Improving Internal BGP Provide Fast Failover in Multi-homing Environment Mobile Backhaul

Wendi Usino¹, Hillman Akhyar Damanik², Merry Anggraeni³
¹Faculty of Information Technology Budi Luhur University, Jakarta, Indonesia
²Dept. Master of Computer Science, Faculty of Information Technology Budi Luhur University, Jakarta, Indonesia
³Dept. Master of Computer Science, Faculty of Information Technology Budi Luhur University, Jakarta, Indonesia

Abstract. Integrating multipath technology with inter-domain routing is one of the most promising trends in building a package routing policy system for the next generation. However, gradually changing the current inter-domain routing from a single path to multipath failover routing is a difficult problem. In the previous study, the application of the BGP algorithm was applied only to connections on external BGP techniques or between AS numbers. In this paper, we study the impact of implementing traffic routing policies on backhaul carrier Ethernet (Metro-E). Implementing iBGP as a medium on Metro-E for failover technology. The purpose of this paper is to propose and implement routing policies including policy expressions, firewall filters, route preferences, policy chain and policy statements in failover links in IGPs on routing deployment policies, using the Juniper Router Operating System. The results of this application show that periodically the intervals in the graph clearly show a direct correlation between the average failover time. When the main link fails, the link will be connected or active, secondary links and tertiary links will be created. Likewise, switching paths from secondary links to tertiary links, and so on will choose random (round robin). So that during the transfer process does not require time, it can be interpreted as 0 seconds, or not packet loss. So if the average failover response for direct implementation requires only 0-1 seconds.

1. Introduction
Interior Gateway Protocol (IGP) protocol BGP policy provides a set of rules that define how the AS Number directs traffic in and out of the Internet. In BGP, only one route is selected and advertised for each specified network destination [1]. IBGP session connects two speakers including the same AS [2]. Basically the IBGP concept is applied to two or more gateways to forward packets originating from the source address in exchanging information. Internet service providers (ISPs) run the Internal Border Gateway Protocol (i-BGP) to distribute inter-domain routing information between their BGP routers to internet [3] [4] [5]. This paper presents the application of IBGP to the Metro-E environment in backhaul transmissions. This research paper presents a transmission scheme on layer 2 local loop Metro-E, and combines with Layer 3 Routing Protocol Internal Border Gateway Protocol (i-BGP). On transmissions that continue data traffic communication, to be able to carry the backhaul communication connection to the destination. The merging scheme applied to the Metro-E and i-BGP lines is used because of its ability to choose the best route to the destination, especially in Multi-homing environments. The fault tolerance connection is highlighted on the challenge to the Multi-
homing network. Application of i-BGP is carried out in multi-homing where the results present limited reliability in guaranteeing customer data. Failover techniques were introduced to provide realization of reliability on the network while traffic sharing techniques were introduced in realizing scalability. The purpose of this paper is to propose and implement routing policies including policy expressions, firewall filters, route preferences, policy chain and policy statements in failover links in IGPs on routing deployment policies, using the Juniper router Operating System. The concept of backhaul communication in the Ethernet carrier environment will be applied to layer 3, Metro-E will work on layer 2 communication, by classifying each i-BGP source address to destination peering. With the same autonomous number, which is 45699. In detail the application will be explained as shown in figure 1 as follows:

![Figure 1. Metro Ethernet and i-BGP Backhaul Architectural Layers End-to-End Scheme](image)

2. Benefits of Integrating multipath Metro Ethernet and i-BGP Backhaul

Transport and transmission networks consist of multiple layers network, technologies, and areas of deployment [6]. For example, integration and combination of layer 2 and layer 3 for implementing fault links. There is a possibility that lead to network failover, such as link or node failure, administration change, setting of IGP is overloaded and path optimization [7]. In addition to link failure, link congestion is also a challenging problem for network service provider [8]. In the implementation of Integrating multipath technology failover this is very beneficial for providers, because with the devices available at the current technology, loses link connectivity on the primary link is replaced by the other link backup [9]. Failover in this paper will be carried out when this failover occurs, the network must converge before traffic will be able to pass to and from the network segment that incurred link failure [10].

