IPv6 Maintenance
Internet-Draft
Updates: 4861 (if approved)
Intended status: Standards Track
Expires: March 6, 2016

F. Baker
Cisco Systems
B. Carpenter
Univ. of Auckland

September 3, 2015

Host routing in a multi-prefix network
draft-baker-6man-multi-homed-host-03

Abstract

This note describes expected IPv6 host behavior in a network that has more than one prefix, each allocated by an upstream network that implements BCP 38 ingress filtering, when the host has multiple routers to choose from. It also applies to other scenarios such as the usage of stateful firewalls that effectively act as address-based filters.

This host behavior may interact with source address selection in a given implementation, but logically follows it. Given that the network or host is, or appears to be, multihomed with multiple provider-allocated addresses, that the host has elected to use a source address in a given prefix, and that some but not all neighboring routers are advertising that prefix in their Router Advertisement Prefix Information Options, this document specifies to which router a host should present its transmission. It updates RFC 4861.

Table of Contents

1. Introduction and Applicability

This note describes the expected behavior of an IPv6 [RFC2460] host in a network that has more than one prefix, each allocated by an upstream network that implements BCP 38 [RFC2827] ingress filtering, and in which the host is presented with a choice of routers. It expects that the network will implement some form of egress routing, so that packets sent to a host outside the local network from a given ISP's prefix will go to that ISP. If the packet is sent to the wrong egress, it is liable to be discarded by the BCP 38 filter. However, the mechanics of egress routing once the packet leaves the host are

out of scope. The question here is how the host interacts with that network.

BCP 38 filtering by ISPs is not the only scenario where such behavior is valuable. The combination of existing recommendations for home gateways [RFC6092] [RFC7084] can also result in such filtering. Another case is when the connections to the upstream networks include stateful firewalls, such that return packets in a stream will be discarded if they do not return via the firewall that created state for the outgoing packets. A similar cause of such discards is unicast reverse path forwarding (uRPF) [RFC3704].

In this document, the term "filter" is used for simplicity to cover all such cases. In any case, one cannot assume the host to be aware whether an ingress filter, a stateful firewall, or any other type of filter is in place. Therefore, the only safe solution is to implement the features defined in this document.

Note that, apart from ensuring that a message with a given source address is given to a first-hop router that appears to know about the prefix in question, this specification is consistent with [RFC4861]. Nevertheless, implementers of Sections 5.2, 6.2.3, 6.3.4 and 8 of RFC 4861 will need to extend their implementations accordingly. This specification is fully consistent with [RFC6724] and implementers will need to add support for its Rule 5.5. Hosts that do not support these features may fail to communicate in the presence of filters as described above.

1.1. Requirements Language

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

2. Sending context expected by the host

2.1. Expectations the host has of the network

A host receives prefixes in a Router Advertisement [RFC4861], which goes on to identify whether they are usable by SLAAC [RFC4862] [RFC4941] [RFC7217]. When no prefixes are usable for SLAAC, the Router Advertisement would normally signal the availability of DHCPv6 [RFC3315] and the host would use it to configure its addresses. In the latter case (or if both SLAAC and DHCPv6 are used on the same link for some reason) it will be generally the case that the configured addresses match one of the prefixes advertised in a Router Advertisement that are supposed to be in that link.

The simplest multihomed network implementation in which a host makes choices among routers might be a LAN with one or more hosts on it and two or more routers, one for each upstream network, or a host that is served by disjoint networks on separate interfaces. In such a network, especially the latter, there is not necessarily a routing protocol, and the two routers may not even know that the other is a router as opposed to a host, or may be configured to ignore its presence. One might expect that the routers may or may not receive each other's RAs and form an address in the other router's prefix (which is not per [RFC4862], but is implemented by some stub router implementations). However, all hosts in such a network might be expected to create an address in each prefix so advertised.

