INTERNET-DRAFT
September 15, 1995

Thomas Narten, IBM
Erik Nordmark, Sun Microsystems
W A Simpson, Daydreamer

# Neighbor Discovery for IP Version 6 (IPv6)

<draft-ietf-ipngwg-discovery-02.txt>

## Status of this Memo

This Internet Draft expires March 15, 1996.

## Abstract

This document specifies the Neighbor Discovery protocol for IP Version 6. IPv6 nodes on the same link use Neighbor Discovery to discover each other's presence, to determine each other's link-layer addresses, to find routers and to maintain reachability information about the paths to active neighbors.

Contents

## 1. INTRODUCTION

This specification defines the Neighbor Discovery (ND) protocol for Internet Protocol Version 6 (IPv6). Nodes (hosts and routers) use Neighbor Discovery to determine the link-layer addresses for neighbors known to reside on attached links and to quickly purge cached values that become invalid. Hosts also use Neighbor Discovery to find neighboring routers that are willing to forward packets on their behalf. Finally, nodes use the protocol to actively keep track of which neighbors are reachable and which are not, and to detect changed link-layer addresses. When a router or the path to a router fails, a host actively searches for functioning alternates.

This document is a revision of <draft-ietf-ipngwg-discovery-01.txt>
which was itself based on the protocol specified in the two documents:
     <draft-simpson-ipv6-discov-formats-02.txt>, and
     <draft-simpson-ipv6-discov-process-02.txt>

The authors would like to acknowledge the contributions the IPNGWG working group an, in particular, (in alphabetical order) Ran Atkinson, Jim Bound, Scott Bradner, Stephen Deering, Robert Hinden, Allison Mankin, Dan McDonald, Charles Perkins, and Sue Thomson.

## 2. TERMINOLOGY

### 2.1. General

IP - Internet Protocol Version 6. The terms IPv4 and IPv6 are used only in contexts where necessary to avoid ambiguity.

ICMP - Internet Message Control Protocol for the Internet Protocol Version 6. The terms ICMPv4 and ICMPv6 are used only in contexts where necessary to avoid ambiguity.

node - a device that implements IP.

router - a node that forwards IP packets not explicitly addressed to itself.

host - any node that is not a router.

upper layer - a protocol layer immediately above IP. Examples are transport protocols such as TCP and UDP, control protocols such as ICMP, routing protocols such as OSPF,

and internet or lower-layer protocols being "tunneled" over (i.e., encapsulated in) IP such as IPX, AppleTalk, or IP itself.

link - a communication facility or medium over which nodes can communicate at the link layer, i.e., the layer immediately below IP. Examples are Ethernets (simple or bridged); PPP links; X.25, Frame Relay, or ATM networks; and internet (or higher) layer "tunnels", such as tunnels over IPv4 or IPv6 itself.

interface - a node's attachment to a link.

neighbors - nodes attached to the same link.

address - an IP-layer identifier for an interface or a set of interfaces.

anycast address
- an identifier for a set of interfaces (typically belonging to different nodes). A packet sent to an anycast address is delivered to one of the interfaces

identified by that address (the "nearest" one, according to the routing protocol's measure of distance).  See [ADDR-ARCH].

link-layer address
- a link-layer identifier for an interface.  Examples include IEEE 802 addresses for Ethernet links and E.164 addresses for ISDN links.

on-link - an address that is assigned to an interface on a specified link.  A node considers an address to be on-link if:

- it is covered by one of the link's prefixes, or
- a neighboring router specifies the address as the target of a Redirect message, or
- a Neighbor Advertisement message is received for the (target) address, or
- a Neighbor Discovery message is received from the address.

off-link - the opposite of "on-link"; an address that is not assigned to any interfaces on the specified link.

reachability

- whether or not the one-way "forward" path to a neighbor is functioning properly.  In particular, whether packets sent to a neighbor are reaching the IP layer on the neighboring machine and are being processed properly by the receiving layer.  For neighboring routers, reachability means that packets sent by a node's IP layer are delivered to the router's IP layer, and the router is indeed forwarding packets (i.e., it is configured as a router, not a host).  For hosts, reachability means that packets sent by a node's IP layer are delivered to the neighbor host's IP layer.

packet - an IP header plus payload.

link MTU - the maximum transmission unit, i.e., maximum packet size in octets, that can be conveyed in one piece over a link.

target - an address about which address resolution information is sought, or an address which is the new first-hop when being redirected.

proxy - a router that responds to Neighbor Discovery query messages on behalf of another node. A router acting on behalf of a mobile node that has moved off-link potentially acts as a proxy for the mobile node.

ICMP destination unreachable indication
- an error indication returned to the original sender of a packet that cannot be delivered for the reasons outlined in [ICMPv6]. If the error occurs on a node other than the node originating the packet, an ICMP error message is generated. If the error occurs on the originating node, an implementation is not required to actually create and send an ICMP error packet to the source, as long as the sender is notified through an appropriate mechanism (e.g., return value from a procedure call). Note, however, that an implementation may find it convenient in some cases to return errors to the sender by taking the offending packet, generating an ICMP error message, and then delivering it (locally) through the generic error handling routines.

## 2.2. Link Types

Different link layers have different properties. The ones of concern to Neighbor Discovery are:

multicast - a link that supports some mechanism at the link layer for sending packets to all (i.e. broadcast) or a subset of all neighbors. Multicast/broadcast can be provided by a variety of link layer mechanisms such as the physical link layer itself (for example, Ethernet), replicated unicast packets sent by the link layer software, or multicast servers (such as in ATM). Note that all point-to-point links are multicast links.

point-to-point - a link that connects exactly two interfaces. A point-to-point link is assumed to have multicast capability and have a link-local address.

non-broadcast multi-access (NBMA)
- a link to which more than two interfaces can attach, but that does not support any form of multicast or broadcast (e.g., X.25).

shared media - a link that allows direct communication among a number of nodes, but attached nodes are configured in such a way that they do not have complete prefix information for all on-link destinations. That is, at the IP level, nodes on the same link may not know that they are neighbors; by default, they communicate through a router. Examples are large (switched) public data networks such as SMDS and B-ISDN. Also known as "large clouds". See [SH-MEDIA].

variable MTU - a link that does not have a well-defined MTU (e.g., IEEE 802.5 token rings). Many links (e.g., Ethernet) have a standard MTU defined by the link-layer protocol.

asymmetric reachability
- a link where non-reflexive and/or non-transitive reachability is part of normal operation. (Non-reflexive reachability means packets from A reach B but packets from B don't reach A. Non-transitive reachability means packets from A reach B, and packets from B reach C, but packets from A don't reach C.) Many radio links exhibit these

properties.

## 2.3. Addresses

Neighbor Discovery makes use of a number of different addresses defined in [ADDR-ARCH], including:

all-nodes multicast address
- the link-local scope address to reach all nodes. FF02::1

all-routers multicast address
- the link-local scope address to reach all routers. FF02::2

solicited-node multicast address
- a link-local scope multicast address that is computed as a function of the solicited target's address. The solicited-node multicast address is formed by taking the low-order 32 bits of the target IP address and appending those bits to the 96-bit prefix FF02:0:0:0:0:1 to produce a multicast address within the range FF02::1:0:0 to FF02::1:FFFF:FFFF. For example, the solicited node multicast address corresponding to the IP address 4037::01:800:200E:8C6C is FF02::1:200E:8C6C. IP addresses that differ only in the high-order bits, e.g. due to multiple high-order prefixes associated with different providers, will map to the same solicited-node address thereby reducing the number of multicast addresses a node must join.

link-local address
- a unicast address having link-only scope that can be used to reach neighbors. All interfaces MUST have a link-local address. Routers MUST NOT forward packets with a link-local source address. See [ADDR-ARCH].

unspecified address
- a reserved address value that indicates the lack of an address (e.g., the address is unknown). It is never used as a destination address, but may be used as a source address if the sender does not (yet) know its own address (e.g., while verifying an address is unused during address autoconfiguration [ADDRCONF]). The unspecified address has a value of 0:0:0:0:0:0:0:0.

## 2.4. Requirements

Throughout this document, the words that are used to define the significance of the particular requirements are capitalized. These words are:

MUST
This word or the adjective "REQUIRED" means that the item is an absolute requirement of this specification.

MUST NOT
This phrase means the item is an absolute prohibition of this specification.

SHOULD
This word or the adjective "RECOMMENDED" means that there may exist valid reasons in particular circumstances to ignore this item, but the full implications should be understood and the case carefully weighed before choosing a different course.

SHOULD NOT
This phrase means that there may exist valid reasons in particular circumstances when the listed behavior is acceptable or even useful, but the full implications should be understood and the case carefully weighted before implementing any behavior described with this label.

MAY This word or the adjective "OPTIONAL" means that this item is truly optional. One vendor may choose to include the item because a particular marketplace requires it or because it enhances the product, for example, another vendor may omit the same item.

## 3. PROTOCOL OVERVIEW

This protocol solves a set of problems related to the interaction between nodes attached to the same link. It defines mechanisms for solving each of the following problems:

Router Discovery: How hosts locate routers that reside on an attached link.

Prefix Discovery: How hosts discover the set of address prefixes that define which destinations are on-link for an attached link. (Nodes use prefixes to distinguish destinations that reside on-link from those only

reachable through a router.)

Parameter Discovery: How a node learns such link parameters as the link MTU or such Internet parameters as the maximum hop limit value to place in outgoing packets.

Address Autoconfiguration: How nodes automatically configure an address for an interface.

Address Resolution: How nodes determine the link-layer address of an on-link destination (e.g., a neighbor) given only the destination's IP address.

Next-hop determination: The algorithm for mapping an IP destination address into the IP address of the neighbor to which traffic for the destination should be sent. The next-hop can be a router or the destination itself.

Neighbor Unreachability Detection: How nodes determine that a neighbor is no longer reachable. For neighbors used as routers, alternate default routers can be tried. For both routers and hosts, address resolution can be performed again.

Duplicate Address Detection: How a node determines that an address it wishes to use is not already in use by another node.

Redirect: How a router informs a host of a better first-hop node to reach a particular destination.

Neighbor Discovery defines five different ICMP packet types: A pair of Router Solicitation and Router Advertisement messages, a pair of Neighbor Solicitation and Neighbor Advertisements messages, and a Redirect message. The messages serve the following purpose:

Router Solicitation: When an interface becomes enabled, hosts may send out Router Solicitations that request routers to generate Router Advertisements immediately rather than at their next scheduled time.

Router Advertisement: Routers advertise their presence together with various link and Internet parameters either periodically, or in response to an explicit Router Solicitation message. Router Advertisements contain prefixes that are used for on-link determination and/or address configuration, a Maximum Hop Limit value, etc.

Neighbor Solicitation: Sent by a node to determine the link-layer address of a neighbor, or to verify that a neighbor is still reachable via a cached link-layer address. Neighbor Solicitations are also used for Duplicate Address Detection.

Neighbor Advertisement: A response to a Neighbor Solicitation message. A node may also send unsolicited Neighbor Advertisements to announce a link-layer address change.

Redirect: Used by routers to inform hosts of a better first hop for a destination.

On multicast-capable links, each router periodically multicasts a Router Advertisement packet announcing its availability. A host receives Router Advertisements from all routers, building a list of default routers. Routers generate Router Advertisements frequently enough that hosts will learn of their presence within a few minutes, but not frequently enough to rely on an absence of advertisements to detect router failure; a separate Neighbor Unreachability Detection algorithm provides failure detection.

Router Advertisements contain a list of prefixes used for on-link determination and/or autonomous address configuration; flags associated with the prefixes specify the intended uses of a particular prefix. Hosts use the advertised on-link prefixes to build and maintain a list that is used in deciding when a packet's destination is on-link or beyond a router. Note that a destination can be on-link even though it is not covered by any advertised on-link prefix. In such cases a router can send a Redirect informing the sender that the destination is a neighbor.

Router Advertisements (and per-prefix flags) allow routers to inform hosts how to perform Address Autoconfiguration. For example, routers can specify whether hosts should use stateful (DHCPv6) and/or autonomous (stateless) address configuration. The exact semantics and usage of the address configuration-related information is specified in [ADDRCONF].

Router Advertisement messages also contain Internet parameters such as the maximum hop that hosts should use in outgoing packets and, optionally, link parameters such as the link MTU. This facilitates centralized administration of critical parameters that can be set on routers and automatically propagated to all attached hosts.

Nodes accomplish Address Resolution by multicasting a Neighbor Solicitation that asks the target node to return its link-layer address. Neighbor Solicitation messages are multicast to the solicited-node

multicast address of the target address. The target returns its link-layer address in a unicast Neighbor Advertisement message. A single request-response pair of packets is sufficient for both the initiator and the target to resolve each other's link-layer addresses; the initiator includes its IP address and link-layer address in the Neighbor Solicitation.

Neighbor Solicitation messages can also be used to determine if more than one node has been configured to use a particular unicast address. The use of Neighbor Solicitation messages for Duplicate Address Detection is specified in [ADDRCONF].

