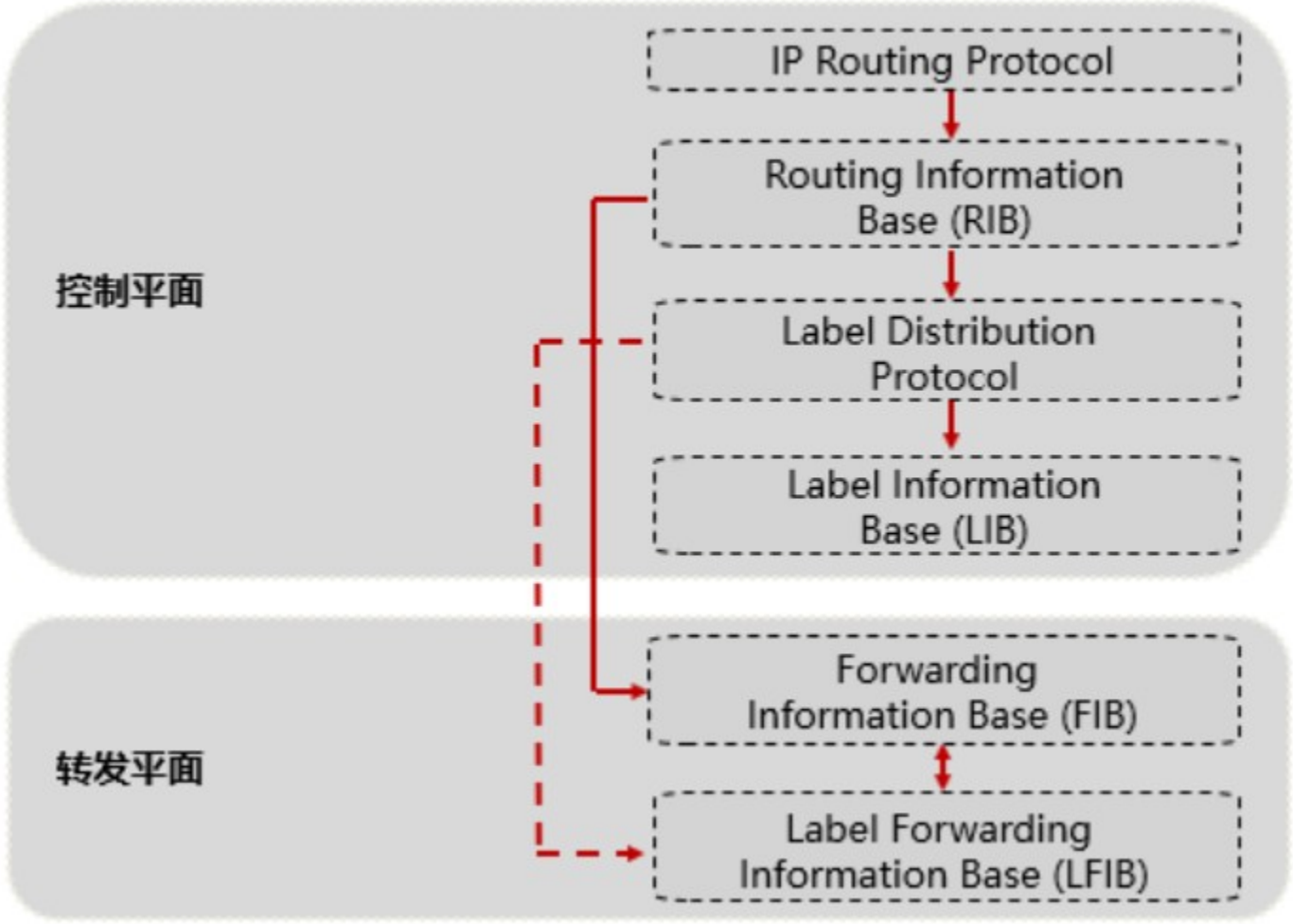
# HCRSE111-MPLS BGP VPN 跨域

MPLS （Multiprotocol Label Switching）多协议标签交换是一种 IP（Internet Protocol）骨干网技术。MPLS 在无连接的 IP 网络上引入面向连接的标签交换概念，将第三层路由技术和第二层交换技术相结合，充分发挥了 IP 路由的灵活性和二层交换的简捷性。

MPLS 起源于 IPv4，其核心技术可扩展到多种网络协议，包括 IPv6、IPX（和 CLNP 等。MPLS 中的“Multiprotocol”指的就是支持多种网络协议。由此可见，MPLS 并不是一种业务或者应用，它实际上是一种隧道技术。这种技术不仅支持多种高层协议与业务，而且在一定程度上可以保证信息传输的安全性。

## MPLS 的体系结构

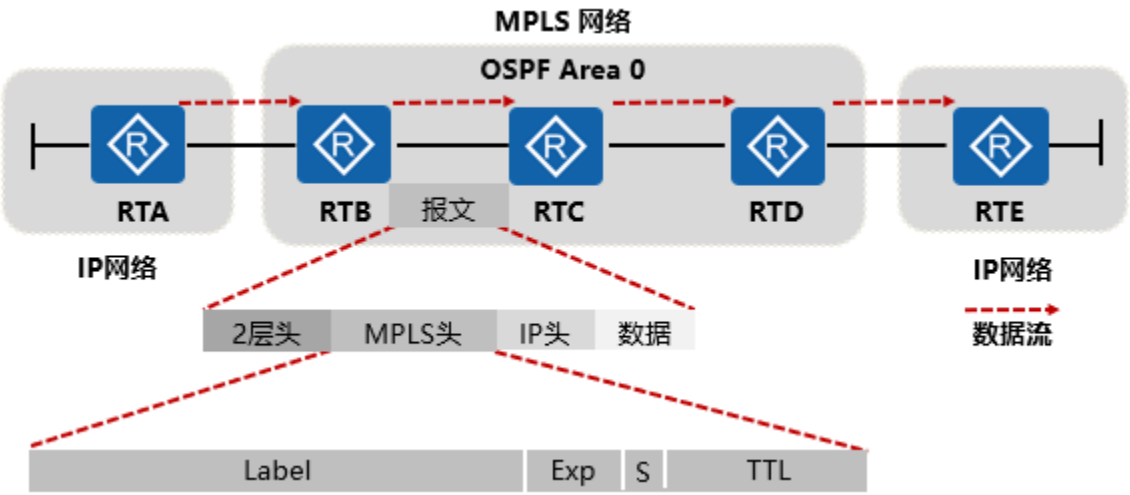



- 控制平面：负责产生和维护路由信息以及标签信息。
- 路由信息表 RIB（Routing Information Base）：由 IP 路由协议（IP Routing Protocol）生成，用于选择路由。
- 标签分发协议 LDP（Label Distribution Protocol）：负责标签的分配、标签转发信息表的建立、标签交换路径的建立、拆除等工作。
- 标签信息表 LIB（Label Information Base）：由标签分发协议生成，用于管理标签信息。
- 转发平面：即数据平面（Data Plane），负责普通 IP 报文的转发以及带 MPLS 标签报文的转发。
- 转发信息表 FIB（Forwarding Information Base）：从 RIB 提取必要的路由信息生成，负责普通 IP 报文的转发。

- 标签转发信息表 LFIB（Label Forwarding Information Base）：简称标签转发表，由标签分发协议建立 LFIB，负责带 MPLS 标签报文的转发。
- MPLS 路由器上，报文的转发过程：
- 当收到普通 IP 报文时，查找 FIB 表，如果 Tunnel ID 为 0x0，则进行普通 IP 转发；如果查找 FIB 表，Tunnel ID 为非 0x0，则进行 MPLS 转发。

当收到带标签的报文时，查找 LFIB 表，如果对应的出标签是普通标签，则进行 MPLS 转发；查找 LFIB 表，如果对应的出标签是特殊标签，如标签 3，则将报文的标签去掉，进行 IP 转发。

## MPLS 数据报文结构




MPLS 标签封装在链路层和网络层之间，可以支持任意的链路层协议，MPLS 标签的封装结构如图所示。

MPLS 标签的长度为 4 个字节，共分 4 个字段：
Label：20bit，标签值域；
Exp：3bit，用于扩展。通常用 CoS（Class of Service），设备发生阻塞时，优先发送优先级高的报文；

S：1bit，栈底标识。MPLS 支持多层标签，即标签嵌套。S 值为 1 时表明为最底层标签；
TTL：8bit，和 IP 报文中的 TTL（Time To Live）意义相同。

```
Frame 41: 106 bytes on wire (848 bits), 106 bytes captured (848 bits) or
Ethernet II, Src: HuaweiTe_32:76:56 (54:89:98:32:76:56), Dst: HuaweiTe_1
MultiProtocol Label Switching Header, Label: 1025, Exp: 0, S: 0, TTL: 25
   0000 0000 0100 0000 0001 .... .... .... = MPLS Label: 1025
   .... .... .... .... .... 000. .... .... = MPLS Experimental Bits: 0
   .... .... .... .... .... ...0 .... .... = MPLS Bottom Of Label Stack:
   .... .... .... .... .... .... 1111 1111 = MPLS TTL: 255
MultiProtocol Label Switching Header, Label: 1029, Exp: 0, S: 1, TTL: 25
Internet Protocol Version 4, Src: 9.9.9.9, Dst: 10.10.10.10
```

标签空间是指标签的取值范围。标签空间划分如下：
0～15：特殊标签。如标签 3，称为隐式空标签，用于倒数第二跳弹出；
16～1023：静态 LSP 和静态 CR-LSP 共享的标签空间；
1024 及以上：LDP、RSVP-TE、MP-BGP 等动态信令协议的标签空间。

建立 LSP 的方式有两种：
静态 LSP：用户通过手工方式为各个转发等价类分配标签建立转发隧道；
动态 LSP：通过标签发布协议动态建立转发隧道。

静态 LSP
配置 MPLS 协议，首先需要配置 LSR ID，全局开启 MPLS，在转发 MPLS 的接口上开启 MPLS 功能
手动建立一条的静态 LSP
配置 LSR ID 用来在网络中唯一标识一个 MPLS 路由器。缺省没有配置 LSR ID，必须手工配置。为了提高网络的可靠性，

推荐使用 LSR 某个 Loopback 接口的地址作为 LSR ID。

```
static-lsp ingress R3toR1 destination 1.1.1.1 32 nexthop 192.168.23.2 out-label 302
static-lsp transit R3toR1 incoming-interface g0/0/1 in-label 302 nexthop 192.168.12.1 out-label 201
static-lsp egress R3toR1   incoming-interface g0/0/0 in-label 201
```

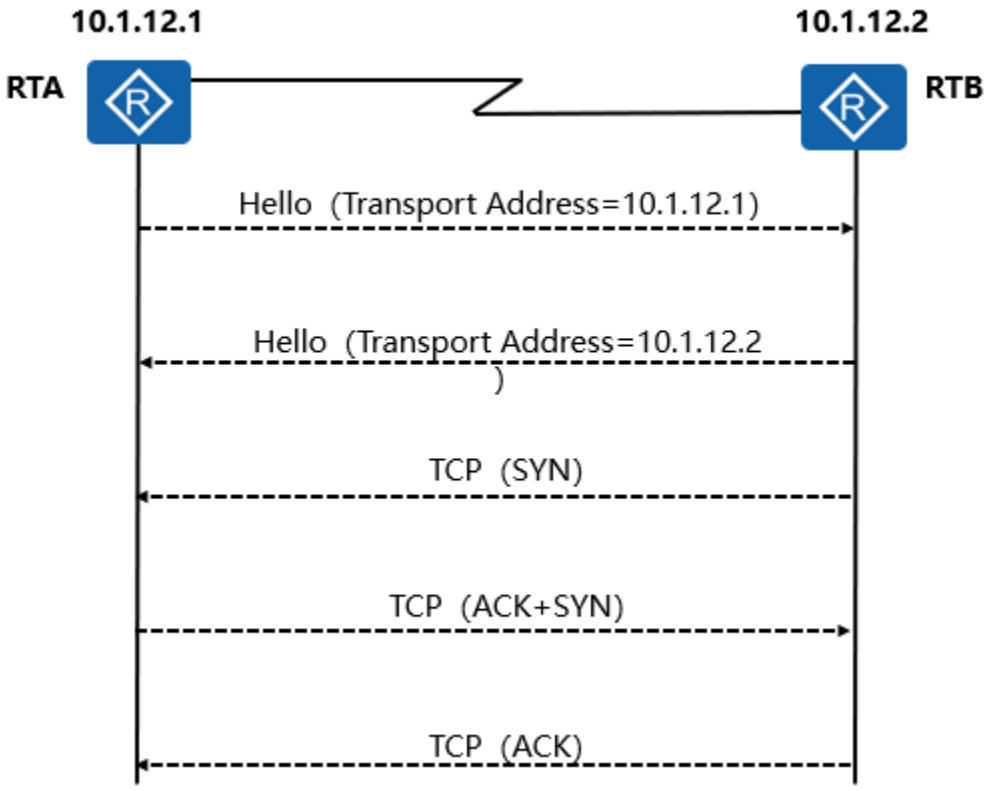
## 动态 LSP

动态 LSP 通过 LDP 协议实现对 FEC 的分类、标签的分配及 LSP 的建立和维护等操作。

动态 LSP 的特点：

组网配置简单，易于管理和维护；

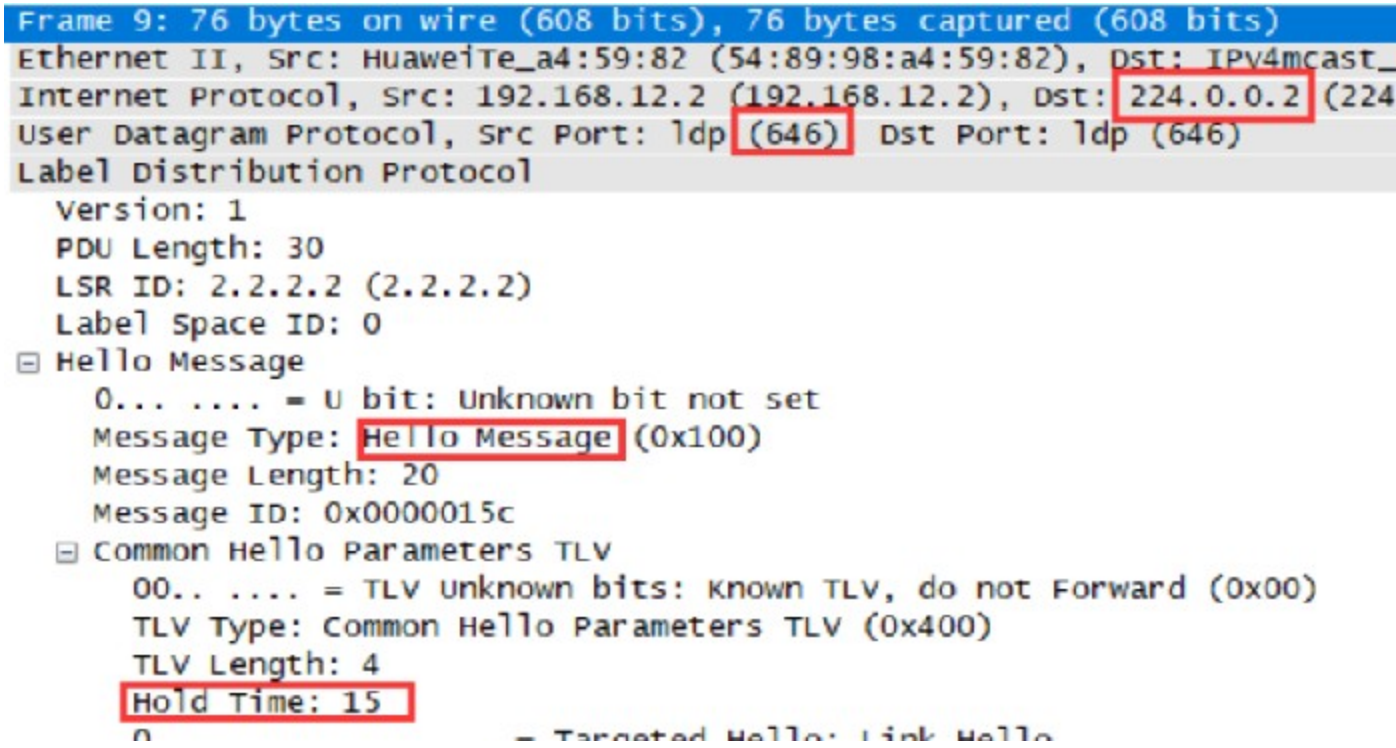
支持基于路由动态建立 LSP，网络拓扑发生变化时，能及时反映网络状况。



LDP 的 hello 时间为 5s ，holdtime 15s，发送的组播地址为 224.0.0.2 ，端口号为 UDP 646

MPLS 路由器通过周期性地发送 LDP 链路 Hello 消息（LDP Link Hello），实现 LDP 邻居的发现，并建立本地 LDP 会话。为了能使开启 LDP 协议的设备快速发现邻居，LDP 的 Hello 消息使用 UDP 封装。UDP 是无连接的协议，为了保证邻居的有效性和可靠性，Hello 消息周期发送，发送周期为 5s，使用组播 224.0.0.2 作为目的 IP 地址，意思是“发送给网络中的所有路由器”。

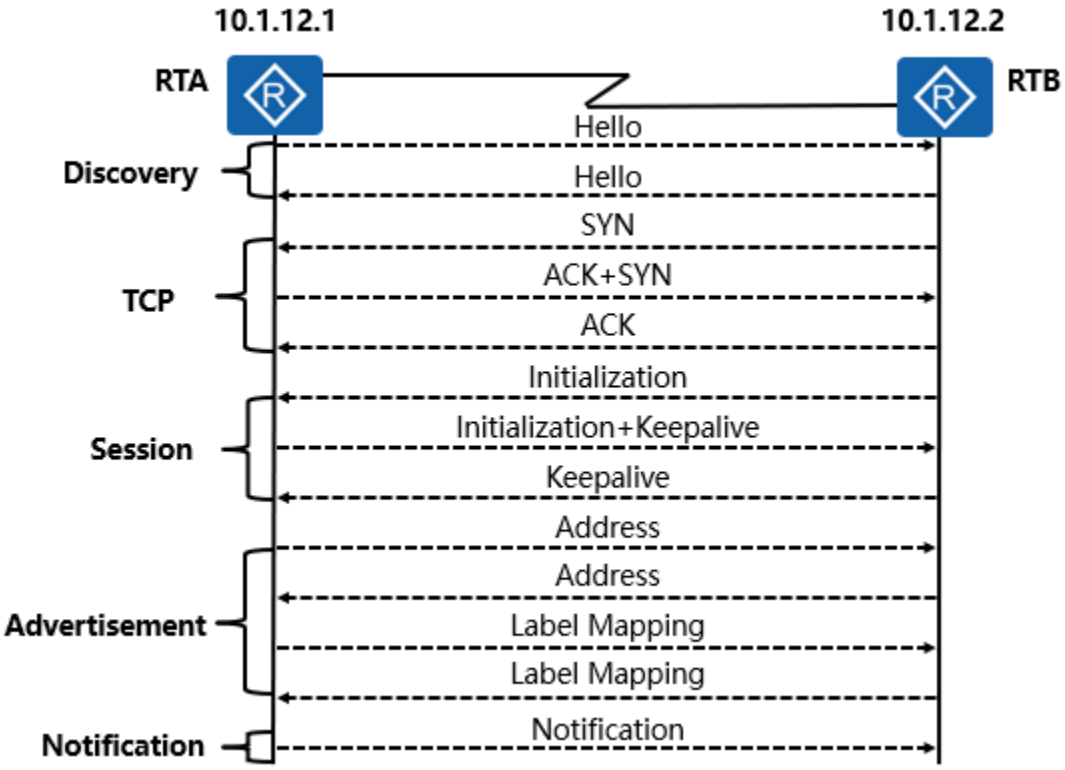
LDP 的 Hello 消息中，携带有 Transport Address 字段，该字段与设备配置的 LSR ID 一致，表明与对端建立邻居关系时所使用的 IP 地址。如果该字段 IP 地址是直连接口 IP 地址，则直接建立邻居关系；如果该字段地址是 LoopBack 接口 IP 地址，保证该接口 IP 地址路由可达，才能建立邻居关系。




LDP 协议四类消息：

发现（Discovery）消息：用于通告和维护网络中邻居的存在，如 Hello 消息。
会话（Session）消息：用于建立、维护和终止 LDP 对等体之间的会话，
通告（Advertisement）消息：用于创建、改变和删除 FEC 的标签映射，
通知（Notification）消息：用于提供建议性的消息和差错通知。

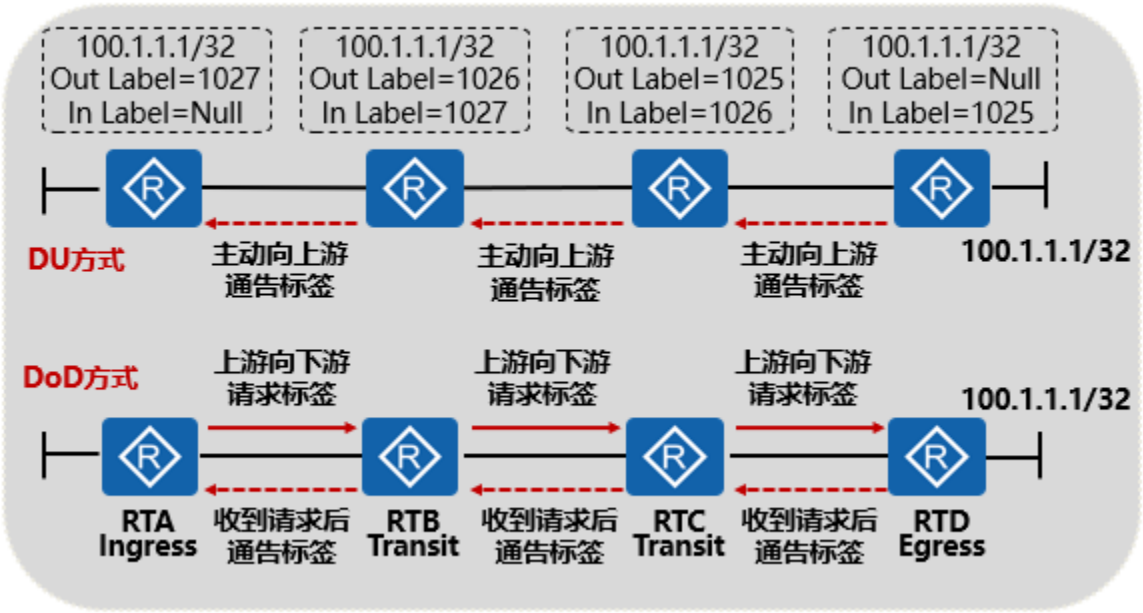
两个 LSR 之间互相发送 Hello 消息。
Hello 消息中携带传输地址，双方使用传输地址建立 LDP 会话。
传输地址较大的一方作为主动方，发起 TCP 连接。
TCP 连接建立成功后，由主动方发送初始化消息，协商建立 LDP 会话的相关参数。
LDP 会话的相关参数包括 LDP 协议版本、标签分发方式、Keepalive 保持定时器的值、最大 PDU 长度和标签空间等。
被动方收到初始化消息后，如果接受相关参数，则发送初始化消息，同时发送 Keepalive 消息给主动方。
如果被动方不能接受相关参数，则发送 Notification 消息终止 LDP 会话的建立。
主动方收到初始化消息后，接受相关参数，则发送 Keepalive 消息给被动方。
如果主动方=不能接受相关参数，则发送 Notification 消息给被动方=终止 LDP 会话的建立。
当双方都收到对端的 Keepalive 消息后，LDP 会话建立成功。
LDP 会话建立成功后，进行 FEC 的创建与标签的分发。




