Journal of Innovative Technology, Vol. 1, No. 2, pp. 27-33, September 2019
http://doi.org/10.29424/JIT.201909_1(2).0005

IMPLEMENTING BACKUP ROUTING FOR OPTICAL NETWORKS VIA A BGP ENHANCEMENT SCHEME

Wen-Fong Wang [1*] and Yen-Ting Chou [2]

ABSTRACT

Border gateway protocol (BGP) is an inter-domain routing protocol that allows an autonomous system (AS) to select the best route and decide whether this route is propagated to other ASes or not. As a network failure occurs, BGP may withdraw the failed path and select immediately an alternative path for backup routing. In this study, we investigate an optical BGP which can give edge-network customers in the optical networks an automatic control access to establish a light path through optical AS domains. This protocol is actually an extension of BGP in the optical network. However, in the previous research, it has indicated that BGP cannot guarantee the stability of backup routing when a system failure occurs; this instability is due to the inconstancy of local policies between two neighboring ASes. In order to create a stable and safe backup routing, we propose a new scheme for backup routing in the AS domains and create guidelines for conducting local policies. We also devise an algorithm to identify the safest backup path for inter-connected optical ASes. To verify the new scheme, the functionality of the backup routing is tested in an experimental environment. The result shows that our new scheme can be an effective backup routing.

Keywords: BGP, OBGP, Backup Routing, Convergence.

1. INTRODUCTION

Optical networks have become more accessible to users on the edge of the networks after fiber optic cables were laid in many communities by network carriers. With the technology of Wavelength Division Multiplexing (WDM), users can create high bandwidth connections to their peer groups by employing the leased links and wavelengths of the optical networks. In order to light up dim fibers, network carriers are willing to set up their own "wavelength cloud" to offer such lightpath service to customers at the edge of the network. This service will require network carriers to manually set up and manage the connectivity.

There are a few ways for managing and configuring wavelengths among network domains. These methods will allow network customers to manage their own lightpaths across several wavelength clouds. By shifting the responsibility of managing lightpaths to customers, customers can customize their own optical wavelengths and avoid paying a high maintenance fee to the network carriers. Researchers show that the BGP protocol can be extended to allow an edge customer to set up a lightpath to their peers across AS domains (Arnaud *et al*. 2001). This approach is called Optical BGP or OBGP. It is a distributed approach which gives more control to the edge users. They can manage their leased objects better. The OBGP scheme can provide an inter-domain routing and a signaling capability that integrate heterogeneous domains into an end-to-end optical network. It can also coexist with most common intra-domain solution.

Related research shows that new attributes and tags carried by UPDATE messages can reserve optical wavelengths in a lightpath setup (Francisco *et al*. 2001). Another study proposes a new "OBGP message". This type of messaging can establish end-to-end signaling and routing in optical networks (Francisco *et al*. 2002). The authors create a wavelength table for each OBGP router to store wavelength availability and setup information. Other authors describe the application and functional requirements of the OBGP scheme and investigate the lightpath provision for inter-domain routing (Arnaud *et al*. 2001; Blanchet *et al*. 2001). 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. Hence, wavelength information can be encoded into the routing information base (RIB) of BGP. Bernstein et al. discuss a broad range of issues related to the requirements for general inter-domain and inter-area routing in optical networks. They review the applicability of various existing routing protocols in the Internet and telecommunications. In our investigation, we develop our new scheme based on the research findings of Arnaud *et al*. (2001) and Blanchet *et al*. (2001), which seem more reasonable and applicable.

One of common weaknesses in most optical networks is that any link or router failure among ASes would result in a significant loss of transmitted data, especially in the OBGP controlled networks. In the Internet, there are thousands of ASes connected with one another, and an AS is a collection of routers and links operating in a single institution. To increase the reliability of networks when the failure of a link or a router happens, the backup routing schemes can be used to withdraw a failed route and to select an alternative path. This path is referred as a backup path. Nevertheless, the back path is not easy to select, and it must be constrained by some commercial relationships among ASes. In some examples of network failure, the backup route will cause a BGP convergence problem (Griffin and Wilfong 1999), which will result in protocol divergence. Some researchers use a general

Manuscript received May 11, 2018; revised October 31, 2018; accepted April 12, 2019.

