A simple host centric solution for a network research multihoming environment

Antonio Tapiador, Tomás de Miguel, María J. Perea, Omar Walid, Marco Hernández, David Fernández

Abstract— **IPv6 routing system has been designed to provide the required scalability to manage the huge address space needed to satisfy future networking requirements. Scalability is based on a strongly hierarchical addressing and routing system. However, this model makes difficult the management of multihomed organizations -the ones maintaining two or more connections to the network. In this article several multihoming scenarios are analyzed and a simple host centric solution that complements the source address selection mechanism is proposed. The proposal has been implemented and integrated in Linux as part of USAGI kernel.**

Keywords— **multihoming, IP routing, IPv6**

I. Introduction

INTERNET has experienced an amazing growth in the last years. Its overall architecture evolves to accommodate new technologies to support the growing number of users, applications and services. This deployment, in addition to the connectivity expectations of multiple devices in the next years, has exceeded the address space estimations made in 1978 for the current Internet network layer protocol, IPv4.

Additionally, many organizations have more than one access connection to Internet, provided by several Internet Service Providers. Each ISP supplies the organization with its own range of IP addresses and probably with its own access link (ADSL, cable, wireless or gigabit Ethernet). The use of more than one connection increases reliability and allows load balancing. Ideally if a site has two connections, you would like roughly half the traffic to go through each connection. You would also like to have failure resistant routing: when one connection goes down, the other one keeps you connected to the Internet. In an ideal scenario, you would like to maintain your connectivity when at least one of your connections is still working.

The goal of multihoming is to provide solutions to manage multiple connections and support load balancing. Multihoming offers redundancy in Internet access if each network node knows where to route the packets for each connection.

Multihoming in IPv4 was traditionally achieved by means of the use of provider independent (PI) addresses ranges announced through all ISP connections of a multihomed site. For example, if an organization is multihomed by A and B ISPs, their PI address prefixes are announced to A and B. In this way, if there is a failure in A ISP, external hosts may connect to the multihomed site through B, being the BGP interdomain routing the responsible to provide failure tolerance.

Dept. of Telematic Systems Engineering (DIT), Technical University of Madrid (UPM)

Fig. 1. Multihoming service provided by several ISPs

A. Next Generation Internet Protocol

IPv6, the new Internet network protocol designed by IETF tries to solve the lack of addresses, increasing widely the address space for the next generation Internet. In addition, it provides multiple features and facilities, such as address auto-configuration, mobility, security or QoS capabilities.

In order to improve the network scalability, IPv6 routing architecture has been redesigned following a strongly hierarchical model. Address aggregation is the key design premise, requiring address delegation from providers to organizations.

Usually, organizations are connected to the network only in one point in the hierarchy, receiving only one address prefix from its provider. However, an important number of organizations need two or more connection points to the network. For example, multinational organizations connected to different countries providers or, as it is UPM case, university departments connected to production and experimental research networks. In this case, organizations inherit one address range from each ISP connection.

Due to the hierarchical nature of the IPv6 routing model, IPv4 multihoming solutions are not valid in IPv6 networks, as it is not possible to announce all the prefixes through all ISP connections.

This behaviour would break address aggregation, as each multihomed organization prefix would be differ-

ent, and could not be aggregated into a single route. This would lead to an unmanageable size of routing tables, especially in *border routers* (routers in the Internet core that do not have a default route in their routing tables), breaking the whole interdomain routing system. Therefore, new solutions to multihoming scenarios are required for the IPv6 deployment, in order to keep addresses aggregation.

The article presents multihoming solutions in IPv6 and describes a host centric solution implemented and tested over different network scenarios. An evaluation environment is also described, that has been deployed in order to create and check easily the proposed solution. This work has been partially carried out in the context of Euro6IX research project.

