Preserving Established Communications in IPv6 Multi-homed Sites with MEX

M. Bagnulo, A. García-Martínez, I. Soto, A. Azcorra, J. F. Rodríguez

Departamento de Ingeniería Telemática - Universidad Carlos III de Madrid
{marcelo,alberto,isoto,azcorra,jrh}@it.uc3m.es

Abstract. A proper support for multimedia communications transport has to provide fault tolerance capabilities such as the preservation of established connections in case of failures. While multi-homing addresses this issue, the currently available solution based in massive BGP route injection presents serious scalability limitations, since it contributes to the exponential growth of the BGP table size. Alternative solutions proposed for IPv6 fail to provide equivalent facilities to the current BGP based solution. In this paper we present MEX (Muti-homing through EXtension header) a novel proposal for the provision of IPv6 multi-homing capabilities. MEX preserves overall scalability by storing alternative route information in end-hosts while at the same time prevents packet losses by allowing routers to re-route in-course packets. This behavior is enabled by conveying alternative route information within packets inside a newly defined Extension Header. The resulting system provides fault tolerance capabilities and preserves scalability, while the incurred costs, namely deployment and packet overhead, are only imposed to those that benefit from it. An implementation of the MEX host and router components is also presented.

1. Introduction

In order to provide production-quality multimedia communications over the Internet, fault tolerance capabilities are required, including the preservation of established connections in case of failure in the transmission path as long as an alternative one is available. To address this need, more and more sites are adopting multiple connections to the Internet, becoming multi-homed. 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 BGP 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, which impose an additional penalty for those adopting the new protocol. Despite its relevance, developing a scalable multi-homing solution has proven to be a problem

extremely hard to solve, basically due to the heterogeneous set of requirements imposed to its design.

In order to be adopted, an IPv6 multi-homing solution has to furnish most of the benefits provided by the current IPv4 multi-homing solution while preserving the scalability of the routing system. The benefits provided by the current IPv4 solution include a high level of fault tolerance support, meaning that communications (including established TCP connections) are not to be interrupted because of an outage as long as at least one path exists between the site and the correspondent node in the global network. Besides, the current solution provides some degree of policing, allowing multi-homed sites to route inbound and outbound traffic through different providers based on administrative criteria. Additionally, since a new multi-homing solution may imply some changes in current implementations, the adoption of a new mechanism must honor legacy implementations, meaning that nodes supporting the new solution must be able to communicate with legacy ones, even if this particular communication does not obtain multi-homing benefits. Finally, new mechanisms must neither introduce new vulnerabilities to the multi-homed sites nor enable new attacks to any other party. For a more detailed description of the requirements imposed to a multi-homing solution, the reader is referred to the work that is being done at the IETF by the multi6 working group [2].

In this article, we will present MEX (Multi-homing through EXtension headers), a novel IPv6 multi-homing solution that achieves equivalent benefits to those provided by current IPv4 multi-homing solution while preserving the route aggregation capabilities provided by the CIDR scheme [4]. MEX is based on including in the packets flowing to a multi-homing site the information needed to re-route them through alternative paths in case that an outage occurs in the currently used path. This information is conveyed into a new Extension Header [3] defined ad hoc. The Extension Header can be processed by intermediate routers when the destination address containing the packets is unreachable. Scalability is granted by the fact that no information about alternative paths is stored in the routing system.

The remainder of this article is organized as follows. The next section describes the design rationale and motivations. In Section 3, the solution is described, starting by the presentation of its components, and following by the detail of its operation in a typical scenario. Cost-benefit analysis is performed in Section 4. Implementation details of a prototype are next presented in Section 5. Section 6 summarizes the related work and Section 7 highlights the most relevant conclusions of the paper.

