MPLS/EVPN/SRv6 异构网络互通技术白皮书

Copyright © 2026 新华三技术有限公司 版权所有，保留一切权利。
非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。
除新华三技术有限公司的商标外，本手册中出现的其它公司的商标、产品标识及商品名称，由各自权利人拥有。
本文中的内容为通用性技术信息，某些信息可能不适用于您所购买的产品。

目 录

1 概述 1

1.1 MPLS、EVPN、SRv6 三种 VPN 技术简介 1

1.1.1 传统的 MPLS VPN 技术 1

1.1.2 EVPN 技术 2

1.1.3 SRv6 VPN 4

1.2 异构网络互通的必要性 5

2 异构网络互通技术综述 5

2.1 异构网络互通实现方式 5

2.2 异构网络互通总览 7

3 SRv6 网络与 MPLS 网络互通（L3VPN） 8

3.1 概述 8

3.2 技术对比 9

3.3 工作机制 10

3.3.1 网络边界设备部署 VPN 实例方式的工作机制 10

3.3.2 Option B 跨域方式的工作机制 14

4 EVPN L3VPN over SRv6 与 IP L3VPN over SRv6 互通 20

4.1 概述 20

4.2 工作机制 20

4.2.1 组网模型 20

4.2.2 路由发布过程 21

4.2.3 报文转发过程 23

5 MPLS L3VPN 网络与 EVPN L3VPN 网络互通 24

5.1 概述 24

5.2 工作机制 25

5.2.1 私网路由发布 25

5.2.2 报文转发过程 26

6 MPLS L3VPN/EVPN L3VPN 与 EVPN VXLAN 互通 27

6.1 概述 27

6.2 工作机制 28

6.2.1 组网模型 28

6.2.2 私网路由发布 29

6.2.3 报文转发过程 30

7 SRv6 网络与 EVPN VXLAN 网络互通 ····· 31
7.1 概述 ····· 31
7.2 IP L3VPN over SRv6 与 EVPN VXLAN 网络互通的工作机制 ····· 31
7.2.1 组网模型 ····· 31
7.2.2 Locator 网段路由发布 ····· 32
7.2.3 私网路由发布 ····· 33
7.2.4 报文转发过程 ····· 34
7.3 EVPN L3VPN over SRv6 与 EVPN VXLAN 网络互通的工作机制 ····· 35
7.3.1 组网模型 ····· 35
7.3.2 Locator 网段路由发布 ····· 35
7.3.3 私网路由发布 ····· 36
7.3.4 报文转发过程 ····· 38
7.4 EVPN VPLS over SRv6 与 EVPN VXLAN 网络互通的工作机制 ····· 39
7.4.1 网络模型 ····· 39
7.4.2 网络互通原理 ····· 39
7.4.3 控制平面工作机制 ····· 40
7.4.4 数据平面工作机制 ····· 44
8 EVPN VXLAN 与 EVPN VPLS 网络互通 ····· 47
8.1 网络模型 ····· 47
8.2 网络互通原理 ····· 47
8.3 控制平面工作机制 ····· 48
8.3.1 PW 及 BUM 广播表建立 ····· 48
8.3.2 VXLAN 隧道及 BUM 广播表建立 ····· 49
8.3.3 MAC 地址学习 ····· 49
8.4 数据平面工作机制 ····· 52
8.4.1 转发已知单播流量 ····· 52
8.4.2 转发 BUM 流量 ····· 53
9 VPLS 与 EVPN VPLS/EVPN VPLS over SRv6/EVPN VXLAN 互通 ····· 55
9.1 概述 ····· 55
9.2 工作机制 ····· 56
9.2.1 建立 PW ····· 56
9.2.2 MAC 地址学习 ····· 57
9.2.3 报文转发过程 ····· 58
10 VPWS 与 EVPN VPWS/EVPN VPWS over SRv6 互通 ····· 59
10.1 概述 ····· 59
10.2 工作机制 ····· 60

10.2.1 建立 PW……60
10.2.2 报文转发过程 ……61

1 概述

1.1 MPLS、EVPN、SRv6三种VPN技术简介

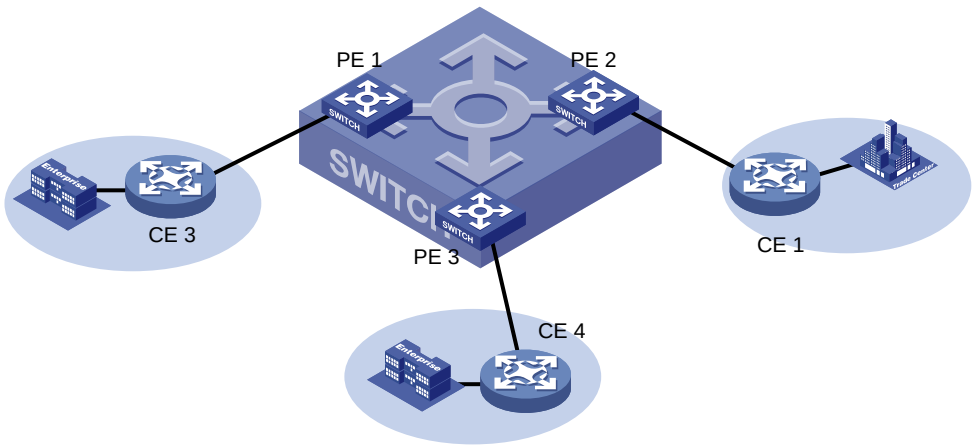
MPLS、EVPN（Ethernet Virtual Private Network，以太网虚拟专用网络）和 SRv6（Segment Routing IPv6，IPv6 段路由）是现代网络中常见的 VPN（Virtual Private Network，虚拟专用网络）技术，共同推动了网络的发展和应用。它们可以为用户提供二层 VPN（L2VPN）和三层 VPN（L3VPN）服务。

1.1.1 传统的 MPLS VPN 技术

MPLS VPN 是基于 MPLS 技术的一种虚拟专有网络（VPN）技术，通过标签来标识私网，并通过标签交换技术实现报文的转发。MPLS VPN 技术分为以下几种：

- BGP/MPLS L3VPN（以下简称 MPLS L3VPN）：构建在 BGP 和 MPLS 基础上的三层 VPN 技术。它借助 BGP 协议进行 VPN 路由发布，使用 VPN 实例实现业务隔离，并通过 MPLS 标签交换技术来实现 VPN 报文的转发。
- VPLS（Virtual Private LAN Service，虚拟专用局域网服务）：用来在 MPLS 骨干网上提供点到多点的 L2VPN 服务。VPLS 结合了以太网技术和 MPLS 技术的优势，是对传统 LAN 全部功能的仿真，其主要目的是通过运营商提供的 IP/MPLS 网络连接地域上隔离的多个由以太网构成的 LAN，使它们像一个 LAN 那样工作。如图 1 所示，服务提供商利用 VPLS 技术在 MPLS 骨干网络上为用户网络模拟了一个以太网桥，基于 MAC 地址或者 MAC 地址+VLAN 来做出转发决策。

  图1 VPLS 模拟以太网桥



- VPWS（Virtual Private Wire Service，虚拟专线服务）：基于 MPLS 的二层 VPN 技术，是 PWE3（Pseudo Wire Emulation Edge-to-Edge，边缘到边缘的伪线仿真）的一种实现方式，用来实现点到点的二层互通。VPWS 将用户的二层数据（如以太网数据帧、ATM 信元等）封装成可以在 IP 或 MPLS 网络中传送的分组，通过 IP 路径或 MPLS 隧道转发，接收端解封装分组后恢复原来的二层数据，从而实现用户二层数据跨越 MPLS 或 IP 网络在不同站点间透明地传送。

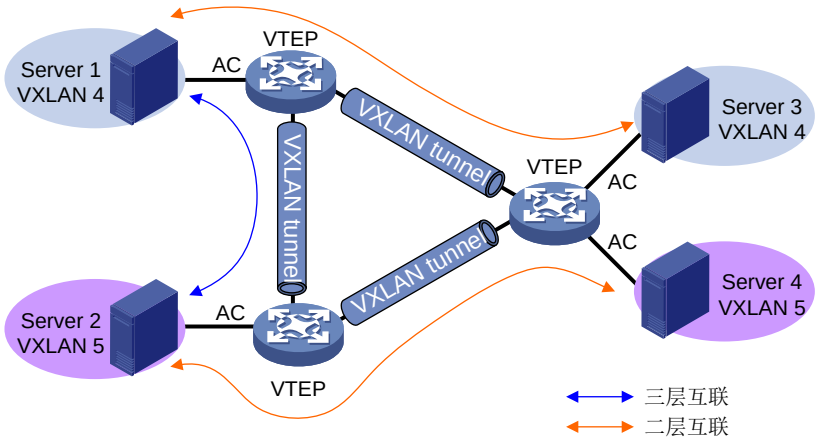
1.1.2 EVPN 技术

EVPN 是一种满足全业务需求的 NVO（Network Virtualization Overlay）解决方案，具有部署简单、扩展性强等优点。在控制平面，EVPN 采用 MP-BGP 协议通告二层或三层可达性信息，通过生成的 MAC 地址表项和路由表项进行二/三层报文转发，以实现为用户提供 L2VPN 和 L3VPN 服务。在数据平面，EVPN 支持 VXLAN、SRv6 或 MPLS 等多种报文转发方式，增加了网络部署的灵活性。

根据数据平面封装方式的不同，EVPN 支持如下几种组网方式：

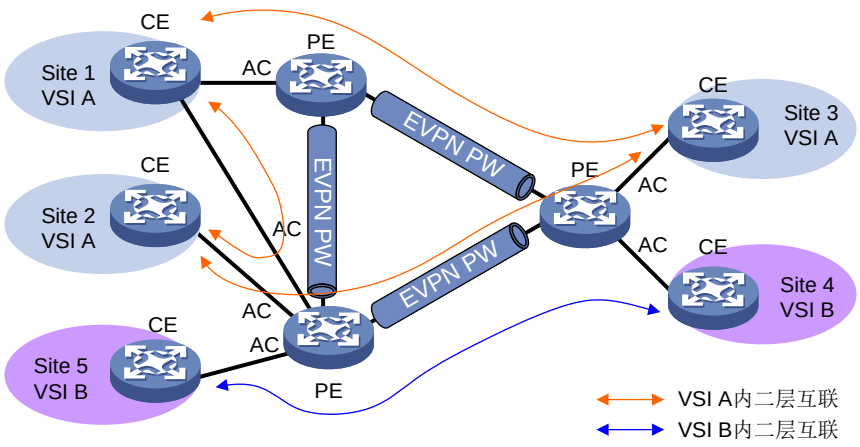
- EVPN VXLAN：数据平面采用 VXLAN 封装。EVPN VXLAN 网络的边缘设备称为 VTEP（VXLAN Tunnel End Point，VXLAN 隧道端点），EVPN 的相关处理均在 VTEP 上完成。EVPN VXLAN 通过在 VTEP 间建立 VXLAN 隧道，透明传输二层数据报文，实现不同站点间的二层互联。通过在 EVPN VXLAN 网络中部署 EVPN 网关，可以实现为同一租户的不同子网提供三层互联，并为其提供与外部网络的三层互联。

  图2 EVPN VXLAN 网络示意图



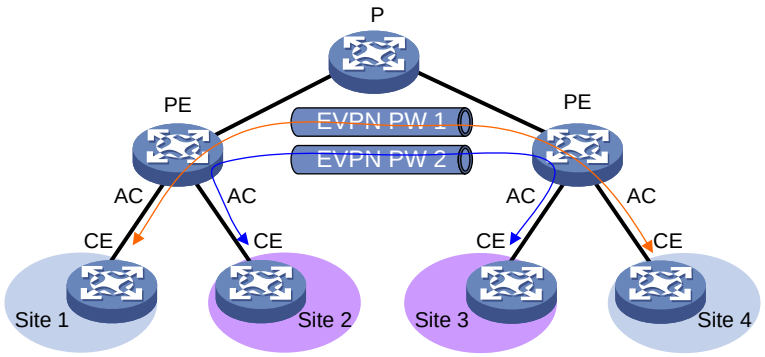
- EVPN VPLS：数据平面采用 MPLS 封装，用来实现用户的点到多点二层互通。EVPN VPLS 组网中，服务提供商网络侧设备 PE 间通过 MP-BGP 协议建立 EVPN PW、学习 MAC 地址表项，并通过查找 MAC 地址表转发报文，以实现多点间的二层互通。

  图3 EVPN VPLS 网络示意图



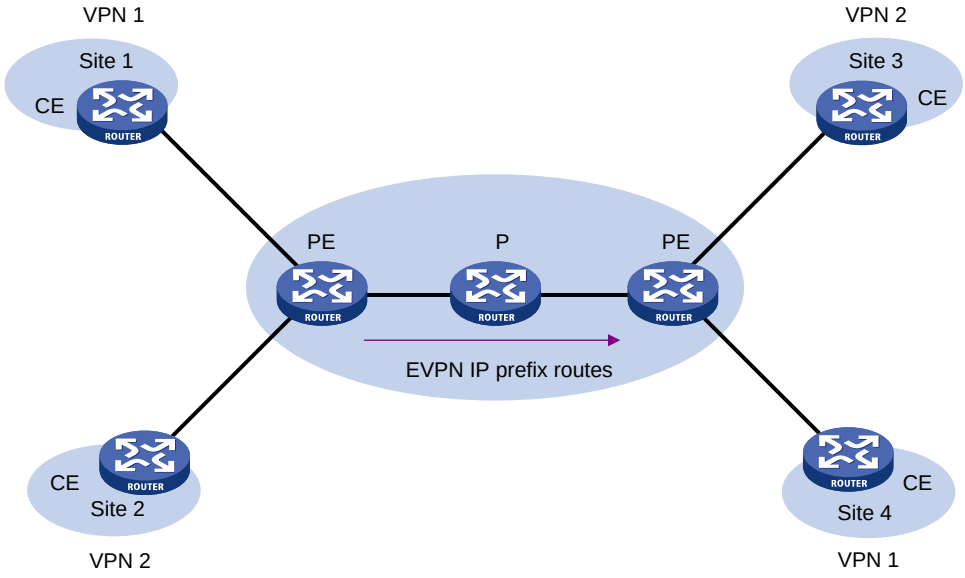
- EVPN VPWS：数据平面采用 MPLS 封装，用来实现用户站点网络的点到点二层互通。EVPN VPWS 组网中，用户网络侧设备 CE 通过 AC（Attachment Circuit，接入电路）接入服务提供商网络侧设备 PE，PE 间通过 MP-BGP 协议建立 EVPN PW，在 PE 上将 AC 与 EVPN PW 关联，即可沿着本端 CE—>本端 AC—>本端 PE—>EVPN PW—>远端 PE—>远端 AC—>远端 CE 的路径实现点到点二层互通。

  图4 EVPN VPWS 网络示意图



- EVPN L3VPN：EVPN 的 IP 前缀路由可以用来发布 VPN 私网路由信息，以实现 MPLS L3VPN 组网，该网络称为 EVPN L3VPN。与 MPLS L3VPN 网络相比，EVPN L3VPN 组网中，在 EVPN 的基础上可以快速部署大二层网络，使得网络同时承载二层 VPN 和三层 VPN 业务。

  图5 EVPN L3VPN 网络示意图



- EVPN VPLS over SRv6：数据平面采用 SRv6 封装，用来实现用户站点网络的多点间二层互通。EVPN VPLS over SRv6 与 EVPN VPLS 类似，不同之处在于 EVPN VPLS over SRv6 组网中，PE 之间建立的虚拟连接为 SRv6 隧道。
- EVPN VPWS over SRv6：数据平面采用 SRv6 封装，用来实现用户站点网络的点到点二层互通。EVPN VPWS over SRv6 与 EVPN VPWS 类似，不同之处在于 EVPN VPWS over SRv6 组网中，PE 之间建立的虚拟连接为 SRv6 隧道。

1.1.3 SRv6 VPN

SRv6 是指基于 IPv6 转发平面实现 SR。SRv6 通过在 IPv6 报文中插入一个路由扩展头 SRH（Segment Routing Header，SRv6 报文头），并在 SRH 中添加路径所要经过的所有段的 SID，即 SID List，来显式指定该报文的转发路径。

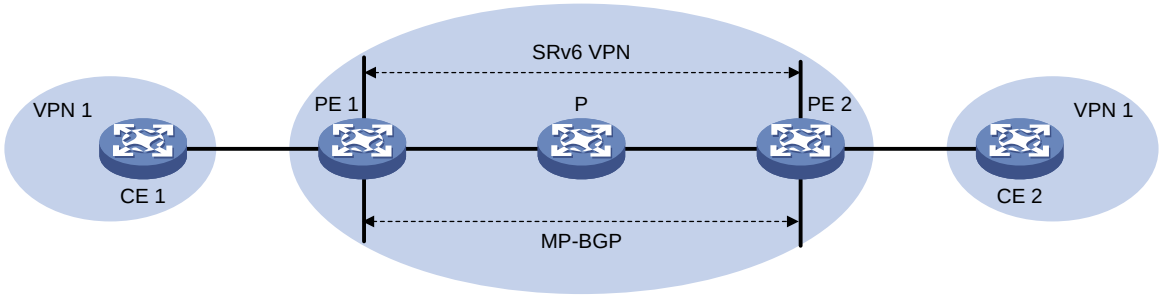
SRv6 VPN 是通过 SRv6 隧道承载 IPv6 网络中的 VPN 业务的技术，控制平面采用 MP-BGP 通告 VPN 路由信息，数据平面采用 SRv6 封装方式转发报文。租户的物理站点分散在不同位置时，SRv6 VPN 可以基于已有的服务提供商或企业 IP 网络，为同一租户的不同物理站点提供二层或三层互联。

根据 VPN 业务种类，SRv6 VPN 分为 L3VPN 业务（IP L3VPN over SRv6 和 EVPN L3VPN over SRv6）、L2VPN 业务（EVPN VPLS over SRv6 和 EVPN VPWS over SRv6）。

- IP L3VPN over SRv6

  如图 6 所示，IP L3VPN over SRv6 通过 MP-BGP 在 IPv6 骨干网上发布用户站点的 IPv4/IPv6 私网路由，使用 PE 间的 SRv6 路径承载私网报文，从而实现通过 IPv6 骨干网连接属于同一个 VPN、位于不同地理位置的用户。

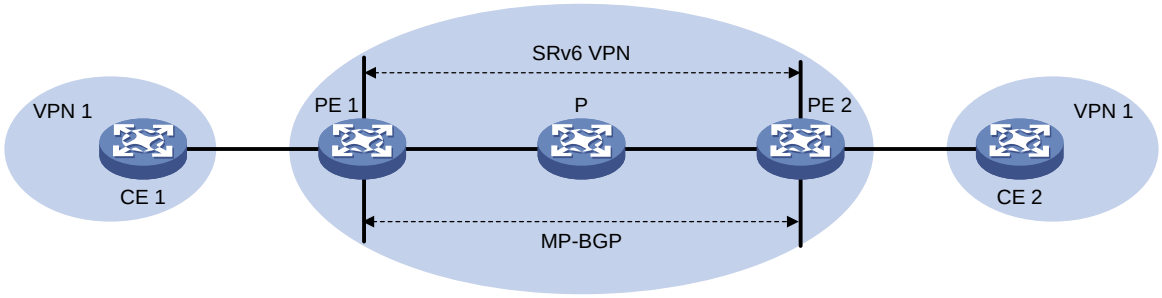
  图6 IP L3VPN over SRv6 组网示意图



- EVPN L3VPN over SRv6

  如图 7 所示，EVPN L3VPN over SRv6 通过 MP-BGP 在 IPv6 骨干网上使用 EVPN 的 IP 前缀路由发布用户站点的 IPv4/IPv6 私网路由，使用 PE 间的 SRv6 路径承载私网报文，从而实现通过 IPv6 骨干网连接属于同一个 VPN、位于不同地理位置的用户。

  图7 EVPN L3VPN over SRv6 组网示意图



- EVPN VPLS over SRv6

详细介绍，请参见“1.1.2 EVPN 技术”。

- EVPN VPWS over SRv6

  详细介绍，请参见“1.1.2 EVPN 技术”。

1.2 异构网络互通的必要性

在当今网络环境中，MPLS VPN 作为传统的 VPN 技术部署广泛，而新兴的 EVPN 和 SRv6 VPN 也越来越受欢迎。这些网络技术之间存在差异性和独特性。分布式、异构的网络资源管理成为网络部署和管理的难点，导致网络的可用性和灵活性有限。为此，需要通过网络互通的方式来实现异构网络之间的通信，促进不同类型的网络之间的集成和协作，以提高整个网络的管理和控制效率，实现更好的网络性能和服务质量。

实现异构网络互通可以带来如下好处：

- 增强网络灵活性：打破了传统网络技术之间的壁垒，可以在不同的网络中使用不同的技术，不必局限于某一种技术，从而更好地满足不同服务的需求。
- 提高网络安全性：通过使用不同的协议来进行互联，可以避免单一技术的漏洞和缺陷，并提高网络的安全性和可靠性。
- 易于管理和维护：由于不同的网络使用不同的技术，可以更好地区分和管理网络资源，避免资源的浪费和冲突，并使维护更加便捷和高效。

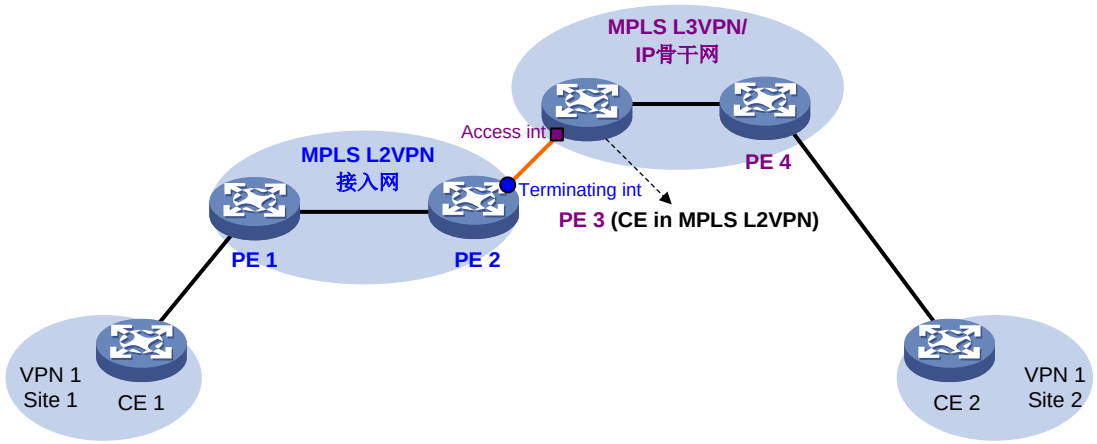
2 异构网络互通技术综述

2.1 异构网络互通实现方式

若要实现异构网络的互通，则需要在异构网络之间“打通”数据传输通道，使得报文能够从一种网络传递到另一种网络。在异构网络之间“打通”数据传输通道的方式主要分为以下几种：

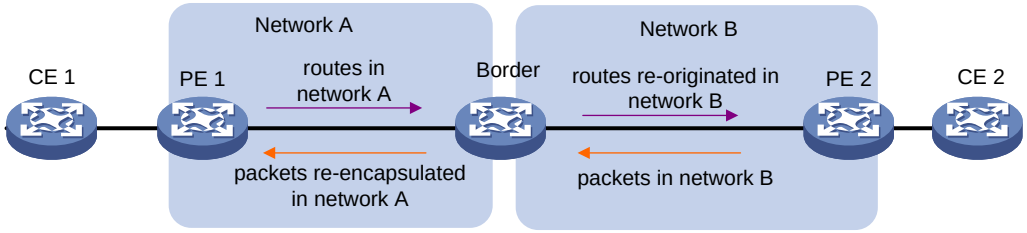
- 终结后重接入方式：在不同网络的边界上部署两台设备，分别用来终结一种网络和接入另一种网络。如图 8 所示，以 MPLS L2VPN 接入 MPLS L3VPN 网络为例，该方式的工作机制为边界设备 PE 2 上终结 MPLS L2VPN 业务，恢复原始报文，原始报文通过 PE 3 接入 MPLS L3VPN 网络。采用该方式时，MPLS L2VPN 和 MPLS L3VPN 之间彼此独立，与各自单独运行时的机制相同。该方式可以实现所有异构网络的互通，但是需要部署额外的设备，且在边界设备上 VPN 信息会丢失，异构的网络独立部署 VPN，无法实现端到端的 VPN 业务隔离。本文不展开介绍该方式。

图8 终结后重接入方式示意图



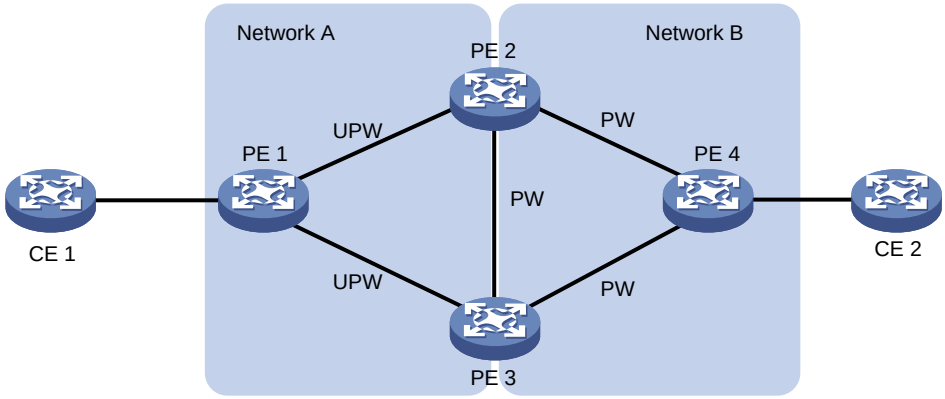
- 修改路由方式：通过路由重生成等手段在不同网络的边界设备上修改路由信息，将在一种网络内发布的路由转变为可以在另一种网络中发布和识别的路由，以实现路由信息的端到端传递。数据报文根据学习到的路由进行转发。边界设备将报文从一种网络转发到另一种网络时，重新对报文进行封装，使得该报文可以在新的网络中顺利转发到目的地。不同的 L3VPN 之间互通时，通常使用这种方式。

图9 路由重生成方式示意图



- UPW 接入方式：将一种网络内的虚拟连接（例如，VPLS 网络中的 PW）看作为另一个网络的 AC（Attachment Circuit，接入电路），该虚拟连接称为 UPW。作为 AC 的 UPW 可以关联另一个网络的 PW 或 VSI。边界设备从 UPW 接收到报文后，解除报文的封装，通过 AC 关联的 PW 转发报文，或在 AC 关联的 VSI 内查表转发报文。边界设备从网络 B 内接收到报文后，解除报文的封装，将其转发给对应的 AC（即 UPW），并为报文添加 UPW 的封装。不同的 L2VPN 之间互通时，通常使用这种方式。

图10 UPW 接入方式示意图



2.2 异构网络互通总览

不同类型网络之间互通的方式有所不同。表 1 罗列了部分网络之间的互通方式。

表1 不同类型网络之间的互通方式汇总表

	VPLS	VPWS	EVPN VXLAN	EVPN VPLS	EVPN VPWS	EVPN L3VPN	EVPN VPLS over SRv6	EVPN VPWS over SRv6	IP L3VPN over SRv6	EVPN L3VPN over SRv6
BGP/ MPLS L3VPN	终结后重接入	终结后重接入	修改路由	终结后重接入	终结后重接入	修改路由	终结后重接入	终结后重接入	修改路由	修改路由
VPLS		暂不支持	UPW接入	UPW接入	终结后重接入	终结后重接入	UPW接入	终结后重接入	终结后重接入	终结后重接入
VPWS			终结后重接入	终结后重接入	UPW接入	终结后重接入	终结后重接入	UPW接入	终结后重接入	终结后重接入
EVPN VXLAN				修改路由	终结后重接入	修改路由	修改路由	终结后重接入	修改路由	修改路由
EVPN VPLS					终结后	终结后重接入	终结后重接入	终结后重接入	终结后重接入	终结后重接入

	VPLS	VPWS	EVPN VXLAN	EVPN VPLS	EVPN VPWS	EVPN L3VPN	EVPN VPLS over SRv6	EVPN VPWS over SRv6	IP L3VPN over SRv6	EVPN L3VPN over SRv6
					重接入					
EVPN VPWS						终结后重接入	终结后重接入	终结后重接入	终结后重接入	终结后重接入
EVPN L3VPN							终结后重接入	终结后重接入	修改路由	修改路由
EVPN VPLS over SRv6								暂不支持	终结后重接入	终结后重接入
EVPN VPWS over SRv6									终结后重接入	终结后重接入
IP L3VPN over SRv6										修改路由

3 SRv6 网络与 MPLS 网络互通（L3VPN）

3.1 概述

当承载私网流量的骨干网穿越了 SRv6 网络与 MPLS 网络时，需要实现 SRv6 网络与 MPLS 网络互通。

SRv6 网络与 MPLS 网络互通可区分为以下两种方式：

- 网络边界设备部署 VPN 实例
- 网络边界设备不部署 VPN 实例（Option B 跨域方式）

上述两种方式的实现原理不相同。

而根据骨干网架构的不同，SRv6 网络与 MPLS 网络互通又可以分为如下四种类型：

- IP L3VPN over SRv6 网络与 MPLS L3VPN 网络互通。
- IP L3VPN over SRv6 网络与 EVPN L3VPN 网络互通。
- EVPN L3VPN over SRv6 网络与 MPLS L3VPN 网络互通。
- EVPN L3VPN over SRv6 网络与 EVPN L3VPN 网络互通。

对于同一种互通方式，上述四种类型的网络互通实现原理类似，区别仅在于：

- IP L3VPN over SRv6 网络与 MPLS L3VPN 网络互通：IP L3VPN over SRv6 网络与 MPLS L3VPN 网络均使用 BGP VPNv4 或 BGP VPNv6 路由承载私网路由信息，网络边界设备无需修改路由类型。
- IP L3VPN over SRv6 网络与 EVPN L3VPN 网络互通：IP L3VPN over SRv6 网络使用 BGP VPNv4 或 BGP VPNv6 路由承载私网路由信息，EVPN L3VPN 网络使用 IP 前缀路由承载私

