International Journal of Engineering and Technology, 1 (2) (2012) 94-104
©Science Publishing Corporation
www.sciencepubco.com/index.php/IJET

Improvement of BGP Session Maintenance

Sadia Amin, Saira Ahthasham , Aftab Ahmed, Ahthasham Sajid, Mirza Aamir Mehmood

Department of Computer Science, BUITEMS, Quetta
sadia_amin@yahoo.com
Islamabad College for Girls, Islamabad
sairaimtiaz@gmail.com
Department of Computer Science, BUITEMS, Quetta
aftab.ahmed@buitms.edu.pk
Department of Computer Science, BUITEMS, Quetta
ahthasham.sajid@buitms.edu.pk
Department of Computer Science, BUITEMS, Quetta
aamir.mehmood@buitms.edu.pk

Abstract

In the early days, the Internet used static routes, but after sometimes updating routing table manually becomes tough task for network administrators to achieve. Later, EGP protocol was introduced for Internet network administrator, but EGP was not scalable so in replacement of EGP a new protocol BGP was introduced to solve all problems which were unaddressed in EGP. BGP stands for Border Gateway Protocol and the most current version is BGP4. BGP is a routing protocol that runs on routers. BGP allows fully decentralized management of the Internet. That means, if BGP router is on the Internet, it can tell all other routers that what networks are available to everyone in the world. BGP calls each routing domain an autonomous system (AS). It selects the best path, through the Internet, by choosing the route that has to traverse the fewest autonomous systems.
As BGP provides information for controlling the flow of packets between AS, the protocol plays a critical role in Internet efficiency, reliability, and security. However, slow convergence and abnormal termination of session are major vulnerability of BGP. Simplifying BGP design complexity helps in research, for the root cause analysis of BGP.
Deriving technique to prevent the loss of data, reduce the convergence time and maintaining the sessions of BGP by instantly diverting the traffic from teardown / flapped link to the backup link considering as active link. It will be a great success to achieve the goal to maintain the sessions and reduce the BGP convergence time.

Keywords: *BGP, TTL, OER, OSPF, EIGRP, IS-IS*

1 Introduction

Border Gateway Protocol (BGP) is an inter-domain routing protocol i.e. between different autonomous systems (AS) that is used to exchange routing information between them. An autonomous system is the collection of networks having the same set of routing policies. In a typical inter-network (and in the Internet) each autonomous system designates one or more routers that run BGP software. BGP routers in each autonomous system are linked to those in one or more other autonomous systems. Service providers usually use Interior Gateway Protocol (IGP) e.g. RIP, OSPF, EIGRP, IS-IS etc for exchange of routing information within their Autonomous system, and with the help of BGP it is made possible to exchange routing information between IGP's and different Autonomous Systems.
An AS must have a minimum of one router running BGP, but can have more than one. It is also possible to use BGP to communicate between BGP routers within the same autonomous system [1]. Figure 1 shows the communication between 2 different autonomous systems using BGP:

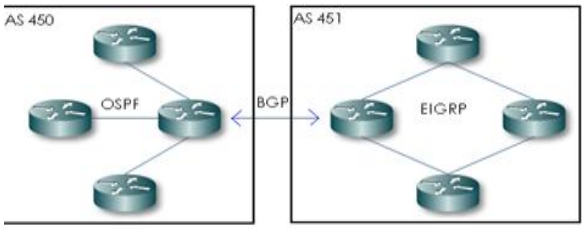

Figure: 1 Communication Between two different AS

Autonomous System 450 and Autonomous System 451 both are using different IGPs for the exchanging of routing information within their networks but BGP is running for the communication between autonomous systems.

The selection of routes by a BGP router can be controlled through a set of BGP policies that specify, for example, whether an AS is willing to carry traffic from other AS or not. However, BGP cannot guarantee the most efficient route to any destination, because it cannot know what happens within each AS and therefore what the cost is to traverse each AS [5].

BGP's operation is based on the exchange of messages that perform different functions. BGP routers use Open messages to contact neighboring routers and establish BGP sessions. They exchange Update messages to communicate information about reachable networks, sending only partial information as needed.

1.1Characteristics & Common usage of BGP

. Some of the BGP characteristics are as follow.

- BGP is a distance vector protocol which shows that BGP will advertise only those networks to its neighbors that are reachable to itself [4]. When

neighbors receive advertised networks they will consider "If that AS can reach those networks, then they can reach them via the AS".

