v6ops J. Linkova
Internet-Draft Google
Intended status: Informational July 5, 2019
Expires: January 6, 2020

Neighbor Cache Entries on First-Hop Routers: Operational Considerations
draft-linkova-v6ops-nd-cache-init-01

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.

Status of This Memo

Copyright Notice

Table of Contents

```
   1.  Introduction  . . . . . . . . . . . . . . . . . . . . . . . .   2
     1.1.  Requirements Language . . . . . . . . . . . . . . . . . .   4
     1.2.  Terminology . . . . . . . . . . . . . . . . . . . . . . .   4
   2.  Potential Solutions . . . . . . . . . . . . . . . . . . . . .   5
     2.1.  Do Nothing  . . . . . . . . . . . . . . . . . . . . . . .   5
       2.1.1.  Pros  . . . . . . . . . . . . . . . . . . . . . . . .   5
       2.1.2.  Cons  . . . . . . . . . . . . . . . . . . . . . . . .   5
     2.2.  Change to the Registration-Based Neighbor Discovery . . .   5
     2.3.  Hosts Explicitly Advertizing Their GUAs Using Existing ND
           Mechanisms  . . . . . . . . . . . . . . . . . . . . . . .   6
       2.3.1.  Host Sending Unsolicited NA . . . . . . . . . . . . .   6
         2.3.1.1.  Pros  . . . . . . . . . . . . . . . . . . . . . .   7
         2.3.1.2.  Cons  . . . . . . . . . . . . . . . . . . . . . .   7
       2.3.2.  Host Sending NS to the Router Address from Its GUA  .   7
         2.3.2.1.  Pros  . . . . . . . . . . . . . . . . . . . . . .   7
         2.3.2.2.  Cons  . . . . . . . . . . . . . . . . . . . . . .   7
       2.3.3.  Host Sending Router Solicitation from its GUA . . . .   8
         2.3.3.1.  Pros  . . . . . . . . . . . . . . . . . . . . . .   8
         2.3.3.2.  Cons  . . . . . . . . . . . . . . . . . . . . . .   8
     2.4.  Initiating Hosts2Routers Communication  . . . . . . . . .   8
       2.4.1.  Pros  . . . . . . . . . . . . . . . . . . . . . . . .   9
       2.4.2.  Cons  . . . . . . . . . . . . . . . . . . . . . . . .   9
     2.5.  Tweaking Probing Algorithms . . . . . . . . . . . . . . .   9
     2.6.  Routers Buffering More Packets  . . . . . . . . . . . . .   9
       2.6.1.  Pros  . . . . . . . . . . . . . . . . . . . . . . . .  10
       2.6.2.  Cons  . . . . . . . . . . . . . . . . . . . . . . . .  10
   3.  Recommendations . . . . . . . . . . . . . . . . . . . . . . .  10
     3.1.  Avoiding Disruption . . . . . . . . . . . . . . . . . . .  10
   4.  IANA Considerations . . . . . . . . . . . . . . . . . . . . .  11
   5.  Security Considerations . . . . . . . . . . . . . . . . . . .  11
   6.  Acknowledgements  . . . . . . . . . . . . . . . . . . . . . .  12
   7.  References  . . . . . . . . . . . . . . . . . . . . . . . . .  12
     7.1.  Normative References  . . . . . . . . . . . . . . . . . .  12
     7.2.  Informative References  . . . . . . . . . . . . . . . . .  13
   Author's Address  . . . . . . . . . . . . . . . . . . . . . . . .  13
```

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 host2host communications which are in most cases bi-directional and it could be expected that if a host A has an ND cache entry for the host B IPv6 address, the host B also has the corresponding ND 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. Here is the most typical scenario when the problem may arise:

1. When a host joins the network it receives an RA packet from the first-hop router (either a periodic unsolicited RA or a response to an RS sent by the host). The RA contains information the host needs to perform SLAAC and to configure its network stack. Among other things the host populates its ND cache with the router link-local address and potentially link-layer address (if included in the RA Source Link-Layer Address option).

2. The host starts opening connections to off-link destinations. Very common use case is a mobile device sending probes to detect the Internet connectivity and/or the captive portals presence on the network. To speed up that process many implementations are using the Optimistic Duplicate Address Detection ([RFC4429]) which allows them to send probes from their GUA before the DAD process is completed. Imprortant point here is that at that moment the device ND cache contains all information required to send those probes (such as the default gateway LLA and the link-layer address). The router ND cache, however, might contain an entry for the device link-local address (if the device has been performing the ND process for the roiter LLA) but there are no entries for the device GUA.

3. Response packets for the probes (or any other traffic sent by the host) are received by the first-hop router. As the router does not have any ND cache entry for the host GUA, the router starts the neighbor discovery process by creating an INCOMPLETE cache entry and then sending an NS to the Solicited Node Multicast Address. Apparently most of the router implementations buffer only one data packet while performing the ND process for its destination. Therefore all packets for the host GUA, except for the very first one are dropped until the address resolution process is completed.

