# H3C MSR 系列路由器

# 三层技术-IP 路由配置指导(V7)

杭州华三通信技术有限公司
http://www.h3c.com.cn

资料版本：6W103-20140512
产品版本：MSR-CMW710-R0105

# 前 言

H3C MSR 系列路由器 配置指导(V7)共分为十五本手册，介绍了 MSR 系列路由器各软件特性的原理及其配置方法，包含原理简介、配置任务描述和配置举例。《三层技术-IP 路由配置指导》主要介绍路由协议的原理和配置，包括 IPv4、IPv6 网络的各种路由学习技术，以及影响路由选择或者路由表生成的策略。

前言部分包含如下内容：



## 适用款型

本手册所描述的内容适用于 MSR 系列路由器中的如下款型：

| 款型 | |
|---|---|
| MSR 2600 | MSR 26-30 |
| MSR 3600 | MSR 36-10 |
| | MSR 36-20 |
| | MSR 36-40 |
| | MSR 36-60 |
| | MSR3600-28 |
| | MSR3600-51 |
| MSR 5600 | MSR 56-60 |
| | MSR 56-80 |

## 读者对象

本手册主要适用于如下工程师：

- 网络规划人员
- 现场技术支持与维护人员
- 负责网络配置和维护的网络管理员

# 本书约定

## 1. 命令行格式约定

| 格　式 | 意　义 |
| --- | --- |
| **粗体** | 命令行关键字（命令中保持不变、必须照输的部分）采用**加粗**字体表示。 |
| *斜体* | 命令行参数（命令中必须由实际值进行替代的部分）采用*斜体*表示。 |
| [] | 表示用“[]”括起来的部分在命令配置时是可选的。 |
| { x \| y \| ... } | 表示从多个选项中仅选取一个。 |
| [ x \| y \| ... ] | 表示从多个选项中选取一个或者不选。 |
| { x \| y \| ... }* | 表示从多个选项中至少选取一个。 |
| [ x \| y \| ... ]* | 表示从多个选项中选取一个、多个或者不选。 |
| &<1-n> | 表示符号&前面的参数可以重复输入1～n次。 |
| # | 由“#”号开始的行表示为注释行。 |

## 2. 各类标志

本书还采用各种醒目标志来表示在操作过程中应该特别注意的地方，这些标志的意义如下：

| | |
| --- | --- |
| 警告 | 该标志后的注释需给予格外关注，不当的操作可能会对人身造成伤害。 |
| 注意 | 提醒操作中应注意的事项，不当的操作可能会导致数据丢失或者设备损坏。 |
| 提示 | 为确保设备配置成功或者正常工作而需要特别关注的操作或信息。 |
| 说明 | 对操作内容的描述进行必要的补充和说明。 |
| 窍门 | 配置、操作、或使用设备的技巧、小窍门。 |

## 3. 图标约定

本书使用的图标及其含义如下：

| | |
| --- | --- |
|  | 该图标及其相关描述文字代表一般网络设备，如路由器、交换机、防火墙等。 |
|  | 该图标及其相关描述文字代表一般意义下的路由器，以及其他运行了路由协议的设备。 |
|  | 该图标及其相关描述文字代表二、三层以太网交换机，以及运行了二层协议的设备。 |

#### 4. 端口编号示例约定

本手册中出现的端口编号仅作示例，并不代表设备上实际具有此编号的端口，实际使用中请以设备上存在的端口编号为准。

## 产品配套资料

H3C MSR 系列路由器的配套资料包括如下部分：

| 大类 | 资料名称 | 内容介绍 |
| --- | --- | --- |
| 产品知识介绍 | 路由器产品彩页 | 帮助您了解产品的主要规格参数及亮点 |
| 硬件描述与安装 | 路由器安装指导 | 帮助您详细了解设备硬件规格和安装方法，指导您对设备进行安装 |
| | MSR 系列路由器接口模块手册 | 帮助您详细了解单板的硬件规格 |
| 业务配置 | MSR 系列路由器配置指导(V7) | 帮助您掌握设备软件功能的配置方法及配置步骤 |
| | MSR 系列路由器命令参考(V7) | 详细介绍设备的命令，相当于命令字典，方便您查阅各个命令的功能 |
| 运行维护 | 路由器版本说明书 | 帮助您了解产品版本的相关信息（包括：版本配套说明、兼容性说明、特性变更说明、技术支持信息）及软件升级方法 |

## 资料获取方式

您可以通过H3C网站（www.h3c.com.cn）获取最新的产品资料：

H3C 网站与产品资料相关的主要栏目介绍如下：

- [服务支持/文档中心]：可以获取硬件安装类、软件升级类、配置类或维护类等产品资料。
- [产品技术]：可以获取产品介绍和技术介绍的文档，包括产品相关介绍、技术介绍、技术白皮书等。
- [解决方案]：可以获取解决方案类资料。
- [服务支持/软件下载]：可以获取与软件版本配套的资料。

## 技术支持

用户支持邮箱：service@h3c.com

技术支持热线电话：400-810-0504（手机、固话均可拨打）

网址：http://www.h3c.com.cn

# 资料意见反馈

如果您在使用过程中发现产品资料的任何问题，可以通过以下方式反馈：

E-mail：info@h3c.com

感谢您的反馈，让我们做得更好！

# 目 录

# 1 IP路由基础

说明

本手册仅介绍单播路由协议，组播路由协议请参见“IP 组播配置指导”。

## 1.1 IP路由简介

在网络中路由器根据所收到的报文的目的地址选择一条合适的路径，并将报文转发到下一个路由器。路径中最后一个路由器负责将报文转发给目的主机。

路由就是报文在转发过程中的路径信息，用来指导报文转发。

根据路由目的地的不同，路由可划分为：

- 网段路由：目的地为网段，子网掩码长度小于 32 位。
- 主机路由：目的地为主机，子网掩码长度为 32 位。

另外，根据目的地与该路由器是否直接相连，路由又可划分为：

- 直接路由：目的地所在网络与路由器直接相连。
- 间接路由：目的地所在网络与路由器非直接相连。

### 1.1.1 路由表

#### 1. 路由表简介

RIB（Routing Information Base，路由信息库），是一个集中管理路由信息的数据库，包含路由表信息以及路由周边信息（路由迭代信息、路由共享信息以及路由扩展信息）等。

路由器通过对路由表进行优选，把优选路由下发到 FIB（Forwarding Information Base，转发信息库）表中，通过 FIB 表指导报文转发。

路由表中保存了各种路由协议发现的路由，根据来源不同，通常分为以下三类：

- 直连路由：链路层协议发现的路由，也称为接口路由。
- 静态路由：网络管理员手工配置的路由。静态路由配置方便，对系统要求低，适用于拓扑结构简单并且稳定的小型网络。其缺点是每当网络拓扑结构发生变化，都需要手工重新配置，不能自动适应。
- 动态路由：路由协议发现的路由。

FIB 表中每条转发项都指明了要到达某子网或某主机的报文应通过路由器的哪个物理接口发送，就可以到达该路径的下一个路由器，或者不需再经过别的路由器便可传送到直接相连的网络中的目的主机。FIB 表的具体内容，请参见“三层技术-IP 业务配置指导”中的“IP 转发基础”。

#### 2. 路由表内容

通过命令 **display ip routing-table** 可以显示路由表的摘要信息，例如：

```
<Sysname> display ip routing-table
```

```
Destinations : 19        Routes : 19

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
0.0.0.0/32          Direct 0    0            127.0.0.1       InLoop0
1.1.1.0/24          Direct 0    0            1.1.1.1         Eth1/1
1.1.1.0/32          Direct 0    0            1.1.1.1         Eth1/1
1.1.1.1/32          Direct 0    0            127.0.0.1       InLoop0
1.1.1.255/32        Direct 0    0            1.1.1.1         Eth1/1
2.2.2.0/24          Static 60   0            12.2.2.2        Eth1/2
80.1.1.0/24         OSPF   10   2            80.1.1.1        Eth1/3
......（省略部分显示信息）
```

路由表中包含了下列关键项：

- Destination：目的地址。用来标识 IP 报文的目的地址或目的网络。
- Mask：网络掩码。与目的地址一起来标识目的主机或路由器所在的网段的地址。将目的地址和网络掩码“逻辑与”后可得到目的主机或路由器所在网段的地址。例如：目的地址为 129.102.8.10、掩码为 255.255.0.0 的主机或路由器所在网段的地址为 129.102.0.0。掩码由若干个连续“1”构成，既可以用点分十进制法表示，也可以用掩码中连续“1”的个数来表示。
- Proto：发现该路由的路由协议类型。可以是直连、静态或者 OSPF 等路由协议。
- Pre：路由优先级。对于同一目的地，可能存在若干条不同下一跳的路由，这些不同的路由可能是由不同的路由协议发现的，也可能是手工配置的静态路由。优先级高（数值小）的路由将成为当前的最优路由。
- Cost：路由的度量值。当到达同一目的地的多条路由具有相同的优先级时，路由的度量值越小的路由将成为当前的最优路由。
- NextHop：下一跳地址。此路由的下一跳 IP 地址。
- Interface：出接口。指明 IP 报文将从该路由器哪个接口转发。

### 1.1.2 路由协议分类

路由协议有自己的路由算法，能够自动适应网络拓扑的变化，适用于具有一定规模的网络拓扑。其缺点是配置比较复杂，对系统的要求高于静态路由，并占用一定的网络资源。

对路由协议的分类可采用以下不同标准。

#### 1. 根据作用范围

- IGP（Interior Gateway Protocol，内部网关协议）：在一个自治系统内部运行，常见的 IGP 协议包括 RIP、OSPF 和 IS-IS。
- EGP（Exterior Gateway Protocol，外部网关协议）：运行于不同自治系统之间，BGP 是目前最常用的 EGP。

---

说明

AS（Autonomous System，自治系统）是在同一技术管理部门下运行的一组路由器。

---

#### 2. 根据使用算法

- 距离矢量（Distance-Vector）协议：包括 RIP 和 BGP。其中，BGP 也被称为路径矢量协议（Path-Vector）。
- 链路状态（Link-State）协议：包括 OSPF 和 IS-IS。

#### 3. 根据目的地址类型

- 单播路由协议：包括 RIP、OSPF、BGP 和 IS-IS 等。
- 组播路由协议：包括 PIM-SM、PIM-DM 等。

#### 4. 根据IP协议版本

- IPv4 路由协议：包括 RIP、OSPF、BGP 和 IS-IS 等。
- IPv6 路由协议：包括 RIPng、OSPFv3、IPv6 BGP 和 IPv6 IS-IS 等。

### 1.1.3 路由优先级

对于相同的目的地，不同的路由协议、直连路由和静态路由可能会发现不同的路由，但这些路由并不都是最优的。为了判断最优路由，各路由协议、直连路由和静态路由都被赋予了一个优先级，具有较高优先级的路由协议发现的路由将成为最优路由。

除直连路由外，各路由协议的优先级都可由用户手工进行配置。另外，每条静态路由的优先级都可以不相同。缺省的路由优先级如 表 1-1 所示，数值越小表明优先级越高。

**表1-1 缺省的路由优先级**

| 路由协议或路由种类 | 缺省的路由优先级 |
|---|---|
| DIRECT（直连路由） | 0 |
| 组播静态路由 | 1 |
| OSPF | 10 |
| IS-IS | 15 |
| 单播静态路由 | 60 |
| RIP | 100 |
| OSPF ASE | 150 |
| OSPF NSSA | 150 |
| IBGP | 255 |
| EBGP | 255 |
| UNKNOWN（来自不可信源端的路由） | 256 |

### 1.1.4 负载分担

对同一路由协议来说，允许配置多条目的地相同且开销也相同的路由。当到同一目的地的路由中，没有更高优先级的路由时，这几条路由都被采纳，在转发去往该目的地的报文时，依次通过各条路径发送，从而实现网络的负载分担。

目前支持负载分担有静态路由/IPv6 静态路由、RIP/RIPng、OSPF/OSPFv3、BGP/IPv6 BGP 和 IS-IS/IPv6 IS-IS。

### 1.1.5 路由备份

使用路由备份可以提高网络的可靠性。用户可根据实际情况，配置到同一目的地的多条路由，其中优先级最高的一条路由作为主路由，其余优先级较低的路由作为备份路由。

正常情况下，路由器采用主路由转发数据。

(1) 当链路出现故障时，主路由变为非激活状态，路由器选择备份路由中优先级最高的转发数据。这样，也就实现了从主路由到备份路由的切换。

(2) 当链路恢复正常时，路由器重新选择路由。由于主路由的优先级最高，路由器选择主路由来发送数据。这就是从备份路由到主路由的切换。

### 1.1.6 路由迭代

对于 BGP 路由（直连 EBGP 路由除外）和静态路由（配置了下一跳）以及多跳 RIP 路由而言，其所携带的下一跳信息可能并不是直接可达，需要找到到达下一跳的直连出接口。路由迭代的过程就是通过路由的下一跳信息来找到直连出接口的过程。

而对于 OSPF 和 IS-IS 等链路状态路由协议而言，其下一跳是直接在路由计算时得到的，不需要进行路由迭代。

路由迭代信息记录并保存路由迭代的结果，包括依赖路由的概要信息、迭代路径、迭代深度等。

### 1.1.7 路由共享

由于各路由协议采用的路由算法不同，不同的路由协议可能会发现不同的路由。如果网络规模较大，当使用多种路由协议时，往往需要在不同的路由协议间能够共享各自发现的路由。

各路由协议都可以引入其它路由协议的路由、直连路由和静态路由，具体内容请参见本手册中各路由协议模块有关引入外部路由的描述。

路由共享信息记录了路由协议之间的引入关系。

### 1.1.8 路由扩展

路由扩展属性主要是指 BGP 路由的扩展团体属性以及 OSPF 路由的区域 ID、路由类型和 Router ID 等。同路由共享一样，路由协议可以引入其它路由协议的路由扩展属性。

路由扩展信息记录了各路由协议的路由扩展属性以及路由协议扩展属性之间的引入关系。

## 1.2 配置路由的最大存活时间

### 1.2.1 配置路由和标签在RIB中的最大存活时间

当协议路由表项较多或协议 GR 时间较长时，由于协议收敛速度较慢，可能会出现协议路由表项提前老化的问题。通过调节路由和标签在 RIB 中的最大存活时间，可以解决上面的问题。

### 1. 配置IPv4 路由和标签在RIB中的最大存活时间

提示

该配置在下一次协议进程倒换或者 RIB 进程倒换时才生效。

表1-2 配置 IPv4 路由和标签在 RIB 中的最大存活时间

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIB视图 | **rib** | - |
| 创建RIB IPv4地址族，并进入RIB IPv4地址族视图 | **address-family ipv4** | 缺省情况下，没有创建RIB IPv4地址族 |
| 配置IPv4路由和标签在RIB中的最大存活时间 | **protocol** *protocol* **lifetime** *seconds* | 缺省情况下，IPv4路由和标签在RIB中的最大存活时间为480秒 |

### 2. 配置IPv6 路由和标签在RIB中的最大存活时间

提示

该配置在下一次协议进程倒换或者 RIB 进程倒换时才生效。

表1-3 配置 IPv6 路由和标签在 RIB 中的最大存活时间

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIB视图 | **rib** | - |
| 创建RIB IPv6地址族，并进入RIB IPv6地址族视图 | **address-family ipv6** | 缺省情况下，没有创建RIB IPv6地址族 |
| 配置IPv6路由和标签在RIB中的最大存活时间 | **protocol** *protocol* **lifetime** *seconds* | 缺省情况下，IPv6路由和标签在RIB中的最大存活时间为480秒 |

## 1.2.2 配置路由在FIB中的最大存活时间

当协议进程倒换或 RIB 进程倒换后，如果协议进程没有配置 GR 或 NSR，需要多保留一段时间 FIB 表项；如果协议进程配置了 GR 或 NSR，需要立刻删除 FIB 表项，避免 FIB 表项长时间存在导致问题。通过调节路由在 FIB 中的最大存活时间，可以解决上面的问题。

### 1. 配置IPv4 路由在FIB中的最大存活时间

表1-4 配置 IPv4 路由在 FIB 中的最大存活时间

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIB视图 | **rib** | - |
| 创建RIB IPv4地址族，并进入RIB IPv4地址族视图 | **address-family ipv4** | 缺省情况下，没有创建RIB IPv4地址族 |
| 配置IPv4路由在FIB中的最大存活时间 | **fib lifetime** *seconds* | 缺省情况下，IPv4路由在FIB中的最大存活时间为600秒 |

### 2. 配置IPv6 路由在FIB中的最大存活时间

表1-5 配置 IPv6 路由在 FIB 中的最大存活时间

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIB视图 | **rib** | - |
| 创建RIB IPv6地址族，并进入RIB IPv6地址族视图 | **address-family ipv6** | 缺省情况下，没有创建RIB IPv6地址族 |
| 配置IPv6路由在FIB中的最大存活时间 | **fib lifetime** *seconds* | 缺省情况下，IPv6路由在FIB中的最大存活时间为600秒 |

## 1.3 路由表显示和维护

在任意视图下执行 **display** 命令可以显示路由表信息。在用户视图下执行 **reset** 命令可以清除路由表的统计信息。

表1-6 路由表显示和维护

| 操作 | 命令 |
| --- | --- |
| 显示路由表的信息（MSR 2600/MSR 3600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] [ **verbose** ] |
| 显示路由表的信息（MSR 5600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示通过指定基本访问控制列表过滤的路由信息（MSR 2600/MSR 3600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] **acl** *acl-number* [ **verbose** ] |
| 显示通过指定基本访问控制列表过滤的路由信息（MSR 5600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] **acl** *acl-number* [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示指定目的地址的路由（MSR 2600/MSR 3600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] *ip-address* [ *mask* \| *mask-length* ] [ **longer-match** ] [ **verbose** ] |

| 操作 | 命令 |
| --- | --- |
| 显示指定目的地址的路由（MSR 5600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] *ip-address* [ *mask* \| *mask-length* ] [ **longer-match** ] [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示指定目的地址范围内的路由（MSR 2600/MSR 3600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] *ip-address1* **to** *ip-address2* [ **verbose** ] |
| 显示指定目的地址范围内的路由（MSR 5600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] *ip-address1* **to** *ip-address2* [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示通过指定前缀列表过滤的路由信息（MSR 2600/MSR 3600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] **prefix-list** *prefix-list-name* [ **verbose** ] |
| 显示通过指定前缀列表过滤的路由信息（MSR 5600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] **prefix-list** *prefix-list-name* [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示指定协议生成或发现的路由信息（MSR 2600/MSR 3600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] **protocol** *protocol* [ **inactive** \| **verbose** ] |
| 显示指定协议生成或发现的路由信息（MSR 5600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] **protocol** *protocol* [ **inactive** \| **verbose** ] [ **standby slot** *slot-number* ] |
| 显示路由表中的综合路由统计信息（MSR 2600/MSR 3600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] **statistics** |
| 显示路由表中的综合路由统计信息（MSR 5600） | **display ip routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] **statistics**[ **standby slot** *slot-number* ] |
| 显示RIB的路由属性信息（MSR 2600/MSR 3600） | **display rib attribute** [ *attribute-id* ] |
| 显示RIB的路由属性信息（MSR 5600） | **display rib attribute** [ *attribute-id* ] [ **standby slot** *slot-number* ] |
| 显示RIB的GR状态信息 | **display rib graceful-restart** |
| 显示RIB的下一跳信息（MSR 2600/MSR 3600） | **display rib nib** [ **self-originated** ] [ *nib-id* ] [ **verbose** ]<br>**display rib nib protocol** *protocol-name* [ **verbose** ] |
| 显示RIB的下一跳信息（MSR 5600） | **display rib nib** [ **self-originated** ] [ *nib-id* ] [ **verbose** ] [ **standby slot** *slot-number* ]<br>**display rib nib protocol** *protocol-name* [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示直连路由下一跳信息 | **display route-direct nib** [ *nib-id* ] [ **verbose** ] |
| 清除路由表中的综合路由统计信息（MSR 2600/MSR 3600） | **reset ip routing-table statistics protocol** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] { *protocol* \| **all** } |
| 清除路由表中的综合路由统计信息（MSR 5600） | **reset ip routing-table statistics protocol** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] { *protocol* \| **all** } [ **standby slot** *slot-number* ] |
| 显示IPv6路由表的信息（MSR 2600/MSR 3600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] [ **verbose** ] |
| 显示IPv6路由表的信息（MSR 5600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] [ **verbose** ] [ **standby slot** *slot-number* ] |

| 操作 | 命令 |
| --- | --- |
| 显示指定目的地址的IPv6路由信息（MSR 2600/MSR 3600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] *ipv6-address* [ *prefix-length* ] [ **longer-match** ] [ **verbose** ] |
| 显示指定目的地址的IPv6路由信息（MSR 5600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] *ipv6-address* [ *prefix-length* ] [ **longer-match** ] [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示通过指定基本IPv6 ACL过滤的IPv6路由信息（MSR 2600/MSR 3600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] **acl** *acl6-number* [ **verbose** ] |
| 显示通过指定基本IPv6 ACL过滤的IPv6路由信息（MSR 5600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] **acl** *acl6-number* [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示指定目的地址范围内的IPv6路由信息（MSR 2600/MSR 3600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] *ipv6-address1* **to** *ipv6-address2* [ **verbose** ] |
| 显示指定目的地址范围内的IPv6路由信息（MSR 5600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] *ipv6-address1* **to** *ipv6-address2* [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示通过指定前缀列表过滤的IPv6路由信息（MSR 2600/MSR 3600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] **prefix-list** *prefix-list-name* [ **verbose** ] |
| 显示通过指定前缀列表过滤的IPv6路由信息（MSR 5600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] **prefix-list** *prefix-list-name* [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示指定协议生成或发现的IPv6路由信息（MSR 2600/MSR 3600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] **protocol** *protocol* [ **inactive** \| **verbose** ] |
| 显示指定协议生成或发现的IPv6路由信息（MSR 5600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] **protocol** *protocol* [ **inactive** \| **verbose** ] [ **standby slot** *slot-number* ] |
| 显示IPv6路由表中的综合路由统计信息（MSR 2600/MSR 3600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] **statistics** |
| 显示IPv6路由表中的综合路由统计信息（MSR 5600） | **display ipv6 routing-table** [ **vpn-instance** *vpn-instance-name* ] **statistics** [ **standby slot** *slot-number* ] |
| 显示IPv6 RIB的路由属性信息（MSR 2600/MSR 3600） | **display ipv6 rib attribute** [ *attribute-id* ] |
| 显示IPv6 RIB的路由属性信息（MSR 5600） | **display ipv6 rib attribute** [ *attribute-id* ] [ **standby slot** *slot-number* ] |
| 显示IPv6 RIB的GR状态信息 | **display ipv6 rib graceful-restart** |
| 显示IPv6 RIB的下一跳信息（MSR 2600/MSR 3600） | **display ipv6 rib nib** [ **self-originated** ] [ *nib-id* ] [ **verbose** ]<br>**display ipv6 rib nib protocol** *protocol-name* [ **verbose** ] |
| 显示IPv6 RIB的下一跳信息（MSR 5600） | **display ipv6 rib nib** [ **self-originated** ] [ *nib-id* ] [ **verbose** ] [ **standby slot** *slot-number* [ **cpu** *cpu-number* ] ]<br>**display ipv6 rib nib protocol** *protocol-name* [ **verbose** ] [ **standby slot** *slot-number* ] |
| 显示IPv6直连路由下一跳信息 | **display ipv6 route-direct nib** [ *nib-id* ] [ **verbose** ] |
| 清除IPv6路由表中的综合路由统计信息（MSR 2600/MSR 3600） | **reset ipv6 routing-table statistics protocol** [ **vpn-instance** *vpn-instance-name* ] { *protocol* \| **all** } |
| 清除IPv6路由表中的综合路由统计信息（MSR 5600） | **reset ipv6 routing-table statistics protocol** [ **vpn-instance** *vpn-instance-name* ] { *protocol* \| **all** } [ **standby slot** *slot-number* ] |

## 1.4 配置路由NSR

MSR 系列路由器各款型对于本节所描述的特性的支持情况有所不同，详细差异信息如下：

| 型号 | 特性 | 描述 |
| --- | --- | --- |
| MSR 2600 | 配置路由NSR | 不支持 |
| MSR 3600 | | 不支持 |
| MSR 5600 | | 支持 |

提示

配置本功能的同时，请配置协议的 GR 或 NSR 功能，否则可能导致路由老化和流量中断。

NSR（Nonstop Routing，不间断路由）将路由信息从主进程备份到备进程，在设备发生主备倒换时保证路由信息不丢失，解决了主备倒换期间引发的路由震荡问题，保证转发业务不中断。

路由 NSR 相对于协议 NSR 功能，主备倒换时路由收敛速度更快。

### 1.4.1 配置IPv4 路由NSR

表1-7 配置 IPv4 路由 NSR

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIB视图 | **rib** | - |
| 创建RIB IPv4地址族，并进入RIB IPv4地址族视图 | **address-family ipv4** | 缺省情况下，没有创建RIB IPv4地址族 |
| 配置IPv4路由NSR | **non-stop-routing** | 缺省情况下，未使能NSR功能 |

### 1.4.2 配置IPv6 路由NSR

表1-8 配置 IPv6 路由 NSR

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIB视图 | **rib** | - |
| 创建RIB IPv6地址族，并进入RIB IPv6地址族视图 | **address-family ipv6** | 缺省情况下，没有创建RIB IPv6地址族 |
| 配置IPv6路由NSR | **non-stop-routing** | 缺省情况下，未使能NSR功能 |

# 目 录

# 1 静态路由

## 1.1 静态路由简介

静态路由是一种特殊的路由，由管理员手工配置。当网络结构比较简单时，只需配置静态路由就可以使网络正常工作。

静态路由不能自动适应网络拓扑结构的变化。当网络发生故障或者拓扑发生变化后，必须由网络管理员手工修改配置。

## 1.2 配置静态路由

### 1.2.1 配置准备

在配置静态路由之前，需完成以下任务：

- 配置相关接口的物理参数
- 配置相关接口的链路层属性
- 配置相关接口的 IP 地址

### 1.2.2 配置静态路由

表1-1 配置静态路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置静态路由 | **ip route-static** *dest-address* { *mask-length* \| *mask* } { *interface-type interface-number* [ *next-hop-address* ] \| *next-hop-address* [ **track** *track-entry-number* ] \| **vpn-instance** *d-vpn-instance-name next-hop-address* [ **track** *track-entry-number* ] } [ **permanent** ] [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 三者选其一<br>缺省情况下，没有配置静态路由 |
| | **ip route-static vpn-instance** *s-vpn-instance-name dest-address* { *mask-length* \| *mask* } { *interface-type interface-number* [ *next-hop-address* ] \| *next-hop-address* [ **public** ] [ **track** *track-entry-number* ] \| **vpn-instance** *d-vpn-instance-name next-hop-address* [ **track** *track-entry-number* ] } [ **permanent** ] [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | |
| | **ip route-static topology** *topo-name dest-address* { *mask* \| *mask-length* } { *next-hop-address* \| *interface-type interface-number* [ *next-hop-address* ] } [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| （可选）配置静态路由的缺省优先级 | **ip route-static default-preference** *default-preference-value* | 缺省情况下，静态路由的缺省优先级为60 |
| （可选）删除所有静态路由 | **delete** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] **static-routes all** | - |

说明

- 通过在 Track 模块和静态路由之间建立联动，可以实现实时监测下一跳的可达性，以便及时判断静态路由是否有效。关于 Track 的详细介绍，请参见“可靠性配置指导”中的“Track”。
- 使用 **undo ip route-static** 命令可以删除一条静态路由，而使用 **delete static-routes all** 命令可以删除包括缺省路由在内的所有静态路由。

## 1.3 配置静态路由与BFD联动

注意

路由振荡时，使能 BFD 功能可能会加剧振荡，请谨慎使用。

BFD（Bidirectional Forwarding Detection，双向转发检测）提供了一个通用的、标准化的、介质无关、协议无关的快速故障检测机制，可以为上层协议（如路由协议、MPLS 等）统一地快速检测两台路由器间双向转发路径的故障。

关于 BFD 的详细介绍，请参见“可靠性配置指导”中的“BFD”。

### 1.3.1 双向检测

双向检测，即本端和对端需要同时进行配置，通过控制报文检测两个方向上的链路状态，实现毫秒级别的链路故障检测。

双向检测支持直连下一跳和非直连下一跳。

#### 1. 直连下一跳

直连下一跳是指下一跳和本端是直连的，配置时必须指定出接口和下一跳。

表1-2 配置静态路由与 BFD 联动（双向检测—直连）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置静态路由与BFD联动 | **ip route-static** *dest-address* { *mask-length* \| *mask* } *interface-type interface-number next-hop-address* **bfd control-packet** [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 二者选其一<br>缺省情况下，没有配置静态路由 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| | **ip route-static vpn-instance** *s-vpn-instance-name dest-address* { *mask-length* \| *mask* } *interface-type interface-number next-hop-address* **bfd control-packet** [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 与BFD联动 |

### 2. 非直连下一跳

非直连下一跳是指下一跳和本端不是直连的，中间还有其它设备。配置时必须指定下一跳和 BFD 源 IP 地址。

**表1-3 配置静态路由与 BFD 联动（双向检测—非直连）**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置静态路由与BFD联动 | **ip route-static** *dest-address* { *mask-length* \| *mask* } { *next-hop-address* **bfd control-packet bfd-source** *ip-address* \| **vpn-instance** *d-vpn-instance-name next-hop-address* **bfd control-packet bfd-source** *ip-address* } [ **preference** *preference-value* ] [ **tag tag-value** ] [ **description** *description-text* ] | 二者选其一<br>缺省情况下，没有配置静态路由与BFD联动 |
| | **ip route-static vpn-instance** *s-vpn-instance-name dest-address* { *mask-length* \| *mask* } { *next-hop-address* **bfd control-packet bfd-source** *ip-address* \| **vpn-instance** *d-vpn-instance-name next-hop-address* **bfd control-packet bfd-source** *ip-address* } [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | |

## 1.3.2 单跳检测

单跳检测，即只需要本端进行配置，通过 echo 报文检测链路的状态。echo 报文的目的地址为本端接口地址，发送给下一跳设备后会直接转发回本端。这里所说的“单跳”是 IP 的一跳。

**表1-4 配置静态路由与 BFD 联动（单跳检测）**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置echo报文的源IP地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置echo报文的源IP地址<br>本命令的详细情况请参见“可靠性命令参考”中的“BFD” |
| 配置静态路由与BFD联动 | **ip route-static** *dest-address* { *mask-length* \| *mask* } *interface-type interface-number next-hop-address* **bfd echo-packet** [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 二者选其一<br>缺省情况下，没有配置静态路由 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| | **ip route-static vpn-instance** *s-vpn-instance-name dest-address* { *mask-length* \| *mask* } *interface-type interface-number next-hop-address* **bfd echo-packet** [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 与BFD联动 |

说明

静态路由的出接口为处于 SPOOFING 状态时，不能使用 BFD 进行检测。

## 1.4 配置静态路由快速重路由功能

提示

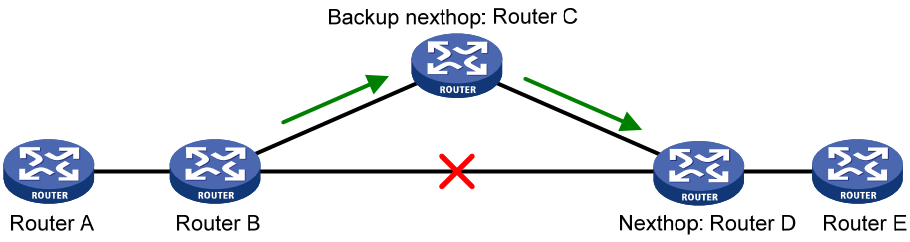
- 静态路由快速重路由功能不能与静态路由 BFD 功能同时使用。
- 等价路由不支持配置静态路由快速重路由功能。

### 1.4.1 功能简介

当网络中的链路或某台路由器发生故障时，需要通过故障链路或故障路由器传输才能到达目的地的报文将会丢失或产生路由环路，数据流量将会被中断。

为了尽可能避免网络故障导致的流量中断，网络管理员可以根据需要配置静态路由快速重路由功能。

图1-1 静态路由快速重路由功能示意图




如 图 1-1 所示，通过配置快速重路由功能，网络管理员可以为路由指定备份下一跳，也可以在存在低优先级静态路由的情况下，使能自动快速重路由功能，查找满足条件的低优先级路由的下一跳作为主路由的备份下一跳，当路由器检测到网络故障时，路由器会使用事先配置好的备份下一跳替换失效下一跳，通过备份下一跳来指导报文的转发，从而避免了流量中断。

### 1.4.2 配置限制和指导

本功能只适合在主链路三层接口 up，主链路由双通变为单通或者不通的情况下使用。在主链路三层接口 down 的情况下，本功能不可用。

单通现象，即一条链路上的两端，有且只有一端可以收到另一端发来的报文，此链路称为单向链路。

### 1.4.3 配置步骤

#### 1. 配置静态路由快速重路由功能

(1) 配置静态路由快速重路由功能（手工指定备份下一跳）

表1-5 配置静态路由快速重路由功能（手工指定备份下一跳）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置静态路由快速重路由功能 | **ip route-static** *dest-address* { *mask-length* \| *mask* } *interface-type interface-number* [ *next-hop-address* [ **backup-interface** *interface-type interface-number* [ **backup-nexthop** *backup-nexthop-address* ] ] ] [ **permanent** ] [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 三者选其一<br>缺省情况下，没有配置静态路由快速重路由功能 |
| | **ip route-static vpn-instance** *s-vpn-instance-name dest-address* { *mask-length* \| *mask* } *interface-type interface-number* [ *next-hop-address* [ **backup-interface** *interface-type interface-number* [ **backup-nexthop** *backup-nexthop-address* ] ] ] [ **permanent** ] [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | |
| | **ip route-static topology** *topo-name dest-address* { *mask* \| *mask-length* } { *next-hop-address* \| *interface-type interface-number* [ *next-hop-address* [ **backup-interface** *interface-type interface-number* **backup-nexthop** *backup-nexthop-address* ] ] } [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | |

说明

- 静态路由配置的备份出接口拔出或者删除时，配置的路由会失效。
- 备份出接口和下一跳不能直接修改，且不能和主出接口和下一跳相同。

(2) 配置静态路由快速重路由功能（自动查找备份下一跳）

表1-6 配置静态路由快速重路由功能（自动查找备份下一跳）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置静态路由自动快速重路由功能 | **ip route-static fast-reroute auto** | 缺省情况下，没有配置静态路由自动快速重路由功能 |

#### 2. 配置静态路由快速重路由支持BFD检测功能

静态路由的快速重路由特性中，主用链路缺省不使用 BFD 进行链路故障检测。配置本功能后，将使用 BFD（Echo 方式）进行检测，可以更快地检测到链路故障。

表1-7 配置静态路由快速重路由支持 BFD 检测功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置BFD Echo报文源地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置BFD Echo报文源地址 |
| 使能静态路由中主用链路的BFD（Echo方式）检测功能 | **ip route-static primary-path-detect bfd echo** | 缺省情况下，静态路由中主用链路的BFD（Echo方式）检测功能处于关闭状态 |

## 1.5 静态路由显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令查看静态路由配置的运行情况并检验配置结果。

表1-8 静态路由显示和维护

| 操作 | 命令 |
| --- | --- |
| 查看静态路由表信息（本命令的详细情况请参见“三层技术-IP路由命令参考”中的“IP路由基础”） | **display ip routing-table protocol static** [ **inactive** \| **verbose** ] |
| 显示静态路由下一跳信息 | **display route-static nib** [ *nib-id* ] [ **verbose** ] |
| 显示静态路由表信息 | **display route-static routing-table** [ **topology** *topo-name* \| **vpn-instance** *vpn-instance-name* ] [ *ip-address* { *mask-length* \| *mask* } ] |

## 1.6 静态路由典型配置举例

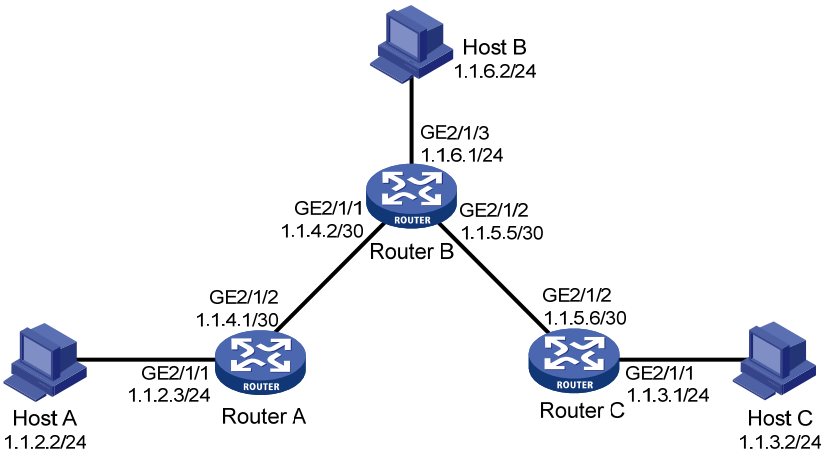
### 1.6.1 静态路由基本功能配置举例

#### 1. 组网需求

路由器各接口及主机的IP地址和掩码如 图 1-2 所示。要求采用静态路由，使图中任意两台主机之间都能互通。

## 2. 组网图

图1-2 静态路由基本功能配置组网图




## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置静态路由

# 在 Router A 上配置缺省路由。

```
<RouterA> system-view
[RouterA] ip route-static 0.0.0.0 0.0.0.0 1.1.4.2
```

# 在 Router B 上配置两条静态路由。

```
<RouterB> system-view
[RouterB] ip route-static 1.1.2.0 255.255.255.0 1.1.4.1
[RouterB] ip route-static 1.1.3.0 255.255.255.0 1.1.5.6
```

# 在 Router C 上配置缺省路由。

```
<RouterC> system-view
[RouterC] ip route-static 0.0.0.0 0.0.0.0 1.1.5.5
```

(3) 配置主机

配置 Host A 的缺省网关为 1.1.2.3，Host B 的缺省网关为 1.1.6.1，Host C 的缺省网关为 1.1.3.1，具体配置过程略。

## 4. 验证配置

# 查看 Router A 的静态路由信息。

```
[RouterA] display ip routing-table protocol static

Summary Count : 1

Static Routing table Status : <Active>
Summary Count : 1

Destination/Mask    Proto  Pre  Cost        NextHop         Interface
0.0.0.0/0           Static 60   0           1.1.4.2         GE2/1/2
```

```
Static Routing table Status : <Inactive>
Summary Count : 0
```

# 查看 Router B 的静态路由信息。

```
[RouterB] display ip routing-table protocol static

Summary Count : 2

Static Routing table Status : <Active>
Summary Count : 2

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
1.1.2.0/24          Static 60   0            1.1.4.1         GE2/1/1
1.1.3.0/24          Static 60   0            1.1.5.6         GE2/1/2

Static Routing table Status : <Inactive>
Summary Count : 0
```

# 在 Host B 上使用 **ping** 命令验证 Host A 是否可达（假定主机安装的操作系统为 Windows XP）。

```
C:\Documents and Settings\Administrator>ping 1.1.2.2

Pinging 1.1.2.2 with 32 bytes of data:

Reply from 1.1.2.2: bytes=32 time=1ms TTL=126
Reply from 1.1.2.2: bytes=32 time=1ms TTL=126
Reply from 1.1.2.2: bytes=32 time=1ms TTL=126
Reply from 1.1.2.2: bytes=32 time=1ms TTL=126

Ping statistics for 1.1.2.2:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 1ms, Maximum = 1ms, Average = 1ms
```

# 在 Host B 上使用 **tracert** 命令验证 Host A 是否可达。

```
C:\Documents and Settings\Administrator>tracert 1.1.2.2

Tracing route to 1.1.2.2 over a maximum of 30 hops

  1    <1 ms    <1 ms    <1 ms  1.1.6.1
  2    <1 ms    <1 ms    <1 ms  1.1.4.1
  3     1 ms    <1 ms    <1 ms  1.1.2.2

Trace complete.
```

## 1.6.2 配置静态路由与BFD联动（直连）

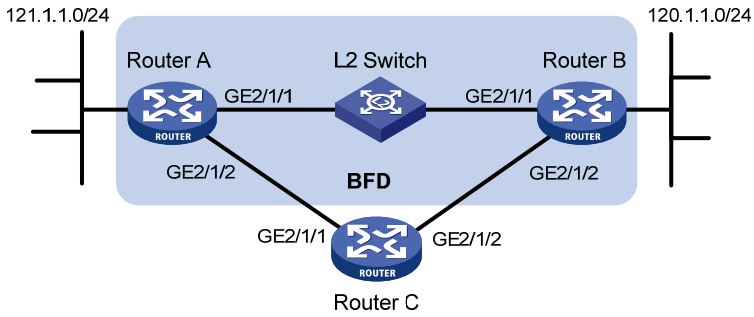
### 1. 组网需求

- 在 Router A 上配置静态路由可以到达 120.1.1.0/24 网段，在 Router B 上配置静态路由可以到达 121.1.1.0/24 网段，并使能 BFD 检测功能。
- 在 Router C 上配置静态路由可以到达 120.1.1.0/24 网段和 121.1.1.0/24 网段。

- 当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时，BFD 能够快速感知，并且切换到 Router C 进行通信。

## 2. 组网图

图1-3 静态路由与 BFD 联动（直连）配置组网图




| 设备 | 接口 | IP地址 | 设备 | 接口 | IP地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 12.1.1.1/24 | Router B | GE2/1/1 | 12.1.1.2/24 |
| | GE2/1/2 | 10.1.1.102/24 | | GE2/1/2 | 13.1.1.1/24 |
| Router C | GE2/1/1 | 10.1.1.100/24 | | | |
| | GE2/1/2 | 13.1.1.2/24 | | | |

## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置静态路由和 BFD

# 在 Router A 上配置静态路由，并使能 BFD 检测功能，使用双向检测方式。

```
<RouterA> system-view
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterA-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterA-GigabitEthernet2/1/1] bfd detect-multiplier 9
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] ip route-static 120.1.1.0 24 gigabitethernet 2/1/1 12.1.1.2 bfd control-packet
[RouterA] ip route-static 120.1.1.0 24 gigabitethernet 2/1/2 10.1.1.100 preference 65
[RouterA] quit
```

# 在 Router B 上配置静态路由，并使能 BFD 检测功能，使用双向检测方式。

```
<RouterB> system-view
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterB-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterB-GigabitEthernet2/1/1] bfd detect-multiplier 9
[RouterB-GigabitEthernet2/1/1] quit
[RouterB] ip route-static 121.1.1.0 24 gigabitethernet 2/1/1 12.1.1.1 bfd control-packet
[RouterB] ip route-static 121.1.1.0 24 gigabitethernet 2/1/2 13.1.1.2 preference 65
[RouterB] quit
```

# 在 Router C 上配置静态路由。

```
<RouterC> system-view
[RouterC] ip route-static 120.1.1.0 24 13.1.1.1
```

```
[RouterC] ip route-static 121.1.1.0 24 10.1.1.102
```

#### 4. 验证配置

下面以 Router A 为例，Router B 和 Router A 类似，不再赘述。

# 查看 BFD 会话，可以看到 BFD 会话已经创建。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv4 Session Working Under Ctrl Mode:

 LD/RD          SourceAddr      DestAddr        State     Holdtime     Interface
 4/7            12.1.1.1        12.1.1.2        Up        2000ms       GE2/1/1
```

# 查看静态路由，可以看到 Router A 经过 L2 Switch 到达 Router B。

```
<RouterA> display ip routing-table protocol static

Summary Count : 1

Static Routing table Status : <Active>
Summary Count : 1

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
120.1.1.0/24        Static 60   0            12.1.1.2        GE2/1/1

Static Routing table Status : <Inactive>
Summary Count : 0
```

当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时：

# 查看静态路由，可以看到 Router A 经过 Router C 到达 Router B。

```
<RouterA> display ip routing-table protocol static

Summary Count : 1

Static Routing table Status : <Active>
Summary Count : 1

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
120.1.1.0/24        Static 65   0            10.1.1.100      GE2/1/2

Static Routing table Status : <Inactive>
Summary Count : 0
```

### 1.6.3 配置静态路由与BFD联动（非直连）

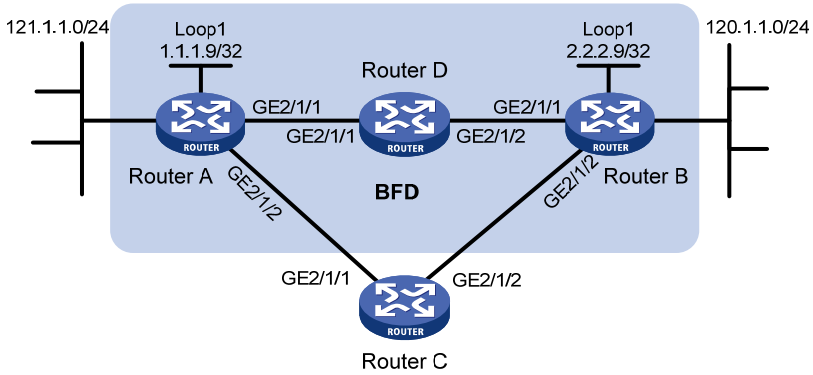
#### 1. 组网需求

- 在 Router A 上配置静态路由可以到达 120.1.1.0/24 网段，在 Router B 上配置静态路由可以到达 121.1.1.0/24 网段，并使能 BFD 检测功能。
- 在 Router C 和 Router D 上配置静态路由可以到达 120.1.1.0/24 网段和 121.1.1.0/24 网段。

- Router A 存在到 Router B 的接口 Loopback1（2.2.2.9/32）的路由，出接口为 GigabitEthernet2/1/1；Router B 存在到 Router A 的接口 Loopback1（1.1.1.9/32）的路由，出接口为 GigabitEthernet2/1/1；Router D 存在到 1.1.1.9/32 的路由，出接口为 GigabitEthernet2/1/1，存在到 2.2.2.9/32 的路由，出接口为 GigabitEthernet2/1/2。
- 当 Router A 和 Router B 通过 Router D 通信的链路出现故障时，BFD 能够快速感知，并且切换到 Router C 进行通信。

## 2. 组网图

图1-4 静态路由与 BFD 联动（非直连）配置组网图




| 设备 | 接口 | IP地址 | 设备 | 接口 | IP地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 12.1.1.1/24 | Router B | GE2/1/1 | 11.1.1.2/24 |
| | GE2/1/2 | 10.1.1.102/24 | | GE2/1/2 | 13.1.1.2/24 |
| | Loop1 | 1.1.1.9/32 | | Loop1 | 2.2.2.9/32 |
| Router C | GE2/1/1 | 10.1.1.100/24 | Router D | GE2/1/1 | 12.1.1.2/24 |
| | GE2/1/2 | 13.1.1.2/24 | | GE2/1/2 | 11.1.1.1/24 |

## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置静态路由和 BFD

# 在 Router A 上配置静态路由，并使能 BFD 检测功能，使用双向检测方式。

```
<RouterA> system-view
[RouterA] bfd multi-hop min-transmit-interval 500
[RouterA] bfd multi-hop min-receive-interval 500
[RouterA] bfd multi-hop detect-multiplier 9
[RouterA] ip route-static 120.1.1.0 24 2.2.2.9 bfd control-packet bfd-source 1.1.1.9
[RouterA] ip route-static 120.1.1.0 24 gigabitethernet 2/1/2 10.1.1.100 preference 65
[RouterA] quit
```

# 在 Router B 上配置静态路由，并使能 BFD 检测功能，使用双向检测方式。

```
<RouterB> system-view
[RouterB] bfd multi-hop min-transmit-interval 500
[RouterB] bfd multi-hop min-receive-interval 500
[RouterB] bfd multi-hop detect-multiplier 9
[RouterB] ip route-static 121.1.1.0 24 1.1.1.9 bfd control-packet bfd-source 2.2.2.9
[RouterB] ip route-static 121.1.1.0 24 gigabitethernet 2/1/2 13.1.1.2 preference 65
[RouterB] quit
```

# 在 Router C 上配置静态路由。

```
<RouterC> system-view
[RouterC] ip route-static 120.1.1.0 24 13.1.1.1
[RouterC] ip route-static 121.1.1.0 24 10.1.1.102
```

# 在 Router D 上配置静态路由。

```
<RouterD> system-view
[RouterD] ip route-static 120.1.1.0 24 11.1.1.2
[RouterD] ip route-static 121.1.1.0 24 12.1.1.1
```

### 4. 验证配置

下面以 Router A 为例，Router B 和 Router A 类似，不再赘述。

# 查看 BFD 会话，可以看到 BFD 会话已经创建。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv4 Session Working Under Ctrl Mode:

 LD/RD          SourceAddr      DestAddr        State    Holdtime    Interface
 4/7            1.1.1.9         2.2.2.9         Up       2000ms      N/A
```

# 查看静态路由，可以看到 Router A 经过 Router D 到达 Router B。

```
<RouterA> display ip routing-table protocol static

Summary Count : 1

Static Routing table Status : <Active>
Summary Count : 1

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
120.1.1.0/24        Static 60   0            12.1.1.2        GE2/1/1

Static Routing table Status : <Inactive>
Summary Count : 0
```

当 Router A 和 Router B 通过 Router D 通信的链路出现故障时：

# 查看静态路由，可以看到 Router A 经过 Router C 到达 Router B。

```
<RouterA> display ip routing-table protocol static

Summary Count : 1

Static Routing table Status : <Active>
Summary Count : 1

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
120.1.1.0/24        Static 65   0            10.1.1.100      GE2/1/2

Static Routing table Status : <Inactive>
Summary Count : 0
```

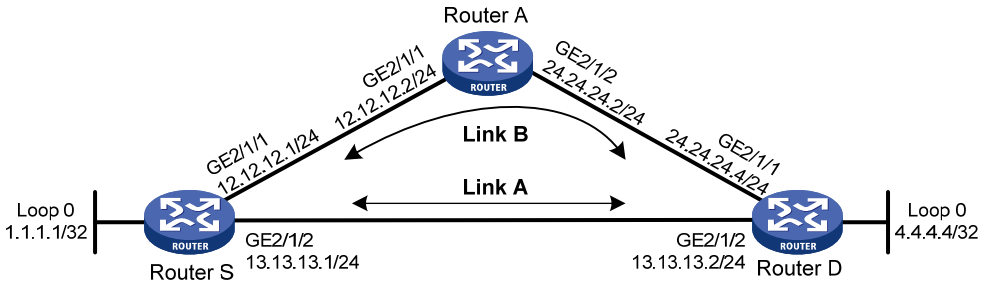
### 1.6.4 静态路由快速重路由配置举例

#### 1. 组网需求

如图 1-5所示，Router S、Router A和Router D通过静态路由实现网络互连。要求当Router S和Router D之间的链路A出现单通故障时，业务可以快速切换到链路B上。

#### 2. 组网图

图1-5 静态路由快速重路由配置组网图




#### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置链路 A 上的静态路由快速重路由

静态路由支持快速重路由配置有两种方法，可以任选一种

方法一：配置静态路由快速重路由功能（手工指定备份下一跳）

# 在 Router S 上配置静态路由，并指定备份出接口和下一跳。

```
<RouterS> system-view
[RouterS] ip route-static 4.4.4.4 32 gigabitethernet 2/1/2 13.13.13.2 backup-interface
gigabitethernet 2/1/1 backup-nexthop 12.12.12.2
```

# 在 Router D 上配置静态路由，并指定备份出接口和下一跳。

```
<RouterD> system-view
[RouterD] ip route-static 1.1.1.1 32 gigabitethernet 2/1/2 13.13.13.1 backup-interface
gigabitethernet 2/1/1 backup-nexthop 24.24.24.2
```

方法二：配置静态路由快速重路由功能（自动查找备份下一跳）

# 在 Router S 上配置静态路由，并配置静态路由自动快速重路由功能。

```
<RouterS> system-view
[RouterS] ip route-static 4.4.4.4 32 gigabitethernet 2/1/2 13.13.13.2
[RouterS] ip route-static 4.4.4.4 32 gigabitethernet 2/1/1 12.12.12.2 preference 70
[RouterS] ip route-static fast-reroute auto
```

# 在 Router D 上配置静态路由，并配置静态路由自动快速重路由功能。

```
<RouterD> system-view
[RouterD] ip route-static 1.1.1.1 32 gigabitethernet 2/1/2 13.13.13.1
[RouterD] ip route-static 1.1.1.1 32 gigabitethernet 2/1/1 24.24.24.2 preference 70
[RouterD] ip route-static fast-reroute auto
```

(3) 配置链路 B 上的静态路由

# 在 Router A 上配置静态路由。

```
<RouterA> system-view
```

```
[RouterA] ip route-static 4.4.4.4 32 gigabitethernet 2/1/2 24.24.24.4
[RouterA] ip route-static 1.1.1.1 32 gigabitethernet 2/1/1 12.12.12.1
```

### 4. 验证配置

# 在 Router S 上查看 4.4.4.4/32 路由，可以看到备份下一跳信息。

```
[RouterS] display ip routing-table 4.4.4.4 verbose

Summary Count : 1

Destination: 4.4.4.4/32
   Protocol: Static          Process ID: 0
  SubProtID: 0x0                    Age: 04h20m37s
       Cost: 0               Preference: 60
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 0.0.0.0
      Flags: 0x1008c        OrigNextHop: 13.13.13.2
      Label: NULL           RealNextHop: 13.13.13.2
    BkLabel: NULL             BkNextHop: 12.12.12.2
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: GigabitEthernet2/1/1
```

# 在 Router D 上查看 1.1.1.1/32 路由，可以看到备份下一跳信息。

```
[RouterS] display ip routing-table 1.1.1.1 verbose

Summary Count : 1

Destination: 1.1.1.1/32
   Protocol: Static          Process ID: 0
  SubProtID: 0x0                    Age: 04h20m37s
       Cost: 0               Preference: 10
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 0.0.0.0
      Flags: 0x1008c        OrigNextHop: 13.13.13.1
      Label: NULL           RealNextHop: 13.13.13.1
    BkLabel: NULL             BkNextHop: 24.24.24.2
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: GigabitEthernet2/1/1
```

# 2 缺省路由

## 2.1 缺省路由简介

缺省路由是在路由器没有找到匹配的路由表项时使用的路由。

如果报文的目的地不在路由表中且没有配置缺省路由，那么该报文将被丢弃，将向源端返回一个ICMP 报文报告该目的地址或网络不可达。

缺省路由有两种生成方式：

- 第一种是网络管理员手工配置。配置请参见 表 1-1，将目的地址与掩码配置为全零（0.0.0.0 0.0.0.0）。
- 第二种是动态路由协议生成（如 OSPF、IS-IS 和 RIP），由路由能力比较强的路由器将缺省路由发布给其它路由器，其它路由器在自己的路由表里生成指向那台路由器的缺省路由。配置请参见各个路由协议手册。

# 目 录

# 1 RIP

## 1.1 RIP简介

RIP（Routing Information Protocol，路由信息协议）是一种较为简单的内部网关协议（Interior Gateway Protocol，IGP），主要用于规模较小的网络中，比如校园网以及结构较简单的地区性网络。对于更为复杂的环境和大型网络，一般不使用 RIP。

由于 RIP 的实现较为简单，在配置和维护管理方面也远比 OSPF 和 IS-IS 容易，因此在实际组网中仍有广泛的应用。

### 1.1.1 RIP的工作机制

#### 1. RIP的基本概念

RIP 是一种基于距离矢量（Distance-Vector）算法的协议，它通过 UDP 报文进行路由信息的交换，使用的端口号为 520。

RIP 使用跳数来衡量到达目的地址的距离，跳数称为度量值。在 RIP 中，路由器到与它直接相连网络的跳数为 0，通过一个路由器可达的网络的跳数为 1，其余依此类推。为限制收敛时间，RIP 规定度量值取 0～15 之间的整数，大于或等于 16 的跳数被定义为无穷大，即目的网络或主机不可达。由于这个限制，使得 RIP 不适合应用于大型网络。

为提高性能，防止产生路由环路，RIP 支持水平分割（Split Horizon）和毒性逆转（Poison Reverse）功能。

#### 2. RIP的路由数据库

每个运行 RIP 的路由器管理一个路由数据库，该路由数据库包含了到所有可达目的地的路由项，这些路由项包含下列信息：

- 目的地址：主机或网络的地址。
- 下一跳地址：为到达目的地，需要经过的相邻路由器的接口 IP 地址。
- 出接口：本路由器转发报文的出接口。
- 度量值：本路由器到达目的地的开销。
- 路由时间：从路由项最后一次被更新到现在所经过的时间，路由项每次被更新时，路由时间重置为 0。
- 路由标记（Route Tag）：用于标识外部路由，在路由策略中可根据路由标记对路由信息进行灵活的控制。关于路由策略的详细信息，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

#### 3. RIP防止路由环路的机制

RIP 协议向邻居通告的是自己的路由表，有可能会发生路由环路，可以通过以下机制来避免：

- 计数到无穷（Counting to infinity）：将度量值等于 16 的路由定义为不可达（infinity）。在路由环路发生时，某条路由的度量值将会增加到 16，该路由被认为不可达。

- 触发更新（Triggered Updates）：RIP 通过触发更新来避免在多个路由器之间形成路由环路的可能，而且可以加速网络的收敛速度。一旦某条路由的度量值发生了变化，就立刻向邻居路由器发布更新报文，而不是等到更新周期的到来。
- 水平分割（Split Horizon）：RIP 从某个接口学到的路由，不会从该接口再发回给邻居路由器。这样不但减少了带宽消耗，还可以防止路由环路。
- 毒性逆转（Poison Reverse）：RIP 从某个接口学到路由后，将该路由的度量值设置为 16（不可达），并从原接口发回邻居路由器。利用这种方式，可以清除对方路由表中的无用信息。

#### 4. RIP的运行过程

RIP 的运行过程如下：

(1) 路由器启动 RIP 后，便会向相邻的路由器发送请求报文（Request message），相邻的 RIP 路由器收到请求报文后，响应该请求，回送包含本地路由表信息的响应报文（Response message）。

(2) 路由器收到响应报文后，更新本地路由表，同时向相邻路由器发送触发更新报文，通告路由更新信息。相邻路由器收到触发更新报文后，又向其各自的相邻路由器发送触发更新报文。在一连串的触发更新广播后，各路由器都能得到并保持最新的路由信息。

(3) 路由器周期性向相邻路由器发送本地路由表，运行 RIP 协议的相邻路由器在收到报文后，对本地路由进行维护，选择一条最佳路由，再向其各自相邻网络发送更新信息，使更新的路由最终能达到全局有效。同时，RIP 采用老化机制对超时的路由进行老化处理，以保证路由的实时性和有效性。

### 1.1.2 RIP的版本

RIP 有两个版本：RIP-1 和 RIP-2。

RIP-1 是有类别路由协议（Classful Routing Protocol），它只支持以广播方式发布协议报文。RIP-1 的协议报文无法携带掩码信息，它只能识别 A、B、C 类这样的自然网段的路由，因此 RIP-1 不支持不连续子网（Discontiguous Subnet）。

RIP-2 是一种无类别路由协议（Classless Routing Protocol），与 RIP-1 相比，它有以下优势：

- 支持路由标记，在路由策略中可根据路由标记对路由进行灵活的控制。
- 报文中携带掩码信息，支持路由聚合和 CIDR（Classless Inter-Domain Routing，无类域间路由）。
- 支持指定下一跳，在广播网上可以选择到最优下一跳地址。
- 支持组播路由发送更新报文，只有 RIP-2 路由器才能收到更新报文，减少资源消耗。
- 支持对协议报文进行验证，并提供明文验证和 MD5 验证两种方式，增强安全性。

RIP-2 有两种报文传送方式：广播方式和组播方式，缺省将采用组播方式发送报文，使用的组播地址为 224.0.0.9。当接口运行 RIP-2 广播方式时，也可接收 RIP-1 的报文。

### 1.1.3 协议规范

与 RIP 相关的协议规范有：

- RFC 1058：Routing Information Protocol
- RFC 1723：RIP Version 2 - Carrying Additional Information

- RFC 1721：RIP Version 2 Protocol Analysis
- RFC 1722：RIP Version 2 Protocol Applicability Statement
- RFC 1724：RIP Version 2 MIB Extension
- RFC 2082：RIP-2 MD5 Authentication
- RFC 2091：Triggered Extensions to RIP to Support Demand Circuits
- RFC 2453：RIP Version 2

## 1.2 RIP配置任务简介

表1-1 RIP 配置任务简介

| 配置任务 | | 说明 | 详细配置 |
|---|---|---|---|
| 配置RIP的基本功能 | 启动RIP | 必选 | 1.3.2 |
| | 配置接口的工作状态 | 可选 | 1.3.3 |
| | 配置RIP版本 | 可选 | 1.3.4 |
| 配置RIP的路由信息控制 | 配置接口附加度量值 | 可选 | 1.4.2 |
| | 配置RIP-2路由聚合 | 可选 | 1.4.3 |
| | 禁止RIP接收主机路由 | 可选 | 1.4.4 |
| | 配置RIP发布缺省路由 | 可选 | 1.4.5 |
| | 配置RIP对接收/发布的路由进行过滤 | 可选 | 1.4.4 |
| | 配置RIP协议优先级 | 可选 | 1.4.7 |
| | 配置RIP引入外部路由 | 可选 | 1.4.8 |
| 调整和优化RIP网络 | 配置RIP定时器 | 可选 | 1.5.2 |
| | 配置水平分割和毒性逆转 | 可选 | 1.5.3 |
| | 配置最大等价路由条数 | 可选 | 1.5.4 |
| | 配置RIP-1报文的零域检查 | 可选 | 1.5.5 |
| | 配置源地址检查 | 可选 | 1.5.6 |
| | 配置RIP-2报文的认证方式 | 可选 | 1.5.7 |
| | 配置RIP邻居 | 可选 | 1.5.8 |
| | 配置RIP和MIB绑定 | 可选 | 1.5.9 |
| | 配置RIP报文的发送速率 | 可选 | 1.5.10 |
| | 配置RIP报文的最大长度 | 可选 | 1.5.11 |
| 配置RIP GR | | 可选 | 1.6 |
| 配置RIP与BFD联动 | | 可选 | 1.7 |
| 配置RIP快速重路由功能 | | 可选 | 1.8 |

## 1.3 配置RIP的基本功能

### 1.3.1 配置准备

在配置 RIP 的基本功能之前，需完成以下任务：

- 配置链路层协议
- 配置接口的网络层地址，使相邻节点的网络层可达

### 1.3.2 启动RIP

提示

- 如果在启动 RIP 前在接口视图下配置了 RIP 相关命令，这些配置只有在 RIP 启动后才会生效。
- RIP 不支持将同一物理接口下的不同网段使能到不同的 RIP 进程中。
- RIP 不支持在同一物理接口下使能多个 RIP 进程。

目前，系统支持 RIP 多进程。当在一台路由器上启动多个 RIP 进程时，需要指定不同的进程号。RIP 进程号是本地概念，不影响与其它路由器之间的报文交换。因此，不同的路由器之间，即使进程号不同也可以进行报文交换。

#### 1. 在指定网段上使能RIP

RIP 只在指定网段的接口上运行，指定网段的同时可以配置反码；对于不在指定网段上的接口，RIP 既不在它上面接收和发送路由，也不将它的接口路由转发出去。因此，RIP 启动后必须指定其工作网段。

表1-2 在指定网段上使能 RIP

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 启动RIP，并进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 缺省情况下，系统没有启动RIP |
| 在指定网段上使能RIP | **network** *network-address* [ *wildcard-mask* ] | 缺省情况下，没有网段使能RIP<br>在单进程情况下，可以使用**network** 0.0.0.0命令用来在所有接口上使能RIP；在多进程情况下，无法使用**network** 0.0.0.0命令 |

#### 2. 在指定接口上使能RIP

RIP 支持在接口下使能 RIP 进程。

表1-3 在指定接口上使能 RIP

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 启动RIP，并进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 缺省情况下，系统没有启动RIP |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 在指定接口上使能RIP | **rip** *process-id* **enable** [ **exclude-subip** ] | 缺省情况下，接口没有使能RIP |

### 1.3.3 配置接口的工作状态

用户可对接口的工作状态进行配置：

- 配置接口工作在抑制状态，即接口只接收 RIP 报文而不发送 RIP 报文
- 配置允许接口接收 RIP 报文
- 配置允许接口发送 RIP 报文

**表1-4 配置 RIP 收发报文控制**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置抑制接口 | **silent-interface** { *interface-type interface-number* \| **all** } | 缺省情况下，允许所有接口发送路由更新报文<br>若抑制接口收到非知名端口的单播请求，需要发送响应报文 |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置允许接口接收RIP报文 | **rip input** | 缺省情况下，允许接口接收RIP报文 |
| 配置允许接口发送RIP报文 | **rip output** | 缺省情况下，允许接口发送RIP报文 |

### 1.3.4 配置RIP版本

用户可以在 RIP 视图下配置 RIP 版本，也可在接口上配置 RIP 版本：

- 当全局和接口都没有进行 RIP 版本配置时，接口发送 RIP-1 广播报文，可以接收 RIP-1 广播/单播报文、RIP-2 广播/组播/单播报文。
- 如果接口上配置了 RIP 版本，以接口配置的为准；如果接口没有进行 RIP 版本配置，接口运行的 RIP 版本将以全局配置的版本为准。

表1-5 配置 RIP 版本号

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置全局RIP版本 | **version** { **1** \| **2** } | 缺省情况下，没有配置全局RIP版本。接口只能发送RIP-1广播报文，可以接收RIP-1广播/单播报文、RIP-2广播/组播/单播报文 |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口运行的RIP版本 | **rip version** { **1** \| **2** [ **broadcast** \| **multicast** ] } | 缺省情况下，没有配置接口运行的RIP版本。接口只能发送RIP-1广播报文，可以接收RIP-1广播/单播报文、RIP-2广播/组播/单播报文 |

## 1.4 配置RIP的路由信息控制

### 1.4.1 配置准备

在实际应用中，有时候需要对 RIP 路由信息进行更为精确的控制以满足复杂网络环境的需要。

在配置之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点网络层可达
- 配置 RIP 的基本功能

### 1.4.2 配置接口附加度量值

附加度量值是在 RIP 路由原来度量值的基础上所增加的度量值（跳数），包括发送附加度量值和接收附加度量值。

- 发送附加度量值：不会改变路由表中的路由度量值，仅当接口发送 RIP 路由信息时才会添加到发送路由上。
- 接收附加度量值：会影响接收到的路由度量值，接口接收到一条合法的 RIP 路由时，在将其加入路由表前会把度量值附加到该路由上，当附加度量值与原路由度量值之和大于 16 时，该条路由的度量值取 16。

表1-6 配置接口附加度量值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口接收RIP路由时的附加度量值 | **rip metricin** [ **route-policy** *route-policy-name* ] *value* | 缺省情况下，接口接收RIP路由时的附加路由度量值为0 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置接口发送RIP路由时的附加度量值 | **rip metricout** [ **route-policy** *route-policy-name* ] *value* | 缺省情况下，接口发送RIP路由时的附加路由度量值为1 |

### 1.4.3 配置RIP-2 路由聚合

路由聚合是指路由器把同一自然网段内的连续子网的路由聚合成一条路由向外发送，如路由表里有10.1.1.0/24、10.1.2.0/24、10.1.3.0/24 三条路由，可以通过配置把它们聚合成一条路由 10.1.0.0/16 向外发送，这样邻居路由器只接收到一条路由 10.1.0.0/16，从而减少了路由表的规模，以及网络上的传输流量。

通过配置路由聚合，可以提高网络的可扩展性以及路由器的处理速度。

RIP-2 将多条路由聚合成一条路由时，聚合路由的 Metric 值将取所有路由 Metric 的最小值。

在 RIP-2 中，有两种路由聚合方式：自动路由聚合和手工配置聚合路由。

#### 1. 自动路由聚合

自动路由聚合是指 RIP-2 将同一自然网段内的不同子网的路由聚合成一条自然掩码的路由向外发送，例如，假设路由表里有 10.1.1.0/24、10.1.2.0/24、10.1.3.0/24 三条路由，使能 RIP-2 自动路由聚合功能后，这三条路由聚合成一条自然掩码的路由 10.0.0.0/8 向外发送。

**表1-7 配置自动路由聚合**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能RIP-2自动路由聚合功能 | **summary** | 缺省情况下，RIP-2自动路由聚合功能处于使能状态<br>如果路由表里的路由子网不连续，则需要取消自动路由聚合功能，使得RIP-2能够向外发布子网路由和主机路由 |

#### 2. 手工配置聚合路由

用户可在指定接口配置 RIP-2 发布一条聚合路由。

聚合路由的目的地址和掩码进行与运算到一个网络地址，RIP-2 将对落入该网段内的路由进行聚合，接口只发布聚合后的路由。

例如，假设路由表里有 10.1.1.0/24、10.1.2.0/24、10.1.3.0/24 三条子网连续的路由，在接口 GigabitEthernet2/1/1 配置发布一条聚合路由 10.1.0.0/16 后，这三条路由聚合成一条路由 10.1.0.0/16 向外发送。

缺省情况下，RIP-2 的路由将按照自然掩码自动聚合，如果用户在指定接口配置发布一条聚合路由，则必须先关闭自动聚合功能。

表1-8 手工配置聚合路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 关闭RIP-2自动路由聚合功能 | **undo summary** | 缺省情况下，RIP-2自动路由聚合功能处于使能状态 |
| 退至系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置发布一条聚合路由 | **rip summary-address** *ip-address* { *mask-length* \| *mask* } | 缺省情况下，没有配置发布一条聚合路由 |

## 1.4.4 禁止RIP接收主机路由

提示

禁止接收主机路由仅对 RIPv2 报文携带的路由有效，对 RIPv1 报文携带的路由无效。

在某些特殊情况下，路由器会收到大量来自同一网段的主机路由。这些路由对于路由寻址没有多少作用，却占用了大量的资源，此时可配置 RIP 禁止接收主机路由，以节省网络资源。

表1-9 禁止 RIP 接收主机路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 禁止RIP接收主机路由 | **undo host-route** | 缺省情况下，允许RIP接收主机路由 |

## 1.4.5 配置RIP发布缺省路由

用户可以配置 RIP 以指定度量值向邻居发布一条缺省路由。

- 用户可以在 RIP 视图下配置 RIP 进程的所有接口向邻居发布缺省路由，也可以在接口下配置指定 RIP 接口向邻居发布缺省路由。
- 如果接口没有进行发布缺省路由的相关配置，则以 RIP 进程下的配置为准，否则将以接口配置为准。
- 如果 RIP 进程配置了发布缺省路由，但希望该进程下的某个接口不发送缺省路由（只发布普通路由），可以通过在接口下配置 **rip default-route no-originate** 命令实现。

表1-10 配置 RIP 发布缺省路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIP发布缺省路由 | **default-route** { **only** \| **originate** } [ **cost** *cost* ] | 缺省情况下，RIP不向邻居发送缺省路由 |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置RIP接口发布缺省路由 | **rip default-route** { { **only** \| **originate** } [ **cost** *cost* ] \| **no-originate** } | 缺省情况下，RIP接口是否发布缺省路由以RIP进程配置的为准 |


说明

配置发布缺省路由的 RIP 路由器不接收来自 RIP 邻居的缺省路由。

### 1.4.6 配置RIP对接收/发布的路由进行过滤

路由器提供路由信息过滤功能，通过指定地址前缀列表，可以配置入口或出口过滤策略，对接收和发布的路由进行过滤。在接收路由时，可以指定只接收来自某个邻居的 RIP 报文。

表1-11 配置 RIP 对接收/发布的路由进行过滤

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 对接收的路由信息进行过滤 | **filter-policy** { *acl-number* \| **gateway** *prefix-list-name* \| **prefix-list** *prefix-list-name* [ **gateway** *prefix-list-name* ] } **import** [ *interface-type interface-number* ] | 缺省情况下，RIP不对接收的路由信息进行过滤<br>本命令对从邻居收到的RIP路由进行过滤，没有通过过滤的路由将不被加入路由表，也不向邻居发布该路由 |
| 对发布的路由信息进行过滤 | **filter-policy** { *acl-number* \| **prefix-list** *prefix-list-name* } **export** [ *protocol* [ *process-id* ] \| *interface-type interface-number* ] | 缺省情况下，RIP不对发布的路由信息进行过滤<br>本命令对本机所有路由的发布进行过滤，包括使用**import-route**引入的路由和从邻居学到的RIP路由 |

### 1.4.7 配置RIP协议优先级

在路由器中可能会运行多个 IGP 路由协议，如果想让 RIP 路由具有比从其它路由协议学来的路由更高的优先级，需要配置小的优先级值。优先级的高低将最后决定 IP 路由表中的路由是通过哪种路由算法获取的最佳路由。

表1-12 配置 RIP 协议优先级

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIP路由的优先级 | **preference** [ **route-policy** *route-policy-name* ] *value* | 缺省情况下，RIP路由的优先级为100 |

### 1.4.8 配置RIP引入外部路由

如果在路由器上不仅运行 RIP，还运行着其它路由协议，可以配置 RIP 引入其它协议生成的路由，如 OSPF、IS-IS、BGP、静态路由或者直连路由。

表1-13 配置 RIP 引入外部路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 引入外部路由 | **import-route** *protocol* [ *process-id* \| **all-processes** \| **allow-ibgp** ] [ **allow-direct** \| **cost** *cost* \| **route-policy** *route-policy-name* \| **tag** *tag* ] * | 缺省情况下，RIP不引入其它路由<br>只能引入路由表中状态为active的路由，是否为active状态可以通过**display ip routing-table protocol**命令来查看 |
| （可选）配置引入路由的缺省度量值 | **default cost** *value* | 缺省情况下，引入路由的缺省度量值为0 |

## 1.5 调整和优化RIP网络

### 1.5.1 配置准备

在某些特殊的网络环境中，需要对 RIP 网络的性能进行调整和优化，在调整和优化 RIP 网络之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点网络层可达
- 配置 RIP 的基本功能

### 1.5.2 配置RIP定时器

提示

定时器值的调整应考虑网络的性能，并在所有运行 RIP 的路由器上进行统一配置，以免增加不必要的网络流量或引起网络路由震荡。

通过调整 RIP 定时器可以改变 RIP 网络的收敛速度。

RIP 受四个定时器的控制，分别是 Update、Timeout、Suppress 和 Garbage-Collect。

- Update 定时器，定义了发送路由更新的时间间隔。
- Timeout 定时器，定义了路由老化时间。如果在老化时间内没有收到关于某条路由的更新报文，则该条路由在路由表中的度量值将会被设置为 16。
- Suppress 定时器，定义了 RIP 路由处于抑制状态的时长。当一条路由的度量值变为 16 时，该路由将进入抑制状态。在被抑制状态，只有来自同一邻居且度量值小于 16 的路由更新才会被路由器接收，取代不可达路由。
- Garbage-Collect 定时器，定义了一条路由从度量值变为 16 开始，直到它从路由表里被删除所经过的时间。在 Garbage-Collect 时间内，RIP 以 16 作为度量值向外发送这条路由的更新，如果 Garbage-Collect 超时，该路由仍没有得到更新，则该路由将从路由表中被彻底删除。

表1-14 配置 RIP 定时器

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIP定时器的值 | **timers** { **garbage-collect** *garbage-collect-value* \| **suppress** *suppress-value* \| **timeout** *timeout-value* \| **update** *update-value* } * | 缺省情况下，Garbage-collect定时器的值为120秒，Suppress定时器的值为120秒，Timeout定时器的值为180秒，Update定时器的值为30秒 |

### 1.5.3 配置水平分割和毒性逆转

提示

如果同时配置了水平分割和毒性逆转，则只有毒性逆转功能生效。

通过配置水平分割或毒性逆转功能可以防止路由环路。

#### 1. 配置水平分割

配置水平分割可以使得从一个接口学到的路由不能通过此接口向外发布，用于避免相邻路由器间的路由环路。

表1-15 配置水平分割

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能水平分割功能 | **rip split-horizon** | 缺省情况下，水平分割功能处于使能状态 |

### 2. 配置毒性逆转

配置毒性逆转后，从一个接口学到的路由还可以从这个接口向外发布，但这些路由的度量值会设置为 16（即不可达），可以用于避免相邻路由器间的路由环路。

表1-16 配置毒性逆转

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能毒性逆转功能 | **rip poison-reverse** | 缺省情况下，毒性逆转功能处于关闭状态 |

## 1.5.4 配置最大等价路由条数

通过配置最大等价路由条数，可以使用多条等价路由对 RIP 网络进行负载分担。

表1-17 配置最大等价路由条数

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入 RIP 视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIP最大等价路由条数 | **maximum load-balancing** *number* | 缺省情况下，RIP支持的最大等价路由条数为32。 |

## 1.5.5 配置RIP-1 报文的零域检查

RIP-1 报文中的有些字段必须为零，称之为零域。用户可配置 RIP-1 在接收报文时对零域进行检查，零域值不为零的 RIP-1 报文将不被处理。如果用户能确保所有报文都是可信任的，则可以不进行该项检查，以节省 CPU 处理时间。

由于 RIP-2 的报文没有零域，此项配置对 RIP-2 无效。

表1-18 配置 RIP-1 报文的零域检查

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入 RIP 视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能RIP-1报文的零域检查功能 | **checkzero** | 缺省情况下，RIP-1报文的零域检查功能处于使能状态 |

### 1.5.6 配置源地址检查

通过配置对接收到的 RIP 路由更新报文进行源 IP 地址检查：

- 对于在接口上接收的报文，RIP 将检查该报文源地址和接收接口的 IP 地址是否处于同一网段，如果不在同一网段则丢弃该报文。
- 对于串口上接收的报文，RIP 检查该报文的源地址是否是对端接口的 IP 地址，如果不是则丢弃该报文。

**表1-19 配置源地址检查**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入 RIP 视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能对接收到的RIP路由更新报文进行源IP地址检查功能 | **validate-source-address** | 缺省情况下，对接收到的RIP路由更新报文进行源IP地址检查功能处于使能状态 |

### 1.5.7 配置RIP-2 报文的认证方式

在安全性要求较高的网络环境中，可以通过配置报文的认证方式来对 RIP-2 报文进行有效性检查和验证。

RIP-2 支持两种认证方式：简单认证和 MD5 认证。

**表1-20 配置 RIP-2 报文的认证方式**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置RIP-2报文的认证方式 | **rip authentication-mode** { **md5** { **rfc2082** { **cipher** *cipher-string* \| **plain** *plain-string* } *key-id* \| **rfc2453** { **cipher** *cipher-string* \| **plain** *plain-string* } } \| **simple** { **cipher** *cipher-string* \| **plain** *plain-string* } } | 缺省情况下，接口没有配置RIP-2的认证方式<br>当RIP的版本为RIP-1时，虽然在接口视图下仍然可以配置验证方式，但由于RIP-1不支持认证，因此该配置不会生效 |

### 1.5.8 配置RIP邻居

通常情况下，RIP 使用广播或组播地址发送报文，如果在不支持广播或组播报文的链路上运行 RIP，则必须手工指定 RIP 的邻居。

表1-21 配置 RIP 邻居

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIP邻居 | **peer** *ip-address* | 缺省情况下，RIP不向任何定点地址发送单播更新报文<br>当RIP邻居与当前设备直连时不推荐使用**peer** *ip-address*命令，因为这样可能会造成对端同时收到同一路由信息的组播（或广播）和单播两种形式的报文 |
| 取消对接收到的RIP路由更新报文进行源IP地址检查操作 | **undo validate-source-address** | 缺省情况下，对接收到的RIP路由更新报文进行源IP地址检查<br>当指定的邻居和本地路由器非直接连接，则必须取消对更新报文的源地址进行检查 |