Neighbor Unreachability Detection detects the failure of a neighbor or the failure of the forward path to the neighbor. Doing so requires positive confirmation that packets sent to a neighbor are actually reaching that neighbor and being processed properly by its IP layer. Neighbor Unreachability Detection uses confirmation from two sources. When possible, upper-layer protocols provide a positive confirmation that a connection is making "forward progress", that is, previously sent data is known to have been delivered correctly (e.g., new acknowledgments were received recently). When positive confirmation is not forthcoming through such "hints", a node sends explicit unicast Neighbor Solicitation messages that solicit Neighbor Advertisements as reachability confirmation from the next hop. To reduce unnecessary network traffic, probe messages are only sent to neighbors to which the node is actively sending packets.

In addition to addressing the above general problems, Neighbor Discovery also handles the following situations:

Link-layer address change - A node that knows its link-layer address has changed can multicast a few (unsolicited) Neighbor Advertisement packets to all nodes to quickly (but unreliably) update cached link-layer addresses that have become invalid. Note that the sending of unsolicited advertisements is a performance enhancement only (e.g., unreliable). The Neighbor Unreachability Detection algorithm ensures that all nodes will reliably discover the new address, though the delay may be somewhat longer.

Inbound load balancing - Nodes with replicated interfaces may want to load balance the reception of incoming packets across multiple network interfaces on the same link. Such nodes have multiple link-layer addresses assigned to the same interface. For example, a single network driver could represent multiple network interface cards as a single logical interface having multiple link-layer addresses. Load balancing is handled by allowing routers to omit the source link-layer address from

Router Advertisement packets, thereby forcing neighbors to use Neighbor Solicitation messages to learn the link-layer addresses. Returned Neighbor Advertisement messages can then contain link-layer addresses that differ depending on who issued the solicitation.

Anycast addresses - Anycast addresses identify one of a set of nodes providing an equivalent service, and multiple nodes on the same link may be configured to recognize the same Anycast address. Neighbor Discovery handles anycasts by having nodes expect to receive multiple Neighbor Advertisements for the same target. All advertisements for anycast addresses are tagged as being "Secondary" advertisements. This invokes specific rules to determine which of potentially multiple advertisements should be used.

Proxy advertisements - A router willing to accept packets on behalf of a target address that is unable to respond to Neighbor

Solicitations can issue Secondary Neighbor Advertisements. There is currently no specified use of proxy, but proxy advertising could potentially be used to handle cases like mobile nodes that have moved off-link. However, it is not intended as a general mechanism to handle nodes that, e.g., do not implement this protocol.

## 3.1. Comparison with IPv4

The IPv6 Neighbor Discovery protocol corresponds to a combination of the IPv4 protocols ARP [ARP], ICMP Router Discovery [RDISC], and ICMP Redirect [ICMPv4]. In IPv4 there is no generally agreed upon protocol or mechanism for Neighbor Unreachability Detection, although Hosts Requirements [HR-CL] does specify some possible algorithms for Dead Gateway Detection (a subset of the problems Neighbor Unreachability Detection tackles).

The Neighbor Discovery protocol provides a multitude of improvements over the IPv4 set of protocols:

Router Discovery is part of the base protocol set; there is no need for hosts to "snoop" the routing protocols.

Router advertisements carry link-layer addresses; no additional packet exchange is needed to resolve the router's link-layer address.

Router advertisements carry prefixes for a link; there is no need

to have a separate mechanism to configure the "netmask".

Router advertisements enable Address Autoconfiguration.

Routers can advertise an MTU for hosts to use on the link, ensuring that all nodes use the same MTU value on links lacking a well-defined MTU.

Address Resolution multicasts are "spread" over 4 billion (2^32) multicast addresses greatly reducing Address Resolution related interrupts on nodes other than the target. Moreover, non-IPv6 machines should not be interrupted at all.

Redirects contain the link-layer address of the new first hop; separate Address Resolution is not needed upon receiving a redirect.

Multiple prefixes can be associated with the same link. By default, hosts learn all on-link prefixes from Router Advertisements. However, routers may be configured to omit some or all prefixes from Router Advertisements. In such cases hosts assume that destinations are off-link and send traffic to routers. A router can then issue redirects as appropriate.

Unlike IPv4, the recipient of an IPv6 redirect assumes that the new next-hop is on-link. In IPv4, a host ignores redirects specifying a next-hop that is not on-link according to the link's network mask. The IPv6 redirect mechanism is analogous to the XRedirect facility specified in [SH-MEDIA]. It is expected to be useful on non-broadcast and shared media links in which it is undesirable or not possible for nodes to know all prefixes for on-link destinations.

Neighbor Unreachability Detection is part of the base, significantly improving the robustness of packet delivery in the presence of failing routers, partially failing or partitioned links and nodes that change their link-layer addresses. For instance, mobile nodes can move off-link without losing any connectivity due to stale ARP caches.

Unlike ARP, Neighbor Discovery detects half-link failures and avoids sending traffic to neighbors with which two-way connectivity is absent.

Placing address resolution at the ICMP layer makes the protocol more media-independent than ARP and makes it possible to use standard IP authentication and security mechanisms as appropriate [IPv6-AUTH, IPv6-ESP].

## 3.2. Supported Link Types

Neighbor Discovery supports links with different properties. In the presence of certain properties only a subset of the ND protocol is available:

point-to-point - Neighbor Discovery handles such links just like multicast links.  (Multicast can be trivially provided on point to point links, and interfaces can be assigned link-local addresses.)

multicast - All aspects of Neighbor Discovery are available.

non-broadcast multiple access (NBMA)
- The only Neighbor Discovery mechanisms available on these links are Redirect handling and Neighbor Unreachability Detection.

If hosts support manual configuration of a list of default routers, the hosts can dynamically acquire the link-layer addresses for their neighbors from Redirect messages.

shared media - The Redirect message is modeled after the XRedirect message in [SH-MEDIA] in order to simplify use of the protocol on shared media links.

This specification does not address shared media issues that only relate to routers, such as:

- How routers exchange reachability information on a shared media link.

- How a router determines the link-layer address of a host, which it needs to send redirect messages to the host.

- How a router determines that it is the first hop router for a received packet.

The protocol is extensible (through the definition of new options) so that other solutions might be possible in the future.

variable MTU - Neighbor Discovery allows routers to specify a MTU for the link, which all nodes then use.  All nodes on a link must use the same MTU (or Maximum Receive

Unit) in order for multicast to work properly. When multicasting, a sender has no way of knowing which nodes will receive the packet, and cannot determine a minimum packet size all receivers can process.

asymmetric reachability
- Neighbor Discovery detects the absence of symmetric reachability; a node avoids paths to a neighbor with which it does not have symmetric connectivity.

The Neighbor Unreachability Detection will typically identify such half-links and the node will refrain from using them.

The protocol can presumably be extended in the future to find viable paths in environments that lack reflexive and transitive connectivity.

## 4. CONCEPTUAL MODEL OF A HOST

This section describes a conceptual model of one possible data structure organization that hosts (and to some extent routers) will maintain in interacting with neighboring nodes. The described organization is provided to facilitate the explanation of how the Neighbor Discovery protocol should behave. This document does not mandate that implementations adhere to this model as long as their behavior is consistent with the protocol specification.

This model is only concerned with the aspects of host behavior directly related to Neighbor Discovery. In particular, it does not concern itself with such issues as source address selection or the selecting of an outgoing interface on a multihomed host.

### 4.1. Conceptual Data Structures

Hosts will need to maintain the following pieces of information about an interface:

Neighbor Cache - A set of entries about individual neighbors to which traffic has been sent recently. Entries are keyed on the neighbor's on-link IP address and contain such information as its link-layer address, a flag indicating whether the neighbor is a router or a host (called "is_router" in this document), a pointer to any queued packets waiting for Address Resolution to

complete, etc.

A Neighbor Cache entry also contains information used by the Neighbor Unreachability Detection algorithm. This includes the reachability state, the number of unanswered probes, and the time the next Neighbor Unreachability Detection event is scheduled to take place.

Destination Cache
- A set of entries for each destination to which traffic has been sent recently. The Destination Cache includes both on-link and off-link destinations and provides a level of indirection into the Neighbor Cache; the Destination Cache maps a destination IP address to the IP address of the next-hop neighbor. Implementations may find it convenient to store additional information not directly related to Neighbor Discovery in Destination Cache entries, such as the Path MTU (PMTU) and round trip timers maintained by transport protocols.

Prefix List
- A list of the prefixes that define a set of addresses that are on-link. Prefix List entries are created from information received in Router Advertisements. Each entry has an associated invalidation timer value (extracted from the advertisement) used to delete prefixes that routers stop advertising. A special "infinity" timer value specifies that a prefix remains valid forever, unless a new (finite) value is received in a subsequent advertisement.

Default Router List
- A list of routers to which packets may be sent. Router list entries point to entries in the Neighbor Cache; the algorithm for selecting a default router favors routers known to be reachable over those whose reachability is suspect. Each entry also has an associated invalidation timer value (extracted from

Router Advertisements) used to delete entries that are no longer advertised.

Note that the above conceptual data structures can be implemented using a variety of techniques. One possible implementation is to use a single longest-match routing table for all of the above data structures. However, in all cases it is important to not duplicate the conceptual

Neighbor Cache entry for a router in order to prevent redundant Neighbor Unreachability Detection probes.

The Neighbor Cache contains information maintained by the Neighbor Unreachability Detection algorithm. A key piece of information is a neighbor's reachability state, which is one of three possible values:

INCOMPLETE Address Resolution is in progress and the link-layer address of the neighbor has not yet been determined.

REACHABLE Roughly speaking, the neighbor is known to have been reachable recently (within tens of seconds ago).

PROBE The neighbor may be reachable, but the last explicit reachability confirmation was received long enough ago that verification is now actively sought.

## 4.2. Conceptual Sending Algorithm

When sending a packet, a node uses a combination of the Destination Cache, the Prefix List, and the Default Router List to determine the IP address of the appropriate next hop, an operation known as "next-hop determination". Once the IP address of the next hop is known, the Neighbor Cache is consulted for link-level information about that neighbor.

Next-hop determination operates as follows for unicast packets. The sender examines the Prefix List to determine whether the packet's destination is on- or off-link. If the destination is on-link, the next-hop address is the same as the packet's destination address. If the destination is off-link, the sender selects a router from the

Default Router List (following the rules described in Section 5.3.4). If the Default Router List is empty, the sender assumes that the destination is on-link.

For multicast packets the next-hop is always the (multicast) destination address.

For efficiency reasons, next-hop determination is not performed on every packet that is sent. Instead, the results of next-hop determination computations are saved in the Destination Cache. When the sending node has a packet to send, it first examines the Destination Cache. If no entry exists for the destination, next-hop determination is invoked to create a Destination Cache entry.

Once the IP address of the next-hop node is known, the sender examines

the Neighbor Cache for link-level information about that neighbor. If no entry exists for a multicast destination, an entry is created using the link specific mapping to a multicast link-layer address (see e.g. [IPv6-ETHER]). If no entry exists for a unicast destination, the sender creates a new one, sets its state to INCOMPLETE, sends a Neighbor Solicitation message, and then queues the data packet pending completion of Address Resolution. When a Neighbor Advertisement response is received, the link-layer addresses is entered in the Neighbor Cache entry and the queued packet is transmitted. The Address Resolution mechanism is described in detail in Section 6.2.

Each time a Neighbor Cache entry is accessed while transmitting a unicast packet, the sender checks Neighbor Unreachability Detection related information according to the Neighbor Unreachability Detection algorithm (Section 6.3), unless the upper-layer has indicated that such checks are not needed. For instance, the Neighbor Discovery protocol itself when sending packets should pass an indication to IP that the packet should not trigger Neighbor Unreachability Detection. This unreachability check might result in the sender transmitting a unicast Neighbor Solicitation to verify that the neighbor is still reachable.

Next-hop determination is done the first time traffic is sent to a destination. As long as subsequent communication to that destination proceeds successfully, the Destination Cache entry continues to be used. If at some point communication ceases to proceed, as determined by the Neighbor Unreachability Detection algorithm, next-hop determination may need to be performed again. For example, traffic through a failed router should be switched to a working router. Likewise, it may be possible to reroute traffic destined for a mobile node to a "mobility agent".

Note that when a node redoes next-hop determination there is no need to discard the complete Destination Cache entry. In fact, it is generally beneficial to retain such cached information as the PMTU and round trip timer values that may also be kept in the Destination Cache entry.

## 4.3. Garbage Collection and Timeout Requirements

The conceptual data structures described above use different mechanisms for discarding potentially stale or unused information.

=46rom the perspective of correctness there is no need to periodically purge Destination and Neighbor Cache entries. Although stale information can potentially remain in the cache indefinitely, the Neighbor Unreachability Detection algorithm ensures that stale information is purged quickly if it is actually being used.

To limit the storage needed for the Destination and Neighbor Caches, a node may need to garbage-collect old entries. However, care must be taken to insure that sufficient space is always present to hold the working set of active entries. A small cache may result in an excessive number of Neighbor Discovery messages if entries are discarded and rebuilt in quick succession. Any LRU-based policy that only reclaims entries that have not been used in some time (e.g., ten minutes or more) should be adequate for garbage-collecting unused entries.