II. Multihoming solutions for IPv6

IPv6 demands new solutions to the multihoming problem, not only from the addresses aggregation point of view, but also to solve other critical issues [1]:

- A site should be able to keep isolated from certain failure modes within one or more transit providers, as well as failures in the interconnection among transit providers.
- The organization should be able to perform traffic balance between its Internet connections attending not only to load balance but also to no-technical reasons.
- The transport level sessions should be preserved, and DNS modification should be affordable.
- The solution should be immune against fake source address packets filtering performed by service providers.
- The impact on hosts and routers should be minimum and isolated from the rest of elements. In the same way, interaction between hosts and routing systems should be simple, scalable and secure.

These are not the unique requirements to be considered. Non technical considerations about simple management or security should also be taken into account in order to select the proper solution.

Multiple IPv6 multihoming solutions have been proposed. They try to cover most of the multihoming requirements described previously, each one having their own advantages and disadvantages.

Multihoming solutions have been classified according to the layer where the problem is dealt with:

- Network layer solutions
- Transport layer solutions
- Complete solutions

A. Network layered solutions

Network based solutions try to solve multihoming as a routing problem, derived from the usage of several network addresses in communications in an environment where source address ingress filtering is used by ISPs, and source address selection issues have to be solved. These solutions do not solve transport session survival.

Multihoming support at site exit routers solution [2] is based on the use of tunnels between site exit routers and ISP access routers. It is able to survive one or more ISP link failures by diverting the traffic through those tunnels. It needs ISPs collaboration with multihoming sites for the tunnel establishment.

The *Host Centric Multihoming* draft [3] tries to find a solution to the multihoming problem that works only within end systems. It explores solutions from two points of view:

- *Site exit issue*: In this case, different ways of avoiding ISPs ingress filters are approached: the (improbable) removal of those anti-spoofing filters by ISPs, the use of source address routing inside the site network, the use of algorithms to allow hosts to find the right source address or the rewriting of packet header by exit routers.
- *Solutions to provide a rapid reaction to topology changes*: The draft discusses about the problem of selecting the most suitable site exit path. The solutions proposed rely on the routing system, or on allowing hosts to explore existing paths and decide. Besides, hybrid approaches where hosts choose between available paths provided by routing system are mentioned.

Finally, the Host Centric draft combines both points of view, proposing different solutions depending on the size of the multihomed organization.

In *Router Renumbering* proposal [4] the routers would deprecate addresses as they become invalid. In this way, this solution prevents the use of invalid source addresses by hosts.

In *NAROS* [5] a centralized server maintains the routing information of the multihomed site. Whenever a new communication is established, hosts query the NAROS server in order to obtain the source address they have to use.

B. Transport layered solutions

Solutions inside this category try to solve the transport session survival problem. As IP addresses are used as transport level connection identifiers, as well as to calculate checksums, any change of the IP address in the middle of a transport session will invalidate packets and provoke the session closing.

There are some proposals in this area:

- *LIN6* [6] defines a transport level identifier which is dynamically mapped to the different locators (IP addresses) that may be used by a host, solving in this way the transport level survival problem.
- *Multiple Address Service for Transport (MAST)* [7] provides a set of messages in order to establish associations between multiple IP addresses during the transport session life time.
- *Host Identity Protocol (HIP)* [8] provides an identifier to which the transport level socket is bound. This identifier does not change during the transport session. Different IP addresses may be used, which must be validated if they are new addresses. A Forwarding Agent is described for situations in which the help from the network is required.

• A similar approach is taken by *Weak Identifier Multihoming Protocol (WIMP)* proposal [9]. Based on opportunistic security principles, it uses a lighter and more efficient cryptographic operations than the ones proposed in HIP, but less secure as well.
• *Stream Control Transport Protocol (SCTP)* [10] provides fault tolerance at networking layer supporting multihoming in both sides of the communication. It supports multiple connections with different source and destination IP addresses.
• Finally, an *extension to Transport Control Protocol (ETCP)* is a modification to the current Internet connection-oriented protocol that allows the use of several IP addresses per transport session.

