带BGP信令的VPLS技术说明

目录

简介
问题
解决方案架构
PE路由器终端标识
VPLS身份和MPLS标签
封装信息
VPLS BGP自动发现前缀和VPLS BGP信令前缀
Cisco IOS软件配置示例
通告标签块
路由标识符和路由目标
PE1的配置示例
检查标签范围
检查标签
检查标签块
检查通告前缀
详细查看前缀
在BGP更新消息中通告、接收和处理标签块
PE2:接收BGP更新
PE2:查找标签
PE2:将前缀发送到PE1
PE1:查找标签
其他验证命令
由PE路由器为一个VFI通告的多个L2VPN VPLS前缀
PE1配置
PE2配置
初始标签块
PE1和PE2交换
PE1和PE2交换分析
前缀详细信息
互操作性

简介

本文档介绍基于边界网关协议(BGP)的自动发现，适用于具有BGP信令的虚拟专用局域网服务(VPLS)。自动发现是提供商边缘(PE)获取哪些远程PE是给定VPLS域的成员的方法。信令是PE获取给定VPLS域的给定远程PE期望的伪线标签的方法。

请参阅以下Internet工程任务组文档：

- RFC 4762使用标签分发协议(LDP)信令的虚拟专用LAN服务(VPLS)描述了使用VPLS的标签分发协议(LDP)信令（也称为Martini）的BGP自动发现。
- RFC 4761使用BGP进行自动发现和信令的虚拟专用局域网服务(VPLS)描述了BGP自动发现和VPLS的BGP信令（也称为Kompella）。

本文档重点介绍RFC 4761。使用RFC 4761,BGP更新的BGP网络层可达性信息(NLRI)保存自动发现和信令信息。当远程PE路由器收到此BGP更新时，它们拥有为VPLS建立全网状伪线所需的所有信息。BGP自动发现和BGP信令使用相同的BGP地址系列。

命令行界面(CLI)和输出来自Cisco IOS®软件。Cisco IOS-XR软件和Cisco NX-OS软件的配置和功能非常相似。

问题

VPLS由一组伪线(PW)组成，采用点对多点方式。到目前为止，LDP用于向PE路由器之间的伪线发出信号。因此，目标LDP会话发出信号，表示要使用哪个标签，一对PE路由器之间的伪线。您可以手动配置参与一个VPLS域的PE路由器集，也可以使用BGP自动发现配置。为了执行此自动发现，BGP通告哪个PE是哪个VPLS域的成员。但是，即使使用BGP自动发现，LDP也用于向多协议标签交换(MPLS)虚电路(VC)标签和伪线ID发出信号。

现在可以使用BGP来向PE路由器之间的伪线发出信号。

当一对路由器之间要设置一条伪线时，其他路由器不需要与此伪线相关的信息。例如，此类信息是要使用的VC标签。

由于LDP是用于设置伪线的信令协议，因此信息仅由这对路由器接收，因为LDP以点对点方式执行信令。

由于BGP作为用于设置伪线的信令协议，因此信息由所有其他路由器接收，因为内部BGP(iBGP)以点对多点方式执行信令。iBGP有全网状要求，因此一台路由器向所有其它iBGP路由器发送iBGP更新。这也可以通过路由反射器完成。

使用iBGP作为信令协议时，发送更新的方法有两种：

1. 每个PE路由器向每个PW的所有iBGP邻居通告一个BGP更新；每次都附加一个MPLS VC标签。因此，一个PE路由器发送的BGP更新数与PE路由器相同。但是，连接到BGP更新的VC标签只能由PE路由器中的一个使用 — PW另一端的PE路由器。
2. 为避免BGP更新数量过多的问题，设计了一种架构，其中一个本地PE路由器向所有远程PE路由器发送一组或一组本地VC标签。每个远程PE路由器选择其中一个VC标签，作为指向本地PE路由器的PW的远程VC标签。远程PE路由器必须以唯一方式选择远程VC标签，以便其他PE路由器不能从通告的标签块中选择相同的VC标签。由于发送了一组标签，因此必须有足够的标签可用来为所有可能设置的PW提供服务，但不应保留太多标签，以至于它们未使用和浪费。

本文档介绍如何使用BGP来发出伪线信号；请注意，BGP也同时用于自动发现。

解决方案架构

因为这是VPLS，所以核心层中仍然需要逐跳信令协议来将标记数据包从PE传输到PE路由器。核心中的此传输功能仍必须由LDP或MPLS流量工程实现。

BGP需要发送必要信息，以便以VPLS所需的点对多点方式设置伪线。此信令信息包括：

- PE路由器端点标识
- VPLS身份
- MPLS标签块
- 封装信息

PE路由器终端标识

PE路由器终端标识由作为更新BGP发送方的PE路由器确定。

与第2层虚拟专用网络(L2VPN)VPLS有关的BGP更新由AFI/SAFI 25/65标识。此地址系列在BGP发送OPEN消息时协商。

VPLS身份和MPLS标签

NLRI（也称为前缀）保存有关VPLS身份和MPLS标签块的信息。其编码的总长度为19字节：

```
 +-------------------------------------+
|  Length (2 octets)                   |
+-------------------------------------+
|  Route Distinguisher  (8 octets)     |
+-------------------------------------+
|  VE ID (2 octets)                    |
+-------------------------------------+
|  VE Block Offset (2 octets)          |
+-------------------------------------+
|  VE Block Size (2 octets)            |
+-------------------------------------+
|  Label Base (3 octets)               |
+-------------------------------------+
```

