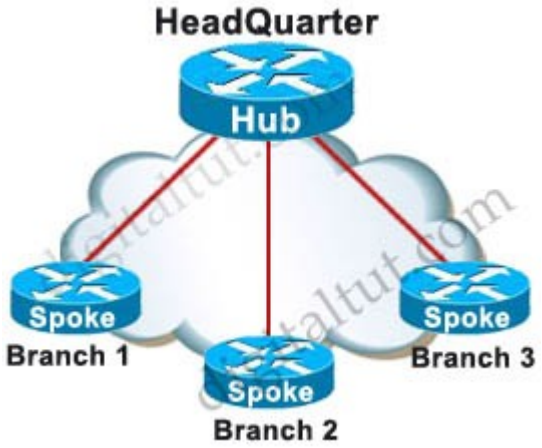
DMVPN Tutorial

February 14th, 2015 in ROUTE Knowledge Go to comments

One of the most popular network topology in practical nowadays is shown below with one HeadQuarter connecting to branch offices at some locations. The main enterprise resources are located in the HeadQuarter.

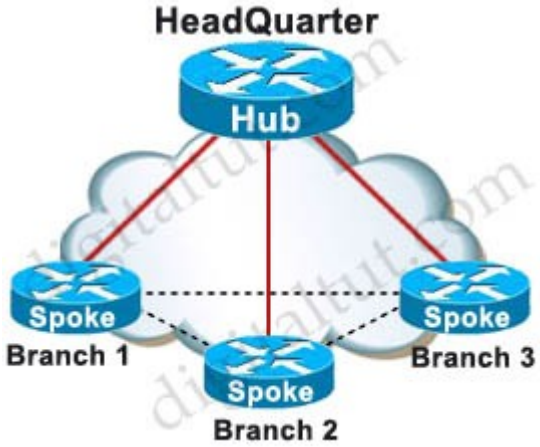


The router at the HeadQuarter undertakes the role of a **Hub** while branch routers take the role of **Spokes**. In this **Hub-and-Spoke** topology, each Branch can access some resources on the HeadQuarter. But there are some disadvantages with this topology:

+ When a spoke wants to communicate with another Spoke, it must go through the Hub which increases the traffic passing through the Hub, increase CPU and memory usage on Hub and can create bottle-neck problem. This also increases latency for time-sensitive applications such as VoIP, video conference…
+ Each site requires a static public IP address if the environment between them are public (like the Internet).
+ The configuration is complex, especially with large network. When a new Spoke is added, additional configuration is required on Hub

Dynamic Multipoint VPN (DMVPN) is a solution of Cisco that can be used to overcome these disadvantages. DMVPN provides the following advantages:

+ Provides full meshed connectivity with simple Hub-and-Spoke topology. The spokes can communicate between each other without going through Hub
+ Only one static public IP address is required on Hub. Spokes can use dynamic (unknown) public IP addresses
+ The configuration is simple even in large network. No additional configuration is required on Hub when new Spokes are added.



DMVPN provides full-meshed connectivity
with Hub-and-Spoke topology

But notice that DMVPN is not a protocol, it is the combination of the following technologies:

+ Multipoint GRE (mGRE)
+ Next-Hop Resolution Protocol (NHRP)
+ Dynamic Routing Protocol (EIGRP, RIP, OSPF, BGP…) (optional)
+ Dynamic IPsec encryption (optional)
+ Cisco Express Forwarding (CEF)

DMVPN combines multiple GRE (mGRE) Tunnels, IPSec encryption and NHRP (Next Hop Resolution Protocol) to perform its job and save the administrator the need to define multiple static crypto maps and dynamic discovery of tunnel endpoints.

To keep this tutorial simple we only mention about mGRE and NHRP.