C. Complete solutions

Multi Homing Aliasing Protocol (MHAP) [11] is a network layer protocol that tries to solve the IPv6 multihoming problem by defining a routing architecture with addresses aliasing. A new address space is defined: multihoming addresses. Multihomed hosts would use only one of this address for communication. Out of its multihomed domain, packets would be translated to singlehomed addresses for global routing. Before reaching the final host, packet would be rewriten to use its original multihoming source address. The whole network infrastructure would need to be deployed.

III. Host Centric Multihomed solution

As mentioned before, whenever a site network is connected to the IPv6 Internet through two or more ISPs, the network inherits several network prefixes. In the case of host based multihoming solutions, that prefix multiplicity reaches the hosts, which have to cope with the fact of having several global IPv6 addresses, apart from site-local ones.

When an application inside a multihomed host tries to establish a new connection, it just provides the destination address, being the IP stack responsible of choosing the source address. If there are multiple possible source addresses, the selection typically depends on the destination address. However, the main problem to be solved here is the definition of the procedure to predict the best source address for each connection.

RFC3484 [12] discusses about default address selection on host centric multihoming scenarios, defining a procedure to select the source address depending on the destination address. That algorithm is implemented inside multihomed source nodes: routers and destination nodes are not affected.

The algorithm for selecting the source address is defined by a set of rules that try to find most suitable source address for each case. Basically, the rules defined in the proposed standard are:
1. Prefer same address
2. Prefer appropriate scope
3. Avoid deprecated addresses
4. Prefer home addresses to care-of addresses
5. Prefer same outgoing interface
6. Prefer matching label in *Policy table*
7. Prefer public addresses to temporary addresses
8. Use longest matching prefix

The sixth rule opens the possibility to add new rules to the decision algorithm. It is based on a Policy Table, which may be configured by the network administrator to associate different destination networks to the different source addresses, in order to define and use specific source address when communicating with some destination network or addresses.

The use of a Policy Table can be useful to solve the multihoming problem in some situations, for example, in the case of UPM. The University gets its Internet access service from the national research network, RedIRIS, but it is also connected to other IPv6 research project networks, such as Euro6IX, which provide its own address space. Research networks are more suitable for certain destinations (typically project partners) than the standard one. Source address selection for this case can be easily configured by adding the adequate entries to the Policy table, as shown below.

A. Source Address Selection implementation

IPv6 capable USAGI [13] kernel implements RFC 3484, in particular, what it is related to the "source address selection by host" solution. This implementation was initially tested, verifying its correct behaviour according to RFC 3484, but finding that it lacks a user space configuration tool in order to allow the modification of the policy table. Without this configuration capability, RFC 3484 becomes insufficient to manage a multihoming scenario which needs some kind of differentiation between its global-scope source addresses.

In the context of the Euro6IX research project, UPM has developed some extensions to the USAGI kernel, together with a user space configuration tool -named *addrlabel*- that allows the dynamic modification of the source address selection policy table. Entries to that table, which are static in the original USAGI kernel, can be added, modified or deleted dynamically using the configuration tool, which internally uses "`ioctl`" calls to communicate whith the kernel.

Besides, support in `/proc` has been added; the Policy Table can be directly read through `/proc/net/addrselect_label_table`.

From the begining we found interesting to integrate the modifications made into the USAGI project kernel. Therefore, a patch was send to USAGI, who included it since August 18th, 2003 snapshot.

In summary, using addrlabel tool the network administrator is able to define and modify the source address selection policies of a multihomed host based on destination network's prefixes.