路由识别器(RD)与VPLS的身份相关。

> **注意：**在Cisco IOS和Cisco NX-OS软件实施中，所有PE路由器必须在同一VPLS域内具有相同的RD。

虚拟扩展(VE)ID、VE块偏移量、VE块大小和标签库(LB)与通告的标签块相关，如下一节所述。

封装信息

封装信息也附加到前缀，并作为BGP更新的扩展社区“Layer2 Info Extended Community”进行编码。该值为0x800A，编码为：

```
 +-------------------------------------+
| Extended community type (2 octets) |
+-------------------------------------+
```

```
     |  Encaps Type (1 octet)             |
     +------------------------------------+
     |  Control Flags (1 octet)           |
     +------------------------------------+
     |  Layer-2 MTU (2 octet)             |
     +------------------------------------+
     |  Reserved (2 octets)               |
     +------------------------------------+
```

VPLS的Encaps Type为19。

控制标志（位矢量）的编码方式如下：

```
       0 1 2 3 4 5 6 7
     +-+-+-+-+-+-+-+-+
     |   MBZ     |C|S|      (MBZ = MUST Be Zero)
     +-+-+-+-+-+-+-+-+
```

名称	价值	含义
C	1	当VPLS数据包发送到此PE时，必须存在控制字。
	0	当VPLS数据包发送到此PE时，控制字不得存在。
S	1	将VPLS数据包发送到此PE时，必须使用帧的顺序传送。
	0	当VPLS数据包发送到此PE时，不得使用帧的顺序传送。

还有路由目标(RT)附加到BGP更新。RT以与MPLS L3VPN相同的方式控制L2VPN的导入和导出。

VPLS BGP自动发现前缀和VPLS BGP信令前缀

VPLS BGP自动发现前缀是/96前缀，而VPLS BGP信令前缀是/136前缀。以下是各项的示例：

```
PE2#show bgp l2vpn vpls all
BGP table version is 264, local router ID is 10.100.1.2
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,
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: 1:100
 *>i 1:100:VEID-1001:Blk-150/136
                      10.100.1.1               0    100      0 ?
*>  1:100:10.100.1.2/96
                      0.0.0.0                            32768 ?

PE2#show bgp l2vpn vpls rd 1:100 ve-id 1001 block-offset 150
BGP routing table entry for 1:100:VEID-1001:Blk-150/136, version 262
Paths: (1 available, best #1, table L2VPN-VPLS-BGP-Table)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   10.100.1.1 (metric 21) from 10.100.1.4 (10.100.1.4)
     Origin incomplete, metric 0, localpref 100, valid, internal, best
     AGI version(0), VE Block Size(50) Label Base(10105)
     Extended Community: RT:1:100 RT:32:64 L2VPN L2:0x0:MTU-1500
     Originator: 10.100.1.1, Cluster list: 10.100.1.4
     rx pathid: 0, tx pathid: 0x0
```

```
PE2#show bgp l2vpn vpls rd 1:100 10.100.1.2
BGP routing table entry for 1:100:10.100.1.2/96, version 43
Paths: (1 available, best #1, table L2VPN-VPLS-BGP-Table)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   0.0.0.0 from 0.0.0.0 (10.100.1.2)
     Origin incomplete, localpref 100, weight 32768, valid, sourced, local,
best, AGI version(0)
     Extended Community: RT:1:100 L2VPN AGI:1:100
     rx pathid: 0, tx pathid: 0x0
```

Cisco IOS软件配置示例

以下是Cisco IOS软件配置示例：

```
!
l2vpn vfi context one
 vpn id 100
 autodiscovery bgp signaling bgp     <<< "signaling ldp" would be RFC 4762
 ve id 1001
 ve range 50
 route-target export 32:64
 route-target import 32:64

mpls label range 10000 20000

!
bridge-domain 1
 member Ethernet0/0 service-instance 100
 member vfi one

!
l2 router-id 10.100.1.1
!

interface Ethernet0/0
 no ip address
 service instance 100 ethernet
 !

!
router bgp 1
 bgp log-neighbor-changes
 neighbor 10.100.1.4 remote-as 1
 neighbor 10.100.1.4 update-source Loopback0
 !
 address-family l2vpn vpls
 neighbor 10.100.1.4 activate
 neighbor 10.100.1.4 send-community extended
 neighbor 10.100.1.4 suppress-signaling-protocol ldp
 exit-address-family
```

通告标签块

一个PE路由器必须通告至少一个标签块。标签块是一组连续的MPLS标签，由远程PE路由器使用，以选择一个远程VC标签。远程标签用于本地和远程PE路由器之间的PW。（PE路由器可以通告多个标签块，如后面各节所述。）

必须在每个PE上配置VE-ID。标识VPLS域内的PE路由器。

VE块大小(VBS)是标签块的大小，其默认值为10。如果配置了“ve range”，则是该值。“ve range”可配置为[11 -100]。

标签基础(LB)是可由PE路由器保留以用于此VPLS域的一组自由标签的第一个标签值。

VE块偏移(VBO)是当PE路由器必须创建多个标签块时要使用的偏移值。VBO使用以下公式计算：VBO = RND(VE-ID/VBS)* VBS

以下是示例计算：

- 如果VBS = 8,VE-ID = 2,VBO = RND(2/8)* 8 = 1
- 如果VBS = 8,VE-ID = 20，则VBO = RND(20/8)* 8 = 16
- 如果VBS = 50,VE-ID = 199,VBO = RND(199/50)* 50 = 150
- 如果VBS = 50,VE-ID = 1002，则VBO = RND(1002/50)* 50 = 1000

通告给远程PE路由器的标签块是{LB， LB + 1, ?,LB + VBS - 1}。标签块由LB和VBS定义；块从LB开始，以(LB + VBS - 1)结束。

