Routing over Large Clouds Working Group
INTERNET-DRAFT
<draft-ietf-rolc-nhrp-02.txt>

Dave Katz
(cisco Systems)
David Piscitello
(Core Competence, Inc.)
August 25, 1994

NBMA Next Hop Resolution Protocol (NHRP)

Status of this Memo

Abstract

This document describes the NBMA Next Hop Resolution Protocol (NHRP). NHRP can be used by a source station (host or router) connected to a Non-Broadcast, Multi-Access (NBMA) network to determine the IP and NBMA network addresses of the "NBMA next hop" towards a destination station. If the destination is connected to the NBMA network, then the NBMA next hop is the destination station itself. Otherwise, the NBMA next hop is the egress router from the NBMA network that is "nearest" to the destination station. Although this document focuses on NHRP in the context of IP, the technique is applicable to other network layer protocols as well.

This document is intended to be a functional superset of the NBMA Address Resolution Protocol (NARP) documented in [1].

1. Introduction

The NBMA Next Hop Resolution Protocol (NHRP) allows a source station (a host or router), wishing to communicate over a Non-Broadcast, Multi-Access (NBMA) network, to determine the IP and NBMA addresses of the "NBMA next hop" toward a destination station. A network can be non-broadcast either because it technically doesn't support broadcasting (e.g., an X.25 network) or because broadcasting is not feasible for one reason or another (e.g., an SMDS broadcast group or an extended Ethernet would be too large). If the destination is connected to the NBMA network, then the NBMA next hop is the destination station itself. Otherwise, the NBMA next hop is the egress router from the NBMA network that is "nearest" to the destination station.

An NBMA network may, in general, consist of multiple logically independent IP subnets (LISs), defined in [3] and [4] as having the following properties:

1) All members of a LIS have the same IP network/subnet number and address mask.

2) All members within a LIS are directly connected to the same NBMA network.

3) All members outside of the LIS are accessed via a router.

IP routing described in [3] and [4] only resolves the next hop address if the destination station is a member of the same LIS as the source station; otherwise, the source station must forward packets to a router that is a member of multiple LIS's. In multi-LIS configurations, hop-by-hop IP routing may not be sufficient to resolve the "NBMA next hop" toward the destination station, and IP packets may traverse the NBMA network more than once.

NHRP describes a routing method that obviates the need for LISs. With NHRP, once the NBMA next hop has been resolved, the source may either start sending IP packets to the destination (in a connectionless NBMA network such as SMDS) or may first establish a connection to the destination with the desired bandwidth and QOS characteristics (in a connection-oriented NBMA network such as ATM).

NHRP in its most basic form provides a simple IP-to-NBMA-address binding service. This may be sufficient for hosts which are directly connected to an NBMA network, allowing for straightforward implementations in NBMA stations. Optional services extend this functionality to include loop detection, sanity checks, diagnostics, security features, and fallback capabilities, providing improved robustness and functionality.

NHRP supports both a server-based style of deployment and a ubiquitous "fabric", consisting of NHRP-capable routers. The

server-based approach requires a smaller number of machines to support NHRP, but requires significantly more manual configuration.

Address resolution techniques such as those described in [3] and [4] may be in use when NHRP is deployed. ARP servers and services over NBMA networks may be required to support hosts that are not capable of dealing with any model for communication other than the LIS model, and deployed hosts may not implement NHRP but may continue to support ARP variants such as those described in [3] and [4]. NHRP is designed to eliminate the suboptimal routing that results from the LIS model, and can be deployed in a non-interfering manner alongside existing ARP services.

2. Protocol Overview

In this section, we briefly describe how a source S (which potentially can be either a router or a host) uses NHRP to determine the "NBMA next hop" to destination D.

For administrative and policy reasons, a physical NBMA network may be partitioned into several, disjoint "Logical NBMA networks". A Logical NBMA network is defined as a collection of hosts and routers that share ulfiltered data link connectivity over an NBMA network. "Unfiltered data link connectivity" refers to the absence of closed user groups, address screening or similar features that may be used to prevent direct communication between stations connected to the same NBMA network. (Hereafter, unless otherwise specified, we use NBMA network to mean logical NBMA network.)

Placed within the NBMA network are one or more entities that implement the NHRP protocol, otherwise known as "Next Hop Servers" (NHSs). Each NHS serves a set of destination hosts, which may or may not be directly connected to the NBMA network. NHSs cooperatively resolve the NBMA next hop within their logical NBMA network. In addition to the NHRP, NHSs participate in protocols used to disseminate routing information across (and beyond the boundaries of) the NBMA network, and may support "classical" ARP service as well.

An NHS maintains a "next-hop resolution" cache, which is a table of address mappings (IP-to-NBMA address). This table can be constructed from information gleaned from NHRP Register packets (see Section 5.4), extracted from NHRP replies that traverse NHS as they are forwarded toward the NHRP request initiator, or through mechanisms outside the scope of this document (examples of such mechanisms include ARP [2, 3, 4] and pre-configured tables).

A host or router that is not an NHRP speaker must be configured with

the identity of the NHS which serves it (see Configuration, Section 4).

[Note: for NBMA networks that offer group or multicast addressing features, it may be desirable to configure stations with a group identity for NHSs, i.e., addressing information that would solicit a response from "all NHSs". The means whereby a group of NHSs divide responsibilities for next hop resolution are not described here.]

