Your Time Is Now
Advanced Concepts of DMVPN (Dynamic Multipoint VPN)

Mike Sullenberger – Distinguished Engineer
BRKSEC-3052
Agenda

• DMVPN Design Overview
  • DMVPN General
  • IWAN Specific

• NHRP Details
  • NHRP Overview
  • NHRP Registrations/Resolutions/Redirects

• Recent and New Features
  • Configuration, Resiliency, Routing and Forwarding, Centralized Control
DMVPN Design Overview
Agenda

• DMVPN Design Overview
  • DMVPN General
  • IWAN Specific

• NHRP Details
  • NHRP Overview
  • NHRP Registrations/Resolutions/Redirects

• Recent and New Features
  • Configuration, Resiliency, Routing and Forwarding, Centralized Control
What is Dynamic Multipoint VPN?

DMVPN is a Cisco IOS software solution for building IPsec+GRE VPNs in an easy, dynamic and scalable manner

- Uses two proven technologies
  - Next Hop Resolution Protocol (NHRP)
    - Creates a distributed mapping database of VPN (tunnel int.) to real (public int.) addresses
  - Multipoint GRE Tunnel Interface
    - Single GRE interface to support multiple GRE/IPsec tunnels and endpoints
    - Simplifies size and complexity of configuration
    - Supports dynamic tunnel creation
DMVPN Major Features

- Configuration reduction and no-touch deployment
- Supports:
  - Passenger protocols (IP(v4/v6) unicast, multicast and dynamic Routing Protocols)
  - Transport protocols (NBMA) (IPv4 and IPv6)
  - Remote peers with dynamically assigned transport addresses.
  - Spoke routers behind dynamic NAT; Hub routers behind static NAT.
- Dynamic spoke-spoke tunnels for partial/full mesh scaling.
- Can be used without IPsec Encryption
- Works with MPLS; GRE tunnels and/or data packets in VRFs and MPLS switching over the tunnels
- Wide variety of network designs and options.
## DMVPN Phases

<table>
<thead>
<tr>
<th>Phase 1 – 12.2(13)T</th>
<th>Phase 2 – 12.3(4)T (Phase 1 +)</th>
<th>Phase 3 – 12.4.(6)T (Phase 2 +)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hub and spoke functionality</td>
<td>Spoke to spoke functionality</td>
<td>More network designs and greater scaling</td>
</tr>
<tr>
<td>mGRE or p-pGRE interface on spokes, mGRE on hubs</td>
<td>mGRE interface on spokes</td>
<td>Same Spoke to Hub ratio</td>
</tr>
<tr>
<td>Simplified and smaller configuration on hubs</td>
<td>Direct spoke to spoke data traffic reduces load on hubs</td>
<td>No hub daisy-chain</td>
</tr>
<tr>
<td>Support dynamically addressed CPEs (NAT)</td>
<td>Hubs must interconnect in daisy-chain</td>
<td>Spokes don’t need full routing table – can summarize</td>
</tr>
<tr>
<td>Support for routing protocols and multicast</td>
<td>Spoke must have full routing table – no summarization</td>
<td>Spoke-spoke tunnel triggered by hubs</td>
</tr>
<tr>
<td>Spokes don’t need full routing table – can summarize on hubs</td>
<td>Spoke-spoke tunnel triggered by spoke itself</td>
<td>Remove routing protocol limitations</td>
</tr>
<tr>
<td>No touch deployment</td>
<td>Routing protocol limitations</td>
<td>NHRP routes/next-hops in RIB (15.2(1)T)</td>
</tr>
</tbody>
</table>
DMVPN How it works

• Spokes build a dynamic permanent GRE/IPsec tunnel to the hub, but not to other spokes. They register as clients of the NHRP server (hub).

• When a spoke needs to send a packet to a destination (private) subnet behind another spoke, it queries via NHRP for the real (outside) address of the destination spoke.

• Now the originating spoke can initiate a dynamic GRE/IPsec tunnel to the target spoke (because it knows the peer address).

• The dynamic spoke-to-spoke tunnel is built over the mGRE interface.

• When traffic ceases then the spoke-to-spoke tunnel is removed.
DMVPN Example

Static known IP address

Dynamic unknown IP addresses

LANs can have private addressing

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

Physical: dynamic
Tunnel0: 10.0.0.12

Spoke A

Physical: dynamic
Tunnel0: 10.0.0.11

Spoke B

192.168.0.0/24

192.168.1.0/24

192.168.2.0/24

192.168.1.0/24

192.168.0.0/24
DMVPN Example

Static Spoke-to-hub tunnels
Dynamic Spoke-to-spoke tunnels

192.168.0.0/24
.1

LANs can have private addressing

Spoke A
Physical: dynamic
Tunnel0: 10.0.0.11

Spoke B
Physical: dynamic
Tunnel0: 10.0.0.12

192.168.2.0/24
.1

192.168.1.0/24
.1

Static known IP address
Dynamic unknown IP addresses

Spoke A
192.168.1.0/24

172.17.0.1

Physical: dynamic
Tunnel0: 10.0.0.1

Physical: 172.17.0.1
DMVPN Example

- **Static Spoke-to-hub tunnels**
- **Dynamic Spoke-to-spoke tunnels**

**Static known IP address**
- Spoke A: Physical 172.17.0.1, Tunnel0: 10.0.0.11
- Spoke B: Physical dynamic, Tunnel0: 10.0.0.11

**Dynamic unknown IP addresses**
- Spoke A: 192.168.1.0/24
- Spoke B: 192.168.2.0/24

**LANs can have private addressing**

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
DMVPN Example

Static Spoke-to-hub tunnels

Dynamic Spoke-to-spoke tunnels

LANs can have private addressing

Static known IP address

Dynamic unknown IP addresses

Spoke A

Spoke B

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

Physical: 192.168.1.0/24 .1

Physical: 172.17.0.1
Tunnel0: 10.0.0.11

Tunnel0: 10.0.0.12

Physical: 192.168.0.0/24 .1

Physical: dynamic Tunnel0: 10.0.0.11

192.168.1.0/24 .1

192.168.0.0/24 .1

192.168.2.0/24 .1

LANs can have private addressing
DMVPN and IPv6

- IPv6 Passenger over DMVPN (IPv4 or IPv6) Transport
  - IPv4 or IPv6 infrastructure transport network
  - Both IPv6 and IPv4 can be over the same DMVPN (IPv4 or IPv6) GRE tunnel
  - IPv6 Addresses:
    - NHRP requires IPv6 Unicast Global
    - Routing Protocol requires IPv6 Link-local
    - NHRP automatically registers both Unicast Global and Link-local Addresses

- (IPv4 and/or IPv6) Passenger over DMVPN IPv6 Transport
  - IPv4 and IPv6 transports require separate DMVPNs (mGRE tunnels)
  - Must use IKEv2 for IPsec encryption key management
  - DMVPN IPv4 ↔ DMVPN IPv6 spoke to spoke via hub
  - Standard IPv6 configuration on Outside (WAN) interface
  - WAN interface may support both IPv4 and IPv6
DMVPN over IPv6 Transport – Configuration

**Hub**

```
crypto ikev2 keyring DMVPN
peer DMVPNv6
    address ::/0
    pre-shared-key cisco123v6
crypto ikev2 profile DMVPN
    match identity remote address ::/0
    authentication local pre-share
    authentication remote pre-share
    keyring DMVPN

crypto ipsec profile DMVPN
    set transform-set DMVPN
    set ikev2-profile DMVPN

... interface Tunnel0
    ip address 10.0.0.1 255.255.255.0
    ... ip nhrp map multicast dynamic
    ip nhrp network-id 100000
    ... ipv6 address 2001:DB8:0:100::1/64
    ... ipv6 nhrp map multicast dynamic
    ipv6 nhrp network-id 100006
    ... tunnel source Serial2/0
    tunnel mode gre multipoint ipv6
tunnel protection ipsec profile DMVPN
!
interface Serial2/0
    ip address 172.17.0.1 255.255.255.252
    ipv6 address 2001:DB8:0:FFFF:1::1/126
!
ipv6 route ::/0 Serial2/0
```

**Spoke**

```
crypto ikev2 keyring DMVPN
peer DMVPNv6
    address ::/0
    pre-shared-key cisco123v6
crypto ikev2 profile DMVPN
    match identity remote address ::/0
    authentication local pre-share
    authentication remote pre-share
    keyring DMVPN
    dpd keepalive 30 5 on-demand
crypto ipsec profile DMVPN
    set transform-set DMVPN
    set ikev2-profile DMVPN

... interface Tunnel0
    ip address 10.0.0.11 255.255.255.0
    ... ip nhrp network-id 100000
    ip nhrp nhs 10.0.0.1 nbma 2001:DB8:0:FFFF:1::1 multicast
    ... ipv6 address 2001:DB8:0:100::B/64
    ... ipv6 nhrp network-id 100006
    ipv6 nhrp nhs 2001:DB8:0:100::1 nbma 2001:DB8:0:FFFF:1::1 multicast
    ... tunnel source Serial1/0
    tunnel mode gre multipoint ipv6
tunnel protection ipsec profile DMVPN
!
interface Serial1/0
    ip address 172.16.1.1 255.255.255.252
    ipv6 address 2001:DB8:0:FFFF:0:1:0:1/126
!
ipv6 route ::/0 Serial1/0
```
DMVPN and IPsec

- IPsec integrated with DMVPN, but not required
- Packets Encapsulated in GRE, then Encrypted with IPsec
  - Supports both IKEv1 (ISAKMP) and IKEv2
- NHRP controls the tunnels, IPsec does the encryption
- Bringing up a tunnel
  - NHRP signals IPsec to setup encryption
  - ISAKMP/IKEv2 authenticates peer, generates SAs
  - IPsec responds to NHRP and the tunnel is activated
  - All NHRP and data traffic is Encrypted
- Bringing down a tunnel
  - NHRP signals IPsec to tear down tunnel
  - If encryption is cleared or lost IPsec can signal NHRP to clear the tunnel
  - ISAKMP/IKEv2 Keepalives monitor remote crypto peers

* BFD over DMVPN
DMVPN Encryption Scaling

SLB Design

- Throughput depends on number and types of hub platforms

- (45 GB) ASR1002-HX, ASR1006+/RP2/ESP200
- (18 GB) ASR1006+/RP2/ESP100
- (7 GB) ASR1006+/RP2/ESP40
- (6 GB) ASR1004+/RP2/ESP20
- (4 GB) ASR100[1,2]-X/Integrated
- (2.5 GB) ASR1004+/RP2/ESP10
- (1 GB) 4451-X
- 4351

IMIX Encryption Throughput at 70% Max CPU
Routing over DMVPN

- Supports all routing protocols, except ISIS
- Hubs are routing neighbors with spokes
  - Receive spoke network routes from spokes
  - Advertise spoke and local networks to all spokes
    - Phase 1 & 3: Can Summarize (except OSPF)
    - Phase 2: Cannot summarize (OSPF limited to 2 hubs)
- Hubs are routing neighbors with other hubs
  - Phase 1: Can use different interface and routing protocol than hub-spoke tunnels
  - Phase 2: Must use same tunnel interface and routing protocol as hub-spoke tunnels
  - Phase 3: Can use different tunnel interface and routing protocol than hub-spoke tunnels
- Spokes are only routing neighbors with hubs, not with other spokes
  - Phase 3: Spoke-spoke NHRP “routes” are added directly to routing table (15.2(1)T)
Routing Table Example (Spoke)

**Phase 1 & 3**
(with summarization)

- C 172.16.1.0/30 is directly connected, Serial1/0
- C 10.0.0.0/24 is directly connected, Tunnel0
- C 192.168.1.0/24 is directly connected, Ethernet0/0
- S* 0.0.0.0/0 is directly connected, Serial1/0
- D 192.168.0.0/16 via 10.0.0.1, 00:00:08, Tunnel0

**Phase 1 & 3**
(without summarization, next-hop not preserved)

- C 172.16.1.0/30 is directly connected, Serial1/0
- C 10.0.0.0/24 is directly connected, Tunnel0
- D 192.168.0.0/24 via 10.0.0.1, 00:02:36, Tunnel0
- C 192.168.1.0/24 is directly connected, Ethernet0/0
- D 192.168.2.0/24 via 10.0.0.1, 00:02:36, Tunnel0
- D 192.168.3.0/24 via 10.0.0.1, 00:02:36, Tunnel0
- S* 0.0.0.0/0 [1/0] via 172.16.1.1

**Phase 2**
(no summarization, next-hop preserved)

- C 172.16.1.0/30 is directly connected, Serial1/0
- C 10.0.0.0/24 is directly connected, Tunnel0
- D 192.168.0.0/24 via 10.0.0.1, 00:42:34, Tunnel0
- C 192.168.1.0/24 is directly connected, Ethernet0/0
- D 192.168.2.0/24 via 10.0.0.1, 00:42:34, Tunnel0
- D 192.168.3.0/24 via 10.0.0.1, 00:42:34, Tunnel0
- S* 0.0.0.0/0 [1/0] via 172.16.1.1
Routing Protocols over DMVPN

EIGRP

• Distance Vector style matches with DMVPN NBMA network style
  • Feasible successor for quick spoke-to-hub convergence
• Good scaling with reasonably fast convergence (hello 5, hold 15)
• Good metric control (automatic and/or manual)
  • Change metrics, route tagging, filtering or summarization at hub and/or spoke
  • Can be used to control load-balancing of spoke ↔ hub(s) traffic
  • Automatic metric increase per DMVPN hop
• Feature additions (Phase 2)
  • Equal Cost MultiPath (15.2(3)T, 15.2(1)S)
  • Add-path (15.3(1)S)
Routing Protocols over DMVPN

BGP

- Base Distance Vector style matches with DMVPN NBMA network style
  - iBGP (recommended)
    - Dynamic Neighbors, MED to control/compare routes;
    - May need iBGP local-as (15.2(2)T, 15.1(3)S)
  - eBGP (okay)
    - AS-Path length to control/compare routes
- Good scaling but with slower convergence (hello 15+, hold 45+)
- Good metric control (manual)
  - Change metrics, route tagging, filtering or summarization at hub and/or spoke
  - Can be used to control load-balancing of spoke ➔ hub(s) traffic
  - Only manual metric increase per DMVPN hop
- Some issues with Equal Cost multi-path (ECMP) route selection
  - Between multiple DMVPNs and preserving correct next-hop
  - Spoke-spoke tunnel load-balancing for spoke sites with multiple spoke routers
Routing Protocols over DMVPN

OSPF

• Link-state style doesn’t match as well with DMVPN NBMA network style

• Area issues – DMVPN requires single Area
  • Area 0 over DMVPN – spoke sites can be in different areas
    • But, area 0 is extended over WAN – possible stability issues for Area 0
  • Non-Area 0 over DMVPN – all spoke sites in this same single area
  • Multi-subnet DMVPN can be used to have multiple OSPF areas
    • Increase in complexity of DMVPN and OSPF design

• More difficult metric control
  • Can only change metrics, filter or summarize at area boundaries
  • Automatic metric increase per DMVPN hop
  • Slight metric issue for failover path between multiple DMVPNs

• No issues with Equal Cost multi-path (ECMP) route selection
Routing Protocol?

• Which routing protocol should I use?
  • In general you would use the same routing protocol over DMVPN that you use in the rest of your network, or over other WAN networks (like MPLS).

• BUT...
  • EIGRP being an advanced distance vector protocol matches really well with DMVPN network topologies
  • BGP, specifically iBGP, runs well over DMVPN, but is more complicated to setup to have it act more like an IGP than an EGP
  • OSPF can run over DMVPN, BUT lower scaling and Area 0 issues can complicate the network
  • RIP can be used, but has longer hold time and limited metric values
  • IS-IS cannot be used since it doesn’t run over IP
Routing Protocol Scaling

SLB design using BGP or EIGRP

<table>
<thead>
<tr>
<th>Number of Branches</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>3000</td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td></td>
</tr>
<tr>
<td>5000</td>
<td></td>
</tr>
<tr>
<td>6000</td>
<td></td>
</tr>
</tbody>
</table>

OSPF
- 3945E
- 4451-X

EIGRP
- 3945E
- 4451-X
- (ASR100x-X) ASR1004+/RP2/ESP20+

BGP
- 3945E
- 4451-X
- ASR1004+/RP2/ESP20+
- (ASR100x-X) ASR1004+/RP2/ESP100+
Redundancy

- Active-active redundancy model – two or more hubs per spoke
  - All configured hubs are active and are routing neighbors with spokes
    - Can use Backup NHS feature to activate a subset of configured hubs
  - Routing protocol routes are used to determine traffic forwarding
    - Single route: one tunnel (hub) at a time – primary/backup mode
    - Multiple routes: multiple tunnels (hubs) – load-balancing mode (CEF, PfR)

- (ISAKMP/IKEv2)/IPsec
  - Cannot use IPsec Stateful failover (NHRP isn’t supported)
  - Invalid SPI recovery is not useful with DMVPN
    - no crypto isakmp invalid-spi-recovery
  - ISAKMP/IKEv2 keepalives on spokes for timely hub recovery*

* BFD over DMVPN
Redundancy (cont)

- Can use single or multiple DMVPNs for redundancy
  - Each mGRE interface is a separate DMVPN network using
    - Same: Tunnel source (optional).
    - Different: NHRP network-id and IP subnet, (or no) Tunnel key
    - When using same tunnel source → different tunnel keys and same IPsec profile (name) with shared
      tunnel protection ipsec profile name shared
  - Can “glue” mGRE interfaces into same DMVPN network (Phase 3 only)
    - Same: NHRP network-id and authentication, Tunnel key (optional)
    - Different: Tunnel source and IP subnet

- Spokes – at least two hubs (NHSs)
  - Phase 1: (Hub-and-spoke)
    - p-pGRE interfaces → two or more DMVPN networks, one hub (NHS) on each
  - Phase 1, 2 or 3: (Hub-and-spoke or Dynamic Mesh)
    - mGRE interface → one DMVPN network, two or more hubs (NHSs)
Redundancy (cont.)

• Hubs – interconnect and routing
  • Phase 1: (Hub and spoke only)
    • Interconnect hubs directly over physical link, p-pGRE or mGRE tunnel
    • Can exchange routing through any of these paths
    • Same or different routing protocol as with spokes
  • Phase 2: (Dynamic Mesh)
    • Must interconnect hubs over same mGRE tunnel as spokes, daisy-chain as NHSs
    • Must exchange routing over DMVPN network
    • Must use same routing protocol as with spokes
  • Phase 3: (Dynamic Mesh)
    • Interconnect hubs over same or different mGRE tunnel (same NHRP Network-id)
    • Must exchange routing over a DMVPN network
    • Same or different routing protocol as with spokes
Spoke-Spoke Tunnels – Considerations

• Resiliency
  • No direct monitoring of spoke-spoke tunnel (use ISAKMP/IKEv2 keepalives)*
    crypto [isakmp keepalive | ikev2 dpd] initial retry [on-demand | periodic]
    crypto [isakmp | ikev2] nat keepalive interval
  • Path Selection
    • NHRP will always build spoke-spoke tunnel
    • No bandwidth/latency measurement of spoke-spoke vs. spoke-hub-spoke paths
    • Can do interesting things with Smart-spoke feature
  • Overloading spoke routers
    • CPU or memory → IKE Call Admission Control (CAC)
      • crypto call admission limit ike {sa | in-negotiation} max-SAs
      • call admission limit percent
      • show crypto call admission statistics
    • Bandwidth → Design for expected traffic
      • Hub-spoke versus Spoke-spoke; Spoke-spoke availability is best effort

  * BFD over DMVPN
Best Practices

• mGRE Tunnel configuration
  • Both Hubs and Spokes
    • `tunnel source <interface>`
    • `bandwidth <WAN-interface>` (as starting point, may adjust)
    • `ip mtu 1400; ip tcp adjust-mss 1360`

• NHRP
  • Spokes
    • `ip nhrp shortcut`
    • `ip nhrp nhs <hub-tunnel> nbma <hub-nbma-ip | hub-fqdn> multicast (12.4(20)T)`
  • Hubs
    • `ip nhrp redirect`
    • `ip nhrp map multicast dynamic`
    • `ip nhrp server-only`
Best Practices (cont)

• Crypto
  • crypto [isakmp keepalive | ikev2 dpd] initial retrans on-demand – Spokes only
    • Initial = 30 (> RP hello); retrans = 5 → 55 seconds for neighbor down

• Routing
  • Phase 2 – RP advertises routes with remote spoke as the next-hop
    • EIGRP: (hubs) no ip [next-hop-self | split-horizon] eigrp as, (all) use delay to adjust metric
    • OSPF: (all) ip ospf network broadcast; (spokes only) ip ospf priority 0
    • BGP: iBGP (hubs) route-reflectors; (spokes) neighbor hub next-hop-self
  • Phase 3 – RP advertises routes with the hub as the next-hop
    • EIGRP: (hubs) no ip split-horizon eigrp <as>
    • OSPF: (all) ip ospf network point-multipoint; prefix-suppression (suppress /32 routes)
    • BGP: iBGP (hubs) route-reflectors; (all) neighbor [hub | spoke] next-hop-self all
  • To manipulate path selection through DMVPN use:
    • EIGRP: delay not bandwidth; OSPF: cost; iBGP: MED, Local-pref
Cisco IOS Code and Platform Support

- 3900(E), 2900, 1900, 890, 819, 880
  - 15.2.4M10, 15.3.3M7*, 15.4.3M5*, 15.5.3M3**
  - 15.4.2T4, 15.5.2T3+, 15.6.2T+

- ASR100[2,4,6](RP2), ASR100[1,2]-X, 4451-X, 4431, 4300
  - (3.10.7S)15.3.3S7**, (3.13.5S)15.4.3S5*, (3.16.3S)15.5.3S3*
  - (3.15.3S)15.5.2S3, (3.17.2S)15.6.1S2

- ASR1002-HX, ASR100[6,9]-X
  - (3.16.3S)15.5.3S3+, (3.17.2S)15.6.1S2+
  - 16.2.1

- CSR1000V
  - (3.13.5S)15.4.3S5*, (3.15.2S)15.5.2S2, (3.16.3S)15.5.3S3*, (3.17.2S)15.6.1S2

* Recommended

+ N/A for 881-887

+ N/A for 4431,4300,ASR1001-X

+ N/A for ASR1002-HX
Basic DMVPN Designs

- Hub-and-spoke – Order(n)
  - Spoke-to-spoke traffic via hub
    - Phase 1: Hub bandwidth and CPU limit VPN
    - SLB: Many “identical” hubs; increases CPU and bandwidth limits
  - Spoke-to-spoke – Order(n) « Order(n^2)
    - Control traffic; Hub and spoke; Hub to hub
      - Phase 2: (single)
      - Phase 3: (hierarchical)
    - Unicast Data traffic; Dynamic mesh
      - Spoke routers support spoke-hub and spoke-spoke tunnels currently in use.
      - Hub supports spoke-hub traffic and overflow from spoke-spoke traffic.

