The bridge to possible

Сетевой марафон Cisco: Классика WAN

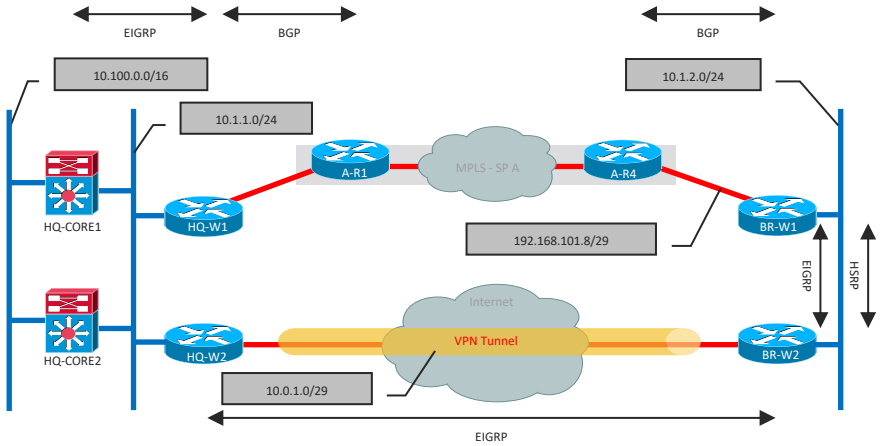
День 2. DMVPN

Денис Коденцев
Старший архитектор, CCIE
20 апреля 2021

DUAL WAN (MPLS + Internet)

PE-CE Protocol: BGP, Tunnel Protocol: EIGRP

- Headquarters WAN Edge
  - W1 learns Branch route via eBGP
  - W2 learns Branch route via EIGRP
- Headquarters Core
  - W1 redistributes eBGP into EIGRP, results in EIGRP external
  - W2 does not require redistribution, results in EIGRP internal
  - Core1, Core2 install Branch route via W2



HQ to Branch Traffic Flows Across Tunnel

```
HQ-W1#show ip route
B    10.1.2.0/24 [20/0] via 192.168.101.2, 05:24:01
```

```
HQ-W2#show ip route
D    10.1.2.0/24 [90/26882560] via 10.0.1.2, 00:00:04, Tunnel1
```

```
HQ-CORE1#show ip route
D    10.1.2.0/24 [90/26882816] via 10.1.1.210, 00:02:32, Vlan10
```

VPN Selection

Use Case/ Solution	DMVPN (mGRE, p-pGRE)	GETVPN (Tunnel-less)	FlexVPN (dVTI, IKEv2)	SSLVPN (TLS)	Easy VPN (IKEv1)	IPsec VPN (CM, sVTI, p-pGRE)
Remote Access	N-R	N-S	R	R	N-R	N-R
Hub-Spoke (HS)	R	N-S	R Non-Cisco Sp	N-R	N-R	N-R
HS + Spoke-Spoke	R	R	N-R	N-R	N-S	N-S
IoT	R	N-R	R	R	N-R	N-R
IWAN	R	N-S	N-S	N-S	N-S	N-S
MPLS over xVPN	R MPLS-o-DMVPN	R MPLS-o-mGRE	N-R MPLS-o-Flex	N-S	N-S	N-R MPLS-o-GRE

R = Recommended

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

Что такое DMVPN?

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

- Relies on two proven technologies
  - Next Hop Resolution Protocol (NHRP)

    Creates a distributed mapping database of
    VPN (tunnel interface) to real (public interface) addresses
  - Multipoint GRE Tunnel Interface

    Single GRE interface to support multiple GRE/IPsec tunnels and endpoints

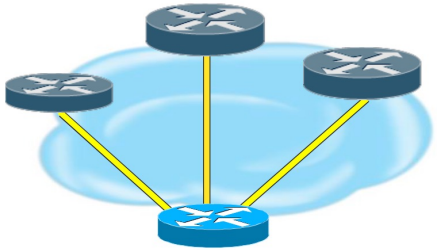
    Simplifies size and complexity of configuration

    Supports dynamic tunnel creation

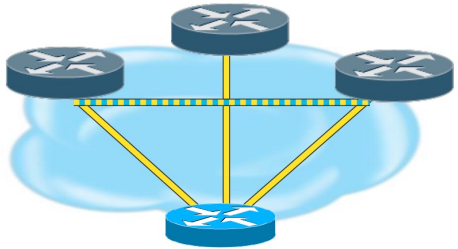
DMVPN: Основные отличительные особенности

- 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.

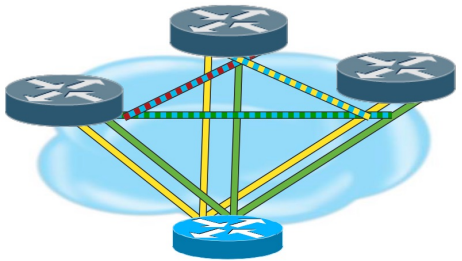
DMVPN: Типы внедрений



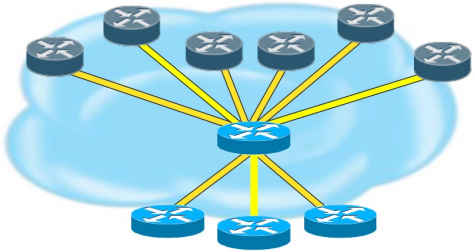
Hub and spoke
(Phase 1)



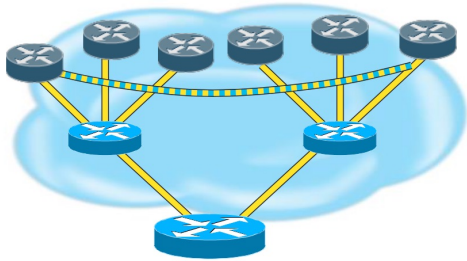
Spoke-to-spoke
(Phase 2)



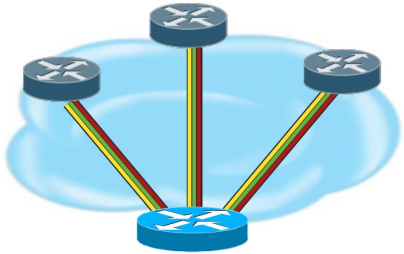
VRF-lite



Server Load Balancing



Hierarchical (Phase 3)

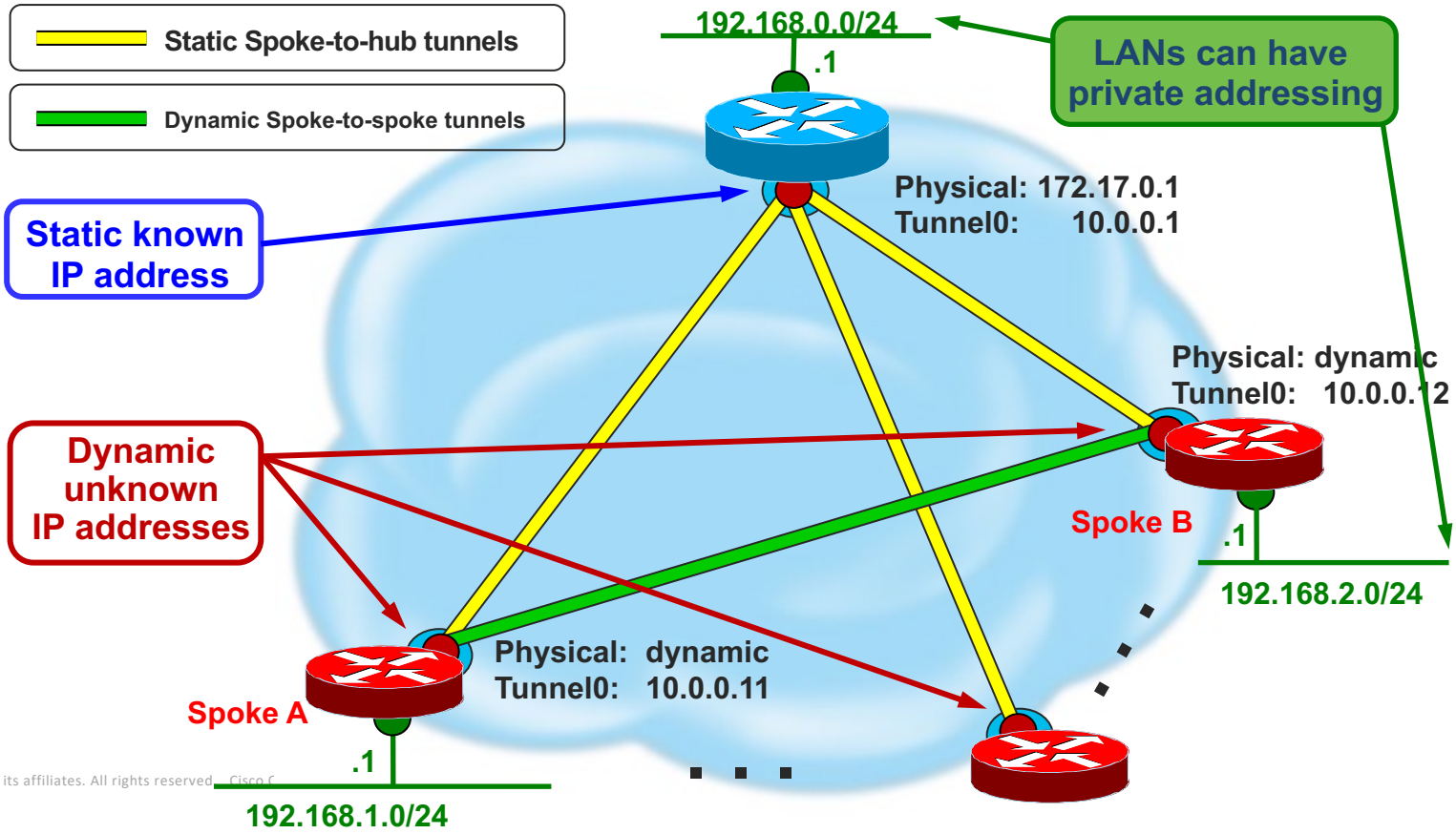


2547oDMVPN

DMVPN: Как он работает

- 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: Пример инфраструктуры



DMVPN: Компоненты

- Next Hop Resolution Protocol (NHRP)

