MPLS VPN 部署与应用

BGP 协议 SoO（Site of Origin）

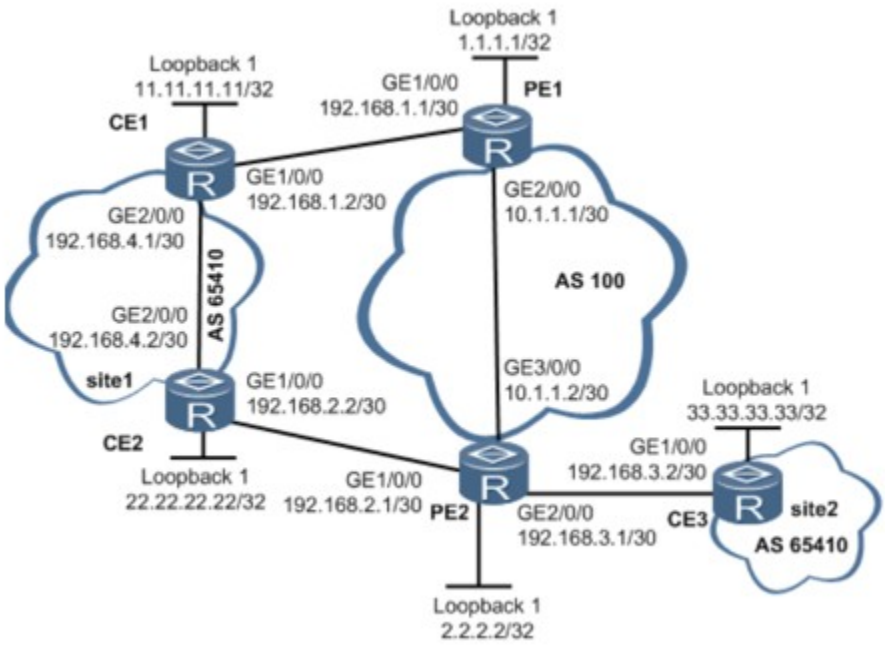
在 BGP/MPLS IP VPN 场景中，如果两个 VPN 站点所处的自治系统使用的是私有 AS 号，可能会出现两个 VPN 站点的 AS 号相同的情况，进而导致同一 VPN 的不同站点之间无法连通。此时需要在 PE 上执行 peer substitute-as 命令使能 AS 号替换功能。

但是使能 AS 号替换功能又会产生另外一个问题。当同一 VPN 站点内的多个 CE 通过 EBGP 接入 BGP/MPLS IP VPN 骨干网的不同 PE 设备，且 CE 间已部署路由协议时，如果 PE 上配置了 AS 号替换功能，则此 VPN 站点的私网路由在 PE 上将会被替换 AS 号，然后可能经过骨干网又通过其它 CE 发送回该 VPN 站点内，产生路由环路。此时可以在 PE 上执行命令 peer soo，对指定 CE 配置 SoO 特性来解决这个问题。

在 PE 上配置该命令后，PE 会为从 CE 收到的路由添加 SoO 属性并发布给其他的 PE 对等体。其他 PE 对等体向接入的 CE 发布这些路由时会检查 VPN 路由携带的 SoO 属性，如果与本地配置的 SoO 属性相同，PE 则不会向 CE 发布该路由，避免了 VPN 站点内路由环路。

注意事项

peer soo 命令仅用于 PE 与 CE 之间建立 EBGP 对等体关系的场景。



CE1、CE2 所在的站点相同，需要在 PE1 和 PE2 上分别针对 CE1 和 CE2 配置相同的 BGP SoO 属性。PE2 上连接了两个 VPN 站点，需要针对不同的 CE 配置不同的 SoO 属性。

```
[PE1] bgp 100
[PE1-bgp] ipv4-family vpn-instance vpna
[PE1-bgp-vpna] peer 192.168.1.2 soo 100:101
```

```
[PE2] bgp 100
[PE2-bgp] ipv4-family vpn-instance vpna
[PE2-bgp-vpna] peer 192.168.2.2 soo 100:101
[PE2-bgp-vpna] peer 192.168.3.2 soo 100:102
```

OSPF 协议 Sham link

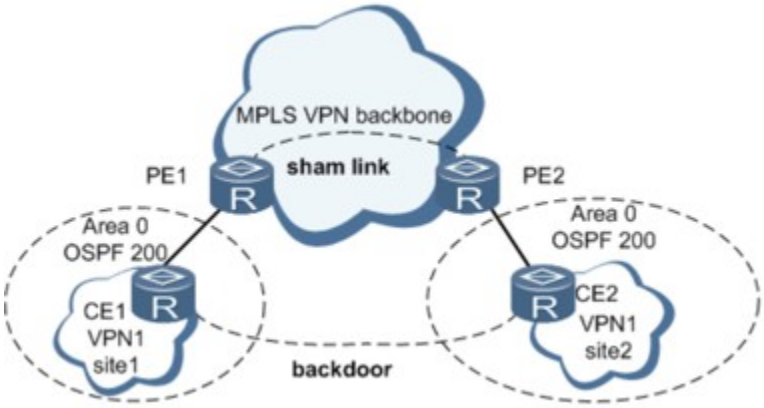
本命令仅在 VPN 场景下使用。

在 OSPF 协议实际应用中，如果两个要互通的 Site 都在相同的 AS 内，那么每个 Site 都应该将另一个 Site 的路由看成区域间路由，而不是 AS 外部路由

使用 sham-link 命令创建伪连接，使 VPN 流量优先经过 VPN 骨干区域的路由，避免在同一个 OSPF 区域内属于同一个 VPN 之间的通信总是通过 OSPF 区域内路由转发。

通常情况下，BGP 对等体之间通过 BGP 扩展团体属性在 MPLS VPN 骨干网上承载路由信息。另一端 PE 上运行的 OSPF 可利用这些信息来生成 PE 到 CE 的 Type 3 LSA，这些路由是区域间路由（Inter_Area route）。
若在 CE1 和 CE2 之间增加一条后门（Backdoor）链路，并且直接运行 OSPF 交互路由。通过后门链路学习到的路由类型为区域内路由（Intra_Area route）。
由于区域内路由优于区域间路由，故后门链路会被优选，若想实现后门链路作为备份链路，可采用 sham link 实现。

Sham link 在两台 PE 之间创建了一条区域内链路。当 LSA 在伪装链路中泛洪，所有的 OSPF 路由类型都不会改变，不会转换成 LSA3 或者 LSA5 的类型。
Sham link 被看成是两个 VPN 实例之间的链路，链路的两端是 PE 上的端点地址，分别作为建立连接时的源和目的地址。伪连接的源地址和目的地址使用 32 位掩码的 Loopback 接口地址，该 Loopback 接口需要绑定到 VPN 实例中，并通过 BGP 发布。



ospf
area 0
sham-link 1.1.1.1 2.2.2.2

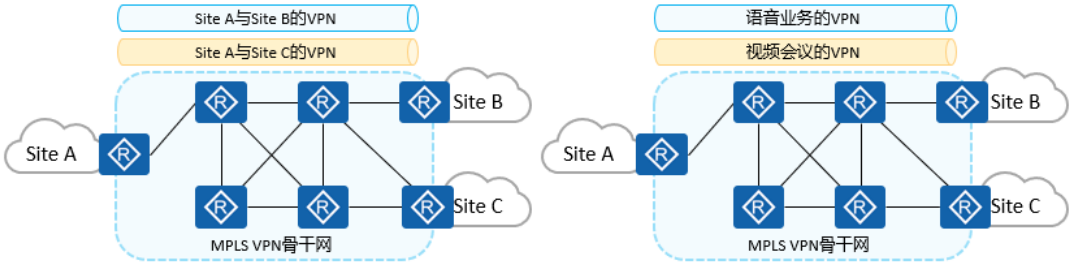
OSPF 伪连接仅应用在属于同一个 OSPF 区域的两个 Site 间存在后门链路的情况，如果 Site 间没有后门链路，则不需要配置 OSPF 伪连接。

==========

- BGP/MPLS IP VPN 因其支持地址空间重叠、组网方式灵活、可扩展性好，并能够方便地支持 MPLS TE 等一系列优点，已经在广域 IP 承载网络得到了广泛的应用。
- 针对不同客户的业务需求以及组网情况，MPLS VPN 的部署方式不尽相同。
- 本课程将介绍几种常见的 MPLS VPN 应用场景与这些场景下 MPLS VPN 的部署方法，此外还将介绍 OSPF 对 MPLS VPN 的扩展特性与功能。

