6man
Internet-Draft
Intended status: Standards Track
Expires: April 11, 2011

M. Kohno
Juniper Networks, Keio University
B. Nitzan
Juniper Networks
R. Bush
Y. Matsuzaki
Internet Initiative Japan
L. Colitti
Google
T. Narten
IBM Corporation
October 8, 2010

Using 127-bit IPv6 Prefixes on Inter-Router Links
draft-kohno-ipv6-prefixlen-p2p-03.txt

Abstract

On inter-router point-to-point links, it is useful for security and other reasons, to use 127-bit IPv6 prefixes. Such a practice parallels the use of 31-bit prefixes in IPv4 [RFC3021]. This document specifies motivation and usages of 127-bit IPv6 prefix lengths on inter-router point-to-point links.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 11, 2011.

Copyright Notice

Table of Contents

1. Conventions Used In This Document

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 RFC 2119 [RFC2119].

2. Introduction

[RFC4291] specifies that interface IDs for all unicast address, except those that start with the binary value 000, are required to be 64 bits long and to be constructed in Modified EUI-64 format. In addition, it defines the Subnet-Router anycast address, which is intended to be used for applications where a node needs to communicate with any one of the set of routers on a link.

Some operators have been using 127-bit prefixes, but this has been discouraged due to conflicts with Subnet-Router anycast [RFC3627]. However, using 64-bit prefixes creates security issues which are particularly problematic on inter-router links, and there are other valid reasons to use prefixes longer than 64 bits, in particular /127 (see Section 5).

This document provides rationale for using 127-bit prefix lengths, reevaluates the reasons why doing so was considered harmful, and specifies how /127 prefixes can be used on inter-router links configured for use as point-to-point links.

3. Scope Of This Memo

This document is applicable to cases where operators assign specific addresses on inter-router point-to-point links and do not rely on link-local addresses. Many operators assign specific addresses for purposes of network monitoring, reverse DNS resolution for traceroute and other management tools, EBGP peering sessions, and so on.

For the purposes of this document, an inter-router point-to-point link is a link to which only two routers and no hosts are attached. This may include Ethernet links which are configured to be point-to-point. In such cases, there is no need to support Neighbor Discovery for address resolution, and other general scenarios like the use of stateless address autoconfiguration are not relevant.

Links between a router and a host, or links to which both routers and hosts are attached, are out of scope of this document.

4. Problems identified with 127-bit prefix lengths in the past

[RFC3627] discourages the use of 127-bit prefix lengths due to conflicts with the Subnet-Router anycast addresses, while stating that the utility of Subnet-Router Anycast for point-to-point links is questionable.

[RFC5375] also says the usage of 127-bit prefix lengths is not valid and should be strongly discouraged, but the stated reason for doing this is to be in compliance with [RFC3627].

Though the analyses in the RFCs are correct, operational experience with IPv6 has shown that /127 prefixes can be used successfully.

5. Reasons for using longer prefixes

There are reasons network operators use IPv6 prefix lengths greater than 64, particularly 127, for inter-router point-to-point links.

5.1. Ping-pong issue

A forwarding loop may occur on a point-to-point link with a prefix length shorter than 127. This does not affect interfaces that perform Neighbor Discovery, but some point-to-point links, which uses medium such as SONET, do not use Neighbor Discovery. As a consequence, configuring any prefix length shorter than 127 bits on these links can create an attack vector in the network.

The pingpong issue happens in case of IPv4 as well. But due to the scarcity of IPv4 address space, the current practice is to assign long prefix lengths such as /30 or /31 [RFC3021] on point-to-point links, thus the problem did not come to the fore.

The latest ICMPv6 specification [RFC4443] mitigates this problem by specifying that a router receiving a packet on a point-to-point link, which is destined to an address within a subnet assigned to that same link (other than one of the receiving router’s own addresses), MUST NOT forward the packet back on that link. Instead, it SHOULD generate an ICMPv6 Destination Unreachable message code 3 in response. This check is on the forwarding processing path, so it may have performance impact.

5.2. Neighbor Cache Exhaustion issue

As described in Section 4.3.2 of [RFC3756], the use of a 64-bit prefix length on an inter-router link that uses Neighbor Discovery (e.g., Ethernet) potentially allows for denial-of-service attacks on

the routers on the link.

Consider an Ethernet link between two routers A and B to which a /64 subnet has been assigned. A packet sent to any address on the /64 (except the addresses of A and B) will cause the router attempting to forward it to create an new cache entry in state INCOMPLETE, send a Neighbor Solicitation message to be sent on the link, start a retransmit timer, and so on [RFC4861].