A node should retain entries in the Default Router List and the Prefix List until their lifetimes expire. However, a node may garbage collect entries prematurely if it is low on memory. If not all routers are kept on the Default Router list, a node should retain at least two entries in the Default Router List (and preferably more) in order to maintain robust connectivity for off-link destinations.

When removing an entry from the Default Router List or the Prefix List there is no need to purge any entries from the Destination or Neighbor Caches. Neighbor Unreachability Detection will efficiently purge any entries in these caches that have become invalid.

## 5. ROUTER AND PREFIX DISCOVERY

This section describes message formats, router behavior and host behavior related to the Router Discovery portion of Neighbor Discovery. Router Discovery is used to locate neighboring routers as well as learn prefixes and configuration parameters related to address autoconfiguration.

Prefix Discovery provides a mechanism through which hosts learn of ranges of IP addresses that reside on-link and thus can be reached directly without going through a router.  Routers advertise a set of prefixes that cover those IP addresses that are on-link.  Prefix discovery is logically separate from Router Discovery.  In practice, prefix information is included in options piggybacked on Router Advertisement messages to reduce network traffic.

Address Autoconfiguration information is also logically separate from Router Discovery.  To reduce network traffic, however, autoconfiguration information is piggybacked on Router Discovery messages.  In fact, the same prefixes can be advertised for on-link determination and address autoconfiguration by specifying the appropriate flags in the Prefix Information options.  This document does not define how autoconfiguration information is processed.  See [ADDRCONF] for details.

### 5.1. Message Formats

### 5.1.1. Router Solicitation Message Format

```
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |     Type      |     Code      |          Checksum             |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                            Reserved                           |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |   Options ...
     +-+-+-+-+-+-+-+-+-+-+-+-
```

IP Fields:

Source Address
MUST be the link-local address assigned to the interface from which this message is sent.

Destination Address
The all-routers link-local multicast address.

Hop Count 1

Authentication Header
If a Security Association for the IP Authentication Header exists between the sender and the destination address, then the sender SHOULD include this header.

Routing Header MUST NOT be sent.

ICMP Fields:

Type 133

Code 0

Checksum The ICMP checksum. See [ICMPv6].

Reserved This field is unused. It MUST be initialized to zero by the sender and ignored by the receiver.

Options:

Source link-layer address
The link-layer address for the sender. This option SHOULD be included on link layers that have addresses so that routers responding to the request can unicast a response without the need to first perform address resolution.

Future versions of this protocol may define new option types.

Receivers MUST skip over and ignore any options they do not recognize and continue processing the message.

### 5.1.2. Router Advertisement Message Format

```
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |     Type      |     Code      |          Checksum             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Max Hop Limit |M|O| Reserved |        Router Lifetime         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                         Reachable Time                        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                          Retrans Timer                        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |   Options ...
 +-+-+-+-+-+-+-+-+-+-+-+-
```

IP Fields:

Source Address
MUST be the link-local address assigned to the interface from which this message is sent.

Destination Address
Either the Source Address of an invoking Router Solicitation or the all-nodes link-local multicast address.

Hop Count 1

Authentication Header
If a Security Association for the IP Authentication Header exists between the sender and the destination address, then the sender SHOULD include this header.

Routing Header MUST NOT be sent.

ICMP Fields:

Type 134

Code 0

Checksum The ICMP checksum. See [ICMPv6].

Max Hop Limit 8-bit unsigned integer. The maximum hop limit that the router suggests that hosts use when sending IP packets. A value of zero means unspecified.

M 1-bit "Managed address configuration" flag. Use the administered (stateful) protocol for address autoconfiguration in addition to any addresses autoconfigured using stateless address autoconfiguration. The use of this flag is described in [ADDRCONF].

O 1-bit "Other configuration" flag. Use the administered (stateful) protocol for autoconfiguration of other (non-address) information. The use of this flag is described in [ADDRCONF].

Reserved A 6-bit unused field. It MUST be initialized to zero by the sender and ignored by the receiver.

Router Lifetime

16-bit unsigned integer. The lifetime associated with the default router in units of seconds. The maximum value corresponds to 18.2 hours. A Lifetime of 0 indicates that the router is not a default router and SHOULD NOT appear on the default router list. The Router Lifetime does not apply to information contained in any options in the message. Options that need time limits for their information include their own lifetime fields.

Reachable Time 32-bit unsigned integer. The time, in milliseconds, that a node assumes a neighbor is reachable after receiving some reachability confirmation. Used by the Neighbor Unreachability Detection algorithm (see Section 6.3). A value of zero means unspecified (by the router).

Retrans Timer 32-bit unsigned integer. The time, in milliseconds, between retransmitted Neighbor Solicitation messages. Used by Address Resolution and the Neighbor Unreachability Detection algorithm (see Sections 6.2 and 6.3). A value of zero means unspecified (by the router).

Options:

Source link-layer address
: The link-layer address of the interface from which the Router Advertisement is sent. Only used on link layers that have addresses. A router MAY omit this option in order to enable inbound load sharing across multiple link-layer addresses.

MTU
: SHOULD be sent on links that have a variable MTU. MAY be sent on other links.

Prefix Information
: These options specify the prefixes that are on-link and/or are used for address autoconfiguration. A router SHOULD include all its on-link prefixes so that multihomed hosts have complete prefix information about on-link destinations for the links to which they attach. If complete information is lacking, a multihomed host may not be able to chose the correct outgoing interface when sending traffic to its neighbors.

Future versions of this protocol may define new option types. Receivers MUST skip over and ignore any options they do not recognize and continue processing the message.

## 5.2. Router Specification

### 5.2.1. Router Configuration Variables

A router MUST allow for the following variables to be configured by system management; default values are specified so as to make it unnecessary to configure any of these variables in many cases.

For each multicast interface:

AdvertiseDefault

A flag indicating whether or not the router should advertise itself as a default router on the interface.

Default: TRUE

ManagedFlag The true/false value to be placed in the "Managed address configuration" field in the Router Advertisement. See [ADDRCONF].

Default: FALSE

OtherFlag The true/false value to be placed in the "Other configuration" field in the Router Advertisement. See [ADDRCONF].

Default: FALSE

LinkMTU The value to be placed in MTU options sent by the router. If the value is set to zero no MTU options are sent.

Default: 0

AdvReachableTime

The value to be placed in the Reachable Time field in the Router Advertisement messages sent by the router. The value zero means unspecified (by this router). MUST be no greater than 3,600,000 milliseconds (1 hour).

Default: REACHABLE_TIME milliseconds

ReachableTime The time a neighbor is considered reachable after receiving a reachability confirmation.

Default: If AdvReachableTime is non-zero (specified) a uniformly-distributed random value between MIN_RANDOM_FACTOR and MAX_RANDOM_FACTOR times AdvReachableTime milliseconds. Otherwise, A uniformly-distributed random value between

```
                        MIN_RANDOM_FACTOR and MAX_RANDOM_FACTOR times
                        REACHABLE_TIME milliseconds.

      RetransTimer      The value to be placed in the Retrans Timer field in
                        the Router Advertisement messages sent by the
                        router.  The value zero means unspecified (by this
                        router).

                        Default: RETRANS_TIMER milliseconds

      MaximumHopLimit
                        The value to be placed in the Max Hop Limit field in
                        the Router Advertisement messages sent by the
```

```
                        router.  The value zero means unspecified (by this
                        router).

                        Default: The value specified in the most recent
                        "Assigned Numbers" RFC [ASSIGNED].

      MaxRtrAdvInterval
                        The maximum time allowed between sending unsolicited
                        multicast Router Advertisements from the interface,
                        in seconds.  MUST be no less than 1 second and no
                        greater than 1800 seconds.

                        Default: 600 seconds

      MinRtrAdvInterval
                        The minimum time allowed between sending unsolicited
                        multicast Router Advertisements from the interface,
                        in seconds.  MUST be no less than 0.1 seconds and no
                        greater than .75 * MaxRtrAdvInterval.

                        Default: 0.33 * MaxRtrAdvInterval

      RtrAdvLifetime
                        The value to be placed in the Router Lifetime field
                        of Router Advertisements sent from the interface, in
                        seconds.  MUST be no less than MaxRtrAdvInterval and
                        no greater than 9000 seconds.
```

Note: if AdvertiseDefault is false, the value of RtrAdvLifetime is irrelevant; a Lifetime value of 0 MUST be placed in outgoing Router Advertisements messages so that hosts do not use the router as a default router.

Default: 3 * MaxRtrAdvInterval

PrefixList

A list of prefixes to be placed in Prefix Information options in Router Advertisement messages sent from the interface.

Default: The PrefixList contains all prefixes that the router advertises via routing protocols as being on-link for the interface from which the advertisement is sent.

Each prefix is associated with:

InvalidationLifetime

The value to be placed in the Invalidation Lifetime in the Prefix Information option, in seconds. The designated value of all 1's (0xffffffff) represents infinity.

Default: infinity.

OnLinkFlag

The value to be placed in the on-link flag ("L-bit") field in the Prefix Information option.

Default: TRUE

Automatic address configuration [ADDRCONF] defines additional information associated with each the prefixes:

DeprecationLifetime

The value to be placed in the Deprecation Lifetime in the Prefix Information option, in seconds. The designated value of all 1's (0xffffffff) represents infinity. See [ADDRCONF].

Default: 604800 seconds (7 days)

AutonomousFlag

The value to be placed in the Autonomous Flag field in the Prefix Information option. See [ADDRCONF].

Default: TRUE

Protocol constants are defined in Section 10.

### 5.2.2. Validation of Router Solicitation Messages

A router MUST silently discard any received Router Solicitation messages that do not satisfy all of the following validity checks:

- IP Source Address is a link-local address.

- IP Destination Address is a link-local address or a multicast address with link-local scope.

- IP Routing Header is not present.

- if the message includes an IP Authentication Header, the message authenticates correctly.

- ICMP Checksum is valid.

- ICMP Code is 0.

- ICMP length (derived from the IP length) is 8 or more octets.

- all included options have a length that is greater than zero.

The contents of the Reserved field, and of any unrecognized options, MUST be ignored.  Future, backward-compatible changes to the protocol may specify the contents of the Reserved field or add new options; backward-incompatible changes may use different Code values.

A solicitation that passes the validity checks is called a "valid solicitation".

Routers MUST also validate Router Advertisements as described in Section 5.3.3.

### 5.2.3.  Router Behavior

A router MUST join the all-routers multicast address on all multicast capable interfaces.

The term "advertising interface" refers to any functioning and enabled interface that has at least one IP address assigned to it.  From each advertising interface, the router transmits periodic, multicast Router Advertisements, containing the following values consistent with the message format above:

- In the Router Lifetime field: the interface's configured RtrAdvLifetime.  If the router's AdvertiseDefault flag is set to false, the Router Lifetime field MUST be set to 0.

- In the M and O flags: the interface's configured ManagedFlag and OtherFlag, respectively.  See [ADDRCONF].

- In the Max Hop Limit field: the interface's configured

MaximumHopLimit.

- In the Reachable Time field: the interface's configured AdvReachableTime.

- In the Retrans Timer field: the interface's configured RetransTimer.

- In the options:

  - o Source Link-Layer Address option: link-layer address of the sending interface. This option MAY be omitted to facilitate in-bound load balancing over replicated interfaces.

  - o MTU option: the MTU value that all nodes should be using.

  - o Prefix Information options: one Prefix Information option for each prefix listed in PrefixList with the option fields set from the information in the PrefixList entry as follows:

    - In the "on-link" flag: the entry's OnLinkFlag.

    - In the Invalidation Lifetime field: the entry's InvalidationLifetime.

    - In the "Autonomous address configuration" flag: the entry's AutonomousFlag.

    - In the Deprecation Lifetime field: the entry's DeprecationLifetime.

Router advertisements are not strictly periodic: the interval between subsequent transmissions is randomized to reduce the probability of synchronization with the advertisements from other routers on the same link [SYNC]. Each advertising interface has its own timer. Whenever a multicast advertisement is sent from an interface, that interface's timer is reset to a uniformly-distributed random value between the interface's configured MinRtrAdvInterval and MaxRtrAdvInterval; expiration of the timer causes the next advertisement to be sent from the interface, and a new random value to be chosen. (It is recommended that routers include some unique value, such as one of their IP or link-layer addresses, in the seed used to initialize their pseudo-random number generators. Although the randomization range is configured in units of seconds, the actual randomly-chosen values should not be in units of whole seconds, but rather in units of the highest available timer resolution.)

For the first few advertisements sent from an interface (up to MAX_INITIAL_RTR_ADVERTISEMENTS), if the randomly chosen interval is greater than MAX_INITIAL_RTR_ADVERT_INTERVAL, the timer SHOULD be set to MAX_INITIAL_RTR_ADVERT_INTERVAL instead. Using this smaller interval for the initial advertisements increases the likelihood of a router being discovered quickly when it first becomes available, in the presence of possible packet loss.

In addition to the periodic, unsolicited advertisements, a router sends advertisements in response to valid solicitations received on any of its advertising interfaces. A router MAY choose to unicast the response directly to the soliciting host's address, or multicast it to the all-nodes address; in the latter case, the interface's interval timer is reset to a new random value, as with unsolicited advertisements. A unicast response MAY be delayed, and a multicast response MUST be delayed, for a small random interval not greater than MAX_RTR_RESPONSE_DELAY, in order to prevent synchronization with other responding routers, and to allow multiple, closely-spaced solicitations to be answered with a single multicast advertisement. A router that chooses to delay responses behaves as follows:

- Upon receipt of a Router Solicitation, start a timer taken from a random value within the range 0-MAX_RTR_RESPONSE_DELAY.