### 1.5.9 配置RIP网管功能

配置 RIP 进程绑定 MIB 功能后，可以通过网管软件对指定的 RIP 进程进行管理。

表1-22 配置 RIP 网管功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置RIP进程绑定MIB | **rip mib-binding** *process-id* | 缺省情况下，MIB绑定在进程号最小的RIP进程上 |

### 1.5.10 配置RIP报文的发送速率

RIP 周期性地将路由信息放在 RIP 报文中向邻居发送。

如果路由表里的路由条目数量很多，同时发送大量 RIP 协议报文有可能会对当前设备和网络带宽带来冲击；因此，路由器将 RIP 协议报文分为多个批次进行发送，并且对 RIP 接口每次允许发送的 RIP 协议报文最大个数做出限制。

用户可根据需要配置接口发送 RIP 报文的时间间隔以及接口一次发送 RIP 报文的最大个数。

表1-23 配置 RIP 报文的发送速率

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入 RIP 视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置 RIP 报文的发送速率 | **output-delay** *time* **count** *count* | 缺省情况下，接口发送 RIP 报文的时间间隔为 20 毫秒，一次最多发送 3 个 RIP 报文 |

### 1.5.11 配置RIP报文的最大长度

提示

由于不同厂商对 RIP 报文最大长度的支持情况不同，要谨慎使用本特性，以免出现不兼容的情况。

RIP 周期性地将路由信息放在 RIP 报文中向邻居发送，根据 RIP 报文的最大长度来计算报文中发送的最大路由数。通过设置 RIP 报文的最大长度，可以合理利用链路带宽。

在配置认证的情况下，如果配置不当可能会造成报文无法发送，建议用户按照下面进行配置：

- 简单验证方式时，RIP 报文的最大长度不小于 52 字节；
- MD5 验证方式（使用 RFC 2453 规定的报文格式）时，RIP 报文的最大长度不小于 56 字节；
- MD5 验证方式（使用 RFC 2082 规定的报文格式）时，RIP 报文的最大长度不小于 72 字节。

表1-24 配置 RIP 报文的最大长度

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置 RIP 报文的最大长度 | **rip max-packet-length** *value* | 缺省情况下，接口发送 RIP 报文的最大长度为 512 字节 |

## 1.6 配置RIP GR

GR（Graceful Restart，平滑重启）是一种在协议重启或主备倒换时 RIP 进行平滑重启，保证转发业务不中断的机制。

GR 有两个角色：

- GR Restarter：发生协议重启或主备倒换事件且具有 GR 能力的设备。
- GR Helper：和 GR Restarter 具有邻居关系，协助完成 GR 流程的设备。

在普通的路由协议重启的情况下，路由器需要重新学习 RIP 路由，并更新 FIB 表，此时会引起网络暂时的中断，基于 RIP 的 GR 可以解决这个问题。

应用了 GR 特性的设备向外发送 RIP 全部路由表请求报文，重新从邻居处学习 RIP 路由，在此期间 FIB 表不变化。在路由协议重启完毕后，设备将重新学到的 RIP 路由下刷给 FIB 表，使该设备的路由信息恢复到重启前的状态。

在作为 GR Restarter 的设备上进行以下配置。启动了 RIP 的设备缺省就是 GR Helper。

表1-25 配置 RIP GR

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能 RIP 协议的 GR 能力 | **graceful-restart** | 缺省情况下，RIP 协议的 GR 能力处于关闭状态 |

## 1.7 配置RIP与BFD联动

RIP 协议依赖周期性发送路由更新请求作为检测机制，当在指定时间内没有收到路由更新回应时，认为此条路由不再生效，这种方式不能快速响应链路故障。使用 BFD（Bidirectional Forwarding Detection，双向转发检测）检测到链路故障时，RIP 能快速撤销失效路由，减少对其他业务的影响。关于 BFD 的介绍和基本功能配置，请参见“可靠性配置指导”中的“BFD”。

目前 RIP 支持 BFD 提供了下面几种检测方式：

- echo 报文单跳检测方式：直连邻居使用。在对端有 RIP 路由发送时才能建立 BFD 会话。
- 指定目的地址的 echo 报文单跳检测方式：直连邻居使用，并且在接口上直接指定 RIP 邻居的 IP 地址。当该接口使能了 RIP 功能，会建立到指定目的 IP 地址的 BFD 会话。
- control 报文双向检测方式：非直连邻居使用。当两端互有 RIP 路由发送时，且使能 BFD 的接口与接收接口为同一接口，邻居之间才能建立 BFD 会话。

### 1.7.1 echo报文单跳检测

表1-26 配置 RIP 与 BFD 联动（echo 报文单跳检测）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置echo报文源地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置echo报文源地址 |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能RIP的BFD功能 | **rip bfd enable** | 缺省情况下，RIP的BFD功能处于关闭状态 |

### 1.7.2 指定目的地址的echo报文单跳检测

提示

本特性只检测本端到 RIP 直连邻居的链路的连通状况。配置本特性时，指定的目的地址只能是 RIP 直连邻居的 IP 地址。

在链路出现单通故障时，为了加快路由收敛速度，可以在本端设备上配置本特性对链路进行检测。链路出现故障时，本端设备不再从该接口收发任何 RIP 报文；链路恢复后，接口将继续发送 RIP 报文。

表1-27 配置 RIP 与 BFD 联动（指定目的地址的 echo 报文单跳检测）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置echo报文源地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置echo报文源地址 |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能RIP指定目的地址的BFD功能 | **rip bfd enable destination** *ip-address* | 缺省情况下，RIP的BFD功能处于关闭状态 |

### 1.7.3 control报文双向检测

表1-28 配置 RIP 与 BFD 联动（control 报文双向检测）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIP邻居 | **peer** *ip-address* | 缺省情况下，RIP不向任何定点地址发送更新报文<br>由于**peer**命令与邻居之间没有对应关系，**undo peer**操作并不能立刻删除邻居，因此不能立刻删除BFD会话 |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能RIP的BFD功能 | **rip bfd enable** | 缺省情况下，RIP的BFD功能处于关闭状态 |

# 1.8 配置RIP快速重路由功能

提示

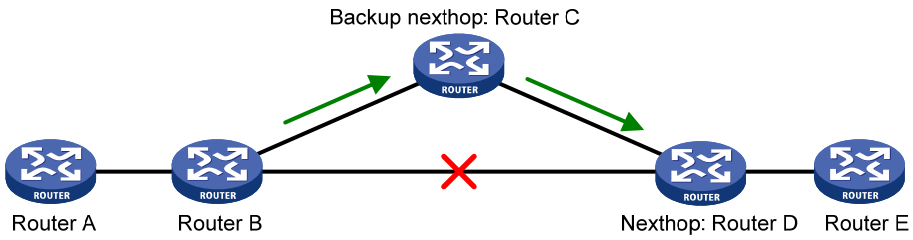
- RIP 快速重路由功能仅对非迭代 RIP 路由（即从直连邻居学到 RIP 路由）有效。
- RIP 快速重路由功能不能与 RIP 的 BFD 功能同时使用，否则可能导致快速重路由功能失效。

## 1.8.1 功能简介

当 RIP 网络中的链路或某台路由器发生故障时，数据流量将会被中断，直到 RIP 根据新的拓扑网络路由收敛完毕后，被中断的流量才能恢复正常的传输。

为了尽可能缩短网络故障导致的流量中断时间，网络管理员可以根据需要配置 RIP 快速重路由功能。

图1-1 RIP 快速重路由功能示意图




如 图 1-1 所示，通过在Router B上配置快速重路由功能，RIP可以为路由指定备份下一跳，当Router B检测到网络故障时，RIP会使用事先获取好的备份下一跳替换失效下一跳，通过备份下一跳来指导报文的转发，从而大大缩短了流量中断时间。在使用备份下一跳指导报文转发的同时，RIP会根据变化后的网络拓扑重新计算路由，网络收敛完毕后，使用新计算出来的最优路由来指导报文转发。

## 1.8.2 配置限制和指导

本功能只适合在主链路三层接口 up，主链路由双通变为单通或者不通的情况下使用。在主链路三层接口 down 的情况下，本功能不可用。

单通现象，即一条链路上的两端，有且只有一端可以收到另一端发来的报文，此链路称为单向链路。

## 1.8.3 配置准备

要配置快速重路由功能，网络管理员需要配置路由策略，通过 **apply fast-reroute backup-interface** 命令在路由策略中指定备份下一跳；关于 **apply fast-reroute backup-interface** 命令以及路由策略的相关配置，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

### 1.8.4 配置步骤

#### 1. 配置RIP快速重路由功能

表1-29 配置 RIP 快速重路由功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIP视图 | **rip** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIP快速重路由功能 | **fast-reroute route-policy** *route-policy-name* | 缺省情况下，RIP快速重路由功能处于关闭状态 |

#### 2. 配置RIP快速重路由支持BFD检测功能

RIP 协议的快速重路由特性中，主用链路缺省不使用 BFD 进行链路故障检测。配置本功能后，将使用 BFD（Echo 方式）进行检测，可以加快 RIP 协议的收敛速度。

表1-30 配置 RIP 快速重路由支持 BFD 检测功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置BFD Echo报文源地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置BFD Echo报文源地址 |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能RIP协议中主用链路的BFD（Echo方式）检测功能 | **rip primary-path-detect bfd echo** | 缺省情况下，RIP协议中主用链路的BFD（Echo方式）检测功能处于关闭状态 |

## 1.9 RIP显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 RIP 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以重启 RIP 进程或清除指定 RIP 进程的统计信息。

表1-31 RIP 显示和维护

| 操作 | 命令 |
| --- | --- |
| 显示RIP的当前运行状态及配置信息 | **display rip** [ *process-id* ] |
| 显示RIP数据库的激活路由 | **display rip** *process-id* **database** [ *ip-address* { *mask-length* \| *mask* } ] |
| 显示RIP的接口信息 | **display rip** *process-id* **interface** [ *interface-type interface-number* ] |
| 显示RIP的路由信息 | **display rip** *process-id* **route** [ *ip-address* { *mask-length* \| *mask* } [ **verbose** ] \| **peer** *ip-address* \| **statistics** ] |
| 重启指定RIP进程 | **reset rip** *process-id* **process** |

| 操作 | 命令 |
| --- | --- |
| 清除RIP进程的统计信息 | **reset rip** *process-id* **statistics** |

# 1.10 RIP典型配置举例

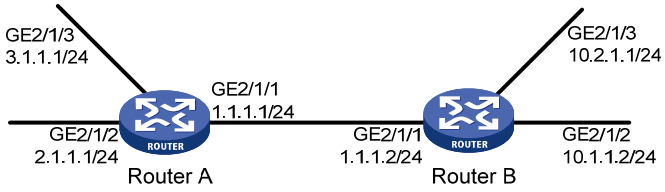
## 1.10.1 配置RIP基本功能

### 1. 组网需求

- 在 Router A 和 Router B 的所有接口上使能 RIP，并使用 RIP-2 进行网络互连。
- 在 Router B 上配置路由出策略，向 Router A 发布的路由中过滤掉 10.2.1.0/24；Router B 上配置入策略，使得 Router B 只接收路由 2.1.1.0/24。

### 2. 组网图

图1-2 RIP 基本功能配置组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 使能 RIP 功能

使能 RIP 功能在下面采用了两种不同配置方式，请根据实际情况进行选择。

# 配置 Router A，在指定网段上使能 RIP。

```
<RouterA> system-view
[RouterA] rip
[RouterA-rip-1] network 1.0.0.0
[RouterA-rip-1] network 2.0.0.0
[RouterA-rip-1] network 3.0.0.0
[RouterA-rip-1] quit
```

# 配置 Router B，在指定接口上使能 RIP。

```
<RouterB> system-view
[RouterB] rip
[RouterB-rip-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouteB-GigabitEthernet2/1/1] rip 1 enable
[RouterB-rip-1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouteB-GigabitEthernet2/1/2] rip 1 enable
[RouterB-rip-1] quit
[RouterB] interface gigabitethernet 2/1/3
```

```
[RouteB-GigabitEthernet2/1/3] rip 1 enable
[RouterB-rip-1] quit
```

# 查看 Router A 的 RIP 路由表。

```
[RouterA] display rip 1 route
 Route Flags: R - RIP
              A - Aging, S - Suppressed, G - Garbage-collect, D – Direct
              O - Optimal, F - Flush to RIB
 ------------------------------------------------------------------------------
 Peer 1.1.1.2 on GigabitEthernet2/1/1
      Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
      10.0.0.0/8              1.1.1.2          1       0       RAOF    9
 Local route
      Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
      1.1.1.0/24              0.0.0.0          0       0       RDOF    -
      2.1.1.0/24              0.0.0.0          0       0       RDOF    -
      3.1.1.0/24              0.0.0.0          0       0       RDOF    -
```

从路由表中可以看出，RIP-1 发布的路由信息使用的是自然掩码。

(3) 配置 RIP 的版本

# 在 Router A 上配置 RIP-2。

```
[RouterA] rip
[RouterA-rip-1] version 2
[RouterA-rip-1] undo summary
[RouterA-rip-1] quit
```

# 在 Router B 上配置 RIP-2。

```
[RouterB] rip
[RouterB-rip-1] version 2
[RouterB-rip-1] undo summary
[RouterB-rip-1] quit
```

# 查看 Router A 的 RIP 路由表。

```
[RouterA] display rip 1 route
 Route Flags: R - RIP
              A - Aging, S - Suppressed, G - Garbage-collect, D – Direct
              O - Optimal, F - Flush to RIB
 ------------------------------------------------------------------------------
 Peer 1.1.1.2  on GigabitEthernet2/1/1
      Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
      10.0.0.0/8              1.1.1.2          1       0       RAOF    87
      10.1.1.0/24             1.1.1.2          1       0       RAOF    19
      10.2.1.0/24             1.1.1.2          1       0       RAOF    19
 Local route
      Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
      1.1.1.0/24              0.0.0.0          0       0       RDOF    -
      2.1.1.0/24              0.0.0.0          0       0       RDOF    -
      3.1.1.0/24              0.0.0.0          0       0       RDOF    -
```

从路由表中可以看出，RIP-2 发布的路由中带有更为精确的子网掩码信息。

说明

由于 RIP 路由信息的老化时间较长，所以在配置 RIP-2 版本后的一段时间里，路由表中还会存在 RIP-1 的路由信息。

# 查看 Router B 的路由表信息。

```
[RouterB] display rip 1 route
 Route Flags: R - RIP
              A - Aging, S - Suppressed, G - Garbage-collect, D – Direct
              O - Optimal, F - Flush to RIB
 ------------------------------------------------------------------------------
 Peer 1.1.1.1 on GigabitEthernet2/1/1
      Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
      2.1.1.0/24              1.1.1.1          1       0       RAOF    19
      3.1.1.0/24              1.1.1.1          1       0       RAOF    19
 Local route
      Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
      1.1.1.0/24              0.0.0.0          0       0       RDOF    -
      10.1.1.0/24             0.0.0.0          0       0       RDOF    -
      10.2.1.0/24             0.0.0.0          0       0       RDOF    -
```

(4) 配置 RIP 路由过滤

# 在 Router B 配置地址前缀列表。

```
[RouterB] ip prefix-list aaa index 10 permit 2.1.1.0 24
[RouterB] ip prefix-list bbb index 10 permit 10.1.1.0 24
[RouterB] rip 1
[RouterB-rip-1] filter-policy prefix-list aaa import
[RouterB-rip-1] filter-policy prefix-list bbb export
[RouterB-rip-1] quit
```

# 查看路由过滤后 Router A 的路由信息。

```
[RouterA] display rip 1 route
 Route Flags: R - RIP
              A - Aging, S - Suppressed, G - Garbage-collect, D – Direct
              O - Optimal, F - Flush to RIB
 ------------------------------------------------------------------------------
 Peer 1.1.1.2 on GigabitEthernet2/1/1
      Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
      10.1.1.0/24             1.1.1.2          1       0       RAOF    19
 Local route
      Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
      1.1.1.0/24              0.0.0.0          0       0       RDOF    -
      10.1.1.0/24             0.0.0.0          0       0       RDOF    -
      10.2.1.0/24             0.0.0.0          0       0       RDOF    -
```

# 查看 Router B 的路由表信息。

```
[RouterB] display rip 1 route
 Route Flags: R - RIP
              A - Aging, S - Suppressed, G - Garbage-collect, D – Direct
```

```
            O - Optimal, F - Flush to RIB
------------------------------------------------------------------------------
Peer 1.1.1.1 on GigabitEthernet2/1/1
     Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
     2.1.1.0/24              1.1.1.1          1       0       RAOF    19
Local route
     Destination/Mask        Nexthop          Cost    Tag     Flags   Sec
     1.1.1.0/24              0.0.0.0          0       0       RDOF    -
     10.1.1.0/24             0.0.0.0          0       0       RDOF    -
     10.2.1.0/24             0.0.0.0          0       0       RDOF    -
```

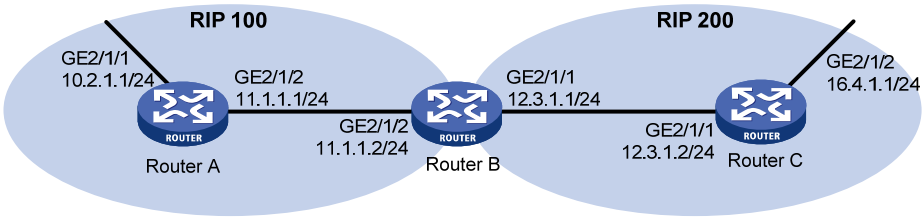
## 1.10.2 配置RIP引入外部路由

### 1. 组网需求：

- Router B 上运行两个 RIP 进程：RIP 100 和 RIP 200。Router B 通过 RIP 100 和 Router A 交换路由信息，通过 RIP 200 和 Router C 交换路由信息。
- 在 Router B 上配置 RIP 进程 200 引入外部路由，引入直连路由和 RIP 进程 100 的路由，使得 Router C 能够学习到达 10.2.1.0/24 和 11.1.1.0/24 的路由，但 Router A 不能学习到达 12.3.1.0/24 和 16.4.1.0/24 的路由。

### 2. 组网图

图1-3 RIP 引入外部路由配置组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 RIP 基本功能

# 在 Router A 上启动 RIP 进程 100，并配置 RIP 版本号为 2。

```
<RouterA> system-view
[RouterA] rip 100
[RouterA-rip-100] network 10.0.0.0
[RouterA-rip-100] network 11.0.0.0
[RouterA-rip-100] version 2
[RouterA-rip-100] undo summary
[RouterA-rip-100] quit
```

# 在 Router B 上启动两个 RIP 进程，进程号分别为 100 和 200，并配置 RIP 版本号为 2。

```
<RouterB> system-view
[RouterB] rip 100
[RouterB-rip-100] network 11.0.0.0
[RouterB-rip-100] version 2
```

```
[RouterB-rip-100] undo summary
[RouterB-rip-100] quit
[RouterB] rip 200
[RouterB-rip-200] network 12.0.0.0
[RouterB-rip-200] version 2
[RouterB-rip-200] undo summary
[RouterB-rip-200] quit
```

# 在 Router C 上启动 RIP 进程 200，并配置 RIP 版本号为 2。

```
<RouterC> system-view
[RouterC] rip 200
[RouterC-rip-200] network 12.0.0.0
[RouterC-rip-200] network 16.0.0.0
[RouterC-rip-200] version 2
[RouterC-rip-200] undo summary
[RouterC-rip-200] quit
```

# 查看 Router C 的路由表信息。

```
[RouterC] display ip routing-table

Destinations : 13        Routes : 13

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
0.0.0.0/32          Direct 0    0            127.0.0.1       InLoop0
12.3.1.0/24         Direct 0    0            12.3.1.2        GE2/1/1
12.3.1.0/32         Direct 0    0            12.3.1.2        GE2/1/1
12.3.1.2/32         Direct 0    0            127.0.0.1       InLoop0
12.3.1.255/32       Direct 0    0            12.3.1.2        GE2/1/1
16.4.1.0/24         Direct 0    0            16.4.1.1        GE2/1/2
16.4.1.0/32         Direct 0    0            16.4.1.1        GE2/1/2
16.4.1.1/32         Direct 0    0            127.0.0.1       InLoop0
16.4.1.255/32       Direct 0    0            16.4.1.1        GE2/1/2
127.0.0.0/8         Direct 0    0            127.0.0.1       InLoop0
127.0.0.0/32        Direct 0    0            127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
127.255.255.255/32  Direct 0    0            127.0.0.1       InLoop0
```

(3) 配置 RIP 引入外部路由

# 在 Router B 配置 RIP 进程 200 引入外部路由，引入直连路由和 RIP 进程 100 的路由。

```
[RouterB] rip 200
[RouterB-rip-200] import-route rip 100
[RouterB-rip-200] import-route direct
[RouterB-rip-200] quit
```

# 查看路由引入后 Router C 的路由表信息。

```
[RouterC] display ip routing-table

Destinations : 15        Routes : 15

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
0.0.0.0/32          Direct 0    0            127.0.0.1       InLoop0
```

```
10.2.1.0/24          RIP    100  1           12.3.1.1        GE2/1/1
11.1.1.0/24          RIP    100  1           12.3.1.1        GE2/1/1
12.3.1.0/24          Direct 0    0           12.3.1.2        GE2/1/1
12.3.1.0/32          Direct 0    0           12.3.1.2        GE2/1/1
12.3.1.2/32          Direct 0    0           127.0.0.1       InLoop0
12.3.1.255/32        Direct 0    0           12.3.1.2        GE2/1/1
16.4.1.0/24          Direct 0    0           16.4.1.1        GE2/1/2
16.4.1.0/32          Direct 0    0           16.4.1.1        GE2/1/2
16.4.1.1/32          Direct 0    0           127.0.0.1       InLoop0
16.4.1.255/32        Direct 0    0           16.4.1.1        GE2/1/2
127.0.0.0/8          Direct 0    0           127.0.0.1       InLoop0
127.0.0.0/32         Direct 0    0           127.0.0.1       InLoop0
127.0.0.1/32         Direct 0    0           127.0.0.1       InLoop0
127.255.255.255/32   Direct 0    0           127.0.0.1       InLoop0
```

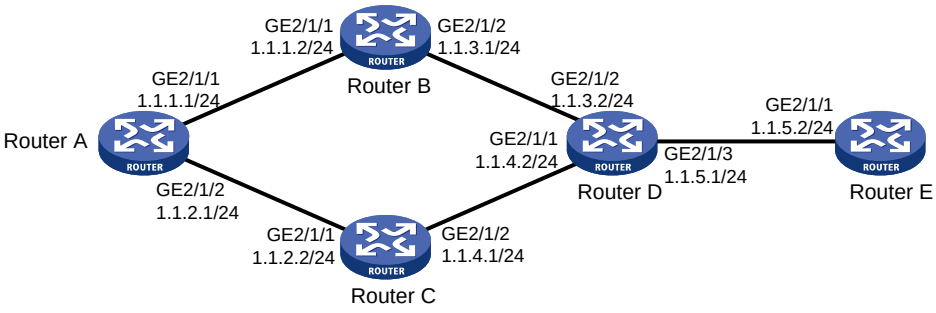
## 1.10.3 配置RIP接口附加度量值

### 1. 组网需求

- 在 Router A、Router B、Router C、Router D 和 Router E 的所有接口上使能 RIP，并使用 RIP-2 进行网络互连。
- Router A 有两条链路可以到达 Router D，其中，通过 Router B 到达 Router D 的链路比通过 Router C 到达 Router D 的链路更加稳定。通过在 Router A 的 GigabitEthernet2/1/2 上配置接口接收 RIP 路由的附加度量值，使得 Router A 优选从 Router B 学到的 1.1.5.0/24 网段的路由。

### 2. 组网图

图1-4 RIP 接口附加度量值配置组网图




### 3. 配置步骤

(1) 配置各接口的地址（略）

(2) 配置 RIP 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] rip
[RouterA-rip-1] network 1.0.0.0
[RouterA-rip-1] version 2
```

```
[RouterA-rip-1] undo summary
[RouterA-rip-1] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] rip
[RouterB-rip-1] network 1.0.0.0
[RouterB-rip-1] version 2
[RouterB-rip-1] undo summary
```

# 配置 Router C。

```
<RouterC> system-view
[RouterB] rip
[RouterC-rip-1] network 1.0.0.0
[RouterC-rip-1] version 2
[RouterC-rip-1] undo summary
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] rip
[RouterD-rip-1] network 1.0.0.0
[RouterD-rip-1] version 2
[RouterD-rip-1] undo summary
```

# 配置 Router E。

```
<RouterE> system-view
[RouterE] rip
[RouterE-rip-1] network 1.0.0.0
[RouterE-rip-1] version 2
[RouterE-rip-1] undo summary
```

# 在 Router A 上查看 RIP 数据库的所有激活路由。

```
[RouterA] display rip 1 database
   1.0.0.0/8, auto-summary
       1.1.1.0/24, cost 0, nexthop 1.1.1.1, RIP-interface
       1.1.2.0/24, cost 0, nexthop 1.1.2.1, RIP-interface
       1.1.3.0/24, cost 1, nexthop 1.1.1.2
       1.1.4.0/24, cost 1, nexthop 1.1.2.2
       1.1.5.0/24, cost 2, nexthop 1.1.1.2
       1.1.5.0/24, cost 2, nexthop 1.1.2.2
```

可以看到，到达网段 1.1.5.0/24 有两条 RIP 路由，下一跳分别是 Router B（IP 地址为 1.1.1.2）和 Router C（IP 地址为 1.1.2.2），cost 值都是 2。

(3) 配置 RIP 接口附加度量值

# 在 Router A 上配置接口 GigabitEthernet2/1/2 的接口附加度量值为 3。

```
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] rip metricin 3
```

# 在 Router A 上查看 RIP 数据库的所有激活路由。

```
[RouterA-GigabitEthernet2/1/2] display rip 1 database
   1.0.0.0/8, auto-summary
       1.1.1.0/24, cost 0, nexthop 1.1.1.1, RIP-interface
       1.1.2.0/24, cost 0, nexthop 1.1.2.1, RIP-interface
```

```
1.1.3.0/24, cost 1, nexthop 1.1.1.2
1.1.4.0/24, cost 2, nexthop 1.1.1.2
1.1.5.0/24, cost 2, nexthop 1.1.1.2
```

可以看到，到达网段 1.1.5.0/24 的 RIP 路由仅有一条，下一跳是 Router B（IP 地址为 1.1.1.2），cost 值为 2。

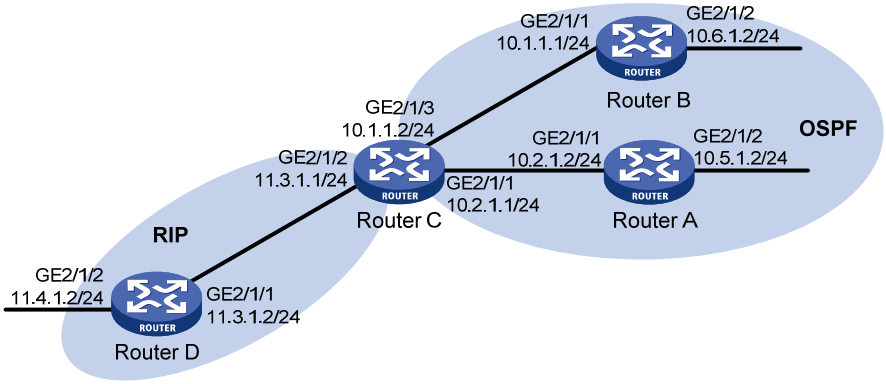
## 1.10.4 配置RIP发布聚合路由

### 1. 组网需求

- Router A、Router B 运行 OSPF，Router D 运行 RIP，Router C 同时运行 OSPF 和 RIP。
- 在 Router C 上配置 RIP 进程引入 OSPF 路由，使 Router D 有到达 10.1.1.0/24、10.2.1.0/24、10.5.1.0/24 和 10.6.1.0/24 网段的路由。
- 为了减小 Router D 的路由表规模，在 Router C 上配置路由聚合，只发布聚合后的路由 10.0.0.0/8。

### 2. 组网图

图1-5 RIP 发布聚合路由配置组网图




### 3. 配置步骤

(1) 配置各接口的地址（略）

(2) 配置 OSPF 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] ospf
[RouterA-ospf-1] area 0
[RouterA-ospf-1-area-0.0.0.0] network 10.5.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] ospf
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
```

```
[RouterB-ospf-1-area-0.0.0.0] network 10.6.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] ospf
[RouterC-ospf-1] area 0
[RouterC-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] quit
[RouterC-ospf-1] quit
```

(3) 配置 RIP 基本功能

# 配置 Router C。

```
[RouterC] rip 1
[RouterC-rip-1] network 11.3.1.0
[RouterC-rip-1] version 2
[RouterC-rip-1] undo summary
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] rip 1
[RouterD-rip-1] network 11.0.0.0
[RouterD-rip-1] version 2
[RouterD-rip-1] undo summary
[RouterD-rip-1] quit
```

# 在 Router C 上配置 RIP 引入外部路由，引入 OSPF 进程 1 的路由和直连路由。

```
[RouterC-rip-1] import-route direct
[RouterC-rip-1] import-route ospf 1
[RouterC-rip-1] quit
```

# 查看 Router D 的路由表信息。

```
[RouterD] display ip routing-table

Destinations : 15        Routes : 15

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
0.0.0.0/32          Direct 0    0           127.0.0.1       InLoop0
10.1.1.0/24         RIP    100  1           11.3.1.1        GE2/1/1
10.2.1.0/24         RIP    100  1           11.3.1.1        GE2/1/1
10.5.1.0/24         RIP    100  1           11.3.1.1        GE2/1/1
10.6.1.0/24         RIP    100  1           11.3.1.1        GE2/1/1
11.3.1.0/24         Direct 0    0           11.3.1.2        GE2/1/1
11.3.1.0/32         Direct 0    0           11.3.1.2        GE2/1/1
11.3.1.2/32         Direct 0    0           127.0.0.1       InLoop0
11.4.1.0/24         Direct 0    0           11.4.1.2        GE2/1/2
11.4.1.0/32         Direct 0    0           11.4.1.2        GE2/1/2
11.4.1.2/32         Direct 0    0           127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0           127.0.0.1       InLoop0
127.0.0.0/32        Direct 0    0           127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0           127.0.0.1       InLoop0
```

```
127.255.255.255/32  Direct 0    0           127.0.0.1       InLoop0
```

(4) 在 Router C 上配置路由聚合，只发布聚合路由 10.0.0.0/8。

```
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] rip summary-address 10.0.0.0 8
```

# 查看 Router D 的路由表信息。

```
[RouterD] display ip routing-table

Destinations : 12        Routes : 12

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
0.0.0.0/32          Direct 0    0           127.0.0.1       InLoop0
10.0.0.0/8          RIP    100  1           11.3.1.1        GE2/1/1
11.3.1.0/24         Direct 0    0           11.3.1.2        GE2/1/1
11.3.1.0/32         Direct 0    0           11.3.1.2        GE2/1/1
11.3.1.2/32         Direct 0    0           127.0.0.1       InLoop0
11.4.1.0/24         Direct 0    0           11.4.1.2        GE2/1/2
11.4.1.0/32         Direct 0    0           11.4.1.2        GE2/1/2
11.4.1.2/32         Direct 0    0           127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0           127.0.0.1       InLoop0
127.0.0.0/32        Direct 0    0           127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0           127.0.0.1       InLoop0
127.255.255.255/32  Direct 0    0           127.0.0.1       InLoop0
```

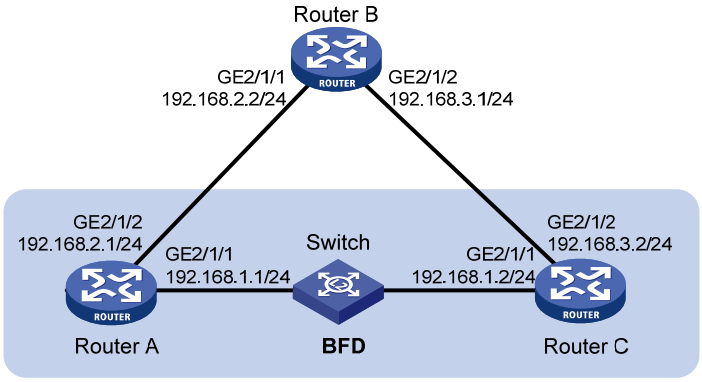
## 1.10.5 配置RIP与BFD联动（echo报文单跳检测）

### 1. 组网需求

- Router A 和 Router C 通过二层交换机互连，它们的接口 GigabitEthernet2/1/1 都运行 RIP 进程 1。并且 Router A 的接口 GigabitEthernet2/1/1 上还使能了 BFD 检测功能。
- Router A 通过 Router B 与 Router C 互连，Router A 的接口 GigabitEthernet2/1/2 运行 RIP 进程 2。Router C 的接口 GigabitEthernet2/1/2、Router B 的接口 GigabitEthernet2/1/1 和 GigabitEthernet2/1/2 上都运行 RIP 进程 1。
- Router C 上配置静态路由，并将静态路由引入 RIP 进程中，使 Router C 有路由发送至 Router A。Router A 上学习到 Router C 发送的静态路由，出接口为与二层交换机相连的接口。
- 在 Router C 和二层交换机之间的链路发生故障后，BFD 能够快速检测链路中断并通告 RIP 协议。RIP 协议响应 BFD 会话 down，删除与 Router C 的邻居，并删除从 Router C 学习的路由。Router A 上学习到 Router C 发送的静态路由，出接口为与 Router B 连接的接口。

## 2. 组网图

图1-6 RIP 与 BFD 联动配置组网图（echo 报文单跳检测）




## 3. 配置步骤

(1) 配置 RIP 基本功能并且在接口上使能 BFD

# 配置 Router A。

```
<RouterA> system-view
[RouterA] rip 1
[RouterA-rip-1] version 2
[RouterA-rip-1] undo summary
[RouterA-rip-1] network 192.168.1.0
[RouterA-rip-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] rip bfd enable
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] rip 2
[RouterA-rip-2] network 192.168.2.0
[RouterA-rip-2] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] rip 1
[RouterB-rip-1] version 2
[RouterB-rip-1] undo summary
[RouterB-rip-1] network 192.168.2.0
[RouterB-rip-1] network 192.168.3.0
[RouterB-rip-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] rip 1
[RouterC-rip-1] version 2
[RouterC-rip-1] undo summary
[RouterC-rip-1] network 192.168.1.0
[RouterC-rip-1] network 192.168.3.0
[RouterC-rip-1] import-route static
```

```
[RouterC-rip-1] quit
```

(2) 配置接口 BFD 参数

# 配置 Router A。

```
[RouterA] bfd session init-mode active
[RouterA] bfd echo-source-ip 11.11.11.11
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterA-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterA-GigabitEthernet2/1/1] bfd detect-multiplier 7
[RouterA-GigabitEthernet2/1/1] return
```

(3) Router C 配置静态路由

```
[RouterC] ip route-static 120.1.1.1 24 null 0
```

## 4. 验证配置

# 查看 Router A 的 BFD 信息。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv4 Session Working Under Echo Mode:

 LD              SourceAddr      DestAddr        State    Holdtime    Interface
 4               192.168.1.1     192.168.1.2     Up       2000ms      GE2/1/1
```

# 查看 Router A 上学到的路由 120.1.1.0/24，可以看到 Router A 经过 L2 Switch 到达 Router C。

```
<RouterA> display ip routing-table 120.1.1.0 24 verbose

Summary Count : 1

Destination: 120.1.1.0/24
   Protocol: RIP             Process ID: 1
  SubProtID: 0x1                    Age: 04h20m37s
       Cost: 1               Preference: 100
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 192.168.1.2
      Flags: 0x1008c        OrigNextHop: 192.168.1.2
      Label: NULL           RealNextHop: 192.168.1.2
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid        BkInterface: N/A
```

当 Router C 和二层交换机之间的链路发生故障时：

# 查看 Router A 上学到的路由 120.1.1.0/24，可以看到 Router A 经过 Router B 到达 Router C。

```
<RouterA> display ip routing-table 120.1.1.0 24 verbose

Summary Count : 1
```

```
Destination: 120.1.1.0/24
   Protocol: RIP             Process ID: 2
  SubProtID: 0x1                    Age: 04h20m37s
       Cost: 1               Preference: 100
        Tag: 0                     State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 192.168.2.2
      Flags: 0x1008c        OrigNextHop: 192.168.2.2
      Label: NULL           RealNextHop: 192.168.2.2
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: N/A
```

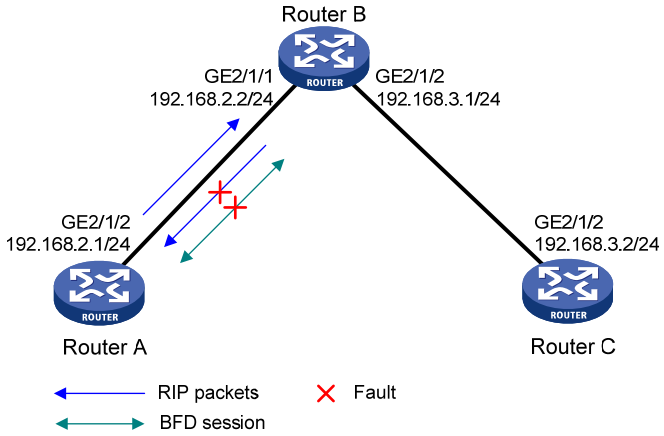
### 1.10.6 配置RIP与BFD联动（指定目的地址的echo报文单跳检测）

#### 1. 组网需求

- Router A 和 Router B 互连，Router A 的接口 GigabitEthernet2/1/2 和 Router B 的接口 GigabitEthernet2/1/1 都运行 RIP 进程 1。Router A 的接口 GigabitEthernet2/1/2 上使能了 BFD 检测功能，指定目的地址为 Router B 的接口 GigabitEthernet2/1/1 的地址。
- Router B 与 Router C 互连，它们的接口 GigabitEthernet2/1/2 都运行 RIP 进程 1。
- Router A 和 Router C 上配置静态路由，并都将静态路由引入 RIP 进程中，其中 Router A 引入路由的 cost 值比 Router C 引入的 cost 值小，这样，当 Router B 上学习到 Router A 和 Router C 发送的路由后，会优选 Router A 的路由，出接口为与 Router A 连接的接口。
- 在 Router A 和 Router B 之间的链路发生单通故障（从 Router A 到 Router B 方向报文是通的，但是从 Router B 到 Router A 方向的链路不通）后，Router A 上 BFD 能够快速检测链路故障并通告 RIP 协议。Router A 上的 RIP 协议响应 BFD 会话 down，删除从 GigabitEthernet2/1/2 口学习到的邻居和路由，并不再从该接口接收和发送 RIP 报文。Router B 上在学自 Router A 的路由老化后，会优选 Router C 发送的静态路由，出接口为与 Router C 连接的接口。

## 2. 组网图

图1-7 RIP 与 BFD 联动配置组网图（指定目的地址的 echo 报文单跳检测）




## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 RIP 基本功能并且在接口上使能 BFD

# 配置 Router A。

```
<RouterA> system-view
[RouterA] rip 1
[RouterA-rip-1] network 192.168.2.0
[RouterA-rip-1] import-route static
[RouterA-rip-1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] rip bfd enable destination 192.168.2.2
[RouterA-GigabitEthernet2/1/2] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] rip 1
[RouterB-rip-1] network 192.168.2.0
[RouterB-rip-1] network 192.168.3.0
[RouterB-rip-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] rip 1
[RouterC-rip-1] network 192.168.3.0
[RouterC-rip-1] import-route static cost 3
[RouterC-rip-1] quit
```

(3) 配置接口 BFD 参数

# 配置 Router A。

```
[RouterA] bfd echo-source-ip 11.11.11.11
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] bfd min-echo-receive-interval 500
```

```
[RouterA-GigabitEthernet2/1/2] quit
```

(4) 配置静态路由

# 配置 Router A。

```
[RouterA] ip route-static 100.1.1.0 24 null 0
```

# 配置 Router C。

```
[RouterC] ip route-static 100.1.1.0 24 null 0
```

### 4. 验证配置

# 显示 Router A 的 BFD 信息。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv4 session working under Echo mode:

 LD              SourceAddr      DestAddr        State    Holdtime     Interface
 3               192.168.2.1     192.168.2.2     Up       2000ms       GE2/1/2
```

# 显示 Router B 上学到的路由 100.1.1.0/24。

```
<RouterB> display ip routing-table 100.1.1.0 24 verbose

Summary Count : 1

Destination: 100.1.1.0/24
   Protocol: RIP             Process ID: 1
  SubProtID: 0x1                    Age: 00h02m47s
       Cost: 1              Preference: 100
        Tag: 0                   State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x12000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 192.168.2.1
      Flags: 0x1008c        OrigNextHop: 192.168.2.1
      Label: NULL           RealNextHop: 192.168.2.1
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid        BkInterface: N/A
```

当 Router A 和 Router B 之间的链路发生故障时：

# 显示 Router B 上学到的路由 100.1.1.0/24。

```
<RouterB> display ip routing-table 100.1.1.0 24 verbose

Summary Count : 1

Destination: 100.1.1.0/24
   Protocol: RIP             Process ID: 1
  SubProtID: 0x1                    Age: 00h21m23s
       Cost: 4              Preference: 100
        Tag: 0                   State: Active Adv
```

```
  OrigTblID: 0x0                OrigVrf: default-vrf
    TableID: 0x2                 OrigAs: 0
      NBRID: 0x12000003          LastAs: 0
     AttrID: 0xffffffff        Neighbor: 192.168.3.2
      Flags: 0x1008c        OrigNextHop: 192.168.3.2
      Label: NULL           RealNextHop: 192.168.3.2
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: N/A
```

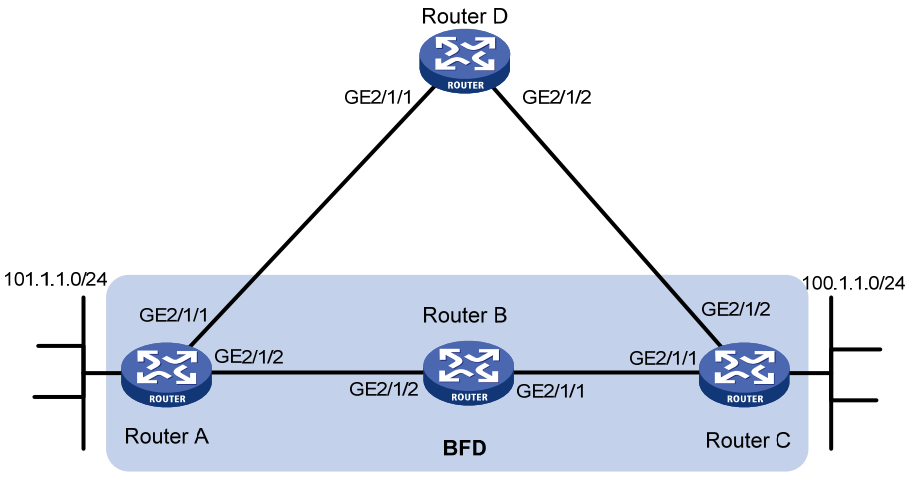
## 1.10.7 配置RIP与BFD联动（control报文双向检测）

### 1. 组网需求

- Router A 通过 Router B 与 Router C 互连。Rourer A 的接口 GigabitEthernet2/1/2 和 Rouer C 的接口 GigabitEthernet2/1/1 上都运行 RIP 进程 1。分别在 Router A 和 Router C 上配置到达对端的静态路由，并在 Rourer A 的接口 GigabitEthernet2/1/2 和 Rouer C 的接口 GigabitEthernet2/1/1 上使能 BFD 检测功能。
- Router A 通过 Router D 与 Router C 互连。Rourer A 的接口 GigabitEthernet2/1/1 运行 RIP 进程 2。Rouer C 的接口 GigabitEthernet2/1/2、Rourer D 的接口 GigabitEthernet2/1/1 和 GigabitEthernet2/1/2 上运行 RIP 进程 1。
- 为使 Router A 与 Router C 互有路由发送，Router A 与 Router C 上的 RIP 协议都要配置引入静态路由。Router A 建立至 Router C 的会话。Router A 上学习到 Router C 发送的静态路由，出接口为与 Router B 连接的接口。
- 在 Router B 与 Router C 之间的链路发生故障后，BFD 能够快速检测链路中断并通告 RIP 协议。RIP 协议响应 BFD 会话 down，删除与 Router C 的邻居，并删除从 Router C 学习的路由。Router A 上学习到 Router C 发送的静态路由，出接口为与 Router D 连接的接口。

### 2. 组网图

图1-8 RIP 与 BFD 联动配置组网图（control 报文双向检测）




| 设备 | 接口 | IP地址 | 设备 | 接口 | IP地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 192.168.3.1/24 | Router B | GE2/1/1 | 192.168.2.1/24 |
| | GE2/1/2 | 192.168.1.1/24 | | GE2/1/2 | 192.168.1.2/24 |
| Router C | GE2/1/1 | 192.168.2.2/24 | Router D | GE2/1/1 | 192.168.3.2/24 |

| | | | | | |
|---|---|---|---|---|---|
| | GE2/1/2 | 192.168.4.2/24 | | GE2/1/2 | 192.168.4.1/24 |

### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 RIP 基本功能，并引入静态路由，使 Router A 与 Router C 互有路由发送

# 配置 Router A。

```
<RouterA> system-view
[RouterA] rip 1
[RouterA-rip-1] version 2
[RouterA-rip-1] undo summary
[RouterA-rip-1] network 192.168.1.0
[RouterA-rip-1] network 101.1.1.0
[RouterA-rip-1] peer 192.168.2.2
[RouterA-rip-1] undo validate-source-address
[RouterA-rip-1] import-route static
[RouterA-rip-1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] rip bfd enable
[RouterA-GigabitEthernet2/1/2] quit
[RouterA] rip 2
[RouterA-rip-2] version 2
[RouterA-rip-2] undo summary
[RouterA-rip-2] network 192.168.3.0
[RouterA-rip-2] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] rip 1
[RouterC-rip-1] version 2
[RouterC-rip-1] undo summary
[RouterC-rip-1] network 192.168.2.0
[RouterC-rip-1] network 192.168.4.0
[RouterC-rip-1] network 100.1.1.0
[RouterC-rip-1] peer 192.168.1.1
[RouterC-rip-1] undo validate-source-address
[RouterC-rip-1] import-route static
[RouterC-rip-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] rip bfd enable
[RouterC-GigabitEthernet2/1/1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] rip 1
[RouterD-rip-1] version 2
[RouterD-rip-1] undo summary
[RouterD-rip-1] network 192.168.3.0
[RouterD-rip-1] network 192.168.4.0
```

```
[RouterD-rip-1] quit
```

(3) 配置接口及 BFD 参数

# 配置 Router A。

```
[RouterA] bfd session init-mode active
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ip address 192.168.3.1 24
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] ip address 192.168.1.1 24
[RouterA-GigabitEthernet2/1/2] bfd min-transmit-interval 500
[RouterA-GigabitEthernet2/1/2] bfd min-receive-interval 500
[RouterA-GigabitEthernet2/1/2] bfd detect-multiplier 7
[RouterA-GigabitEthernet2/1/2] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] ip address 192.168.1.2 24
[RouterB-GigabitEthernet2/1/2] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ip address 192.168.2.1 24
```

# 配置 Router C。

```
[RouterC] bfd session init-mode active
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ip address 192.168.2.2 24
[RouterC-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterC-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterC-GigabitEthernet2/1/1] bfd detect-multiplier 6
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] ip address 192.168.4.2 24
[RouterC-GigabitEthernet2/1/2] quit
```

# 配置 Router D。

```
[RouterD] interface gigabitethernet 2/1/2
[RouterD-GigabitEthernet2/1/2] ip address 192.168.4.1 24
[RouterD-GigabitEthernet2/1/2] quit
[RouterD] interface gigabitethernet 2/1/1
[RouterD-GigabitEthernet2/1/1] ip address 192.168.3.2 24
[RouterD-GigabitEthernet2/1/1] quit
```

(4) 配置静态路由

# 配置 Router A。

```
[RouterA] ip route-static 192.168.2.0 24 GigabitEthernet2/1/2 192.168.1.2
[RouterA] quit
```

# 配置 Router C。

```
[RouterC] ip route-static 192.168.1.0 24 GigabitEthernet2/1/1 192.168.2.1
```

### 4. 验证配置

# 显示 Router A 的 BFD 信息。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv4 session working under Ctrl mode:

 LD/RD           SourceAddr      DestAddr        State    Holdtime    Interface
 513/513         192.168.1.1     192.168.2.2     Up       1700ms      GE2/1/2
```

# 显示 Router A 上学到的路由 100.1.1.0/24。

```
<RouterB> display ip routing-table 100.1.1.0 24 verbose

Summary Count : 1

Destination: 100.1.1.0/24
   Protocol: RIP             Process ID: 1
  SubProtID: 0x1                    Age: 00h04m02s
       Cost: 1               Preference: 100
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x12000002         LastAs: 0
     AttrID: 0xffffffff        Neighbor: 192.168.2.2
      Flags: 0x1008c        OrigNextHop: 192.168.2.2
      Label: NULL           RealNextHop: 192.168.1.2
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: N/A
```

Router B 和 Router C 之间的链路发生故障后：

# 显示 Router A 上学到的路由 100.1.1.0/24。

```
<RouterA> display ip routing-table 100.1.1.0 verbose

Summary Count : 1

Destination: 100.1.1.0/24
   Protocol: RIP             Process ID: 2
  SubProtID: 0x1                    Age: 00h10m35s
       Cost: 2               Preference: 100
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x12000003         LastAs: 0
     AttrID: 0xffffffff        Neighbor: 192.168.3.2
      Flags: 0x1008c        OrigNextHop: 192.168.3.2
      Label: NULL           RealNextHop: 192.168.3.2
    BkLabel: NULL             BkNextHop: N/A
```

```
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: N/A
```

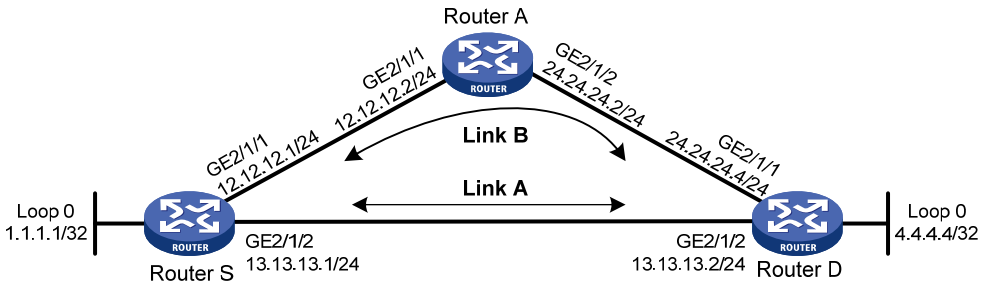
## 1.10.8 配置RIP快速重路由

### 1. 组网需求

如 图 1-9 所示，Router S、Router A和Router D通过RIPv2 协议实现网络互连。要求当Router S和Router D之间的链路出现单通故障时，业务可以快速切换到链路B上。

### 2. 组网图

图1-9 RIP 快速重路由配置组网图




### 3. 配置步骤

(1) 配置各路由器接口的 IP 地址和 RIPv2 协议

请按照上面组网图配置各接口的 IP 地址和子网掩码，具体配置过程略。

配置各路由器之间采用 RIPv2 协议进行互连，确保 Router S、Router A 和 Router D 之间能够在网络层互通，并且各路由器之间能够借助 RIPv2 协议实现动态路由更新。

具体配置过程略。

(2) 配置 RIP 快速重路由

# 配置 Router S。

```
<RouterS> system-view
[RouterS] ip prefix-list abc index 10 permit 4.4.4.4 32
[RouterS] route-policy frr permit node 10
[RouterS-route-policy-frr-10] if-match ip address prefix-list abc
[RouterS-route-policy-frr-10] apply fast-reroute backup-interface gigabitethernet 2/1/1
backup-nexthop 12.12.12.2
[RouterS-route-policy-frr-10] quit
[RouterS] rip 1
[RouterS-rip-1] fast-reroute route-policy frr
[RouterS-rip-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] ip prefix-list abc index 10 permit 1.1.1.1 32
[RouterD] route-policy frr permit node 10
[RouterD-route-policy-frr-10] if-match ip address prefix-list abc
[RouterD-route-policy-frr-10] apply fast-reroute backup-interface gigabitethernet 2/1/1
backup-nexthop 24.24.24.2
```

```
[RouterD-route-policy-frr-10] quit
[RouterD] rip 1
[RouterD-rip-1] fast-reroute route-policy frr
[RouterD-rip-1] quit
```

### 4. 验证配置

# 在 Router S 上查看 4.4.4.4/32 路由，可以看到备份下一跳信息：

```
[RouterS] display ip routing-table 4.4.4.4 verbose

Summary Count : 1

Destination: 4.4.4.4/32
   Protocol: RIP             Process ID: 1
  SubProtID: 0x1                    Age: 04h20m37s
       Cost: 1              Preference: 100
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 13.13.13.2
      Flags: 0x1008c        OrigNextHop: 13.13.13.2
      Label: NULL           RealNextHop: 13.13.13.2
    BkLabel: NULL             BkNextHop: 12.12.12.2
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: GigabitEthernet2/1/1
```

# 在 Router D 上查看 1.1.1.1/32 路由，可以看到备份下一跳信息：

