IPv6 Operations (v6ops) Working Group
Internet Draft
Intended status: Informational
Expires: Jan. 2023

X. Xiao
E. Vasilenko
Huawei Technologies
E. Metz
KPN
G. Mishra
Verizon Inc.

July 1, 2022

Selectively Applying Host Isolation to Simplify IPv6 First-hop Deployment
draft-xiao-v6ops-nd-deployment-guidelines-02

Abstract

Neighbor Discovery (ND) is the key protocol of IPv6 first-hop. ND uses multicast extensively and trusts all hosts. In some scenarios like wireless networks, multicast can be inefficient. In other scenarios like public access networks, hosts may not be trustable. Consequently, ND issues may happen in various scenarios. The issues and solutions are documented in more than 30 RFCs. It is difficult to keep track of all these issues and solutions, and how the various solutions fit together. Therefore, deployment guidelines are needed.

This document firstly summarizes the known ND issues and optimization solutions into a one-stop reference. Analyzing these solutions reveals an insight: isolating hosts is effective in solving ND issues. Four isolation methods are proposed and their applicability is discussed. Guidelines are then described for selecting a suitable isolation method based on the deployment scenario.

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 in Dec. 2022.

Table of Contents

1. Introduction.....3
1.1. Terminology.....4
2. Review of ND Issues.....5
2.1. Multicast Causes Performance and Reliability Issues.....5
2.2. Trusting-all-hosts Causes On-link Security Issues.....6
2.3. Router-NCE-on-Demand Causes Performance, NCE Exhaustion and Lack of Subscriber Management Issues.....6
2.4. Summary of ND Issue.....7
3. Review of ND Solutions.....8
3.1. ND Solution in Mobile Broadband IPv6.....8
3.2. ND Solution in Fixed Broadband IPv6.....8
3.3. Unique IPv6 Prefix per Host.....10
3.4. Wireless ND.....10
3.5. Scalable Address Resolution Protocol.....11
3.6. ARP and ND Optimization for Transparent Interconnection of Lots of Links (TRILL):.....11
3.7. Proxy ARP/ND in EVPN.....11
3.8. Gratuitous Neighbor Discovery.....12
3.9. Reducing Router Advertisements.....12
3.10. Source Address Validation Improvement and Router Advertisement Guard.....12
3.11. Dealing with Off-link Attack that May Cause Router NCE Exhaustion.....13
3.12. Enhanced DAD.....13

3.13. ND Mediation for IP Interworking of Layer 2 VPNs.....14
3.14. ND Solutions Defined before the Latest Versions of ND...14
3.14.1. SeND.....14
3.14.2. Cryptographically Generated Addresses (CGA).....14
3.14.3. ND Proxy.....15
3.14.4. Optimistic DAD.....15
3.15. Observations on the Solutions and an Insight Learned....16
4. Isolating Hosts to Simplify First-hop Deployments.....19
4.1. Applicability of P2P Link and Subnet Isolation.....20
4.2. Applicability of P2MP Link and Subnet Isolation.....21
4.3. Applicability of GUA Isolation.....21
4.4. Applicability of Proxy Isolation.....22
4.5. Deployment Guidelines.....22
4.6. Impact of Host Isolation to Other Protocols in IPv6 First-hops.....25
5. Security Considerations.....25
6. IANA Considerations.....25
7. References.....26
7.1. Informative References.....26
8. Acknowledgments.....29

1. Introduction

Neighbor Discovery [ND] is specified in RFC 4861. It defines how hosts and routers in the link interact with each other. ND contains seven main procedures:

1. Hosts generate Link Local Addresses (LLAs) and use multicast Neighbor Solicitations (NSs) for Duplicate Address Detection (DAD).
2. Hosts send multicast Router Solicitations (RSs) to discover first-hop routers. Routers respond with unicast Router Advertisements (RAs) with prefixes and other information. Routers also send unsolicited multicast RAs from time to time.
3. Hosts form Global Unicast Address (GUA) or Unique Local Address (ULA) and use multicast Neighbor Solicitations (NSs) for DAD.
4. When a packet is to be sent, routers use multicast NSs to perform address resolution for the destination host.
5. When a packet is to be sent, hosts use multicast NSs to perform address resolution for the destination host.
6. Hosts/routers use unicast NS for Node Unreachability Detection (NUD).
7. Hosts may use multicast NSs to announce link layer address changes.

Due to multicast, trusting all hosts, etc, ND issues can happen in some scenarios. Various ND issues and solutions have been described in more than 30 RFCs. These include: ND Trust Models and Threats [RFC3756], Secure ND [SeND], Cryptographically Generated Addresses [CGA], ND Proxy [RFC4389], Optimistic ND [RFC4429], ND for mobile broadband [RFC6459][RFC7066], ND for fixed broadband [TR177], ND Mediation [RFC6575], Operational ND Problems [RFC6583], Wireless ND (WiND) [RFC6775][RFC8505][RFC8928][RFC8929], DAD Proxy [RFC6957], Source Address Validation Improvement [SAVI], Router Advertisement Guard [RA-Guard][RA-Guard+], Enhanced Duplicate Address Detection [RFC7527], Scalable ARP [RFC7586], Reducing Router Advertisements [RFC7772], Unique Prefix Per Host [RFC8273], ND Optimization for TRILL [RFC8302], Gratuitous Neighbor Discovery [GRAND], Proxy ARP/ND for EVPN [RFC9161]. It is difficult to understand all these issues and solutions, and how they fit together. Consequently, IPv6 first-hop deployment may become complicated. This document summarizes the issues and solutions to provide a big picture, and provide guidelines for selecting the proper solutions based on the deployment scenarios.

1.1. Terminology

Some important terms are defined in this section.

MAC - To avoid confusion with link local address, link layer address is called MAC in this document.