[1*] Associate Professor (corresponding author), Department of Computer Science & Information Engineering, National Yunlin University of Science and Technology, Douliu, Yunlin, Taiwan, 64002, R.O.C. (e-mail: wwf@yuntech.edu.tw).

[2] Master Student, Department of Computer Science & Information Engineering, National Yunlin University of Science and Technology, Douliu, Yunlin, Taiwan, 64002, R.O.C.

model to backup routing, and it will allow each AS to apply local routing policies that are consistent with the commercial relationships with the neighbors (Gao *et al*. 2001). The authors prove that their model is inherently safe, and the global system also remains safe in any combination of link and router failures. A safe network refers to the sets of routing policies that will never lead to a BGP divergence.

In this study, an enhancement in OBGP is applied. Our new scheme is called the OBGP with backup routing (OBGP-BR). In our approach, several guidelines are proposed for an AS to follow its routing policies, and an algorithm is devised for OBGP to find the best, safest backup path. Our OBGP-BR will be able to restore transmission quickly and minimize data loss, when an optical link or router fails. Because of this inherent failure-proof property, we can assure that OBGP-BR is a convergent inter-domain routing scheme in optical networks.

Throughout this investigation, the two words, path and route, are used interchangeably. The paper is organized as follows. Section 2 in this paper describes the OBGP and its incorporated architecture with optical cross connects (OXCs). Section 3 specifies a safety model for backup routing, which can avoid the convergent problem of BGP as local policies are applied in a network failure (Griffin and Wilfong 2000; Griffin *et al*. 1999; Griffin *et al*. 2002). In Section 4, our new scheme of OBGP-BR will be depicted. We sketch and formalize the guidelines for the new properties of OBGP, and these guidelines will govern how AS applies routing policies. In Section 5, the implementation of OBGP-BR in a functional testing environment will be elaborated. Finally, our work will be concluded in Section 6.

2. OBGP SYSTEM ARCHITECTURE

BGP can reach different AS domains by using designated AS paths according to path vector routing. The usage of AS paths will enable routing decisions to avoid routing loops. Having full path visibility is useful for BGP to set up a lightpath from one AS to another. Allowing BGP to carry information through the lightpath, the OPEN and UPDATE messages can be used to manage lightpath setup and the reachability (Blanchet *et al*. 2001; RFC 1771, 1995). There are two possible ways to perform lightpath reservation in the use of OPEN and UPDATE messages. First, carrying a lightpath reservation request between OBGP speaking devices and next propagating the status of lightpath reservation information throughout the network.

OXCs are non-blocking, reconfigurable optical switches where an optical signal enters any input port and is redirected to any target output port. In WDM networks, the OXCs may be combined with other optical components, such as multiplexers, demultiplexers, and optical filters. In OBGP, research points out that OXCs can be integrated with BGP routers (Arnaud *et al*. 2001). As it is showed in Fig. 1(a), Router B, a new OBGP router, is combined with another BGP router in OXCs.

Usually, two pairs of input and output ports are required to construct a bidirectional link in order to connect 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). The basic concept of a virtual BGP router is to bind each optical cross connection with a separate BGP process through a bidirectional optical channel (Arnaud *et al*. 2001). With the use of wavelengths in an OBGP router, mapping can be created between the wavelengths and

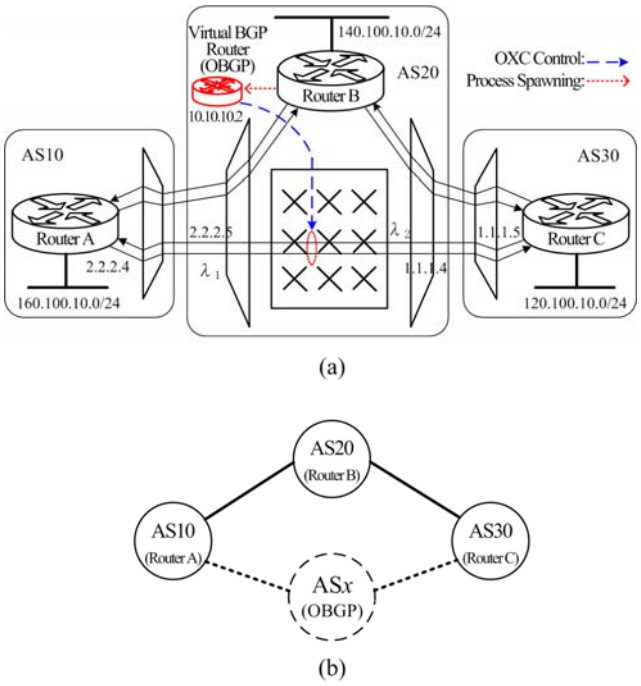

Fig. 1 (a) Integration of an OXC and a BGP Router. (b) The abstract AS model.

