CHAPTER 6

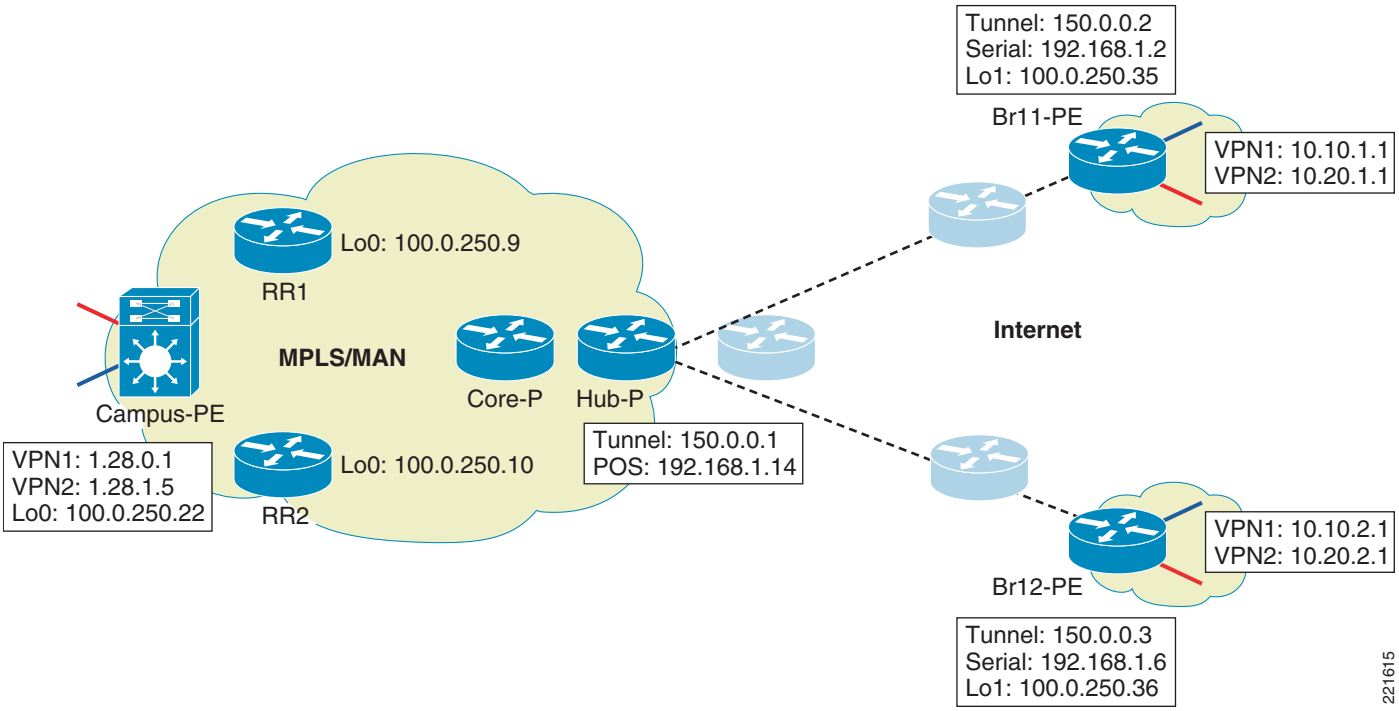
WAN Edge—MPLS VPN over DMVPN—2547oDMVPN (Hub and Spoke Only)

DMVPN provides two key advantages for extending MPLS VPNs to the branches, bulk encryption and, more importantly, a scalable overlay model. Since the assumption here is that the branches in this deployment are connected to the hub through a Layer 3 SP service, a tunneled model using GRE is needed to extend MPLS to the branches. Coupled with the fact that there is large number of existing DMVPN deployments, this solution becomes an attractive deployment option.

DMVPN allows the hub to have a single multipoint GRE tunnel interface to support large numbers of spokes. The spokes can be point-to-point or multipoint GRE tunnels depending on the requirement of direct spoke-to-spoke communication. Spoke-to-Spoke Communication (via Hub) discusses the advantages of point-to-point GRE tunnels at the spokes in the context of the current implementation of MPLS VPN over DMVPN.

To seamlessly extend the enterprise MPLS/Layer 3VPN MAN network to the remote branches, the WAN edge router (also the DMVPN hub in this case) should be a P device to label switching packets between the hub and the branches. As shown in Figure 6-1, the WAN hub router acts as a MPLS/Layer 3VPN P router to establish the LDP neighbor relationship and label switch packet with branch routers which act as a MPL3/Layer 3VPN PE router. The single IGP process is running on the entire enterprise MAN/WAN network to enable the branch routers to establish the MP-iBGP session with RRs in the enterprise MPLS MAN network.

Figure 6-1 2547oDMVPN Deployment



Platforms

Only 7200VXR is supported as the hub router. NPE-G1/G2 is recommended, along with VAM2/VAM2+/VSA modules for encryption. ISRs are recommended as spoke devices. The following images were tested in the lab:

- 7200VXR with NPE-G1/G2—12.4(11)T1
- ISRs—12.4(11)T1

Hub and Spoke Communication

The hub and spoke communication is straightforward as it follows the normal P-PE forwarding mechanism. The example below gives the typical configurations for a DMVPN hub router used as a MPLS P device and a DMVPN spoke router used as a MPLS PE device.

Hub11:

```
!
hostname ngwan-hub11
!
crypto isakmp policy 1
 encr 3des
 authentication pre-share
crypto isakmp key Cisco123 address 0.0.0.0 0.0.0.0
crypto isakmp keepalive 10 5
!
!
```

```
crypto ipsec transform-set T1 esp-3des
 mode transport
!
crypto ipsec profile P1
 set transform-set T1
!
interface Loopback1
 ip address 100.0.250.33 255.255.255.255
 ip pim sparse-mode
!
interface Tunnel1
 bandwidth 1500
 ip address 150.0.0.1 255.255.255.0
 no ip redirects
 ip mtu 1368
 ip pim sparse-mode
 ip nhrp map multicast dynamic
 ip nhrp network-id 1
 ip nhrp cache non-authoritative
 ip ospf network point-to-multipoint
 ip ospf priority 100
 load-interval 30
 mpls ip
 tunnel source 192.168.1.14
 tunnel mode gre multipoint
 tunnel key 777
 tunnel protection ipsec profile P1
!
router ospf 1
 mpls traffic-eng router-id Loopback1
 mpls traffic-eng area 0
 router-id 100.0.250.33
 log-adjacency-changes
 network 100.0.250.33 0.0.0.0 area 0
 network 100.0.0.0 0.0.255.255 area 0
 network 150.0.0.0 0.0.0.255 area 0
!
router ospf 100
 log-adjacency-changes
network 192.168.1.12 0.0.0.3 area 3
!
ip pim ssm range 1
access-list 1 permit 239.232.0.0 0.0.255.255
```

