

NHRP for Destinations off the NBMA Subnetwork

[draft-ietf-rolc-r2r-nhrp-00.txt](#)

1. Status of this Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as ``work in progress.''

To learn the current status of any Internet-Draft, please check the ``1id-abstracts.txt' listing contained in the Internet-Drafts Shadow Directories on ftp.is.co.za (Africa), nic.nordu.net (Europe), munnari.oz.au (Pacific Rim), ds.internic.net (US East Coast), or ftp.isi.edu (US West Coast).

2. Abstract

The NBMA Next Hop Resolution Protocol (NHRP) [[NHRP](#)] specifies a mechanism that allows a station (e.g., a host or a router) on an NBMA subnetwork to find the NBMA subnetwork address of a destination station when the destination station is connected to the NBMA subnetwork. For the case where the destination station is off the NBMA subnetwork the mechanism described in [[NHRP](#)] allows to determine the NBMA subnetwork address of an egress router from the NBMA subnetwork that is ``nearest' to the destination station. However, [[NHRP](#)] constrains the ability of determining the egress router to the destinations that are directly connected to the egress router.

This document describes extensions to the NHRP that allow a station to acquire and maintain the information about the egress router without constraining the destination(s) to be directly connected to the egress router.

Internet Draft [draft-ietf-rolc-r2r-nhrp-00.txt](https://www.ietf.org/archive/id/draft-ietf-rolc-r2r-nhrp-00.txt)

January 1996

3. Definitions

The mechanism described in this document allows to find an egress router for either a single destination, or a set of destinations (where the set is expressed as a single address prefix). Since a single destination is just a special case of a set of destinations, for the rest of the document we will always talk about a set of destinations, and will refer to this set as an ``NHRP target''.

The NHRP target is carried in the NHRP Request, Reply, and Purge messages as an address prefix using the Destination Prefix Length extension. This document requires that the NHRP target shall not be modified by the routers that forward the messages.

In general a router may maintain in its Forwarding Information Base (FIB) routes whose Network Layer Reachability Information (NLRI) exhibits a subset relation. Such routes are called overlapping routes.

A route (from a local FIB) whose NLRI forms a minimal superset of all the destinations covered by the NHRP target is called an ``NHRP forwarding route''. Observe, that by definition the set of destinations covered by an NHRP target always exhibits a subset relation to the set of destinations covered by the NHRP forwarding route associated with the target.

We will refer to the information acquired via NHRP as a ``shortcut''. We will refer to the entity that originates an NHRP Request and the entity that replies to that Request as the ``ends of the shortcut''.

To provide correct forwarding in the presence of overlapping routes this document constrains an NHRP target by prohibiting the NHRP target (carried by a Request) to form a superset of the destinations covered by any of the routes in the local FIB. The constraint applies both to the station that originates an NHRP Request and to the routers that propagate the Request. A station can originate an NHRP Request, and a router can propagate an NHRP Request only if the NHRP target of the Request does not violate the NHRP target constraint. For the rest of the document we'll refer to this constraint as the ``NHRP target constraint''.

The NHRP target constraint guarantees that within a given station forwarding to all the destinations covered by the NHRP target would

This document defines the following values for the Protocol Type field:

RIP	1
RIP-2	2
OSPF	3
Dual IS-IS	4
BGP	5

[5.](#) Processing NHRP Request

Processing of an NHRP Request by routers is covered by two sets of rules: the first set is independent of a particular routing domain, the second set is specific to a particular routing domains.

[5.1.](#) Domain-independent rules

When a router receives an NHRP Request, the router uses the NHRP target and the NHRP route information carried in the Request to check whether (a) the NHRP target constraint is satisfied, (b) the router it is in the same routing domain as the originator of the Request, and if yes, then whether (c) it is a border router for that domain.

If the NHRP target constraint is violated, the router reports an error to the originator of the Request (by sending to the originator the NHRP Error Indication message) and terminates the query. The message should indicate that the NHRP target constraint was violated. If the constraint is not violated, the router determines the NHRP forwarding route associated with the NHRP target carried by the Request. This route is used by the domain-specific rules (see [Section 5.2](#)) to determine whether the router is in the same routing domain as the originator of the Request, and whether the router is a border router for the routing domain that the originator of the Request is in.

