IPv6 Maintenance
Internet-Draft
Updates: 4861 (if approved)
Intended status: Standards Track
Expires: January 26, 2021

J. Linkova
Google

July 25, 2020

Gratuitous Neighbor Discovery: Creating Neighbor Cache Entries on First-Hop Routers

draft-ietf-6man-grand-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 updates RFC4861 to allow routers to proactively create a Neighbor Cache entry when a new IPv6 address is assigned to a node. It also updates RFC4861 and recommends nodes to send unsolicited Neighbor Advertisements upon assigning a new IPv6 address. The proposed change will minimize the delay and packet loss when a node initiate connections to off-link destination from a new IPv6 address.

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 January 26, 2021.

Copyright Notice

Table of Contents

1. Introduction

The Neighbor Discovery state machine defined in [RFC4861] assumes that communications between IPv6 nodes are in most cases bi-directional and if a node A is trying to communicate to its neighbor, neighbor B, the return traffic flows could be expected. So when the node A starts the address resolution process, the target node would also create an entry for A address in its neighbor cache. That entry will be used for sending the return traffic to A.

However when a host sends traffic to off-link destinations a different scenario is observed. After receiving a Router Advertisement the host populates its neighbor cache with the default router IPv6 and link-layer addresses and is able to send traffic to off-link destinations. At the same time the router does not have any cache entries for the host global addresses yet and only starts address resolution upon receiving the first packet of the return traffic flow. While waiting for the resolution to complete routers only keep a very small number of packets in the queue, as recommended in Section 7.2.2 [RFC4861]. All subsequent packets arriving before the resolution process finishes are likely to be dropped. It might cause user-visible packet loss and performance degradation.

The detailed problem statement and the various solution approaches could be found in [I-D.ietf-v6ops-nd-cache-init]. This document summarizes the proposed neighbor discovery updates to address the issue.

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

Node: a device that implements IP, [RFC4861].

Host: any node that is not a router, [RFC4861].

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. Proposed Changes to Neighbor Discovery

The following changes are proposed to minimize the delay in creating new entries in a router neighbor cache

- A node sends unsolicited NAs upon assigning a new IPv6 address to its interface.
- A router creates a new cache entry upon receiving an unsolicited NA from a host.

The following sections discuss these changes in more detail.

2.1. Nodes Sending Gratuitous Neighbor Advertisements

The section 7.2.6 of [RFC4861] discusses using unsolicited Neighbor Advertisement to inform node neighbors of the new link-layer address quickly. The same mechanism could be used to notify the node neighbors about the new network-layer address as well: the node can send gratuitous unsolicited Neighbor Advertisements upon assigning a new IPv6 address to its interface.

To minimize the potential disruption in case of duplicate addresses the node should not set the Override flag for a preferred address and must not set the Override flag if the address is in Optimistic [RFC4429] state.

As the main purpose of sending unsolicited NAs upon configuring a new address is to proactively create a Neighbor Cache entry on the first-hop routers, the gratuitous NAs are sent to all-routers multicast address (ff02::2). Limiting the recipients to routers only would help reduce the multicast noise level. If the link-layer devices are performing MLD snooping [RFC4541] then those unsolicited NAs will be only sent to onlink routers instead of being flooded to all nodes.

It should be noted that the proposed mechanism does not cause any significant increase in the multicast traffic. The additional multicast unsolicited NA would proactively create a STALE cache entry on routers as discussed below. When the router receives the return

traffic flows it does not need to send multicast NSes to the solicited node multicast address but would be sending unicast NSes instead. Therefore total amount of multicast traffic should not increase.

2.2. Routers Creating Cache Entries Upon Receiving Unsolicited Neighbor Advertisements

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".

The reasoning behind dropping unsolicited Neighbor Advertisements ("the recipient has apparently not initiated any communication with the target") is valid for onlink host-to-host communication but, as discussed in [I-D.ietf-v6ops-nd-cache-init] it does not really apply for the scenario when the host is announcing its address to routers. Therefore it would be beneficial to allow routers creating new entries upon receiving an unsolicited Neighbor Advertisement.

This document updates [RFC4861] so that routers create a new Neighbor Cache entry upon receiving an unsolicited Neighbor Advertisement. The proposed changes do not modify routers behaviour specified in [RFC4861] for the scenario when the corresponding Neighbor Cache entry already exists.