- BGP has a very sophisticated method of calculating the shortest path by using its attributes. If two different paths are available to reach the same destination then the shortest path is used. The mechanism for calculating the distance is called metric.

- BGP has the characteristic of utilizing the bandwidth effectively. It sends the routing updates to its neighbors using a reliable transport protocol [3]. This is very powerful feature because there is no need for periodic updates or refreshing routing information. In BGP, only information that has changed is transmitted.

- The reliable transport protocol that is used by BGP is TCP (Transmission Control Protocol). It is an application protocol that uses both TCP and IP protocols for the reliable connection. Because of this feature periodic updates become unnecessary.

- The router that has received the reach-ability information from the neighboring BGP router must ensure that neighbor router is available. Otherwise the router will not route traffic towards the next hop which is no longer available, it will cause loss of packets. TCP does not signal that the neighbor has been lost unless some application data is actually transmitted between the neighbors.

Therefore, BGP detects the presence of neighbors by periodically sending BGP keep-alive packets to them and the neighbor router must reply with BGP keep-alive packet.

1.2: Common BGP Uses:

- Routers connected to more than one service provider.
- Service provider networks.
- Service providers exchanging traffic at an exchange point.
- Network cores of large enterprise consumers.

2 Single Homing & Multi-Homing

2.1 Single Homing:

In Single homing concept a small network is connected to the internet through single ISP (Internet Service Provider).

Generally BGP is not used in such types of scenarios. Normal internet access to a single ISP does not require BGP. Instead of BGP, static routes can be used in such types of scenarios. But when small ISPs get connect with other ISP by using their own AS number then BGP needs to play a vital role for connectivity.

Small ISPs used this type of network design when they are planning to connect with more than one ISP in future. In case when they are not planning for connecting with multiple ISPs then static route is always a preferred method as shown in Figure 2 Router A is connected with Router B using BGP protocol it will be an easy way to be connected via static route to access internet.

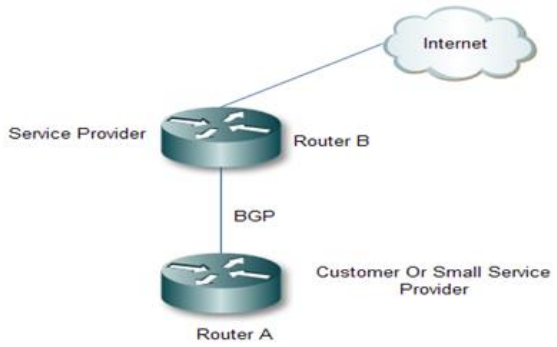

Figure: 2 Single Homing

2.2: Multi Homing:

In Figure 3 two different ISPs are connected with multi-homed small ISP by using BGP. The Small ISP has its own AS number and is responsible for advertising its own network to both ISPs. Both ISPs will forward all routes received from internet to the Multi-homed small ISP. The small ISP should not forward any routing information received from one ISP to the other.

If any one of the two links fails then traffic will get stop and no reach-ability information will be passed through the failed link but BGP reach-ability information will be advertised by the small ISP on the other link.

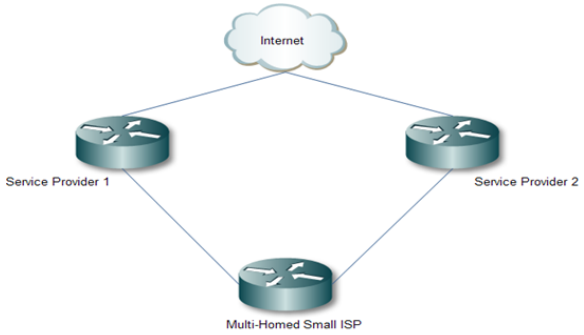

Figure: 3 Multi Homing

2.3: Transit Autonomous System

BGP is most commonly implemented in service provider networks to ensure connectivity between different ISP's and the rest of the internet. In Figure 3.3 ISP can use static routing or it may use BGP to exchange routes. Major ISP is connected with other ISPs to forward the routes that are received from small ISP's / Multi-homed/ other service providers to rest of the internet. As a result, data traffic will start to flow between them and rest of the world as shown in figure 4 [4].

Such a network providing transit services to traffic that is originated in other networks is called "***transit autonomous system***". A transit AS is an AS that exchanges BGP routing information with other AS.

