Configure Inter-AS Option C MPLS VPN With Cisco IOS and Cisco IOS-XR

Contents

Introduction

This document describes how to configure and verify the Inter-AS Layer 3 Multiprotocol Label Switching (MPLS) VPN, option C feature. Cisco IOS[®] and Cisco IOS-XR platforms are used for explanation and verification. A sample network scenario and its configuration and outputs are shown for a better understanding.

Prerequisites

Requirements

There are no specific requirements for this document. However, basic knowledge of MPLS and a working knowledge of the Cisco IOS-XR platform will be helpful.

Components Used

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

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

Background Information

MPLS is widely deployed across Internet Service Providers (ISPs) worldwide. ISPs offer a large range of services to customers and one such service is MPLS Layer 3 VPN. MPLS Layer 3 VPNs mainly stretch the routing boundaries of a customer from one geographical location to another. ISP is mainly used as a transit. Peering with ISP on one geographical location and on the other geographical location is completed, then the customer specific routes are received on the Customer Edge (CE) device from the PE (Provider Edge/ISP) device.
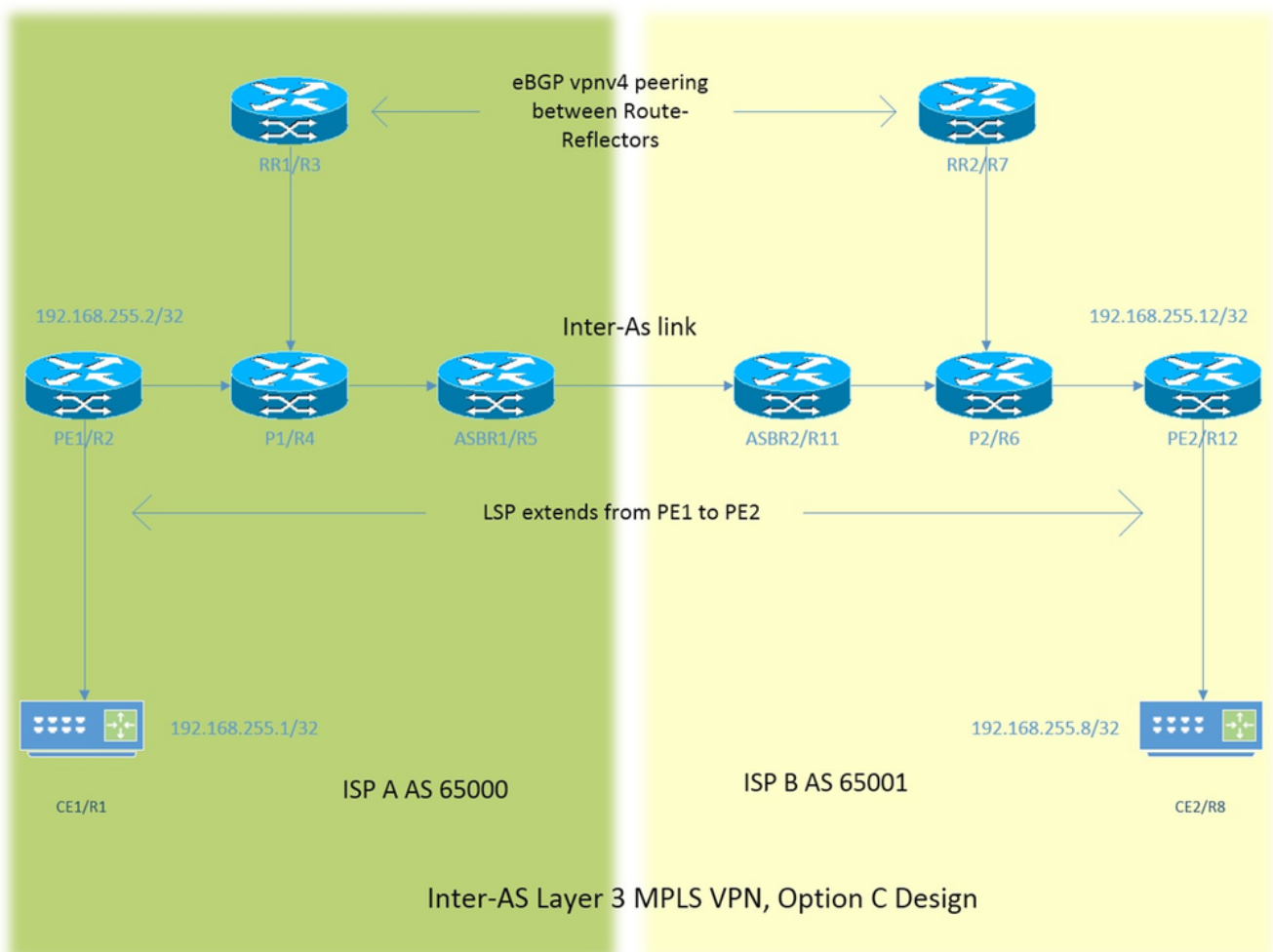
If the requirement is to stretch routing boundaries for a customer for two different geographical locations where two different ISPs have presence, then the two ISPs need to coordinate so that the MPLS Layer 3 VPN is provided to the end customer. Such a solution is called Inter-AS Layer 3 MPLS VPN.