3. 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 a 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 addresses 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.

The following sections are discussing the address collision scenario when a host sends an unsolicited NA for an address in the Optimistic state, while another host has the same address assigned already.

3.1. Neighbor Cache Entry Exists in Any State Other That INCOMPLETE

If the router Neighbor Cache entry for the target address already exists in any state other than INCOMPLETE, then as per section 7.2.5 of [RFC4861] an unsolicited NA with the Override flag cleared would change the entry state from REACHABLE to STALE but would not update the entry in any other way. Therefore even if the host sends an unsolicited NA from the its Optimistic address the router cache entry would not be updated with the new Link-Layer address and no impact to the traffic for the rightful address owner is expected.

3.2. Neighbor Cache Entry is in INCOMPLETE state

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 link-layer address 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 with the Override flag set. Upon receiving a solicited NA with the Override flag the cache entry will be updated with the TLLA supplied and (as the NA has the Solicited flag set), the entry state will be set to REACHABLE. It would recover the cache entry and set the link-layer address to the one of the rightful owner. The only potential impact would be for packets arriving to the router 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 those packets would have been dropped anyway as until the solicited NA is received the router can not send the traffic.

3.3. Neighbor Cache Entry Does Not Exist

There are two distinct scenarios which can lead to the situation when the router does not have a NC entry for the IPv6 address:

1. The rightful owner of the address has not been using it for communication.

2. The rightful owner just started sending packets from that address but the router has not received any return traffic yet.

The impact on the rightful owner's traffic flows would be different in those cases.

3.3.1. The Rightful Owner Is Not Sending Packets From The Address

In this scenario the following events are expected to happen:

1. The host configures the address and sets its state to Optimistic.

2. The host sends an unsolicited NA with the Override flag set to zero and starts sending traffic from the Optimistic address.

3. The router creates a STALE entry for the address and the host link-layer address.

4. The host starts DAD and detects the address duplication.

5. The router receives the return traffic for the duplicated address. As the NC entry is STALE it sends traffic using that entry, changes it to DELAY and wait up to DELAY_FIRST_PROBE_TIME ([RFC4861]) seconds.

6. The router changes the NC entry state to PROBE and sends up to MAX_UNICAST_SOLICIT ([RFC4861]) unicast NSes separated by RetransTimer milliseconds ([RFC4861]) to the host link-layer address.

7. As the host has detected the address conflict already it does not respond to the unicast NSes.

8. The router sends a multicast NS to the solicited node multicast address, the rightful owner responds and the router NC entry is updated with the rightful owner link-local address.

The rightful owner is not experiencing any disruption as it does not send/receive any traffic. If after step 7 the router keeps receiving any return traffic for communication initiated at step 2, those packets would be forwarded to the rightful owner. However the same behaviour would be observed if changes proposed in this document are implemented: if the host starts sending packets from its Optimistic address but then changed the address state to Duplicated, almost all return traffic would be forwarded to the rightful owner of the said address. Therefore it's safe to conclude that the proposed changes do not cause any disruption for the rightful owner.

3.3.2. The Rightful Owner Has Started Sending Packets From The Address

In this scenario the following events are happening:

1. The rightful owner starts sending traffic from the address (e.g. the address has just been configured or has not been recently used).

2. The host configures the address and sets its state to Optimistic.

3. The host sends an unsolicited NA with the Override flag set to zero and starts sending traffic from the Optimistic address.

4. The router creates a STALE entry for the address and the host link-layer address.

5. The host starts DAD and detects the address duplication.

6. The router receives the return traffic flows for both the rightful owner of the duplicated address and the new host. As the NC entry is STALE it sends traffic using that entry, changes it to DELAY and wait up to DELAY_FIRST_PROBE_TIME ([RFC4861]) seconds.

7. The router changes the NC entry state to PROBE and sends up to MAX_UNICAST_SOLICIT ([RFC4861]) unicast NSes separated by RetransTimer milliseconds ([RFC4861]) to the host link-layer address.

8. As the host has detected the address conflict already it does not respond to the unicast NSes.

9. The router sends a multicast NS to the solicited node multicast address, the rightful owner responds and the router NC entry is updated with the rightful owner link-local address.