2. Motivation and Rationale

Back in early 90's, the Classless Inter-Domain Routing address allocation strategy [4] 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 aggregate route that summarizes the reachability information to 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). CIDR aggregation efficiency is granted as long as the underlying network topology is coupled to address allocation, providing maximum aggregation efficiency when the network graph is a tree, with providers at the nodes of the tree and end-sites at the leafs (Figure1a). However, the actual network topology does present a fair amount of exceptions to the ideal tree topology since it is tending to become a denser connectivity mesh [1].
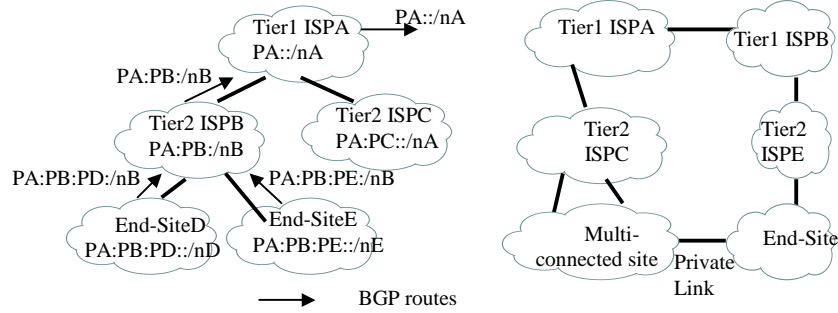

Figure 1.a: Provider Aggregation   Figure 1.b: Local Exceptions

Not all of the exceptions to the tree topology that can be found in the current Internet have impact in the global BGP routing table. For instance, a private link between two customers of different providers (Figure 1b), requires the propagation of routing information with a scope limited to the involved parties. Since this information is not intended to be globally visible, it does not generate an additional entry in the global BGP routing table. A similar situation occurs when considering multi-connected sites (Figure 1b). A *multi-connected site* can be defined as an end-site, i.e. a site that does not provide packet transit service for other sites, that has two or more different links to the same service provider. We reserve the term *multi-homed site* for an end-site that obtains global IP connectivity through two or more different service providers. While a multi-connected site is an exception to the tree topology, it can be handled locally at the service provider without imposing additional information to the global routing system. However, since CIDR inception, it is known that multi-homed sites are an exception to the tree topology that cannot be handled locally by the current routing system, 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 of the network core will be increased as the number of providers plus the number of multi-homers, which seemed to be somehow acceptable until the number of multi-homed sites started to grow exponentially in 1999 [1].

Without the limitations imposed by IPv4 address scarcity, provider aggregation efficiency can be guaranteed in IPv6 by assigning multiple prefixes to a multi-homed site, each one of them corresponding to a different provider [5]. 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 of the established connections throughout an outage in the provider that was being used when the communication was initiated because the ends of a transport and upper layer connections are identified through the initial IP address.

Additional mechanisms can be introduced in order to overcome the above detailed problems. Such mechanisms include some modification in end-hosts, in order to recognize packets carrying multiple source and destination addresses as belonging to the same communication. In addition, end hosts need also to be furnished with mechanisms that allow them to detect that the path currently used is no longer available. These mechanisms involve some kind of explicit or implicit feedback about network status. For instance, routers should send an error message (e.g. ICMP Destination Unreachable [7]) back to the source whenever a packet cannot be forwarded. End hosts can use this information as an indication to change the destination address that is being used, in case that an alternative one is available. On the other hand, end-hosts can detect that a route has become invalid simply by noting that packets do not flow anymore through it. In this case, there is no explicit unreachability information from the network devices. In order to diminish detection latency, explicit checks can be performed by using keep-alive messages, expressly generated for this purpose. It should be noted that in both cases, the end-host based fault detection mechanism capabilities are limited to reactive measures, meaning that actions are undertaken after the fault occurred and its effects are visible, most commonly implying packet loss. This is because once the end host sends a packet to a given destination, there is nothing that end hosts can do to change its path, even if they find out that this path is no longer valid, since the packet belongs to the routing system realm. Because of the very nature of the network functional architecture, a solution based on the routing system is capable of providing better performance during an outage than end-host solutions, since it would be capable of re-routing packets whose current path had become unavailable. As we concluded earlier, with the currently available tools, the routing system needs to store alternative route information, leading to scalability challenges.

In the present article, we will explore the possibility of obtaining the best of both worlds, with a mechanism that stores route information for alternative path in end hosts, assuring system scalability, while transferring recovery responsibilities to the routing devices which are actually handling the packets, allowing the re-routing of packets, and avoiding packet loss.