  Creates a distributed (NHRP) mapping database of all the spoke's tunnel to real addresses

- Multipoint GRE Tunnel Interface (MGRE)

  Single GRE interface to support multiple GRE/IPsec tunnels and endpoints

  Simplifies size and complexity of configuration

  Supports dynamic tunnel creation

- IPsec tunnel protection

  Dynamically creates and applies encryption policies

- Routing

  Dynamic advertisement of branch networks; almost all routing protocols (EIGRP, RIP, OSPF, BGP, ODR) are supported

DMVPN: Три фазы

Phase 1	Phase 2	Phase 3
• Hub and spoke functionality • Interfaces: p-pGRE on spokes, mGRE on hubs • Simplified and smaller configuration on hubs • Support dynamically addressed CPEs (NAT) • Support for routing protocols and multicast • Spokes don’t need full routing table – can summarize on hubs	• Spoke to spoke functionality • mGRE interface on spokes • Direct spoke to spoke data traffic reduces load on hubs • Hubs must interconnect in daisy-chain • Spoke must have full routing table – no summarization • Spoke-spoke tunnel triggered by spoke itself • Routing protocol limitations	• Increase architecture designs and scaling • Same Spoke to Hub ratio • No hub daisy-chain • Spokes don’t need full routing table – can summarize • Spoke-spoke tunnel triggered by hubs • Remove routing protocol limitations

Next-Hop Resolution Protocol (NHRP)

- NHRP Registrations
  - Spoke (NHC) dynamically register its VPN to NBMA address mapping with hub (NHS).

    Static NHRP mappings on spokes for Hub (NHS)

    Needed to “start the game”

    Builds hub-and-spoke control plane network
- NHRP Resolutions
  - Dynamically resolve spoke to spoke VPN to NBMA mapping to build spoke-spoke tunnels.
  - Single instead of multiple tunnel hops across NBMA network
  - NHRP Resolution requests/replies sent via hub-and-spoke control plane path

```
interface Tunnel0
 bandwidth 1000
 ip address 10.0.0.2 255.255.255.0
 ip mtu 1400

! The following line must match on all nodes that want to use this mGRE tunnel:
ip nhrp authentication donttell

! Definition of NHRP server at the hub (10.0.0.1), which is permanently mapped to the static public address of the hub (172.17.0.1).
ip nhrp map 10.0.0.1 172.17.0.1

! Sends multicast packets to the hub router, and enables the use of a dynamic routing protocol between the spoke and the hub.
ip nhrp map multicast 172.17.0.1

! The following line must match on all nodes that want to use this mGRE tunnel:
ip nhrp network-id 99
ip nhrp holdtime 300

! Configures the hub router as the NHRP next-hop server.
ip nhrp nhs 10.0.0.1

ip tcp adjust-mss 1360
delay 1000
tunnel source Gigabitethernet 0/0/0 tunnel mode gre
multipoint
```

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

Hub

```
interface Tunnel0
   ...
   ip nhrp authentication test
   ip nhrp map multicast dynamic
   ip nhrp network-id 100000
   ip nhrp holdtime 600
   ip nhrp nhs 10.0.0.1 nbma 172.17.0.1 multicast
   ip nhrp redirect
   ...
!
```

Spoke

```
interface Tunnel0
   ...
   ip nhrp authentication test
   ip nhrp network-id 100000
   ip nhrp holdtime 600
   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
   ip nhrp registration no-unique
   ip nhrp shortcut
   ...
!
```

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
!
```


15

DMVPN и IPsec

- IPsec integrated with DMVPN, but not required
- Packets Encapsulated in GRE, then Encrypted with IPsec
- NHRP controls the tunnels, IPsec does encryption
- All DMVPN traffic, data and control, protected by encryption
- ISAKMP Keepalives monitor state of spoke-spoke tunnels

DMVPN и IPsec (продолжение)

- Bringing up a tunnel
  - NHRP signals IPsec to setup encryption
  - ISAKMP 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
  - IPsec can signal NHRP if encryption is cleared or lost

Маршрутизация

- Supports all routing protocols, except ISIS
- Best routing protocols are EIGRP and BGP
- 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)
    - All Phases: Turn off split-horizon (EIGRP, RIP)

Маршрутизация (продолжение)

- Hubs are routing neighbors with other hubs
  - Phase 1 & 3: Can use different routing protocol than on hub-spoke tunnels
  - Phase 2: Must use same routing protocol as on 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)

Пример таблицы маршрутизации (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 [90/2841600] via 10.0.0.1, 00:00:08, Tunnel0
```

Phase 1 & 3
(without summarization)

```
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 [90/297372416] 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 [90/297321216] via 10.0.0.1, 00:02:36, Tunnel0
D    192.168.3.0/24 [90/297321216] 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)