网路由信息。网络边界设备需要将来自 IP L3VPN over SRv6 网络的 BGP VPNv4/VPNv6 路由转换成 IP 前缀路由后发往 EVPN L3VPN 网络；以及需要将来自 EVPN L3VPN 网络的 IP 前缀路由转换成 BGP VPNv4/VPNv6 路由后发往 IP L3VPN over SRv6 网络。

- EVPN L3VPN over SRv6 网络与 MPLS L3VPN 网络互通：EVPN L3VPN over SRv6 网络使用 IP 前缀路由承载私网路由信息，MPLS L3VPN 网络使用 BGP VPNv4 或 BGP VPNv6 路由承载私网路由信息。网络边界设备需要将来自 EVPN L3VPN over SRv6 网络的 IP 前缀路由转换成 BGP VPNv4/VPNv6 路由后发往 MPLS L3VPN 网络；以及需要将来自 MPLS L3VPN 网络的 BGP VPNv4/VPNv6 路由转换成 IP 前缀路由后发往 EVPN L3VPN over SRv6 网络。
- EVPN L3VPN over SRv6 网络与 EVPN L3VPN 网络互通：EVPN L3VPN over SRv6 网络与 EVPN L3VPN 网络均使用 IP 前缀路由承载私网路由信息，网络边界设备无需修改路由类型。

在上述四种类型的网络互通中，除路由类型区别外，私网路由信息发布过程中的路由关键信息携带、修改机制，以及数据报文转发机制基本相同。本文的后续介绍均仅以 IP L3VPN over SRv6 网络与 MPLS L3VPN 网络互通为例。

3.2 技术对比

不同场景下 SRv6 网络与 MPLS 网络互通方式的对比，如表 2 所示。

表2 网络边界设备存在 VPN 实例和不存在 VPN 实例情况对比

场景条件	描述
是否需要跨越AS域	• 网络边界设备部署 VPN 实例方式：无需跨域 AS 域 • Option B 跨域方式：需要跨越 AS 域
是否需要重生成私网路由	• 网络边界设备存在 VPN 实例方式：网络边界设备必须对收到的私网路由进行重生成。重生成后的私网路由无法保留源路由100%的路由信息，如携带的 RD、RT 属性等信息都可能发生改变 • Option B 跨域方式：无需重生成私网路由信息，可以最大限度地保留私网路由的信息完整性

其他优势	• 网络边界设备存在 VPN 实例方式：
	  ○ 网络边界设备可以将属于不同站点、不同 VPN 实例的私网路由信息统一收集到一个 VPN 实例，为这些私网路由在同一个 VPN 实例内分配 SRv6 SID 后再进行转发，可以节省 SRv6 SID 资源
	  ○ 可兼容 HoVPN 架构，在 HoVPN 架构网络中实现 SPE 和 UPE 跨越不同类型的网络互通
	• Option B 跨域方式：网络边界设备无需部署 VPN 实例，可以单纯作为私网路由信息的“中转站”，部署难度低

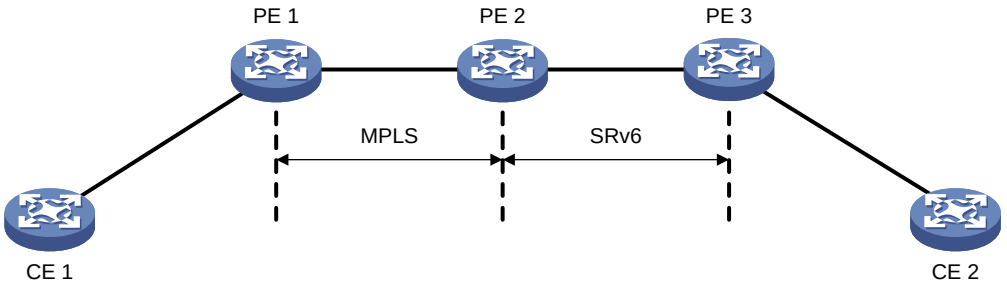
3.3 工作机制

3.3.1 网络边界设备部署 VPN 实例方式的工作机制

1. 组网模型

如图 11 所示，PE 1 和 PE 2 之间部署 MPLS L3VPN 网络，建立 BGP VPNv4 会话，PE 2 和 PE 3 之间部署 IP L3VPN over SRv6 网络，建立 BGP VPNv4 会话。部署 SRv6 与 MPLS 网络互通后，CE 1 和 CE 2 连接的私网用户之间才能互通。

图11 网络边界设备部署 VPN 实例方式的 SRv6 网络与 MPLS 网络互通组网

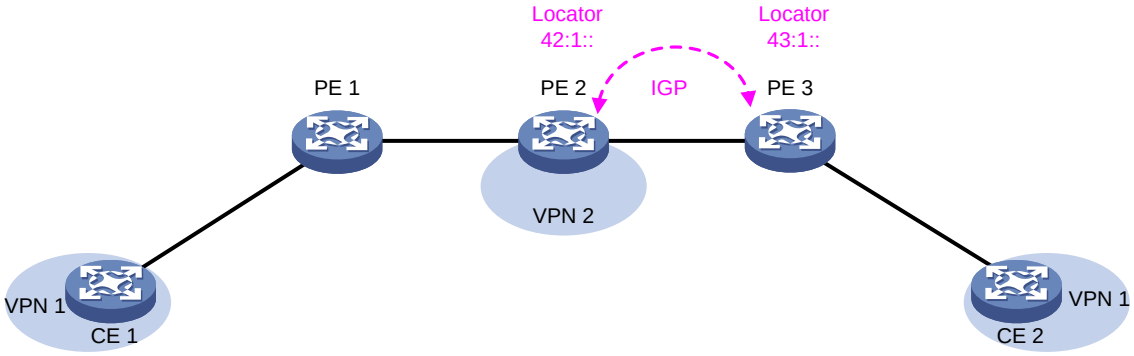


互通主要在 SRv6 网络和 MPLS 网络的边界设备 PE 2 上实现，下面将从路由发布和报文转发两个方面来介绍 SRv6 网络和 MPLS 网络互通的工作机制。

2. Locator 网段路由发布

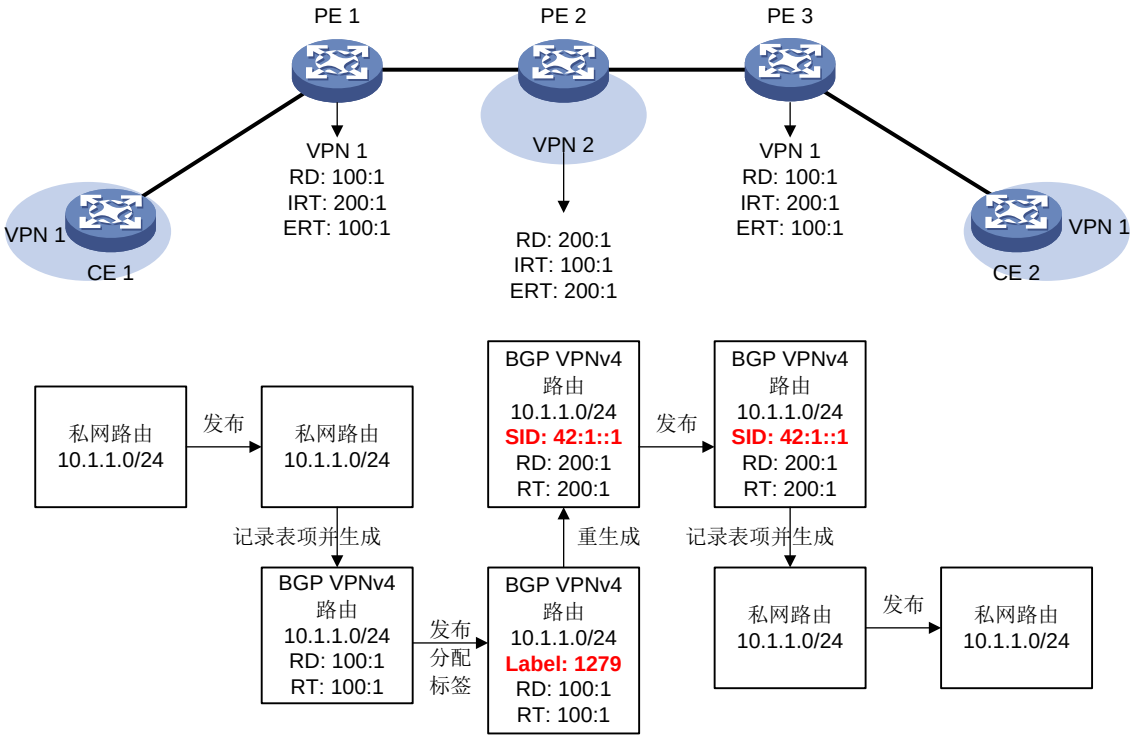
如图 12 所示，PE 2 和 PE 3 需要通过 IGP 协议向对方发布各自的 Locator 网段路由信息。PE 2 发布的 Locator 网段路由前缀为 42:1::/64，PE 3 发布的 Locator 网段路由前缀为 43:1::/64。

图12 网络边界设备部署 VPN 实例方式中 Locator 网段路由发布



3. 私网路由发布

图13 网络边界设备部署 VPN 实例方式中 CE 1 到 CE 2 的路由发布过程



如图 13 所示，PE 1 和 PE 2 均部署 VPN 实例 1，通过 VPN 实例 1 接入 CE 1 和 CE 2，PE 2 部署 VPN 实例 2。私网路由从 CE 1 发往 CE 2 的过程为：

(1) CE 1 通过 IGP、BGP 或静态路由等方式将本地的私网路由 10.1.1.0/24 发布给 PE 1。

(2) PE 1 接收到私网路由信息后，将其添加到 VPN 实例 1 的 IP 路由表中，并为其增加 RD 100:1 和 RT 100:1 属性，生成 BGP VPNv4 路由。

(3) PE 1 为私网路由信息生成的 BGP VPNv4 路由分配私网标签 1279，并将该路由通过 BGP VPNv4 会话发布给 PE 2。私网标签 1279 绑定的 VPN 实例为 VPN 实例 1。

(4) PE 2 本地 VPN 实例 2 配置的 IRT 100:1 与 BGP VPNv4 路由 10.1.1.0/24 携带的 RT 属性匹配，所以 PE 2 接收该路由、将其引入到 VPN 实例 2 并添加到 VPN 实例 2 的 IP 路由表中。

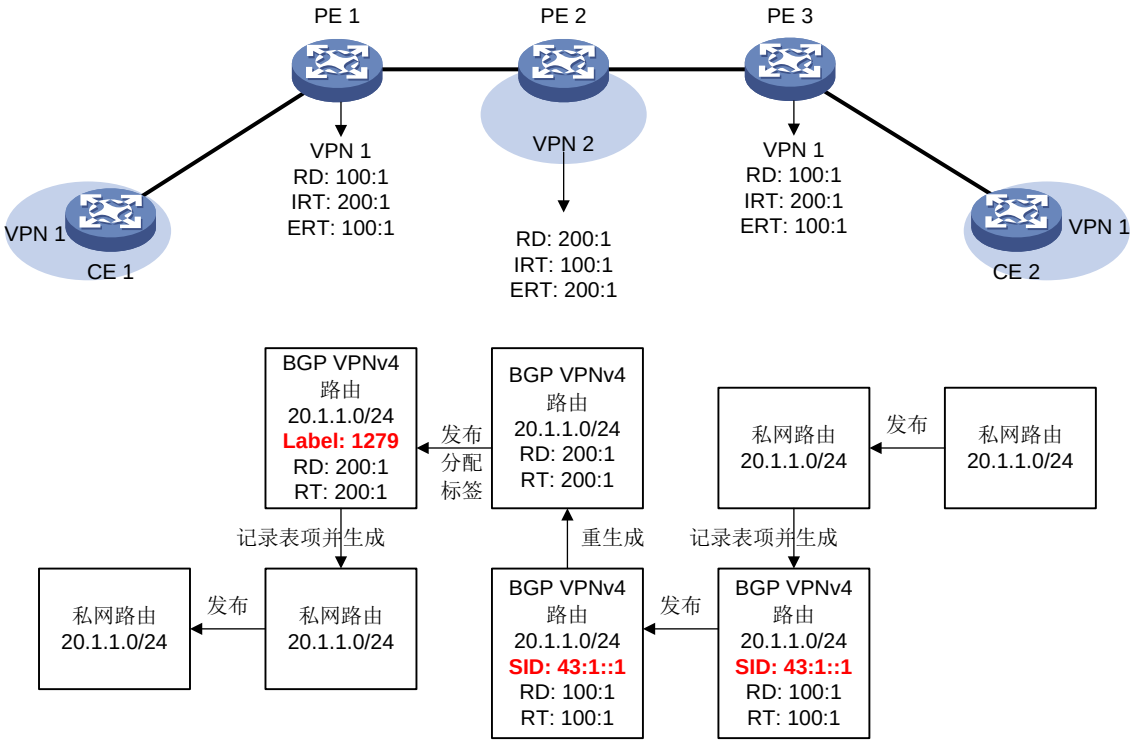
(5) PE 2 重生成引入到 VPN 实例 2 中的 BGP VPNv4 路由 10.1.1.0/24，重生成的 BGP VPNv4 路由 10.1.1.0/24 携带 VPN 实例 2 的 RD 200:1 以及 ERT 200:1，并且，PE 2 为其在 VPN 实例 2 中申请 SRv6 SID 42:1::1（SRv6 SID 类型可以为 End.DT4 SID、End.DT6 SID、End.DT46 SID、End.DX4 SID 或 End.DX6 SID，本例中的 SRv6 SID 为 End.DT4 SID）。

(6) PE 2 将重生成的 BGP VPNv4 路由 10.1.1.0/24 发布给 PE 3。

(7) PE 3 本地 VPN 实例 1 配置的 IRT 200:1 与 BGP VPNv4 路由 10.1.1.0/24 携带的 RT 属性匹配，所以 PE 3 接收该路由、将其引入到 VPN 实例 1 并添加到 VPN 实例 1 的 IP 路由表中。

(8) PE 3 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 10.1.1.0/24 发布给 CE 2。

图14 网络边界设备部署 VPN 实例方式中 CE 2 到 CE 1 的路由发布过程



如图 14 所示，私网路由从 CE 2 发往 CE 1 的过程为：

(1) CE 2 通过 IGP、BGP 或静态路由等方式将本地的私网路由 20.1.1.0/24 发布给 PE 3。

(2) PE 3 接收到私网路由信息后，将其添加到 VPN 实例 1 的 IP 路由表中，并为其增加 RD 100:1 和 RT 100:1 属性，生成 BGP VPNv4 路由。

(3) PE 3 为私网路由信息生成的 BGP VPNv4 路由在 VPN 实例 1 内申请 SRv6 SID 43:1::1（SRv6 SID 类型可以为 End.DT4 SID、End.DT6 SID、End.DT46 SID、End.DX4 SID 或 End.DX6 SID，本例中的 SRv6 SID 为 End.DT4 SID），并将该路由通过 BGP VPNv4 会话发布给 PE 2。

(4) PE 2 本地 VPN 实例 2 配置的 IRT 100:1 与 BGP VPNv4 路由 20.1.1.0/24 携带的 RT 属性匹配，所以 PE 2 接收该路由、将其引入到 VPN 实例 2 并添加到 VPN 实例 2 的 IP 路由表中。

(5) PE 2 重生成引入到 VPN 实例 2 中的 BGP VPNv4 路由 20.1.1.0/24，重生成的 BGP VPNv4 路由 20.1.1.0/24 携带 VPN 实例 2 的 RD 200:1 以及 ERT 200:1，并且不再携带 SRv6 SID。

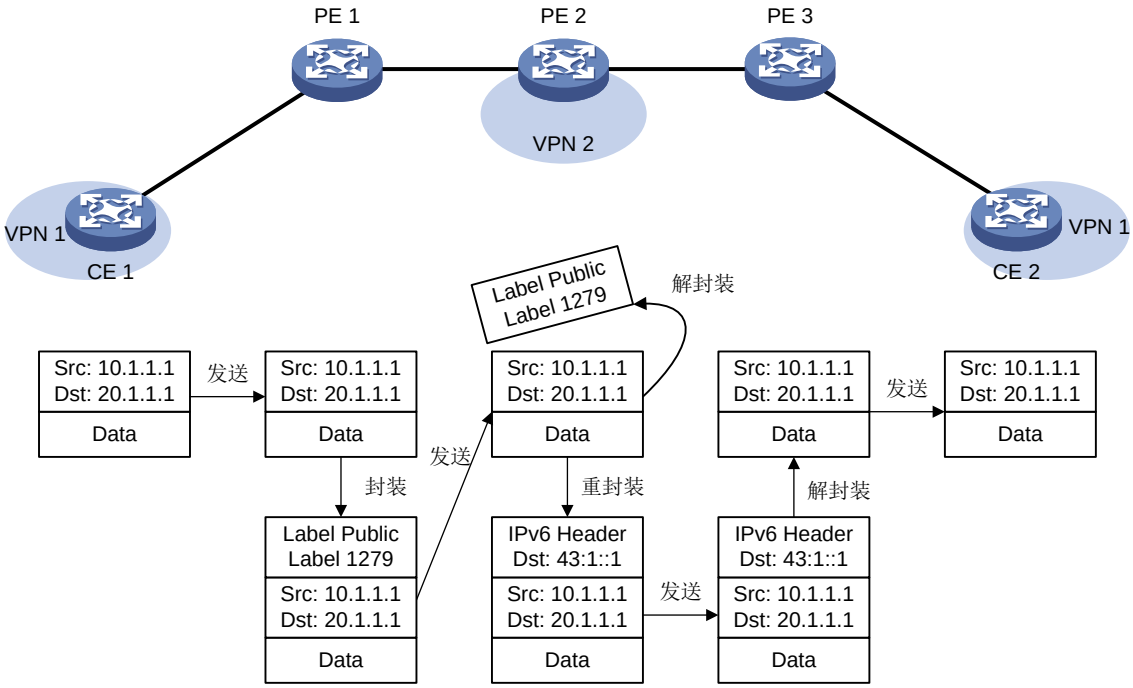
(6) PE 2 为重生成的 BGP VPNv4 路由 20.1.1.0/24 分配私网标签 1279 后将其发布给 PE 1。私网标签 1279 绑定的 VPN 实例为 VPN 实例 2。

(7) PE 1 本地 VPN 实例 1 配置的 IRT 200:1 与 BGP VPNv4 路由 20.1.1.0/24 携带的 RT 属性匹配，所以 PE 1 接收该路由、将其引入到 VPN 实例 1 并添加到 VPN 实例 1 的 IP 路由表中。

(8) PE 1 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 20.1.1.0/24 发布给 CE 1。

4. 报文转发过程

图15 网络边界设备部署 VPN 实例方式中 CE 1 到 CE 2 报文转发过程



如图 15 所示，以 SRv6 BE 转发方式为例，私网报文从 CE 1 发往 CE 2 的过程为：

(1) CE 1 在私网数据上封装 IPv4 报文头，源地址为 10.1.1.1，目的地址为 20.1.1.1，形成私网报文发送给 PE 1。

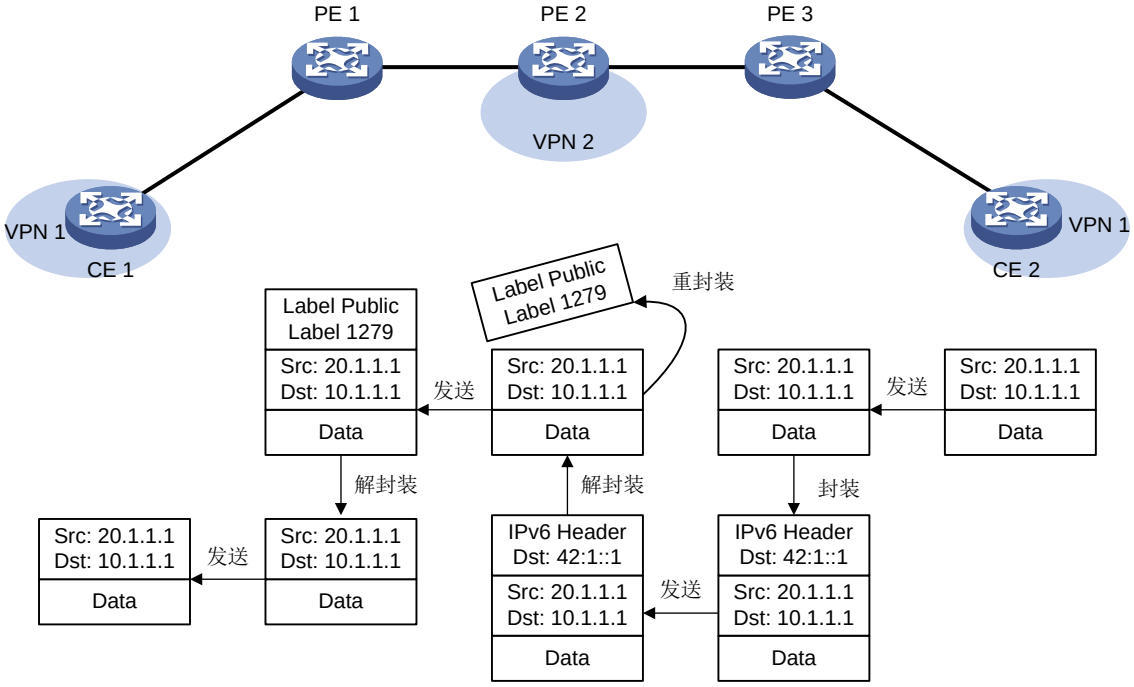
(2) PE 1 收到私网报文后，查找 VPN 实例 1 的 IP 路由表，查找到前缀为 20.1.1.0/24 的路由，找到该路由对应的私网标签 1279，然后为私网报文封装内层私网标签 1279 以及外层公网标签，将报文沿着标签转发路径转发给 PE 2。

(3) PE 2 收到私网报文后，将外层公网标签弹出，根据内层私网标签 1279 判断报文属于 VPN 实例 2，然后在 VPN 实例 2 中查找路由表。PE 2 查找到前缀为 20.1.1.0/24 的路由，找到该路由对应的 SRv6 SID 43:1::1，然后将私网报文的标签全部弹出后重新为其封装外层 IPv6 报文头，封装的外层 IPv6 报文头的目的 IPv6 地址为该 SRv6 SID。

(4) PE 2 根据重新封装后的私网报文外层目的 IPv6 地址查找公网 IP 路由表，找到 PE 3 发布的 Locator 网段路由，然后根据该路由的指导将私网报文发送给 PE 3。

(5) PE 3 收到私网报文后，发现外层目的 IPv6 地址是本地的 End.DT4 SID，然后解封装该报文的外层 IPv6 头，在 End.DT4 SID 43:1::1 对应的 VPN 实例 1 中查找 IP 路由表，找到来自 CE 2 的前缀为 20.1.1.0/24 的路由，根据该路由将私网报文转发给 CE 2。

图16 网络边界设备部署 VPN 实例方式中 CE 2 到 CE 1 报文转发过程



如图 16 所示，以 SRv6 BE 转发方式为例，私网报文从 CE 2 发往 CE 1 的过程为：

(1) CE 2 在私网数据上封装 IPv4 报文头，源地址为 20.1.1.1，目的地址为 10.1.1.1，形成私网报文发送给 PE 3。

(2) PE 3 收到私网报文后，查找 VPN 实例 1 的 IP 路由表，查找到前缀为 10.1.1.0/24 的路由，找到该路由对应的 SRv6 SID 42:1::1，然后为私网报文封装外层 IPv6 报文头，封装的外层 IPv6 报文头的目的 IPv6 地址为该 SRv6 SID。

(3) PE 3 根据重新封装后的私网报文外层目的 IPv6 地址查找公网 IP 路由表，找到 PE 2 发布的 Locator 网段路由，然后根据该路由的指导将私网报文发送给 PE 2。

(4) PE 2 收到私网报文后，发现外层目的 IPv6 地址是本地的 End.DT4 SID，然后解封装该报文的外层 IPv6 头，在 End.DT4 SID 42:1::1 对应的 VPN 实例 2 中查找 IP 路由表。PE 2 找到来自 PE 1 的前缀为 10.1.1.0/24 的路由，根据该路由携带的私网标签 1279，为私网报文封装内层私网标签 1279 以及外层公网标签，将报文沿着标签转发路径转发给 PE 1。

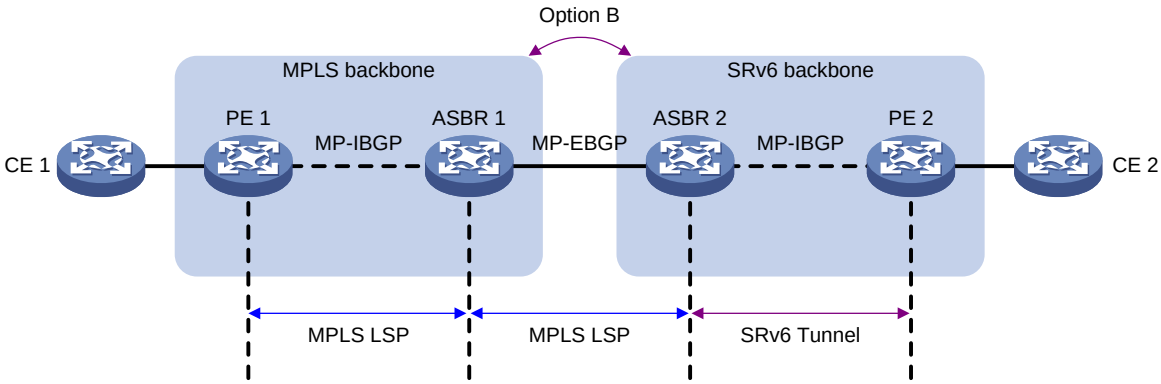
(5) PE 1 收到私网报文后，将外层公网标签弹出，根据内层私网标签 1279 判断报文属于 VPN 实例 1，然后在 VPN 实例 1 中查找路由表。PE 1 查找到来自 CE 1 的前缀为 10.1.1.0/24 的路由，然后根据该路由的指导将私网报文的标签全部弹出后转发给 CE 1。

3.3.2 Option B 跨域方式的工作机制

1. 组网模型

如图 17 所示，PE 1 和 ASBR 1 之间部署 MPLS L3VPN 网络，建立 IBGP VPNv4 会话；ASBR 1 和 ASBR 2 之间部署 MPLS L3VPN 网络，建立 EBGP VPNv4 会话；ASBR 2 和 PE 2 之间部署 IP L3VPN over SRv6 网络，建立 IBGP VPNv4 会话。部署 SRv6 与 MPLS 网络互通后，CE 1 和 CE 2 连接的私网用户之间才能互通。

图17 Option B 跨域方式组网



互通主要在 MPLS 网络和 SRv6 网络的边界设备 ASBR 2 上实现，下面将从路由发布和报文转发两个方面来介绍 MPLS 网络和 SRv6 网络互通的工作机制。

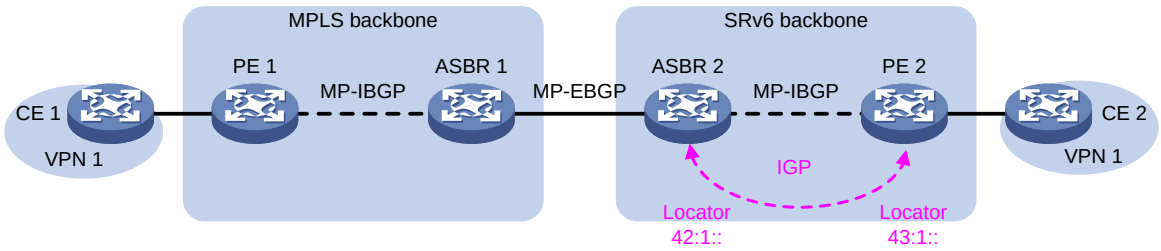
2. End.T SID

SRv6 网络与 MPLS 网络互通的 Option B 跨域方式需要借助一类特殊 SID——End.T SID 来实现。SRv6 网络中的 ASBR 会为来自 MPLS 网络的私网路由分配并添加 End.T SID。End.T SID 对应的转发动作为剥掉外层 IPv6 头，并根据 End.T SID 查找 IPv6 FIB 表转发报文。

