Identifier and Locator Separation Based Site Multi-homing Path Failure Recovery

Feng Liu, Cheng Gong, Rui Tu and Huatao Tang

1 School of Electronics Engineering, Naval University of Engineering, Wuhan, Hubei, China
2 No. 91001 Armed Forces, Beijing, China
Email: trbomb@163.com

Abstract. Multi-homing is one of the effective ways to defeat path failure and increase the reliability of site network service. However, limited by the TCP/IP architecture, multi-homing has not been well deployment. In this paper, we proposed a sit multi-homing path failure recovery method based on “Locator/Identifier Split” by adding mapping service and RTT cooperative detection mechanism to the edge router. In order to reduce the probing’s cost of the router’s resource and link bandwidth, we also optimized the probing algorithm by reducing unnecessary active probing. The results of simulation test proved that this method could effectively detect the path failure and performance drop including packets delay and packets loss, and made a fast path switch to ensure the normal running of the upper application services (e.g., FTP). We also gave a theory analysis of the overhead brought by the sit multi-homing path failure recovery method. It shows that the addition bandwidth cost is 0.036% of the simulation link bandwidth and our approach is practicable.

1. Introduction
The path failure will cause the network application interruption and quality of service decline. The probable reasons are path failure, network equipment failure, DoS even the packet loss and delay increase caused by traffic burst. There are two types of path failure: complete failure and performance decline. The first one is a serious path failure which will cause most of the applications break down, the other one will cause performance degrade that still in the allowance range (e.g. throughput rate decline, delay increase and packet loss rate increase).

It is inefficient for the routing architecture based on routing protocol to resolve path failure. For example, it will cost BGP several minutes to recovery the routing from a path failure [1-2]. During the time, the upper application service will be interrupted. Moreover, the routing protocol concentrate more on the complete failure than on the performance decline closely related with the path quality (e.g. delay and packet loss). When the routing protocol selects a path, it is probably not the optimal path for the application.

The Multi-homing is one of the main methods to reduce the path failure influence on the application service. It can automatically switch to the backup path to ensure the application not to interrupt by maintaining several backup paths. But it is hard to implement and deploy Multi-homing due to the limitation of the TCP/IP network architecture. One of the main reasons is the IP Overload problem which implicated both the network node’s topological location and identifier. The IP Overloading problem hampered the application of some new network technologies such as multi-
homing, traffic engineering [3-4] and caused the critical challenges to network scalability, mobility and security [5].

In order to resolve the IP Overloading problem, the IAB proposed the “Locator/Identifier Split” method [6] to import two name space to represent the node’s location and identifier separately. After importing “Locator/Identifier Split” method, the users in the edge network can use PI (Provider Independent) address to implement multi-homing, traffic engineering and anycast. Some renumbering costs caused by PA (Provider Allocated) address will be avoided, e.g. changing the access control policy and configuration based on IP address.

There are two types of the multi-homing: site multi-homing and host multi-homing. The site multi-homing can support large hosts number and complex routing architecture, which satisfies the needs of the big commercial organization. But the end users can’t setup the traffic engineering and network configuration according to their needs. The host multi-homing is a single host method, the end user can configure the host and edge network to connect to the preferred core network. It is feasible to implement and deploy edge network multi-homing based on “Locator/Identifier Split” method. Moreover, a path failure detection and recovery method are proposed to ensure the reliability of the inner site services.

This paper had six chapters: the second chapter introduced the related works of the path failure fast detection and recovery, the third chapter introduced the “Locator/Identifier Split” naming and addressing architecture, the fourth chapter gave a detailed description of the site multi-homing path failure recovery, the fifth chapter gave a function test and performance analysis with simulation, the last chapter concluded with summary and new expectation of the works.

2. Related Works

In order to improve the reliability of the application service, on one hand people tried to accelerate the convergence of the routing protocol, on the other hand people tried to implement path failure fast detection and recovery by overlay and multi-homing.

