Next Hop Resolution Protocol (NHRP)

Yakov Rekhter

Cisco Systems

yakov@cisco.com

Tradeoffs

- Path optimality (number of IP hops across NBMA)
- Volume of control traffic
- Amount of state maintained
  - distribution of state among routers
- Adaptability

Extra Hop Elimination - destination is off the NBMA

- Ingress router to egress router
- Host to egress router
- NHRP address mapping functionally is insufficient

How to find an egress router for a destination ?

- Requestor (e.g., host, router) sends an NHRP Request destined to that destination (or a group of destinations)
- The Request is forwarded along the best path to the destination (the same path as used for data forwarding)
- When the Request reaches the egress router (at the border of the NBMA subnetwork), the router sends back an NHRP Reply
    - Reply includes router's IP and NBMA addresses
- When the requestor receives the Reply, the requestor has all the necessary information for establishing the shortcut

How to maintain correct next hop information ?

- The next hop information reflects the state of the routing system **at the time the NHRP Request was issued**
- As the state of the routing system changes, the information acquired via the Request/ Reply may no longer be correct, but rather:
  - may result in a path to a “black hole”
  - may result in a path with a forwarding loop

How to maintain correct next hop information (cont.) ?

- Periodic verification:
  - Requestor periodically reissues NHRP Request
  - Adaptability to routing changes is proportional to the frequency of Requests
  - Tradeoffs between the adaptability and the overhead associated with processing NHRP messages

How to maintain correct next hop information (cont.) ?

- Triggered verification:
  - Requestor reissues NHRP Request based on certain events
  - Generating events requires to maintain additional state in the network
  - Tradeoffs between adaptability and the amount of the extra state required

How to avoid “black holes” ?

- How to detect when an egress router goes down ?
  - Very similar to a well-known problem of detecting a dead next-hop router
  - Common problem, not specific to NBMA and/or NHRP

How to avoid “black holes” (cont.) ?

- Use the Link Layer information (e.g., SVC goes down)
  - Very little overhead, but may not always be available
- Use the information carried by the traffic in the reverse direction
  - Very little overhead, but may not always be available
- Use the information from the routing system
  - Very little overhead, but may not always be available

How to avoid “black holes” (cont.) ?

- Use periodic NHRP Request/Reply
  - Overhead of processing NHRP Request/Reply
- Introduce NHRP KeepAlive messages:
  - Exchanged **directly** between the requestor and the egress router
  - Less overhead than NHRP Request/Reply
  - Would allow to send KeepAlives frequently enough to provide good adaptability

How to avoid forwarding loops ?

- Does not apply to the host to egress router case
  - By definition a host does not forward
- Applies only to the ingress router to egress router case

How to avoid forwarding loops (cont.) ?

- Intra-domain shortcuts
  - Allows to avoid extra hops only within a routing domain
- Inter-domain shortcuts
  - Allows to avoid extra hops across several routing domains

Intra-domain shortcuts

- NHRP Request carries sufficient information to detect routing domain boundaries
- Maintaining state at the requestor and the egress router of the requestor's domain is sufficient to detect changes in the routing system within a domain for protocols like OSPF, RIP

Inter-domain shortcuts

- The technique of the intra-domain shortcuts allows to maintain loop-free shortcut information within the requestor's domain
- In the absence of aggregation BGP preserves correct (loop-free) next hop information
  - Provides shortcut information within the BGP domain

Inter-domain shortcuts (cont.)

- By “leaking” the BGP next hop information at the requestor’s domain boundary, we could have the (loop-free) information needed to establish a shortcut that spans multiple routing domains

Inter-domain shortcuts (cont.)

- Observe that a shortcut could traverse at most 3 routing domains:
  - The requestor's routing domain
    - Apply the technique of intra-domain shortcuts
  - BGP routing domain
  - The egress router's routing domain
    - Apply the technique of intra-domain shortcuts

Inter-domain shortcuts (cont.)

- "Leaking" of the shortcut information across domain boundaries would provide the necessary information:
  - From the egress router's domain into the BGP domain
  - From the BGP domain to the requestor's domain

BGP Routing Domain

- BGP already provides the shortcut information that is guaranteed to be loop free, *but*
- We need to figure out how to deal with the loss of this information due to aggregation (at BGP level)

BGP Routing Domain (cont.)

- How to deal with routing information aggregation within a BGP routing domain ?
  - Use a sequence of concatenated NHRP Requests/Replies segments
  - Each segment within the sequence is terminated at the router that performs aggregation
  - Introduces extra state on the routers that perform aggregation
- Alternative is not to maintain state on BGP routers that perform the aggregation