```
[RouterS] display ip routing-table 1.1.1.1 verbose

Summary Count : 1

Destination: 1.1.1.1/32
   Protocol: RIP             Process ID: 1
  SubProtID: 0x1                    Age: 04h20m37s
       Cost: 1              Preference: 100
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 13.13.13.1
      Flags: 0x1008c        OrigNextHop: 13.13.13.1
      Label: NULL           RealNextHop: 13.13.13.1
    BkLabel: NULL             BkNextHop: 24.24.24.2
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: GigabitEthernet2/1/1
```

# 目 录

# 1 OSPF

## 1.1 OSPF简介

OSPF（Open Shortest Path First，开放最短路径优先）是 IETF（Internet Engineering Task Force，互联网工程任务组）组织开发的一个基于链路状态的内部网关协议。目前针对 IPv4 协议使用的是 OSPF Version 2。

下文中所提到的 OSPF 均指 OSPF Version 2。

### 1.1.1 OSPF的特点

OSPF 具有如下特点：

- 适应范围广：支持各种规模的网络，最多可支持几百台路由器。
- 快速收敛：在网络的拓扑结构发生变化后立即发送更新报文，使这一变化在自治系统中同步。
- 无自环：由于 OSPF 根据收集到的链路状态用最短路径树算法计算路由，从算法本身保证了不会生成自环路由。
- 区域划分：允许自治系统的网络被划分成区域来管理。路由器链路状态数据库的减小降低了内存的消耗和 CPU 的负担；区域间传送路由信息的减少降低了网络带宽的占用。
- 等价路由：支持到同一目的地址的多条等价路由。
- 路由分级：使用 4 类不同的路由，按优先顺序来说分别是：区域内路由、区域间路由、第一类外部路由、第二类外部路由。
- 支持验证：支持基于区域和接口的报文验证，以保证报文交互和路由计算的安全性。
- 组播发送：在某些类型的链路上以组播地址发送协议报文，减少对其他设备的干扰。

### 1.1.2 OSPF报文类型

OSPF 协议报文直接封装为 IP 报文，协议号为 89。

OSPF 有五种类型的协议报文：

- Hello 报文：周期性发送，用来发现和维持 OSPF 邻居关系，以及进行 DR（Designated Router，指定路由器）/BDR（Backup Designated Router，备份指定路由器）的选举。
- DD（Database Description，数据库描述）报文：描述了本地 LSDB（Link State DataBase，链路状态数据库）中每一条 LSA（Link State Advertisement，链路状态通告）的摘要信息，用于两台路由器进行数据库同步。
- LSR（Link State Request，链路状态请求）报文：向对方请求所需的 LSA。两台路由器互相交换 DD 报文之后，得知对端的路由器有哪些 LSA 是本地的 LSDB 所缺少的，这时需要发送 LSR 报文向对方请求所需的 LSA。
- LSU（Link State Update，链路状态更新）报文：向对方发送其所需要的 LSA。
- LSAck（Link State Acknowledgment，链路状态确认）报文：用来对收到的 LSA 进行确认。

### 1.1.3 LSA类型

OSPF 中对链路状态信息的描述都是封装在 LSA 中发布出去，常用的 LSA 有以下几种类型：

- Router LSA（Type-1）：由每个路由器产生，描述路由器的链路状态和开销，在其始发的区域内传播。
- Network LSA（Type-2）：由 DR 产生，描述本网段所有路由器的链路状态，在其始发的区域内传播。
- Network Summary LSA（Type-3）：由 ABR（Area Border Router，区域边界路由器）产生，描述区域内某个网段的路由，并通告给其他区域。
- ASBR Summary LSA（Type-4）：由 ABR 产生，描述到 ASBR（Autonomous System Boundary Router，自治系统边界路由器）的路由，通告给相关区域。
- AS External LSA（Type-5）：由 ASBR 产生，描述到 AS（Autonomous System，自治系统）外部的路由，通告到所有的区域（除了 Stub 区域和 NSSA 区域）。
- NSSA External LSA（Type-7）：由 NSSA（Not-So-Stubby Area）区域内的 ASBR 产生，描述到 AS 外部的路由，仅在 NSSA 区域内传播。
- Opaque LSA：用于 OSPF 的扩展通用机制，目前有 Type-9、Type-10 和 Type-11 三种。其中，Type-9 LSA 仅在本地链路范围进行泛洪，用于支持 GR（Graceful Restart，平滑重启）的 Grace LSA 就是 Type-9 的一种类型；Type-10 LSA 仅在区域范围进行泛洪，用于支持 MPLS TE 的 LSA 就是 Type-10 的一种类型；Type-11 LSA 可以在一个自治系统范围进行泛洪。

### 1.1.4 OSPF区域

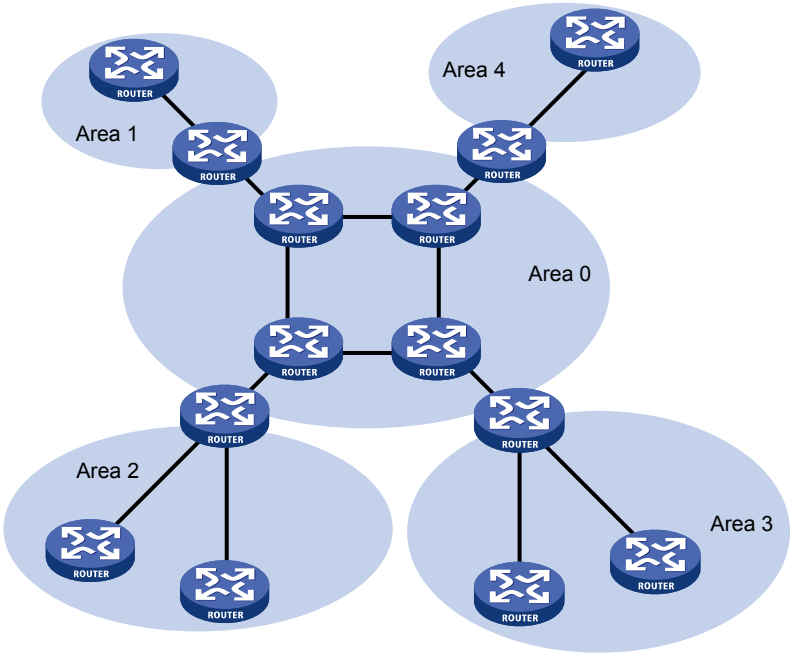
#### 1. 区域划分

随着网络规模日益扩大，当一个大型网络中的路由器都运行 OSPF 协议时，LSDB 会占用大量的存储空间，并使得运行 SPF（Shortest Path First，最短路径优先）算法的复杂度增加，导致 CPU 负担加重。

在网络规模增大之后，拓扑结构发生变化的概率也增大，网络会经常处于“振荡”之中，造成网络中会有大量的 OSPF 协议报文在传递，降低了网络的带宽利用率。更为严重的是，每一次变化都会导致网络中所有的路由器重新进行路由计算。

OSPF协议通过将自治系统划分成不同的区域来解决上述问题。区域是从逻辑上将路由器划分为不同的组，每个组用区域号来标识。如 图 1-1 所示。

图1-1 OSPF 区域划分




区域的边界是路由器，而不是链路。一个路由器可以属于不同的区域，但是一个网段（链路）只能属于一个区域，或者说每个运行 OSPF 的接口必须指明属于哪一个区域。划分区域后，可以在区域边界路由器上进行路由聚合，以减少通告到其他区域的 LSA 数量，还可以将网络拓扑变化带来的影响最小化。

### 2. 骨干区域与虚连接

(1) 骨干区域（Backbone Area）

OSPF 划分区域之后，并非所有的区域都是平等的关系。其中有一个区域是与众不同的，它的区域号是 0，通常被称为骨干区域。骨干区域负责区域之间的路由，非骨干区域之间的路由信息必须通过骨干区域来转发。对此，OSPF 有两个规定：

- 所有非骨干区域必须与骨干区域保持连通；
- 骨干区域自身也必须保持连通。

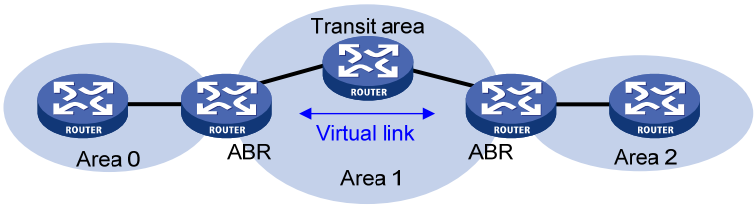
在实际应用中，可能会因为各方面条件的限制，无法满足上面的要求。这时可以通过配置 OSPF 虚连接予以解决。

(2) 虚连接（Virtual Link）

虚连接是指在两台 ABR 之间通过一个非骨干区域而建立的一条逻辑上的连接通道。它的两端必须是 ABR，而且必须在两端同时配置方可生效。为虚连接两端提供一条非骨干区域内部路由的区域称为传输区（Transit Area）。

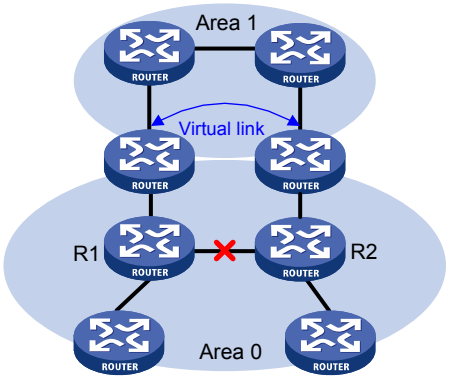
在 图 1-2 中，Area2 与骨干区域之间没有直接相连的物理链路，但可以在ABR上配置虚连接，使Area2通过一条逻辑链路与骨干区域保持连通。

图1-2 虚连接示意图之一




虚连接的另外一个应用是提供冗余的备份链路，当骨干区域因链路故障不能保持连通时，通过虚连接仍然可以保证骨干区域在逻辑上的连通性。如 图 1-3 所示。

图1-3 虚连接示意图之二




虚连接相当于在两个 ABR 之间形成了一个点到点的连接，因此，在这个连接上，和物理接口一样可以配置接口的各参数，如发送 Hello 报文间隔等。

两台 ABR 之间直接传递 OSPF 报文信息，它们之间的 OSPF 路由器只是起到一个转发报文的作用。由于协议报文的目的地址不是中间这些路由器，所以这些报文对于它们而言是透明的，只是当作普通的 IP 报文来转发。

### 3. Stub区域和Totally Stub区域

Stub 区域是一些特定的区域，该区域的 ABR 会将区域间的路由信息传递到本区域，但不会引入自治系统外部路由，区域中路由器的路由表规模以及 LSA 数量都会大大减少。为保证到自治系统外的路由依旧可达，该区域的 ABR 将生成一条缺省路由 Type-3 LSA，发布给本区域中的其他非 ABR 路由器。

为了进一步减少 Stub 区域中路由器的路由表规模以及 LSA 数量，可以将区域配置为 Totally Stub（完全 Stub）区域，该区域的 ABR 不会将区域间的路由信息和自治系统外部路由信息传递到本区域。为保证到本自治系统的其他区域和自治系统外的路由依旧可达，该区域的 ABR 将生成一条缺省路由 Type-3 LSA，发布给本区域中的其他非 ABR 路由器。

### 4. NSSA区域和Totally NSSA区域

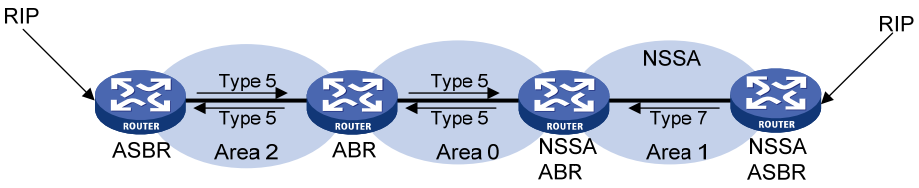
NSSA（Not-So-Stubby Area）区域是 Stub 区域的变形，与 Stub 区域的区别在于 NSSA 区域允许引入自治系统外部路由，由 ASBR 发布 Type-7 LSA 通告给本区域。当 Type-7 LSA 到达 NSSA 的 ABR 时，由 ABR 将 Type-7 LSA 转换成 Type-5 LSA，传播到其他区域。

可以将区域配置为 Totally NSSA（完全 NSSA）区域，该区域的 ABR 不会将区域间的路由信息传递到本区域。为保证到本自治系统的其他区域的路由依旧可达，该区域的 ABR 将生成一条缺省路由 Type-3 LSA，发布给本区域中的其他非 ABR 路由器。

如 图 1-4 所示，运行OSPF协议的自治系统包括 3 个区域：区域 0、区域 1 和区域 2，另外两个自治系统运行RIP协议。区域 1 被定义为NSSA区域，区域 1 接收的RIP路由传播到NSSA ASBR后，由NSSA ASBR产生Type-7 LSA在区域 1 内传播，当Type-7 LSA到达NSSA ABR后，转换成Type-5 LSA传播到区域 0 和区域 2。

另一方面，运行 RIP 的自治系统的 RIP 路由通过区域 2 的 ASBR 产生 Type-5 LSA 在 OSPF 自治系统中传播。但由于区域 1 是 NSSA 区域，所以 Type-5 LSA 不会到达区域 1。

图1-4 NSSA 区域




## 1.1.5 路由器类型

OSPF 路由器根据在 AS 中的不同位置，可以分为以下四类：

### 1. 区域内路由器（Internal Router）

该类路由器的所有接口都属于同一个 OSPF 区域。

### 2. 区域边界路由器ABR

该类路由器可以同时属于两个以上的区域，但其中一个必须是骨干区域。ABR 用来连接骨干区域和非骨干区域，它与骨干区域之间既可以是物理连接，也可以是逻辑上的连接。

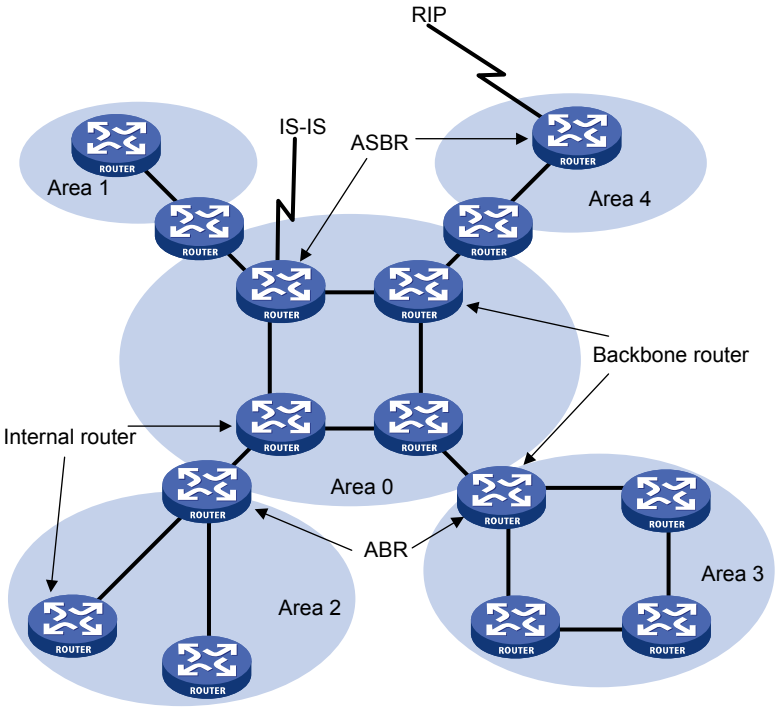
### 3. 骨干路由器（Backbone Router）

该类路由器至少有一个接口属于骨干区域。因此，所有的 ABR 和位于 Area0 的内部路由器都是骨干路由器。

### 4. 自治系统边界路由器ASBR

与其他 AS 交换路由信息的路由器称为 ASBR。ASBR 并不一定位于 AS 的边界，它有可能是区域内路由器，也有可能是 ABR。只要一台 OSPF 路由器引入了外部路由的信息，它就成为 ASBR。

图1-5 OSPF 路由器的类型




## 1.1.6 路由类型

OSPF 将路由分为四类，按照优先级从高到低的顺序依次为：

- 区域内路由（Intra Area）
- 区域间路由（Inter Area）
- 第一类外部路由（Type1 External）：这类路由的可信程度较高，并且和 OSPF 自身路由的开销具有可比性，所以到第一类外部路由的开销等于本路由器到相应的 ASBR 的开销与 ASBR 到该路由目的地址的开销之和。
- 第二类外部路由（Type2 External）：这类路由的可信度比较低，所以 OSPF 协议认为从 ASBR 到自治系统之外的开销远远大于在自治系统之内到达 ASBR 的开销。所以计算路由开销时将主要考虑前者，即到第二类外部路由的开销等于 ASBR 到该路由目的地址的开销。如果计算出开销值相等的两条路由，再考虑本路由器到相应的 ASBR 的开销。

区域内和区域间路由描述的是 AS 内部的网络结构，外部路由则描述了应该如何选择到 AS 以外目的地址的路由。

## 1.1.7 OSPF路由的计算过程

同一个区域内，OSPF 路由的计算过程可简单描述如下：

- 每台 OSPF 路由器根据自己周围的网络拓扑结构生成 LSA，并通过更新报文将 LSA 发送给网络中的其它 OSPF 路由器。
- 每台 OSPF 路由器都会收集其它路由器通告的 LSA，所有的 LSA 放在一起便组成了 LSDB。LSA 是对路由器周围网络拓扑结构的描述，LSDB 则是对整个自治系统的网络拓扑结构的描述。

- OSPF 路由器将 LSDB 转换成一张带权的有向图，这张图便是对整个网络拓扑结构的真实反映。各个路由器得到的有向图是完全相同的。
- 每台路由器根据有向图，使用 SPF 算法计算出一棵以自己为根的最短路径树，这棵树给出了到自治系统中各节点的路由。

### 1.1.8 OSPF的网络类型

OSPF 根据链路层协议类型将网络分为下列四种类型：

- 广播（Broadcast）类型：当链路层协议是 Ethernet、FDDI 时，缺省情况下，OSPF 认为网络类型是 Broadcast。在该类型的网络中，通常以组播形式（OSPF 路由器的预留 IP 组播地址是 224.0.0.5；OSPF DR 的预留 IP 组播地址是 224.0.0.6）发送 Hello 报文、LSU 报文和 LSAck 报文；以单播形式发送 DD 报文和 LSR 报文。
- NBMA（Non-Broadcast Multi-Access，非广播多路访问）类型：当链路层协议是帧中继、ATM 或 X.25 时，缺省情况下，OSPF 认为网络类型是 NBMA。在该类型的网络中，以单播形式发送协议报文。
- P2MP（Point-to-MultiPoint，点到多点）类型：没有一种链路层协议会被缺省的认为是 P2MP 类型。P2MP 必须是由其他的网络类型强制更改的，常用做法是将 NBMA 网络改为 P2MP 网络。在该类型的网络中，缺省情况下，以组播形式（224.0.0.5）发送协议报文。可以根据用户需要，以单播形式发送协议报文。
- P2P（Point-to-Point，点到点）类型：当链路层协议是 PPP、HDLC 时，缺省情况下，OSPF 认为网络类型是 P2P。在该类型的网络中，以组播形式（224.0.0.5）发送协议报文。

NBMA 与 P2MP 网络之间的区别如下：

- NBMA 网络是全连通的；P2MP 网络并不需要一定是全连通的。
- NBMA 网络中需要选举 DR 与 BDR；P2MP 网络中没有 DR 与 BDR。
- NBMA 网络采用单播发送报文，需要手工配置邻居；P2MP 网络采用组播方式发送报文，通过配置也可以采用单播发送报文。

### 1.1.9 DR/BDR

#### 1. DR/BDR简介

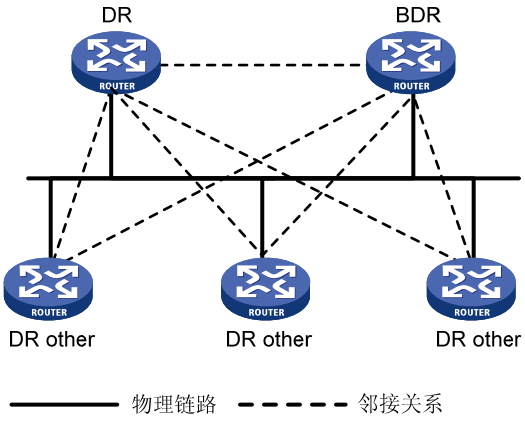
在广播网和 NBMA 网络中，任意两台路由器之间都要交换路由信息。如果网络中有 n 台路由器，则需要建立 n（n-1）/2 个邻接关系。这使得任何一台路由器的路由变化都会导致多次传递，浪费了带宽资源。为解决这一问题，OSPF 提出了 DR 的概念，所有路由器只将信息发送给 DR，由 DR 将网络链路状态发送出去。

另外，OSPF 提出了 BDR 的概念。BDR 是对 DR 的一个备份，在选举 DR 的同时也选举 BDR，BDR 也和本网段内的所有路由器建立邻接关系并交换路由信息。当 DR 失效后，BDR 会立即成为新的 DR。

OSPF 网络中，既不是 DR 也不是 BDR 的路由器为 DR Other。DR Other 仅与 DR 和 BDR 建立邻接关系，DR Other 之间不交换任何路由信息。这样就减少了广播网和 NBMA 网络上各路由器之间邻接关系的数量，同时减少网络流量，节约了带宽资源。

如 图 1-6 所示，进行DR/BDR选举后，5 台路由器之间只需要建立 7 个邻接关系就可以了。

图1-6 DR 和 BDR 示意图




说明

在 OSPF 中，邻居（Neighbor）和邻接（Adjacency）是两个不同的概念。路由器启动后，会通过接口向外发送 Hello 报文，收到 Hello 报文的路由器会检查报文中所定义的参数，如果双方一致就会形成邻居关系。只有当双方成功交换 DD 报文，交换 LSA 并达到 LSDB 同步之后，才形成邻接关系。

### 2. DR/BDR选举过程

DR/BDR 是由同一网段中所有的路由器根据路由器优先级和 Router ID 通过 Hello 报文选举出来的，只有优先级大于 0 的路由器才具有选举资格。

进行 DR/BDR 选举时每台路由器将自己选出的 DR 写入 Hello 报文中，发给网段上每台运行 OSPF 协议的路由器。当处于同一网段的两台路由器同时宣布自己是 DR 时，路由器优先级高者胜出。如果优先级相等，则 Router ID 大者胜出。

需要注意的是：

- 只有在广播或 NBMA 网络中才会选举 DR；在 P2P 或 P2MP 网络中不需要选举 DR。
- DR 是某个网段中的概念，是针对路由器的接口而言的。某台路由器在一个接口上可能是 DR，在另一个接口上有可能是 BDR，或者是 DR Other。
- DR/BDR 选举完毕后，即使网络中加入一台具有更高优先级的路由器，也不会重新进行选举，替换该网段中已经存在的 DR/BDR 成为新的 DR/BDR。DR 并不一定就是路由器优先级最高的路由器接口；同理，BDR 也并不一定就是路由器优先级次高的路由器接口。

## 1.1.10 协议规范

与 OSPF 相关的协议规范有：

- RFC 1765：OSPF Database Overflow
- RFC 2328：OSPF Version 2
- RFC 3101：OSPF Not-So-Stubby Area (NSSA) Option
- RFC 3137：OSPF Stub Router Advertisement

- RFC 4811：OSPF Out-of-Band LSDB Resynchronization
- RFC 4812：OSPF Restart Signaling
- RFC 4813：OSPF Link-Local Signaling

## 1.2 OSPF配置任务简介

无论是哪种类型的路由器，都必须先使能 OSPF，否则 OSPF 协议将无法正常运行。在进行各项配置的时候应该先做好网络规划，错误的配置可能会导致相邻路由器之间无法相互传递信息，甚至导致路由信息的阻塞或者产生路由环路。

表1-1 OSPF 配置任务简介

| 配置任务 | | 说明 | 详细配置 |
|---|---|---|---|
| 使能OSPF功能 | | 必选 | 1.3 |
| 配置OSPF区域 | 配置Stub区域 | 可选 | 1.4.2 |
| | 配置NSSA区域 | 可选 | 1.4.3 |
| | 配置虚连接 | 可选 | 1.4.4 |
| 配置OSPF的网络类型 | 配置OSPF接口网络类型为广播 | 可选 | 1.5.2 |
| | 配置OSPF接口网络类型为NBMA | 可选 | 1.5.3 |
| | 配置OSPF接口网络类型为P2MP | 可选 | 1.5.4 |
| | 配置OSPF接口网络类型为P2P | 可选 | 1.5.5 |
| 配置OSPF的路由信息控制 | 配置OSPF路由聚合 | 可选 | 1.6.2 |
| | 配置OSPF对通过接收到的LSA计算出来的路由信息进行过滤 | 可选 | 1.6.3 |
| | 配置过滤Type-3 LSA | 可选 | 1.6.4 |
| | 配置OSPF接口的开销值 | 可选 | 1.6.5 |
| | 配置OSPF最大等价路由条数 | 可选 | 1.6.6 |
| | 配置OSPF协议的优先级 | 可选 | 1.6.7 |
| | 配置OSPF引入外部路由 | 可选 | 1.6.8 |
| | 配置发布一条主机路由 | 可选 | 1.6.9 |
| 调整和优化OSPF网络 | 配置OSPF报文定时器 | 可选 | 1.7.2 |
| | 配置接口传送LSA的延迟时间 | 可选 | 1.7.3 |
| | 配置OSPF路由计算的时间间隔 | 可选 | 1.7.4 |
| | 配置LSA重复到达的最小时间间隔 | 可选 | 1.7.5 |
| | 配置LSA重新生成的时间间隔 | 可选 | 1.7.6 |
| | 禁止接口收发OSPF报文 | 可选 | 1.7.7 |
| | 配置Stub路由器 | 可选 | 1.7.8 |
| | 配置OSPF验证 | 可选 | 1.7.9 |

| 配置任务 | | 说明 | 详细配置 |
| --- | --- | --- | --- |
| | 配置DD报文中的MTU | 可选 | 1.7.10 |
| | 配置报文的DSCP值 | 可选 | 1.7.11 |
| | 配置LSDB中External LSA的最大数量 | 可选 | 1.7.12 |
| | 配置OSPF尝试退出overflow状态的定时器时间间隔 | 可选 | 1.7.13 |
| | 配置兼容RFC1583的外部路由选择规则 | 可选 | 1.7.14 |
| | 配置邻居状态变化的输出开关 | 可选 | 1.7.15 |
| | 配置OSPF网管功能 | 可选 | 1.7.16 |
| | 配置接口发送LSU报文的时间间隔和一次发送LSU报文的最大个数 | 可选 | 1.7.17 |
| | 配置ISPF | 可选 | 1.7.18 |
| | 配置前缀抑制 | 可选 | 1.7.19 |
| | 配置前缀按优先权收敛 | 可选 | 1.7.20 |
| | 配置PIC | 可选 | 1.7.21 |
| | 配置OSPF的日志信息个数 | 可选 | 1.7.22 |
| 配置OSPF GR | 配置GR Restarter | 可选 | 1.8.1 |
| | 配置GR Helper | 可选 | 1.8.2 |
| | 重启OSPF GR进程 | 可选 | 1.8.3 |
| 配置OSPF NSR | | 可选 | 1.9 |
| 配置OSPF与BFD联动 | | 可选 | 1.10 |
| 配置OSPF快速重路由 | | 可选 | 1.11 |

## 1.3 使能OSPF功能

在 OSPF 的各项配置任务中，必须先使能 OSPF 功能，其它功能的配置才能生效。

### 1.3.1 配置准备

在使能 OSPF 功能之前，需完成以下任务：

- 配置链路层协议，保证链路层通信正常
- 配置接口的网络层地址，使各相邻节点网络层可达

### 1.3.2 使能OSPF功能

要在路由器上使能 OSPF 功能，必须先创建 OSPF 进程、指定该进程关联的区域以及区域包括的网段；对于当前路由器来说，如果某个路由器的接口 IP 地址落在某个区域的网段内，则该接口属于这个区域并使能了 OSPF 功能，OSPF 将把这个接口的直连路由宣告出去。

Router ID 用来在一个自治系统中唯一地标识一台路由器，一台路由器如果要运行 OSPF 协议，则必须存在 Router ID。

- 用户可以在创建 OSPF 进程的时候指定 Router ID，配置时，必须保证自治系统中任意两台路由器的 ID 都不相同。通常的做法是将路由器的 ID 配置为与该路由器某个接口的 IP 地址一致。
- 如果在创建 OSPF 进程的时候没有指定 Router ID，则缺省使用全局 Router ID。建议用户在创建 OSPF 进程的时候指定 Router ID。

目前，系统支持 OSPF 多进程和 OSPF 多实例：

- 当在一台路由器上启动多个 OSPF 进程时，需要指定不同的进程号。OSPF 进程号是本地概念，不影响与其它路由器之间的报文交换。因此，不同的路由器之间，即使进程号不同也可以进行报文交换。
- 可以指定 OSPF 进程所属的 VPN。如果未指定 VPN，则表示 OSPF 位于公网中。VPN 的相关内容请参见“MPLS 配置指导”中的“MPLS L3VPN”。

提示

- 接口配置优先，接口使能 OSPF 优于命令 **netwok** 的配置。
- 接口使能 OSPF 时，如果不存在进程和区域，则创建对应的进程和区域；接口去使能 OSPF 时，不删除已经创建的进程和区域。.

### 1. 在指定网段上使能OSPF

表1-2 使能 OSPF 功能

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| （可选）配置全局 Router ID | **router id** *router-id* | 缺省情况下，未配置全局Router ID<br>如果没有配置全局路由器ID，则按照下面的规则进行选择：<br>(1) 如果存在配置 IP 地址的 Loopback 接口，则选择 Loopback 接口地址中最大的作为 Router ID<br>(2) 如果没有配置 IP 地址的 Loopback 接口，则从其他接口的 IP 地址中选择最大的作为 Router ID（不考虑接口的 up/down 状态） |
| 启动OSPF，并进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | 缺省情况下，系统没有运行OSPF |
| （可选）配置OSPF进程描述 | **description** *description* | 缺省情况下，没有配置进程描述<br>建议用户为每个OSPF进程配置进程描述信息，帮助识别进程的用途，以便于记忆和管理 |
| 配置OSPF区域，进入OSPF区域视图 | **area** *area-id* | 缺省情况下，没有配置OSPF区域 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| （可选）配置区域描述 | **description** *description* | 缺省情况下，没有配置区域描述<br>建议用户为每个区域配置区域描述信息，帮助识别区域的用途，以便于记忆和管理 |
| 配置区域所包含的网段并在指定网段的接口上使能OSPF | **network** *ip-address wildcard-mask* | 缺省情况下，接口不属于任何区域且OSPF功能处于关闭状态<br>一个网段只能属于一个区域 |

### 2. 在指定接口上使能OSPF

表1-3 在指定接口上使能 OSPF

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口使能OSPF | **ospf** *process-id* **area** *area-id* [ **exclude-subip** ] | 缺省情况下，未配置接口使能OSPF |

## 1.4 配置OSPF区域

网络管理员对整个网络划分区域完毕后，可以根据组网需要进一步将区域配置成 Stub 区域或 NSSA 区域。

当非骨干区域不能与骨干区域保持连通，或者骨干区域因为各方面条件的限制无法保持连通时，可以通过配置 OSPF 虚连接予以解决。

### 1.4.1 配置准备

在配置 OSPF 的区域之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点网络层可达
- 使能 OSPF 功能

### 1.4.2 配置Stub区域

对于位于 AS 边缘的一些非骨干区域，我们可以在该区域的所有路由器上配置 **stub** 命令，把该区域配置为 Stub 区域。这样，描述自治系统外部路由的 Type-5 LSA 不会在 Stub 区域里泛洪，减小了路由表的规模。ABR 生成一条缺省路由，所有到达自治系统外部的报文都交给 ABR 进行转发。

如果想进一步减少 Stub 区域路由表规模以及路由信息传递的数量，那么在 ABR 上配置 **stub** 命令时指定 **no-summary** 参数，可以将该区域配置为 Totally Stub 区域。这样，自治系统外部路由和区域间的路由信息都不会传递到本区域，所有目的地是自治系统外和区域外的报文都交给 ABR 进行转发。

Stub 区域和 Totally Stub 区域内不能存在 ASBR，即自治系统外部的路由不能在本区域内传播。

表1-4 配置 Stub 区域

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPF区域视图 | **area** *area-id* | - |
| 配置当前区域为Stub区域 | **stub** [ **default-route-advertise-always** \| **no-summary** ] * | 缺省情况下，没有区域被设置为Stub区域 |
| （可选）配置ABR发送到Stub区域缺省路由的开销 | **default-cost** *cost* | 缺省情况下，ABR发送到Stub区域缺省路由的开销为1<br>本命令只有在Stub区域和Totally Stub区域的ABR上配置才能生效 |

配置时需要注意以下几点：

- 骨干区域不能配置成 Stub 区域或 Totally Stub 区域。
- 如果要将一个区域配置成 Stub 区域，则该区域中的所有路由器必须都要配置 **stub** 命令。
- 如果要将一个区域配置成 Totally Stub 区域，该区域中的所有路由器必须配置 **stub** 命令，该区域的 ABR 路由器需要配置 **stub no-summary** 命令。

### 1.4.3 配置NSSA区域

Stub 区域不能引入外部路由，为了在允许将自治系统外部路由通告到 OSPF 路由域内部的同时，保持其余部分的 Stub 区域的特征，网络管理员可以将区域配置为 NSSA 区域。NSSA 区域也是位于 AS 边缘的非骨干区域。

配置 **nssa** 命令时指定 **no-summary** 参数可以将该区域配置为 Totally NSSA 区域，该区域的 ABR 不会将区域间的路由信息传递到本区域。

表1-5 配置 NSSA 区域

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPF区域视图 | **area** *area-id* | - |
| 配置当前区域为NSSA区域 | **nssa** [ **default-route-advertise** [ **cost** *cost* \| **nssa-only** \| **route-policy** *route-policy-name* \| **type** *type* ] * \| **no-import-route** \| **no-summary** \| **suppress-fa** \| [ **translate-always** \| **translate-never** ] \| **translator-stability-interval** *value* ] * | 缺省情况下，没有区域被设置为NSSA区域 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| （可选）配置发送到NSSA区域缺省路由的开销 | **default-cost** *cost* | 缺省情况下，发送到NSSA区域的缺省路由的开销为1<br>本命令只有在NSSA区域和Totally NSSA区域的ABR/ASBR上配置才能生效 |

配置时需要注意以下几点：

- 骨干区域不能配置成 NSSA 区域或 Totally NSSA 区域。
- 如果要将一个区域配置成 NSSA 区域，则该区域中的所有路由器必须都要配置 **nssa** 命令。
- 如果要将一个区域配置成 Totally NSSA 区域，该区域中的所有路由器必须配置 **nssa** 命令，该区域的 ABR 路由器需要配置 **nssa no-summary** 命令。

### 1.4.4 配置虚连接

在划分区域之后，非骨干区域之间的 OSPF 路由更新是通过骨干区域来完成交换的。对此，OSPF 要求所有非骨干区域必须与骨干区域保持连通，并且骨干区域自身也要保持连通。

但在实际应用中，可能会因为各方面条件的限制，无法满足这个要求。这时可以通过在 ABR 上配置 OSPF 虚连接予以解决。

虚连接不能穿过 Stub 区域和 Totally Stub 区域；虚连接不能穿过 NSSA 区域和 Totally NSSA 区域。

**表1-6 配置虚连接**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPF区域视图 | **area** *area-id* | - |
| 创建并配置虚连接 | **vlink-peer** *router-id* [ **dead** *seconds* \| **hello** *seconds* \| { { **hmac-md5** \| **md5** } *key-id* { **cipher** *cipher-string* \| **plain** *plain-string* } \| **simple** { **cipher** *cipher-string* \| **plain** *plain-string* } } \| **retransmit** *seconds* \| **trans-delay** *seconds* ] * | 缺省情况下，没有虚链接<br>为使虚连接生效，在虚连接的两端都需配置此命令，并且两端配置的**hello**、**dead**参数必须一致 |

## 1.5 配置OSPF的网络类型

OSPF 的网络类型有四种：广播、NBMA、P2MP 和 P2P。

当接口封装的链路层协议不同时，OSPF 接口网络类型的缺省情况也不同：

- 广播：当接口封装的链路层协议是 Ethernet、FDDI 时，接口网络类型缺省值为广播；
- NBMA：当接口封装的链路层协议是 ATM、帧中继或 X.25 时，接口网络类型缺省值为 NBMA；
- P2P：当接口封装的链路层协议是 PPP、LAPB、HDLC 时，接口网络类型缺省值为 P2P。

用户可以根据需要更改接口的网络类型，例如：

- 当 NBMA 网络通过配置地址映射成为全连通网络时（即网络中任意两台路由器之间都存在一条虚电路而直接可达），可以将网络类型更改为广播，不需要手工配置邻居，简化配置。
- 当广播网络中有部分路由器不支持组播时，可以将网络类型更改为 NBMA。
- NBMA 网络要求必须是全连通的，即网络中任意两台路由器之间都必须有一条虚电路直接可达；如果 NBMA 网络不是全连通而是部分连通时，可以将网络类型更改为 P2MP，达到简化配置、节省网络开销的目的。
- 如果路由器在 NBMA 网络中只有一个对端，也可将接口类型配置为 P2P，节省网络开销。

如果接口配置为广播、NBMA 或者 P2MP 网络类型，只有双方接口在同一网段才能建立邻居关系。

### 1.5.1 配置准备

在配置 OSPF 的网络类型之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点之间网络层可达
- 使能 OSPF 功能

### 1.5.2 配置OSPF接口网络类型为广播

表1-7 配置 OSPF 接口网络类型为广播

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置OSPF接口网络类型为广播 | **ospf network-type broadcast** | 缺省情况下，接口的网络类型根据接口封装的链路层协议而定 |
| （可选）配置OSPF接口的路由器优先级 | **ospf dr-priority** *priority* | 缺省情况下，接口的路由器优先级为1 |

### 1.5.3 配置OSPF接口网络类型为NBMA

把接口类型配置为 NBMA 后，需要进行一些特殊的配置。

由于无法通过广播 Hello 报文的形式动态发现相邻路由器，必须手工为接口指定相邻接口的 IP 地址、该相邻接口是否有选举权等（*dr-priority* 参数的值仅表示路由器是否具有 DR 选举权，为 0 表示不具有 DR 选举权，大于 0 时表示具有 DR 选举权）。

表1-8 配置 OSPF 接口网络类型为 NBMA

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置OSPF接口的网络类型为NBMA | **ospf network-type nbma** | 缺省情况下，接口的网络类型根据物理接口而定 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| （可选）配置OSPF接口的路由器优先级 | **ospf dr-priority** *priority* | 缺省情况下，接口的路由器优先级为1<br>本命令设置的优先级用于实际的DR选举 |
| 退回系统视图 | **quit** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置NBMA网络的邻居 | **peer** *ip-address* [ **dr-priority** *dr-priority* ] | 缺省情况下，没有配置邻居<br>本命令设置的优先级用于表示邻居是否具有选举权。如果在配置邻居时将优先级指定为0，则本地路由器认为该邻居不具备选举权，不向该邻居发送Hello报文，这种配置可以减少在DR和BDR选举过程中网络上的Hello报文数量。但如果本地路由器是DR或BDR，它也会向优先级为0的邻居发送Hello报文，以建立邻接关系 |

### 1.5.4 配置OSPF接口网络类型为P2MP

表1-9 配置 OSPF 接口网络类型为 P2MP

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置OSPF接口的网络类型为P2MP | **ospf network-type p2mp** [ **unicast** ] | 缺省情况下，接口的网络类型根据物理接口而定<br>当把接口类型配置为P2MP单播后，OSPF协议在该接口上发送的报文均为单播报文。由于无法通过广播Hello报文的形式动态发现相邻路由器，必须手工为接口指定相邻接口的IP地址 |
| 退回系统视图 | **quit** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| （可选）配置P2MP单播网络的邻居 | **peer** *ip-address* [ **cost** *value* ] | 缺省情况下，没有配置邻居<br>如果接口类型为P2MP单播，必选 |

### 1.5.5 配置OSPF接口网络类型为P2P

表1-10 配置 OSPF 接口网络类型为 P2P

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置OSPF接口的网络类型为P2P | **ospf network-type p2p [ peer-address-check ]** | 缺省情况下，接口的网络类型根据物理接口而定 |

## 1.6 配置OSPF的路由信息控制

通过本节的配置，可以控制 OSPF 的路由信息的发布与接收，并引入其他协议的路由。

### 1.6.1 配置准备

在配置 OSPF 路由信息控制之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点之间网络层可达
- 使能 OSPF 功能
- 如果对路由信息进行过滤，则需要配置对应的过滤列表

### 1.6.2 配置OSPF路由聚合

路由聚合是指 ABR 或 ASBR 将具有相同前缀的路由信息聚合，只发布一条路由到其它区域。

AS 被划分成不同的区域后，每一个区域通过 OSPF 边界路由器（ABR）相连，区域间可以通过路由聚合来减少路由信息，减小路由表的规模，提高路由器的运算速度。

ABR 在计算出一个区域的区域内路由之后，根据聚合相关设置，将其中多条 OSPF 路由聚合成一条发送到区域之外。例如，某个区域内有三条区域内路由 19.1.1.0/24，19.1.2.0/24，19.1.3.0/24，如果在 ABR 上配置了路由聚合，将三条路由聚合成一条 19.1.0.0/16，则 ABR 就只生成一条聚合后的 LSA，并发布给其它区域的路由器。

#### 1. 配置区域边界路由器（ABR）路由聚合

如果区域里存在一些连续的网段，则可以在 ABR 上配置路由聚合，将这些连续的网段聚合成一个网段，ABR 向其它区域发送路由信息时，以网段为单位生成 Type-3 LSA。

这样 ABR 只发送一条聚合后的 LSA，所有属于聚合网段范围的 LSA 将不再会被单独发送出去，既可以减少其它区域中 LSDB 的规模，也减小了因为网络拓扑变化带来的影响。

表1-11 配置 ABR 路由聚合

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPF区域视图 | **area** *area-id* | - |
| 配置OSPF的ABR路由聚合 | **abr-summary** *ip-address* { *mask-length* \| *mask* } [ **advertise** \| **not-advertise** ] [ **cost** *cost* ] | 缺省情况下，ABR不会对路由进行聚合<br>此命令只有在ABR上配置才会有效 |

### 2. 配置自治系统边界路由器（ASBR）对引入的路由进行聚合

ASBR 引入外部路由后，每一条路由都会放在单独的一条 ASE LSA 中向外宣告；通过配置路由聚合，路由器只把聚合后的路由放在 ASE LSA 中向外宣告，减少了 LSDB 中 LSA 的数量。

在 ASBR 上配置路由聚合后，将对聚合地址范围内的 Type-5 LSA 进行聚合。

如果 ASBR 在 NSSA 区域里面，将对聚合地址范围内的 Type-7 LSA 进行聚合，当本地路由器同时是 ASBR 和 ABR 时，将对由 Type-7 LSA 转化成的 Type-5 LSA 进行聚合处理。

表1-12 配置 ASBR 路由聚合

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF的ASBR路由聚合 | **asbr-summary** *ip-address* { *mask-length* \| *mask* } [ **cost** *cost* \| **not-advertise** \| **nssa-only** \| **tag** *tag* ] * | 缺省情况下，ASBR不会对引入的路由进行聚合<br>此命令只有在ASBR上配置才会有效 |

## 1.6.3 配置OSPF对通过接收到的LSA计算出来的路由信息进行过滤

OSPF 是基于链路状态的动态路由协议，路由信息是根据接收到的 LSA 计算出来的，可以对通过接收到的 LSA 计算出来的 OSPF 路由信息进行过滤。

一共有四种过滤方式：

- 基于要加入到路由表的路由信息的目的地址进行过滤，可以通过配置访问控制列表或 IP 地址前缀列表来指定过滤条件；
- 基于要加入到路由表的路由信息的下一跳进行过滤，可以通过在命令中配置 gateway 参数来指定过滤条件；
- 基于要加入到路由表的路由信息的目的地址和下一跳进行过滤，可以通过配置访问控制列表或 IP 地址前缀列表指定过滤目的地址的条件，同时配置 gateway 参数来指定过滤下一跳的条件；
- 基于路由策略对要加入到路由表的路由信息进行过滤，可以通过在命令中配置 route-policy 参数来指定过滤条件。

表1-13 配置 OSPF 对通过接收到的 LSA 计算出来的路由信息进行过滤

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF对通过接收到的LSA计算出来的路由信息进行过滤 | **filter-policy** { *acl-number* [ **gateway** *prefix-list-name* ] \| **gateway** *prefix-list-name* \| **prefix-list** *prefix-list-name* [ **gateway** *prefix-list-name* ] \| **route-policy** *route-policy-name* } **import** | 缺省情况下，OSPF不对通过接收到的LSA计算出来的路由信息进行过滤 |

### 1.6.4 配置过滤Type-3 LSA

通过在 ABR 上配置 Type-3 LSA 过滤，可以对进入 ABR 所在区域或 ABR 向其它区域发布的 Type-3 LSA 进行过滤。

表1-14 配置过滤 Type-3 LSA

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPF区域视图 | **area** *area-id* | - |
| 配置对Type-3 LSA进行过滤 | **filter** { *acl-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } { **export** \| **import** } | 缺省情况下，不对Type-3 LSA进行过滤 |

### 1.6.5 配置OSPF接口的开销值

OSPF 有两种方式来配置接口的开销值：

- 在接口视图下直接配置开销值；
- 配置接口的带宽参考值，OSPF 根据带宽参考值自动计算接口的开销值，计算公式为：接口开销＝带宽参考值÷接口期望带宽（接口期望带宽通过命令 **bandwidth** 进行配置，具体情况请参见接口分册命令参考中的介绍）。当计算出来的开销值大于 65535 时，开销取最大值 65535；当计算出来的开销值小于 1 时，开销取最小值 1。

如果没有在接口视图下配置此接口的开销值，OSPF 会根据该接口的带宽自动计算其开销值。

### 1. 配置接口的开销值

表1-15 配置 OSPF 接口的开销值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 设置OSPF接口的开销值 | **ospf cost** *value* | 缺省情况下，接口按照当前的带宽自动计算接口运行OSPF协议所需的开销。对于Loopback接口，缺省值为0 |

### 2. 配置带宽参考值

表1-16 配置带宽参考值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *instance-name* ] * | - |
| 配置带宽参考值 | **bandwidth-reference** *value* | 缺省情况下，带宽参考值为100Mbps |

## 1.6.6 配置OSPF最大等价路由条数

如果到一个目的地有几条开销相同的路径，可以实现等价路由负载分担，IP 报文在这几个链路上负载分担，以提高链路利用率。该配置用以设置 OSPF 协议的最大等价路由条数。

表1-17 配置 OSPF 最大等价路由条数

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF最大等价路由条数 | **maximum load-balancing** *maximum* | 缺省情况下，OSPF支持的最大等价路由条数为32。 |

## 1.6.7 配置OSPF协议的优先级

由于路由器上可能同时运行多个动态路由协议，就存在各个路由协议之间路由信息共享和选择的问题。系统为每一种路由协议设置一个优先级，在不同协议发现同一条路由时，优先级高的路由将被优先选择。

表1-18 配置 OSPF 协议的优先级

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF协议的路由优先级 | **preference** [ **ase** ] [ **route-policy** *route-policy-name* ] *value* | 缺省情况下，OSPF内部路由的优先级为10，OSPF外部路由的优先级为150 |

## 1.6.8 配置OSPF引入外部路由

### 1. 配置OSPF引入其它协议的路由

提示

**import-route bgp** 命令表示只引入 EBGP 路由；**import-route bgp allow-ibgp** 命令表示将 IBGP 路由也引入，容易引起路由环路，请慎用。

如果在路由器上不仅运行 OSPF，还运行着其它路由协议，可以配置 OSPF 引入其它协议生成的路由，如 RIP、IS-IS、BGP、静态路由或者直连路由，将这些路由信息通过 Type5 LSA 或 Type7 LSA 向外宣告。

OSPF 还可以对引入的路由进行过滤，只将满足过滤条件的外部路由转换为 Type5 LSA 或 Type7 LSA 发布出去。

表1-19 配置 OSPF 引入其它协议的路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF引入其它协议的路由 | **import-route** *protocol* [ *process-id* \| **all-processes** \| **allow-ibgp** ] [ **allow-direct** \| **cost** *cost* \| **nssa-only** \| **route-policy** *route-policy-name* \| **tag** *tag* \| **type** *type* ] * | 缺省情况下，不引入其他协议的路由信息<br>只能引入路由表中状态为active的路由，是否为active状态可以通过**display ip routing-table protocol**命令来查看 |
| （可选）配置对引入的路由进行过滤 | **filter-policy** { *acl-number* \| **prefix-list** *prefix-list-name* } **export** [ *protocol* [ *process-id* ] ] | 缺省情况下，不对引入的路由信息进行过滤 |

### 2. 配置OSPF引入缺省路由

OSPF 不能通过 **import-route** 命令从其它协议引入缺省路由，如果想把缺省路由引入到 OSPF 路由区域，必须要使用下面命令配置 OSPF 引入缺省路由。

表1-20 配置 OSPF 引入缺省路由

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF引入缺省路由 | **default-route-advertise** [ [ [ **always** \| **permit-calculate-other** ] \| **cost** *cost* \| **route-policy** *route-policy-name* \| **type** *type* ] * \| **summary cost** *cost* ] | 缺省情况下，不引入缺省路由<br>**default-route-advertise summary cost**命令仅在VPN中应用，以Type-3 LSA引入缺省路由，PE路由器会将引入的缺省路由发布给CE路由器 |

### 3. 配置引入路由的相关参数

当 OSPF 引入外部路由时，还可以配置一些开销、路由数量、标记和类型等参数的缺省值。路由标记可以用来标识协议相关的信息，如 OSPF 从 BGP 引入路由时，可以用来标记自治系统的编号。

表1-21 配置引入路由时的相关参数

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置引入外部路由时的参数缺省值（开销、路由数量、标记、类型） | **default** { **cost** *cost* \| **tag** *tag* \| **type** *type* } * | 缺省情况下，OSPF引入的外部路由的度量值为1，引入的外部路由的标记为1，引入的外部路由类型为2 |

## 1.6.9 配置发布一条主机路由

表1-22 配置发布一条主机路由

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPF区域视图 | **area** *area-id* | - |
| 配置并发布一条主机路由 | **host-advertise** *ip-address cost* | 缺省情况下，OSPF不发布所包含网段之外的主机路由 |

## 1.7 调整和优化OSPF网络

用户可以从以下几个方面来调整和优化 OSPF 网络：

- 通过改变 OSPF 的报文定时器，可以调整 OSPF 网络的收敛速度以及协议报文带来的网络负荷。在一些低速链路上，需要考虑接口传送 LSA 的延迟时间。
- 通过调整 SPF 计算间隔时间，可以抑制由于网络频繁变化带来的资源消耗问题。
- 在安全性要求较高的网络中，可以通过配置 OSPF 验证特性，来提高 OSPF 网络的安全性。

### 1.7.1 配置准备

在调整和优化 OSPF 网络之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点之间网络层可达
- 使能 OSPF 功能

### 1.7.2 配置OSPF报文定时器

用户可以在接口上配置下列 OSPF 报文定时器：

- Hello 定时器：接口向邻居发送 Hello 报文的时间间隔，OSPF 邻居之间的 Hello 定时器的值要保持一致。
- Poll 定时器：在 NBMA 网络中，路由器向状态为 down 的邻居路由器发送轮询 Hello 报文的时间间隔。
- 邻居失效时间：在邻居失效时间内，如果接口还没有收到邻居发送的 Hello 报文，路由器就会宣告该邻居无效。
- 接口重传 LSA 的时间间隔：路由器向它的邻居通告一条 LSA 后，需要对方进行确认。若在重传间隔时间内没有收到对方的确认报文，就会向邻居重传这条 LSA。

表1-23 配置 OSPF 报文定时器

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置Hello定时器 | **ospf timer hello** *seconds* | 缺省情况下，P2P、Broadcast类型接口发送Hello报文的时间间隔为10秒，P2MP、NBMA类型接口发送Hello报文的时间间隔为30秒<br>修改了网络类型后，Hello定时器将恢复缺省值 |
| 配置Poll定时器 | **ospf timer poll** *seconds* | 缺省情况下，发送轮询Hello报文的时间间隔为120秒<br>轮询Hello报文的时间间隔至少应为Hello时间间隔的4倍 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置邻居失效时间 | **ospf timer dead** *seconds* | 缺省情况下，P2P、Broadcast类型接口的OSPF邻居失效时间为40秒，P2MP、NBMA类型接口的OSPF邻居失效时间为120秒<br>邻居失效时间应至少为Hello时间间隔的4倍<br>修改了网络类型后，邻居失效时间将恢复缺省值 |
| 配置接口重传LSA的时间间隔 | **ospf timer retransmit** *seconds* | 缺省情况下，时间间隔为5秒<br>相邻路由器重传LSA时间间隔的值不要设置得太小，否则将会引起不必要的重传。通常应该大于一个报文在两台路由器之间传送一个来回的时间 |

### 1.7.3 配置接口传送LSA的延迟时间

考虑到 OSPF 报文在链路上传送时也需要花费时间，所以 LSA 的老化时间（age）在传送之前要增加一定的延迟时间，在低速链路上需要对该项配置进行重点考虑。

**表1-24 配置接口传送 LSA 的延迟时间**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口传送LSA的延迟时间 | **ospf trans-delay** *seconds* | 缺省情况下，接口传送LSA的延迟时间为1秒 |

### 1.7.4 配置OSPF路由计算的时间间隔

当 OSPF 的 LSDB 发生改变时，需要重新计算最短路径。如果网络频繁变化，且每次变化都立即计算最短路径，将会占用大量系统资源，并影响路由器的效率。通过调节路由计算的时间间隔，可以抑制由于网络频繁变化带来的影响。

本命令在网络变化不频繁的情况下将连续路由计算的时间间隔缩小到 *minimum-interval*，而在网络变化频繁的情况下可以进行相应惩罚，增加 *incremental-interval*×$2^{n-2}$（n 为连续触发路由计算的次数），将等待时间按照配置的惩罚增量延长，最大不超过 *maximum-interval*。

**表1-25 配置 SPF 计算时间间隔**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置OSPF路由计算的时间间隔 | **spf-schedule-interval** *maximum-interval* [ *minimum-interval* [ *incremental-interval* ] ] | 缺省情况下，OSPF路由计算的最大时间间隔为5秒，最小时间间隔为50毫秒，时间间隔惩罚增量为200毫秒 |

### 1.7.5 配置LSA重复到达的最小时间间隔

如果在重复到达的最小时间间隔内连续收到一条 LSA 类型、LS ID、生成路由器 ID 均相同的 LSA 则直接丢弃，这样就可以抑制网络频繁变化可能导致的占用过多带宽资源和路由器资源。

**表1-26 配置 LSA 的重复接收最小间隔**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置LSA重复到达的最小时间间隔 | **lsa-arrival-interval** *interval* | 缺省情况下，LSA重复到达的最小时间间隔为1000毫秒 |

说明

建议 **lsa-arrival-interval** 命令配置的 *interval* 小于或等于 **lsa-generation-interval** 命令所配置的 *minimum-interval*。

### 1.7.6 配置LSA重新生成的时间间隔

通过调节 LSA 重新生成的时间间隔，可以抑制网络频繁变化可能导致的占用过多带宽资源和路由器资源。

本命令在网络变化不频繁的情况下将 LSA 重新生成时间间隔缩小到 *minimum-interval*，而在网络变化频繁的情况下可以进行相应惩罚，增加 *incremental-interval*×$2^{n-2}$（n 为连续触发路由计算的次数），将等待时间按照配置的惩罚增量延长，最大不超过 *maximum-interval*。

**表1-27 配置 LSA 发送间隔**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置LSA重新生成的时间间隔 | **lsa-generation-interval** *maximum-interval* [ *minimum-interval* [ *incremental-interval* ] ] | 缺省情况下，最大时间间隔为5秒，最小时间间隔为50毫秒，惩罚增量为200毫秒 |

### 1.7.7 禁止接口收发OSPF报文

如果要使 OSPF 路由信息不被某一网络中的路由器获得，可以禁止接口收发 OSPF 报文。

将运行 OSPF 协议的接口指定为 Silent 状态后，该接口的直连路由仍可以由同一路由器的其它接口通过 Router-LSA 发布出去，但 OSPF 报文将被阻塞，接口上无法建立邻居关系。这样可以增强 OSPF 的组网适应能力，减少系统资源的消耗。

表1-28 禁止接口收发 OSPF 报文

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 禁止接口收发OSPF报文 | **silent-interface** { *interface-type interface-number* \| **all** } | 缺省情况下，允许接口收发OSPF报文<br>不同的进程可以对同一接口禁止收发OSPF报文，但本命令只对本进程已经使能的OSPF接口起作用，对其它进程的接口不起作用 |

### 1.7.8 配置Stub路由器

Stub 路由器用来控制流量，它告知其他 OSPF 路由器不要使用这个 Stub 路由器来转发数据，但可以拥有一个到 Stub 路由器的路由。

通过将当前路由器配置为 Stub 路由器，在该路由器发布的 Router-LSA 中，当链路类型取值为 3 表示连接到 Stub 网络时，链路度量值不变；当链路类型为 1、2、4 分别表示通过 P2P 链路与另一路由器相连、连接到传送网络、虚连接时，链路度量值将设置为最大值 65535。通过增加 **include-stub** 参数可以将路由器发布的 Router-LSA 中，链路类型为 3 的 Stub 链路度量值设置为最大值 65535。这样其邻居计算出这条路由的开销就会很大，如果邻居上有到这个目的地址开销更小的路由，则数据不会通过这个 Stub 路由器转发。

表1-29 配置 Stub 路由器

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置当前路由器为Stub路由器 | **stub-router** [ **external-lsa** [ *max-metric-value* ] \| **include-stub** \| **on-startup** { *seconds* \| **wait-for-bgp** [ *seconds* ] } \| **summary-lsa** [ *max-metric-value* ] ] * | 缺省情况下，当前路由器没有被配置为Stub路由器<br>Stub路由器与Stub区域无关 |

## 1.7.9 配置OSPF验证

从安全性角度来考虑，为了避免路由信息外泄或者 OSPF 路由器受到恶意攻击，OSPF 提供报文验证功能。

OSPF 路由器建立邻居关系时，在发送的报文中会携带配置好的口令，接收报文时进行密码验证，只有通过验证的报文才能接收，否则将不会接收报文，不能正常建立邻居。

如果区域验证和接口验证都进行了配置，以接口验证的配置为准。

### 1. 配置区域验证

一个区域中所有路由器的验证模式和验证密码必须一致。

表1-30 配置区域验证

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPF区域视图 | **area** *area-id* | - |
| 配置OSPF区域的验证模式 | **authentication-mode** { **hmac-md5** \| **md5** } *key-id* { **cipher** \| **plain** } *password* | 二者选其一<br>缺省情况下，没有配置区域验证模式 |
| | **authentication-mode simple** { **cipher** \| **plain** } *password* | |

### 2. 配置接口验证

邻居路由器两端接口的验证模式和验证密码必须一致。

表1-31 配置接口验证

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置OSPF接口的验证模式 | **ospf authentication-mode simple** { **cipher** *cipher-string* \| **plain** *plain-string* } | 二者选其一<br>缺省情况下，接口不对OSPF报文进行验证 |
| | **ospf authentication-mode** { **hmac-md5** \| **md5** } *key-id* { **cipher** *cipher-string* \| **plain** *plain-string* } | |

## 1.7.10 配置DD报文中的MTU

一般情况下，接口发送 DD 报文时不使用接口的实际 MTU 值，而是用 0 代替。进行此配置后，将使用接口的实际 MTU 值填写 DD 报文 Interface MTU 字段。

表1-32 配置 DD 报文中的 MTU

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置DD报文中MTU域的值为发送该报文接口的MTU值 | **ospf mtu-enable** | 缺省情况下，接口发送的DD报文中MTU域的值为0 |

### 1.7.11 配置OSPF发送协议报文的DSCP优先级

表1-33 配置 OSPF 发送协议报文的 DSCP 优先级

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF发送协议报文的DSCP优先级 | **dscp** *dscp-value* | 缺省情况下，OSPF发送协议报文的DSCP优先级值为48 |

### 1.7.12 配置LSDB中External LSA的最大数量

表1-34 配置 LSDB 中 External LSA 的最大数量

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置LSDB中External LSA的最大数量 | **lsdb-overflow-limit** *number* | 缺省情况下，不对LSDB中External LSA的最大条目数进行限制 |

### 1.7.13 配置OSPF尝试退出overflow状态的定时器时间间隔

网络中出现过多 LSA，会占用大量系统资源。当设置的 LSDB 中 External LSA 的最大数量达到上限时，LSDB 会进入 overflow 状态，在 overflow 状态中，不再接收 External LSA，同时删除自己生成的 External LSA，对于已经收到的 External LSA 则不会删除。这样就可以减少 LSA 从而节省系统资源。

通过配置可以调整 OSPF 退出 overflow 状态的时间。

表1-35 配置 OSPF 尝试退出 overflow 状态的定时器时间间隔

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF尝试退出overflow状态的定时器时间间隔 | **lsdb-overflow-interval** *interval* | 缺省情况下，OSPF尝试退出overflow定时器间隔是300秒，配置为0时，表示不退出Overflow状态 |

## 1.7.14 配置兼容RFC 1583 的外部路由选择规则

当有多条路径可以到达同一个外部路由时，在选择最优路由的问题上，RFC 2328 中定义的选路规则与 RFC 1583 的有所不同，进行此配置可以兼容 RFC 1583 中定义的规则。

具体的选路规则如下：

(1) 当 RFC 2328 兼容 RFC 1583 时，所有到达 ASBR 的路由优先级相同。当 RFC 2328 不兼容 RFC 1583 时，非骨干区的区域内路由优先级最高，区域间路由与骨干区区域内路由优先级相同，优选非骨干区的区域内路由，尽量减少骨干区的负担；

(2) 若存在多条优先级相同的路由时，按开销值优选，优先开销值小的路由；

(3) 若存在多条开销值相同路由时，按路由来源区域的区域 ID 选择，优选区域 ID 大的路由。

为了避免路由环路，同一路由域内的路由器建议统一配置相同选择规则。

**表1-36 配置兼容 RFC 1583 的外部路由选择规则**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置兼容RFC 1583的外部路由选择规则 | **rfc1583 compatible** | 缺省情况下，使能兼容RFC 1583的选路规则 |

## 1.7.15 配置邻居状态变化的输出开关

打开邻居状态变化的输出开关后，OSPF 邻居状态变化时会生成日志信息发送到设备的信息中心，通过设置信息中心的参数，最终决定日志信息的输出规则（即是否允许输出以及输出方向）。（有关信息中心参数的配置请参见“网络管理和监控配置指导”中的“信息中心”。）

**表1-37 配置邻居状态变化的输出开关**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 打开邻居状态变化的输出开关 | **log-peer-change** | 缺省情况下，邻居状态变化的输出开关处于打开状态 |

### 1.7.16 配置OSPF网管功能

配置 OSPF 进程绑定 MIB 功能后，可以通过网管软件对指定的 OSPF 进程进行管理。

开启 OSPF 模块的告警功能后，该模块会生成告警信息，用于报告该模块的重要事件。生成的告警信息将发送到设备的 SNMP 模块，通过设置 SNMP 中告警信息的发送参数，来决定告警信息输出的相关属性。（有关告警信息的详细介绍，请参见“网络管理和监控配置指导”中的“SNMP”。）

通过调整 OSPF 在指定时间间隔内允许输出的告警信息条数，可以避免网络出现大量告警信息时对资源的消耗。

表1-38 配置 OSPF 网管功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置OSPF进程绑定MIB | **ospf mib-binding** *process-id* | 缺省情况下，MIB绑定在进程号最小的OSPF进程上 |
| 开启OSPF的告警功能 | **snmp-agent trap enable ospf** [ **authentication-failure** \| **bad-packet** \| **config-error** \| **grhelper-status-change** \| **grrestarter-status-change** \| **if-state-change** \| **lsa-maxage** \| **lsa-originate** \| **lsdb-approaching-overflow** \| **lsdb-overflow** \| **neighbor-state-change** \| **nssatranslator-status-change** \| **retransmit** \| **virt-authentication-failure** \| **virt-bad-packet** \| **virt-config-error** \| **virt-retransmit** \| **virtgrhelper-status-change** \| **virtif-state-change** \| **virtneighbor-state-change** ] * | 缺省情况下，OSPF的告警功能处于开启状态 |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF在指定时间间隔内允许输出的告警信息条数 | **snmp trap rate-limit interval** *trap-interval* **count** *trap-number* | 缺省情况下，OSPF模块的10秒允许输出7条告警信息 |

### 1.7.17 配置接口发送LSU报文的时间间隔和一次发送LSU报文的最大个数

如果路由器路由表里的路由条目很多，在与邻居进行 LSDB 同步时，可能需要发送大量 LSU，有可能会对当前设备和网络带宽带来影响；因此，路由器将 LSU 报文分为多个批次进行发送，并且对 OSPF 接口每次允许发送的 LSU 报文的最大个数做出限制。

用户可根据需要配置OSPF接口发送LSU报文的时间间隔以及接口一次发送LSU报文的最大个数。

表1-39 配置接口发送 LSU 报文的时间间隔和一次发送 LSU 报文的最大个数

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置接口发送LSU报文的时间间隔和一次发送LSU报文的最大个数 | **transmit-pacing interval** *interval* **count** *count* | 缺省情况下，OSPF接口发送LSU报文的时间间隔为20毫秒，一次最多发送3个LSU报文 |

## 1.7.18 配置ISPF

ISPF（Incremental Shortest Path First，增量最短路径优先）是对 OSPF 中最短路径树的增量计算，当网络的拓扑结构发生变化，即影响到最短路径树的结构时，只对受影响的部分节点进行重新计算拓扑结构，只对最短路径树中受影响的部分进行修正，而不需要重建整棵最短路径树。

**表1-40 配置增量 SPF 计算**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能增量SPF计算功能 | **ispf enable** | 缺省情况下，使能增量SPF计算功能 |

## 1.7.19 配置前缀抑制

OSPF 使能网段时会将接口上匹配该网段的所有网段路由与主机路由都通过 LSA 发布，但有些时候主机路由或网段路由是不希望被发布的。通过前缀抑制配置，可以减少 LSA 中携带不需要的前缀，即不发布某些网段路由和主机路由，从而提高网络安全性，加快路由收敛。

当使能前缀抑制时，具体情况如下：

- P2P 或 P2MP 类型网络：Router LSA 中不发布接口的主地址，即 Router LSA 中链路类型为 3 的 Stub 链路被抑制，不生成接口路由，但其他路由信息可以正常计算，不会影响流量转发。
- 广播类型或者 NBMA 网络：DR 发布的 Network LSA 的掩码字段会填成 32 位，即不生成网段路由，但其他路由信息可以正常计算，不会影响流量转发。另外，如果没有邻居，发布的 Router LSA 中也不发布接口的主地址，即 Router LSA 中链路类型为 3 的 Stub 链路被抑制。

提示

如果需要抑制前缀发布，建议整个 OSPF 网络都配置本命令。

### 1. 配置全局前缀抑制

全局配置不能抑制从地址、Loopback 接口以及处于抑制状态的接口对应的前缀。如果想对 Loopback 接口或处于抑制状态的接口进行抑制，可以通过配置接口前缀抑制来实现。

表1-41 配置全局前缀抑制

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置前缀抑制功能 | **prefix-suppression** | 缺省情况下，不抑制OSPF进程进行前缀发布 |

### 2. 配置接口前缀抑制

接口配置不抑制从地址对应的前缀。

表1-42 配置接口前缀抑制

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的前缀抑制功能 | **ospf prefix-suppression [ disable ]** | 缺省情况下，不抑制接口进行前缀发布 |

## 1.7.20 配置OSPF的前缀按优先权收敛功能

通过策略指定优先权，不同前缀按优先权顺序下发，由高到低分为 4 个优先权（Critical、High、Medium 和 Low），如果一条路由符合多个收敛优先权的匹配规则，则这些收敛优先权中最高者当选为路由的收敛优先权。

OSPF 路由的 32 位主机路由为 Medium 优先权，其它为 Low 优先权。

表1-43 配置 OSPF 的前缀按优先权收敛功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能OSPF的前缀按优先权快速收敛功能 | **prefix-priority route-policy** *route-policy-name* | 缺省情况下，OSPF的前缀按优先权快速收敛功能处于关闭状态 |

## 1.7.21 配置PIC

提示

- PIC 和 OSPF 快速重路由功能同时配置时，OSPF 快速重路由功能生效。.
- 目前只支持区域间路由以及外部路由的 PIC 功能。

PIC（Prefix Independent Convergence，前缀无关收敛），即收敛时间与前缀数量无关，加快收敛速度。传统的路由计算快速收敛都与前缀数量相关，收敛时间与前缀数量成正比。

### 1. 配置PIC

表1-44 配置 PIC

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能PIC功能 | **pic** [ **additional-path-always** ] | 缺省情况下，使能PIC功能 |

### 2. 配置PIC支持BFD检测功能

OSPF 协议的 PIC 特性中，主用链路缺省不使用 BFD 进行链路故障检测。配置本功能后，将使用 BFD（Echo 方式）进行检测，可以加快 OSPF 协议的收敛速度。

表1-45 配置 PIC 支持 BFD 检测功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置BFD Echo报文源地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置BFD Echo报文源地址 |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能OSPF协议中主用链路的BFD（Echo方式）检测功能 | **ospf primary-path-detect bfd echo** | 缺省情况下，OSPF协议中主用链路的BFD（Echo方式）检测功能处于关闭状态 |

## 1.7.22 配置OSPF的日志信息个数

OSPF 的日志信息包括路由计算日志信息和邻居日志信息。

表1-46 配置 OSPF 的日志信息个数

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF的日志信息个数 | **event-log** { **peer** \| **spf** } **size** *count* | 缺省情况下，路由计算和邻居的日志信息个数为10 |

# 1.8 配置OSPF GR

说明

设备充当 GR Restarter 后不能再配置 OSPF NSR 功能。

GR（Graceful Restart，平滑重启）是一种通过备份 OSPF 配置信息，在协议重启或主备倒换时 OSPF 进行平滑重启，从邻居那里获得邻居关系，并对 LSDB 进行同步，从而保证转发业务不中断的机制。

GR 有两个角色：

- GR Restarter：发生协议重启或主备倒换事件且具有 GR 能力的设备。
- GR Helper：和 GR Restarter 具有邻居关系，协助完成 GR 流程的设备。

目前有两种方式实现 OSPF GR 技术：

- 一种是基于 IETF 标准，GR Restarter 通过向 GR Helper 发送一种称为 Grace LSA 的 9 类 Opaque LSA 来控制 GR 的交互过程。
- 另外一种是非 IETF 标准，GR Restarter 与 GR Helper 之间是通过相互发送携带 LLS 与 OOB 扩展信息的 OSPF 报文来完成 GR 的交互过程。

一台设备可以同时充当 GR Restarter 和 GR Helper。

## 1.8.1 配置GR Restarter

可以在 GR Restarter 上配置基于 OSPF 的 IETF 标准或非 IETF 标准的 GR 能力。在作为 GR Restarter 的设备上进行如下配置：

### 1. 配置IETF标准GR Restarter

表1-47 配置 IETF 标准 GR Restarter

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 启动OSPF，进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能Opaque LSA发布接收能力 | **opaque-capability enable** | 缺省情况下，OSPF的Opaque LSA发布接收能力处于开启状态 |
| 使能OSPF协议的IETF标准GR能力 | **graceful-restart ietf** [ **global** \| **planned-only** ] * | 缺省情况下，OSPF协议的IETF标准GR能力处于关闭状态 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| （可选）配置OSPF协议的GR重启间隔时间 | **graceful-restart interval** *interval-value* | 缺省情况下，OSPF协议的GR重启间隔时间为120秒 |

### 2. 配置非IETF标准GR Restarter

表1-48 配置非 IETF 标准 GR Restarter

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 启动OSPF，进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能OSPF本地链路信令能力 | **enable link-local-signaling** | 缺省情况下，OSPF本地链路信令能力处于关闭状态 |
| 使能OSPF带外同步能力 | **enable out-of-band-resynchronization** | 缺省情况下，OSPF带外同步能力处于关闭状态 |
| 使能OSPF协议的非IETF标准GR能力 | **graceful-restart** [ **nonstandard** ] [ **global** \| **planned-only** ] * | 缺省情况下，OSPF协议的非IETF标准GR能力处于关闭状态 |
| （可选）配置OSPF协议的GR重启间隔时间 | **graceful-restart interval** *interval-value* | 缺省情况下，OSPF协议的GR重启间隔时间为120秒 |

## 1.8.2 配置GR Helper

可以在作为 GR Helper 的设备上配置基于 OSPF 的 IETF 标准或非 IETF 标准的 GR Helper 能力。在作为 GR Helper 的设备上进行如下配置：

### 1. 配置IETF标准GR Helper

表1-49 配置 IETF 标准 GR Helper

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 启动OSPF，进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能Opaque LSA发布接收能力 | **opaque-capability enable** | 缺省情况下，OSPF的Opaque LSA发布接收能力处于开启状态 |
| （可选）使能GR Helper能力 | **graceful-restart helper enable** [ **planned-only** ] | 缺省情况下，OSPF的GR Helper能力处于开启状态 |
| （可选）配置GR Helper 严格检查LSA能力 | **graceful-restart helper strict-lsa-checking** | 缺省情况下，OSPF协议的GR Helper严格LSA检查能力处于关闭状态 |

### 2. 配置非IETF标准GR Helper

表1-50 配置非 IETF 标准 GR Helper

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 启动OSPF，进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能OSPF本地链路信令能力 | **enable link-local-signaling** | 缺省情况下，OSPF本地链路信令能力处于关闭状态 |
| 使能OSPF带外同步能力 | **enable out-of-band-resynchronization** | 缺省情况下，OSPF带外同步能力处于关闭状态 |
| （可选）使能GR Helper能力 | **graceful-restart helper enable** | 缺省情况下，OSPF的GR Helper能力处于开启状态 |
| （可选） 配置GR Helper 严格检查LSA能力 | **graceful-restart helper strict-lsa-checking** | 缺省情况下，OSPF协议的GR Helper严格LSA检查能力处于关闭状态 |

## 1.8.3 以GR方式重启OSPF进程

设备进行主备倒换或者进行如下操作均可以以 GR 方式重启 OSPF 进程。

表1-51 以 GR 方式重启 OSPF 进程

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 以GR方式重启OSPF进程 | **reset ospf** [ *process-id* ] **process graceful-restart** | 请在用户视图下执行该命令 |

# 1.9 配置OSPF NSR

MSR 系列路由器各款型对于本节所描述的特性的支持情况有所不同，详细差异信息如下：

| 型号 | 特性 | 描述 |
| --- | --- | --- |
| MSR 2600 | 配置OSPF NSR | 不支持 |
| MSR 3600 | | 不支持 |
| MSR 5600 | | 支持 |

说明

设备配置了 OSPF NSR 功能后不能再充当 GR Restarter。

NSR（Nonstop Routing，不间断路由）通过将 OSPF 链路状态信息从主进程备份到备进程，使设备在发生主备倒换时可以自行完成链路状态的恢复和路由的重新生成，邻接关系不会发生中断，从而避免了主备倒换对转发业务的影响。

GR 特性需要周边设备配合才能完成路由信息的恢复，在网络应用中有一定的限制。NSR 特性不需要周边设备的配合，网络应用更加广泛。

表1-52 配置 OSPF NSR

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能OSPF NSR功能 | **non-stop-routing** | 缺省情况下，OSPF NSR功能处于关闭状态 |

## 1.10 配置OSPF与BFD联动

BFD（Bidirectional Forwarding Detection，双向转发检测）能够为 OSPF 邻居之间的链路提供快速检测功能。当邻居之间的链路出现故障时，加快 OSPF 协议的收敛速度。关于 BFD 的介绍和基本功能配置，请参见“可靠性配置指导”中的“BFD”。

OSPF 使用 BFD 来进行快速故障检测时，提供两种检测方式：

- control 报文双向检测：需要建立 OSPF 邻居的两端设备均支持 BFD 配置。
- echo 报文单跳检测：仅需要一端设备支持 BFD 配置。

### 1.10.1 control报文双向检测

表1-53 配置 OSPF 与 BFD 联动（control 报文双向检测）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能OSPF的BFD功能 | **ospf bfd enable** | 缺省情况下，OSPF的BFD功能处于关闭状态<br>创建BFD会话的通信双方必须处于特定区域的同一网段 |

### 1.10.2 echo报文单跳检测

表1-54 配置 OSPF 与 BFD 联动（echo 报文单跳检测）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置echo报文源地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置echo报文源地址 |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能OSPF的BFD功能 | **ospf bfd enable echo** | 缺省情况下，OSPF的BFD功能处于关闭状态 |

## 1.11 配置OSPF快速重路由

提示

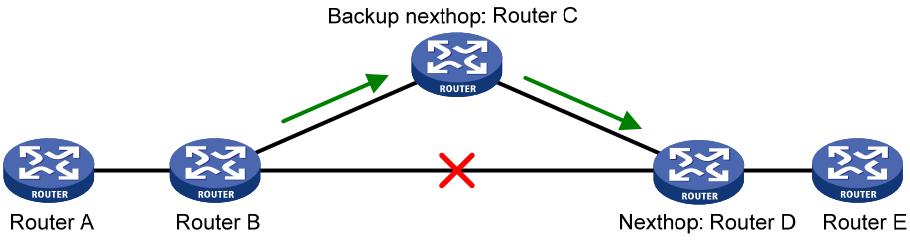
- OSPF 快速重路由功能不能与 OSPF 的 BFD 功能同时使用，否则可能导致快速重路由功能失效。
- OSPF 快速重路由功能（通过 LFA 算法选取备份下一跳信息）不能与 **vlink-peer**、**sham-link**（MPLS 命令参考/MPLS L3VPN）命令同时使用。
- OSPF 快速重路由功能和 PIC 同时配置时，OSPF 快速重路由功能生效。

### 1.11.1 功能简介

当 OSPF 网络中的链路或某台路由器发生故障时，需要通过故障链路或故障路由器传输才能到达目的地的报文将会丢失或产生路由环路，数据流量将会被中断，直到 OSPF 根据新的拓扑网络路由收敛完毕后，被中断的流量才能恢复正常的传输。

为了尽可能缩短网络故障导致的流量中断时间，网络管理员可以根据需要配置 OSPF 快速重路由功能。

图1-7 OSPF 快速重路由功能示意图




如 图 1-7 所示，通过在Router B上使能快速重路由功能，OSPF将为路由计算或指定备份下一跳，当Router B检测到网络故障时，OSPF会使用事先获取的备份下一跳替换失效下一跳，通过备份下一跳来指导报文的转发，从而大大缩短了流量中断时间。在使用备份下一跳指导报文转发的同时，OSPF会根据变化后的网络拓扑重新计算最短路径，网络收敛完毕后，使用新计算出来的最优路由来指导报文转发。

网络管理员可以配置给所有 OSPF 路由通过 LFA（Loop Free Alternate）算法选取备份下一跳，也可以在路由策略中指定备份下一跳，为符合过滤条件的路由指定备份下一跳。

### 1.11.2 配置准备

在配置 OSPF 快速重路由特性之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点网络层可达
- 使能 OSPF 功能

### 1.11.3 配置步骤

#### 1. 配置OSPF支持快速重路由功能

(1) 配置 OSPF 支持快速重路由功能（通过 LFA 算法选取备份下一跳信息）

**表1-55 配置 OSPF 支持快速重路由功能（自动计算备份下一跳）**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| （可选）指定接口上使能参与LFA计算 | **ospf fast-reroute lfa-backup** | 缺省情况下，使能接口参与LFA计算，能够被选为备份接口 |
| 退回系统视图 | **quit** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF支持快速重路由功能（通过LFA算法选取备份下一跳信息） | **fast-reroute lfa** [ **abr-only** ] | 缺省情况下，没有配置OSPF快速重路由功能<br>**abr-only**表示仅选取到ABR设备的路由作为备份下一跳 |

(2) 配置 OSPF 支持快速重路由功能（通过路由策略指定备份下一跳）

网络管理员可以通过 **apply fast-reroute backup-interface** 命令在路由策略中指定备份下一跳，为符合过滤条件的路由指定备份下一跳，关于 **apply fast-reroute backup-interface** 命令以及路由策略的相关配置，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

**表1-56 配置 OSPF 支持快速重路由功能（通过路由策略指定备份下一跳）**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入OSPF视图 | **ospf** [ *process-id* \| **router-id** *router-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPF支持快速重路由功能（通过路由策略指定备份下一跳） | **fast-reroute route-policy** *route-policy-name* | 缺省情况下，没有配置OSPF快速重路由功能 |

#### 2. 配置OSPF快速重路由支持BFD检测功能

OSPF 协议的快速重路由特性中，主用链路缺省不使用 BFD 进行链路故障检测。配置本功能后，将使用 BFD（Echo 方式）进行检测，可以加快 OSPF 协议的收敛速度。

表1-57 配置 OSPF 快速重路由支持 BFD 检测功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置BFD Echo报文源地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置BFD Echo报文源地址 |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能OSPF协议中主用链路的BFD（Echo方式）检测功能 | **ospf primary-path-detect bfd echo** | 缺省情况下，OSPF协议中主用链路的BFD（Echo方式）检测功能处于关闭状态 |

## 1.12 OSPF显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 OSPF 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 OSPF 的统计信息、重启 OSPF 进程或重新向 OSPF 引入外部路由。

表1-58 OSPF 显示和维护

| 操作 | 命令 |
| --- | --- |
| 显示OSPF的进程信息 | **display ospf** [ *process-id* ] [ **verbose** ] |
| 显示OSPF进程的GR状态信息 | **display ospf** [ *process-id* ] **graceful-restart** [ **verbose** ] |
| 显示区域中FRR备份下一跳候选列表 | **display ospf** [ *process-id* ] [ **area** *area-id* ] **fast-reroute lfa-candidate** |
| 显示OSPF的LSDB信息 | **display ospf** [ *process-id* ] **lsdb** [ **area** *area-id* \| **brief** \| [ { **ase** \| **router** \| **network** \| **summary** \| **asbr** \| **nssa** \| **opaque**-link \| **opaque-area** \| **opaque**-as } [ *link-state-id* ] ] [ **originate-router** *advertising-router-id* \| **self-originate** ] ] |
| 显示进程中的下一跳信息 | **display ospf** [ *process-id* ] **nexthop** |
| 显示OSPF邻居的信息 | **display ospf** [ *process-id* ] **peer** [ **verbose** ] [ *interface-type interface-number* ] [ *neighbor-id* ] |
| 显示OSPF各区域邻居的统计信息 | **display ospf** [ *process-id* ] **peer statistics** |
| 显示OSPF路由表的信息 | **display ospf** [ *process-id* ] **routing** [ *ip-address* { *mask-length* \| *mask* } ] [ **interface** *interface-type interface-number* ] [ **nexthop** *nexthop-address* ] [ **verbose** ] |
| 显示OSPF区域中的拓扑信息 | **display ospf** [ *process-id* ] [ **area** *area-id* ] **spf-tree** [ **verbose** ] |
| 显示OSPF的统计信息 | **display ospf** [ *process-id* ] **statistics** [ **error** ] |
| 显示OSPF虚连接信息 | **display ospf** [ *process-id* ] **vlink** |
| 显示OSPF请求列表 | **display ospf** [ *process-id* ] **request-queue** [ *interface-type interface-number* ] [ *neighbor-id* ] |

| 操作 | 命令 |
| --- | --- |
| 显示OSPF重传列表 | **display ospf** [ *process-id* ] **retrans-queue** [ *interface-type interface-number* ] [ *neighbor-id* ] |
| 显示OSPF ABR及ASBR信息 | **display ospf** [ *process-id* ] **abr-asbr** [ **verbose** ] |
| 显示OSPF的ABR聚合信息 | **display ospf** [ *process-id* ] [ **area** *area-id* ] **abr-summary** [ *ip-address* { *mask-length* \| *mask* } ] [ **verbose** ] |
| 显示OSPF接口信息 | **display ospf** [ *process-id* ] **interface** [ *interface-type interface-number* \| **verbose** ] |
| 显示OSPF路由计算的日志信息 | **display ospf** [ *process-id* ] **event-log** { **peer** \| **spf** } |
| 显示OSPF ASBR聚合信息 | **display ospf** [ *process-id* ] **asbr-summary** [ *ip-address* { *mask-length* \| *mask* } ] |
| 显示全局Router ID | **display router id** |
| 清除OSPF的统计信息 | **reset ospf** [ *process-id* ] **statistics** |
| 清除OSPF的日志信息 | **reset ospf** [ *process-id* ] **event-log** [ **peer** \| **spf** ] |
| 重启OSPF进程 | **reset ospf** [ *process-id* ] **process** [ **graceful-restart** ] |
| 重新向OSPF引入外部路由 | **reset ospf** [ *process-id* ] **redistribution** |

# 1.13 典型配置举例

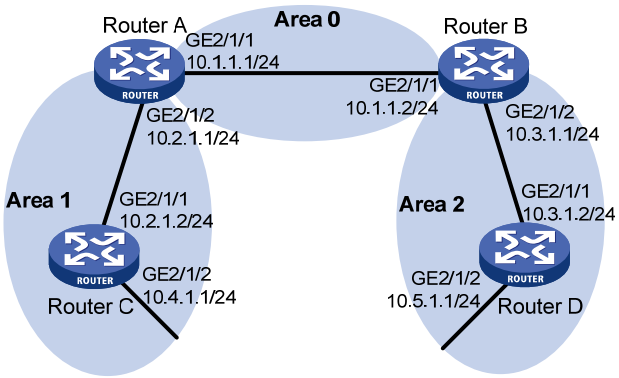
## 1.13.1 配置OSPF基本功能

### 1. 组网需求

- 所有的路由器都运行 OSPF，并将整个自治系统划分为 3 个区域。
- 其中 Router A 和 Router B 作为 ABR 来转发区域之间的路由。
- 配置完成后，每台路由器都应学到 AS 内的到所有网段的路由。

### 2. 组网图

图1-8 OSPF 基本功能配置组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 OSPF 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] router id 10.2.1.1
[RouterA] ospf
[RouterA-ospf-1] area 0
[RouterA-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] quit
[RouterA-ospf-1] area 1
[RouterA-ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.1] quit
[RouterA-ospf-1] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] router id 10.3.1.1
[RouterB] ospf
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
[RouterB-ospf-1] area 2
[RouterB-ospf-1-area-0.0.0.2] network 10.3.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.2] quit
[RouterB-ospf-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] router id 10.4.1.1
[RouterC] ospf
[RouterC-ospf-1] area 1
[RouterC-ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.1] network 10.4.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.1] quit
[RouterC-ospf-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] router id 10.5.1.1
[RouterD] ospf
[RouterD-ospf-1] area 2
[RouterD-ospf-1-area-0.0.0.2] network 10.3.1.0 0.0.0.255
[RouterD-ospf-1-area-0.0.0.2] network 10.5.1.0 0.0.0.255
[RouterD-ospf-1-area-0.0.0.2] quit
[RouterD-ospf-1] quit
```

### 4. 验证配置

# 查看 Router A 的 OSPF 邻居。

```
[RouterA] display ospf peer verbose

          OSPF Process 1 with Router ID 10.2.1.1
```

```
                Neighbors

 Area 0.0.0.0 interface 10.1.1.1(GigabitEthernet2/1/1)'s neighbors
 Router ID: 10.3.1.1         Address: 10.1.1.2         GR State: Normal
   State: Full  Mode: Nbr is Master  Priority: 1
   DR: 10.1.1.1  BDR: 10.1.1.2  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 37  sec
   Neighbor is up for 06:03:59
   Authentication Sequence: [ 0 ]
   Neighbor state change count: 5

 Area 0.0.0.1 interface 10.2.1.1(GigabitEthernet2/1/2)'s neighbors
 Router ID: 10.4.1.1         Address: 10.2.1.2         GR State: Normal
   State: Full  Mode: Nbr is Master  Priority: 1
   DR: 10.2.1.1  BDR: 10.2.1.2  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 32  sec
   Neighbor is up for 06:03:12
   Authentication Sequence: [ 0 ]
   Neighbor state change count: 5
```

# 查看 Router A 的 OSPF 路由信息。

```
[RouterA] display ospf routing

          OSPF Process 1 with Router ID 10.2.1.1
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 10.2.1.0/24        1        Transit 10.2.1.1        10.2.1.1        0.0.0.1
 10.3.1.0/24        2        Inter   10.1.1.2        10.3.1.1        0.0.0.0
 10.4.1.0/24        2        Stub    10.2.1.2        10.4.1.1        0.0.0.1
 10.5.1.0/24        3        Inter   10.1.1.2        10.3.1.1        0.0.0.0
 10.1.1.0/24        1        Transit 10.1.1.1        10.2.1.1        0.0.0.0

 Total Nets: 5
 Intra Area: 3  Inter Area: 2  ASE: 0  NSSA: 0
```

# 查看 Router D 的 OSPF 路由信息。

```
[RouterD] display ospf routing

          OSPF Process 1 with Router ID 10.5.1.1
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 10.2.1.0/24         3       Inter   10.3.1.1        10.3.1.1        0.0.0.2
 10.3.1.0/24         1       Transit 10.3.1.2        10.3.1.1        0.0.0.2
 10.4.1.0/24         4       Inter   10.3.1.1        10.3.1.1        0.0.0.2
```

```
10.5.1.0/24         1       Stub    10.5.1.1        10.5.1.1        0.0.0.2
10.1.1.0/24         2       Inter   10.3.1.1        10.3.1.1        0.0.0.2

Total Nets: 5
Intra Area: 2  Inter Area: 3  ASE: 0  NSSA: 0
```

# 在 Router D 上使用 Ping 测试连通性。

```
[RouterD] ping 10.4.1.1
Ping 10.4.1.1 (10.4.1.1): 56 data bytes, press CTRL_C to break
56 bytes from 10.4.1.1: icmp_seq=0 ttl=253 time=1.549 ms
56 bytes from 10.4.1.1: icmp_seq=1 ttl=253 time=1.539 ms
56 bytes from 10.4.1.1: icmp_seq=2 ttl=253 time=0.779 ms
56 bytes from 10.4.1.1: icmp_seq=3 ttl=253 time=1.702 ms
56 bytes from 10.4.1.1: icmp_seq=4 ttl=253 time=1.471 ms

--- Ping statistics for 10.4.1.1 ---
5 packet(s) transmitted, 5 packet(s) received, 0.0% packet loss
round-trip min/avg/max/std-dev = 0.779/1.408/1.702/0.323 ms
```

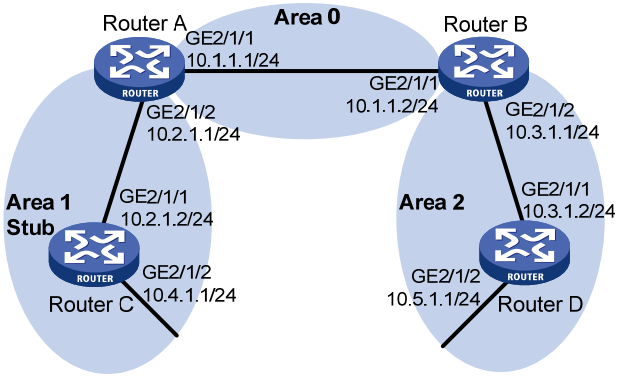
## 1.13.2 配置OSPF引入自治系统外部路由

### 1. 组网需求

- 所有的路由器都运行 OSPF，整个自治系统划分为 3 个区域。
- 其中 Router A 和 Router B 作为 ABR 来转发区域之间的路由。
- 在 Router C 上配置为 ASBR 引入外部路由(静态路由)，且路由信息可正确的在 AS 内传播。

### 2. 组网图

图1-9 OSPF 引入自治系统外部路由配置组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置OSPF（同前例 1.13.1 ）

(3) 配置引入自治系统外部路由

# 在 Router C 上配置一条到目的网段 3.1.2.0/24 的静态路由。

```
<RouterC> system-view
[RouterC] ip route-static 3.1.2.1 24 10.4.1.2
```

# 在 Router C 上配置 OSPF 引入静态路由。

```
[RouterC] ospf 1
[RouterC-ospf-1] import-route static
```

#### 4. 验证配置

# 查看 Router D 的 ABR/ASBR 信息。

```
<RouterD> display ospf abr-asbr

          OSPF Process 1 with Router ID 10.5.1.1
                  Routing Table to ABR and ASBR

 Type    Destination     Area            Cost     Nexthop         RtType
 Intra   10.3.1.1        0.0.0.2         10       10.3.1.1        ABR
 Inter   10.4.1.1        0.0.0.2         22       10.3.1.1        ASBR
```

# 查看 Router D 的 OSPF 路由信息。

```
<RouterD> display ospf routing

          OSPF Process 1 with Router ID 10.5.1.1
                   Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 10.2.1.0/24        22       Inter   10.3.1.1        10.3.1.1        0.0.0.2
 10.3.1.0/24        10       Transit 10.3.1.2        10.3.1.1        0.0.0.2
 10.4.1.0/24        25       Inter   10.3.1.1        10.3.1.1        0.0.0.2
 10.5.1.0/24        10       Stub    10.5.1.1        10.5.1.1        0.0.0.2
 10.1.1.0/24        12       Inter   10.3.1.1        10.3.1.1        0.0.0.2

 Routing for ASEs
 Destination        Cost     Type    Tag         NextHop         AdvRouter
 3.1.2.0/24         1        Type2   1           10.3.1.1        10.4.1.1

 Total Nets: 6
 Intra Area: 2  Inter Area: 3  ASE: 1  NSSA: 0
```

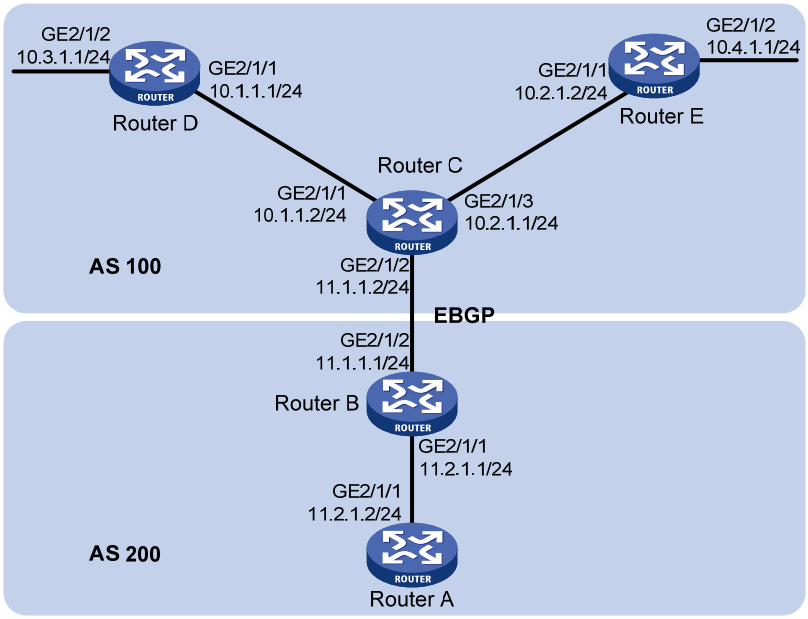
### 1.13.3 配置OSPF发布聚合路由

#### 1. 组网需求

- Router A 和 Router B 位于 AS200 内，AS200 内使用 OSPF 作为 IGP 协议。
- Router C、Router D 和 Router E 位于 AS100 内，AS100 内使用 OSPF 作为 IGP 协议。
- Router B 和 Router C 之间建立 EBGP 连接，配置 BGP 引入 OSPF 和直连路由，配置 OSPF 进程引入 BGP 路由。
- 为了减小 Router A 的路由表规模，在 Router B 上配置路由聚合，只发布聚合后的路由 10.0.0.0/8。

## 2. 组网图

图1-10 OSPF 发布聚合路由配置组网图




## 3. 配置步骤

(1) 配置接口的 IP 地址（略）

(2) 配置 OSPF

# 配置 Router A。

```
<RouterA> system-view
[RouterA] router id 11.2.1.2
[RouterA] ospf
[RouterA-ospf-1] area 0
[RouterA-ospf-1-area-0.0.0.0] network 11.2.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] quit
[RouterA-ospf-1] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] router id 11.2.1.1
[RouterB] ospf
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 11.2.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
[RouterB-ospf-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] router id 11.1.1.2
[RouterC] ospf
[RouterC-ospf-1] area 0
[RouterC-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
```

```
[RouterC-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] quit
[RouterC-ospf-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] router id 10.3.1.1
[RouterD] ospf
[RouterD-ospf-1] area 0
[RouterD-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[RouterD-ospf-1-area-0.0.0.0] network 10.3.1.0 0.0.0.255
[RouterD-ospf-1-area-0.0.0.0] quit
```

# 配置 Router E。

```
<RouterE> system-view
[RouterE] router id 10.4.1.1
[RouterE] ospf
[RouterE-ospf-1] area 0
[RouterE-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[RouterE-ospf-1-area-0.0.0.0] network 10.4.1.0 0.0.0.255
[RouterE-ospf-1-area-0.0.0.0] quit
[RouterE-ospf-1] quit
```

(3) 配置 BGP，引入 OSPF 和直连路由

# 配置 Router B。

```
[RouterB] bgp 200
[RouterB-bgp] peer 11.1.1.2 as-number 100
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] import-route ospf
[RouterB-bgp-ipv4] import-route direct
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
```

# 配置 Router C。

```
[RouterC] bgp 100
[RouterC-bgp] peer 11.1.1.1 as-number 200
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] import-route ospf
[RouterC-bgp-ipv4] import-route direct
[RouterB-bgp-ipv4] quit
[RouterC-bgp] quit
```

(4) 在 Router B 和 Router C 上配置 OSPF 引入 BGP 路由

# 在 Router B 上配置 OSPF 引入 BGP 路由。

```
[RouterB] ospf
[RouterB-ospf-1] import-route bgp
```

# 在 Router C 上配置 OSPF 引入 BGP 路由。

```
[RouterC] ospf
[RouterC-ospf-1] import-route bgp
```

# 查看 Router A 的路由表信息。

```
[RouterA] display ip routing-table
```

```
Destinations : 16        Routes : 16

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
0.0.0.0/32          Direct 0    0            127.0.0.1       InLoop0
10.1.1.0/24         OSPF   150  1            11.2.1.1        GE2/1/1
10.2.1.0/24         OSPF   150  1            11.2.1.1        GE2/1/1
10.3.1.0/24         OSPF   150  1            11.2.1.1        GE2/1/1
10.4.1.0/24         OSPF   150  1            11.2.1.1        GE2/1/1
11.2.1.0/24         Direct 0    0            11.2.1.2        GE2/1/1
11.2.1.0/32         Direct 0    0            11.2.1.2        GE2/1/1
11.2.1.2/32         Direct 0    0            127.0.0.1       InLoop0
11.2.1.255/32       Direct 0    0            11.2.1.2        GE2/1/1
127.0.0.0/8         Direct 0    0            127.0.0.1       InLoop0
127.0.0.0/32        Direct 0    0            127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
127.255.255.255/32  Direct 0    0            127.0.0.1       InLoop0
224.0.0.0/4         Direct 0    0            0.0.0.0         NULL0
224.0.0.0/24        Direct 0    0            0.0.0.0         NULL0
255.255.255.255/32  Direct 0    0            127.0.0.1       InLoop0
```

(5) 在 Router B 上配置路由聚合，只发布聚合路由 10.0.0.0/8。

```
[RouterB-ospf-1] asbr-summary 10.0.0.0 8
```

# 查看 Router A 的路由表信息。

```
[RouterA] display ip routing-table

Destinations : 13        Routes : 13

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
0.0.0.0/32          Direct 0    0            127.0.0.1       InLoop0
10.0.0.0/8          OSPF   150  2            11.2.1.1        GE2/1/1
11.2.1.0/24         Direct 0    0            11.2.1.2        GE2/1/1
11.2.1.0/32         Direct 0    0            11.2.1.2        GE2/1/1
11.2.1.2/32         Direct 0    0            127.0.0.1       InLoop0
11.2.1.255/32       Direct 0    0            11.2.1.2        GE2/1/1
127.0.0.0/8         Direct 0    0            127.0.0.1       InLoop0
127.0.0.0/32        Direct 0    0            127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
127.255.255.255/32  Direct 0    0            127.0.0.1       InLoop0
224.0.0.0/4         Direct 0    0            0.0.0.0         NULL0
224.0.0.0/24        Direct 0    0            0.0.0.0         NULL0
255.255.255.255/32  Direct 0    0            127.0.0.1       InLoop0
```

可以看出，路由 10.1.1.0/24、10.2.1.0/24、10.3.1.0/24、10.4.1.0/24 已经聚合为一条路由 10.0.0.0/8。

## 1.13.4 配置OSPF的Stub区域

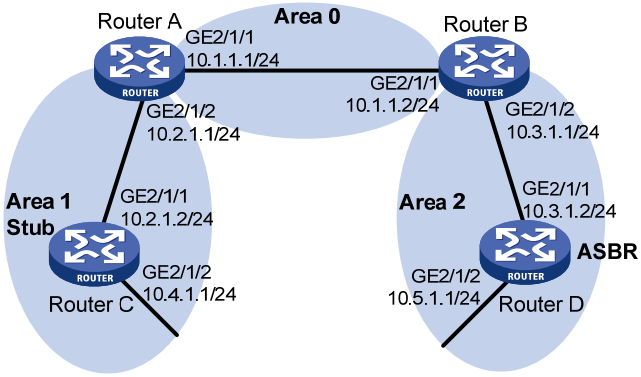
### 1. 组网需求

- 所有的路由器都运行 OSPF，整个自治系统划分为 3 个区域。

- 其中 Router A 和 Router B 作为 ABR 来转发区域之间的路由，Router D 作为 ASBR 引入了外部路由（静态路由）。
- 要求将 Area1 配置为 Stub 区域，减少通告到此区域内的 LSA 数量，但不影响路由的可达性。

## 2. 组网图

图1-11 OSPF Stub 区域配置组网图




## 3. 配置步骤

(1) 配置接口的 IP 地址（略）

(2) 配置OSPF（同前例 1.13.1 ）

(3) 配置 Router D 引入静态路由

```
<RouterD> system-view
[RouterD] ip route-static 3.1.2.1 24 10.5.1.2
[RouterD] ospf
[RouterD-ospf-1] import-route static
[RouterD-ospf-1] quit
```

# 查看 Router C 的 ABR/ASBR 信息。

```
<RouterC> display ospf abr-asbr

          OSPF Process 1 with Router ID 10.4.1.1
                  Routing Table to ABR and ASBR

 Type        Destination     Area            Cost  Nexthop         RtType
 Intra       10.2.1.1        0.0.0.1         3     10.2.1.1        ABR
 Inter       10.5.1.1        0.0.0.1         7     10.2.1.1        ASBR
```

# 查看 Router C 的 OSPF 路由表。

```
<RouterC> display ospf routing

          OSPF Process 1 with Router ID 10.4.1.1
                   Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 10.2.1.0/24        3        Transit 10.2.1.2        10.2.1.1        0.0.0.1
 10.3.1.0/24        7        Inter   10.2.1.1        10.2.1.1        0.0.0.1
```

```
10.4.1.0/24        3        Stub    10.4.1.1        10.4.1.1        0.0.0.1
10.5.1.0/24        17       Inter   10.2.1.1        10.2.1.1        0.0.0.1
10.1.1.0/24        5        Inter   10.2.1.1        10.2.1.1        0.0.0.1

Routing for ASEs
Destination        Cost     Type    Tag         NextHop         AdvRouter
3.1.2.0/24         1        Type2   1           10.2.1.1        10.5.1.1

Total Nets: 6
Intra Area: 2  Inter Area: 3  ASE: 1  NSSA: 0
```

说明

当 Router C 所在区域为普通区域时，可以看到路由表中存在 AS 外部的路由。

(4) 配置 Area1 为 Stub 区域

# 配置 Router A。

```
<RouterA> system-view
[RouterA] ospf
[RouterA-ospf-1] area 1
[RouterA-ospf-1-area-0.0.0.1] stub
[RouterA-ospf-1-area-0.0.0.1] quit
[RouterA-ospf-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] ospf
[RouterC-ospf-1] area 1
[RouterC-ospf-1-area-0.0.0.1] stub
[RouterC-ospf-1-area-0.0.0.1] quit
[RouterC-ospf-1] quit
```

# 显示 Router C 的路由表。

```
[RouterC] display ospf routing

          OSPF Process 1 with Router ID 10.4.1.1
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 0.0.0.0/0          4        Inter   10.2.1.1        10.2.1.1        0.0.0.1
 10.2.1.0/24        3        Transit 10.2.1.2        10.2.1.1        0.0.0.1
 10.3.1.0/24        7        Inter   10.2.1.1        10.2.1.1        0.0.0.1
 10.4.1.0/24        3        Stub    10.4.1.1        10.4.1.1        0.0.0.1
 10.5.1.0/24        17       Inter   10.2.1.1        10.2.1.1        0.0.0.1
 10.1.1.0/24        5        Inter   10.2.1.1        10.2.1.1        0.0.0.1

 Total Nets: 6
 Intra Area: 2  Inter Area: 4  ASE: 0  NSSA: 0
```