Spoke br11:

```
hostname ngwan-br11
!
ip vrf vpn1
 rd 100:10
 route-target export 100:110
 route-target import 100:110
 mdt default 239.232.1.1
 mdt data 239.232.1.128 0.0.0.127 threshold 10
!
ip vrf vpn2
 rd 100:20
 route-target export 100:120
 route-target import 100:120
 mdt default 239.232.2.1
 mdt data 239.232.2.128 0.0.0.127 threshold 10
!
ip multicast-routing
ip multicast-routing vrf vpn1
```

```
ip multicast-routing vrf vpn2
!
crypto isakmp policy 1
 encr 3des
 authentication pre-share
crypto isakmp key Cisco123 address 0.0.0.0 0.0.0.0
crypto isakmp keepalive 10 5
!
!
crypto ipsec transform-set T1 esp-3des
 mode transport
!
crypto ipsec profile P1
 set transform-set T1
!
interface Loopback1
 ip address 100.0.250.35 255.255.255.255
 ip pim sparse-mode
!
interface Tunnel1
 bandwidth 1500
 ip address 150.0.0.2 255.255.255.0
 no ip redirects
 ip mtu 1368
 ip pim sparse-mode
 ip nhrp map 150.0.0.1 192.168.1.14
 ip nhrp map multicast 192.168.1.14
 ip nhrp network-id 1
 ip nhrp nhs 150.0.0.1
 ip nhrp cache non-authoritative
 ip ospf network point-to-multipoint
 ip ospf priority 0
 load-interval 30
 mpls ip
 qos pre-classify
 tunnel source 192.168.1.2
 tunnel mode gre multipoint
 tunnel key 777
 tunnel protection ipsec profile P1
!
interface POS5/0
 ip address 192.168.1.2 255.255.255.252
 load-interval 30
 crc 32
 clock source internal
 service-policy output wan-edge
!
router ospf 100
 log-adjacency-changes
 network 192.168.1.0 0.0.0.3 area 3
!
router ospf 1
 mpls traffic-eng router-id Loopback1
 mpls traffic-eng area 0
 router-id 100.0.250.35
 log-adjacency-changes
 network 100.0.250.35 0.0.0.0 area 0
 network 150.0.0.0 0.0.0.255 area 0
 network 150.0.1.0 0.0.0.255 area 0
!
router bgp 64512
 bgp log-neighbor-changes
 neighbor RRs peer-group
 neighbor RRs remote-as 64512
```

```
 neighbor RRs update-source Loopback1
 neighbor 100.0.250.9 peer-group RRs
 neighbor 100.0.250.10 peer-group RRs
 !
 address-family ipv4
 neighbor 100.0.250.9 activate
 neighbor 100.0.250.10 activate
 no auto-summary
 no synchronization
 exit-address-family
 !
address-family ipv4 vrf vpn2
 redistribute connected
 no synchronization
 exit-address-family
 !
 address-family ipv4 vrf vpn1
 redistribute connected
 no synchronization
 exit-address-family
 !
ip pim ssm range 1
access-list 1 permit 239.232.0.0 0.0.255.255

ngwan-br11#sh ip bgp vpn vrf vpn1 1.28.0.1
ngwan-br11#sh ip bgp vpnv4 vrf vpn1 1.28.0.1
BGP routing table entry for 100:10:1.28.0.0/30, version 269
Paths: (1 available, best #1, table vpn1)
  Not advertised to any peer
  Local, imported path from 100:180:1.28.0.0/30
    100.0.250.22 (metric 69) from 100.0.250.9 (100.0.250.9)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      Extended Community: SoO:100:110 RT:100:110
      Originator: 100.0.250.22, Cluster list: 0.0.0.1
      mpls labels in/out nolabel/21
ngwan-br11#sh ip cef vrf vpn1 1.28.0.1
1.28.0.0/30, version 12, epoch 0
0 packets, 0 bytes
  tag information set
    local tag: VPN-route-head
    fast tag rewrite with Tu1, 150.0.0.1, tags imposed: {63 21}
  via 100.0.250.22, 0 dependencies, recursive
    next hop 150.0.0.1, Tunnel1 via 100.0.250.22/32
    valid adjacency
    tag rewrite with Tu1, 150.0.0.1, tags imposed: {63 21}
ngwan-br11#

ngwan-br11#ping vrf vpn1 1.28.0.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 1.28.0.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/4 ms
ngwan-br11#
```

Notice that there are two labels assigned to the destination address in one of the VRFs within the enterprise MPLS MAN. This is exactly the expected behavior for using MPLS to extend the network segmentation to remote branches. Further test with ping shows the LSP and successfully established.

Spoke-to-Spoke Communication (via Hub)

In branch aggregation scenarios, while most traffic is typically between the hub and the spokes, there is always a requirement for spoke-to-spoke communication. VOIP traffic is a good example of peer-to-peer traffic. In a normal DMVPN scenario this would have been achieved by dynamically creating direct spoke-to-spoke tunnels. While there are obvious advantages to this approach, there are issues as well:

- Depending on the underlying physical connectivity, in certain cases the spoke-to-spoke path may not necessarily be better then the spoke-hub-spoke path which could be a problem for latency sensitive traffic such as VOIP.
- There is a possibility of receiving out-of-order packets as during the initial tunnel setup time the traffic traverses the hub, but once the spoke-to-spoke tunnel is setup it switches it over.
- Depending on the number of spoke-to-spoke tunnels that need to created/maintained simultaneously, this can put scale pressures on the spoke router especially if it is a low-end CPE.

Additionally, the MPLS network requires packets to be label switched all the way between source PEs and destination PEs. Running MPLS over DMVPN tunnels makes the remote branch router a full function PE router, which means label imposition is done in the branch router and label switching must be performed all the way between spokes. This requirement make the direct spoke-spoke communication impossible due to the lack of a label allocation mechanism on the dynamically created spoke-spoke tunnels. However, label switching between spoke PE routers can easily be done if spoke-hub-spoke switching path is implemented. With this approach, the hub router act as a MPLS P router, maintains the LDP neighbor relationship, and exchanges label allocation information with all spoke routers. The hub router label switches the packets in-and-out the mGRE interface between the spokes. Since it is done in the fast path (whether encrypted or not), there should be minimal performance implications other then the increase in the hub traffic.

While this solution breaks the benefit of dynamically building spoke-to-spoke tunnels, it provides a acceptable and often more deterministic path for spoke-to-spoke communications and meets the segmentation requirement. It is a very attractive solution when the large enterprise needs to extend their MPLS-segmented data center or large campus to remote branches.

Configuration Example:

The following example shows the two VPNs in the two remote branches (br11 and br12) communicate to each other via the DMVPN hub router (hub11). The router hub11's configuration is the same as shown in Hub and Spoke Communication. The VPN naming and address scheme, along with the address of hub router and branch routers, are illustrated in Figure 6-1.

Br11:

```
hostname ngwan-br11
!
ip vrf vpn1
 rd 100:10
 route-target export 100:110
 route-target import 100:110
 mdt default 239.232.1.1
 mdt data 239.232.1.128 0.0.0.127 threshold 10
!
ip vrf vpn2
 rd 100:20
 route-target export 100:120
 route-target import 100:120
 mdt default 239.232.2.1
 mdt data 239.232.2.128 0.0.0.127 threshold 10
!
ip multicast-routing
```

```
ip multicast-routing vrf vpn1
ip multicast-routing vrf vpn2
!
crypto isakmp policy 1
 encr 3des
 authentication pre-share
crypto isakmp key Cisco123 address 0.0.0.0 0.0.0.0
crypto isakmp keepalive 10 5
!
!
crypto ipsec transform-set T1 esp-3des
 mode transport
!
crypto ipsec profile P1
 set transform-set T1
!
interface Loopback1
 ip address 100.0.250.35 255.255.255.255
 ip pim sparse-mode
!
interface Tunnel1
 bandwidth 1500
 ip address 150.0.0.2 255.255.255.0
 no ip redirects
 ip mtu 1368
 ip pim sparse-mode
 ip nhrp map 150.0.0.1 192.168.1.14
 ip nhrp map multicast 192.168.1.14
 ip nhrp network-id 1
 ip nhrp nhs 150.0.0.1
 ip nhrp cache non-authoritative
 ip ospf network point-to-multipoint
 ip ospf priority 0
 load-interval 30
 mpls ip
 qos pre-classify
 tunnel source 192.168.1.2
 tunnel mode gre multipoint
 tunnel key 777
 tunnel protection ipsec profile P1
!
interface POS5/0
 ip address 192.168.1.2 255.255.255.252
 load-interval 30
 crc 32
 clock source internal
 service-policy output wan-edge
!
router ospf 100
 log-adjacency-changes
 network 192.168.1.0 0.0.0.3 area 3
!
router ospf 1
 mpls traffic-eng router-id Loopback1
 mpls traffic-eng area 0
 router-id 100.0.250.35
 log-adjacency-changes
 network 100.0.250.35 0.0.0.0 area 0
 network 150.0.0.0 0.0.0.255 area 0
 network 150.0.1.0 0.0.0.255 area 0
!
router bgp 64512
 bgp log-neighbor-changes
 neighbor RRs peer-group
```

```
 neighbor RRs remote-as 64512
 neighbor RRs update-source Loopback1
 neighbor 100.0.250.9 peer-group RRs
 neighbor 100.0.250.10 peer-group RRs
 !
 address-family ipv4
 neighbor 100.0.250.9 activate
 neighbor 100.0.250.10 activate
 no auto-summary
 no synchronization
 exit-address-family
 !
address-family ipv4 vrf vpn2
 redistribute connected
 no synchronization
 exit-address-family
 !
 address-family ipv4 vrf vpn1
 redistribute connected
 no synchronization
 exit-address-family
 !
ip pim ssm range 1
access-list 1 permit 239.232.0.0 0.0.255.255
```