The protocol proceeds as follows. An event occurs triggering station S to want to resolve the NBMA address of a path to D. This is most likely to be when data packet addressed to station D is to be emitted from station S (either because station S is a host, or station S is a transit router), but could also be triggered by other means (a resource reservation request, for example). Station S first determines the next hop to station D through normal routing processes (for a host, this may simply be the default router; for routers, this is the "next hop" to the destination IP address). If the next hop is reachable through its NBMA interface, S constructs an NHRP request packet (see Section 5.2) containing station D's IP address as the (target) destination address, S's own IP address as the source address (NHRP request initiator), and station S's NBMA addressing information. Station S may also indicate whether it prefers an authoritative reply (i.e., station S only wishes to receive a reply from the NHS-speaker that maintains the NBMA-to-IP address mapping for this destination). Station S encapsulates the NHRP request packet in an IP packet containing as its destination address the IP address of its NHS. This IP packet is emitted across the NBMA interface to the NBMA address of the NHS.

If the NHRP request is triggered by a data packet, station S may choose to dispose of the data packet While awaiting an NHRP reply in one of the following ways:

(a) Drop the packet
(b) Retain the packet until the reply arrives and a more optimal path is available
(c) Forward the packet along the routed path toward D

The choice of which of the above to perform is a local policy matter, though option (c) is an attractive default.

When the NHS receives an NHRP request, it checks to see if it "serves" station D, i.e., the NHS checks to see if it has a "next hop" entry for D in its next-hop resolution cache. If so, the NHS resolves station D's NBMA address. The NHS then generates a positive NHRP reply on D's behalf. The NHRP reply packet contains the next hop IP and NBMA address for station D and is sent back to S. The reply

generated in this case is marked as "authoritative". (Note that if station D is not on the NBMA network, the next hop IP address will be that of the egress router through which packets for station D are forwarded.)

If the NHS does not serve D, the NHS forwards the NHRP request to another NHS. (Mechanisms for determining how to forward the NHRP request are discussed in Section 3, Modes of Deployment.) If this NHS serves D, it generates a positive NHRP reply on D's behalf. (NHRP replies in this scenario are always marked as "authoritative".) NHRP replies usually traverse the same sequence of NHSs as the NHRP request (in reverse order). This is typically a consequence of having symmetric routing. An NHS receiving an NHRP reply may cache the NBMA next hop information contained therein. To a subsequent NHRP request, this NHS may respond with the cached, non-authoritative, NBMA next hop information or with cached negative information. Non-authoritative NHRP replies are distinguished from authoritative replies so that if a communication attempt based on non-authoritative information fails, a source station can choose to send an authoritative NHRP request. NHSs MUST never respond to authoritative NHRP requests with cached information.

[Note: An NHRP reply can be returned directly to the NHRP request initiator, i.e., without traversing the list of NHSs that forwarded the request, if all of the following criteria are satisfied:

(a) Direct communication is available via datagram transfer (e.g., SMDS) or the NHS has an existing virtual circuit connection to the NHRP request initiator or is permitted to open one.
(b) The NHRP request initiator has not included the NHRP Reverse NHS record Option (see Section 5.6.5).
(c) The NHRP request initiator has not included the destination mask option (see Section 5.6.1).
(d) The authentication policy in force permits direct communication between the NHS and the NHRP request initiator.

The purpose of allowing an NHS to reply directly is to reduce response time. A consequence of allowing a direct reply is that NHSs that would under normal circumstances be traversed by the reply would not cache next hop information contained therein.]

The process of forwarding the NHRP request is repeated until the request is satisfied, or an error occurs (e.g., no NHS in the NBMA network can resolve the request.) If the determination is made that station D's next hop cannot be resolved, a negative reply is returned. This occurs when (a) no next-hop resolution information is

available for station D from any NHS, or (b) an NHS is unable to forward the NHRP request (e.g., connectivity is lost).

NHRP requests and replies MUST never cross the borders of a logical NBMA network (an explicit NBMA network identifier may be included as an option in the NHRP request, see section 5.6.2). Thus, IP traffic out of and into a logical NBMA network always traverses an IP router at its border. Network layer filtering can then be implemented at these border routers.

NHRP optionally provides a mechanism to aggregate NBMA next hop information in NHS caches. Suppose that router X is the NBMA next hop from station S to station D. Suppose further that X is an egress router for all stations sharing an IP address prefix with station D. When an NHRP reply is generated in response to a request, the responder may augment the IP address of station D with a mask defining this prefix (see Section 5.6.1). The prefix to egress router mapping in the reply MUST be cached in all NHSs on the path of the reply. A subsequent (non-authoritative) NHRP request for some destination that shares an IP address prefix with D can be satisfied with this cached information.

To dynamically detect link-layer filtering in NBMA networks (e.g., X.25 closed user group facility, or SMDS address screens), as well as to provide loop detection and diagnostic capabilities, NHRP optionally incorporates a "Route Record" in requests and replies (see Sections 5.6.4 and 5.6.5). The Route Record options contain the network (and link layer) addresses of all intermediate NHSs between source and destination (in the forward direction) and between destination and source (in the reverse direction). When a source station is unable to communicate with the responder, it may attempt to do so successively with other link layer addresses in the Route Record until it succeeds (if authentication policy permits such action). This approach can find the optimal best hop in the presence of link-layer filtering (which may be source/destination sensitive, for instance, without necessarily creating separate logical NBMA networks) or link-layer congestion (especially in connection-oriented media).