- Network Virtualization
  - VRF-lite; Multiple DMVPNs (one per VRF)
  - MPLS over DMVPN (2547oDMVPN); Single DMVPN (many VRFs)
Basic DMVPN Designs

Dual DMVPN Single Hub
Single mGRE tunnel on Hub,
two p-pGRE tunnels on Spokes

Single DMVPN Dual Hub
Single mGRE tunnel on all nodes
Multiple DMVPNs versus Single DMVPN

• Multiple DMVPNs
  • Best for Hub-and-spoke only
    • Easier to manipulate RP metrics between DMVPNs for Load-sharing
      • EIGRP – Route tags, Delay; iBGP – Communities, MED; OSPF – Cost
    • Performance Routing (PfR) selects between interfaces
  • Load-balancing over multiple ISPs (physical paths)
    • Load-balance data flows over tunnels → Better statistical load-balancing

• Single DMVPN
  • Best for spoke-spoke DMVPN
    • Can only build spoke-spoke within a DMVPN not between DMVPNs*
    • Slightly more difficult to manipulate RP metrics within DMVPN for Load-sharing
      • EIGRP – Route tags, delay; iBGP – Communities, MED; OSPF – Can’t do
  • Load-balancing over multiple ISPs (physical paths)
    • Load-balance tunnel destinations over physical paths → Worse statistical load-balancing
DMVPN Combination Designs

Retail/Franchise

Spoke-to-hub tunnels
Spoke-to-spoke tunnels

Dual ISP

ISP 1
ISP 2

Spoke-to-hub tunnels
Spoke-to-spoke tunnels
Spoke-hub-hub-spoke tunnel
DMVPN Combination Designs (cont)

Hierarchical

Server Load Balancing

- Spoke-to-hub tunnels
- Spoke-to-spoke tunnels

- Spoke-to-hub tunnels
- Spoke-to-spoke tunnels
- Hub-to-hub tunnel
Network Virtualization
Separate DMVPN mGRE tunnel per VRF (VRF-lite)

- Hub routers handle all DMVPNs
  - Multiple Hub routers for redundancy and load
- IGP used for routing protocol over DMVPNs on Spokes and Hubs
  - Address family per VRF
  - Routing neighbor per spoke per VRF
- BGP used only on the hub
  - Redistribute between IGP and BGP for import/export of routes between VRFs
  - “Internet” VRF for Internet access and routing between VRFs
- Global routing table used for routing DMVPN tunnel packets
Network Virtualization
MPLS over DMVPN – 2547oDMVPN

- MPLS VPN over DMVPN
  - Single DMVPN/mGRE tunnel on all routers
  - Multiple Hub routers for redundancy and load

- MPLS configuration – routers are PEs
  - Spoke to spoke via hub and direct shortcut
  - MPLS labels via NHRP, ‘mpls nhrp’ (15.4(1)S, 15.4(2)T)
    - Replaces ‘mpls ip’; No LDP

- Routing
  - Global for routing DMVPN tunnel packets
  - IGP for routing outside of DMVPN
  - MP-BGP for routing over DMVPN
    - Redistribute between IGP and BGP for over DMVPN
    - Import/export routes between VRFs and Global (or Internet VRF)
  - One routing neighbor per spoke
# VPN Selection

<table>
<thead>
<tr>
<th>Use Case/ Solution</th>
<th>DMVPN (mGRE, p-pGRE)</th>
<th>GETVPN (Tunnel-less)</th>
<th>FlexVPN (dVTI, IKEv2)</th>
<th>SSLVPN (TLS)</th>
<th>Easy VPN (IKEv1)</th>
<th>IPsec VPN (CM, sVTI, p-pGRE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote Access</td>
<td>N-R</td>
<td>N-S</td>
<td>R</td>
<td>R</td>
<td>N-R</td>
<td>N-R</td>
</tr>
<tr>
<td>Hub-Spoke (HS)</td>
<td>R</td>
<td>N-S</td>
<td>R</td>
<td>N-R</td>
<td>N-R</td>
<td>N-R</td>
</tr>
<tr>
<td>HS + Spoke-Spoke</td>
<td>R</td>
<td>R</td>
<td>N-R</td>
<td>N-R</td>
<td>N-S</td>
<td>N-S</td>
</tr>
<tr>
<td>IoT</td>
<td>R</td>
<td>N-R</td>
<td>R</td>
<td>R</td>
<td>N-R</td>
<td>N-R</td>
</tr>
<tr>
<td>IWAN</td>
<td>R</td>
<td>N-S</td>
<td>N-S</td>
<td>N-S</td>
<td>N-S</td>
<td>N-S</td>
</tr>
<tr>
<td>MPLS over xVPN</td>
<td>R MPLS-o-DMVPN</td>
<td>R MPLS-o-mGRE</td>
<td>N-R MPLS-o-Flex</td>
<td>N-S</td>
<td>N-S</td>
<td>N-R MPLS-o-GRE</td>
</tr>
</tbody>
</table>

R = Recommended  
N-R = Not Recommended  
N-S = Not Supported
Agenda

• DMVPN Design Overview
  • DMVPN General
  • IWAN Specific

• NHRP Details
  • NHRP Overview
  • NHRP Registrations/Resolutions/Redirects

• Recent and New Features
  • Configuration, Resiliency, Routing and Forwarding, Centralized Control
Intelligent WAN Solution Components

**Transport Independent**
- Consistent operational model
- Simple provider migrations
- Scalable and modular design
- DMVPN IPsec overlay design

**Intelligent Path Control**
- Application best path based on delay, loss, jitter, path preference
- Load balancing for full utilization of all bandwidth
- Improved network availability
- Performance Routing (PfR)

**Application Optimization**
- Application monitoring with Application Visibility and Control (AVC)
- Application Acceleration and bandwidth savings with WAAS

**Secure Connectivity**
- Certified strong encryption
- Comprehensive threat defense with ASA and IOS firewall/IPS
- Cloud Web Security (CWS) for scalable secure direct Internet access
Flexible Secure WAN Design Over Any Transport
Dynamic Multipoint VPN

Transport-Independent

Simplifies WAN Design
- Easy multi-homing over any carrier service offering
- Single routing control plane with minimal peering to the provider

Flexible

Dynamic Full-Meshed Connectivity
- Consistent design over all transports
- Automatic site-to-site IPsec tunnels
- Zero-touch hub configuration for new spokes

Secure

Proven Robust Security
- Certified crypto and firewall for compliance
- Scalable design with high-performance cryptography in hardware
DMVPN design with IWAN

- Multiple DMVPNs
  - One per physical transport network
  - Path diversity
  - Separate failure domains
- Each Phase 3 DMVPN
  - Single layer hub-and-spoke; hierarchical not currently supported
  - Physical WAN interface in f-VRF
  - Single Hub; Multi-Hub
    - PfRv3 Multi-NH and Multi-DC feature (15.5(3)S, 15.5(3)M)
  - Spoke-Spoke dynamic tunnels
  - Per-Tunnel QOS
- PfRv3 interoperability
  - Dynamic path selection
    - Per application
    - Load Balancing
    - Brownout circumvention
  - Communicates with NHRP via RIB
    - Triggers secondary spoke-spoke tunnels
- Single Overlay Routing Domain
  - Simplified operations and support
  - Simple ECMP load-balancing and primary path provisioning
  - EIGRP or BGP
    - PfRv3 gets secondary path directly from RP
Basic DMVPN Design for IWAN

Dual DMVPN Dual Hub

192.168.10.0/24

Physical: 172.16.0.1
Tunnel0: 10.0.0.1
Loop0: 172.18.0.1

Dual DMVPN

192.168.100.0/24

Physical: 172.16.0.5
Tunnel0: 10.0.0.2
Loop0: 172.18.1.1

Internet DMVPN

192.168.20.0/24

Physical: 172.17.0.1
Tunnel0: 10.0.1.2
Loop0: 172.18.1.2

MPLS DMVPN

Spoke A

192.168.1.0/24

Physical: (dynamic)
Tunnel0: 10.0.0.11
Tunnel1: 10.0.1.11
Loop0: 172.18.0.11

Physical: (dynamic)
Tunnel0: 10.0.0.12
Tunnel1: 10.0.1.12
Loop0: 172.18.0.12

Physical: 192.168.3.0/24

Support coming in March (MTT)

Spoke B1

192.168.2.0/24

Spoke B2

192.168.10.0/24

192.168.100.0/24

192.168.20.0/24

192.168.3.0/24
NHRP Details
Agenda

- DMVPN Design Overview
  - DMVPN General
  - IWAN Specific
- NHRP Details
  - NHRP Overview
  - NHRP Registrations/Resolutions/Redirects
- Recent and New Features
  - Configuration, Resiliency, Routing and Forwarding, Centralized Control
NHRP Message Types

• Registration
  • Build base hub-and-spoke network for control and data traffic
    (Phase 1 and 2 – single layer, Phase 3 – hierarchical)

• Resolution – Phase 2 and 3
  • Get mapping to build dynamic spoke-spoke tunnels

• Traffic Indication (Redirect) – Phase 3
  • Trigger resolution requests at previous GRE tunnel hop

• Purge
  • Clear out stale dynamic NHRP mappings

• Error
  • Signal error conditions
NHRP Main Functionality

• NHRP Registrations
  • Static NHRP mappings on spokes for Hub (NHS)
  • Spoke (NHC) dynamically registers its VPN to NBMA address mapping with hub (NHS)

• NHRP Resolutions – Phase 2 and 3
  • Dynamically resolve spoke to spoke VPN to NBMA mapping for spoke-spoke tunnels
    • Phase 2 – NHC self triggers to send NHRP Resolution request
    • Phase 3 – NHC triggered by first hop NHS to send NHRP Resolution request
  • NHRP Resolution requests sent via hub-and-spoke or direct spoke-spoke path
  • NHRP Resolution replies sent via direct spoke-spoke path

• NHRP Redirects (Traffic Indication) – Phase 3
  • Data packets forwarded via NHS, which “hairpins” data packets back onto DMVPN
  • NHS sends redirect message to “trigger” NHC to resolve direct spoke-spoke path
NHRP Message Extension Types

• Responder Address Extension:
  • Address mapping for Responding node (Reply messages)

• Forward Transit NHS Record Extension:
  • List of NHSs that NHRP request message traversed – copied to reply message

• Reverse Transit NHS Record Extension:
  • List of NHSs that NHRP reply message traversed

• Authentication Extension:
  • NHRP Authentication (clear-text)

• NAT Address Extension:
  • Address mapping: For peer (Registration request/reply); For self (Resolution request/reply)

• Cisco Vendor Extension
  • NHRP Group name
  • Smart-spoke attributes (name; value)
NHRP Mapping Entries

- **Static**
  - Both host (/32, /128) and network (/<x>) mappings

- **Dynamic**
  - Registered (/32, /128)
    - From NHRP Registration
    - NAT – record both inside and outside address
  - Learned (/32, /128 or /<x>)
    - From NHRP Resolution
    - NAT – record both inside and outside address

- **Incomplete (/32, /128)**
  - Rate-limit NHRP Resolution Requests
  - Data packets process-switched via NHS while building spoke-spoke tunnels. (Phase 2)

- **Temporary (/32) (12.4(22)T)**
  - Same as “Incomplete” mapping except that NBMA is set to Hub
  - Data packets CEF-switched via NHS while building spoke-spoke tunnels. (Phase 2)

- **Local (/32, /128 or /<x>)**
  - Mapping for local network sent in an NHRP Resolution Reply
  - Record which nodes were sent this mapping

- **(no socket)**
  - Not used to forward data packets
  - Do not trigger IPsec encryption
  - Set on Local entries
NHRP Mapping Entries

Static

10.0.0.1/32 via 10.0.0.1, Tunnel0 created 01:20:10, never expire
Type: static, Flags: used
NBMA address: 172.17.0.9

Registered

10.0.0.19/32 via 10.0.0.19, Tunnel0 created 01:20:08, expire 00:05:51
Type: dynamic, Flags: unique registered used
NBMA address: 172.16.3.1

NAT

10.0.0.18/32 via 10.0.0.18, Tunnel0 created 00:16:09, expire 00:05:50
Type: dynamic, Flags: unique registered used
NBMA address: 172.18.0.2
(Claimed NBMA address: 172.16.2.1)

Resolution

192.168.23.0/24 via 10.0.0.19, Tunnel0 created 00:00:11, expire 00:05:48
Type: dynamic, Flags: router used
NBMA address: 172.16.3.1

Incomplete

10.0.0.45/32, Tunnel0 created 00:00:21, expire 00:02:43
Type: incomplete, Flags: negative
Cache hits: 2

Temporary

10.0.0.17/32 via 10.0.2.17, Tunnel0 created 00:00:09, expire 00:02:55
Type: dynamic, Flags: used temporary
NBMA address: 172.17.0.9

Local (no-socket)

192.168.15.0/24 via 10.0.0.11, Tunnel0 created 00:05:39, expire 00:05:50
Type: dynamic, Flags: router unique local
NBMA address: 172.16.1.1
(no-socket)
## NHRP Mapping Flags

<table>
<thead>
<tr>
<th>Flag</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>unique</td>
<td>Mapping entry is unique, don’t allow overwrite with new NBMA</td>
</tr>
<tr>
<td>registered</td>
<td>Mapping entry from an NHRP registration</td>
</tr>
<tr>
<td>authoritative</td>
<td>Mapping entry can be used to answer NHRP resolution requests</td>
</tr>
<tr>
<td>used</td>
<td>Mapping entry was used in last 60 seconds to forward data traffic</td>
</tr>
<tr>
<td>router</td>
<td>Mapping entry for remote router</td>
</tr>
<tr>
<td>implicit</td>
<td>Mapping entry from source information in NHRP resolution request packet</td>
</tr>
<tr>
<td>local</td>
<td>Mapping entry for a local network, record remote requester</td>
</tr>
<tr>
<td>nat</td>
<td>Remote peer supports the NHRP NAT extension</td>
</tr>
<tr>
<td>rib</td>
<td>Routing Table entry created</td>
</tr>
<tr>
<td>nho</td>
<td>Next-Hop-Override Routing Table entry created</td>
</tr>
<tr>
<td>nhop</td>
<td>Explicit Next-Hop route out tunnel interface added to RIB/FIB</td>
</tr>
<tr>
<td>nf</td>
<td>Non-forwarding Entry (No Socket)</td>
</tr>
</tbody>
</table>
Agenda

- DMVPN Design Overview
  - DMVPN General
  - IWAN Specific

- NHRP Details
  - NHRP Overview
  - NHRP Registrations/Resolutions/Redirects

- Recent and New Features
  - Configuration, Resiliency, Routing and Forwarding, Centralized Control
Hub-and-Spoke – Features

- GRE, NHRP and IPsec configuration
  - p-pGRE or mGRE on spokes; mGRE on hubs
  - ISAKMP/IKEv2 Authentication
    - Certificate (PKI), (Pairwise/Wildcard) Pre-shared Key (PSK)

- NHRP Registration
  - Spoke has static NHRP mapping for Hubs
  - Hub dynamically learns Spoke’s NHRP mapping
    - Handles dynamically addressed spokes (DHCP, NAT, …)
  - NAT detection support
NHRP Registration

• Builds base hub-and-spoke network
  • Hub-and-spoke data traffic
  • Control traffic; NHRP, Routing protocol, IP multicast
  • Phase 2 – Single layer hub-and-spoke
  • Phase 3 – Hierarchical hub-and-spoke (tree).

• Next Hop Client (NHC) has static mapping for Next Hop Servers (NHSs)
  • NHC dynamically registers own mapping with NHS
    • Supports spokes with dynamic NBMA addresses or NAT
    • Reports outside address of Hub (if Hub behind NAT)
    • NHRP-group for per-Tunnel QoS (H→S)
    • IPv6: Includes both Unicast-Global and Link-local spoke mappings

• NHS registration reply gives liveliness of NHS
  • Supplies outside NAT address of spoke (Spoke behind NAT?)
  • NHRP-group for per-Tunnel QoS (H→S)
  • IPv6: Includes link-local address hub mapping (needed by EIGRP; OSPF)
NHRP Registration

Before Building Spoke-Hub Tunnels

- **Spoke A**
  - Physical: 172.17.0.1
  - Tunnel0: 10.0.0.1

- **Spoke B**
  - Physical: (dynamic)
  - Tunnel0: 10.0.0.11
  - Tunnel0: 10.0.0.12

- **10.0.0.1 ➔ 172.17.0.1**
- **192.168.1.1/24 ➔ Conn.**
- **192.168.0.0/24 ➔ Conn.**
- **192.168.2.0/24 ➔ Conn.**

- **NHRP mapping**
- **Routing Table**

- **192.168.0.1/24 ➔**
- **192.168.0.0/24 ➔ Conn.**

= Dynamic permanent IPsec tunnels
NHRP Registration

Building Spoke-Hub Tunnels

NHRP Registration

192.168.0.1/24
10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1
192.168.0.0/24 → Conn.

NHRP mapping

Routing Table

Spoke A
Physical: 172.17.0.1
Tunnel0: 10.0.0.1
192.168.1.1/24
10.0.0.1 → 172.16.1.1
192.168.1.0/24 → Conn.

Spoke B
Physical: 172.16.2.1
Tunnel0: 10.0.0.12
192.168.2.1/24
10.0.0.1 → 172.17.0.1
192.168.2.0/24 → Conn.

192.168.0.1/24
10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1
192.168.0.0/24 → Conn.

= Dynamic permanent IPsec tunnels
NHRP Registration
Building Spoke-Hub Tunnels

NHRP Registration
NHRP mapping
Routing Table

192.168.0.1/24

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

192.168.0.0/24 → Conn.

Physical: 172.16.1.1
Tunnel0: 10.0.0.11

Spoke A

192.168.1.1/24

10.0.0.11 → 172.17.0.1

192.168.1.0/24 → Conn.

10.0.0.11

10.0.0.1

172.17.0.1

172.16.1.1

172.16.2.1

10.0.0.12

Spoke B

192.168.2.0/24 → Conn.

192.168.2.0/24

10.0.0.1

172.17.0.1

10.0.0.11

172.16.1.1

172.16.2.1

10.0.0.12

192.168.2.2/24

NHRP mapping
Routing Table

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

Spoke A

192.168.1.1/24

10.0.0.11 → 172.17.0.1

192.168.1.0/24 → Conn.

10.0.0.1

172.17.0.1

10.0.0.11

172.16.1.1

172.16.2.1

10.0.0.12

Spoke B

192.168.2.0/24 → Conn.

192.168.2.0/24

10.0.0.1

172.17.0.1

10.0.0.11

172.16.1.1

172.16.2.1

10.0.0.12

Physical: 172.16.1.1
Tunnel0: 10.0.0.11

Physical: (dynamic)
Tunnel0: (dynamic)

NHRP mapping
Routing Table

Physical: (dynamic)
Tunnel0: (dynamic)

10.0.0.1

172.17.0.1

10.0.0.11

172.16.1.1

172.16.2.1

10.0.0.12

10.0.0.1

= Dynamic permanent IPsec tunnels
NHRP Registration

Routing Adjacency

Routing packet

NHRP mapping

Routing Table

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

192.168.0.1/24

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

192.168.0.0/24 → Conn.
192.168.0.0/16 → Summ.

Spoke A

Physical: 172.16.1.1
Tunnel0: 10.0.0.11

192.168.1.1/24

10.0.0.1 → 172.17.0.1
192.168.1.0/24 → Conn.

Spoke B

Physical: 172.16.2.1
Tunnel0: 10.0.0.12

192.168.2.1/24

10.0.0.1 → 172.17.0.1
192.168.2.0/24 → Conn.

= Dynamic permanent IPsec tunnels
NHRP Registration
Routing Adjacency

Routing packet
NHRP mapping
Routing Table

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

192.168.0.1/24
Physical: 172.16.1.1
Tunnel0: 10.0.0.11

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

192.168.0.0/24 → Conn.
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12
192.168.0.0/16 → Summ.

192.168.1.1
192.168.2.1

10.0.0.11 → 172.17.0.1
192.168.0.0/16 → 10.0.0.1
192.168.1.0/24 → Conn.

Spoke A

192.168.1.1/24

10.0.0.1
172.17.0.1

tunnel0: 10.0.0.1

Spoke B

192.168.2.1/24

10.0.0.12
172.16.2.1

tunnel0: 10.0.0.12

= Dynamic permanent IPsec tunnels

NHRP mapping
Routing Table
NHRP Registration
Routing Adjacency

Routing packet

NHRP mapping

Routing Table

Physical: 172.16.1.1
Tunnel0: 10.0.0.1

192.168.0.1/24

192.168.0.0/16

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

192.168.0.0/24 → Conn.
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12
192.168.0.0/16 → Summ.

192.168.0.1/24

192.168.1.0/24 → Conn.
192.168.2.0/24 → Conn.

192.168.1.0/24

192.168.2.0/24

10.0.0.1 → 172.17.0.1

192.168.0.0/16

10.0.0.1 → 172.17.0.1

192.168.0.0/16 → 10.0.0.1
192.168.1.0/24 → Conn.

= Dynamic permanent IPsec tunnels

Spoke A

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

192.168.1.1/24

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

192.168.0.0/24 → Conn.
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12
192.168.0.0/16 → Summ.

192.168.1.1/24

192.168.2.1/24

Spoke B

Physical: 172.16.2.1
Tunnel0: 10.0.0.12

192.168.0.1/24

192.168.0.0/24

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

192.168.0.0/24 → Conn.
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12
192.168.0.0/16 → Summ.
Hub-and-Spoke

Data Packet Forwarding

- Process-switching
  - Routing table selects outgoing interface and IP next-hop
  - NHRP looks up packet IP destination to select IP next-hop, overriding IP next-hop from routing table.
    - Could attempt to trigger spoke-spoke tunnel
      - ‘tunnel destination …’ ➔ Can only send to hub
      - ‘ip nhrp server-only’ ➔ Don’t send NHRP resolution request
  - If no matching NHRP mapping then send to NHS (hub)

- CEF switching
  - IP Next-hop from FIB table (Routing table)
    - IP Next-hop ➔ Hub ➔ data packets send to Hub
  - Adjacency will be complete so CEF switch packet to hub
    - NHRP not involved
Agenda

- DMVPN Design Overview
  - DMVPN General
  - IWAN Specific

- NHRP Details
  - NHRP Overview
  - NHRP Registrations/Resolutions/Redirects

- Recent and New Features
  - Configuration, Resiliency, Routing and Forwarding, Centralized Control
Phase 2 – Features

• Single mGRE interface with ‘tunnel protection …’
  • On Hubs and Spokes
  • Hubs must be inter-connected in a “Daisy chain” over same mGRE tunnel
  • IKE authentication information (Certificates, Wildcard Pre-shared Keys)

• Spoke-spoke data traffic direct
  • Reduced load on hub
  • Reduced latency
    • Single IPsec encrypt/decrypt

• Routing Protocol
  • Still hub-and-spoke
  • Hub cannot summarize spoke routes
  • Routes on spokes must have IP next-hop of remote spoke (preserve next-hop)
Phase 3 – Features

• Increase scale
  • Increase number of spokes, with same spoke/hub ratio
  • Distribution hubs off load central hub
    • Manage local spoke-spoke tunnels
    • IP multicast and routing protocol

• No hub daisy-chain
  • Use RIB to forward NHRP packets through NHSs
  • Reduces complexity and load for routing protocol

• OSPF not limited to 2 hubs
  • Network point-multipoint mode
  • Single OSPF area; No summarization

• Spokes don’t need full routing tables
  • Can summarize routes at the hub
  • Reduce space and load on spokes
  • Reduce routing protocol load on hub
    • 1000 spokes, 1 route per spoke;
    • hub advertises 1 route to 1000 spokes → 1000 advertisements

• Don’t recommend mixing Phase 2 and 3 on same DMVPN
  • Build separate Phase 3 DMVPN
  • Migrate spokes from Phase 2 DMVPN to Phase 3 DMVPN
  • Remove Phase 2 DMVPN
Phase 3 – Building Spoke-spoke Tunnels

- Originating spoke
  - IP Data packet is forwarded out tunnel interface to destination via Hub (NHS)
- Hub (NHS)
  - Receives and forwards data packet on tunnel interfaces with same NHRP Network-id.
  - Sends NHRP Redirect message to originating spoke.
- Originating spoke
  - Receives NHRP redirect message
  - Sends NHRP Resolution Request for Data IP packet destination
- Destination spoke
  - Receives NHRP Resolution Request
  - Builds spoke-spoke tunnel
  - Sends NHRP Resolution Reply over spoke-spoke tunnel
Phase 3 – NHRP Redirects

Data packet
NHRP Redirect
NHRP Resolution

NHRP mapping
CEF FIB Table
CEF Adjacency

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

192.168.0.1/24

192.168.0.0/24 → Conn.
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

192.168.0.1/24

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

Spoke A

192.168.1.1/24

10.0.0.11 → 172.16.1.1
192.168.1.0/24 → 10.0.0.1

192.168.1.0/24 → 10.0.0.1
192.168.0.0/16 → 10.0.0.1

Spoke B

192.168.2.1/24

10.0.0.1 → 172.17.0.1

10.0.0.1 → 172.16.2.1

10.0.0.1 → 172.17.0.1

192.168.2.0/24 → Conn.
192.168.0.0/16 → 10.0.0.1

10.0.0.1 → 172.17.0.1

10.0.0.1 → 172.17.0.1

CEF Adjacency
Phase 3 – NHRP Redirect Processing

• Sender
  • Insert (GRE IP header source, packet destination IP address) in NHRP redirect table – used to rate-limit NHRP redirect messages ‘show ip nhrp redirect’
  • Send NHRP redirect to GRE/IP header source (previous tunnel hop)
  • Time out rate-limit entries from the NHRP redirect table

• Receiver
  • Check data IP source address from data IP header in redirect
  • If routing to the IP source is out:
    • A GRE tunnel interface with the same NHRP Network-id
      • then drop redirect
    • Another interface, ‘ip nhrp shortcut’ is configured on inbound tunnel and the IP destination is permitted by ‘ip nhrp interest ACL’ (if configured)
      • then trigger an NHRP resolution request to data IP destination from data IP header in redirect
      • otherwise drop redirect
Phase 3 – NHRP Resolution Request

- **Data packet**
- **NHRP Redirect**
- **NHRP Resolution**

**NHRP mapping**

- **CEF FIB Table**
- **CEF Adjacency**

**Spoke A**
- Physical: 172.17.0.1
- Tunnel0: 10.0.0.1

**Spoke B**
- Physical: (dynamic)
- Tunnel0: 10.0.0.11

**CEF Adjacency**

- 10.0.0.1 → 172.17.0.1
- 10.0.0.11 → 172.16.1.1
- 10.0.0.12 → 172.16.2.1
- 192.168.0.0/24 → Conn.
- 192.168.1.0/24 → 10.0.0.11
- 192.168.2.0/24 → 10.0.0.12