说明

当把 Router C 所在区域配置为 Stub 区域时，已经看不到 AS 外部的路由，取而代之的是一条缺省路由。

# 在 ABR 上配置禁止向区域通告 Type3 LSA。

```
[RouterA] ospf
[RouterA-ospf-1] area 1
[RouterA-ospf-1-area-0.0.0.1] stub no-summary
[RouterA-ospf-1-area-0.0.0.1] quit
```

# 查看 Router C 的 OSPF 路由表。

```
[RouterC] display ospf routing

          OSPF Process 1 with Router ID 10.4.1.1
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 0.0.0.0/0          4        Inter   10.2.1.1        10.2.1.1        0.0.0.1
 10.2.1.0/24        3        Transit 10.2.1.2        10.4.1.1        0.0.0.1
 10.4.1.0/24        3        Stub    10.4.1.1        10.4.1.1        0.0.0.1

 Total Nets: 3
 Intra Area: 2  Inter Area: 1  ASE: 0  NSSA: 0
```

说明

禁止向 Stub 区域通告 Summary LSA 后，Stub 路由器的路由表项进一步减少，只保留了一条通往区域外部的缺省路由。

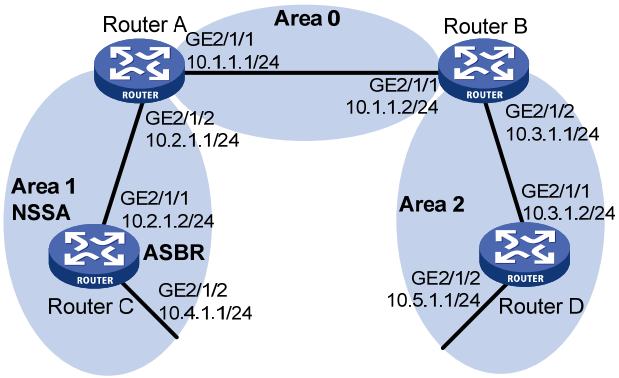
## 1.13.5 配置OSPF的NSSA区域

### 1. 组网需求

- 所有的路由器都运行 OSPF，整个自治系统划分为 3 个区域。
- 其中 Router A 和 Router B 作为 ABR 来转发区域之间的路由。
- 要求将 Area1 配置为 NSSA 区域，同时将 Router C 配置为 ASBR 引入外部路由（静态路由），且路由信息可正确的在 AS 内传播。

## 2. 组网图

图1-12 OSPF NSSA 区域配置组网图




## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置OSPF（同前例 1.13.1 ）

(3) 配置 Area1 区域为 NSSA 区域

# 配置 Router A。

```
<RouterA> system-view
[RouterA] ospf
[RouterA-ospf-1] area 1
[RouterA-ospf-1-area-0.0.0.1] nssa
[RouterA-ospf-1-area-0.0.0.1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] ospf
[RouterC-ospf-1] area 1
[RouterC-ospf-1-area-0.0.0.1] nssa
[RouterC-ospf-1-area-0.0.0.1] quit
[RouterC-ospf-1] quit
```

说明

- 如果 NSSA 区域内路由器（Router C）需要获取通往 AS 内其他区域的路由，ABR（Router A）上必须配置 **default-route-advertise** 参数，这样 Router C 才可以获取到缺省路由。
- 建议在 ABR( Router A )上配置 **no-summary** 参数，这样可以减少 NSSA 路由器的路由表数量。其他 NSSA 路由器只需配置 **nssa** 命令就可以。

# 查看 Router C 的 OSPF 路由表。

```
[RouterC] display ospf routing

          OSPF Process 1 with Router ID 10.4.1.1
                  Routing Tables
```

```
 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 10.2.1.0/24        3        Transit 10.2.1.2        10.4.1.1        0.0.0.1
 10.3.1.0/24        7        Inter   10.2.1.1        10.2.1.1        0.0.0.1
 10.4.1.0/24        3        Stub    10.4.1.1        10.4.1.1        0.0.0.1
 10.5.1.0/24        17       Inter   10.2.1.1        10.2.1.1        0.0.0.1
 10.1.1.0/24        5        Inter   10.2.1.1        10.2.1.1        0.0.0.1

 Total Nets: 5
 Intra Area: 2  Inter Area: 3  ASE: 0  NSSA: 0
```

(4) 配置 Router C 引入静态路由

```
[RouterC] ip route-static 3.1.2.1 24 10.4.1.2
[RouterC] ospf
[RouterC-ospf-1] import-route static
[RouterC-ospf-1] quit
```

# 查看 Router D 的 OSPF 路由表。

```
<RouterD> display ospf routing

          OSPF Process 1 with Router ID 10.5.1.1
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 10.2.1.0/24        22       Inter   10.3.1.1        10.3.1.1        0.0.0.2
 10.3.1.0/24        10       Transit 10.3.1.2        10.3.1.1        0.0.0.2
 10.4.1.0/24        25       Inter   10.3.1.1        10.3.1.1        0.0.0.2
 10.5.1.0/24        10       Stub    10.5.1.1        10.5.1.1        0.0.0.2
 10.1.1.0/24        12       Inter   10.3.1.1        10.3.1.1        0.0.0.2

 Routing for ASEs
 Destination        Cost     Type    Tag         NextHop         AdvRouter
 3.1.2.0/24         1        Type2   1           10.3.1.1        10.2.1.1

 Total Nets: 6
 Intra Area: 2  Inter Area: 3  ASE: 1  NSSA: 0
```

说明

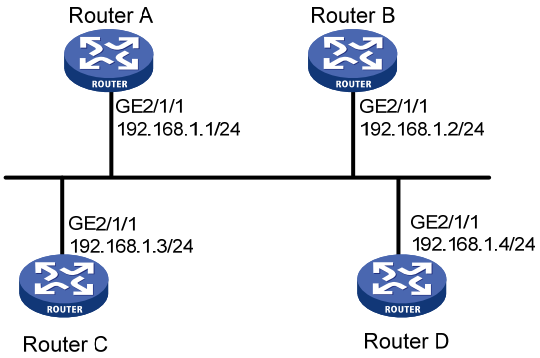
在 Router D 上可以看到 NSSA 区域引入的一条 AS 外部的路由。

## 1.13.6 配置OSPF的DR选择

### 1. 组网需求

- Router A、Router B、Router C、Router D 在同一网段，运行 OSPF 协议。
- 配置 Router A 为 DR，Router C 为 BDR。

图1-13 OSPF 的 DR 选择配置组网图




## 2. 配置思路

- 配置各接口的 IP 地址；
- 配置 OSPF 基本功能；
- 改变路由器接口的路由器优先级使 Router A 成为 DR，Router C 成为 BDR。

## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 OSPF 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] router id 1.1.1.1
[RouterA] ospf
[RouterA-ospf-1] area 0
[RouterA-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] quit
[RouterA-ospf-1] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] router id 2.2.2.2
[RouterB] ospf
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
[RouterB-ospf-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] router id 3.3.3.3
[RouterC] ospf
[RouterC-ospf-1] area 0
[RouterC-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] quit
[RouterC-ospf-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] router id 4.4.4.4
[RouterD] ospf
[RouterD-ospf-1] area 0
[RouterD-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[RouterD-ospf-1-area-0.0.0.0] quit
[RouterD-ospf-1] return
```

# 查看 Router A 的邻居信息。

```
[RouterA] display ospf peer verbose

          OSPF Process 1 with Router ID 1.1.1.1
                  Neighbors

 Area 0.0.0.0 interface 192.168.1.1(GigabitEthernet2/1/1)'s neighbors
 Router ID: 2.2.2.2          Address: 192.168.1.2      GR State: Normal
   State: 2-Way  Mode: None  Priority: 1
   DR: 192.168.1.4  BDR: 192.168.1.3  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 38  sec
   Neighbor is up for 00:01:31
   Authentication Sequence: [ 0 ]

 Router ID: 3.3.3.3          Address: 192.168.1.3      GR State: Normal
   State: Full  Mode: Nbr is Master  Priority: 1
   DR: 192.168.1.4  BDR: 192.168.1.3  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 31  sec
   Neighbor is up for 00:01:28
   Authentication Sequence: [ 0 ]

 Router ID: 4.4.4.4          Address: 192.168.1.4      GR State: Normal
   State: Full  Mode: Nbr is Master  Priority: 1
   DR: 192.168.1.4  BDR: 192.168.1.3  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 31  sec
   Neighbor is up for 00:01:28
   Authentication Sequence: [ 0 ]
```

可以看到 Router D 为 DR，Router C 为 BDR。

(3) 配置接口上的路由器优先级

# 配置 Router A。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospf dr-priority 100
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B。

```
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospf dr-priority 0
[RouterB-GigabitEthernet2/1/1] quit
```

# 配置 Router C。

```
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ospf dr-priority 2
[RouterC-GigabitEthernet2/1/1] quit
```

# 查看 Router D 的邻居信息。

```
<RouterD> display ospf peer verbose

          OSPF Process 1 with Router ID 4.4.4.4
                  Neighbors

 Area 0.0.0.0 interface 192.168.1.4(GigabitEthernet2/1/1)'s neighbors
 Router ID: 1.1.1.1      Address: 192.168.1.1      GR State: Normal
   State: Full  Mode:Nbr is  Slave  Priority: 100
   DR: 192.168.1.4  BDR: 192.168.1.3  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 31  sec
   Neighbor is up for 00:11:17
   Authentication Sequence: [ 0 ]

 Router ID: 2.2.2.2      Address: 192.168.1.2      GR State: Normal
   State: Full  Mode:Nbr is  Slave  Priority: 0
   DR: 192.168.1.4  BDR: 192.168.1.3  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 35  sec
   Neighbor is up for 00:11:19
   Authentication Sequence: [ 0 ]

 Router ID: 3.3.3.3      Address: 192.168.1.3      GR State: Normal
   State: Full  Mode:Nbr is  Slave  Priority: 2
   DR: 192.168.1.4  BDR: 192.168.1.3  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 33  sec
   Neighbor is up for 00:11:15
   Authentication Sequence: [ 0 ]
```

可以看到，网络中 DR/BDR 并没有改变。

说明

网络中 DR/BDR 已经存在的情况下，接口上的路由器优先级的配置并不会立即生效。

(4) 重启 OSPF 进程

# 重启 Router D 的进程。

```
<RouterD> reset ospf 1 process
Warning : Reset OSPF process? [Y/N]:y
```

# 查看 Router D 的邻居信息。

```
<RouterD> display ospf peer verbose
```

```
          OSPF Process 1 with Router ID 4.4.4.4
                  Neighbors

 Area 0.0.0.0 interface 192.168.1.4(GigabitEthernet2/1/1)'s neighbors
 Router ID: 1.1.1.1          Address: 192.168.1.1      GR State: Normal
   State: Full  Mode: Nbr is Slave  Priority: 100
   DR: 192.168.1.1  BDR: 192.168.1.3  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 39  sec
   Neighbor is up for 00:01:40
   Authentication Sequence: [ 0 ]

 Router ID: 2.2.2.2          Address: 192.168.1.2      GR State: Normal
   State: 2-Way  Mode: None  Priority: 0
   DR: 192.168.1.1  BDR: 192.168.1.3  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 35  sec
   Neighbor is up for 00:01:44
   Authentication Sequence: [ 0 ]

 Router ID: 3.3.3.3          Address: 192.168.1.3      GR State: Normal
   State: Full  Mode: Nbr is Slave  Priority: 2
   DR: 192.168.1.1  BDR: 192.168.1.3  MTU: 0
   Options is 0x02 (-|-|-|-|-|-|E|-)
   Dead timer due in 39  sec
   Neighbor is up for 00:01:41
   Authentication Sequence: [ 0 ]
```

可以看到 Router A 成为 DR，Router C 为 BDR。

说明

当路由器的邻居关系稳定后：

- 如果邻居的状态是 Full，这说明它和邻居之间形成了邻接关系；
- 如果邻居的状态是 2-Way，则说明它们都不是 DR 或 BDR，两者之间不需要交换 LSA。

# 查看 OSPF 接口的状态。

```
[RouterA] display ospf interface

          OSPF Process 1 with Router ID 1.1.1.1
                  Interfaces

 Area: 0.0.0.0
 IP Address      Type      State   Cost  Pri   DR             BDR
 192.168.1.1     Broadcast DR      1     100   192.168.1.1    192.168.1.3
[RouterB] display ospf interface

          OSPF Process 1 with Router ID 2.2.2.2
```

```
                Interfaces

Area: 0.0.0.0
IP Address      Type       State     Cost  Pri   DR             BDR
192.168.1.2     Broadcast  DROther   1     0     192.168.1.1    192.168.1.3
```

说明

如果 OSPF 接口的状态是 DROther，则说明它既不是 DR，也不是 BDR。

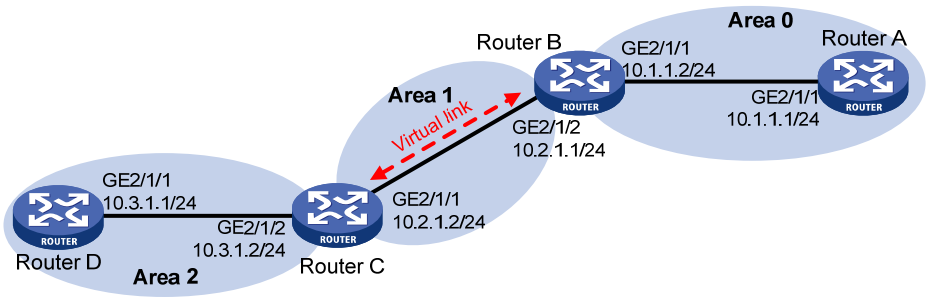
## 1.13.7 配置OSPF虚连接

### 1. 组网需求

- Area2 没有与 Area0 直接相连。Area1 被用作传输区域（Transit Area）来连接 Area2 和 Area0。Router B 和 Router C 之间配置一条虚连接。
- 配置完成后，Router B 能够学到 Area2 中的路由。

### 2. 组网图

图1-14 OSPF 虚连接配置组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 OSPF 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] ospf 1 router-id 1.1.1.1
[RouterA-ospf-1] area 0
[RouterA-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] ospf 1 router-id 2.2.2.2
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
```

```
[RouterB-ospf-1] area 1
[RouterB–ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[RouterB–ospf-1-area-0.0.0.1] quit
[RouterB-ospf-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] ospf 1 router-id 3.3.3.3
[RouterC-ospf-1] area 1
[RouterC-ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.1] quit
[RouterC-ospf-1] area 2
[RouterC–ospf-1-area-0.0.0.2] network 10.3.1.0 0.0.0.255
[RouterC–ospf-1-area-0.0.0.2] quit
[RouterC-ospf-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] ospf 1 router-id 4.4.4.4
[RouterD-ospf-1] area 2
[RouterD-ospf-1-area-0.0.0.2] network 10.3.1.0 0.0.0.255
[RouterD-ospf-1-area-0.0.0.2] quit
```

# 查看 Router B 的 OSPF 路由表。

```
[RouterB] display ospf routing

          OSPF Process 1 with Router ID 2.2.2.2
                   Routing Tables
 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 10.2.1.0/24        2        Transit 10.2.1.1        3.3.3.3         0.0.0.1
 10.1.1.0/24        2        Transit 10.1.1.2        2.2.2.2         0.0.0.0
 Total Nets: 2
 Intra Area: 2  Inter Area: 0  ASE: 0  NSSA: 0
```

说明

由于 Area0 没有与 Area2 直接相连，所以 Router B 的路由表中没有 Area 2 的路由。

(3) 配置虚连接

# 配置 Router B。

```
[RouterB] ospf
[RouterB-ospf-1] area 1
[RouterB-ospf-1-area-0.0.0.1] vlink-peer 3.3.3.3
[RouterB-ospf-1-area-0.0.0.1] quit
[RouterB-ospf-1] quit
```

# 配置 Router C。

```
[RouterhC] ospf
[RouterC-ospf-1] area 1
[RouterC-ospf-1-area-0.0.0.1] vlink-peer 2.2.2.2
```

```
[RouterC-ospf-1-area-0.0.0.1] quit
```

# 查看 Router B 的 OSPF 路由表。

```
[RouterB] display ospf routing

          OSPF Process 1 with Router ID 2.2.2.2
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 10.2.1.0/24        2        Transit 10.2.1.1        3.3.3.3         0.0.0.1
 10.3.1.0/24        5        Inter   10.2.1.2        3.3.3.3         0.0.0.0
 10.1.1.0/24        2        Transit 10.1.1.2        2.2.2.2         0.0.0.0

 Total Nets: 3
 Intra Area: 2  Inter Area: 1  ASE: 0  NSSA: 0
```

可以看到，Router B 已经学到了 Area2 的路由 10.3.1.0/24。

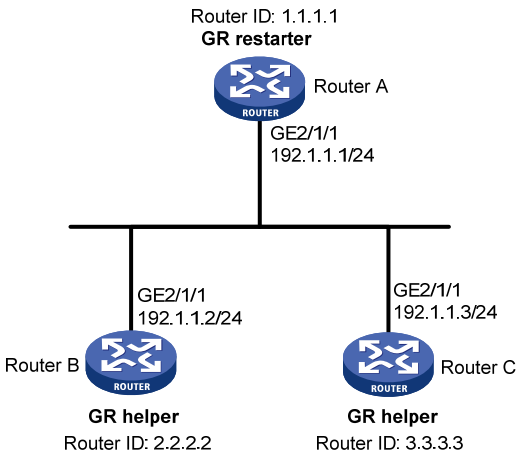
## 1.13.8 OSPF GR配置举例

### 1. 组网需求

- Router A、Router B 和 Router C 既属于同一自治系统，也属于同一 OSPF 域，通过 OSPF 协议实现网络互连，并提供 GR 机制。
- Router A 作为非 IETF 标准 GR Restarter，Router B 和 Router C 作为 GR Helper 并且通过 GR 机制与 Router A 保持带外同步。

### 2. 组网图

图1-15 OSPF GR 配置组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 OSPF 基本功能

# 配置 Router A。

```
<RouterA> system-view
```

```
[RouterA] router id 1.1.1.1
[RouterA] ospf 100
[RouterA-ospf-100] area 0
[RouterA-ospf-100-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[RouterA-ospf-100-area-0.0.0.0] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] router id 2.2.2.2
[RouterB] ospf 100
[RouterB-ospf-100] area 0
[RouterB-ospf-100-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[RouterB-ospf-100-area-0.0.0.0] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] router id 3.3.3.3
[RouterC] ospf 100
[RouterC-ospf-100] area 0
[RouterC-ospf-100-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[RouterC-ospf-100-area-0.0.0.0] quit
```

(3) 配置 OSPF GR

# 配置 Router A 作为非 IETF 标准 GR Restarter，即使能 OSPF 进程 100 的本地链路信令能力、OSPF 带外同步能力和非 IETF 标准 GR 能力。

```
[RouterA-ospf-100] enable link-local-signaling
[RouterA-ospf-100] enable out-of-band-resynchronization
[RouterA-ospf-100] graceful-restart
[RouterA-ospf-100] return
```

# 配置 Router B 作为 GR Helper，即使能 OSPF 进程 100 的本地链路信令能力和 OSPF 带外同步能力。

```
[RouterB-ospf-100] enable link-local-signaling
[RouterB-ospf-100] enable out-of-band-resynchronization
```

# 配置 Router C 作为 GR Helper，即使能 OSPF 进程 100 的本地链路信令能力和 OSPF 带外同步能力。

```
[RouterC-ospf-100] enable link-local-signaling
[RouterC-ospf-100] enable out-of-band-resynchronization
```

### 4. 验证配置

# 运行稳定后，打开 Router A 的 OSPF 平滑启动事件调试信息开关。在 Router A 上以 GR 方式重启 OSPF 进程。

```
<RouterA> debugging ospf event graceful-restart
<RouterA> terminal monitor
<RouterA> terminal logging level 7
<RouterA> reset ospf 100 process graceful-restart
Reset OSPF process? [Y/N]:y
%Oct 21 15:29:28:727 2011 RouterA OSPF/5/OSPF_NBR_CHG: -MDC=1; OSPF 100 Neighbor
192.1.1.2(GigabitEthernet2/1/1) from Full to Down.
%Oct 21 15:29:28:729 2011 RouterA OSPF/5/OSPF_NBR_CHG: -MDC=1; OSPF 100 Neighbor
192.1.1.3(GigabitEthernet2/1/1) from Full to Down.
```

```
*Oct 21 15:29:28:735 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF 100 nonstandard GR Started for OSPF Router
*Oct 21 15:29:28:735 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF 100 created GR wait timer,timeout interval is 40(s).
*Oct 21 15:29:28:735 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF 100 created GR Interval timer,timeout interval is 120(s).
*Oct 21 15:29:28:758 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF 100 created OOB Progress timer for neighbor 192.1.1.3.
*Oct 21 15:29:28:766 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF 100 created OOB Progress timer for neighbor 192.1.1.2.
%Oct 21 15:29:29:902 2011 RouterA OSPF/5/OSPF_NBR_CHG: -MDC=1; OSPF 100 Neighbor
192.1.1.2(GigabitEthernet2/1/1) from Loading to Full.
*Oct 21 15:29:29:902 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF 100 deleted OOB Progress timer for neighbor 192.1.1.2.
%Oct 21 15:29:30:897 2011 RouterA OSPF/5/OSPF_NBR_CHG: -MDC=1; OSPF 100 Neighbor
192.1.1.3(GigabitEthernet2/1/1) from Loading to Full.
*Oct 21 15:29:30:897 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF 100 deleted OOB Progress timer for neighbor 192.1.1.3.
*Oct 21 15:29:30:911 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF GR: Process 100 Exit Restart,Reason : DR or BDR change,for neighbor : 192.1.1.3.
*Oct 21 15:29:30:911 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF 100 deleted GR Interval timer.
*Oct 21 15:29:30:912 2011 RouterA OSPF/7/DEBUG: -MDC=1;
OSPF 100 deleted GR wait timer.
%Oct 21 15:29:30:920 2011 RouterA OSPF/5/OSPF_NBR_CHG: -MDC=1; OSPF 100 Neighbor
192.1.1.2(GigabitEthernet2/1/1) from Full to Down.
%Oct 21 15:29:30:921 2011 RouterA OSPF/5/OSPF_NBR_CHG: -MDC=1; OSPF 100 Neighbor
192.1.1.3(GigabitEthernet2/1/1) from Full to Down.
%Oct 21 15:29:33:815 2011 RouterA OSPF/5/OSPF_NBR_CHG: -MDC=1; OSPF 100 Neighbor
192.1.1.3(GigabitEthernet2/1/1) from Loading to Full.
%Oct 21 15:29:35:578 2011 RouterA OSPF/5/OSPF_NBR_CHG: -MDC=1; OSPF 100 Neighbor
192.1.1.2(GigabitEthernet2/1/1) from Loading to Full.
```

从上面的信息可以看出 Router A 完成了 GR。

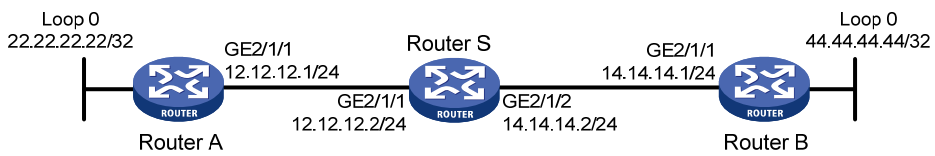
## 1.13.9 OSPF NSR配置举例

### 1. 组网需求

Router S、Router A、Router B 属于同一 OSPF 区域，通过 OSPF 协议实现网络互连。要求对 Router S 进行主备倒换时，Router A 和 Router B 到 Route S 的邻居没有中断，Router A 到 Router B 的流量没有中断。

### 2. 组网图

图1-16 OSPF NSR 配置组网图



### 3. 配置步骤

(1) 配置各路由器接口的 IP 地址和 OSPF 协议

请按照上面组网图配置各接口的 IP 地址和子网掩码，具体配置过程略。

配置各路由器之间采用 OSPF 协议进行互连，确保 Router S、Router A 和 Router B 之间能够在网络层互通，并且各路由器之间能够借助 OSPF 协议实现动态路由更新。具体配置过程略。

(2) 配置 OSPF NSR

# 使能 Router S 的 OSPF NSR 功能。

```
<RouterS> system-view
[RouterS] ospf 100
[RouterS-ospf-100] ospf non-stop-routing
[RouterS-ospf-100] quit
```

### 4. 验证配置

# Router S 分别与 Router A 和 Router B 建立邻接关系后，三台路由器开始交换路由信息。当网络稳定后，Router S 进行主备倒换。在 Route S 主备倒换期间，使用 **display ospf peer** 命令查看 Router A 和 Router B 上到 Router S 的邻居是否发生任何变化；使用 **display ospf routing** 命令查看 Router A 上是否有 Route B 上 Loopback 接口的路由，Router B 上是否有 Route A 上 Loopback 接口的路由。

# Router S 主备倒换。

```
[RouterS] placement reoptimize
Predicted changes to the placement
Program                           Current location       New location
-------------------------------------------------------------------------
lb                                0/0                    0/0
lsm                               0/0                    0/0
slsp                              0/0                    0/0
rib6                              0/0                    0/0
routepolicy                       0/0                    0/0
rib                               0/0                    0/0
staticroute6                      0/0                    0/0
staticroute                       0/0                    0/0
eviisis                           0/0                    0/0
ospf                              0/0                    1/0
Continue? [y/n]:y
Re-optimization of the placement start. You will be notified on completion
Re-optimization of the placement complete. Use 'display placement' to view the new placement
```

# 查看 Router A 上 OSPF 协议的邻居和路由。

```
<RouterA> display ospf peer

          OSPF Process 1 with Router ID 2.2.2.1
               Neighbor Brief Information

 Area: 0.0.0.0
 Router ID       Address         Pri Dead-Time  State             Interface
 3.3.3.1         12.12.12.2      1   37         Full/BDR          GE2/1/1
<RouterA> display ospf routing
```

```
          OSPF Process 1 with Router ID 2.2.2.1
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 44.44.44.44/32     2        Stub    12.12.12.2      4.4.4.1         0.0.0.0
 14.14.14.0/24      2        Transit 12.12.12.2      4.4.4.1         0.0.0.0
 22.22.22.22/32     0        Stub    22.22.22.22     2.2.2.1         0.0.0.0
 12.12.12.0/24      1        Transit 12.12.12.1      2.2.2.1         0.0.0.0

 Total Nets: 4
 Intra Area: 4  Inter Area: 0  ASE: 0  NSSA: 0
```

# 查看 Router B 上 OSPF 协议的邻居和路由。

```
<RouterB> display ospf peer

          OSPF Process 1 with Router ID 4.4.4.1
               Neighbor Brief Information

 Area: 0.0.0.0
 Router ID       Address         Pri Dead-Time  State             Interface
 3.3.3.1         14.14.14.2      1   39         Full/BDR          GE2/1/1
<RouterB> display ospf routing

          OSPF Process 1 with Router ID 4.4.4.1
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 44.44.44.44/32     0        Stub    44.44.44.44     4.4.4.1         0.0.0.0
 14.14.14.0/24      1        Transit 14.14.14.1      4.4.4.1         0.0.0.0
 22.22.22.22/32     2        Stub    14.14.14.2      2.2.2.1         0.0.0.0
 12.12.12.0/24      2        Transit 14.14.14.2      2.2.2.1         0.0.0.0

 Total Nets: 4
 Intra Area: 4  Inter Area: 0  ASE: 0  NSSA: 0
```

通过上面信息可以看出在 Router S 发生主备倒换的时候，Router A 和 Router B 的邻居和路由信息保持不变，从 Router A 到 Router B 的流量转发没有受到主备倒换的影响。

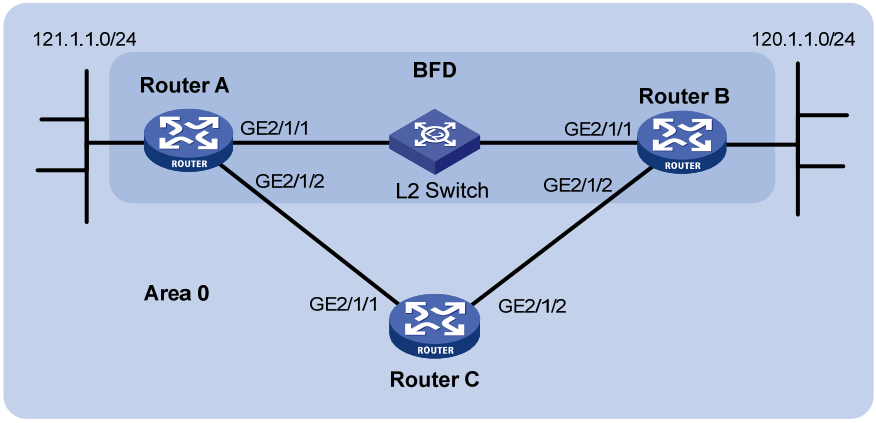
## 1.13.10 配置OSPF与BFD联动

### 1. 组网需求

- Router A、Router B 和 Ruter C 上运行 OSPF，网络层相互可达。
- 当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时 BFD 能够快速感知通告 OSPF 协议，并且切换到 Router C 进行通信。

## 2. 组网图

图1-17 OSPF 与 BFD 联动配置组网图




| 设备 | 接口 | IP地址 | 设备 | 接口 | IP地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 192.168.0.102/24 | Router B | GE2/1/1 | 192.168.0.100/24 |
| | GE2/1/2 | 10.1.1.102/24 | | GE2/1/2 | 13.1.1.1/24 |
| Router C | GE2/1/1 | 10.1.1.100/24 | | | |
| | GE2/1/2 | 13.1.1.2/24 | | | |

## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 OSPF 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] ospf
[RouterA-ospf-1] area 0
[RouterA-ospf-1-area-0.0.0.0] network 192.168.0.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] network 121.1.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] quit
[RouterA-ospf-1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] ospf cost 2
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] ospf
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 192.168.0.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] network 13.1.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] network 120.1.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
[RouterB-ospf-1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] ospf cost 2
```

# 配置 Router C。

```
<RouterC> system-view
```

```
[RouterC] ospf
[RouterC-ospf-1] area 0
[RouterC-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] network 13.1.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] quit
[RouterC-ospf-1] quit
```

(3) 配置 BFD 功能

# 在 Router A 上使能 BFD 检测功能，并配置 BFD 参数。

```
[RouterA] bfd session init-mode active
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospf bfd enable
[RouterA-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterA-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterA-GigabitEthernet2/1/1] bfd detect-multiplier 7
[RouterA-GigabitEthernet2/1/1] return
```

# 在 Router B 上使能 BFD 检测功能，并配置 BFD 参数。

```
[RouterB] bfd session init-mode active
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospf bfd enable
[RouterB-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterB-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterB-GigabitEthernet2/1/1] bfd detect-multiplier 6
```

### 4. 验证配置

下面以 Router A 为例，Router B 和 Router A 类似，不再赘述。

# 查看 BFD 信息。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv4 Session Working Under Ctrl Mode:

 LD/RD          SourceAddr      DestAddr        State    Holdtime    Interface
 3/1            192.168.0.102   192.168.0.100   Up       1700ms      GE2/1/1
```

# 在 Router A 上查看 120.1.1.0/24 的路由信息，可以看出 Router A 和 Router B 是通过 L2 Switch 进行通信的。

```
<RouterA> display ip routing-table 120.1.1.0 verbose

Summary Count : 1

Destination: 120.1.1.0/24
   Protocol: OSPF            Process ID: 1
  SubProtID: 0x1                  Age: 04h20m37s
       Cost: 2              Preference: 10
        Tag: 0                   State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
```

```
      NBRID: 0x26000002          LastAs: 0
     AttrID: 0xffffffff        Neighbor: 0.0.0.0
      Flags: 0x1008c        OrigNextHop: 192.168.0.100
      Label: NULL           RealNextHop: 192.168.0.100
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid        BkInterface: N/A
```

当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时：

# 查看 120.1.1.0/24 的路由信息，可以看出 Router A 和 Router B 已经切换到 Router C 进行通信。

```
<RouterA> display ip routing-table 120.1.1.0 verbose

Summary Count : 1

Destination: 120.1.1.0/24
   Protocol: OSPF            Process ID: 1
  SubProtID: 0x1                    Age: 04h20m37s
       Cost: 4              Preference: 10
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002          LastAs: 0
     AttrID: 0xffffffff        Neighbor: 0.0.0.0
      Flags: 0x1008c        OrigNextHop: 10.1.1.100
      Label: NULL           RealNextHop: 10.1.1.100
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: N/A
```

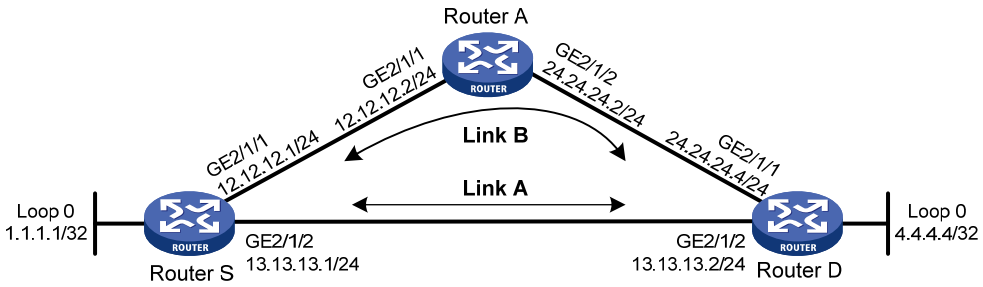
## 1.13.11 OSPF快速重路由配置举例

### 1. 组网需求

如 图 1-18 所示，Router S 、Router A和Router D属于同一OSPF区域，通过OSPF协议实现网络互连。要求当Router S和Router D之间的链路出现故障时，业务可以快速切换到链路B上。

### 2. 组网图

图1-18 OSPF 快速重路由配置组网图



### 3. 配置步骤

(1) 配置各路由器接口的 IP 地址和 OSPF 协议

请按照上面组网图配置各接口的 IP 地址和子网掩码，具体配置过程略。

配置各路由器之间采用 OSPF 协议进行互连，确保 Router S、Router A 和 Router D 之间能够在网络层互通，并且各路由器之间能够借助 OSPF 协议实现动态路由更新。

具体配置过程略。

(2) 配置 OSPF 快速重路由

OSPF 支持快速重路由配置有两种配置方法，可以任选一种。

方法一：使能 Router S 和 Router D 的 OSPF 快速重路由功能（通过 LFA 算法选取备份下一跳信息）

# 配置 Router S。

```
<RouterS> system-view
[RouterS] ospf 1
[RouterS-ospf-1] fast-reroute lfa
[RouterS-ospf-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] ospf 1
[RouterD-ospf-1] fast-reroute lfa
[RouterD-ospf-1] quit
```

方法二：使能 Router S 和 Router D 的 OSPF 快速重路由功能（通过路由策略指定备份下一跳）

# 配置 Router S。

```
<RouterS> system-view
[RouterS] ip prefix-list abc index 10 permit 4.4.4.4 32
[RouterS] route-policy frr permit node 10
[RouterS-route-policy-frr-10] if-match ip address prefix-list abc
[RouterS-route-policy-frr-10] apply fast-reroute backup-interface gigabitethernet 2/1/1
backup-nexthop 12.12.12.2
[RouterS-route-policy-frr-10] quit
[RouterS] ospf 1
[RouterS-ospf-1] fast-reroute route-policy frr
[RouterS-ospf-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] ip prefix-list abc index 10 permit 1.1.1.1 32
[RouterD] route-policy frr permit node 10
[RouterD-route-policy-frr-10] if-match ip address prefix-list abc
[RouterD-route-policy-frr-10] apply fast-reroute backup-interface gigabitethernet 2/1/1
backup-nexthop 24.24.24.2
[RouterD-route-policy-frr-10] quit
[RouterD] ospf 1
[RouterD-ospf-1] fast-reroute route-policy frr
[RouterD-ospf-1] quit
```

### 4. 验证配置

# 在 Router S 上查看 4.4.4.4/32 路由，可以看到备份下一跳信息。

```
[RouterS] display ip routing-table 4.4.4.4 verbose

Summary Count : 1

Destination: 4.4.4.4/32
   Protocol: OSPF            Process ID: 1
  SubProtID: 0x1                   Age: 04h20m37s
       Cost: 1               Preference: 10
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 0.0.0.0
      Flags: 0x1008c        OrigNextHop: 13.13.13.2
      Label: NULL           RealNextHop: 13.13.13.2
    BkLabel: NULL             BkNextHop: 12.12.12.2
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: GigabitEthernet2/1/1
```

# 在 Router D 上查看 1.1.1.1/32 路由，可以看到备份下一跳信息。

```
[RouterD] display ip routing-table 1.1.1.1 verbose

Summary Count : 1

Destination: 1.1.1.1/32
   Protocol: OSPF            Process ID: 1
  SubProtID: 0x1                   Age: 04h20m37s
       Cost: 1               Preference: 10
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 0.0.0.0
      Flags: 0x1008c        OrigNextHop: 13.13.13.1
      Label: NULL           RealNextHop: 13.13.13.1
    BkLabel: NULL             BkNextHop: 24.24.24.2
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: GigabitEthernet2/1/1
```

## 1.14 常见配置错误举例

### 1.14.1 OSPF邻居无法建立

#### 1. 故障现象

OSPF 邻居无法建立。

#### 2. 分析

如果物理连接和下层协议正常，则检查接口上配置的 OSPF 参数，必须保证与相邻路由器的参数一致，区域号相同，网段与掩码也必须一致（点到点与虚连接的网段与掩码可以不同）。

### 3. 处理过程

(1) 使用 **display ospf peer** 命令查看 OSPF 邻居状态。

(2) 使用 **display ospf interface** 命令查看 OSPF 接口的信息。

(3) 检查物理连接及下层协议是否正常运行，可通过 **ping** 命令测试。若从本地路由器 Ping 对端路由器不通，则表明物理连接和下层协议有问题。

(4) 检查 OSPF 定时器，在同一接口上邻居失效时间应至少为 Hello 报文发送时间间隔的 4 倍。

(5) 如果是 NBMA 网络，则应该使用 **peer ip-address** 命令手工指定邻居。

(6) 如果网络类型为广播网或 NBMA，则至少有一个接口的路由器优先级大于零。

## 1.14.2 OSPF路由信息不正确

### 1. 故障现象

OSPF 不能发现其他区域的路由。

### 2. 分析

应保证骨干区域与所有的区域相连接。若一台路由器配置了两个以上的区域，则至少有一个区域应与骨干区域相连。骨干区域不能配置成 Stub 区域。

在 Stub 区域内的路由器不能接收外部 AS 的路由。如果一个区域配置成 Stub 区域，则与这个区域相连的所有路由器都应将此区域配置成 Stub 区域。

### 3. 处理过程

(1) 使用 **display ospf peer** 命令查看 OSPF 邻居状态。

(2) 使用 **display ospf interface** 命令查看 OSPF 接口的信息。

(3) 使用 **display ospf lsdb** 查看 LSDB 的信息是否完整。

(4) 使用 **display current-configuration configuration ospf** 命令查看区域是否配置正确。若配置了两个以上的区域，则至少有一个区域与骨干区域相连。

(5) 如果某区域是 Stub 区域，则该区域中的所有路由器都要配置 **stub** 命令；如果某区域是 NSSA 区域，则该区域中的所有路由器都要配置 **nssa** 命令。

(6) 如果配置了虚连接，使用 **display ospf vlink** 命令查看 OSPF 虚连接是否正常。

# 目 录

# 1 IS-IS

## 1.1 IS-IS简介

IS-IS（Intermediate System-to-Intermediate System，中间系统到中间系统）最初是ISO（International Organization for Standardization，国际标准化组织）为它的CLNP（Connection-Less Network Protocol，无连接网络协议）设计的一种动态路由协议。

为了提供对IP的路由支持，IETF（Internet Engineering Task Force，互联网工程任务组）在RFC 1195中对IS-IS进行了扩充和修改，使它能够同时应用在TCP/IP和OSI环境中，称为集成化IS-IS（Integrated IS-IS或Dual IS-IS）。

IS-IS属于IGP（Interior Gateway Protocol，内部网关协议），用于自治系统内部。IS-IS是一种链路状态协议，使用SPF（Shortest Path First，最短路径优先）算法进行路由计算。

### 1.1.1 基本概念

#### 1. IS-IS路由协议的基本术语

- IS（Intermediate System）：中间系统。相当于TCP/IP中的路由器，是IS-IS协议中生成路由和传播路由信息的基本单元。在下文中IS和路由器具有相同的含义。
- ES（End System）：终端系统。相当于TCP/IP中的主机系统。ES不参与IS-IS路由协议的处理，ISO使用专门的ES-IS协议定义终端系统与中间系统间的通信。
- RD（Routing Domain）：路由域。在一个路由域中多个IS通过相同的路由协议来交换路由信息。
- Area：区域，路由域的细分单元，IS-IS允许将整个路由域分为多个区域。
- LSDB（Link State DataBase）：链路状态数据库。网络内所有链路的状态组成了链路状态数据库，在每一个IS中都至少有一个LSDB。IS使用SPF算法，利用LSDB来生成自己的路由。
- LSPDU（Link State Protocol Data Unit）：链路状态协议数据单元，简称LSP。在IS-IS中，每一个IS都会生成LSP，此LSP包含了本IS的所有链路状态信息。
- NPDU（Network Protocol Data Unit）：网络协议数据单元，是OSI中的网络层协议报文，相当于TCP/IP中的IP报文。
- DIS（Designated IS）：广播网络上选举的指定中间系统，也可以称为指定IS。
- NSAP（Network Service Access Point）：网络服务接入点，即OSI中网络层的地址，用来标识一个抽象的网络服务访问点，描述OSI模型的网络地址结构。

#### 2. IS-IS地址结构

(1) NSAP

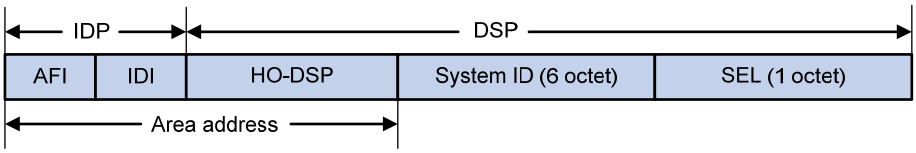
如 图 1-1 所示，NSAP由IDP（Initial Domain Part）和DSP（Domain Specific Part）组成。IDP相当于IP地址中的主网络号，DSP相当于IP地址中的子网号和主机地址。

IDP部分是ISO规定的，它由AFI（Authority and Format Identifier）与IDI（Initial Domain Identifier）组成，AFI表示地址分配机构和地址格式，IDI用来标识域。

DSP 由 HO-DSP（High Order Part of DSP）、SystemID 和 SEL 三个部分组成。HO-DSP 用来分割区域，SystemID 用来区分主机，SEL 指示服务类型。

IDP 和 DSP 的长度都是可变的，NSAP 总长最多是 20 个字节，最少 8 个字节。

图1-1 IS-IS 协议的地址结构示意图




(2) 区域地址

IDP 和 DSP 中的 HO-DSP 一起，既能够标识路由域，也能够标识路由域中的区域，被称为区域地址。两个不同的路由域中不允许有相同的区域地址。

一般情况下，一台路由器只需要配置一个区域地址，且同一区域中所有节点的区域地址都要相同。为了支持区域的平滑合并、分割及转换，一台路由器最多可配置 3 个区域地址。

(3) System ID

System ID 用来在区域内唯一标识主机或路由器。它的长度固定为 48 比特。

在实际应用中，一般使用 Router ID 与 System ID 进行对应。假设一台路由器使用接口 Loopback0 的 IP 地址 168.10.1.1 作为 Router ID，则它在 IS-IS 使用的 System ID 可通过如下方法转换得到：

- 将 IP 地址 168.10.1.1 的每一部分都扩展为 3 位，不足 3 位的在前面补 0；
- 将扩展后的地址 168.010.001.001 重新划分为 3 部分，每部分由 4 位数字组成，得到的 1680.1000.1001 就是 System ID。

实际 System ID 的指定可以有不同的方法，但要保证能够唯一标识主机或路由器。

(4) SEL

SEL 有时也写成 N-SEL（NSAP Selector），它的作用类似 IP 中的“协议标识符”，不同的传输协议对应不同的 SEL。在 IP 中，SEL 均为 00。

(5) 路由方式

由于这种地址结构明确的定义了区域，Level-1 路由器很容易识别出发往它所在的区域之外的报文，这些报文是需要转交给 Level-1-2 路由器的。

- Level-1 路由器利用 System ID 进行区域内的路由，如果发现报文的目的地址不属于自己所在的区域，就将报文转发给最近的 Level-1-2 路由器。
- Level-2 路由器根据区域地址进行区域间的路由。

### 3. NET

NET（Network Entity Title，网络实体名称）指示的是 IS 本身的网络层信息，不包括传输层信息，可以看作是一类特殊的 NSAP，即 SEL 为 0 的 NSAP 地址。因此，NET 的长度与 NSAP 的相同，为 8～20 个字节。

NET 由三部分组成：

- 区域 ID：它的长度可变的，为 1～13 个字节。
- System ID：用来在区域内唯一标识主机或路由器，它的长度固定为 6 个字节。
- SEL：为 0，它的长度固定为 1 个字节。

例如 NET 为：ab.cdef.1234.5678.9abc.00，则其中区域 ID 为 ab.cdef，System ID 为 1234.5678.9abc，SEL 为 00。

通常情况下，一台路由器配置一个 NET 即可，当区域需要重新划分时，例如将多个区域合并，或者将一个区域划分为多个区域，这种情况下配置多个 NET 可以在重新配置时仍然能够保证路由的正确性。由于一台路由器最多可配置 3 个区域地址，所以最多也只能配置 3 个 NET。在配置多个 NET 时，必须保证它们的 System ID 都相同。

### 1.1.2 IS-IS区域

#### 1. 两级结构

为了支持大规模的路由网络，IS-IS 在路由域内采用两级的分层结构。一个大的路由域通常被分成多个区域（Areas）。一般来说，我们将 Level-1 路由器部署在区域内，Level-2 路由器部署在区域间，Level-1-2 路由器部署在 Level-1 路由器和 Level-2 路由器的中间。

#### 2. Level-1 与Level-2

(1) Level-1 路由器

Level-1 路由器负责区域内的路由，它只与属于同一区域的 Level-1 和 Level-1-2 路由器形成邻居关系，维护一个 Level-1 的 LSDB，该 LSDB 包含本区域的路由信息，到区域外的报文转发给最近的 Level-1-2 路由器。

属于不同区域的 Level-1 路由器不能形成邻居关系。

(2) Level-2 路由器

Level-2 路由器负责区域间的路由，可以与同一区域或者其它区域的 Level-2 和 Level-1-2 路由器形成邻居关系，维护一个 Level-2 的 LSDB，该 LSDB 包含区域间的路由信息。所有 Level-2 路由器和 Level-1-2 路由器组成路由域的骨干网，负责在不同区域间通信，骨干网必须是物理连续的。

Level-2 路由器是否形成邻居关系与区域无关。

(3) Level-1-2 路由器

同时属于 Level-1 和 Level-2 的路由器称为 Level-1-2 路由器，可以与同一区域的 Level-1 和 Level-1-2 路由器形成 Level-1 邻居关系，也可以与同一区域或者其他区域的 Level-2 和 Level-1-2 路由器形成 Level-2 的邻居关系。Level-1 路由器必须通过 Level-1-2 路由器才能连接至其他区域。Level-1-2 路由器维护两个 LSDB，Level-1 的 LSDB 用于区域内路由，Level-2 的 LSDB 用于区域间路由。

图 1-2 为一个运行IS-IS协议的网络，其中Area 1 是骨干区域，该区域中的所有路由器均是Level-2 路由器。另外 4 个区域为非骨干区域，它们都通过Level-1-2 路由器与骨干路由器相连。

图1-2 IS-IS 拓扑结构图之一

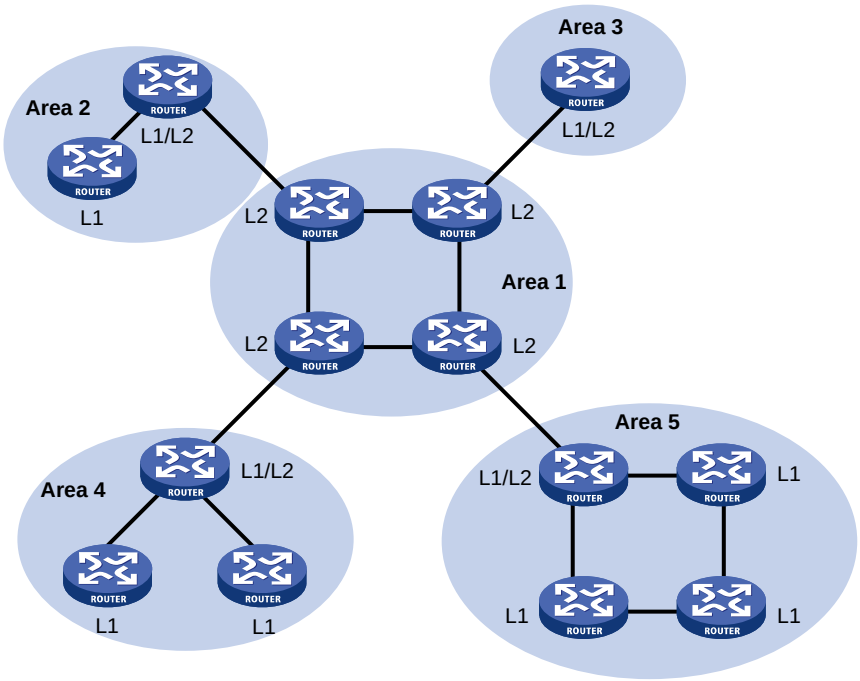
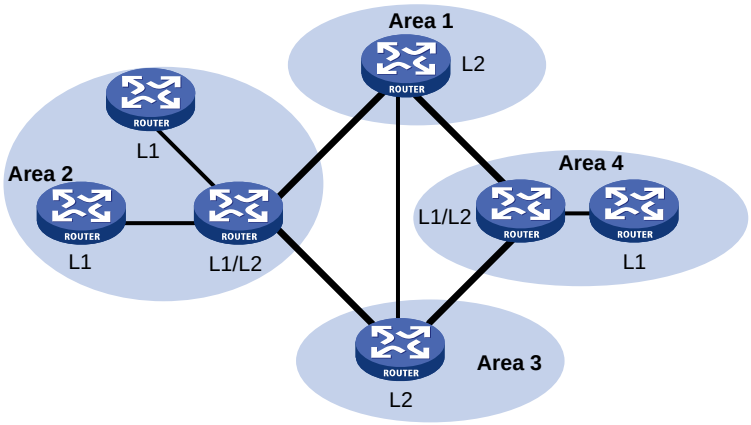



图 1-3 是IS-IS的另外一种拓扑结构图。其中Level-1-2 路由器不仅仅用来连接Level-1 和Level-2 路由器，而且还与其它Level-2 路由器一起构成了IS-IS的骨干网。在这个拓扑中，并没有规定哪个区域是骨干区域。所有Level-2 路由器和Level-1-2 路由器构成了IS-IS的骨干网，它们可以属于不同的区域，但必须是物理连续的。IS-IS的骨干网（Backbone）指的不是一个特定的区域。

图1-3 IS-IS 拓扑结构图之二




IS-IS 不论是 Level-1 还是 Level-2 路由，都采用 SPF 算法，分别生成最短路径树（Shortest Path Tree，SPT）。

### 3. 路由渗透

通常情况下，区域内的路由通过 Level-1 的路由器进行管理。所有的 Level-2 路由器和 Level-1-2 路由器构成一个 Level-2 区域。因此，一个 IS-IS 的路由域可以包含多个 Level-1 区域，但只有一个 Level-2 区域。

Level-1 区域必须且只能与 Level-2 区域相连，不同的 Level-1 区域之间并不相连。

Level-1 区域内的路由信息通过 Level-1-2 路由器发布到 Level-2 区域，因此，Level-2 路由器知道整个 IS-IS 路由域的路由信息。但是，在缺省情况下，Level-2 路由器并不将自己知道的其它 Level-1 区域以及 Level-2 区域的路由信息发布到 Level-1 区域。这样，Level-1 路由器将不了解本区域以外的路由信息，Level-1 路由器只将去往其它区域的报文发送到最近的 Level-1-2 路由器，所以可能导致对本区域之外的目的地址无法选择最佳的路由。

为解决上述问题，IS-IS 提供了路由渗透功能，使 Level-1-2 路由器可以将已知的其它 Level-1 区域以及 Level-2 区域的路由信息发布到指定的 Level-1 区域。

## 1.1.3 IS-IS的网络类型

### 1. 网络类型

IS-IS 只支持两种类型的网络，根据物理链路不同可分为：

- 广播链路：如 Ethernet、Token-Ring 等。
- 点到点链路：如 PPP、HDLC 等。

说明

对于 NBMA（Non-Broadcast Multi-Access）网络，如 ATM，需对其配置子接口，并将子接口类型配置为点到点网络或广播网络。IS-IS 不能在点到多点（Point to MultiPoint，P2MP）链路上运行。

### 2. DIS和伪节点

在广播网络中，IS-IS 需要在所有的路由器中选举一个路由器作为 DIS。

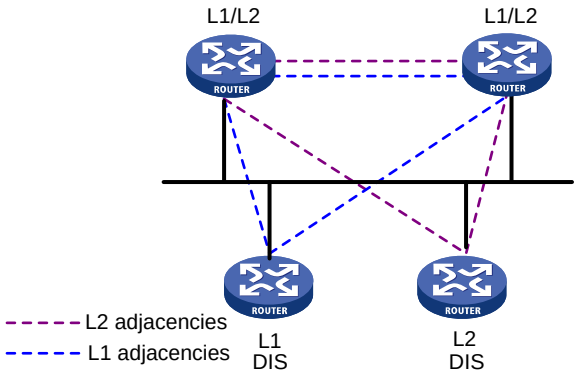
Level-1 和 Level-2 的 DIS 是分别选举的，用户可以为不同级别的 DIS 选举设置不同的优先级。DIS 优先级数值越高，被选中的可能性就越大。如果优先级最高的路由器有多台，则其中 SNPA（Subnetwork Point of Attachment，子网连接点）地址（广播网络中的 SNPA 地址是 MAC 地址）最大的路由器会被选中。不同级别的 DIS 可以是同一台路由器，也可以是不同的路由器。

与 OSPF 的不同点：

- 优先级为 0 的路由器也参与 DIS 的选举；
- 当有新的路由器加入，并符合成为 DIS 的条件时，这个路由器会被选中成为新的 DIS，此更改会引起一组新的 LSP 泛洪。

在IS-IS广播网中，同一网段上的同一级别的路由器之间都会形成邻接关系，包括所有的非DIS路由器之间也会形成邻接关系。如 图 1-4 所示。

图1-4 IS-IS 广播网的 DIS 和邻接关系




DIS 用来创建和更新伪节点（Pseudonodes），并负责生成伪节点的 LSP，用来描述这个网络上有哪些路由器。

伪节点是用来模拟广播网络的一个虚拟节点，并非真实的路由器。在 IS-IS 中，伪节点用 DIS 的 System ID 和一个字节的 Circuit ID（非 0 值）标识。

使用伪节点可以简化网络拓扑，减少 SPF 的资源消耗。

说明

IS-IS 广播网络上所有的路由器之间都形成邻接关系，但 LSDB 的同步仍然依靠 DIS 来保证。

### 1.1.4 IS-IS报文

#### 1. PDU

IS-IS报文是直接封装在数据链路层的帧结构中的。PDU（Protocol Data Unit，协议数据单元）可以分为两个部分，报文头和变长字段部分。其中报文头又可分为通用报头和专用报头。对于所有PDU来说，通用报头都是相同的，但专用报头根据PDU类型不同而有所差别，如 图 1-5 所示。

图1-5 PDU 格式


表1-1 PDU 类型对应关系表

| 类型值 | PDU 类型 | 简称 |
|---|---|---|
| 15 | Level-1 LAN IS-IS Hello PDU | L1 LAN IIH |
| 16 | Level-2 LAN IS-IS Hello PDU | L2 LAN IIH |
| 17 | Point-to-Point IS-IS Hello PDU | P2P IIH |
| 18 | Level-1 Link State PDU | L1 LSP |
| 20 | Level-2 Link State PDU | L2 LSP |

| 类型值 | PDU 类型 | 简称 |
|---|---|---|
| 24 | Level-1 Complete Sequence Numbers PDU | L1 CSNP |
| 25 | Level-2 Complete Sequence Numbers PDU | L2 CSNP |
| 26 | Level-1 Partial Sequence Numbers PDU | L1 PSNP |
| 27 | Level-2 Partial Sequence Numbers PDU | L2 PSNP |

### 2. Hello报文

Hello 报文：用于建立和维持邻居关系，也称为 IIH（IS-to-IS Hello PDUs）。其中，广播网中的 Level-1 路由器使用 Level-1 LAN IIH，广播网中的 Level-2 路由器使用 Level-2 LAN IIH，点到点网络中的路由器则使用 P2P IIH。

### 3. LSP报文

LSP 报文：用于交换链路状态信息。LSP 分为两种：Level-1 LSP 和 Level-2 LSP。Level-1 路由器传送 Level-1 LSP，Level-2 路由器传送 Level-2 LSP，Level-1-2 路由器则可传送以上两种 LSP。

### 4. SNP报文

SNP（Sequence Number PDUs，时序报文）通过描述全部或部分数据库中的 LSP 来同步 LSDB，从而维护 LSDB 的完整和同步。

SNP 包括 CSNP（Complete SNP，全时序报文）和 PSNP（Partial SNP，部分时序报文），进一步又可分为 Level-1 CSNP、Level-2 CSNP、Level-1 PSNP 和 Level-2 PSNP。

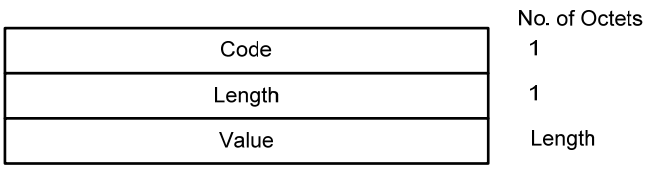
CSNP 包括 LSDB 中所有 LSP 的概要信息，从而可以在相邻路由器间保持 LSDB 的同步。在广播网络上，CSNP 由 DIS 定期发送（缺省的发送周期为 10 秒）；在点到点链路上，CSNP 只在第一次建立邻接关系时发送。

PSNP 只列举最近收到的一个或多个 LSP 的序列号，它能够一次对多个 LSP 进行确认。当发现 LSDB 不同步时，也用 PSNP 来请求邻居发送新的 LSP。

### 5. CLV

PDU中的变长字段部分是多个CLV（Code-Length-Value）三元组。其格式如 图 1-6 所示：

图1-6 CLV 格式




不同PDU类型所包含的CLV是不同的，如 表 1-2 所示。

表1-2 PDU 类型和包含的 CLV 名称

| CLV Code | 名称 | 所应用的 PDU 类型 |
|---|---|---|
| 1 | Area Addresses | IIH、LSP |
| 2 | IS Neighbors（LSP） | LSP |
| 4 | Partition Designated Level-2 IS | L2 LSP |

| CLV Code | 名称 | 所应用的 PDU 类型 |
|---|---|---|
| 6 | IS Neighbors（MAC Address） | LAN IIH |
| 7 | IS Neighbors（SNPA Address） | LAN IIH |
| 8 | Padding | IIH |
| 9 | LSP Entries | SNP |
| 10 | Authentication Information | IIH、LSP、SNP |
| 128 | IP Internal Reachability Information | LSP |
| 129 | Protocols Supported | IIH、LSP |
| 130 | IP External Reachability Information | L2 LSP |
| 131 | Inter-Domain Routing Protocol Information | L2 LSP |
| 132 | IP Interface Address | IIH、LSP |
| 222 | MT-ISN | LSP |
| 229 | M-Topologies | IIH,、LSP |
| 235 | MT IP. Reach | LSP |
| 237 | MT IPv6 IP. Reach | LSP |

其中，Code 值从 1 到 10 的 CLV 在 ISO 10589 中定义（有 2 类未在上表中列出），128 到 132 的 CLV 在 RFC 1195 中定义，多拓扑相关 CLV 在 RFC 5120 中定义。

### 1.1.5 协议规范

与 IS-IS 相关的协议规范有：

- ISO 10589：ISO IS-IS Routing Protocol
- ISO 9542：ES-IS Routing Protocol
- ISO 8348：Ad2 Network Services Access Points
- RFC 1195：Use of OSI IS-IS for Routing in TCP/IP and Dual Environments
- RFC 2763：Dynamic Hostname Exchange Mechanism for IS-IS
- RFC 2966：Domain-wide Prefix Distribution with Two-Level IS-IS
- RFC 2973：IS-IS Mesh Groups
- RFC 3277：IS-IS Transient Blackhole Avoidance
- RFC 3358：Optional Checksums in ISIS
- RFC 3373：Three-Way Handshake for IS-IS Point-to-Point Adjacencies
- RFC 3567：Intermediate System to Intermediate System (IS-IS) Cryptographic Authentication
- RFC 3719：Recommendations for Interoperable Networks using IS-IS
- RFC 3786：Extending the Number of IS-IS LSP Fragments Beyond the 256 Limit
- RFC 3787：Recommendations for Interoperable IP Networks using IS-IS
- RFC 3847：Restart signaling for IS-IS

- RFC 4444：Management Information Base for Intermediate System to Intermediate System (IS-IS)
- RFC 5120：Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)
- RFC 5303：Three-Way Handshake for IS-IS Point-to-Point Adjacencies
- RFC 5310：IS-IS Generic Cryptographic Authentication

## 1.2 IS-IS配置任务简介

表1-3 IS-IS 配置任务简介

| 配置任务 | | 说明 | 详细配置 |
|---|---|---|---|
| 配置IS-IS基本功能 | 使能IS-IS功能 | 必选 | 1.3.2 |
| | 配置路由器的Level级别和接口的链路邻接关系类型 | 可选 | 1.3.3 |
| | 配置接口网络类型 | 可选 | 1.3.4 |
| 配置IS-IS路由信息控制 | 配置IS-IS链路开销 | 可选 | 1.4.2 |
| | 配置IS-IS路由优先级 | 可选 | 1.4.3 |
| | 配置IS-IS最大等价路由条数 | 可选 | 1.4.4 |
| | 配置IS-IS路由聚合 | 可选 | 1.4.5 |
| | 配置IS-IS发布缺省路由 | 可选 | 1.4.6 |
| | 配置IS-IS引入外部路由 | 可选 | 1.4.7 |
| | 配置IS-IS路由过滤 | 可选 | 1.4.8 |
| | 配置IS-IS路由渗透 | 可选 | 1.4.9 |
| 调整和优化IS-IS网络 | 配置Hello报文发送时间间隔 | 可选 | 1.5.2 |
| | 配置Hello报文失效数目 | 可选 | 1.5.3 |
| | 配置CSNP报文发送时间间隔 | 可选 | 1.5.4 |
| | 配置DIS优先级 | 可选 | 1.5.5 |
| | 配置在PPP接口上建立邻接关系必须在同一网段的检查功能 | 可选 | 1.5.6 |
| | 禁止接口发送和接收IS-IS报文 | 可选 | 1.5.7 |
| | 配置接口发送小型Hello报文 | 可选 | 1.5.8 |
| | 配置LSP参数 | 可选 | 1.5.9 |
| | 配置SPF参数 | 可选 | 1.5.10 |
| | 配置优先级参数 | 可选 | 1.5.11 |
| | 配置LSDB过载标志位 | 可选 | 1.5.12 |
| | 配置ATT连接位 | 可选 | 1.5.13 |

| 配置任务 | | 说明 | 详细配置 |
| --- | --- | --- | --- |
| | 配置接口的Tag值 | 可选 | 1.5.14 |
| | 配置IS-IS主机名映射 | 可选 | 1.5.15 |
| | 配置邻接状态变化的输出开关 | 可选 | 1.5.16 |
| | 配置IS-IS ISPF | 可选 | 1.5.17 |
| | 配置前缀抑制 | 可选 | 1.5.18 |
| | 配置IS-IS网管功能 | 可选 | 1.5.19 |
| | 配置PIC | 可选 | 1.5.20 |
| 提高IS-IS网络的安全性 | 配置邻居关系验证 | 可选 | 1.6.2 |
| | 配置区域验证 | 可选 | 1.6.3 |
| | 配置路由域验证 | 可选 | 1.6.4 |
| 配置IS-IS GR | | 可选 | 1.7 |
| 配置IS-IS NSR | | 可选 | 1.8 |
| 配置IS-IS与BFD联动 | | 可选 | 1.9 |
| 配置IS-IS快速重路由 | | 可选 | 1.10 |
| 配置IS-IS支持IPv4单播拓扑 | | 可选 | 1.11 |

# 1.3 配置IS-IS基本功能

## 1.3.1 配置准备

在配置 IS-IS 基本功能之前，需完成以下任务：

- 配置链路层协议
- 配置接口的网络层地址，使相邻节点网络层可达

## 1.3.2 使能IS-IS功能

表1-4 使能 IS-IS 功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 创建一个IS-IS进程，并进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 缺省情况下，系统没有运行任何IS-IS进程 |
| 配置网络实体名称 | **network-entity** *net* | 缺省情况下，没有配置NET |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置指定接口上使能IS-IS路由进程 | **isis enable** [ *process-id* ] | 缺省情况下，IS-IS功能在接口上处于关闭状态，且没有任何IS-IS进程与其关联 |

### 1.3.3 配置路由器的Level级别和接口的链路邻接关系类型

建议用户在配置 IS-IS 时配置路由器类型：

- 如果只有一个区域，建议用户将所有路由器设置为 Level-1 或者 Level-2，因为没有必要让所有路由器同时维护两个完全相同的 LSDB。
- 在 IP 网络中使用时，建议将所有的路由器都设置为 Level-2，这样有利于以后的扩展。

当路由器类型是 Level-1（Level-2）时，接口的链路邻接类型只能为 Level-1（Level-2），当路由器类型是 Level-1-2 时，接口的链路邻接类型缺省为 Level-1-2，当路由器只需要与对端建立 Level-1（Level-2）的邻接关系时，可以将接口的链路邻接类型配置为 Level-1（Level-2）来限制接口上所能建立的邻接关系，如Level-1的接口只能建立Level-1的邻接关系，Level-2的接口只能建立Level-2的邻接关系，让接口只发送和接收 Level-1（Level-2）类型的 Hello 报文，既减少了路由器的处理时间又节省了带宽。

**表1-5 配置路由器的 Level 级别和接口的链路邻接关系类型**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置路由器的Level级别 | **is-level** { **level-1** \| **level-1-2** \| **level-2** } | 缺省情况下，路由器的Level级别为Level-1-2 |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的链路邻接关系类型 | **isis circuit-level** [ **level-1** \| **level-1-2** \| **level-2** ] | 缺省情况下，接口既可以建立Level-1的邻接关系，也可以建立Level-2的邻接关系 |

### 1.3.4 配置接口网络类型

接口网络类型不同，其工作机制也略微不同，如：当网络类型为广播网时，需要选举 DIS、通过泛洪 CSNP 报文来实现 LSDB 同步；当网络类型为 P2P 时，不需要选举 DIS，LSDB 同步机制也不同。

当只有两台路由器接入到同一个广播网时，通过将接口网络类型配置为 P2P 可以使 IS-IS 按照 P2P 而不是广播网的工作机制运行，避免 DIS 选举以及 CSNP 的泛洪，既可以节省网络带宽，又可以加快网络的收敛速度。

表1-6 配置接口网络类型

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的网络类型为P2P | **isis circuit-type p2p** | 缺省情况下，路由器接口网络类型根据物理接口决定，交换机VLAN接口网络类型为Broadcast<br>仅当接口的网络类型为广播网，且只有两台路由器接入该广播网时才需要进行该项配置，并且两台路由器都要进行此项配置 |

# 1.4 配置IS-IS路由信息控制

## 1.4.1 配置准备

在配置 IS-IS 路由信息控制之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点网络层可达
- 使能 IS-IS 功能

## 1.4.2 配置IS-IS链路开销

IS-IS 有三种方式来配置接口的链路开销值，按照选择顺序依次为：

- 在接口视图下为指定接口配置的链路开销值。
- 在系统视图下全局配置的链路开销值，该配置将对该 IS-IS 进程关联的接口同时生效。
- 自动计算开销值：将根据带宽参考值自动计算接口的链路开销值。当开销值的类型为 **wide** 或 **wide-compatible** 时，可以根据公式“开销=（带宽参考值÷接口期望带宽）×10”计算接口的链路开销值，取值范围为 1～16777214。当开销值类型为其他类型时，具体情况如下：接口带宽≤10Mbps 时，值为 60；接口带宽≤100Mbps 时，值为 50；接口带宽≤155Mbps 时，值为 40；接口带宽≤622Mbps 时，值为 30；接口带宽≤2500Mbps 时，值为 20；接口带宽>2500Mbps 时，值为 10。

如果没有采用上述三种方式中的任一种进行开销值的配置，接口的链路开销值将取系统设置的缺省值 10。

提示

接口期望带宽通过命令 **bandwidth** 进行配置，具体情况请参见接口分册命令参考中的介绍。

### 1. 接口配置IS-IS链路开销值

表1-7 接口配置 IS-IS 链路开销值

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS视图 | | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| （可选）配置IS-IS开销值的类型 | | **cost-style** { **narrow** \| **wide** \| **wide-compatible** \| { **compatible** \| **narrow-compatible** } [ **relax-spf-limit** ] } | 缺省情况下，IS-IS开销值的类型为narrow |
| 退回至系统视图 | | **quit** | - |
| 进入接口视图或接口IPv4单播拓扑视图 | 进入接口视图 | **interface** *interface-type interface-number* | 二者选其一 |
| | 进入接口IPv4单播拓扑视图 | **interface** *interface-type interface-number* | |
| | | **topology ipv4** [ **unicast** ] *topo-name* | |
| 配置IS-IS接口的链路开销值 | | **isis cost** *value* [ **level-1** \| **level-2** ] | 缺省情况下，没有配置IS-IS接口的链路开销值 |

### 2. 全局配置IS-IS链路开销值

表1-8 全局配置 IS-IS 链路开销值

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 全局配置IS-IS的链路开销值 | | **circuit-cost** *value* [ **level-1** \| **level-2** ] | 缺省情况下，没有全局配置IS-IS的链路开销值 |

### 3. 配置IS-IS自动计算链路开销值

表1-9 配置 IS-IS 自动计算链路开销值

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS视图或IS-IS | 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| IPv4单播拓扑视图 | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 使能自动计算接口链路开销值功能 | | **auto-cost enable** | 缺省情况下，自动计算接口链路开销值功能处于关闭状态 |
| （可选）配置IS-IS自动计算链路开销值时依据的带宽参考值 | | **bandwidth-reference** *value* | 缺省情况下，带宽参考值为100Mbps |

### 1.4.3 配置IS-IS路由优先级

一台路由器可同时运行多个路由协议，当多个路由协议都发现到同一目的地的路由时，将选用高优先级路由协议所发现的路由。

以下配置用来为 IS-IS 路由设置优先级，使用路由策略可以为特定的路由设置特定的优先级，路由策略的相关知识请参见“三层技术-IP 路由配置指导”中的“路由策略”。

表1-10 配置 IS-IS 路由优先级

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置IS-IS路由优先级 | | **preference** { *preference* \| **route-policy** *route-policy-name* } * | 缺省情况下，IS-IS路由优先级为15 |

### 1.4.4 配置IS-IS最大等价路由条数

如果到一个目的地有几条开销相同的路径，可以通过等价路由负载分担来提高链路利用率。该配置用以设置 IS-IS 协议的最大等价路由条数。

表1-11 配置 IS-IS 最大等价路由条数

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置在负载分担方式下IS-IS最大等价路由条数 | | **maximum load-balancing** *number* | 缺省情况下，IS-IS支持的最大等价路由条数为32。 |

## 1.4.5 配置IS-IS路由聚合

通过配置路由聚合，可以减小路由表规模，还可以减少本路由器生成的 LSP 报文大小和 LSDB 的规模。其中，被聚合的路由可以是 IS-IS 协议发现的路由，也可以是引入的外部路由。

路由器只对本地生成的 LSP 中的路由进行聚合。

表1-12 配置 IS-IS 路由聚合

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置聚合路由 | | **summary** *ip-address* { *mask-length* \| *mask* } [ **avoid-feedback** \| **generate_null0_route** \| [ **level-1** \| **level-1-2** \| **level-2** ] \| **tag** *tag* ] * | 缺省情况下，没有对路由进行聚合<br>聚合后路由的开销值取所有被聚合路由中最小的开销值 |

### 1.4.6 配置IS-IS发布缺省路由

对于运行 IS-IS 的路由器来说，无法引入缺省路由，因此也无法通过将目的地为 0.0.0.0/0 的路径信息（即缺省路由）通过 LSP 发布给其它路由器，可以通过配置发布一条缺省路由，将目的地为 0.0.0.0/0 的路径信息通过 LSP 发布出去，其它同级别的路由器中将在自己的路由表中新增一条缺省路由。

表1-13 配置 IS-IS 发布缺省路由

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置IS-IS发布缺省路由 | | **default-route-advertise** [ [ **level-1** \| **level-1-2** \| **level-2** ] \| **route-policy** *route-policy-name* ] * | 缺省情况下，此功能关闭<br>产生的缺省路由只被发布到同级别的路由器 |

### 1.4.7 配置IS-IS引入外部路由

IS-IS 将其它路由协议发现的路由当作外部路由处理。在引入其它协议路由时，可指定引入路由的缺省开销。还可以通过配置对引入路由进行过滤。

在实际组网环境中，每台路由器的性能即处理能力不同，如果在处理能力强的高端设备上引入大量外部路由，那么可能会对网络上其它低端设备的性能造成较大的冲击，网络管理员可以通过配置支持的最大引入路由条数，限制引入外部路由的条数，从而最终限制发布路由的数量。

表1-14 配置 IS-IS 引入外部路由

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| | 单播拓扑视图 | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 从其它路由协议或其它IS-IS进程引入路由信息 | | **import-route** *protocol* [ *process-id* \| **all-processes** \| **allow-ibgp** ] [ **allow-direct** \| **cost** *cost* \| **cost-type** { **external** \| **internal** } \| [ **level-1** \| **level-1-2** \| **level-2** ] \| **route-policy** *route-policy-name* \| **tag** *tag* ] * | 缺省情况下，IS-IS不引入其它协议的路由信息<br>如果**import-route**命令中不指定引入的级别，则默认为引入路由到Level-2路由表中<br>只能引入路由表中状态为active的路由，是否为active状态可以通过**display ip routing-table protocol**命令来查看 |
| （可选）配置引入Level1/Level2的IPv4路由最大条数 | | **import-route limit** *number* | （number 的缺省情况如下）<br>MSR 2600：200000<br>MSR 36-10、MSR 3600-28/ MSR 3600-51：200000<br>其他MSR 3600：500000<br>MSR 5600：1000000 |

## 1.4.8 配置IS-IS路由过滤

路由过滤就是通过对 ACL、IP 地址前缀列表或路由策略等规则的引用对路由信息的生成进行更加严格的控制，包括对接收的路由是否加入 IP 路由表进行过滤和对引入的路由信息进行过滤。

### 1. 配置IS-IS对接收的路由是否加入IP路由表进行过滤

运行 IS-IS 的路由器会把从邻居收到的 LSP 保存到自己维护的链路状态数据库中，使用 SPF 算法计算出以自己为根的最短路径树，并把计算好的路由信息加入到 IS-IS 路由表中，最终把最优路由加入到 IP 路由表中。

通过 ACL、IP 地址前缀列表或路由策略可以对将要加入到 IP 路由表中的路由进行过滤，满足条件则加入到 IP 路由表中，否则将不能加入到 IP 路由表中。没有加入 IP 路由表的路由仍然在 IS-IS 路由表中，可以通过 LSP 发布出去。

**表1-15 配置 IS-IS 对接收的路由是否加入 IP 路由表进行过滤**

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 播拓扑视图 | 拓扑视图 | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置IS-IS对接收的路由是否加入IP路由表进行过滤 | | **filter-policy** { *acl-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } **import** | 缺省情况下，没有配置该过滤功能 |

### 2. 配置IS-IS对引入的路由信息进行过滤

IS-IS 可以从其它路由协议或其它 IS-IS 进程引入路由信息，把它直接加入到 IS-IS 的路由表中并通过 LSP 发布出去。

通过 ACL、IP 地址前缀列表或路由策略可以对引入的路由信息进行过滤，满足条件加入到 IS-IS 路由表中，否则将不能加入到 IS-IS 路由表中。没有加入 IS-IS 路由表的路由将不会通过 LSP 发布出去。

表1-16 配置 IS-IS 对引入的路由信息进行过滤

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置IS-IS对引入的路由信息进行过滤 | | **filter-policy** { *acl-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } **export** [ **protocol** [ *process-id* ] ] | 缺省情况下，没有配置该过滤功能 |

## 1.4.9 配置IS-IS路由渗透

通过 IS-IS 路由渗透功能（Level-2 to Level-1），可以将 Level-2 级别的路由信息和其他区域的 level-1 级别的路由信息渗透到 Level-1 区域。

通过控制 IS-IS 路由渗透（Level-1 to Level-2），可以控制 Level-1 区域的 IS-IS 路由信息不向 Level-2 渗透，达到有效控制 Level-2 级别的路由信息的目的。

表1-17 配置 IS-IS 路由渗透

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置将Level-1区域的路由信息引入到Level-2区域 | | **import-route isis level-1 into level-2** [ **filter-policy** { *acl-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } \| **tag** *tag* ] * | 缺省情况下，Level-1区域的路由信息向Level-2区域发布 |
| 配置将Level-2区域的路由信息引入到Level-1区域 | | **import-route isis level-2 into level-1** [ **filter-policy** { *acl-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } \| **tag** *tag* ] * | 缺省情况下，Level-2区域的路由信息不向Level-1区域发布 |

# 1.5 调整和优化IS-IS网络

## 1.5.1 配置准备

在配置 IS-IS 调整和优化之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点网络层可达
- 使能 IS-IS 功能

## 1.5.2 配置Hello报文发送时间间隔

如果路由器在邻居关系保持时间内（即 Hello 报文失效数目与 Hello 报文发送时间间隔的乘积）没有收到来自邻居路由器的 Hello 报文时将宣告邻居关系失效。通过设置 Hello 报文失效数目和 Hello 报文的发送时间间隔，可以调整邻居关系保持时间，即邻居路由器要花多长时间能够监测到链路已经失效并重新进行路由计算。

表1-18 配置 Hello 报文发送时间间隔

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置Hello报文的发送时间间隔 | **isis timer hello** *seconds* [ **level-1** \| **level-2** ] | 缺省情况下，Hello报文的发送时间间隔为10秒<br>DIS发送Hello报文的时间间隔是**isis timer hello**设置的时间的1/3 |

### 1.5.3 配置Hello报文失效数目

Hello 报文失效数目，即宣告邻居失效前 IS-IS 没有收到的邻居 Hello 报文的数目。

如果路由器在邻居关系保持时间内（即 Hello 报文失效数目与 Hello 报文发送时间间隔的乘积）没有收到来自邻居路由器的 Hello 报文时将宣告邻居关系失效。通过设置 Hello 报文失效数目和 Hello 报文的发送时间间隔，可以调整邻居关系保持时间，即邻居路由器要花多长时间能够监测到链路已经失效并重新进行路由计算。

