TOC

v6man	M. Kohno
Internet-Draft	Juniper Networks, Keio University
Obsoletes: 3627 (if approved)	B. Nitzan
Updates: 4291, 5375	Juniper Networks
(if approved)	R. Bush
Intended status: Standards Track	Y. Matsuzaki
Expires: September 5, 2010	Internet Initiative Japan
	L. Colitti
	Google
	March 04, 2010

Using 127-bit IPv6 Prefixes on Inter-Router Links
draft-kohno-ipv6-prefixlen-p2p-01.txt

Abstract

In many environments it is useful, for security and other reasons, to use 127-bit IPv6 prefixes on inter-router links. This document outlines some of these reasons and proposes that 127-bit IPv6 prefix lengths be allowed on these links.

Status of this Memo

Copyright Notice

Table of Contents

1. Conventions Used In This Document
2. Introduction
3. Scope Of This Memo
4. Problems identified with 127-bit prefix lengths in the past
5. Reasons for using longer prefixes
    5.1. Ping-pong issue
    5.2. Neighbor Cache Exhaustion issue
    5.3. Other reasons
6. Recommendations
7. Security Considerations
8. IANA Considerations
9. Contributors
10. Acknowledgments
11. References
    11.1. Normative References
    11.2. Informative References
§ Authors' Addresses

---

1. Conventions Used In This Document

TOC

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 (Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” March 1997.) [RFC2119].

TOC

2. Introduction

RFC 4291 (Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture," February 2006.) [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] (Savola, P., "Use of /127 Prefix Length Between Routers Considered Harmful," September 2003.). 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 (Reasons for using longer prefixes)).

This document provides rationale for using 127-bit prefix lengths, reevaluates the reasons why doing so was considered harmful, and proposes that /127 prefixes may be used on inter-router links if certain conditions are met.

3. Scope Of This Memo

TOC

This document is applicable to cases where operators assign specific addresses on inter-router 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 link is a link to which only routers and no hosts are attached. Thus, 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

TOC

[RFC3627] (Savola, P., "Use of /127 Prefix Length Between Routers Considered Harmful," September 2003.) discourages the use of 127-bit prefix lengths due to conflicts with the Subnet-Router anycast addresses. However, the RFC also states that the utility of Subnet-Router Anycast for point-to-point links is questionable.

[RFC5375] (Van de Velde, G., Popoviciu, C., Chown, T., Bonness, O., and C. Hahn, "IPv6 Unicast Address Assignment Considerations," December 2008.) 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] (Savola, P., “Use of /127 Prefix Length Between Routers Considered Harmful,” September 2003.).
Given that Subnet-Router Anycast is not currently widely implemented, an alternative solution to this problem could have been to recommend that Subnet-Router Anycast be disabled on prefixes that are 127 bits long.

5. Reasons for using longer prefixes

TOC

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

5.1. Ping-pong issue

TOC

As described in [PINGPONG] (Hagino, H. and T. Jimmei, “Avoiding ping-pong packets on point-to-point links,” .), 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, such as SONET, do not use Neighbor Discovery. As a consequence, configuring any prefix length other than 127 bits on these links creates an attack vector in the network.
The latest ICMPv6 specification [RFC4443] (Conta, A., Deering, S., and M. Gupta, “Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification,” March 2006.) provides a solution to this issue by specifying that a router receiving a packet on a point-to-point link sent to an address on the link itself must not forward the packet back onto the same link, instead generating an ICMPv6 Destination Unreachable (Code 3) message. However, checking all traffic for this condition is likely to affect performance, as it doubles the number of routing lookups required to forward every packet.

5.2. Neighbor Cache Exhaustion issue

TOC

As described in Section 4.3.2 of [RFC3756] (Nikander, P., Kempf, J., and E. Nordmark, “IPv6 Neighbor Discovery (ND) Trust Models and Threats,” May 2004.), 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] (Narten, T., Nordmark, E., Simpson, W., and H. Soliman, “Neighbor Discovery for IP version 6 (IPv6),” September 2007.).

By sending a continuous stream of packets to a large number of the 2^64 - 3 addresses on the link that are not assigned to the routers (one for each router and one for Subnet-Router Anycast), an attacker can cause the routers to create a large number of neighbor cache entries and send a large number of Neighbor Solicitation packets which will never receive replies, possibly consuming a disproportionate amount of memory and processing resources. Sending the packets to one of the 2^24 addresses on the link that have the same Solicited-Node multicast address as of one of the routers also causes the other router to spend disproportionate 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] (Narten, T., Nordmark, E., Simpson, W., and H. Soliman, “Neighbor Discovery for IP version 6 (IPv6),” September 2007.) 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.

5.3. Other reasons

TOC

1. With the use of 127-bit or other long prefix lengths, interface IDs are simpler and easier to remember (e.g., the Interface ID is 0 or 1).

2. Using 64-bit prefixes for inter-router links leaves a large number of unused addresses that an attacker with physical access to a link could use to insert a node onto the link without having to compromise the routing protocols used on the link. If 127-bit prefix lengths are in use, this is not possible.

3. Though address space conservation considerations are less important for IPv6 than they are in IPv4, it may still be desirable to use the smallest possible prefix to number links (and thus use, for example, /127 for point-to-point links). For example, a large end-site that is assigned a /48 of IPv6 space may not want to reserve a full /64 for every point-to-point link to avoid renumbering in the future.

6. Recommendations

TOC

In light of the above reasons, this document proposes that inter-router links MAY be assigned 127-bit prefix lengths. If such a prefix is assigned to a link, Subnet-Router Anycast MUST be disabled for the prefix.

7. Security Considerations

TOC

This draft addresses, among other things, various security issues.

8. IANA Considerations

TOC

None.

9. Contributors

TOC

Chris Morrow, morrowc@google.com
Pekka Savola, pekkas@netcore.fi
Remi Despres, remi.despres@free.fr

10. Acknowledgments

TOC

We'd like to thank Ron Bonica, Pramod Srinivasan, Olivier Vautrin, Tomoya Yoshida and Warren Kumari for their helpful inputs.

11. References

TOC

11.1. Normative References

TOC

[RFC2119]	Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels,” BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).
[RFC4291]	Hinden, R. and S. Deering, “IP Version 6 Addressing Architecture,” RFC 4291, February 2006 (TXT).
[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 (TXT).
[RFC4861]	Narten, T., Nordmark, E., Simpson, W., and H. Soliman, “Neighbor Discovery for IP version 6 (IPv6),” RFC 4861, September 2007 (TXT).

11.2. Informative References

TOC

[PINGPONG]	Hagino, H. and T. Jimmei, “Avoiding ping-pong packets on point-to-point links.”
[RFC3627]	Savola, P., “Use of /127 Prefix Length Between Routers Considered Harmful,” RFC 3627, September 2003 (TXT).
[RFC3756]	Nikander, P., Kempf, J., and E. Nordmark, “IPv6 Neighbor Discovery (ND) Trust Models and Threats,” RFC 3756, May 2004 (TXT).
[RFC5375]	Van de Velde, G., Popoviciu, C., Chown, T., Bonness, O., and C. Hahn, “IPv6 Unicast Address Assignment Considerations,” RFC 5375, December 2008 (TXT).

Authors' Addresses

TOC

	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