multi6   J. Abley
Internet-Draft   ISC
Expires: October 16, 2003   B. Black
                                                          Layer8 Networks
                                                                 V. Gill
                                                          AOL Time Warner
                                                          April 17, 2003

Goals for IPv6 Site-Multihoming Architectures
draft-ietf-multi6-multihoming-requirements-05

Status of this Memo

Copyright Notice

Abstract

Site-multihoming, i.e. connecting to more than one IP service provider, is an essential component of service for many sites which are part of the Internet.

This document outlines a set of goals for proposed new IPv6 site-multihoming architectures.

1. Introduction

Current IPv4 site-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.

However, it appears that this hierarchy is being supplanted by a dense mesh of interconnections [7]. Additionally, there has been an enormous growth in the number of multihomed sites. For purposes of redundancy and load-sharing, the multihomed address blocks, which are almost always a longer prefix than the provider aggregate, are announced along with the larger, covering aggregate originated by the provider. This contributes to the rapidly-increasing size of the global routing table. This explosion places significant stress on the inter-provider routing system.

Continued growth of both the Internet and the practice of site-multihoming will seriously exacerbate this stress. The site-multihoming architecture for IPv6 should allow the routing system to scale more pleasantly.

2. Terminology

A "site" is an entity autonomously operating a network using IP and, in particular, determining the addressing plan and routing policy for that network. This definition is intended to be equivalent to "enterprise" as defined in [2].

A "transit provider" operates a site which directly provides connectivity to the Internet to one or more external sites. The connectivity provided extends beyond the transit provider's own site. A transit provider's site is directly connected to the sites for which it provides transit.

A "multihomed" site is one with more than one transit provider. "Site-multihoming" is the practice of arranging a site to be multihomed.

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

3. Multihoming Goals

3.1 Capabilities of IPv4 Multihoming

The following capabilities of current IPv4 multihoming practices should be supported by an IPv6 multihoming architecture.

3.1.1 Redundancy

By multihoming, a site should be able to insulate itself from certain failure modes within one or more transit providers, as well as failures in the network providing interconnection among one or more transit providers.

Infrastructural commonalities below the IP layer may result in connectivity which is apparently diverse sharing single points of failure. For example, two separate DS3 circuits ordered from different suppliers and connecting a site to independent transit providers may share a single conduit from the street into a building; in this case backhoe-fade of both circuits may be experienced due to a single incident in the street. The two circuits are said to "share fate".

The multihoming architecture should accommodate (in the general case, issues of shared fate notwithstanding) continuity of connectivity during the following failures:

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

3.1.2 Load Sharing

By multihoming, a site should be able to distribute both inbound and

outbound traffic between multiple transit providers. This requirement is for concurrent use of the multiple transit providers, not just the usage of one provider over one interval of time and another provider over a different interval.

3.1.3 Performance

By multihoming, a site should be able to protect itself from performance difficulties directly between the site's transit providers.

For example, suppose site E obtains transit from transit providers T1 and T2, and there is long-term congestion between T1 and T2. The

multihoming architecture should allow E to ensure that in normal operation none of its traffic is carried over the congested interconnection T1-T2. The process by which this is achieved should be a manual one.

A multihomed site should be able to distribute inbound traffic from particular multiple transit providers according to the particular address range within their site which is sourcing or sinking the traffic.

3.1.4 Policy

A customer may choose to multihome for a variety of policy reasons beyond technical scope (e.g. cost, acceptable use conditions, etc.) For example, customer C homed to ISP A may wish to shift traffic of a certain class or application, NNTP, for example, to ISP B as matter of policy. A new IPv6 multihoming proposal should provide support for site-multihoming for external policy reasons.

3.1.5 Simplicity

As any proposed multihoming solution must be deployed in real networks with real customers, simplicity is paramount. The current multihoming solution is quite straightforward to deploy and maintain. A new IPv6 multihoming proposal should not be substantially more complex to deploy and operate than current IPv4 multihoming practices.

