath setup information in addition to reachability information [4], [11]. There are two possible ways to perform lightpath reservation using OPEN and UPDATE messages, first carrying a lightpath reservation request between OBGP speaking devices and secondly propagating the status of lightpath reservation information throughout the network.

Optical cross connects (OXCs) are non-blocking, reconfigurable optical switches where an optical signal entering any input port can be directed to any desired output port. In WDM networks, the OXCs may be combined with other optical components, such as optical multiplexers and demultiplexers, optical filters, etc., to fulfill wavelength routing. For OBGP, it is proposed that OXCs can be integrated with BGP routers [1]. As Router B in Fig. 1(a), a new router, called an OBGP router, combines a BGP router with OXCs.

Usually, two pairs of input and output ports are necessary to form a bidirectional link of connecting two routers via an OXC. For the bidirectional link, the two pairs of input and output ports with the connections inside the OXC constitute an optical cross connect as shown in Fig. 1(a). In [1], the basic concept of a virtual BGP router is to bind each optical cross connect with a separate BGP process and administer the bidirectional optical channel; and together for the wavelengths used in an OBGP router, assume mapping can be made between the wavelengths and IP addresses. The use of a virtual BGP router for each cross connect can allow the use of standard BGP routing with virtually no modifications necessary to support optical lightpaths. As for tunable lasers and filters, which have a limited range of wavelengths, different IP suffixes can be used to indicate the appropriate wavelength range. In addition, the virtual BGP router could be assigned its own private (or public) AS for inter-domain routing. The main purpose of OBGP routers is to be able to announce routes, perform route filtering and classification, and provide enhanced BGP capabilities to other OBGP peers.

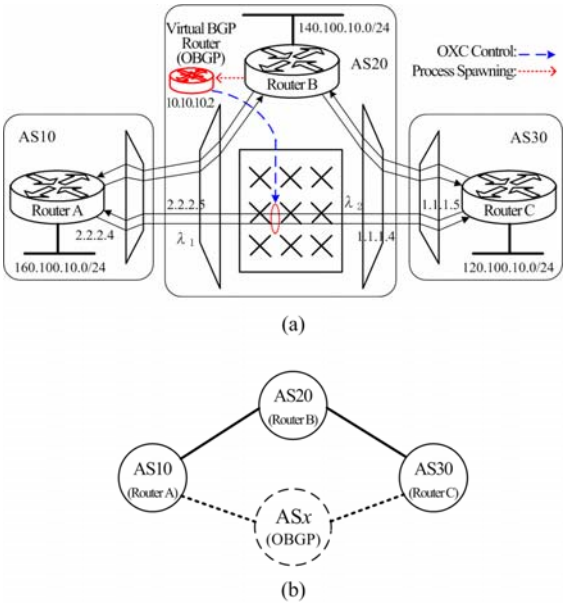

Figure 1. (a) Integration of an OXC and a BGP Router; (b) The abstract AS model

To explain the operation of virtual BGP routers, suppose that Router B receives BGP OPEN messages from Routers A and C (Fig. 1(b)) asynchronously. It can decide to set up an optical cross connect between the two routers if the information about wavelengths (IP addresses), the framing protocol, the preferred destination, etc. is equal in the optional fields of the OPEN messages. Rather than modifying the existing BGP code on Router B, it is envisaged that upon detecting the optional fields in the OPEN messages from Routers A and C, a process, called Lightpath Route Arbiter (LRA), in Router B would spawn a virtual BGP router process that would establish the optical cross connect and BGP peering sessions between Routers A and C through specific input and output ports of the OXC.

Router B's LRA spawns the virtual router process on its own CPU and then creates a configuration file for the virtual router from the information it received in the OPEN messages from Routers A and C. The configuration file for the virtual router might look like as shown in Fig. 2. While Router B is configuring its new virtual router, the LRA processes in routers A and C will update their configuration statements using the information provided in the options field of the OPEN message from Router B. For example, in Fig. 2, it shows the initial configuration of the virtual BGP router. The

loop back address of interface is defined as 10.10.10.2. The wavelength λ_1 is defined to be with suffix *x.x.x.*4 and λ_2 with suffix *x.x.x.*5 as shown in Fig. 1(a). Notably, the symbol *x.x.x* means the address prefix of the shared network between neighboring ASes. Therefore, the suffix *x.x.x.*4 indicates that λ_1 can pass from AS10 through the OXC to AS30 with the fixed identifier 4 and for λ_2, vice versa. If the establishment of BGP peering sessions with Routers A and C is successful, the BGP UPDATE messages would be used for exchanging routing information; otherwise, Router B can either decide to leave the virtual BGP router in IDLE mode or close it entirely.

