Advanced site-to-site IPsec VPN

BRKSEC-3006

Frédéric Detienne

Cisco Networkers 2007

HOUSEKEEPING

- We value your feedback, don't forget to complete your online session evaluations after each session and complete the Overall Conference Evaluation which will be available online from Friday.
- Visit the World of Solutions on Level -01!
- Please remember this is a 'No Smoking' venue!
- Please switch off your mobile phones!
- Please remember to wear your badge at all times including the Party!
- Do you have a question? Feel free to ask them during the Q&A section or write your question on the Question form given to you and hand it to the Room Monitor when you see them holding up the Q&A sign.

Sessions objectives

- We first introduce the latest DMVPN enhancements
- With those, we will target the creation of large IPsec VPN meshes
  - DMVPN is best fit to achieve our goal
  - Enhancements are needed to easily and efficiently scale

Agenda

- Introducing DMVPN phases
- DMVPN phase 2 recollection
  - DMVPN phase 2 resolution forwarding
  - DMVPN phase 2 data forwarding (CEF vs. Process)
- DMVPN phase 3 enhancements
  - Shortcut Switching
  - CEF and process switching resolution
  - NHRP forwarding
  - NAT improvements
  - Troubleshooting enhancements
- Designing with DMVPN phase 3
  - Basic design – passing the 1,000 nodes barrier with a single hub
  - Dual homed design – hub resilience
  - Very large scale DMVPN design – limitless aggregation

DMVPN phases

DMVPN phases

- DMVPN was introduced in IOS 12.2(13)T
  - All the features were present but it suffered from bugs
  - Cisco recommended hub&spoke topologies only
  - This was DMVPN phase 1
- With 12.3(1) Mainline, spoke-to-spoke was stable
  - Cisco gave green light to dynamic meshes
  - This was DMVPN phase 2

DMVPN phase 2

- DMVPN phase 2 applies to

  12.3 mainline; i.e. 12.3(1) → 12.3(…)

  12.4 mainline; i.e. 12.4(1) → 12.4(…)

  12.4(1) T → 12.4(4) T

- DMVPN phase 2 has been widely deployed

  Successfully: networks up to 13,000 nodes were deployed

  Both in hub & spoke and spoke-to-spoke

DMVPN phase 2 “recollection”

- Phase 2 shows discrepancies in the switching paths
  - CEF and process do not work the same at all
- Multi-hubs are difficult to configure
  - Resolutions follow the NHS path
  - → Need for hub daisy chaining → tricky
- Routing protocols are difficult to scale
  - No way to summarize in spoke-to-spoke mode
  - → can’t redistribute into other protocols at hubs
- DMVPN phase 2 worked well but
  - There was room for improvement

DMVPN phase 3

- From 12.4(6)T onward, a series of features appeared that facilitate DMVPN scalability
  - More natural CEF vs. process comparisons
  - Resolution request forwarding changes
  - CEF Shortcut Switching
  - NAT handling improvements
- 12.4(9)T also introduces troubleshooting aids

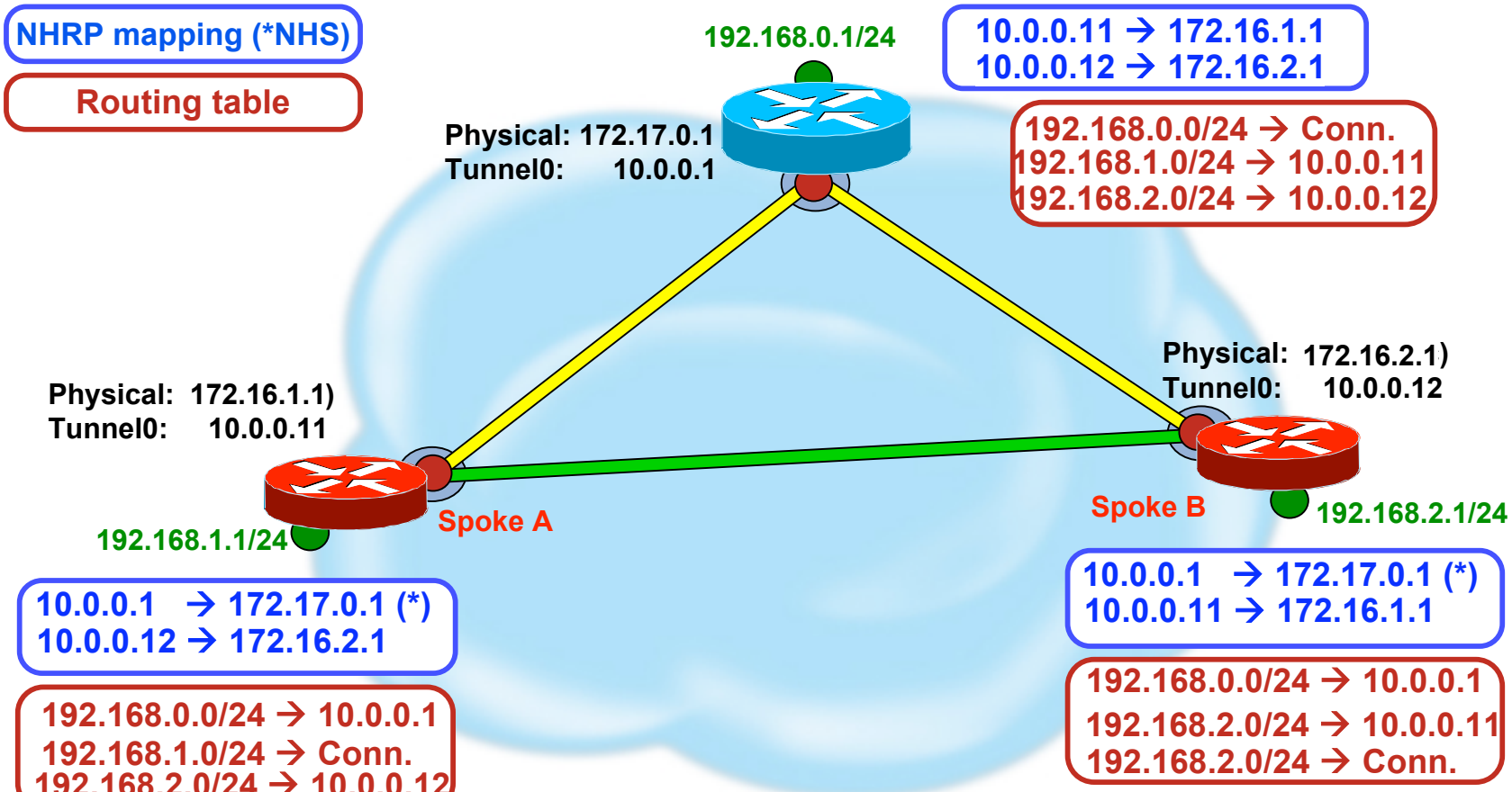
Phase 3 focuses on

scalability and maintainability of dynamic meshes

DMVPN phase 2 – data forwarding CEF vs process

DMVPN phase 2 – Design Style

- In DMVPN phase 2 the spoke routing table determines when to build spoke-to-spoke tunnels



Dynamic Spoke-Spoke Tunnels
Phase 2 data forwarding

- DMVPN phase 2 behaves differently when in CEF switching or Process switching
- CEF switching

  IP Next-hop from routing table

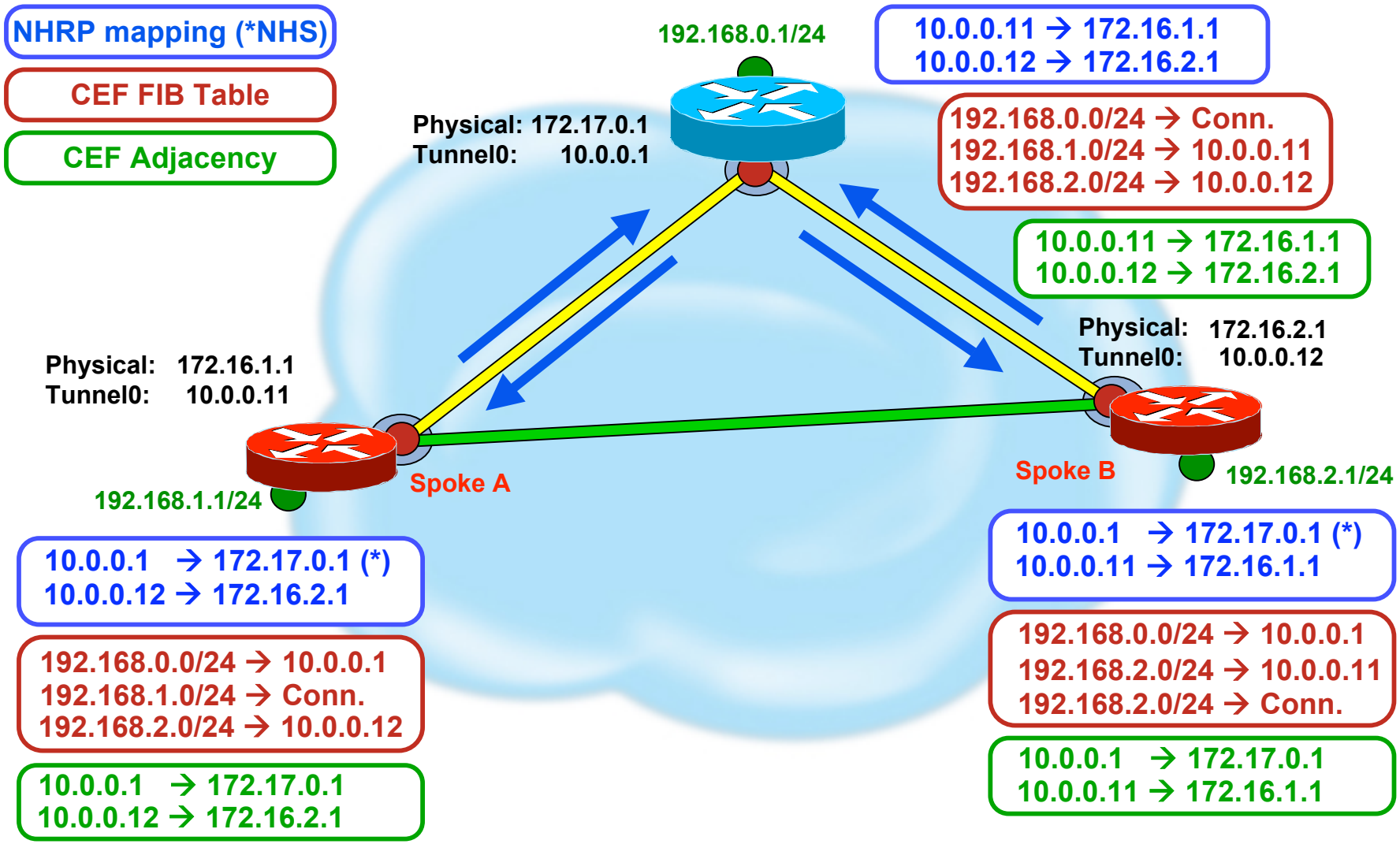
  Next-hop → hub ➔ data packets via hub

  Next-hop → spoke ➔ data packets direct
- Process-switching

  Routing selects outgoing interface and IP next-hop

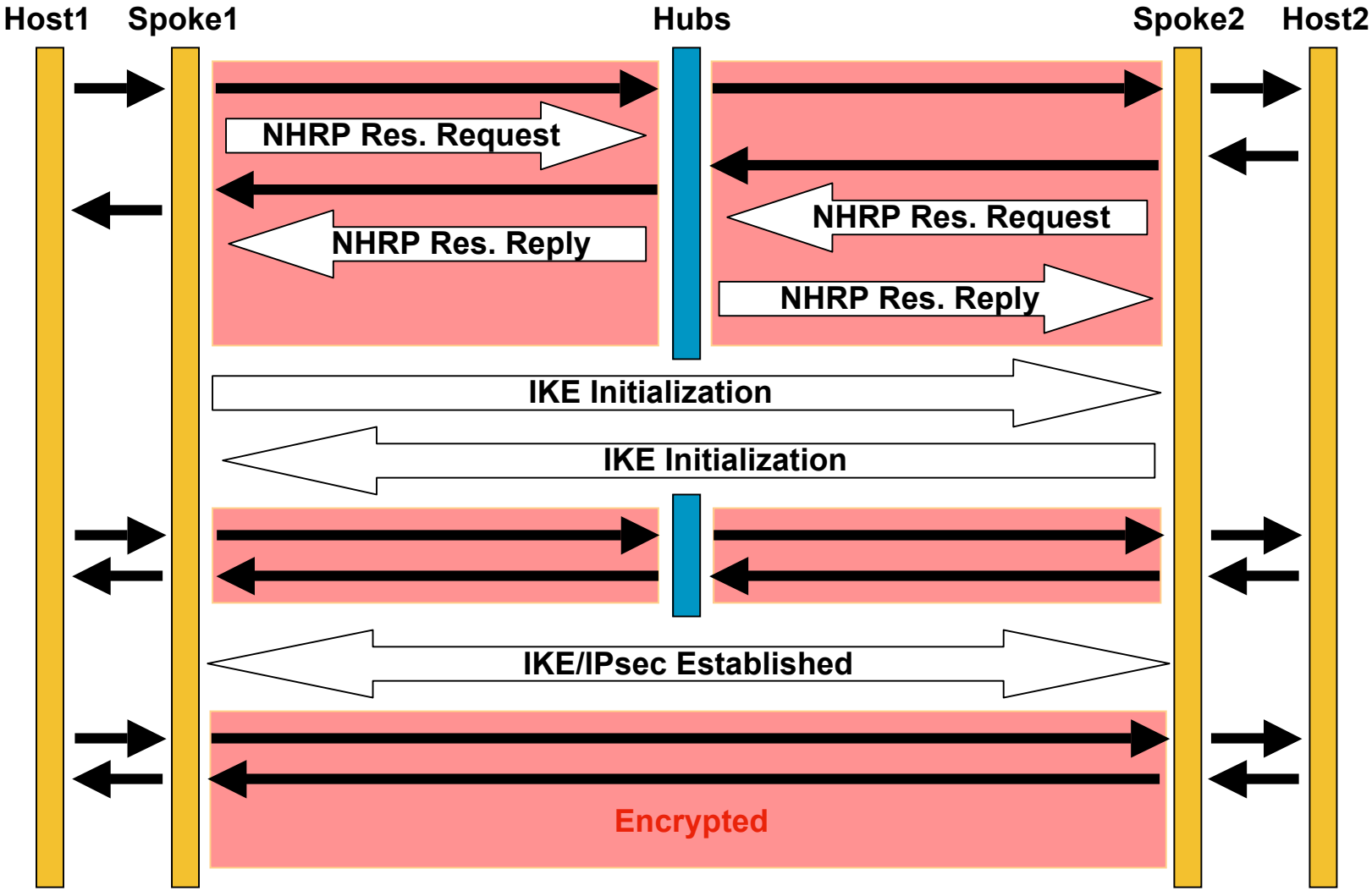
  NHRP overrides IP next-hop from routing by snooping the destination address of the packet
- Data packets via hub while spoke-spoke tunnel is coming up, then direct. Temporary punting to the hub is done in process switching

NHRP Resolution
CEF Switching

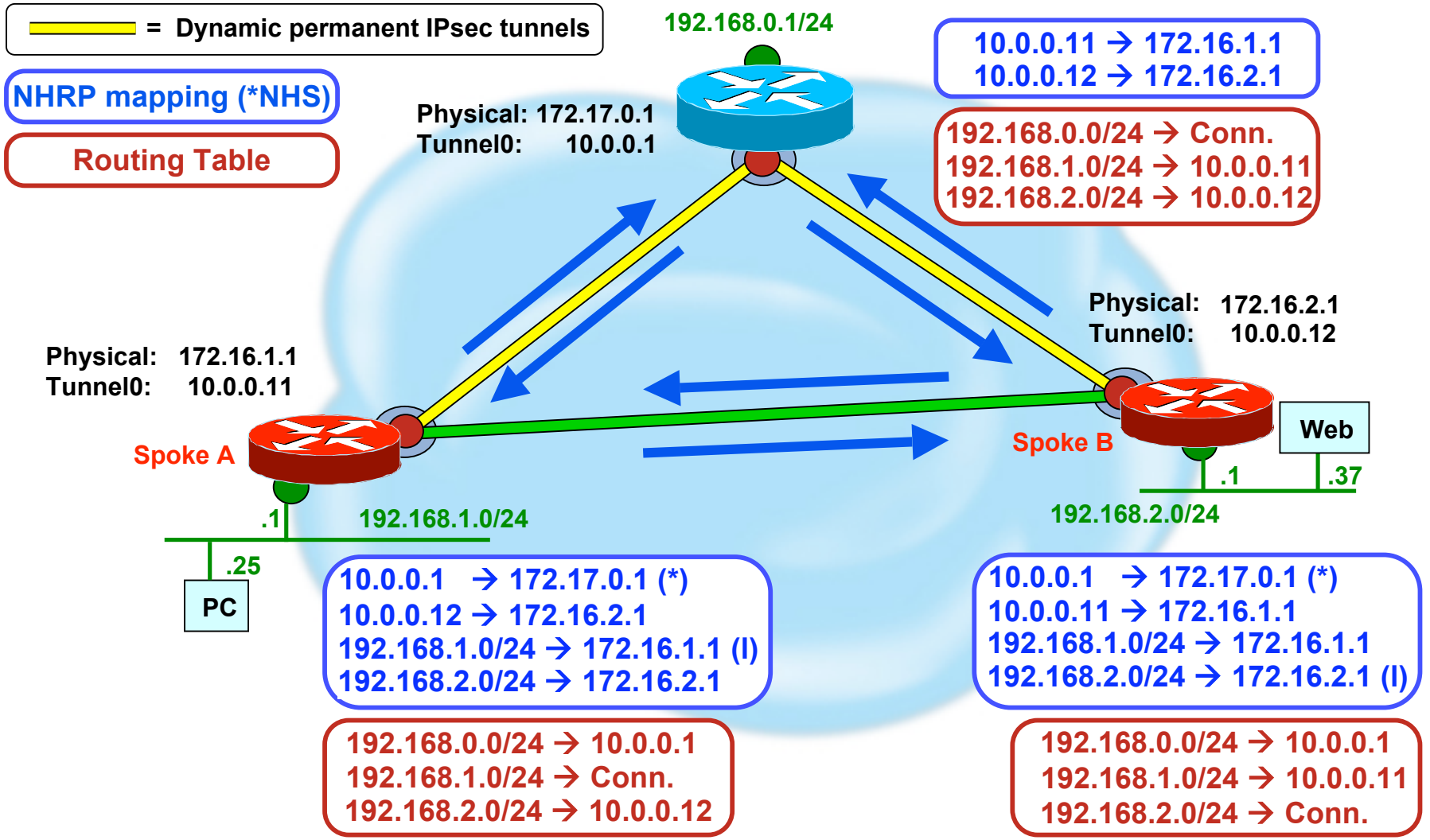


Building Spoke-Spoke tunnels CEF Switching

For your reference

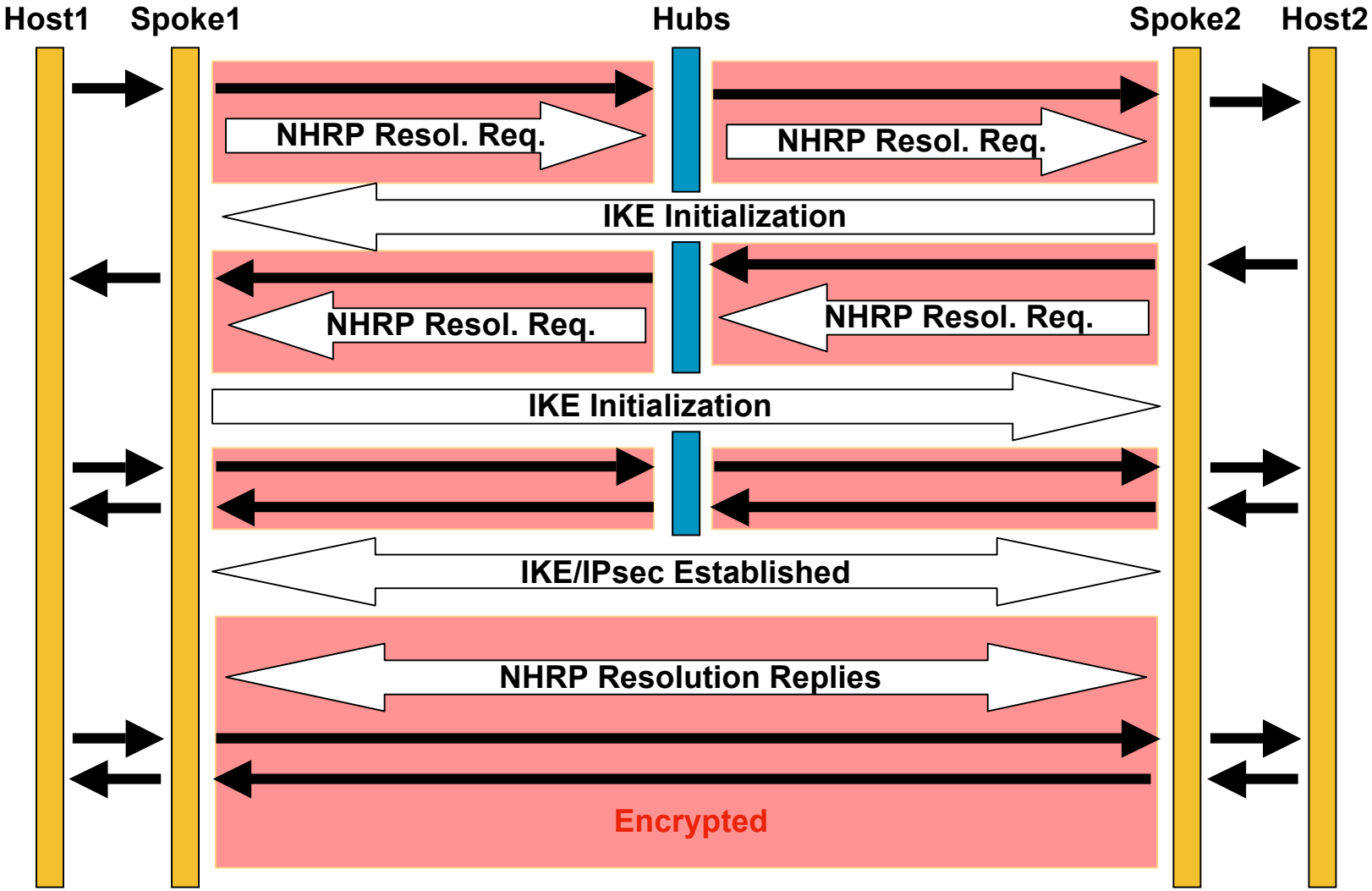


NHRP Resolution Process Switching



Building Spoke-Spoke tunnels Process Switching

For your reference



DMVPN phase 2 – Multi hubs

DMVPN phase 2
NHRP resolution forwarding

- A unique hub may not be able to aggregate all spokes

  Large environments require multiple hub

- To build spoke-to-spoke tunnels, a spokes need the NBMA address of the remote spoke