==========

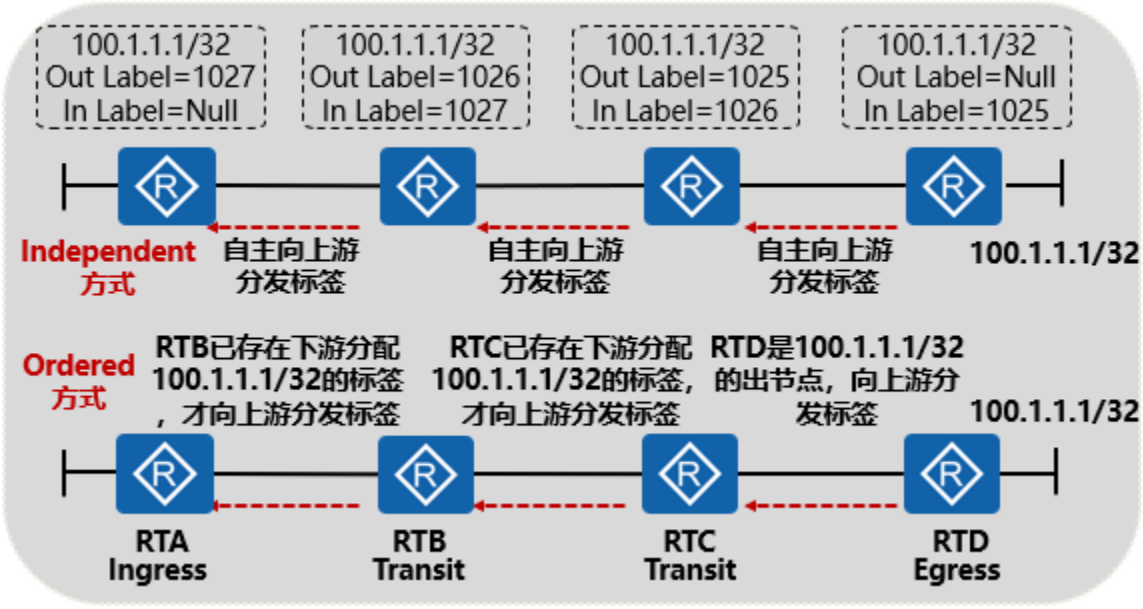
## 标签管理

### 标签的发布方式：华为采用 DU 下游自主



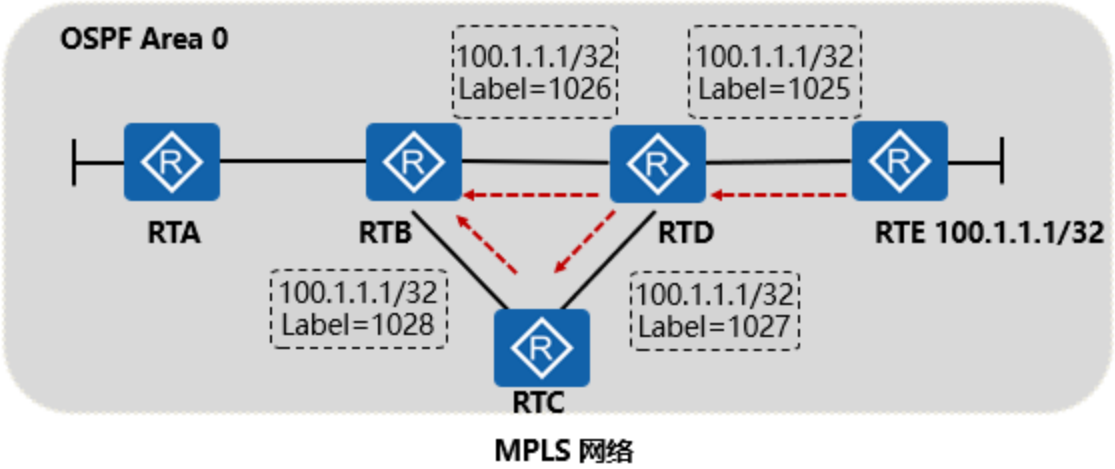
DU（Downstream Unsolicited，下游自主方式）：
对于一个到达同一目地址报文的分组，LSR 无需从上游获得标签请求消息即可进行标签分配与分发。
DoD（Downstream on Demand，下游按需方式）：
对于一个到达同一目的地址报文的分组，LSR 获得标签请求消息之后才进行标签分配与分发。

标签的分配控制方式：华为采用 Ordered 有序标签分配控制
采用 Ordered 方式，只有当该 LSR 已经具有此 IP 分组的下一跳的标签时，才能向上游分发标签，这样是为了避免下游 IP 分组的标签未分配或收敛时间较长，上游的标签已分配，数据开始转发造成的数据丢失。




Independent（独立标签分配控制方式）：本地 LSR 可以自主地分配一个标签绑定到某个 IP 分组，并通告给上游 LSR，而无需等待下游的标签。
Ordered（有序标签分配控制方式）：只有当该 LSR 已经具有此 IP 分组的下一跳的标签，或者该 LSR 就是该 IP 分组的出节点时，该 LSR 才可以向上游发送此 IP 分组的标签。

标签的保持方式：华为采用　Liberal 自由标签保持




Liberal（自由标签保持方式）：对于从邻居 LSR 收到标签映射，无论邻居 LSR 是不是自己的下一跳都保留。

Conservative（保守标签保持方式）：对于从邻居 LSR 收到的标签映射，只有当邻居 LSR 是自己的下一跳时才保留。

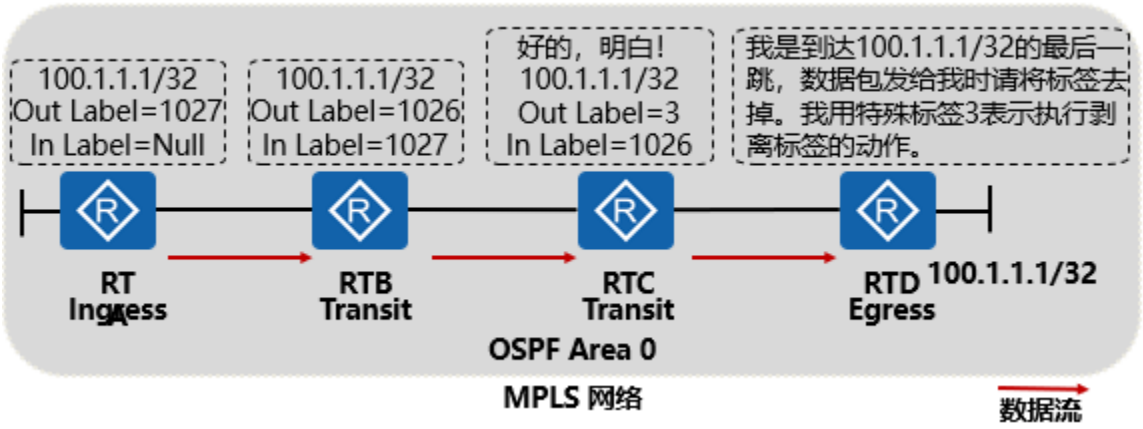
当网络拓扑变化引起下一跳邻居改变时：

使 Liberal，LSR 直接利用原来非下一跳邻居发来的标签，迅速重建 LSP，但需要更多的内存和标签空间。

使 Conservative，LSR 只保留来自下一跳邻居的标签，节省内存和标签空间，但 LSP 的重建会比较慢。

PHP（Penultimate Hop Popping，倒数第二跳弹出）

如果 MPLS 网络中的业务量很大，则每次数据包在 Egress 节点都要进行两次处理才能进行正确的路由转发，这样会导致 Egress 节点的处理压力增加，路由器的处理性能降低。我们希望在 Egress 节点上只处理一次就能将数据包正确转发，以提高 Egress 的转发性能，所以提出了 PHP 技术。

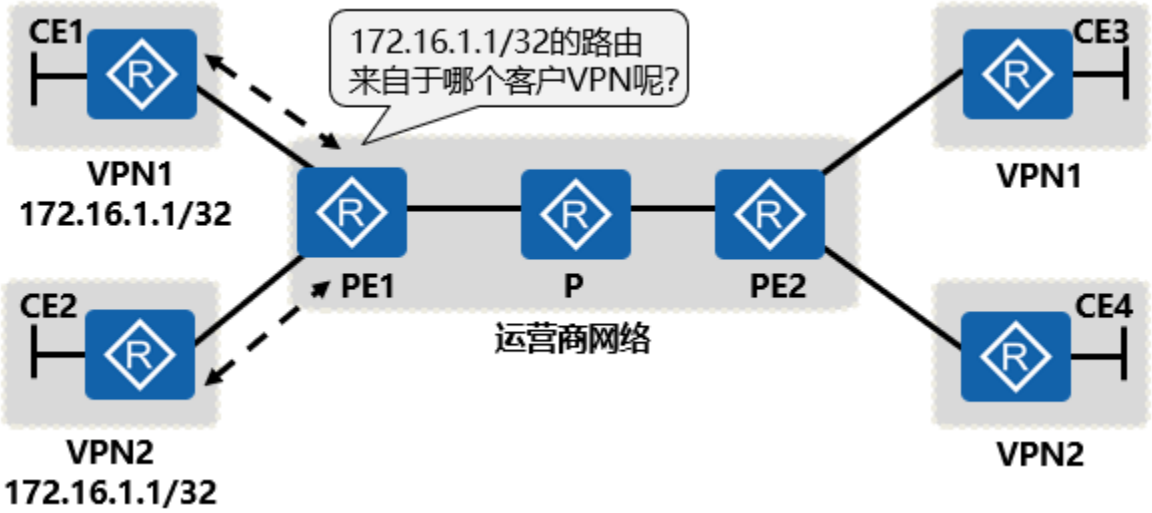



倒数第二跳弹出具体过程如下：
RTC 收到 RTB 发送的带标签 1026 的报文，查找 LFIB 表，发现分配的出标签为隐式空标签 3，于是执行弹出标签的动作，并将 IP 数据包转发给下游路由器 RTD；
RTD 收到 RTC 发送的 IP 报文，直接查找自己的 FIB 表，根据 FIB 表中的出接口进行 IP 数据的封装并转发。

==========

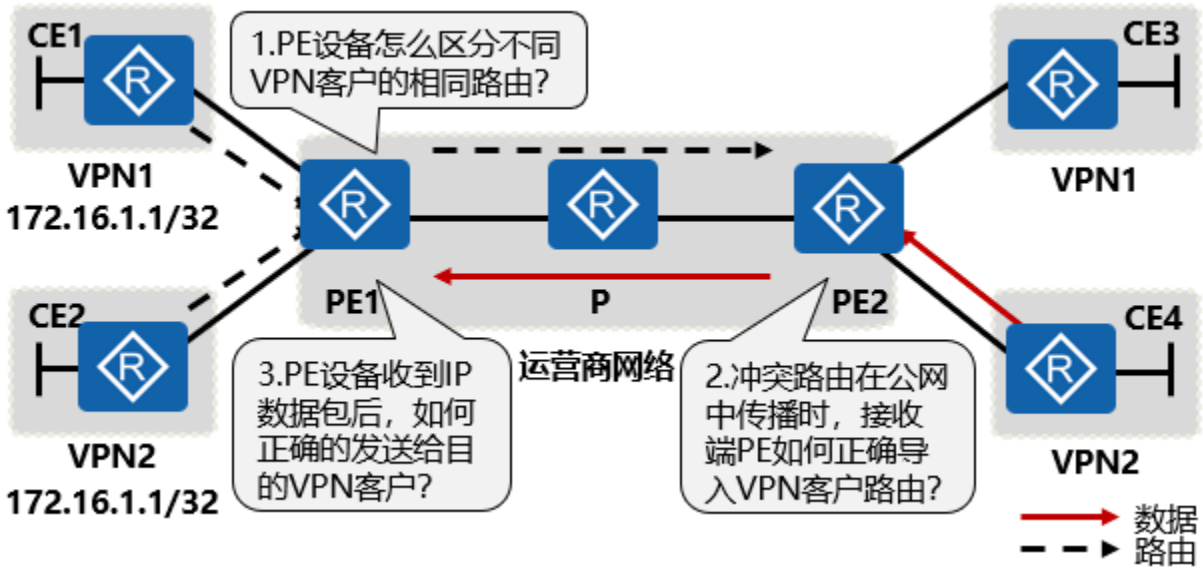
## MPLS VPN

传统的 VPN 技术存在一些固有的缺陷，导致客户组网时的很多需求无法得到满足，并且实施比较复杂，MPLS VPN 的出现解决了传统 VPN 技术的固有缺陷——地址空间的重叠问题。



两个客户的 VPN 存在相同的地址空间，传统 VPN 网络结构中的设备无法区分客户重叠的路由信息。

## 解决地址空间重叠

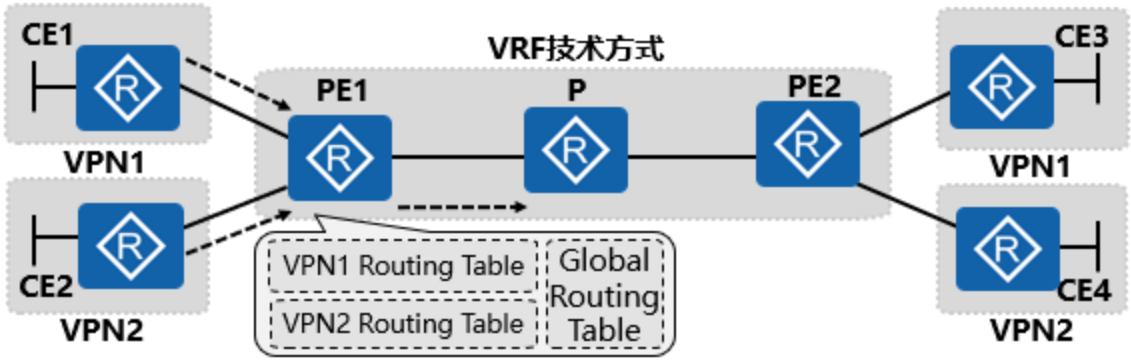



由于 BGP 的诸多优点对技术难点的解决提供了思路：
公共网络上的 VPN 路由数量庞大，BGP 是唯一支持大量路由的协议；
BGP 的报文基于 TLV 的结构，便于扩展；
BGP 可以承载附加在路由后面的任何信息，并作为可选属性传递给其他邻居。

上面提到的 3 个技术难点迎刃而解：
本地路由冲突的问题：可以通过在同一台 PE 设备上为不同的 VPN 建立单独的路由，这样冲突的的路由就被隔离开来；
在路由传递过程中，为不同的 VPN 路由添加不同的标识，以示区别。这些标识可以作为 BGP 属性进行传递；
由于 IP 报文不可更改，可以在 IP 报文头前加一些信息。由始

发路由器打上标记，接收路由器在收到带标记的的数据包时，根据标记转发给正确的 VPN。

本地路由冲突

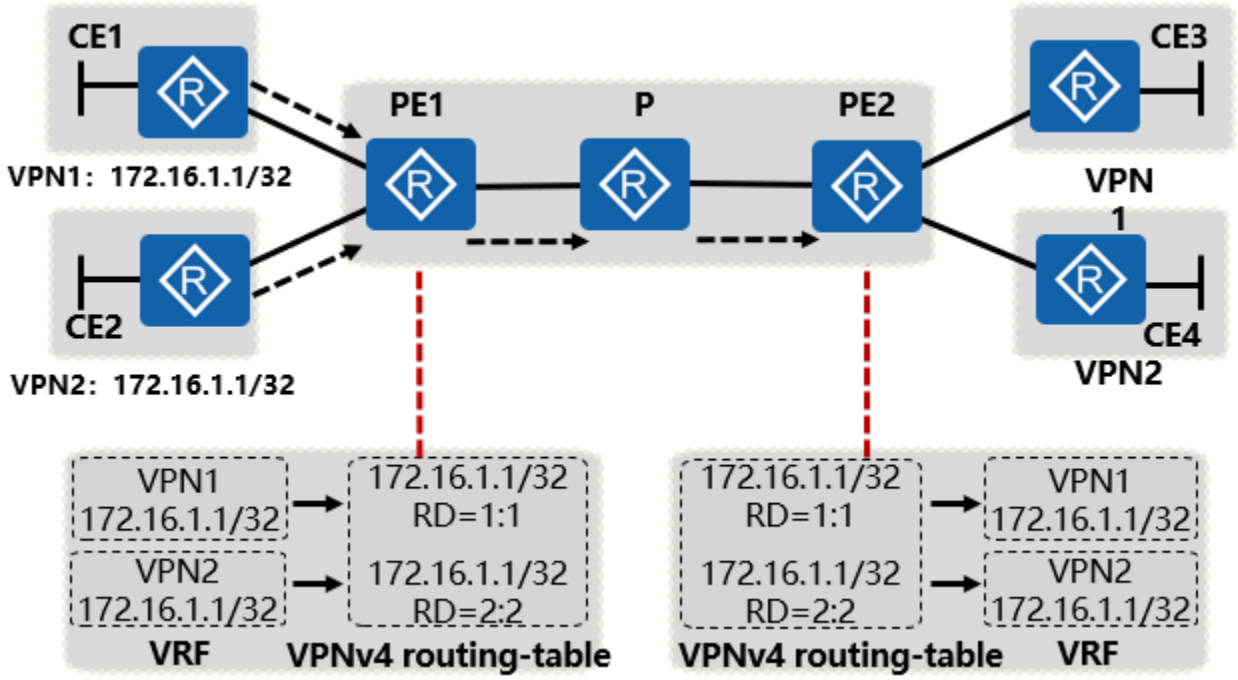



共享 PE 设备上实现重叠路由的隔离就是在 PE 设备上将来自每个 VPN 的路由放入自己对应的 VPN Routing Table 中，每个 VPN Routing Table 只记录对应 VPN 中学来的路由，就像是专用 PE 一样。这个 VPN Routing Table 称谓 VRF（VPN Routing and Forwarding table），即 VPN 路由转发表。
每一个 VRF 都需要对应一个 VPN instance，VPN 用户对应的接口绑定到 VPN instance 中。
对于每个 PE，可以维护一个或多个 VPN instance ，同时维护一个公网的路由表（也叫全局路由表），多个 VPN instance 实例相互独立且隔离。其实实现 VPN instance 并不困难，关键在于如何在 PE 上使用特定的策略规则来协调各 VPN instance 和全局路由表之间的关系。

网络传递过程中区分冲突路由
RD：区分私网路由，标识不同的 vpn 实例　　VPNv4 路由表
通过 MP-BGP 传递
RT：对路由的喜好，接收或不接收　　封装在 BG

## P 的扩展 Community 属性

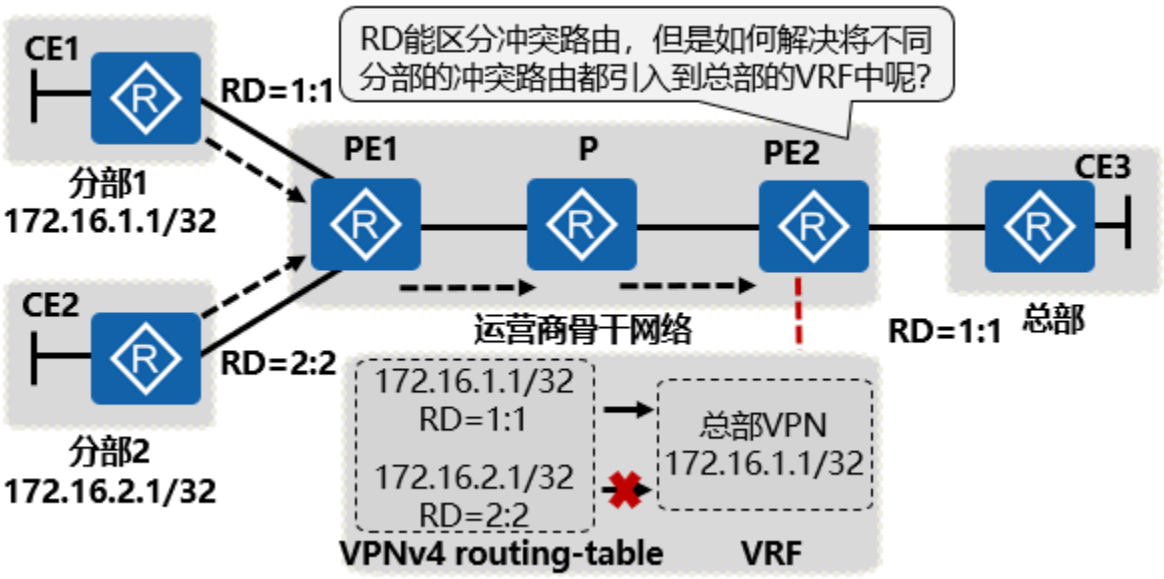
将 VPN 路由发布到全局路由表之前，使用一个全局唯一的标识和路由绑定，以区分冲突的私网路由。这个标识被称为 RD（Route Distinguisher）



