

# 了解MPLS L2VPN伪线

## 目录

---

[简介](#)

[背景信息](#)

[L2VPN 概述](#)

[为什么需要L2VPN](#)

[MPLSL2VPN型号](#)

[技术选项](#)

[1. VPWS 服务](#)

[2. VPLS 服务](#)

[3. EVPN](#)

[4. PBB-EVPN](#)

[VPWS 伪线参考模型](#)

[第2层VPN启用程序：伪线](#)

[AToM 架构](#)

[基于MPLS的L2传输](#)

[VPWS 流量封装](#)

[伪线信令](#)

[控制字](#)

[转发平面处理](#)

[操作](#)

[PW 状态信令](#)

[AToM 基本配置](#)

[伪线数据包分析](#)

[拓扑](#)

[L2VPN 交互操作](#)

[可能的交互操作](#)

[相关信息](#)

---

## 简介

本文档介绍了基于多协议标签交换 (MPLS) 的 L2 虚拟专用网络 (L2VPN) 伪线。

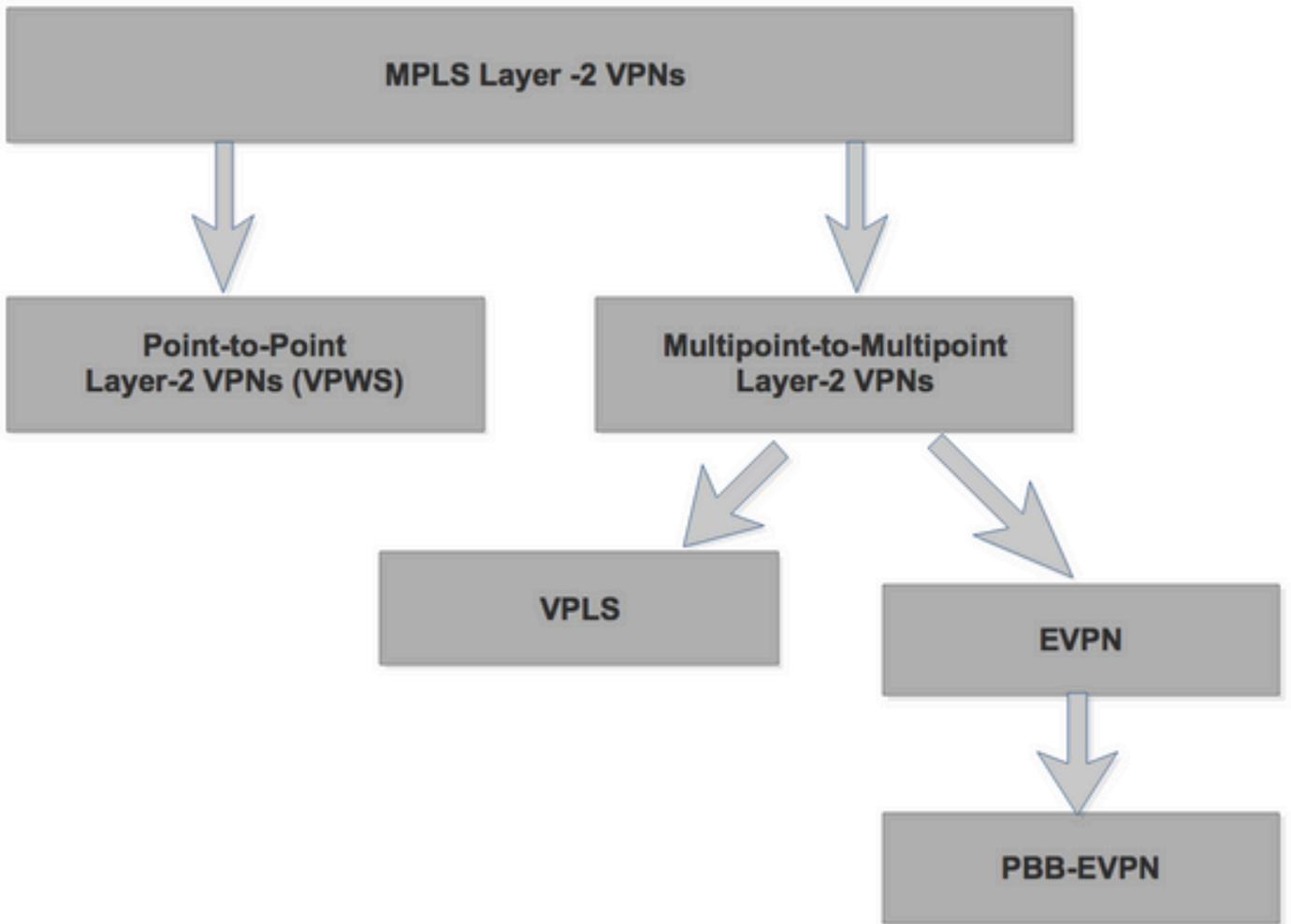
## 背景信息

本章介绍了Cisco IOS®、Cisco IOS® XE中伪线信令和数据包分析，以说明其行为。

## L2VPN 概述

对等连接电路(例如以太网到以太网、PPP到PPP、高级数据链路控制(HDLC)等)已经存在基于





### 1. VPWS 服务

- 点对点•称为“伪线”(PW)

