Site Multihoming in IPv6 (multi6)
Internet-Draft
Expires: November 30, 2004

J. Abley
ISC
B. Black
Layer8 Networks
V. Gill
AOL
K. Lindqvist
Netnod Internet Exchange

June 2004

IPv4 Multihoming Motivation, Practices and Limitations
draft-ietf-multi6-v4-multihoming-01

Status of this Memo

Copyright Notice

Abstract

Multihoming is an essential component of service for enterprises which are part of the Internet. This draft describes some of the motivations, practices and limitations of multihoming as it is achieved in the IPv4 world today. The analysis is done in order to

serve as underlaying documentation to the discussions in the "Site multihoming for IPv6" workinggroup of the IETF, who are working to a longerterm solution to some of the issues that arise from doing multihoming in the ways as are described here.

1. Introduction

Multihoming is an important component of service for many enterprises which are part of the Internet. Current IPv4 multihoming practices have been added on to the CIDR architecture [1], which assumes that routing table entries can be aggregated based upon a hierarchy of customers and service providers.

Multihoming is a mechanism by which enterprises can currently satisfy a number of high-level requirements, and is widely used in the IPv4 network today. There are some practical limitations, however, including concerns of how well (or, if) the current practice will scale as the network continues to grow.

The preferred way to multihome in IPv4 is to announce an independent block of address space over two or more ISPs using BGP. Until the mid-1990s this was relatively easy to accomplish, as the maximum generally accepted prefix length in the global routing table was a /24, and little justification was needed to receive a /24. However, in 1995 the growth of the global routing table became a problem once again, and as a result some providers decided to start filtering prefixes it accepted from peers based on prefix length. This broke the expectation that a multihomed network announcing a /24, regardless of where in the IPv4 address space this /24 was taken from, would be globally reachable.

This practice has two advantages and one disadvantage for the multihomed network. The first advantage is that they can obtain a much smaller block of address space from an ISP than from a RIR. (Would-be multihomers still often optimize their networks for qualifying for at least a /24 by adopting accepted but relatively wasteful address deployment strategies.) The second advantage is that even if their announcement is filtered, they are still reachable over the primary ISP by virtue of the aggregate announced by this ISP. Even when the circuit to the primary ISP is down, this often works because the primary ISP will generally accept the announcement over the secondary ISP, so traffic flows from the filtering network to the primary ISP and then to the secondary ISP in order to arrive at the multihomed network. While this is common, it is also not universally true.

The disadvantage is that the multihomed network must depend on the primary ISP for the aggregate. If the primary ISP goes down, this will impact reachability to networks that filter. And when the multihomed network leaves the primary ISP, they are generally expected to return the address space because otherwise this ISP would have to route traffic for a non-customer. Most ISPs will cooperate with this "punching holes in an aggregate" solution to multihoming,

but some are reluctant.

2. Terminology

An "enterprise" is an entity autonomously operating a network using TCP/IP and, in particular, determining the addressing plan and address assignments within that network. This is the definition of "enterprise" used in [2].

A "transit provider" is an enterprise which provides connectivity to the Internet to one or more other enterprises. The connectivity provided extends beyond the transit provider's own network.

A "multi-homed" enterprise is one with more than one transit provider. "Multihoming" is the practice of being multi-homed.

A "multi-attached" enterprise is one with more than one point of layer-3 interconnection to a single transit provider.

The term "re-homing" denotes a transition of an enterprise between two states of connectedness, due to a change in the connectivity between the enterprise and its transit providers.

3. Motivations for Multihoming

3.1 Redundancy

By multihoming, an enterprise can insulate itself from certain failure modes within one or more transit providers, as well as failures in the network providing interconnection with one or more transit providers.

Examples of failure modes from which an enterprise can obtain some degree of protection by multi-homing are:

- Physical link failure, such as a fiber cut or router failure,
- Logical link failure, such as a misbehaving router interface,
- Routing protocol failure, such as a BGP peer reset,
- Transit provider failure, such as a backbone-wide IGP failure, and
- Exchange failure, such as a BGP reset on an inter-provider peering.

Some of these failure modes may be protected against by multi-attaching to a single transit provider, rather than multi-homing.