Link isolation for hosts - isolating hosts in L2. This includes 2 flavors: P2P link isolation and P2MP link isolation.

P2P link isolation - connecting each host in a P2P link to the router. The router has a separate interface for each host. Consequently, any L2 message from a host can only reach the router, not other hosts. Similarly, any L2 message from the router can only reach one host.

P2MP link isolation - connecting multiple hosts in a P2MP link to the router. The router has a single interface for all hosts. Example P2MP links are Private VLAN [PVLAN] and Wi-Fi with Wireless Isolation [W-Iso]. Consequently, any L2 message from a host can only reach the router, not other hosts. But an L2 multicast message from the router can reach multiple hosts simultaneously.

Subnet isolation for hosts - assigning a unique prefix per host so each host is in its own subnet [RFC8273].

GUA/ULA isolation for hosts - setting PIO L-bit=0 so that other hosts appear off-link [ND]. There will be no GUA/ULA resolution for other hosts in the link, and all GUA/ULA traffic will be sent via the router. Therefore hosts appear isolated from a GUA/ULA perspective. To be simple, this is also called GUA Isolation in this document.

Proxy isolation for hosts - using an ND proxy device to represent the hosts behind it, and effectively isolate such hosts from other hosts.

2. Review of ND Issues

2.1. Multicast Causes Performance and Reliability Issues

ND uses multicast extensively for Node Solicitations (NSs), Router Solicitations (RSs) and Router Advertisements (RAs). Multicast can be inefficient in some scenarios, e.g. large L2 networks or wireless networks.

In large L2 networks, e.g. DC networks involving many Virtual Machines (VMs), ND multicast can create a large amount of protocol traffic. This can consume network bandwidth, create a processing burden, and reduce network performance [RFC7342].

In wireless networks, to ensure that the multicast messages reach even the remotest hosts, multicast messages are sent at the lowest modulation rate. This prolongs receiving time and consumes more power of the hosts. Some low-power or remote hosts may not receive or respond to multicast messages. In addition, multicast messages are not acknowledged at L2. Consequently, multicast in wireless networks reduces not only network performance but also protocol reliability [RFC9119]. For example, ND uses no response as an indication of no duplication in Duplicate Address Detection (DAD). If the DAD multicast messages are lost, DAD may fail.

ND uses the following multicast messages. Their impact on performance and reliability is summarized below:

- Hosts' LLA DAD: may cause a performance issue, and a reliability issue in wireless networks.
- Router's periodic unsolicited RAs: may cause performance issue if it is sent frequently [RFC7772].
- Hosts' GUA (or ULA) DAD: may cause a performance issue, and a reliability issue in wireless networks.

. Router's address resolution for hosts: in a large network of N hosts, there can be N such multicast messages. This may cause a performance issue.
. Hosts address resolution for hosts: in a large network of N hosts, there can be N-square such multicast messages. This may cause the largest performance issue.
. Hosts' MAC change NAs: this type of multicast messages is rare and will not cause a performance issue. It will not be further discussed.

2.2. Trusting-all-hosts Causes On-link Security Issues

ND trusts all hosts. In some scenarios like public access networks, some hosts may not be trustable. An attacker host in the link can cause the following security issues [RFC3756][RFC9099]:

. Source IP address spoofing: an attacker can use a victim host's IP address as the source address of its ND message to pretend to be the victim. The attacker can then launch Redirect or Denial of Service (DoS) attacks on the victim.
. DAD denial: an attacker can repeatedly reply to a victim's DAD messages, causing the victim's address configuration procedure to fail. That is a DoS attack.
. Fake RAs: an attacker can send RAs to other hosts to claim to be a router and also preempt the real router. This is a Redirect attack.
. Fake Redirects: an attacker can pretend to be the router and send Redirects to other hosts to redirect their traffic to the router to itself. This is a Redirect attack.
. Replay attacks: an attacker captures valid ND messages and replays them later.

2.3. Router-NCE-on-Demand Causes Performance, NCE Exhaustion and Lack of Subscriber Management Issues

In ND, the router does not maintain (IP, MAC) binding (i.e. NCE) for a host until it is needed. This is called Router-NCE-on-Demand. When the router is to forward a packet to an on-link host, it will use address resolution to find out the MAC of the host. This can cause three issues:

. The packet has to be queued before the router finds out the MAC of the destination host. This reduces forwarding performance and may be an issue in high-performance computing environment, e.g. DCs. This is called "Router-NCE-on-Demand Performance Degradation" in this document.

. The way ND does address resolution is the node will create an NCE entry first and set its state to INCOMPLETE, the node will then multicast NSs to all the hosts and wait for the destination host to reply with its MAC. This creates a security vulnerability. If an attacker sends a large number of packets destined to non-existing IP addresses to the router, the router will create a large amount of NCEs with INCOMPLETE state while trying to resolve the MACs. The router may run out of resources and stop functioning. This is called "Rt NCE Exhaustion" in this document. Note that in this case, the attacker can be off-link. So this is different from the on-link security issues.
. Without an NCE, a router does not know the IP address of a host when SLAAC is used rather than [DHCPv6]. In a service provider network, subscribers are generally managed by their IP addresses, because MAC addresses are only present in the first-hop. Consequently, if the router does not know a host's IP address, the service provider cannot manage the subscriber. This is an issue for public access networks.

2.4. Summary of ND Issue

The ND issues discussed in Sections 2.1 to 2.3 are summarized below. It is worth noting that these issues originate from three causes: multicast, trusting all hosts and Router-NCE-on-Demand. If the causes can be reduced, the issues will also be reduced. This points out the directions for ND optimization.

