

目 录

第 1 章	MPLS概述.....	1-1
1.1	MPLS概述.....	1-1
1.1.1	MPLS简介.....	1-1
1.1.2	MPLS网络简介.....	1-4
1.1.3	MPLS与路由协议简介.....	1-5
1.1.4	MPLS应用简介.....	1-5
1.1.5	MPLS PHP.....	1-7
第 2 章	LDP.....	2-1
2.1	LDP简介.....	2-1
2.1.1	LDP基本概念.....	2-1
2.1.2	LDP报文格式简介.....	2-2
2.1.3	LDP标签管理.....	2-4
2.1.4	LDP会话.....	2-7
2.1.5	LDP环路检测.....	2-8
2.2	LDP配置.....	2-8
2.3	LDP典型案例.....	2-13
2.4	LDP排错帮助.....	2-14
第 3 章	MPLS VPN.....	3-1
3.1	BGP/MPLS VPN简介.....	3-1
3.1.1	BGP/MPLS VPN网络结构.....	3-1
3.1.2	BGP/MPLS VPN基本概念.....	3-2
3.1.3	BGP/MPLS VPN报文转发.....	3-4
3.1.4	BGP/MPLS VPN组网方案.....	3-5
3.1.5	BGP/MPLS VPN路由发布.....	3-8
3.1.6	跨域VPN简介.....	3-8
3.2	BGP MPLS VPN配置.....	3-9
3.3	BGP MPLS VPN典型案例.....	3-15
3.3.1	PE-CE间使用EBGP建立BGP MPLS VPN.....	3-15
3.3.2	PE-CE间使用OSPF建立BGP MPLS VPN.....	3-19
3.3.3	PE-CE间使用RIP建立BGP MPLS VPN.....	3-22
3.3.4	PE-CE间使用静态路由建立BGP MPLS VPN.....	3-25
3.4	MPLS BGP VPN排错帮助.....	3-28
第 4 章	MPLS VPN访问公网.....	4-1
4.1	访问公网简介.....	4-1
4.1.1	非VRF Internet接入方式.....	4-1
4.1.2	VRF Internet接入方式一.....	4-2

4.1.3	VRF Internet接入方式三	4-2
4.2	访问公网配置	4-3
4.3	访问公网典型案例	4-5
4.3.1	非VRF Internet接入方式	4-5
4.3.2	VRF Internet接入方式一	4-9
4.4	访问公网排错帮助	4-13
第 5 章	VPLS	5-1
5.1	VPLS简介	5-1
5.1.1	VPLS基本概念	5-1
5.1.2	VPLS基本网络模型	5-2
5.1.3	VPLS的优点	5-2
5.1.4	PE全连接的基本VPLS网络模型	5-3
5.1.5	分层VPLS模型	5-3
5.1.6	VPLS的报文转发	5-6
5.2	VPLS配置	5-7
5.3	VPLS典型案例	5-10
5.3.1	全连接的VPLS配置	5-10
5.3.2	LSP方式接入的H-VPLS	5-13
5.3.3	QinQ方式接入的H-VPLS	5-17
5.3.4	VPWS配置	5-21
5.4	VPLS排错帮助	5-23
第 6 章	MAC-in-MAC	6-1
6.1	MAC-in-MAC简介	6-1
6.1.1	MAC-in-MAC基本概念	6-1
6.1.2	MAC-in-MAC基本网络模型	6-2
6.1.3	MAC-in-MAC报文封装	6-3
6.1.4	MAC-in-MAC报文转发	6-4
6.1.5	MAC-in-MAC的优点	6-5
6.2	MAC-in-MAC配置	6-5
6.3	MAC-in-MAC典型案例	6-7
6.3.1	基本的MAC-in-MAC应用场景	6-7
6.4	MAC-in-MAC排错帮助	6-9

第1章 MPLS概述

1.1 MPLS概述

MPLS (Multiprotocol Label Switching, 多协议标签交换) 起源于 IPv4, 最初是为了提高转发速度而提出的, 其核心技术可扩展到多种网络协议, 包括 IPv6 (Internet Protocol version 6)、IPX (Internet Packet Exchange)、Appletalk、DECnet、CLNP (Connectionless Network Protocol) 等。MPLS 中的 “Multiprotocol” 指的就是支持多种网络协议。MPLS 技术集二层的快速交换和三层的路由转发于一体, 可以满足各种新应用对网络的要求。

1.1.1 MPLS简介

转发等价类

MPLS 作为一种分类转发技术, 将具有相同转发处理方式的分组归为一类, 称为 FEC (Forwarding Equivalence Class, 转发等价类)。相同 FEC 的分组在 MPLS 网络中将获得完全相同的处理。FEC 是一组三层报文, 它们在同样的路径上、按照相同的转发待遇、以相同的模式被转发。转发决定可以分为两步:

- ☞ 分析分组头并将分组分成 FEC;
- ☞ 将 FEC 映射到下一跳。

传统 IP 转发网络中, 每个路由器对相同分组都要进行 FEC 分类和选择下一跳。FEC 可以包含一个或多个 FEC 单元, 每个 FEC 单元是一组可以映射到相同 LSP 的三层报文。

目前, 定义了两种类型的 FEC 单元:

- ☞ 地址前缀型(Address Prefix): 使用地址前缀作为 FEC 单元, 其长度为从 0 到完整的地址。前缀型 FEC 单元对应于一个目标子网;
- ☞ 主机地址型(Host Address): 使用终端地址作为一个 FEC 单元, 对应于一个终端地址。

FEC 的划分方式非常灵活, 可以是源地址、目的地址、源端口、目的端口、协议类型、VPN 等的任意组合。例如, 在传统的采用最长匹配算法的 IP 转发中, 到同一个目的地址的所有报文就是一个 FEC。

标签

MPLS 网络中, 在边缘 LSR 处每一特定 FEC 都被编码为一个短而定长的值, 称为标签, 标记加在分组前成为标签分组, 再转发到下一跳。标签中, 除了表示 FEC 的标记, 还可以(但不是必须)包含 COS 字段。此时, 标签就成为代表 FEC、优先级和服务类型的综合体。LSR 可以将到达不同端口的分组分配到不同的 FEC, 这就形成了建立 VPN 的基础。当 LSR 创建一个新的 FEC, LSR 就创建一个对应的标签, 并把标签分发给其所有 peer。LSR 维护输入和输出两个方向的标签。假如要进行负载分担, 对应一个 FEC 可能会有多个标签, 但是一个标签只能代表一个 FEC。

标签由报文的头部所携带, 不包含拓扑信息, 只具有局部意义。标签的长度为 4 个字节, 封装结构如图 1-1 所示:



图 1-1 标签的封装结构

标签共有 4 个域：

- ☞ Label: 标签值字段，长度为 20bits，用于转发的指针；
- ☞ Exp: 3bits，用于 QoS；
- ☞ S: 1bit，MPLS 支持标签的分层结构，即多重标签。值为 1 时表明为最底层标签；
- ☞ TTL: 8bits，和 IP 分组中的 TTL（Time To Live，生存时间）意义相同。

标签与 ATM 的 VPI/VCI 以及 Frame Relay 的 DLCI 类似，是一种连接标识符。如果链路层协议具有标签域，如 ATM 的 VPI/VCI 或 Frame Relay 的 DLCI，则标签封装在这些域中；如果不支持，则标签封装在链路层和 IP 层之间的一个垫层中。这样，标签能够被任意的链路层所支持。

标签空间

LSR 可以基于输入端口为一个 FEC 分配不同的标签。这样，从不同端口输出的分组就可以独立转发，这是实现 VPN 的基础。为了提高标签利用率，MPLS 提出了标签空间的概念。标签空间是一个标签前缀，在分配标签时，对于不同标签空间的转发等价类，可以分配相同的标签，这样就相当于扩展了标签的范围。标签空间只对标签的分配有意义，对于标签转发无意义。

标签交换

在非边缘 LSR，不再需要分析分组头，而是用标签作为指针，指向下一跳的输出端口和一个新的标签，标签分组用新标签替代旧标签后经指定的输出端口转发。标签交换不仅可以简化转发、提高速度，还可以实现 VPN，QOS，流量工程等应用。

标签交换路由器

LSR（Label Switching Router，标签交换路由器）是 MPLS 网络中的基本元素，所有 LSR 都支持 MPLS 技术。

LSR 是一种可以基于分组中的标签值来转发分组的设备。连接 IP 路由网和 MPLS 交换网的 LSR，被称为边缘 LSR，它对 IP 报文添加标签，然后按照 LSP 转发数据；或者删除 MPLS 分组标签，按照 IP 路由转发数据。每个 LSR 必须分配一个全局唯一的标识符(LSR ID)，通常取 LSR 一个接口的 IP 地址。假设，LSR Ru 和 Rd 对标签 L 和 FEC F 之间的映射关系达成一致，分组可以利用标签 L 从 Ru 转发到 Rd，则把 Ru 称为上游 LSR，Rd 称为下游 LSR。也就是说，分组总是从上游 LSR 向下游 LSR 转发。

标签交换路径

一个转发等价类在 MPLS 网络中经过的路径称为标签交换路径 LSP（Label Switched Path）。在一条 LSP 上，沿数据传送的方向，相邻的 LSR 分别称为上游 LSR 和下游 LSR。如图 1-2 中，R2 为 R1 的下游 LSR，相应的，R1 为 R2 的上游 LSR。

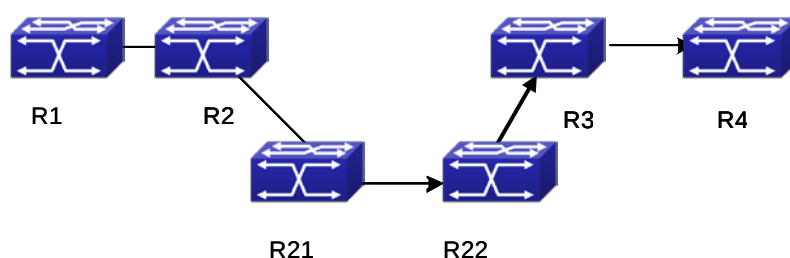


图 1-2 标签交换路径 LSP

LSP 在功能上与 ATM 和帧中继（Frame Relay）的虚电路相同，是从 MPLS 网络的入口到出口的一个单向路径。LSP 中的每个节点由 LSR 组成。

下游 LSR 向上游 LSR 分发标签，连成一串的标签和 LSR 序列就构成 LSP。LSP 是把 IP 层路由信息映射到链路层的一条交换路径。LSP 是一条单向分组转发路径，分组总是从上游 LSR 转发到下游 LSR。需要反方向转发的分组，必须建立另外一条完全独立的 LSP。LDP 总是把 FEC 和 LSP 关联起来，FEC 和 LSP 的关联关系被称为报文映射到 LSP。

1. 映射报文到 LSP 需要遵循的规则：

- (1) 如果只有一个 LSP，它包含一个终端地址型的 FEC 单元和报文的地址相同，则报文被映射到这个 LSP；
- (2) 如果存在多个 LSP 满足条件 1，则报文被映射到这些 LSP 中的一个；
- (3) 如果只有一个 LSP，报文可以匹配其包含的地址前缀型 FEC 单元，则报文被映射到这个 LSP；
- (4) 如果存在多个 LSP 满足条件 3，则根据最长匹配原则选择 LSP；
- (5) 如果能够确定一个分组必须通过某个特定出口 LSR，并且存在一条 LSP，它绑定的 FEC 的前缀型成员就是出口 LSR 的地址，则报文被映射到这条 LSP。

2. 附加规则：

- (1) 如果报文地址没有匹配任何 LSP，报文可以在与其出口节点(Egress Router)地址相同的包含地址前缀型 FEC 单元的 LSP 上发送；
- (2) 如果报文匹配了两个 LSP，一个 LSP 包含终端地址型的 FEC 单元，一个 LSP 包含地址前缀型 FEC 单元，报文总是映射到前者；
- (3) 如果报文没有匹配包含终端地址型的 FEC 单元的 LSP，即使终端地址型的 FEC 单元与其出口节点地址相同，也不能在这个 LSP 上发送。
- (4) LSP 的创建是基于连接的，这种连接是根据拓扑信息，而不是数据流的要求建立的。也就是说，不管是否存在通过该路由转发的数据，LSP 都将被建立。

标签合并

LSR 把多个入标签映射到相同的 FEC，这些入标签对应的出标签和出接口都将是相同的。这样当携带不同标签的分组到达 LSR 时，输出分组将携带相同的标签，这种情况被称为标签合并。标签合并可以减少 MPLS 域的标签数量，但可能丢失分组入接口信息。

如果 LSR 不支持标签合并，对于多个标签请求，无论它们的 FEC 是否相同，都将向下游 LSR 发起新的标签请求。如果 LSR 支持标签合并，对于 FEC 的相同标签请求，LSR 只向下游 LSR 发起一次标签请求。

标签分发协议

标签分发协议 LDP（Label Distribution Protocol）是 MPLS 的控制协议，它相当于传统网络中的信令协议，负责 FEC 的分类、标签的分配以及 LSP 的建立和维护等一系列操作。

MPLS 可以使用多种标签发布协议，包括专为标签发布而制定的协议，例如：LDP、CR-LDP（Constraint-Based Routing using LDP，基于约束路由的 LDP）；也包括现有协议扩展后支持标签发布的协议，例如：BGP（Border Gateway Protocol，边界网关协议）、RSVP（Resource Reservation Protocol，资源预留协议）。同时，还可以手工配置静态 LSP。

LSP 隧道技术

MPLS 支持 LSP 隧道技术。一条 LSP 的上游 LSR 和下游 LSR，尽管它们之间的路径可能并不在路由协议所提供的路径上，但是 MPLS 允许在它们之间建立一条新的 LSP，这样，上游 LSR 和下游 LSR 分别就是这条 LSP 的起点和终点。这时，上游 LSR 和下游 LSR 间的 LSP 就是 LSP 隧道，它避免了采用传统的网络层封装隧道。如图 1-2 中 LSP<R2→R21→R22→R3>就是 R2、R3 间的一条隧道。

如果隧道经由的路由与从路由协议取得的路由一致，这种隧道就称为逐跳路由隧道（Hop-by-Hop Routed Tunnel）；否则称为显式路由隧道（Explicitly Routed Tunnel）。

多层标签栈

如果分组在超过一层的 LSP 隧道中传送，就会有多层标签，形成标签栈（Label Stack）。在每一隧道的入口和出口处，进行标签的入栈（PUSH）和出栈（POP）操作。

标签栈按照“后进先出”（Last-In-First-Out）方式组织标签，MPLS 从栈顶开始处理标签。

MPLS 对标签栈的深度没有限制。若一个分组的标签栈深度为 m ，则位于栈底的标签为 1 级标签，位于栈顶的标签为 m 级标签。未压入标签的分组可看作标签栈为空（即标签栈深度为零）的分组。

1.1.2 MPLS 网络简介

MPLS 网络结构

如图 1-3 所示，MPLS 网络的基本构成单元是 LSR，由 LSR 构成的网络称为 MPLS 域。

位于 MPLS 域边缘、连接其它用户网络的 LSR 称为 LER（LER，Label Edge Router，边缘 LSR），区域内部的 LSR 称为核心 LSR。核心 LSR 可以是支持 MPLS 的路由器，也可以是由 ATM 交换机等升级而成的 ATM-LSR。域内部的 LSR 之间使用 MPLS 通信，MPLS 域的边缘由 LER 与传统 IP 技术进行适配。

分组在入口 LER 被压入上标签后，沿着由一系列 LSR 构成的 LSP 传送，其中，入口 LER 被称为 Ingress，出口 LER 被称为 Egress，中间的节点则称为 Transit。

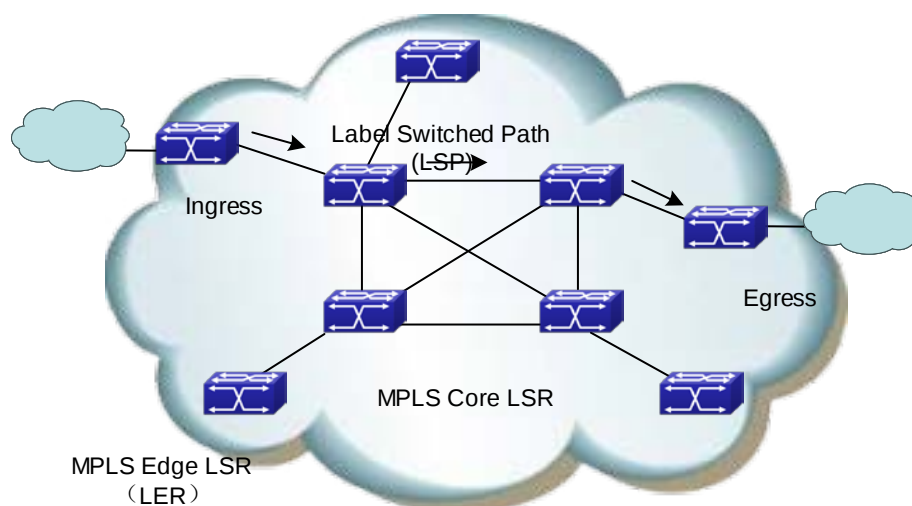


图 1-3 MPLS 网络结构

结合图 1-3 简要介绍 MPLS 的基本工作过程：

首先，LDP 和传统路由协议（如 OSPF、ISIS 等）一起，在各个 LSR 中为有业务需求的 FEC 建立路由表和标签信息表（Label Information Base, LIB）；

入口 LER 接收分组，完成第三层功能，判定分组所属的 FEC，并给分组加上标签，形成 MPLS 标签分组；

接下来，在 LSR 构成的网络中，LSR 根据分组上的标签以及标签转发表（Label Forwarding Information Base, LFIB）进行转发，不对标签分组进行任何第三层处理；

最后，在 MPLS 出口 LER 去掉分组中的标签，继续进行后面的 IP 转发。

由此可以看出，MPLS 并不是一种业务或者应用，它实际上是一种隧道技术，也是一种将标签交换转发和网络层路由技术集于一身的路由与交换技术平台。这个平台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。

1.1.3 MPLS与路由协议简介

LDP 通过逐跳方式建立 LSP 时，利用沿途各 LSR 路由转发表中的信息来确定下一跳，而路由转发表中的信息一般是通过 IGP、BGP 等路由协议收集的。LDP 并不直接和各种路由协议关联，只是间接使用路由信息。

另一方面，通过对 BGP、RSVP 等已有协议进行扩展，也可以支持 MPLS 标签的分发。

在 MPLS 的应用中，也可能需要对某些路由协议进行扩展。例如，基于 MPLS 的 VPN 应用需要对 BGP 进行扩展，使 BGP 能够传播 VPN（Virtual Private Network，虚拟专用网）的路由信息；基于 MPLS 的 TE（Traffic Engineering，流量工程）需要对 OSPF 或 IS-IS 协议进行扩展，以携带链路状态信息。

1.1.4 MPLS应用简介

最初，MPLS 技术结合了二层交换技术和三层路由技术，提高了路由查找速度。但是，随着 ASIC（Application-Specific Integrated Circuit，专用集成电路）技术的发展，路由查找速度已经不成阻碍网络发展的瓶颈。这使得 MPLS 在提高转发速度方面不具备明显的优势。

但由于 MPLS 结合了 IP 网络强大的三层路由功能和传统二层网络高效的转发机制，在转发平面采用面向连接方式，与现有二层网络转发方式非常相似，这些特点使得 MPLS 能够很容易地实现 IP 与 ATM、帧中继等二层网络的无缝融合，并为 QoS（Quality of Service，服务质量）、TE、VPN 等应用提供更好的解决方案。

基于 MPLS 的 VPN

传统的 VPN 一般是通过 GRE、L2TP、PPTP 等隧道协议来实现私有网络间数据流在公网上的传送，LSP 本身就是公网上的隧道，因此，用 MPLS 来实现 VPN 有天然的优势。

基于 MPLS 的 VPN 就是通过 LSP 将私有网络的不同分支联结起来，形成一个统一的网络。基于 MPLS 的 VPN 还支持对不同 VPN 间的互通控制。

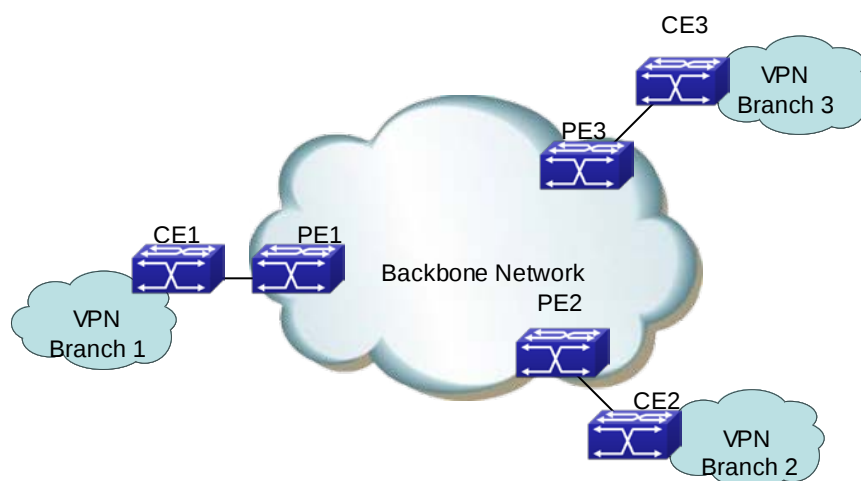


图 1-4 基于 MPLS 的 VPN

图 1-4 是基于 MPLS 的 VPN 的基本结构：CE（Customer Edge，用户边缘设备）可以是路由器，也可以是交换机或主机；PE（Provider Edge，服务商边缘路由器）位于骨干网络。

PE 负责对 VPN 用户进行管理、建立各 PE 间 LSP 连接、同一 VPN 用户各分支间路由分派。PE 间的路由分派通常是用 LDP 或扩展的 BGP 协议实现。

基于 MPLS 的 VPN 支持不同分支间 IP 地址复用，并支持不同 VPN 间互通。与传统的路由相比，VPN 路由中需要增加分支和 VPN 的标识信息，这就需要对 BGP 协议进行扩展，以携带 VPN 路由信息。

基于 MPLS 的流量工程

基于 MPLS 的 TE 和差分服务 Diff-serv 特性，在保证网络高利用率的同时，可以根据不同数据流的优先级实现差别服务，从而为语音、视频等数据流提供有带宽保证的低延时、低丢包率的服务。

由于全网实施流量工程的难度比较大，因此，在实际的组网方案中往往通过差分服务模型来实施 QoS。

Diff-Serv 的基本机制是在网络边缘，根据业务的服务质量要求将该业务映射到一定的业务类别中，利用 IP 分组中的 DS 字段（由 ToS 域而来）唯一的标记该类业务，然后，骨干网络中的各节点根据该字段对各种业务采取预先设定的服务策略，保证相应的服务质量。

Diff-Serv 的这种对服务质量的分类和标签机制和 MPLS 的标签分配十分相似，事实上，

基于 MPLS 的 Diff-Serv 就是通过将 DS 的分配与 MPLS 的标签分配过程结合来实现的。

1.1.5 MPLS PHP

在 MPLS 网络中，核心 LSR 根据分组上的标签进行转发；在 Egress 节点（出口 LER）去掉分组中的标签，继续进行 IP 转发。

实际上，在比较简单的 MPLS 应用中，Egress 节点只需要进行 IP 转发，标签已经没有使用价值。这种情况下，可以利用倒数第二跳弹出特性 PHP（Penultimate Hop Popping），在倒数第二个节点处将标签弹出，Egress 节点就不再进行标签操作了。

第2章 LDP

2.1 LDP简介

LDP 协议是在 MPLS 标签交换环境中用于标签分配的协议。LDP 协议只运行在具有标签交换能力的网络上。LDP 与传统路由算法相结合, 通过在 TCP 连接上传送各种消息, 分配标签、发布 <标签, FEC>映射, 建立维护标签转发表和标签交换路径(LSP)。在 MPLS VPN 环境下, LDP 协议用于分配公网标签。

LDP 协议本身不产生路由, 它从系统获得路由信息, 为路由分配标签, 并把标签通告给其上游路由器。同时, 对于具有下游的 FEC, LDP 协议从其从下游接收标签作为出标签, 从而建立标签交换的通路, 即建立一个把入标签交换为出标签的表项。如果下游分配的标签为 3, 则 LDP 产生一个弹出标签的表项。

LDP 协议定义在 RFC3036, 最新的标准为 RFC5036, 它通过对等体间的 TCP 连接交换各自的标签和路由的对应关系, 其发现邻居的方式包括基本方式(自动发现)和扩展方式(指定)两种。对等体的自动发现使用发往所有路由器(224.0.0.2)的 UDP 组播报文, 其使用的 TCP 和 UDP 报文的端口号都是 646。

其主要运行过程如下:

- 邻居发现与维护: 在全局和指定接口上启动 LDP 协议后, LDP 协议在指定接口上向外发送组播 Hello 报文(除非配置了不采用组播报文发现邻居)向网络通告自己的存在。Hello 报文中携带了自己的传输地址, 即用于建立 TCP 连接的地址。在收到其它 LSR 发出的 Hello 报文后, 创建邻居关系; 邻居关系依靠周期性发送 Hello 报文来维护。
- 会话建立与维护: LDP 会话是建立在 TCP 连接之上的; 首先比较对端 Hello 报文中携带的传输地址和本端的传输地址, 地址比较大的成为 ACTIVE, 地址比较小的成为 PASSIVE, ACTIVE 的路由器主动发起连接请求以建立 TCP 连接(这是为了避免类似 BGP 邻居的连接冲突问题)。TCP 连接建立后, 连接双方发送初始化报文以协商会话参数, 协商一致建立会话关系。会话关系后互相把本地的接口地址列表以及标签信息发送给自己的邻居。没有信息时通过发送 KEEPALIVE 消息以保持连接。
- LSP 的建立与维护: 每两个 LSR 对等体之间需要一个会话关系来交换标签信息, LSR 之间通过交换 FEC 和标签绑定消息来建立 LSP。
- 会话的撤销: 若长时间收不到对端的消息, LDP 会中断连接并发出 notification 消息通知对端关闭会话连接。

值得注意的是, LDP 协议不会为默认路由分配标签, 另外, 除非特别指定, LDP 协议也不会为 BGP 路由分配标签。

2.1.1 LDP基本概念

LDP 对等体

LDP 分配标签给 FEC 的时候, 需要把此标签及其意义在 MPLS 网络内传播, 从而建立

LSP。采用 LDP 交换标签信息的 LSR 被称为 LDP 对等体(LSP Peer)。LDP 对等体通过它们之间的 LDP 会话获得对方的标签映射消息以及其他消息。

LDP 会话

交换 LDP Discovery Hello 报文之后，LSR 之间就会建立 LDP 会话。LDP 会话用于在 LSR 之间交换标签映射、释放等各种消息。

LDP 会话的建立可以分为两步：

- ✎ 建立传输连接
- ✎ 初始化会话

LDP 会话可以分为两种类型：

- ✎ 本地 LDP 会话（Local LDP Session）：建立会话的两个 LSR 之间是直连的；
- ✎ 远端 LDP 会话（Remote LDP Session）：建立会话的两个 LSR 之间是非直连的。

LDP 消息类型

LDP 协议主要使用四类消息：

- ✎ 发现（Discovery）消息：用于通告和维护网络中 LSR 的存在；
- ✎ 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、改变和删除标签到 FEC 的映射；
- ✎ 通知（Notification）消息：用于提供建议性的消息和差错通知。

为保证 LDP 消息的可靠发送，除了发现阶段使用 UDP 传输外，LDP 的 Session 消息、Advertisement 消息和 Notification 消息都使用 TCP 传输。

2.1.2 LDP报文格式简介

LDP PDU

LDP PDU 包括 LDP 报文头和若干个 LDP 消息。LDP 报文头格式如下：

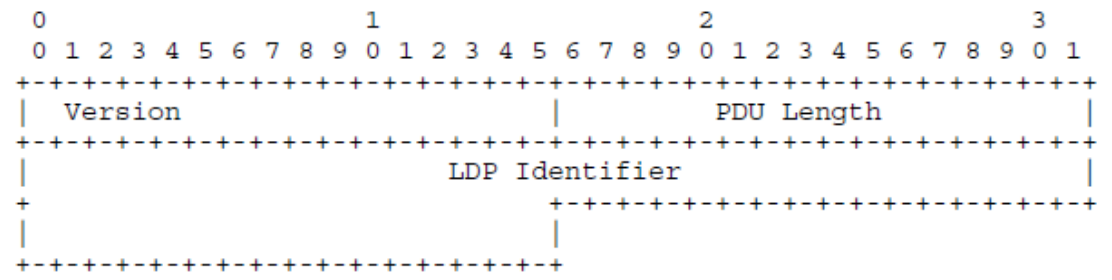


图 2-1 LDP 的 TLV 格式

- ✎ Version: 版本号，1 字节。当前 LDP 版本号为 1；
- ✎ PDU Length: LDP 报文的总长度，2 字节。单位是字节；
- ✎ LDP ID: LDP 标识符，6 字节。前 4 字节是全局唯一的 LSR ID，后 2 字节是标签空间标识符。对于全局标签空间，这个字段为 0。

TLV 编码

LDP 采用 TLV(Type-Length-Value)方式封装 LDP 消息中的参数。LDP TLV 格式如下：

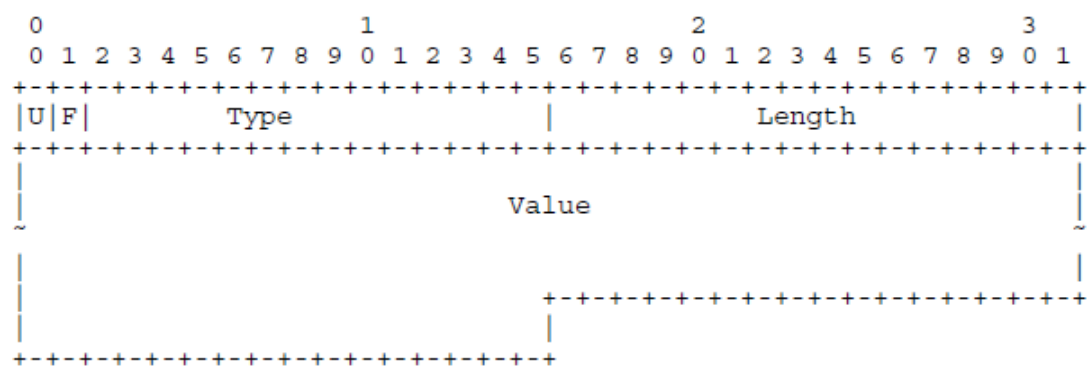


图 2-2 LDP 的 TLV 格式

- U bit: 未知标志，1bit。如果 U 标志置 0，LSR 必须向发送报文的源 LSR 发送通知，并忽视整个消息；如果 U 标志置 1，忽视此 TLV 参数，正常解析其它 TLV 参数；
- F bit: 转发未知 TLV 标志，1。此标志只适用于未知 TLV 且 U bit 置 1 的 LDP 消息。如果 F 标志置 0，不再转发未知 TLV 参数；如果 F 标志置 1，转发未知 TLV 参数；
- Type: 类型，14bits；
- Length: 长度，1 字节。TLV 数值字段的长度；
- Value: 数值字段。其长度由参数 Length 定义；
- TLV 的数值字段还可以包含 TLV 参数，也就是说 TLV 可以嵌套。TLV 的第一个字节病不要求字节对齐。

目前定义的 TLV 类型：

表 2-1 LDP 报文中 TLV 类型

TLV	Type
FEC	0x0100
Address List	0x0101
Hop Count	0x0103
Path Vector	0x0104
Generic Label	0x0200
ATM Label	0x0201
Frame Relay Label	0x0202
Status	0x0300
Extended Status	0x0301
Returned PDU	0x0302
Returned Message	0x0303
Common Hello Parameters	0x0400
IPv4 Transport Address	0x0401
Configuration Sequence Number	0x0402

