验证MPLS第3层VPN转发

目录

简介

本文档介绍验证跨MPLS第3层VPN核心网络的端到端连接的过程。

先决条件

要求

Cisco 建议您了解以下主题：

- 基本IP路由的相关知识
- 了解Cisco IOS® XE和Cisco IOS® XR命令行

使用的组件

本文档中的信息基于以下软件和硬件版本：

- 使用Cisco IOS XR软件的路由器
- 使用Cisco IOS XE软件的路由器

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

背景信息

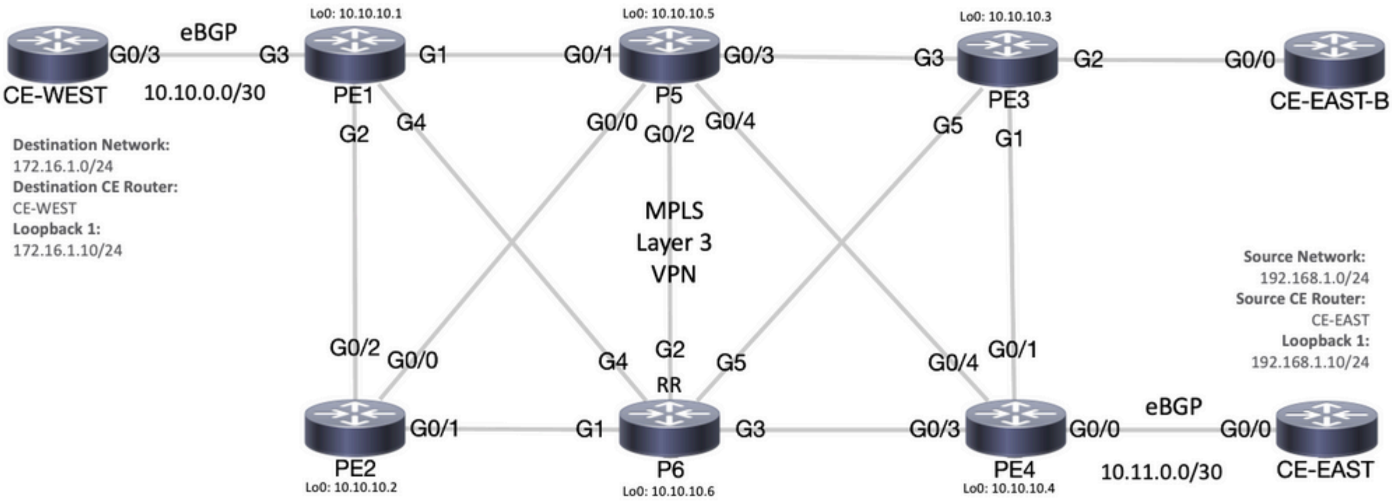
本文档旨在演示基本的验证和故障排除步骤，以检查通过混合使用Cisco IOS XE和Cisco IOS XR路

由器作为PE（提供商边缘）和P（提供商）路由器的MPLS第3层VPN核心网络与BGP（边界网关协议）互连的两台CE（客户边缘）路由器之间的连接和转发。

规则

有关文件规则的更多信息请参见“ Cisco技术提示规则”。

拓扑



MPLS拓扑图

故障排除

初始信息

源网络：192.168.1.0/24

源CE路由器：CE-EAST

目的网络：172.16.1.0/24

目的CE路由器：CE-WEST

根据初始信息和拓扑，在路由器CE-EAST上由Loopback1表示的源地址192.168.1.10和路由器CE-WEST上由Loopback1表示的目的地地址172.16.1.10之间必须能够成功通信：

<#root>

```
CE-EAST#

show run interface loopback1

Building configuration...

Current configuration : 66 bytes
```

```
!
interface Loopback1
 ip address 192.168.1.10 255.255.255.0
end
```

CE-WEST#

show run interface loopback 1

```
Building configuration...

Current configuration : 65 bytes
!
interface Loopback1
 ip address 172.16.1.10 255.255.255.0
end
```

ICMP可达性和traceroute用于开始检查这些源地址和目标地址之间的连接，但是从接下来的输出中可看到此操作不成功：

<#root>

CE-EAST#

ping 172.16.1.10 source loopback1

```
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.1.10, timeout is 2 seconds:
Packet sent with a source address of 192.168.1.10
.....
Success rate is 0 percent (0/5)
```

CE-EAST#

traceroute 172.16.1.10 source loop1 probe 1 numeric

```
Type escape sequence to abort.
Tracing the route to 172.16.1.10
VRF info: (vrf in name/id, vrf out name/id)
  1 10.11.0.2 2 msec
  2  *
  3 10.10.0.2 [MPLS: Label 16 Exp 0] 9 msec
  4  *
  5  *
  6  *
  7  *
  8  *
  9  *
 10  *
 11  *
 12  *
 13  *
 14  *
 15  *
 16  *
 17  *
 18  *
 19  *
 20  *
 21  *
```

```
 22  *
 23  *
 24  *
 25  *
 26  *
 27  *
 28  *
 29  *
 30  *
CE-EAST#
```

注意：在排除故障时，在连接到MPLS网络时使用traceroute可能不太有效，因为一些服务提供商倾向于在Cisco IOS XE中配置no mpls ip propagate-ttl forward命令或在Cisco IOS XR中配置mpls ip-ttl-propagate disable forwarded命令，以隐藏核心中的所有LSR（标签交换路由器）（但入口和出口PE路由器除外）。

在查看源CE路由器的状态时，由于此路由器没有任何VRF（虚拟路由转发）并且不支持MPLS感知，您需要验证RIB（路由信息库）、CEF（思科快速转发）和BGP。在接下来的输出中，可以看到

有一个路由条目通过BGP获知到目标子网172.16.1.0/24，可通过接口GigabitEthernet0/0到达：

<#root>

CE-EAST#

show ip route 172.16.1.10

```
Routing entry for 172.16.1.0/24
```

Known via "bgp 65001", distance 20, metric 0

```
          <<<<<
  Tag 65500, type external
  Last update from 10.11.0.2 3d01h ago
  Routing Descriptor Blocks:
  * 10.11.0.2, from 10.11.0.2, 3d01h ago
      Route metric is 0, traffic share count is 1
      AS Hops 2
      Route tag 65500
      MPLS label: none

CE-EAST#
```

show ip cef 172.16.1.10

```
172.16.1.0/24
```

nexthop 10.11.0.2 GigabitEthernet0/0

```
          <<<<<
CE-EAST#
```

由于源CE-EAST路由器在RIB中安装了通往目标的路由，因此查看提供商边缘路由器PE4（入口PE）的时间较长，如拓扑所示。此时，配置了VRF和路由区分符以及路由目标导入和导出，如以下输出所示：