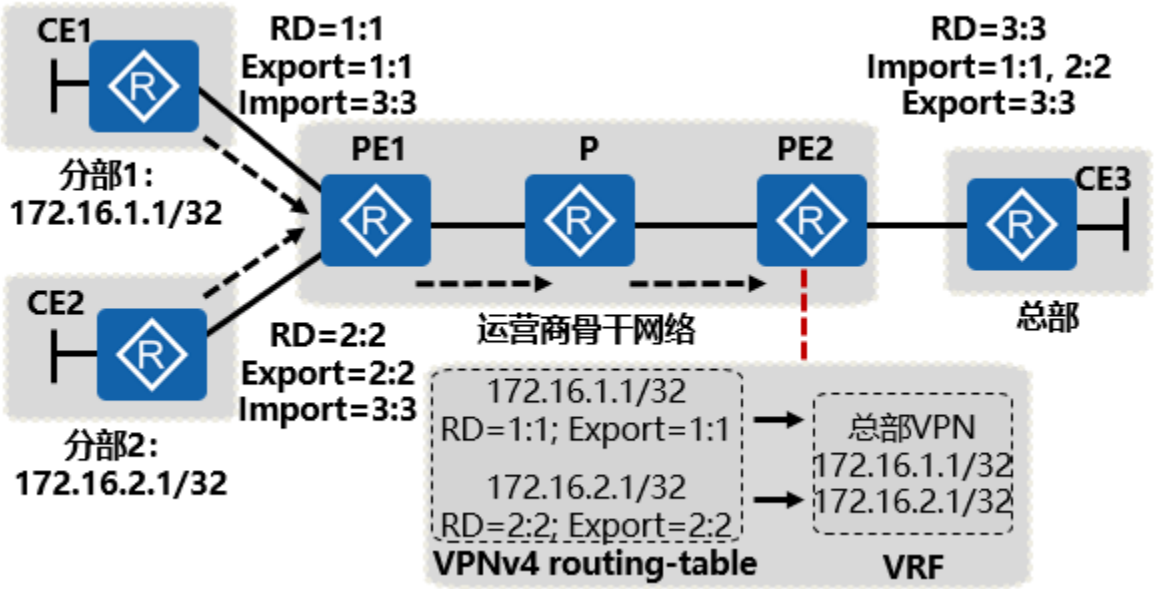

RD 即 VPN 路由标识符，由 8 字节组成，配置时同一 PE 设备上分配给每个 VPN 的 RD 必须唯一。

RD 用于区分使用相同地址空间的 IPv4 前缀，增加了 RD 的 IPv4 地址称为 VPN-IPv4 地址（即 VPNv4 地址）。

运营商设备采用 BGP 协议作为承载 VPN 路由的协议，并将 BGP 协议进行了扩展，称为 MP-BGP。PE 从 CE 接收到客户的 IPv4 私网路由后，将客户的私网路由添加各种标识信息后变为 VPNv4 路由放入 MP-BGP 的 VPNv4 路由表中，并通过 MP-BGP 协议在公网上传递。




RD 不能解决 VPN 路由正确引入 VPN 的问题。
我们需要一种类似于 Tag 的标识。这个标识由人工分配，发送端 PE 发送时打上标识，接收端 PE 收到后，根据需要将带有相应标识的路由引入 VPN。



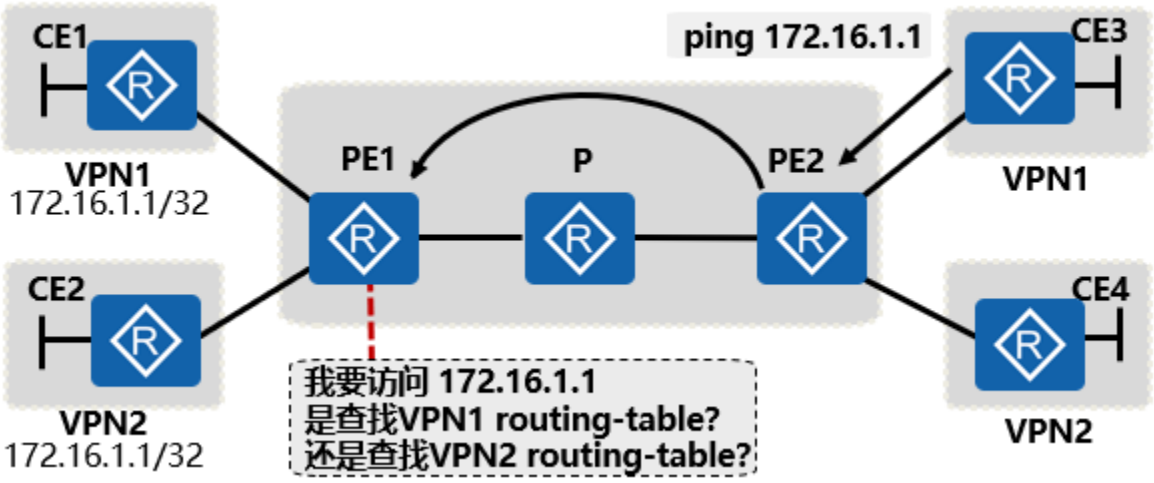

RT 属性用于将路由正确引入 VPN，有两类 VPN Target 属性，Import Target 和 Export Target，分别用于 VPN 路由的导出与导入。

使用 RT 实现本端与对端的路由正确引入 VPN，原则如下：
本端的 Export Target=对端的 Import Target，本端的 Import Target=对端的 Export Target。

RT（Route Target）封装在 BGP 的扩展 Community 属性中，在路由传递过程中作为可选可传递属性进行传递。RT 的本质是每个 VRF 表达自己的路由取舍及喜好的属性，有两类 VPN Target 属性：
Export Target：本端的路由在导出 VRF，转变为 VPNv4 的路由时，标记该属性；
Import Target：对端收到路由时，检查其 Export Target 属性。当此属性与 PE 上某个 VPN 实例的 Import Target 匹配时，PE 就把路由加入到该 VPN 实例中。

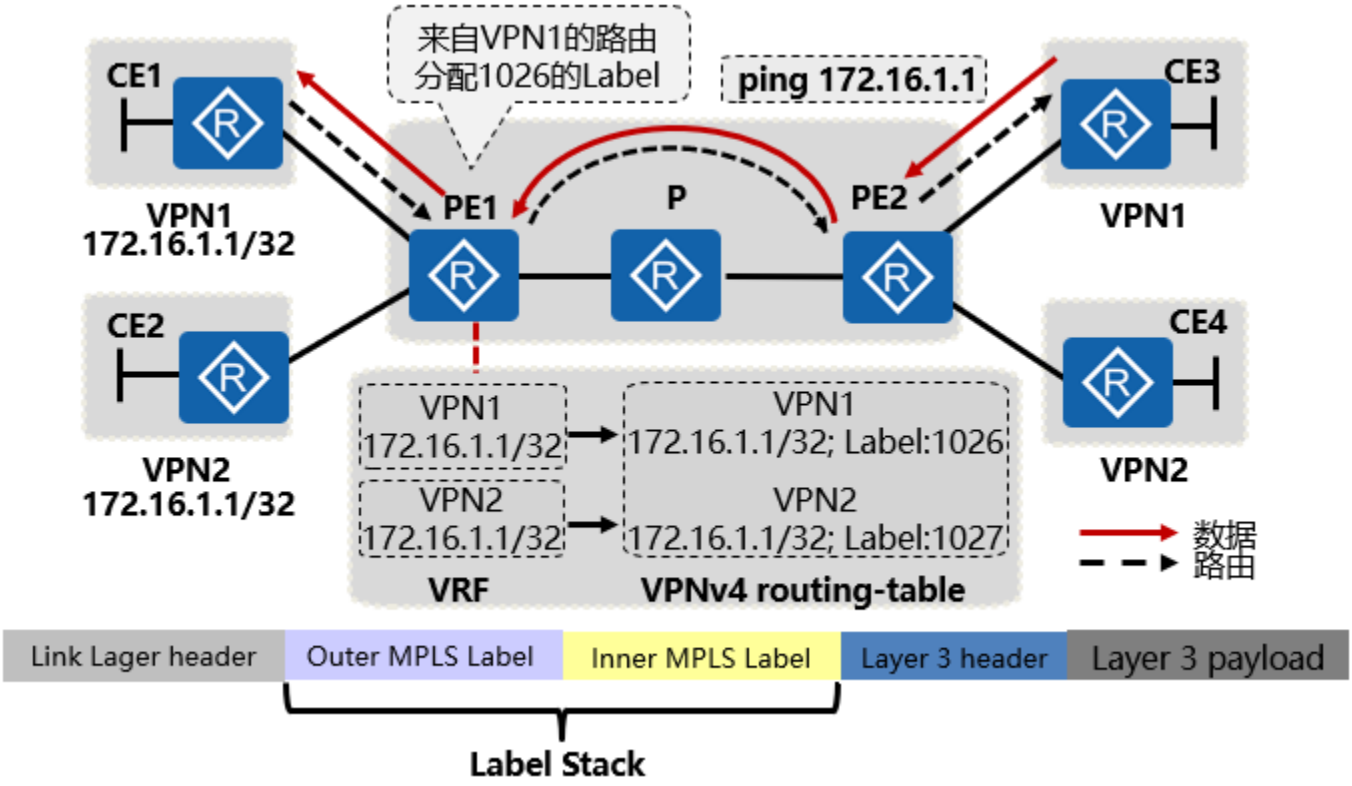
## 数据转发过程中冲突路由的查找

因为数据包没有携带任何标识，所以在 ICMP 的数据包到达 PE1 时，PE1 并不知道该查找哪个 VPN 的路由表找到正确的目标地址。




解决该问题的方案有两种：采用第二种
在数据包中增加标识信息，并且使用 RD 作为区分数据包所属 VPN 的标识符，数据转发时也携带 RD 信息。缺点是由于 RD

由 8 字节组成，额外增大数据包，会导致转发效率降低。
借助公网中已经实施的 MPLS 协议建立的标签隧道，采用标签作为数据包正确转发的标识，MPLS 标签支持嵌套，可以将区分数据包所属 VPN 的标签封装在公网标签内。




MPLS 标签嵌套的应用
使用标签嵌套解决数据转发过程中冲突路由的查找问题。
Outer MPLS Label 在 MPLS VPN 中被称为公网标签，用于 MPLS 网络中转发数据。一般公网标签会在到达 PE 设备时已被倒数第二跳剥掉，漏出 Inner Label。Inner MPLS Label 在 MPLS VPN 中被称为私网标签，用于将数据正确发送到相应的 VPN 中，PE 依靠 Inner Label 区分数据包属于哪个 VPN。

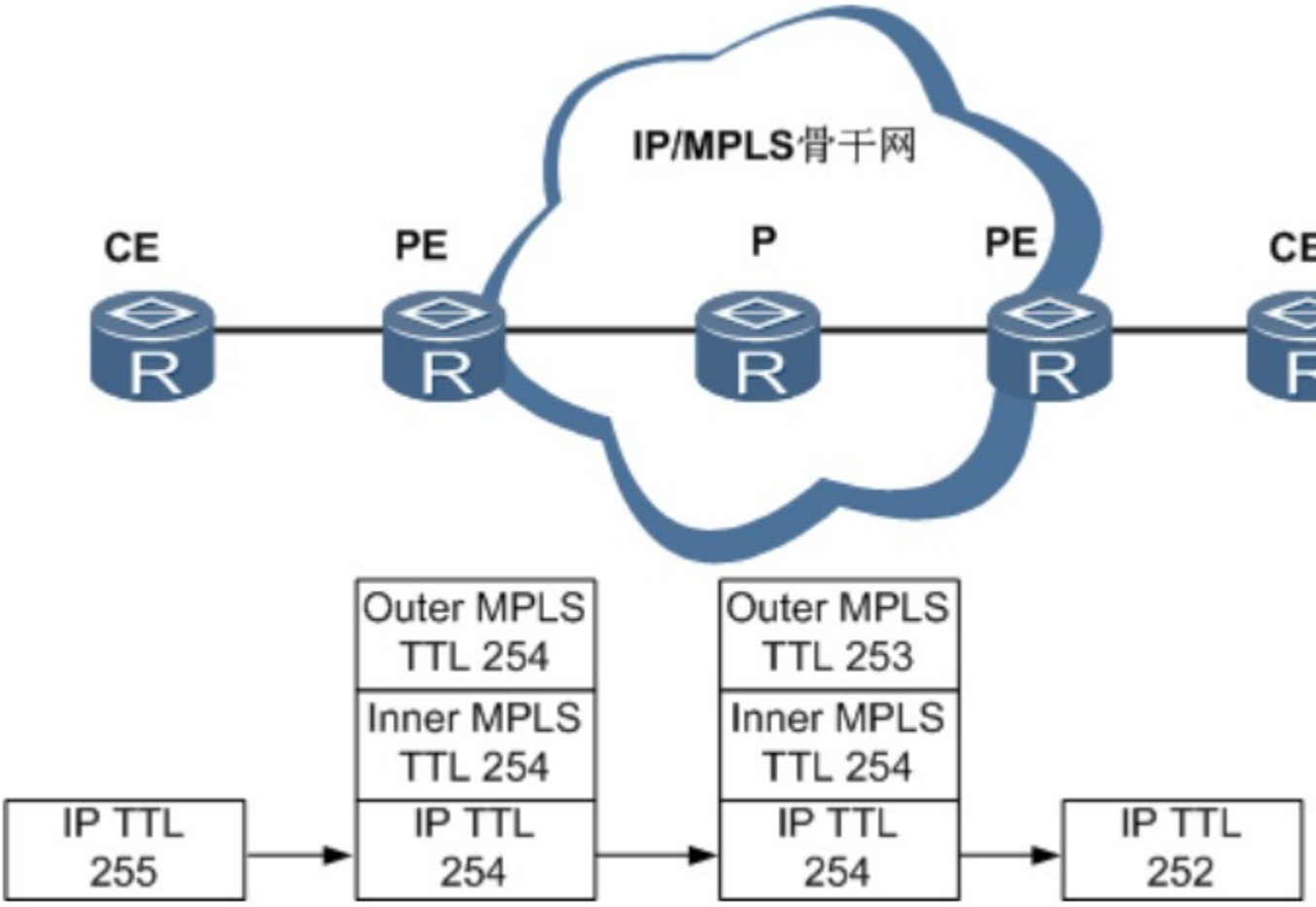
## MPLS 对 TTL 的处理

MPLS 标签中包含一个 8 比特的 TTL 字段，其含义与 IP 头中的 TTL 域相同。MPLS 对 TTL 的处理除了用于防止产生路由环路外，也用于实现 Traceroute 功能。
定义了两种 MPLS 对 TTL 的处理模式：Uniform 和 Pipe。缺省情况下，为 Uniform。

缺省情况下，MPLS 报文中 TTL 传播模式是 Uniform 统一模式
undo ttl propagate，配模式为 Pipe 管道模式。
ttl propagate，　　配置模式为 Uniform 统一模式。

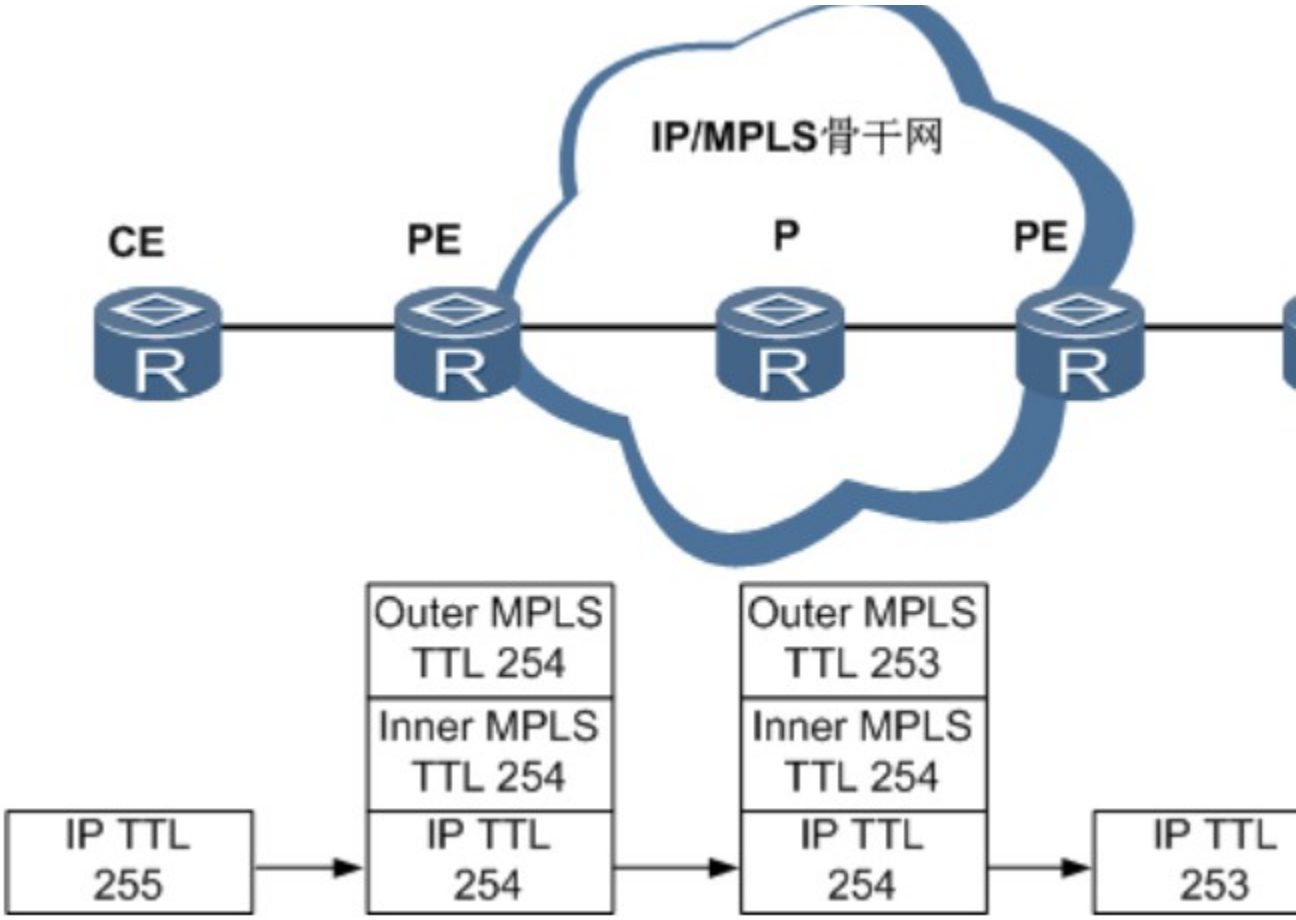
## Uniform 模式

IP 报文经过 MPLS 网络时，在入节点，IP TTL 减 1 映射到 MPLS TTL 字段，此后报文在 MPLS 网络中按照标准的 TTL 处理方式处理。在出节点将 MPLS TTL 减 1 后映射到 IP TTL 字段。




Pipe 模式

在入节点，IP TTL 值减 1，MPLS TTL 字段为固定值，此后报文在 MPLS 网络中按照标准的 TTL 处理方式处理。在出节点会将 IP TTL 字段的值减 1。即 IP 分组经过 MPLS 网络时，无论经过多少跳，IP TTL 只在入节点和出节点分别减 1。

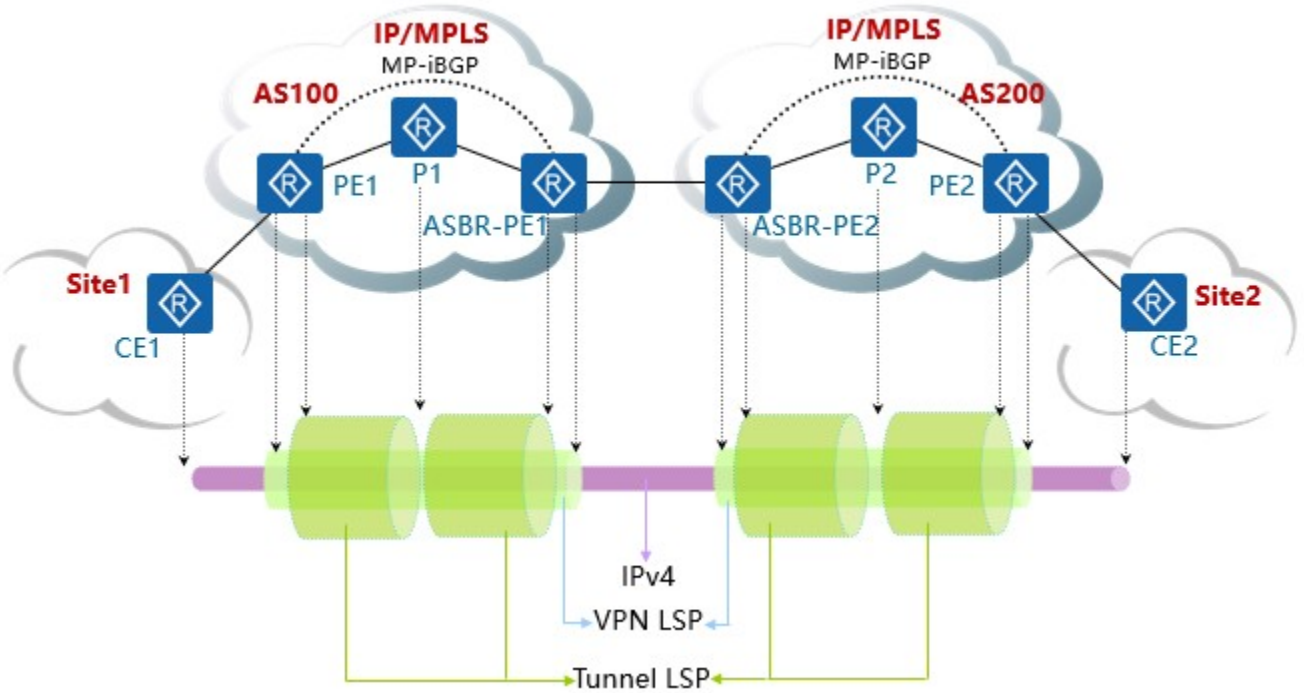



在 MPLS VPN 应用中，出于网络安全的考虑，需要隐藏 MPLS 骨干网络的结构，这种情况下，对于私网报文，Ingress 上使用 Pipe 模式。

## 跨域 VPN-OptionA 方式

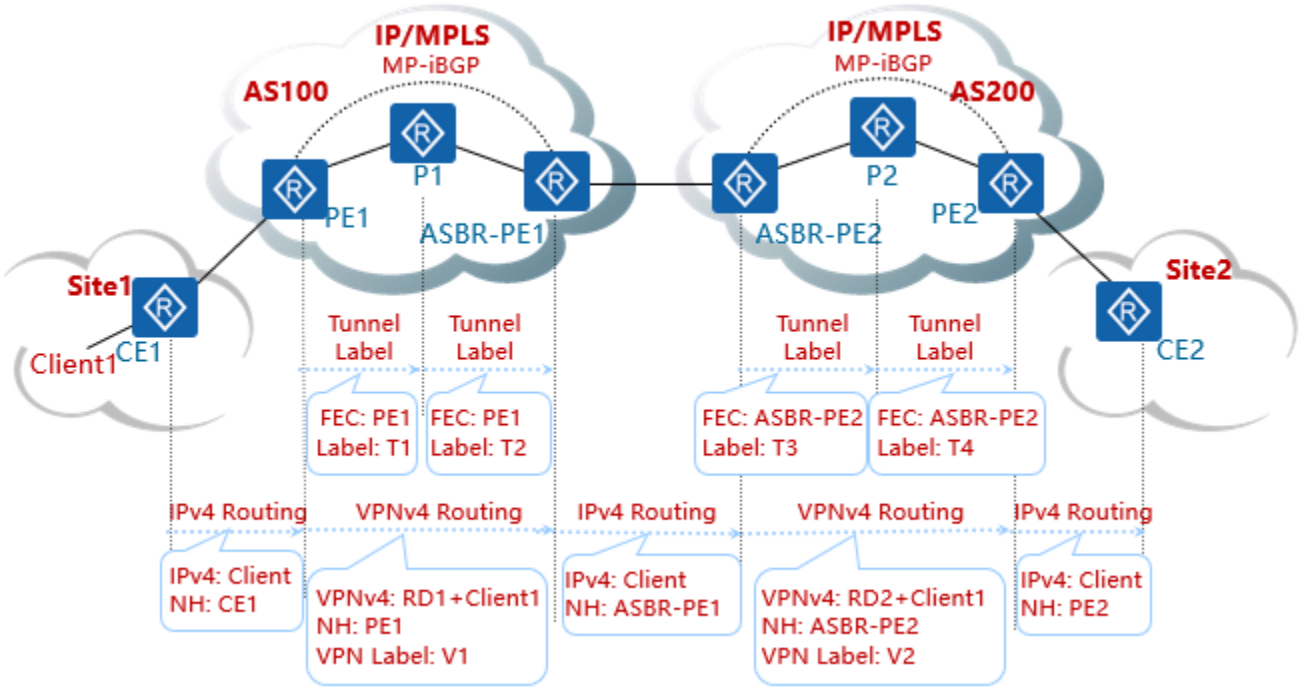
是基本 BGP/MPLS IP VPN 在跨域环境下的应用，ASBR 之间不需要运行 MPLS，也不需要为跨域进行特殊配置。这种方式下，两个 AS 的边界路由器 ASBR 直接相连，ASBR 同时也是各自所在自治系统的 PE。两个 ASBR 都把对端 ASBR 看作自己的 CE 设备，使用 EBGP 方式向对端发布 IPv4 路由。在 ASBR 间通过专用的接口管理自己的 VPN 路由，也称为 VRF-to-