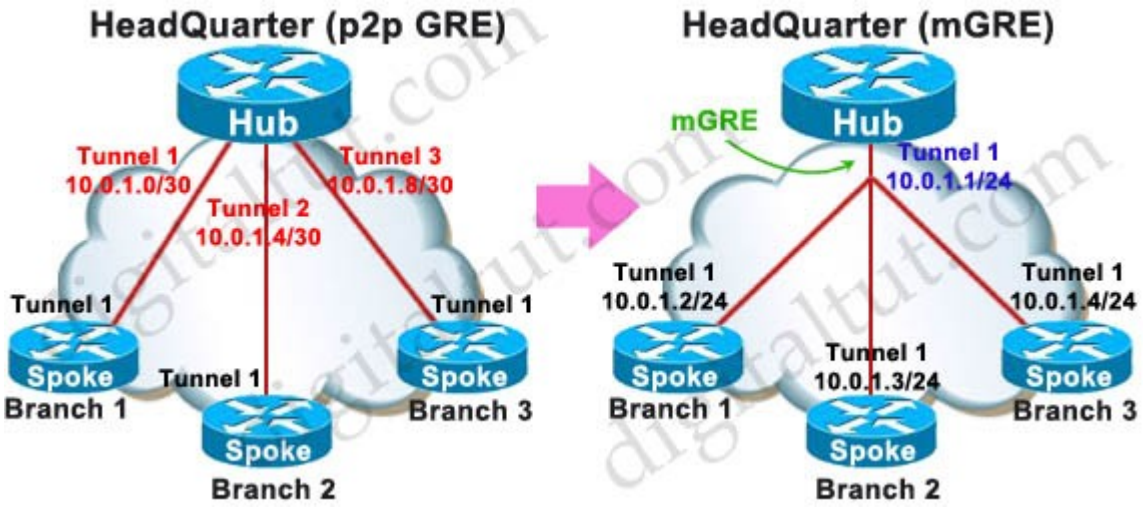
Multipoint Generic Routing Encapsulation (mGRE)

Before taking about mGRE we should learn why we have to run GRE on DMVPN. The answer is simple: because we want to run IPSec on it. And why we need IPSec? Because we want to utilize the power of cheap but insecure Internet (and other insecure public) connections at our sites.

As you may know, IPSec is a framework consisting of protocols and algorithms for protecting data through an untrusted IP network, such as the internet. Although IPSec provides a secure tunneling method but it does not support multicast and broadcast traffic so popular routing protocol (OSPF, EIGRP, …) run based on multicast cannot be used with IPSec. So we have to use GRE to “wrap” these multicast traffic. As a result, all traffic (including unicast, multicast and broadcast) between sites are encapsulated into GRE packets before being encrypted and sent over the network.

Now we knew why GRE should be used here. But traditional GRE (sometimes called point-to-point or p2p GRE) also has its limitation: for each connection to the Spoke, Hub router needs to establish a separate GRE tunnel. So when the number of Spokes increases, Hub must increase the number of tunnels at the same rate -> lots of configuration on Hub. So it is the time when mGRE takes part in.

An mGRE tunnel inherits the concept of a classic GRE tunnel but an mGRE tunnel does not require a unique tunnel interface for each connection between Hub and spoke like traditional GRE. One mGRE can handle multiple GRE tunnels at the other ends. Unlike classic GRE tunnels, the tunnel destination for a mGRE tunnel does not have to be configured; and all tunnels on Spokes connecting to mGRE interface of the Hub can use the same subnet.



mGRE tunnel is treated as a non-broadcast multi-access (NBMA) environment. mGRE tunnel does not have to be configured with a tunnel destination so we need another protocol to take care of the destination addresses. In this case NHRP is used for NBMA environment.

Note: Besides the Tunnel IP address, each Spoke and Hub will have a NBMA IP address, which is a public IP address used as the tunnel source IP address. We post the configuration here as an example to help you understand more about the difference of these two IP addresses:

Hub	**Spoke (Branch 3)**
interface fa0/0 ip address 11.11.11.1 255.255.255.0 interface tunnel 1 ip address 192.168.100.1 255.255.255.0 -> Tunnel IP address (private IP) tunnel source fa0/0 -> NBMA IP address (public IP)	interface fa0/0 ip address 13.13.13.3 255.255.255.0 interface tunnel 1 ip address 192.168.100.3 255.255.255.0 -> Tunnel IP address (private IP) tunnel source fa0/0 -> NBMA IP address (public IP)

So the Tunnel address is the address configured under “interface tunnel” while the NBMA address is the address used as source of the tunnel.