<#root>

RP/0/0/CPU0:PE4#

show run vrf EAST

```
Mon Sep 11 20:01:54.454 UTC
vrf EAST
 address-family ipv4 unicast
```

import route-target 65000:1 65001:1 65001:2 ! export route-target 65001:1

```
  !
 !
!

RP/0/0/CPU0:PE4#
```

show run router bgp

```
Mon Sep 11 20:06:48.164 UTC
router bgp 65500
 address-family ipv4 unicast
 !
 address-family vpnv4 unicast
 !
 neighbor 10.10.10.6
  remote-as 65500
  update-source Loopback0
  address-family vpnv4 unicast
  !
 !
 vrf EAST
```

rd 65001:1

```
  address-family ipv4 unicast
  !
  neighbor 10.11.0.1
   remote-as 65001
   address-family ipv4 unicast
    route-policy PASS in
    route-policy PASS out
   !
  !
 !
!

RP/0/0/CPU0:PE4#
```

从先前的输出中可以看到，VRF名称“EAST”是使用65000：1的路由-目标导入定义的，现在可以检查VRF路由表，这有助于确定PE4是否具有通往目标IP地址172.16.1.10的路由：

<#root>

RP/0/0/CPU0:PE4#

show route vrf EAST 172.16.1.10

```
Mon Sep 11 19:58:28.128 UTC

Routing entry for 172.16.1.0/24
  Known via "bgp 65500", distance 200, metric 0
  Tag 65000, type internal
  Installed Sep  8 18:28:46.303 for 3d01h
  Routing Descriptor Blocks
    10.10.10.1, from 10.10.10.6
      Nexthop in Vrf: "default", Table: "default", IPv4 Unicast, Table Id: 0xe0000000
      Route metric is 0
  No advertising protos.
RP/0/0/CPU0:PE4#
```

因为此PE是思科IOS XR设备，所以可在show route vrf <name>命令末尾使用“detail”关键字来查看

其他一些信息，例如由MP-BGP（多协议BGP）强加的VPNv4标签和源RD（路由区分符）与前缀：

<#root>

RP/0/0/CPU0:PE4#

show route vrf EAST 172.16.1.10 detail

```
Mon Sep 11 20:21:48.492 UTC

Routing entry for 172.16.1.0/24
  Known via "bgp 65500", distance 200, metric 0
  Tag 65000, type internal
  Installed Sep  8 18:28:46.303 for 3d01h
  Routing Descriptor Blocks
    10.10.10.1, from 10.10.10.6
      Nexthop in Vrf: "default", Table: "default", IPv4 Unicast, Table Id: 0xe0000000
      Route metric is 0
```

Label: 0x10 (16)

```
          <<<<<
      Tunnel ID: None
      Binding Label: None
      Extended communities count: 0
```

Source RD attributes: 0x0000:65000:1

```
         <<<<<
      NHID:0x0(Ref:0)
  Route version is 0x5 (5)
  No local label
  IP Precedence: Not Set
  QoS Group ID: Not Set
  Flow-tag: Not Set
  Fwd-class: Not Set
  Route Priority: RIB_PRIORITY_RECURSIVE (12) SVD Type RIB_SVD_TYPE_REMOTE
  Download Priority 3, Download Version 36
  No advertising protos.
RP/0/0/CPU0:PE4#
```

现在，让我们看看导入到VRF中的BGP VPNv4前缀，观察到，这是从上一个输出中得到的相同标签16，并且它还具有扩展社区65000：1。此外，请注意10.10.10.1是PE4需要对其执行路由递归的下一跳，下一个地址“from 10.10.10.6”是PE4用于获取此前缀的BGP对等体（在本场景中是路由反射器P6）：

<#root>

RP/0/0/CPU0:PE4#

show bgp vpnv4 unicast vrf EAST 172.16.1.10

```
Mon Sep 11 22:42:28.114 UTC
```

```
BGP routing table entry for 172.16.1.0/24, Route Distinguisher: 65001:1
Versions:
  Process           bRIB/RIB  SendTblVer
  Speaker                 48          48
Last Modified: Sep  8 18:28:46.314 for 3d04h
Paths: (1 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  65000
```

10.10.10.1 (metric 20) from 10.10.10.6 (10.10.10.1)

<<<<<

Received Label 16

```
      Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate, importe
      Received Path ID 0, Local Path ID 0, version 48
```

Extended community: RT:65000:1

<<<<<

```
      Originator: 10.10.10.1, Cluster list: 10.10.10.6
```

Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 65000:1

<<<<<

通过在VRF级别使用exact-route关键字查看CEF，您可以了解数据包的送出接口。此命令还可以提供一些重要的详细信息，因为它显示施加到前缀24001和16的两个标签，原因是标签16来自BGP VPNv4，标签24001来自LDP（标签分发协议）：

<#root>

RP/0/0/CPU0:PE4#

show cef vrf EAST exact-route 192.168.1.10 172.16.1.10

```
Mon Sep 11 22:48:15.241 UTC
172.16.1.0/24, version 36, internal 0x5000001 0x0 (ptr 0xa12dc74c) [1], 0x0 (0x0), 0x208 (0xa155b1b8)
 Updated Sep  8 18:28:46.323
 local adjacency 10.0.0.16
 Prefix Len 24, traffic index 0, precedence n/a, priority 3
   via GigabitEthernet0/0/0/4
   via 10.10.10.1/32, 3 dependencies, recursive [flags 0x6000]
    path-idx 0 NHID 0x0 [0xa15c3f54 0x0]
    recursion-via-/32
    next hop VRF - 'default', table - 0xe0000000
    next hop 10.10.10.1/32 via 24010/0/21
```

next hop 10.0.0.16/32 Gi0/0/0/4 labels imposed {24001 16}

<<<<<

下一步，使用show bgp vpnv4 unicast命令检查由此PE获取的VPNv4路由。此输出显示将VPNv4前缀导入VRF之前的信息，请记住，配置的RT（路由目标）(例如，导入的RT是65000：1、65001：1、65001：2)表示导入的路由和导入的VRF：

