Network Working Group   J. Luciani
Request for Comments: 2332   Bay Networks
Category: Standards Track   D. Katz
                                                        cisco Systems
                                                        D. Piscitello
                                                 Core Competence, Inc.
                                                              B. Cole
                                                     Juniper Networks
                                                         N. Doraswamy
                                                         Bay Networks
                                                           April 1998

NBMA Next Hop Resolution Protocol (NHRP)

Status of this Memo

Copyright Notice

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) subnetwork to determine the
   internetworking layer address and NBMA subnetwork addresses of the
   "NBMA next hop" towards a destination station. If the destination is
   connected to the NBMA subnetwork, then the NBMA next hop is the
   destination station itself. Otherwise, the NBMA next hop is the
   egress router from the NBMA subnetwork that is "nearest" to the
   destination station. NHRP is intended for use in a multiprotocol
   internetworking layer environment over NBMA subnetworks.

   Note that while this protocol was developed for use with NBMA
   subnetworks, it is possible, if not likely, that it will be applied
   to BMA subnetworks as well. However, this usage of NHRP is for
   further study.

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

Operation of NHRP as a means of establishing a transit path across an
NBMA subnetwork between two routers will be addressed in a separate
document (see [13]).

1. **Introduction**

The keywords MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD,
SHOULD NOT, RECOMMENDED, MAY, and OPTIONAL, when they appear in this
document, are to be interpreted as described in [15].

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) subnetwork, to determine the internetworking
layer addresses and NBMA addresses of suitable "NBMA next hops"
toward a destination station. A subnetwork can be non-broadcast
either because it technically doesn't support broadcasting (e.g., an
X.25 subnetwork) or because broadcasting is not feasible for one
reason or another (e.g., an SMDS multicast group or an extended
Ethernet would be too large). If the destination is connected to the
NBMA subnetwork, then the NBMA next hop is the destination station
itself. Otherwise, the NBMA next hop is the egress router from the
NBMA subnetwork that is "nearest" to the destination station.

One way to model an NBMA network is by using the notion of logically
independent IP subnets (LISs). LISs, as defined in [3] and [4], have
the following properties:

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

    2) All members of a LIS are directly connected to the same
        NBMA subnetwork.

    3) All hosts and routers outside of the LIS are accessed via
        a router.

    4) All members of a LIS access each other directly (without
        routers).

Address resolution as 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 address resolution may not be sufficient
to resolve the "NBMA next hop" toward the destination station, and IP
packets may have multiple IP hops through the NBMA subnetwork.

Another way to model NBMA is by using the notion of Local Address
Groups (LAGs) [10]. The essential difference between the LIS and the
LAG models is that while with the LIS model the outcome of the
"local/remote" forwarding decision is driven purely by addressing
information, with the LAG model the outcome of this decision is
decoupled from the addressing information and is coupled with the
Quality of Service and/or traffic characteristics. With the LAG
model any two entities on a common NBMA network could establish a
direct communication with each other, irrespective of the entities'
addresses.

Support for the LAG model assumes the existence of a mechanism that
allows any entity (i.e., host or router) connected to an NBMA network
to resolve an internetworking layer address to an NBMA address for
any other entity connected to the same NBMA network. This resolution
would take place regardless of the address assignments to these
entities. Within the parameters described in this document, NHRP
describes such a mechanism. For example, when the internetworking
layer address is of type IP, once the NBMA next hop has been
resolved, the source may either start sending IP packets to the
destination (in a connectionless NBMA subnetwork such as SMDS) or may
first establish a connection to the destination with the desired
bandwidth (in a connection-oriented NBMA subnetwork such as ATM).

Use of NHRP may be sufficient for hosts doing address resolution when
those hosts are directly connected to an NBMA subnetwork, allowing
for straightforward implementations in NBMA stations. NHRP also has
the capability of determining the egress point from an NBMA
subnetwork when the destination is not directly connected to the NBMA
subnetwork and the identity of the egress router is not learned by
other methods (such as routing protocols). Optional extensions to
NHRP provide additional robustness and diagnosability.

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 subnetworks 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 intended to reduce or eliminate the extra router hops required by
the LIS model, and can be deployed in a non-interfering manner with
existing ARP services [14].

The operation of NHRP to establish transit paths across NBMA
subnetworks between two routers requires additional mechanisms to
avoid stable routing loops, and will be described in a separate
document (see [13]).

2. Overview

2.1 Terminology

The term "network" is highly overloaded, and is especially confusing
in the context of NHRP. We use the following terms:

   Internetwork layer--the media-independent layer (IP in the case of
   TCP/IP networks).

   Subnetwork layer--the media-dependent layer underlying the
   internetwork layer, including the NBMA technology (ATM, X.25, SMDS,
   etc.)

   The term "server", unless explicitly stated to the contrary, refers
   to a Next Hop Server (NHS). An NHS is an entity performing the
   Next Hop Resolution Protocol service within the NBMA cloud. An NHS
   is always tightly coupled with a routing entity (router, route
   server or edge device) although the converse is not yet guaranteed
   until ubiquitous deployment of this functionality occurs. Note
   that the presence of intermediate routers that are not coupled with
   an NHS entity may preclude the use of NHRP when source and
   destination stations on different sides of such routers and thus
   such routers may partition NHRP reachability within an NBMA
   network.

   The term "client", unless explicitly stated to the contrary, refers
   to a Next Hop Resolution Protocol client (NHC). An NHC is an
   entity which initiates NHRP requests of various types in order to
   obtain access to the NHRP service.

   The term "station" generally refers to a host or router which
   contains an NHRP entity. Occasionally, the term station will
   describe a "user" of the NHRP client or service functionality; the
   difference in usage is largely semantic.

2.2 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 subnetwork may
be partitioned into several, disjoint "Logical NBMA subnetworks". A
Logical NBMA subnetwork is defined as a collection of hosts and
routers that share unfiltered subnetwork connectivity over an NBMA
subnetwork. "Unfiltered subnetwork 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 subnetwork. (Hereafter, unless otherwise
specified, we use the term "NBMA subnetwork" to mean *logical* NBMA
subnetwork.)

Placed within the NBMA subnetwork are one or more entities that
implement the NHRP protocol. Such stations which are capable of
answering NHRP Resolution Requests are 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 subnetwork. NHSs cooperatively
resolve the NBMA next hop within their logical NBMA subnetwork. In
addition to NHRP, NHSs may support "classical" ARP service; however,
this will be the subject of a separate document [14].

An NHS maintains a cache which contains protocol layer address to
NBMA subnetwork layer address resolution information. This cache can
be constructed from information obtained from NHRP Register packets
(see Section 5.2.3 and 5.2.4), from NHRP Resolution Request/Reply
packets, or through mechanisms outside the scope of this document
(examples of such mechanisms might include ARP[3] and pre-configured
tables). Section 6.2 further describes cache management issues.

For a station within a given LIS to avoid providing NHS
functionality, there must be one or more NHSs within the NBMA
subnetwork which are providing authoritative address resolution
information on its behalf. Such an NHS is said to be "serving" the
station. A station on a LIS that lacks NHS functionality and is a
client of the NHRP service is known as NHRP Client or just NHCs. If
a serving NHS is to be able to supply the address resolution
information for an NHC then NHSs must exist at each hop along all
routed paths between the NHC making the resolution request and the
destination NHC. The last NHRP entity along the routed path is the
serving NHS; that is, NHRP Resolution Requests are not forwarded to
destination NHCs but rather are processed by the serving NHS.

An NHC also maintains a cache of protocol address to NBMA address
resolution information. This cache is populated through information
obtained from NHRP Resolution Reply packets, from manual
configuration, or through mechanisms outside the scope of this
document.

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 a 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 the address resolution could also
be triggered by other means (a routing protocol update packet, for
example). Station S first determines the next hop to station D
through normal routing processes (for a host, the next hop may simply
be the default router; for routers, this is the "next hop" to the
destination internetwork layer address). If the destination's
address resolution information is already available in S's cache then
that information is used to forward the packet. Otherwise, if the
next hop is reachable through one of its NBMA interfaces, S
constructs an NHRP Resolution Request packet (see Section 5.2.1)
containing station D's internetwork layer address as the (target)
destination address, S's own internetwork layer address as the source
address (Next Hop Resolution Request initiator), and station S's NBMA
addressing information. Station S may also indicate that it prefers
an authoritative NHRP Resolution Reply (i.e., station S only wishes
to receive an NHRP Resolution Reply from an NHS serving the
destination NHC). Station S emits the NHRP Resolution Request packet
towards the destination.

If the NHRP Resolution Request is triggered by a data packet then S
may, while awaiting an NHRP Resolution Reply, choose to dispose of
the data packet in one of the following ways:

  (a) Drop the packet
  (b) Retain the packet until the NHRP Resolution 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 the recommended default, since it may allow data
to flow to the destination while the NBMA address is being resolved.
Note that an NHRP Resolution Request for a given destination MUST NOT
be triggered on every packet.

When the NHS receives an NHRP Resolution Request, a check is made to
see if it serves station D. If the NHS does not serve D, the NHS
forwards the NHRP Resolution Request to another NHS. Mechanisms for
determining how to forward the NHRP Resolution Request are discussed
in Section 3.

If this NHS serves D, the NHS resolves station D's NBMA address
information, and generates a positive NHRP Resolution Reply on D's
behalf. NHRP Resolution Replies in this scenario are always marked
as "authoritative". The NHRP Resolution Reply packet contains the

address resolution information for station D which is to be sent back
to S. Note that if station D is not on the NBMA subnetwork, the next
hop internetwork layer address will be that of the egress router
through which packets for station D are forwarded.

A transit NHS receiving an NHRP Resolution Reply may cache the
address resolution information contained therein. To a subsequent
NHRP Resolution Request, this NHS may respond with the cached, "non-
authoritative" address resolution information if the NHS is permitted
to do so (see Sections 5.2.2 and 6.2 for more information on non-
authoritative versus authoritative NHRP Resolution Replies). Non-
authoritative NHRP Resolution Replies are distinguished from
authoritative NHRP Resolution Replies so that if a communication
attempt based on non-authoritative information fails, a source
station can choose to send an authoritative NHRP Resolution Request.
NHSs MUST NOT respond to authoritative NHRP Resolution Requests with
cached information.

If the determination is made that no NHS in the NBMA subnetwork can
reply to the NHRP Resolution Request for D then a negative NHRP
Resolution Reply (NAK) 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 Resolution Request (e.g.,
connectivity is lost).

NHRP Registration Requests, NHRP Purge Requests, NHRP Purge Replies,
and NHRP Error Indications follow a routed path in the same fashion
that NHRP Resolution Requests and NHRP Resolution Replies do.
Specifically, "requests" and "indications" follow the routed path
from Source Protocol Address (which is the address of the station
initiating the communication) to the Destination Protocol Address.
"Replies", on the other hand, follow the routed path from the
Destination Protocol Address back to the Source Protocol Address with
the following exceptions: in the case of a NHRP Registration Reply
and in the case of an NHC initiated NHRP Purge Request, the packet is
always returned via a direct VC (see Sections 5.2.4 and 5.2.5); if
one does not exists then one MUST be created.

