Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 9, 2009

H. Singh
W. Beebee
Cisco Systems, Inc.
E. Nordmark
Sun Microsystems

October 6, 2008

IPv6 Subnet Model: the Relationship between Links and Subnet Prefixes
draft-ietf-6man-ipv6-subnet-model-02

Abstract

IPv6 specifies a model of a subnet that is different than the IPv4 subnet model. The subtlety of the differences has resulted in incorrect implementations that do not interoperate. This document spells out the most important difference; that an IPv6 address isn't automatically associated with an IPv6 on-link prefix. This document also invalidates (partially due to security concerns) a part of the definition of on-link from [RFC4861].

Table of Contents

1. Introduction

IPv4 implementations typically associate a netmask with an address when an IPv4 address is assigned to an interface. That netmask together with the IPv4 address designates an on-link prefix. Addresses that are covered by this prefix are viewed as on-link i.e., traffic to these addresses is not sent to a router. See section 3.3.1 in [RFC1122]. Prior to the deployment of Classless Intern-Domain Routing (CIDR), an address's netmask could be derived directly from the address. In the absence of specifying a specific netmask when assigning a address, some implementations would fall back to deriving the netmask from the class of the address.

The behavior of IPv6 as specified in Neighbor Discovery [RFC4861] is quite different. The on-link determination is separate from the address assignment. A host can have IPv6 addresses without any related on-link prefixes or have on-link prefixes that are not related to any IPv6 addresses that are assigned to the host. Any assigned address on an interface should initially be considered as having no internal structure as shown in [RFC4291].

In IPv6, by default, a host treats only the link-local prefix as on-link.

The reception of a Prefix Information Option (PIO) with the L-bit set [RFC4861] and a non-zero valid lifetime creates an entry (or updates the valid lifetime for an existing entry) in the Prefix List. All the prefixes that are on the Prefix List, i.e., have not yet timed out, are on-link.

The on-link definition in the Terminology section of [RFC4861], as modified by this document, defines the complete list of cases where an address is considered on-link. Note, in particular, that Redirect Messages can also indicate an address is off-link. Individual address entries can be expired by the Neighbor Unreachability Detection mechanism.

A host only performs address resolution for IPv6 addresses that are on-link. Packets to any other address are sent to a default router. If there is no default router, then the node should send an ICMPv6 Destination Unreachable indication as specified in [RFC4861] - more details are provided in the Host Behavior and Rules section. (Note that [RFC4861] changed the behavior when the Default Router List is empty. The behavior in the old version of Neighbor Discovery [RFC2461] was different when there were no default routers.)

Failure of host implementations to correctly implement the IPv6 subnet model can result in lack of IPv6 connectivity. See the

Observed Incorrect Implementation Behavior section for details.

Host behavior is clarified in the Host Behavior and Rules section.

2. Host Behavior and Rules

A correctly implemented IPv6 host MUST adhere to the following rules:

1. By default only the link-local prefix is on-link.

2. The configuration of an IPv6 address, whether through IPv6 stateless address autoconfiguration [RFC4862], DHCPv6 [RFC3315], or manual configuration MUST NOT implicitly cause a prefix derived from that address to be treated as on-link. A host considers a prefix to be on-link only through explicit means, such as those specified in the on-link definition in the Terminology section of [RFC4861], as modified by this document, or via manual configuration. Note that the requirement for manually configured addresses is not explicitly mentioned in [RFC4861].

3. Note that the following items (from the definition of on-link in [RFC4861]):

   - a Neighbor Advertisement message is received for the (target) address, or

   - any Neighbor Discovery message is received from the address.

   are not sufficient to consider an address to be on-link and will be removed in a future update to [RFC4861]. A literal reading of the second test would allow a neighboring intruder to generate bogus ND messages that result in a spoofed address being improperly treated as on-link. This vulnerability is a specific instance of the broad set of attacks that are possible by an on-link neighbor [RFC3756]. The threat is particularly problematic in the case of routers which allow such a spoofed message to update their forwarding tables (which can happen if a neighbor cache entry can update the forwarding table). Only addresses that are covered by the modified on-link definition should be treated as on-link from a sending or forwarding perspective, and it should be noted that routers should generally obtain on-link information from sources other than RAs and Redirects.

