DYNAMIC MULTIPOINT VPN HUB AND SPOKE INTRODUCTION

NOVEMBER 2004
INTRODUCTION
What is Dynamic Multipoint VPN?

- Dynamic Multipoint VPN (DMVPN) is a combination of GRE, NHRP, and IPsec
- NHRP allows the peers to have dynamic addresses (ie: Dial and DSL) with GRE / IPsec tunnels
- Backbone is a hub and spoke topology
- Allows direct spoke to spoke tunneling by auto leveling to a partial mesh
Site-to-Site, DMVPN: mGRE/IPsec/NHRP Integration, Only HUB address Is Known

LANs can have private addressing

Static known IP address

Dynamic unknown IP addresses

SPOKE

HUB

Static known IP address = Static spoke-to-hub IPsec tunnels

= Dynamic&Temporary Spoke-to-spoke IPsec tunnels

10.0.0.0 255.255.255.0

10.0.3.0 255.255.255.0

10.0.2.0 255.255.255.0

10.0.1.0 255.255.255.0

10.0.1.1

10.0.0.1

10.0.3.1

10.0.2.1
Terminology Pause

- DMVPN is a partial dynamic mesh
  - **Spoke**: all the devices that contact a central router called “hub”
  - **Node**: any hub or a spoke
This Presentation

• This presentation concentrate on hub and spoke to explain how DMVPN works
DMVPN
GRE Tunnels

- A GRE tunnel is a simple non-negotiated tunnel; GRE only needs tunnel endpoints
- GRE encapsulate frames or packets into another IP packet + IP header
- GRE has only 4 to 8 bytes of overhead
- GRE tunnels exist in two main flavors:
  - Point-to-point (GRE)
  - Point-to-multipoint (mGRE)
GRE multipoint and DMVPN

• A GRE interface definition includes
  - An IP address
  - A tunnel source
  - A tunnel destination
  - An optional tunnel key

• An mGRE interface definition includes
  - An IP address
  - A tunnel source
  - A tunnel key

• mGRE interfaces do not have a tunnel destination

```
interface Tunnel 0
  ip address 10.0.0.1 255.0.0.0
tunnel source Dialer1
tunnel destination 172.16.0.2
tunnel key 1
```
Terminology Pause

• The **tunnel address** is the ip address defined on the tunnel interface

• The Non-Broadcast Multiple Access (NBMA) address is the ip address used as tunnel source (or destination)

• Example… on router A, one configures

```plaintext
interface Ethernet0/0
  ip address 172.16.0.1 255.255.255.0
interface Tunnel0
  ip address 10.0.0.1 255.0.0.0
  tunnel source Ethernet0/0
[...]
10.0.0.1 is router A’s **tunnel address**
172.16.0.1 is router A’s **NBMA address**
```
mGRE Tunnels

- Single tunnel interface (mp)
  - Non-Broadcast Multi-Access (NBMA) Network
  - Multiple (dynamic) tunnel destinations
  - Multicast/broadcast support

- Next Hop Resolution Protocol (NHRP)
  - VPN IP to NBMA IP address mapping
GRE Encapsulation
DMVPN GRE Interfaces

• In DMVPN, the hub **must** have a point to mGRE

• Spokes **can** have a point to point GRE interface or an mGRE interface

• This presentation will use mGRE everywhere for consistency

• Note that point-to-point GRE interfaces prevent spoke to spoke direct tunneling
mGRE Talking to a Peer

- Because mGRE tunnels do **not** have a tunnel destination defined, they cannot be used alone.
- **NHRP** tells mGRE where to send the packets to.
- NHRP is defined in RFC 2332.
What is NHRP?

- NHRP is a layer two resolution protocol and cache like ARP or Reverse ARP (Frame Relay)
- It is used in DMVPN to map a tunnel IP address to an NBMA address
- Like ARP, NHRP can have static and dynamic entries
- NHRP has worked fully dynamically since Release 12.2(13)T
How mGRE Uses NHRP

- When a packet is routed, it is passed to the mGRE interface along with a next-hop
- The next-hop is the tunnel address of a remote peer
- mGRE looks up the **NHRP cache** for the next-hop address and retrieves the NBMA address of the remote peer
- mGRE encapsulates the packet into a GRE/IP payload
- The new packet destination is the NBMA address
- Multicast packets are only sent to **specific remote peers** identified in the NHRP configuration
mGRE/NHRP Path