3. Modes of Deployment

NHRP supports two deployment modes of operation: "server" and "fabric" modes. The two modes differ only in the way NHRP packets are propagated, which is driven by differences in configuration.

It is desirable that hosts attached directly to the NBMA network have no knowledge of whether NHRP is deployed in "server" or "fabric"

modes, so that a change in deployment strategy can be done within a single administration, transparently to hosts. For this reason, host configuration is invariant between the two cases. Note that irrespective of which mode is deployed, NHRP clients must nominally be configured with the NBMA (and IP) address of at least one NHS. In practice, a host's default router should also be its NHS.

Server Mode

In "server" mode, the expectation is that a small number of NHSs will be fielded in an NBMA network. This may be appropriate in networks containing routers that do not support NHRP, or networks that have large numbers of directly-attached hosts (and relatively few routers). Server mode assumes that NHRP is very loosely coupled with IP routing, and that the path taken by NHRP requests has little to do with the path taken by IP data packets routed to the desired destination.

[Note: This is the likely scenario for initial deployment of NHRP. It is also likely that single and Multi-LIS configurations using either group-addressed ARP (in the case of SMDS) or ARP servers (in the case of ATM or SMDS) may already be in place.]

Server mode uses static configuration of NHS identity. The client station must be configured with the IP address of one or more NHSs, and there must be a path to that NHS (either directly, in which case the NHS's NBMA address must be known, or indirectly, through a router whose NBMA address is known). If there are multiple NHSs, they must be configured with each others' addresses, the identities of the destinations that each of them serves, and optionally a logical NBMA network identifier. (This static configuration requirement, which may involve authentication as well as addressing information, tends to limit such deployments to a very small number of NHSs.)

If the NBMA network offers a group addressing or multicast feature, the client (station) may be configured with a group address assigned to the group of next-hop servers. The client might then submit NHRP requests to the group address, eliciting a response from one or more NHSs, depending on the response strategy selected. Note that the constraints described in Section 2 regarding direct replies may apply.

The servers can also be deployed with the group or multicast address of their peers, and an NHS might use this as a means of forwarding NHRP requests it cannot satisfy to its peers. This might elicit a response (to the NHS) from one or more NHSs, depending on the response strategy. The NHS would then forward the

NHRP reply to the NRHP request originator. The purpose of using group addressing or a similar multicast mechanism in this scenario would be to eliminate the need to preconfigure each NHS in a logical NBMA network with both the individual identities of other NHSs as well as the destinations they serve. It reduces the number of NHSs that might be traversed to process an NHRP request (in those configurations where NHSs either respond or forward via the multicast, only two NHSs would be traversed), and allows the NHS that serves the NHRP request originator to cache next hop information associated with the reply (again, within the constraints described in Section 2).

The NHRP request packet's destination IP address is set by the source station to the first-hop NHS's IP address. If the addressed NHS does not serve the destination, the NHRP request is forwarded to the IP address of the NHS that serves the destination.

The responding NHS uses the source address from within the NHRP packet (not the source address of the IP packet) as the IP destination of the NHRP reply.

Fabric Mode

In "fabric" mode, it is expected that NHRP-capable routers are ubiquitous throughout the NBMA network, and that NHSs acquire knowledge about destinations other NHSs serve as a direct consequence of participating in intradomain and interdomain routing protocol exchange. In particular, it is expected that an NHS serving a particular destination is guaranteed to lie along the routed path to that destination. In practice, this means that all egress routers must double as NHSs serving the destinations beyond them, and that hosts on the NBMA network are served by routers that double as NHSs.

Fabric mode leverages a routed infrastructure that "overlays" the NBMA network. The source station passes the NHRP request to the router which serves as the next hop toward the destination. Each router in turn forwards the NHRP request toward the destination. Eventually, the NHRP request arrives at a router that is acting as an NHS serving the destination (or the destination itself, if it is an NHRP-speaker), which generates the NHRP reply.

If the source station is a host, it sets the IP destination address of the NHRP request to the first-hop NHS/router (so that hosts needn't know the mode in which the network is running). If the source station is a router, the destination IP address may be set either to the next-hop router or to the ultimate destination being

resolved. Each NHS/router examines the NHRP request packet on its way toward the destination, optionally modifying it on the way (such as updating the Forward Record option). If an NHS/router receives an NHRP packet addressed to itself to which it cannot reply (because it does not serve the destination directly), it will forward the NHRP request with the destination IP address set to the ultimate destination (thus allowing invariant host behavior). Eventually, the NHRP packet will arrive at the router/NHS that serves the destination (which will return a positive NHRP reply) or it will arrive at a router/NHS that has no route to the destination (which will return a negative NHRP reply), or it may arrive at a router/NHS that cannot reach the NHS that serves the destination due to a loss of reachability among the NHSs (in which case the router will return a negative NHRP reply).

The procedural difference between server mode and fabric mode is reduced to deciding how to update the destination address in the IP packet carrying the NHRP request.

Note that addressing the NHRP request to the ultimate destination allows for networks that do not have NHSs deployed in all routers; typically a very large NBMA network might only deploy NHSs in egress routers, and not in transit routers.