NHRP

Next Hop Resolution Protocol (NHRP), defined in RFC 2332, is a Layer 2 address resolution protocol and cache, like Address Resolution Protocol (ARP). NHRP is used by a branch router connected to a non-broadcast, multi-access (NBMA) sub-network to determine the IP address of the “NBMA next hop”; in this case, the headend router or the destination IP address of another branch router.

NHRP is used to map tunnel IP addresses to “physical” or “real” IP addresses, used by endpoint routers. It resolves private addresses (those behind mGRE and optionally IPSEC) to a public address. NHRP is layer 2 resolution protocol and cache, much like Address Resolution Protocol (ARP) or Reverse ARP (Frame Relay).

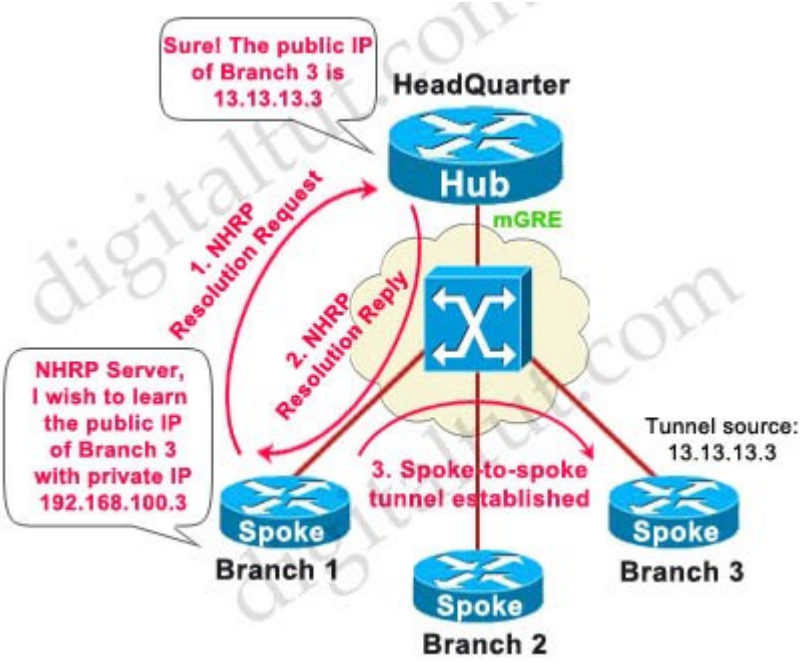
In order for DMVPN to work correctly, DMVPN relies on NHRP to create a mapping database of all spoke tunnels to real (public) IP addresses. When a Spoke joins a DMVPN network it will register itself with the Hub via NHRP. The **NHRP Registration Process** is described below:

+ When a Spoke joins a DMVPN network, it sends a Registration Request to the Hub whose IP address has already been configured on the Spoke (via the “ip nhrp nhs <Hub IP address>” command)
+ The Registration Request contains the Spoke’s Tunnel and NBMA addresses along with the hold time -> Hub does not have to statically configure Spoke IP -> simplify Hub configuration
+ Hub then create an NHRP mapping entry in its NHRP cache (just like an ARP cache) to keep the mapping between Spoke’s Tunnel and NBMA addresses. The hold time of this mapping equals to the hold time in the Registration Request.
+ Hub sends a NHRP Registration Reply to the Spoke to complete the process

NHRP Registration Process

Note:
+ The Spoke who sends NHRP Registration Request is called NHRP Client (NHC) while the Hub who replies the request is called NHRP Server (NHS).
+ The Spoke’s NBMA address is often its public IP and obtained dynamically while the Spoke’s Tunnel address is the private IP
+ NHRP mapping can be statically configured on both Spoke and Hub

A cool advantage of NHRP is the ability to help DMVPN establish direct Spoke-to-Spoke communication without going through Hub. Let’s see how NHRP works in this case.



