MPLS VPN 原理与配置

MPLS（Multi-Protocol Label Switching）多协议标签交换
VPN　　（Virtual Private Network）虚拟专用网

==========

CE（Customer Edge）：用户网络边缘设备，有接口直接与服务提供商 SP（Service Provider）网络相连。CE 可以是 SVN 或交换机，也可以是一台主机。通常情况下，CE“感知”不到 VPN 的存在，也不需要支持 MPLS。
PE（Provider Edge）：服务提供商边缘设备，是服务提供商网络的边缘设备，与 CE 直接相连。在 MPLS 网络中，对 VPN 的所有处理都发生在 PE 上。
P（Provider）：服务提供商网络中的骨干设备，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。

传统 VPN 模型（Overlay VPN　，Peer to peer VPN）

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

3 个需要解决的问题：

1.本地路由冲突问题，即：在同一台 PE 上如何区分不同 VPN 的相同路由。

PE 设备怎么区分不同 VPN 客户的相同路由？

2.路由在网络中的传播问题，两条相同的路由，都在网络中传播，对于接收者如何分辨彼此？

冲突路由在公网中传播时，接收端 PE 如何正确导入 VPN 客户路由？

3.报文的转发问题，即使成功的解决了路由表的冲突，但是当 PE 接收到一个 IP 报文时，他又如何能够知道该发给那个 VPN？因为 IP 报文头中唯一可用的信息就是目的地址。而很多 VPN 中都可能存在这个地址。

PE 设备收到 IP 数据包后，如何正确的发送给目的 VPN 客户？

解决的方法：

1.本地路由冲突问题，可以通过在同一台路由器上创建不同的路由表解决，而不同的接口可以分属不同的路由表中，这就相当于将一台共享 PE 模拟成多台专用 PE。

可以通过在同一台 PE 设备上为不同的 VPN 建立单独的路由，这样冲突的的路由就被隔离开来；

VRF（VPN Routing and Forwarding table）VPN 路由转发表

2.在路由传递过程中，为不同的 VPN 路由添加不同的标识，以示区别。这些标识可以作为 BGP 属性进行传递；

RD　RT

增加了 RD 的 IPv4 地址称为 VPN-IPv4 地址，即 VPNv4 地址

RT（Route Target）封装在 BGP 的扩展 Community 属性中，

RD（Route Distinguisher）路由标识符
将 VPN 路由发布到全局路由表之前，使用一个全局唯一的标识和路由绑定，以区分冲突的私网路由。
RT（Route Target）路由目标
使用 RT 实现本端与对端的路由正确引入 VPN

3 由于 IP 报文不可更改，可以在 IP 报文头前加一些信息。由始发路由器打上标记，接收路由器在收到带标记的的数据包时，根据标记转发给正确的 VPN。

MP-BGP 分发内层标签

==========

RD (Route Distinguisher)路由区分符
用于标识 PE 上不同 VPN 实例，全局唯一，其主要作用是实现 VPN 实例之间地址复用，与 IP 地址一起构成 12 Bytes 的 VPNv4 地址。
RD 与路由一起被携带在 BGP Update 报文中发送给对端。
RD 不具有选路能力，不影响路由的发送与接受。
RD 用来区分本地 VRF，本地有效。

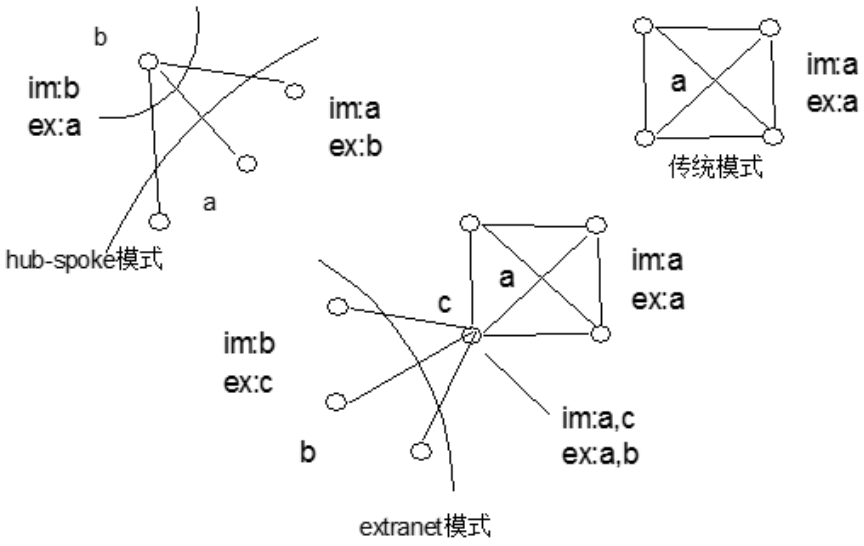
为了防止一台 PE 接收到远端 PE 发来的不同 VRF 的相同路由时不知所措，而加在路由前面的特殊信息。在 PE 发布路由时加上，在远端 PE 接收到路由后放在本地路由表中，用来与后来接收到的路由进行比较。

RT (Route Target) 路由目标
RT 是 VPNv4 路由携带的一个重要属性，它决定 VPN 路由的收发和过滤，PE 依靠 RT 属性区分不同 VPN 之间路由。

当从 VRF 表中导出 VPN 路由时，要用 Export RT 对 VPN 路由进行标记。
当往 VRF 表中导入 VPN 路由时，只有所带 RT 标记与 VRF 表中任意一个 Import RT 相符的路由才会被导入到 VRF 表中。
表明了一个 VRF 的路由喜好，通过他可以实现不同 VRF 之间的路由互通。他的本质就是 BGP 的 community 属性。

RT的灵活应用

由于每个RT Export Target与import Target都可以配置多个属性，例如：我对红色或者蓝色的路由都感兴趣。接收时是“或”操作，红色的、蓝色的以及同时具备两种颜色的路由都会被接受。所以就可以实现非常灵活的VPN访问控制。



从不同 PE 收到的相同路由靠 RD 区别
路由条目本端接受与否看 RT

举个生活的例子，RD 就是身份证，RT 就是护照。身份证只能有一张，护照可以很多张。
护照相同，就能进入相同的局域网。用在运营商的边界路由器上，RD 是一个 VRF 的身份证。RT 是这个 VRF 的护照，他可以导入很多不同的 RT。

==========

概念总结

VRF：在一台 PE 上虚拟出来的一个路由器，包括一些特定的接口，一张路由表，一个路由协议，一个 RD 和一组 RT 规则。

RD：为了防止一台 PE 接收到远端 PE 发来的不同 VRF 的相同路由时不知所措，而加在路由前面的特殊信息。在 PE 发布路由时加上，在远端 PE 接收到路由后放在本地路由表中，用来与后来接收到的路由进行比较。

RT：表明了一个 VRF 的路由喜好，通过他可以实现不同 VRF 之间的路由互通。他的本质就是 BGP 的 community 属性。

Label：为了防止一台 PE 接收到远端 PE 发给本地不同 VRF 的相同地址的主机时不知所措，而加在报文前面的特殊信息。由本地 PE 在发布路由时加上，远端 PE 接收到保存在相应的 VRF 中。

SITE：一个 VRF 加上与其相连的所有的 CE 的集合。

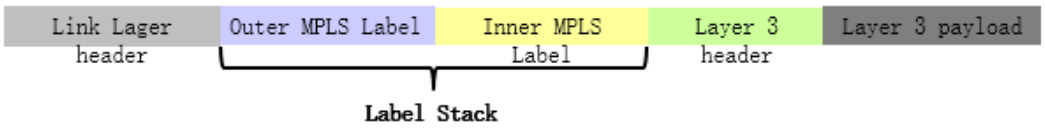
VPN：是一些 SITE 的集合，这些 SITE 由于共享了相同的路由信息可以互通。