3. Description of MEX

The proposed solution assumes the usage of currently adopted PA address allocation schemes to preserve routing system scalability. Therefore, multi-homed sites are supposed to obtain one PA address block from each of its providers. So, to avoid the scalability limitations caused by storing tree topology exception information in the routing system, MEX stores information linking the multiple addresses available for a

given host in the host itself. In order to prevent packet loss, re-routing of packets to alternative available addresses is to be performed by the routing system, imposing the need to convey alternative address information from hosts to routers. This is done through a newly defined Extension Header that carries information about alternative addresses, so that if the address contained in the Destination Address field of the IPv6 header becomes unreachable, it is swapped with an alternative address extracted from the Extension Header, and then the packet is re-routed to the new destination. It must be noted that the extension header carrying alternative address information must be included in packets flowing towards the multi-homed site, which means that the header must be created by the correspondent node (the other end of the communication). Since alternative address information is stored in the hosts of the multi-homed site, a new Destination option [3] is defined to convey alternative address information from the multi-homed host to the correspondent node.

3.1. MEX Components

3.1.1. Alternative Prefix Destination Option

A new *Alternative Prefix* (hereafter AP) *Destination Option* is defined in order to convey information about multiple alternative address from where it is stored i.e. hosts in the multi-homed site to where the packets are created i.e. the correspondent node.

It is assumed that hosts in the multi-homed site will configure multiple prefixes per interface in order to enable multi-homing benefits in communications through this interface. Furthermore, it is assumed that, in general, multiple addresses assigned to the same interface will share the same Interface Identifier part and will differ in the prefix part. This is considered the most natural configuration since it is the output of the Stateless Address Auto-configuration procedure as specified in [8]. Therefore, both the Destination Option and the Extension Header will only carry alternative prefix information, instead of full alternative address information. The exact format of the Destination Option is outside the scope of this paper.

3.1.2. Alternative Prefix Extension Header

As it has been previously stated, the fundamental component of MEX is the new *Alternative Prefix* (AP) *Extension Header* that carries alternative prefix information within packets flowing to the multi-homed destination, so that alternative prefixes carried within it can be used in case that the address contained in the Destination Address field of the IPv6 header [3] becomes unreachable.

While the detailed format of the Extension Header is outside the scope of this paper, it is relevant to note that the new Extension Header will carry an *Alternative Prefix* field containing alternative prefixes assigned to the destination interface other than the one included in the Destination Address field of the IPv6 header. It will also contain a *Pleft* field that carries the number of A*lternative Prefixes* left, i.e. the number of Prefixes that has not been used in the Destination Address field of the IPv6 header for reaching the final destination and a *Hdr Ext Len* (Extension Header

Length) field that contains the total number of Alternative Prefixes carried in the Extension Header.

The intended usage of the AP Extension Header is the following:

1. If a router receives a packet and it has no route to the address contained in the Destination Address field, the router must look for an AP Extension Header.
2. If such header exists, and the value of *Pleft* is non zero, then the router must swap the 64 most significant bits of the Destination Address with the Prefix located in the AP Extension Header at the position number *i*, being *i* equal to *Ext Hdr Len* minus *Pleft*.
3. Then the router must decrement *Pleft*.
4. The router must try to forward the packet to the new destination address. In case that there is no route to the new destination, processing is resumed from step 2.
5. If there is no AP Extension Header or the *Pleft* value is zero, the packet must be discarded.

A formal description of this procedure is the following:

while (No Route to Destination) AND (Exists AP Extension Header) **{**
   if (*Pleft* = 0) **{**Discard packet; **}**
   else {
     if (*Pleft* > *Hdr Ext Len*) **{**
       send (ICMP message to the Source Address, pointing to the *Pleft* field);
       discard the packet; **}**
     else {
      Pleft = *Pleft* - 1;
      i = *Hdr Ext Len* – *Pleft*;
      swap (prefix of the Destination Address, Alternative Prefix #i);
      resubmit the packet to the IPv6 module for transmission; **}**
    }
  }