- 10.0.0.1 → 172.17.0.1
- 10.0.0.11 → 172.16.1.1
- 10.0.0.12 → 172.16.2.1
- 192.168.0.0/24 → Conn.
- 192.168.1.0/24 → 10.0.0.11
- 192.168.2.0/24 → 10.0.0.12

**CEF FIB Table**

- 172.16.1.1
- 172.16.2.1

- NHRP mapping
- 192.168.1.0/24 → Conn.
- 10.0.0.1 → 172.17.0.1
- 192.168.1.0/24 → 10.0.0.1

- 10.0.0.1 → 172.17.0.1
Phase 3 – NHRP Resolution Processing

• Spoke (NHC) routing table has Hub (NHS) as IP next-hop for networks behind remote Spoke
  • If routing table has IP next-hop of remote spoke then process as in Phase 2
• Data packets are forwarded (CEF-switched) via routed path
  • Redirect message sent by every tunnel hop on routed path
  • Redirect for data packet triggers resolution request only on source spoke
• Send resolution request for IP destination from data packet header in redirect
• Resolution requests forwarded via routed path
• Resolution replies forwarded over direct tunnel
  • Direct tunnel initiated from remote → local spoke
• Forward data packets over direct tunnel after receipt of resolution reply.
Phase 3 – NHRP Resolution Reply

- **NHRP mapping**
- **CEF FIB Table**
- **CEF Adjacency**

**Spoke A**
- Physical: 172.17.0.1
- Tunnel0: 10.0.0.1

**Spoke B**
- Physical: (dynamic)
- Tunnel0: 10.0.0.11
- Tunnel0: 10.0.0.12

---

**CEF FIB Table**
- 172.16.1.1 ➔ 172.16.2.1
- 172.16.2.1 ➔ 172.16.1.1

**CEF Adjacency**
- 10.0.0.1 ➔ 172.17.0.1
- 10.0.0.12 ➔ 172.16.2.1
- 10.0.0.11 ➔ 172.16.1.1
- 10.0.0.12 ➔ 172.16.2.1

---

**Data packet**
- NHRP Redirect
- NHRP Resolution

---

**NHRP Resolution**
- 192.168.0.0/24 ➔ Conn.
- 192.168.1.0/24 ➔ 10.0.0.11
- 192.168.2.0/24 ➔ 10.0.0.12

---

**NHRP Redirect**
- 192.168.0.0/24 ➔ Conn.
Phase 3 – NHRP and Routing Table

Data Packet Forwarding

- When NHRP resolution is received
  - Insert mapping information in mapping table replacing Incomplete/Temporary mapping
  - Insert NHRP routing entry in Routing Table (RT)
    - NHRP NET/Mask is more specific than RT Net/Mask
      - Add new route owned by NHRP (**Type = H**)
    - NHRP Net/Mask is equal to RT Net/Mask
      - Add Override Alternate Next-hop (**% flag**)
      - Route still owned by original owner
    - NHRP Net/Mask is less specific than RT Net/Mask
      - Reduce NHRP mask to = RT Mask
      - Add Override Alternate Next-hop (**% flag**)
      - Route still owned by original owner
  - Insert connected route for tunnel next-hop of NHRP parent mapping (**nhop flag**)
Phase 3 – NHRP and RT

Routing Table

#show ip route
H  192.168.11.0/24 [250/1] via 10.0.1.11, 00:01:02
D  % 192.168.128.0/24 [90/3200000] via 10.0.2.16, 00:50:56, Tunnel0

#show ip route next-hop-override | section H%
H  192.168.11.0/24 [250/1] via 10.0.1.11, 00:01:02
D  % 192.168.128.0/24 [90/3200000] via 10.0.2.16, 00:50:56, Tunnel0

Routing entry for 192.168.11.0/24
Known via "nhrp", distance 250, metric 1
Last update from 10.0.1.11 00:05:29 ago
Routing Descriptor Blocks:
  * 10.0.1.11, from 10.0.1.11, 00:05:29 ago
    Route metric is 1, traffic share count is 1

Routing entry for 192.168.128.0/24
Known via "eigrp 1", distance 90, metric 3200000, type internal
Redistributing via eigrp 1
Last update from 10.0.2.16 on Tunnel0, 00:43:44 ago
Routing Descriptor Blocks:
  * 10.0.2.16, from 10.0.2.16, 00:43:44 ago, via Tunnel0
    Route metric is 3200000, traffic share count is 1

... [NHO]10.0.0.1, from 10.0.0.1, 00:05:57 ago, via Tunnel0
    Route metric is 1, traffic share count is 1
...
Phase 3 – NHRP and Routing Table

NHRP Parent Route Rules

- Insert NHRP routing entry in Routing Table (RIB)
  - NHRP follows the rules outlined above for inserting RIB routes
  - **BUT**
    - NHRP also makes sure to **not contradict** routing protocol routes
- Check for “parent” route
  - Parent – next route with mask prefix less than or equal to NHRP route
  - If Parent route via:
    - same tunnel interface → add NHRP route
    - another interface → do not add NHRP route
- After adding NHRP route → Watch Parent route
  - If Parent route changed or removed (attach to next parent route)
  - If Parent route now via:
    - same tunnel interface → leave NHRP route
    - another interface → remove NHRP route
- Override with ‘no nhrp route-watch’ – can misroute or black-hole traffic
Phase 3 – Refresh or Remove Dynamic Mappings

• Dynamic NHRP mapping entries have finite lifetime
  • Controlled by ‘ip nhrp holdtime …’ on source of mapping (remote spoke)
  • Two types of mapping entries
    • Master entry – Remote Spoke Tunnel IP address
    • Child entries – Remote Network address(es) behind remote-spoke

• Background process checks mapping entries every 60 seconds
  • Master entry: Timing out* → mark CEF adjacency stale
    • If CEF adjacency is then used
      • Refresh Master entry and for each child entry that is also timing out* → queue for immediate refresh

• Refreshing entries
  • Send another Resolution request and reply
    • Resolution request/reply sent over direct tunnel

• If entry expires it is removed
  • If using IPsec and last entry using this NBMA address
    • Trigger IPsec to remove IPsec and ISAKMP/IKEv2 SAs

* Expire timer < 120 seconds
NHRP Purge Messages

- Used to clear invalid NHRP mapping information from the network
- NHRP “local”, “(no socket)” mapping entries
  - Created when sending an NHRP resolution reply
  - Copy of mapping information sent in reply
  - Entry tied to corresponding entry in routing table
  - Keeps list of nodes where resolution reply was sent – ‘show ip nhrp detail’
- If routing table changes so that local mapping entry is no longer valid
  - Purge message is sent to each NHRP node in list
  - NHRP nodes clear that mapping from their table
  - Purge messages forwarded over direct tunnel if available, otherwise sent via routed path
Recent and New Features
Agenda

• DMVPN Design Overview
  • DMVPN General
  • IWAN Specific

• NHRP Details
  • NHRP Overview
  • NHRP Registrations/Resolutions/Redirects

• Recent and New Features
  • Configuration, Resiliency, Routing and Forwarding, Centralized Control
Configuration Reduction

• Issue
  • CLI commands need to be configured to recommended values because defaults are not very useful.

• Solution
  • Change CLI command defaults to recommended values
  • Set other CLI commands as default so that they don’t have to be configured at all
  • Derive CLI command values from other parts of the configuration so they don’t have to be configured.
Configuration

• New defaults (IOS/XE 16.3)
  • NHRP
    • Spoke: (ip/ipv6)
      • nhrp holdtime 600
      • nhrp shortcut
      • nhrp registration no-unique
    • Hub: (ip/ipv6)
      • nhrp holdtime 600
      • nhrp map multicast dynamic
      • nhrp max-send 10000 every 10
        (15.5(3)[S,M]2)

• Future Defaults & Auto-config.
  • NHRP
    • ip/ipv6 nhrp network-id #
      • 1st: tunnel key #
      • 2nd: Interface tunnel #
  • Tunnel Defaults
    • tunnel vrf <tunnel-source-vrf>
  • Miscellaneous Defaults
    • ip mtu
    • ip tcp adjust-mss
    • bandwidth (inherit)
NHRP Original Configuration

Hub

```
interface Tunnel0
  bandwidth 1000
  ip address 10.0.0.1 255.255.255.0
  no ip redirects
  ip mtu 1400
  ip nhrp authentication test
  ip nhrp map multicast dynamic
  ip nhrp map multicast 172.17.0.1
  ip nhrp map 10.0.0.1 172.17.0.1
  ip nhrp network-id 100000
  ip nhrp holdtime 600
  ip nhrp nhs 10.0.0.1
  ip nhrp redirect
  ip tcp adjust-mss 1360
  delay 1000
  tunnel source Serial2/0
  tunnel mode gre multipoint
  tunnel key 100000
  tunnel vrf Outside
  tunnel protection ipsec profile DMVPN
```
NHRP NHS Configuration Reduction – IOS 12.4(20)

• Main use of NHRP mapping is to create static mapping for NHS.
• Combine associated NHRP mapping and NHS commands into a single line.
• Can still configure separate NHRP mappings for other purposes.
NHRP NHS Configuration Reduction – IOS 12.4(20)

- Main use of NHRP mapping is to create static mapping for NHS.
- Combine associated NHRP mapping and NHS commands into a single line.
- Can still configure separate NHRP mappings for other purposes.

```plaintext
interface Tunnel0
    ...
    ip nhrp map multicast 172.17.0.1
    ip nhrp map 10.0.0.1 172.17.0.1
    ...
    ip nhrp nhs 10.0.0.1 nbma 172.17.0.1 multicast
    ...
```

**Hub**

```plaintext
interface Tunnel0
    ...
    ip nhrp map multicast 172.17.0.1
    ip nhrp map 10.0.0.1 172.17.0.1
    ...
    ip nhrp nhs 10.0.0.1 nbma 172.17.0.1 multicast
    ...
```

**Spoke**

```plaintext
interface Tunnel0
    ...
    ip nhrp map multicast 172.17.0.1
    ip nhrp map 10.0.0.1 172.17.0.1
    ip nhrp map multicast 172.17.0.5
    ip nhrp map 10.0.0.2 172.17.0.5
    ...
    ip nhrp nhs 10.0.0.1 nbma 172.17.0.1 multicast
    ip nhrp nhs 10.0.0.2 nbma 172.17.0.5 multicast
    ...
```
NHRP Configuration New Defaults – IOS/XE 16.3

- Spoke: (ip/ipv6)
  - nhrp holdtime 600
  - nhrp shortcut
  - nhrp registration no-unique
- Hub: (ip/ipv6)
  - nhrp holdtime 600
  - nhrp map multicast dynamic
Tunnel Configuration Automatic Settings – Future

- NHRP network-id
  - Set to tunnel key <value> if configured
  - Otherwise, set to tunnel interface <#>

- Tunnel VRF
  - Set to VRF of tunnel source interface

- MTU
  - Set to 1400 bytes
  - Use tunnel source <interface>
    (IPv4/IPv6) MTU – (100/120) bytes

- MSS
  - Set to (IPv4/IPv6) MTU – (40/60) bytes
NHRP Final Configuration

**Hub**

```
interface Tunnel0
  bandwidth 1000
  ip address 10.0.0.1 255.255.255.0
  ip nhrp authentication test
  ip nhrp nhs 10.0.0.1 nbma 172.17.0.1 multicast
  ip nhrp redirect
  delay 1000
  tunnel source Serial2/0
  tunnel mode gre multipoint
  tunnel key 100000
  tunnel protection ipsec profile DMVPN
```

**Spoke**

```
interface Tunnel0
  bandwidth 1000
  ip address 10.0.0.11 255.255.255.0
  ip nhrp authentication test
  ip nhrp nhs 10.0.0.1 nbma 172.17.0.1 multicast
  ip nhrp nhs 10.0.0.2 nbma 172.17.0.5 multicast
  delay 1000
  tunnel source Serial1/0
  tunnel mode gre multipoint
  tunnel key 100000
  tunnel protection ipsec profile DMVPN
```
Agenda

- DMVPN Design Overview
  - DMVPN General
  - IWAN Specific
- NHRP Details
  - NHRP Overview
  - NHRP Registrations/Resolutions/Redirects
- Recent and New Features
  - Configuration, **Resiliency**, Routing and Forwarding, Centralized Control
Resiliency

• Issues
  • Many backup NHSs configured, but don’t want them all up
  • Quickly failover all spokes to alternate hubs when a hub fails
  • Quickly failover a spoke to alternate hub when spoke-hub tunnel fails

• Solutions
  • Backup and FQDN NHS
  • Fast Hub Failover using BGP (BFD between hubs)
  • BFD over DMVPN (BFD on spoke-hub and spoke-spoke tunnels)
Backup and FQDN NHS – 15.1(2)T

• Issue
  • Backup NHSs only needed when primary NHSs are down
  • Backup NHSs can be over subscribed

• Solution
  • Set NHS ‘max-connections’
    • Can set NHS priority (default=0 (best)) – Can have multiple hubs at the same priority
    • Can group NHSs into clusters (default=0) – Separate max-connection value per cluster
    • Configuration reduction – Single line NHS configuration and FQDN NHS
  • Functionality
    • NHSs are brought up in priority order, until cluster max-connections
    • Down NHS at same priority is probed if not at max-connections
    • Down NHS at a lower priority than an active NHS is probed even when max-connections is reached
    • FQDN resolved when bringing up NHS
Backup and FQDN NHS (cont.)

```
interface Tunnel0
  ip nhrp map 10.0.0.1 172.17.0.1
  ip nhrp map multicast 172.17.0.1
  ip nhrp map 10.0.0.2 172.17.0.5
  ip nhrp map multicast 172.17.0.5
  ip nhrp map 10.0.0.3 172.17.0.9
  ip nhrp map multicast 172.17.0.9
  ip nhrp map 10.0.0.4 172.17.0.13
  ip nhrp map multicast 172.17.0.13
  ip nhrp nhs cluster 0 max-connections 2
  ip nhrp nhs 10.0.0.1
  ip nhrp nhs 10.0.0.2
  ip nhrp nhs 10.0.0.3
  ip nhrp nhs 10.0.0.4
  ip nhrp nhs cluster 0 max-connections 2

#show ip nhrp
  10.0.0.1/32 via 10.0.0.1  Tunnel0: Type: static, Flags: used
     NBMA address: 172.17.0.1
  10.0.0.2/32 via 10.0.0.2  Tunnel0: Type: static, Flags: used
     NBMA address: 172.17.0.5
  10.0.0.3/32 via 10.0.0.3  Tunnel0: Type: static, Flags: used
     NBMA address: 172.17.0.9
  10.0.0.4/32 via 10.0.0.4  Tunnel0: Type: static, Flags: used
     NBMA address: 172.17.0.13

#show ip nhrp nhs
  Legend: E=Expecting replies, R=Responding, W=Waiting
  Tunnel0:
  10.0.0.1  RE priority = 0 cluster = 0
  10.0.0.2  RE priority = 0 cluster = 0
  10.0.0.3  W priority = 0 cluster = 0
  10.0.0.4  W priority = 0 cluster = 0
```

```
interface Tunnel0
  ip nhrp map 10.0.0.1 172.17.0.1
  ip nhrp map multicast 172.17.0.1
  ip nhrp map 10.0.0.2 172.17.0.5
  ip nhrp map multicast 172.17.0.5
  ip nhrp map 10.0.0.3 172.17.0.9
  ip nhrp map multicast 172.17.0.9
  ip nhrp map 10.0.0.4 172.17.0.13
  ip nhrp map multicast 172.17.0.13
  ip nhrp nhs cluster 0 max-connections 2

#show ip nhrp nhs
  Legend: E=Expecting replies, R=Responding, W=Waiting
  Tunnel0:
  10.0.0.1  RE priority = 0 cluster = 0
  10.0.0.2  RE priority = 0 cluster = 0
  10.0.0.3  W priority = 0 cluster = 0
  10.0.0.4  W priority = 0 cluster = 0
```
Fast Hub Failover using BGP

• Normal forwarding
  • Few summary routes advertised to spokes
    • Covering all spoke site networks
    • May have separate summary for Hub site networks
    • Use MED to load balance or prefer one hub over the other
  • Covering all spoke site networks

• Hubs “watch” each other
  • Use BFD on physical link or tunnel link between hubs
  • Special trigger route advertised only between hubs over BFD link
    • Example: 1.0.0.[1,2]/32 on Hub[1,2]

• Hubs “watch” each other (cont.)
  • Track loss of trigger route
    • When lost
      • Install static null0 route with special tag for the summary routes
      • Use BGP route-map to increase the Local-Pref on tagged routes
      • Spokes use Local-Pref over MED
  • Recovery
    • Remove static null0 route with special tag
    • Local-Pref reverts back to normal
    • Spokes go back to using MED
Fast Hub Failover using BGP

BGP Configuration

```
router bgp 1
  bgp listen range 10.0.0.0/24 peer-group spokes
  neighbor spokes peer-group
  neighbor spokes remote-as 1
  neighbor spokes timers 20 60
  neighbor 10.0.0.<x> remote-as 1
  neighbor 10.0.0.<x> timers 20 60
  neighbor 192.168.0.<x> remote-as 1
  neighbor 192.168.0.<x> timers 20 60
  neighbor 192.168.0.<x> fall-over bfd
!
address-family ipv4
  network 1.0.0.<x> mask 255.255.255.255
  network 192.168.0.0
  network 192.168.0.0 mask 255.255.0.0 route-map Local-Pref
  network 192.168.252.0 mask 255.255.252.0 route-map Local-Pref
  aggregate-address 192.168.0.0 255.255.0.0 summary-only suppress-map BGP-LEAK
  neighbor spokes activate
  neighbor spokes route-reflector-client
  neighbor spokes next-hop-self all
  neighbor spokes route-map Block-Special out
  neighbor 10.0.0.<x> activate
  neighbor 10.0.0.<x> next-hop-self all
  neighbor 10.0.0.<x> route-map Add-Metric-Hub in
  neighbor 10.0.0.<x> route-map Block-Special out
  neighbor 192.168.0.<x> activate
  neighbor 192.168.0.<x> next-hop-self all
  neighbor 192.168.0.<x> route-map Add-Metric-Hub in
  distance bgp 20 150 150
exit-address-family
```

Enable BFD for BGP

Trigger Route

Modify Local-Pref when adding routes to BGP

Block Trigger route to Spokes and to other Hub over tunnel
Fast Hub Failover using BGP
Tracking, Route-maps and Routes Configuration

Track trigger route

Turn on BFD on Ethernet (250 ms × 4 = 1 sec)

Enable BFD for EIGRP (used on LAN)

Static routes with tag 200 and tracking object 2

Block Trigger route to Spokes and to other Hub over tunnel

Change Local-Pref when route tag is 200

track timer ip route msec 500
track 1 ip route 1.0.0.<y> 255.255.255.255 reachability
track 2 list boolean and
object 1 not

interface Ethernet0/0
ip address 192.168.0.<x> 255.255.255.0
delay 1000
bfd interval 250 min_rx 250 multiplier 4

router eigrp 1
bfd interface Ethernet0/0
default-metric 1000 1000 255 1 1500
network 192.168.0.0
redistribute bgp 1 route-map BGP-EIGRP

ip route 192.168.0.0 255.255.255.0 Null0 tag 200 track 2
ip route 192.168.252.0 255.255.255.0 Null0 tag 200 track 2
ip route 1.0.0.<x> 255.255.255.255 Null0
ip route 192.168.0.0 255.255.255.0 Null0
ip route 192.168.252.0 255.255.255.0 Null0

access-list 1 permit 1.0.0.0 0.0.0.3
route-map Block-Special deny 10
match ip address 1
route-map Block-Special permit 20
route-map Local-Pref permit 10
match tag 200
set local-preference 200
route-map Local-Pref permit 20

Hub1 ↦ <x>=1,<y>=2
Hub2 ↦ <x>=2,<y>=1
Fast Hub Failover using BGP (normal)

RIB
192.168.253.0/24
192.168.254.0/24
192.168.255.0/24

BGP

192.168.0.0/24

DMVPN
10.0.0.0/24

S  1.0.0.1   Null0
B  1.0.0.2 [150/51200]  192.168.0.2
C  10.0.0.0/24  Tunnel0
B  192.168.0.0/16  Null0
C  192.168.0.0/24  Ethernet0/0
B  192.168.1.0/24 [150/0]  10.0.0.11
B  192.168.2.0/24 [150/12]  10.0.0.12

S  194.168.11.0/24 [150/332800]  10.0.0.11
B  194.168.12.0/24 [150/332800]  10.0.0.12

S  192.168.252.0/22  Null0
D  192.168.253.0/24 [90/537600]  192.168.0.3

Spoke1
192.168.1.0/24

192.168.11.0/24

RS1
192.168.12.0/24

B  192.168.2.0/24 [90/537600]  192.168.2.2
B  192.168.252.0/22 [200/0]  10.0.0.1

Spoke2
192.168.2.0/24

192.168.253.0/24

RS2
192.168.11.0/24

B  194.168.11.0/24 [90/307200]  10.0.0.11
B  194.168.12.0/24 [90/332800]  10.0.0.12

S  192.168.252.0/22  Null0
D  192.168.253.0/24 [90/537600]  192.168.0.3

Hub1
192.168.0.0/24

192.168.253.0/24

Hub2
192.168.0.0/24

192.168.254.0/24

DMVPN
10.0.0.0/24

C  10.0.0.0/24
B  192.168.0.0/16 [200/0]  10.0.0.1
B  192.168.0.0/24 [200/0]  10.0.0.1
C  192.168.1.0/24  Ethernet0/0
D  192.168.11.0/24 [90/307200]  10.0.0.1
D  192.168.12.0/24 [90/332800]  192.168.2.2
B  192.168.252.0/22 [200/0]  10.0.0.1

C  10.0.0.0/24
B  192.168.0.0/16 [200/0]  10.0.0.1
B  192.168.0.0/24 [200/0]  10.0.0.1
C  192.168.1.0/24  Ethernet0/0
D  192.168.11.0/24 [90/307200]  10.0.0.1
D  192.168.12.0/24 [90/332800]  192.168.2.2
B  192.168.252.0/22 [200/0]  10.0.0.1

C  10.0.0.0/24
B  192.168.0.0/16 [200/0]  10.0.0.1
B  192.168.0.0/24 [200/0]  10.0.0.1
C  192.168.1.0/24  Ethernet0/0
D  192.168.11.0/24 [90/307200]  10.0.0.1
D  192.168.12.0/24 [90/332800]  192.168.2.2
B  192.168.252.0/22 [200/0]  10.0.0.1

C  10.0.0.0/24
B  192.168.0.0/16 [200/0]  10.0.0.1
B  192.168.0.0/24 [200/0]  10.0.0.1
C  192.168.1.0/24  Ethernet0/0
D  192.168.11.0/24 [90/307200]  10.0.0.1
D  192.168.12.0/24 [90/332800]  192.168.2.2
B  192.168.252.0/22 [200/0]  10.0.0.1
Fast Hub Failover using BGP
Hub2 Debugs

00:47:08.732: BFD-DEBUG Event: V1 FSM ld:17 handle:1 event:ECHO FAILURE state:UP (0)
00:47:08.732: BFD-DEBUG Event: notify client(BGP) IP:192.168.0.1, ld:17, handle:1, event:DOWN, cp independent failure (0)

00:47:08.744: %BGP-5-NBR_RESET: Neighbor 192.168.0.1 reset (BFD adjacency down)
00:47:08.756: %BGP-5-ADJCHANGE: neighbor 192.168.0.1 Down BFD adjacency down
00:47:08.756: %BGP_SESSION-5-ADJCHANGE: neighbor 192.168.0.1 IPv4 Unicast topology base removed from session BFD adjacency down

00:47:08.756: BFD-DEBUG EVENT: bfd_session_destroyed, proc:BGP, handle:1 act
00:47:08.756: BFD-DEBUG Event: V1 FSM ld:17 handle:1 event:Session delete state:DOWN (0)

00:47:08.756: RT: del 1.0.0.1 via 192.168.0.1, bgp metric [150/51200]
00:47:08.756: RT: delete subnet route to 1.0.0.1/32

00:47:09.104: %TRACK-6-STATE: 1 ip route 1.0.0.1/32 reachability Up -> Down
00:47:09.884: %TRACK-6-STATE: 2 list boolean and Down -> Up

00:47:09.888: RT: updating static 192.168.0.0/16 (0x0) : via 0.0.0.0 Nu0 1048578
00:47:09.888: RT: updating static 192.168.252.0/22 (0x0) : via 0.0.0.0 Nu0 104878
Fast Hub Failover using BGP

