Design and Implementation of Overlaying Multi-Homing Architecture

Satoshi Uda   Nobuo Ogashiwa
School of Information Science
Japan Advanced Institute of Science and Technology
1–1 Asahidai, Tatsunokuchi, Ishikawa 923–1292, Japan
{zin, n-ogashi}@jaist.ac.jp

Kenichi Nagami   Kuniaki Kondo   Ikuo Nakagawa
Intec NetCore, Inc.
1–3–3 Shinsuna, Koto-ku, Tokyo 136–0075, Japan
{nagami, kuniaki, ikuo}@inetcore.com

Yoichi Shinoda
Center for Information Science
Japan Advanced Institute of Science and Technology
1–1 Asahidai, Tatsunokuchi, Ishikawa 923–1292, Japan
shinoda@jaist.ac.jp

Hiroshi Esaki
Graduate School of Information Science and Technology
The University of Tokyo
7–3–1 Hongo, Bunkyo-ku, Tokyo 113–8656, Japan
hiroshi@wide.ad.jp

Abstract

As a result of the vast Internet growth, a multi-homing technology is being extensively utilized to keep a connectivity to important networks. However, the traditional multi-homing architecture severely impacts routing performance of the Internet making it nonscalable. Furthermore, although it can enhance a reachability in case of connectivity anomalies, it can not be used to enhance an utilization because it can not assign in-coming traffic to multiple links.

In this paper, we propose a new multi-homing architecture which is based on the overlay networking model, providing solutions to these problems. Our architecture improves scalability for increasing multi-homing users and realizes a dynamic and flexible line selection for in-coming traffic, while retaining the original advantage of the multi-homing scheme, which is continuous connectivity to the Internet. Of course, our proposed architecture can be applied not only to IPv4 networks but also to IPv6 networks.

1. Introduction

The Internet has grown drastically, and has become a major infrastructure for communication network. Many networks are being connected to the Internet, and the importance of these networks varies from one to another. Important networks are usually protected by the multi-homing technology, in which a network is connected to multiple ISPs using multiple links to retain reachability to and from the Internet in case of the link failures.

Traditional multi-homing architectures provided redundancy of connectivity to user network by utilizing basic properties of existing routing technologies, in which one link is selected from multiple available links based on a route selection mechanism. They lack flexibility because of the indirect nature of the method, so it is hard to meet user’s requirement. For example, it is especially hard to control traffic to user’s network. Additional problem with traditional architectures is that they accelerate route entry increase on the Internet. Therefore, they are not appropriate for general users of the Internet to benefit from multi-
homing.
In this paper, we will first summarize problems of traditional multi-homing architectures, and propose a new multi-homing architecture based on overlay networking model as a solution to these problems. We will also present experiences from a prototype implementation of our architecture and its operations.

2. Multi-homing Technology

2.1. Traditional Multi-homing Technology

Traditional multi-homing technologies generally provides redundancy in reachability by utilizing basic properties of existing routing technologies. A multi-homing network advertises its address blocks to the global Internet using routing protocols, and packets for this network are delivered through one of multi-homing links based on route information propagated by this advertisement.

The following problems are known to be inherent to this approach.

- Increasing route entries in global Internet.
- Difficulty to efficiently use multiple links.
- Difficulty to operate multi-homed networks.

We will examine these problems in detail.

Increase of route entries First, we discuss a problem about increasing route entries in the global Internet. The number of route entries has been running over 110,000 entries, and it is still increasing. This situation is incurring a grave issue, because a large number of route entries put pressure on a memory space and increase a packet forwarding cost on routers on the Internet backbone. The provider route aggregation technology based on CIDR (Classless Inter-Domain Routing) [6, 4] is been widely deployed for restraining increasing number of route entries.

In multi-homing, each user’s network needs to advertise its address block to all ISPs connected to the multi-homed user for ensuring route reachabilities through any ISPs. So ISPs need to advertise route information for multi-homed users as is, without aggregating. Therefore, the number of route entries in the Internet is increasing one by one, whenever the multi-homed user is increasing (Figure 1).