在广播链路上，Level-1 和 Level-2 Hello 报文会分别发送，Hello 报文失效数目需要分别设置；在点到点链路中，Level-1 和 Level-2 的 Hello 报文是在同一个点到点 Hello 报文中发送，因此不需要指定 Level-1 或 Level-2。

**表1-19 配置 Hello 报文失效数目**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置Hello报文失效数目 | **isis timer holding-multiplier** *value* [ **level-1** \| **level-2** ] | 缺省情况下，Hello报文失效数目为3 |

### 1.5.4 配置CSNP报文发送时间间隔

当网络类型为广播网时，DIS 使用 CSNP 报文来进行 LSDB 同步，因此只有在被选举为 DIS 的路由器上进行该项配置才有效。

**表1-20 配置 CSNP 报文发送时间间隔**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置DIS在广播网络上发送CSNP报文的时间间隔 | **isis timer csnp** *seconds* [ **level-1** \| **level-2** ] | 缺省情况下，CSNP报文的发送时间间隔为10秒 |

### 1.5.5 配置接口的DIS优先级

在广播网络中，IS-IS 需要在所有的路由器中选举一个路由器作为 DIS。

对于 IS-IS，Level-1 和 Level-2 的 DIS 是分别选举的，可以为不同级别的 DIS 选举设置不同的优先级。优先级数值越高，被选中的可能性就越大。如果所有路由器的 DIS 优先级相同，将会选择 MAC 地址最大的路由器作为 DIS。

表1-21 配置接口的 DIS 优先级

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的DIS优先级 | **isis dis-priority** *value* [ **level-1** \| **level-2** ] | 缺省情况下，接口的DIS优先级为64 |

### 1.5.6 配置在PPP接口上建立邻接关系必须在同一网段的检查功能

当接口封装 PPP 协议时，对端的 IP 地址与当前接口不在同一网段也可以建立邻接关系。通过配置与对端路由器建立邻接关系必须在同一网段的检查功能，即在 PPP 协议接口上接收 Hello 报文时，对端的 IP 地址与当前接口必须在同一网段才可以建立邻接关系。

表1-22 配置在 PPP 接口上建立邻接关系必须在同一网段的检查功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置在PPP接口上建立邻接关系必须在同一网段的检查功能 | **isis peer-ip-check** | 缺省情况下，协议类型为PPP的接口要与对端路由器建立邻接关系，双方可以不在同一网段<br>该命令只能在协议类型为PPP的接口上配置 |

### 1.5.7 禁止接口发送和接收IS-IS报文

通过禁止接口发送和接收 IS-IS 报文，禁止了该接口与相邻路由器建立邻居关系，但仍然可以把该接口直连网络的路由信息放在 LSP 中从其它接口宣告出去。由于不用建立邻居关系，可以节省带宽和路由器处理时间，同时，其它路由器也可以知道到达该接口直连网络的路由信息。

表1-23 禁止接口发送和接收 IS-IS 报文

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 禁止接口发送和接收IS-IS报文 | **isis silent** | 缺省情况下，接口既发送也接收IS-IS报文 |

### 1.5.8 配置接口发送小型Hello报文

IS-IS 协议报文直接封装在链路层报文头后面，无法实现协议报文在 IP 层的自动分片。因此，运行 IS-IS 的路由器与对端路由器建立邻居关系时，会发送达到链路 MTU 大小的 Hello 报文，双方进行 MTU 大小的通信协商，来保证建立邻居双方接口 MTU 的一致性，从而避免双方 MTU 大小不一致导致较小的 PDU 可以通过，但是较大的 PDU 无法通过。

当邻居路由器双方 MTU 大小一样的时候，为了避免发送过大的 Hello 报文浪费带宽，可以配置接口发送不加入填充 CLV 的小型 Hello 报文。

表1-24 配置接口发送小型 Hello 报文

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口发送不加入填充CLV的小型Hello报文 | **isis small-hello** | 缺省情况下，接口发送标准Hello报文 |

### 1.5.9 配置LSP参数

#### 1. 配置LSP时间参数

(1) 配置 LSP 最大生存时间

每个 LSP 都有一个最大生存时间，随着时间的推移最大生存时间将逐渐减小，当 LSP 的最大生存时间为 0 时，IS-IS 将启动清除过期 LSP 的过程。用户可根据网络规模对 LSP 的最大生存时间进行调整。

表1-25 配置 LSP 最大生存时间

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置LSP最大生存时间 | **timer lsp-max-age** *seconds* | 缺省情况下，LSP最大生存时间为1200秒 |

(2) 配置 LSP 刷新周期和 LSP 重新生成的时间间隔

路由器必须定时刷新自己生成的 LSP，防止 LSP 的最大生存时间减小为 0。另外，通过定时刷新 LSP 可以使整个区域中的 LSP 保持同步。用户可对 LSP 的刷新周期进行配置，提高 LSP 的刷新频率可以加快网络收敛速度，但是将占用更多的带宽。

除了定时刷新可以重新生成 LSP 外，当网络拓扑发生变化，如邻居路由器 up 或 down，接口 Metric 值、System ID 或区域地址发生变化等，将触发路由器重新生成 LSP。为了防止网络拓扑频繁变化而导致 LSP 频繁重新生成，用户可配置 LSP 生成时间间隔，以抑制网络变化频繁导致占用过多的带宽资源和路由器资源。

本命令在网络变化不频繁的情况下将 LSP 重新生成时间间隔缩小到 *minimum-interval*，而在网络变化频繁的情况下可以进行相应惩罚，增加 *incremental-interval*×$2^{n-2}$（n 为连续触发路由计算的次数），将等待时间按照配置的惩罚增量延长，最大不超过 *maximum-interval*。

表1-26 配置 LSP 刷新周期和 LSP 重新生成的时间间隔

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置LSP刷新周期 | **timer lsp-refresh** *seconds* | 缺省情况下，LSP刷新周期为900秒 |
| 配置LSP重新生成的时间间隔 | **timer lsp-generation** *maximum-interval* [ *minimum-interval* [ *incremental-interval* ] ] [ **level-1** \| **level-2** ] | 缺省情况下，LSP重新生成的最大时间间隔为5秒，最小时间间隔为20毫秒，时间间隔惩罚增量为200毫秒 |

(3) 配置 LSP 发送时间间隔

当 LSDB 的内容发生变化时，IS-IS 将把发生变化的 LSP 扩散出去，用户可以对 LSP 的最小发送时间间隔进行调节。

请合理配置 LSP 发送时间间隔，当存在大量 IS-IS 接口或大量路由时，会发送大量的 LSP 报文，导致 LSP 风暴的出现。

在点到点链路上，发送的 LSP 需要得到对端的应答，否则将在指定的时间间隔内重新发送该 LSP，重传时间间隔决定了当一个 LSP 在 P2P 链路上丢失时它被重传需要等待的时间。

表1-27 配置 LSP 发送时间间隔

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置发送LSP的最小时间间隔以及一次最多可以发送的LSP报文数目 | **isis timer lsp** *time* [ **count** *count* ] | 缺省情况下，LSP的发送最小时间间隔为33毫秒，一次最多可以发送5个LSP报文 |
| 配置LSP在点到点链路上的重传时间间隔 | **isis timer retransmit** *seconds* | 缺省情况下，LSP在点到点链路上的重传时间间隔为5秒 |

## 2. 配置LSP报文长度

IS-IS 协议报文直接封装在链路层报文头后面，无法实现协议报文在 IP 层的自动分片。

为了不影响 LSP 的正常扩散，要求同一区域内所有 IS-IS 路由器生成 LSP 报文的最大长度不能超过该区域内所有路由器 IS-IS 接口 MTU 的最小值。

如果 IS-IS 运行的区域中各 IS-IS 接口的 MTU 值不一致，建议用户对 IS-IS 生成 LSP 报文的最大长度进行配置，将同一区域内所有 IS-IS 路由器生成 LSP 报文的最大长度配置为该区域内所有路由器 IS-IS 接口 MTU 的最小值。如果不进行配置，系统将根据当前设备 IS-IS 接口最小 MTU 值的变化而自动重启 IS-IS 进程动态调整生成 LSP 报文的最大长度，会在一定程度上影响业务的正常运行。

表1-28 配置 LSP 报文长度

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置生成的Level-1 LSP和Level-2 LSP的最大长度 | **lsp-length originate** *size* [ **level-1** \| **level-2** ] | 缺省情况下，生成的Level-1 LSP和Level-2 LSP的最大长度为1497字节 |
| 配置可以接收LSP的最大长度 | **lsp-length receive** *size* | 缺省情况下，接收的LSP报文的最大长度为1497字节 |

### 3. 配置LSP快速扩散功能

通过使能 LSP 快速扩散功能，当 LSP 发生变化而导致 SPF 重新计算时，在 SPF 重新计算前，把导致 SPF 重新计算的 LSP 快速扩散出去，将大大缩短路由器之间由于进行 LSP 同步而导致 LSDB 不一致的时间，提高全网的快速收敛性能。

表1-29 配置 LSP 快速扩散功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置LSP快速扩散功能 | **flash-flood** [ **flood-count** *flooding-count* \| **max-timer-interval** *flooding-interval* \| [ **level-1** \| **level-2** ] ] * | 缺省情况下，禁止LSP快速扩散功能 |

### 4. 配置LSP分片扩展功能

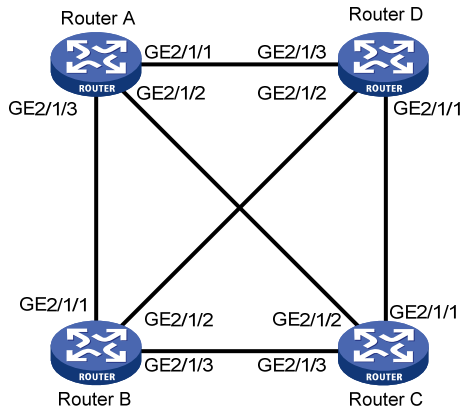
表1-30 配置 LSP 分片扩展功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能IS-IS进程的LSP分片扩展功能 | **lsp-fragments-extend** [ **level-1** \| **level-1-2** \| **level-2** ] | 缺省情况下，LSP分片扩展功能处于关闭状态<br>使能分片扩展功能后，使能该IS-IS进程的所有接口的MTU不能小于512，否则LSP分片扩展功能将不会生效 |
| 配置IS-IS进程的虚拟系统ID | **virtual-system** *virtual-system-id* | 缺省情况下，没有配置IS-IS进程的虚拟系统ID<br>为了使路由器生成扩展LSP分片，应至少配置一个虚拟System ID |

### 5. 限制LSP泛洪

在ATM、FR等NBMA网络中，如果网络联通程度较高、网络中存在多条点到点链路时，如 图 1-7 所示，Router A、Router B、Router C和Router D均使能了IS-IS，Router A新生成一个LSP时，将把该LSP分别从GigabitEthernet2/1/1、GigabitEthernet2/1/2 和GigabitEthernet2/1/3 泛洪出去，Router D从GigabitEthernet2/1/3 收到Router A发送的LSP后，也会把该LSP从GigabitEthernet2/1/1 和 GigabitEthernet2/1/2 发送给Router B和Router C，而Router B和Router C已经从GigabitEthernet2/1/1、GigabitEthernet2/1/2 收到了Router A发送的LSP。LSP的重复扩散会导致带宽的浪费。

图1-7 联通程度较高网络示意图




为了避免这种情况的发生，可以将一些接口配置属于一个 Mesh-Group，也可以配置接口阻塞。

- 将设备的几个接口配置属于一个 Mesh-Group 后，如果从其中的一个接口接收到一个新的 LSP，IS-IS 只把该 LSP 扩散到没有配置 Mesh-Group 的接口以及与当前接口不属于同一个 Mesh-Group 的接口，而不会扩散到同 Mesh-Group 中的其它接口。
- 配置接口阻塞后，只有该接口从邻居路由器收到要求发送 LSP 的请求时才会发送 LSP，否则不会主动向外发送 LSP。

设置接口加入 Mesh-Group 或对接口进行阻塞时应注意保留一定的冗余度，以免由于链路故障影响 LSP 报文的正常扩散。

表1-31 限制 LSP 泛洪

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口属于Mesh-Group | **isis mesh-group** *mesh-group-number* | 二者选其一<br>缺省情况下，接口不属于任何Mesh-Group且接口不阻塞<br>只对点到点类型链路的接口起作用 |
| 配置接口阻塞 | **isis mesh-group mesh-blocked** | |

## 1.5.10 配置SPF参数

根据本地维护的 LSDB，运行 IS-IS 协议的路由器通过 SPF 算法计算出以自己为根的最短路径树，并根据这一最短路径树决定到目的网络的下一跳。通过调节 SPF 的计算间隔，可以抑制网络频繁变化可能导致的占用过多带宽资源和路由器资源。

本命令在网络变化不频繁的情况下将连续路由计算的时间间隔缩小到 *minimum-interval*，而在网络变化频繁的情况下可以进行相应惩罚，增加 *incremental-interval*×$2^{n-2}$（n 为连续触发路由计算的次数），将等待时间按照配置的惩罚增量延长，最大不超过 *maximum-interval*。

**表1-32 配置 SPF 参数**

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS视图或IS-IS IPv4单播地址族视图 | 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置IS-IS路由计算时间间隔 | | **timer spf** *maximum-interval* [ *minimum-interval* [ *incremental-interval* ] ] | 缺省情况下，IS-IS路由计算的最大时间间隔为5秒，最小时间间隔为50毫秒，时间间隔惩罚增量为200毫秒 |

## 1.5.11 配置优先级参数

IS-IS 协议中，当网络拓扑发生变化时，路由要重新收敛。IS-IS 路由收敛的优先级由高到低包括：

- critical：最高优先级。
- high：高优先级。
- medium：中优先级。
- 低优先级：缺省优先级。需要注意的是，IS-IS 主机路由的缺省优先级为中优先级。

IS-IS 路由的优先级越高收敛的速度越快。

**表1-33 配置优先级参数**

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| | 扑视图 | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置指定IS-IS路由收敛的优先级 | | **prefix-priority** { **critical** \| **high** \| **medium** } { **prefix-list** *prefix-list-name* \| **tag** *tag-value* }<br>**prefix-priority route-policy** *route-policy-name* | 缺省情况下，IS-IS路由收敛的优先级为低优先级 |

## 1.5.12 配置LSDB过载标志位

通过配置 LSDB 过载标志位，IS-IS 将在其发送的 LSP 报文中把 OL 位置位，以通知其它路由器当前路由器发生了问题，无法正确的执行路由选择和报文转发。

当运行 IS-IS 的路由器因为内存不足或其它原因无法记录完整的 LSDB 时，将会导致区域路由的计算错误，在故障排除过程中，通过给怀疑有问题的路由器设置过载标志位，可以将其从 IS-IS 网络中暂时隔离，便于进行故障定位。

表1-34 配置 LSDB 过载标志位

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 设置过载标志位 | | **set-overload** [ **on-startup** [ [ **start-from-nbr** *system-id* [ *timeout1* [ *nbr-timeout* ] ] ] \| *timeout2* ] [ **allow** { **external** \| **interlevel** } * ] | 缺省情况下，不设置过载标志位 |

## 1.5.13 配置ATT连接位

ATT 连接位由 L1/L2 路由器产生，但仅与 L1 LSP 有关，表示产生此 LSP 的路由器（L1/L2 路由器）与多个区域相连接。

### 1. 配置IS-IS不采用ATT位计算缺省路由

表1-35 配置 IS-IS 不采用 ATT 位计算缺省路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS不采用ATT位计算缺省路由 | **ignore-att** | 缺省情况下，IS-IS采用ATT位计算缺省路由 |

2. 设置系统自身发布的Level-1 LSP的ATT位

表1-36 设置系统自身发布的 Level-1 LSP 的 ATT 位

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 设置系统自身发布的Level-1 LSP的ATT位 | **set-att** { **always** \| **never** } | 缺省情况下，没有设置系统自身发布的Level-1 LSP的ATT位 |

## 1.5.14 配置接口的Tag值

当 cost-sytle 为 wide、wide-compatible 或 compatible 时，如果发布可达的 IP 地址前缀具有 tag 属性，IS-IS 会将 tag 加入到该前缀的 IP 可达信息 TLV 中。

表1-37 配置接口的 Tag 值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的Tag值 | **isis tag** *tag* | 缺省情况下，没有配置接口的Tag值 |

## 1.5.15 配置IS-IS主机名映射

IS-IS 用 System ID 来在区域内唯一标识主机或路由器，System ID 长度固定为 6 字节。当网络管理员检查 IS-IS 邻居关系的状态、IS-IS 路由表以及 LSDB 中的内容时，十六进制表示的 System ID 以及 LSP 标识符不够直观，查看也不方便。

主机名映射提供了一种将 System ID 映射到主机名的服务，运行 IS-IS 的路由器维护一个主机名到 System ID 的映射关系表，在维护和管理以及网络故障诊断时，使用主机名比使用 System ID 会更直观，也更容易记忆。

可以通过静态配置和动态生成两种方式生成和维护此关系映射表，需要注意的是：

- 只有使能动态主机名映射功能后，使用 **display isis lsdb** 等命令才可以看到路由器的主机名而不是 System ID。

- 倘若网络中的一台路由器使能了动态主机名映射功能且在当前路由器也通过静态方式为那台路由器配置了主机名，动态配置的主机名将覆盖当前路由器为其静态配置的主机名称。

### 1. 配置IS-IS静态主机名映射

网络管理员为远端 IS 手工配置 System ID 与主机名称的映射关系。

表1-38 配置 IS-IS 静态主机名映射

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 为远端IS配置System ID与主机名称的映射关系 | **is-name map** *sys-id map-sys-name* | 每个System ID只能对应一个主机名称 |

### 2. 配置IS-IS动态主机名映射

静态配置关系映射表要求网络中的每一台路由器为其它路由器配置 System ID 和主机名的映射关系，当网络中路由器数目增多时，网络中每新增一台路由器或修改某台路由器的主机名映射关系，其它路由器都要做相应配置，增加了维护工作量。

使能动态主机名映射功能后，IS-IS 网络中的每台路由器只需要在本机上配置自己的主机名称即可，配置的主机名称将通过动态主机名 CLV 发布出去，最后 IS-IS 网络中使能动态主机名映射功能的路由器都将收集到其它路由器 System ID 与主机名称的映射关系并生成映射表。

同时还可以为广播网中的 DIS 配置局域网名称来代表这个广播网中的伪节点，便于网络管理员查看 LSDB 内容时判断 LSP 是由哪个 DIS 产生的。

表1-39 配置 IS-IS 动态主机名映射

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能动态主机名映射功能并为当前路由器配置主机名称 | **is-name** *sys-name* | 缺省情况下，动态主机名映射功能处于关闭状态且没有为当前路由器配置主机名称 |
| 退回至系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置本地局域网名称 | **isis dis-name** *symbolic-name* | 缺省情况下，没有配置本地局域网名称<br>该命令只有在使能了动态主机名进程的路由器上有效。该命令在点到点链路的接口上无效 |

### 1.5.16 配置邻接状态变化的输出开关

打开邻接状态输出开关后，IS-IS 邻接状态变化时会生成日志信息发送到设备的信息中心，通过设置信息中心的参数，最终决定日志信息的输出规则（即是否允许输出以及输出方向）。（有关信息中心参数的配置请参见“网络管理和监控配置指导”中的“信息中心”。）

表1-40 配置邻接状态变化的输出开关

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 打开邻接状态变化的输出开关 | **log-peer-change** | 缺省情况下，邻接状态变化的输出开关处于打开状态 |

### 1.5.17 配置IS-IS ISPF

ISPF（Incremental Shortest Path First，增量最短路径优先）计算是对 IS-IS 中最短路径树的增量计算，当网络的拓扑结构发生变化，即影响到最短路径树的结构时，只对受影响的部分节点进行重新计算拓扑结构，对最短路径树中受影响的部分进行修正，而不需要重建整棵最短路径树。

表1-41 配置 IS-IS ISPF

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 使能IS-IS ISPF功能，即增量SPF计算功能 | | **ispf enable** | 缺省情况下，使能IS-IS ISPF功能 |

### 1.5.18 配置前缀抑制

接口上配置本功能后，禁止此接口的前缀在 LSP 中携带，屏蔽内部节点被发布，提高安全性，加快路由收敛。

表1-42 配置前缀抑制

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的前缀抑制功能 | **isis prefix-suppression** | 缺省情况下，未配置接口的前缀抑制功能<br>本命令对接口子地址同样生效 |

## 1.5.19 配置IS-IS网管功能

配置 IS-IS 进程绑定 MIB 功能后，可以通过网管软件对指定的 IS-IS 进程进行管理。

开启 IS-IS 模块的告警功能后，该模块会生成告警信息，用于报告该模块的重要事件。生成的告警信息将发送到设备的 SNMP 模块，通过设置 SNMP 中告警信息的发送参数，来决定告警信息输出的相关属性。（有关告警信息的详细介绍，请参见“网络管理和监控配置指导”中的“SNMP”。）

TRILL 使用 IS-IS 的 MIB（Management Information Base，管理信息库）对 NMS（Network Management System，网络管理系统）提供 TRILL 对象的管理，但标准 IS-IS MIB 中定义的 MIB 为单实例管理对象，无法同时对 IS-IS 和 TRILL 进行管理。因此，参考 RFC 4750 中对 OSPF 多实例的管理方法，为管理 TRILL 定义一个上下文名称，以区分来自 NMS 的 SNMP 请求是要对 IS-IS 还是 TRILL 进行管理。需要注意的是，由于上下文名称只是 SNMPv3 独有的概念，因此对于 SNMPv1/v2c，会将团体名映射为上下文名称以对不同协议进行区分。

表1-43 配置 IS-IS 网管功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置IS-IS进程绑定MIB | **isis mib-binding** *process-id* | 缺省情况下，MIB绑定在进程号最小的IS-IS进程上 |
| 开启IS-IS的告警功能 | **snmp-agent trap enable isis** [ **adjacency-state-change** \| **area-mismatch** \| **authentication** \| **authentication-type** \| **buffsize-mismatch** \| **id-length-mismatch** \| **lsdboverload-state-change** \| **lsp-corrupt** \| **lsp-parse-error** \| **lsp-size-exceeded** \| **manual-address-drop** \| **max-seq-exceeded** \| **maxarea-mismatch** \| **own-lsp-purge** \| **protocol-support** \| **rejected-adjacency** \| **skip-sequence-number** \| **version-skew** ] * | 缺省情况下，IS-IS的告警功能处于开启状态 |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置管理IS-IS的SNMP实体所使用的上下文名称 | **snmp context-name** *context-name* | 缺省情况下，没有配置管理IS-IS的SNMP实体所使用的上下文名称 |

### 1.5.20 配置PIC

提示

- PIC 和 IS-IS 快速重路由功能同时配置时，IS-IS 快速重路由功能生效。
- 邻居发送的 LSP 才会进行 PIC。

PIC（Prefix Independent Convergence，前缀无关收敛），即收敛时间与前缀数量无关，加快收敛速度。传统的路由计算快速收敛都与前缀数量相关，收敛时间与前缀数量成正比。

#### 1. 配置PIC

表1-44 配置 PIC

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能前缀无关收敛功能 | **pic** [ **additional-path-always** ] | 缺省情况下，使能前缀无关收敛功能 |

#### 2. 配置PIC支持BFD检测功能

IS-IS 协议的 PIC 特性中，主用链路缺省不使用 BFD 进行链路故障检测。配置本功能后，将使用 BFD（Echo 方式）进行检测，可以加快 IS-IS 协议的收敛速度。

表1-45 配置 PIC 支持 BFD 检测功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置BFD Echo报文源地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置BFD Echo报文源地址 |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能IS-IS协议中主用链路的BFD（Echo方式）检测功能 | **isis primary-path-detect bfd echo** | 缺省情况下，IS-IS协议中主用链路的BFD（Echo方式）检测功能处于关闭状态 |

## 1.6 提高IS-IS网络的安全性

在安全性要求较高的网络中，可以通过配置 IS-IS 验证来提高 IS-IS 网络的安全性。IS-IS 验证特性分为邻居关系的验证和区域或路由域的验证。

### 1.6.1 配置准备

在配置 IS-IS 验证功能之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点网络层可达
- 使能 IS-IS 功能

### 1.6.2 配置邻居关系验证

配置邻居关系验证后，验证密码将会按照设定的方式封装到 Hello 报文中，并对接收到的 Hello 报文进行验证密码的检查，通过检查才会形成邻居关系，否则将不会形成邻居关系，用以确认邻居的正确性和有效性，防止与无法信任的路由器形成邻居。

两台路由器要形成邻居关系必须配置相同的验证方式和验证密码。

切换密码时可以通过配置发送报文携带验证信息，接收报文时不进行验证实现认证密码无缝切换。

**表1-46 配置邻居关系验证**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置邻居关系验证方式和验证密码 | **isis authentication-mode** { **md5** \| **simple** \| **gca** *key-id* { **hmac-sha-1** \| **hmac-sha-224** \| **hmac-sha-256** \| **hmac-sha-384** \| **hmac-sha-512** } } { **cipher** *cipher-string* \| **plain** *plain-string* } [ **level-1** \| **level-2** ] [ **ip** \| **osi** ] | 缺省情况下，接口没有配置邻居关系验证方式和验证密码 |
| （可选）配置对收到的Hello报文忽略认证信息检查 | **isis authentication send-only** [ **level-1** \| **level-2** ] | 缺省情况下，如果配置了接口验证方式和验证密码，对收到的报文执行认证信息检查 |

### 1.6.3 配置区域验证

通过配置区域验证，可以防止将从不可信任的路由器学习到的路由信息加入到本地 Level-1 的 LSDB 中。

配置区域验证后，验证密码将会按照设定的方式封装到 Level-1 报文（LSP、CSNP、PSNP）中，并对收到的 Level-1 报文进行验证密码的检查。

同一区域内的路由器必须配置相同的验证方式和验证密码。

切换密码时可以通过配置发送报文携带验证信息，接收报文时不进行验证实现认证密码无缝切换。

**表1-47 配置区域验证**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置区域验证方式和验证密码 | **area-authentication-mode** { **md5** \| **simple** \| **gca** *key-id* { **hmac-sha-1** \| **hmac-sha-224** \| **hmac-sha-256** \| **hmac-sha-384** \| **hmac-sha-512** } } { **cipher** *cipher-string* \| **plain** *plain-string* } [ **ip** \| **osi** ] | 缺省情况下，系统没有配置区域验证方式和验证密码 |
| （可选）配置对收到的Level-1报文（包括LSP、CSNP、PSNP）忽略认证信息检查 | **area-authentication send-only** | 缺省情况下，如果配置了区域验证方式和验证密码，对收到的报文执行认证信息检查 |

### 1.6.4 配置路由域验证

通过配置路由域验证，可以防止将不可信的路由信息注入当前路由域。

配置路由域验证后，验证密码将会按照设定的方式封装到 Level-2 报文（LSP、CSNP、PSNP）中，并对收到的 Level-2 报文进行验证密码的检查。

所有骨干层（Level-2）路由器必须配置相同的验证方式和验证密码。

切换密码时可以通过配置发送报文携带验证信息，接收报文时不进行验证实现认证密码无缝切换。

**表1-48 配置路由域验证**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置路由域验证方式和验证密码 | **domain-authentication-mode** { **md5** \| **simple** \| **gca** *key-id* { **hmac-sha-1** \| **hmac-sha-224** \| **hmac-sha-256** \| **hmac-sha-384** \| **hmac-sha-512** } } { **cipher** *cipher-string* \| **plain** *plain-string* } [ **ip** \| **osi** ] | 缺省情况下，系统没有配置路由域验证方式和验证密码 |
| （可选）配置对收到的Level-2报文（包括LSP、CSNP、PSNP）忽略认证信息检查 | **domain-authentication send-only** | 缺省情况下，如果配置了路由域验证方式和验证密码，对收到的报文执行认证信息检查 |

## 1.7 配置IS-IS GR


提示

IS-IS GR 特性与 IS-IS NSR 特性互斥，不能同时配置。

GR（Graceful Restart，平滑重启）是一种通过备份 IS-IS 配置信息，在协议重启或主备倒换时 IS-IS 进行平滑重启，保持邻接关系，并对 LSDB 进行同步，从而保证转发业务不中断的机制。

GR 有两个角色：

- GR Restarter：发生协议重启或主备倒换事件且具有 GR 能力的设备。

- GR Helper：和 GR Restarter 具有邻居关系，协助完成 GR 流程的设备。

只需要在作为 GR Restarter 的设备上进行以下配置，设备缺省都是 GR Helper。

**表1-49 配置 IS-IS GR**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 使能IS-IS路由进程，进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能IS-IS协议的GR能力 | **graceful-restart** | 缺省情况下，IS-IS协议的GR能力处于关闭状态 |
| （可选）配置重启时抑制SA位 | **graceful-restart suppress-sa** | 缺省情况下，重启时不抑制SA位<br>配置重启时抑制SA（Suppress-Advertisement）位，即在重启路由器的Hello PDU中设置抑制发布SA位，重启路由器的邻居将继续发布该邻接关系 |
| （可选）配置T1定时器 | **graceful-restart t1** *seconds* **count** *count* | 缺省情况下，T1定时器的超时值为3秒，超时次数为10次<br>T1定时器用来控制发送带有RR标志位的Restart TLV的次数。重启路由器发送带有RR标志位的Restart TLV，如果在超时时间内收到对端回复的带有RA标志的Restart TLV，才能正常进入GR流程；否则GR流程失败 |
| （可选）配置T2定时器 | **graceful-restart t2** *seconds* | 缺省情况下，T2定时器的超时值为60秒<br>T2定时器用来控制LSDB同步时间。每个LSDB都有一个T2定时器，对于Level-1-2路由器来说，就需要有两个T2定时器，一个为Level-1的T2定时器，另外一个为Level-2的T2定时器。如果Level-1和Level-2的T2定时器都超时，LSDB同步还没有完成，则GR失败 |
| （可选）配置T3定时器 | **graceful-restart t3** *seconds* | 缺省情况下，T3定时器的超时值为300秒<br>T3定时器用来控制路由器的重启时间间隔。重启时间间隔在IS-IS的Hello PDU中设置为保持时间，这样在该路由器重启的时间内邻居不会断掉与其的邻接关系。如果T3定时器超时后GR还没有完成，则GR失败 |

## 1.8 配置IS-IS NSR

MSR 系列路由器各款型对于本节所描述的特性的支持情况有所不同，详细差异信息如下：

| 型号 | 特性 | 描述 |
| --- | --- | --- |
| MSR 2600 | 配置IS-IS NSR | 不支持 |
| MSR 3600 | | 不支持 |
| MSR 5600 | | 支持 |

提示

IS-IS NSR 特性与 IS-IS GR 特性互斥，不能同时配置。

GR 特性存在一些缺陷，如主备倒换期间需要周边设备配合才能完成路由信息的恢复，在网络应用中有一定的限制；而且在主备倒换后 IS-IS 进程重新学习所有的路由，如果在主备倒换期间拓扑发生变化，删除的路由不能及时更新，容易造成黑洞路由。

NSR 就是为了解决 GR 特性的一些缺陷和使用场景限制而实现的一种新特性。NSR 将 IS-IS 链路状态信息从主进程备份到备进程，在发生主备倒换时不需要周边设备配合就可以完成链路状态的恢复和路由的重新生成。

表1-50 配置 IS-IS NSR

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能IS-IS协议的NSR功能 | **non-stop-routing** | 缺省情况下，IS-IS协议的NSR功能处于关闭状态 |

## 1.9 配置IS-IS与BFD联动

BFD（Bidirectional Forwarding Detection，双向转发检测）能够为 IS-IS 邻居之间的链路提供快速检测功能。当邻居之间的链路出现故障时，加快 IS-IS 协议的收敛速度。关于 BFD 的介绍和基本功能配置，请参见“可靠性配置指导”中的“BFD”。

表1-51 配置 IS-IS 与 BFD 联动

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 指定接口上使能IS-IS | **isis enable** [ *process-id* ] | - |
| 指定接口上使能BFD | **isis bfd enable** | 缺省情况下，运行IS-IS的接口不使用BFD提供的链路检测功能 |

# 1.10 配置IS-IS快速重路由

提示

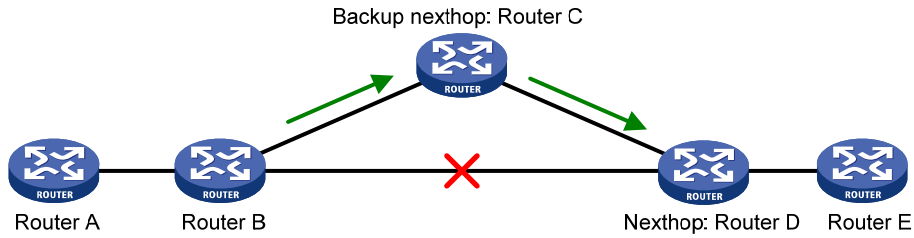
- IS-IS 支持快速重路由功能不能与 IS-IS 的 BFD 功能同时使用，否则可能导致快速重路由功能失效。
- IS-IS 支持快速重路由自动计算备份下一跳功能与 IS-IS TE 特性互斥。

## 1.10.1 功能简介

当 IS-IS 网络中的链路或某台路由器发生故障时，需要通过故障链路或故障路由器传输才能到达目的地的报文将会丢失或产生路由环路，数据流量将会中断，直到 IS-IS 根据新的拓扑网络路由收敛完毕后，被中断的流量才能恢复正常的传输。

为了尽可能缩短网络故障导致的流量中断时间，网络管理员可以配置 IS-IS 快速重路由功能。

图1-8 IS-IS 快速重路由功能示意图




如 图 1-8 所示，通过在Router B上使能快速重路由功能，IS-IS将为路由计算或指定备份下一跳，当Router B检测到网络故障时，IS-IS会使用事先获取的备份下一跳替换失效下一跳，通过备份下一跳来指导报文的转发，从而大大缩短了流量中断时间。在使用备份下一跳指导报文转发的同时，IS-IS会根据变化后的网络拓扑重新计算最短路径，网络收敛完毕后，使用新计算出来的最优路由来指导报文转发。

网络管理员可以配置给所有 IS-IS 路由自动计算备份下一跳，也可以在路由策略中指定备份下一跳，为符合过滤条件的路由指定备份下一跳。

## 1.10.2 配置准备

在配置 IS-IS 快速重路由特性之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点网络层可达
- 使能 IS-IS 功能

## 1.10.3 配置步骤

### 1. 配置IS-IS支持快速重路由功能

(1) 配置 IS-IS 支持快速重路由功能（自动计算备份下一跳）

表1-52 配置 IS-IS 支持快速重路由功能（自动计算备份下一跳）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入接口视图 | | **interface** *interface-type interface-number* | - |
| （可选）去使能接口LFA计算功能 | | **isis fast-reroute lfa-backup exclude** | 缺省情况下，接口参与LFA计算，能够被选为备份接口 |
| 退回系统视图 | | **quit** | - |
| 进入IS-IS IPv4单播地址族视图或IS-IS IPv4单播拓扑视图 | 进入IS-IS IPv4单播地址族视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置IS-IS支持快速重路由功能（自动计算备份下一跳） | | **fast-reroute auto** | 缺省情况下，IS-IS支持快速重路由功能处于关闭状态 |

(2) 配置 IS-IS 支持快速重路由功能（通过路由策略指定备份下一跳）

网络管理员可以通过 **apply fast-reroute backup-interface** 命令在路由策略中指定备份下一跳，为符合过滤条件的路由指定备份下一跳，关于 **apply fast-reroute backup-interface** 命令以及路由策略的相关配置，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

表1-53 配置 IS-IS 支持快速重路由功能（通过路由策略指定备份下一跳）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入接口视图 | | **interface** *interface-type interface-number* | - |
| （可选）去使能接口LFA计算功能 | | **isis fast-reroute lfa-backup exclude** | 缺省情况下，接口参与LFA计算，能够被选为备份接口 |
| 退回系统视图 | | **quit** | - |
| 进入IS-IS IPv4单播地 | 进入IS-IS IPv4单 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 二者选其一 |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 址族视图或IS-IS IPv4单播拓扑视图 | 播地址族视图 | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入IS-IS IPv4单播拓扑视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | |
| | | **cost-style** { **wide** \| **wide-compatible** } | |
| | | **address-family ipv4** [ **unicast** ] | |
| | | **topology** *topo-name* **tid** *tid* | |
| 配置IS-IS支持快速重路由功能（通过路由策略指定备份下一跳） | | **fast-reroute route-policy** *route-policy-name* | 缺省情况下，IS-IS支持快速重路由功能处于关闭状态 |

### 2. 配置IS-IS快速重路由支持BFD检测功能

IS-IS 协议的快速重路由特性中，主用链路缺省不使用 BFD 进行链路故障检测。配置本功能后，将使用 BFD（Echo 方式）进行检测，可以加快 IS-IS 协议的收敛速度。。

表1-54 配置 IS-IS 快速重路由支持 BFD 检测功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置BFD Echo报文源地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置BFD Echo报文源地址 |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能IS-IS协议中主用链路的BFD（Echo方式）检测功能 | **isis primary-path-detect bfd echo** | 缺省情况下，IS-IS协议中主用链路的BFD（Echo方式）检测功能处于关闭状态 |

## 1.11 配置IS-IS支持IPv4单播拓扑

说明

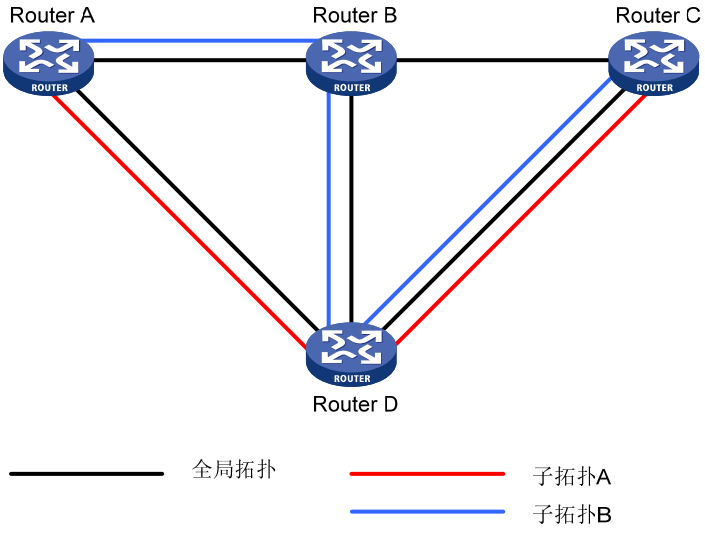
IS-IS 支持 IPv6 单播拓扑的相关内容，请参见“三层技术-IP 路由配置指导”中的“IPv6 IS-IS”。

### 1.11.1 概述

对于传统的路由技术，一个物理拓扑是依靠路由来建立逻辑结构的，同一目的的不同业务报文必然是通过相同的链路来进行转发。虽然可以通过策略路由来改变下一跳，或者是通过 TE 来进行流量的规划，但是 MTR 提供了另一个选择。与策略路由相比，MTR 的优势在于是基于拓扑而不是下一跳。与 TE 相比，MTR 的部署要更方便一些。MTR 的实现是在某一个地址族中（如：IPv4），进行

一个全拓扑计算的同时，按照拓扑号将全拓扑分为多子拓扑，对于不同的流量可以通过不同的子拓扑进行转发。

图1-9 IS-IS 支持 IPv4 单播拓扑功能示意图




如 图 1-9 所示，可以根据需要对全局拓扑进行划分，分为多个子拓扑，这样不同的流量就可以走不同的拓扑。例如，语音流可以走子拓扑A，视频流可以走子拓扑B。

对于子拓扑 A 而言，Router B 并不存在；而对于子拓扑 B 而言，它认为 Router A 和 Router D 之间，以及 Router B 和 Router C 之间并没有可用的链路相连。每一个单独的拓扑都根据路由协议计算出自己的路由，属于本拓扑的流量则根据本拓扑的路由表进行转发。

## 1.11.2 配置IS-IS支持IPv4 单播拓扑

### 1. 配置准备

在配置 IS-IS 支持 IPv4 单播拓扑之前，需完成以下任务：

- 配置 IS-IS 基本功能，使网络建立 IS-IS 邻居，有基本拓扑；
- 配置 MTR。关于 MTR 的详细配置过程，请参见“三层技术-IP 路由配置指导”中的“MTR”。

### 2. 配置IS-IS支持IPv4 单播拓扑

表1-55 配置 IS-IS 支持 IPv4 单播拓扑

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS的Wide Metric属性 | **cost-style** { **wide** \| **wide-compatible** \| **compatible** } | 缺省情况下，IS-IS只收发采用narrow方式的报文 |
| 进入IPv4地址族视图 | **address-family ipv4** [ **unicast** ] | - |
| 配置IS-IS支持IPv4单播拓扑 | **topology** *topo-name* **tid** *tid* | 缺省情况下，IS-IS不支持任何IPv4单播拓扑 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 进入接口IPv4单播拓扑视图 | **topology ipv4** [ **unicast** ] *topo-name* | - |
| 配置接口的指定拓扑中使能IS-IS功能 | **isis topology enable** | 缺省情况下，接口不关联到任何拓扑 |

## 1.12 IS-IS显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 IS-IS 的运行情况，用户可以通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除 IS-IS 进程所有的数据结构信息。

表1-56 IS-IS 显示和维护

| 操作 | 命令 |
| --- | --- |
| 显示IS-IS的进程信息（MSR 2600/MSR 3600） | **display isis** [ *process-id* ] |
| 显示IS-IS备份的进程信息（MSR 5600） | **display isis** [ *process-id* ] [ **standby slot** *slot-number* ] |
| 显示IS-IS GR日志信息（MSR 2600/MSR 3600） | **display isis graceful-restart event-log** |
| 显示IS-IS GR日志信息（MSR 5600） | **display isis graceful-restart event-log slot** *slot-number* |
| 显示IS-IS协议的GR状态 | **display isis graceful-restart status** [ **level-1** \| **level-2** ] [ *process-id* ] |
| 显示IS-IS的接口信息（MSR 2600/MSR 3600） | **display isis interface** [ [ *interface-type interface-number* ] [ **verbose** ] \| **statistics** ] [ *process-id* ] |
| 显示IS-IS备份的接口信息（MSR 5600） | **display isis interface** [ *interface-type interface-number* ] [ **verbose** ] [ *process-id* ] [ **standby slot** *slot-number* ] |
| 显示IS-IS的链路状态数据库信息（MSR 2600/MSR 3600） | **display isis lsdb** [ [ **level-1** \| **level-2** ] \| **local** \| **lsp-id** *lspid* \| [ **lsp-name** *lspname* ] \| **verbose** ] * [ *process-id* ] |
| 显示IS-IS备份的链路状态数据库信息（MSR 5600） | **display isis lsdb** [ [ **level-1** \| **level-2** ] \| **local** \| [ **lsp-id** *lspid* \| **lsp-name** *lspname* ] \| **verbose** ] * [ *process-id* ] [ **standby slot** *slot-number* ] |
| 显示IS-IS Mesh-Group的配置信息 | **display isis mesh-group** [ *process-id* ] |
| 显示系统ID到主机名称的映射关系表 | **display isis name-table** [ *process-id* ] |
| 显示IS-IS NSR日志信息（MSR 5600） | **display isis non-stop-routing event-log slot** *slot-number* |
| 显示IS-IS的NSR状态 | **display isis non-stop-routing status** |

| 操作 | 命令 |
| --- | --- |
| 显示IS-IS的邻居信息（MSR 2600/MSR 3600） | **display isis peer** [ **statistics \| verbose** ] [ *process-id* ] |
| 显示IS-IS备份的邻居信息（MSR 5600） | **display isis peer** [ **statistics** \| **verbose** ] [ *process-id* ] [ **standby slot** *slot-number* ] |
| 显示IS-IS引入路由的信息 | **display isis redistribute** [ **ipv4** [ **topology** *topo-name* ] [ *ip-address mask-lengh* ] ] [ **level-1** \| **level-2** ] [ *process-id* ] |
| 显示IS-IS的IPv4路由信息 | **display isis route** [ **ipv4** [ **topology** *topo-name* ] [ *ip-address mask-length* ] ] [ [ **level-1** \| **level-2** ] \| **verbose** ] * [ *process-id* ] |
| 显示IS-IS的IPv4拓扑信息 | **display isis spf-tree** [ **ipv4** [ **topology** *topo-name* ] ] [ [ **level-1** \| **level-2** ] \| **verbose** ] * [ *process-id* ] |
| 显示IS-IS的统计信息 | **display isis statistics** [ **ipv4** [ **topology** *topo-name* ] ] [ **level-1** \| **level-1-2** \| **level-2** ] [ *process-id* ] |
| 显示OSI连接的信息（MSR 2600/MSR 3600） | **display osi** |
| 显示OSI连接的信息（MSR 5600） | **display osi** [ **slot** *slot-number* ] |
| 显示OSI连接的报文统计信息（MSR 2600/MSR 3600） | **display osi statistics** |
| 显示OSI连接的报文统计信息（MSR 5600） | **display osi statistics** [ **slot** *slot-number* ] |
| 清除IS-IS进程所有的数据结构信息 | **reset isis all** [ *process-id* ] [ **graceful-restart** ] |
| 清除IS-IS GR的日志信息（MSR 2600/MSR 3600） | **reset isis graceful-restart event-log** |
| 清除IS-IS GR的日志信息（MSR 5600） | **reset isis graceful-restart event-log slot** *slot-number* |
| 清除IS-IS NSR的日志信息（MSR 2600/MSR 3600） | **reset isis non-stop-routing event-log slot** *slot-number* |
| 清除IS-IS NSR的日志信息（MSR 5600） | **reset isis non-stop-routing event-log slot** *slot-number* |
| 清除IS-IS指定邻居的数据结构信息 | **reset isis peer** *system-id* [ *process-id* ] |
| 清除OSI连接的报文统计信息 | **reset osi statistics** |

# 1.13 IS-IS典型配置举例

## 1.13.1 IS-IS基本功能配置举例

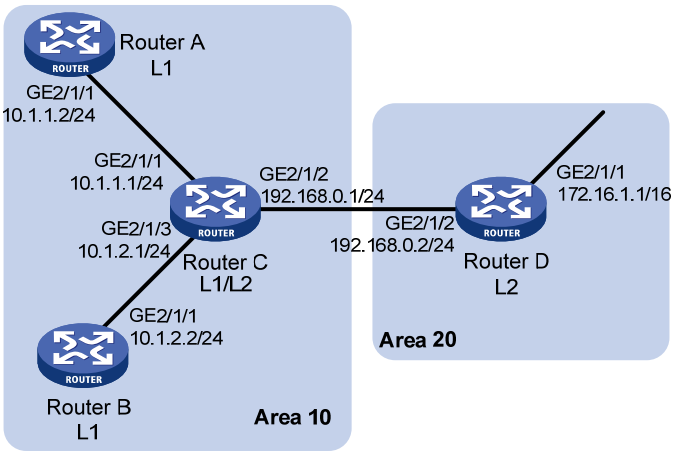
### 1. 组网需求

如 图 1-10 所示，Router A、Router B、Router C和Router D属于同一自治系统，要求它们之间通过IS-IS协议达到IP网络互连的目的。

Router A 和 Router B 为 Level-1 路由器，Router D 为 Level-2 路由器，Router C 作为 Level-1-2 路由器将两个区域相连。Router A、Router B 和 Router C 的区域号为 10，Router D 的区域号为 20。

## 2. 组网图

图1-10 IS-IS 基本功能配置组网图




## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 IS-IS

# 配置 Router A。

```
<RouterA> system-view
[RouterA] isis 1
[RouterA-isis-1] is-level level-1
[RouterA-isis-1] network-entity 10.0000.0000.0001.00
[RouterA-isis-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis enable 1
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] isis 1
[RouterB-isis-1] is-level level-1
[RouterB-isis-1] network-entity 10.0000.0000.0002.00
[RouterB-isis-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis enable 1
[RouterB-GigabitEthernet2/1/1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] isis 1
[RouterC-isis-1] network-entity 10.0000.0000.0003.00
[RouterC-isis-1] quit
[RouterC] interface gigabitethernet 2/1/3
[RouterC-GigabitEthernet2/1/3] isis enable 1
[RouterC-GigabitEthernet2/1/3] quit
[RouterC] interface gigabitethernet 2/1/1
```

```
[RouterC-GigabitEthernet2/1/1] isis enable 1
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] isis enable 1
[RouterC-GigabitEthernet2/1/2] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] isis 1
[RouterD-isis-1] is-level level-2
[RouterD-isis-1] network-entity 20.0000.0000.0004.00
[RouterD-isis-1] quit
[RouterD] interface gigabitethernet 2/1/1
[RouterD-GigabitEthernet2/1/1] isis enable 1
[RouterD-GigabitEthernet2/1/1] quit
[RouterD] interface gigabitethernet 2/1/2
[RouterD-GigabitEthernet2/1/2] isis enable 1
[RouterD-GigabitEthernet2/1/2] quit
```

## 4. 验证配置

# 显示各路由器的 IS-IS LSDB 信息。

```
[RouterA] display isis lsdb

                      Database information for IS-IS(1)
                      ---------------------------------

                         Level-1 Link State Database
                         ---------------------------

LSPID                 Seq Num      Checksum      Holdtime      Length  ATT/P/OL
-------------------------------------------------------------------------------
0000.0000.0001.00-00* 0x00000004   0xdf5e        1096          68      0/0/0
0000.0000.0002.00-00  0x00000004   0xee4d        1102          68      0/0/0
0000.0000.0002.01-00  0x00000001   0xdaaf        1102          55      0/0/0
0000.0000.0003.00-00  0x00000009   0xcaa3        1161          111     1/0/0
0000.0000.0003.01-00  0x00000001   0xadda        1112          55      0/0/0

    *-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload
[RouterB] display isis lsdb

                      Database information for IS-IS(1)
                      ---------------------------------

                         Level-1 Link State Database
                         ---------------------------

LSPID                 Seq Num      Checksum      Holdtime      Length  ATT/P/OL
-------------------------------------------------------------------------------
0000.0000.0001.00-00  0x00000006   0xdb60        988           68      0/0/0
```

```
0000.0000.0002.00-00* 0x00000008   0xe651        1189          68      0/0/0
0000.0000.0002.01-00* 0x00000005   0xd2b3        1188          55      0/0/0
0000.0000.0003.00-00  0x00000014   0x194a        1190          111     1/0/0
0000.0000.0003.01-00  0x00000002   0xabdb        995           55      0/0/0

    *-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload
[RouterC] display isis lsdb

                        Database information for IS-IS(1)
                        ---------------------------------

                          Level-1 Link State Database
                          ---------------------------

LSPID                 Seq Num      Checksum      Holdtime      Length  ATT/P/OL
-------------------------------------------------------------------------------
0000.0000.0001.00-00  0x00000006   0xdb60        847           68      0/0/0
0000.0000.0002.00-00  0x00000008   0xe651        1053          68      0/0/0
0000.0000.0002.01-00  0x00000005   0xd2b3        1052          55      0/0/0
0000.0000.0003.00-00* 0x00000014   0x194a        1051          111     1/0/0
0000.0000.0003.01-00* 0x00000002   0xabdb        854           55      0/0/0

    *-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload

                          Level-2 Link State Database
                          ---------------------------

LSPID                 Seq Num      Checksum      Holdtime      Length  ATT/P/OL
-------------------------------------------------------------------------------
0000.0000.0003.00-00* 0x00000012   0xc93c        842           100     0/0/0
0000.0000.0004.00-00  0x00000026   0x331         1173          84      0/0/0
0000.0000.0004.01-00  0x00000001   0xee95        668           55      0/0/0

    *-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload
[RouterD] display isis lsdb

                        Database information for IS-IS(1)
                        ---------------------------------

                          Level-2 Link State Database
                          ---------------------------

LSPID                 Seq Num      Checksum      Holdtime      Length  ATT/P/OL
-------------------------------------------------------------------------------
0000.0000.0003.00-00  0x00000013   0xc73d        1003          100     0/0/0
0000.0000.0004.00-00* 0x0000003c   0xd647        1194          84      0/0/0
0000.0000.0004.01-00* 0x00000002   0xec96        1007          55      0/0/0
```

```
    *-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload
```

# 显示各路由器的 IS-IS 路由信息。Level-1 路由器的路由表中应该有一条缺省路由，且下一跳为 Level-1-2 路由器，Level-2 路由器的路由表中应该有所有 Level-1 和 Level-2 的路由。

```
[RouterA] display isis route

                        Route information for IS-IS(1)
                        ------------------------------

                        Level-1 IPv4 Forwarding Table
                        ------------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 10.1.1.0/24          10         NULL    GE2/1/1         Direct          D/L/-
 10.1.2.0/24          20         NULL    GE2/1/1         10.1.1.1        R/-/-
 192.168.0.0/24       20         NULL    GE2/1/1         10.1.1.1        R/-/-
 0.0.0.0/0            10         NULL    GE2/1/1         10.1.1.1        R/-/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
[RouterC] display isis route

                        Route information for IS-IS(1)
                        ------------------------------

                        Level-1 IPv4 Forwarding Table
                        ------------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 10.1.1.0/24          10         NULL    GE2/1/1         Direct          D/L/-
 10.1.2.0/24          10         NULL    GE2/1/3         Direct          D/L/-
 192.168.0.0/24       10         NULL    GE2/1/2         Direct          D/L/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set

                        Level-2 IPv4 Forwarding Table
                        ------------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 10.1.1.0/24          10         NULL                                    D/L/-
 10.1.2.0/24          10         NULL                                    D/L/-
 192.168.0.0/24       10         NULL                                    D/L/-
 172.16.0.0/16        20         NULL    GE2/1/2         192.168.0.2     R/-/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
[RouterD] display isis route
```

```
                  Route information for IS-IS(1)
                  ----------------------------

                  Level-2 IPv4 Forwarding Table
                  ----------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 192.168.0.0/24       10         NULL    GE2/1/2         Direct          D/L/-
 10.1.1.0/24          20         NULL    GE2/1/2         192.168.0.1     R/-/-
 10.1.2.0/24          20         NULL    GE2/1/2         192.168.0.1     R/-/-
 172.16.0.0/16        10         NULL    GE2/1/1         Direct          D/L/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
```

## 1.13.2 配置IS-IS的DIS选择

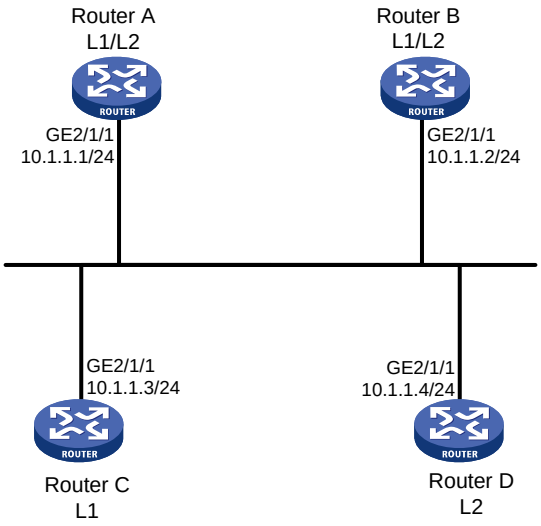
### 1. 组网需求

如 图 1-11 所示，Router A、Router B、Router C和Router D都运行IS-IS路由协议以实现互连，它们属于同一区域 10，网络类型为广播网（以太网）。

Router A 和 Router B 是 Level-1-2 路由器，Router C 为 Level-1 路由器，Router D 为 Level-2 路由器。要求通过改变接口的 DIS 优先级，将 Router A 配置为 Level-1-2 的 DIS 路由器。

### 2. 组网图

图1-11 配置 IS-IS 的 DIS 选择组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 IS-IS

# 配置 Router A。

```
<RouterA> system-view
```

```
[RouterA] isis 1
[RouterA-isis-1] network-entity 10.0000.0000.0001.00
[RouterA-isis-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis enable 1
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] isis 1
[RouterB-isis-1] network-entity 10.0000.0000.0002.00
[RouterB-isis-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis enable 1
[RouterB-GigabitEthernet2/1/1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] isis 1
[RouterC-isis-1] network-entity 10.0000.0000.0003.00
[RouterC-isis-1] is-level level-1
[RouterC-isis-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] isis enable 1
[RouterC-GigabitEthernet2/1/1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] isis 1
[RouterD-isis-1] network-entity 10.0000.0000.0004.00
[RouterD-isis-1] is-level level-2
[RouterD-isis-1] quit
[RouterD] interface gigabitethernet 2/1/1
[RouterD-GigabitEthernet2/1/1] isis enable 1
[RouterD-GigabitEthernet2/1/1] quit
```

# 查看 Router A 的 IS-IS 邻居信息。

```
[RouterA] display isis peer

                        Peer information for IS-IS(1)
                        ---------------------------

  System Id: 0000.0000.0002
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0003.01
  State: Up     HoldTime: 21s        Type: L1(L1L2)     PRI: 64

  System Id: 0000.0000.0003
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0003.01
  State: Up     HoldTime: 6s         Type: L1           PRI: 64

  System Id: 0000.0000.0002
```

```
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0004.01
  State: Up     HoldTime: 23s        Type: L2(L1L2)     PRI: 64

  System Id: 0000.0000.0004
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0004.01
  State: Up     HoldTime: 23s        Type: L2           PRI: 64
```

# 显示 Router A 的 IS-IS 接口信息。

```
[RouterA] display isis interface

                     Interface information for IS-IS(1)
                     --------------------------------

  Interface: GigabitEthernet2/1/1
  Id      IPv4.State      IPv6.State     MTU   Type   DIS
  001         Up              Down        1497  L1/L2  No/No
```

# 显示 Router C 的 IS-IS 接口信息。

```
[RouterC] display isis interface

                     Interface information for IS-IS(1)
                     --------------------------------

  Interface: GigabitEthernet2/1/1
  Id      IPv4.State      IPv6.State     MTU   Type   DIS
  001         Up              Down        1497  L1/L2  Yes/No
```

# 显示 Router D 的 IS-IS 接口信息。

```
[RouterD] display isis interface

                     Interface information for IS-IS(1)
                     --------------------------------

  Interface: GigabitEthernet2/1/1
  Id      IPv4.State      IPv6.State     MTU   Type   DIS
  001         Up              Down        1497  L1/L2  No/Yes
```

从接口信息中可以看到，在使用缺省 DIS 优先级的情况下，Router C 为 Level-1 的 DIS，Router D 为 Level-2 的 DIS。Level-1 和 Level-2 的伪节点分别是 0000.0000.0003.01 和 0000.0000.0004.01。

# 配置 Router A 的 DIS 优先级。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis dis-priority 100
```

# 查看 Router A 的 IS-IS 邻居信息。

```
[RouterA] display isis peer

                        Peer information for IS-IS(1)
                        ---------------------------

  System Id: 0000.0000.0002
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 29s        Type: L1(L1L2)     PRI: 64
```

```
  System Id: 0000.0000.0003
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 22s        Type: L1           PRI: 64

  System Id: 0000.0000.0002
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 22s        Type: L2(L1L2)     PRI: 64

  System Id: 0000.0000.0004
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 22s        Type: L2           PRI: 64
```

# 查看 Router A 的 IS-IS 接口信息。

```
[RouterA] display isis interface

                     Interface information for IS-IS(1)
                     --------------------------------

  Interface: GigabitEthernet2/1/1
  Id      IPv4.State      IPv6.State     MTU   Type   DIS
  001         Up             Down        1497  L1/L2  Yes/Yes
```

从上述信息中可以看到，在改变 IS-IS 接口的 DIS 优先级后，Router A 立即成为 Level-1-2 的 DIS，且伪节点是 0000.0000.0001.01。

# 显示 Router C 的 IS-IS 邻居和接口信息。

```
[RouterC] display isis peer

                        Peer information for IS-IS(1)
                        ----------------------------

  System Id: 0000.0000.0001
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 7s         Type: L1           PRI: 100

  System Id: 0000.0000.0002
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 23s        Type: L1           PRI: 64
[RouterC] display isis interface

                     Interface information for IS-IS(1)
                     --------------------------------

  Interface: GigabitEthernet2/1/1
  Id      IPv4.State      IPv6.State     MTU   Type   DIS
  001         Up             Down        1497  L1/L2  No/No
```

# 显示 Router D 的 IS-IS 邻居和接口信息。

```
[RouterD] display isis peer
```

```
                        Peer information for IS-IS(1)
                        ----------------------------

  System Id: 0000.0000.0001
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 7s         Type: L2           PRI: 100

  System Id: 0000.0000.0002
  Interface: GigabitEthernet2/1/1             Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 26s        Type: L2           PRI: 64
[RouterD] display isis interface

                     Interface information for IS-IS(1)
                     ----------------------------------

  Interface: GigabitEthernet2/1/1
  Id      IPv4.State      IPv6.State     MTU   Type   DIS
  001         Up              Down       1497  L1/L2  No/No
```

## 1.13.3 配置IS-IS引入外部路由

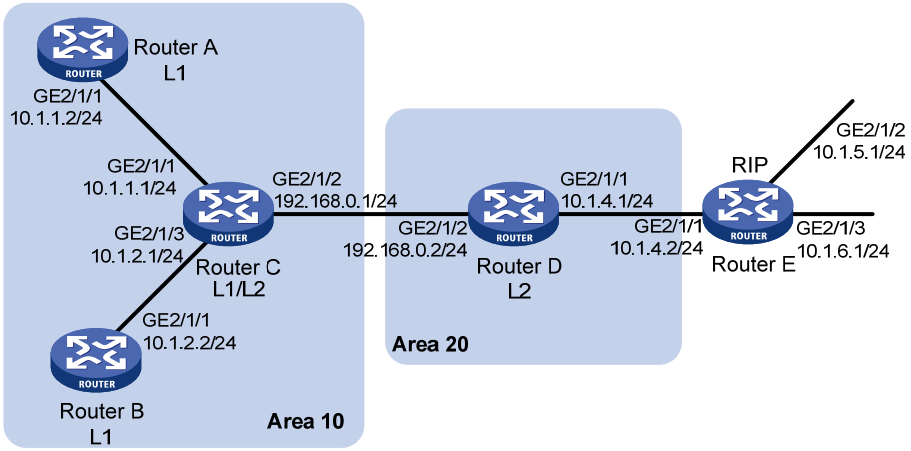
### 1. 组网需求

如 图 1-12 所示：

- Router A、Router B、Router C 和 Router D 属于同一自治系统，要求它们之间通过 IS-IS 协议达到 IP 网络互连的目的。
- Router A 和 Router B 为 Level-1 路由器，Router D 为 Level-2 路由器，Router C 作为 Level-1-2 路由器将两个区域相连。Router A、Router B 和 Router C 的区域号为 10，Router D 的区域号为 20。
- 在 Router D 的 IS-IS 进程中引入 RIP 路由。

### 2. 组网图

图1-12 配置 IS-IS 引入外部路由组网图



### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 IS-IS 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] isis 1
[RouterA-isis-1] is-level level-1
[RouterA-isis-1] network-entity 10.0000.0000.0001.00
[RouterA-isis-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis enable 1
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] isis 1
[RouterB-isis-1] is-level level-1
[RouterB-isis-1] network-entity 10.0000.0000.0002.00
[RouterB-isis-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis enable 1
[RouterB-GigabitEthernet2/1/1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] isis 1
[RouterC-isis-1] network-entity 10.0000.0000.0003.00
[RouterC-isis-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] isis enable 1
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] isis enable 1
[RouterC-GigabitEthernet2/1/2] quit
[RouterC] interface gigabitethernet 2/1/3
[RouterC-GigabitEthernet2/1/3] isis enable 1
[RouterC-GigabitEthernet2/1/3] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] isis 1
[RouterD-isis-1] is-level level-2
[RouterD-isis-1] network-entity 20.0000.0000.0004.00
[RouterD-isis-1] quit
[RouterD] interface gigabitethernet 2/1/1
[RouterD-GigabitEthernet2/1/1] isis enable 1
[RouterD-GigabitEthernet2/1/1] quit
```

# 显示各路由器的 IS-IS 路由信息。

```
[RouterA] display isis route
```

```
                        Route information for IS-IS(1)
                        ------------------------------

                        Level-1 IPv4 Forwarding Table
                        ------------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 10.1.1.0/24          10         NULL    GE2/1/1         Direct          D/L/-
 10.1.2.0/24          20         NULL    GE2/1/1         10.1.1.1        R/-/-
 192.168.0.0/24       20         NULL    GE2/1/1         10.1.1.1        R/-/-
 0.0.0.0/0            10         NULL    GE2/1/1         10.1.1.1        R/-/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
[RouterC] display isis route

                        Route information for IS-IS(1)
                        ------------------------------

                        Level-1 IPv4 Forwarding Table
                        ------------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 10.1.1.0/24          10         NULL    GE2/1/1         Direct          D/L/-
 10.1.2.0/24          10         NULL    GE2/1/3         Direct          D/L/-
 192.168.0.0/24       10         NULL    GE2/1/2         Direct          D/L/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set

                        Level-2 IPv4 Forwarding Table
                        ------------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 10.1.1.0/24          10         NULL                                    D/L/-
 10.1.2.0/24          10         NULL                                    D/L/-
 192.168.0.0/24       10         NULL                                    D/L/-
 172.16.0.0/16        20         NULL    GE2/1/2         192.168.0.2     R/-/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
[RouterD] display isis route

                        Route information for IS-IS(1)
                        ------------------------------

                        Level-2 IPv4 Forwarding Table
```

```
                        ----------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
------------------------------------------------------------------------------
 192.168.0.0/24       10         NULL    GE2/1/2         Direct          D/L/-
 10.1.1.0/24          20         NULL    GE2/1/2         192.168.0.1     R/-/-
 10.1.2.0/24          20         NULL    GE2/1/2         192.168.0.1     R/-/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
```

(3) 在 Router D 和 Router E 之间运行 RIPv2，在 Router D 上配置 IS-IS 进程引入 RIP 路由。

# 在 Router D 上配置 RIPv2。

```
[RouterD] rip 1
[RouterD-rip-1] network 10.0.0.0
[RouterD-rip-1] version 2
[RouterD-rip-1] undo summary
```

# 在 Router E 上配置 RIPv2。

```
[RouterE] rip 1
[RouterE-rip-1] network 10.0.0.0
[RouterE-rip-1] version 2
[RouterE-rip-1] undo summary
```

# 在 Router D 上配置 IS-IS 进程引入 RIP 进程的路由。

```
[RouterD-rip-1] quit
[RouterD] isis 1
[RouterD–isis-1] address-family ipv4
[RouterD–isis-1-ipv4] import-route rip level-2
```

# 显示 Router C 的 IS-IS 路由信息。

```
[RouterC] display isis route

                        Route information for IS-IS(1)
                        ----------------------------

                        Level-1 IPv4 Forwarding Table
                        ----------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
------------------------------------------------------------------------------
 10.1.1.0/24          10         NULL    GE2/1/1         Direct          D/L/-
 10.1.2.0/24          10         NULL    GE2/1/3         Direct          D/L/-
 192.168.0.0/24       10         NULL    GE2/1/2         Direct          D/L/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set

                        Level-2 IPv4 Forwarding Table
                        ----------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
------------------------------------------------------------------------------
```

```
10.1.1.0/24          10         NULL                                    D/L/-
10.1.2.0/24          10         NULL                                    D/L/-
192.168.0.0/24       10         NULL                                   D/L/-
10.1.4.0/24          10         NULL    GE2/1/2         192.168.0.2     R/L/-
10.1.5.0/24          20         NULL    GE2/1/2         192.168.0.2     R/L/-
10.1.6.0/24          20         NULL    GE2/1/2         192.168.0.2     R/L/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
```

## 1.13.4 IS-IS验证配置举例

### 1. 组网需求

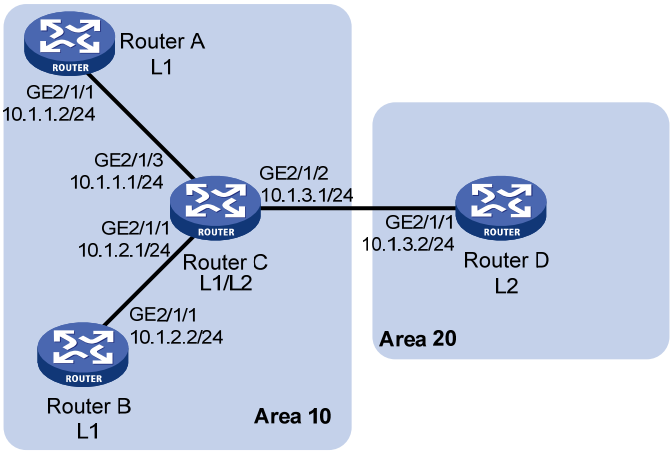
如 图 1-13 所示，Router A、Router B、Router C和Router D属于同一路由域，要求它们之间通过IS-IS协议达到IP网络互连的目的。

其中，Router A、Router B 和 Router C 属于同一个区域，区域号为 10，Router D 属于另外一个区域，区域号为 20。

在区域 10 内配置区域验证，防止不可信任的路由信息加入到区域 10 的 LSDB 中；在 Router C 和 Router D 上配置路由域验证，防止将不可信的路由信息注入当前路由域；分别在 Router A、Router B、Router C 和 Router D 上配置邻居关系验证。

### 2. 组网图

图1-13 IS-IS 验证配置举例图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 IS-IS 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] isis 1
[RouterA-isis-1] network-entity 10.0000.0000.0001.00
[RouterA-isis-1] is-level level-1
[RouterA-isis-1] quit
[RouterA] interface gigabitethernet 2/1/1
```

```
[RouterA-GigabitEthernet2/1/1] isis enable 1
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] isis 1
[RouterB-isis-1] network-entity 10.0000.0000.0002.00
[RouterB-isis-1] is-level level-1
[RouterB-isis-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis enable 1
[RouterB-GigabitEthernet2/1/1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] isis 1
[RouterC-isis-1] network-entity 10.0000.0000.0003.00
[RouterC-isis-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] isis enable 1
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] isis enable 1
[RouterC-GigabitEthernet2/1/2] quit
[RouterC] interface gigabitethernet 2/1/3
[RouterC-GigabitEthernet2/1/3] isis enable 1
[RouterC-GigabitEthernet2/1/3] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] isis 1
[RouterD-isis-1] network-entity 20.0000.0000.0001.00
[RouterD-isis-1] quit
[RouterD] interface gigabitethernet 2/1/1
[RouterD-GigabitEthernet2/1/1] isis enable 1
[RouterD-GigabitEthernet2/1/1] quit
```

(3) 在 Router A、Router B、Router C 和 Router D 之间建立邻居关系验证

# 分别在 Router A 的 GigabitEthernet2/1/1、Router C 的 GigabitEthernet2/1/3 配置邻居关系验证，验证方式为 MD5 明文，验证密码为“eRg”。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis authentication-mode md5 plain eRg
[RouterA-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/3
[RouterC-GigabitEthernet2/1/3] isis authentication-mode md5 plain eRg
[RouterC-GigabitEthernet2/1/3] quit
```

# 分别在 Router B 的 GigabitEthernet2/1/1、Router C 的 GigabitEthernet2/1/1 配置邻居关系验证，验证方式为 MD5 明文，验证密码为“t5Hr”。

```
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis authentication-mode md5 plain t5Hr
[RouterB-GigabitEthernet2/1/1] quit
```

```
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] isis authentication-mode md5 plain t5Hr
[RouterC-GigabitEthernet2/1/1] quit
```

# 分别在 Router C 的 GigabitEthernet2/1/2、Router D 的 GigabitEthernet2/1/1 配置邻居关系验证，验证方式为 MD5 明文，验证密码为“hSec”。

```
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] isis authentication-mode md5 plain hSec
[RouterC-GigabitEthernet2/1/2] quit
[RouterD] interface gigabitethernet 2/1/1
[RouterD-GigabitEthernet2/1/1] isis authentication-mode md5 plain hSec
[RouterD-GigabitEthernet2/1/1] quit
```

(4) 在 Router A、Router B 和 Router C 上配置区域验证，验证方式为 MD5 明文验证，验证密码为“10Sec”。

```
[RouterA] isis 1
[RouterA-isis-1] area-authentication-mode md5 plain 10Sec
[RouterA-isis-1] quit
[RouterB] isis 1
[RouterB-isis-1] area-authentication-mode md5 plain 10Sec
[RouterB-isis-1] quit
[RouterC] isis 1
[RouterC-isis-1] area-authentication-mode md5 plain 10Sec
[RouterC-isis-1] quit
```

(5) 在 Router C 和 Router D 上配置路由域验证，验证方式为 MD5 明文验证，验证密码为“1020Sec”。

```
[RouterC] isis 1
[RouterC-isis-1] domain-authentication-mode md5 plain 1020Sec
[RouterC-isis-1] quit
[RouterD] isis 1
[RouterD-isis-1] domain-authentication-mode md5 plain 1020Sec
```

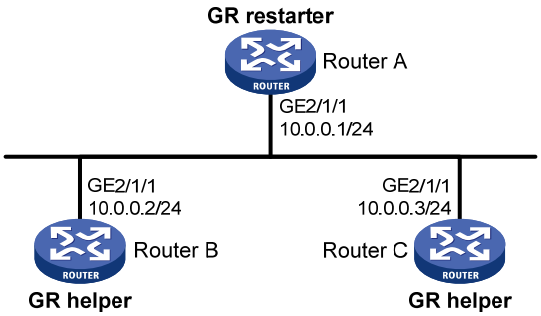
## 1.13.5 IS-IS GR配置举例

### 1. 组网需求

如 图 1-14 所示，Router A、Router B和Router C属于同一域。这三台路由器都运行IS-IS协议以实现路由互连。

## 2. 组网图

图1-14 IS-IS GR 配置组网图




## 3. 配置步骤

(1) 配置各路由器接口的 IP 地址和 IS-IS 协议

请按照 图 1-14 配置各接口的IP地址和子网掩码，具体配置过程略。

配置各路由器之间采用 IS-IS 协议进行互连，确保 Router A、Router B 和 Router C 之间能够在网络层互通，并且各路由器之间能够借助 IS-IS 协议实现动态路由更新，具体配置过程略。

(2) 配置 IS-IS GR

# 使能 Router A 的 IS-IS 协议的 GR 能力。

```
<RouterA> system-view
[RouterA] isis 1
[RouterA-isis-1] graceful-restart
[RouterA-isis-1] return
```

## 4. 验证配置

Router A 分别与 Router B 和 Router C 建立邻接关系后，三台路由器开始交换路由信息。Router A 的 IS-IS 协议重启，进入重启模式后，通过 GR 机制向邻居重新发送连接请求，同步数据库。使用 **display isis graceful-restart status** 命令，可查看 Router A 上 IS-IS 协议的 GR 状态。

# 重启 Router A 的 IS-IS 进程。

```
<RouterA> reset isis all 1 graceful-restart
Reset IS-IS process? [Y/N]:y
```

# 查看 Router A 上 IS-IS 协议的 GR 状态。

```
<RouterA> display isis graceful-restart status

                      Restart information for IS-IS(1)
                      --------------------------------
Restart status: COMPLETE
Restart phase: Finish
Restart t1: 3, count 10; Restart t2: 60; Restart t3: 300
SA Bit: supported

                        Level-1 restart information
                        ---------------------------
Total number of interfaces: 1
```

```
Number of waiting LSPs: 0

                          Level-2 restart information
                          ---------------------------
Total number of interfaces: 1
Number of waiting LSPs: 0
```

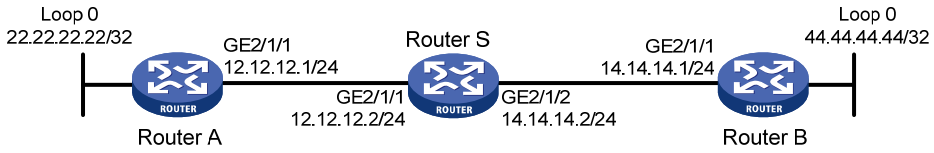
## 1.13.6 IS-IS NSR配置举例

### 1. 组网需求

如 图 1-15 所示，Router S、Router A、Router B属于同一IS-IS区域，通过IS-IS协议实现网络互连。要求对Router S进行主备倒换时，Router A和Router B到Route S的邻居没有中断，Router A到Router B的流量没有中断。

### 2. 组网图

图1-15 IS-IS NSR 配置组网图




### 3. 配置步骤

(1) 配置各路由器接口的 IP 地址和 IS-IS 协议

请按照 图 1-15 配置各接口的IP地址和子网掩码，具体配置过程略。

配置各路由器之间采用 IS-IS 协议进行互连，确保 Router S、Router A 和 Router B 之间能够在网络层互通，并且各路由器之间能够借助 IS-IS 协议实现动态路由更新。具体配置过程略。

(2) 配置 IS-IS NSR

# 使能 Router S 的 IS-IS NSR 功能。

```
<RouterS> system-view
[RouterS] isis 1
[RouterS-isis-1] non-stop-routing
[RouterS-isis-1] return
```

### 4. 验证配置

Router S 分别与 Router A 和 Router B 建立邻接关系后，三台路由器开始交换路由信息。当网络稳定后，Router S 进行主备倒换。在 Route S 主备倒换期间，使用 **display isis peer** 命令查看 Router A 和 Router B 上到 Router S 的邻居是否发生任何变化；使用 **display isis route** 命令查看 Router A 上是否有 Route B 上 Loopback 接口的路由，Router B 上是否有 Route A 上 Loopback 接口的路由。

# Router S 主备倒换。

```
<RouterS> system-view
[RouterS] placement reoptimize
Predicted changes to the placement
Program                          Current location       New location
-------------------------------------------------------------------------
```

```
syslog                              0/0                   0/0
diagusageratio                      0/0                   0/0
l3vpn                               0/0                   0/0
fc                                  0/0                   0/0
dns                                 0/0                   0/0
lauth                               0/0                   0/0
aaa                                 0/0                   0/0
lsm                                 0/0                   0/0
rm                                  0/0                   0/0
rm6                                 0/0                   0/0
track                               0/0                   0/0
ip6addr                             0/0                   0/0
ipaddr                              0/0                   0/0
rpm                                 0/0                   0/0
trange                              0/0                   0/0
tunnel                              0/0                   0/0
lagg                                0/0                   0/0
bfd                                 0/0                   0/0
acl                                 0/0                   0/0
slsp                                0/0                   0/0
usr6                                0/0                   0/0
usr                                 0/0                   0/0
qos                                 0/0                   0/0
fczone                              0/0                   0/0
ethbase                             0/0                   0/0
ipcim                               0/0                   0/0
ip6base                             0/0                   0/0
ipbase                              0/0                   0/0
eth                                 0/0                   0/0
eviisis                             0/0                   0/0
ifnet                               NA                    NA
isis                                0/0                   1/0
Continue? [y/n]:y
Re-optimization of the placement start. You will be notified on completion
Re-optimization of the placement complete. Use 'display placement' to view the n
ew placement
```

# 查看 Router A 上 IS-IS 协议的邻居和路由。

```
<RouterA> display isis peer

                        Peer information for ISIS(1)
                        ----------------------------

 System Id: 0000.0000.0001
 Interface: GE2/1/1                 Circuit Id: 0000.0000.0001.01
 State: Up     HoldTime:  23s       Type: L1(L1L2)     PRI: 64

 System Id: 0000.0000.0001
 Interface: GE2/1/1                 Circuit Id: 0000.0000.0001.01
```

```
 State: Up     HoldTime:  28s       Type: L2(L1L2)     PRI: 64
<RouterA> display isis route

                        Route information for ISIS(1)
                        ----------------------------

                        Level-1 IPv4 Forwarding Table
                        ----------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
-------------------------------------------------------------------------------
 12.12.12.0/24        10         NULL    GE2/1/1         Direct          D/L/-
 22.22.22.22/32       10         NULL    Loop0           Direct          D/-/-
 14.14.14.0/32        10         NULL    GE2/1/1         12.12.12.2      R/L/-
 44.44.44.44/32       10         NULL    GE2/1/1         12.12.12.2      R/L/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set

                        Level-2 IPv4 Forwarding Table
                        ----------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
-------------------------------------------------------------------------------
 12.12.12.0/24        10         NULL    GE2/1/1         Direct          D/L/-
 22.22.22.22/32       10         NULL    Loop0           Direct          D/-/-
 14.14.14.0/32        10         NULL
 44.44.44.44/32       10         NULL

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
```