```
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 [90/297372416] 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 [90/297321216] via 10.0.0.12, 00:42:34, Tunnel0
D    192.168.3.0/24 [90/297321216] via 10.0.0.13, 00:42:34. Tunnel0
. . .
S*   0.0.0.0/0 [1/0] via 172.16.1.1
```

Протоколы маршрутизации в 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
  - 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

Полезные функции протокола маршрутизации EIGRP

- Equal Cost MultiPath
  - Destination network is reachable via more than one DMVPN (mGRE tunnel) and the ip next-hop needs to be preserved over both paths (Phase 2).
    - ‘no ip next-hop-self eigrp <as> [no-ecmp-mode]’
- Add-path
  - Spoke site has multiple DMVPN spoke routers and want to be able to load-balance spoke-spoke tunnels going into this spoke site (Phase 2).
    - Requires new “named” EIGRP router configuration
      - router eigrp addpath
        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

Протоколы маршрутизации в DMVPN: BGP

- Base Distance Vector style matches with DMVPN NBMA network style
- iBGP
  - Allows use of MED to control/compare routing
  - Dynamic Neighbors; May need to use “local-as” for iBGP
- eBGP is okay
  - AS-Path length is only thing to control/compare routing
- Good scaling but with slower convergence (hello 15+, hold 45+)
- Good metric control
  - 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

Полезные функции протокола маршрутизации BGP

- iBGP Local-AS
  - 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
    - Also use BGP Dynamic Neighbors to reduce configuration on hub

Протоколы маршрутизации в 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
    - Area 0 extended over WAN – possible stability issues for Area 0
  - Non-Area 0 over DMVPN
    - All spokes sites in same 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 Equal Cost multi-path (ECMP) route selection issues

Маршрутизация в DMVPN: Подводя итоги

- 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
- BUT...
  - EIGRP being an advanced distance vector protocol matches really well with DMVPN network topologies
  - BGP, specifically iBGP, can run well over DMVPN, but it is more complicated to setup and to have it act more like an IGP rather than a EGP.
  - OSPF can run over DMVPN, BUT lesser scaling and Area 0 issues really complicate the network.

Резервирование

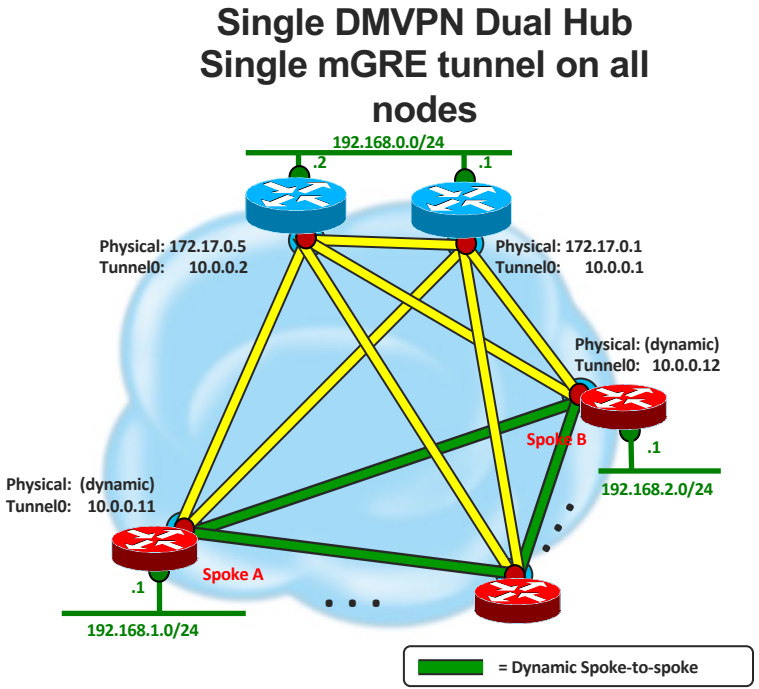
- Active-active redundancy model – two or more hubs per spoke
  - All configured hubs are active and are routing neighbors with spokes
  - Routing protocol routes are used to determine traffic forwarding
- ISAKMP/IPsec
  - Cannot use IPsec Stateful failover (NHRP isn't supported)
  - ISAKMP keepalives on spokes for timely hub recovery
- Can use single or multiple DMVPNs for redundancy
  - Each mGRE interface is a separate DMVPN network
  - 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)
- Hubs – interconnect and routing
  - Hubs exchange routing over DMVPN network

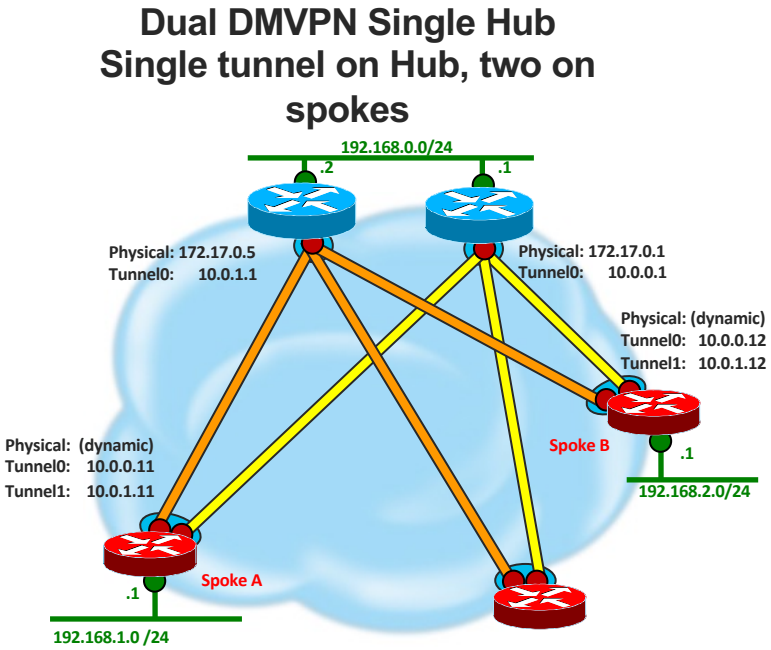
Организация Spoke-Spoke туннелей

- Resiliency
  - No monitoring of spoke-spoke tunnel (use ISAKMP keepalives)
- Path Selection
  - NHRP will always build spoke-spoke tunnel
  - No latency measurement of spoke-spoke vs spoke-hub-spoke paths
- Overloading spoke routers
  - CPU or memory → IKE Call Admission Control (CAC)
  - Bandwidth → Design for expected traffic
    - Hub-spoke versus Spoke-spoke; Spoke-spoke availability is best effort

Дизайн