B. Testing Enviroment

In order to test and demonstrate the extensions to the source address selection mechanism implemented a web demonstration scenario [14] based on the VNUML

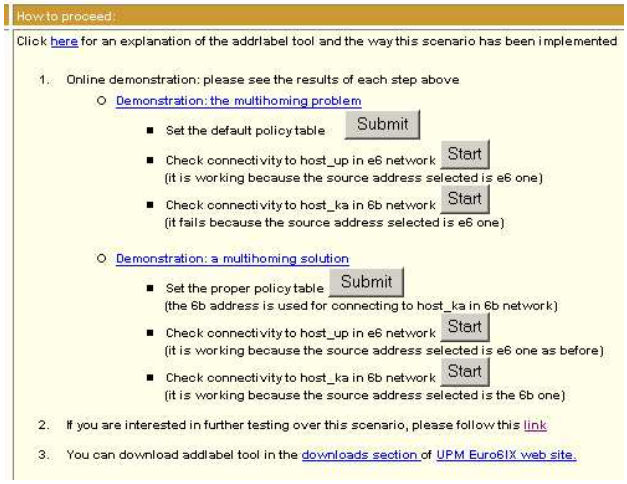

Fig. 2. DIT-UPM Euro6IX multihoming Looking Glass detail.

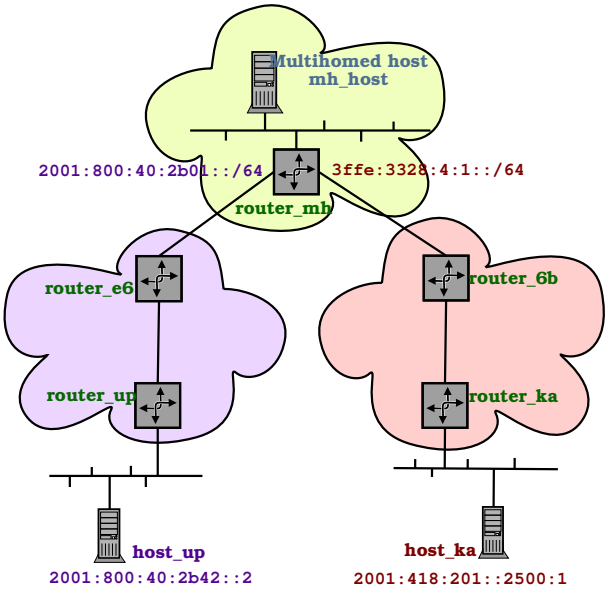

Fig. 3. Simulation scenario

network emulation tool [15] was developed. VNUML tool allows the creation of Linux based virtual network scenarios inside one machine. VNUML is based on User Mode Linux (UML) [16] which allows running a completely functional Linux kernel run as a conventional user process inside a standard Linux kernel. Several such virtual machines can be run concurrently inside the same physical machine. Each one can be configured with the desired number of network interfaces, as well as the topology that interconnects them. Besides, interconnection with external networks through the physical interface of the hosting machine is also possible.

The main advantage of UML is its high flexibility and transparency: processes running in UML behave as if they were running in a real scenario. The main disadvantage is the penalty in performance (in terms of CPU usage, physical memory. and storage capability of the hosting box) due to the necessity of having an underlying kernel in order to run a process in the emulated scenario.

Although UML is a powerful general-purpose tool, its use is too complex for a user to build scenarios including many virtual machines and complex virtual network topologies. Furthermore, good knowledge of some Linux operating system details (tap devices, UNIX sockets, virtual bridging, etc.) is needed to start an emulated scenario "by hand". In order to make the use of UML easier for emulating network scenarios, we developed Virtual Network UML (VNUML) tool [17].

Using VNUML, we created the testing scenario shown in Figure 3, made of eight virtual machines (hosts and routers) and several virtual networks interconnecting them. The whole scenario is defined using the VNUML XML-based specification language and can be started and stopped easily by just executing one command.

Besides, we modified our previously developed Looking Glass application to allow users to access the virtual testing scenarios form a web page and to be able to directly interact with the extensions to the USAGI kernel and the user space configuration tool.