The RON [7] is proposed by David Andersen in MIT, which is a routing infrastructure based on overlay. The RON makes the distribution application be aware of the path failure and periodic performance decline, made a fast recovery in a few seconds. It provides more elastic and fault-tolerant service by RON path selection. The RON is an application layer overlay built on Internet network layer. The RON nodes can monitor the path quality between them and decides whether to forward the IP packets through Internet or by other RON nodes.

David Andersen measured two group RON in the Internet and proved the advantage of the RON. The test implemented a RON wish 12 nodes and 132 paths and lasted 64 hours. During the time, there were 32 serious path failures happened. The test result showed that the RON routing could detect failure, find new path and recover routing in only 20 seconds.

The SCTP (Stream Control Transmission Protocol) [8] is a host multi-homing method. The difference between the SCTP and normal transmission protocols (e.g. TCP, UDP) is that the SCTP provides host based multi-homing and multi-traffic. In order to support multi-homing, the SCTP provides a set of peer node’s transmission addresses for the end node, e.g. a combination of SCTP port and multiple IP address. The association is a new concept in the SCTP. When establishing communication link between two hosts, the association must be built firstly by the cooperation of each host’s multiple interfaces. The multi-homing host has multiple network interfaces to connect multiple path and such paths can be combined into one association. The SCTP uses built-in heartbeat mechanism to monitor all the paths in the association. After detecting a path failure, the SCTP will forward the packets by another path. Such path switching is transparent to upper application.

The SCTP also has some defects. It must modify the host protocol stack to adapt to the changes of the transmission layer protocols. The association is built on a set of static IP address, which is not fit for mobility situation and can’t support traffic engineering and multicast.

The SHIM6 [9] is a host multi-homing method and supports multicast. It uses IPv6 address space to build Locator and Identifier without importing new address space. The SHIM6 uses multiple IPv6
address to build Locator set to the multi-homing host and selects a IPv6 address as the ULID (Up layer ID). The Locator is used to routing on the host interface. The ULID is a static id used to establishing connection for the up layer session (TCP/UDP) between the transmission layer and the network layer. The up layer application uses the existing IPv6 address as the ULID. According to the link’s load, the host’s SHIM6 layer selects the Locator from the Locator set to choose the provider and path. Because the ULID uses he existing address space and never changes, it is not necessary to modify the up layer socket API and the communication session will not be interrupted even the source and destination’s location is changed. The SHIM6 also provides a path fail detection and recovery mechanism [10]. The forwarding plane of the SHIM6 uses REAP protocol to detect path failure and avoids the fault link by reselecting the source and destination’s Locator and ISP.

At present, the SHIM6 can only support IPv6. Moreover, the SHIM6 is too complex and must update host, which hampers its application.

Our path failure recovery method imports the “Locator/Identifier Split” idea of SCTP and SHIM6 with the difference that our method provides network equipment base site multi-homing path failure recovery. We also import the RON’s path failure detection and periodic performance test mechanism to monitor the link quality of the site’s multiple paths.

3. Locator/Identifier Split Architecture
The Locator/Identifier Split naming and addressing architecture is as figure 1 showed. The network is divided two parts: core network and edge network. The core network uses Locator name space and the edge network uses Identifier name space. The hosts can still use the existing address space (e.g. IPv4/IPv6 address) as the identifier not supporting global routing. The host’s up layer communication session is established on the static Identifier with its changeable corresponding Locator. This architecture makes the “Locator/Identifier Split” to be transparent to the end hosts. The host’s protocol stack and application can still use the original address format, which avoids updating large number of hosts.

![Locator/Identifier Split Architecture](image)

Figure 1. Locator/Identifier split naming and addressing architecture.

The network equipment (e.g., edge router) performs the transformation between the Identifier space and the Locator space by querying the distributed mapping service system. When the host moves (Locator changes), the edge router will quickly update the mapping record in the distributed mapping service system.

This architecture is implemented by “mapping + encapsulation” method. After receiving the packet from the host, the edge router queries the mapping service system to get the Locator record according to the packet’s Identifier. The edge router adds a new packet header included the Locator to the
original packet and forwards the new packet through the GRE tunnel (Generic Routing Encapsulation) [11] between the routers. In the new packet, the inner packet header’s source and destination address is Identifier (IP address) and the outer packet header’s source and destination address is Locator. When the encapsulated packet arrives the destination edge router, the router decapsulates the packet and forwards it to the destination according to the Identifier.