VRF；背对背




在本方案中，ASBR-PE 之间直接相连。两台 ASBR-PE 之间用多个接口(包括子接口)互连，每个接口关联一个 VPN，每个 ASBR-PE 都把对端当成 CE。因此，ASBR-PE 相连的接口(包括子接口)需要绑定 VRF，并通过 eBGP 邻居关系把 VPNv4 路由转变成普通 IPv4 路由从一个 AS 传递到另一个 AS。因此，两个 ASBR 相连，但不需要启用 MPLS。此方案在 MPLS BGP VPN 业务属性上没有做扩展。

## OptionA 方式 - 控制平面




我们只通过单方向来解释控制平面的工作过程，同时假设在站点 Site1 有主机 Client1，如上图，现在需要把 Client1 这条路由从 CE1 穿过 AS100 和 AS200 传递到 CE2：

在 AS100 中，通过运行 LDP 协议，PE1 分配一个与去往 PE1 的路由相关联的隧道标签(外层标签)T1 给 P1。

在 AS100 中，通过运行 LDP 协议，P1 分配一个与去往 PE1 的路由相关联的隧道标签(外层标签)T2 给 ASBR-PE1。

在 AS200 中，同样通过运行 LDP 协议，ASBR-PE2 分配一个与去往 ASBR-PE2 的路由相关联的隧道标签(外层标签)T3 给 P2。

在 AS200 中，通过运行 LDP 协议，P2 分配一个与去往 ASBR-PE2 的路由相关联的隧道标签(外层标签)T4 给 PE2。

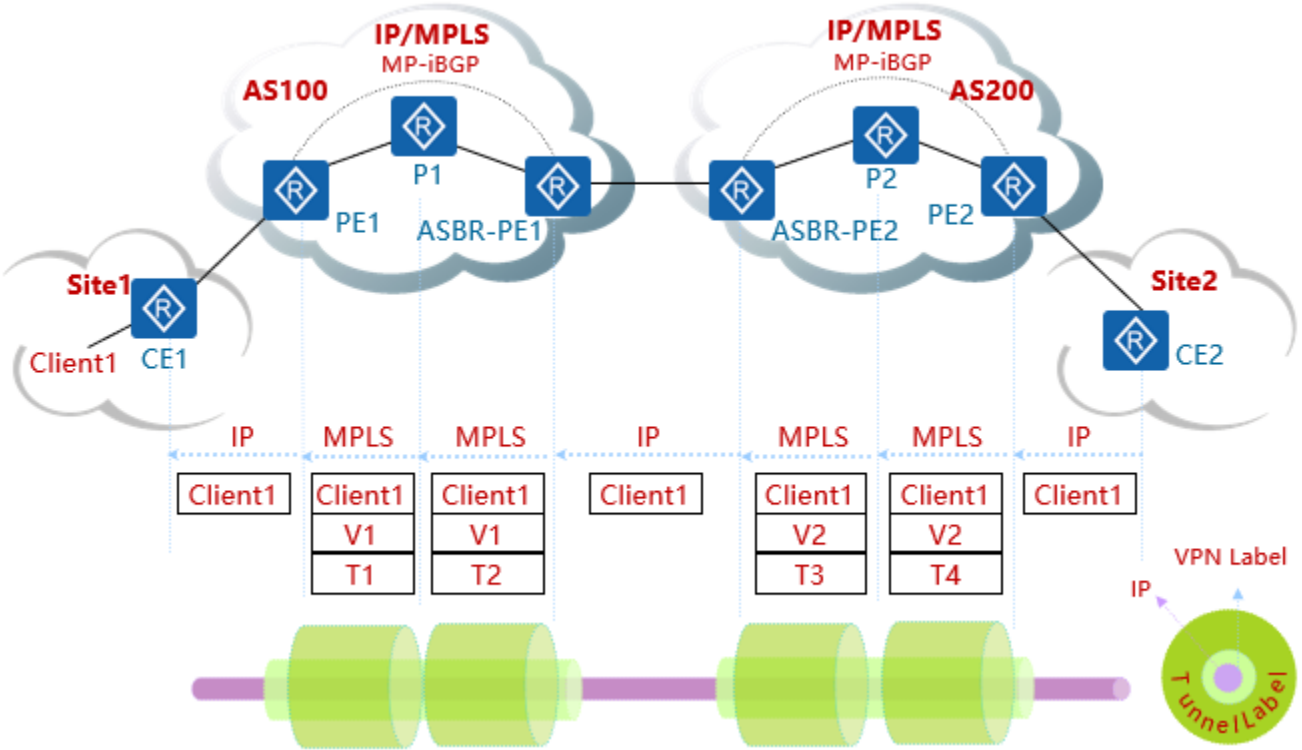
CE1 通告路由 Client1 给 PE1，路由的下一跳为 CE1 的接口地址。

PE1 将 Client1 的 IPv4 路由，封装为 VPNv4 路由、通过 MP-BGP，并且下一跳改为 PE1，分配一个 VPN 标签 V1，通告给 ASBR-PE1。

ASBR-PE1 将 VPNv4 路由变为 IPv4 路由，把 IPv4 路由 Clien

t1 通告给 ASBR-PE2，并且下一跳指向 ASBR-PE1。
ASBR-PE2 将 IPv4 路由 Client1 通过 MP-BGP 重发布为 VPN v4 路由，并且下一跳为 ASBR-PE2，为该路由分配一个 VPN 标签 V2，将其通告给 PE2。
PE2 将 VPNv4 路由转变为 IPv4 路由 Client1，把路由 Client1 通告给 CE2，并且下一跳指向 PE2。

## OptionA 方式 - 转发平面




从反向来分析转发平面的工作过程，即 CE2 要发送一个目的地为 Client1 的 IP 报文给 CE1，如上图所示：
CE2 发送一个目的地为 Client1 的 IP 报文给 PE2。
PE2 收到 IP 报文后进行 MPLS 标签的封装，先封装 VPN 标签 V2，再封装外层标签 T4，然后将此报文发送给 P2。
P2 进行标签交换，把外层标签 T4 换成 T3，然后将此报文发送给 ASBR-PE2。

ASBR-PE2 去掉所有标签，将报文(普通 IP 报文)转发给 ASBR-PE1。
ASBR-PE1 收到 IP 报文后进行 MPLS 标签的封装，先封装 VPN 标签 V1，再封装外层标签 T2，然后将此报文发送给 P1。
P1 进行标签交换，把外层标签 T2 换成 T1，然后将此报文发送给 PE1。
PE1 收到后去掉所有标签，将报文(普通 IP 报文)转发给 CE1。

## 跨域 VPN-OptionA 方式的特点

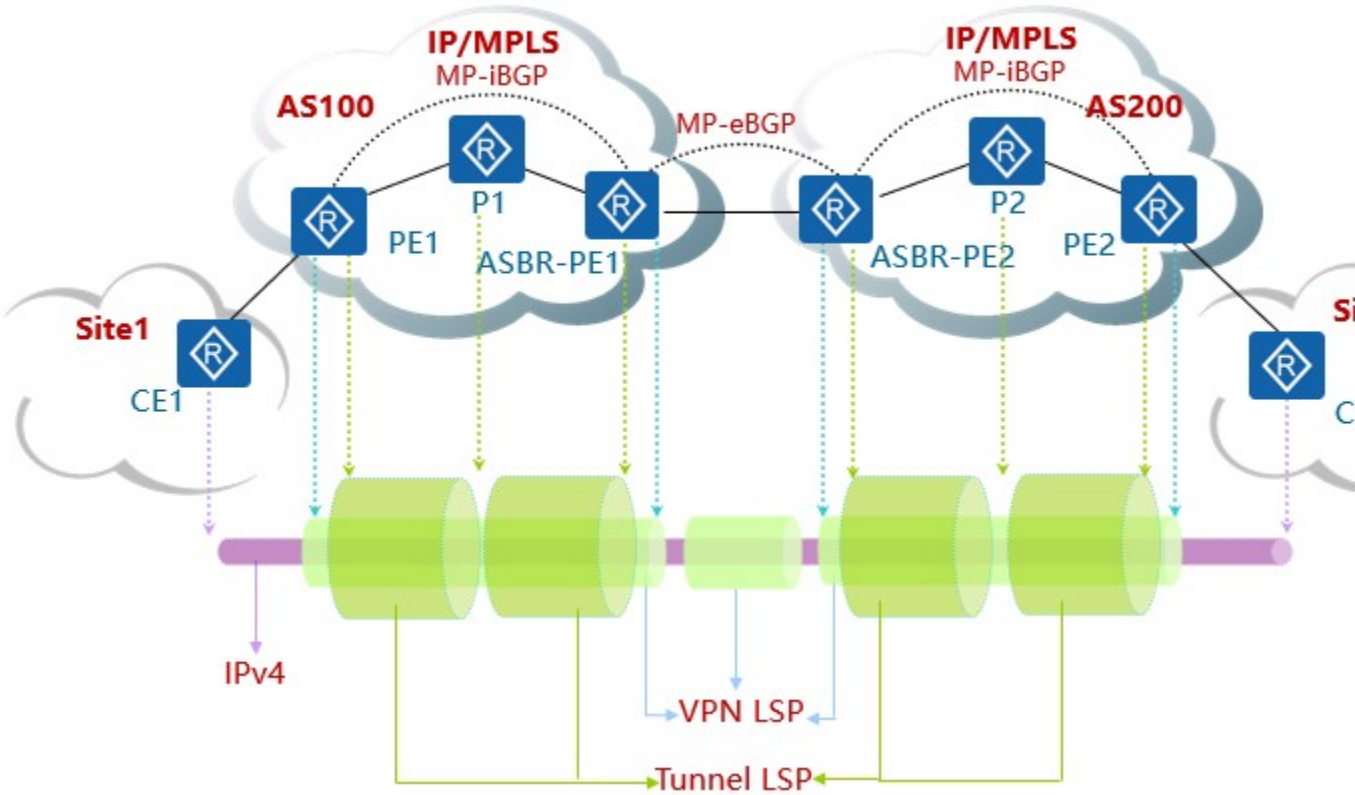
优点是配置简单：由于 ASBR 之间不需要运行 MPLS，也不需要为跨域进行特殊配置。
缺点是可扩展性差：由于 ASBR 需要管理所有 VPN 路由，为每个 VPN 创建 VPN 实例。这将导致 ASBR 上的 VPN-IPv4 路由数量过大。并且，由于 ASBR 间是普通的 IP 转发，要求为每个跨域的 VPN 使用不同的接口，从而提高了对 PE 设备的要求。如果跨越多个自治域，中间域必须支持 VPN 业务，不仅配置量大，而且对中间域影响大。在需要跨域的 VPN 数量比较少的情况，可以优先考虑使用。

=========

## OptionB 方式

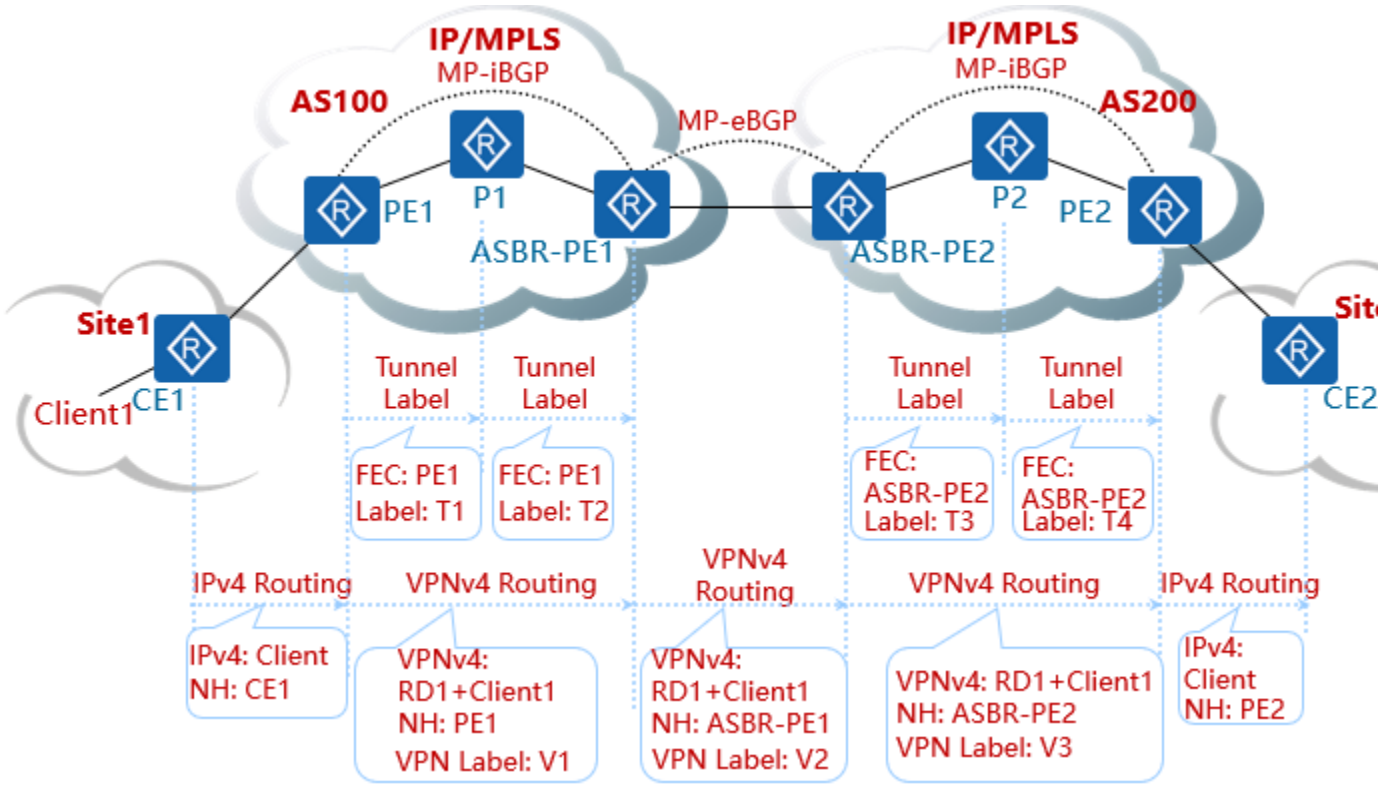
两个 ASBR 通过 MP-EBGP 交换它们从各自 AS 的 PE 设备接收的标签 VPN-IPv4 路由。
跨域 VPN-OptionB 方案中，ASBR 接收本域内和域外传过来的所有跨域 VPN-IPv4 路由，再把 VPN-IPv4 路由发布出去。但 MPLS VPN 的基本实现中，PE 上只保存与本地 VPN 实例的 VPN Target 相匹配的 VPN 路由。因此，可以在 ASBR 上配置不做 RT 过滤来传递路由，因此无需在 ASBR 创建 VPN 实例，无需绑定任何接口。可以在网络中叠加部署 RR 设备，专门负责客户侧 VPN 路由的传递。

bgp 100
ipv4-family vpnv4
undo policy vpn-target




PE 通过 MP-iBGP 将 VPNv4 路由通告给 ASBR-PE 或是 VPN RR(其中 ASBR-PE 是其客户机)。ASBR-PE 再通过 MP-eBGP 将 VPNv4 通告给另一个 AS 的 ASBR-PE，再由这个 ASBR-PE 将 VPNv4 路由通告给该 AS 内的 PE。

控制平面 (无 RR 场景)




我们只通过单方向来解释控制平面的工作过程，同时假设在站点 Site1 有一 VPN 路由 Client1 连接：

1、在 AS100 中，通过运行 LDP 协议，PE1 分配一个与去往 PE1 的路由相关联的隧道标签(外层标签)T1 给 P1

2、在 AS100 中，通过运行 LDP 协议，P1 分配一个与去往 PE1 的路由相关联的隧道标签(外层标签)T2 给 ASBR-PE1

3、在 AS200 中，同样通过运行 LDP 协议，ASBR-PE2 分配一个与去往 ASBR-PE2 的路由相关联的隧道标签(外层标签)T3 给 P2

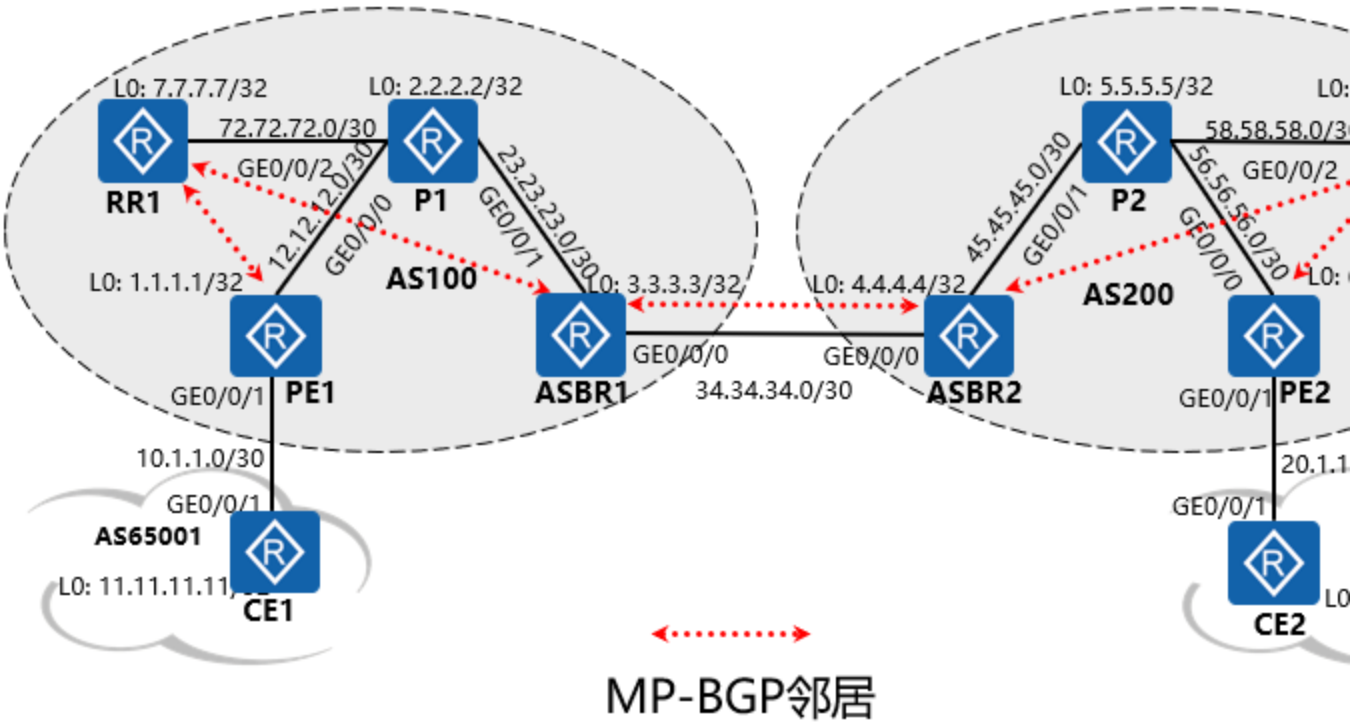
4、在 AS200 中，通过运行 LDP 协议，P2 分配一个与去往 ASBR-PE2 的路由相关联的隧道标签(外层标签)T4 给 PE2

5、CE1 通告路由 Client1 给 PE1，路由的下一跳为 CE1 的接口地址

6、PE1 将 Client1 的 IPv4 路由，封装为 VPNv4 路由、通过 MP-BGP，并且下一跳改为 PE1，分配一个 VPN 标签 V1，然后通告给 ASBR-PE1

7、ASBR-PE1 通过 MP-EBGP 将 Client1 的 VPNv4 路由通告给 ASBR-PE2，将下一跳改为 ASBR-PE1，并重新分配一个 VPN 标签 V2
8、ASBR-PE2 将收到的 Client1 的 VPNv4 路由通过 MP-IBGP 通告给 PE2，将下一跳指向自己，并重新分配一个 VPN 标签 V3
9、PE2 将 Client1 的 VPNv4 路由变为 IPv4 路由，把路由 Client1 通告给 CE2，并且下一跳改为 PE2