Br12:

```
!
hostname ngwan-br12


ip vrf vpn1
 rd 100:10
 route-target export 100:110
 route-target import 100:110
 mdt default 239.232.1.1
 mdt data 239.232.1.128 0.0.0.127 threshold 10
!
ip vrf vpn2
 rd 100:20
 route-target export 100:120
 route-target import 100:120
 mdt default 239.232.2.1
 mdt data 239.232.2.128 0.0.0.127 threshold 10
!
mpls label protocol ldp
!
crypto isakmp policy 1
 encr 3des
 authentication pre-share
crypto isakmp key Cisco123 address 0.0.0.0 0.0.0.0
crypto isakmp keepalive 10 5
!
!
crypto ipsec transform-set T1 esp-3des
 mode transport
!
crypto ipsec profile P1
 set transform-set T1
!
interface Loopback1
 ip address 100.0.250.36 255.255.255.255
```

```
 ip pim sparse-mode
!
interface Tunnel1
 ip address 150.0.0.3 255.255.255.0
 no ip redirects
 ip mtu 1368
 ip pim sparse-mode
 ip nhrp authentication spe
 ip nhrp map 150.0.0.1 192.168.1.14
 ip nhrp map multicast 192.168.1.14
 ip nhrp network-id 1
 ip nhrp holdtime 360
 ip nhrp nhs 150.0.0.1
 ip nhrp cache non-authoritative
 ip ospf network point-to-multipoint
 ip ospf priority 0
 load-interval 30
 mpls label protocol ldp
 mpls ip
 qos pre-classify
 tunnel source Serial2/0
 tunnel mode gre multipoint
 tunnel key 777
 tunnel protection ipsec profile P1

interface Serial2/0
 ip address 192.168.1.6 255.255.255.252
 load-interval 30
 dsu bandwidth 44210
 service-policy output wan-edge
!
router ospf 100
 log-adjacency-changes
 network 192.168.1.4 0.0.0.3 area 3
!
router ospf 1
 mpls traffic-eng router-id Loopback1
 mpls traffic-eng area 0
 log-adjacency-changes
 network 100.0.250.36 0.0.0.0 area 0
 network 150.0.0.0 0.0.0.255 area 0
 network 150.0.1.0 0.0.0.255 area 0
!
router bgp 64512
 bgp log-neighbor-changes
 neighbor RRs peer-group
 neighbor RRs remote-as 64512
 neighbor RRs update-source Loopback1
 neighbor 100.0.250.9 peer-group RRs
 neighbor 100.0.250.10 peer-group RRs
 !
 address-family ipv4
 neighbor 100.0.250.9 activate
 neighbor 100.0.250.10 activate
 no auto-summary
 no synchronization
 exit-address-family
 !
 address-family vpnv4
 neighbor RRs send-community extended
 neighbor 100.0.250.9 activate
 neighbor 100.0.250.10 activate
 exit-address-family
```

```
 !
 address-family ipv4 vrf vpn2
 redistribute connected
 no synchronization
 exit-address-family
 !
 address-family ipv4 vrf vpn1
 redistribute connected
 no synchronization
 exit-address-family
!
ip pim ssm range 1!
access-list 1 permit 239.232.0.0 0.0.255.255
!
```

The following command shows the seamless integrating of remote spokes with the Enterprise MPLS network. VPN routes are distributed by RRs in the enterprise MPLS network via MP-iBGP:

```
ngwan-br11#sh ip bgp vpnv4 vrf vpn1 10.10.2.1
BGP routing table entry for 100:10:10.10.2.0/24, version 3267
Paths: (2 available, best #2, table vpn1)
  Not advertised to any peer
  Local
    100.0.250.36 (metric 133) from 100.0.250.10 (100.0.250.10)
      Origin incomplete, metric 0, localpref 100, valid, internal
      Extended Community: RT:100:110
      Originator: 100.0.250.36, Cluster list: 0.0.0.1
      mpls labels in/out nolabel/76
  Local
    100.0.250.36 (metric 133) from 100.0.250.9 (100.0.250.9)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      Extended Community: RT:100:110
      Originator: 100.0.250.36, Cluster list: 0.0.0.1
      mpls labels in/out nolabel/76
ngwan-br11#
```

The following command shows two labels are allocated for the VPN routes in spoke routers and hub is in the middle of the LSP:

```
ngwan-br11#sh ip cef vrf vpn1 10.10.2.1 detail
10.10.2.0/24, version 423, epoch 0
0 packets, 0 bytes
  tag information set
    local tag: VPN-route-head
    fast tag rewrite with Tu1, 150.0.0.1, tags imposed: {78 76}
  via 100.0.250.36, 0 dependencies, recursive
    next hop 150.0.0.1, Tunnel1 via 100.0.250.36/32
    valid adjacency
    tag rewrite with Tu1, 150.0.0.1, tags imposed: {78 76}
ngwan-br11#
```

The following command shows the VPN traffic between spokes are label switched via hub router:

```
ngwan-br11#traceroute vrf vpn1 10.10.2.1

Type escape sequence to abort.
Tracing the route to 10.10.2.1

  1 150.0.0.1 [MPLS: Labels 78/76 Exp 0] 4 msec 0 msec 0 msec
  2 10.10.2.1 4 msec *  8 msec
ngwan-br11#
```

Connecting to the Core MPLS Network

The core MPLS network and the DMVPN-based MPLS network are fully integrated together when the DMVPN hub router act as a MPLS P router. The normal MPLS/LDP configuration applies here when it connects to the enterprise MPLS core networks.

Building Redundancy

Redundancy can be built at various points within the networks:

- Use of multiple routers and HSRP/GLBP with the Enhanced Object Tracking at the branch
- Multiple hub routers at the headend
- Hub connects to multiple core routers

Ideally all three should be used to provide a robust end-to-end connectivity solution. For cost reasons, it may not be feasible to have two routers at every branch; however it should be implemented at least at the large branches when the high-available requirement is a must. While the loss of a spoke router may not be critical to the network, loss of the hub may mean loss of multiple sites and hence more critical. Thus each spoke should be connected to multiple hubs via GRE tunnels which are maintained in active/standby state by controlling the route metrics.

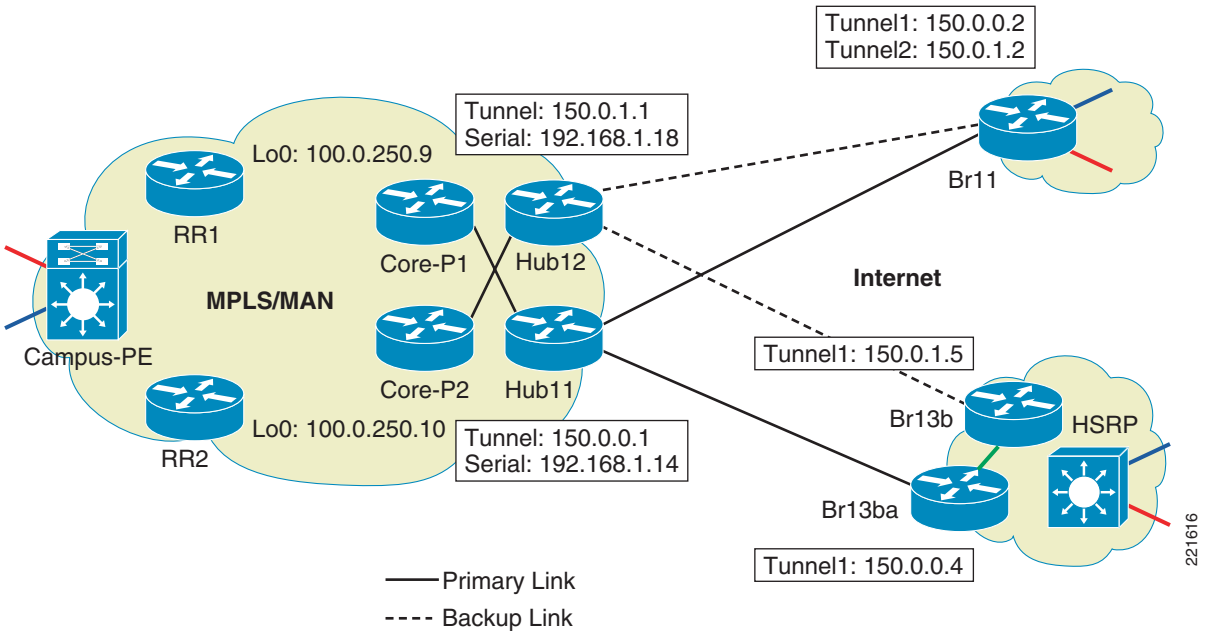
While it may seem desirable to keep both (or all) the hubs as active/active and allow the traffic to be load balanced, we do not see any true advantage in doing so. Keeping the tunnels as active/standby allows the hubs to be better engineered for steady state performance. It also reduces the load on the spoke routers while allowing more deterministic traffic path characteristics.