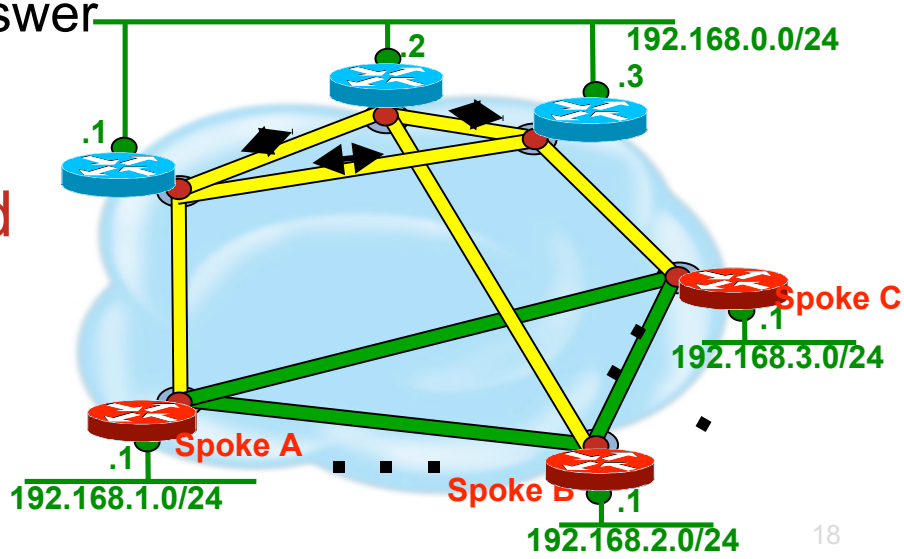
  ➔ NHRP resolution requests are sent to the NHS

- The target spoke may not be known on the hub

  An other hub may know the answer

- Resolution requests

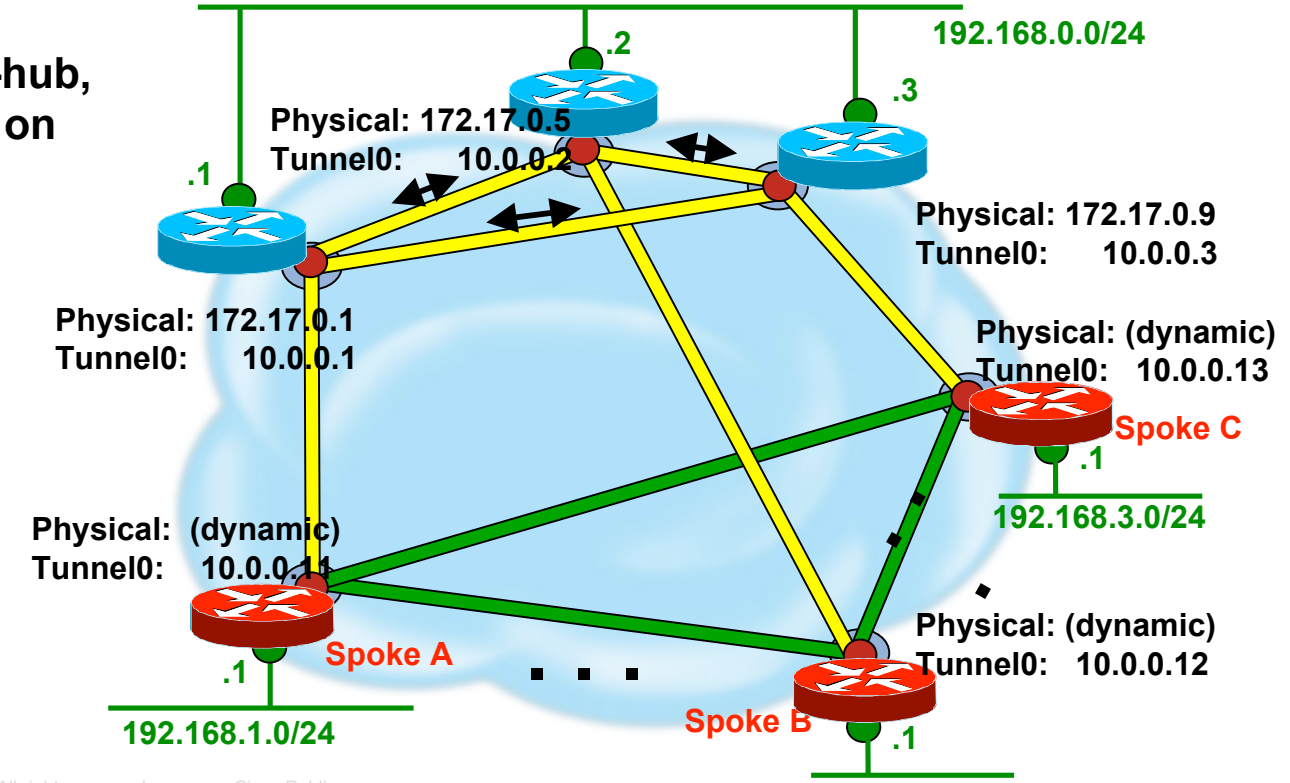
  must be forwarded



DMVPN phase 2
NHRP resolution forwarding (cont.)

- Resolution requests are forwarded from NHS to NHS
- Hub routers must point to other hub routers as NHSs in a “daisy-chain” fashion

Single DMVPN Multi-hub, Single mGRE tunnel on all nodes



DMVPN phase 2 Multi-Hub Hub Daisy Chaining

For your reference

- Single daisy chain through all hubs

  Loss of Hub breaks daisy chain

  ip nhrp nhs <hub<x+1>>

- Two layer daisy chain

  Can lose of every other Hub without splitting DMVPN network.

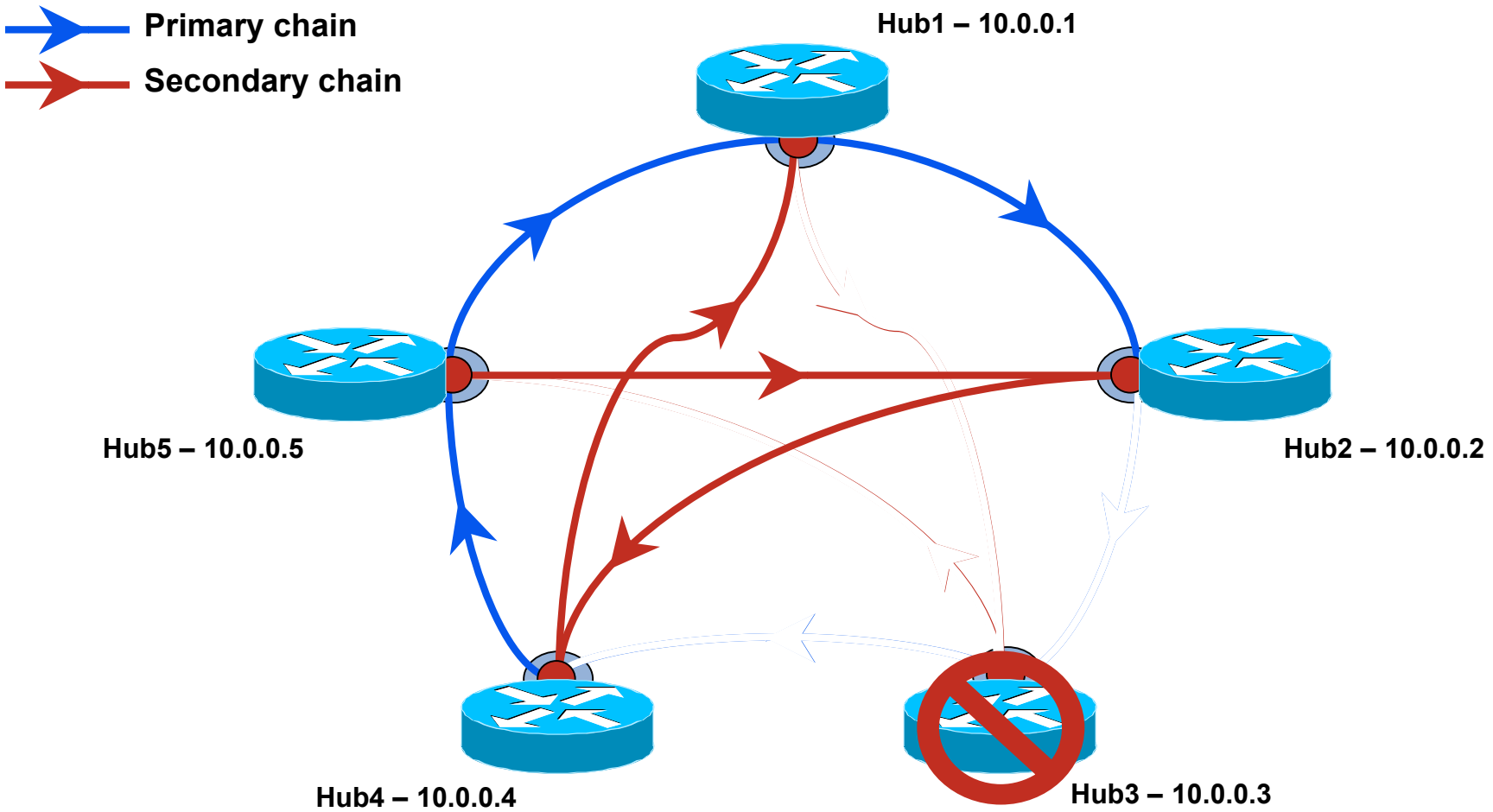
  ip nhrp nhs <hub<x+1>>
  ip nhrp nhs <hub<x+2>>

- Three layer daisy chain

  Can handle losing more hubs, but greater complexity

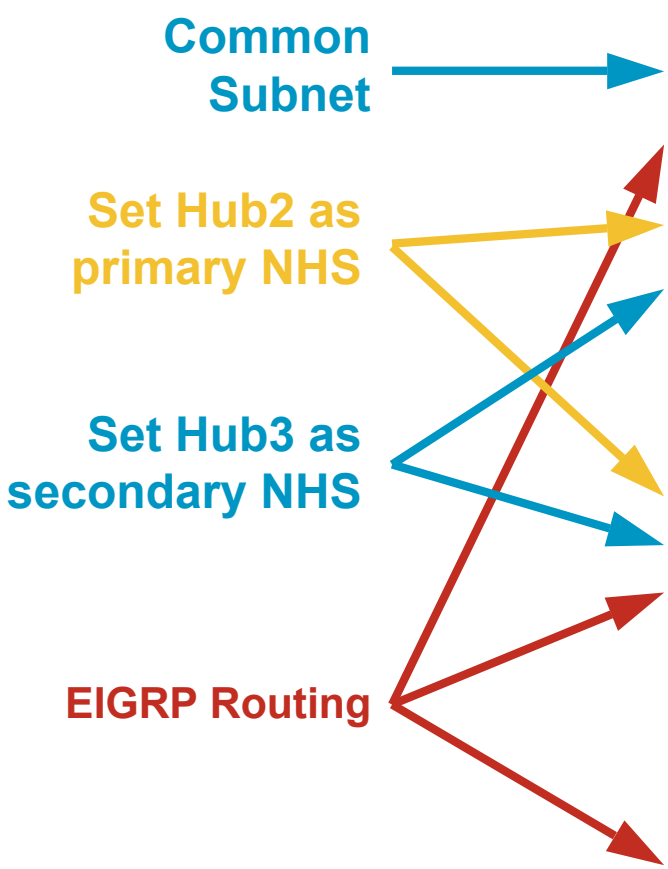
  ip nhrp nhs <hub<x+1>>
  ip nhrp nhs <hub<x+2>>
  ip nhrp nhs <hub<x+3>>

DMVPN Phase 2 Multi-Hub Hub Daisy Chaining (cont.)



DMVPN phase 2 Multi-Hub Hub 1

For your reference




```
interface Tunnel0
      bandwidth 1000
      ip address 10.0.0.1 255.255.255.0
      ip mtu 1400
      no ip next-hop-self eigrp 1
      ip nhrp authentication test
      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 multicast dynamic
      ip nhrp network-id 100000
      ip nhrp holdtime 360
      ip tcp adjust-mss 1360
      ip nhrp nhs 10.0.0.2
      ip nhrp nhs 10.0.0.3
      no ip split-horizon eigrp 1
      ip tcp adjust-mss 1360
      delay 1000
      tunnel source Serial1/0
      tunnel mode gre multipoint
      tunnel key 100000
      tunnel protection ipsec profile vpnprof
!
router eigrp 1
      network 10.0.0.0 0.0.0.255
      network 192.168.0.0
      no auto-summary
```

DMVPN phase 2 Multi-Hub Hub 2 and Hub 3

For your reference

Hub 2

Common Subnet

Set Hub3 as primary NHS

Set Hub1 as secondary NHS

```
interface Tunnel0
     ip address 10.0.0.2 255.255.255.0
     …
     ip nhrp map 10.0.0.3 172.17.0.9
     ip nhrp map multicast 172.17.0.9
     ip nhrp map 10.0.0.1 172.17.0.1
     ip nhrp map multicast 172.17.0.1
     …
     ip nhrp nhs 10.0.0.3
     ip nhrp nhs 10.0.0.1
     …
```

Hub 3

Common Subnet

Set Hub1 as primary NHS

Set Hub2 as secondary NHS

```
interface Tunnel0
     ip address 10.0.0.3 255.255.255.0
     …
     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 nhs 10.0.0.1
     ip nhrp nhs 10.0.0.2
     …
```

DMVPN phase 2 Multi-Hub Spoke A

For your reference

Hub1 as NHS

EIGRP Routing

```
interface Tunnel0
     bandwidth 1000
     ip address 10.0.0.11 255.255.255.0
     ip mtu 1400
     ip nhrp authentication test
     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 360
     ip nhrp nhs 10.0.0.1
     ip tcp adjust-mss 1360
     delay 1000
     tunnel source Serial1/0
     tunnel mode gre multipoint
     tunnel key 100000
     tunnel protection ipsec profile vpnprof
!
router eigrp 1
     network 10.0.0.0 0.0.0.255
     network 192.168.1.0 0.0.0.255
     no auto-summary
```

DMVPN phase 2 Multi-Hub Spoke B and Spoke C

For your reference

Spoke B

Common Subnet

Hub2 as primary NHS

EIGRP Routing

```
interface Tunnel0
     ip address 10.0.0.12 255.255.255.0
     …
     ip nhrp map 10.0.0.2 172.17.0.5
     ip nhrp map multicast 172.17.0.5
     …
     ip nhrp nhs 10.0.0.2
…
     network 192.168.2.0 0.0.0.255
```

Spoke C

Common Subnet

Hub3 as primary NHS

EIGRP Routing

```
interface Tunnel0
     ip address 10.0.0.13 255.255.255.0
     …
     ip nhrp map 10.0.0.3 172.17.0.9
     ip nhrp map multicast 172.17.0.9
     …
     ip nhrp nhs 10.0.0.3
…
     network 192.168.3.0 0.0.0.255
```

DMVPN phase 2
Dynamic Mesh Routing summary

- Spokes are routing neighbors with hubs
  - Spokes advertise local network to hubs
  - Spokes are not routing neighbors with each other
- Hubs are routing neighbors with spokes
  - Hubs advertise spoke and local networks to all spokes
    - Turn off split-horizon (EIGRP, RIP)
    - Hub must preserve original IP next-hop
      - EIGRP (no ip next-hop-self), OSPF (network broadcast)
    - Cannot summarize
    - OSPF: Single area, no summarization and only two hubs
- Hubs must be routing neighbors with other hubs over the same DMVPN network
  - Must use the same routing protocol between hubs as is used between hubs and spokes (same AS, domain, etc.)

Dynamic Routing in spoke-to-spoke environments (Configuration)

For your reference

- EIGRP

```
no ip next-hop-self eigrp <as>          (on hub – tunnel interface)
no ip split-horizon eigrp <as>
router eigrp <as>
  no auto-summary                       (cannot summarize on hubs)
```

- OSPF

```
ip ospf network broadcast (on hub & spoke – tunnel interface)
ip ospf priority [2 (hub) | 0 (spoke)]
```

- BGP (Hub is route-reflector)

```
router bgp <AS>
  neighbor <spoke-tunnel-ip> <AS>       (on hub – one for each spoke)
  neighbor <spoke-tunnel-ip> route-reflector-client
  no next-hop self
```

Dynamic Routing Scaling

- 1 mGRE interface/hub/DMVPN
- Must use same routing protocol for spoke-hub and hub-hub neighbors
- EIGRP

  350 Spokes/Hub
- OSPF

  400 Spokes/Hub

  Single OSPF Area/DMVPN

  Maximum of 2 Hubs
- BGP

  1000 Spokes/Hub

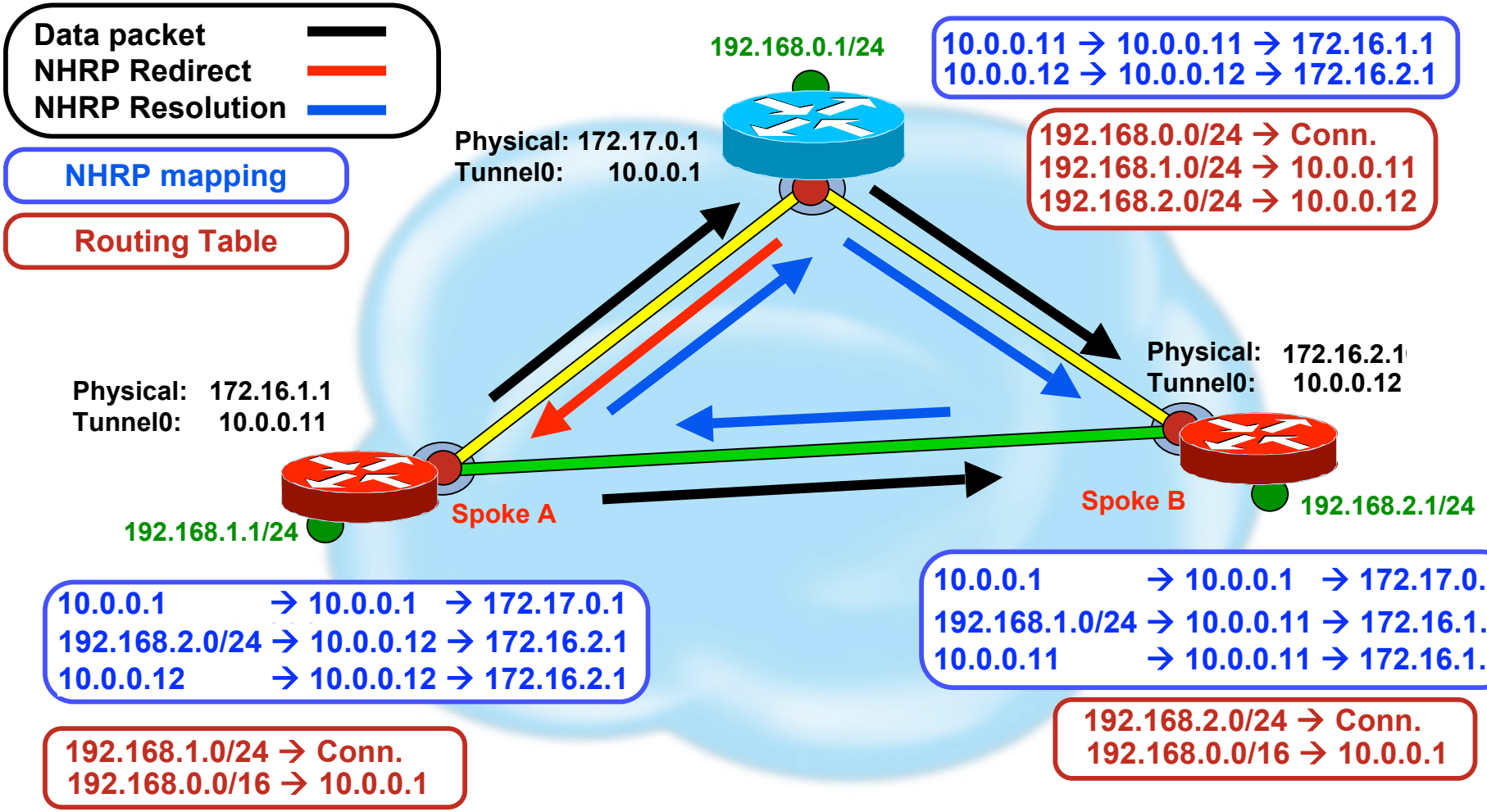
All testing done under clean lab conditions, you may not be able to get to these numbers in real world conditions (Internet).