Routing Table
192.168.1.0/24 -> Tunnel 0, via 10.0.0.2

NHRP Table
10.0.0.2 -> 172.16.0.2

Tunnel address:
10.0.0.1/24

NBMA address:
172.16.0.1/24

NBMA address:
172.16.0.2/24

Tunnel address:
10.0.0.2/24

IP
s=172.16.0.1, d=172.16.0.2

GRE
IP
s=192.168.0.1, dst=192.168.1.1

Payload
IP
s=192.168.0.1, dst=192.168.1.1

Payload
How NHRP Works

• mGRE uses NHRP, but how does NHRP work?
• This presentation will introduce a network topology and illustrate the associated NHRP commands
NHRP Registration
Dynamically Addressed Spokes

DMVPN Hub and Spoke, 11/04
© 2004 Cisco Systems, Inc. All rights reserved.
Basic NHRP Configuration

- In order to configure an mGRE interface to use NHRP, the following command is necessary:
  \[ \text{ip nhrp network-id <id>} \]
- Where \(<id>\) is a unique number (same on hub and all spokes)
- \(<id>\) has nothing to do with tunnel key
- The network ID defines an NHRP domain
  Several domains can co-exist on the same router
Populating the NHRP Cache

- Three ways to populate the NHRP cache:
  - Manually add static entries
  - Hub learns via registration requests
  - Spokes learn via resolution requests
- We will now study “static” and “registration”
- “Resolution” is for spoke to spoke
Initial NHRP Caches

- Initially, the hub has an empty cache
- The spoke has one static entry mapping the hub’s tunnel address to the hub’s NBMA address:
  ```
  ip nhrp map 10.0.0.1 172.17.0.1
  ```
- Multicast traffic must be sent to the hub
  ```
  ip nhrp map multicast 172.17.0.1
  ```
The Spokes Must Register To The Hub

- In order for the spokes to register themselves to the hub, the hub must be declared as a Next Hop Server (NHS):

  `ip nhrp nhs 10.0.0.1`

  `ip nhrp holdtime 3600 (optional)`

  `ip nhrp registration no-unique (optional)`

- **Spokes control the cache on the hub**
Registration Process

- The spokes send **Registration-requests** to the hub
- The request contains the spoke’s **Tunnel** and **NBMA** addresses as well as the hold time and some flags
- The hub creates an entry in its NHRP cache
- The entry will be valid for the duration of the **hold time defined in the registration**
- The NHS returns a **registration reply** (acknowledgement)
Multicast Packets from the Hub

- The hub must also send multicast traffic to all the spokes that registered to it
- This **must** be done dynamically (possible since Release 12.2(13)T)
- This is **not** the default
  
  `ip nhrp map multicast dynamic`
NHRP Registration Request

**NHRP Table**
- 192.168.1.1/24
  - Physical: 172.17.0.1
  - Tunnel0: 10.0.0.1
- 10.0.0.11 -> 172.16.1.1
  - (dynamic, mcast, hold=3600, no-unique)

**NHRP Table**
- 10.0.0.1 -> 172.17.0.1
  - (static, mcast)

**IP**
- s=172.16.1.1, d=172.17.0.1

**GRE**
- NHRP-Registration
  - Tunnel=10.0.0.11
  - s=10.0.0.11
  - d=172.16.1.1
  - dst=10.0.0.1
  - Hold=3600, no-unique
NHRP Registration Reply

27.168.0.1/24

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

NHRP Table
10.0.0.11 -> 172.16.1.1
(dynamic, mcast, hold=3600, no-unique)

IP
s=172.16.1.1,
d=172.17.0.1

GRE
NHRP-Registration Reply
s=10.0.0.11
dst=10.0.0.1
Code = Successful

Physical: (dynamic)172.16.1.1
Tunnel0: 10.0.0.11

Spoke A

192.168.1.1/24

NHRP Table
10.0.0.1 -> 172.17.0.1
(static, mcast)
NHRP Functionality

- Address mapping/resolution
  - Static NHRP mapping
  - Next Hop Client (NHC) registration with Next Hop Server (NHS)

- Packet Forwarding
  - Resolution of VPN to NBMA mapping
    - Routing: IP destination ➔ Tunnel IP next-hop
    - NHRP: Tunnel IP next-hop ➔ NBMA address
Routing Protocol