### 2. VPLS 服务

- 多点

### 3. EVPN

- xEVPN 系列推出新一代以太网服务解决方案

a.用于以太网网段和MPLS核心上的MAC分配和学习的BGP控制平面

b. IP VPN的相同原理和操作经验

- 不使用伪线

a.将MP2P隧道用于单播

b.通过入口复制 ( 通过MP2P隧道 ) 或LSM的多目标帧传输

- 符合 IETF 标准化的多供应商解决方案

#### 4. PBB-EVPN

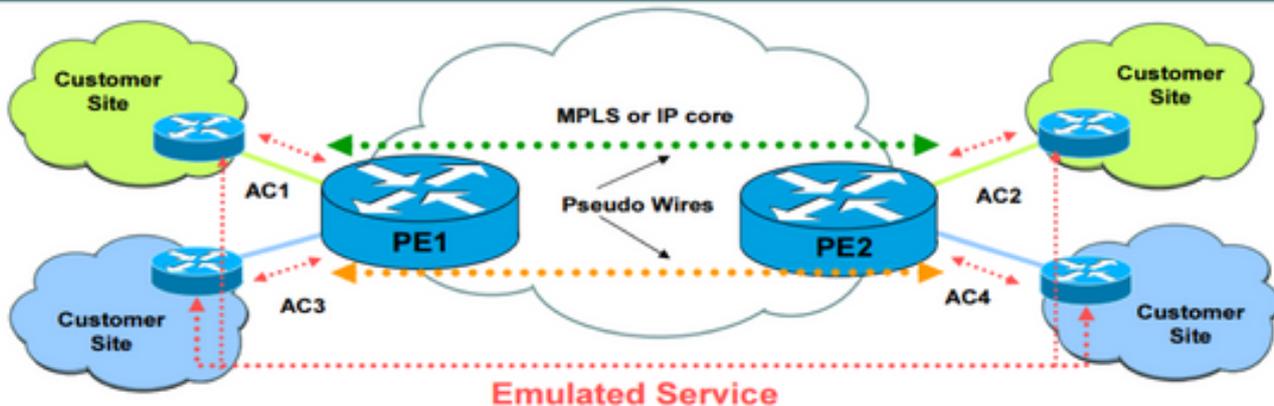
- 将 PBB 的扩展工具 ( 又名 MAC-in-MAC ) 与 EVPN 基于 BGP 的 MAC 学习相结合

EVPN 和提供商主干网桥 EVPN (PBB-EVPN) 是基于 BGP 控制平面在核心上进行 MAC 分发或学习的下一代 L2VPN 解决方案，旨在满足以下要求：

- 实现单流冗余和负载均衡
- 简化调配和操作
- 实现最优转发
- 快速收敛
- 扩展 MAC 地址

### VPWS 伪线参考模型

1. PW是两个PE设备之间的连接，连接两个传输L2帧的AC。
2. MPLS 上的任何传输 (AToM) 是思科针对 IP/MPLS 网络实施的 VPWS。
3. 接入电路 (AC) 是连接 CE 与 PE 的物理或虚拟电路，可以是 ATM、帧中继、HDLC、PPP 等。
4. Edge(CE)设备将PW视为非共享链路或电路。