# 查看 Router B 上 IS-IS 协议的邻居和路由。

```
<RouterB> display isis peer

                        Peer information for ISIS(1)
                        ----------------------------

 System Id: 0000.0000.0001
 Interface: GE2/1/1                 Circuit Id: 0000.0000.0001.01
 State: Up     HoldTime:  23s       Type: L1(L1L2)     PRI: 64

 System Id: 0000.0000.0001
 Interface: GE2/1/1                 Circuit Id: 0000.0000.0001.01
 State: Up     HoldTime:  28s       Type: L2(L1L2)     PRI: 64
<RouterB> display isis route

                        Route information for ISIS(1)
                        ----------------------------

                        Level-1 IPv4 Forwarding Table
```

```
                          ----------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
------------------------------------------------------------------------------
 14.14.14.0/24        10         NULL    GE2/1/1         Direct          D/L/-
 44.44.44.44/32       10         NULL    Loop0           Direct          D/-/-
 12.12.12.0/32        10         NULL    GE2/1/1         14.14.14.4      R/L/-
 22.22.22.22/32       10         NULL    GE2/1/1         14.14.14.4      R/L/-

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set

                       Level-2 IPv4 Forwarding Table
                       ----------------------------

 IPv4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
------------------------------------------------------------------------------
 14.14.14.0/24        10         NULL    GE2/1/1         Direct          D/L/-
 44.44.44.44/32       10         NULL    Loop0           Direct          D/-/-
 12.12.12.0/32        10         NULL
 22.22.22.22/32       10         NULL

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
```

通过上面信息可以看出 Router A 和 Router B 的邻居和路由信息保持不变，即 NSR 特性使周边设备无法感知 Router S 的主备倒换。

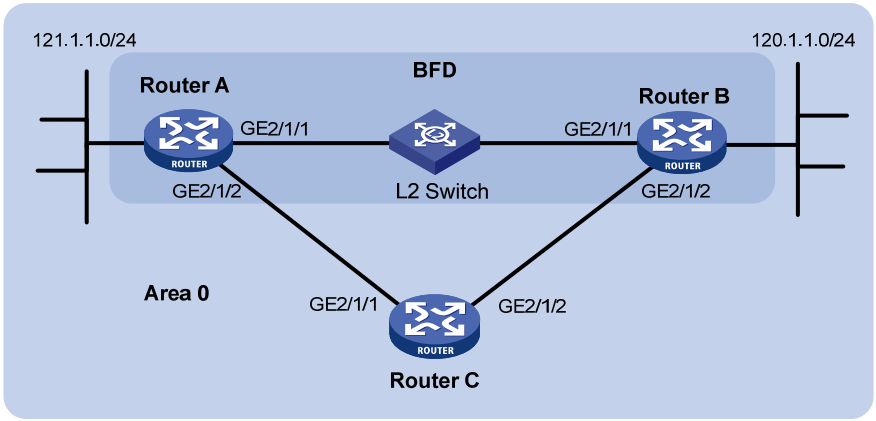
## 1.13.7 配置IS-IS与BFD联动

### 1. 组网需求

- Router A、Router B 和 Ruter C 上运行 IS-IS，网络层相互可达。
- 当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时 BFD 能够快速感知通告 IS-IS 协议，并且切换到 Router C 进行通信。

### 2. 组网图

图1-16 配置 IS-IS 与 BFD 联动组网图




| 设备 | 接口 | IP地址 | 设备 | 接口 | IP地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 192.168.0.102/24 | Router B | GE2/1/1 | 192.168.0.100/24 |

| | | | | | |
|---|---|---|---|---|---|
| | GE2/1/2 | 10.1.1.102/24 | | GE2/1/2 | 13.1.1.1/24 |
| Router C | GE2/1/1 | 10.1.1.100/24 | | | |
| | GE2/1/2 | 13.1.1.2/24 | | | |

### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(1) 配置 IS-IS 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] isis
[RouterA-isis-1] network-entity 10.0000.0000.0001.00
[RouterA-isis-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis enable
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] isis enable
[RouterA-GigabitEthernet2/1/2] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] isis
[RouterB-isis-1] network-entity 10.0000.0000.0002.00
[RouterB-isis-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis enable
[RouterB-GigabitEthernet2/1/1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] isis enable
[RouterB-GigabitEthernet2/1/2] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] isis
[RouterC-isis-1] network-entity 10.0000.0000.0003.00
[RouterC-isis-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] isis enable
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] isis enable
[RouterC-GigabitEthernet2/1/2] quit
```

(2) 配置 BFD 功能

# 在 Router A 上使能 BFD 检测功能，并配置 BFD 参数。

```
[RouterA] bfd session init-mode active
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis bfd enable
[RouterA-GigabitEthernet2/1/1] bfd min-receive-interval 500
```

```
[RouterA-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterA-GigabitEthernet2/1/1] bfd detect-multiplier 7
```

# 在 Router B 上使能 BFD 检测功能，并配置 BFD 参数。

```
[RouterB] bfd session init-mode active
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis bfd enable
[RouterB-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterB-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterB-GigabitEthernet2/1/1] bfd detect-multiplier 8
```

### 4. 验证配置

下面以 Router A 为例，Router B 和 Router A 类似，不再赘述。

# 显示 Router A 的 BFD 信息。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv4 Session Working Under Ctrl Mode:

 LD/RD          SourceAddr      DestAddr        State    Holdtime    Interface
 3/1            192.168.0.102   192.168.0.100   Up       1700ms      GE2/1/1
```

# 在 Router A 上查看 120.1.1.0/24 的路由信息，可以看出 Router A 和 Router B 是通过 L2 Switch 进行通信的。

```
<RouterA> display ip routing-table 120.1.1.0 verbose

Summary Count : 1

Destination: 120.1.1.0/24
   Protocol: ISIS            Process ID: 1
  SubProtID: 0x1                   Age: 04h20m37s
       Cost: 10              Preference: 10
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0                OrigVrf: default-vrf
    TableID: 0x2                 OrigAs: 0
      NBRID: 0x26000002          LastAs: 0
     AttrID: 0xffffffff        Neighbor: 0.0.0.0
      Flags: 0x1008c        OrigNextHop: 192.168.0.100
      Label: NULL           RealNextHop: 192.168.0.100
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid        BkInterface: N/A
```

当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时：

# 在 Router A 上查看 120.1.1.0/24 的路由信息，可以看出 Router A 和 Router B 已经切换到 Router C 进行通信。

```
<RouterA> display ip routing-table 120.1.1.0 verbose

Summary Count : 1
```

```
Destination: 120.1.1.0/24
   Protocol: ISIS            Process ID: 1
  SubProtID: 0x1                   Age: 04h20m37s
       Cost: 20              Preference: 10
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 0.0.0.0
      Flags: 0x1008c        OrigNextHop: 10.1.1.100
      Label: NULL           RealNextHop: 10.1.1.100
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid        BkInterface: N/A
```

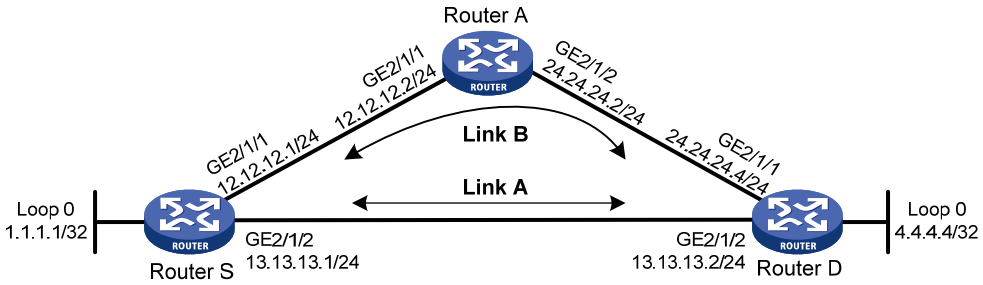
## 1.13.8 IS-IS快速重路由配置举例

### 1. 组网需求

如 图 1-17 所示，Router S、Router A和Router D属于同一IS-IS区域，通过IS-IS协议实现网络互连。要求当Router S和Router D之间的链路出现故障时，业务可以快速切换到链路B上。

### 2. 组网图

图1-17 IS-IS 快速重路由配置组网图




### 3. 配置步骤

(1) 配置各路由器接口的 IP 地址和 IS-IS 协议

请按照 图 1-17 配置各接口的IP地址和子网掩码，具体配置过程略。

配置各路由器之间采用 IS-IS 协议进行互连，确保 Router A、Router D 和 Router S 之间能够在网络层互通，并且各路由器之间能够借助 IS-IS 协议实现动态路由更新。

具体配置过程略。

(2) 配置 IS-IS 快速重路由

IS-IS 支持快速重路由配置有两种配置方法，一种是自动计算，另一种是通过策略指定，两种方法任选一种。

方法一：使能 Router S 和 Router D 的 IS-IS 协议的自动计算快速重路由能力。

# 配置 Router S。

```
<RouterS> system-view
```

```
[RouterS] isis 1
[RouterS-isis-1] address-family ipv4
[RouterS-isis-1-ipv4] fast-reroute auto
[RouterS-isis-1-ipv4] quit
[RouterS-isis-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] isis 1
[RouterD-isis-1] address-family ipv4
[RouterS-isis-1-ipv4] fast-reroute auto
[RouterD-isis-1-ipv4] quit
[RouterD-isis-1] quit
```

方法二：使能 Router S 和 Router D 的 IS-IS 协议的指定路由策略快速重路由能力。

# 配置 Router S。

```
<RouterS> system-view
[RouterS] ip prefix-list abc index 10 permit 4.4.4.4 32
[RouterS] route-policy frr permit node 10
[RouterS-route-policy-frr-10] if-match ip address prefix-list abc
[RouterS-route-policy-frr-10] apply fast-reroute backup-interface gigabitethernet 2/1/1
backup-nexthop 12.12.12.2
[RouterS-route-policy-frr-10] quit
[RouterS] isis 1
[RouterS-isis-1] address-family ipv4
[RouterS-isis-1-ipv4] fast-reroute route-policy frr
[RouterS-isis-1-ipv4] quit
[RouterS-isis-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] ip prefix-list abc index 10 permit 1.1.1.1 32
[RouterD] route-policy frr permit node 10
[RouterD-route-policy-frr-10] if-match ip address prefix-list abc
[RouterD-route-policy-frr-10] apply fast-reroute backup-interface gigabitethernet 2/1/1
backup-nexthop 24.24.24.2
[RouterD-route-policy-frr-10] quit
[RouterD] isis 1
[RouterD-isis-1] address-family ipv4
[RouterS-isis-1-ipv4] fast-reroute route-policy frr
[RouterD-isis-1-ipv4] quit
[RouterD-isis-1] quit
```

### 4. 验证配置

# 在 Router S 上查看 4.4.4.4/32 路由，可以看到备份下一跳信息。

```
[RouterS] display ip routing-table 4.4.4.4 verbose

Summary Count : 1

Destination: 4.4.4.4/32
   Protocol: ISIS            Process ID: 1
```

```
  SubProtID: 0x1                  Age: 04h20m37s
       Cost: 10             Preference: 10
        Tag: 0                   State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 0.0.0.0
      Flags: 0x1008c       OrigNextHop: 13.13.13.2
      Label: NULL          RealNextHop: 13.13.13.2
    BkLabel: NULL            BkNextHop: 12.12.12.2
  Tunnel ID: Invalid         Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid       BkInterface: GigabitEthernet2/1/1
```

# 在 Router D 上查看 1.1.1.1/32 路由，可以看到备份下一跳信息。

```
[RouterD] display ip routing-table 1.1.1.1 verbose

Summary Count : 1

Destination: 1.1.1.1/32
   Protocol: ISIS            Process ID: 1
  SubProtID: 0x1                  Age: 04h20m37s
       Cost: 10             Preference: 10
        Tag: 0                   State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x26000002         LastAs: 0
     AttrID: 0xffffffff       Neighbor: 0.0.0.0
      Flags: 0x1008c       OrigNextHop: 13.13.13.1
      Label: NULL          RealNextHop: 13.13.13.1
    BkLabel: NULL            BkNextHop: 24.24.24.2
  Tunnel ID: Invalid         Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid       BkInterface: GigabitEthernet2/1/1
```

# 目 录

# 1 BGP

## 1.1 BGP简介

BGP（Border Gateway Protocol，边界网关协议）是一种既可以用于不同 AS（Autonomous System，自治系统）之间，又可以用于同一 AS 内部的动态路由协议。当 BGP 运行于同一 AS 内部时，被称为 IBGP（Internal BGP）；当 BGP 运行于不同 AS 之间时，称为 EBGP（External BGP）。AS 是拥有同一选路策略，属于同一技术管理部门的一组路由器。

当前使用的 BGP 版本是 BGP-4。BGP-4 作为 Internet 外部路由协议标准，被 ISP（Internet Service Provider，互联网服务提供商）广泛应用。

BGP 具有如下特点：

- BGP 是一种 EGP（Exterior Gateway Protocol，外部网关协议），与 OSPF、RIP 等 IGP（Interior Gateway Protocol，内部网关协议）不同，其着眼点不在于发现和计算路由，而在于控制路由的传播和选择最佳路由。
- BGP 使用 TCP 作为其传输层协议（端口号 179），提高了协议的可靠性。
- BGP 是一种路径矢量（Path-Vector）路由协议，它采用到达目的地址所经过的 AS 列表来衡量到达目的地址的距离。
- BGP 支持 CIDR（Classless Inter-Domain Routing，无类域间路由）。
- 路由更新时，BGP 只发送更新的路由，大大减少了 BGP 传播路由所占用的带宽，适用于在 Internet 上传播大量的路由信息。
- BGP 路由通过携带 AS 路径信息彻底解决路由环路问题。
- BGP 提供了丰富的路由策略，能够对路由实现灵活的过滤和选择。
- BGP 易于扩展，能够适应网络新的发展。

### 1.1.1 BGP发言者和BGP对等体

运行 BGP 协议的路由器称为 BGP 发言者。BGP 发言者接收或产生路由信息，并将路由信息发布给其它 BGP 发言者。

相互之间存在 TCP 连接、相互交换路由信息的 BGP 发言者互为 BGP 对等体。根据对等体所在的 AS，对等体分为以下几种：

- IBGP 对等体：对等体与本地路由器位于同一 AS。
- EBGP 对等体：对等体与本地路由器位于不同 AS。

### 1.1.2 BGP的消息类型

BGP 定义了以下几种消息类型：

- Open：TCP 连接建立后发送的第一个消息，用于在 BGP 对等体之间建立会话。
- Update：用于在对等体之间交换路由信息。一条 Update 消息可以发布具有相同路径属性的多条可达路由，也可以同时撤销多条不可达路由。
- Keepalive：BGP 周期性地向对等体发送 Keepalive 消息，以保持会话的有效性。

- Route-refresh：用来要求对等体重新发送指定地址族的路由信息。
- Notification：当 BGP 检测到错误状态时，就向对等体发出 Notification 消息，之后 BGP 会话会立即中断。

### 1.1.3 BGP的路由属性

BGP 路由属性是跟随路由一起发布出去的一组参数。它对特定的路由进行了进一步的描述，使得路由接收者能够根据路由属性值对路由进行过滤和选择。下面将介绍几种常见的路由属性。

#### 1. 源（ORIGIN）属性

ORIGIN 属性定义了路由信息的来源，标记一条 BGP 路由是怎么生成的。它有以下三种类型：

- IGP：优先级最高，表示路由产生于本 AS 内。
- EGP：优先级次之，表示路由通过 EGP 学到。
- Incomplete：优先级最低，表示路由的来源无法确定。例如，从其它路由协议引入的路由信息。

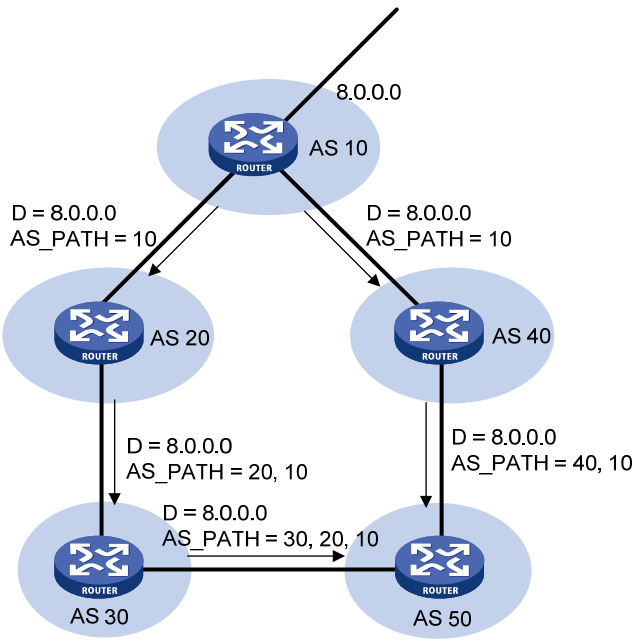
#### 2. AS路径（AS_PATH）属性

AS_PATH 属性记录了某条路由从本地到目的地址所要经过的所有 AS 号。当 BGP 路由器将一条路由通告到其他 AS 时，会把本地 AS 号添加在 AS_PATH 列表中。收到此路由的 BGP 路由器根据 AS_PATH 属性就可以知道到达目的地址所要经过的 AS。

AS_PATH 属性有以下两种类型：

- AS_SEQUENCE：AS号按照一定的顺序排列。如 图 1-1 所示，离本地AS最近的相邻AS号排在前面，其他AS号按顺序依次排列。
- AS_SET：AS 号只是经过的 AS 的简单罗列，没有顺序要求。

图1-1 AS_PATH 属性




AS_PATH 属性具有如下用途：

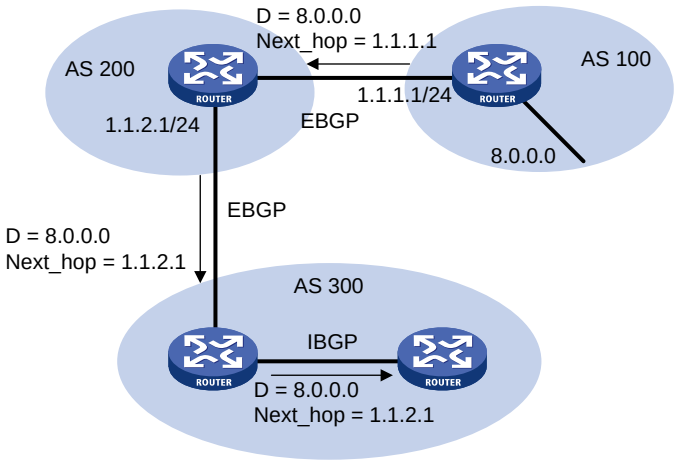
- 避免路由环路的形成：缺省情况下，如果 BGP 路由器接收到的路由的 AS_PATH 属性中已经包含了本地的 AS 号，则 BGP 路由器认为出现路由环路，不会接受该路由。
- 影响路由的选择：在其他因素相同的情况下，BGP会优先选择路径较短的路由。比如在 图 1-1 中，AS 50 中的BGP路由器会选择经过AS 40 的路径作为到目的地址 8.0.0.0 的最优路由。用户可以使用路由策略来人为地增加AS路径的长度，以便更为灵活地控制BGP路径的选择。路由策略的详细介绍，请参见“三层技术-IP路由配置指导”中的“路由策略”。
- 对路由进行过滤：通过配置 AS 路径过滤列表，可以针对 AS_PATH 属性中所包含的 AS 号来对路由进行过滤。AS 路径过滤列表的详细介绍，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

### 3. 下一跳（NEXT_HOP）属性

BGP的NEXT_HOP属性取值不一定是邻居路由器的IP地址。如 图 1-2 所示，NEXT_HOP属性取值情况分为几种：

- BGP 发言者把自己产生的路由发给所有邻居时，将该路由信息的 NEXT_HOP 属性设置为自己与对端连接的接口地址；
- BGP 发言者把接收到的路由发送给 EBGP 对等体时，将该路由信息的 NEXT_HOP 属性设置为自己与对端连接的接口地址；
- BGP发言者把从EBGP邻居得到的路由发给IBGP邻居时，并不改变该路由信息的NEXT_HOP属性。如果配置了负载分担，等价路由被发给IBGP邻居时则会修改NEXT_HOP属性。关于“负载分担”的概念请参见“1.1.6 BGP负载分担”。

图1-2 NEXT_HOP 属性

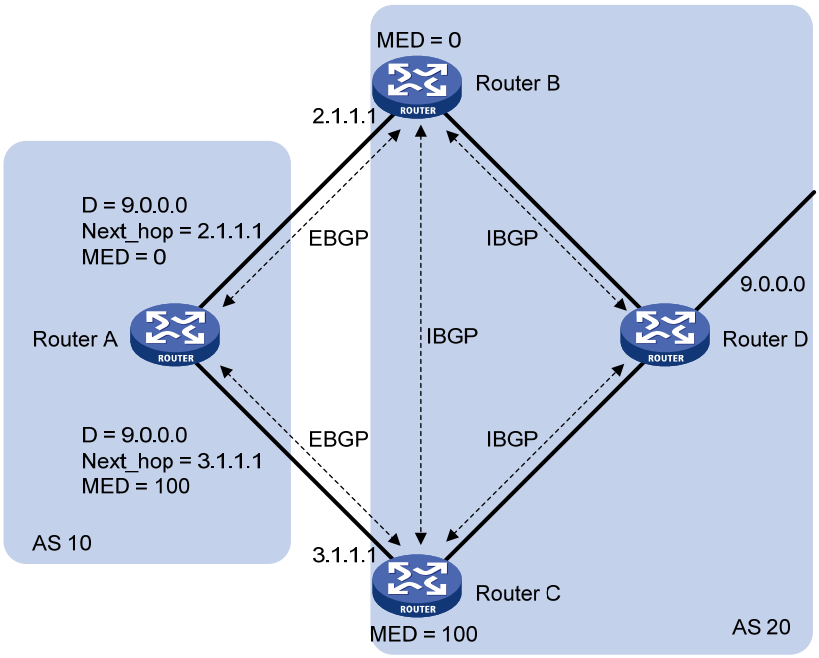



### 4. MED（Multi-Exit Discriminator，多出口区分）属性

MED 属性仅在相邻两个 AS 之间交换，收到此属性的 AS 不会再将其通告给其它 AS。

MED属性相当于IGP使用的度量值（metrics），它用于判断流量进入AS时的最佳路由。当一个BGP路由器通过不同的EBGP对等体得到目的地址相同但下一跳不同的多条路由时，在其它条件相同的情况下，将优先选择MED值较小者作为最佳路由。如 图 1-3 所示，从AS 10 到AS 20 的流量将选择Router B作为入口。

图1-3 MED 属性



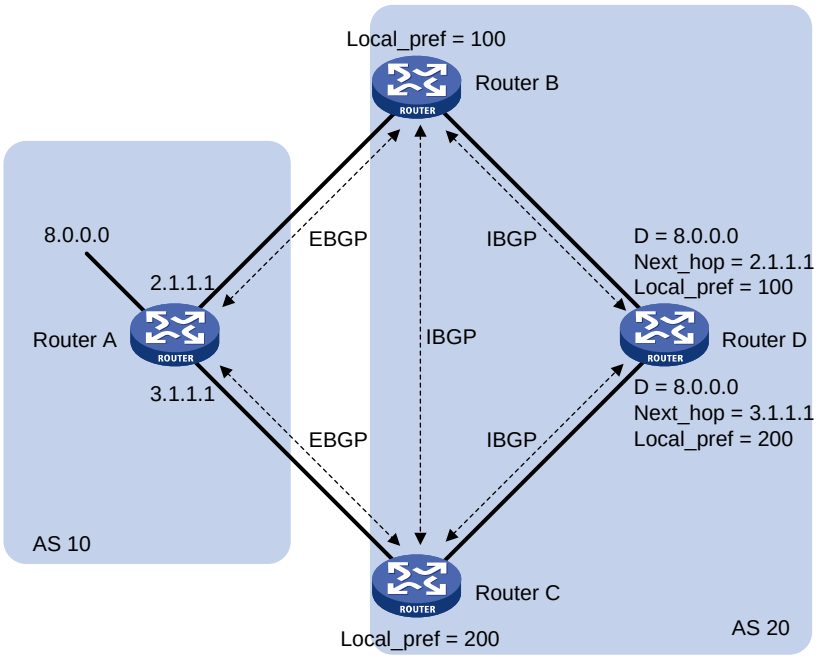

通常情况下，BGP 只比较来自同一个 AS 的路由的 MED 属性值。在某些特殊的应用中，用户也可以通过配置 **compare-different-as-med** 命令，强制 BGP 比较来自不同 AS 的路由的 MED 属性值。

### 5. 本地优先（LOCAL_PREF）属性

LOCAL_PREF 属性仅在 IBGP 对等体之间交换，不通告给其他 AS。它表明 BGP 路由器的优先级。

LOCAL_PREF属性用于判断流量离开AS时的最佳路由。当BGP路由器通过不同的IBGP对等体得到目的地址相同但下一跳不同的多条路由时，将优先选择LOCAL_PREF属性值较高的路由。如 图 1-4 所示，从AS 20 到AS 10 的流量将选择Router C作为出口。

图1-4 LOCAL_PREF 属性




### 6. 团体（COMMUNITY）属性

BGP 将具有相同特征的路由归为一组，称为一个团体，通过在路由中携带团体属性标识路由所属的团体。团体没有物理上的边界，不同 AS 的路由可以属于同一个团体。

根据需要，一条路由可以携带一个或多个团体属性值（每个团体属性值用一个四字节的整数表示）。接收到该路由的路由器可以通过比较团体属性值对路由作出适当的处理（比如决定是否发布该路由、在什么范围发布等），而不需要匹配复杂的过滤规则（如 ACL），从而简化路由策略的应用和降低维护管理的难度。

公认的团体属性有：

- INTERNET：缺省情况下，所有的路由都属于 INTERNET 团体。具有此属性的路由可以被通告给所有的 BGP 对等体。
- NO_EXPORT：具有此属性的路由在收到后，不能被发布到本地AS之外。如果使用了联盟，则不能被发布到联盟之外，但可以发布给联盟中的其他子AS（关于联盟的定义请参见“1.1.7 6. 联盟”）。
- NO_ADVERTISE：具有此属性的路由被接收后，不能被通告给任何其他的 BGP 对等体。
- NO_EXPORT_SUBCONFED：具有此属性的路由被接收后，不能被发布到本地 AS 之外，也不能发布到联盟中的其他子 AS。

除了公认的团体属性外，用户还可以使用团体属性列表自定义团体属性，以便更为灵活地控制路由策略。

### 7. 扩展团体属性

随着团体属性的应用日益广泛，原有四字节的团体属性无法满足用户的需求。因此，BGP 定义了新的路由属性——扩展团体属性。扩展团体属性与团体属性有如下不同：

- 扩展团体属性为八字节，提供了更多的属性值。

- 扩展团体属性可以划分类型。在不同的组网应用中，可以使用不同类型的扩展团体属性对路由进行过滤和控制。与不区分类型、统一使用同一个属性值空间的团体属性相比，扩展团体属性的配置和管理更为简单。

目前，设备支持的扩展团体属性有 VPN Target 属性和 SoO（Site of Origin，源站点）属性。VPN Target 属性的详细介绍，请参见“MPLS 配置指导”中的“MPLS L3VPN”。

SoO 扩展团体属性用来标识路由的原始站点。路由器不会将带有 SoO 属性的路由发布给该 SoO 标识的站点，确保来自某个站点的路由不会再被发布到该站点，从而避免路由环路。在 AS 路径信息丢失时，可以通过 SoO 属性来避免发生环路。

SoO 属性有三种格式：

- 16 位自治系统号:32 位用户自定义数，例如：101:3。
- 32 位 IP 地址:16 位用户自定义数，例如：192.168.122.15:1。
- 32 位自治系统号:16 位用户自定义数字，其中的自治系统号最小值为 65536。例如：65536:1。

### 1.1.4 BGP的选路规则

目前，BGP 选择路由的过程为：

(1) 丢弃下一跳（NEXT_HOP）不可达的路由；

(2) 优选首选值（Preferred-value）最大的路由；

(3) 优选本地优先级（LOCAL_PREF）最高的路由；

(4) 依次选择 **network** 命令生成的路由、**import-route** 命令引入的路由、聚合路由；

(5) 优选 AS 路径（AS_PATH）最短的路由；

(6) 依次选择 ORIGIN 类型为 IGP、EGP、Incomplete 的路由；

(7) 优选 MED 值最低的路由；

(8) 依次选择从 EBGP、联盟 EBGP、联盟 IBGP、IBGP 学来的路由；

(9) 优选下一跳 Cost 值最低的路由；

(10) 优选 CLUSTER_LIST 长度最短的路由；

(11) 优选 ORIGINATOR_ID 最小的路由；

(12) 优选 Router ID 最小的路由器发布的路由；

(13) 优选 IP 地址最小的对等体发布的路由。

说明

- CLUSTER_ID 为路由反射器的集群 ID，CLUSTER_LIST 由 CLUSTER_ID 序列组成，路由反射器将自己的 CLUSTER_ID 加入 CLUSTER_LIST 中。若路由反射器收到路由中的 CLUSTER_LIST 包含自己的 CLUSTER_ID，则丢弃该路由，从而避免集群内发生环路。
- 如果配置了负载分担，并且有多条到达同一目的地的路由，则根据配置的路由条数选择多条路由进行负载分担。

### 1.1.5 BGP发布路由的策略

BGP 发布路由时采用如下策略：

- 存在多条有效路由时，BGP 发言者只将最优路由发布给对等体。如果配置了 **advertise-rib-active** 命令，则 BGP 发布 IP 路由表中的最优路由；否则，发布 BGP 路由表中的最优路由。
- BGP 发言者只把自己使用的路由发布给对等体。
- BGP 发言者会将从 EBGP 获得的路由发布给它的所有 BGP 对等体（包括 EBGP 对等体和 IBGP 对等体）。
- BGP 发言者会将从 IBGP 获得的路由发布给它的 EBGP 对等体，但不会发布给它的 IBGP 对等体。
- 会话一旦建立，BGP 发言者将把满足上述条件的所有 BGP 路由发布给新对等体。之后，BGP 发言者只在路由变化时，向对等体发布更新的路由。

### 1.1.6 BGP负载分担

BGP 可以通过如下两种方式实现负载分担：

- 基于迭代路由实现负载分担
- 通过改变 BGP 选路规则实现负载分担

#### 1. 基于迭代路由实现BGP负载分担

由于 BGP 协议本身的特殊性，它产生的路由的下一跳地址可能不是当前路由器直接相连的邻居。常见的一个原因是：IBGP 之间发布路由信息时不改变下一跳。这种情况下，为了能够将报文正确转发出去，路由器必须先找到一个直接可达的地址（查找 IGP 建立的路由表项），通过这个地址到达路由表中指示的下一跳。在上述过程中，去往直接可达地址的路由被称为依赖路由，BGP 路由依赖于这些路由指导报文转发。根据下一跳地址找到依赖路由的过程就是路由迭代。

目前系统支持基于迭代的 BGP 负载分担，即如果依赖路由本身是负载分担的（假设有三个下一跳地址），则 BGP 也会生成与依赖路由数量相同的下一跳地址来指导报文转发。需要说明的是，基于迭代的 BGP 负载分担并不需要命令配置，这一特性在系统上始终启用。

#### 2. 通过改变BGP选路规则实现负载分担

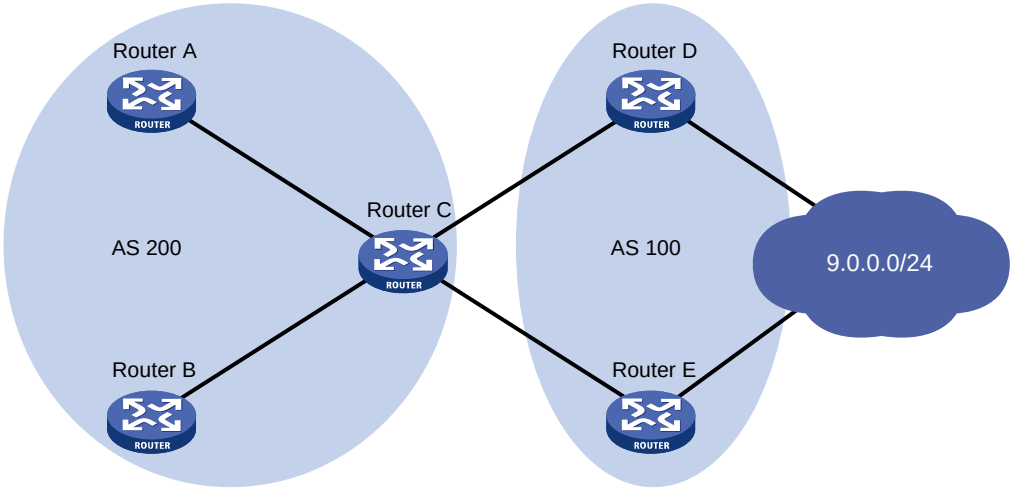
在实现方法上，BGP 的负载分担与 IGP 的负载分担有所不同：

- IGP（如 RIP、OSPF）是通过协议定义的路由算法，对到达同一目的地址的不同路由，根据计算结果，将度量值（metric）相等的路由进行负载分担，选择的标准很明确（按 metric）。
- BGP 本身并没有路由计算的算法，它只是一个选路的路由协议，因此，不能根据一个明确的度量值决定是否对路由进行负载分担，但 BGP 有丰富的选路规则，可以在对路由进行一定的选择后，有条件地进行负载分担，也就是将负载分担加入到 BGP 的选路规则中去。

采用本方式进行负载分担时，BGP不再按照“1.1.4 BGP的选路规则”中的规则选择路由，当路由同时满足如下条件时，即在这些路由间进行负载分担：

- 路由的 AS_PATH 属性、ORIGIN 属性、LOCAL_PREF 属性和 MED 属性完全相同。
- 同为经过路由反射器反射的路由，或同为未经路由反射器反射的路由。

图1-5 BGP 负载分担示意图




在 图 1-5 中，Router A和Router B是Router C的IBGP对等体。当Router D和Router E同时向Router C通告到达同一目的地的路由时，如果用户在Router C上配置了进行负载分担的BGP路由条数为 2，则当满足一定的选路规则后，并且两条路由具有相同的AS_PATH属性、ORIGIN属性、LOCAL_PREF属性和MED属性时，Router C就把接收的两条路由同时加入到转发表中，实现BGP路由的负载分担。Router C只向Router A和Router B转发一次该路由，AS_PATH不变，但NEXT_HOP属性改变为Router C的地址，而不是原来的EBGP对等体地址。Router C通告给Router A和Router B的路由中，其它的BGP路由属性为最佳路由的属性。

说明

BGP 负载分担特性适用于 EBGP、IBGP 以及联盟之间。

### 1.1.7 大规模BGP网络所遇问题的解决方法

在大规模 BGP 网络中，对等体的数目众多，路由表庞大，配置和维护极为不便。通过如下方法，可以降低管理难度，提高路由发布效率。

#### 1. 路由聚合

在大规模的网络中，BGP 路由表十分庞大，使用路由聚合（Routes Aggregation）可以大大减小 BGP 路由表的规模。

路由聚合实际上是将多条路由合并的过程。这样 BGP 在向对等体通告路由时，可以只通告聚合后的路由，而不是将所有的具体路由都通告出去。

目前系统支持自动聚合和手动聚合方式。使用后者还可以控制聚合路由的属性，以及决定是否发布具体路由。

#### 2. 路由衰减

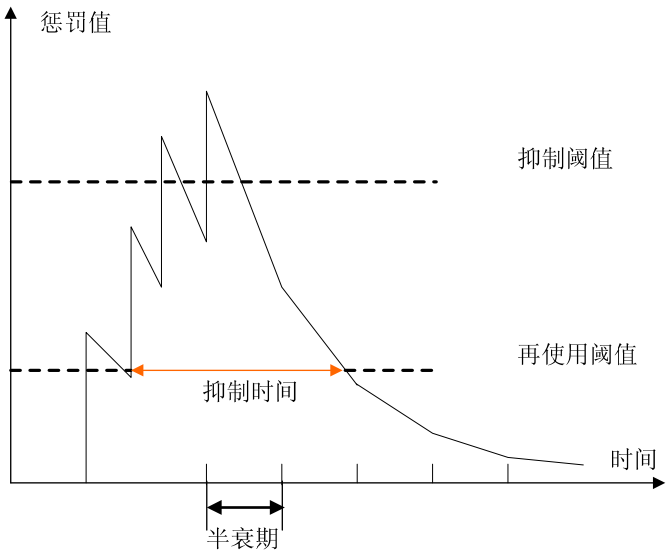
路由发生变化时，路由协议会向邻居发布路由更新，收到路由更新的路由器需要重新计算路由并修改路由表。如果发生路由振荡，即路由不稳定，路由表中的某条路由反复消失和重现，则会消耗大量的带宽资源和 CPU 资源，严重时会影响到网络的正常工作。

在多数情况下，BGP 协议都应用于复杂的网络环境中，路由变化十分频繁。为了防止持续的路由振荡带来的不利影响，BGP 使用衰减来抑制不稳定的路由。

BGP 衰减使用惩罚值来衡量一条路由的稳定性，惩罚值越高说明路由越不稳定。路由每次从可达状态变为不可达状态，或者可达路由的属性每次发生变化时，BGP 给此路由增加一定的惩罚值（系统固定为 1000，不可修改）。当惩罚值超过抑制阈值时，此路由被抑制，不参与路由优选。惩罚值达到设置的上限后，不再继续增加。

发生振荡的路由如果没有再次振荡，则路由的惩罚值会逐渐减少。每经过一段时间，惩罚值便会减少一半，这个时间称为半衰期（Half-life）。当惩罚值低于再使用阈值时，此路由变为可用路由，参与路由优选。

图1-6 BGP 路由衰减示意图




### 3. 对等体组

在大规模 BGP 网络中，对等体的数量很多，其中很多对等体具有相同的策略，在配置时会重复使用一些命令。此时，将这些对等体加入一个对等体组，可以简化配置。

对等体组是具有某些相同属性的对等体的集合。当一个对等体加入对等体组时，此对等体将获得与所在对等体组相同的配置。当对等体组的配置改变时，组内成员的配置也相应改变。

### 4. 团体

在大规模的网络中，如果通过地址前缀列表、ACL、AS_PATH等实现对路由的控制，不仅配置复杂，而且不方便维护。利用团体属性和扩展团体属性，可以提高路由策略配置的灵活度，简化路由策略的管理，从而降低维护管理的难度。团体属性和扩展团体属性的介绍请参见“1.1.3 BGP的路由属性”。

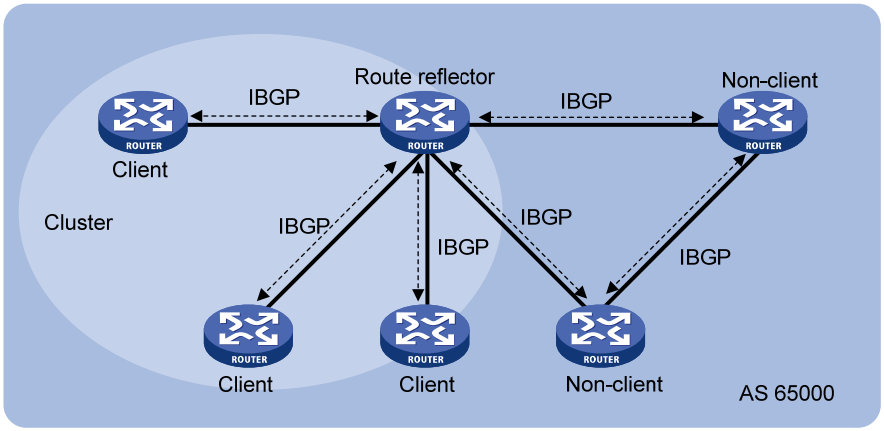
### 5. 路由反射器

为保证 IBGP 对等体之间的连通性，需要在 IBGP 对等体之间建立全连接关系。假设在一个 AS 内部有 n 台路由器，那么应该建立的 IBGP 连接数就为 n(n-1)/2。当 IBGP 对等体数目很多时，对网络资源和 CPU 资源的消耗都很大。

利用路由反射可以解决这一问题。在一个 AS 内，其中一台路由器作为 RR（Route Reflector，路由反射器），作为客户机（Client）的路由器与路由反射器之间建立 IBGP 连接。路由反射器从客户机接收到路由后，将其传递（反射）给所有其他的客户机，从而保证客户机之间不需要建立 BGP 连接，就可以学习到彼此的路由。

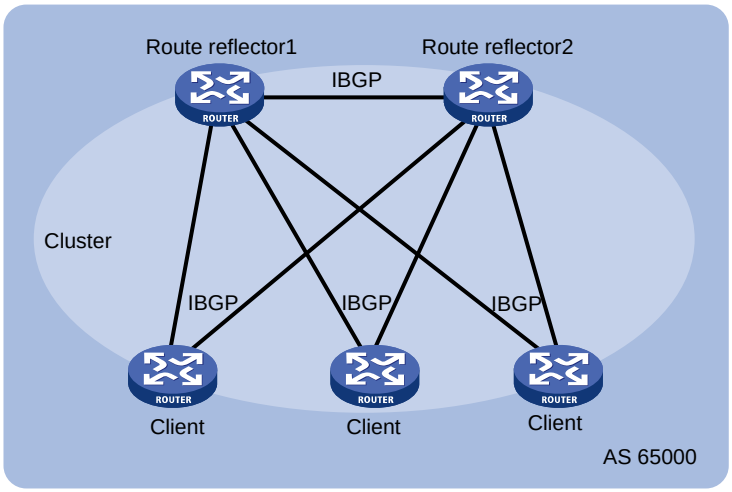
既不是路由反射器也不是客户机的BGP路由器被称为非客户机（Non-client）。非客户机与路由反射器之间，以及所有的非客户机之间仍然必须建立全连接关系。其示意图如 图 1-7 所示。

图1-7 路由反射器示意图




路由反射器及其客户机形成了一个集群。通常情况下，一个集群中只有一个路由反射器，该反射器的Router ID就作为集群ID，用于识别该群。如 图 1-8 所示，为了提高网络的可靠性、避免单点故障，一个集群中可以设置多个路由反射器。此时，集群中所有路由反射器上都需要配置相同的集群ID，以便集群具有统一的标识，避免路由环路的产生。

图1-8 多路由反射器




如果配置了路由反射器后，由于组网需要在路由反射器的客户机之间又建立了全连接，则客户机之间可以直接交换路由信息，客户机到客户机之间的路由反射是没有必要的。此时，不需要修改网络配置或改变网络拓扑，只需在路由反射器上通过相关命令禁止其在客户机之间反射路由，就可以避免路由反射，减少占用的带宽资源。

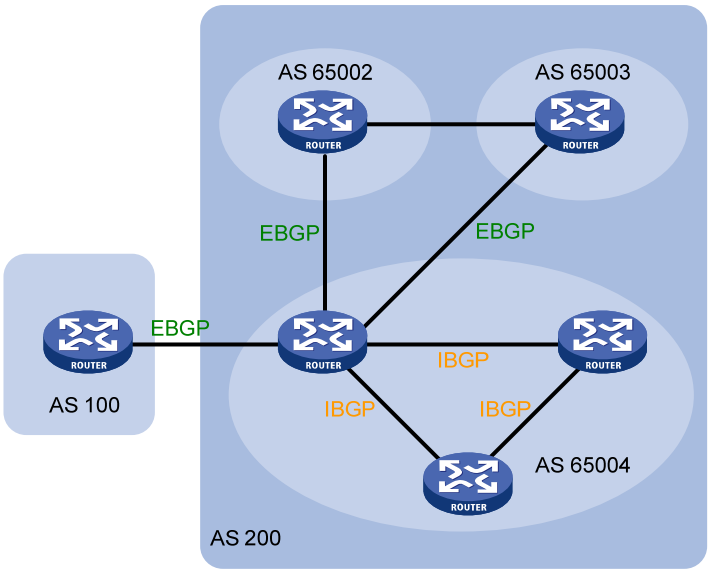
说明

禁止客户机之间的路由反射后，客户机到非客户机之间的路由仍然可以被反射。

### 6. 联盟

联盟（Confederation）是处理自治系统内部的IBGP网络连接激增的另一种方法，它将一个自治系统划分为若干个子自治系统，每个子自治系统内部的IBGP对等体建立全连接关系，子自治系统之间建立联盟内部EBGP连接关系。其示意图如 图 1-9 所示。

图1-9 联盟示意图




在不属于联盟的 BGP 发言者看来，属于同一个联盟的多个子自治系统是一个整体，外界不需要了解内部的子自治系统情况，联盟 ID 就是标识联盟这一整体的自治系统号，如上图中的 AS 200 就是联盟 ID。

联盟的缺陷是从非联盟方案向联盟方案转变时，要求路由器重新进行配置，逻辑拓扑也要改变。

在大型 BGP 网络中，路由反射器和联盟可以被同时使用。

## 1.1.8 MP-BGP

### 1. MP-BGP概述

BGP-4 只能传递 IPv4 单播的路由信息，不能传递其它网络层协议（如 IPv6 等）的路由信息。

为了提供对多种网络层协议的支持，IETF 对 BGP-4 进行了扩展，形成 MP-BGP（Multiprotocol Border Gateway Protocol，多协议边界网关协议）。MP-BGP 可以为多种网络层协议传递路由信息，如 IPv6 单播、IPv4 组播、IPv6 组播、VPNv4、VPNv6、L2VPN、IPv4 MDT 等。

- IPv6 单播

  通过 MP-BGP 发布和维护 IPv6 单播路由前缀信息。

- IPv4 组播/IPv6 组播

组播路由协议 PIM（Protocol Independent Multicast，协议无关组播）根据单播静态路由或者任意单播路由协议（包括 RIP、OSPF、IS-IS、BGP 等）所生成的单播路由表进行 RPF（Reverse Path Forwarding，逆向路径转发）检查，以创建组播路由表项，从而进行组播报文的转发。组播转发路径与单播转发路径是一致的。但是，在某些情况下，组播网络拓扑和单播网络拓扑有可能不同；有些用户希望组播转发路径不同于单播转发路径，以便分别对组播流量和单播流量进行管理和控制。

MP-BGP 对 IPv4 组播/IPv6 组播的扩展，称为 MBGP（Multicast BGP，组播 BGP），它通过 MP-BGP 传递用于 RPF 检查的路由信息，并将该信息保存在独立的组播 BGP 路由表中，以实现单播转发和组播转发的隔离，使得组播转发路径可以不同于单播转发路径。

有关组播、PIM 和 RPF 检查的详细介绍，请参见“IP 组播配置指导”。

- VPNv4/VPNv6

  通过 MP-BGP 发布和维护 VPNv4/VPNv6 路由前缀信息，详细介绍请参见“MPLS 配置指导”中的“MPLS L3VPN”和“IPv6 MPLS L3VPN”。

- L2VPN

  通过 MP-BGP 发布和维护标签块信息和远端邻居信息，详细介绍请参见“MPLS 配置指导”中的“MPLS L2VPN”和“VPLS”。

- IPv4 MDT

  通过 MP-BGP 发布和维护包含 PE 地址及 PE 所在的 Default-group 等信息的 MDT 信息，以便组播 VPN 根据 MDT 信息在公网上建立以 PE 为根（即组播源）的 Default-MDT。详细介绍请参见“IP 组播配置指导”中的“组播 VPN”。

### 2. MP-BGP的扩展属性

路由信息中与网络层协议相关的关键信息包括路由前缀和下一跳地址。BGP-4 通过 Update 消息中的 NLRI（Network Layer Reachability Information，网络层可达性信息）字段携带可达路由的前缀信息，Withdrawn Routes 字段携带不可达路由的前缀信息，NEXT_HOP 属性携带下一跳地址信息。NLRI 字段、Withdrawn Routes 字段和 NEXT_HOP 属性不易于扩展，无法携带多种网络层协议的信息。

为实现对多种网络层协议的支持，MP-BGP 定义了两个新的路径属性：

- MP_REACH_NLRI（Multiprotocol Reachable NLRI，多协议可达 NLRI）：用于携带多种网络层协议的可达路由前缀及下一跳地址信息，以便向邻居发布该路由。
- MP_UNREACH_NLRI（Multiprotocol Unreachable NLRI，多协议不可达 NLRI）：用于携带多种网络层协议的不可达路由前缀信息，以便撤销该路由。

MP-BGP 通过上述两个路径属性传递不同网络层协议的可达路由和不可达路由信息。不支持 MP-BGP 的 BGP 发言者接收到带有这两个属性的 Update 消息后，忽略这两个属性，不把它们传递给其它邻居。

### 3. 地址族

MP-BGP 采用地址族（Address Family）和子地址族（Subsequent Address Family）来区分 MP_REACH_NLRI 属性、MP_UNREACH_NLRI 属性中携带路由信息所属的网络层协议。例如，如果 MP_REACH_NLRI 属性中 AFI（Address Family Identifier，地址族标识符）为 2、SAFI（Subsequent Address Family Identifier，子地址族标识符）为 1，则表示该属性中携带的是 IPv6 单播路由信息。关于地址族的一些取值可以参考 RFC 1700。

### 1.1.9 BGP相关视图介绍

设备为 BGP 定义了多种视图，分别用来管理不同地址族、不同 VPN 实例的路由信息。BGP 支持 VPN 多实例，可以为不同的 VPN 实例维护独立的路由表。

大多数BGP配置命令可以在多个视图下执行，不同视图下命令的作用范围有所不同，详细介绍如 表1-1 所示。

表1-1 BGP 相关视图介绍

| 视图名称 | 进入视图方法 | 说明 |
| --- | --- | --- |
| BGP视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] | 该视图下有些配置对公网和所有VPN实例内所有地址族的路由和对等体生效（如联盟、GR、日志功能的配置等），有些配置只对公网内所有地址族的路由和对等体生效 |
| BGP IPv4单播地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] address-family ipv4 unicast<br>[Sysname-bgp-ipv4] | 该视图下的配置对公网内的IPv4单播路由和对等体生效 |
| BGP IPv6单播地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] address-family ipv6 unicast<br>[Sysname-bgp-ipv6] | 该视图下的配置对公网内的IPv6单播路由和对等体生效 |
| BGP IPv4组播地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] address-family ipv4 multicast<br>[Sysname-bgp-mul-ipv4] | 该视图下的配置对IPv4组播路由和对等体生效 |
| BGP IPv6组播地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] address-family ipv6 multicast<br>[Sysname-bgp-mul-ipv6] | 该视图下的配置对IPv6组播路由和对等体生效 |
| BGP VPNv4地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] address-family vpnv4<br>[Sysname-bgp-vpnv4] | 该视图下的配置对VPNv4路由和对等体生效<br>BGP VPNv4地址族视图的配置请参见“MPLS配置指导”中的“MPLS L3VPN” |
| BGP VPNv6地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] address-family vpnv6<br>[Sysname-bgp-vpnv6] | 该视图下的配置对VPNv6路由和对等体生效<br>BGP VPNv6地址族视图的配置请参见“MPLS配置指导”中的“IPv6 MPLS L3VPN” |

| 视图名称 | 进入视图方法 | 说明 |
| --- | --- | --- |
| BGP L2VPN地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] address-family l2vpn<br>[Sysname-bgp-l2vpn] | 该视图下的配置对L2VPN对等体和L2VPN信息生效<br>BGP L2VPN地址族视图的配置请参见“MPLS配置指导”中的“MPLS L2VPN”和“VPLS” |
| BGP-VPN实例视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] ip vpn-instance vpn1<br>[Sysname-bgp-vpn1] | 该视图下的配置对指定VPN实例内所有地址族的路由和对等体生效 |
| BGP-VPN IPv4单播地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] ip vpn-instance vpn1<br>[Sysname-bgp-vpn1] address-family ipv4 unicast<br>[Sysname-bgp-ipv4-vpn1] | 该视图下的配置对指定VPN实例内的IPv4单播路由和对等体生效 |
| BGP-VPN IPv6单播地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] ip vpn-instance vpn1<br>[Sysname-bgp-vpn1] address-family ipv6 unicast<br>[Sysname-bgp-ipv6-vpn1] | 该视图下的配置对指定VPN实例内的IPv6单播路由和对等体生效 |
| BGP-VPN VPNv4地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] ip vpn-instance vpn1<br>[Sysname-bgp-vpn1] address-family vpnv4<br>[Sysname-bgp-vpnv4-vpn1] | 该视图下的配置对指定VPN实例内的VPNv4路由和对等体生效<br>BGP-VPN VPNv4地址族视图的配置请参见“MPLS配置指导”中的“MPLS L3VPN” |
| BGP IPv4 MDT地址族视图 | <Sysname> system-view<br>[Sysname] bgp 100<br>[Sysname-bgp] address-family ipv4 mdt<br>[Sysname-bgp-mdt] | 该视图下的配置对IPv4 MDT路由和对等体生效 |

### 1.1.10 协议规范

与 BGP 相关的协议规范有：

- RFC 1700：ASSIGNED NUMBERS
- RFC 1771：A Border Gateway Protocol 4 (BGP-4)
- RFC 1997：BGP Communities Attribute
- RFC 2439：BGP Route Flap Damping
- RFC 2796：BGP Route Reflection
- RFC 2858：Multiprotocol Extensions for BGP-4
- RFC 2918：Route Refresh Capability for BGP-4
- RFC 3065：Autonomous System Confederations for BGP

- RFC 3392：Capabilities Advertisement with BGP-4
- RFC 4271：A Border Gateway Protocol 4 (BGP-4)
- RFC 4360：BGP Extended Communities Attribute
- RFC 4724：Graceful Restart Mechanism for BGP
- RFC 4760：Multiprotocol Extensions for BGP-4
- RFC 5082：The Generalized TTL Security Mechanism (GTSM)
- RFC 6037：Cisco Systems' Solution for Multicast in BGP MPLS IP VPNs

## 1.2 BGP配置任务简介

在最基本的 BGP 网络中，只需完成如下配置：

- 启动 BGP。
- 配置 BGP 对等体或对等体组。如果分别对对等体组和对等体组中的对等体进行了某项 BGP 配置，则以最后一次配置为准。
- 控制 BGP 路由信息的生成。

如果在 BGP 网络中，需要对 BGP 路由信息的发布、BGP 路径的选择等进行控制，则可以根据需要进行其他配置。

**表1-2 BGP 配置任务简介（IPv4 单播/IPv4 组播）**

| 配置任务 | | 说明 | 详细配置 |
|---|---|---|---|
| 配置BGP基本功能 | 启动BGP | 必选 | 1.3.1 |
| | 配置BGP对等体 | 二者必选其一<br>建议在大规模的BGP网络中选择“配置BGP对等体组”，以便简化配置 | 1.3.2 |
| | 配置BGP对等体组 | | 1.3.3 |
| | 配置建立TCP连接使用的源地址 | 可选 | 1.3.4 |
| 控制BGP路由信息的生成 | 配置BGP发布本地路由 | 二者至少选其一 | 1.4.1 |
| | 配置BGP引入IGP路由协议的路由 | | 1.4.2 |
| 控制BGP路由信息的发布与接收 | 配置BGP路由聚合 | 可选 | 1.5.1 |
| | 配置发布IP路由表中的最优路由 | | 1.5.2 |
| | 配置向对等体/对等体组发送缺省路由 | | 1.5.3 |
| | 限制从BGP对等体/对等体组接收的路由数量 | | 1.5.4 |
| | 配置BGP路由信息的发布/接收策略 | | 1.5.5 |
| | 配置BGP路由衰减 | | 1.5.6 |
| 控制BGP路径的选择 | 为接收路由分配首选值 | 可选 | 1.6.1 |
| | 配置BGP的路由优先级 | | 1.6.2 |

| 配置任务 | | 说明 | 详细配置 |
| --- | --- | --- | --- |
| | 配置本地优先级的缺省值 | | 1.6.3 |
| | 配置MED属性 | | 1.6.4 |
| | 配置NEXT_HOP属性 | | 1.6.5 |
| | 配置AS_PATH属性 | | 1.6.6 |
| | 配置SoO属性 | | 1.6.7 |
| 调整和优化BGP网络 | 配置BGP会话的存活时间间隔与保持时间 | 可选 | 1.7.1 |
| | 配置发布同一路由的时间间隔 | | 1.7.2 |
| | 配置允许同非直连邻居建立EBGP会话 | | 1.7.3 |
| | 使能直连EBGP会话快速复位功能 | | 1.7.4 |
| | 使能4字节AS号抑制功能 | | 1.7.5 |
| | 配置BGP的MD5认证 | | 1.7.6 |
| | 配置BGP负载分担 | | 1.7.7 |
| | 禁止与对等体/对等体组建立会话 | | 1.7.9 |
| | 配置BGP GTSM功能 | | 1.7.10 |
| | 配置BGP软复位 | | 1.7.11 |
| | 配置系统进入二级内存门限告警状态后不断开EBGP对等体 | | 1.7.12 |
| 配置大规模BGP网络 | 配置BGP团体 | 可选 | 1.8.1 |
| | 配置BGP路由反射 | | 1.8.2 |
| | 配置BGP联盟 | | 1.8.3 |
| 配置BGP GR | | 可选 | 1.9 |
| 配置BGP NSR | | 可选 | 1.10 |
| 开启告警功能 | | 可选 | 1.11 |
| 使能BGP日志功能 | | 可选 | 1.12 |
| 配置BGP与BFD联动 | | 可选 | 1.13 |
| 配置BGP快速重路由 | | 可选 | 1.14 |

表1-3 BGP 配置任务简介（IPv6 单播/IPv6 组播）

| 配置任务 | | 说明 | 详细配置 |
| --- | --- | --- | --- |
| 配置BGP基本功能 | 启动BGP | 必选 | 1.3.1 |
| | 配置BGP对等体 | 二者必选其一 | 1.3.2 |

| 配置任务 | | 说明 | 详细配置 |
| --- | --- | --- | --- |
| | 配置BGP对等体组 | 建议在大规模的BGP网络中选择“配置BGP对等体组”，以便简化配置 | 1.3.3 |
| | 配置建立TCP连接使用的源地址 | 可选 | 1.3.4 |
| 控制BGP路由信息的生成 | 配置BGP发布本地路由 | 二者至少选其一 | 1.4.1 |
| | 配置BGP引入IGP路由协议的路由 | | 1.4.2 |
| 控制BGP路由信息的发布与接收 | 配置BGP路由聚合 | 可选 | 1.5.1 |
| | 配置向所有对等体发布IP路由表中的最优路由 | | 1.5.2 |
| | 配置向对等体/对等体组发送缺省路由 | | 1.5.3 |
| | 限制从BGP对等体/对等体组接收的路由数量 | | 1.5.4 |
| | 配置BGP路由信息的发布/接收策略 | | 1.5.5 |
| | 配置BGP路由衰减 | | 1.5.6 |
| 控制BGP路径的选择 | 为接收路由分配首选值 | 可选 | 1.6.1 |
| | 配置BGP的路由优先级 | | 1.6.2 |
| | 配置本地优先级的缺省值 | | 1.6.3 |
| | 配置MED属性 | | 1.6.4 |
| | 配置NEXT_HOP属性 | | 1.6.5 |
| | 配置AS_PATH属性 | | 1.6.6 |
| | 配置SoO属性 | | 1.6.7 |
| 调整和优化BGP网络 | 配置BGP会话的存活时间间隔与保持时间 | 可选 | 1.7.1 |
| | 配置发布同一路由的时间间隔 | | 1.7.2 |
| | 配置允许同非直连邻居建立EBGP会话 | | 1.7.3 |
| | 使能直连EBGP会话快速复位功能 | | 1.7.4 |
| | 使能4字节AS号抑制功能 | | 1.7.5 |
| | 配置BGP的MD5认证 | | 1.7.6 |
| | 配置BGP负载分担 | | 1.7.7 |
| | 配置通过IPsec保护IPv6 BGP报文 | | 1.7.8 |
| | 禁止与对等体/对等体组建立会话 | | 1.7.9 |
| | 配置BGP GTSM功能 | | 1.7.10 |
| | 配置BGP软复位 | | 1.7.11 |

| 配置任务 | | 说明 | 详细配置 |
| --- | --- | --- | --- |
| | 配置系统进入二级内存门限告警状态后不断开EBGP对等体 | | 1.7.12 |
| 配置大规模BGP网络 | 配置BGP团体 | 可选 | 1.8.1 |
| | 配置BGP路由反射 | | 1.8.2 |
| | 配置BGP联盟 | | 1.8.3 |
| 配置BGP GR | | 可选 | 1.9 |
| 配置BGP NSR | | 可选 | 1.10 |
| 开启告警功能 | | 可选 | 1.11 |
| 使能BGP日志功能 | | 可选 | 1.12 |
| 配置BGP与BFD联动 | | 可选 | 1.13 |
| 配置BGP快速重路由 | | 可选 | 1.14 |
| 配置6PE | | 可选<br>IPv6组播不支持本功能 | 1.15 |

# 1.3 配置BGP基本功能

## 1.3.1 启动BGP

一台路由器如果要运行 BGP 协议，则必须存在 Router ID。Router ID 用来在一个自治系统中唯一的标识一台路由器。

- 用户可以在启动 BGP 进入 BGP 视图后指定 Router ID，配置时，必须保证自治系统中任意两台路由器的 ID 都不相同。通常的做法是将路由器的 ID 配置为与该路由器某个接口的 IP 地址一致，为了增加网络的可靠性，建议将 Router ID 配置为 Loopback 接口的 IP 地址。
- 如果没有在 BGP 视图下配置 Router ID，则使用全局 Router ID。
- BGP 的 Router ID 一旦确定为非零值后不会随着系统视图下 **router id** 命令配置的改变而改变。只能在 BGP 视图下通过 **router-id** 命令改变 BGP 的 Router ID。
- 如果是在 BGP 视图下配置的 Router ID，则 Router ID 所在接口被删除时路由器不会重新选择 Router ID，只有在 BGP 视图下使用 **undo router-id** 命令删除手工配置的 Router ID 后，路由器才会重新选择 Router ID。

表1-4 启动 BGP

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 配置全局Router ID | | **router id** *router-id* | 缺省情况下，未配置全局Router ID<br>如果没有配置全局Router ID，则按照下面的规则进行选择：<br>(1) 如果存在配置 IP 地址的 Loopback 接口，则选择 Loopback 接口地址中最大的作为 Router ID<br>(2) 如果所有 Loopback 接口都没有配置 IP 地址，则从其他接口的 IP 地址中选择最大的作为 Router ID（不考虑接口的 up/down 状态） |
| 启动BGP，并进入BGP视图或BGP-VPN实例视图 | 启动BGP，并进入BGP视图 | **bgp** *as-number* | 缺省情况下，系统没有运行BGP<br>一台路由器只能位于一个AS内，一台路由器上只能启动一个BGP进程 |
| | 启动BGP，并进入BGP-VPN实例视图 | **bgp** *as-number* | 进入BGP-VPN实例视图时，指定的VPN实例必须已经创建，且VPN实例内必须配置RD（Route Distinguisher，路由标识符）。VPN实例的详细介绍，请参见“MPLS配置指导”中的“MPLS L3VPN” |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置路由器的Router ID | | **router-id** { *router-id* \| **auto-select** } | 缺省情况下，BGP路由器的Router ID与全局Router ID相同<br>仅在BGP-VPN实例视图下支持**auto-select**关键字 |

## 1.3.2 配置BGP对等体

表1-5 配置 BGP 对等体（IPv4 单播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建IPv4 BGP对等体，并指定对等体的AS号 | | **peer** *ip-address* **as-number** *as-number* | 缺省情况下，设备上不存在任何IPv4 BGP对等体 |
| （可选）配置对等体的描述信息 | | **peer** *ip-address* **description** *description-text* | 缺省情况下，对等体没有描述信息 |
| 创建BGP IPv4单播地址族或BGP-VPN IPv4单播地址族，并进入相应地址族视图 | | **address-family ipv4** [ **unicast** ] | 缺省情况下，没有创建BGP IPv4单播地址族和BGP-VPN IPv4单播地址族 |
| 使能本地路由器与指定对等体交换IPv4单播路由信息的能力 | | **peer** *ip-address* **enable** | 缺省情况下，本地路由器不能与对等体交换IPv4单播路由信息 |

表1-6 配置 BGP 对等体（IPv6 单播）

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建IPv6 BGP对等体，并指定对等体的AS号 | | **peer** *ipv6-address* **as-number** *as-number* | 缺省情况下，设备上不存在任何IPv6 BGP对等体 |
| （可选）配置对等体的描述信息 | | **peer** *ipv6-address* **description** *description-text* | 缺省情况下，对等体没有描述信息 |
| 创建BGP IPv6单播地址族或BGP-VPN IPv6单播地址族，并进入相应地址族视图 | | **address-family ipv6** [ **unicast** ] | 缺省情况下，没有创建BGP IPv6单播地址族和BGP-VPN IPv6单播地址族 |
| 使能本地路由器与指定对等体交换IPv6单播路由信息的能力 | | **peer** *ipv6-address* **enable** | 缺省情况下，本地路由器不能与对等体/对等体组交换IPv6单播路由信息 |

表1-7 配置 BGP 对等体（IPv4 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 创建IPv4 BGP对等体，并指定对等体的AS号 | **peer** *ip-address* **as-number** *as-number* | 缺省情况下，设备上不存在任何IPv4 BGP对等体 |
| （可选）配置对等体的描述信息 | **peer** *ip-address* **description** *description-text* | 缺省情况下，对等体没有描述信息 |
| 创建BGP IPv4组播地址族，并进入BGP IPv4组播地址族视图 | **address-family ipv4 multicast** | 缺省情况下，没有创建BGP IPv4组播地址族 |
| 使能本地路由器与指定对等体交换用于RPF检查的IPv4单播路由信息的能力 | **peer** *ip-address* **enable** | 缺省情况下，本地路由器不能与对等体交换用于RPF检查的IPv4单播路由信息 |

表1-8 配置 BGP 对等体（IPv6 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 创建IPv6 BGP对等体，并指定对等体的AS号 | **peer** *ipv6-address* **as-number** *as-number* | 缺省情况下，设备上不存在任何IPv6 BGP对等体 |
| （可选）配置对等体的描述信息 | **peer** *ipv6-address* **description** *description-text* | 缺省情况下，对等体没有描述信息 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 创建BGP IPv6组播地址族，并进入BGP IPv6组播地址族视图 | **address-family ipv6 multicast** | 缺省情况下，没有创建BGP IPv6组播地址族 |
| 使能本地路由器与指定对等体交换用于RPF检查的IPv6单播路由信息的能力 | **peer** *ipv6-address* **enable** | 缺省情况下，本地路由器不能与对等体/对等体组交换用于RPF检查的IPv6单播路由信息 |

### 1.3.3 配置BGP对等体组

对等体组是具有相同更新策略的对等体的集合。

在大型 BGP 网络中，对等体的数量会很多，其中，很多对等体需要配置相同的策略，通过配置对等体组并将对等体加入到对等体组，可以使对等体获得与所在对等体组相同的配置，而且当对等体组的配置改变时，组内成员的配置也相应改变，从而简化配置。

根据对等体所在的 AS，对等体组可分为：

- IBGP 对等体组：对等体组中的对等体与当前路由器位于同一 AS。
- EBGP 对等体组：对等体组中的对等体与当前路由器位于不同 AS。

#### 1. 配置IBGP对等体组

创建 IBGP 对等体组后，系统在将对等体加入 IBGP 对等体组时，会自动在 BGP 视图下创建该对等体，并设置其 AS 号为本地 AS 号。

表1-9 配置 IBGP 对等体组（IPv4 单播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建IBGP对等体组 | | **group** *group-name* [ **internal** ] | 缺省情况下，设备上不存在任何对等体组 |
| 向对等体组中添加指定的IPv4 BGP对等体 | | **peer** *ip-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和本地的AS号一致 |
| （可选）配置对等体组的描述信息 | | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv4单播地址族或BGP-VPN IPv4单播地址族，并进入相应地址族视图 | | **address-family ipv4** [ **unicast** ] | 缺省情况下，没有创建BGP IPv4单播地址族和BGP-VPN IPv4单播地址族 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 使能本地路由器与指定对等体组中的对等体交换IPv4单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换IPv4单播路由信息 |

表1-10 配置 IBGP 对等体组（IPv6 单播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建IBGP对等体组 | | **group** *group-name* [ **internal** ] | 缺省情况下，设备上不存在任何对等体组 |
| 向对等体组中添加指定的IPv6 BGP对等体 | | **peer** *ipv6-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和本地的AS号一致 |
| （可选）配置对等体组的描述信息 | | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv6单播地址族或BGP-VPN IPv6单播地址族，并进入相应地址族视图 | | **address-family ipv6** [ **unicast** ] | 缺省情况下，没有创建BGP IPv6单播地址族和BGP-VPN IPv6单播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换IPv6单播路由信息的能力 | | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换IPv6单播路由信息 |

表1-11 配置 IBGP 对等体组（IPv4 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 创建IBGP对等体组 | **group** *group-name* [ **internal** ] | 缺省情况下，设备上不存在任何对等体组 |
| 向对等体组中添加指定的IPv4 BGP对等体 | **peer** *ip-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和本地的AS号一致 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 创建BGP IPv4组播地址族，并进入BGP IPv4组播地址族视图 | **address-family ipv4 multicast** | 缺省情况下，没有创建BGP IPv4组播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换用于RPF检查的IPv4单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换用于RPF检查的IPv4单播路由信息 |

表1-12 配置 IBGP 对等体组（IPv6 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 创建IBGP对等体组 | **group** *group-name* [ **internal** ] | 缺省情况下，设备上不存在任何对等体组 |
| 向对等体组中添加指定的IPv6 BGP对等体 | **peer** *ipv6-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和本地的AS号一致 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv6组播地址族，并进入BGP IPv6组播地址族视图 | **address-family ipv6 multicast** | 缺省情况下，没有创建BGP IPv6组播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换用于RPF检查的IPv6单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换用于RPF检查的IPv6单播路由信息 |

### 2. 配置EBGP对等体组

根据对等体组中的对等体是否属于同一个外部 AS，EBGP 对等体组又可以分为纯 EBGP 对等体组和混合 EBGP 对等体组。如果对等体组中的对等体属于同一个外部 AS，该对等体组就是纯 EBGP 对等体组；如果对等体组中的对等体属于不同外部 AS，该对等体组就是混合 EBGP 对等体组。

用户有三种方式配置 EBGP 对等体组：

- 第一种方式是创建对等体组后，先指定对等体组的 AS 号，再将对等体加入到对等体组中，该方式下加入的对等体具有相同的 AS 号，均为对等体组的 AS 号。对等体加入对等体组之前可以配置 AS 号，且为对等体配置的 AS 号必须与对等体组的 AS 号相同。
- 第二种方式是创建对等体组后，先配置对等体的 AS 号，再将对等体加入对等体组中。该方式下，对等体组中对等体的 AS 号可以相同也可以不同。
- 第三种方式是创建对等体组后，将对等体加入对等体组的同时指定 AS 号。该方式下，对等体组中对等体的 AS 号可以相同也可以不同。

(1) 配置 EBGP 对等体组方式一

表1-13 配置 EBGP 对等体组方式一（IPv4 单播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建EBGP对等体组 | | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 指定对等体组的AS号 | | **peer** *group-name* **as-number** *as-number* | 缺省情况下，没有指定对等体组的AS号<br>如果对等体组中已经存在对等体，则不能改变该对等体组的AS号，也不能使用**undo**命令删除已指定的AS号 |
| 向对等体组中添加指定的IPv4 BGP对等体 | | **peer** *ip-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和**peer** *group-name* **as-number** *as-number*命令中配置的一致 |
| （可选）配置对等体组的描述信息 | | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv4单播地址族或BGP-VPN IPv4单播地址族，并进入相应地址族视图 | | **address-family ipv4** [ **unicast** ] | 缺省情况下，没有创建BGP IPv4单播地址族和BGP-VPN IPv4单播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换IPv4单播路由信息的能力 | | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换IPv4单播路由信息 |

表1-14 配置 EBGP 对等体组方式一（IPv6 单播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建EBGP对等体组 | | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 指定对等体组的AS号 | **peer** *group-name* **as-number** *as-number* | 缺省情况下，没有指定对等体组的AS号<br>如果对等体组中已经存在对等体，则不能改变该对等体组的AS号，也不能使用**undo**命令删除已指定的AS号 |
| 向对等体组中添加指定的IPv6 BGP对等体 | **peer** *ipv6-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和**peer** *group-name* **as-number** *as-number*命令中配置的一致 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv6单播地址族或BGP-VPN IPv6单播地址族，并进入相应地址族视图 | **address-family ipv6** [ **unicast** ] | 缺省情况下，没有创建BGP IPv6单播地址族和BGP-VPN IPv6单播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换IPv6单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换IPv6单播路由信息 |

表1-15 配置 EBGP 对等体组方式一（IPv4 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 创建EBGP对等体组 | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 指定对等体组的AS号 | **peer** *group-name* **as-number** *as-number* | 缺省情况下，没有指定对等体组的AS号<br>如果对等体组中已经存在对等体，则不能改变该对等体组的AS号，也不能使用**undo**命令删除已指定的AS号 |
| 向对等体组中添加指定的IPv4 BGP对等体 | **peer** *ip-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和**peer** *group-name* **as-number** *as-number*命令中配置的一致 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv4组播地址族，并进入BGP IPv4组播地址族视图 | **address-family ipv4 multicast** | 缺省情况下，没有创建BGP IPv4组播地址族 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 使能本地路由器与指定对等体组中的对等体交换用于RPF检查的IPv4单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换用于RPF检查的IPv4单播路由信息 |

表1-16 配置 EBGP 对等体组方式一（IPv6 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 创建EBGP对等体组 | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 指定对等体组的AS号 | **peer** *group-name* **as-number** *as-number* | 缺省情况下，没有指定对等体组的AS号<br>如果对等体组中已经存在对等体，则不能改变该对等体组的AS号，也不能使用**undo**命令删除已指定的AS号 |
| 向对等体组中添加指定的IPv6 BGP对等体 | **peer** *ipv6-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和**peer** *group-name* **as-number** *as-number*命令中配置的一致 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv6组播地址族，并进入BGP IPv6组播地址族视图 | **address-family ipv6 multicast** | 缺省情况下，没有创建BGP IPv6组播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换用于RPF检查的IPv6单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换用于RPF检查的IPv6单播路由信息 |

(2) 配置 EBGP 对等体组方式二

表1-17 配置 EBGP 对等体组方式二（IPv4 单播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建EBGP对等体组 | | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 创建IPv4 BGP对等体，并指定对等体的AS号 | **peer** *ip-address* **as-number** *as-number* | 缺省情况下，设备上不存在任何IPv4 BGP对等体 |
| 向对等体组中添加指定的IPv4 BGP对等体 | **peer** *ip-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和**peer** *ip-address* **as-number** *as-number*命令中配置的一致 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv4单播地址族或BGP-VPN IPv4单播地址族，并进入相应地址族视图 | **address-family ipv4** [ **unicast** ] | 缺省情况下，没有创建BGP IPv4单播地址族和BGP-VPN IPv4单播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换IPv4单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换IPv4单播路由信息 |

表1-18 配置 EBGP 对等体组方式二（IPv6 单播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建EBGP对等体组 | | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 创建IPv6 BGP对等体，并指定对等体的AS号 | | **peer** *ipv6-address* **as-number** *as-number* | 缺省情况下，设备上不存在任何IPv6 BGP对等体 |
| 向对等体组中添加指定的IPv6 BGP对等体 | | **peer** *ipv6-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和**peer** *ipv6-address* **as-number** *as-number*命令中配置的一致 |
| （可选）配置对等体组的描述信息 | | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv6单播地址族或BGP-VPN IPv6单播地址族，并进入相应地址族视图 | | **address-family ipv6** [ **unicast** ] | 缺省情况下，没有创建BGP IPv6单播地址族和BGP-VPN IPv6单播地址族 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 使能本地路由器与指定对等体组中的对等体交换IPv6单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换IPv6单播路由信息 |

表1-19 配置 EBGP 对等体组方式二（IPv4 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 创建EBGP对等体组 | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 创建IPv4 BGP对等体，并指定对等体的AS号 | **peer** *ip-address* **as-number** *as-number* | 缺省情况下，设备上不存在任何IPv4 BGP对等体 |
| 向对等体组中添加指定的IPv4 BGP对等体 | **peer** *ip-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和**peer** *ip-address* **as-number** *as-number*命令中配置的一致 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv4组播地址族，并进入BGP IPv4组播地址族视图 | **address-family ipv4 multicast** | 缺省情况下，没有创建BGP IPv4组播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换用于RPF检查的IPv4单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换用于RPF检查的IPv4单播路由信息 |

表1-20 配置 EBGP 对等体组方式二（IPv6 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 创建EBGP对等体组 | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 创建IPv6 BGP对等体，并指定对等体的AS号 | **peer** *ipv6-address* **as-number** *as-number* | 缺省情况下，设备上不存在任何IPv6 BGP对等体 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 向对等体组中添加指定的IPv6 BGP对等体 | **peer** *ipv6-address* **group** *group-name* [ **as-number** *as-number* ] | 缺省情况下，对等体组中不存在任何对等体<br>**as-number** *as-number*参数可选可不选，如果选择则必须和**peer** *ipv6-address* **as-number** *as-number*命令中配置的一致 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv6组播地址族，并进入BGP IPv6组播地址族视图 | **address-family ipv6 multicast** | 缺省情况下，没有创建BGP IPv6组播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换用于RPF检查的IPv6单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换用于RPF检查的IPv6单播路由信息 |

(3) 配置 EBGP 对等体组方式三

表1-21 配置 EBGP 对等体组方式三（IPv4 单播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建EBGP对等体组 | | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 向对等体组中添加指定的对等体 | | **peer** *ip-address* **group** *group-name*<br>**as-number** *as-number* | 缺省情况下，对等体组中不存在任何对等体 |
| （可选）配置对等体组的描述信息 | | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv4单播地址族或BGP-VPN IPv4单播地址族，并进入相应地址族视图 | | **address-family ipv4** [ **unicast** ] | 缺省情况下，没有创建BGP IPv4单播地址族和BGP-VPN IPv4单播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换IPv4单播路由信息的能力 | | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换IPv4单播路由信息 |

表1-22 配置 EBGP 对等体组方式三（IPv6 单播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 创建EBGP对等体组 | | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 向对等体组中添加指定的IPv6 BGP对等体 | | **peer** *ipv6-address* **group** *group-name* **as-number** *as-number* | 缺省情况下，对等体组中不存在任何对等体 |
| （可选）配置对等体组的描述信息 | | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv6单播地址族或BGP-VPN IPv6单播地址族，并进入相应地址族视图 | | **address-family ipv6** [ **unicast** ] | 缺省情况下，没有创建BGP IPv6单播地址族和BGP-VPN IPv6单播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换IPv6单播路由信息的能力 | | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换IPv6单播路由信息 |

