v6ops
Internet-Draft
Intended status: Informational
Expires: March 10, 2021

J. Linkova
Google

September 6, 2020

Neighbor Cache Entries on First-Hop Routers: Operational Considerations

draft-ietf-v6ops-nd-cache-init-05

Abstract

Neighbor Discovery (RFC4861) is used by IPv6 nodes to determine the link-layer addresses of neighboring nodes as well as to discover and maintain reachability information. This document discusses how the neighbor discovery state machine on a first-hop router is causing user-visible connectivity issues when a new (not being seen on the network before) IPv6 address is being used. The various approaches to mitigate the problem are described, with the proposed solution fully documented in I-D.ietf-6man-grand.

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 March 10, 2021.

Copyright Notice

Table of Contents

1. Introduction

The section 7.2.5 of [RFC4861] states: "When a valid Neighbor Advertisement is received (either solicited or unsolicited), the Neighbor Cache is searched for the target's entry. If no entry exists, the advertisement SHOULD be silently discarded. There is no need to create an entry if none exists, since the recipient has apparently not initiated any communication with the target."

This approach is perfectly suitable for host-to-host communications, which are in most cases bi-directional, and it could be expected that if a host A has an neighbor cache entry for the host B IPv6 address, host B also has the corresponding entry for the host A address in its cache. However when a host communicates to off-link destinations via its first-hop router, that logic does not apply. The most typical scenario when the problem may arise is a host joining the network, forming a new address and using that address for accessing the Internet:

1. A host joins the network and receives a Router Advertisement (RA) packet from the first-hop router (either a periodic unsolicited RA or a response to a Router Solicitation sent by the host). The RA contains information the host needs to perform Stateless Address Autoconfiguration ([RFC4862]) and to configure its network stack. As in most cases the RA also contains the link-layer address of the router, the host can populate its Neighbor Cache with the router's link-local and link-layer addresses.

2. The host starts opening connections to off-link destinations. A very common use case is a mobile device sending probes to detect the Internet connectivity and/or the presence of a captive portal on the network. To speed up that process many implementations use Optimistic Duplicate Address Detection [RFC4429] which allows them to send probes before the Duplicate Address Detection (DAD) process is completed. At that moment the device neighbor cache contains all information required to send those probes (such as the default router link-local the link-layer addresses). The router neighbor cache, however, might contain an entry for the device link-local address (if the device has been performing the address resolution for the router link-local address), but there are no entries for the device global addresses.

3. Return traffic is received by the first-hop router. As the router does not have any cache entry for the host global address yet, the router starts the neighbor discovery process by creating an INCOMPLETE cache entry and then sending a Neighbor Solictation to the Solicited Node Multicast Address. Most router implementations buffer only one data packet while resolving the packet destination address, so it would drop all subsequent packets for the host global address, until the address resolution process is completed.

4. If the host sends multiple probes in parallel, it would consider all but one of them failed. That leads to user-visible delay in connecting to the network, especially if the host implements some form of backoff mechanism and does not retransmit the probes as soon as possible.

This scenario illustrates the problem occurring when the device connects to the network for the first time or after a timeout long enough for the device address to be removed from the router's neighbor cache. However, the same sequence of events happen when the host starts using a new global address previously unseen by the router, such as a new privacy address [RFC4941] or if the router's Neighbor Cache has been flushed.

While in dual-stack networks this problem might be hidden by Happy Eyeballs [RFC8305] it manifests quite clearly in IPv6-only environments, especially wireless ones, leading to poor user experience and contributing to a negative perception of IPv6-only solutions as unstable and non-deployable.

This document discusses the operational implications of not proactively creating Neighbor Cache entries on first-hop routers and summarizes various approaches to mitigate the problem. The document provides an overview of the proposed solution which is fully described in [I-D.ietf-6man-grand].

1.1. 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.

1.2. Terminology

ND: Neighbor Discovery, [RFC4861].

SLAAC: IPv6 Stateless Address Autoconfiguration, [RFC4862].

NS: Neighbor Solicitation, [RFC4861].

NA: Neighbor Advertisement, [RFC4861].

RS: Router Solicitation, [RFC4861].