The various options discussed here will are illustrated using an example.

Example:

In the following example (Figure 6-2), two remote spoke sites br11 and br13 (br13a, br13b) are connected to two hubs (hub11 and hub12). Br11 is considered as a single-tier branch, where it has only one WAN router with two DMVPN tunnels terminated at hub11 and hub12. Hub11 is the primary hub and hub12 is the backup. Br13 is considered as a Dual-tier branch, representing a large branch with two WAN routers with the dual DMVPN tunnel which provides the WAN link redundancy. HSRP with enhanced object tracking is used to provide the network resiliency for the clients on the branch.

Figure 6-2 2547oDMVPN Redundancy



Note The encryption configuration is shown here as an example and standard best practices for IPSec should be followed in an actual deployment.

Below is the configuration example for br11, which use dual tunnels on the same router for the WAN redundancy. OSPF cost is used to tune the routing matrix to select which hub is the primary one.

Spoke B21a:

```
hostname ngwan-br11
!
crypto isakmp policy 1
 encr 3des
 authentication pre-share
crypto isakmp key Cisco123 address 0.0.0.0 0.0.0.0
crypto isakmp keepalive 10 5
!
crypto ipsec transform-set T1 esp-3des
 mode transport
!
crypto ipsec profile P1
 set transform-set T1
!
interface Tunnel1
 bandwidth 1500
 ip address 150.0.0.2 255.255.255.0
 no ip redirects
 ip mtu 1368
 ip pim sparse-mode
 ip nhrp authentication ngwan
 ip nhrp map 150.0.0.1 192.168.1.14
 ip nhrp map multicast 192.168.1.14
 ip nhrp network-id 1
 ip nhrp nhs 150.0.0.1
 ip nhrp cache non-authoritative
 ip ospf network point-to-multipoint
```

```
 ip ospf priority 0
 load-interval 30
 mpls ip
 qos pre-classify
 tunnel source 192.168.1.2
 tunnel mode gre multipoint
 tunnel key 777
 tunnel protection ipsec profile P1 shared
!!
interface Tunnel2
 ip address 150.0.1.2 255.255.255.0
 no ip redirects
 ip mtu 1368
 ip pim sparse-mode
 ip nhrp authentication ngwan
 ip nhrp map 150.0.0.1 192.168.1.18
 ip nhrp map multicast 192.168.1.18
 ip nhrp network-id 2
 ip nhrp nhs 150.0.1.1
 ip nhrp cache non-authoritative
 ip ospf network point-to-multipoint
 ip ospf cost 100
 ip ospf priority 0
 load-interval 30
 qos pre-classify
 mpls ip
 tunnel source POS5/0
 tunnel mode gre multipoint
 tunnel key 888
 tunnel protection ipsec profile P1 shared

!
```

Spokes Br13a and Br13b:

Spoke Br13a has the DMVPN tunnel connected to Hub11 while br13b's DMVPN tunnel connects to hub12. Br13a is selected as the HSRP active router (higher priority and preemption enabled) and br13b is configured as the standby, This also makes the hub11 as the primary hub and hub12 as the backup hub for this remote site. Working together with the HSRP, Enhanced object tracking is also configured to track two objects, the line protocol on the tunnel interface and the reachability to the hub itself (tunnel destination address). The latter is a more reliable object to track since there are scenarios where the line protocol on the tunnel interface may not go down.

Since the rest of the configuration is similar to other spokes, only the HSRP relevant configuration is shown here:

```
hostname ngwan-br13a
!
track 1 interface Tunnel1 line-protocol
track 2 ip route 192.168.1.14.255.255.255.252 reachability
!
hostname ngwan-br13a
!
track 1 interface Tunnel1 line-protocol
 delay up 50
track 2 ip route 192.168.1.14.255.255.255.252 reachability
 delay up 50
!
interface Vlan110
 ip vrf forwarding vpn1
 ip address 10.10.4.2 255.255.255.0
 standby 1 ip 10.10.4.3
```

```
 standby 1 timers 1 3
 standby 1 priority 105
 standby 1 preempt
 standby 1 track 1 decrement 10
 standby 1 track 2 decrement 10
!
```

```
hostname ngwan-br13b
!
track 1 interface Tunnel1 line-protocol
 delay up 50
track 2 ip route 192.168.1.18.255.255.255.252 reachability
 delay up 50
!
interface Vlan110
 ip vrf forwarding vpn1
 ip address 10.10.4.2 255.255.255.0
 standby 1 ip 10.10.4.3
 standby 1 timers 1 3
 standby 1 preempt
 standby 1 track 1 decrement 10
 standby 1 track 2 decrement 10
!
```

Understanding Convergence

We focus on traffic convergence for the hub <-> spoke traffic. As seen in the redundancy section, there are two major backup options available for two types of branch architecture—single-tier branch which having multiple tunnels originating on the same router and dual-tier branch which use two separate routers with HSRP at the branches.

Single-Tier Branches—Backup Tunnel on the Same Router

When the backup tunnel is on the same router, the traffic convergence is primarily dependent on the IGP. By keeping the default timers, following test has conducted to know the network convergence time.

The failure/recovery is simulated by shut/no shut of the link connecting the hub to the SP (doing it on the SP router).

Table 6-1 Convergence When the Primary Tunnel is Down

Iteration	1		2		3	
	Down	Up	Down	Up	Down	Up
Spoke-to-hub traffic	5.5s	0s	5s	0s	4s	0s
Hub-to-spoke traffic	6s	0s	5s	0s	5s	1s

Table 6-2 Convergence When the Primary Router is Reloaded

Iteration	1		2		3	
	Down	Up	Down	Up	Down	Up
Spoke-to-hub traffic	1.5s	0s	2s	4s	2s	1s
Hub-to-spoke traffic	15s	0s	1.5s	3s	2s	0.6s

As can be seen tuning down the BGP timers provides a much faster convergence. While tuning down to (1s, 3s) provides the best end-to-end convergence performance, it is not recommended in a scaled environment due to the additional overhead on the hub router. At the very least the performance impact needs to be studied in a scaled environment before tuning it down to such levels.

Dual-Tier Branches—Backup Tunnel on Different Routers

With two WAN routers used in the dual-tier branch, the primary and backup tunnels are configured on two different WAN routers (HSRP enabled). Since the branch routers is actually a MPLS PE device maintaining the MP-iBGP session with RRs, the MP-iBGP convergence time is a big factor when the failover happens. In addition, two other factor need to be considered as well—object tracking detection of absence of the DMVPN tunnel availability and HSRP switchover.

We test the failure as in the previous case by shutting the SP link to the hub1 as well as reloading the primary hub. Two sets of BGP timer are tested: BGP default timer and BGP keepalive timers set to (2s, 6s). Shorter BGP timers like (2s, 6s) are not recommended for large-scale deployments without additional testing and is provided here for comparison only.

