IPv6 Working Group   R. Hinden
INTERNET-DRAFT   Nokia
October 18, 2004   D. Thaler
Expires April 2005   Microsoft

IPv6 Host to Router Load Sharing
<draft-ietf-ipv6-host-load-sharing-03.txt>

Status of this Memo

Copyright Notice

Abstract

The original IPv6 conceptual sending algorithm does not do load-sharing among equivalent IPv6 routers, and suggests schemes which can be problematic in practice. This document updates the conceptual sending algorithm so that traffic to different destinations can be distributed among routers in an efficient fashion.

1. Introduction

In the conceptual sending algorithm in [ND] and in the optional extension in [ROUTERSEL], a next hop is chosen when no destination cache entry exists for an off-link destination or when communication through an existing router is failing. Normally a router is selected the first time traffic is sent to a specific destination IP address. Subsequent traffic to the same destination address continues to use the same router unless there is some reason to change to a different router (e.g., a redirect message is received, or the router is found to be unreachable).

In addition, as described in [ADDRSEL], the choice of next hop may also affect the choice of source address, and hence indirectly (and to a lesser extent) may affect the router used for inbound traffic as well.

In both the base sending algorithm and in the optional extension, sometimes a host has a choice of multiple equivalent routers for a destination. That is, all other factors are equal and a host must break a tie via some implementation-specific means.

It is typically desirable when there is more than one equivalent router that hosts distribute their outgoing traffic among these routers. This shares the load among multiple routers and provides better performance for the host's traffic.

On the other hand, load sharing can be undesirable in situations where sufficient capacity is available through a single router and the traffic patterns could be more predictable by using a single router; in particular, this helps to diagnose connectivity problems beyond the first-hop routers.

[ND] does not require any particular behavior in this respect. It specifies that an implementation may always choose the same router (e.g., the first in the list) or may cycle through the routers in a round-robin manner. Both of these suggestions are problematic.

Clearly, always choosing the same router does not provide load sharing. Some problems with load sharing using naive tie-breaking techniques such as round-robin and random are discussed in [MULTIPATH]. While the destination cache provides some stability since the determination is not per-packet, cache evictions or timeouts can still result in unstable or unpredictable paths over time, lowering the performance and making it harder to diagnose problems. Round-robin selection may also result in synchronization issues among hosts, where in the worst case the load is concentrated on one router at a time.

In the remainder of this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in [RFC2119].

2. Load Sharing

When a host chooses from multiple equivalent routers, it SHOULD support choosing using some method which distributes load for different destinations among the equivalent routers rather than always choosing the same router (e.g., the first in the list). This memo takes no stance on whether the support for load sharing should be turned on or off by default. Furthermore, a host that does attempt to distribute load among routers SHOULD use a hash-based scheme which takes the destination IP address into account, such as those described in [MULTIPATH], for choosing a router to use.

Note that traffic for a given destination address will use the same router as long as the Destination Cache Entry for the destination address is not deleted. With a hash-based scheme, traffic for a given destination address will use the same router over time even if the Destination Cache Entry is deleted, as long as the list of equivalent routers remains the same.

3. Security Considerations

As mentioned in [MULTIPATH], when next-hop selection is predictable, an application can synthesize traffic that will all hash the same, making it possible to launch a denial-of-service attack against the load sharing algorithm, and overload a particular router. This can even be done by a remote application that can cause a host to respond to a given destination address. A special case of this is when the same (single) next-hop is always selected, such as in the algorithm allowed by [ND]. Introducing hashing can make such an attack more difficult; the more unpredictable the hash is, the harder it becomes to conduct a denial-of-service attack against any single router.

However, a malicious local application can bypass the algorithm for its own traffic by using mechanisms such as raw sockets, and remote attackers can still overload the routers directly. Hence, the mechanisms discussed herein have no significant incremental impact on Internet infrastructure security.

4. IANA Considerations

This document has no actions for IANA.

5. Acknowledgments

The authors of this document would like to thank Erik Nordmark, Brian Haberman, Steve Deering, Aron Silverton, Christian Huitema, and Pekka Savola.

6. Normative References

[ND] Narten, T., Nordmark, E. and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[RFC2119]
Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, BCP0014, March 1997.

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

7. Informative References

[MULTIPATH]
Thaler, D. and C. Hopps, "Multipath Issues in Unicast and Multicast Next-Hop Selection", RFC 2991, November 2000.

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

8. Authors' Addresses

Robert Hinden
Nokia
313 Fairchild Drive
Mountain View, CA 94043
Phone: +1 650 625-2004
Email: bob.hinden@nokia.com

Dave Thaler
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
Phone: +1 425 703 8835
EMail: dthaler@microsoft.com

9. Full Copyright Statement

10. Intellectual Property