RA: Router Advertisement, [RFC4861].

SLLAO: Source link-layer Address Option, an option in the ND packets containing the link-layer address of the sender of the packet, [RFC4861].

GUA: Global Unicast Address, [RFC4291].

DAD: Duplicate Address Detection, [RFC4862].

Optimistic DAD: a modification of DAD, [RFC4429].

FCFS SAVI: First-Come, First-Served Source Address Validation, [RFC6620].

2. Proposed Solution

2.1. Solution Requirements

It would be highly desirable to improve the Neighbor Discovery mechanics so routers have a usable cache entry for a host address by the time the router receives the first packet for that address. In particular:

- If the router does not have a Neighbor Cache entry for the address, a STALE entry needs to be created.
- The solution needs to work for Optimistic addresses as well. Devices implementing the Optimistic DAD usually attempt to minimize the delay in connecting to the network and therefore are more likely to be affected by the problem described in this document.
- In case of duplicate addresses present in the network, the proposed solution MUST NOT override the existing entry.
- In topologies with multiple first-hop routers the cache needs to be updated on all of them, as traffic might be asymmetric: outgoing flows leaving the network via one router while the return traffic enters the segment via another one.

In addition the solution MUST NOT exacerbate issues described in [RFC6583] and MUST be compatible with the recommendations provided in [RFC6583].

2.2. Solution Overview

The Neighbor Discovery is designed to allow IPv6 nodes to discover neighboring nodes' reachability and learn IPv6 to link-layer addresses mapping. Therefore ND seems to be the most appropriate tool to inform the first-hop routers about addresses the host is going to use.

Section 4.4 of [RFC4861] says:

"A node sends Neighbor Advertisements in response to Neighbor Solicitations and sends unsolicited Neighbor Advertisements in order to (unreliably) propagate new information quickly."

Propagating information about new GUA as quickly as possible is exactly what is required to solve the problem outlined in this document. Therefore the host might send an unsolicited NA with the

target link-layer address option to advertise its GUA as soon as the said address enters Optimistic or Preferred state.

The proposed solution is discussed in [I-D.ietf-6man-grand]. In summary, the following changes to [RFC4861] are suggested:

- A node SHOULD send up to MAX_NEIGHBOR_ADVERTISEMENT unsolicited NA packets with the Override flag cleared to all-routers multicast address (ff02::2) as soon as one of the following events happens:
  - (if Optimistic DAD is used): a new Optimistic address is assigned to the node interface.
  - (if Optimistic DAD is not used): an address changes the state from tentative to preferred.
- Routers SHOULD create a new STALE ND cache entry upon receiving unsolicited NAs.

It should be noted that some routing and switching platforms have implemented such behaviour already. Administrators could enable the creation of neighbor discovery cache entries based on unsolicited NA packets sent from the previously unknown neighbors on that interface.

Network devices implementing FCFS SAVI might drop Neighbor Advertisements received through a Validating Port which is in the TENTATIVE state (see Section 2.3.2 of[RFC6620]). Therefore hosts using Optimistic DAD might not benefit from the proposed solution if FCFS SAVI is implemented on the network infrastructure. [I-D.ietf-6man-grand] discusses in more details how the proposed solution interacts with SAVI.

3. Solutions Considered but Discarded

The problem could be addressed from different angles. Possible approaches are:

- Just do nothing.
- Migrate from the "reactive" Neighbor Discovery ([RFC4861]) to the registration-based mechanisms ([RFC8505]).
- The router creates new entries in its Neighbor Cache by gleaning from Neighbor Discovery DAD messages.
- The host initiates bidirectional communication to the router using the host GUA.

- Making the probing logic on hosts more robust.
- Increasing the buffer size on routers.
- Transit dataplane traffic from an unknown address (an address w/o the corresponding neighbor cache entry) triggers an address resolution process on the router.

It should be noted that some of those options are already implemented by some vendors. The following sections discuss those approaches and the reasons they were discarded.

3.1. Do Nothing

One of the possible approaches might be to declare that everything is working as intended and let the upper-layer protocols to deal with packet loss. The obvious drawbacks include:

- Unhappy users.
- Many support tickets.
- More resistance to deploy IPv6 and IPv6-Only networks.