- When the timer expires, send out the Router Advertisement. If no other Router Solicitation was received while waiting for the timer to expire, unicast the advertisement. Otherwise, multicast the response and reset the interface timer to a new random value, as is done when multicasting an unsolicited response.

Note that a router is permitted to send multicast Router Advertisements more frequently than indicated by the MinRtrAdvInterval configuration variable if the additional advertisements are responses to explicit solicitations. In all cases, however, unsolicited multicast advertisements MUST NOT be sent more frequently than indicated by MinRtrAdvInterval.

When a router receives a Router Solicitation it records that the source of the packet is a neighbor. If the solicitation contains a Source Link-Layer Address option, and the router has a Neighbor Cache entry for the neighbor, the link-layer address SHOULD be updated in the Neighbor Cache and the entry's "is_router" flag SHOULD be set to false. If a Neighbor Cache entry is created for the source its reachability state MUST be set to PROBE as specified in Section 6.3.2.

It should be noted that an interface may become an advertising interface at times other than system startup, as a result of recovery from an interface failure or through actions of system management such as:

- enabling the interface, if it had been administratively disabled, and its AdvertiseDefault flag is TRUE, or

- enabling IP forwarding capability (i.e., changing the system from being a host to being a router), when the interface's AdvertiseDefault flag is TRUE, or

- changing the AdvertiseDefault flag from FALSE to TRUE.

In such cases the router MUST commence transmission of periodic advertisements on the new advertising interface, limiting the first few advertisements to intervals no greater than MAX_INITIAL_RTR_ADVERT_INTERVAL. In the case of a host becoming a router, the system MUST also join the all-routers IP multicast group on all interfaces on which the router supports IP multicast (whether or not they are advertising interfaces).

An interface may also cease to be an advertising interface, through actions of system management such as:

- administratively disabling the interface, or

- shutting down the system, or disabling the IP forwarding capability (i.e., changing the system from being a router to being a host), or

- setting the AdvertiseDefault flag of the interface to FALSE.

In such cases the router SHOULD transmit a final multicast Router Advertisement on the interface with a Router Lifetime field of zero. In the case of a router becoming a host, the system MUST also depart from the all-routers IP multicast group on all interfaces on which the router supports IP multicast (whether or not they had been advertising interfaces). In addition, the host MUST insure that subsequent Neighbor Advertisement messages sent from the interface have the Router flag set to zero.

The information advertised in Router Advertisements may change through actions of system management. For instance, the lifetime of advertised prefixes may change, the advertised MTU may change, etc. In such cases, the router MAY transmit a few (no more than MAX_INITIAL_RTR_ADVERTISEMENTS) Router Advertisements separated by an interval of MAX_INITIAL_RTR_ADVERT_INTERVAL.

A router might want to send Router Advertisements without advertising itself as being a default router. For instance, a router might advertise prefixes for address autoconfiguration while not wishing to forward packets. Such a router MUST set the Router Lifetime field to zero in its advertisements.

A router MAY choose not to include all Prefix Information options in every Router Advertisement. For example, if prefix lifetimes are much longer than RtrAdvLifetime, including them every few advertisements may be sufficient. However, when responding to a Router Solicitation the router SHOULD transmit all prefixes to allow hosts to quickly discover the prefixes during system initialization.

### 5.2.4. Router Advertisement Consistency

Routers SHOULD inspect valid Router Advertisements sent by other routers on the link and verify that the routers are advertising consistent information. Detected inconsistencies indicate that one or more routers might be misconfigured and SHOULD be logged to system or network management. The minimum set of information that should be checked includes:

- Max Hop Limit values (except for the unspecified value of zero).

- Values of the M or O flags.

- Reachable Time values (except for the unspecified value of zero).

- Retrans Timer values (except for the unspecified value of zero).

- Values in the MTU options.

- Invalidation Lifetimes for the same prefix.

- Deprecation Lifetimes for the same prefix.

Note that it is not an error for different routers to advertise different sets of prefixes.  Also, some routers might leave some fields as unspecified i.e. with the value zero.  The logging of errors SHOULD be restricted to conflicting information that causes hosts to continually switch from one value to another.

In addition, routers can optionally examine the source address of Router Advertisements to determine which of a neighboring router's addresses is its link-local address.

Any other action on reception of Router Advertisement messages by a router is beyond the scope of this document.

### 5.2.5.  Link-local Address Change

The link-local address on a router SHOULD change infrequently.  Nodes receiving Neighbor Discovery messages use the source address to identify the sender.  If multiple packets from the same router contain different source addresses, nodes will assume they come from different nodes, leading to undesirable behavior.  For example, a node will ignore Redirect messages that are believed to have been sent by a router other than the current first-hop router.  Thus the source address used in Router Advertisements must be identical to the target address in a Redirect message when redirecting to the router.

Using the link-local address to uniquely identify routers on the link has the benefit that the link-local address does not change when a site renumbers.

If a router changes the link-local address for one of its interfaces, it SHOULD inform hosts of this change.  The router SHOULD multicast a few Router Advertisements with Router Lifetime field set to zero for the old link-local address and also multicast a few Router Advertisements for the new link-local address.  The exact procedures SHOULD be the same as when an interface ceases being an advertising interface, and when an interface becomes an advertising interface, respectively.

A router MUST be able to determine the link-local address for each of its neighboring routers in order to ensure that the target address in a Redirect message identifies the neighbor router by its link-local address. This may require that routing protocols exchange link-local addresses. Alternatively, routers could listen to Router Advertisements messages to determine link-local addresses of neighboring routers. However, doing so only works if all routers are sending out Router Advertisements.

### 5.3. Host Specification

#### 5.3.1. Host Configuration Variables

None.

#### 5.3.2. Host Variables

A host maintains certain Neighbor Discovery related variables in addition to the data structures defined in Section 4.1. These variables have default values that are overridden by information received in Router Advertisement messages. The default values are used when there is no router on the link, or when all received Router Advertisements have left a particular value unspecified.

For each interface:

LinkMTU The MTU of the link.

Default: The valued defined in the specific document that describe how IPv6 operates over the particular link layer (e.g., [IPv6-ETHER]).

MaximumHopLimit
The maximum Hop Count to be used when sending IP packets.

Default: The value specified in the most recent

```
                     "Assigned Numbers" RFC [ASSIGNED].

      ReachableTime  The time a neighbor is considered reachable after
                     receiving a reachability confirmation.

                     Default: A uniformly-distributed random value
                     between MIN_RANDOM_FACTOR and MAX_RANDOM_FACTOR
                     times REACHABLE_TIME milliseconds.


      RetransTimer   The time between retransmissions of Neighbor
                     Solicitation messages to a neighbor when resolving
                     the address or when probing the reachability of a
                     neighbor.

                     Default: RETRANS_TIMER milliseconds
```

### 5.3.3. Validation of Router Advertisement Messages

A node MUST silently discard any received Router Advertisement messages that do not satisfy all of the following validity checks:

- IP Source Address is a link-local address.

- IP Destination Address is a link-local address or a multicast address with link-local scope.

- IP Routing Header is not present.

- if the message includes an IP Authentication Header, the message authenticates correctly.

- ICMP Checksum is valid.

- ICMP Code is 0.

- ICMP length (derived from the IP length) is 16 or more octets.

- all included options have a length that is greater than zero.

The contents of the Reserved field, and of any unrecognized options, MUST be ignored.  Future, backward-compatible changes to the protocol may specify the contents of the Reserved field or add new options; backward-incompatible changes may use different Code values.

An advertisement that passes the validity checks is called a "valid advertisement".

A host MUST silently discard any received Router Solicitation messages.

### 5.3.4.  Host Behavior

The host joins the all-nodes multicast address on all multicast capable interfaces.

A host MUST NOT send a Router Advertisement message at any time.

To process a valid Router Advertisement, a host extracts the source address of the packet and does the following:

- If the address is not already present in the host's Default Router List, and the advertisement's Router Lifetime is non-zero, create a new entry in the list, and initialize its timer value from the advertisement's Router Lifetime field.

- If the address is already present in the host's Default Router List as a result of a previously-received advertisement, reset its timer to the Router Lifetime value in the newly-received advertisement.

- If the address is already present in the host's Default Router List

and the received Router Lifetime value is zero, time-out the entry immediately and remove it from the Default Router list.

If the received Max Hop Limit value is non-zero the host SHOULD set its MaximumHopLimit variable to the received value.  Hosts use the last Max Hop Limit value they have received; routers should be configured to advertise identical values to avoid hosts switching between different values.

The host SHOULD set its ReachableTime variable based on the Reachable Time field, if the received value is non-zero. The value is computed as a uniformly-distributed random value between MIN_RANDOM_FACTOR and MAX_RANDOM_FACTOR times the value received in the Reachable Time field. Reception of another Router Advertisement causes a new random value to be chosen. This avoids any synchronization of Neighbor Unreachability Detection messages.

The RetransTimer SHOULD be set to the Retrans Timer field, if the received value is non-zero.

Hosts use the last Reachable Time and Retrans Timer values they have received; routers should be configured to advertise identical values to avoid having hosts switch between different values as they receive advertisements from different routers.

After extracting information from the fixed part of the Router Advertisement message, the advertisement MUST be scanned for valid options. If the advertisement contains a source link-layer address option the link-layer address MUST be recorded in the Neighbor Cache entry for the router (creating an entry if necessary) and the "is_router" flag in the Neighbor Cache entry MUST be set to true. The "is_route" flag is used by Neighbor Unreachability Detection to determine when a router changes to being a host (i.e. no longer capable of forwarding packets). If a Neighbor Cache entry is created for the router its reachability state MUST be set to PROBE as specified in Section 6.3.2.

Received MTU options are handled as specified in Section 8.4.

For each Prefix Information option that has the "on-link" (L) flag set, the host does the following:

- If the prefix is not already present in the Prefix List, create a new entry for the prefix and initialize its invalidation timer to the Invalidation Lifetime value in the Prefix Information option.

- If the prefix is already present in the host's Prefix List as the result of a previously-received advertisement, reset its invalidation timer to the Invalidation Lifetime value in the Prefix Information option. If the new Lifetime value is zero, time-out the prefix immediately.

- If the received Invalidation Lifetime value is zero, and the prefix is not present in the host's Prefix List, silently ignore the option.

Note: Implementations can choose to process the on-link aspects of the prefixes separately from the address autoconfiguration aspects of the prefixes by e.g. passing a copy of each valid Router Advertisement message to both an "on-link" and an "addrconf" function. Each function can then operate independently on the prefixes that have the appropriate flag set.

Whenever the invalidation timer expires for a Prefix List entry, that entry is discarded. No existing Destination Cache entries are affected, however.

Whenever a timer expires for an entry in the Default Router List, that entry is discarded. Any entries in the Destination Cache going through that router will continue to be used. Neighbor Unreachability Detection will purge them if appropriate.

To limit the storage needed for the Default Router List, a host MAY choose not to store all of the router addresses discovered via advertisements. However, a host MUST retain at least two router addresses and SHOULD retain more. Default router selections are made whenever communication to a destination appears to be failing. Thus, the more routers on the list, the more likely an alternative working router can be found quickly (e.g., without having to wait for the next advertisement to arrive).

The algorithm for selecting a router depends in part on whether or not a router is known to be reachable. The exact details of how a node keeps track of a neighbor's reachability state are covered in Section 6.3. The algorithm for selecting a default router is invoked only when a Destination Cache entry is incomplete or when communication through an existing router appears to be failing. Under normal conditions, a router would be selected the first time traffic is sent to a destination, with subsequent traffic for that destination using the same router as indicated in the Destination Cache. The policy for selecting routers from the Default Router List is as follows:

1) Routers reachable or probably reachable (e.g., in the REACHABLE or

PROBE state) MUST be preferred over routers whose reachability is unknown or suspect. An implementation may choose to always return the same router or cycle through the router list in a round-robin fashion as long as it always returns a reachable or probably reachable router when one is available.

2) When no routers on the list are known to be reachable or probably reachable, routers SHOULD be selected in a round-robin fashion, so that subsequent requests for a default router do not return the same router until all other routers have been selected.

   Cycling through the router list in this case ensures that all available routers are actively probed by the Neighbor Unreachability Detection algorithm. A request for a default router is made in conjunction with the sending of a packet to a router, and the selected router will be probed for reachability as a side effect.

3) If the Default Router List is empty, assume that the destination is on-link as specified in Section 4.2.

A host is permitted (but not required) to transmit up to MAX_RTR_SOLICITATIONS Router Solicitation messages from any of its multicast interfaces after any of the following events:

- The interface is initialized at system startup time.

- The interface is reinitialized after a temporary interface failure or after being temporarily disabled by system management.

- The system changes from being a router to being a host, by having its IP forwarding capability turned off by system management.

- The host is re-attached to a link after being detached for some time.