==========

两层标签

内层标签由 MP-BGP 生成并在 VPNV4 的邻居中传递，用于区分不同的 VPN 流量；
外层标签由 MPLS 的 ldp 协议生成，用于解决传输的可达性问

题

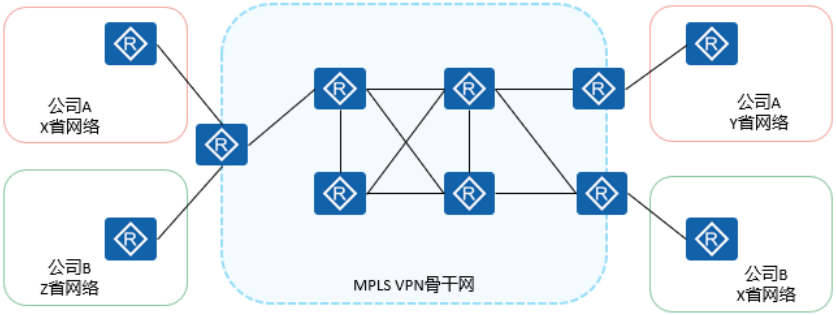


- 使用标签嵌套解决数据转发过程中冲突路由的查找问题。

- VPN（Virtual Private Network，虚拟专用网络）指的是在一个公共网络中实现虚拟的专用网络，从而使得用户能够基于该专用网络实现通信的技术。
- MPLS VPN 也是 VPN 技术中的一种。需要强调的是，本课程所介绍的 MPLS VPN 指的是 BGP/MPLS IP VPN，这是一种被业界广泛使用的三层 VPN。
- 本课程将介绍 MPLS VPN 的基本概念、工作过程以及典型配置方法。

MPLS VPN定义

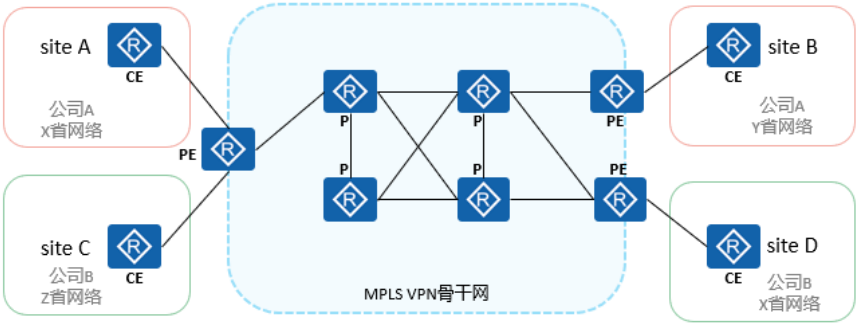
- BGP/MPLS IP VPN网络一般由**运营商**搭建，**VPN用户购买**VPN服务来实现用户网络之间的路由传递、数据互通等。
- MPLS VPN使用**BGP**在运营商骨干网（**IP网络**）上发布VPN路由，使用**MPLS**在运营商骨干网上转发VPN报文。BGP/MPLS IP VPN又被简称为MPLS VPN，是一种常见的L3VPN（Layer 3 VPN）技术。



- 注：MPLS VPN 的骨干网也可以由企业自行搭建，技术层面与运营商搭建基本一致，本课程仅讨论企业购买运营商 MPLS VPN 服务的场景。

MPLS VPN网络架构

- MPLS VPN网络架构由三部分组成：CE（Customer Edge）、PE（Provider Edge）和P（Provider），其中PE和P是运营商设备，CE是MPLS VPN用户设备。
- 站点（site）就是MPLS VPN的用户，由CE和其他用户设备构成。



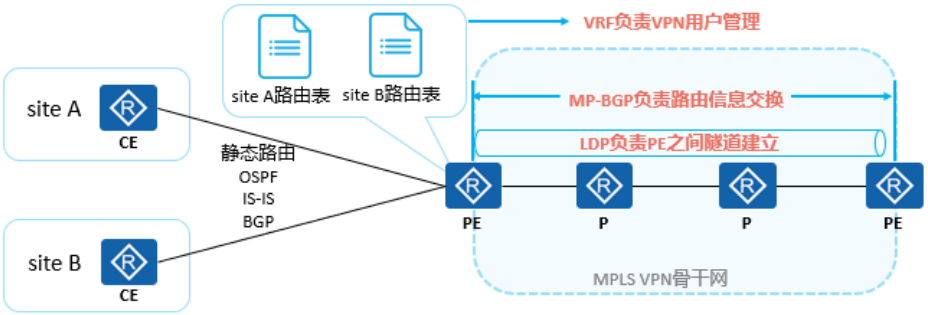
- CE：用户网络边缘设备，有接口直接与运营商网络相连。CE 可以是路由器或交换机，也可以是一台主机。通常情况下，CE“感知”不到 VPN 的存在，也不需要支持 MPLS。
- PE：运营商边缘路由器，是运营商网络的边缘设备，与 CE 直接相连。在 MPLS 网络中，对 VPN 的所有处理都发生在 PE 上，对 PE 性能要求较高。
- P：运营商网络中的骨干路由器，不与 CE 直接相连。P 设备只需要具备基本 MPLS 转发能力，不维护 VPN 相关信息。
- 站点的含义可以从下述几个方面理解：
- 站点是指相互之间具备 IP 连通性的一组 IP 系统，并且这组 IP 系统的 IP 连通性不需通过运营商网络实现。
- 站点的划分是根据设备的拓扑关系，而不是地理位置。如图所示，公司 A 的 X 省网络和公司 A 的 Y 省网络需要通过运营商的骨干网进行互联，所以它们被划分为两个站点。若在当前 X 省网络和 Y 省网络的 CE 之间增加一条物理专线，不需要通过运营商网络就可以互通，那么这两张网络就是一个站点。
- 站点与 VPN 的关系：
- 对于多个连接到同一服务提供商网络的站点，通过制定策略，可以将它们划分为不同的集合，只有属于相同集合的站

点之间才能通过服务提供商网络互访，这种集合就是 VPN。

- 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。
- 注：也有可能出现站点为一台主机的情况，此时该主机就是 CE 设备。本课程仅讨论站点是一个或多个子网，用路由器或交换机作为 CE 的情况。

MPLS VPN技术架构

- MPLS VPN不是单一的一种VPN技术，是多种技术结合的综合解决方案，主要包含下列技术：
  - MP-BGP：负责在PE与PE之间传递站点内的路由信息。
  - LDP：负责PE与PE之间的隧道建立
  - VRF：负责PE的VPN用户管理。
  - 静态路由、IGP、BGP：负责PE与CE之间的路由信息交换。



- MP-BGP（MultiProtocol BGP）：拓展的 BGP 协议，可提供对多种地址族的支持，后续课程将会详细介绍。

为什么要选择 MPLS VPN

- 对 VPN 客户而言：
- “感知”不到 VPN 的存在，不需要部署和维护 VPN，降低企业运维难度和成本。
- 一般部署在运营商的 MPLS VPN 专网上，有一定的安全性保障。
- 对于运营商而言：
- MPLS 在无连接的 IP 网络中增加了面向连接的控制平面，

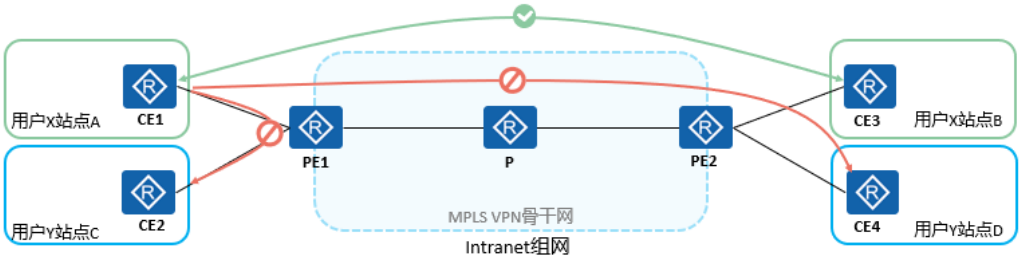
为 IP 网络增添了管理和运营的手段。

- 支持地址空间重叠、支持重叠 VPN、组网方式灵活、可扩展性好。
- 能够方便地支持 MPLS TE 合理调控现有网络资源，最大限度的节省运营商成本。

- MPLS TE（MPLS Traffic Engineering，MPLS 流量工程）：基于一定约束条件 LSP 隧道，并将流量引入到这些隧道中进行转发，使网络流量按照指定的路径进行传输。可以在不进行硬件升级的情况下对现有网络资源进行合理调配和利用，并对网络流量提供带宽和 QoS 保证，最大限度的节省成本。