Spoke1 Debugs

00:47:10.025: RT: updating bgp 192.168.0.0/16 (0x0) : via 10.0.0.2 1048577
00:47:10.025: RT: closer admin distance for 192.168.0.0, flushing 1 routes
00:47:10.025: RT: add 192.168.0.0/16 via 10.0.0.2, bgp metric [200/51200]

00:47:10.025: RT: updating bgp 192.168.252.0/22 (0x0) : via 10.0.0.2 1048577
00:47:10.025: RT: closer admin distance for 192.168.252.0, flushing 1 routes
00:47:10.025: RT: add 192.168.252.0/22 via 10.0.0.2, bgp metric [200/51200]

00:48:00.725: %BGP-3-NOTIFICATION: sent to neighbor 10.0.0.1 4/0 (hold time expired) 0 bytes
00:48:00.725: %BGP-5-NBR_RESET: Neighbor 10.0.0.1 reset (BGP Notification sent)
00:48:00.725: %BGP-5-ADJCHANGE: neighbor 10.0.0.1 Down BGP Notification sent
00:48:00.725: %BGP_SESSION-5-ADJCHANGE: neighbor 10.0.0.1 IPv4 Unicast topology base removed from session BGP Notification sent

00:48:00.725: RT: updating bgp 192.168.0.0/24 (0x0) : via 10.0.0.2 1048577
00:48:00.725: RT: closer admin distance for 192.168.0.0, flushing 1 routes
00:48:00.725: RT: add 192.168.0.0/24 via 10.0.0.2, bgp metric [200/25600]

Switch routing to Hub2 (~1.5 secs)

BGP Hub1 neighbor down (60 secs)
Fast Hub Failover using BGP (failover) (after 1-2 secs)

```
RIB | 192.168.253.0/24  
    | 192.168.254.0/24  
    | 192.168.255.0/24
BGP

Spoke1
  192.168.1.0/24
  192.168.11.0/24
RS1
  192.168.11.0/24

Hub1
  192.168.0.0/24

DMVPN
  10.0.0.0/24

Hub2
  192.168.0.0/24

Spoke2
  192.168.2.0/24

RS2
  192.168.12.0/24

S  1.0.0.2 ➔ Null0
C  10.0.0.0/24 ➔ Tunnel0
B  192.168.0.0/16 ➔ Null0, Tag 200
C  192.168.0.0/24 ➔ Ethernet0/0
B  192.168.1.0/24 ➔ [150/25600] 10.0.0.11
B  192.168.2.0/24 ➔ [150/25600] 10.0.0.12
D  192.168.252.0/22 ➔ Null0, Tag 200
```

```
S  192.168.252.0/22 ➔ Null0, Tag 200
D  192.168.253.0/24 ➔ [90/537600] 192.168.3
...
```
Fast Hub Failover using BGP (failover)

(S) 1.0.0.2 -> Null0
C 10.0.0.0/24 -> Tunnel0
B 192.168.0.0/16 -> Null0, Tag 200
C 192.168.0.0/24 -> Ethernet0/0
B 192.168.1.0/24 [150/25600] -> 10.0.0.11
B 192.168.2.0/24 [150/25600] -> 10.0.0.12
...
B 194.168.11.0/24 [150/332800] -> 10.0.0.11
B 194.168.12.0/24 [150/332800] -> 10.0.0.12
...
S 192.168.252.0/22 -> Null0, Tag 200
D 192.168.253.0/24 [90/537600] -> 192.168.0.3

DMVPN 10.0.0.0/24

C 10.0.0.0/24 -> Tunnel0
B 192.168.0.0/16 [200/51200] -> 10.0.0.2
B 192.168.0.0/24 [200/25600] -> 10.0.0.2
C 192.168.1.0/24 -> Ethernet0/0
D 192.168.11.0/24 [90/307200] -> 192.168.1.2
B 192.168.252.0/22 [200/51200] -> 10.0.0.2

* > 192.168.0.2 25600 100 0 i
* > 192.168.0.2 51200 200 0 i
* > 192.168.0.2 51200 200 0 i

C 10.0.0.0/24 -> Tunnel0
B 192.168.0.0/16 [200/51200] -> 10.0.0.2
B 192.168.0.0/24 [200/25600] -> 10.0.0.2
C 192.168.1.0/24 -> Ethernet0/0
D 192.168.11.0/24 [90/307200] -> 192.168.1.2
B 192.168.252.0/22 [200/51200] -> 10.0.0.2

* > 10.0.0.2 25600 100 0 i
* > 10.0.0.2 51200 200 0 i
* > 10.0.0.2 51200 200 0 i
BFD over DMVPN

- BFD configured on mGRE tunnel interface
  - Use Echo mode
  - BFD maximum probe interval increased to 10 seconds (9999 msec)
  - Spoke-hub tunnel → Only Spoke sends/receives BFD probes*
  - Spoke-spoke tunnel → Both spokes send/receive BFD probes

- NHRP is a BFD client
  - BFD notifies NHRP when tunnel endpoint is down

- NHRP provides a registry for other applications (RP, PfR, IPsec, ...)
  - Applications register with NHRP for a tunnel endpoint (peer, neighbor) address
  - NHRP notifies application when tunnel endpoint is down

```bash
bfd-template single-hop DMVPN
  interval min-tx 2000 min-rx 2000 multiplier 3
  echo

interface Tunnel0
  bfd template DMVPN

Echo mode BFD
  2/6 second keepalive/hold

Apply on Tunnel interface
```

* Currently both Hub and Spoke will send/receive separate BFD probe sets
BFD over DMVPN

BFD session for NHRP static peer (hub)

• If the BFD session is reporting the static peer as down NHRP will:
  • Notify upper layer applications (RP, PFR, …).
  • Initiate NHRP registration requests, if the peer is an NHS
    • If NHRP registration reply is received, peer is up
      • BFD should reflect this state
      • Upper layers should have reset and re-attached to the peer
      • No change in lower layer (IKE/IPsec) crypto session stayed up
    • If NHRP registration reply is not received after 3 retransmissions (~15 seconds)
      • Notify Lower layers (IKE/IPsec) to tear down the crypto session
      • NHRP continues to send registration requests – trigger (IKE/IPsec) crypto session back up
  • Eventually an NHRP registration reply is received
    • The upper layer application sessions (RP and PFR) come back up
  • Note, BFD session is not cleared
BFD over DMVPN

BFD session for NHRP dynamic peer (spoke)

- If the BFD session is reporting the dynamic peer as down NHRP will:
  - Notify upper layer applications (RP, PFR, …)
  - Notify lower layer applications (IPsec) to clear the crypto session
  - Clear the BFD session, NHRP mapping and associated RIB routes
  - Routing will revert back to spoke-hub-spoke
  - A new spoke-spoke tunnel will be attempted if there is more data traffic

- Detect when spoke-spoke tunnel is no longer used for data packets
  - Use packet count estimates to detect when only BFD probes are using tunnel
  - NHRP mapping times out normally
    - Clear NHRP mapping, BFD session, RIB routes and IKE/IPsec session
BFD over DMVPN

Spoke-Hub tunnel

18:13:56.096: BFD-DEBUG Event: V1 FSM Id:1 handle:2 event:DELETE TIMER EXPIRED state:UP (0)
18:13:56.096: BFD-DEBUG Event: notify client(NHRP) IP:10.0.0.1, Id:1, handle:2, event:DOWN, (0)
18:13:56.096: BFD-DEBUG Event: notify client(EIGRP) IP:10.0.0.1, Id:1, handle:2, event:DOWN, (0)

18:13:56.097: %DUAL-5-NBRCHANGE: EIGRP-IPv4 1: Neighbor 10.0.0.1 (Tunnel0) is down: BFD peer down notified
18:13:56.097: RT: delete route to 192.168.0.0 via 10.0.0.1, eigrp metric [90/15360000]
18:13:56.097: RT: add 192.168.0.0/16 via 10.0.0.2, eigrp metric [90/15360015]

18:13:59.059: NHRP: Setting retrans delay to 2 for nhs dst 10.0.0.1
18:13:59.060: NHRP: Send Registration Request via Tunnel0 vrf global(0x0), packet size: 104 src: 10.0.0.11, dst: 10.0.0.1

18:14:02.771: NHRP: Setting retrans delay to 8 for nhs dst 10.0.0.1
18:14:02.771: NHRP: Send Registration Request via Tunnel0 vrf global(0x0), packet size: 104 src: 10.0.0.11, dst: 10.0.0.1

18:14:10.092: NHRP: Setting cache expiry for 172.17.0.1 to 1 milliseconds in cache
18:14:10.092: NHRP: Setting retrans delay to 16 for nhs dst 10.0.0.1

18:14:10.103: IKEv2:(SESSION ID = 1,SA ID = 2):Sending DELETE INFO message for IPsec SA [SPI: 0xAC54C857]
18:14:10.103: IKEv2:(SESSION ID = 1,SA ID = 2):Sending Packet [To 172.17.0.1:500/From 172.16.1.1:500/VRF i0:f0]

18:14:10.105: IKEv2:(SESSION ID = 1,SA ID = 1):Generating IKE_SA_INIT message
18:14:10.105: IKEv2:(SESSION ID = 1,SA ID = 1):Sending Packet [To 172.17.0.1:500/From 172.16.1.1:500/VRF i0:f0]

Switch routing to Hub2
Trigger NHRP Registrations
Reset Crypto
BFD over DMVPN

Spoke-Spoke tunnel

18:46:52.695: NHRP: Receive Traffic Indication via Tunnel0 vrf global(0x0), packet size: 96
18:46:52.784: NHRP: Receive Resolution Request via Tunnel0 vrf global(0x0), packet size: 104
18:46:52.839: %BFD-6-BFD_SESS_CREATED: bfd_session_created, neigh 10.0.0.12 proc:NHRP, idb:Tunnel0 handle:7 act
18:46:52.839: NHRP: Send Resolution Reply via Tunnel0 vrf global(0x0), packet size: 132
18:46:52.875: %BFD-6-BFD_SESS_UP: BFD session ld:2 handle:7 is going UP
18:46:52.875: NHRP: Receive Resolution Reply via Tunnel0 vrf global(0x0), packet size: 132

18:56:52.875: %BFD-6-BFD_SESS_DESTROYED: bfd_session_destroyed, ld:2 neigh proc:NHRP, handle:7 act

19:19:04.622: NHRP: Receive Traffic Indication via Tunnel0 vrf global(0x0), packet size: 96
19:19:04.703: NHRP: Receive Resolution Request via Tunnel0 vrf global(0x0), packet size: 104
19:19:04.734: %BFD-6-BFD_SESS_CREATED: bfd_session_created, neigh 10.0.0.12 proc:NHRP, idb:Tunnel0 handle:7 act
19:19:04.734: NHRP: Send Resolution Reply via Tunnel0 vrf global(0x0), packet size: 132
19:19:04.771: NHRP: Receive Resolution Reply via Tunnel0 vrf global(0x0), packet size: 132
19:19:04.782: %BFDFSM-6-BFD_SESS_UP: BFD session ld:10 handle:7 is going UP

19:19:24.209: %BFDFSM-6-BFD_SESS_DOWN: BFD session ld:10 handle:7 is going Down Reason: DETECT TIMER EXPIRED
19:19:24.211: NHRP: Calling for delete of Tunnel Endpoints (VPN: 10.0.0.12, NBMA: 172.16.2.1)

19:19:24.800: NHRP: Receive Traffic Indication via Tunnel0 vrf global(0x0), packet size: 96

Normal tunnel down (no data traffic) (10 min)
Abnormal tunnel down (BFD triggered) (20 sec)
Agenda

- DMVPN Design Overview
  - DMVPN General
  - IWAN Specific
- NHRP Details
  - NHRP Overview
  - NHRP Registrations/Resolutions/Redirects
- Recent and New Features
  - Configuration, Resiliency, Routing and Forwarding, Centralized Control
NHRP routes and routing

• Issues
  • Can’t control NHRP short-cut routes on spokes
    • Can’t prefer/order routes using multiple short-cut tunnels
    • Can’t summarize NHRP short-cut routes like you can with RP routes
  • Routing protocol limits scale of DMVPN on hubs (IoT)
  • Need separate DMVPN hub router per Cloud (Transport)

• Solutions
  • NHRP Route Metric control
  • NHRP Route Summarization
  • NHRP Route Advertisement
  • Multiple Tunnel Termination (MTT) on Hub routers
Controlling NHRP routes

NHRP route metric control per tunnel interface

- Egress Load-balancing or Ingress traffic engineering
- Peer NHS path preference used to calculate NHRP route metric
- Preference value is (1-255); best = 1; default = 255
- NHRP route metric = \( \frac{255^2}{\text{preference}} \)
  - Examples: (preference = 16 \( \rightarrow \) metric = 4064); (preference = 255 \( \rightarrow \) metric = 255)
- Strict preference: \( (p_1 > 16 \times p_2) \)
  - Example: \( (p_1=32, p_2=1) \rightarrow \frac{2032}{65025} = \frac{1}{32} \)
- Unequal load-balancing: \( (p_1 \leq 16 \times p_2) \)
  - Example: \( (p_1=16, p_2=4) \rightarrow \frac{4064}{16256} = \frac{1}{4} \)

```
# show ip route nhrp
...
192.168.11.0/27 is subnetted, 1 subnets
H 192.168.11.32 [250/4064] via 10.0.0.11, 00:00:09, Tunnel0
```
```bash
#show ip nhrp
...
192.168.11.0/24 via 10.0.0.11
   Tunnel0 created 00:01:46, expire 00:08:13
   Type: dynamic, Flags: router rib
   NBMA address: 172.16.1.1
```
Controlling NHRP routes

NHRP summarization

• Current Behavior
  • NHRP answers resolution request with most specific RIB network/mask
  • Ability to summarize NHRP mappings and routes like RP routes
    • `ip nhrp summary-map { network/mask-length | network mask }
    • Used in resolution responses instead of matching RIB network/mask*
    • Similar to a summary route for a Routing Protocol

• Use Cases
  • Summary of spoke subnets for NHRP resolution replies
  • Fixes 1\textsuperscript{st} subnet of summary route use at spoke, spoke-spoke refresh issue
  • Default (0/0) \(\rightarrow\) NHRP /32 resolution replies mitigation rather than static routes

```
interface Tunnel0
  ip nhrp summary-map 192.168.11.0/24
```

```
#show ip nhrp 192.168.11.0
192.168.11.0/24 via 10.0.0.11
  Tunnel0 created 00:00:07, never expire
  Type: static, Flags: local
  NBMA address: 172.16.1.1 (no-socket)
```
NHRP Summary Map

Without Summary-Map

**NHRP mapping**

**RIB Table**

10.0.0.1/32 -> 172.17.0.1

C 10.0.0.0/24 -> Tunnel0
B 192.168.0.0/16 -> Null0
C 192.168.2.0/24 -> Ethernet0/0
C 192.168.0.0/24 -> Ethernet0/0
B 192.168.1.0/24 -> 10.0.0.11
B 192.168.2.0/24 -> 10.0.0.12
B 192.168.11.0/24 -> 10.0.0.11
B 192.168.12.0/24 -> 10.0.0.12

10.0.0.1/32 -> 172.17.0.1

C 192.168.0.0/24 -> Tunnel0
B 192.168.0.0/16 -> 10.0.0.1
C 192.168.1.0/24 -> Ethernet0/0
B 192.168.11.0/24 -> Null0
D 192.168.11.0/27 -> 192.168.1.2
D 192.168.11.32/27 -> 192.168.1.2
D 192.168.11.64/27 -> 192.168.1.2
D 192.168.11.96/27 -> 192.168.1.2
D 192.168.11.128/27 -> 192.168.1.2
D 192.168.11.192/27 -> 192.168.1.2
D 192.168.11.224/27 -> 192.168.1.2

Spoke A

Physical: 172.168.1.1
Tunnel0: 10.0.0.1

Physical: 172.17.0.1

Spoke B

Physical: 172.168.2.1
Tunnel0: 10.0.0.12

Physical: 172.168.1.1

Tunnel0: 10.0.0.1

192.168.0.1/24

192.168.2.0/24

192.168.1.1/24

192.168.2.1/24

192.168.12.0/24
NHRP Summary Map

Without Summary-Map

NHRP mapping

RIB Table

10.0.0.1/32  →  172.17.0.1
192.168.11.0/27  →  172.16.1.1 (I)

A 10.0.0.0/24  →  Tunnel0
B 192.168.0.0/16  →  Null0
C 192.168.2.0/24  →  Ethernet0/0
C 192.168.0.0/24  →  Ethernet0/0
B 192.168.1.0/24  →  10.0.0.11
B 192.168.2.0/24  →  10.0.0.12
B 192.168.11.0/24  →  10.0.0.11
B 192.168.12.0/24  →  10.0.0.12

Physical: 172.17.0.1
Tunnel0:  10.0.0.11

Physical: 172.16.1.1
Tunnel0:  10.0.0.1

Physical: 172.16.2.1
Tunnel0:  10.0.0.12

Spoke A

192.168.1.1/24
192.168.0.1/24

Spoke B

192.168.11.0/24
192.168.12.0/24

10.0.0.0/24  →  Tunnel0
10.0.0.11/32  →  172.16.1.1

10.0.0.0/24  →  10.0.0.1
192.168.0.0/16  →  10.0.0.1

192.168.1.0/24  →  Ethernet0/0
C 192.168.1.0/24  →  Ethernet0/0

B 192.168.11.0/24  →  10.0.0.11
B 192.168.12.0/24  →  10.0.0.12

192.168.11.0/27  →  172.16.1.1 (I)
192.168.11.32/27  →  192.168.1.2
192.168.11.64/27  →  192.168.1.2
192.168.11.96/27  →  192.168.1.2
192.168.11.128/27 →  192.168.1.2
192.168.11.160/27 →  192.168.1.2
192.168.11.192/27 →  192.168.1.2
192.168.11.224/27 →  192.168.1.2

192.168.12.0/24  →  192.168.2.2

10.0.0.0/24  →  Tunnel0
10.0.0.11/32  →  172.16.1.1
192.168.11.0/27  →  172.16.1.1
NHRP Summary Map

Without Summary-Map

NHRP mapping

RIB Table

10.0.0.1/32 → 172.17.0.1
192.168.11.0/27 → 172.16.1.1 (I)
192.168.11.32/27 → 172.16.1.1 (I)
192.168.11.64/27 → 172.16.1.1 (I)
192.168.11.96/27 → 172.16.1.1 (I)
192.168.11.128/27 → 172.16.1.1 (I)
192.168.11.192/27 → 172.16.1.1 (I)
192.168.11.224/27 → 172.16.1.1 (I)

C 10.0.0.0/24 → Tunnel0
B 192.168.0.0/16 → Null0
C 192.168.2.0/24 → Ethernet0/0
C 192.168.0.0/24 → Ethernet0/0
B 192.168.1.0/24 → 10.0.0.11
B 192.168.2.0/24 → 10.0.0.12
B 192.168.11.0/24 → 10.0.0.11
B 192.168.12.0/24 → 10.0.0.12

10.0.0.11/32 → 172.16.1.1
10.0.0.12/32 → 172.16.2.2

Physical: 172.16.1.1
Tunnel0: 10.0.0.1

Physical: 172.16.2.1
Tunnel0: 10.0.0.12

C 10.0.0.0/24 → Tunnel0
B 192.168.0.0/16 → 10.0.0.1
C 192.168.2.0/24 → Ethernet0/0
H 192.168.11.0/27 → 10.0.0.11
H 192.168.11.32/27 → 10.0.0.11
H 192.168.11.64/27 → 10.0.0.11
H 192.168.11.96/27 → 10.0.0.11
H 192.168.11.128/27 → 10.0.0.11
H 192.168.11.160/27 → 10.0.0.11
H 192.168.11.192/27 → 10.0.0.11
H 192.168.11.224/27 → 10.0.0.11
D 192.168.12.0/24 → 192.168.2.2

Physical: 172.17.0.1

Spoke A

192.168.1.1/24
192.168.0.1/24

Spoke B

192.168.1.1/24
192.168.2.1/24

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved.  Cisco Public
NHRP Summary Map

Without Summary-Map

NHRP mapping

RIB Table

10.0.0.0/24 → Tunnel0
B 192.168.0.0/16 → Null0
C 192.168.2.0/24 → Ethernet0/0
C 192.168.0.0/24 → Ethernet0/0
B 192.168.1.0/24 → 10.0.0.11
B 192.168.2.0/24 → 10.0.0.12
B 192.168.11.0/24 → 10.0.0.11
B 192.168.12.0/24 → 10.0.0.12

192.168.0.0/16 → 172.17.0.1
192.168.1.0/24 → Ethernet0/0
B 192.168.11.0/24 → Null0

D 192.168.11.0/27 → 192.168.1.2
D 192.168.11.32/27 → 192.168.1.2
D 192.168.11.64/27 → 192.168.1.2
D 192.168.11.96/27 → 192.168.1.2
D 192.168.11.128/27 → 192.168.1.2
D 192.168.11.160/27 → 192.168.1.2
D 192.168.11.192/27 → 192.168.1.2
D 192.168.11.224/27 → 192.168.1.2

192.168.12.0/24 → 192.168.2.2
192.168.11.0/27 → 172.16.1.1 (l)
192.168.11.32/27 → 172.16.1.1 (l)
192.168.11.64/27 → 172.16.1.1 (l)
192.168.11.96/27 → 172.16.1.1 (l)
192.168.11.128/27 → 172.16.1.1 (l)
192.168.11.160/27 → 172.16.1.1 (l)
192.168.11.192/27 → 172.16.1.1 (l)
192.168.11.224/27 → 172.16.1.1 (l)

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

10.0.0.11/32 → 172.17.0.1
192.168.11.0/27 → 172.16.1.1 (l)
192.168.11.32/27 → 172.16.1.1 (l)
192.168.11.64/27 → 172.16.1.1 (l)
192.168.11.96/27 → 172.16.1.1 (l)
192.168.11.128/27 → 172.16.1.1 (l)
192.168.11.160/27 → 172.16.1.1 (l)
192.168.11.192/27 → 172.16.1.1 (l)
192.168.11.224/27 → 172.16.1.1 (l)

Physical: 172.16.1.1
Tunnel0: 10.0.0.11

192.168.0.0/16 → 10.0.0.1
C 10.0.0.0/24 → Tunnel0
B 192.168.0.0/16 → 10.0.0.1
C 192.168.1.0/24 → Ethernet0/0
B 192.168.11.0/24 → Tunnel0
D 192.168.11.0/27 → 192.168.1.2
D 192.168.11.32/27 → 192.168.1.2
D 192.168.11.64/27 → 192.168.1.2
D 192.168.11.96/27 → 192.168.1.2
D 192.168.11.128/27 → 192.168.1.2
D 192.168.11.160/27 → 192.168.1.2
D 192.168.11.192/27 → 192.168.1.2
D 192.168.11.224/27 → 192.168.1.2

10.0.0.1/32 → 172.17.0.1
10.0.0.11/32 → 172.16.1.1
192.168.11.0/27 → 172.16.1.1
192.168.11.32/27 → 172.16.1.1
192.168.11.64/27 → 172.16.1.1
192.168.11.96/27 → 172.16.1.1
192.168.11.128/27 → 172.16.1.1
192.168.11.160/27 → 172.16.1.1
192.168.11.192/27 → 172.16.1.1
192.168.11.224/27 → 172.16.1.1

Spoke A

Spoke B

192.168.12.0/24 → 192.168.2.2

Physical: 172.16.2.1
Tunnel0: 10.0.0.12

10.0.0.1/32 → 172.17.0.1
10.0.0.11/32 → 172.16.1.1
192.168.11.0/27 → 172.16.1.1
192.168.11.32/27 → 172.16.1.1
192.168.11.64/27 → 172.16.1.1
192.168.11.96/27 → 172.16.1.1
192.168.11.128/27 → 172.16.1.1
192.168.11.160/27 → 172.16.1.1
192.168.11.192/27 → 172.16.1.1
192.168.11.224/27 → 172.16.1.1

192.168.1.2

192.168.1.1/24
NHRP Summary Map

With Summary-Map

NHRP mapping

RIB Table

10.0.0.11/32  ➔  172.16.1.1
10.0.0.12/32  ➔  172.16.2.2

10.0.0.0/24  ➔  Tunnel0
B 192.168.0.0/16  ➔  Null0
C 192.168.2.0/24  ➔  Ethernet0/0
C 192.168.0.0/24  ➔  Ethernet0/0
B 192.168.1.0/24  ➔  10.0.0.11
B 192.168.2.0/24  ➔  10.0.0.12
B 192.168.11.0/24  ➔  10.0.0.11
B 192.168.12.0/24  ➔  10.0.0.12

