A MIPv6-based multi-homing solution

Marcelo Bagnulo, Alberto García-Martínez, Ignacio Soto, Arturo Azcorra

***Abstract*--Global adoption of IPv6 requires the provision of a scalable support for multi-homed sites. This article proposes a multi-homing solution based on the usage of the currently available MIPv6 protocol. It is shown that the MIPv6 protocol is suitable to be used for the provision of IPv6 multi-homing support without modifications in the packet formats defined in the specification, and that it only requires minor changes in node behavior as defined in MIPv6. The resulting solution provides transport layer connection survivability while preserving routing system scalability.**

***Index terms*—IPv6, multi-homing, MIPv6, fault tolerance**

I. INTRODUCTION

As more organizations depend on critical applications built over the Internet, access links are becoming a vital resource for them. As a result, many network sites improve their Internet connection through *multi-homing*, i.e. the achievement of global connectivity through several connections, possibly supplied by different Internet Service Providers (ISPs). However, the extended usage of the currently available IPv4 multi-homing solution is jeopardizing the future of the Internet, since it has become a major contributor to the post-CIDR exponential growth in the number of global routing table entries [1].

Taking this into account, a cornerstone of the design of IPv6 was routing system scalability, which initially resulted in the prohibition of massive route injection into core routers. As a result of this policy, direct adoption of IPv4 multi-homing techniques into IPv6 world was inhibited, so new mechanisms were needed. However, currently available IPv6 multi-homing solutions fail to provide IPv4 multi-homing equivalent benefits, imposing an additional penalty for those adopting the new protocol. This handicap prevents IPv6 adoption in critical environments, and it also provides an excellent excuse to reluctant users. Despite its relevance, developing a scalable multi-homing solution has proven to be a problem extremely hard to solve.

In a different area, during the last few years the Internet community has invested an important amount of effort in providing support for mobile users in IPv6. The result is the almost completed specification of *Mobile IPv6* [2](hereafter MIPv6), which preserves established communication with moving devices.

Universidad Carlos III de Madrid. Departamento de Ingeniería Telemática. Av. Universidad, 30 - Madrid - España.
Email:{marcelo,alberto,isoto,azcorra}@it.uc3m.es

It is relevant to note that the common goal of the multi-homing and mobility solutions is to provide a communicating node with uninterrupted connectivity throughout changes in the point that it is using for attaching to the public IP network. In the mobility scenario, the communicating end-point changes its attachment point because of its movement. In the multi-homing scenario, the attachment point used by the end-point for communication varies because of topological changes caused by outages in the communication elements. While the causes are different, there seems to be enough similarities between the two scenarios to consider common solutions.

Considering that the basic specification for MIPv6 is almost finished, and that the appropriate approach for IPv6 multi-homing support is far from reaching consensus, this paper analyzes the application of the mobility solution to the multi-homing problem.

The remainder of the paper is structured as follows: in section 2 the IPv6 multi-homing problem is characterized. In section 3, the application scenario is presented. Next, in section 4, the MIPv6-based multi-homing solution is described. Details of the resulting behavior of the solution follow on section 5, and section 6 includes related work. Finally, section 7 is devoted to conclusions and future work.

II. THE MULTI-HOMING PROBLEM

Back in early 90's, the *Classless Inter-Domain Routing* (CIDR) address allocation strategy [3] was created in order to cope with the BGP routing table size explosion problem. CIDR proposes the allocation of IP address blocks to transit providers so that customers obtain its address allocation directly from their service provider, instead of obtaining it from a central allocation authority. This strategy allows providers to announce one single aggregated route that describes all their customers, reducing the number of routes in the global BGP routing table. Addresses allocated following the above-described policy are called *Provider Aggregatable* (PA) (see Acronyms Appendix). CIDR provides maximum aggregation efficiency when the network graph is a tree, with providers at the nodes of the tree and end-sites at the leafs. However, the actual network topology does present a fair amount of exceptions to the ideal tree topology, because it is tending to become a dense connectivity mesh [1]. Among the several exceptions to the tree topology that can be found in the current Internet, multi-homed sites are a case that has a direct impact in the routing table size, since multiple available routes to the multi-homed sites must be announced globally in order to obtain multi-

homing benefits. This implies that the size of the BGP routing table will be increased as the number of multi-homers plus the number of providers, which seemed to be somehow acceptable until the number of multi-homed sites started to grow exponentially in 1999 [1].