Convergence Time When BGP Default Timer is Used

Two scenarios, primary link down/up and primary router reload, have been tested.

Table 6-3 Convergence When the Primary Tunnel is Down

Iteration	1		2		3	
	Down	Up	Down	Up	Down	Up
Spoke-to-hub traffic	4s	0s	6s	2s	5s	1s
Hub-to-spoke traffic	3s	0s	5s	0s	5s	1s

Table 6-4 Convergence When the Primary Router is Reload

Iteration	1		2		3	
	Down	Up	Down	Up	Down	Up
Spoke-to-hub traffic	8s	1.5s	10s	1.2s	8s	1.5s
Hub-to-spoke traffic	8s	1s	10s	1.5s	8s	1s

While HSRP provides better redundancy and fast LAN convergence, the convergence on the WAN side is affected by other factors. With object tracking we are watching for the hub's tunnel source interface, which comes up first (in the case of up convergence), but BGP itself takes much longer since it has to wait for the Tunnel itself to come up. Thus even though HSRP has switched over to the original active router (because of preempt), BGP convergence takes longer. This issue can be addressed by delaying the HSRP switchover time when BGP is converging, which can be done by adding the delay statement for the objects that have been tracked. The delay timer implemented in the network needs to adjust based on the BGP convergence time, which needs to be tested in a scenario that's very close to the production network.

From the network design perspective, how the network redundancy is implemented in the remote branch is an integrated part of the overall branch architecture. As shown above, different redundancy approaches yield to different convergence time and which one should be used needs to be evaluated from the overall branch architecture point of view.

Note HSRP timers were kept at 1s for hello and 3s for holdtime in both the cases.

Implementing Multicast

Multicast VPN (MVPN) is the technique used to delivery multicast traffic across the MPLS network for different VPNs (user groups). From a multicast perspective, DMVPN is treated as any other transport media although we do have to account for its multipoint nature in the design.

Assuming that the enterprise MAN MPLS network at the headend is already MVPN enabled, then it is a matter of extending the functionality to the branches. In our example, each VRF is set up with static anycast RP with MSDP enabled, which provides simplicity and redundancy. The RPs typically reside closer to the source and this case the RP is configured in campus CE device at the enterprise MAN data center for each VPN (user group), where the source is connected. All the VPNs in each of the spokes has reachability to the RP and source, so from a VRF perspective the setup looks similar to a normal multicast network.

In the global space mVPN need to be implemented across the entire MPLS network where multicast is required. PIM-SSM or PIM-Bidir are the recommended protocols for the core. Default MDT is used to maintain the control plan traffic and low rate data traffic as well. Data MDT is created automatically when the traffic exceeds the configured threshold. Data MDTs are even more important in the DMVPN network because without them the hub would end up replicating the multicast traffic for all the spokes that are attached to it irrespective of whether they have receivers or not. With Data MDTs the spokes would only join the specific (S,G) if they had receivers for it. This saves CPU resources on the headend device and bandwidth at the hub and the spokes. Two conditions need to be met for the Data MDTs to be initiated:

- The traffic threshold needs to be low enough to enable (set it to 1kbps for almost instantaneous initiation).
- (S,G) entries need to exist within the VRF.

Additionally, PIM NBMA mode needs to be configured on the mGRE interface. This creates the spoke specific entry in the Multicast Output Interface List (OIL).

Caveats:

- CSCse05807 identifies problems with multicast forwarding—received mvpn traffic is process switched on ISRs. This problem is observed when the ISRs are used as a PE (irrespective of using 2547oDMVPN).

Configuration Example:

The following is built on our earlier example (Figure 6-2) and only the multicast-relevant configurations are shown here. The VPN and address scheme is illustrated in Figure 6-1. The multicast traffic is delivered from the campus to remote branches.

Hub11:

```
Hostname ngwan-hub11
!
ip multicast-routing
!
interface Tunnel1
 ip address 150.0.0.1 255.255.255.0
 ip pim nbma-mode
 ip pim sparse-mode
!
ip pim ssm range 1
access-list 1 permit 239.232.0.0 0.0.255.255
!
```

Br11:

```
ip vrf vpn1
 !
 mdt default 239.232.2.1
 mdt data 239.232.2.128 0.0.0.127 threshold 10
!
ip multicast-routing
ip multicast-routing vrf vpn1
!
interface Tunnel1
 ip address 150.0.0.2 255.255.255.0
 no ip redirects
 ip pim sparse-mode
 !

interface Loopback1
 ip address 100.0.250.35 255.255.255.255
 ip pim sparse-mode
!
!
router bgp 64512

 !
 address-family vpnv4
 neighbor RRs send-community extended
 neighbor 100.0.250.9 activate
 neighbor 100.0.250.10 activate
 exit-address-family
!
ip pim spt-threshold infinity
ip pim ssm range 1
ip pim vrf vpn1 rp-address 1.28.103.1
access-list 1 permit 239.232.0.0 0.0.255.255
!
```

Below are the show commands that illustrate the packet flow from the source PE in enterprise MPLS campus to the receiving PEs in the remote branches.

Let's first check the mroute table in both global and VPN space on the campus PE connecting to the source:

```
campus-pe1#sh ip mrou 100.0.250.22 239.232.2.128
IP Multicast Routing Table
```

```
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group
       V - RD & Vector, v - Vector
Outgoing interface flags: H - Hardware switched, A - Assert winner
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(100.0.250.22, 239.232.2.128), 02:19:15/00:03:20, flags: sTz
  Incoming interface: Loopback0, RPF nbr 0.0.0.0, RPF-MFD
  Outgoing interface list:
    GigabitEthernet4/15, Forward/Sparse, 02:19:15/00:02:57, H

campus-pe1#
ngden-7606-pe1#sh ip mrou 239.232.2.128 ac
Active IP Multicast Sources - sending >= 4 kbps

Group: 239.232.2.128, (?)
   Source: 100.0.250.22 (?)
     Rate: 2840 pps/1681 kbps(1sec), 1681 kbps(last 10 secs), 1671 kbps(life avg)
campus-pe1#

campus-pe1#sh ip mrou vrf vpn2 224.2.253.249
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group
       V - RD & Vector, v - Vector
Outgoing interface flags: H - Hardware switched, A - Assert winner
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 224.2.253.249), 02:07:15/00:03:27, RP 1.28.103.1, flags: S
  Incoming interface: GigabitEthernet4/3.1121, RPF nbr 1.28.1.1, RPF-MFD
  Outgoing interface list:
    Tunnel2, Forward/Sparse, 02:07:15/00:03:27, H

(1.28.101.2, 224.2.253.249), 02:07:03/00:03:25, flags: Ty
  Incoming interface: GigabitEthernet4/3.1121, RPF nbr 1.28.1.1, RPF-MFD
  Outgoing interface list:
    Tunnel2, Forward/Sparse, 02:07:15/00:03:27, H

campus-pe1#sh ip mrou vrf vpn2 224.2.253.249 ac
Active IP Multicast Sources - sending >= 4 kbps

Group: 224.2.253.249, (?)
   Source: 1.28.101.2 (?)
     Rate: 2840 pps/1045 kbps(1sec), 1045 kbps(last 0 secs), 1038 kbps(life avg)
campus-pe1#
```

Now let's take a look the mroute table in the DMVPN hub router, which is a P device in the MPLS network:

```
ngwan-hub11#sh ip mrou 100.0.250.22 239.232.2.128
IP Multicast Routing Table
```

```
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group
Outgoing interface flags: H - Hardware switched, A - Assert winner
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(100.0.250.22, 239.232.2.128), 02:22:34/00:03:20, flags: sT
  Incoming interface: GigabitEthernet0/2, RPF nbr 100.0.33.1
  Outgoing interface list:
    Tunnel1, Forward/Sparse, 02:22:34/00:03:05

ngwan-hub11#sh ip mrou 239.232.2.128 ac
Active IP Multicast Sources - sending >= 4 kbps
 a negative (-) Rate counts pps being fast-dropped

Group: 239.232.2.128, (?)
   Source: 100.0.250.22 (?)
     Rate: 2394 pps/1340 kbps(1sec), 1340 kbps(last 0 secs), 357 kbps(life avg)
ngwan-hub11#
```