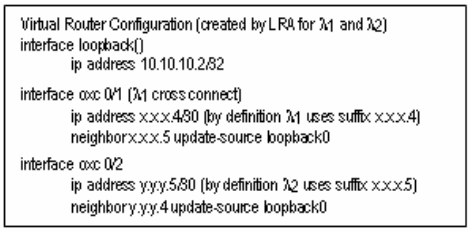

Figure 2. Configuration of a virtual BGP router

Contrary to a normal BGP multi-router configuration, the virtual BGP router would not establish any internal BGP connectivity even though it might be within Router B's AS. It would behave as an independent router carrying its own set of routes, metrics, etc. and advertise itself independently with its own loop back address and its own set of IP addresses for its interfaces.

III. POLICY-BASED CONVERGENT BACKUP ROUTING

In BGP, ASes are allowed to apply local policies for selecting paths and propagating routing information without divulging their policies or internal topology to others. The policies reflect the commercial relationships between neighboring ASes under economic incentive. Typically, the relationship of AS pairs can be customer-provider or peer-peer. To improve the reliability of inter-domain routing, a local backup relationship between ASes can be arranged to prevent link or node failure. There are two kinds of backup arrangements commonly used: multi-homed backup and peer-peer backup [14]. For multi-homed backup, it includes using a secondary customer-provider link as the link to the primary provider fails. For peer-peer backup, an existing peer-peer link is used for backup under a link failure.

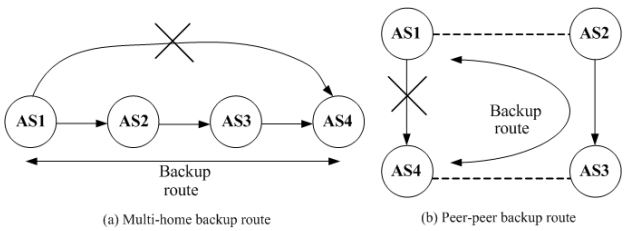

Figure 3. (a) Multi-home backup route; (b) peer-peer backup route

If a path fails, an AS should withdraw the path immediately and select a backup path to recover the interrupted services. Fig. 3 shows two examples, where the provider-customer relationship is represented as a solid line with an arrow pointing from a provider to its customer and the peer-peer relationship as a dotted line without using an arrow. Given a link failure between AS1 and AS4 in Fig. 3(a), AS4 can choose the backup path via AS3, the secondary provider. For the peer-peer backup in Fig. 3(b), if the link between AS1 and AS4 fails, the backup path can be chosen through the peer-peer links from AS1 to AS2 and AS3 to AS4. In this example, AS3 must advertise backup paths, learned from AS2, to AS4.

Indeed, local backup arrangements bring neighboring ASes more path advertisements to announce backup paths. These additional advertisements would cause global BGP convergence problems [7], [10]; and conflicting local backup policies among a collection of ASes could incur BGP route oscillations [8], [9]. To solve the issues of the BGP routing under the interaction of local backup policies, an abstract model for BGP routing policies in the context of the Stable Paths Problem (SPP) can be considered [10].

A. Stable Paths and Simple Path Vector Routing

Path advertisements in BGP are sent between ASes. These advertisements include attributes **nlri** (network layer reachability information), **next_hop**, **as_path**, **local_pref** (local preference), etc. For the path selection process of BGP, the attributes are used by import and export policies at each router to implement its local routing policies. For example, as a BGP advertisement moves from AS *x* to AS *y*, *x* applies its export policies. If the **as_path** of the advertisement contains *y*, *x* filters out the advertisement; if the path advertisement is not filtered out, then *x* is added to the **as_path**. Finally, the import policies of *y* are applied to the advertisement. This is where a **local_pref** value is assigned or modified.