In IPv6, without the limitations imposed by IPv4 address scarcity, it is possible to guarantee provider aggregation efficiency by assigning multiple prefixes to a multi-homed site, each one corresponding to a different provider [4]. In this configuration, providers serving multi-homed sites only announce their aggregate in the BGP routing table, and multi-homed sites obtain as many prefixes as providers they have, implying that a multi-homed site is represented in the address space as multiple single-homed sites. In order to benefit from multi-homing, nodes within the multi-homed site must configure multiple addresses (one per provider) in each interface. This configuration allows these interfaces to be reachable through the multiple providers. However, this arrangement does not provide by itself survivability for the established connections throughout an outage in the provider that was being used when the communication was initiated, as it will be described in detail in the next section. So, additional mechanisms are needed in order to provide an IPv6 multi-homing solution compatible with PA addressing that could fulfill current IPv4 fault tolerance level.

III. Application Scenario

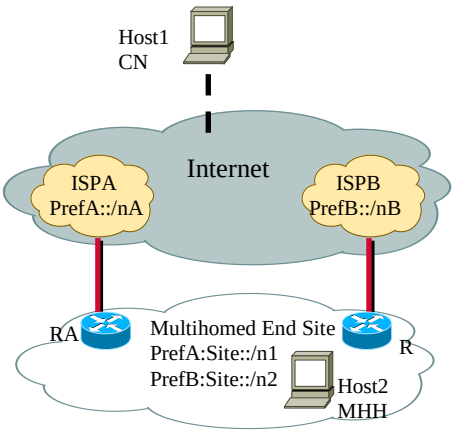

Figure 1: Scenario Topology

The application scenario consists of a multi-homed end-site that obtains global connectivity through two (or more) ISPs i.e. *ISPA* and *ISPB*. Since the end-site is multi-homed and provider aggregatable addresses are being used, the site has obtained two address ranges: one delegated from *ISPA* address range (i.e. *PrefA:Site::/48*) and the other one from *ISPB* address space (i.e. *PrefB:Site::/48*). Furthermore, in order to benefit from multi-homing, hosts within the site have to be reachable through both ISPs. This implies that hosts have to configure one address from each ISP address range that the site has obtained. However, this configuration does not allow the preservation of established connections through an outage. This is the result of:

- Most connections established at the transport level and above identify the nodes involved in the communication by their IP addresses, imposing that they must remain unchanged during the lifetime of the connection.
- In order to preserve aggregation benefits, PA addresses delegated to the end-site by an ISP are to be routed through this ISP.

These constraints imply that if a connection between *Host1* and *Host2* is established using *PrefA:Site:Host2* as the address for *Host2*, packets flowing to *Host2* will be routed through *ISPA*, and only through *ISPA*. Then, if during the lifetime of this connection an outage occurs in *ISPA*, the connection will be dropped, even if a path between *Host1* and *Host2* is available.

IV. A MIPv6-based multi-homing solution

A. *Required capabilities*

In order to preserve established connections throughout an outage, the following capabilities are required:

1. A protocol to inform the other end of the communication about the alternative path that is to be used. Since PA addresses are used, alternative paths (alternative ISPs) are represented by alternative destination addresses. So the protocol is used for conveying alternative destination address.
2. A mechanism that allows packets carrying the alternative address as destination address to be recognized as belonging to the established connection, which means that the original destination address has to be restored when the final destination is reached.
3. A path failure detection mechanism, that enables end-hosts to detect outages in the path that is currently being used.
4. Tools to ensure compatibility with ingress filtering mechanisms. Since an alternative ISP will be used when an outage occurs, packets carrying the original source address would be incompatible with ingress filtering mechanisms.

B. *Solution Rationale*

In this section, we will propose a multi-homing solution that profits from MIPv6 message exchanges for preserving established connections. We will see that the MIPv6 protocol provides the signaling needed to convey alternative path information. However, we will also present that some modification in the behavior of the nodes is needed. It should be noted that most of the required modifications are imposed to nodes within the multi-homed site, while only minor modifications are required to external nodes.

In order to apply the MIPv6 protocol to the considered scenario, the first step is to map the multi-homing scenario into a mobility scenario. Since the multi-homed

host (*MHH*) has the need to use multiple alternative addresses in a given connection, it will assume the role of Mobile Node (*MN*), while the node that it is communicating with it will have the role of Correspondent Node (*CN*). It is assumed that *CNs* support route optimization. Home Agent capabilities are not required.