Phase 3 – Shortcut Switching

Introducing Shortcut Switching

- Spokes can now have a summary route
  - Supernet of private networks → hub
- All the packets are initially sent to the hub
- The hub routes-back into the DMVPN
- … then sends and NHRP indirection message back
  - New message type introduced for the purpose
- The spoke receives the indirection message and resolves the private network NBMA address
- A spoke-spoke direct tunnel is created

DMVPN phase 3 – Design Style



DMVPN Shortcut Switching Hub configuration revisited

For your reference

Common Subnet

Turn on Redirects

ODR Routing

Prevent recursive routing

```
interface Tunnel0
      bandwidth 1000
      ip address 10.0.0.1 255.255.255.0
      ip mtu 1400
      ip nhrp map multicast dynamic
      ip nhrp network-id 100000
      ip nhrp redirect
      tunnel source Serial1/0
      tunnel mode gre multipoint
      tunnel protection ipsec profile vpnprof
      cdp enable
!
interface serial1/0
      ip address 172.17.0.1 255.255.255.252
!
router odr
      distribute-list 1 in

access-list 1 permit 192.168.0.0 0.0.255.255
```

DMVPN Shortcut Switching
Spoke configuration revisited

For your reference

Common Subnet

Hub as NHS

Allow shortcuts

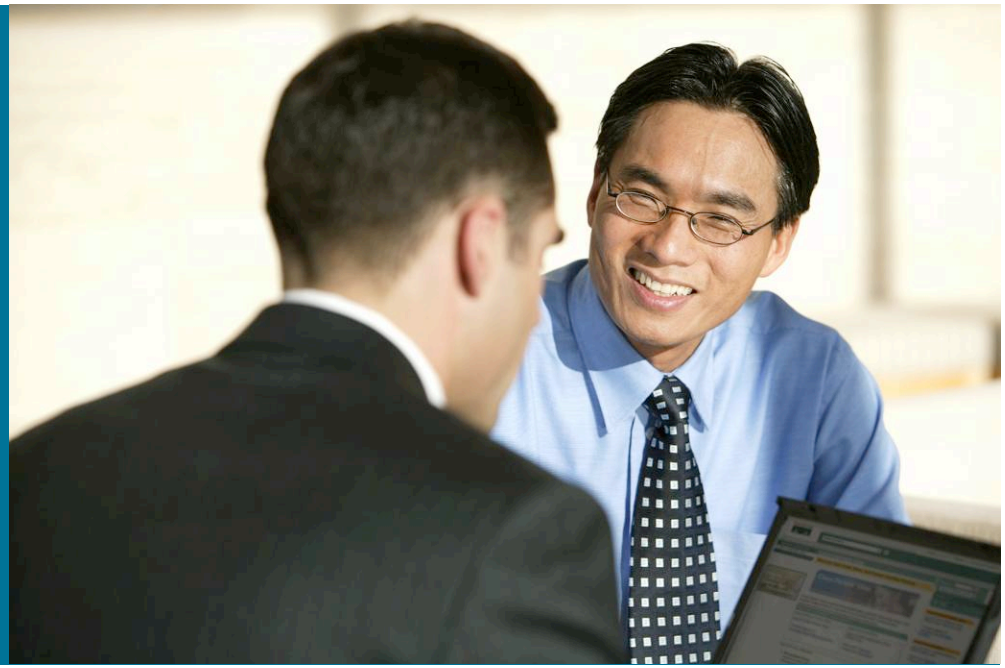
ODR Routing

Private network summary → hub

ESP Packets → ISP

```
interface Tunnel0
      bandwidth 1000
      ip address 10.0.0.11 255.255.255.0
      ip mtu 1400
      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 nhs 10.0.0.1
      ip nhrp shortcut
      ip tcp adjust-mss 1360
      tunnel source Serial1/0
      tunnel mode gre multipoint
      tunnel protection ipsec profile vpnprof
      cdp enable
!
interface Ethernet0/0
      ip address 192.168.11.1 255.255.255.0

ip route 192.168.0.0 255.255.0.0 10.0.0.1


ip route 0.0.0.0 0.0.0.0 Serial1/0
```

DMVP Phase 3 Forwarding enhancements



DMVPN Phase 3
next-hop resolution and forwarding

- DMVPN phase 3 always tries to resolve for next-hop
- No difference between CEF and Process Switching

  Shortcut switching, normal forwarding, resolutions… all the same!

- During a resolution request, packets are forwarded to the hub
- Phase 3 design style ➔ packets are CEF switched during resolution

Resolution Requests Forwarding Algorithm 12.4(6)T – Phase 3

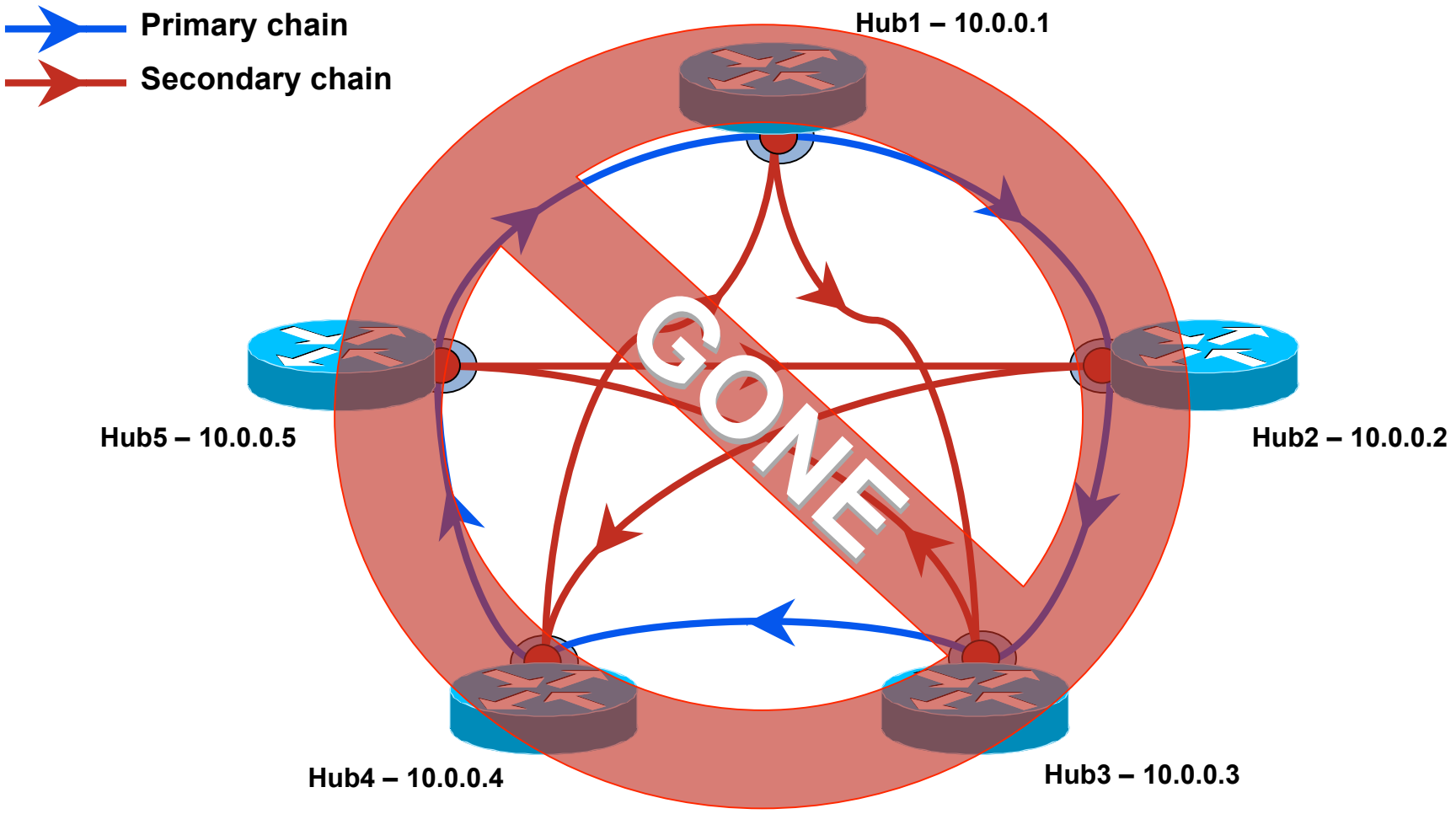
- Router receives a resolution request for address X
- Route lookup performed for X

  next-hop and output interface are determined
- If output interface is not in the DMVPN

  → reply to resolution with our NBMA/tunnel addresses
- If output interface is in the DMVPN

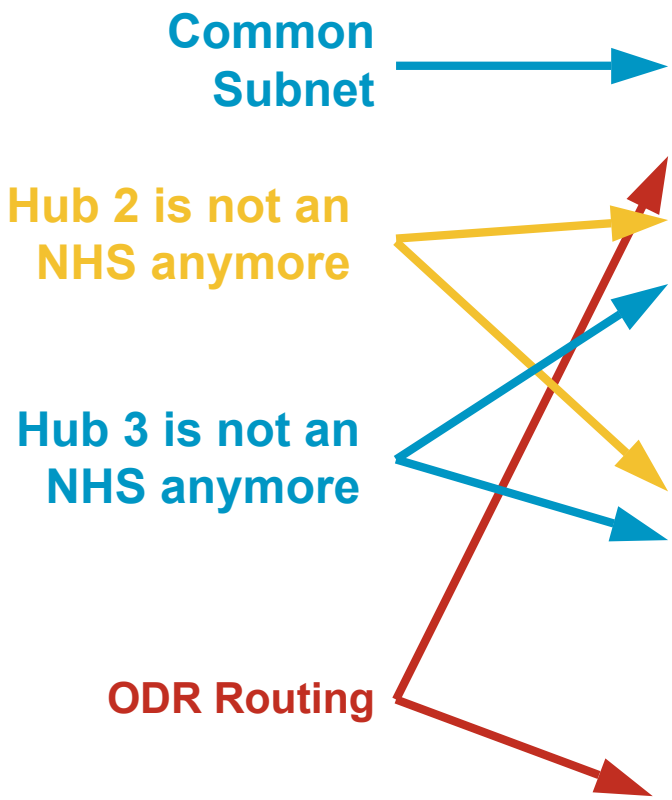
  → forward the request to next-hop as if it was an NHS

Spoke-to-Spoke (MH) Phase 3
Hub Daisy Chaining revisited



DMVPN Multi-Hub Phase 3
Hub1 revisited

For your reference




```
interface Tunnel0
      bandwidth 1000
      ip address 10.0.0.1 255.255.255.0
      ip mtu 1400
      cdp enable
      ip nhrp authentication test
      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 multicast dynamic
      ip nhrp network-id 100000
      ip nhrp holdtime 360
      ip tcp adjust-mss 1360
      ip nhrp nhs 10.0.0.2
      ip nhrp nhs 10.0.0.3
      tunnel source Serial1/0
      tunnel mode gre multipoint
      tunnel protection ipsec profile vpnprof
!
router odr
      distribute-list 1 in

access-list 1 permit 192.168.0.0 0.0.255.255
```

DMVPN Multi-Hub Phase 3
Hub2 and Hub3 revisited

For your reference

Hub 2

Common Subnet

Hub 3 is not an NHS anymore

Hub 1 is not an NHS anymore

```
interface Tunnel0
    ip address 10.0.0.2 255.255.255.0
    …
    ip nhrp map 10.0.0.3 172.17.0.9
    ip nhrp map multicast 172.17.0.9
    ip nhrp map 10.0.0.1 172.17.0.1
    ip nhrp map multicast 172.17.0.1
    …
    ip nhrp nhs 10.0.0.3
    ip nhrp nhs 10.0.0.1
    …
```

Hub 3

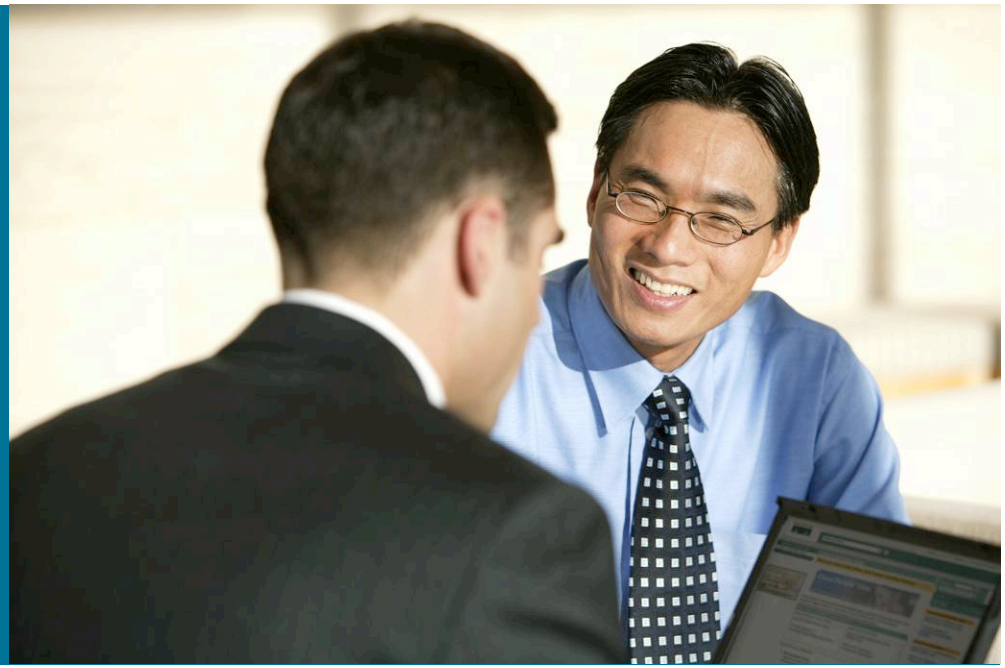
Common Subnet

Hub 1 is not an NHS anymore

Hub 2 is not an NHS anymore

```
interface Tunnel0
    ip address 10.0.0.3 255.255.255.0
    …
    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 nhs 10.0.0.1
    ip nhrp nhs 10.0.0.2
    …
```

DMVPN and NAT



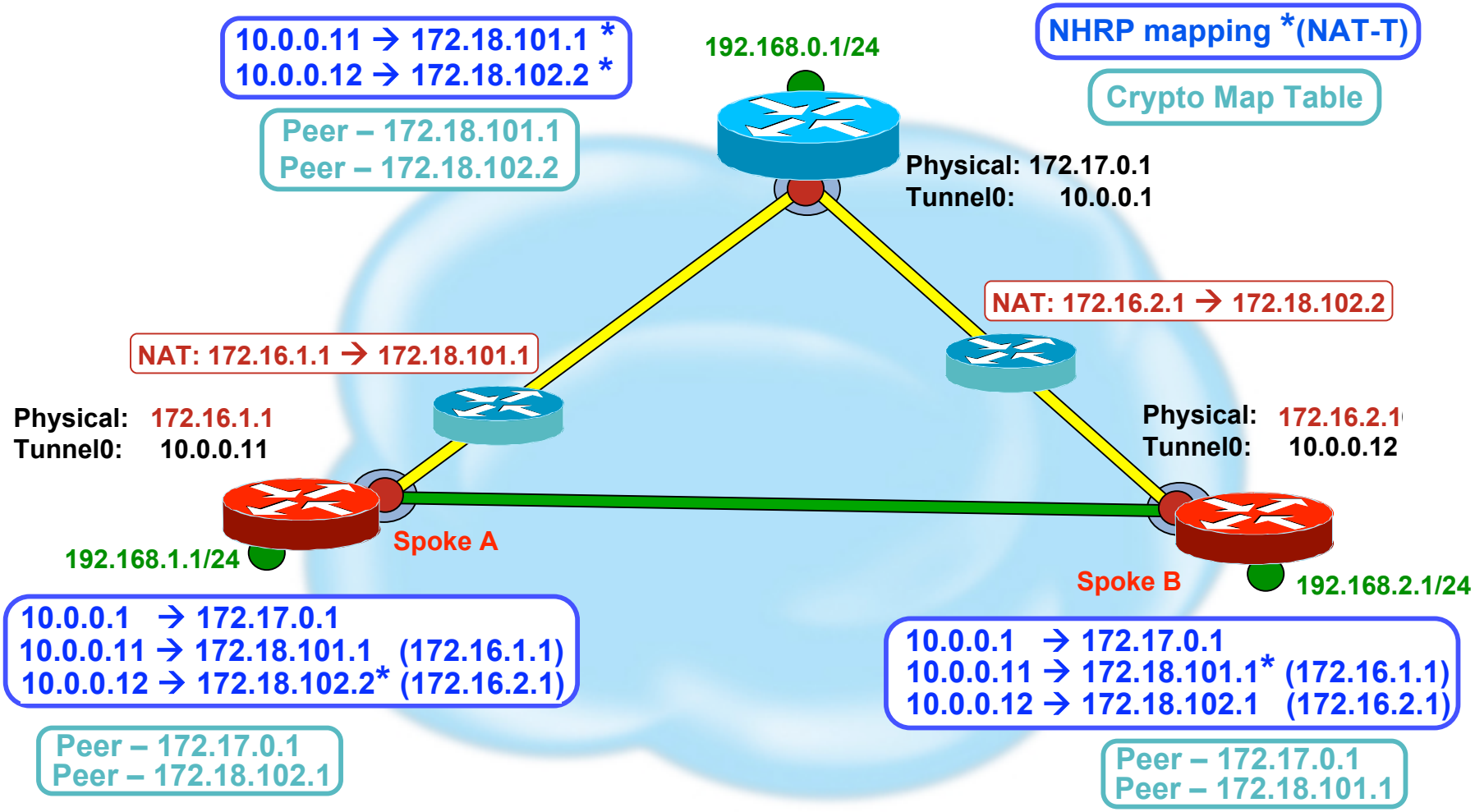
DMVPN NAT Enhancements 12.4(6)T

- Spoke-spoke dynamic tunnels are now supported to/from NAT translated spokes

  Hub reports spoke's outside NAT IP address back to spoke in NHRP registration reply.

- Spoke outside NAT IP address passed in NHRP resolution request and reply packets
- Spokes use remote spoke outside NAT IP address to build spoke-to-spoke tunnel.

DMVPN NAT Enhancements 12.4(6)T (Cont.)



DMVPN NAT
General remarks

- Two spokes behind the same NAT node must be NAT translated to unique outside NAT IP address
  - In general, this really means NAT, not PAT !
  - This is true for hub&spoke and spoke-to-spoke
- If spoke-spoke tunnel will not come up, traffic will continue to be forwarded via the hub.

Troubleshooting Aids

DMVPN
Data Structures

- NHRP Mapping Table

  Maps VPN and Tunnel IP addresses to NBMA (Physical address)

  show ip nhrp { brief | <address> },
  debug nhrp { packet | cache | extension }

- Crypto Socket Table

  Mapping between NHRP and IPsec

  show crypto socket, debug crypto socket,
  show crypto ipsec profile, debug tunnel { protection }

- Crypto Map Table

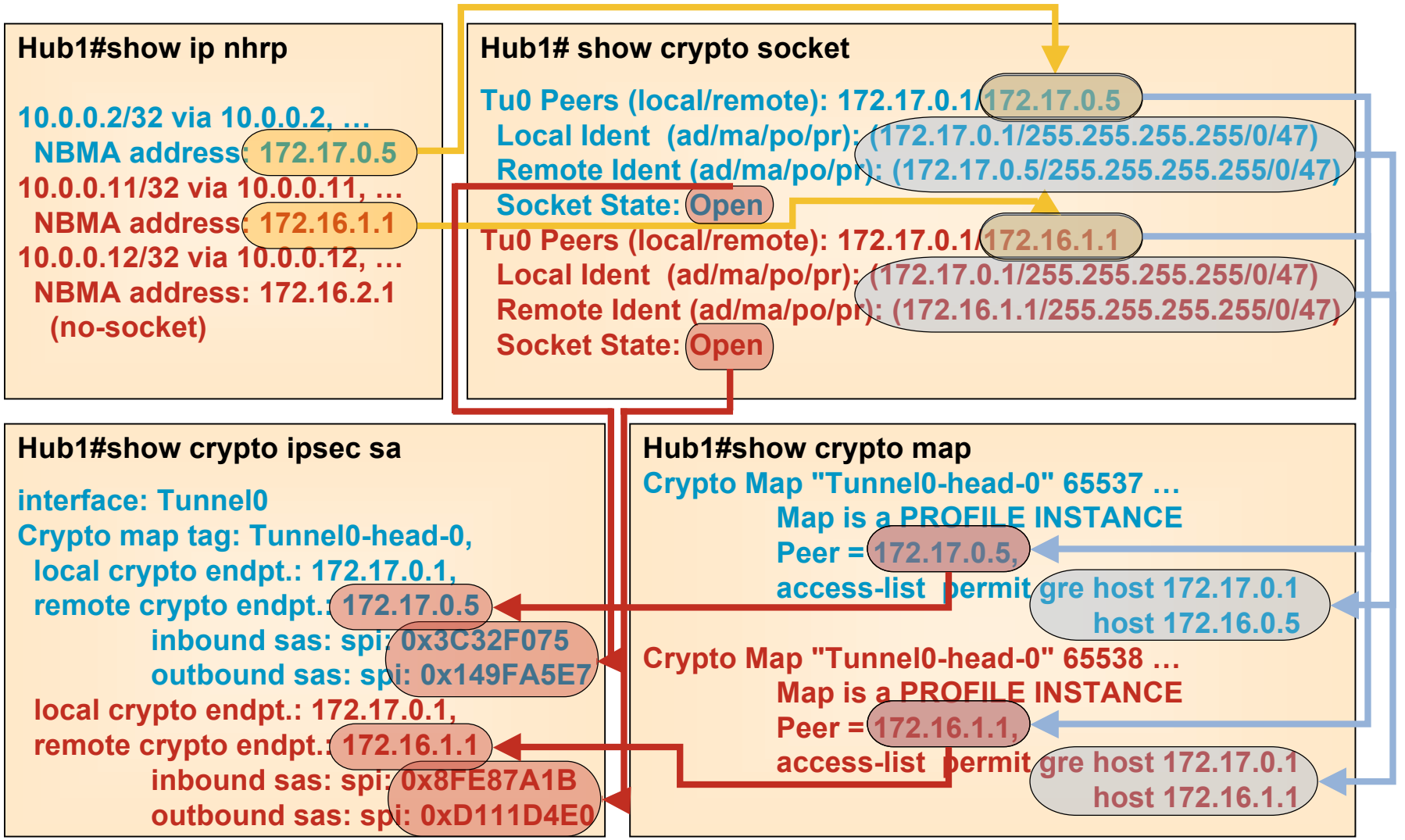
  Dynamic Crypto map for each mGRE tunnel (tunnel protection …)
  or for each IPsec profile (… shared)

  show crypto map

- ISAKMP and IPsec SA Table

  show crypto session { detail }, show crypto isakmp sa { detail }
  show crypto ipsec sa { | include Tag|peer|spi|endpt }

DMVPN Data Structures
Complex interaction



DMVPN Show and Debug Commands Introduced in 12.4(9)T

- Show

  show dmvpn

  [peer {{{ nbma | tunnel } ***ip_address*** } |

  { network ***ip_address mask*** } | { interface ***tunnel#*** } |

  { vrf ***vrf_name*** }}]

  [detail] [static]

- Debug

  debug dmvpn [{ error | event | detail | packet | all }

  { nhrp | crypto | tunnel | socket | all }]

  debug dmvpn condition [peer

  {{{ nbma | tunnel } ***ip_address*** } | { network ***ip_address mask*** } |

  { interface ***tunnel#*** } | { vrf ***vrf_name*** }}]

- Logging

  dmvpn logging { enable | interval < ***0-3600*** > }

Show dmvpn detail – example

For your reference

```
Router# show dmvpn detail

Legend: Attrb --> S - Static, D - Dynamic, I - Incomplete
        N - NATed, L - Local, X - No Socket
        # Ent --> Number of NHRP entries with same NBMA peer
-------------- Interface Tunnel1 info: --------------
Intf. is up, Line Protocol is up, Addr. is 192.0.2.5
   Source addr: 192.0.2.229, Dest addr: MGRE
  Protocol/Transport: "multi-GRE/IP", Protect "gre_prof",
Tunnel VRF "" ip vrf forwarding ""

NHRP Details: NHS: 192.0.2.10 RE 192.0.2.11  E

Type: Spoke, NBMA Peers: 4
# Ent  Peer NBMA Addr Peer Tunnel Add State  UpDn Tm Attrb    Target Network
----- --------------- --------------- ----- -------- ----- -----------------
    2        192.0.2.21      192.0.2.116    UP 00:14:59 D      192.0.2.118/24
                                            UP 00:14:59 D      192.0.2.116/32

  IKE SA: local 192.0.2.229/500 remote 192.0.2.21/500 Active
          Capabilities:(none) connid:1031 lifetime:23:45:00
  Crypto Session Status: UP-ACTIVE
  fvrf: (none)
  IPSEC FLOW: permit 47 host 192.0.2.229 host 192.0.2.21
        Active SAs: 2, origin: crypto map
        Inbound:  #pkts dec'ed 1 drop 0 life (KB/Sec) 4494994/2700
        Outbound: #pkts enc'ed 1 drop 0 life (KB/Sec) 4494994/2700
   Outbound SPI : 0xD1EA3C9B, transform : esp-3des esp-sha-hmac
    Socket State: Open
```

Show dmvpn detail (cont.)