If the router is in a different routing domain than the originator of the Request, then the router reports an error to the originator of

the Request (by sending to the originator the NHRP Error Indication message) and terminates the query.

If the router is within the same routing domain as the originator of the Request, and the router determines that it is a border router for that domain (using the domain-specific rules), then the router terminates the query and sends back an NHRP Reply. The information carried in the Reply may be either (a) IP and NBMA addresses of the router itself, or (b) IP and NBMA addresses of some other router that the router acquires via either NHRP or some other procedures (see [Section 7](#)). The former allows to establish shortcuts within a single routing domain. The latter allows to establish shortcuts that cross domain's boundary. The choice between (a) and (b) is a local to the router matter.

If the router is within the same routing domain as the originator of the Request, and the router performs routing information aggregation, then it could be possible for the NHRP forwarding route associated with the NHRP target to be a local aggregate (constructed by the

router as a result of routing information aggregation). In this case the router must terminate the query and send back an NHRP Reply with its own IP and NBMA addresses as the next hop.

[5.2](#). Domain-specific rules

The following describes NHRP handling rules specific to particular routing domains (e.g., RIP domain, OSPF domain).

[5.2.1](#). RIP, OSPF, Dual IS-IS Domains

If the routing protocol by which the NHRP forwarding route was acquired is the same as the protocol indicated by the Protocol Type field in the NHRP Route Information Extension carried by the Request, then the router handles the Request following the procedures described in [[NHRP](#)]. Otherwise, the router is a border router.

[5.2.2](#). RIP-2 Domain

If the routing protocol by which the NHRP forwarding route was acquired is the same as the protocol indicated by the Protocol Type field in the NHRP Route Information Extension carried by the Request, and the Route Tag of the route is the same as carried in the NHRP Route Information Extension, then the router handles the Request following the procedures described in [[NHRP](#)]. Otherwise, the router is a border router.

6. Maintaining correct shortcut information

Once a station that originates an NHRP Request acquires an address of an egress router along a path to a destination, it is essential for the station to be able to detect any changes that would affect the correctness of this information. The following measures are intended to provide the correctness.

Both ends of a shortcut should monitor the status of the route that was associated with the shortcut (the NHRP forwarding route). If the status changes at the router that generated the NHRP Reply (the egress router), this router should send a Purge message, so that the NHRP Requester would issue another NHRP. If the status changes at the Requester, the Requester must issue another NHRP Request. This allows to ensure that when both ends of a shortcut are up, any changes in routing that impact forwarding to any of the destination covered by

the NHRP target would result in a revalidation (via NHRP) of the shortcut.

Once a shortcut is established, the Requester needs to have some mechanism(s) to ensure that the other end of the shortcut is alive. This is needed to suppress black holes if the next hop router in the shortcut (the router that generated Reply) goes down. Among the possible mechanisms are: (a) indications from the Data Link layer, (b) presence of traffic in the reverse direction that comes with the Link Layer address of the other end, (c) information gleaned from routing protocol(s), (d) NHRP itself.

A Requester should establish a shortcut only after the Requester has a reasonable assurance that the information provided by NHRP is fairly stable. This is necessary in order to avoid initiating

shortcuts that are based on transients routing information, and thus would need to be revalidated almost immediately anyway. A router should not propagate an NHRP Request if the propagation is based on the routing information that the router views as transient. Likewise, a router should not construct an NHRP Reply based on such information.

7. Multi-domains shortcuts

While the NHRP mechanism described above is constrained to the routers within a single routing domain, the information provided by this mechanism could be sufficient to establish shortcuts that would span multiple domains.

7.1. Using the ``third-party'' next hop information

Certain routing protocols (e.g., BGP) allow a router to advertise a route with some other (than the router) entity as the next hop. This feature could be used to acquire the shortcut information that crosses domain's boundary.