4. Configuration

Stations

To participate in NHRP, a station connected to an NBMA network should be configured with the IP and NBMA address(es) of its NHS(s) (alternatively, it should be configured with a means of acquiring them, i.e., the group address that members of a NHS group use for the purpose of address or next-hop resolution.) The NHS(s) may be physically located on the stations's default or peer routers, so their addresses may be obtained from the station's IP forwarding table. If the station is attached to several link layer networks (including logical NBMA networks), the station should also be configured to receive routing information from its NHS(s) and peer routers so that it can determine which IP networks are reachable through which link layer networks.

Next Hop Servers

An NHS is configured with its own identity, a set of IP address prefixes that correspond to the IP addresses of the stations it serves, a logical NBMA network identifier (see Section 5.6.2), and

in the case of "server" mode, the identities of other NHSs in the same logical NBMA network. If a served station is attached to several link layer networks, the NHS may also need to be configured to advertise routing information to such stations.

If an NHS acts as an egress router for stations connected to other link layer networks than the NBMA network, the NHS must, in addition to the above, be configured to exchange routing information between the NBMA network and these other link layer networks.

In all cases, routing information is exchanged using conventional intra-domain and/or inter-domain routing protocols.

The NBMA addresses of the stations served by the NHS may be learned via NHRP Register packets or manual configuration.

5. Packet Formats

This section describes the format of NHRP packets.

An NHRP packet consists of a Fixed Part, a Mandatory Part, and an Options Part. The Fixed Part is common to all NHRP packet types. The Mandatory Part must be present, but varies depending on packet type. The Options Part also varies depending on packet type, and need not be present.

The length of the Fixed Part is fixed at 8 octets. The length of the Mandatory Part is carried in the Fixed Part. The length of the Options Part is implied by the total packet length (Internet datagram total length minus IP header length minus NHRP fixed part length minus NHRP mandatory part length).

Note that since the lengths of all fields are self-encoding, it is permissible to pad the Mandatory and Options parts with arbitrary numbers of trailing zero octets to achieve any desired alignment. Note however that any padding in the Mandatory Part must be included in the Mandatory Part Length.

NHRP packets are carried in IP packets as protocol type 54 (decimal). NHSs may increase the size of an NHRP packet as a result of option processing. IP datagrams containing NHRP packets must have the Don't Fragment bit set.

Fields marked "unused" must be set to zero on transmission, and ignored on receipt.

Most packet types have both network layer protocol-independent fields and protocol-specific fields. The protocol-independent fields always come first in the packet, and the Protocol ID field qualifies the format of the protocol-specific fields. The protocol-specific fields defined in this document are for IPv4 only; formats of protocol-specific fields for other protocols are for further study.

5.1 NHRP Fixed Header

The NHRP Fixed Header is present in all NHRP packets. It contains the basic information needed to parse the rest of the packet.

```
       0                   1                   2                   3
       0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |    Version    |   Hop Count   |          Checksum             |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |     Type      |    Unused     |    Mandatory Part Length      |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Version

The NHRP version number. Currently this value is 1.

Hop Count

The Hop count indicates the maximum number of NHSs that an NHRP packet is allowed to traverse before being discarded.

Checksum

The standard IP checksum over the entire NHRP packet (starting with the fixed header). If only the hop count field is changed, the checksum is adjusted without full recomputation. The checksum is completely recomputed when other header fields are changed.

Type

The NHRP packet type: Request, Response, Register, or Error Indication (see below).

Mandatory Part Length

The length in octets of the Mandatory Part. This length does not include the Fixed Header.

5.2 NHRP Request

The NHRP Request packet has a Type code of 1. The Mandatory Part has the following format:

```
       0                   1                   2                   3
       0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |Q|S|A|P|       Unused          |       Protocol ID             |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                         Request ID                            |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

                               (IPv4-Specific)
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                  Destination IP address                       |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                    Source IP address                          |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |         Holding Time          |    Unused     | Address Type  |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      | NBMA Length   |    Source NBMA Address (variable length)      |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Q

Set if the Requestor is a router; clear if the requestor is a host.

S

Unused (zero on transmit)

A

A response to an NHRP request may contain cached information. If an authoritative answer is desired, then this bit ("Authoritative") should be set. If non-authoritative (cached) information is acceptable, this bit should be clear.

P

Unused (zero on transmit)

Protocol ID

Specifies the network layer protocol for which we are obtaining routing information. This value also qualifies the structure of the remainder of the Mandatory Part. For IPv4, the Protocol ID is hexadecimal 800 (decimal 2048). Protocol ID values for other network layer protocols are for future study.

Request ID

A value chosen by the source to aid in matching requests with replies. This value could also be sent across a virtual circuit (in SVC environments) to aid in matching NHRP transactions with virtual circuits (this use is for further study).

Destination and Source IP Addresses

Respectively, these are the IP addresses of the station for which the NBMA next hop is desired, and the NHRP request initiator.

Source Holding Time, Address Type, NBMA Length, and NBMA Address

The Holding Time field specifies the number of seconds for which the source NBMA information is considered to be valid. Cached information shall be discarded when the holding time expires.

The Address Type field specifies the type of NBMA address (qualifying the NBMA address). Possible address types are <TBD>.

The NBMA length field is the length of the NBMA address of the source station in bits. The NBMA address field itself is zero-filled to the nearest 32-bit boundary.

