Site Multihoming: Practices, Mechanisms and Perspective¹

Jun Bi, Ping Hu, and Lizhong Xie Network Research Center, Tsinghua University, Beijing 100084, China junbi@cernet.edu.cn

Abstract

Multihoming to upstream ISPs is a requirement for most IPv4 networks in today's Internet, since multihoming brings benefits such as fault tolerance, traffic engineering, and policy selection to the sites. This paper firstly summarizes the IPv4 site multihoming practices and limitations, and then surveys the current IPv6 site multihoming approaches and point out that the shim6 is the most promising multihoming solution. In addition, this paper discusses the opportunities and challenges that multihoming brings to mobility and the security issues in multihoming environment.

1. Introduction

Multihoming (MH) refers to the phenomenon that one network end node accesses to the Internet through multiple network paths. The network end node here can be a host or a site (such as enterprise network, campus network, etc.). Since the host has not abundant requirements for MH, this paper mainly discusses the site MH, where a site gets multiple IP connectivity from several different ISPs. This paper does not involve the "multi-attaching" phenomenon that a site gets multiple IP connectivity from the same ISP, since this problem can be easily dealt with.

MH to upstream ISPs is a requirement today for most IPv4 networks in the Internet, especially for the enterprise networks. With the deployment of IPv6 network, IPv6 MH will be a very common phenomenon undoubtedly.

Three main motivations for MH are [1]: (1) Redundancy. A site can obtain backup network path to the Internet via MH to several upstream ISPs. In the event of failure of the link to one ISP or failure of the ISP, the site can remain its IP connectivity to the Internet via the backup network path. (2) Traffic engineering (TE). Traffic engineering is an ability to control the path of the inbound and/or outbound traffic. For example, in order to increase the throughput, the site can distribute the traffic on the links attached to different ISPs. (3) Policy selection. A multihomed site can allocate different type of traffic to different ISPs according to its own policy. For example, certain ISP may offer a cheaper price than others for the VoIP service, thus, the multihomed site can allocate its VoIP traffic to that certain ISP.

2. IPv4 MH practices

2.1. The MH approaches for IPv4

BGP based Approach. In this approach, the IPv4 networks obtain a PI (Provider Independent) address space, and then announce this PI prefix as distinct routing prefix into the inter-domain routing system [2]. The main process is as follows: (1) The site gets connectivity to multiple ISPs; (2) The site obtains a PI address space from LIR (Local Internet Registries); (3) The site establishes BGP sessions to the attached ISPs, and propagate its PI prefix to the top level hierarchy of the routing system know as the default free zone (DFZ). This approach can satisfy the most requirements of MH and can deploy with an economical way, thus, it becomes the preferred MH approach for IPv4 networks. However, this approach is completely non-scalable. As the number of multihomed sites in the Internet grows, the number of routing prefixes that are injected into the global routing system increases linearly. This will lead to an unacceptable number of routing prefixes to manage in the DFZ of the Internet.

The MH approach using NAT. This approach uses the PA (Provider Assignment) address assigned by each ISP to which the site is attached. And then the site uses NAT (Network Address Translation) to translate the multiple provider addresses into a single set of private-use addresses within the site [3]. In this way, the hosts within the site do not need to do any changes when the site switches to another ISP, since the hosts just use the private addresses.This MH approach requires no PI addresses and imposes no additional load on the Internet's global routing system. The main problem of this approach is that if one path fails, existing TCP connections will break. Of course, the

¹ Funded by Basic Research Foundation of School of Information Science and Technology of Tsinghua (Grant # SIST2029)

problems of NAT itself are also bought into this MH approach and some complex applications require explicit support for re-mapping the addresses and /or ports.

2.2. The limitations

As described above, the BGP-based approach has been the most popular MH method for the IPv4 networks. Only a few IPv4 networks use the NATbased MH approach. However, a common opinion is that: the BGP-based MH approach is completely nonscalable.