MPLS VPN常见组网

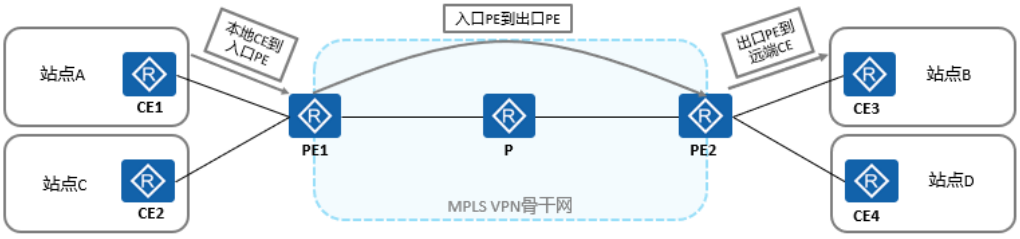
- 根据VPN用户的需求不同，可采用以下几种常见的组网方案：
  - Intranet：一个VPN中的所有用户形成闭合用户群，同一VPN站点之间可以互访，不同VPN站点间不能互访。
  - Extranet：适用于一个VPN用户希望提供部分本VPN的站点资源给其他VPN的用户访问的场景。
  - Hub&Spoke：如果希望在VPN中设置中心访问控制设备，其它用户的互访都通过中心访问控制设备进行，可采用Hub&Spoke组网方案。



- Intranet 组网是最简单也是最典型的 MPLS VPN 组网方案，后续课程将基于该组网方案对 MPLS VPN 技术原理展开介绍。

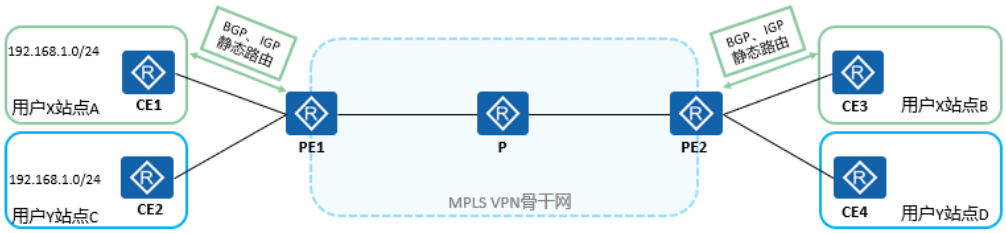
MPLS VPN路由发布概述

- 若想实现同一个VPN的不同站点之间的通信，首先需要完成不同站点之间的路由交互。在基本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。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



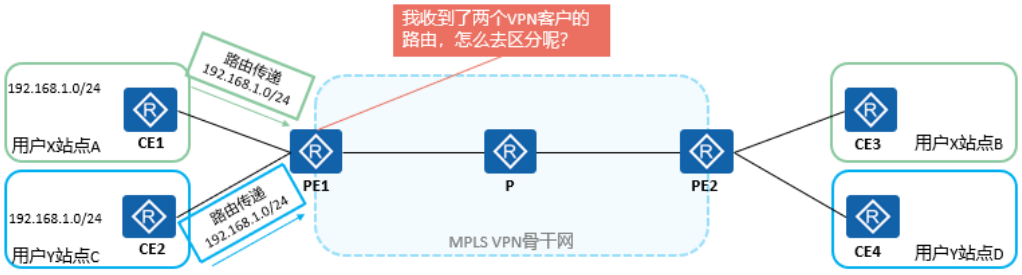
CE与PE之间的路由信息交换

- 如图，客户X和客户Y属于不同的VPN，分别拥有两个站点，现需要实现站点间的路由信息交互。
- CE与PE之间可以使用静态路由、OSPF、IS-IS或BGP交换路由信息。无论使用哪种路由协议，CE和PE之间交换的都是**标准的IPv4**路由。
- 本地CE到入口PE和出口PE到远端CE的路由信息交换原理完全相同。



入口PE到出口PE路由传递 (1)

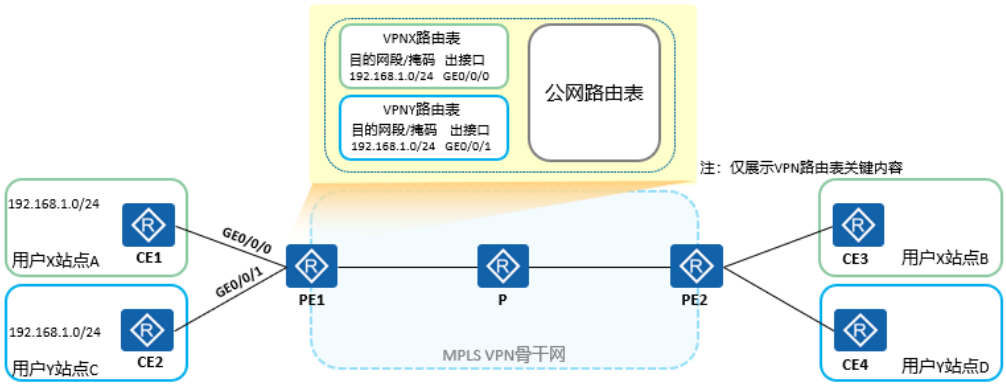
PE在接收到CE传递来的路由之后，需要独立保存不同VPN的路由，且需要解决不同的客户使用重叠IP地址空间的问题。



- VPN 是一种私有网络，不同的 VPN 独立管理自己的地址范围，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，例如图中用户 X 和用户 Y 都使用 192.168.1.0/24 网段地址，这就发生了地址空间的重叠（address spaces overlapping）。以下两种情况允许 VPN 使用重叠的地址空间：
- 两个 VPN 没有共同的站点；
- 两个 VPN 有共同的站点，但此共同站点中的设备不与两个 VPN 中使用重叠地址空间的设备互访。

VRF

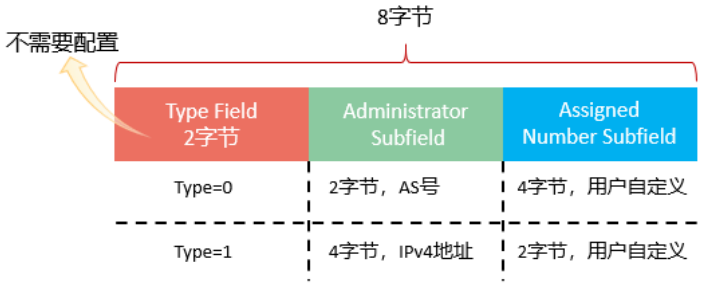
- VRF（Virtual Routing and Forwarding，虚拟路由转发），又称VPN实例，是MPLS VPN架构中的关键技术，每个VPN实例使用独立的路由转发表项，实现VPN之间的逻辑隔离。



- 有关 VRF 技术更详细的内容，请参考 HCIP-Datacom-Core 课程相关内容。

RD

- PE收到不同VPN的CE发来的IPv4地址前缀，本地根据VPN实例配置去区分这些地址前缀。但是VPN实例只是一个本地的概念，PE无法将VPN实例信息传递到对端PE，故有了RD（Route Distinguisher，路由标识符）。
  - RD长8字节，用于区分使用相同地址空间的IPv4前缀。
  - PE从CE接收到IPv4路由后，在IPv4前缀前加上RD，转换为全局唯一的VPN-IPv4路由。

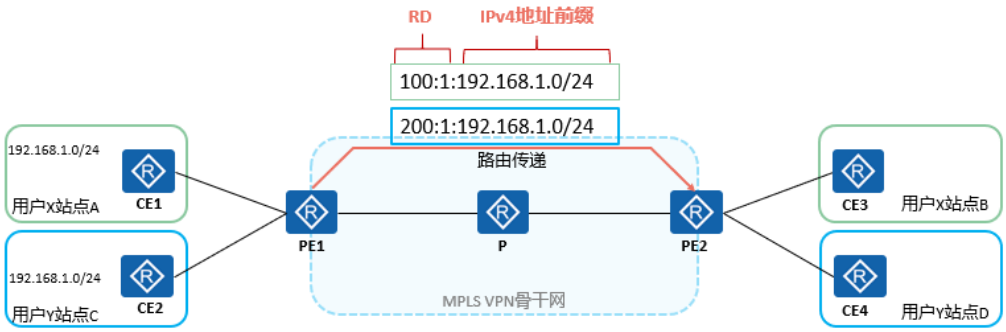


- 配置 RD 时，只需要指定 RD 的 Administrator 子字段和 Assigned Number 子字段。
- RD 的配置格式有四种，常用的两种如下：
- 16bits 自治系统号:32bits 用户自定义数字（例如：100:1）。
- 32bits IPv4 地址:16bits 用户自定义数字（例如：172.1.1.1:1）。