1) Required capability 1: Protocol for conveying alternative path information.

The most natural application of the MIPv6 protocol for this problem is the usage of Binding Update (*BU*) messages to inform the *CN* that an alternative address is to be used for the established communication, fulfilling requirement 1. So, when the *MHH* detects that the currently used path becomes unavailable, it would send a *BU* message to the *CN*, informing that an alternative address is to be used. In MIPv6 terminology, the original address would be the Home Address (*HoA*) and the new alternative address would be the Care-of Address (*CoA*) (*Figure 2*). However, MIPv6 security requirements impose the performance of the Return Routability (*RR*) procedure to enable the required *BU* message authorization. Such procedure, illustrated in figure 2, implies the exchange of Home Test Init (*HoTI*) and Home Test (*HoT*) messages using the *HoA*, and the exchange of Care-of Test Init (*CoTI*) and Care-of Test (*CoT*) messages using the *CoA*. These exchanges are designed to verify that the host reachable through both the *CoA* and the *HoA* is the same. This means that the *MHH* needs to be reachable through both paths, implying that these exchanges cannot be performed successfully once an outage has occurred. So, the *RR* procedure should be performed when a connection with a new *CN* is established allowing the protection of this connection during its lifetime.

However, the data used for the generation of authorization information has a limited lifetime, imposing periodical *RR* checks, in order to ensure that valid *BU* authorization information is available when an outage occurs. Time constraints imposed by the MIPv6 specification are that authorization data must remain valid for at least *MAX_TOKEN_LIFE* (210 sec.) after it has been used for the *RR* procedure, but no longer than *MAX_NONCE_LIFE* (240 sec).

These constraints impose the performance of the *RR* procedure every *MAX_TOKEN_LIFE* minus the time required to perform the procedure, which would include the Round Trip Time (RTT) and the processing time. If the typical value used for TCP connection establishment timeout (75 sec) is accepted as a reasonable upper bound to the RTT, the *RR* procedure needs to be performed every 135 sec.

2) Required capability 2: Original connection packets recognition capability.

Once that the availability of the information needed to authorize *BU* messages is guaranteed, the *MHH* is prepared to re-route its connections through an alternative address when an outage occurs. So, when the outage is detected, the *MHH* will send a *BU* message to the *CN*, informing that a new address (*CoA*) is to be used. Then, the *CN* will address packets to the new *CoA* as long as the Binding Cache Entry (*BCE*) that links the *HoA* with the particular *CoA* remains valid (discussed below). However, packets addressed to *CoA* have to be recognized as belonging to the established connection. This can be achieved by using the Type 2 Routing Header specified in MIPv6, in the same way that it is used for supporting mobility. That is, after receiving a valid *BU*, the *CN* addresses packets to the *MHH* to the new *CoA*, and it also includes a Type 2 Routing Header carrying the *HoA*. The resulting behavior is that when these packets reach the *MHH* through the *CoA*, the Routing Header is processed and the *HoA* is restored and packets are presented to the transport and above layers as being addressed to the *HoA*, preserving established connections.

As we stated above, communication can be preserved as long as the correspondent *BCE* in the *CN* remains valid. It is stated in MIPv6 specification that *BCEs* that have been authorized using the *RR* procedure have a maximum lifetime of *MAX_RR_BINDING_LIFE* (420 sec). If this *CoA* is to be used to reach the *HoA* after this period, a new *BU* message binding the *HoA* and the *CoA* has to be sent. However, as it has been presented earlier, the *BU* authorization data has to be acquired using the *RR* procedure, which implies communication through both the *CoA* and the *HoA*. In the multi-homing scenario, the *RR* procedure cannot be performed once that an outage occurred along the initial path, since the *HoA* is unreachable. This constraint means that if the MIPv6 protocol is used for preserving an established communication in multi-homed environments, such communication can only be preserved for 7 minutes after an outage occurs which seems to be a severe limitation for this application. It is then deemed necessary to overcome this limitation if MIPv6 is to be applied to solve the multi-homing problem. However, the *BCE* lifetime has been limited to a few minutes in order to limit the possibility of time shifting attacks, as it is presented in [5]. The goal of MIPv6 security is to avoid the introduction of new security hazards that are not present in non-MIPv6 enabled environments. In particular, the goal of the *RR* procedure is the reduction of the set of potential attackers to those who can intercept packets flowing between the *MN* and the *CN*. This procedure forces the attacker to be present somewhere along the path between the mobile node and the correspondent node in order to acquire valid authorization data needed to generate forged *BU* messages.