IPv6 Transport Address	0x0403
Common Session Parameters	0x0500
ATM Session Parameters	0x0501
Frame Relay Session Parameters	0x0502
Label Request Message ID	0x0600
Vendor-Private	0x3E00- 0x3EFF
Experimental	0x3F00- 0x3FFF

2.1.3 LDP 标签管理

在 MPLS 体系中，由下游 LSR 决定将标签分配给特定 FEC，再通知上游 LSR。即，标签由下游指定，标签的分配按从下游到上游的方向。

标签分发模式（Label Advertisement Mode）

在 MPLS 域中，分组从上游 LSR 经过标签交换打上下游 LSR 的标签转发至下游 LSR。下游 LSR 为 FEC 分配的标签只在本地和其上游 LSR 之间有意义，下游 LSR 分配的标签必需传播到上游 LSR。对于标签从下游 LSR 到上游 LSR 的传播方式问题，MPLS 定义了两种标签分发模式：

- ☞ 下游按需模式 DoD（Downstream On Demand）：对于一个特定的 FEC，LSR 从上游获得标签请求消息之后才进行标签分配与分发；
- ☞ 下游自主模式 DU（Downstream Unsolicited）：对于一个特定的 FEC，LSR 无须从上游获得标签请求消息即进行标签分配与分发，会主动发送标签映射消息，通知上游 LSR。

这两种模式可以混合适用，LSR 的每个接口可以独立配置为这两种标签分发模式之一。在初始化时，上游 LSR 和下游 LSR 必须交换标签分发模式信息，就标签分发模式达成一致，否则 LSP 无法正常建立。

如下图 2-3 为 LDP 标签分发示意图。

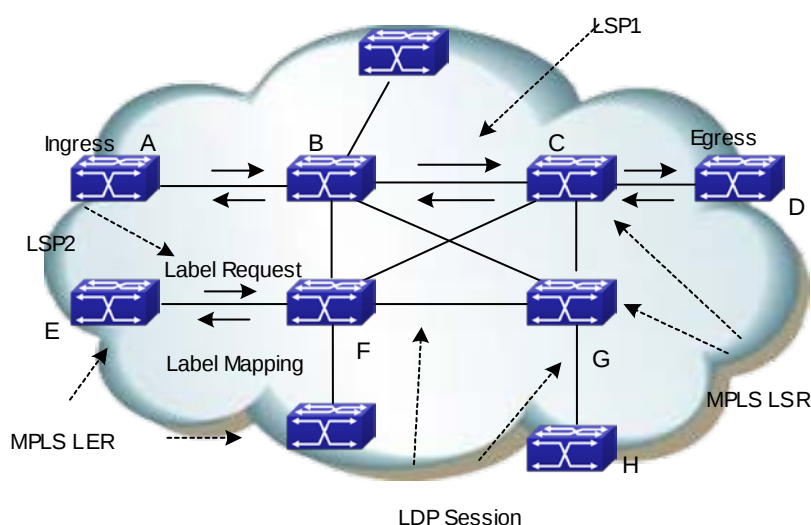


图 2-3 标签分发过程

例如，图 2-3 中的 LSP1 上，LSR B 为 LSR C 的上游 LSR，LSR C 则是 LSR B 的下游 LSR。

标签的分发过程有两种模式，主要区别在于标签映射的发布是上游请求（DoD）还是下游主动发布（DU）。

下面分别详细描述这两种模式的标签分发过程：

(1) DoD（downstream-on-demand）模式分发过程

上游 LSR 向下游 LSR 发送标签请求消息（Label Request Message），其中包含 FEC 的描述信息。下游 LSR 为此 FEC 分配标签，并将映射的标签通过标签映射消息（Label Mapping Message）反馈给上游 LSR。

下游 LSR 何时反馈标签映射消息，取决于该 LSR 采用的标签分配控制方式。

- 1) 采用 Ordered 方式时，只有收到它的下游返回的标签映射消息后，才向其上游发送标签映射消息；
- 2) 采用 Independent 方式时，不管有没有收到它的下游返回的标签映射消息，都立即向其上游发送标签映射消息。

上游 LSR 一般是根据其路由表中的信息来选择下游 LSR。在图 2-3 中，LSP1 沿途的 LSR 都采用 Ordered 方式，LSP2 上的 LSR F 则采用 Independent 方式。

(2) DU（downstream unsolicited）模式分发过程

下游 LSR 在 LDP 会话建立成功后，主动向其上游 LSR 发布标签映射消息。上游 LSR 保存标签映射信息，并根据保有模式来处理收到的标签映射信息。

标签控制模式（Label Distribution Control Mode）

在 MPLS 域中，LSR 之间通过交换标签建立一条从输入节点（ingress router）到输出节点（egress router）的标签转发路径 LSP。LSP 是沿着 MPLS 域中 IGP 建立的路由转发路径建立的，只有一条完整的路径才对分组转发有意义。建立 LSP 的过程，也就是 LSR 分发标签的过程；控制 LSP 的建立，就是控制 LSR 的标签分发。对于下游 LSR 何时向上游 LSR 分发标签的问题，MPLS 定义了两种 LSP 控制模式：

- ☞ 有序模式(Ordered)：对于 LSR 上某个 FEC 的标签映射，只有当该 LSR 已经具有此 FEC 下一跳的标签映射消息、或者该 LSR 就是此 FEC 的出口节点时，该 LSR

才可以向上游发送此 FEC 的标签映射。一个流的标签分发从这个 FEC 流所属的出口节点开始，由下游向上游逐级绑定，这样可以保证整个网络内标记与流的映射完整一致。有序方式提供了更好的环路预防能力。

- ☞ 独立模式(Independent): LSR 不必等待 FEC 下一跳的标签就可以把 FEC 标签分发给它的 peer。LSR 可以在任意时间向与它连接的 LSR 通告标签映射。这种方式可能导致在收到下游标签之前就向上游发布了标签。独立方式可以提高 LSP 的建立和汇聚速度。

LSR 作为出口节点的条件:

- ☞ FEC 引用了 LSR 本身的地址;
- ☞ FEC 的下一跳路由器位于标签交换网之外;
- ☞ FEC 的单元经过路由区域, 例如另一个 OSPF SUMMARY 域, 或者 OSPF、BGP 的另一个自治系统。

标签保有模式 (Label Retention Mode)

标签保有(retention)模式指 LSR 对收到的、但目前暂时用不到的标签到 FEC 映射的处理方式。在下游主动标签分发模式下, 上游 LSR 可能接收到下游 LSR 通知的大量<FEC、标签>映射组, 只有当映射组中的 FEC 是上游 LSR 本地 FEC 的下一跳时, 这个映射组才对标签转发有意义。对于如何处理暂时没有转发意义的映射组, MPLS 定义了两种标签保有模式:

- ☞ 保守模式(Conservative): 对于从邻居 LSR 收到的标签映射, 无论邻居 LSR 是不是自己的下一跳都保留。保守模式的优点是: 只创建和维护对数据转发有意义的标签, 在标签空间有限的情况下(ATM 交换), 非常重要。
- ☞ 自由模式(Liberal): 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。自由模式的优点是: 处理路由变化的开销非常小; 缺点是: 通告和维护了许多暂时无效的标签。

使用自由标签保持方式, LSR 能够迅速适应路由变化; 而使用保守标签保持方式, LSR 可以分配和保存较少的标签数量。保守标签保持方式通常与 DoD 方式一起, 用于对于标签空间有限的 LSR。

标签交换中的几个基本概念

- ☞ NHLFE (Next Hop Label Forwarding Entry): 下一跳标签转发项。用于描述对标签执行的操作, 包括入栈 (Push) 操作和交换 (Swap) 操作。
- ☞ FTN (FEC to NHLFE map): 在 Ingress 节点将转发等价类 FEC 映射到 NHLFE 的过程。
- ☞ ILM (Incoming Label Map): 入标签映射。对于接收的标签分组, LSR 将标签映射到 NHLFE 的过程。

标签交换的过程

入口 LER (Ingress) 将进入网络的分组划分成转发等价类 FEC。属于相同 FEC 的分组在 MPLS 域中将经过相同的路径, 即使用同一条 LSP。LSR 对到来的 FEC 分组分配一个标签, 然后从相应的接口转发出去。

标签交换的具体操作描述如下:

- ☞ LSP 沿途的所有 LSR 都先建立 ILM, 入标签是根据 ILM 表项进行映射的;
- ☞ 对于接收到的标签分组, LSR 将标签映射到 NHLFE;
- ☞ LSR 只需根据标签从标签信息表中找到相应的 NHLFE, 用新的标签替换原来的标签, 然后对标签分组继续转发。

2.1.4 LDP 会话

按照先后顺序，LDP 会话建立过程主要包括以下四个阶段：

- ☞ 发现阶段
- ☞ 会话建立与维护
- ☞ LSP 建立
- ☞ 会话撤销

发现阶段

在这一阶段，希望建立会话的 LSR 向相邻 LSR 周期性地发送 Hello 消息，通知相邻节点自己的存在。在基本发现机制下，通过这一过程，LSR 可以自动发现它的 LDP 对等体，而无需进行手工配置。LDP 有两种发现机制：

- ☞ 基本发现机制

基本发现机制用于发现本地的 LDP 对等体，即通过链路层直接相连的 LSR，建立本地 LDP 会话。这种方式下，LSR 周期性以 UDP 报文形式从接口发送 LDP 链路 Hello 消息（LDP Link Hello），发往标识“子网内所有路由器”的组播地址。

LDP 链路 Hello 消息带有接口的 LDP 标识符及其他相关信息，如果 LSR 在某个接口收到了 LDP 链路 Hello 消息，则表明在该接口（链路层）存在 LDP 对等体。

- ☞ 扩展发现机制

扩展发现机制用于发现远端的 LDP 对等体，即不通过链路层直接相连的 LSR，建立远端 LDP 会话。这种方式下，LSR 周期性以 UDP 报文形式向指定的 IP 地址发送 LDP 目标 Hello 消息（LDP Targeted Hello）。

LDP 目标 Hello 消息带有 LSR 的 LDP 标识符及其他相关信息，如果 LSR 收到 LDP 目标 Hello 消息，则表明在网络层存在 LDP 对等体。

会话建立与维护

在第一阶段发现 LDP 对等体之后，LSR 开始建立会话。这一过程又可分为两步：

- ☞ 建立传输层连接，即，在 LSR 之间建立 TCP 连接；
- ☞ 随后对 LSR 之间的会话进行初始化，协商会话中涉及的各种参数，如 LDP 版本、标签分发方式、定时器值、标签空间等。协商成功之后，就在 LSR 之间建立了会话连接

会话建立后，通过不断地发送 Keepalive 消息来维护这个会话。

LSP 建立

LSP 的建立过程实际就是将 FEC 和标签进行映射，并将这种映射通告 LSP 上相邻 LSR。这个过程是通过 LDP 实现的，以 DoD 模式为例，主要步骤如下：

- (1) 当网络的路由改变时，如果有一个边缘节点发现自己的路由表中出现了新的目的地址，并且这一地址不属于任何现有的 FEC，则该边缘节点需要为这一目的地址建立一个新的 FEC。边缘 LSR 决定该 FEC 将要使用的路由，向其下游 LSR 发起标签请求消息，并指明是要为哪个 FEC 申请标签；
- (2) 收到标签请求消息的下游 LSR 记录这一请求消息，根据本地的路由表找出对应该 FEC 的下一跳，继续向下游 LSR 发出标签请求消息；
- (3) 当标签请求消息到达目的节点或 MPLS 网络的出口节点时，如果此节点尚有可供分配的标签，并且判定上述标签请求消息合法，则该节点为 FEC 分配标签，并向上游发出标签映射消息，标签映射消息中包含分配的标签等信息；
- (4) 收到标签映射消息的 LSR 检查本地存储的标签请求消息状态。对于某一 FEC 的标

签映射消息，如果数据库中记录了相应的标签请求消息，LSR 将为该 FEC 进行标签分配，并在其标签转发表中增加相应的条目，然后向上游 LSR 发送标签映射消息；

- (5) 当入口 LSR 收到标签映射消息时，它也需要在标签转发表中增加相应的条目。这时，就完成了 LSP 的建立。

会话撤销

LDP 通过检测 Hello 消息来维持邻接关系。同时 LDP 也通过检测 Keepalive 消息来维持会话的连接，如果一定时间内没有收到 Keepalive 消息，则 LDP 会话就会关闭连接。

每个 LDP 会话可以包含一个或者多个 Hello 邻接。LDP 依靠周期性的 Hellos 报文维持 Hello 邻接关系。如果长时间没有接收到 LDP Discovery Hellos 报文，LDP 就关闭 Hello 邻接。当关闭 LDP 会话中的最后一个 Hello 邻接时，LDP 就发送通知消息，并关闭传输连接。

2.1.5 LDP环路检测

在 MPLS 域中建立 LSP 也要防止产生环路，LDP 环路检测机制可以检测 LSP 环路的出现，并避免发生环路。

如果对 MPLS 域进行环路检测，则必须在所有 LSR 上都配置环路检测。但在建立 LDP 会话时，并不要求双方的环路检测配置一致。

LDP 环路检测有两种方式：

最大跳数

指标签消息(包括标签映射和标签请求)经过的 LSR 的数量。当 LSR 传播包含跳数参数的标签信息时，首先把跳数加 1。当跳数达到配置的最大数值时，即认为出现环路，LSP 建立失败。如果跳数为 0，就表示跳数未知，这类消息，保持跳数为 0 不变。最大跳数的默认值为 255。

路径向量

在传递标签映射（或者标签请求）的消息中记录路径信息，每经过一跳，LSR 就会检查自己的 LSR ID 是否在此记录中。在以下条件之一时认为出现环路，LSP 建立失败：

- ☞ 路径向量记录表中已有本 LSR 的记录；
- ☞ 路径的跳数超过设置的最大值。

如果记录中没有自己的 LSR ID，就会将其添加到该记录中。路径向量的最大值与最大跳数相等。

2.2 LDP配置

LDP 配置任务序列：

1. 全局启动 MPLS（必选）
2. 启动 LDP 协议（必选）
 - (1) 启动 LDP 模块/关闭 LDP 模块
 - (2) 在接口上启动 label-switching/关闭 label-switching
 - (3) 在接口上启动 LDP 模块/关闭 LDP 模块
3. 配置 LDP 协议参数（可选）
 - (1) 配置 LDP 标签管理模式
 - 1) 配置 LDP 标签保有模式

- 2) 配置 LDP 标签分发模式
- 3) 配置 LDP 控制模式
- (2) 配置 LDP 环路检测
 - 1) 配置 LDP 是否进行环路检测
 - 2) 配置 LDP 环路检测时允许的跳数
- (3) 配置 LDP 指定对等体
- (4) 配置 LDP 协议其它参数
 - 1) 配置 LDP 各计时器老化时间或时间间隔
 - 2) 配置 LDP 路由器 ID
 - 3) 配置 LDP 用于 TCP 连接的接口地址
 - 4) 配置 LDP 是否通过组播 Hello 发现对等体
 - 5) 配置 LDP 是否引入 BGP 路由
 - 6) 配置 LDP 是否使能标签合并能力
 - 7) 设置 LDP 是否传播标签释放消息
 - 8) 设置 LDP 在申请标签被拒绝后是否重试
 - 9) 配置 LDP 接收指定目标的 Hello
- 4. 清除 LDP 的连接或邻接关系

1. 全局启动 MPLS（必选）

命令	解释
全局配置模式	
mpls enable no mpls enable	启动 mpls 协议；本命令的 no 操作关闭 MPLS 协议。

2. 启动 LDP 协议（必选）

在 DCNOS 中运行 LDP 协议的基本配置很简单，通常只需打开 LDP 开关、并在准备运行 LDP 的接口配置启动 LDP 即可，但应当注意的是，配置启动 LDP 的接口应当使能标签交换。

命令	解释
全局配置模式	
router ldp no router ldp	打开/关闭 LDP 协议；缺省情况下未使能 LDP。
接口配置模式	
label-switching no label-switching	使能标签交换功能；缺省情况下未使能标签交换。
mpls proxy loopback-group <1-max_agg_num> no mpls proxy loopback-group	启动 MPLS 代理功能，缺省不启动；只在 MPLS 单播报文入方向的板卡不支持 MPLS 功能时使能 MPLS 代理。
ldp {enable disable}	接口上启动/关闭 LDP 协议；接口缺省情况下未使能 LDP。

3. 配置 LDP 协议参数（可选）

(1) 配置 LDP 标签管理模式

- 1) 配置 LDP 标签保有模式
- 2) 配置 LDP 标签分发模式
- 3) 配置 LDP 控制模式

命令	解释
路由器配置模式	
label-retention-mode {conservative liberal}	配置全局标签保有模式，可选择保守模式或自由模式，缺省情况下是自由模式 liberal 。
advertisement-mode {downstream-on-demand downstream- unsolicited}	配置全局标签分发模式，可选择下游主动或下游按需。该模式与另两种模式有搭配关系，标签分发模式改变会同时改变标签保有模式和全局标签路径控制模式。缺省情况下标签分发模式是下游主动模式 downstream-unsolicited 。
control-mode {ordered independent}	配置全局标签路径控制模式，可选择有序模式和独立模式。缺省情况下是独立模式 independent 。
接口配置模式	
ldp label-retention-mode {conservative liberal}	配置所在接口的标签保有模式。缺省情况与全局模式相同。如果配置与全局模式不符，以接口上的配置为准。
ldp advertisement-mode {downstream-on-demand downstream-unsolicited}	配置所在接口的标签分发模式。缺省情况与全局模式相同。如果配置与全局模式不符，以接口上的配置为准。

(2) 配置 LDP 环路检测

- 1) 配置 LDP 是否进行环路检测
- 2) 配置 LDP 环路检测时允许的跳数

命令	解释
路由器配置模式	
loop-detection no loop-detection	设定 LDP 进行环路检测。其 NO 形式取消进行环路检测。
loop-detection-count <count> no loop-detection-count <count>	设定 LDP 进行环路检测时允许的跳数，缺省为 255，其 no 形式恢复为缺省值。

(3) 配置 LDP 指定对等体

命令	解释
路由器配置模式	
targeted-peer <ip-addr> no targeted-peer <ip-addr>	配置 LDP 指定目的地的远端邻居或远端对等体。

(4) 配置 LDP 协议其它参数

- 1) 配置 LDP 各计时器老化时间或时间间隔
- 2) 配置 LDP 路由器 ID
- 3) 配置 LDP 用于 TCP 连接的接口地址
- 4) 配置 LDP 是否通过组播 Hello 发现对等体
- 5) 配置 LDP 是否引入 BGP 路由
- 6) 配置 LDP 是否使能标签合并能力
- 7) 配置 LDP 是否传播标签释放消息
- 8) 配置 LDP 在申请标签被拒绝后是否重试
- 9) 配置 LDP 接收指定目标的 Hello

命令	解释
路由器配置模式	
keepalive-interval <interval> no keepalive-interval <interval>	配置 LDP 保活报文发送的时间间隔, 缺省为 10 秒, 其 no 形式恢复为缺省值。
keepalive-timeout <time-val> no keepalive-timeout <time-val>	配置 LDP 会话的连接保持时间, 缺省为 30 秒, 其 no 形式恢复为缺省值。
Hello-interval <Hello-interval> no Hello-interval <Hello-interval>	配置 LDP 发送组播 HELLO 的时间间隔, 缺省为 5 秒, 其 no 形式恢复为缺省值。
hold-time <hold-time > no hold-time <hold-time >	配置 LDP 组播对等体的保持时间, 缺省为 15 秒, 其 no 形式恢复为缺省值。
targeted-peer-Hello-interval <Hello-interval> no targeted-peer-Hello-interval <Hello-interval>	配置 LDP 发送指定目的地的 HELLO 的时间间隔, 缺省为 15 秒, 其 no 形式恢复为缺省值。
targeted-peer-hold-time <hold-time> no targeted-peer-hold-time <hold-time>	配置 LDP 指定目的地对等体的保持时间, 缺省为 45 秒, 其 no 形式恢复为缺省值。
接口配置模式	
ldp keepalive-interval <interval> no ldp keepalive-interval <interval>	针对指定接口配置 LDP 保活报文发送的时间间隔; 其 no 形式恢复为缺省值。
ldp keepalive-timeout <time-val> no ldp keepalive-timeout <time-val>	针对指定接口配置 LDP 会话的连接保持时间; 其 no 形式恢复为缺省值。
ldp Hello-interval <Hello-interval> no ldp Hello-interval <Hello-interval>	针对指定接口配置 LDP 发送组播 HELLO 的时间间隔; 其 no 形式恢复为缺省值。
ldp hold-time <hold-time> no ldp hold-time <hold-time>	针对指定接口配置 LDP 组播对等体的保持时间; 其 no 形式恢复为缺省值。
ldp targeted-peer-Hello-interval <Hello-interval> no ldp targeted-peer-Hello-interval <Hello-interval>	针对指定接口配置 LDP 发送指定目的地的 HELLO 的时间间隔; 其 no 形式恢复为缺省值。

ldp targeted-peer-hold-time <hold-time> no ldp targeted-peer-hold-time <hold-time>	针对指定接口配置 LDP 指定目的地对等体的保持时间；其 no 形式恢复为缺省值。
路由器配置模式	
router-id <ip-addr> no router-id <ip-addr>	配置 LDP 使用的路由器 ID,缺省为自动获得，其 no 形式取消手动配置的路由器 ID,并自动获得一个满足条件的接口的 IP 地址作为路由器 ID。
transport-address <ip-addr> no transport-address <ip-addr>	配置 LDP 用于进行 TCP 连接的 IP 地址，需要注意的是，该地址必须是主 VRF 上一个环回接口的地址。其 no 形式取消手动的配置而让 LDP 自动选择进行 TCP 连接的地址。
multicast-Hellos no multicast-Hellos	配置 LDP 使用组播 HELLO 来发现对等体，其 no 形式配置不使用组播 HELLO 来发现对等体。缺省使用组播 HELLO。
import-bgp-routes no import-bgp-routes	配置 LDP 引入 BGP 路由，其 no 形式取消引入 BGP 路由。缺省不引入 BGP 路由。
global-merge-capability {merge-capable non-merge-capable} no global-merge-capability {merge-capable non-merge-capable}	配置使能或不使能全局标签合并能力，缺省为使能，其 no 形式恢复为缺省值。
propagate-release no propagate-release	配置标签释放向对等体传播，其 no 形式配置不传播标签的释放，缺省不进行传播。
request-retry no request-retry	配置标签请求失败后进行 5 次重试，缺省不进行重试，其 no 形式取消重试的配置。
request-retry-timeout <time-val> no request-retry-timeout <time-val>	配置进行重试的时间间隔，缺省为 5 秒，其 no 形式恢复为缺省值。
targeted-peer-Hello-receipt no targeted-peer-Hello-receipt	配置 LDP 接收指定目的地的 HELLO,即使本端并未配置该对等体，缺省不接收这样的 HELLO。其 no 形式恢复为不接收。注意如果配置了指定的 LDP 对等体，那么一定要配置 targeted-peer-Hello-receipt

4. 清除 LDP 的连接或邻接关系（可选）

命令	解释
特权模式	

clear ldp adjacency {<ip-addr> *}	清除指定的 LDP 邻接关系，“*”表示所有邻接关系。
clear ldp session {<ip-addr> *}	清除指定的 LDP 会话，“*”表示所有会话。

2.3 LDP 典型案例

LDP 的一些设计是为了适应不同网络环境进行的，在典型的以太网环境的配置十分简单，另外由于硬件系统的发展，特别是三层交换机的普遍应用，纯粹的 MPLS 网络已经没有什么明显的价值，而更多的应用在 MPLS VPN 环境下。

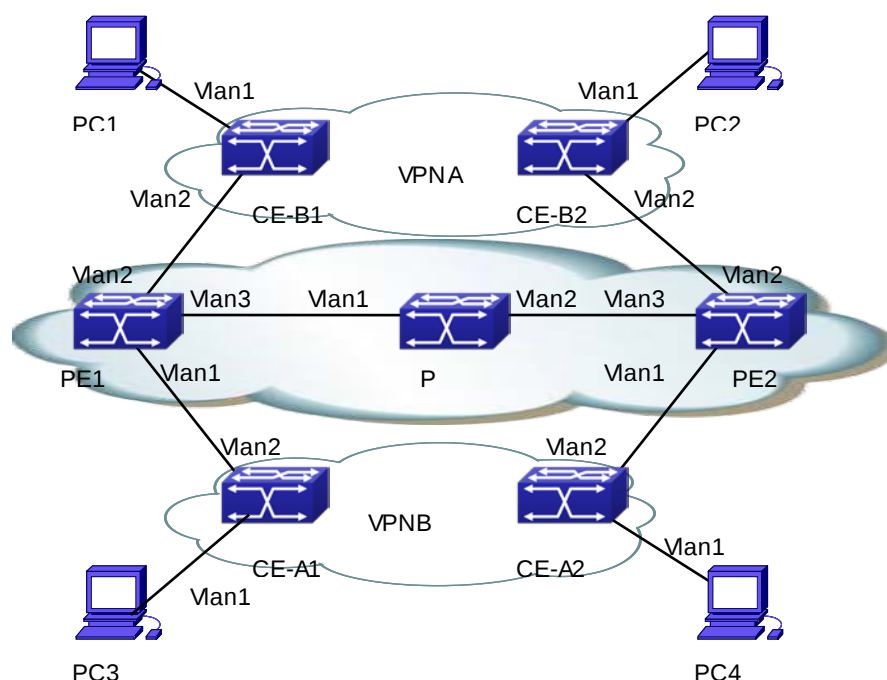


图 2-4 MPLS VPN 典型案例

如图 2-4 是一个典型的 MPLS VPN 案例，其中 PE1、P、PE2 组成了公网区域，也就是使用 MPLS 交换的区域，CE-A1 与 CE-A2 组成了 VPN-A，CE-B1 与 CE-B2 组成了 VPN-B。两个 VPN 都通过公网的标签交换通信，在公网区域，需要配置 LDP 协议来进行标签的分发、分配。为了保证路由的可达性，我们使用 OSPF 协议公布他们各自的路由。

PE1 的 LDP 配置如下：

```

PE1#config
PE1(config)#mpls enable
PE1(config)# router ldp
PE1(config-router)#exit
PE1(config)#interface vlan 3
PE1(config-if-Vlan3)#ip address 202.200.1.2 255.255.255.0
PE1(config-if-Vlan3)#ldp enable
PE1(config-if-Vlan3)#label-switching

```

```
PE1(config-if-Vlan3)#exit
PE1(config)#router ospf
PE1(config-router)#network 200.200.1.1/32 area 0
PE1(config-router)#network 202.200.1.0/24 area 0
PE1(config-router)#exit
```

P 交换机的 LDP 配置如下：

```
P#config
P(config)#mpls enable
P(config)# router ldp
P(config-router)#exit
P(config)#interface vlan 1
P(config-if-Vlan1)#ip address 202.200.1.1 255.255.255.0
P(config-if-Vlan1)#ldp enable
P(config-if-Vlan1)#label-switching
P(config-if-Vlan1)#exit
P(config)#interface vlan 2
P(config-if-Vlan2)#ip address 202.200.2.1 255.255.255.0
P(config-if-Vlan2)#ldp enable
P(config-if-Vlan2)#label-switching
P(config-if-Vlan2)#exit
P(config)#router ospf
P(config-router)#network 202.200.1.0/24 area 0
P(config-router)#network 202.200.2.0/24 area 0
P(config-router)#exit
```

PE1 的 LDP 配置如下：

```
PE2#config
PE2(config)#mpls enable
PE2(config)# router ldp
PE2(config-router)#exit
PE2(config)#interface vlan 3
PE2(config-if-Vlan3)#ip address 202.200.2.2 255.255.255.0
PE2(config-if-Vlan3)#ldp enable
PE2(config-if-Vlan3)#label-switching
PE2(config-if-Vlan3)#exit
PE2(config)#router ospf
PE2(config-router)#network 200.200.1.2/32 area 0
PE2(config-router)#network 202.200.2.0/24 area 0
PE2(config-router)#exit
```

关于 BGP 的配置，详见 BGP VPN 典型案例。

2.4 LDP排错帮助

在配置、使用 LDP 协议时，可能会由于物理连接、配置错误等原因导致 LDP 协议未能

正常运行。因此，用户应注意以下要点：

- ☞ 首先应该保证系统全局启动了 LDP 并且在活动接口上正确启动了 LDP，注意接口使能 LDP 的前提是全局使能了 MPLS；
- ☞ 其次，在连接正确后，使用 `show ldp interface` 命令查看是否相应的接口上 LDP 已经正常启动，如果确信接口上正确配置了启动 LDP，但显示却没有看到，可能是接口没有处于 UP 状态或没有配置接口的标签交换（命令 `label-switching`）；
- ☞ 然后，确认相邻的接口处在同一网段，并通过 `show ldp adjacency` 查看 LDP 是否可以正常发现对等体并建立邻接关系。如果没有发现对等体或者没有成功建立邻接关系，可能是接口没有处在同一网段或者本端和远端邻居之一配置了不发送组播 HELLO。另外在建立 TCP 连接时，缺省使用的地址是 LSR ID，因此，必须确认将 LSR ID 的路由发布给对端。
- ☞ 接着，通过 `show ldp session` 查看 LDP 会话是否建立并处于 operational 状态，只有处于 operational 状态的 LDP 会话才能交换消息。如果 LDP 会话不能建立，使用 `show ldp` 命令查看正在双方使用的用于 TCP 连接的地址并查看路由表，确认对端路由可达。
- ☞ 最后，在上述过程都正确的情况下，使用 `show ldp fec` 命令查看 LDP 引入了哪些路由及这些路由的相关信息，也可以通过 `show mpls ftn` 和 `show mpls ilm` 命令查看其建立的表项。
- ☞ 另外，如果是配置了指定目的地的 LDP 对等体，则要确保对端也配置了以本机为目的的 LDP 对等体或者配置了可以接收指定目的地的 HELLO。并且双方指定的地址要路由可达。

第3章 MPLS VPN

3.1 BGP/MPLS VPN简介

3.1.1 BGP/MPLS VPN网络结构

BGP/MPLS VPN是提供商 VPN 解决方案中一种基于 PE 的 L3VPN 技术,它使用 BGP 在服务提供商骨干网上发布 VPN 路由,使用 MPLS 在服务提供商骨干网上转发 VPN 报文。

BGP/MPLS VPN 组网方式灵活、可扩展性好,并能够方便地支持 MPLS QoS 和 MPLS TE, 因此得到越来越多的应用。

BGP/MPLS VPN 模型由三部分组成: CE、PE 和 P。

- ☞ P 路由器 (Provide Router): 供应商路由器。位于 MPLS 域的内部。可以基于标签交换快速转发 MPLS 数据流。P 路由器接收 MPLS 报文,交换标签后,输出 MPLS 报文。
- ☞ PE 路由器 (Provide Edge Router): 供应商边界路由器。位于 MPLS 域的边界,用于转换 IP 报文和 MPLS 报文。PE 路由器接收 IP 报文,压入 MPLS 标签后,输出 MPLS 报文;并且接收 MPLS 报文,弹出标签之后,输出 IP 报文。PE 路由器上,与其它 P 路由器或者 PE 路由器连接的端口被称为“公网端口”,配置公网 IP 地址;与 CE 路由器连接的端口被称为“私网端口”,配置私网 IP 地址。
- ☞ CE 路由器 (Customer Edge Router): 用户边界路由器。位于用户 IP 域边界,直接和 PE 路由器连接,用于汇聚用户数据,并把用户 IP 域的路由信息转发到 PE 路由器。

图 3-1 是一个 BGP/MPLS VPN 组网方案的示意图:

