Workgroup: 6MAN
Published: 24 November 2020
Intended Status: Informational
Expires: 28 May 2021
Authors: P. Thubert, Ed.
Cisco Systems

IPv6 Neighbor Discovery on Wireless Networks

Abstract

This document describes how the original IPv6 Neighbor Discovery and Wireless ND (WiND) can be applied on various abstractions of wireless media.

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 28 May 2021.

Table of Contents

1. Introduction

[IEEE Std. 802.1] Ethernet Bridging provides an efficient and reliable broadcast service for wired networks; applications and protocols have been built that heavily depend on that feature for their core operation. Unfortunately, Low-Power Lossy Networks (LLNs) and Wireless Local Area Networks (WLANs) generally do not benefit from the same reliable and cheap broadcast capabilities as Ethernet links.

As opposed to unicast transmissions, the broadcast transmissions over wireless links are not subject to automatic retries (ARQ) and can be very unreliable. Reducing the speed at the physical (PHY) layer for broadcast transmissions can increase the reliability, at the expense of a higher relative cost of broadcast on the overall available bandwidth. As a result, protocols designed for bridged networks that rely on broadcast transmissions often exhibit disappointing behaviours when employed unmodified on a local wireless medium (see [MCAST PROBLEMS]).

Like Transparent Bridging, the IPv6 [RFC8200] Neighbor Discovery [RFC4861] [RFC4862] Protocol (ND-Classic) is reactive, and relies on on-demand Network Layer multicast to locate an on-link correspondent (Address Resolution, AR) and ensure the uniqueness of an IPv6

address (Duplicate Address Detection, DAD). On Ethernet LANs and most WLANs and Low-Power Personal Area Networks (LoWPANs), the Network Layer multicast operation is typically implemented as a Link Layer broadcast for the lack of an adapted and scalable Link Layer multicast operation.

It results that on wireless, an ND-Classic multicast message is typically broadcasted. So even though there are very few nodes subscribed to the Network Layer multicast group, and there is at most one intended Target, the broadcast is received by many wireless nodes over the whole subnet (e.g., the ESS fabric). And yet, the broadcast transmission being unreliable, the intended Target may effectively have missed the packet.

On paper, a Wi-Fi station must keep its radio turned on to listen to the periodic series of broadcast frames, which for the most part will be dropped when they reach Network Layer. In order to avoid this waste of energy and increase its battery life, a typical battery-operated device such as an IoT sensor or a smartphone will blindly ignore a ratio of the broadcasts, making ND-Classic operations even less reliable.

Wi-Fi [IEEE Std. 802.11] Access Points (APs) deployed in an Extended Service Set (ESS) act as [IEEE Std. 802.1] bridges between the wireless stations (STA) and the wired backbone. As opposed to the classical Transparent (aka Learning) Bridge operation that installs the forwarding state reactively to traffic, the bridging state in the AP is established proactively, at the time of association. This protects the wireless medium against broadcast-intensive Transparent Bridging lookups. The association process registers the Link Layer (MAC) Address (LLA) of the STA to the AP proactively, i.e., before it is needed. The AP maintains the list of the associated addresses and blocks the lookups for destinations that are not registered. This solves the broadcast issue for the Link Layer lookups, but the Network Layer problem remains.

Though ND-Classic was the state of the art when designed for an Ethernet wire at the end of the twentieth century, it must be reevaluated for the new technologies, such as wireless and overlays, that evolved since then. This document discusses the applicability of ND-Classic over wireless links, as compared with routing-based alternatives such as prefix-per node and multi-link subnets (MLSN), and with Wireless ND (WiND), that is similar to the Wi-Fi association and reduces the need for Network Layer multicast.

2. Acronyms

This document uses the following abbreviations:

6BBR: 6LoWPAN Backbone Router
6LN: 6LoWPAN Node
6LR: 6LoWPAN Router
ARO: Address Registration Option
DAC: Duplicate Address Confirmation
DAD: Duplicate Address Detection
DAR: Duplicate Address Request
EDAC: Extended Duplicate Address Confirmation
EDAR: Extended Duplicate Address Request
MLSN: Multi-link subnet
LLN: Low-Power and Lossy Network
LoWPAN: Low-Power Wireless Personal Area Network
NA: Neighbor Advertisement
NBMA: Non-Broadcast Multi-Access
NCE: Neighbor Cache Entry
ND: Neighbor Discovery
NDP: Neighbor Discovery Protocol
NS: Neighbor Solicitation
RPL: IPv6 Routing Protocol for LLNs
RA: Router Advertisement
RS: Router Solicitation
VLAN: Virtual Local Area Network
WiND: Wireless Neighbor Discovery
WLAN: Wireless Local Area Network
WPAN: Wireless Personal Area Network

3. ND-Classic, Wireless ND and ND-Proxies

The ND-Classic Neighbor Solicitation (NS) [RFC4861] message is used as a multicast IP packet for Address Resolution (AR) and Duplicate Address Detection (DAD) [RFC4862]. In those cases, the NS message is sent at the Network Layer to a Solicited-Node Multicast Address (SNMA) [RFC4291] and should in theory only reach a very small group of nodes. It is intended for one Target, that may or may not be present in the network, but it is often turned into a MAC-Layer broadcast and effectively reaches most of the nodes on link.

DAD was designed for the efficient broadcast operation of Ethernet. Experiments show that DAD often fails to discover the duplication of IPv6 addresses in large wireless access networks [DAD ISSUES]. In practice, IPv6 addresses very rarely conflict, not because the address duplications are detected and resolved by the DAD operation, but thanks to the entropy of the 64-bit Interface IDs (IIDs) that makes a collision quasi-impossible for randomized IIDs.