MPLS VPN典型应用

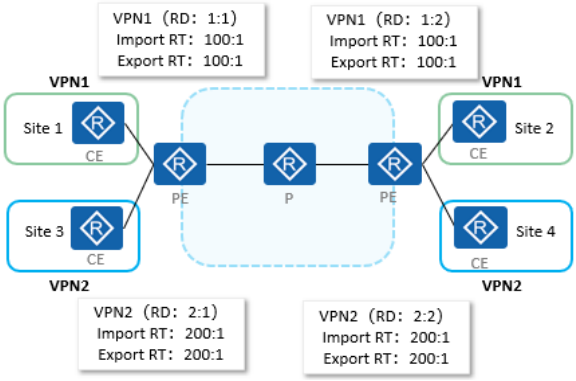
- 目前，MPLS VPN的主要应用包括企业互连和虚拟业务网络。
  - 企业互连应用：可通过MPLS VPN将分布在各地的分支机构、出差员工和合作伙伴的IP网络连接在一起；
  - 虚拟业务网络：可在同一物理网络上运行多种业务，如VoIP、IPTV等，为每个业务建立一个VPN，实现业务隔离。



- MPLS VPN 的主要优点包括但不限于以下几项：
- 可以实现“一点接入，全网连通”，支持异种介质的互连。而不像传统专线那样在每一对用户设备间采用同样的介质连接，可方便地提供普遍服务。
- 可以实现“弹性带宽”，采用流量监管技术，在保证用户基本带宽的同时，对突发流量尽力而为，同时基本带宽也可以“软扩容”，即根据用户的需求在一个范围内连续选择。
- 在资源隔离或隧道绑定的 MPLS VPN 技术保证下，充分保证每个 VPN 的专有带宽，满足各类业务有不同的用户，不同的流量模型，不同的 QoS 要求。

MPLS VPN基本组网 - Intranet

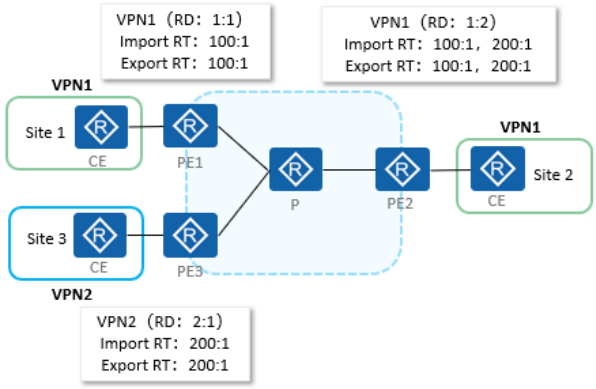
当采用Intranet组网方案时，一个VPN中的所有用户形成闭合用户群，相互之间能够进行流量转发，VPN中的用户不能与任何本VPN以外的用户通信，其站点通常是属于同一个组织。



- PE需要为每个站点创建VPN实例，并配置全网唯一的RD。
- PE通过配置Import RT和Export RT来控制不同VPN的站点做到无法互访

MPLS VPN基本组网 - Extranet

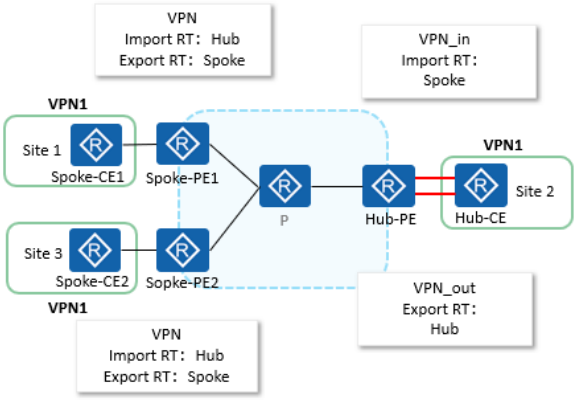
当采用Extranet组网方案时，VPN用户可将部分站点中的网络资源给其他VPN用户进行访问。



- 如图，Site 2作为能被VPN1和VPN2访问的共享站点，需要保证：
  - PE2能够接收PE1和PE3发布的VPNv4路由；
  - PE2发布的VPNv4路由能够被PE1和PE3接收；
  - PE2不把从PE1接收的VPNv4路由发布给PE3，也不把从PE3接收的VPNv4路由发布给PE1。

MPLS VPN基本组网 - Hub&Spoke (1)

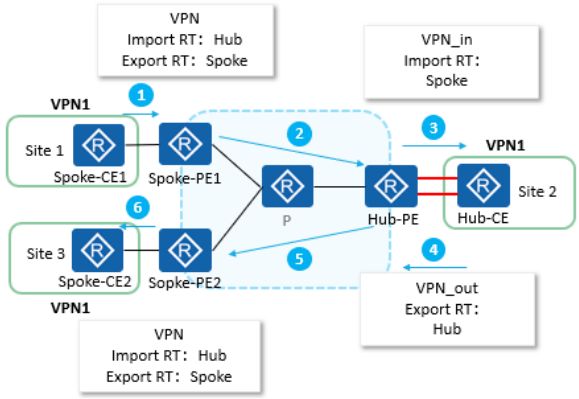
当采用Hub&Spoke方案时，可以将多个站点中的一个站点设置为Hub站点，其余站点为Spoke站点。站点间的互访必须通过Hub站点，通过Hub站点集中管控站点间的数据传输。



- Spoke站点需要把路由发布给Hub站点，再通过Hub站点发布给其他Spoke站点。Spoke站点之间不直接交互路由信息。
- Spoke-PE需要设置Export Target为“Spoke”，Import Target为“Hub”；
- Hub-PE上需要使用两个接口或子接口（创建两个VPN实例），一个用于接收Spoke-PE发来的路由，其VPN实例的Import Target为“Spoke”；另一个用于向Spoke-PE发布路由，其VPN实例的Export Target为“Hub”。

MPLS VPN基本组网 - Hub&Spoke (2)

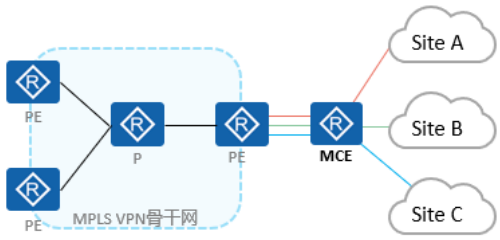
从Site1到Site2的路由发布过程如下：



1. Spoke-CE1发布路由给Spoke-PE1。
2. Spoke-PE1通过IBGP将该路由发布给Hub-PE。
3. Hub-PE通过VPN实例（VPN_in）的Import Target属性将该路由引入VPN_in路由表，并发布给Hub-CE。
4. Hub-CE学习到该路由，并将该路由发布给Hub-PE的VPN实例（VPN_out）。
5. Hub-PE通过VPN_out发布该路由给Spoke-PE2（携带VPN_out的Export Target属性）。
6. Spoke-PE2该路由发布给Spoke-CE2。

MCE组网

- 当一个私网需要根据业务或者网络划分VPN时，不同VPN用户间的业务需要完全隔离。此时，为每个VPN单独配置一台CE将增加用户的设备开支和维护成本。
- 具有MCE（Multi-VPN-Instance，CE多实例CE）功能的CE设备可以在MPLS VPN组网应用中承担多个VPN实例的CE功能，减少用户网络设备的投入。

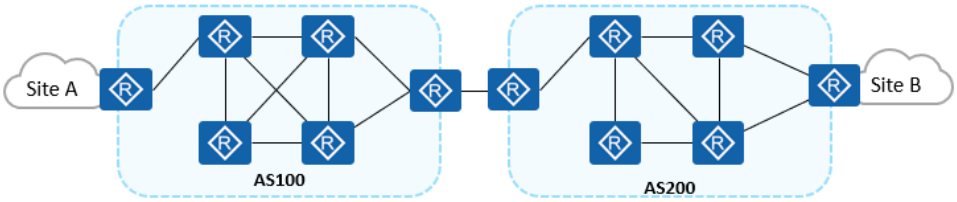


- MCE将PE的部分功能扩展到CE设备，通过将不同的接口与VPN绑定，并为每个VPN创建和维护独立的路由转发表（Multi-VRF）。
- MCE与对应的PE之间可以通过物理接口、子接口或者逻辑接口进行互联，PE上需要将这些接口绑定到对应的VPN实例。

...