IP addresses. The use of a virtual BGP router in each cross connection can support optical lightpaths with no additional modifications. As for tunable lasers and filters, which have a limited range of wavelengths, different IP suffixes can be assigned to identify the specific wavelength range. In addition, the virtual BGP router can assign its own private or public AS in the inter-domain routing. The main purposes of OBGP routers are to assign routes, to perform filtering and classification, and to provide enhanced BGP capabilities to other OBGP peers.

To explain the operation of virtual BGP routers, we may suppose that Router B receives BGP OPEN messages from Routers A and C asynchronously in Fig. 1 (b). Router B can decide how to set up a lightpath with another router if the information of wavelengths, the framing protocol, and the preferred destination are equal in the optical 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.

```
Virtual Router Configuration (created by LRA for λ1 and λ2)
interface loopback()
        ip address 10.10.10.2/32
interface oxc 0/1 (λ1 cross connect)
        ip address x.x.x.4/30 (by definition λ1 uses suffix x.x.x.4)
        neighbor x.x.x.5 update-source loopback0
interface oxc 0/2
        ip address y.y.y.5/30 (by definition λ2 uses suffix x.x.x.5)
        neighbor y.y.y.4 update-source loopback0
```

Fig. 2 Configuration of a virtual BGP router

The LRA in Router B will start the virtual router process on its own CPU and then creates a configuration file for the virtual router based on the OPEN messages received from Router A and C. The configuration file for the virtual router can be illustrated in Fig. 2. While Router B is configuring its new virtual router, the LRA processes in Router A and C will also update their configurations by using the information provided from Router B. The address of the loopback interface is defined as 10.10.10.2. Suppose the wavelength λ_1 is assigned with the suffix x.x.x.4 and λ_2 with the suffix x.x.x.5 as it is shown in Fig. 1(a). The symbol x.x.x means the address prefix of the shared network among each neighboring ASes. Therefore, the suffix x.x.x.4 indicates that λ_1 can pass through AS10 through the OXC to AS30 with fixed identifier 4. The same principle will also work for λ_2. If the establishment of BGP peering sessions with Routers A and C is successful, the BGP UPDATE messages can be used to exchange routing information; otherwise, Router B can either decide to leave the virtual BGP router in IDLE mode or close the path completely.

Contrary to a normal BGP multi-router configuration, the virtual BGP router will not establish any internal BGP connectivity even though it might be within the AS of Router B. It will behave as an independent router carrying its own set of routes and metrics, advertising itself independently, and having its own loopback and IP addresses among interfaces.

Globally, the operations of OBGP can be divided into two phases:

- The first phase is to exchange the routing information about lightpath reachability and the topology of ASes.
- The second phase is about lightpath signaling and setup.

During the first phase, an OBGP router advertises the IP addresses assigned to the optical ports of an optical cross connection. They will become available lightpaths in the local OXC. The information can be encoded by using multiprotocol BGP extensions (RFC 2858, 2000) and BGP extended community (Sangli, et al., 2004). Afterward, the OBGP router will build up a lightpath RIB to determine if there is a feasible lightpath across a number of OXCs on different OBGP sites. In the second phase, it uses the information received from the previous phase and then adopts a BGP UPDATE message to establish the channel of lightpaths through the OXCs on the routing path. All the information will be encoded with multiprotocol BGP extensions and BGP extended community.

There are several advantages of using OBGP to set up and control lightpaths in the optical networks. First, a virtual BGP router can reconfigure its optical cross connection easily to interconnect with other OBGP neighbors if the traffic is changed. Second, the number of the virtual BGP router in processes is scalable. Third, OBGP can give customers the authority to manage their virtual optical resources. In other words, customers will have the ability to configure and manage their own lightpaths.