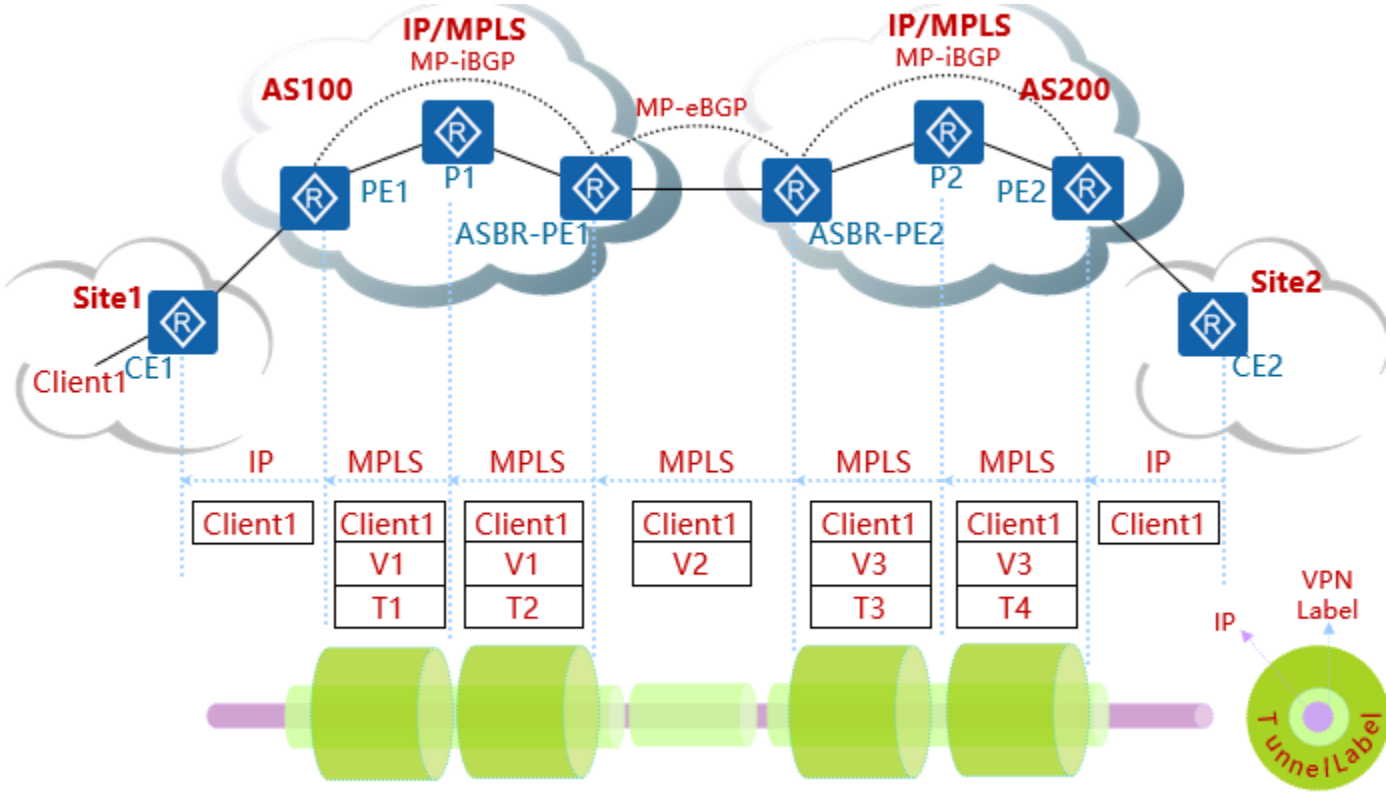
## 控制平面 (带 RR 场景)




当 VPN 实例数量较多时，可以部署专门的 RR 设备。如图，AS 内的 PE 和 ASBR 设备只与 RR 设备建立 MP-BGP 邻居关系，由 RR 负责路由的反射传递，PE 和 ASBR 之间无需建立 BGP 邻居。
RR 只负责控制平面的 VPNv4 路由传递，数据转发时，流量

不经过 RR。

## 转发平面




从反向来分析转发平面的工作过程：
CE2 发送一个目的地为 Client1 的 IP 报文给 PE2。
PE2 收到 IP 报文后进行 MPLS 标签的封装，先封装 VPN 标签 V3，再封装外层标签 T4，然后将此报文发送给 P2。
P2 进行标签交换，把外层标签 T4 换成 T3，然后将此报文发送给 ASBR-PE2。
ASBR-PE2 去掉外层标签，将 VPN 标签 V3 交换为 V2，再将其转发给 ASBR-PE1（此时报文仅带有一层私网标签）。
ASBR-PE 交换 VPN 标签 V2 成 V1，再加一个外层标签 T2，并将报文转发给 P1。
P1 进行标签交换，把外层标签 T2 换成 T1，然后将此报文发送给 PE1。
PE1 收到后去掉所有标签，将报文(普通 IP 报文)转发给 CE1。

## 跨域 VPN-OptionB 方式的特点

不同于 OptionA，OptionB 方案不受 ASBR 之间互连链路数目的限制。

局限性：VPN 的路由信息是通过 AS 之间的 ASBR 来保存和扩散的，当 VPN 路由较多时，ASBR 负担重，容易成为故障点。因此在 MP-EBGP 方案中，需要维护 VPN 路由信息的 ASBR 一般不再负责公网 IP 转发。

=========

## OptionC 方式 方案一

底层标签是由对端 PE 分配的与 VPN 路由相关联的 VPN 标签，中间的标签是 ASBR 分配的与去往对端 PE 的路由相关联的标签，

外层标签则是与去往下一跳 ASBR 的路由相关联的标签。

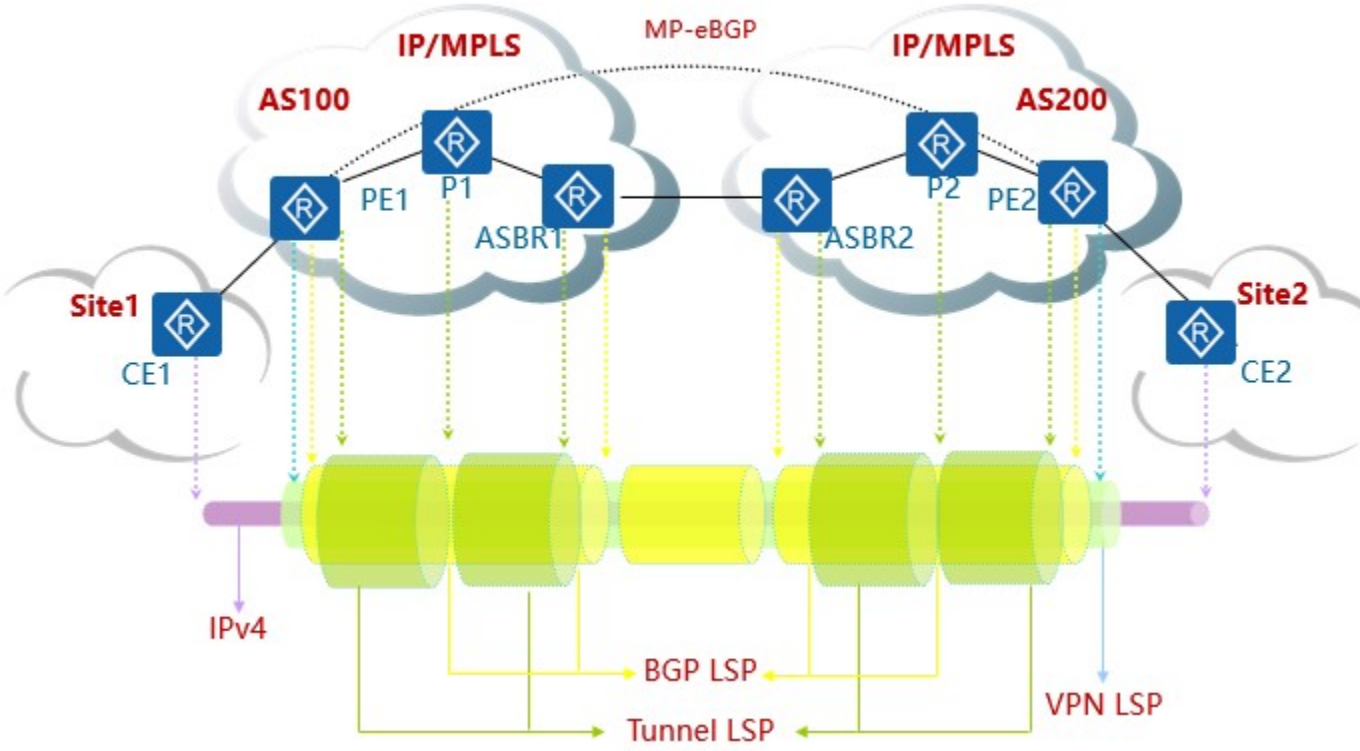
Option C 在 Option B 的基础上进一步对 BGP 协议进行了扩展，并且将其应用在了 AS 边界，从而实现了 MBGP 部署与传统 BGPv4 部署的“完全分离”，域间 PE 之间（或者 RR 之间）建立 Multi-MPEBGP

邻居关系传递私网路由，ASBR 之间建立“普通”的 EBGP 邻居关系，ASBR 不需要同时维护 BGP IPv4 路由和 VPNv4 路由，显著降低设备性能的影响，从而更适合 VPN 数量较多的应用场景。

此外，由于 VPN 的路由信息只出现在 PE 设备上，而 ASBR 设备只负责报文的转发，这样就使中间域的设备可以不支持 MPLS VPN 业务，就是一个普通的支持 IP 转发路由器。在充当以上跨域的前提下，还可以同时支持普通的 IP 业务，尤其是在跨越多个域时优势更为明显，而且这个方案更适合支持

MPLS VPN 的负载分担等功能，ASBR 不会再成为网络中的性能瓶颈。不过由于这种解决方案需要对普通的 BGP 做扩展，且隧道的生成也是有别于普通的 MPLS VPN 结构，因此维护和理解起来难度较大。

ASBR 通过 MP-IBGP 向各自 AS 内的 PE 设备发布标签 IPv4 路由，并将到达本 AS 内 PE 的标签 IPv4 路由通告给它在对端 AS 的 ASBR 对等体，过渡自治系统中的 ASBR 也通告带标签的 IPv4 路由。这样，在入口 PE 和出口 PE 之间建立一条 BGP LSP。
不同 AS 的 PE 之间建立 Multihop 方式的 EBGP 连接，交换 VPNv4 路由。
ASBR 上不保存 VPN-IPv4 路由，相互之间也不通告 VPNv4 路由。
当网络规模较大时，可以在方案中部署 RR 设备，专门负责用户侧路由的传递。即，PE 与 RR 建立 MP-IBGP 邻居，RR1 与 RR2 建立 MP-EBGP 邻居，路由传递为 PE1-RR1-RR2-PE2，PE 之间无需直接建立 BGP 邻居关系，当 VPN 数量较多时，引入 RR 的方式可以减轻 PE 的工作负担。



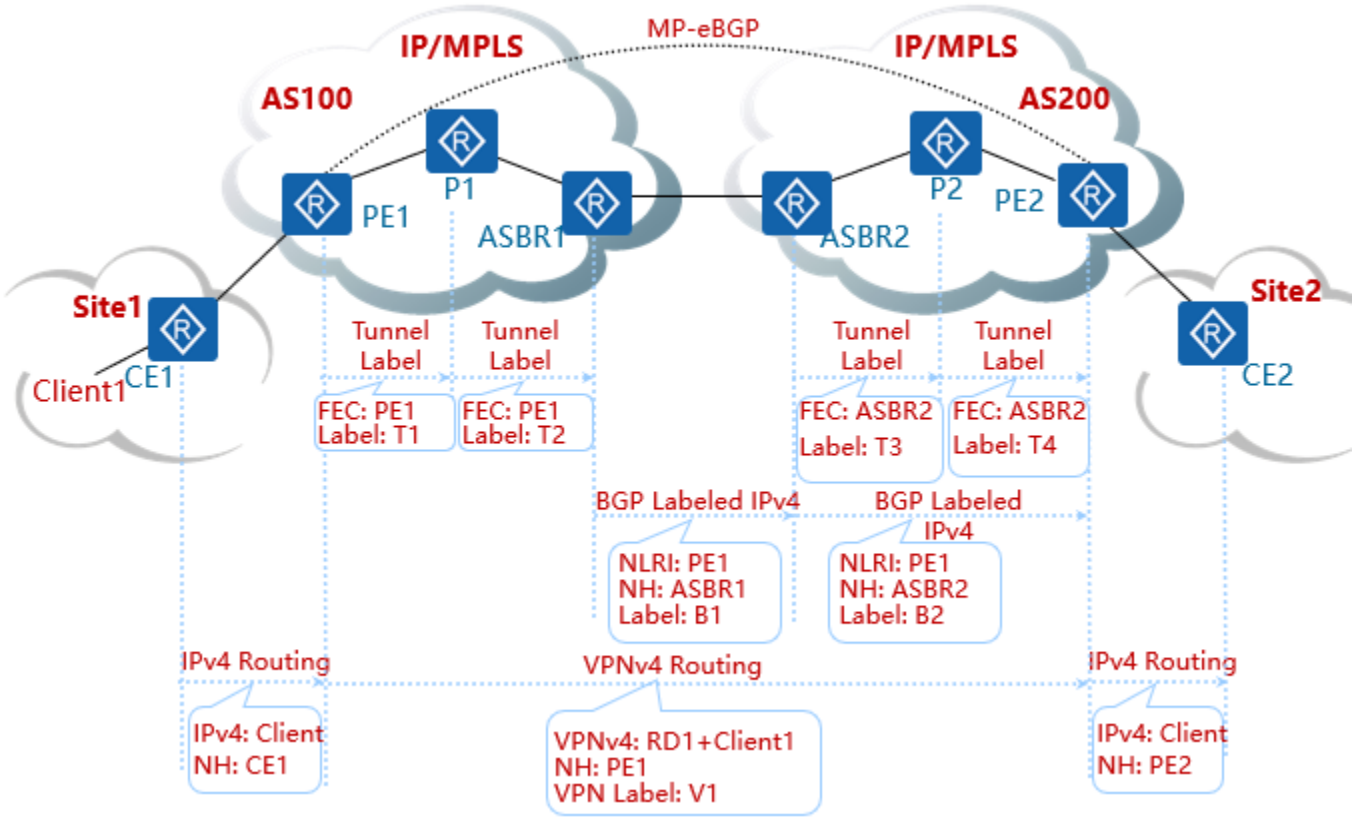

ASBR 不再维护或是通告 VPNv4 路由（所以如上图，此处将 ASBR-PE 路由器改名为 ASBR）。ASBR 只需要维护所有去往 PE 的带标签路由，并通过 EBGP 通告给对端 AS。在 transit AS 内的 ASBR 也同样需要使用 EBGP 通告这些带标签的 IPv4 路由。这样在不同 AS 的 PE 之间给会建立一条 LSP，从而可以建立起 PE 之间的多跳 MP-EBGP 连接并进行 VPNv4 路由的通告。

如果每个 AS 的 P 路由器都能够知道去往其他 AS 的 PE 路由器的路由，那情况会比较简单。但是如果 P 不知道，那么当 PE 收到从 CE 收到 VPN 数据时，就要加上三层标签，底层标签是由对端 PE 分配的与 VPN 路由相关联的 VPN 标签，中间的标签是 ASBR 分配的与去往对端 PE 的路由相关联的标签，外层标签则是与去往下一跳 ASBR 的路由相关联的标签。

为了进一步扩展性能，多跳 MP-EBGP 会话可以建立在不同的 AS 的 VPN RR 之间。并且当这些 VPN RR 通告 VPNv4 路由时不改变下一跳信息。PE 只与 VPN RR 建立 MP-iBGP 会

话。
注意：为了方便，如上图，使用的是对称的 LSP 进行示意，但是实际上在控制平面和数据平面的工作过程上，两端 AS 的 LSP 结构是不对称的。

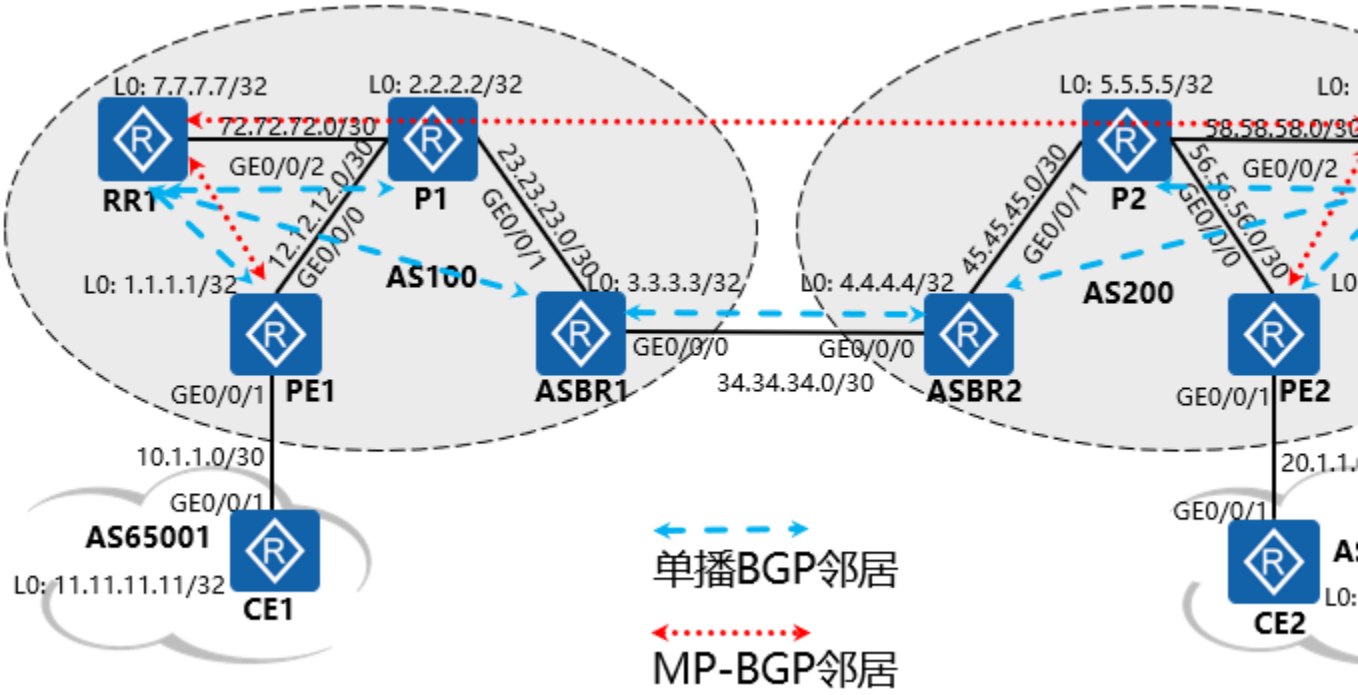
## 控制平面 (无 RR 场景)




通过单方向来解释控制平面的工作过程，同时假设在站点 Site 1 有一 VPN 路由 Client1 连接，并且 P1 与 P2 路由器都没有去往另一个 AS 的 PE 的路由，以上图为例：
在 AS100 中，通过运行 LDP 协议，PE1 分配一个与去往 PE1 的路由相关联的隧道标签(外层标签)T1 给 P1
在 AS100 中，通过运行 LDP 协议，P1 分配一个与去往 PE1 的路由相关联的隧道标签(外层标签)T2 给 ASBR-PE1

在 AS200 中，同样通过运行 LDP 协议，ASBR-PE2 分配一个与去往 ASBR-PE2 的路由相关联的隧道标签(外层标签)T3 给 P2
在 AS200 中，通过运行 LDP 协议，P2 分配一个与去往 ASBR-PE2 的路由相关联的隧道标签(外层标签)T4 给 PE2
ASBR1 通过 EBGP 会话通告一条去往 PE1 的带标签的 IPv4 路由给 ASBR2，其中下一跳为 ASBR1，标签为 BGP 标签，值为 B1
ASBR2 通过 BGP 会话通告一条去往 PE1 的带标签的 IPv4 路由给 PE2，其中下一跳为 ASBR2，标签为 BGP 标签，值为 B2。
注意：这里假设 PE2 与 ASBR1 所在的 AS 已经为去往它们的路由分配了隧道标签(公网标签)，并且去往 PE2 的带标签路由也已经被通告。
PE1 与 PE2 建立起 MP-EBGP 会话
CE1 通告路由 Client1 给 PE1，路由的下一跳为 CE1 的接口地址
PE1 将 Client1 的 IPv4 路由，封装为 VPNv4 路由、通过 MP-BGP，并且下一跳改为 PE1，分配一个 VPN 标签 V1，将其通告给 PE2
PE2 将 VPNv4 路由变为 IPv4 路由，把 IPv4 路由 Client1 通告给 CE2，并且下一跳改为 PE2

## 控制平面 (带 RR 场景)




VPNv4 邻居：

本端 PE 只与本端 RR 建立 VPNv4 邻居，本端 RR 与对端 RR 建立 VPNv4 邻居，实现了跨域 VPN 路由的传递。

ASBR,P,PE 同 RR 建立 BGP 单播 IPv4 邻居：

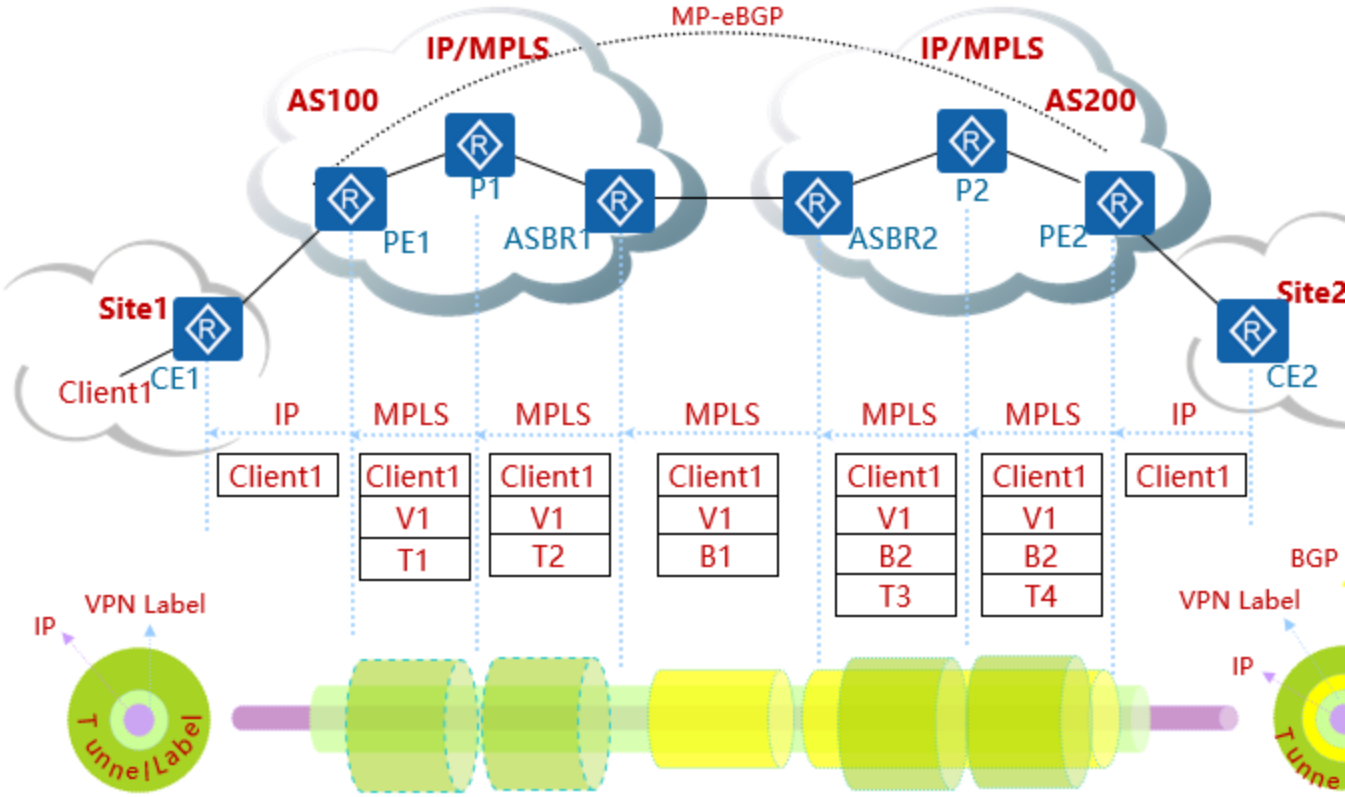
ASBR 通过 ipv4 邻居学习将从对端 ASBR 学到的 RR 的 loopback，传递给本端 RR，用于本端 RR 与对端 RR 建立 vpnv4 邻居。