MPLS VPN跨域组网

- 随着MPLS VPN解决方案的广泛应用，服务的终端用户的规格和范围也在增长，在一个企业内部的站点数目越来越大，某个地理位置与另外一个服务提供商相连的需求变得非常的普遍，例如国内运营商的不同城域网之间，或相互协作的运营商的骨干网之间都存在着跨越不同自治系统（AS，Autonomous System）的情况。
- 一般的MPLS VPN体系结构都是在一个AS内运行，任何VPN的路由信息都是只能在一个AS内按需扩散。AS之间的MPLS VPN部署需要通过跨域（Inter-AS） MPLS VPN解决方案来实现。



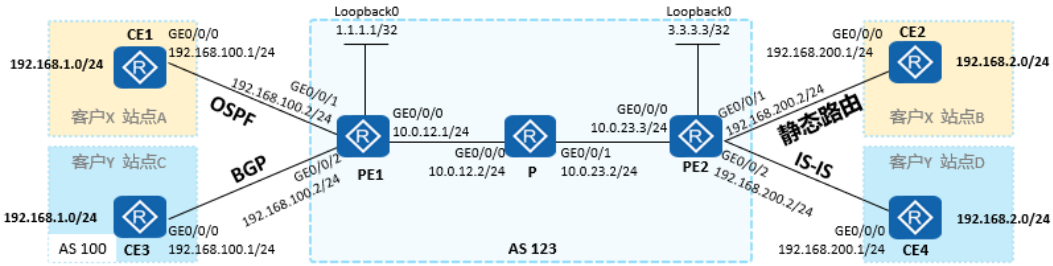
- RFC2547 中提出了三种跨域 VPN 解决方案，分别是：
- 跨域 VPN-OptionA（Inter-Provider Backbones Option A）方式：需要跨域的 VPN 在 ASBR（AS Boundary Router）间通过专用的接口管理自己的 VPN 路由，也称为 VRF-to-VRF；
- 跨域 VPN-OptionB（Inter-Provider Backbones Option B）方式：ASBR 间通过 MP-EBGP 发布标签 VPN-IPv4 路由，也称为 EBGP redistribution of labeled VPN-IPv4 routes；
- 跨域 VPN-OptionC（Inter-Provider Backbones Option C）方式：PE 间通过 Multi-hop MP-EBGP 发布标签 VPN-IPv

4 路由，也称为 Multihop EBGP redistribution of labeled VP N-IPv4 routes。

- 更多跨域 MPLS VPN 相关内容，请参考 HCIE-Datacom 相关课程。

部署Intranet场景的MPLS VPN

- 如图所示，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，CE与PE通过如图的协议或方法交换路由信息；



- 注：本课程仅讨论非跨域的 MPLS VPN 部署场景。

部署思路

1. 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。

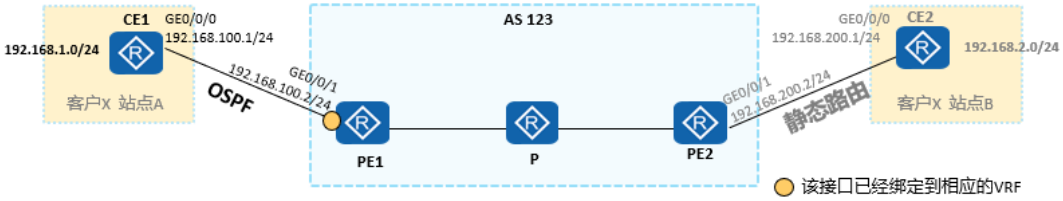
2. VPN用户接入配置

   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

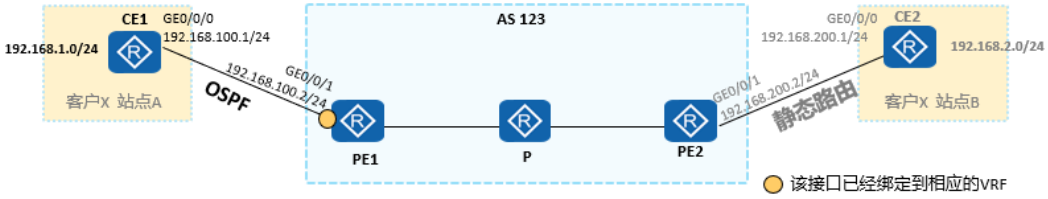
PE-CE之间部署OSPF (1)




```
[CE1] ospf 1
[CE1-ospf-1] area 0
[CE1-ospf-1-area-0.0.0.0] network 192.168.100.0 0.0.0.255
[CE1-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
```

CE1的OSPF配置还是传统的OSPF配置，CE1无需支持VRF。

PE-CE之间部署OSPF (2)



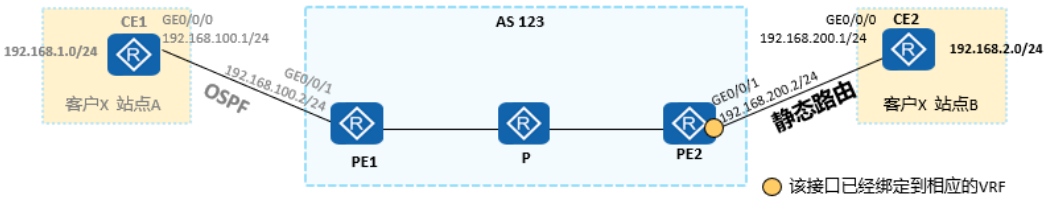

```
[PE1] ospf 1 vpn-instance VPNX
[PE1-ospf-1] area 0
[PE1-ospf-1-area-0.0.0.0] network 192.168.100.0 0.0.0.255
[PE1-ospf-1-area-0.0.0.0] quit
[PE1-ospf-1] import bgp
```

PE1用于跟CE1对接的OSPF进程必须与对应的VPN实例绑定。
将PE1的VPN实例VPNX的路由表中的BGP路由（主要是PE1通过BGP获知的、到达站点B的客户路由）引入OSPF，以便将这些路由通过OSPF通告给CE1。

```
[PE1] bgp 123
[PE1-bgp] ipv4-family vpn-instance VPNX
[PE1-bgp] import-route ospf 1
```

将PE1的VPN实例VPNX的路由表中通过OSPF进程1学习到的OSPF路由引入BGP，从而将到达站点A的客户路由转换成BGP的VPNv4路由，以便通告给远端的PE2。

PE-CE之间部署静态路由



[CE2] **ip route-static** 192.168.1.0 24 192.168.200.2
[CE2] **ip route-static** 192.168.100.0 24 192.168.200.2

CE2需配置到达站点A内的各个网段的静态路由。

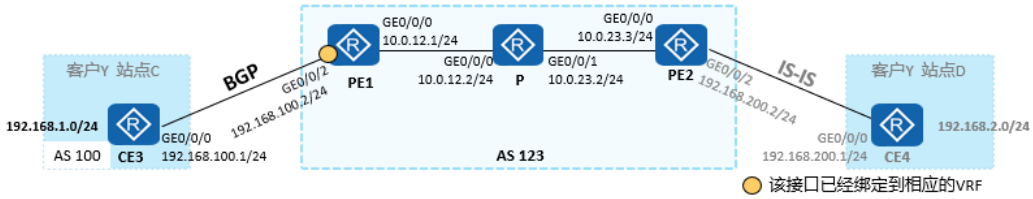
[PE2] **ip route-static vpn-instance** VPNX 192.168.2.0 24 192.168.200.1

PE2需配置到达站点B内各个网段的静态路由。

[PE2] **bgp** 123
[PE2-bgp] **ipv4-family vpn-instance** VPNX
[PE2-bgp] **import-route static**

将PE2的VPN实例VPNX的路由表中的静态路由引入BGP，从而将客户路由转换成BGP的VPNv4路由，以便通告给远端的PE1。

PE-CE之间部署EBGP



[CE3] **bgp** 100
[CE3-bgp] **peer** 192.168.100.2 **as-number** 123
[CE3-bgp] **network** 192.168.1.0 24