Inter-AS Layer 3 MPLS VPNS can be deployed in four different ways, known as Option A, Option B, Option C, and Option D. Implementation with Option C is explained in this document.

Configure

Network Diagram

The topology for the Inter-AS Option C exchange as shown in this image.

Inter-AS Layer 3 MPLS VPN, Option C Design

The addressing scheme is very simple. Every router has loopback1 interface described as 192.168.255.X, where X=1 when the router 1 is under concern. The interface addressing is of the type 192.168.XY.X . Suppose R1 and R2 are in under consideration, configuration of the interface under router R1 is 192.168.12.1 (here X =1, Y = 2).

CE - Customer Edge

PE - Provider Edge

RR - Route Reflector

ASBR - Autonomous System Boundary Router

Throughout the document, the term CE denotes both the Customer Edge devices. If a specific reference has to be made for a particular device then it will be referenced as CE1. This applies to PE, RR, and ASBR as well.

All the devices run Cisco IOS, however ASBR2/R11 and PE2/R12 run Cisco IOS-XR.

Two ISPs are referenced with Autonomous System (AS) 65000 and AS 65001. ISP with AS 65000 is on the left side of the topology and is referenced as ISP A and ISP with AS 65001 is on the right side of the topology and is referenced as ISP B.

Configurations

The configurations of the devices are described.

CE1

```
interface Loopback1                        #Customer Edge configuration.
ip address 192.168.255.1 255.255.255.255                              !
interface FastEthernet0/0
ip address 192.168.12.1 255.255.255.0
!
router eigrp 1
network 0.0.0.0
!
```

PE1

```
vrf definition A                      #Provider Edge Configuration.
rd 192.168.255.2:65000
!
address-family ipv4
route-target export 99:99
route-target import 99:99
exit-address-family
!
interface Loopback1
ip address 192.168.255.2 255.255.255.255
ip ospf 1 area 0
!
interface FastEthernet0/0
vrf forwarding A
ip address 192.168.12.2 255.255.255.0
```

```
!
interface FastEthernet1/0
ip address 192.168.24.2 255.255.255.0
ip ospf 1 area 0
mpls ip
!
router eigrp 65000                        #EIGRP is PE-CE routing
!                                         #protocol.
address-family ipv4 vrf A autonomous-system 1
redistribute bgp 65000 metric 10000 10 255 1 1500
network 192.168.12.2 0.0.0.0
exit-address-family
!
router ospf 1
!
router bgp 65000
bgp log-neighbor-changes
no bgp default ipv4-unicast
neighbor 192.168.255.3 remote-as 65000
neighbor 192.168.255.3 update-source Loopback1
!
address-family ipv4
exit-address-family
!
address-family vpnv4                      #Advertising vpnv4 routes
neighbor 192.168.255.3 activate           #from PE1 to RR1.
neighbor 192.168.255.3 send-community both
exit-address-family
!
address-family ipv4 vrf A
redistribute eigrp 1
exit-address-family
!
```