Suppose an AS domain is represented by a virtual network node. Consider an AS network as an undirected graph $G = (V, E)$, where $V = \{0, 1, 2, \ldots, n\}$ is the set of nodes and E the set of edges. An edge in G is denoted by (i, j), where $i, j \in V$. For any node u, its neighbors is defined by $neighbors(u) = \{v | (u, v) \in E\}$, which can be further partitioned into three subsets: $providers(u)$, $customers(u)$, and $peers(u)$, the sets of the providers, customers, and peers of u, respectively. A path in G is a sequence of nodes $(v_k v_{k-1} \ldots v_0)$, such that $(v_i, v_{i-1}) \in E$, $1 \le i \le k$; and it has the direction from v_k to v_0. An empty path is denoted by ε. Nonempty paths $P = (v_1 v_2 \ldots v_k)$ and $Q = (w_1 w_2 \ldots w_n)$ can be concatenated if v_k is the same as w_1. Then PQ denotes the path formed by the *concatenation* of the paths. If $Q = \varepsilon$, we have $P\varepsilon = \varepsilon P = P$. For example, (123)(345) represents the path (12345), and ε(456) the path (456).

In SPP, there is an *origin* node $o \in V$, which is the destination to which all other nodes are trying to establish a path. For each node $v \in V$, it has the corresponding set of permitted paths from v to the origin (node o), denoted by P^v. Let P be the union of all sets P^v. There is a non-negative, integer-valued ranking function λ^v, defined over P^v, which represents the degree of preference to the permitted path. If P_1, $P_2 \in P^v$, and $\lambda^v(P_1) < \lambda^v(P_2)$, then P_2 is said to be preferred over P_1. Let $\Lambda = \{\lambda^v | v \in V\text{-}\{o\}\}$. We say that $S = (G, P, \Lambda)$ is an instance of SPP with a graph, the set of permitted paths

from each node to the origin, and the ranking functions for each node.

A Simple Path Vector Protocol (SPVP) [9], [10] is a distributed algorithm to solve SPP. SPVP can be thought of as an abstract model of BGP. There are two desirable properties for the SPVP with an instance of SPP:

- *Safety* ─ If the protocol SPVP can never diverge, then we say an instance of SPP is safe.
- *Inherent safety* ─ If SPP is safe, and remains safe after removing any node, edge, or permitted path, then we say an instance of SPP is inherently safe.

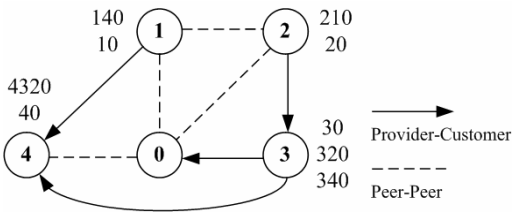

Figure 4. A bad backup arrangement: the routing protocol diverging if link (30) fails

Fig. 4 presents a bad backup arrangement, which is not inherently safe. Assume that in Fig. 4, the vertical list next to each node (except node 0) is the set of permitted paths to the common sink, *i.e.* the node 0 and the paths in each list are ranked from top to bottom for path preference. In this case, the SPVP is safe; it has a set of stable path vectors, {(140), (20), (30), (40)}, to node 0 from all other nodes. If the link (30) fails, one of the paths (320) and (340) must be chosen as the backup path. Nevertheless, the successive path advertisements for dropping the failed route and selecting a new backup route will cause the SPVP divergence.

B. Safe Backup Routing

Due to conflicting local policies, AS paths may be filtered out by neighboring BGP speakers besides the removal of AS paths due to link or node failures. To study the inherent safety of AS networks to guarantee the safety of backup routing, a specialized SPP under commercial relationships must be considered [7].

In AS domains, transit traffic (non-local traffic) must be constrained by the commercial relationship, which is either customer-provider or peer-peer, of an AS pair. Figs. 3 and 4 show the examples of AS graphs for the specialized SPP with the constraints of commercial relationships. In Fig. 4, the path (1430) is not allowable since node 4, which is a customer AS, cannot transit non-local traffic between node 1 and node 3, its providers. In this situation, we say that the path (1430) has a valley—a provider-customer edge, e.g. edge (1, 4), followed by one or more customer-provider edges. For a path with valleys inside, it is not allowed to pass transit traffic. In addition, the paths with one or more edges of customer-provider relationships (or provider-customer relationships) are allowed to pass transit traffic.