From the situation of current AS number assignment, we can see what a big challenge the scalability problem is. As we known, the AS number is 16-bit long, that is, there are 65,535 AS numbers aggregately. Up to February 2005, 35,000 AS numbers have been assigned. Of these assignments, 18,900 (54%) are announced into the global routing system. But, of the 18,900 AS numbers, about 13,300 (70%) are assigned to enterprise networks [4]. Apparently, it is a common phenomenon to use BGP to achieve MH. [4] estimates the potential number of enterprises who want to be multihomed. Taking the USA for example, there are 650,000 enterprises of over 50 employees. If we assume that enterprises with at least 50 employees would require being multihomed, this would result in the explosion of global routing entry number. Therefore, when we develop the MH approach for IPv6 networks, solutions without having the scalability problem inherent in the BGP-based approach should be emphasized.

3. IPv6 MH approaches

BGP-based IPv6 MH approach has not been feasible due to IPv6 address allocation policies and BGP advertisement prefix length restriction. There is no NAT in IPv6 at all, thus the NAT-based MH approach makes no sense for IPv6 networks. So, the major IPv4 MH solutions are not suitable for IPv6.

3.1. Site-level MH approaches





This kind MH approach is transparent to the hosts within the multihomed site. The hosts within the

multihomed site do not need to do any change in their network protocol stacks. IETF proposes several IPv6 site-level MH approaches, including: IPv6 MH with route aggregation [6], IPv6 MH support at site exit routers [7], IPv6 MH with router renumbering [8] and IPv6 MH with routing support [9]. In this paper, we will only analyze two representative IPv6 site-level MH solutions: IPv6 MH support at site exit routers and IPv6 MH with router renumbering.

IPv6 MH support at site exit routers. RFC3178 [7] proposes an IPv6 MH solution with support at site exit routers. As shown in Figure 1, the multihomed site is connected to the Internet through 2 ISPs (ISPA and ISPB) using link1 and link2 respectively. Each of the ISP allocates an address space to the site. Besides, secondary links (link3 and link4) are established between the site and the ISPs. The site exit router RA of ISPA uses lin3 to connect with the border router BRB of ISPB, whereas the site exit router RB of ISPB uses lin4 to connect with the border router BRA of ISPA. Secondary links are usually implemented as IP over IP tunnels rather than real physical links. In normal conditions, the site uses the primary links (link1 and link2) to communicate. Once one primary link (e.g. link1) fails, the secondary link (e.g. link4) is set up to keep communicating. This approach can preserve established TCP connections through link failure. And there is no need to advertise the site's prefixes to the global routing system, thus this approach is fine scalable. However, this approach does not provide fault tolerance in case of ISP failure. For example, in Figure 1, if ISPA fails, the primary link1 and the secondary link4 also fail. Besides, this approach can not satisfy another two important requirements of MH: traffic engineering and policy selection.



Figure 2. IPv6 MH with routing renumbering

IPv6 MH with routing renumbering. This approach is firstly proposed in [8]. It uses the routing renumbering protocol[10] to switch network path. As shown in Figure 2, the multihomed site obtains two separate address spaces (PrefA:Prefsite::/n, PrefB:Prefsite::/n) from ISPA and ISPB respectively. In case of failure of certain ISP and/or certain link

(taking ISPB and/or link2 as example), the site exit router RB would detect the failure and immediately would use the Router Renumbering protocol to propagate the information to other routers, such as RC, so this addresses (PrefB:Prefsite::/n) will not be used in any new connection. This approach provides both link fault tolerance and ISP fault tolerance. However, once the ISP or link fails, the existing TCP connections will be broken. If the site is in large-scale, the cost and difficulties of routing renumbering are both high.

3.2. Host-level MH approaches

As has been emphasized, host-level MH approaches discussed here are still used to achieve site MH. In this kind of approach, the hosts within the site need to do some changes in their network protocol stacks to achieve MH.