By sending a continuous stream of packets to a large number of the 2^64 - 3 unassigned addresses on the link (one for each router and one for Subnet-Router Anycast), an attacker can create a large number of neighbor cache entries and send a large number of Neighbor Solicitation packets which will never receive replies, thereby consuming large amounts of memory and processing resources. Sending the packets to one of the 2^24 addresses on the link which has the same Solicited-Node multicast address as one of the routers also causes the victim to spend large amounts of processing time discarding useless Neighbor Solicitation messages.

Careful implementation and rate-limiting can limit the impact of such an attack, but are unlikely to neutralize it completely. Rate-limiting neighbor solicitation messages will reduce CPU usage, and following the garbage-collection recommendations in [RFC4861] will maintain reachability, but if the link is down and neighbor cache entries have expired while the attack is ongoing, legitimate traffic (for example, BGP sessions) over the link might never be re-established because the routers cannot resolve each others' IPv6 addresses to MAC addresses.

This attack is not specific to point-to-point links, but is particularly harmful in the case of point-to-point backbone links, which may carry large amounts of traffic to many destinations over long distances.

While there are a number of ways to mitigate this kind of issue, assigning /127 subnets eliminates it completely.

5.3. Other reasons

Though address space conservation considerations are less important for IPv6 than they are in IPv4, some operators prefer not to assign /64s to individual point-to-point links. Instead, they may be able to number all of their point-to-point links out of a single (or small number of) /64s.

6. Recommendations

Routers MUST support the assignment of /127 prefixes on point-to-point inter-router links.

When assigning and using any /127 prefixes, the following considerations apply.Some addresses have special meanings, in particular addresses corresponding to reserved anycast addresses. When assigning prefixes (and addresses) to links, care should be taken to ensure that addresses reserved for such purposes aren't inadvertently assigned and used as unicast addresses. Otherwise, nodes may receive packets that they are not intended to receive. Specifically, assuming that a number of point-to-point links will be numbered out of a single /64 prefix:

a) Addresses with all zeros in the rightmost 64 bits SHOULD NOT be assigned as unicast addresses, to avoid colliding with the Subnet-Router anycast address. [RFC4291]

b) Addresses in which the rightmost 64 bits are assigned the highest 128 values SHOULD NOT be used as unicast addresses, to avoid colliding with Reserved Subnet Anycast Addresses. [RFC2526]

7. Security Considerations

Section 5.1 and 5.2 discuss about seurity related issues.

8. IANA Considerations

None.

9. Contributors

Chris Morrow, morrowc@google.com

Pekka Savola, pekkas@netcore.fi

Remi Despres, remi.despres@free.fr

Seiichi Kawamura, karamucho@mesh.ad.jp

10. Acknowledgments

We'd like to thank Ron Bonica, Pramod Srinivasan, Olivier Vautrin,

Tomoya Yoshida, Warren Kumari and Tatsuya Jinmei for their helpful inputs.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

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

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

11.2. Informative References

[RFC2526] Johnson, J. and S. Deering, "Reserved IPv6 Subnet Anycast Addresses", RFC 2526, March 1999.

[RFC3021] Retana, A., White, R., and V. Fuller, "Using 31-Bit Prefixes on IPv4 Point-to-Point Links", December 2000.

[RFC3627] Savola, P., "Use of /127 Prefix Length Between Routers Considered Harmful", RFC 3627, September 2003.

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

[RFC4443] Conta, A., Deering, S., and M. Gupta, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 4443, March 2006.

[RFC5375] Van de Velde, G., Popoviciu, C., Chown, T., Bonness, O., and C. Hahn, "IPv6 Unicast Address Assignment Considerations", RFC 5375, December 2008.

Authors' Addresses

Miya Kohno
Juniper Networks, Keio University
Shinjuku Park Tower, 3-7-1 Nishishinjuku
Shinjuku-ku, Tokyo 163-1035
Japan

Email: mkohno@juniper.net

Becca Nitzan
Juniper Networks
1194 North Marhilda Avenue
Sunnyvale, CA 94089
USA

Email: nitzan@juniper.net

Randy Bush
Internet Initiative Japan
5147 Crystal Springs
Bainbridge Island, WA 98110
USA

Email: randy@psg.com

Yoshinobu Matsuzaki
Internet Initiative Japan
Jinbocho Mitsui Building,
1-105 Kanda Jinbo-cho, Tokyo 101-0051
Japan

Email: maz@iij.ad.jp

Lorenzo Colitti
Google
1600 Amphitheatre Parkway,
Mountain View, CA 94043
USA

Email: lorenzo@google.com

Thomas Narten
IBM Corporation
3039 Cornwallis Ave.
PO Box 12195 - BRQA/502 Research Triangle Park, NC 27709-2195
USA

Email: narten@us.ibm.com