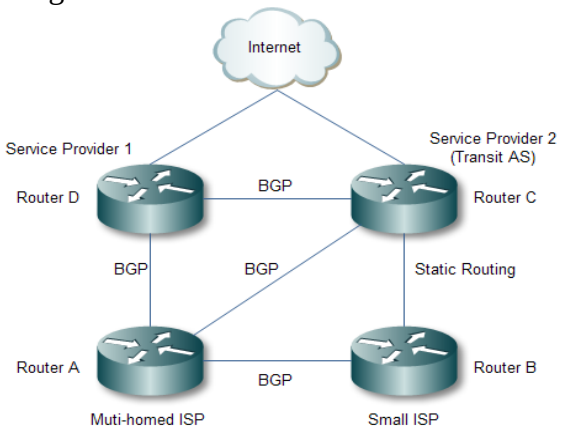

Figure: 4 Transit Autonomous System

2.4: Applying AS-path Filters

In Figure 5 AS-path filters are configured router selects only those routes that are allowed. Behavior of the selected routes is as under: When the selection is applied on the incoming routes received from neighbor, selected routes will enter in the local BGP table. The routes that are not selected will be silently. The selected routes will be sent to the neighbor when selection is applied on the outgoing routes. The routes not selected will be used locally but never sent to the neighbor.

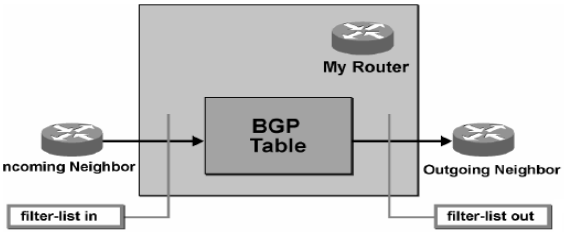

Figure: 5 AS Path Filters

3 BGP Sessions & Route Selection Criteria

BGP is an exterior gateway protocol (EGP) that is designed for scalability and policy control. BGP protocol must be pre-configured between neighboring routers for exchange of network route information. BGP differs from Interior Gateway Protocols such as Enhanced Interior Gateway Protocol (EIGRP) and Open Shortest Path First (OSPF) that discover its neighbors by using broadcast packets or hello protocol.

Unlike other routing protocols, BGP does not detect its neighbors automatically. Initially TCP session is established between two neighboring BGP routers. Once TCP session is successfully established, BGP start to exchange routing information. Attempting to open the session must be manually configured with neighbor information. If the neighboring router is not configured then it will only receive the incoming connection attempts and will not respond them. To be successful in the connection attempts both router must be configured to reach each other.

3.1: BGP Route Selection Criteria

To decide which route is best router use some BGP attributes. The selection criteria checked in the order that is described in the following steps:

1. The router checks whether the IP address that is indicated in "next hop attribute" is reachable or not. It is not necessary to have a direct connection with next hop. If the next hop is not reachable then the router does not consider the BGP route as the best [2].

2. The router prefers the route which has high value of "weight". This is not in the update; it is assigned to a router manually and is considered only within the router itself.

3. If the local preference attributes are different for different routers then the router having higher value of local preference will be selected as best.

4. If the local router injects the route into BGP table then this route will be preferred.

5. The length of AS paths are compared and the route with the shortest length is selected.

6. If the AS-path lengths are same then the origin code is checked and the path which has lowest origin is preferred. IGP (Interior Gateway Protocol) is lower than EGP (Exterior Gateway Protocol).

7. The router next compares MED values and the routes with a lower MED are preferred.

8. If the router receives all alternative routes from peer router in local AS, each of them will indicate an exit point and the closest exit is used.
9. If the router receives all alternatives from EBGP (External Border Gateway Protocol) neighbors then the most stable path which is the oldest path is preferred.

10. If the router still cannot make decision and cannot select the best path then it checks the BGP sessions and chooses the route that was received on the session for which the peer router has the lowest BGP router-ID.

4 Implementation of Session Maintenance Topology

In Figure 6 Session maintenance of BGP is successfully implemented using routers and multi-homing technique.

Company A is a multi-homed network with Autonomous System Number 10. Company A have to communicate with the Company D which have the Autonomous System Number 12. In between Company A and Company D there are 2 Companies. Company B and Company C and both of them have equal hop count to reach Company D. Based on the bandwidth and delay, Company A selects the Company B as a best path to reach Company D.