### 第2层VPN启用程序：伪线

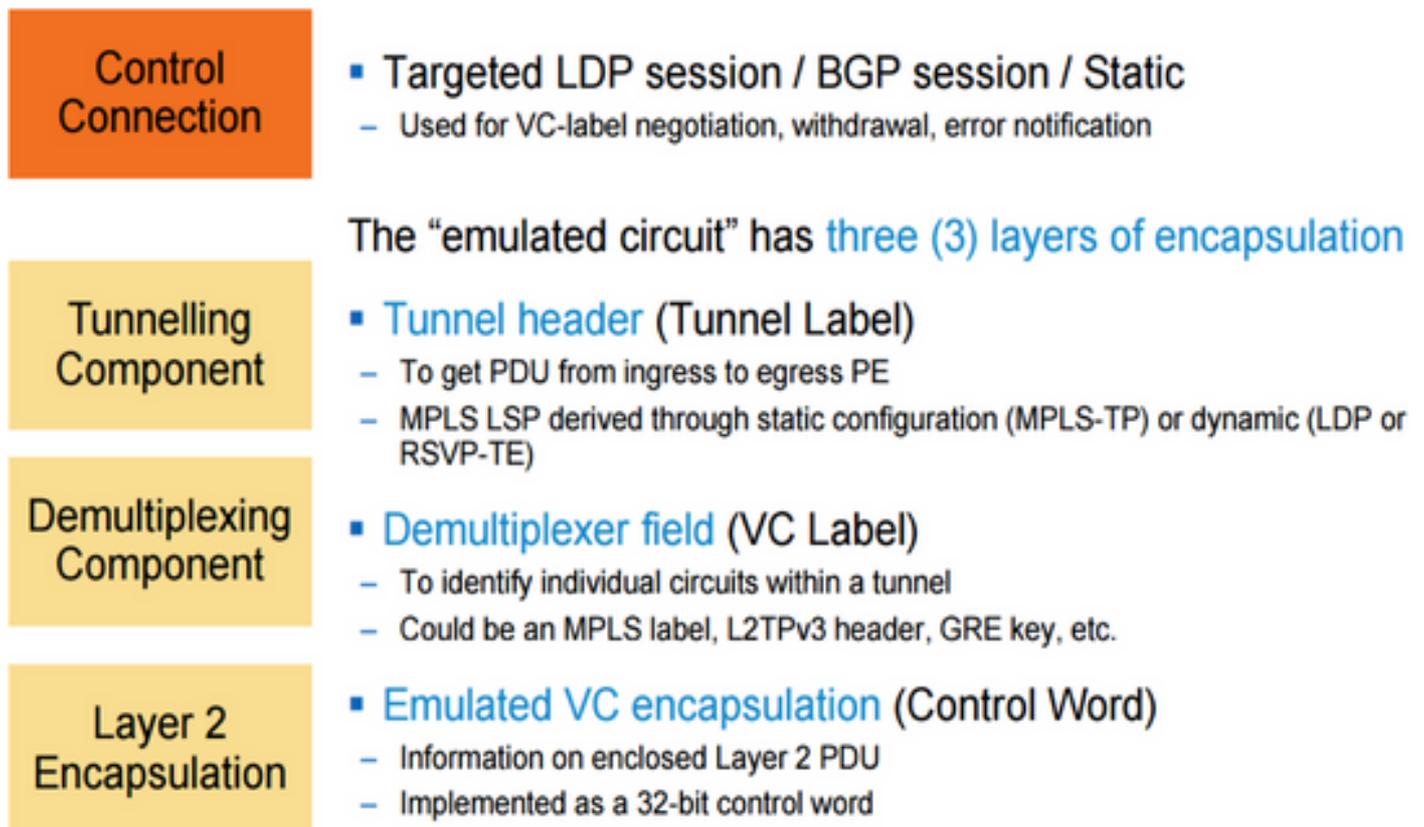
L2VPN采用伪线(PW)技术构建。

- PW提供通用的中间格式来通过数据包交换网络(PSN) ( 转发数据包的网络 ) ( IPv4、IPv6、MPLS、以太网 ) 传输多种类型的网络服务。
- PW技术提供同类传输和互通(IW)。
- AC 上的 PE 路由器接收的帧会经过封装，然后通过 PSW 发送至远程 PE 路由器。
- 出口 PE 路由器可接收来自伪线的数据包并删除其封装。
- 出口 PE 可提取帧并将其转发给 AC。