3. Locator 网段路由发布

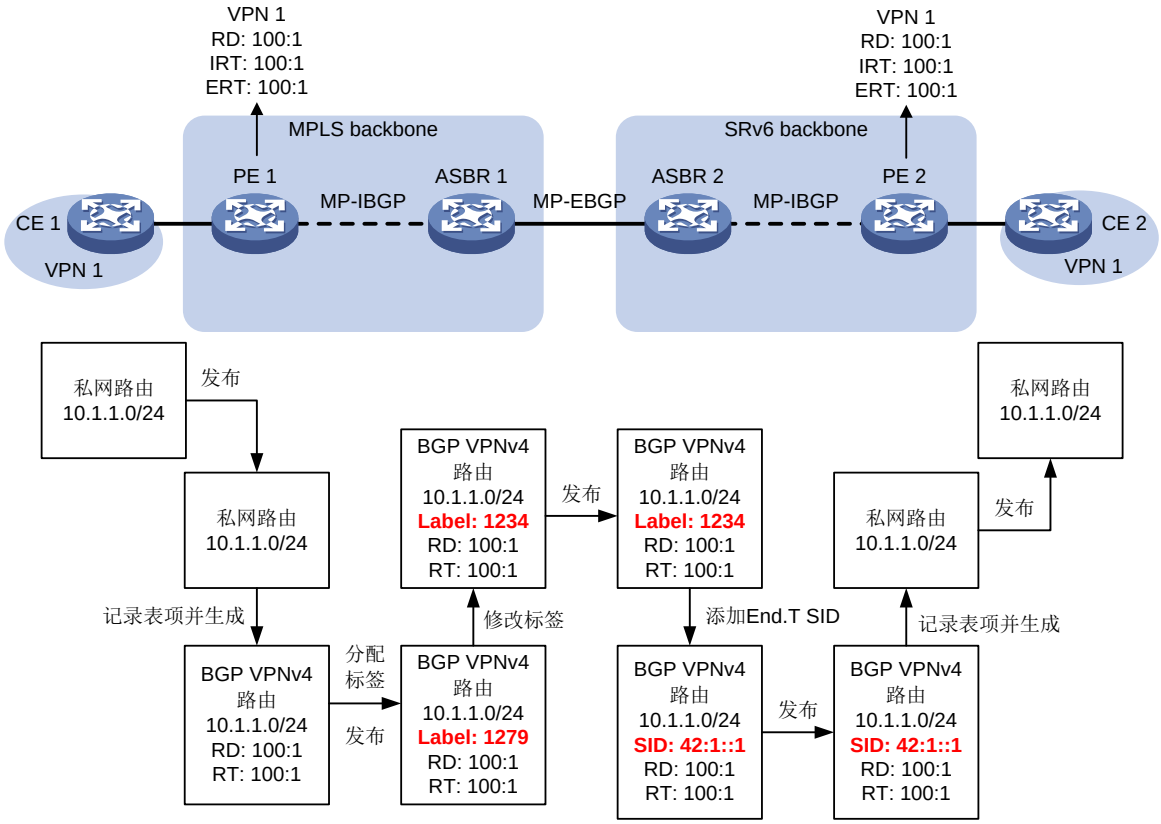
如图 18 所示，ASBR 2 和 PE 2 需要通过 IGP 协议向对方发布各自的 Locator 网段路由信息。ASBR 2 发布的 Locator 网段路由前缀为 42:1::/64，PE 2 发布的 Locator 网段路由前缀为 43:1::/64。

图18 Option B 方式中 Locator 网段路由发布



4. 私网路由发布

图19 Option B 跨域方式中 CE 1 到 CE 2 路由发布过程



如图 19所示，私网路由从 CE 1 发往 CE 2 的过程为：

(1) CE 1 通过 IGP、BGP 或静态路由等方式将本地的私网路由 10.1.1.0/24 发布给 PE 1。

(2) PE 1 接收到私网路由信息后，将其添加到 VPN 实例 1 的 IP 路由表中，并为其增加 RD 100:1 和 RT 100:1 属性，生成 BGP VPNv4 路由。

(3) PE 1 为私网路由信息生成的 BGP VPNv4 路由分配私网标签 1279，并将该路由通过 IBGP VPNv4 会话发布给 ASBR 1。私网标签 1279 绑定的 VPN 实例为 VPN 实例 1。

(4) ASBR 1 接收 PE 1 发布的 BGP VPNv4 路由 10.1.1.0/24，将路由的下一跳修改为自身的地址，并为路由分配新的私网标签 1234 来替代原私网标签 1279。

(5) ASBR 1 在本地的 ILM 表项信息中增加 BGP VPNv4 路由 10.1.1.0/24 新旧标签的映射表项：入标签为新标签 1234 时，将其替换为旧标签 1279 并发往 PE 1。

(6) ASBR 1 将携带私网标签 1234 的 BGP VPNv4 路由 10.1.1.0/24 通过 EBGP VPNv4 会话发布给 ASBR 2。

(7) ASBR 2 从 ASBR 1 接收到 BGP VPNv4 路由后，将路由的下一跳修改为自身的地址，为路由分配 End.T SID 42:1::1，并生成如下所示的 IPv6 FIB 表项：

```
Destination: 42:1::1                                    Prefix length: 128
Nexthop    : ::                                         Flags: UGH
Time stamp : 0x11                                       Label: 1234
Interface  : NULL0                                      Token: 1
```

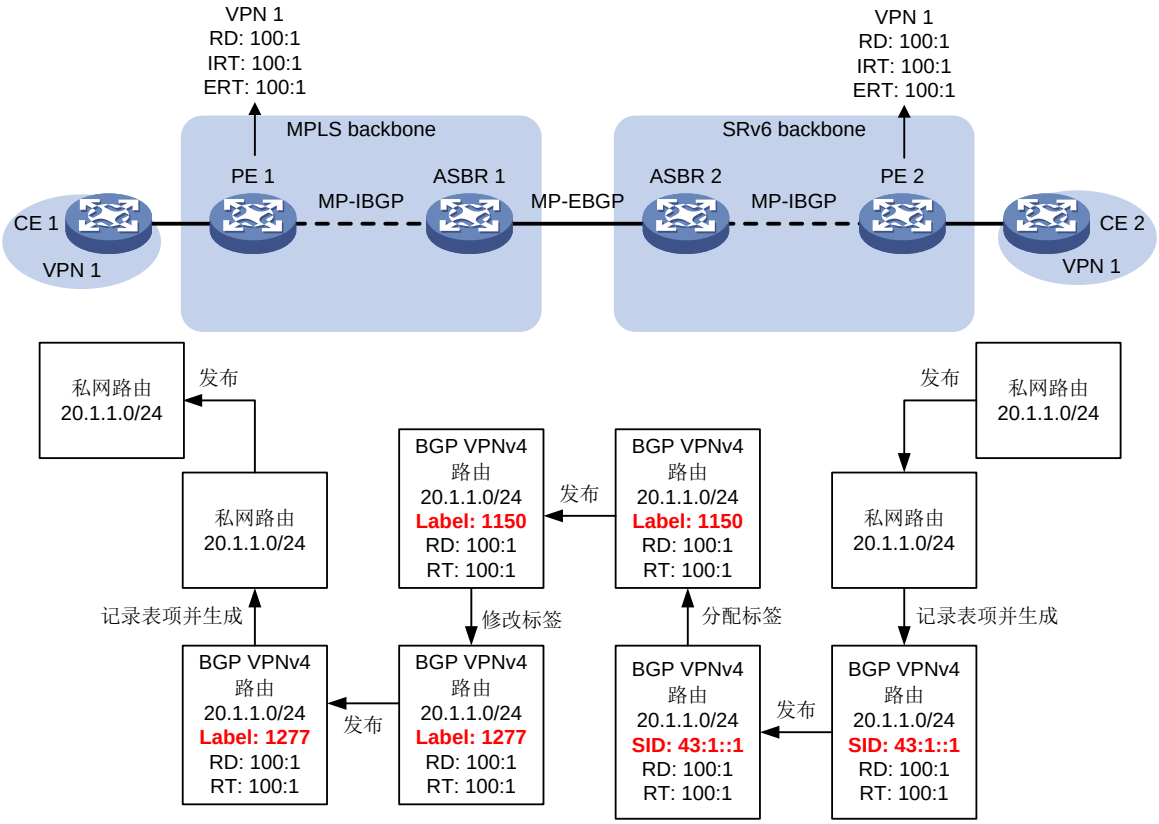
可以看到，通过 IPv6 FIB 表项，路由携带的 End.T SID 42:1::1 与路由携带的私网标签 1234 产生了关联。根据此标签以及私网标签 1234 对应的 NHLFE 表项，目的地址为 End.T SID 42:1::1 的报文在 ASBR 2 上会被封装上私网标签 1234 发往 ASBR 1。

(8) ASBR 2 将添加了 End.T SID 42:1::1 后的 BGP VPNv4 路由 10.1.1.0/24 通过 IBGP VPNv4 会话发布给 PE 2。

(9) PE 2 本地 VPN 实例 1 配置的 IRT 100:1 与 BGP VPNv4 路由 10.1.1.0/24 携带的 RT 属性匹配，所以 PE 2 接收该路由、将其引入到 VPN 实例 1 并添加到 VPN 实例 1 的 IP 路由表中。

(10) PE 2 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 10.1.1.0/24 发布给 CE 2。

图20 Option B 跨域方式中 CE 2 到 CE 1 的路由发布过程



如图 20 所示，私网路由从 CE 2 发往 CE 1 的过程为：

(1) CE 2 通过 IGP、BGP 或静态路由等方式将本地的私网路由 20.1.1.0/24 发布给 PE 2。

(2) PE 2 接收到私网路由信息后，将其添加到 VPN 实例 1 的 IP 路由表中，并为其增加 RD 100:1 和 RT 100:1 属性，生成 BGP VPNv4 路由。

(3) PE 2 为私网路由信息生成的 BGP VPNv4 路由在 VPN 实例 1 内申请 SRv6 SID 43:1::1（SRv6 SID 类型可以为 End.DT4 SID、End.DT6 SID、End.DT46 SID、End.DX4 SID 或 End.DX6 SID，本例中的 SRv6 SID 为 End.DT4 SID），并将该路由通过 IBGP VPNv4 会话发布给 ASBR 2。

(4) ASBR 2 接收 PE 2 发布的 BGP VPNv4 路由 20.1.1.0/24，将路由的下一跳修改为自身的地址，为路由分配私网标签 1150，并在本地的 ILM 表项信息中增加私网标签和 SRv6 SID 的如下映射表项：

```
InLabel Oper    VRF   Flag SwapInfo  Forwarding Info
```

```
------------------------------------------------------------------------------
1150       SWAP      0         NA        43:1::1
                                         GE0/0/1
                                         FE80::CA2:4BFF:FE97:407
```

可以看到，通过 ILM 表项，ASBR 2 将路由携带的 End.DT4 SID 43:1::1 与为该路由分配的私网标签 1150 产生了关联，携带私网标签 1150 的报文在 ASBR 2 上会被重新封装上目的 IPv6 地址为 End.DT4 SID 43:1::1 的 SRv6 封装发往 PE 2。

(5) ASBR 2 将携带私网标签 1150 的 BGP VPNv4 路由 20.1.1.0/24 通过 EBGP VPNv4 会话发布给 ASBR 1。

(6) ASBR 1 从 ASBR 2 接收到 BGP VPNv4 路由后，将路由的下一跳修改为自身的地址，并为路由分配新的私网标签 1277 来替代原私网标签 1150。

(7) ASBR 1 在本地的 ILM 表项信息中增加 BGP VPNv4 路由 20.1.1.0/24 新旧标签的映射表项：入标签为新标签 1277 时，将其替换为旧标签 1150 并发往 ASBR 2。

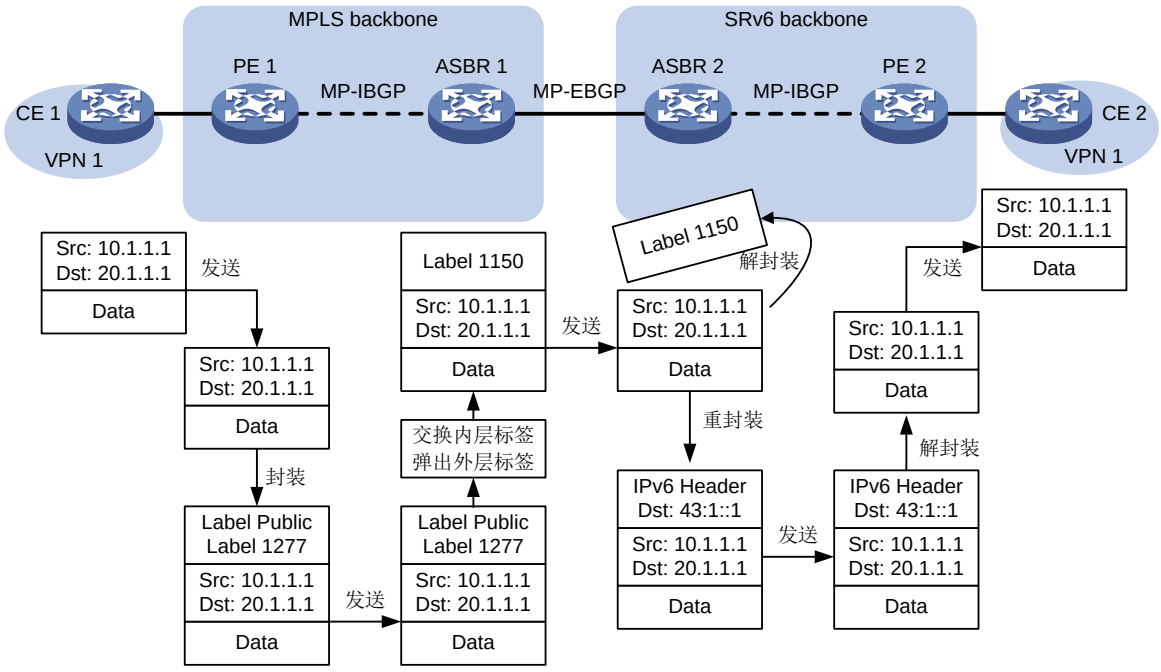
(8) ASBR 1 将携带私网标签 1277 的 BGP VPNv4 路由 20.1.1.0/24 通过 IBGP VPNv4 会话发布给 PE 1。

(9) PE 1 本地 VPN 实例 1 配置的 IRT 100:1 与 BGP VPNv4 路由 20.1.1.0/24 携带的 RT 属性匹配，所以 PE 1 接收该路由、将其引入到 VPN 实例 1 并添加到 VPN 实例 1 的 IP 路由表中。

(10) PE 1 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 20.1.1.0/24 发布给 CE 1。

5. 报文转发过程

图21 Option B 跨域方式中 CE 1 到 CE 2 报文转发过程



如图 21 所示，以 SRv6 BE 转发方式为例，私网报文从 CE 1 发往 CE 2 的过程为：

(1) CE 1 在私网数据上封装 IPv4 报文头，源地址为 10.1.1.1，目的地址为 20.1.1.1，形成私网报文发送给 PE 1。

(2) PE 1 收到私网报文后，查找 VPN 实例 1 的 IP 路由表，查找到前缀为 20.1.1.0/24 的路由，找到该路由对应的私网标签 1277，然后为私网报文封装内层私网标签 1277 以及外层公网标签，将报文沿着标签转发路径转发给 ASBR 1。

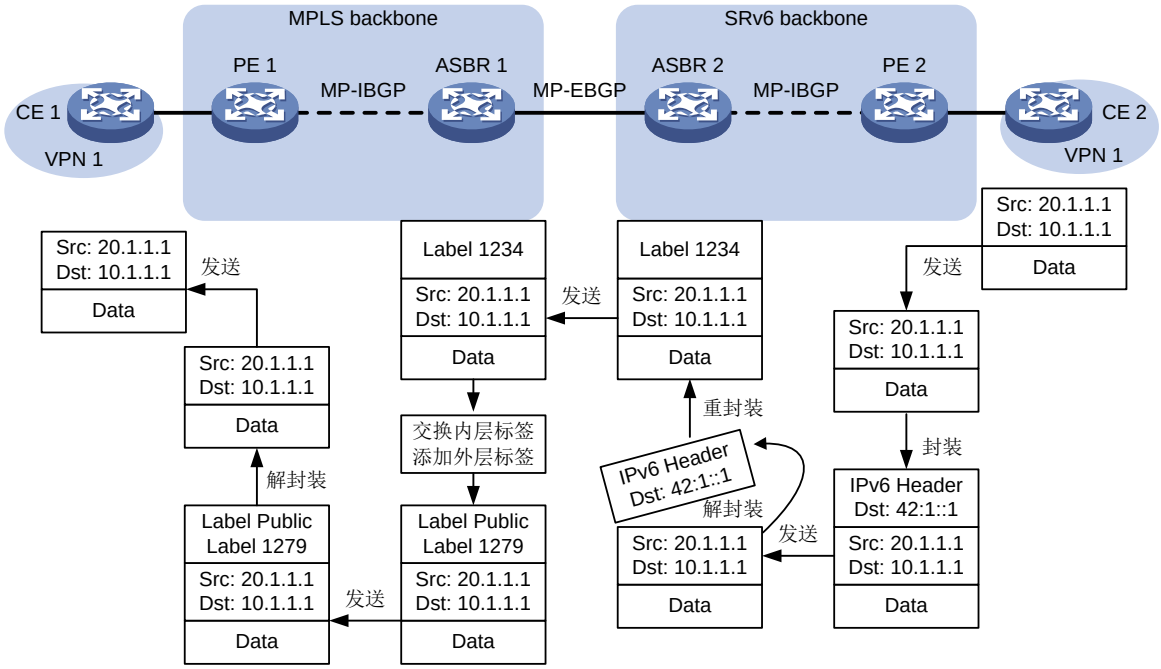
(3) ASBR 1 收到私网报文后，将外层公网标签弹出，根据内层私网标签 1277 对应的 ILM 表项将该私网标签替换为私网标签 1150。由于 ASBR 1 和 ASBR 2 只存在一跳标签转发路径，ASBR 1 将仅携带一层私网标签 1150 的私网报文直接转发给 ASBR 2。

(4) ASBR 2 收到私网报文后，根据私网标签 1150 对应的 ILM 表项，将私网标签 1150 弹出后重新为其封装外层 IPv6 报文头，封装的外层 IPv6 报文头的目的 IPv6 地址为私网标签 1150 映射的 End.DT4 SID 43:1::1。

(5) ASBR 2 根据重新封装后的私网报文外层目的 IPv6 地址查找公网 IP 路由表，找到 PE 2 发布的 Locator 网段路由，然后根据该路由的指导将私网报文发送给 PE 2。

(6) PE 2 收到私网报文后，发现外层目的 IPv6 地址是本地的 End.DT4 SID，然后解封装该报文的外层 IPv6 头，在 End.DT4 SID 43:1::1 对应的 VPN 实例 1 中查找 IP 路由表，找到来自 CE 2 的前缀为 20.1.1.0/24 的路由，根据该路由将私网报文转发给 CE 2。

图22 Option B 跨域方式中 CE 2 到 CE 1 报文转发过程



如图 22 所示，以 SRv6 BE 转发方式为例，私网报文从 CE 2 发往 CE 1 的过程为：

(1) CE 2 在私网数据上封装 IPv4 报文头，源地址为 20.1.1.1，目的地址为 10.1.1.1，形成私网报文发送给 PE 2。

(2) PE 2 收到私网报文后，查找 VPN 实例 1 的 IP 路由表，查找到前缀为 10.1.1.0/24 的路由，找到该路由对应的 SRv6 SID 42:1::1，然后为私网报文封装外层 IPv6 报文头，封装的外层 IPv6 报文头的目的 IPv6 地址为该 SRv6 SID。

(3) PE 2 根据重新封装后的私网报文外层目的 IPv6 地址查找公网 IP 路由表，找到 ASBR 2 发布的 Locator 网段路由，然后根据该路由的指导将私网报文发送给 ASBR 2。

(4) ASBR 2 收到私网报文后，发现外层目的 IPv6 地址是本地的 End.T SID，然后解封装该报文的外层 IPv6 头，在本地查找 IPv6 FIB 表。根据查找到的 IPv6 FIB 表项，ASBR 2 为私网报文封装内层私网标签 1234。由于 ASBR 1 和 ASBR 2 只存在一跳标签转发路径，ASBR 2 将仅携带一层私网标签 1234 的私网报文直接转发给 ASBR 1。

(5) ASBR 1 收到私网报文后，根据私网标签 1234 对应的 ILM 表项将该私网标签替换为私网标签 1279，并为私网报文额外添加外层公网标签后，将私网报文沿着标签转发路径转发给 PE 1。

(6) PE 1 收到私网报文后，将外层公网标签弹出，根据内层私网标签 1279 判断报文属于 VPN 实例 1，然后在 VPN 实例 1 中查找路由表。PE 1 查找到来自 CE 1 的前缀为 10.1.1.0/24 的路由，然后根据该路由的指导将私网报文的标签全部弹出后转发给 CE 1。

4 EVPN L3VPN over SRv6 与 IP L3VPN over SRv6 互通

4.1 概述

数据平面统一使用 SRv6 来承载业务的 IP 承载网或骨干网中，某个 AS 域内采用 BGP EVPN 来传递私网路由，而另一 AS 域内采用 BGP VPNv4 或 BGP VPNv6 来传递私网路由，在 AS 域间要实现私网业务端到端的互通，这种场景就是 EVPN L3VPN over SRv6 网络与 IP L3VPN over SRv6 网络互通。

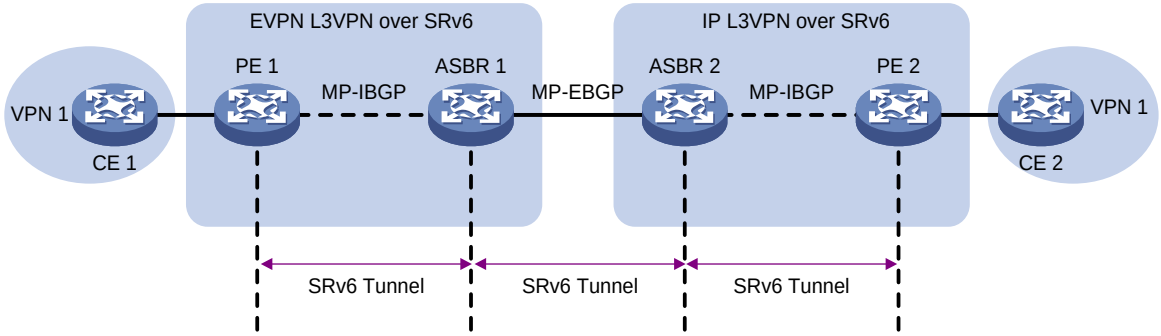
在该场景中，ASBR 上需要将 BGP EVPN 的 IP 前缀路由和 BGP VPNv4/BGP VPNv6 路由相互转换后，发布到对方域内。因此，管理员需要配置 **advertise l3vpn route** 和 **advertise evpn route** 命令来重发布路由。

4.2 工作机制

4.2.1 组网模型

如图 23 所示，PE 1 和 ASBR 1 属于同一个 AS 域 100，两者之间部署 EVPN L3VPN over SRv6 承载 CE 1 上的 VPN1 的私网业务。PE 2 和 ASBR 2 之间属于同一个 AS 域 200，两者之间部署 IP L3VPN over SRv6 网络承载 CE 2 上的 VPN1 的私网业务。EVPN L3VPN over SRv6 与 IP L3VPN over SRv6 网络互通后，CE 1 和 CE 2 连接的私网用户之间可以端到端互通。

图23 EVPN L3VPN over SRv6 网络与 IP L3VPN over SRv6 网络互通组网



异构网络的互通主要在边界设备 ASBR 1 和 ASBR 2 上实现，下面将从路由发布和报文转发两个方面来介绍 EVPN L3VPN over SRv6 网络与 IP L3VPN over SRv6 网络互通的工作机制。

4.2.2 路由发布过程

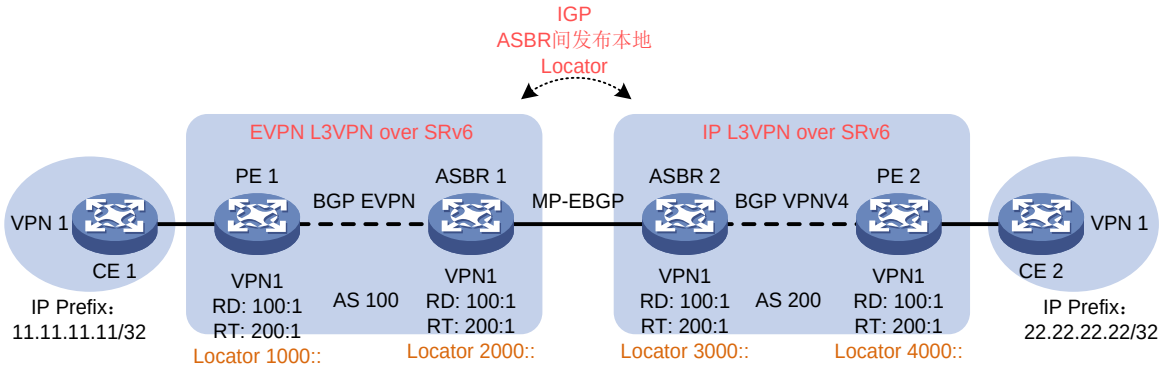
1. 路由的规划和准备

如图 24 所示，AS 100 域内为 EVPN L3VPN over SRv6 网络，AS 200 域内为 IP L3VPN over SRv6 网络，在域内分别采用 IGP 协议实现路由互通，且域内 PE 和 ASBR 的 SRv6 Locator 相互可达。

在 ASBR 1 和 ASBR 2 之间另起 IGP 进程，使 ASBR 1 和 ASBR 2 间用于建立 MP-EBGP 邻居的 Loopback 口地址可达，且 ASBR 1 和 ASBR 2 需要互相发布本地 SRv6 Locator，使域间 MP-EBGP 发布的路由可以通过 SRv6 隧道承载。

PE 1、ASBR 1、ASBR 2 和 PE 2 均部署 VPN 实例 1，CE 1 和 CE 2 通过 VPN 实例 1 分别接入 PE 1 和 PE 2。

图24 组网规划示意图



2. 私网路由从 CE 1 发往 CE 2 的过程

如图 25 所示，私网路由从 CE 1 发往 CE 2 的过程为：

(1) CE 1 通过 IGP、BGP 或静态路由等方式将本地的私网路由 11.11.11.11/32 发布给 PE 1。

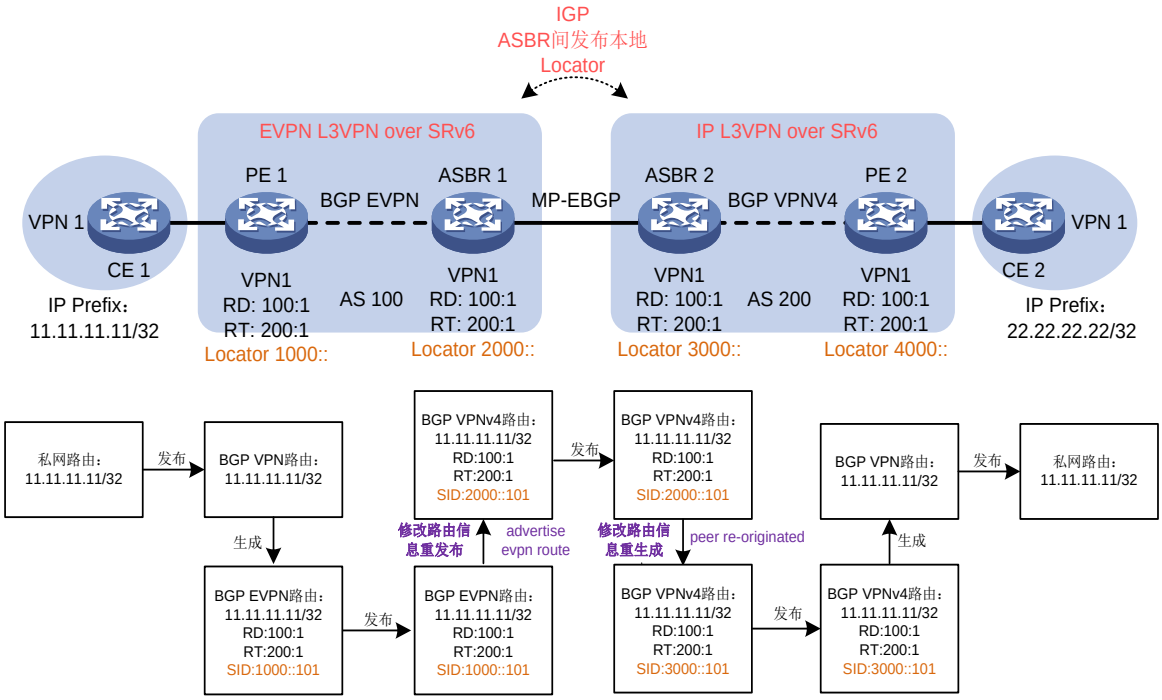
(2) PE 1 接收到私网路由信息后，根据连接 CE 1 接口所绑定的 VPN 实例，将其添加到 VPN 实例 1 的 IP 路由表中，并生成 BGP VPN 路由表。PE 1 和 ASBR 1 建立 BGP EVPN 邻居，因此，PE 1 为 BGP VPN 路由表中的私网路由添加 RD 值 100:1，扩展团体属性 RT 200:1，并从本地 SRv6 Locator 段中为私网路由分配 SRv6 SID 1000::101 后，生成 BGP EVPN 的 IP 前缀路由，再将 IP 前缀路由发往其 BGP EVPN 邻居 ASBR 1，修改下一跳为 PE 1 用于建立 BGP EVPN 邻居的 Loopback 口地址。

(3) ASBR 1 接收到从 BGP EVPN 邻居发布的 IP 前缀路由后，根据 IP 前缀路由的 RT 值过滤路由，并将私网路由学习到本地 VPN 路由表中。基于 VPN 路由表的私网路由重新生成 BGP VPNv4 路由，此时，先删除 IP 前缀路由的 SRv6 SID，再从本地 SRv6 Locator 段中重新为私网路由分配 SRv6 SID 2000::101。ASBR 1 与 ASBR 2 建立 MP-EBGP 邻居，ASBR 1 向 ASBR 2 发布 BGP VPNv4 路由，并修改下一跳为 ASBR 1 用于建立 MP-EBGP 邻居的 Loopback 口地址。

(4) ASBR 2 接收到从 ASBR 1 发布的 BGP VPNv4 路由后，根据 BGP VPNv4 路由的 RT 值过滤路由，并将私网路由学习到本地 VPN 路由表中。基于 VPN 路由表的私网路由重新生成 BGP VPNv4 路由，此时，先删除原 BGP VPNv4 路由的 SRv6 SID 2000::101，再从本地 SRv6 Locator 段中重新为私网路由分配 SRv6 SID 3000::101。ASBR 2 与 PE 2 建立 BGP VPNv4 邻居，ASBR 2 向 PE 2 发布 BGP VPNv4 路由，并修改下一跳为 ASBR 2 用于建立 BGP VPNv4 邻居的 Loopback 口地址。

(5) PE 2 接收到 BGP VPNv4 路由后，根据 BGP VPNv4 路由的 RT 值过滤路由，并将私网路由学习到本地 BGP VPN 路由表，再加入 VPN 路由表中。PE 2 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 11.11.11.11/32 发布给 CE 2。