For any reason, if the link fails between Company A and Company B then the communication between Company A and Company D will now start through Company C but there will be at least 2 packets loss and within 1 second traffic will divert to Company C. To minimize the packet loss and delay, BGP and OER (Optimized Edge Routing) is configured.

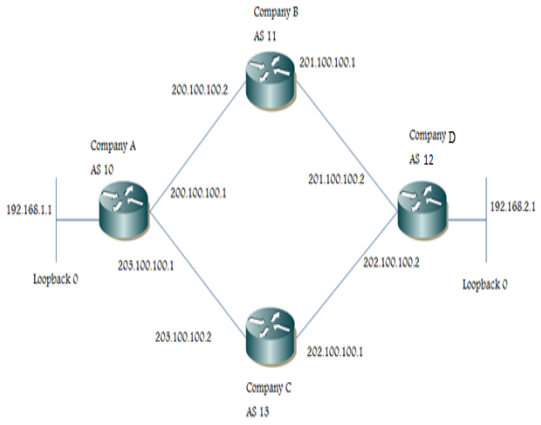

Figure 6: Session Maintenance Topology

4.1: OER (Optimized Edge Routing)

Optimized Edge Routing (OER) is intended for sites using multiple Internet or WAN Service Providers. OER is used to optimize traffic load distribution and cost minimization on links. To prevent instability, traffic shifts cannot be instantaneous, but are generally in response to conditions lasting one or more second. OER uses a router as a master controller and it communicates with one or more border routers that connect to the ISP or WAN providers. In our case, master controller and border router is a single Company A router. The links to the controller should be reliable and high-speed for proper OER operation. By implementing OER traffic never switch-back at the time of restoration of broken link. It will switch back when BGP table get cleared forcefully or current link get flapped or tear down. It is an advantage because if traffic gets switched back instantly the session will also get tear down.

OER can respond to policy violations concerning:

- response time
- packet loss
- path availability
- traffic load distribution

5 Implementation of Solution

Multi-homing concept along with OER technique really helped us to improve the session maintenance. In the topology we have to declare Company A router as Master and border router because there is BGP running in parallel OER.

First Company A router is to be declared as Border and Master router then we have to trim down the advertisement timer at minimum level (value 1:3). It should be handled very carefully and must be defined at interface level through which WAN links are connected. Both links must contain equal bandwidth to share the load easily incase if one link tear off. If it is not equally balanced then company A will face bandwidth choke and slow communication. When one WAN link tear down BGP instantly switch the traffic to the second WAN link.
Here OER play key role by keeping second WAN link as primary for the BGP traffic and never let BGP to switch back in case of restoration of first WAN link. Rapid switching between links will badly affect the session maintenance. After successfully passing the peak hours we can force fully divert traffic to primary WAN link instead of running secondary link as primary. Default routes must be configured for the unknown destined packets.

Company B and C will simply act as ISP running just BGP simple features forwarding the packets to the destination router containing routes of both source and destination routers. Same advertisement timers will be configured on

Company B router, Company C and Company D router interfaces as advertisement timer values configured on Company router A.

First time it will take time to make adjacencies between routers then in case of interruption in primary link it will switch to secondary link with-in 1 second. Most important thing is when all routers configuration get done all BGP tables must get cleared forcefully for fresh adjacencies. Once adjacencies get completed BGP is ready for session building with protection. This solution will really help people to maintain the sessions of BGP easily. Though it is a costly solution but uninterruptable communication is possible for highly demanding corporations.

6 Discussion

In the early days, the Internet used static routes, but very quickly network administrators couldn't keep manually updating their routing tables. Later, Internet network administrator used a protocol called EGP but it was not scalable so BGP was introduced to solve all problems. BGP stands for Border Gateway Protocol and the current version is BGP4.

BGP is a routing protocol that runs on routers. BGP allows fully decentralized management of the Internet. That means, if BGP router is on the Internet, it can tell all other routers that what networks are available to everyone in the world. BGP calls each routing domain an autonomous system (AS). It selects the best path, through the Internet, by choosing the route that has to navigate the fewest autonomous systems. In short, Internet world is incomplete without running BGP but there are major vulnerabilities exist in BGP such as slow convergence and abnormal termination of session. However, Simplifying BGP design complexity helps in research, for the root cause analysis.