4. To maintain consistency with the invalidation of the last two bullets of the on-link definition in [RFC4861], the following text from section 7.2.3 of [RFC4861] will also be augmented:

> If the Source Address is not the unspecified address and, on link layers that have addresses, the solicitation includes a Source Link-Layer Address option, then the recipient SHOULD create or update the Neighbor Cache entry for the IP Source Address of the solicitation.

changes to:

> If the Source Address is not the unspecified address and, on link layers that have addresses, the solicitation includes a Source Link-Layer Address option, then the recipient SHOULD create or update the Neighbor Cache entry for the IP Source Address of the solicitation provided that the source address of the NS is deemed on-link through other indications.

5. In the absence of other sources of on-link information, including Redirects, if the RA advertises a prefix with the on-link(L) bit set and later the Valid Lifetime expires, the host MUST then consider addresses of the prefix to be off-link, as specified by the PIO paragraph of section 6.3.4 of [RFC4861].

6. Newer implementations, which are compliant with [RFC4861] MUST adhere to the following rules. Older implementations, which are compliant with [RFC2461] but not [RFC4861] may remain as is. If the Default Router List is empty and there is no other source of on-link information about any address or prefix:

   1. The host MUST NOT assume that all destinations are on-link.

   2. The host MUST NOT perform address resolution for non-link-local addresses.

   3. Since the host cannot assume the destination is on-link, and off-link traffic cannot be sent to a default router (since the Default Router List is empty), address resolution cannot be performed. This case is specified in the last paragraph of section 4 of [RFC4943]: when there is no route to destination, the host should send an ICMPv6 Destination Unreachable indication (for example, a locally delivered error message) as specified in the Terminology section of [RFC4861].

   On-link information concerning particular addresses and prefixes can make those specific addresses and prefixes on-link, but does not change the default behavior mentioned above for addresses and prefixes not specified. [RFC4943] provides justification for these rules.

Using cached on-link determination information without first verifying that the information is still valid after IPv6 interface re-initialization may lead to lack of IPv6 network connectivity. For example, a host receives an RA from a router with on-link prefix A. The host reboots. During the reboot, the router sends out prefix A with on-link bit set and a zero lifetime to indicate a renumbering. The host misses the renumbering. The host comes online. Then, the router sends an RA with no PIO. The host uses cached on-link prefix A and issues NS's instead of sending traffic to a default router. The "Observed Incorrect Implementation Behavior" section below describes how this can result in lack of IPv6 connectivity.

3. Observed Incorrect Implementation Behavior

One incorrect implementation behavior illustrates the severe consequences when the IPv6 subnet model is not understood by the implementers of several popular host operating systems. In an access concentrator network ([RFC4388]), a host receives a Router Advertisement Message with no on-link prefix advertised. The host incorrectly assumes an invented prefix is on-link and performs address resolution when the host should send all non-link-local traffic to a default router. Neither the router nor any other host will respond to the address resolution, preventing this host from sending IPv6 traffic.

4. Conclusion

This document clarifies and summarizes the relationship between links and subnet prefixes described in [RFC4861]. Configuration of an IPv6 address does not imply the existence of corresponding on-link prefixes. One should also look at API considerations for prefix length as described in last paragraph of section 4.2 of [RFC4903]. This document also invalidates a part of the definition of on-link from [RFC4861].

5. Security Considerations

This document addresses a security concern present in [RFC4861]. As a result, the last two bullets of the on-link definition in [RFC4861] have been invalidated.

6. IANA Considerations

None.