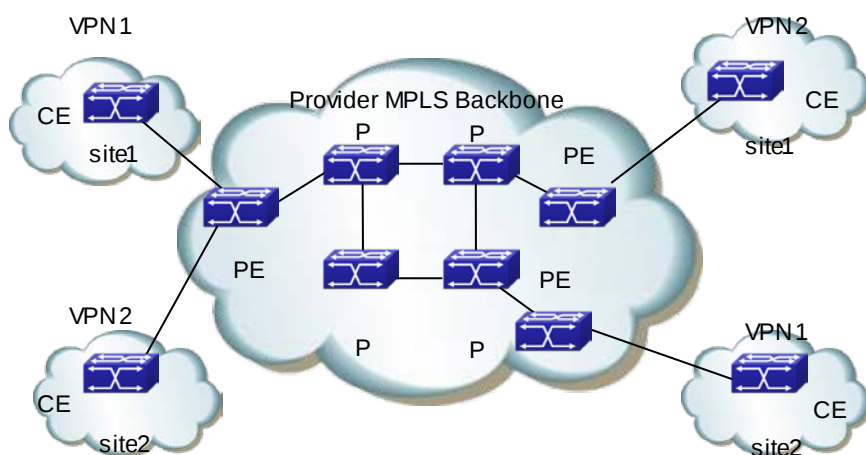


图 3-1 BGP/MPLS VPN 组网

CE 和 PE 的划分主要是根据 SP 与用户的管理范围, CE 和 PE 是两者管理范围的边界。

CE 设备通常是一台路由器, 当 CE 与直接相连的 PE 建立邻接关系后, CE 把本站点的 VPN 路由发布给 PE, 并从 PE 学到远端 VPN 的路由。CE 与 PE 之间使用 BGP/IGP 交换路由信息, 也可以使用静态路由。

PE 从 CE 学到 CE 本地的 VPN 路由信息后, 通过 BGP 与其它 PE 交换 VPN 路由信息。PE 路由器只维护与它直接相连的 VPN 的路由信息, 不维护服务提供商网络中的所有 VPN 路由。

P 路由器只维护到 PE 的路由, 不需要了解任何 VPN 路由信息。

当在 MPLS 骨干网上传输 VPN 流量时, 入口 PE 做为 Ingress LSR (Label Switch Router), 出口 PE 做为 Egress LSR, P 路由器则做为 Transit LSR。

3.1.2 BGP/MPLS VPN基本概念

Site

在介绍 VPN 时经常会提到 “site”, site (站点) 的含义可以从下述几个方面理解:

- ☞ site 是指相互之间具备 IP 连通性的一组 IP 系统, 并且, 这组 IP 系统的 IP 连通性不需通过服务提供商网络实现;
- ☞ site 的划分是根据设备的拓扑关系, 而不是地理位置, 尽管在大多数情况下一个 site 中的设备地理位置相邻;
- ☞ 一个 site 中的设备可以属于多个 VPN, 换言之, 一个 site 可以属于多个 VPN;
- ☞ site 通过 CE 连接到服务提供商网络, 一个 site 可以包含多个 CE, 但一个 CE 只属于一个 site。

对于多个连接到同一服务提供商网络的 sites, 通过制定策略, 可以将它们划分为不同的集合 (set), 只有属于相同集合的 sites 之间才能通过服务提供商网络互访, 这种集合就是 VPN。

VRF

VRF 即 VPN Routing & Forwarding Instance, VPN 路由转发实例, 由 VPN IP 路由表和 VPN IP 转发表(转发表包含了 MPLS 封装信息)组成, 是实现 MPLS VPN 分组转发的核心表项。一个 VPN 有自己独立的一个 VRF, 不同 VPN 的 VRF 地址空间可以重叠。

在 MPLS VPN 网络中一个 PE/P 路由器通常包含有多个独立运作的 VRF。**地址空间重叠**

VPN 是一种私有网络, 不同的 VPN 独立管理自己使用的地址范围, 也称为地址空间 (Address Space)。

不同 VPN 的地址空间可能会在一定范围内重合, 比如, VPN1 和 VPN2 都使用了 10.110.10.0/24 网段的地址, 这就发生了地址空间重叠 (Overlapping Address Spaces)。

VPN 实例

在 MPLS VPN 中, 不同 VPN 之间的路由隔离通过 VPN 实例 (VPN-instance) 实现。

PE 为每个直接相连的 Site 建立并维护专门的 VPN 实例。VPN 实例中包含对应 Site 的 VPN 成员关系和路由规则。如果一个 Site 中的用户同时属于多个 VPN, 则该 Site 的 VPN 实例中将包括所有这些 VPN 的信息。

为保证 VPN 数据的独立性和安全性, PE 上每个 VPN 实例都有相对独立的路由表和 LFIB (Label Forwarding Information Base, 标签转发表)。

具体来说, VPN 实例中的信息包括: 标签转发表、IP 路由表、与 VPN 实例绑定的接口、以及 VPN 实例的管理信息 (包括 RD、路由过滤策略、成员接口列表等)。

VPN-IPv4 地址

传统 BGP 无法正确处理地址空间重叠的 VPN 的路由。假设 VPN1 和 VPN2 都使用了 10.110.10.0/24 网段的地址，并各自发布了一条去往此网段的路由，BGP 将只会选择其中一条路由，从而导致去往另一个 VPN 的路由丢失。

PE 路由器之间使用 MP-BGP 来发布 VPN 路由，并使用 VPN-IPv4 地址族来解决上述问题。

VPN-IPv4 地址共有 12 个字节，包括 8 字节的 RD（Route Distinguisher）和 4 字节的 IPv4 地址前缀，如图 3-2 所示：

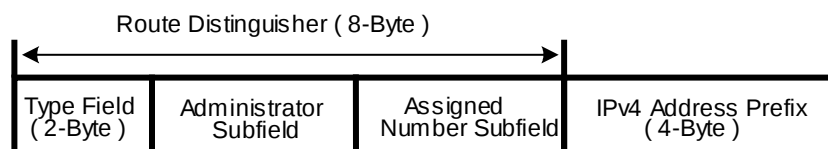


图 3-2 VPN-IPv4 地址结构

PE 从 CE 接收到普通 IPv4 路由后，需要将这些私网 VPN 路由发布给对端 PE。私网路由的独立性是通过为这些路由附加 RD 实现的。

SP 可以独立地分配 RD，但必须保证 RD 的全局唯一性。这样，即使来自不同服务提供商的 VPN 使用了同样的 IPv4 地址空间，PE 路由器也可以向各 VPN 发布不同的路由。

建议为 PE 上每个 VPN 实例配置专门的 RD，以保证到达同一 CE 的路由都使用相同的 RD。RD 为 0 的 VPN-IPv4 地址相当于全局唯一的 IPv4 地址。

RD 的作用是添加到一个特定的 IPv4 前缀，使之成为全局唯一的 VPN IPv4 前缀。

RD 或者是与自治系统号（ASN）相关的，在这种情况下，RD 是由一个自治系统号和一个任意的数组成；或者是与 IP 地址相关的，在这种情况下，RD 是由一个 IP 地址和一个任意的数组成。

RD 有两种格式，通过 2 字节的 Type 字段区分：

- ☞ Type 为 0 时，Administrator 子字段占 2 字节，Assigned Number 子字段占 4 字节，格式为：16bits 自治系统号:32bits 用户自定义数字。例如：100:1
- ☞ Type 为 1 时，Administrator 子字段占 4 字节，Assigned Number 子字段占 2 字节，格式为：32bits IPv4 地址:16bits 用户自定义数字。例如：172.1.1.1:1

为保证 RD 的全局唯一性，建议不要将 Administrator 子字段的值设置为私有 AS 号或私有 IP 地址。

VPN Target 属性

BGP/MPLS VPN 使用 32 位的 BGP 扩展团体属性——VPN Target（也称为 Route Target）来控制 VPN 路由信息的发布。

PE 路由器上的 VPN 实例有两类 VPN Target 属性：

- ☞ Export Target 属性：本地 PE 在把从与自己直接相连的 Site 学到的 VPN-IPv4 路由发布给其它 PE 前，为这些路由设置 Export Target 属性；
- ☞ Import Target 属性：PE 在接收到其它 PE 路由器发布的 VPN-IPv4 路由时，检查其 Export Target 属性，只有当此属性与 PE 上 VPN 实例的 Import Target 属性匹配时，才把路由加入到相应的 VPN 路由表中。

也就是说，VPN Target 属性定义了，一条 VPN-IPv4 路由可以为哪些 Site 所接收，PE 路由器可以接收哪些 Site 发送来的路由。

与 RD 类似，VPN Target 也有两种格式：

- ☞ 16bits 自治系统号:32bits 用户自定义数字，例如：100:1。
- ☞ 32bits IPv4 地址:16bits 用户自定义数字，例如：172.1.1.1:1。

MP-BGP

MP-BGP（Multiprotocol extensions for BGP-4）在 PE 路由器之间传播 VPN 组成信息和路由。MP-BGP 向下兼容，既可以支持传统的 IPv4 地址族，又可以支持其他地址族（比如 VPN-IPv4 地址族）。使用 MP-BGP 确保 VPN 的私网路由只在 VPN 内发布，并实现 MPLS VPN 成员间的通信。

路由策略（Routing Policy）

在通过入口、出口扩展团体来控制 VPN 路由发布的基础上，如果需要更精确地控制 VPN 路由的引入和发布，可以使用入方向或出方向路由策略。

入方向路由策略根据路由的 VPN Target 属性进一步过滤可引入到 VPN 实例的路由，它可以拒绝接收引入列表中的团体选定的路由，而出方向路由策略则可以拒绝发布输出列表中的团体选定的路由。

VPN 实例创建完成后，可以选择是否需要配置入方向或出方向路由策略。

隧道策略（Tunneling Policy）

隧道策略用于选择给特定 VPN 实例的报文使用的隧道。

隧道策略是可选配的，VPN 实例创建完成后，就可以配置隧道策略。缺省情况下，选择 LSP 做为隧道，不进行负载分担（负载分担条数为 1）。另外，隧道策略只在同一 AS 域内生效。

3.1.3 BGP/MPLS VPN 报文转发

在基本 L3VPN 应用中（不包括跨域的情况），VPN 报文转发采用两层标签方式：

- ☞ 第一层（外层）标签在骨干网内部进行交换，指示从 PE 到对端 PE 的一条 LSP。VPN 报文利用这层标签，可以沿 LSP 到达对端 PE；
- ☞ 第二层（内层）标签在从对端 PE 到达 CE 时使用，指示报文应被送到哪个 Site，或者更具体一些，到达哪一个 CE。这样，对端 PE 根据内层标签可以找到转发报文的接口。

特殊情况下，属于同一个 VPN 的两个 Site 连接到同一个 PE，这种情况下只需要知道如何到达对端 CE。

以图 3-3 为例，说明 VPN 报文的转发：

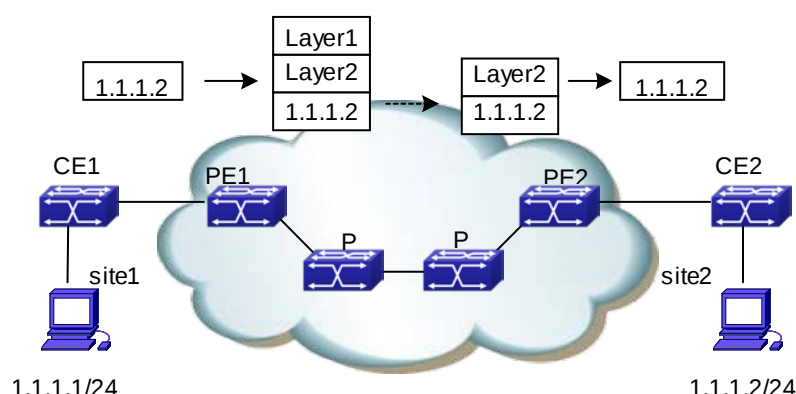


图 3-3 VPN 报文转发示意图

- (1) Site1 发出一个目的地址为 1.1.1.2 的 IP 报文，由 CE1 将报文发送至 PE1。
- (2) PE1 根据报文到达的接口及目的地址查找 VPN-instance 表项，匹配后将报文转发出去，同时打上内层和外层两个标签。
- (3) MPLS 网络利用报文的外层标签，将报文传送到 PE2（报文在到达 PE2 前一跳时已经被剥离外层标签，仅含内层标签）。
- (4) PE2 根据内层标签和目的地址查找 VPN-instance 表项，确定报文的出接口，将报文转发至 CE2。
- (5) CE2 根据正常的 IP 转发过程将报文传送到目的地。

3.1.4 BGP/MPLS VPN组网方案

在 BGP/MPLS VPN 网络中，通过 VPN Target 属性来控制 VPN 路由信息在各 site 之间的发布和接收。VPN Export Target 和 Import Target 的设置相互独立，并且都可以设置多个值，能够实现灵活的 VPN 访问控制，从而实现多种 VPN 组网方案。

基本的 VPN

最简单的情况下，一个 VPN 中的所有用户形成闭合用户群，相互之间能够进行流量转发，VPN 中的用户不能与任何本 VPN 以外的用户通信。

对于这种组网，需要为每个 VPN 分配一个 VPN Target，作为该 VPN 的 Export Target 和 Import Target，并且，此 VPN Target 不能被其他 VPN 使用。

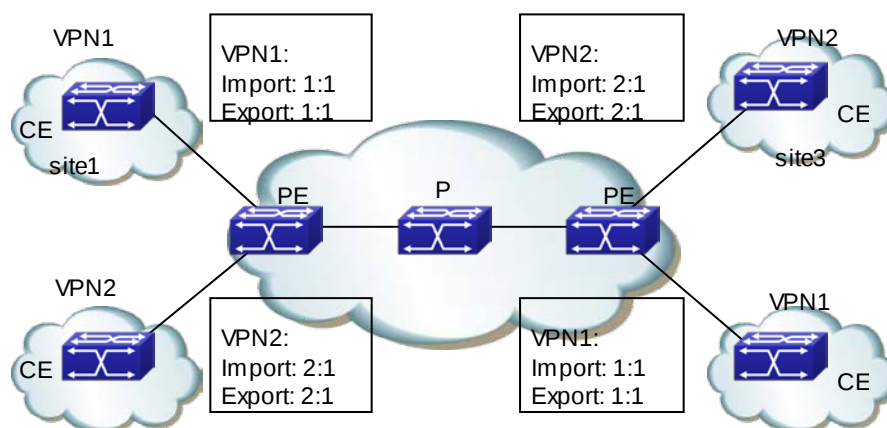


图 3-4 基本的 VPN 组网方案

在图 3-4 中, PE 上为 VPN1 分配的 VPN Target 值为 100:1, 为 VPN2 分配的 VPN Target 值为 200:1。VPN1 的两个 site 之间可以互访, VPN2 的两个 site 之间也可以互访, 但 VPN1 和 VPN2 的 site 之间不能互访。

Hub&Spoke VPN

如果希望在 VPN 中设置中心访问控制设备, 其它用户的互访都通过中心访问控制设备进行, 可以使用 Hub&Spoke 组网方案, 从而实现中心设备对两端设备之间的互访进行监控和过滤等功能。

对于这种组网, 需要设置两个 VPN Target, 一个表示“Hub”, 另一个表示“Spoke”。

各 site 在 PE 上的 VPN 实例的 VPN Target 设置规则为:

- ☞ 连接 Spoke 站点的 (Spoke-PE): Export Target 为“Spoke”, Import Target 为“Hub”;
- ☞ 连接 Hub 站点的 (Hub-PE): Hub-PE 上需要使用两个接口或子接口, 一个用于接收 Spoke-PE 发来的路由, 其 VPN 实例的 Import Target 为“Spoke”; 另一个用于向 Spoke-PE 发布路由, 其 VPN 实例的 Export Target 为“Hub”。

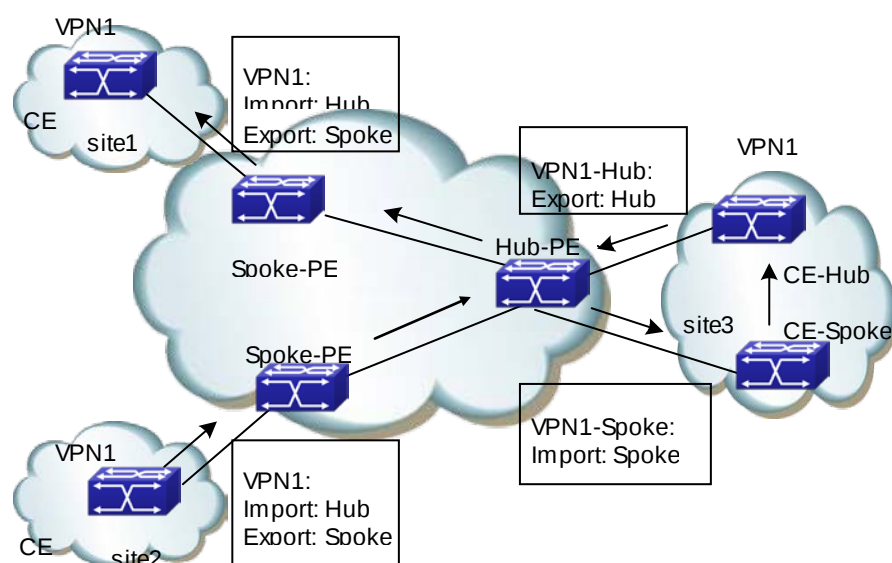


图 3-5 Hub&Spoke 组网方案

在图 3-5 中，Spoke 站点之间的通信通过 Hub 站点进行（图中箭头所示为 site2 的路由向 site1 的发布过程）：

- ☞ Hub-PE 能够接收所有 Spoke-PE 发布的 VPN-IPv4 路由；
- ☞ Hub-PE 发布的 VPN-IPv4 路由能够为所有 Spoke-PE 接收；
- ☞ Hub-PE 将从 Spoke-PE 学到的路由发布给其他 Spoke-PE，因此，Spoke 站点之间可以通过 Hub 站点互访。
- ☞ 任意 Spoke-PE 的 Import Target 属性不与其它 Spoke-PE 的 Export Target 属性相同。因此，任意两个 Spoke-PE 之间不直接发布 VPN-IPv4 路由，Spoke 站点之间不能直接互访。

Extranet VPN

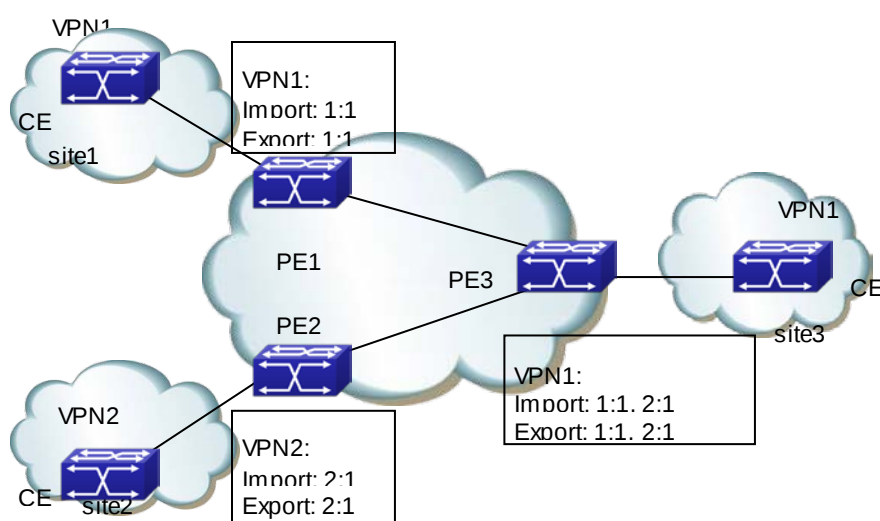


图 3-6 Extranet 组网方案

如果一个 VPN 用户希望提供部分本 VPN 的站点资源给非本 VPN 的用户访问, 可以使用 Extranet 组网方案。

对于这种组网, 如果某个 VPN 需要访问共享站点, 则该 VPN 的 Export Target 必须包含在共享站点的 VPN 实例的 Import Target 中, 而其 Import Target 必须包含在共享站点 VPN 实例的 Export Target 中。

在图 3-6 中, VPN1 的 site3 能够被 VPN1 和 VPN2 访问:

- ☞ PE3 能够接受 PE1 和 PE2 发布的 VPN-IPv4 路由;
- ☞ PE3 发布的 VPN-IPv4 路由能够为 PE1 和 PE2 接受;
- ☞ 基于以上两点, VPN1 的 site1 和 site3 之间能够互访, VPN2 的 site2 和 VPN1 的 site3 之间能够互访;

PE3 不把从 PE1 接收的 VPN-IPv4 路由发布给 PE2, 也不把从 PE2 接收的 VPN-IPv4 路由发布给 PE1 (IBGP 邻居学来的条目是不会再发送给别的 IBGP 邻居), 因此, VPN1 的 site1 和 VPN2 的 site2 之间不能互访。

3.1.5 BGP/MPLS VPN路由发布

在基本 BGP/MPLS VPN 组网中, VPN 路由信息的发布涉及 CE 和 PE, P 路由器只维护骨干网的路由, 不需要了解任何 VPN 路由信息。PE 路由器也只维护与它直接相连的 VPN 的路由信息, 不维护所有 VPN 路由。因此, BGP/MPLS VPN 网络具有良好的可扩展性。

VPN 路由信息的发布过程包括三部分: 本地 CE 到入口 PE、入口 PE 到出口 PE、出口 PE 到远端 CE。完成这三部分后, 本地 CE 与远端 CE 之间将建立可达路由, VPN 私网路由信息能够在骨干网上发布。

下面分别对这三部分进行介绍。

本地 CE 到入口 PE 的路由信息交换

CE 与直接相连的 PE 建立邻接关系后, 把本站点的 VPN 路由发布给 PE。

CE 与 PE 之间可以使用静态路由、RIP、OSPF、IS-IS 或 EBGp。无论使用哪种路由协议, CE 发布给 PE 的都是标准的 IPv4 路由。

入口 PE 到出口 PE 的路由信息交换

PE 从 CE 学到 VPN 路由信息后, 为这些标准 IPv4 路由增加 RD 和 VPN Target 属性, 形成 VPN-IPv4 路由, 存放到为 CE 创建的 VPN 实例中。

入口 PE 通过 MP-BGP 把 VPN-IPv4 路由发布给出口 PE。出口 PE 根据 VPN-IPv4 路由的 Export Target 属性与自己维护的 VPN 实例的 Import Target, 决定是否将该路由加入到 VPN 实例的路由表。

PE 之间通过 IGP 来保证内部的连通性。

出口 PE 到远端 CE 的路由信息交换

远端 CE 有多种方式可以从出口 PE 学习 VPN 路由, 包括静态路由、RIP、OSPF、IS-IS 和 EBGp, 与本地 CE 到入口 PE 的路由信息交换相同。

3.1.6 跨域VPN简介

实际组网应用中, 某用户一个 VPN 的多个 Site 可能会连接到使用不同 AS 号的多个服务提供商, 或者连接到一个服务提供商的多个 AS。这种 VPN 跨越多个自治系统的应用方式被称为跨域 VPN (Multi-AS VPN)。RFC 2547 中提出了三种跨域 VPN 解决方案, 分别是:

- ☞ VRF-to-VRF: ASBR 间使用 VRF 接口建立 EBGP 邻居并管理 VPN 路由, 也称为 Inter-Provider Option A;
- ☞ EBGP Redistribution of labeled VPN-IPv4 routes: ASBR 间通过 MP-EBGP 发布标签 VPN-IPv4 路由, 也称为 Inter-Provider Option B;
- ☞ Multihop EBGP redistribution of labeled VPN-IPv4 routes: PE 间通过 Multi-hop MP-EBGP 发布标签 VPN-IPv4 路由, 也称为 Inter-Provider Option C。

目前我们能够实现第一种 VRF-to-VRF 的跨域 VPN 解决方案。

跨域 VPN 解决方案

如图 3-7 所示, 在这种方式下, 两个 AS 的 PE 路由器直接相连, PE 路由器同时也是各自所在自治系统的边界路由器 ASBR。作为 ASBR 的 PE 之间通过 VRF 接口相连, 并使其 IMPORT 所有本侧想要对方所知的 RT, EXPORT 所有本侧想从对方获得的 RT, 并在双方的 VRF 上建立 EBGP 的连接, 从而使各自服务的 CE 象在同一 AS 内一样达到相互通信和隔离的目的。此时两个 PE 都把对方作为自己的 CE 设备对待, 报文在 AS 内部作为 VPN 报文, 采用两层标签转发方式; 在 ASBR 之间则采用普通 IP 转发方式。

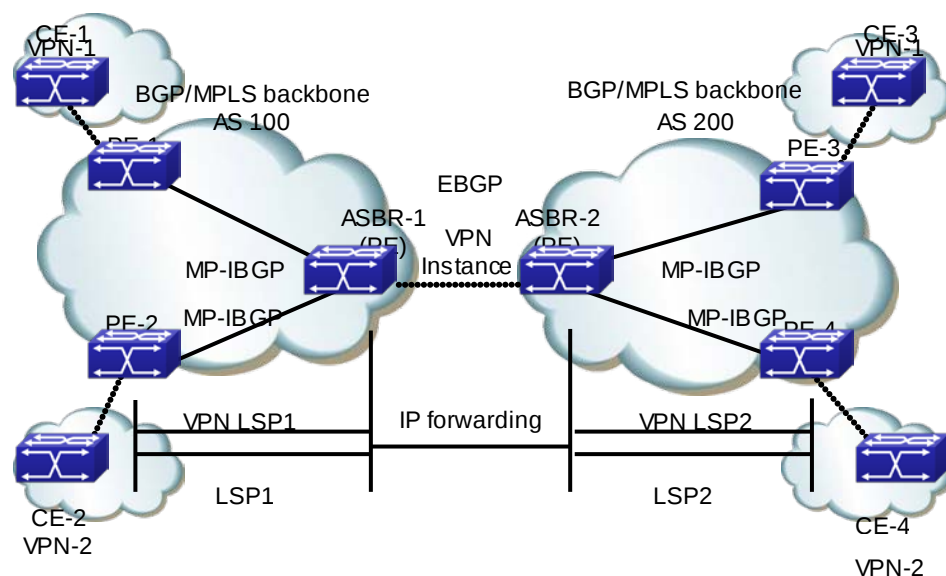


图 3-7 跨域 VPN 组网图

- ☞ 使用这种方式实现跨域 VPN 的优点是实现简单: 两个作为 ASBR 的 PE 之间不需要为跨域进行特殊配置。
- ☞ 缺点是可扩展性差: 作为 ASBR 的 PE 需要管理所有 VPN 路由, 为每个 VPN 创建 VPN 实例。这将导致 PE 上的 VPN-IPv4 路由数量过于庞大。

3.2 BGP MPLS VPN配置

BGP MPLS VPN 配置任务序列:

1. 全局启动 MPLS (必选)
2. 配置 VPN 实例 (必选)
 - (1) 创建 VPN 实例, 并进入 VPN 实例视图

- (2) 配置 VPN 实例的 RD
- (3) 配置 VPN 实例的 RT
- (4) 配置 VPN 实例与接口关联
3. 配置基本 MPLS VPN (必选)
 - (1) 配置 PE-CE 间使用 EBGp
 - 1) 配置对端 PE 为公网 VPNv4 邻居
 - 2) 进入 BGP-VPN 实例视图
 - 3) 将 CE 配置为 VPN 私网邻居
 - 4) 发布本端的私网路由
 - (2) 配置 PE-CE 间使用 OSPF
 - 1) 配置对端 PE 为公网 VPNv4 邻居
 - 2) 创建 PE-CE 间的 OSPF 实例，并进入 Router OSPF 视图
 - 3) 使能 PE-CE 间网段的 OSPF
 - 4) 配置重发布 BGP 路由
 - 5) 进入 BGP-VPN 实例视图
 - 6) 配置重发布 OSPF 路由
 - 7) 发布本端的私网路由
 - (3) 配置 PE-CE 间使用 RIP
 - 1) 配置对端 PE 为公网 VPNv4 邻居
 - 2) 进入 RIP VPN 实例视图
 - 3) 使能 PE-CE 间网段的 RIP
 - 4) 配置重发布 BGP 路由
 - 5) 进入 BGP-VPN 实例视图
 - 6) 配置重发布 RIP 路由
 - 7) 发布本端的私网路由
 - (4) 配置 PE-CE 间使用静态路由
 - 1) 配置对端 PE 为公网 VPNv4 邻居
 - 2) 配置静态 VPN 路由
 - 3) 进入 BGP-VPN 实例视图
 - 4) 配置重发布静态路由
 - 5) 发布本端的私网路由

1. 全局启动 MPLS (必选)

命令	解释
全局配置模式	
mpls enable no mpls enable	启动 mpls 协议；本命令的 no 操作关闭 MPLS 协议。

2. 配置 VPN 实例 (必选)

- (1) 创建 VPN 实例，并进入 VPN 实例视图
- (2) 配置 VPN 实例的 RD
- (3) 配置 VPN 实例的 RT
- (4) 配置 VPN 实例与接口关联

命令	解释
全局配置模式	
ip vrf <vrf-name> no ip vrf <vrf-name>	创建 VPN 实例; 缺省情况下未创建 VPN 实例。
VRF 配置模式	
rd <ASN:nn_or_IP-address:nn> no rd <ASN:nn_or_IP-address:nn>	配置 VPN 实例的 RD。缺省情况下未创建 RD。
route-target {import export both} <rt-value> no route-target {import export both} <rt-value>	配置 VPN 实例的 RT。
接口配置模式	
ip vrf forwarding <vrf-name> no ip vrf forwarding <vrf-name>	配置 VPN 实例与接口关联。
mpls proxy loopback-group <1-max_agg_num> no mpls proxy loopback-group	启动 MPLS 代理, 缺省不启动。注: 只在 VRF 入方向的板卡不支持 MPLS 功能时使能 MPLS 代理功能。
ip address <ip-address> <mask> no ip address <ip-address> <mask>	配置 PE-CE 间直连接口的私网 IP 地址。

3 配置基本 MPLS VPN (必选)

- (1) 配置 PE-CE 间使用 EBGP (必选)
 - 1) 配置对端 PE 为公网 VPNv4 邻居
 - 2) 进入 BGP-VPN 实例视图
 - 3) 将 CE 配置为 VPN 私网邻居
 - 4) 发布本端的私网路由

命令	解释
BGP 协议配置模式	
neighbor <ip-address> remote-as <as-num>	配置对端 PE 为公网 VPNv4 邻居。建议选取 loopback 接口建立公网 PE 间的 BGP 邻居。
neighbor <ip-address> update-source <as-num>	指定建立邻居的本地 loopback 接口。
进入 BGP-VPNv4 视图	
address-family vpnv4 [unicast]	创建 BGP 协议 VPNv4。缺省情况下未创建任何 VPNv4。
neighbor <ip-address> active no neighbor <ip-address> active	激活 VPNv4 下的邻居。缺省情况下 VPNv4 视图下所有邻居都是未激活的。
BGP 协议配置模式	

address-family ipv4 {unicast multicast vrf <vrf-name>} no address-family ipv4 {unicast multicast vrf <vrf-name>}	创建 BGP 协议 IPv4 并进入 BGP-VPN 实例视图。缺省情况下未创建任何 IPv4。
BGP-VPN 实例视图	
neighbor <ip-address> remote-as <as-num> no neighbor <ip-address> remote-as <as-num>	将 CE 配置为 VPN 私网邻居。缺省情况下未配置任何私网邻居。
neighbor <ip-address> active no neighbor <ip-address> active	激活 IPv4 下的邻居。缺省情况下所有邻居都是激活的。
redistribute {connected ospf rip static} no redistribute {connected ospf rip static}	配置重发布直连路由和其他协议的路由。缺省情况下未重发布任何路由。

(2) 配置 PE-CE 间使用 OSPF

- 1) 配置对端 PE 为公网 VPNv4 邻居
- 2) 创建 PE-CE 间的 OSPF 实例，并进入 OSPF VPN 实例视图
- 3) 使能 PE-CE 间网段的 OSPF
- 4) 配置重发布 BGP 路由
- 5) 进入 BGP-VPN 实例视图
- 6) 配置重发布 OSPF 路由
- 7) 发布本端的私网路由