每台PE路由器可根据需要创建多个标签块。路由器必须确保其是一组连续的空闲标签。

路由标识符和路由目标

PE1的配置示例

```
router bgp 1

l2vpn vfi context one
 vpn id 100
 autodiscovery bgp signaling bgp
 ve id 1001
 ve range 50
 route-target export 32:64
 route-target import 32:64

mpls label range 10000 20000
```

以下是配置值的说明：

- VPN ID配置为100。
- RD取自[ASN:vpn id]，除非RD已明确配置。这里，RD是1:100
- 导入/导出路由目标为32:64。
- LB来自范围[10000 20000]。LB的确切值取决于第一组自由连续本地标签，这些标签足够大，足以容纳VBS确定的所有标签。
- VE-ID配置为1001。
- VBS配置为50。
- VBO计算为：VBO = RND(VE-ID/VBS)* VBS或RND(1001/50)* 50 = 1000。

检查标签范围

您可以使用show mpls label range命令**检查标签范**围：

```
PE1#show mpls label range
Downstream Generic label region: Min/Max label: 10000/20000
```

有默认标签范围（按平台），您可以使用mpls label range命**令进行更**改。

检查标签

您可以使用show mpls forwarding-table命令检查标签转发信息库(LFIB)中一个标签块的实际**已使用的标**签。

```
PE1#show mpls forwarding-table
Local      Outgoing   Prefix           Bytes Label   Outgoing   Next Hop   Label
Label      or Tunnel Id     Switched      interface
10000      No Label   lbl-blk-id(1:0)  0             drop
10001      No Label   lbl-blk-id(1:1)  0             drop
10002      No Label   lbl-blk-id(1:2)  0             drop
?
10048      No Label   lbl-blk-id(1:48) 0             drop
10049      No Label   lbl-blk-id(1:49) 0             drop
10050      Pop Label  10.100.1.4/32    0             Et1/0      10.1.1.4
```

在本例中，本地路由器PE1为标签块保留了50个本地标签。“lbl-blk-id(1:0)”表示块ID为1，块实例为0，用于标识块的第一个标签。此块的最后一个标签是标签10049。

只要没有为该本地标签设置PW，LFIB中的“传出”接口就是“丢弃”。如果设置了PW，则“传出”接口为“none point2point”。

检查标签块

当配置“service internal”时，也可以使用**show mpls infrastructure lfd block-database summary命**令检查分配的标签块。

```
PE1#show mpls infrastructure lfd block-database summary
Block-DB entry for block-id : 0x1
Block-size : 50, App-Key type : AToM PWID, Labels : 10000 - 10049
```

LB是10000。在本例中，标签块是从LB到(LB + VBS - 1)或从10000到(10000 + 50 - 1)= 10049。

检查通告前缀

可以使用show bgp l2vpn vpls rd 1:100**命令检查通告的前**缀：

```
PE1#show bgp l2vpn vpls rd 1:100
BGP table version is 3, local router ID is 10.100.1.1
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,
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: 1:100
 *>  1:100:VEID-1001:Blk-1000/136
                      0.0.0.0                          32768 ?
```

详细查看前缀

要详细查看此前缀，请使用**show bgp l2vpn vpls rd 1:100 ve-id 1001 block-offset 1000** 命令。请注意，您指定VE-ID和标签块，可在NLRI(Blk-1000)中找到。

```
PE1#show bgp l2vpn vpls rd 1:100 ve-id 1001 block-offset 1000
BGP routing table entry for 1:100:VEID-1001:Blk-1000/136, version 3
Paths: (1 available, best #1, table L2VPN-VPLS-BGP-Table)
 Advertised to update-groups:
    1
 Refresh Epoch 1
 Local
   0.0.0.0 from 0.0.0.0 (10.100.1.1)
     Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
     AGI version(0), VE Block Size(50) Label Base(10000)
     Extended Community: RT:1:100 RT:32:64 L2VPN L2:0x0:MTU-1500
     rx pathid: 0, tx pathid: 0x0
```

NLRI显示RD为1:100、VE-ID为1001、VBO为1000、VBS为50和LB为10000。

第2层信息扩展社区包含以下信息：

- 封装类型为19(VPLS)
- 控制标志：C = 0（不能设置控制字）；S = 0（无帧的顺序传输）
- MTU为1500

RT扩展社区包含以下信息：

- RT 1:100
- RT 32:64

注意：默认VBS(10)很小，因此不会浪费本地标签。

在BGP更新消息中通告、接收和处理标签块

当本地PE路由器通告L2VPN VPLS前缀/标签块时，每个远程PE路由器必须尝试从该范围选择一个标签，以便用作远程VC标签。

- 如果远程PE路由器成功，它会使用该远程VC标签并在数据平面中对其进行编程。BGP不再发送信令。
- 如果远程PE路由器发生故障，它必须等待本地PE路由器通告另一个L2VPN VPLS前缀，然后尝试从该标签块中选择另一个远程VC标签。

假设PE1是具有上一配置的本地PE，而PE2是具有此配置的远程PE:

```
l2vpn vfi context one
 vpn id 100
```

```
 autodiscovery bgp signaling bgp
 ve id 1002
 ve range 50
!
mpls label range 3000 60000
```