图25 私网路由从 EVPN L3VPN over SRv6 网络发布到 IP L3VPN over SRv6 网络示意图



3. 私网路由从 CE 2 发往 CE 1 的过程

如图 26 所示，私网路由从 CE 2 发往 CE 1 的过程为：

(1) CE 2 通过 IGP、BGP 或静态路由等方式将本地的私网路由 22.22.22.22/32 发布给 PE 2。

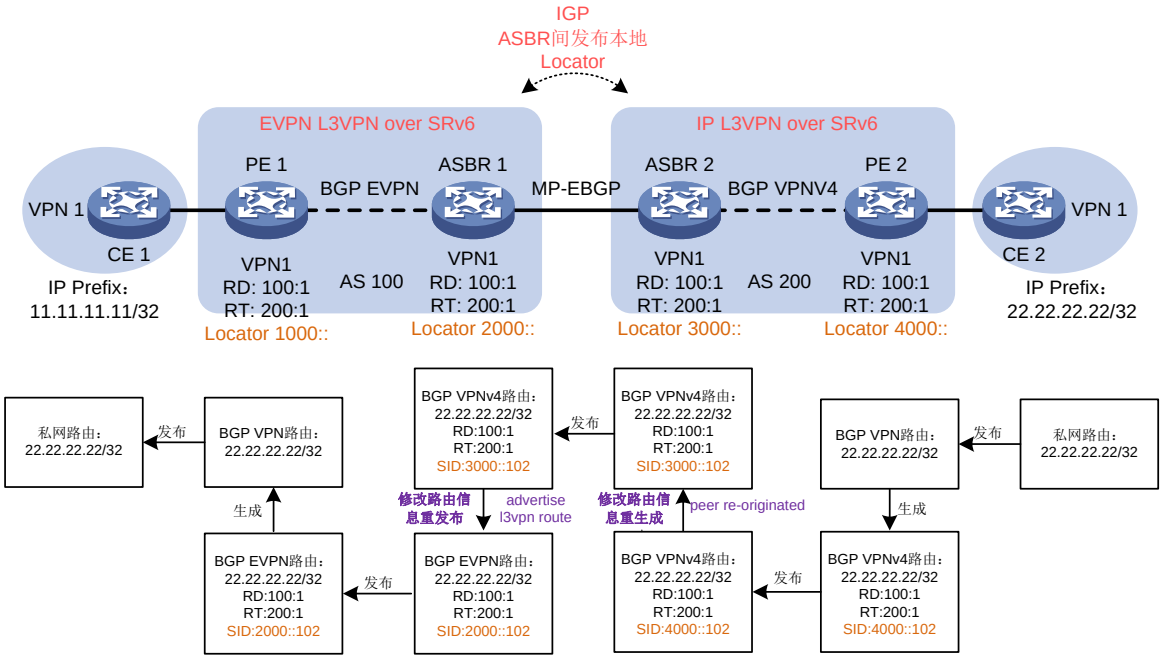
(2) PE 2 接收到私网路由信息后，根据连接 CE 2 接口所绑定的 VPN 实例，将其添加到 VPN 实例 1 的 IP 路由表中，并生成 BGP VPN 路由表。PE 2 和 ASBR 2 建立 BGP VPNv4 邻居，因此，PE 2 为 BGP VPN 路由表中的私网路由添加 RD 值 100:1，扩展团体属性 RT 200:1，并从本地 SRv6 Locator 段中为私网路由分配 SRv6 SID 4000::102 后，生成 BGP VPNv4 路由，再将 BGP VPNv4 路由发往其 BGP VPNv4 邻居 ASBR 2，修改下一跳为 PE 2 用于建立 BGP VPNv4 邻居的 Loopback 口地址。

(3) ASBR 2 接收到 BGP VPNv4 路由后，根据 RT 值过滤路由，并将私网路由学习到本地 VPN 路由表中。基于 VPN 路由表的私网路由重新生成 BGP VPNv4 路由，此时，先删除原 BGP VPNv4 路由的 SRv6 SID，再从本地 SRv6 Locator 段中重新为私网路由分配 SRv6 SID 3000::102。ASBR 1 与 ASBR 2 建立 MP-EBGP 邻居，ASBR 2 向 ASBR 1 发布 BGP VPNv4 路由，并修改下一跳为 ASBR 2 用于建立 MP-EBGP 邻居的 Loopback 口地址。

(4) ASBR 1 接收到从 ASBR 2 发布的 BGP VPNv4 路由后，根据 BGP VPNv4 路由的 RT 值过滤路由，并将私网路由学习到本地 VPN 路由表中。基于 VPN 路由表的私网路由重新生成 BGP EVPN 路由，此时，先删除原 BGP VPNv4 路由的 SRv6 SID 3000::102，再从本地 SRv6 Locator 段中重新为私网路由分配 SRv6 SID 2000::102。ASBR 1 与 PE 1 建立 BGP EVPN 邻居，ASBR 1 向 PE 1 发布 BGP EVPN 的 IP 前缀路由，并修改下一跳为 ASBR 1 用于建立 BGP EVPN 邻居的 Loopback 口地址。

(5) PE 1 接收到 BGP EVPN 的 IP 前缀路由后，根据 RT 值过滤路由，并将私网路由学习到本地 BGP VPN 路由表，再加入 VPN 路由表中。PE 1 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 22.22.22.22/32 发布给 CE 1。

图26 私网路由从 IP L3VPN over SRv6 网络发布到 EVPN L3VPN over SRv6 网络示意图



4.2.3 报文转发过程

如图 27 所示，完成路由学习后，以私网报文从 EVPN L3VPN over SRv6 网络发往 IP L3VPN over SRv6 网络的转发过程为例，端到端的报文转发过程为：

(1) 在 CE 1 上源地址为 11.11.11.11/32，目的地址为 22.22.22.22/32 的私网报文查询 FIB 转发表项，发现出接口为 CE 1 与 PE 1 的互联接口，将报文发送给 PE 1。

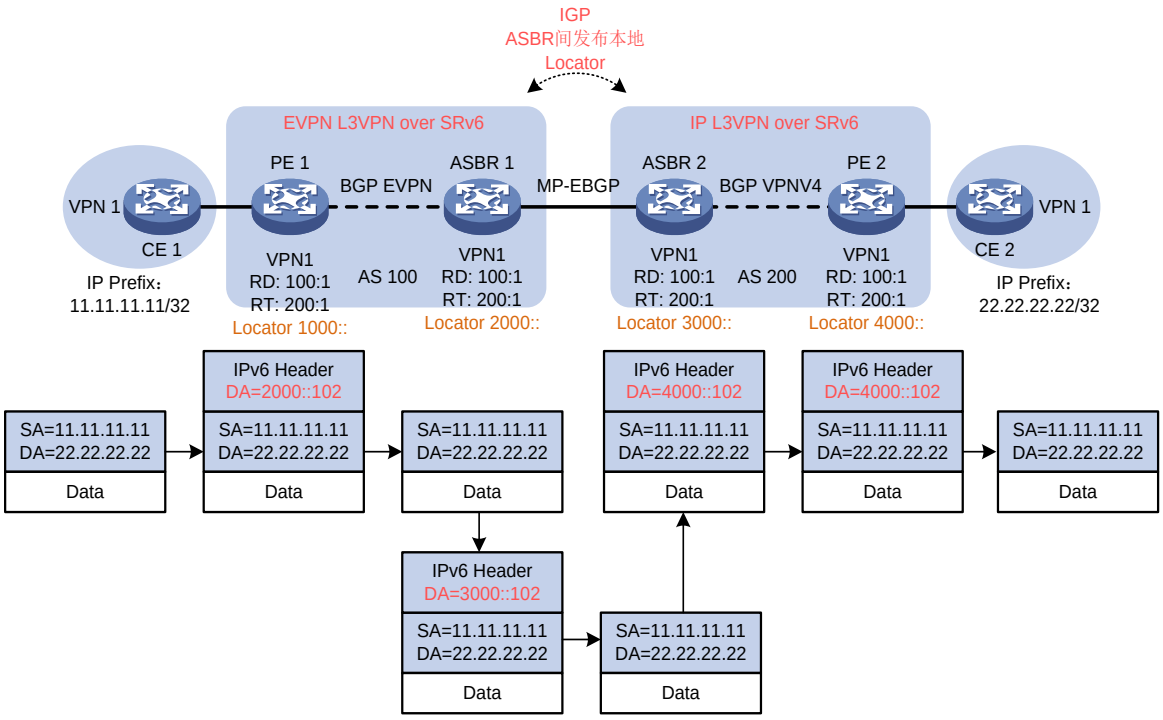
(2) PE 1 从绑定了 VPN 实例 1 的私网接口接收到报文后，查找 VPN 实例 1 的 IP 路由表，私网路由前缀为 22.22.22.22/32，找到该路由的下一跳为 ASBR 1 为私网路由 22.22.22.22/32 分配的 SRv6 SID 2000::102，因此，私网报文通过 SRv6 公网隧道转发到 ASBR 1，PE 1 为私网报文封装外层 IPv6 报文头，目的 IPv6 地址为 SRv6 SID 2000::102。

(3) ASBR 1 接收到目的 IPv6 地址为 SRv6 SID 2000::102，私网路由为 22.22.22.22/32 的报文，发现该 SRv6 SID 2000::102 为本地的 End.DT4 SID，解封装报文的外层 IPv6 头，在 End.DT4 SID 2000::102 对应的 VPN 实例 1 中查找转发信息，报文下一跳为 ASBR 2 为私网路由 22.22.22.22/32 分配的 SRv6 SID 3000::102，因此，私网报文通过 SRv6 公网隧道转发到 ASBR 2，ASBR 1 为私网报文封装外层 IPv6 报文头，目的 IPv6 地址为 SRv6 SID 3000::102。

(4) ASBR 2 接收到目的 IPv6 地址为 SRv6 SID 3000::102，私网路由为 22.22.22.22/32 的报文，发现该 SRv6 SID 3000::102 为本地的 End.DT4 SID，解封装报文的外层 IPv6 头，在 End.DT4 SID 3000::102 对应的 VPN 实例 1 中查找转发信息，报文下一跳为 PE 2 为私网路由 22.22.22.22/32 分配的 SRv6 SID 4000::102，因此，私网报文通过 SRv6 公网隧道转发到 PE 2，ASBR 2 为私网报文封装外层 IPv6 报文头，目的 IPv6 地址为 SRv6 SID 4000::102。

(5) PE 2 接收到目的 IPv6 地址为 SRv6 SID 4000::102，私网路由为 22.22.22.22/32 的报文，发现该 SRv6 SID 4000::102 为本地的 End.DT4 SID，解封装报文的外层 IPv6 头，在 End.DT4 SID 4000::102 对应的 VPN 实例 1 中查找转发信息，报文下一跳为 CE 2 和 PE 2 连接的接口地址，出接口为绑定 VPN 实例 1 的接口，因此，PE 2 将目的地址为 22.22.22.22/32 私网报文转发给 CE 2。

图27 EVPN L3VPN over SRv6 与 IP L3VPN over SRv6 网络互通的报文转发图



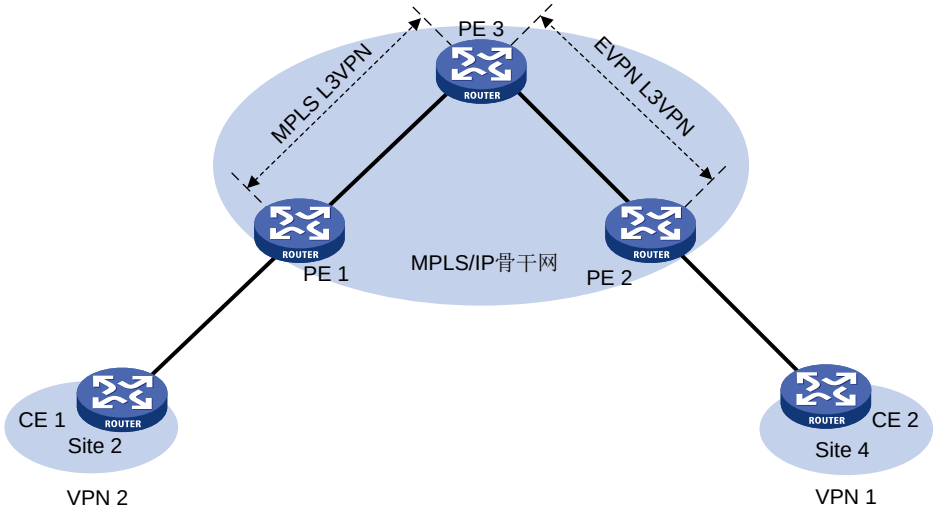
EVPN L3VPN over SRv6 网络和 IP L3VPN over SRv6 网络均采用 SRv6 隧道承载私网业务，私网报文从 EVPN L3VPN over SRv6 网络发往 IP L3VPN over SRv6 网络与报文从 IP L3VPN over SRv6 网络发往 EVPN L3VPN over SRv6 网络的转发过程类似，此处不再赘述。

5 MPLS L3VPN 网络与 EVPN L3VPN 网络互通

5.1 概述

如图 28 所示，将现有的 MPLS L3VPN 网络改造成 EVPN L3VPN 网络的过程中，会存在两种类型网络对接的情况。MPLS L3VPN 网络中，PE 设备之间通过 BGP VPNv4 或 BGP VPNv6 路由传递私网路由信息；EVPN L3VPN 网络中，PE 设备之间通过 BGP EVPN 路由传递私网路由信息。为了实现 MPLS L3VPN 和 EVPN L3VPN 网络之间的互通，边界设备（PE 3）上需要进行 BGP VPNv4（或 BGP VPNv6）与 BGP EVPN 路由的相互转换，以使私网路由能够在两种网络之间传递。这样不同网络之间的互通就能够得以实现。

图28 MPLS L3VPN 与 EVPN L3VPN 对接示意图



5.2 工作机制

5.2.1 私网路由发布

为了实现私网路由在 MPLS L3VPN 和 EVPN L3VPN 网络之间的跨网络传递，边界设备上需要进行如下路由转换：

- 从 MPLS L3VPN 网络接收到 VPNv4 或 VPNv6 路由后，将其转换为 EVPN 的 IP 前缀路由，并发布到 EVPN L3VPN 网络。
- 从 EVPN L3VPN 网络接收到 EVPN 的 IP 前缀路由或携带主机路由信息的 MAC/IP 发布路由后，将其转换为 VPNv4 或 VPNv6 路由，并发布到 MPLS L3VPN 网络。

如图 29所示，以 IPv4 MPLS L3VPN 为例，MPLS L3VPN 与 EVPN L3VPN 互通组网中，私网路由从 PE 1 发布到 PE 2 的过程为：

(1) PE 1 通过 IGP、BGP 或静态路由等方式获得 VPN 实例 1 用户的私网路由 10.1.1.0/24。

(2) PE 1 为本地 VPN 实例 1 的私网路由 10.1.1.0/24 增加 RD 100:1、RT 200:1，并为其分配私网标签 1000，生成 EVPN 的 IP 前缀路由。

(3) PE 1 将 IP 前缀路由 10.1.1.0/24 通过 BGP EVPN 路由会话发布给 Border（PE 3）。

(4) Border 本地 VPN 实例 1 配置的 RT 200:1 与 IP 前缀路由 10.1.1.0/24 携带的 RT 属性匹配，所以 Border 接收该路由，并将其添加到 VPN 实例 1 的路由表中。

(5) Border 将添加到 VPN 实例 1 的路由 10.1.1.0/24 重生成为 BGP VPNv4 路由，并为该路由分配私网标签 2000。Border 在路由原有私网标签 1000 和新分配的私网标签 2000 之间建立关联。

说明

Border 将 IP 前缀路由重生成为 VPNv4 路由时，既可以保留 IP 前缀路由的 RD 和 RT，也可以将 RD 和 RT 修改为本地 VPN 实例的 RD 和 RT。通过配置来决定是否修改路由的 RD 和 RT。

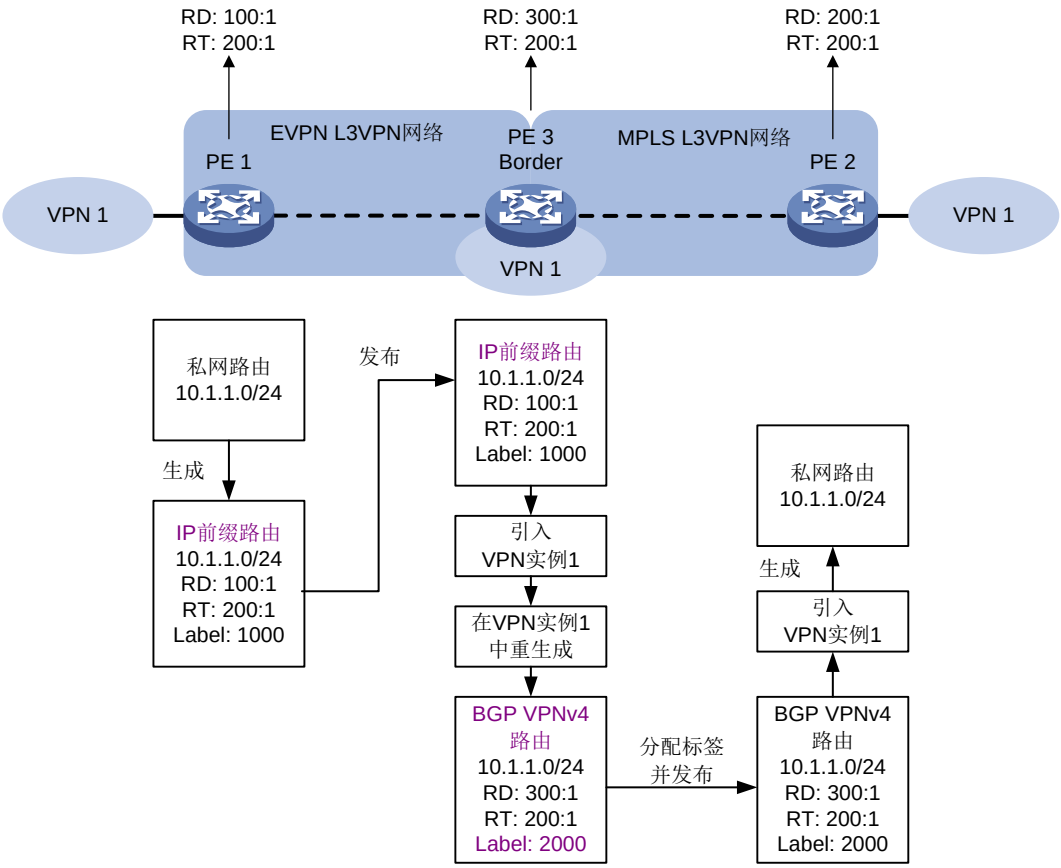
(6) Border 将重生成的 VPNv4 路由通过 BGP VPNv4 会话发布给 PE 2。

(7) PE 2 本地 VPN 实例 1 配置的 RT 200:1 与 BGP VPNv4 路由 10.1.1.0/24 携带的 RT 属性匹配，所以 PE 2 接收该路由，并将其添加到 VPN 实例 1 的路由表中。

(8) PE 2 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 10.1.1.0/24 发布给 VPN 实例 1 的用户。

私网路由从 PE 2 发往 PE 1 的过程与上述过程类似，本文不再赘述。

图29 MPLS L3VPN 和 EVPN L3VPN 网络互通的私网路由发布过程图



5.2.2 报文转发过程

如图 30 所示，私网报文从 PE 2 连接的 VPN 实例 1 内用户发往 PE 1 连接的 VPN 实例 1 内用户的过程为：

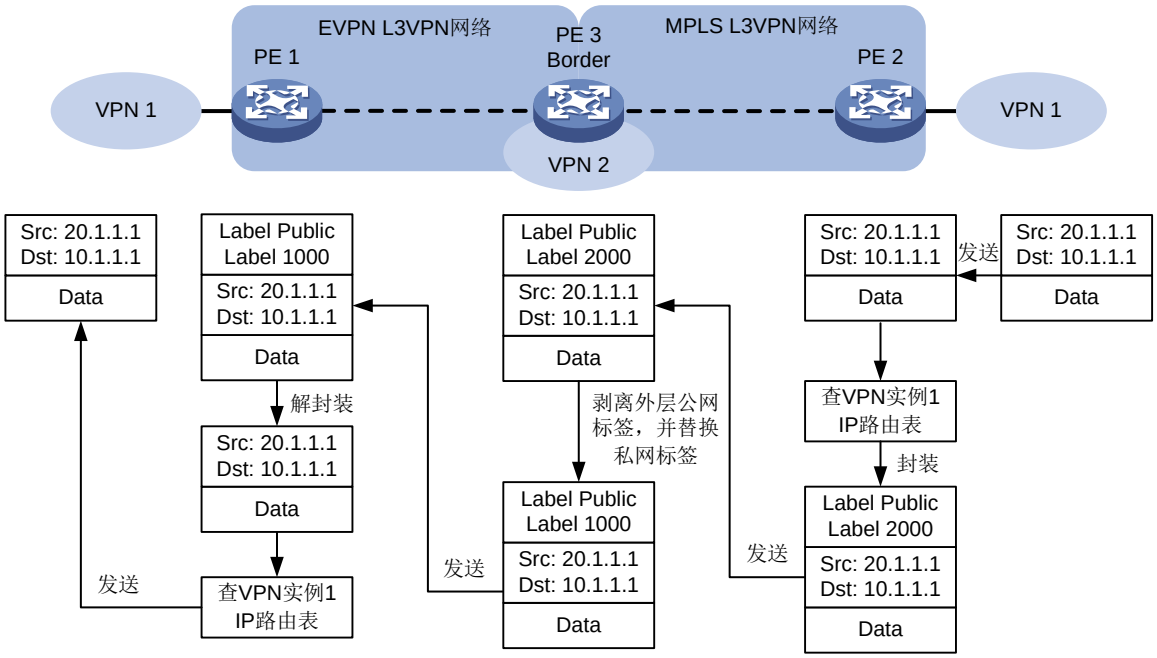
(1) VPN 实例 1 内用户将源地址为 20.1.1.1，目的地址为 10.1.1.1 的私网报文发送给 PE 2。

(2) PE 2 收到私网报文后，查找 VPN 实例 1 的 IP 路由表，找到前缀为 10.1.1.0/24 的路由，该路由对应的私网标签为 2000。PE 2 为私网报文封装私网标签和公网标签，将其转发给 Border。

(3) Border 接收到报文后，剥离报文的外层公网标签，将私网标签 2000 替换为与其关联的私网标签 1000，并添加新的公网标签后，将封装后的报文转发给 PE 1。

(4) PE 1 接收到报文后，剥离公网标签和私网标签，根据私网标签 1000 判断报文所属的 VPN 实例，并在该 VPN 实例内查找 IP 路由表，将其发送给 VPN 实例 1 内的用户。

图30 MPLS L3VPN 和 EVPN L3VPN 网络互通的报文转发过程图



6 MPLS L3VPN/EVPN L3VPN 与 EVPN VXLAN 互通

6.1 概述

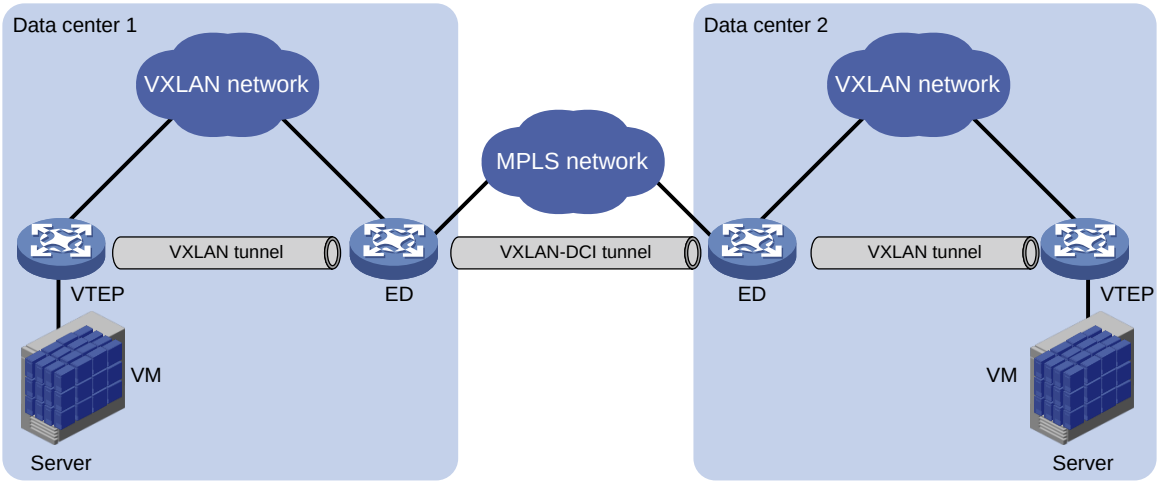
根据骨干网架构的不同，L3VPN 网络与 EVPN VXLAN 网络互通又可以分为如下两种类型：

- MPLS L3VPN 网络与 EVPN VXLAN 网络互通。
- EVPN L3VPN 网络与 EVPN VXLAN 网络互通。

上述两种类型的网络互通实现原理类似，本文的后续介绍均仅以 MPLS L3VPN 网络与 EVPN VXLAN 网络互通为例。

如图 31 所示，不同的 EVPN VXLAN 网络（通常是数据中心）之间通过 MPLS L3VPN 网络互连时，不同数据中心的 VM 需要跨域 EVPN VXLAN 网络以及 MPLS L3VPN 网络才能互相传递数据。因此，需要 ED 作为 MPLS L3VPN 网络和 EVPN VXLAN 网络的边界设备，对 BGP 路由进行重生成使得路由信息可以在不同网络中顺利地传递，以实现数据中心之间的互通。

图31 通过 MPLS L3VPN 网络连接不同的数据中心

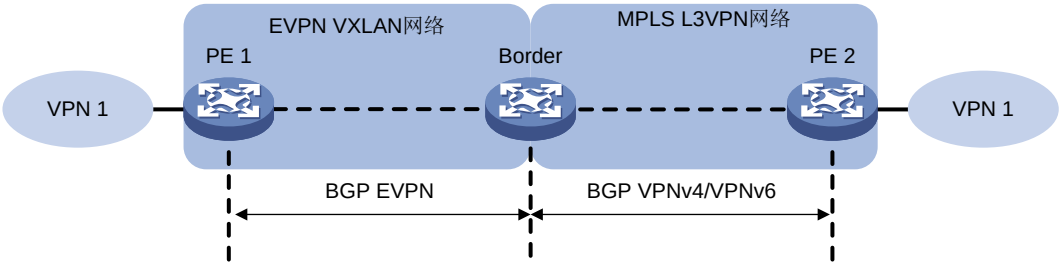


6.2 工作机制

6.2.1 组网模型

本章节将 MPLS L3VPN 网络与 EVPN VXLAN 网络互通场景抽象为如图 32 所示的组网进行介绍。Border 作为 MPLS L3VPN 网络与 EVPN VXLAN 网络的边界设备，与位于 EVPN VXLAN 网络的 PE 1 建立 BGP EVPN 会话，与位于 MPLS L3VPN 网络的 PE 2 建立 BGP VPNv4 或 BGP VPNv6 会话。部署 MPLS L3VPN 网络与 EVPN VXLAN 网络互通后，VPN 实例 1 内接入 PE 1 和接入 PE 2 的用户之间可以互通。

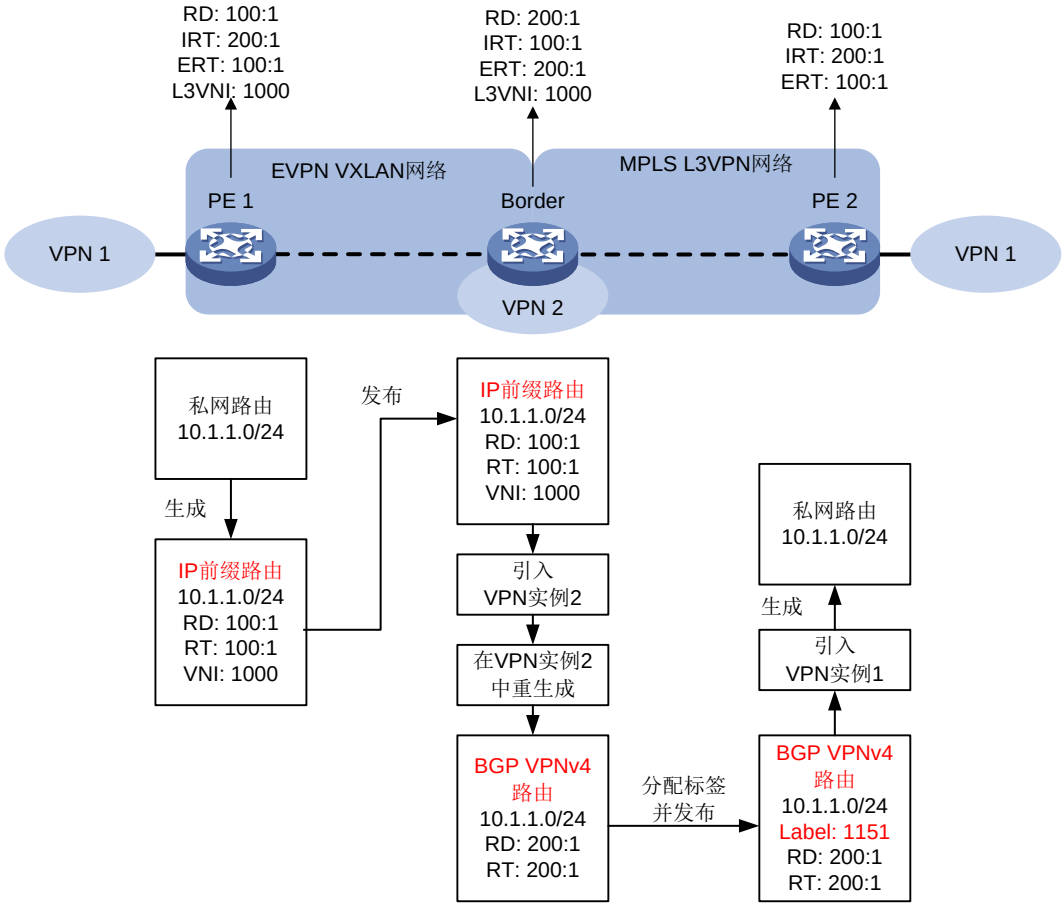
图32 MPLS L3VPN 网络与 EVPN VXLAN 网络互通组网图



互通主要在 Border 上实现，下面将从路由发布和报文转发两个方面来介绍 EVPN VXLAN 网络和 MPLS 网络互通的工作机制。

6.2.2 私网路由发布

图33 MPLS L3VPN 网络与 EVPN VXLAN 网络互通私网路由发布过程



如图 33所示，PE 1 和 PE 2 均通过 VPN 实例 1 接入私网用户，Border 部署 VPN 实例 2。PE 1 的 VPN 实例 1 关联 L3VNI 1000，Border 的 VPN 实例 2 关联 L3VNI 1000。以私网路由从 PE 1 发往 PE 2，Border 与 PE 2 建立 BGP VPNv4 会话为例，私网路由的发布过程为：