NHRP Requests and NHRP Replies do NOT cross the borders of a NBMA
subnetwork however further study is being done in this area (see
Section 7).   Thus, the internetwork layer data traffic out of and
into an NBMA subnetwork always traverses an internetwork layer router
at its border.

NHRP optionally provides a mechanism to send a NHRP Resolution Reply
which contains aggregated address resolution information. For
example, suppose that router X is the next hop from station S to
station D and that X is an egress router for all stations sharing an

internetwork layer address prefix with station D. When an NHRP
Resolution Reply is generated in response to a NHRP Resolution
Request, the responder may augment the internetwork layer address of
station D with a prefix length (see Section 5.2.0.1). A subsequent
(non-authoritative) NHRP Resolution Request for some destination that
shares an internetwork layer address prefix (for the number of bits
specified in the prefix length) with D may be satisfied with this
cached information. See section 6.2 regarding caching issues.

To dynamically detect subnetwork-layer filtering in NBMA subnetworks
(e.g., X.25 closed user group facility, or SMDS address screens), to
trace the routed path that an NHRP packet takes, or to provide loop
detection and diagnostic capabilities, a "Route Record" may be
included in NHRP packets (see Sections 5.3.2 and 5.3.3). The Route
Record extensions are the NHRP Forward Transit NHS Record Extension
and the NHRP Reverse Transit NHS Record Extension. They contain the
internetwork (and subnetwork layer) addresses of all intermediate
NHSs between source and destination and between destination and
source respectively. When a source station is unable to communicate
with the responder (e.g., an attempt to open an SVC fails), it may
attempt to do so successively with other subnetwork layer addresses
in the NHRP Forward Transit NHS Record Extension until it succeeds
(if authentication policy permits such action). This approach can
find a suitable egress point in the presence of subnetwork-layer
filtering (which may be source/destination sensitive, for instance,
without necessarily creating separate logical NBMA subnetworks) or
subnetwork-layer congestion (especially in connection-oriented
media).

3. Deployment

NHRP Resolution Requests traverse one or more hops within an NBMA
subnetwork before reaching the station that is expected to generate a
response. Each station, including the source station, chooses a
neighboring NHS to which it will forward the NHRP Resolution Request.
The NHS selection procedure typically involves applying a destination
protocol layer address to the protocol layer routing table which
causes a routing decision to be returned. This routing decision is
then used to forward the NHRP Resolution Request to the downstream
NHS. The destination protocol layer address previously mentioned is
carried within the NHRP Resolution Request packet. Note that even
though a protocol layer address was used to acquire a routing
decision, NHRP packets are not encapsulated within a protocol layer
header but rather are carried at the NBMA layer using the
encapsulation described in Section 5.

Each NHS/router examines the NHRP Resolution Request packet on its
way toward the destination. Each NHS which the NHRP packet traverses
on the way to the packet's destination might modify the packet (e.g.,
updating the Forward Record extension). Ignoring error situations,
the NHRP Resolution Request eventually arrives at a station that is
to generate an NHRP Resolution Reply. This responding station
"serves" the destination. The responding station generates an NHRP
Resolution Reply using the source protocol address from within the
NHRP packet to determine where the NHRP Resolution Reply should be
sent.

Rather than use routing to determine the next hop for an NHRP packet,
an NHS may use other applicable means (such as static configuration
information) in order to determine to which neighboring NHSs to
forward the NHRP Resolution Request packet as long as such other
means would not cause the NHRP packet to arrive at an NHS which is
not along the routed path. The use of static configuration
information for this purpose is beyond the scope of this document.

The NHS serving a particular destination must 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 subnetwork are served by routers that double
as NHSs. Also, this implies that forwarding of NHRP packets within
an NBMA subnetwork requires a contiguous deployment of NHRP capable
routers. It is important that, in a given LIS/LAG which is using
NHRP, all NHSs within the LIS/LAG have at least some portion of their
resolution databases synchronized so that a packet arriving at one
router/NHS in a given LIS/LAG will be forwarded in the same fashion
as a packet arriving at a different router/NHS for the given LIS/LAG.
One method, among others, is to use the Server Cache Synchronization
Protocol (SCSP) [12]. It is RECOMMENDED that SCSP be the method used
when a LIS/LAG contains two or more router/NHSs.

During migration to NHRP, it cannot be expected that all routers
within the NBMA subnetwork are NHRP capable. Thus, NHRP traffic
which would otherwise need to be forwarded through such routers can
be expected to be dropped due to the NHRP packet not being
recognized. In this case, NHRP will be unable to establish any
transit paths whose discovery requires the traversal of the non-NHRP
speaking routers. If the client has tried and failed to acquire a
cut through path then the client should use the network layer routed
path as a default.

If an NBMA technology offers a group, an anycast, or a multicast
addressing feature then the NHC may be configured with such an
address (appropriate to the routing realm it participates in) which
would be assigned to all NHS serving that routing realm. This

address can then be used for establishing an initial connection to an
NHS to transmit a registration request. This address may not be used
for sending NHRP requests. The resulting VC may be used for NHRP
requests if and only if the registration response is received over
that VC, thereby indicating that one happens to have anycast
connected to an NHS serving the LIS/LAG. In the case of non-
connection oriented networks, or of multicast (rather than anycast)
addresses, the addres MUST NOT be used for sending NHRP resolution
requests.

When an NHS "serves" an NHC, the NHS MUST send NHRP messages destined
for the NHC directly to the NHC. That is, the NHRP message MUST NOT
transit through any NHS which is not serving the NHC when the NHRP
message is currently at an NHS which does serve the NHC (this, of
course, assumes the NHRP message is destined for the NHC). Further,
an NHS which serves an NHC SHOULD have a direct NBMA level connection
to that NHC (see Section 5.2.3 and 5.2.4 for examples).

With the exception of NHRP Registration Requests (see Section 5.2.3
and 5.2.4 for details of the NHRP Registration Request case), an NHC
MUST send NHRP messages over a direct NBMA level connection between
the serving NHS and the served NHC.

It may not be desirable to maintain semi-permanent NBMA level
connectivity between the NHC and the NHS.   In this case, when NBMA
level connectivity is initially setup between the NHS and the NHC (as
described in Section 5.2.4), the NBMA address of the NHS should be
obtained through the NBMA level signaling technology. This address
should be stored for future use in setting up subsequent NBMA level
connections. A somewhat more information rich technique to obtain
the address information (and more) of the serving NHS would be for
the NHC to include the Responder Address extension (see Section
5.3.1) in the NHRP Registration Request and to store the information
returned to the NHC in the Responder Address extension which is
subsequently included in the NHRP Registration Reply. Note also
that, in practice, a client's default router should also be its NHS;
thus a client may be able to know the NBMA address of its NHS from
the configuration which was already required for the client to be
able to communicate. Further, as mentioned in Section 4, NHCs may be
configured with the addressing information of one or more NHSs.

4. Configuration

Next Hop Clients

   An NHC connected to an NBMA subnetwork MAY be configured with the
   Protocol address(es) and NBMA address(es) of its NHS(s). The
   NHS(s) will likely also represent the NHC's default or peer

routers, so their NBMA addresses may be obtained from the NHC's
existing configuration. If the NHC is attached to several
subnetworks (including logical NBMA subnetworks), the NHC should
also be configured to receive routing information from its NHS(s)
and peer routers so that it can determine which internetwork layer
networks are reachable through which subnetworks.

Next Hop Servers

An NHS is configured with knowledge of its own internetwork layer
and NBMA addresses. An NHS MAY also be configured with a set of
internetwork layer address prefixes that correspond to the
internetwork layer addresses of the stations it serves. The NBMA
addresses of the stations served by the NHS may be learned via NHRP
Registration packets.

If a served NHC is attached to several subnetworks, the
router/route-server coresident with the serving NHS may also need
to be configured to advertise routing information to such NHCs.

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

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

5. NHRP Packet Formats

This section describes the format of NHRP packets. In the following,
unless otherwise stated explicitly, the unqualified term "request"
refers generically to any of the NHRP packet types which are
"requests". Further, unless otherwise stated explicitly, the
unqualified term "reply" refers generically to any of the NHRP packet
types which are "replies".

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

The length of the Fixed Part is fixed at 20 octets. The length of
the Mandatory Part is determined by the contents of the extensions
offset field (ar$extoff). If ar$extoff=0x0 then the mandatory part
length is equal to total packet length (ar$pktsz) minus 20 otherwise
the mandatory part length is equal to ar$extoff minus 20. The length

of the Extensions Part is implied by ar$pktsz minus ar$extoff. NHSs
may increase the size of an NHRP packet as a result of extension
processing, but not beyond the offered maximum packet size of the
NBMA network.

NHRP packets are actually members of a wider class of address mapping
and management protocols being developed by the IETF. A specific
encapsulation, based on the native formats used on the particular
NBMA network over which NHRP is carried, indicates the generic IETF
mapping and management protocol. For example, SMDS networks always
use LLC/SNAP encapsulation at the NBMA layer [4], and an NHRP packet
is preceded by the following LLC/SNAP encapsulation:

[0xAA-AA-03] [0x00-00-5E] [0x00-03]

The first three octets are LLC, indicating that SNAP follows. The
SNAP OUI portion is the IANA's OUI, and the SNAP PID portion
identifies the mapping and management protocol. A field in the Fixed
Header following the encapsulation indicates that it is NHRP.

ATM uses either LLC/SNAP encapsulation of each packet (including
NHRP), or uses no encapsulation on VCs dedicated to a single protocol
(see [7]). Frame Relay and X.25 both use NLPID/SNAP encapsulation or
identification of NHRP, using a NLPID of 0x0080 and the same SNAP
contents as above (see [8], [9]).

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

Most packet types (ar$op.type) have both internetwork layer
protocol-independent fields and protocol-specific fields. The
protocol type/snap fields (ar$pro.type/snap) qualify the format of
the protocol-specific fields.