PE2:接收BGP更新

PE2从PE1接收此BGP更新：

```
PE2#show bgp l2vpn vpls rd 1:100 ve-id 1001 block-offset 1000
BGP routing table entry for 1:100:VEID-1001:Blk-1000/136, version 5
Paths: (1 available, best #1, table L2VPN-VPLS-BGP-Table)
 Not advertised to any peer
 Refresh Epoch 2
 Local
   10.100.1.1 (metric 21) from 10.100.1.4 (10.100.1.4)
     Origin incomplete, metric 0, localpref 100, valid, internal, best
     AGI version(0), VE Block Size(50) Label Base(10000)
     Extended Community: RT:1:100 RT:32:64 L2VPN L2:0x0:MTU-1500
     Originator: 10.100.1.1, Cluster list: 10.100.1.4
     rx pathid: 0, tx pathid: 0x0
```

PE2:查找标签

PE2需要找到一个标签，它可以用作PW到PE1的远程VC标签。

PE2必须首先确定VBO是否在其配置范围内。PE2使用计算VBO <= VE-ID < VBO + VBS根据PE1通告的范围检查其VE-ID。在本例中，1000 <= 1002 < 1000 + 50，因此PE2成功。

然后，PE2需要选择远程VC标签。远程PE使用的解复用器(VC)标签计算为(LB+VE-ID-VBO)。

从早期的前缀中，LB为10000,VBO为1000。VE-ID是来自PE2的VE-ID，是1002。因此，PE2选择标签(LB + VE-ID - VBO)=(10000 + 1002 - 1000)= 10002。

使用**show l2vpn vfi name one**命令验证以下情况：

```
PE2#show l2vpn vfi name one
Legend: RT=Route-target, S=Split-horizon, Y=Yes, N=No

VFI name: one, state: up, type: multipoint, signaling: BGP
 VPN ID: 100, VE-ID: 1002, VE-SIZE: 50
 RD: 1:100, RT: 1:100
 Bridge-Domain 100 attachment circuits:
 Pseudo-port interface: pseudowire100001
 Interface          Peer Address    VE-ID  Local Label  Remote Label    S
 pseudowire100002   10.100.1.1      1001   3101         10002           Y
```

PE2:将前缀发送到PE1

然后，PE2将其前缀发送到PE1:

```
PE1#show bgp l2vpn vpls rd 1:100 ve-id 1002 block-offset 1000
```

```
BGP routing table entry for 1:100:VEID-1002:Blk-1000/136, version 4
Paths: (1 available, best #1, table L2VPN-VPLS-BGP-Table)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   10.100.1.2 (metric 21) from 10.100.1.4 (10.100.1.4)
     Origin incomplete, metric 0, localpref 100, valid, internal, best
     AGI version(0), VE Block Size(50) Label Base(3100)
     Extended Community: RT:1:100 L2VPN L2:0x0:MTU-1500
     Originator: 10.100.1.2, Cluster list: 10.100.1.4
     rx pathid: 0, tx pathid: 0x0
```

PE1:查找标签

PE1现在是远程PE，需要找到一个标签，它可以用作PW到PE2的远程VC标签。

PE1必须首先确定VBO是否在其配置范围内。PE1使用计算VBO <= VE-ID < VBO + VBS根据PE2通告的范围检查其VE-ID。在本例中，1000 <= 1001 < 1000 + 50，因此PE1成功。

然后，PE1需要选择远程VC标签。远程PE使用的解复用器(VC)标签计算为(LB+VE-ID-VBO)。

在前缀中，LB为3100,VBO为1000。VE-ID是来自PE1的VE-ID，是1001。因此，PE1选择标签(LB + VE-ID - VBO)=(3100 + 1001 - 1000)= 3101。

使用**show l2vpn vfi name one**命令验证以下情况：

```
PE1#show l2vpn vfi name one
 Legend: RT=Route-target, S=Split-horizon, Y=Yes, N=No

VFI name: one, state: up, type: multipoint, signaling: BGP
 VPN ID: 100, VE-ID: 1001, VE-SIZE: 50
 RD: 1:100, RT: 1:100, 32:64
 Bridge-Domain 1 attachment circuits:
 Pseudo-port interface: pseudowire100001
 Interface          Peer Address    VE-ID  Local Label  Remote Label    S
 pseudowire100002   10.100.1.2      1002   10002        3101            Y
```