命令	解释
BGP 协议配置模式	
neighbor <ip-address> remote-as <as-num>	配置对端 PE 为公网 VPNv4 邻居。建议选取 loopback 接口建立公网 PE 间的 BGP 邻居。
neighbor <ip-address> update-source <as-num>	指定建立邻居的本地 loopback 接口。
进入 BGP-VPNv4 视图	
address-family vpnv4 [unicast]	创建 BGP 协议 VPNv4。缺省情况下未创建任何 VPNv4。
neighbor <ip-address> active no neighbor <ip-address> active	可选 激活 VPNv4 下的邻居。缺省情况下所有邻居都是激活的。
全局配置模式	

router ospf [<process_id> [<vrf-nam>]] no router ospf [<process_id> [<vrf-nam>]]	可选 创建 PE-CE 间的 OSPF 实例，并进入 OSPF VPN 实例视图
OSPF VPN 实例视图	
network {<network> <mask> <network>/<prefix>} area <area_id> no network {<network> <mask> <network>/<prefix>} area <area_id>	可选 使能 PE-CE 间网段的 OSPF。缺省情况下未使能任何网段
redistribute {bgp connected static rip kernel} [metric-type {1 2}] [tag <tag>] [metric <cost_value>] [router-map <WORD>] no redistribute {bgp connected static rip kernel} [metric-type {1 2}] [tag <tag>] [metric <cost_value>] [router-map <WORD>]	可选 配置重发布 BGP 路由。缺省情况下未重发布任何路由
BGP 协议配置模式	
address-family ipv4 {unicast multicast vrf <vrf-nam>} no address-family ipv4 {unicast multicast vrf <vrf-nam>}	可选 创建 BGP 协议 IPv4 并进入 BGP-VPN 实例视图。缺省情况下未创建任何 IPv4
BGP-VPN 实例视图	
redistribute {connected ospf rip static} no redistribute {connected ospf rip static}	可选 配置重发布直连路由和其他协议路由。缺省情况下未重发布任何路由。

(3) 配置 PE-CE 间使用 RIP

- 1) 配置对端 PE 为公网 VPNv4 邻居
- 2) 进入 RIP VPN 实例视图
- 3) 使能 PE-CE 间网段的 RIP
- 4) 配置重发布 BGP 路由
- 5) 进入 BGP-VPN 实例视图
- 6) 配置重发布 RIP 路由
- 7) 发布本端的私网路由

命令	解释
BGP 协议配置模式	

neighbor <ip-address> remote-as <as-num>	配置对端 PE 为公网 VPNv4 邻居。建议选取 loopback 接口建立公网 PE 间的 BGP 邻居。
neighbor <ip-address> update-source <as-num>	指定建立邻居的本地 loopback 接口
进入 BGP-VPNv4 视图	
address-family vpnv4 [unicast]	创建 BGP 协议 VPNv4。缺省情况下未创建任何 VPNv4。
neighbor <ip-address> active no neighbor <ip-address> active	可选 激活 VPNv4 下的邻居。缺省情况下所有邻居都是激活的。
RIP 协议配置模式	
address-family ipv4 vrf <vrf-name> no address-family ipv4 vrf <vrf-name>	可选 创建 PE-CE 间的 RIP IPv4 协议族, 并进入 RIP VPN 实例视图
RIP VPN 实例视图	
network {A.B.C.D/M ifname vlan<id> loopback <1-1024> } no network {A.B.C.D/M ifname vlan<id> loopback <1-1024> }	可选 使能 PE-CE 间网段的 RIP
redistribute {kernel connected static ospf isis bgp} [metric <value>] [route-map <word>] no redistribute {kernel connected static ospf isis bgp} [metric <value>] [route-map <word>]	可选 配置重发布 BGP 路由。缺省情况下未重发布任何路由
BGP 协议配置模式	
address-family ipv4 {unicast multicast vrf <vrf-nam>} no address-family ipv4 {unicast multicast vrf <vrf-nam>}	可选 创建 BGP 协议 IPv4 并进入 BGP-VPN 实例视图。缺省情况下未创建任何 IPv4
BGP-VPN 实例视图	
redistribute {connected ospf rip static} no redistribute {connected ospf rip static}	可选 配置重发布直连路由和其他协议路由。缺省情况下未重发布任何路由。

(4) 配置 PE-CE 间使用静态路由

1) 配置对端 PE 为公网 VPNv4 邻居

- 2) 配置静态 VPN 路由
- 3) 进入 BGP-VPN 实例视图
- 4) 配置重发布静态路由
- 5) 发布本端的私网路由

命令	解释
BGP 协议配置模式	
neighbor <ip-address> remote-as <as-num>	配置对端 PE 为公网 VPNv4 邻居。建议选取 loopback 接口建立公网 PE 间的 BGP 邻居。
neighbor <ip-address> update-source <as-num>	指定建立邻居的本地 loopback 接口。
进入 BGP-VPNv4 视图	
address-family vpnv4 [unicast]	创建 BGP 协议 VPNv4。缺省情况下未创建任何 VPNv4。
neighbor <ip-address> active no neighbor <ip-address> active	可选 激活 VPNv4 下的邻居。缺省情况下所有邻居都是激活的。
全局配置模式	
ip route vrf <vrf-name> {<ip-prefix> <mask> <ip-prefix/<prefix-length>} <gateway-address> null0} no ip route vrf <vrf-name> <ip-prefix> <mask> <ip-prefix/<prefix-length>} <gateway-address> null0}	可选 手动配置 PE-CE 间的静态 VPN 路由。缺省情况下没有配置任何静态路由
BGP 协议配置模式	
address-family ipv4 {unicast multicast vrf <vrf-name>} no address-family ipv4 {unicast multicast vrf <vrf-name>}	可选 创建 BGP 协议 IPv4 并进入 BGP-VPN 实例视图。缺省情况下未创建任何 IPv4
BGP-VPN 实例视图	
redistribute {connected ospf rip static} no redistribute {connected ospf rip static}	可选 配置重发布静态路由、直连路由和其他协议路由。缺省情况下未重发布任何路由。

3.3 BGP MPLS VPN典型案例

3.3.1 PE-CE间使用EBGP建立BGP MPLS VPN

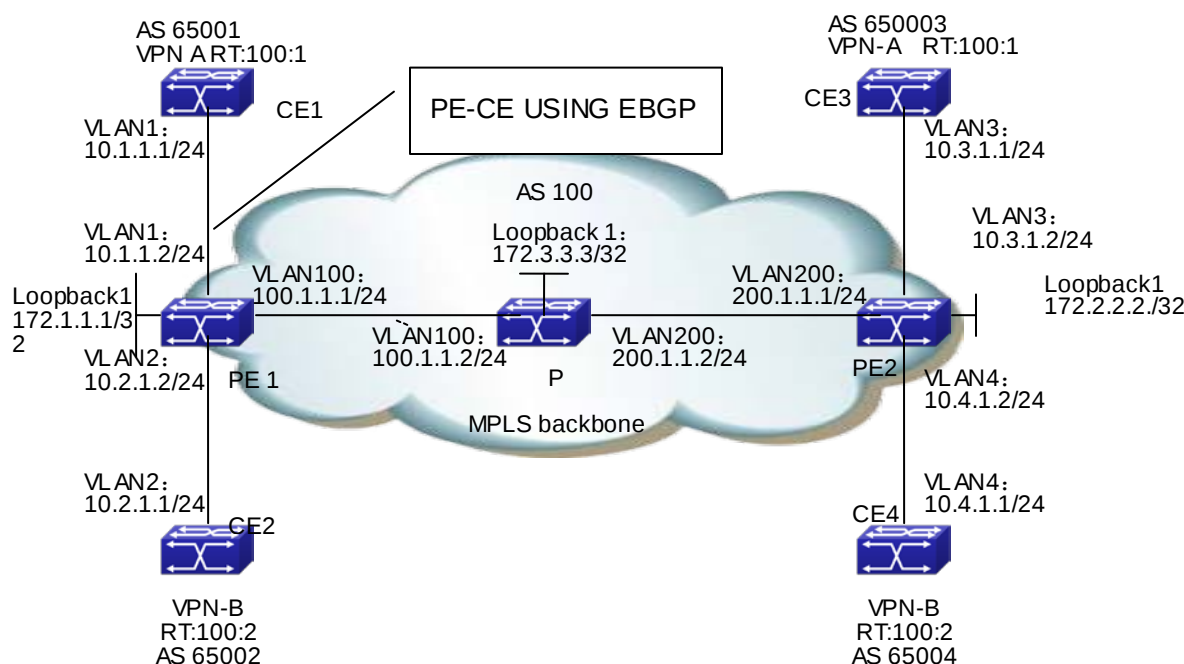


图 3-8 PE-CE 间使用 EBGP 建立 BGP MPLS VPN

CE1 的配置如下: (CE2~CE4 另外三个设备配置与 CE1 类似, 所以省略)

```
CE1#config
CE1(config)# interface vlan 1
CE1(config-if-Vlan1)#ip address 10.1.1.1 255.255.255.0
CE1(config-if-Vlan1)#exit
CE1(config)# router bgp 65001
CE1(config-router)#neighbor 10.1.1.2 remote-as 100
CE1(config-router)#redistribute connect
CE1(config-router)#exit
```

PE1 交换机的 MPLS BGP 配置如下:

(1) 配置 VPN 实例

```
PE1#config
PE1(config)#ip vrf vpna
PE1(config-vrf)#rd 100:1
PE1(config-vrf)#route-target both 100:1
PE1(config)#ip vrf vpnb
PE1(config-vrf)#rd 100:2
PE1(config-vrf)#route-target both 100:2
```

(2) 配置接口绑定 VPN 实例

```
PE1(config)# interface vlan 1
PE1(config-if-Vlan1)# ip vrf forwarding vpna
```

```
PE1(config-if-Vlan1)#ip address 10.1.1.2 255.255.255.0
```

```
PE1(config-if-Vlan1)#exit
```

```
PE1(config)# interface vlan 2
```

```
PE1(config-if-Vlan2)# ip vrf forwarding vpnb
```

```
PE1(config-if-Vlan2)#ip address 10.2.1.2 255.255.255.0
```

```
PE1(config-if-Vlan2)#exit
```

(3) 全局启动 MPLS 和 LDP

```
PE1(config)#mpls enable
```

```
PE1(config)#router ldp
```

```
PE1(config-router)#exit
```

(4) 配置接口以及使能 LDP

```
PE1(config)# interface loopback 1
```

```
PE1(config-if-Loopback1)# ip address 172.1.1.1 255.255.255.255
```

```
PE1(config-if-Loopback1)# exit
```

```
PE1(config)# interface vlan 100
```

```
PE1(config-if-Vlan100)#ip address 100.1.1.1 255.255.255.0
```

```
PE1(config-if-Vlan100)#label-switching
```

```
PE1(config-if-Vlan100) #ldp enable
```

```
PE1(config-if-Vlan100)#exit
```

(5) 启动 OSPF 协议传送内网路由

```
PE1(config)#router ospf
```

```
PE1(config-router)# ospf router-id 172.1.1.1
```

```
PE1(config-router)# network 0.0.0.0/0 area 0
```

```
PE1(config-router)# redistribute connected
```

(6) 配置 BGP

```
PE1(config)# router bgp 100
```

```
PE1(config-router)#neighbor 172.2.2.2 remote-as 100
```

```
PE1(config-router)#neighbor 172.2.2.2 update-source 172.1.1.1
```

```
PE1(config-router)#address-family vpnv4
```

```
PE1(config-router-af)#neighbor 172.2.2.2 activate
```

```
PE1(config-router-af)#exit
```

```
PE1(config-router)# address-family ipv4 vrf vpna
```

```
PE1(config-router-af)#neighbor 10.1.1.1 remote-as 65001
```

```
PE1(config-router-af)#redistribute connected
```

```
PE1(config-router-af)#exit
```

```
PE1(config-router)# address-family ipv4 vrf vpnb
```

```
PE1(config-router-af)#neighbor 10.2.1.1 remote-as 65002
```

```
PE1(config-router-af)#redistribute connected
```

```
PE1(config-router-af)#exit
```

```
PE1(config-router)#exit
```

P 交换机的配置如下：

(1) 全局使能 MPLS 并且配置相关接口的 LDP

```
P#config
```

```
P(config)#mpls enable
```

```
P(config)#router ldp
P(config-router)#exit
P(config)# interface loopback 1
P(config-if-Loopback1)# ip address 172.3.3.3 255.255.255.255
P(config-if-Loopback1)# exit
P(config)#interface vlan 100
P(config-if-Vlan100)#ip address 100.1.1.2 255.255.255.0
P(config-if-Vlan100)#label-switching
P(config-if-Vlan100)#ldp enable
P(config-if-Vlan100)#exit
P(config)#interface vlan200
P(config-if-Vlan200)#ip address 200.1.1.2 255.255.255.0
P(config-if-Vlan200)#label-switching
P(config-if-Vlan200)#ldp enable
P(config-if-Vlan200)#exit
(2) 配置 OSPF 协议
P(config)#router ospf
P(config-router)# ospf router-id 172.3.3.3
P(config-router)# network 0.0.0.0/0 area 0
P(config-router)# redistribute connected
```

PE2 交换机的配置如下：

```
(1) 配置 VPN 实例
PE2#config
PE2(config)#ip vrf vpna
PE2(config-vrf)#rd 100:1
PE2(config-vrf)#route-target both 100:1
PE2(config)#ip vrf vpnb
PE2(config-vrf)#rd 100:2
PE2(config-vrf)#route-target both 100:2
(2) 配置接口绑定 VPN 实例
PE2(config)# interface vlan 3
PE2(config-if-Vlan3)# ip vrf forwarding vpna
PE2(config-if-Vlan3)#ip address 10.3.1.2 255.255.255.0
PE2(config-if-Vlan3)#exit
PE2(config)# interface vlan 4
PE2(config-if-Vlan4)# ip vrf forwarding vpnb
PE2(config-if-Vlan4)#ip address 10.4.1.2 255.255.255.0
PE2(config-if-Vlan4)#exit
(3) 全局启动 MPLS 和 LDP
PE2(config)#mpls enable
PE1(config)#router ldp
PE1(config-router)#exit
(4) 配置接口以及使能 LDP
PE2(config)# interface loopback 1
```

```
PE2(config-if-Loopback1)# ip address 172.2.2.2 255.255.255.255
PE2(config-if-Loopback1)# exit
PE2(config)# interface vlan 200
PE2(config-if-Vlan200)#ip address 200.1.1.1 255.255.255.0
PE2(config-if-Vlan200)#label-switching
PE2(config-if-Vlan200) #ldp enable
PE2(config-if-Vlan200)#exit
(5) 启动 OSPF 协议传送内网路由
PE2(config)#router ospf
PE2(config-router)# ospf router-id 172.2.2.2
PE2(config-router)# network 0.0.0.0/0 area 0
PE2(config-router)# redistribute connected
(6) 配置 BGP
PE2(config)# router bgp 100
PE2(config-router)#neighbor 172.1.1.1 remote-as 100
PE2(config-router)#neighbor 172.1.1.1 update-source 172.2.2.2
PE2(config-router)#address-family vpnv4
PE2(config-router-af)#neighbor 172.1.1.1 activate
PE2(config-router-af)#exit
PE2(config-router)# address-family ipv4 vrf vpna
PE2(config-router-af)#neighbor 10.3.1.1 remote-as 65003
PE2(config-router-af)#redistribute connected
PE2(config-router-af)#exit
PE2(config-router)# address-family ipv4 vrf vpnb
PE2(config-router-af)#neighbor 10.4.1.1 remote-as 65004
PE2(config-router-af)#redistribute connected
PE2(config-router-af)#exit
PE2(config-router)#exit
```

3.3.2 PE-CE间使用OSPF建立BGP MPLS VPN

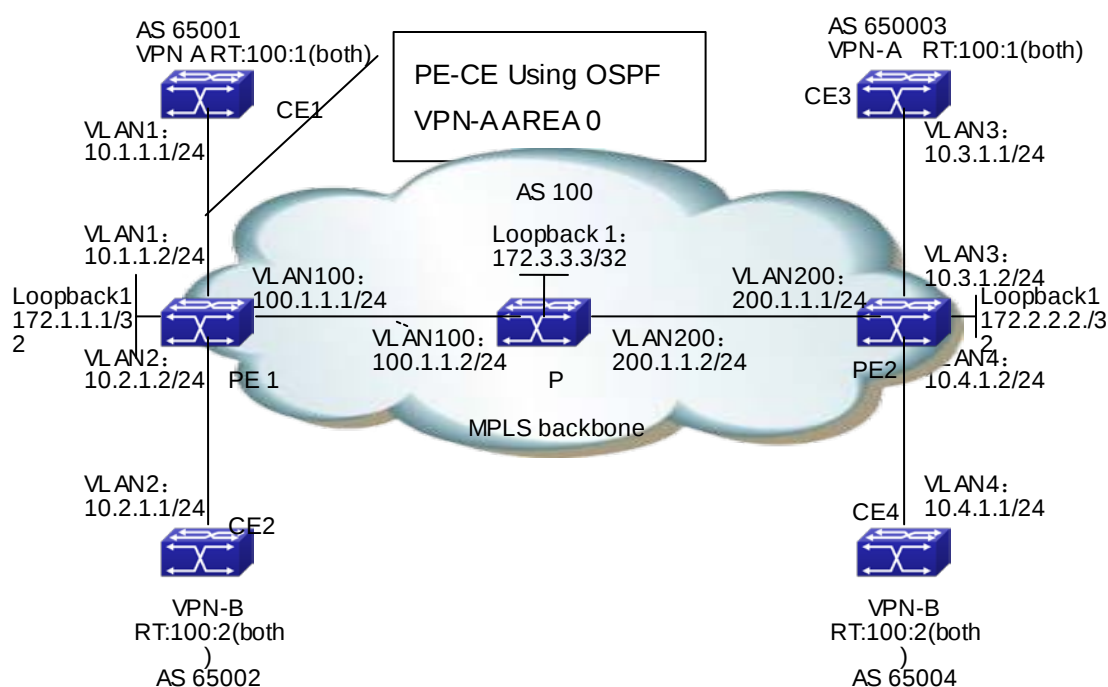


图 3-9 PE-CE 间使用 OSPF 建立 BGP MPLS VPN

CE1 的配置如下: (CE2~CE4 另外三个设备配置与 CE1 类似, 所以省略)

```
CE1#config
CE1(config)# interface vlan 1
CE1(config-if-Vlan1)#ip address 10.1.1.1 255.255.255.0
CE1(config-if-Vlan1)#exit
CE1(config)# router ospf
CE1(config-router)#network 0.0.0.0/0 area 0
CE1(config-router)#redistribute connect
CE1(config-router)#exit
```

PE1 交换机的 MPLS BGP 配置如下: (PE2 配置类似, 所以省略)

(1) 配置 VPN 实例

```
PE1#config
PE1(config)#ip vrf vpna
PE1(config-vrf)#rd 100:1
PE1(config-vrf)#route-target both 100:1
PE1(config)#ip vrf vpnb
PE1(config-vrf)#rd 100:2
PE1(config-vrf)#route-target both 100:2
```

(2) 配置接口绑定 VPN 实例

```
PE1(config)# interface vlan 1
PE1(config-if-Vlan1)# ip vrf forwarding vpna
PE1(config-if-Vlan1)#ip address 10.1.1.2 255.255.255.0
PE1(config-if-Vlan1)#exit
```

```
PE1(config)# interface vlan 2
PE1(config-if-Vlan2)# ip vrf forwarding vpnb
PE1(config-if-Vlan1)#ip address 10.2.1.2 255.255.255.0
PE1(config-if-Vlan1)#exit
(3) 全局启动 MPLS 和 LDP
PE1(config)#mpls enable
PE1(config)#router ldp
PE1(config-router)#exit
(4) 配置接口以及使能 LDP
PE1(config)# interface loopback 1
PE1(config-if-Loopback1)# ip address 172.1.1.1 255.255.255.255
PE1(config-if-Loopback1)# exit
PE1(config)# interface vlan 100
PE1(config-if-Vlan100)#ip address 100.1.1.1 255.255.255.0
PE1(config-if-Vlan100)# label-switching
PE1(config-if-Vlan100) #ldp enable
PE1(config-if-Vlan100)#exit
(5) 启动 OSPF 协议传送内网路由
PE1(config)#router ospf
PE1(config-router)# ospf router-id 172.1.1.1
PE1(config-router)# network 0.0.0.0/0 area 0
PE1(config-router)# redistribute connected
PE1(config-router)#exit
(6) 配置 OSPF VRF 传送私网路由
PE1(config)#router ospf 1 vpna
PE1(config-router)# network 0.0.0.0/0 area 0
PE1(config-router)#redistribute connected
PE1(config-router)#redistribute bgp
PE1(config-router)#exit
PE1(config)#router ospf 1 vpnb
PE1(config-router)# network 0.0.0.0/0 area 0
PE1(config-router)#redistribute connected
PE1(config-router)#redistribute bgp
PE1(config-router)#exit
(7) 配置 BGP
PE1(config)# router bgp 100
PE1(config-router)#neighbor 172.2.2.2 remote-as 100
PE1(config-router)#neighbor 172.2.2.2 update-source 172.1.1.1
PE1(config-router)#address-family vpnv4
PE1(config-router-af)#neighbor 172.2.2.2 activate
PE1(config-router-af)#exit
PE1(config-router)# address-family ipv4 vrf vpna
PE1(config-router-af)#redistribute connected
PE1(config-router-af)#redistribute ospf
```

```
PE1(config-router-af)#exit
PE1(config-router)# address-family ipv4 vrf vpnb
PE1(config-router-af)#redistribute connected
PE1(config-router-af)#redistribute ospf
PE1(config-router-af)#exit
PE1(config-router)#exit
```

P 交换机的配置如下:

(1) 全局使能 MPLS 并且配置相关接口的 LDP

```
P#config
P(config)#mpls enable
P(config)#router ldp
P(config-router)#exit
P(config)# interface loopback 1
P(config-if-Loopback1)# ip address 172.3.3.3 255.255.255.255
P(config-if-Loopback1)# exit
P(config)#interface vlan 100
P(config-if-Vlan100)#ip address 100.1.1.2 255.255.255.0
P(config-if-Vlan100)#label-switching
P(config-if-Vlan100)#ldp enable
P(config-if-Vlan100)#exit
P(config)#interface vlan200
P(config-if-Vlan200)#ip address 200.1.1.2 255.255.255.0
P(config-if-Vlan200)#label-switching
P(config-if-Vlan200)#ldp enable
P(config-if-Vlan200)#exit
(2) 配置 OSPF 协议
P(config)#router ospf
P(config-router)# ospf router-id 172.3.3.3
P(config-router)# network 0.0.0.0/0 area 0
P(config-router)# redistribute connected
```

3.3.3 PE-CE间使用RIP建立BGP MPLS VPN

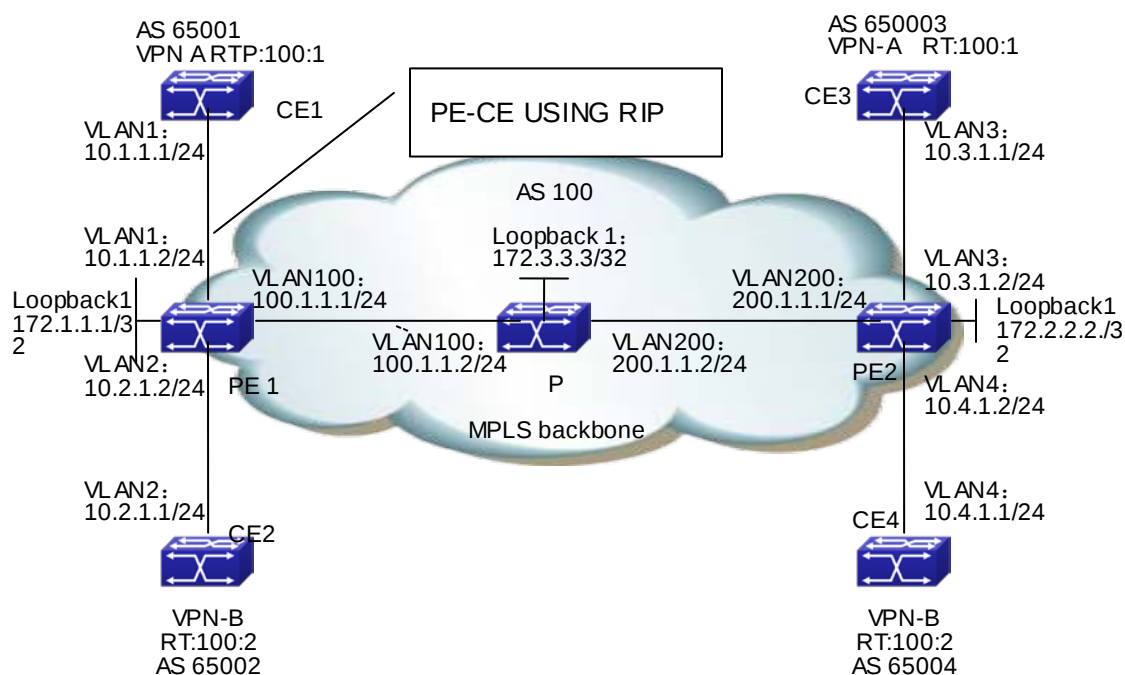


图 3-10 PE-CE 间使用 RIP 建立 BGP MPLS VPN

CE1 的配置如下: (CE2~CE4 另外三个设备配置与 CE1 类似, 所以省略)

```
CE1#config
CE1(config)# interface vlan 1
CE1(config-if-Vlan1)#ip address 10.1.1.1 255.255.255.0
CE1(config-if-Vlan1)#exit
CE1(config)# router rip
CE1(config-router)#network 0.0.0.0/0
CE1(config-router)#redistribute connect
CE1(config-router)#exit
```

PE1 交换机的 MPLS BGP 配置如下: (PE2 配置类似, 所以省略)

```
(1) 配置 VPN 实例
PE1#config
PE1(config)#ip vrf vpna
PE1(config-vrf)#rd 100:1
PE1(config-vrf)#route-target both 100:1
PE1(config)#ip vrf vpnb
PE1(config-vrf)#rd 100:2
PE1(config-vrf)#route-target both 100:2

(2) 配置接口绑定 VPN 实例
PE1(config)# interface vlan 1
PE1(config-if-Vlan1)# ip vrf forwarding vpna
PE1(config-if-Vlan1)#ip address 10.1.1.2 255.255.255.0
PE1(config-if-Vlan1)#exit
PE1(config)# interface vlan 2
```

```
PE1(config-if-Vlan2)# ip vrf forwarding vpnb
PE1(config-if-Vlan1)#ip address 10.2.1.2 255.255.255.0
PE1(config-if-Vlan1)#exit
(3) 全局启动 MPLS 和 LDP
PE1(config)#mpls enable
PE1(config)#router ldp
PE1(config-router)#exit
(4) 配置接口以及使能 LDP
PE1(config)# interface loopback 1
PE1(config-if-Loopback1)# ip address 172.1.1.1 255.255.255.255
PE1(config-if-Loopback1)# exit
PE1(config)# interface vlan 100
PE1(config-if-Vlan100)#ip address 100.1.1.1 255.255.255.0
PE1(config-if-Vlan100)#label-switching
PE1(config-if-Vlan100) #ldp enable
PE1(config-if-Vlan100)#exit
(5) 启动 OSPF 协议传送内网路由
PE1(config)#router ospf
PE1(config-router)# ospf router-id 172.1.1.1
PE1(config-router)# network 0.0.0.0/0 area 0
PE1(config-router)# redistribute connected
PE1(config-router)#exit
(6) 配置 RIP VRF 传送私网路由
PE1(config)#router rip
PE1(config-router)#address-family ipv4 vrf vpna
PE1(config-router-af)#network 0.0.0.0/0
PE1(config-router-af)#redistribute connected
PE1(config-router-af)#redistribute bgp
PE1(config-router-af)#exit
PE1(config-router)#address-family ipv4 vrf vpnb
PE1(config-router-af)#network 0.0.0.0/0
PE1(config-router-af)#redistribute connected
PE1(config-router-af)#redistribute bgp
PE1(config-router-af)#exit
PE1(config-router)#exit
(7) 配置 BGP
PE1(config)# router bgp 100
PE1(config-router)#neighbor 172.2.2.2 remote-as 100
PE1(config-router)#neighbor 172.2.2.2 update-source 172.1.1.1
PE1(config-router)#address-family vpnv4
PE1(config-router-af)#neighbor 172.2.2.2 activate
PE1(config-router-af)#exit
PE1(config-router)# address-family ipv4 vrf vpna
PE1(config-router-af)#redistribute connected
```

```
PE1(config-router-af)#redistribute ospf
PE1(config-router-af)#exit
PE1(config-router)# address-family ipv4 vrf vpnb
PE1(config-router-af)#redistribute connected
PE1(config-router-af)#redistribute ospf
PE1(config-router-af)#exit
PE1(config-router)#exit
```

P 交换机的配置如下：

(1) 全局使能 MPLS 并且配置相关接口的 LDP

```
P#config
P(config)#mpls enable
P(config)#router ldp
P(config-router)#exit
P(config)# interface loopback 1
P(config-if-Loopback1)# ip address 172.3.3.3 255.255.255.255
P(config-if-Loopback1)# exit
P(config)#interface vlan 100
P(config-if-Vlan100)#ip address 100.1.1.2 255.255.255.0
P(config-if-Vlan100)#label-switching
P(config-if-Vlan100)#ldp enable
P(config-if-Vlan100)#exit
P(config)#interface vlan200
P(config-if-Vlan200)#ip address 200.1.1.2 255.255.255.0
P(config-if-Vlan200)#label-switching
P(config-if-Vlan200)#ldp enable
P(config-if-Vlan200)#exit
(2) 配置 OSPF 协议
P(config)#router ospf
P(config-router)# ospf router-id 172.3.3.3
P(config-router)# network 0.0.0.0/0 area 0
P(config-router)# redistribute connected
```

3.3.4 PE-CE间使用静态路由建立BGP MPLS VPN

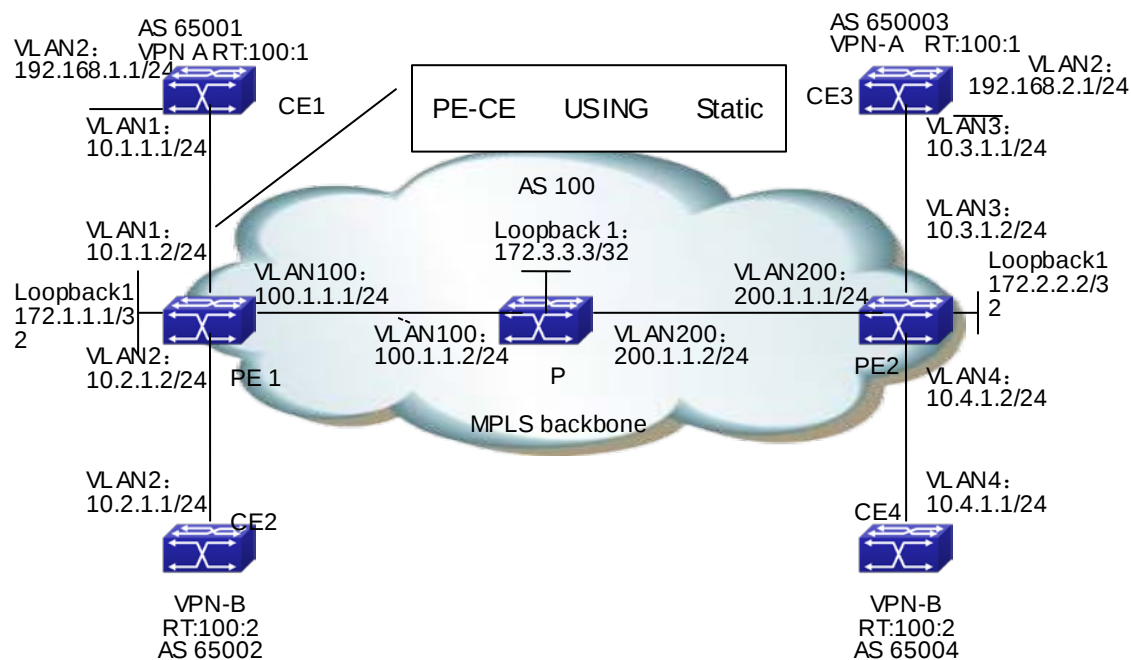


图 3-11 PE-CE 间使用静态路由建立 BGP MPLS VPN

CE1 的配置如下: (CE2~CE4 另外三个设备配置与 CE1 类似, 所以省略)

```
CE1#config
CE1(config)# interface vlan 1
CE1(config-if-Vlan1)#ip address 10.1.1.1 255.255.255.0
CE1(config-if-Vlan1)#exit
CE1(config)# interface loopback 1
CE1(config-if-Vlan1)#ip address 192.168.1.1 255.255.255.0
CE1(config-if-Vlan1)# exit
CE1(config)# ip route vrf vpna 192.168.2.1/24 10.1.1.2
```