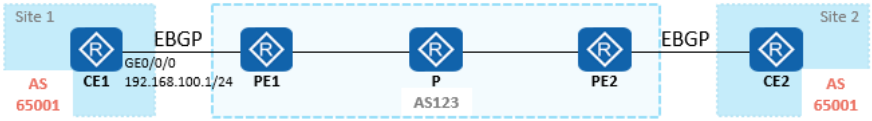
CE3只需要执行普通BGP配置，且无需支持VRF。

[PE1] **bgp** 123
[PE1-bgp] **ipv4-family vpn-instance** VPNY
[PE1-bgp-VPNY] **peer** 192.168.100.1 **as-number** 100

当PE与CE之间使用BGP交互客户路由时，无需在PE上手工执行路由重分发操作。在本例中，PE1通过BGP从CE3学习到的客户路由后，PE1会自动将这些路由转换成VPNv4路由并通告给PE2；而PE1通过BGP从PE2获知到达站点D的路由后，会自动将它们转换成IPv4路由并通告给CE3。

特殊场景下的BGP配置 - AS号替换

在MPLS VPN场景中，若PE与CE之间运行EBGP交互路由信息，则可能会出现两个站点的AS号相同的情况。



- 若CE1通过EBGP向PE1发送一条私网路由，并经过PE2发送到CE2，则CE2会由于AS号重复丢弃这条路由，导致属于同一VPN的Site 1和Site 2之间无法连通。
- 可以在PE上执行**peer substitute-as**命令使能AS号替换功能，即PE用本地AS号替换收到的私网路由中CE所在VPN站点的AS号，这样对端CE就不会因为AS号重复而丢弃路由了。

```
[PE1] bgp 123
[PE1-bgp] ipv4-family vpn-instance vpn1
[PE1-bgp-vpn1] peer 192.168.100.1 substitute-as
```

PE1在向CE1发送BGP路由时，若发现AS_Path中包含65001，则会用本地AS号，也就是123去替换65001。所以，若有一条路由从CE2传给PE2，再由PE2传给PE1，当PE1再传递给CE1，此时BGP路由的AS_Path属性为{123,123}。

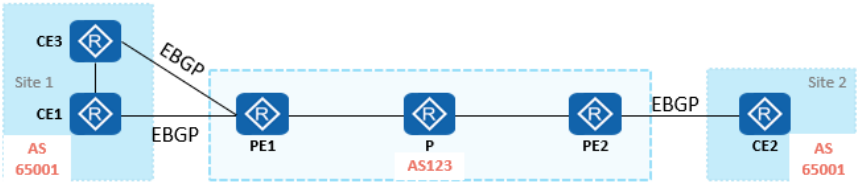
特殊场景下的BGP配置 - SoO

- 在CE多归属场景，若使能了BGP的AS号替换功能，可能会引起路由环路，需要SoO（Site of Origin）特性来避免环路。
  - CE1与CE3处于同一个VPN站点1，CE2位于站点Site2，Site1和Site2站点所在的AS号都为65001。PE与CE之间运行的都是EBGP路由协议，为了Site 1和Site 2之间的路由可以正常学习，需要在PE1和PE2上配置AS号替换功能。
  - CE1传递站点内的路由给PE1，PE1传递该路由给CE3，由于配置AS号替换，CE3会接收该路由，可能会导致产生路由环路。

```
[PE1] bgp 123
[PE1-bgp] ipv4-family vpn-instance vpn1
[PE1-bgp-vpn1] peer 192.168.100.1 soo 200:1
[PE1-bgp-vpn1] peer 192.168.200.1 soo 200:1
```

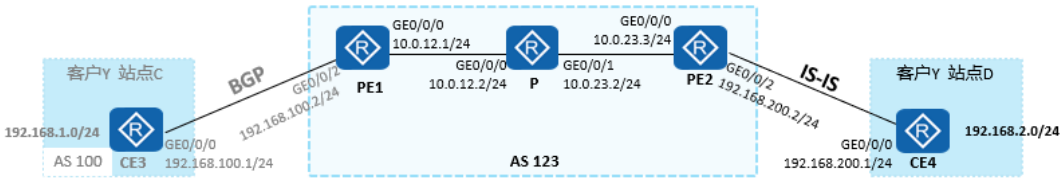
配置了BGP邻居的SoO后：

- 接收到该邻居的BGP路由时，会在路径属性中携带该SoO属性并通告给其他BGP邻居。
- 向该邻居通告BGP路由时，会检查路由中的SoO属性是否与配置的SoO值相同，若相同则不通告，避免引起环路。



- 注：192.168.100.1 和 192.168.200.1 分别是 CE1 和 CE3 上和 PE1 建立 BGP 对等体的接口地址。

PE-CE之间部署IS-IS

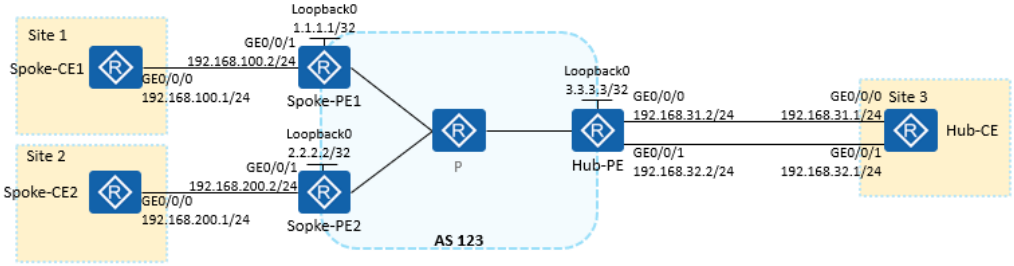



```
[CE4] isis 1
[CE4-isis-1] network-entity 49.0001.0000.0000.1111.00
[CE4-isis-1] is-level level-2
[CE4-isis-1] quit
[CE4] interface GigabitEthernet 0/0/0
[CE4-GigabitEthernet0/0/0] isis enable 1
[CE4-GigabitEthernet0/0/0]quit
[CE4] interface GigabitEthernet 0/0/1
[CE4-GigabitEthernet0/0/1] isis enable 1
#GE0/0/1接口是192.168.2.0/24网段所在接口
```

```
[PE2] isis 1 vpn-instance VPNY
[PE2-isis-1] network-entity 49.0002.0000.0000.2222.00
[PE2-isis-1] is-level level-2
[PE2-isis-1] import-route bgp level-2
[PE2-isis-1] quit
[PE2] interface GigabitEthernet 0/0/2
[PE2-GigabitEthernet0/0/2] isis enable 1
[PE2] bgp 123
[PE2-bgp] ipv4-family vpn-instance VPNY
[PE2-bgp] import-route isis 1
```

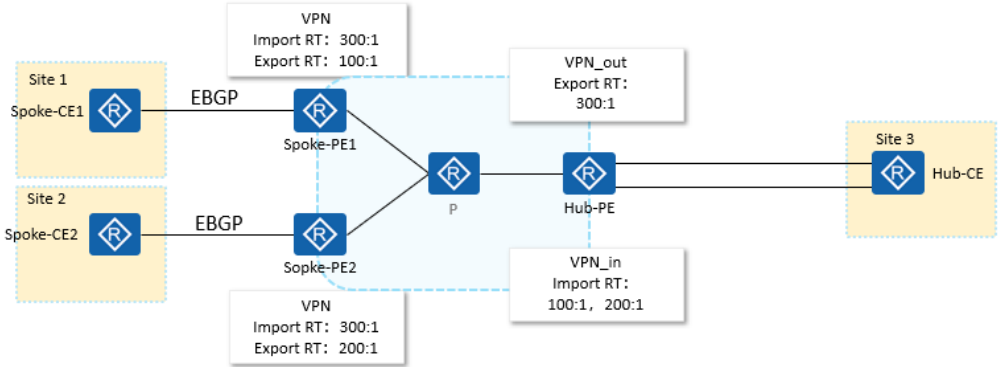
部署Hub&Spoke场景的MPLS VPN

- Hub&Spoke有以下组网方案：
  - 方式一：Hub-CE与Hub-PE，Spoke-PE与Spoke-CE使用EBGP
  - 方式二：Hub-CE与Hub-PE，Spoke-PE与Spoke-CE使用IGP
  - 方式三：Hub-CE与Hub-PE使用EBGP，Spoke-PE与Spoke-CE使用IGP
- **无法**通过Hub-CE与Hub-PE使用IGP，Spoke-PE与Spoke-CE使用EBGP来部署Hub&Spoke组网的MPLS VPN。