As a result the traffic for the address rightful owner would be sent to the host with the duplicated address instead. The duration of the disruption can be estimated as DELAY_FIRST_PROBE_TIME*1000 + (MAX_UNICAST_SOLICIT - 1)*RetransTimer milliseconds. As per the constants defined in Section 10 of [RFC4861] this interval is equal to 5*1000 + (3 - 1)*1000 = 7000ms or 7 seconds.

However it should be noted that the probability of such scenario is rather low as it would require the following things to happen almost simultaneously (within tens of milliseconds):

- One host starts using a new IPv6 address and sending traffic.

- Another host configures the same IPv6 address in Optimistic mode before the router receives the return traffic for the first host.

4. Modifications to RFC-Mandated Behavior

All normative text in this memo is contained in this section.

4.1. Modification to RFC4861 Neighbor Discovery for IP version 6 (IPv6)

4.1.1. Modification to the section 7.2.5

This document proposes the following changes to the section 7.2.5 of [RFC4861]:

OLD TEXT:

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.

NEW TEXT:

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, hosts SHOULD silently discard the advertisement. There is no need to create an entry if none exists, since the recipient has apparently not initiated any communication with the target. Routers SHOULD create a new entry for the target address with the link-layer address set to the Target link-layer address option (if supplied). The entry its reachability state MUST also be set to STALE. If the received Neighbor Advertisement does not contain the Target link-layer address option the advertisement SHOULD be silently discarded.

4.1.2. Modification to the section 7.2.6

This document proposes the following changes to the section 7.2.6 of [RFC4861]:

OLD TEXT:

Also, a node belonging to an anycast address MAY multicast unsolicited Neighbor Advertisements for the anycast address when the node's link-layer address changes.

NEW TEXT:

Also, a node belonging to an anycast address MAY multicast unsolicited Neighbor Advertisements for the anycast address when the node's link-layer address changes.

A node may also wish to notify its first-hop routers when it configures a new global IPv6 address so the routers can proactively populate their neighbor caches with the corresponding entries. In such cases a node SHOULD send up to MAX_NEIGHBOR_ADVERTISEMENT Neighbor Advertisement messages. If the address is preferred then the Override flag SHOULD NOT be set. If the address is in the Optimistic state then the Override flag MUST NOT be set. The destination address SHOULD be set to the all-routers multicast address. These advertisements MUST be separated by at least RetransTimer seconds. The first advertisement SHOULD be sent as soon as one of the following events happens:

- if Optimistic DAD [RFC4429] 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.

---

5. IANA Considerations

This memo asks the IANA for no new parameters.

6. Security Considerations

One of the potential attack vectors 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 an attack: the attacker could do it, for example, by sending a NS or RS packet with SLLAO included. All recommendations from [RFC6583] still apply.

Announcing a new address to all-routers multicast address may inform an on-link attacker about IPv6 addresses assigned to the host. However hiding information about the specific IPv6 address should not be considered a security measure as such information is usually disclosed via DAD to all nodes anyway. Network administrators can also mitigate this issue by enabling MLD snooping on the link-layer devices to prevent IPv6 link-local multicast packets being flooded to all onlink nodes. If peer-to-peer onlink communications are not desirable for the given network segment they should be prevented by proper layer2 security mechanisms. Therefore the risk of allowing hosts to send unsolicited Neighbor Advertisements to all-routers multicast address is low.

It should be noted that the proposed mechanism allows hosts to proactively inform their routers about global IPv6 addresses existing on-link. Routers could use that information to distinguish between used and unused addresses to mitigate ND cache exhaustion DoS attacks described in Section 4.3.2 [RFC3756] and [RFC6583].

7. Acknowledgements

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

8. References

8.1. 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>.

[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>.

8.2. Informative References

[I-D.ietf-v6ops-nd-cache-init] Linkova, J., "Neighbor Cache Entries on First-Hop Routers: Operational Considerations", draft-ietf-v6ops-nd-cache-init-03 (work in progress), July 2020.

[RFC3756] Nikander, P., Ed., Kempf, J., and E. Nordmark, "IPv6 Neighbor Discovery (ND) Trust Models and Threats", RFC 3756, DOI 10.17487/RFC3756, May 2004, <https://www.rfc-editor.org/info/rfc3756>.

[RFC4541] Christensen, M., Kimball, K., and F. Solensky, "Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches", RFC 4541, DOI 10.17487/RFC4541, May 2006, <https://www.rfc-editor.org/info/rfc4541>.

[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