Основные дизайны DMVPN

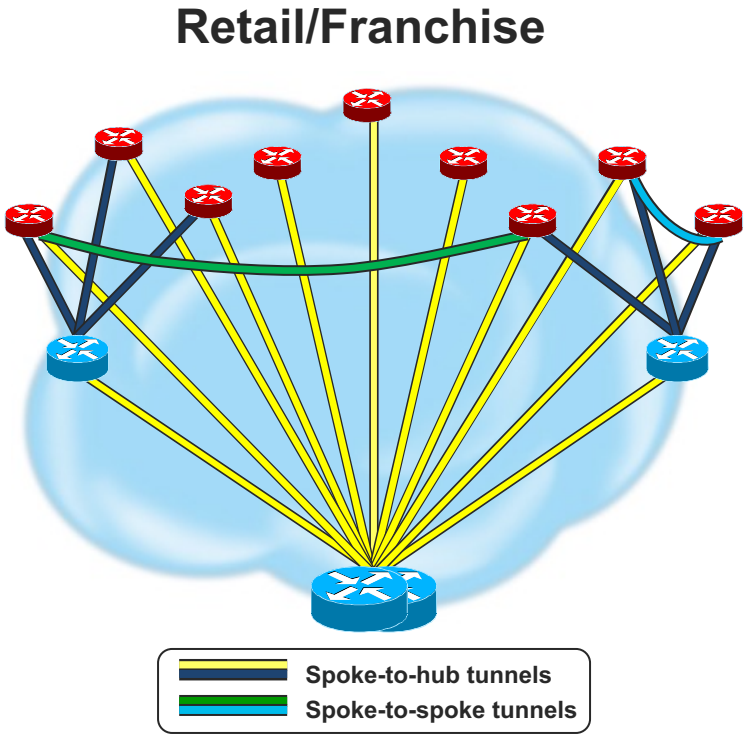


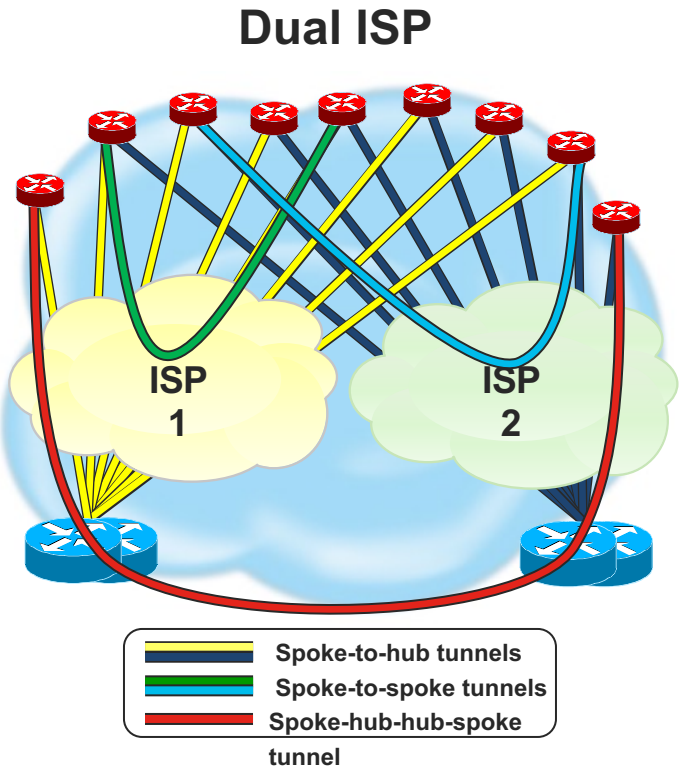


Несколько DMVPN против одного DMVPN

- Multiple DMVPNs
  - Best for Hub-and-spoke only
    - Easier to manipulate RP metrics between DMVPNs for Load-sharing
      - EIGRP – Delay on tunnel, BGP – Communities; OSPF – Cost
    - Performance Routing (PfR) selects between interfaces
  - Load-balancing over multiple ISPs (physical paths)
    - Load-balance data flows over tunnels → Better statistical balancing
- Single DMVPN
  - Best for spoke-spoke DMVPN
    - Can only build spoke-spoke within a DMVPN not between DMVPNs
    - More difficult to manipulate RP metrics within DMVPN for Load-sharing
      - EIGRP – Route tagging; BGP – Communities; OSPF – Can’t do
  - Load-balancing over multiple ISPs (physical paths)
    - Load-balance tunnel destinations or physical → Worse statistical balancing

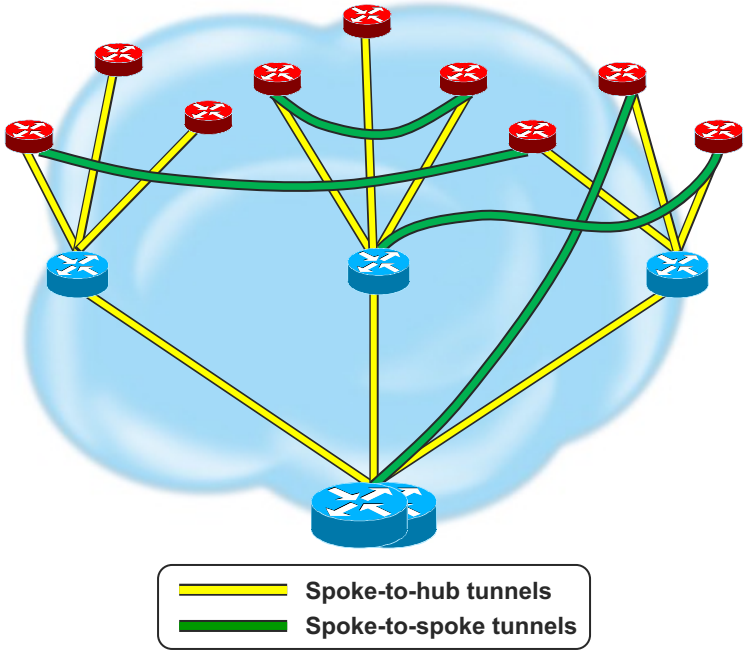
Комбинация дизайнов DMVPN



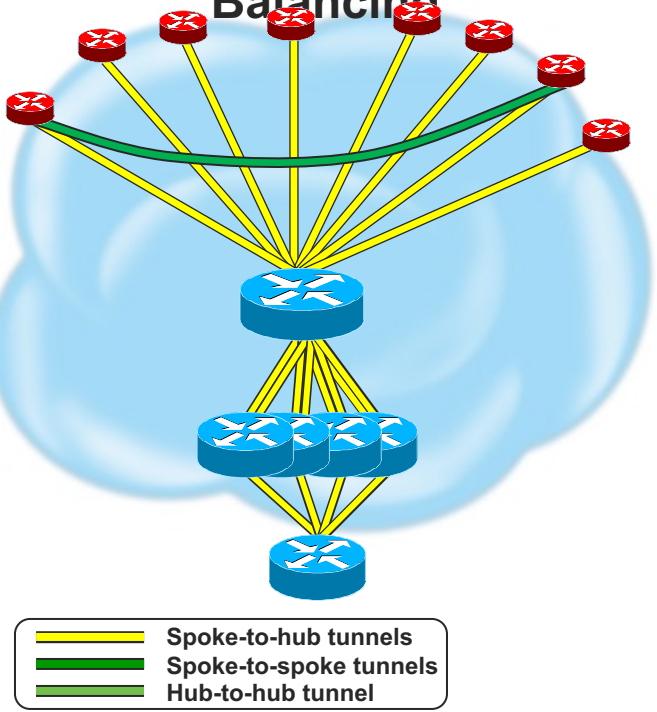


Комбинация дизайнов DMVPN (продолжение_

Hierarchical



Server Load Balancing



Виртуализация сети: Разделение DMVPN – VRF-lite

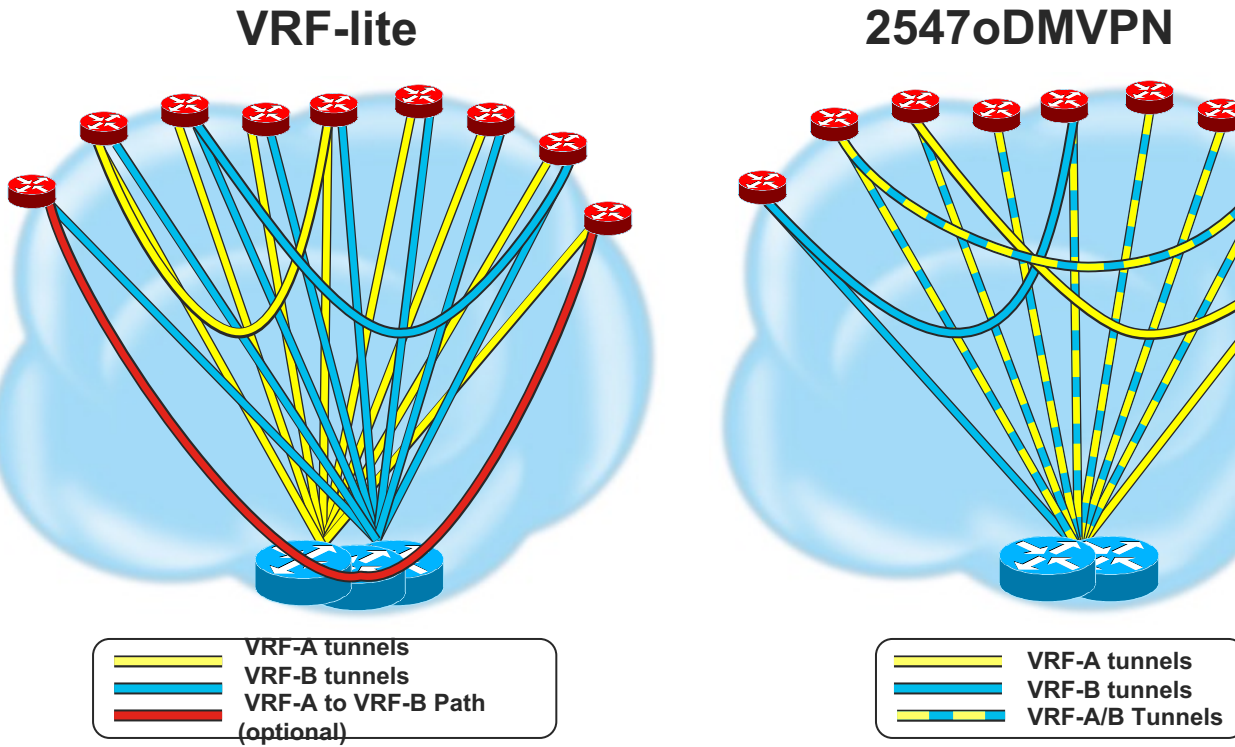
- Separate DMVPN mGRE tunnel per VRF
- Hub routers handle all DMVPNs
  - Multiple Hub routers for redundancy and load
- IGP used for routing protocol outside of and over DMVPNs on Spokes and Hubs
  - Address family per VRF
  - Routing neighbor per spoke per VRF
- BGP used only on the hub
  - Redistribute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

MPLS over DMVPN (ex. 2547oDMVPN)

Network Virtualization
MPLS over DMVPN – 2547oDMVPN

- Single DMVPN (Hub-and-spoke Only)
  - MPLS VPN over DMVPN
  - Single mGRE tunnel on all routers
- MPLS configuration
  - Hub and Spoke routers are MPLS PEs
- Multiple Hub routers for redundancy and load
- IGP is used for routing outside of DMVPN network
- BGP used for routing protocol over DMVPN
  - Redistribute between IGP and BGP for transport over DMVPN
  - Import/export of routes between VRFs and Global (or Internet VRF)
  - One routing neighbor per spoke
- Global routing table for routing DMVPN tunnel packets

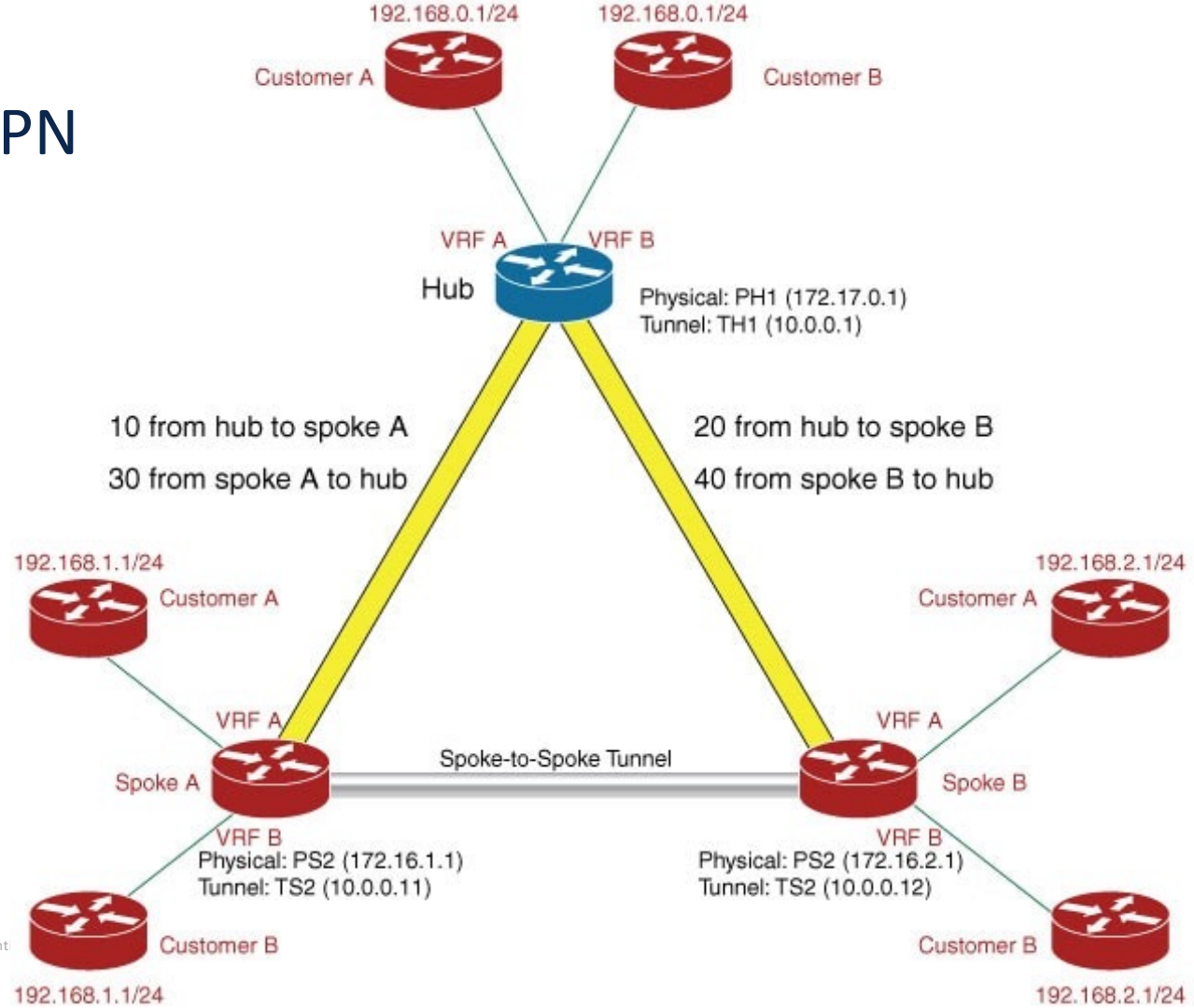
DMVPN Network Virtualization Designs



MPLS over DMVPN (ex. 2547oDMVPN)

- **NHRP**—Next Hop Resolution Protocol (NHRP) resolves the remote overlay address and dynamically discovers the transport end point needed to establish a secure tunnel.
- **MPLS**—Multiprotocol label switching (MPLS) enables MPLS tag switching for data packets. Label Distribution Protocol (LDP) is not enabled between spokes.
- **MFI**—Multicast Forwarding Information (MFI) allocates and releases labels assigned to tunnels.
- **MP-BGP**—Multiprotocol BGP (MP-BGP) distributes overlay labels for the customer network on different VRFs.

MPLS over DMVPN



Дополнительная информация

Per-tunnel QoS

- QoS per tunnel (spoke) on hub
  - Dynamically selected Hierarchical (parent/child) QoS Policy
    - Spoke: Configure NHRP group name
    - Hub: NHRP group name mapped to QoS template policy
  - Multiple spokes with same NHRP group mapped to individual instances of same QoS template policy
- QoS policy applied at outbound physical interface
  - Classification done before GRE encapsulation by tunnel
    - ACL match against Data IP packet
    - ‘qos pre-classify’ not configured on tunnel interface
  - Shaping/policing done on physical after IPsec encryption
  - Can’t have separate aggregate QoS policy on physical

Per-tunnel QoS Configurations

Hub