Multicast NS transmissions may occur when a node joins the network, moves, or wakes up and reconnects to the network. Over a very large fabric, this can generate hundreds of broadcasts per second. If the broadcasts were blindly copied over Wi-Fi, the MAC-layer broadcast

traffic associated to ND IP-layer multicast could consume enough bandwidth to cause a substantial degradation to the unicast service [MCAST EFFICIENCY]. To protect their bandwidth, some networks throttle ND-related broadcasts, which reduces the capability for the ND protocol to operate as expected.

This problem can be alleviated by reducing the size of the broadcast domain that encompasses wireless access links. This has been done in the art of IP subnetting by partitioning the subnets and by routing between them, at the extreme by assigning a /64 prefix to each wireless node (see [RFC8273]).

Another way to split the broadcast domain within a subnet is to proxy at the boundary of the wired and wireless domains the Network Layer protocols that rely on Link Layer broadcast operations. [IEEE Std. 802.11] recommends to deploy proxies for the IPv4 Address Resolution Protocol (ARP) and IPv6 ND at the APs. This requires the exhaustive list of the IP addresses for which proxying is provided. Forming and maintaining that knowledge a hard problem in the general case of radio connectivity, which keeps changing with movements and variations in the environment that alter the range of transmissions.

[SAVI] suggests to discover the addresses by snooping the ND-Classic protocol, but that can also be unreliable. An IPv6 address may not be discovered immediately due to a packet loss. It may never be discovered in the case of a "silent" node that is not currently using one of its addresses, e.g., a printer that waits in wake-on-lan state. A change of anchor, e.g. due to a movement, may be missed or misordered, leading to unreliable connectivity and an incomplete list of addresses.

Wireless ND (WiND) introduces a new approach to IPv6 Neighbor Discovery that is designed to apply to the WLANs and LoWPANs types of networks, as well as other Non-Broadcast Multi-Access (NBMA) networks such as Data-Center overlays. WiND applies routing inside the subnets, which enables to form potentially large MLSNs without creating a large broadcast domain at the Link Layer. In a fashion similar to a Wi-Fi Association, IPv6 Hosts register their addresses to their serving router(s), using [RFC8505]. With the registration, the routers have a complete knowledge of the hosts they serve and in return, hosts obtain routing services for their registered addresses. The registration is abstract to the routing service, and it can be protected to prevent impersonation attacks with [RFC8928].

The routing service can be a simple reflexion in a Hub-and-Spoke subnet that emulates an IEEE Std. 802.11 Infrastructure BSS at the Network Layer. It can also be a full-fledge routing protocol, in particular RPL [RFC6550], which is designed to adapt to various LLNs such as WLAN and WPAN radio meshes. Finally, the routing service can

also be an ND proxy that emulates an IEEE Std. 802.11 Infrastructure ESS at the Network Layer, as specified in the IPv6 Backbone Router [RFC8929].

On the one hand, WiND avoids the use of broadcast operation for DAD and AR, and on the other hand, WiND supports use cases where subnet and Link Layer domains are not congruent, which is common in wireless networks unless a specific Link Layer emulation is provided. More details on WiND can be found in Section 5.1.

4. IP Models

4.1. Physical Broadcast Domain

At the physical (PHY) Layer, a broadcast domain is the set of nodes that may receive a transmission that one sends over an interface, in other words the set of nodes in range of the radio transmission. This set can comprise a single peer on a serial cable used as point-to-point (P2P) link. It may also comprise multiple peer nodes on a broadcast radio or a shared physical resource such as the Ethernet wires and hubs for which ND-Classic was initially designed.

On WLAN and LoWPAN radios, the physical broadcast domain is defined relative to a particular transmitter, as the set of nodes that can receive what this transmitter is sending. Literally every frame defines its own broadcast domain since the chances of reception of a given frame are statistical. In average and in stable conditions, the broadcast domain of a particular node can be still be seen as mostly constant and can be used to define a closure of nodes on which an upper Layer abstraction can be built.

A PHY Layer communication can be established between two nodes if the physical broadcast domains of their unicast transmissions overlap. On WLAN and LoWPAN radios, that relation is usually not reflexive, since nodes disable the reception when they transmit; still they may retain a copy of the transmitted frame, so it can be seen as reflexive at the MAC Layer. It is often symmetric, meaning that if B can receive a frame from A, then A can receive a frame from B. But there can be asymmetries due to power levels, interferers near one of the receivers, or differences in the quality of the hardware (e.g., crystals, PAs and antennas) that may affect the balance to the point that the connectivity becomes mostly uni-directional, e.g., A to B but practically not B to A.

It takes a particular effort to place a set of devices in a fashion that all their physical broadcast domains fully overlap, and that specific situation can not be assumed in the general case. In other words, the relation of radio connectivity is generally not

transitive, meaning that A in range with B and B in range with C does not necessarily imply that A is in range with C.

4.2. Link Layer Broadcast Emulations

We call Direct MAC Broadcast (DMB) the transmission mode where the broadcast domain that is usable at the MAC layer is directly the physical broadcast domain. [IEEE Std. 802.15.4] and [IEEE Std. 802.11] OCB (for Out of the Context of a BSS) are examples of DMB radios. DMB networks provide mostly symmetric and non-transitive transmission. This contrasts with a number of Link Layer Broadcast Emulation (LLBE) schemes that are described in this section.