Spoke A

192.168.11.0/24  ➔  172.16.1.1 (s,l)
D 192.168.11.0/24  ➔  192.168.1.2
D 192.168.11.0/27  ➔  192.168.1.2
D 192.168.11.32/27  ➔  192.168.1.2
D 192.168.11.64/27  ➔  192.168.1.2
D 192.168.11.96/27  ➔  192.168.1.2
D 192.168.11.128/27  ➔  192.168.1.2
D 192.168.11.160/27  ➔  192.168.1.2
D 192.168.11.192/27  ➔  192.168.1.2
D 192.168.11.224/27  ➔  192.168.1.2

Spoke B

10.0.0.0/32  ➔  172.17.0.1
192.168.11.0/24  ➔  172.16.1.1

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

Physical: 172.16.1.1
Tunnel0: 10.0.0.11

Physical: 172.16.2.1
Tunnel0: 10.0.0.12

192.168.16.1/24

Physical: 172.16.2.1
Tunnel0: 10.0.0.12

Spoke B

10.0.0.0/24  ➔  Tunnel0
B 192.168.0.0/16  ➔  10.0.0.1
C 192.168.2.0/24  ➔  Ethernet0/0
C 192.168.0.0/24  ➔  Ethernet0/0
B 192.168.1.0/24  ➔  10.0.0.11
B 192.168.2.0/24  ➔  10.0.0.12
B 192.168.11.0/24  ➔  10.0.0.11
B 192.168.12.0/24  ➔  10.0.0.12

192.168.12.0/24  ➔  192.168.2.2

192.168.11.0/24  ➔  172.16.1.1 (s,l)
NHRP Summary Map

With Summary-Map

NHRP mapping

RIB Table

10.0.0.11/32 → 172.16.1.1
10.0.0.12/32 → 172.16.2.2

10.0.0.1/32 → 172.17.0.1
192.168.11.0/24 → 172.16.1.1 (s,l)

C 10.0.0.0/24 → Tunnel0
B 192.168.0.0/16 → Null0
C 192.168.1.0/24 → Ethernet0/0
B 192.168.1.0/24 → 10.0.0.11
B 192.168.1.10/24 → 10.0.0.11
B 192.168.1.12/24 → 10.0.0.12
B 192.168.1.2/24 → 10.0.0.12
B 192.168.1.32/27 → 192.168.1.2
B 192.168.1.48/27 → 192.168.1.2
B 192.168.1.64/27 → 192.168.1.2
B 192.168.1.96/27 → 192.168.1.2
B 192.168.1.128/27 → 192.168.1.2
B 192.168.1.160/27 → 192.168.1.2
B 192.168.1.192/27 → 192.168.1.2
B 192.168.1.224/27 → 192.168.1.2

Physical: 172.16.1.1
Tunnel0: 10.0.0.1

Physical: 172.16.2.1
Tunnel0: 10.0.0.12

192.168.1.1/24

Spoke A

Spoke B

192.168.0.1/24

192.168.0.0/16 → 10.0.0.1
C 192.168.2.0/24 → Ethernet0/0
C 192.168.0.0/24 → Ethernet0/0
B 192.168.1.0/24 → 10.0.0.11
B 192.168.2.0/24 → 10.0.0.12
B 192.168.1.10/24 → 10.0.0.11
B 192.168.12.0/24 → 10.0.0.12

192.168.11.0/24

10.0.0.0/24 → Tunnel0
B 192.168.0.0/16 → 10.0.0.1
C 192.168.0.0/24 → Ethernet0/0
H 192.168.11.0/24 → 10.0.0.11
D 192.168.12.0/24 → 192.168.2.2

192.168.11.0/24

10.0.0.1/32 → 172.17.0.1
NHRP Summary Map

With Summary-Map

192.168.0.0/16 → 10.0.0.1
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12
192.168.11.0/24 → 10.0.0.11
192.168.12.0/24 → 10.0.0.12
192.168.0.0/16 → Tunnel0
192.168.1.0/24 → Null0
192.168.2.0/24 → Ethernet0/0
192.168.11.0/24 → Ethernet0/0
192.168.12.0/24 → 10.0.0.21
192.168.1.0/24 → 172.16.1.1
192.168.0.0/24 → 172.16.1.1
192.168.11.0/24 → 172.16.1.1
192.168.12.0/24 → 172.16.1.1

Spoke A

Physical: 172.168.1.1
Tunnel0: 10.0.0.11

Spoke B

Physical: 172.168.2.1
Tunnel0: 10.0.0.12

192.168.11.254/27 → 192.168.1.2
192.168.11.32/27 → 192.168.1.2
192.168.11.64/27 → 192.168.1.2
192.168.11.96/27 → 192.168.1.2
192.168.11.128/27 → 192.168.1.2
192.168.11.160/27 → 192.168.1.2
192.168.11.192/27 → 192.168.1.2
192.168.11.224/27 → 192.168.1.2

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
NHRP Route Advertisement

• Route advertisement between hub and spoke in NHRP registration message
• Ability to redistribute routes between NHRP and other routing protocols
  • redistribute nhrp ...
• Control NHRP routing using standard ‘router nhrp ...’ CLI construct
• Not a replacement for regular routing protocols (EIGRP, BGP, ...)
  • RPs handle much more complex networks
• For simple hub-spoke and spoke-spoke DMVPNs (IoT)
  • 10,000s small spoke sites with one or few subnets
  • 100,000s of spokes sites in hub-spoke IoT networks
  • Preliminary scaling to 10,000 → 30,000+ spokes per hub (CSR)
Multiple Tunnel Termination (MTT)

- Issue
  - Multiple DMVPN clouds (IWAN Transports) terminating on the same Hub
  - Spoke-spoke tunnels don’t always get built
    - Data packets CEF switched between DMVPNs
      → No NHRP Redirect sent
    - NHRP Resolution (NHRP) switched between DMVPNs
      → Hub answers NHRP resolution
  - Spoke-spoke traffic continues to traverse the hub

- Solution
  - Forward NHRP and Data traffic out the same DMVPN on which it arrived
    - Part 1: NHRP traffic; controlled by NHRP control plane
    - Part 2: Data plane traffic forwarding; controlled by FIB/CEF (future)
DMVPN without MTT
Routing preferred via MPLS

192.168.1.0/24 (10) → 10.0.0.11
192.168.2.0/24 (10) → 10.0.0.12

10.0.0.0/24
10.0.1.0/24

DMVPN Tunnel0
DMVPN Tunnel1

CEF
NHRP

Global
INET
MPLS
DMVPN without MTT

Routing preferred via MPLS

- CEF MPLS → MPLS

192.168.1.0/24 (10) → 10.0.0.11
192.168.2.0/24 (10) → 10.0.0.12
DMVPN without MTT

Routing preferred via MPLS

- CEF MPLS $\rightarrow$ MPLS
  - Send NHRP Redirect $\rightarrow$ MPLS

192.168.1.0/24 (10) $\rightarrow$ 10.0.0.11
192.168.2.0/24 (10) $\rightarrow$ 10.0.0.12
DMVPN without MTT

Routing preferred via MPLS

- CEF MPLS → MPLS
  - Send NHRP Redirect → MPLS
  - Forward NHRP Resolution Request MPLS → MPLS
- S
  poke-spoke
DMVPN without MTT

Routing preferred via MPLS

- CEF MPLS → MPLS
  - Send NHRP Redirect → MPLS
  - Forward NHRP Resolution Request MPLS → MPLS
    - Spoke-spoke

- CEF INET → MPLS
  - Don’t send NHRP Redirect
    - No spoke-spoke
DMVPN without MTT

ECMP routing via MPLS and INET

- CEF inbound MPLS
  - Forward: 50% → MPLS
    - Send NHRP Redirect → MPLS
    - NHRP Resolution Request
      - 50% Forward → MPLS (spoke-spoke)
      - 50% Hub answers (no spoke-spoke)
DMVPN without MTT

ECMP routing via MPLS and INET

- CEF inbound MPLS
  - Forward: 50% → MPLS
    - Send NHRP Redirect → MPLS
    - NHRP Resolution Request
      - 50% Forward → MPLS (spoke-spoke)
      - 50% Hub answers (no spoke-spoke)
  - Forward: 50% → INET
    - Don’t send NHRP Redirect
    - No spoke-spoke
- CEF inbound INET (similar)
DMVPN with MTT (Part 1)

ECMP routing via MPLS and INET

• CEF inbound MPLS
  • Forward: 50% → MPLS
  • Send NHRP Redirect → MPLS
DMVPN with MTT (Part 1)

ECMP routing via MPLS and INET

192.168.1.0/24 (10) → 10.0.0.11
(10) → 10.0.1.11
192.168.2.0/24 (10) → 10.0.0.12
(10) → 10.0.1.12

- CEF inbound MPLS
  - Forward: 50% → MPLS
    - Send NHRP Redirect → MPLS
    - NHRP Resolution Request
      - 100% Forward → MPLS (spoke-spoke)
DMVPN with MTT (Part 1)

ECMP routing via MPLS and INET

- CEF inbound MPLS
  - Forward: 50% → MPLS
    - Send NHRP Redirect → MPLS
    - NHRP Resolution Request
      - 100% Forward → MPLS (spoke-spoke)
  - Forward: 50% → INET
    - Don’t send NHRP Redirect
    - No spoke-spoke
- CEF inbound INET (similar)
DMVPN with MTT (Part 2)

Routing via MPLS and INET (ECMP or Un-equal cost)

- CEF inbound MPLS
  - Forward: 100% → MPLS
    - Send NHRP Redirect → MPLS
    - NHRP Resolution Request
      - 100% Forward → MPLS (spoke-spoke)

- CEF inbound INET (similar)
DMVPN with Multiple Tunnel Termination (MTT)

• Part 1
  • With DMVPN – ECMP RIB routes over both tunnels
    • Statistical per flow (src-IP, dst-IP) whether spoke-spoke tunnel is triggered
    • All flows use spoke-spoke tunnel when built
  • With IWAN – Preferred route over one tunnel (all routes in RIB)
    • Data traffic over primary (preferred) route – spoke-spoke triggered by:
      • Data packets for primary tunnel
      • PfR probes for secondary tunnel
    • Data traffic over secondary (non-preferred) route – spoke-spoke triggered by:
      • PfR probes for both tunnels; up to 10 second delay for secondary tunnel

• Part 2 (future)
  • Preferred route over one tunnel (all routes in RIB)
  • Data traffic or PfR Probes trigger spoke-spoke tunnels over both tunnels
Agenda

- DMVPN Design Overview
  - DMVPN General
  - IWAN Specific
- NHRP Details
  - NHRP Overview
  - NHRP Registrations/Resolutions/Redirects
- Recent and New Features
  - Configuration, Resiliency, Routing and Forwarding, Centralized Control
Centralized Control

Separating Control and Data Planes

• Issues – Converged control and data planes on hubs
  • Routing Protocol scaling limits scale of DMVPN hubs
  • ISP managed DMVPN – want Hubs in ISP network
    • Data traffic traverses DMVPN hubs while short-cut tunnel is built
    • Multicast traffic (replication*) goes through DMVPN hubs

• Solution – Separate control plane from data plane
  • Separate DMVPN Control Plane Hub (CPH) and Data Plane Hub (DPH)
    • Scale CPH and DPH independently of each other
    • Data traffic only ever goes through DPHs never CPH
  • Other central control services at CPH
    • Routing Protocol (BGP, EIGRP*); Key Management (ESON*)

• ISP managed DMVPN → CPH in ISP network; DPHs in customer network
Centralized Routing and NHS
Separating the Control and Data planes

- Control Plane Hub (CPH)
  - Routing – peer with DPHs and spokes
    - iBGP (route-reflector), in future EIGRP (OTP route-reflector)
  - NHS → NHRP registrations and resolution request processing
  - Future:
    - Smart data plane hub selection pushed to spokes
    - Optional Centralized Key Server (ESON)

- Data Plane Hub (DPH)
  - Provide data path for spoke-hub-spoke
  - Routing – peer with CPH
    - Advertise network and/or regional summaries to CPH
  - NHS → NHRP redirect; NHRP registrations (for no-drop); Backup CPH
Centralized Routing and NHS

CPH

Phy: 172.31.0.1
Tu0: 10.0.0.254

Spoke A

Phy: 172.16.1.1
Tu0: 10.0.0.11

192.168.1.1/24

192.168.0.1/24

DPH

Phy: 172.17.0.1
Tu0: 10.0.0.1

Spoke B

Phy: 172.16.2.1
Tu0: 10.0.0.12

192.168.2.1/24

= Control Plane tunnels
= Data Plane tunnels
Centralized Routing and NHS

CPH Configuration

interface Tunnel0
  ip address 10.0.0.254 255.255.255.0
  ip nhrp authentication test
  ip nhrp network-id 100000
  tunnel source Serial2/0
  tunnel mode gre multipoint
  tunnel key 100000
  tunnel protection ipsec profile DMVPN
end

router bgp 1
  bgp listen range 10.0.0.0/24 peer-group spokes
  neighbor spokes peer-group
  neighbor spokes remote-as 1
  neighbor spokes timers 20 60
  neighbor 10.0.0.1 remote-as 1
  neighbor 10.0.0.1 timers 20 60
  address-family ipv4
    neighbor spokes route-reflector-client
    neighbor spokes route-map SUMMARY-ONLY out
  exit-address-family

Although DMVPN is the focus, you can use another tunnel protocol, like GRE, which has the flexibility to provide tunneling over a variety of protocols, such as IP, IPv6, and Ethernet. GRE tunnels can also be configured to carry multiple protocols simultaneously, which is useful for environments where different protocols need to be supported over a single tunnel connection.

**SPKSEC**

- **DPH is a regular neighbor**
- **Spokes are dynamic and route-reflector-clients**
- **Only send summary to Spokes**
- **Send everything to DPH**
- **Don't reset next-hop**

The configuration in the image is designed to ensure that the Spokes are properly configured as dynamic and route-reflector-clients. The summary to Spokes is sent without resetting the next-hop, which is essential for maintaining the correct routing paths. The use of route-reflector-clients ensures that the Spokes receive the necessary updates to their routing table, allowing for efficient and accurate data transmission.
Centralized Routing and NHS

DPH Configuration

```
interface Tunnel0
  ip address 10.0.0.1 255.255.255.0
  ip nhrp authentication test
  ip nhrp network-id 100000
  ip nhrp nhs 10.0.0.254 nbma 172.31.0.1 multicast
  no ip nhrp send-routed
  ip nhrp redirect
  tunnel source Serial2/0
  tunnel mode gre multipoint
  tunnel key 100000
  tunnel protection ipsec profile DMVPN
end!

router bgp 1
  bgp log-neighbor-changes
  neighbor 10.0.0.254 remote-as 1
  neighbor 10.0.0.254 timers 20 60!
  address-family ipv4
    bgp redistribute-internal
    network 192.168.0.0
    aggregate-address 192.168.0.0 255.255.0.0 attribute-map SUMMARY-CMNTY
    neighbor 10.0.0.254 activate
    neighbor 10.0.0.254 send-community
    neighbor 10.0.0.254 next-hop-self all
    exit-address-family!

  ip bgp-community new-format route-map SUMMARY-CMNTY permit 10 set community 1:255
```
Centralized Routing and NHS
Spoke Configuration

interface Tunnel0
  ip address 10.0.0.11 255.255.255.0
  ip nhrp authentication test
  ip nhrp network-id 100000
  ip nhrp nhs 10.0.0.254 nbma 172.31.0.1 multicast
  ip nhrp nhs 10.0.0.1 nbma 172.17.0.1 multicast priority 16
  no ip nhrp send-routed
  tunnel source Serial1/0
  tunnel mode gre multipoint
  tunnel key 100000
  tunnel protection ipsec profile DMVPN

! router bgp 1
  bgp log-neighbor-changes
  neighbor 10.0.0.254 remote-as 1
  neighbor 10.0.0.254 timers 20 60

! address-family ipv4
  bgp redistribute-internal
  network 192.168.1.0
  neighbor 10.0.0.254 activate
  neighbor 10.0.0.254 next-hop-self all
  exit-address-family

CPH is main NHS
DPH is secondary NHS
NHRP Res. Req. via CPH

Neighbor only with CPH

Set next-hop to Self
Centralized Routing and NHS
NHRP Registration and Routing

Spoke A: 192.168.1.1/24
  Phy: 172.17.0.1
  Tu0: 10.0.0.1

Spoke B: 192.168.2.1/24
  Phy: 172.16.1.1
  Tu0: 10.0.0.11

10.0.0.254 → 172.31.0.1
10.0.0.1 → 172.17.0.1
10.0.0.1 → 172.17.0.1
10.0.0.254 → 172.31.0.1
192.168.0.0/16 → NULL0
192.168.0.0/24 → Ethernet0/0
192.168.1.0/24 → Ethernet0/0
192.168.2.0/24 → Ethernet0/0

NHRP mapping
Routing Table
Centralized Routing and NHS
NHRP Registration and Routing

10.0.0.1 → 172.17.0.1
192.168.0.0/16 → 10.0.0.1
192.168.0.0/24 → 10.0.0.1

CPH
Phy: 172.17.0.1
Tu0: 10.0.0.1

192.168.0.1/24

Spoke A
Phy: 172.16.1.1
Tu0: 10.0.0.11
192.168.1.1/24

Routing Table

10.0.0.254 → 172.31.0.1
192.168.0.0/16 → Null0
192.168.0.0/24 → Ethernet0/0

Spoke B
Phy: 172.16.2.1
Tu0: 10.0.0.12
192.168.2.1/24

192.168.2.0/24 → Ethernet0/0

10.0.0.1 → 172.17.0.1
10.0.0.254 → 172.31.0.1

192.168.1.0/24 → Ethernet0/0

NHRP mapping

10.0.0.1 → 172.17.0.1
10.0.0.254 → 172.31.0.1

192.168.0.0/24 → 10.0.0.1
192.168.0.0/16 → 10.0.0.1

= Control Plane tunnels
= Data Plane tunnels
Centralized Routing and NHS

NHRP Registration and Routing

**Spoke A**
- Phy: 172.16.1.1
- Tu0: 10.0.0.11
- 192.168.1.1/24

**Spoke B**
- Phy: 172.16.2.1
- Tu0: 10.0.0.12

**Centralized Routing (CPH)**
- Phy: 172.17.0.1
- Tu0: 10.0.0.1

**Centralized Routing (DPH)**
- Phy: 172.31.0.1
- Tu0: 10.0.0.254

**Routing Table**
- 192.168.0.0/16 → Null0
- 192.168.0.0/24 → Ethernet0/0

**NHRP Mapping**
- 10.0.0.254 → 172.31.0.1
- 192.168.0.0/16 → 10.0.0.1
- 192.168.1.0/24 → 10.0.0.11
- 192.168.2.0/24 → 10.0.0.12

**Control Plane Tunnels**
- Yellow

**Data Plane Tunnels**
- Green

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
Centralized Routing and NHS

NHRP Registration and Routing

10.0.0.1 → 172.17.0.1
10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1
192.168.0.0/16 → 10.0.0.1
192.168.0.0/24 → 10.0.0.1
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

192.168.1.1/24
Phy: 172.16.1.1
Tu0: 10.0.0.11

192.168.2.1/24
Phy: 172.16.2.1
Tu0: 10.0.0.12

192.168.0.1/24
Phy: 172.17.0.1
Tu0: 10.0.0.1

192.168.2.0/24
Phy: 172.31.0.1
Tu0: 10.0.0.254

0x0

= Control Plane tunnels
= Data Plane tunnels

192.168.0.0/16 → 10.0.0.1
192.168.1.0/24 → Ethernet0/0

192.168.0.0/16 → Null0
192.168.0.0/24 → Ethernet0/0
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1
10.0.0.254 → 172.31.0.1
Centralized Routing and NHS

Data packets via DPH

10.0.0.1 → 172.17.0.1
10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

192.168.0.0/16 → 10.0.0.1
192.168.0.0/24 → 10.0.0.1
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

10.0.0.1 → 172.17.0.1 (16)
10.0.0.254 → 172.31.0.1
12.168.2.1 → ???
192.168.0.0/16 → 10.0.0.1
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

10.0.0.1 → 172.17.0.1 (16)
10.0.0.254 → 172.31.0.1
192.168.0.0/16 → 10.0.0.1
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1
10.0.0.254 → 172.31.0.1

192.168.0.0/16 → Null0
192.168.1.0/24 → Ethernet0/0
192.168.2.0/24 → 10.0.0.12

Routing Table

= Control Plane tunnels
= Data Plane tunnels

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
Centralized Routing and NHS

Control packets via CPH

- **10.0.0.1** → 172.17.0.1
- **10.0.0.11** → 172.16.1.1
- **10.0.0.12** → 172.16.2.1
- **192.168.0.0/16** → 10.0.0.1
- **192.168.0.0/24** → 10.0.0.1
- **192.168.1.0/24** → 10.0.0.11
- **192.168.2.0/24** → 10.0.0.12

- **10.0.0.254/32** → 172.31.0.1
- **192.168.0.0/16** → 10.0.0.1
- **192.168.0.0/24** → 10.0.0.1
- **192.168.1.0/24** → 10.0.0.11
- **192.168.2.0/24** → 10.0.0.12

- **192.168.0.0/16** → 172.17.0.1 (16)
- **10.0.0.1/32** → 172.17.0.1
- **192.168.0.254/32** → 172.31.0.1
- **192.168.2.1/32** → ?

- **192.168.0.0/16** → 10.0.0.1
- **192.168.1.0/24** → Ethernet0/0
- **192.168.2.0/24** → Ethernet0/0

- **192.168.0.0/24** → 10.0.0.1
- **192.168.0.0/16** → 10.0.0.1
- **192.168.2.0/24** → 10.0.0.1

**NHRP mapping**

**Routing Table**

- **10.0.0.11** → 172.16.1.1
- **10.0.0.12** → 172.16.2.1
- **10.0.0.254** → 172.31.0.1

- **192.168.0.0/16** → Null0
- **192.168.0.0/24** → Ethernet0/0
- **192.168.1.0/24** → 10.0.0.11
- **192.168.2.0/24** → 10.0.0.12

**Spoke A**

- **192.168.1.1/24**
- **Phy:** 172.31.0.1
- **Tu0:** 10.0.0.254

**Spoke B**

- **192.168.2.1/24**
- **Phy:** 172.31.0.1
- **Tu0:** 10.0.0.254

**CPH**

- **192.168.0.1/24**
- **Phy:** 172.17.0.1
- **Tu0:** 10.0.0.1

**DPH**

- **192.168.0.1/24**
- **Phy:** 172.16.2.1
- **Tu0:** 10.0.0.12

**BRKSEC - 3052**

= Control Plane tunnels

= Data Plane tunnels

= Dynamic spoke-spoke tunnel
Centralized Routing and NHS

Control packets via CPH

- **Centralized Routing and NHS**

- **Control packets via CPH**

  - **Spoke A**
    - `192.168.0.1/24`
    - `192.168.2.1/24`
    - `10.0.0.1/32`
    - `10.0.0.254/32`
    - `192.168.1.0/24`
    - `192.168.2.0/24`

  - **Spoke B**
    - `192.168.0.1/24`
    - `192.168.2.0/24`

- **NHRP mapping**

- **Routing Table**

  - **CPH**
    - Phy: `172.31.0.1`
    - Tu0: `10.0.0.254`

  - **DPH**
    - Phy: `172.16.1.1`
    - Tu0: `172.16.1.1`

  - **Spoke A**
    - 10.0.0.1/32
    - 10.0.0.254/32
    - 192.168.0.0/16

  - **Spoke B**
    - 10.0.0.11
    - 10.0.0.12

- **BRKSEC - 3052**

  - Control Plane tunnels
  - Data Plane tunnels
  - Dynamic spoke-spoke tunnel

- **NHRP mapping**

- **Routing Table**
Centralized Routing and NHS
Control packets via CPH

192.168.0.0/16 → 10.0.0.1
192.168.0.0/24 → 10.0.0.1
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

10.0.0.1/32 → 172.17.0.1 (16)
10.0.0.12/32 → 172.16.2.1
10.0.0.254/32 → 172.31.0.1

192.168.0.0.1/32 → 172.16.0.1
192.168.0.254/32 → 172.31.0.1

10.0.0.1 → 172.17.0.1
10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

NHRP mapping
Routing Table

= Control Plane tunnels
= Data Plane tunnels
= Dynamic spoke-spoke tunnel
Centralized Routing and NHS

Control packets via CPH

- 10.0.0.1 → 172.17.0.1
- 10.0.0.11 → 172.16.1.1
- 10.0.0.12 → 172.16.2.1

192.168.0.0/16 → 10.0.0.1
192.168.0.0/24 → 10.0.0.1
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

10.0.0.1/32 → 172.17.0.1 (16)
10.0.0.12/32 → 172.16.2.1
10.0.0.254/32 → 172.31.0.1
192.168.2.0/24 → 172.16.2.1

192.168.2.0/24 → Ethernet0/0
10.0.0.1/32 → 172.17.0.1 (16)
10.0.0.12/32 → 172.16.2.1
192.168.1.1/32 → 172.16.1.1
192.168.2.0/24 → 10.0.0.12