3. Research Method

The methodology of this paper is to propose, develop and implement routing policy on the internal Border Gateway Protocol (i-BGP). In routing the distribution and carrying backhaul traffic that stands on the Metro-E network, using i-BGP technology itself, namely policy expressions, firewall filters, route preferences, policy chain and policy statements in failover links in IGP on routing deployment policies.

3.1 Topology Network Configuration Schema and Method

3.1.1 Evaluating Complex Cases Using Policy Chains and Routing Decisions

The Concept of implementation that will be applied is evaluating complex cases using policy chains and subroutines. Figure 2 shows how the routing policy chain is evaluated. This routing policy consists of several terms. Each term consists of match conditions and actions to be applied to a suitable route. Each route is evaluated against the policy as follows:
Procedure:

1. The route is evaluated first term: If it matches the first route, the specified action will be taken. If the action accepts or rejects the route, the action is taken and the route evaluates and ends. If the next term action is determined, if no action will be determined, or if the route does not match, then the evaluation continues as described in Step 2. If the next policy action is determined, the act of accepting or rejecting the terms specified will be skipped, all terms remaining in the policy will be skipped, all other actions taken, and the evaluation continues as explained in Step 3.

2. The route is evaluate the second term: If the routing matches, then action will be taken. If the next action is to accept or reject the route, the action will be taken and the evaluation and route will end. If the next term no action is to be determined, or if routing does not match, the evaluation will continue in the same way in the first routing policy. If the next policy action will be determined, the act of accepting or rejecting what is specified in this term will be skipped, all remaining terms in this policy will be skipped, all other actions taken, and then continued as explained in Step 3.

3. The route is evaluated third term: If route does not match the term or match, with the term with the action to be performed, at the next policy route in the first routing policy, the next will be evaluated against the first term in the second routing policy.

4. The route is evaluated fourth term: Evaluation will continue until the route matches a term with the action of accepting or rejecting a predetermined one or until there is no more policy on each route to be evaluated. If there is no more policy on routing, then accept or reject action that will be determined by the default policy is taken.

The implementation of failover for routing decisions will be used by using the preference value. The value route preference (Administrative Distance) is a value $2^{32} - 1$. In table 1 it contains several routing rules and priority rules with multiple i-BGP gateways.

| Destination | Routing Decision Gateway | Distance Preference | Priority | Term       |
|-------------|--------------------------|---------------------|----------|------------|
| 10.17.0.0/16 | 172.30.0.178/30          | 100                 | 1        | Primary Link |
| 10.17.0.0/16 | 172.30.0.126/30          | 150                 | 2        | Secondary Link |
| 10.17.0.0/16 | 192.168.252.6/30         | 200                 | 3        | Tertiary Link |
Figure 3 Routing rules and priority is to determine which routing path is a priority and which becomes a backup path. The smaller the distance, the rule will be prioritized. A routing rules and priority proposed can be a graphical representation of preference distance values.

![Figure 3. Routing Rules and Priority](image)

3.1.2 Route Filters and Radix Tree
A route filter is a collection of matching prefixes address. When determining the start of a match route, can then determine the exact match with a particular route. Then configure the general actions that apply to all lists or actions associated with each prefix address. The route-filter option is used to match an incoming route address to destination match prefixes address of any type except for unicast source addresses. In a route filter and detailed radix tree for the route 10.17.0.0/16 specify actions in three ways procedure:

| Procedure |
|-----------|
| 1. source-address-filter option—these actions are taken immediately after a match occurs, and the then statement is accept and reject. |
| 2. then statement—these actions are taken after a match occurs but no actions are specified for the route-filter or source-address-filter option. |
| 3. Route filter match types prefix list is exact (radix tree) schema: |
| - route address shares the same most-significant bits as the match prefix address (destination-prefix or source-prefix). Number of significant bits is described by the prefix-length component of the match prefix address. |
| - prefix-length component address of the match prefix is equal to the route prefix length address. |
| - route-filter or source-address-filter statements: |
| [edit policy-options policy-statement policy-name term term-name from] route-filter destination-prefix match-type exact { actions; } |