```
class-map match-all typeA_voice
  match access-group 100
class-map match-all typeB_voice
  match access-group 100
class-map match-all typeA_Routing
  match ip precedence 6
class-map match-all typeB_Routing
  match ip precedence 6

policy-map typeA
  class typeA_voice
    priority 1000
  class typeA_Routing
    bandwidth percent 20

policy-map typeB
  class typeB_voice
    priority percent 20
  class typeB_Routing
    bandwidth percent 10

policy-map typeA_parent
  class class-default
    shape average 3000000
    service-policy typeA

policy-map typeB_parent
  class class-default
    shape average 2000000
    service-policy typeB
```

Hub (cont)

```
interface Tunnel0
  ip address 10.0.0.1 255.255.255.0
  …
  ip nhrp map group typeA service-policy output typeA_parent
  ip nhrp map group typeB service-policy output typeB_parent
  …
  ip nhrp redirect
  no ip split-horizon eigrp 100
  ip summary-address eigrp 100 192.168.0.0 255.255.192.0 5
  …
```

Spoke1

```
interface Tunnel0
  ip address 10.0.0.11 255.255.255.0
  …
  ip nhrp group typeA
  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
  …
```

Spoke2

```
interface Tunnel0
  ip address 10.0.0.12 255.255.255.0
  …
  ip nhrp group typeB
  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
  …
```

Spoke3

```
interface Tunnel0
  ip address 10.0.0.13 255.255.255.0
  …
  ip nhrp group typeA
  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
  …
```

Per-tunnel QoS Output

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
     10.0.0.11/172.16.1.1
     10.0.0.13/172.16.3.1
  NHRP group: typeB
     QoS policy: typeB_parent
     Tunnels using the QoS policy:
     Tunnel destination overlay/transport address
     10.0.0.12/172.16.2.1
```

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
```

Tunnel Health Monitoring

- 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

```
interface Tunnel0
 ip address 10.0.0.11 255.255.255.0
 …
 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
 ip nhrp nhs 10.0.0.2
 …
 if-state nhrp
 …
```

Tunnel Health Monitoring (продолжение)

```
#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
*Apr 19 21:34:42 %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel0, changed state to up
*Apr 19 21:34:42 NHRP: if_up: Tunnel0 proto 0

#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     E req-sent 135  req-failed 0  repl-recv 79 (00:04:09 ago)