IPv6 MH using mobile IPv6. There is a natural link between MH and mobility. In Mobile IPv6 (MIPv6) [11] mechanism, when a mobile node moves from one network to another, its IP address also needs to switch to the one assigned by the network to which the mobile node moves. This process is very similar to the address switch process in case of failure of one ISP or link attached by the multihomed site. Therefore, a very natural idea is that we can utilize the mobile IPv6 mechanism to achieve MH. [8] describes the MH approach using mobile IPv6. This approach is to use the CoA (care-of-address) assignment mechanism to switch addresses if one ISP or link fails. As shown in Figure 3, suppose that the host A within the multihomed site communicates with host B through link1 and **ISPA** and uses the address PrefA:Prefsite:hostA as the HoA (home address). If ISPA or link1 fails, the address switch process is as follows: (1) When host A sends packets to host B, it fills the PrefB:Prefsite:hostA into the source address field of the packets. Besides, the packet carry a destination option which contain the HoA (PrefA:Prefsite:hostA). So, the devices on the path from host A to host B just see that the source address of these packets are PrefB:Prefsite:hostA. Only the destination host B replaces the source address by PrefA:Prefsite:hostA. (2) Host A sends a BU (Binding

Updating) message to host B in order to notify PrefB:Prefsite:hostA as its CoA. (3) After sending a binding acknowledgement to host A, the destination address fields of the packets sent from host B to host A are filled with PrefB:Prefsite:hostA. Besides, these packets carry routing headers which indicate that the final destination address is PrefA:Prefsite:hostA. Consequently, all packets are sent towards host A using ISPB, and when packets reach host A, the destination addresses are replaced from PrefB:Prefsite:hostA to PrefA:Prefsite:hostA. This approach utilizes the existing protocol MIPv6 to achieve MH, and it can provide both link fault tolerance and ISP fault tolerance. The main problem of this approach is that how to guarantee the security of the BU message, which is implemented by Return Routability Check in MIPv6. However, Return Routability Check depends on the reachability of home agent and the home address. In the MH scenario, if the current using ISP/link fails, the home agent and home address also fail to reach. Thus, Return Routability Check can not work in the MH scenario. Some new security mechanism should be developed, which raises the complexity.





layer. With the shim6, hosts have to deploy multiple provider-assigned IP address prefixes from multiple ISPs. These IP addresses are used by applications and if a session becomes inoperational, shim6 sub-layer can switch to using a different address pair. The switch is transparent to applications as the SHIM6 layer rewrites and restores the addresses at the sending and receiving host. For the purpose of transport layer communication survivability, the shim6 approach separates the identity and location functions for IPv6 addresses. In shim6, the identifier is used to uniquely identify endpoints in the Internet, while the locator is used to perform the role of routing. There is a one-tomore relationship between the identifier and locator. The shim6 layer performs the mapping function between the identifier and the locator consistently at the sender and the receiver. The upper layers above the shim6 sub-layer just use the unique identifier to identify the communication peer, even though the locator of the peer has changed. Hence, when the multihomed host switches to another locator, the current transport layer communication does not break up since the identifier is not changed. Currently, the shim6 mechanism is the most promising MH approach in the IETF's viewpoint. This approach can provide complete fault tolerance. However, it also brings a lot of problems. For example: some protocols, such as ICMP6 [20] or Flow Control Protocol [21], can not work properly since the routers on the path can't see the identifier of the host. Besides, shim6 itself can not do traffic engineering, which is a very important requirement of large-scale site. In order to provide traffic engineering, shim6 WG proposes a solution in which the exit routers rewrite the source addresses of the packets originated within the multihomed site [23]. Anyhow, shim6 is a good MH solution for mid or small sites.

IPv6 MH using transport layer modification. We have seen that there are suitable host MH solutions at the IP layer; support for address changes added to the typical transport protocols is an alternative MH solution. The Stream Control Transmission Protocol (SCTP) [26] is one of these kinds of solutions. SCTP is a TCP-like reliable transport protocol that provides network-level fault tolerance by supporting host MH. Two multihomed hosts inform each other about all of their IPv6 addresses at the beginning of SCTP connection establishment. A SCTP connection regards each IPv6 address of its peer as a "transmission path" towards it. Each SCTP connection chooses one of these addresses as the primary transmission path, upon which data exchange will normally occur. Each end of the SCTP connection monitors all the transmission paths to its peer by sending HEARTBEAT chunks on every path that is not being used to exchange data