4. Path Failure Recovery
The path failure recovery method can monitor the path condition and performance, which is based on the site multi-homing. On one hand, the up layer communication session can switch on the multiple transmission paths, which makes a fault-tolerant network; On the other hand, the path selection is based on the packet delay and packet loss which are closely related with the application performance. So we can select the optimum flexibly transmission path according to the application’s needs.

4.1. Framework
The Locator/Identifier Split naming and addressing architecture provides the foundation of the site multi-homing as figure 2 showed. The site edge router is connected to multiple ISPs and gains multiple Locators. The edge router can build multi-homing mapping records for the site inner hosts in the distributed mapping service system. That is a site host Identifier is mapped to multiple core network Locator. In the multi-homing mapping records, one Locator is the prime Locator and the rest of the locators are backup. So the site has multiple forwarding paths across different ISPs. Because the Locators are assigned by different ISPs, the corresponding paths have few overlaps, which ensures the multiple paths’ heterogeneity and provides the foundation for the path switching.

![Figure 2. Site multi-homing based path failure recovery.](image)

When the edge router queries the mapping records and finds that the destination host has multiple Identifier’s corresponding Locators, it will use the prime Locator to forward packet and monitor every Locator’s forwarding path condition and quality. As the figure 2 showed, the edge route 1 queries the host B’s Identifier (IDB) and finds that the host B is in a multi-homing site. Then the edge router 1 adds the host B’s two Locators (Locator2_ISP1, Locator2_ISP2) to the monitor list and sends probe packet to monitor the two paths across the ISP1 and ISP2.

When the prime Locator’s path is interrupted, the current forwarding path will switch to the backup Locator’s forwarding path as figure 2 showed. Assume the Locator2_ISP1 is destination host B’s prime Locator, the packets between the host A and B are forwarded by ISP1 normally. If the edge router 1 finds that there are faults or performance decline in the ISP1’s forwarding path and the ISP2’s forwarding paths is still work, it will encapsulate the packet with the Locator2_ISP2 as the host B’s Locator. So all the packets will be forwarded through the ISP2’s path.
4.2. Edge Router

4.2.1. Basic Structure and Probing Mechanism. In our method, the component and function of the core router are as same as the current core router, but the edge router has made great changes. The edge router is the core of the path failure recovery method. It sends probe packet to test the link quality of the path between it and the multi-homing site and selects the best path to forward packet.

As figure 3 showed, the edge router not only provides packets forwarding but also adds the mapping service query/setting, probing and probe member management function module. The mapping service query and setting module can communicate with the distribute mapping service system to get or update the mapping records. The probing module is used to monitor the link quality of the destination multi-homing’s multiple paths. The probe member management module can maintain the probe member list (Locator) and save the detail information (e.g. delay, packet loss) of the multiple paths. The packet forwarding module determines the forwarding path according to the mapping information and the link quality information.

In order to control the size of the probe members and reduce the probing cost of the edge router, the path failure recovery method uses a data driven probing mechanism which only probe the active multi-homing site’s locators. When the edge router forwards packets, after querying the Identifier and knowing that the destination is a multi-homing site, it adds the site’s multiple Locators to the probing list. Each probing Locator has its life time. If there is no Identifier query about one Locator during its whole life time, then the Locator will be deleted from the probing member list and stops the monitor on the locator’s path.

4.2.2. Optimization of the Probing Mechanism. With the wide application of the multi-homing, both sides of the communication need to probe the link state and quality and perform the path selection. The path failure recovery method tests the path state and performance by measuring the RTT of the probing packets. If both sides of the communication are multi-homing site which has multiple Locators, then both sites must send probing packets to measure RTT. So a RTT cooperation probing mechanism is presented.

