Troubleshooting LSP Failure in MPLS VPN

Document ID: 23565

Contents

Introduction

This document assumes you have a prior understanding of basic Multiprotocol Label Switching (MPLS) concepts. MPLS−switched packets are forwarded based on information contained in the Label Forwarding Information Base (LFIB). A packet leaving a router over a label−switched interface will receive labels with values specified by the LFIB. Labels are associated with destinations in the LFIB according to Forwarding Equivalence Classes (FECs). A FEC is a grouping of IP packets which travel over the same path and receive the same forwarding treatment. The most simple example of a FEC is all packets traveling to a certain subnet. Another example could be all packets with a given IP precedence going to an Interior Gateway Protocol (IGP) next hop associated with a group of Border Gateway Protocol (BGP) routes.

The Label Information Base (LIB) is a structure which stores labels received from all Label Distribution Protocol (LDP) or Tag Distribution Protocol (TDP) neighbors. For Cisco implementation, labels are sent for all routes in a given router's routing table (with the exception of BGP routes), to all LDP or TDP neighbors. All labels received from neighbors are retained in the LIB, whether or not they are used. If the labels are received from a downstream neighbor for their FEC, then the labels stored in the LIB are used for packet forwarding by the LFIB. Meaning the labels used for forwarding are those received from a router's next hop to a destination, according to the router's Cisco Express Forwarding (CEF) and routing tables.

If label bindings are received from a downstream neighbor for prefixes (including subnet mask) which do not appear in a router's routing and CEF tables, these bindings will not be used. In a similar manner, if a router advertises labels for a subnet/subnet mask pair, which do not correspond to the routing updates also advertised by this router for the same subnet/subnet mask pair, these labels will not be used by upstream neighbors and the Label Switched Path (LSP) between these devices will fail.

This document gives an example of this kind of LSP failure and several possible solutions. The document covers one scenario wherein label bindings received by a router are not used to forward MPLS−switched packets. However, the steps used to diagnose and correct this problem are applicable to any problem involving label bindings and the LFIB on routers configured for MPLS.

Prerequisites

Requirements

There are no specific requirements for this document.
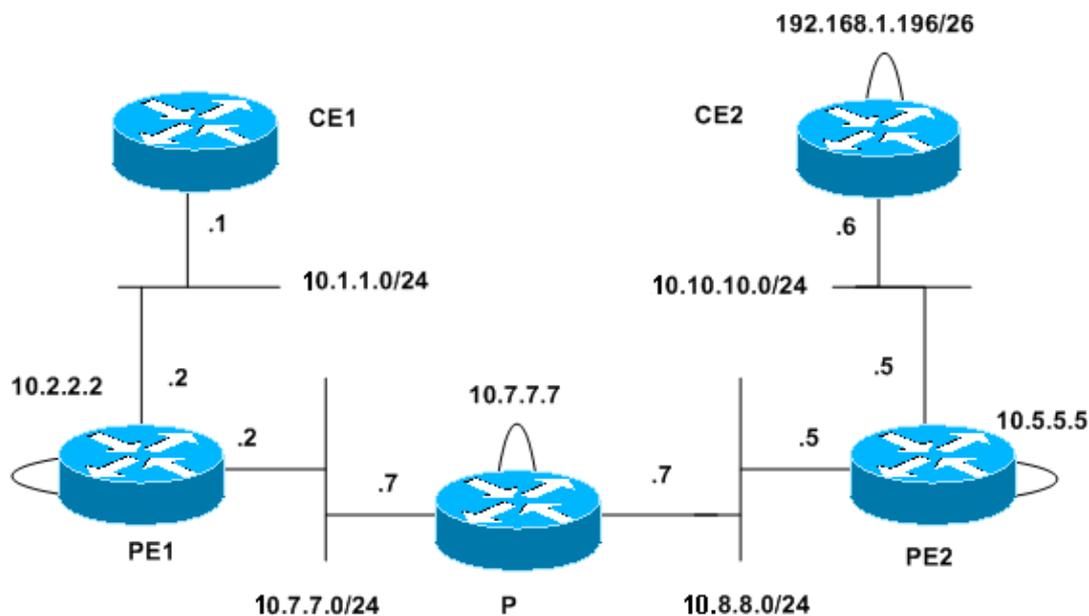
Components Used

The information in this document is based on this software version:

- Cisco IOS® Software release version 12.0(21)ST2

Conventions

Refer to Cisco Technical Tips Conventions for more information on document conventions.