PE1 交换机的 MPLS BGP 配置如下: (PE2 配置类似, 所以省略)

(1) 配置 VPN 实例

```
PE1#config
PE1(config)#ip vrf vpna
PE1(config-vrf)#rd 100:1
PE1(config-vrf)#route-target both 100:1
PE1(config)#ip vrf vpnb
PE1(config-vrf)#rd 100:2
PE1(config-vrf)#route-target both 100:2
```

(2) 配置接口绑定 VPN 实例

```
PE1(config)# interface vlan 1
PE1(config-if-Vlan1)# ip vrf forwarding vpna
PE1(config-if-Vlan1)#ip address 10.1.1.2 255.255.255.0
PE1(config-if-Vlan1)#exit
PE1(config)# interface vlan 2
```

```
PE1(config-if-Vlan2)# ip vrf forwarding vpnb
PE1(config-if-Vlan1)#ip address 10.2.1.2 255.255.255.0
PE1(config-if-Vlan1)#exit
(3) 全局启动 MPLS 和 LDP
PE1(config)#mpls enable
PE1(config)#router ldp
PE1(config-router)#exit
(4) 配置接口以及使能 LDP
PE1(config)# interface loopback 1
PE1(config-if-Loopback1)# ip address 172.1.1.1 255.255.255.255
PE1(config-if-Loopback1)# exit
PE1(config)# interface vlan 100
PE1(config-if-Vlan100)#ip address 100.1.1.1 255.255.255.0
PE1(config-if-Vlan100) #ldp enable
PE1(config-if-Vlan100)#exit
(5) 启动 OSPF 协议传送内网路由
PE1(config)#router ospf
PE1(config-router)# ospf router-id 172.1.1.1
PE1(config-router)# network 0.0.0.0/0 area 0
PE1(config-router)# redistribute connected
PE1(config-router)#exit
(6) 配置静态私网路由
PE1(config)# ip route vrf vpna 192.168.1.1/24 10.1.1.2
PE1(config)# ip route vrf vpnb 192.168.2.1/24 10.1.1.2
PE1(config-router)#address-family ipv4 vrf vpna
PE1(config-router-af)#network 0.0.0.0/0
PE1(config-router-af)#redistribute connected
PE1(config-router)#exit
(7) 配置 BGP
PE1(config)# router bgp 100
PE1(config-router)#neighbor 172.2.2.2 remote-as 100
PE1(config-router)#neighbor 172.2.2.2 update-source 172.1.1.1
PE1(config-router)#address-family vpnv4
PE1(config-router-af)#neighbor 172.2.2.2 activate
PE1(config-router-af)#exit
PE1(config-router)# address-family ipv4 vrf vpna
PE1(config-router-af)#redistribute connected
PE1(config-router-af)#redistribute static
PE1(config-router-af)#exit
PE1(config-router)# address-family ipv4 vrf vpnb
PE1(config-router-af)#redistribute connected
PE1(config-router-af)# redistribute static
PE1(config-router-af)#exit
PE1(config-router)#exit
```

P 交换机的配置如下：

(1) 全局使能 MPLS 并且配置相关接口的 LDP

```
P#config
P(config)#mpls enable
P(config)#router ldp
P(config-router)#exit
P(config)# interface loopback 1
P(config-if-Loopback1)# ip address 172.3.3.3 255.255.255.255
P(config-if-Loopback1)# exit
P(config)#interface vlan 100
P(config-if-Vlan100)#ip address 100.1.1.2 255.255.255.0
P(config-if-Vlan100)#ldp enable
P(config-if-Vlan100)#exit
P(config)#interface vlan200
P(config-if-Vlan200)#ip address 200.1.1.2 255.255.255.0
P(config-if-Vlan100)#ldp enable
P(config-if-Vlan200)#exit
(2) 配置 OSPF 协议
P(config)#router ospf
P(config-router)# ospf router-id 172.3.3.3
P(config-router)# network 0.0.0.0/0 area 0
P(config-router)# redistribute connected
```

3.4 MPLS BGP VPN排错帮助

在配置、使用 MPLS BGP VPN 协议时，可能会由于物理连接、配置错误等原因导致 VPN 未能正常运行。因此，用户应注意以下要点：

- ☞ 首先应该保证 PE1、P、PE2 之间的 OSPF 邻居建立正常，包括 loopback 口在内的路由发布正常，并且 PE 之间的 BGP 邻居成功建立。
- ☞ 其次应该保证 PE1、P、PE2 上全局启动了 LDP 并且在活动接口上正确启动了 LDP，检查下 PE1、P、PE2 上的 LDP session 建立是否正常；
- ☞ 然后，确认选取建立 VPN 所采用的 PE-CE 间的传送路由方式，以及对应的配置是否正确，是否在 CE 上发布了相关的私网路由到对端 PE，同时需要注意 CE 端是不需要配置 VRF 实例的。如果使用 EBGP 传送私网路由，注意 CE1 和 CE2 之间的 BGP 自治系统号不要相同，否则 BGP 的环路检测会过滤对应的私网路由；
- ☞ 接着，确认 PE 上 BGP VPN 实例的配置正确，如果采用 OSPF 或者 RIP 建立传送 PE-CE 间私网路由时，注意要引入 BGP 路由，并且在 BGP VPN 实例下引入对应的 OSPF 和 RIP 路由，正确配置后在 PE1 上执行 `show ip bgp vpnv4 all` 可以看到 CE1、CE2 的路由信息。之后在 PE 上执行 `show mpls vrf-table` 可以看对应的私网路由已经分配了标签，并且其状态是 UP 的，如果对应的私网路由的 vrf-table 中 Oper Status 状态为 down，那么执行 `show mpls ftn-table` 查看对应的 FEC 是否生成了 ftn；
- ☞ 最后，在上述过程都正确的情况下，在 CE1 和 CE2 上使用 `show ip route` 命令可以查看到自己 VPN 内的正确路由信息。除非已经对 BGP MPLS VPN 非常熟悉，

我们不推荐用户采用静态路由的方式建立 VPN;

- ☞ 另外，如果正确的配置下保存重启设备后，CE 上没有立刻看到对端 CE 的路由，请耐心等待一段时间，OSPF、LDP、BGP 的连接建立以及发布路由都需要时间。

第4章 MPLS VPN访问公网

4.1 访问公网简介

VPN 访问公网是指 VPN 站点能够访问公共 Internet，协议基本规范由 RFC4364 定义。RFC4364 定义了几种 VPN 访问 Internet 的方式：

- ☞ 非 VRF Internet 接入
- ☞ VRF Internet 接入方式一
- ☞ VRF Internet 接入方式三

4.1.1 非VRF Internet接入方式

非 VRF Internet 接入方式中，如图 4-1，PE 路由器与 Internet 网关通过非 VRF 接口通信，VPN 站点访问 Internet 的往返流量经过 PE 路由器的全局路由表转发。具有访问 Internet 功能的 CE 路由器与 PE 路由器存在两条连接，一条连接与 PE 的公网接口相连（公网连接），另一条连接与 PE 的私网接口相连（私网连接）。PE 路由器的全局路由表可以包含完整的 Internet 路由或者一部分，或者只有一条指向 Internet 网关的默认路由，CE 路由器通过公网连接学习 Internet 路由，并将 VPN 站点内全球注册的 IP 地址子网路由通过公网连接发布给 PE，PE 再将这些路由发布到 Internet 网关，最后发布到 Internet。VPN 站点访问 Internet 的往返流量也通过公网连接。CE 和 PE 之间的私网连接用于 CE 路由器学习和发布 VPN 内的私网路由，VPN 站点间也通过私网连接通信，经由 PE 路由器 VRF 路由表转发。在这种方式中，PE 路由器的全局路由表和 VRF 路由表之间完全隔离，VPN 路由分发和 Internet 路由分发过程完全独立。

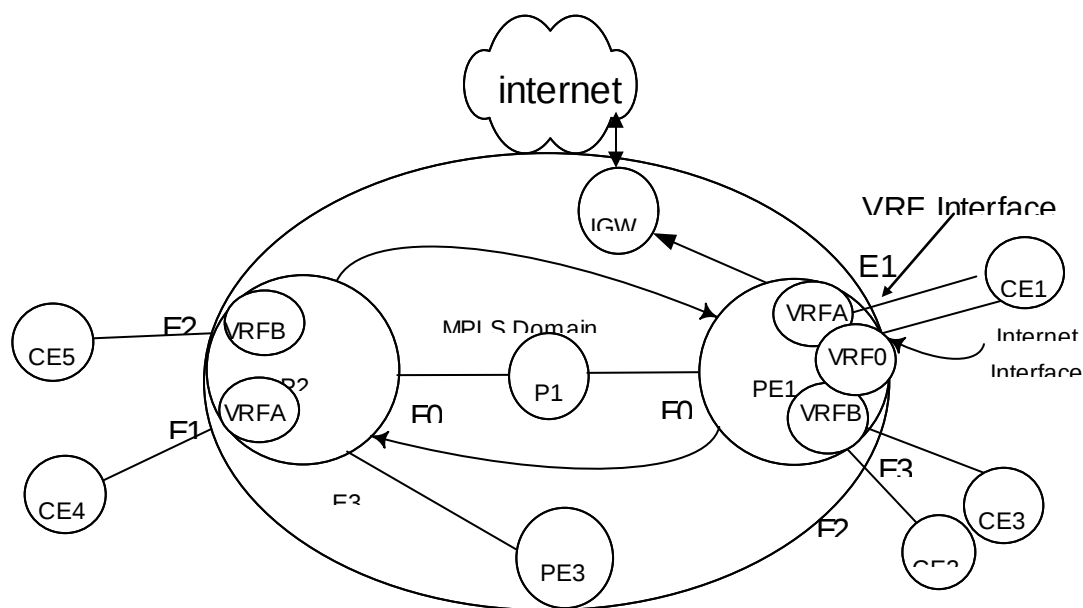


图 4-1 非 VRF Internet 接入

4.1.2 VRF Internet接入方式一

VRF Internet 接入方式中,如图 4-2, PE 路由器与 Internet 网关通过非 VRF 接口通信, VPN 站点访问其他站点的 VPN 往返流量和访问 Internet 的往返流量均经过 CE 和 PE 的私网连接。PE 路由器包含完整的 Internet 路由或者只有一条到达 Internet 网关的缺省路由。VPN 访问 Internet 的 IP 报文到达 PE 的 VRF 接口时,如果在 VRF 路由表中查找失败,可以跳转到全局路由表中查找,如果命中,则报文可转发到 Internet 网关,再通过 Internet 网关转发到 Internet。为了实现 Internet 主机回访 VPN 站点,需要在 PE 全局路由表创建一条特殊的静态路由,该路由的目的网段为 VPN 站点内全球注册的 IP 地址子网,出接口为指向 VPN 站点的私网接口,下一跳为 CE 路由器,这条静态路由通过 PE 发布到 Internet 网关,经过 Internet 网关发布到 Internet。Internet 主机回访 VPN 站点的报文到达 PE 的公网接口,查找 PE 的全局路由表,如果匹配指向 VPN 站点的静态路由,则将报文通过私网接口转发到下一跳。在这种方式中, PE 路由器的全局路由表表和 VRF 路由表并不是完全隔离的, PE 路由器的全局路由表包含了部分的 VPN 路由。

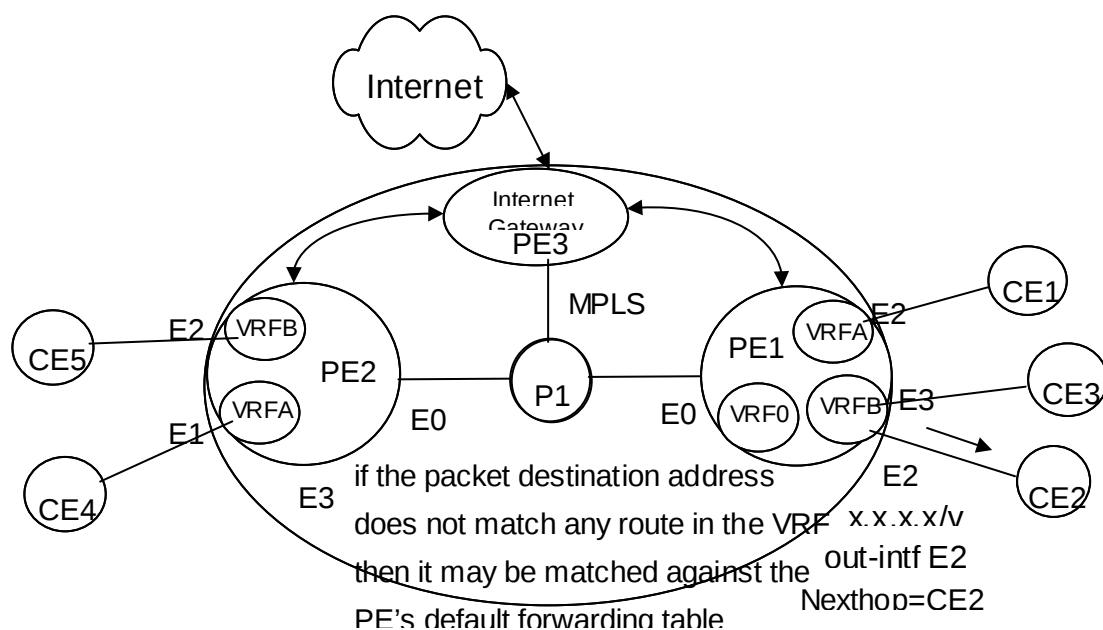


图 4-2 VRF Internet 接入方式一

4.1.3 VRF Internet接入方式三

VRF Internet 接入 3 方式中,如图 4-3, VPN 站点通过 PE 和 CE 间的私网连接访问 Internet。PE 路由器的 VRF 路由表包含 Internet 路由。这些 Internet 路由是通过与 Internet 网关相连的 PE 路由器(称为 Internet PE)学到的, Internet PE 创建一个 Internet VRF,通过绑定 Internet VRF 的接口与 Internet 网关相连,这样 Internet 网关可将 Internet 路由发布给 Internet PE 的 VRF 路由表中,这些路由以 VPNv4 路由发布到其他 PE 路由器的 VRF 中。需访问 Internet 的 VPN 站点相连的 PE 路由器将对应的 VRF 路由(最好仅为目的网段

为 VPN 站点内全球注册的 IP 地址子网的路由) 通过 VPNv4 路由发布到 Internet PE, 安装在 Internet VRF 中, Internet 网关再将这些路由发布到 Internet。这些路由的引入引出策略依赖于 MBGP 和 vrf 配置的 route-target。注意的是, 在这种方式中, 可以访问 Internet 的 VPN 站点是不允许出现地址或路由重叠的。

值得注意的是本种方式需要在 PE 上引入大量 Internet 路由, 不建议用户使用此种方式实现访问公网。

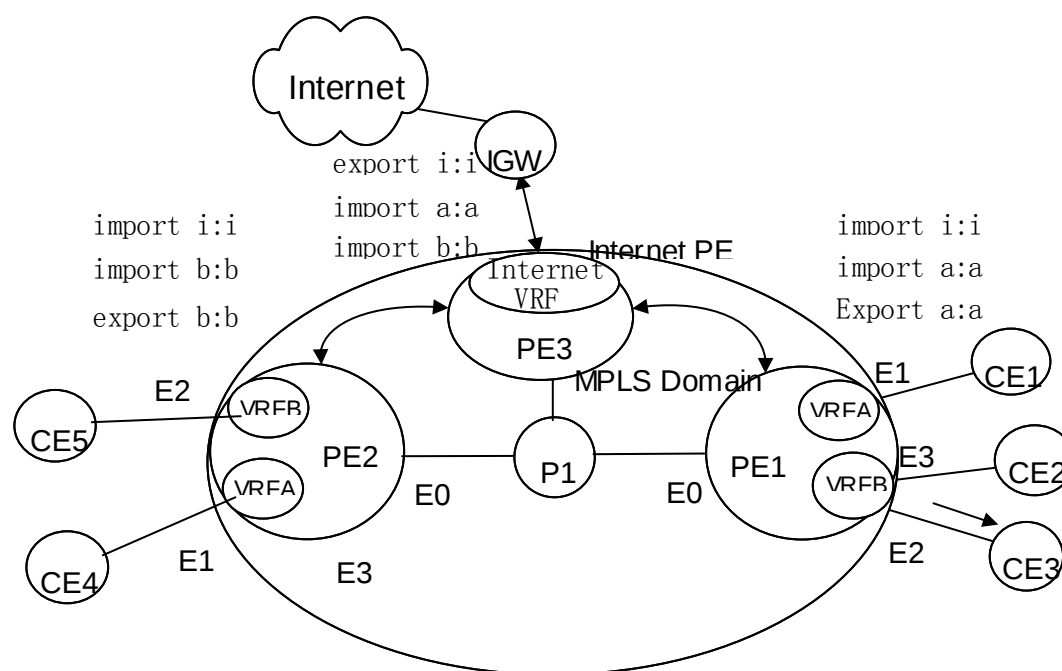


图 4-3 VRF Internet 接入 3

4.2 访问公网配置

访问公网配置任务序列:

1. 配置非 VRF Internet 接入
 - (1) 配置普通的 L3VPN
 - (2) 在 CE 和 PE 上增加一条公网连接, 连接接口为 non-VRF 接口。
 - (3) 在 CE 上进行路由过滤, 将公网路由通过公网连接发送到 PE, 私网路由通过私网连接发送到 PE, 分别在全局路由表和相应的 VRF 表中维护。
 - (4) 在 PE 和 CE 的公网接口上配置合适的过滤策略, 将源地址和目的地址为私网地址的报文过滤掉。
 - (5) 配置缺省路由
 - 1) IGW 将缺省路由引入 BGP 协议
 - 2) PE 将缺省路由通过公网连接发布到 CE
 - 3) CE 通过私网连接将缺省路由发布到 PE, 进而发布到其他 CE 上
 - (6) 配置静态路由
 - 1) 在 CE1 上配置静态路由, 指向 Internet。
 - 2) 在 PE1 上配置静态路由, 指向 CE 的公网接口。

2. 配置 VRF Internet 接入方式一

- (1) 配置普通的 L3VPN
- (2) PE 的私网接口配置 `ip vrf forwarding VPNA fallback global`。
- (3) 配置 3 条静态路由，具体如下：
 - 1) 在 CE 上配置一条默认路由，下一跳为代理服务器。
 - 2) 在 PE 上增加一条到 Internet 的默认路由，下一跳为 IGW。PE 通过 OSPF 通告一条缺省路由，下一跳为自己。
 - 3) 在 PE 全局路由表上增加一条从 Internet 回到代理服务器的静态路由，目的为 VPN 公网地址，下一跳为代理服务器，并将该路由通过 OSPF 通告给其他 PE。

配置非 VRF Internet 接入

该配置不涉及单独的命令行，按照配置序列完成配置即可。具体的命令参见相应功能的配置说明。

配置 VRF Internet 接入方式一

1. 配置 VRF Internet 接入方式一

- (1) 配置普通的 L3VPN
- (2) PE 的私网接口配置 `ip vrf forwarding VPNA fallback global`。
- (3) 配置 3 条静态路由，具体如下：
 - 1) 在 CE 上配置一条默认路由，下一跳为代理服务器。
 - 2) 在 PE 上增加一条到 Internet 的默认路由，下一跳为 IGW。PE 通过 OSPF 通告一条缺省路由，下一跳为自己。
 - 3) 在 PE 全局路由表上增加一条从 Internet 回到代理服务器的静态路由，目的为 VPN 公网地址，下一跳为代理服务器，并将该路由通过 OSPF 通告给其他 PE。

命令	解释
配置普通的 L3VPN	参见 BGP MPLS VPN 配置部分
接口配置模式	
<code>ip vrf forwarding <vrf_name> fallback global no ip vrf forwarding <vrf_name> fallback global</code>	必选 配置 VRF 路由表的全局二次查找功能。缺省情况下未配置，在配置二次查找以前需要先取消接口视图下 VRF 的配置。
全局配置模式	
<code>ip route vrf <vrf-name> {<ip-prefix> <mask> <ip-prefix/prefix-length>} {<gateway-address> null0} no ip route vrf <vrf-name> {<ip-prefix> <mask> <ip-prefix/<prefix-length>} {<gateway-address> null0}</code>	必选 配置静态路由。这里需要配置三条，一条是 CE1 的默认路由，一条是 PE3 上到 Internet 的默认路由，最后一条是 PE1 上从 Internet 回到代理服务器的静态路由。

4.3 访问公网典型案例

4.3.1 非VRF Internet接入方式

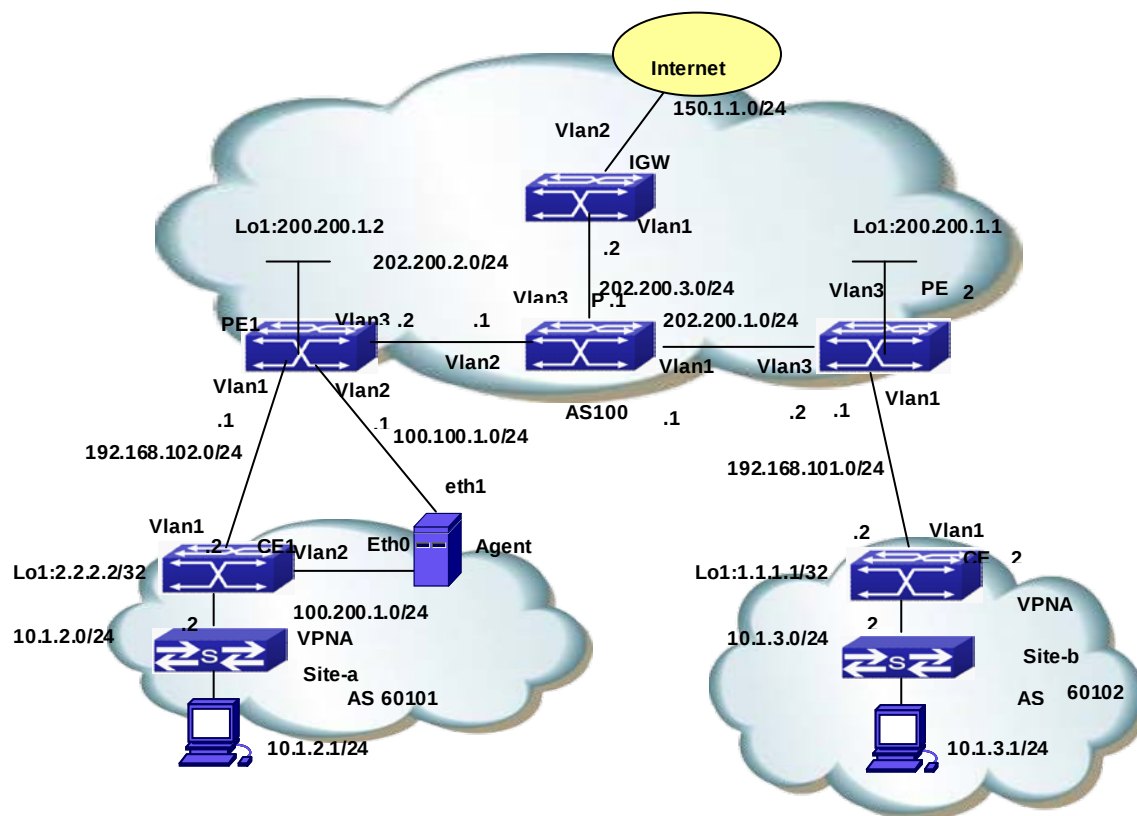


图 4-4 非 VRF Internet 接入方式

CE1 的配置如下:

```
CE1#config
CE1(config)#access-list 1 deny 100.100.1.0 0.0.0.255
CE1(config)#access-list 1 deny 100.200.1.0 0.0.0.255
CE1(config)#access-list 1 permit any-source
CE1(config)#access-list 2 permit 10.1.1.0 0.0.0.255
CE1(config)#access-list 2 permit 10.1.2.0 0.0.0.255
CE1(config)#access-list 2 deny any-source
CE1(config)# interface vlan 1
CE1(config-if-Vlan1)#ip address 192.168.102.2 255.255.255.0
CE1(config-if-Vlan1)#exit
CE1(config)# interface vlan 2
CE1(config-if-Vlan1)#ip address 100.200.1.2 255.255.255.0
CE1(config-if-Vlan1)#exit
CE1(config)# interface vlan 3
CE1(config-if-Vlan1)#ip address 10.1.2.2 255.255.255.0
```

```
CE1(config-if-Vlan1)#exit
CE1(config)# interface loopback 1
CE1(config-if-Vlan1)#ip address 2.2.2.2 255.255.255.255
CE1(config-if-Vlan1)# exit
CE1(config)#router bgp 60102
CE1(config-router)#network 120.1.1.0/24
CE1(config-router)#network 120.1.2.0/24
CE1(config-router)#network 10.1.2.0/24
CE1(config-router)#redistribute connected
CE1(config-router)#neighbor 100.100.1.1 remote-as 100
CE1(config-router)#neighbor 100.100.1.1 distribute-list 2 out
CE1(config-router)#neighbor 192.168.102.1 remote-as 100
CE1(config-router)#neighbor 192.168.102.1 default-originate
CE1(config-router)#neighbor 192.168.102.1 distribute-list 1 out
CE1(config-router)#exit
CE1(config)# ip route 100.100.1.1 255.255.255.0 100.200.1.1
CE1(config)# ip route 0.0.0.0/0 100.200.1.1
CE1(config)# exit
```

PE1 配置如下:

```
PE1#config
PE1(config)#access-list 100 deny ip 10.1.2.0 0.0.0.255 any-destination
PE1(config)#access-list 100 deny ip 10.1.2.0 0.0.0.255 any-destination
PE1(config)#access-list 100 deny ip 10.1.3.0 0.0.0.255 any-destination
PE1(config)#access-list 100 deny ip anysource 200.200.1.0 0.0.0.255
PE1(config)#access-list 100 deny ip anysource 202.200.0.0 0.0.255.255
PE1(config)#firewall enable
PE1(config-vrf)#ip vrf VRF-A
PE1(config-vrf)#rd 100:10
PE1(config-vrf)#route-target both 100:10
PE1(config-vrf)#exit
PE1(config)#interface vlan1
PE1(config-if-Vlan1)#ip vrf forwarding VRF-A
PE1(config-if-Vlan1)#ip address 192.168.102.1 255.255.255.0
PE1(config-if-Vlan1)#exit
PE1(config)#interface vlan2
PE1(config-if-Vlan2)#ip address 100.100.1.1 255.255.255.0
PE1(config-if-Vlan2)#ip access-group 1 in
PE1(config-if-Vlan2)#exit
PE1(config)# interface vlan3
PE1(config-if-Vlan3)#label-switching
PE1(config-if-Vlan3)#ldp enable
PE1(config-if-Vlan3)#ip address 202.200.2.2 255.255.255.0
PE1(config-if-Vlan3)#exit
PE1(config)#interface Loopback1
```

```
PE1(config)#ip address 200.200.1.2 255.255.255.255
PE1(config)#router ospf
PE1(config-router)#network 200.200.1.2/32 area 0
PE1(config-router)#network 202.200.2.0/24 area 0
PE1(config-router)#exit
PE1(config)#router bgp 100
PE1(config-router)#neighbor 100.200.1.2 remote-as 60102
PE1(config-router)#neighbor 200.200.1.1 remote-as 100
PE1(config-router)#neighbor 202.200.3.2 remote-as 100
PE1(config-router)#neighbor 202.200.3.2 next-hop-self
PE1(config-router)#address-family vpnv4 unicast
PE1(config-router-af)#neighbor 200.200.1.1 activate
PE1(config-router-af)#exit-address-family
PE1(config-router)#address-family ipv4 vrf VRF-A
PE1(config-router-af)#neighbor 192.168.102.2 remote-as 60102
PE1(config-router-af)#no neighbor 192.168.102.2 send-community extended
PE1(config-router-af)#exit-address-family
PE1(config-router)#exit
PE1(config)# router ldp
PE1(config-router)#ip route 100.200.1.2 255.255.255.0 100.100.1.2
```

P 配置如下

```
P#config
P(config)#interface Vlan1
P(config-if-Vlan1)#label-switching
P(config-if-Vlan1)#ldp enable
P(config-if-Vlan1)#ip address 202.200.1.1 255.255.255.0
P(config-if-Vlan1)#exit
P(config)#interface Vlan2
P(config-if-Vlan2)#label-switching
P(config-if-Vlan2)#ldp enable
P(config-if-Vlan2)#ip address 202.200.2.1 255.255.255.0
P(config-if-Vlan2)#exit
P(config)#interface Vlan3
P(config-if-Vlan3)#ip address 202.200.3.1 255.255.255.0
P(config-if-Vlan3)#exit
P(config)#router ospf
P(config-router)#network 202.200.1.0/24 area 0
P(config-router)#network 202.200.2.0/24 area 0
P(config-router)#network 202.200.3.0/24 area 0
P(config-router)#exit
P(config)#router ldp
```

PE2 如下配置:

```
PE2#config
PE2(config)#ip vrf VRF-A
```

```
PE2(config-vrf)#rd 100:10
PE2(config-vrf)#route-target both 100:10
PE2(config-vrf)#exit
PE2(config)#interface Vlan1
PE2(config-if-Vlan1)#ip vrf forwarding VRF-A
PE2(config-if-Vlan1)#ip address 192.168.101.1 255.255.255.0
PE2(config-if-Vlan1)#exit
PE2(config)#interface Vlan2
PE2(config-if-Vlan2)#label-switching
PE2(config-if-Vlan2)#ldp enable
PE2(config-if-Vlan2)#ip address 202.200.1.2 255.255.255.0
PE2(config-if-Vlan2)#exit
PE2(config)#interface Loopback1
PE2(config-if-loopback1)#ip address 200.200.1.1 255.255.255.255
PE2(config-if-loopback1)#exit
PE2(config)#router ospf
PE2(config-router)#network 200.200.1.1/32 area 0
PE2(config-router)#network 202.200.1.0/24 area 0
PE2(config-router)#exit
PE2(config)#router bgp 100
PE2(config-router)#address-family vpnv4 unicast
PE2(config-router-af)#neighbor 200.200.1.1 activate
PE2(config-router-af)#exit-address-family
PE2(config-router)#address-family ipv4 vrf VRF-A
PE2(config-router-af)#neighbor 192.168.101.2 remote-as 60101
PE2(config-router-af)#no neighbor 192.168.101.2 send-community extended
PE2(config-router-af)#exit-address-family
PE2(config-router)#exit
PE2(config)#router ldp
```

CE2 配置如下:

```
CE2#config
CE2(config)#interface vlan 1
CE2(config-if-Vlan1)#ip address 192.168.101.2 255.255.255.0
CE2(config-if-Vlan1)#exit
CE2(config)#interface Loopback1
CE2(config-if-Loopback1)#ip address 1.1.1.1 255.255.255.255
CE2(config-if-Loopback1)#exit
CE2(config)#router bgp 60101
CE2(config-router)#network 10.1.3.0/24
CE2(config-router)#neighbor 192.168.101.1 remote-as 100
```