The RTT cooperation probing process is as figure 4 showed. The edge router probes one of the multi-homing site’s Locators and the initial probing packets carry the ID. After the peer edge router received the packet, it sends the response packet which includes the initial packet’ IDi and the response packet’s IDR. When the router receives the response packet, it can compute the probing packet’s RTT (8 seconds) by comparing the packet’s ID and the send time. The router regards the received response packet as the probing packet and also sends the response packet to the multi-homing site. So the site can also compute the RTT (9 seconds). Both sides of the communication can compute the RTT separately with 3 packets, which is fit for the both sides are the multi-homing site has path failure recovery capability. On one hand, the duplicate probing packets can be reduced. If both sides perform the independent probing, then it will send at least 4 packets to measure the RTT. On the other hand, the time synchronization can be avoided, both sides compute the RTT with their local time.

In order to further reduce duplicate probing, we analyzed the different probing attributes of the RTT cooperation probing. Some terms are imported as followed:

Term 1 Active Probing: actively probes the Locators in the probe member list on the set frequency.
Term 2 Passive Probing: passively response other router’s active probing to send probing packets.
Term 3 Initial Probing Packet: the first packet of the active probing which only carries itself ID.
Term 4 Response Packet: the packet as the response to the initial probing packet and the other response packet which carries not only itself ID but also responded packet’s ID.

Assume the router A and B are multi-homing site’s edge routers and there exists traffic between the two sites. The router A and B perform path failure recovery probing at each other and the active probing time interval is $\Delta T_A$ and $\Delta T_B$. So in the period of time T, the RTT update count of A and B is $T/\Delta T_A + T/\Delta T_B$. The actual update count is far beyond the count of A or B itself.
In order to reduce the probing’s cost of the router’s resource and link bandwidth, unnecessary active probing should be avoided. So each router needs to maintain two Locator probing list: active probing list and passive probing list. When the router receives the peer router’s active probing packets, it adds the peer router’s Locator to the active probing list. Before the active probing, the router queries the passive probing list. If the Locator is in the list, the router will not perform active probing and not add the Locator to the active probing list. So there is only one active/passive probing between the couple multi-homing site routers, which reduces the resource costs.

### 4.3. Path Selection

The path selection mechanism is used to fulfil the path performance requirements of the up layer application. The edge router probes and maintains the destination multi-homing site’s multiple path state information and decides whether the primary path is active. Moreover, the edge router estimates the backup path quality and selects the best path according to the path state and application requirements. The default two path quality index are delay and packet loss rate. The up layer applications can customize the path quality index based on their needs.

The probing packet is built to probe the path state and quality of the forwarding path. The probing packet is based on UDP and is sent to the peer multi-homing site with the setting frequency. When the peer router receives the probing packet, it returns the ACK message immediately. This method is used to measure the accurate RTT of multi-homing site path and performs real time monitoring on the usability and packets loss of the path. The probing packet’s sending frequency can be changed according to user’s needs.

After the source edge router sends probing packet to the destination router through multiple paths (different Locators), it will wait the destination’s ACK message. During the setting time, if the ACK message is received, then the transmission of the probing packet is completed. If the ACK is timed out, it means that the packets are lost or occurs network congestion. If one path occurs transmission time out repeatedly and reaches the count of Path_Max_Retrans, then the path will be marked as unavailable. As a result, all the packets sent to the destination site will be switched to the backup path.

The packet delay of one path is determined by measuring the RTT. Each probing gets a new delay value and the current path average delay is computer with equation (1). The edge router selects the best path to forward packets by comparing the average delay of multiple paths and the delay limits of the up layer application.

$$Latency_{\text{average}} = (1 - \alpha)Latency_{\text{average}} + \alpha(Latency_{up})$$  \hspace{1cm} (1)

The packets loss rate is computed by counting the probing packets sent or dropped through one path during a setting time interval. The path selection is based on the up layer application’s packet loss rate limit.
5. Performance Evaluation
At first, we tested the fast detection and recovery capability during the path failure by NS-2 simulation. Secondly, we tested the path reselecting capability which can improve the delay and packet loss when the link quality declines. At last, we analyzed the network bandwidth cost by theory computation.

