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 October 28, 2019.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.
1. Introduction

2. IP Models

2.1. Physical Broadcast Domain

At the physical (PHY) Layer, a broadcast domain is the set of all peers that may receive a datagram that one sends over an interface. 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 legacy Ethernet shared wire.

On WLAN and WPAN radios, the physical broadcast domain is defined by a particular transmitter, as the set of nodes that can receive what this transmitter is sending. Literally every datagram defines its own broadcast domain since the chances of reception of a given datagram are statistical. In average and in stable conditions, the broadcast domain of a particular node can 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 2 nodes if their physical broadcast domains overlap.
On WLAN and WPAN radios, this property is usually reflexive, meaning that if B can receive a datagram from A, then A can receive a datagram 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 unidirectional, e.g., A to B but not 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 it can not be assumed in the general case. In other words, the property of radio connectivity is generally not transitive, meaning that A may talk to B and B may talk to C does not necessarily imply that A can talk to C.

We define MAC-Layer Direct Broadcast (DMC) a transmission mode where the broadcast domain that is usable at the MAC layer is directly the physical broadcast domain. IEEE 802.15.4 [IEEE802154] and IEEE 802.11 [IEEE80211] OCB (for Out of the Context of a BSS) are examples of DMC radios. This constrasts with a number of MAC-layer Broadcast Emulation schemes that are described in the next section.

2.2. MAC-Layer Broadcast Emulations

While a physical broadcast domain is constrained to a single shared wire, Ethernet Briding 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 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 mapping server. 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.

An IEEE Std 802.11 Infrastructure Basic Service Set (BSS) also provides a closure of nodes as defined by the broadcast domain of a central Access Point (AP). The AP relays both unicast and broadcast packets and ensures a reflexive and transitive emulation of the shared wire between the associated nodes, with the capability to signal link-up/link-down to the upper layer. Within an Infrastructure 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. For that reason, special efforts are made to ensure...
that all nodes in the BSS receive the broadcast transmission. To achieve this, the transmission is sent at the highest power and slowest PHY speed. This translates into maximum co-channel interferences for others and longest occupancy of the medium, for a duration that can be 100 times that of a unicast. For that reason, upper layer protocols should tend to avoid the use of broadcast when operating over Wi-Fi.

In an IEEE Std 802.11 Infrastructure Extended Service Set (ESS), the process of the association also prepares a bridging state proactively at the AP, so as to avoid the reactive broadcast lookup that takes place in the process of transparent bridging over a spanning tree. This model provides a more reliable operation than the reactive transparent bridging and avoid the need of multicast, and it is only logical that IPv6 ND evolved towards proposes similar methods at Layer-3 for its operation.

In some cases of WLAN and WPAN radios, a mesh-under technology (e.g., an IEEE 802.11s or IEEE 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 Layer-2. When operating on a single frequency, this operation is known to interfere with itself, forcing deployment to introduce delays that dampen the collisions. All in all, the mechanism is slow, inefficient and expensive.

Going down the list of cases above, the cost of a broadcast transmission becomes increasingly expensive, and there is a push to rethink the upper-layer protocols so as to reduce the dependency on broadcast operations.

There again, a MAC-layer communication can be established between 2 nodes if their MAC-layer broadcast domains overlap. In the absence of a MAC-layer emulation such as a mesh-under or an Infrastructure BSS, the MAC-layer broadcast domain is congruent with that of the PHY-layer and inherits its properties for reflexivity and transitivity. IEEE 802.11p, which operates Out of the Context of a BSS (DMC radios) is an example of a network that does not have a MAC-Layer broadcast domain emulation, which means that it will exhibit mostly reflexive and mostly non-transitive transmission properties.

2.3. 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 IPv6 Neighbor Discovery (ND) Duplicate Address Detection (DAD) process leverages a multicast transmission to ensure that an IPv6 address is unique as long as the owner of the
address is connected to the broadcast domain. It must be noted that in all the cases in this specification, the Layer-3 multicast operation is always a MAC_Layer broadcast for the lack of a Layer-2 multicast operation that could handle a possibly very large number of groups in order to make the unicast efficient. This means that for every multicast packet regardless of the destination group, all nodes will receive the packet and process it all the way to Layer-3.