In the case of Ethernet, while a physical broadcast domain is constrained to a single shared wire, the [IEEE Std. 802.1] bridging function emulates the broadcast properties of that wire over a whole physical mesh of Ethernet links. For the upper layer, the qualities of the shared wire are essentially conserved, with a reliable and cheap broadcast operation over a transitive closure of nodes defined by their connectivity to the emulated wire.

In large switched fabrics, overlay techniques enable a limited connectivity between nodes that are known to a Map Resolver. The emulated broadcast domain is configured to the system, e.g., with a VXLAN network identifier (VNID). Broadcast operations on the overlay can be emulated but can become very expensive, and it makes sense to proactively install the relevant state in the mapping server as opposed to rely on reactive broadcast lookups to do so.

An [IEEE Std. 802.11] Infrastructure Basic Service Set (BSS) also provides a transitive closure of nodes as defined by the broadcast domain of a central AP. The AP relays both unicast and broadcast packets and provides the symmetric and transitive emulation of a shared wire between the associated nodes, with the capability to signal link-up/link-down to the upper layer. Within a BSS, the physical broadcast domain of the AP serves as emulated broadcast domain for all the nodes that are associated to the AP. Broadcast packets are relayed by the AP and are not acknowledged. To increase the chances that all nodes in the BSS receive the broadcast transmission, AP transmits at the slowest PHY speed. This translates into maximum co-channel interferences for others and the longest occupancy of the medium, for a duration that can be a hundred times that of the unicast transmission of a frame of the same size.

For that reason, upper layer protocols should tend to avoid the use of broadcast when operating over Wi-Fi. To cope with this problems, APs may implement strategies such as turn a broadcast into a series of unicast transmissions, or drop the message altogether, which may impact the upper layer protocols. For instance, some APs may not

copy Router Solicitation (RS) messages under the assumption that there is no router across the wireless interface. This assumption may be correct at some point of time and may become incorrect in the future. Another strategy used in Wi-Fi APS is to proxy protocols that heavily rely on broadcast, such as the Address Resolution in ARP and ND-Classic, and either respond on behalf or preferably forward the broadcast frame as a unicast to the intended Target.

In an [IEEE Std. 802.11] Infrastructure Extended Service Set (ESS), infrastructure BSSes are interconnected by a bridged network, typically running Transparent Bridging and the Spanning tree Protocol or a more advanced Layer 2 Routing (L2R) scheme. In the original model of learning bridges, the forwarding state is set by observing the source MAC address of the frames. When a state is missing for a destination MAC address, the frame is broadcasted with the expectation that the response will populate the state on the reverse path. This is a reactive operation, meaning that the state is populated reactively to the need to reach a destination. It is also possible in the original model to broadcast a gratuitous frame to advertise self throughout the bridged network, and that is also a broadcast.

The process of the Wi-Fi association prepares a bridging state proactively at the AP, which avoids the need for a reactive broadcast lookup over the wireless access. In an ESS, the AP may also generate a gratuitous broadcast sourced at the MAC address of the STA to prepare or update the state in the learning bridges so they point towards the AP for the MAC address of the STA. WiND emulates that proactive method at the Network Layer for the operations of AR, DAD and ND proxy.

In some instances of WLANs and LoWPANs, a Mesh-Under technology (e.g., a IEEE Std. 802.11s or IEEE Std. 802.15.10) provides meshing services that are similar to bridging, and the broadcast domain is well-defined by the membership of the mesh. Mesh-Under emulates a broadcast domain by flooding the broadcast packets at the Link Layer. When operating on a single frequency, this operation is known to interfere with itself, and requires inter-frame gaps to dampen the collisions, which reduces further the amount of available bandwidth.

As the cost of broadcast transmissions becomes increasingly expensive, there is a push to rethink the upper Layer protocols to reduce the dependency on broadcast operations.

4.3. Mapping the IPv6 link Abstraction

IPv6 defines a concept of Link, link Scope and Link-Local Addresses (LLA), an LLA being unique and usable only within the Scope of a

Link. The ND-Classic [RFC4861] DAD [RFC4862] process uses a multicast transmission to detect a duplicate address, which requires that the owner of the address is connected to the Link Layer broadcast domain of the sender.

On wired media, the link is often confused with the physical broadcast domain because both are determined by the serial cable or the Ethernet shared wire. Ethernet Bridging reinforces that illusion with a Link Layer broadcast domain that emulates a physical broadcast domain over the mesh of wires. But the difference shows on legacy Non-Broadcast Multi-Access (NBMA) networks such as ATM and Frame-Relay, on shared links and on newer types of NBMA networks such as radio and composite radio-wires networks. It also shows when private VLANs or Link Layer cryptography restrict the capability to read a frame to a subset of the connected nodes.

In Mesh-Under and Infrastructure BSS, the IP link extends beyond the physical broadcast domain to the emulated Link Layer broadcast domain. Relying on Multicast for the ND operation remains feasible but becomes highly detrimental to the unicast traffic, and becomes less and less energy-efficient and reliable as the network grows.

On DMB radios, IP links between peers come and go as the individual physical broadcast domains of the transmitters meet and overlap. The DAD operation cannot provide once and for all guarantees over the broadcast domain defined by one radio transmitter if that transmitter keeps meeting new peers on the go.