(1) PE 1 通过 IGP、BGP 或静态路由等方式获得 VPN 实例 1 用户的私网路由 10.1.1.0/24。

(2) PE 1 为本地 VPN 实例 1 的私网路由 10.1.1.0/24 增加 RD 100:1、RT 100:1 以及 L3VNI 1000 属性，生成 IP 前缀路由。

(3) PE 1 将 IP 前缀路由 10.1.1.0/24 通过 BGP EVPN 会话发布给 Border。

(4) Border 本地 VPN 实例 2 配置的 IRT 100:1 与 IP 前缀路由 10.1.1.0/24 携带的 RT 属性匹配，所以 Border 接收该路由、将其引入到 VPN 实例 2 并添加到 VPN 实例 2 的 IP 路由表中。

(5) Border 重生成引入到 VPN 实例 2 的 IP 前缀路由 10.1.1.0/24，重生成的路由为 BGP VPNv4 路由，携带 Border 的 VPN 实例 2 的 RD 200:1 以及 ERT 200:1。

(6) Border 为重生成的 BGP VPNv4 路由 10.1.1.0/24 分配私网标签 1151，并将该路由通过 BGP VPNv4 会话发布给 PE 2。私网标签 1151 绑定的 VPN 实例为 VPN 实例 2。

(7) PE 2 本地 VPN 实例 1 配置的 IRT 200:1 与 BGP VPNv4 路由 10.1.1.0/24 携带的 RT 属性匹配，所以 PE 2 接收该路由、将其引入到 VPN 实例 1 并添加到 VPN 实例 1 的 IP 路由表中。

(8) PE 2 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 10.1.1.0/24 发布给 VPN 实例 1 的用户。

私网路由从 PE 2 发往 PE 1 的机制类似，区别仅在于：

- 私网路由从 PE 1 发往 PE 2：Border 将来自 PE 1 的 IP 前缀路由重生成为 BGP VPNv4 路由。
- 私网路由从 PE 2 发往 PE 1：Border 将来自 PE 2 的 BGP VPNv4 路由重生成为 IP 前缀路由。

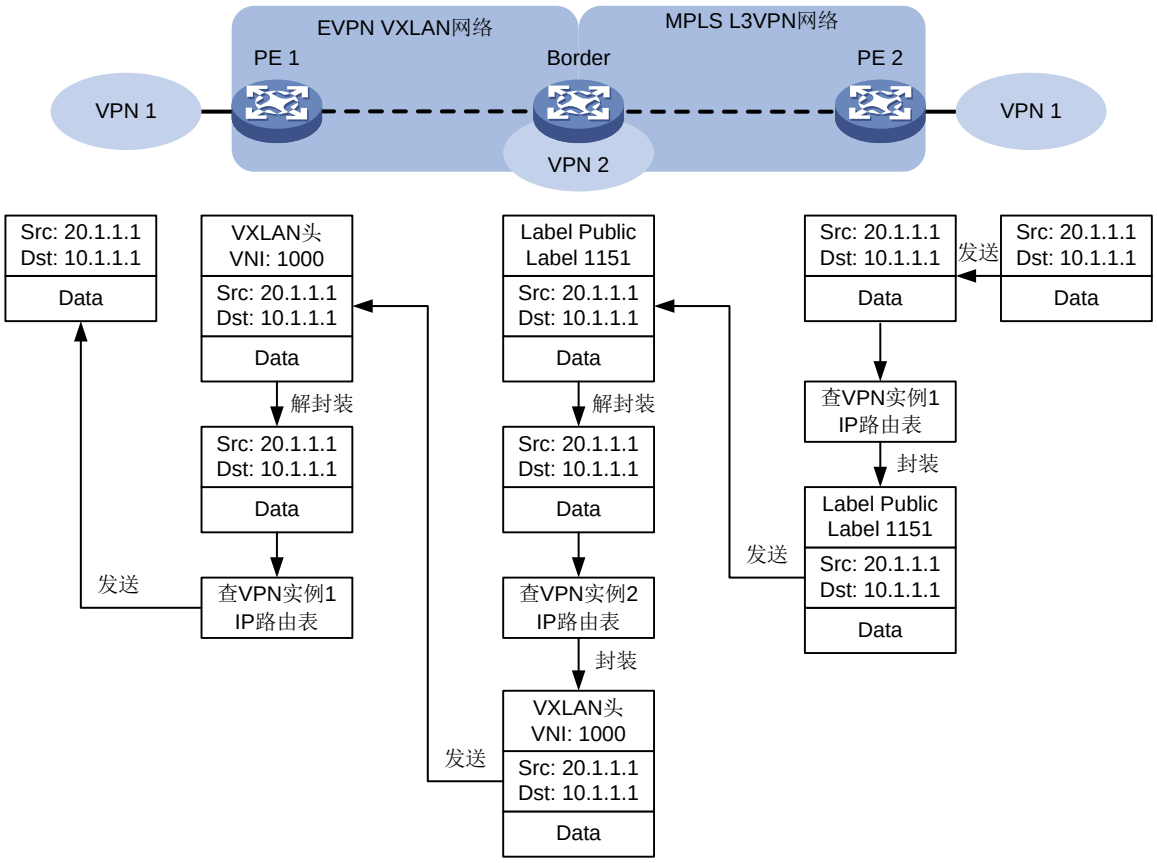
本章节不再赘述私网路由从 PE 2 发往 PE 1 的具体过程。

说明

- Border 与 PE 2 建立的是 BGP VPNv6 会话时，路由发布机制与建立 BGP VPNv4 会话时相同，区别仅在于承载私网路由信息的路由类型为 BGP VPNv6 路由。
- 在 EVPN L3VPN 网络与 EVPN VXLAN 网络互通组网中，EVPN L3VPN 网络通过 IP 前缀路由发布私网路由。

6.2.3 报文转发过程

图34 MPLS L3VPN 网络与 EVPN VXLAN 网络互通报文转发过程



如图 34 所示，私网报文从 PE 2 的 VPN 实例 1 内用户发往 PE 1 的 VPN 实例 1 内用户的过程为：

(1) VPN 实例 1 内用户将源地址为 20.1.1.1，目的地址为 10.1.1.1 的私网报文发送给 PE 2。

(2) PE 2 收到私网报文后，查找 VPN 实例 1 的 IP 路由表，查找到前缀为 10.1.1.0/24 的路由，找到该路由对应的私网标签 1151，为私网报文封装内层私网标签 1151 以及外层公网标签，将报文沿着标签转发路径转发给 Border。

(3) Border 收到私网报文后，将外层公网标签弹出，根据内层私网标签 1151 判断报文属于 VPN 实例 2，然后在 VPN 实例 2 中查找路由表。Border 查找到来自 PE 1 的前缀为 10.1.1.0/24 的路由，然后根据该路由关联的 L3VNI 1000，为私网报文封装 VXLAN 头，将私网报文通过 VXLAN 隧道发送给 PE 1。（VXLAN 隧道与 L3VNI 的关系为：L3VNI 关联 VSI 虚接口，VSI 虚接口关联 VSI，VSI 关联 VXLAN，VXLAN 关联 VXLAN 隧道）

(4) PE 1 收到私网报文后，根据 L3VNI 1000 关联的 VPN 实例 1，将私网报文的 VXLAN 封装解除后发送给 VPN 实例 1 内的用户。

私网报文从 PE 1 发往 PE 2 的机制类似，区别仅在于：

- 私网报文从 PE 1 发往 PE 2：Border 解封装来自 PE 1 的携带 VXLAN 头的私网报文，为报文重新封装 MPLS 标签后转发。
- 私网报文从 PE 2 发往 PE 1：Border 解封装来自 PE 2 的携带 MPLS 标签的私网报文，为报文重新封装 VXLAN 头后转发。

本章节不再赘述私网报文从 PE 1 发往 PE 2 的具体过程。

7 SRv6 网络与 EVPN VXLAN 网络互通

7.1 概述

当承载私网流量的骨干网穿越了 SRv6 网络与 EVPN VXLAN 网络时，需要实现 SRv6 网络与 EVPN VXLAN 网络互通。

根据骨干网架构的不同，SRv6 网络与 EVPN VXLAN 网络可以分为如下几种类型：

- IP L3VPN over SRv6 网络与 EVPN VXLAN 网络互通
- EVPN L3VPN over SRv6 网络与 EVPN VXLAN 网络互通
- EVPN VPLS over SRv6 网络与 EVPN VXLAN 网络互通

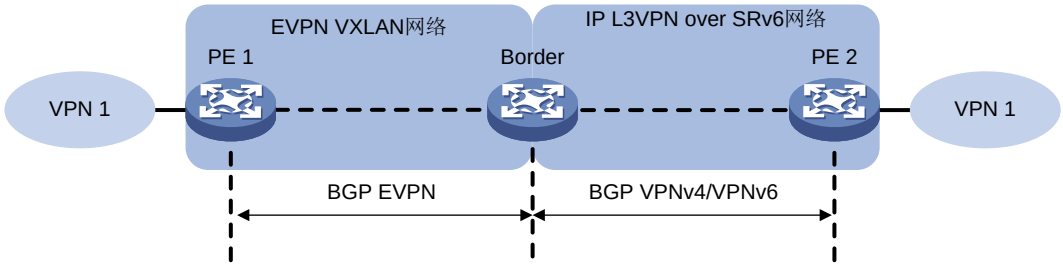
下文将分别介绍不同类型网络互通的工作机制。

7.2 IP L3VPN over SRv6与EVPN VXLAN网络互通的工作机制

7.2.1 组网模型

本章节将 IP L3VPN over SRv6 网络与 EVPN VXLAN 网络互通场景抽象为如图 35 所示的组网进行介绍。Border 作为 EVPN VXLAN 网络与 IP L3VPN over SRv6 网络的边界设备，与位于 EVPN VXLAN 网络的 PE 1 建立 BGP EVPN 会话，与位于 IP L3VPN over SRv6 网络的 PE 2 建立 BGP VPNv4 或 BGP VPNv6 会话。部署 IP L3VPN over SRv6 与 EVPN VXLAN 网络互通后，VPN 实例 1 内接入 PE 1 和 PE 2 的用户之间可以互通。

图35 IP L3VPN over SRv6 网络与 EVPN VXLAN 网络互通组网图

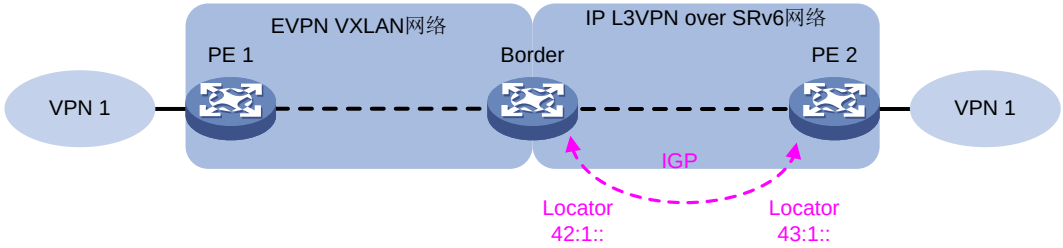


互通主要在 Border 上实现，下面将从路由发布和报文转发两个方面来介绍 IP L3VPN over SRv6 网络与 EVPN VXLAN 网络互通的工作机制。

7.2.2 Locator 网段路由发布

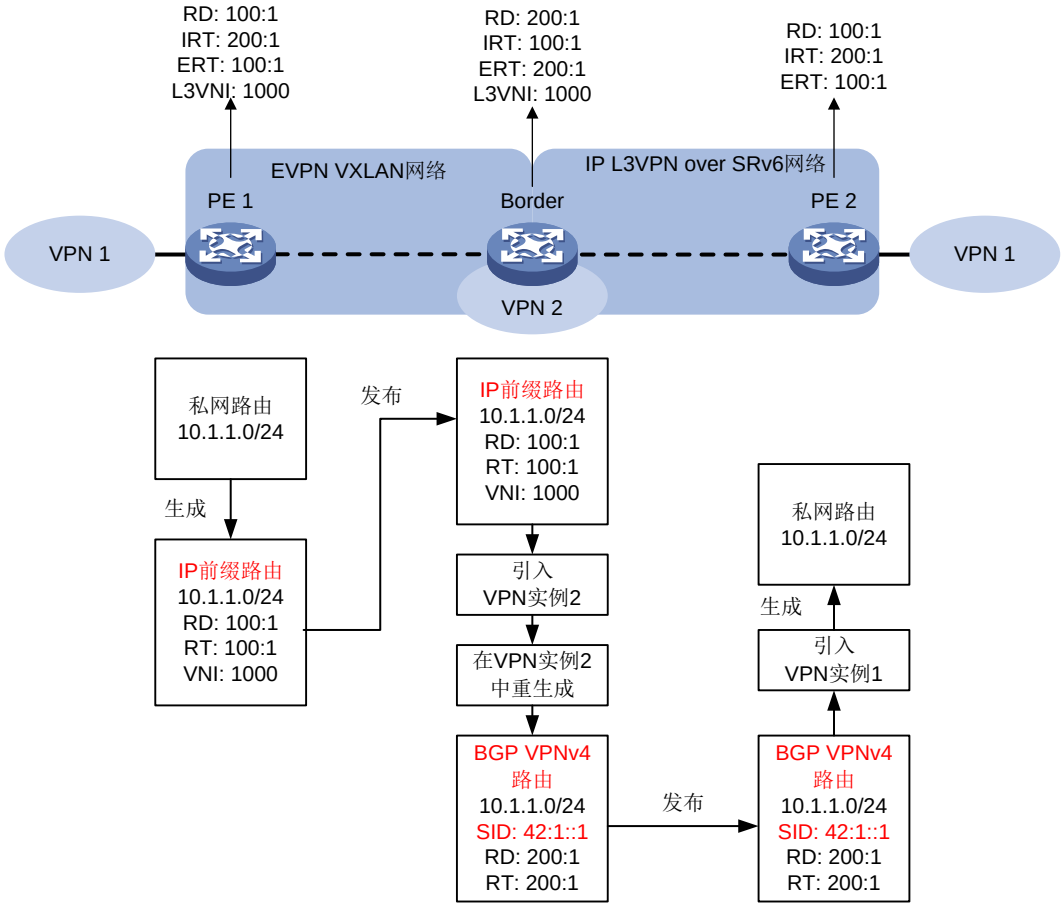
如图 36 所示，Border 和 PE 2 需要通过 IGP 协议向对方发布各自的 Locator 网段路由信息。Border 发布的 Locator 网段路由前缀为 42:1::/64，PE 2 发布的 Locator 网段路由前缀为 43:1::/64。

图36 IP L3VPN over SRv6 网络与 EVPN VXLAN 网络互通中 Locator 网段路由发布



7.2.3 私网路由发布

图37 IP L3VPN over SRv6 网络与 EVPN VXLAN 网络互通私网路由发布过程



如图 37 所示，PE 1 和 PE 2 均通过 VPN 实例 1 接入私网用户，Border 部署 VPN 实例 2。PE 1 的 VPN 实例 1 关联 L3VNI 1000，Border 的 VPN 实例 2 关联 L3VNI 1000。以私网路由从 PE 1 发往 PE 2，Border 与 PE 2 建立 BGP VPNv4 会话为例，私网路由的发布过程为：

(1) PE 1 通过 IGP、BGP 或静态路由等方式获得 VPN 实例 1 用户的私网路由 10.1.1.0/24。

(2) PE 1 为本地 VPN 实例 1 的私网路由 10.1.1.0/24 增加 RD 100:1、RT 100:1 以及 L3VNI 1000 属性，生成 IP 前缀路由。

(3) PE 1 将 IP 前缀路由 10.1.1.0/24 通过 BGP EVPN 会话发布给 Border。

(4) Border 本地 VPN 实例 2 配置的 IRT 100:1 与 IP 前缀路由 10.1.1.0/24 携带的 RT 属性匹配，所以 Border 接收该路由、将其引入到 VPN 实例 2 并添加到 VPN 实例 2 的 IP 路由表中。

(5) Border 重生成引入到 VPN 实例 2 的 IP 前缀路由 10.1.1.0/24，重生成的路由为 BGP VPNv4 路由，携带 PE 2 的 VPN 实例 2 的 RD 200:1 以及 ERT 200:1，并为其分配 SRv6 SID 42:1::1（SRv6 SID 类型可以为 End.DT4 SID、End.DT6 SID、End.DT46 SID、End.DX4 SID 或 End.DX6 SID，本例中的 SRv6 SID 为 End.DT4 SID）。

(6) Border 将重生成的 BGP VPNv4 路由 10.1.1.0/24 通过 BGP VPNv4 会话发布给 PE 2。

(7) PE 2 本地 VPN 实例 1 配置的 IRT 200:1 与 BGP VPNv4 路由 10.1.1.0/24 携带的 RT 属性匹配，所以 PE 2 接收该路由、将其引入到 VPN 实例 1 并添加到 VPN 实例 1 的 IP 路由表中。

(8) PE 2 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 10.1.1.0/24 发布给 VPN 实例 1 的用户。

私网路由从 PE 2 发往 PE 1 的机制类似，区别仅在于：

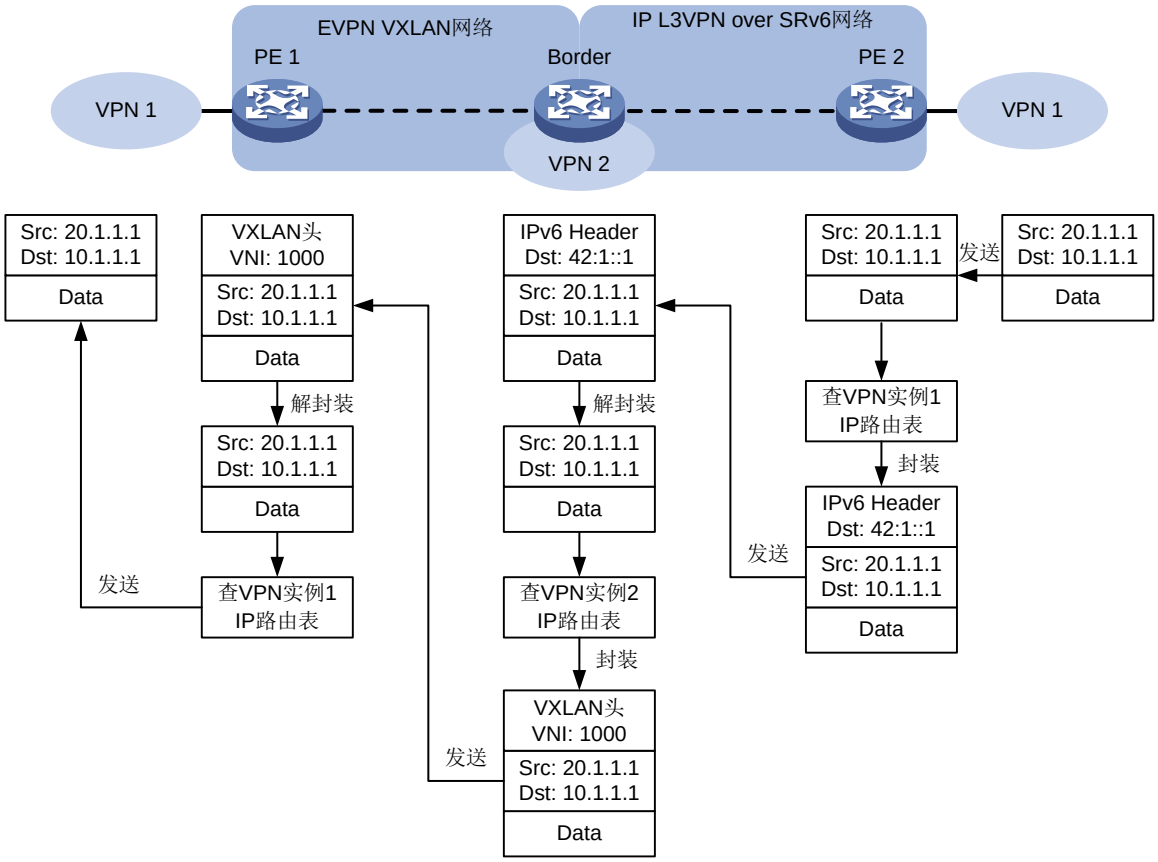
- 私网路由从 PE 1 发往 PE 2：Border 将来自 PE 1 的 IP 前缀路由重生成为携带 SRv6 SID 的 BGP VPNv4 路由。
- 私网路由从 PE 2 发往 PE 1：Border 将来自 PE 2 的携带 SRv6 SID 的 BGP VPNv4 路由重生成为 IP 前缀路由。

本章节不再赘述私网路由从 PE 2 发往 PE 1 的具体过程。

Border 与 PE 2 建立的是 BGP VPNv6 会话时，路由发布机制与建立 BGP VPNv4 会话时相同，区别仅在于承载私网路由信息的路由类型为 BGP VPNv6 路由。

7.2.4 报文转发过程

图38 IP L3VPN over SRv6 网络与 EVPN VXLAN 网络互通中 CE 1 到 CE 2 报文转发过程



如图 38 所示，以 SRv6 BE 转发方式为例，私网报文从 PE 2 的 VPN 实例 1 内用户发往 PE 1 的 VPN 实例 1 内用户的过程为：

(1) VPN 实例 1 内用户将源地址为 20.1.1.1，目的地址为 10.1.1.1 的私网报文发送给 PE 2。

(2) PE 2 收到私网报文后，查找 VPN 实例 1 的 IP 路由表，查找到前缀为 10.1.1.0/24 的路由，找到该路由对应的 SRv6 SID 42:1::1，为私网报文封装外层 IPv6 报文头，封装的外层 IPv6 报文头的目的 IPv6 地址为 SRv6 SID 42:1::1。

(3) PE 2 根据重新封装后的私网报文外层目的 IPv6 地址查找公网 IP 路由表，找到 Border 发布的 Locator 网段路由，然后根据该路由的指导将私网报文发送给 Border。

(4) Border 收到私网报文后，外层目的 IPv6 地址是本地的 End.DT4 SID，然后解封装该报文的外层 IPv6 头，在 End.DT4 SID 42:1::1 对应的 VPN 实例 2 中查找 IP 路由表。Border 查找到来自 PE 1 的前缀为 10.1.1.0/24 的路由，然后根据该路由关联的 L3VNI 1000，为私网报文封装 VXLAN 头，将私网报文通过 VXLAN 隧道发送给 PE 1。（VXLAN 隧道与 L3VNI 的关系为：L3VNI 关联 VSI 虚接口，VSI 虚接口关联 VSI，VSI 关联 VXLAN，VXLAN 关联 VXLAN 隧道）

(5) PE 1 收到私网报文后，根据 L3VNI 1000 关联的 VPN 实例 1，将私网报文的 VXLAN 封装解除后发送给 VPN 实例 1 内的用户。

私网报文从 PE 1 发往 PE 2 的机制类似，区别仅在于：

- 私网报文从 PE 1 发往 PE 2：Border 解封装来自 PE 1 的携带 VXLAN 头的私网报文，为报文进行 SRv6 封装后转发。
- 私网报文从 PE 2 发往 PE 1：Border 解封装来自 PE 2 的 SRv6 封装私网报文，为报文重新封装 VXLAN 头后转发。

本章节不再赘述私网报文从 PE 1 发往 PE 2 的具体过程。

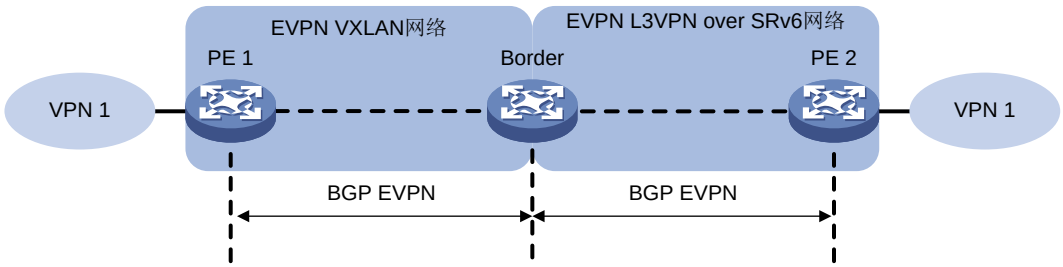
7.3 EVPN L3VPN over SRv6与EVPN VXLAN网络互通的工作机制

7.3.1 组网模型

本章节将 EVPN L3VPN over SRv6 网络与 EVPN VXLAN 网络互通场景抽象为如图 39 所示的组网进行介绍。Border 作为 EVPN VXLAN 网络与 EVPN L3VPN over SRv6 网络的边界设备，与位于 EVPN VXLAN 网络的 PE 1 和位于 EVPN L3VPN over SRv6 网络的 PE 2 均建立 BGP EVPN 会话。

缺省情况下，EVPN VXLAN 网络与 EVPN L3VPN over SRv6 网络中发布的 BGP EVPN 路由封装类型不同，导致不同网络间不能互通。部署 EVPN L3VPN over SRv6 与 EVPN VXLAN 网络互通后，VPN 实例 1 内接入 PE 1 和 PE 2 的用户之间可以互通。

图39 EVPN L3VPN over SRv6 网络与 EVPN VXLAN 网络互通组网图

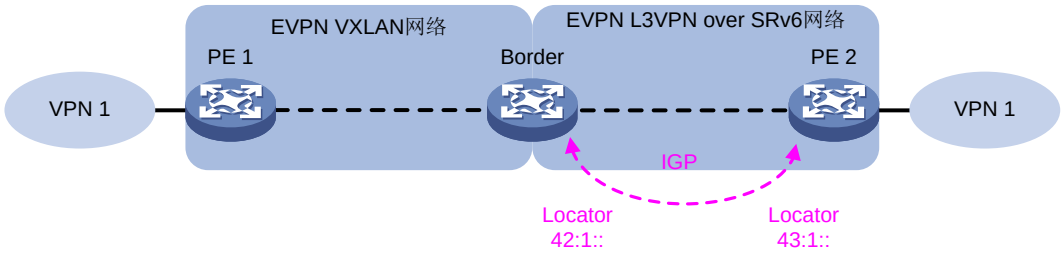


互通主要在 Border 上实现，下面将从路由发布和报文转发两个方面来介绍 EVPN L3VPN over SRv6 网络与 EVPN VXLAN 网络互通的工作机制

7.3.2 Locator 网段路由发布

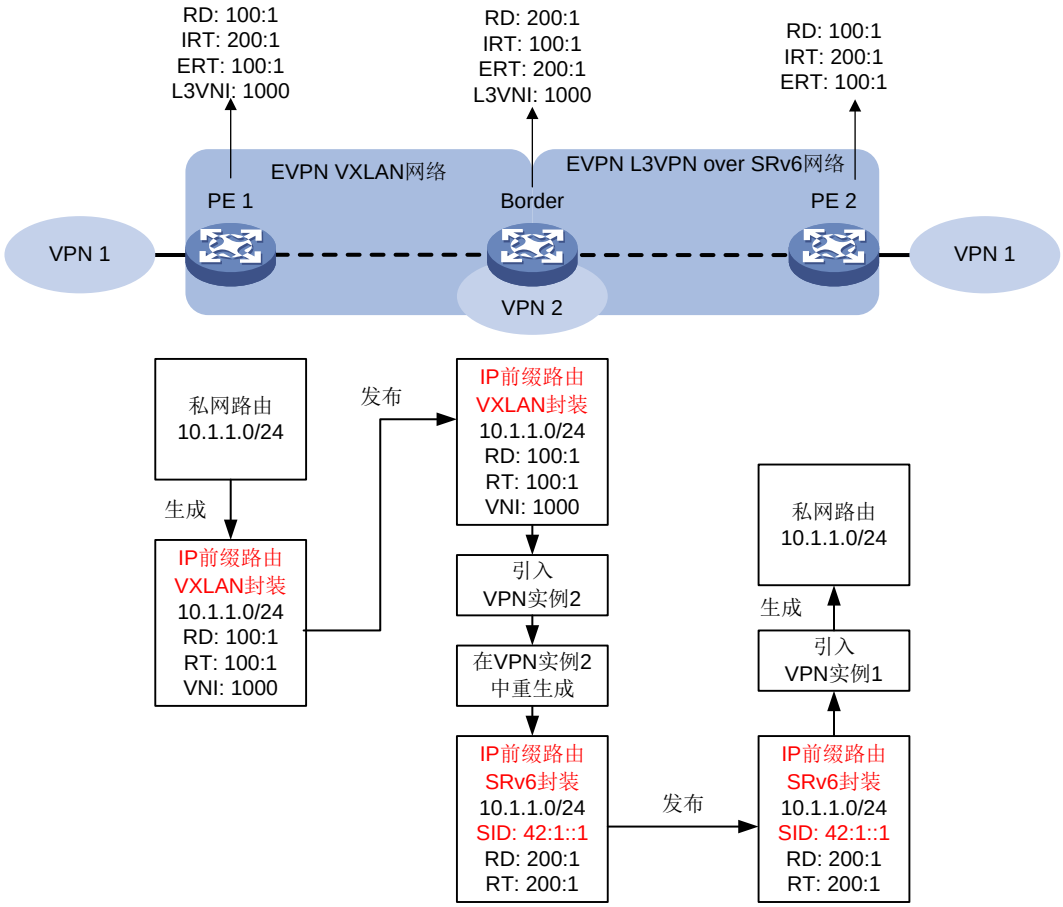
如图 40 所示，Border 和 PE 2 需要通过 IGP 协议向对方发布各自的 Locator 网段路由信息。Border 发布的 Locator 网段路由前缀为 42:1::/64，PE 2 发布的 Locator 网段路由前缀为 43:1::/64。