. Performance issues caused by multicast
    - o LLA DAD degrading performance
    - o Unsolicited RA degrading performance
    - o GUA (or ULA) DAD degrading performance
    - o Router address resolution for hosts degrading performance
    - o Host Address resolution for other hosts degrading performance
    - o Host MAC change announcement degrading performance (minor issue, no further discussion)
. Reliability issues caused by multicast
    - o LLA DAD not reliable for wireless networks
    - o GUA (or ULA) DAD not reliable for wireless networks
. On-link security issues caused by trusting all hosts
    - o Source IP address spoofing
    - o DAD denial
    - o Fake RAs
    - o Fake Redirect
    - o Replay attacks
. Off-link security issues caused by Router-NCE-on-Demand

    - Router NCE exhaustion
  - Performance issue caused by Router-NCE-on-Demand
    - NCE on demand degrading performance
  - Subscriber management issue caused by Router-NCE-on-Demand
    - Lack of subscriber management using ND with SLAAC

3. Review of ND Solutions

This section reviews the ND optimization solutions developed over the years so that network administrators can get an idea of what solutions are available for which issues. The solutions are reviewed in an order that helps to reveal a theme: isolating hosts to solve ND issues. This theme will be further analyzed in Section 3.15 after all the solutions are reviewed.

3.1. ND Solution in Mobile Broadband IPv6

Mobile Broadband IPv6 (MBBv6) is defined in "IPv6 in 3GPP EPS" [RFC6459] and "IPv6 for 3GPP Cellular Hosts" [RFC7066]. The solution key points are:

- Putting every host, i.e. the mobile User Equipment (UE), in a P2P link with the router, i.e. the mobile gateway. MBBv6 also simplifies ND to take advantage of this P2P architecture. As a result:
  - All multicast is effectively turned into unicast
  - The P2P links in MBB do not have link layer address. Therefore, Router-NCE-on-Demand is not needed.
  - Trusting-all-host is only relevant to the router. By applying some filtering at the router, e.g. dropping RAs from the host, even malicious hosts cannot cause security harm.
- Assigning a unique /64 prefix to each host, as each host is a separate link and subnet.
- Maintaining (prefix, interface) binding at the router for forwarding purpose.

Since all the three causes of ND issues are addressed, MBBv6 solves all ND issues.

3.2. ND Solution in Fixed Broadband IPv6

FBBv6 is defined in "IPv6 in the context of TR-101" [TR177]. FBBv6 has two flavors:

. P2P: every host, i.e. the Residential Gateway (RG), is in a P2P link with the router, i.e. the Broadband Network Gateway (BNG). In this case, the solution is essentially the same as MBBv6. All ND issues are solved.
. P2MP: all hosts on an access device, e.g. the Optical Line Terminal (OLT), are in a P2MP link with the router. This is implemented by aggregating all hosts into a single VLAN at the router and implementing Split Horizon at the access device to prevent direct host communication.

The solution key points of FBBv6-P2MP [TR177] are:

. Putting all hosts in a P2MP link with the router, and implementing DAD Proxy. P2MP architecture with Split Horizon breaks normal ND's DAD procedure. Because all hosts are in the same interface from the router's perspective, the router must ensure that the hosts have different LLAs and GUAs. Otherwise, the router will not be able to distinguish them. But because hosts cannot reach each other, normal DAD will not work. Therefore, the router must participate in the hosts' DAD process and help hosts resolve duplication. This is called DAD Proxy [RFC6957]. With P2MP link and DAD Proxy:
  o All upstream multicast from hosts to the router is effectively turned into unicast, as every host can only reach the router.
  o Trusting-all-host is only relevant to the router. By applying some simple filtering at the router, e.g. dropping RAs from the host, even malicious hosts cannot cause security harm.
. Assigning a unique /64 prefix to each host. As a result:
  o When a prefix is assigned to the host, the router can proactively create (IP prefix, MAC) binding and use it for forwarding. There is no need for Router-NCE-on-Demand.
  o Since different hosts are in different subnets, hosts will send traffic to other hosts via the router. There is no address resolution for other hosts.
  o Without address resolution, downstream multicast to hosts consists only of unsolicited RAs. Because every host is in its own subnet, unsolicited RAs will be sent individually to each host with the "host's MAC replacing the multicast MAC" approach specified in [RFC6085]. Therefore, downstream multicast is turned into unicast.

Since all the three causes of ND issues are addressed, FBBv6-P2MP solves all ND issues.

3.3. Unique IPv6 Prefix per Host

Unique IPv6 Prefix per Host is specified in [RFC8273]. The purpose is to "improve host isolation and enhanced subscriber management on shared network segments" such as Wi-Fi or Ethernet. The solution key points are:

. Assigning a unique prefix to each host with SLAAC. As a result:
    - o When a prefix is assigned to the host, the router can proactively create (Prefix, MAC) binding and use it for forwarding. There is no need for Router-NCE-on-Demand.
    - o Since different hosts are in different subnets, hosts will send traffic to other hosts via the router. There is no host to host address resolution.
    - o Without address resolution, downstream multicast to hosts consists only of unsolicited RAs. They will be sent host by host in unicast because the prefix for every host is different.

RFC 8273 believes that "A network implementing a unique IPv6 prefix per host can simply ensure that devices cannot send packets to each other except through the first-hop router". But this may not be true when hosts are on a certain shared medium like Ethernet. In that case, hosts can still reach each other in L2 with their LLAs. So on-link security issues will remain. LLA-DAD-not-reliable issue can still exist for wireless media too. RFC 8273 solves other ND issues discussed in Section 2.

3.4. Wireless ND

Wireless ND (WiND) is specified in a series of RFCs [RFC6775][RFC8505][RFC8928][RFC8929]. WiND defines a new ND solution for Low-Power and Lossy Networks (LLNs) [RFC7102]. WiND changes host and router behaviors to use multicast only for router discovery. The solution key points are (please check if you agree):

. Hosts use unicast to proactively register their addresses at the routers. Routers use unicast to communicate with hosts and become the central address register and arbitrator for the hosts.
. The router also proactively installs Neighbor Cache Entries (NCEs) for the hosts. This avoids the need for address resolution for the hosts.
. The router sets PIO L-bit to 0. Each host communicates only with the router.
. Other functionalities that are relevant only to LLNs.