In an AS path, a mixture of peer-peer, customer-provider, and provider-customer edges will constrain the ability of relaying transit traffic. To analyze the mixture of commercial relationships in AS paths, consider a path $P_1(uv)P_2$, where (u, v) is a peer-peer edge and P_1 and P_2 might be ε. Edge (u, v) is called a step if either the last edge of P_1 is a peer-peer or provider-customer edge, or the first edge in P_2 is a customer-provider edge. For instance, in Fig. 4, the path (41230) contains no step, but the path (4120) has a step (20), the path (140) a step (40), and the path (304) a step (04). AS paths with steps should not be permitted as far as possible since valleys might exist in them and cause them the violation of commercial relationships. However, peer-peer backup arrangements often involve steps such as the case in Fig. 3(b). Instead, we need to define a slightly weaker notion of reachability, where the set of permitted paths can include paths with steps for backup routing.

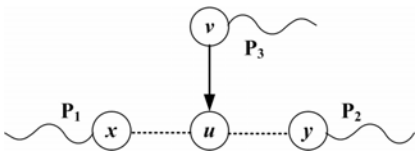

Figure 5. Conditions of permitted backup paths with a step

Fig. 5 shows the conditions of permitted backup paths with a step. Suppose that nodes x and u have peer-peer backup relationship. There are four types of peer-peer backup paths: $(vux)\mathbf{P_1}$, $(xuv)\mathbf{P_3}$, $(xuy)\mathbf{P_2}$, and $(yux)\mathbf{P_1}$, as drawn in Fig. 5. For example, in Fig. 3(b), AS4-3-2-1 is a backup path of the second type in Fig. 5, since the link between AS3 and AS4 is a step. It is worthy to recognize three points for backup paths. First, if a path P is a backup path, then $(uv)P$ is also a backup path. Next, a backup path may have one or more steps. Last, a backup path should not be used unless all primary paths are unavailable. More specifically, if path P_1 has no steps and path P_2 has one or more steps, then $\lambda(P_2)<\lambda(P_1)$. Ranking backup paths lower is essential for the safety of SPP.

To select the best backup path for recovering from failures, each node needs to rank among permitted backup paths. In [7], an effective technique is employed to sort permitted backup paths by avoidance levels. The idea of using the avoidance levels is based on counting the number of steps in a path. To utilize the avoidance levels in path selection, a non-negative function $\kappa(P)$, called avoidance classifier that is step aware for a backup path P, is devised. The value of an avoidance level is within the range of κ. In principle, an avoidance classifier κ obeys the rules below.

- As a path traverses additional edges, its avoidance level increases; for instance, if X, Y, and YX are permitted paths, then $\kappa(YX) \geq \kappa(X)$.
- κ is step aware; for any P permitted at v and $(xuv)P$ permitted at x and $(xuv)P$ being one of the above four types of peer-peer backup paths, we have $\kappa((xuv)P) > \kappa((uv)P)$.

By including the notion of avoidance classifiers to the specialized SPP under commercial relationships, the following rules must apply to the path selection process for this new SPP:

- A path with a lower avoidance level is preferred over a path with higher avoidance level; that is, if X and Y are paths permitted at a node and $\kappa(Y) \geq \kappa(X)$, then $\lambda(X) \geq \lambda(Y)$.
- With the same avoidance level, customer paths are preferred over peer and provider paths; for X and Y both permitted at u with $\kappa(X) = \kappa(Y)$, if X is a path through one of *customers*(u) and Y is not, then $\lambda(X) > \lambda(Y)$.

With the above generalization to the specialized SPP under commercial relationships, permitted paths with steps can be included for save backup routing. In summary, if the specialized SPP S that has the no-valley property, a step aware avoidance classifier κ, and preferring customers with respect to κ, then S is inherently safe.

IV. Stable Path Selection for Backup Routing

In Section II, the OBGP architecture to provision and manage lightpaths through the optical ASes has been described. However, OBGP inherits the stable convergence issue from BGP in case of link or node failures [1], [4]. In this section, a stable path selection algorithm against the convergence issue for backup routing is considered to guarantee the safety property of OBGP.