10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1
10.0.0.254 → 172.31.0.1

192.168.0.0/16 → Null0
192.168.0.0/24 → Ethernet0/0
192.168.0.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

192.168.0.1/24 → Ethernet0/0
192.168.2.0/24 → 10.0.0.12

= Control Plane tunnels
= Data Plane tunnels
= Dynamic spoke-spoke tunnel
Centralized Routing and NHS

Control packets via CPH

10.0.0.1 → 172.17.0.1
10.0.0.11 → 172.16.1.1
10.0.0.12 → 172.16.2.1

192.168.0.0/16 → 10.0.0.1
192.168.0.0/24 → 10.0.0.1
192.168.1.0/24 → 10.0.0.11
192.168.2.0/24 → 10.0.0.12

10.0.0.1/32 → 172.17.0.1 (16)
10.0.0.12/32 → 172.16.2.1
10.0.0.254/32 → 172.31.0.1
192.168.2.0/24 → 172.16.2.1

192.168.0.0/16 → 10.0.0.1
192.168.1.0/24 → Ethernet0/0
192.168.2.0/24 → 10.0.0.12

10.0.0.1/32 → 172.17.0.1 (16)
10.0.0.11/32 → 172.16.1.1
10.0.0.254/32 → 172.31.0.1
192.168.1.0/24 → 172.16.1.1

192.168.0.0/16 → 10.0.0.1
192.168.1.0/24 → Ethernet0/0
192.168.2.0/24 → 10.0.0.12

= Control Plane tunnels
= Data Plane tunnels
= Dynamic spoke-spoke tunnel
Centralized Routing and NHS

• Summary
  • Separation of Control and Data Planes
    • ISP Managed DMVPN Service (CPH in ISP network, DPHs in customer network)
    • Separate scaling for CPH (RP peers) and DPH (Encryption throughput)
    • Uses the same DMVPN/mGRE infrastructure
      • Main NHS at CPH, Natural backup NHS at DPH

• Future
  • Download from CPH to spokes, NHS summary-map configuration for DPH
    ```
    {ip | ipv6} nhrp summary-map {all-routed [nbma] | prefix [[nbma [preference pref]]] [multicast] [resolve]
      [match {group group_name | geo-location geo-location | topo-location topo-location | attribute attr_type attr_value}]
    ```
    • All-routed: RP advertises summary → temporary map to use NBMA as DPH
    • Prefix: Default/summary prefix passed to spokes
    • Resolve: Prefix is specified, but not NBMA → forces resolution for all packets; hub-less model
    • Match: Push different summary maps depending on attributes from spoke registration to CPH
Centralized Control
Extensible Security for Overlay Network (ESON)

- A Centralized Key Server Solution with pairwise key capability.
  - Centralized management of policy & pairwise and group keys for IPsec overlay VPNs
  - Leverages GetVPN control plane (GDOI/G-IKEv2) as underlying infrastructure
    - GM-KS: G-IKEv2 Registrations for initial pull of policy & keying material
    - KS-GM: KS pushes periodic rekeys (unicast/multicast)
    - KS-KS: Multiple KSs for redundancy using COOP over IKEv2
- IKEv2 is not used between GMs (no Diffee Hellman (DH))
- Peer Introduction Protocol (PIP) is lightweight control plane (2 messages) between GMs
  - GM-GM: Exchanges cryptographic identities and nonces for pair-wise key generation and detects NAT between Peers
DMVPN with ESON

G-IKEv2 based centralized management of pairwise and group IPsec session keys

**Control Plane**
- **GM – KS**
  - G-IKEv2 (KEK, TEK and key material from KS)
- **GM – GM**
  - PIP* (TEK): (Encrypted with TEK key)

**Control Plane Redundancy**
- **KS – KS**
  - COOP over IKEv2

**Data Plane**
- **GM – GM**
  - IPsec *(GM1-GM2 Pairwise key)*

**Data Plane Redundancy**
- **GM – GM**
  - Redundant Hubs

**Legend**
- **TEK**: Traffic Encryption Key
- **KEK**: Key Encryption Key
DMVPN with ESON

G-IKEv2 based centralized management of pairwise and group IPsec session keys

**Control Plane**
- **GM – KS**
  - G-IKEv2 (KEK, TEK and key material from KS)
- **GM – GM**
  - PIP*(TEK): (Encrypted with TEK key)

**Control Plane Redundancy**
- **KS – KS**
  - COOP over IKEv2

**Data Plane**
- **GM – GM**
  - IPsec *(GM1-GM2 Pairwise key)

**Data Plane Redundancy**
- **GM – GM**
  - Redundant Hubs

- **TEK**: Traffic Encryption Key
- **KEK**: Key Encryption Key

Group Keys: TEK, KEK
- **GM1 Key Material**
- **GM1 Identity**

Group Keys: TEK, KEK
- **GM2 Key Material**
- **GM2 Identity**
DMVPN with ESON

G-IKEv2 based centralized management of pairwise and group IPsec session keys

**Control Plane**
- **GM – KS**
  - G-IKEv2 (KEK, TEK and key material from KS)
- **GM – GM**
  - PIP*(TEK): (Encrypted with TEK key)

**Control Plane Redundancy**
- **KS – KS**
  - COOP over IKEv2

**Data Plane**
- **GM – GM**
  - IPsec *(GM1-GM2 Pairwise key)

**Data Plane Redundancy**
- **GM – GM**
  - Redundant Hubs

---

**TEK:** Traffic Encryption Key
**KEK:** Key Encryption Key
DMVPN with ESON

G-IKEv2 based centralized management of pairwise and group IPsec session keys

Control Plane

GM – KS
- G-IKEv2 (KEK, TEK and key material from KS)

GM – GM
- PIP*(TEK): (Encrypted with TEK key)

Control Plane Redundancy

KS – KS
- COOP over IKEv2

Data Plane

GM – GM
- IPsec *(GM1-GM2 Pairwise key)

Data Plane Redundancy

GM – GM
- Redundant Hubs

TEK: Traffic Encryption Key
KEK: Key Encryption Key

Group Keys: TEK, KEK
GM1 Key Material
GM1 Identity
GM1-GM2 Pairwise key

Group Keys: TEK, KEK
GM2 Key Material
GM2 Identity
GM1-GM2 Pairwise key
DMVPN with ESON

G-IKEv2 based centralized management of pairwise and group IPsec session keys

Control Plane
- **GM – KS**
  - G-IKEv2 (KEK, TEK and key material from KS)
  - **GM – GM**
    - PIP*(TEK): (Encrypted with TEK key)

Control Plane Redundancy
- **KS – KS**
  - COOP over IKEv2

Data Plane
- **GM – GM**
  - IPsec *(GM1-GM2 Pairwise key)

Data Plane Redundancy
- **GM – GM**
  - Redundant Hubs

---

**TEK**: Traffic Encryption Key
**KEK**: Key Encryption Key

**GM1** / DMVPN Hub/Spoke
- *(GM1-GM2 pairwise key)*

**GM2** / DMVPN Hub/Spoke
- *(GM1-GM2 pairwise key)*

**Internet/WAN**

**KS1**(DC)
- **KS1**(DC)

**KS2**(DR)
- **KS2**(DR)

---

Group Keys: TEK, KEK
- GM1 Key Material
- GM1 Identity
- GM1-GM2 Pairwise key

Group Keys: TEK, KEK
- GM2 Key Material
- GM2 Identity
- GM1-GM2 Pairwise key
DMVPN with ESON - Value Proposition

- Centralized key server and management
  - Centralized authentication & authorization of GMs (DMVPN Hub/spoke)
  - Centralized management of crypto policy and keys
  - Crypto Control-plane/Data-plane separation, no IKEv2 or DH between GMs

- Easier to manage

- Elasticity of scale; Reduced setup latency; Virtualized Key Server

- Faster & more effective removal of compromised GMs

- Better enforcement of enterprise security policy & centralized trust management

- Allows varying key management schemes
  - Group keys: Control Plane (PIP); Data Plane (Native Multicast)
  - Pairwise keys for better security – Data Plane (Unicast)
  - Various rekey policies/schemes are possible
DMVPN Other Recent and Future Features

- **Recently Available**
  - Metadata (CMD, NSH) over DMVPN
    - PfR; TrustSec (SGT)

- **Coming Next**
  - IKEv2 cluster load-balancing
  - VXLAN-GPE encapsulation for DMVPN
    - Support for multiple spokes behind NPAT
  - DMVPN extended authentication
    - Strong NHRP authentication using HMAC
    - NHRP spoke authentication using Radius
    - Dynamic Tunnel Key on spokes

- **Future**
  - NHRP route advertisement
  - Limited spoke-spoke multicast support
    - Large spoke to many small spokes
    - PIM based signaling
    - NHRP based signaling
  - Native Multicast over DMVPN
    - Tunnel packets with multicast destination
    - ISP network does replication
    - ESON KS (Group and pair-wise keys)
  - GRE tunnel sub-interfaces
  - EVN WAN using DMVPN
Thank You
Complete Your Online Session Evaluation

• Please complete your Online Session Evaluations after each session

• Complete 4 Session Evaluations & the Overall Conference Evaluation (available from Thursday) to receive your Cisco Live T-shirt

• All surveys can be completed via the Cisco Live Mobile App or the Communication Stations

Don’t forget: Cisco Live sessions will be available for viewing on-demand after the event at CiscoLive.com/Online
Use Cisco Spark to communicate with the Speaker and fellow participants after the session

Download the Cisco Spark app from iTunes or Google Play

1. Go to the Cisco Live Berlin 2017 Mobile app
2. Find this session
3. Click the Spark button under Speakers in the session description
4. Enter the room, room name = BRKSEC-3052
5. Join the conversation!

The Spark Room will be open for 2 weeks after Cisco Live
Continue Your Education

- Demos in the Cisco campus
- Walk-in Self-Paced Labs
  - LABIOT-2012 Implementing Dynamic Multipoint VPN
- Lunch & Learn
- Meet the Engineer 1:1 meetings
- Related sessions
Recommended Reading

Cisco Intelligent
WAN (IWAN)

Brad Edgeworth, CCIE No. 31574
Jean-Marc Barozet
David Prall, CCIE No. 6508
Anthony Lockhart
Thank You
Your Time Is Now
Extras
Tunnel Health Monitoring

Interface State – 15.0(1)M

• Issue
  • mGRE tunnel Interface is always “up”
  • Can’t use standard backup/recovery mechanisms
    • backup interface, static interface routes, …

• Solution
  • New Command ‘if-state nhrp’
  • Monitor NHRP registration replies
    • If all NHSs are “down” then set tunnel interface up/down
    • Continue to send NHRP registration requests
    • If a single NHS is “up” then set tunnel interface up/up
  • Combine with ‘backup interface …’
    • Backup (tunnel) interface only up when main interface is down.
Tunnel Health Monitoring – Interface State (cont.)

```plaintext
#show ip nhrp nhs detail
  10.0.0.1  RE req-sent 100  req-failed 0  repl-recv 90 (00:01:38 ago)
  10.0.0.2  RE req-sent 125  req-failed 0  repl-recv 79 (00:01:38 ago)

#show interface tunnel0
  Tunnel0 is up, line protocol is up
  *Apr 19 21:32:52 NHRP: NHS-DOWN: 10.0.0.1
  *Apr 19 21:32:52 NHRP: NHS 10.0.0.1 Tunnel0 vrf 0 Cluster 0 Priority 0 Transitioned to 'E' from 'RE'
  *Apr 19 21:32:53 NHRP: NHS-DOWN: 10.0.0.2
  *Apr 19 21:32:53 NHRP: NHS 10.0.0.2 Tunnel0 vrf 0 Cluster 0 Priority 0 Transitioned to 'E' from 'RE'

  *Apr 19 21:33:02 %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel0, changed state to down
  *Apr 19 21:33:02 NHRP: if_down: Tunnel0 proto IPv4

#show ip nhrp nhs detail
  10.0.0.1  E req-sent 105  req-failed 0  repl-recv 90 (00:02:12 ago)
  10.0.0.2  E req-sent 130  req-failed 0  repl-recv 79 (00:02:12 ago)

#show interface tunnel0
  Tunnel0 is up, line protocol is down
  *Apr 19 21:33:12 NHRP: Send Registration Request via Tunnel0 vrf 0, packet size: 92
  *Apr 19 21:33:13 NHRP: Send Registration Request via Tunnel0 vrf 0, packet size: 92

  *Apr 19 21:34:36 NHRP: NHS 10.0.0.1 Tunnel0 vrf 0 Cluster 0 Priority 0 Transitioned to 'RE' from 'E'
  *Apr 19 21:34:36 NHRP: NHS-UP: 10.0.0.1

#show ip nhrp nhs detail
  10.0.0.1  RE req-sent 110  req-failed 0  repl-recv 96 (00:00:19 ago)
  10.0.0.2  RE req-sent 135  req-failed 0  repl-recv 79 (00:04:09 ago)

#show interface tunnel0
  Tunnel0 is up, line protocol is up
```

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
Extras

• Recent Features
  • IKEv2 with DMVPN
  • DMVPN IPv6 Transport
  • Routing protocol
IKEv2 with DMPVN

- DMVPN works with ISAKMP (IKEv1) and/or IKEv2
  - Transparent to DMVPN
  - Node can be responder for both ISAKMP and IKEv2
    - Both ISAKMP and IKEv2 are configured.
  - Node can be Initiator for either ISAKMP or IKEv2 not both
    - Configure under the ‘crypto ipsec profile …’

```plaintext
crypto isakmp policy 2
  encr aes
  authentication pre-share
group 2

crypto ikev2 keyring DMVPN
  peer DMVPN
  address 0.0.0.0 0.0.0.0
  pre-shared-key cisco123

crypto ikev2 profile DMVPN
  match identity remote address 0.0.0.0
  authentication local pre-share
  authentication remote pre-share
  keyring DMVPN

crypto isakmp key cisco123 address 0.0.0.0 0.0.0.0

crypto ipsec transform-set DMVPN esp-aes esp-sha-hmac
  mode transport [require]

crypto ipsec profile DMVPN
  set transform-set DMVPN
  set ikev2-profile DMVPN

interface Tunnel0
  ...tunnel protection ipsec profile DMVPN

With → initiate IKEv2
Without → initiate IKEv1
```
DMVPN over IPv6 Transport – 15.2(1)T

- IPv6 and IPv4 packets over DMVPN IPv6 tunnels
  - Introduced in IOS 15.2(1)T, 15.3(1)S
  - IPv6 infrastructure network
  - IPv6 and/or IPv4 data packets over same IPv6 GRE tunnel
  - NHRP modifies Routing Table
- Can run both DMVPN IPv4 and IPv6
  - Separate DMVPNs (mGRE tunnel)
  - DMVPN IPv4 ↔ IPv6 spoke to spoke via hub

- Configuration
  - Standard IPv6 configuration on Outside (WAN) interface
  - Small change on mGRE tunnel
  - Must use IKEv2 for IPsec encryption
- Split-tunneling
  - Enterprise versus ISP assigned IPv6 addresses at spoke
  - No NAT66
**DMVPN over IPv6 Transport – Configuration**

**Hub**

```
crypto ikev2 keyring DMVPN
  peer DMVPNv6
    address ::/0
    pre-shared-key cisco123v6
crypto ikev2 profile DMVPN
  match identity local-pre-share
tunnel source Serial2/0
tunnel mode multipoint ipv6
tunnel protection ipsec profile DMVPN

interface Tunnel0
  ip address 10.0.0.1 255.255.255.0

interface Serial2/0
  ip address 172.17.0.1 255.255.255.252
  ipv6 address 2001:DB8:0:100::1/64
  ipv6 route ::/0 Serial2/0
```

**Spoke**

```
crypto ikev2 keyring DMVPN
  peer DMVPNv6
    address ::/0
    pre-shared-key cisco123v6
crypto ikev2 profile DMVPN
  match identity remote-address ::/0
  authentication local-pre-share
  authentication remote-pre-share
  keyring DMVPN
dpd keepalive 30 5 on-demand
crypto ipsec profile DMVPN
  set transform-set DMVPN
  set ikev2-profile DMVPN

interface Tunnel0
  ip address 10.0.0.11 255.255.255.0

interface Serial2/0
  ip address 172.16.1.1 255.255.255.252
  ipv6 address 2001:DB8:0:FFFF:1::1/126
  ipv6 route ::/0 Serial1/0
```
DMVPN over IPv6 Transport – Data Structures

Hub1# show ip nhrp
10.0.0.11/32 via 10.0.0.11
Tunnel0 created 22:26:55, expire 00:03:37
Type: dynamic, Flags: unique registered used
NBMA address: 2001:DB8:0:FFFF:0:1:0:1

Hub1# show ipv6 nhrp
2001:DB8:0:100::B/128 via 2001:DB8:0:100::B
Tunnel0 created 22:27:52, expire 00:03:39
Type: dynamic, Flags: unique registered
NBMA address: 2001:DB8:0:FFFF:0:1:0:1
FE80::A8BB:CCFF:FE00:C800/128 via 2001:DB8:0:100::B
Tunnel0 created 22:27:52, expire 00:03:39
Type: dynamic, Flags: unique registered
NBMA address: 2001:DB8:0:FFFF:0:1:0:1

Hub1# show crypto session
Interface: Tunnel0; Session status: UP-ACTIVE
Peer: 2001:DB8:0:FFFF:0:1:0:1 port 500
IKEv2 SA: local 2001:DB8:0:FFFF:1::1/500
remote 2001:DB8:0:FFFF:0:1:0:1/500 Active
IPSEC FLOW: permit 47 host 2001:DB8:0:FFFF:1::1 host 2001:DB8:0:FFFF:0:1:0:1
Active SAs: 2, origin: crypto map
Routing Protocol Features – BGP

- iBGP Local-AS (15.2(2)T, 15.1(3)S (CSCtj48063))
  - Run iBGP over DMVPN
    - Tunnel end-point routers may have different native BGP ASs
    - Allows ‘neighbor ... local-as #’ and ‘neighbor ... remote-as #’ to be the same (iBGP)
    - ‘neighbor ... local-as #’ is different from local native BGP AS, ‘router bgp #’
      - Almost like eBGP within the router between the native AS and the AS over DMVPN
  - BGP Dynamic Neighbors to reduce configuration on hub
    - Added IPv6 Dynamic Neighbor support in 16.3, 15.6(3)M

```
router bgp 65000
  bgp listen range 10.0.0.0/24 peer-group spokes
  ...
  neighbor spokes peer-group
  neighbor spokes remote-as 65001
  neighbor spokes local-as 65001
  ...
```

- BRKSEC-3052
- © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
Routing Protocol Features – EIGRP

• Equal Cost MultiPath (15.2(3)T, 15.2(1)S (CSCsj31328))
  • Destination network is reachable via more than one DMVPN (mGRE tunnel) and the ip next-hop needs to be preserved (Phase 2).

  no ip next-hop-self eigrp <as> [no-ecmp-mode]

• Add-path (15.3(1)S (CSCtw86791))
  • Spoke site has multiple DMVPN spoke routers and want to be able to load-balance spoke-spoke tunnels (Phase 2).
  • Requires new “named” EIGRP router configuration

  router eigrp <name>
  address-family ipv4 unicast autonomous-system 1
  af-interface Tunnel0
  no next-hop-self
  add-path <paths> (<paths> = number of extra paths)
  no split-horizon
  ...

MPLS over DMVPN – 2547oDMVPN

• Single DMVPN to support network virtualization
  • Single mGRE tunnel on all routers

• Simplified MPLS configuration
  • Still adds complexity for managing and troubleshooting

• Routing:
  • EIGRP is used for routing outside the DMVPN
  • MP-BGP used for routing protocol over DMVPN
    • Redistribute EIGRP to/from BGP for transport over DMVPN, and Import/export of VRF routes

• Support:
  • DMVPN Phase 1 – hub-and-spoke only
  • DMVPN Phase 2 – spoke-spoke only after shortcut tunnel is up
  • DMVPN Phase 3 – full spoke-spoke support (15.4(1)S, 15.4(2)T)
**MPLS over DMVPN Phase 3**

- New support in NHRP to
  - Keep track of NHRP mapping table entries per VRF
  - Transport MPLS forwarding labels
    - MPLS LDP *not* used over DMVPN
    - MP-BGP still propagates VPN labels
- New CLI
  - `mpls nhrp` replaces `mpls ip` on the tunnel interface, provides
    - Tag switching over the Tunnel interface
    - Tag switching on spoke-spoke tunnels
    - Hub router must PE can be P/PE
  - `mpls mtu ...` applied before MPLS encap.
  - Per-tunnel QoS
    - MPLS experimental bits (15.5(3)M,S)
MPLS over DMVPN Phase 3 (cont)

# show ip nhrp

10.0.0.1/32 via 10.0.0.1
   Tunnel0 created 1d22h, never expire
   Type: static, Flags: used
   NBMA address: 172.17.0.1

10.0.0.13/32 via 10.0.0.13
   Tunnel0 created 00:00:08, expire 00:03:51
   Type: dynamic, Flags: router nhop rib
   NBMA address: 172.16.3.1

192.168.11.0/24 (CompA) via 10.0.0.11
   Tunnel0 created 00:00:07, expire 00:03:51
   Type: dynamic, Flags: router unique local
   NBMA address: 172.16.1.1
   (no-socket)

192.168.13.0/24 (CompA) via 10.0.0.13
   Tunnel0 created 00:00:07, expire 00:03:51
   Type: dynamic, Flags: router rib nho
   NBMA address: 172.16.3.1

# show ip route vrf CompA next-hop-over

Routing Table: CompA

... 192.168.1.0/24 is variably subnetted, 2 subnets, 2 masks
C  192.168.1.0/24 is directly connected, Ethernet0/0
L  192.168.1.1/32 is directly connected, Ethernet0/0
B  192.168.3.0/24 [200/0] via 10.0.0.1, 13:13:51
D  192.168.11.0/24 [200/307200] via 192.168.1.2, 13:14:01, Ethernet0/0
   [NHO][200/1] via 10.0.0.13, 00:00:29, Tunnel0

# show mpls forwarding

... 27 No Label 192.168.1.0/24[V] 0 aggregate/CompA
29 No Label 192.168.11.0/24[V] 2850 Et0/0 192.168.1.2
30 35 192.168.13.0/24[V] 0 Tu0 10.0.0.13
31 20 192.168.101.0/24[V] 0 Tu0 10.0.0.1
...
Per-tunnel QoS

(hub→spoke) 12.4(22)T; (spoke→hub, spoke→spoke) 15.5(1)S,T

- QoS per tunnel on hub and spokes
  - Dynamically select Hierarchical (parent/child) QoS Policy
    - **Receiving Node:** Configure NHRP group name on tunnel
    - **Sending Node:** Configure QoS template policies; Map NHRP group name to QoS template policy
  - Nodes with same NHRP group name are mapped to separate instances of QoS policy
  - Same policy used for both IPv4 and IPv6

- QoS policy applied at outbound physical interface
  - Classification done **before** GRE encapsulation by tunnel
    - ACL matches against Data IP packet
    - Don't configure 'qos pre-classify' on tunnel interface
  - Shaping/policing done on physical after IPsec encryption
  - On physical may have separate aggregate QoS policy
    - With only a class-default shaper (15.2(2)T,S)

- CPU intensive; can reduce hub scaling by about 50% on software forwarding platforms
# Per-tunnel QoS – Configurations