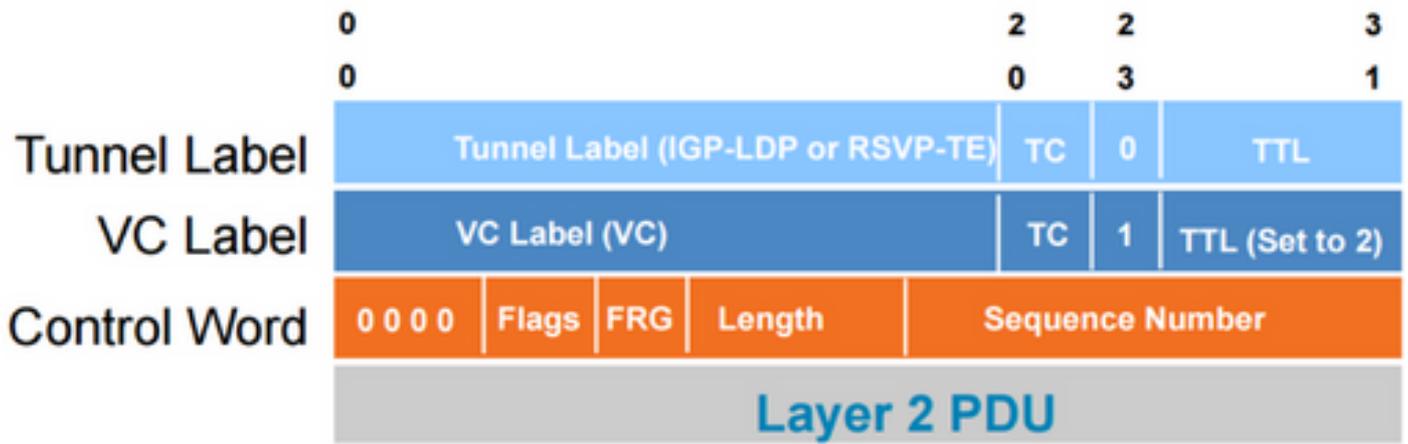
## AToM 架构

- 在 AToM 网络中，SP 的所有路由器均运行 MPLS，且 PE 路由器的 AC 可通往 CE 路由器。
- 就 AToM 而言，PSN 隧道只不过是两台 PE 路由器之间的标签交换路径 (LSP)。
- 因此，与该 LSP 关联的标签在 AToM 上下文中称为“隧道标签”。
- 首先，LDP 用于在 PE 之间逐跳传递信令。
- 其次，LSP 可以是 MPLS TE 隧道，RSVP 会使用 TE 所需的扩展传递信令。
- 使用此隧道标签，您可以确定帧所承载的 PSN 隧道属于哪个 PSN 隧道。
- 此隧道标签还可通过 MPLS 主干网将帧从本地或入口 PE 传送至远程或出口 PE。
- 要将多条伪线复用到一条 PSN 隧道中，PE 路由器需使用其他标签来标识伪线。
- 此标签因可标识多路复用帧所到的 VC 或 PW 而被称为 VC 或 PW 标签。

## 基于 MPLS 的 L2 传输



## VPWS 流量封装



1. 使用了三级封装。
2. 使用隧道标签在 PE 之间交换数据包。
3. VC 标签标识 PW。
4. 在 PE 之间传递 VC 标签的信令。
5. 可选控制字 (CW) 承载第 2 层控制位并启用排序。

Control Word	
Encap.	Required
ATM N:1 Cell Relay	No
ATM AAL5	Yes
Ethernet	No
Frame Relay	Yes
HDLC	No
PPP	No
SAToP	Yes
CESoPS N	Yes

## 伪线 伪线

- PE 路由器之间使用 TLDP 会话传递伪线中的信令。
- PE 路由器之间使用 T-LDP 会话，通告与 PSW 关联的 VC 标签。
- 此标签在使用下游主动标签通告模式的标签映射消息中通告。
- 出口 PE 通过 TLDP 会话向 AC 的入口 PE 通告 VC 标签。 # TLDP 通告的 VC 标签
- LDP向入口PE通告出口PE路由器的隧道标签。 # LDP通告的隧道标签

请注意，出口 PE 通告的是标签 3，这表明使用了 PHP。

TLDP 会话所通告的标签映射消息中包含以下 TLV：

# LDP Label Mapping message:

IP Header

TCP Header (Port 646)