Finally let's examine the mroute and active multicast traffic in both global and VPN space in the remote spokes (receiving PEs) attached with multicast receiver.

```
ngwan-br12#sh ip mrou 100.0.250.22 239.232.2.128
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group
Outgoing interface flags: H - Hardware switched, A - Assert winner
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(100.0.250.22, 239.232.2.128), 02:55:26/00:02:57, flags: sTIZ
  Incoming interface: Tunnel1, RPF nbr 150.0.0.1
  Outgoing interface list:
    MVRF vpn2, Forward/Sparse, 00:10:10/00:01:50

ngwan-br12#sh ip mrou 239.232.2.128 ac
Active IP Multicast Sources - sending >= 4 kbps
 a negative (-) Rate counts pps being fast-dropped

Group: 239.232.2.128, (?)
   Source: 100.0.250.22 (?)
     Rate: 4787 pps/2680 kbps(1sec), 2654 kbps(last 30 secs), 270 kbps(life avg)
ngwan-br12#


ngwan-br12#sh ip mrou vrf vpn2 224.2.253.249
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
```

```
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group
Outgoing interface flags: H - Hardware switched, A - Assert winner
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 224.2.253.249), 02:51:45/stopped, RP 1.28.103.1, flags: SJC
  Incoming interface: Tunnel3, RPF nbr 100.0.250.22
  Outgoing interface list:
    GigabitEthernet0/1.112, Forward/Sparse, 02:51:42/00:02:01

(1.28.101.2, 224.2.253.249), 02:51:42/00:02:55, flags: JTY
  Incoming interface: Tunnel3, RPF nbr 100.0.250.22,
MDT:[100.0.250.22,239.232.2.128]/00:02:42
  Outgoing interface list:
    GigabitEthernet0/1.112, Forward/Sparse, 02:51:42/00:02:01

ngwan-br12#
ngwan-br12#sh ip mrou vrf vpn2 224.2.253.249 ac
Active IP Multicast Sources - sending >= 4 kbps
 a negative (-) Rate counts pps being fast-dropped

Group: 224.2.253.249, (?)
   Source: 1.28.101.2 (?)
     Rate: 2357 pps/867 kbps(1sec), 879 kbps(last 40 secs), 53 kbps(life avg)
ngwan-br12#
```

Configuration Checklist:

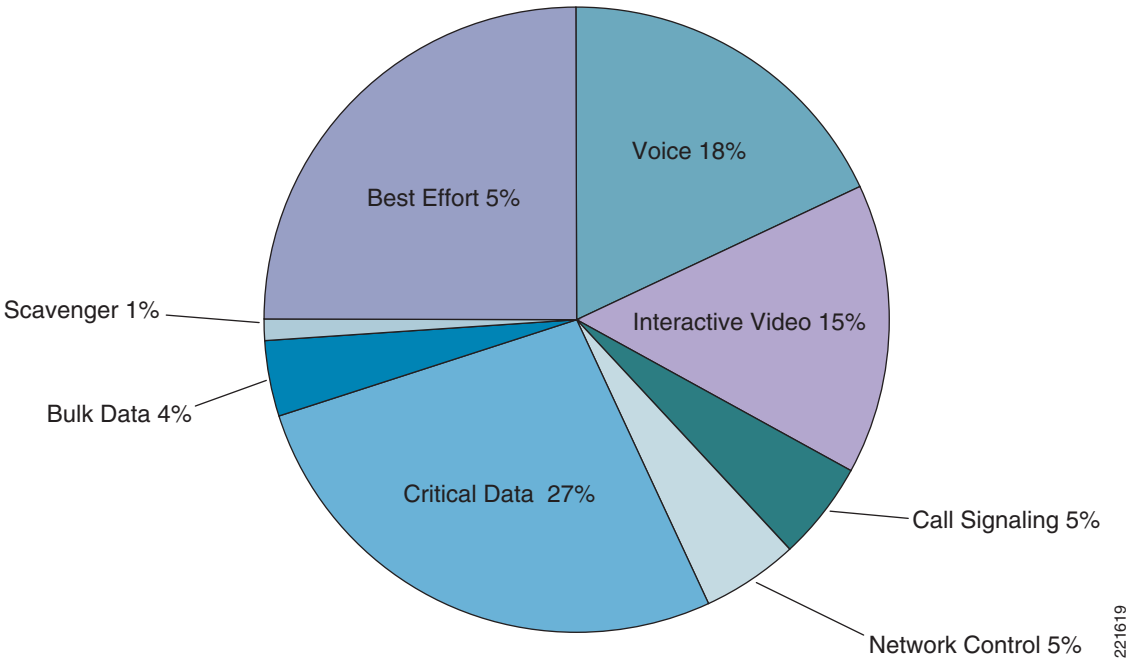
- Configure PIM nbma-mode on the mGRE interface at the headend.
- Ensure that the MDT switchover threshold is set to the lowest value to enable the data MDTs.

Implementing QoS

A basic assumption of this implementation is that the enterprise is getting a Layer 3 VPN service from a provider. Thus the level of QoS service from a provider becomes important. Typically, a service with 3-5 classes of service can be expected. We do not focus on SP service in this design guide as this has been discussed extensively in the Consumer Guidance WP (http://www.cisco.com/application/pdf/en/us/guest/netsol/ns465/c654/cdccont_0900aecd80375d78.pdf). We focus on the aspects of QoS that are within enterprise control, primarily on WAN Edge QoS at the DMVPN headend and the branches.

QoS policies required on the headend include queuing, shaping, selective dropping, and link-efficiency policies in the outbound direction of the WAN link. Traffic is assumed to be correctly classified and marked (at Layer 3) before it ingresses the headend router. At the headend the expectation is that a interface with high link speed is used (DS3/OC3/GE range). At these speeds, link-efficiency policies such as LFI and cRTP are not required. The Enterprise QoS SRND recommends 5-11 classes at the WAN edge. The choice would be dependent on the existing core QoS deployment. We use a 8-class model in our example. The typical bandwidth allocation for a 8-class model is shown in Figure 6-3 (from the SRND).

Figure 6-3 8 Class QoS Model

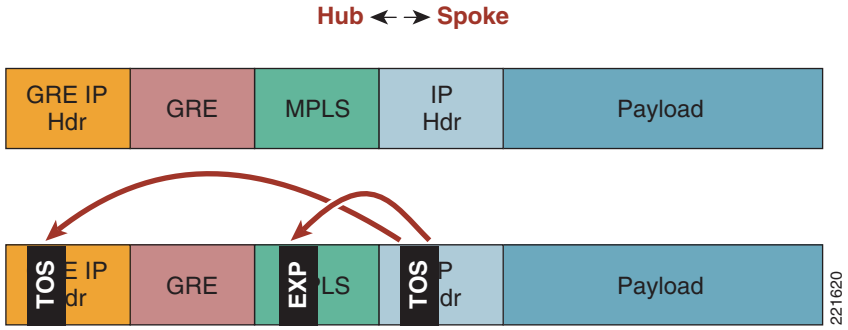


At the branches, the PE could be configured to map the COS to DSCP, but in our example we assume that the packets are already marked with the appropriate DSCP. If the branches have slow/medium speed links (<T1) then a 3-5 class model is recommended. One option could be to match the model used by the SP providing the Layer 3 VPN service. We assume that the branches have higher speed links (>T1) and implement a 8-class model as well (similar to the hub).

Overall the WAN QoS recommendation made in the Enterprise QoS SRND remain for 2547oDMVPN as well. This is because the labeled packets are encapsulated in the GRE header and at the outgoing interface the packet looks like a normal IP packet which can be treated under existing guidelines.

For packets going from hub to spoke and vice-versa, as shown in Figure 6-4, the DSCP/TOS from original IP packet is copied to the MPLS EXP (automatic IP Precedence to EXP mapping) as well as to outer headers DSCP/TOS field.