P1

```
interface Loopback1                       #P router configuration.
ip address 192.168.255.4 255.255.255.255
ip ospf 1 area 0
!
interface FastEthernet0/0
ip address 192.168.24.4 255.255.255.0
ip ospf 1 area 0
duplex half
mpls ip
!
interface FastEthernet1/0
ip address 192.168.34.4 255.255.255.0
ip ospf 1 area 0
mpls ip
!
interface FastEthernet1/1
ip address 192.168.45.4 255.255.255.0
ip ospf 1 area 0
mpls ip
!
router ospf 1
!
```

RR1

```
interface Loopback1                       #Route-Reflector configuration.
```

```
ip address 192.168.255.3 255.255.255.255
ip ospf 1 area 0
!
interface FastEthernet0/0
ip address 192.168.34.3 255.255.255.0
ip ospf 1 area 0
mpls ip
!
router ospf 1
!
router bgp 65000
bgp log-neighbor-changes
neighbor 192.168.255.2 remote-as 65000
neighbor 192.168.255.2 update-source Loopback1
neighbor 192.168.255.7 remote-as 65001
neighbor 192.168.255.7 ebgp-multihop 255        #EBGP-Multihop vpnv4
neighbor 192.168.255.7 update-source Loopback1 #peering with RR2.

!
address-family vpnv4
neighbor 192.168.255.2 activate
neighbor 192.168.255.2 send-community both
neighbor 192.168.255.2 route-reflector-client
neighbor 192.168.255.7 activate
neighbor 192.168.255.7 send-community both
neighbor 192.168.255.7 next-hop-unchanged
exit-address-family
!
```

ASBR1

```
interface Loopback1                          #Autonomous-System boundary-
ip address 192.168.255.5 255.255.255.255  #router configuration.
ip ospf 1 area 0
!
interface FastEthernet0/0
ip address 192.168.45.5 255.255.255.0
ip ospf 1 area 0
mpls ip
!
interface FastEthernet1/0
ip address 192.168.115.5 255.255.255.0
mpls bgp forwarding
!
router ospf 1
redistribute bgp 65000 subnets route-map REDISTRIBUTE_IN_IGP
!                                        #Redistributing the loopbacks of
router bgp 65000                         #RR2 and PE2 in AS 65000.
bgp log-neighbor-changes
network 192.168.255.2 mask 255.255.255.255
network 192.168.255.3 mask 255.255.255.255
neighbor 192.168.115.11 remote-as 65001
neighbor 192.168.115.11 send-label
!
ip prefix-list FOREIGN_PREFIXES seq 5 permit 192.168.255.12/32
ip prefix-list FOREIGN_PREFIXES seq 10 permit 192.168.255.7/32
!
route-map REDISTRIBUTE_IN_IGP permit 10
match ip address prefix-list FOREIGN_PREFIXES
!
```

ASBR2

```
interface Loopback1                           #Autonomous System boundary
ipv4 address 192.168.255.11 255.255.255.255 #configuration.
!
interface GigabitEthernet0/0/0/0
ipv4 address 192.168.115.11 255.255.255.0
!
interface GigabitEthernet0/0/0/1
ipv4 address 192.168.116.11 255.255.255.0
!
prefix-set FOREIGN_PREFIXES
192.168.255.2/32,
192.168.255.3/32
end-set
!
route-policy DEFAULT
pass
end-policy
!
route-policy REDISTRIBUTE_IN_IGP
if destination in FOREIGN_PREFIXES then
pass
endif
end-policy
!
router static
address-family ipv4 unicast
192.168.115.5/32 GigabitEthernet0/0/0/0
!
router ospf 1
redistribute bgp 65001 route-policy REDISTRIBUTE_IN_IGP
area 0                                    #Redistributing the loopback
interface Loopback1                       #of RR1 and PE1 in AS 65001.
!
interface GigabitEthernet0/0/0/1
!
router bgp 65001
address-family ipv4 unicast
network 192.168.255.7/32
network 192.168.255.12/32
allocate-label all
!
neighbor 192.168.115.5
remote-as 65000
address-family ipv4 labeled-unicast
route-policy DEFAULT in
route-policy DEFAULT out
!
mpls ldp
address-family ipv4
!
interface GigabitEthernet0/0/0/1
!
```