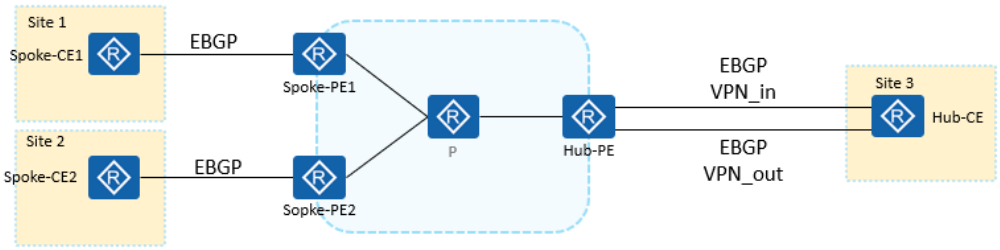
VRF配置

- Spoke-PE上创建一个VPN实例，RT配置如图。
- Hub-PE上创建VPN_in和VPN_out两个VPN实例，分别用于从Spoke-PE接收私网路由或向Spoke-PE发布私网路由，RT配置如图。



方式一部署 - 路由发布过程

- Spoke-CE与Spoke-PE之间通过EBGP交互路由信息，建立EBGP连接后，把相关的路由发布到BGP即可。
- Hub-PE与Hub-CE之间建立两条EBGP连接，分别用来发布和接收私网路由。

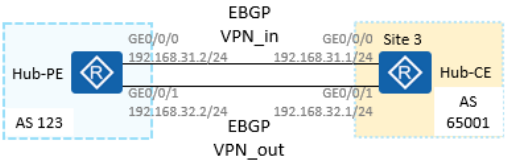


- 以路由从 Spoke-CE1 发布到 Spoke-CE2 为例，大体过程如下：
- Spoke-CE1 通过 EBGP 将路由发布给 Spoke-PE1。
- Spoke-PE1 通过 IBGP 将该路由发布给 Hub-PE。
- Hub-PE 通过 VPN 实例（VPN_in）的 Import Target 属性将该路由引入 VPN_in 路由表，并通过 EBGP 发布给 Hub-CE。
- Hub-CE 通过 EBGP 连接学习到该路由，并通过另一个 EBGP 连接将该路由发布给 Hub-PE 的 VPN 实例（VPN_out）。

- Hub-PE 发布携带 VPN_out 的 Export Target 属性的路由给所有 Spoke-PE。
- Spoke-PE2 通过 EBGP 将该路由发布给 Spoke-CE2。

方式一部署 - Hub-PE与Hub-CE间配置

- Hub-PE通过VPN_in对应的EBGP连接将从Spoke站点学习的路由发布到Hub站点。
- Hub-CE通过VPN_out对应的EBGP将这些路由发布到Spoke站点。



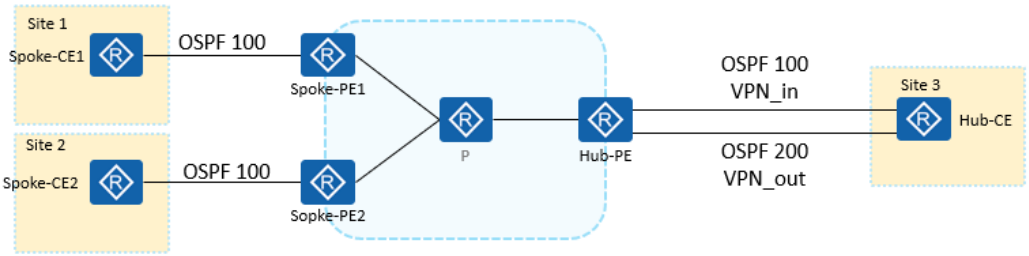

```
# Hub-PE与Hub-CE建立两条EBGP连接
[Hub-PE] bgp 123
[Hub-PE-bgp] ipv4-family vpn-instance VPN_in
[Hub-PE-bgp-VPN_in] peer 192.168.31.1 as-number 65001
[Hub-PE-bgp-VPN_in]quit
[Hub-PE-bgp] ipv4-family vpn-instance VPN_out
[Hub-PE-bgp-VPN_out] peer 192.168.32.1 as-number 65001
[Hub-PE-bgp-VPN_out] peer 192.168.32.1 allow-as-loop
```

```
# Hub-CE与Hub-PE建立两条EBGP连接
[Hub-CE] bgp 65001
[Hub-CE-bgp] peer 192.168.31.2 as-number 123
[Hub-CE-bgp] peer 192.168.32.2 as-number 123
```

由于Hub-CE通过VPN_out对应的EBGP连接发送给Hub-PE的路由可能带有AS 123，则这些路由将会被Hub-PE丢弃，故Hub-PE上必须手工配置允许本地AS编号重复。

方式二部署 - 路由发布过程

- 以选用OSPF作为IGP协议为例：
  - Spoke-CE与Spoke-PE之间通过OSPF（进程100）邻居关系交互路由信息。
  - Hub-PE通过两个OSPF进程与Hub-CE建立OSPF邻居，分别负责私网路由的发送和接收。



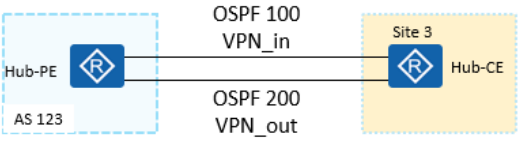
- 以路由从 Spoke-CE1 发布到 Spoke-CE2 为例，大体过程如下：
- Spoke-CE1 通过 OSPF100 将路由发布给 Spoke-PE1。
- Spoke-PE1 通过 IBGP 将路由发布给 Hub-PE。
- Hub-PE 通过 VPN 实例（VPN_in）的 Import Target 属性将该路由引入 VPN_in 路由表；通过将 BGP 引入 OSPF100

的配置进而将从 Spoke-PE1 传递来的路由发布给 Hub-CE。

- Hub-CE 通过 OSPF100 接收该路由；并通过配置路由引入将路由发布到 OSPF200，OSPF200 再将路由发布给 Hub-PE。
- Hub-PE 的 BGP-VPN 实例（VPN_out）引入 OSPF200 多实例的路由，将携带 Export Target 属性的路由发布给所有 Spoke-PE。
- Spoke-PE2 通过 OSPF100 将该路由发布给 Spoke-CE2。

方式二部署 - Hub-PE与Hub-CE间配置

- Hub-PE通过VPN_in对应的OSPF（进程100）将从Spoke站点学习的路由发布到Hub站点。
- Hub-CE通过VPN_out对应的OSPF（进程200）将这些路由发布到Hub-PE，进而发布给所有Spoke站点。

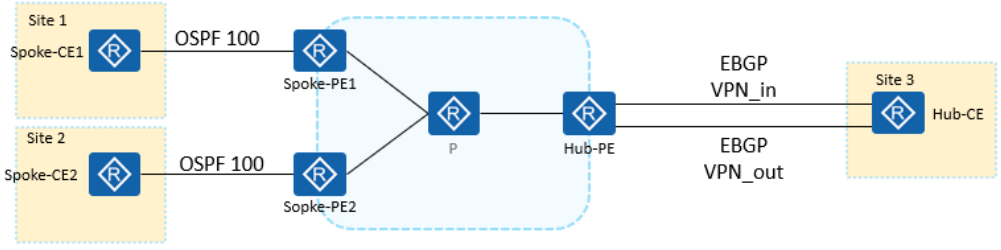



```
# Hub-PE上执行OSPF和BGP的相互引入
[Hub-PE] OSPF 100 vpn-instance VPN_in
[Hub-PE-ospf-100]import-route bgp
[Hub-PE-ospf-100]quit
[Hub-PE]bgp 100
[Hub-PE-bgp]ipv4-family vpn-instance VPN_out
[Hub-PE-bgp-VPN_out]import-route ospf 200
```

```
# Hub-CE上OSPF 200到OSPF 100的路由引入
[Hub-CE]OSPF 200
[Hub-CE-ospf-200]import-route OSPF 100
```

