v6ops
Internet-Draft
Intended status: Informational
Expires: May 7, 2020

J. Palet Martinez
The IPv6 Company
November 4, 2019

IPv6 Point-to-Point Links
draft-palet-v6ops-p2p-links-04

Abstract

This document describes different alternatives for configuring IPv6 point-to-point links, considering the prefix size, numbering choices and prefix pool to be used.

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 https://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 May 7, 2020.

Table of Contents

1. Introduction

There are different alternatives for numbering IPv6 point-to-point links, and from an operational perspective, there may have different advantages or disadvantages that need to be taken in consideration under the scope of each specific network architecture design.

[RFC6164] describes using /127 prefixes for inter-router point-to-point links, using two different address pools, one for numbering the point-to-point links and another one for delegating the prefixes at the end of the point-to-point link. However, this doesn't exclude other choices.

This document describes alternative approaches, for the prefix size, the numbering of the link and the prefix pool.

The proposed approaches are suitable for those point-to-point links connecting ISP to customers, but not limited to those cases, and in fact, all them are being used by a relevant number of networks worldwide, in several different scenarios (service providers,

enterprise networks, etc.).

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. The Ping-Pong Problem in Point-to-Point Links

Some point-to-point links may present the ping-pong problem, (a forwarding loop). The fundamental root cause of this problem is an IPv6 implementations not performing full Neighbor Discovery (NS/NA) on addresses that the prefix says could exist on the link.

IPv6 implementations are assuming that all addresses within the prefix must exist at the other end of the point-to-point link, and send the traffic straight onto the link. If the address doesn't exist, and there is a covering route back in the other direction, the ping-pong problem occurs.

Full Neighbor Discovery is doing more than just resolving the link-layer address of an IPv6 address. Neighbor Discovery is also determining if the address exists. Even if a point-to-point link doesn't have link-layer addresses to resolve, ND determining if an address exists on the link is very beneficial because it will prevent the ping-pong problem occurring entirely regardless of the IPv6 prefix length being used on the link.

4. Prefix Size Choices

[RFC7608] already discusses about the IPv6 prefix length recommendations for forwarding, and the need for routing and forwarding implementations to ensure that longest-prefix-match works on any prefix length. So, in this document, we concentrate in the most commonly used choices, not excluding other options.

4.1. Rationale for using /64

The IPv6 Addressing Architecture ([RFC4291]) specifies that all the Interface Identifiers for all the unicast addresses (except for 000/3) are required to be 64 bits long and to be constructed in Modified EUI-64 format.

The same document also mandates the usage of the predefined subnet-router anycast address, which has cleared to zero all the bits that

do not form the subnet prefix.

Using /64 is the most common scenario and currently the best practice by the number of service providers using this approach compared to others.

Using a /64 has the advantage of being future proof and avoids renumbering, in the event that new standards take advantage of the 64 bits for other purposes, or the link becomes a point-to-multipoint, or there is a need to use more addresses in the link (e.g., monitoring equipment, managed bridges).

It has been raised also the issue of some hardware having limitations in using prefixes longer then /64, for example using extra hardware resources.

Section 5. of [RFC6164] describes possible issues when using /64 for the point-to-point links, such as the ping-pong and the neighbor cache exhaustion. However, it also states that they can be mitigated by other means, including the latest ICMPv6 [RFC4443] ND [RFC4861]. Indeed, considering the publication date of that document, those issues should not be any longer a concern. The fact is that many operators worldwide, today use /64 without any concerns, as vendors have taken the necessary code updates.

Consequently, we shall conclude that it is a valid approach to use /64 prefixes for the point-to-point links.

4.2. Rationale for using /127

[RFC6164] already do a complete review of reasons why /127 is a good approach vs other options. However, it needs to be considered that it was published a number of years ago, and most of the hardware today already incorporate mitigations.

It should be noted that, when using a /127 prefix, configuration of each of the addresses within the /127 prefix, at each respective end of the link, must be actively validated by the network operator. A missing /127 address from one end of the link, with a local route pointing out that end of the link that covers the missing /127 address, such as a default route, causes a "ping-pong" scenario to exist for the missing /127 address. The link could still be successfully carrying transit traffic, and IPv6 will not report any errors, because IPv6 doesn't require or nor check to ensure all interfaces attached to a link has addresses from all prefixes assigned to the link, excepting the Link-Local prefix per [RFC4291].