IGW 如下配置:

```
IGW#config
IGW(config)#interface Vlan1
IGW(config-if-Vlan1)#ip address 202.200.3.2 255.255.255.0
```

```

IGW(config-if-Vlan1)#exit
IGW(config)#interface Vlan2
IGW(config-if-Vlan2)#ip address 150.1.1.1 255.255.255.0
IGW(config-if-Vlan2)#exit
IGW(config)#router ospf
IGW(config-router)#network 202.200.3.0 0.0.0.255 area 0
IGW(config-router)#exit
IGW(config)#router bgp 100
IGW(config-router)#neighbor 202.200.2.2 remote-as 100
IGW(config-router)#neighbor 202.200.2.2 default-originate

```

4.3.2 VRF Internet接入方式一

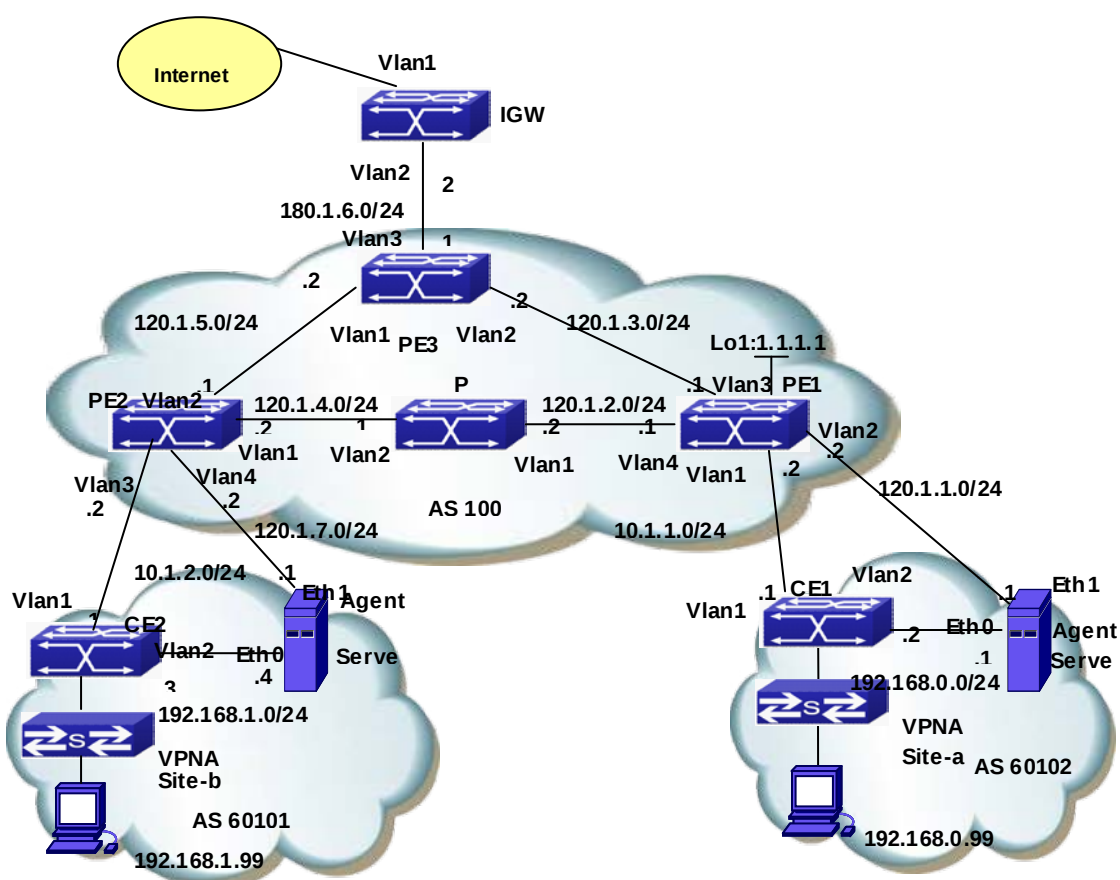


图 4-5 VRF Internet 接入方式一

site-a 和 site-b 属于 VPNA, site-a 和 site-b 之间的用户可以互访, site-a 和 site-b 的用户均有访问 Internet 的需求, 在 site-a 和 site-b 内单独配置代理服务器实现 site-a 和 site-b 用户使用私网地址访问 Internet 进行 NAT 转换。

CE1 的配置如下

```

CE1#config
CE1(config)#interface Vlan1
CE1(config-if-Vlan1)#ip address 10.1.1.1 255.255.255.0

```

```
CE1(config-if-Vlan1)#exit
CE1(config)#interface Vlan2
CE1(config-if-Vlan2)#ip address 192.168.0.2 255.255.255.0
CE1(config-if-Vlan2)#exit
CE1(config)#interface loopback1
CE1(config-if-Loopback1)#ip address 11.11.11.11 255.255.255.255
CE1(config-if-Loopback1)#exit
CE1(config)#ip route 0.0.0.0/0 192.168.0.1
CE1(config)#router bgp 60101
CE1(config-router)#neighbor 10.1.1.2 remote-as 100
CE1(config-router)#network 192.168.0.0/24
```

PE1 的配置文件

```
PE1#config
PE1(config)#ip vrf VPNA
PE1(config-vrf)#rd 100:10
PE1(config-vrf)#route-target both 100:10
PE1(config-vrf)#exit
PE1(config)#interface Vlan1
PE1(config-if-Vlan1)#ip vrf forwarding VPNA
PE1(config-if-Vlan1)#ip address 10.1.1.1 255.255.255.0
PE1(config-if-Vlan1)#exit
PE1(config)#interface Vlan2
PE1(config-if-Vlan2)#ip vrf forwarding VPNA fallback global
PE1(config-if-Vlan2)#ip address 120.1.1.2 255.255.255.0
PE1(config-if-Vlan2)#exit
PE1(config)#interface Vlan3
PE1(config-if-Vlan3)#ip address 120.1.3.1 255.255.255.0
PE1(config-if-Vlan2)#exit
PE1(config)#interface Vlan4
PE1(config-if-Vlan4)#label-switching
PE1(config-if-Vlan4)#ldp enable
PE1(config-if-Vlan4)#ip address 202.200.1.2 255.255.255.0
PE1(config-if-Vlan4)#exit
PE1(config)#interface loopback1
PE1(config-if-Loopback1)#ip address 1.1.1.1 255.255.255.255
PE1(config-if-Loopback1)#exit
PE1(config)#router ospf
PE1(config-router)#redistribute static
PE1(config-router)#network 1.1.1.1/32 area 0
PE1(config-router)#network 120.1.2.0/24 area 0
PE1(config-router)#network 120.1.3.0/24 area 0
PE1(config-router)#exit
PE1(config)#router bgp 100
PE1(config-router)#neighbor 2.2.2.2 remote-as 100
```

```
PE1(config-router)#neighbor 2.2.2.2 update-source 1.1.1.1
PE1(config-router)#address-family vpnv4 unicast
PE1(config-router-af)#neighbor 2.2.2.2 activate
PE1(config-router-af)#exit-address-family
PE1(config-router)#address-family ipv4 vrf VPNA
PE1(config-router-af)#network 120.1.1.0/24
PE1(config-router-af)#neighbor 10.1.1.1 remote-as 60101
PE1(config-router-af)#no neighbor 10.1.1.1 send-community extended
PE1(config-router-af)#exit-address-family
PE1(config-router)#exit
PE1(config)#router ldp
PE1(config-router)#exit
PE1(config)#ip route 120.1.1.0/24 vlan 2 120.1.1.1
```

P 的配置文件

```
P#config
P(config)#interface Vlan1
P(config-if-Vlan1)#label-switching
P(config-if-Vlan1)#ldp enable
P(config-if-Vlan1)#ip address 120.1.2.2 255.255.255.0
P(config-if-Vlan1)#exit
P(config)#interface Vlan2
P(config-if-Vlan2)#label-switching
P(config-if-Vlan2)#ldp enable
P(config-if-Vlan2)#ip address 120.1.4.1 255.255.255.0
P(config-if-Vlan2)#exit
P(config)#router ospf
P(config-router)#network 0.0.0.0/0 area 0
P(config-router)#exit
P(config)#router ldp
```

PE2 的配置文件

```
PE2#config
PE2(config)#ip vrf VPNA
PE2(config-vrf)#rd 100:10
PE2(config-vrf)#route-target both 100:10
PE2(config-vrf)#exit
PE2(config)#interface Vlan1
PE2(config-if-Vlan1)#label-switching
PE2(config-if-Vlan1)#ldp enable
PE2(config-if-Vlan1)#ip address 120.1.4.2 255.255.255.0
PE2(config-if-Vlan1)#exit
PE2(config)#interface Vlan2
PE2(config-if-Vlan2)#ip address 120.1.5.1 255.255.255.0
PE2(config-if-Vlan2)#exit
PE2(config)#interface Vlan3
```

```
PE2(config-if-Vlan3)#ip vrf forwarding VPNA
PE2(config-if-Vlan1)#ip address 10.1.2.2 255.255.255.0
PE2(config-if-Vlan1)#exit
PE2(config)#interface Vlan4
PE2(config-if-Vlan4)#ip vrf forwarding VPNA fallback global
PE2(config-if-Vlan4)#ip address 120.1.7.2 255.255.255.0
PE2(config-if-Vlan4)#exit
PE2(config)#interface Loopback1
PE2(config-if-Loopback1)#ip address 2.2.2.2 255.255.255.255
PE2(config-if-Loopback1)#exit
PE2(config)#router ospf
PE2(config-router)#redistribute static
PE2(config-router)#network 2.2.2.2/32 area 0
PE2(config-router)#network 120.1.4.0/24 area 0
PE2(config-router)#network 120.1.5.0/24 area 0
PE2(config-router)#exit
PE2(config)#router bgp 100
PE2(config-router)#neighbor 1.1.1.1 remote-as 100
PE2(config-router)#neighbor 1.1.1.1 update-source 2.2.2.2
PE2(config-router)#address-family vpnv4 unicast
PE2(config-router-af)#neighbor 1.1.1.1 activate
PE2(config-router-af)#exit-address-family
PE2(config-router)#address-family ipv4 vrf VPNA
PE2(config-router-af)#network 120.1.7.0/24
PE2(config-router-af)#neighbor 10.1.2.1 remote-as 60102
PE2(config-router-af)#no neighbor 10.2.1.1 send-community extended
PE2(config-router-af)#exit-address-family
PE2(config-router)#exit
PE2(config)#router ldp
PE2(config-router)#exit
PE2(config)#ip route 120.1.7.0/24 vln 4 120.1.7.1
```

PE3 的配置文件

```
PE3#config
PE3(config)#interface Loopback1
PE3(config-if-Loopback1)#ip address 3.3.3.3 255.255.255.255
PE3(config-if-Loopback1)#exit
PE3(config-if-Vlan1)#interface Vlan1
PE3(config-if-Vlan1)#ip address 120.1.5.2 255.255.255.0
PE3(config-if-Vlan1)#exit
PE3(config)#interface Vlan2
PE3(config-if-Vlan2)#ip address 120.1.3.2 255.255.255.0
PE3(config-if-Vlan2)#exit
PE3(config)#interface Vlan3
PE3(config-if-Vlan3)#ip address 180.1.6.1 255.255.255.0
```

```
PE3(config-if-Vlan3)#exit
PE3(config)#router ospf 1
PE3(config-router)#default-information originate
PE3(config-router)# network 0.0.0.0/0 area 0
PE3(config-router)#exit
PE3(config)#router bgp 100
PE3(config-router)#network 120.1.1.0 mask 255.255.255.0
PE3(config-router)#network 120.1.7.0 mask 255.255.255.0
PE3(config-router)#neighbor 180.1.6.2 remote-as 200
PE3(config-router)#exit
PE3(config)#ip route 0.0.0.0/0 180.1.6.2
```

CE2 配置如下

```
CE2#config
CE2(config)#interface Vlan1
CE2(config-if-Vlan1)#ip address 10.1.2.1 255.255.255.0
CE2(config-if-Vlan1)#exit
CE2(config)#interface Vlan2
CE2(config-if-Vlan2)#ip address 192.168.1.3 255.255.255.0
CE2(config-if-Vlan2)#exit
CE2(config-if-Loopback1)#interface Loopback1
CE2(config-if- Loopback1)#ip address 22.22.22.22 255.255.255.255
CE2(config-if- Loopback1)#exit
CE2(config)#ip route 0.0.0.0/0 192.168.1.4
CE2(config)#router bgp 60101
CE2(config-router)#neighbor 10.1.2.2 remote-as 100
CE2(config-router)#network 192.168.1.0/24
CE2(config-router)#exit
```

IGW 的配置文件

```
IGW#config
IGW(config)#interface Vlan1
IGW(config-if-Vlan1)#ip address 180.1.5.2 255.255.255.0
IGW(config-if-Vlan1)#exit
IGW(config)#interface Vlan2
IGW(config-if-Vlan2)#ip address 180.1.6.2 255.255.255.0
IGW(config-if-Vlan2)#exit
IGW(config)#router bgp 200
IGW(config-router)#neighbor 180.1.6.1 remote-as 100
IGW(config-router)#exit
```

4.4 访问公网排错帮助

在配置、使用访问公网功能时，可能会由于物理连接、配置错误等原因导致未能正常运行。因此，用户应注意以下要点：

- ☞ 首先应该保证普通的 MPLS BGP VPN 功能正常运行，私网间可以互通，如果 VPN 不通请参考第三章 MPLS BGP VPN 的排错帮助。
- ☞ 其次应该确认所采用的访问公网是 VRF 接入方式还是非 VRF 接入方式，二者配置有较大的不同；
- ☞ 然后，如果是非 VRF 接入方式请注意一定要在 PE-CE 的非 VRF 接口上配置过滤规则，禁止私网路由和流量通过公网接口进入 PE 中，否则会有安全隐患；此外注意默认路由的发布以及 IGW 的 NAT 配置是否正确；
- ☞ 接着，如果是 VRF 接入方式一请注意配置私网接口时需要使用 `ip vrf forwarding vrf_name fallback global` 命令，确保在私网路由查找不到的时候才会二次查找全局路由表；此外注意公网回来的路由发布是否正确以及 IGW 的 NAT 配置是否正确；
- ☞ 最后，在上述过程都正确的情况下，在 CE 上都可以访问 Internet，如上所述的组网中，不管哪种方式其他 CE 都是通过 VPN 将流量发布到 PE 上以后访问 Internet 的，从 Internet 返回的流量也必须经过 PE 后转发。

第5章 VPLS

5.1 VPLS简介

- ✎ 当今 IP 网络已经遍布全球，利用现有 IP 网络为企业低成本专网逐渐成为各大运营商的关注点。因此，一种在 IP 网上提供 VPN 服务、可方便设定速率、配置简单的技术应运而生，这种技术即 MPLS VPN 技术。基于 MPLS 的 VPN 技术有两种，分别是 MPLS L3VPN 和 MPLS L2VPN。MPLS L3VPN 需要介入用户的内部路由管理，运营商的管理比较复杂。传统 VLL 方式的 MPLS L2VPN，在公网中提供一种点到点的 L2VPN 业务，可以让两个站点之间的连接效果像直接用链路连接一样，但它不能直接在服务提供者处进行多点间的交换。VPLS 在传统 MPLS L2VPN 方案的基础上发展而成，它可以实现多点到多点的 VPN 组网。VPLS 为运营商提供了一种更加完备的解决方案

VPLS，即 Virtual Private LAN Service（虚拟专用 LAN 业务），是一种在 MPLS 网络上提供类似 LAN 的一种业务，实际上是一种基于以太网的 L2VPN 技术。VPLS 集以太网和 MPLS 的优点于一体，他使分散在不同地理位置上的用户网络可以互相通信，就像它们直接相互连接在一起一样，VPLS 能使用户延伸他们的 LAN 到 MAN，甚至 WAN 上。

VPWS 是 VPLS 的一个特例，而 VPLS 是 VPWS 的扩展。VPLS 提供一个多点到多点的 VPN 组网，从这个角度来看，VPLS 网络群对用户 CE 而言类似一个交换机。而 VPWS 则相当于一根网线，它仅能提供基于点到点的 L2VPN 解决方案。

VPLS 可选择使用 LDP 信令和 BGP 信令来构建 PW，基于 LDP 协议的信令通过在每一对 PE 之间建立点到点的 LDP 会话来建立虚电路，由于 LDP 协议提出的较早，应用简单，目前有许多厂商的产品支持这一机制，所以从兼容性的角度考虑我们实现的是基于 LDP 的信令机制。

5.1.1 VPLS基本概念

- ✎ **VPLS（Virtual Private LAN Service）**
虚拟专用 LAN 服务。是一种在 MPLS 网络上提供类似 LAN 的业务，它可以使使用户从多个地理位置分散的点同时接入网络，相互访问，就像这些点直接接入到 LAN 上一样。VPLS 使用户延伸他们的 LAN 到 MAN，甚至 WAN 上。
- ✎ **VC（Virtual Circuit）**
虚电路。在两个节点之间的单向逻辑连接。一个 PW 有一对反向 VC 组成。
- ✎ **S-TAG(Service Tag)、S-VID(Service VLAN ID)**
服务提供商网络为了区分用户而压入的“服务界定符”。
- ✎ **P-TAG(Provider Tag)、P-VID(Provider VLAN ID)**
对端 PE 期望的 VLAN Tag。
- ✎ **VFI（Virtual Forward Instance）**
每个 VFI 提供单独的 VPLS 服务，VFI 实现以太网桥接功能，通过 VFI，可以将 VPLS 的实际接入链路映射到各条 PW 上，反过来并能够终结 PW，将 PW 映射到实际接入链路上。
- ✎ **UPE（User facing-Provider Edge）**

面向用户侧的 PE 设备，主要作为用户接入 VPN 的汇聚设备。

☞ **NPE (Network Provider Edge)**

网络核心 PE 设备，处于 VPLS 网络的核心域边缘，提供在核心网之间的 VPLS 透明传输服务。

☞ **MTU (Maximum Transmission Unit)**

最大传输单元。

☞ **QinQ (802.1Q in 802.1Q)**

一种基于 802.1Q 封装的隧道协议，能够提供点到多点的 L2VPN 服务。它将用户私网 VLAN Tag 封装到公网 VLAN Tag 中，最终让报文携带两层 Tag 穿越服务提供商的骨干网络，从而为用户提供一种较为简单的二层 VPN 隧道。

☞ **PW Signaling**

PW 信令，用于创建和维护 PW，是 VPLS 的实现基础。PW 信令协议还可以用于自动发现 VFI 的对端 PE 设备。目前 PW 信令协议主要有 LDP 和 BGP。本实现仅支持 LDP。

☞ **VPWS (Virtual Private Wire Service)**

虚拟专用线服务，是一种点对点的 L2VPN 服务，对于一端（AC 或 PW）收到的报文直接转发到另一端（PW 或 AC）。VPWS 相对于 VPLS 来说不需要进行 MAC 地址的学习、查找，也不存在广播、多播等情况，所以效率更高。

5.1.2 VPLS基本网络模型

VPLS 模型主要有五部分组成：CE、PE、P、AC 和 PW。其基本概念如下。

☞ **CE (Custom Edge)**

直接和服务提供商相连的用户边缘设备。CE 可以是路由器或交换机，也可以是一台主机。

☞ **PE (Provider Edge)**

服务提供商网络上的边缘设备，与 CE 相连，主要负责 VPN 业务的接入。它完成报文从私网到公网隧道，并从公网隧道到私网的映射和转发。

☞ **P (Provider)**

服务提供商网络中的骨干路由器，不与 CE 直接相连。P 设备只需要具备基本 MPLS 转发能力，不维护 VPN 信息。

☞ **AC (Attachment Circuit)**

接入电路。在 L2VPN 中，CE 通过 AC 接入到 PE。AC 可以是物理链路也可以是逻辑链路，AC 用于在 CE 和 PE 之间传输帧。

☞ **PW (Pseudo Wire)**

虚链路，简单的说就是 VC 加隧道，隧道可以是 LSP、GRE、或者是 CR-LSP。PW 对于 VPLS 来说，就像是一条本地 AC 到对端 AC 之间的一条直连通道，完成用户的二层数据透传。

VPLS 相关的 rfc4762 中提供了两种 VPLS 网络模型，分别是：PE 全连接的 VPLS 模型和分层 VPLS 模型（H-VPLS）。如图 5-1、5-2 所示。

5.1.3 VPLS的优点

VPLS 具有以下优点：

- ☞ VPLS 在面向用户网一侧使用以太网接口，简化了 LAN/WAN 边界，可以支持快速

和灵活的服务部署

- ☞ VPLS 将用户网络的路由策略控制和维护的权利交给了用户，简化了运营商网络的管理
- ☞ VPLS 服务内的所有用户 CE 是相同子网的一部分，简化了 IP 地址的规划
- ☞ VPLS 服务不需要参与 IP 寻址和路由

5.1.4 PE全连接的基本VPLS网络模型

接入 VPLS 的各站点的 PE 设备之间逻辑全连接，PE 设备能在多点之间进行 MAC 地址学习以及数据包的转发。MPLS 网络提供隧道来透传 VPN 站点间的报文，网络中的 P 设备作用类似于 L3VPN 中的 P 设备，它不参与 MAC 地址的学习和交换，只对 MPLS 报文进行转发。PE 上各 VPN 之间的转发表相互独立，使得各 VPN 间的 MAC 地址可以重叠。

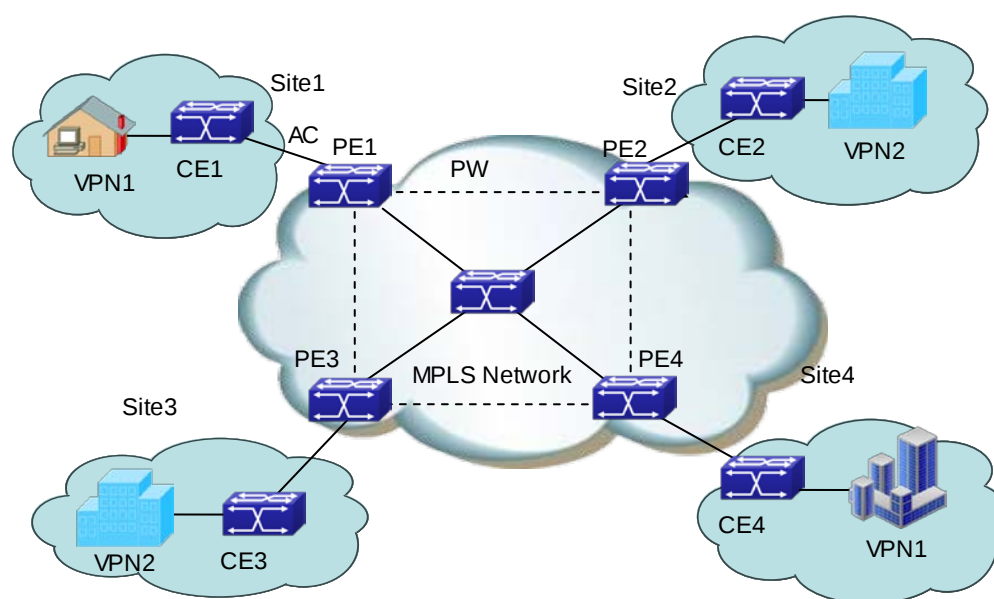


图 5-1 PE 全连接 VPLS 模型

☞ PE 全连接基本 VPLS 模型环路避免方法

VPLS 全连接模型是通过水平分割转发来避免环路，每个 PE 设备如果从 PW 上收到报文，那么这个报文将不再向这个 VFI 关联的其它 PW 上转发，也就是说要求任意两个 PE 之间通过直接相连的 PW 通信，而不能通过第三个 PE 设备中转报文，这也就是 VPLS 的 VFI 之间要求建立全连接 PW 的原因。

5.1.5 分层VPLS模型

所有的 NPE 设备之间逻辑全连接，UPE 只与最近的 NPE 建立虚连接，通过 NPE 与对端 VPN 站点进行报文交换，这样层次化了网络拓扑。UPE 主要用于 VPN 业务的接入，对性能要求比较低，NPE 主要用于流量的汇聚，对性能要求比较高。另外，还可以在 UPE 和 NPE 之间增加链路备份，保证网络的健壮性。UPE 和 NPE 之间的虚连接可以使用 QinQ 或者 LDP 来建立。

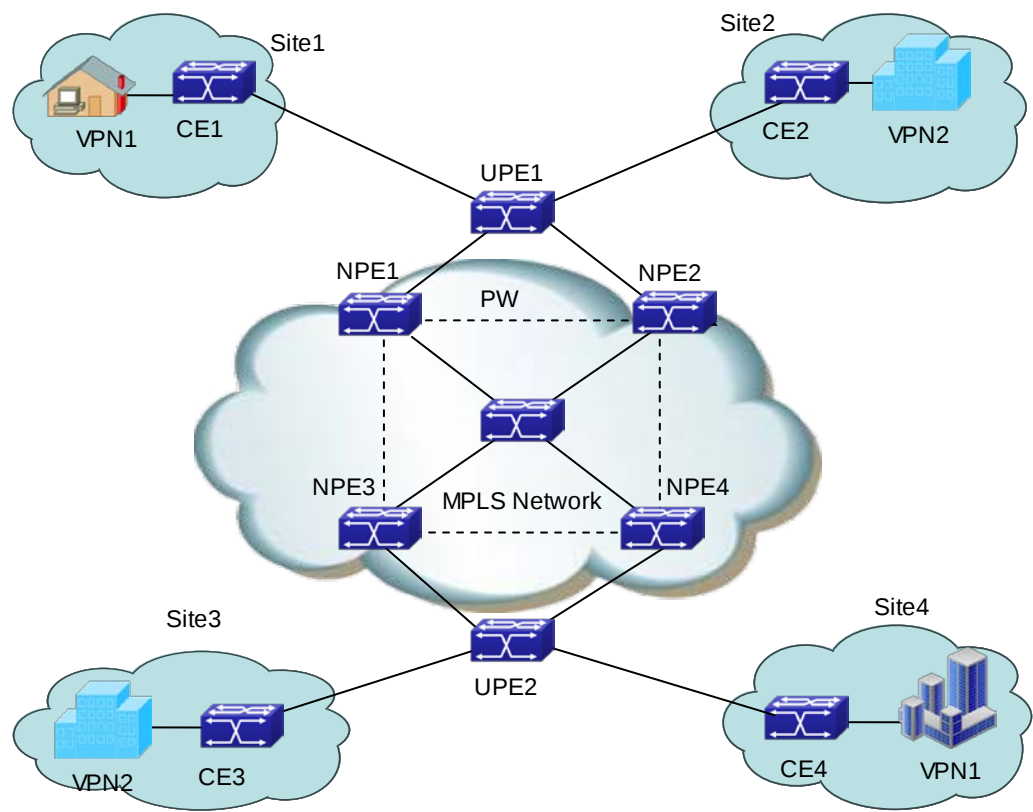


图 5-2 分层 VPLS 模型

根据 UPE 和 NPE 之间连接方式的不同，H-VPLS 分为：LSP 接入方式和 QinQ 接入方式

☞ LSP 接入方式

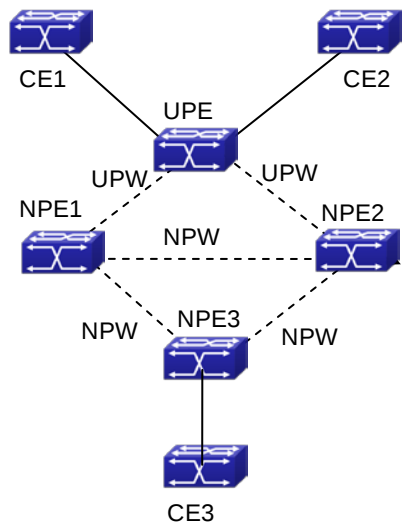


图 5-3 LSP 接入方式

如图 5-3 所示，UPE 作为汇聚设备，它与 NPE1 和 NPE2 之间通过 LSP 接入，需要在 UPE 和 NPE1、NPE2 之间分别建立虚连接 U-PW（建立 U-PW 需要在 UPE 和 NPE 设备上分别创建 VFI，指定对等体，并要求两台设备的 PWID 必须相同），UPE 跟

其他所有的对端都不建立虚连接。

LSP 接入方式的数据转发流程如下：

- (1) UPE 负责将 CE 上送的报文发给 NPE1，同时打上 U-PW 对应的 VC 标签（NPE1 分配的 VC 标签，作为多路 PW 复用分离标记）；
- (2) NPE1 收到报文后，先根据 VC 标签判断报文所属的 VFI，再根据该报文的目的地 MAC 压入 N-PW 对应的 VC 标签，然后转发该报文；
- (3) NPE1 从 N-PW 侧收到报文后，打上 U-PW 对应的 VC 标签，将报文发送给 UPE，UPE 再将报文转发给 CE。

CE1 与 CE2 之间的数据交换为本地 CE 之间交换时，如果 UPE 本身具有桥接功能，UPE 将直接完成两者间的报文转发，而无需将报文上送给 NPE1。不过对于目的地 MAC 未知的第一个数据报文或广播报文，UPE 在将数据通过桥广播到 CE2 的同时，仍然会通过 U-PW 转发给 NPE1，由 NPE1 来完成报文的复制并转发到各个对端 CE。

QinQ 接入方式

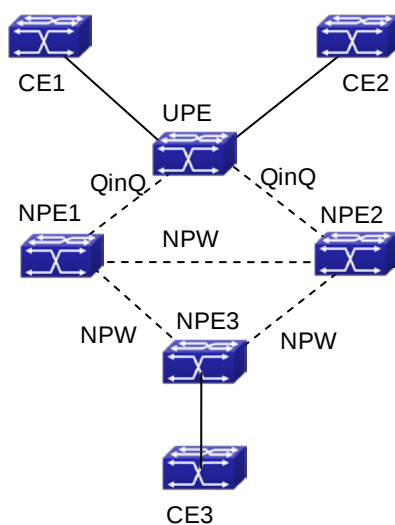


图 5-4 QinQ 接入方式

如图 5-4 所示，UPE 为标准的桥接设备，在 UPE 和 PE1 和 PE2 之间建立点到点的以太网 QinQ 连接（即在 UPE 面向 CE 的接口上使能 QinQ，在与 UPE 直连的 PE1 上使用 VLAN 接入方式）。UPE 从 CE 上收到的报文都将被打上一层外层 VLAN 标记，报文转发到 PE1 时，根据配置的 VLAN 接入方式，外层 VLAN 标记被解释为服务提供商 VLAN 标记，也就是服务提供商为用户分配的服务界定符，根据这个界定符，报文被映射到相应的 VFI，由这个 VFI 决策出报文如何转发（单播或广播）。

QinQ 接入方式的数据转发流程如下：

- (1) 在 CE 接入端口使能 QinQ，为收到的报文添加压入 VLAN TAG 作为多路复用分离标记，在 UPE 与 PE1 之间通过 QinQ 隧道将报文透明传输到 PE1；
- (2) PE1 先根据报文携带的 VLAN TAG 判断所属的 VFI，再根据该报文的目的地 MAC 为其压入 PW 对应的 PW 标签，然后将其转发；
- (3) PE1 从 PW 侧收到报文后，根据 PW 标签，判断报文所属的 VFI，再根据用户报文的目的地 MAC 打上 VLAN TAG 通过 QinQ 隧道将报文转发给 UPE，由 UPE 将报文转发给 CE。

如果 CE1 与 CE2 之间的数据交换为本地 CE 之间交换，由于 UPE 本身具有桥接功能，UPE 将直接完成两者间的报文转发，而无需将报文上送给 PE1。不过对于目的地

MAC 未知的第一个数据报文或广播报文，UPE 在通过桥广播到 CE2 的同时，仍然会通过 QinQ 隧道转发给 PE1，由 PE1 来完成报文的复制并转发到各个对端 CE。