RR2

```
interface Loopback1                       #Route-Refector Configuration.
ip address 192.168.255.7 255.255.255.255
ip ospf 1 area 0
!
interface FastEthernet0/0
ip address 192.168.67.7 255.255.255.0
ip ospf 1 area 0
```

```
mpls ip
!
router ospf 1
!
router bgp 65001
bgp log-neighbor-changes
neighbor 192.168.255.3 remote-as 65000     #EBGP-Multihop vpnv4 peering
neighbor 192.168.255.3 ebgp-multihop 255   #with RR1 in AS 65000.
neighbor 192.168.255.3 update-source Loopback1
neighbor 192.168.255.12 remote-as 65001
neighbor 192.168.255.12 update-source Loopback1
!
address-family vpnv4
neighbor 192.168.255.3 activate
neighbor 192.168.255.3 send-community both
neighbor 192.168.255.3 next-hop-unchanged
neighbor 192.168.255.12 activate
neighbor 192.168.255.12 send-community both
neighbor 192.168.255.12 route-reflector-client
exit-address-family
!
```

P2

```
interface Loopback1                         #P router configuration.
ip address 192.168.255.6 255.255.255.255
ip ospf 1 area 0
!
interface FastEthernet0/0
ip address 192.168.116.6 255.255.255.0
ip ospf 1 area 0
mpls ip
!
interface FastEthernet1/0
ip address 192.168.67.6 255.255.255.0
ip ospf 1 area 0
mpls ip
!
interface FastEthernet1/1
ip address 192.168.126.6 255.255.255.0
ip ospf 1 area 0
mpls ip
!
router ospf 1
!
```

PE2

```
vrf A                                  #Provider Edge Configuration.
address-family ipv4 unicast
import route-target
99:99
!
export route-target
99:99
!
!
interface Loopback1
ipv4 address 192.168.255.12 255.255.255.255
!
interface GigabitEthernet0/0/0/0
```

```
ipv4 address 192.168.126.12 255.255.255.0
!
interface GigabitEthernet0/0/0/1
vrf A
ipv4 address 192.168.128.2 255.255.255.0
!
router ospf 1
address-family ipv4
area 0
interface Loopback1
!
interface GigabitEthernet0/0/0/0
!
router bgp 65001
address-family vpnv4 unicast
!
neighbor 192.168.255.7                    #Advertising vpnv4 routes from
remote-as 65001                           #PE2 to RR2.
update-source Loopback1
address-family vpnv4 unicast
!
!
vrf A
rd 192.168.255.12:65001
address-family ipv4 unicast
redistribute eigrp 1
!
mpls ldp
address-family ipv4
!
interface GigabitEthernet0/0/0/0
!
router eigrp 65001                         #EIGRP as PE-CE protocol
vrf A
address-family ipv4
autonomous-system 1
redistribute bgp 65001
interface GigabitEthernet0/0/0/1
!
```

CE2

```
interface Loopback1                        #Customer-Edge Configuration.
ip address 192.168.255.8 255.255.255.255
!
interface FastEthernet1/0
ip address 192.168.128.8 255.255.255.0
!
router eigrp 1
network 0.0.0.0
!
```

Explanation

- Enhanced Interior Gateway Routing Protocol (EIGRP) as the PE-CE routing protocol is being deployed.
- Open Shortest Path First (OSPF) is used as the Interior Gateway Protocol (IGP) for the ISP core. On both the ISPs on all the physical links Label Distribution Protocol (LDP) + IGP is deployed. LDP + IGP is not configured on the Inter-AS link between ASBR1 and ASBR2.
- Redistribution of EIGRP under VRF A into Border Gateway Protocol (BGP) and vice versa is performed on PE.