Consider an example where an NHRP Request was originated within a particular routing domain A, and the NHRP target of the Request is in some other routing domain B. Further assume that the border routers in both A and B participate in a single common instance of BGP. Since BGP preserves the next hop information across an NBMA network, the routing information available at the border routers in A would contain the next hop IP information that may identify a router in some other routing domain along the path to B, perhaps even in B itself. Therefore, when a border router in A receives the Request, the router could use this information (rather than its own IP and

NBMA addresses) to construct an NHRP Reply. This way the Reply would carry the next hop information that is associated with a router in some other routing domain, thus providing to the Requester the information needed to establish a shortcut that spans multiple routing domains.

Since BGP does not carry the NBMA address information for the next hop, a router that uses the next hop information from a BGP-learned

route should use NHRP to acquire the NBMA address of the entity identified by the next hop.

7.2. Chaining/Leaking NHRP information across domain's boundary

While the ability of BGP to preserve the next hop information could reduce the number of IP hops along a path, the information, by itself, may not be sufficient to form a single IP hop across an NBMA network. However, we could observe that once a router (e.g., a border router) acquires a shortcut information, then as long as the router has sufficient assurances that the information is correct, the router could pass this information to other routers in response to NHRP Requests by using this information to construct NHRP Replies. In effect the router would ``leak'' the NHRP-learned information.

Since a border router (by definition) belongs to multiple routing domains, passing the NHRP information through the border router from one domain to another would be sufficient for establishing shortcuts that span multiple routing domains.

For example, assume that a border router X within a given domain A acquired the information needed to form a shortcut within A for a given NHRP target (the target may be either within A or outside of A). Further assume that X is also in some other routing domain B, and there is a router Y in B that would like to acquire the shortcut information for exactly the same NHRP target. If the NHRP Request originated by Y would reach X, then when X receives the Request rather than indicating itself as the next hop, X would use the shortcut information it already has to specify the next hop in the Reply. This way Y would get the information needed to construct a shortcut that crosses domain's boundary.

If X would detect any changes in the information associated with the shortcut (either due to local changes, or as a result of receiving a Purge message), then X would reissue the NHRP Request, and also would send a Purge message to Y. When Y would receive the Purge message from X, Y would reissue the NHRP Request as well.

7.3. Chaining/Leaking NHRP information with BGP

Additional complexity in handling multi-domains shortcuts arises if the routing information gets aggregated at the border routers (which certainly happens in practice). Since BGP is the major protocol that is used to exchange routing information across multiple routing domains, the following assumes that the routing information exchange across domains' boundary is controlled by BGP.

If both the source and the destination domains are on a common NBMA network, and a path between these two domains is also fully within the same NBMA network, then we have only three routing domains to deal with: the source routing domain, the BGP routing domain, and the destination routing domain. If the destination domain is not on the same NBMA as the source domain, then we need to deal only with two domains – the source and the BGP. [Note that we treat all routers that participate in a single (common) instance of BGP as a single BGP routing domain, even if these routers participate in different intra-domain routing protocols, or in different instances of the same intra-domain routing protocol.]

To simplify the presentation we decompose the problem into the following three subproblems:

- (a) how a border router in the domain that the originator of the Request is in handles the Request (crossing IGP/BGP boundary),
- (b) how the Request is handled across the BGP domain,
- (c) how a border router in the domain where the NHRP target is in handles the Request (crossing BGP/IGP boundary).

7.3.1. Handling NHRP Request at the border router in the source domain

When a border router receives an NHRP Request originated from within its own (IGP) routing domain, the border router determines the NHRP forwarding route for the NHRP target carried by the Request. If the router already has the shortcut information for the forwarding route, then the router uses this information to construct a Reply to the source of the NHRP Request. Otherwise, the router originates its own NHRP Request. The Request contains exactly the same NHRP target, as was carried by the original (received) Request; the NHRP Route Information extension contains the Protocol Type (BGP) of the NHRP forwarding route. The newly originated Request is sent to the next hop of the NHRP forwarding route. Once the border router receives a Reply to its own Request, the border router uses the next hop

information from the Reply to construct its own Reply to the source of the original NHRP Request.

If later on the border router receives a Purge message for the NHRP forwarding route, the border router treats this event as if there was a local change to the NHRP forwarding route (even if there was no changes to the route).