其他验证命令

```
PE1#show mpls l2transport vc detail
Local interface: VFI one vfi up
 Interworking type is Ethernet
 Destination address: 10.100.1.2, VC ID: 100, VC status: up
   Output interface: Et1/0, imposed label stack {17 3101}
   Preferred path: not configured
   Default path: active
   Next hop: 10.1.1.4
 Create time: 02:06:08, last status change time: 02:06:08
   Last label FSM state change time: 02:06:08
 Signaling protocol: BGP
   Status TLV support (local/remote)   : Not Applicable
     LDP route watch                   : Not Applicable
     Label/status state machine        : established, LruRru
     Last local dataplane   status rcvd: No fault
     Last BFD dataplane     status rcvd: Not Applicable
     Last BFD peer monitor  status rcvd: Not Applicable
```

```
      Last local AC  circuit status rcvd: No fault
      Last local AC  circuit status sent: No fault
      Last local PW i/f circ status rcvd: No fault
      Last local LDP TLV     status sent: Not Applicable
      Last remote LDP TLV    status rcvd: Not Applicable
      Last remote LDP ADJ    status rcvd: Not Applicable
    MPLS VC labels: local 10002, remote 3101
    Group ID: local 0, remote 0
    MTU: local 1500, remote 1500
  Control Word: Off
  Dataplane:
    SSM segment/switch IDs: 8195/4097 (used), PWID: 3
  VC statistics:
    transit packet totals: receive 0, send 0
    transit byte totals:   receive 0, send 0
    transit packet drops:  receive 0, seq error 0, send 0

PE1#show mpls infrastructure lfd block-database id 1
Block-DB entry for block-id : 0x1
Block-size : 50, App-Key type : AToM PWID
App-Key entries:
 l2ckt(1)  10000
 l2ckt(2)  10001
 l2ckt(3)  10002
 l2ckt(4)  10003
 l2ckt(5)  10004
 l2ckt(6)  10005
 l2ckt(7)  10006
 l2ckt(8)  10007
 l2ckt(9)  10008
 l2ckt(10)  10009
 l2ckt(11)  10010
 l2ckt(12)  10011
 l2ckt(13)  10012
 l2ckt(14)  10013
 l2ckt(15)  10014
 l2ckt(16)  10015
 l2ckt(17)  10016
 l2ckt(18)  10017
 l2ckt(19)  10018
 l2ckt(20)  10019
 l2ckt(21)  10020
 l2ckt(22)  10021
 l2ckt(23)  10022
 l2ckt(24)  10023
 l2ckt(25)  10024
 l2ckt(26)  10025
 l2ckt(27)  10026
 l2ckt(28)  10027
 l2ckt(29)  10028
 l2ckt(30)  10029
 l2ckt(31)  10030
 l2ckt(32)  10031
 l2ckt(33)  10032
 l2ckt(34)  10033
 l2ckt(35)  10034
 l2ckt(36)  10035
 l2ckt(37)  10036
 l2ckt(38)  10037
 l2ckt(39)  10038
 l2ckt(40)  10039
 l2ckt(41)  10040
 l2ckt(42)  10041
```

```
 l2ckt(43)  10042
 l2ckt(44)  10043
 l2ckt(45)  10044
 l2ckt(46)  10045
 l2ckt(47)  10046
 l2ckt(48)  10047
 l2ckt(49)  10048
 l2ckt(50)  10049

PE1#show l2vpn atom vc destination 10.100.1.2

                                      Service
Interface Dest Address    VC ID      Type   Name                      Status
--------- --------------- ---------- ------ ------------------------- ----------
pw100002  10.100.1.2      100        vfi    one                       UP

PE1#show l2vpn atom vc destination 10.100.1.2 detail
pseudowire100002 is up, VC status is up PW type: Ethernet
 Create time: 02:11:13, last status change time: 02:11:13
   Last label FSM state change time: 02:11:13
 Destination address: 10.100.1.2 VC ID: 100
   Output interface: Et1/0, imposed label stack {17 3101}
   Preferred path: not configured
   Default path: active
   Next hop: 10.1.1.4
 Member of vfi service one
   Bridge-Domain id: 1
   Service id: 0xe7000001
 Signaling protocol: BGP
   Local VE ID: 1001, Remote VE ID: 1002
   Status TLV support (local/remote)         : Not Applicable
     LDP route watch                         : Not Applicable
     Label/status state machine              : established, LruRru
     Local dataplane status received         : No fault
     BFD dataplane status received           : Not Applicable
     BFD peer monitor status received        : Not Applicable
     Status received from access circuit     : No fault
     Status sent to access circuit           : No fault
     Status received from pseudowire i/f     : No fault
     Status sent to network peer             : Not Applicable
     Status received from network peer       : Not Applicable
     Adjacency status of remote peer         : Not Applicable
 Bindings
   Parameter    Local                          Remote
   ------------ ------------------------------ ------------------------------
   Label        10002                          3101
   Group ID     0                              0
   Interface
   MTU          1500                           1500
   Control word off                            off
   PW type      Ethernet                       Ethernet
   VCCV CV type 0x32                           0x32
                  LSPV [2], BFD/Raw [5]          LSPV [2], BFD/Raw [5]
                  BFD/Raw + sig [6]              BFD/Raw + sig [6]
   VCCV CC type 0x07                           0x07
                  CW [1], RA [2], TTL [3]       CW [1], RA [2], TTL [3]
   Status TLV   disabled                       N/A
 Dataplane:
   SSM segment/switch IDs: 8195/4097 (used), PWID: 3
 Rx Counters
   0 input transit packets, 0 bytes
   0 drops, 0 seq err
 Tx Counters
```

```
   0 output transit packets, 0 bytes
   0 drops

PE1#show l2vpn signaling rib rd 1:100
+- Origin of entry                                (i=iBGP/e=eBGP)
| +- Provisioned                                  (Yes/No)?
| | +- Stale entry                                (Yes/No)?
| | |
v v v
O P S     RD             VE-ID    VBO     VBS       LB       Next-Hop
-+-+-+-----------------+-------+-------+-------+---------+-----------------+
i Y N  1:100             1002    1000    50      3100       10.100.1.2

PE1#show l2vpn signaling rib rd 1:100 detail
Route 1:100:1002 (epoch:0) from iBGP peer 10.100.1.2
Provisioned (Y) Stale (N)
 Route-Target: 1:100
NLRI [FF000001]
   VE-ID:1002 VBO:1000 VBS:50 LB:3100
   MTU: 1500 Control Word: off
RIB Filter [27000002]
   RD: 1:100
   VE-ID: 1001, VBO: 1000, VBS: 50 LB: 10000
Forwarder [58000001] VFI one

PE1#show l2vpn atom pwid
AToM Pseudowire IDs: In use: 50, In holddown: 0

Label  Peer-Address    VCID       PWID       In-Use FirstUse ResuedAt FreedAt
------ --------------- ---------- ---------- ------ -------- -------- --------
10000  0.0.0.0         0          1          Yes    00:00:15 Never    Never
10001  0.0.0.0         0          2          Yes    00:00:15 Never    Never
10002  10.100.1.2      100        3          Yes    00:00:15 Never    Never
10003  0.0.0.0         0          4          Yes    00:00:15 Never    Never
10004  0.0.0.0         0          5          Yes    00:00:15 Never    Never

PE1#show l2vpn atom summary
Destination address: 10.100.1.2, total number of vc: 1
 0 unknown, 1 up, 0 down, 0 admin down, 0 recovering, 0 standby, 0 hotstandby
 1 active vc on MPLS interface Et1/0
```