chunks. The peer SCTP hosts acknowledge each HEARTBEAT chunk with a HEARTBEAT-ACK chunk. If the primary path becomes inactive, the sending host may automatically choose a new primary path or the user may instruct the local SCTP connection to use a new primary path. SCTP already is a quite mature protocol. The sites are more willing to adopt the mature protocol to achieve MH. However, transport-layer MH solutions have some common problems. Firstly, this kind solution has to design the separate MH mechanism for each transport protocol. For example, SCTP can not apply to UDP. Secondly, the multihomed sites have to develop and use new applications to suit for this kind protocol (e.g. SCTP), which raises the complexity and the difficulty of the deployment of this kind MH solution. Nevertheless, transport-layer MH approach provides a choice of MH solution for certain applications between the multihomed sites.





NAROS. Reviewing the above three host-level MH solutions, the common issue of them is that they can not provide traffic engineering since it is hard to know the status of inbound/outbound link of the site to the hosts within the site. To solve this problem, [27] proposes a mechanism called NAROS (Name, Address and ROute System), which is a solution to perform traffic engineering in an IPv6 multihomed site. NAROS is not intended to preserve TCP connections. Other mechanisms (e.g. shim6) can be used for this purpose. The rationale of NAROS is that the burden of selecting the source address is delegated to the NAROS server, which has access to all the requirements and informations about the actual environnement. An IPv6 multihomed host inquires a NAROS server to determine the source address to use to contact a destination. The NAROS server can use a round-robin or another scheme such as traffic cost when selecting the best address. Figure 5 depicts the scenario that the NAROS server instructs the hosts select the best address according to the traffic cost.

We have seen that there are so many MH solutions for IPv6 networks. How to select a suitable approach for the site that wish to take advantage of MH? We make a comparison of several typical MH approaches. Except the three main requirements of MH: fault tolerance, traffic engineering and policy selection, we take the scalability as another evaluation criterion because the scalability has significant impact on the whole Internet. Besides, in view of the fact that the multihomed sites need not only the ISP/link fault tolerance but also the TCP survivability, we also take the TCP survivability as an evaluation criterion. Table 1 is a brief comparison of the IPv6 MH approaches this paper discussed.

From table 1 we can see, no one approach can fit all requirements of MH. The site should select the appropriate MH approach according to its certain demands. From the perspective of the IETF, shim6 is undoubtedly the most promising MH solution. As for shim6 itself can not achieve traffic engineering, another auxiliary solution in which the exit routers rewrite the source addresses of the packets originated within the multihomed site [23] can realize traffic engineering in a certain extent. Therefore, although shim6 still needs to improve continually, it is worthwhile to be the first choice.

	1		11	
MH	Redund	Session	Traffic	Policy
Approaches	-ancy	Survivability	Engineering	Selection
Exit Routers	yes	yes	no	no
Mobile IPv6	yes	yes	no	no
Shim6	yes	yes	no	no
Transport	yes	yes	no	no
Layer				

Table 1. Comparison of IPv6 MH approaches

4. Related Issues

In most cases, multihomed hosts may connect to the Internet via multiple interfaces. If these interfaces can attach heterogeneous networks, for example, one interface attaches WLAN and another attaches GPRS, in this way, the multihomed mobile node can roam through heterogeneous networks, which is impossible for the single-homed mobile node. In addition, if multihomed hosts connect to the Internet via multiple interfaces, the handoff latency can be reduced effectively [28]. In the traditional mobile IP mechanism, when the mobile node moves from one network to another, the handoff latency consists of two parts: (1) Link layer switch delay. This is the delay in establishing a connection with a new network at link layer when the connection quality with the existing network deteriorates or connectivity is lost. (2) Network layer switch delay. Once the link layer handoff is completed, the network layer switch process is triggered. This process includes that the mobile node discovers the FA (foreign agent) and a number of exchanges of messages. In traditional single-homed mobile IP mechanism, the link layer switch is a "breakbefore-make" process, that is, it has to break the connection with the previous network before establish new connection with the new network. In the MH environment, the mobile node can try to establish the new connection while it still uses the old connection to communicate. In this way, delay of link layer switch can be reduced. Besides, the improved mobile IP mechanism [29] supports the registration of CoA (Care-of-Address) while there are two link connections. Thus, the delay of network layer switch also can be reduced. However, MH also brings some challenges to mobile IP. In MH environment, the mobile node may have multiple CoAs (Care-of-Address) and HoAs (Home Address), thus, there are multiple tunnels between the CoAs and HoAs. In this case, how to create, select and modify the tunnels is an important research problem. IETF sets up Monami6 WG [30] to study this problem.