- RD 的结构使得每个运营商可以独立地分配 RD，但为了在某些应用场景下保证路由正常，必须保证 RD 全局唯一。

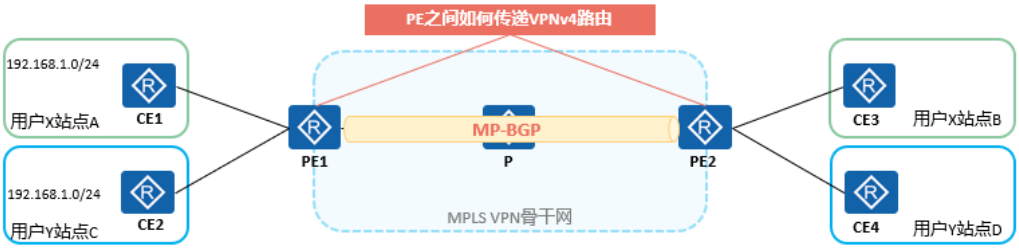
VPN-IPv4地址

VPN-IPv4地址又被称为VPNv4地址： VPNv4地址共有12个字节，包括8字节的路由标识符RD （Route Distinguisher）和4字节的IPv4地址前缀。



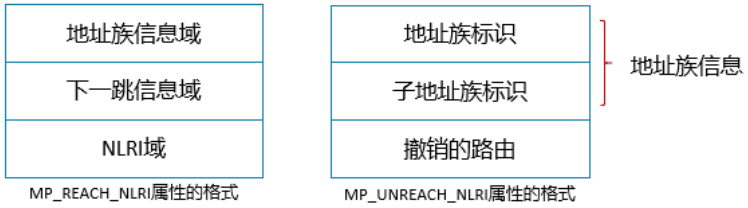
入口PE到出口PE路由传递 (2)

- PE之间建立BGP邻居关系，并通过BGP进行路由传递。为什么采用BGP呢?
  - BGP使用TCP作为其传输层协议，提高了协议的可靠性。可以跨路由器的两个PE设备之间直接交换路由。
  - BGP拓展性强，为PE间传播VPN路由提供了便利。
  - PE之间需要传送的路由条目可能较大，BGP只发送更新的路由，提高传递路由数量的同时不占用过多链路带宽。
- 传统的BGP-4**不支持**处理VPNv4路由。



MP-BGP

- 为了正确处理VPN路由，MPLS VPN使用RFC2858（Multiprotocol Extensions for BGP-4）中规定的MP-BGP，即BGP-4的多协议扩展。
- MP-BGP采用地址族（Address Family）来区分不同的网络层协议，既可以支持传统的IPv4地址族，又可以支持其它地址族（比如VPN-IPv4地址族、IPv6地址族等）。
- MP-BGP新增了两种路径属性：
  - MP_REACH_NLRI：Multiprotocol Reachable NLRI，多协议可达NLRI。用于发布可达路由及下一跳信息。
  - MP_UNREACH_NLRI：Multiprotocol Unreachable NLRI，多协议不可达NLRI。用于撤销不可达路由。

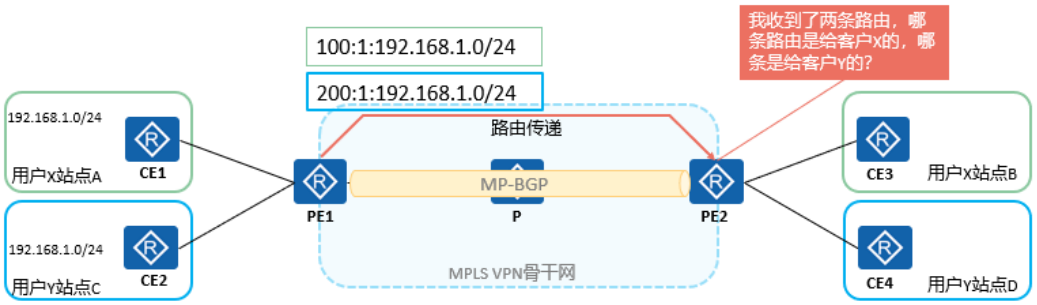


- NLRI：Network Layer Reachability Information，网络层可达信息。
- 关于地址族的一些取值请参考 RFC3232（Assigned Numbers）。
- MP_REACH_NLRI 用于发布可达路由及下一跳信息。该属性由一个或多个三元组<地址族信息、下一跳信息、网络可达性信息>组成，格式如下：
- 地址族信息（Address Family Information）域：由 2 字节的地址族标识 AFI（Address Family Identifier）和 1 字节的子地址族标识 SAFI（Subsequent Address Family Identifier）组成。
- AFI 标识网络层协议，对应 RFC3232 的“Address Family Number”所定义的地址族值。例如 IPv4 的值是 1，IPv6 的值是 2。
- SAFI 表示 NLRI 的类型。AFI 值为 1，SAFI 值为 128 表示 NLRI 中的地址为 MPLS-labeled VPN-IPv4 地址。
- 下一跳信息（Next Hop Network Address Information）域：由一字节的下一跳网络地址长度和可变长度的下一跳网络地址组成。

- 网络层可达性信息（NLRI）域：由一个或多个三元组<长度、标签、前缀>组成，该部分内容将在后面的课程里详细介绍。
- MP_UNREACH_NLRI 用于通知对等体删除不可达的路由。该属性的格式如下：
- 地址族标识 AFI：与 MP_REACH_NLRI 属性中的相同。
- 子地址族标识 SAFI：与 MP_REACH_NLRI 属性中的相同，表示 NLRI 的类型。
- 撤销路由（Withdrawn Routes）：不可达路由列表，也是由一个或多个 NLRI 组成。BGP 发言者可以通过在撤销路由域中携带与之前发布的可达路由中相同的 NLRI 来撤销路由。
- MP-BGP 的报文类型、VPNv4 路由发布策略仍与普通 BGP 相同。

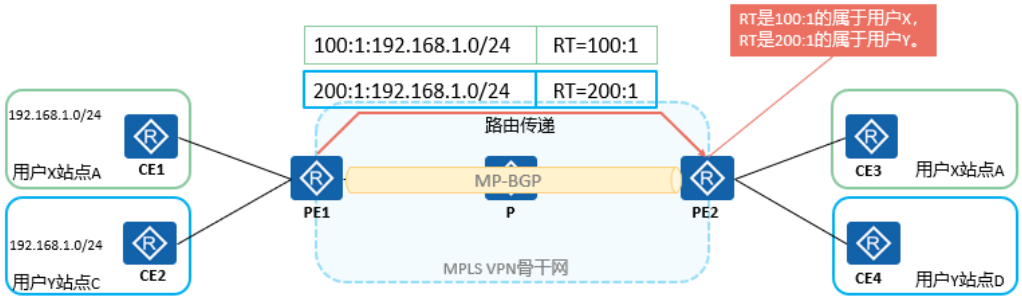
入口PE到出口PE路由传递 (3)

- MP-BGP将VPNv4传递到远端PE之后，远端PE需要将VPNv4路由导入正确的VPN实例。
- 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路由信息的发布与接收。
- 本地PE在发布VPNv4路由前附上RT属性，对端RT在接到到VPNv4路由后根据RT将路由导入对应的VPN实例。



RT

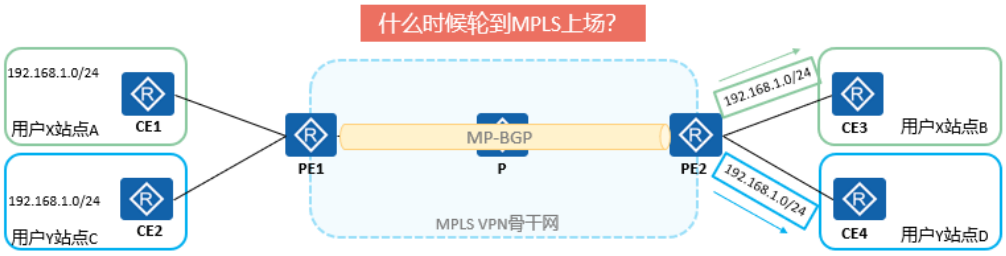
- 在PE上，每一个VPN实例都会与**一个或多个**VPN Target属性绑定，有两类VPN Target属性：
  - Export Target（ERT）：本地PE从直接相连站点学到IPv4路由后，转换为VPN IPv4路由，并为这些路由添加Export Target属性。Export Target属性作为BGP的扩展团体属性随路由发布。
  - Import Target（IRT）：PE收到其它PE发布的VPN-IPv4路由时，检查其Export Target属性。当此属性与PE上某个VPN实例的Import Target匹配时，PE就把路由加入到该VPN实例的路由表。