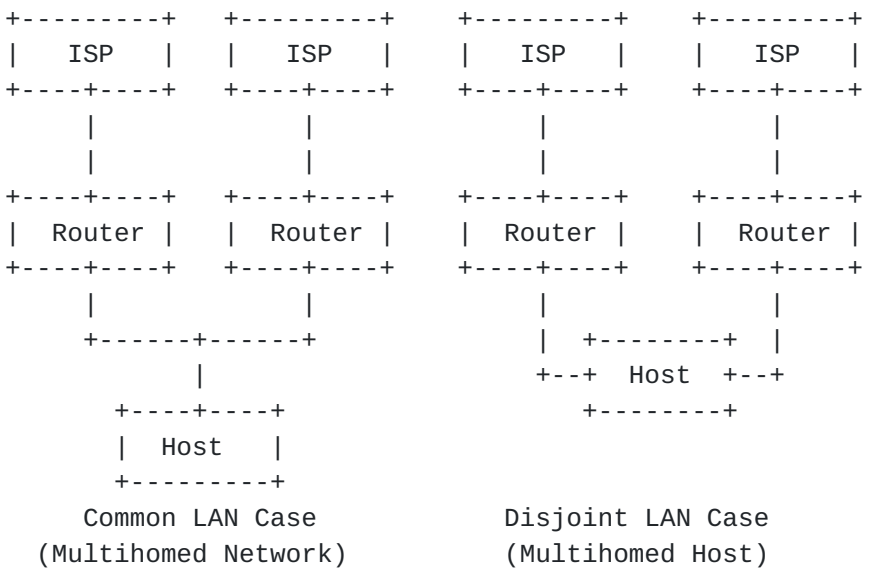

Figure 1: Two simple networks

If there is no routing protocol among those routers, there is no mechanism by which packets can be deterministically forwarded between the routers (as described in BCP 84 [RFC3704]) in order to avoid filters. Even if there was routing, it would result in an indirect route, rather than a direct route originating with the host; this is not "wrong", but can be inefficient. Therefore the host would do well to select the appropriate router itself.

Since the host derives fundamental default routing information from the Router Advertisement, this implies that, in any network with hosts using multiple prefixes, each prefix SHOULD be advertised via a Prefix Information Option (PIO) [RFC4861] by one of the attached routers, even if addresses are being assigned using DHCPv6. A router that advertises a prefix indicates that it is able to appropriately route packets with source addresses within that prefix, regardless of

the setting of the L and A flags in the PIO. In some circumstances both L and A might be zero.

Although this does not violate the existing standard [RFC4861], such a PIO has not previously been common, and it is possible that existing host implementations simply ignore such a PIO or that a router implementation rejects such a PIO as a configuration error. Newer implementations that support this mechanism will need to be updated accordingly: a host SHOULD NOT ignore a PIO simply because both L and A flags are cleared; a router SHOULD be able to send such a PIO.

2.2. Expectations of multihomed networks

The direct implication of Section 2.1 is that routing protocols used in multihomed networks SHOULD be capable of source-prefix based egress routing, and that multihomed networks SHOULD deploy them.

3. Reasonable expectations of the host

3.1. Default Router Selection

Default Router Selection is modified as follows: A host SHOULD select default routers for each prefix it is assigned an address in. Routers that have advertised the prefix in its Router Advertisement message SHOULD be preferred over routers that do not advertise the prefix.

As a result of doing so, when a host sends a packet using a source address in one of those prefixes and has no history directing it otherwise, it SHOULD send it to the indicated default router. In the "simplest" network described in Section 2.1, that would get it to the only router that is directly capable of getting it to the right ISP. This will also apply in more complex networks, even when more than one physical or virtual interface is involved.

In more complex cases, wherein routers advertise RAs for multiple prefixes whether or not they have direct or isolated upstream connectivity, the host is dependent on the routing system already. If the host gives the packet to a router advertising its source prefix, it should be able to depend on the router to do the right thing.

3.2. Source Address Selection

There is an interaction with Default Address Selection [RFC6724]. Rule 5.5 of that specification states that the source address used to send to a given destination address should if possible be chosen from

a prefix known to be advertised by the first-hop router for that destination. This selection rule would be applicable in a host following the recommendation in the previous paragraph.

3.3. Redirects

There is potential for adverse interaction with any off-link Redirect (Redirect for a GUA destination that is not on-link) message sent by a router in accordance with Section 8 of [RFC4861]. Hosts SHOULD apply off-link redirects only for the specific pair of source and destination addresses concerned, so the host's Destination Cache may need to contain appropriate source-specific entries.

3.4. History

Some modern hosts maintain history, in terms of what has previously worked or not worked for a given address or prefix and in some cases the effective window and MSS values for TCP or other protocols. This might include a next hop address for use when a packet is sent to the indicated address.