由PE路由器为一个VFI通告的多个L2VPN VPLS前缀

一个PE可能需要为一个虚拟转发实例(VFI)通告多个标签块。

如果远程PE的VE-ID不在本地PE通告的范围内，则远程PE无法为PW选择远程标签。此计算（如前所述）是VBO <= VE-ID < VBO + VBS。

如果此检查失败，则远程PE的VE-ID超出范围。远程PE忽略从本地PE接收的前缀。本地PE在接收远程PE通告的前缀时获悉远程PE超出范围。本地PE需要确定该远程PE路由器使用的远程标签。本地PE还向远程PE发送新的第二前缀，用于新的本地标签块，远程PE应该能够使用该前缀来选择远程标签。

PE1配置

上一个示例在此处继续；PE1仍具有：

```
l2vpn vfi context one
 vpn id 100
 autodiscovery bgp signaling bgp
 ve id 1001
 ve range 50
 route-target export 32:64
 route-target import 32:64
!
mpls label range 10000 20000
```

PE2配置

PE2的VE-ID现在为1002，此配置：

```
l2vpn vfi context one
 vpn id 100
 autodiscovery bgp signaling bgp
 ve id 10002
 ve range 50
!
mpls label range 3000 60000
```

初始标签块

PE1和PE2都以这些初始标签块开头。

```
PE1#show bgp l2vpn vpls rd 1:100
BGP table version is 2, local router ID is 10.100.1.1
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,
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: 1:100
 *>  1:100:VEID-1001:Blk-1000/136
                      0.0.0.0                            32768 ?

PE2#show bgp l2vpn vpls rd 1:100
BGP table version is 3, local router ID is 10.100.1.2
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,
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: 1:100
 *>  1:100:VEID-10002:Blk-10000/136
                      0.0.0.0                            32768 ?
```

PE1和PE2交换

使用**debug bgp l2vpn vpls updates**命令查看PE1和PE2交换，然后使用**show bgp l2vpn vpls rd**

1:100命令查看详细信息。

```
PE1#
%BGP-5-ADJCHANGE: neighbor 10.100.1.4 Up
BGP(9): update formatted for 1:100:VEID-1001:Blk-1000:VBS-50:LB-10000/136 VE ID
1001 VE Block Offset 1000 VE Block Size 50 Label Base 10000 /136
BGP(9): (base) 10.100.1.4 send UPDATE (format) 1:100:VEID-1001:Blk-1000:VBS-50:
LB-10000/136, next 10.100.1.1, metric 0, path Local, extended community RT:1:100
RT:32:64 L2VPN L2:0x0:MTU-1500
BGP(9): 10.100.1.4 rcvd UPDATE w/ attr: nexthop 10.100.1.2, origin ?,
localpref 100, metric 0, originator 10.100.1.2, clusterlist 10.100.1.4, extended
community RT:1:100 L2VPN L2:0x0:MTU-1500
BGP(9): 10.100.1.4 rcvd 1:100:VEID-10002:Blk-10000:VBS-50:LB-3000/136
BGP(9): bump net 1:100:VEID-10002:Blk-10000:VBS-50:LB-3000/136, non bpath  added
BGP(9): nettable_walker called for 1:100:VEID-10002:Blk-10000:VBS-50:LB-3000/136
BGP(9): best path[0] 1:100:VEID-10002:Blk-10000:VBS-50:LB-3000/136 source
10.100.1.1 nh 10.100.1.2 vpls-id: L2VPN L2:0x0:MTU-1500
BGP(9): add XC RIB route 1:100:VEID-10002:Blk-10000:VBS-50:LB-3000/136 masklen 136
L2VPN L2:0x0:MTU-1500 pathcount: 1 [0] LDP source:10.100.1.1 nexthop:10.100.1.2
RT:1:100
BGP(9): bump net 1:100:VEID-1001:Blk-10000:VBS-50:LB-10053/136, non bpath  added
BGP(9): nlri update add VBS 50 LB 10053
BGP(9): nlri update add export extcomm count 4
BGPSSA ssacount is 0
BGP(9): update formatted for 1:100:VEID-10002:Blk-10000:VBS-50:LB-3000/136 VE ID
10002 VE Block Offset 10000 VE Block Size 50 Label Base 3000 /136
BGP(9): nettable_walker called for 1:100:VEID-1001:Blk-10000:VBS-50:LB-10053/136
BGP(9): nettable_walker 1:100:VEID-1001:Blk-10000:VBS-50:LB-10053/136 route sourced
locally
BGP(9): update formatted for 1:100:VEID-1001:Blk-10000:VBS-50:LB-10053/136 VE ID
1001 VE Block Offset 10000 VE Block Size 50 Label Base 10053 /136
BGP(9): (base) 10.100.1.4 send UPDATE (format) 1:100:VEID-1001:Blk-10000:VBS-50:
LB-10053/136, next 10.100.1.1, metric 0, path Local, extended community RT:1:100
RT:32:64 L2VPN L2:0x0:MTU-1500 L2VPN L2:0x0:MTU-1500
BGP(9): 10.100.1.4 rcvd UPDATE w/ attr: nexthop 10.100.1.2, origin ?, localpref 100,
metric 0, originator 10.100.1.2, clusterlist 10.100.1.4, extended community
RT:1:100 L2VPN L2:0x0:MTU-1500
BGP(9): 10.100.1.4 rcvd 1:100:VEID-10002:Blk-1000:VBS-50:LB-3053/136
BGP(9): bump net 1:100:VEID-10002:Blk-1000:VBS-50:LB-3053/136, non bpath  added
BGP(9): nettable_walker called for 1:100:VEID-10002:Blk-1000:VBS-50:LB-3053/136
BGP(9): best path[0] 1:100:VEID-10002:Blk-1000:VBS-50:LB-3053/136 source 10.100.1.1
nh 10.100.1.2 vpls-id: L2VPN L2:0x0:MTU-1500
BGP(9): add XC RIB route 1:100:VEID-10002:Blk-1000:VBS-50:LB-3053/136 masklen 136
L2VPN L2:0x0:MTU-1500 pathcount: 1 [0] LDP source:10.100.1.1 nexthop:10.100.1.2
RT:1:100
BGP(9): update formatted for 1:100:VEID-10002:Blk-1000:VBS-50:LB-3053/136 VE ID
10002 VE Block Offset 1000 VE Block Size 50 Label Base 3053 /136
BGPSSA ssacount is 0

PE1#show bgp l2vpn vpls rd 1:100
BGP table version is 5, local router ID is 10.100.1.1
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,
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: 1:100
 *>  1:100:VEID-1001:Blk-1000/136
                      0.0.0.0                            32768 ?
 *>  1:100:VEID-1001:Blk-10000/136
```

```
                      0.0.0.0                          32768 ?
 *>i 1:100:VEID-10002:Blk-1000/136
                      10.100.1.2               0    100      0 ?
 *>i 1:100:VEID-10002:Blk-10000/136
                      10.100.1.2               0    100      0 ?

PE2#show bgp l2vpn vpls rd 1:100
BGP table version is 6, local router ID is 10.100.1.2
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,
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: 1:100
 *>i 1:100:VEID-1001:Blk-1000/136
                      10.100.1.1               0    100      0 ?
 *>i 1:100:VEID-1001:Blk-10000/136
                      10.100.1.1               0    100      0 ?
 *>  1:100:VEID-10002:Blk-1000/136
                      0.0.0.0                          32768 ?
 *>  1:100:VEID-10002:Blk-10000/136
                      0.0.0.0                          32768 ?
```