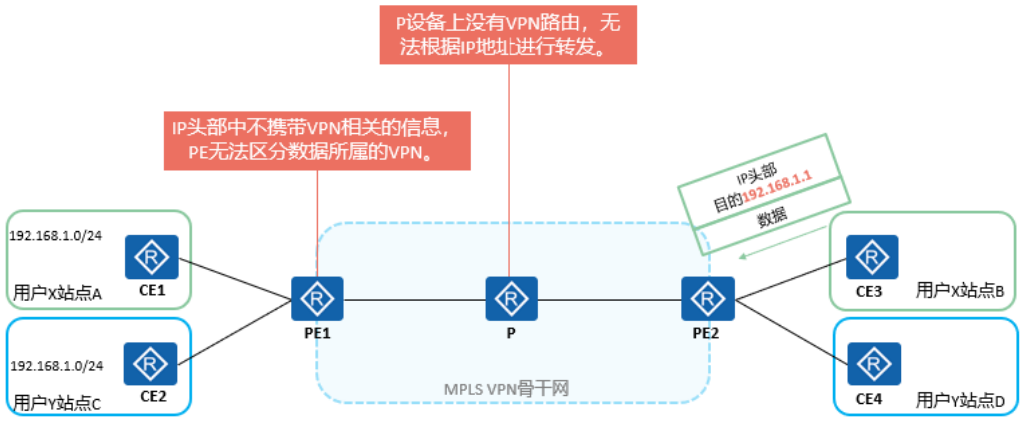
- 与 RD 相同，RT 由 Type、Administrator 和 Assigned Number 三个字段构成，长度也是 8 字节。
- 配置 VPN-Target 时，只需要指定 VPN-Target 的 Administrator 子字段和 Assigned Number 子字段。VPN-Target 的配置格式与 RD 格式一致。

入口PE到出口PE路由传递 (4)

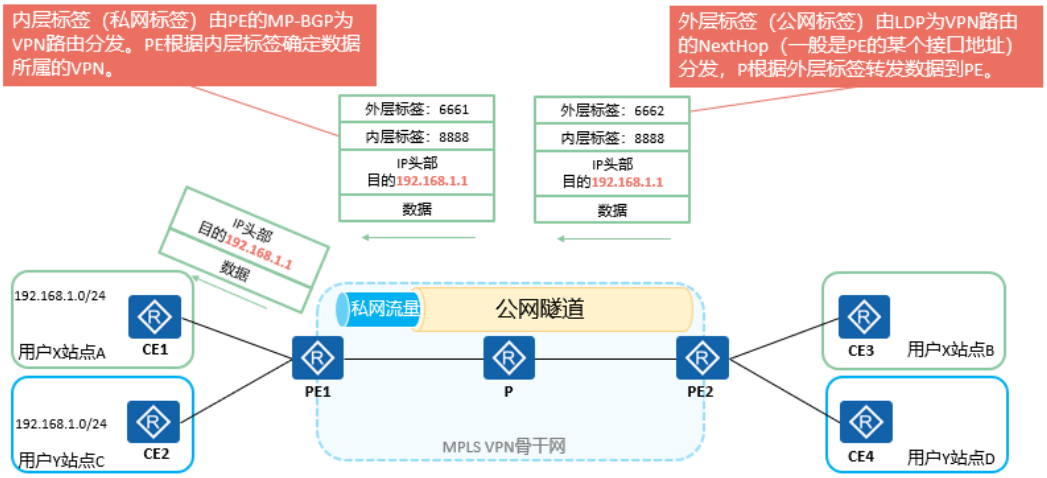
- PE根据VPNv4路由所携带的RT将路由导入正确的VPN实例之后，VPNv4路由的RD值剥除，将**IPv4路由**通告给相应的客户的CE设备。
- 站点B和站点D的CE设备就能学习到去往各自远端站点的路由。同理，通过一系列的操作，可以实现同一用户（同一VPN）不同站点之间的路由互通。



数据转发时遇到的问题

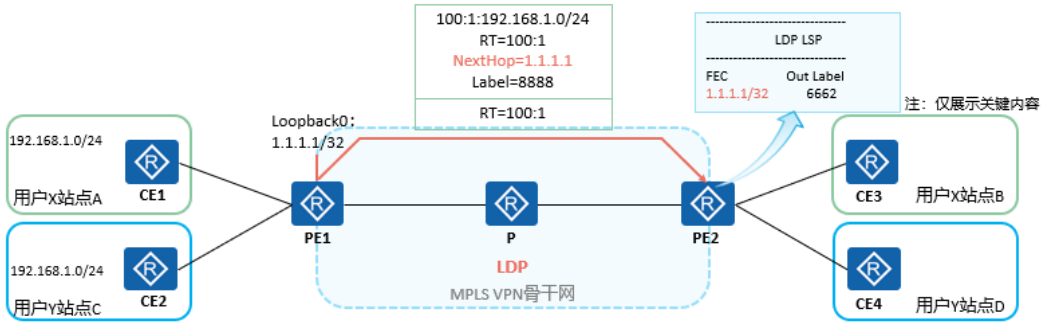


通过标签解决问题



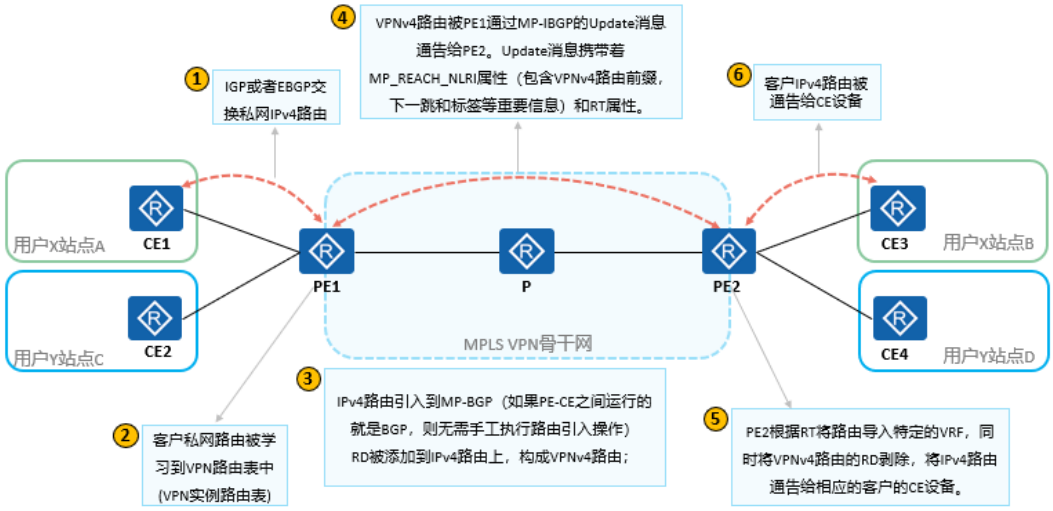
入口PE到出口PE路由传递 (5)

- PE和P设备之间运行LDP，交换公网标签，建立PE之间的LSP隧道（公网隧道）。
- 入口PE在通过MP-BGP传递VPNv4路由时，会携带私网标签，用于区分不同VPN的数据。
- 出口PE在接收到VPNv4路由后，需要执行私网路由交叉和隧道迭代来选择路由。



- PE 上分配私网标签的方法有如下两种：
- 基于路由的 MPLS 标签分配：为 VPN 路由表的每一条路由分配一个标签（one label per route）。这种方式的缺点是：当路由数量比较多时，设备入标签映射表 ILM（Incoming Label Map）需要维护的表项也会增多，从而提高了对设备容量的要求。
- 基于 VPN 实例的 MPLS 标签分配：为整个 VPN 实例分配一个标签，该 VPN 实例里的所有路由都共享一个标签。使用这种分配方法的好处是节约了标签。
- 私网路由交叉：VPNv4 路由与本地 VPN 实例的 VPN-Target 进行匹配的过程称为私网路由交叉。PE 在收到 VPNv4 路由后，既不进行优选，也不检查隧道是否存在，直接将其与本地的 VPN 实例进行交叉。
- 隧道迭代：为了将私网流量通过公网传递到另一端，需要有一条公网隧道承载这个私网流量。因此私网路由交叉完成后，需要根据目的 IPv4 前缀进行路由迭代，即该 IPv4 路由的下一跳有对应的 LSP 存在；只有隧道迭代成功，该路由才被放入对应的 VPN 实例路由表。