5.3 NHRP Reply

The NHRP Reply packet has a type code of 2. The Mandatory Part has the following format:

```
        0                   1                   2                   3
        0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |Q|S|A|P|       Unused          |       Protocol ID             |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                          Request ID                           |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

                               (IPv4-Specific)
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                    Destination IP address                     |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                    Source IP address                          |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                    Next-hop IP address                        |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |         Holding Time          |  Preference   | Address Type  |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       | NBMA Length   |       NBMA Address (variable length)          |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                    ...
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                    Next-hop IP address                        |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |         Holding Time          |  Preference   | Address Type  |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       | NBMA Length   |       NBMA Address (variable length)          |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Q

Copied from the NHRP Request. Set if the Requestor is a router; clear if the requestor is a host.

S

Set if the next hop identified in the reply is a router; clear if the next hop is a host.

A

Set if the reply is authoritative; clear if the reply is non-authoritative.

P

Set if the reply is positive; clear if the reply is negative.

An NHS is not allowed to reply to an NHRP request for authoritative information with cached information, but may do so for an NHRP

Request which indicates a request for non-authoritative information. An NHS may reply to an NHRP request for non-authoritative information with authoritative information.

Protocol ID

Specifies the network layer protocol for which we are obtaining routing information. This value also qualifies the structure of the remainder of the Mandatory Part. For IPv4, the Protocol ID is hexadecimal 800 (decimal 2048). Protocol ID values for other network layer protocols are for future study.

Request ID

Copied from the NHRP Request.

Destination IP Address

The address of the target station (copied from the corresponding NHRP Request).

Source IP Address

The address of the initiator of the request (copied from the corresponding NHRP Request).

Next-hop entry

A Next-hop entry consists of the following fields: a 32-bit Next-hop IP Address, a 16-bit Holding Time, an 8-bit Preference, an 8-bit Address Type, an 8-bit NBMA Length, and an NBMA Address whose length is the value of the NBMA length field.

The Next-hop IP Address specifies the IP address of the next hop. This will be the address of the destination host if it is directly attached to the NBMA network, or the egress router if it is not directly attached.

The Holding Time field specifies the number of seconds for which the associated Next-hop entry inforamtion is considered to be valid. Cached information shall be discarded when the holding time expires. (Holding time is to be specified for both positive and negative replies).

The Preference field specifies the preference of the Next-hop entry, relative to other Next-hop entries in this NHRP Reply packet. Higher values indicate more preferable Next-hop entries.

The Address Type field specifies the type of NBMA address (qualifying the NBMA address). Possible address types are <TBD>.

The NBMA length field specifies the length of the NBMA address of the destination station in bits. The NBMA address field itself is

zero-filled to the nearest 32-bit boundary. For negative replies, the Holding Time field is relevant; however, the preference, Address Type, and NBMA length fields must be zero, and the NBMA Address shall not be present.

There may be multiple Next-hop entries returned in the reply (as implied by the Mandatory Part Length). The preference values are used to select the preferred entry. The same next-hop IP address may be associated with multiple NBMA addresses. Load-splitting may be performed over the addresses, given equal preference values, and the alternative addresses may be used in case of connectivity failure in the NBMA network (such as a failed call attempt in connection-oriented NBMA networks).

5.4 NHRP Register

The NHRP Register packet is sent from a station to an NHS to notify the NHS of the station's NBMA address. It has a Type code of 3. The Mandatory Part has the following format:

```
       0                   1                   2                   3
       0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |          Unused               |          Protocol ID          |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

                               (IPv4-Specific)
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                    Source IP address                          |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |         Holding Time          |    Unused     | Address Type  |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      | NBMA Length   |       NBMA Address (variable length)          |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Protocol ID

Specifies the network layer protocol for which we are obtaining routing information. This value also qualifies the structure of the remainder of the Mandatory Part. For IPv4, the Protocol ID is hexadecimal 800 (decimal 2048). Protocol ID values for other network layer protocols are for future study.

Source IP Address

The IP address of the station wishing to register its NBMA address with an NHS.

Source Holding Time, Address Type, NBMA Length, and NBMA Address

The Holding Time field specifies the number of seconds for which the source NBMA information is considered to be valid. Cached information shall be discarded when the holding time expires.

The Address Type field specifies the type of NBMA address (qualifying the NBMA address). Possible address types are <TBD>.

The NBMA length field is the length of the NBMA address of the source station in bits. The NBMA address itself is zero-filled to the nearest 32-bit boundary.

This packet is used to register a station's IP and NBMA addresses with its configured NHS. This allows static configuration information to be reduced; the NHSs need not be configured with the identities of all of the stations that they serve.

It is possible that a misconfigured station will attempt to register with the wrong NHS (i.e., one that cannot serve it due to policy constraints or routing state). If this is the case, the NHS must reply with an Error Indication of type Can't Serve This Address.

If an NHS cannot serve a station due to a lack of resources, the NHS must reply with an Error Indication of type Registration Overflow.

5.5 NHRP Error Indication

The NHRP Error Indication is used to convey error indications to the initiator of an NHRP Request packet. It has a type code of 4. The Mandatory Part has the following format:

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |           Error Code          |        Error Offset           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                                                               |
   +-+-+-+-+-+-+-+-+  Contents of NHRP Packet in error +-+-+-+-+-+-+-+
   |                                                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Error Code
  An error code indicating the type of error detected, chosen from the following list:

    1 - Unrecognized Option
    2 - Network ID Mismatch

```
3 - NHRP Loop Detected
4 - Can't Serve This Address
5 - Registration Overflow
6 - Server Unreachable
7 - Protocol Error
8 - NHRP fragmentation failure
```

Error Offset
The offset in octets into the original NHRP packet, starting at the NHRP Fixed Header, at which the error was detected.