WiND is a totally new ND solution. It solves all ND issues in LLNs.

3.5. Scalable Address Resolution Protocol

Scalable Address Resolution Protocol (SARP) is specified in [RFC7586]. The usage scenario is Data Centers (DCs) where large L2 domains spanned across multiple sites. In each site, multiple hosts are connected to a switch. The hosts can be Virtual Machines (VMs) so the number can be large. The switches are interconnected by a native or overlay L2 network.

The switch will snoop and install (IP, MAC) proxy table for the local hosts. The switch will also reply to address resolution requests from other sites to its hosts with its own MAC. This way, all hosts in a site will appear to have a single MAC to other sites. Therefore, a switch only needs to build a MAC table for the local hosts and the remote switches, not for all the hosts in the L2 domain. The MAC table size of the switches is therefore significantly reduced. A switch will also add the (IP, MAC) replies from remote switches to its proxy ND table so that it can reply to future address resolution requests for such IPs directly. This greatly reduces the number of address resolution multicast in the network.

Unlike MBBv6, FBBv6 and RFC 8372 which try to solve all ND issues, SARP focuses on reducing address resolution multicast to improve performance and scalability of large L2 domains in DCs.

3.6. ARP and ND Optimization for Transparent Interconnection of Lots of Links (TRILL):

ARP and ND Optimization for TRILL is specified in [RFC8302]. The solution is very similar to SARP discussed in Section 3.5. It can be considered as an application of SARP in the TRILL environment.

Like SARP, ARP and ND Optimization for TRILL focuses on reducing address resolution multicast.

3.7. Proxy ARP/ND in EVPN

Proxy ARP/ND in EVPN is specified in [RFC9161]. The usage scenario is Data Centers (DCs) where large L2 domains spanned across multiple sites. In each site, multiple hosts are connected to a Provider Edge (PE) router acting as a switch. The PEs are interconnected by an overlay network.

PE of each site snoops the local address resolution NAs to build (IP, MAC) Proxy ND table entries. PEs then propagate such Proxy ND entries to other PEs via BGP EVPN. Each PE also snoops address resolution NSs from its hosts. If an entry exists in its Proxy ND table for the specified destination IP address, the PE will reply directly. Consequently, the number of multicast address resolution messages is significantly reduced.

Like SARP, Proxy ARP/ND in EVPN also focuses on reducing address resolution multicast.

3.8. Gratuitous Neighbor Discovery

Gratuitous Neighbor Discovery is specified in [GRAND]. GRAND changes router and host behaviors to allow routers to proactively create an NCE when a new IPv6 address is assigned to a host, and to recommend that hosts send unsolicited Neighbor Advertisements upon having a new IPv6 address. It can be considered as the IPv6 equivalent of Gratuitous ARP in IPv4.

GRAND solves the Router-NCE-on-Demand issue.

3.9. Reducing Router Advertisements

[RFC7772] specifies a solution for reducing RAs. The key points are:

- The router should respond to RS with unicast RA if the host's source IP address is not unspecified (i.e. the RS is not the first RS before GUA DAD) and the host's MAC is valid.
- The router should reduce multicast RA frequency.
- Sleeping hosts that process unicast packets while asleep must also process multicast RAs while asleep.
- Sleeping hosts that do not intend to maintain IPv6 connectivity while asleep should either disconnect from the network and clear all IPv6 configuration, or perform Detecting Network Attachment in IPv6 (DNAv6) procedures [RFC6059] when waking up.

RFC 7772 alleviates the multicast RA issue.

3.10. Source Address Validation Improvement and Router Advertisement Guard

Source Address Validation Improvement is specified in [SAVI]. Router Advertisement Guard is specified in [RA-Guard][RA-Guard+]. SAVI binds an address to a port and rejects claims from other ports for that address. Therefore, a node cannot spoof the IP address of

another node. RA-Guard and RA-Guard+ only allow RAs from a port that a router is connected to. Therefore, nodes on other ports cannot pretend to be a router.

SAVI, RA-Guard and RA-Guard+ solve the on-link security issues.

3.11. Dealing with Off-link Attack that May Cause Router NCE Exhaustion

Router NCE Exhaustion handling is described in [RFC6583]. This is to deal with the off-link security issue discussed in Section 2.3. The solution key points are:

. For operators:
    - o Filtering of unused address space so that messages to such addresses can be dropped rather than triggering NCE creation;
    - o Minimizing subnet size so that there are fewer potential NCEs to create;
    - o Rate-limiting the NDP queue to avoid CPU/memory overflow.
. For vendors:
    - o Prioritizing NDP processing for existing NCEs over creating new NCEs

RFC 6583 acknowledges that "some of these options are 'kludges', and can be operationally difficult to manage". RFC 6583 partially solves the Router NCE Exhaustion issue.

3.12. Enhanced DAD

Enhanced DAD is specified in [RFC7527]. Enhanced DAD solves a DAD failure issue in a specific situation: looped back interface. DAD will fail in a looped back interface because the sending host will receive the DAD message back and will interpret it as another host is trying to use the same address. The solution is to include a Nonce option (defined in [SeND]) in each DAD message so that the sending host can detect that the looped back DAD message is sent by itself.

Enhanced DAD does not solve any of the ND issues discussed in Section 2. It extends ND to work in a new scenario: looped back interface. It is reviewed here for completeness but will not be further discussed.