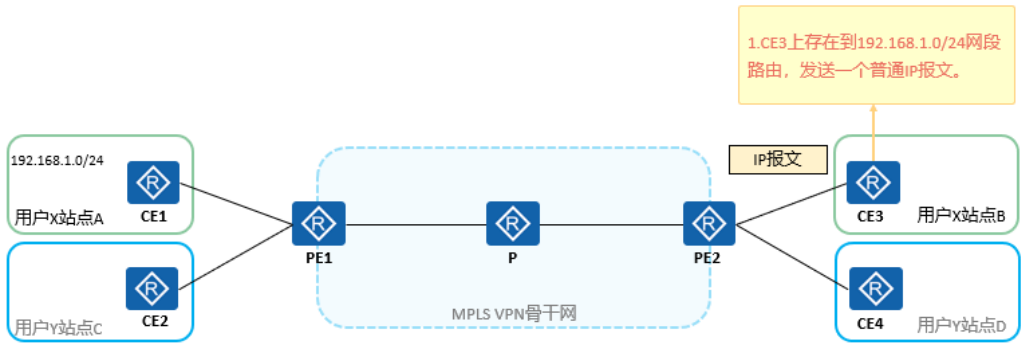
MPLS VPN中的路由交互全过程



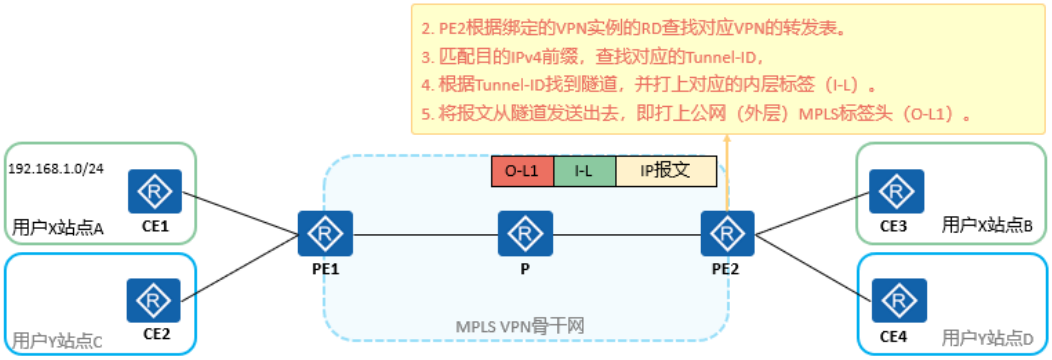
- 在 MPLS VPN 中，PE 与 CE，PE 与 PE 之间需要进行 VPN 路由信息的传递。
- PE 与 CE 之间可以采用 BGP、IGP 以及静态路由方式交互 IPv4 路由信息。
- PE 与 PE 之间通过 MP-BGP 交互 VPNv4 路由信息，包含
- RD：与 IPv4 前缀组合组成 VPNv4 前缀。
- RT：用于控制 PE 之间路由信息的接收和发布。
- 标签：数据转发时的内层（私网）标签，在 PE 上用来区分不同 VPN 的数据。

报文转发过程 (1)

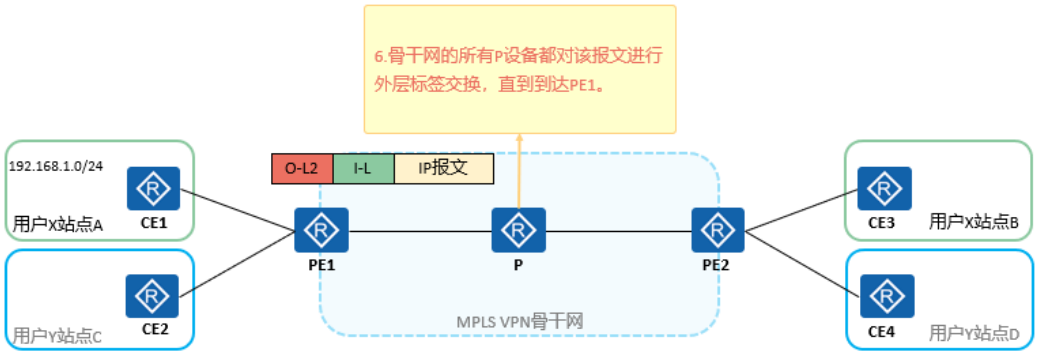
- 以图中用户X的站点B访问站点A的192.168.1.0/24网段为例，报文转发过程如下



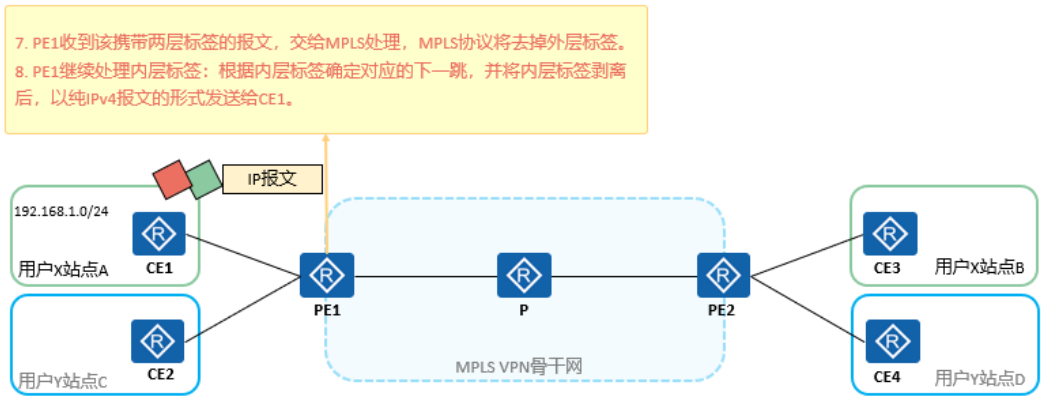
报文转发过程 (2)



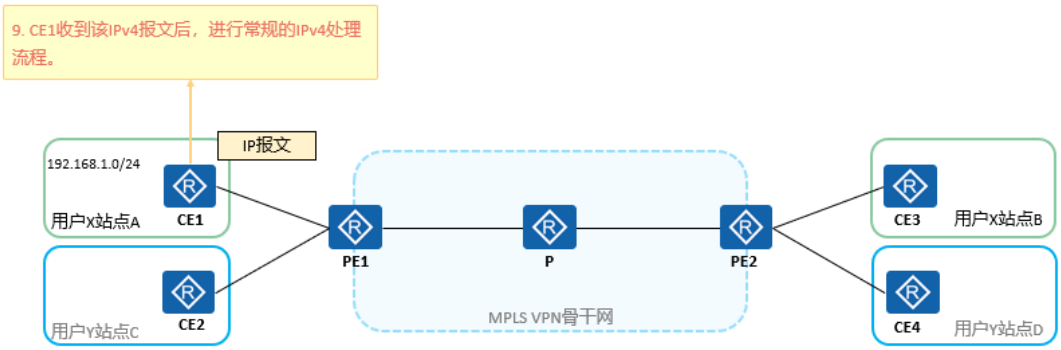
报文转发过程 (3)



报文转发过程 (4)



报文转发过程 (5)



配置命令 - VPN实例配置 (1)

1. 创建VPN实例/进入VPN实例视图

   [PE] **ip vpn-instance** *vpn-instance-name*

   缺省情况下，未配置VPN实例。

2. 使能VPN实例的IPv4地址族/进入VPN实例IPv4地址族视图

   [PE-vpn-instance-InstanceName] **ipv4-family**

   缺省情况下，未使能VPN实例的IPv4地址族。

3. 为VPN实例地址族配置路由标识RD

   [PE-vpn-instance-InstanceName] **route-distinguisher** *route-distinguisher*

   RD常见格式有两种：

   - 2字节自治系统号:4字节用户自定义数，例如100:1。
   - IPv4地址:2字节用户自定义数，例如192.168.122.15:1。

   无论采用哪种格式，必须保证RD值全局唯一。

   VPN实例地址族下一旦配置RD后，RD将不能被修改或删除。如果要修改，需要去使能VPN实例相应的地址族或者删除VPN实例，然后再重新配置。

配置命令 - VPN实例配置 (2)