A feature of MH is that the network path of packets can change. This feature not only provides the fault tolerance but also brings some security problems. RFC4218 [31] summarizes the security issues of MH:

(1) Using redirection attacks to cause packets to be sent to the attacker. In this way, the attacker can inspect and/or modify the payload.

(2) Using redirection attacks to cause packets to be sent to a black hole. The black hole is an address that is nonexistent or unreachable. This attack causes the packets to be dropped by the network somewhere.

(3) DoS/DDoS attack. An attacker can perform redirection to cause overload on an third party.

(4) Influence on ingress filtering. Ingress filtering [32] is that the stub ISP just allows the packets with the source address prefix assigned by this ISP to traverse. Ingress filtering has widely deployed in the current Internet. However, the multihomed site always obtains multiple IP prefixes assigned by several ISPs. Thus, in order to support MH, the stub ISP may allow the IP prefixes that are not assigned by itself to traverse, which may bring some potential security problems. [33] discusses this problem in detail.

(5) Privacy consideration. For the identifier/locator two space MH solutions, it is expected that the identifier can keep constant in a long time, which cause that the behavior of the multihomed hosts can be monitored easily.

5. Summary

After analyzing the IPv4 MH practices and limitations, this paper discusses the representative IPv6 MH approaches proposed in recent time. The present state is that shim6 becomes the most promising MH approach from the viewpoint of IETF. Shim6 is a hostlevel MH solution. It combines the advantages of a number of abandoned host-level MH approaches (e.g. WIMP-F [24], NOID [25], etc.) and it can achieve complete fault tolerance. Traffic engineering and policy selection can be achieved in a certain extent using the auxiliary solution in which the exit routers rewrite the source addresses of the packets originated within the multihomed site. Therefore, shim6 can be a first choice of the multihomed sites.

How to do traffic engineering in MH environment is a research point. [34] points out that in the ideal case, the performance (response time) of the MH site that performs traffic engineering can enhance 40% over the single-homed site. [35] proposes a practical traffic engineering algorithm, which can make the performance of the multihomed site enhance on 15-25%. [36] firstly proposes a traffic engineering algorithm which regards both performance and cost as the optimized factors. Besides, how to apply MH to seamless mobility and how to eliminate the security problems of MH are also worthy to study.

6. References

[1] J. Abley, B. Black, and V. Gill, Goals for IPv6 Site-Multihoming Architectures, IETF RFC 3582, Aug. 2003

[2] J. Abley, et. al, IPv4 Multihoming Practices and Limitations, IETF RFC4116, July 2005

[3] Y. Rekhter, et. al, Address Allocation for Private Internets, RFC 1918, Feb. 1996

[4] P. Savola, Examining Site Multihoming in Finnish Networks, Master's thesis, Helsinki University of Technology, Finland, 2003

[5] H. Berkowitz, Router Renumbering Guide, RFC 2072, Jan. 1997.

[6] J. Yu, IPv6 Multihoming with Route Aggregation, IETF Internet Draft draft-ietf-ipngwgipv6multihome-with-aggr-01.txt, Aug. 2000

[7] J. Hagino, H. Snyder, IPv6 Multihoming Support at Site Exit Routers, IETF RFC 3178, Oct. 2001.

[8] F. Dupont, Multihomed routing domain issues, draft-ietfipngwg-multi-isp-00.txt, Sep. 1999.

[9] N. Bragg, Routing support for IPv6 multi-homing, draftbragg-ipv6-multihoming-00.txt, Nov. 2000.

[10] M. Crawford, Router Renumbering for IPv6, RFC 2894, Aug. 2000

[11] D. B. Johnson, C. E. Perkins, J. Arkko, Mobility Support in IPv6, IETF RFC 3775, June2004.