PE1和PE2交换分析

PE1和PE2现在已相互通告两个标签块。

PE1首先向PE2通告初始BGP更新：

```
BGP(9): update formatted for 1:100:VEID-1001:Blk-1000:VBS-50:LB-10000/136 VE ID
1001 VE Block Offset 1000 VE Block Size 50 Label Base 10000 /136
BGP(9): (base) 10.100.1.4 send UPDATE (format) 1:100:VEID-1001:Blk-1000:VBS-50:
LB-10000/136, next 10.100.1.1, metric 0, path Local, extended community
RT:1:100 RT:32:64 L2VPN L2:0x0:MTU-1500
```

此更新根据PE1上的配置设置了NLRI。

然后，PE1从PE2接收初始BGP更新。

```
BGP(9): 10.100.1.4 rcvd UPDATE w/ attr: nexthop 10.100.1.2, origin ?, localpref
100, metric 0, originator 10.100.1.2, clusterlist 10.100.1.4, extended
community RT:1:100 L2VPN L2:0x0:MTU-1500
BGP(9): 10.100.1.4 rcvd 1:100:VEID-10002:Blk-10000:VBS-50:LB-3000/136
```

PE2通告初始前缀，其值为VE-ID 10002,VBO = 10000,VBS = 50,LB = 3000。

PE1发现PE2超出范围，因为PE1从标签块LB开始到(LB + VBS - 1)或从10000到(10000 + 50 - 1)= 10049。

PE1必须确定VBO是否在其配置范围内。因此，需要根据PE1通告的范围检查PE2的VE-ID。计算为VBO <= VE-ID < VBO + VBS。在本例中，1000 <= 10002 < 1000 + 50，不正确。因此，PE1需要发送新的标签块以适应PE2的超出范围的VE-ID。为了响应来自PE2的初始更新，PE1格式化并向PE2发送新的附加BGP更新。PE1现在使用新的VBO 10000。

```
BGP(9): update formatted for 1:100:VEID-1001:Blk-10000:VBS-50:LB-10053/136
```

```
VE ID 1001 VE Block Offset 10000 VE Block Size 50 Label Base 10053 /136
BGP(9): (base) 10.100.1.4 send UPDATE (format) 1:100:VEID-1001:Blk-10000:
VBS-50:LB-10053/136, next 10.100.1.1, metric 0, path Local, extended
community RT:1:100 RT:32:64 L2VPN L2:0x0:MTU-1500 L2VPN L2:0x0:MTU-1500
```

对于PE1,VBO为10000,VBS为50,LB为10053。PE2的检查是VBO <= VE-ID < VBO + VBS。在本例中，为10000 <= 10002 < 10000 + 50，这是正确的。PE2可以从PE1的此新标签块[10053 - 10102]中选择远程标签。换句话说，PE1添加了新标签块以容纳PE2并发送了两条BGP更新消息。

相反的方向也发生了同样的情况。PE2从PE1接收初始BGP更新。此更新的值为VE-ID 1001,VBO = 1000,VBS = 50,LB = 10000。

PE2发现PE1的VE-ID与PE2的初始更新超出范围。PE1的检查是VBO <= VE-ID < VBO + VBS或10000 <= 1001 < 10000 + 50。作为响应，PE2发送此第二个BGP更新，其中新标签块[3053 - 3102]包含VE-ID 1PE1的001，因为PE1的检查是VBO <= VE-ID < VBO + VBS或1000 <= 1001 < 1000 + 50。

```
BGP(9): 10.100.1.4 rcvd UPDATE w/ attr: nexthop 10.100.1.2, origin ?,
localpref 100, metric 0, originator 10.100.1.2, clusterlist 10.100.1.4,
extended community RT:1:100 L2VPN L2:0x0:MTU-1500
BGP(9): 10.100.1.4 rcvd 1:100:VEID-10002:Blk-1000:VBS-50:LB-3053/136
```