3.2 Load Sharing

By multihoming, an enterprise can distribute both inbound and outbound traffic between multiple transit providers.

Sometimes it is not possible to increase transit capacity to a single transit provider because that provider does not have sufficient spare capacity to sell. In this case a solution is to acquire additional transit capacity through a different provider. This scenario is common in bandwidth-starved stubs of the Internet where, for example, transit demand outpaces under-sea cable deployment.

3.3 Performance

By multihoming, an enterprise can protect itself from performance difficulties between transit providers.

For example, suppose enterprise E obtains transit from transit providers T1 and T2, and there is long-term congestion between T1 and T2. By multihoming between T1 and T2, E is able to ensure that in normal operation none of its traffic is carried over the congested interconnection T1-T2.

3.4 Policy

An enterprise may choose to load-share for a variety of policy

reasons outside technical scope (e.g. cost, acceptable use conditions, etc).

For example, enterprise E homed to transit provider T1 may be able to identify a particular range of addresses within its network that correspond to non-real-time traffic (e.g. a network containing mail and Usenet/NNTP servers). It may be advantageous to shift inbound traffic destined for that range of addresses to transit-provider T2, since T2 charges less for traffic.

3.5 Independency

Enterprises might also choose to multihome in order to achive independence of some sort. Independence can here mean policy, financial or administrative. This need for independence vary, and so does the reasons for it. Some common examples are

- Ease of (or not having to) renumbering.
- Avoiding upstream peering policy in order to have other/shorter paths.
- Stronger negotiation position with upstreams due to easier migration.
- Appearance. Large enterprises might have "marketing" reasons to show independence from any given provider.

4. Current methods used for IPv4 multihoming

In IPv4 there are today a number of ways that an enterprise that wants to do multihoming can achieve this. These methods can broadly be split into five categories as described below

4.1 Multihoming with your own addresses and AS

The most commonly used method is to multihome to two or more providers, announcing provider independent (PI) IP addresses, or addresses allocated (PA) from a regional Internet registry (RIR) to the enterprise. These addresses are announced as sourced from an autonomos system (AS) number that belongs to the entrprise.

4.2 Multihoming with your own AS, but PA addresses

The most likely secondly most common approach used today is to use an autonomous system number that belongs to the entrprise, and using that announcing addresses that belongs to one of the upstreams. That is, the enterprise gets allocated an addressblock from one of it's upstreams. The enterprise then announce those addresses, as a more specific route than the providers aggregated address block. This route is announced to all teh upstreams of the enterprise, including the provider that allocated the addresses.

4.3 Multihoming with your own addresses, and private AS

A third possible way of multihoming is with addresses owned by the enterprise whising to multihome, but advertising them without having a public AS, or with no AS at all. This is done with the enterprise either sourcing the prefixes in a private AS [3], and having their upstreams remove those on announcement to the rest of the world, or the upstreams simply sourcing the prefixes in their AS and then routing to the organization.

4.4 Multiple attachments to the same ISP

Fourth option is to have multiple connection to the same ISP. This is fairly popular, but will not have an impact on the global routing table as both paths are covered by the ISPs aggregate route. An enterprise that have solved their multihoming needs in this way is commonly reffered to as "multi-attached".

4.5 NAT or RFC2260 based multihoming

This last method might very well be the most commonly used method in terms of volume. Simply because this is what most residential users are normally referred to. This method uses addresses from each of

the upstream that an organization is connected to. Either the addresses are allocated to nodes inside the network according to the proposal in [4], or the enterprise uses NAT to translate into private addresses insde the enterprise.

5. Features of IPv4 Multihoming

The following section analyses some of the features driving the choices for various multihoming approaches in todays IPv4 Internet. As the "Site multihoming for IPv6" working group progresses, they will have to take similiar considerations into approach, learning from IPv4. This considerations are listed in [5], and some of the operational considerations that needs to be thought of for new multihoming mechanisms can be found in [6].

5.1 Simplicity

The current methods used as multihoming solutions are not all without complexity, but in practice it quite straightforward to deploy and maintain by virtue of the fact that they are well-known, tried and tested.