ASBR 通过 ipv4 邻居学习将从对端 ASBR 学到的 RR 和 PE 的 loopback，传递给本端 RR，本端 RR 再将其反射给 P，用于跨域 bgp 路由的递归查询。

ASBR 通过 ipv4 邻居学习将从对端 ASBR 学到的 RR 和 PE 的 loopback，传递给本端 RR，本端 RR 再将其反射给 PE，用于跨域之间的 PE 建立 BGP LSP。

带 RR 场景中，RR 负责控制平面 IPv4 的路由反射、VPNv4 路由的传递，转发平面的流量不经过 RR。

转发平面




从反向来分析转发平面的工作过程：

CE2 发送一个目的地为 Client1 的 IP 报文给 PE2。

PE2 收到 IP 报文后进行 MPLS 标签的封装，先封装 VPN 标签 V1，由于去往 Client1 的下一跳 PE1 不是直连邻居，通过查表发现去往 PE1 的 BGP 路由是带标签的路由，因此加上分配的 BGP 标签 B2 做为中间标签，最后，由于去往 PE1 的路由的下一跳 ASBR2 也不是直连邻居，通过查表发现去往 ASBR2 也有关联的标签 T4，因此，封装上外层标签 T4。

P2 进行标签交换，把外层标签 T4 换成 T3，然后将此报文发送给 ASBR-PE2。

ASBR2 去掉外层标签，将 BGP 标签 B2 交换为 B1，再将其转发给 ASBR1。

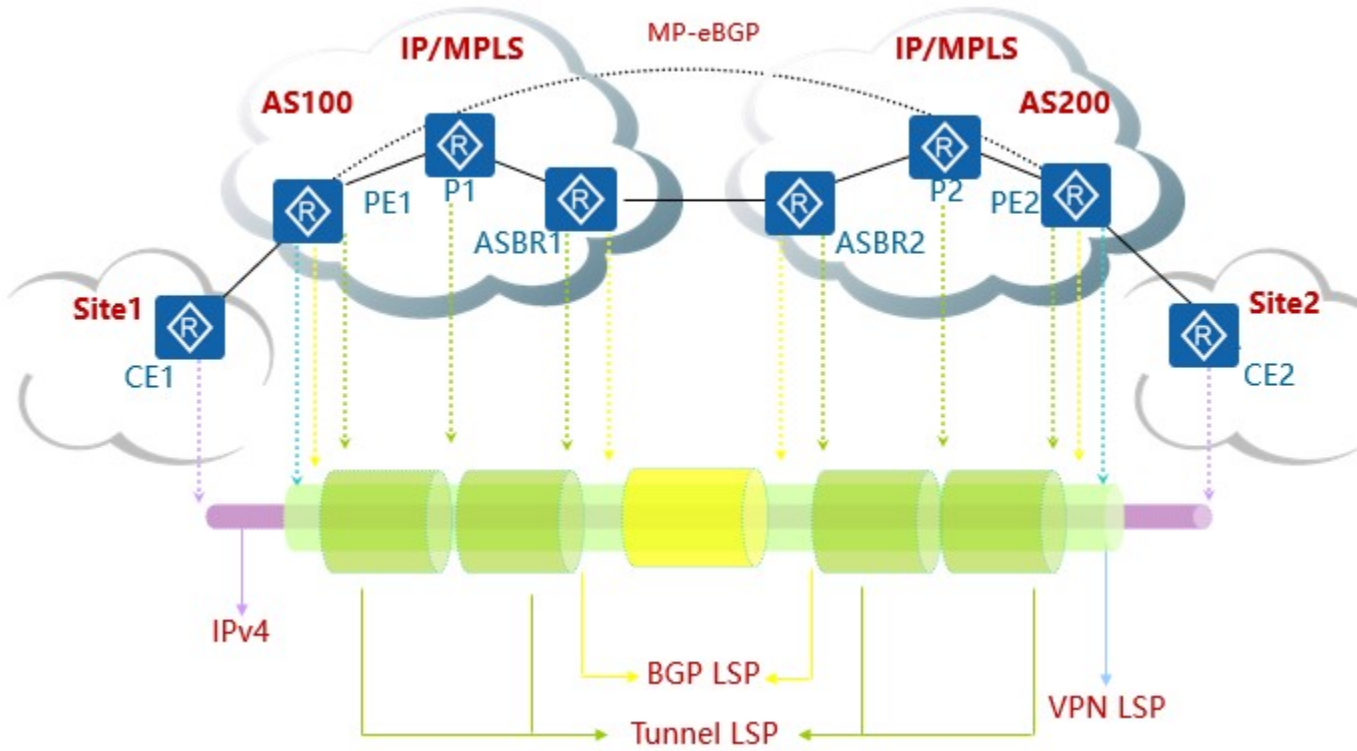
当 ASBR1 收到报文后，发现 B1 是它分配的，所以去掉 B1 进一步查表转发，发现此时去往 PE1 的路由有一个关联的标签

T2，因此，ASBR1 将其加在栈顶，并转发给 P1。
P1 进行标签交换，把外层标签 T2 换成 T1，然后将此报文发送给 PE1。
PE1 收到后去掉所有标签，将报文(普通 IP 报文)转发给 CE1。

=========

## OptionC 方式 方案二

跨域 VPN-OptionC 方案二与方案一大体相似。
不同之处在于，方案一中，需要使用三层标签，即 VPN label，BGP LSP， Tunnel LSP 来承载流量，而方案二只需要两层。
方案一，ASBR 在收到对端 ASBR 发来的 BGP 标签路由后，需要配置策略产生一个新的标签并发布给 AS 内的 PE 或者 RR 设备，以建立一条完整的 BGP LSP。方案二中，ASBR 需要配置 MPLS 触发为 BGP 标签路由分发标签，因此在 AS 内的 PE 上可以看到去往对端 PE 的 LDP LSP，而非 BGP LSP。
同理，方案二支持 RR 设备的部署。

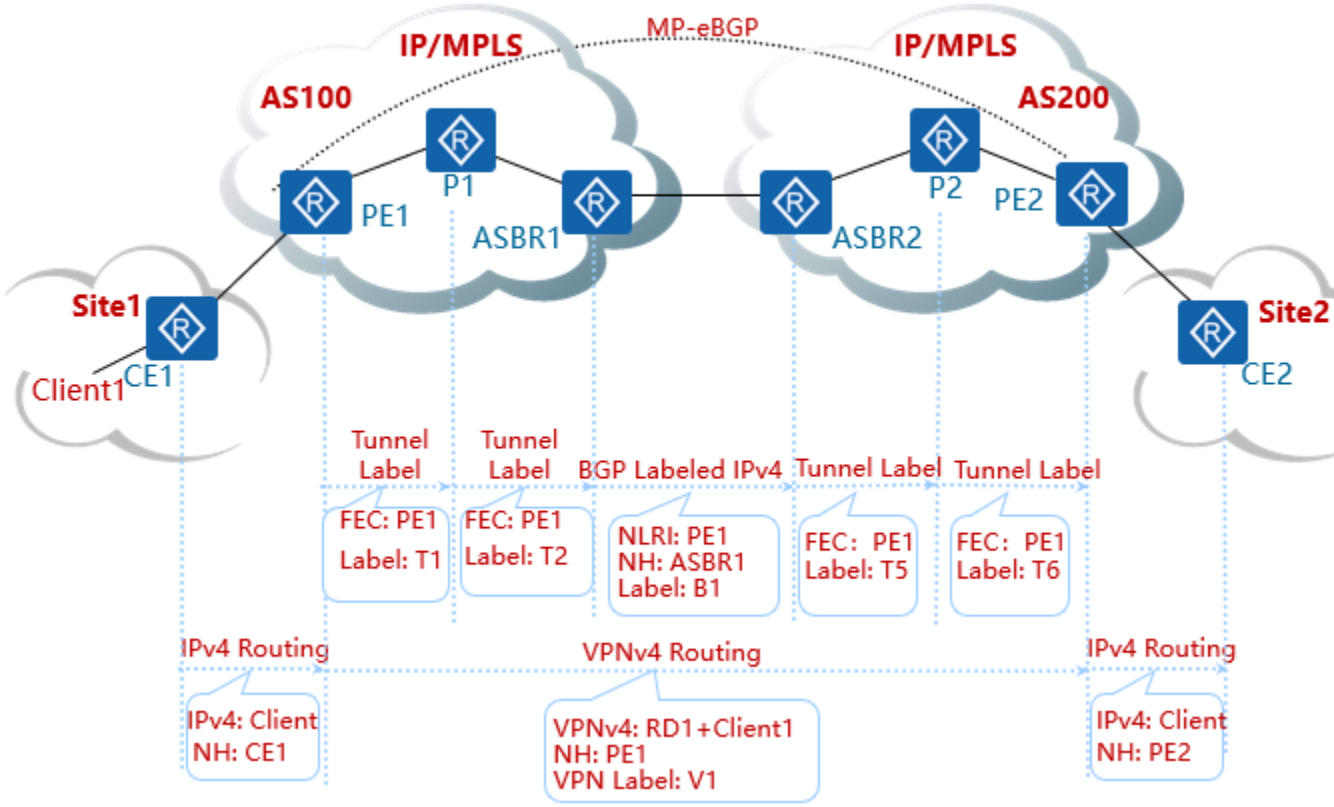



在此方案中，ASBR 不再维护或是通告 VPNv4 路由。ASBR 只需要维护所有去往 PE 的带标签路由，并通过 EBGP 通告给对端 ASBR。

本端 ASBR 收到带 BGP 标签路由后，MPLS LDP 会触发为该 BGP 标签路由产生标签，并在 AS 内的 LDP 邻居间传递。因此，在 PE 上可以看到去往对端 PE 的 LDP LSP。

为了进一步扩展性能，多跳 MP-EBGP 会话可以建立在不同的 AS 的 VPN RR 之间，本 AS 内的 PE 只需要与 RR 建立 MP-IBGP 即可。这些 VPN RR 通告 VPNv4 路由时不改变下一跳信息，进而当对端 PE 转发流量时，可以迭代至正确的隧道。

## 控制平面 (无 RR 场景)




通过单方向来解释控制平面的工作过程，同时假设在站点 Site 1 有一 VPN 路由 Client1 连接，并且 P1 与 P2 路由器都没有去往另一个 AS 的 PE 的路由，以上图为例：

在 AS100 中，通过运行 LDP 协议，PE1 分配一个与去往 PE1 的路由相关联的隧道标签(外层标签)T1 给 P1。

在 AS100 中，通过运行 LDP 协议，P1 分配一个与去往 PE1 的路由相关联的隧道标签(外层标签)T2 给 ASBR1。

在 AS200 中，同样通过运行 LDP 协议，ASBR2 分配一个与去往 ASBR2 的路由相关联的隧道标签(外层标签)T3 给 P2。

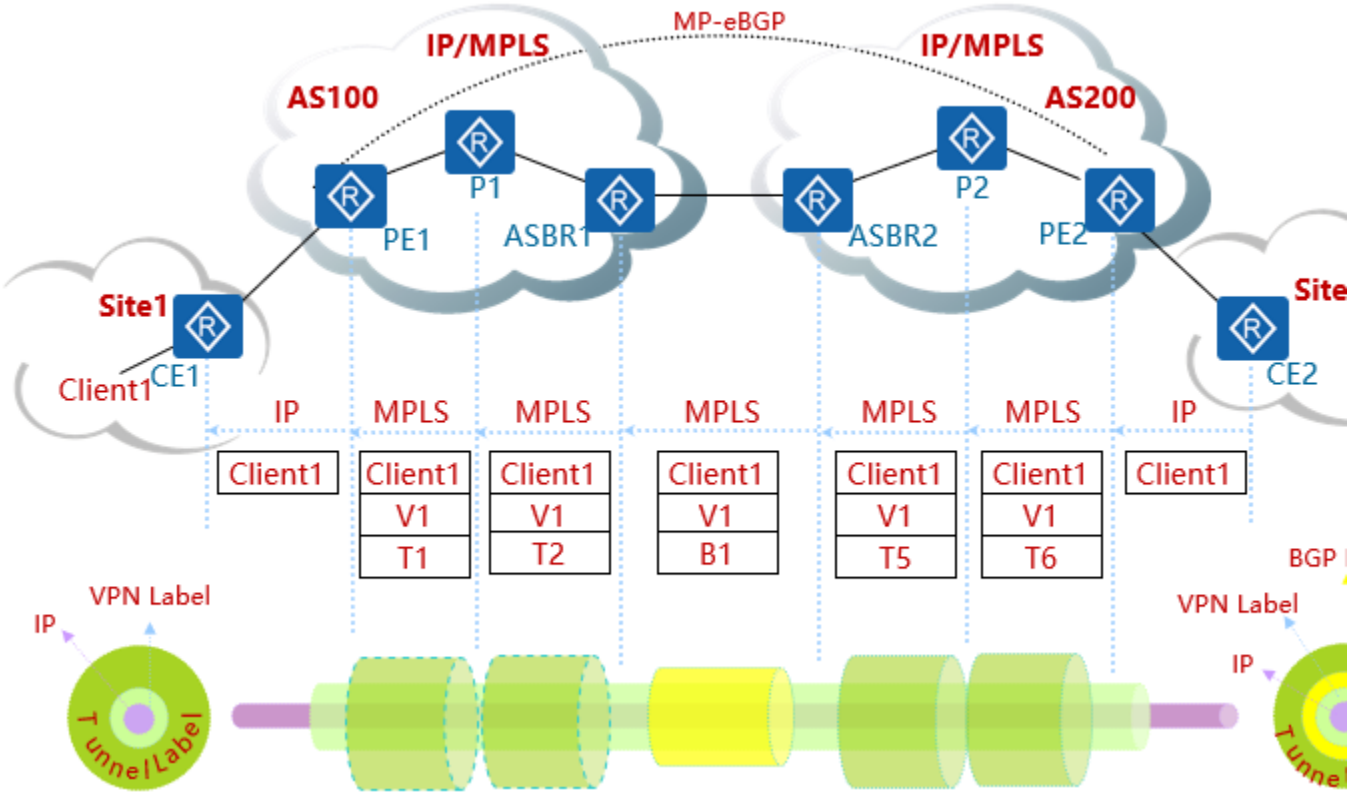
在 AS200 中，通过运行 LDP 协议，P2 分配一个与去往 ASBR2 的路由相关联的隧道标签(外层标签)T4 给 PE2。

ASBR1 通过 EBGP 会话通告一条去往 PE1 的带标签的 IPv4 路由给 ASBR2，其中下一跳为 ASBR1，标签为 BGP 标签，值为 B1。

ASBR2 为这条 BGP 标签路由触发建立 LSP，分发 LDP 标签

T5 至 P2，P2 进而分发 T6 至 PE2。
PE1 与 PE2 建立起 MP-EBGP 会话。
CE1 通告路由 Client1 给 PE1，路由的下一跳为 CE1 的接口地址。
PE1 将 Client1 的 IPv4 路由，封装为 VPNv4 路由、通过 MP-BGP，并且下一跳改为 PE1，分配一个 VPN 标签 V1，将其通告给 PE2。
PE2 将 VPNv4 路由变为 IPv4 路由，把 IPv4 路由 Client1 通告给 CE2，并且下一跳改为 PE2。

## 转发平面



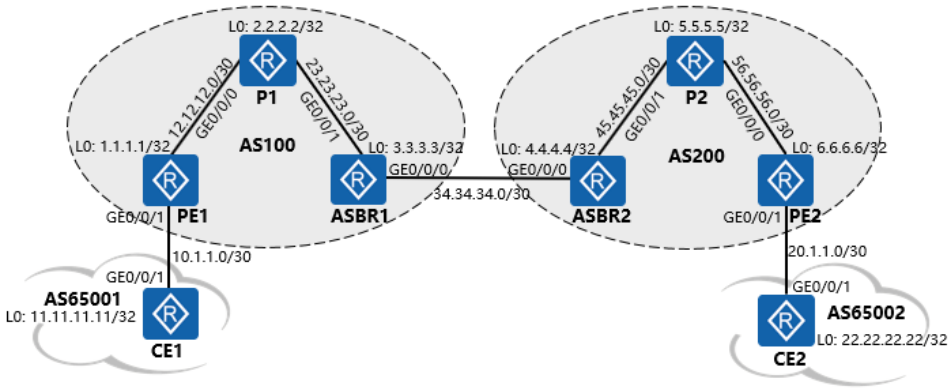

从反向来分析转发平面的工作过程：
CE2 发送一个目的地为 Client1 的 IP 报文给 PE2。
PE2 收到 IP 报文后进行 MPLS 标签的封装，先封装 VPN 标签 V1，由于去往 Client1 的下一跳 PE1 不是直连邻居，通过

查表发现去往 PE1 的标签为 T6，打上 T6。
P2 进行标签交换，把外层标签 T6 换成 T5，然后将此报文发送给 ASBR2。
ASBR2 去掉外层标签，将 T5 交换为 B1，再将其转发给 ASBR1。
当 ASBR1 收到报文后，发现 B1 是它分配的，所以去掉 B1 进一步查表转发，发现此时去往 PE1 的路由有一个关联的标签 T2，因此，ASBR1 将其加在栈顶，并转发给 P1。
P1 进行标签交换，把外层标签 T2 换成 T1，然后将此报文发送给 PE1。
PE1 收到后去掉所有标签，将报文(普通 IP 报文)转发给 CE1。

## 跨域 VPN-OptionC 方式的特点

VPN 路由在入口 PE 和出口 PE 之间直接交换，不需要中间设备的保存和转发。
VPN 的路由信息只出现在 PE 设备上，而 P 和 ASBR 只负责报文的转发，使得中间域的设备可以不支持 MPLS VPN 业务，只需支持 MPLS 转发，ASBR 设备不再成为性能瓶颈。因此跨域 VPN-OptionC 更适合在跨越多个 AS 时使用。
更适合支持 MPLS VPN 的负载分担。
缺点是维护一条端到端的 PE 连接管理代价较大。

# 跨域VPN-OptionA方式拓扑介绍




- 配置任务：使用跨域VPN-OptionA的方式实现两端的CE设备互访。注：本实例使用OSPF作为IGP。

- 如上图所示，共分为 4 个 AS，AS100 和 AS200 作为 ISP，PE1 和 ASBR1 属于 AS100，PE2 和 ASBR2 属于 AS200。CE1 和 CE2 属于同一个 VPN，CE1 通过 AS100 的 PE1 接入，CE2 通过 AS200 的 PE2 接入。
- 每台路由器的 IP 地址规划详见拓扑图。

# 跨域VPN-OptionA方式配置步骤

| 步骤 | 任务描述 | 参考命令 |
|---|---|---|
| 1 | 配置各接口的IP地址 | 参考基础配置 |
| 2 | 在公网的PE，P和ASBR上配置OSPF | 参考普通OSPF配置 |
| 3 | 在公网的PE，P和ASBR上配置MPLS和LDP协议 | mpls lsr-id lsr-id<br>mpls<br>mpls ldp |
| 4 | 在PE1与ASBR1，PE2与ASBR2之间配置IBGP<br>并在PE1与ASBR1，PE2与ASBR2之间使能MP-IBGP | 参考普通BGP配置<br>ipv4-family vpnv4<br>peer { group-name \| ipv4-address \| ipv6-address } enable |
| 5 | 在PE和ASBR上配置VPN实例。配置RD和RT | ip vpn-instance vpn-instance-name<br>route-distinguisher route-distinguisher<br>vpn-target vpn-target [ both \|export-extcommunity \|<br>importextcommunity] |
| 6 | 在PE，ASBR上配置接口与VPN 实例关联 | ip binding vpn-instance vpn-instance-name |
| 7 | 在PE与CE，ASBR1与ASBR2之间在VPN实例中配置EBGP | 参考BGP配置 |

# 跨域VPN-OptionA方式配置举例 (1)

1. 搭建拓扑，配置IP：
   - 按照拓扑要求搭建试验拓扑，并配置IP地址。
2. 配置OSPF路由协议：
   - 在PE1，P1和ASBR1，PE2，P2和ASBR2上配置OSPF协议。
3. 在公网的PE，P和ASBR上配置MPLS和LDP协议。

- 搭建拓扑，配置 IP：按照拓扑要求搭建试验拓扑，并配置 IP 地址。
- CE1 需要配置 Loopback 0，GE0/0/1；
- PE1 需要配置 Loopback 0，GE0/0/0；
- P1 需要配置 Loopback 0，GE0/0/0，GE0/0/1；
- ASBR1 需要配置 Loopback 0，GE0/0/1；
- ASBR2 需要配置 Loopback 0，GE0/0/1；
- P2 需要配置 Loopback 0，GE0/0/0，GE0/0/1；
- PE2 需要配置 Loopback 0，GE0/0/0；
- CE2 需要配置：Loopback 0，GE0/0/1。
- 配置 OSPF 路由协议：在 PE1、P1 和 ASBR1，PE2、P2 和 ASBR2 上配置 OSPF 协议
- 在 PE1 上宣告网络 1.1.1.1/32， 12.12.12.0/30；
- 在 P1 上宣告网络 2.2.2.2/32，12.12.12.0/30，23.23.23.0/30；
- 在 ASBR1 上宣告网络 3.3.3.3/32，23.23.23.0/30；
- 在 ASBR2 上宣告网络 4.4.4.4/32，45.45.45.0/30；
- 在 P2 上宣告网络 5.5.5.5/32，45.45.45.0/30，56.56.56.0/30；
- 在 PE2 上宣告网络 6.6.6.6/32， 56.56.56.0/30。