3.13. ND Mediation for IP Interworking of Layer 2 VPNs

ND mediation is specified in [RFC6575]. When two Attachment Circuits (ACs) are interconnected by a Virtual Private Wired Service (VPWS), and the two ACs are of different medium (e.g. one is Ethernet while the other is Frame Relay), the two Provider Edges (PEs) must interwork to provide mediation service so that a Customer Edge (CE) can resolve the link layer address of the remote end. RFC 6575 specifies such a solution.

ND Mediation does not solve any of the ND issues discussed in Section 2. It extends ND to work in a new scenario: two ACs of different media interconnected by a VPWS. It is reviewed here for completeness but will not be further discussed.

3.14. ND Solutions Defined before the Latest Versions of ND

The latest versions of [ND] and [SLAAC] are specified in RFCs 4861 and 4862. Several ND optimization solutions are based on the older version of ND and SLAAC. They are reviewed in this section for completeness but will not be further discussed.

3.14.1. SeND

Secure Neighbor Discovery [SeND] is specified in RFC 3971. The purpose is to ensure that hosts and routers are trustable. SeND defined three new ND options (i.e. Cryptographically Generated Addresses [CGA], RSA public-key cryptosystem, Timestamp/Nonce), an authorization delegation discovery process, an address ownership proof mechanism, and requirements for the use of these components in NDP.

SeND solves the on-link and off-link security issues. But it has high requirements on the hosts and routers, especially to maintain the keys. SeND is rarely deployed and will not be further discussed in this document.

3.14.2. Cryptographically Generated Addresses (CGA)

Cryptographically Generated Addresses [CGA] is specified in RFC 3972. The purpose is to associate a cryptographic public key with an IPv6 address in [SeND]. The solution key point is to generate the Interface Identifier (IID) of the IPv6 address by computing a cryptographic hash of the public key. The resulting IPv6 address is called a CGA. The corresponding private key can then be used to sign messages sent from the address.

CGA uses the fact that a legitimate host does not care about the bit combination of IID that would be created as a result of some hash procedure. The attacker needs an exact IID to impersonate the legitimate hosts but then the attacker is challenged to do a reverse hash calculation that is a strong mathematical challenge.

CGA is part of SeND. It is rarely deployed and will not be further discussed in this document.

3.14.3. ND Proxy

ND Proxy is specified in [RFC4389]. The purpose is to enable multiple links joined by an ND-Proxy device to work as a single link. The ND-Proxy acts like a bridge. The solution key points are:

. When it receives an ND request from a host in a link, it will "proxy" the message out from the "best" outgoing interface. How to determine the "best" interface is explained later. If there is no "best" interface, the ND-Proxy will "proxy" the message to all other links. Here "proxy" means acting as if the ND message originates from the ND-Proxy itself. That is, the ND-Proxy will change the ND message's source IP and source MAC to the ND-Proxy's outgoing interface's IP and MAC, and create an NCE entry at the outgoing interface accordingly.
. When ND-Proxy receives an ND reply, it will act as if the ND message is destined to itself, and update the NCE entry state at the receiving interface. Based on such state information, the ND-Proxy can determine the "best" outgoing interface for future ND requests. The ND-Proxy then "proxy" the ND message back to the requesting host.

ND Proxy does not solve any of the ND issues discussed in Section 2. It extends ND to work in a new scenario: multiple links joined by a device that is not a bridge but acting like a bridge.

The idea of ND Proxy is widely used in SARP, ND Optimization for TRILL and Proxy ARP/ND in EVPN which are discussed in Sections 3.4 to 3.6.

3.14.4. Optimistic DAD

Optimistic DAD is specified in [RFC4429]. The purpose is to minimize address configuration delays in the successful case and to reduce disruption as far as possible in the failure case. Optimistic DAD modified the original ND (RFC 2461) and SLAAC (RFC 2462) but the

solution was not incorporated into the latest specification of [ND] and [SLAAC].

Optimistic DAD does not solve any of the ND issues discussed in Section 2. It tries to enhance ND's performance for DAD. But the changes are big and the benefits are not significant. Optimistic DAD has not been widely deployed.

3.15. Observations on the Solutions and an Insight Learned

First, which ND solution solves which ND issues is tabulated below for reference later.

There are thirteen ND issues as summarized in Section 2.4:

. Performance issues caused by multicast
    - o I1: LLA DAD degrading performance
    - o I2: Unsolicited RA degrading performance
    - o I3: GUA (or ULA) DAD degrading performance
    - o I4: Router address resolution for hosts degrading performance
    - o I5: Host Address resolution for other hosts degrading performance
. Reliability issues caused by multicast
    - o I6: LLA DAD not reliable for wireless networks
    - o I7: GUA DAD not reliable for wireless networks
. On-link security issues caused by trusting all hosts
    - o I8: Source IP address spoofing
    - o I9: DAD denial
    - o I10: Fake RAs
    - o I11: Fake Redirect
    - o I12: Replay attacks
. Off-link security issues caused by Router-NCE-on-Demand
    - o I13: Router NCE exhaustion
. Performance issue caused by Router-NCE-on-Demand
    - o I14: NCE on demand degrading performance
. Subscriber management issue caused by Router-NCE-on-Demand
    - o I15: Lack of subscriber management using ND with SLAAC

	Multicast performance	Reli-ability	On-link security	Off-link security	NCE on Demand	Sub Mgmt.
Issue	1	2	3	4	5	6	7	8-12	13	14	15
MBBv6	All issues solved										
FBBv6	All issues solved										
8273		X	X	X	X		X		X	X	X
WiND	All issues solved										
SARP					X						
ND TRILL					X						
ND EVPN					X						

7772		X								
GRAND				X						
SAVI/RAG G+							X			
6583								X		

Table 1. Which solution solves which issue(s)

Although the various ND solutions look unrelated, dividing them into four groups will help to reveal a common theme: isolating hosts to solve issues.