3. POLICY-BASED CONVERGENT BACKUP ROUTING

In BGP, ASes is allowed to apply local policies, select paths, and propagate routing information without divulging their policies or internal topology to others. The policies reflect the commercial relationships between neighboring ASes and economic incentives. Typically, the relationship between an AS pairs can be defined as customer-provider or peer-peer. To improve the reliability of inter-domain routing, a local backup relationship in 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 (RFC 1998, 1996). In the multi-homed backup, the network will activate a secondary customer-provider link when the link to the primary network provider fails. In the peer-peer backup, an existing peer-peer link is used as a backup route when a link failure happens.

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 presented in a solid line with an arrow pointing from a provider to its customer, and the peer-peer relationship is presented in a dotted line without 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 from AS3 to AS4. In this example, AS3 must advertise backup paths, learned from AS2, to AS4.

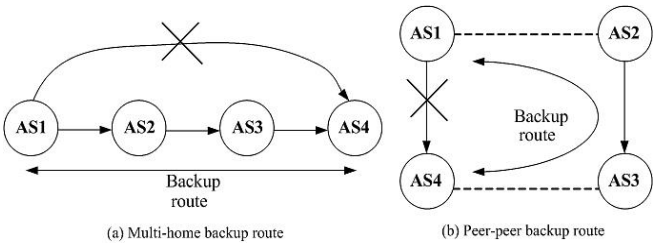

Fig. 3 Two kinds of backup routes

Indeed, local backup arrangements can bring neighboring ASes more path advertisements and also announce backup paths. These additional advertisements will cause global BGP convergent problems (Gao *et al*. 2001; Griffin *et al*. 2002). Conflicts in local backup policies among ASes can incur GBP route oscillations (Griffin and Wilfong 2000; Griffin *et al*. 1999). In order to solve the issues in the interaction of local backup policies, Grinffin *et al*. (2002) has proposed a new abstract model for BGP routing policies in the hope of tackling Stable Paths Problem (SPP). SPP is a static formalism, and it provides a formal semantics for BGP policies. Therefore, BGP becomes a distributed means to solve SPP.

3.1 Stable Paths and Simple Path Vector Routing

Path advertisements in BGP are sent between ASes. These advertisements include attributes of nlri (network layer reachability information), next_hop, as_path, and local_pref (local preference), etc. In the path selection process of BGP, the attributes are used by import and export policies in each router. As a BGP advertisement moves from AS x to AS y, x will apply its export policies. If the as_path of the advertisement contains y, x will filters out the advertisement; if the path advertisement is not filtered out, then x will be added to the as_path. Finally, the import policies of y are will also be applied in the advertisement. This is where a local_pref value will be 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 $V = \{0, 1, 2, \cdots, n\}$ is the set of nodes and E is the set of edges. An edge in G is denoted by (i, j), where $i, j \in V$. For any node u, its neighbors are defined by neighbors(u) = $\{v| (u, v) \in E\}$, which can be further partitioned into three subsets: providers(u), customers(u), and peers(u), respectively. A path in G is a sequence of nodes $(v_k v_{k-1} \cdots v_0)$, such that $(v_i, v_{i-1}) \in E$, $1 \le i \le k$; and it has the direction from vk to v0. An empty path is denoted by ε. Nonempty paths $P = (v_1 v_2 \cdots v_k)$ and $Q = (w_1 w_2 \cdots w_n)$ 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. Suppose $\Lambda = \{\lambda^v \mid v \in V\text{-}\{o\}\}$. We can 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.

In previous studies (Griffin *et al*. 1999; Griffin *et al*. 2002), a Simple Path Vector Protocol (SPVP) is a distributed algorithm to solve SPP. SPVP can be considered as an abstract model of BGP. There are two desirable properties of the SPVP in an instance of SPP:

Safety — If the protocol SPVP will never diverge, then we say an instance of SPP is safe.

Inherent safety — If SPP is safe, and it remains safe after removing any node, edge, or permitted path, then we can say that an instance of SPP is inherently safe.

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 in path preferences. 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 link (30) fails, one of the paths (320) and (340) can be chosen as the backup path. Nevertheless, the successive path advertisements for dropping the failed route and the process of selecting a new backup route will cause the SPVP divergence.