<#root>

RP/0/0/CPU0:PE4#

show bgp vpnv4 unicast

```
Fri Sep 15 02:15:15.463 UTC
BGP router identifier 10.10.10.4, local AS number 65500
BGP generic scan interval 60 secs
Non-stop routing is enabled
BGP table state: Active
Table ID: 0x0   RD version: 0
BGP main routing table version 85
BGP NSR Initial initsync version 1 (Reached)
BGP NSR/ISSU Sync-Group versions 0/0
BGP scan interval 60 secs

Status codes: s suppressed, d damped, h history, * valid, > best
              i - internal, r RIB-failure, S stale, N Nexthop-discard
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network            Next Hop            Metric LocPrf Weight Path
```

Route Distinguisher: 65000:1

***>i172.16.1.0/24 10.10.10.1 0 100 0 65000 i**

```
         <<<<<
*>i172.16.2.0/24      10.10.10.1               0    100      0 65000 i
Route Distinguisher: 65001:1 (default for vrf EAST)
* i0.0.0.0/0          10.10.10.3               0    100      0 65001 i
*>                    10.11.0.1                0             0 65001 i
*>i172.16.1.0/24      10.10.10.1               0    100      0 65000 i
*>i172.16.2.0/24      10.10.10.1               0    100      0 65000 i
*> 192.168.1.0/24     10.11.0.1                0             0 65001 i
*>i192.168.2.0/24     10.10.10.3               0    100      0 65001 i
*> 192.168.3.0/24     10.11.0.1                0             0 65001 i
Route Distinguisher: 65001:2
*>i0.0.0.0/0          10.10.10.3               0    100      0 65001 i
*>i192.168.2.0/24     10.10.10.3               0    100      0 65001 i

Processed 10 prefixes, 11 paths
```

在本例中，VPNv4表可能很小，但在生产环境中，您可以使用以下命令将验证范围缩小到特定RD和前缀，而不是查看所有VPNv4前缀：

<#root>

RP/0/0/CPU0:PE4#

show bgp vpnv4 unicast rd 65000:1 172.16.1.10

```
Mon Sep 11 22:54:04.967 UTC
BGP routing table entry for 172.16.1.0/24, Route Distinguisher: 65000:1
Versions:
  Process           bRIB/RIB  SendTblVer
  Speaker                 46          46
Last Modified: Sep  8 18:28:46.314 for 3d04h
Paths: (1 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  65000
    10.10.10.1 (metric 20) from 10.10.10.6 (10.10.10.1)
      Received Label 16
      Origin IGP, metric 0, localpref 100, valid, internal, best, group-best, import-candidate, not-in-
      Received Path ID 0, Local Path ID 0, version 46
      Extended community: RT:65000:1
      Originator: 10.10.10.1, Cluster list: 10.10.10.6
```

此时，MP-BGP控制平面具有目标前缀以及LDP和VPNv4标签{24001 16}，此流量的送出接口似乎是Gi0/0/0/4，需要转发流量的下一跳是10.10.10.1。但是，是否还有其它选项可用于检验首选送出接口？是时候了解MPLS转发表或LFIB（标签转发信息库）了。使用命令show mpls forwarding将显示指向10.10.10.1目标（来自PE1的Loopback0）的两个条目，一条路径具有传出接口Gi0/0/0/4和下一跳10.0.0.16（路由器P5），其中强加的传出标签为24001，另一条路径通过Gi0/0/0/3，下一跳10.0.13（路由器P6）和传出标签23：

<#root>

RP/0/0/CPU0:PE4#

show mpls forwarding

```
Mon Sep 11 23:28:33.425 UTC
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
24000  Unlabelled  192.168.1.0/24[V]  Gi0/0/0/0    10.11.0.1       1096
24001  Unlabelled  192.168.3.0/24[V]  Gi0/0/0/0    10.11.0.1       56056
24002  Unlabelled  0.0.0.0/0[V]       Gi0/0/0/0    10.11.0.1       0
24003  Pop         10.10.10.6/32      Gi0/0/0/3    10.0.0.13       7778512
24004  Pop         10.0.0.4/31        Gi0/0/0/3    10.0.0.13       0
24005  Pop         10.0.0.8/31        Gi0/0/0/3    10.0.0.13       0
24006  Pop         10.10.10.5/32      Gi0/0/0/4    10.0.0.16       3542574
24007  Pop         10.0.0.10/31       Gi0/0/0/3    10.0.0.13       0
       Pop         10.0.0.10/31       Gi0/0/0/4    10.0.0.16       0
24008  Pop         10.0.0.6/31        Gi0/0/0/4    10.0.0.16       0
24009  Pop         10.0.0.0/31        Gi0/0/0/4    10.0.0.16       0
```

24010 23 10.10.10.1/32 Gi0/0/0/3 10.0.0.13 22316

<<<<<

24001 10.10.10.1/32 Gi0/0/0/4 10.0.0.16 42308

```
       <<<<<
24011  18          10.10.10.2/32      Gi0/0/0/3    10.0.0.13       0
       24003       10.10.10.2/32      Gi0/0/0/4    10.0.0.16       0
24012  17          10.0.0.2/31        Gi0/0/0/3    10.0.0.13       0
```

```
       24005       10.0.0.2/31        Gi0/0/0/4    10.0.0.16       0
24013  Pop         10.10.10.3/32      Gi0/0/0/1    10.0.0.20       3553900
24014  Pop         10.0.0.14/31       Gi0/0/0/1    10.0.0.20       0
       Pop         10.0.0.14/31       Gi0/0/0/4    10.0.0.16       0
24015  Pop         10.0.0.18/31       Gi0/0/0/1    10.0.0.20       0
       Pop         10.0.0.18/31       Gi0/0/0/3    10.0.0.13       0
```

RP/0/0/CPU0:PE4#

show mpls forwarding prefix 10.10.10.1/32

```
Mon Sep 11 23:30:54.685 UTC
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
24010  23          10.10.10.1/32      Gi0/0/0/3    10.0.0.13       3188
       24001       10.10.10.1/32      Gi0/0/0/4    10.0.0.16       6044
```