The first group contains MBBv6, FBBv6, Unique Prefix Per Host and WiND. These solutions all isolate hosts individually in some way, and they solve all or most ND issues.

The second group contains SARP, ND Optimization for TRILL, and Proxy ND in EVPN. They use a proxy device to represent the hosts behind it, and effectively isolate such hosts from other hosts. The solutions alleviate the biggest ND issue - address resolution among hosts.

The third group contains Reducing RAs, SAVI, RA-Guard, RA-Guard+, and Dealing with Off-link Security Issue. They do not try to isolate hosts to solve many issues. They focus on solving a specific ND issue instead.

The fourth group contains the solutions designated "will not be further discussed". They are not relevant to the discussion here.

This theme reveals an insight: isolating hosts is effective in solving ND issues. The stronger hosts are isolated, the more ND

issues can be solved. This is natural because isolating hosts reduces multicast and hosts to trust, two of the causes of ND issues.

This insight can be used to formulate guidelines to simplify future ND deployment in IPv6 first-hops.

4. Isolating Hosts to Simplify First-hop Deployments

This section describes how to isolate hosts and the advantages and disadvantages of doing so. It also provides some guidelines on how to select a suitable isolation method based on the deployment scenario.

The solution review in Section 3 reveals four different host isolation methods:

. Link isolation has 2 flavors:
    - o P2P link isolation, used in MBBv6 and FBBv6-PPPoE
    - o P2MP link isolation, used in FBBv6-IPoE
. Subnet isolation (i.e. Unique Prefix Per Host), used in MBBv6, FBBv6 and RFC 8273
. GUA isolation (i.e. setting PIO L-bit=0), used in WiND and RFC 8273
    - o GUA isolation is different from link isolation in that there can be multiple hosts in the link. It is just that each host treats other hosts as off-link and does not perform address resolution for other hosts' GUA. The host will send messages with GUA via the router instead. But all messages with LLA can still reach other hosts. In link isolation, there is only one host in each link. A host cannot send messages with LLAs to other hosts.
. Proxy isolation, used in SARP, ND Optimization for TRILL, Proxy ND in EVPN

These different isolation methods are not independent:

First, [RFC4291] stated that "IPv6 continues the IPv4 model in that a subnet prefix is associated with one link". Therefore, link isolation and subnet isolation should be used together, so that the link and the subnet are congruent in scope: both have just one host. Otherwise, additional ND issues will appear and the solution will be more complicated:

. L2 isolation without subnet isolation, which creates a Multi-Link SubNet (MLSN), violates [RFC4291] by making the subnet

associate with multiple links (i.e., the subnet is bigger than the links). The consequence is GUA DAD may not work unless the router provides DAD Proxy. [RFC4903] documented the concerns about MLSN.

. Subnet isolation without L2 isolation, when used for multiple hosts on a shared medium, also violates [RFC4291] by making the subnet not associate with any link (i.e., the subnet with just one host is smaller than the link with multiple hosts). The consequence is, on-link security issues will remain. For example, LLA DAD denial may happen.

Second, link isolation and subnet isolation automatically imply GUA isolation. When there is only one host in a link/subnet, setting PIO L-bit to 1 has the same effect as setting it to 0, because all communication will go through the router.

Third, proxy isolation and other isolation methods are mutually exclusive. Proxy isolation uses a proxy to represent multiple hosts at a site. In other words, hosts are not isolated individually like in link/subnet/GUA isolation.

Therefore, these different isolation methods only produce four meaningful combinations:

. P2P Link and Subnet Isolation
. P2MP Link and Subnet Isolation
. GUA Isolation (without link or subnet isolation)
. Proxy Isolation

These isolation methods are listed from the highest degree of isolation to the lowest. Their applicability is discussed below.

4.1. Applicability of P2P Link and Subnet Isolation

The advantages of applying P2P link and subnet isolation are:

o All ND issues are solved

The disadvantages are:

o The hosts must be able to set up P2P links with the router.

o Many interfaces will be needed at the router, one per host.

o Many prefixes will be needed, one per host.

- This is probably not an issue for IPv6. Today, any company can get a /29 from a Regional Internet Registry (RIR) [RIPE738]. This contains 32 billion /64 prefixes and should be sufficient for any scenarios. The fact that MBBv6 assigns a /64 to every mobile UE [RFC6459], and FBBv6 assigns a /56 to every routed RG [TR177] is evidence.

- All hosts will communicate through the router, and the router may become a bottleneck. So this cannot be used in a high-performance computing environment like DCs.

- Services relying on multicast, e.g. mDNS, will not work.

4.2. Applicability of P2MP Link and Subnet Isolation

The applicability of P2MP Link and Subnet Isolation is the same as P2P, except that:

- DAD Proxy is required in P2MP.

- Hosts do not need the capability to set up P2P links with the router. [PVLAN] or Wireless Isolation [W-Iso] must be configured to enable the P2MP link instead.

- Only one interface is needed at the router

4.3. Applicability of GUA Isolation

The advantages of GUA Isolation are:

- No address resolution for GUA or ULA among hosts. This eliminates the largest source of multicast in ND.

- This is normal ND behavior. No ND optimization solution is needed.

The disadvantages are:

- Only multicast address resolution for GUA or ULA among hosts is eliminated. All other ND issues remain. Consequently, other solutions may be needed to solve such issues.

- All host communication with GUA or ULA will go through the router, and the router may become a bottleneck. So this cannot be used in a high-performance computing environment like DCs.

4.4. Applicability of Proxy Isolation

The advantages of Proxy Isolation are:

- Reduced address resolution for GUA among hosts behind different proxies. This reduces the largest source of multicast in ND.

- Hosts can communicate directly without going through the router. This can be used in a high-performance computing environment like DCs.

The disadvantages are:

- Only multicast address resolution for GUA among hosts behind different proxies is reduced. All other ND issues remain. Consequently, other solutions may be needed to solve such issues.