NHRP Resolution Process

1. Before a spoke can directly send traffic to another spoke, it must still query the Hub to get the NBMA address of the destination spoke. To do this, Spoke must send a NHRP Resolution Request to the Hub asking for the NBMA address of the destination spoke.
2. The Hub replies with the NBMA (public) address of Spoke 3 (which is 13.13.13.3 in this case). If the Hub does not known NBMA of Spoke 3 it will query Spoke 3 first.
3. The direct IPsec tunnel between two spokes is built only after that. But the spoke-to-spoke tunnel is only temporary and is torn down after a pre-configured period of inactivity to save resources.

Note:
+ In case NHS does not have an entry in its cache for the NHC's query, NHS returns an error and the spoke will install an entry pointing to the NHS. So traffic must flow through the Hub
+ Instead of asking NHS, the destination spoke IP can be statically configured on the NHC.
+ "Resolution" is only used for spoke to spoke communication

Now let's see the whole picture of how NHRP takes part in the routing process.

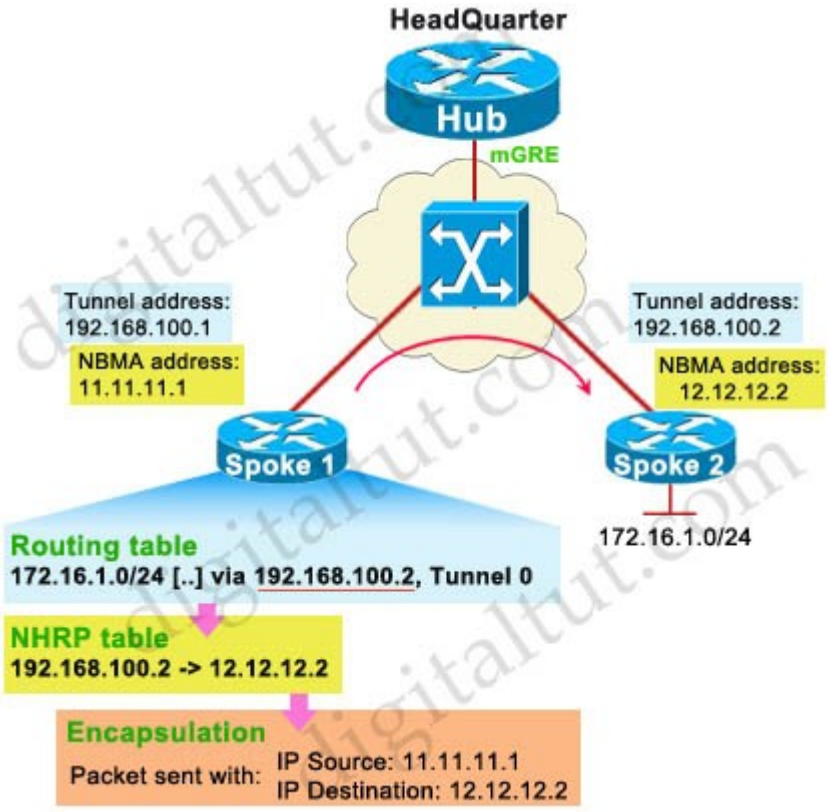
1. Suppose Spoke 1 wants to send traffic to network 172.16.1.0/24 behind Spoke 2. It will look up its routing table and see an entry like this:

172.16.1.0/24 … via 192.168.100.2, Tunnel0

(means this subnet was learned from next-hop 192.168.100.2 via its Tunnel0)

2. Spoke 1 looks up its NHRP mapping table to search for the NBMA address of 192.168.100.2. If it can't find one, it will send an NHRP Resolution Request to get the mapping information from the Hub. Suppose the NBMA address of 192.168.100.2 configured on Spoke 2 is 12.12.12.2.
3. Now Spoke 1 has enough information to encapsulate original packets. It will encapsulate

packets with IP source of 11.11.11.1 (its NBMA address) and IP destination of 12.12.12.2 (Spoke 2's NBMA address) then send to the destination.