In Fig. 1(b), it shows an abstract AS model, which allows ASx containing a virtual BGP router. Suppose that a carrier (represented by AS20) leases ports of the OXC and dark fibers to customers (represented by AS10 and AS30). Then the virtual BGP router is created by Router B to establish an optical cross connect for the backup connection between AS10 and AS30. We can classify this new commercial relationship to the peer-peer relationship. The reason for the classification is that customers rent optical equipment for their private applications such as connections to their peer groups or for backup. As shown in Fig. 1(b), this peer-peer relationship consists of two peer-peer links between AS10 and ASx and between ASx and AS30, respectively. In this situation, the path AS10-x-30 contains one step (AS10-x or ASx-30). More generally, except Router B, if Router A or C itself controls OXCs, the same approach of the new peer-peer relationship can be applied to the connections to more OBGP routers.

According to the specialized SPP under the peer-peer backup relationship stated in Section III.B, the results can be extended to OBGP. Then, we convert formally the properties of the new SPP with the OBGP peer-peer relationship into Properties 4.1 to 4.4. The goal of the first guideline is to include permitted backup paths with OBGP to the set of permitted paths. The other properties can be used to ensure the scheme inherent safety.

Property 4.1 (*obgp peers*) ─ if a path $(v_k \ldots v_1 v_0) \in$ P and v_j contains only a virtual BGP router for $j=k-1,\ldots,1$, then $v_{j+1}, v_{j-1} \in$ *peers*(v_j) and the path has at least one step.

Property 4.2 (*no valley*) ─ if a path $(v_k \ldots v_1 v_0) \in$ P and $v_{j-1} \in$ *customers*(v_j) for some $j=k,\ldots,1$, then $v_{i-1} \notin$ *providers*(v_i) for all $i=j-1,\ldots,1$.

Property 4.3 (*step aware*) ─ any avoidance classifier κ must satisfy the following condition; for nodes x, u, and v, if $P \in P^v$, $(xuv)P \in P^x$, and (xuv) has a step (see Fig. 5), then $\kappa((xuv)P) > \kappa((uv)P)$.

Property 4.4 (*prefer customer*) ─ if $v \in$ *customers*(u) and $w \in$ *providers*(u) $\cup$ *peers*(u) and $\kappa((uv)P_1) = \kappa((uw)P_2)$, then $\lambda((uv)P_1) > \lambda((uw)P_2)$ for all paths P_1 and P_2.

The stable path selection algorithm for OBGP convergent backup routing can be divided into three phases. In the first phase, translate the AS graph indicated by the BGP RIB and local policies of a router to an instance of the new SPP using Property 4.1. In the second phase, delete the permitted paths of violating Property 4.2 and update the avoidance level of the remaining permitted paths by following Property 4.3. In the last phase, select the best backup path from the remaining permitted paths according to Property 4.4. The details of notations and the algorithm are described below.

Abbreviations:

o, V, E, and G: as defined in Section III.A;
AS_{local}: the local AS;
k: a finite integer;

Stable_Path_Selection( )
{ // Phase-I
 {designate $AS_{local} \rightarrow o$;
  construct G from the BGP RIB and local policies;
  for each $u \in V \wedge u \neq o$
    with *Property 4.1*,
      enumerate every $(uv_k \ldots v_1 o)$, such that $v_k,\ldots,v_1 \in V$,
          $v_k \in$ *neighbors*(u), $v_1 \in$ *neighbors*(o),
          $(v_i, v_{i-1}) \in E$, $i=k,\ldots,2$, and $v_k \neq \ldots v_2 \neq v_1$;
      include $(uv_k \ldots v_1 o)$ to P^u and P; }
  // Phase-II
  {**for each** $(v_k \ldots v_1 v_0) \in$ P
    along $(v_k \ldots v_1 v_0)$, //check *Property 4.2*
      if $(v_{j-1} \in$ *customers*(v_j), $\exists\ j=k,\ldots,1) \wedge (v_{i-1} \in$
          providers(v_i), $\forall\ i=j-1,\ldots,1)$
        delete $(v_k \ldots v_1 v_0)$ from P^{v_k} and P;
      else //follow *Property 4.3*
        if $(v_{j+1}, v_{j-1} \in$ *peers*$(v_j)) \vee$
          $((v_{j+1} \in$ *peers*$(v_j)) \wedge (v_{j-1} \in$ *providers*$(v_j))) \vee$
          $((v_{j+1} \in$ *providers*$(v_j)) \wedge (v_{j-1} \in$ *peers*$(v_j)))$,
          $\forall\ i=k-1,\ldots,1$
          //increase the avoidance level of $(v_k \ldots v_1 v_0)$;
          apply $\kappa((v_k \ldots v_1 v_0))$; }
  // Phase-III
  {**for each** $u \in V \wedge u \neq o$
    with *Property 4.4*,
      apply BGP path selection process [11] to P for the best
          backup path;
      mark the best backup path in the BGP RIB; }
}