It is a valid approach to use /127 for the point-to-point links,

however is not future proof considering the comments from the previous section, and older equipment may not support it.

4.3. Rationale for using /126 and Other Options

/126 was considered by [RFC3627], and despite this document has been obsoleted, because was considering /127 as harmful, the considerations in Section 4.3 are still valid.

The same document describes options such as /112 and /120, and all those are commonly used in worldwide IPv6 deployments [IPv6-Survey], though in a lesser degree than /64 or /127.

Consequently, we shall conclude that /126, /120 and /112 are valid approaches for the point-to-point links.

4.4. A Possible Middle-Term Choice

A possible "middle-term" approach, will be to allocate a /64 for each point-to-point link, but use just one /127 out of it, making it future proof and at the same time avoiding possible issues indicated in the previous sections.

5. Numbering Choices

IPv6 provides different unicast addressing scopes which can be considered when numbering a point-to-point link.

It has been reported that certain hardware may consume resources when using numbered links. This is a very specific situation that may need to be consider on a case by case basis.

5.1. GUA (Global Unicast Addresses)

Using GUA is the most common approach. It provides full functionality for both end-points of the point-to-point link and consequently, facilitates troubleshooting.

5.2. ULA (Unique Local Addresses)

Some networks use ULAs for numbering the point-to-point links. This approach may cause numerous problems when carrying Internet traffic and therefore, is strongly discouraged. For example, if the CE needs to send an ICMPv6 message to a host outside that network (to the Internet), the packet with ULA source address will not get thru and PMTUD will break, which in turn will completely break that IPv6 connection when the MTU is not the same for all the path.

ULAs are IPv6 private addresses, not intended to be used as source or destination addresses across the Internet. This issue also exists in IPv4 when using [RFC1918] addresses on links carrying IPv4 Internet traffic. [RFC6752] discusses this issue for IPv4, with much of the discussion applying similarly to IPv6 and ULAs.

However, this approach is valid if, following Section 2.2 of [RFC4443], and despite using ULA for the point-to-point link, the router is configured with at least one GUA and the source of the ICMPv6 messages are always a GUA, per the IPv6 Default Address Selection algorithm [RFC6724].

5.3. Link-Local Addresses Only

Some networks leave the point-to-point links with only Link-Local addresses used at both ends of the link. This is sometimes improperly referred as "unnumbered", because the Link-Local addresses are also "numbers". Furthermore, [RFC4291] requires that all interfaces attached to a link have at least a Link-Local "number" or address from the Link-Local prefix.

[RFC7404] (Using Only Link-Local Addressing inside an IPv6 Network) discusses pros and cons of this alternative, which in general apply for the point-to-point links.

While this choice might work if the point-to-point link is terminated in a router, which typically will get configured with a suitable routable GUA or ULA, it will not work for devices that can't be further configured, for example if they do not support DHCPv6-PD. This is the case for hosts, when the Operating System is not expected to be a DHCPv6-PD client and are therefore left without any usable GUA to allow traffic forwarding.

In the case of a router, the route for the assigned prefix is pointed towards the link-local address on the router WAN port and the default route on the router is pointed towards the link-local address on the upstream network equipment port.

This choice seems easier to implement, compared the previous ones, but it also brings some drawbacks, such as difficulties with troubleshooting and monitoring. For example, link local addresses do not appear in traceroute, so it makes more difficult to locate the exact point of failure.

It is more useful in scenarios where it is known that only a router will be attached to the point-to-point link, and where the configured address of the router is known. Non-routers connecting to a network, which can't initiate DHCPv6-PD might experience problems and will

stay unnumbered upon connection, if a /64 prefix is not used to number the link. This may be also the case for routers, which will not be able to complete the DHCPv6-PD in unnumbered links.

The considerations indicated in the previous section, regarding not using ULA as source address of ICMPv6 messages, and instead ensuring there is at least one GUA configured for that, also apply if link-locals are used for the point-to-point link.

6. Prefix Pool Choices

The logic choice seems to use a dedicated pool of IPv6 addresses, as this is the way we are "used to" with IPv4. Actually, this is done often by means of different IPv6 pools at every PoP in a service provider network.

A possible benefit of using a dedicated IPv6 pool, is that allows applying security policies without harming the customers. This is only true if customers always have a CE at their end of the WAN link.

However, the fact that the default IPv6 link size is /64 and commonly multiple /64's are assigned to a single customer, provides an interesting alternative approach for combining "best practices" described in the precedent sections.