# 跨域VPN-OptionA方式配置举例 (2)

3. 在公网的PE,P和ASBR上配置MPLS和LDP协议。
   - 以PE1配置为例：

```
<PE1>system-view
[PE1]mpls lsr-id 1.1.1.1
[PE1]mpls
Info: Mpls starting, please wait... OK!
[PE1-mpls]mpls ldp
[PE1-mpls-ldp]quit
[PE1]interface GigabitEthernet0/0/0
[PE1-GigabitEthernet0/0/0]mpls
[PE1-GigabitEthernet0/0/0]mpls ldp
[PE1-GigabitEthernet0/0/0]quit
```

- P1，P2，PE2，ASBR1，ASBR2的配置参考PE1。

# 跨域VPN-OptionA方式配置举例 (3)

- 上述配置完成后，同一AS的P和PE、P和ASBR之间应该建立起LDP对等体，在各P设备上执行display mpls ldp session命令可以看到显示结果中Session State项为“Operational”。
- 结果验证，以P1为例：

```
[P1]display mpls ldp session
 LDP Session(s) in Public Network
 Codes: LAM(Label Advertisement Mode), SsnAge Unit(DDDD:HH:MM)
 A '*' before a session means the session is being deleted.
 ------------------------------------------------------------------------------
 PeerID          Status      LAM  SsnRole  SsnAge       KASent/Rcv
 ------------------------------------------------------------------------------
 1.1.1.1:0       Operational DU   Active   0000:02:21   568/568
 3.3.3.3:0       Operational DU   Passive  0000:02:21   567/567
 ------------------------------------------------------------------------------
 TOTAL: 2 session(s) Found.
```

# 跨域VPN-OptionA方式配置举例 (4)

4. 配置PE1与ASBR1，PE2与ASBR2之间配置IBGP和MP-IBGP。

```
[PE1]bgp 100
[PE1-bgp]peer 3.3.3.3 as-number 100
                //配置PE1与ASBR1之间的IBGP邻居关系
[PE1-bgp]peer 3.3.3.3 connect-interface loopback 0
                //配置建立对等体的接口为LOOPBACK 0
[PE1-bgp]ipv4-family vpnv4
                //进入BGP的VPNv4视图
[PE1-bgp-af-vpnv4]peer 3.3.3.3 enable
                //使能PE1与ASBR1的MP-IBGP邻居
[PE1-bgp-af-vpnv4]quit
```

- PE2，ASBR1，ASBR2的配置参考PE1。

# 跨域VPN-OptionA方式配置举例 (5)

- 结果验证：

```
[PE1] display bgp vpnv4 all peer
BGP local router ID : 1.1.1.1
Local AS number : 100 Total number of peers : 1
Peers in established state : 1
Peer         V    AS  MsgRcvd  MsgSent  OutQ  Up/Down       State PrefRcv  3.3.3.3        4   100
3       7     0 00:01:36  Established     0
```

# 跨域VPN-OptionA方式配置举例 (6)

5. 在PE和ASBR上配置VPN实例。

```
[PE1]ip vpn-instance huawei
[PE1-vpn-instance-huawei]route-distinguisher 100:1
[PE1-vpn-instance-huawei]vpn-target 100:1 both
[PE1-vpn-instance-huawei]quit
```

```
[ASBR1]ip vpn-instance huawei
[ASBR1-vpn-instance-huawei]route-distinguisher 100:1
[ASBR1-vpn-instance-huawei]vpn-target 100:1 both
[ASBR1-vpn-instance-huawei]quit
```

- PE2与ASBR2的配置参考PE1和ASBR1。

# 跨域VPN-OptionA方式配置举例 (7)

6. 在PE，ASBR上配置接口与VPN实例关联

```
[PE1]interface GigabitEthernet0/0/1
[PE1-GigabitEthernet0/0/1]ip binding vpn-instance huawei
[PE1-GigabitEthernet0/0/1]ip address 10.1.1.2 30
[PE1-GigabitEthernet0/0/1]quit
```

```
[ASBR1] interface GigabitEthernet0/0/0
[ASBR1-GigabitEthernet0/0/0]ip binding vpn-instance huawei
[ASBR1-GigabitEthernet0/0/0]ip address 34.34.34.1 30
[ASBR1-GigabitEthernet0/0/0]quit
//创建VPN实例，并将此实例绑定到连接ASBR2的接口（ASBR1认为ASBR2是自己的CE）。
```

- PE2，ASBR2的配置参考以上。

# 跨域VPN-OptionA方式配置举例 (8)

7. 在PE与CE，ASBR1与ASBR2之间在VPN实例中配置EBGP。

```
[PE1]bgp 100
[PE1-bgp]ipv4-family vpn-instance huawei
[PE1-bgp-huawei]peer 10.1.1.1 as-number 65001
//配置PE1与CE1的BGP邻居关系
```

```
[CE1]bgp 65001
[CE1-bgp]peer 10.1.1.2 as-number 100
[CE1-bgp]network 11.11.11.11 32
```

```
[ASBR1]bgp 100
[ASBR1-bgp]ipv4-family vpn-instance huawei
[ASBR1-bgp-huawei]peer 34.34.34.2 as-number 200
//配置ASBR1与ASBR2的BGP邻居关系
```

- 其他设备的配置参考以上。

# 跨域VPN-OptionA方式配置举例 (9)

- 结果验证：

```
[PE1]display bgp vpnv4 vpn-instance huawei peer
BGP local router ID : 1.1.1.1
Local AS number : 100
Total number of peers : 1
Peers in established state : 1
Peer      V  AS  MsgRcvd  MsgSent  OutQ  Up/Down     State      PrefRcv  10.1.1.1  4 65001
10        10     0 00:07:10 Established       0
```

# 跨域VPN-OptionA方式配置验证 (1)

- 跨域VPN-OptionA方式配置验证：
  - 上述配置完成后，CE之间能学习到对方的接口路由，CE1和CE2能够相互ping通。
  - 以CE1的显示为例：

```
[CE1]dis ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 6        Routes : 6

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

10.1.1.0/30         Direct  0    0           D   10.1.1.1        GigabitEthernet0/0/1
10.1.1.1/32         Direct  0    0           D   127.0.0.1       GigabitEthernet0/0/1
11.11.11.11/32      Direct  0    0           D   127.0.0.1       LoopBack0
22.22.22.22/32      EBGP    255  0           D   10.1.1.2        GigabitEthernet0/0/1
127.0.0.0/8         Direct  0    0           D   127.0.0.1       InLoopBack0
127.0.0.1/32        Direct  0    0           D   127.0.0.1       InLoopBack0
```

# 跨域VPN-OptionA方式配置验证 (2)

- 在ASBR上执行display ip routing-table vpn-instance命令，可以看到ASBR上为VPN维护的路由表。

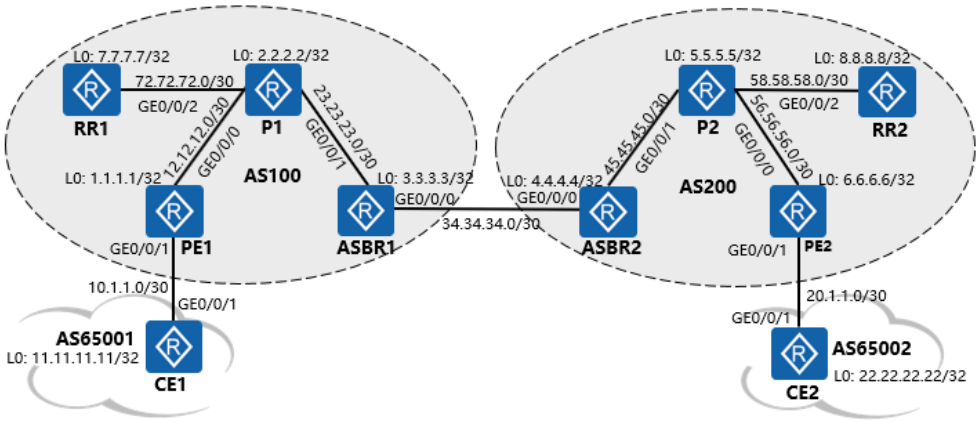
```
[ASBR1]dis ip routing-table vpn-instance huawei
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: huawei
         Destinations : 4        Routes : 4

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

11.11.11.11/32      IBGP    255  0           RD  1.1.1.1         GigabitEthernet0/0/1
22.22.22.22/32      EBGP    255  0           D   34.34.34.2      GigabitEthernet0/0/0
34.34.34.0/30       Direct  0    0           D   34.34.34.1      GigabitEthernet0/0/0
34.34.34.1/32       Direct  0    0           D   127.0.0.1       GigabitEthernet0/0/0
```

# 跨域VPN-OptionA方式配置验证 (3)

- 在ASBR上执行display bgp vpnv4 all routing-table命令，可看到ASBR上的VPNv4路由。

```
[ASBR1]dis bgp vpnv4 all routing-table

BGP Local router ID is 3.3.3.3
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? – incomplete

Total number of routes from all PE: 2
Route Distinguisher: 100:1
     Network            NextHop    MED    LocPrf    PrefVal     Path/Ogn

*>i  11.11.11.11/32     1.1.1.1    0      100       0           65001i
*>   22.22.22.22/32     34.34.34.2                  0           200 65002i

VPN-Instance huawei, Router ID 3.3.3.3:
Total Number of Routes: 2
     Network            NextHop    MED    LocPrf    PrefVal     Path/Ogn
*>i  11.11.11.11/32     1.1.1.1    0      100       0           65001i
*>   22.22.22.22/32     34.34.34.2                  0           200 65002i
```

# 跨域VPN-OptionB方式拓扑介绍




- 配置任务：使用跨域VPN-OptionB的方式实现两端的CE设备互访。注：本实例使用OSPF作为IGP。

# 跨域VPN-OptionB方式配置步骤 (1)

| 步骤 | 任务描述 | 参考命令 |
| --- | --- | --- |
| 1 | 配置各接口的IP地址 | 参考基础配置 |
| 2 | 在公网的PE，P，ASBR和RR上配置OSPF | 参考普通OSPF配置 |
| 3 | 在公网的PE，P和ASBR上配置MPLS和LDP协议 | mpls lsr-id lsr-id<br>mpls<br>mpls ldp |
| 4 | 在PE1&ASBR1与RR1，PE2&ASBR2与RR2之间配置IBGP<br>在PE，RR，ASBR之间在VPNv4视图下使能MP-IBGP | 参考普通BGP配置<br>ipv4-family vpnv4<br>peer { group-name \| ipv4-address \| ipv6-address } enable |
| 5 | 在RR1与RR2上配置VPN路由反射 | ipv4-family vpnv4<br>peer { group-name \| ipv4-address \| ipv6-address } reflect-client |
| 6 | 在PE上配置VPN实例。配置RD和RT | ip vpn-instance vpn-instance-name<br>route-distinguisher route-distinguisher<br>vpn-target vpn-target [ both \|export-extcommunity \|<br>importextcommunity] |

# 跨域VPN-OptionB方式配置步骤 (2)

| 步骤 | 任务描述 | 参考命令 |
| --- | --- | --- |
| 7 | 在PE上配置接口与VPN 实例关联 | ip binding vpn-instance vpn-instance-name |
| 8 | 在PE的VPN实例中配置与CE的EBGP | 参考BGP配置 |
| 9 | 在ASBR1与ASBR2之间相连的接口上使能MPLS | mpls |
| 10 | 在ASBR1与ASBR2之间在VPNv4视图下配置EBGP（即MP-EBGP） | ipv4-family vpnv4<br>peer { group-name \| ipv4-address \| ipv6-address }<br>enable<br>undo policy vpn-target |
| 说明 | 1. ASBR1和ASBR2之间需使能MPLS，LDP不用使能。<br>2. 采用OptionB方式时，不同AS的PE的VPN实例的VPN-Target需要匹配。<br>3. 不再需要在ASBR1，ASBR2上配置VPN实例和接口绑定。<br>4. ASBR1，ASBR2，RR1，RR2上需要进行特殊配置：在VPNV4地址族中配置undo policy vpn-target。 | |

# 跨域VPN-OptionB方式配置举例 (1)

- 第1-4步的配置参考前面OptionA方式的部分。

5. 在RR上配置VPN路由反射，并设置不对接收的VPN路由进行VPN-target过滤。

```
[RR1-bgp]ipv4-family vpnv4
[RR1-bgp-af-vpnv4]undo policy vpn-target
[RR1-bgp-af-vpnv4]peer 1.1.1.1 reflect-client
[RR1-bgp-af-vpnv4]peer 3.3.3.3 reflect-client
```

6. 在PE上配置VPN实例，配置RD（100:1）和RT（100:1）。

```
[PE1]ip vpn-instance huawei
[PE1-vpn-instance-huawei]route-distinguisher 100:1
[PE1-vpn-instance-huawei]vpn-target 100:1 both
[PE1-vpn-instance-huawei]quit
```

- RR2配置参考RR1，PE2的配置参考PE1。

# 跨域VPN-OptionB方式配置举例 (2)

7. 在PE上配置接口与VPN 实例关联。

```
[PE1]interface GigabitEthernet0/0/1
[PE1-GigabitEthernet0/0/1]ip binding vpn-instance huawei
[PE1-GigabitEthernet0/0/1]ip address 10.1.1.2 30
[PE1-GigabitEthernet0/0/1]quit
```

- PE2的配置参考PE1。

# 跨域VPN-OptionB方式配置举例 (3)

8. 在PE的VPN实例中配置与CE的EBGP。

```
[PE1]bgp 100
[PE1-bgp]ipv4-family vpn-instance huawei
[PE1-bgp-huawei]peer 10.1.1.1 as-number 65001
//配置PE1与CE1的BGP邻居关系
```

```
[CE1]bgp 65001
[CE1-bgp]peer 10.1.1.2 as-number 100
[CE1-bgp]network 11.11.11.11 32
```

- PE2与CE2的BGP配置参考PE1与CE1。

# 跨域VPN-OptionB方式配置举例 (4)

9. 在ASBR1与ASBR2之间相连的接口上使能MPLS。

```
[ASBR1]interface GigabitEthernet0/0/0
[ASBR1-GigabitEthernet0/0/0]ip address 34.34.34.1 24
[ASBR1-GigabitEthernet0/0/0]mpls
//在ASBR1与ASBR2之间相连的接口上使能MPLS
```

- ASBR2的配置参考ASBR1。

# 跨域VPN-OptionB方式配置举例 (5)

10. 在ASBR之间的VPNv4视图下配置EBGP (即MP-EBGP)。

```
[ASBR1]bgp 100
[ASBR1-bgp]peer 34.34.34.2 as-number 200
[ASBR1-bgp]ipv4-family vpnv4
[ASBR1-bgp-af-vpnv4]peer 34.34.34.2 enable
//与ASBR2建立MP-EBGP对等体关系
[ASBR1-bgp-af-vpnv4]undo policy vpn-target
//不对接收的VPNv4路由进行VPN-target过滤。
```

- ASBR2的配置参考ASBR1。

# 跨域VPN-OptionB方式配置验证 (1)

- 配置验证：上述配置完成后，CE之间能学习到对方的接口路由，CE1和CE2能够相互ping通。
- 以CE1的显示为例：

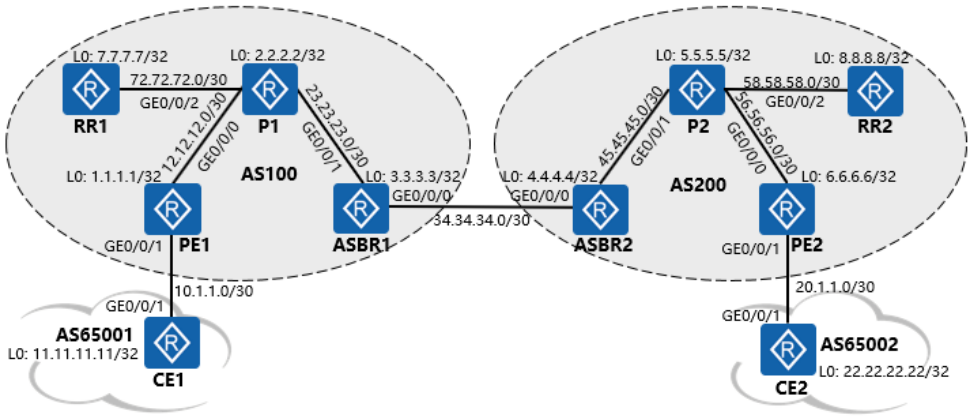
```
[CE1]dis ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 6        Routes : 6

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

10.1.1.0/30         Direct  0    0           D   10.1.1.1        GigabitEthernet0/0/1
10.1.1.1/32         Direct  0    0           D   127.0.0.1       GigabitEthernet0/0/1
11.11.11.11/32      Direct  0    0           D   127.0.0.1       LoopBack0
22.22.22.22/32      EBGP    255  0           D   10.1.1.2        GigabitEthernet0/0/1
127.0.0.0/8         Direct  0    0           D   127.0.0.1       InLoopBack0
127.0.0.1/32        Direct  0    0           D   127.0.0.1       InLoopBack0
```

## 跨域VPN-OptionB方式配置验证 (2)

- 在ASBR上执行display bgp vpnv4 all routing-table命令，可以看到ASBR上的VPNv4路由。

```
[ASBR1]dis bgp vpnv4 all routing-table
 BGP Local router ID is 23.23.23.2
 Status codes: * - valid, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
 Total number of routes from all PE: 2
 Route Distinguisher: 100:1

      Network            NextHop        MED        LocPrf    PrefVal  Path/Ogn
 *>i  11.11.11.11/32     1.1.1.1        0          100       0        65001i
 *>   22.22.22.22/32     34.34.34.2                           0        200  65002i
```

## 跨域VPN-OptionC方式拓扑介绍 (方案一)




- 配置任务：使用跨域VPN-OptionC的方式实现两端的CE设备互访。注：本实例使用OSPF作为IGP。

- 如上图所示，共分为 4 个 AS，AS100 和 AS200 作为 ISP，PE1，P1，RR1 和 ASBR1 属于 AS100，PE2，P2，RR2 和 ASBR2 属于 AS200。CE1 和 CE2 属于同一个 VPN，CE1 通过 AS100 的 PE1 接入，CE2 通过 AS200 的 PE2 接入。
- 每台路由器的 IP 地址规划详见拓扑图。
- 本例为 OptionC 方式实现方案一，可以采用 PE1 与 PE2 直接建立 MP-EBGP（不带 RR）来传递跨域 VPN 路由，也可以采用 RR1 与 RR2 建立 MP-EBGP 邻居（PE1 与 RR1，PE2 与 RR2 建立 MP-IBGP 邻居）传递跨域 VPN 路由，两者相似，

本例采用 RR 方式实现 OptionC 方案一。

## 跨域VPN-OptionC方式配置步骤 (1)

| 步骤 | 任务描述 | 参考命令 |
| --- | --- | --- |
| 1 | 配置各接口的IP地址 | 参考基础配置 |
| 2 | 在公网的PE,P,RR,ASBR上配置OSPF | 参考普通OSPF配置 |
| 3 | 在公网的PE,P,ASBR上配置MPLS和LDP协议 | mpls lsr-id lsr-id<br>mpls<br>mpls ldp |
| 4 | 在PE,P,ASBR与RR之间配置IBGP，并设置其为RR的反射客户体 | 参考普通BGP配置 |
| 5 | 在ASBR1与ASBR2之间配置EBGP | 参考普通BGP配置 |
| 6 | 在PE上配置VPN实例，配置RD（100:1）和RT（100:1） | ip vpn-instance vpn-instance-name<br>route-distinguisher route-distinguisher<br>vpn-target vpn-target [ both \|export-extcommunity \| importextcommunity] |
| 7 | 在PE上配置接口与VPN 实例关联 | ip binding vpn-instance vpn-instance-name |
| 8 | 在PE的VPN实例中配置与CE的EBGP | 参考BGP配置 |
| 9 | 在RR1与RR2之间在VPNv4视图下配置EBGP（即多跳的MP-EBGP）<br>在PE1与RR1之间在VPNv4视图下配置MP-IBGP<br>在PE2与RR2之间在VPNv4视图下配置MP-IBGP<br>设置RR在传递路由时不改变下一跳 | peer { group-name \| ipv4-address \| ipv6-address } ebgp-max-hop [ hop-count ]<br>ipv4-family vpnv4<br>peer { group-name \| ipv4-address \| ipv6-address } enable<br>peer next-hop-invariable |

## 跨域VPN-OptionC方式配置步骤 (2)

| 步骤 | 任务描述 | 参考命令 |
| --- | --- | --- |
| 10 | 使能PE与RR，RR与ASBR，ASBR与ASBR之间相互交换标签IPv4路由的能力 | peer { group-name \| ipv4-address } label-route-capability |
| 11 | 配置本端ASBR向远端ASBR，本端ASBR向本端RR发布的路由应用路由策略 | route-policy route-policy-name { permit \| deny } node node<br>peer { group-name \| ipv4-address \| ipv6-address } route-policy route-policy-name { import \| export } |
| 说明 | 1. ASBR1和ASBR2之间需使能MPLS，LDP可以不使能。<br>2. 不需要在ASBR1，ASBR2上配置VPN实例和接口绑定。<br>3. 在以下路由器之间配置能够交换带标签的IPv4路由：PE1与RR1、RR1与ASBR1、PE2与RR2、RR2与ASBR2、ASBR1与ASBR2。<br>4. 在RR传递VPN路由至远端RR，及从远端RR接收路由进一步传递至本端PE时，需要设置下一跳不改变。<br>5. 在ASBR1和ASBR2上配置路由策略：对于从本AS内的RR接收的路由，在向对端AS的ASBR发布时，分配MPLS标签；对于向本AS内的RR发布的路由，如果是带标签的IPv4路由，为其分配新的MPLS标签。 | |

## 跨域VPN-OptionC方式配置举例 (1)

- 第1-3步的配置参考前面OptionA方式的部分。

4. 配置PE，P，ASBR与RR之间的IBGP邻居，并设置其为RR的反射客户体。

```
[RR1-bgp]peer 1.1.1.1 as-number 100
[RR1-bgp]peer 1.1.1.1 connect-interface Loopback0
[RR1-bgp]peer 1.1.1.1 reflect-client
[RR1-bgp]peer 2.2.2.2 as-number 100
[RR1-bgp]peer 2.2.2.2 connect-interface Loopback0
[RR1-bgp]peer 2.2.2.2 reflect-client
[RR1-bgp]peer 3.3.3.3 as-number 100
[RR1-bgp]peer 3.3.3.3 connect-interface Loopback0
[RR1-bgp]peer 3.3.3.3 reflect-client
```