图40 EVPN L3VPN over SRv6 网络与 EVPN VXLAN 网络互通 Locator 网段路由发布



7.3.3 私网路由发布

图41 EVPN L3VPN over SRv6 网络与 EVPN VXLAN 网络互通私网路由发布过程



如图 41 所示，PE 1 和 PE 2 均通过 VPN 实例 1 接入私网用户，Border 部署 VPN 实例 2。PE 1 的 VPN 实例 1 关联 L3VNI 1000，Border 的 VPN 实例 2 关联 L3VNI 1000。以私网路由从 PE 1 发往 PE 2 为例，私网路由的发布过程为：

(1) PE 1 通过 IGP、BGP 或静态路由等方式获得 VPN 实例 1 用户的私网路由 10.1.1.0/24。

(2) PE 1 为本地 VPN 实例 1 的私网路由 10.1.1.0/24 增加 RD 100:1、RT 100:1 以及 L3VNI 1000 属性，生成 VXLAN 封装的 IP 前缀路由。

(3) PE 1 将 VXLAN 封装的 IP 前缀路由 10.1.1.0/24 通过 BGP EVPN 会话发布给 Border。

(4) Border 本地 VPN 实例 2 配置的 IRT 100:1 与 IP 前缀路由 10.1.1.0/24 携带的 RT 属性匹配，所以 Border 接收该路由、将其引入到 VPN 实例 2 并添加到 VPN 实例 2 的 IP 路由表中。

(5) Border 重生成引入到 VPN 实例 2 的 IP 前缀路由 10.1.1.0/24，重生成的路由为 SRv6 封装的 IP 前缀路由，携带 PE 2 的 VPN 实例 2 的 RD 200:1 以及 ERT 200:1，并为其分配 SRv6 SID 42:1::1（SRv6 SID 类型可以为 End.DT4 SID、End.DT6 SID、End.DT46 SID、End.DX4 SID 或 End.DX6 SID，本例中的 SRv6 SID 为 End.DT4 SID）。

(6) Border 将重生成的 SRv6 封装 IP 前缀路由 10.1.1.0/24 通过 BGP EVPN 会话发布给 PE 2。

(7) PE 2 本地 VPN 实例 1 配置的 IRT 200:1 与 BGP EVPN 路由 10.1.1.0/24 携带的 RT 属性匹配，所以 PE 2 接收该路由、将其引入到 VPN 实例 1 并添加到 VPN 实例 1 的 IP 路由表中。

(8) PE 2 通过 IGP、BGP 或静态路由等方式将 VPN 实例 1 的私网路由 10.1.1.0/24 发布给 VPN 实例 1 的用户。

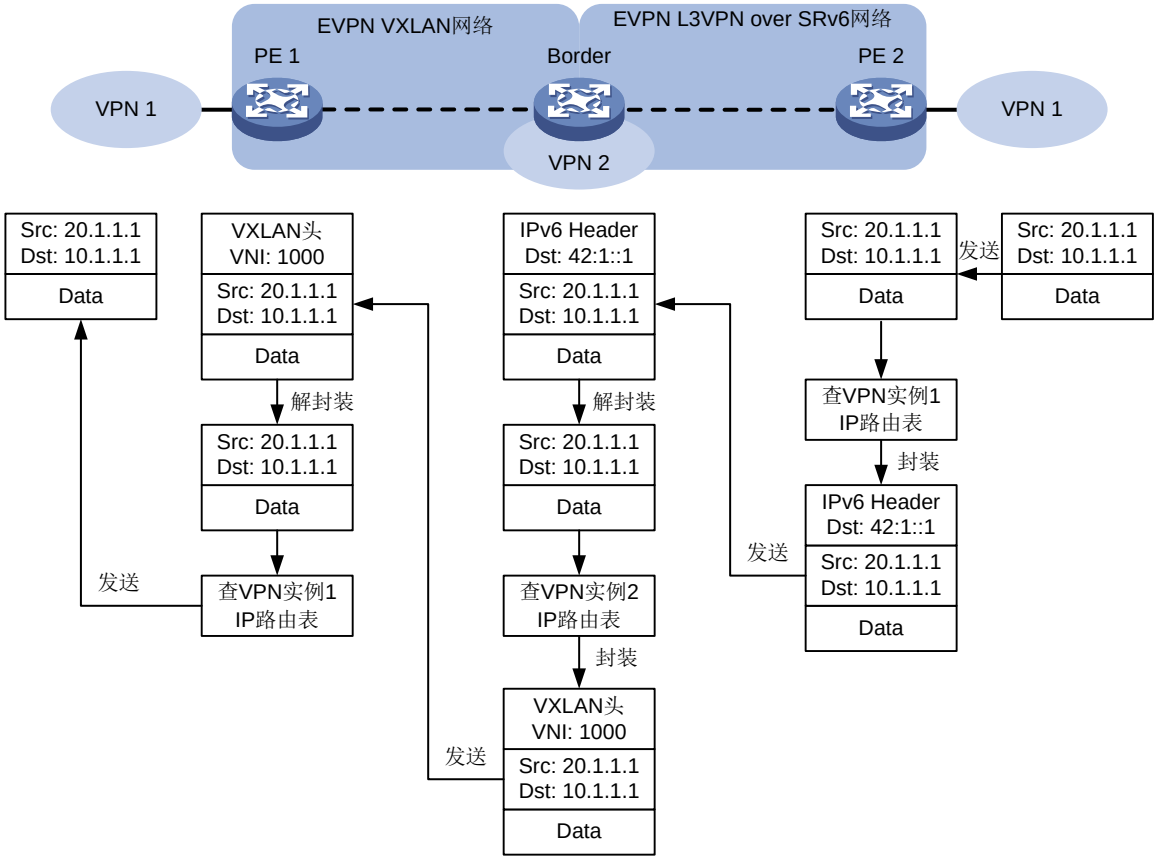
私网路由从 PE 2 发往 PE 1 的机制类似，区别仅在于：

- 私网路由从 PE 1 发往 PE 2：Border 将来自 PE 1 的携带 L3VNI 的 VXLAN 封装 IP 前缀路由重生成为携带 SRv6 SID 的 SRv6 封装 IP 前缀路由。
- 私网路由从 PE 2 发往 PE 1：Border 将来自 PE 2 的携带 SRv6 SID 的 SRv6 封装 IP 前缀路由重生成为携带 L3VNI 的 VXLAN 封装 IP 前缀路由。

本章节不再赘述私网路由从 PE 2 发往 PE 1 的具体过程。

7.3.4 报文转发过程

图42 EVPN L3VPN over SRv6 网络与 EVPN VXLAN 网络互通中 CE 1 到 CE 2 报文转发过程



如图 42 所示，以 SRv6 BE 转发方式为例，私网报文从 PE 2 的 VPN 实例 1 内用户发往 PE 1 的 VPN 实例 1 内用户的过程为：

(1) VPN 实例 1 内用户将源地址为 20.1.1.1，目的地址为 10.1.1.1 的私网报文发送给 PE 2。

(2) PE 2 收到私网报文后，查找 VPN 实例 1 的 IP 路由表，查找到前缀为 10.1.1.0/24 的路由，找到该路由对应的 SRv6 SID 42:1::1，为私网报文封装外层 IPv6 报文头，封装的外层 IPv6 报文头的目的 IPv6 地址为 SRv6 SID 42:1::1。

(3) PE 2 根据重新封装后的私网报文外层目的 IPv6 地址查找公网 IP 路由表，找到 Border 发布的 Locator 网段路由，然后根据该路由的指导将私网报文发送给 Border。

(4) Border 收到私网报文后，外层目的 IPv6 地址是本地的 End.DT4 SID，然后解封装该报文的外层 IPv6 头，在 End.DT4 SID 42:1::1 对应的 VPN 实例 2 中查找 IP 路由表。Border 查找到来自 PE 1 的前缀为 10.1.1.0/24 的路由，然后根据该路由关联的 L3VNI 1000，为私网报文封装 VXLAN 头，将私网报文通过 VXLAN 隧道发送给 PE 1。（VXLAN 隧道与 L3VNI 的关系为：L3VNI 关联 VSI 虚接口，VSI 虚接口关联 VSI，VSI 关联 VXLAN，VXLAN 关联 VXLAN 隧道）

(5) PE 1 收到私网报文后，根据 L3VNI 1000 关联的 VPN 实例 1，将私网报文的 VXLAN 封装解除后发送给 VPN 实例 1 内的用户。

私网报文从 PE 1 发往 PE 2 的机制类似，区别仅在于：

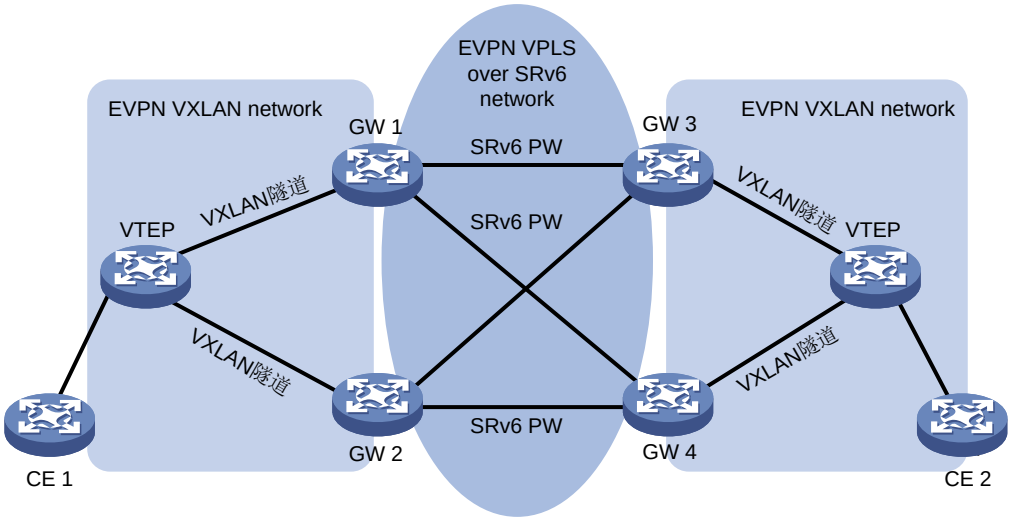
- 私网报文从 PE 1 发往 PE 2：Border 解封装来自 PE 1 的携带 VXLAN 头的私网报文，为报文进行 SRv6 封装后转发。
- 私网报文从 PE 2 发往 PE 1：Border 解封装来自 PE 2 的 SRv6 封装私网报文，为报文重新封装 VXLAN 头后转发。

本章节不再赘述私网报文从 PE 1 发往 PE 2 的具体过程。

7.4 EVPN VPLS over SRv6与EVPN VXLAN网络互通的工作机制

7.4.1 网络模型

图43 EVPN VXLAN 网络与 EVPN VPLS over SRv6 网络互通典型应用场景



EVPN VXLAN 网络与 EVPN VPLS over SRv6 网络互通功能的典型应用场景如图 43 所示。两个 EVPN VXLAN 网络通过 EVPN VPLS over SRv6 网络互联。GW 1、GW 2、GW 3 和 GW 4 为 EVPN VXLAN 网络和 EVPN VPLS over SRv6 网络的边界设备，通过在 GW 上重生成 EVPN 路由，实现 EVPN VXLAN 网络和 EVPN VPLS over SRv6 网络之间互通。

7.4.2 网络互通原理

EVPN VXLAN 网络和 EVPN VPLS over SRv6 网络之间实现互通，依赖于网络的边界设备上进行 MAC/IP 发布路由的重生成，具体过程为如下两种：

- GW 从 EVPN VXLAN 网络接收到 MAC/IP 发布路由后，GW 重生成该路由，即为该路由分配并添加 SRv6 SID、将路由的封装类型修改为 SRv6 封装、修改路由的 RD 和 RT，之后将该路由发送到 EVPN VPLS over SRv6 网络。
- GW 从 EVPN VPLS over SRv6 网络接收到 MAC/IP 发布路由后，GW 重生成该路由，即查找该路由对应的 VXLAN ID 并为路由添加 VXLAN ID、将路由的封装类型修改为 VXLAN 封装、修改路由的 RD 和 RT，之后将该路由发送到 EVPN VXLAN 网络。

完成路由重生成后，边界网关上会形成 SRv6 SID、VXLAN ID 与 VSI 实例的映射关系。在报文转发过程中，当 GW 收到 VXLAN 封装的报文时，对报文解封装，然后重新查表，再使用 VXLAN ID

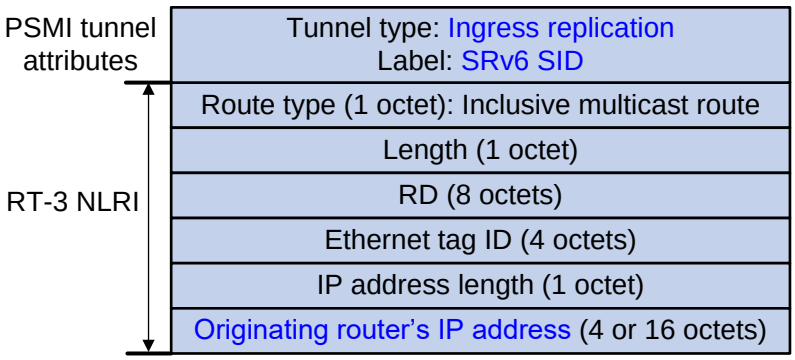
对应的 SRv6 SID 对报文进行重新封装。同样的，当 GW 收到 IPv6 或 SRv6 封装的报文时，对报文解封装，然后重新查表，再使用 SRv6 SID 对应的 VXLAN ID 对报文进行重新封装。

7.4.3 控制平面工作机制

1. SRv6 PW 及 BUM 广播表建立

如图 44 所示，EVPN VPLS over SRv6 组网中 PE 1 和 PE 2 依靠 RT-3（Inclusive Multicast Ethernet Tag Route）路由自动发现 PE 站点、建立 SRv6 PW 和 BUM 广播表。

图44 RT-3 路由消息格式



如图 45 所示，SRv6 PW 及 BUM 广播表建立的具体过程为：

(1) PE 1 通过 RT-3 向 PE 2 通告自身的 IP 地址，以及 PE 1 为 VSI 实例 A 分配的用于转发 BUM 流量的 End.DT2M SID。

(2) PE 2 收到 PE 1 发送的 RT-3 路由后，若路由中 VPN Target 属性与 PE 2 本地配置的 Import Target 属性匹配，则 PE 2 建立一条到达 PE 1 的 IP 地址的单向 SRv6 隧道，其出 SID 为 PE 1 携带的 End.DT2M SID；若路由中 VPN Target 属性与 PE 2 本地配置的 Import Target 属性不匹配，则不建立 SRv6 PW。

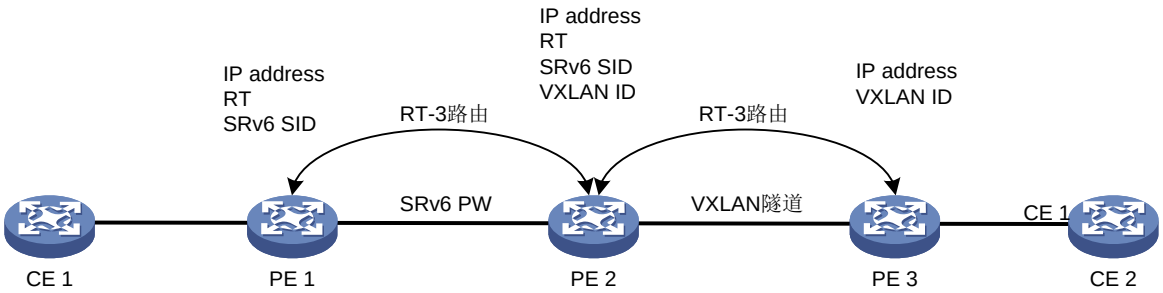
(3) PE 2 通过 RT-3 向 PE 1 通告自身的 IP 地址，以及 PE 2 为 VSI 实例 A 分配的用于转发 BUM 流量的 End.DT2M SID。

(4) PE 1 收到 PE 2 发送的 RT-3 路由后，若路由中 VPN Target 属性与 PE 1 本地配置的 Import Target 属性匹配，则 PE 1 建立一条到达 PE 2 的 IP 地址的单向 SRv6 隧道，其出 SID 为 PE 2 携带的 End.DT2M SID；若路由中 VPN Target 属性与 PE 1 本地配置的 Import Target 属性不匹配，则不建立 SRv6 PW。

(5) 两条方向相反的 SRv6 隧道，构成一条 SRv6 PW。

(6) 对于每个 VSI 而言，所有这些建立并关联的 SRv6 PW 就形成了 BUM 广播表。

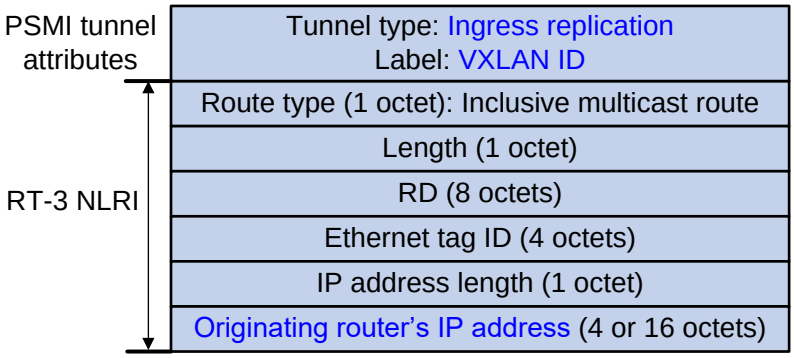
图45 SRv6 PW 及 BUM 广播表建立过程示意图



2. VXLAN 隧道及 BUM 广播表建立

如图 46 所示，EVPN VXLAN 组网中 PE 2 和 PE 3 依靠 RT-3（Inclusive Multicast Ethernet Tag Route）路由自动发现 PE 站点、建立 VXLAN 隧道和 BUM 广播表。

图46 RT-3 路由消息格式



VXLAN 隧道及 BUM 广播表建立的具体过程为：

(1) PE 2 通过 RT-3 向 PE 3 通告自己所属的 VXLAN ID 及自身的 IP 地址。

(2) PE 3 收到 PE 2 发送的 RT-3 路由后，若路由中 VPN Target 属性与 PE 3 本地配置的 Import Target 属性匹配，则获取 PE 2 的 VXLAN 信息以及 VXLAN 和下一跳的关系；若路由中 VPN Target 属性与 PE 3 本地配置的 Import Target 属性不匹配，则忽略该路由。

(3) PE 3 查看自己的 VXLAN 信息，若存在与 PE 2 相同的 VXLAN，则与 RT-3 路由中携带的下一跳自动建立 VXLAN 隧道，并将此 VXLAN 隧道与 VXLAN 关联。

(4) PE 3 通过 RT-3 向 PE 2 通告自己所属的 VXLAN ID 及自身的 IP 地址。

(5) PE 2 收到 PE 3 发送的 RT-3 路由后，若路由中 VPN Target 属性与 PE 2 本地配置的 Import Target 属性匹配，则获取 PE 3 的 VXLAN 信息以及 VXLAN 和下一跳的关系；若路由中 VPN Target 属性与 PE 2 本地配置的 Import Target 属性不匹配，则忽略该路由。

(6) PE 2 查看自己的 VXLAN 信息，若存在与 PE 3 相同的 VXLAN，则与 RT-3 路由中携带的下一跳自动建立 VXLAN 隧道，并将此 VXLAN 隧道与 VXLAN 关联。

(7) 对于每个 VXLAN 而言，所有这些建立并关联的 VXLAN 隧道就形成 BUM 广播表。

3. MAC 地址学习

如图 47 所示，EVPN VXLAN 与 EVPN VPLS over SRv6 互通组网中，PE 之间通过 MAC/IP 发布路由（即 RT-2，二类路由）通告和学习站点的 MAC 地址。

图47 RT-2 路由格式

Route type (1 octet): MAC/IP advertisement route
Length (1 octet)
RD (8 octets)
Ethernet segment identifier (10 octets)
Ethernet tag ID (4 octets)
MAC address length (1 octet)
MAC address (6 octets)
IP address length (1 octet)
IP address (0, 4, or 16 octets)
L2VNI (3 octets)
L3VNI (0 or 3 octets)

如图 48 所示，MAC 地址通告和学习过程为：

(1) PE 1 接收到 CE 1 发送的报文后，通过以太网报文的源 MAC 地址学习获得 CE 1 的 MAC 地址。即将 CE 1 的 MAC 地址添加到接收报文的 AC 口关联的 VSI 实例 A 的 MAC 地址表中，其出接口为连接 CE 1 的 AC 口。

(2) PE 1 通过 RT-2 路由将 CE 1 的 MAC 地址以及 PE 1 为 VSI 实例 A 分配的 End.DT2U SID 发布给 PE 2。

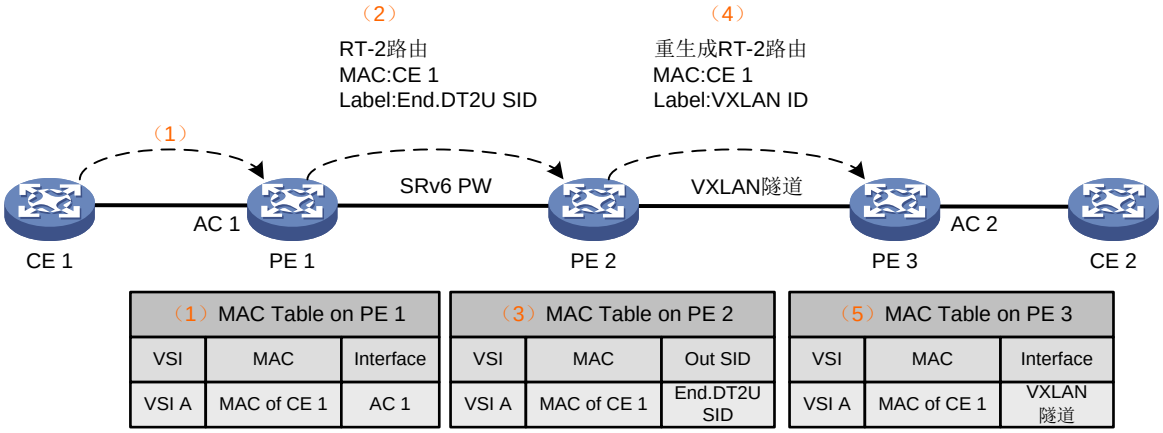
(3) PE 2 接收到 RT-2 路由后，如果路由中的 VPN Target 属性与 PE 2 本地配置的 Import target 属性匹配，则 PE 2 接收该路由，并将 CE 1 的 MAC 地址加入到本地 VSI 实例 A 的 MAC 地址表中，其出 SID 为 PE 1 通告的 End.DT2U SID；如果路由中的 VPN Target 属性与 PE 2 本地配置的 Import target 属性不匹配，则 PE 2 丢弃该路由。

(4) PE 2 将从 PE 1 接收到 RT-2 路由的下一跳地址修改为自身的地址、将报文封装类型修改为 VXLAN 方式，并添加 VSI 实例 A 关联的 VXLAN ID，然后通过 RT-2 路由发送给 PE 3。此外，PE 2 上还会建立 End.DT2U SID、VXLAN ID 与 VSI 实例的映射关系。

(5) PE 3 接收到 RT-2 路由后，如果路由中的 VPN Target 属性与 PE 3 本地配置的 Import target 属性匹配，则将 MAC 地址加入到本地 VSI 实例 A 的 MAC 地址表中，其出接口为接收该路由的 VXLAN 隧道。

完成 RT-2 路由发布后，PE 1、PE 2 和 PE 3 上均有 CE 1 的 MAC，PE 2 上还具有 End.DT2U SID、VXLAN ID 与 VSI 实例的映射关系。用户二层数据报文可以通过查找 MAC 地址表在 EVPN VXLAN 网络和 EVPN VPLS over SRv6 网络之间转发。

图48 MAC 地址通告和学习过程



（1）MAC Table on PE 1		
VSI	MAC	Interface
VSI A	MAC of CE 1	AC 1

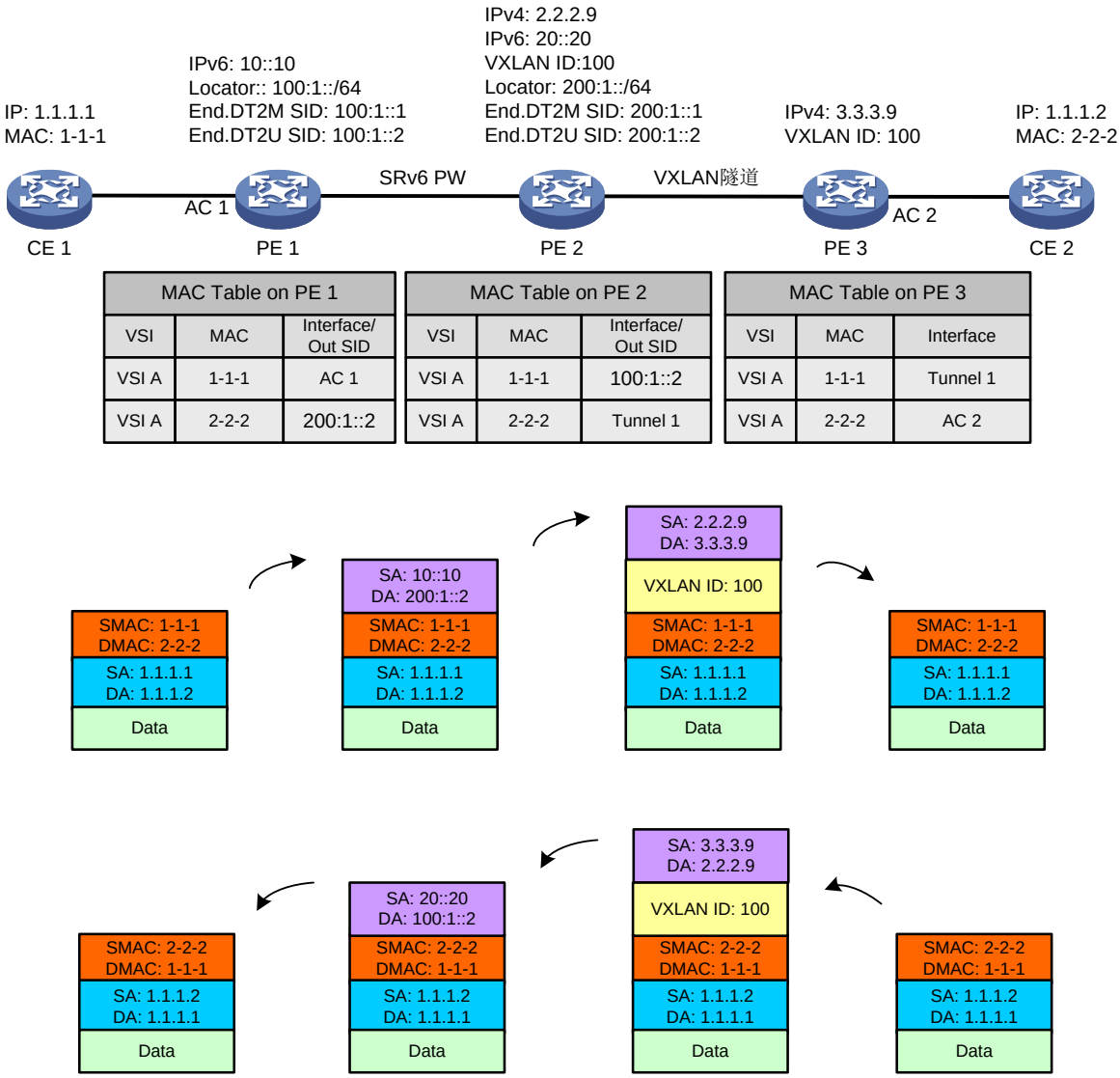
（3）MAC Table on PE 2		
VSI	MAC	Out SID
VSI A	MAC of CE 1	End.DT2U SID

（5）MAC Table on PE 3		
VSI	MAC	Interface
VSI A	MAC of CE 1	VXLAN 隧道

7.4.4 数据平面工作机制

1. 转发已知单播流量

图49 转发已知单播流量



如图 49 所示，PE 完成 MAC 地址表项的学习后，以 CE 1 访问 CE 2 为例，已知单播报文从 EVPN VPLS over SRv6 网络转发到 EVPN VXLAN 网络的转发过程为：

(1) PE 1 从 AC 接收到 CE 1 发送的报文后，在 AC 关联的 VSI 实例内查找 MAC 地址表，找到对应的出接口 SRv6 PW，并获取该 VSI 实例的 End.DT2U SID。

(2) PE 1 根据 SRv6 BE 或 SRv6 TE 方式为报文封装 IPv6 报文头，然后查找 IPv6 路由表将报文发送到 PE 2。

(3) PE 2 接收到报文后，对报文解封装，去掉 IPv6 报文头，根据 End.DT2U SID 找到关联的 VSI 实例，在 VSI 实例内查找 MAC 地址表，找到对应的 VXLAN 隧道，然后为报文进行 VXLAN 封装，通过 VXLAN 隧道将报文发送给 PE 3。

(4) PE 3 接收到报文后，对报文解封装，根据 VXLAN 报头中的 VXLAN ID 找到关联的 VSI，并在该 VSI 实例内查找 MAC 地址表，根据查表结果将报文转发给 CE 2。

以 CE 2 访问 CE 1 为例，已知单播报文从 EVPN VXLAN 网络转发到 EVPN VPLS over SRv6 网络的转发过程为：

(5) PE 3 从 AC 接收到 CE 2 发送的报文后，在 AC 关联的 VSI 实例内查找 MAC 地址表，找到对应的出接口 VXLAN 隧道，并获取 VSI 实例关联的 VXLAN ID。