Network Diagram

Router Configurations

<table>
<tr><td align="center">PE1 Router Configuration</td></tr>
</table>

```
ip vrf aqua
 rd 100:1
 route-target export 1:1
 route-target import 1:1
!
interface Loopback0
 ip address 10.2.2.2 255.255.255.255
 no ip directed-broadcast
!
interface Ethernet2/0/1
 ip vrf forwarding aqua
 ip address 10.1.1.2 255.255.255.0
 no ip directed-broadcast
 ip route-cache distributed

!--- The VPN Routing and Forwarding (VRF) interface
!--- toward the customer edge (CE) router.
```

```
interface Ethernet2/0/2
 ip address 10.7.7.2 255.255.255.0
 no ip directed-broadcast
 ip route-cache distributed
 tag-switching ip
!
router ospf 1
 log-adjacency-changes
 network 0.0.0.0 255.255.255.255 area 0
!
router bgp 1
 bgp log-neighbor-changes
 neighbor 10.5.5.5 remote-as 1
 neighbor 10.5.5.5 update-source Loopback0
 no auto-summary
 !
 address-family vpnv4
 neighbor 10.5.5.5 activate
 neighbor 10.5.5.5 send-community extended
 exit-address-family
 !
 address-family ipv4
 neighbor 10.5.5.5 activate
 no auto-summary
 no synchronization
 exit-address-family
 !
 address-family ipv4 vrf aqua
 redistribute connected
 no auto-summary
 no synchronization
 exit-address-family
```

P Router Configuration

```
interface Loopback0
 ip address 10.7.7.7 255.255.255.255
 no ip directed-broadcast
!
interface Ethernet2/0
 ip address 10.8.8.7 255.255.255.0
 no ip directed-broadcast
 tag-switching ip
!
interface Ethernet2/1
 ip address 10.7.7.7 255.255.255.0
 no ip directed-broadcast
 tag-switching ip
!
router ospf 1
 log-adjacency-changes
 network 0.0.0.0 255.255.255.255 area 0


!--- BGP is not run on this router.

```

PE2 Router Configuration

```
ip vrf aqua
 rd 100:1
 route-target export 1:1
 route-target import 1:1
!
```

```
interface Loopback0
 ip address 10.5.5.5 255.255.255.0
 no ip directed-broadcast
!
interface Ethernet0/0
 ip vrf forwarding aqua
 ip address 10.10.10.5 255.255.255.0
 no ip directed-broadcast
```

!--- The VRF interface toward the CE router.

```
!
interface Ethernet0/3
 ip address 10.8.8.5 255.255.255.0
 no ip directed-broadcast
 tag-switching ip
!
router ospf 1
 log-adjacency-changes
 network 0.0.0.0 255.255.255.255 area 0
!
router rip
 version 2
 !
 address-family ipv4 vrf aqua
 version 2
 network 10.0.0.0
 no auto-summary
 exit-address-family
!
router bgp 1
 bgp log-neighbor-changes
 neighbor 10.2.2.2 remote-as 1
 neighbor 10.2.2.2 update-source Loopback0
 no auto-summary
 !
 address-family vpnv4
 neighbor 10.2.2.2 activate
 neighbor 10.2.2.2 send-community extended
 exit-address-family
 !
 address-family ipv4
 neighbor 10.2.2.2 activate
 no auto-summary
 no synchronization
 exit-address-family
 !
 address-family ipv4 vrf aqua
 redistribute connected
 redistribute rip
 no auto-summary
 no synchronization
 exit-address-family
```

CE2 Router Configuration

```
interface Loopback0
 ip address 192.168.1.196 255.255.255.192
 no ip directed-broadcast
!
interface Ethernet1
 ip address 10.10.10.6 255.255.255.0
 no ip directed-broadcast
!
router rip
```

```
 version 2
 network 10.0.0.0
 network 192.168.1.0
 no auto-summary

!--- Routing Information Protocol (RIP) is used for the advertisement
!--- of routes between the CE and the provider edge (PE) router.

!
ip route 0.0.0.0 0.0.0.0 10.10.10.5
```

Note: The CE1 configuration has been omitted. The configuration consists of only IP addressing on the Ethernet interface and a static default route to 10.2.2.2.