The destination IP address of an NHRP Error Indication shall be set to the IP address of the initiator of the original NHRP Request (as extracted from the NHRP Request or NHRP Reply).

An Error Indication packet shall never be generated in response to another Error Indication packet. When an Error Indication packet is generated, the offending NHRP packet shall be discarded. In no case should more than one Error Indication packet be generated for a single NHRP packet.

5.6 Options Part

The Options Part, if present, carries one or more options in {Type, Length, Value} triplets. Options are only present in a Reply if they were present in the corresponding Request; therefore, minimal NHRP station implementations that do not act as an NHS and do not transmit options need not be able to receive them. Such an implementation that receives a packet with options shall return an Error Indication of type Unrecognized Option.

Options are typically protocol-specific, as noted.

Options have the following format:

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |O|           Type              |        Length                 |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                          Value...                             |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

O
"Optional." If set, and the NHS does not recognize the type code, the option may safely be ignored. If clear, and the NHS does not recognize the type code, the NHRP request is considered in error. (See below for details.)

Type
The option type code (see below). The option type is not qualified by the Optional bit, but is orthogonal to it.

Length
The length in octets of the value (not including the Type and Length fields; a null option will have only an option header and a length of zero).

Each option is padded with zero octets to a 32 bit boundary. This padding is not included in the Length field.

Options may occur in any order, but any particular option type may occur only once in an NHRP packet.

The Optional bit provides for a means to extend the option set. If it is clear, the NHRP request cannot be satisfied if the option is unrecognized, so the responder must return an Error Indication of type Unrecognized Option. If set, the option can be safely ignored. In this case, the offending option should simply be returned unchanged in the NHRP Reply.

If a transit NHS (one which is not going to generate a reply) detects an unrecognized option, it shall ignore the option, and if the Optional bit is clear, must not cache the information (in the case of a reply) and must not identify itself as an egress router (in the Forward Record or Reverse Record options). Effectively, this means that a transit NHS that doesn't understand an option with the Optional bit clear must not participate in any way in the protocol exchange, other than acting as a forwarding agent for the request.

5.6.1 Destination Mask Option (IPv4-Specific)

Optional = 0
Type = 1
Length = 4

This option is used to indicate that the information carried in an NHRP Reply pertains to an equivalence class of destinations rather than just the destination IP address specified in the request. All addresses that match the destination IP address in the bit positions

for which the mask has a one bit are part of the equivalence class.

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                    Destination Mask                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

If an initiator would like to receive this equivalence information, it shall add this option to the NHRP Request with a value of 255.255.255.255. The responder shall copy the option to the NHRP Reply and modify the mask appropriately.

5.6.2 NBMA Network ID Option (Protocol-Independent)

Optional = 0
Type = 2
Length = variable

This option is used to carry one or more identifiers for the NBMA network. This can be used as a validity check to ensure that the request does not leave a particular NBMA network. The option is placed in an NHRP Request packet by the initiator with an ID value of zero; the first NHS fills in the field with the identifier(s) for the NBMA network.

Multiple NBMA Network IDs may be used as a transition mechanism while NBMA Networks are being split or merged.