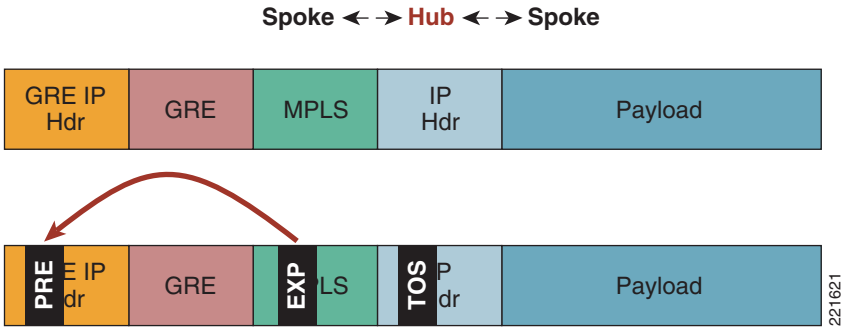
Figure 6-4 Headers for QoS—Hub and Spoke Traffic



Packets traverse the hub in the case of spoke-to-spoke communication. In this case the source branch behavior remains same as above. At the hub, the GRE header is stripped off before a forwarding decision is taken, which in this case requires adding another GRE header before forwarding it back out to the

destination branch. As shown in Figure 6-5, the EXP gets copied to the outgoing GRE IP header as Precedence (3 bits only). Thus the outgoing policy on the hub should account for DSCP as well as Precedence.

Figure 6-5 Headers for QoS—Spoke to Spoke Traffic



The original IP headers marking are always preserved in either case.

Note At the hub, since all the traffic is placed inside the same GRE tunnel, per spoke QoS is not supported.

Example:

Hub1 has a OC3 ATM connection to the SP and spoke B21a has a high speed Ethernet connection. LLQ is used for Voice and Interactive Video traffic. The rest of the classes are provided a bandwidth percentage. DSCP-based WRED is enabled for both Critical Data and Bulk Data.

Hub1:

```
class-map match-any Bulk-Data
 match ip dscp af11
 match ip dscp af12
 match ip precedence 1
class-map match-any Interactive-Video
 match ip dscp af41
 match ip dscp af41
 match ip precedence 4
class-map match-any Network-Control
 match ip dscp cs6
 match ip dscp cs2
 match ip precedence 6
class-map match-any Critical-Data
 match ip dscp af21
 match ip dscp af22
 match ip precedence 2
class-map match-any Call-Signaling
 match ip dscp cs3
 match ip dscp af31
 match ip precedence 3
class-map match-any Voice
 match ip dscp ef
 match ip precedence 5
class-map match-any Scavenger
 match ip dscp cs1
!
policy-map WAN-EDGE
 class Interactive-Video
  priority percent 15
 class Call-Signaling
```

```
  bandwidth percent 5
 class Network-Control
  bandwidth percent 5
 class Critical-Data
  bandwidth percent 27
  random-detect dscp-based
 class Bulk-Data
  bandwidth percent 4
  random-detect dscp-based
 class Scavenger
  bandwidth percent 1
 class Voice
  priority percent 18
 class class-default
  bandwidth percent 25
  random-detect
!
interface ATM5/0
 ip address 135.0.13.2 255.255.255.252
 pvc 1/1
  vbr-nrt 44209 44209
  service-policy output WAN-EDGE
  max-reserved-bandwidth 100
```

Note Scavenger traffic is mapped to Bulk Data at the hub for spoke-to-spoke communication.

Spoke B21a:

```
class-map match-all Bulk-Data
 match ip dscp af11  af12
class-map match-all Interactive-Video
 match ip dscp af41  af42
class-map match-any Network-Control
 match ip dscp cs6
 match ip dscp cs2
class-map match-all Critical-Data
 match ip dscp af21  af22
class-map match-any Call-Signaling
 match ip dscp cs3
 match ip dscp af31
class-map match-all Voice
 match ip dscp ef
class-map match-all Scavenger
 match ip dscp cs1
!
policy-map WAN-EDGE
 class Voice
  priority percent 18
 class Interactive-Video
  priority percent 15
 class Call-Signaling
  bandwidth percent 5
 class Network-Control
  bandwidth percent 5
 class Critical-Data
  bandwidth percent 27
  random-detect dscp-based
 class Bulk-Data
  bandwidth percent 4
  random-detect dscp-based
 class Scavenger
  bandwidth percent 1
```

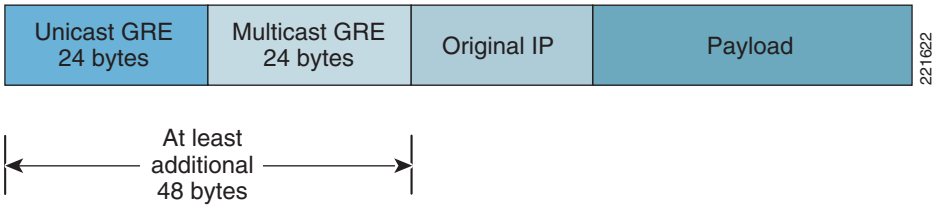
```
 class class-default
  bandwidth percent 25
  random-detect
!
interface FastEthernet1/1
 description To SP
 ip address 135.0.5.1 255.255.255.252
 max-reserved-bandwidth 100
 service-policy output WAN-EDGE
```

Note QoS policy is applied to the outgoing physical interface only. No policy is required on the mGRE interface.

MTU Issues

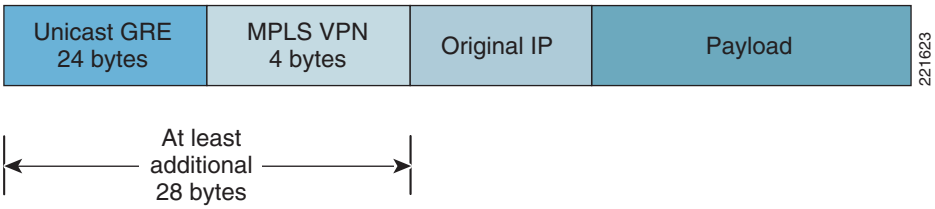
As with any tunneled implementation, MTU size can become an issue. This becomes particular acute with MVPN over 2547oDMVPN when the underlying service is a Layer 3 VPN service from a provider and mVPN is used to delivery the multicast packets. As can be seen in Figure 6-6, at the hub the original IP multicast packet is encapsulated in a multicast GRE header (MTI) which is encapsulated in the unicast GRE header (DMVPN) before being sent to the SP (with appropriate Layer 2 header added). This means that the original IP packet ends up with an additional overhead of at least 48 bytes (without encryption).

Figure 6-6 2547oDMVPN—Multicast Packet Overhead



In the case of unicast, things are a little better. The original IP packet has a MPLS label attached to it before being encapsulated into the unicast GRE (DMVPN). As shown in Figure 6-7, the additional overhead can be expected to be at least 28 bytes (without encryption).

Figure 6-7 2547oDMVPN—Unicast Packet Overhead



The safest MTU (the worst case MTU) for tunnel interface to avoid fragmentation is 1400 bytes. Taking into consideration that each MPLS label is 4 bytes, the safest MTU on 2547oDVMPN tunnel is 1392 Bytes for unicast traffic and 1368 for multicast traffic.

"ip tcp adjust-mss <value>" can also be used to inform the end device to use the correct MSS for TCP transmissions. The MSS must be set to a value that equals the interface MTU minus the size of IP, TCP, GRE, and MPLS headers.

The GRE interface can also be configured with "mpls mtu" which sets the MTU for the labeled packets. MPLS MTU is derived by adding (label stack x 4bytes) to the interface MTU.