3.2. MEX Operation

A typical scenario where MEX can be adopted is depicted in the figure below.
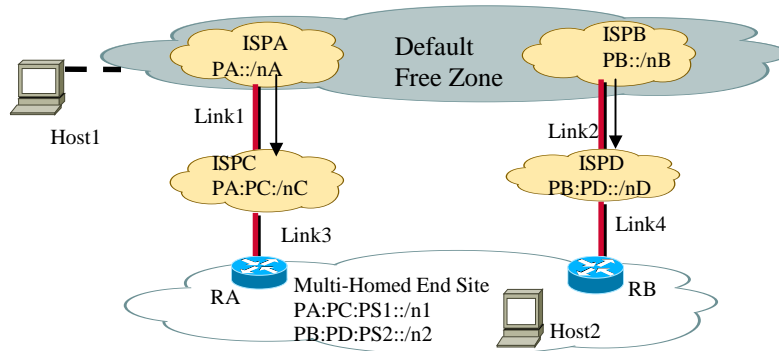

Figure 2: Scenario topology

A Multi-homed end-site obtains global connectivity through two ISPs: *ISPC* and *ISPD*. These ISPs do not belong to the Default Free Zone (i.e. they have a default route to its upstream provider) and they buy transit from *ISPA* and *ISPB* respectively. *ISPA* and *ISPB* do belong to the Default Free Zone, i.e. at least one of their routers has full BGP routing information.

Since the end-site is multi-homed, it has received two address ranges from its providers, one delegated from *ISPC* address range i.e. *PA:PC:PS1::/n1* and another one delegated from *ISPD* address space i.e. *PB:PD:PS2::/n2*.

ISPC and *ISPD* have obtained a range of the address space from the address range assigned to their respective providers, i.e. *ISPA* and *ISPB*. So, *ISPA* has delegated the range *PA:PC::/nC* to *ISPC* and *ISPB* has delegated the range *PB:PD::/nD* to *ISPD*.

3.2.1. Normal operation
In this section we will consider the case of a given host in the Internet (*Host1*) communicating with a host belonging to the multi-homed end-site considered (*Host2*).

Host2 belongs to the multi-homed site, so it has at least two addresses published in the DNS: *PA:PC:PS1:PL1:IIdHost2* and *PB:PD:PS2:PL2:IIdHost2*.

If the communication is initiated by *Host2*, it obtains *Host1* address through the DNS as usual, and then it sends a packet (*packet1*) to *Host1* address, including in it an AP Destination Option with all the different prefixes it is willing to use to receive replies to this packet. Then, *Host1* replies sending *packet2* to *Host2*, addressing it to the source address included in *packet1* and also including an AP Extension Header with the prefixes included in the AP Destination Option of *packet1*. When *Host2* receives *packet2*, it verifies that the destination address and all the prefixes included in the AP Extension Header belong to the list of addresses assigned to its interfaces. If at least one of the derived addresses is not assigned to any of the interfaces, the packet is discarded, because of the security issues considered below.

If the communication is initiated by *Host1*, it performs an AAAA-type query to the DNS and obtains *PA:PC:PS1:PL1:IIdHost2* and *PB:PD:PS2:PL2:IIdHost2*. *Host1* uses one of the obtained addresses as destination address and it includes the other address in an AP Extension Header. The communication will continue as in the previous case.

3.2.2. Fault Tolerance Support
In this section, we will present MEX response to an outage along the currently used path. We will consider the case where *Host1* sends packets to *Host2*, addressing them to *PA:PC:PS1:PL1:IIdHost2*, and *Link1* in figure 2 fails. In this case, *ISPA* routers will not be able to route subsequent packets of this communication, since there will be no route to this destination in its routing tables. Then, the MEX capable router (a router that is capable of processing the AP Extension Header), this router will look for the Extension Header in those packets whose destination address is unreachable, including those addressed to *PA:PC:PS1:PL1:IIdHost2*. If such header is found, it will be processed and the prefix of the destination address will be replaced with the alternative one, and the packet will follow the alternative route toward its destination.