5.1 NHRP Fixed Header

The Fixed Part of the NHRP packet contains those elements of the NHRP
packet which are always present and do not vary in size with the type
of 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
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |            ar$afn              |            ar$pro.type        |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                           ar$pro.snap                         |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |  ar$pro.snap  |    ar$hopcnt  |            ar$pktsz           |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |           ar$chksum           |            ar$extoff          |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     | ar$op.version |   ar$op.type  |    ar$shtl    |    ar$sstl    |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
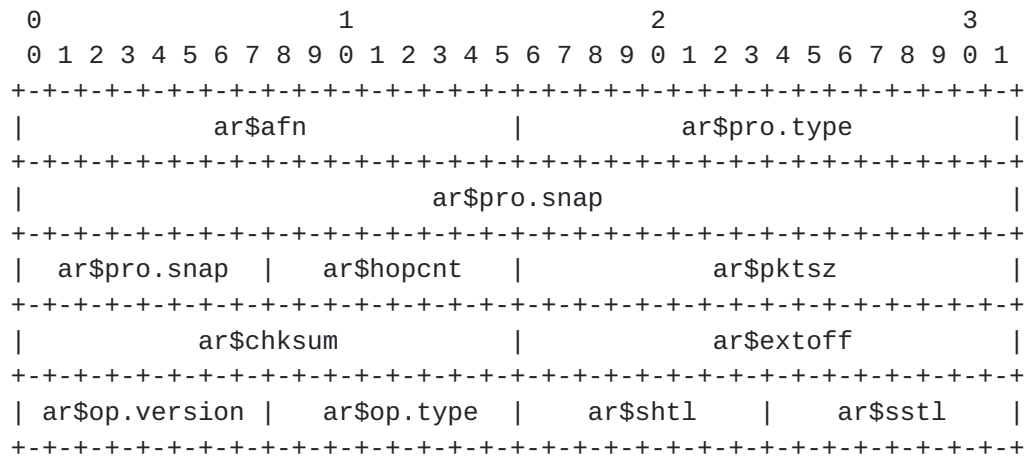
   ar$afn
     Defines the type of "link layer" addresses being carried. This
     number is taken from the 'address family number' list specified in
     [6]. This field has implications to the coding of ar$shtl and
     ar$sstl as described below.

   ar$pro.type
     field is a 16 bit unsigned integer representing the following
     number space:

        0x0000 to 0x00FF Protocols defined by the equivalent NLPIDs.
        0x0100 to 0x03FF Reserved for future use by the IETF.
        0x0400 to 0x04FF Allocated for use by the ATM Forum.
        0x0500 to 0x05FF Experimental/Local use.
        0x0600 to 0xFFFF Protocols defined by the equivalent Ethertypes.

     (based on the observations that valid Ethertypes are never smaller
     than 0x600, and NLPIDs never larger than 0xFF.)

   ar$pro.snap
     When ar$pro.type has a value of 0x0080, a SNAP encoded extension is
     being used to encode the protocol type. This snap extension is
     placed in the ar$pro.snap field. This is termed the 'long form'
     protocol ID. If ar$pro != 0x0080 then the ar$pro.snap field MUST be
     zero on transmit and ignored on receive. The ar$pro.type field
     itself identifies the protocol being referred to. This is termed
     the 'short form' protocol ID.

     In all cases, where a protocol has an assigned number in the
     ar$pro.type space (excluding 0x0080) the short form MUST be used
     when transmitting NHRP messages; i.e., if Ethertype or NLPID
     codings exist then they are used on transmit rather than the

   ethertype.   If both Ethertype and NLPID codings exist then when
   transmitting NHRP messages, the Ethertype coding MUST be used (this
   is consistent with RFC 1483 coding). So, for example, the
   following codings exist for IP:

     SNAP: ar$pro.type = 0x00-80, ar$pro.snap = 0x00-00-00-08-00
     NLPID: ar$pro.type = 0x00-CC, ar$pro.snap = 0x00-00-00-00-00
     Ethertype: ar$pro.type = 0x08-00, ar$pro.snap = 0x00-00-00-00-00

   and thus, since the Ethertype coding exists, it is used in
   preference.

 ar$hopcnt
   The Hop count indicates the maximum number of NHSs that an NHRP
   packet is allowed to traverse before being discarded. This field
   is used in a similar fashion to the way that a TTL is used in an IP
   packet and should be set accordingly. Each NHS decrements the TTL
   as the NHRP packet transits the NHS on the way to the next hop
   along the routed path to the destination. If an NHS receives an
   NHRP packet which it would normally forward to a next hop and that
   packet contains an ar$hopcnt set to zero then the NHS sends an
   error indication message back to the source protocol address
   stating that the hop count has been exceeded (see Section 5.2.7)
   and the NHS drops the packet in error; however, an error
   indication is never sent as a result of receiving an error
   indication. When a responding NHS replies to an NHRP request, that
   NHS places a value in ar$hopcnt as if it were sending a request of
   its own.

 ar$pktsz
   The total length of the NHRP packet, in octets (excluding link
   layer encapsulation).

 ar$chksum
   The standard IP checksum over the entire NHRP packet starting at
   the fixed header. If the packet is an odd number of bytes in
   length then this calculation is performed as if a byte set to 0x00
   is appended to the end of the packet.

 ar$extoff
   This field identifies the existence and location of NHRP
   extensions. If this field is 0 then no extensions exist otherwise
   this field represents the offset from the beginning of the NHRP
   packet (i.e., starting from the ar$afn field) of the first
   extension.

ar$op.version
  This field indicates what version of generic address mapping and
  management protocol is represented by this message.

      0   MARS protocol [11].
      1   NHRP as defined in this document.
      0x02 - 0xEF   Reserved for future use by the IETF.
      0xF0 - 0xFE   Allocated for use by the ATM Forum.
      0xFF   Experimental/Local use.

ar$op.type
  When ar$op.version == 1, this is the NHRP packet type: NHRP
  Resolution Request(1), NHRP Resolution Reply(2), NHRP Registration
  Request(3), NHRP Registration Reply(4), NHRP Purge Request(5), NHRP
  Purge Reply(6), or NHRP Error Indication(7). Use of NHRP packet
  Types in the range 128 to 255 are reserved for research or use in
  other protocol development and will be administered by IANA as
  described in Section 9.

ar$shtl
  Type & length of source NBMA address interpreted in the context of
  the 'address family number'[6] indicated by ar$afn. See below for
  more details.

ar$sstl
  Type & length of source NBMA subaddress interpreted in the context
  of the 'address family number'[6] indicated by ar$afn. When an
  NBMA technology has no concept of a subaddress, the subaddress
  length is always coded ar$sstl = 0 and no storage is allocated for
  the subaddress in the appropriate mandatory part. See below for
  more details.

Subnetwork layer address type/length fields (e.g., ar$shtl, Cli Addr
T/L) and subnetwork layer subaddresses type/length fields (e.g.,
ar$sstl, Cli SAddr T/L) are coded as follows:

 7 6 5 4 3 2 1 0
+-+-+-+-+-+-+-+-+
|0|x| length |
+-+-+-+-+-+-+-+-+

The most significant bit is reserved and MUST be set to zero. The
second most significant bit (x) is a flag indicating whether the
address being referred to is in:

  - NSAP format (x = 0).
  - Native E.164 format (x = 1).

For NBMA technologies that use neither NSAP nor E.164 format
addresses, x = 0 SHALL be used to indicate the native form for the
particular NBMA technology.

If the NBMA network is ATM and a subaddress (e.g., Source NBMA
SubAddress, Client NBMA SubAddress) is to be included in any part of
the NHRP packet then ar$afn MUST be set to 0x000F; further, the
subnetwork layer address type/length fields (e.g., ar$shtl, Cli Addr
T/L) and subnetwork layer subaddress type/length fields (e.g.,
ar$sstl, Cli SAddr T/L) MUST be coded as in [11]. If the NBMA
network is ATM and no subaddress field is to be included in any part
of the NHRP packet then ar$afn MAY be set to 0x0003 (NSAP) or 0x0008
(E.164) accordingly.

The bottom 6 bits is an unsigned integer value indicating the length
of the associated NBMA address in octets. If this value is zero the
flag x is ignored.

5.2.0 Mandatory Part

The Mandatory Part of the NHRP packet contains the operation specific
information (e.g., NHRP Resolution Request/Reply, etc.) and variable
length data which is pertinent to the packet type.

5.2.0.1 Mandatory Part Format

Sections 5.2.1 through 5.2.6 have a very similar mandatory part.
This mandatory part includes a common header and zero or more Client
Information Entries (CIEs). Section 5.2.7 has a different format
which is specified in that section.

The common header looks like the following:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Src Proto Len | Dst Proto Len |            Flags              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Request ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Source NBMA Address (variable length)              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Source NBMA Subaddress (variable length)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Source Protocol Address (variable length)           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Destination  Protocol Address (variable length)       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

And the CIEs 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Code       | Prefix Length |            unused             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Maximum Transmission Unit     |         Holding Time          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Cli Addr T/L | Cli SAddr T/L | Cli Proto Len |   Preference   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Client NBMA Address (variable length)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Client NBMA Subaddress (variable length)           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Client Protocol Address (variable length)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                     ....................
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Code       | Prefix Length |            unused             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Maximum Transmission Unit     |         Holding Time          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Cli Addr T/L | Cli SAddr T/L | Cli Proto Len |   Preference   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Client NBMA Address (variable length)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Client NBMA Subaddress (variable length)           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Client Protocol Address (variable length)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The meanings of the fields are as follows:

Src Proto Len
  This field holds the length in octets of the Source Protocol
  Address.

Dst Proto Len
  This field holds the length in octets of the Destination Protocol
  Address.

Flags
  These flags are specific to the given message type and they are
  explained in each section.

Request ID
  A value which, when coupled with the address of the source,
  provides a unique identifier for the information contained in a
  "request" packet. This value is copied directly from an "request"
  packet into the associated "reply". When a sender of a "request"
  receives "reply", it will compare the Request ID and source address
  information in the received "reply" against that found in its
  outstanding "request" list. When a match is found then the
  "request" is considered to be acknowledged.

  The value is taken from a 32 bit counter that is incremented each
  time a new "request" is transmitted. The same value MUST be used
  when resending a "request", i.e., when a "reply" has not been
  received for a "request" and a retry is sent after an appropriate
  interval.

  It is RECOMMENDED that the initial value for this number be 0. A
  node MAY reuse a sequence number if and only if the reuse of the
  sequence number is not precluded by use of a particular method of
  synchronization (e.g., as described in Appendix A).

The NBMA address/subaddress form specified below allows combined
E.164/NSAPA form of NBMA addressing. For NBMA technologies without a
subaddress concept, the subaddress field is always ZERO length and
ar$sstl = 0.

Source NBMA Address
  The Source NBMA address field is the address of the source station
  which is sending the "request". If the field's length as specified
  in ar$shtl is 0 then no storage is allocated for this address at
  all.

Source NBMA SubAddress
  The Source NBMA subaddress field is the address of the source
  station which is sending the "request". If the field's length as
  specified in ar$sstl is 0 then no storage is allocated for this
  address at all.

For those NBMA technologies which have a notion of "Calling Party
Addresses", the Source NBMA Addresses above are the addresses used
when signaling for an SVC.