LDP PDU

LDP Header

LDP Message: Label Mapping

FEC TLV

PW ID FEC Element 128: Interface Parameters

Generic Label TLV

数据包的控制字长度不为 0，则路由器便会知晓已添加填充内容，并在转发帧之前正确删除填充内容。

2. 保留传输帧的顺序：使用此序列号时，接收方可以检测数据包：

发送到 PW 的第一个数据包的序列号为 1，后续每个数据包序列号均会递增 1，直至达到 65535

如果接收器检测到失序序列，则会将其丢弃，不会对无序 AToM 数据包重新排序。

默认情况下，排序处于禁用状态。

3. 负载均衡：

路由器负责执行 MPLS 负载检查，并在此基础上决定如何实现流量负载均衡。

路由器会查看第一个半字节，如果第一个半字节为 4，则为 IPV4 数据包。通用控制字以值为 0 的半字节开头，而使用 OAM 数据的控制字则以值 1 开头。

4. 促进分段和重组：

可用于指示负载分段

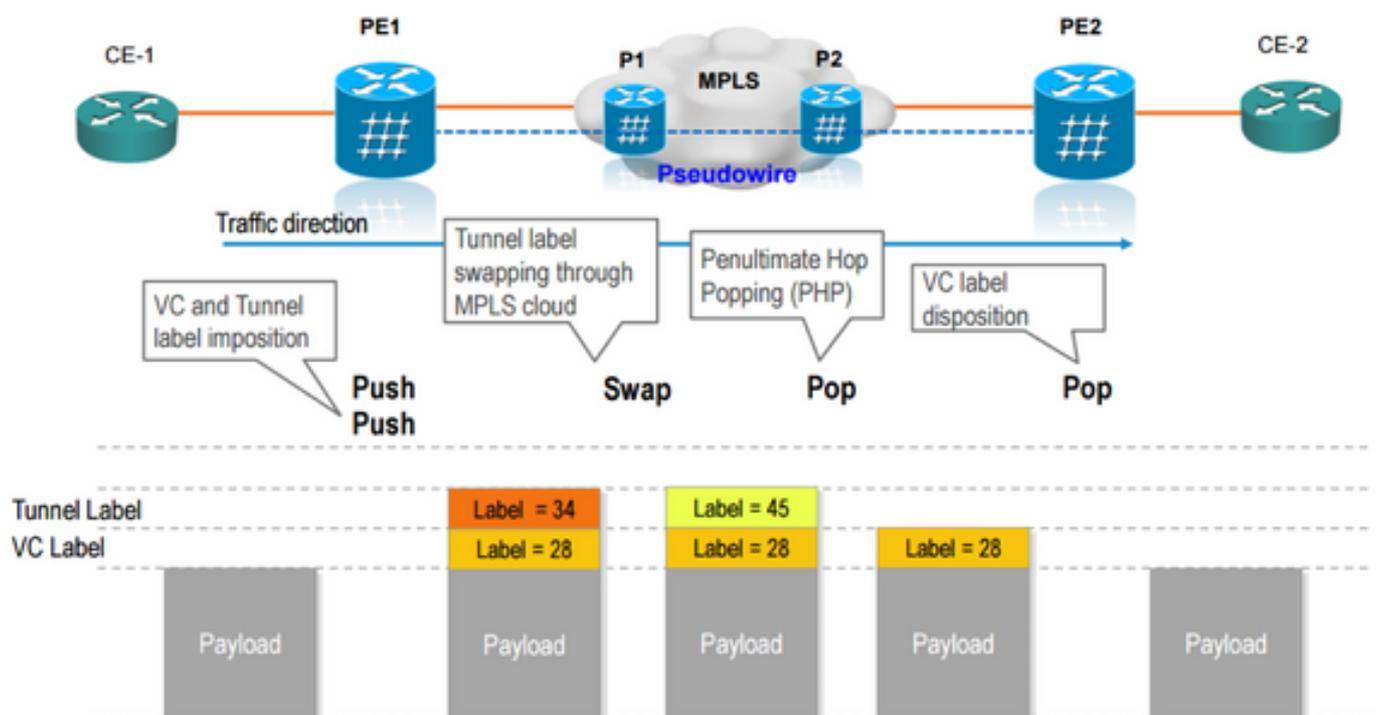
00 = 未分段

01 = 第一个分段

10 = 最后一个分片

11 = 中间分段

## 转发平面处理



入口 PE 收到来自 CE 的帧时，会通过 MPLS 主干网将帧转发到带有如下两个标签的出口 LSR：

1. 隧道标签（顶部标签）– 告知所有 LSR 和出口 PE 转发帧的目标位置。
2. VC 标签（底部标签）– 标识出口 PE 上的出口 AC。

在 AToM 网络中，每对 PE 路由器必须在相互之间进行目标 LDP 会话。

TLDP 会话可传递伪线图表的信令，最重要的是，通告 VC 标签。

## 操作

步骤1: 入口PE路由器首先将VClablel推送到帧上。然后推送隧道标签。

第二步：隧道标签是与标识远程PE的IGP前缀关联的标签。而前缀是配置 AToM 的指定位。

第三步：然后，将根据隧道标签逐跳转发MPLS数据包，直到数据包到达出口PE2。

第四步：数据包到达出口 PE 时，隧道标签已被删除。这是因为最后一个P路由器和出口PE之间的PHP行为。

第五步：出口 PE 在转发信息库中查找 VC 标签，将其剥离，然后将帧转发给正确的 AC。