Problem

The connectivity between CE1 and the loopback interface of CE2 has been lost, as shown in the following example.

```
CE1#ping 192.168.1.196

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.1.196, timeout is 2 seconds:
.....
Success rate is 0 percent (0/5)
```

However, CE1 has a valid routing entry for this destination, as shown in the following example.

```
CE1#show ip route 0.0.0.0
Routing entry for 0.0.0.0/0, supernet
  Known via "static", distance 1, metric 0, candidate default path
  Redistributing via ospf 100
  Routing Descriptor Blocks:
  * 10.1.1.2
      Route metric is 0, traffic share count is 1
```

At PE1 (the PE router attached to CE1), you can check MPLS VPN specific information. The following examples show that a valid route to the destination is present in the VRF table for this VPN.

```
PE1#show ip route vrf aqua 192.168.1.196
Routing entry for 192.168.1.192/26
  Known via "bgp 1", distance 200, metric 1, type internal
  Last update from 10.5.5.5 00:09:52 ago
  Routing Descriptor Blocks:
  * 10.5.5.5 (Default-IP-Routing-Table), from 10.5.5.5, 00:09:52 ago
      Route metric is 1, traffic share count is 1
      AS Hops 0, BGP network version 0

PE1#show tag-switching forwarding-table vrf aqua 192.168.1.196 detail
Local  Outgoing     Prefix            Bytes tag  Outgoing    Next Hop
tag    tag or VC    or Tunnel Id      switched   interface
None   16           192.168.1.192/26  0          Et2/0/2     10.7.7.7
        MAC/Encaps=14/22, MTU=1496, Tag Stack{16 32}
        00603E2B0241006083588742884 7 0001000000020000
        No output feature configured

PE1#show ip bgp vpnv4 vrf aqua 192.168.1.192
BGP routing table entry for 100:1:192.168.1.192/26, version 43
Paths: (1 available, best #1, table aqua)
  Not advertised to any peer
  Local
    10.5.5.5 (metric 21) from 10.5.5.5 (10.5.5.5)
```

```
              Origin incomplete, metric 1, localpref 100, valid, internal, best
              Extended Community: RT:1:1

      PE1#show tag-switching forwarding-table 10.5.5.5 detail
      Local  Outgoing    Prefix            Bytes tag  Outgoing    Next Hop
      tag    tag or VC   or Tunnel Id      switched   interface
      18     16          10.5.5.5/32       0          Et2/0/2     10.7.7.7
              MAC/Encaps=14/18, MTU=1500, Tag Stack{16}
              00603E2B0241006083587428847 00010000
              No output feature configured
           Per-packet load-sharing, slots: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
```

As shown in this example, PE1 does not have a route for the BGP next hop with the correct mask.

```
      PE1#
      PE1#show ip route 10.5.5.5 255.255.255.0
      % Subnet not in table
      PE1#show ip route 10.5.5.5 255.255.255.255
      Routing entry for 10.5.5.5/32
        Known via "ospf 1", distance 110, metric 21, type intra area
        Last update from 10.7.7.7 on Ethernet2/0/2, 00:38:55 ago
        Routing Descriptor Blocks:
        * 10.7.7.7, from 10.5.5.5, 00:38:55 ago, via Ethernet2/0/2
            Route metric is 21, traffic share count is 1
```

The IGP routing information used by PE1 to reach this BGP next hop is received from the P router. As shown in the following example, this router also shows an incorrect mask for the PE2 loopback and does not have a route for this prefix with the correct mask.

```
      P#show ip route 10.5.5.5
      Routing entry for 10.5.5.5/32
        Known via "ospf 1", distance 110, metric 11, type intra area
        Last update from 10.8.8.5 on Ethernet2/0, 00:47:48 ago
        Routing Descriptor Blocks:
        * 10.8.8.5, from 10.5.5.5, 00:47:48 ago, via Ethernet2/0
            Route metric is 11, traffic share count is 1

      P#show ip route 10.5.5.5 255.255.255.0
      % Subnet not in table
```

Cause of the LSP Failure

The LFIB and tag bindings on the P router show the cause of the LSP failure between this router and PE2. There is no outgoing label for 10.5.5.5. When the packet leaves PE1 it carries two labels, the BGP next hop label generated by the P router (16) and the VPN label generated by PE2 (32). Because this entry on the P router shows untagged, label–switched packets for this destination, it will be sent out without any labels. Since the VPN label 32 was lost, it will never be received by PE2, and PE2 will not have the correct information to forward the packet to the proper VPN destination.

```
      P#show tag-switching forwarding-table 10.5.5.5 detail
      Local  Outgoing    Prefix            Bytes tag  Outgoing    Next Hop
      tag    tag or VC   or Tunnel Id      switched   interface
      16     Untagged    10.5.5.5/32       5339       Et2/0       10.8.8.5
              MAC/Encaps=0/0, MTU=1504, Tag Stack{}
              No output feature configured
           Per-packet load-sharing, slots: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
```