However, this mechanism by itself only imposes that the attacker has to be present on the path the time needed to intercept the messages that carry authorization information. Once that the attacker has intercepted valid authorization information, he can leave his position along the path and still perform attacks using such information. These are called time shifting attacks since an attacker that once was on-path intercepting packets can perform attacks in the future, when he is no longer on the communication path. Time shifting attacks can be

achieved in the following way: the attacker placed along the communication path intercepts authorization information and generates a forged *BU* message. The attacker leaves the position, but the attack continues since the traffic is still diverted to the *CoA* contained in the fake *BU* message. The effect of this attack is limited by reducing *BCE* lifetime in the *CN* to 7 minutes, imposing the generation of a new *BU* message in order to restore the *BCE*.

In order to enable a MIPv6 based multi-homing solution, *BCE* lifetime has to susceptible to be extended but alternative protection must be provided. So, the following modification in the *CN* behavior is proposed. The *CN* must send a Binding Refresh Request (*BRR*) to the *HoA* included in the *BCE*, if the *BCE* has not been refreshed by the mobile node for 420 sec. In case that this is a legitimate *BCE*, the mobile node will send a *BU* to refresh it. If this is a fake *BCE*, the fixed node will reply to the *BRR* message with an ICMP Parameter Problem [6]. Then, upon the reception of an ICMP Parameter Problem message, the *CN* has to delete the correspondent fake *BCE* from the cache, terminating the attack. However, if this address has became unavailable because of an outage, an ICMP Destination Unreachable message containing a No Route to Destination code is to be sent. In this case, it is proposed that when the *CN* node receives such a message, it must verify that it is the reply to the *BRR* message by contrasting it with the information contained within the ICMP message payload. If the verification is successful, the *CN* recognizes the multi-homing application of the protocol and extends the lifetime of the *BCE* for another 420 sec. The proposed modification achieves the initial goal that was to avoid attacks coming from a host that is not present along the used path.

3) Required capability 3: Path failure detection mechanism

Additionally, a failure detection mechanism that triggers the generation of *BU* messages is required to provide a complete solution. A path failure detection mechanism can be based on the exchange of *HoTI/HoT* messages. The *RR* procedure needs to be performed periodically, implying message exchange between the *MHH* and the *CN*. However, in order to provide a failure detection mechanism, the message exchange frequency has to be increased, not only because its period of 135 seconds may be deemed as unacceptable for certain applications, but because valid authorization information is required for sending the *BU* message. Since a failure is indicated by at least one keep-alive message lost, it is necessary that after such event valid *BU* authorization information is still available, which imply that the information acquired during the previous message exchange is still valid. Then, assuming that a failure is indicated by the lost of two consecutive keep-alive packets, *HoTI* messages have to be generated by the *MHH* every *MAX_TOKEN_LIFE*/3 seconds, i.e. 70 seconds. Then if two *HoTI* messages are lost, that is, if no reply is received 140 sec. after a *HoTI* was sent, a *BU* message is generated and an alternative route is used. The simple mechanism presented provides the minimum required functionality while honoring the timing constraints imposed by the *RR* procedure parameters. Failure response time can be improved by increasing the message exchange frequency. Moreover, adaptive mechanism, such as the TCP time-out calculation mechanism can also be considered, as long as they respect the timing constraints imposed by MIPv6

4) Required capability 4: Tools to ensure compatibility with ingress filtering mechanisms.

When a *MHH* has multiple PA addresses configured in its interface, source address selection implies the selection of the ISP to be used in the return path. Moreover, because of ISP ingress filtering mechanism, source address selection also imposes the ISP to be used in the forward path, requiring additional functionalities at the multi-homed site to guarantee the appropriate ISP selection as discussed in [7]. Besides, when host based path failure detection mechanisms are used, the only party that has the information needed for selecting the path to be used is the host itself. So, in order to guarantee the compatibility with ingress filtering mechanisms, the *MHH* can select the exit ISP by means of a Routing Header. In order to simplify ISP selection, the Site Exit Anycast Address (*SEAA*) defined in [7] can be used. Then, after performing source address selection, the *MHH* addresses packets to the *SEAA* corresponding to the ISP that has assigned the address used as source address and it includes the final destination address in a Routing Header.