The scope on which the uniqueness of an LLA must be checked is each new pair of nodes for the duration of their conversation. As long as there's no conflict, a node may use the same LLA with multiple peers but it has to perform DAD again with each new peer. A node may need to form a new LLA to talk to a new peer, and multiple LLAs may be present in the same radio interface to talk to different peers. In practice, each pair of nodes defines a temporary P2P link, which can be modeled as a sub-interface of the radio interface.

The DAD and AR procedures in ND-Classic expect that a node in a subnet is reachable within the broadcast domain of any other node in the subnet when that other node attempts to form an address that would be a duplicate or attempts to resolve the MAC address of this node. This is why ND is applicable for P2P and transit links, but requires extensions for more complex topologies.

4.4. Mapping the IPv6 subnet Abstraction

IPv6 also defines the concept of a subnet for Global and Unique Local Addresses (GLA and ULA). All the addresses in a subnet share the same prefix, and by extension, a node belongs to a subnet if it

has an address that derives from the prefix of the subnet. That address must be topologically correct, meaning that it must be installed on an interface that is connected to the subnet.

Unless intently replicated in different locations for very specific purposes, a subnet prefix is unique within a routing system; for ULAs, the routing system is typically a limited domain, whereas for GLAs, it is the whole Internet.

For that reason, it is sufficient to validate that an address that is formed from a subnet prefix is unique within the scope of that subnet to guarantee that it is globally unique within the whole routing system. Note that a subnet may become partitioned due to the loss of a wired or wireless link, so even that operation is not necessarily obvious, more in [DAD APPROACHES].

The IPv6 aggregation model relies on the property that a packet from the outside of a subnet can be routed to any router that belongs to the subnet, and that this router will be able to either resolve the destination Link Layer address and deliver the packet, or, in the case of an MLSN, route the packet to the destination within the subnet.

If the subnet is known as on-link, then any node may also resolve the destination Link Layer address and deliver the packet, but if the subnet is not on-link, then a host in the subnet that does not have a Neighbor Cache Entry (NCE) for the destination will also need to pass the packet to a router, more in [RFC5942].

On Ethernet, an IP subnet is often congruent with an IP link because both are determined by the physical attachment to a shared wire or an IEEE Std. 802.1 bridged domain. In that case, the connectivity over the link is both symmetric and transitive, the subnet can appear as on-link, and any node can resolve a destination MAC address of any other node directly using ND-Classic.

But an IP link and an IP subnet are not always congruent. In the case of a Shared Link, individual subnets may each encompass only a subset of the nodes connected to the link. Conversely, in Route-Over Multi-link subnets (MLSN) [RFC4903], routers federate the links between nodes that belong to the subnet, the subnet is not on-link and it extends beyond any of the federated links.

5. Wireless Neighbor Discovery

5.1. Introduction to Wireless ND

WiND [RFC6775][RFC8505][RFC8929][RFC8928] defines a new operation for ND that is based on 2 major paradigm changes, proactive address registration by hosts to their attachment routers and routing to

host routes (/128) within the subnet. This allows WiND to avoid the expectations of transit links and subnet-wide broadcast domains.

WiND is agnostic to the method used for Address Assignment, e.g., Stateless Address Autoconfiguration (SLAAC) [RFC4862] or DHCPv6 [RFC8415]. It does not change the IPv6 addressing [RFC4291] or the current practices of assigning prefixes, typically a /64, to a subnet. But the DAD operation is performed as a unicast exchange with a central registrar, using new ND Extended Duplicate Address messages (EDAR and EDAC) [RFC6775][RFC8505]. This modernizes ND for application in overlays with Map Resolvers and enables unicast lookups [UNICAST AR] for addresses registered to the resolver.

The proactive address registration is performed with a new option in NS/NA messages, the Extended Address Registration Option (EARO) defined in [RFC8505]. This method allows to prepare and maintain the host routes in the routers and avoids the reactive Address Resolution in ND-Classic and the associated Link Layer broadcasts transmissions.

The EARO provides information to the router that is independent to the routing protocol and routing can take multiple forms, from a traditional IGP to a collapsed Hub-and-Spoke model where only one router owns and advertises the prefix. [RFC8505] is already referenced as the registrtaion interface to "RIFT: Routing in Fat Trees" [I-D.ietf-rift-rift] and "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks" [RFC6550] with [RPL UNAWARE LEAVES].

WiND also enables to span a subnet over an MLSN that federates edge wireless links with a high-speed, typically Ethernet, backbone. This way, nodes can form any address they want and move freely from a wireless edge link to another, without renumbering. Backbone Routers (6BBRs) placed along the wireless edge of the Backbone handle IPv6 Neighbor Discovery and forward packets over the backbone on behalf of the registered nodes on the wireless edge.

For instance, a 6BBR in bridging proxy mode (more in [RFC8929]) can operate as a Layer-3 AP to serve wireless IPv6 hosts that are Wi-Fi STAs and maintain the reachability for Global Unicast and Link-LOcal Addresses within the federated MLSN.

5.2. links and Link-Local Addresses

For Link-Local Addresses, DAD is typically performed between communicating pairs of nodes and an NCE can be populated with a single unicast exchange. In the case of a bridging proxies, though, the Link-Local traffic is bridged over the backbone and the DAD must proxied there as well.

For instance, in the case of Bluetooth Low Energy (BLE) [RFC7668] [IEEEstd802151], the uniqueness of Link-Local Addresses needs only to be verified between the pair of communicating nodes, the central router and the peripheral host. In that example, 2 peripheral hosts connected to the same central router can not have the same Link-Local Address because the addresses would collision at the central router which could not talk to both over the same interface. The DAD operation from WiND is appropriate for that use case, but the one from ND is not, because the peripheral hosts are not on the same broadcast domain.