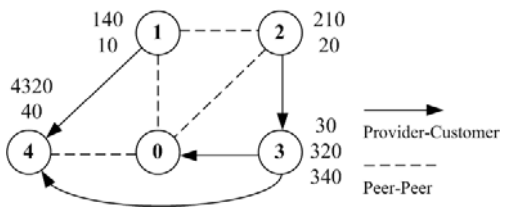

Fig. 4 A bad backup arrangement: the routing protocol diverging if link (30) fails

3.2 Safe backup routing

Due to the conflicts in local policies, AS paths may be filtered out by neighboring BGP speakers, in addition to the removal of AS paths caused by link or node failures. In order to study the inherent safety of AS networks and to guarantee the safety of backup routing, a specialized SPP in commercial relationships must be considered (Gao *et al*. 2001).

In AS domains, transit traffic (non-local traffic) must be constrained by the commercial relationship, which is either customer-provider or peer-peer as a pair in AS. 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 allowed, since node 4, a customer AS, cannot transit non-local traffic between node 1 and node 3, which are the providers. In this situation, we say that the path (1430) has a valley—a provider-customer edge, which is edge (1, 4), followed by one or more customer-provider edges. On a path with valley inside, transit traffic will not be able to pass through. However, transit traffic will be able to pass through the paths with one or more edges of customer-provider or provider-customer relationships.

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 may be ε. Edge (u, v) is denoted as a step if either the last edge of P1 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 since valleys might exist among them and create the violation in commercial relationships. However, peer-peer backup arrangements often involve steps, as it is shown in Fig. 3(b). Instead, we need to define a slightly more general notion of reachability, where the requirements of permitted paths can also tolerate steps in backup routing.

Fig. 5 shows the conditions of permitted backup paths with a step. Suppose that nodes x and u have a peer-peer backup relationship. There are four types of peer-peer backup paths: (vux) P1, (xuv) P3, (xuy) P2, and (yux) P1, as drawn in Fig. 5. For example, in Fig. 3(b), AS4-3-2-1 is a backup path corresponding to the second type of path in Fig. 5, since the link between AS3 and AS4 is a step. It is worthy to mention three points in the 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 P1 has no steps, and path P2 has one or more steps, then $\lambda(P_2) < \lambda(P_1)$. Ranking backup paths lower is essential for the safety of SPP.

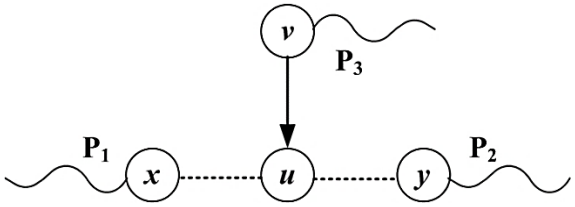

Fig. 5 Conditions of permitted backup paths with a step

In order to select the best backup path and recover from network failures, each node needs to be ranked among permitted backup paths. An effective technique is employed to sort permitted backup paths by avoidance levels (Gao *et al*. 2001). The idea of using the avoidance levels is to count the number of steps in a path. In order to do so, a non-negative function $\kappa(P)$ is devised as an avoidance classifier for a backup path P. 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 will increase; 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 will have $\kappa((xuv)P) > \kappa((uv)P)$.

By including the notion of avoidance classifiers in the specialized SPP among commercial relationships, the following rules must be applied 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 among commercial relationships, permitted paths with steps can be included to save backup routing. In summary, if the specialized SPP S that has the no-valley property, S will be inherently safe, considering a step aware avoidance classifier κ and the preferred customer path in relation to κ.

4. THE OBGP-BR SCHEME

While lightpaths are provided and managed through optical ASes, the stable convergence issue of OBGP and the cases of link or node failures were not addressed in the previous research (Arnaud *et al*. 2001; Blanchet *et al*. 2001). Therefore, the OBGP-BR scheme is devised to cope with the convergence issue.

Fig. 1(b) shows an abstract AS model which allows ASx to contain 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). Customers will be able to manage their rent-a-equipment. Then, the virtual BGP router is created by Router B to establish and control an optical cross connection in the new commercial relationship. As for the connection established by the virtual BGP router between AS10 and AS30, we can classify this new commercial relationship as the peer-peer relationship. The reason for conducting the classification is that customers usually rent optical equipment for their private applications, such as the connections to their peer groups or for the purpose of backing up data. 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). In other words, 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 with more OBGP routers.

