

# Implement DMVPN Phase3 Multi-subnet Design

## Contents

[Introduction](#)

[Prerequisites](#)

[Requirements](#)

[Components Used](#)

[Background Information](#)

[Problem](#)

[Problem Details](#)

[Solution](#)

[Related Information](#)

## Introduction

This document describes routing considerations in a DMVPN phase3 multi-subnet design to ensure direct spoke to spoke tunnels are correctly built.

## Prerequisites

## Requirements

Cisco recommends that you have knowledge of these topics:

- [Basic knowledge of Dynamic Multipoint VPN \(DMVPN\)](#)
- [Basic knowledge of Next-Hop Resolution Protocol \(NHRP\)](#)

## Components Used

This document is not restricted to specific software and hardware versions.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.

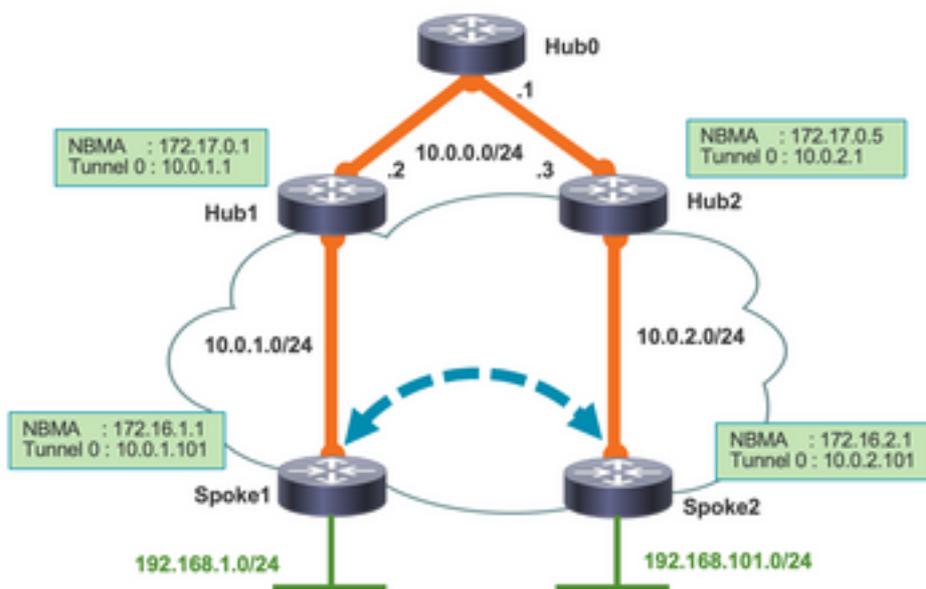
## Background Information

Both DMVPN phase2 and phase3 implementations allow a spoke device to build a direct spoke to spoke tunnel and not need to go through the hub. However, DMVPN phase3 provides much better scalability by with the NHRP redirect mechanism for the spoke to dynamically discover the remote networks through NHRP, and subsequently install NHRP (H) routes into the routing table. This removes the phase2 restriction of requiring each spoke to have specific network prefixes for the remote networks in its routing table. With phase3, the NHRP Resolution Reply from the target

spoke (NBMA exit point) must go over the direct spoke-to-spoke tunnel. However, special consideration must be given to a multi-subnet phase3 design so that spoke-to-spoke tunnel can be built correctly. This article discusses these requirements in detail.

## Problem

DMVPN phase3 can be implemented in either single subnet overlay or multi-subnet overlay. In a single-subnet overlay topology, the hub and all the spoke router tunnel addresses are allocated out of a single logical IP subnet; whereas in a multi-subnet design, spoke to spoke tunnel needs to be built for spokes with their tunnel addresses in different IP subnets. The latter is a common scenario used in a hierarchical DMVPN design shown in the image below.



## DMVPN Phase3 Multi-subnet Topology

### Problem Details

With DMVPN phase3, it is commonly understood that when the NHRP Resolution Request is received, the target spoke initiates the IPsec tunnel to the source spoke, and subsequently, sends the Resolution Reply over that tunnel. However, this is only the case with a single subnet overlay. When the spokes' tunnel interfaces are in different IP logical subnets, the NHRP control packets can traverse via the spoke-hub-spoke path instead of with the direct spoke to spoke tunnel. Here is the sequence of events when Spoke1 sends a Resolution Request to Spoke2 after it receives an NHRP redirect from Hub1:

#### 1. Spoke2 receives Resolution Request

```
*Feb 7 20:57:22.272: NHRP: Receive Resolution Request via Tunnel0 vrf global(0x0), packet
size: 144
*Feb 7 20:57:22.272: (F) afn: AF_IP(1), type: IP(800), hop: 252, ver: 1
*Feb 7 20:57:22.272: sht1: 4(NSAP), sst1: 0(NSAP)
*Feb 7 20:57:22.272: pktsz: 144 extoff: 52
*Feb 7 20:57:22.272: (M) flags: "router auth src-stable nat ", reqid: 5
*Feb 7 20:57:22.272: src NBMA: 172.16.1.1
*Feb 7 20:57:22.272: src protocol: 10.0.1.101, dst protocol: 192.168.101.1
```

```
*Feb 7 20:57:22.272: (C-1) code: no error(0)
*Feb 7 20:57:22.272: prefix: 32, mtu: 17912, hd_time: 900
*Feb 7 20:57:22.272: addr_len: 0(NSAP), subaddr_len: 0(NSAP), proto_len: 0, pref:
255
```