[12] LIN6 homepage, http://www.lin6.net/

[13] F. Teraoka, M. Ishiyama, M. Kunshi, A. Shionozaki, LIN6: A Solution to Mobility and Multi-Homing in IPv6, IETF Internet Draft draft-teraoka-ipng-lin6-01.txt, Aug. 2001 [14] Arifumi Matsumoto, et. al, Basic Socket API Extensions for LIN6 End-to-End Multihoming, IETF Internet Draft draft-arifumi-lin6-multihome-api-00.txt, June 2003 [15] P. Nikander, R. Moskowitz, Host Identity Protocol IETF Internet Draft draft-moskowitz-hip-07.txt, June 2003

[16] P. Nikander, J. Arkko, P. Jokela, End-Host Mobility and Multi-Homing with Host Identity Protocol, IETF Internet Draft draft-nikander-hip-mm-00.txt, June 2003.

[17] G. Huston, Architectural Commentary on Site Multihoming using a Level 3 Shim IETF Internet Draft draft-ietfshim6-arch-00.txt, July 2005

[18] Erik Nordmark, et. al, Multihoming L3 Shim Approach, IETF Internet Draft draft-ietf-shim6-13shim-00.txt, July 2005
[19] E. Nordmark, Level 3 multihoming shim protocol, draft-ietf-shim6-proto-03.txt, Dec. 2005

[20] J. Postel, INTERNET CONTROL MESSAGE PROTOCOL, IETF RFC729, Sep. 1981

[21] J. Rajahalme, et. al, IPv6 Flow Label Specification, IETF Internet Draft draft-ietf-ipv6-flow-label-02.txt, Dec. 2002

[22] Pekka Savola and Tim Chown, A Survey of IPv6 Site Multihoming Proposals, ConTEL 2005

[23] E. Nordmark, Extended Shim6 Design for ID/loc split and Traffic Engineering, IETF Internet Draft draftnordmark-shim6-esd-00.txt, Sep. 2006

[24] J. Ylitalo, V. Torvinen, E. Nordmark, Weak Identifier Multihoming Protocol Framework (WIMP-F), IETF Internet Draft draft-ylitalo-multi6-wimp-01.txt, June 2004.

[25] E. Nordmark, Multihoming without IP Identifiers, IETF Internet Draft draft-nordmark-multi6-noid-02.txt, July 2004.

[26] R. Stewart, et al, Stream Control Transmission Protocol, IETF RC 2960, Oct. 2000.

[27] C. de Launois, O. Bonaventure, NAROS: Host-Centric IPv6 Multihoming with Traffic Engineering, IETF Internet Draft draft-de-launois-multi6-naros-00.txt, May 2003.

[28] V. Kaulgud, Exploiting Multihoming for Low Latency Handoff in Heterogeneous Networks, ConTEL 2005.

[29] K. Malki, et. al, A. Singh, H. Soliman and S. Thalanany, Low Latency Handoffs in Mobile IPv4, Internet Draft, Mobile IP Working Group, 2002.

[30]Monami6 WG http://www.ietf.org/html.charters/mona--mi6-charter.html.

[31] E. Nordmark and T Li , Threats Relating to IPv6 Multihoming Solutions, IETF RFC 4218, Oct. 2005.

[32] P. Ferguson and D. Senie, Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing, IETF RFC 2827, May 2000.

[33] C. Huitema, R. Draves and M. Bagnulo ,Ingress filtering compatibility for IPv6 multihomed sites, IETF Internet Draft draft-huitema-shim6-ingress-filtering-00.txt, Oct. 2005.

[34] Aditya Akella, et. al, A Measurement-Based Analysis of Multi-homing, SIGCOMM 2003.

[35] Aditya Akella, Srinivasan Seshan and Anees Shaikh, Multihoming Performance Benefits: An Experimental Evaluation of Practical Enterprise Strategies, USENIX 2004.
[36] Lili Qiu, et. al, Optimizing Cost and Performance for Multihoming, SIGCOMM 2004.