For your reference

```
# Ent  Peer NBMA Addr Peer Tunnel Add State  UpDn Tm Attrb    Target Network
----- --------------- --------------- ----- -------- ----- -----------------
    1     192.0.2.229       192.0.2.5    UP 00:15:00 DLX        192.0.2.5/32

# Ent  Peer NBMA Addr Peer Tunnel Add State  UpDn Tm Attrb    Target Network
----- --------------- --------------- ----- -------- ----- -----------------
    1     192.0.2.102      192.0.2.11 NHRP 02:55:47  S         192.0.2.11/32

  IKE SA: local 192.0.2.229/4500 remote 192.0.2.102/4500 Active
          Capabilities:N connid:1028 lifetime:11:45:37
  Crypto Session Status: UP-ACTIVE
  fvrf: (none)
  IPSEC FLOW: permit 47 host 192.0.2.229 host 192.0.2.102
        Active SAs: 2, origin: crypto map
        Inbound:  #pkts dec'ed 199056 drop 393401 life (KB/Sec) 4560270/1524
        Outbound: #pkts enc'ed 416631 drop 10531 life (KB/Sec) 4560322/1524
   Outbound SPI : 0x9451AF5C, transform : esp-3des esp-sha-hmac

     Socket State: Open
[…]
```

Show dmvpn detail (cont.)

For your reference

```
Pending DMVPN Sessions:

!There are no pending DMVPN sessions.

The following example shows example configured conditions displays for DMVPN debugging:
Router# show dmvpn debug-condition

NBMA addresses under debug are:


Interfaces under debug are:

Tunnel101,

Crypto DMVPN filters:

Interface = Tunnel101


DMVPN Conditional debug context unmatched flag: OFF
```

Phase 3: Platform Support Summary

Cisco IOS Code and Platform Support

- IOS Code
  - Phase 1 & 2
    - 12.3(17), 12.3(14)T6, 12.4(7), 12.4(4)T
  - Phase 1, 2 & 3
    - 12.4(6)T
- Platforms
  - 6500/7600 (12.2(18)SXF4) with VPN-SPA + sup720
    - No Phase 3 capability yet
  - 7301, 7204/6, 38xx, 37xx, 36xx, 28xx, 26xx,
  - 18xx, 17xx, 87x, 83x
    - Phase 1, 2 & 3

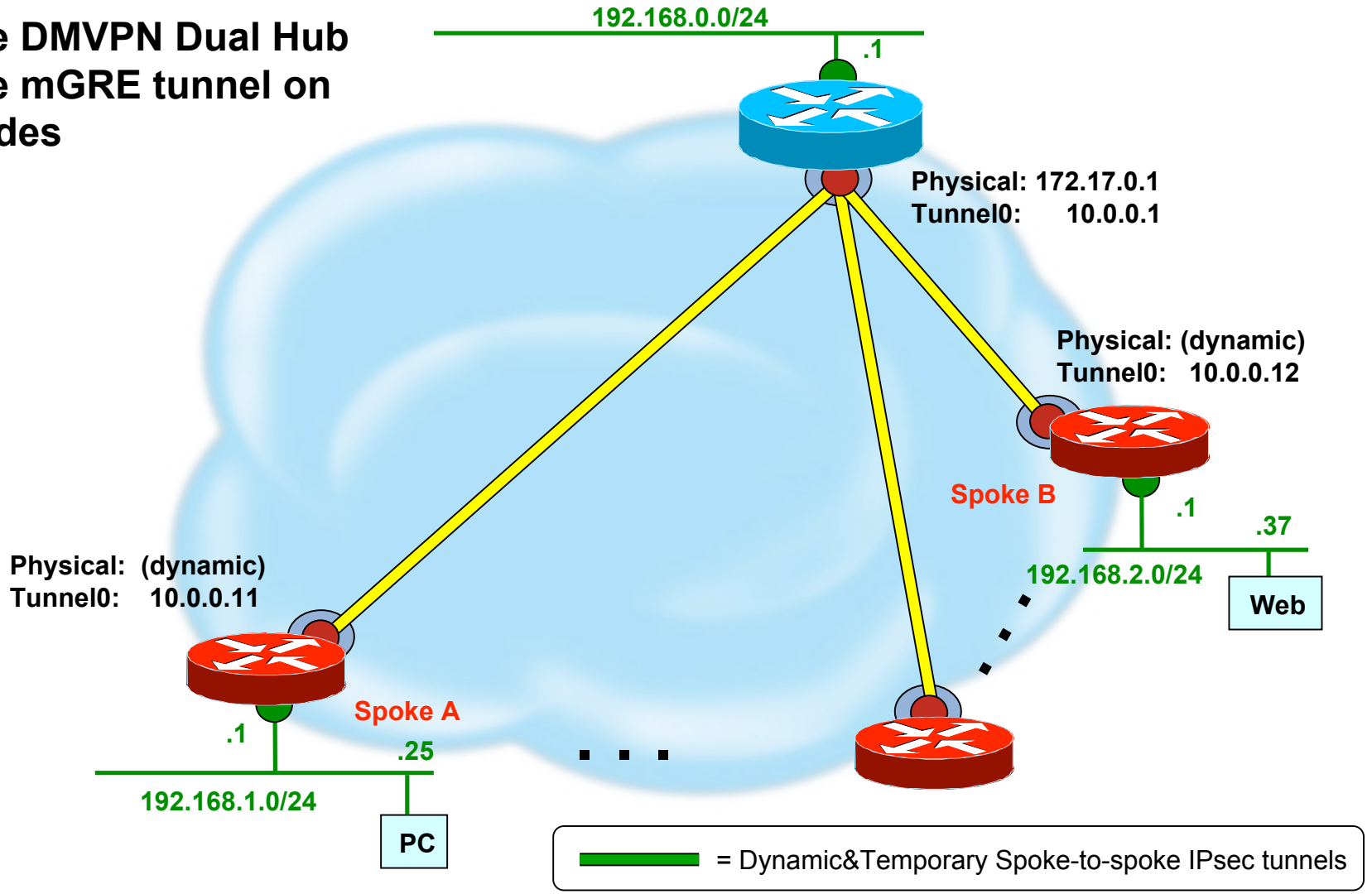
Basic DMVPN Spoke – Spoke

Basic DMVPN spoke-spoke

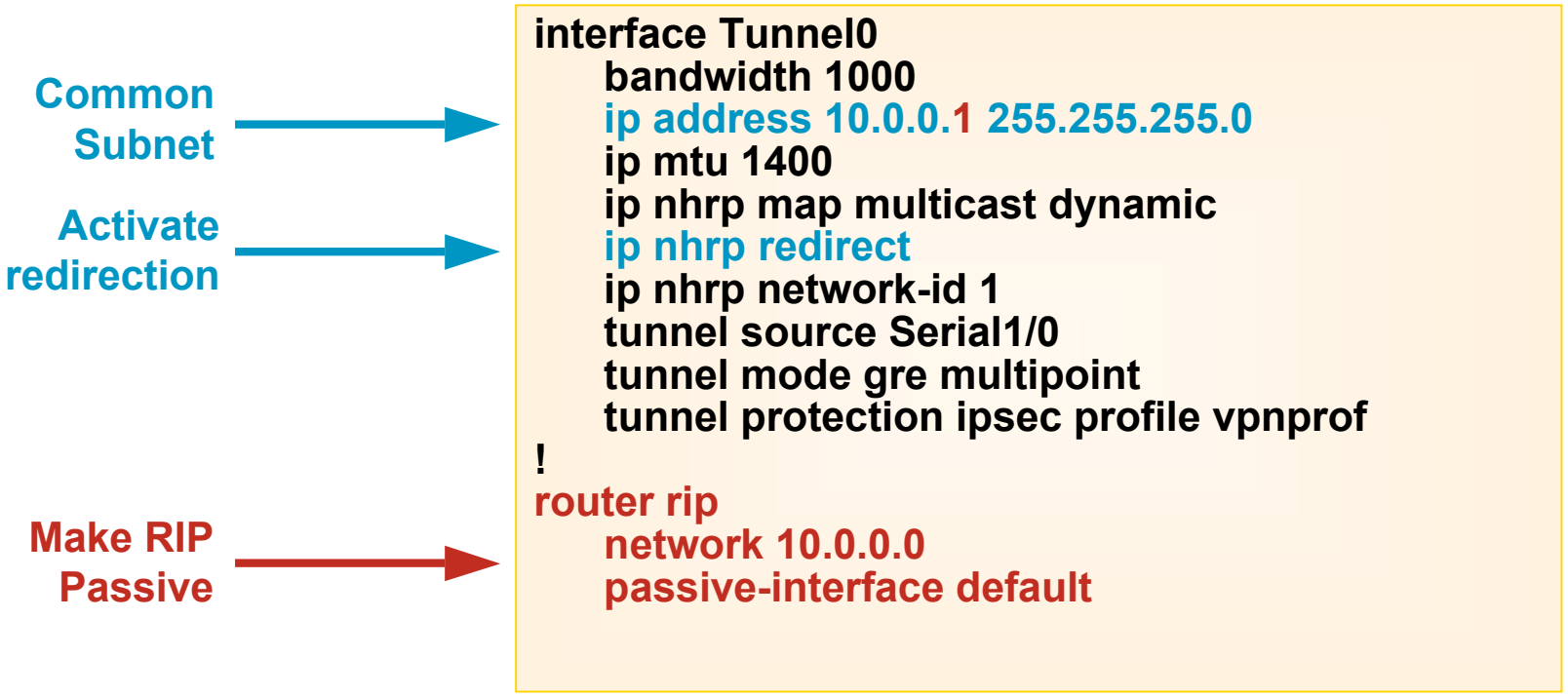
- Phase 3 enhancements helps scaling DMVPN
- The 1000 nodes/hub barrier with spoke-spoke is broken
- Let's start with a very simple spoke-spoke design
- We will use a mix of two protocols
  - RIP passive to scale spoke→hub routing propagation
    - Up to 1500 spokes on a single 7200/VAM2+
  - NHRP shortcut switching to offer spoke-spoke

Basic DMVPN spoke-spoke topology

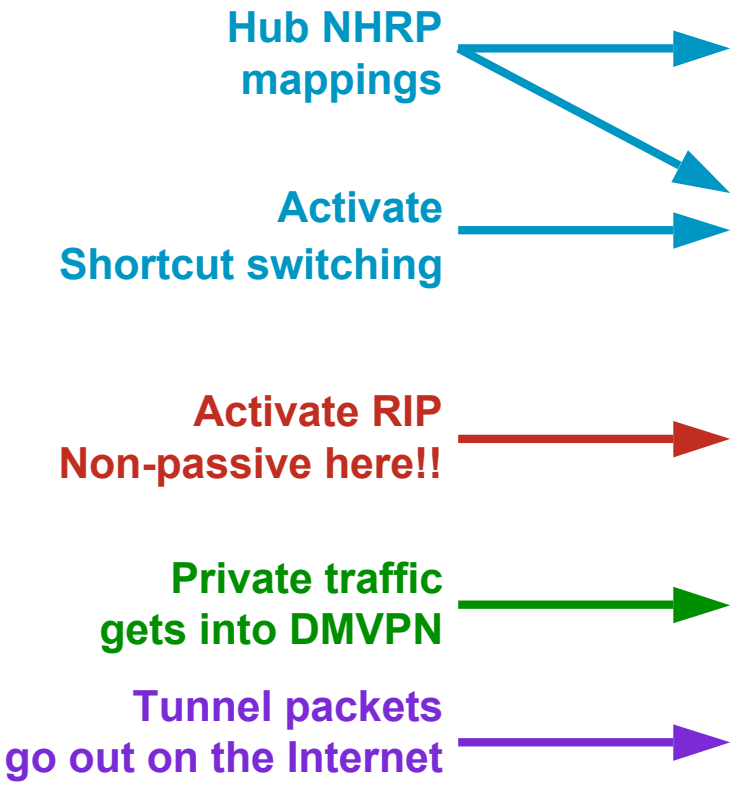
Single DMVPN Dual Hub Single mGRE tunnel on all nodes



Basic DMVPN spoke-spoke Hub configuration



Basic DMVPN spoke-spoke
Spoke configuration




```
interface Tunnel0
    ip vrf forwarding Private
    bandwidth 1000
    ip address 10.0.0.<x> 255.255.255.0        ! <x> = 11,12,…
    ip mtu 1400
    ip nhrp map multicast 172.17.0.1
    ip nhrp map 10.0.0.1 172.17.0.1
    ip nhrp network-id 1
    ip nhrp holdtime 360
    ip nhrp nhs 10.0.0.1
    ip nhrp shortcut
    tunnel source Serial1/0
    tunnel mode gre multipoint
    tunnel protection ipsec profile vpnprof

router rip
    network 10.0.0.0
    network 192.168.<x>.0                      ! <x> = 1,2,…

ip route 192.168.0.0 255.255.0.0 10.0.0.1
ip route 10.0.0.0 255.0.0.0 10.0.0.1


ip route 0.0.0.0 0.0.0.0 Serial 1/0
```

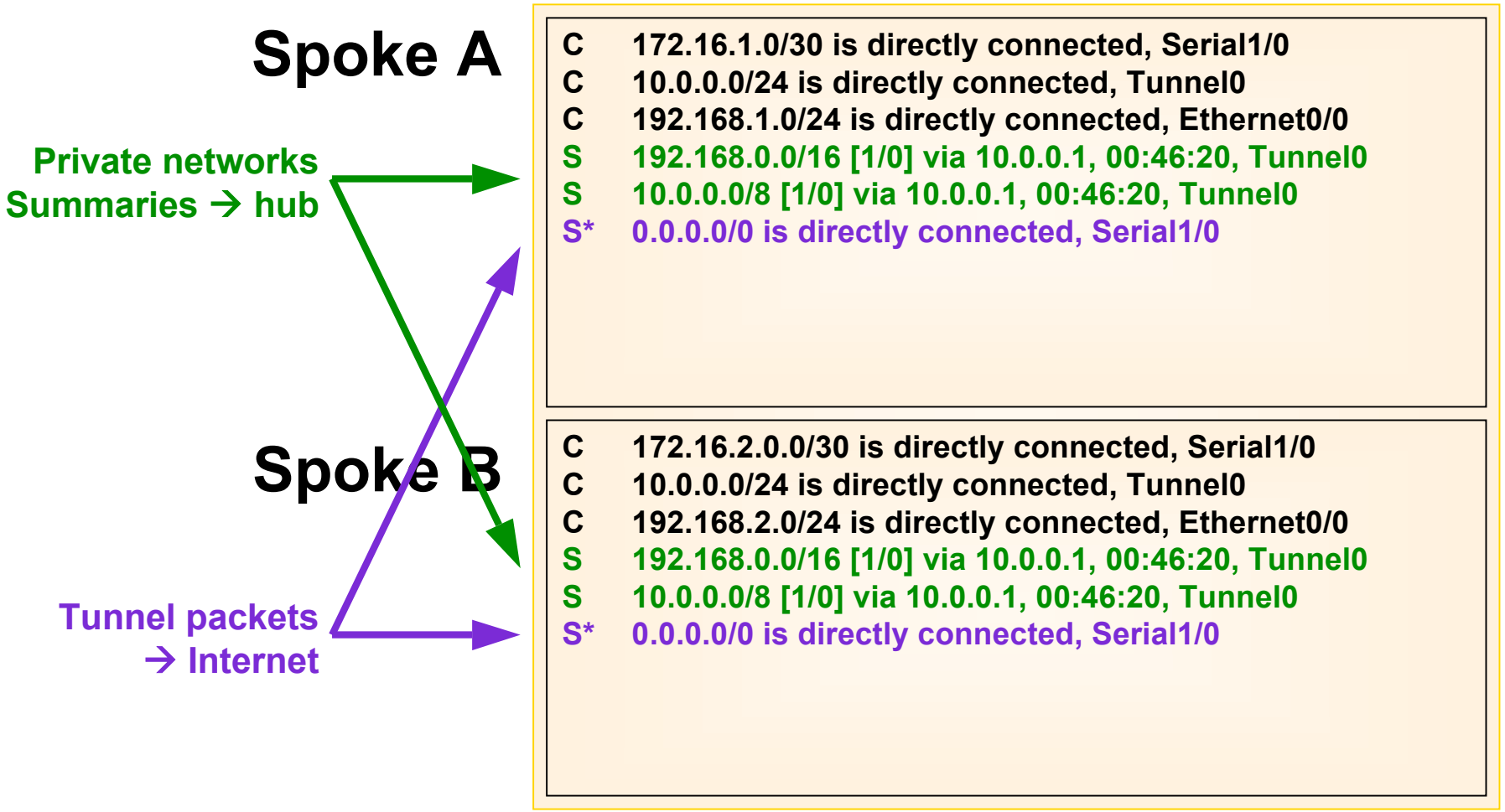
Basic DMVPN spoke-spoke Hub Routing Table

Hub

RIP learned routes

```
C    172.17.0.0/30 is directly connected, Serial1/0
C    10.0.0.0/24 is directly connected, Tunnel0
C    192.168.0.0/24 is directly connected, Ethernet0/0
R    192.168.1.0/24 [120/1] via 10.0.0.11, 00:36:53, Tunnel0
R    192.168.2.0/24 [120/1] via 10.0.0.12, 00:37:58, Tunnel0
      . . .
S*   0.0.0.0/0 [1/0] via 172.17.0.2
```

Basic DMVPN spoke-spoke Spokes Routing Tables



Basic DMVPN spoke-spoke Hub NHRP Tables

Hub

Learned via registration

```
10.0.0.11/32 via 10.0.0.11, Tunnel0 created 02:51:46, expire 00:04:13
  Type: dynamic, Flags: authoritative unique registered used
  NBMA address: 172.16.1.1
10.0.0.12/32 via 10.0.0.12, Tunnel0 created 02:51:26, expire 00:04:33
  Type: dynamic, Flags: authoritative unique registered used
  NBMA address: 172.16.2.1
```

Basic DMVPN spoke-spoke Spokes NHRP Tables

Spoke A

```
10.0.0.1/32 via 10.0.0.1, Tunnel0 created 02:51:20, never expire
  Type: static, Flags: authoritative used
  NBMA address: 172.17.0.1

10.0.0.12/32 via 10.0.0.12, Tunnel0 created 00:00:06, expire 00:05:05
  Type: dynamic, Flags: router unique used
  NBMA address: 172.16.2.1
```

Learned via resolution →

Spoke B

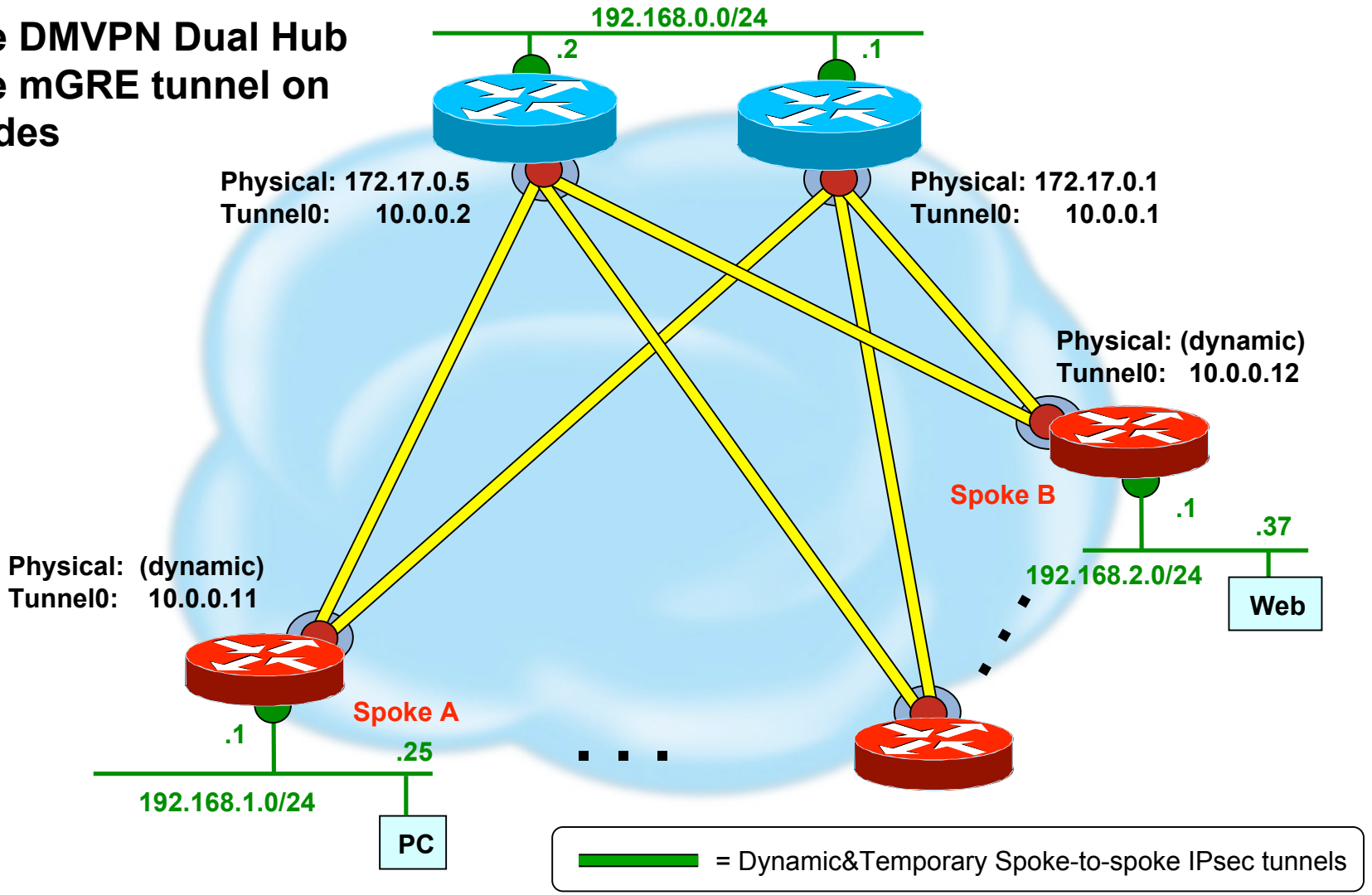
```
10.0.0.1/32 via 10.0.0.1, Tunnel0 created 02:51:18, never expire
  Type: static, Flags: authoritative used
  NBMA address: 172.17.0.1

10.0.0.11/32 via 10.0.0.11, Tunnel0 created 00:00:24, expire 00:04:27
  Type: dynamic, Flags: router unique used
  NBMA address: 172.16.1.1
```