7. Acknowledgements

Thanks (in alphabetical order) to Adeel Ahmed, Jari Arkko, Ralph Droms, Alun Evans, Dave Forster, Prashanth Krishnamurthy, Suresh Krishnan, Josh Littlefield, David Miles, Thomas Narten, Madhu Sudan, Jinmei Tatuya, Dave Thaler, Bernie Volz, and Vlad Yasevich for their consistent input, ideas and review during the production of this document. The security problem that provides one reason for invalidating a part of the on-link definition was found by David Miles. Thomas Narten has provided substantial guidance to the production of this document.

8. References

8.1. Normative References

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.

8.2. Informative References

[RFC1122] Braden, R., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, October 1989.

[RFC2461] Narten, T., Nordmark, E., and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

[RFC3756] Nikander, P., Kempf, J., and E. Nordmark, "IPv6 Neighbor Discovery (ND) Trust Models and Threats", RFC 3756, May 2004.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

[RFC4388] Woundy, R. and K. Kinnear, "Dynamic Host Configuration Protocol (DHCP) Leasequery", RFC 4388, February 2006.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, September 2007.

[RFC4903] Thaler, D., "Multi-Link Subnet Issues", RFC 4903,

June 2007.

[RFC4943] Roy, S., Durand, A., and J. Paugh, "IPv6 Neighbor Discovery On-Link Assumption Considered Harmful", RFC 4943, September 2007.

Appendix A. CHANGE HISTORY

[NOTE TO RFC EDITOR: PLEASE REMOVE THIS SECTION UPON PUBLICATION.]

Changes in draft-ietf-6man-ipv6-subnet-model-02.txt since -01.txt are:

- Augmented Abstract to say an important change to [RFC4861] is being made by this document.

- Removed the following sentence at the end of the Introduction section: "Finally, this document mainly restates and clarifies [RFC4861]."

- Added new bullet three to the "Host Behavior and Rules" section where the bullet invalidates bullets three and four from the on-link definition from [RFC4861].

- Added new bullet four to the "Host Behavior and Rules" section where the bullet proposes changes to text in section 7.2.3 of [RFC4861].

- The security section has been modified to reflect the important invalidation proposed by this document.

- Modified minor text in the "Observed Incorrect Implementation Behavior" section to explain what the prefix is in the second sentence.

- Changed bullet 3 from a new rule with normative language to just a paragraph of text describing behavior for a host blindly caching on-link determination and a possible severe consequence of that. The text also includes a solution for the problem. The new text lies at the end of section 2 as a new paragraph.

- The title of section 2 has been changed to Host Behavior and Rules. Also changed Host Behavior Rules to Host Behavior and Rules in two places in the Introduction section.

Changes in draft-ietf-6man-ipv6-subnet-model-01.txt since -00.txt are:

- Changed Introduction section to remove any mention of src address of ND message as a means for on-link determination. Also reworded first paragraph of Introduction section.

- Reworded bullet 2 of section 2 and added text to clarify on-link definition.

- Added text to bullet 3 of section 2 to make explicit that this is a new rule.

- Reworded bullet 5 of section 2 to clearly explain where ICMPv6 Destination Unreachable is sent to.

Authors' Addresses

Hemant Singh
Cisco Systems, Inc.
1414 Massachusetts Ave.
Boxborough, MA 01719
USA

Phone: +1 978 936 1622
Email: shemant@cisco.com
URI: http://www.cisco.com/

Wes Beebee
Cisco Systems, Inc.
1414 Massachusetts Ave.
Boxborough, MA 01719
USA

Phone: +1 978 936 2030
Email: wbeebee@cisco.com
URI: http://www.cisco.com/

Erik Nordmark
Sun Microsystems
17 Network Circle
Menlo Park, CA 94025
USA

Phone: +1 650 786 2921
Email: erik.nordmark@sun.com

Full Copyright Statement

Intellectual Property