The operation process on the route filter to be used, the router device will be configured by classifying matching binary numbers known as the radix tree (trix radix). Figure 4: the radix tree tree using binary search to identify the IP Address (route) address for the route filter. The proposed radix tree can be a graphical representation of these binary numbers.
3.1.3 Routing Policy Match Conditions

A policy match condition defines the criteria that a route must match. Match conditions include communities, prefix list address and AS path. Policy match conditions, each term can consist of statements, from that define match conditions: In the from statement, we define criteria that an incoming internal group address must match. And we specify one or more match conditions. If specify more than one, all conditions must match the route internal group for a match to occur. Each term in a routing policy can include statements from to define the conditions that a route must match for the policy to apply:

```
from {
  family family-name;
  match-conditions;
  policy subroutine-policy-name;
  prefix-list name;
  route-filter destination-prefix match-type <actions>;
  source-address-filter source-prefix match-type <actions>;
}
```

3.2 Proposed Design Integration Internal BGP Provide Failover in Multi-homing Environment Backhaul Metro-E

Improving Internal BGP Provide Fast Failover in Multi-homing Environment Mobile Backhaul, and paper design research specifications are modeled as Figure 5. We study the impact of implementing routing policies on Ethernet carrier (Metro-E) by implementing i-BGP as a medium over Metro-E for failover technology. The design and configuration that will be modeled is end to end from the subscriber site to the mobile backhaul subscriber core. In the Ethernet carrier link the bandwidth width allocation is 50Mbps, which will bring subscriber site traffic to the core subscriber.

In the context of this paper, Layer 2 E-Line will be used for backhaul traffic from the customer's site to the Layer 3 service as shown in Figure 5. The standard Ethernet carrier configuration for each port will be setup tagged port that will bring traffic to the communication i-BGP. In each router ASN 45699 will be set up. Then the failover configuration will be setup by internal BGP peering session, evaluating complex cases using policy chains and subroutines, route filter and radix tree and routing policy match conditions. These methods will result in a failover link on the Ethernet carrier, to bring backhaul traffic.
3.3 Failover Functionality and Recovery

One of the main objectives of this research paper is to provide results on how the failover function is available in carrier Ethernet backhaul. Preserve connectivity across the backhaul network. For testing purposes, network failover events are limited to link failures and nodes, on the Metro-Ethernet interface port and sub-interface ge-0/0/0.665-ge-0/0/0.667-ge-0/0/0.669.

There are two options for testing, on link failures in the failover technique that is applied. The first option is to deactivate the ports on the network interface from the interface. The second option is to issue a "shutdown or deactivate" command on a particular network interface (sub-interface ge-0/0/5-ge-0/0/6-ge-0/0/7) through the router command line interface. Both options are tested and the results will be described in the Excel table and mapped on the graph.

The two testing techniques that will be carried out are the implementation in the real, when one link fails. The table below shows procedures for simulating links and node failures and recovery.

| Procedure | Details |
|-----------|---------|
| 1. | First, the routers on each internal BGP peering will choose the most specific routing rule with dst-address. |
| 2. | Then the Router will see the value in the preference parameter of each routing rule, the smaller the preference, the rule will be used. |
| 3. | If there are several routing rules with specific dst-addresses and similar preferences, then the Router will choose Random (round robin). |
Table 2 below shows the procedures to test conducted a link and node failure (Primary Link: interface ge-0/0/5) as well as recovery.