4.5. ND Deployment Guidelines

Given the applicability analysis above, network administrators can decide where to apply which isolation method.

The guidelines below start from the strongest isolation method. This solves the most ND issues, and therefore, requires fewest additional solutions for the remaining issues. The overall solution will likely be the simplest. But the strongest isolation also has the highest entry requirements and the fewest applicable scenarios. If the strongest isolation is not possible, the next level of isolation is tried, until no isolation is applied. Therefore, network administrators can likely find the most suitable isolation method for their deployment scenarios.

1. If P2P Link and Subnet Isolation is feasible:

    a) Applicable scenarios:

        1) Direct host to host communication is not required.

        2) A P2P architecture is feasible.

        3) Multicast is not desirable (implying mDNS is not needed) for performance or reliability reasons, or

        4) Hosts may not be trustable, or

        5) Subscriber management is needed.

Examples are public access networks such as MBBv6 or FBBv6 PPPoE

b) Entry requirements:

1) Hosts must be able to set up P2P links with the router.

2) The router must have an optimized ND solution that avoids downstream multicast (i.e. DADs, unsolicited RAs, address resolution for hosts), like MBBv6 or FBBv6 or RFC 8273.

c) Remaining issues and solutions:

1) All ND issues are solved

2) Filtering may be needed at the router to discard malicious/erroneous ND messages from hosts, e.g. RAs.

2. Otherwise, if P2MP Link and Subnet Isolation is feasible

a) Applicable scenarios:

1) Same as the P2P scenarios, except that a P2P architecture is not possible while a P2MP architecture is possible.

Examples: FBBv6 IPoE, public Wi-Fi access

b) Entry requirements:

1) The L2 media supports a P2MP architecture (e.g. with PVLAN on Ethernet, or with wireless isolation on Wi-Fi).

2) DAD Proxy must be added on top of the P2P-aware ND optimization solution.

c) Remaining issues and solutions

1) Same as the P2P case

3. Otherwise, if GUA Isolation (i.e. setting PIO L-bit=0) is feasible

a) Applicable scenarios:

1) Direct host to host communication is not required.

2) Multicast is needed, or the L2 medium is not feasible to support P2P/P2MP architecture

Examples: [HomeNet] where mDNS is desired, LLNs.

b) Entry requirements:

1) For LLNs, WiND is required.

2) For other scenarios, no other requirement than ND.

c) Remaining issues and solutions:

1) If WiND is used, all ND issues are solved, as WiND modified ND significantly to solve the issues.

2) If normal ND is used, only multicast address resolution for GUA among hosts is eliminated. All other ND issues may happen. Depending on the specific deployment scenario, only a subset of issues may actually happen.

3) Use Table 1 to pick the solutions for the issues that will actually happen

4. Otherwise, if Proxy Isolation is feasible

a) Applicable scenarios:

1) Direct host to host communication is required.

Examples: large scale DC involving a large number of VMs, and the link spanned across multiple sites interconnected by PEs

b) Entry requirements:

1) A Proxy Isolation solution like SARP, ND Optimization for TRILL or Proxy ND in EVPN

c) Remaining issues and solutions:

1) Only multicast address resolution for GUA among hosts is reduced. All other ND issues may happen. Depending on the specific deployment scenario, only a subset of issues may actually happen.

2) Use Table 1 to pick the solutions for the issues that will actually happen

5. Otherwise, no isolation to apply

a) Applicable scenarios:

1) Small scale and low requirement scenarios

b) Entry requirements:

1) None

c) Remaining issues and solutions

1) All ND issues may happen. Depending on the specific deployment scenario, only some issues may actually happen, and even fewer issues may be of concern, because this is a small scale and low requirement scenario.

2) Use Table 1 to pick the solutions for the issues that are of concern.

4.6. Impact of Host Isolation on Other Protocols in IPv6 First-hops

The impact (i.e. the disadvantages) of various isolation methods has been discussed in the applicability sections. The guidelines have considered such applicability in selecting a suitable isolation method. Therefore, the guidelines will have no negative impact on other protocols in IPv6 first-hops.

Since the guidelines simplify the ND-related part of IPv6 first-hops, and have no negative impact on other protocols, the guidelines simplify the whole IPv6 first-hops.

5. Security Considerations

This document provides guidelines on how to select a suitable isolation method depending on the deployment scenario. When an isolation method is selected, the security considerations of the used solutions apply. This document does not introduce any new solutions. Therefore, it does not introduce new security issues.

6. IANA Considerations

This document has no request to IANA.

7. References

7.1. Informative References

[CGA] T. Aura, "Cryptographically Generated Addresses (CGA)", RFC3972

[DHCPv6] T. Mrugalski M. Siodelski B. Volz A. Yourtchenko M. Richardson S. Jiang T. Lemon T. Winters, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 8415.

[GRAND] J. Linkova, "Gratuitous Neighbor Discovery: Creating Neighbor Cache Entries on First-Hop Routers", RFC 9131

[HomeNet] T. Chown, J. Arkko, A. Brandt, O. Troan, J. Weil, "IPv6 Home Networking Architecture Principles", RFC 7368, DOI 10.17487/RFC7368, October 2014, <https://www.rfc-editor.org/info/rfc7368>.

[mDNS] S. Cheshire, M. Krochmal, "Multicast DNS", RFC 6762.

[ND] 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>.

[PVLAN] https://en.wikipedia.org/wiki/Private_VLAN

[RA-Guard] E. Levy-Abegnoli, G. Van de Velde, C. Popoviciu, J. Mohacsi, "IPv6 Router Advertisement Guard", RFC 6105, DOI 10.17487/RFC6105, February 2011, <https://www.rfc-editor.org/info/rfc6105>.