It may be argued that AP Extension Header processing imposes an unacceptable load in routers, especially in those located at the core of the network. Another issue

that could be raised is that deploying MEX imposes the need for upgrading all the routers of the ISP in order to be able to process the AP Extension Header. A workaround for these issues is to limit the Extension Header processing to specific upgraded routers connected to the ISP network. The proposed configuration would operate in the following way: These upgraded routers announce a default route within the ISP network; in figure 2, the upgraded router is connected to the *ISPA* network and announces a route to 0/0. Then, if *link1* is working properly, packets will flow through *link1* because of the longest prefix match rule. If *link1* is down, there will be no more-specific route in the routing tables, so the default route will prevail, making packets flow to the upgraded router. This device will process the AP Extension Header, swapping prefix information. Once this is done, it will forward the packet to the *ISPA* network, and then to the alternative route.

A slightly different approach is needed to provide a sink route for packets with an unreachable destination address when *link3* fails. Since *ISPC* obtains a default route from its provider *ISPA*, it is not possible to announce a default route to sink packets with unreachable destination, as presented above. In this case, the upgraded routers announce a route to the address range allocated to the ISP; in the figure above, the upgraded router is connected to the *ISPC* network and announces a route to *PA:PC::/nC*. Then if *link3* is working properly, packets will flow through *link3* due to the longest prefix match rule. If *link3* is down, packets will be forwarded to the upgraded router, where the AP Extension Header will be processed, swapping prefix information. Once this is done, the MEX capable router will forward the packet to the *ISPC* network, and then to the alternative route. Eventually, packets will reach *Host2*, where the original destination address is restored based on the information contained in the *Ext Hdr Len* field and the *Pleft* field.

4. Cost-Benefit analysis

The adoption of MEX imposes essentially two costs, namely, additional overhead and a considerable impact in the installed base of equipment, as it will be detailed next.

Overhead. The usage of the AP Extension Header and Destination Option introduces additional overhead in the packets exchanged by the multi-homed site. Furthermore, as the overhead increases linearly with the number of providers, MEX becomes less attractive. So, the proposed mechanism is not considered to be suitable for large sites with several providers, while it is considered to be attractive for sites with a few providers, such as dual-homed sites. In this case, the introduced overhead is limited to 128 bits per packet. Despite the fact that this solution may not be attractive for all scenarios, we should note that at this point it is not clear that a one-size-fits-all solution will emerge covering all the imposed requirements.

Impact on the installed base. In order to obtain multi-homing benefits, both ends of the communication must be capable of processing the new Extension Header and Destination Option defined. This imposes the upgrade of not only the hosts within the multi-homed site but also of the correspondent nodes. While this is considered to be a great challenge, it should be noted that several new features that should be supported by the IPv6 stack are still being introduced, e.g. Mobile IP [6] header processing.

Furthermore, it is relevant to recall that the solution preserves backward compatibility with nodes that can not process the new Header and Option, since communication between MEX enabled nodes and non MEX enabled nodes is possible; the imposed penalty is that the particular communication will not benefit from multi-homing. On the other hand, most routers would not need to be upgraded in order to support MEX as it has already been mentioned.

Among the detected benefits of adopting MEX we can highlight the following:

Zero packet loss. MEX preserves established communications without packet loss when outages occur because every packet contains all the needed information to be re-routed to alternative paths.

Scalability. MEX presents good scalability features, since information about multiple paths toward multi-homed sites is stored in hosts within the multi-homed site and transmitted to correspondent nodes only when needed. Besides, Extension Header processing can be located in selected up-graded devices, using the sink-route mechanism explained above. This allows load sharing among as many devices as necessary in order to support the required load.

Robustness. No state information is required by MEX capable routers in order to process the AP Extension Header, since the alternative route information is contained in the packet itself. Provided that the additional state information introduced by MEX is stored only in end-hosts, this solution satisfies the fate-sharing principle presented in [9]. This means that no new single point of failure is introduced by the mechanism in the network, since any MEX capable router can process any Extension Header independently, because no previous information is required for its processing. The absence of critical state in the network allows the mechanism to be extremely simple.