Current situation is that route entries for small address blocks (prefix length \(\geq 24\)-bits) has grown up to about 60,000 entries. This means that small address blocks including those for multi-homed users occupy over a half of the entire route entries in the global Internet.

Difficulty of efficient utilization Next, we explain a difficulty associated with efficient use of multiple links. It is one of serious weakness in the traditional multi-homing architecture that it is difficult for multi-homed users to control in-coming traffic.

The BGP-4 (Border Gateway Protocol version 4) [7], which is the standard inter-AS routing protocol, is usually used as a routing protocol between multi-homed user and ISPs. However, with this protocol, it is difficult to advertise fine requirements for in-coming packets to peers. For example, it is very hard to control in-coming traffic so that it is distributed evenly among multiple links. Furthermore, link selection would only use source and destination address of the packet, because the selection is based on the IP routing system. For example, if one link from multi-homing has different quality from the other links (e.g. one has wide bandwidth but low quality, the other has high quality but narrow bandwidth), users want to choose a different link according to a kind of applications. It is difficult for the traditional multi-homing technology to comply with these requirements.

Difficulty of operation Finally, we discuss a difficulty of operating a multi-homed network. It is necessary to advertise multi-homing users’ address blocks to the Internet, and the BGP-4 is globally used for this as discussed above. The BGP-4 is a highly functional and complex routing protocol used on backbone networks of the Internet, so it is required for operators to have rich knowledge about Internet routing systems. Therefore, the technical barrier to get advantage of the multi-homing is high, because it is necessary to operate the BGP-4 on the users’ network.

Furthermore, it is necessary to exquisitely adjust the parameters of route advertisements on BGP-4 in order to utilize multiple links effectively. The traffic engineering on
the BGP-4 largely depends on intuition even now, so highly technical backgrounds and experiences are required for operators.

2.2. Multi-homing on IPv6 networks

In response to the problem of increasing route entries in IPv4 networks, route aggregation behavior has been considered carefully in IPv6 (Internet Protocol version 6) [3]. When a site is multi-homing on IPv6 networks, the site is assigned address prefixes from all upstream ISPs. Nodes in the site keep multiple addresses corresponding to these prefixes, and select one address to use when they establish a new session. This way, upstream ISPs can aggregate route entries of downstream sites whether the site is multi-homing or not, solving the problem of increasing route entries.

However, there are weak points in this multi-homing architecture. When a link between a multi-homed site and an upstream ISP fails, connections using addresses with a prefix corresponding to the failed link as a local address cannot sustain the connection. There is a proposed solution for this problem, based on IP over IP tunneling [5], but it should be noted that upstream ISPs support is required in this solution.

Another problem is difficulty to control traffic for a site. In IPv4 networks, administrators of multi-homed sites decide policies of how each packet is delivered to/from the site. But in the IPv6 multi-homing, the routes are decided based on address selection at each end node, making it difficult for site administrators to intervene in route selection.

Therefore, the site-multihoming, e.g. making multi-homing as a site and letting end nodes be unaware of multiple links, is desirable on IPv6 networks also. Requirements and goals for the site-multihoming are discussed in detail in [1].

3. Overlaying Multi-homing Architecture

In this section, we propose the brand new architecture for multi-homing to solve problems of traditional multi-homing technologies, that is not a straight application of routing technologies.

3.1. Overview of the Proposed Architecture

In our architecture, we introduce a route selection mechanism for overlay networking on backbone networks. A user network is connected to route selection mechanisms using tunnels. An address block which is used in user networks is assigned from a CIDR block that is assigned for route selection mechanism. Route information for the address is aggregated and advertised to the Internet from the route selection mechanism. With this scheme, route entry growth for the global Internet can be significantly reduced.

A user network and a route selection mechanism are connected using tunnels for each multihome connection. For example, the user network which is connected to two ISPs (ISP-A and ISP-B) use two virtual links (via ISP-A and via ISP-B) to connect to a route selection mechanism. End points of these virtual links which are user-side of virtual links have IP addresses which are assigned from each ISPs connected to the user network. These IP addresses are assigned from ISP-A for the ISP-A side and from ISP-B for the ISP-B side. Consequently, from the viewpoint of each ISPs, a user network seems to be using IP addresses for each ISP, so that each ISP does not have to be aware of the user being multi-homing.