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                     NBMA Network ID                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                    ...
```

Each identifier consists of a 32 bit globally unique value assigned to the NBMA network. This value should be chosen from the IP address space administered by the operators of the NBMA network. This value is used for identification only, not for routing or any other purpose.

Each NHS processing an NHRP Request shall verify these values. If none of the values matches the NHS's NBMA Network ID, the NHS shall return an Error Indication of type "Network ID Mismatch" and discard the NHRP Request.

When an NHS is building an NHRP Reply and the NBMA Network ID option is present in the NHRP Request, the NBMA Network ID option shall be copied from the Request to the Reply.

Each NHS processing an NHRP Reply shall verify the value carried in the NBMA Network ID option, if present. If none of the values matches the NHSs NBMA Network ID, the NHS shall return an Error Indication of type "Network ID Mismatch" and discard the NHRP Reply.

5.6.3 Responder Address Option (IPv4-Specific)

Optional = 0
Type = 3
Length = 4

This option is used to determine the IP address of the NHRP Responder, that is, the entity that generates the NHRP Reply packet.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Responder's IP Address                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

If a requestor desires this information, it shall include this option, with a value of zero, in the NHRP Request packet.

If an NHS is generating an NHRP Reply packet in response to a request containing this option, it shall include this option, containing its IP address, in the NHRP Reply. If an NHS has more than one IP address, it shall use the same IP address consistently in all of the Responder Address, Forward NHS Record, and Reverse NHS Record options.

If an NHRP Reply packet being forwarded by an NHS contains the IP address of that NHS in the Responder Address Option, the NHS shall generate an Error Indication of type "NHRP Loop Detected" and discard the Reply.

If an NHRP Reply packet is being returned by an intermediate NHS based on cached data, it shall place its own address in this option (differentiating it from the address in the Next-hop field).

5.6.4 NHRP Forward NHS Record Option (IPv4-Specific)

Optional = 0
Type = 4
Length = variable

The NHRP forward NHS record is a list of NHSs through which an NHRP request traverses. Each NHS shall append a Next-hop element containing its IP address to this option.

In addition, NHSs that are willing to act as egress routers for packets from the source to the destination shall include information about their NBMA Address.

Each Next-hop element is formatted as follows:

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                    IP address                                 |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |         Holding Time          |    Unused     | Address Type  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | NBMA Length   |       NBMA Address (variable length)          |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IP address
  The IP address of the NHS.

Holding Time
  The number of seconds for which this information is valid. If a station chooses to use this information as a next-hop entry, it may not be used once the holding timer expires.

Address Type, NBMA Length, and NBMA Address
  The Address Type field specifies the type of NBMA address (qualifying the NBMA address). Possible address types are <TBD>.

  The NBMA length field is the length of the NBMA address of the destination station in bits. The NBMA address itself is zero-filled to the nearest 32-bit boundary.

  NHSs that are not egress routers shall specify an NBMA Length of zero and shall not include an NBMA Address.

If a requestor wishes to obtain this information, it shall include this option with a length of zero.

Each NHS shall append an appropriate Next-hop element to this option when processing an NHRP Request. The option length field and NHRP checksum shall be adjusted as necessary.

The last-hop NHS (the one that will be generating the NHRP Reply) shall not update this option (since this information will be in the reply).

If an NHS has more than one IP address, it shall use the same IP address consistently in all of the Responder Address, Forward NHS Record, and Reverse NHS Record options.

If an NHRP Request packet being forwarded by an NHS contains the IP address of that NHS in the Forward NHS Record Option, the NHS shall generate an Error Indication of type "NHRP Loop Detected" and discard the Request.

5.6.5 NHRP Reverse NHS Record Option (IPv4-Specific)

Optional = 0
Type = 5
Length = variable

The NHRP reverse NHS record is a list of NHSs through which an NHRP reply traverses. Each NHS shall append a Next-hop element containing its IP address to this option.

In addition, NHSs that are willing to act as egress routers for packets from the source to the destination shall include information about their NBMA Address.

Each Next-hop element is formatted as follows:

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                    IP address                                 |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |         Holding Time          |    Unused     | Address Type  |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    | NBMA Length   |       NBMA Address (variable length)          |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IP address
  The IP address of the NHS.

Holding Time

The number of seconds for which this information is valid. If a station chooses to use this information as a next-hop entry, it may not be used once the holding timer expires.

Address Type, NBMA Length, and NBMA Address

The Address Type field specifies the type of NBMA address (qualifying the NBMA address). Possible address types are <TBD>.

The NBMA length field is the length of the NBMA address of the destination station in bits. The NBMA address itself is zero-filled to the nearest 32-bit boundary.

NHSs that are not egress routers shall specify an NBMA Length of zero and shall not include an NBMA Address.

If a requestor wishes to obtain this information, it shall include this option with a length of zero.

Each NHS shall append an appropriate Next-hop element to this option when processing an NHRP Reply. The option length field and NHRP checksum shall be adjusted as necessary.

The NHS generating the NHRP Reply shall not update this option.

If an NHS has more than one IP address, it shall use the same IP address consistently in all of the Responder Address, Forward NHS Record, and Reverse NHS Record options.

If an NHRP Reply packet being forwarded by an NHS contains the IP address of that NHS in the Reverse NHS Record Option, the NHS shall generate an Error Indication of type "NHRP Loop Detected" and discard the Reply.

Note that this information may be cached at intermediate NHSs; if so, the cached value shall be used when generating a reply. Note that the Responder Address option may be used to disambiguate the set of NHSs that actually processed the reply.

5.6.6 NHRP QoS Option

Optional = 0
Type = 6
Length = variable

The NHRP QoS Option is carried in NHRP Request packets to indicate the desired QoS of the path to the indicated destination. This

information may be used to help select the appropriate NBMA next hop.

It may also be carried in NHRP Register packets to indicate the QoS to which the registration applies.

The syntax and semantics of this option are TBD; alignment with resource reservation may be useful.

5.6.7 NHRP Authentication Option

Optional = 0
Type = 7
Length = variable

The NHRP Authentication Option is carried in NHRP packets to convey authentication information between NHRP speakers. The semantics and encoding of the authentication option is for further study.