RP/0/0/CPU0:PE4#

show mpls forwarding prefix 10.10.10.1/32 detail hardware egress

```
Mon Sep 11 23:36:06.504 UTC
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
24010  23          10.10.10.1/32      Gi0/0/0/3    10.0.0.13       N/A
     Updated: Sep  8 20:27:26.596
     Version: 39, Priority: 3
     Label Stack (Top -> Bottom): { 23 }
     NHID: 0x0, Encap-ID: N/A, Path idx: 0, Backup path idx: 0, Weight: 0
     MAC/Encaps: 14/18, MTU: 1500
     Outgoing Interface: GigabitEthernet0/0/0/3 (ifhandle 0x000000a0)
     Packets Switched: 0

       24001       10.10.10.1/32      Gi0/0/0/4    10.0.0.16       N/A
     Updated: Sep  8 20:27:26.596
     Version: 39, Priority: 3
     Label Stack (Top -> Bottom): { 24001 }
     NHID: 0x0, Encap-ID: N/A, Path idx: 1, Backup path idx: 0, Weight: 0
     MAC/Encaps: 14/18, MTU: 1500
     Outgoing Interface: GigabitEthernet0/0/0/4 (ifhandle 0x000000c0)
     Packets Switched: 0
```

从前面的输出可以清楚地看到，流量可以负载均衡的路径选项有两个，但是有几种方法可以帮助确定哪一个是首选路径。一种方式是使用show cef exact-route <source IP> <destination IP>命令，方法是添加源PE的Loopback0和目标PE的Loopback0。如下一个输出所示，首选路径是通过Gi0/0/0/4：

<#root>

RP/0/0/CPU0:PE4#

show cef exact-route 10.10.10.4 10.10.10.1

```
Mon Sep 11 23:49:44.558 UTC
```

```
10.10.10.1/32, version 39, internal 0x1000001 0x0 (ptr 0xa12dbdbc) [1], 0x0 (0xa12c18c0), 0xa28 (0xa185
 Updated Sep  8 20:27:26.596
 local adjacency 10.0.0.16
 Prefix Len 32, traffic index 0, precedence n/a, priority 3
   via GigabitEthernet0/0/0/4
```

via 10.0.0.16/32, GigabitEthernet0/0/0/4, 9 dependencies, weight 0, class 0 [flags 0x0]

```
         <<<<<
    path-idx 1 NHID 0x0 [0xa16765bc 0x0]
    next hop 10.0.0.16/32
    local adjacency
     local label 24010      labels imposed {24001}
```

另一个选项是首先验证LIB（标签信息库）并通过使用show mpls ldp bindings <prefix/mask>命令获取目标Loopback0 （属于出口PE的10.10.10.1）的LDP绑定，然后从该输出中找到本地绑定标签后，使用该标签值在show mpls forwarding exact-route label <label> ipv4 <source IP> <destination IP> detail命令中查找首选路径：

<#root>

RP/0/0/CPU0:PE4#

show mpls ldp bindings 10.10.10.1/32

```
Wed Sep 13 17:18:43.007 UTC
10.10.10.1/32, rev 29
```

Local binding: label: 24010

```
         <<<<<
        Remote bindings: (3 peers)
            Peer                Label
            -----------------   ---------
            10.10.10.3:0        24
            10.10.10.5:0        24001
            10.10.10.6:0        23
```

RP/0/0/CPU0:PE4#

show mpls forwarding exact-route label 24010 ipv4 10.10.10.4 10.10.10.1 detail

```
Wed Sep 13 17:20:06.342 UTC
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
```

24010 24001 10.10.10.1/32 Gi0/0/0/4 10.0.0.16 N/A

```
                  <<<<<
     Updated: Sep 12 14:15:37.009
     Version: 198, Priority: 3
     Label Stack (Top -> Bottom): { 24001 }
     NHID: 0x0, Encap-ID: N/A, Path idx: 1, Backup path idx: 0, Weight: 0
     Hash idx: 1
     MAC/Encaps: 14/18, MTU: 1500
     Outgoing Interface: GigabitEthernet0/0/0/4 (ifhandle 0x000000c0)
```

```
Packets Switched: 0

Via: Gi0/0/0/4, Next Hop: 10.0.0.16
Label Stack (Top -> Bottom): { 24001 }
NHID: 0x0, Encap-ID: N/A, Path idx: 1, Backup path idx: 0, Weight: 0
Hash idx: 1
MAC/Encaps: 14/18, MTU: 1500
Outgoing Interface: GigabitEthernet0/0/0/4 (ifhandle 0x000000c0)
```

接下来，必须检查数据平面中的下一跳路由器，在本例中，要验证的路由器是P5（即接口为10.0.0.16）。首先要了解的是MPLS转发表，其中前缀10.10.10.1的本地标签必须24001：

<#root>

RP/0/0/CPU0:P5#

show mpls forwarding

```
Thu Sep 14 20:07:16.455 UTC
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
24000  Pop         10.10.10.6/32      Gi0/0/0/2    10.0.0.11       361906
```

24001 Pop 10.10.10.1/32 Gi0/0/0/1 10.0.0.0 361002

```
               <<<<<
24002  Pop         10.0.0.4/31        Gi0/0/0/1    10.0.0.0        0
       Pop         10.0.0.4/31        Gi0/0/0/2    10.0.0.11       0
24003  Pop         10.10.10.2/32      Gi0/0/0/0    10.0.0.6        360940
24004  Pop         10.0.0.8/31        Gi0/0/0/0    10.0.0.6        0
       Pop         10.0.0.8/31        Gi0/0/0/2    10.0.0.11       0
24005  Pop         10.0.0.2/31        Gi0/0/0/0    10.0.0.6        0
       Pop         10.0.0.2/31        Gi0/0/0/1    10.0.0.0        0
24006  Pop         10.10.10.4/32      Gi0/0/0/4    10.0.0.17       361230
24007  Pop         10.0.0.12/31       Gi0/0/0/2    10.0.0.11       0
       Pop         10.0.0.12/31       Gi0/0/0/4    10.0.0.17       0
24008  Pop         10.10.10.3/32      Gi0/0/0/3    10.0.0.15       361346
24009  Pop         10.0.0.20/31       Gi0/0/0/3    10.0.0.15       0
       Pop         10.0.0.20/31       Gi0/0/0/4    10.0.0.17       0
24010  Pop         10.0.0.18/31       Gi0/0/0/2    10.0.0.11       0
       Pop         10.0.0.18/31       Gi0/0/0/3    10.0.0.15       0

RP/0/0/CPU0:P5#
```

show mpls forwarding labels 24001

```
Thu Sep 14 20:07:42.584 UTC
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
```