| Table 2. Procedures to test conducted a link and node failure |
|---------------------------------------------------------------|
| :configure :configure :configure |
| Entering configuration mode Entering configuration mode Entering configuration mode |
| :deactivate interfaces ge-0/0/5 :deactivate interfaces ge-0/0/6 :deactivate interfaces ge-0/0/7 |
| :show interfaces ge-0/0/5 :show interfaces ge-0/0/6 :show interfaces ge-0/0/5 |
| ## ## ## | ## | ## |
| **# inactive: interfaces ge-0/0/5** | **# inactive: interfaces ge-0/0/6** | **# inactive: interfaces ge-0/0/7** |
| **#** | **#** | **#** |
| description Primary-Link-Preference-100 | description Secondary-Link Preference-150 | description Tertiary-Link Preference-200 |
| unit 0 { | unit 0 { | unit 0 { |
| family ethernet-switching { | family ethernet-switching { | family ethernet-switching { |
| port-mode trunk; | port-mode trunk; | port-mode trunk; |
| vlan { | vlan { | vlan { |
| members; | members; | members; |
| } | } | } |
| } | } | } |
| | | |

Testing is done with the date and time details in the following table 3:

| Table 3. Date and time recovery |
|---------------------------------|
| Link Metro-E Backbone | Date and Time Failure | Status | Status Recovery | Date and Time Recovery |
|----------------------------|------------------------|--------|-----------------|------------------------|
| Primary Link | 2018-12-21 10:00:00 | Ping Timeout | Running Preference 150 (Secondary Link) | 2018-12-21 10:00:00 |
| Secondary Link | 2018-12-21 22:00:00 | Ping Timeout | Running Preference 200 (Tertiary Link) | 2018-12-21 22:00:00 |

In testing the response failover time, how long it will be tested response failover time, or the time the data path used is used from the primary link (preference 100) to the secondary link (preference 150) and tertiary link (preference 200) with the condition, only one Metro-E link serves to channel data traffic and other Metro-E links as backups if the primary link (100 preference) fails.

4. Evaluation Result and Discussion

4.1 Traffic Monitoring Packet Latency
Based on the readings collected, each for the three failover links and recovery time in each test case, testing traffic monitoring, with ICMP packets and peering sessions viewed by date and time on the
primary link, experiencing link failure. In experiments each set of results is followed by detailed analysis. ICMP packet sending mechanism will be based on the primary path as a headend, and where the backup path will only be established after a link failure or node, on the primary link, then the secondary link will recover and if the secondary link path occurs a link or node failure, then the link tertiary will do recovery automatically. The status of the reserve path will be active and standby, with the concept of an internal BGP peering session (i-BGP).

Figure 7. ∑ Ping Latency (Primary Link, Secondary Link and Tertiary Link)

Figure 7 above shows the average failure time, at the primary link and recovery time with respect to the time interval in the log i-BGP state transition events. The graph clearly shows a direct correlation between average failover time, on ICMP packet sending. The failover time for the default i-BGP peering session with the default value of the waiting interval is 5 seconds. The failure time on the primary link does not affect the secondary link and tertiary link. When the primary link fails, the link will be connected or active, the secondary link and tertiary link will be established. In this state, BGP exchanges update packets with its peers and the hold timer is restarted at the receipt of the update or keep the message when it is not set to zero. If errors are found, the notification is sent to the peer, and the state falls back to idle session. So that the link path from the primary link to the secondary link doesn't take time. Likewise, the displacement of the path from the secondary link to the tertiary link, and so on will choose Random (round robin). So that during the transfer process does not require time, it can be interpreted as 0 seconds, or not the packet loss. So if averaged response failover for direct implementation only takes 0–1 seconds.

4.2 Traffic Monitoring ∑ inbound and Outbound

Figure 8. ∑ Index Traffic Bandwidth Operational Primary Link (Link Traffic Drop)
Figure 8 above shows in 2018-12-21 10:00:00, inbound and outbound traffic on primary link drop. Periodically, secondary intervals and tertiary links will do recovery, on primary i-BGP links by forming TCP sessions with neighboring routers called peers. BGP uses Finite State Machine (FSM) to maintain tables of all BGP peers and their operational status. i-BGP secondary and tertiary link sessions are specified as established status. In this condition of recovery, i-BGP sessions are established, on secondary and tertiary links. The i-BGP neighbor exchanges routes through update message. When update and the keepalive message are received, hold timer is reset.