(6) PE 3 为报文进行 VXLAN 封装，并通过 VXLAN 隧道将报文发送给 PE 2。

(7) PE 2 接收到报文后，对报文进行解封装，根据 VXLAN 报头中的 VXLAN ID 找到关联的 VSI，并在该 VSI 实例内查找 MAC 地址表，获取对应的 End.DT2U SID。

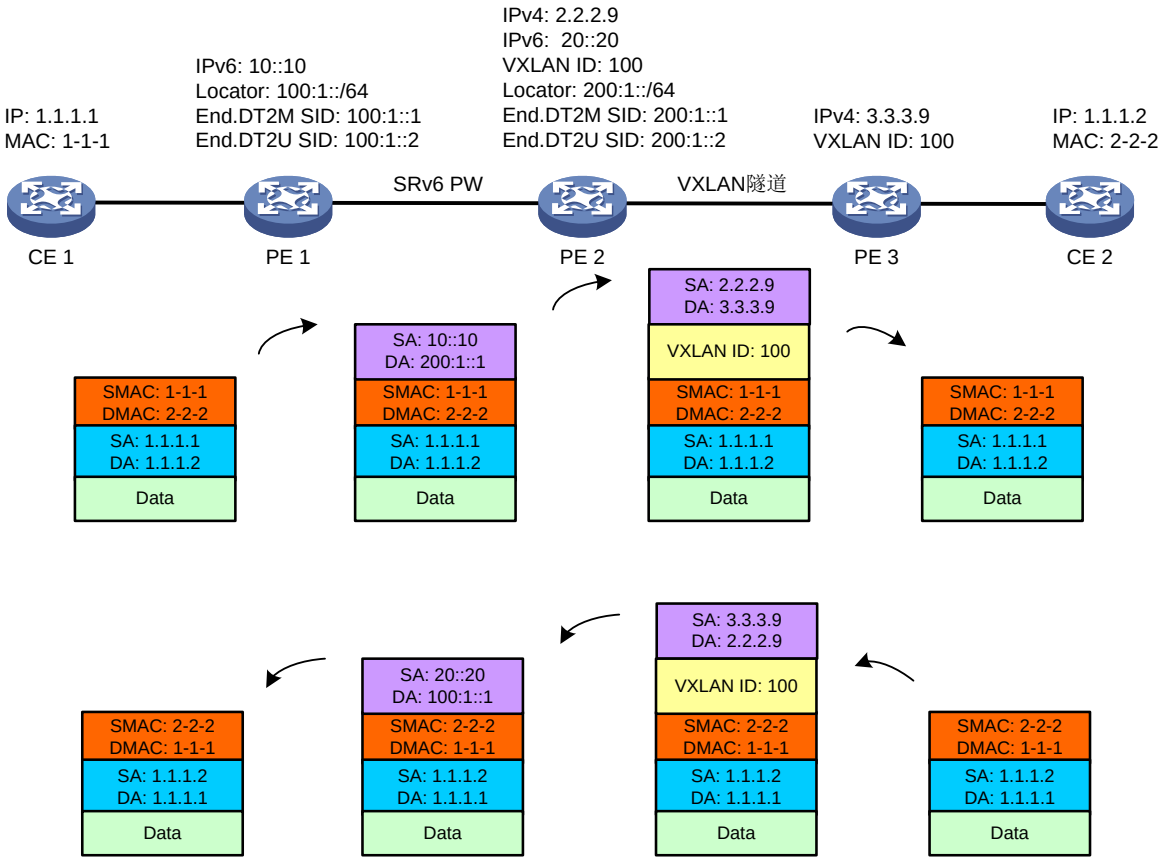
(8) PE 2 根据 SRv6 BE 或 SRv6 TE 方式为报文封装 IPv6 报文头，然后查找 IPv6 路由表将报文发送到 PE 1。

(9) PE 1 接收到报文后，对报文解封装，去掉 IPv6 报文头，根据 End.DT2U SID 找到关联的 VSI 实例，在该 VSI 实例内查找 MAC 地址表，根据查表结果将报文转发给 CE 1。

2. 转发 BUM 流量

除了单播流量转发，EVPN VPLS over SRv6 与 EVPN VXLAN 网络互通组网中还会转发广播、未知组播与未知单播流量，即 BUM 流量。

图50 转发 BUM 流量



如图 50 所示，以 CE 1 侧发送未知单播报文为例，报文转发过程为：

(1) PE 1 从 AC 接收到 CE 1 发送的 BUM 报文后，在 AC 关联的 VSI 的 MAC 地址表中未找到匹配的 MAC 地址表项，则在 VSI 中查找所有远端 PE 分配的所有 End.DT2M SID，此例中只有 PE 2。

(2) PE 1 根据 SRv6 BE 或 SRv6 TE 方式为报文封装 IPv6 报文头，然后查找 IPv6 路由表将报文发送到 PE 2。如果存在多个远端 PE，则会将报文复制多份并为报文封装不同的 IPv6 报文头，然后查表发给多个远端 PE。

(3) PE 2 接收到报文后，对报文解封装，去掉 IPv6 报文头，根据 End.DT2M SID 找到关联的 VSI 实例，在 VSI 实例内查找绑定的 VXLAN 隧道，然后为报文进行 VXLAN 封装，通过 VXLAN 隧道发送给 PE 3。如果存在多个远端 PE，则会对报文进行复制并进行 VXLAN 封装，然后通过多个 VXLAN 隧道将报文转发到远端 PE。

(4) PE 3 接收到报文后，对报文解封装，根据 VXLAN 报头中的 VXLAN ID 找到关联的 VSI，在该 VSI 实例内进行广播，即通过 VSI 实例关联的所有 AC 将报文转发到 CE。

以 CE 2 侧发送未知单播报文为例，报文转发过程为：

(1) PE 3 从 AC 收到 CE 2 发送的 BUM 报文后，在 AC 关联的 VSI 的 MAC 地址表中未找到匹配的 MAC 地址表项，则在该 VXLAN 内除接收接口外的所有本地 AC 口和 VXLAN 隧道转发该报文。

(2) PE 3 为报文进行 VXLAN 封装，并通过 VXLAN 隧道将报文发送给远端 PE 2。如果存在多个远端 PE，则会对报文进行复制并进行 VXLAN 封装，然后通过多个 VXLAN 隧道将报文转发到远端 PE。

(3) PE 2 接收到报文后，对报文进行解封装，根据 VXLAN 报头中的 VXLAN ID 找到关联的 VSI，并在该 VSI 实例内查找所有远端 PE 分配的所有 End.DT2M SID，此例中只有 PE 1。

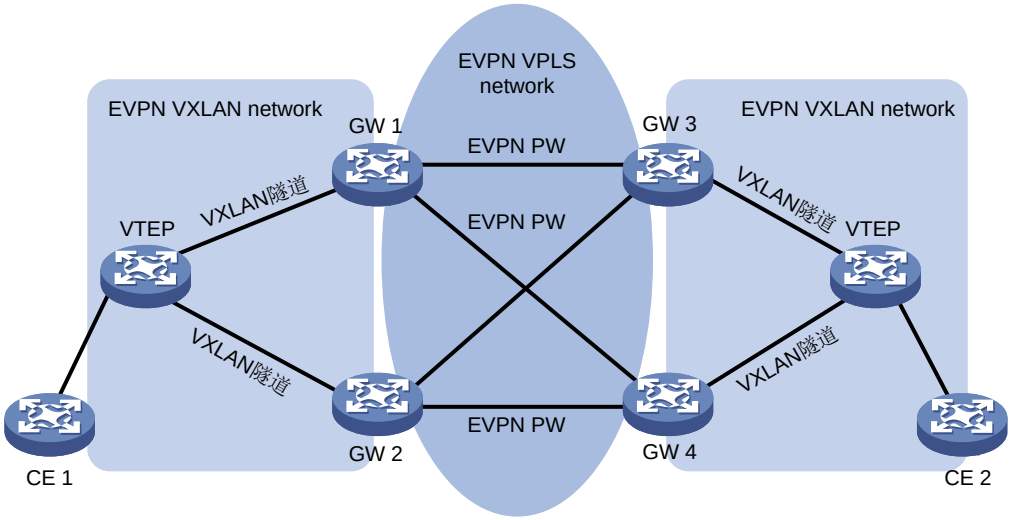
(4) PE 2 根据 SRv6 BE 或 SRv6 TE 方式为报文封装 IPv6 报文头，然后查找 IPv6 路由表将报文发送到 PE 1。如果存在多个远端 PE，则会将报文复制多份并为报文封装不同的 IPv6 报文头，然后查表发给多个远端 PE。

(5) PE 1 接收到报文后，对报文解封装，去掉 IPv6 报文头，根据 End.DT2M SID 找到关联的 VSI 实例，在该 VSI 实例内进行广播，即通过 VSI 实例关联的所有 AC 将报文转发到 CE。

8 EVPN VXLAN 与 EVPN VPLS 网络互通

8.1 网络模型

图51 EVPN VXLAN 网络与 EVPN VPLS 网络互通典型应用场景



EVPN VXLAN 网络与 EVPN VPLS 网络互通功能的典型应用场景如图 51 所示。两个 EVPN VXLAN 网络通过 EVPN VPLS 网络互联。GW 1、GW 2、GW 3 和 GW 4 为 EVPN VXLAN 网络和 EVPN VPLS 网络的边界设备，通过在 GW 上重生成 EVPN 路由，实现 EVPN VXLAN 网络和 EVPN VPLS 网络之间互通。

8.2 网络互通原理

EVPN VXLAN 网络和 EVPN VPLS 网络之间实现互通，依赖于网络的边界设备上进行 MAC/IP 发布路由的重生成，具体过程为如下两种：

- 如果 GW 从 EVPN VXLAN 网络接收到 MAC/IP 发布路由，则 GW 将重生成该路由，即为该路由分配并添加 MPLS 标签、将路由的封装类型修改为 MPLS 封装、修改路由的 RD 和 RT，之后将该路由发送到 EVPN VPLS 网络。
- 如果 GW 从 EVPN VPLS 网络接收到 MAC/IP 发布路由，则 GW 将重生成该路由，即查找该路由对应的 VXLAN ID 并为路由添加 VXLAN ID、将路由的封装类型修改为 VXLAN 封装、修改路由的 RD 和 RT，之后将该路由发送到 EVPN VXLAN 网络。

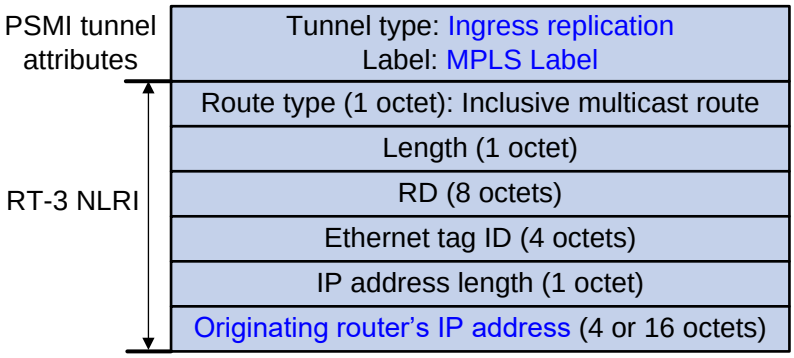
通过 MAC/IP 发布路由完成 MAC 地址表项学习后，GW 将根据学习到的表项为接收到的报文添加 VXLAN 或 MPLS 封装，将报文转发到 EVPN VXLAN 或 EVPN VPLS 网络。

8.3 控制平面工作机制

8.3.1 PW 及 BUM 广播表建立

如图 52 图 44 所示，EVPN VPLS 组网中 PE 1 和 PE 2 依靠 RT-3（Inclusive Multicast Ethernet Tag Route）路由自动发现 PE 站点、建立 PW 和 BUM 广播表。

图52 RT-3 路由消息格式



如图 53 所示，EVPN VPLS 组网中，PW 的建立过程为：

(1) PE 为每个 VSI 实例分配两个 PW 标签，分别用于转发已知单播报文和 BUM（Broadcast/Unknown unicast/Unknown Multicast，广播/未知单播/未知组播）报文。

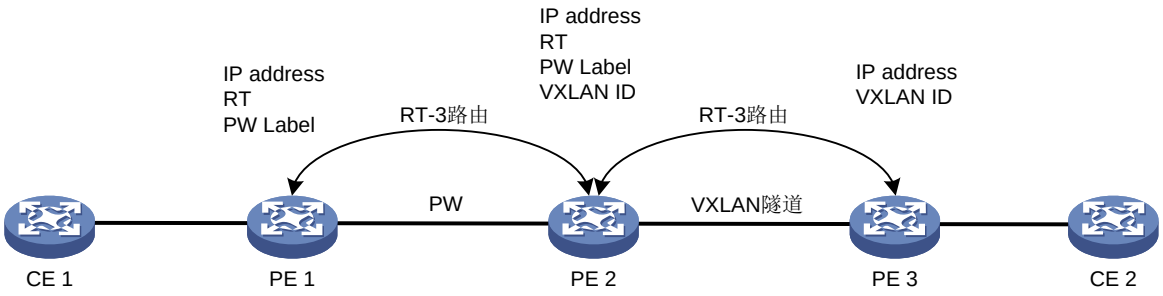
(2) 本端 PE 通过 MAC/IP 发布路由将转发已知单播报文的 PW 标签通告给远端 PE；通过 IMET 路由将转发 BUM 报文的 PW 标签通告给远端 PE。路由中携带 VPN Target 属性。

(3) 远端 PE 接收到 MAC/IP 发布路由或 IMET 路由后，将路由中的 VPN Target 属性与 EVPN 实例的 Import Target 进行匹配，如果一致则根据路由中携带的 PE 地址（对于 MAC/IP 发布路由，为路由的下一跳地址；对于 IMET 路由，为 PSMI tunnel attributes 中 Tunnel Identifier 字段携带的地址）、PW 标签等信息建立一条单向的虚拟连接。

(4) 当两端的 PE 间建立了两条方向相反的单向虚拟连接，则 PW 建立完成。

(5) 对于每个 VSI 而言，所有这些建立并关联的 PW 就形成了 BUM 广播表。

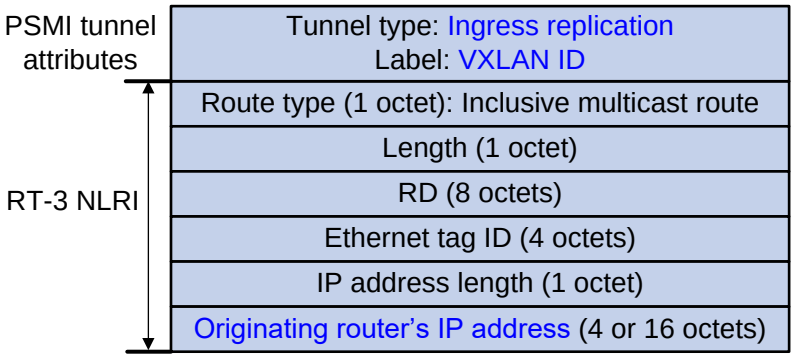
图53 PW 的建立过程示意图



8.3.2 VXLAN 隧道及 BUM 广播表建立

如图 54 所示，EVPN VXLAN 组网中 PE 2 和 PE 3 依靠 RT-3（Inclusive Multicast Ethernet Tag Route）路由自动发现 PE 站点、建立 VXLAN 隧道和 BUM 广播表。

图54 RT-3 路由消息格式



VXLAN 隧道及 BUM 广播表建立的具体过程为：

(1) PE 2 通过 RT-3 向 PE 3 通告自己所属的 VXLAN ID 及自身的 IP 地址。

(2) PE 3 收到 PE 2 发送的 RT-3 路由后，若路由中 VPN Target 属性与 PE 3 本地配置的 Import Target 属性匹配，则获取 PE 2 的 VXLAN 信息以及 VXLAN 和下一跳的关系；若路由中 VPN Target 属性与 PE 3 本地配置的 Import Target 属性不匹配，则忽略该路由。

(3) PE 3 查看自己的 VXLAN 信息，若存在与 PE 2 相同的 VXLAN，则与 RT-3 路由中携带的下一跳自动建立 VXLAN 隧道，并将此 VXLAN 隧道与 VXLAN 关联。

(4) PE 3 通过 RT-3 向 PE 2 通告自己所属的 VXLAN ID 及自身的 IP 地址。

(5) PE 2 收到 PE 3 发送的 RT-3 路由后，若路由中 VPN Target 属性与 PE 2 本地配置的 Import Target 属性匹配，则获取 PE 3 的 VXLAN 信息以及 VXLAN 和下一跳的关系；若路由中 VPN Target 属性与 PE 2 本地配置的 Import Target 属性不匹配，则忽略该路由。

(6) PE 2 查看自己的 VXLAN 信息，若存在与 PE 3 相同的 VXLAN，则与 RT-3 路由中携带的下一跳自动建立 VXLAN 隧道，并将此 VXLAN 隧道与 VXLAN 关联。

(7) 对于每个 VXLAN 而言，所有这些建立并关联的 VXLAN 隧道就形成 BUM 广播表。

8.3.3 MAC 地址学习

如图 55 所示，EVPN VXLAN 与 EVPN VPLS 互通组网中，PE 之间通过 MAC/IP 发布路由（即 RT-2，二类路由）通告和学习站点的 MAC 地址。

图55 RT-2 路由格式

Route type (1 octet): MAC/IP advertisement route
Length (1 octet)
RD (8 octets)
Ethernet segment identifier (10 octets)
Ethernet tag ID (4 octets)
MAC address length (1 octet)
MAC address (6 octets)
IP address length (1 octet)
IP address (0, 4, or 16 octets)
L2VNI (3 octets)
L3VNI (0 or 3 octets)

如图 56 所示，MAC 地址通告和学习过程为：

(1) PE 1 接收到 CE 1 发送的报文后，通过以太网报文的源 MAC 地址学习获得 CE 1 的 MAC 地址。即将 CE 1 的 MAC 地址添加到接收报文的 AC 口关联的 VSI 实例 A 的 MAC 地址表中，其出接口为连接 CE 1 的 AC 口。

(2) PE 1 通过 RT-2 路由将 CE 1 的 MAC 地址以及 PE 1 为 VSI 实例 A 分配的 PW 标签（用于转发已知单播报文的 PW 标签）发布给 PE 2。

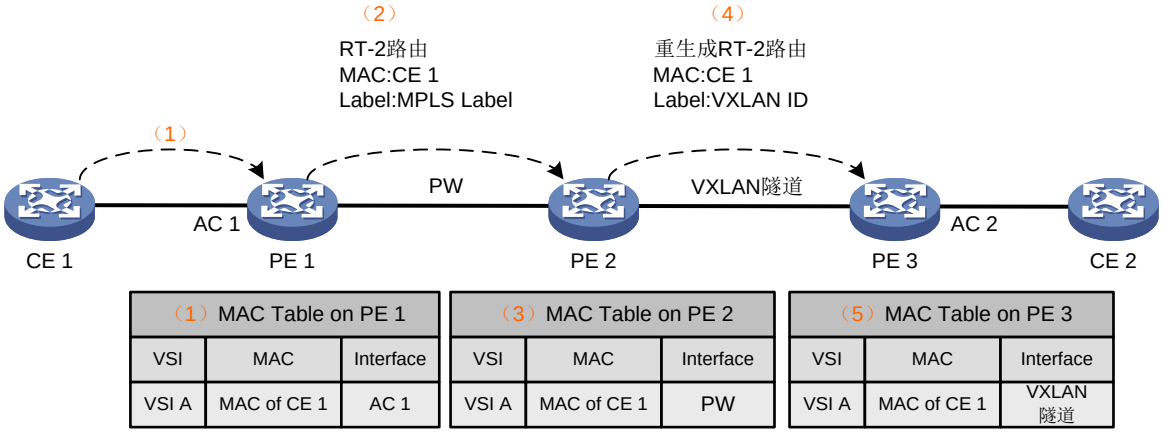
(3) PE 2 接收到 RT-2 路由后，如果路由中的 VPN Target 属性与 PE 2 本地配置的 Import target 属性匹配，则 PE 2 接收该路由，并将 CE 1 的 MAC 地址加入到本地 VSI 实例 A 的 MAC 地址表中，其出标签为 PE 1 通告的 PW 标签；如果路由中的 VPN Target 属性与 PE 2 本地配置的 Import target 属性不匹配，则 PE 2 丢弃该路由。

(4) PE 2 将从 PE 1 接收到 RT-2 路由的下一跳地址修改为自身的地址、将报文封装类型修改为 VXLAN 方式，并添加 VSI 实例 A 关联的 VXLAN ID，然后通过 RT-2 路由发送给 PE 3。

(5) PE 3 接收到 RT-2 路由后，如果路由中的 VPN Target 属性与 PE 3 本地配置的 Import target 属性匹配，则将 MAC 地址加入到本地 VSI 实例 A 的 MAC 地址表中，其出接口为接收该路由的 VXLAN 隧道。

完成 RT-2 路由发布后，PE 1、PE 2 和 PE 3 上均有 CE 1 的 MAC 地址表。用户二层数据报文可以通过查找 MAC 地址表在 EVPN VXLAN 网络和 EVPN VPLS 网络之间转发。

图56 MAC 地址通告和学习过程



（1）MAC Table on PE 1		
VSI	MAC	Interface
VSI A	MAC of CE 1	AC 1

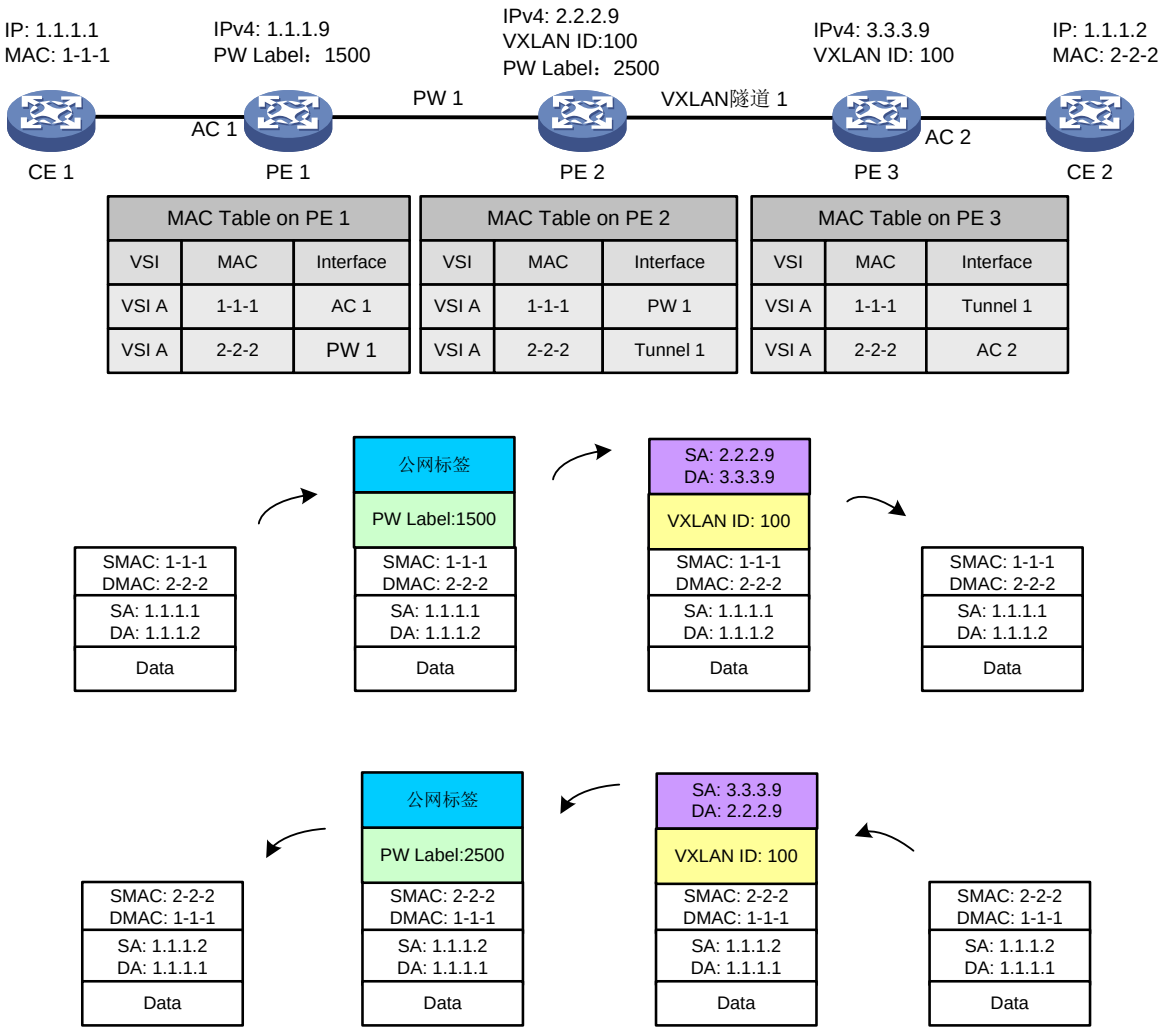
（3）MAC Table on PE 2		
VSI	MAC	Interface
VSI A	MAC of CE 1	PW

（5）MAC Table on PE 3		
VSI	MAC	Interface
VSI A	MAC of CE 1	VXLAN 隧道

8.4 数据平面工作机制

8.4.1 转发已知单播流量

图57 转发已知单播流量



如图 57所示，PE 完成 MAC 地址表项的学习后，以 CE 1 访问 CE 2 为例，已知单播报文从 EVPN VPLS 网络转发到 EVPN VXLAN 网络的转发过程为：

(1) PE 1 从 AC 接收到 CE 1 发送的报文后，在 AC 关联的 VSI 实例内查找 MAC 地址表，找到对应的出接口 PW，并获取该 VSI 实例用于转发已知单播报文的 PW 标签。

(2) PE 1 为报文封装用于转发已知单播报文的 PW 标签，然后通过 MAC 地址表的出接口（PW）将报文发送到 PE 2。



说明

如果公网隧道为 LSP 或 MPLS TE 隧道，则通过 PW 转发报文时将为报文封装两层标签：内层标签为私网标签，用来决定报文转发给相应的 VSI；外层标签为公网 LSP 或 MPLS TE 隧道标签，用来保证报文在 PE 之间正确传送。

(3) PE 2 接收到报文后，对报文解封装，去掉公网标签，根据 PW 标签找到关联的 VSI 实例，在 VSI 实例内查找 MAC 地址表，找到对应的 VXLAN 隧道，然后为报文进行 VXLAN 封装，通过 VXLAN 隧道将报文发送给 PE 3。

(4) PE 3 接收到报文后，对报文解封装，根据 VXLAN 报头中的 VXLAN ID 找到关联的 VSI，并在该 VSI 实例内查找 MAC 地址表，根据查表结果将报文转发给 CE 2。

以 CE 2 访问 CE 1 为例，已知单播报文从 EVPN VXLAN 网络转发到 EVPN VPLS 网络的转发过程为：

(5) PE 3 从 AC 接收到 CE 2 发送的报文后，在 AC 关联的 VSI 实例内查找 MAC 地址表，找到对应的出接口 VXLAN 隧道，并获取 VSI 实例关联的 VXLAN ID。

(6) PE 3 为报文进行 VXLAN 封装，并通过 VXLAN 隧道将报文发送给 PE 2。

(7) PE 2 接收到报文后，对报文进行解封装，根据 VXLAN 报头中的 VXLAN ID 找到关联的 VSI，并在该 VSI 实例内查找 MAC 地址表，获取用于转发已知单播报文的 PW 标签。

(8) PE 2 为报文封装用于转发已知单播报文的 PW 标签，然后通过 MAC 地址表的出接口（PW）将报文发送到 PE 1。



说明

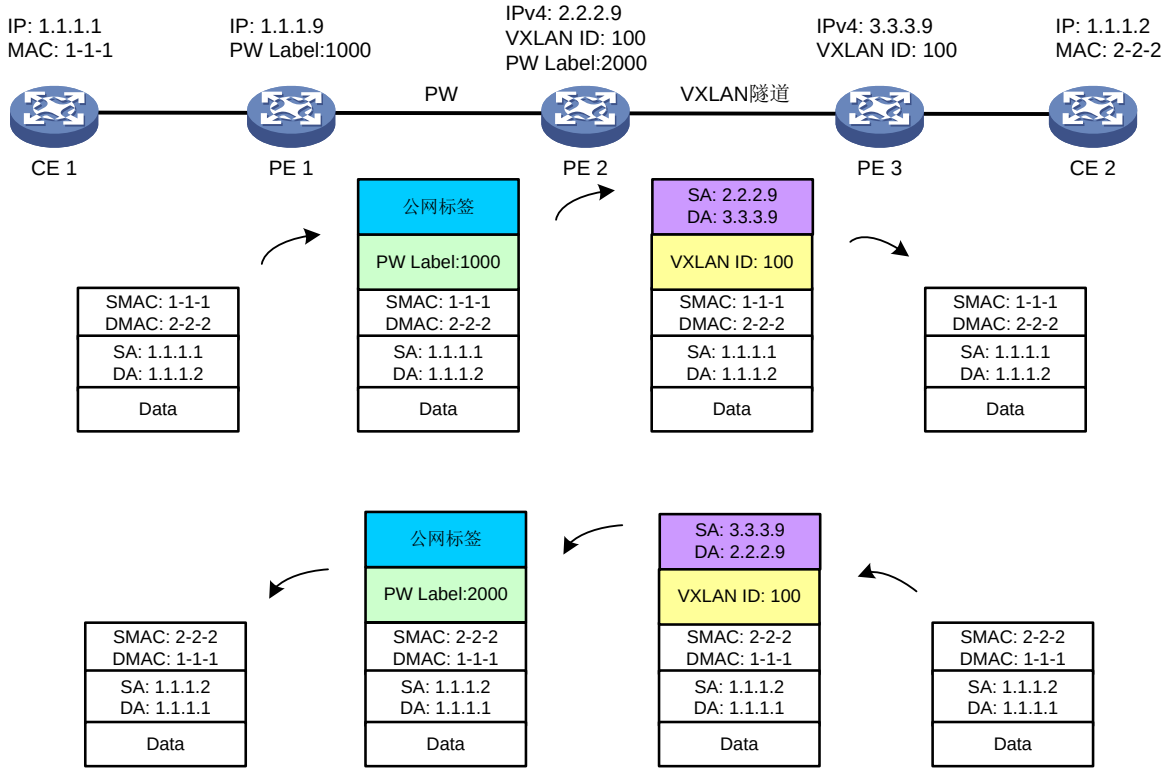
如果公网隧道为 LSP 或 MPLS TE 隧道，则通过 PW 转发报文时将为报文封装两层标签：内层标签为私网标签，用来决定报文转发给相应的 VSI；外层标签为公网 LSP 或 MPLS TE 隧道标签，用来保证报文在 PE 之间正确传送。