5.1. Simulation Scenario Design
The simulation is run by NS-2 software. We designed the multi-homing site scenario by taking the SCTP’s multi-homing host scenario as a reference. In the NS-2, single node has no multi-homing attribute, so we built a virtual multi-homing node based on several real nodes. In the simulation scenario, the site has two out paths and the edge is the virtual multi-homing node. As figure 5 showed, the virtual multi-homing node has 3 nodes: one master node and two slave nodes. Each slave node connects to one path and transmits real traffic. The master node decides that which slave node is used to forward packets. The master node and slave node are connected through a one-way link. The master node uses this link to control the slave nodes and there is no real packets forwarding in the link. The recovery agent implements the path failure recovery function and is connected with each site node.

5.2. Function Test
5.2.1. Path Failure Detection. As figure 5 showed, n1-n2 is primary packets forwarding path and n1-n3 is the backup path. The experiment simulates an FTP server is running on the multi-homing site with path failure recovery capability. We simulate the primary path failure by cutting the n1-n2 link.

The experiment simulates the path failure scenario and the variety of the FTP throughput rate when close or open the path failure recovery function. As figure 6 showed, the experiment starts for a while and the n1-n2 link is broken. After probing for a period of time, the n1-na path failure is detected and the n1-n3 path is selected to forward packets. When close the path failure recovery function, in case of the primary path failure, the FTP service is down. When open the path failure recovery function, after a short time interruption, the FTP service is recovery. So the path failure recovery method can detect the path failure and resumes the packets transmission shortly, which increase the reliability of the application.

5.2.2. Path Performance Test. The experiment simulates the path performance decline (delay increase, packets loss) caused by network congestion. As figure 5 showed, we add a stable additional traffic (CBR) to the ns-n2 path and make the real traffic is beyond the bandwidth. As a result, the FTP service meets packets loss and delay increase.

As figure 7 showed, when the additional traffic enters into the n1-n2 path, the network congestion occurs, which causes the FTP packets delay increase rapidly. After a period of time probing, the path
failure recovery method finds that the packet delay of n1-n2 path exceeds the setting threshold value and selects the n1-n3 path to forward packets. When network congestion occurs, the FTP service is in the high packet delay situation during the whole time without the path failure recovery mechanism. But when open the path failure recovery function, after a short time high packet delay, the packet delay value returns to normal.

The network congestion not only brings the packet delay but also cause the packet loss. As figure 8 showed, when the n1-n2 path occurs network congestion, the FTP packet loss increases rapidly to reach a peak value and keeps stable. After a period of time probing, the path failure recovery method finds that the packet loss of n1-n2 path exceeds the setting threshold value and selects the n1-n3 path to forward packets. In case of no path failure recovery function, when network congestion occurs, the FTP service is in the packet loss situation during the whole experiment. When open the path failure recovery function, after a short time packet loss, there is no FTP packet loss any more.

\[ \text{Detect}_\text{time}_{\text{max}} = \frac{1}{4} \text{Pr}_\text{obe}_\text{Interval} + \frac{1}{4} \text{Pr}_\text{obe}_\text{Interval} + 3 \times \text{Pr}_\text{obe}_\text{Timeout} \]  

(2)

The average detection time is

\[ \text{Detect}_\text{time}_{\text{average}} = \frac{1}{2} \times \frac{9}{8} \text{Pr}_\text{obe}_\text{Interval} + 3 \times \text{Pr}_\text{obe}_\text{Timeout} \]  

(3)

5.3. Analysis of Probing Performance and Costs

The path failure recovery method monitors the path state and quality by sending probing packet. So the probing frequency is closely related with the path failure detection’s efficiency and costs. Assume each edge router’s probing interval time is \( \text{Probe}_\text{Interval} \) seconds. In order to avoid the self-synchronization of multiple edge routers which can cause probing packets storm, we add a random value between 0 and \( 1/4 \text{Probe}_\text{Interval} \) seconds to the static probing interval time. So the average probing interval time between two probing is \( 9/8 \text{Probe}_\text{Interval} \) seconds.