Figure 9 above shows that at the same time, the primary link has a link failure, the status of the secondary link is recovering, by forming a TCP session on peers inet.0: 9/10/9/0 (172.30.0.178). In figure 9 above, too, shows 2018-12-21 22:00:00, inbound and outbound traffic on secondary link drop. Periodically the tertiary link interval will do a recovery, like Figure 10. By forming a TCP session on peers inet.0: 0/10/9/0 (192.168.252.6).

Figure 10
5. Conclusion

The application and testing that we did in this paper, we present a quick link failure recovery mechanism or a quick switchover mechanism to handle link failures and connect congestion in the Metro-E backhaul network, at layer 3 by implementing peering sessions on i-BGP. In a fast failover mechanism, we use the value of internal BGP advertisements, namely default value is 5 seconds for internal peers. The value of distance in preference will calculate the parameters of each routing rule, the smaller the preference, the rule will be used. This value will calculate periodically, for the path for each source-destination pair and proactively install flow entries and group entries in the corresponding switch. When a link or link fails, each session (neighbors) on i-BGP can restore the affected stream to the backup path. The experimental results show that the average recovery time of the failover mechanism is significantly faster. The failure time on the primary link does not affect the secondary and tertiary link. When the primary link fails, the link will be connected or active, the secondary link and tertiary link will be established. In this case, the BGP algorithm protocol will swap the update package with its peers and the time delay starts again when it receives an update or saves the message when it is not set to zero. If an error is found, a notification is sent to the peer, and the state returns to Idle status. So that the link path from the primary link to the secondary link doesn't take time. Likewise, the displacement of the path from the secondary link to the tertiary link, and so on will choose random (round robin). So that during the transfer process does not require time, it can be interpreted as 0 seconds, or not the packet loss. So if averaged response failover for direct implementation only takes 0-1 seconds.

References

[1] Abu Hena Al Muktadir, Kenji Fujikawa, Hiroaki Harai, "Route advertisement policies for border gateway protocol with provider aggregatable addressing" High Performance Switching and Routing (HPSR) 2016 IEEE 17th International Conference on, pp. 42-48, 2016.
[2] Vissicchio, S., Cittadini, L., & Di Battista, G. (2015). On iBGP Routing Policies. IEEE/ACM Transactions on Networking, 23(1), 227–240.
[3] Santhosh, S., Dakshayini, M., Tech, M., & Student, C. N. E. (2016). Effect of Route Reflection on iBGP Convergence and an approach to reduce convergence time, 4(8), 4530–4535.
[4] Border Gateway Protocol To Enhance Failover In Multihoming Environment By Muhammed Zaharadeen Ahmed A dissertation submitted in fulfilment of the requirement for the degree of Master of Science (Computer and Information Engineering) Kulliyyah of Engineering International Islamic University Malaysia. (2017), (March).
[5] Pelsser, C., Takeda, T., Oki, E., & Shiomoto, K. (n.d.). Improving Route Diversity through the Design of iBGP Topologies.
[6] Mirkhanzadeh, B., Shakeri, A., Shao, C., Razo, M., Tacca, M., Galimberti, G. M., Fumagalli, A. An SDN-enabled multi-layer protection and restoration mechanism, (2018).
[7] Shukri, A., Noor, M., & Mat, M. (2013). Fail-stop failure recovery in neighbor replica environment. Procedia - Procedia Computer Science, 19, 1040–1045.
[8] Lin, Y.-D., Teng, H.-Y., Hsu, C.-R., Liao, C.-C., & Lai, Y.-C. (2016). Fast failover and switchover for link failures and congestion in software defined networks. 2016 IEEE International Conference on Communications.
[9] Peer, A. M., & Registration, C. (2007). A Transparent Failover Mechanism for a Mobile Network with Multiple Mobile Routers, 11(7), 604–606.
[10] Dwyer, J., & Jasani, H. (2013). An Analysis of Convergence Delay Caused by Link Failures in Autonomous Systems, 1–6.