6. Security Considerations

Security considerations are for further study.

7. Discussion

The result of an NHRP request depends on how routing is configured among the NHSs of an NBMA network. If the destination station is directly connected to the NBMA network and the NHSs always prefer NBMA routes over routes via other link layer networks, the NHRP replies always return the NBMA address of the destination station itself rather than the NBMA address of some egress router. For destinations outside the NBMA network, egress routers and routers in the other link layer networks should exchange routing information so that the optimal egress router is always found.

When the NBMA next hop toward a destination is not the destination station itself, the optimal NBMA next hop may change dynamically. This can happen, for instance, when an egress router nearer to the destination becomes available. This change can be detected in a number of ways. First of all, the source station will need to periodically reissue the NHRP Request at a minimum just prior to the expiration of the holding timer, and most likely more aggressively than that. Alternatively, the source can be configured to receive routing information from its NHSs. When it detects an improvement in the route to the destination, the source can reissue the NHRP request to obtain the current optimal NBMA next hop. Source stations that

are routers may choose to establish a routing association with the egress router, allowing the egress router to explicitly inform the source about changes in routing (and providing additional routing information, authentication, etc.)

The dynamic nature of routing impacts caching strategies as well, since cached information may not be up-to-date. This is especially an issue when NHSs are deployed in server mode, since the NHSs may not be privy to routing information. However, stale cached information may only cause suboptimal routing (choosing the wrong egress point and taking extra hops across the NBMA network) rather than causing black holes. Cache management strategies are for further study.

In addition to NHSs, an NBMA station could also be associated with one or more regular routers that could act as "connectionless servers" for the station. The station could then choose to resolve the NBMA next hop or just send the IP packets to one of its connectionless servers. The latter option may be desirable if communication with the destination is short-lived and/or doesn't require much network resources. The connectionless servers could, of course, be physically integrated in the NHSs by augmenting them with IP switching functionality.

NHRP supports portability of NBMA stations. A station can be moved anywhere within the NBMA network and still keep its original IP address as long as its NHS(s) remain the same. Requests for authoritative information will always return the correct link layer address.

8. Protocol Operation

In this section, we discuss certain operational considerations of NHRP.

8.1 Router-to-Router Operation

In practice, the initiating and responding stations may be either hosts or routers. However, there is a possibility under certain conditions that a stable routing loop may occur if NHRP is used between two routers. This situation can be avoided if there are no "back door" paths between the entry and egress router outside of the NBMA network, and can be ameliorated by periodically reissuing the NHRP request. If these conditions cannot be satisfied, the use of NHRP between routers is not recommended.

8.2 Handling of IP Destination Address Field

NHRP packets are self-contained in terms of the IP addressing information needed for protocol operation--the IP source and destination addresses in the encapsulating IP header are not used. However, the setting of the IP destination address field does impact how NHRP requests are forwarded.

There are essentially three choices in how to set the destination IP address field at any particular point in the forwarding of an NHRP request: the ultimate destination being resolved, the next-hop IP router on the path to the destination, and the next-hop NHS (which might not be adjacent to the NHS forming the packet header).

The first case, addressing the packet to the destination being resolved (in the hopes that an NHS lies along the path) is desirable for at least two reasons. It simplifies configuration (since the identity of the next NHS need not be known explicitly), and it simplifies deployment (since the packet will pass silently through routers that are not NHSs). However, it assumes that the serving NHS lies along the path to the destination, and it requires NHSs along the path to examine the packet even though it is not addressed to them.

The second case, addressing the packet to the next-hop router, is similar to the first in that it follows the path to the destination, thus reducing configuration complexity. It furthermore only requires NHSs to process the packet if they are directly addressed. It too assumes that the responding NHS is on the path to the destination. However, it requires that all routers along the path are also NHSs.

The third case, addressing the packet to the next-hop NHS, allows the NHSs to be independent of routing, and requires only addressed NHSs to examine the packet. However, there is no reasonable way, other than manual configuration, to determine the identity of the next hop NHS if it is not also the next hop IP router (making it option two).

In order to balance all of these issues, the following rules shall be used when constructing IP packets to carry NHRP requests.

Stations

Stations shall address NHRP packets to the NHS by which they are served, regardless of whether NHRP has been deployed in Server mode or Fabric mode.

NHSs

If an NHS receives an NHRP packet in which the IP destination address does not match any of its own IP addresses, it shall process the NHRP packet as appropriate, and if it must forward the NHRP packet to another NHS, shall transmit the packet with the same IP destination address with which it was received.

If an NHS receives an NHRP packet in which the IP destination address matches one of its own IP addresses, it shall process the NHRP packet as appropriate, and if it must forward the NHRP packet to another NHS, shall set the destination IP address in one of the following ways:

If there is a configured next-hop NHS for the destination being resolved (Server mode), it shall transmit the packet with the IP destination address set to the next-hop NHS.

If there is no configured next-hop NHS (Fabric Mode), it shall transmit the packet with the IP destination address set to the address of the destination being resolved, and shall include the Router Alert option [5] so that intermediate NHS/routers can examine the NHRP packet.

8.3 Pseudocode

TBD.

References

[1] NBMA Address Resolution Protocol (NARP), Juha Heinanen and Ramesh Govindan, draft-ietf-rolc-nbma-arp-00.txt.

[2] Address Resolution Protocol, David C. Plummer, RFC 826.

[3] Classical IP and ARP over ATM, Mark Laubach, Internet Draft.

[4] Transmission of IP datagrams over the SMDS service, J. Lawrence and D. Piscitello, RFC 1209.

[5] IP Router Alert Option, Dave Katz, draft-katz-router-alert-00.txt.

Acknowledgements

We would like to thank Juha Heinenan of Telecom Finland and Ramesh Govidan of ISI for their work on NBMA ARP and the original NHRP

draft, which served as a basis for this work. John Burnett of Adaptive, Dennis Ferguson of ANS, Joel Halpern of Network Systems, Paul Francis of NTT, and Tony Li of cisco should also be acknowledged for comments and suggestions that improved this work substantially.

Authors' Addresses

Dave Katz
cisco Systems
1525 O'Brien Dr.
Menlo Park, CA 94025 USA

Phone: +1 415 688 8284
Email: dkatz@cisco.com

David Piscitello
Core Competence
1620 Tuckerstown Road
Dresher, PA 19025 USA

Phone: +1 215 830 0692
Email: dave@corecom.com