24001 Pop 10.10.10.1/32 Gi0/0/0/1 10.0.0.0 361060

```
RP/0/0/CPU0:P5#
```

从先前的输出中可以看到，前缀10.10.10.1/32的LFIB条目显示“Pop”作为传出标签，这意味着此路由器是倒数第二跳跳跳跃(PHP)。它还显示必须根据LFIB信息通过Gi0/0/0/1发送流量，在查看CEF时也可以进行验证。下一个CEF exact-route输出显示隐式空标签作为强加的标签，这同样是因为在Gi0/0/0/1连接的下一跳是标签交换机路径中的最后一个路由器，也是面向目标站点(CE-WEST)的PE。这也是路由器P5删除数据包且不向数据包施加其他标签的原因，由于此过程，出口路由器PE1将接收不带LDP标签的数据包：

<#root>

RP/0/0/CPU0:P5#

show cef exact-route 10.10.10.4 10.10.10.1

```
Thu Sep 14 20:25:57.269 UTC
10.10.10.1/32, version 192, internal 0x1000001 0x0 (ptr 0xa1246394) [1], 0x0 (0xa122b638), 0xa20 (0xa15
 Updated Sep 12 14:15:38.009
 local adjacency 10.0.0.0
 Prefix Len 32, traffic index 0, precedence n/a, priority 3
   via GigabitEthernet0/0/0/1
   via 10.0.0.0/32, GigabitEthernet0/0/0/1, 9 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 NHID 0x0 [0xa166e280 0xa166e674]
    next hop 10.0.0.0/32
    local adjacency
```

local label 24001 labels imposed {ImplNull}

```
         <<<<<
```

检验标签交换机路径的最后一点是PE1。在查看MPLS转发表时，可以注意到LFIB中没有前缀10.10.10.1/32的条目：

<#root>

PE1#

show mpls forwarding-table

```
Local      Outgoing   Prefix           Bytes Label   Outgoing   Next Hop
Label      Label      or Tunnel Id     Switched      interface
16         No Label   172.16.1.0/24[V] 12938         Gi3        10.10.0.1
17         No Label   172.16.2.0/24[V] 0             Gi3        10.10.0.1
18         Pop Label  10.0.0.6/31      0             Gi1        10.0.0.1
           Pop Label  10.0.0.6/31      0             Gi2        10.0.0.3
19         Pop Label  10.0.0.8/31      0             Gi2        10.0.0.3
           Pop Label  10.0.0.8/31      0             Gi4        10.0.0.5
20         Pop Label  10.0.0.10/31     0             Gi1        10.0.0.1
           Pop Label  10.0.0.10/31     0             Gi4        10.0.0.5
21         Pop Label  10.0.0.12/31     0             Gi4        10.0.0.5
22         Pop Label  10.0.0.14/31     0             Gi1        10.0.0.1
23         Pop Label  10.0.0.16/31     0             Gi1        10.0.0.1
24         Pop Label  10.0.0.18/31     0             Gi4        10.0.0.5
25         24009      10.0.0.20/31     0             Gi1        10.0.0.1
           22         10.0.0.20/31     0             Gi4        10.0.0.5
26         Pop Label  10.10.10.2/32    0             Gi2        10.0.0.3
```

```
27          24008      10.10.10.3/32    0              Gi1        10.0.0.1
            24         10.10.10.3/32    0              Gi4        10.0.0.5
28          24006      10.10.10.4/32    0              Gi1        10.0.0.1
            25         10.10.10.4/32    0              Gi4        10.0.0.5
29          Pop Label  10.10.10.5/32    0              Gi1        10.0.0.1
Local       Outgoing   Prefix           Bytes Label    Outgoing   Next Hop
Label       Label      or Tunnel Id     Switched       interface
30          Pop Label  10.10.10.6/32    0              Gi4        10.0.0.5
31     [T]  Pop Label  1/1[TE-Bind]     0              drop

[T]     Forwarding through a LSP tunnel.
        View additional labelling info with the 'detail' option
```

如您所知，此行为的原因在于前缀(10.10.10.1/32)属于PE1，并且路由器还为此连接的前缀分配了一个隐式null标签。这可以通过使用show mpls ldp bindings命令进行验证：

<#root>

PE1#

show run interface loopback 0

```
Building configuration...

Current configuration : 66 bytes
!
interface Loopback0
 ip address 10.10.10.1 255.255.255.255
end
```

PE1#

show mpls ldp bindings 10.10.10.1 32

```
  lib entry: 10.10.10.1/32, rev 24
```

local binding: label: imp-null

```
        remote binding: lsr: 10.10.10.6:0, label: 23
        remote binding: lsr: 10.10.10.5:0, label: 24001
        remote binding: lsr: 10.10.10.2:0, label: 24000
```

由于PE1是Cisco IOS XE路由器，使用命令show bgp vpnv4 unicast all或show bgp vpnv4 unicast rd <value> <destination IP>有助于识别和确认是否通过MP-BGP正确获取了目标前缀172.16.1.0/24。导出后，这些命令的输出会显示前缀：

<#root>

PE1#

show bgp vpnv4 unicast all

```
BGP table version is 61, local router ID is 10.10.10.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
```

```
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
              t secondary path, L long-lived-stale,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 65000:1 (default for vrf WEST)
 *>i  0.0.0.0          10.10.10.3               0    100      0 65001 i
 *bi                   10.10.10.4               0    100      0 65001 i
```

***> 172.16.1.0/24 10.10.0.1 0 0 65000 i**

```
        <<<<<
 *>   172.16.2.0/24    10.10.0.1                0             0 65000 i
 *>i  192.168.1.0      10.10.10.4               0    100      0 65001 i
 *>i  192.168.2.0      10.10.10.3               0    100      0 65001 i
 *>i  192.168.3.0      10.10.10.4               0    100      0 65001 i
Route Distinguisher: 65001:1
 *>i  0.0.0.0          10.10.10.4               0    100      0 65001 i
 *>i  192.168.1.0      10.10.10.4               0    100      0 65001 i
 *>i  192.168.3.0      10.10.10.4               0    100      0 65001 i
Route Distinguisher: 65001:2
     Network          Next Hop            Metric LocPrf Weight Path
 *>i  0.0.0.0          10.10.10.3               0    100      0 65001 i
 *>i  192.168.2.0      10.10.10.3               0    100      0 65001 i

PE1#
```

show bgp vpnv4 unicast rd 65000:1 172.16.1.10

```
BGP routing table entry for 65000:1:172.16.1.0/24, version 2
Paths: (1 available, best #1, table WEST)
  Additional-path-install
  Advertised to update-groups:
     6
  Refresh Epoch 2
  65000
```