表1-23 配置 EBGP 对等体组方式三（IPv4 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 创建EBGP对等体组 | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 向对等体组中添加指定的对等体 | **peer** *ip-address* **group** *group-name* **as-number** *as-number* | 缺省情况下，对等体组中不存在任何对等体 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv4组播地址族，并进入BGP IPv4组播地址族视图 | **address-family ipv4 multicast** | 缺省情况下，没有创建BGP IPv4组播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换用于RPF检查的IPv4单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换用于RPF检查的IPv4单播路由信息 |

表1-24 配置 EBGP 对等体组方式三（IPv6 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 创建EBGP对等体组 | **group** *group-name* **external** | 缺省情况下，设备上不存在任何对等体组 |
| 向对等体组中添加指定的IPv6 BGP对等体 | **peer** *ipv6-address* **group** *group-name* **as-number** *as-number* | 缺省情况下，对等体组中不存在任何对等体 |
| （可选）配置对等体组的描述信息 | **peer** *group-name* **description** *description-text* | 缺省情况下，对等体组没有描述信息 |
| 创建BGP IPv6组播地址族，并进入BGP IPv6组播地址族视图 | **address-family ipv6 multicast** | 缺省情况下，没有创建BGP IPv6组播地址族 |
| 使能本地路由器与指定对等体组中的对等体交换用于RPF检查的IPv6单播路由信息的能力 | **peer** *group-name* **enable** | 缺省情况下，本地路由器不能与对等体交换用于RPF检查的IPv6单播路由信息 |

### 1.3.4 配置建立TCP连接使用的源地址

BGP 使用 TCP 作为其传输层协议。缺省情况下，BGP 使用到达 BGP 对等体的最佳路由出接口的主 IP 地址或 IPv6 地址与对等体/对等体组建立 TCP 连接。在如下场合可以通过本配置指定建立 TCP 连接使用的源地址或源接口（即采用指定源接口的 IP 地址/IPv6 地址与对等体/对等体组建立 TCP 连接）：

- 当指定的对等体的 IP 地址/IPv6 地址不是本地路由器与对等体之间直连接口的 IP 地址/IPv6 地址时，需要在对等体上通过本配置将建立 TCP 连接使用的源接口指定为对等体 IP 地址/IPv6 地址所在的接口。例如，本端设备通过接口 A 和对端设备的接口 B 相连，在本端使用 **peer** x.x.x.x **as-number** *as-number* 命令将对端指定为自己的对等体，但是 x.x.x.x 不是接口 B 的 IP 地址时，需要在对端设备上使用 **peer connect-interface** 命令配置源接口，指定源接口为 IP 地址 x.x.x.x 所在的接口。
- 当建立 BGP 会话的路由器之间存在冗余链路时，如果路由器上的一个接口发生故障，链路状态变为 down，建立 TCP 连接的源地址可能会随之发生变化，导致 BGP 需要重新建立 TCP 连接，造成网络震荡。为了避免该情况的发生，建议网络管理员将建立 TCP 连接所使用的源地址配置为 Loopback 接口的地址，或将源接口配置为 Loopback 接口，以提高 TCP 连接的可靠性和稳定性。
- 当 BGP 对等体之间同时建立多条 BGP 会话时，如果没有明确指定建立 TCP 连接的源地址，可能会导致根据最优路由选择 TCP 连接源地址错误，并影响 BGP 会话的建立。如果多条 BGP 会话基于不同接口的 IP 地址建立，则建议用户在配置 BGP 对等体时，通过配置源接口或源地址明确指定每个 BGP 会话的 TCP 连接源地址；如果多条 BGP 会话基于同一接口的不同 IP 地址建立，则建议用户通过配置源地址，明确指定每个 BGP 会话的 TCP 连接源地址。

**表1-25 配置建立 TCP 连接使用的源地址（IPv4 单播/IPv4 组播）**

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视 | 进入BGP视图 | **bgp** *as-number* | - |

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 图或BGP-VPN实例视图 | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 指定与对等体/对等体组创建BGP会话时建立TCP连接使用的源IP地址 | | **peer** { *group-name* \| *ip-address* } **source-address** *source-ip-address* | 缺省情况下，BGP使用到达BGP对等体的最佳路由出接口的主IP地址与对等体/对等体组建立TCP连接 |
| 配置与对等体/对等体组创建BGP会话时建立TCP连接使用的源接口 | | **peer** { *group-name* \| *ip-address* } **connect-interface** *interface-type interface-number* | |

表1-26 配置建立 TCP 连接使用的源地址（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 指定与对等体/对等体组创建BGP会话时建立TCP连接使用的源IPv6地址 | | **peer** { *group-name* \| *ipv6-address* } **source-address** *source-ipv6-address* | 缺省情况下，BGP使用到达BGP对等体的最佳路由出接口的IPv6地址与对等体/对等体组建立TCP连接 |
| 配置与对等体/对等体组创建BGP会话时建立TCP连接使用的源接口 | | **peer** { *group-name* \| *ipv6-address* } **connect-interface** *interface-type interface-number* | |

## 1.4 控制BGP路由信息的生成

要生成 BGP 路由，可以通过以下一种或几种途径：

- 配置 BGP 发布本地路由
- 配置 BGP 引入 IGP 路由协议的路由

### 1.4.1 配置BGP发布本地路由

通过本配置可以将本地路由表中指定网段的路由添加到 BGP 路由表中，以便通过 BGP 发布该网段路由。通过该种方式发布的路由的 ORIGIN 属性为 IGP。网络管理员还可以通过使用路由策略更为灵活地控制所发布的路由。

本配置中指定的网段路由必须存在于本地的 IP 路由表中，且处于 Active 状态，否则无法将该网段路由添加到 BGP 路由表中。

表1-27 配置 BGP 发布本地路由（IPv4 单播/IPv4 组播）

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 将本地路由表中指定网段的路由添加到BGP路由表中 | | **network** *ip-address* [ *mask* \| *mask-length* ] [ **route-policy** *route-policy-name* ] | 缺省情况下，BGP不发布任何本地的网段路由 |

表1-28 配置 BGP 发布本地路由（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 将本地路由表中指定网段的路由添加到IPv6 BGP路由表中 | | **network** *ipv6-address prefix-length* [ **route-policy** *route-policy-name* ] | 缺省情况下，BGP不发布任何本地的网段路由 |

### 1.4.2 配置BGP引入IGP路由协议的路由

BGP 可以向邻居 AS 发送本地 AS 内部网络的路由信息，但 BGP 不是自己去发现 AS 内部的路由信息，而是将 IGP 路由协议的路由信息引入到 BGP 路由表中，并发布给对等体。在引入 IGP 路由协议的路由时，可以针对不同的路由协议来对路由信息进行过滤。

缺省情况下，BGP 引入 IGP 路由协议的路由时，不会引入该协议的缺省路由。用户可以通过配置，指定 BGP 引入 IGP 路由协议的路由时，允许将缺省路由引入到 BGP 路由表中。

通过引入方式发布的路由的 ORIGIN 属性为 Incomplete。

只能引入路由表中状态为 active 的路由，是否为 active 状态可以通过 **display ip routing-table protocol** 命令或 **display ipv6 routing-table protocol** 命令来查看。这两条命令的详细介绍，请参见“三层技术-IP 路由命令参考”中的“IP 路由基础”。

表1-29 配置 BGP 引入 IGP 路由协议的路由（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 将IGP路由协议的路由信息引入到BGP路由表中 | | **import-route** *protocol* [ { *process-id* \| **all-processes** } [ **allow-direct** \| **med** *med-value* \| **route-policy** *route-policy-name* ] * ] | 缺省情况下，BGP不会引入IGP路由协议的路由信息 |
| （可选）允许将缺省路由引入到BGP路由表中 | | **default-route imported** | 缺省情况下，BGP不允许将缺省路由引入到BGP路由表中 |

表1-30 配置 BGP 引入 IGP 路由协议的路由（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 将IGP路由协议的路由信息引入到IPv6 BGP路由表中 | | **import-route** *protocol* [ { *process-id* \| **all-processes** } [ **allow-direct** \| **med** *med-value* \| **route-policy** *route-policy-name* ] * ] | 缺省情况下，BGP不会引入IGP路由协议的路由信息 |
| （可选）允许将缺省路由引入到IPv6 BGP路由表中 | | **default-route imported** | 缺省情况下，BGP不允许将缺省路由引入到IPv6 BGP路由表中 |

# 1.5 控制BGP路由信息的发布与接收

## 1.5.1 配置BGP路由聚合

在中型或大型 BGP 网络中，在向对等体发布路由信息时，可以配置路由聚合，减少发布的路由数量，并减小路由表的规模。IPv4 BGP 支持自动聚合和手动聚合两种聚合方式，同时配置时，手动聚合的优先级高于自动聚合的优先级。IPv6 BGP 只支持手动聚合。

提示

BGP 路由表中创建的聚合路由的出接口为 Null0 接口，聚合后可以减少向 BGP 对等体发布的路由数目。在使用中应注意不要使这条聚合路由成为本设备的优选路由，否则会导致报文转发失败。如果聚合路由的子网掩码长度和被聚合的某一条具体路由完全相同，且聚合路由优先级高于具体路由，则聚合路由会成为优选路由，这种情况下需要通过修改路由优先级等方式，来确保优选的路由为具体路由。

### 1. 配置路由自动聚合

配置自动聚合功能后，BGP 将对通过 **import-route** 命令引入的 IGP 子网路由进行聚合，不再发布子网路由，而是发布聚合的自然网段的路由。

**表1-31 配置路由自动聚合（IPv4 单播/IPv4 组播）**

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置对引入的子网路由进行自动聚合 | | **summary automatic** | 缺省情况下，不对引入的子网路由进行自动聚合 |

### 2. 配置路由手动聚合

自动聚合是按照自然网段进行聚合，而且只能对 IGP 引入的子网路由进行聚合。

通过配置手动聚合，用户可以同时对从 IGP 路由协议引入的子网路由和用 **network** 命令发布的路由进行聚合，而且还可以根据需要定义聚合路由的子网掩码长度。

表1-32 配置路由手动聚合（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 在BGP路由表中创建一条聚合路由 | | **aggregate** *ip-address* { *mask* \| *mask-length* } [ **as-set** \| **attribute-policy** *route-policy-name* \| **detail-suppressed** \| **origin-policy** *route-policy-name* \| **suppress-policy** *route-policy-name* ] * | 缺省情况下，不会进行路由聚合 |

表1-33 配置路由手动聚合（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 在IPv6 BGP路由表中创建一条聚合路由 | | **aggregate** *ipv6-address prefix-length* [ **as-set** \| **attribute-policy** *route-policy-name* \| **detail-suppressed** \| **origin-policy** *route-policy-name* \| **suppress-policy** *route-policy-name* ] * | 缺省情况下，不会进行路由聚合 |

## 1.5.2 配置发布IP路由表中的最优路由

缺省情况下，BGP 发布 BGP 路由表中的最优路由，不管该路由在 IP 路由表中是否为最优路由。通过本配置可以保证 BGP 发送出去的路由是 IP 路由表中的最优路由，以减少 BGP 发送的路由数量。

表1-34 配置向所有对等体发布 IP 路由表中的最优路由

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 配置向所有对等体发布IP路由表中的最优路由 | **advertise-rib-active** | 缺省情况下，BGP向所有对等体发布BGP路由表中的最优路由 |

表1-35 配置向对等体/对等体组发布 IP 路由表中的最优路由（IPv4）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 进入BGP IPv4单播地址族视图或BGP-VPN IPv4单播地址族视图 | | **address-family ipv4** [ **unicast** ] | - |
| 向对等体/对等体组发布IP路由表中的最优路由 | | **advertise-rib-active** | 缺省情况下，BGP向对等体/对等体组发布BGP路由表中的最优路由 |

表1-36 配置向对等体/对等体组发布 IPv6 路由表中的最优路由（IPv6）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 进入BGP IPv6单播地址族视图或BGP-VPN IPv6单播地址族视图 | | **address-family ipv6** [ **unicast** ] | - |
| 向对等体/对等体组发布IP路由表中的最优路由 | | **advertise-rib-active** | 缺省情况下，BGP向对等体/对等体组发布BGP路由表中的最优路由 |

### 1.5.3 配置向对等体/对等体组发送缺省路由

执行本配置后，设备将向指定对等体/对等体组发布一条下一跳地址为本地地址的缺省路由。

表1-37 配置向对等体/对等体组发送缺省路由（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播 | **bgp** *as-number* | |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| | 地址族视图 | **address-family ipv4 multicast** | |
| 向对等体/对等体组发送缺省路由 | | **peer** { *group-name* \| *ip-address* } **default-route-advertise** [ **route-policy** *route-policy-name* ] | 缺省情况下，不向对等体/对等体组发送缺省路由 |

表1-38 配置向对等体/对等体组发送缺省路由（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 向对等体/对等体组发送缺省路由 | | **peer** { *group-name* \| *ipv6-address* } **default-route-advertise** [ **route-policy** *route-policy-name* ] | 缺省情况下，不向对等体/对等体组发送缺省路由 |

## 1.5.4 限制从BGP对等体/对等体组接收的路由数量

通过本配置可以避免攻击者向路由器发送大量的 BGP 路由，对路由器进行攻击。

当路由器从指定对等体/对等体组接收的路由数量超过指定的最大值时，可以选择以下处理方式：

- 路由器中断与该对等体/对等体组的 BGP 会话，不再尝试重建会话。
- 路由器保持与该对等体/对等体组的 BGP 会话，可以继续接收路由，仅打印日志信息。
- 路由器保持与该对等体/对等体组的 BGP 会话，丢弃超出限制的路由，并打印日志信息。
- 路由器中断与该对等体/对等体组的 BGP 会话，经过指定的时间后自动与对等体/对等体组重建会话。

执行本配置任务时，还可以指定路由器产生日志信息的阈值，即路由器接收的路由数量与配置的最大值的百分比达到指定的阈值时，路由器将产生日志信息。

表1-39 限制从 BGP 对等体/对等体组接收的路由数量（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播 | **bgp** *as-number* | |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| | 地址族视图 | **address-family ipv4 multicast** | |
| 配置允许从对等体/对等体组接收的路由的最大数量 | | **peer** { *group-name* \| *ip-address* } **route-limit** *prefix-number* [ { **alert-only** \| **discard** \| **reconnect** *reconnect-time* } \| *percentage-value* ] * | 缺省情况下，不限制从对等体/对等体组接收的路由数量 |

表1-40 限制从 BGP 对等体/对等体组接收的路由数量（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置允许从对等体/对等体组接收的路由的最大数量 | | **peer** { *group-name* \| *ipv6-address* } **route-limit** *prefix-number* [ { **alert-only** \| **discard** \| **reconnect** *reconnect-time* } \| *percentage-value* ] * | 缺省情况下，不限制从对等体/对等体组接收的路由数量 |

## 1.5.5 配置BGP路由信息的发布/接收策略

### 1. 配置准备

配置 BGP 路由信息的发布/接收策略前，根据采取的策略，需要配置下列过滤器：

- 访问控制列表，详细配置过程请参见“ACL 和 QoS 配置指导”中的“ACL”。
- 地址前缀列表，详细配置过程请参见“三层技术-IP 路由配置指导”中的“路由策略”。
- 路由策略，详细配置过程请参见“三层技术-IP 路由配置指导”中的“路由策略”。
- AS 路径过滤列表，详细配置过程请参见“三层技术-IP 路由配置指导”中的“路由策略”。

### 2. 配置BGP路由信息的发布策略

可以通过以下几种方式配置 BGP 路由信息的发布策略：

- 使用访问控制列表或地址前缀列表对向所有对等体发布的路由信息进行过滤。
- 向指定对等体或对等体组发布路由时，使用路由策略、访问控制列表、AS 路径过滤列表或地址前缀列表对发布给该对等体或对等体组的路由信息进行过滤。

用户可以根据需求选择过滤策略。如果同时配置了几种过滤策略，则按照如下顺序过滤发布的路由信息：

- **filter-policy export**
- **peer filter-policy export**
- **peer as-path-acl export**
- **peer prefix-list export**
- **peer route-policy export**

只有通过前面的过滤策略，才能继续执行后面的过滤策略；只有通过所有配置的过滤策略后，路由信息才能被发布。

表1-41 配置 BGP 路由信息的发布策略（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 对向所有对等体发布的路由信息进行过滤 | | **filter-policy** { *acl-number* \| **prefix-list** *prefix-list-name* } **export** [ **direct** \| **isis** *process-id* \| **ospf** *process-id* \| **rip** *process-id* \| **static** ] | 至少选其一<br>缺省情况下，不对发布的路由信息进行过滤 |
| 为对等体/对等体组设置基于路由策略的路由发布过滤策略 | | **peer** { *group-name* \| *ip-address* } **route-policy** *route-policy-name* **export** | |
| 为对等体/对等体组设置基于ACL的路由发布过滤策略 | | **peer** { *group-name* \| *ip-address* } **filter-policy** *acl-number* **export** | |
| 为对等体/对等体组设置基于AS路径过滤列表的路由发布过滤策略 | | **peer** { *group-name* \| *ip-address* } **as-path-acl** *as-path-acl-number* **export** | |
| 为对等体/对等体组设置基于IPv4地址前缀列表的路由发布过滤策略 | | **peer** { *group-name* \| *ip-address* } **prefix-list** *prefix-list-name* **export** | |

表1-42 配置 BGP 路由信息的发布策略（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 对向所有IPv6 BGP对等体发布的路由信息进行过滤 | **filter-policy** { *acl6-number* \| **prefix-list** *ipv6-prefix-name* } **export** [ **direct** \| **isisv6** *process-id* \| **ospfv3** *process-id* \| **ripng** *process-id* \| **static** ] | 至少选其一<br>缺省情况下，不对发布的路由信息进行过滤<br>BGP-VPN IPv6单播地址族视图下不支持**peer as-path-acl**命令 |
| 为对等体/对等体组设置基于路由策略的路由发布过滤策略 | **peer** { *group-name* \| *ipv6-address* } **route-policy** *route-policy-name* **export** | |
| 为对等体/对等体组设置基于ACL的路由发布过滤策略 | **peer** { *group-name* \| *ipv6-address* } **filter-policy** *acl6-number* **export** | |
| 为对等体/对等体组设置基于AS路径过滤列表的路由发布过滤策略 | **peer** { *group-name* \| *ipv6-address* } **as-path-acl** *as-path-acl-number* **export** | |
| 为对等体/对等体组设置基于IPv6地址前缀列表的路由发布过滤策略 | **peer** { *group-name* \| *ipv6-address* } **prefix-list** *ipv6-prefix-name* **export** | |

### 3. 配置BGP路由信息的接收策略

可以通过以下几种方式配置 BGP 路由信息的接收策略：

- 使用访问控制列表或地址前缀列表对从所有对等体接收的路由信息进行过滤。
- 从指定对等体或对等体组接收路由时，使用路由策略、访问控制列表、AS 路径过滤列表或地址前缀列表对从该对等体或对等体组接收的路由信息进行过滤。

用户可以根据需求选择过滤策略。如果同时配置了几种过滤策略，则按照如下顺序过滤接收的路由：

- **filter-policy import**
- **peer filter-policy import**
- **peer as-path-acl import**
- **peer prefix-list import**
- **peer route-policy import**

只有通过前面的过滤策略，才能继续执行后面的过滤策略；只有通过所有配置的过滤策略后，路由信息才能被接收。

表1-43 配置 BGP 路由信息的接收策略（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 对从所有对等体接收的路由信息进行过滤 | **filter-policy** { *acl-number* \| **prefix-list** *prefix-list-name* } **import** | 至少选其一<br>缺省情况下，不对接收的路由信息进行过滤 |
| 为对等体/对等体组设置基于路由策略的路由接收过滤策略 | **peer** { *group-name* \| *ip-address* } **route-policy** *route-policy-name* **import** | |
| 为对等体/对等体组设置基于ACL的路由接收过滤策略 | **peer** { *group-name* \| *ip-address* } **filter-policy** *acl-number* **import** | |
| 为对等体/对等体组设置基于AS路径过滤列表的路由接收过滤策略 | **peer** { *group-name* \| *ip-address* } **as-path-acl** *as-path-acl-number* **import** | |
| 为对等体/对等体组设置基于IPv4地址前缀列表的路由接收过滤策略 | **peer** { *group-name* \| *ip-address* } **prefix-list** *prefix-list-name* **import** | |

表1-44 配置 BGP 路由信息的接收策略（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 对从所有IPv6 BGP对等体接收的路由信息进行过滤 | | **filter-policy** { *acl6-number* \| **prefix-list** *ipv6-prefix-name* } **import** | 至少选其一<br>缺省情况下，不对接收的路由信息进行过滤<br>BGP-VPN IPv6单播地址族视图下不支持**peer as-path-acl**命令 |
| 为对等体/对等体组设置基于路由策略的路由接收过滤策略 | | **peer** { *group-name* \| *ipv6-address* } **route-policy** *route-policy-name* **import** | |
| 为对等体/对等体组设置基于ACL的路由接收过滤策略 | | **peer** { *group-name* \| *ipv6-address* } **filter-policy** *acl6-number* **import** | |
| 为对等体/对等体组设置基于AS路径过滤列表的路由接收过滤策略 | | **peer** { *group-name* \| *ipv6-address* } **as-path-acl** *as-path-acl-number* **import** | |
| 为对等体/对等体组设置基于IPv6地址前缀列表的路由接收过滤策略 | | **peer** { *group-name* \| *ipv6-address* } **prefix-list** *ipv6-prefix-name* **import** | |

### 1.5.6 配置BGP路由衰减

通过配置 BGP 路由衰减，可以抑制不稳定的路由信息，不允许这类路由参与路由优选。

本配置只对 EBGP 路由生效，对 IBGP 路由无效。

表1-45 配置 BGP 路由衰减（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置BGP路由衰减 | | **dampening** [ *half-life-reachable half-life-unreachable reuse suppress ceiling* \| **route-policy** *route-policy-name* ] * | 缺省情况下，没有配置BGP路由衰减 |

表1-46 配置 BGP 路由衰减（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置IPv6 BGP路由衰减 | | **dampening** [ *half-life-reachable half-life-unreachable reuse suppress ceiling* \| **route-policy** *route-policy-name* ] * | 缺省情况下，没有配置IPv6 BGP路由衰减 |

## 1.6 控制BGP路径的选择

BGP 具有很多路由属性，通过配置这些属性可以控制 BGP 路径的选择。

### 1.6.1 为接收路由分配首选值

BGP 选择路由时首先丢弃下一跳不可达的路由，其次优选 Preferred-value 值最大的路由。通过本配置，可以修改路由的 Preferred-value，以便控制 BGP 路径的选择。

缺省情况下，从对等体/对等体组学到的路由的首选值为 0，网络管理员可以为从某个对等体/对等体组接收的路由配置首选值，从而提高从指定对等体/对等体学到的路由的优先级。

表1-47 为接收路由分配首选值（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 为从对等体/对等体组接收的路由分配首选值 | | **peer** { *group-name* \| *ip-address* }<br>**preferred-value** *value* | 缺省情况下，从对等体/对等体组接收的路由的首选值为0 |

表1-48 为接收路由分配首选值（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 为从IPv6 BGP对等体/对等体组接收的路由分配首选值 | | **peer** { *group-name* \| *ipv6-address* }<br>**preferred-value** *value* | 缺省情况下，从IPv6 BGP对等体/对等体组接收的路由的首选值为0 |

### 1.6.2 配置BGP的路由优先级

路由器上可能同时运行多个动态路由协议，存在各个路由协议之间路由信息共享和选择的问题。系统为每一种路由协议设置一个优先级，在不同协议发现同一条路由时，优先级高的路由将被优先选择。

用户可以通过 **preference** 命令修改 EBGP 路由、IBGP 路由以及本地产生的 BGP 路由的优先级，或应用路由策略为通过匹配规则过滤的特定路由配置优先级，没有通过过滤的路由使用缺省优先级。

缺省情况下，EBGP 路由的优先级低于本地产生的 BGP 路由的优先级。设备上存在到达某一目的网络的 EBGP 路由和本地产生的 BGP 路由时，不会选择 EBGP 路由。通过执行 **network short-cut** 命令将一条 EBGP 路由配置成 short-cut，可以使得指定 EBGP 路由的优先级与本地产生的 BGP 路由的优先级相同，从而提高该 EBGP 路由成为最佳路由的可能性。

表1-49 配置 BGP 的路由优先级（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置BGP路由的优先级 | | **preference** { *external-preference internal-preference local-preference* \| **route-policy** *route-policy-name* } | 缺省情况下，EBGP路由的优先级为255，IBGP路由的优先级为255，本地产生的BGP路由的优先级为130 |
| 提高接收到的指定EBGP路由的路由优先级 | | **network** *ip-address* [ *mask* \| *mask-length* ] **short-cut** | 缺省情况下，接收到的EBGP路由的路由优先级为255 |

表1-50 配置 BGP 的路由优先级（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置BGP路由的优先级 | **preference** { *external-preference internal-preference local-preference* \| **route-policy** *route-policy-name* } | 缺省情况下，EBGP路由的优先级为255，IBGP路由的优先级为255，本地产生的BGP路由的优先级为130 |
| 提高接收到的指定EBGP路由的路由优先级 | **network** *ipv6-address prefix-length* **short-cut** | 缺省情况下，接收到的EBGP路由的路由优先级为255 |

### 1.6.3 配置本地优先级的缺省值

本地优先级用来判断流量离开 AS 时的最佳路由。当 BGP 路由器通过不同的 IBGP 对等体得到目的地址相同但下一跳不同的多条路由时，将优先选择本地优先级较高的路由。

用户可以通过本配置改变 BGP 路由器向 IBGP 对等体发送的路由本地优先级的缺省值。

表1-51 配置本地优先级的缺省值（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置本地优先级的缺省值 | | **default local-preference** *value* | 缺省情况下，本地优先级的缺省值为100 |

表1-52 配置本地优先级的缺省值（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置本地优先级的缺省值 | **default local-preference** *value* | 缺省情况下，本地优先级的缺省值为100 |

### 1.6.4 配置MED属性

MED 用来判断流量进入 AS 时的最佳路由。当一个 BGP 路由器通过不同的 EBGP 对等体得到目的地址相同但下一跳不同的多条路由时，在其它条件相同的情况下，将优先选择 MED 值较小者作为最佳路由。

#### 1. 配置MED缺省值

表1-53 配置 MED 缺省值（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置MED的缺省值 | | **default med** *med-value* | 缺省情况下，MED的缺省值为0 |

表1-54 配置 MED 缺省值（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置MED的缺省值 | | **default med** *med-value* | 缺省情况下，MED的缺省值为0 |

### 2. 配置允许比较来自不同AS路由的MED属性值

缺省情况下，BGP 只比较来自同一个 AS 的路由的 MED 属性值。

通过配置 **compare-different-as-med** 命令，可以强制 BGP 比较来自不同 AS 的路由的 MED 属性值。

表1-55 配置允许比较来自不同 AS 路由的 MED 属性值（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置允许比较来自不同AS路由的MED属性值 | | **compare-different-as-med** | 缺省情况下，不允许比较来自不同AS路由的MED属性值 |

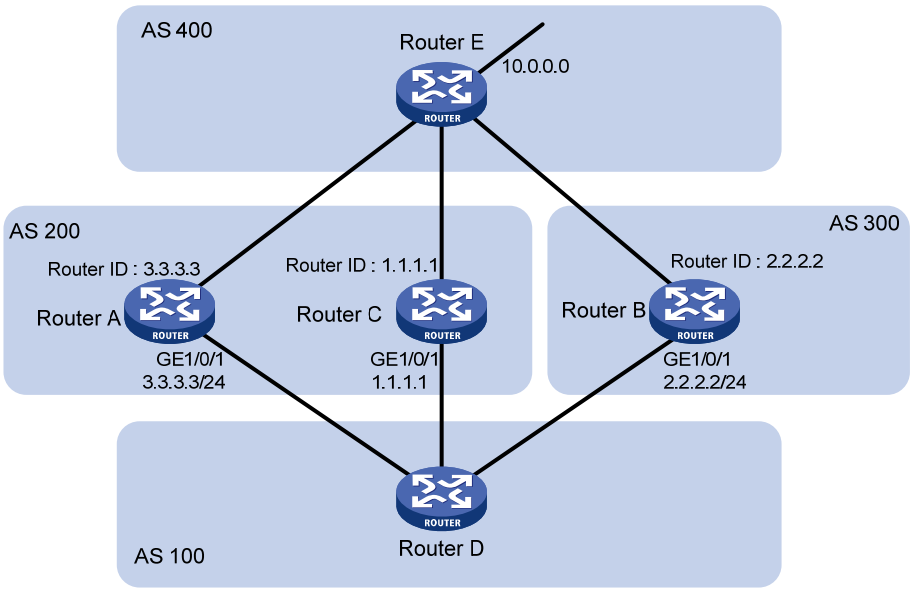
表1-56 配置允许比较来自不同 AS 路由的 MED 属性值（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置允许比较来自不同AS路由的MED属性值 | | **compare-different-as-med** | 缺省情况下，不允许比较来自不同AS路由的MED属性值 |

### 3. 配置对来自同一AS的路由进行MED排序优选

缺省情况下，BGP 选择最优路由时是将新的路由和当前 BGP 路由表中的最优路由进行比较，只要新的路由比当前 BGP 路由表中的最优路由更优，新的路由将成为最优路由，路由学习的顺序有可能会影响最优路由的选择结果。

图1-10 MED 排序优选示意图（以 IPv4 为例）




如上图所示，Router D 分别从 Router A 和 Router B 学习到了到达网段 10.0.0.0 的路由，由于 Router B 的 Router ID 值较小，因此，从 Router B 学来的路由被选为最优路由：

```
     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn
 *>i  10.0.0.0           2.2.2.2         50                    0       300e
 * i                     3.3.3.3         50                    0       200e
```

当 Router D 再从 Router C 学习到到达 10.0.0.0 网段的路由时，它只和当前路由表的最优路由进行比较。由于 Router C 和 Router B 位于不同的 AS，选择路由时不会比较 MED 值，而 Router C 的 Router ID 值更小，相对更优，它将成为最优路由。

```
     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn
 *>i  10.0.0.0           1.1.1.1         60                    0       200e
 * i  10.0.0.0           2.2.2.2         50                    0       300e
 * i                     3.3.3.3         50                    0       200e
```

但是如果将这条路由与从 Router A 学习到的路由进行比较，那么由于两条路由来自同一个 AS，且从 Router C 学习到的路由 MED 值更大，则从 Router C 学习到的路由应该视为无效路由。

在 Router D 上配置 **bestroute compare-med** 命令后，Router D 学习到新的路由时，会首先按照路由来自的 AS 分组，对来自同一 AS 的路由根据 MED 值的大小进行优选，选出 MED 值最小的路由，然后再对优选出来的、来自不同 AS 的路由进行优选，从而避免路由优选结果的不确定性。配置对来自同一 AS 的路由进行 MED 排序优选后，Router D 上的 BGP 路由表如下所示，从 Router B 学习到的到达 10.0.0.0 网段的路由将成为最优路由。

```
     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn
 *>i  10.0.0.0           2.2.2.2         50                    0       300e
 * i                     3.3.3.3         50                    0       200e
 * i                     1.1.1.1         60                    0       200e
```

表1-57 配置对来自同一 AS 的路由进行 MED 排序优选（IPv4 单播/IPv4 组播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置对来自同一AS的路由进行MED排序优选 | | **bestroute compare-med** | 缺省情况下，不会对来自同一AS的路由进行MED排序优选 |

表1-58 配置对来自同一 AS 的路由进行 MED 排序优选（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置对来自同一AS的路由进行MED排序优选 | | **bestroute compare-med** | 缺省情况下，不会对来自同一AS的路由进行MED排序优选 |

### 4. 配置允许比较来自同一联盟不同子自治系统邻居路由的MED属性值

表1-59 配置允许比较来自同一联盟不同子自治系统邻居路由的 MED 属性值（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置允许比较来自同一联盟不同子自治系统邻居路由的MED属性值 | | **bestroute med-confederation** | 缺省情况下，不比较来自同一联盟不同子自治系统邻居路由的MED属性值 |

表1-60 配置允许比较来自同一联盟不同子自治系统邻居路由的 MED 属性值（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置允许比较来自同一联盟不同子自治系统邻居路由的MED属性值 | | **bestroute med-confederation** | 缺省情况下，不比较来自同一联盟不同子自治系统邻居路由的MED属性值 |

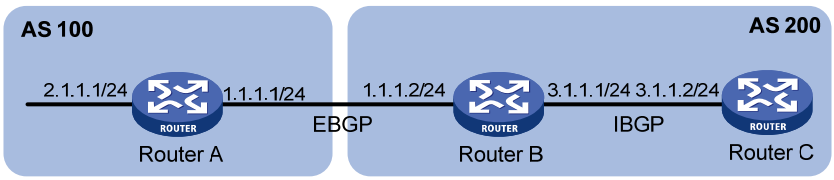
说明

只有 AS_PATH 里不包含联盟体外的自治系统编号时，才会比较来自同一联盟不同子自治系统邻居路由的 MED 属性值。例如，联盟中包含的子自治系统为 65006、65007 和 65009。如果存在三条路由，它们的 AS-PATH 值分别为 65006 65009、65007 65009 和 65008 65009，MED 值分别为 2、3、1，由于第三条路由包含了联盟体外的自治系统编号，因此在选择最优路由时第一条路由将成为最优路由。

### 1.6.5 配置NEXT_HOP属性

缺省情况下，路由器向 IBGP 对等体/对等体组发布路由时，不将自身地址作为下一跳，但有的时候为了保证 IBGP 邻居能够找到下一跳，可以配置将自身地址作为下一跳。以下图为例，Router A 与 Router B 建立 EBGP 邻居关系，Router B 与 Router C 建立 IBGP 邻居关系，Router B 在向 Router C 发布从 Router A 学到的 BGP 路由时，如果 Router C 上没有到达 1.1.1.1/24 的路由，可以在 Router B 上配置 **peer next-hop-local** 命令将 3.1.1.1/24 作为下一跳，这样，Router C 就能找到下一跳。

图1-11 配置 BGP NEXT_HOP 属性应用组网图一




在一些比较特殊的组网环境中（即两个 BGP 连接在同一网段的广播网），路由器向 EBGP 对等体/对等体组发布路由时不会将自身地址作为下一跳，以下图为例：Router A 与 Router B 建立 EBGP 邻居关系，Router B 与 Router C 建立 IBGP 邻居关系，两个 BGP 连接都位于同一个广播网 1.1.1.0/24 中，Router B 向 Router A 发布 EBGP 路由时不会将自身地址 1.1.1.2/24 作为下一跳，但如果用户有需要，也可以通过配置 **peer next-hop-local** 命令实现将自身地址 1.1.1.2/24 作为下一跳。

图1-12 配置 BGP NEXT_HOP 属性应用组网图二

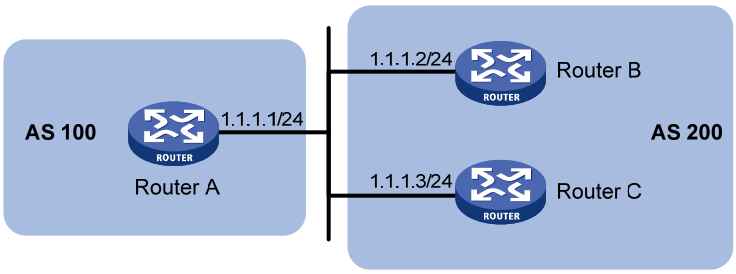



表1-61 配置 BGP 的 NEXT_HOP 属性（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置向对等体/对等体组发布路由时，将下一跳属性修改为自身的地址 | | **peer** { *group-name* \| *ip-address* } **next-hop-local** | 缺省情况下，向EBGP对等体/对等体组发布路由时，将下一跳属性修改为自身的地址；向IBGP对等体/对等体组发布路由时，不修改下一跳属性 |

表1-62 配置 BGP 的 NEXT_HOP 属性（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置向对等体/对等体组发布路由时，将下一跳属性修改为自身的地址 | | **peer** { *group-name* \| *ipv6-address* } **next-hop-local** | 缺省情况下，向EBGP对等体/对等体组发布路由时，将下一跳属性修改为自身的地址；向IBGP对等体/对等体组发布路由时，不修改下一跳属性 |

说明

如果配置了 BGP 负载分担，则不论是否配置了 **peer next-hop-local** 命令，本地路由器向 IBGP 对等体/对等体组发布路由时都先将下一跳地址改变为自身地址。

## 1.6.6 配置AS_PATH属性

### 1. 配置允许本地AS号出现的次数

通常情况下，BGP 会检查对等体发来的路由的 AS_PATH 属性，如果其中已存在本地 AS 号，则 BGP 会忽略此路由，以免形成路由环路。

但是，在某些特殊的组网环境下（如 MPLS L3VPN 的 Hub&Spoke 组网），需要允许本地 AS 号在接收路由的 AS_PATH 属性中出现，否则无法正确发布路由。通过本配置，可以允许本地 AS 号在所接收的路由的 AS_PATH 属性中出现，并可同时配置允许出现的次数。

表1-63 配置允许本地 AS 号出现的次数（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置对于从对等体/对等体组接收的路由，允许本地AS号在接收路由的AS_PATH属性中出现，并配置允许出现的次数 | | **peer** { *group-name* \| *ip-address* }<br>**allow-as-loop** [ *number* ] | 缺省情况下，不允许本地AS号在接收路由的AS_PATH属性中出现 |

表1-64 配置允许本地 AS 号出现的次数（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组 | **bgp** *as-number* | |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| | 播地址族视图 | **address-family ipv6 multicast** | |
| 配置对于从对等体/对等体组接收的路由，允许本地AS号在接收路由的AS_PATH属性中出现，并配置允许出现的次数 | | **peer** { *group-name* \| *ipv6-address* }<br>**allow-as-loop** [ *number* ] | 缺省情况下，不允许本地AS号在接收路由的AS_PATH属性中出现 |

### 2. 禁止路由器将AS_PATH当作选路算法中的一个因素

路由器在选择最优路由时会优选 AS 路径最短的路由，通过如下配置可以禁止路由器将 AS_PATH 当作选路算法中的一个因素。

表1-65 禁止路由器将 AS_PATH 当作选路算法中的一个因素（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 禁止路由器将AS_PATH当作选路算法中的一个因素 | | **bestroute as-path-neglect** | 缺省情况下，路由器将AS_PATH当作选路算法中的一个因素 |

表1-66 禁止路由器将 AS_PATH 当作选路算法中的一个因素（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 禁止路由器将AS_PATH当作选路算法中的一个因素 | | **bestroute as-path-neglect** | 缺省情况下，路由器将AS_PATH当作选路算法中的一个因素 |

### 3. 为对等体/对等体组指定一个虚拟的自治系统号

进行系统移植时，例如，Router A 原来位于 AS 2，现在将它移植到 AS 3 里，网络管理员需要在 Router A 的所有 EBGP 对等体上修改 Router A 所在的 AS 号。通过在 Router A 上为 EBGP 对等体/对等体组配置一个虚拟的本地自治系统号 2，可以将本地真实的 AS 号 3 隐藏起来。在 EBGP 对等体看来 Router A 始终位于 AS 2，不需要改变 EBGP 对等体上的配置。

表1-67 为对等体/对等体组指定一个虚拟的自治系统号（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 为对等体/对等体组指定一个虚拟的本地自治系统号 | | **peer** { *group-name* \| *ip-address* } **fake-as** *as-number* | 缺省情况下，没有为对等体/对等体组配置虚拟的本地自治系统号<br>本命令只适用于EBGP对等体和对等体组 |

表1-68 为对等体/对等体组指定一个虚拟的自治系统号（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 为对等体/对等体组指定一个虚拟的本地自治系统号 | | **peer** { *group-name* \| *ipv6-address* } **fake-as** *as-number* | 缺省情况下，没有为对等体/对等体组配置虚拟的本地自治系统号<br>本命令只适用于EBGP对等体和对等体组 |

### 4. 配置AS号替换功能

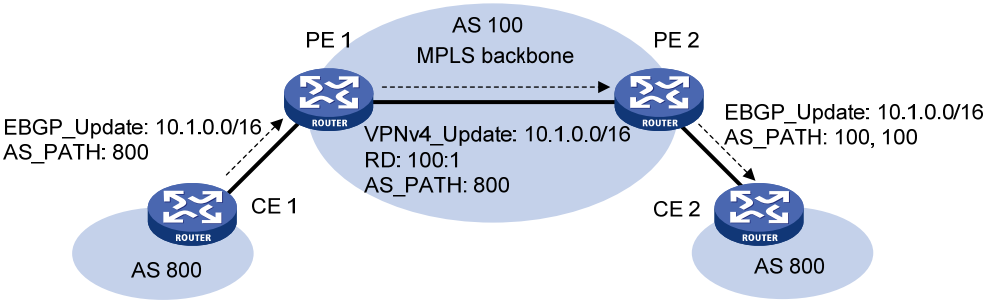
注意

本配置仅用于特定的组网环境。通常情况下，建议不要使用本配置，否则可能会引起路由环路。

在 MPLS L3VPN 中，如果 PE 和 CE 之间运行 EBGP，由于 BGP 使用 AS 号检测路由环路，为保证路由信息的正确发送，需要为物理位置不同的站点分配不同的 AS 号。

如果物理位置不同的 CE 复用相同的 AS 号，则需要在 PE 上配置 BGP 的 AS 号替换功能。当 PE 向指定对等体（CE）发布路由时，如果路由的 AS_PATH 中存在 CE 所在的 AS 号，则 PE 将该 AS 号替换成 PE 的 AS 号后，再发布该路由，以保证私网路由能够正确发布。

图1-13 BGP AS 号替换应用示意图（以 IPv4 为例）




如 图 1-13 所示，CE 1 和CE 2 都使用AS号 800，在PE 2 上使能针对CE 2 的AS号替换功能。当CE 1 发来的Update信息从PE 2 发布给CE 2 时，PE 2 发现AS_PATH中存在与CE 2 相同的AS号 800，就把它替换为自己的AS号 100。如果需要完全的连接性，PE 1 上也需要做类似的配置。

表1-69 配置 AS 号替换功能（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置用本地AS号替换AS_PATH属性中指定对等体/对等体组的AS号 | | **peer** { *group-name* \| *ip-address* } **substitute-as** | 缺省情况下，不会用本地AS号替换AS_PATH属性中指定对等体/对等体组的AS号 |

表1-70 配置 AS 号替换功能（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置用本地AS号替换AS_PATH属性中指定对等体/对等体组的AS号 | | **peer** { *group-name* \| *ipv6-address* } **substitute-as** | 缺省情况下，不会用本地AS号替换AS_PATH属性中指定对等体/对等体组的AS号 |

### 5. 配置发送BGP更新消息时AS_PATH属性中不携带私有AS号

私有 AS 号是内部使用的 AS 号，范围为 64512～65535。私有 AS 号主要用于测试网络，一般情况下不需要在公共网络中传播。

通过本配置，可以指定如果向 EBGP 对等体/对等体组发送的 BGP 更新消息中 AS_PATH 属性只包括私有 AS 号，则删除私有 AS 号后，将 BGP 更新消息发送给对等体/对等体组。

表1-71 配置发送 BGP 更新消息时 AS_PATH 属性中不携带私有 AS 号（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置向指定EBGP对等体/对等体组发送BGP更新消息时只携带公有AS号，不携带私有AS号 | | **peer** { *group-name* \| *ip-address* } **public-as-only** | 缺省情况下，向EBGP对等体/对等体组发送BGP更新消息时，既可以携带公有AS号，又可以携带私有AS号<br>本命令只适用于EBGP对等体和对等体组 |

表1-72 配置发送 BGP 更新消息时 AS_PATH 属性中不携带私有 AS 号（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置向指定EBGP对等体/对等体组发送BGP更新消息时只携带公有AS号，不携带私有AS号 | | **peer** { *group-name* \| *ipv6-address* } **public-as-only** | 缺省情况下，向EBGP对等体/对等体组发送BGP更新消息时，既可以携带公有AS号，又可以携带私有AS号<br>本命令只适用于EBGP对等体和对等体组 |

### 6. 配置不检测EBGP路由的第一个AS号

缺省情况下，BGP 会检查对等体发来的路由的 AS_PATH 属性。如果第一个 AS 号不是 BGP 邻居的 AS 号，则丢弃此路由。

通过配置 **ignore-first-as** 命令，可以忽略对 EBGP 路由第一个 AS 号的检测。

表1-73 配置 EBGP 不检测路由的第一个 AS 号

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 配置不检测EBGP路由的第一个AS号 | **ignore-first-as** | 缺省情况下，系统收到EBGP路由后，会检测路由的第一个AS号 |

## 1.6.7 配置SoO属性

为 BGP 对等体/对等体组配置 SoO 属性后，从该 BGP 对等体/对等体组接收路由时设备会为路由增加 SoO 属性，并且向该 BGP 对等体/对等体组发布路由时设备会检查路由的 SoO 属性，如果路由中携带的 SoO 属性与为对等体/对等体组配置的 SoO 属性相同，则不会将该路由发布给对等体/对等体组，从而避免路由环路。

表1-74 配置 SoO 属性（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 为BGP对等体/对等体组配置SoO属性 | | **peer** { *group-name* \| *ip-address* } **soo** *site-of-origin* | 缺省情况下，没有为BGP对等体/对等体组配置SoO属性 |

表1-75 配置 SoO 属性（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播 | **bgp** *as-number* | |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| | 地址族视图 | **address-family ipv6 multicast** | |
| 为BGP对等体/对等体组配置SoO属性 | | **peer** { *group-name* \| *ipv6-address* } **soo** *site-of-origin* | 缺省情况下，没有为BGP对等体/对等体组配置SoO属性 |

# 1.7 调整和优化BGP网络

## 1.7.1 配置BGP会话的存活时间间隔与保持时间

当对等体间建立了 BGP 会话后，它们定时向对端发送 Keepalive 消息，以防止路由器认为 BGP 会话已中断。Keepalive 消息的发送时间间隔称为存活时间间隔。

若路由器在设定的会话保持时间（Holdtime）内未收到对端的 Keepalive 消息或 Update 消息，则认为此 BGP 会话已中断，从而断开此 BGP 会话。

用户可以全局配置当前路由器上所有 BGP 会话的存活时间间隔与保持时间，也可以配置与指定对等体/对等体组建立的 BGP 会话的存活时间间隔和保持时间。如果同时配置了两者，则为指定对等体/对等体组配置的值具有较高的优先级。

存活时间间隔、会话保持时间的协商及计算方法如下：

- 如果当前路由器上配置的保持时间与对端设备（对等体）上配置的保持时间不一致，则数值较小者作为协商后的保持时间。协商的保持时间为 0 时，不向对等体发送 Keepalive 消息，与对等体之间的会话永远不会超时断开。
- 存活时间间隔为 0，协商的保持时间不为 0 时，以协商的保持时间的三分之一作为存活时间间隔；存活时间间隔不为 0 时，将协商的保持时间的三分之一与配置的存活时间间隔比较，取最小值作为存活时间间隔。

表1-76 配置 BGP 会话的存活时间间隔与保持时间（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置所有BGP会话的存活时间间隔和保持时间 | | **timer keepalive** *keepalive* **hold** *holdtime* | 二者选其一<br>缺省情况下，BGP会话的存活时间间隔为60秒，保持时间为180秒 |
| 配置本地路由器与指定对等体/对等体组之间BGP会话的存活时间间隔和保持时间 | | **peer** { *group-name* \| *ip-address* } **timer keepalive** *keepalive* **hold** *holdtime* | 配置**timer**命令后，不会影响已建立的BGP会话，只对新建立的会话生效<br>配置**timer**和**peer timer**命令后，不会马上断开会话，而是等到其他条件触发会话重建（如复位BGP会话）时，再以配置的保持时间协商建立会话 |

表1-77 配置 BGP 会话的存活时间间隔与保持时间（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置所有BGP会话的存活时间间隔和保持时间 | | **timer keepalive** *keepalive* **hold** *holdtime* | 二者选其一<br>缺省情况下，BGP会话的存活时间间隔为60秒，保持时间为180秒 |
| 配置本地路由器与指定IPv6 BGP对等体/对等体组之间BGP会话的存活时间间隔和保持时间 | | **peer** { *group-name* \| *ipv6-address* } **timer keepalive** *keepalive* **hold** *holdtime* | 配置**timer**命令后，不会影响已建立的BGP会话，只对新建立的会话生效<br>配置**timer**和**peer timer**命令后，不会马上断开会话，而是等到其他条件触发会话重建（如复位BGP会话）时，再以配置的保持时间协商建立会话 |

## 1.7.2 配置发布同一路由的时间间隔

BGP 路由发生变化时，BGP 路由器会发送 Update 消息通知对等体。如果同一路由频繁变化，BGP 路由器会频繁发送 Update 消息更新路由，导致路由震荡。通过本配置指定向对等体/对等体组发布同一路由的时间间隔，可以避免每次路由变化都发送 Update 消息，避免路由震荡。

表1-78 配置发布同一路由的时间间隔（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置向指定对等体/对等体组发布同一路由的时间间隔 | | **peer** { *group-name* \| *ip-address* } **route-update-interval** *interval* | 缺省情况下，向IBGP对等体发布同一路由的时间间隔为15秒，向EBGP对等体发布同一路由的时间间隔为30秒 |

表1-79 配置发布同一路由的时间间隔（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置向指定IPv6 BGP对等体/对等体组发布同一路由的时间间隔 | **peer** { *group-name* \| *ipv6-address* } **route-update-interval** *interval* | 缺省情况下，向IBGP对等体发布同一路由的时间间隔为15秒，向EBGP对等体发布同一路由的时间间隔为30秒 |

### 1.7.3 配置允许同非直连邻居建立EBGP会话

提示

配置 BGP GTSM 功能后，只要本地设备和指定的对等体通过了 GTSM 检查，就允许在二者之间建立 EBGP 会话，不管二者之间的跳数是否超过 **peer ebgp-max-hop** 命令指定的跳数范围。

当前路由器要与另外一个路由器建立 EBGP 会话，它们之间必须具有直连的物理链路，如果不满足这一要求，则必须使用 **peer ebgp-max-hop** 命令允许它们经过多跳建立 EBGP 会话。

直连 EBGP 邻居使用 Loopback 接口建立 BGP 会话时，不需要配置 **peer ebgp-max-hop** 命令。

表1-80 配置允许同非直连邻居建立 EBGP 会话（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置允许本地路由器同非直连网络上的邻居建立EBGP会话，同时指定允许的最大跳数 | | **peer** { *group-name* \| *ip-address* } **ebgp-max-hop** [ *hop-count* ] | 缺省情况下，不允许同非直连网络上的邻居建立EBGP会话 |

表1-81 配置非直接相连的邻居建立 EBGP 连接（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置允许本地路由器同非直连网络上的邻居建立EBGP会话，同时指定允许的最大跳数 | | **peer** { *group-name* \| *ipv6-address* } **ebgp-max-hop** [ *hop-count* ] | 缺省情况下，不允许同非直连网络上的邻居建立EBGP会话 |

### 1.7.4 使能直连EBGP会话快速复位功能

如果没有使能本功能，则连接直连 EBGP 对等体的链路 down 后，本地路由器不会立即断开与 EBGP 对等体的会话，而是等待会话保持时间（Holdtime）超时后，才断开该会话。没有使能本功能时，链路震荡不会影响 EBGP 会话的状态。

如果使能了本功能，则连接直连 EBGP 对等体的链路 down 后，本地路由器会立即断开与 EBGP 对等体的会话，并重新与该对等体建立 EBGP 会话，从而实现快速发现链路故障，快速重建会话。

**表1-82 使能直连 EBGP 会话快速复位功能**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 使能直连EBGP会话快速复位功能 | **ebgp-interface-sensitive** | 缺省情况下，直连EBGP会话快速复位功能处于使能状态 |

### 1.7.5 使能 4 字节AS号抑制功能

**提示**

如果对端设备支持 4 字节 AS 号，请不要使能 4 字节 AS 号抑制功能，否则会导致 BGP 会话无法建立。

设备支持 4 字节的 AS 号，即 AS 号取值占用 4 字节，取值范围为 1～4294967295。缺省情况下，设备在与对端设备建立 BGP 会话时，通过 Open 消息通告对端设备本端支持 4 字节的 AS 号。如果对端设备不支持 4 字节 AS 号（只支持 2 字节 AS 号），则会导致会话协商失败。此时，在本端与对端设备之间使能 4 字节 AS 号抑制功能，可以使得本端设备通过 Open 消息向对端设备谎称自己不支持 4 字节的 AS 号，从而确保本端和对端设备之间可以成功建立 BGP 会话。

**表1-83 使能 4 字节 AS 号抑制功能（IPv4 单播/IPv4 组播）**

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 使能4字节AS号抑制功能 | | **peer** { *group-name* \| *ip-address* } **capability-advertise suppress-4-byte-as** | 缺省情况下，设备没有使能4字节AS号抑制功能 |

表1-84 使能 4 字节 AS 号抑制功能（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 使能4字节AS号抑制功能 | | **peer** { *group-name* \| *ipv6-address* } **capability-advertise suppress-4-byte-as** | 缺省情况下，设备没有使能4字节AS号抑制功能 |

## 1.7.6 配置BGP的MD5 认证

通过为 BGP 对等体配置 BGP 的 MD5 认证，可以在以下两方面提高 BGP 的安全性：

- 为 BGP 建立 TCP 连接时进行 MD5 认证，只有两台路由器配置的密钥相同时，才能建立 TCP 连接，从而避免与非法的 BGP 路由器建立 TCP 连接。
- 传递 BGP 报文时，对封装 BGP 报文的 TCP 报文段进行 MD5 运算，从而保证 BGP 报文不会被篡改。

表1-85 配置 BGP 的 MD5 认证（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置BGP的MD5认证 | | **peer** { *group-name* \| *ip-address* } **password** { **cipher** \| **simple** } *password* | 缺省情况下，不进行BGP的MD5认证 |

表1-86 配置 BGP 的 MD5 认证（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置BGP的MD5认证 | | **peer** { *group-name* \| *ipv6-address* } **password** { **cipher** \| **simple** } *password* | 缺省情况下，不进行BGP的MD5认证 |

### 1.7.7 配置BGP负载分担

通过改变 BGP 选路规则实现负载分担时，设备根据 **balance** 命令配置的进行 BGP 负载分担的路由条数，选择指定数目的路由进行负载分担，以提高链路利用率。

表1-87 配置 BGP 负载分担（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置进行BGP负载分担的路由条数 | | **balance** [ **ebgp** \| **eibgp** \| **ibgp** ] *number* | 缺省情况下，不会进行BGP负载分担 |

表1-88 配置 BGP 负载分担（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图、BGP-VPN IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP-VPN IPv6单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置进行BGP负载分担的路由条数 | | **balance** [ **ebgp** \| **eibgp** \| **ibgp** ] *number* | 缺省情况下，不会进行BGP负载分担 |

### 1.7.8 配置通过IPsec保护IPv6 BGP报文

为了避免路由信息外泄或者非法者对设备进行恶意攻击，可以利用 IPsec 安全隧道对 IPv6 BGP 报文进行保护。通过 IPsec 提供的数据机密性、完整性、数据源认证等功能，确保 IPv6 BGP 报文不会被侦听或恶意篡改，并避免非法者构造 IPv6 BGP 报文对设备进行攻击。

在互为 IPv6 BGP 邻居的两台设备上都配置通过 IPsec 保护 IPv6 BGP 报文后，一端设备在发送 IPv6 BGP 报文时通过 IPsec 对报文进行加封装，另一端设备接收到报文后，通过 IPsec 对报文进行解封

装。如果解封装成功，则接收该报文，正常建立 IPv6 BGP 对等体关系或学习 IPv6 BGP 路由；如果设备接收到不受 IPsec 保护的 IPv6 BGP 报文，或 IPv6 BGP 报文解封装失败，则会丢弃该报文。

表1-89 配置通过 IPsec 保护 IPv6 BGP 报文（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 配置IPsec安全提议和手工方式的IPsec安全框架 | | 配置方法请参见“安全配置指导”中的“IPsec” | 缺省情况下，设备上不存在IPsec安全提议和IPsec安全框架 |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 为IPv6 BGP对等体/对等体组应用IPsec安全框架 | | **peer** { *group-name* \| *ipv6-address* }<br>**ipsec-profile** *profile-name* | 缺省情况下，IPv6 BGP对等体/对等体组没有应用IPsec安全框架<br>应用的安全框架必须是手工方式的IPsec安全框架 |

## 1.7.9 禁止与对等体/对等体组建立会话

由于网络升级维护等原因，需要暂时断开与某个对等体/对等体组的 BGP 会话时，可以通过本配置禁止与该对等体/对等体组建立会话。当网络恢复后，通过执行 **undo peer ignore** 命令恢复与对等体/对等体组的会话。这样，网络管理员无需删除并重新进行对等体/对等体组相关配置，减少了网络维护的工作量。

表1-90 禁止与对等体/对等体组建立会话（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 禁止与对等体/对等体组建立会话 | | **peer** { *group-name* \| *ip-address* } **ignore** | 缺省情况下，允许与BGP对等体/对等体组建立会话 |

表1-91 禁止与对等体/对等体组建立会话（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实 | **bgp** *as-number* | |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 实例视图 | 例视图 | **ip vpn-instance** *vpn-instance-name* | |
| 禁止与IPv6 BGP对等体/对等体组建立会话 | | **peer** { *group-name* \| *ipv6-address* } **ignore** | 缺省情况下，允许与BGP对等体/对等体组建立会话 |

## 1.7.10 配置BGP GTSM功能

提示

- 执行本配置后，只要本地设备和指定的对等体通过了 GTSM 检查，就允许在二者之间建立 EBGP 会话，不管二者之间的跳数是否超过 **peer ebgp-max-hop** 命令指定的跳数范围。
- 使用 BGP GTSM 功能时，要求本设备和对等体设备上同时配置本特性，指定的 *hop-count* 值可以不同，只要能够满足合法性检查即可。

GTSM（Generalized TTL Security Mechanism，通用 TTL 安全保护机制）是一种简单易行的、对基于 IP 协议的上层业务进行保护的安全机制。GTSM 通过检查接收到的 IP 报文头中的 TTL 值是否在一个预先定义好的范围内，来判断 IP 报文是否合法，避免攻击者向网络设备发送大量有效的 IP 报文时对网络设备造成的 CPU 利用（CPU-utilization）等类型的攻击。

配置 BGP GTSM 功能时，用户可以指定本地设备到达某个对等体的最大跳数为 *hop-count*，则从该对等体接收到的 BGP 报文的合法 TTL 范围为 255-“*hop-count*”+1 到 255。只有来自该对等体的报文 TTL 值在该合法范围内时，才将报文上送 CPU 处理；否则，直接丢弃报文。另外，配置 BGP GTSM 功能后，设备会将发送报文的初始 TTL 设置为 255。

对于直连 EBGP 对等体，GTSM 可以提供最佳的保护效果；对于非直连 EBGP 或 IBGP 对等体，由于中间设备可能对 TTL 值进行篡改，GTSM 的保护效果受到中间设备安全性的限制。

表1-92 配置 BGP GTSM 功能（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 为对等体/对等体组使能BGP报文的GTSM安全检测功能 | | **peer** { *group-name* \| *ip-address* } **ttl-security hops** *hop-count* | 缺省情况下，BGP报文的GTSM安全检测功能处于关闭状态 |

表1-93 配置 BGP GTSM 功能（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视 | 进入BGP视图 | **bgp** *as-number* | - |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 图或BGP-VPN实例视图 | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 为对等体/对等体组使能BGP报文的GTSM安全检测功能 | | **peer** { *group-name* \| *ipv6-address* }<br>**ttl-security hops** *hop-count* | 缺省情况下，BGP报文的GTSM安全检测功能处于关闭状态 |

## 1.7.11 配置BGP软复位

BGP 的选路策略改变（即影响 BGP 路由选择的配置，如路由首选值等，发生变化）后，为了使新的策略生效，必须复位 BGP 会话，即删除并重新建立 BGP 会话，以便重新发布路由信息，并应用新的策略对路由信息进行过滤。复位 BGP 会话时，会造成短暂的 BGP 会话中断。

通过 BGP 软复位，可以实现在不中断 BGP 会话的情况下，对 BGP 路由表进行更新，并应用新的选路策略。

BGP 软复位的方法有以下三种：

- 通过 Route-refresh 功能实现 BGP 软复位：如果 BGP 的选路策略发生了变化，则本地路由器会向 BGP 对等体发送 Route-refresh 消息，收到此消息的对等体将其路由信息重新发给本地路由器，本地路由器根据新的路由策略对接收到的路由信息进行过滤。采用这种方式时，要求当前路由器和对等体都支持 Route-refresh 功能。
- 通过将所有路由更新信息保存在本地的方式实现 BGP 软复位：将从对等体接收的所有原始路由更新信息保存在本地，当选路策略发生改变后，对保存在本地的所有路由使用新的路由策略重新进行过滤。采用这种方式时，不要求当前路由器和对等体都支持 Route-refresh 功能，但是保存路由更新需要占用较多的内存资源。
- 手工软复位 BGP 会话：执行 **refresh bgp** 命令手工触发本地路由器将本地路由信息发送给 BGP 对等体或向 BGP 对等体发送 Route-refresh 消息，收到 Route-refresh 消息的对等体将其路由信息重新发给本地路由器，以便本地路由器根据新的路由策略对接收到的路由信息进行过滤。采用这种方式时，要求当前路由器和对等体都支持 Route-refresh 功能。

### 1. 通过Route-refresh功能实现BGP软复位

表1-94 通过 Route-refresh 功能实现 BGP 软复位（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 使能本地路由器与指定对等体/对等体组的BGP路由刷新功能 | | **peer** { *group-name* \| *ip-address* }<br>**capability-advertise route-refresh** | 二者选其一<br>缺省情况下，BGP路由刷新功能处于使能状态 |
| 使能本地路由器与指定BGP对等体/对等体组的BGP路由刷新和多协议扩展功能 | | **undo peer** { *group-name* \| *ip-address* }<br>**capability-advertise conventional** | |

表1-95 通过 Route-refresh 功能实现 BGP 软复位（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 使能本地路由器与指定IPv6 BGP对等体/对等体组的BGP路由刷新功能 | | **peer** { *group-name* \| *ipv6-address* } **capability-advertise route-refresh** | 二者选其一<br>缺省情况下，BGP路由刷新功能处于使能状态 |
| 使能本地路由器与指定IPv6 BGP对等体/对等体组的BGP路由刷新和多协议扩展功能 | | **undo peer** { *group-name* \| *ipv6-address* } **capability-advertise conventional** | |

## 2. 通过将所有路由更新信息保存在本地实现BGP软复位

表1-96 通过将所有路由更新信息保存在本地实现 BGP 软复位（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 保存所有来自指定对等体/对等体组的原始路由更新信息 | | **peer** { *group-name* \| *ip-address* } **keep-all-routes** | 缺省情况下，不保存来自对等体/对等体组的原始路由更新信息<br>本命令只对执行该命令后接收到的路由生效 |

表1-97 通过将所有路由更新信息保存在本地实现 BGP 软复位（IPv6 单播/IPv6 组播）

| 步骤 | | 操作 | 命令 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图或BGP IPv6 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地 | **bgp** *as-number* | |

| 步骤 | | 操作 | 命令 |
| --- | --- | --- | --- |
| 组播地址族视图 | 址族视图 | **address-family ipv6 multicast** | |
| 保存所有来自指定IPv6 BGP对等体/对等体组的原始路由更新信息 | | **peer** { *group-name* \| *ipv6-address* } **keep-all-routes** | 缺省情况下，不保存来自对等体/对等体组的原始路由更新信息<br>本命令只对执行该命令后接收到的路由生效 |

### 3. 手工软复位BGP会话

表1-98 手工软复位 BGP 会话（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 使能本地路由器与指定对等体/对等体组的BGP路由刷新功能 | | **peer** { *group-name* \| *ip-address* } **capability-advertise route-refresh** | 二者选其一<br>缺省情况下，BGP路由刷新功能处于使能状态 |
| 使能本地路由器与指定BGP对等体/对等体组的BGP路由刷新和多协议扩展功能 | | **undo peer** { *group-name* \| *ip-address* } **capability-advertise conventional** | |
| 退回用户视图 | | **return** | - |
| 手工对BGP会话进行软复位 | | **refresh bgp** { *ip-address* \| **all** \| **external** \| **group** *group-name* \| **internal** } { **export** \| **import** } **ipv4** { **multicast** \| [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] } | - |

表1-99 手工软复位 BGP 会话（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 使能本地路由器与IPv6 BGP指定对等体/对等体组的BGP路由刷新功能 | | **peer** { *group-name* \| *ipv6-address* } **capability-advertise route-refresh** | 二者选其一<br>缺省情况下，BGP路由刷新功能 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 使能本地路由器与指定IPv6 BGP对等体/对等体组的BGP路由刷新和多协议扩展功能 | **undo peer** { *group-name* \| *ipv6-address* } **capability-advertise conventional** | 处于使能状态 |
| 退回用户视图 | **return** | - |
| 手工对BGP会话进行软复位 | **refresh bgp** { *ipv6-address* \| **all** \| **external** \| **group** *group-name* \| **internal** } { **export** \| **import** } **ipv6** { **multicast** \| [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] } | - |

## 1.7.12 配置系统进入二级内存门限告警状态后不断开EBGP对等体

当系统进入二级内存门限告警状态后，BGP 会周期性地选择一个 EBGP 对等体，断开与该对等体之间的 BGP 会话，直到系统内存恢复为止。用户可以通过本配置来避免在二级内存门限告警状态下，断开与指定 EBGP 对等体/对等体组之间的 BGP 会话，以达到对特定 EBGP 对等体/对等体组进行保护的目的。

内存告警门限的详细介绍，请参见“基础配置指导”中的“设备管理”。

表1-100 配置系统进入二级内存门限告警状态后不断开 EBGP 对等体（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置系统进入二级内存门限告警状态后，不断开与指定EBGP对等体/对等体组之间的会话 | | **peer** { *group-name* \| *ip-address* } **low-memory-exempt** | 缺省情况下，系统在二级内存门限告警状态下，会周期性地选择EBGP对等体，并断开与该对等体之间的BGP会话 |

表1-101 配置系统进入二级内存门限告警状态后不断开 EBGP 对等体（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置系统进入二级内存门限告警状态后，不断开与指定EBGP对等体/对等体组之间的会话 | | **peer** { *group-name* \| *ipv6-address* } **low-memory-exempt** | 缺省情况下，系统在二级内存门限告警状态下，会周期性地选择EBGP对等体，并断开与该对等体之间的BGP会话 |

# 1.8 配置大规模BGP网络

在大规模BGP网络中，对等体的数目众多，配置和维护极为不便，可以根据组网需要，配置对等体组、团体、路由反射器或联盟，以降低管理难度和提高路由发布效率。对等体组的配置方法，请参见“1.3.3 配置BGP对等体组”。

## 1.8.1 配置BGP团体

缺省情况下，本地路由器不向对等体/对等体组发布团体属性和扩展团体属性。如果接收到的路由中携带团体属性或扩展团体属性，则本地路由器删除该团体属性或扩展团体属性后，再将路由发布给对等体/对等体组。

通过本配置可以允许本地路由器在向对等体发布路由时携带团体属性或扩展团体属性，以便根据团体属性或扩展团体属性对路由进行过滤和控制。本配置和路由策略配合使用，可以灵活地控制路由中携带的团体属性和扩展团体属性值，例如在路由中添加团体属性或扩展团体属性、修改路由中原有的团体属性或扩展团体属性值。路由策略的详细介绍，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

表1-102 配置 BGP 团体（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置向对等体/对等体组发布团体属性或扩展团体属性 | 配置向对等体/对等体组发布团体属性 | **peer** { *group-name* \| *ip-address* } **advertise-community** | 缺省情况下，不向对等体/对等体组发布团体属性和扩展团体属性 |
| | 配置向对等体/对等体组发布扩展团体属性 | **peer** { *group-name* \| *ip-address* } **advertise-ext-community** | |
| （可选）对发布给对等体/对等体组的路由指定路由策略 | | **peer** { *group-name* \| *ip-address* } **route-policy** *route-policy-name* **export** | 缺省情况下，不指定对等体/对等体组的路由策略 |

表1-103 配置 BGP 团体（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP | 进入BGP IPv6单播地址 | **bgp** *as-number* | - |

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| IPv6单播地址族视图或BGP IPv6组播地址族视图 | 族视图 | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置向IPv6 BGP对等体/对等体组发布团体属性或扩展团体属性 | 配置向对等体/对等体组发布团体属性 | **peer** { *group-name* \| *ipv6-address* } **advertise-community** | 缺省情况下，不向IPv6 BGP对等体/对等体组发布团体属性和扩展团体属性 |
| | 配置向对等体/对等体组发布扩展团体属性 | **peer** { *group-name* \| *ipv6-address* } **advertise-ext-community** | |
| （可选）对发布给IPv6 BGP对等体/对等体组的路由指定路由策略 | | **peer** { *group-name* \| *ipv6-address* } **route-policy** *route-policy-name* **export** | 缺省情况下，不指定对等体/对等体组的路由策略 |

## 1.8.2 配置BGP路由反射

### 1. 配置BGP路由反射器

如果同一个 AS 内有多个 BGP 路由器，为了减少在同一 AS 内建立的 IBGP 连接数，可以把几个 BGP 路由器划分为一个集群，将其中的一台路由器配置为路由反射器，其它路由器作为客户机，通过路由反射器在客户机之间反射路由。

为了增加网络的可靠性和防止单点故障，可以在一个集群中配置一个以上的路由反射器，这时，网络管理员必须给位于相同集群中的每个路由反射器配置相同的集群 ID，以避免路由环路。

表1-104 配置 BGP 路由反射器（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv4单播地址族视图、BGP-VPN IPv4单播地址族视图或BGP IPv4组播地址族视图 | 进入BGP IPv4单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP-VPN IPv4单播地址族视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| | | **address-family ipv4** [ **unicast** ] | |
| | 进入BGP IPv4组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv4 multicast** | |
| 配置本机作为路由反射器，对等体/对等体组作为路由反射器的客户机 | | **peer** { *group-name* \| *ip-address* } **reflect-client** | 缺省情况下，没有配置路由反射器及其客户机 |
| 允许路由反射器在客户机之间反射路由 | | **reflect between-clients** | 缺省情况下，允许路由反射器在客户机之间反射路由 |
| （可选）配置路由反射器的集群ID | | **reflector cluster-id** { *cluster-id* \| *ip-address* } | 缺省情况下，每个路由反射器都使用自己的Router ID作为集群ID |

表1-105 配置 IPv6 BGP 路由反射器（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP IPv6单播地址族视图或BGP IPv6组播地址族视图 | 进入BGP IPv6单播地址族视图 | **bgp** *as-number* | - |
| | | **address-family ipv6** [ **unicast** ] | |
| | 进入BGP IPv6组播地址族视图 | **bgp** *as-number* | |
| | | **address-family ipv6 multicast** | |
| 配置本机作为路由反射器，对等体/对等体组作为路由反射器的客户机 | | **peer** { *group-name* \| *ipv6-address* } **reflect-client** | 缺省情况下，没有配置路由反射器及其客户机 |
| 允许路由反射器在客户机之间反射路由 | | **reflect between-clients** | 缺省情况下，允许路由反射器在客户机之间反射路由 |
| （可选）配置路由反射器的集群ID | | **reflector cluster-id** { *cluster-id* \| *ip-address* } | 缺省情况下，每个路由反射器都使用自己的Router ID作为集群ID |

### 2. 配置忽略BGP路由的ORIGINATOR_ID属性

路由反射器从某个对等体接收到路由后，在反射该路由之前为其添加 ORIGINATOR_ID 属性，标识该路由在本 AS 内的起源。ORIGINATOR_ID 属性的值为该对等体的 Router ID。BGP 路由器接收到路由后，将路由中的 ORIGINATOR_ID 属性值与本地的 Router ID 进行比较，如果二者相同则丢弃该路由，从而避免路由环路。

在某些特殊的组网中（如防火墙组网），如果需要接收 ORIGINATOR_ID 属性值与本地 Router ID 相同的路由，则需要通过本配置忽略 BGP 路由的 ORIGINATOR_ID 属性。

表1-106 配置忽略 BGP 路由的 ORIGINATOR_ID 属性（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置忽略BGP路由的ORIGINATOR_ID属性 | | **peer** { *group-name* \| *ip-address* } **ignore-originatorid** | 缺省情况下，BGP路由器不会忽略BGP路由的ORIGINATOR_ID属性<br>请谨慎使用本命令。如果无法确保执行本命令后网络中不会产生环路，请不要执行本命令<br>执行本命令后，BGP路由的CLUSTER_LIST属性也会被忽略 |

表1-107 配置忽略 BGP 路由的 ORIGINATOR_ID 属性（IPv6 单播/IPv6 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置忽略BGP路由的ORIGINATOR_ID属性 | | **peer** { *group-name* \| *ipv6-address* } **ignore-originatorid** | 缺省情况下，BGP路由器不会忽略BGP路由的ORIGINATOR_ID属性<br>请谨慎使用本命令。如果无法确保执行本命令后网络中不会产生环路，请不要执行本命令<br>执行本命令后，BGP路由的CLUSTER_LIST属性也会被忽略 |

## 1.8.3 配置BGP联盟

联盟是处理 AS 内部的 IBGP 网络连接激增的另一种方法，它将一个自治系统划分为若干个子自治系统，每个子自治系统内部的 IBGP 对等体建立全连接关系，子自治系统之间建立 EBGP 连接关系。

### 1. BGP联盟基本配置

网络管理员将一个自治系统划分为若干个子自治系统后，如果路由器位于联盟中的某个子自治系统中，需要在路由器上做如下配置：

(1) 启动BGP，并指定该路由器所属的子自治系统号，配置方法请参见“1.3.1 启动BGP”；

(2) 配置联盟 ID。在不属于联盟的 BGP 发言者看来，属于同一个联盟的多个子自治系统是一个整体，联盟 ID 就是标识联盟这一整体的自治系统号；

(3) 如果该路由器与该联盟的其它子自治系统建立 EBGP 邻居关系，需要在该路由器上指定该联盟体中除了自己还包含哪些子自治系统。

一个联盟最多可包括 32 个子自治系统，配置属于联盟的子自治系统时使用的 *as-number* 仅在联盟内部有效。

表1-108 配置 BGP 联盟

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 配置联盟的ID | **confederation id** *as-number* | 缺省情况下，未配置联盟的ID |
| 指定一个联盟中包含的子自治系统 | **confederation peer-as** *as-number-list* | 缺省情况下，未指定属于联盟的子自治系统 |

### 2. 配置联盟兼容性

如果其他路由器的联盟实现机制不同于 RFC 3065 标准，可以通过如下配置与未采用 RFC 3065 配置的 AS 联盟兼容。

**表1-109 配置联盟兼容性**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 配置设备可以与未遵循RFC 3065实现联盟的路由器互通 | **confederation nonstandard** | 缺省情况下，设备不能与未遵循RFC 3065实现联盟的路由器互通 |

## 1.9 配置BGP GR

BGP GR（Graceful Restart，平滑重启）是一种在主备倒换或 BGP 协议重启时保证转发业务不中断的机制。GR 有两个角色：

- GR Restarter：发生主备倒换或协议重启，且具有 GR 能力的设备。
- GR Helper：和 GR Restarter 具有邻居关系，协助完成 GR 流程的设备。GR Helper 也具有 GR 能力。

设备既可以作为 GR Restarter，又可以作为 GR Helper。设备的角色由该设备在 BGP GR 过程中的作用决定。

BGP GR 的工作过程为：

(1) GR Restarter 和 GR Helper 通过 Open 消息交互 GR 能力。只有双方都具有 GR 能力时，建立起的 BGP 会话才具备 GR 能力。GR Restarter 还会通过 Open 消息，将本端通过 **graceful-restart timer restart** 命令配置的对端等待重建 BGP 会话时间通告给 GR Helper。

(2) 建立具备 GR 能力的 BGP 会话后，如果 GR Restarter 进行主备倒换或 BGP 协议重启，GR Restarter 不会删除转发表项，仍然按照原有的转发表项转发报文。GR Helper 发现 GR Restarter 进行主备倒换或 BGP 协议重启后，GR Helper 不会删除从该 GR Restarter 学习到的路由，而是将这些路由标记为失效路由，仍按照这些路由转发报文，从而确保在 GR Restarter 进行主备倒换或 BGP 协议重启的过程中，报文转发不会中断。

(3) GR Restarter 主备倒换或 BGP 协议重启完成后，它会重新与 GR Helper 建立 BGP 会话。如果在 GR Restarter 通告的 BGP 会话重建时间内没有成功建立 BGP 会话，则 GR Helper 会删除标记为失效的路由。

(4) 如果在 GR Restarter 通告的 BGP 会话重建时间内成功建立 BGP 会话，则 GR Restarter 和 GR Helper 在建立的 BGP 会话上进行路由信息交互，以便 GR Restarter 恢复路由信息、GR Helper 根据学习到的路由删除路由的失效标记。

(5) 在 GR Restarter 和 GR Helper 上可以通过 **graceful-restart timer wait-for-rib** 命令配置本端等待 End-Of-RIB（End of Routing-Information-Base，路由信息库结束）标记的时间，以控制路由信息收敛的速度。如果在该命令指定的时间内没有完成路由信息的交互，则 GR Restarter

不再接收新的路由，根据已经学习到的 BGP 路由信息更新路由表和转发表，完成 BGP 协议收敛；GR Helper 则删除标记为失效的路由。

说明

- End-Of-RIB 标记用来标识路由更新发送的结束。
- 本端配置的等待 End-Of-RIB 标记的时间不会通告给对端，只用来控制本端路由信息交互的时间，即 GR Restarter 上配置的时间只用来控制 GR Restarter 从 GR Helper 接收路由更新的时间，GR Helper 上配置的时间只用来控制 GR Helper 从 GR Restarter 接收路由更新的时间。

在作为 GR Restarter 和 GR Helper 的设备上均需要进行以下配置。

表1-110 配置 BGP GR

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 使能BGP协议的GR能力 | **graceful-restart** | 缺省情况下，BGP协议的GR能力处于关闭状态 |
| 配置对端等待重建BGP会话的时间 | **graceful-restart timer restart** *timer* | 缺省情况下，对端等待重建BGP会话的时间为150秒<br>对端等待重建BGP会话的时间应小于Open消息中的Holdtime时间 |
| 配置本端等待End-Of-RIB标记的时间 | **graceful-restart timer wait-for-rib** *timer* | 缺省情况下，本端等待End-Of-RIB标记的时间为180秒 |

## 1.10 配置BGP NSR

MSR 系列路由器各款型对于本节所描述的特性的支持情况有所不同，详细差异信息如下：

| 型号 | 特性 | 描述 |
| --- | --- | --- |
| MSR 2600 | 配置BGP NSR | 不支持 |
| MSR 3600 | | 不支持 |
| MSR 5600 | | 支持 |

### 1. 功能简介

BGP NSR（Nonstop Routing，不间断路由）是一种通过在 BGP 协议主备进程之间备份必要的协议状态和数据（如 BGP 邻居信息和路由信息），使得 BGP 协议的主进程中断时，备份进程能够无缝地接管主进程的工作，从而确保对等体感知不到 BGP 协议中断，保持 BGP 路由，并保证转发不会中断的技术。