<table>
<thead>
<tr>
<th>Class-map Match-all TypeA_VOICE</th>
<th>Hub and Spokes</th>
<th>Policy-map TypeA</th>
</tr>
</thead>
<tbody>
<tr>
<td>match all typeA_voice</td>
<td></td>
<td>class typeA_voice</td>
</tr>
<tr>
<td>match access-group 100</td>
<td></td>
<td>priority 1000</td>
</tr>
<tr>
<td>class-map match-all TypeB_VOICE</td>
<td></td>
<td>class typeB_voice</td>
</tr>
<tr>
<td>match access-group 100</td>
<td></td>
<td>priority percent 20</td>
</tr>
<tr>
<td>class-map match-all TypeA_Routing</td>
<td></td>
<td>class typeB_Routing</td>
</tr>
<tr>
<td>match ip precedence 6</td>
<td></td>
<td>bandwidth percent 20</td>
</tr>
<tr>
<td>class-map match-all TypeB_Routing</td>
<td></td>
<td>class typeB_Routing</td>
</tr>
<tr>
<td>match ip precedence 6</td>
<td></td>
<td>bandwidth percent 10</td>
</tr>
<tr>
<td>policy-map TypeA_parent</td>
<td></td>
<td>class class-default</td>
</tr>
<tr>
<td>class default</td>
<td></td>
<td>shape average 3000000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>service-policy typeA</td>
</tr>
<tr>
<td>policy-map TypeB_parent</td>
<td></td>
<td>class class-default</td>
</tr>
<tr>
<td></td>
<td></td>
<td>shape average 2000000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>service-policy typeB</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Policy-map TypeB</th>
<th></th>
<th>Policy-map TypeA_child</th>
</tr>
</thead>
<tbody>
<tr>
<td>class typeA_voice</td>
<td></td>
<td>class typeA_voice</td>
</tr>
<tr>
<td>priority percent 20</td>
<td></td>
<td>priority percent 20</td>
</tr>
<tr>
<td>class typeB_Routing</td>
<td></td>
<td>class typeB_Routing</td>
</tr>
<tr>
<td>bandwidth percent 20</td>
<td></td>
<td>bandwidth percent 10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hub and Spokes</th>
<th>Hub</th>
<th>Spoke1,3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface Tunnel0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ip address 10.0.0.1 255.255.255.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ip nhrp map multicast dynamic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nhrp group typeB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nhrp map group typeA service-policy output typeA_parent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nhrp map group typeB service-policy output typeB_parent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ip nhrp redirect</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hub and Spokes</th>
<th>Spoke2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface Tunnel0</td>
<td></td>
</tr>
<tr>
<td>ip address 10.0.0.0.12 255.255.255.0</td>
<td></td>
</tr>
<tr>
<td>nhrp group typeB</td>
<td></td>
</tr>
<tr>
<td>nhrp map group typeA service-policy output typeA_parent</td>
<td></td>
</tr>
<tr>
<td>nhrp map group typeB service-policy output typeB_parent</td>
<td></td>
</tr>
<tr>
<td>ip nhrp nhs 10.0.0.1 nbma 172.17.0.1 multicast</td>
<td></td>
</tr>
<tr>
<td>interface Tunnel0</td>
<td></td>
</tr>
<tr>
<td>ip address 10.0.0.0.12 255.255.255.0</td>
<td></td>
</tr>
<tr>
<td>nhrp group typeB</td>
<td></td>
</tr>
<tr>
<td>nhrp map group typeA service-policy output typeA_parent</td>
<td></td>
</tr>
<tr>
<td>nhrp map group typeB service-policy output typeB_parent</td>
<td></td>
</tr>
<tr>
<td>ip nhrp nhs 10.0.0.1 nbma 172.17.0.1 multicast</td>
<td></td>
</tr>
</tbody>
</table>

Spoke1,3

Spoke2
Per-tunnel QoS – QoS Output on Hub

```
Hub#show ip nhrp
10.0.0.11/32 via 10.0.0.11
  Tunnel0 created 21:24:03, expire 00:04:01
  Type: dynamic, Flags: unique registered
  NBMA address: 172.16.1.1
  Group: typeA
10.0.0.12/32 via 10.0.0.12
  Tunnel0 created 21:22:33, expire 00:05:30
  Type: dynamic, Flags: unique registered
  NBMA address: 172.16.2.1
  Group: typeB
10.0.0.13/32 via 10.0.0.13
  Tunnel0 created 00:09:04, expire 00:04:05
  Type: dynamic, Flags: unique registered
  NBMA address: 172.16.3.1
  Group: typeA

Hub#show ip nhrp group-map
Interface: Tunnel0
  NHRP group: typeA
    QoS policy: typeA_parent
    Tunnels using the QoS policy:
    Tunnel destination overlay/transport address
    10.0.0.11/32: 172.16.1.1
    10.0.0.12/32: 172.16.2.1
  NHRP group: typeB
    QoS policy: typeB_parent
    Tunnels using the QoS policy:
    Tunnel destination overlay/transport address
    10.0.0.13/32: 172.16.3.1

Hub#show policy-map multipoint tunnel 0 <spoke> output
Interface Tunnel0 ↔ 172.16.1.1
  Service-policy output: typeA_parent
    Class-map: class-default (match-any)
      19734 packets, 6667163 bytes
      shape (average) cir 3000000, bc 12000, be 12000
    Service-policy : typeA
      Class-map: typeA_voice (match-all)
        3737 packets, 4274636 bytes
      Class-map: typeA_Routing (match-all)
        14424 packets, 1269312 bytes
      Class-map: class-default (match-any)
        1573 packets, 1123215 bytes
  Interface Tunnel0 ↔ 172.16.2.1
  Service-policy output: typeB_parent
    Class-map: class-default (match-any)
      11420 packets, 1076898 bytes
      shape (average) cir 2000000, bc 8000, be 8000
    Service-policy : typeB
      Class-map: typeB_voice (match-all)
        1005 packets, 128640 bytes
      Class-map: typeB_Routing (match-all)
        10001 packets, 880088 bytes
      Class-map: class-default (match-any)
        414 packets, 68170 bytes
  Interface Tunnel0 ↔ 172.16.3.1
  Service-policy output: typeA_parent
    Class-map: class-default (match-any)
      5458 packets, 4783903 bytes
      shape (average) cir 3000000, bc 12000, be 12000
    Service-policy : typeA
      Class-map: typeA_voice (match-all)
        4914 packets, 4734392 bytes
      Class-map: typeA_Routing (match-all)
        523 packets, 46004 bytes
      Class-map: class-default (match-any)
        21 packets, 14995 bytes
```
Agenda

- DMVPN Design Overview
  - DMVPN General
  - IWAN Specific

- NHRP Details
  - NHRP Overview
  - NHRP Registrations/Resolutions/Redirects
Phase 2 – Features

• Single mGRE interface with ‘tunnel protection …’
  • On Hubs and Spokes
  • Hubs must be inter-connected in a “Daisy chain” over same mGRE tunnel
  • IKE authentication information (Certificates, Wildcard Pre-shared Keys)

• Spoke-spoke data traffic direct
  • Reduced load on hub
  • Reduced latency
    • Single IPsec encrypt/decrypt

• Routing Protocol
  • Still hub-and-spoke
  • Cannot summarize spoke routes on hub
  • Routes on spokes must have IP next-hop of remote spoke (preserve next-hop)
Phase 2 – Process switching

- IP Data packet is forwarded out tunnel interface to IP next-hop from routing table
- NHRP looks in mapping table for IP destination
  - If Entry Found
    - Forward to NBMA from mapping table – overriding IP next-hop
  - If No Entry Found
    - Forward to IP next-hop (if in NHRP table) otherwise to NHS
    - If arriving interface was not tunnel interface
      - Initiate NHRP Resolution Request for IP next-hop and send via NHS path (first up NHS)
  - If (no socket) Entry Found
    - If arriving interface is not tunnel interface – convert entry to (socket)
    - Trigger IPsec to bring up crypto socket
    - Forward to IP next-hop (if in NHRP table) otherwise to NHS
Phase 2 – CEF Switching

• IP Data packet is forwarded out tunnel interface to IP next-hop from FIB table
  • If adjacency is of type Valid
    • Packet is encapsulated and forwarded by CEF out tunnel interface
    • NHRP is not involved
  • If adjacency is of type Glean or Incomplete
    • Punt packet to process switching
    • If original arriving interface was not this tunnel interface
    • Initiate NHRP Resolution Request for IP next-hop
      • Send resolution request for IP next-hop (tunnel IP address) of remote Spoke
      • Resolution request forwarded via NHS path (first up NHS)
      • Resolution reply is used to create NHRP mapping and to complete the Adjacency
Phase 2 – NHRP Resolution Request

- **Spoke A**: 192.168.1.1/24
  - Physical: 172.17.0.1
  - Tunnel0: 10.0.0.1
  - 10.0.0.11 → 172.16.1.1
  - 10.0.0.12 → 172.16.2.1
  - 10.0.0.0/24 → Conn.
  - 192.168.0.0/24 → 10.0.0.12
  - 192.168.1.0/24 → 10.0.0.11
  - 192.168.2.0/24 → 10.0.0.11

- **Spoke B**: 192.168.2.1/24
  - Physical: (dynamic)
  - Tunnel0: (dynamic)
  - 10.0.0.11 → 172.16.1.1
  - 10.0.0.12 → 172.16.2.1
  - 192.168.0.0/24 → Conn.
  - 192.168.1.0/24 → 10.0.0.11
  - 192.168.2.0/24 → 10.0.0.12

- **CEF FIB Table**: Data packet
- **CEF Adjacency**: NHRP mapping

---

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public 185
Phase 2 – NHRP Resolution Request

**Data packet**

**NHRP Resolution**

**NHRP mapping**

**CEF FIB Table**

**CEF Adjacency**

---

**Spoke A**

192.168.1.1/24

- Physical: 172.17.0.1
- Tunnel0: 10.0.0.1

- 10.0.0.11 \(\rightarrow\) 172.16.1.1
- 10.0.0.12 \(\rightarrow\) 172.16.2.1

**Spoke B**

192.168.2.1/24

- Physical: 172.16.2.1
- Tunnel0: 10.0.0.12

- 10.0.0.11 \(\rightarrow\) 172.16.1.1
- 10.0.0.12 \(\rightarrow\) 172.16.2.1

---

**CEF FIB Table**

**NHRP mapping**

**CEF Adjacency**

---

**Spoke A**

192.168.1.1/24

- 10.0.0.1 \(\rightarrow\) 172.17.0.1 (*)

- 192.168.0.0/24 \(\rightarrow\) 10.0.0.1
- 192.168.1.0/24 \(\rightarrow\) Conn.
- 192.168.2.0/24 \(\rightarrow\) 10.0.0.12

- 10.0.0.1 \(\rightarrow\) 172.17.0.1 (*)

**Spoke B**

192.168.2.1/24

- 10.0.0.1 \(\rightarrow\) 172.16.2.1

- 192.168.0.0/24 \(\rightarrow\) 10.0.0.1
- 192.168.1.0/24 \(\rightarrow\) 10.0.0.11
- 192.168.2.0/24 \(\rightarrow\) Conn.

- 10.0.0.1 \(\rightarrow\) 172.16.2.1
- 10.0.0.11 \(\rightarrow\) incomplete
- 10.0.0.12 \(\rightarrow\) incomplete
Phase 2 – NHRP Resolution Request

Data packet
NHRP Resolution

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping

CEF FIB Table

CEF Adjacency

Data packet
NHRP Resolution

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency
Phase 2 – NHRP Resolution Reply
Phase 2 – NHRP Resolution Reply
Phase 2 – NHRP Resolution Reply

CEF FIB Table
CEF Adjacency
NHRP mapping
Data packet
NHRP Resolution

Physical: 172.17.0.1
Tunnel0: 10.0.0.1

192.168.0.1/24

10.0.0.11 \(\rightarrow\) 172.16.1.1
10.0.0.12 \(\rightarrow\) 172.16.2.1

192.168.0.0/24 \(\rightarrow\) Conn.
192.168.1.0/24 \(\rightarrow\) 10.0.0.11
192.168.2.0/24 \(\rightarrow\) 10.0.0.12

10.0.0.11 \(\rightarrow\) 172.16.1.1
10.0.0.12 \(\rightarrow\) 172.16.2.1

10.0.0.1 \(\rightarrow\) 172.16.1.1
10.0.0.11 \(\rightarrow\) 172.16.1.1
10.0.0.12 \(\rightarrow\) 172.16.2.1 (l)

10.0.0.1 \(\rightarrow\) 172.17.0.1 (*)

Spoke A

192.168.1.1/24

10.0.0.1 \(\rightarrow\) 172.17.0.1 (*)
10.0.0.12 \(\rightarrow\) 172.16.2.1

192.168.0.0/24 \(\rightarrow\) 10.0.0.1
192.168.1.0/24 \(\rightarrow\) Conn.
192.168.2.0/24 \(\rightarrow\) 10.0.0.12

10.0.0.1 \(\rightarrow\) 172.17.0.1
10.0.0.12 \(\rightarrow\) 172.16.2.1

Spoke B

192.168.2.1/24

192.168.2.0/24 \(\rightarrow\) Conn.
192.168.1.0/24 \(\rightarrow\) 10.0.0.11
192.168.2.0/24 \(\rightarrow\) Conn.

10.0.0.1 \(\rightarrow\) 172.17.0.1
10.0.0.11 \(\rightarrow\) 172.16.1.1
10.0.0.12 \(\rightarrow\) 172.16.2.1

10.0.0.1 \(\rightarrow\) 172.16.2.1 (l)

Tunnel0: 10.0.0.11

Physical: (dynamic)
Tunnel0: 10.0.0.11

192.168.1.0/24 \(\rightarrow\) Conn.
192.168.0.0/24 \(\rightarrow\) 10.0.0.1
192.168.2.0/24 \(\rightarrow\) Conn.

10.0.0.1 \(\rightarrow\) 172.17.0.1 (*)

192.168.2.0/24 \(\rightarrow\) Conn.
192.168.1.0/24 \(\rightarrow\) 10.0.0.11
192.168.2.0/24 \(\rightarrow\) Conn.

10.0.0.1 \(\rightarrow\) 172.17.0.1
10.0.0.11 \(\rightarrow\) 172.16.1.1
10.0.0.12 \(\rightarrow\) 172.16.2.1

10.0.0.1 \(\rightarrow\) 172.17.0.1 (*)
10.0.0.11 \(\rightarrow\) 172.16.1.1
10.0.0.12 \(\rightarrow\) 172.16.2.1 (l)
Phase 2 – NHRP Resolution Reply

Data packet
NHRP Resolution

NHRP mapping
CEF FIB Table
CEF Adjacency

CEF FIB Table

CEF Adjacency

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
Phase 2 – NHRP Resolution Response Processing

• Receive NHRP Resolution reply
  • If using IPsec (tunnel protection …) then
    • Trigger IPsec to setup ISAKMP and IPsec SAs for tunnel
    • Data packets still forwarded via spoke-hub-…-hub-spoke path
    • IPsec triggers back to NHRP when done

• Install new mapping in NHRP mapping table

• Send trigger to CEF to complete corresponding CEF adjacency
  • Data packets now forwarded via direct spoke-spoke tunnel by CEF
  • NHRP no longer involved
Phase 2 – Refresh or Remove Dynamic mappings

- Dynamic NHRP mapping entries have finite lifetime
  - Controlled by ‘ip nhrp holdtime …’ on source of mapping (spoke)
- Background process checks mapping entry every 60 seconds
  - Process-switching
    - Used flag set each time mapping entry is used
    - If used flag is set and expire time < 120 seconds then refresh entry, otherwise clear used flag
  - CEF-switching
    - If expire time < 120 seconds, CEF Adjacency entry marked “stale”
    - If “stale” CEF Adjacency entry is then used, signal to NHRP to refresh entry
- Another resolution request is sent to refresh entry
  - Resolution request via NHS path; reply via direct tunnel
- If entry expires it is removed
  - If using IPsec → Trigger IPsec to remove IPsec/ISAKMP SAs
Phase 3 – NHRP Resolution Reply

(Prior to 15.2(1)T – ISR, 7200)
Phase 3 – CEF Switching

Data Packet Forwarding

• IP Data packet is forwarded out tunnel interface
  1. IP next-hop from CEF FIB mapped to Adjacency
     If adjacency is:
     • Glean or Incomplete → Punt to process switching
     • Valid → Select adjacency for the packet
  2. NHRP in Outbound CEF Feature path
     Look up packet IP destination in NHRP mapping table
     • Matching entry: Reselect adjacency → use direct spoke-spoke tunnel
     • No matching entry: Leave CEF adjacency → packet goes to hub
• If packet arrived on and is forwarded out the same tunnel interface
  • Forward data packet
  • If ‘ip nhrp redirect’ is on inbound tunnel then send NHRP redirect
• Packet is encapsulated, encrypted and forwarded

(Prior to 15.2(1)T – ISR, 7200)
Interaction with IWAN
Agenda

- DMVPN Design Overview
  - General and IWAN Specific
- NHRP Details
  - NHRP Overview
  - NHRP Registrations/Resolutions/Redirects
- Interaction with IWAN
  - f-VRFs
  - NHRP the RIB and PfR
- Recent and New Features
DMVPN with IWAN f-VRFs

- **MPLS f-VRF**
- **INTERNET f-VRF**
- **Global**

**Router**

- ISAKMP and IKE packets in respective f-VRF WAN

**DMVPN Tunnel0**

**DMVPN Tunnel1**

**Data Packets using Global**

**Data packets encapsulated in f-VRF tunnel**
DMVPN with IWAN f-VRFs

• Create VRF for each transport WAN interface (Ex: INTERNET, MPLS)
  • vrf definition <fvrf-name>

• “Outside” of tunnel is in front-door VRF (f-VRF)
  • interface tunnel<x>; tunnel vrf <fvrf-name>

• WAN (transport) interface is in f-VRF
  • interface <wan-interface>; vrf forwarding <fvrf-name>

• Crypto – ISAKMP/IKEv2 are also in f-VRFs
  • ISAKMP – need keyring for each f-VRF
  • IKEv2 – need keyring, IKEv2 profile and IPsec profile
    • Separate one for each f-VRF
    Or
    • Single one for all fVRFs by using ‘match fvrf any’ in IKEv2 profile
DMVPN with IWAN f-VRFs

f-VRF Configuration

```
vrf definition INTERNET ...
vrf definition MPLS ...
!
crypto ikev2 keyring DMVPN
  peer ANY
  address 0.0.0.0 0.0.0.0
  pre-shared-key cisco123
!
crypto ikev2 profile DMVPN
  match fvrf any
  match identity remote address 0.0.0.0
  authentication remote pre-share
  authentication local pre-share
  keyring local DMVPN
  dpd 20 5 on-demand ! Spokes only
!
crypto ipsec transform-set DMVPN esp-aes 256 esp-sha256-hmac
  mode transport
!
crypto ipsec profile DMVPN
  set transform-set DMVPN
  set ikev2-profile DMVPN
```

```
interface Tunnel0
  ip address 10.0.0.11 255.255.255.0 ...
  tunnel source FastEthernet0
  tunnel key 100000
  tunnel vrf INTERNET
  tunnel protection ipsec profile DMVPN
interface Tunnel1
  ip address 10.0.1.11 255.255.255.0 ...
  tunnel source FastEthernet1
  tunnel key 100001
  tunnel vrf MPLS
  tunnel protection ipsec profile DMVPN
!
interface FastEthernet0
  vrf forwarding INTERNET
  ip address 172.16.1.1 255.255.255.240
!
interface FastEthernet1
  vrf forwarding MPLS
  ip address 172.17.1.1 255.255.255.240
!
ip route vrf MPLS 0.0.0.0 0.0.0.0 172.17.1.2
ip route vrf INTERNET 0.0.0.0 0.0.0.0 172.16.1.2
```
DMVPN with IWAN f-VRFs

Routing

Spoke1#show ip route vrf *

D*EX 0.0.0.0/0 [170/2918400] via 10.0.1.2, 00:00:04, Tunnel1
10.0.0.0/8 is variably subnetted, 5 subnets, 3 masks
C 10.0.0.0/24 is directly connected, Tunnel0
C 10.0.1.0/24 is directly connected, Tunnel1
D 192.168.0.0/24 [90/2892800] via 10.0.1.2, 00:00:04, Tunnel1
192.168.1.0/24 is variably subnetted, 2 subnets, 2 masks
C 192.168.1.0/24 is directly connected, Ethernet0/0
D 192.168.10.0/24 [90/2918400] via 10.0.1.2, 00:32:39, Tunnel1

Routing Table: INTERNET
Gateway of last resort is 172.16.1.2 to network 0.0.0.0

S* 0.0.0.0/0 [1/0] via 172.16.1.2
172.16.0.0/16 is variably subnetted, 2 subnets, 2 masks
C 172.16.1.0/24 is directly connected, FastEthernet0

Routing Table: MPLS
Gateway of last resort is 172.17.1.2 to network 0.0.0.0

S* 0.0.0.0/0 [1/0] via 172.17.1.2
172.17.0.0/16 is variably subnetted, 2 subnets, 2 masks
C 172.17.1.0/24 is directly connected, FastEthernet1

Crypto

Spoke1#show crypto ikev2 session

Session-id:1845, Status:UP-ACTIVE, IKE count:1, CHILD count:1

<table>
<thead>
<tr>
<th>T-id</th>
<th>Local</th>
<th>Remote</th>
<th>fvrf/ivrf</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>172.16.1.1/500</td>
<td>172.16.0.1/500</td>
<td>INTERNET/none</td>
<td>READY</td>
</tr>
</tbody>
</table>

Life/Active Time: 86400/1263 sec
Child sa: local selector 172.16.1.1/0 - 172.16.1.1/65535
remote selector 172.16.0.1/0 - 172.16.0.1/65535
ESP spi in/out: 0x86D2651B/0x1B72FEB6

Session-id:1844, Status:UP-ACTIVE, IKE count:1, CHILD count:1

<table>
<thead>
<tr>
<th>T-id</th>
<th>Local</th>
<th>Remote</th>
<th>fvrf/ivrf</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>172.17.1.1/500</td>
<td>172.17.0.5/500</td>
<td>MPLS/none</td>
<td>READY</td>
</tr>
</tbody>
</table>

Life/Active Time: 86400/1290 sec
Child sa: local selector 172.17.1.1/0 - 172.17.1.1/65535
remote selector 172.17.0.5/0 - 172.17.0.5/65535
ESP spi in/out: 0xF8C63D42/0x66DEA87D
DMVPN with IWAN DIA

DIA packets “route” between Global and f-VRF
DMVPN with IWAN DIA

• Outbound
  • Block learning default through tunnel
    • Access-list: deny default; match everything else
    • Route-map: if match “learn” route
    • Apply route-map in Routing Protocol
      • EIGRP: use “distribute-list ... in <tunnel-interface>
      • BGP: use “neighbor ... in”
  • Static default route in global table forwarding out Internet WAN interface
    • ip route 0.0.0.0 0.0.0.0 <Internet-WAN> <next-hop>|dhcp <admin-distance>

• Inbound
  • Policy-based routing (PBR)
    • access-list: match internal networks
    • route-map: if match use global routing table
DMVPN with IWAN DIA

Inbound

```plaintext
interface FastEthernet0
description INTERNET
vrf forwarding INTERNET
ip address 172.16.1.1 255.255.255.240
ip policy route-map INET-INTERNAL
!
ip access-list extended INTERNAL-NETS
   permit ip any 10.0.0.0 0.0.1.255
   permit ip any 192.168.0.0 0.0.255.255
   permit ip any 172.20.0.0 0.0.255.255
!
route-map INET-INTERNAL permit 10
   match ip address INTERNAL-NETS
   set global
```

Outbound

```plaintext
router eigrp 1
distribute-list route-map BLOCK-DEFAULT in Tunnel0
   [distribute-list route-map BLOCK-DEFAULT in Tunnel1]
   network 10.0.0.0 0.0.1.255
   network 192.168.1.0
!
ip access-list standard ALL-EXCEPT-DEFAULT
   deny 0.0.0.0
   permit any
!
route-map BLOCK-DEFAULT permit 10
   match ip address ALL-EXCEPT-DEFAULT
!
ip route 0.0.0.0 0.0.0.0 FastEthernet0 172.16.1.2 10
```

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
### DMVPN with IWAN DIA

#### Before

<table>
<thead>
<tr>
<th>Command</th>
<th>Output</th>
</tr>
</thead>
</table>
| Spoke1#show ip eigrp topology | P 192.168.10.0/24, 1 successors, FD is 2918400  
via 10.0.1.2 (2918400/332800), Tunnel1  
via 10.0.0.1 (3020800/332800), Tunnel0  
P 172.20.1.0/24, 1 successors, FD is 409600  
via 192.168.1.2 (409600/128256), Ethernet0/0  
P 192.168.0.0/21, 1 successors, FD is 2892800  
via 10.0.1.2 (2892800/307200), Tunnel1  
via 10.0.0.1 (2995200/307200), Tunnel0  
P 192.168.1.0/24, 1 successors, FD is 281600  
via Connected, Ethernet0/0  
P 0.0.0.0/0, 1 successors, FD is 2918400  
via 10.0.1.2 (2918400/2636800), Tunnel1  
via 10.0.0.1 (3020800/2636800), Tunnel0 |

<table>
<thead>
<tr>
<th>Command</th>
<th>Output</th>
</tr>
</thead>
</table>
| Spoke1#show ip route | D 172.20.1.0 [90/409600] via 192.168.1.2, 01:47:00, Ethernet0/0  
D 192.168.0.0/21 [90/2892800] via 10.0.1.2, 00:20:27, Tunnel1  
C 192.168.1.0/24 is directly connected, Ethernet0/0  
D 192.168.10.0/24 [90/2918400] via 10.0.1.2, 00:32:39, Tunnel1 |

#### After

<table>
<thead>
<tr>
<th>Command</th>
<th>Output</th>
</tr>
</thead>
</table>
| Spoke1#show ip eigrp topology | P 192.168.10.0/24, 1 successors, FD is 2918400  
via 10.0.1.2 (2918400/332800), Tunnel1  
via 10.0.0.1 (3020800/332800), Tunnel0  
P 172.20.1.0/24, 1 successors, FD is 409600  
via 192.168.1.2 (409600/128256), Ethernet0/0  
P 192.168.0.0/21, 1 successors, FD is 2892800  
via 10.0.1.2 (2892800/307200), Tunnel1  
via 10.0.0.1 (2995200/307200), Tunnel0  
P 192.168.1.0/24, 1 successors, FD is 281600  
via Connected, Ethernet0/0  
P 0.0.0.0/0, 0 successors, FD is Infinity  
via 10.0.1.2 (2918400/2636800), Tunnel1 |