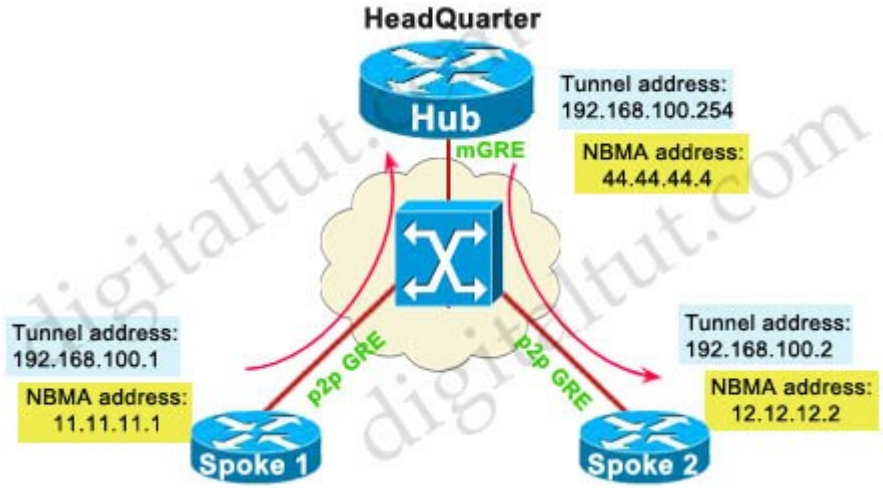


In the next part we will learn how to configure DMVPN

Configuring DMVPN

DMVPN can be configured in three different methods, each method is often called a "phase":

1) DMVPN Phase I (Spoke-to-Hub only):
+ mGRE is configured on Hub, p2p GRE is configured on Spokes
+ Traffic flows between Spoke & Hub only (Spokes talk to each other through hub). No spoke-to-spoke direct communication



DMVPN Phase I

DMVPN Phase I – Static Mapping **Hub** interface tunnel 1 ip address 192.168.100.254 255.255.255.0 tunnel source 44.44.44.4 tunnel mode gre multipoint ip nhrp network 10 ip nhrp map 192.168.100.1 11.11.11.1 ip nhrp map 192.168.100.2 12.12.12.2 **Spoke 1** interface tunnel 1 ip address 192.168.100.1 255.255.255.0 tunnel source 11.11.11.1 tunnel destination 44.44.44.4 ip nhrp network 10 ip nhrp map 192.168.100.254 44.44.44.4 ! ! ! **Spoke 2** interface tunnel 1 ip address 192.168.100.2 255.255.255.0 tunnel source 12.12.12.2 tunnel destination 44.44.44.4 ip nhrp network 10 ip nhrp map 192.168.100.254 44.44.44.4	**DMVPN Phase I – Dynamic Mapping** **Hub** interface tunnel 1 ip address 192.168.100.254 255.255.255.0 tunnel source 44.44.44.4 tunnel mode gre multipoint ip nhrp network 10 (Notice there are no “ip nhrp map …” commands in Hub, since mapping will be dynamic) **Spoke 1** interface tunnel 1 ip address 192.168.100.1 255.255.255.0 tunnel source 11.11.11.1 tunnel destination 44.44.44.4 ip nhrp network 10 ip nhrp map 192.168.100.254 44.44.44.4 ip nhrp nhs 192.168.100.254 (“ip nhrp nhs …” command send registration request to hub, tells our spoke router who the Next Hop Server is) **Spoke 2** interface tunnel 1 ip address 192.168.100.2 255.255.255.0 tunnel source 12.12.12.2 tunnel destination 44.44.44.4 ip nhrp network 10 ip nhrp map 192.168.100.254 44.44.44.4 ip nhrp nhs 192.168.100.254