V. Experimentation

We implemented an experimental environment (see Fig. 6) and tested the functionality of controlling OXC by using OBGP for wavelength routing. Actually, this experiment is difficult to cover all features of the OBGP scheme due to the

scale and complexity of emulating real networks, which may include many optical links. Consequently, our goal is to build a prototype implementation and verify the basic functions of the scheme.

The experimental network structure of Fig. 6 is very similar to Fig. 1(b). The role of AS20 is a service provider for customers AS10 and AS30; and, AS10 is a peer AS of AS30, and vice versa. In AS20 of Fig. 6, a virtual BGP router will be spawned by Router B2 and controls an OXC (DiCon GP700), which is used to support optical cross connections between different ASes (i.e. AS10 and AS30). Routers A and C are equipped with both ordinary Ethernet and optical gigabit Ethernet, and the remaining routers are linked by ordinary Ethernet with twisted pair cables. The testing optical channel is formed by connecting the optical Ethernet interface of Routers A and C to the I/O ports of the OXC with optical cables. Fig. 6 also shows the network configuration, including IP addresses and prefixes, and those experimental routers are implemented by personal computers with the Quagga routing software [16] installed. Furthermore, in Fig. 6, two personal computers, PCs A and B, are used to establish an FTP (File Transfer Protocol) service connection for testing and observing the change of routing information.

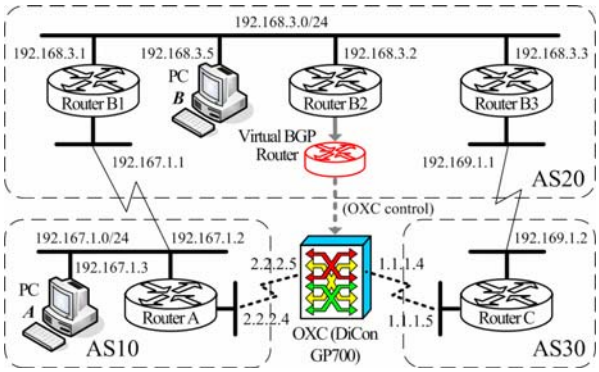

Figure 6. Experimental Environment for OBGP

In the implementation of the OBGP scheme, we develop three software modules, i.e. LRA, the stable path selection algorithm, and the OXC LabVIEW [17] driver, to integrate into the BGP protocol software. As described in Section II, the LRA is responsible to create virtual BGP routers configured according to the example shown in Fig. 2. For the virtual BGP router, its daemon process exchanges the information of lightpath reachability and finally establishes the route of lightpaths through a number of optical cross connects. The establishment of optical cross connects along an optical route is done by giving commands to the OXC driver module, coded by LabVIEW, to control the connection of input and output ports of OXCs in each OBGP node. Subsequently, using BGP UPDATE messages, the daemon process of virtual routers advertises the completed optical routes to its neighbors. For example, in Fig. 6, the optical path AS10-30 will be included in the BGP RIB of AS10 and AS30 eventually. For the stable path selection algorithm, it can be used to find out the inherently safe optical backup path for the local AS by following the properties presented in Section IV.

VI. CONCLUSION

OBGP is a distributed mechanism, which gives managing authority to users for setting up lightpaths to their peers across optical AS domains. As we considered the convergence issue, the leased commercial relationship of wavelengths and dim fibers has been extended to OBGP; and, this extension has been turn into a local policy for BGP routing. Combined with other local policies, we draw the four properties for the inherently safe backup routing of OBGP. In this study, we proposed a stable path selection scheme to cope with the convergent issue of OBGP in case failures occurred in inter-domain optical routing. To verify our approach, an OBGP prototype and an experimental environment have been implemented to conduct a functional testing. From observing the testing activities, we found that the MRAI timer [11] can influence the time for OBGP to converge. This point is very interesting for future investigation.