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 by providing a MAC-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) 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 Layer-2 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 MAC-Layer broadcast domain. Relying on Multicast for the ND operation remains feasible but becomes detrimental to unicast traffic, energy-inefficient and unreliable, and its use is discouraged.

On DMC 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 on the broadcast domain defined by one radio transmitter if that transmitter keeps meeting new peers on the go. The nodes may need to form new LLAs to talk to one another and the scope where LLA uniqueness can be dynamically checked is that pair of nodes. As long as there’s no conflict a node may use the same LLA with multiple peers but it has to revalidate DAD with every new peer node. In practice, each pair of nodes defines a temporary P2P link, which can be modeled as a sub-interface of the radio interface.

Wireless Neighbor Discovery (WiND) [RFC6775][RFC8505][I-D.ietf-6lo-backbone-router][I-D.ietf-6lo-ap-nd] defines a new ND operation that derives from the IEEE 802.11 Infrastructure mode. For LLAs, DAD is performed between communicating pairs of nodes. It is carried out as part of a registration process that is based on a NS/NA exchange that transports an Extended Address Registration Option (EARO). During that process, the DAD is validated and a Neighbor Cache Entry (NCE) is populated with a single unicast exchange.
2.4. Mapping the IPv6 Subnet Abstraction

IPv6 also defines a concept of Subnet for Glocal and Unique Local Addresses. Addresses in a same Subnet share a same prefix and by extension, a node belongs to a Subnet if it has an interface with an address on that Subnet. A Subnet prefix is Globally Unique so it is sufficient to validate that an address that is formed from a Subnet prefix is unique within that Subnet to guarantee that it is globally unique. IPv6 aggregation 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 MAC address and deliver the packet, or route the packet to the destination within the Subnet. If the Subnet is known as onlink, then any node may also resolve the destination MAC address and deliver the packet, but if the Subnet is not onlink, then a host that does not have an NCE for the destination will need to pass the packet to a router.

On IEEE Std. 802.3, a Subnet is often congruent with an IP Link because both are determined by the physical attachment to an Ethernet shared wire or an IEEE Std. 802.1 bridged broadcast domain. In that case, the connectivity over the Link is transitive, the Subnet can appear as onlink, and any node can resolve a destination MAC address of any other node directly using IPv6 Neighbor Discovery.

But an IP Link and an IP Subnet are not always congruent. In a shared Link situation, a Subnet may encompass only a subset of the nodes connected to the Link. In Route-Over Multi-Link Subnets (MLSN), routers federate the Links between nodes that belong to the Subnet, the Subnet is not onlink and it extends beyond any of the federated Links.

The DAD and lookup procedures in IPv6 ND expects 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 only applicable for P2P and transit links, and requires extensions for other topologies.

WiND extends IPv6 ND for Hub-and-Spoke MLSN (e.g., a central and some peripheral nodes in Bluetooth low energy (BTLE) [RFC7668]) and Route-Over MLSN (e.g., [I-D.ietf-6tisch-architecture] leveraging [RFC6550]) topologies.

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 6LR, the Hub advertises the prefix as not-onlink to the
spokes in RA messages Prefix Information Options (PIO). Acting as 6LNs, the Spokes autoconfigure addresses from that prefix and register them to the Hub with a corresponding lifetime. Acting as a 6LBR, 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. Acting as 6LR, the Hub also maintains an NCE for the registered addresses and can deliver a packet to any of them for their respective lifetimes. It can be observed that this design builds a form of Layer-3 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 6LBR 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 [draft unaware leaf]. In a degraded mode, all the Hubs are connected to a same high speed backbone such as an Ethernet bridging domain where IPv6 ND is operated. In that case, it is possible to federate the Hub, Spoke and Backbone nodes as a single Subnet, operating IPv6 ND proxy operations [I-D.ietf-6lo-backbone-router] at the Hubs, acting as 6BBRs. It can be observed that this latter design builds a form of Layer-3 Infrastructure ESS.