On the other hand, the uniqueness of Global and Unique-Local Addresses is validated at the subnet Level, using a logical registrar that is global to the subnet.

5.3. subnets and Global Addresses

WiND extends ND-Classic for Hub-and-Spoke (e.g., BLE) and Route-Over (e.g., RPL) Multi-link subnets (MLSNs).

In the Hub-and-Spoke case, each Hub-Spoke pair is a distinct IP Link, and a subnet can be mapped on a collection of links that are connected to the Hub. The subnet prefix is associated to the Hub.

Acting as a router, the Hub advertises the prefix as not-on-link to the spokes in RA messages Prefix Information Options (PIO). Acting as hosts, the Spokes autoconfigure addresses from that prefix and register them to the Hub with a corresponding lifetime.

Acting as a registrar, the Hub maintains a binding table of all the registered IP addresses and rejects duplicate registrations, thus ensuring a DAD protection for a registered address even if the registering node is sleeping.

The Hub also maintains an NCE for the registered addresses and can deliver a packet to any of them during their respective lifetimes. It can be observed that this design builds a form of Network Layer Infrastructure BSS.

A Route-Over MLSN is considered as a collection of Hub-and-Spoke where the Hubs form a connected dominating set of the member nodes of the subnet, and IPv6 routing takes place between the Hubs within the subnet. A single logical registrar is deployed to serve the whole mesh.

The registration in [RFC8505] is abstract to the routing protocol and provides enough information to feed a routing protocol such as RPL as specified in [RPL UNAWARE LEAVES]. In a degraded mode, all the Hubs are connected to a same high speed backbone such as an Ethernet bridging domain where ND-Classic is operated. In that case,

it is possible to federate the Hub, Spoke and Backbone nodes as a single subnet, operating ND proxy operations [RFC8929] at the Hubs, acting as 6BBRs. It can be observed that this latter design builds a form of Network Layer Infrastructure ESS.

6. WiND Applicability

WiND applies equally to P2P links, P2MP Hub-and-Spoke, Link Layer Broadcast Domain Emulation such as Mesh-Under and Wi-Fi BSS, and Route-Over meshes.

There is an intersection where link and subnet are congruent and where both ND and WiND could apply. These includes P2P, the MAC emulation of a PHY broadcast domain, and the particular case of always on, fully overlapping physical radio broadcast domain. But even in those cases where both are possible, WiND is preferable vs. ND because it reduces the need of broadcast.

This is discussed in more details in the introduction of [RFC8929].

There are also a number of practical use cases in the wireless world where links and subnets are not congruent:

* The IEEE Std. 802.11 infrastructure BSS enables one subnet per AP, and emulates a broadcast domain at the Link Layer. The Infrastructure ESS extends that model over a backbone and recommends the use of an ND proxy [IEEE Std. 802.11] to interoperate with Ethernet-connected nodes. WiND incorporates an ND proxy to serve that need, which was missing so far.

* BlueTooth is Hub-and-Spoke at the Link Layer. It would make little sense to configure a different subnet between the central and each individual peripheral node (e.g., sensor). Rather, [RFC7668] allocates a prefix to the central node acting as router, and each peripheral host (acting as a host) forms one or more address(es) from that same prefix and registers it.

* A typical Smartgrid networks puts together Route-Over MLSNs that comprise thousands of IPv6 nodes. The 6TiSCH architecture [I-D.ietf-6tisch-architecture] presents the Route-Over model over an IEEE Std. 802.15.4 Time-Slotted Channel-Hopping (TSCH) [IEEEstd802154] mesh, and generalizes it for multiple other applications.

  Each node in a Smartgrid network may have tens to a hundred others nodes in range. A key problem for the routing protocol is which other node(s) should this node peer with, because most of the possible peers do not provide added routing value. When both energy and bandwidth are constrained, talking to them is a waste of resources and most of the possible P2P links are not even

used. Peerings that are actually used come and go with the dynamics of radio signal propagation. It results that allocating prefixes to all the possible P2P links and maintain as many addresses in all nodes is not even considered.

6.1. Case of LPWANs

LPWANs are by nature so constrained that the addresses and subnets are fully pre-configured and operate as P2P or Hub-and-Spoke. This saves the steps of neighbor Discovery and enables a very efficient stateful compression of the IPv6 header.

6.2. Case of Infrastructure BSS and ESS

In contrast to IPv4, IPv6 enables a node to form multiple addresses, some of them temporary to elusive, and with a particular attention paid to privacy. Addresses may be formed and deprecated asynchronously to the association.

Snooping protocols such as ND-Classic and DHCPv6 and observing data traffic sourced at the STA provides an imperfect knowledge of the state of the STA at the AP. Missing a state or a transition may result in the loss of connectivity for some of the addresses, in particular for an address that is rarely used, belongs to a sleeping node, or one in a situation of mobility. This may also result in undesirable remanent state in the AP when the STA ceases to use an IPv6 address while remaining associated. It results that snooping protocols is not a recommended technique and that it should only be used as last resort, when the WiND registration is not available to populate the state.

The recommended alternative method is to use the WiND Registration for IPv6 Addresses. This way, the AP exposes its capability to proxy ND to the STA in Router Advertisement messages. In turn, the STA may request proxy ND services from the AP for all of its IPv6 addresses, using the Extended Address Registration Option, which provides the following elements:

* The registration state has a lifetime that limits unwanted state remanence in the network.