Learned via resolution →

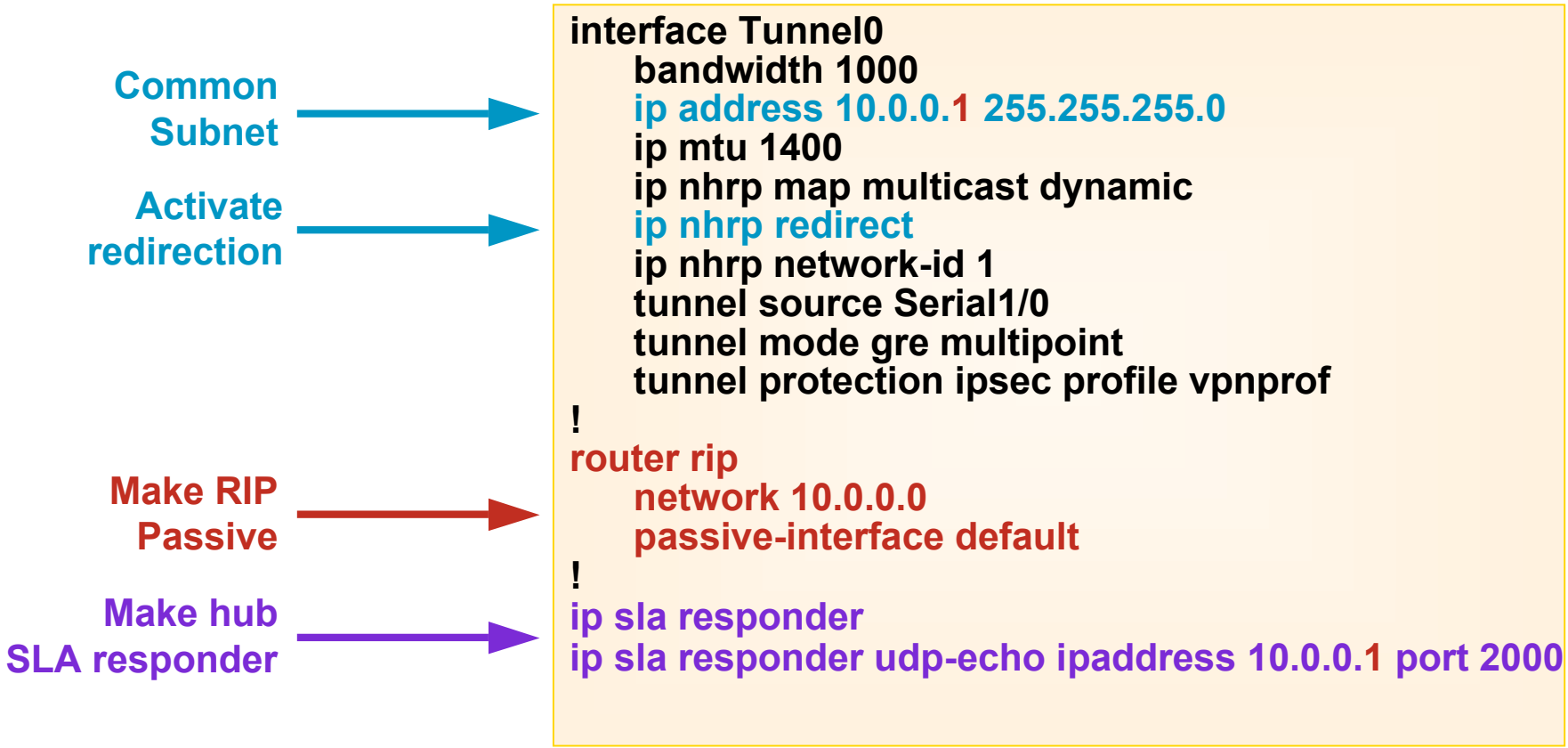
Dual homed spokes

Dual homed DMVPN spokes

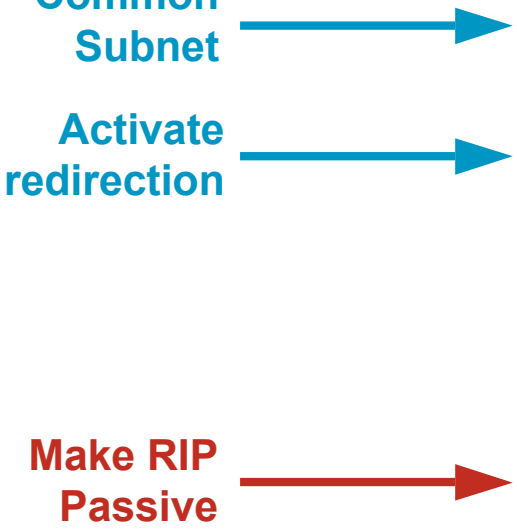
Single DMVPN Dual Hub
Single mGRE tunnel on all nodes



Dual homed DMVPN spokes
Hub1

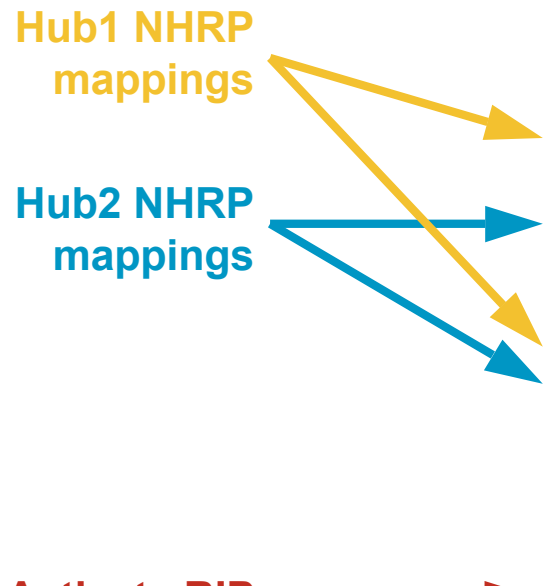


Dual homed DMVPN spokes
Hub2



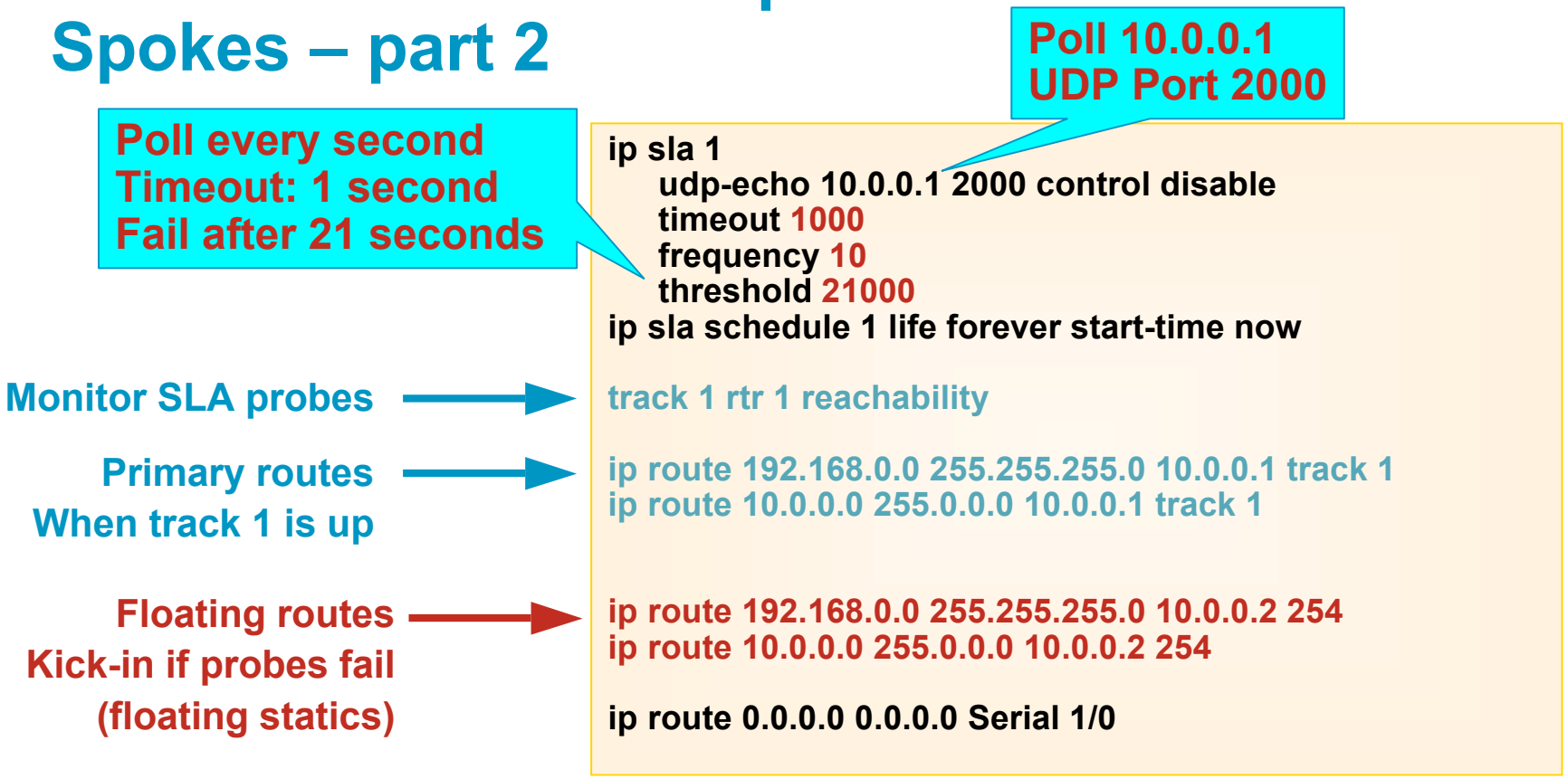

```
interface Tunnel0
     bandwidth 1000
     ip address 10.0.0.2 255.255.255.0
     ip mtu 1400
     ip nhrp map multicast dynamic
     ip nhrp redirect
     ip nhrp network-id 1
     tunnel source Serial1/0
     tunnel mode gre multipoint
     tunnel protection ipsec profile vpnprof
!
router rip
     network 10.0.0.0
     passive-interface default
```

Dual homed DMVPN spokes
Spokes – part 1




```
interface Tunnel0
      bandwidth 1000
      ip address 10.0.0.<x> 255.255.255.0      ! <x> = 11,12,…
      ip mtu 1400
      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 network-id 1
      ip nhrp holdtime 360
      ip nhrp nhs 10.0.0.1
      ip nhrp nhs 10.0.0.2
      tunnel source Serial1/0
      tunnel mode gre multipoint
      tunnel protection ipsec profile vpnprof

router rip
      network 10.0.0.0
      network 192.168.<x>.0                    !<x> = 1,2,…
```

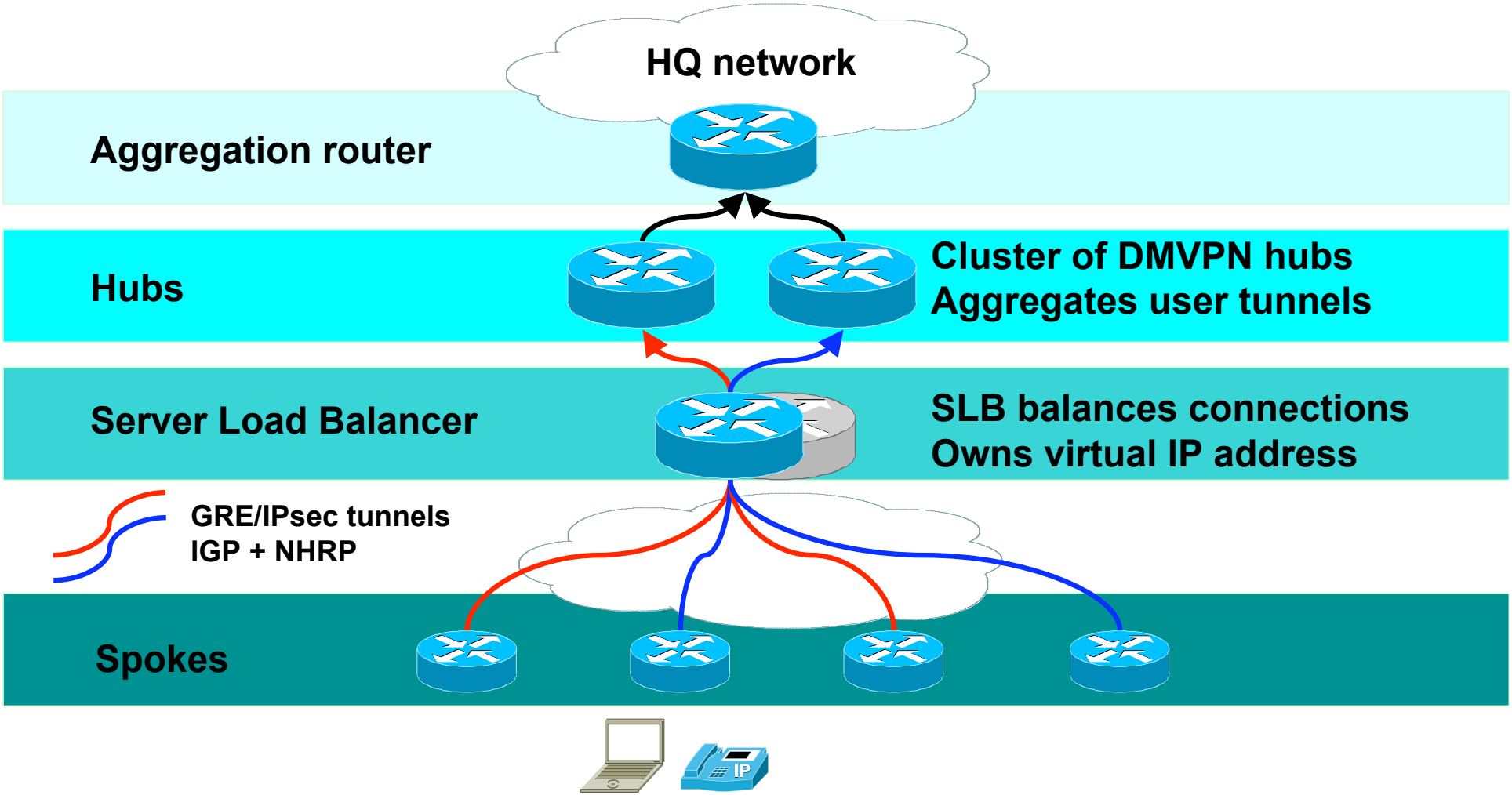
Dual homed DMVPN spokes
Spokes – part 2



- Model shown here makes hub1 primary, hub2 backup
- Track both hubs to make active-active if desired

Large Scale DMVPN Hub & Spoke

Overall solution



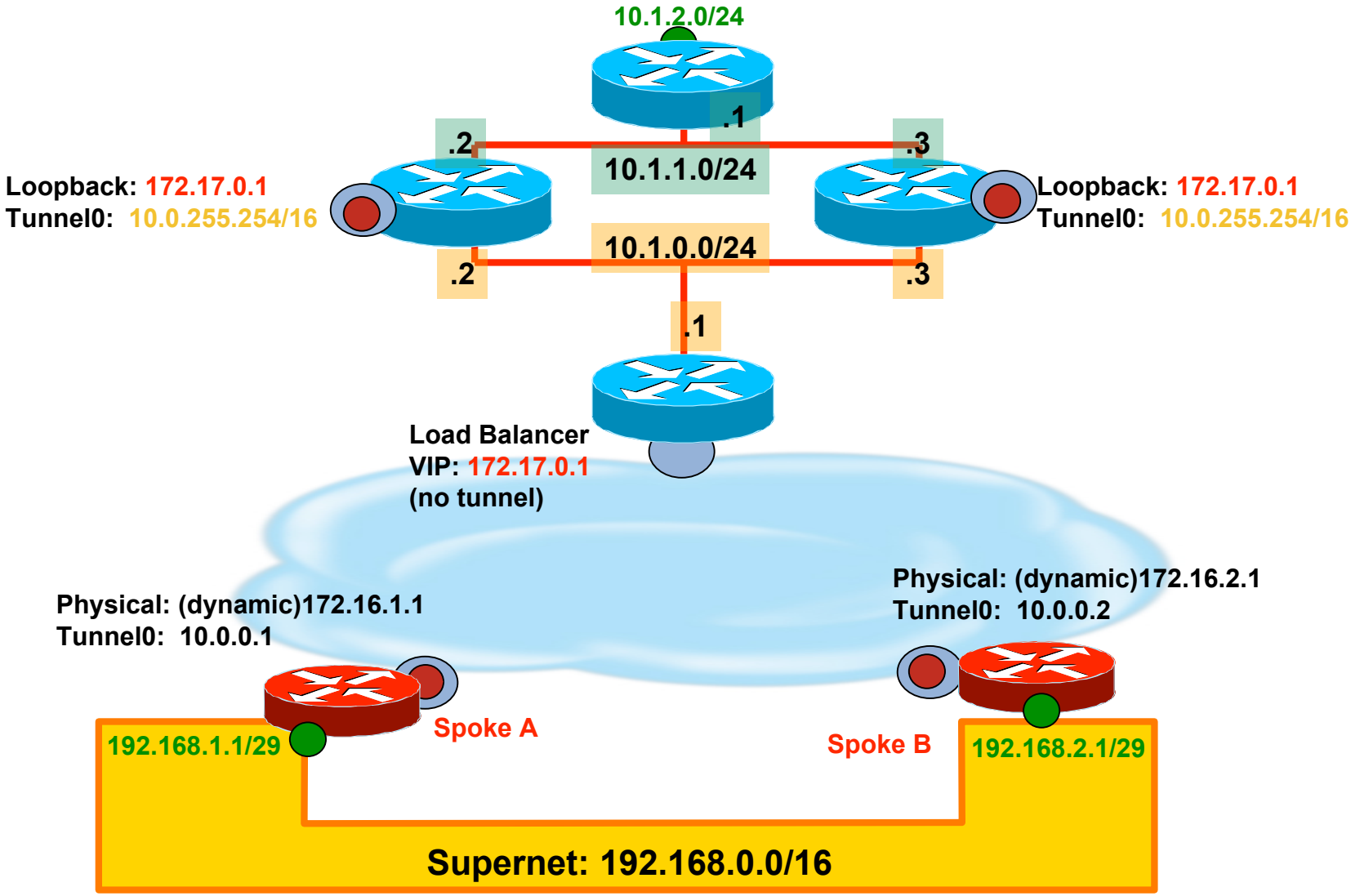
High level description

- Spokes believe there is a single hub
- NHRP map points to the Load Balancer's Virtual IP Address
- The Load Balancer is configured in forwarding mode (no NAT)
- All the hubs have the same DMVPN configuration
  - Same Tunnel interface address
  - Same Loopback address (equal to the VIP)
- All the spokes have the same DMVPN configuration
  - Same hub NBMA address
  - Same NHS

The Load Balancer in general

- The Load Balancer owns a Virtual IP Address (VIP)
- When IKE or ESP packets are targeted at the VIP, the LB chooses a hub
- The hub choice is policy (predictor) based:
  - weighted round-robin
  - least-connections
  - …
- When hub chosen for a “tunnel”, all packets go to the same hub
  - → **stickyness**
- Once a decision is made for IKE, the same is made for ESP
  - → **buddying**