REFERENCES

[1] B. St. Arnaud, R. Hatem, W. Hong, M. Blanchet, and F. Parent, "Optical BGP networks," revised draft discussion paper, Mar. 2001, http://www.canarie.ca/canet4/library/ canet4design.html.

[2] M. J. Francisco, S. Simpson, L. Pezoulas, C. Hwang, I. Lambadaris, "Interdomain routing in optical networks," Opticomm 2001, *Proceeding of SPIE*, Vol.4599, pp. 120-129, Aug. 2001.

[3] M. Francisco, L. Pezoulas, C. Huang, and I. Lambadaris, "End-to-end signaling and routing for optical networks," *ICC'2002*, Vol.5, pp. 2870-2875, Apr. 2002

[4] M. Blanchet, F. Parent, and B. St. Arnaud, "Optical BGP (OBGP): interAS lightpath provisioning," Internet draft, <draft-parent-obgp-01.txt>, Aug. 2001.

[5] G. Bernstein, L. Ong, et al., "Optical inter domain routing considerations," Internet draft, <draft-bernstein-obgp-01.txt>, Jul. 2001.

[6] T.G. Griffin and G. Wilfong, "An analysis of BGP convergence properties," *Proceedings of the ACM SIGCOMM*, Aug. 1999.

[7] L. Gao, T.G. Griffin, and J. Rexford, "Inherently safe backup routing with BGP," *IEEE INFOCOM*, Vol. 1, pp.547-556, Apr. 2001.

[8] T. Griffin and G. Wilfong, "A safe path vector protocol," *IEEE INFOCOM*, Mar. 2000.

[9] T.G. Griffin, F.B. Shepherd, and G. Wilfong, "Policy disputes in pathvector protocols," *Intl. Conf. on Network Protocols*, Nov. 1999.

[10] T.G. Griffin, F.B. Shepherd, and G. Wilfong, "The stable paths problem and interdomain routing," *IEEE/ACM Trans. Networking*, Vol. 10, Issue 2 , Apr. 2002, pp. 232–243.

[11] Y. Rekhter and T. Li, "A border gateway protocol," RFC 1771 (BGP version 4), Mar. 1995.

[12] T. Bates, et al., "Multiprotocol extensions for BGP4," RFC 2858, June 2000.

[13] S.R. Sangli, D. Tappan, et al. "BGP extended communities attribute," Internet draft, <draft-ietf-idr-bgp-ext-community-07.txt>, Mar. 2004

[14] E. Chen, and T. Bates, "An application of the BGP community attribute in multi-home routing," RFC 1998, Aug. 1996.

[15] R. Chandra, P. Traina, and T. Li, "BGP communities attribute," RFC 1997, Aug. 1996.

[16] K. Ishikuro, et al., "Quagga – A routing software package for TCP/IP networks," Ver. 0.98.6, Jun. 2005.

[17] A. McDonough, "LabVIEW: Data Acquisition & Analysis for Movement Sciences," Upper Saddle River, NJ, USA: Prentice Hall, 2001.

Wen-Fong Wang received the B.S. degree in Applied Math. from National Chen-Chi University (Taipei, Taiwan) in 1983, and the M.S. degree in Information Engineering and the Ph.D. degree in Electrical Engineering from National Cheng-Kung University (Tainan, Taiwan) in 1989 and 1998, respectively. During 1989-2001, he was a researcher in Telecommunications Laboratory (TL), Chung-Hua Telecommunications Inc., Taiwan. He is now an Assistant Professor in the Department of Computer and Communications Engineering, National Yunlin University of Science & Technology (Touliu, Taiwan). His current research focuses on switching technology, metropolitan area networks, and optical networking.

Lih-Chyau Wuu received her B.S. degree in the Department of Information Engineering from National Taiwan University (Taipei, Taiwan) in 1982, and her Ph.D. degree in the Department of Computer Science from National Tsing Hua University (HsinChu, Taiwan) in 1994. She is currently a Professor in the Department of Electronic Engineering, National Yunlin Institute of Technology (Touliu, Taiwan). Her research interests include network security, high speed network and distributed self-stabilizating systems.