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

Next-Hop Resolution Protocol: Детали

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

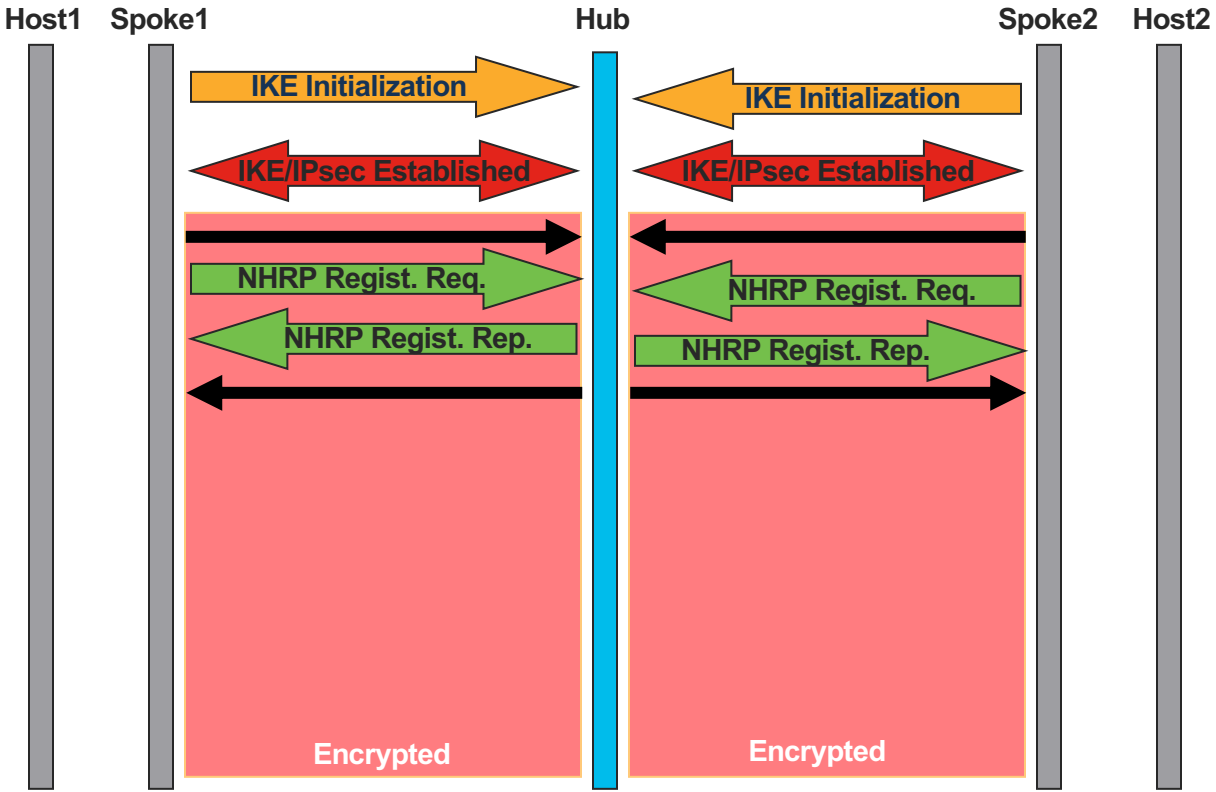
NHRP: Регистрация

- Builds base hub-and-spoke network
  - Hub-and-spoke data traffic
  - Control traffic; NHRP, Routing protocol, IP multicast
  - Phase 2 – Single level hub-and-spoke
  - Phase 3 – Hierarchical hub-and-spoke (tree).
- Next Hop Client (NHC) has static mapping for Next Hop Servers (NHSs)

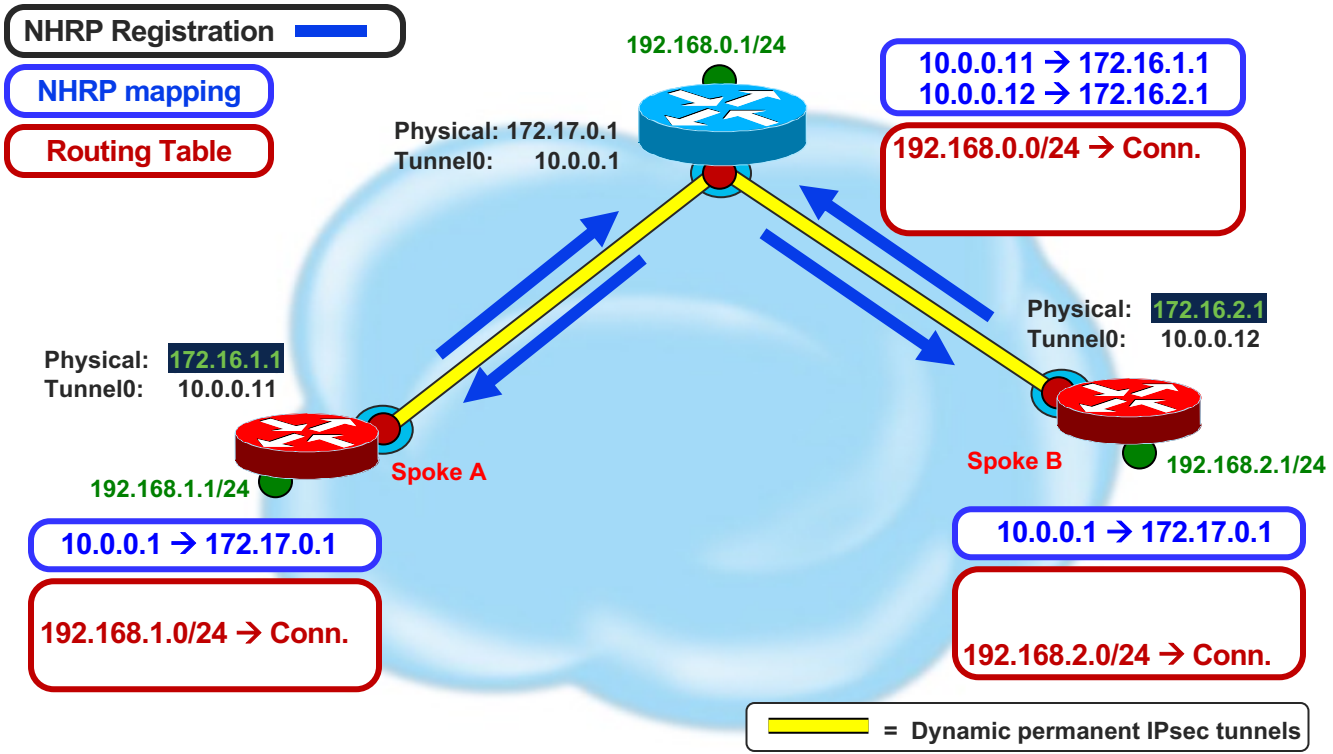
NHRP: Регистрация (продолжение)

- NHC dynamically registers own mapping with NHS
  - Supports spokes with dynamic NBMA addresses or NAT
  - Supplies outside NAT address of Hub
  - NHRP-group for per-Tunnel QoS
- NHS registration reply gives liveliness of NHS
  - Supplies outside NAT address of spoke