According to the specialized SPP in the peer-peer backup relationship stated in Section 3.2, the results can be applied to OBGP. Then, we can create Guidelines 4.1 to 4.4 according to the properties we find in the new SPP and in the peer-peer OBGP. The goal of the first guideline is to include permitted backup paths in OBGP. The other guidelines can be used to ensure inherent safety in the OBGP-BR scheme.

Guideline 4.1

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

Guideline 4.2

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

Guideline 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 $\kappa((xuv)P) > \kappa((uv)P)$.

Guideline 4.4

(*prefer customer*) — if $v \in$ customers(u) and $w \in$ providers(u) 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 backup path finding algorithm of OBGP-BR is divided into three phases. In the first phase, the network will translate the AS graph indicated by the BGP RIB and local policies of a router into an instance of the new SPP, using Guideline 4.1. In the second phase, the network will delete the permitted paths which violates Guideline 4.2 and updates the avoidance level of the remaining permitted paths by following Guideline 4.3. In the last phase, the network will select the best backup path from the remaining permitted paths according to Guideline 4.4. The details of the notations and the algorithm are provided below.

Abbreviations:

o, V, E, and G: as defined in section 3.1;
AS_{local}: the local AS;
k: a finite integer;

Backup_Path_Finding( )

```
{ // Phase One
   {designate AS_local → o;
   construct G from the BGP RIB and local policies;
   for each u ∈ V ∧ u ≠ o with Guideline 4.1,   {
         enumerate every (uv_k⋯v_1o), such that v_k,⋯,v_1 ∈ V, v_k
         ∈ neighbors(u), v_1 ∈ neighbors(o), (v_i, v_i–1) ∈ E, i =
         k, ⋯, 2, and v_k ≠ ⋯v_2 ≠ v_1; include (uv_k…v_1o) to P^u
         and P;   }
   }
 // Phase Two
   {for each (v_k⋯v_1v_0) ∈ P along (v_k⋯v_1v_0),      {
     //check Guideline 4.2.
         if (v_j–1 ∈ customers(v_j), ∃ j = k, ⋯, 1 ) ∧ (v_i–1
     providers(v_i), ∀ i = j–1,⋯,1)   {
               delete (v_k…v_1v_0) from P^vk and P;   }
          else   {                    //follow Guideline 4.3.
```

if (v_{j+1}, v_{j-1} ∈ *peers*(v_j)) ∨ ((v_{j+1} ∈ *peers*(v_j)) ∧ (v_{j-1} ∈ *providers*(v_j))) ∨ ((v_{j+1} ∈ providers(v_j)) ∧ (v_{j-1} ∈ peers(v_j))), ∀ $I = k$-1,⋯,1 {
//increase the avoidance level of ($v_k \ldots v_1 v_0$).
apply κ(($v_k \cdots v_1 v_0$)); } }
}
// Phase Three
{**for each** $u \in V \wedge u \neq o$ with *Guideline 4.4*, {
apply BGP path selection process to P for the best backup path; mark the best backup path in the BGP RIB; }
}
}

In an AS network, a node can use UPADTE messages to exchange reachability information with its neighbors. Eventually, it will have complete topology information and include the primary and backup paths to each destination in its BGP RIB when the backup path finding algorithm is applied. Thus, the node can request the establishment of the backup lightpath to one destination by sending an extended UPDATE messages. Once a node or a link failure occurs in the primary path, the node can choose the backup path to recover the transmission of the primary path to the original destination.

Community	LOCAL_PREF	Category
549 : 100	set local pref 100	Customer Routes
549 : 90	set local pref 90	Customer Backup Routes
549 : 80	set local pref 80	Other ISP Routes
549 : 70	set local pref 70	Other ISP Provided Backup Routes

Fig. 6 Defining and mapping local_pref to community

To recognize a backup lightpath between ASes, the BGP community attribute (RFC 1997, 1996) can be used to exchange routing information. First, a service provider needs to coordinate with its customers and has a set of communities to be mapped according to certain BGP local_pref values (RFC 1998, 1996). The provider can apply a uniform BGP configuration to all its customers that will capture routes with the community values, and sets up the proper local_pref values accordingly. A customer who requires customization in its BGP configuration can simply send the appropriate community values in its route advertisement. Fig. 6 shows the community values which are defined with particular meanings. For example, AS 549 has defined several community values that can be used by customers to tag routes so that the appropriate local_pref values can be configured. In this example, customer routes are preferred over other provider routes; backup routes have lower local_pref than primary routes. These conditions match the guidelines we propose in this study.