4. 配置VPN实例的VPN-Target属性

   [PE-vpn-instance-InstanceName] **vpn-target** *vpn-target* &<1-8> [**both** | **export-extcommunity** | **import-extcommunity**]

   vpn-target命令用来配置VPN实例地址族入方向或出方向的VPN-Target扩展团体属性。

   - VPN-Target的格式与RD一致。
   - 一条**vpn-target**命令一次最多可配置8个VPN-Target。如果希望配置更多的VPN-Target，可以多次使用**vpn-target**命令。

5. 将接口绑定到VPN实例

   [PE-GigabitEthernet0/0/0]**ip binding vpn-instance** *vpn-instance-name*

   ip binding vpn-instance命令用来将PE上的接口与VPN实例绑定。缺省情况下，接口不与任何VPN实例绑定，属于根实例。配置接口与VPN实例绑定后，或取消接口与VPN实例的绑定，都会清除该接口的IP地址、三层特性和IP相关的路由协议，如果需要应重新配置。

配置命令 - MP-BGP配置

1. BGP基本配置

```
[PE] bgp { as-number-plain | as-number-dot }
[PE-bgp] peer ipv4-address as-number as-number
[PE-bgp] peer ipv4-address connect-interface loopback interface-number
```

PE之间必须使用32位掩码的Loopback接口地址来建立MP-IBGP对等体关系，以便能够迭代到隧道。

2. 使能与指定MP-BGP对等体之间交换VPNv4路由信息

```
[PE-bgp] ipv4-family vpnv4 [ unicast ]
[PE-bgp-af-vpnv4] peer ipv4-address enable
```

缺省情况下，只有BGP-IPv4单播地址族的对等体是自动使能的。即在BGP视图下配置**peer as-number**命令后，系统会自动配置相应的**peer enable**命令。其他地址族视图下都必须手动使能。

3. 配置对VPNv4路由进行过滤

```
[PE-bgp-af-vpnv4] policy vpn-target
```

policy vpn-target命令用来对接收到的VPN路由根据VPN-Target进行过滤。缺省情况下，该功能已经使能。在某些特定组网场景下，需要手动关闭过滤。

配置命令 - PE与CE间路由配置

1. PE与CE间采用EBGP传递路由

```
[PE-bgp] ipv4-family vpn-instance vpn-instance-name
[PE-bgp-InstanceName] peer ipv4-address as-number as-number
```

PE上需进入VPN实例IPv4地址族视图，并将CE配置为VPN私网对等体。

CE上配置与普通EBGP配置相同，并将VPN路由通过import或network的方式引入BGP。

2. PE与CE间采用IGP传递路由（以OSPF为例）

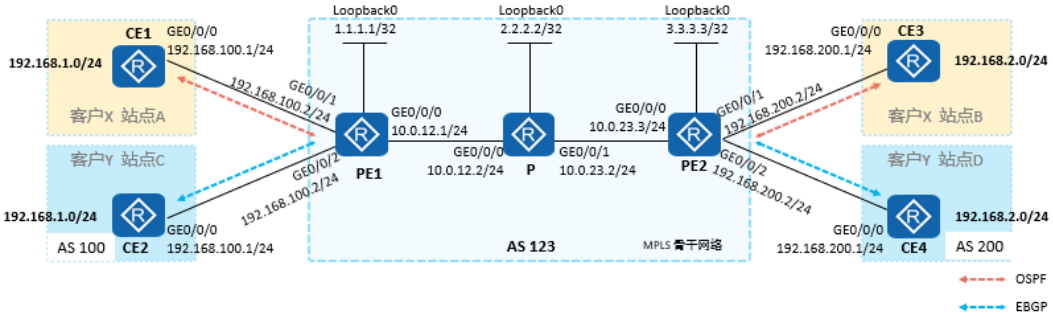
```
[PE] ospf process-id [ router-id router-id ] vpn-instance vpn-instance-name
[PE-ospf-processid] import-route bgp [ permit-ibgp ] [ cost cost | route-policy route-policy-name | tag tag | type type ] *
[PE-bgp] ipv4-family vpn-instance vpn-instance-name
[PE-bgp] import-route ospf process-id [ med med | route-policy route-policy-name ] *
```

PE上需要创建与VPN实例绑定的OSPF进程，并与CE建立OSPF邻居关系。同时需要在OSPF和BGP之间互相引入路由。

CE上配置与普通OSPF配置相同。

MPLS VPN配置示例 - 背景介绍

- 客户X及Y各自有2个站点，现需要通过MPLS VPN实现站点之间的互联，分别对应VPNX和VPNY；
- 互联接口、AS号及IP地址信息如图；
- 客户X站点与PE之间采用OSPF交互路由信息，客户Y站点与PE之间采用BGP交互路由信息。



MPLS VPN 配置示例 - 配置思路

- MPLS VPN 骨干网配置

  1.1 IGP 配置，实现骨干网的 IP 连通性。

  1.2 MPLS 与 MPLS LDP 配置，建立 MPLS LSP 公网隧道，传输 VPN 数据。

  1.3 MP-BGP 配置，建立后续传递 VPNv4 路由的 MP-BGP 对等体关系。

- VPN 用户接入配置

  2.1 创建 VPN 实例并配置参数（RT、RD）

  2.2 将接口加入 VPN 实例

  2.3 配置 PE 与 CE 之间的路由交换

MPLS VPN配置示例 - 数据规划

- MPLS骨干网采用单区域OSPF实现路由互通，所有PE和P互联接口均使能MPLS LDP功能。
- PE上的VPN相关配置如表格：

配置项	描述			
	PE1		PE2	
VPN名称	VPNX	VPNY	VPNX	VPNY
RD	100:1	200:1	100:1	200:1
IRT	100:321	200:234	100:321	200:432
ERT	100:123	200:432	100:123	200:234
接口	GE0/0/1	GE0/0/2	GE0/0/1	GE0/0/2
MP-BGP	源接口：Loopback0		源接口：Loopback0	

MPLS VPN骨干网配置 (1)

1.1 在MPLS VPN骨干网络内部署OSPF，MPLS VPN骨干网络内部署的OSPF用于实现骨干网络内部的路由互通。

以PE1节点的OSPF配置为例。

```
[PE1]ospf 100 router-id 1.1.1.1
[PE1-ospf-100]area 0
[PE1-ospf-100-area-0.0.0.0]network 10.0.12.1 0.0.0.0
[PE1-ospf-100-area-0.0.0.0]network 1.1.1.1 0.0.0.0
```

1.2 在PE1、P、PE2节点配置MPLS及LDP，以PE1为例。

```
[PE1]mpls lsr-id 1.1.1.1
[PE1]mpls
Info: Mpls starting, please wait... OK!
[PE1-mpls]mpls ldp
[PE1-mpls-ldp]Interface GigabitEthernet 0/0/0
[PE1-GigabitEthernet0/0/0]mpls
[PE1-GigabitEthernet0/0/0]mpls ldp
```

MPLS VPN骨干网配置 (2)

1.3 在PE1及PE2之间建立MP-BGP对等体关系，以PE1为例。

```
[PE1]bgp 123
[PE1-bgp]router-id 1.1.1.1
[PE1-bgp]peer 3.3.3.3 as-number 123
[PE1-bgp]peer 3.3.3.3 connect-interface LoopBack 0
#进入BGP-VPNv4地址族视图，并使能与对等体3.3.3.3的VPNv4地址族能力。
[PE1-bgp]ipv4-family vpnv4 unicast
[PE1-bgp-af-vpnv4]peer 3.3.3.3 enable
```