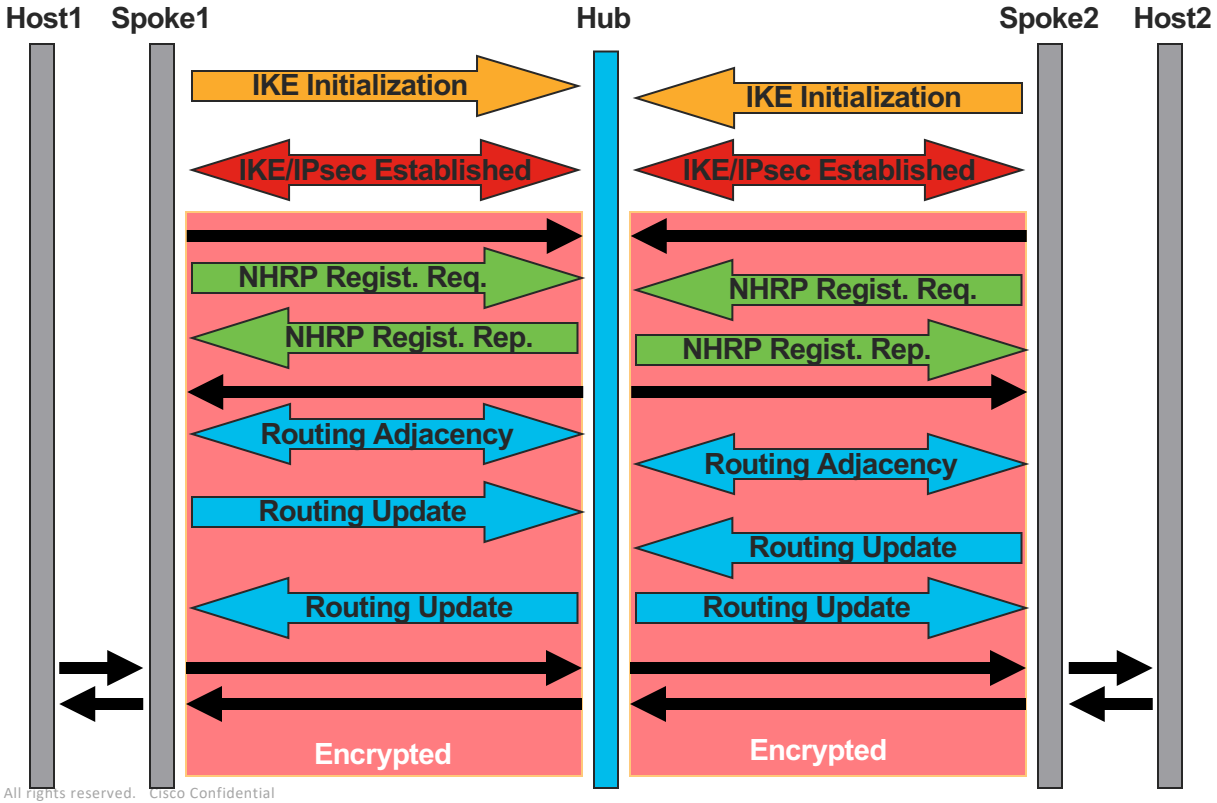
NHRP: Построение туннелей Spoke → Hub



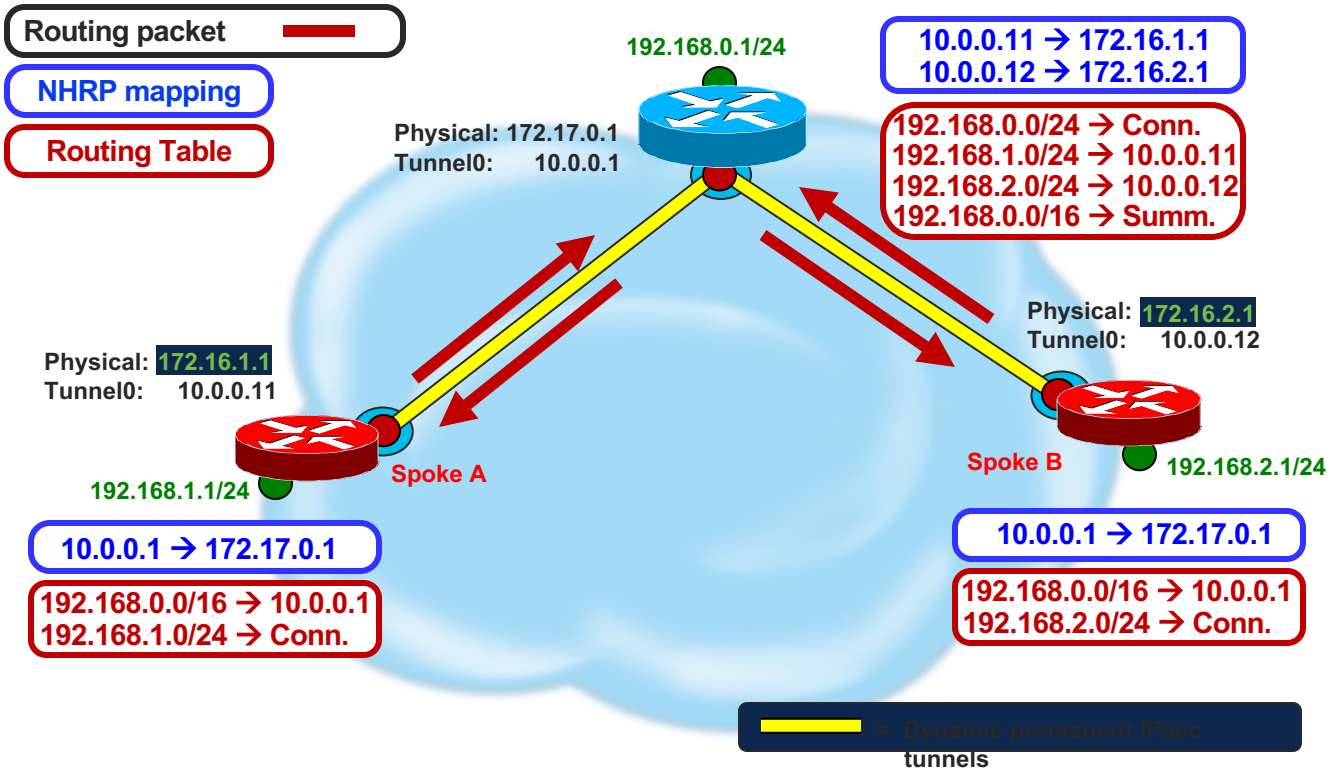
NHRP: Построение туннелей Spoke → Hub



NHRP: Routing Adjacency



NHRP: Routing Adjacency



Hub-and-Spoke: Передача данных

- **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

DMVPN Phase 3

Phase 3 – Features

- Used to increase scale of DMVPN networks
  - 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 routing and CEF switching to forward data and NHRP packets optimally through hubs
  - Reduces complexity and load for routing protocol
- OSPF routing protocol not limited to 2 hubs
  - Network point-multipoint mode
  - Still single OSPF area and no summarization

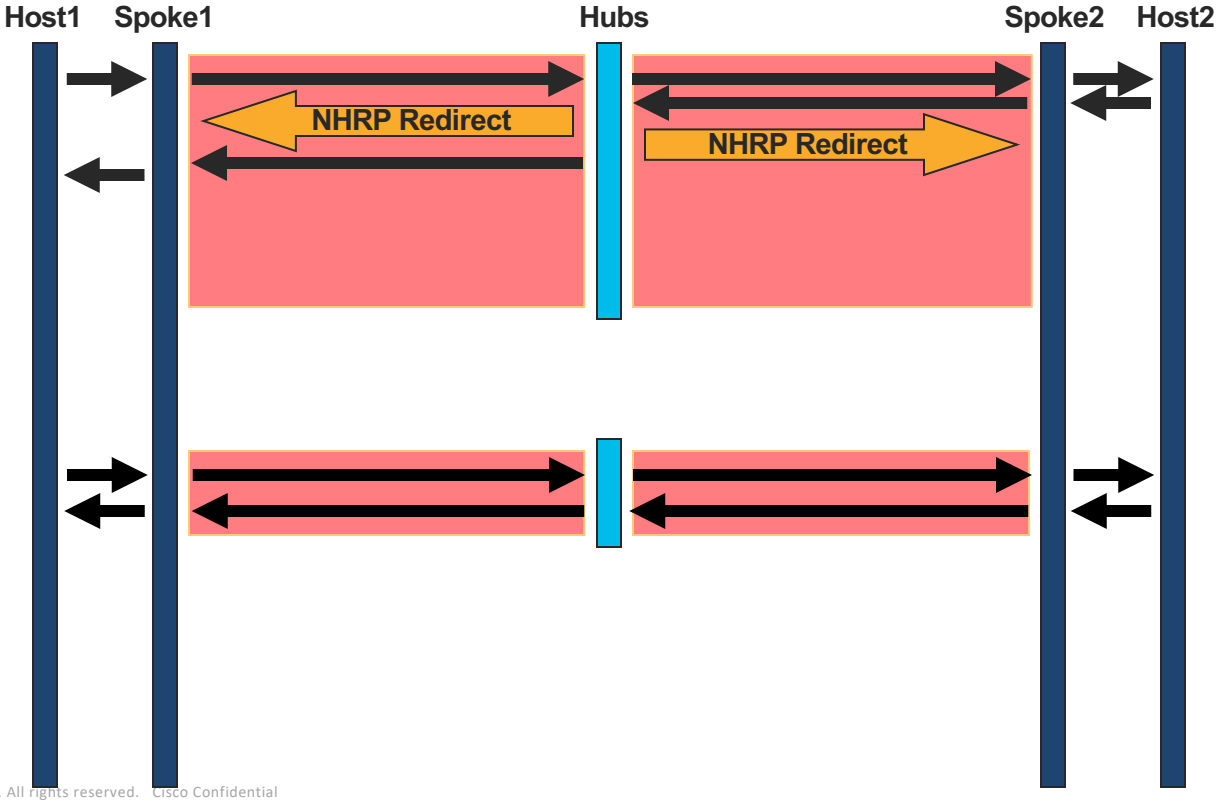
Phase 3 – Features (cont)

- Spokes do not need full routing tables
  - Can summarize routes at the hub
  - Reduced space and load on small spokes
  - Reduced routing protocol load on hub
    - 1000 spokes, 1 route per spoke;
    - hub advertises 1 route to 1000 spokes -> 1000 advertisements
- Not recommended to mix Phase 2 and Phase 3 on same DMVPN
  - Migrate spokes from Phase 2 DMVPN to Phase 3 DMVPN

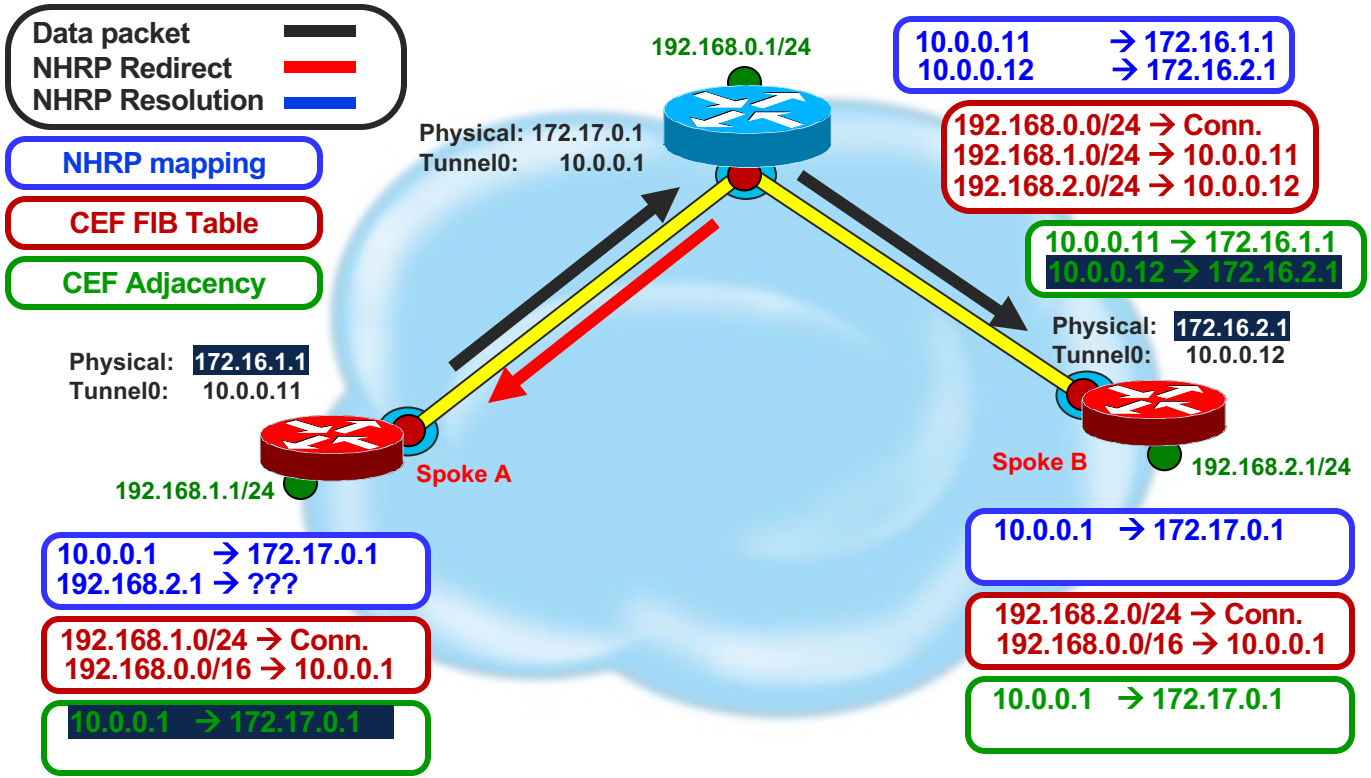
Phase 3
Building Spoke-spoke Tunnels

- 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 via NHS
- Destination spoke
  - Receives NHRP Resolution Request
  - Builds spoke-spoke tunnel
  - Sends NHRP Resolution Reply over spoke-spoke tunnel