* The registration is optionally secured using [RFC8928] to prevent address theft and impersonation.

* The registration carries a sequence number, which enables to figure the order of events in a fast mobility scenario without loss of connectivity.

The ESS mode requires a proxy ND operation at the AP. The proxy ND operation must cover Duplicate Address Detection, Neighbor

Unreachability Detection, Address Resolution and Address Mobility to transfer a role of ND proxy to the AP where a STA is associated following the mobility of the STA.

The WiND proxy ND specification that associated to the Address Registration is [RFC8929]. With that specification, the AP participates to the protocol as a Backbone Router, typically operating as a bridging proxy though the routing proxy operation is also possible. As a bridging proxy, the backbone router either replies to NS lookups with the MAC address of the STA, or preferably forwards the lookups to the STA as Link Layer unicast frames to let the STA answer. For the data plane, the backbone router acts as a normal AP and bridges the packets to the STA as usual. As a routing proxy, the backbone router replies with its own MAC address and then routes to the STA at the IP layer. The routing proxy reduces the need to expose the MAC address of the STA on the wired side, for a better stability and scalability of the bridged fabric.

6.3. Case of Mesh Under Technologies

The Mesh-Under provides a broadcast domain emulation with symmetric and Transitive properties and defines a transit link for IPv6 operations. It results that the model for IPv6 operation is similar to that of a BSS, with the root of the mesh operating as an Access Point does in a BSS/ESS.

While it is still possible to operate ND-Classic, the inefficiencies of the flooding operation make the associated operations even less desirable than in a BSS, and the use of WiND is highly recommended.

6.4. Case of DMB radios

IPv6 over DMB radios uses P2P links that can be formed and maintained when a pair of DMB radios transmitters are in range from one another.

6.4.1. Using ND-Classic only

DMB radios do not provide MAC level broadcast emulation. An example of that is IEEE Std. 802.11 OCB which uses IEEE Std. 802.11 MAC/PHYs but does not provide the BSS functions.

It is possible to form P2P IP links between each individual pairs of nodes and operate ND-Classic over those links with Link-Local addresses. DAD must be performed for all addresses on all P2P IP links.

If special deployment care is taken so that the physical broadcast domains of a collection of the nodes fully overlap, then it is also

possible to build an IP subnet within that collection of nodes and operate ND-Classic.

If an external mechanism avoids duplicate addresses and if the deployment ensures the connectivity between peers, a non-transit Hub-and-Spoke deployment is also possible where the Hub is the only router in the subnet and the Prefix is advertised as not on-link.

6.4.2. Using Wireless ND

Though this can be achieved with ND-Classic, WiND is the recommended approach since it uses unicast communications which are more reliable and less impacting for other users of the medium.

The routers send RAs with a SLLAO at a regular period. The period can be indicated in the RA-Interval Option [RFC6275]. If available, the message can be transported in a compressed form in a beacon, e.g., in OCB Basic Safety Messages (BSM) that are nominally sent every 100ms.

An active beaconing mode is possible whereby the Host sends broadcast RS messages to which a router can answer with a unicast RA.

A router that has Internet connectivity and is willing to serve as an Internet Access may advertise itself as a default router [RFC4191] in its RA messages. The RA is sent over an unspecified link where it does not conflict to anyone, so DAD is not necessary at that stage.

The host instantiates a link where the router's address is not a duplicate. To achieve this, it forms an LLA that does not conflict with that of the router and registers to the router using [RFC8505]. If the router sent an RA(PIO), the host can also autoconfigure an address from the advertised prefix and register it.

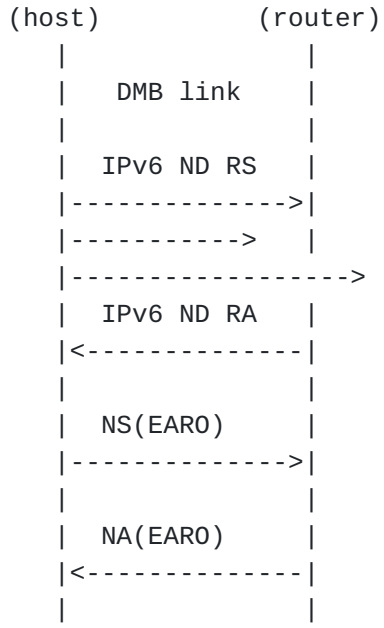

Figure 1: Initial Registration Flow

The lifetime in the registration should start with a small value (X=RMin, TBD), and exponentially grow with each re-registration to a larger value (X=Rmax, TBD). The IP link is considered down when (X=NbBeacons, TDB) expected messages are not received in a row. It must be noted that the link flapping does not affect the state of the registration and when a link comes back up, the active registrations (i.e., registrations for which lifetime is not elapsed) are still usable. Packets should be held or destroyed when the link is down.

P2P links may be federated in Hub-and-Spoke and then in Route-Over MLSNs as illustrated in Figure 2. More details on the operation of WiND and RPL over the MLSN can be found in section 3.1, 3.2, 4.1 and 4.2.2 of [I-D.ietf-6tisch-architecture].