The IP destination address of the solicitations is the all-routers multicast address. The IP source address MUST be one of the interface's addresses and MUST be a link-local address. The Source Link-Layer Address option is set to the host's link-layer address.

If a host does choose to send a solicitation after one of the above events, it SHOULD delay that transmission for a random amount of time between 0 and MAX_RTR_SOLICITATION_DELAY. This serves to alleviate congestion when many hosts start up on a link at the same time, such as might happen after recovery from a power failure. (It is recommended

that hosts include some unique value, such as one of their IP or link-layer addresses, in the seed used to initialize their pseudo-random number generators.) Although the randomization range is specified in units of seconds, the actual randomly-chosen values should not be in units of whole seconds, but rather in units of the highest available timer resolution.

If a host has performed a random delay earlier during the system startup (e.g. as part of Duplicate Address Detection [ADDRCONF]) there is no need to randomly delay the first Router Solicitation message.

A host MAY also choose to further postpone its solicitations, subsequent to one of the above events, until the first time it needs to use a default router.

Upon receiving a valid advertisement with a non-zero Lifetime, the host MUST desist from sending any solicitations on that interface (even if none have been sent yet), until the next time one of the above events occurs. The small number of retransmissions of a solicitation, which are permitted if no such advertisement is received, SHOULD be sent at intervals of RTR_SOLICITATION_INTERVAL seconds, without randomization.

# 6. ADDRESS RESOLUTION AND NEIGHBOR UNREACHABILITY DETECTION

This section describes the functions related to the Neighbor Solicitation and Neighbor Advertisement messages and includes descriptions of Address Resolution and the Neighbor Unreachability Detection algorithm.

These messages are also used for Duplicate Address Detection as specified by [ADDRCONF]. In particular, Duplicate Address Detection sends Neighbor Solicitation messages using an unspecified source address targeting its own address. This will generate a multicast Neighbor Advertisement from any node(s) that have been configured with the same address.

## 6.1. Message Formats

### 6.1.1. Neighbor Solicitation Message Format

Nodes send Neighbor Solicitations to request the link-layer address of a target node while also providing their own link-layer address to the target. Neighbor Solicitations are multicast when the node needs to resolve an address and unicast when the node seeks to verify the reachability of a neighbor.

```
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |     Type      |     Code      |          Checksum             |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                           Reserved                            |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                                                               |
     +                                                               +
     |                                                               |
     +                       Sender Address                          +
     |                                                               |
     +                                                               +
     |                                                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                                                               |
     +                                                               +
     |                                                               |
     +                       Target Address                          +
     |                                                               |
     +                                                               +
     |                                                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |   Options ...
     +-+-+-+-+-+-+-+-+-+-+-+-
```

IP Fields:

Source Address
MUST be either the link-local address assigned to the interface from which this message is sent, or the

unspecified address. Use of the unspecified address directs the target node to multicast the resultant Neighbor Advertisement as required by Duplicate Address Detection in [ADDRCONF].

Destination Address
Either the solicited-node link-local multicast address corresponding to the target address, or the target address. Packets unicast to the target address are used to verify reachability.

Hop Count 1

Authentication Header
If a Security Association for the IP Authentication Header exists between the sender and the destination

address, then the sender SHOULD include this header.

Routing Header MUST NOT be sent.

ICMP Fields:

Type 135

Code 0

Checksum The ICMP checksum. See [ICMPv6].

Reserved This field is unused. It MUST be initialized to zero by the sender and ignored by the receiver.

Sender Address
An IP address assigned to the interface from which the solicitation is sent. If the source address of the data packet prompting the solicitation is the same of one of the sending interface's addresses, that address SHOULD be used. Doing so ensures that the receiver of the solicitation places the data packet's source address in its Neighbor Cache, eliminating the need for address resolution in the likely case that reverse traffic for that destination will follow.

Target Address
The IP address of the target of the solicitation. It MUST NOT be a multicast address.

Options:

Source link-layer address
The link-layer address for the sender. MUST NOT be included in unicast solicitations, in order to prevent off-link senders from creating or modifying cached link-layer addresses. For multicast solicitations sent on link layers that have addresses it SHOULD be included.

Future versions of this protocol may define new option types. Receivers MUST skip over and ignore any options they do not recognize and continue processing the message.

### 6.1.2. Neighbor Advertisement Message Format

A node MUST send a Neighbor Advertisement in response to a Neighbor Solicitation for a target IP address that matches an assigned address on the receiving interface. A node MAY also send an unsolicited Neighbor Advertisement if wishes to advertise that its link-layer address has changed.

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Code      |          Checksum             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|S|N|                     Reserved                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                                                               +
|                                                               |
+                       Target Address                          +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Options ...
+-+-+-+-+-+-+-+-+-+-+-+-
```

IP Fields:

Source Address
MUST be the link-local address assigned to the interface from which this message is sent.

Destination Address
The Source Address of an invoking Neighbor Solicitation or, if the source address in the solicitation is the unspecified address, the all-nodes link-local multicast address.  For an unsolicited advertisement the destination is typically the all-nodes link-local multicast address.

Hop Count 1

Authentication Header
If a Security Association for the IP Authentication Header exists between the sender and the destination address, then the sender SHOULD include this header.

Routing Header MUST NOT be sent.

ICMP Fields:

Type 136

Code 0

Checksum The ICMP checksum.  See [ICMPv6].

R Router flag.  When set, the R-bit indicates that the sender is a router.  The R-bit is used by Neighbor Unreachability Detection to detect a router that changes to a host.

S — Solicited flag. When set, the S-bit indicates that the advertisement was sent in response to a Neighbor Solicitation from the Destination address. It MUST be zero in a multicast advertisement and in an unsolicited unicast advertisement.

N — Secondary Advertisement flag. When set, the N-bit indicates that the advertisement should only be used if no other advertisement has been received i.e. the advertisement will not update a cached link-layer address. It SHOULD be set in solicited advertisements for anycast addresses and in solicited proxy advertisements. It SHOULD be zero in other solicited advertisements and in unsolicited advertisements.

Reserved — 29-bit unused field. It MUST be initialized to zero by the sender and ignored by the receiver.

Target Address

The address from the Target Address field in the Neighbor Solicitation message that prompted this advertisement. For an unsolicited advertisement, the address whose link-layer address has changed. The Target Address MUST NOT be a multicast address.

Options:

Target link-layer address

The link-layer address for the target. MUST be included on link layers that have addresses.

Future versions of this protocol may define new option types. Receivers MUST skip over and ignore any options they do not recognize and continue processing the message.

## 6.2. Address Resolution

Address Resolution provides the mechanism through which a node determines the link-layer address of a neighbor. Address Resolution is only used for destinations that are determined to be on-link and for which the sender does not know the corresponding link-layer address. Address resolution is never used for multicast destinations.

### 6.2.1. Node Specification

When a multicast-capable interface is initialized the node MUST join the all-nodes multicast address on that interface, as well as the solicited-node multicast address corresponding to each of the IP addresses assigned to the interface.

The operation of automatic address configuration [ADDRCONF] may, over time, change the set of addresses assigned to an interface; new addresses might be added and old addresses might be removed. In such case the node MUST join and leave the solicited-node multicast address corresponding to the new and old addresses, respectively. Note that multiple assigned addresses might correspond to the same solicited-node multicast address; a node MUST NOT leave the solicited-node multicast group until all assigned addresses corresponding to that multicast address have been removed.

### 6.2.2. Sending Neighbor Solicitations

When a node has a unicast packet to send, but does not know the next-hop's link-layer address, it performs address resolution by creating a Neighbor Cache entry in the INCOMPLETE state and transmitting a Neighbor Solicitation message targeted at the neighbor. The solicitation must be sent to the solicited-node multicast address of the target address.

The sender SHOULD include its link-layer address (if it has one) in the multicast solicitation as a Source Link-Layer Address option. If the source address of the packet prompting the solicitation is the same as one of the addresses assigned to the outgoing interface, that address SHOULD be placed in the ICMP Sender Address of the outgoing solicitation. Otherwise, the interface's link-local address should be used. Using the prompting packet's source address when possible insures that the recipient of the Neighbor Solicitation installs in its Neighbor Cache the IP address that is highly likely to be used in subsequent traffic belonging to the prompting packet's "connection".

While waiting for address resolution to complete, the sender MUST retain packets waiting for address resolution to complete in a small queue.

The queue MUST hold at least one packet, and MAY contain more. However, the number of queued packets per neighbor SHOULD be limited to some small value. When a queue overflows, the new arrival SHOULD replace the oldest entry. Once address resolution completes, all queued packets SHOULD be transmitted.

While awaiting a response, the sender MUST retransmit Neighbor Solicitation messages approximately every RetransTimer milliseconds, even in the absence of additional traffic to the neighbor. Retransmissions MUST be rate-limited for each neighbor to at most one solicitation every RetransTimer milliseconds.

If no advertisement is received after MAX_MULTICAST_SOLICIT solicitations, address resolution has failed. The sender MUST return ICMP destination unreachable indications with code 3 (Address Unreachable) for each packet queued awaiting address resolution.

### 6.2.3. Validation of Neighbor Solicitations

A node MUST silently discard any received Neighbor Solicitation messages that do not satisfy all of the following validity checks:

- IP Source Address is a link-local address or the unspecified address.

- if the IP Destination Address is a multicast address, its scope is link-local.

- IP Routing Header is not present.

- if the message includes an IP Authentication Header, the message authenticates correctly.

- ICMP Checksum is valid.

- ICMP Code is 0.

- ICMP length (derived from the IP length) is 40 or more octets.

- Target Address is not a multicast address.

- if the Source Address is the unspecified address or the Destination Address is a unicast address, there is no Source Link-layer Address option.

- all included options have a length that is greater than zero.

- the Target Address matches an address assigned to the receiving interface.

The contents of the Reserved field, and of any unrecognized options, MUST be ignored. Future, backward-compatible changes to the protocol may specify the contents of the Reserved field or add new options; backward-incompatible changes may use different Code values.

A Neighbor Solicitation that passes the validity checks is called a "valid solicitation".

### 6.2.4. Receipt of Neighbor Solicitations

If and the Source Link-Layer Address option is present, the recipient SHOULD update the Neighbor Cache entries for both the IP Source Address and the ICMP Sender Address of the solicitation. In those cases where a corresponding entry does not already exist, the node SHOULD create a new one and set its reachability state to PROBE as specified in Section 6.3.2. In all cases the source link-layer address option in the received advertisement SHOULD replace any cached link-layer addresses.

A Neighbor Solicitation that is being used for Duplicate Address Detection, i.e. with an unspecified source address, can not contain a source link-layer address option thus it has no effect on the Neighbor Cache.

### 6.2.5. Sending Solicited Neighbor Advertisements

A Neighbor Advertisement is sent in response to a valid Neighbor Solicitation. The Target Address of the advertisement is copied from the Target Address of the Solicitation. The Target Link-Layer Address option SHOULD be included, using as its value the interface's link-layer address. If the node is a router, it MUST set the Router flag to one; otherwise it MUST set the flag to zero.

If the Target Address is either an anycast address or a unicast address for which the node is providing proxy service, the Secondary

Advertisement flag SHOULD be set to one. Otherwise, it SHOULD be set to 0. Proper setting of the Secondary Advertisement flag insures that nodes give preference to "primary" advertisements, even when received after "secondary" advertisements.

If the source of the solicitation is the unspecified address, the node MUST set the Solicited flag to zero and multicast the advertisement to the all-nodes address. Otherwise, the node MUST set the Solicited flag

to one and unicast the advertisement to the link-local Source Address of the solicitation.

### 6.2.6. Validation of Neighbor Advertisements

A node MUST silently discard any received Neighbor Advertisement messages that do not satisfy all of the following validity checks:

- IP Source Address is a link-local address.

- IP Destination Address is a link-local address or a multicast address with link-local scope.

- IP Routing Header is not present.

- if the message includes an IP Authentication Header, the message authenticates correctly.

- ICMP Checksum is valid.

- ICMP Code is 0.

- ICMP length (derived from the IP length) is 24 or more octets.

- Target Address is not a multicast address.

- if the Destination Address is a multicast address the Solicited flag is zero.

- all included options have a length that is greater than zero.

The contents of the Reserved field, and of any unrecognized options, MUST be ignored.  Future, backward-compatible changes to the protocol may specify the contents of the Reserved field or add new options; backward-incompatible changes may use different Code values.

A Neighbor Advertisements that passes the validity checks is called a "valid advertisement".

### 6.2.7.  Receipt of Neighbor Advertisments

When a valid Neighbor Advertisement is received (either solicited or unsolicited), the Neighbor Cache is searched for the target's entry.  If no entry exists, the advertisement SHOULD be silently discarded.  There is no need to create an entry in this case, since the recipient has

apparently not initiated any communication with the target.

Once the appropriate Neighbor Cache entry has been located, the specific actions taken depend on the state of the Neighbor Cache entry.  In particular, if no link-layer address is cached for the target (e.g., it is in the INCOMPLETE state), the first received advertisement would be used.  On the other hand, if we already have a cached link-layer address, we can safely be more selective about what information is used in received advertisements.

If the target's Neighbor Cache entry is in the INCOMPLETE state, the advertisement is the first response to a solicitation.  The receiving node MUST record the link-layer address in the Neighbor Cache entry and send any packets queued for the neighbor awaiting address resolution. If the Solicited flag is set, the reachability state for the neighbor MUST be set to REACHABLE; otherwise it MUST be set to PROBE. (A more detailed explanation of reachability state is described in Section 6.3.2).  The Secondary Advertisement flag is ignored if the entry is in the INCOMPLETE state.

If the target's Neighbor Cache entry is in the REACHABLE or PROBE state, the Secondary Advertisement flag is examined.  If set, the entry's state should be set to PROBE, and the packet SHOULD be silently discarded; no other changes are made to the Neighbor Cache entry.

If the Secondary Advertisement flag is not set, the link-layer address in the Target Link-Layer Address option should be copied into the Neighbor Cache entry. Furthermore, if the Solicited flag is set, the entry's state should be set to REACHABLE. Otherwise, the entry's state should be set to PROBE.

Finally, the receiving node MUST examine the Router flag in the received advertisement and update the "is_router" flag in the Neighbor Cache entry to reflect whether the node is a host or router. In those cases where the neighbor was previously used as a router, but the advertisement's Router flag is now set to zero, the node MUST remove that router from the Default Router List and update the Destination Cache entries for all destinations using that neighbor as a router as specified in Section 6.3.2.

### 6.2.8. Sending Unsolicited Neighbor Advertisements

In some cases a node may be able to determine that its link-layer address has changed (e.g., hot-swap of an interface card) and may wish to inform its neighbors of the new link-layer address quickly. In such cases a node MAY send up to MAX_NEIGHBOR_ADVERTISEMENT unsolicited Neighbor Advertisement messages to the all-nodes multicast address.

These advertisements MUST be separated by at least MIN_NEIGHBOR_ADVERT_INTERVAL seconds.

The Target Address field in the unsolicited advertisement is set to an IP address of the interface, and the Target Link-Layer Address option is filled with the new link-layer address. The Solicited flag MUST be set to zero, in order to avoid confusing the Neighbor Unreachability Detection algorithm. If the node is a router, it MUST set the Router flag to one; otherwise it MUST set it to zero. The Secondary Advertisement flag MAY be either set or cleared. In either case, neighboring nodes will immediately change the state of their Neighbor Cache entries for the Target Address to PROBE, prompting them to verify the path for reachability. If the Secondary Advertisement is set, neighboring nodes will install the new link-layer address in their caches. Otherwise, they will ignore the new link-layer address, choosing instead to probe the cached address instead.

A node that has multiple IP addresses assigned to an interface MAY multicast a separate Neighbor Advertisement for each address. In such a case the node SHOULD introduce a small delay between the sending of each advertisement to reduce the probability of the advertisements being lost due to congestion.

A proxy MAY multicast Neighbor Advertisements when its link-layer address changes or when it is configured (by system management or other mechanisms) to proxy for an address. If there are multiple nodes that are providing proxy services for the same set of addresses the proxies SHOULD provide a mechanism that prevents multiple proxies from multicasting advertisements for any one address, in order to reduce the risk of excessive multicast traffic.

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

Note that because unsolicited Neighbor Advertisements do not reliably update caches in all nodes (the advertisements might not be received by all nodes), they should only be viewed as a performance optimization to quickly update the caches in most neighbors. The Neighbor Unreachability Detection algorithm ensures that all nodes reliably obtain the new link-layer address, though the delay may be slightly longer.

### 6.2.9. Anycast Neighbor Advertisements

A node belonging to an anycast address MUST join the solicited-node multicast address that corresponds to the anycast address.

When a node responds to a Neighbor Solicitation for an anycast address, it MUST respond with an Neighbor Advertisement that has the Secondary Advertisement flag set to one. In addition, the sender should delay sending a response for a random time between 0 and MAX_ANYCAST_DELAY_TIME seconds.

Neighbor Unreachability Detection ensures that a node quickly detects when the current binding for an anycast address becomes invalid.

### 6.2.10. Proxy Neighbor Advertisements

Under limited circumstances, a router MAY proxy for one or more other nodes, that is, through Neighbor Advertisements indicate that it is willing to accept packets not explicitly addressed to itself. For example, a router might potentially accept packets on behalf of a mobile node that has moved off-link. The mechanisms used by proxy are identical to the mechanisms needed for anycast addresses.

A proxy MUST join the solicited-node multicast address(es) that correspond to the IP address(es) assigned to the node for which it is proxying.

All solicited proxy Neighbor Advertisement messages MUST have the Secondary Advertisement flag set to one. This ensures that if the node itself is present on the link its Neighbor Advertisement (with the Secondary flag set to zero) will take precedence of any advertisement received from a proxy. A proxy MAY send unsolicited advertisements with the Secondary Advertisement flag set to zero as specified in Section 6.2.8, but doing so may cause the proxy advertisement to override a valid entry created by the node itself.

Finally, when sending a proxy advertisement in response to a Neighbor Solicitation, the sender should delay its response by a random time between 0 and MAX_ANYCAST_DELAY_TIME seconds.

## 6.3. Neighbor Unreachability Detection

Communication to or through a neighbor may fail for numerous reasons at any time, including hardware failure, hot-swap of an interface card, etc. If the destination has failed, no recovery is possible and communication fails. On the other hand, if it is the path that has failed, recovery may be possible. Thus, a node actively tracks the reachability "state" for the neighbors to which it is sending packets.

Neighbor Unreachability Detection is used for all paths between hosts and neighboring nodes, including host-to-host, host-to-router, and

router-to-host communication. Neighbor Unreachability Detection may also be used between routers, but is not required if an equivalent mechanism is available, for example, as part of the routing protocols. The conceptual model allows an upper-layer to indicate to IP that Neighbor Unreachability Detection is not needed for a packet being sent.

This is used by Neighbor Discovery to skip these checks when sending Neighbor Discovery messages.

When a path to a neighbor appears to be failing, the specific recovery procedure depends on how the neighbor is being used. For example, the specific recovery procedure used when the neighbor is used as a router differs from that used when the neighbor is the destination.

Neighbor Unreachability Detection is performed only for neighbors to which unicast packets are sent; it is not used when sending to multicast addresses.

### 6.3.1. Reachability Confirmation

A neighbor is considered reachable if the node has recently received a confirmation that packets sent recently to the neighbor were received by its IP layer. Positive confirmation can be gathered in two ways: hints from upper layer protocols that indicate a connection is making "forward progress", or receipt of a Neighbor Advertisement message that is a response to an explicit Neighbor Solicitation probe.

A connection makes "forward progress" if the packets received from a remote peer can only be arriving if recent packets sent to that peer are actually reaching it. For example, receipt of a (new) acknowledgement indicates that previously sent data reached the peer. Likewise, the arrival of a new (non-duplicate) packet indicates that earlier acknowledgements are being delivered to the remote peer. If packets are reaching the peer, they must also be reaching the sender's next-hop neighbor; thus "forward progress" is a confirmation that the next-hop neighbor is reachable. For off-link destinations, forward progress implies that the first-hop router is reachable. When available, this upper-layer information SHOULD be used.

In some cases (e.g., UDP-based protocols and routers forwarding packets to hosts) such reachability information may not be readily available from upper-layer protocols. When no hints are available and a node is sending packets to a neighbor, the node actively probes the neighbor using unicast Neighbor Solicitation messages to verify that the forward path is still working.

The receipt of a solicited Neighbor Advertisement that is a response to a Neighbor Solicitation probe serves as reachability confirmation, since advertisements with the Solicited flag set to one are sent only in response to a solicitation. Receipt of other Neighbor Discovery messages such as Router Advertisements and Neighbor Advertisement with the Solicited flag set to zero MUST NOT be treated as a reachability confirmation. Receipt of such unsolicited messages only confirm the one-way path from the neighbor to the recipient node. In contrast, Neighbor Unreachability Detection requires that the forward path from the sender to the neighbor be working. Note that an advertisement sent in response to an explicit solicitation confirms that a path is working in both directions; the solicitation reached the neighbor, prompting it to generate an advertisement, and the advertisement reached the querying node. However, from the perspective of Neighbor Unreachability Detection, only the reachability of the forward path is of interest.

### 6.3.2. Node Behavior

Neighbor Unreachability Detection operates in parallel with the sending of packets to a neighbor. While reasserting a neighbor's reachability, a node continues sending packets to that neighbor using the cached link-layer address.

A Neighbor Cache entry can be in one of three states:

```
   INCOMPLETE  Address resolution is being performed on the entry.
               Specifically, a Neighbor Solicitation has been sent to
               the solicited-node multicast address of the target, but
               the corresponding Neighbor Advertisement has not yet been
               received.

   REACHABLE   Positive confirmation was received within the last
               ReachableTime milliseconds that the forward path to the
               neighbor was functioning properly.  While REACHABLE, no
               special action takes place as packets are sent.

   PROBE More than ReachableTime milliseconds have elapsed since the
               last positive confirmation was received that the forward
               path was functioning properly.  Upon entering the PROBE
               state, no Neighbor Solicitation is sent.  However, a
               timer is set to expire DELAY_FIRST_PROBE_TIME seconds
               later, and a Neighbor Solicitation probe is sent if the
               entry is still in a PROBE state when the timer expires.
               Delaying the sending of the initial Neighbor Solicitation
               gives the upper layers additional time to provide
               reachability confirmation information.  After the initial
               delay, Neighbor Solicitations are retransmitted every
               RetransTimer milliseconds until a reachability
               confirmation is received.
```

When an entry is created as a result of needing to perform address resolution, a Neighbor Solicitation is sent to the solicited-node multicast address of the target, a timer is started to expire RETRANS_TIMER milliseconds later and the entry's state is set to INCOMPLETE.

As specified in Section 6.2.2, when in the INCOMPLETE state, Neighbor Solicitation messages are retransmitted every RETRANS_TIMER milliseconds until a response is received. If no response is received within RETRANS_TIMER milliseconds after sending MAX_MULTICAST_SOLICIT probes to the solicited-node multicast address, address resolution fails. Upon failure, ICMP destination unreachable indications with code 3 (Address unreachable) are returned for any queued packets and the entry is deleted. Note that deleting the entry implies that all destinations using that neighbor must perform next-hop resolution again before sending a subsequent packet. Thus, if the neighbor is a router, an alternate router may be selected. Alternatively, a destination previously thought to be on-link, may now only be reachable through a router.

Unreachability detection changes a neighbor's state from REACHABLE to PROBE only on-demand, as a side effect of sending a data packet to that neighbor. If no traffic is sent to a neighbor, no probes are sent either. Note that an entry may technically no longer be in a REACHABLE state, but the condition need not be checked or acted upon until a packet is sent to the neighbor.

The first time a Neighbor Cache entry is referenced and more than ReachableTime milliseconds have passed since receipt of the last reachability confirmation, its state changes to PROBE. However, no Neighbor Solicitation probe is sent. Probing is deferred for an additional DELAY_FIRST_PROBE_TIME seconds, an optimization that gives the upper-layer protocol additional time to provide a reachability confirmation in those cases where ReachableTime milliseconds have passed since the last confirmation due to lack of recent traffic. Without this optimization the opening of a TCP connection after a traffic lull would initiate probes even though the subsequent three-way handshake would provide a reachability confirmation almost immediately.

If no reachability confirmation is received within DELAY_FIRST_PROBE_TIME seconds after entering the PROBE state, a unicast Neighbor Solicitation message is sent to the neighbor using the cached link-layer address. In addition, the sender starts a timer to retransmit probe messages every RetransTimer milliseconds until the desired solicitation is received. Subsequent probes are retransmitted even if no additional packets are sent to the neighbor. If no response is received after waiting RetransTimer milliseconds after sending the MAX_UNICAST_SOLICIT solicitations, retransmissions cease and the entry

SHOULD be deleted. Subsequent traffic to that neighbor recreates the entry and performs address resolution again.

Note that all Neighbor Solicitations are rate-limited on a per-neighbor basis. A node MUST NOT send Neighbor Solicitations to the same neighbor more frequently than once every RetransTimer milliseconds.

A Neighbor Cache entry also enters the PROBE state when created as a result of receiving packets other than solicited Neighbor Advertisements (e.g., Router Solicitations, Router Advertisements, Redirects, and Neighbor Solicitations). These packets contain the link-layer address of either the sender or, in the case of Redirect, the redirection target. However, receipt of these link-layer addresses does not confirm reachability of the forward-direction path to that node. Placing a newly created Neighbor Cache entry for which the link-layer address is known in the PROBE state provides assurance that path failures are detected quickly. As always, when entering the PROBE state, the first probe is delayed for DELAY_FIRST_PROBE_TIME to give the upper layer some time to provide a reachability confirmation thereby suppressing the sending of a probe.

To detect the case where a router switches from being a router to being a host (e.g., by having its IP forwarding capability turned off by system management), a node MUST compare the Router flag field in all received Neighbor Advertisement messages with the "is_router" flag recorded in the Neighbor Cache entry. When a node detects that a neighbor has changed from being a router to being a host, the node MUST remove that router from the Default Router List and update the Destination Cache so that all entries using that neighbor as a router switch to another router. Note that a router may not be listed in the Default Router List, even though a Destination Cache entry is using it (e.g., the a host was redirected to it).

In some cases, link-specific information may indicate that a path to a neighbor has failed (e.g., the resetting of a virtual circuit). In such cases, link-specific information may be used to purge Neighbor Cache entries before the Neighbor Unreachability Detection would do so. However, link-specific information MUST NOT be used to confirm the reachability of a neighbor; such information does not provide end-to-end confirmation between neighboring IP layers.

## 7.  REDIRECT FUNCTION

This section describes the functions related to the sending and processing of Redirect messages.

### 7.1.  Redirect Message Format

A Redirect packet is sent from a router to a host to inform the host of a better first-hop node on the path to a destination.