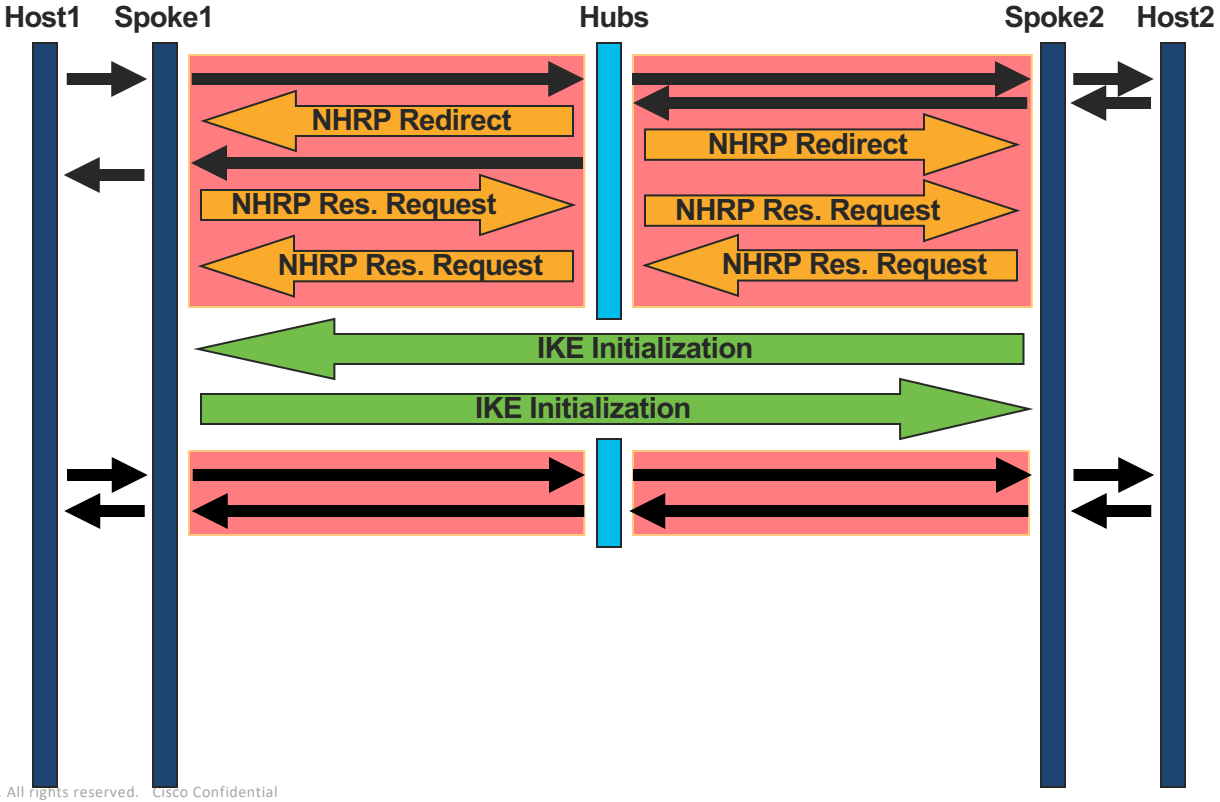
Phase 3
NHRP Redirects



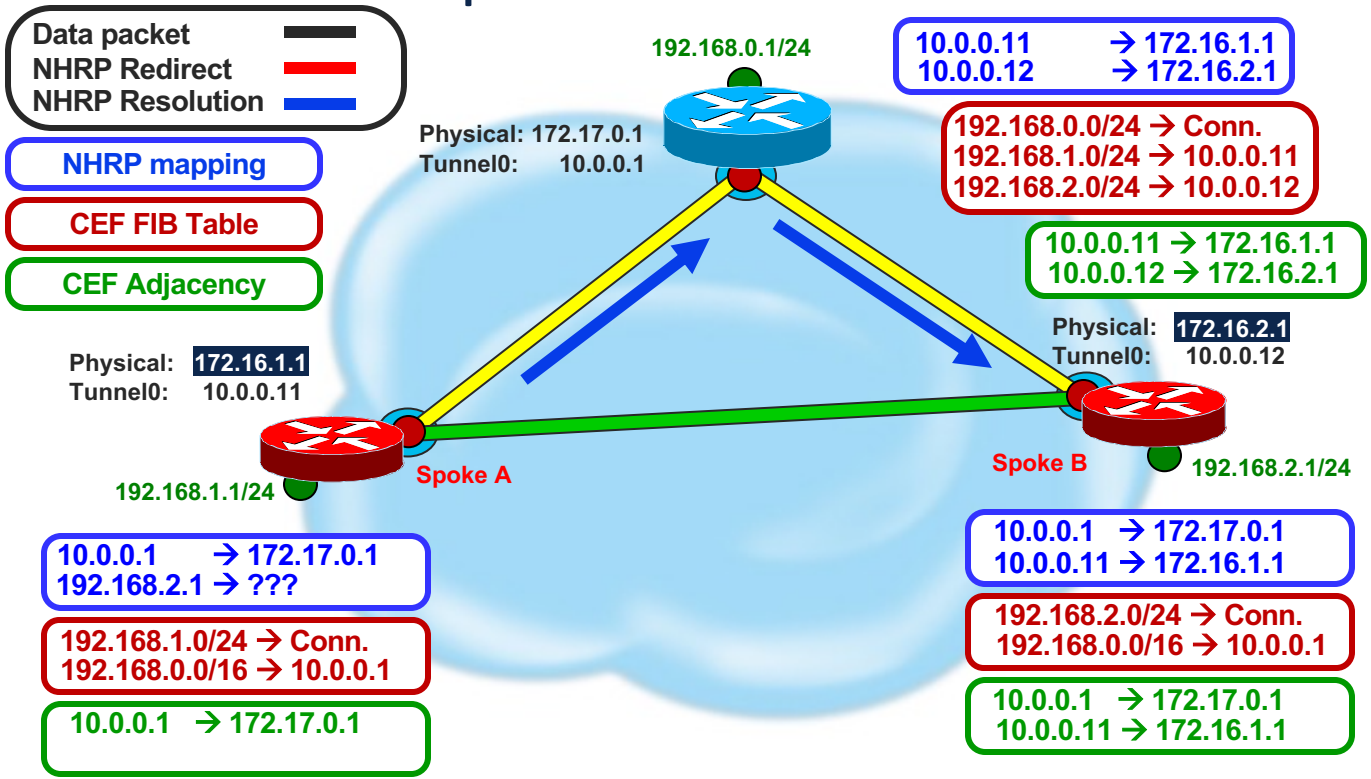
Phase 3
NHRP Redirects



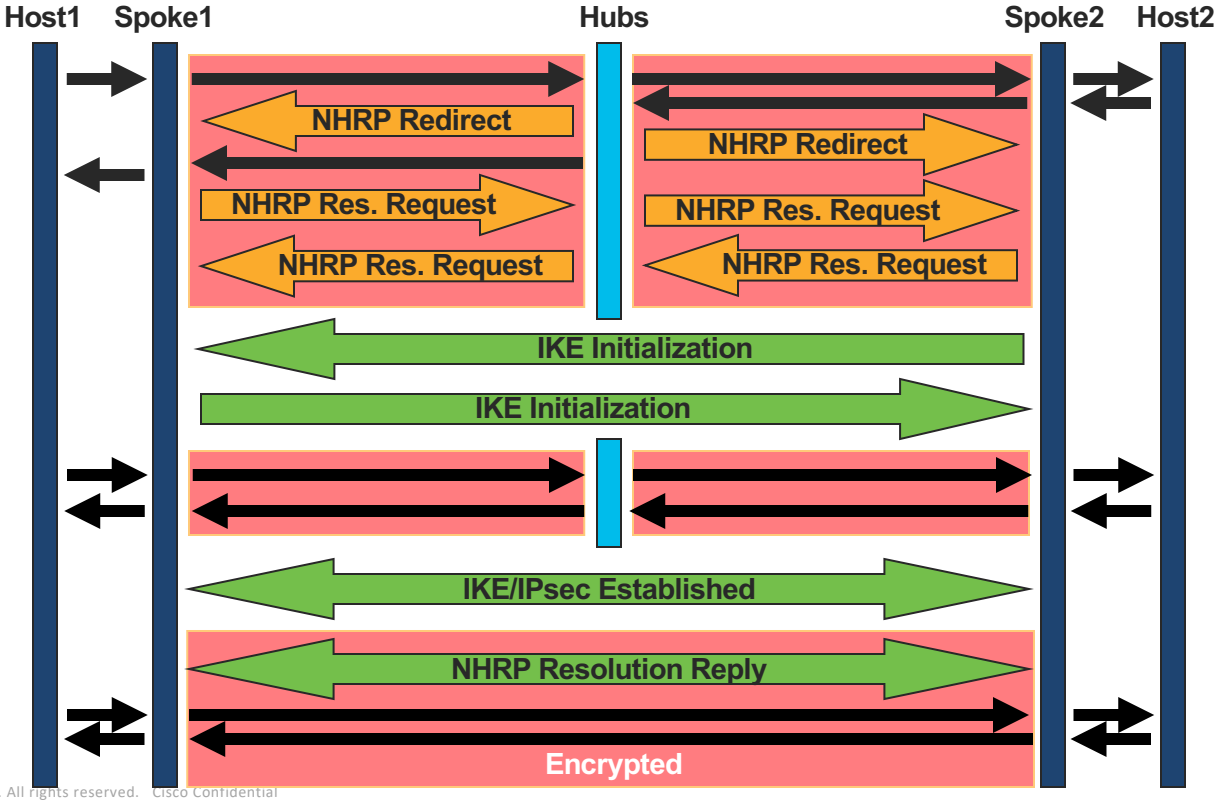
Phase 3
NHRP Resolution Request



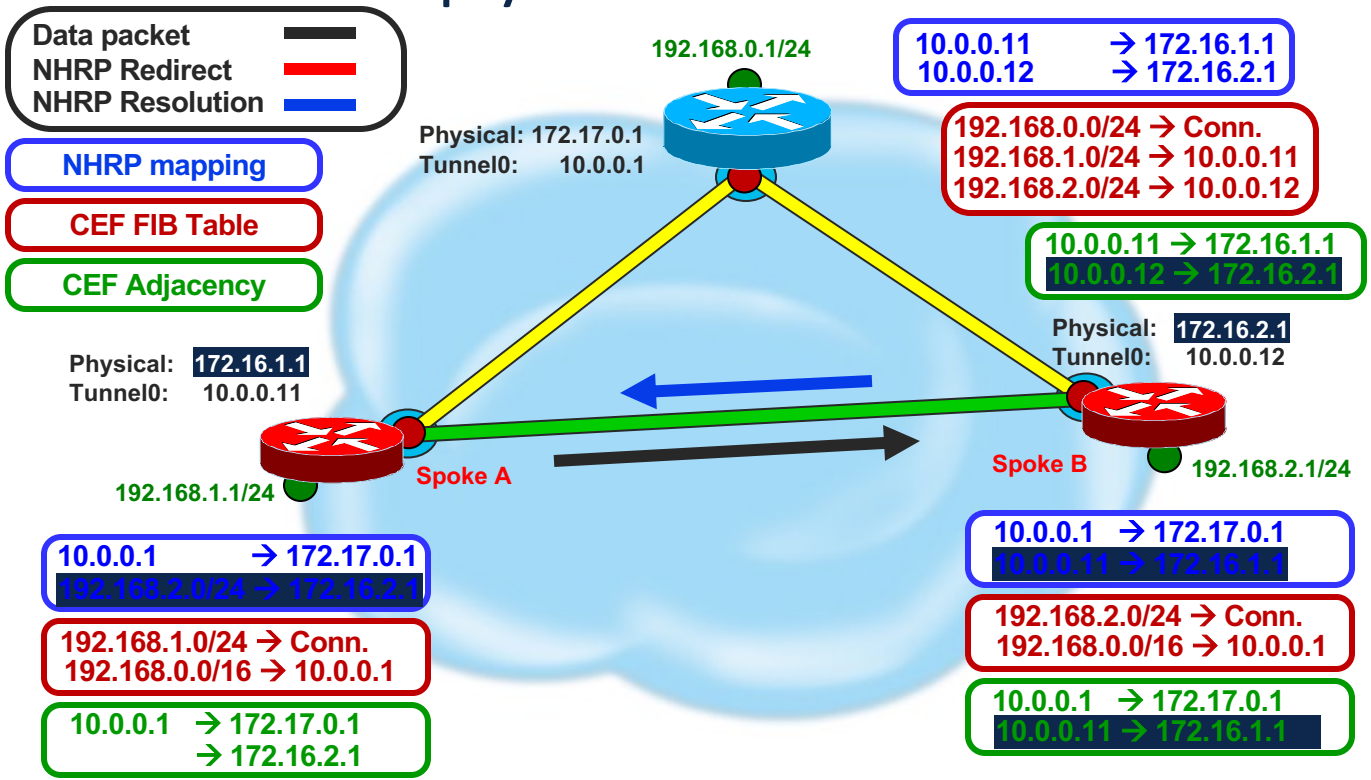
Phase 3
NHRP Resolution Request



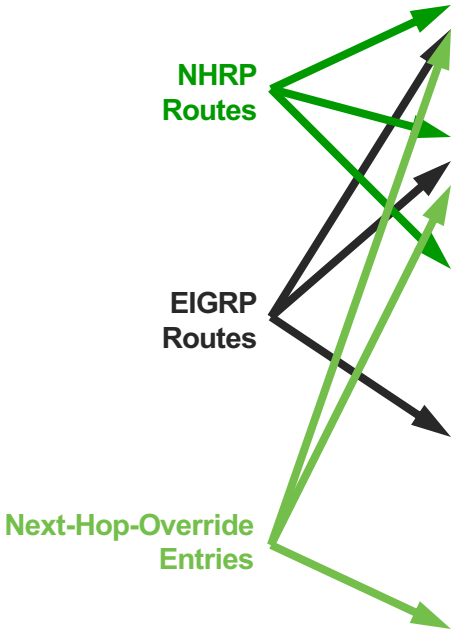
Phase 3
NHRP Resolution Reply



Phase 3
NHRP Resolution Reply



Phase 3: NHRP and RT Routing Table



(ASR1k; 15.2(1)T – ISR, 7200)

```
#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
           [NHO][90/1] via 10.0.0.1, 00:00:40, 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
       ...
```

BFD over DMVPN

16.3.1,
15.6(3)M3, 15.6(2)S

- 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 appli

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

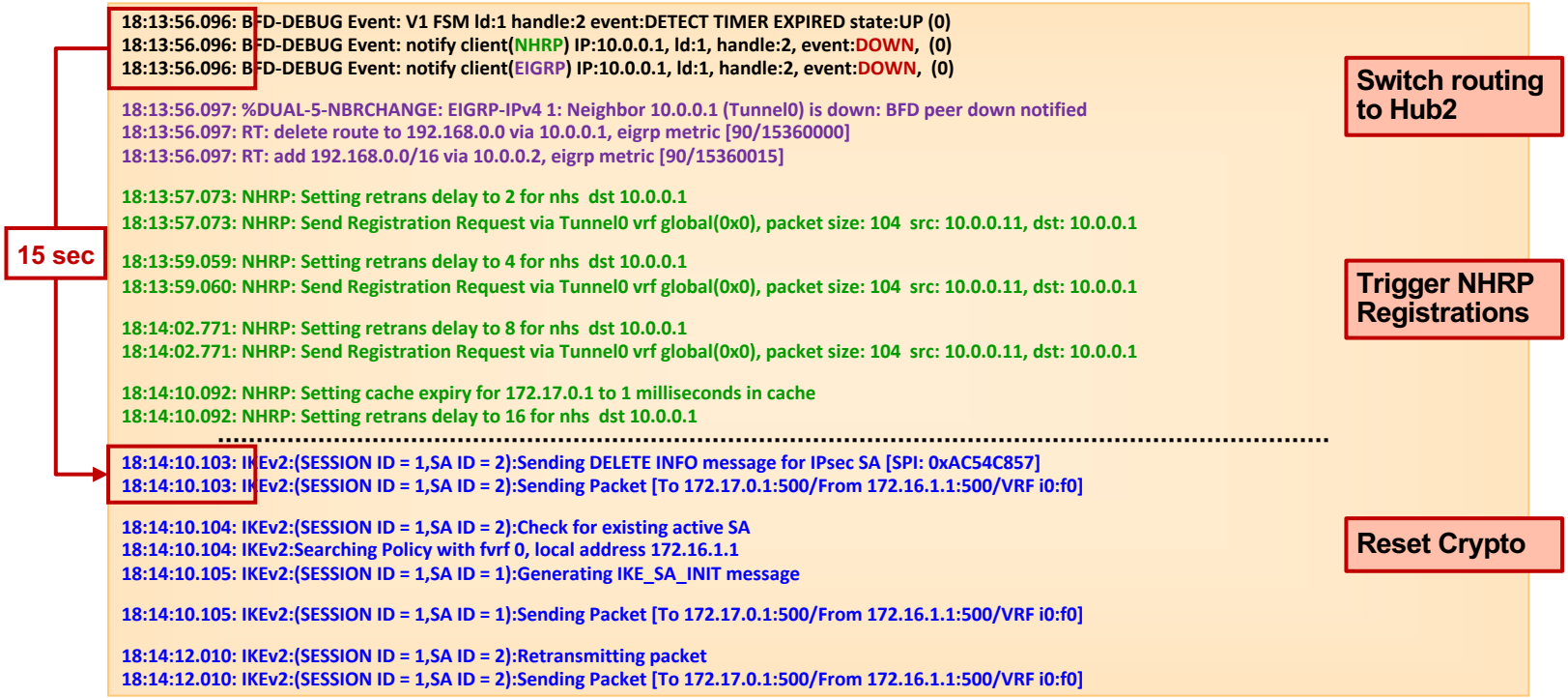
```
bfd-template single-hop DMVPN
 interval min-tx 2000 min-rx 2000 multiplier 3
 echo
!
interface Tunnel0
 ...
 bfd template DMVPN
 ...
```

Echo mode BFD 2/6 second keepalive/hold

Apply on Tunnel interface

BFD over DMVPN

Spoke-Hub tunnel



BFD over DMVPN

Spoke-Spoke tunnel

```
18:46:52.695: NHRP: Receive Traffic Indication via Tunnel0 vrf global(0x0), packet size: 96
18:46:52.705: NHRP: Send Resolution Request for dest: 192.168.12.1 to nexthop: 192.168.12.1 src: 10.0.0.11
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: %BFDFSM-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
```

Normal tunnel down (no data traffic) (10 min)

```
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
19:19:04.632: NHRP: Send Resolution Request for dest: 192.168.12.1 to nexthop: 192.168.12.1 using our src: 10.0.0.11
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
19:19:24.209: BFD-DEBUG Event: notify client(NHRP) IP:10.0.0.12, ld:10, handle:7, event:DOWN,  (0)
19:19:24.211: NHRP: Calling for delete of Tunnel Endpoints (VPN: 10.0.0.12, NBMA: 172.16.2.1)
19:19:24.211: %BFD-6-BFD_SESS_DESTROYED: bfd_session_destroyed,  ld:10 neigh proc:NHRP, handle:7 act

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

Abnormal tunnel down (BFD triggered) (20 sec)

The bridge to possible