Topology and addresses



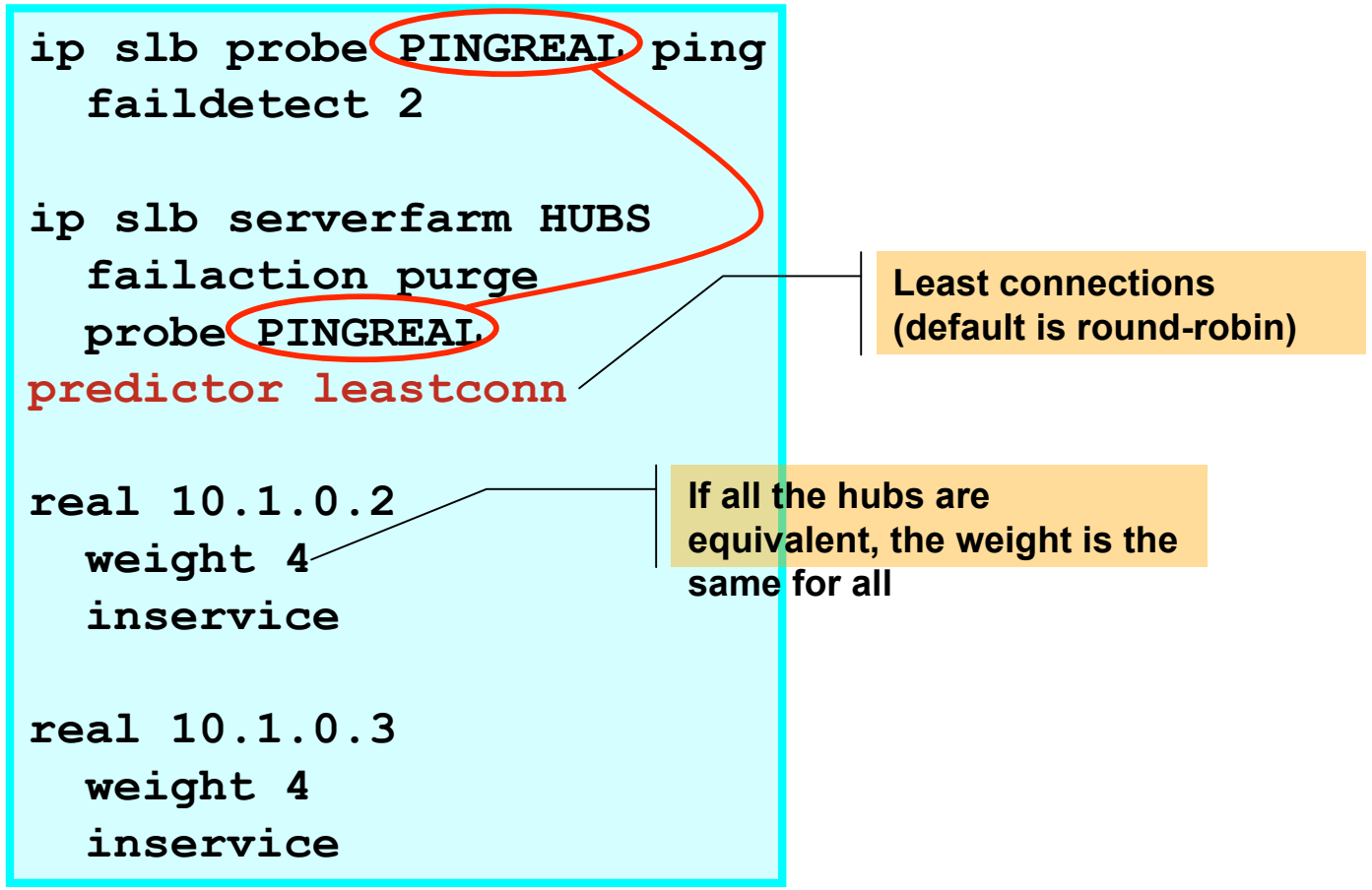
Load Balancer

- We will use an IOS-SLB
  - IOS SLB runs on top of c7200 or Catalyst6500
  - As of today, opt for 12.2S or 12.1E releases
- The LB must be able to do layer 3 and 4 load balancing. Upper layers are useless (encrypted)
- Content Switching Module 3.1 or above will work too but we do not need most of its features (layer 5+)
- ACE is ok but need to disable NAT-T
- Any SLB will do…

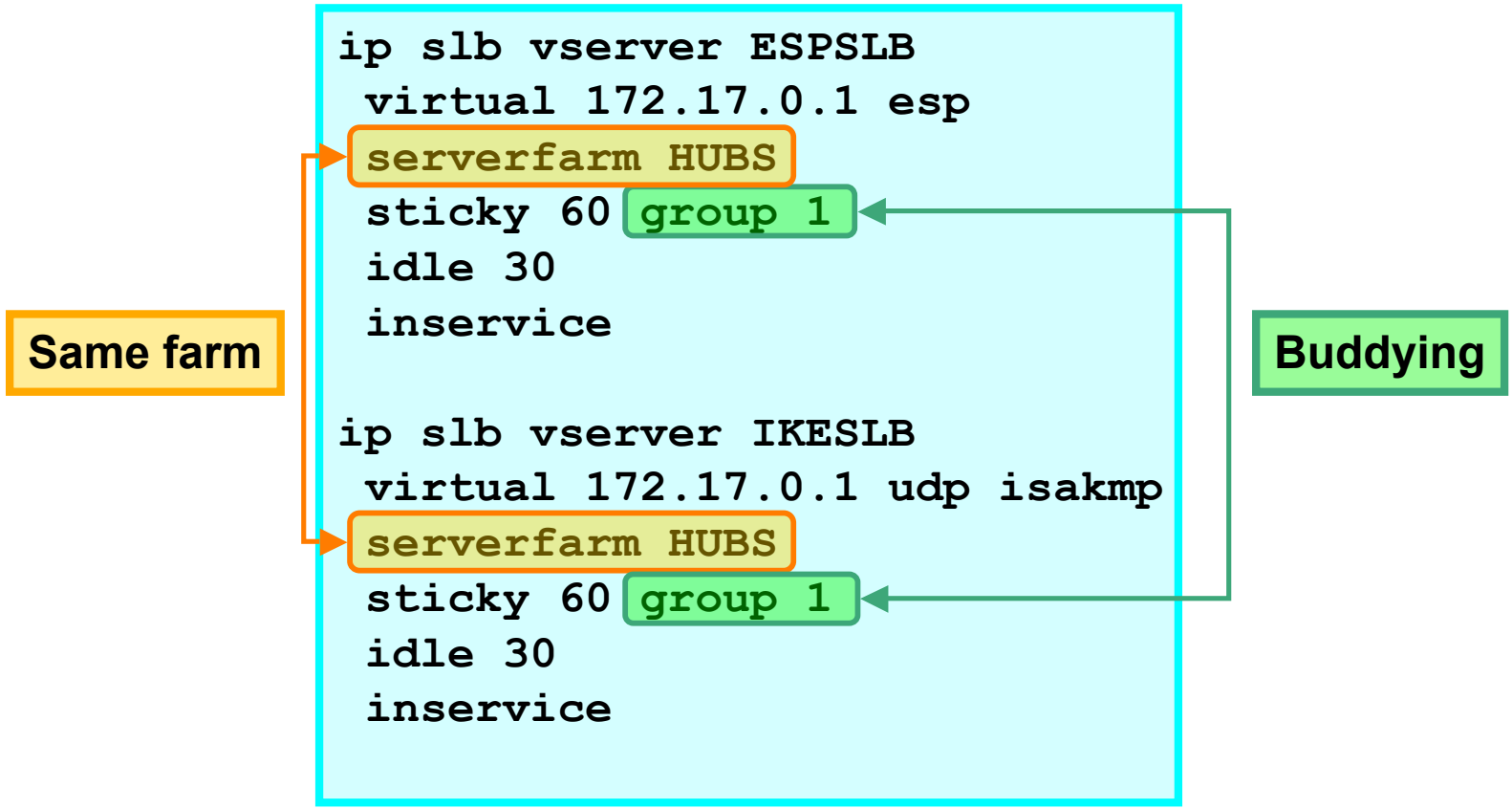
IOS SLB performances

- IOS SLB on a Cat6500 (MSFC-2)
  - Can manage 1M connections w/ 128MB RAM
  - Can create 20,000 connections per second
  - Switches packets at 10Gbps (64 bytes)
- IOS SLB on a c7200 (NPE-400)
  - Can create 5,000 connections per second
  - Switches packets at ½ the CEF rate (depending on other features)
- Typically not a bottleneck

IOS SLB cluster definition



IOS SLB VIP definition

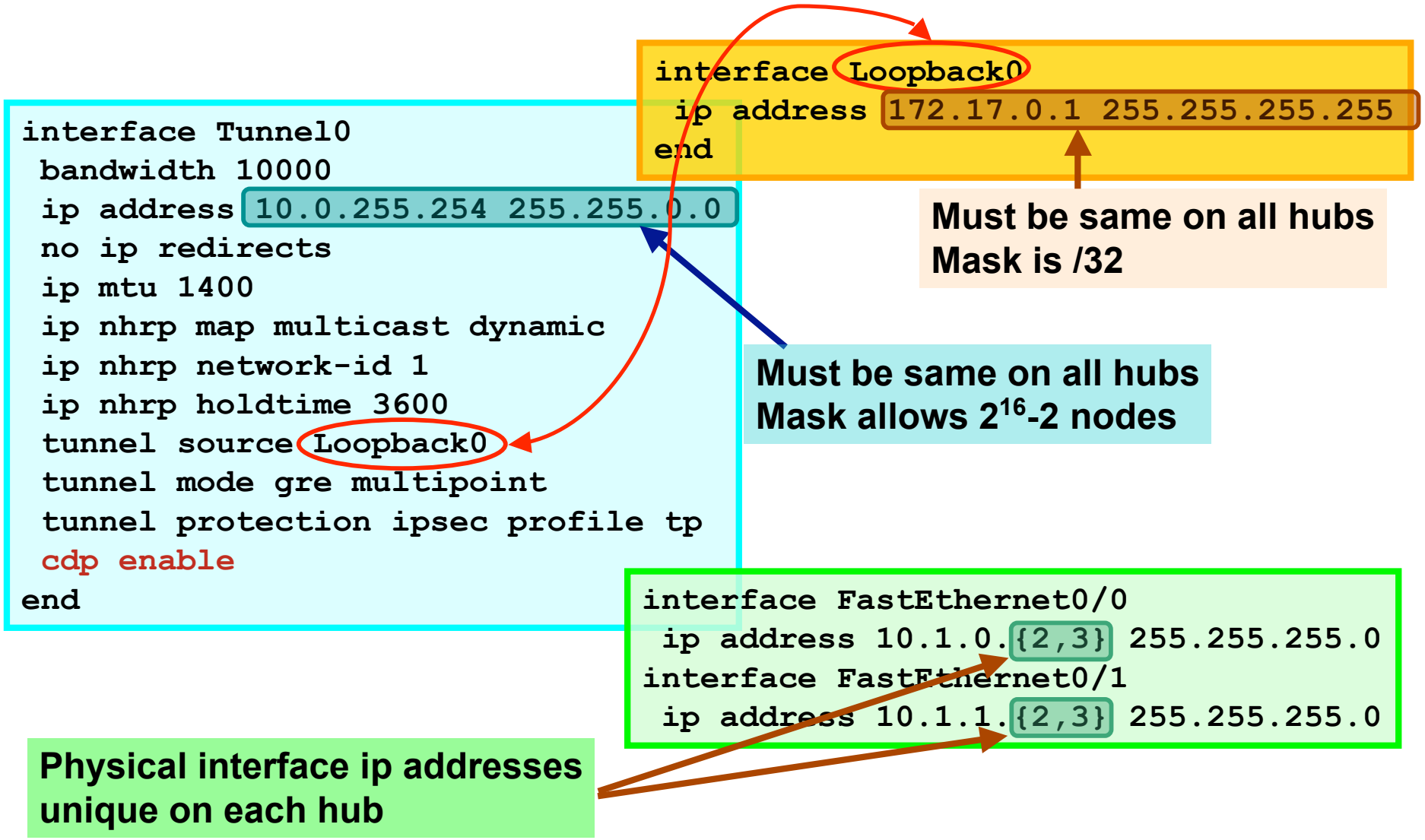


Monitoring and managing

For your reference

```
SLB-7200#sh ip slb connections

vserver          prot client                 real                   state      nat
-------------------------------------------------------------------------------
-
IKESLB           UDP  64.103.8.8:500        10.1.0.2               ESTAB
none
ESPSLB           ESP  217.136.116.189:0     10.1.0.2               ESTAB
none
IKESLB           UDP  213.224.65.3:500      10.1.0.2               ESTAB
none
ESPSLB           ESP  80.200.49.217:0       10.1.0.2               ESTAB
none
ESPSLB           ESP  217.136.132.202:0     10.1.0.3               ESTAB
none

SLB-7200#clear ip slb connections ?
  firewallfarm  Clear connections for a firewallfarm
  serverfarm    Clear connections for a specific serverfarm
  vserver       Clear connections for a specific virtual server
  <cr>

SLB-7200#sh ip slb reals

real             farm name       weight  state          conns
```

Hub Tunnel configuration



Spoke tunnel configuration

- Basic DMVPN / ODR configuration

```
interface Tunnel0
  ip address 10.0.0.1 255.255.255.0
  ip nhrp map 10.0.255.254 172.17.0.1
  ip nhrp nhs 10.0.255.254
  …
```

- Remember…

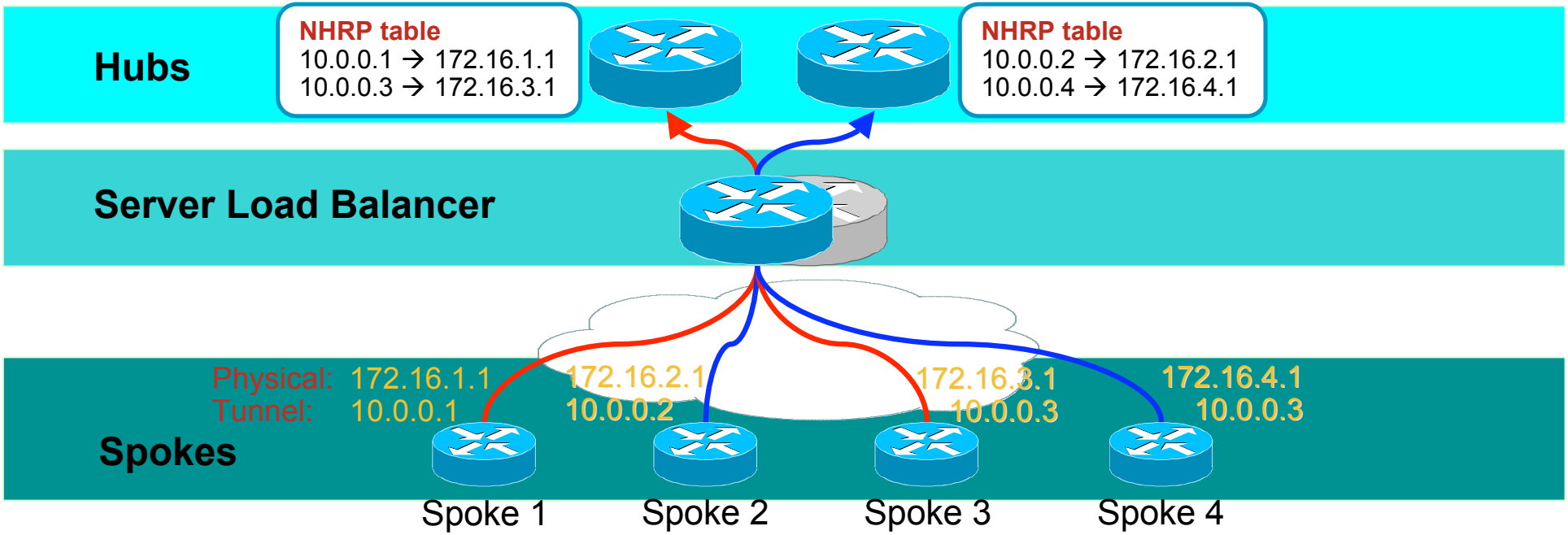
  All the spokes have the same configuration

Current status – Tunnel setup

- We now allow spokes to

  build a DMVPN tunnel to a virtual hub

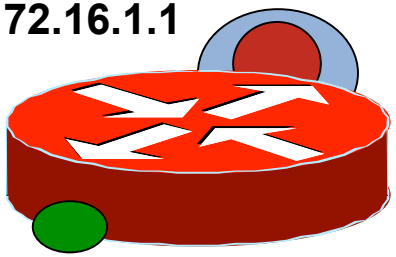
  NHRP-register to their assigned hub



Spoke routing configuration

	`interface Tunnel0`
Activate ODR over tunnel	`cdp enable`
Tunnel packet → physical	`ip route 0.0.0.0 0.0.0.0 Dialer0`
Private traffic (summary) → Tunnel 0	`ip route 192.168.0.0 255.255.0.0 10.0.0.1` `ip route 10.0.0.0 255.0.0.0 10.0.0.1`

Physical: (dynamic)172.16.1.1
Tunnel0: 10.0.0.11



Spoke A

192.168.1.1/29

Hub Routing Protocol configuration

- Only allow private networks in the routing table
- Prevents recursive routing

Activate ODR over tunnel

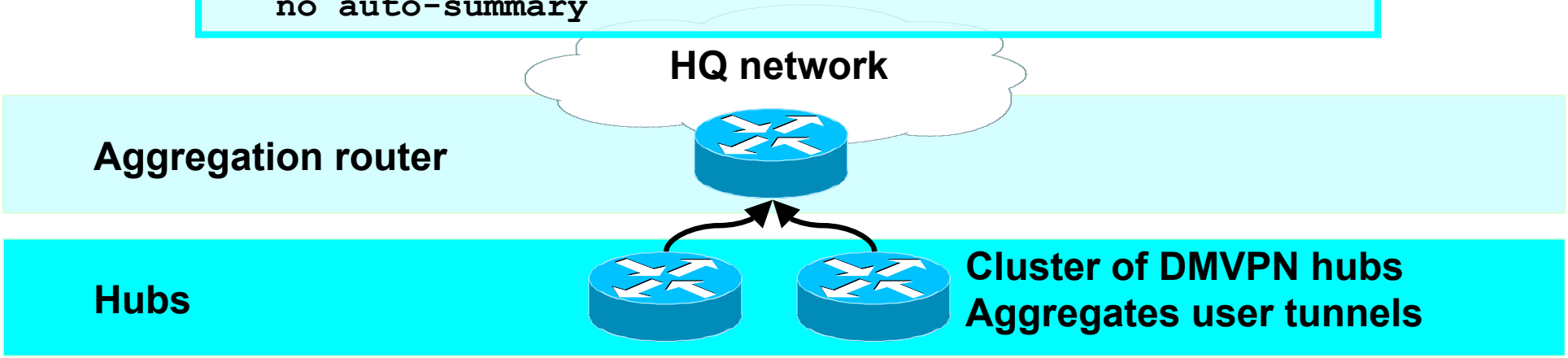
Tunnel packet → physical

Redistribute
ODR → BGP
Send information to aggregation router

```
interface Tunnel0
  cdp enable
router odr
  distribute-list 1 in
access-list 1 permit 192.168.0.0 0.0.255.255
router bgp 1
  redistribute odr
  neighbor 10.1.1.1 remote-as 1
  neighbor 10.1.1.1 next-hop-self
```

HQ Edge BGP configuration

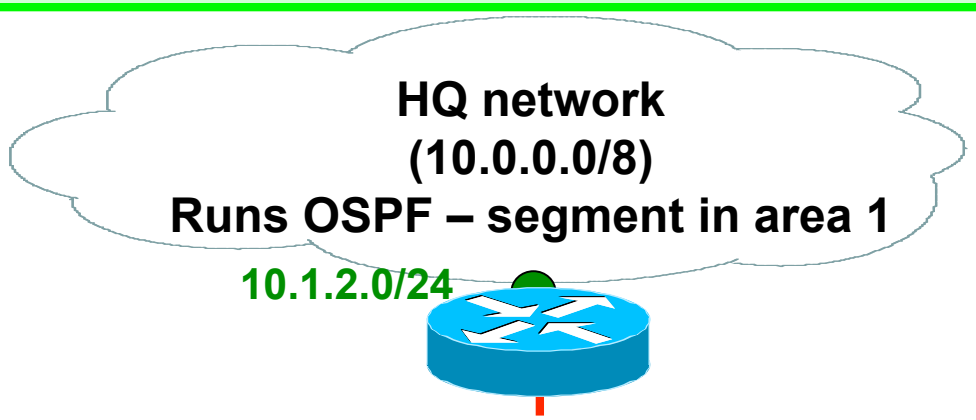
```
router bgp 1
   no synchronization
   bgp log-neighbor-changes
   aggregate-address 10.0.0.0 255.0.0.0 summary-only
   aggregate-address 192.168.0.0 255.0.0.0 summary-only
   neighbor HUB peer-group
   neighbor HUB remote-as 1
   neighbor 10.1.1.2 peer-group HUB
   neighbor 10.1.1.3 peer-group HUB
   neighbor <other hubs> peer-group HUB
   no auto-summary
```



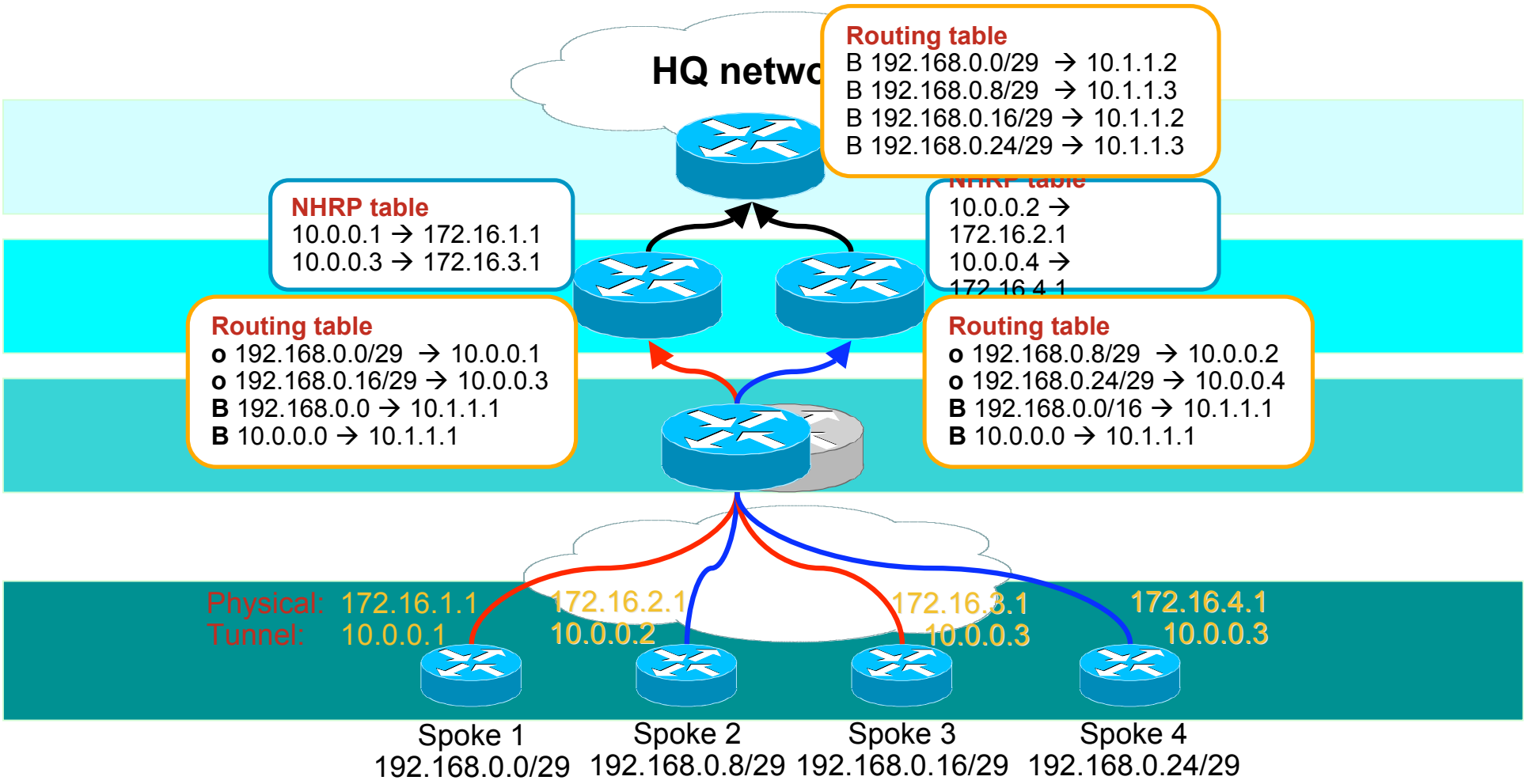
Edge router OSPF configuration

- OSPF attracts traffic from the HQ → DMVPN
- Floating static route to Null0 discards packets to unconnected spokes