方式三部署 - 路由发布过程

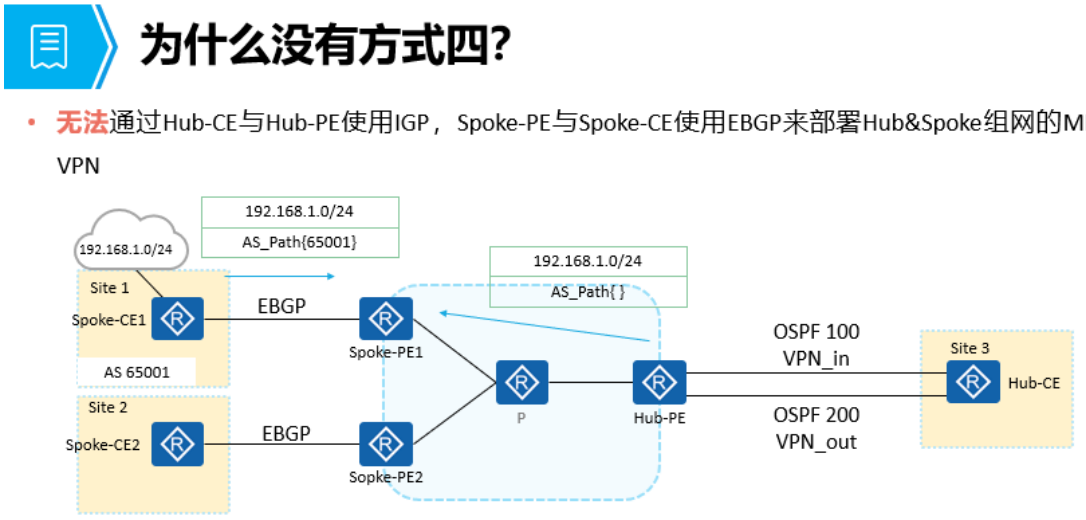
- 以选用OSPF作为IGP协议为例，Spoke-CE与Spoke-PE之间通过OSPF（进程100）邻居关系交互路由信息。
- Hub-PE与Hub-CE之间建立两条EBGP连接，分别用来发布和接收私网路由，Hub-PE与Hub-CE的配置与方式一相同。



- 以路由从 Spoke-CE1 发布到 Spoke-CE2 为例，大体过

程如下：

- Spoke-CE1 通过 OSPF100 将路由发布给 Spoke-PE1。
- Spoke-PE1 通过 IBGP 将路由发布给 Hub-PE。
- Hub-PE 通过 VPN 实例（VPN_in）的 Import Target 属性将该路由引入 VPN_in 路由表，并通过 EBGP 发布给 Hub-CE。
- Hub-CE 通过 EBGP 连接学习到该路由，并通过另一个 EBGP 连接将该路由发布给 Hub-PE 的 VPN 实例（VPN_out）。
- Hub-PE 发布携带 VPN_out 的 Export Target 属性的路由给 Spoke-PE2。
- Spoke-PE2 通过 OSPF100 将该路由发布给 Spoke-CE2。



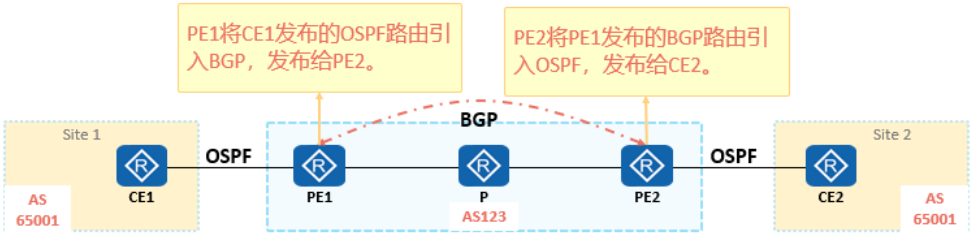
- 以从 Spoke-CE1 向 Spoke-CE2 发布路由（目的地址为 192.168.1.0/24）为例，大体过程如下：
- Spoke-CE1 通过 EBGP 将路由发布给 Spoke-PE1。
- Spoke-PE1 通过 IBGP 将路由发布给 Hub-PE。
- Hub-PE 通过 BGP-VPN 实例（VPN_in）的 Import Target 属性将该路由引入 VPN_in 路由表；并通过 OSPF100 多实

例发布给 Hub-CE。

- Hub-CE 通过 OSPF100 学习到该路由；并通过 OSPF200 将路由发布给 Hub-PE。
- Hub-PE 的 BGP-VPN 实例（VPN_out）引入 OSPF200 多实例路由，并将携带 VPN_out 的 Export Target 属性的路由发布给所有 Spoke-PE。
- Spoke-PE2 的 VPN 实例根据 Import Target 属性引入该路由；Spoke-PE2 通过 EBGP 发布给本地 Spoke-CE2。
- Hub-PE 的 BGP-VPN 实例（VPN_out）通过 Export Target 属性将路由发布给 Spoke-PE2 的同时，也会将该路由发布给 Spoke-PE1。此时，这条路由是 Hub-PE 通过 IGP（OSPF200 多实例）引入的，由于 IGP 路由不携带 AS-PATH 属性，AS_Path 为空；而原来从 Spoke-CE1 来的 192.168.1.0/24 路由，其 AS_Path 不为空，所以从 Hub-PE 返回的路由会优于从 Spoke-CE1 来的路由。这样会引起路由振荡，其过程如下：
- Spoke-CE1 发来的路由因为 AS_Path 变成非最佳路由
- Spoke-PE1 发布 Update 撤销路由的报文给 Hub-PE 来撤销 192.168.1.0/24 路由
- Hub-PE（通过撤销相应的 OSPF LSA）撤销发给 Hub-CE 的路由
- Hub-CE（原理同上）撤销发给 Hub-PE 的路由
- Hub-PE 发布 Update 撤销路由撤销发给 Spoke-PE1 的路由
- 于是在 Spoke-PE1 上从 Spoke-CE1 来的路由又变成最佳路由。Spoke-PE1 又通过 IBGP 将路由发布给 Hub-PE。Hub-PE 又会返回该路由，从 Spoke-CE1 来的路由又变成非最佳路由。如此反复。

MPLS VPN中的OSPF/BGP

- 当PE-CE间部署OSPF交互路由信息时，若在PE上使用标准BGP/OSPF过程（简称为BGP/OSPF互操作）互来传递路由信息，则远端PE在将BGP引入VPN实例的OSPF进程时，会直接产生Type5 LSA，不同站点都会将其他站点的路由视为自治系统外部路由（AS_external）。
- 为了解决标准BGP/OSPF的互操作导致的OSPF路由信息丢失的问题，BGP和OSPF都做了相应的拓展。



Site 1向Site 2发布路由过程图

- 在实际应用中，如果两个要互通的 Site 都在相同的 AS 内，那么每个 Site 都应该将另一个 Site 的路由看成区域间路由，而不是 AS 外部路由。

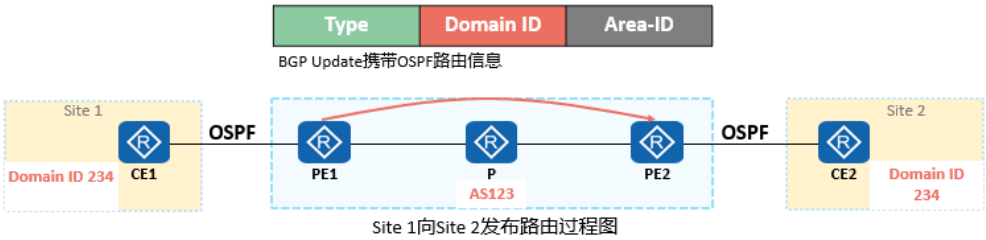
BGP 扩展团体属性

- 为了保留 OSPF 的路由信息，BGP 新增了部分可携带 OSPF 路由信息的团体属性：
- Domain ID：域标识符用来标识和区分不同的域。
- Route Type：包含被引入到 BGP 的 OSPF 路由的 Area-ID 以及 Route Type
- Area-ID：PE 的 VPN 实例的 OSPF 进程与 CE 建立邻接关系的区域号
- Route Type：被引入的 OSPF 路由的类型
- 1 或 2：表示路由的类型为区域内部路由，也就是 PE 根据 Type-1 及 Type-2 LSA 所计算出来的路由。
- 3：表示路由的类型为区域间路由。
- 5：表示路由的类型为 OSPF 外部路由，也就是 PE 通过 Type-5 LSA 计算得出的路由。当 Route-Type 字段的值为 5 时，Area-ID 字段的值需为 0.0.0.0。

