Dual Homing Experiment

Christian Huitema, David Kessens

1 Simple dual homing problem statement

In order to make progress towards an IPv6 site multi-homing solution, we propose to start with a simple problem: how to multi-home the simplest site, i.e. a site composed of a single link, e.g. a switched Ethernet network.

We believe that addressing the problem of simple, small networks is both urgent and useful. Providing a simple solution for the most numerous networks drastically reduces the danger that generalized multi-homing poses to the Internet. A small network solution may or may not be easily extensible to larger networks, but we know that there are order of magnitude less large networks than small. If we have a solution for the small networks, the scope of the remaining problem will "only" be the hundreds of thousands of larger networks, rather than the hundreds of millions of home networks.

Many small business networks use a simple pattern of multi-homing today: the very small business network is composed of a switched Ethernet, combining fixed line and wireless links; it connects to the Internet through two independent connections, such as two DSL lines. There is no coordination between the two providers. In a typical IPv4 set-up, each Internet connection terminates in a different NAT/Firewall; one connection is used as back-up, and is turned on if the other connection fails.

Other examples of this simple multi-homing pattern include home networks connected to a broadband service but also maintaining a dial-up connection for back-up; and home networks connected to two broadband connections such as DSL and cable, or perhaps wireless services.

2 Define a simple set of requirements

For the small networks, the basic motivation for multi-homing is redundancy and reliability. Reliability is defined, as a minimum, as the possibility to maintain operation in cases of failure of one the Internet provider connections.

We will expect the site to contain both "simple" and "smart" hosts. Simple hosts have an implementation of IPv6 conformant to RFC 2641; smart hosts are supposed to be upgraded to become "site multi-homing aware". The reliability expectation is slightly different for the two kinds of hosts. In case of failure of a provider connection, the simple hosts may suffer from a temporary loss of connectivity; the expectation is that operation will shortly resume over the remaining connection; this matches the expectation of current IPv4 "back-up" services. On the contrary, the smart hosts are expected to obtain better services: established TCP-IP connections should survive the loss of one of the provider connections; if connections cannot be maintained, operation should at least resume immediately; it should be possible to perform some amount of load balancing between the two connections.

We assume that it is entirely acceptable that the two Internet providers each allocate a "provider aggregated" prefix to the small site. Simple hosts are expected to auto-configure addresses using at least one of these prefixes; smart hosts are expected to auto-configure addresses using both. It is reasonable to expect that hosts have to be renumbered after a connectivity event, although we also expect that addresses configured with one prefix will remain valid as long as this prefix is valid.

The following section details specific problems that must be solved by the simple multi-homing solution.

3 Multi-homing issues

The five multihoming issues that we want to solve are ingress filtering, avoidance of a "dead connection" for outbound connections, avoidance of a dead address for inbound connections, connection maintenance and exit selection. We believe that the solution to these problems is in the realm of engineering, not research. For each problem, we give hint of one or several possible solutions. We also believe that relatively simple solutions are possible, and should be preferred to "innovative" constructs.

3.1 Ingress filtering

We assume that each of the ISP may perform ingress filtering, and will reject packets whose source address does not belong to the prefix allocated to the network by that ISP. This leads to a possible failure scenario:

- Host choose source address A
- Default route leads to router B
- Router B forwards the packet to ISP B
- Packet is dropped by ingress filtering.

The multi-homing proposal must present a solution to this problem. We believe that this solution can be engineered by simple improvements in the small site exit routers, such as checking the source address of the Internet bound packets and redirecting these packets to the "right" exit router in case of problem. This solution should not depend on any particular host behavior, beyond what is mandated by RFC 2641; however, smart hosts may improve on the basic solution, e.g. by selecting a next hop exit router as a function of the source address.

3.2 Avoid using dead router for outbound connections

The purpose of multi-homing is to provide redundancy, and to avoid the following failure scenario:

- Router A is dead, or link to ISP A is dead
- Host continue sending using source address A
- Because of ingress filtering, router B cannot forward the packets

The multi-homing proposal must present a solution to this problem, and the solution must work for both simple hosts and smart hosts. We believe that a solution can be engineered by an appropriate use of "router announce" messages, such as stopping advertisement of a prefix if the connection is unavailable, or possibly advertising this prefix as "deprecated." Smart host may improve on that solution by implementing detecting the loss of connectivity independently of the router advertisements, and automatically preferring the other address for new connections.

3.3 Avoid using dead address for outbound connections

A variation of the previous failure scenario occurs when a service is hosted on the small network:

- Service advertises both address A and address B in the DNS
- Router A is dead, or link to ISP A is dead
- A remote client selects address A to reach the service
- The connection fails

The multi-homing proposal must present a solution to this problem, and the solution must work for both simple hosts and smart hosts. A possible solution is to expect remote hosts to retry the connection attempt using the second address; another solution is to somehow update the DNS.

3.4 Maintaining connections

Solving the ingress filtering and dead router problem provides small site with a functional multi-homing solution, but does not resolve the following failure scenario:

- Host has connection with peer P using address A
- Router A or ISP A fails
- Existing connections break

We accept the loss of connections in the case of the simple hosts; the simple hosts will simply have to reestablish their connections. However, the multi-homing proposal should provide smart hosts with a solution to this problem. We believe that it is possible to engineer this solution by using a variation of Mobile IPv6.

3.5 Use the right exit/entrance

Even if we are able to provide redundancy and reliability, the dual homed networks are still exposed to the following failure mode:

- Peer picks address A rather than address B
- Resulting traffic is routed on a slower path or on a congested connection
- Performance looks terrible

The multi-homing proposal should provide a solution to this problem, and this solution should be available at least to smart hosts. The goal is not to obtain an absolute optimization of network usage, but rather to avoid the most obnoxious results of load imbalance. We believe that it is possible to engineer a solution using either a smart naming service (a.k.a. DNS load balancing), a routing level optimization such as a randomized choice of exit routers, or possibly a variation of Mobile IPv6 to move existing connections towards a less loaded path.

4 Further study

The goal of the simple dual homing exercise is to provide a simple solution to the simple sites. Once this goal is achieved, we believe that the solution can be extended in two directions: extend the solution to medium size sites; and extend the solution to accommodate a form of provider independent addressing.

The extension to medium size sites can be thought as a progress on a scale of complexity: first, extend the solution to multilink subnets, where all exit routers are on the same link; second, extend the solution to routed network, where all exit routers are on the same link;

and third, extend the solution to routed network, where exit routers are present at different locations.

Provider independent addressing is often used today by large IPv4 sites, which simply obtain a prefix form a registry, use this prefix for internal addresses, and insure that the prefix is distributed in the global routing tables. We believe that similar solutions will be available for large IPv6 sites, but we are also very aware that they cannot be extended to the majority of smaller IPv6 sites. Yet, the small IPv6 sites may benefit from some form of provider independent addresses, e.g. in order to advertise addresses that don't depend on a particular provider configuration. It may be possible for these sites to use a "virtual IP" solution, in which connectivity through a PI address is overlaid on top of the regular provider based addressing. This can be thought of as an extension of the simple solution.