Network Working Group
Request for Comments: 3627
Category: Informational

P. Savola
CSC/FUNET

September 2003

Use of /127 Prefix Length Between Routers Considered Harmful

Status of this Memo

Copyright Notice

Abstract

In some cases, the operational decision may be to use IPv6 /127 prefix lengths, especially on point-to-point links between routers. Under certain situations, this may lead to one router claiming both addresses due to subnet-router anycast being implemented. This document discusses the issue and offers a couple of solutions to the problem; nevertheless, /127 should be avoided between two routers.

1. Introduction

[ADDRARCH] defines Subnet-router anycast address: in a subnet prefix of n bits, the last 128-n bits are all zero. It is meant to be in use of any one router in the subnet.

Even though having prefix length longer than /64 is forbidden by [ADDRARCH] section 2.4 for non-000/3 unicast prefixes, using /127 prefix length has gained a lot of operational popularity; it seems like that these prefix lengths are being used heavily in point-to-point links. The operational practise has often been to use the least amount of address space especially in the presence of a large number of point-to-point links; it may be unlikely that all of these links would start to use /64's. Using /127 has also other operational benefits: you always know which address the other end uses, and there is no "ping-pong" [PINGPONG] problem with older ICMP implementations (fixed now in [ICMPv3]).

2. Scope of this Memo

This memo does not advocate the use of long prefixes, but brings up problems for those that do want to use them, for one reason or another.

Detailed discussion on what is the "right" solution is out of the scope; it is not the goal of this memo to try to find the "best" addressing solution for everyone.

3. Problem with /127 and Two Routers

Note that this problem does not exist between a router and a host, assuming the PREFIX::0/127 address is assigned to the router.

Using /127 can be especially harmful on a point-to-point link when Subnet-router anycast address is implemented. Consider the following sequence of events:

1. Router A and Router B are connected by a point-to-point link.

2. Neither has anything configured or set up on this link.

3. 3ffe:ffff::1/127 address is added to Router A; now it performs Duplicate Address Detection (DAD) [NDISC] for 3ffe:ffff::1. Router A also adds the Subnet-router anycast address 3ffe:ffff::0/127. (DAD is not performed for anycast addresses.)

4. Now Router B has been planned and configured to use 3ffe:ffff::0/127 as its unicast IPv6 address, but adding it will fail DAD, and Router B does not have any address.

Similar scenarios also happen during router reboots, crashes and such.

The usability of subnet-router anycast address between two routers on a point-to-point link is very questionable, but it is still a mandated feature of [ADDRARCH]. Workarounds for this are presented in the next section.

As of yet, this kind of unexpected behavior hasn't been seen at large perhaps because the Subnet-router anycast address hasn't been implemented or too widely used.

4. Solutions

1. One could use /64 for subnets, including point-to-point links.

2. One could use only link-local addresses, but that may make network maintenance and debugging impractical at least in bigger networks; for example, "traceroute" can only return a list of nodes on the path, not the links which would have been used.

3. Failing that, /126 does not have this problem, and it can be used safely on a point-to-point link (e.g., using the 2nd and the 3rd address for unicast). This is analogous to using /30 for IPv4. Using two /128 addresses is also one, though often cumbersome, approach. Naturally, not much would be lost if even a shorter prefix was used, e.g., /112 or /120.

   The author feels that if /64 cannot be used, /112, reserving the last 16 bits for node identifiers, has probably the least amount of drawbacks (also see section 3).

4. [ADDRARCH] could be revised to state that Subnet-router anycast address should not be used if the prefix length of the link is not /64 (or even longer than /120). This does not seem like a good approach, as we should avoid making assumptions about prefix lengths in the specifications, to maintain future flexibility. Also, in some cases, it might be usable to have a Subnet-router anycast address in some networks with a longer prefix length.

   A more conservative (implementation) approach would be not using Subnet-router anycast addresses in subnets with a prefix length of /127 if there are only two routers on the link: this can be noticed with [NDISC] 'Router' bit in Neighbor Advertisement messages. However, this seems to overload the functionality of 'R' bit, so it does not look like a good approach in the long run.

5. It's also possible to improve implementations: if /127 is used on a point-to-point link, never claim two addresses. This has the drawback that even if the router using the combined unicast and anycast address is down, the packets to subnet-router anycast address will be lost as the other cannot claim the address. This approach might lead to unpredictability which would be hard to trace when debugging problems. However, this would normally be an issue only when the Subnet-router anycast address is used from outside of the link; usually, this cannot be done reliably as the prefix length or EUI64 u/g bits cannot be known for certain. There are other problems with an address being anycast and unicast

too: use of it as a source address, whether to use unicast or anycast semantics in [NDISC], and others: allowing this behavior would seem to only add a lot of complexity to the implementations.

1) is definitely the best solution, wherever it is possible. 2) may be usable in some scenarios, but in larger networks (where the most often the desire would be to use longer prefix length) it may be deemed very impractical. There are some situations where one of these may not be an option; then an operational work-around for this operational problem, that is 3), appears to be the best course of action. This is because it may be very difficult to know whether all implementations implement some checks, like ones described in 4) or 5).

5. Other Problems with Long Prefixes

These issues are not specific to /127.

One should note that [ADDRARCH] specifies universal/local bits (u/g), which are the 70th and 71st bits in any address from non-000/3 range. When assigning prefixes longer than 64 bits, these should be taken into consideration; in almost every case, u should be 0, as the last 64 bits of a long prefix is very rarely unique. 'G' is still unspecified, but defaults to zero. Thus, all prefixes with u or g=1 should be avoided.

[MIPV6] specifies "Mobile IPv6 Home-Agents" anycast address which is used for Home Agent Discovery. In consequence, 7 last bits of have been reserved in [ANYCAST] of every non-000/3 non-multicast address, similar to [ADDRARCH]. Thus, at least /120 would seem to make sense. However, as the sender must know the destination's prefix length, this "reserved anycast addresses" mechanism is only applicable when the sender knows about the link and expects that there is a service it needs there. In the case of e.g., /126 between routers, the only to node to be found on this link would be the other router, so the mechanism does not seem useful. At least, Mobile IPv6 Home Agent Discovery should not be performed if the prefix length is longer than /120.

6. References

6.1. Normative References

[ADDRARCH] Hinden, R. and S. Deering, "IP Version 6 (IPv6) Addressing Architecture", RFC 3513, April 2003.

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

6.2. Informative References

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

[MIPV6] Johnson, D., Perkins, C., Arkko, J., "Mobility Support in IPv6", Work in Progress.

[ICMPv3] Conta, A., Deering, S., "Internet Control Message Protocol (ICMPv6)", Work in Progress.

[PINGPONG] Hagino, J., Jinmei, T., Zill, B., "Avoiding ping-pong packets on point-to-point links", Work in Progress.

7. Security Considerations

Beyond those already existing in other specifications, solution 4) might lead to denial of service in the case that one router is down: the packet to subnet-router anycast address would be lost.

8. Acknowledgements

Thanks to Robert Elz and many others on the IPv6 Working Group for discussion, and Alain Durand for pointing out [ADDRARCH] requirements for prefix lengths. Charles Perkins pointed out MIPv6 HA requirements. Randy Bush and Ole Troan commented on the document extensively, and Erik Nordmark pointed out issues with u-bit.

9. Author's Address

Pekka Savola
CSC/FUNET
Espoo, Finland

EMail: psavola@funet.fi

10. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.