When such a host makes a successful exchange with a remote destination using a particular address pair, and the host has previously received a PIO that matches the source address, then the host SHOULD include the prefix in such history, whatever the setting of the L and A flags in the PIO. On subsequent attempts to communicate with that destination, if it has an address in that prefix at that time, a host MAY use an address in the remembered prefix for the session.

4. Residual issues

Consider a network where routers on a link run a routing protocol and are configured with the same information. Thus, on each link all routers advertise all prefixes on the link. The assumption that packets will be forwarded to the appropriate egress by the local routing system might cause at least one extra hop in the local network (from the host to the wrong router, and from there to another router on the same link).

In a slightly more complex situation such as the disjoint LAN case of Figure 1, which happens to be one of the authors' home plus corporate home-office configuration, the two upstream routers might be on different LANs and therefore different subnets (e.g., the host is itself multi-homed). In that case, there is no way for the "wrong" router to detect the existence of the "right" router, or to route to it.

In such a case it is particularly important that hosts take the responsibility to memorize and select the best first-hop as described in Section 3.

5. IANA Considerations

This memo asks the IANA for no new parameters.

6. Security Considerations

This document does not create any new security or privacy exposures. It is intended to avoid connectivity issues in the presence of BCP 38 ingress filters or stateful firewalls combined with multihoming.

There might be a small privacy improvement, however: with the current practice, a multihomed host that sends packets with the wrong address to an upstream router or network discloses the prefix of one upstream to the other upstream network. This practice reduces the probability of that occurrence.

7. Acknowledgements

Comments were received from Jinmei Tatuya and Ole Troan, who have suggested important text, plus Mikael Abrahamsson, Steven Barth, Juliusz Chroboczek, Toerless Eckert, David Farmer, Pierre Pfister, Mark Smith, Dusan Mudric, and James Woodyatt.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <http://www.rfc-editor.org/info/rfc2460>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <http://www.rfc-editor.org/info/rfc4861>.

[RFC6724] Thaler, D., Ed., Draves, R., Matsumoto, A., and T. Chown, "Default Address Selection for Internet Protocol Version 6 (IPv6)", RFC 6724, DOI 10.17487/RFC6724, September 2012, <http://www.rfc-editor.org/info/rfc6724>.

8.2. Informative References

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, DOI 10.17487/RFC2827, May 2000, <http://www.rfc-editor.org/info/rfc2827>.

[RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, DOI 10.17487/RFC3315, July 2003, <http://www.rfc-editor.org/info/rfc3315>.

[RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", BCP 84, RFC 3704, DOI 10.17487/RFC3704, March 2004, <http://www.rfc-editor.org/info/rfc3704>.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <http://www.rfc-editor.org/info/rfc4862>.

[RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, DOI 10.17487/RFC4941, September 2007, <http://www.rfc-editor.org/info/rfc4941>.

[RFC6092] Woodyatt, J., Ed., "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", RFC 6092, DOI 10.17487/RFC6092, January 2011, <http://www.rfc-editor.org/info/rfc6092>.

[RFC7084] Singh, H., Beebee, W., Donley, C., and B. Stark, "Basic Requirements for IPv6 Customer Edge Routers", RFC 7084, DOI 10.17487/RFC7084, November 2013, <http://www.rfc-editor.org/info/rfc7084>.

[RFC7217] Gont, F., "A Method for Generating Semantically Opaque Interface Identifiers with IPv6 Stateless Address Autoconfiguration (SLAAC)", RFC 7217, DOI 10.17487/RFC7217, April 2014, <http://www.rfc-editor.org/info/rfc7217>.

Appendix A. Change Log

Initial Version: 2015-08-05

Version 01: Update text on PIOs, added text on Redirects, and clarified the concept of a "simple" network, 2015-08-13.

Version 02: Clarifications after WG discussions, 2015-08-19.

Version 03: More clarifications after more WG discussions, especially adding stateful firewalls, uRPF, and more precise discussion of RFC 4861, 2015-09-03.

Authors' Addresses

Fred Baker
Cisco Systems
Santa Barbara, California 93117
USA

Email: fred@cisco.com

Brian Carpenter
Department of Computer Science
University of Auckland
PB 92019
Auckland 1142
New Zealand

Email: brian.e.carpenter@gmail.com