☞ H-VPLS 模型环路避免方法

和全连接相比，分层 VPLS 模型环路避免方法略有不同，由于 H-VPLS 只需在 NPE 之间建立全连接，UPE 和 NPE 之间不需要全连接，因此每个 NPE 设备上，从与 NPE 连接的 PW 上收到的报文，不再向这个 VFI 关联的，与其它 NPE 连接的 PW 转发，但可以向与 UPE 连接的 PW 转发。每个 NPE 设备上，从与 UPE 连接的 PW 上收到的报文，可以向这个 VFI 关联的所有与其它 NPE 连接的 PW 转发。

5.1.6 VPLS 的报文转发

在 VPLS 模型中的 AC 和 PW 上的报文封装方式分别有两种，其中 AC 上的是 VLAN 接入和 Ethernet 接入两种。含义分别如下：

☞ VLAN 接入

CE 发送给 PE 或 PE 发送给 CE 的以太网帧头带有一个 VLAN TAG，该 TAG 是一个服务提供商网络为了区分用户而压入的“服务界定符”。服务界定符一般是服务提供商设备添加的，我们把这个作为服务界定符的 TAG 称为 S-TAG。

☞ Ethernet 接入

CE 发送给 PE 或 PE 发送给 CE 的以太网帧头中没有服务界定符，如果此时帧头中有 VLAN TAG，则说明它只是用户报文的内部 VLAN TAG，对于 PE 设备没有意义。这种用户内部 VLAN 的 TAG 称为 U-TAG。

PW 上的报文封装方式分别是 Raw 模式和 Tagged 模式。含义分别如下：

☞ Raw 模式

PW 上传输的帧不能带 S-TAG：对于 CE 侧的报文，如果收到带有服务界定符的报文，则将其去除后再压入 PW 标签和隧道标签后转发；如果收到不带服务界定符的报文，则直接压入 PW 标签和隧道标签后转发。对于 PE 侧的下行报文，根据实际配置选择添加或不添加服务界定符后转发给 CE，但是它不允许重写或删除已经存在的任何 TAG。

☞ Tagged 模式

PW 上传输的帧必须带 S-TAG：对于 CE 侧的报文，如果收到带有服务界定符的报文，则通过与对端 PE 发送过来的期望的 VLAN TAG 进行比较，如果相同，则保留 S-TAG，否则将 S-TAG 改写为对端 PE 期望的 VLAN TAG，如果对端 PE 没有传递期望的 VLAN TAG，则填写空 TAG（TAG 值为 0），然后再压入 PW 标签和隧道标签后转发；如果收到不带服务界定符的报文，则添加一个对端 PE 发送过来的期望的 VLAN TAG 或空 TAG 后，再压入 PW 标签和隧道标签后转发。对于 PE 侧的下行报文，根据实际配置选择重写、去除或保留服务界定符后转发给 CE。

VPLS 报文的转发方式根据 AC 接入方式和 PW 上报文封装方式的不同，可以分为以下四种：

☞ Ethernet 接入 Raw 模式的报文转发

以图 5-1 的网络模型为例，AC 采用 Ethernet 接入方式，PW 上的报文封装模式为 Raw 模式，从 CE1 到 CE4 的报文转发过程如下：

1. CE1 将携带用户所属 VLAN 信息（U-Tag）的报文发送给 PE1；
2. PE1 根据用户目的 MAC 地址，选择合适的 PW，并在报文中添加 PW 对应的 VC 标签；
3. 为了在公网上利用 MPLS 隧道转发报文，PE1 在报文中添加公网隧道标

签，通过公网隧道将报文传递给 PE4；

4. PE4 收到报文后，根据 VC 标签找到报文所属的 VFI，并将携带 U-TAG 的报文发送给 CE4。

✎ Ethernet 接入 Tagged 模式的报文转发

以图 5-1 的网络模型为例，AC 采用 Ethernet 接入方式，PW 上的报文封装模式为 Tagged 模式时，报文的转发过程与 Ethernet 接入方式、Raw 模式的报文转发过程类似。所不同的是，PW 上传输的帧必须带 S-TAG。PE1 从 CE1 收到不带 S-TAG 的报文后，首先添加一个对端 PE 发送过来的期望的 VLAN TAG 或空 TAG，再压入两层 MPLS 标签转发。PE4 接收到报文后，去除 MPLS 两层标签和 S-TAG，再将报文转发给 CE4。

✎ VLAN 接入 Raw 模式的报文转发

以图 5-1 的网络模型为例，AC 采用 VLAN 接入方式，PW 上的报文封装模式为 Raw 模式，从 CE1 到 CE4 的报文转发过程如下：

1. CE1 发送的报文应携带服务界定符 S-TAG，否则，如果 CE1 发送的报文不带 VLAN TAG 或者所携带的 VLAN TAG 和 S-TAG 不匹配，则走普通的二层报文转发流程。

2. PE1 接收到报文后去除 S-TAG，并添加两层 MPLS 标签，通过公网 MPLS 隧道，将报文发送给 PE4。

3. PE4 去除接收到报文中的两层 MPLS 标签，增加 S-TAG 后，将报文转发给 CE4。

✎ VLAN 接入 Tagged 模式的报文转发

以图 5-1 的网络模型为例，AC 采用 VLAN 接入方式，PW 上的报文封装模式为 Tagged 模式时，报文的转发过程与 VLAN 接入方式 Raw 模式的报文转发过程类似。二者的区别是，PW 上传输的帧携带 S-TAG，PE 接收到报文后，根据和对端 PE 发送过来的所期望的 VLAN TAG 相比较，如果相同，则保留 S-TAG，否则将 S-TAG 改写为对端 PE 期望的 VLAN TAG，如果对端 PE 没有传递期望的 VLAN TAG，则填写空 TAG（TAG 值为 0）。

5.2 VPLS配置

VPLS 配置任务序列：

1. 全局启动路由协议（必选）
2. MPLS 基本配置（必选）
 - (1) 全局使能 MPLS
 - (2) 在接口使能标签交换能力
3. 配置 LDP 会话（必选）
 - (1) 配置模式下进入 ldp 视图，同时也代表开启了 LDP 功能
 - (2) 配置远端 LDP 邻居，使用 targeted-peer 命令
 - (3) 在相应的三层接口视图下使能 LDP 命令
4. 创建 PW 模板（可选）
 - (1) 创建 pw-class 并进入 pw 视图
 - (2) 设定 pw-class 的传输模式
5. 配置 VFI（必选）
 - (1) 创建 VFI，指定 VFI ID，并进入 VFI 视图

- (2) 配置远端 PW 的终结点（一般情况与 LDP 的 targeted peer 相同）
- (3) 可以为每个 peer 指定 pw-class
- (4) 可以设置 VFI 的 PW 封装模式
- 6. 配置用户接入方式并绑定 VFI（VPLS 必选）
 - (1) 进入端口视图
 - (2) 配置端口绑定 VFI
- 7. 配置端口绑定 PW（VPWS 必选）
 - (1) 创建 L2VC，配置 VPWS 的对端 PE
 - (2) 进入端口视图
 - (3) 将端口与指定 PW 进行绑定，并指定接入模式

1. 全局启动路由协议（必选）

命令	解释
全局配置模式	
router ospf	进入 OSPF 路由配置模式。
OSPF 路由配置模式	
network 0.0.0.0/0 area 0	配置使能 OSPF 的接口网段地址，缺省情况下接口上 OSPF 未使能。

2. 全局 MPLS（必选）

命令	解释
全局配置模式	
mpls enable no mpls enable	启动 MPLS 协议；本命令的 no 操作关闭 MPLS 协议。
三层接口视图	
label-switching enable no label-switching	使能标签交换功能；缺省情况下未使能标签交换。

3. 配置 LDP 会话（必选）

- (1) 配置模式下进入 ldp 视图，同时也代表开启了 LDP 功能
- (2) 配置远端 LDP 邻居，使用 targeted-peer 命令
- (3) 在相应的三层接口视图下使能 LDP 命令

命令	解释
全局配置模式	
router ldp no router ldp	打开 LDP 协议；本命令的 no 操作为关闭 LDP 协议，缺省情况下未使能 LDP。
targeted-peer 1.1.1.1	配置远端 LDP 邻居；缺省没有远端邻居
接口配置模式	
ldp {enable disable}	接口上启动 LDP 协议；本命令的 no 操作为关闭接口上的 LDP 协议，接口缺省情况下未使能 LDP。

4. 创建 PW 模板（可选）

- (1) 创建 pw-class 并进入 pw 视图
- (2) 设定 pw-class 的传输模式

命令	解释
全局配置模式	
pw-class <pw-class-name> no pw-class <pw-class-name>	创建 pw-class；缺省情况下不存在任何 PW 模板。
pw 模板配置模式	
transport-mode {ethernet vlan}	配置 PW 模板的 PW 上的报文封装模式，ethernet 对应 raw 模式，vlan 对应 tagged 模式。

5. 配置 VFI（必选）

- (1) 创建 VFI，指定 VFI ID，并进入 VFI 视图
- (2) 配置远端 PW 的终结点（一般情况与 LDP 的 targeted peer 相同）
- (3) 可以为每个 peer 指定 pw-class
- (4) 可以设置 VFI 的 PW 封装模式

命令	解释
全局配置模式	
vfi <vfi-name> <vfi-id> no vfi <vfi-name>	创建 VFI 并指定 VFI ID；缺省情况下不存在任何 VFI。
VFI 配置模式	
peer ip-address [pw-id pw-id] [no-split-horizon] [pw-class class-name]	配置 VPLS 实例包含的对端 PE 可以配置是否使能水平分割(默认使能)和 pw 模板。
transport-mode {ethernet vlan}	配置 PW 模板的 PW 上的报文封装模式。

6. 配置用户接入方式并绑定 VFI（VPLS 必选）

- (1) 进入端口视图
- (2) 配置端口绑定 VFI

命令	解释
端口配置模式	
xconnect vfi vfi-id [mode {ethernet vlan [svid <svid>]]]	配置用来将一个端口绑定到一个 VFI，并配置 AC 的接入模式。缺省情况下端口不绑定任何 VFI。

7. 配置端口绑定 PW（VPWS 必选）

- (1) 创建 L2VC，配置 VPWS 的对端 PE
- (2) 进入端口视图
- (3) 将端口与指定 PW 进行绑定，并指定接入模式

命令	解释
全局配置模式	
l2-vc ip-address pw-id pw-id [group group-id] [pw-class class-name]	创建 L2VC，配置 VPWS 的对端 PE。缺省情况下没有配置对端 PE。

进入端口视图	
xconnect l2-vc pw-id <pw-id> [mode {ethernet vlan [svid <svid>]]]	将端口与指定 PW 进行绑定，并指定接入模式，缺省情况下端口不绑定到任何 PW。

5.3 VPLS典型案例

5.3.1 全连接的VPLS配置

5.3.1.1组网需求

- CE1 和 CE2 分属于两个站点，同属于一个 VPN1
- CE1 和 CE2 分别通过端口 Eth1/1 接入到 PE 设备
- CE1 到 PE1 的接入链路配置为 Ethernet 接入，CE2 和 PE2 之间的接入链路配置为 VLAN 接入 Svid 为 200
- PE1 和 PE2 之间 PW 上的报文封装模式配置为 RAW(Ethernet)模式，PE1 和 PE3，PE2 和 PE3 之间的报文封装模式配置为 Tagged(Vlan)模式
- 要求通过配置 VPLS 的方式建立一个二层 VPN1，使得 CE1 和 CE2 之间可以进行二层互访

5.3.1.2组网图

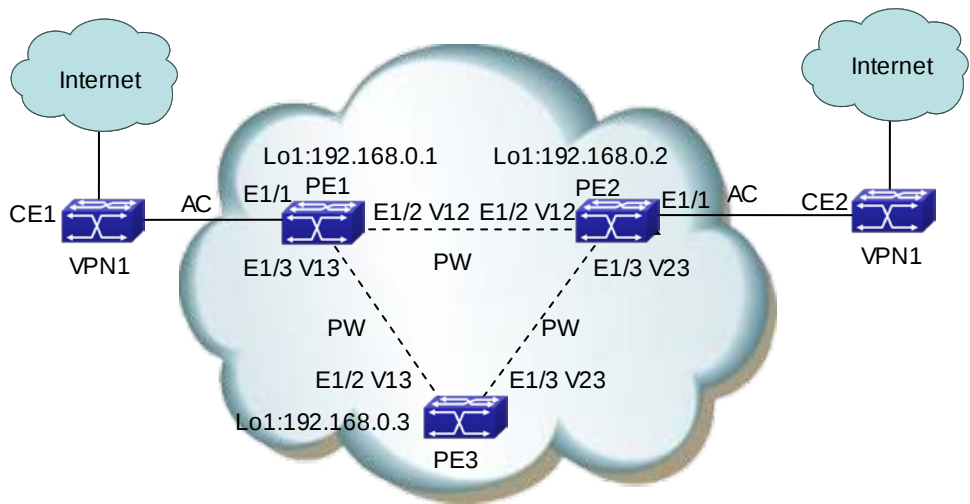


图 5-5 PE 全连接 VPLS 模型

5.3.1.3配置步骤

请按照图 5-5 配置各接口 IP 地址和掩码，包括 VLAN 接口和 Loopback 接口，具体配

置过程略。本配置只给出 LDP 方式下 VPLS 实例相关的配置。

(1) PE1 的配置

#创建 PW 模板

```
PE1(config)#pw-class c1
PE1(config-class) #transport-mode ethernet
PE1(config-class)#exit
PE1(config)#pw-class c2
PE1(config-class)#transport-mode vlan
PE1(config-class)#exit
```

#配置路由

```
PE1(config)#router ospf
PE1(config-router)#router-id 192.168.0.1
PE1(config-router)#network 0.0.0.0/0 area 0
PE1(config-router)#exit
```

#配置 MPLS 能力和标签交换能力(非 H-VPLS 情况下与 CE 相连的接口不用使能标签交换能力)

```
PE1(config)#mpls enable
PE1(config)#int vlan 12
PE1(config-if-vlan12)#label-switching
PE1(config-if-vlan12)#exit
PE1(config)#int vlan 13
PE1(config-if-vlan13)#label-switching
PE1(config-if-vlan13)#exit
```

#配置 LDP

```
PE1(config)#router ldp
PE1(config-router)#router-id 192.168.0.1
PE1(config-router)#targeted-peer 192.168.0.2
PE1(config-router)#targeted-peer 192.168.0.3
PE1(config-router)#exit
PE1(config)#int vlan 12
PE1(config-if-vlan12)#ldp enable
PE1(config-if-vlan12)#exit
PE1(config)#int vlan 13
PE1(config-if-vlan13)#ldp enable
PE1(config-if-vlan13)#exit
```

#配置 LDP 方式下的 VPLS 实例 V1 以及实例中的远端 PE

```
PE1(config)#vfi v1 100
```

```
PE1(config-vfi)#peer 192.168.0.2 pw-class c1
```

```
PE1(config-vfi)#peer 192.168.0.3 pw-class c2
```

```
PE1(config-vfi)#exit
```

#在面向 CE1 的端口上绑定 vpls 实例模式为 ethernet

```
PE1(config-if-ethernet1/1)#xconnect vfi 100 mode ethernet
```

(2) PE2 的配置

#创建 PW 模板

```
PE2(config)#pw-class c1
```

```
PE2(config-class)#transport-mode ethernet
```

```
PE2(config-class)#exit
```

```
PE2(config)#pw-class c2
```

```
PE2(config-class)#transport-mode vlan
```

```
PE2(config-class)#exit
```

#配置路由

```
PE2(config)#router ospf
```

```
PE2(config-router)#router-id 192.168.0.2
```

```
PE2(config-router)#network 0.0.0.0/0 area 0
```

```
PE2(config-router)#exit
```

#配置 MPLS 能力和标签交换能力(非 H-VPLS 情况下与 CE 相连的接口不用使能标签交换能力)

```
PE2(config)#mpls enable
```

```
PE2(config)#int vlan 12
```

```
PE2(config-if-vlan12)#label-switching
```

```
PE2(config-if-vlan12)#exit
```

```
PE2(config)#int vlan 23
```

```
PE2(config-if-vlan23)#label-switching
```

```
PE2(config-if-vlan23)#exit
```

#配置 LDP

```
PE2(config)#router ldp
```

```
PE2(config-router)#router-id 192.168.0.2
```

```
PE2(config-router)#targeted-peer 192.168.0.1
PE2(config-router)#targeted-peer 192.168.0.3
PE2(config-router)#exit
PE2(config)#int vlan 12
PE2(config-if-vlan12)#ldp enable
PE2(config-if-vlan12)#exit
PE2(config)#int vlan 23
PE2(config-if-vlan23)#ldp enable
PE2(config-if-vlan23)#exit
```

#配置 LDP 方式下的 VPLS 实例 V1 以及实例中的远端 PE

```
PE2(config)#vfi v1 100
PE2(config-vfi)#peer 192.168.0.1 pw-class c1
PE2(config-vfi)#peer 192.168.0.3 pw-class c2
PE2(config-vfi)#exit
```

#在面向 CE2 的端口上绑定 VPLS 实例模式为 VLAN，Svid 为 200

```
PE2(config-if-ethernet1/1)#xconnect vfi 100 mode vlan svid 200
```

(3) PE3 的配置

同 PE1 和 PE2 类似。

5.3.2 LSP方式接入的H-VPLS

5.3.2.1组网需求

- ☞ CE1 和 CE2 分属于两个站点，同属于一个 VPN1
- ☞ CE1 和 CE2 分别通过端口 Eth1/1 接入到 PE 设备
- ☞ N-PE1、N-PE2、N-PE3 之间构建网络侧全连接的 VPLS 网络
- ☞ U-PE 作为用户侧的接入设备，以分层的 VPLS PW 方式接入 N-PE1
- ☞ CE1 到 U-PE 的接入链路配置为 Ethernet 接入，CE2 和 N-PE3 之间的接入链路配置为 VLAN 接入 Svid 为 200
- ☞ N-PE 之间 PW 上的报文封装模式配置为 RAW(Ethernet)模式
- ☞ U-PE 和 N-PE1 之间为 PW 连接
- ☞ 要求通过配置 VPLS 的方式建立一个二层 VPN1，使得 CE1 和 CE2 之间可以进行二层互访

5.3.2.2 组网图

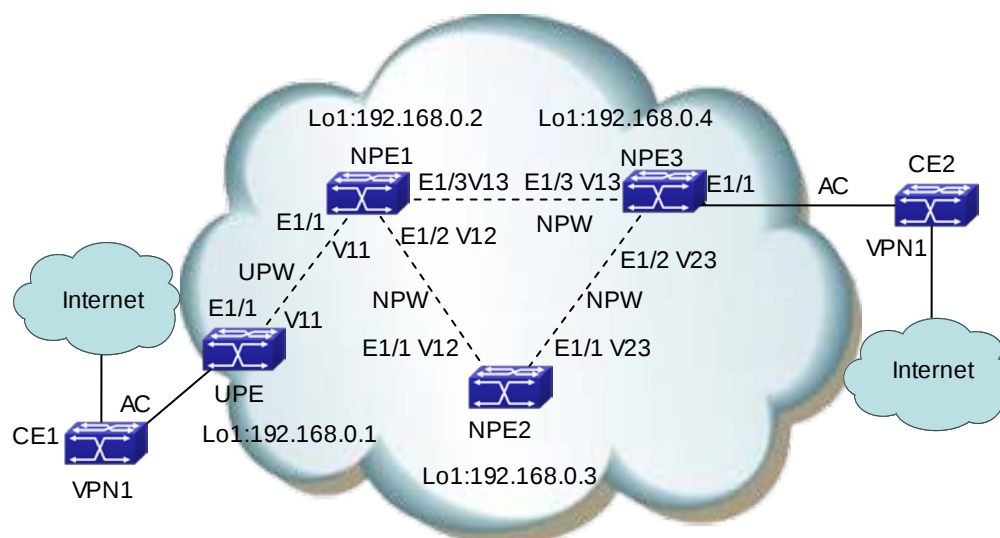


图 5-6 LSP 接入的 H-VPLS 应用组网

5.3.2.3 配置步骤

请按照图 5-6 配置各接口 IP 地址和掩码，包括 VLAN 接口和 Loopback 接口，具体配置过程略。本配置只给出 LDP 方式下实现 H-VPLS 相关的配置。

(1) UPE 的配置

#创建 PW 模板

```
UPE(config)#pw-class c1
UPE(config-class)#transport-mode ethernet
UPE(config-class)#exit
```

#配置路由

```
UPE(config)#router ospf
UPE(config-router)#router-id 192.168.0.1
UPE(config-router)#network 0.0.0.0/0 area 0
UPE(config-router)#exit
```

#配置 MPLS 能力和标签交换能力(与 CE 相连的接口不用使能标签交换能力)

```
UPE(config)#mpls enable
UPE(config)#int vlan 11
UPE(config-if-vlan11)#label-switching
UPE(config-if-vlan11)#exit
```

#配置 LDP

```
UPE(config)#router ldp
UPE(config-router)#router-id 192.168.0.1
UPE(config-router)#targeted-peer 192.168.0.2
UPE(config-router)#exit
UPE(config)#int vlan 11
UPE(config-if-vlan11)#ldp enable
UPE(config-if-vlan11)#exit
```

#配置 LDP 方式下的 VPLS 实例 V1 以及实例中的远端 PE

```
UPE(config)#vfi v1 100
UPE(config-vfi)#peer 192.168.0.2 pw-class c1
UPE(config-vfi)#exit
```

#在面向 CE1 的端口上绑定 vpls 实例模式为 ethernet

```
UPE(config-if-ethernet1/1)#xconnect vfi 100 mode ethernet
```

(2) NPE1 的配置

#创建 PW 模板

```
NPE1(config)#pw-class c1
NPE1(config-class)#transport-mode ethernet
NPE1(config-class)#exit
```

#配置路由

```
NPE1(config)#router ospf
NPE1(config-router)#router-id 192.168.0.2
NPE1(config-router)#network 0.0.0.0/0 area 0
NPE1(config-router)#exit
```

#配置 MPLS 能力和标签交换能力

```
NPE1(config)#mpls enable
NPE1(config)#int vlan 11
NPE1(config-if-vlan11)#label-switching
NPE1(config-if-vlan11)#exit
NPE1(config)#int vlan 12
NPE1(config-if-vlan12)#label-switching
NPE1(config-if-vlan12)#exit
NPE1(config)#int vlan 13
NPE1(config-if-vlan13)#label-switching
NPE1(config-if-vlan13)#exit
```

#配置 LDP

```
NPE1(config)#router ldp
NPE1(config-router)#router-id 192.168.0.2
NPE1(config-router)#targeted-peer 192.168.0.1
NPE1(config-router)#targeted-peer 192.168.0.3
NPE1(config-router)#targeted-peer 192.168.0.4
NPE1(config-router)#exit
NPE1(config)#int vlan 11
NPE1(config-if-vlan11)#ldp enable
NPE1(config-if-vlan11)#exit
NPE1(config)#int vlan 12
NPE1(config-if-vlan12)#ldp enable
NPE1(config-if-vlan12)#exit
NPE1(config)#int vlan 13
NPE1(config-if-vlan13)#ldp enable
NPE1(config-if-vlan13)#exit
```

配置 LDP 方式下的 VPLS 实例 V1 以及实例中的远端 PE(UPE 的 peer 需要关闭水平分割功能)

```
NPE1(config)#vfi v1 100
NPE1(config-vfi)#peer 192.168.0.3 pw-class c1
NPE1(config-vfi)#peer 192.168.0.4 pw-class c1
NPE1(config-vfi)#peer 192.168.0.1 no-split-horizon pw-class c1
NPE1(config-vfi)#exit
```

(3) NPE2 的配置

与 NPE1 类似，区别在于没有与 UPE 的 peer，不用关闭水平分割

(4) NPE3 的配置

#创建 PW 模板

```
NPE3(config)#pw-class c1
NPE3(config-class)#transport-mode ethernet
NPE3(config-class)#exit
```

#配置路由

```
NPE3(config)#router ospf
NPE3(config-router)#router-id 192.168.0.4
NPE3(config-router)#network 0.0.0.0/0 area 0
```

```
NPE3(config-router)#exit
```

```
#配置 MPLS 能力和标签交换能力
```

```
NPE3(config)#mpls enable
```

```
NPE3(config)#int vlan 13
```

```
NPE3(config-if-vlan11)#label-switching
```

```
NPE3(config-if-vlan11)#exit
```

```
NPE3(config)#int vlan 23
```

```
NPE3(config-if-vlan12)#label-switching
```

```
NPE3(config-if-vlan12)#exit
```

```
#配置 LDP
```

```
NPE3(config)#router ldp
```

```
NPE3(config-router)#router-id 192.168.0.4
```

```
NPE3(config-router)#targeted-peer 192.168.0.2
```

```
NPE3(config-router)#targeted-peer 192.168.0.3
```

```
NPE3(config-router)#exit
```

```
NPE3(config)#int vlan 13
```

```
NPE3(config-if-vlan11)#ldp enable
```

```
NPE3(config-if-vlan11)#exit
```

```
NPE3(config)#int vlan 23
```

```
NPE3(config-if-vlan12)#ldp enable
```

```
NPE3(config-if-vlan12)#exit
```

```
#配置 LDP 方式下的 VPLS 实例 V1 以及实例中的远端 PE
```

```
NPE3(config)#vfi v1 100
```

```
NPE3(config-vfi)#peer 192.168.0.2 pw-class c1
```

```
NPE3(config-vfi)#peer 192.168.0.3 pw-class c1
```

```
NPE3(config-vfi)#exit
```

```
#在面向 CE2 的端口上绑定 vpls 实例
```

```
N-PE3(config-if-ethernet1/1)#xconnect vfi 100 mode vlan svid 200
```

5.3.3 QinQ方式接入的H-VPLS

5.3.3.1组网需求

☞ CE1 和 CE2 分属于两个站点，同属于一个 VPN1

- ☞ CE1 和 CE2 分别通过端口 Eth1/2 和 Eth1/1 接入到 PE 设备
- ☞ N-PE1、N-PE2、N-PE3 之间构建成网络侧全连接的 VPLS 网络
- ☞ U-PE 作为用户侧的接入设备，以分层的 QinQ 方式接入 N-PE1，Svid 为 100
- ☞ CE2 上采用正常的 VPLS 接入方式模式为 VLAN，Svid 为 200
- ☞ N-PE 之间 PW 上的报文封装模式配置为 Tagged 模式
- ☞ U-PE 和 N-PE1 之间为 QinQ 连接
- ☞ 要求通过配置 VPLS 的方式建立一个二层 VPN1，使得 CE1 和 CE2 之间可以进行二层互访

5.3.3.2 组网图

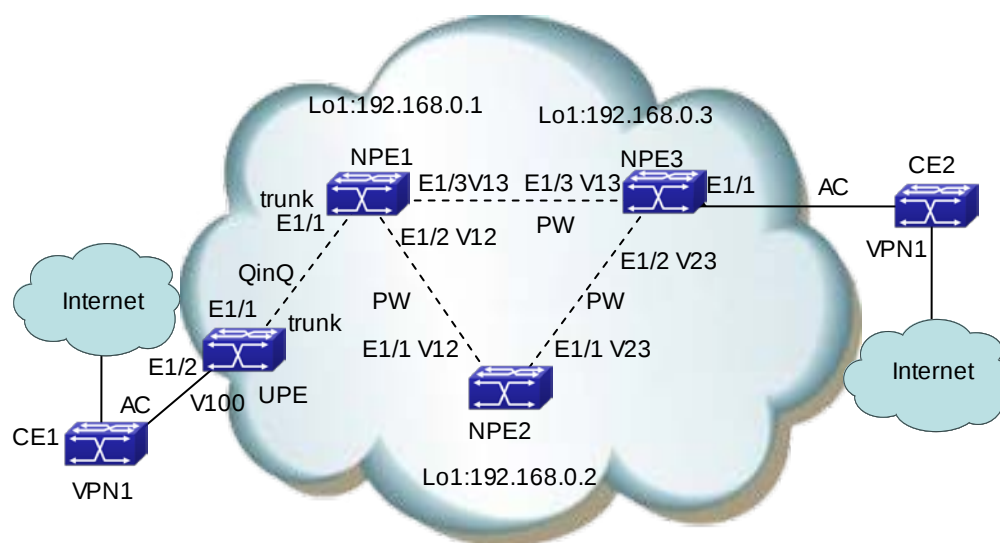


图 5-7 QinQ 接入的 H-VPLS 应用组网

5.3.3.3 配置步骤

请按照图 5-7 配置各接口 IP 地址和掩码，包括 VLAN 接口和 Loopback 接口，具体配置过程略。本配置只给出 QinQ 方式下实现 H-VPLS 相关的配置。

(1) UPE 的配置

#在 CE1 的接入端口上使能 QinQ，该端口所属的 VLAN 为 100

```
UPE(config-if-ethernet1/2)#switchport access vlan 100
```

```
UPE(config-if-ethernet1/2)#dot1q-tunnel enable
```

#把与 NPE1 连接的端口配置为 Trunk，默认情况下 trunk 端口允许所有 VLAN 通过

```
UPE(config-if-ethernet1/1)#switchport mode trunk
```

(2) NPE1 的配置

#创建 PW 模板

```
NPE1(config)#pw-class c1
NPE1(config-class)#transport-mode Vlan
NPE1(config-class)#exit

#配置路由
NPE1(config)#router ospf
NPE1(config-router)#router-id 192.168.0.1
NPE1(config-router)#network 0.0.0.0/0 area 0
NPE1(config-router)#exit
```

```
#配置 MPLS 能力和标签交换能力
NPE1(config)#mpls enable
NPE1(config)#int vlan 12
NPE1(config-if-vlan12)#label-switching
NPE1(config-if-vlan12)#exit
NPE1(config)#int vlan 13
NPE1(config-if-vlan13)#label-switching
NPE1(config-if-vlan13)#exit
```

```
#配置 LDP
NPE1(config)#router ldp
NPE1(config-router)#router-id 192.168.0.1
NPE1(config-router)#targeted-peer 192.168.0.2
NPE1(config-router)#targeted-peer 192.168.0.3
NPE1(config-router)#exit
NPE1(config)#int vlan 12
NPE1(config-if-vlan12)#ldp enable
NPE1(config-if-vlan12)#exit
NPE1(config)#int vlan 13
NPE1(config-if-vlan13)#ldp enable
NPE1(config-if-vlan13)#exit
```

配置 LDP 方式下的 VPLS 实例 V1 以及实例中的远端 PE

```
NPE1(config)#vfi v1 100
NPE1(config-vfi)#peer 192.168.0.2 pw-class c1
NPE1(config-vfi)#peer 192.168.0.3 pw-class c1
NPE1(config-vfi)#exit
```

配置与 UPE QinQ 方式接入的端口上绑定 vpls 实例模式为 VLAN, Svid 为 100

```
NPE1(config-if-ethernet1/2)#switchport mode trunk
```

```
NPE1 (config-if-ethernet1/2)#xconnect vfi 100 mode vlan svid 100
```

(3) NPE2 的配置

与 NPE1 类似

(4) NPE3 的配置

#创建 PW 模板

```
NPE3(config)#pw-class c1
```

```
NPE3(config-class)#transport-mode Vlan
```

```
NPE3(config-class)#exit
```

#配置路由

```
NPE3(config)#router ospf
```

```
NPE3(config-router)#router-id 192.168.0.3
```

```
NPE3(config-router)#network 0.0.0.0/0 area 0
```

```
NPE3(config-router)#exit
```

#配置 MPLS 能力和标签交换能力

```
NPE3(config)#mpls enable
```

```
NPE3(config)#int vlan 13
```

```
NPE3(config-if-vlan11)#label-switching
```

```
NPE3(config-if-vlan11)#exit
```

```
NPE3(config)#int vlan 23
```

```
NPE3(config-if-vlan12)#label-switching
```

```
NPE3(config-if-vlan12)#exit
```