Additional complexity results when an outage occurs. In this case, an alternative ISP is to be used for coursing packets. Source address filtering mechanisms of the alternative ISP precludes the flow of packets carrying the address originally used, i.e. the *HoA*. However, the *CN* only recognizes packets as belonging to the established connection if they carry the original *HoA*. In order to overcome this issue, the Home Address Destination Option is to be used, so that the source address corresponding to the alternative ISP (i.e. the *CoA*) is carried in the Source Address field of the IPv6 header and the original address (i.e. the *HoA*) is carried within the Home Address Option. When the packet is received by the *CN*, it processes the Home Address Option and restores the *HoA* as the Source Address.

V. RESULTING BEHAVIOR

In this section, the complete operation of the solution is described.

The communication established between the *MHH* and the *CN* can be initiated by any of the parties. Suppose that the communication is initiated by the *CN* (the case where the communication is initiated by the *MHH* is analogous): It will first obtain at least one of the *MHH*'s addresses, for instance using the DNS. If all the *MHH*'s addresses are listed in the DNS, the *CN* will pick one and try to initiate the communication using this address. If a failure has

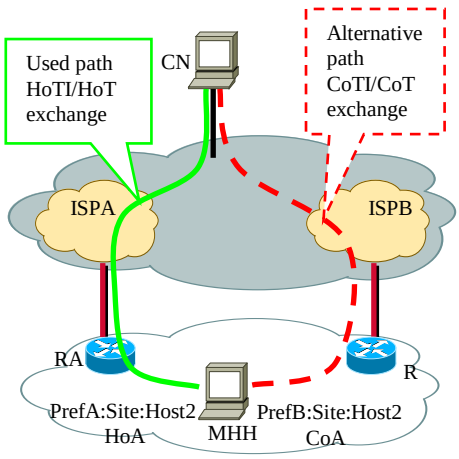

Figure 2: Solution Behavior

occurred along the path, the attempt to initiate the communication will fail, and the *CN* will try another address. Eventually, a packet from the *CN* will reach the *MHH*. The *CN* will then reply to the received packet using the addresses contained within the packet. The *MHH* attempts to communicate with the *CN* using the selected source address. In order to avoid that the discarding of the packet by the ISP ingress filtering mechanism, the *MHH* addresses the packet to the *SEAA* of the ISP that assigned the source address selected and includes the *CN* address within a Routing Header.

Once that the *MHH* starts sending packets to the *CN*, different address roles have been set: the address used as Source Address in the first packet flowing from *MHH* to *CN* will be the *HoA*, and the other available addresses of *MHH* will be *CoA*s. It should be noted that these roles are assigned when the communication is established, and they are not predetermined. In the application scenario, we suppose that the first packet flowing from *MHH* to *CN* has *PrefA:Site:Host2* as source address, so that: *PrefA:Site:Host2* is the *HoA* and *PrefB:Site:Host2* is the *CoA*.

Once the first packet is carried from *MHH* to *CN*, the *MHH* has to perform the *RR* procedure in order to obtain valid authorization data. The *RR* procedure consists on exchanging *HoTI/HoT* messages using the *HoA*, and *CoTI/CoT* messages using *CoA*. These messages also have to include a Routing Header to select the appropriate exit ISP. The *HoTI/HoT* message exchange is also used as a path failure detection mechanism, imposing a *HoTI/HoT* exchange every 70 sec.

If no outage occurs, the communication continues as it is, and *HoTI/HoT* and *CoTI/CoT* message exchanges continue until the communication is finished.

If an outage occurs, it will be detected by the failure detection mechanism and an alternative path will be used. If two consecutive packets *HoTI* are not replied within 140 sec. after the first message was sent, a failure will be assumed. If this is the case, a *BU* message is sent, informing the *CN* that the *CoA* will be used to exchange packets. This *BU* message will carry the authorization information obtained through the last successful *HoTI/HoT* and *CoTI/CoT* message exchanges.