3.1.6 Transport-Layer Survivability

Multihoming solutions should provide re-homing transparency for transport-layer sessions; i.e. exchange of data between devices on the multihomed site and devices elsewhere on the Internet may proceed with no greater interruption than that associated with the transient packet loss during the re-homing event.

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

Transport-layer sessions include those involving transport-layer protocols such as TCP, UDP and SCTP over IP. Applications which communicate over raw IP and other network-layer protocols may also enjoy re-homing transparency.

3.1.7 Impact on DNS

Multi-homing solutions either should be compatible with the observed dynamics of the current DNS system, or the solutions should

demonstrate that the modified name resolution system required to support them are readily deployable.

3.1.8 Packet Filtering

Multihoming solutions should not preclude filtering packets with forged or otherwise inappropriate source IP addresses at the administrative boundary of the multihomed site.

3.2 Additional Requirements

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

A new IPv6 multihoming architecture should scale to accommodate

orders of magnitude more multihomed sites without imposing unreasonable requirements on the routing system.

3.2.2 Impact on Routers

The solution may require changes to IPv6 router implementations, but these changes should be either minor, or in the form of logically separate functions added to existing functions.

Such changes should not prevent normal single-homed operation, and routers implementing these changes should be able to interoperate fully with hosts and routers not implementing them.

3.2.3 Impact on Hosts

The solution should not destroy IPv6 connectivity for a legacy host implementing RFC 3513 [3], RFC 2460 [5], RFC 2553 [6] and other basic IPv6 specifications current in April 2003. That is to say, if a host can work in a single-homed site, it should still be able to work in a multihomed site, even if it cannot benefit from site-multihoming.

It would be compatible with this goal for such a host to lose connectivity if a site lost connectivity to one transit provider, despite the fact that other transit provider connections were still operational.

If the solution requires changes to the host stack, these changes

should be either minor, or in the form of logically separate functions added to existing functions.

If the solution requires changes to the socket API and/or the transport layer, it should be possible to retain the original socket API and transport protocols in parallel, even if they cannot benefit from multihoming.

The multihoming solution may allow host or application changes if that would enhance session survivability.

3.2.4 Interaction between Hosts and the Routing System

The solution may involve interaction between a site's hosts and its

routing system; such an interaction should be simple, scaleable and securable.

3.2.5 Operations and Management

It should be posssible for staff responsible for the operation of a site to monitor and configure the site's multihoming system.

3.2.6 Cooperation between Transit Providers

A multihoming strategy may require cooperation between a site and its transit providers, but should not require cooperation (relating specifically to the multihomed site) directly between the transit providers.

3.2.7 Multiple Solutions

There may be more than one approach to multihoming, provided all approaches are orthogonal (i.e. each approach addresses a distinct segment or category within the site multihoming problem). Multiple solutions will incur a greater management overhead, however, and the adopted solutions should attempt to cover as many multihoming scenarios and goals as possible.

4. Security Considerations

A multihomed site should not be more vulnerable to security breaches than a traditionally IPv4-multihomed site.

Any changes to routing practices made to accommodate multihomed sites should not cause non-multihomed sites to become more vulnerable to security breaches.

Normative 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] Hinden, R. and S. Deering, "Internet Protocol Version 6 (IPv6) Addressing Architecture", RFC 3513, April 2003.

[4] Hinden, R., O'Dell, M. and S. Deering, "An IPv6 Aggregatable Global Unicast Address Format", RFC 2374, July 1998.

[5] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[6] Gilligan, R., Thomson, S., Bound, J. and W. Stevens, "Basic Socket Interface Extensions for IPv6", RFC 2553, March 1999.

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

Authors' Addresses

Joe Abley
Internet Software Consortium
10805 Old River Road
Komoka, ON N0L 1R0
Canada

Phone: +1 519 641 4368
EMail: jabley@isc.org

Benjamin Black
Layer8 Networks

EMail: ben@layer8.net

Vijay Gill
AOL Time Warner

EMail: vijaygill9@aol.com

Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on the IETF's procedures with respect to rights in standards-track and standards-related documentation can be found in BCP-11. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementors or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

Full Copyright Statement

Acknowledgement

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