3. IPv6 Over Wireless

3.1. 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.

3.2. 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. Even if the knowledge of IPv6 addresses used by a STA can be obtained by snooping protocols such as IPv6 ND and DHCPPv6, or by observing data traffic sourced at the STA, such methods provide only an imperfect knowledge of the state of the STA at the AP. This may result in a loss of connectivity for some IPv6 addresses, in particular for addresses rarely used and in a situation of mobility. This may also result in undesirable remanent state in the AP when a STA ceases to use an IPv6 address. It results
that snooping protocols is not a recommended technique and that it should only be used as last resort.

The recommended alternate is to use the IPv6 Registration method specified in p. By that method, 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 one or more IPv6 addresses, using an Address Registration Option. The Registration state has a lifetime that limits unwanted state remanence in the network. The registration is optionally secured using [I-D.ietf-6lo-ap-nd] to prevent address theft and impersonation. The registration carries a sequence number, which enables a fast mobility without a 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 proxy ND specification associated to the address registration is [I-D.ietf-6lo-backbone-router]. 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 proxy replies to NS lookups with the MAC address of the STA, and then bridges packets to the STA normally; as a routing proxy, it 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.

3.3. Case of Mesh Under Technologies

The Mesh-Under provides a broadcast domain emulation with reflexive 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 an Access Point does in a BSS/ESS. While it is still possible to operate IPv6 ND, the inefficiencies of the flooding operation make the IPv6 ND operations even less desirable than in a BSS, and the use of WiND is highly recommended.

3.4. Case of DMC radios

IPv6 over DMC radios uses P2P Links that can be formed and maintained when a pair of DMC radios transmitters are in range from one another.
3.4.1. Using IPv6 ND only

By definition, DMC radios uses IEEE Std. 802.11 transmissions but does not provide the BSS functions.

It is possible to form P2P IP Links between each individual pairs of nodes and operate IPv6 ND 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 IPv6 ND.

The model can be stretched beyond the scope of IPv6 ND if an external mechanism avoids duplicate addresses and if the deployment ensures the connectivity between peers. This can be achieved for instance in a Hub-and-Spoke deployment if the Hub is the only router in the Subnet and the Prefix is advertised as not onlink.

3.4.2. Using Wireless ND

Though this can be achieved with IPv6 ND, WiND is the recommended approach since it uses more unicast communications which are more reliable and less impacting for other users of the medium.

Router and Hosts respectively send a compressed RA/NA with a SLLAO at a regular period. The period can be indicated in a RA as in an 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. The NA/RA is sent over an Unspecified Link where it does not conflict to anyone, so DAD is not necessary at that stage.

The receiver instantiates a Link where the sender’s address is not a duplicate. To achieve this, it forms an LLA that does not conflict with that of the sender and registers to the sender using [RFC 8505]. If the sender sent an RA(PIO) the receiver can also autoconfigure an address from the advertised prefix and register it.
6LoWPAN Node (RPL leaf) 6LR (router)

<table>
<thead>
<tr>
<th>LLN link</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv6 ND RS</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>IPv6 ND RA</td>
</tr>
<tr>
<td>&lt;-----------------</td>
</tr>
<tr>
<td>NS(EARO)</td>
</tr>
<tr>
<td>&lt;-----------------</td>
</tr>
<tr>
<td>NA(EARO)</td>
</tr>
</tbody>
</table>

Figure 1: Initial Registration Flow

The lifetime in the registration should start with a small value (X=RMin, TBD), and exponentially grow with each reregistration 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 -lifetime not elapsed- registrations 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 described above. 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].
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.

4. IANA Considerations

This specification does not require IANA action.
5. Security Considerations

This specification refers to the security sections of IPv6 ND and WiND, respectively.

6. Acknowledgments

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

7. References

7.1. Normative References

[I-D.ietf-6lo-ap-nd]

[I-D.ietf-6lo-backbone-router]


7.2. Informative References


[IEEE802154] 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".


Author’s Address

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

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