Then, an alternative ISP will be used to course packets between the *MHH* and the *CN*. Packets from the *CN* to the *MHH* will contain a Type 2 Routing Header in order to be routed through the alternative ISP. Packets from the *MHH* to the *CN* will carry the Home Address Destination Option and a Routing Header to select the exit ISP:

The communication can continue using this route while the binding established at the *CN* remains valid. The *CN* sends periodical *BRR* messages, and if an ICMP Destination Unreachable message containing a No Route to Destination code is replied, it extends the *BCE* lifetime.

VI. Related Work

In this section, we will consider alternative approaches proposed to tackle the IPv6 multi-homing problem. An approach compatible with PA addressing is presented in [8]. If we apply this mechanism to the multi-homed site depicted in *Figure 1*, the solution consists in building a tunnel between an *ISPA* exit router and *RB*, and another tunnel between *ISPB* exit router and *RA*. Then if a link is down, packets are forwarded through the tunnel. In this case, alternative route information is confined to routers connecting ISPs with multi-homed sites, so the scalability of the global routing system is preserved. However, this solution presents limited fault tolerance capabilities.

The Host Centric Multi-homing proposal that is being developed in [7] provides some of the multi-homing benefits through available tools such as Router Advertisements and Router Renumbering. It also deals with the problem caused by ingress filtering. This problem is basically caused when packets containing a source address from the *ISPA* block are coursed from the multi-homed site through *ISPB* (*figure 1*). In this case, ingress filtering configured in the *ISPB* ingress router will discard those packets, because their source address is considered to be spoofed. Host Centric Multi-homing solution proposes several options to deal with this issue, ranging from source address routing to redirecting packets to appropriate site exit routers. However, this proposal does not include mechanisms to preserve established communications through an outage in the used route. So, the authors consider that both proposals complement each other, since they address different aspects of the multi-homing problem.

VII. Conclusions and Future Work

In this paper we have presented a multi-homing solution for IPv6 sites based on the usage of MIPv6 protocol. The main benefit of the solution is its compatibility with the existent technology. Changes needed involve mainly hosts within the multi-homed site, which must perform a failure detection mechanism exchanging MIPv6 messages. The solution only implies a minor change in external nodes, which must be modified so that they extend *BCE* lifetime

upon the reception of an ICMP Destination Unreachable message containing a No Route to Destination Code as a reply of a *BRR* message. Besides, it should be noted that the solution is compatible with the PA scheme, granting the scalability of the routing system.

Multiple additional optimizations can be done to enhance the solution. Some of which are presented next. MIPv6 specification includes the possibility of piggybacking binding related messages in data packets as a future extension of the protocol to reduce the overhead.

Other possible optimization that can be performed is related to the failure detection mechanism. The proposed mechanism provides minimum facilities. Improved algorithms can be proposed so that faster detection is provided. For instance adaptive mechanisms such as the ones used by TCP can be adopted.

APPENDIX: ACRONYMS

BA: Binding Acknowledgment
BCE: Binding Cache Entry
BRR: Binding Refresh Request
BU: Binding Update
CN: Correspondent Node
CoA: Care-of Address
CoT: Care-of Test
CoTI: Care-of Test Init
HoA: Home Address
HoT: Home Test
HoTI: Home Test Init
MHH: Multi-Homed Host
MN: Mobile Node
PA: Provider Aggregatable
RR: Return Routability
SEAA: Site Exit Anycast Address
SEAAA: Site Exit Anycast Address corresponding to ISPA
SEAAB: Site Exit Anycast Address corresponding to ISPB

REFERENCES

[1] G. Huston, "Commentary on Inter-Domain Routing in the Internet", RFC 3221, 2001.

[2] Johnson, D et al. "Mobility Support in IPv6", Internet Draft, Work in progress, Feb 2003.

[3] Fuller, V. et al."Classless Inter-Domain Routing (CIDR): An Address Assignment and Aggregation Strategy", RFC 1519, 1993.

[4] R. Hinden et al. "An IPv6 Aggregatable Global Unicast Address Format", RFC 2374, 1998.

[5] Nikander et al. "Mobile IP version 6 (MIPv6) Route Optimization Security Design Background", Internet Draft, Work in progress, March 2003.

[6] Conta, A. et al."Internet Control Message Protocol for the Internet Protocol Version 6 Specification", RFC 2463, December 1998.

[7] Huitema, C. et al. "Host-Centric IPv6 Multihoming", Internet Draft, Work in progress, June 2002.

[8] J. Hagino, et al. “IPv6 Multihoming Support at Site Exit Routers”, RFC 3178, 2001.