#配置 LDP

```
NPE3(config)#router ldp
```

```
NPE3(config-router)#router-id 192.168.0.3
```

```
NPE3(config-router)#targeted-peer 192.168.0.1
```

```
NPE3(config-router)#targeted-peer 192.168.0.2
```

```
NPE3(config-router)#exit
```

```
NPE3(config)#int vlan 13
```

```
NPE3(config-if-vlan11)#ldp enable
```

```
NPE3(config-if-vlan11)#exit
```

```
NPE3(config)#int vlan 23
```

```
NPE3(config-if-vlan12)#ldp enable
```

```
NPE3(config-if-vlan12)#exit
```

#配置 LDP 方式下的 VPLS 实例 V1 以及实例中的远端 PE

```
NPE3(config)#vfi v1 100
```

```
NPE3(config-vfi)#peer 192.168.0.1 pw-class c1
```

```
NPE3(config-vfi)#peer 192.168.0.2 pw-class c1
```

```
NPE3(config-vfi)#exit
```

#在面向 CE2 的端口上绑定 VPLS 实例

```
N-PE3(config-if-ethernet1/1)#xconnect vfi 100 mode vlan svid 200
```

5.3.4 VPWS配置

5.3.4.1组网需求

- ☞ CE1 和 CE2 分属于两个站点，同属于一个 VPN1
- ☞ CE1 和 CE2 分别通过端口 Eth1/1 接入到 PE 设备
- ☞ CE1 到 PE1 的接入链路配置为 Ethernet 接入，CE2 和 PE2 之间的接入链路配置为 VLAN 接入 Svid 为 200
- ☞ PE1 和 PE2 之间 PW 上的报文封装模式配置为 RAW(Ethernet)模式
- ☞ 要求通过配置 VPLS 的方式建立一个 VPWS 方式的二层 VPN1，使得 CE1 和 CE2 之间可以进行二层互访

5.3.4.2组网图

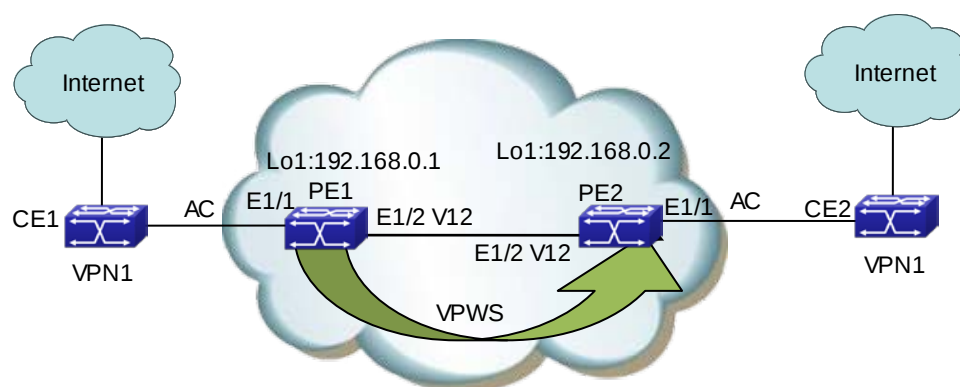


图 5-8 VPWS 组网图

5.3.4.3配置步骤

请按照图 5-8 配置各接口 IP 地址和掩码，包括 VLAN 接口和 Loopback 接口，具体配置过程略。本配置只给出 VPWS 相关的配置。

(1) PE1 的配置

#创建 PW 模板

```
PE1(config)#pw-class c1
PE1(config-class)#transport-mode ethernet
PE1(config-class)#exit
```

#配置路由

```
PE1(config)#router ospf
PE1(config-router)#router-id 192.168.0.1
PE1(config-router)#network 0.0.0.0/0 area 0
PE1(config-router)#exit
```

#配置 MPLS 能力和标签交换能力(非 H-VPLS 情况下与 CE 相连的接口不用使能标签交换能力)

```
PE1(config)#mpls enable
PE1(config)#int vlan 12
PE1(config-if-vlan12)#label-switching
PE1(config-if-vlan12)#exit
```

#配置 LDP

```
PE1(config)#router ldp
PE1(config-router)#router-id 192.168.0.1
PE1(config-router)#targeted-peer 192.168.0.2
PE1(config-router)#exit
```

#配置 VPWS 的对端 PE

```
PE1(config)#l2-vc 192.168.0.2 pw-id 1 pw-class c1
```

#在面向 CE1 的端口上绑定 VPWS 模式为 ethernet

```
PE1(config-if-ethernet1/1)#xconnect l2-vc pw-id 1 mode Ethernet
```

(2) PE2 的配置

#创建 PW 模板

```
PE2(config)#pw-class c1
PE2(config-class)#transport-mode vlan
PE2(config-class)#exit
```

#配置路由

```
PE2(config)#router ospf
```

```
PE2(config-router)#router-id 192.168.0.2
```

```
PE2(config-router)#network 0.0.0.0/0 area 0
```

```
PE2(config-router)#exit
```

#配置 MPLS 能力和标签交换能力(非 H-VPLS 情况下与 CE 相连的接口不用使能标签交换能力)

```
PE2(config)#mpls enable
```

```
PE2(config)#int vlan 12
```

```
PE2(config-if-vlan12)#label-switching
```

```
PE2(config-if-vlan12)#exit
```

#配置 LDP

```
PE2(config)#router ldp
```

```
PE2(config-router)#router-id 192.168.0.2
```

```
PE2(config-router)#targeted-peer 192.168.0.1
```

```
PE2(config-router)#exit
```

```
PE2(config)#int vlan 12
```

```
PE2(config-if-vlan12)#ldp enable
```

```
PE2(config-if-vlan12)#exit
```

#配置 VPWS 的对端 PE

```
PE2(config)#l2-vc 192.168.0.1 pw-id 1 pw-class c1
```

#在面向 CE2 的端口上绑定 VPWS 模式为 VLAN, Svid 为 200

```
PE2(config-if-ethernet1/1)#xconnect l2-vc pw-id 1 mode vlan svid 200
```

5.4 VPLS排错帮助

在配置、使用 VPLS 时,可能会由于物理连接、配置错误等原因导致 L2 VPN 未能正常运行。因此,用户应注意以下要点:

- ☞ 首先应该保证 PE1、P、PE2 之间的 OSPF 邻居建立正常,包括 loopback 口在内的路由发布正常,即检查所有 PE 设备上是否都存在对方 PE、P 的路由;
- ☞ 其次应该保证 PE1、P、PE2 上全局启动了 MPLS 和 LDP 并且在活动接口上正确启动了 Label-switching 和 LDP,检查下 PE1、P、PE2 上的 LDP session 建立是否正常,确认 LDP 的远端邻居即 targeted peer 地址是否配置正确,是否正确建立了邻居,同时检查 LDP 远端 peer 是否和配置的 VFI 上或者 L2VC 上的 peer 地址

是否一致；

- ☞ 然后，通过 **show vpls peer xxx** 命令查看 PE 上 PW 是否建立成功，对应的 Peer 其 pw 状态应该为 UP，如果没有 UP 请检查两端配置的 PW 封装方式(ethernet 或者 vlan)是否一致，VFI ID 是否一致，MTU、传输模式等参数协商是否一致；
- ☞ 接着，判断绑定到 VFI 的端口是否绑定到正确的 VFI 上，其接入模式是否正确，Svid 是否存在错误，如果为 VPWS 方式，检查端口上绑定的 pw-id 是否正确，两端 PE 上绑定的 pw-id 是否相同；
- ☞ 最后，在上述过程都正确的情况下，在 CE1 和 CE2 上使用 **show vfi mac-addresses-table** 可以查看到自己 VPN 内的正确 MAC 地址信息，且两端都可以显示对方 CE 上的 MAC 地址信息；
- ☞ 另外，如果正确的配置下保存重启设备后，CE 上没有立刻看到对端 CE 的 MAC 地址流量也没有立即恢复，请耐心等待一段时间，OSPF、LDP、PW 的连接建立都需要一定时间；
- ☞ 注意： PE 端口在绑定到 VFI 以后建议该端口不要打开 802.1X、STP、QinQ 等可能会影响用户侧的网络功能，另外该端口所属的三层 VLAN 端口如果有路由(例如 OSPF、RIP)、组播(IGMP、MLD)等协议可能会造成运营商的部分 IP 等信息泄漏到 CE 一侧，所以建议该端口不要加入公网上正在使用的三层接口中。

第6章 MAC-in-MAC

6.1 MAC-in-MAC简介

MAC-in-MAC 协议由 IEEE 802.1ah 所定义,又称为 PBB(Provider Backbone Bridge, 运营商骨干网桥)协议。作为一种二层 VPN (Virtual Private Network, 虚拟专用网络)技术,它通过 MAC 地址嵌套,将用户 MAC 地址封装在运营商 MAC 地址之中,从而对不同的用户业务进行隔离。

6.1.1 MAC-in-MAC基本概念

1. PBBN

我们将采用了 MAC-in-MAC 协议的网络称为 PBBN (Provider Backbone Bridge Network, 运营商骨干桥接网),也称为 MAC-in-MAC 网络。从用户的角度来看, PBBN 就是一个二层交换网络,可以在不同节点之间建立二层连接。

2. PBN

我们将连接用户网络和 PBBN 的网络称为 PBN (Provider Bridge Network, 运营商桥接网)。用户网络既可以直接接入 PBBN,也可以通过 PBN 接入 PBBN。

3. MAC-in-MAC 报文

我们将经过 MAC-in-MAC 封装的报文简称为 MAC-in-MAC 报文。

4. BEB

BEB (Backbone Edge Bridge, 骨干网边缘网桥)设备是 PBBN 的边缘设备,相当于 MPLS 网络中的 PE 设备。它负责将来自用户网络的报文进行 MAC-in-MAC 封装,并转发到 PBBN 中;或者将来自 PBBN 的 MAC-in-MAC 报文进行解封装,并转发到用户网络中。

5. BCB

BCB (Backbone Core Bridge, 骨干网核心网桥)设备是 PBBN 的核心设备,相当于 MPLS 网络中的 P 设备。它负责按照 B-MAC 和 B-VLAN 转发 MAC-in-MAC 报文。BCB 设备只需要转发报文和学习骨干网中的 MAC 地址,不需要学习用户网络中大量的 MAC 地址,从而降低了网络部署的成本,也为 PBBN 提供了更好的可扩展性。

6. B-MAC/B-VLAN

BEB 设备在对用户报文进行封装时,会为其打上运营商分配的 MAC 地址和 VLAN。这个由运营商分配的 MAC 地址和 VLAN 就分别称为 B-MAC (Backbone MAC, 骨干网 MAC)

和 B-VLAN（Backbone VLAN，骨干网 VLAN）。在 PBBN 中，BCB 设备就是按照 B-MAC 和 B-VLAN 转发 MAC-in-MAC 报文的。

B-MAC 包括源 B-MAC 和目的 B-MAC，BEB 设备在对用户报文进行封装时，会将自己的 MAC 作为源 B-MAC、将 PBBN 隧道目的端 BEB 设备的 MAC 作为目的 B-MAC 进行封装。

7. 上行口/下行口

BEB 设备上连接 PBBN 的端口称为上行口，连接用户网络的端口称为下行口。来自用户网络的报文被封装成 MAC-in-MAC 报文后，会通过 BEB 设备相应的上行口转发出去；而来自 PBBN 的 MAC-in-MAC 报文被解封装后，会按照用户 MAC 地址通过 BEB 设备相应的下行口转发出去。

8. MAC-in-MAC 实例和 I-SID

在 PBBN 中，一个 MAC-in-MAC 实例代表了运营商提供的一类业务，I-SID（Backbone ServiceInstance Identifier，骨干网服务实例编号）就是这个 MAC-in-MAC 实例的唯一编号。

6.1.2 MAC-in-MAC 基本网络模型

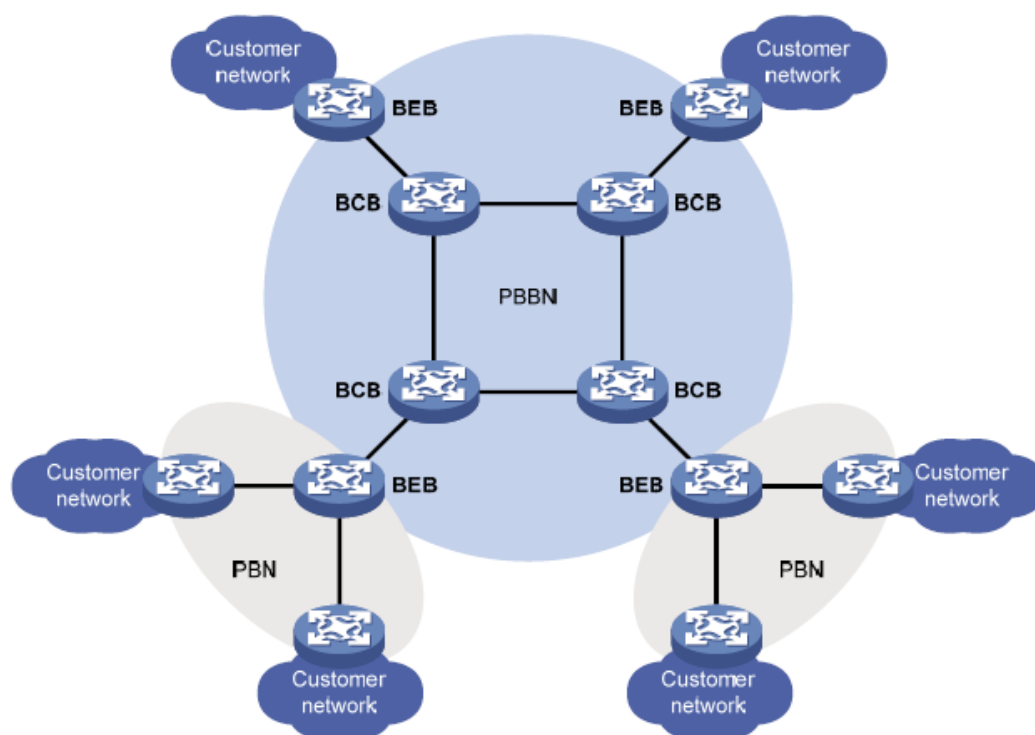


图 6-1 MAC-in-MAC 基本组网模型

MAC-in-MAC 模型主要有 3 部分组成：CE、BEB、BCB。其基本概念如下。

☞ CE（Custom Edge）

直接和服务提供商相连的用户边缘设备。CE 可以是路由器或交换机，也可以是一台主

机。

BEB

BEB（Backbone Edge Bridge，骨干网边缘网桥）设备是 PBBN 的边缘设备，相当于 MPLS 网络中的 PE 设备。它负责将来自用户网络的报文进行 MAC-in-MAC 封装，并转发到 PBBN 中；或者将来自 PBBN 的 MAC-in-MAC 报文进行解封装，并转发到用户网络中。

BCB

BCB（Backbone Core Bridge，骨干网核心网桥）设备是 PBBN 的核心设备，相当于 MPLS 网络中的 P 设备。它负责按照 B-MAC 和 B-VLAN 转发 MAC-in-MAC 报文。BCB 设备只需要转发报文和学习骨干网中的 MAC 地址，不需要学习用户网络中大量的 MAC 地址，从而降低了网络部署的成本，也为 PBBN 提供了更好的可扩展性。

6.1.2.1H-PBBN基本组网模型

H-PBBN（Hierarchical Provider Backbone Bridging Network）是将一个低级别的 PBB 网络在一个更高级别的 PBB 网络中透传，在低级别的 PBB 网络的 MAC-in-MAC 报文基础上再封装一层 MIM 头，形成二层 MIM 头嵌套。

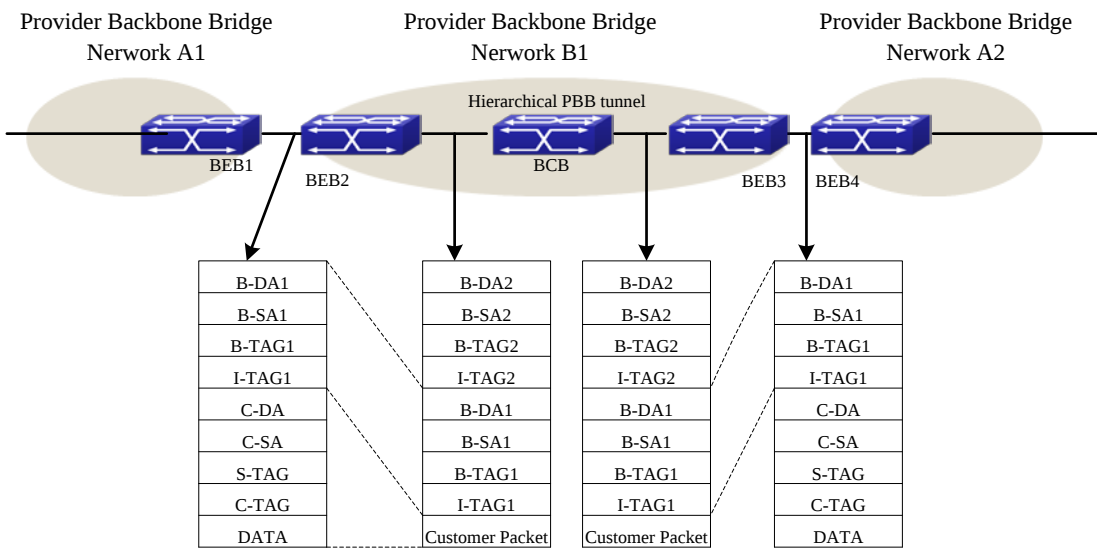


图 6-2 H-PBBN 基本模型

6.1.3 MAC-in-MAC报文封装

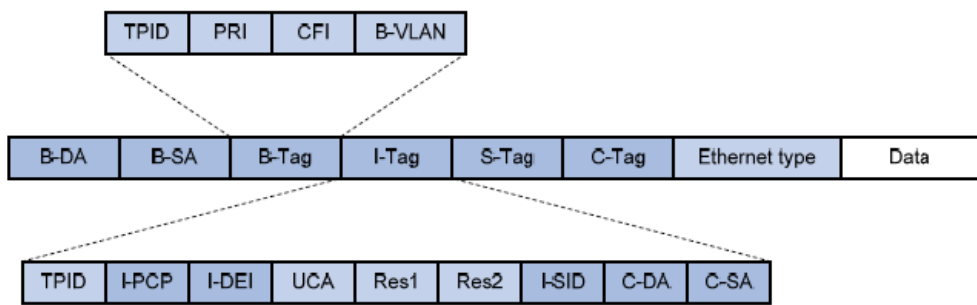


图 6-3 MAC-in-MAC 报文封装格式

字段	全称	含义
B-DA	Backbone Destination MAC address (骨干网目的 MAC 地址)	即目的 B-MAC, 是 MAC-in-MAC 封装的外层目的 MAC 地址, 为 PBBN 隧道目的端 BEB 设备的 MAC 地址, 与 B-SA 合称为 B-MAC
B-SA	Backbone Source MAC address (骨干网源 MAC 地址)	即源 B-MAC, 是 MAC-in-MAC 封装的外层源 MAC 地址, 为 PBBN 隧道源端 BEB 设备的 MAC 地址, 与 B-DA 合称为 B-MAC
B-Tag	Backbone VLAN Tag (骨干网 VLAN 标签)	即 B-VLAN Tag, 是 MAC-in-MAC 封装的外层 VLAN Tag, 用来标识报文在 PBBN 中的 VLAN 和优先级信息, 其 TPID 值固定为 0x8100
I-Tag	Backbone Service Instance Tag (骨干网服务实例标签)	MAC-in-MAC 封装中的业务标记, 包括报文在 BEB 设备上处理时的传送优先级 I-PCP 和丢弃优先级 I-DEI、标识业务实例的 I-SID, 以及用户报文原始的目的 MAC 地址 (即 C-DA) 和源 MAC 地址 (即 C-SA), 其 TPID 值固定为 0x88E7
S-Tag	Service VLAN Tag (服务 VLAN 标签)	报文在 PBN 中的外层 VLAN Tag, 用来标识报文在 PBN 中的 VLAN 和优先级信息
C-Tag	Customer VLAN Tag (用户 VLAN 标签)	报文在 PBN 中的内层 VLAN Tag, 用来标识报文在用户网络中的 VLAN 和优先级信息

表 6-1 MAC-in-MAC 报文封装重点字段含义

6.1.4 MAC-in-MAC 报文转发

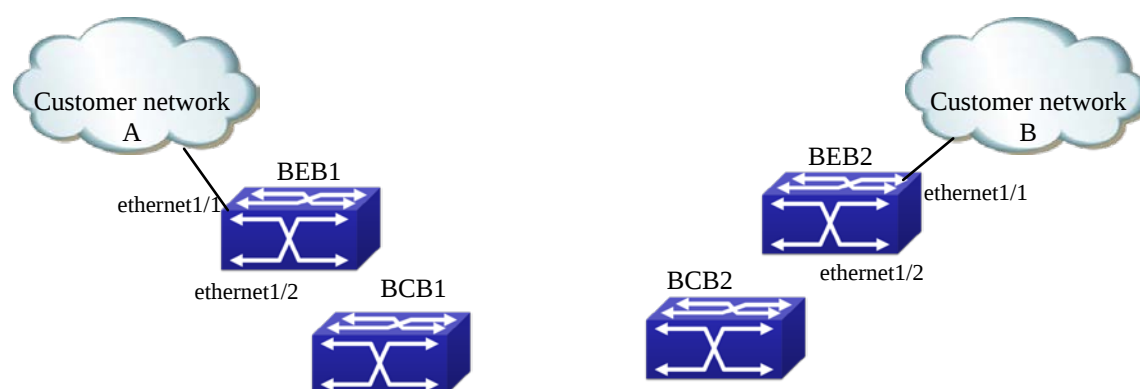


图 6-4 PBBN 的应用组网

6.1.4.1 下行口的接入方式

在 PBBN 基本模型中的下行口接入 BEB 方式有两种, 分别是 VLAN 接入和 Ethernet 接入两种。含义分别如下:

- VLAN 接入: CE 发送给 BEB 或 BEB 发送给 CE 的以太网帧头带有一个 VLAN TAG, 该 TAG 是一个服务提供商网络为了区分用户而压入的“服务界定符”。服务界定符一般

是服务提供商设备添加的，我们把这个作为服务界定符的 TAG 称为 S-TAG。

- Ethernet 接入：CE 发送给 BEB 或 BEB 发送给 CE 的以太网帧头中没有服务界定符，如果此时帧头中有 VLAN TAG，则说明它只是用户报文的内部 VLAN TAG，对于 BEB 设备没有意义。这种用户内部 VLAN 的 TAG 称为 C-TAG。

6.1.4.2 Ethernet接入的报文转发

以图 6-4 的基本模型为例，BEB 的下行端口采用 Ethernet 接入方式，从 customer A 到 customer B 的报文转发过程如下：

1. customer A 将携带用户所属 VLAN 信息（U-Tag）的报文发送给 BEB1；
2. BEB1 根据用户目的 MAC 地址，选择合适的 B-DA，并在报文添加外层 MAC 头；
3. 为了在公网上传输，BEB1 在外层 MAC 头中添加公网的 B-VLAN，通过公网隧道将报文传递给 BEB2；
4. BEB2 收到报文后，根据 B-SA+B-VLAN 找到报文所属的 VFI，并将携带 U-TAG 的报文发送给 customer B。

6.1.4.3 VLAN接入的报文转发

以图 6-4 的基本模型为例，BEB 的下行端口采用 VLAN 接入方式，从 customer A 到 customer B 的报文转发过程如下：

1. customer A 发送的报文应携带服务定符 S-TAG，否则，如果 CE1 发送的报文不带 VLAN TAG 或者所携带的 VLAN TAG 和 S-TAG 不匹配，则走普通的二层报文转发流程；
2. BEB1 根据用户目的 MAC 地址，选择合适的 B-DA，并在报文中添加外层 MAC 头；
3. 为了在公网上传输，BEB1 在外层 MAC 头中添加公网的 B-VLAN，通过公网隧道将报文传递给 BEB2；
4. BEB2 收到报文后，根据 B-SA+B-VLAN 找到报文所属的 VFI，并将携带 S-TAG 的报文发送给 customer B。

6.1.5 MAC-in-MAC的优点

MIM 具有以下优点：

- MIM 解决了 QinQ 只有 4096 个 SVLAN 的限制；
- MIM 隔离了用户的 MAC，增强了以太网业务的安全性；
- MIM 扩展了 MAC 地址空间。

6.2 MAC-in-MAC配置

MAC-in-MAC 配置任务序列：

1. 配置 VFI 及参数（必选）

- (1) 创建 VFI，指定 VFI ID 以及 ISID，并进入 VFI 视图
- (2) 配置 VFI 的 BVLAN（可选）
- (3) 配置 VFI 的默认组播组（可选）
2. 配置下联端口绑定到 VFI（必选）
 - (1) 进入端口视图
 - (2) 将下联端口与指定 VFI 进行绑定，并指定接入模式
3. 配置上联端口绑定到 VFI（必选）
 - (1) 进入端口视图
 - (2) 将该端口配置为 VFI 的上联端口

1. 配置 VFI 及参数（必选）

- (1) 创建 VFI，指定 VFI ID 以及 ISID，并进入 VFI 视图
- (2) 配置 VFI 的 BVLAN（可选）
- (3) 配置 VFI 的默认组播组（可选）

命令	解释
全局配置模式	
vfi <vfi-name> <vfi-id> mim <i-sid> no vfi <vfi-name>	创建一个 MIM 实例并进入到 VFI 模式。在创建实例时必须指定全局唯一 MIM 实例名和 VFI-ID 以及 ISID。本命令的 no 操作删除相应的 VFI。
VFI 配置模式	
mim bvlan <vlan-id> no mim bvlan <vlan-id>	本命令用来给 VFI 实例指定一个 BVLAN，本命令的 no 操作恢复 BVLAN 为默认值。
mim address destination default <mac-addr> no mim address destination default <mac-addr>	为 MIM 实例指定默认的远端目的 B-DA。本命令的 no 操作恢复远端的 B-DA 为默认值。

2. 配置下联端口绑定到 VFI（必选）

- (1) 进入端口视图
- (2) 将下联端口与指定 VFI 进行绑定，并指定接入模式

命令	解释
端口配置模式	
xconnect vfi <vfi-id> [mode {ethernet vlan [svid <svid>]}] no xconnect vfi <vfi-id> [mode {ethernet vlan [svid <svid>]}]	本命令用于将下行端口与 MIM 实例进行绑定，并指定接入模式。本命令的 no 操作删除端口绑定的 MIM 实例。

3. 配置上联端口绑定到 VFI（必选）

- (1) 进入端口视图
- (2) 将该端口配置为 VFI 的上联端口

命令	解释
端口模式	
mim uplink vfi <vfi-id> no mim uplink vfi <vfi-id>	将端口指定为 MAC-in-MAC 实例的上行口。本命令的 no 操作取消端口为指定 MAC-in-MAC 实例的上行口。

6.3 MAC-in-MAC 典型案例

6.3.1 基本的 MAC-in-MAC 应用场景

6.3.1.1 组网需求

- ☞ CE1 和 CE2 分别属于 CustomerA 和 CustomerB 中的两个站点，同属于一个 VPN1
- ☞ CE1 和 CE2 分别通过端口 Eth1/1 接入到 BEB 设备
- ☞ CE1 到 BEB1 的接入链路配置为 Ethernet 接入，CE2 和 BEB2 之间的接入链路配置为 VLAN 接入 Svid 为 200
- ☞ 要求通过配置 MAC-in-MAC 的方式建立一个二层 VPN1，使得 CE1 和 CE2 之间可以进行二层互访

6.3.1.2 组网图

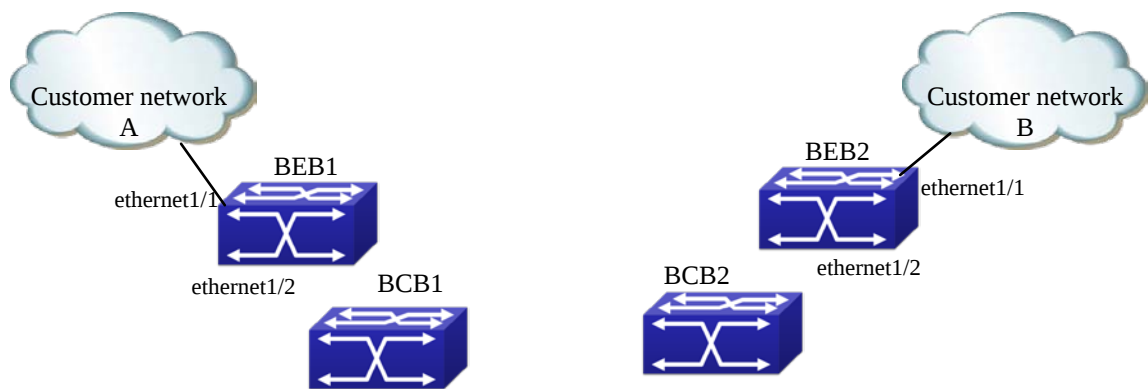


图 6-5 PBBN 的应用组网

6.3.1.3 配置步骤

- (1) BEB1 的配置

#创建 VFI

```
BEB1(config)#vfi a 1 mim 17
```

```
BEB1(config-vfi)#mim bvlan 2
```

```
BEB1(config-vfi)#exit
```

```
BEB1(config)#
```

#配置下联口

```
BEB1(config)#interface ethernet 1/1
```

```
BEB1(config-if-ethernet1/1)#xconnect vfi 1
```

```
BEB1(config-if-ethernet1/1)#exit
```

```
BEB1(config)#
```

#配置上联口

```
BEB1(config)#interface ethernet 1/2
```

```
BEB1(config-if-ethernet1/2)#switchport mode trunk
```

```
BEB1(config-if-ethernet1/2)#mim uplink vfi 1
```

```
BEB1(config-if-ethernet1/2)#exit
```

```
BEB1(config)#
```

(2) BEB2 的配置

#创建 VFI

```
BEB1(config)#vfi a 1 mim 17
```

```
BEB1(config-vfi)#mim bvlan 2
```

```
BEB1(config-vfi)#exit
```

```
BEB1(config)#
```

#配置下联口

```
BEB1(config)#interface ethernet 1/1
```

```
BEB1(config-if-ethernet1/1)#xconnect vfi 1 mode vlan svid 200
```

```
BEB1(config-if-ethernet1/1)#exit
```

```
BEB1(config)#
```

#配置上联口

```
BEB1(config)#interface ethernet 1/2
```

```
BEB1(config-if-ethernet1/2)#switchport mode trunk
```

```
BEB1(config-if-ethernet1/2)#mim uplink vfi 1
```

```
BEB1(config-if-ethernet1/2)#exit
```

```
BEB1(config)#
```

6.4 MAC-in-MAC排错帮助

在配置、使用 MAC-in-MAC 时，可能会由于物理连接、配置错误等原因导致 L2 VPN 未能正常运行。因此，用户应注意以下要点：

- ☞ 首先应该保证 BEB1、BCB、BEB2 之间的物理链路畅通；
- ☞ 其次在创建 VFI 时必须保证同一个 VPN 内不同的 BEB 的 ISID 相同，否则无法通信；
- ☞ 然后应该保证在 BEB 的 VFI 上面添加了上联口和下联口，并且上联口必须为 trunk 或 hybrid 口；
- ☞ 接着，保证 VFI 内配置的 BVLAN 和 VFI 上联口的 native vlan 不同，否则无法通信；
- ☞ 接着，如果下联口的接入方式为 vlan 接入时，需要保证 BEB 收到的报文携带的 tag 与下联口的 svid 一致，否则无法通信；
- ☞ 接着，需要保证下联口是以 tag 方式加入到 BVLAN 中，否则可能会导致从下联口出来的报文没有带上预期的 tag，进而导致通信不正常；
- ☞ 最后，在上述过程都正确的情况下，在 BEB1 和 BEB2 上使用 `show vfi mac-addresses-table` 可以查看到自己 VFI 内的正确 MAC 地址信息，且两端都可以显示对方 CE 上的 MAC 地址信息。