5. IMPLEMENTATION AND TESTING

We have implemented an experimental environment (see Fig. 7) and tested the functionality of the OBGP-BR scheme. Actually, it is difficult to cover all the features of the OBGP-BR scheme in this experiment due to the scale and complexity of emulating real networks, which may include many optical links. Therefore, our goal is to build an implementation prototype and verify the basic functions of the scheme.

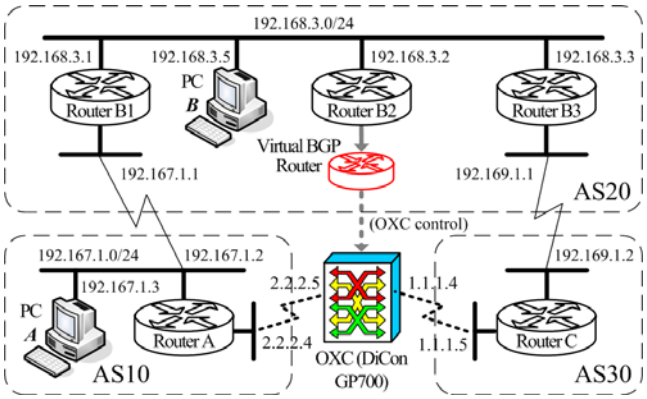

Fig. 7 Experimental Environment for OBGP-BR

The experimental network structure in Fig. 7 is very similar to Fig. 1(b). The role of AS20 is a service provider. AS10 and AS30 are the customers. AS10 is a peer AS of AS30, and vice versa. AS20 in Fig. 7 is a virtual BGP router that will be spawned by Router B2. It 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. The remaining routers are linked by ordinary Ethernet with twisted pair cables. The testing optical channel is constructed by connecting the optical Ethernet interface of Routers A and C to the I/O ports of the OXC. Fig. 7 also shows the network configuration, including IP addresses and prefixes, and those experimental routers are implemented by personal computers with the Quagga routing software (Ishkuro, et al., 2005) installed. Furthermore, in Fig. 7, two personal computers, PCs A and B, are used to establish an FTP (File Transfer Protocol) service connection for testing and observing the exchange of routing information.

In the implementation of the OBGP-BR scheme, we develop three software modules, i.e., LRA, the backup path finding algorithm, and the OXC LabVIEW driver (McDonough, 2001), to be integrated into the BGP protocol software. As described in Section 2, the LRA is responsible to create virtual BGP router configurations according to the example shown in Fig. 2. Since a virtual BGP router works in two phases presented in Section 2, its daemon process will exchange the information of lightpath reachability and establish the route of lightpaths through a number of optical cross connections. The establishment of optical cross connections along an optical route is done by giving commands to the OXC driver module, coded by LabVIEW, in order 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 will advertise the completed optical routes to its neighbors. For example, in Fig. 7, the optical path of AS10-30 will be eventually included in the BGP RIB of AS10 and AS30. The core of the OBGP-BR scheme is the backup path finding module, which utilizes the backup path finding algorithm described in Section 4. This module can find the inherently safe optical backup path for the local AS to follow the guidelines presented in Section 4.

To substantiate the operation of the OBGP-BR scheme, we verify that in the experimental environment (Fig. 7), the optical backup AS path AS10-30-20 will be selected to replace the primary AS path AS10-20, in case of failure in the primary path.

There are two parts in this testing. For the first part, it is the preamble of the testing, and it is used to ensure that the optical path AS10-30 has been included in the BGP RIB of Routers A and C after the setup of the experiment. To complete the actions in the preamble, three manual steps are required.

Step 1: Connect and configure the network for the testing according to Fig. 7.

Step 2: Initiate each LRA in Routers A, B2, and C.

Step 3: Use the software tools of Quagga. Observe and verify that the optical path AS10-30 has been included in the BGP RIB of Routers A and C after the startup of the testing.

For the second part, we need to confirm that the optical backup AS path AS10-30-20 has replaced the primary AS path AS10-20. The required actions are listed as follows.