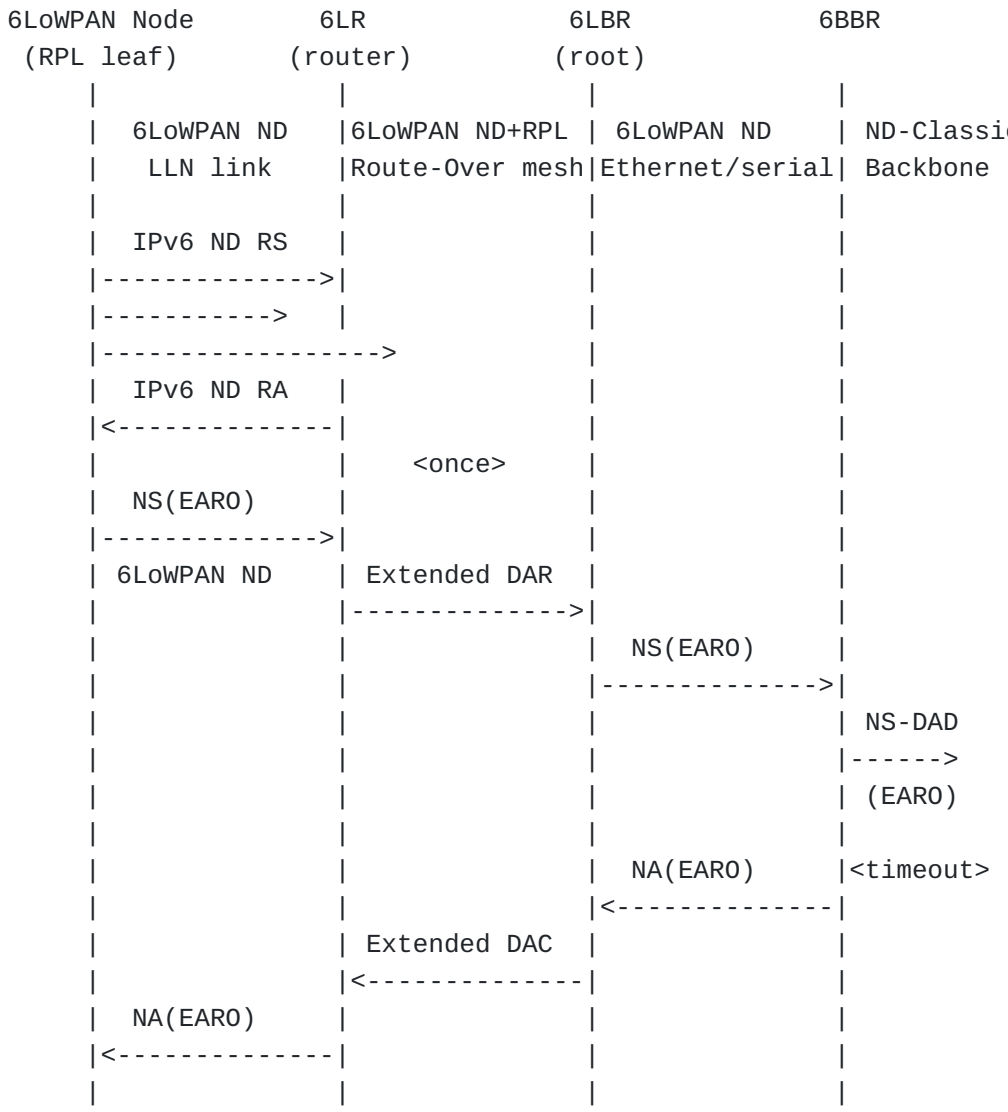

Figure 2: Initial Registration Flow over Multi-link subnet

An example Hub-and-Spoke is an OCB Road-Side Unit (RSU) that owns a prefix, provides Internet connectivity using that prefix to On-Board Units (OBUs) within its physical broadcast domain. An example of Route-Over MLSN is a collection of cars in a parking lot operating RPL to extend the connectivity provided by the RSU beyond its physical broadcast domain. Cars may then operate NEMO [RFC3963] for their own prefix using their address derived from the prefix of the RSU as CareOf Address.

7. IANA Considerations

This specification does not require IANA action.

8. Security Considerations

This specification refers to the security sections of ND-Classic and WiND, respectively.

9. Acknowledgments

Many thanks to the participants of the 6lo WG where a lot of the work discussed here happened. Also ROLL, 6TiSCH, and 6LoWPAN.

10. Normative References

[RFC3963] Devarapalli, V., Wakikawa, R., Petrescu, A., and P. Thubert, "Network Mobility (NEMO) Basic Support Protocol", RFC 3963, DOI 10.17487/RFC3963, January 2005, <https://www.rfc-editor.org/info/rfc3963>.

[RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and More-Specific Routes", RFC 4191, DOI 10.17487/RFC4191, November 2005, <https://www.rfc-editor.org/info/rfc4191>.

[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>.

[RFC5942] Singh, H., Beebee, W., and E. Nordmark, "IPv6 Subnet Model: The Relationship between Links and Subnet Prefixes", RFC 5942, DOI 10.17487/RFC5942, July 2010, <https://www.rfc-editor.org/info/rfc5942>.

[RFC6275] Perkins, C., Ed., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, DOI 10.17487/RFC6275, July 2011, <https://www.rfc-editor.org/info/rfc6275>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[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>.

[RFC8928] Thubert, P., Ed., Sarikaya, B., Sethi, M., and R. Struik, "Address-Protected Neighbor Discovery for Low-Power and Lossy Networks", RFC 8928, DOI 10.17487/RFC8928, November 2020, <https://www.rfc-editor.org/info/rfc8928>.

[RFC8929] Thubert, P., Ed., Perkins, C.E., and E. Levy-Abegnoli, "IPv6 Backbone Router", RFC 8929, DOI 10.17487/RFC8929, November 2020, <https://www.rfc-editor.org/info/rfc8929>.