3.2. Change to the Registration-Based Neighbor Discovery

The most radical approach would be to move away from the reactive ND as defined in [RFC4861] and expand the registration-based ND ([RFC6775], [RFC8505]) used in Low-Power Wireless Personal Area Networks (6LoWPANs) to the rest of IPv6 deployments. This option requires some investigation and discussions and seems to be excessive for the problem described in this document.

3.3. Host Sending NS to the Router Address from Its GUA

The host could force creating a STALE entry for its GUA in the router ND cache by sending the following Neighbor Solicitation message:

- The NS source address is the host GUA.
- The destination address is the default router IPv6 address.
- The Source Link-Layer Address option contains the host link-layer address.
- The target address is the host default router address (the default router address the host received in the RA).

The main disadvantages of this approach are:

- Would not work for Optimistic addresses as section 2.2 of [RFC4429] explicitly prohibits sending Neighbor Solicitations from an Optimistic Address.

- If first-hop redundancy is deployed in the network, the NS would reach the active router only, so all backup routers (or all active routers except one) would not get their neighbor cache updated.

- Some wireless devices are known to alter ND packets and perform various non-obvious forms of ND proxy actions. In some cases, unsolicited NAs might not even reach the routers.

3.4. Host Sending Router Solicitation from its GUA

The host could send a router solicitation message to 'all routers' multicast address, using its GUA as a source. If the host link-layer address is included in the Source Link-Layer Address option, the router would create a STALE entry for the host GUA as per the section 6.2.6 of [RFC4861]. However, this approach can not be used if the GUA is in optimistic state: section 2.2 of [RFC4429] explicitly prohibits using an Optimistic Address as the source address of a Router Solicitation with a SLLAO as it might disrupt the rightful owner of the address in the case of a collision. So for the optimistic addresses the host can send an RS without SLLAO included. In that case the router may respond with either a multicast or a unicast RA (only the latter would create a cache entry).

This approach has the following drawbacks:

- If the address is in the Optimistic state the RS can not contain SLLAO. As a result the router would only create a cache entry if the solicited RAs is sent as as a unicast. Routers sending solicited RAs as multicast would not create a new cache entry as they do not need to send a unicast packet back to the host.

- There might be a random delay between receiving an RS and sending a unicast RA back (and creating a cache entry) which might undermine the idea of creating the cache entry proactively.

- Some wireless devices are known to intercept ND packets and perform various non-obvious forms of ND proxy actions. In some cases the RS might not even reach the routers.

3.5. Routers Populating Their Caches by Gleaning From Neighbor Discovery Packets

Routers may be able to learn about new addresses by gleaning from the DAD Neighbor Solicitation messages. The router could listen to all solicited node multicast address groups and upon receiving a Neighbor Solicitation from the unspecified address search its Neighbor Cache for the solicitation's Target Address. If no entry exists, the router may create an entry, set its reachability state to 'INCOMPLETE' and start the address resolution for that entry.

The same solution was proposed in [I-D.halpern-6man-nd-pre-resolve-addr]. Some routing vendors support such optimization already. However, this approach has a number of drawbacks and therefore should not be used as the only solution:

- Routers need to receive all multicast Neighbor Discovery packets which might negatively impact the routers CPU.

- If the router starts the address resolution as soon as it receives the DAD Neighbor Solicitation the host might be still performing DAD and the target address might be tentative. In that case, the host SHOULD silently ignore the received Neighbor Solicitation from the router as per the Section 5.4.3 of [RFC4862]. As a result the router might not be able to complete the address resolution before the return traffic arrives.

3.6. Initiating Hosts-to-Routers Communication

The host may force the router to start address resolution by sending a data packet such as ping or traceroute to its default router link-local address, using the GUA as a source address. As the RTT to the default router is lower than RTT to any off-link destinations it's quite likely that the router would start the neighbor discovery process for the host GUA before the first packet of the returning traffic arrives.

This approach has the following drawbacks:

- Data packets to the router link-local address could be blocked by security policy or control plane protection mechanism.

- It introduces an additional overhead for routers control plane (in addition to processing ND packets, the data packet needs to be processed as well).