The multihoming demonstration scenario is organized in the following way (see Figure 2):
1. Problem: the multihoming problem is presented: the multihomed host `mh_host` can reach `host_up` but is not able to reach `host_ka`. As `mh_host` follows the basic "Source address selection algorithm", it selects the wrong source address because it is the longest matching prefix address with `host_ka` address, as specified in rule 8.
2. Solution: to solve the problem, a new suitable entry in the policy table of the multihomed host (rule 6) is added, to tell the host it must use `3ffe:3328:4:1::2` address when communicating with `2001:410:201::/64` network, despite `2001:800:40:2b01::2` source address is available.
3. Test: after adding the entry to the Policy Table, the multihomed host is able to reach the destination, since the source address selected has been set correctly. Now the multihomed host can simultaneously reach both remote hosts.

As mentioned before, the demonstration can be accessed through a standard web browser. Remote users can either follow the predefined demo or directly play with the kernel Policy table over a more flexible interface. The results of all the operations made are displayed on web pages.

IV. Advantages

In a static multihoming scenario, the use of the modified USAGI kernel with the configuration tool addrlabel allows the establishment of the proper rules on the Policy table, resolving the problems derived from the source address selection and providing an usable multihoming solution.

It is a ready-to-use solution that only requires the installation of the modified USAGI kernel and a sim-

ple tool in end systems. The network administrator may define a configuration file containing the adequate source address selection rules. This file has to be copied to all multihomed hosts and loaded from a script. Such Policy table rules would be set in all network hosts.

Nevertheless, as network state is not static (link states and access service characteristics depend on the time, for example, whenever a network congestion problem or link failure arises) the configuration of the Policy table should follow network variations.

This would create dependency between the validity of the solution and how quickly a network administrator changes the Policy table, according to network changes. Changes in Policy table could be automated somehow, but in any case, it would still require the network administration intervention, which is costly, inefficient and unreliable.

Another disadvantage of the solution adopted is the lack of compliance with transport survival requirement. This would probably be achieved by joining the adopted solution together with one of the proposed solution that tries to solve this issue.

V. Future Works

The first objective would be the definition of a procedure to automatically update the rules defined in the Policy Table of a host, according to the network state or the network administrator needs.

Currently, we are working on a XML specification, including two multihoming problem aspects: host address selection and site ingress filtering problem. With this language, the network administration would be able to define the general policies in multihoming, considering both technical and non-technical issues.

Once the multihomed site policy is defined, the problem is how to provide this information to hosts.

One approach would consist on using Router Advertisements, maybe by means of a new option as proposed in draft "Default Router Preferences, More-Specific Routes, and Load Sharing" [18], which defines a new option in Router Advertisements frames that instruct the hosts about the routers they should use to reach certain destinations. A similar option could be defined in order to indicate the source address to use with each destination.

A configured router or a centralized server could be used to organize the Router Advertisements generation.

Another interesting improvement is to fix the problems that arise in a scenario where hosts receive Router Advertisements from two or more routers, thus having two or more possible gateways.

It seems acceptable to use as gateway the router that announced the prefix elected as source address for the outgoing packet. It is supposed that, if a source address seems to be the most appropriate, the outgoing path defined by the ISP that provides the prefix will be the most appropriate too, given the host is selecting the source address according to routing information.

In this situation, the main concern is the size of the Policy Table, because it may become extremely large in the case we try to contain the whole Internet.

We may think that a way to solve this, similar to the way routers correct a host when it makes the wrong routing decision is the use of redirect packets. A similar packet may be defined, informing the hosts about the best source address to use the next time they want to communicate with a given destination.

Some steps have been taken on this direction with the proposal of a new "destination unreachable" code for ICMPv6: "source address failed ingress policy" in the discussion derived in the IETF IPv6 working group list[1]. A router may correct the source address used by the host sending a "source address failed" ICMP packet which source address prefix matches the suitable source address.

With these router corrections, the Policy table would have a general configuration and routers would be left to do the small fixes.

Nevertheless, all this would not solve the transport survival problem. A change in the source address in the middle of a transport session would break it, because as mentioned source address is used in packet checksum calculation and as a session identifier.