- 7：表示路由的类型为 NSSA 路由，也就是 PE 通过 Type-7 LSA 计算得出的路由。

Domain ID

- 在PE上将OSPF引入BGP时，PE将根据本地的配置为BGP路由增加域ID属性，域ID作为BGP的扩展团体属性传播。
- 在PE将BGP路由引入OSPF时，若BGP路由携带的Domain ID与本地相同，则认为两个站点属于同一个OSPF路由域。若不相同，则认为不在同一个路由域。



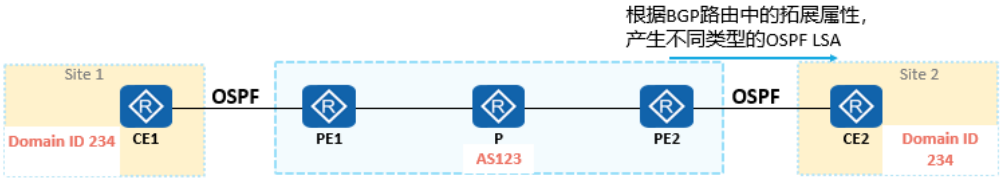
Site 1向Site 2发布路由过程图

- Domain ID 需要在绑定到 VRF 的 OSPF 进程视图下使用命令 **domain-id** 配置。
- 缺省情况下，Domain ID 的值为 0（NULL）。如果不同 OSPF 域都使用 NULL 作为 Domain ID，将无法区分 OSPF 域，因此它们之间的路由将被当作区域内路由。
- 如果一个 OSPF 域配置了非 0（即非 NULL）的 Domain ID，NULL 不再是该 OSPF 域的 Domain ID。
- 建议与同一个 VPN 相关的所有 OSPF 实例都使用相同的 Domain ID，或者都使用缺省的 Domain ID。

Domain ID与Route Type

根据BGP路由中的Domain ID与Route Type属性，PE将产生不同类型的OSPF LSA类型发布到VRF的OSPF进程中

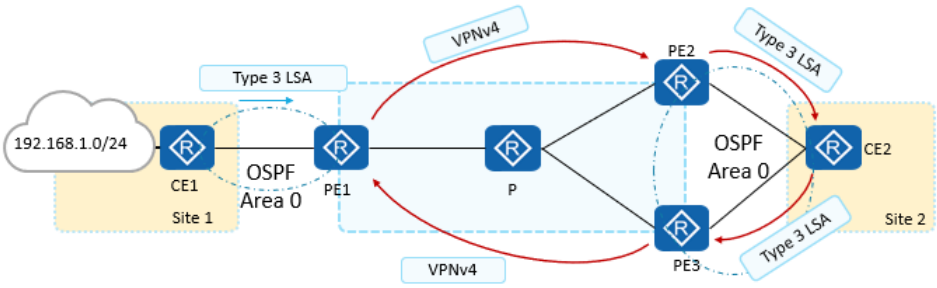
Domain ID与本地是否相同	Route Type	PE生成的OSPF LSA类型
是	1、2、3	3
	5、7	5、7
否	1、2、3、5、7	5、7



Site 1向Site 2发布路由过程图

Type3路由防环 - 案例

- 如图是Type3 LSA路由产生环路的一个例子：
  - 其中站点1和站点2都属于VPN1。
  - 站点1通过OSPF Area0接入骨干网的PE1；
  - 站点2通过OSPF Area0分别接入骨干网的PE2和PE3（双归属负载分担场景）。

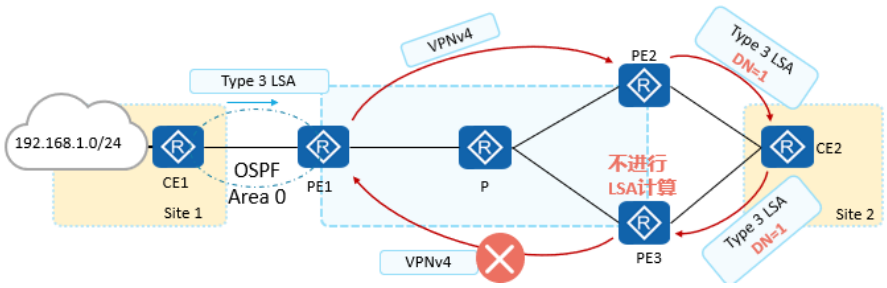


- 环路产生过程如下：
- Site1 的 CE1 通过 Type3 LSA 发布到 192.168.1.0/24 的路由给 PE1。
- PE1 向 BGP 引入 OSPF VPN1 进程，通过 MP-IBGP 将该路由发布给 PE2 和 PE3。
- 由于 PE2 上配置了 BGP 到 OSPF 的路由引入，故 PE2 将产生 Type3 LSA 给 CE2，CE2 将来自 PE2 的 Type3 LSA 发布给 PE3。

- 此时 PE3 收到两条到达 192.168.1.0/24 的路由：一条是 PE1 发布的，另一条是 PE2 上路由引入产生的。由于缺省情况下，IGP（OSPF）路由优先级高于 IBGP 路由，PE3 将选择 OSPF 路由。
- PE3 将优选的学自 OSPF 通过 MP-IBGP 发布给 PE1。
- 此时，PE1 上存在两条到达目的地 192.168.1.0/24 网段的路由，一条通过 OSPF 学习自 CE1，另一条通过 MP-IBGP 学习自 PE3。可能会导致以下问题：
- PE1 撤销 192.168.1.0/24 这条路由，但由于 PE1 与 PE2 之间的链路阻塞，BGP Update（撤销报文）无法及时发送给 PE2，导致 PE3 发给 PE1 的路由依然存在（正常情况下会随着 PE1 发送给 PE2 的 Update 报文被撤销），PE1 上到达目的地 192.168.1.0/24 的下一跳为 PE3。此时路由环路产生。
- 若 PE1 上 MP-IBGP 的路由优先级高于 OSPF，则 PE1 会优选 PE3 通告的 BGP 路由，此时 PE1 需要撤销发给 PE2 的 BGP 路由，导致 PE3 撤销发布给 PE1 的路由，PE1 上的 OSPF 再次被优选。如此反复，形成路由震荡。

Type3路由防环 - DN位

- 为了防止3类LSA环路，OSPF多实例进程使用LSA Options域中一个原先未使用的比特作为标志位，称为DN位。使用DN位可以防止Type3 LSA环路。
- PE路由器的OSPF实例进程在进行SPF计算时，忽略DN置位的Type3 LSA。

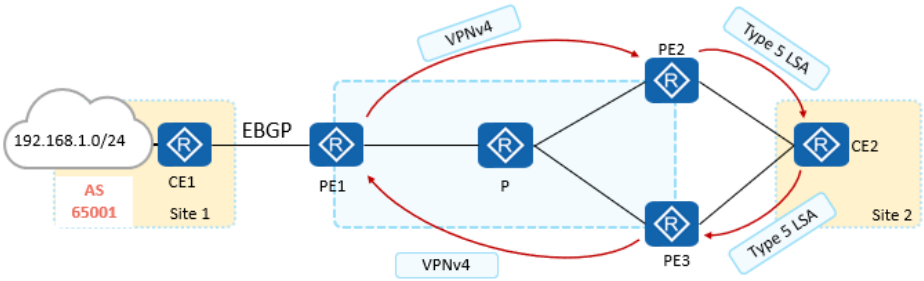


- 缺省情况下，OSPF 产生的 LSA 的 DN 位设置为 1。可

通过命令 **dn-bit-set disable** 来禁止设置 OSPF LSA 的 DN 位。

Type5/7路由防环 - 案例

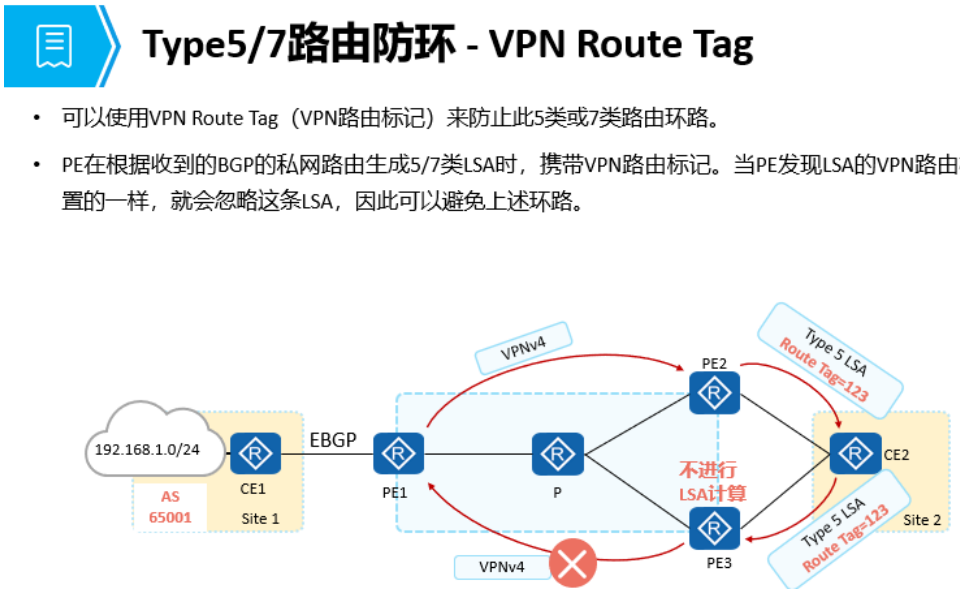
- 如图是Type5 LSA路由产生环路的一个例子：
  - 其中站点1和站点2都属于VPN1。
  - 站点1通过EBGP接入骨干网的PE1；
  - 站点2通过OSPF分别接入骨干网的PE2和PE3。