- The spoke needs to advertise its private network to the hub
- Can use BGP, EIGRP, OSPF, RIP or ODR; however, this presentation will focus on EIGRP
- Must consider several caveats
Spoke Hellos

- Spoke has all it needs to send hellos immediately:
  - Tunnel is defined
  - Static NHRP entry to hub is present
  - NHRP entry is marked for multicast
- So the spoke never waits…
Hub hello’s

- With its basic tunnel definition, the hub **cannot send anything** (including hellos) to anyone
- It must wait NHRP for registrations to arrive
- As soon as the spokes have registered, the NHRP is marked “Multicast” due to
  ```
  ip nhrp map multicast dynamic
  ```
- The hub sends hellos to **all the registered spokes simultaneously**
Hub sending EIGRP hello

192.168.0.1/24

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

NHRP Table
10.0.0.11 -> 172.16.1.1
(dynamic, mcast, hold=3600, no-unique)

IP
s=172.17.0.1,
d=172.16.1.1

GRE

IP
s=10.0.0.1,
d=224.0.0.10

EIGRP hello

Spoke A

Physical: (dynamic)172.16.1.1
Tunnel0: 10.0.0.11

192.168.1.1/24

NHRP Table
10.0.0.1

EIGRP neighbor
10.0.0.1

NHRP Table
10.0.0.1 -> 172.17.0.1
(static, mcast)
GRE and EIGRP

- The **default** bandwidth of a GRE tunnel is 9Kbps
- This has no influence on the traffic but...
- **EIGRP** will take \( \frac{1}{2} \) the interface bandwidth maximum (4.5 Kbps) – this is too low

`bandwidth 1000`
Spoke EIGRP configuration

- Nothing special on the spoke
- EIGRP stub should be considered
Hub EIGRP Configuration

- There are many options...
- If you want a spoke to see other spokes:
  ```
  no ip split-horizon eigrp 1
  ```
- Summarization is to be considered
- Setting the **bandwidth** is crucial in the hub to spoke direction
- Best-practice: Set the bandwidth the same on all nodes
IPsec Protection

- GRE/NHRP can build a fully functional overlay network
- GRE is insecure; ideally, it must be protected
- The good old crypto map configuration is rather cumbersome; DMVPN introduced tunnel protection
- Still need to define an IPsec security level
The IPsec Security Policy

- A transform set must be defined:
  ```
  crypto ipsec transform-set ts esp-sha-hmac esp-3des
  mode transport
  ```
- An IPsec profile replaces the crypto map
  ```
  crypto ipsec profile prof
  set transform-set ts
  ```
- The IPsec profile is like a crypto map without “set peer” and “match address”
Protecting the tunnel

- The profile must be applied on the tunnel
  `tunnel protection ipsec profile prof`

- Internally Cisco IOS® Software will treat this as a dynamic crypto map and it derives the `local-address`, `set peer` and `match address` parameters from the `tunnel parameters` and the `NHRP cache`

- This must be configured on the hub and spoke tunnels
Relation Between GRE, NHRP and IPsec

• For each NHRP cache unique NBMA address, Cisco IOS Software will create an internal crypto map that protects
  GRE traffic
  From tunnel source (NBMA) address
  To NHRP entry NBMA address

• The SAs will be negotiated as soon as the cache entry is created (static and resolved)
Relationship (cont’d.)

- NHRP registration will be triggered
  - When the Tunnel interface comes up/up
  - When the tunnel source address changes
  - When IPsec finishes negotiating the phase 2 related to the tunnel protection
  - When the registration timer expires
NHRP Registration
Dynamically Addressed Spokes

192.168.1.0/24
192.168.2.0/24
192.168.0.0/24

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1
192.168.1.0/24 → Conn.
192.168.0.0/24 → Conn.

172.16.1.1
172.16.2.1

192.168.1.1
192.168.2.1

10.0.0.1 → 172.16.0.1

NHRP mapping
Routing Table

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

Spoke A
192.168.1.1/24

Spoke B
192.168.2.1/24

10.0.0.11
10.0.0.12

192.168.1.0/24
192.168.2.0/24
192.168.0.0/24
Building Hub-and-Spoke tunnels
NHRP Registration

IKE Initialization
IKE/IPsec Established
NHRP Regist. Req.
Routing Adjacency
Routing Update
Routing Update

Encrypted

IKE Initialization
IKE/IPsec Established
NHRP Regist. Req.
Routing Adjacency
Routing Update
Routing Update

Encrypted