Cost distribution. We have previously presented the costs of MEX as being the additional overhead introduced and the required upgrade of the involved devices (hosts and routers). At this point, we would like to focus on how these costs are distributed. Currently deployed BGP multi-homing solution has been commonly called the "Tragedy of the Commons", since the provision of a multi-homing solution for a few sites negatively affects all the Internet community. This is definitely not the case with MEX, since multi-homing costs are strictly paid by those who benefit from it, without affecting other parties. On one hand, the additional overhead introduced by MEX is limited to communications that obtain multi-homing benefits, so both ends of the communication are free to decide if they are willing to pay for it. On the other hand, AP Extension Header processing is exclusively performed by the same ISPs who are actually carrying the traffic from/to multi-homed sites, meaning that they obtain some form of economical benefit from doing it.

Incremental deployment. As it has been previously stated, MEX preserves backward compatibility with non MEX capable hosts, allowing legacy hosts to communicate with MEX capable hosts, which enables an incremental deployment of the solution. Also, not all routers need to be upgraded in order to support the solution, since Extension Header processing will be placed in specific devices.

Policy. MEX is based on the usage of one prefix per provider. This means that addresses with a given prefix will be routed through the correspondent provider. By selecting the prefix used, the provider and the route are selected, enabling a per host policy definition. Consequently, while MEX does not provide explicit mechanisms to

express policy, its multi-address nature provide means to route identification, which enables rich policy expression.

5. Implementation

In order to validate the presented solution and to demonstrate the simplicity of the proposal, a prototype of the mechanism has been implemented. There are two roles to be implemented to provide a full MEX implementation, namely, the end-host role and the router role. The MEX capable end-host must be able to generate and receive packets with both AP Destination Option and AP Extension Header. The MEX capable router must be able to process an AP Extension Header when the destination address of the packet to be forwarded is unreachable.

The following functions have been implemented into a KAME-FreeBSD 4.5 kernel in order to build a MEX capable end-host:

The *IPv6_input.c* module receives and parses IPv6 packets. This module has been updated to be able to process incoming packets carrying the newly defined Destination Option, so that alternative prefix information is extracted from the Destination Option and cached for future packets. This multiple Prefix information will be used for building the AP Extension Header when packets are sent to the host that has generated the Destination Option. Additionally, when the module receives a packet carrying the AP Extension Header, it extracts the prefix information and verifies that all prefixes contained in the Extension Header are assigned to its interface. If this is the case, the packet processing continues, otherwise the packet is discarded.

The *IPv6_output.c* module, which is the module that conforms IPv6 exit packets, has been modified to include the AP Destination Option, when multiple addresses have been configured in an interface. Additionally, the module has been modified so that if additional prefixes are cached for a given destination, they are included in an AP Extension Header within the packet.

Alternative prefix information can also be obtained through the DNS, when multiple addresses sharing the same Interface Identifier part but with multiple global prefixes are returned from a AAAA-query. To cope with this, the *getaddrinfo.c* function has been modified so that such information is stored in the mentioned cache through a new system call to the OS kernel.

The MEX router functionality has been integrated into the *IPv6_forward.c* module, which is the module that performs IPv6 packet forwarding. The added mechanism is triggered when a packet with an unreachable destination address is found. In this case, the module inspects the packet looking for an AP Extension Header. If this header is found, it is processed by swapping the prefix contained in the Destination Address field of the IPv6 header with the prefix information contained in the Extension Header. Then the modified packet is forwarded to the new destination.

As it can be seen from the description of the changes required to provide MEX functionality, the implementation effort is low. The prototype implementation has been tested in a local testbed comprising several FreeBSD boxes working as routers and hosts, showing that the behavior of the solution is as expected. No side-effects

have been detected due to the changes performed in the host and router implementations. Trials involving larger environments and complex topologies are required for further functional validation. Besides, the solution performance should also be evaluated in more demanding environments.

6. Related Work.

In this section, we will consider alternative approaches proposed to tackle the IPv6 multi-homing problem. A straightforward option is to extend the currently used IPv4 multi-homing techniques to IPv6. However, these techniques have already exhibited scalability limitations in the IPv4 Internet. Considering that IPv6 extended address space will foster the growth of the number of sites with public addresses, it is reasonable to expect that the IPv6 BGP routing table will be larger than the IPv4 one if no aggressive address aggregation mechanisms applied.