This architecture is useful because not only it reduces the number of route entries of the Internet, but it also is capable of controlling of incoming traffic to user networks. The traffic for user networks transit the route selection mechanism because the route information is advertised from the route selection mechanism. The route selection mechanism selects and uses appropriate virtual link for each packet based on destination addresses. This route selection process may incorporate results of deep packet analysis, in addition to the standard selection algorithm that is based on destination addresses. For example, selecting a low jitter route for real time application, or keeping traffic load balances of two links are enabled.

Keeping reachability is a crucial matter in case of troubles in multi-homing. In our architecture, a route selecting mechanism and state of virtual links are always watched. When a route selection mechanism detects an amputation of a virtual link, it changes its own route selection rules to stop using an amputated virtual link, so that the amputated virtual link is not used and the connectivity of the user network is kept.

3.2. System Elements

In this section, we show system elements on our proposed multi-homing architecture. Figure 2 illustrates overview of our proposed multi-homing architecture. There are two major elements in our architecture, which are DR (Distribution Router) and UR (User Router) respectively.

- **DR: Distribution Router**
  The route selection mechanism consists of DRs. There are placed on backbone networks, and they are working with cooperating with one another.

- **UR: User Router**
  There are placed on each multi-homed user network, and it terminates virtual links from the route selection
mechanism (i.e., DRs). Furthermore, this router is accepting user requirements such as route selection policy of multi-links, and directing this to the route selection mechanism.

DRs, which are on backbone networks, are advertising route information for each user’s address block, and packets for the multi-homed user network are drawn in a nearest DR (Figure 2). When each DR has received a packet for user network, the DR selects one route from more than one route, and is forwarding the packet to UR on user network through a virtual link on the route.

In the followings, DR and UR are explained in detail with attention to their functionalities.

3.2.1 DR: Distribution Router

A DR is the router that works as a route selection mechanism. Major functions of the DR are as follows.

- Advertise route information for user address blocks.
- Analyze each in-coming packet, and decide a path for the packet.
- Send packets toward user network through virtual links.
- Observe status of each virtual links for user network.
- Filter or shape packets.

Each DR is advertising route information for a user network with aggregating their address blocks. Because more than one DR will be advertising a same address block, packets for the user network are drawn to closest DR from source nodes. In case of DR failure, packets for the user network will be drawn to next closest DR, because the failed DR cease to advertise route information.

When each DR has received a packet for user network, the DR analyzes characteristics of the packet. DRs extract not only destination IP address, but also source IP address, protocol and port numbers, and other deep packet information. Route selection is made according to these information, consulting user specified route selection policies. In this selection process, we can also use status and statistics of each virtual link as a metric. Therefore, for example, it is also possible to control in-coming traffic to keep utilization of multiple links balanced.

Virtual links between DR–UR can be implemented using various technologies such as IP over IP tunneling and MPLS LSP (Label Switched Path) [8]. It is necessary for DRs to provide functions to terminate virtual links, such as establishing or releasing each virtual link, and encapsulating each packet for virtual links. Specific functionalities required on each DR depends on what technology has been adopted.

As mentioned above, persistent confirmation process watches for failures in each virtual link. This process is running on DRs in collaborations with URs if necessary. Furthermore, it enables route selection with considering each virtual link status to notify this detected links status to route selection system in each DR.

It is possible on route selection system on DR not only to select a route for user network but also to discard packets or limit bandwidth of specified flow, based on each packet specification. This is effective for opposing DoS (Denial of Service) attacking to user networks, owing to protecting the links connected multi-homed user network to upstream ISPs from a risk of congestion by filtering any attack packets at DRs. DRs are placed on backbone networks widely, and packets from attackers are delivered separately to closest DR from the source of the packet, so attacking packets of DDoS (Distributed DoS) do not concentrate on a few DRs. This function helps protecting multi-homing users network from the menace of almost all DDoS attacks.