10.10.0.1 (via vrf WEST) from 10.10.0.1 (172.16.2.10)

```
         <<<<<
      Origin IGP, metric 0, localpref 100, valid, external, best
```

Extended Community: RT:65000:1 , recursive-via-connected

```
         <<<<<
      mpls labels in/out 16/nolabel
      rx pathid: 0, tx pathid: 0x0
      Updated on Sep 15 2023 18:27:23 UTC
```

类似地，查看VRF上的BGP VPNv4前缀，即CE-WEST接收的前缀，使用命令show bgp vpnv4 unicast vrf <name> <prefix>时，输出显示一直传输到入口PE4的MP-BGP标签16以及配置了65000：1的RT导出：

<#root>

PE1#

show bgp vpnv4 unicast vrf WEST 172.16.1.10

```
BGP routing table entry for 65000:1:172.16.1.0/24, version 2
Paths: (1 available, best #1, table WEST)
  Additional-path-install
  Advertised to update-groups:
     6
  Refresh Epoch 2
  65000
    10.10.0.1 (via vrf WEST) from 10.10.0.1 (172.16.2.10)
      Origin IGP, metric 0, localpref 100, valid, external, best
```

Extended Community: RT:65000:1 , recursive-via-connected

```
         <<<<<
```

mpls labels in/out 16/nolabel

```
        <<<<<
      rx pathid: 0, tx pathid: 0x0
      Updated on Sep 15 2023 18:27:23 UTC

PE1#
```

show run vrf WEST

```
Building configuration...

Current configuration : 478 bytes
vrf definition WEST
 rd 65000:1
```

route-target export 65000:1

```
         <<<<<
 route-target import 65000:1
 route-target import 65001:1
 route-target import 65001:2
 !
 address-family ipv4
 exit-address-family
!
!
interface GigabitEthernet3
 vrf forwarding WEST
 ip address 10.10.0.2 255.255.255.252
 negotiation auto
 no mop enabled
 no mop sysid
!
router bgp 65500
 !
 address-family ipv4 vrf WEST
  neighbor 10.10.0.1 remote-as 65000
  neighbor 10.10.0.1 activate
 exit-address-family
!
end
```

在此PE上要检查的最后信息是到目标IP的VRF级别的RIB和CEF条目，与在PE4上看到的条目相比，RIB上没有前缀172.16.1.0/24的标签，原因是这是来自CE的路由，并且通过eBGP获知该路由并将其插入到VRF路由表中，然后将此前缀导出到VPNv4。这可以通过使用show ip route vrf <name> <prefix>和show ip cef vrf <name> <prefix>（如下所示）命令进行验证：

<#root>

```
PE1#

show ip route vrf WEST 172.16.1.10


Routing Table: WEST
Routing entry for 172.16.1.0/24
  Known via "bgp 65500", distance 20, metric 0
  Tag 65000, type external
  Last update from 10.10.0.1 1w0d ago
  Routing Descriptor Blocks:
  * 10.10.0.1, from 10.10.0.1, 1w0d ago, recursive-via-conn
      opaque_ptr 0x7F8B4E3E1D50
      Route metric is 0, traffic share count is 1
      AS Hops 1
      Route tag 65000
      MPLS label: none

PE1#

show ip cef vrf WEST 172.16.1.10


172.16.1.0/24
  nexthop 10.10.0.1 GigabitEthernet3
```

这时，已确认目的地前缀172.16.1.0/24是由流量CE (CE-EAST)的源正确获知的，它通过MP-BGP正确传播，并且来自PE和Ps环回的标签也是在标签交换机路径上获知的。但是，源/目标之间的可达性仍然不成功，而且还有最后一个路由器用于检验CE-WEST。要在此路由器中检查的第一件事是路由表，请记住，源IP前缀192.168.1.0/24应显示在此表中：

<#root>

```
CE-WEST#

show ip route 192.168.1.10

% Network not in table

CE-WEST#
```

“Network not in table”显然是个问题，BGP表也可以验证，但在查找前缀后同样不存在：

<#root>

CE-WEST#

show ip bgp

```
BGP table version is 41, local router ID is 172.16.2.10
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
              t secondary path,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 *>   172.16.1.0/24    0.0.0.0                  0         32768 i
 *>   172.16.2.0/24    0.0.0.0                  0         32768 i
CE-WEST#
```

后退一步，您可以验证此提供商边缘路由器(PE1)是否正在向eBGP邻居CE-WEST通告前缀，这可以使用命令show bgp vpnv4 unicast vrf <name> neighbors <neighbor IP> advertised-routes完成，如下所示：

<#root>

PE1#

show bgp vpnv4 unicast vrf WEST neighbors 10.10.0.1 advertised-routes

```
BGP table version is 61, local router ID is 10.10.10.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
              t secondary path, L long-lived-stale,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 65000:1 (default for vrf WEST)
 *>i  0.0.0.0          10.10.10.3               0    100      0 65001 i
```

***>i 192.168.1.0 10.10.10.4 0 100 0 65001 i**

```
         <<<<<
 *>i  192.168.2.0      10.10.10.3               0    100      0 65001 i
 *>i  192.168.3.0      10.10.10.4               0    100      0 65001 i

Total number of prefixes 4
```

根据之前步骤，可以确认PE1路由器正在向CE-WEST正确通告前缀，因此是时候查看CE端的BGP邻居了：

<#root>

CE-WEST#

show ip bgp neighbors

```
BGP neighbor is 10.10.0.2,  remote AS 65500, external link
  BGP version 4, remote router ID 10.10.10.1
  BGP state = Established, up for 1w4d
  Last read 00:00:40, last write 00:00:43, hold time is 180, keepalive interval is 60 seconds
  Neighbor sessions:
    1 active, is not multisession capable (disabled)
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Four-octets ASN Capability: advertised and received
    Address family IPv4 Unicast: advertised and received
    Enhanced Refresh Capability: advertised and received
    Multisession Capability:
    Stateful switchover support enabled: NO for session 1
  Message statistics:
    InQ depth is 0
    OutQ depth is 0

                         Sent       Rcvd
    Opens:                  1          1
    Notifications:          0          0
    Updates:                3         17
    Keepalives:         19021      18997
    Route Refresh:          2          0
    Total:              19029      19019
  Do log neighbor state changes (via global configuration)
  Default minimum time between advertisement runs is 30 seconds

 For address family: IPv4 Unicast
  Session: 10.10.0.2
  BGP table version 41, neighbor version 41/0
  Output queue size : 0
  Index 3, Advertise bit 0
  3 update-group member
  Inbound path policy configured
```