A more restrictive approach, compatible with PA addressing, is presented in [10]. If we apply this mechanism to the multi-homed site depicted in Figure 2, the solution consists on building a tunnel between an *ISPC* exit router and *RB*, and another tunnel between *ISPD* exit router and *RA*. Then if, for instance, *link3* is down, packets are forwarded through the tunnel to *RB*. In this case, alternative route information is only stored in routers connecting ISPs with multi-homed sites, so scalability of the global routing system is preserved. However, this solution presents limited fault tolerance capabilities, since it only preserves established communications when directly connected links fail (*link3* or *link4*), but it does not protects the multi-homed site in case of another failure mode.

The Host Centric Multi-homing proposal that is being developed in [11] provides some of the multi-homing benefits through proper use of available tools. It also deals with the problem caused by ingress filtering to multi-address solutions. This is basically caused when packets containing a source address from the ISPC block are coursed from the multi-homed site through ISPD (in the example of figure 2). In this case, ingress filtering configured in ISPD ingress router will discard those packets because their source address is considered to be spoofed. The Host Centric Multi-homing approach 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, we consider that both proposals complement each other, since they address different aspects of the multi-homing problem.

7. Conclusions

In this article we have presented MEX, a novel approach to provide IPv6 multi-homing facilities, based on the transmission of the information needed to re-route packets through alternative paths in the packets themselves. The proposed solution presents fault tolerance capabilities, being able to preserve established communications through outages in the currently used path, without packet loss,

providing the quality needed to support multimedia communications. This functionality can only be provided by packet re-routing, which in turn can only be performed by the routing system. However, re-routing of packets requires the alternative route information to be available at the router involved. The currently deployed IPv4 multi-homing solution stores the alternative route information in the routing system, presenting important scalability limitations. MEX instead grants overall scalability by storing alternative route information in the end-hosts involved, and conveys this information to the routers through the AP Extension Header included in the packets flowing to the multi-homed site. In the long term, the cost of this approach is the additional overhead introduced by the Extension Header. The trade-off is then established between bandwidth and global routing table space, in the sense that, in order to provide a solution capable of surviving outages without packet loss, alternative route information must be either carried in packets or stored in routers. Global routing table space is a scarce and expensive resource, as the Internet community has so painfully learnt back in the early 90's. Nowadays, global table routing size is more and more critical, since because of its own size, BGP reconvergence times had become higher than retransmission timeouts of typical transport layers and applications [12], implying packet losses and connection timeouts when an outage occurs. Considering the ever-increasing tendency of available bandwidth, it is the authors' opinion that trading bandwidth by routing system stability is a sensible trade-off. Moreover, bandwidth consumption is limited to the involved parties, i.e. parties that are obtaining the multi-homing benefits, while the cost of storing alternative route information in the global routing table is paid by the whole Internet Community.

References

[1] G. Huston, "Commentary on Inter-Domain Routing in the Internet", RFC 3221, 2001.
[2] B. Black et al. "Goals for IP Multihoming Architectures", Internet Draft, work-in-progress, 2002.
[3] S. Deering et al."Internet Protocol, Version 6 Specification", RFC 2460, December 1998.
[4] V. Fuller et al. "Classless Inter-Domain Routing (CIDR): An Address Assignment and Aggregation Strategy", RFC 1519, 1993.
[5] R. Hinden et al."An IPv6 Aggregatable Global Unicast Address Format", RFC 2374, 1998.
[6] D. Johnson et al. "Mobility Support in IPv6", Internet Draft-Work-in-progress, 2003.
[7] A. Conta et al. "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 2463, 1998.
[8] S. Thomson et al. "IPv6 Stateless Address Autoconfiguration", RFC 2462, 1998
[9] B. Carpenter, "Architectural Principles of the Internet", RFC 1958, 1996
[10] J. Hagino et al. "IPv6 Multihoming Support at Site Exit Routers", RFC 3178, 2001.
[11] C. Huitema et al. "Host-Centric IPv6 Multihoming", Internet Draft-Work-in-progress, 2002.
[12] C. Labovitz et al."Delayed Internet Routing Convergence", ACM SIGCOMM 2000, 2000.