- 缺省情况下，只有 BGP-IPv4 单播地址族的对等体是自动使能的。即在 BGP 视图下配置 peer as-number 命令后，系

统会自动配置相应的 peer enable 命令。其他地址族视图下都必须手动使能。

MPLS VPN骨干网配置 - 配置验证

查看公网隧道建立情况

```
[PE1]display mpls lsp
------------------------------------------------------------------------------
                    LSP Information: LDP LSP
------------------------------------------------------------------------------
FEC                 In/Out Label  In/Out IF       Vrf Name
3.3.3.3/32          NULL/1025     -/GE0/0/0
1.1.1.1/32          3/NULL        -/-
```

```
[PE2]display mpls lsp
------------------------------------------------------------------------------
                    LSP Information: LDP LSP
------------------------------------------------------------------------------
FEC                 In/Out Label  In/Out IF       Vrf Name
3.3.3.3/32          3/NULL                       -/-
1.1.1.1/32          NULL/1024     -/GE0/0/0
```

查看MP-BGP邻居状态，以PE1为例

```
[PE1]display bgp vpnv4 all peer
BGP local router ID : 1.1.1.1
Local AS number : 123
Total number of peers : 1                 Peers in established state : 1

 Peer        V          AS  MsgRcvd  MsgSent  OutQ    Up/Down       State Pre              fRcv
 3.3.3.3     4         123  16        18        0 00:14:20  Established  0
```

VPN用户接入配置 (1)

2.1 创建VPN实例并按照规划配置RD与RT参数，以PE1为例

```
[PE1]ip vpn-instance VPNX
[PE1-vpn-instance-VPNX]route-distinguisher 100:1
[PE1-vpn-instance-VPNX-af-ipv4] vpn-target 100:321 import-extcommunity
 IVT Assignment result:
Info: VPN-Target assignment is successful.
[PE1-vpn-instance-VPNX-af-ipv4] 100:123 export-extcommunity
 EVT Assignment result:
Info: VPN-Target assignment is successful.
[PE1-vpn-instance-VPNX-af-ipv4] quit
[PE1-vpn-instance-VPNX]quit
[PE1]ip vpn-instance VPNY
[PE1-vpn-instance-VPNY]route-distinguisher 200:1
[PE1-vpn-instance-VPNY-af-ipv4]vpn-target 200:234 import-extcommunity
[PE1-vpn-instance-VPNY-af-ipv4]vpn-target 200:432 export-extcommunity
[PE1-vpn-instance-VPNY-af-ipv4]quit
[PE1-vpn-instance-VPNY]quit
```

VPN用户接入配置 (2)

2.2 将接口绑定到VPN实例。

```
[PE1]interface GigabitEthernet 0/0/1
[PE1-GigabitEthernet0/0/1]ip binding vpn-instance VPNX
Info: All IPv4 related configurations on this interface are removed!
Info: All IPv6 related configurations on this interface are removed!
[PE1-GigabitEthernet0/0/1]ip address 192.168.100.2 24
[PE1-GigabitEthernet0/0/1]interface GigabitEthernet 0/0/2
[PE1-GigabitEthernet0/0/2]ip binding vpn-instance VPNY
Info: All IPv4 related configurations on this interface are removed!
Info: All IPv6 related configurations on this interface are removed!
[PE1-GigabitEthernet0/0/2]ip address 192.168.100.2 24
```

VPN用户接入配置 (3)

2.3.1 部署CE1-PE1、CE3-PE2间的路由信息交互，以PE1为例

```
#创建与实例绑定的OSPF进程
[PE1]ospf 2 vpn-instance VPNX
[PE1-ospf-2]area 0
[PE1-ospf-2-area-0.0.0.0]network 192.168.100.0 0.0.0.255
[PE1-ospf-2-area-0.0.0.0]quit
```

```
#配置OSPF进程与MP-BGP之间的路由双向引入
[PE1]ospf 2 vpn-instance VPNX
[PE1-ospf-2]import-route bgp
[PE1-ospf-2]quit
[PE1]bgp 123
[PE1-bgp]ipv4-family vpn-instance VPNX
[PE1-bgp-VPNX]import-route ospf 2
```

2.3.2 部署CE2-PE1、CE4-PE2间的路由信息交互，以CE2和PE1为例

```
#配置CE2上的EBGP，并引入直连路由192.168.1.0/24
[CE2]BGP 200
[CE2-bgp]peer 192.168.100.2 as-number 123
[CE2-bgp]network 192.168.1.0 24
[CE2-bgp]quit
```

```
#配置PE1上VPN实例的EBGP对等体
[PE1]bgp 123
[PE1-bgp]ipv4-family vpn-instance VPNY
[PE1-bgp-VPNY]peer 192.168.100.1 as-number 200
```

配置验证 (1)

查看VPNX用户的CE路由学习情况

```
[CE1]display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

192.168.1.0/24      Direct  0    0           D   192.168.1.254   GigabitEthernet0/0/1
192.168.2.0/24      OSPF    10   4           D   192.168.100.2   GigabitEthernet0/0/0
192.168.100.0/24    Direct  0    0           D   192.168.100.1   GigabitEthernet0/0/0
192.168.200.0/24    O_ASE   150  1           D   192.168.100.2   GigabitEthernet0/0/0
```

```
[CE3]dis ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

192.168.1.0/24      OSPF    10   4           D   192.168.200.2   GigabitEthernet0/0/0
192.168.2.0/24      Direct  0    0           D   192.168.2.254   GigabitEthernet0/0/1
192.168.100.0/24    O_ASE   150  1           D   192.168.200.2   GigabitEthernet0/0/0
192.168.200.0/24    Direct  0    0           D   192.168.200.1   GigabitEthernet0/0/0
```

配置验证 (2)

查看VPNY用户的CE路由学习情况

```
[CE2]display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface
192.168.1.0/24      Direct  0    0           D   192.168.1.254   GigabitEthernet0/0/1
192.168.2.0/24      EBGP    255  0           D   192.168.100.2   GigabitEthernet0/0/0
192.168.100.0/24    Direct  0    0           D   192.168.100.1   GigabitEthernet0/0/0
```

```
[CE4]display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface
192.168.2.0/24      Direct  0    0           D   192.168.1.254   GigabitEthernet0/0/1
192.168.1.0/24      EBGP    255  0           D   192.168.100.2   GigabitEthernet0/0/0
192.168.200.0/24    Direct  0    0           D   192.168.100.1   GigabitEthernet0/0/0
```

配置验证 (3)

```
[PE2] display bgp vpnv4 vpn-instance VPNX routing-table 192.168.1.0 24

BGP local router ID : 3.3.3.3
Local AS number : 123
VPN-Instance VPNX, Router ID 3.3.3.3:
Paths:   1 available, 1 best, 1 select
BGP routing table entry information of 192.168.1.0/24:
Label information (Received/Applied): 1026/NULL
From: 1.1.1.1 (1.1.1.1)
Relay token: 0x1
Original nexthop: 1.1.1.1
```

```
[PE2]display mpls lsp
------------------------------------------------------------------------------
LSP Information: LDP LSP
------------------------------------------------------------------------------
FEC             In/Out Label  In/Out IF       Vrf Name
1.1.1.1/32                    NULL/1024       -/GE0/0/0
1.1.1.1/32                    1024/1024       -/GE0/0/0
```

以192.168.2.0/24网段到192.168.1.0/24网段的数据为例，外层标签为1024，由MPLS LDP分配。内层标签为1026，由MP-BGP分配。

```
No.     Time        Source                Destination           Protocol Length Info
5       12.109000   192.168.2.254192.168.1.254ICMP      102             Echo (ping) request
Frame 5: 102 bytes on wire (816 bits), 102 bytes captured (816 bits) on interface 0
Ethernet II, Src: HuaweiTe_b1:15:3e (00:e0:fc:b1:15:3e), Dst: HuaweiTe_49:20:bb (00:e0:fc:49:20:bb)
MultiProtocol Label Switching Header, Label: 1024, Exp: 0, S: 0, TTL: 254
MultiProtocol Label Switching Header, Label: 1026, Exp: 0, S: 1, TTL: 254
Internet Protocol Version 4, Src: 192.168.2.254, Dst: 192.168.1.254
Internet Control Message Protocol
```

思考题：

- （单选）MP-BGP 在传递 VPNv4 路由时，携带哪种 Route Target ？（ ）

  A. Export RT

  B. Implied RT

  C. Import RT

  D. Extended RT

- （多选）基本 MPLS VPN 组网中，需要对 VPN 实例做哪些配置（ ）

- Import RT
- Export RT
- 配置 RD
- 配置与 VPN 实例绑定的接口

参考答案：

- A
- ABCD
-