11. Informative References

[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>.

[RFC4903] Thaler, D., "Multi-Link Subnet Issues", RFC 4903, DOI 10.17487/RFC4903, June 2007, <https://www.rfc-editor.org/info/rfc4903>.

[RFC6550] Winter, T., Ed., Thubert, P., Ed., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, JP., and R. Alexander, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks", RFC 6550, DOI 10.17487/RFC6550, March 2012, <https://www.rfc-editor.org/info/rfc6550>.

[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>.

[RFC7668] Nieminen, J., Savolainen, T., Isomaki, M., Patil, B., Shelby, Z., and C. Gomez, "IPv6 over BLUETOOTH(R) Low Energy", RFC 7668, DOI 10.17487/RFC7668, October 2015, <https://www.rfc-editor.org/info/rfc7668>.

[RFC8273] Brzozowski, J. and G. Van de Velde, "Unique IPv6 Prefix per Host", RFC 8273, DOI 10.17487/RFC8273, December 2017, <https://www.rfc-editor.org/info/rfc8273>.

[RFC8415] Mrugalski, T., Siodelski, M., Volz, B., Yourtchenko, A., Richardson, M., Jiang, S., Lemon, T., and T. Winters, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)",

RFC 8415, DOI 10.17487/RFC8415, November 2018, <https://www.rfc-editor.org/info/rfc8415>.

[I-D.ietf-rift-rift]
Przygienda, T., Sharma, A., Thubert, P., Rijsman, B., and D. Afanasiev, "RIFT: Routing in Fat Trees", Work in Progress, Internet-Draft, draft-ietf-rift-rift-12, 26 May 2020, <https://tools.ietf.org/html/draft-ietf-rift-rift-12>.

[RPL UNAWARE LEAVES] Thubert, P. and M. Richardson, "Routing for RPL Leaves", Work in Progress, Internet-Draft, draft-ietf-roll-unaware-leaves-23, 10 November 2020, <https://tools.ietf.org/html/draft-ietf-roll-unaware-leaves-23>.

[DAD ISSUES] Yourtchenko, A. and E. Nordmark, "A survey of issues related to IPv6 Duplicate Address Detection", Work in Progress, Internet-Draft, draft-yourtchenko-6man-dad-issues-01, 3 March 2015, <https://tools.ietf.org/html/draft-yourtchenko-6man-dad-issues-01>.

[MCAST EFFICIENCY] Vyncke, E., Thubert, P., Levy-Abegnoli, E., and A. Yourtchenko, "Why Network-Layer Multicast is Not Always Efficient At Datalink Layer", Work in Progress, Internet-Draft, draft-vyncke-6man-mcast-not-efficient-01, 14 February 2014, <https://tools.ietf.org/html/draft-vyncke-6man-mcast-not-efficient-01>.

[I-D.ietf-6tisch-architecture]
Thubert, P., "An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4", Work in Progress, Internet-Draft, draft-ietf-6tisch-architecture-29, 27 August 2020, <https://tools.ietf.org/html/draft-ietf-6tisch-architecture-29>.

[MCAST PROBLEMS]
Perkins, C., McBride, M., Stanley, D., Kumari, W., and J. Zuniga, "Multicast Considerations over IEEE 802 Wireless Media", Work in Progress, Internet-Draft, draft-ietf-mboned-ieee802-mcast-problems-12, 26 October 2020, <https://tools.ietf.org/html/draft-ietf-mboned-ieee802-mcast-problems-12>.

[SAVI] Bi, J., Wu, J., Wang, Y., and T. Lin, "A SAVI Solution for WLAN", Work in Progress, Internet-Draft, draft-bi-savi-wlan-20, 14 November 2020, <https://tools.ietf.org/html/draft-bi-savi-wlan-20>.

[UNICAST AR] Thubert, P. and E. Levy-Abegnoli, "IPv6 Neighbor Discovery Unicast Lookup", Work in Progress, Internet-

Draft, draft-thubert-6lo-unicast-lookup-00, 25 January 2019, <https://tools.ietf.org/html/draft-thubert-6lo-unicast-lookup-00>.

[DAD APPROACHES] Nordmark, E., "Possible approaches to make DAD more robust and/or efficient", Work in Progress, Internet-Draft, draft-nordmark-6man-dad-approaches-02, 19 October 2015, <https://tools.ietf.org/html/draft-nordmark-6man-dad-approaches-02>.

[IEEE Std. 802.15.4] IEEE standard for Information Technology, "IEEE Std. 802.15.4, Part. 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks".

[IEEE Std. 802.11] IEEE standard for Information Technology, "IEEE Standard for Information technology -- Telecommunications and information exchange between systems Local and metropolitan area networks-- Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications".

[IEEEstd802151] IEEE standard for Information Technology, "IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements. - Part 15.1: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs)".

[IEEEstd802154] IEEE standard for Information Technology, "IEEE Standard for Local and metropolitan area networks -- Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs)".

[IEEE Std. 802.1] IEEE standard for Information Technology, "IEEE Standard for Information technology -- Telecommunications and information exchange between systems Local and metropolitan area networks Part 1: Bridging and Architecture".

Author's Address

Pascal Thubert (editor)
Cisco Systems, Inc
Building D
45 Allee des Ormes - BP1200
06254 Mougins - Sophia Antipolis
France

Phone: +33 497 23 26 34
Email: pthubert@cisco.com