"Requests" and "indications" follow the routed path from Source
Protocol Address to the Destination Protocol Address. "Replies", on
the other hand, follow the routed path from the Destination Protocol
Address back to the Source Protocol Address with the following

exceptions: in the case of a NHRP Registration Reply and in the case
of an NHC initiated NHRP Purge Request, the packet is always returned
via a direct VC (see Sections 5.2.4 and 5.2.5).

Source Protocol Address
  This is the protocol address of the station which is sending the
  "request". This is also the protocol address of the station toward
  which a "reply" packet is sent.

Destination Protocol Address
  This is the protocol address of the station toward which a
  "request" packet is sent.

Code
  This field is message specific. See the relevant message sections
  below. In general, this field is a NAK code; i.e., when the field
  is 0 in a reply then the packet is acknowledging a request and if
  it contains any other value the packet contains a negative
  acknowledgment.

Prefix Length
  This field is message specific. See the relevant message sections
  below. In general, however, this fields is used to indicate that
  the information carried in an NHRP message pertains to an
  equivalence class of internetwork layer addresses rather than just
  a single internetwork layer address specified. All internetwork
  layer addresses that match the first "Prefix Length" bit positions
  for the specific internetwork layer address are included in the
  equivalence class. If this field is set to 0x00 then this field
  MUST be ignored and no equivalence information is assumed (note
  that 0x00 is thus equivalent to 0xFF).

Maximum Transmission Unit
  This field gives the maximum transmission unit for the relevant
  client station. If this value is 0 then either the default MTU is
  used or the MTU negotiated via signaling is used if such
  negotiation is possible for the given NBMA.

Holding Time
  The Holding Time field specifies the number of seconds for which
  the Next Hop NBMA information specified in the CIE is considered to
  be valid. Cached information SHALL be discarded when the holding
  time expires. This field must be set to 0 on a NAK.

Cli Addr T/L
  Type & length of next hop NBMA address specified in the CIE. This
  field is interpreted in the context of the 'address family
  number'[6] indicated by ar$afn (e.g., ar$afn=0x0003 for ATM).

Cli SAddr T/L
  Type & length of next hop NBMA subaddress specified in the CIE.
  This field is interpreted in the context of the 'address family
  number'[6] indicated by ar$afn (e.g., ar$afn=0x0015 for ATM makes
  the address an E.164 and the subaddress an ATM Forum NSAP address).
  When an NBMA technology has no concept of a subaddress, the
  subaddress is always null with a length of 0. When the address
  length is specified as 0 no storage is allocated for the address.

Cli Proto Len
  This field holds the length in octets of the Client Protocol
  Address specified in the CIE.

Preference
  This field specifies the preference for use of the specific CIE
  relative to other CIEs. Higher values indicate higher preference.
  Action taken when multiple CIEs have equal or highest preference
  value is a local matter.

Client NBMA Address
  This is the client's NBMA address.

Client NBMA SubAddress
  This is the client's NBMA subaddress.

Client Protocol Address
  This is the client's internetworking layer address specified.

   Note that an NHS may cache source address binding information from an
   NHRP Resolution Request if and only if the conditions described in
   Section 6.2 are met for the NHS. In all other cases, source address
   binding information appearing in an NHRP message MUST NOT be cached.

5.2.1 NHRP Resolution Request

   The NHRP Resolution Request packet has a Type code of 1. Its
   mandatory part is coded as described in Section 5.2.0.1 and the
   message specific meanings of the fields are as follows:

   Flags - The flags field is coded as follows:

```
 0                       1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Q|A|D|U|S|        unused        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Q

  Set if the station sending the NHRP Resolution Request is a
  router; clear if the it is a host.

A

  This bit is set in a NHRP Resolution Request if only
  authoritative next hop information is desired and is clear
  otherwise. See the NHRP Resolution Reply section below for
  further details on the "A" bit and its usage.

D

  Unused (clear on transmit)

U

  This is the Uniqueness bit. This bit aids in duplicate address
  detection. When this bit is set in an NHRP Resolution Request
  and one or more entries exist in the NHS cache which meet the
  requirements of the NHRP Resolution Request then only the CIE in
  the NHS's cache with this bit set will be returned. Note that
  even if this bit was set at registration time, there may still be
  multiple CIEs that might fulfill the NHRP Resolution Request
  because an entire subnet can be registered through use of the
  Prefix Length in the CIE and the address of interest might be
  within such a subnet. If the "uniqueness" bit is set and the
  responding NHS has one or more cache entries which match the
  request but no such cache entry has the "uniqueness" bit set,
  then the NHRP Resolution Reply returns with a NAK code of "13 -
  Binding Exists But Is Not Unique" and no CIE is included. If a
  client wishes to receive non- unique Next Hop Entries, then
  the client must have the "uniqueness" bit set to zero in its NHRP
  Resolution Request. Note that when this bit is set in an NHRP
  Registration Request, only a single CIE may be specified in the
  NHRP Registration Request and that CIE must have the Prefix
  Length field set to 0xFF.

S

  Set if the binding between the Source Protocol Address and the
  Source NBMA information in the NHRP Resolution Request is
  guaranteed to be stable and accurate (e.g., these addresses are
  those of an ingress router which is connected to an ethernet stub
  network or the NHC is an NBMA attached host).

Zero or one CIEs (see Section 5.2.0.1) may be specified in an NHRP
Resolution Request. If one is specified then that entry carries the
pertinent information for the client sourcing the NHRP Resolution
Request. Usage of the CIE in the NHRP Resolution Request is
described below:

  Prefix Length
    If a CIE is specified in the NHRP Resolution Request then the
    Prefix Length field may be used to qualify the widest acceptable
    prefix which may be used to satisfy the NHRP Resolution Request.
    In the case of NHRP Resolution Request/Reply, the Prefix Length
    specifies the equivalence class of addresses which match the
    first "Prefix Length" bit positions of the Destination Protocol
    Address. If the "U" bit is set in the common header then this
    field MUST be set to 0xFF.

  Maximum Transmission Unit
    This field gives the maximum transmission unit for the source
    station. A possible use of this field in the NHRP Resolution
    Request packet is for the NHRP Resolution Requester to ask for a
    target MTU.

  Holding Time
    The Holding Time specified in the one CIE permitted to be
    included in an NHRP Resolution Request is the amount of time
    which the source address binding information in the NHRP
    Resolution Request is permitted to cached by transit and
    responding NHSs. Note that this field may only have a non-zero
    value if the S bit is set.

  All other fields in the CIE MUST be ignored and SHOULD be set to 0.

The Destination Protocol Address in the common header of the
Mandatory Part of this message contains the protocol address of the
station for which resolution is desired. An NHC MUST send the NHRP
Resolution Request directly to one of its serving NHSs (see Section 3
for more information).

5.2.2 NHRP Resolution Reply

The NHRP Resolution Reply packet has a Type code of 2. CIEs
correspond to Next Hop Entries in an NHS's cache which match the
criteria in the NHRP Resolution Request. Its mandatory part is coded
as described in Section 5.2.0.1. The message specific meanings of
the fields are as follows:

Flags - The flags field is coded as follows:

```
 0                       1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Q|A|D|U|S|        unused        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Q

  Copied from the NHRP Resolution Request. Set if the NHRP
  Resolution Requester is a router; clear if it is a host.

A

  Set if the next hop CIE in the NHRP Resolution Reply is
  authoritative; clear if the NHRP Resolution Reply is non-
  authoritative.

  When an NHS receives a NHRP Resolution Request for authoritative
  information for which it is the authoritative source, it MUST
  respond with a NHRP Resolution Reply containing all and only
  those next hop CIEs which are contained in the NHS's cache which
  both match the criteria of the NHRP Resolution Request and are
  authoritative cache entries. An NHS is an authoritative source
  for a NHRP Resolution Request if the information in the NHS's
  cache matches the NHRP Resolution Request criteria and that
  information was obtained through a NHRP Registration Request or
  through synchronization with an NHS which obtained this
  information through a NHRP Registration Request. An
  authoritative cache entry is one which is obtained through a NHRP
  Registration Request or through synchronization with an NHS which
  obtained this information through a NHRP Registration Request.

  An NHS obtains non-authoritative CIEs through promiscuous
  listening to NHRP packets other than NHRP Registrations which are
  directed at it. A NHRP Resolution Request which indicates a
  request for non-authoritative information should cause a NHRP
  Resolution Reply which contains all entries in the replying NHS's
  cache (i.e., both authoritative and non-authoritative) which
  match the criteria specified in the request.

D

  Set if the association between destination and the associate next
  hop information included in all CIEs of the NHRP Resolution Reply
  is guaranteed to be stable for the lifetime of the information
  (the holding time). This is the case if the Next Hop protocol
  address in a CIE identifies the destination (though it may be
  different in value than the Destination address if the
  destination system has multiple addresses) or if the destination
  is not connected directly to the NBMA subnetwork but the egress
  router to that destination is guaranteed to be stable (such as

when the destination is immediately adjacent to the egress router
through a non-NBMA interface).

U

This is the Uniqueness bit. See the NHRP Resolution Request
section above for details. When this bit is set, only one CIE is
included since only one unique binding should exist in an NHS's
cache.

S

Copied from NHRP Resolution Request message.

One or more CIEs are specified in the NHRP Resolution Reply. Each CIE
contains NHRP next hop information which the responding NHS has
cached and which matches the parameters specified in the NHRP
Resolution Request. If no match is found by the NHS issuing the NHRP
Resolution Reply then a single CIE is enclosed with the a CIE Code
set appropriately (see below) and all other fields MUST be ignored
and SHOULD be set to 0. In order to facilitate the use of NHRP by
minimal client implementations, the first CIE MUST contain the next
hop with the highest preference value so that such an implementation
need parse only a single CIE.

Code

If this field is set to zero then this packet contains a
positively acknowledged NHRP Resolution Reply. If this field
contains any other value then this message contains an NHRP
Resolution Reply NAK which means that an appropriate
internetworking layer to NBMA address binding was not available
in the responding NHS's cache. If NHRP Resolution Reply contains
a Client Information Entry with a NAK Code other than 0 then it
MUST NOT contain any other CIE. Currently defined NAK Codes are
as follows:

4 - Administratively Prohibited

An NHS may refuse an NHRP Resolution Request attempt for
administrative reasons (due to policy constraints or routing
state). If so, the NHS MUST send an NHRP Resolution Reply
which contains a NAK code of 4.

5 - Insufficient Resources

If an NHS cannot serve a station due to a lack of resources
(e.g., can't store sufficient information to send a purge if
routing changes), the NHS MUST reply with a NAKed NHRP
Resolution Reply which contains a NAK code of 5.

      12 - No Internetworking Layer Address to NBMA Address Binding
           Exists

         This code states that there were absolutely no internetworking
         layer address to NBMA address bindings found in the responding
         NHS's cache.

      13 - Binding Exists But Is Not Unique

         This code states that there were one or more internetworking
         layer address to NBMA address bindings found in the responding
         NHS's cache, however none of them had the uniqueness bit set.