- Initiate an FTP connection between PCs A and B.
- Disconnect intentionally the link between Router A (in AS10) and Router B1 (in AS20) in Fig. 7.
- Observe whether the optical backup path is selected to recover the FTP service within a period of time.
- Employ the tools of Quagga to display and examine the BGP RIB in the corresponding routers.

We have performed the above testing for the OBGP-BR scheme. Two observations are worth mentioning. First, the duration set for the BGP timer MRAI (Minimum Route Advertisement Interval) can affect significantly the convergent delay of obtaining stable routing information in OBGP-BR. Usually, this timer is set around 30 seconds with a jitter (RFC 1771, 1995). Nevertheless, in order to reach a fast convergence, this timer may be set for a shorter time interval. Second, since the BGP community attribute has been used for exchanging routing information in the testing, we will employ the Quagga's tools to examine the BGP RIB of Routers A and C. It is found that the optical backup path has the local_pref value lower than the primary AS path. This finding implies that the backup path finding algorithm works as we have expected.

6. CONCLUSION

OBGP is a distributed mechanism, which can give managing authority to users for setting up lightpaths to their peers across AS domains. In this study, we have proposed the OBGP-BR scheme to cope with the convergent issue of OBGP in case of the failure in inter-domain optical routing. As we considered the convergent issue, the leased commercial relationship between wavelengths and dim fibers has been extended to OBGP, and this extension has been turning into a local policy for BGP routing. Combined with other local policies, we draw the four guidelines for the inherently safe backup routing in OBGP. Also, we have presented a backup path finding algorithm for OBGP to find the best safety backup path. 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 find that the MRAI timer can influence the time for OBGP-BR to converge. This finding is very intriguing to the future investigation. Definitely, the OBGP-BR scheme is a feasible solution to guarantee the stable backup routing in optical networks.

REFERENCES

Arnaud, B. St., Hatem, R., Hong, W., Blanchet, M., and Parent, F. (2001). Optical BGP networks, Mar. 2001, discussion paper presented in http://www.canarie.ca/canet4/library/ canet4design.html.

Bernstein, G., and Ong, L., *et al*. (2001). Optical inter domain routing considerations, Internet draft, <draft-bernstein-obgp-01.txt>.

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

Francisco, M.J., Simpson, S., Pezoulas, L., Hwang, C., and Lambadaris, I. (2001). Interdomain routing in optical networks, Opticomm 2001, Proceeding of SPIE, Vol.4599, pp. 120-129.

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

Gao, L., Griffin, T.G., and Rexford, J. (2001). Inherently safe backup routing with BGP, IEEE INFOCOM 2001, Vol. 1, pp. 547-556.

Griffin, T.G. and Wilfong, G. (1999). An analysis of BGP convergence properties, Proceedings of the ACM SIGCOMM 1999, pp. 277-288.

Griffin, T.G., Shepherd, F.B., and Wilfong, G. (1999). Policy disputes in pathvector protocols, Intl. Conf. on Network Protocols, ICNP 1999, pp. 21-30.

Griffin, T.G. and Wilfong, G. (2000). A safe path vector protocol, IEEE INFOCOM 2000, pp. 490-499.

Griffin, T.G., Shepherd, F.B., and Wilfong, G. (2002). The stable paths problem and interdomain routing, IEEE/ACM Trans. Networking 10, Issue 2, 232–243.

Ishikuro, K., *et al*. (2005). Quagga – A routing software package for TCP/IP networks, Ver. 0.98.6, Jun. 2005.

McDonough, A. (2001). LabVIEW: Data Acquisition & Analysis for Movement Sciences, Prentice Hall, Upper Saddle River, NJ, USA.

RFC 1771, Rekhter, Y., Li, T., (1995). A border gateway protocol, BGP version 4, Mar. 1995.

RFC 1997, Chandra, R., and Traina, P., Li, T., (1996). BGP communities attribute, Aug. 1996.

RFC 1998, Chen, E. and Bates, T. (1996). An application of the BGP community attribute in multi-home routing, Aug. 1996.

RFC 2858, Bates, T., *et al*. (2000). Multiprotocol extensions for BGP4, June 2000.

Sangli, S.R., Tappan, D., *et al*. (2004). BGP extended communities attribute, Internet draft, < draft-ietf-idr-bgp-ext-community-07. txt> , Mar. 2004.