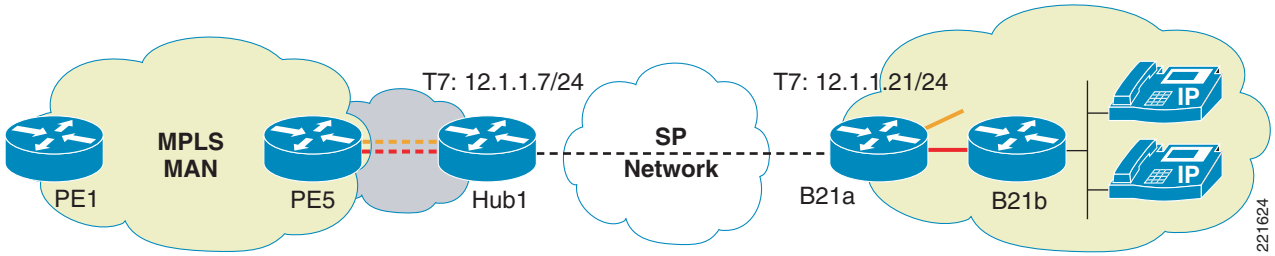
Voice and VRFs

Typically voice traffic has no dependency on the network type since they are just transported as IP packets and require correct QoS behavior applied to them. An exception is when routers are used as gateways for voice services because a lot of voice features and protocols deployed at the branches are not VRF aware (for example, SRST, CME, etc. Thus just getting the voice traffic in a VRF could be a challenge. This is apart from larger issues of having the voice in a VRF; while you can have the IP phones within a VRF, other services such as softphones VT advantage may be in a different VRF. There are challenges in implementing Inter-VRF IP communications. These are not discussed here as its part of the larger virtualization architecture issue. The current recommendation is to keep voice within the global space especially at the branches. At the hub they could remain in the global space or would have to be placed within its own VRF. We look at both options, getting the voice in the VRF at the branch as well keeping it in the global table at the branch.

Voice in a VRF at the Branch

If we need to put the voice in the VRF and still want to use voice features such as CME, then the only way to currently do this is by having two separate routers at the branch. The branch edge router still has a voice VRF configured but treats it like any other VRF. It has a second router (such as a low end ISR) connected to its voice VRF VLAN. The second router, as shown in Figure 6-8, has all the phones attached to it. Since it requires two routers at every such branch, this can be a expensive proposition.

Figure 6-8 2547oDMVPN—Voice in a VRF at the Branch



Example:

Branch 21 has the voice VRF configured on B21a. B21a has 21b connected to it within VRF red-voice. B21b provides the connection support for IP/Pots phones within the branch.

B21a:

```
ip vrf red-voice
 rd 10:104
 route-target export 10:104
 route-target import 10:104
!
interface GigabitEthernet0/0
 description to voice-B21b
```

```
 ip vrf forwarding red-voice
 ip address 125.1.14.129 255.255.255.128
!
router ospf 2 vrf red-voice
 log-adjacency-changes
 redistribute bgp 1 subnets
 network 125.1.14.0 0.0.0.255 area 0
!
router bgp 1
<snip>
!
address-family ipv4 vrf red-voice
redistribute connected
redistribute ospf 2 vrf red-voice match internal external 1 external 2
no synchronization
exit-address-family
```

Voice Global at the Branch

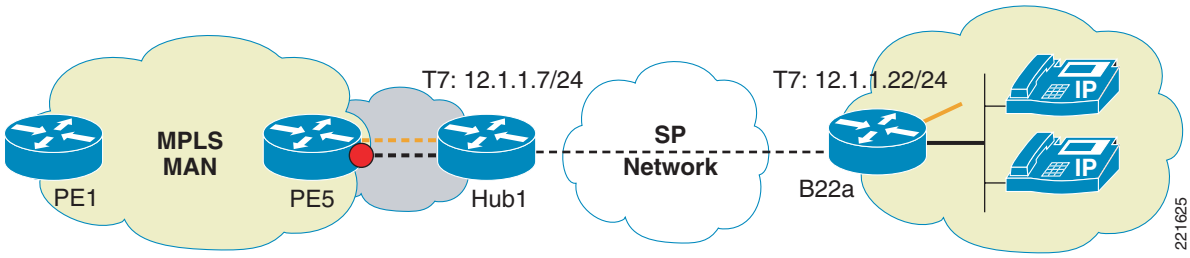
If we choose to keep the voice global at the branch then a single router would suffice. The voice VLAN is connected to the branch router but remains in the global space. It is carried across the same GRE that is used to carry labeled packets but as normal IPv4 traffic. At the hub router, it remains in the global space. It can be forwarded to the next hop router on a global VLAN or can be forwarded to core MPLS PE where it can be placed in its own VRF.

Note To get the voice traffic routed, a routing protocol needs to be used over the GRE tunnels. This potentially has a scaleability impact on the DMVPN setup (relying solely on MP-BGP vs. a combination of MP-BGP and IGP for route distribution).

Example:

B22a in our example is a PE which uses MP-BGP for all the VRFs except voice. Voice traffic exists in global space and hence voice VRF is not defined on it. It runs OSPF with the hub. At the hub the traffic remains in the global space but it has a OSPF peering with PE5. On PE5 the traffic is placed within voice VRF as shown in Figure 6-9.

Figure 6-9 2547oDMVPN—Voice Global at the Branch



B22a:

```
interface Tunnel7
 ip address 12.1.1.22 255.255.255.0
 ip pim sparse-mode
 ip nhrp authentication spe
 ip nhrp network-id 7
 ip nhrp nhs 12.1.1.7
 ip ospf network broadcast
```

```
 ip ospf priority 0
 load-interval 30
 mpls ip
 tunnel source 135.0.3.1
 tunnel destination 135.0.13.2
 tunnel key 777
 tunnel protection ipsec profile P1
!
interface GigabitEthernet0/1.2
 description voice in global
 encapsulation dot1Q 222
 ip address 125.1.15.1 255.255.255.0
 ip pim sparse-mode
 no snmp trap link-status
!
router ospf 3
 log-adjacency-changes
 network 12.1.1.0 0.0.0.255 area 0
 network 121.1.1.0 0.0.0.255 area 0
 network 125.1.15.0 0.0.0.255 area 0
```

Hub1:

```
interface Tunnel7
 ip address 12.1.1.7 255.255.255.0
 no ip redirects
 ip pim nbma-mode
 ip pim sparse-mode
 ip nhrp authentication spe
 ip nhrp map multicast dynamic
 ip nhrp network-id 7
 ip ospf network broadcast
 ip ospf priority 100
 load-interval 30
 mpls ip
 qos pre-classify
 tunnel source 135.0.13.2
 tunnel mode gre multipoint
 tunnel key 777
 tunnel protection ipsec profile P1
!
interface GigabitEthernet0/2.4
 description To PE5
 encapsulation dot1Q 104
 ip address 125.1.103.110 255.255.255.252
!
router ospf 3
 log-adjacency-changes
 network 12.1.1.0 0.0.0.255 area 0
 network 125.1.103.108 0.0.0.3 area 0
 network 125.1.125.22 0.0.0.0 area 0
```

PE5:

```
interface GigabitEthernet1/6.4
 description To 7200-hub1
 encapsulation dot1Q 104
 ip vrf forwarding red-voice
 ip address 125.1.103.109 255.255.255.252
!
```

```
router ospf 2 vrf red-voice
 log-adjacency-changes
 redistribute bgp 1 subnets
 network 125.1.103.108 0.0.0.3 area 0
!
router bgp 1
<snip>
address-family ipv4 vrf red-voice
redistribute ospf 2 vrf red-voice match internal external 1 external 2
no auto-summary
no synchronization
exit-address-family
```

Scale Considerations

The 2547oDMVPN hub can support at least 500 remote spokes as we have tested this scenario in the lab with a simulated customer-representative environment. OSPF is used as the routing protocol between DMVPN hub and spoke. Hence the following scalability numbers apply to this setup; these numbers will satisfy many of the customer scaling requirements:

- 500 OSPF neighbors on hub router
- 500 LDP sessions on hub router
- 500 NHRP entries on hub router
- 500 IPSec sessions on hub router
- 500 MP-iBGP sessions on RRs[1]

Solution Caveats Summary

2547oDMVPN provides a scalable solution for both greenfield virtualization deployments and established DMVPN deployments moving towards virtualization. The following list summarizes the caveats for the deployment model discussed here:

- No direct spoke-to-spoke tunnels can be established. Spoke-to-spoke communication has to happen through the hub.
- For multicast ensure that PIM-NBMA mode is configured on the tunnel interface.
- Use Data MDTs where possible to avoid unnecessary flows to the bandwidth-sensitive remote spokes.
- MTU overhead due to MPLS labels and GRE headers need to be considered.

1. The RRs also handle the other MP-iBGP session for the rest of the network.