The following section depicts this alternative.

7. /64 from Customer Prefix for point-to-point links

Using a /64 from the customer prefix, in addition to the advantages already indicated when using /64, simplifies the addressing plan.

The use of /64 also facilitates an easier way for routing the shorter aggregated prefix into the point-to-point link. Consequently it simplifies the "view" of a more unified addressing plan, providing an easier path for following up any issue when operating IPv6 networks and typically, will have a great impact in saving expensive hardware resources (lower usage of TCAM, typically by half).

This mechanism would not work in broadcast layer two media that rely on ND, because it will try ND for all the addresses within the shorter prefix that is being routed thru the point-to-point link.

7.1. Numbering Interfaces

Often, in point-to-point links, hardware tokens are not available, or there is the need to keep certain bits (u, g) cleared, so the links can be manually numbered sequentially with most of the bits cleared

to zero. This numbering makes as well easier to remember the interfaces, which typically will become numbered as 0 (with 63 leading zero bits) for the provider side and 1 (with 63 leading zero bits) for the customer side.

Using interface identifiers as 0 and 1 is not only a very simple approach, but also a very common practice. Other different choices can as well be used as required in each case.

On the other hand, using the EUI-64, makes it more difficult to remember and handle the interfaces, but provides an additional degree of protection against port (actually address) scanning as described at [RFC7707].

7.2. Routing Aggregation of the Point-to-Point Links

Following this approach and assuming that a shorter prefix is typically delegated to a customer, for example a /48, it is possible to simplify the routing aggregation of the point-to-point links. Towards this, the point-to-point link may be numbered using the first /64 of the /48 delegated to the customer.

Let's see a practical example:

- A service provider uses the prefix 2001:db8::/32 and is using 2001:db8:aaaa::/48 for a given customer.

- Instead of allocating the point-to-point link from a different addressing pool, it may use 2001:db8:aaaa::/64 (which is the first /64 subnet from the 2001:db8:aaaa::/48) to number the link.

- This means that, in the case the non-EUI-64 approach is used, the point-to-point link may be numbered as 2001:db8:aaaa::1/64 for the provider side and 2001:db8:aaaa::2/64 for the customer side.

- Note that using the first /64 and interface identifiers 1 and 2 is a very common practice. However other values may be chosen according to each case specific needs.

In this way, as the same address pool is being used for both, the prefix and the point-to-point link, one of the advantages of this approach is to make very easy the recognition of the point-to-point link that belongs to a given customer prefix, or in the other way around, the recognition of the prefix that is linked by a given point-to-point link.

For example, making a trace-route to debug any issue to a given address in the provider network, will show a straight view, and it

becomes unnecessary one extra step to check a database that correlate an address pool for the point-to-point links and the customer prefixes, as all they are the same.

Moreover, it is possible to use the shorter prefix as the provider side numbering for the point-to-point link and keep the /64 for the customer side. In our example, it will become:

- Point-to-point link at provider side: 2001:db8:aaaa::1/48

- Point-to-point link at customer side: 2001:db8:aaaa::2/64

This provides one additional advantage as in some platforms the configuration may be easier saving one step for the route of the delegated prefix (no need for two routes to be configured, one for the delegated prefix, one for the point-to-point link). It is possible because the longest-prefix-match rule.

The behavior of this type of configuration has been successfully deployed in different operator and enterprise networks, using commonly available implementations with different routing protocols, including RIP, BGP, IS-IS, OSPF, along static routing, and no failures or interoperability issues have been reported.

7.3. DHCPv6 Considerations

As stated in [RFC3633], "the requesting router MUST NOT assign any delegated prefixes or subnets from the delegated prefix(es) to the link through which is received the DHCP message from the delegating router", however the approach described in this document is still useful in other DHCPv6 scenarios or non-DHCPv6 scenarios.

Furthermore, [RFC3633] was updated by Prefix Exclude Option for DHCPv6-based Prefix Delegation ([RFC6603]), precisely to define a new DHCPv6 option, which covers the case described by this document.

Moreover, [RFC3769] has no explicit requirement that avoids the approach described in this document.

7.4. Router Considerations

This approach is being used by operators in both, residential/SOHO and enterprise networks, so the routers at the customer end for those networks MUST support [RFC6603] if DHCPv6-PD is used.