As shown in the following example, the label binding table of the P router shows that PE2 (tsr: 10.8.8.5:0) only advertises a binding for 10.5.5.5 with a /24 mask. A label for the /32 route is advertised by the P router and PE1 (tsr: 10.2.2.2:0), but not PE2. Because the binding advertised by PE2 does not match the route it also advertises, no label is present in the LFIB of the P router to forward packets to this destination.

```
P#show tag-switching tdp bindings detail

  tib entry: 10.5.5.0/24, rev 67(no route)
       remote binding: tsr: 10.8.8.5:0, tag: imp-null
  tib entry: 10.5.5.5/32, rev 62
       local binding:  tag: 16
         Advertised to:
         10.2.2.2:0              10.8.8.5:0
       remote binding: tsr: 10.2.2.2:0, tag: 18
```

The reason for the discrepancy between the routing updates and label bindings advertised by PE2 can be seen in the routing table and tag binding table of this router. The directly connected loopback shows the correct /24 mask, this is used by the router in generating the label binding. Because this network uses Open Shortest Path First (OSPF), the router advertises this interface with a /32 mask, as shown in the following example.

```
PE2#show ip route 10.5.5.5
Routing entry for 10.5.5.0/24
  Known via "connected", distance 0, metric 0 (connected, via interface)
  Routing Descriptor Blocks:
  * directly connected, via Loopback0
      Route metric is 0, traffic share count is 1

PE2#show tag-switching tdp bindings detail

  tib entry: 10.5.5.0/24, rev 142
       local binding:  tag: imp-null
         Advertised to:
         10.7.7.7:0
  tib entry: 10.5.5.5/32, rev 148
       remote binding: tsr: 10.7.7.7:0, tag: 16

PE2#show ip ospf interface loopback 0
Loopback0 is up, line protocol is up
  Internet Address 10.5.5.5/24, Area 0
  Process ID 1, Router ID 10.5.5.5, Network Type LOOPBACK, Cost: 1
  Loopback interface is treated as a stub Host


!--- OSPF advertises all interfaces of Network Type LOOPBACK as host
!--- routes (/32).
```

Solutions

Because the failure of the LSP between the P router and PE1 was caused by a mismatch between the route advertised for the loopback and the label binding generated by PE1, the most simple solution is to change the mask of the loopback to conform to the mask advertised by OSPF for all networks of the LOOPBACK type.

Solution 1: Change of Subnet Mask on PE2

```
PE2#configure terminal
  Enter configuration commands, one per line.  End with CNTL/Z.
  PE2(config)#int lo 0
  PE2(config-if)#ip add 10.5.5.5 255.255.255.255
  PE2(config-if)#end
  PE2#
```

The information on PE1 appears the same as in the scenario where LSP failure occurs, as shown in the following example.

```
PE1#show tag-switching forwarding-table vrf aqua 192.168.1.196 detail
```

```
              Local   Outgoing     Prefix                   Bytes tag   Outgoing    Next Hop
              tag     tag or VC    or Tunnel Id             switched    interface
              None    16                 192.168.1.192/26   0           Et2/0/2     10.7.7.7
                      MAC/Encaps=14/22, MTU=1496, Tag       Stack{16 32}
                      00603E2B02410060835887428847 0001000000020000
                      No output feature configured


      PE1#show tag-switching forwarding-table 10.5.5.5 detail
              Local   Outgoing     Prefix                   Bytes tag   Outgoing    Next Hop
              tag     tag or VC    or Tunnel Id             switched    interface
              18      16                 10.5.5.5/32         0           Et2/0/2     10.7.7.7
                      MAC/Encaps=14/18, MTU=1500, Tag       Stack{16}
                      00603E2B02410060835887428847 00010000
                      No output feature configured
             Per-packet load-sharing, slots: 0 1 2 3 4 5 6 7 8 9 10    11 12 13 14 15
```

The P router shows that the conditions which caused the LSP failure are no longer present. The outgoing label is now a pop tag. This means that the top label for the BGP next hop will be popped as the packets traverse the router, but the packets will still have the second VPN label (the packets are no longer sent out untagged).