5. 在ASBR1与ASBR2之间配置EBGP。

```
[ASBR1-bgp]peer 34.34.34.2 as-number 200
```

- 步骤 4，RR2 的配置参考 RR1。PE，P，ASBR 的配置参考普通 BGP 邻居配置。
- 步骤 5，ASBR2 的配置参考 ASBR1。
- ASBR 之间建立单播 EBGP 邻居，并将本端 RR 和 PE 的 loopback 发布给对方。
- ASBR1 在向 ASBR2 发布 RR1 ,PE1 的 loopback 时，为其分配 MPLS 标签；ASBR2 向 RR2 传递 RR1 和 PE1 的 loopback 路由时，为其分配新的 MPLS 标签。
- ASBR,PE 同本端 RR 建立邻居时使能交换标签能力。
- ASBR,P,PE 同 RR 建立 ipv4 邻居：
- ASBR 通过 ipv4 邻居学习将从对端 ASBR 学到的 RR 的 loopback，传递给本端 RR，用于本端 RR 与对端建立 vpnv4 邻居。
- ASBR 通过 ipv4 邻居学习将从对端 ASBR 学到的 RR 和 PE 的 loopback，传递给本端 RR，本端 RR 再将其反射给 P，用于跨域 bgp 路由的递归查询。
- ASBR 通过 ipv4 邻居学习将从对端 ASBR 学到的 RR 和 PE 的 loopback，传递给本端 RR，本端 RR 再将其反射给 PE，用于跨域之间的 PE 建立 BGP LSP。

# 跨域VPN-OptionC方式配置举例 (2)

6. 在PE上配置VPN实例，配置RD (100:1)和RT (100:1)。

```
[PE1]ip vpn-instance huawei
[PE1-vpn-instance-huawei]route-distinguisher 100:1
[PE1-vpn-instance-huawei-af-ipv4]vpn-target 100:1 both
```

7. 在PE上配置接口与VPN实例关联。

```
[PE1]interface GigabitEthernet0/0/1
[PE1-GigabitEthernet0/0/1]ip binding vpn-instance huawei
[PE1-GigabitEthernet0/0/1]ip address 10.1.1.2 30
[PE1-GigabitEthernet0/0/1]quit
```

- PE2的配置参考PE1。

# 跨域VPN-OptionC方式配置举例 (3)

8. 在PE的VPN实例中配置与CE的EBGP。

```
[PE1]bgp 100
[PE1-bgp]ipv4-family vpn-instance huawei
[PE1-bgp-huawei]peer 10.1.1.1 as-number 65001
//配置PE1与CE1的BGP邻居关系
```

```
[CE1]bgp 65001
[CE1-bgp]peer 10.1.1.2 as-number 100
[CE1-bgp]network 11.11.11.11 32
```

- PE2与CE2的配置参考PE1与CE1。

## 跨域VPN-OptionC方式配置举例 (4)

9. 在RR1与RR2之间的VPNv4视图下配置MP-EBGP，并配置传递路由时下一跳不改变。

```
[RR1]bgp 100
[RR1-bgp]peer 8.8.8.8 as-number 200
[RR1-bgp]peer 8.8.8.8 connect-interface LoopBack 0
[RR1-bgp]peer 8.8.8.8 ebgp-max-hop 10
[RR1-bgp]ipv4-family vpnv4
[RR1-bgp-af-vpnv4]peer 8.8.8.8 enable
[RR1-bgp-af-vpnv4]peer 8.8.8.8 next-hop-invariable
[RR1-bgp-af-vpnv4]undo policy vpn-target
```

- 在PE与RR之间在VPNv4视图下配置MP-IBGP，并配置传递路由时下一跳不改变。

```
[RR1]bgp 100
[RR1-bgp]peer 1.1.1.1 as-number 100
[RR1-bgp]peer 1.1.1.1 connect-interface LoopBack 0
[RR1-bgp]ipv4-family vpnv4
[RR1-bgp-af-vpnv4]peer 1.1.1.1 enable
[RR1-bgp-af-vpnv4]peer 1.1.1.1 next-hop-invariable
```

- PE2 与 RR2 的 MP-IBGP 配置参考 PE1 与 RR1 的配置。
- 配置“undo policy vpn-target”原理与 Opiton B 中相同，即 RR 上不用 RT 来过滤路由。
- 配置“peer X.X.X.X next-hop-invariable”，保证对端 PE 可以在迭代路由，用于流量传输时，通往本端 PE 的 BGP LSP。

- RR 之间在 vpnv4 视图下建立 MP-EBGP 邻居，向对端传递路由时不改变下一跳。即对端 PE 学到的 VPNv4 路由的下一跳是本端 PE。
- RR 与本端 PE 建立 vpnv4 邻居，RR 向本端 PE 传递路由时不改变下一跳，即本端 PE 学到的 VPNv4 路由的下一跳是对端 PE。
- 本端 PE 只与本端 RR 建立 VPNv4 邻居，本端 RR 与对端 RR 建立 VPNv4 邻居，实现了跨域 VPN 路由的传递。

# 跨域VPN-OptionC方式配置举例 (5)

10. 使能PE与RR，RR与ASBR，ASBR与ASBR之间相互交换标签IPv4路由的能力。

```
[PE1]bgp 100
[PE1-bgp]peer 7.7.7.7 as-number 100
[PE1-bgp]peer 7.7.7.7 label-route-capability
```

```
[RR1]bgp 100
[RR1-bgp]peer 1.1.1.1 as-number 100
[RR1-bgp]peer 1.1.1.1 label-route-capability
[RR1-bgp]peer 3.3.3.3 as-number 100
[RR1-bgp]peer 3.3.3.3 label-route-capability
```

```
[ASBR1]bgp 100
[ASBR1-bgp]peer 7.7.7.7 as-number 100
[ASBR1-bgp]peer 7.7.7.7 label-route-capability
[ASBR1-bgp]peer 34.34.34.2 as-number 200
[ASBR1-bgp]peer 34.34.34.2 label-route-capability
```

- PE2,RR2,ASBR2 的配置分别参考 PE1,RR1 和 ASBR1。

# 跨域VPN-OptionC方式配置举例 (6)

11. 配置本端ASBR向远端ASBR，本端ASBR向本端RR发布的路由应用路由策略。

```
[ASBR1]interface GigabitEthernet0/0/0
[ASBR1-GigabitEthernet0/0/0]ip address 34.34.34.1 30
[ASBR1-GigabitEthernet0/0/0]mpls
[ASBR1-GigabitEthernet0/0/0]quit
[ASBR1]route-policy policy1 permit node 10
[ASBR1-route-policy]apply mpls-label //为匹配条件的路由分配标签
[ASBR1-route-policy]quit
[ASBR1]route-policy policy2 permit node 10
[ASBR1-route-policy]if-match mpls-label
[ASBR1-route-policy]apply mpls-label//如果路由带有标签，则为其分配标签
[ASBR1-route-policy]quit
//在ASBR1上创建2个路由策略
```

# 跨域VPN-OptionC方式配置举例 (7)

```
[ASBR1]bgp 100
[ASBR1-bgp]peer 7.7.7.7 route-policy policy2 export
//在ASBR1上对向RR1发布的路由应用路由策略，对于向本AS内的RR发布的路由，如果是带标签的IPv4路由，为其分配新的MPLS标签

[ASBR1-bgp]peer 34.34.34.2 as-number 200
[ASBR1-bgp]peer 34.34.34.2 route-policy policy1 export
[ASBR1-bgp]peer 34.34.34.2 label-route-capability
[ASBR1-bgp]quit
//配置ASBR1：对向ASBR2发布的路由应用路由策略，对于从本AS内的RR接收的路由，在向对端AS的ASBR发布时，分配MPLS标签

[ASBR1]bgp 100
[ASBR1-bgp]network 1.1.1.1 32
[ASBR1-bgp]network 7.7.7.7 32
//配置ASBR1：将PE1和RR1的Loopback地址发布给ASBR2，进而发布给RR2和PE2
[ASBR1-bgp]quit
```

- ASBR2上的配置参考ASBR1。

# 跨域VPN-OptionC方式配置验证 (1)

- 配置验证：上述配置完成后，CE之间能学习到对方的环回口路由，CE1和CE2能够相互ping通。
- 以CE1的显示为例：

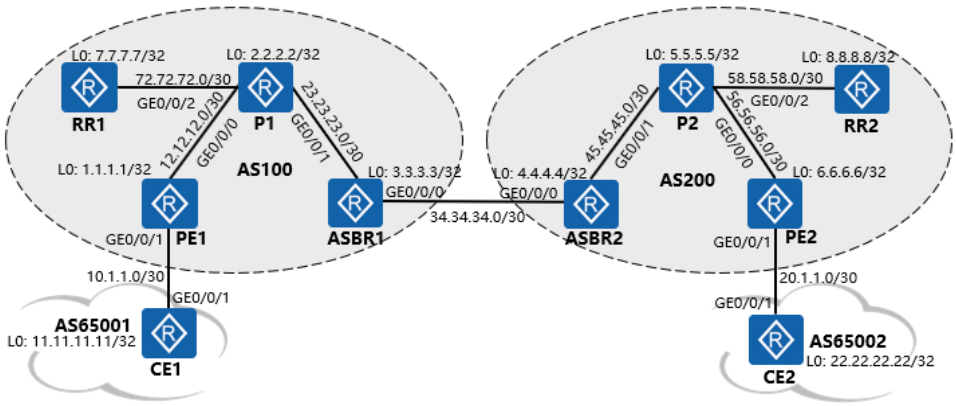
```
[CE1]dis ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 6        Routes : 6

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

10.1.1.0/30         Direct  0    0           D   10.1.1.1        GigabitEthernet0/0/1
10.1.1.1/32         Direct  0    0           D   127.0.0.1       GigabitEthernet0/0/1
11.11.11.11/32      Direct  0    0           D   127.0.0.1       LoopBack0
22.22.22.22/32      EBGP    255  0           D   10.1.1.2        GigabitEthernet0/0/1
127.0.0.0/8         Direct  0    0           D   127.0.0.1       InLoopBack0
127.0.0.1/32        Direct  0    0           D   127.0.0.1       InLoopBack0
```

## 跨域VPN-OptionC方式配置验证 (2)

- 在ASBR上执行display bgp routing-table label命令，可以看到路由的标签信息。

```
[ASBR1]dis bgp routing-table label

BGP Local router ID is 23.23.23.2
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete

Total Number of Routes: 4
      Network        NextHop        In/Out Label

*>    1.1.1.1        23.23.23.1     1080/NULL
*>    6.6.6.6        34.34.34.2     1077/1071
*>    7.7.7.7        23.23.23.1     1081/NULL
*>    8.8.8.8        34.34.34.2     1076/1070
```

## 跨域VPN-OptionC方式拓扑介绍 (方案二)




- 配置任务：使用跨域VPN-OptionC的方式实现两端的CE设备互访。注：本实例使用OSPF作为IGP。

- 如上图所示，共分为 4 个 AS，AS100 和 AS200 作为 ISP，PE1，P1，RR1 和 ASBR1 属于 AS100，PE2，P2，RR2 和 ASBR2 属于 AS200。CE1 和 CE2 属于同一个 VPN，CE1 通过 AS100 的 PE1 接入，CE2 通过 AS200 的 PE2 接入。
- 每台路由器的 IP 地址规划详见拓扑图。
- 本例为 OptionC 方式实现方案二。方案二的实现与方案一大致相似，主要区别在于，当本端 ASBR 收到对端 ASBR 传递来的 labeled-IPv4-Route 后，触发 LDP 为 BGP 标签路由分标签。

# 跨域VPN-OptionC方式配置步骤 (1)

| 步骤 | 任务描述 | 参考命令 |
| --- | --- | --- |
| 1 | 配置各接口的IP地址 | 参考基础配置 |
| 2 | 在公网的PE,P,RR,ASBR上配置OSPF | 参考普通OSPF配置 |
| 3 | 在公网的PE,P,ASBR上配置MPLS和LDP协议 | mpls lsr-id lsr-id<br>mpls<br>mpls ldp |
| 4 | 在ASBR1与ASBR2之间配置EBGP<br>在ASBR通告各自AS内RR的loopback接口，引入到IGP进程 | 参考普通BGP配置<br>import-route bgp |
| 5 | 在PE上配置VPN实例。配置RD（100:1）和RT（100:1） | ip vpn-instance vpn-instance-name<br>route-distinguisher route-distinguisher<br>vpn-target vpn-target [ both \|export-extcommunity \| importextcommunity] |
| 6 | 在PE上配置接口与VPN 实例关联 | ip binding vpn-instance vpn-instance-name |
| 7 | 在PE的VPN实例中配置与CE的EBGP | 参考BGP配置 |
| 8 | 在RR1与RR2之间在VPNv4视图下配置EBGP（即多跳的MP-EBGP）<br>在PE1与RR1之间在VPNv4视图下配置MP-IBGP<br>在PE2与RR2之间在VPNv4视图下配置MP-IBGP<br>设置RR在传递路由时不改变下一跳 | peer { group-name \| ipv4-address \| ipv6-address } ebgp-max-hop [ hop-count ]<br>ipv4-family vpnv4<br>peer { group-name \| ipv4-address \| ipv6-address } enable<br>peer next-hop-invariable |

# 跨域VPN-OptionC方式配置步骤 (2)

| 步骤 | 任务描述 | 参考命令 |
| --- | --- | --- |
| 9 | 使能ASBR间交换标签IPv4路由的能力 | peer { group-name \| ipv4-address } label-route-capability |
| 10 | 在ASBR上使能LDP为带标签的公网BGP路由分标签的能力 | lsp-trigger bgp-label-route |
| 11 | 配置本端ASBR向远端ASBR发布的路由应用路由策略 | route-policy route-policy-name { permit \| deny } node node<br>peer { group-name \| ipv4-address \| ipv6-address } route-policy route-policy-name { import \| export } |
| 说明 | 1. ASBR1和ASBR2之间需使能MPLS，LDP可以不使能；<br>2. 不需要在ASBR1，ASBR2上配置VPN实例和接口绑定；<br>3. 在以下路由器之间配置能够交换带标签的IPv4路由：ASBR1与ASBR2；<br>4. 在RR传递VPN路由至远端RR，及从远端RR接收路由进一步传递至本端PE时，需要设置下一跳不改变；<br>5. 在ASBR1和ASBR2上配置路由策略：在向对端AS的ASBR发布时，分配MPLS标签；<br>6. ASBR上需要配置MPLS触发建立BGP标签路由LSP的能力。 | |

## 跨域VPN-OptionC方式配置举例 (1)

- 第1-3步的配置参考前面OptionA方式的部分。

4. 在ASBR1与ASBR2之间配置EBGP。在ASBR通告各自AS内RR的loopback接口，引入到IGP进程。

```
[ASBR1]bgp 100
[ASBR1-bgp]peer 34.34.34.2 as-number 200
[ASBR1-bgp]network 7.7.7.7 255.255.255.255
[ASBR1-bgp]quit
[ASBR1]ospf
[ASBR1-ospf-1]import-route bgp
```

```
[ASBR2]bgp 200
[ASBR2-bgp]peer 34.34.34.1 as-number 100
[ASBR2-bgp]network 8.8.8.8 255.255.255.255
[ASBR2-bgp]quit
[ASBR2]ospf
[ASBR2-ospf-1]import-route bgp
```

- 在 OSPF 进程中引入 BGP 路由，是为了让 RR1 与 RR2 能够顺利建立 EBGP 邻居关系，进而传递 VPN 路由。建议在 OSPF 中引入 BGP 路由时配置路由策略，做好精确的明细引入，减少不必要的路由进入 IGP 域。

## 跨域VPN-OptionC方式配置举例 (2)

5. 在PE上配置VPN实例，配置RD (100:1)和RT (100:1)。

```
[PE1]ip vpn-instance huawei
[PE1-vpn-instance-huawei]route-distinguisher 100:1
[PE1-vpn-instance-huawei-af-ipv4]vpn-target 100:1 both
```

6. 在PE上配置接口与VPN实例关联。

```
[PE1]interface GigabitEthernet0/0/1
[PE1-GigabitEthernet0/0/1]ip binding vpn-instance huawei
[PE1-GigabitEthernet0/0/1]ip address 10.1.1.2 30
[PE1-GigabitEthernet0/0/1]quit
```

- PE2的配置参考PE1。

## 跨域VPN-OptionC方式配置举例 (3)

7. 在PE的VPN实例中配置与CE的EBGP。

```
[PE1]bgp 100
[PE1-bgp]ipv4-family vpn-instance huawei
[PE1-bgp-huawei]peer 10.1.1.1 as-number 65001
//配置PE1与CE1的BGP邻居关系
```

```
[CE1]bgp 65001
[CE1-bgp]peer 10.1.1.2 as-number 100
[CE1-bgp]network 11.11.11.11 32
```

- PE2与CE2的配置参考PE1与CE1。

## 跨域VPN-OptionC方式配置举例 (4)

8. 在RR1与RR2之间的VPNv4视图下配置MP-EBGP,并配置传递路由时下一跳不改变。

```
[RR1]bgp 100
[RR1-bgp]peer 8.8.8.8 as-number 200
[RR1-bgp]peer 8.8.8.8 connect-interface LoopBack 0
[RR1-bgp]peer 8.8.8.8 ebgp-max-hop 10
[RR1-bgp]ipv4-family vpnv4
[RR1-bgp-af-vpnv4]peer 8.8.8.8 enable
[RR1-bgp-af-vpnv4]peer 8.8.8.8 next-hop-invariable
[RR1-bgp-af-vpnv4]undo policy vpn-target
```

- 在PE与RR之间在VPNv4视图下配置MP-IBGP, 并配置传递路由时下一跳不改变。

```
[RR1]bgp 100
[RR1-bgp]peer 1.1.1.1 as-number 100
[RR1-bgp]peer 1.1.1.1 connect-interface LoopBack 0
[RR1-bgp]ipv4-family vpnv4
[RR1-bgp-af-vpnv4]peer 1.1.1.1 enable
[RR1-bgp-af-vpnv4]peer 1.1.1.1 next-hop-invariable
```

- PE2 与 RR2 的 MP-IBGP 配置参考 PE1 与 RR1 的配置。
- 配置“undo policy vpn-target”原理与 Opiton B 中相同，即 RR 上不用 RT 来过滤路由。
- 配置“peer X.X.X.X next-hop-invariable”，保证对端 PE 可以在流量传输时迭代到通往本端 PE 的 BGP LSP。

- RR 之间建立 vpnv4 邻居，向对端传递路由时不改变下一跳。即对端 PE 学到的 VPNv4 路由的下一跳是本端 PE。
- RR 与本端 PE 建立 vpnv4 邻居，RR 向本端 PE 传递路

由时不改变下一跳，即本端 PE 学到的 VPNv4 路由的下一跳是对端 PE。

- 本端 PE 只与本端 RR 建立 VPNv4 邻居，本端 RR 与对端 RR 建立 VPNv4 邻居，实现了跨域 VPN 路由的传递。

## 跨域VPN-OptionC方式配置举例 (5)

9. 使能ASBR间交换标签IPv4路由的能力。

```
[ASBR1]bgp 100
[ASBR1-bgp]peer 34.34.34.2 as-number 200
[ASBR1-bgp]peer 34.34.34.2 label-route-capability
```

```
[ASBR2]bgp 200
[ASBR2-bgp]peer 34.34.34.1 as-number 100
[ASBR2-bgp]peer 34.34.34.1 label-route-capability
```

## 跨域VPN-OptionC方式配置举例 (6)

10. 在ASBR上配置MPLS触发建立BGP标签路由LSP的能力。

```
[ASBR1]mpls
[ASBR1-mpls]lsp-trigger bgp-label-route
```

```
[ASBR2]mpls
[ASBR2-mpls]lsp-trigger bgp-label-route
```

# 跨域VPN-OptionC方式配置举例 (7)

11. 配置本端ASBR向远端ASBR发布路由时应用路由策略。

```
[ASBR1]interface GigabitEthernet0/0/0
[ASBR1-GigabitEthernet0/0/0]ip address 34.34.34.1 30
[ASBR1-GigabitEthernet0/0/0]mpls
[ASBR1-GigabitEthernet0/0/0]quit
[ASBR1]route-policy policy1 permit node 10
[ASBR1-route-policy]apply mpls-label //为匹配条件的路由分配标签
[ASBR1-route-policy]quit
[ASBR1]bgp 100
[ASBR1-bgp]peer 34.34.34.2 as-number 200
[ASBR1-bgp]peer 34.34.34.2 route-policy policy1 export
[ASBR1-bgp]network 1.1.1.1 32
```

- ASBR2的配置参考ASBR1。

# 跨域VPN-OptionC方式配置验证 (1)

- 配置验证：上述配置完成后，CE之间能学习到对方的环回口路由，CE1和CE2能够相互ping通。
- 以CE1的显示为例：

```
[CE1]dis ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 6        Routes : 6

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

10.1.1.0/30         Direct  0    0           D   10.1.1.1        GigabitEthernet0/0/1
10.1.1.1/32         Direct  0    0           D   127.0.0.1       GigabitEthernet0/0/1
11.11.11.11/32      Direct  0    0           D   127.0.0.1       LoopBack0
22.22.22.22/32      EBGP    255  0           D   10.1.1.2        GigabitEthernet0/0/1
127.0.0.0/8         Direct  0    0           D   127.0.0.1       InLoopBack0
127.0.0.1/32        Direct  0    0           D   127.0.0.1       InLoopBack0
```

# 跨域VPN-OptionC方式配置验证 (2)

- 在ASBR上执行display bgp routing-table label命令，可以看到路由的标签信息。

```
[ASBR1]dis bgp routing-table label

BGP Local router ID is 34.34.34.1
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete

Total Number of Routes: 4
      Network         NextHop       In/Out Label

*>    1.1.1.1         23.23.23.1    1047/NULL
*>    6.6.6.6         34.34.34.2    NULL/1041
*>    7.7.7.7         23.23.23.1    1043/NULL
*>    8.8.8.8         34.34.34.2    NULL/1042
```

# 跨域VPN-OptionC方式配置验证 (3)

- 在PE上执行display mpls lsp命令，可以看到去往对方PE的LDP LSP。

```
[PE1]dis mpls lsp
------------------------------------------------------------------------------
                LSP Information: BGP  LSP
------------------------------------------------------------------------------
FEC                 In/Out Label      In/Out IF            Vrf Name
11.11.11.11/32      1042/NULL         -/-                  huawei
------------------------------------------------------------------------------
                LSP Information: LDP LSP
------------------------------------------------------------------------------
FEC                 In/Out Label      In/Out IF            Vrf Name
2.2.2.2/32          NULL/3            -/GE0/0/0
2.2.2.2/32          1037/3            -/GE0/0/0
6.6.6.6/32          NULL/1033         -/GE0/0/0
6.6.6.6/32          1039/1033         -/GE0/0/0
.............
```

## 思考题

1. 以下哪种方案ASBR不用保存用户CE侧路由信息（　　）
   A. Option-A
   B. Option-B
   C. Option-C
2. 只通过LDP/BGP分发标签的情况下，报文转发过程中，可能用到三层标签的是（　）
   A. Option-A
   B. Option-B
   C. Option-C

1 答案：C
2 答案：C