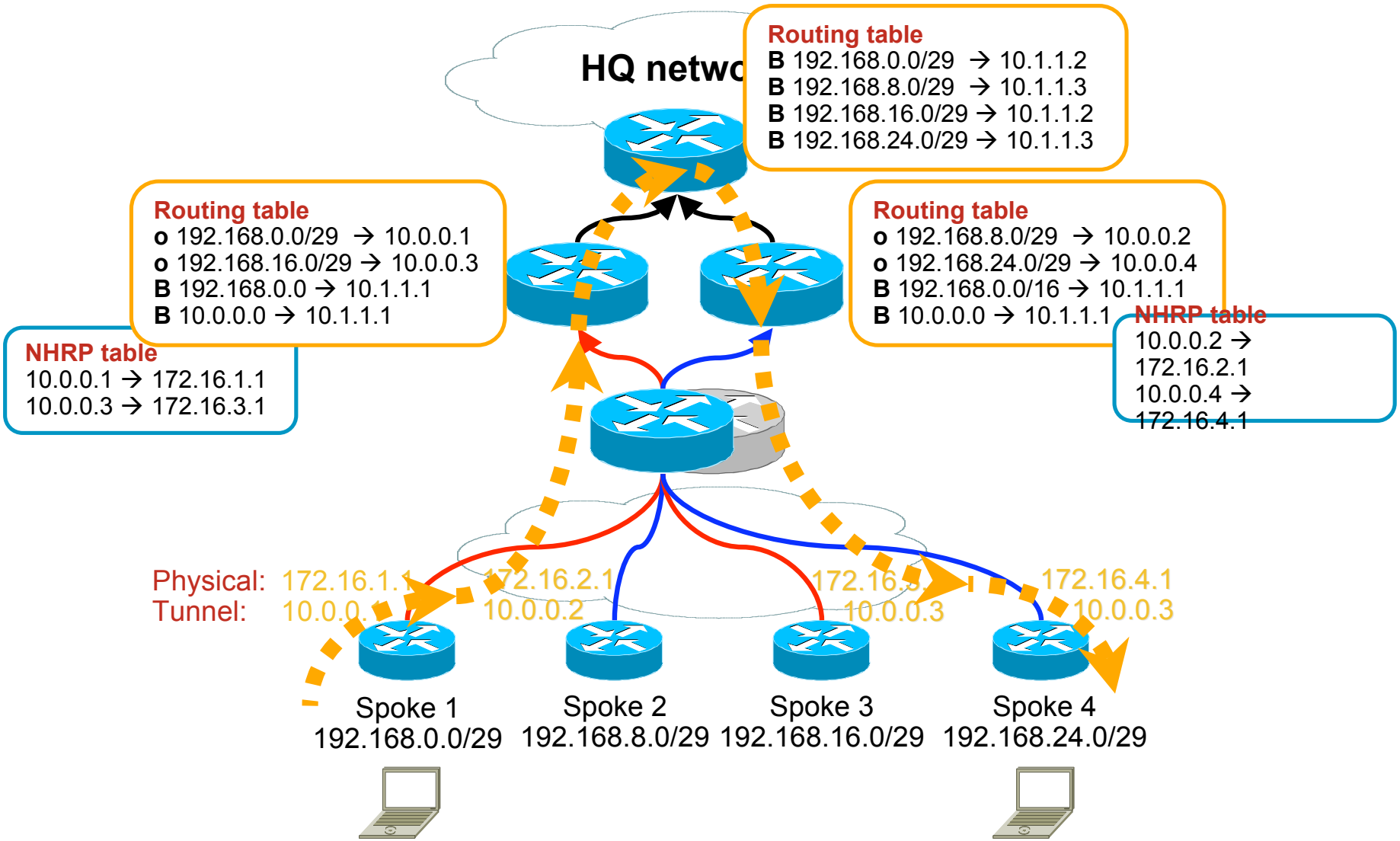
```
ip route 192.168.0.0 255.255.255.127 Null0 254

router ospf 1
 redistribute static
 network 10.1.2.0 0.0.0.255 area 1
```



Routing protocols
Route Propagation spoke → aggregation



Hub&Spoke packet flow



Large Scale DMVPN Spoke – Spoke

Shortcut switching

- Spoke configurations get a single extra line:

  interface Tunnel0

  ip nhrp shortcut ! ← that's it!!

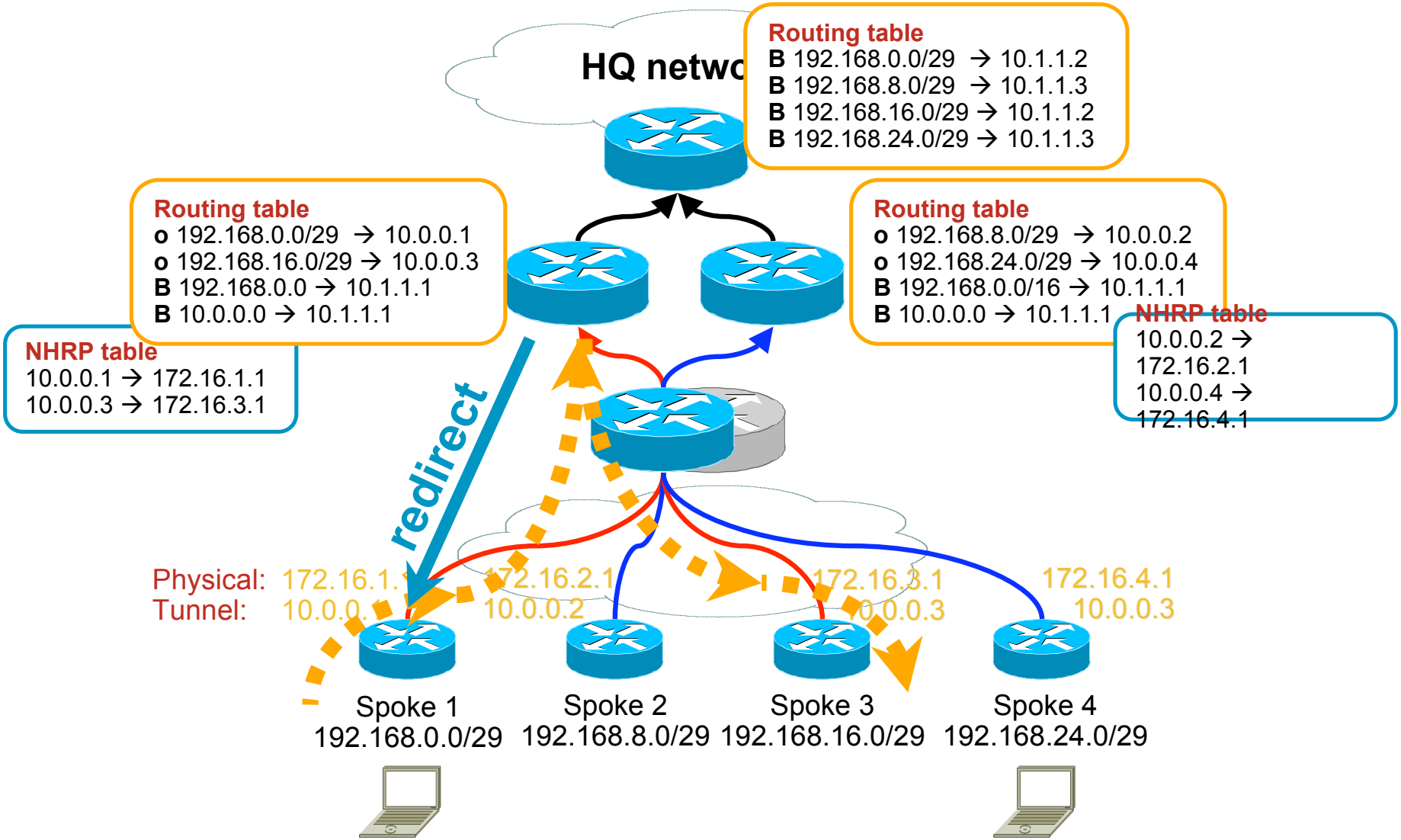
- Hub get an extra line:

  interface Tunnel0

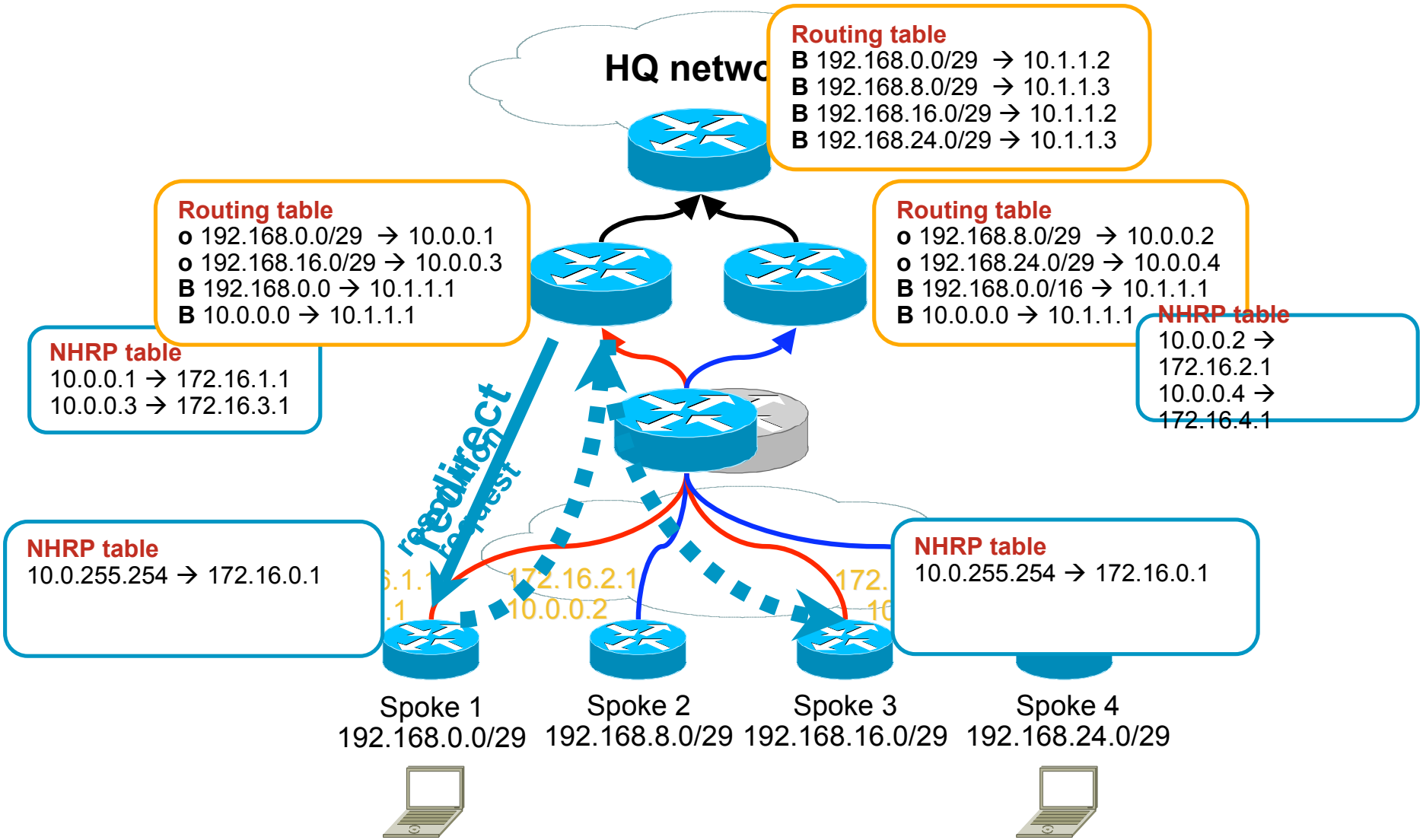
  ip nhrp redirect ! ← that's it!!

- Spokes on a given hub will create direct tunnels
- Spokes on different hubs will NOT create tunnels

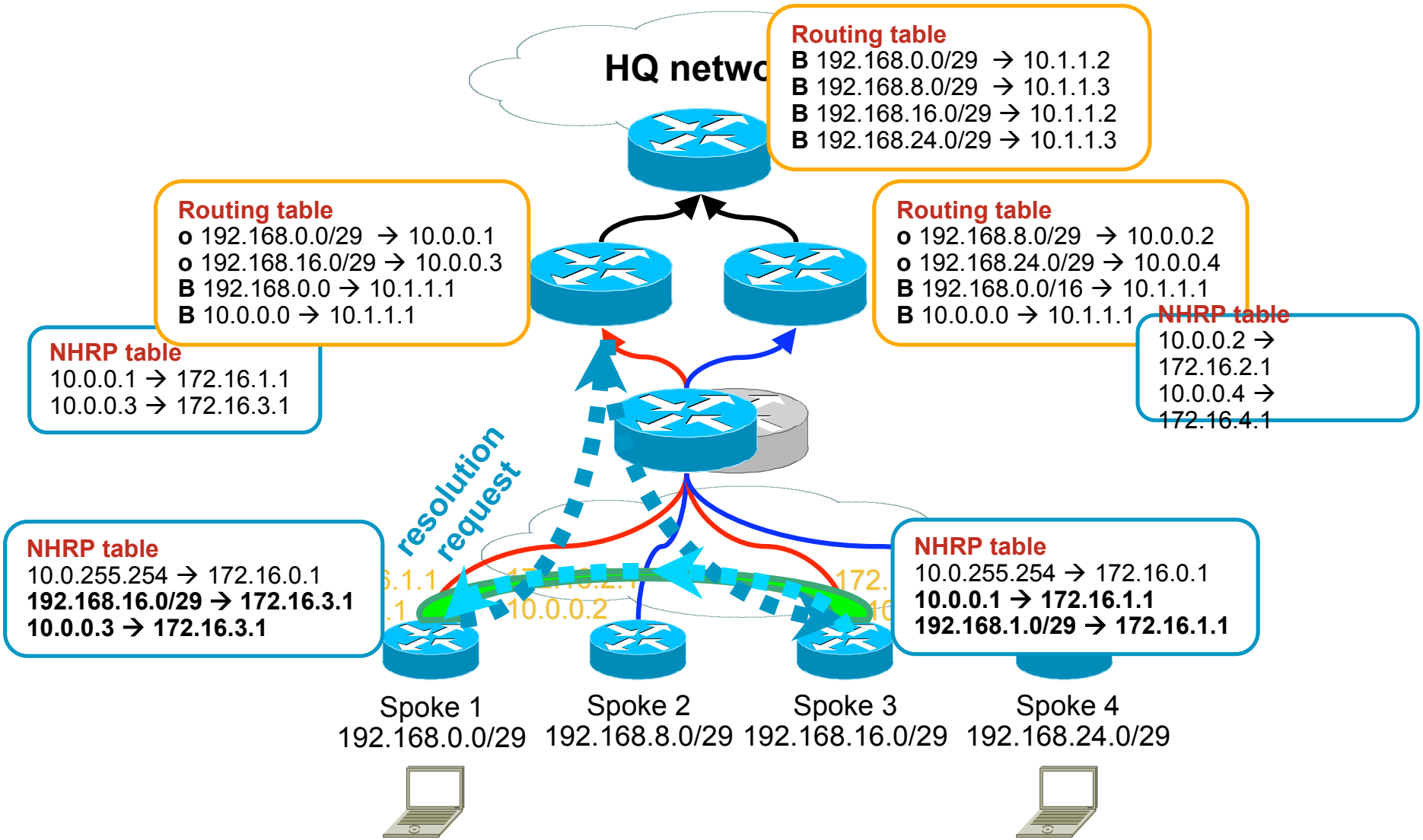
Basic spoke-spoke packet flow



Basic spoke-spoke packet flow



Basic spoke-spoke packet flow



Cross-hubs spoke-spoke tunnels

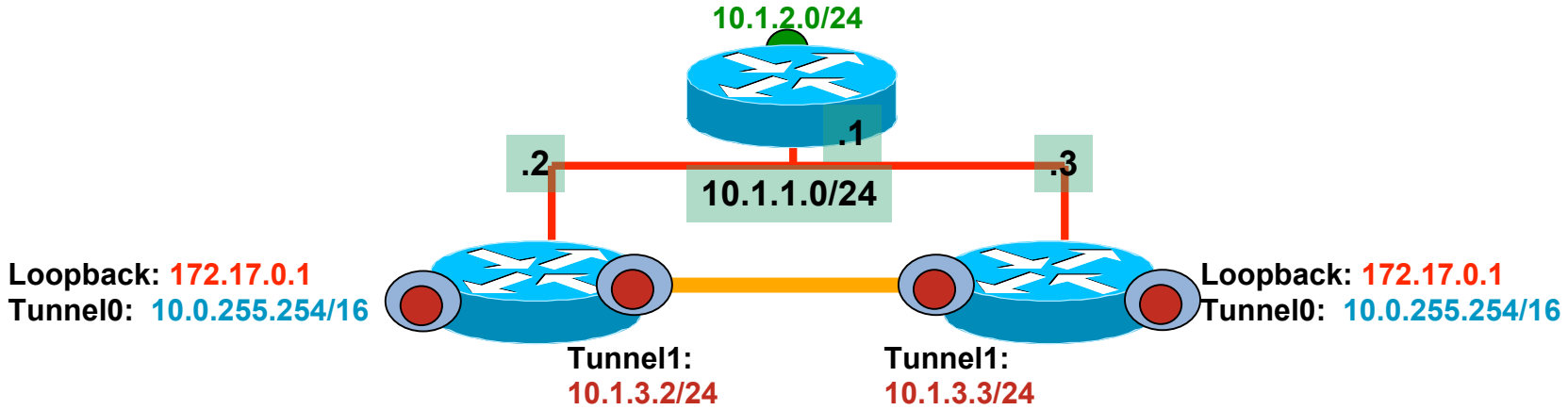
- We want spokes to create direct tunnels even if they are on different hubs
- For this, we link the hubs via a DMVPN
- NOT a daisy chain!!!

Linking the hubs

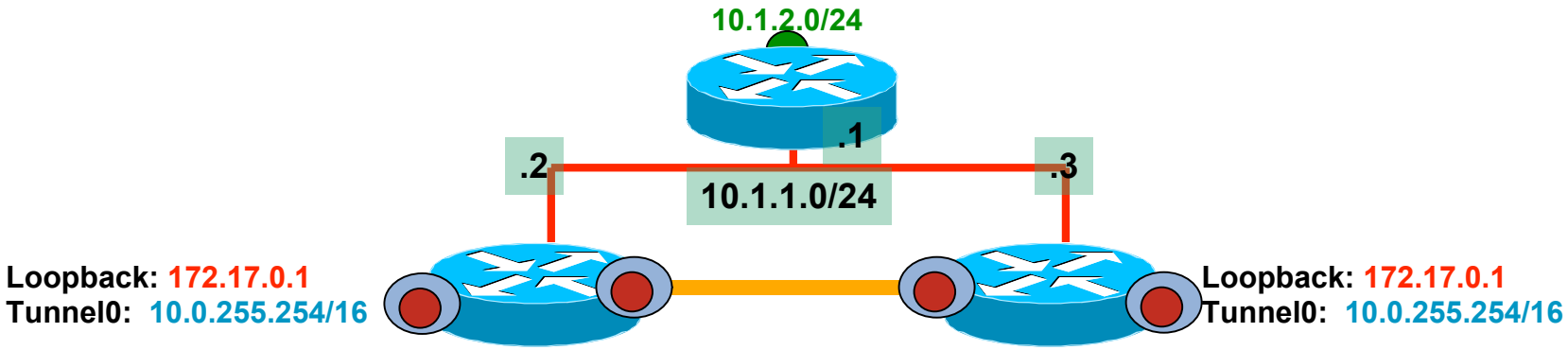
```
interface Tunnel1
 ip address 10.1.3.2 255.255.255.0
 no ip redirects
 ip mtu 1400
 ip nhrp network-id 1
 ip nhrp redirect
 ip nhrp map 10.1.3.3 10.1.0.3
 tunnel source FastEthernet0/1
end
```

Same network ID as Tunnel0 !!