4. As many implementations send multiple probes in parallel it's very likely that all probes ex. the first one would be considered failed. If the host implements an exponential backoff for

probing it leads to user-noticeable delay in detecting network connectivity/reporting the network as usable.

The above-mentioned scenario illustrates the problem happening when the device connects to the network for the first time/after a long timeout. However the same sequence of events happen when the host starts using the new (previously unseen by the router or) GUA (e.g. a new privacy address [RFC4941]) or if the router Neighbor Cache has been flushed.

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

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].

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

TLLA: Target link-layer Address, an option in the ND packets containing the link-layer address of the target ([RFC4861]).

GUA: Global Unicast Address ([RFC4291]).

DAD: Duplicate Address Detection, [RFC4862].

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

2. Potential Solutions

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 host explicitly advertizes its GUAs using Neighbor Discovery mechanisms.

- 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.

The following sections discuss those approaches in more detail.

2.1. Do Nothing

One of the possible approaches might be to declare that everything is working as intended.

2.1.1. Pros

- No work required.

2.1.2. Cons

- Unhappy users.

- Many support tickets.

- More resistance to deploy IPv6 and IPv6-Only networks.

2.2. 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 required some investigation and discussions. More text will be added here in the following revision of the draft.

2.3. Hosts Explicitly Advertizing Their GUAs Using Existing ND Mechanisms

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. The following sections discuss potential apptoaches in more detail.

2.3.1. Host Sending Unsolicited NA

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 to advertize its GUA as soon as the said address enters Optimistic or Preferred state. The NA should include the target link-layer address option. To ensure that all first-hop routers receive the advertisement it could be sent to all-routers multicast address (ff02::2).

As it's been mentioned, [RFC4861] explicitly states that receiving a NA should not create a new NC entry. However the justification for that requirement ("There is no need to create an entry if none exists, since the recipient has apparently not initiated any communication with the target.") clearly does not apply for the case discussed. As per [RFC2119] "there may exist valid reasons in particular circumstances to ignore a particular item, but the full implications must be understood and carefully weighed before choosing a different course.". Therefore routers creating a new NC entry upon receiving an unsolicited NA would still be compliant with [RFC4861].

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

2.3.1.1. Pros

- Already implemented on some platforms.

- In accordance with [RFC4861].

2.3.1.2. Cons

- Allows a malicious host to execute an ND cache exhaustion attack. It's recommended that this functionality is configurable and recommendations from [RFC6583] are taken into account.

- Requires hosts to send unsolicited NA (changes to the hosts).

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

2.3.2. 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 Source Link-Layer Address option contains the host link-layer address.

- The target address is the host default gateway address (the default router address the host received in the RA).

The main disadvantage of this approach is that it would not work if the GUA the host needs to advertise is still in the Optimistic state. The section 2.2 of [RFC4429] explicitly prohibits sending Neighbor Solicitations from an Optimistic Address.

2.3.2.1. Pros

- Router implementations which follow recommendations made in [RFC6583] might prioritize responding to NS packets to own addresses.

2.3.2.2. Cons

- Does not work for Optimistic addresses (see section 2.2 of [RFC4429]).

- 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 ex. one) would not get their neighbor cache updated.

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

2.3.3. 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 (see the section 6.2.6 of [RFC4861]). However this approach can not be used if the GUA is in optimistic state: the 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).

2.3.3.1. Pros

- Unlike NS packets, RS packets would reach all routers on link, allowing all routers to update their neighbor caches and preventing packet loss in case of asymmetric routing.

2.3.3.2. Cons

- As for the Optimistic addresses SLLAO can not be included into RS packets, the cache entry for the optimistic address would be created only if the router sends solicited RAs as unicast. In addition, 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 fiddle with ND packets and perform various non-obvious forms of ND proxy actions. In some cases RSes might not even reach the routers.

2.4. Initiating Hosts2Routers Communication

Every time the host configures a new GUA (when the address enters the Optimistic state or, if the optimistic DAD is not used, as soon as it changes the state from tentative to preferred) the host can a ping or traceroute packet to the default gateway LLA. As the RTT to the

default gateway 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. There are pretty good chances that the process would be completed before the actual data traffic reaches the router.

2.4.1. Pros

- As data packets are involved, there is no potential impact caused by smart wireless infrastructure performing ND proxy.

- Full compliance with existing standards.

2.4.2. Cons

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

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

- If the first hop redundancy is implemented in the network the host ping/traceroute packet would reach the active router only. All backup routers would not receive it and therefore would not start populating the cache. So in the case of asymmetric traffic flow (packets leave the network via one router while the return flow is going via another) the backup router(s) still would not have the cache entry. (A hacky way to overcome this limitation would be sending ping/traceroute packet to 'all routers' ff02::2 multicast address).