Route map for incoming advertisements is FILTER

```
         <<<<<
  Slow-peer detection is disabled
  Slow-peer split-update-group dynamic is disabled
                               Sent       Rcvd
  Prefix activity:               ----       ----
    Prefixes Current:               2          0
    Prefixes Total:                 4         23
    Implicit Withdraw:              2         13
    Explicit Withdraw:              0         10
    Used as bestpath:             n/a          0
    Used as multipath:            n/a          0
    Used as secondary:            n/a          0

                                   Outbound    Inbound
  Local Policy Denied Prefixes:    --------    -------
    route-map:                            0          4
    Bestpath from this peer:             18        n/a
    Total:                               18          4
  Number of NLRIs in the update sent: max 2, min 0
  Last detected as dynamic slow peer: never
  Dynamic slow peer recovered: never
  Refresh Epoch: 3
```

```
  Last Sent Refresh Start-of-rib: 4d23h
  Last Sent Refresh End-of-rib: 4d23h
  Refresh-Out took 0 seconds
  Last Received Refresh Start-of-rib: 4d23h
  Last Received Refresh End-of-rib: 4d23h
  Refresh-In took 0 seconds
                                    Sent       Rcvd
        Refresh activity:           ----       ----
          Refresh Start-of-RIB        1          2
          Refresh End-of-RIB          1          2

  Address tracking is enabled, the RIB does have a route to 10.10.0.2
  Route to peer address reachability Up: 1; Down: 0
    Last notification 1w5d
  Connections established 3; dropped 2
  Last reset 1w4d, due to Peer closed the session of session 1
  External BGP neighbor configured for connected checks (single-hop no-disable-connected-check)
  Interface associated: GigabitEthernet0/3 (peering address in same link)
  Transport(tcp) path-mtu-discovery is enabled
  Graceful-Restart is disabled
  SSO is disabled
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Connection is ECN Disabled, Mininum incoming TTL 0, Outgoing TTL 1
Local host: 10.10.0.1, Local port: 179
Foreign host: 10.10.0.2, Foreign port: 39410
Connection tableid (VRF): 0
Maximum output segment queue size: 50

Enqueued packets for retransmit: 0, input: 0  mis-ordered: 0 (0 bytes)

Event Timers (current time is 0x4D15FD56):
Timer          Starts    Wakeups            Next
Retrans         19027          1             0x0
TimeWait            0          0             0x0
AckHold         19012      18693             0x0
SendWnd             0          0             0x0
KeepAlive           0          0             0x0
GiveUp              0          0             0x0
PmtuAger            0          0             0x0
DeadWait            0          0             0x0
Linger              0          0             0x0
ProcessQ            0          0             0x0

iss: 1676751051  snduna: 1677112739  sndnxt: 1677112739
irs: 2109012892  rcvnxt: 2109374776

sndwnd:  16061  scale:      0  maxrcvwnd:  16384
rcvwnd:  15890  scale:      0  delrcvwnd:    494

SRTT: 1000 ms, RTTO: 1003 ms, RTV: 3 ms, KRTT: 0 ms
minRTT: 0 ms, maxRTT: 1000 ms, ACK hold: 200 ms
uptime: 1036662542 ms, Sent idletime: 40725 ms, Receive idletime: 40925 ms
Status Flags: passive open, gen tcbs
Option Flags: nagle, path mtu capable
IP Precedence value : 6

Datagrams (max data segment is 1460 bytes):
Rcvd: 37957 (out of order: 0), with data: 19014, total data bytes: 361883
Sent: 37971 (retransmit: 1, fastretransmit: 0, partialack: 0, Second Congestion: 0), with data: 19027,

 Packets received in fast path: 0, fast processed: 0, slow path: 0
 fast lock acquisition failures: 0, slow path: 0
```

```
TCP Semaphore      0x0F3194AC  FREE
```

前面的输出显示，有一个路由映射应用于名为“FILTER”的传入通告，查看路由映射配置后，它显示了一个指向192.168.0.0/16上带有permit语句的前缀列表的match子句，但这是不正确的，因为prefix-list仅允许该特定前缀，而不允许该范围内可以包含的所有前缀：

<#root>

```
CE-WEST#

show route-map FILTER

route-map FILTER, permit, sequence 10
  Match clauses:

ip address prefix-lists: FILTER

  Set clauses:
  Policy routing matches: 0 packets, 0 bytes

CE-WEST#

show ip prefix-list FILTER

ip prefix-list FILTER: 1 entries

seq 5 permit 192.168.0.0/16

         <<<<<

CE-WEST#

show run | i ip prefix-list

ip prefix-list FILTER seq 5 permit 192.168.0.0/16
```

通过对前缀列表配置进行细微更改，指向192.168.1.10的路由现在已安装到RIB中：

<#root>

```
CE-WEST#

show run | i ip prefix-list

ip prefix-list FILTER seq 5 permit 192.168.0.0/16 le 32

         <<<<<

CE-WEST#

show ip bgp

BGP table version is 44, local router ID is 172.16.2.10
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
```

```
              x best-external, a additional-path, c RIB-compressed,
              t secondary path,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 *>   172.16.1.0/24    0.0.0.0                  0         32768 i
 *>   172.16.2.0/24    0.0.0.0                  0         32768 i
```

***> 192.168.1.0 10.10.0.2 0 65500 65001 i**

```
         <<<<<
 *>   192.168.2.0      10.10.0.2                              0 65500 65001 i
 *>   192.168.3.0      10.10.0.2                              0 65500 65001 i

CE-WEST#
```