## PW 状态信令

在PE路由器设置了伪线后，PE可以向远程PE发出伪线状态信号。有两种方法：

### 1. 标签撤销（较旧的方法）

- PE 路由器可以发送标签撤销消息或标签映射释放消息，进而撤销标签映射。
- 如果 AC 关闭，PE 路由器会向远程 PE 发送标签撤销消息，以此传递 PW 状态信令。
- 如果物理接口关闭，则包含组 ID 的标签撤销消息可传递此接口的所有 AC 均已关闭的信令。

### 2. PW状态TLV

- 当伪线被单接到时，PW状态TLV使用LDP标签映射TLV。这表明 PE 路由器希望使用第二种方法。
- 如果另一台 PE 路由器不支持 PW 状态 TLV 方法，则两台 PE 路由器均恢复为使用“标签撤销”方法。
- 选中伪线后，LDP 通知消息中含有 PW 状态 TLV。PW 状态 TLV 包含 32 位状态码字段。

## AToM 基本配置

步骤1: 选择封装类型。

第 2 步：在面向 CE 的接口上启用 specifying the connect 命令。

```
xconnect peer-router-id vcid encapsulation mpls
```

Peer-router-id：远程PE路由器的LDP路由器id。

VCID：您分配给PW的标识符。

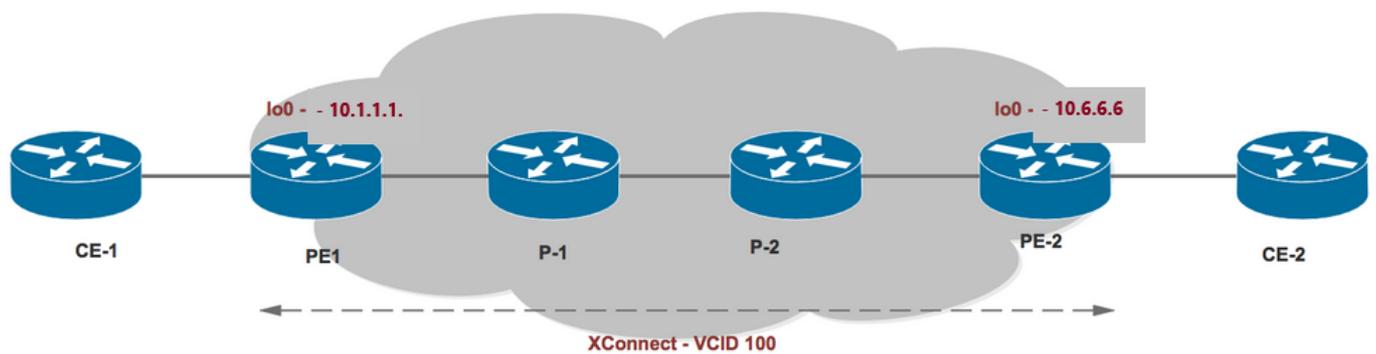
步骤3.一旦在两台PE路由器中配置了xconnect，即会在PE路由器之间建立目标LDP会话。

## 伪线数据包分析

启动从入口PE到出口PE的伪线ping。

通过点对点伪线发送 MPLS 回应请求和应答数据包。

## 拓扑



从PE1 ping PE2:

```
R1#ping mpls pseudowire 10.6.6.6 100
```

```
Sending 5, 100-byte MPLS Echos to 10.6.6.6,
```

```
timeout is 2 seconds, send interval is 0 msec:
```

```
Type escape sequence to abort.
```

```
!!!!
```

```
Success rate is 100 percent (5/5), round-trip min/avg/max = 48/61/80 ms
```