   Prefix Length
      In the case of NHRP Resolution Reply, the Prefix Length specifies
      the equivalence class of addresses which match the first "Prefix
      Length" bit positions of the Destination Protocol Address.

   Holding Time
      The Holding Time specified in a CIE of an NHRP Resolution Reply
      is the amount of time remaining before the expiration of the
      client information which is cached at the replying NHS. It is
      not the value which was registered by the client.

   The remainder of the fields for the CIE for each next hop are
   filled out as they were defined when the next hop was registered
   with the responding NHS (or one of the responding NHS's
   synchronized servers) via the NHRP Registration Request.

Load-splitting may be performed when more than one Client Information
Entry is returned to a requester when equal preference values are
specified. Also, the alternative addresses may be used in case of
connectivity failure in the NBMA subnetwork (such as a failed call
attempt in connection-oriented NBMA subnetworks).

Any extensions present in the NHRP Resolution Request packet MUST be
present in the NHRP Resolution Reply even if the extension is non-
Compulsory.

If an unsolicited NHRP Resolution Reply packet is received, an Error
Indication of type Invalid NHRP Resolution Reply Received SHOULD be
sent in response.

When an NHS that serves a given NHC receives an NHRP Resolution Reply
destined for that NHC then the NHS must MUST send the NHRP Resolution
Reply directly to the NHC (see Section 3).

5.2.3 **NHRP Registration Request**

   The NHRP Registration Request is sent from a station to an NHS to
   notify the NHS of the station's NBMA information. It has a Type code
   of 3. Each CIE corresponds to Next Hop information which is to be
   cached at an NHS. The mandatory part of an NHRP Registration Request
   is coded as described in Section 5.2.0.1. The message specific
   meanings of the fields are as follows:

   Flags - The flags field is coded as follows:

```
    0                   1
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |U|         unused              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   U

      This is the Uniqueness bit. When set in an NHRP Registration
      Request, this bit indicates that the registration of the protocol
      address is unique within the confines of the set of synchronized
      NHSs. This "uniqueness" qualifier MUST be stored in the NHS/NHC
      cache. Any attempt to register a binding between the protocol
      address and an NBMA address when this bit is set MUST be rejected
      with a Code of "14 - Unique Internetworking Layer Address Already
      Registered" if the replying NHS already has a cache entry for the
      protocol address and the cache entry has the "uniqueness" bit
      set. A registration of a CIE's information is rejected when the
      CIE is returned with the Code field set to anything other than
      0x00. See the description of the uniqueness bit in NHRP
      Resolution Request section above for further details. When this
      bit is set only, only one CIE MAY be included in the NHRP
      Registration Request.

   Request ID
     The request ID has the same meaning as described in Section
     5.2.0.1. However, the request ID for NHRP Registrations which is
     maintained at each client MUST be kept in non-volatile memory so
     that when a client crashes and reregisters there will be no
     inconsistency in the NHS's database. In order to reduce the
     overhead associated with updating non-volatile memory, the actual
     updating need not be done with every increment of the Request ID
     but could be done, for example, every 50 or 100 increments. In
     this scenario, when a client crashes and reregisters it knows to
     add 100 to the value of the Request ID in the non-volatile memory
     before using the Request ID for subsequent registrations.

One or more CIEs are specified in the NHRP Registration Request.
Each CIE contains next hop information which a client is attempting
to register with its servers. Generally, all fields in CIEs enclosed
in NHRP Registration Requests are coded as described in Section
5.2.0.1. However, if a station is only registering itself with the
NHRP Registration Request then it MAY code the Cli Addr T/L, Cli
SAddr T/L, and Cli Proto Len as zero which signifies that the client
address information is to be taken from the source information in the
common header (see Section 5.2.0.1). Below, further clarification is
given for some fields in a CIE in the context of a NHRP Registration
Request.

   Code
     This field is set to 0x00 in NHRP Registration Requests.

   Prefix Length

     This field may be used in a NHRP Registration Request to register
     equivalence information for the Client Protocol Address specified
     in the CIE of an NHRP Registration Request In the case of NHRP
     Registration Request, the Prefix Length specifies the equivalence
     class of addresses which match the first "Prefix Length" bit
     positions of the Client Protocol Address. If the "U" bit is set
     in the common header then this field MUST be set to 0xFF.

The NHRP Registration Request is used to register an NHC's NHRP
information with its NHSs. If an NHC is configured with the protocol
address of a serving NHS then the NHC may place the NHS's protocol
address in the Destination Protocol Address field of the NHRP
Registration Request common header otherwise the NHC must place its
own protocol address in the Destination Protocol Address field.

When an NHS receives an NHRP Registration Request which has the
Destination Protocol Address field set to an address which belongs to
a LIS/LAG for which the NHS is serving then if the Destination
Protocol Address field is equal to the Source Protocol Address field
(which would happen if the NHC put its protocol address in the
Destination Protocol Address) or the Destination Protocol Address
field is equal to the protocol address of the NHS then the NHS
processes the NHRP Registration Request after doing appropriate error
checking (including any applicable policy checking).

When an NHS receives an NHRP Registration Request which has the
Destination Protocol Address field set to an address which does not
belong to a LIS/LAG for which the NHS is serving then the NHS
forwards the packet down the routed path toward the appropriate
LIS/LAG.

When an NHS receives an NHRP Registration Request which has the
Destination Protocol Address field set to an address which belongs to
a LIS/LAG for which the NHS is serving then if the Destination
Protocol Address field does not equal the Source Protocol Address
field and the Destination Protocol Address field does not equal the
protocol address of the NHS then the NHS forwards the message to the
appropriate NHS within the LIS/LAG as specified by Destination
Protocol Address field.

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 a NAK-ed Registration Reply 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 a NAK-ed Registration Reply of type Registration
Overflow.

In order to keep the registration entry from being discarded, the
station MUST re-send the NHRP Registration Request packet often
enough to refresh the registration, even in the face of occasional
packet loss. It is recommended that the NHRP Registration Request
packet be sent at an interval equal to one-third of the Holding Time
specified therein.

5.2.4 NHRP Registration Reply

The NHRP Registration Reply is sent by an NHS to a client in response
to that client's NHRP Registration Request. If the Code field of a
CIE in the NHRP Registration Reply has anything other than zero in it
then the NHRP Registration Reply is a NAK otherwise the reply is an
ACK. The NHRP Registration Reply has a Type code of 4.

An NHRP Registration Reply is formed from an NHRP Registration
Request by changing the type code to 4, updating the CIE Code field,
and filling in the appropriate extensions if they exist. The message
specific meanings of the fields are as follows:

Attempts to register the information in the CIEs of an NHRP
Registration Request may fail for various reasons. If this is the
case then each failed attempt to register the information in a CIE of
an NHRP Registration Request is logged in the associated NHRP
Registration Reply by setting the CIE Code field to the appropriate
error code as shown below:

CIE Code

   0 - Successful Registration

      The information in the CIE was successfully registered with the
      NHS.

   4 - Administratively Prohibited

      An NHS may refuse an NHRP Registration Request attempt for
      administrative reasons (due to policy constraints or routing
      state). If so, the NHS MUST send an NHRP Registration Reply
      which contains a NAK code of 4.

   5 - Insufficient Resources

      If an NHS cannot serve a station due to a lack of resources,
      the NHS MUST reply with a NAKed NHRP Registration Reply which
      contains a NAK code of 5.

   14 - Unique Internetworking Layer Address Already Registered
      If a client tries to register a protocol address to NBMA
      address binding with the uniqueness bit on and the protocol
      address already exists in the NHS's cache then if that cache
      entry also has the uniqueness bit on then this NAK Code is
      returned in the CIE in the NHRP Registration Reply.

Due to the possible existence of asymmetric routing, an NHRP
Registration Reply may not be able to merely follow the routed path
back to the source protocol address specified in the common header of
the NHRP Registration Reply. As a result, there MUST exist a direct
NBMA level connection between the NHC and its NHS on which to send
the NHRP Registration Reply before NHRP Registration Reply may be
returned to the NHC. If such a connection does not exist then the
NHS must setup such a connection to the NHC by using the source NBMA
information supplied in the common header of the NHRP Registration
Request.

5.2.5 **NHRP Purge Request**

The NHRP Purge Request packet is sent in order to invalidate cached
information in a station. The NHRP Purge Request packet has a type
code of 5. The mandatory part of an NHRP Purge Request is coded as
described in Section 5.2.0.1. The message specific meanings of the
fields are as follows:

Flags - The flags field is coded as follows:

```
 0                   1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|N|          unused             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

N

   When set, this bit tells the receiver of the NHRP Purge Request
   that the requester does not expect to receive an NHRP Purge
   Reply. If an unsolicited NHRP Purge Reply is received by a
   station where that station is identified in the Source Protocol
   Address of the packet then that packet must be ignored.

One or more CIEs are specified in the NHRP Purge Request. Each CIE
contains next hop information which is to be purged from an NHS/NHC
cache. Generally, all fields in CIEs enclosed in NHRP Purge Requests
are coded as described in Section 5.2.0.1. Below, further
clarification is given for some fields in a CIE in the context of a
NHRP Purge Request.

  Code
    This field is set to 0x00 in NHRP Purge Requests.

  Prefix Length

    In the case of NHRP Purge Requests, the Prefix Length specifies
    the equivalence class of addresses which match the first "Prefix
    Length" bit positions of the Client Protocol Address specified in
    the CIE. All next hop information which contains a protocol
    address which matches an element of this equivalence class is to
    be purged from the receivers cache.

  The Maximum Transmission Unit and Preference fields of the CIE are
  coded as zero. The Holding Time should be coded as zero but there
  may be some utility in supplying a "short" holding time to be
  applied to the matching next hop information before that
  information would be purged; this usage is for further study. The
  Client Protocol Address field and the Cli Proto Len field MUST be
  filled in. The Client Protocol Address is filled in with the
  protocol address to be purged from the receiving station's cache
  while the Cli Proto Len is set the length of the purged client's
  protocol address. All remaining fields in the CIE MAY be set to
  zero although the client NBMA information (and associated length
  fields) MAY be specified to narrow the scope of the NHRP Purge
  Request if requester desires. However, the receiver of an NHRP
  Purge Request may choose to ignore the Client NBMA information if
  it is supplied.

An NHRP Purge Request packet is sent from an NHS to a station to
cause it to delete previously cached information. This is done when
the information may be no longer valid (typically when the NHS has
previously provided next hop information for a station that is not
directly connected to the NBMA subnetwork, and the egress point to
that station may have changed).