[RA-Guard+] F. Gont, "Implementation Advice for IPv6 Router Advertisement Guard (RA-Guard)", RFC 7113, DOI 10.17487/RFC7113, February 2014, <https://www.rfc-editor.org/info/rfc7113>.

[RFC3756] P. Nikander, J. Kempf, E. Nordmark, "IPv6 Neighbor Discovery (ND) Trust Models and Threats", RFC 3756.

[RFC4291] R. Hinden, S.Deering, "IP Version 6 Addressing Architecture", RFC 4291.

[RFC4389] D. Thaler, M. Talwar, C. Patel, "Neighbor Discovery Proxies (ND Proxy)", RFC 4389.

[RFC4429] N. Moore, "Optimistic Duplicate Address Detection (DAD) for IPv6", RFC 4429.

[RFC4903] D. Thaler, "Multi-Link Subnet Issues", RFC 4903.

[RFC6459] J. Korhonen, J. Soininen, B. Patil, T. Savolainen, G. Bajko, K. Iisakkila, "IPv6 in 3rd Generation Partnership Project (3GPP) Evolved Packet System (EPS)", RFC 6459.

[RFC6059] S. Krishnan, G. Daley, "Simple Procedures for Detecting Network Attachment in IPv6", RFC 6059.

[RFC6085] S. Gundavelli, M. Townsley, O. Troan, W. Dec, "Address Mapping of IPv6 Multicast Packets on Ethernet", RFC 6085.

[RFC6575] H. Shah, E. Rosen, G. Heron, V. Kompella, "Address Resolution Protocol (ARP) Mediation for IP Interworking of Layer 2 VPNs", RFC 6575.

[RFC6583] I. Gashinsky, J. Jaeggli, W. Kumari, "Operational Neighbor Discovery Problems", RFC 6583.

[RFC6775] Z. Shelby, S. Chakrabarti, E. Nordmark, C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775.

[RFC6957] F. Costa, J-M. Combes, X. Pougnard, H. Li, "Duplicate Address Detection Proxy", RFC 6957

[RFC7066] J. Korhonen, J. Arkko, T. Savolainen, S. Krishnan, "IPv6 for Third Generation Partnership Project (3GPP) Cellular Hosts", RFC 7066.

[RFC7102] JP. Vasseur, "Terms Used in Routing for Low-Power and Lossy Networks", RFC 7102.

[RFC7342] L. Dunbar, W. Kumari, I. Gashinsky, "Practices for Scaling ARP and Neighbor Discovery (ND) in Large Data Centers", RFC 7342.

[RFC7527] R. Asati, H. Singh, W. Beebee, C. Pignataro, E. Dart, W. George, "Enhanced Duplicate Address Detection", RFC 7527.

[RFC7586] Y. Nachum, L. Dunbar, I. Yerushalmi, T. Mizrahi, "The Scalable Address Resolution Protocol (SARP) for Large Data Centers", RFC7586.

[RFC7772] A. Yourtchenko, L. Colitti, "Reducing Energy Consumption of Router Advertisements", RFC 7772.

[RFC8273] J. Brzozowski, G. Van de Velde, "Unique IPv6 Prefix per Host", RFC 8273.

[RFC8302] Y. Li, D. Eastlake 3rd, L. Dunbar, R. Perlman, M. Umair, "Transparent Interconnection of Lots of Links (TRILL): ARP and Neighbor Discovery (ND) Optimization", RFC 8302.

[RFC8505] P. Thubert, E. Nordmark, S. Chakrabarti, C. Perkins, "Registration Extensions for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Neighbor Discovery", RFC 8505.

[RFC8928] P. Thubert, B. Sarikaya, M. Sethi, R. Struik, "Address-Protected Neighbor Discovery for Low-Power and Lossy Networks", RFC 8928.

[RFC8929] P. Thubert, C.E. Perkins, E. Levy-Abegnoli, "IPv6 Backbone Router", RFC 8929.

[RFC9099] E. Vyncke, K. Chittimaneni, M. Kaeo, E. Rey, "Operational Security Considerations for IPv6 Networks", RFC 9099.

[RFC9119] C. Perkins, M. McBride, D. Stanley, W. Kumari, JC. Zuniga, "Multicast Considerations over IEEE 802 Wireless Media", RFC 9119.

[RFC9161] J. Rabadan, S. Sathappan, K. Nagaraj, G. Hankins, T. King, "Operational Aspects of Proxy ARP/ND in Ethernet Virtual Private Networks", RFC 9161

[RIPE738] IPv6 Address Allocation and Assignment Policy, https://www.ripe.net/publications/docs/ripe-738

[SAVI] J. Wu, J. Bi, M. Bagnulo, F. Baker, C. Vogt, "Source Address Validation Improvement (SAVI) Framework", RFC 7039

[SeND] J. Arkko, J. Kempf, B. Zill, P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC3971

[SLAAC] 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>.

[TR177] S. Ooghe, B. Varga, W. Dec, D. Allan, "IPv6 in the context of TR-101", Broadband Forum, TR-177.

[W-Iso] Wireless Isolation, https://www.quora.com/What-is-wireless-isolation

8. Acknowledgments

The authors would like to thank Pascal Thubert, Ole Troan, Brian Carpenter, David Thaler, Jen Linkova, Eric Vyncke, Lorenzo Colitti for the discussion and input.

Authors' Addresses

XiPeng Xiao
Huawei Technologies Dusseldorf
Hansaallee 205, 40549 Dusseldorf, Germany

Email: xipengxiao@huawei.com

Eduard Vasilenko
Huawei Technologies
17/4 Krylatskaya st, Moscow, Russia 121614

Email: vasilenko.eduard@huawei.com

Eduard Metz
KPN N.V.
Maanplein 55, 2516CK The Hague, The Netherlands

Email: eduard.metz@kpn.com

Gyan Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com