(9) PE 1 接收到报文后，对报文解封装，去掉公网标签，根据 PW 标签找到关联的 VSI 实例，在该 VSI 实例内查找 MAC 地址表，根据查表结果将报文转发给 CE 1。

8.4.2 转发 BUM 流量

除了单播流量转发，EVPN VPLS 与 EVPN VXLAN 网络互通组网中还会转发广播、未知组播与未知单播流量，即 BUM 流量。

图58 转发 BUM 流量



如图 58 所示，以 CE 1 侧发送未知单播报文为例，报文转发过程为：

(1) PE 1 从 AC 接收到 CE 1 发送的 BUM 报文后，在 AC 关联的 VSI 的 MAC 地址表中未找到匹配的 MAC 地址表项，则在 VSI 中查找所有远端 PE 分配的转发 BUM 报文的 PW 标签，此例中只有 PE 2。

(2) PE 1 根据 VSI 内所有用于转发 BUM 流量的 PW 标签，为报文分别封装这些 PW 标签后，将该报文泛洪给所有远端 PE。

说明

如果公网隧道为 LSP 或 MPLS TE 隧道，则通过 PW 转发报文时将为报文封装两层标签：内层标签为私网标签，用来决定报文转发给相应的 VSI；外层标签为公网 LSP 或 MPLS TE 隧道标签，用来保证报文在 PE 之间正确传送。

(3) PE 2 接收到报文后，对报文解封装，去掉公网标签，根据 PW 标签找到关联的 VSI 实例，在 VSI 实例内查找绑定的 VXLAN 隧道，然后为报文进行 VXLAN 封装，通过 VXLAN 隧道发送给 PE 3。如果存在多个远端 PE，则会对报文进行复制并进行 VXLAN 封装，然后通过多个 VXLAN 隧道将报文转发到远端 PE。

(4) PE 3 接收到报文后，对报文解封装，根据 VXLAN 报头中的 VXLAN ID 找到关联的 VSI，在该 VSI 实例内进行广播，即通过 VSI 实例关联的所有 AC 将报文转发到 CE。

以 CE 2 侧发送未知单播报文为例，报文转发过程为：

(5) PE 3 从 AC 收到 CE 2 发送的 BUM 报文后，在 AC 关联的 VSI 的 MAC 地址表中未找到匹配的 MAC 地址表项，则在该 VXLAN 内除接收接口外的所有本地 AC 口和 VXLAN 隧道转发该报文。

(6) PE 3 为报文进行 VXLAN 封装，并通过 VXLAN 隧道将报文发送给远端 PE 2。如果存在多个远端 PE，则会对报文进行复制并进行 VXLAN 封装，然后通过多个 VXLAN 隧道将报文转发到远端 PE。

(7) PE 2 接收到报文后，对报文进行解封装，根据 VXLAN 报头中的 VXLAN ID 找到关联的 VSI，并在该 VSI 实例内查找所有用于转发 BUM 流量的 PW 标签，此例中只有 PE 1。

(8) PE 2 为报文分别封装这些 PW 标签后，将该报文泛洪给所有远端 PE。

说明

如果公网隧道为 LSP 或 MPLS TE 隧道，则通过 PW 转发报文时将为报文封装两层标签：内层标签为私网标签，用来决定报文转发给相应的 VSI；外层标签为公网 LSP 或 MPLS TE 隧道标签，用来保证报文在 PE 之间正确传送。

(9) PE 1 接收到报文后，对报文解封装，去掉公网标签，根据 PW 标签找到关联的 VSI 实例，在该 VSI 实例内进行广播，即通过 VSI 实例关联的所有 AC 将报文转发到 CE。

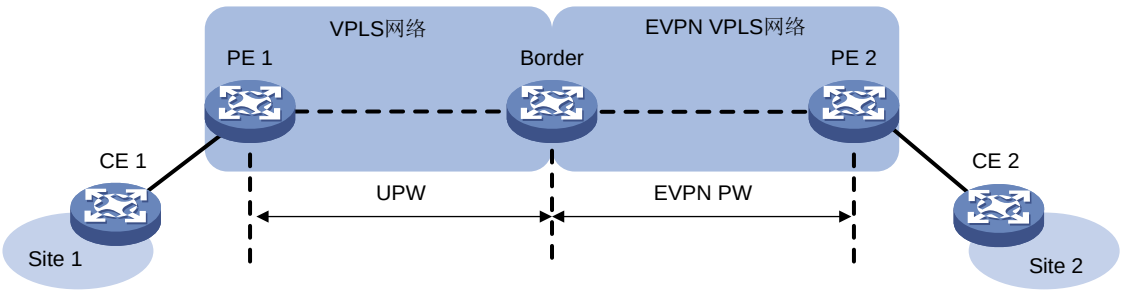
9 VPLS 与 EVPN VPLS/EVPN VPLS over SRv6/EVPN VXLAN 互通

9.1 概述

在实际组网中，可能会存在传统的 VPLS 网络与 EVPN VPLS（或 EVPN VPLS over SRv6、EVPN VXLAN）网络共存的情况。为了实现站点间流量跨越 VPLS 网络和 EVPN VPLS（或 EVPN VPLS over SRv6、EVPN VXLAN）网络传递，VPLS 网络与 EVPN VPLS（或 EVPN VPLS over SRv6、EVPN VXLAN）网络需要互通。

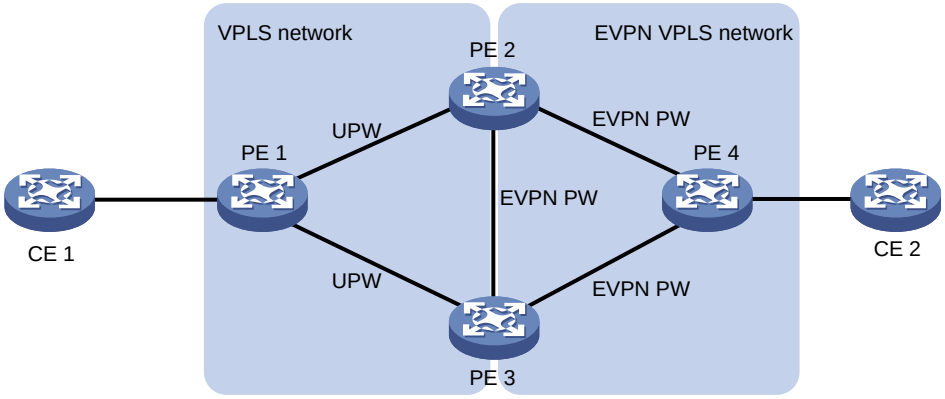
通过将 VPLS 网络中的 LDP PW 或静态 PW 看作 EVPN VPLS（或 EVPN VPLS over SRv6、EVPN VXLAN）网络的 AC（该 PW 称为 UPW），可以实现报文在 EVPN PW（或 SRv6 PW、VXLAN 隧道）与 UPW 之间相互转发，从而实现 VPLS 网络与 EVPN VPLS（或 EVPN VPLS over SRv6、EVPN VXLAN）网络的互通。

图59 VPLS 网络与 EVPN VPLS 网络互通示意图



如图 60 所示，VPLS 与 EVPN VPLS（或 EVPN VPLS over SRv6、EVPN VXLAN）互通组网不仅支持一条 LDP PW 或静态 PW 接入一条 EVPN PW（或 SRv6 PW、VXLAN 隧道），还支持将两条 LDP PW 或静态 PW 多归属接入两条 EVPN PW（或 SRv6 PW、VXLAN 隧道）。

图60 LDP PW 或静态 PW 多归属接入示意图



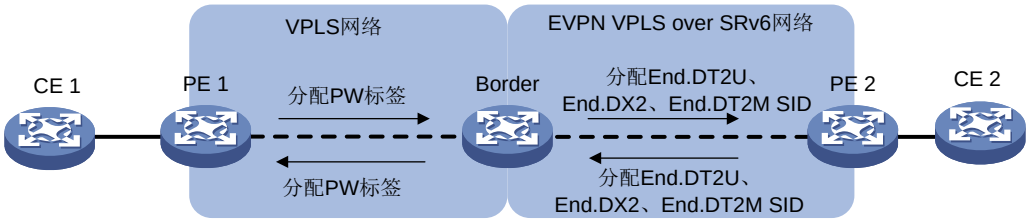
9.2 工作机制

VPLS 网络与 EVPN VPLS 网络互通、VPLS 网络与 EVPN VPLS over SRv6 网络互通、VPLS 网络与 EVPN VXLAN 网络互通的工作机制类似。下文将以 VPLS 网络与 EVPN VPLS over SRv6 网络互通为例，说明该工作机制。

9.2.1 建立 PW

如图 61 所示，VPLS 与 EVPN VPLS over SRv6 互通组网中，Border 设备同时运行 VPLS 和 EVPN VPLS over SRv6，与 PE 1 建立 VPLS 的 PW、与 PE 2 建立 SRv6 PW。其中，Border 和 PE 1 之间的 PW 为 UPW，在 EVPN VPLS over SRv6 网络中将该 UPW 当作 AC 来转发流量。

图61 VPLS 与 EVPN VPLS over SRv6 互通组网中建立 PW



说明

PW 和 SRv6 PW 的建立方式与 VPLS 网络、EVPN VPLS over SRv6 网络单独部署时相同，本文不再赘述。

9.2.2 MAC 地址学习

VPLS 网络和 EVPN VPLS over SRv6 网络均为二层 VPN 网络。在这些网络中 PE 需要学习用户站点的 MAC 地址，以便在用户站点所属的 VSI 内查找 MAC 地址表，对数据帧进行二层转发。

VPLS 与 EVPN VPLS over SRv6 互通组网中，不同网络的 MAC 地址学习方式为：

- VPLS 根据数据帧的源 MAC 地址学习 MAC 地址表项。
- EVPN VPLS over SRv6 根据数据帧的源 MAC 地址学习本地站点的 MAC 地址表项，并将该 MAC 地址通过 EVPN 的 MAC/IP 发布路由（二类路由）发布给远端 PE。

以图 62 为例，VPLS 网络与 EVPN VPLS over SRv6 网络互通的 MAC 地址学习过程为：

(1) CE 1 广播发送 ARP 请求报文获取 CE 2 的 MAC 地址。

(2) PE 1 接收到 ARP 请求后，根据 ARP 请求报文的源 MAC 地址，将 CE 1 的 MAC 地址学习到 CE 1 所属 VSI 实例的 MAC 地址表中。PE 1 在 VSI 实例内广播该 ARP 请求报文，为 ARP 请求报文添加 PW 标签和公网标签后，通过 PW 将其发送给 Border 设备。

(3) Border 设备接收到 ARP 请求报文后，执行如下操作：
   - 根据 PW 标签判断 ARP 请求报文所属的 VSI 实例，对 ARP 请求报文进行解封装，并根据 ARP 请求报文的源 MAC 地址，将 CE 1 的 MAC 地址学习到该 VSI 实例的 MAC 地址表中。
   - 通过 EVPN 的 MAC/IP 发布路由将 CE 1 的 MAC 地址通告给 PE 2。
   - 将从 UPW 上接收到的 ARP 请求报文当作为 AC 上的报文，在 UPW（即 AC）所属的 VSI 实例（根据 PW 标签判断所属 VSI 实例）内广播 ARP 请求，为 ARP 请求添加 End.DT2M SID 后，通过 SRv6 PW 将其发送给 PE 2。

(4) PE 2 接收到 EVPN 路由后，根据路由的 Route Target 属性判断路由对应的 VSI 实例，并将 CE 1 的 MAC 地址学习到该 VSI 实例的 MAC 地址表中。PE 2 接收到携带 End.DT2M SID 的 ARP 请求报文后，解封装该报文，并在 VSI 实例内广播该报文，将其发送给 CE 2。

(5) CE 2 单播回复 ARP 应答报文。

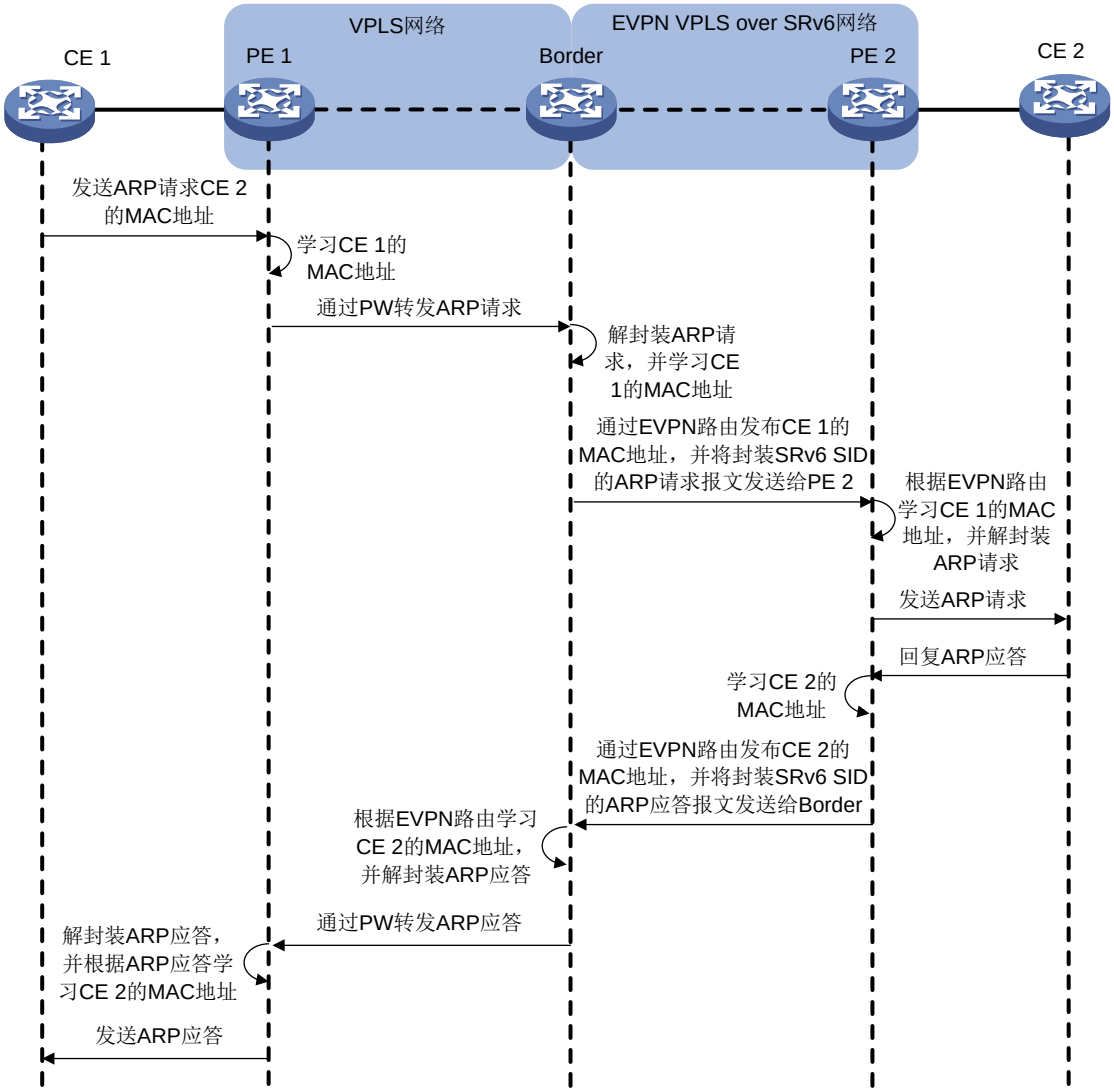
(6) PE 2 接收到 ARP 应答报文后，执行如下操作：
   - 根据 ARP 应答报文的源 MAC 地址，将 CE 2 的 MAC 地址学习到 CE 2 所属 VSI 实例的 MAC 地址表中。
   - 通过 EVPN 的 MAC/IP 发布路由将 CE 2 的 MAC 地址通告给 Border。
   - 根据 ARP 应答报文的目的 MAC 地址（CE 1 的 MAC 地址）查找 MAC 地址表项，找到出 SRv6 SID（Border 设备分配的 End.DT2U 或 End.DX2 SID）。PE 1 为 ARP 应答报文封装 SRv6 SID，通过 SRv6 PW 将其发送给 Border。

(7) Border 设备接收到 EVPN 路由后，根据路由的 Route Target 属性判断路由对应的 VSI 实例，并将 CE 2 的 MAC 地址学习到该 VSI 实例的 MAC 地址表中。Border 接收到携带 End.DT2U 或 End.DX2 SID 的 ARP 应答报文后，解封装该报文，并执行 SRv6 SID 标识的动作：
   - End.DT2U SID：解封装报文去掉 IPv6 报文头，并在 End.DT2U SID 所属的 VSI 内查找 MAC 地址表，根据查表结果，为报文添加 PW 标签和公网标签，将报文通过 PW 转发给 PE 1。
   - End.DX2 SID：解封装报文去掉 IPv6 报文头，并将报文转发给 End.DX2 SID 关联的 AC，即 UPW。为报文添加 PW 标签和公网标签，将报文通过 PW 转发给 PE 1。

(8) PE 1 从 PW 上接收到报文后，执行如下操作：

- 根据 PW 标签判断 ARP 应答报文所属的 VSI 实例，对 ARP 应答报文进行解封装，并根据 ARP 应答报文的源 MAC 地址，将 CE 2 的 MAC 地址学习到该 VSI 实例的 MAC 地址表中。
- 根据 ARP 应答报文的目的 MAC 地址（CE 1 的 MAC 地址）查找 MAC 地址表项，找到出接口为连接 CE 1 的 AC，通过该 AC 将 ARP 应答报文转发给 CE 1。

图62 VPLS 网络与 EVPN VPLS over SRv6 网络互通的 MAC 地址学习过程



9.2.3 报文转发过程

完成 PW 建立和 MAC 地址学习后，以图 63 的 SRv 6 BE 方式为例，CE 2 访问 CE 1 的报文转发过程为：

(1) CE 2 将访问 CE 1 的数据帧发送给 PE 2。

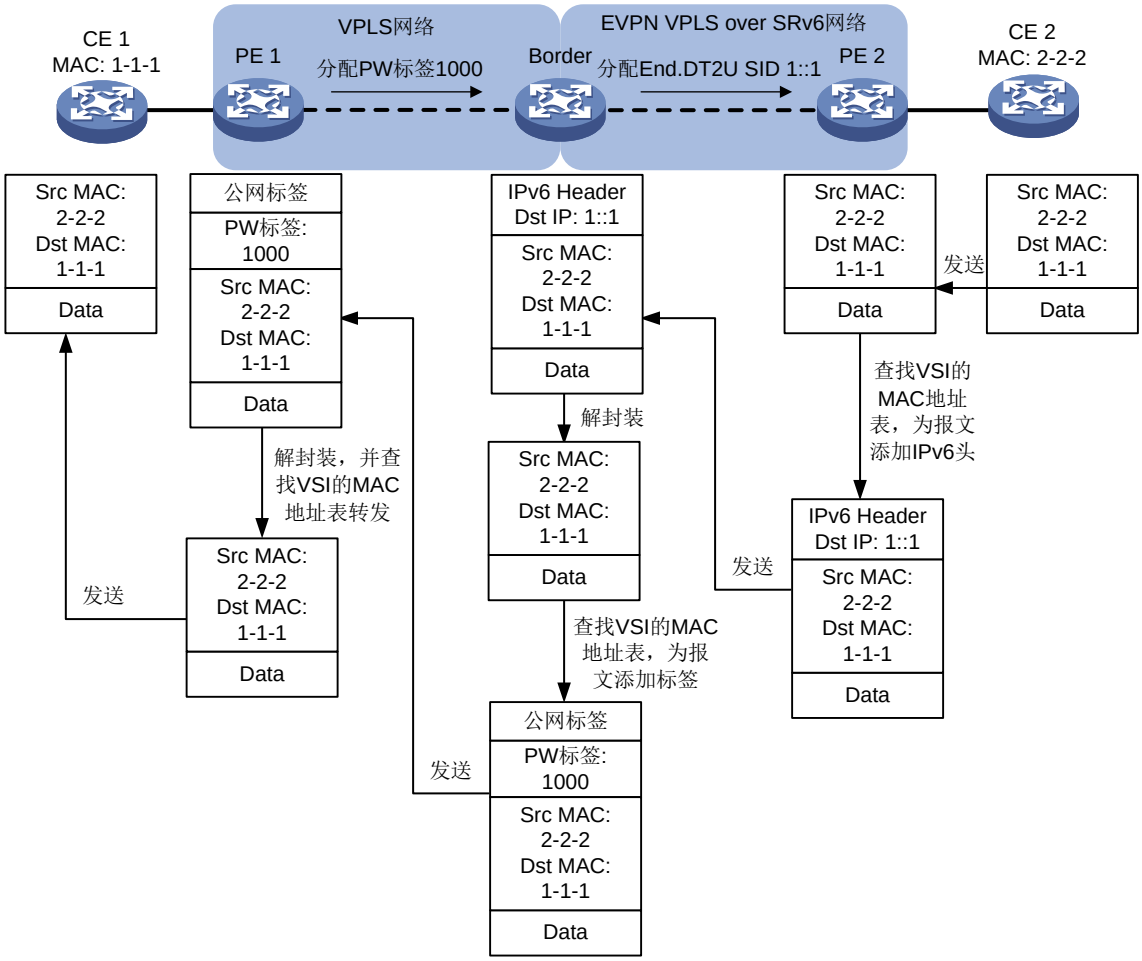
(2) PE 2 接收到数据帧后，在 VSI 实例内查找 MAC 地址表，找到 Border 分配的 End.DT2U SID。PE 2 为数据帧添加 IPv6 头，IPv6 头的目的地址为 End.DT2U SID 1::1。PE 2 根据目的 IPv6 地址 1::1 查找 IPv6 路由表，通过最优路由将其发送给 Border。

(3) Border 设备接收到携带 End.DT2U SID 的报文后，将其解封装，并在 End.DT2U SID 所属的 VSI 实例内查找 MAC 地址表，找到数据帧的出接口为看作 AC 的 UPW。

(4) Border 设备为数据帧添加 UPW 的 PW 标签 1000，并添加公网标签，将其转发给 PE 1。

(5) PE 1 根据 PW 标签 1000 判断数据帧所属的 VSI 实例，对报文进行解封装后，根据数据帧的目的 MAC 地址查找对应 VSI 的 MAC 地址表，找到出接口为连接 CE 1 的 AC，并通过该 AC 将报文转发给 CE 1。

图63 VPLS 网络与 EVPN VPLS over SRv6 网络互通的报文转发过程



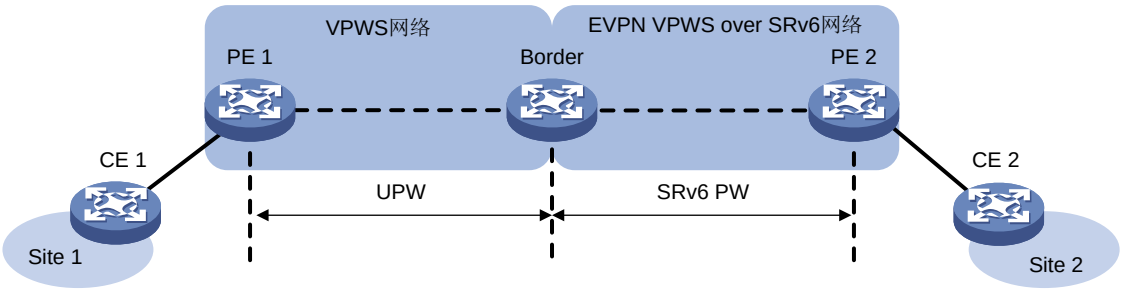
10 VPWS 与 EVPN VPWS/EVPN VPWS over SRv6 互通

10.1 概述

在实际组网中，可能会存在传统的 VPWS 网络与 EVPN VPWS（或 EVPN VPWS over SRv6）网络共存的情况。为了实现站点间流量跨越 VPWS 网络和 EVPN VPWS（或 EVPN VPWS over SRv6）网络传递，VPWS 网络与 EVPN VPWS（或 EVPN VPWS over SRv6）网络需要互通。

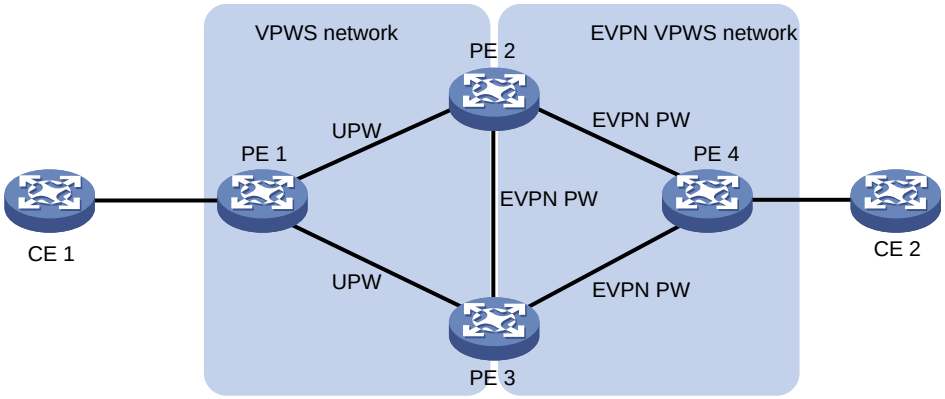
通过将 VPWS 网络中的 LDP PW 或静态 PW 看作 EVPN VPWS（或 EVPN VPWS over SRv6）网络的 AC（该 PW 称为 UPW），可以实现报文在 EVPN PW（或 SRv6 PW）与 UPW 之间相互转发，从而实现 VPWS 网络与 EVPN VPWS（或 EVPN VPWS over SRv6）网络的互通。

图64 VPWS 网络与 EVPN VPWS 网络互通示意图



VPWS 与 EVPN VPWS（或 EVPN VPWS over SRv6）互通组网不仅支持一条 LDP PW 或静态 PW 接入一条 EVPN PW（或 SRv6 PW），还支持将两条 LDP PW 或静态 PW 多归属接入两条 EVPN PW（或 SRv6 PW）。

图65 LDP PW 或静态 PW 多归属接入示意图



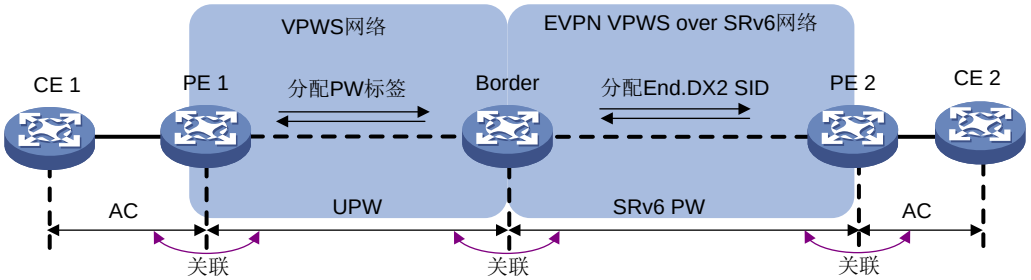
10.2 工作机制

VPWS 网络与 EVPN VPWS 网络互通、VPWS 网络与 EVPN VPWS over SRv6 网络互通的工作机制类似。下文将以 VPWS 网络与 EVPN VPWS over SRv6 网络互通为例，说明该工作机制。

10.2.1 建立 PW

如图 66 所示，VPWS 与 EVPN VPWS over SRv6 互通组网中，Border 设备同时运行 VPWS 和 EVPN VPWS over SRv6，与 PE 1 建立 VPWS 的 PW、与 PE 2 建立 SRv6 PW。其中，Border 和 PE 1 之间的 PW 为 UPW，在 EVPN VPWS over SRv6 网络中将该 UPW 当作 AC，将 UPW 与 SRv6 PW 关联，从而实现报文沿着 AC—UPW—SRv6 PW—AC 的路径，跨越 VPWS 网络与 EVPN VPWS over SRv6 网络转发。

图66 VPWS 与 EVPN VPWS over SRv6 互通组网中建立 PW



说明

PW 和 SRv6 PW 的建立方式与 VPWS 网络、EVPN VPWS over SRv6 网络单独部署时相同，本文不再赘述。

10.2.2 报文转发过程

完成 PW 建立、AC 与 PW 关联后，以图 67 的 SRv 6 BE 方式为例，CE 2 访问 CE 1 的报文转发过程为：

(1) CE 2 将访问 CE 1 的数据帧发送给 PE 2。

(2) PE 2 从 AC 接收到数据帧后，通过与 AC 关联的 SRv6 PW 转发该数据帧。SRv6 PW 的出 End.DX2 SID 为 1::1。PE 2 为数据帧添加 IPv6 头，IPv6 头的目的地址为 End.DX2 SID 1::1。PE 2 根据目的 IPv6 地址 1::1 查找 IPv6 路由表，通过最优路由将其发送给 Border。

(3) Border 设备接收到携带 End.DX2 SID 的报文后，将其解封装，并查找 End.DX2 SID 关联的 AC（即 UPW）。

(4) Border 设备为数据帧添加 UPW 的 PW 标签 1000，并添加公网标签，将其转发给 PE 1。

(5) PE 1 查找 PW 关联的 AC，并通过该 AC 将报文转发给 CE 1。

图67 VPWS 网络与 EVPN VPWS over SRv6 网络互通的报文转发过程