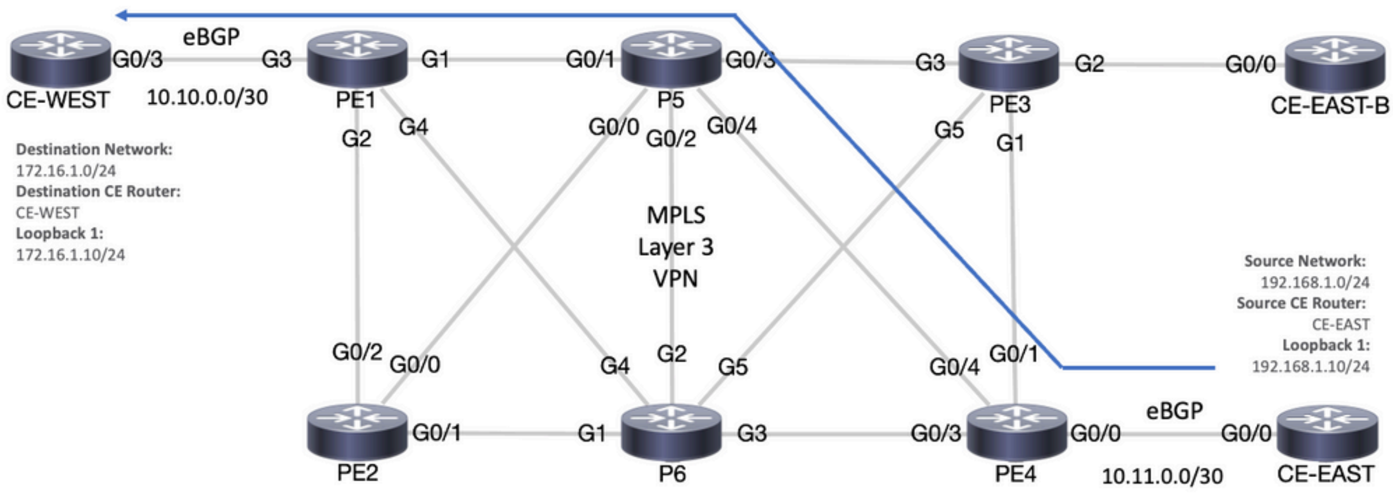
show ip route 192.168.1.10

Routing entry for 192.168.1.0/24

```
         <<<<<
  Known via "bgp 65000", distance 20, metric 0
  Tag 65500, type external
  Last update from 10.10.0.2 00:00:37 ago
  Routing Descriptor Blocks:
  * 10.10.0.2, from 10.10.0.2, 00:00:37 ago
      Route metric is 0, traffic share count is 1
      AS Hops 2
      Route tag 65500
      MPLS label: none
```

确认

现在，源和目标之间的可达性是成功的，并且可以确认traceroute通过与MPLS网络中跟踪的同一标签交换机路径：



转发路径

<#root>

CE-EAST#

ping 172.16.1.10 source loopback 1

```
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.16.1.10, timeout is 2 seconds:
Packet sent with a source address of 192.168.1.10
!!!!!
```

Success rate is 100 percent (5/5), round-trip min/avg/max = 7/7/9 ms

<<<<<

CE-EAST#

traceroute 172.16.1.10 source loop1 probe 1 numeric

```
Type escape sequence to abort.
Tracing the route to 172.16.1.10
VRF info: (vrf in name/id, vrf out name/id)
  1 10.11.0.2 2 msec
  2 10.0.0.16 [MPLS: Labels 24001/16 Exp 0] 9 msec
  3 10.10.0.2 [MPLS: Label 16 Exp 0] 8 msec
  4 10.10.0.1 9 msec
```

RP/0/0/CPU0:P5#

show ipv4 interface brief

```
Wed Sep 20 18:23:47.158 UTC

Interface                      IP-Address      Status          Protocol Vrf-Name
Loopback0                      10.10.10.5      Up              Up       default
MgmtEth0/0/CPU0/0              unassigned      Shutdown        Down     default
GigabitEthernet0/0/0/0         10.0.0.7        Up              Up       default
```

GigabitEthernet0/0/0/1 10.0.0.1 Up Up default

<<<<<

```
GigabitEthernet0/0/0/2         10.0.0.10       Up              Up       default
GigabitEthernet0/0/0/3         10.0.0.14       Up              Up       default
```

GigabitEthernet0/0/0/4 10.0.0.16 Up Up default

<<<<<

```
RP/0/0/CPU0:P5#
```

Cisco IOS XE验证命令

<#root>

MPLS/LDP

```
show mpls interfaces
show mpls forwarding-table
show mpls ldp bindings [destination prefix]
```

```
show mpls ldp neighbor [neighbor address]
clear mpls ldp neighbor [neighbor address|*]
```

RIB and CEF

```
show ip vrf [detail]
show run vrf
show ip route [destination prefix]
show ip route vrf <name> [destination prefix]
show ip cef vrf <name> [destination prefix]
show ip cef exact-route <source IP> <destination IP>
show ip cef vrf <name> exact-route <source IP> <destination IP>
```

BGP/VPNv4

```
show ip bgp [neighbors] <neighbor address>
show bgp vpnv4 unicast all [summary|destination prefix]
show bgp vpnv4 unicast all neighbor <neighbor address> advertised-routes
show bgp vpnv4 unicast vrf <name> neighbors <neighbor IP> advertised-routes
show bgp vpnv4 unicast vrf <name> <prefix>
show bgp vpnv4 unicast rd <value> <destination IP>
```

思科IOS XR验证命令

<#root>

MPLS/LDP

```
show mpls interfaces
show mpls forwarding
show mpls ldp bindings [destination prefix/mask]
show mpls ldp neighbor [neighbor address]
show mpls forwarding prefix [destination prefix/mask]
show mpls forwarding prefix [destination prefix/mask] detail hardware egress
clear mpls ldp neighbor [neighbor address]
```

RIB and CEF

```
show vrf [name|all]
show run vrf [name]
show route [destination prefix]
show route vrf <name> [destination prefix]
show cef vrf <name> [destination prefix]
show cef exact-route <source IP> <destination IP>
show cef vrf <name> exact-route <source IP> <destination IP>
```

BGP/VPNv4

```
show bgp vpnv4 unicast [summary|destination prefix/mask]
show bgp vpnv4 unicast neighbors <neighbor address> advertised-routes
show bgp vpnv4 unicast vrf <name> [prefix]
show bgp vrf <name> neighbors <neighbor IP> advertised-routes
show bgp vpnv4 unicast rd [value|all] [destination IP]
```

相关信息

- MPLS基本MPLS配置指南
- 配置基本 MPLS VPN 网络
- 如何排除 MPLS VPN 故障
- 验证网段路由SP之间的端到端连接

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。