前缀详细信息

以下是PE1产生的两个前缀的详细信息：

```
PE1#show bgp l2vpn vpls rd 1:100 ve-id 1001 block-offset 1000
BGP routing table entry for 1:100:VEID-1001:Blk-1000/136, version 2
Paths: (1 available, best #1, table L2VPN-VPLS-BGP-Table)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   0.0.0.0 from 0.0.0.0 (10.100.1.1)
     Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
     AGI version(0), VE Block Size(50) Label Base(10000)
     Extended Community: RT:1:100 RT:32:64 L2VPN L2:0x0:MTU-1500
     rx pathid: 0, tx pathid: 0x0

PE1#show bgp l2vpn vpls rd 1:100 ve-id 1001 block-offset 10000
BGP routing table entry for 1:100:VEID-1001:Blk-10000/136, version 4
Paths: (1 available, best #1, table L2VPN-VPLS-BGP-Table)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   0.0.0.0 from 0.0.0.0 (10.100.1.1)
     Origin incomplete, localpref 100, weight 32768, valid, sourced, local, best
     AGI version(0), VE Block Size(50) Label Base(10053)
     Extended Community: RT:1:100 RT:32:64 L2VPN L2:0x0:MTU-1500
       L2VPN L2:0x0:MTU-1500
     rx pathid: 0, tx pathid: 0x0
```

此处，两个PE路由器具有不连续的编号方案，这会导致每个PE发送两个BGP更新。如果有许多PE路由器采用不连续的编号方案，BGP更新的数量会迅速增长。

www.cisco.com说："例如，VE-ID编号序列（如1、2、3或501、502、503）良好，因为VE-ID是连续的。100、200、300等编号方案不良，因为它是不连续的。"

1、2、3或501、502、503的第一个示例是连续的数字，因此每个PE路由器只需发送一个L2VPN VPLS前缀。在第三个示例(100、200、300)中，每个PE必须发送许多L2VPN VPLS前缀。对于非连续数字，足够大的VE范围会使要通告的前缀数量保持较低。但是，保留（浪费）标签的数量仍然较大。

互操作性

如果BGP路由反射器(RR)运行的软件不理解RFC 4761，但支持RFC 4762，则RR上需要特殊的BGP neighbor x.x.x.x prefix-length-size 2 配置命令，以便它能够反映用于RFC 47的RFC更新61.

前缀的发送长度通常为1字节。Cisco IOS软件实施了草案“draft-ietf-l2vpn-signaling-08”，后来成为RFC 6074。当时选择了1个字节的长度字段，表示以位为单位的长度。

RFC 6074第2层虚拟专用网络(L2VPN)中的调配、自动发现和信令（RFC 6074 Provisioning、Auto-Discovery和Signaling）指定BGP自动发现的NLRI编码应为2个字节。2个字节表示可变长度前缀中跟随的前缀字节数。

RFC 6074第7部分“BGP-AD和VPLS-BGP互操作性”中的说明：

“BGP-AD和VPLS-BGP [RFC4761]都使用相同的AFI/SAFI。要使BGP-AD和VPLS-BGP共存，必须将NLRI长度用作解复用器。

BGP-AD NLRI的NLRI长度为12字节，仅包含8字节RD和4字节VSI-ID。VPLS-BGP [RFC4761]使用17字节NLRI长度。因此，BGP-AD的实施必须忽略大于12字节的NLRI。”

如果RR**上不存在neighbor x.x.x.x prefix-length-size 2**命令，则BGP邻居不会出现，并且RR将长度字段解释为仅1字节。此通知显示在RR:

```
%BGP-3-NOTIFICATION: sent to neighbor 10.100.1.2 3/10 (illegal network) 1 bytes FF
%BGP-4-MSGDUMP: unsupported or mal-formatted message received from 10.100.1.2:
FFFF FFFF FFFF FFFF FFFF FFFF FFFF FFFF 005E 0200 0000 4780 0E1C 0019 4104 0A64
0102 0000 1100 0000 0100 0000 6427 1227 1000 3200 BB80 4001 0102 4002 0080 0404
*Feb 15 12:14:11.561: %BGP_SESSION-5-ADJCHANGE: neighbor 10.100.1.2 L2VPN Vpls
topology base removed from session  BGP Notification sent
*Feb 15 12:14:11.561: %BGP_SESSION-5-ADJCHANGE: neighbor 10.100.1.2 IPv4 Unicast
topology base removed from session  BGP Notification sent
```

PE路由器上显示以下通知：

```
%BGP-3-NOTIFICATION: received from neighbor 10.100.1.4 3/10 (illegal network)
1 bytes FD
```

这是因为在Cisco IOS软件中BGP自动发现的原始实施中，长度字段为1字节。

如果在RR上**输入neighbor x.x.x.x prefix-length-size 2命**令，则不会显示通知。

```
router bgp 1
 neighbor 10.100.1.2 remote-as 1
 neighbor 10.100.1.2 update-source Loopback0
!
 address-family l2vpn vpls
 neighbor 10.100.1.2 activate
 neighbor 10.100.1.2 send-community extended
```

```
neighbor 10.100.1.2 prefix-length-size 2
neighbor 10.100.1.2 route-reflector-client
```