- Unless the data packet is sent to 'all routers' ff02::2 multicast address, if the network provides a first-hop redundancy then only the active router would create a new cache entry.

3.7. Transit Dataplane Traffic From a New Address Triggering Address Resolution

When a router receives a transit packet, it might check the presence of the neighbor cache entry for the packet source address and if the entry does not exist start address resolution process. This approach does ensure that a Neighbor Cache entry is proactively created every time a new, previously unseen GUA is used for sending offlink traffic. However this approach has a number of limitations, in particular:

- If traffic flows are asymmetrical the return traffic might not transit the same router as the original traffic which triggered the address resolution. So the neighbor cache entry is created on the "wrong" router, not the one which actually needs the neighbor cache entry for the host address.

- The functionality needs to be limited to explicitly configured networks/interfaces, as the router needs to distinguish between onlink addresses (ones the router needs to have Neighbor Cache entries for) and the rest of the address space.

- Implementing such functionality is much more complicated than all other solutions as it would involve complex data-control planes interaction.

4. IANA Considerations

This memo asks the IANA for no new parameters.

5. Security Considerations

This memo documents the operational issue and does not introduce any new security considerations. Security considerations of the proposed solution are discussed in the corresponding section of [I-D.ietf-6man-grand].

6. Acknowledgements

Thanks to the following people (in alphabetical order) for their review and feedback: Mikael Abrahamsson, Stewart Bryant, Lorenzo Colitti, Owen DeLong, Igor Gashinsky, Fernando Gont, Tatuya Jinmei, Erik Kline, Warren Kumari, Barry Leiba, Jordi Palet Martinez, Michael

Richardson, Dave Thaler, Pascal Thubert, Loganaden Velvindron, Eric Vyncke.

7. References

7.1. Normative References

[I-D.ietf-6man-grand]
Linkova, J., "Gratuitous Neighbor Discovery: Creating Neighbor Cache Entries on First-Hop Routers", draft-ietf-6man-grand-01 (work in progress), July 2020.

[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>.

[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>.

[RFC4429] Moore, N., "Optimistic Duplicate Address Detection (DAD) for IPv6", RFC 4429, DOI 10.17487/RFC4429, April 2006, <https://www.rfc-editor.org/info/rfc4429>.

[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>.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <https://www.rfc-editor.org/info/rfc4862>.

[RFC6583] Gashinsky, I., Jaeggli, J., and W. Kumari, "Operational Neighbor Discovery Problems", RFC 6583, DOI 10.17487/RFC6583, March 2012, <https://www.rfc-editor.org/info/rfc6583>.

[RFC6620] Nordmark, E., Bagnulo, M., and E. Levy-Abegnoli, "FCFS SAVI: First-Come, First-Served Source Address Validation Improvement for Locally Assigned IPv6 Addresses", RFC 6620, DOI 10.17487/RFC6620, May 2012, <https://www.rfc-editor.org/info/rfc6620>.

[RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <https://www.rfc-editor.org/info/rfc6775>.

[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>.

[RFC8305] Schinazi, D. and T. Pauly, "Happy Eyeballs Version 2: Better Connectivity Using Concurrency", RFC 8305, DOI 10.17487/RFC8305, December 2017, <https://www.rfc-editor.org/info/rfc8305>.

[RFC8505] Thubert, P., Ed., Nordmark, E., Chakrabarti, S., and C. Perkins, "Registration Extensions for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Neighbor Discovery", RFC 8505, DOI 10.17487/RFC8505, November 2018, <https://www.rfc-editor.org/info/rfc8505>.

7.2. Informative References

[I-D.halpern-6man-nd-pre-resolve-addr] Chen, I. and J. Halpern, "Triggering ND Address Resolution on Receiving DAD-NS", draft-halpern-6man-nd-pre-resolve-addr-00 (work in progress), January 2014.

[RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, DOI 10.17487/RFC4941, September 2007, <https://www.rfc-editor.org/info/rfc4941>.

Author's Address

Jen Linkova
Google
1 Darling Island Rd
Pyrmont, NSW 2009
AU

Email: furry@google.com