3.2.2 UR: User Router

The UR is the router terminating multi-homing links of at user network. Followings are major functions of UR.

- Terminate virtual links from DRs.
- Monitor status of each virtual link.
- Monitor statistics of each virtual link and each physical link.
- Route selection and transmission for each out-going packet from user network.
- Accepting user’s requests, and relaying them to the route selection mechanisms (DRs).

\[\text{Figure 2. Overview of our proposed multi-homing architecture}\]
Any packets for user network from any source in the Internet is relayed by DR through a virtual link between DR and UR. The termination router of this virtual link on user network side is the UR. The UR has functions for encapsulating each packet, receiving the packet, and functions for establishing or releasing virtual links if necessary.

As discussed with DR’s functions, it is necessary to monitor status of each virtual link. This monitoring may be able to be done solely with DRs, but there may be a case in which it must be carried out by collaborations of DRs and URs. In the latter case, the functions for virtual link status monitoring are required on the UR.

The UR needs to monitor traffic statistics of each physical links used to connect user network to upstream ISPs. Statistics of each virtual links may be monitored to realize fine traffic control. These statistics information is notified to DRs, and used as metrics in route selection for each packet.

In this paper, we have been focusing on controlling incoming traffic to users’ networks so far. Here we mention briefly about controlling out-going traffic from users’ networks. It is much easier to controlling out-going traffic than in-coming cases in the multi-homed network. This type of control is generally possible only on a border router in users network, using techniques such as static policy routing. In our architecture, the UR is the border router in user network, so this function should be in the UR. That is, the UR decide a route for each out-going packet by analyzing the packet to extract packet specifications such as source and destination IP address, protocol type, port numbers, etc. and applying administrator’s policies. Of course, we can use traffic statistics of each link as a metric for route selections. The UR is going to transmit the packet to upstream network by a result of these route selections.

In addition to these functions, the UR also plays a role of user interface. The UR accepts requests such as route selection policies from users or possibly from applications in the user’s network. These requests are relayed to appropriate boxes (DRs, etc.), and will be reflected to packet forwarding processes.

4. Evaluation

4.1. Comparison with the Traditional Multi-homing Method

In this section, we evaluate our multi-homing architecture based on comparison it with the traditional multi-homing architecture. We have selected following points as metrics for comparison. One is from the fundamental purpose of multi-homing:

- Improvement of reachability in case of failures.

And the others are from the current multi-homing users’ demands discussed in Section 2;

- Impacts of increasing multi-homing users to the routing system in the Internet.
- Possibility to efficiently use multiple links.
- Operational costs of multi-homed networks.

First, we compare the effect in an improvement of reachability when a trouble is occurred on upstream networks or links, which is the fundamental purpose of multi-homing. With traditional multi-homing architecture, alternative route is chosen as a result of BGP peering being breaking down and route information regarding the failing link being expired. However, this process would take anywhere between 10 seconds to several minutes, because the BGP-4 is not designed to instantaneously respond to link status changes. In our proposed architecture, the failure on upstream networks or links is detected by using subsistence confirmation between DR and UR. After a detecting the failure, the DR is going to change a path selection table to the user network to leave the failure link out. This process has would take only a few seconds, depending on an interval of subsistence confirmations.

Next, we compare impacts of increasing multi-homing users to the routing system in the Internet. In traditional methods, number of route entries increases by more than one as multi-homing user increase, as discussed in Section 2. In our architecture, it is possible to aggregate route information related multi users when advertising these route information on the route selection mechanism. Therefore, it has robustness in scalability for increasing multi-homing users, because increasing multi-homing users influence scarcely the number of route entries in the Internet.

Third, we show a possibility to efficiently use multiple links that consist a multi-homing. We can only control incoming traffic by regulating metrics for each route if we advertise route information in traditional methods. That is, for example, it is not possible to distribute in-coming traffic among all links equally, nor use an alternative link according to a request from each application. In our proposed architecture, it is possible to flexibly meet the demands from users in link selection, due to flexibility of controlling traffic at route selection mechanism (e.g. policy routing applying user’s request). And, it is also possible to distribute incoming traffic according with feed-back from traffic monitoring mechanism at UR.