5.2 Transport-Layer Survivability

The current multihoming solution provides session survivability for transport-layer protocols; i.e. exchange of data between devices on the multi-homed enterprise network and devices elsewhere on the Internet may proceed with no greater interruption than that associated with the transient packet loss during a re-homing event.

New transport-layer sessions are able to be created following a re-homing event.

5.3 Inter-Provider Traffic Engineering

A multi-homed enterprise may influence routing decisions beyond its immediate transit providers by advertising a strategic mixture of carefully-aimed long prefixes and covering shorter-prefix routes. This precise effects of such egress policy are often difficult to predict, but an approximation of the desired objective is often easy to accomplish. This can provide a similar mechanisms to that described in Section 3.3, except that the networks whose traffic is being influenced are not transit providers of the enterprise itself.

5.4 Load Control

The current multihoming solution places control of traffic flow in the hands of the enterprise responsible for the multi-homed interconnections with transit providers. A single-homed customer of a multi-homed enterprise may vary the demand for traffic that they impose on the enterprise, and may influence differential traffic load between transit providers; however, the basic mechanisms for congestion control and route propagation are in the hands of the

enterprise, not the customer.

5.5 Impact on Routers

The routers at the boundary of a multi-homed enterprise are usually required to participate in BGP sessions with the interconnected routers of transit providers. Other routers within the enterprise have no special requirements beyond those of single-homed enterprises' routers.

5.6 Impact on Hosts

There are no requirements of hosts beyond those of single-homed enterprises' hosts.

5.7 Interactions between Hosts and the Routing System

There are no requirements for interaction between routers and hosts beyond those of single-homed enterprises' routers and hosts.

6. Limitations of IPv4 Multihoming

6.1 Scalability

Current IPV4 multihoming practices contribute to the significant growth currently observed in the state held in the global inter-provider routing system; this is a concern both because of the hardware requirements it imposes and also because of the impact on the stability of the routing system.

These mechanisms also add to the consumption of public AS number resources, when small enterprises wishing to multihome obtain an AS number specifically for only that purpose. Using a different mechanism would help to conserve the 16-bit AS number space, and avoid the move to 32-bit AS numbers.

This issue is discussed in great detail in [7].

7. Security Considerations

This memo analyzes the IPv4 multihoming practices. This analysis only includes the description of the mechanisms and partially how they affect the availability of the enterprise deploying the IPv4 multihoming mechanism. Other security properties of the IPv4 multihoming mechanisms are not analyzed.

8. Acknowledgements

Thanks goes to Pekka Savola and Iljitsch van Beijnum for providing feedback and suggestions on the text as well as text.

9 Informative references

[1] Fuller, V., Li, T., Yu, J. and K. Varadhan, "Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy", RFC 1519, September 1993.

[2] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G. and E. Lear, "Address Allocation for Private Internets", RFC 1918, February 1996.

[3] Hawkinson, J. and T. Bates, "Guidelines for creation, selection, and registration of an Autonomous System (AS)", RFC 1930, March 1996.

[4] Bates, T. and Y. Rekhter, "Scalable Support for Multi-homed Multi-provider Connectivity", RFC 2260, January 1998.

[5] Black, B., Gill, V. and J. Abley, "Goals for IP Multihoming Architectures", RFC 3582, August 2003.

[6] Lear, E., "Things MULTI6 Developers should think about", Internet-Drafts draft-ietf-multi6-things-to-think-about-00, June 2004.

[7] Huston, G., "Analyzing the Internet's BGP Routing Table", January 2001.

Authors' Addresses

Joe Abley
ISC
2204 Pembroke Court
Burlington, ON L7P 3X8
Canada

Phone: +1 905 319 9064
EMail: jabley@isc.org

Benjamin Black
Layer8 Networks

EMail: ben@layer8.net

Vijay Gill
AOL
12100 Sunrise Valley Dr
Reston, VA 20191
US

Phone: +1 410 336 4796
EMail: vgill@vijaygill.com

Kurt Erik Lindqvist
Netnod Internet Exchange
Bellmansgatan 30
Stockholm S-118 47
Sweden

Phone: +46 8 615 85 70
EMail: kurtis@kurtis.pp.se

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.