观察结果：

## 1. 回应请求：

带有2个标签 — VPN和传输

作为带有PW LABEL的标记数据包发送。可以标签交换（使用传输标签）。

标签：2

SRC IP：环回 IP（在目标 LDP 邻域中使用）

DST IP：127.0.0.1

L4类型：UDP

SRC 端口：3503

DST 端口：3505

TOS字节：关闭

MPLS EXP：关闭

DF位：开

IPv4选项字段正在使用中：路由器警报选项字段（从PUNT到CPU）

UDP负载可以是MPLS标签交换回应请求

概述：

4 0.203148 10.1.1.1 10.0.0.1 MPLS E... 130 MPLS Echo Request

```
Frame 2: 130 bytes on wire (1040 bits), 130 bytes captured (1040 bits) on interface 0
Ethernet II, Src: ca:01:1b:c0:00:06 (ca:01:1b:c0:00:06), Dst: ca:04:13:5c:00:06 (ca:04:13:5c:00:06)
MultiProtocol Label Switching Header, Label: 24, Exp: 0, S: 0, TTL: 255   Transport label
MultiProtocol Label Switching Header, Label: 28, Exp: 0, S: 1, TTL: 1   VPN label
PW Associated Channel Header
Internet Protocol Version 4, Src: 10.1.1.1 Dst: 10.0.0.1
User Datagram Protocol, Src Port: 3503 (3503), Dst Port: 3503 (3503)
Multiprotocol Label Switching Echo
```

第 2 层/标签：

```

> Frame 4: 130 bytes on wire (1040 bits), 130 bytes captured (1040 bits) on interface 0
v Ethernet II, Src: ca:01:1b:c0:00:06 (ca:01:1b:c0:00:06), Dst: ca:04:13:5c:00:06 (ca:04:13:5c:00:06)
  > Destination: ca:04:13:5c:00:06 (ca:04:13:5c:00:06)
  > Source: ca:01:1b:c0:00:06 (ca:01:1b:c0:00:06)
  Type: MPLS Label switched packet (0x8847)
v MultiProtocol Label Switching Header, Label: 24, Exp: 0, S: 0, TTL: 255
  0000 0000 0000 0001 1000 .... = MPLS Label: 24
  .... 000. .... = MPLS Experimental Bits: 0
  .... 0 .... = MPLS Bottom Of Label Stack: 0
  .... 1111 1111 = MPLS TTL: 255
v MultiProtocol Label Switching Header, Label: 28, Exp: 0, S: 1, TTL: 1
  0000 0000 0000 0001 1100 .... = MPLS Label: 28
  .... 000. .... = MPLS Experimental Bits: 0
  .... 1 .... = MPLS Bottom Of Label Stack: 1
  .... 0000 0001 = MPLS TTL: 1
v PW Associated Channel Header
  .... 0000 = Channel Version: 0
  Reserved: 0x00
  Channel Type: IPv4 packet (0x0021)
> Internet Protocol Version 4, Src: 10.1.1.1, Dst: 10.0.0.1
> User Datagram Protocol, Src Port: 3503 (3503), Dst Port: 3503 (3503)
> Multiprotocol Label Switching Echo

```

L3/L4 :

```

v PW Associated Channel Header
  .... 0000 = Channel Version: 0
  Reserved: 0x00
  Channel Type: IPv4 packet (0x0021)
v Internet Protocol Version 4, Src: 10.1.1.1, Dst: 10.0.0.1
  0100 .... = Version: 4
  .... 0110 = Header Length: 24 bytes
  > Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
  Total Length: 104
  Identification: 0xfd8f (64911)
  v Flags: 0x02 (Don't Fragment)
    0... .... = Reserved bit: Not set
    .1.. .... = Don't fragment: Set
    ..0. .... = More fragments: Not set
  Fragment offset: 0
  > Time to live: 1
  Protocol: UDP (17)
  > Header checksum: 0x65ee [validation disabled]
  Source: 10.1.1.1
  Destination: 10.0.0.1
  [Source GeoIP: unknown]
  [Destination GeoIP: Unknown]
  v Options: (4 bytes), Router Alert
    v Router Alert (4 bytes): Router shall examine packet (0)
      > Type: 148
      Length: 4
      Router Alert: Router shall examine packet (0)
v User Datagram Protocol, Src Port: 3503 (3503), Dst Port: 3503 (3503)
  Source Port: 3503
  Destination Port: 3503
  Length: 80
  > Checksum: 0x029f [validation disabled]
  [Stream index: 0]
> Multiprotocol Label Switching Echo