An NHRP Purge Request packet may also be sent from an NHC to an NHS
with which the NHC had previously registered. This allows for an NHC
to invalidate its registration with NHRP before it would otherwise
expire via the holding timer. If an NHC does not have knowledge of a
protocol address of a serving NHS then the NHC must place its own
protocol address in the Destination Protocol Address field and
forward the packet along the routed path. Otherwise, the NHC must
place the protocol address of a serving NHS in this field.

Serving NHSs may need to send one or more new NHRP Purge Requests as
a result of receiving a purge from one of their served NHCs since the
NHS may have previously responded to NHRP Resolution Requests for
that NHC's NBMA information. These purges are "new" in that they are
sourced by the NHS and not the NHC; that is, for each NHC that
previously sent a NHRP Resolution Request for the purged NHC NBMA
information, an NHRP Purge Request is sent which contains the Source
Protocol/NBMA Addresses of the NHS and the Destination Protocol
Address of the NHC which previously sent an NHRP Resolution Request
prior to the purge.

The station sending the NHRP Purge Request MAY periodically
retransmit the NHRP Purge Request until either NHRP Purge Request is
acknowledged or until the holding time of the information being
purged has expired. Retransmission strategies for NHRP Purge Requests
are a local matter.

When a station receives an NHRP Purge Request, it MUST discard any
previously cached information that matches the information in the
CIEs.

An NHRP Purge Reply MUST be returned for the NHRP Purge Request even
if the station does not have a matching cache entry assuming that the
"N" bit is off in the NHRP Purge Request.

If the station wishes to reestablish communication with the
destination shortly after receiving an NHRP Purge Request, it should
make an authoritative NHRP Resolution Request in order to avoid any
stale cache entries that might be present in intermediate NHSs (See
section 6.2.2.). It is recommended that authoritative NHRP
Resolution Requests be made for the duration of the holding time of
the old information.

5.2.6 NHRP Purge Reply

   The NHRP Purge Reply packet is sent in order to assure the sender of
   an NHRP Purge Request that all cached information of the specified
   type has been purged from the station sending the reply. The NHRP
   Purge Reply has a type code of 6.

   An NHRP Purge Reply is formed from an NHRP Purge Request by merely
   changing the type code in the request to 6. The packet is then
   returned to the requester after filling in the appropriate extensions
   if they exist.

5.2.7 NHRP Error Indication

   The NHRP Error Indication is used to convey error indications to the
   sender of an NHRP packet. It has a type code of 7. 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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | Src Proto Len | Dst Proto Len |            unused             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |          Error Code           |         Error Offset          |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |             Source NBMA Address (variable length)            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |           Source NBMA Subaddress (variable length)           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |           Source Protocol Address (variable length)          |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |        Destination  Protocol Address (variable length)       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |        Contents of NHRP Packet in error (variable length)    |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   Src Proto Len
     This field holds the length in octets of the Source Protocol
     Address.

   Dst Proto Len
     This field holds the length in octets of the Destination Protocol
     Address.

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

    1 - Unrecognized Extension

      When the Compulsory bit of an extension in NHRP packet is set,
      the NHRP packet cannot be processed unless the extension has
      been processed. The responder MUST return an NHRP Error
      Indication of type Unrecognized Extension if it is incapable of
      processing the extension. However, if a transit NHS (one which
      is not going to generate a reply) detects an unrecognized
      extension, it SHALL ignore the extension.

    3 - NHRP Loop Detected

      A Loop Detected error is generated when it is determined that
      an NHRP packet is being forwarded in a loop.

    6 - Protocol Address Unreachable

      This error occurs when a packet it moving along the routed path
      and it reaches a point such that the protocol address of
      interest is not reachable.

    7 - Protocol Error

      A generic packet processing error has occurred (e.g., invalid
      version number, invalid protocol type, failed checksum, etc.)

    8 - NHRP SDU Size Exceeded

      If the SDU size of the NHRP packet exceeds the MTU size of the
      NBMA network then this error is returned.

    9 - Invalid Extension

      If an NHS finds an extension in a packet which is inappropriate
      for the packet type, an error is sent back to the sender with
      Invalid Extension as the code.

    10 - Invalid NHRP Resolution Reply Received

      If a client receives a NHRP Resolution Reply for a Next Hop
      Resolution Request which it believes it did not make then an
      error packet is sent to the station making the reply with an
      error code of Invalid Reply Received.

   11 - Authentication Failure

      If a received packet fails an authentication test then this
      error is returned.

   15 - Hop Count Exceeded

      The hop count which was specified in the Fixed Header of an
      NHRP message has been exceeded.

Error Offset
  The offset in octets into the original NHRP packet in which an
  error was detected. This offset is calculated starting from the
  NHRP Fixed Header.

Source NBMA Address
  The Source NBMA address field is the address of the station which
  observed the error.

Source NBMA SubAddress
  The Source NBMA subaddress field is the address of the station
  which observed the error. If the field's length as specified in
  ar$sstl is 0 then no storage is allocated for this address at all.

Source Protocol Address
  This is the protocol address of the station which issued the Error
  packet.

Destination Protocol Address
  This is the protocol address of the station which sent the packet
  which was found to be in error.

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

If an NHS sees its own Protocol and NBMA Addresses in the Source NBMA
and Source Protocol address fields of a transiting NHRP Error
Indication packet then the NHS will quietly drop the packet and do
nothing (this scenario would occur when the NHRP Error Indication
packet was itself in a loop).

Note that no extensions may be added to an NHRP Error Indication.

5.3 Extensions Part

The Extensions Part, if present, carries one or more extensions in
{Type, Length, Value} triplets.

Extensions 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|C|u|        Type            |              Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Value...                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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

u
  Unused and must be set to zero.

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

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

When extensions exist, the extensions list is terminated by the Null
TLV, having Type = 0 and Length = 0.

Extensions may occur in any order, but any particular extension type
may occur only once in an NHRP packet unless explicitly stated to the
contrary in the extensions definition. For example, the vendor-
private extension may occur multiple times in a packet in order to
allow for extensions which do not share the same vendor ID to be
represented. It is RECOMMENDED that a given vendor include no more
than one Vendor Private Extension.

An NHS MUST NOT change the order of extensions. That is, the order
of extensions placed in an NHRP packet by an NHC (or by an NHS when
an NHS sources a packet) MUST be preserved as the packet moves
between NHSs. Minimal NHC implementations MUST only recognize, but

not necessarily parse, the Vendor Private extension and the End Of
Extensions extension. Extensions are only present in a "reply" if
they were present in the corresponding "request" with the exception
of Vendor Private extensions. The previous statement is not intended
to preclude the creation of NHS-only extensions which might be added
to and removed from NHRP packets by the same NHS; such extensions
MUST not be propagated to NHCs.

The Compulsory bit provides for a means to add to the extension set.
If the bit is set in an extension then the station responding to the
NHRP message which contains that extension MUST be able to understand
the extension (in this case, the station responding to the message is
the station that would issue an NHRP reply in response to a NHRP
request). As a result, the responder MUST return an NHRP Error
Indication of type Unrecognized Extension. If the Compulsory bit is
clear then the extension can be safely ignored; however, if an
ignored extension is in a "request" then it MUST be returned,
unchanged, in the corresponding "reply" packet type.

If a transit NHS (one which is not going to generate a "reply")
detects an unrecognized extension, it SHALL ignore the extension. If
the Compulsory bit is set, the transit NHS MUST NOT cache the
information contained in the packet and MUST NOT identify itself as
an egress router (in the Forward Record or Reverse Record
extensions). Effectively, this means, if a transit NHS encounters an
extension which it cannot process and which has the Compulsory bit
set then that NHS MUST NOT participate in any way in the protocol
exchange other than acting as a forwarding agent.

The NHRP extension Type space is subdivided to encourage use outside
the IETF.

```
  0x0000 - 0x0FFF          Reserved for NHRP.
  0x1000 - 0x11FF          Allocated to the ATM Forum.
  0x1200 - 0x37FF          Reserved for the IETF.
  0x3800 - 0x3FFF          Experimental use.
```

IANA will administer the ranges reserved for the IETF as described in
Section 9. Values in the 'Experimental use' range have only local
significance.

5.3.0 The End Of Extensions

```
 Compulsory = 1
 Type = 0
 Length = 0
```

When extensions exist, the extensions list is terminated by the End
Of Extensions/Null TLV.

5.3.1 Responder Address Extension

     Compulsory = 1
     Type = 3
     Length = variable

This extension is used to determine the address of the NHRP
responder; i.e., the entity that generates the appropriate "reply"
packet for a given "request" packet. In the case of an NHRP
Resolution Request, the station responding may be different (in the
case of cached replies) than the system identified in the Next Hop
field of the NHRP Resolution Reply. Further, this extension may aid
in detecting loops in the NHRP forwarding path.

This extension uses a single CIE with the extension specific meanings
of the fields set as follows:

The Prefix Length fields MUST be set to 0 and ignored.

CIE Code
  5 - Insufficient Resources
    If the responder to an NHRP Resolution Request is an egress point
    for the target of the address resolution request (i.e., it is one
    of the stations identified in the list of CIEs in an NHRP
    Resolution Reply) and the Responder Address extension is included
    in the NHRP Resolution Request and insufficient resources to
    setup a cut-through VC exist at the responder then the Code field
    of the Responder Address Extension is set to 5 in order to tell
    the client that a VC setup attempt would in all likelihood be
    rejected; otherwise this field MUST be coded as a zero. NHCs MAY
    use this field to influence whether they attempt to setup a cut-
    through to the egress router.

Maximum Transmission Unit
  This field gives the maximum transmission unit preferred by the
  responder. If this value is 0 then either the default MTU is used
  or the MTU negotiated via signaling is used if such negotiation is
  possible for the given NBMA.

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

"Client Address" information is actually "Responder Address"
information for this extension. Thus, for example, Cli Addr T/L is
the responder NBMA address type and length field.

If a "requester" desires this information, the "requester" SHALL
include this extension with a value of zero. Note that this implies
that no storage is allocated for the Holding Time and Type/Length
fields until the "Value" portion of the extension is filled out.

If an NHS is generating a "reply" packet in response to a "request"
containing this extension, the NHS SHALL include this extension,
containing its protocol address in the "reply". If an NHS has more
than one protocol address, it SHALL use the same protocol address
consistently in all of the Responder Address, Forward Transit NHS
Record, and Reverse Transit NHS Record extensions. The choice of
which of several protocol address to include in this extension is a
local matter.

If an NHRP Resolution Reply packet being forwarded by an NHS contains
a protocol address of that NHS in the Responder Address Extension
then that NHS SHALL generate an NHRP Error Indication of type "NHRP
Loop Detected" and discard the NHRP Resolution Reply.

If an NHRP Resolution Reply packet is being returned by an
intermediate NHS based on cached data, it SHALL place its own address
in this extension (differentiating it from the address in the Next
Hop field).