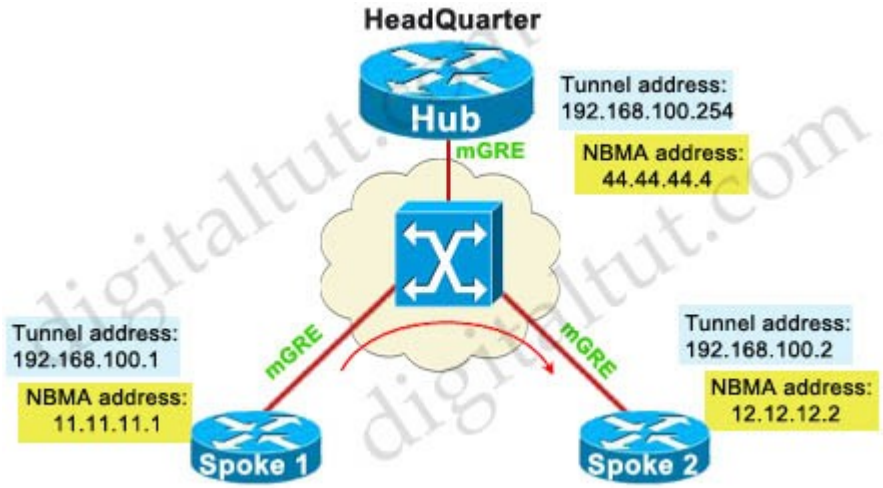
First we notice in the Hub configuration there is no "tunnel destination" command because the tunnel destination is derived from NHRP database. If we are running dynamic routing protocols based on multicast (like RIP, OSPF, EIGRP …) we have to add the command "**ip nhrp map multicast dynamic**" in Hub to replicate all multicast traffic to all dynamic entries in the NHRP table (multicast will be proceeded as unicast traffic).

"ip nhrp network 10" uniquely identifies the DMVPN network; tunnels will not form between routers with different NHRP network IDs.

2) DMVPN Phase II (Spoke-to-Spoke):

In this phase every hub and spoke is configured with mGRE interface so we can create dynamic spoke-to-spoke connectivity, no more static tunnel destination's will be configured.

+ Hub use mGRE tunnel
+ Spokes use mGRE tunnel
+ Spokes talk to each other directly



DMVPN Phase II

DMVPN Phase II Configuration

DMVPN Phase II – Static Mapping	**DMVPN Phase II – Dynamic Mapping**
Hub	**Hub**
interface tunnel 1	interface tunnel 1
ip address 192.168.100.254 255.255.255.0	ip address 192.168.100.254 255.255.255.0
tunnel source 44.44.44.4	tunnel source 44.44.44.4
tunnel mode gre multipoint	tunnel mode gre multipoint
ip nhrp network 10	ip nhrp network 10
ip nhrp map 192.168.100.1 11.11.11.1	!
ip nhrp map 192.168.100.2 12.12.12.2	!
Spoke 1	**Spoke 1**
interface tunnel 1	interface tunnel 1
ip address 192.168.100.1 255.255.255.0	ip address 192.168.100.1 255.255.255.0
tunnel source 11.11.11.1	tunnel source 11.11.11.1

tunnel mode gre multipoint ip nhrp network 10 ip nhrp map 192.168.100.1 11.11.11.1 ip nhrp map 192.168.100.2 12.12.12.2 ip nhrp map 192.168.100.254 44.44.44.4 **Spoke 2** interface tunnel 1 ip address 192.168.100.2 255.255.255.0 tunnel source 12.12.12.2 tunnel mode gre multipoint ip nhrp network 10 ip nhrp map 192.168.100.1 11.11.11.1 ip nhrp map 192.168.100.2 12.12.12.2 ip nhrp map 192.168.100.254 44.44.44.4	tunnel mode gre multipoint ip nhrp network 10 ip nhrp map 192.168.100.254 44.44.44.4 ip nhrp nhs 192.168.100.254 ! **Spoke 2** interface tunnel 1 ip address 192.168.100.2 255.255.255.0 tunnel source 12.12.12.2 tunnel mode gre multipoint ip nhrp network 10 ip nhrp map 192.168.100.254 44.44.44.4 ip nhrp nhs 192.168.100.254

Note: Although Phase II – Dynamic Mapping is "dynamic" but we still need to add a static entry for the hub because without that entry, the NHRP registration cannot be sent.