- These redistributed routes are advertised as VPNv4 routes to the route reflector (RR).
- The route-reflector RR1 peers with PE1 and reflects these routes learned via PE1 to RR2 via eBGP VPNv4 multihop peering.
- This eBGP VPNv4 multihop peering is between two RRs in distinct ASs.
- It is important that LSP (Label Switch Path) should exist between the two RRs.
- In order to achieve a LSP between the two RRs located in a different AS, it is needed to leak the specific routes between the ASs.
- The ASBR1 and ASBR2 leak the specific routes, basically the loopback1 of the PE and RR of its own AS. Leaking is done via advertising the route in normal eBGP peering between the ASBRs.
- The ASBRs mutually receive each other's advertised loopback1 prefixes of RR and PE routers. Next, the received routes are redistributed in IGP (OSPF here). The redistribution is specific in nature, only the two prefixes, that is, the loopback1 of remote RR and PE are redistributed.
- The redistribution of routes from BGP to OSPF and matching the routes to be redistributed in OSPF is slightly different in Cisco IOS-XR and needs the knowledge of prefix-set and route-policy configurations. Prefix-set is similar to prefix-list in Cisco IOS and route-policy is equivalent to route-map.
- Now a LSP exists between RR1 and RR2 and as well as PE1 and PE2.
- The next-hop-unchanged for eBGP VPNv4 peers is used in RRs. It has to be noted that next hop of the VPNv4 route defines the LSP. Now, if an update is originated from PE2 and is sent to RR2 (iBGP peering) next hop is preserved. When RR2 reflects this update to RR1, since this is an eBGP peering, by normal scenario RR2 will set itself as the next hop for the update and advertise it to RR1. RR1 will reflect this update to PE1. So, PE1 will install the update and will see the next hop of the update as RR2. As already mentioned, the next hop of the VPNv4 route defines the LSP. Hence for PE1 to get to PE2, RR2 is the next hop. Hence, two LSPs are needed, one from PE1 to RR2 and other from RR2 to PE2. The drawback in such a design is that the traffic may traverse the same link twice (as in this topology) and RRs also lie in the transit path of traffic.
- In order to overcome such a design issue, next-hop-unchanged is used. When RR2 gets an update from PE2 and reflects the update to RR1, the next hop in the update will still be PE2 and when RR1 reflects this to PE1, PE1 installs the update with the next hop of PE2. This means a single LSP from PE1 to PE2 and no RR in transit.
- It has to be noted that on the Inter-AS link, no MPLS or LDP is deployed. ASBRs used BGP to send labels. XR needs to enable IPv4 labeled unicast address-family.
- When the eBGP labeled unicast peering comes up on the ASBR1 (Cisco IOS) with the Cisco IOS-XR device, automatically "MPLS BGP forwarding" is configured on the Inter-AS link. Exchange of the labels with ASBR2 is accomplished, not via LDP but via BGP. Cisco IOS also automatically adds a connected /32 route to ASBR2's interface so that MPLS label is bound to a /32 route and label switching is properly done.
- For Cisco IOS-XR over Inter-AS link, there is a different logic as compared to that of Cisco IOS. It is required to configure a static /32 route to ASBR1's interface, so that MPLS label is bound for a /32 prefix. If this is not done then the control plane will come up, but the traffic will not be forwarded.

Verify

Ping from CE1 to CE2 and Vice Versa

The output of ping from CE1 to CE2 with the loopback1 interface as the source is:

```
R1#ping 192.168.255.8 source lo1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.255.8, timeout is 2 seconds:
Packet sent with a source address of 192.168.255.1
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 104/300/420 ms
```

The output of ping from CE2 to CE1 with the loopback1 interface as the source is:

```
R8#ping 192.168.255.1 source lo1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.255.1, timeout is 2 seconds:
Packet sent with a source address of 192.168.255.8
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 168/303/664 ms
```