The solution to this problem is more complicated, as it suggests changes related to the Internet protocol structure. There are some suggestions to solve this problem, by inserting a new level. It can be a session level that would allow a source address change in the middle of the session: it would maintain several transport sessions being used according to network state.

Or it could be a wedge level between net and transport, which, regardless the network address used in each moment, it would provide a constant identifier to transport level.

Other possibility comes up by the use of a new transport protocol that takes into account the multiple addresses available, like SCTP.

Anyway, these are long term solutions, as they involve changes in Internet architecture, and the deployment of the chosen solution in every communicating host.

VI. Summary

Multihoming is a frequently demanded capability in today's networks. Therefore, it is a must to design flexible and efficient solutions to that problem.

But the multihoming solutions used for the old Internet protocol are no longer desirable as they break address aggregation and hierarchical routing. New IPv6 specific multihoming solutions must be found that meet several requirements, like traffic policy routing or transport session survival.

[1]ICMPv6: New destination unreachable codes. https//www1.ietf.org/mail-archive/working-groups/ipv6/current/msg01431.html

There are multiple IPv6 solutions being proposed, each one solving certain problem, all having its own advantages and problems.

UPM's contribution has focused on the implementation of a default address selection solution valid for static scenarios that solves a specific multihoming problem commonly found in network research contexts. However, a few enhancements and improvements may be added to this solution, in order to get a dynamic and fully scalable one.

Acknowledgments

This work presented in this article has been partially funded by the European Union in the context of the European IPv6 Internet Exchanges Backbone (IST Euro6IX) project.

References

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

[2] J. Hagino, H. Snyder, *IPv6 Multihoming Support at Site Exit Routers*, IETF RFC 3178, October 2001.

[3] C. Huitema, R. Draves, M. Bagnulo, *Host-Centric IPv6 Multihoming*, IETF draft-huitema-multi6-hosts-03, February 2004. Work in progress.

[4] M. Crawford, *Router Renumbering for IPv6*, IETF RFC 2894, August 2000.

[5] C. de Launois, O. Bonaventure and M. Lobelle, *The NAROS Approach for IPv6 Multihoming with Traffic Engineering*, http://www.info.ucl.ac.be/people/delaunoi/naros/

[6] Location Independent Networking for IPv6. http://www.lin6.net/

[7] D. Crocker, *Multiple Address Service for Transport*, IETF draft-crocker-mast-proposal-02, October 2003. Work in progress.

[8] R. Moskowitz, P. Nikander, P. Jokela, T. Henderson, *Host Identity Protocol*, IETF draft-moskowitz-hip-09, February 2004. Work in progress.

[9] J. Ylitalo, V. Torvinen, E. Nordmark, *Weak Identifier Multihoming Protocol*, IETF draft-ylitalo-multi6-wimp-00, January 2004. Work in progress.

[10] Stream Control Transmission Protocol (SCTP). http://www.sctp.org/

[11] M. Py, *Multi Homing Aliasing Protocol (MHAP) intro*, draft-py-mhap-intro-00.txt, March 2003.

[12] R. Draves, *Default Address Selection for Internet Protocol version 6 (IPv6)*, IETF RFC 3484, February 2003.

[13] USAGI Project Linux IPv6 Development Project, http://www.linux-ipv6.org/

[14] Looking Glass Multihoming Test, http://www.upm.euro6ix.org/cgi-bin/looking-glass-upm-v0.6/ntools.pl

[15] Virtual Network User Mode Linux (VNUML) emulation tool, http://www.dit.upm.es/vnuml

[16] J. Dike, *User Mode Linux*, Proc. 5th Annual Linux Showcase & Conf., Oakland CA, 2001.

[17] F. Galan, D. Fernandez and T. de Miguel, *Study and Emulation of IPv6 Internet-Exchange-Based Addressing Models*, IEEE Communications Magazine, pp 105-112. January 2004.

[18] R. Draves, D. Thaler, *Default Router Preferences and More-Specific Routes*, IETF draft-ietf-ipv6-router-selection-03, December 2003. Work in progress.