To verify the DMVPN configuration we can use the "show dmvpn" or "show ip nhrp" command. The outputs of these commands are shown below:

On Hub:

```
Hub#show dmvpn
Legend: Attrb --> S - Static, D - Dynamic, I - Incomplete
        N - NATed, L - Local, X - No Socket
        # Ent --> Number of NHRP entries with same NBMA peer
        NHS Status: E --> Expecting Replies, R --> Responding, W -->
Waiting
        UpDn Time --> Up or Down Time for a Tunnel
==========================================================================

Interface: Tunnel1, IPv4 NHRP Details
Type:Hub, NHRP Peers:2,

 # Ent  Peer NBMA Addr Peer Tunnel Add State  UpDn Tm Attrb
 ----- --------------- --------------- ----- -------- -----
     1 11.11.11.1          192.168.100.1  UP 00:03:08     D
     1 12.12.12.2          192.168.100.2  UP 00:03:16     D
Hub#show ip nhrp
192.168.100.1/32 via 192.168.100.1
   Tunnel1 created 00:28:51, expire 01:48:59
   Type: dynamic, Flags: unique registered used nhop
   NBMA address: 11.11.11.1
192.168.100.2/32 via 192.168.100.2
   Tunnel1 created 00:26:47, expire 01:48:57
   Type: dynamic, Flags: unique registered used nhop
   NBMA address: 12.12.12.2
```

On Spoke:

```
Spoke1#show dmvpn
Legend: Attrb --> S - Static, D - Dynamic, I - Incomplete
        N - NATed, L - Local, X - No Socket
        # Ent --> Number of NHRP entries with same NBMA peer
```

```
        NHS Status: E --> Expecting Replies, R --> Responding, W -->
Waiting
        UpDn Time --> Up or Down Time for a Tunnel
==========================================================================

Interface: Tunnel1, IPv4 NHRP Details
Type:Spoke, NHRP Peers:2,

 # Ent  Peer NBMA Addr Peer Tunnel Add    State  UpDn Tm Attrb
 ----- --------------- ---------------    ----- -------- -----
     1 44.44.44.4          192.168.100.254   UP 00:03:40     S
     1 12.12.12.2          192.168.100.2     UP 00:03:20     D
Spoke1#show ip nhrp
192.168.100.254/32 via 192.168.100.254
   Tunnel1 created 00:11:35, never expire
   Type: static, Flags: used
   NBMA address: 44.44.44.4
192.168.100.2/32 via 192.168.100.2
   Tunnel1 created 00:11:16, expire 01:48:43
   Type: dynamic, Flags: router used nhop
   NBMA address: 12.12.12.2
192.168.100.1/32 via 192.168.100.1
   Tunnel1 created 00:11:16, expire 01:48:45
   Type: dynamic, Flags: router unique local
   NBMA address: 11.11.11.1
    (no-socket)
```

3) DMVPN Phase III:

Same as Phase 2 but removes some restrictions and complexities of Phase 2. Also allows greater variety of DMVPN network designs we use:
+ **ip nhrp redirect** in hub: tells the initiator spoke to look for a better path to the destination spoke than through the Hub. Upon receiving the NHRP redirect message the spokes communicate with each other over the hub and they have their NHRP replies for the NHRP Resolution Requests that they sent out.
+ **ip nhrp shortcut** in spokes: overwrite the CEF table on the spoke. It basically overrides the next-hop value for a remote spoke network from the default initial hub tunnel IP address to the NHRP resolved remote spoke tunnel IP address)

Note: From the configuration above we can quickly find out which phase of DMVPN is being used when checking an existing DMVPN configuration by looking at the Spoke configuration. If the Spoke's tunnel is configured as mGRE (with the command "tunnel mode gre multipoint") then it is using DMVPN Phase II or Phase III. Next check if the Spokes has the command "ip nhrp shortcut" then it is running DMVPN Phase III.