<table>
<thead>
<tr>
<th>Command</th>
<th>Output</th>
</tr>
</thead>
</table>
| Spoke1#show ip route | S 0.0.0.0/0 [170/2918400] via 172.16.1.2, 00:00:04, Tunnel1  
D 172.20.1.0 [90/409600] via 192.168.1.2, 01:47:00, Ethernet0/0  
D 192.168.0.0/21 [90/2892800] via 10.0.1.2, 00:20:27, Tunnel1  
C 192.168.1.0/24 is directly connected, Ethernet0/0  
D 192.168.10.0/24 [90/2918400] via 10.0.1.2, 00:32:39, Tunnel1 |
Agenda

• DMVPN Design Overview
  • General and IWAN Specific

• NHRP Details
  • NHRP Overview
  • NHRP Registrations/Resolutions/Redirects

• Interaction with IWAN
  • f-VRFs
    • NHRP the RIB and PfRv3

• Recent and New Features
Routing Protocol (RP), NHRP and PfRv3

- Routing protocol (RP) – destinations outside of the DMVPN
  - Advertises reachability of these destinations over any/all DMVPNs
  - Sets base forwarding within DMVPNs via the RIB

- PfRv3 – optimize forwarding of flows over different DMVPN paths
  - PfR RIB used to control forwarding of flows
  - Lookup alternate paths directly in RP database (except OSPF)
  - Bring up alternate paths, with probe traffic

- NHRP – optimizes forwarding within a single DMVPN
  - Shortcut (spoke-spoke) tunnels
    - Triggered by data traffic, including PfRv3 probe traffic
    - Changes forwarding by making changes in the RIB
    - Tracks RIB RP entries to control adding/removing shortcut tunnel
Basic DMVPN Design for IWAN

```
- Physical: (dynamic)
- Tunnel0: 10.0.0.11
- Tunnel1: 10.0.1.11
- Loop0: 172.18.0.11

Spoke A
- 192.168.1.0/24
- 192.168.11.0/24

Hub1
- Physical: 172.16.0.1
- Tunnel0: 10.0.0.1
- Loop0: 172.18.0.1

Hub2
- Physical: 172.17.0.5
- Tunnel0: 10.0.1.1
- Loop0: 172.18.0.2

MC
- Physical: 192.168.10.3
- Loop0: 172.18.0.10

Internet DMVPN
- 192.168.1.0/24
- 192.168.11.0/24
- 192.168.3.0/24
- 192.168.13.0/14

Internet
- 192.168.2.0/24
- 192.168.12.0/24

MPLS
- 192.168.10.0/24

Hub1
- Tunnel0: 10.0.0.1

Spoke B1
- Tunnel0: 10.0.0.12

Spoke B2
- Tunnel0: 10.0.1.12

Spoke B
- Tunnel1: 10.0.1.13

Spoke C
- Tunnel1: 10.0.1.1

Physical:
- 172.16.0.1
- 10.0.0.11
- 10.0.1.11
- 172.18.0.11

Physical: (dynamic)
- Tunnel0: 10.0.0.11
- Tunnel1: 10.0.1.11
- Loop0: 172.18.0.11

Physical:
- 172.17.0.5
- 10.0.1.1

Physical:
- 192.168.10.3
- 172.18.0.10

Physical:
- 192.168.3.0/24
- 192.168.13.0/14
```
DMVPN with Routing Protocol

Routing Protocol – Both paths

SpokeA# show ip eigrp topology

<table>
<thead>
<tr>
<th>Path</th>
<th>Successors</th>
<th>Metric</th>
<th>Next-Hop Address</th>
<th>Route Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>P 0.0.0.0/0, 0 successors, FD is Infinity</td>
<td>0</td>
<td>-</td>
<td>10.0.1.2 (1769472000/1048576000), Tunnel1</td>
<td></td>
</tr>
<tr>
<td>P 10.0.1.0/24, 1 successors, FD is 137625600</td>
<td>1</td>
<td>-</td>
<td>Connected, Tunnel1</td>
<td></td>
</tr>
<tr>
<td>P 10.0.0.0/24, 1 successors, FD is 1638400000</td>
<td>1</td>
<td>-</td>
<td>Connected, Tunnel0</td>
<td></td>
</tr>
<tr>
<td>P 192.168.0.0/21, 1 successors, FD is 1703936000</td>
<td>1</td>
<td>-</td>
<td>Connected, Tunnel1</td>
<td></td>
</tr>
<tr>
<td>P 192.168.1.0/24, 1 successors, FD is 131072000</td>
<td>1</td>
<td>-</td>
<td>Connected, Ethernet0/0</td>
<td></td>
</tr>
<tr>
<td>P 192.168.10.0/24, 1 successors, FD is 1769472000</td>
<td>1</td>
<td>-</td>
<td>Connected, Tunnel1</td>
<td></td>
</tr>
<tr>
<td>P 192.168.11.0/24, 1 successors, FD is 1966080000</td>
<td>1</td>
<td>-</td>
<td>Connected, Ethernet0/0</td>
<td></td>
</tr>
<tr>
<td>P 192.168.13.0/24, 1 successors, FD is 2228224000</td>
<td>1</td>
<td>-</td>
<td>Connected, Tunnel1</td>
<td></td>
</tr>
</tbody>
</table>

- **Default over MPLS**
- **Tunnel subnets**
- **Data Summary Route**
- **Local Subnet**
- **Data Specific Routes**
- **Not including MC/BR Loopback Routes**

| In RIB | Not in RIB
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>INET</td>
<td></td>
</tr>
</tbody>
</table>

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
DMVPN with Routing Protocol

RIB – Path via MPLS

SpokeA# show ip route

Gateway of last resort is 172.16.1.2 to network 0.0.0.0

S* 0.0.0.0/0 [10/0] via 172.16.1.2, Serial1/0
10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
C 10.0.0.0/24 is directly connected, Tunnel0
C 10.0.1.0/24 is directly connected, Tunnel1
172.18.0.0/32 is subnetted, 8 subnets
D 172.18.0.1 [90/12800640] via 10.0.0.1, 01:10:55, Tunnel0
D 172.18.0.2 [90/10752640] via 10.0.1.2, 01:10:55, Tunnel1
D 172.18.0.10 [90/13312640] via 10.0.1.2, 01:10:55, Tunnel1
C 172.18.0.11 is directly connected, Loopback0
D 172.18.0.13 [90/16384640] via 10.0.1.2, 01:10:55, Tunnel1
D 192.168.0.0/21 [90/13312000] via 10.0.1.2, 01:10:55, Tunnel1
192.168.1.0/24 is variably subnetted, 2 subnets, 2 masks
C 192.168.1.0/24 is directly connected, Ethernet0/0
D 192.168.10.0/24 [90/13824000] via 10.0.1.2, 01:10:55, Tunnel1
D 192.168.11.0/24 [90/1536000] via 192.168.1.2, 01:10:55, Ethernet0/0
D 192.168.13.0/24 [90/17408000] via 10.0.1.2, 01:10:55, Tunnel1

- Static Default for DIA
- Tunnel Interface Subnets
- MC/BR Loopback Subnet
- Data Summary Route
- Data Specific Routes
Forwarding over Primary DMVPN

Dual DMVPN

Hub1
- Physical: 172.16.0.1
- Tunnel0: 10.0.0.1
- Loop0: 172.18.0.1

Hub2
- Physical: 172.17.0.5
- Tunnel1: 10.0.1.1
- Loop0: 172.18.0.2

MC
- Physical: 192.168.10.3
- Loop0: 172.18.0.10

Internet DMVPN

MPLS

Spoke A
- Physical: (dynamic)
- Tunnel0: 10.0.0.11
- Tunnel1: 10.0.1.11
- Loop0: 172.18.0.11

Spoke C
- Physical: 192.168.3.0/24
- Tunnel0: 10.0.1.1
- Loop0: 172.18.0.13

Physical: 192.168.1.0 /24
192.168.11.0/24
192.168.3.0/24
192.168.5.0/14
192.168.10.0/24
192.168.13.0/14
Forwarding over Primary DMVPN

NHRP

SpokeA# show ip nhrp

10.0.1.13/32 via 10.0.1.13
   Tunnel1 created 00:04:23, expire 00:04:19
   Type: dynamic, Flags: router nhop rib
   NBMA address: 172.17.3.1
192.168.1.0/24 via 10.0.1.11
   Tunnel1 created 00:04:25, expire 00:01:36
   Type: dynamic, Flags: router unique local
   NBMA address: 172.17.1.1
   (no-socket)
192.168.3.0/24 via 10.0.1.13
   Tunnel1 created 00:01:40, expire 00:04:19
   Type: dynamic, Flags: router rib
   NBMA address: 172.17.3.1
192.168.11.0/24 via 10.0.1.11
   Tunnel1 created 00:04:02, expire 00:01:57
   Type: dynamic, Flags: router unique local
   NBMA address: 172.17.1.1
   (no-socket)
192.168.13.0/24 via 10.0.1.13
   Tunnel1 created 00:04:02, expire 00:01:57
   Type: dynamic, Flags: router rib nho
   NBMA address: 172.17.3.1

RIB

SpokeA# show ip route

   10.0.0.0/8 is variably subnetted, 5 subnets, 2 masks
   C   10.0.0.0/24 is directly connected, Tunnel0
   L   10.0.0.11/32 is directly connected, Tunnel0
   C   10.0.1.0/24 is directly connected, Tunnel1
   L   10.0.1.11/32 is directly connected, Tunnel1
   H   10.0.1.13/32 is directly connected, 00:05:28, Tunnel1
   D   192.168.0.0/21 [90/13312000] via 10.0.1.2, 00:11:02, Tunnel1
       192.168.1.0/24 is variably subnetted, 2 subnets, 2 masks
       C   192.168.1.0/24 is directly connected, Ethernet0/0
       L   192.168.1.1/32 is directly connected, Ethernet0/0
       H   192.168.3.0/24 [250/1] via 10.0.1.13, 00:03:06, Tunnel1
       D   192.168.10.0/24 [90/13824000] via 10.0.1.2, 00:11:02, Tunnel1
       D   192.168.11.0/24 [90/1536000] via 192.168.1.2, 00:11:02, Ethernet0/0
       D   192.168.13.0/24 [90/17408000] via 10.0.1.2, 00:11:02, Tunnel1
       [NHO][90/1] via 10.0.1.13, 00:05:28, Tunnel1

%
Forwarding over Secondary DMVPN

Dual DMVPN

Hub1
- Physical: 172.16.0.1
- Tunnel0: 10.0.0.1
- Loop0: 172.18.0.1

Hub2
- Physical: 172.17.0.5
- Tunnel1: 10.0.1.1
- Loop0: 172.18.0.2

Spoke A
- 192.168.1.0 /24
- 192.168.11.0/24

Spoke C
- 192.168.3.0/24
- 192.168.13.0/14

MC
- Physical: 192.168.10.3
- Loop0: 172.18.0.10

Internet DMVPN
- Internet

MPLS DMVPN
- MPLS

Primary path
- (dynamic)

Tunnel0:
- 10.0.0.11
- 10.0.1.11
- 172.18.0.11

Tunnel1:
- 10.0.0.13
- 10.0.1.13
- 172.18.0.13

Hub1
- Physical: (dynamic)
- Tunnel0: 10.0.0.11
- Tunnel1: 10.0.1.11
- Loop0: 172.18.0.11

Hub2
- Physical: (dynamic)
- Tunnel0: 10.0.0.13
- Tunnel1: 10.0.1.13
- Loop0: 172.18.0.13

Spoke C
- 192.168.3.0/24
- 192.168.13.0/14

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
Forwarding over Secondary DMVPN (nhrp route-watch)

NHRP

SpokeA# show ip nhrp

10.0.0.13/32 via 10.0.0.13
   Tunnel0 created 00:01:01, expire 00:05:07
   Type: dynamic, Flags: router nhop
   NBMA address: 172.16.3.1
192.168.1.0/24 via 10.0.0.11
   Tunnel0 created 00:01:01, expire 00:04:58
   Type: dynamic, Flags: router unique local
   NBMA address: 172.16.1.1
   (no-socket)
192.168.3.0/24 via 10.0.0.13
   Tunnel0 created 00:01:00, expire 00:04:59
   Type: dynamic, Flags: router
   NBMA address: 172.16.3.1
192.168.11.0/24 via 10.0.0.11
   Tunnel0 created 00:00:52, expire 00:05:07
   Type: dynamic, Flags: router unique local
   NBMA address: 172.16.1.1
   (no-socket)
192.168.13.0/24 via 10.0.0.13
   Tunnel0 created 00:00:52, expire 00:05:07
   Type: dynamic, Flags: router
   NBMA address: 172.16.3.1

RIB

SpokeA# show ip route

10.0.0.0/8 is variably subnetted, 4 subnets, 2 masks
C   10.0.0.0/24 is directly connected, Tunnel0
L   10.0.0.11/32 is directly connected, Tunnel0
C   10.0.1.0/24 is directly connected, Tunnel1
L   10.0.1.11/32 is directly connected, Tunnel1
D   192.168.0.0/21 [90/13312000] via 10.0.1.2, 00:04:38, Tunnel1
   192.168.1.0/24 is variably subnetted, 2 subnets, 2 masks
   C   192.168.1.0/24 is directly connected, Ethernet0/0
   L   192.168.1.1/32 is directly connected, Ethernet0/0
   D   192.168.10.0/24 [90/13624000] via 10.0.1.2, 00:04:38, Tunnel1
   D   192.168.11.0/24 [90/1536000] via 192.168.1.2, 00:04:38, Ethernet0/0
   D   192.168.13.0/24 [90/17408000] via 10.0.1.2, 00:04:38, Tunnel1

NHRP mapping entries not in RIB
No matching Parent Route
Forwarding over Secondary DMVPN

(no nhrp route-watch)

- Internet DMVPN
- MPLS DMVPN
- Primary path
- nhrp route-watch
- no nhrp route-watch

### Hub1
- Physical: 172.16.0.1
- Tunnel0: 10.0.0.1
- Tunnel1: 10.0.1.1
- Loop0: 172.18.0.11

### Hub2
- Physical: 172.17.0.5
- Tunnel0: 10.0.0.13
- Tunnel1: 10.0.1.13
- Loop0: 172.18.0.13

### MC
- Physical: 192.168.10.3
- Loop0: 172.18.0.10

### Spoke A
- Physical: (dynamic)
- Tunnel0: 10.0.0.11
- Tunnel1: 10.0.1.11
- Loop0: 172.18.0.11

### Spoke C
- Physical: (dynamic)
- Tunnel0: 10.0.0.13
- Tunnel1: 10.0.1.13
- Loop0: 172.18.0.13

### Internet DMVPN
- 192.168.1.0/24
- 192.168.11.0/24
- 192.168.3.0/24
- 192.168.13.0/14

### MPLS
- 192.168.10.0/24

### Primary path

### Secondary path

---

BRKSEC-3052 © 2017 Cisco and/or its affiliates. All rights reserved. Cisco Public
Forwarding over Secondary DMVPN

NHRP

SpokeA# show ip nhrp
10.0.0.13/32 via 10.0.0.13
  Tunnel0 created 00:00:36, expire 00:05:25
  Type: dynamic, Flags: router nhop rib
  NBMA address: 172.16.3.1
192.168.1.0/24 via 10.0.0.11
  Tunnel0 created 00:00:35, expire 00:05:24
  Type: dynamic, Flags: router unique local
  NBMA address: 172.16.1.1
    (no-socket)
192.168.3.0/24 via 10.0.0.13
  Tunnel0 created 00:00:34, expire 00:05:25
  Type: dynamic, Flags: router rib
  NBMA address: 172.16.3.1
192.168.11.0/24 via 10.0.0.11
  Tunnel0 created 00:00:24, expire 00:05:35
  Type: dynamic, Flags: router unique local
  NBMA address: 172.16.1.1
    (no-socket)
192.168.13.0/24 via 10.0.0.13
  Tunnel0 created 00:00:24, expire 00:05:35
  Type: dynamic, Flags: router rib nho
  NBMA address: 172.16.3.1

RIB

SpokeA# show ip route
10.0.0.0/8 is variably subnetted, 5 subnets, 2 masks
 C  10.0.0.0/24 is directly connected, Tunnel0
 L  10.0.0.11/32 is directly connected, Tunnel0
 H  10.0.0.1/32 is directly connected, Tunnel0
 C  10.1.0.0/24 is directly connected, Tunnel1
 L  10.1.1.1/32 is directly connected, Tunnel1
 D  192.168.0.0/21 [90/13312000] via 10.0.1.2, 00:11:02, Tunnel1
    192.168.1.0/24 is variably subnetted, 2 subnets, 2 masks
    C  192.168.1.0/24 is directly connected, Ethernet0/0
    L  192.168.1.1/32 is directly connected, Ethernet0/0
 H  192.168.3.0/24 [250/1] via 10.0.0.13, 00:00:34, Tunnel0
 D  192.168.10.0/24 [90/13824000] via 10.0.1.2, 00:11:02, Tunnel1
 D  192.168.11.0/24 [90/1536000] via 192.168.1.2, 00:11:02, Ethernet0/0
 D  192.168.13.0/24 [90/17408000] via 10.0.1.2, 00:11:02, Tunnel1
    [NHO][90/1] via 10.0.0.13, 00:00:28, Tunnel0

No Check for Parent Routes
Building spoke-spoke tunnels with NHRP and PfRv3

- PfRv3 Controlled Data flows
  - Forwards data flows over both primary and secondary DMVPN
    - PfR controls any load-balancing
  - Uses PfR Loopback as next-hop (Ex: 172.18.0.x)
    - NHRP triggered to build spoke-spoke tunnel over both DMVPNs
      - NHRP mapping entries to Loopback (Ex: 172.18.0.x)
      - NHRP modifies RIB for Loopback next-hop
        - If routing changes → PfR controlled flows quickly rerouted

- PfRv3 Uncontrolled Data flows
  - Data flows forwarded via the RIB
  - Uses primary DMVPN
  - Need ECMP routes to load-balancing over both DMVPNs
Building spoke-spoke tunnels with NHRP and PfRv3
## Forwarding over Primary and Secondary DMVPN

### NHRP

<table>
<thead>
<tr>
<th>Target</th>
<th>Via</th>
<th>NBMA</th>
<th>Mode</th>
<th>Intfc</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0.0.1/32</td>
<td>10.0.0.1</td>
<td>172.16.0.1</td>
<td>static</td>
<td>Tu0</td>
</tr>
<tr>
<td>10.0.0.11/32</td>
<td>10.0.0.11</td>
<td>172.16.1.1</td>
<td>dyn,loc</td>
<td>Tu0</td>
</tr>
<tr>
<td>10.0.0.13/32</td>
<td>10.0.0.13</td>
<td>172.16.3.1</td>
<td>dyn,rib</td>
<td>Tu0</td>
</tr>
<tr>
<td>172.18.0.11/32</td>
<td>10.0.0.11</td>
<td>172.16.1.1</td>
<td>dyn,loc</td>
<td>Tu0</td>
</tr>
<tr>
<td><strong>172.18.0.13/32</strong></td>
<td><strong>10.0.0.13</strong></td>
<td><strong>172.16.3.1</strong></td>
<td>dyn,nho</td>
<td><strong>Tu0</strong></td>
</tr>
<tr>
<td>10.0.1.12/32</td>
<td>10.0.1.2</td>
<td>172.17.0.5</td>
<td>static</td>
<td>Tu1</td>
</tr>
<tr>
<td>10.0.1.11/32</td>
<td>10.0.1.11</td>
<td>172.17.1.1</td>
<td>dyn,loc</td>
<td>Tu1</td>
</tr>
<tr>
<td><strong>10.0.1.13/32</strong></td>
<td><strong>10.0.1.13</strong></td>
<td><strong>172.17.3.1</strong></td>
<td>dyn,rib</td>
<td><strong>Tu1</strong></td>
</tr>
<tr>
<td>172.18.0.11/32</td>
<td>10.0.1.11</td>
<td>172.17.1.1</td>
<td>dyn,loc</td>
<td>Tu1</td>
</tr>
<tr>
<td><strong>172.18.0.13/32</strong></td>
<td><strong>10.0.1.13</strong></td>
<td><strong>172.17.3.1</strong></td>
<td>dyn,nho</td>
<td><strong>Tu1</strong></td>
</tr>
<tr>
<td>192.168.1.0/24</td>
<td>10.0.1.11</td>
<td>172.17.1.1</td>
<td>dyn,loc</td>
<td>Tu1</td>
</tr>
<tr>
<td><strong>192.168.3.0/24</strong></td>
<td><strong>10.0.1.13</strong></td>
<td><strong>172.17.3.1</strong></td>
<td>dyn,rib</td>
<td><strong>Tu1</strong></td>
</tr>
<tr>
<td>192.168.11.0/24</td>
<td>10.0.1.11</td>
<td>172.17.1.1</td>
<td>dyn,loc</td>
<td>Tu1</td>
</tr>
<tr>
<td><strong>192.168.13.0/24</strong></td>
<td><strong>10.0.1.13</strong></td>
<td><strong>172.17.3.1</strong></td>
<td>dyn,nho</td>
<td><strong>Tu1</strong></td>
</tr>
</tbody>
</table>

### RIB

<table>
<thead>
<tr>
<th>Target</th>
<th>Via</th>
<th>Subnet</th>
<th>Mode</th>
<th>Intfc</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.0.0.0/8</td>
<td></td>
<td>is variably subnetted, 6 subnets, 2 masks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C 10.0.0.0/24</td>
<td>is directly connected, <strong>Tunnel0</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L 10.0.0.11/32</td>
<td>is directly connected, <strong>Tunnel0</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H 10.0.1.13/32</td>
<td>is directly connected, 00:08:40, <strong>Tunnel0</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C 10.0.1.10/24</td>
<td>is directly connected, <strong>Tunnel1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L 10.0.1.11/32</td>
<td>is directly connected, <strong>Tunnel1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H 10.0.1.13/32</td>
<td>is directly connected, 00:09:05, <strong>Tunnel1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 172.18.0.0/32</td>
<td>is subnetted, 8 subnets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 172.18.0.1 [90/12800640] via 10.0.0.1, 02:07:25, <strong>Tunnel0</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 172.18.0.2 [90/10752640] via 10.0.1.2, 02:07:25, <strong>Tunnel1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 172.18.0.10 [90/13312640] via 10.0.1.2, 02:07:25, <strong>Tunnel1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C 172.18.0.11 is directly connected, Loopback0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D % 172.18.0.13 [90/16384640] via 10.0.1.2, 02:04:46, <strong>Tunnel1</strong> [NHO][90/1] via 10.0.0.1, 00:02:19, <strong>Tunnel0</strong> [NHO][90/1] via 10.0.1.13, 00:08:40, <strong>Tunnel1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 192.168.0.0/21 [90/13312000] via 10.0.0.1, 02:07:25, <strong>Tunnel1</strong> 192.168.1.0/24 is variably subnetted, 2 subnets, 2 masks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C 192.168.1.0/24 is directly connected, Ethernet0/0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L 192.168.1.1/32 is directly connected, Ethernet0/0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H 192.168.3.0/24 [250/1] via 10.0.1.13, 00:09:05, <strong>Tunnel1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 192.168.10.0/24 [90/13824000] via 10.0.1.2, 02:04:46, <strong>Tunnel1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 192.168.11.0/24 [90/1536000] via 192.168.1.2, 02:07:25, Ethernet0/0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D % 192.168.13.0/24 [90/17408000] via 10.0.1.2, 02:04:46, <strong>Tunnel1</strong> [NHO][90/1] via 10.0.1.13, 00:08:59, <strong>Tunnel1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

Routing Protocol (RP), NHRP and PfRv3

• Routing protocol (RP) – destinations outside of the DMVPN
  • Sets base forwarding for IWAN
  • Set preference for one DMVPN or can setup up ECMP routes

• PfRv3 – optimize forwarding of flows over different DMVPN paths
  • Find paths directly in RP database (except OSPF)
  • PfR RIB forwards flows over paths to MC/BR Loopback next-hop
  • Probe traffic over alternate paths

• NHRP – optimizes forwarding within a single DMVPN
  • Shortcut (spoke-spoke) tunnels
    • Triggered by data traffic and/or PfRv3 probe traffic
    • Use ‘no nhrp route-watch’ to enable shortcut tunnels over alternate paths
    • NHRP mapping/routes to MC/BR Loopback addresses
Thank You
Your Time Is Now