And the last, we compared costs for operating multi-homing network. We have already discussed that traditional methods required deep knowledge about the Internet routing system. In our proposed architecture, multi-homing users leave hard operations such as routing engineering to DR’s operator (e.g. operators in multi-homing
Table 1. Comparison the proposed architecture with the traditional method

<table>
<thead>
<tr>
<th></th>
<th>Traditional</th>
<th>Our Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redundancy</td>
<td>Good</td>
<td>Very Good</td>
</tr>
<tr>
<td>Load Sharing</td>
<td>Difficult</td>
<td>Good</td>
</tr>
<tr>
<td>Policy</td>
<td>Difficult</td>
<td>Good</td>
</tr>
<tr>
<td>Simplicity</td>
<td>Good</td>
<td>Fair</td>
</tr>
<tr>
<td>L4 Survivability</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Packet Filtering</td>
<td>Difficult</td>
<td>Good</td>
</tr>
<tr>
<td>Scalability</td>
<td>Bad</td>
<td>Good</td>
</tr>
<tr>
<td>Impact on Routers</td>
<td>Big</td>
<td>Little</td>
</tr>
<tr>
<td>Impact on Hosts</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Operation and Management</td>
<td>Difficult</td>
<td>Easy</td>
</tr>
<tr>
<td>Cooperation with ISPs</td>
<td>Fair</td>
<td>No Need</td>
</tr>
</tbody>
</table>

Table 2. A Number of Route Entries on the Internet (by each prefix length)

<table>
<thead>
<tr>
<th>prefix length</th>
<th>entries</th>
<th>prefix length</th>
<th>entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>17</td>
<td>1,666</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>18</td>
<td>2,991</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>19</td>
<td>8,446</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>20</td>
<td>8,458</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>21</td>
<td>6,011</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>22</td>
<td>8,997</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>23</td>
<td>8,498</td>
</tr>
<tr>
<td>8</td>
<td>21</td>
<td>24</td>
<td>59,378</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>25</td>
<td>78</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>26</td>
<td>87</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>27</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>53</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td>13</td>
<td>98</td>
<td>29</td>
<td>31</td>
</tr>
<tr>
<td>14</td>
<td>256</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>469</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>7,400</td>
<td>32</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 2 shows a number of route entries in the global Internet, sorted by prefix length. We have counted these values by having received routing advertisements from three different ISPs and averaged them. This table show that there are many route entries with prefix length of 23-24 bits, and accounting for a half of all entries. On the other hand, the RIRs (Regional Internet Registries) are assigning global IP address blocks to ISPs which allocation unit is generally wider than 20 bits prefix length block. Consequently, we can easily imagine that ISPs are advertising more specific route entries (punching hole), in addition to the route entries corresponding to assigned address blocks.

The punching hole route advertisements are mainly caused by multi-homing, because it is required for the traditional multi-homing methods to advertise users’ address blocks to the global Internet explicitly without aggregating routing entries.

Table 3 shows an example of an actual situation of the punching holes. This table is made by correcting full-routes at the border router in AS#17932, picking up prefixes which address block has been assigned based on the CIDR allocation, and comparing a prefix length of each route with the minimum allocation size of each blocks on RIRs’ regulation. This table shows there are many “more specific” routes, many punching holes in other words, in the blocks in which the minimum allocation size is 20 bits in length. On the other hand, there are few “more specific” routes in the blocks with over 24 bits minimum allocation size. Therefore, most of these blocks are for assign addresses to multi-homing users, the minimum allocation size of these block was reduced to assign small address blocks to multi-homing user in other words. In either case, multi-homing increases route information for small address blocks, and thus helping increase the number of route entries in the global Internet.