- 环路产生过程如下：
- CE1 通过 EBGP 发布到 192.168.1.0/24 的路由给 PE1，AS_Path 为 65001。
- PE1 通过 MP-IBGP 将该路由发布给 PE2 和 PE3。
- PE2 在 OSPF VPN1 实例进程中引入 BGP 路由，发布到 192.168.1.0/24 的 Type5 LSA 给 CE2；
- CE2 将该 Type5 LSA 发布给 PE3。
- PE3 将优选择 OSPF 路由（OSPF 优先级高于 IBGP），并通过 MP-IBGP 发布 Update 消息给 PE1。
- PE1 收到 PE3 发送的 MP-IBGP Update 消息。由于其中的路由是 PE3 的 BGP 引入的 IGP（OSPF）路由，其 AS_Path 为空，因此 PE3 发布的 MP-IBGP 路由优先级比从 CE1 发布的 EBGP 路由优先级高，PE1 将优选用 PE3 发布的路由。
- 此时形成一条路由环路：PE3—>CE2—>PE2—>PE1—>PE3。
- 由于 PE1 不再优选 CE1 学来的路由，故 PE1 会撤销发给 PE2 的路由，PE2 中的 OSPF VPN 实例进程进程中也要对应撤销引入的 BGP 路由，继而 CE2、PE3 都相继撤销该 OS

PF 路由。PE3 向 PE1 发布的 BGP 路由也被撤销，在 PE1 上，从 CE1 学来的路由又变成最优路由。这样，形成了路由振荡。

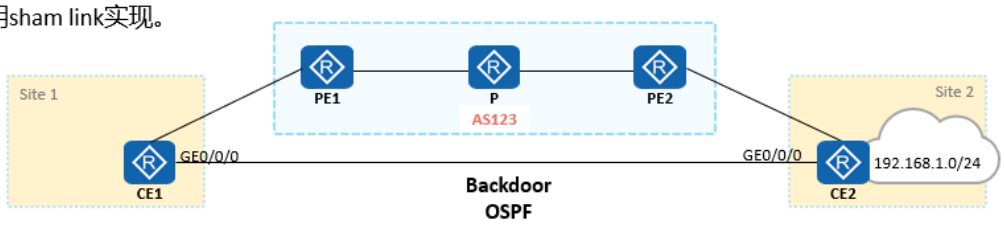
- Type7 LSA 环路的产生和消除的过程与 Type5 类似，此处不再赘述。



- VPN 路由标记不在 BGP 的扩展团体属性中传递，只是本地概念，只在收到 BGP 路由并且产生 OSPF LSA 的 PE 设备上有意义。
- 缺省情况下，VPN 的路由标记是根据 BGP 的 AS 号计算得到的。如果没有配置 BGP，则默认值为 0。
- 可以通过指令 **route-tag** 配置 VPN 的路由标记。

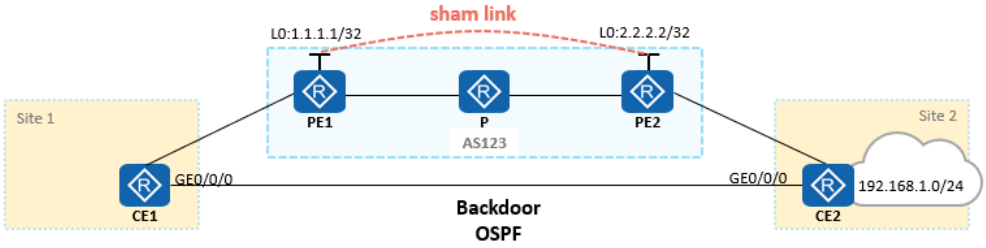
Sham link的应用场景

- 通常情况下，BGP对等体之间通过BGP扩展团体属性在MPLS VPN骨干网上承载路由信息。另一端PE上运行的OSPF可利用这些信息来生成PE到CE的Type 3 LSA，这些路由是区域间路由（Inter_Area route）。
- 若在CE1和CE2之间增加一条后门（Backdoor）链路，并且直接运行OSPF交互路由。通过后门链路学习到的路由类型为区域内路由（Intra_Area route）。
- 由于区域内路由优于区域间路由，故后门链路会被优选，若想实现后门链路作为备份链路，可采用sham link实现。



Sham link的工作机制

- Sham link在两台PE之间创建了一条区域内链路。当LSA在伪装链路中泛洪，所有的OSPF路由类型都不会改变，不会转换成LSA3或者LSA5的类型。
- Sham link被看成是两个VPN实例之间的链路，链路的两端是PE上的端点地址，分别作为建立连接时的源和目的地址。伪连接的源地址和目的地址使用32位掩码的Loopback接口地址，该Loopback接口需要绑定到VPN实例中，并通过BGP发布。



Sham link的配置示例

1. 在PE上创建用于建立sham Link的接口（PE2配置类似）。

```
[PE1]interface LoopBack0
[PE1-LoopBack0] ip binding vpn-instance VPNA
[PE1-LoopBack0]ip address 1.1.1.1 32
#在BGP的VPN地址族中发布出去
[PE1-bgp-VPNA]network  1.1.1.1 32
```

2. 在PE节点上配置sham Link（PE2节点配置类似）。

```
[PE1-ospf-1]area 0
[PE1-ospf-1-area-0.0.0.0]sham-link 1.1.1.1 2.2.2.2
```

3. 调整cost值，确保后门链路的cost要大于sham Link上的cost。

```
[CE1-GigabitEthernet0/0/0]ospf cost  1000
```

- 在配置 Sham link 时可以指定 Sham link 的路由开销。缺省值为 1。

sham link的配置验证

1. OSPF VPN常规配置请参照前述的配置示例 （略）。

```
<PE1>display ospf sham-link  area  0

      OSPF Process 1 with Router ID 1.1.1.1

 Sham-Link: 1.1.1.1 --> 2.2.2.2
 Neighbor ID: 2.2.2.2 ,  State: Full ,  GR status: Normal
 Area: 0.0.0.0
 Cost: 1 , State: P-2-P , Type: Sham
 Timers: Hello 10 , Dead 40 , Retransmit 5 , Transmit Delay 1
```

2. 在CE1节点上查看OSPF路由，可以看到对端路由是作为区域内路由学习到的。

```
<CE1>display  ospf routing

 Routing for Network
 Destination        Cost      Type       NextHop          AdvRouter        Area
 1.1.1.1/32         0         Stub       1.1.1.1          10.1.12.1        0.0.0.0
 10.1.12.0/24       1         Transit    10.1.12.1        10.1.12.1        0.0.0.0
 192.168.1.0/24     3         Stub       10.0.12.1        2.2.2.2          0.0.0.0
 10.1.23.0/24       3         Transit    10.1.12.2        10.1.23.2        0.0.0.0
```

思考题：

- （多选题）在 MPLS VPN 组网中，当 PE 向 OSPF 引入从其他 PE 学习来的 VPN 路由时，可能会产生以下哪几类 LSA（　　）。
- Type 1 LSA
- Type 3 LSA

- Type 5 LSA
- Type 7 LSA
- （判断）CE 通过 BGP 传递路由给 PE 时，可能会携带 SoO 属性。（　　）。
- 正确
- 错误

参考答案：

- BCD
- B
-