2.5. Tweaking Probing Algorithms

While tweaking the probing logic on devices might make the problem less visible it would be still desirable to avoid packet loss everytime the new GUA is used by a host. It would be quite tricky to adjust every probing algorithm to find the right balance between prompt detection of network connectivity and false positives in IPv6-only mode.

2.6. Routers Buffering More Packets

Another way to mitigate the issue, at least partially, would be increasing the number of packets the router could buffer while performing the neighbor discovery process for the INCOMPLETE cache entry. However it would be against recommendations made in the section 7.2.2 of [RFC4861] and [RFC6583].

2.6.1. Pros

- Does not require changes on hosts.

2.6.2. Cons

- This approach makes the routers even more vulnerable to attack vectors described in [RFC6583]. In particular, it would amplify the impact of any scanning attack.

- Against the recommendations from the section 7 of [RFC6583].

- Requires router vendors support.

3. Recommendations

- Hosts SHOULD send at least one unsolicited NA packet to all-routers multicast address (ff02::2) as soon as one of the following events happens:

  * (if Optimistic DAD is used): a new Optimistic GUA is assigned to the host interface.

  * (if Optimistic DAD is not used): a GUA changes the state from tentative to preferred.

- Routers SHOULD have a configuration knob to enable creating ND cache entry upon receiving unsolicited NAs on a specific interface. This document does not change the behavior if the ND cache entry already exists when receiving an unsolicited NA.

3.1. Avoiding Disruption

If hosts following the recommendations in this document are using the DAD mechanism defined in [RFC4862], they would send unsolicited NA as soon as the address changes the state from tentative to preferred (after its uniqueness has been verified). However hosts willing to minimize network stack configuration delays might be using optimistic addresses, which means there is possibility of the address not being unique on the link. The section 2.2 of [RFC4429] discusses measures to ensure that ND packets from the optimistic address do not override any existing neighbor cache entries as it would cause traffic interruption of the rightful address owner in case of address conflict.

As hosts willing to speed up their network stack configuration are most likely to be affected by the problem outlined in this document it seems reasonable for such hosts to advertise their optimistic GUAs

by sending unsolicited NAs. The main question to consider is the potential risk of overriding the cache entry for the rightful address owner if the optimistic address happens to be duplicated.

As per section 7.2.5 of [RFC4861] if the Neighbor Cache entry for the target address already exists and is in in any state other than INCOMPLETE then the only change the unsolicited NA could cause is to change the entry from REACHABLE to STALE. It would not cause any traffic interruption for the rightful address owner.

If there is no entry then it would be created/updated with the supplied LLA and its state set to STALE. In that case as soon as the entry is used for sending traffic to the host, the entry state will be changed to DELAY and the Neighbor Unreachability Detection would be started and the rightful owner LLA will be entered in the cache. So in the scenario when the rightful owner does not use the address for communication then it might be a short (a few seconds) period of time when the data packets sent from the outside could reach the host with the optimistic address. However it seems likely that hosts using Optimistic DAD would start sending/receiving traffic right away, so the first return packet would trigger the NUD process and rewrite the cache.

Another corner case is the INCOMPLETE cache entry for the address. If the host sends an unsolicited NA from the Optimistic address it would update the entry with the host LLA and set the entry to the STALE state. As the INCOMPLETE entry means that the router has started the ND process for the address and the multicast NS has been sent, the rightful owner is expected to reply with solicited NA which would recover the cache entry and set the LLA to the rightful owner's one. The risk here:

o The data packet arrives after the unsolicited NA from the host but before the rightful owner responded with the solicited NA. Those packets would be sent to the host with the optimistic address instead of its rightful owner. However without the unsolicited NA those packets would have been dropped anyway (as the entry was in INCOMPLETE state).

4. IANA Considerations

This memo asks the IANA for no new parameters.

5. Security Considerations

One of the potential attack vector to consider is a cache spoofing when the attacker might try to install a cache entry for the victim's IPv6 address and the attacker's Link-Layer address. However it

should be noted that this document does not propose any changes for the scenario when the ND cache for the given IPv6 address already exists. Therefore it is not possible for the attacker to override any existing cache entry.

A malicious host could attempt to exhaust the neighbor cache on the router by creating a large number of STALE entries. However this attack vector is not new and this document does not increase the risk of such attack: the attacker could do it, for example, by sending a NS or RS packet with SLLAO included. All recommendations from [RFC6583] still apply.

6. Acknowledgements

Thanks to the following people (in alphabetical order) for their review and feedback: Lorenzo Colitti, Tatuya Jinmei, Erik Kline, Warren Kumari, Michael Richardson, Pascal Thubert, Loganaden Velvindron, Eric Vyncke.

7. References

7.1. 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>.

[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>.

[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

[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>.

[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>.

Author's Address

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

Email: furry@google.com