Explanation of Updates Exchanged and MPLS Labels

- On CE1, the **show ip route** command gives the route for loopback1 of the CE2 on the other end.
  ```
  R1#show ip route 192.168.255.8
  Routing entry for 192.168.255.8/32
  Known via "eigrp 1", distance 90, metric 156416, type internal
  ```
- The traffic flow with MPLS labels imposed/disposed along the path CE1 to CE2 is discussed here, along with how reachability is obtained when it goes from source loopback1 of CE1 to loopback1 of CE2.
- In MPLS Layer 3 VPN designs, it should be remembered that during the label switch operation the transport label is swapped and the VPN label is untouched. The VPN label is exposed when Penultimate Hop Popping (PHP) occurs and traffic reaches PE or when a Label Switched Path (LSP) is terminated.
- On PE1, the loopback1 of CE2 is learned via BGP VPNv4 update and redistributed into to VRF aware EIGRP. The loopback1 learned via CE1 via EIGRP is redistributed into BGP and it also becomes a VPNv4 route.
  ```
  R2#show bgp vpnv4 unicast all labels
  Network            Next Hop        In label/Out label
  Route Distinguisher: 192.168.255.2:65000 (A)
  192.168.12.0       0.0.0.0         24/nolabel(A)
  192.168.128.0      192.168.255.12  nolabel/24000
  192.168.255.1/32   192.168.12.1    25/nolabel
  192.168.255.8/32   192.168.255.12  nolabel/24007
  ```
- From the previous output, it can be concluded that to reach the 192.168.255.8/32; that is, the loopback1 of CE2, an outgoing label of 24007 is learned via BGP VPNv4 update. In a similar way, PE1 advertises the reachability to CE1's loopback1 via VPN label of 25.
  ```
  R2#show mpls forwarding-table
  Local   Outgoing  Prefix             Bytes Label  Outgoing   Next Hop
  Label   Label     or Tunnel Id       Switched     interface

  22      20        192.168.255.12/32  0            Fa1/0      192.168.24.4

  25      No Label  192.168.255.1/32[V]5976         Fa0/0      192.168.12.1
  ```
- The next hop to reach 192.168.255.8/32 is 192.168.255.12 and the next hop decides the LSP.

The MPLS forwarding table shows 20 as the outgoing label to reach 192.168.255.12. Hence traffic from CE1 going to CE2's loopback 1 will have 20 as the transport label and 24007 as the VPN label.

- For the return traffic destined to CE1's loopback1 the PHP operation would already have occurred on P1 as 192.168.255.1/32 belongs to CE1. The traffic destined to 192.168.255.1/32 will hit PE1 with a VPN label of 25 and this label will be removed and this packet will be sent to fa0/0 interface; that is, to CE1.

- The VPNv4 labels on RR1 reconfirm the same.

```
R3#show bgp vpnv4 unicast all labels
Network             Next Hop         In label/Out label
Route Distinguisher: 192.168.255.2:65000
192.168.255.1/32    192.168.255.2    nolabel/25
Route Distinguisher: 192.168.255.12:65001
192.168.255.8/32    192.168.255.12   nolabel/24007
```

- On P1 the traffic from CE1 destined to CE2 will hit with a transport label of 20.

```
R4#show mpls forwarding-table
Local  Outgoing  Pefix          Bytes Label Outgoing   Next Hop
Label  Label     or Tunnel Id   Switched    interface
20     22        192.168.255.12/32 5172      Fa1/1      192.168.45.5
```

- Now the traffic from CE1 destined to CE2 will hit ASBR1 with a transport label of 22.

```
R5#show mpls forwarding-table
Local  Outgoing  Prefix         Bytes Label Outgoing   Next Hop
Label  Label     or Tunnel Id   Switched    interface
22     24002     192.168.255.12/32 5928      Fa1/0      192.168.115.11
```

- Now the traffic from CE1 destined to CE2 will hit ASBR2 with a transport label of 24002.

```
RP/0/0/CPU0:ios#show mpls forwarding
Local    Outgoing   Prefix         Outgoing   Next Hop      Bytes
Label    Label      or ID          Interface                Switched
24002    19         192.168.255.12/32 Gi0/0/0/1  192.168.116.6  7092
```

- Now the traffic from CE1 destined to CE2 will hit P2 with a transport label of 19.

```
R6#show mpls forwarding-table
Local    Outgoing   Prefix         Bytes Label Outgoing  Next Hop
Label    Label      or Tunnel Id   Switched    interface
19       Pop Label  192.168.255.12/32 9928      Fa1/1     192.168.126.12
```