Multi-homing technique is used widely to overcome such loopholes but simply implementing Multi-homing technique is not quite enough, it requires more to handle these issues properly. By implementing multi-homing rapid switching between links will get done that will badly affect sessions maintenance. It requires to be aggregated with some other techniques to get proper benefit from it. According to my research implementing Multi-homing OER must be embedded with reduction of convergence timing will really help to overcome major vulnerabilities.

By default convergence time of BGP neighbors are 60:180 seconds that cause a very long delay in corporate running environment and adversely affect the sessions. Whenever link is out of order will take same specific time interval for the convergence which results in session loss and loss of data. It is important to reduce the convergence time, up to 1:3 seconds between neighbors to overcome the loophole of BGP.

By manipulating the time intervals of BGP in parallel using OER technique because if we don't implement OER, rapid switching between links will generate same affect on sessions and data will be lost. OER helps BGP to stick on backup running link despite jumping over primary link when it gets restored. It is the nature of BGP it sense if primary link get restored it will break all sessions and build new sessions through primary link and traffic will go on which cause loss of data so OER is a mandatory part of this solution. Basically it is a triangle **"Manipulation with timers + OER + Multi-homing "**every technique used in the solution is necessary for each other. Skipping one of it will lead to the failure it is important that each technique must be used very carefully and in proper manner to get fruitful result.

7 Conclusion

In this Era BGP becomes very sophisticated and complex protocol and it is believed that BGP is playing vital role for the connectivity of internet all over the world. Without BGP it is impossible to maintain connectivity of such a mesh and worldwide network. Utilizing the BGP in best optimal way there so many advantages but unfortunately there are some dark areas discovered in BGP that are affecting the performance of protocol directly & indirectly. There are so many Major vulnerabilities of BGP but we are discussing here some of them i.e. (slow convergence, abnormal termination of sessions) these can be triumph over by the implementation of different techniques in appropriate manner such as Multi-homing, Single-homing, and OER .In several condition it is necessary to have multiple paths to internet. When the applications are accessible over the internet companies mostly use multi-homed BGP networks. If any one of the two links fails then traffic will get stop and no reach-ability information will be passed through the failed link but BGP reach-ability information will be advertised by the small ISP on the other link. OER uses a router as a master controller and it communicate with one or more border router that to ISP or WAN provider. By default BGP timer keep-alive is 60 seconds and hold –time is 180 seconds .These timers work fine in most scenarios but if any reason quick response is needed than the timer can be reduce.

According to this research major vulnerabilities of BGP by embedding Multi-homing with OER and manipulating the timers in an appropriate manner by default convergence time of BGP neighbors are 16:180 seconds that cause a very long delay and affect the session. With Timer, multi-homing and OER technique we reduce the convergence time up to 1:3 second between neighbors. It should be handled very carefully and must be defined at interface level through which links are connected. Links must contain equal bandwidth to share the load easily incase if one link tear off. OER play a key role for the BGP traffic and never let BGP to switch back in case of restoration of first WAN link. After successfully passing the peak hours we can force fully divert traffic to primary WAN instead of running secondary link as primary. Theory of suppressing the flaws of BGP has

been successfully implemented on simulator GNS3 to prove "***session maintenance and swift convergence of BGP*** ". It will really help the world wide to get or transfer data fearlessly and minimum time interruption will occur, sessions will be restored in no time.

References:

[1] BGP Student Guide Vol.1 version 3.2, Cisco Press.2010

[2] Cisco Systems, Inc. *Compatible Systems Setup Guides: BGP Configuration Guide2010*,http://www.cisco.com/warp/public/707/cscsupport/setup_guides/bgp.html#bgpRouteMap, Accessed on 10th Feb, 2011.

[3] P. Francois, O. Bonaventure, B. Decraene, and P-A. Coste, "Avoiding disruptions during maintenance operations on BGP sessions," in IEEE TNSM, vol. 4, no. 3, 2010.

[4] C. Pelsser, T. Takeda, E. Oki, and K. Shiomoto, "Improving route diversity through the design of iBGP tolopogies," in Proc. IEEE ICC, pp. 5732-5738, 2010.

[5] L.Wang, M. Saranu, J.Gottlieb, and D.Pei. Understanding BGP Session Failures in a Large ISP . In INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE, pages 348-356, 2010.