7.3.2. Handling NHRP Request within the BGP domain

When a BGP router (e.g., a border router at the source domain) originates an NHRP Request, this Request would be sent to a router that is either:

- (a) an egress router from an NBMA network (since in the absence of aggregation BGP preserves the next hop information), or
- (b) a border router within the domain that contains all the destinations carried by the NHRP target, or
- (c) a router that aggregates NLRI carried by the NHRP route information of the Request.

With case (a) when the router receives the Request, the router sends back an NHRP Reply and terminates the query. Case (b) is handled as described in the next section.

With case (c) when a router that receives a Request determines that it performs routing information aggregation for the NHRP target, the router could either (i) initiate another NHRP Request, and use the information received in response to this Request to construct an NHRP Reply for the original Request, or (ii) find the NHRP forwarding route associated with the NHRP target and forward the Request to the next hop of the NHRP forwarding route. The choice between options (i) and (ii) is a local to the router matter.

If the router selects option (i), then when the router receives the Request, the router determines the NHRP forwarding route for the NHRP target carried by the Request and originates its own NHRP Request. The Request contains exactly the same NHRP target, as was carried by the original request; the NHRP route information contains the Protocol Type (BGP) of the NHRP forwarding route. The newly originated Request is sent to the next hop of the NHRP forwarding route. Once the router receives a Reply to its own Request, the

router uses the next hop information from the Reply to construct its own Reply to the source of the original NHRP Request. If the router

later on receives a Purge message for the NHRP forwarding route, the router treats this event as if there was a change to the NHRP forwarding route (even if there was no changes to the route).

7.3.3. Handling NHRP Request at the destination domain border router

When a border router receives an NHRP Request from a BGP speaker, and the border router determines that all the destinations covered by the NHRP target of the Request are within the (IGP) domain of that border router, the border router determines the NHRP forwarding route for the NHRP target carried by the Request. The newly formed Request contains exactly the same NHRP target as the received Request; the NHRP route information contains the Protocol Type of the NHRP forwarding route. The newly originated Request is sent to the next hop of the NHRP forwarding route. Once the border router receives a Reply to its own Request, the border router uses the next hop information from the Reply to construct its own Reply to the source of the original NHRP Request.

If the border router later on receives a Purge message for the NHRP forwarding route, the border router treats this event as if there was a change to the NHRP forwarding route (even if there was no changes to the route).

7.4. Tradeoffs between state, messages, and path optimality

It should be possible to reduce the number of Purge messages and subsequent NHRP messages (caused by the Purge messages) by maintaining more state on the border routers at the source and destination domains, and the BGP routers that perform routing information aggregation along the path from the source to the destination.

Specifically, on each such router it would be necessary to keep the information about all the NHRP targets for which the router maintains the shortcut information. This way when the router determines that

the NHRP forwarding route (for which the router maintains the shortcut information) changes due to some local routing changes, the router could check whether these local changes impact forwarding to the destinations covered by the NHRP targets. The router would send Purge messages only for the targets that are impacted by the changes.

Upon some introspection one could realize that the shortcut information across a BGP domain could be used for as long as the NHRP forwarding route at both ends of the shortcut stays the same (even in

the presence of aggregation along the shortcut). Such information would provide loop-free forwarding, but may result in a potentially suboptimal path (if a router that performs aggregation along the path selects another (better) route for forming the aggregate). This way there is no need to maintain an additional state on the BGP routers that perform routing information aggregation, and there will be no additional NHRP traffic when these routers change the way they construct their aggregates, provided that the aggregated routes would stay the same.

[8.](#) Security Considerations

Security issues are not discussed in this document.

[9.](#) References

[NHRP] Katz, D., Piscitello, D., Cole, B., Luciani, J., ``NBMA Next Hop Resolution Protocol (NHRP)', Internet Draft, December 1995

[10.](#) Acknowledgements

To be supplied.

[11.](#) Author Information

Yakov Rekhter
cisco Systems, Inc.

170 Tasman Dr.
San Jose, CA 95134
Phone: (914) 528-0090
email: yakov@cisco.com