5.3.2 NHRP Forward Transit NHS Record Extension

```
 Compulsory = 1
 Type = 4
 Length = variable
```

The NHRP Forward Transit NHS record contains a list of transit NHSs
through which a "request" has traversed. Each NHS SHALL append to
the extension a Forward Transit NHS element (as specified below)
containing its Protocol address. The extension length field and the
ar$chksum fields SHALL be adjusted appropriately.

The responding NHS, as described in Section 5.3.1, SHALL NOT update
this extension.

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.

This extension uses a single CIE per NHS Record element with the
extension specific meanings of the fields set as follows:

The Prefix Length fields MUST be set to 0 and ignored.

CIE Code
  5 - Insufficient Resources
    If an NHRP Resolution Request contains an NHRP Forward Transit
    NHS Record Extension and insufficient resources to setup a cut-
    through VC exist at the current transit NHS then the CIE Code
    field for NHRP Forward Transit NHS Record Extension is set to 5
    in order to tell the client that a VC setup attempt would in all
    likelihood be rejected; otherwise this field MUST be coded as a
    zero. NHCs MAY use this field to influence whether they attempt
    to setup a cut-through as described in Section 2.2. Note that
    the NHRP Reverse Transit NHS Record Extension MUST always have
    this field set to zero.

Maximum Transmission Unit
  This field gives the maximum transmission unit preferred by the
  transit NHS. If this value is 0 then either the default MTU is
  used or the MTU negotiated via signaling is used if such
  negotiation is possible for the given NBMA.

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

"Client Address" information is actually "Forward Transit NHS
Address" information for this extension. Thus, for example, Cli Addr
T/L is the transit NHS NBMA address type and length field.

If a "requester" wishes to obtain this information, it SHALL include
this extension with a length of zero. Note that this implies that no
storage is allocated for the Holding Time and Type/Length fields
until the "Value" portion of the extension is filled out.

If an NHS has more than one Protocol address, it SHALL use the same
Protocol address consistently in all of the Responder Address,
Forward NHS Record, and Reverse NHS Record extensions. The choice of
which of several Protocol addresses to include in this extension is a
local matter.

If a "request" that is being forwarded by an NHS contains the
Protocol Address of that NHS in one of the Forward Transit NHS
elements then the NHS SHALL generate an NHRP Error Indication of type
"NHRP Loop Detected" and discard the "request".

5.3.3 NHRP Reverse Transit NHS Record Extension

     Compulsory = 1
     Type = 5
     Length = variable

The NHRP Reverse Transit NHS record contains a list of transit NHSs
through which a "reply" has traversed. Each NHS SHALL append a
Reverse Transit NHS element (as specified below) containing its
Protocol address to this extension. The extension length field and
ar$chksum SHALL be adjusted appropriately.

The responding NHS, as described in Section 5.3.1, SHALL NOT update
this extension.

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.

This extension uses a single CIE per NHS Record element with the
extension specific meanings of the fields set as follows:

The CIE Code and Prefix Length fields MUST be set to 0 and ignored.

Maximum Transmission Unit
  This field gives the maximum transmission unit preferred by the
  transit NHS. If this value is 0 then either the default MTU is
  used or the MTU negotiated via signaling is used if such
  negotiation is possible for the given NBMA.

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

"Client Address" information is actually "Reverse Transit NHS
Address" information for this extension. Thus, for example, Cli Addr
T/L is the transit NHS NBMA address type and length field.

If a "requester" wishes to obtain this information, it SHALL include
this extension with a length of zero. Note that this implies that no
storage is allocated for the Holding Time and Type/Length fields
until the "Value" portion of the extension is filled out.

If an NHS has more than one Protocol address, it SHALL use the same
Protocol address consistently in all of the Responder Address,
Forward NHS Record, and Reverse NHS Record extensions. The choice of
which of several Protocol addresses to include in this extension is a
local matter.

If a "reply" that is being forwarded by an NHS contains the Protocol
Address of that NHS in one of the Reverse Transit NHS elements then
the NHS SHALL generate an NHRP 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.

5.3.4 NHRP Authentication Extension

Compulsory = 1 Type = 7 Length = variable

The NHRP Authentication Extension is carried in NHRP packets to
convey authentication information between NHRP speakers. The
Authentication Extension may be included in any NHRP "request" or
"reply" only.

The authentication is always done pairwise on an NHRP hop-by-hop
basis; i.e., the authentication extension is regenerated at each
hop. If a received packet fails the authentication test, the station
SHALL generate an Error Indication of type "Authentication Failure"
and discard the packet. Note that one possible authentication failure
is the lack of an Authentication Extension; the presence or absence
of the Authentication Extension is a local matter.

5.3.4.1 Header Format

The authentication header 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Reserved                  | Security Parameter Index (SPI)|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Src Addr...                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                              |
+-+-+-+-+-+-+-+-+-+-+ Authentication Data... -+-+-+-+-+-+-+-+-+-+
|                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Security Parameter Index (SPI) can be thought of as an index into a
table that maintains the keys and other information such as hash
algorithm. Src and Dst communicate either offline using manual keying
or online using a key management protocol to populate this table. The
sending NHRP entity always allocates the SPI and the parameters
associated with it.

Src Addr a variable length field is the address assigned to the
outgoing interface. The length of the addr is obtained from the
source protocol length field in the mandatory part of the NHRP
header. The tuple <spi, src addr> uniquely identifies the key and
other parameters that are used in authentication.

The length of the authentication data field is dependent on the hash
algorithm used. The data field contains the keyed hash calculated
over the entire NHRP payload. The authentication data field is zeroed
out before the hash is calculated.

5.3.4.2 SPI and Security Parameters Negotiation

SPI's can be negotiated either manually or using an Internet Key
Management protocol. Manual keying MUST be supported. The following
parameters are associated with the tuple <SPI, src>- lifetime,
Algorithm, Key. Lifetime indicates the duration in seconds for which
the key is valid. In case of manual keying, this duration can be
infinite. Also, in order to better support manual keying, there may
be multiple tuples active at the same time (Dst being the same).

Algorithm specifies the hash algorithm agreed upon by the two
entities. HMAC-MD5-128 [16] is the default algorithm. Other
algorithms MAY be supported by defining new values. IANA will assign
the numbers to identify the algorithm being used as described in
Section 9.

Any Internet standard key management protocol MAY so be used to
negotiate the SPI and parameters.

5.3.4.3 **Message Processing**

   At the time of adding the authentication extension header, src looks
   up in a table to fetch the SPI and the security parameters based on
   the outgoing interface address. If there are no entries in the table
   and if there is support for key management, the src initiates the key
   management protocol to fetch the necessary parameters. The src
   constructs the Authentication Extension payload and calculates the
   hash by zeroing authentication data field. The result replaces in the
   zeroed authentication data field. The src address field in the
   payload is the IP address assigned to the outgoing interface.

   If key management is not supported and authentication is mandatory,
   the packet is dropped and this information is logged.

   On the receiving end, dst fetches the parameters based on the SPI and
   the ip address in the authentication extension payload. The
   authentication data field is extracted before zeroing out to
   calculate the hash. It computes the hash on the entire payload and if
   the hash does not match, then an "abnormal event" has occurred.

5.3.4.4 **Security Considerations**

   It is important that the keys chosen are strong as the security of
   the entire system depends on the keys being chosen properly and the
   correct implementation of the algorithms.

   The security is performed on a hop by hop basis. The data received
   can be trusted only so much as one trusts all the entities in the
   path traversed. A chain of trust is established amongst NHRP entities
   in the path of the NHRP Message . If the security in an NHRP entity
   is compromised, then security in the entire NHRP domain is
   compromised.

   Data integrity covers the entire NHRP payload. This guarantees that
   the message was not modified and the source is authenticated as well.
   If authentication extension is not used or if the security is
   compromised, then NHRP entities are liable to both spoofing attacks,
   active attacks and passive attacks.

   There is no mechanism to encrypt the messages. It is assumed that a
   standard layer 3 confidentiality mechanism will be used to encrypt
   and decrypt messages. It is recommended to use an Internet standard
   key management protocol to negotiate the keys between the neighbors.
   Transmitting the keys in clear text, if other methods of negotiation
   is used, compromises the security completely.

   Any NHS is susceptible to Denial of Service (DOS) attacks that cause
   it to become overloaded, preventing legitimate packets from being
   acted upon properly. A rogue host can send request and registration
   packets to the first hop NHS. If the authentication option is not
   used, the registration packet is forwarded along the routed path
   requiring processing along each NHS. If the authentication option is
   used, then only the first hop NHS is susceptible to DOS attacks
   (i.e., unauthenticated packets will be dropped rather than forwarded
   on). If security of any host is compromised (i.e., the keys it is
   using to communicate with an NHS become known), then a rogue host can
   send NHRP packets to the first hop NHS of the host whose keys were
   compromised, which will then forward them along the routed path as in
   the case of unauthenticated packets. However, this attack requires
   that the rogue host to have the same first hop NHS as that of the
   compromised host. Finally, it should be noted that denial of service
   attacks that cause routers on the routed path to expend resources
   processing NHRP packets are also susceptable to attacks that flood
   packets at the same destination as contained in an NHRP packet's
   Destination Protocol Address field.

5.3.5 **NHRP Vendor-Private Extension**

    Compulsory = 0
    Type = 8
    Length = variable

   The NHRP Vendor-Private Extension is carried in NHRP packets to
   convey vendor-private information or NHRP extensions between NHRP
   speakers.

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                   Vendor ID            | Data....      |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   Vendor ID
     802 Vendor ID as assigned by the IEEE [6]

   Data
     The remaining octets after the Vendor ID in the payload are
     vendor-dependent data.

   This extension may be added to any "request" or "reply" packet and it
   is the only extension that may be included multiple times. If the
   receiver does not handle this extension, or does not match the Vendor

   ID in the extension then the extension may be completely ignored by
   the receiver. If a Vendor Private Extension is included in a
   "request" then it must be copied to the corresponding "reply".

6. Protocol Operation

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

6.1 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. In particular, attempting to establish an NHRP
   path across a boundary where information used in route selection is
   lost may result in a routing loop. Such situations include the loss
   of BGP path vector information, the interworking of multiple routing
   protocols with dissimilar metrics (e.g, RIP and OSPF), etc. In such
   circumstances, NHRP should not be used. This situation can be
   avoided if there are no "back door" paths between the entry and
   egress router outside of the NBMA subnetwork. Protocol mechanisms to
   relax these restrictions are under investigation.

   In general it is preferable to use mechanisms, if they exist, in
   routing protocols to resolve the egress point when the destination
   lies outside of the NBMA subnetwork, since such mechanisms will be
   more tightly coupled to the state of the routing system and will
   probably be less likely to create loops.

6.2 Cache Management Issues

   The management of NHRP caches in the source station, the NHS serving
   the destination, and any intermediate NHSs is dependent on a number
   of factors.