- It is observed on the P2 router that PHP operation takes place and transport label is popped. When the traffic hits PE2, it will hit with the VPN label of 24007 as discussed previously. It should also be observed that PE2 would be advertising reachability to CE2's loopback1 via VPN label of 24007.

```
RP/0/0/CPU0:ios#show mpls forwarding
Local    Outgoing    Prefix          Outgoing   Next Hop      Bytes
Label    Label       or ID           Interface                Switched
24007    Unlabelled  192.168.255.8/32[V] Gi0/0/0/1  192.168.128.6 7992
24008    18          192.168.255.2/32    Gi0/0/0/0  192.168.126.6 673200


RP/0/0/CPU0:ios#show bgp vpnv4 unicast labels
Network             Next Hop         Rcvd Label Local Label
Route Distinguisher: 192.168.255.12:65001 (default for vrf A)
*>i192.168.255.1/32  192.168.255.2    25         nolabel
*> 192.168.255.8/32  192.168.128.8    nolabel    24007
```

- It can be observed here that traffic from CE1 to CE2 hits PE2 with a VPN label to 24007, the traffic is sent to Gi/0/0/0/1 where CE2 is located and the VPN label is popped off. It is also observed that PE2 advertises the reachability to 192.168.255.8/32 via the VPN label of 24007. This same information was learned on PE1 earlier. Similarly the reachability to 192.168.255.1/32 was advertised by PE1 via the VPN label of 25 and the same information is learned here. In order to reach 192.168.255.1/32 on CE1 from CE2, a VPN label of 25 and

transport label of 18 will be used, since the next hop 192.168.255.2 is reachable via label 18.

Verification via Traceroutes

- The labels can be seen in the traceroute and they are exactly the same as discussed.
- Next hop in the VPNv4 update controls the label switch path and hence the transport label.
- In both of the traceroutes shown next, it can be observed that the VPN label remains consistent at all hops throughout the LSP. Only the transport label is swapped.
- When PE1 learns an update originated from PE2 then the next hop is PE2, not any RR or ASBR. This causes the LSP to be terminated at PE2, which results in a single LSP throughout the transit path from AS 65000 to AS 65001 and vice versa.

Traceroute from CE1 to CE2

```
R1#traceroute 192.168.255.8 source lo1
 Type escape sequence to abort.
 Tracing the route to 192.168.255.8
 VRF info: (vrf in name/id, vrf out name/id)
 1 192.168.12.2 8 msec 36 msec 16 msec
 2 192.168.24.4 [MPLS: Labels 20/24007 Exp 0] 828 msec 628 msec 2688 msec
 3 192.168.45.5 [MPLS: Labels 22/24007 Exp 0] 1456 msec * 1528 msec
 4 192.168.115.11 [MPLS: Labels 24002/24007 Exp 0] 1544 msec 2452 msec  2164 msec
 5 192.168.116.6 [MPLS: Labels 19/24007 Exp 0] 1036 msec 908 msec 1648 msec
 6 192.168.126.12 [MPLS: Label 24007 Exp 0] 2864 msec 1676 msec 1648 msec
 7 192.168.128.8 2008 msec 400 msec 572 msec
```
The VPN label 24007 remains consistent throughout the LSP.

Traceroute from CE2 to CE1

```
R8#traceroute 192.168.255.1 source lo1
 Type escape sequence to abort.
 Tracing the route to 192.168.255.1
 VRF info: (vrf in name/id, vrf out name/id)
 1 192.168.128.2 1228 msec 68 msec 152 msec
 2 192.168.126.6 [MPLS: Labels 18/25 Exp 0] 1188 msec 816 msec 1316 msec
 3 192.168.116.11 [MPLS: Labels 24007/25 Exp 0] 1384 msec 1816 msec 504  msec
 4 192.168.115.5 [MPLS: Labels 23/25 Exp 0] 284 msec 900 msec 972 msec
 5 192.168.45.4 [MPLS: Labels 17/25 Exp 0] 436 msec 608 msec 292 msec
 6 192.168.12.2 [MPLS: Label 25 Exp 0] 292 msec 108 msec 536 msec
 7 192.168.12.1 224 msec 212 msec 620 msec
```
The VPN label 25 remains consistent throughout the LSP.

Troubleshoot

There is currently no specific troubleshooting information available for this configuration.