```

实际 MPLS 负载：

```

v Multiprotocol Label Switching Echo
  Version: 1
  > Global Flags: 0x0000
  Message Type: MPLS Echo Request (1)
  Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
  Return Code: No return code (0)
  Return Subcode: 0
  Sender's Handle: 0xc7735d85
  Sequence Number: 284
  Timestamp Sent: Feb  3, 2017 10:41:23.998999000 UTC
  Timestamp Received: Jan  1, 1970 00:00:00.000000000 UTC
  v Vendor Private
    Type: Vendor Private (64512)
    Length: 12
    Vendor Id: ciscoSystems (9)
    Value: 0001000400000004
  v Target FEC Stack
    Type: Target FEC Stack (1)
    Length: 20
    v FEC Element 1: FEC 128 Pseudowire (new)
      Type: FEC 128 Pseudowire (new) (10)
      Length: 14
      Sender's PE Address: 10.1.1.1
      Remote PE Address: 10.6.6.6
      VC ID: 100
      Encapsulation: Ethernet (5)
      MBZ: 0x0000
      Padding: 0000

```

2. 回应应答：

可携带1标签 — 运输。

作为单播数据包发送。由于核心中存在LDP，因此可以进行标签交换（使用传输标签）。

标签：1

源IP：送出接口IP地址（本例中为10.1.6.2）

DST IP：回应请求中发现的源IP - 源路由器的环回

L4类型：UDP

SRC 端口：3503

DST 端口：3505

TOS字节：关闭

MPLS EXP：关闭

DF位：开

UDP负载可以是MPLS标签交换回应应答

MPLS EXP 已开启且设置为 6

“DF 位”已开启

VC 详细信息以供参考：

<#root>

```
R1#sh mpls l2transport vc detail
```

```
Local interface: Fa2/0 up, line protocol up, Ethernet up
```

```
Destination address: 10.6.6.6
```

```
,
```

```
VC ID: 100, VC status: up
```

```
Output interface: Fa0/1, imposed label stack {24 28}
```

```
Preferred path: not configured
```

```
Default path: active
```

```
Next hop: 10.1.1.2
```

```
Create time: 2d17h, last status change time: 2d17h
```

```
Last label FSM state change time: 2d17h
```

```
Signaling protocol: LDP, peer 10.6.6.6:0 up
```

```
Targeted Hello: 10.1.1.1(LDP Id) -> 10.6.6.6, LDP is UP
```

```
Status TLV support (local/remote) : enabled/supported
```

```
LDP route watch : enabled
```

```
Label/status state machine : established, LruRru
```

```
Last local dataplane status rcvd: No fault
```

```
Last BFD dataplane status rcvd: Not sent
```

```
Last BFD peer monitor status rcvd: No fault
```

```
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: No fault
```

```
Last remote LDP TLV status rcvd: No fault
```

```
Last remote LDP ADJ status rcvd: No fault
```

```
MPLS VC labels: local 28, remote 28
```

Group ID: local 0, remote 0

MTU: local 1500, remote 1500

Remote interface description:

Sequencing: receive enabled, send enabled

Sequencing resync disabled

Control Word: On (configured: autosense)

Dataplane:

SSM segment/switch IDs: 4097/4096 (used), PWID: 1

VC statistics:

transit packet totals: receive 1027360, send 1027358

transit byte totals: receive 121032028, send 147740215

transit packet drops: receive 0, seq error 0, send 0

## L2VPN 交互操作

L2VPN 交互操作基于上述功能，允许连接不同的接入电路。交互操作功能可促进不同第 2 层封装之间的转换。思科系列路由器早期版本仅支持网桥交互操作，也称为“以太网交互操作”。

到目前为止，网络两端的 AC 都属于相同的封装类型，也称为“同类功能”。

L2VPN 交互操作具有 AToM 特性，允许在 AToM 网络两端使用不同的封装类型

- 这需要将两个异质接入电路 (AC) 互连。
- Cisco IOS 软件支持以下两大 L2VPN 交互操作 (IW) 功能：

1. IP/路由：在 MPLS 云的一端删除（并替换为 MPLS 标签）MAC 报头，并在另一个 PE 构建新的 MAC 报头。IP 报头按原样保留。

2. 以太网/桥接：MAC 报头未删除。MPLS 标签会加在 MAC 报头之上，而 MAC 报头按原样传送到 MPLS 云的另一端。

### 可能的交互操作

a. 帧中继到以太网

b. FR 到 PPP

c. 帧中继到 ATM

d. 以太网到 VLAN

e.以太网到PPP

## 相关信息

- [RFC编辑器4664](#)
- [RFC编辑器4667](#)
- [技术支持和文档 - Cisco Systems](#)

## 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）。