The tag binding table shows a label (imp–null) is advertised by PE2 (tsr: 10.8.8.5:0) for the /32 route.

```
      P#show tag-switching forwarding-table 10.5.5.5 detail
          Local   Outgoing     Prefix              Bytes tag   Outgoing    Next Hop
          tag     tag or VC    or Tunnel Id        switched    interface
          16      Pop tag      10.5.5.5/32         3493        Et2/0          10.8.8.5
                   MAC/Encaps=14/14, MTU=1504, Tag Stack{}
                   006009E08B0300603E2B02408847
                   No output feature configured

          Per-packet load-sharing, slots: 0 1 2 3 4 5 6 7 8 9 10 11    12 13 14 15


      P#show tag-switching tdp bindings detail

          tib entry: 10.5.5.5/32, rev 71
               local binding:  tag: 16
                 Advertised to:
                  10.2.2.2:0                       10.8.8.5:0
               remote binding: tsr: 10.2.2.2:0,      tag: 18
               remote binding: tsr: 10.8.8.5:0,      tag: imp-null
```

Solution 2: OSPF Network Type Change

The second solution is to change the OSPF network type of the loopback interface. When the OSPF network type of PE2's loopback interface is changed to point–to–point, the loopback prefix is no longer automatically advertised with a /32 mask. This means that the label binding generated by PE2, when referencing the directly–connected subnet in its routing table (containing a /24 subnet mask), will match the OSPF route on the P router received from PE2 (containing a /24 subnet mask for this prefix).

The **ip ospf network point–to–point** command can be used to change the network type on the PE2 loopback interface, as shown in the following example.

```
      PE2#configure terminal
      Enter configuration commands, one per line.  End with CNTL/Z.
      PE2(config)#interface loopback 0
      PE2(config-if)#ip ospf network point-to-point
      PE2(config-if)#
```

As shown below, the tag forwarding table on PE1 contains an entry for the BGP next hop, which is consistent with the actual mask of the loopback interface on PE2. The routing table shows the OSPF route associated with this forwarding entry is also correct.

```
PE1#show tag-switching forwarding-table 10.5.5.5 detail
Local  Outgoing    Prefix            Bytes tag  Outgoing   Next Hop
tag    tag or VC   or Tunnel Id      switched   interface
22     17          10.5.5.0/24       0          Et2/0/2    10.7.7.7
         MAC/Encaps=14/18, MTU=1500, Tag Stack{17}
         00603E2B0241006083588742847 00011000
         No output feature configured
      Per-packet load-sharing, slots: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

PE1#show ip route 10.5.5.5
Routing entry for 10.5.5.0/24
  Known via "ospf 1", distance 110, metric 21, type intra area
  Last update from 10.7.7.7 on Ethernet2/0/2, 00:36:53 ago
  Routing Descriptor Blocks:
  * 10.7.7.7, from 10.5.5.5, 00:36:53 ago, via Ethernet2/0/2
      Route metric is 21, traffic share count is 1
```

In the example below, the tag forwarding entry of the P router shows the outgoing tag as a pop tag, as in Solution 1, as shown in example below. Once again, the top label for the BGP next hop will be popped as the packet traverses this router, but the second VPN label will be retained and the LSP will not fail. The binding showing the correct subnet mask is also present.

```
P#show tag-switching forwarding-table 10.5.5.5 detail
Local  Outgoing    Prefix            Bytes tag  Outgoing   Next Hop
tag    tag or VC   or Tunnel Id      switched   interface
17     Pop tag     10.5.5.0/24       4261       Et2/0      10.8.8.5
         MAC/Encaps=14/14, MTU=1504, Tag Stack{}
         006009E08B0300603E2B02408847
         No output feature configured
      Per-packet load-sharing, slots: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15


P#show tag-switching tdp bindings detail

  tib entry: 10.5.5.0/24, rev 68
        local binding:  tag: 17
          Advertised to:
          10.2.2.2:0              10.8.8.5:0
        remote binding: tsr: 10.8.8.5:0, tag: imp-null
        remote binding: tsr: 10.2.2.2:0, tag: 22
```

As shown below, the output of this command confirms that the network type has been changed to point−to−point. Full connectivity is present from CE1 to the loopback interface of CE2.

```
PE2#show ip ospf interface loopback 0
Loopback0 is up, line protocol is up
  Internet Address 10.5.5.5/24, Area 0
  Process ID 1, Router ID 10.5.5.5, Network Type POINT_TO_POINT, Cost: 1
  Transmit Delay is 1 sec, State POINT_TO_POINT,
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
  Index 3/3, flood queue length 0
  Next 0x0(0)/0x0(0)
  Last flood scan length is 0, maximum is 0
  Last flood scan time is 0 msec, maximum is 0 msec
  Neighbor Count is 0, Adjacent neighbor count is 0
  Suppress hello for 0 neighbor(s)

CE1#ping 192.168.1.196

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.1.196, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 4/4/4 ms
CE1.
```

Related Information

- **IP Routing Technology Support**
- **How to Troubleshoot the MPLS VPN**
- **MPLS Technology Support**
- **Technical Support & Documentation – Cisco Systems**

Updated: Jan 18, 2008   Document ID: 23565