In the case of Customer Edge Routers there is a specific requirement ([RFC7084]) WPD-8 (Prefix delegation Requirements), marked as SHOULD for [RFC6603]. However, in a scenario where the approach described

in this document is followed, together with DHCPv6-PD, the CE Router MUST support [RFC6603].

8. Security Considerations

This document does not have any new specific security considerations.

9. IANA Considerations

This document does not have any new specific IANA considerations.

10. Acknowledgements

The author would like to acknowledge the inputs of Mikael Abrahamsson, Brian Carpenter, Eric Vyncke, Mark Smith and TBD.

Acknowledge is also due to my co-authors of RIPE-690 (Best Current Operational Practice for Operators: IPv6 prefix assignment for end-users - persistent vs non-persistent, and what size to choose, https://www.ripe.net/publications/docs/ripe-690) and global community, which provided valuable inputs which have been key for this document.

Acknowledgement to co-authors, Cesar Olvera and Miguel Angel Diaz, of a previous related document (draft-palet-v6ops-point2point, 2006), as well as inputs for that version from Alain Durand, Chip Popoviciu, Daniel Roesen, Fred Baker, Gert Doering, Olaf Bonness, Ole Troan, Pekka Savola and Vincent Jardin, are also granted.

11. References

11.1. 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, <https://www.rfc-editor.org/info/rfc2119>.

[RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, DOI 10.17487/RFC3633, December 2003, <https://www.rfc-editor.org/info/rfc3633>.

[RFC3769] Miyakawa, S. and R. Droms, "Requirements for IPv6 Prefix Delegation", RFC 3769, DOI 10.17487/RFC3769, June 2004, <https://www.rfc-editor.org/info/rfc3769>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89, RFC 4443, DOI 10.17487/RFC4443, March 2006, <https://www.rfc-editor.org/info/rfc4443>.

[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, <https://www.rfc-editor.org/info/rfc4861>.

[RFC6603] Korhonen, J., Ed., Savolainen, T., Krishnan, S., and O. Troan, "Prefix Exclude Option for DHCPv6-based Prefix Delegation", RFC 6603, DOI 10.17487/RFC6603, May 2012, <https://www.rfc-editor.org/info/rfc6603>.

[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, <https://www.rfc-editor.org/info/rfc6724>.

[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, <https://www.rfc-editor.org/info/rfc7084>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

11.2. Informative References

[IPv6-Survey] Palet Martinez, J., "IPv6 Deployment Survey (Residential/ Household Services)", January 2018, <https://indico.uknof.org.uk/event/41/contribution/5/material/slides/0.pdf>.

[RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, DOI 10.17487/RFC1918, February 1996, <https://www.rfc-editor.org/info/rfc1918>.

[RFC3627] Savola, P., "Use of /127 Prefix Length Between Routers Considered Harmful", RFC 3627, DOI 10.17487/RFC3627, September 2003, <https://www.rfc-editor.org/info/rfc3627>.

[RFC6164] Kohno, M., Nitzan, B., Bush, R., Matsuzaki, Y., Colitti, L., and T. Narten, "Using 127-Bit IPv6 Prefixes on Inter-Router Links", RFC 6164, DOI 10.17487/RFC6164, April 2011, <https://www.rfc-editor.org/info/rfc6164>.

[RFC6752] Kirkham, A., "Issues with Private IP Addressing in the Internet", RFC 6752, DOI 10.17487/RFC6752, September 2012, <https://www.rfc-editor.org/info/rfc6752>.

[RFC7404] Behringer, M. and E. Vyncke, "Using Only Link-Local Addressing inside an IPv6 Network", RFC 7404, DOI 10.17487/RFC7404, November 2014, <https://www.rfc-editor.org/info/rfc7404>.

[RFC7608] Boucadair, M., Petrescu, A., and F. Baker, "IPv6 Prefix Length Recommendation for Forwarding", BCP 198, RFC 7608, DOI 10.17487/RFC7608, July 2015, <https://www.rfc-editor.org/info/rfc7608>.

[RFC7707] Gont, F. and T. Chown, "Network Reconnaissance in IPv6 Networks", RFC 7707, DOI 10.17487/RFC7707, March 2016, <https://www.rfc-editor.org/info/rfc7707>.

Author's Address

Jordi Palet Martinez
The IPv6 Company
Molino de la Navata, 75
La Navata - Galapagar, Madrid 28420
Spain

EMail: jordi.palet@theipv6company.com
URI: http://www.theipv6company.com/