2. Spoke2 adds an implicit NHRP cache entry for 10.0.1.101 by snooping the Resolution Request packet.
3. Spoke2 adds an adjacency for 10.0.1.101 for Tunnel0 with Spoke1's NBMA address.
4. Spoke2 responds with Resolution Reply. Notice at this point, the route for the requesting spoke tunnel address points to Hub2:

```
Spoke2#show ip route 10.0.1.101
Routing entry for 10.0.1.0/24
  Known via "eigrp 1", distance 90, metric 3609600, type internal
  Redistributing via eigrp 1
  Last update from 10.0.2.1 on Tunnel0, 00:17:44 ago
  Routing Descriptor Blocks:
  * 10.0.2.1, from 10.0.2.1, 00:17:44 ago, via Tunnel0
    Route metric is 3609600, traffic share count is 1
    Total delay is 41000 microseconds, minimum bandwidth is 1000 Kbit
    Reliability 255/255, minimum MTU 1400 bytes
    Loading 1/255, Hops 3
Spoke2#
Spoke2#
Spoke2#show ip cef 10.0.1.101
10.0.1.0/24
  nexthop 10.0.2.1 Tunnel0
```

Since the NHRP Control Packet are forwarded along the routed path, it is sent to Hub2 instead of the newly created spoke to spoke tunnel to Spoke1:

```
*Feb 7 20:57:22.360: NHRP: Send Resolution Reply via Tunnel0 vrf global(0x0), packet size:
172
*Feb 7 20:57:22.360: src: 10.0.2.101, dst: 10.0.1.101
*Feb 7 20:57:22.360: (F) afn: AF_IP(1), type: IP(800), hop: 255, ver: 1
*Feb 7 20:57:22.360: shtl: 4(NSAP), sstl: 0(NSAP)
*Feb 7 20:57:22.360: pktsz: 172 extoff: 60
*Feb 7 20:57:22.360: (M) flags: "router auth dst-stable unique src-stable nat ", reqid: 5
*Feb 7 20:57:22.360: src NBMA: 172.16.1.1
*Feb 7 20:57:22.360: src protocol: 10.0.1.101, dst protocol: 192.168.101.1
*Feb 7 20:57:22.360: (C-1) code: no error(0)
*Feb 7 20:57:22.360: prefix: 24, mtu: 17912, hd_time: 900
*Feb 7 20:57:22.360: addr_len: 4(NSAP), subaddr_len: 0(NSAP), proto_len: 4, pref:
255
*Feb 7 20:57:22.360: client NBMA: 172.16.2.1
*Feb 7 20:57:22.360: client protocol: 10.0.2.101
*Feb 7 20:57:22.360: NHRP: Encapsulation succeeded. Sending NHRP Control Packet NBMA
Address: 172.17.0.5
```

In theory, for as long all the intermediate hub(s) can route the NHRP control packet back to Spoke1's tunnel, then everything must still work. But this is not always necessarily the case. If the Resolution Reply cannot be forwarded back to Spoke1, then the direct spoke to spoke tunnel cannot be built.

With a single subnet overlay, this is not an issue because each spoke has a directly connected route to the tunnel network. This results in an adjacency lookup for the requesting spoke tunnel

address before the Resolution Reply is sent back. In a multi-subnet overlay network, since the spokes' tunnel addresses are not on the same IP subnet, the Resolution Reply packet is not guaranteed to be sent over the direct spoke to spoke tunnel.

## Solution

For a multi-subnet DMVPN phase3 design, it is recommended for a spoke to have a routing entry that points out the tunnel interface for any remote spoke tunnel subnet to which it needs to build a direct spoke to spoke tunnel. For example:

```
Spoke2#show run | in ip route
ip route 10.0.101.0 255.255.255.0 Tunnel0
```

This allows the spoke to try to resolve the adjacency for the requesting spoke tunnel address, and subsequently, send the Resolution Reply over the spoke to spoke tunnel.

Alternatively, the Resolution Reply can traverse the spoke-hub-spoke tunnels. In such a case, all the intermediate hub(s) must have a route to the requesting spoke tunnel subnet to ensure the NHRP control packets can be delivered end to end.

**Note:** A bug enhancement was opened to explore options to have the Resolution Reply sent over the direct tunnel even without an explicit static route. Cisco bug ID [CSCvo02022](#) - Enhancement: NHRP must send resolution reply over spoke to spoke tunnel for multi-subnet DMVPN.

**Note:** Only registered Cisco clients can access internal Cisco bug information and tools.

## Related Information

- [Cisco Technical Support & Downloads](#)