In our architecture, multiple route information of multiple multi-homing user network can be aggregated before distribution. By this manner, small route information such as 24-bit in length for each multi-homing network are suppressed, and same level of aggregation is possible compared to ordinary aggregations in service providers. Number multi-homing networks are still increasing, and contributing to the growth of number of route entries in the Internet around. Our architecture can avoid an increase of number of route entries which is caused by multi-homing networks.

4.2. Effect of Route Aggregation

By using our proposed multi-homing architecture, it is possible to aggregate route information on route advertisements. In this section, we evaluate a current impact of multi-homing users to number of route entries in the Internet, and show an effect of route aggregation with our multi-homing architecture.

5. Prototype Implementation and Experimentation

We have implemented a prototype of the overlaying multi-homing architecture to confirm its feasibility. In the implementation, we have focused primarily on providing control of in-coming traffic to a user network, which be-
came possible with our architecture. We have limited the function such as traffic analyses and a management function, and implemented a traffic control function which is a core function of our overlaying architecture.

The prototype implementation consists of UR and DR implementation. A virtual link between DR and UR is achieved by GRE tunneling technology that is one of IP over IP tunneling technologies. A DR provide functions for a route advertisement of a user network address space, a tunnel management for a traffic control, and a split traffic for each tunnels. A UR provide a functions for receiving packets that are sent via tunnels.

We have implemented the prototype on the NetBSD operating system that is extensively known as a software router. The prototype consists of a kernel patch and user land applications.

We have conducted an experiment using the prototype and experimental network shown as illustrated in figure 3. By this experiment, we have confirmed the fundamental behavior of our architecture that packets with destination address for a user network are sent to a UR via a DR using tunnels. Furthermore, we have confirmed the function of route selection using characteristics of packets by configuring find grained rules at DR.

Currently, we are extending our prototype with following
functions.

- Adding a monitoring function to monitor states of virtual links
- Design and implementation of control protocol between DR and UR

Furthermore, to obtain an experience of operation of our system, we are planning a wide area operating experiment in which our system will be deployed in a commodity network and handling a commodity traffic.

6. Considerations and Future Directions

In our architecture, packets for user networks are always transmitted via route selection mechanisms. Consequently, packets detour via route selection mechanisms and a round trip time would increase in comparison to a round trip time in case of a transmission along a shortest path. However, we are assuming that route selection mechanisms will be deployed around the Internet backbone networks. If sufficient number of DRs are deployed, round trip increase will not be a problem, but actual relationship between DR deployment and increase in round trip time should be evaluated in a future study.

Security issues must be considered deeply. In our architecture, a user network is connected to DR with virtual links and control of DR is done by UR. Consequently, a control of DR must be done only by an authorized UR. If authorities of a control of DR or virtual links are taken over, then entire of user network may be controlled by intruders. To avoid this problem, security must be maintained in the design of control protocol between UR and DR. We are designing the control protocol between DR and UR with such security functions.

Furthermore, we have to study effects of use of virtual links. In our prototype implementation, we employed GRE tunnels as virtual links. In encapsulation methods such as GRE tunnels, shrinking of tunnel interface’s MTU (Maximum Transfer Unit) size is problematic. In many cases, MTU sizes of tunnel interface are less than 1500 (shrunk by the size of the IP header). There are devices in the Internet with bad configuration, which can not deal with exceeded MTU size correctly. Current practice to keep connectivity through these devices is that to keep MTU so that it is not smaller than 1500. Consequently, our proposal in which a MTU size will be less than 1500 will be difficult to be accepted in the Internet. To avoid this problem, we are currently discussing tunnel methods which does not shrink a size of MTU, such as IP header compression taking advantage of the fact that address space of the user network is small.

7. Conclusion

In this paper, we described problems related to current multi-homing technologies and we proposed a new multi-homing architecture that is based on a overlay networking technology. This architecture not only solves the problem of increasing route information on the Internet, but also offers a possibility to archive some new services which are link selection complying with user application requirements and so on.

We have implemented the prototype implementation to control in-coming traffic to a user network, and have verified its behavior. We have also discussed a design and extension of our implementation for wide area examination.

Acknowledgment

This work was supported by the ministry of public management, home affairs, posts and telecommunications.

References