Send indirection notifications



Routing across hubs

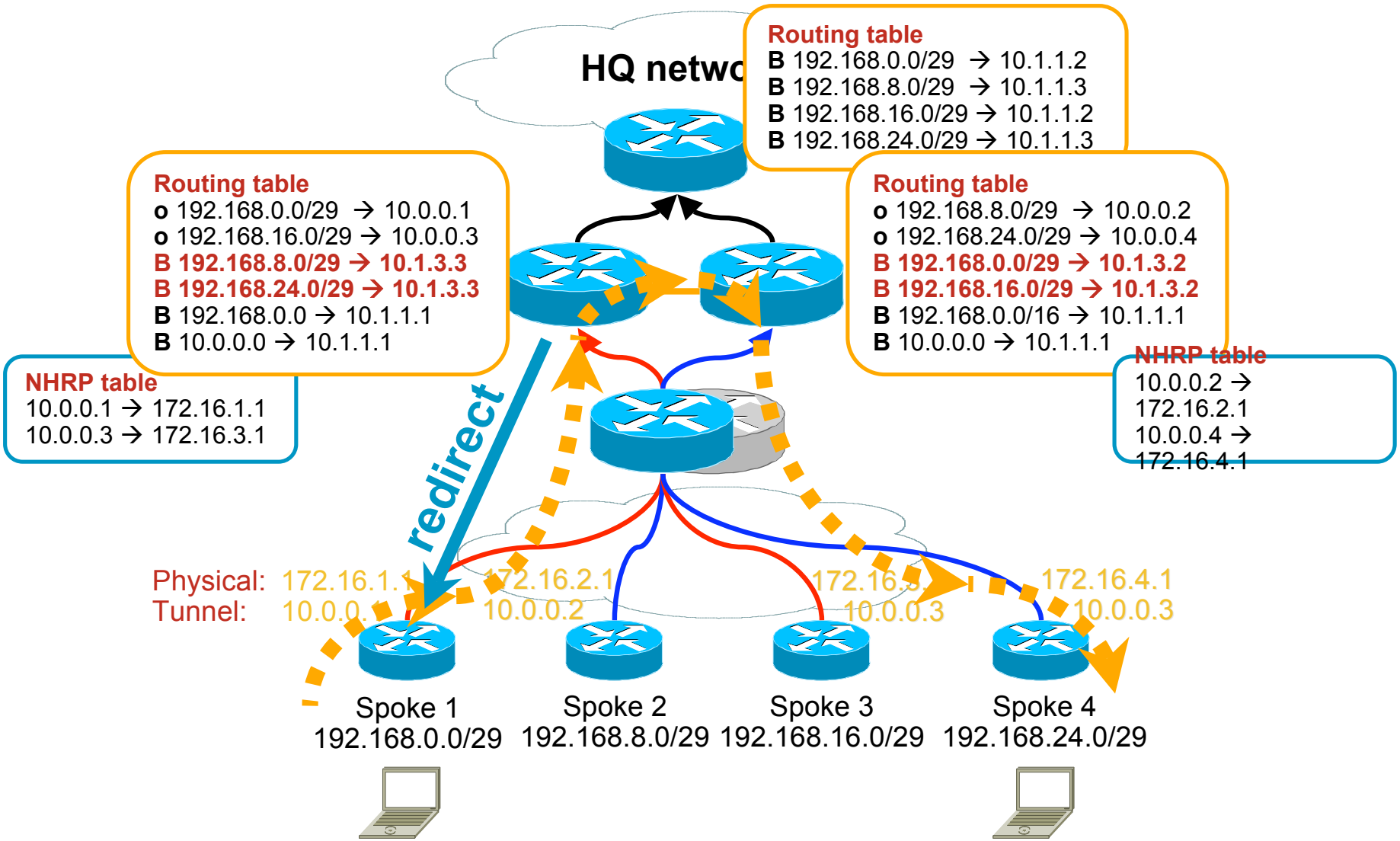



```
router bgp 1
  neighbor 10.1.3.3 remote-as 1
  neighbor 10.1.3.3 next-hop-self
```

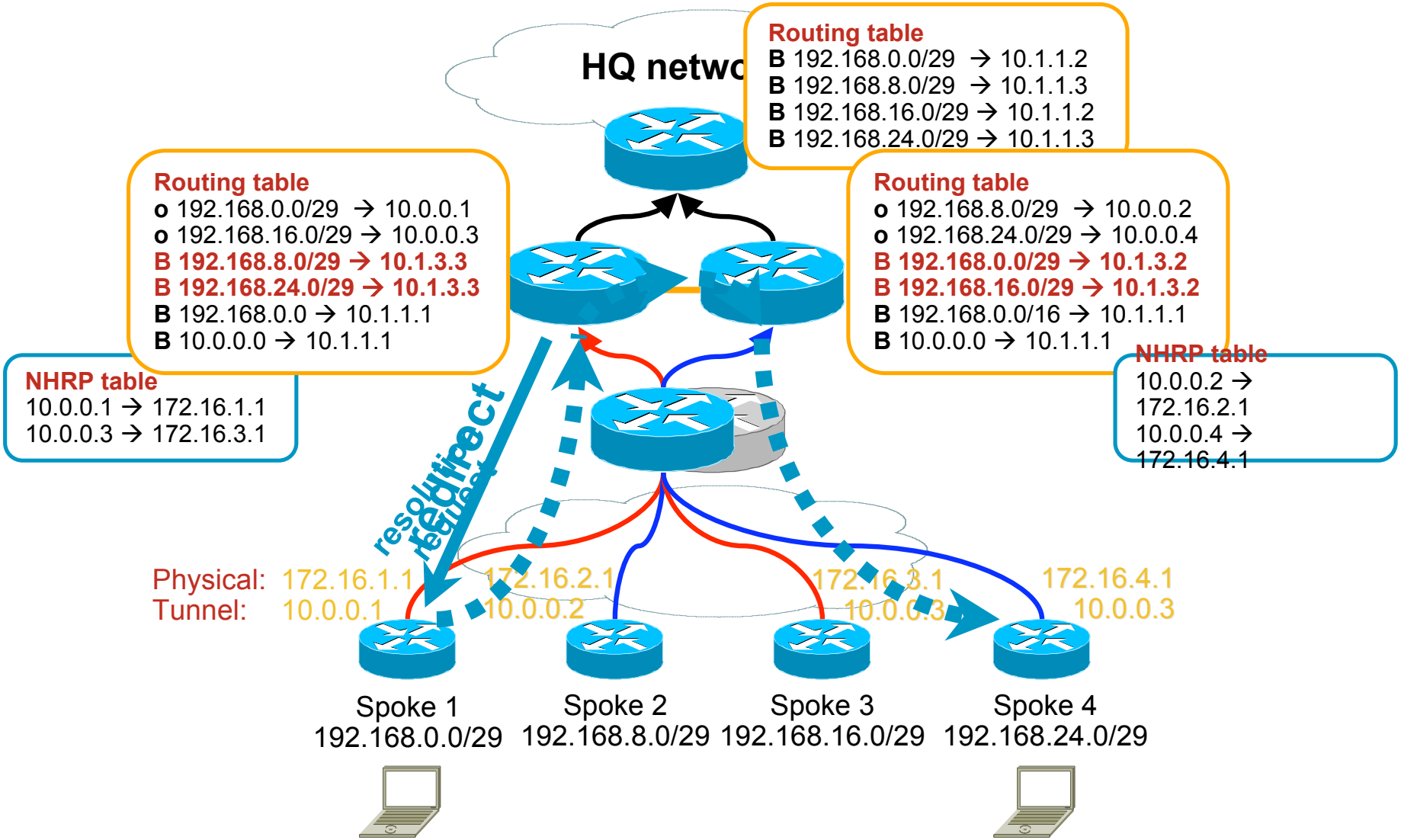
```
router bgp 1
  neighbor 10.1.3.2 remote-as 1
  neighbor 10.1.3.2 next-hop-self
```

- Hubs exchange their ODR information directly via BGP
- The exchange occurs over the inter-hub DMVPN

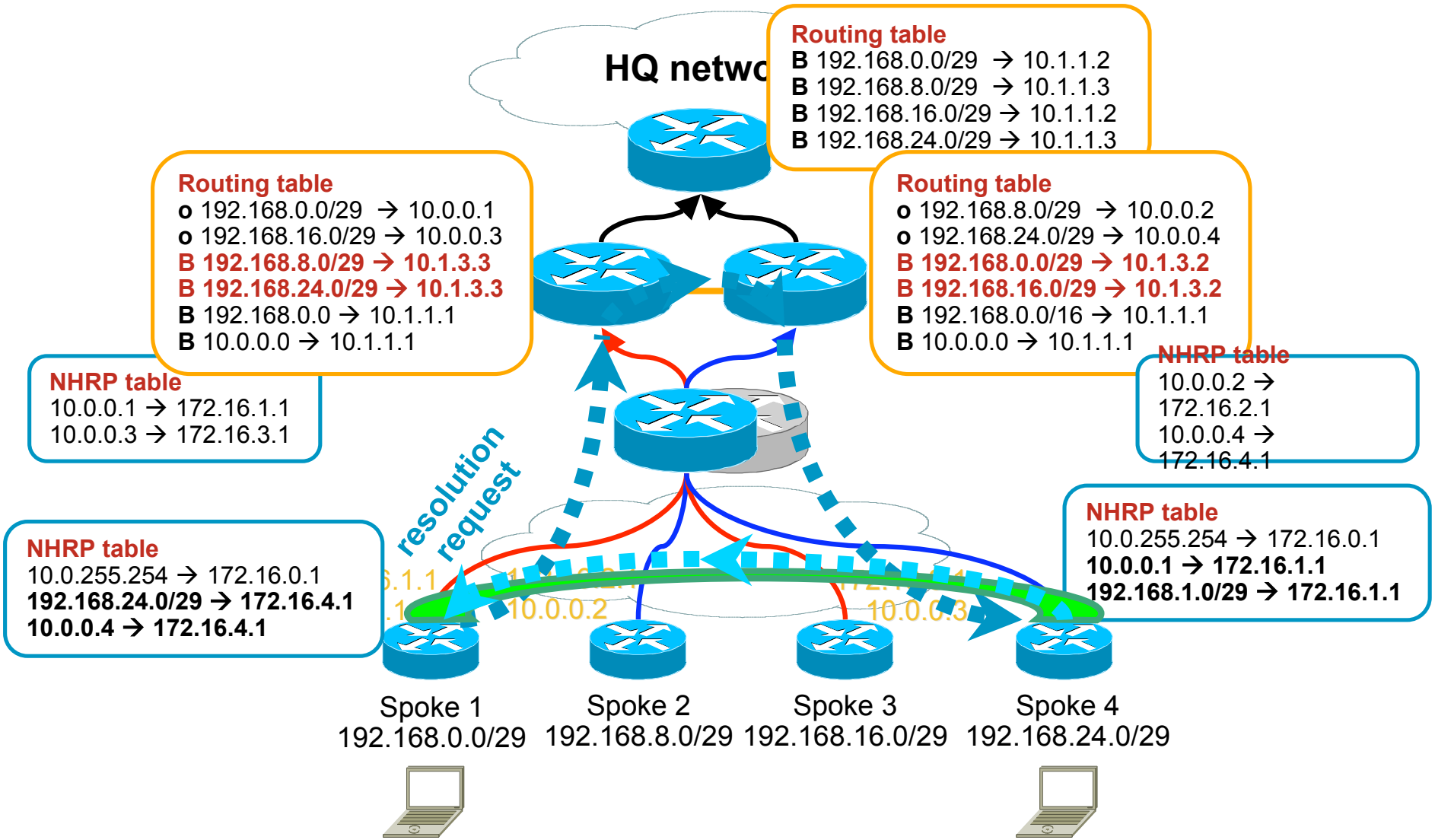
Hub&Spoke packet flow



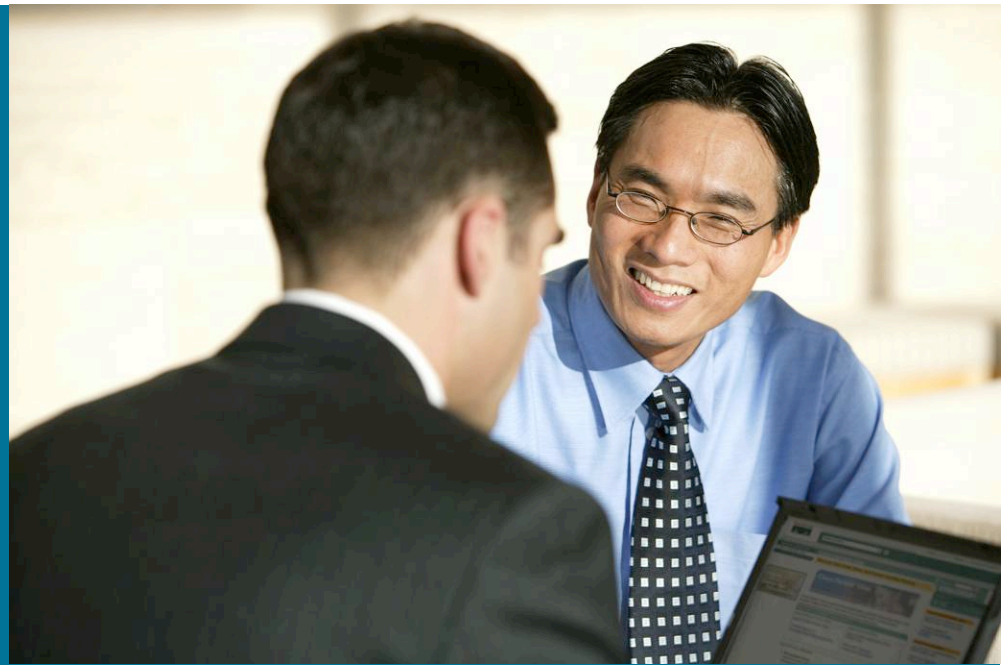
Hub&Spoke packet flow



Hub&Spoke packet flow



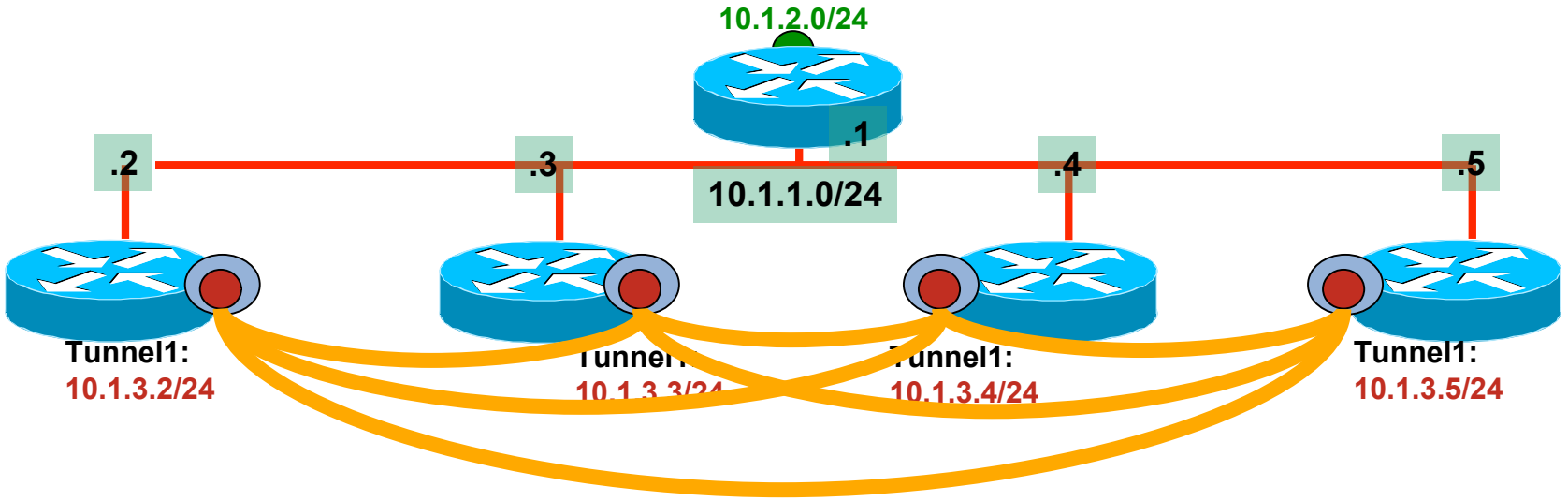
Adding hubs



Linking the hubs – option 1

```
interface Tunnel1
 ip address 10.1.3.2 255.255.255.0
 . . .
 ip nhrp map 10.1.3.3 10.1.0.3
 ip nhrp map 10.1.3.4 10.1.0.4
 ip nhrp map 10.1.3.5 10.1.0.5
 . . .
end
```

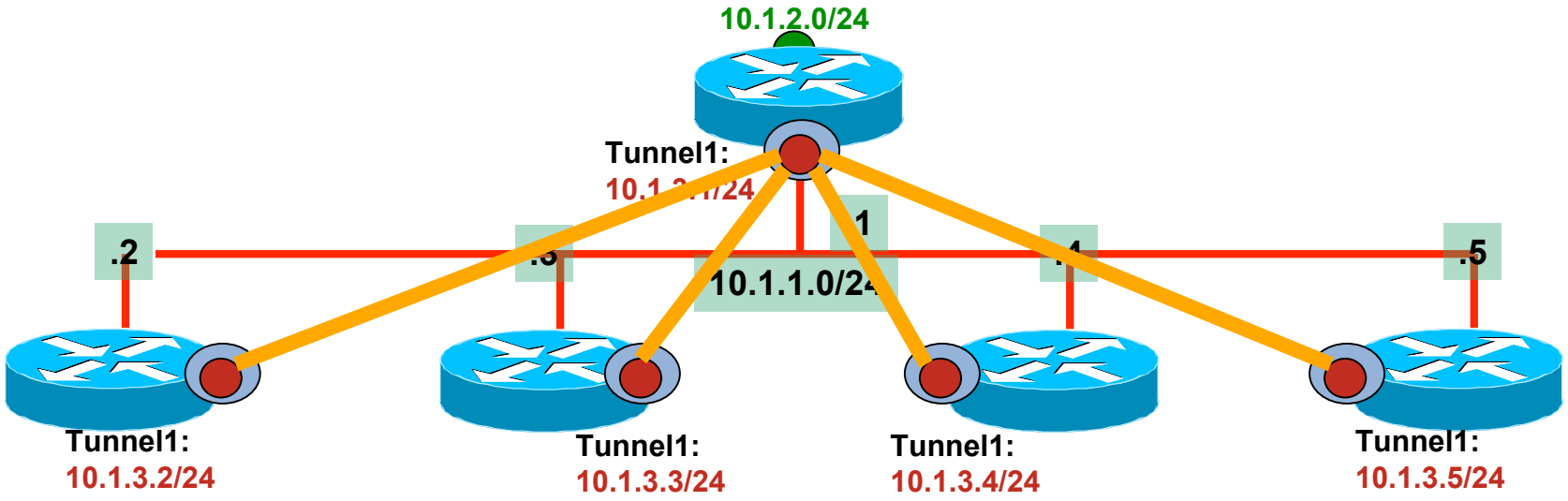
Create a manual full mesh
Do the same with BGP…



Linking the hubs – option 2

```
interface Tunnel1
 ip address 10.1.3.2 255.255.255.0
 ip nhrp network-id 1
 ip nhrp redirect
 ip nhrp map 10.1.3.1 10.1.0.1
 ip nhrp nhs 10.1.3.1
end
```

Use the edge router as NHRP hub
Use the edge as a RR



Large Scale Design Summary

- Virtually limitless scaling w/ automatic load management
- Load balancing AND resilience
- Multiply performances by number of hubs
  - Tunnel creation rate, speed, max SA's
- Resilience in N+1
- No need to touch the hubs while adding a spoke
- All spokes have the same configuration
- New hubs can be added/removed on the fly
  - BGP needs to be told about the new hub
  - EIGRP may be used instead of BGP ➔ full automatic

Session Summary

DMVPN phase 3 NHRP/CEF Enhancements 12.4(6)T onward

For your reference

Previous Limitation	New Feature & Associated Benefits
Large routing tables at spokes sometimes caused network instability.	**Shortcut switching introduced** ▪ Route summarization now possible ▪ Higher scalability
Delays in setting up voice calls between spokes.	**Packets CEF switched via hub** ▪ Reduced latency during call setup.
Complex interconnection of Hubs to expand DMVPN Spoke-to-Spoke Networks.	**NHRP resolution requests forwarding** ▪ Simplified hub network design ▪ Improved resiliency. ▪ Failure of single hub will not affect rest of DMVPN network.
NAT/PAT not possible in spoke-spoke designs	**NAT and static PAT now supported**

Shortcut switching
Routing protocols revisited

- OSPF does not bring anything new
  - Same requirements as in phase 2
- EIGRP can be tuned to summarize routes to spokes
  - Number of neighbors does not increase
- ODR can now be used for spoke-to-spoke configs
  - 1200 neighbors possible
- RIP passive can now be used for spoke-to-spoke
  - 1500 neighbors possible
- Different protocols can be used between hubs and between hub-spoke

Troubleshooting enhancements

- IOS 12.4(9)T offers significant troubleshooting enhancements to IPsec VPN's
- There are more to come…

  Newer images will allow even better troubleshooting

  ➔ better support can be offered

  ➔ more audacious network can be deployed and fixed

  ➔ foresee RAM and Flash ☺

Meet the Experts
Security

- Andres Gasson
  Consulting Systems Engineer
- Christophe Paggen
  Technical Marketing Engineer
- Eric Vyncke
  Distinguished Consulting Engineer
- Erik Lenten
  Technical Marketing Engineer
- Fredéric Detienne
  CA Technical Leader
- Luc Billot
  Consulting Engineer






Meet the Experts
Security

- Michael Behringer
  Distinguished System Engineer

- Olivier Dupont
  Corporate Dev Consulting Engineer

- Peter Matthews
  Technical Marketing Engineer

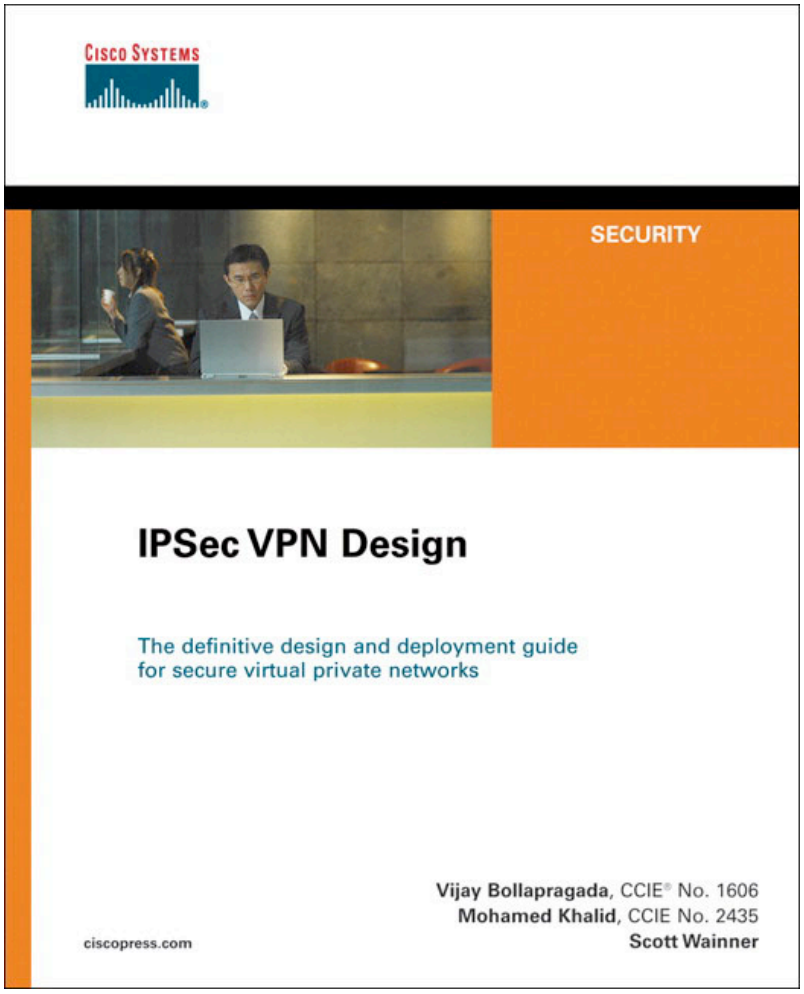
- Scott Wainner
  Distinguished System Engineer

- Steinthor Bjarnason
  Consulting Engineer

Recommended Reading

BRKSEC - 3006

- IPSec VPN Design



Available in the Cisco Company Store

Q and A