```
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |     Type      |     Code      |          Checksum             |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                           Reserved                            |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                                                               |
     +                                                               +
     |                                                               |
     +                       Target Address                          +
     |                                                               |
     +                                                               +
     |                                                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                                                               |
     +                                                               +
     |                                                               |
     +                     Destination Address                       +
     |                                                               |
     +                                                               +
     |                                                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |   Options ...
     +-+-+-+-+-+-+-+-+-+-+-+-
```

IP Fields:

Source Address
MUST be the link-local address assigned to the interface from which this message is sent.

Destination Address
The Source Address of the packet that triggered the redirect.

Hop Count 1

Authentication Header

If a Security Association for the IP Authentication Header exists between the sender and the destination address, then the sender SHOULD include this header.

Routing Header MUST NOT be sent.

ICMP Fields:

Type 5

Code 0

Checksum The ICMP checksum. See [ICMPv6].

Reserved This field is unused. It MUST be initialized to zero by the sender and ignored by the receiver.

Target Address An IP address of the node to which traffic for the Destination SHOULD be sent. When the target is a router, the Target Address MUST be the router's link-local address so that hosts can uniquely identify routers. When the target is the actual endpoint of communication, the target address field MUST contain the same value as the Destination Address field.

Destination Address
The IP address of the destination which is redirected to the target.

Options:

Target link-layer address
The link-layer address for the target. It MUST be included on non-broadcast links, since the host can not use the multicast Neighbor Solicitation to resolve the address. If known by the router, it SHOULD be included on all link layers that have addresses.

Redirected Header
As much as possible of the IP packet that triggered the sending of the Redirect without making the redirect packet exceed 576 octets.

Future versions of this protocol may define new option types. Receivers MUST skip over and ignore any options they do not recognize and continue processing the message.

## 7.2. Router Specification

A router SHOULD send a redirect message, subject to rate limiting, whenever it forwards a packet in which:

- the Source Address field of the packet identifies a neighbor, and
- the router determines that a better first-hop node resides on the same link as the sending node for the Destination Address of the packet being forwarded, and
- the Destination Address of the packet is not a multicast address, and

- the packet is not source routed through the router, i.e. the destination address (when the packet was received by the router) did not match one of the router's addresses. Other source routed packets, not explicitly source routed through the router, can be redirected.

The transmitted redirect packet contains, consistent with the above message format:

- In the Target Address field: the address to which subsequent packets for the destination SHOULD be sent. If the target is a router, that router's link-local address MUST be used. If the target is a host the target address field MUST be set to the same value as the Destination Address field.

- In the Destination Address field: the destination address of the invoking IP packet.

- In the options:

  - o Target Link-Layer Address option: link-layer address of the target, if known.

  - o Redirected Header: as much of the forwarded packet as can fit without the redirect packet exceeding 576 octets in size.

A router MUST limit the rate at which Redirect messages are sent, in order to limit the bandwidth and processing costs incurred by the Redirect messages when the source does not correctly respond to the Redirects, or the source chooses to ignore unauthenticated Redirect messages. More details on the rate-limiting of ICMP error messages can

be found in [ICMPv6].

A router MUST NOT update its routing tables upon receipt of a Redirect.

### 7.3. Host Specification

### 7.3.1. Validation of Redirect Messages

A host MUST silently discard any received Redirect messages that do not satisfy all of the following validity checks:

- IP Source Address is a link-local address.

- IP Routing Header is not present.

- if the message includes an IP Authentication Header, the message authenticates correctly.

- ICMP Checksum is valid.

- ICMP Code is 0.

- ICMP length (derived from the IP length) is 40 or more octets.

- the IP source address of the Redirect is the same as the current first-hop router for the specified destination.

- the Target Address of the redirect is not a multicast address.

- the Destination Address field in the redirect message does not contain a multicast address.

- all included options have a length that is greater than zero.

The contents of the Reserved field, and of any unrecognized options MUST be ignored.  Future, backward-compatible changes to the protocol may specify the contents of the Reserved field or add new options; backward-incompatible changes may use different Code values.

A host MUST NOT consider a redirect invalid just because the Target Address of the redirect is not covered under one of the link's prefixes. That is, identical values in the Target and Destination Address fields indicates that the target destination is on-link.

A redirect that passes the validity checks is called a "valid redirect".

### 7.3.2. Host Behavior

A host receiving a valid redirect SHOULD update its routing information accordingly. When a redirect is received, the host updates the Destination Cache entry for the destination to use to the specified target as the new next-hop. If no Destination Cache entry exists for the destination, such an entry is created (placing it in the PROBE state).

If the redirect contains a Target Link-Layer Address option the host either creates or updates the Neighbor Cache entry for the target. The link-layer address in the Neighbor Cache entry MUST be copied from the Target Link-Layer Address option into the appropriate Neighbor Cache entry. If a Neighbor Cache entry is created for the target its reachability state MUST be set to PROBE as specified in Section 6.3.2. In addition, if the Target Address is the same as the Destination Address, the host MUST treat the destination as on-link and set the "is_router" field in the corresponding Neighbor Cache entry to false. Otherwise it MUST set to true.

A host MAY have a configuration switch that can be set to make it ignore a Redirect message that does not have an IP Authentication header.

A host MUST NOT send Redirect messages.

## 8. OPTIONS

Options provide a mechanism for encoding variable length fields, fields that may appear multiple times in the same packet, or information that is optional and may not appear in all packets. Options can also be used to add additional functionality to future versions of ND.

In order to ensure that future extensions properly coexist with current implementations, all nodes MUST silently ignore any options they do not recognize in received ND packets and continue processing the packet. All options specified in this document MUST be recognized. A node MUST NOT ignore valid options just because the ND message contains unrecognized ones.

The current set of options is defined in such a way that receivers can process multiple options in the same packet independently of each other. In order to maintain these properties future options SHOULD follow the simple rule:

> The option MUST NOT depend on the presence or absence of any other options. The semantics of an option should depend only on the information in the fixed part of the ND packet and on the

information contained in the option itself.

Adhering to the above rule has the following benefits:

1) Receivers can process options independently of one another. For example, an implementation can choose to process the Prefix Information option contained in a Router Advertisement message in a user-space process while the link-layer address in the same message is processed by routines in the kernel.

2) Should the number of options cause a packet to exceed a link's MTU, multiple packets can carry subsets of the options without any change in semantics.

3) Senders MAY send a subset of options in different packets. For instance, if the prefix Invalidation Lifetime is high it might not be necessary to include the Prefix Information option in every Router Advertisement. In addition, different routers might send different sets of options. Thus, a receiver MUST NOT associate any action with the absence of an option in a particular packet. This protocol specifies that receivers should only act on the expiration of timers and on the information that is received in the packets.

When multiple options are present in a Neighbor Discovery packet, they may appear in any order; receivers MUST be prepared to process them independently of their order. There can also be multiple instances of the same option in a message, for instance Prefix Information options.

The length of all options is a multiple of 8 octets, ensuring appropriate alignment without any "pad" options. The fields in the options, as well as the fields in ND packets, are defined to align them on their natural boundaries (e.g. a 16-bit field is aligned on a 16-bit boundary) with the exception of the 128-bit IP addresses/prefixes, which are aligned on a 64-bit boundary.

The link-layer address field contains an uninterpreted octet string; it is aligned on an 8-bit boundary.

All options are of the form:

```
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |     Type      |    Length     |              ...              |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    ~                              ...                              ~
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Fields:

Type — 8-bit identifier of the type of option. The options defined in this document are:

| Option Name | Type |
|---|---|
| Source Link-Layer Address | 1 |
| Target Link-Layer Address | 2 |
| Prefix Information | 3 |
| Redirected Header | 4 |
| MTU | 5 |

Length — 8-bit unsigned integer. The length of the option in units of 8 octets. The value 0 is invalid. Nodes MUST silently discard an ND packet that contains an option with length zero.

The size of an ND packet including the IP header is limited to the link MTU (which is at least 576 octets). When adding options to an ND packet a node MUST NOT exceed the link MTU.

The only ND packets that can potentially exceed the link MTU are Router Advertisements and Redirects; the former due to a large number of Prefix Information options and the latter due to the Redirected Header option.

If there are more Prefix Information options than can fit in a single Router Advertisement packet the router MUST send multiple separate advertisements that each contain a subset of the set of prefixes.

The amount of data to include in the Redirected Header option MUST be limited so that the entire redirect packet does not exceed 576 octets.

## 8.1. Source/Target Link-layer Address

```
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |     Type      |    Length     |    Link-Layer Address ...
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Fields:

```
   Type
                  1 for Source Link-layer Address
                  2 for Target Link-layer Address
```

```
   Length         The length of the option in units of 8 octets.  For
                  example, the length for IEEE 802 addresses is 1
                  [IPv6-ETHER].

   Link-Layer Address
                  The variable length link-layer address.

                  The content and format of this field is expected to be
                  specified in specific documents that describe how IPv6
                  operates over different link layers.  For instance,
                  [IPv6-ETHER].

Description
                  The Source Link-Layer address option contains the
                  link-layer address of the sender of the packet.  It is
                  used in the Neighbor Solicitation, Router
                  Solicitation, and Router Advertisement packets.

                  The Target Link-Layer address option contains the
                  link-layer address of the target.  It is used in
                  Neighbor Advertisement and Redirect packets.
```

## 8.2. Prefix Information

```
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |     Type      |    Length     | Prefix Length |L|A| Reserved1 |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                     Invalidation Lifetime                     |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                     Deprecation Lifetime                      |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                           Reserved2                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    +                                                               +
    |                                                               |
    +                            Prefix                             +
    |                                                               |
    +                                                               +
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Fields:

Type 3

Length 4

Prefix Length 8-bit unsigned integer. The number of leading bits in the Prefix that are valid. The value ranges from 0 to 128.

L 1-bit on-link flag. When set, indicates that this prefix can be used for on-link determination.

A 1-bit autonomous address-configuration flag. When set indicates that this prefix can used for autonomous address configuration as specified in [ADDRCONF].

Reserved1 6-bit unused field. It MUST be initialized to zero by the sender and ignored by the receiver.

Invalidation Lifetime

32-bit unsigned integer. The lifetime of the prefix in seconds for the purpose of on-link determination. A value of all one bits (0xffffffff) represents infinity. This lifetime is also used by [ADDRCONF].

Deprecation Lifetime
32 bits reserved for autonomous address configuration. A value of all one bits (0xffffffff) represents infinity. See [ADDRCONF].

Reserved2 This field is unused. It MUST be initialized to zero by the sender and ignored by the receiver.

Prefix An IP address or a prefix of an IP address. The prefix length field contains the number of valid leading bits in the prefix.

Description

The Prefix Information option is only used in Router Advertisement packets. It provide hosts with on-link prefixes and prefixes for Address Autoconfiguration.

Implementations can choose to process the on-link aspects of the prefixes separately from the address autoconfiguration aspects of the prefixes e.g. by passing a copy of each valid Router Advertisement message to both an "on-link" and an "addrconf" function. Each function can then operate on the prefixes that have the appropriate flag set.

## 8.3. Redirected Header

```
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |     Type      |    Length     |            Reserved           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                           Reserved                            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 ~                       IP header + data                        ~
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Fields:

Type 4

Length         The length of the option in units of 8 octets.

Reserved       These fields are unused.  They MUST be initialized to zero by the sender and ignored by the receiver.

IP header + data

The original packet truncated to ensure that the size of the redirect message does not exceed 576 octets.

Description

The Redirected Header option MUST be included in Redirect packets.

## 8.4.  MTU

```
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |     Type      |    Length     |           Reserved            |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                              MTU                              |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Fields:

Type           5

Length         1

Reserved       This field is unused.  It MUST be initialized to zero by the sender and ignored by the receiver.

MTU            32-bit unsigned integer.  The recommended MTU for the link.

Description

The MTU option SHOULD be included in Router Advertisement packets when the link has no well-known MTU and it MAY be included on links with a well-known MTU.

Hosts MUST handle this option by setting the LinkMTU variable for the interface to the received value.  If the routers on the link are advertising different MTU values this will result in hosts switching between the different MTUs.  Therefore, routers SHOULD verify the consistency between the MTU they and other routers advertise, logging a network management event when contradictory advertisements are detected.

When a host or its interface is initialized the LinkMTU of the interface SHOULD be set to the predefined value for that type of link.  If the host receives no MTU option it MUST continue to use that predefined value.  The MTU option can be used by routers to both increase and decrease the MTU.