The probing timeout value is \( \text{Probe}_\text{Timeout} \). If there is no ACK of the probing packet within the \( \text{Probe}_\text{Timeout} \), the packet is regarded as a lost packet. After detecting the packet loss, the edge router sends another probing packet two times at the most. If all of the three packets are lost, the path is regard as unavailable path. So the shortest detection time of the path failure is \( 3\times\text{Probe}_\text{Timeout} \) and the longest detection time is

\[ \text{Detect}_\text{time}_{\text{min}} = \frac{1}{4} \text{Pr}_\text{obe}_\text{Interval} + \frac{1}{4} \text{Pr}_\text{obe}_\text{Interval} + 3 \times \text{Pr}_\text{obe}_\text{Timeout} \]  

(2)

The average detection time is

\[ \text{Detect}_\text{time}_{\text{average}} = \frac{1}{2} \times \frac{9}{8} \text{Pr}_\text{obe}_\text{Interval} + 3 \times \text{Pr}_\text{obe}_\text{Timeout} \]  

(3)
In our prototype system, the Probe_Interval is 5 seconds and the Probe_Timeout is 3 seconds. From the equation (3), the average path failure detection time is 11.8125 seconds. The path performance detection is also related with the application performance requirements. The detection timeliness is different with different application performance requirements. Assume the size of probing packet is Packet_Size and the monitor path number is n, the addition bandwidth cost is

$$\text{Band} \_\text{Width} = \frac{2 \times \text{Packet} \_\text{Size}}{\text{obe} \_\text{Interval}} \times \frac{n}{9} - \frac{\text{Probe} \_\text{Interval}}{8}$$ (4)

In our prototype system, the Packet_Size is 64 bytes, the Probe_Interval is 5 seconds and n is 2. From the equation (4), the addition bandwidth cost is 0.364 Kbps which is 0.036% of the simulation link bandwidth (1M). So the probing packet’s bandwidth cost is very small.

6. Conclusion
The multi-homing is one of the effective methods to resolve path failure and improve site reliability. But the multi-homing is still not widely adopted because of the limits of the TCP/IP architecture. The idea of Locator/Identifier Split if well fit for multi-homing, so a site multi-homing and path failure recovery method based on the A Locator/Identifier Split is proposed. The simulation experiment verifies that our method can effectively detect the path failure and performance decline and makes fast path switch, which ensure the up layer application’s normal running. Moreover, the cost of the path failure recovery method is very small according to the theory computation. In the future, we will deploy the prototype system in the real commercial network to test its performance index.

References
[1] Labovitz C, Bose A and Jahanian F 2000 Delayed internet routing convergence SIGCOMM'00 (Stockholm, Sweden) pp 175-187.
[2] Mateen A, Hussain M and Yahya M 2015 Bandwidth and delay issues on the network routing Networking and Internet Architecture (28) 365-370.
[3] Farinacci D, Fuller V, Meyer D and Lewis D 2013 The locator identity separation protocol (LISP) The Internet RFC 6830.
[4] Henderson T, Vogt C and Arkko J 2017 Host multihoming with host identity protocol Internet RFC 8047.
[5] Tu R, Su J, Meng Z and Zhao F 2008 UCEN: User centric enterprise network IEEE ICACT 2008 (Phoenix Park, Korea) pp 66-71.
[6] Meyer D and Fall K 2006 Report from the IAB workshop on routing and addressing Internet Draft.
[7] Sontag D, Zhang Y and Ohanishayee A 2009 Scaling all-pairs overlay routing CoNEXT (Rome, Italy).
[8] Ong L Y J 2007 An introduction to the stream control transmission protocol (SCTP) RFC 4960.
[9] Barre S, Ronan J and Bonaventure O 2011 Implementation and evaluation of the Shim6 protocol in the Linux Kernel Computer Communications (34) 1685-1695.
[10] Brian C and Habib N 2014 Putting Shim6 into practice ATNAC 2014.
[11] Farinacci D L T, Hanks S, Mayer D and Traina P 2000 Generic routing encapsulation (GRE) RFC 2784.