6.2.1 Caching Requirements

   Source Stations

      Source stations MUST cache all received NHRP Resolution Replies
      that they are actively using. They also must cache "incomplete"
      entries, i.e., those for which a NHRP Resolution Request has been
      sent but those for which an NHRP Resolution Reply has not been
      received. This is necessary in order to preserve the Request ID

for retries, and provides the state necessary to avoid triggering
NHRP Resolution Requests for every data packet sent to the
destination.

Source stations MUST purge expired information from their caches.
Source stations MUST purge the appropriate cached information upon
receipt of an NHRP Purge Request packet.

When a station has a co-resident NHC and NHS, the co-resident NHS
may reply to NHRP Resolution Requests from the co-resident NHC with
information which the station cached as a result of the co-resident
NHC making its own NHRP Resolution Requests as long as the co-
resident NHS follows the rules for Transit NHSs as seen below.

Serving NHSs

  The NHS serving the destination (the one which responds
  authoritatively to NHRP Resolution Requests) SHOULD cache protocol
  address information from all NHRP Resolution Requests to which it
  has responded if the information in the NHRP Resolution Reply has
  the possibility of changing during its lifetime (so that an NHRP
  Purge Request packet can be issued). The internetworking to NBMA
  binding information provided by the source station in the NHRP
  Resolution Request may also be cached if and only if the "S" bit is
  set, the NHRP Resolution Request has included a CIE with the
  Holding Time field set greater than zero (this is the valid Holding
  Time for the source binding), and only for non-authoritative use
  for a period not to exceed the Holding Time.

Transit NHSs

  A Transit NHS (lying along the NHRP path between the source station
  and the responding NHS) may cache source binding information
  contained in NHRP Resolution Request packets that it forwards if
  and only if the "S" bit is set, the NHRP Resolution Request has
  included a CIE with the Holding Time field set greater than zero
  (this is the valid Holding Time for the source binding), and only
  for non-authoritative use for a period not to exceed the Holding
  Time.

  A Transit NHS may cache destination information contained in NHRP
  Resolution Reply CIE if only if the D bit is set and then only for
  non-authoritative use for a period not to exceed the Holding Time
  value contained in the CIE. A Transit NHS MUST NOT cache source
  binding information contained in an NHRP Resolution Reply.

Further, a transit NHS MUST discard any cached information when the
prescribed time has expired. It may return cached information in
response to non-authoritative NHRP Resolution Requests only.

6.2.2 Dynamics of Cached Information

NBMA-Connected Destinations

NHRP's most basic function is that of simple NBMA address
resolution of stations directly attached to the NBMA subnetwork.
These mappings are typically very static, and appropriately chosen
holding times will minimize problems in the event that the NBMA
address of a station must be changed. Stale information will cause
a loss of connectivity, which may be used to trigger an
authoritative NHRP Resolution Request and bypass the old data. In
the worst case, connectivity will fail until the cache entry times
out.

This applies equally to information marked in NHRP Resolution
Replies as being "stable" (via the "D" bit).

Destinations Off of the NBMA Subnetwork

If the source of an NHRP Resolution Request is a host and the
destination is not directly attached to the NBMA subnetwork, and
the route to that destination is not considered to be "stable," the
destination mapping may be very dynamic (except in the case of a
subnetwork where each destination is only singly homed to the NBMA
subnetwork). As such the cached information may very likely become
stale. The consequence of stale information in this case will be a
suboptimal path (unless the internetwork has partitioned or some
other routing failure has occurred).

6.3 Use of the Prefix Length field of a CIE

A certain amount of care needs to be taken when using the Prefix
Length field of a CIE, in particular with regard to the prefix length
advertised (and thus the size of the equivalence class specified by
it). Assuming that the routers on the NBMA subnetwork are exchanging
routing information, it should not be possible for an NHS to create a
black hole by advertising too large of a set of destinations, but
suboptimal routing (e.g., extra internetwork layer hops through the
NBMA) can result. To avoid this situation an NHS that wants to send
the Prefix Length MUST obey the following rule:

The NHS examines the Network Layer Reachability Information (NLRI)
associated with the route that the NHS would use to forward towards
the destination (as specified by the Destination internetwork layer

address in the NHRP Resolution Request), and extracts from this
NLRI the shortest address prefix such that: (a) the Destination
internetwork layer address (from the NHRP Resolution Request) is
covered by the prefix, (b) the NHS does not have any routes with
NLRI which form a subset of what is covered by the prefix. The
prefix may then be used in the CIE.

The Prefix Length field of the CIE should be used with restraint, in
order to avoid NHRP stations choosing suboptimal transit paths when
overlapping prefixes are available. This document specifies the use
of the prefix length only when all the destinations covered by the
prefix are "stable". That is, either:

   (a) All destinations covered by the prefix are on the NBMA network,
       or
   (b) All destinations covered by the prefix are directly attached to
       the NHRP responding station.

Use of the Prefix Length field of the CIE in other circumstances is
outside the scope of this document.

6.4 Domino Effect

One could easily imagine a situation where a router, acting as an
ingress station to the NBMA subnetwork, receives a data packet, such
that this packet triggers an NHRP Resolution Request. If the router
forwards this data packet without waiting for an NHRP transit path to
be established, then when the next router along the path receives the
packet, the next router may do exactly the same - originate its own
NHRP Resolution Request (as well as forward the packet). In fact
such a data packet may trigger NHRP Resolution Request generation at
every router along the path through an NBMA subnetwork. We refer to
this phenomena as the NHRP "domino" effect.

The NHRP domino effect is clearly undesirable. At best it may result
in excessive NHRP traffic. At worst it may result in an excessive
number of virtual circuits being established unnecessarily.
Therefore, it is important to take certain measures to avoid or
suppress this behavior. NHRP implementations for NHSs MUST provide a
mechanism to address this problem. One possible strategy to address
this problem would be to configure a router in such a way that NHRP
Resolution Request generation by the router would be driven only by
the traffic the router receives over its non-NBMA interfaces
(interfaces that are not attached to an NBMA subnetwork). Traffic
received by the router over its NBMA-attached interfaces would not
trigger NHRP Resolution Requests. Such a router avoids the NHRP
domino effect through administrative means.

7. NHRP over Legacy BMA Networks

   There would appear to be no significant impediment to running NHRP
   over legacy broadcast subnetworks. There may be issues around
   running NHRP across multiple subnetworks. Running NHRP on broadcast
   media has some interesting possibilities; especially when setting up
   a cut-through for inter-ELAN inter-LIS/LAG traffic when one or both
   end stations are legacy attached. This use for NHRP requires further
   research.

8. Discussion

   The result of an NHRP Resolution Request depends on how routing is
   configured among the NHSs of an NBMA subnetwork. If the destination
   station is directly connected to the NBMA subnetwork and the routed
   path to it lies entirely within the NBMA subnetwork, the NHRP
   Resolution Replies always return the NBMA address of the destination
   station itself rather than the NBMA address of some egress router.
   On the other hand, if the routed path exits the NBMA subnetwork, NHRP
   will be unable to resolve the NBMA address of the destination, but
   rather will return the address of the egress router. For
   destinations outside the NBMA subnetwork, egress routers and routers
   in the other subnetworks should exchange routing information so that
   the optimal egress router may be found.

   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 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
   internetwork layer switching functionality.

9. IANA Considerations

   IANA will take advice from the Area Director appointed designated
   subject matter expert, in order to assign numbers from the various
   number spaces described herein. In the event that the Area Director
   appointed designated subject matter expert is unavailable, the
   relevant IESG Area Director will appoint another expert. Any and all
   requests for value assignment within a given number space will be
   accepted when the usage of the value assignment documented. Possible
   forms of documentantion include, but is not limited to, RFCs or the
   product of another cooperative standards body (e.g., the MPOA and
   LANE subworking group of the ATM Forum).

References

   [1] Heinanen, J., and R. Govindan, "NBMA Address Resolution Protocol
   (NARP)", RFC 1735, December 1994.

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

   [3] Laubach, M., and J. Halpern, "Classical IP and ARP over ATM", RFC
   2225, April 1998.

   [4] Piscitello,, D., and J. Lawrence, "Transmission of IP datagrams
   over the SMDS service", RFC 1209, March 1991.

   [5] Protocol Identification in the Network Layer, ISO/IEC TR
   9577:1990.

   [6] Reynolds, J., and J. Postel, "Assigned Numbers", STD 2, RFC 1700,
   October 1994.

   [7] Heinanen, J., "Multiprotocol Encapsulation over ATM Adaptation
   Layer 5", RFC 1483, July 1993.

   [8] Malis, A., Robinson, D., and R. Ullmann, "Multiprotocol
   Interconnect on X.25 and ISDN in the Packet Mode", RFC 1356, August
   1992.

   [9] Bradley, T., Brown, C., and A. Malis, "Multiprotocol Interconnect
   over Frame Relay", RFC 1490, July 1993.

   [10] Rekhter, Y., and D. Kandlur, ""Local/Remote" Forwarding Decision
   in Switched Data Link Subnetworks", RFC 1937, May 1996.

   [11] Armitage, G., "Support for Multicast over UNI 3.0/3.1 based ATM
   Networks", RFC 2022, November 1996.

   [12] Luciani, J., Armitage, G., and J. Halpern, "Server Cache
   Synchronization Protocol (SCSP) - NBMA", RFC 2334, April 1998.

   [13] Rekhter, Y., "NHRP for Destinations off the NBMA Subnetwork",
   Work In Progress.

   [14] Luciani, J., et. al., "Classical IP and ARP over ATM to NHRP
   Transition", Work In Progress.

   [15] Bradner, S., "Key words for use in RFCs to Indicate Requirement
   Levels", BCP 14, RFC 2119, March 1997.

   [16] Krawczyk, H., Bellare, M., and R. Canetti, "HMAC: Keyed Hashing
        for Message Authentication", RFC 2104, February 1997.

Acknowledgments

Authors' Addresses

   James V. Luciani   Dave Katz
   Bay Networks   cisco Systems
   3 Federal Street   170 W. Tasman Dr.
   Mail Stop: BL3-03   San Jose, CA 95134 USA
   Billerica, MA 01821   Phone: +1 408 526 8284
   Phone: +1 978 916 4734   EMail: dkatz@cisco.com
   EMail: luciani@baynetworks.com

   David Piscitello   Bruce Cole
   Core Competence   Juniper Networks
   1620 Tuckerstown Road   3260 Jay St.
   Dresher, PA 19025 USA   Santa Clara, CA 95054
   Phone: +1 215 830 0692   Phone: +1 408 327 1900
   EMail: dave@corecom.com   EMail: bcole@jnx.com

   Naganand Doraswamy
   Bay Networks, Inc.
   3 Federal Street
   Mail Stop: Bl3-03
   Billerica, MA 01801
   Phone: +1 978 916 1323
   EMail: naganand@baynetworks.com

Full Copyright Statement