In configurations in which heterogeneous technologies are bridged together, the maximum supported MTU may differ from one segment to another.  If the bridges do not generate ICMP Packet Too Big messages, communicating nodes will be unable to use Path MTU to dynamically determine the appropriate MTU on a per-neighbor basis.  In such cases, routers use the MTU option to specify an MTU value supported by all segments.

## 9. MULTIHOMED HOSTS

There are a number of complicating issues that arise when Neighbor Discovery is used by hosts that have multiple interfaces.  This section does not attempt to define the proper operation of multihomed hosts with regard to Neighbor Discovery.  Rather, it identifies issues that require further study.  Implementors are encouraged to experiment with various approaches to making Neighbor Discovery work on multihomed hosts and to report their experiences.

If a multihomed host receives Router Advertisements on all of its

interfaces, it will (probably) have learned on-link prefixes for the addresses residing on each link. When a packet must be sent through a router, however, selecting the "wrong" router can result in a suboptimal or non-functioning path. There are number of issues to consider:

1) In order for a router to send a redirect, it must determine that the packet it is forwarding originates from a neighbor. The standard test for this case is to compare the source address of the packet to the list of on-link prefixes associated with the interface on which the packet was received. If the originating host is multihomed, however, the source address it uses may belong to an interface other than the interface from which it was sent. In such cases, a router will not send redirects, and suboptimal routing is likely. In order to be redirected, the sending host must always send packets out the interface corresponding to the outgoing packet's source address. Note that this issue never arises with non-multihomed hosts; they only have one interface.

2) If the selected first-hop router does not have a route at all for the destination, it will be unable to deliver the packet. However, the destination may be reachable through a router on one of the other interfaces. Neighbor Discovery does not address this scenario; it does not arise in the non-multihomed case.

3) Even if the first-hop router does have a route for a destination, there may be a better route via another interface. No mechanism exists for the multihomed host to detect this situation.

If a multihomed host fails to receive Router Advertisements on one or more of its interfaces, it will not know (in the absence of configured information) which destinations are on-link on the affected interface(s). This leads to a number of problems:

1) If no Router Advertisement is received on any interfaces, a multihomed host will have no way of knowing which interface to send packets out on, even for on-link destinations. Under similar conditions in the non-multihomed host case, a node treats all destinations as residing on-link, and communication proceeds. In the multihomed case, however, additional information is needed to select the proper outgoing interface. Alternatively, a node could attempt to perform address resolution on all interfaces, a step involving significant complexity that is not present in the non-multihomed host case.

2) If Router Advertisements are received on some, but not all interfaces, a multihomed host could choose to only send packets out on the interfaces on which it has received Router Advertisements.

A key assumption made here, however, is that routers on those other interfaces will be able to route packets to the ultimate destination, even when those destinations reside on the subnet to which the sender connects, but has no on-link prefix information. Should the assumption be false, communication would fail. Even if the assumption holds, packets will traverse a sub-optimal path.

## 10. PROTOCOL CONSTANTS

Router constants:

| | |
|---|---|
| MAX_INITIAL_RTR_ADVERT_INTERVAL | 16 seconds |
| MAX_INITIAL_RTR_ADVERTISEMENTS | 3 transmissions |
| MAX_RTR_RESPONSE_DELAY | 6 seconds |

Host constants:

| | |
|---|---|
| MAX_RTR_SOLICITATION_DELAY | 1 second |
| RTR_SOLICITATION_INTERVAL | 3 seconds |
| MAX_RTR_SOLICITATIONS | 3 transmissions |

Node constants:

| | |
|---|---|
| MAX_MULTICAST_SOLICIT | 3 transmissions |
| MAX_UNICAST_SOLICIT | 3 transmissions |
| MAX_ANYCAST_DELAY_TIME | 1 second |
| MAX_NEIGHBOR_ADVERTISEMENT | 3 transmissions |
| MIN_NEIGHBOR_ADVERT_INTERVAL | 16 seconds |
| REACHABLE_TIME | 30,000 milliseconds |
| RETRANS_TIMER | 10,000 milliseconds |
| DELAY_FIRST_PROBE_TIME | 5 seconds |
| MIN_RANDOM_FACTOR | .5 |

```
      MAX_RANDOM_FACTOR                   1.5
```

Additional protocol constants are defined with the message formats in Section 5.1, 6.1, and 7.1.

All protocol constants are subject to change in future revisions of the protocol.

## 11. FUTURE EXTENSIONS

Possible extensions for future study are:

- o Using dynamic timers to be able to adapt to links with widely varying delay. Measuring round trip times, however, requires acknowledgments and sequence numbers in order to match received Neighbor Advertisements with the actual Neighbor Solicitation that triggered the advertisement. Implementors wishing to experiment with such a facility could do so in a backwards-compatible way by defining a new option carrying the necessary information. Nodes not understanding the option would simply ignore it.

- o Adding capabilities to facilitate the operation over links that currently require hosts to register with an address resolution server. This could for instance enable routers to ask hosts to send them periodic unsolicited advertisements. Once again this can be added using a new option sent in the Router Advertisements.

- o Adding additional procedures for links where asymmetric and non-transitive reachability is part of normal operations. Such procedures might allow hosts and routers to find usable paths on, e.g., radio links.

## 12. OPEN ISSUES

- o Should the routers listed in Router Advertisements include a precedence metric? What are the semantics of such metrics (e.g., "router preferences" vs. "default router preferences").

## 13. SECURITY CONSIDERATIONS

Neighbor Discovery is subject attacks that cause IP packets to flow to unexpected places. Such attacks can be used to cause denial of service but also allow nodes to intercept and optionally modify packets destined for other nodes.

The protocol reduces the exposure to such threats in the absence of

authentication by designing ND packets that modify neighbor state (e.g. cached link-layer addresses) in such a way that routers cannot or will not forward them. Limiting the scope of ND packets to a particular link makes the protocol more robust against the accidental sending of ND messages with a hop count larger than one. Specifically:

- the source address of all packets have link-local scope. Routers MUST NOT forward such packets. See [ADDR-ARCH].

- with the exception of Redirects, the destination address in all ND packets that can modify any state in the recipient node have link-local scope; routers will be unable to forward them.

- packets containing a Routing Header are ignored upon receipt. If Routing Headers were allowed, it would be possible to forward packets through routers, even if the packet's ultimate destination has link-local scope.

Note that the use of link-local destination address makes the checks for link-local source address somewhat redundant for ND messages other than Redirects. The Redirect message is the only message type sent to a global unicast address that can modify the state in the receiving node. Thus proper robustness for Redirect messages requires that routers not forward packets with link-local source addresses.

The trust model for redirects is the same as in IPv4. A redirect is accepted only if received from the same router that is currently being used for that destination. It is natural to trust the routers on the link. If a host has been redirected to another node (i.e. the destination is on-link) there is no way to prevent the target from issuing another redirect to some other destination. However, this exposure is no worse than it was; the target host, once subverted, could always act as a hidden router to forward traffic elsewhere.

The protocol contains no mechanism to determine which nodes are authorized to send Router Advertisements; any node, presumably even in the presence of authentication, can send Router Advertisement messages thereby being able to cause denial of service. Furthermore, any node can send proxy Neighbor Advertisements as well as unsolicited Neighbor Advertisements as a potential denial of service attack.

Neighbor Discovery protocol packet exchanges can be authenticated using the IP Authentication Header [IPv6-AUTH]. A node SHOULD include an Authentication Header when sending Neighbor Discovery packets if a security association for use with the IP Authentication Header exists for the destination address. The security associations may have been created through manual configuration or through the operation of some

key management protocol.

Received Authentication Headers in Neighbor Discovery packets MUST be verified for correctness and packets with incorrect authentication MUST be ignored.

It SHOULD be possible for the system administrator to configure a node to ignore any Neighbor Discovery messages that are not authenticated using either the Authentication Header or Encapsulating Security Payload. The configuration technique for this MUST be documented. Such a switch SHOULD default to allowing unauthenticated messages.

Confidentiality issues are addressed by the IP Security Architecture and the IP Encapsulating Security Payload documents [IPv6-SA, IPv6-ESP].

REFERENCES

[ADDRCONF] S. Thomson, "IPv6 Address Autoconfiguration", Internet Draft.

[ADDR-ARCH] S. Deering, R. Hinden, Editors, "IP Version 6 Addressing Architecture", Internet Draft.

[ANYCST] C. Partridge, T. Mendez, and W. Milliken, "Host Anycasting Service", RFC 1546, November 1993.

[ARP] D. Plummer, "An Ethernet Address Resolution Protocol", STD 37, RFC 826, November 1982.

[HR-CL] R. Braden, Editor, "Requirements for Internet Hosts -- Communication Layers", STD 3, RFC 1122, October 1989.

[ICMPv4] J. Postel, "Internet Control Message Protocol", STD 5, RFC 792, September 1981.

[ICMPv6] A. Conta, and S. Deering, "ICMP for the Internet Protocol Version 6 (IPv6)", Internet Draft.

[IPv6] S. Deering, R. Hinden, Editors, "Internet Protocol, Version 6 (IPv6) Specification", Internet Draft.

[IPv6-ETHER] M. Crawford. "A Method for the Transmission of IPv6 Packets over Ethernet Networks", Internet Draft.

[IPv6-SA] R. Atkinson. "Security Architecture for the Internet Protocol". RFC 1825, August 1995.

[IPv6-AUTH] R. Atkinson. "IP Authentication Header", RFC 1826, August 1995.

[IPv6-ESP] R. Atkinson. "IP Encapsulating Security Payload (ESP)", RFC 1827, August 1995.

[RDISC] S. Deering, "ICMP Router Discovery Messages", RFC 1256, September 1991.

[SH-MEDIA] R. Braden, J. Postel, Y. Rekhter, "Internet Architecture Extensions for Shared Media", RFC 1620, May 1994.

[ASSIGNED] J. Reynolds, J. Postel, "ASSIGNED NUMBERS", RFC 1700, October 1994.

[SYNC] S. Floyd, V. Jacobsen, "The Synchronization of Periodic Routing Messages", IEEE/ACM Transactions on Networking, April 1994. ftp://ftp.ee.lbl.gov/papers/sync_94.ps.Z

## AUTHORS' ADDRESSES

| | |
|---|---|
| Erik Nordmark | Thomas Narten |
| Sun Microsystems, Inc. | IBM Corporation |

2550 Garcia Ave                P.O. Box 12195
Mt. View, CA 94041             Research Triangle Park, NC 27709-2195
USA                            USA

phone: +1 415 336 2788         phone: +1 919 254 7798
fax:   +1 415 336 6015         fax:   +1 919 254 4027
email: nordmark@sun.com        email: narten@vnet.ibm.com

William Allen Simpson
Daydreamer
Computer Systems Consulting Services
1384 Fontaine
Madison Heights, Michigan  48071
USA

email: Bill.Simpson@um.cc.umich.edu
       bsimpson@MorningStar.com

## CHANGES SINCE PREVIOUS DOCUMENT

There are several changes since the previous version documented in:
<draft-ietf-ipngwg-discovery-01.txt>

based on feedback from the working group:

- o Link-local source address required for Neighbor Solicitation and Neighbor Advertisement messages. This change implied adding an ICMP Sender Address field to the Neighbor Solicitation message and a Secondary Advertisement flag to the Neighbor Advertisement message. This change improves the robustness of the protocol - it is no longer possible for off-link nodes to send ND messages to a link.

- o Made the ReachableTime value random to avoid synchronizing Neighbor Unreachability Detection messages when there is more of less "constant" traffic (i.e. packets are sent with spacing that is very short compared to the ReachableTime value). Without such randomization the NUD probes from all nodes on the link would be sent with almost the same spacing which can result in synchronization. There is no need for any additional randomization elsewhere in the protocol since there is no long-term periodic behavior - at most 3 packets are transmitted.

- o Added definitions for MUST, SHOULD, and MAY.

- o Made NUD and address resolution use the same retransmission timer (which can be specified in the Router Advertisements). Increased the default value of this timer from 3 seconds to 10 seconds.

- o Restricted ReachableTime so that it can not be set to more than 1 hour to prevent misconfiguration that would make ND not detect e.g. changed link-layer addresses.

- o Added text about the support for links with multiple MTUs (e.g. bridged Ethernet and FDDI). With unmodified bridges the routers must send MTU options containing the smaller (smallest) MTU. If the bridges are made aware of IPv6 they can participate in path MTU discovery (for unicast and multicast) and send ICMP packet too big errors for IPv6 packets that cross the bridge.

- o Added additional validatity checks for Router Solicitation and Router Advertisement messages: the destination address must be a link local address or a multicast address with link-local scope.

- Added validity check for all messages: no routing header is allowed.

- Changed the use of the "designated address" term to using "link-local address".

- Clarified authentication header text.

- Clarified how multicast packets are handled in the conceptual model.

- Added text that ND messages are themselves not subject to NUD probes. This avoids an observed problem where a NUD NS/NA exchange would result in a subsequent NUD NS/NA exchange of packets.

- Clarified that the neighbor cache entries generated by unsolicited information (RS, RA, NA, Redirect) do still get DELAY_FIRST_PROBE_TIME seconds before a probe is sent (in order to benefit from upper-layer advise).

- Solicited Proxy/anycast advertisements are delayed 0-1 second to avoid creating a load on the network and/or receiver.