导致 BGP 主进程中断的事件包括以下几种：

- BGP 主进程重启

- BGP 主进程所在的主控板发生故障
- BGP 主进程所在的主控板进行 ISSU（In-Service Software Upgrade，不中断业务升级）

BGP NSR 与 BGP GR 具有如下区别，请根据实际情况选择合适的方式确保数据转发不中断：

- 对设备要求不同：BGP 协议的主进程和备进程运行在不同的主控板上，因此要运行 BGP NSR 功能，设备上必须有两个或两个以上的主控板。要运行 BGP GR 功能，设备上可以只有一个主控板。
- 对 BGP 对等体的要求不同：使用 BGP NSR 功能时，BGP 对等体不会感知本地设备发生了 BGP 进程的异常重启或主备倒换等故障，不需要 BGP 对等体协助恢复 BGP 路由信息。BGP GR 要求 BGP 对等体具有 GR 能力，并且在 BGP 会话中断恢复时，BGP 对等体能够作为 GR helper 协助本地设备恢复 BGP 路由信息。

### 2. 配置限制和指导

如果在设备上同时配置了 BGP NSR 和 BGP GR 功能，则二者的关系如下：

- BGP NSR 优先级高于 BGP GR，即 BGP 主进程中断时通过 BGP NSR 确保转发不中断，设备不会作为 GR Restarter 启动 GR 过程。
- GR Helper 协助 GR Restarter 恢复重启前状态时，如果 GR Helper 发生了主备进程倒换，则即便 GR Helper 上配置了 BGP NSR，也无法保证 GR 过程成功。

在 MPLS L3VPN 组网中，使能 BGP NSR 功能的同时，需要使能 RIB NSR 功能，以确保流量转发不会中断。关于 RIB NSR 功能的详细介绍，请参见“三层技术-IP 路由配置指导”中的“RIB”。

### 3. 配置步骤

**表1-111 配置 BGP NSR**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 使能BGP NSR功能 | **non-stop-routing** | 缺省情况下，BGP NSR功能处于关闭状态 |

## 1.11 开启告警功能

开启 BGP 模块的告警功能后，当 BGP 的邻居状态变化时 BGP 会产生 RFC 4273 中规定的告警信息，该信息包含邻居地址、最近一次出现错误的错误码和错误子码、当前的邻居状态。生成的告警信息将发送到设备的 SNMP 模块，通过设置 SNMP 中告警信息的发送参数，来决定告警信息输出的相关属性。

有关告警信息的详细介绍，请参见“网络管理和监控配置指导”中的“SNMP”。

**表1-112 开启告警功能**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 开启BGP模块的告警功能 | **snmp-agent trap enable bgp** | 缺省情况下，BGP模块的告警功能处于开启状态 |

## 1.12 使能BGP日志功能

使能 BGP 日志记录功能后，BGP 邻居关系建立以及断开时会生成日志信息，通过 **display bgp peer ipv4 unicast log-info** 命令或 **display bgp peer ipv6 unicast log-info** 命令可以查看记录的日志信息。生成的日志信息还将被发送到设备的信息中心，通过设置信息中心的参数，决定日志信息的输出规则（即是否允许输出以及输出方向）。

有关信息中心参数的配置请参见“网络管理和监控配置指导”中的“信息中心”。

表1-113 使能 BGP 日志功能

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 全局使能BGP日志功能 | **log-peer-change** | 缺省情况下，全局BGP日志功能处于使能状态 |

## 1.13 配置BGP与BFD联动

注意

配置 BGP GR 功能后，请慎用 BGP 与 BFD 联动功能。因为当链路故障时，系统可能还没来得及启用 GR 处理流程，BFD 已经检测到链路故障了，从而导致 GR 失败。如果设备上同时配置了 BGP GR 和 BGP BFD，则在 BGP GR 期间请勿去使能 BGP BFD，否则可能导致 GR 失败。

BGP 协议通过存活时间（Keepalive）定时器和保持时间（Holdtime）定时器来维护邻居关系。但这些定时器都是秒级的，而且根据协议规定，设置的保持时间应该至少为存活时间间隔的三倍。这样使得 BGP 邻居关系的检测比较慢，对于报文收发速度快的接口会导致大量报文丢失。通过配置 BGP 与 BFD 联动，可以使用 BFD 来检测本地路由器和 BGP 对等体之间的链路。当本地路由器和 BGP 对等体之间的链路出现故障时，BFD 可以快速检测到该故障，从而加快 BGP 协议的收敛速度。有关 BFD 的介绍和详细配置，请参见“可靠性配置指导”中的“BFD”。

配置通过 BFD 检测本地路由器和指定 BGP 对等体之间的链路之前，需要先在本地路由器和指定 BGP 对等体之间建立 BGP 会话。

表1-114 配置 BGP 与 BFD 联动（IPv4 单播/IPv4 组播）

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置通过BFD检测本地路由器和指定BGP对等体之间的链路 | **peer** *ip-address* **bfd** [ **multi-hop** \| **single-hop** ] | 缺省情况下，不使用BFD检测本地路由器和BGP对等体之间的链路 |

表1-115 配置 BGP 与 BFD 联动（IPv6 单播/IPv6 组播）

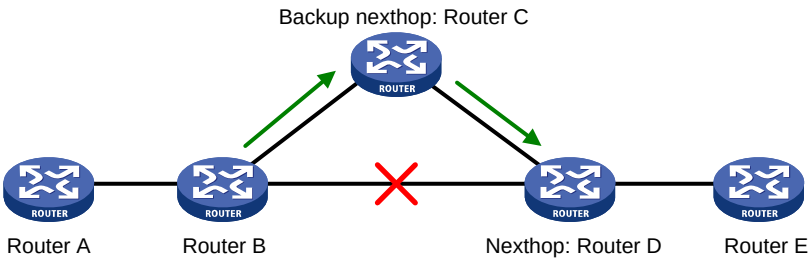
| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入BGP视图或BGP-VPN实例视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN实例视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 配置通过BFD检测本地路由器和指定IPv6 BGP对等体之间的链路 | | **peer** *ipv6-address* **bfd** [ **multi-hop** \| **single-hop** ] | 缺省情况下，不使用BFD检测本地路由器和IPv6 BGP对等体之间的链路 |

## 1.14 配置BGP快速重路由

当 BGP 网络中的链路或某台路由器发生故障时，需要通过故障链路或故障路由器传输才能到达目的地的报文将会丢失或产生路由环路，数据流量将会被中断。直到 BGP 根据新的网络拓扑路由收敛后，被中断的流量才能恢复正常的传输。

为了尽可能缩短网络故障导致的流量中断时间，网络管理员可以开启 BGP 快速重路由功能。

图1-14 BGP 快速重路由功能示意图




如 图 1-14 所示，在Router B上开启快速重路由功能后，BGP将为主路由生成备份下一跳。IPv4 组网中BGP通过ARP或Echo方式的BFD会话检测主路由的下一跳是否可达，IPv6 组网中BGP通过ND（Neighbor Discovery，邻居发现）协议检测主路由的下一跳是否可达。当Router B检测到主路由的下一跳不可达后，BGP会使用备份下一跳替换失效下一跳，通过备份下一跳来指导报文的转发，从而大大缩短了流量中断时间。在使用备份下一跳指导报文转发的同时，BGP会重新进行路由优选，优选完毕后，使用新的最优路由来指导报文转发。

开启 BGP 快速重路由功能的方法有如下两种：

- 在 BGP 地址族视图下执行 **pic** 命令开启当前地址族的 BGP 快速重路由功能。采用这种方法时，BGP 会为当前地址族的所有 BGP 路由自动计算备份下一跳，即只要从不同 BGP 对等体学习到了到达同一目的网络的路由，且这些路由不等价，就会生成主备两条路由。
- 在 BGP 地址族视图下执行 **fast-reroute route-policy** 命令指定快速重路由引用的路由策略，并在引用的路由策略中，通过 **apply [ ipv6 ] fast-reroute backup-nexthop** 命令指定备份下一跳的地址。采用这种方式时，只有为主路由计算出的备份下一跳地址与指定的地址相同时，才会为其生成备份下一跳；否则，不会为主路由生成备份下一跳。在引用的路由策略中，还可以配置 **if-match** 子句，用来决定哪些路由可以进行快速重路由保护，BGP 只会为通过 **if-match** 子句过滤的路由生成备份下一跳。

引用路由策略方式的优先级高于通过 **pic** 命令开启 BGP 快速重路由方式。

IPv4 单播路由和 IPv6 单播路由支持 BGP 快速重路由功能；IPv4 组播路由和 IPv6 组播路由不支持 BGP 快速重路由功能。

表1-116 配置 BGP 快速重路由（IPv4 单播）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置echo报文的源IP地址 | **bfd echo-source-ip** *ip-address* | 缺省情况下，没有配置echo报文的源IP地址<br>通过Echo方式的BFD会话检测主路由的下一跳是否可达时，必须执行本配置<br>echo报文的源IP地址用户可以任意指定。建议配置echo报文的源IP地址不属于该设备任何一个接口所在网段<br>本命令的详细介绍，请参见“可靠性命令参考”中的“BFD” |
| 创建路由策略，并进入路由策略视图 | **route-policy** *route-policy-name* **permit node** *node-number* | 缺省情况下，不存在任何路由策略<br>通过引用路由策略的方式开启BGP快速重路由功能时，必须执行本配置<br>本命令的详细介绍，请参见“三层技术-IP路由命令参考”中的“路由策略” |
| 配置快速重路由的备份下一跳地址 | **apply fast-reroute backup-nexthop** *ip-address* | 缺省情况下，没有指定快速重路由的备份下一跳地址<br>通过引用路由策略的方式开启BGP快速重路由功能时，必须执行本配置<br>本命令的详细介绍，请参见“三层技术-IP路由命令参考”中的“路由策略” |
| 退回系统视图 | **quit** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| （可选）配置通过Echo方式的BFD会话检测主路由的下一跳是否可达 | **primary-path-detect bfd echo** | 缺省情况下，通过ARP检测主路由的下一跳是否可达 |
| （可选）进入BGP-VPN视图 | **ip vpn-instance** *vpn-instance-name* | - |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入BGP IPv4单播地址族视图或BGP-VPN IPv4单播地址族视图 | | **address-family ipv4** [ **unicast** ] | - |
| （二者选其一）开启BGP快速重路由功能 | 开启当前地址族的BGP快速重路由功能 | **pic** | 缺省情况下，BGP快速重路由功能处于关闭状态<br>在某些组网情况下，执行**pic**命令为所有BGP路由生成备份下一跳后，可能会导致路由环路，请谨慎使用本命令 |
| | 在当前地址族视图下指定BGP快速重路由引用的路由策略 | **fast-reroute route-policy** *route-policy-name* | 缺省情况下，BGP快速重路由未引用任何路由策略<br>引用的路由策略中，只有**apply fast-reroute backup-nexthop**和**apply ipv6 fast-reroute backup-nexthop**命令生效，其他**apply**子句不会生效 |

表1-117 配置 BGP 快速重路由（IPv6 单播）

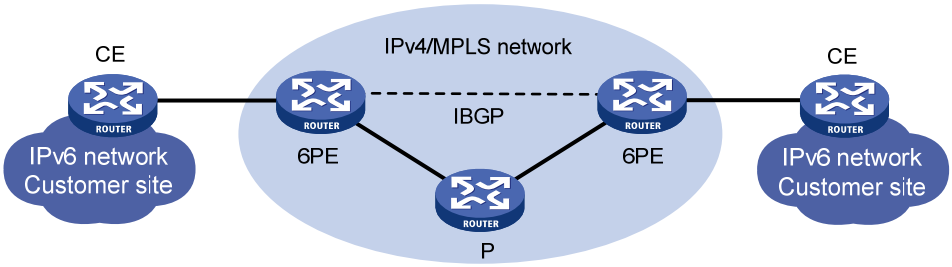
| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 创建路由策略，并进入路由策略视图 | | **route-policy** *route-policy-name* **permit node** *node-number* | 缺省情况下，不存在任何路由策略<br>通过引用路由策略的方式开启BGP快速重路由功能时，必须执行本配置<br>本命令的详细介绍，请参见“三层技术-IP路由命令参考”中的“路由策略” |
| 配置快速重路由的备份下一跳地址 | | **apply ipv6 fast-reroute backup-nexthop** *ipv6-address* | 缺省情况下，没有指定快速重路由的备份下一跳地址<br>通过引用路由策略的方式开启BGP快速重路由功能时，必须执行本配置<br>本命令的详细介绍，请参见“三层技术-IP路由命令参考”中的“路由策略” |
| 退回系统视图 | | **quit** | - |
| 进入BGP视图或BGP-VPN视图 | 进入BGP视图 | **bgp** *as-number* | - |
| | 进入BGP-VPN视图 | **bgp** *as-number* | |
| | | **ip vpn-instance** *vpn-instance-name* | |
| 进入BGP IPv6单播地址族视图或BGP-VPN IPv6单播地址族视图 | | **address-family ipv6** [ **unicast** ] | - |
| （二者选其一）开启BGP快速重路由功能 | 开启当前地址族的BGP快速重路由功能 | **pic** | 缺省情况下，BGP快速重路由功能处于关闭状态<br>在某些组网情况下，执行**pic**命令为所有BGP路由生成备份下一跳后，可能会导致路由环路，请谨慎使用本命令 |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| | 在当前地址族视图下指定BGP快速重路由引用的路由策略 | **fast-reroute route-policy** *route-policy-name* | 缺省情况下，BGP快速重路由未引用任何路由策略<br>引用的路由策略中，只有**apply fast-reroute backup-nexthop**和**apply ipv6 fast-reroute backup-nexthop**命令生效，其他**apply**子句不会生效 |

## 1.15 配置6PE

如 图 1-15 所示，6PE（IPv6 Provider Edge，IPv6 供应商边缘）是一种过渡技术，它采用MPLS（Multiprotocol Label Switching，多协议标签交换）技术实现通过IPv4 骨干网连接隔离的IPv6 用户网络。当ISP希望在自己原有的IPv4/MPLS骨干网的基础上，为用户网络提供IPv6 流量转发能力时，可以采用 6PE技术方便地达到该目的。

图1-15 6PE 组网图




6PE 的主要思想是：

- 6PE 设备从 CE（Customer Edge，用户网络边缘）设备接收到用户网络的 IPv6 路由信息后，为该路由信息分配标签，通过 MP-BGP 会话将带有标签的 IPv6 路由信息发布给对端的 6PE 设备。对端 6PE 设备将接收到的 IPv6 路由信息扩散到本地连接的用户网络。从而，实现 IPv6 用户网络之间的路由信息发布。
- 为了隐藏 IPv6 报文、使得 IPv4 骨干网中的设备能够转发 IPv6 用户网络的报文，在 IPv4 骨干网络中需要建立公网隧道。公网隧道可以是 GRE 隧道、MPLS LSP、MPLS TE 隧道等。
- 6PE 设备转发 IPv6 报文时，先为 IPv6 报文封装 IPv6 路由信息对应的标签（内层标签），再为其封装公网隧道对应的标签（外层标签）。骨干网中的设备根据外层标签转发报文，意识不到该报文为 IPv6 报文。对端 6PE 设备接收到报文后，删除内层和外层标签，将原始的 IPv6 报文转发到本地连接的用户网络。

MPLS、MPLS TE、CE 设备、P（Provider，服务提供商网络）设备的详细介绍，请参见“MPLS 配置指导”。GRE 的详细介绍，请参见“三层技术-IP 业务配置指导”中的“GRE”。

说明

为了实现IPv6路由信息的交互，CE和6PE之间可以配置IPv6静态路由、运行IPv6 IGP协议或IPv6 BGP协议。

## 1.15.2 配置 6PE基本功能

6PE 组网环境中，需要进行如下配置：

- 在 IPv4 骨干网上建立公网隧道。具体配置请参见“三层技术-IP 业务配置指导”中的“GRE”或“MPLS 配置指导”。
- 在 6PE 设备上配置 MPLS 基本能力。具体配置请参见“MPLS 配置指导”中的“MPLS 基础”。
- 在 6PE 设备上配置 BGP 相关功能，以便通过 BGP 会话发布带有标签的 IPv6 路由信息。本文只介绍 6PE 设备上的 BGP 相关配置。

表1-118 配置 6PE 基本功能

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 指定6PE对等体/对等体组的AS号 | **peer** { *group-name* \| *ip-address* } **as-number** *as-number* | 缺省情况下，未指定6PE对等体/对等体组的AS号 |
| 进入BGP IPv6单播地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 使能本地路由器与6PE对等体/对等体组交换IPv6单播路由信息的能力 | **peer** { *group-name* \| *ip-address* } **enable** | 缺省情况下，本地路由器不能与6PE对等体/对等体组交换IPv6单播路由信息 |
| 使能与6PE对等体/对等体组交换带标签IPv6路由的能力 | **peer** { *group-name* \| *ip-address* } **label-route-capability** | 缺省情况下，不具有与6PE对等体/对等体组交换带标签IPv6路由的能力 |

## 1.15.3 配置 6PE可选功能

表1-119 配置 6PE 可选功能

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入BGP视图 | **bgp** *as-number* | - |
| 进入BGP IPv6单播地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 配置向6PE对等体/对等体组发布团体属性 | **peer** { *group-name* \| *ip-address* } **advertise-community** | 缺省情况下，不向6PE对等体/对等体组发布团体属性 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置向6PE对等体/对等体组发布扩展团体属性 | **peer** { *group-name* \| *ip-address* } **advertise-ext-community** | 缺省情况下，不向6PE对等体/对等体组发布扩展团体属性 |
| 配置对于从6PE对等体/对等体组接收的路由，允许本地AS号在接收路由的AS_PATH属性中出现，并配置允许出现的次数 | **peer** { *group-name* \| *ip-address* } **allow-as-loop** [ *number* ] | 缺省情况下，不允许本地AS号在接收路由的AS_PATH属性中出现 |
| 为6PE对等体/对等体组配置基于AS路径过滤列表的路由过滤策略 | **peer** { *group-name* \| *ip-address* } **as-path-acl** *as-path-acl-number* { **export** \| **import** } | 缺省情况下，没有为6PE对等体/对等体组配置基于AS路径过滤列表的路由接收过滤策略 |
| 为6PE对等体/对等体组配置基于IPv6 ACL的过滤策略 | **peer** { *group-name* \| *ip-address* } **filter-policy** *acl6-number* { **export** \| **import** } | 缺省情况下，没有为6PE对等体/对等体组配置基于IPv6 ACL的过滤策略 |
| 为6PE对等体/对等体组配置基于IPv6地址前缀列表的路由过滤策略 | **peer** { *group-name* \| *ip-address* } **prefix-list** *ipv6-prefix-name* { **export** \| **import** } | 缺省情况下，没有为6PE对等体/对等体组配置基于IPv6前缀列表的路由过滤策略 |
| 为6PE对等体/对等体组设置基于路由策略的路由过滤策略 | **peer** { *group-name* \| *ip-address* } **route-policy** *route-policy-name* { **export** \| **import** } | 缺省情况下，没有为6PE对等体/对等体组配置基于路由策略的路由过滤策略 |
| 向6PE对等体/对等体组发送缺省路由 | **peer** { *group-name* \| *ip-address* } **default-route-advertise** [ **route-policy** *route-policy-name* ] | 缺省情况下，不向6PE对等体/对等体组发送缺省路由 |
| 保存所有来自指定6PE对等体/对等体组的原始路由更新信息 | **peer** { *group-name* \| *ip-address* } **keep-all-routes** | 缺省情况下，不保存来自6PE对等体/对等体组的原始路由更新信息 |
| 配置向指定6PE对等体/对等体组发送BGP更新消息时只携带公有AS号，不携带私有AS号 | **peer** { *group-name* \| *ip-address* } **public-as-only** | 缺省情况下，向4PE对等体/对等体组发送BGP更新消息时，既可以携带公有AS号，又可以携带私有AS号 |
| 配置允许从6PE对等体/对等体组接收的路由的最大数量 | **peer** { *group-name* \| *ip-address* } **route-limit** *prefix-number* [ { **alert-only** \| **discard** \| **reconnect** *reconnect-time* } \| *percentage-value* ] * | 缺省情况下，不限制从6PE对等体/对等体组接收的路由数量 |
| 为从指定6PE对等体/对等体组接收的路由分配首选值 | **peer** { *group-name* \| *ip-address* } **preferred-value** *value* | 缺省情况下，从6PE对等体/对等体组接收的路由的首选值为0 |
| 配置本机作为路由反射器，对等体/对等体组作为路由反射器的客户机 | **peer** { *group-name* \| *ip-address* } **reflect-client** | 缺省情况下，没有配置路由反射器及其客户机 |
| 为BGP对等体/对等体组配置SoO属性 | **peer** { *group-name* \| *ip-address* } **soo** *site-of-origin* | 缺省情况下，没有为BGP对等体/对等体组配置SoO属性 |
| 退回用户视图 | **return** | - |
| 显示6PE对等体/对等体组信息 | **display bgp peer ipv6** [ **unicast** ] [ **group-name** *group-name* **log-info** \| *ip-address* { **log-info** \| **verbose** } \| **verbose** ] | **display**命令可在任意视图下执行 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 显示向指定的6PE对等体发送或者从指定的对等体收到的路由信息 | **display bgp routing-table ipv6** [ **unicast** ] **peer** *ip-address* { **advertised-routes** \| **received-routes** } [ *network-address prefix-length* \| **statistics** ] | **display**命令可在任意视图下执行 |
| 软复位指定的BGP 6PE连接 | **refresh bgp** *ip-address* { **export** \| **import** } **ipv6** [ **unicast** ] | **refresh**命令在用户视图下执行 |
| 复位指定的BGP 6PE连接 | **reset bgp** *ip-address* **ipv6** [ **unicast** ] | **reset**命令在用户视图下执行 |

# 1.16 BGP显示和维护

## 1.16.1 显示BGP

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 BGP 的运行情况，通过查看显示信息验证配置的效果。

表1-120 BGP 配置显示（IPv4 单播）

| 操作 | 命令 |
| --- | --- |
| 显示BGP NSR的运行状态 | **display bgp non-stop-routing status** |
| 显示BGP IPv4单播对等体组的信息 | **display bgp group ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ **group-name** *group-name* ] |
| 显示BGP IPv4单播对等体或对等体组的信息（MSR 2600/MSR 3600） | **display bgp peer ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ { *ip-address* \| **group-name** *group-name* } **log-info** \| [ *ip-address* ] **verbose** ] |
| 显示BGP IPv4单播对等体或对等体组的信息（MSR 5600） | **display bgp peer ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ { *ip-address* \| **group-name** *group-name* } **log-info** \| [ [ *ip-address* ] **verbose** ] [ **standby slot** *slot-number* ] ] |
| 显示BGP IPv4单播路由信息（MSR 2600/MSR 3600） | **display bgp routing-table ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address* [ { *mask* \| *mask-length* } [ **longest-match** ] ] \| *network-address* [ *mask* \| *mask-length* ] **advertise-info** \| **as-path-acl** *as-path-acl-number* \| **community-list** { { *basic-community-list-number* \| *comm-list-name* } [ **whole-match** ] \| *adv-community-list-number* } \| **peer** *ip-address* { **advertised-routes** \| **received-routes** } [ *network-address* [ *mask* \| *mask-length* ] \| **statistics** ] \| **statistics** ] |
| 显示BGP IPv4单播路由信息（MSR 5600） | **display bgp routing-table ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address* [ { *mask* \| *mask-length* } [ **longest-match** ] ] \| *network-address* [ *mask* \| *mask-length* ] **advertise-info** \| **as-path-acl** *as-path-acl-number* \| **community-list** { { *basic-community-list-number* \| *comm-list-name* } [ **whole-match** ] \| *adv-community-list-number* } \| **peer** *ip-address* { **advertised-routes** \| **received-routes** } [ *network-address* [ *mask* \| *mask-length* ] \| **statistics** ] \| **statistics** ] [ **standby slot** *slot-number* ] |
| 显示衰减的BGP IPv4单播路由信息 | **display bgp routing-table dampened ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] |
| 显示BGP IPv4单播路由的路由衰减参数 | **display bgp dampening parameter ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] |

| 操作 | 命令 |
| --- | --- |
| 显示BGP IPv4单播路由的震荡统计信息 | **display bgp routing-table flap-info ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address* [ { *mask* \| *mask-length* } [ **longest-match** ] ] \| **as-path-acl** *as-path-acl-number* ] |
| 显示通过**network**命令发布的路由信息和通过**network short-cut**命令配置的Short-cut路由信息 | **display bgp network ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] |
| 显示BGP的路由属性信息 | **display bgp paths** [ *as-regular-expression* ] |
| 显示BGP IPv4单播地址族下打包组的相关信息 | **display bgp update-group ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *ip-address* ] |

表1-121 BGP 配置显示（IPv6 单播）

| 操作 | 命令 |
| --- | --- |
| 显示BGP NSR的运行状态 | **display bgp non-stop-routing status** |
| 显示BGP IPv6单播对等体组的信息 | **display bgp group ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ **group-name** *group-name* ] |
| 显示BGP IPv6单播对等体或对等体组的信息（MSR 2600/MSR 3600） | **display bgp peer ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ { *ipv6-address* \| **group-name** *group-name* } **log-info** \| [ *ipv6-address* ] **verbose** ]<br>**display bgp peer ipv6** [ **unicast** ] [ *ip-address* **log-info** \| [ *ip-address* ] **verbose** ] |
| 显示BGP IPv6单播对等体或对等体组的信息（MSR 5600） | **display bgp peer ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ { *ipv6-address* \| **group-name** *group-name* } **log-info** \| [ [ *ipv6-address* ] **verbose** ] [ **standby slot** *slot-number* ] ]<br>**display bgp peer ipv6** [ **unicast** ] [ *ip-address* **log-info** \| [ [ *ip-address* ] **verbose** ] [ **standby slot** *slot-number* ] ] |
| 显示BGP IPv6单播路由信息（MSR 2600/MSR 3600） | **display bgp routing-table ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address prefix-length* [ **advertise-info** ] \| **as-path-acl** *as-path-acl-number* \| **community-list** { { *basic-community-list-number* \| *comm-list-name* } [ **whole-match** ] \| *adv-community-list-number* } \| **peer** *ipv6-address* { **advertised-routes** \| **received-routes** } [ *network-address prefix-length* \| **statistics** ] \| **statistics** ]<br>**display bgp routing-table ipv6** [ **unicast** ] **peer** *ip-address* { **advertised-routes** \| **received-routes** } [ *network-address prefix-length* \| **statistics** ] |
| 显示BGP IPv6单播路由信息（MSR 5600） | **display bgp routing-table ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address prefix-length* [ **advertise-info** ] \| **as-path-acl** *as-path-acl-number* \| **community-list** { { *basic-community-list-number* \| *comm-list-name* } [ **whole-match** ] \| *adv-community-list-number* } \| **peer** *ipv6-address* { **advertised-routes** \| **received-routes** } [ *network-address prefix-length* \| **statistics** ] \| **statistics** ] [ **standby slot** *slot-number* ]<br>**display bgp routing-table ipv6** [ **unicast** ] **peer** *ip-address* { **advertised-routes** \| **received-routes** } [ *network-address prefix-length* \| **statistics** ] [ **standby slot** *slot-number* ] |

| 操作 | 命令 |
| --- | --- |
| 显示衰减的BGP IPv6单播路由信息 | **display bgp routing-table dampened ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] |
| 显示BGP IPv6单播路由的路由衰减参数 | **display bgp dampening parameter ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] |
| 显示BGP IPv6单播路由的震荡统计信息 | **display bgp routing-table flap-info ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address prefix-length* \| **as-path-acl** *as-path-acl-number* ] |
| 显示BGP IPv6单播路由的入标签信息 | **display bgp routing-table ipv6** [ **unicast** ] **inlabel** |
| 显示BGP IPv6单播路由的出标签信息（MSR 2600/MSR 3600） | **display bgp routing-table ipv6** [ **unicast** ] **outlabel** |
| 显示BGP IPv6单播路由的出标签信息（MSR 5600） | **display bgp routing-table ipv6** [ **unicast** ] **outlabel** [ **standby slot** *slot-number* ] |
| 显示通过**network**命令发布的路由信息和通过**network short-cut**命令配置的Short-cut路由信息 | **display bgp network ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] |
| 显示BGP的路由属性信息 | **display bgp paths** [ *as-regular-expression* ] |
| 显示BGP IPv6单播地址族下打包组的相关信息 | **display bgp update-group ipv6** [ **unicast** ] [ *ip-address* \| *ipv6-address* ]<br>**display bgp update-group ipv6** [ **unicast** ] **vpn-instance** *vpn-instance-name* [ *ipv6-address* ] |

表1-122 BGP 配置显示（IPv4 组播）

| 操作 | 命令 |
| --- | --- |
| 显示BGP NSR的运行状态 | **display bgp non-stop-routing status** |
| 显示BGP IPv4组播对等体组的信息 | **display bgp group ipv4 multicast** [ **group-name** *group-name* ] |
| 显示BGP IPv4组播对等体或对等体组的信息（MSR 2600/MSR 3600） | **display bgp peer ipv4 multicast** [ { *ip-address* \| **group-name** *group-name* } **log-info** \| [ *ip-address* ] **verbose** ] |
| 显示BGP IPv4组播对等体或对等体组的信息（MSR 5600） | **display bgp peer ipv4 multicast** [ { *ip-address* \| **group-name** *group-name* } **log-info** \| [ [ *ip-address* ] **verbose** ] [ **standby slot** *slot-number* ] ] |
| 显示BGP IPv4组播路由信息（MSR 2600/MSR 3600） | **display bgp routing-table ipv4 multicast** [ *network-address* [ { *mask* \| *mask-length* } [ **longest-match** ] ] \| *network-address* [ *mask* \| *mask-length* ] **advertise-info** \| **as-path-acl** *as-path-acl-number* \| **community-list** { { *basic-community-list-number* \| *comm-list-name* } [ **whole-match** ] \| *adv-community-list-number* } \| **peer** *ip-address* { **advertised-routes** \| **received-routes** } [ *network-address* [ *mask* \| *mask-length* ] \| **statistics** ] \| **statistics** ] |

| 操作 | 命令 |
| --- | --- |
| 显示BGP IPv4组播路由信息（MSR 5600） | **display bgp routing-table ipv4 multicast** [ *network-address* [ { *mask* \| *mask-length* } [ **longest-match** ] ] \| *network-address* [ *mask* \| *mask-length* ] **advertise-info** \| **as-path-acl** *as-path-acl-number* \| **community-list** { { *basic-community-list-number* \| *comm-list-name* } [ **whole-match** ] \| *adv-community-list-number* } \| **peer** *ip-address* { **advertised-routes** \| **received-routes** } [ *network-address* [ *mask* \| *mask-length* ] \| **statistics** ] \| **statistics** ] [ **standby slot** *slot-number* ] |
| 显示衰减的BGP IPv4组播路由信息 | **display bgp routing-table dampened ipv4 multicast** |
| 显示BGP IPv4组播路由的路由衰减参数 | **display bgp dampening parameter ipv4 multicast** |
| 显示BGP IPv4组播路由的震荡统计信息 | **display bgp routing-table flap-info ipv4 multicast** [ *network-address* [ { *mask* \| *mask-length* } [ **longest-match** ] ] \| **as-path-acl** *as-path-acl-number* ] |
| 显示通过**network**命令发布的路由信息和通过**network short-cut**命令配置的Short-cut路由信息 | **display bgp network ipv4 multicast** |
| 显示BGP的路由属性信息 | **display bgp paths** [ *as-regular-expression* ] |
| 显示BGP IPv4组播地址族下打包组的相关信息 | **display bgp update-group ipv4 multicast** [ *ip-address* ] |

表1-123 BGP 配置显示（IPv6 组播）

| 操作 | 命令 |
| --- | --- |
| 显示BGP NSR的运行状态 | **display bgp non-stop-routing status** |
| 显示BGP IPv6组播对等体组的信息 | **display bgp group ipv6 multicast** [ **group-name** *group-name* ] |
| 显示BGP IPv6组播对等体或对等体组的信息（MSR 2600/MSR 3600） | **display bgp peer ipv6 multicast** [ { *ipv6-address* \| **group-name** *group-name* } **log-info** \| [ *ipv6-address* ] **verbose** ] |
| 显示BGP IPv6组播对等体或对等体组的信息（MSR 5600） | **display bgp peer ipv6 multicast** [ { *ipv6-address* \| **group-name** *group-name* } **log-info** \| [ [ *ipv6-address* ] **verbose** ] [ **standby slot** *slot-number* ] ] |
| 显示BGP IPv6组播路由信息（MSR 2600/MSR 3600） | **display bgp routing-table ipv6 multicast** [ *network-address prefix-length* [ **advertise-info** ] \| **as-path-acl** *as-path-acl-number* \| **community-list** { { *basic-community-list-number* \| *comm-list-name* } [ **whole-match** ] \| *adv-community-list-number* } \| **peer** *ipv6-address* { **advertised-routes** \| **received-routes** } [ *network-address prefix-length* \| **statistics** ] \| **statistics** ] |
| 显示BGP IPv6组播路由信息（MSR 5600） | **display bgp routing-table ipv6 multicast** [ *network-address prefix-length* [ **advertise-info** ] \| **as-path-acl** *as-path-acl-number* \| **community-list** { { *basic-community-list-number* \| *comm-list-name* } [ **whole-match** ] \| *adv-community-list-number* } \| **peer** *ipv6-address* { **advertised-routes** \| **received-routes** } [ *network-address prefix-length* \| **statistics** ] \| **statistics** ] [ **standby slot** *slot-number* ] |
| 显示衰减的BGP IPv6组播路由信息 | **display bgp routing-table dampened ipv6 multicast** |
| 显示BGP IPv6组播路由的路由衰减参数 | **display bgp dampening parameter ipv6 multicast** |

| 操作 | 命令 |
| --- | --- |
| 显示BGP IPv6组播路由的震荡统计信息 | **display bgp routing-table flap-info ipv6 multicast** [ *network-address prefix-length* \| **as-path-acl** *as-path-acl-number* ] |
| 显示通过**network**命令发布的路由信息和通过**network short-cut**命令配置的Short-cut路由信息 | **display bgp network ipv6 multicast** |
| 显示BGP的路由属性信息 | **display bgp paths** [ *as-regular-expression* ] |
| 显示BGP IPv6组播地址族下打包组的相关信息 | **display bgp update-group ipv6 multicast** [ *ipv6-address* ] |

## 1.16.2 复位BGP会话

当 BGP 路由策略或协议发生变化后，如果需要通过复位 BGP 会话使新的配置生效，请在用户视图下进行下列配置。

表1-124 复位 BGP 会话

| 操作 | 命令 |
| --- | --- |
| 复位所有BGP会话 | **reset bgp all** |
| 复位IPv4单播地址族下的BGP会话 | **reset bgp** { *as-number* \| *ip-address* \| **all** \| **external** \| **group** *group-name* \| **internal** } **ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] |
| 复位IPv6单播地址族下的BGP会话 | **reset bgp** { *as-number* \| *ipv6-address* \| **all** \| **external** \| **group** *group-name* \| **internal** } **ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ]<br>**reset bgp** *ip-address* **ipv6** [ **unicast** ] |
| 复位IPv4组播地址族下的BGP会话 | **reset bgp** { *as-number* \| *ip-address* \| **all** \| **external** \| **group** *group-name* \| **internal** } **ipv4 multicast** |
| 复位IPv6组播地址族下的BGP会话 | **reset bgp** { *as-number* \| *ipv6-address* \| **all** \| **external** \| **group** *group-name* \| **internal** } **ipv6 multicast** |

## 1.16.3 清除BGP信息

在用户视图下，执行 **reset** 命令可以清除 BGP 相关统计信息。

表1-125 清除 BGP 信息

| 操作 | 命令 |
| --- | --- |
| 清除BGP IPv4单播路由的衰减信息，并解除对BGP IPv4单播路由的抑制 | **reset bgp dampening ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address* [ *mask* \| *mask-length* ] ] |
| 清除BGP IPv4单播路由的振荡统计信息 | **reset bgp flap-info ipv4** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address* [ *mask* \| *mask-length* ] \| **as-path-acl** *as-path-acl-number* \| **peer** *ipv4-address* ] |
| 清除BGP IPv6单播路由的衰减信息，并解除对BGP IPv6单播路由的抑制 | **reset bgp dampening ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address prefix-length* ] |

| 操作 | 命令 |
| --- | --- |
| 清除BGP IPv6单播路由的振荡统计信息 | **reset bgp flap-info ipv6** [ **unicast** ] [ **vpn-instance** *vpn-instance-name* ] [ *network-address prefix-length* \| **as-path-acl** *as-path-acl-number* \| **peer** *ipv6-address* ] |
| 清除BGP IPv4组播路由的衰减信息，并解除对BGP IPv4组播路由的抑制 | **reset bgp dampening ipv4 multicast** [ *network-address* [ *mask* \| *mask-length* ] ] |
| 清除BGP IPv4组播路由的振荡统计信息 | **reset bgp flap-info ipv4 multicast** [ *network-address* [ *mask* \| *mask-length* ] \| **as-path-acl** *as-path-acl-number* \| **peer** *ipv4-address* ] |
| 清除BGP IPv6组播路由的衰减信息，并解除对BGP IPv6组播路由的抑制 | **reset bgp dampening ipv6 multicast** [ *network-address prefix-length* ] |
| 清除BGP IPv6组播路由的振荡统计信息 | **reset bgp flap-info ipv6 multicast** [ *network-address prefix-length* \| **as-path-acl** *as-path-acl-number* \| **peer** *ipv6-address* ] |

# 1.17 IPv4 BGP典型配置举例

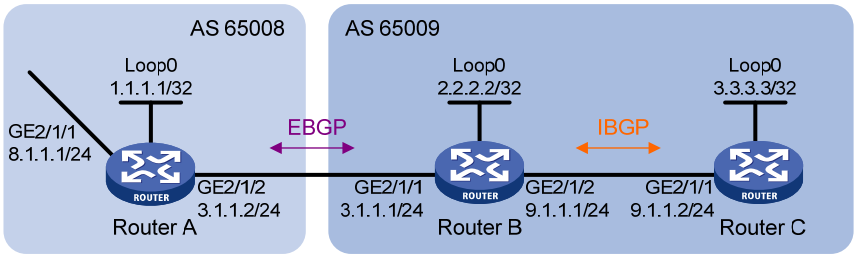
## 1.17.1 BGP基本配置

### 1. 组网需求

如 图 1-16 所示，所有路由器均运行BGP协议。要求Router A和Router B之间建立EBGP连接，Router B和Router C之间建立IBGP连接，使得Router C能够访问Router A直连的 8.1.1.0/24 网段。

### 2. 组网图

图1-16 BGP 基本配置组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 IBGP 连接

- 为了防止端口状态不稳定引起路由震荡，本举例使用 Loopback 接口来创建 IBGP 对等体。
- 使用 Loopback 接口创建 IBGP 对等体时，因为 Loopback 接口不是两对等体实际连接的接口，所以，必须使用 **peer connect-interface** 命令将 Loopback 接口配置为 BGP 连接的源接口。
- 在AS 65009内部，使用OSPF协议，保证Router B到Router C的Loopback接口路由可达。

# 配置 Router B。

```
<RouterB> system-view
```

```
[RouterB] bgp 65009
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 3.3.3.3 as-number 65009
[RouterB-bgp] peer 3.3.3.3 connect-interface loopback 0
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] peer 3.3.3.3 enable
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
[RouterB] ospf 1
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 2.2.2.2 0.0.0.0
[RouterB-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
[RouterB-ospf-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] bgp 65009
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 2.2.2.2 as-number 65009
[RouterC-bgp] peer 2.2.2.2 connect-interface loopback 0
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 2.2.2.2 enable
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
[RouterC] ospf 1
[RouterC-ospf-1] area 0
[RouterC-ospf-1-area-0.0.0.0] network 3.3.3.3 0.0.0.0
[RouterC-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] quit
[RouterC-ospf-1] quit
[RouterC] display bgp peer ipv4

 BGP local router ID : 3.3.3.3
 Local AS number : 65009
 Total number of peers : 1                 Peers in established state : 1

  Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State

  2.2.2.2              65009        7       10    0       0 00:06:09 Established
```

以上显示信息表明 Router B 和 Router C 之间的 IBGP 连接已经建立。

(3) 配置 EBGP 连接

- EBGP 邻居关系的两台路由器（通常属于两个不同运营商），处于不同的 AS 域，对端的 Loopback 接口一般路由不可达，所以一般使用直连地址建立 EBGP 邻居。
- 因为要求 Router C 能够访问 Router A 直连的 8.1.1.0/24 网段，所以，建立 EBGP 连接后，需要将 8.1.1.0/24 网段路由通告到 BGP 路由表中。

# 配置 Router A。

```
<RouterA> system-view
```

```
[RouterA] bgp 65008
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 3.1.1.1 as-number 65009
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 3.1.1.1 enable
[RouterA-bgp-ipv4] network 8.1.1.0 24
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
[RouterB] bgp 65009
[RouterB-bgp] peer 3.1.1.2 as-number 65008
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] peer 3.1.1.2 enable
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
```

# 查看 Router B 的 BGP 对等体的连接状态。

```
[RouterB] display bgp peer ipv4

 BGP local router ID : 2.2.2.2
 Local AS number : 65009
 Total number of peers : 2                 Peers in established state : 2

  Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State

  3.3.3.3              65009       12       10    0       3 00:09:16 Established
  3.1.1.2              65008        3        3    0       1 00:00:08 Established
```

可以看出，Router B 与 Router C、Router B 与 Router A 之间的 BGP 连接均已建立。

# 查看 Router A 的 BGP 路由表。

```
[RouterA] display bgp routing-table ipv4

 Total number of routes: 1

 BGP local router ID is 1.1.1.1
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >  8.1.1.0/24         8.1.1.1         0                     32768   i
```

# 显示 Router B 的 BGP 路由表。

```
[RouterB] display bgp routing-table ipv4

 Total number of routes: 1

 BGP local router ID is 2.2.2.2
 Status codes: * - valid, > - best, d - dampened, h - history,
```

```
              s - suppressed, S - stale, i - internal, e - external
              Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >e 8.1.1.0/24         3.1.1.2         0                     0       65008i
```

# 显示 Router C 的 BGP 路由表。

```
[RouterC] display bgp routing-table ipv4

 Total number of routes: 1

 BGP local router ID is 3.3.3.3
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

   i 8.1.1.0/24         3.1.1.2         0          100        0       65008i
```

从路由表可以看出，Router A 没有学到 AS 65009 内部的任何路由，Router C 虽然学到了 AS 65008 中的 8.1.1.0 的路由，但因为下一跳 3.1.1.2 不可达，所以也不是有效路由。

(4) 配置 BGP 引入直连路由

在 Router B 上配置 BGP 引入直连路由，以便 Router A 能够获取到网段 9.1.1.0/24 的路由，Router C 能够获取到网段 3.1.1.0/24 的路由。

# 配置 Router B。

```
[RouterB] bgp 65009
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] import-route direct
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
```

# 显示 Router A 的 BGP 路由表。

```
[RouterA] display bgp routing-table ipv4

 Total number of routes: 4

 BGP local router ID is 1.1.1.1
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >e 2.2.2.2/32         3.1.1.1         0                     0       65009?
* >e 3.1.1.0/24         3.1.1.1         0                     0       65009?
* >  8.1.1.0/24         8.1.1.1         0                     32768   i
* >e 9.1.1.0/24         3.1.1.1         0                     0       65009?
```

以上显示信息表明，在 Router B 上引入直连路由后，Router A 新增了到达 2.2.2.2/32 和 9.1.1.0/24 的两条路由。

# 显示 Router C 的 BGP 路由表。

```
[RouterC] display bgp routing-table ipv4

 Total number of routes: 4

 BGP local router ID is 3.3.3.3
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >i 2.2.2.2/32         2.2.2.2         0          100        0       ?
* >i 3.1.1.0/24         2.2.2.2         0          100        0       ?
* >i 8.1.1.0/24         3.1.1.2         0          100        0       65008i
* >i 9.1.1.0/24         2.2.2.2         0          100        0       ?
```

以上显示信息表明，到 8.1.1.0 的路由变为有效路由，下一跳为 Router A 的地址。

#### 4. 验证配置

# 使用 Ping 进行验证。

```
[RouterC] ping 8.1.1.1
Ping 8.1.1.1 (8.1.1.1): 56 data bytes, press CTRL_C to break
56 bytes from 8.1.1.1: icmp_seq=0 ttl=255 time=2.000 ms
56 bytes from 8.1.1.1: icmp_seq=1 ttl=255 time=0.000 ms
56 bytes from 8.1.1.1: icmp_seq=2 ttl=255 time=0.000 ms
56 bytes from 8.1.1.1: icmp_seq=3 ttl=255 time=0.000 ms
56 bytes from 8.1.1.1: icmp_seq=4 ttl=255 time=1.000 ms

--- Ping statistics for 8.1.1.1 ---
5 packet(s) transmitted, 5 packet(s) received, 0.0% packet loss
round-trip min/avg/max/std-dev = 0.000/0.600/2.000/0.800 ms
```

### 1.17.2 BGP与IGP交互配置

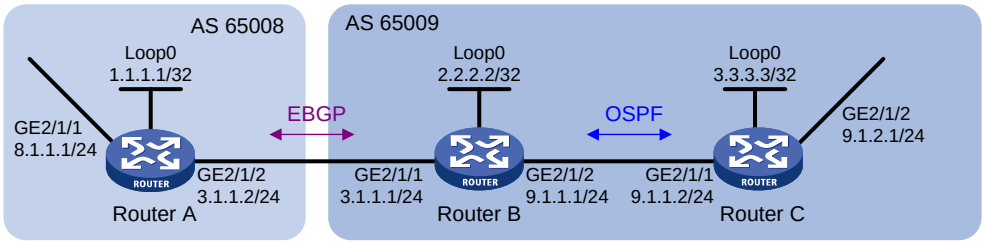
#### 1. 组网需求

如 图 1-17 所示，公司A的所有设备在AS 65008 内，公司B的所有设备在AS 65009 内，AS 65008 和AS 65009 通过设备Router A和Router B相连。

现要求实现 Router A 能够访问 AS 65009 内的网段 9.1.2.0/24，Router C 能够访问 AS 65008 内的网段 8.1.1.0/24。

## 2. 组网图

图1-17 BGP 与 IGP 交互配置组网图




## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 OSPF

在 AS 65009 内配置 OSPF，使得 Router B 能获取到到 9.1.2.0/24 网段的路由。

# 配置 Router B。

```
<RouterB> system-view
[RouterB] ospf 1
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 2.2.2.2 0.0.0.0
[RouterB-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
[RouterB-ospf-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] ospf 1
[RouterC-ospf-1] import-route direct
[RouterC-ospf-1] area 0
[RouterC-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] quit
[RouterC-ospf-1] quit
```

(3) 配置 EBGP 连接

配置 EBGP 连接，并在 Router A 上将 8.1.1.0/24 网段通告到 BGP 路由表中，以便 Router B 获取到网段 8.1.1.0/24 的路由。

# 配置 Router A。

```
<RouterA> system-view
[RouterA] bgp 65008
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 3.1.1.1 as-number 65009
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 3.1.1.1 enable
[RouterA-bgp-ipv4] network 8.1.1.0 24
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
[RouterB] bgp 65009
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 3.1.1.2 as-number 65008
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] peer 3.1.1.2 enable
```

(4) 配置 BGP 与 IGP 交互

- 在 Router B 上配置 BGP 引入 OSPF 路由，以便 Router A 能够获取到到 9.1.2.0/24 网段的路由。
- 在 Router B 上配置 OSPF 引入 BGP 路由，以便 Router C 能够获取到到 8.1.1.0/24 网段的路由。

# 在 Router B 上配置 BGP 和 OSPF 互相引入路由。

```
[RouterB-bgp-ipv4] import-route ospf 1
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
[RouterB] ospf 1
[RouterB-ospf-1] import-route bgp
[RouterB-ospf-1] quit
```

# 查看 Router A 的 BGP 路由表。

```
[RouterA] display bgp routing-table ipv4

 Total number of routes: 3

 BGP local router ID is 1.1.1.1
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >e 3.3.3.3/32         3.1.1.1         1                     0       65009?
* >  8.1.1.0/24         8.1.1.1         0                     32768   i
* >e 9.1.2.0/24         3.1.1.1         1                     0       65009?
```

# 查看 Router C 的 OSPF 路由表。

```
[RouterC] display ospf routing

          OSPF Process 1 with Router ID 3.3.3.3
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter       Area
 9.1.1.0/24         1        Transit 9.1.1.2         3.3.3.3         0.0.0.0
 2.2.2.2/32         1        Stub    9.1.1.1         2.2.2.2         0.0.0.0

 Routing for ASEs
 Destination        Cost     Type    Tag         NextHop         AdvRouter
 8.1.1.0/24         1        Type2   1           9.1.1.1         2.2.2.2
```

```
 Total Nets: 3
 Intra Area: 2  Inter Area: 0  ASE: 1  NSSA: 0
```

#### 4. 验证配置

# 使用 Ping 进行验证。

```
[RouterA] ping -a 8.1.1.1 9.1.2.1
Ping 9.1.2.1 (9.1.2.1) from 8.1.1.1: 56 data bytes, press CTRL_C to break
56 bytes from 9.1.2.1: icmp_seq=0 ttl=254 time=10.000 ms
56 bytes from 9.1.2.1: icmp_seq=1 ttl=254 time=12.000 ms
56 bytes from 9.1.2.1: icmp_seq=2 ttl=254 time=2.000 ms
56 bytes from 9.1.2.1: icmp_seq=3 ttl=254 time=7.000 ms
56 bytes from 9.1.2.1: icmp_seq=4 ttl=254 time=9.000 ms

--- Ping statistics for 9.1.2.1 ---
5 packet(s) transmitted, 5 packet(s) received, 0.0% packet loss
round-trip min/avg/max/std-dev = 2.000/8.000/12.000/3.406 ms
[RouterC] ping -a 9.1.2.1 8.1.1.1
Ping 8.1.1.1 (8.1.1.1) from 9.1.2.1: 56 data bytes, press CTRL_C to break
56 bytes from 8.1.1.1: icmp_seq=0 ttl=254 time=9.000 ms
56 bytes from 8.1.1.1: icmp_seq=1 ttl=254 time=4.000 ms
56 bytes from 8.1.1.1: icmp_seq=2 ttl=254 time=3.000 ms
56 bytes from 8.1.1.1: icmp_seq=3 ttl=254 time=3.000 ms
56 bytes from 8.1.1.1: icmp_seq=4 ttl=254 time=3.000 ms

--- Ping statistics for 8.1.1.1 ---
5 packet(s) transmitted, 5 packet(s) received, 0.0% packet loss
round-trip min/avg/max/std-dev = 3.000/4.400/9.000/2.332 ms
```

### 1.17.3 BGP路由聚合配置

#### 1. 组网需求

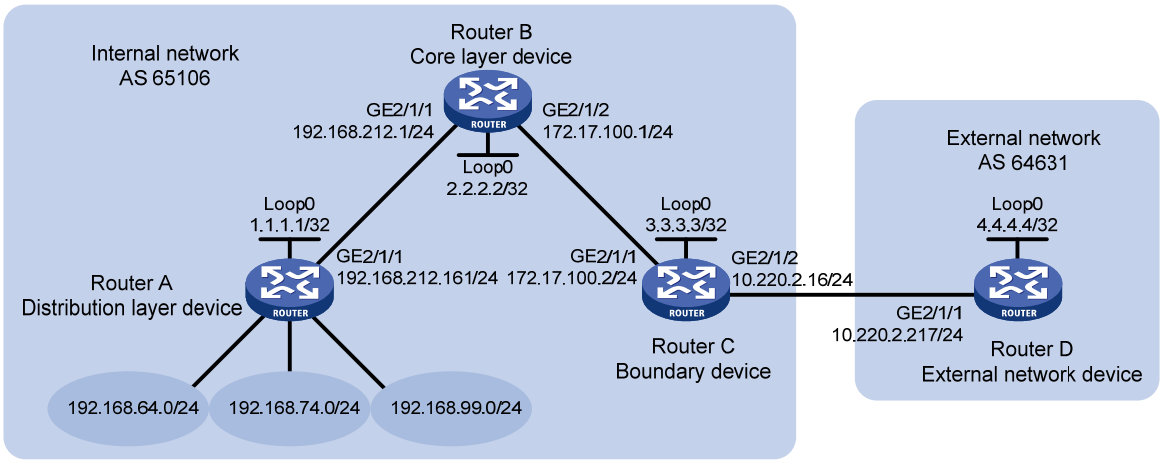
通过在边界设备 Router C 和外部网络设备 Router D 之间建立 EBGP 连接，实现公司内部网络与外部网络的互通。

在公司内部，核心层设备 Router B 与汇聚层设备 Router A 之间配置静态路由，Router B 与 Router C 之间配置 OSPF，并在 OSPF 路由中引入静态路由，以实现公司内部网络的互通。

公司内部网络包括三个网段：192.168.64.0/24、192.168.74.0/24 和 192.168.99.0/24。在 Router C 上配置路由聚合，将这三个网段的路由聚合为一条路由，以减少通过 BGP 发布的路由数量。

## 2. 组网图

图1-18 BGP 路由聚合组网图




## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 在 Router A 和 Router B 之间配置静态路由

# 在 Router A 上配置缺省路由，下一跳为 Router B。

```
<RouterA> system-view
[RouterA] ip route-static 0.0.0.0 0 192.168.212.1
```

# 在 Router B 上配置静态路由，到达目的网络 192.168.64.0/24、192.168.74.0/24 和 192.168.99.0/24 的路由下一跳均为 Router A。

```
<RouterB> system-view
[RouterB] ip route-static 192.168.64.0 24 192.168.212.161
[RouterB] ip route-static 192.168.74.0 24 192.168.212.161
[RouterB] ip route-static 192.168.99.0 24 192.168.212.161
```

(3) 在 Router B 和 Router C 之间配置 OSPF，并引入静态路由

# 在 Router B 上配置 OSPF 发布本地网段路由，并引入静态路由。

```
[RouterB] ospf
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 172.17.100.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
[RouterB-ospf-1] import-route static
[RouterB-ospf-1] quit
```

# 在 Router C 上配置 OSPF 发布本地网段路由。

```
[RouterC] ospf
[RouterC-ospf-1] area 0
[RouterC-ospf-1-area-0.0.0.0] network 172.17.100.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] network 10.220.2.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] quit
[RouterC-ospf-1] quit
```

# 在 Router C 上查看路由表信息，可以看到 Router C 通过 OSPF 学习到了到达 192.168.64.0/24、192.168.74.0/24 和 192.168.99.0/24 网段的路由。

```
[RouterC] display ip routing-table protocol ospf

Summary Count : 5

OSPF Routing table Status : <Active>
Summary Count : 3

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
192.168.64.0/24     OSPF   150  1            172.17.100.1    GE2/1/1
192.168.74.0/24     OSPF   150  1            172.17.100.1    GE2/1/1
192.168.99.0/24     OSPF   150  1            172.17.100.1    GE2/1/1

OSPF Routing table Status : <Inactive>
Summary Count : 2

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
10.220.2.0/24       OSPF   10   1            10.220.2.16     GE2/1/2
172.17.100.0/24     OSPF   10   1            172.17.100.2    GE2/1/1
```

(4) 在 Router C 和 Router D 之间配置 BGP，并引入 OSPF 路由

# 在 Router C 上配置 Router D 为其 EBGP 对等体，并引入 OSPF 路由。

```
[RouterC] bgp 65106
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 10.220.2.217 as-number 64631
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 10.220.2.217 enable
[RouterC-bgp-ipv4] import-route ospf
```

# 在 Router D 上配置 Router C 为其 EBGP 对等体。

```
[RouterD] bgp 64631
[RouterD-bgp] router-id 4.4.4.4
[RouterD-bgp] peer 10.220.2.16 as-number 65106
[RouterD-bgp] address-family ipv4 unicast
[RouterD-bgp-ipv4] peer 10.220.2.16 enable
[RouterD-bgp-ipv4] quit
[RouterD-bgp] quit
```

# 在 Router D 上查看路由表信息，可以看到 Router D 通过 BGP 学习到了到达 192.168.64.0/24、192.168.74.0/24 和 192.168.99.0/24 三个网段的路由。

```
[RouterD] display ip routing-table protocol bgp

Summary Count : 3

BGP Routing table Status : <Active>
Summary Count : 3

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
192.168.64.0/24     BGP    255  1            10.220.2.16     GE2/1/1
192.168.74.0/24     BGP    255  1            10.220.2.16     GE2/1/1
192.168.99.0/24     BGP    255  1            10.220.2.16     GE2/1/1
```

```
BGP Routing table Status : <Inactive>
Summary Count : 0
```

完成上述配置后，在 Router D 上可以 ping 通 192.168.64.0/24、192.168.74.0/24 和 192.168.99.0/24 网段内的主机。

(5) 在 Router C 上配置路由聚合

# 在 Router C 上将路由 192.168.64.0/24、192.168.74.0/24 和 192.168.99.0/24 聚合为 192.168.64.0/18，并抑制发布具体路由。

```
[RouterC-bgp-ipv4] aggregate 192.168.64.0 18 detail-suppressed
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
```

#### 4. 验证配置

# 在 Router C 上查看路由表信息，可以看到 Router C 上产生了一条聚合路由 192.168.64.0/18，该聚合路由的出接口为 Null0。

```
[RouterC] display ip routing-table | include 192.168
192.168.64.0/18      BGP    130  0            127.0.0.1       NULL0
192.168.64.0/24      OSPF   150  1            172.17.100.1    GE2/1/1
192.168.74.0/24      OSPF   150  1            172.17.100.1    GE2/1/1
192.168.99.0/24      OSPF   150  1            172.17.100.1    GE2/1/1
```

# 在 Router D 上查看路由表信息，可以看到 Router D 上到达公司内部三个网络的路由聚合为一条路由 192.168.64.0/18。

```
[RouterD] display ip routing-table protocol bgp

Summary Count : 1

BGP Routing table Status : <Active>
Summary Count : 1

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
192.168.64.0/18     BGP    255  0            10.220.2.16     GE2/1/1

BGP Routing table Status : <Inactive>
Summary Count : 0
```

完成上述配置后，成功实现了路由聚合。并且，在 Router D 上可以 ping 通 192.168.64.0/24、192.168.74.0/24 和 192.168.99.0/24 网段内的主机。

### 1.17.4 BGP负载分担配置

#### 1. 组网需求

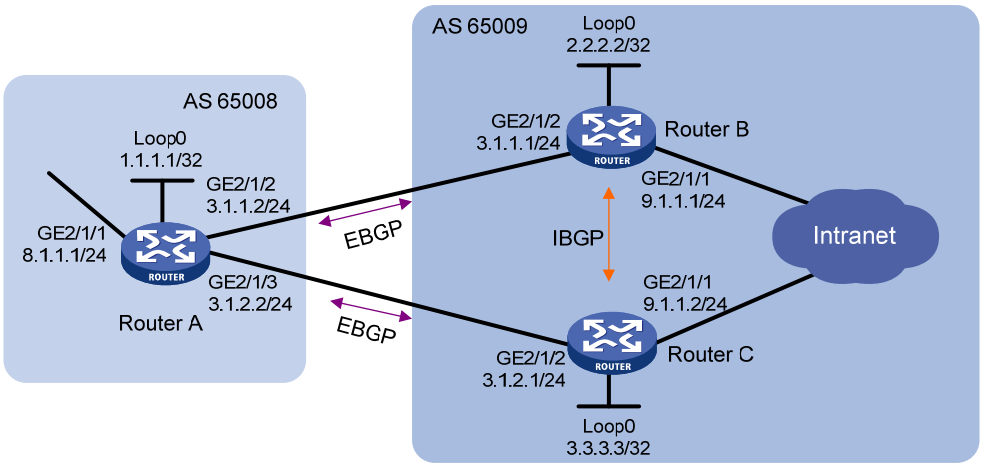
所有路由器都配置 BGP，Router A 在 AS 65008 中，Router B 和 Router C 在 AS 65009 中。

Router A 与 Router B、Router C 之间运行 EBGP，Router B 和 Router C 之间运行 IBGP。

在 Router A 上配置负载分担的路由条数为 2，以提高链路利用率。

## 2. 组网图

图1-19 BGP 负载分担配置组网图




## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 BGP 连接

- 在 Router A 上与 Router B、Router C 分别建立 EBGP 连接，并将 8.1.1.0/24 网段的路由通告给 Router B 和 Router C，以便 Router B 和 Router C 能够访问 Router A 的内部网络。
- 在 Router B 上与 Router A 建立 EBGP 连接，与 Router C 建立 IBGP 连接，并将 9.1.1.0/24 网段的路由通告给 Router A，以便 Router A 能够通过 Router B 访问内部网络。同时，在 Router B 上配置一条到 Router C Loopback0 接口的静态路由（也可以用 OSPF 等协议来实现），以便使用 Loopback 接口建立 IBGP 连接。
- 在 Router C 上与 Router A 建立 EBGP 连接，与 Router B 建立 IBGP 连接，并将 9.1.1.0/24 网段的路由通告给 Router A，以便 Router A 能够通过 Router C 访问内部网络。同时，在 Router C 上配置一条到 Router B Loopback0 接口的静态路由（也可以用 OSPF 等协议来实现），以便使用 Loopback 接口建立 IBGP 连接。

# 配置 Router A。

```
<RouterA> system-view
[RouterA] bgp 65008
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 3.1.1.1 as-number 65009
[RouterA-bgp] peer 3.1.2.1 as-number 65009
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 3.1.1.1 enable
[RouterA-bgp-ipv4] peer 3.1.2.1 enable
[RouterA-bgp-ipv4] network 8.1.1.0 24
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] bgp 65009
```

```
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 3.1.1.2 as-number 65008
[RouterB-bgp] peer 3.3.3.3 as-number 65009
[RouterB-bgp] peer 3.3.3.3 connect-interface loopback 0
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] peer 3.1.1.2 enable
[RouterB-bgp-ipv4] peer 3.3.3.3 enable
[RouterB-bgp-ipv4] network 9.1.1.0 24
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
[RouterB] ip route-static 3.3.3.3 32 9.1.1.2
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] bgp 65009
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 3.1.2.2 as-number 65008
[RouterC-bgp] peer 2.2.2.2 as-number 65009
[RouterC-bgp] peer 2.2.2.2 connect-interface loopback 0
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 3.1.2.2 enable
[RouterC-bgp-ipv4] peer 2.2.2.2 enable
[RouterC-bgp-ipv4] network 9.1.1.0 24
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
[RouterC] ip route-static 2.2.2.2 32 9.1.1.1
```

# 查看 Router A 的路由表。

```
[RouterA] display bgp routing-table ipv4

 Total number of routes: 3

 BGP local router ID is 1.1.1.1
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >  8.1.1.0/24         8.1.1.1         0                     32768   i
* >e 9.1.1.0/24         3.1.1.1         0                     0       65009i
*  e                    3.1.2.1         0                     0       65009i
```

- 从 BGP 路由表中可以看出，到目的地址 9.1.1.0/24 有两条有效路由，其中下一跳为 3.1.1.1 的路由前有标志“>”，表示它是当前有效的最优路由（因为 Router B 的路由器 ID 要小一些）；而下一跳为 3.1.2.1 的路由前有标志“*”，表示它是当前有效的路由，但不是最优的。
- 使用 **display ip routing-table** 命令查看 IP 路由表项，可以看出到达目的地址 9.1.1.0/24 的路由只有一条，下一跳地址为 3.1.1.1，出接口为 GigabitEthernet2/1/2。

(3) 配置负载分担

因为 Router A 有两条路径到达 AS 65009 的内部网络，所以，可以在 Router A 配置负载分担的路由条数为 2，以提高链路利用率。

# 配置 Router A。

```
[RouterA] bgp 65008
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] balance 2
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

#### 4. 验证配置

# 查看 Router A 的 BGP 路由表。

```
[RouterA] display bgp routing-table ipv4

 Total number of routes: 3

 BGP local router ID is 1.1.1.1
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >  8.1.1.0/24         8.1.1.1         0                     32768   i
* >e 9.1.1.0/24         3.1.1.1         0                     0       65009i
* >e                    3.1.2.1         0                     0       65009i
```

- 从 BGP 路由表中可以看到，BGP 路由 9.1.1.0/24 存在两个下一跳，分别是 3.1.1.1 和 3.1.2.1，两条路由前都有标志“>”，表明它们都是当前有效的最优路由。
- 使用 **display ip routing-table** 命令查看 IP 路由表项，可以看出到达目的地址 9.1.1.0/24 的路由有两条，其中一条的下一跳地址为 3.1.1.1，出接口为 GigabitEthernet2/1/2；另一条的下一跳地址为 3.1.2.1，出接口为 GigabitEthernet2/1/3。

### 1.17.5 BGP团体配置

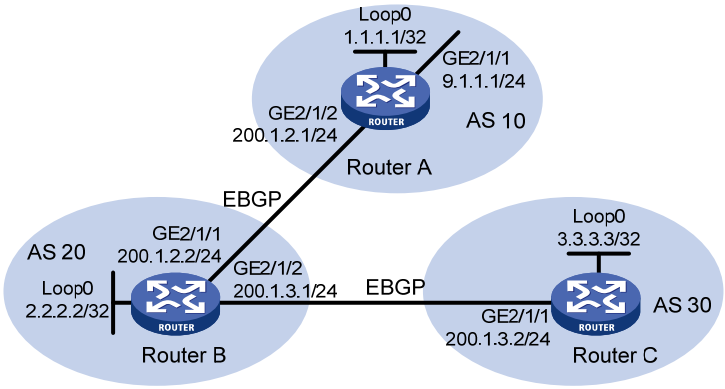
#### 1. 组网需求

Router B 分别与 Router A、Router C 之间建立 EBGP 连接。

通过在 Router A 上配置 NO_EXPORT 团体属性，使得 AS 10 发布到 AS 20 中的路由，不会再被 AS 20 发布到其他 AS。

## 2. 组网图

图1-20 BGP 团体组网图




## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 EBGP

# 配置 Router A。

```
<RouterA> system-view
[RouterA] bgp 10
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 200.1.2.2 as-number 20
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 200.1.2.2 enable
[RouterA-bgp-ipv4] network 9.1.1.0 255.255.255.0
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] bgp 20
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 200.1.2.1 as-number 10
[RouterB-bgp] peer 200.1.3.2 as-number 30
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] peer 200.1.2.1 enable
[RouterB-bgp-ipv4] peer 200.1.3.2 enable
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] bgp 30
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 200.1.3.1 as-number 20
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 200.1.3.1 enable
```

```
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
```
# 查看 Router B 的路由表。
```
[RouterB] display bgp routing-table ipv4 9.1.1.0

 BGP local router ID: 2.2.2.2
 Local AS number: 20

 Paths:   1 available, 1 best

 BGP routing table information of 9.1.1.0/24:
 From            : 200.1.2.1 (1.1.1.1)
 Rely nexthop    : 200.1.2.1
 Original nexthop: 200.1.2.1
 OutLabel        : NULL
 AS-path         : 10
 Origin          : igp
 Attribute value : pref-val 0
 State           : valid, external, best,
 IP precedence   : N/A
 QoS local ID    : N/A
```
# 查看 Router B 的路由发送信息。
```
[RouterB] display bgp routing-table ipv4 9.1.1.0 advertise-info

 BGP local router ID: 2.2.2.2
 Local AS number: 20

 Paths:   1 best

 BGP routing table information of 9.1.1.0/24:
 Advertised to peers (1 in total):
    200.1.3.2
```
可以看出，Router B 能够把到达目的地址 9.1.1.0/24 的路由通过 BGP 发布出去。

# 查看 Router C 的 BGP 路由表。
```
[RouterC] display bgp routing-table ipv4

 Total number of routes: 1

 BGP local router ID is 3.3.3.3
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >e 9.1.1.0/24         200.1.3.1                             0       20 10i
```
可以看出，Router C 从 Router B 那里学到了目的地址为 9.1.1.0/24 的路由。

(3) 配置 BGP 团体属性

# 配置路由策略。

```
[RouterA] route-policy comm_policy permit node 0
[RouterA-route-policy-comm_policy-0] apply community no-export
[RouterA-route-policy-comm_policy-0] quit
```

# 应用路由策略。

```
[RouterA] bgp 10
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 200.1.2.2 route-policy comm_policy export
[RouterA-bgp-ipv4] peer 200.1.2.2 advertise-community
```

### 4. 验证配置

# 查看 Router B 的路由表。

```
[RouterB] display bgp routing-table ipv4 9.1.1.0

 BGP local router ID: 2.2.2.2
 Local AS number: 20

 Paths:   1 available, 1 best

 BGP routing table information of 9.1.1.0/24:
 From            : 200.1.2.1 (1.1.1.1)
 Rely nexthop    : 200.1.2.1
 Original nexthop: 200.1.2.1
 OutLabel        : NULL
 Community       : No-Export
 AS-path         : 10
 Origin          : igp
 Attribute value : pref-val 0
 State           : valid, external, best,
 IP precedence   : N/A
 QoS local ID    : N/A
```

# 查看 Router B 的路由发送信息。

```
[RouterB] display bgp routing-table ipv4 9.1.1.0 advertise-info

 BGP local router ID: 2.2.2.2
 Local AS number: 20

 Paths:   1 best

 BGP routing table information of 9.1.1.0/24:
 Not advertised to any peers yet
```

# 查看 Router C 的 BGP 路由表。

```
[RouterC] display bgp routing-table ipv4

 Total number of routes: 0
```

在 Router B 的 BGP 路由表中可以看到配置的团体属性，Router B 不会通过 BGP 将到达目的地址 9.1.1.0/24 的路由发布出去。

## 1.17.6 BGP路由反射器配置

### 1. 组网需求

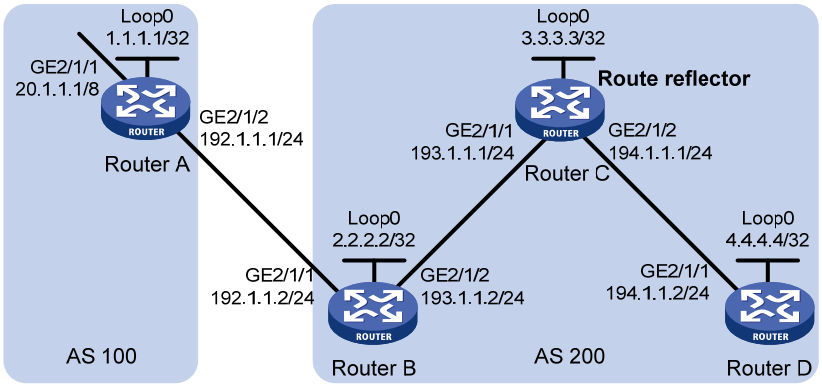
所有路由器运行 BGP 协议，Router A 与 Router B 建立 EBGP 连接，Router C 与 Router B 和 Router D 之间建立 IBGP 连接。

Router C 作为路由反射器，Router B 和 Router D 为 Router C 的客户机。

Router D 能够通过 Router C 学到路由 20.0.0.0/8。

### 2. 组网图

图1-21 配置 BGP 路由反射器的组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 BGP 连接

# 配置 Router A。

```
<RouterA> system-view
[RouterA] bgp 100
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 192.1.1.2 as-number 200
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 192.1.1.2 enable
```

# 通告 20.0.0.0/8 网段路由到 BGP 路由表中。

```
[RouterA-bgp-ipv4] network 20.0.0.0
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] bgp 200
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 192.1.1.1 as-number 100
```

```
[RouterB-bgp] peer 193.1.1.1 as-number 200
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] peer 192.1.1.1 enable
[RouterB-bgp-ipv4] peer 193.1.1.1 enable
[RouterB-bgp-ipv4] peer 193.1.1.1 next-hop-local
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] bgp 200
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 193.1.1.2 as-number 200
[RouterC-bgp] peer 194.1.1.2 as-number 200
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 193.1.1.2 enable
[RouterC-bgp-ipv4] peer 194.1.1.2 enable
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] bgp 200
[RouterD-bgp] router-id 4.4.4.4
[RouterD-bgp] peer 194.1.1.1 as-number 200
[RouterD-bgp] address-family ipv4 unicast
[RouterD-bgp-ipv4] peer 194.1.1.1 enable
[RouterD-bgp-ipv4] quit
[RouterD-bgp] quit
```

(3) 配置路由反射器

# 配置 Router C。

```
[RouterC] bgp 200
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 193.1.1.2 reflect-client
[RouterC-bgp-ipv4] peer 194.1.1.2 reflect-client
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
```

## 4. 验证配置

# 查看 Router B 的 BGP 路由表。

```
[RouterB] display bgp routing-table ipv4

 Total number of routes: 1

 BGP local router ID is 2.2.2.2
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn
```

```
* >e 20.0.0.0           192.1.1.1       0                    0       100i
```

# 查看 Router D 的 BGP 路由表。

```
[RouterD] display bgp routing-table ipv4

 Total number of routes: 1

 BGP local router ID is 4.4.4.4
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

   i 20.0.0.0           193.1.1.2       0          100        0       100i
```

可以看出，Router D 从 Router C 已经学到了 20.0.0.0/8 路由。

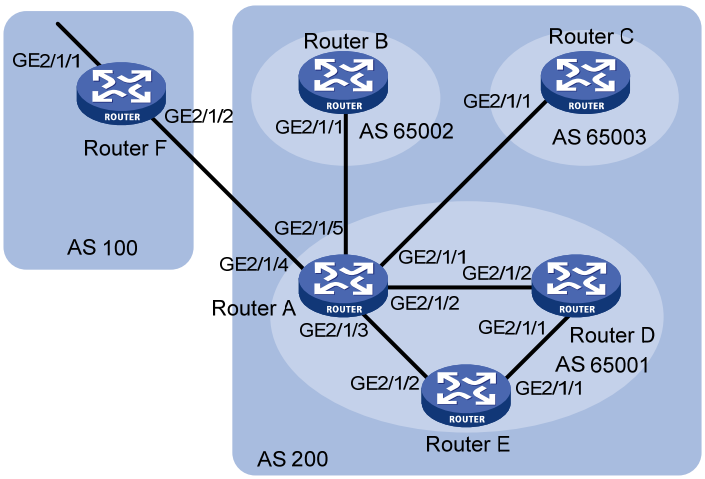
## 1.17.7 BGP联盟配置

### 1. 组网需求

AS 200 中有多台 BGP 路由器，为了减少 IBGP 的连接数，现将他们划分为 3 个子自治系统：AS 65001、AS 65002 和 AS 65003。其中 AS 65001 内的三台路由器建立 IBGP 全连接。

### 2. 组网图

图1-22 配置联盟组网图




| 设备 | 接口 | IP地址 | 设备 | 接口 | IP地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 10.1.2.1/24 | Router D | GE2/1/1 | 10.1.5.1/24 |
| | GE2/1/2 | 10.1.3.1/24 | | GE2/1/2 | 10.1.3.2/24 |
| | GE2/1/3 | 10.1.4.1/24 | Router E | GE2/1/1 | 10.1.5.2/24 |
| | GE2/1/4 | 200.1.1.1/24 | | GE2/1/2 | 10.1.4.2/24 |
| | GE2/1/5 | 10.1.1.1/24 | Router F | GE2/1/1 | 9.1.1.1/24 |
| Router B | GE2/1/1 | 10.1.1.2/24 | | GE2/1/2 | 200.1.1.2/24 |
| Router C | GE2/1/1 | 10.1.2.2/24 | | | |

### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 BGP 联盟

# 配置 Router A。

```
<RouterA> system-view
[RouterA] bgp 65001
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] confederation id 200
[RouterA-bgp] confederation peer-as 65002 65003
[RouterA-bgp] peer 10.1.1.2 as-number 65002
[RouterA-bgp] peer 10.1.2.2 as-number 65003
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 10.1.1.2 enable
[RouterA-bgp-ipv4] peer 10.1.2.2 enable
[RouterA-bgp-ipv4] peer 10.1.1.2 next-hop-local
[RouterA-bgp-ipv4] peer 10.1.2.2 next-hop-local
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] bgp 65002
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] confederation id 200
[RouterB-bgp] confederation peer-as 65001 65003
[RouterB-bgp] peer 10.1.1.1 as-number 65001
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] peer 10.1.1.1 enable
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] bgp 65003
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] confederation id 200
[RouterC-bgp] confederation peer-as 65001 65002
[RouterC-bgp] peer 10.1.2.1 as-number 65001
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 10.1.2.1 enable
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
```

(3) 配置 AS 65001 内的 IBGP 连接

# 配置 Router A。

```
[RouterA] bgp 65001
[RouterA-bgp] peer 10.1.3.2 as-number 65001
[RouterA-bgp] peer 10.1.4.2 as-number 65001
[RouterA-bgp] address-family ipv4 unicast
```

```
[RouterA-bgp-ipv4] peer 10.1.3.2 enable
[RouterA-bgp-ipv4] peer 10.1.4.2 enable
[RouterA-bgp-ipv4] peer 10.1.3.2 next-hop-local
[RouterA-bgp-ipv4] peer 10.1.4.2 next-hop-local
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] bgp 65001
[RouterD-bgp] router-id 4.4.4.4
[RouterD-bgp] confederation id 200
[RouterD-bgp] peer 10.1.3.1 as-number 65001
[RouterD-bgp] peer 10.1.5.2 as-number 65001
[RouterD-bgp] address-family ipv4 unicast
[RouterD-bgp-ipv4] peer 10.1.3.1 enable
[RouterD-bgp-ipv4] peer 10.1.5.2 enable
[RouterD-bgp-ipv4] quit
[RouterD-bgp] quit
```

# 配置 Router E。

```
<RouterE> system-view
[RouterE] bgp 65001
[RouterE-bgp] router-id 5.5.5.5
[RouterE-bgp] confederation id 200
[RouterE-bgp] peer 10.1.4.1 as-number 65001
[RouterE-bgp] peer 10.1.5.1 as-number 65001
[RouterE-bgp] address-family ipv4 unicast
[RouterE-bgp-ipv4] peer 10.1.4.1 enable
[RouterE-bgp-ipv4] peer 10.1.5.1 enable
[RouterE-bgp-ipv4] quit
[RouterE-bgp] quit
```

(4) 配置 AS 100 和 AS 200 之间的 EBGP 连接

# 配置 Router A。

```
[RouterA] bgp 65001
[RouterA-bgp] peer 200.1.1.2 as-number 100
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 200.1.1.2 enable
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router F。

```
<RouterF> system-view
[RouterF] bgp 100
[RouterF-bgp] router-id 6.6.6.6
[RouterF-bgp] peer 200.1.1.1 as-number 200
[RouterF-bgp] address-family ipv4 unicast
[RouterF-bgp-ipv4] peer 200.1.1.1 enable
[RouterF-bgp-ipv4] network 9.1.1.0 255.255.255.0
[RouterF-bgp-ipv4] quit
```

```
[RouterF-bgp] quit
```

### 4. 验证配置

# 查看 Router B 的 BGP 路由表。Router C 的 BGP 路由表与此类似。

```
[RouterB] display bgp routing-table ipv4

 Total number of routes: 1

 BGP local router ID is 2.2.2.2
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >i 9.1.1.0/24         10.1.1.1        0          100        0       (65001)
                                                                      100i
[RouterB] display bgp routing-table ipv4 9.1.1.0

 BGP local router ID: 2.2.2.2
 Local AS number: 65002

 Paths:   1 available, 1 best

 BGP routing table information of 9.1.1.0/24:
 From            : 10.1.1.1 (1.1.1.1)
 Rely nexthop    : 10.1.1.1
 Original nexthop: 10.1.1.1
 OutLabel        : NULL
 AS-path         : (65001) 100
 Origin          : igp
 Attribute value : MED 0, localpref 100, pref-val 0, pre 255
 State           : valid, external-confed, best,
 IP precedence   : N/A
 QoS local ID    : N/A
```

# 查看 Router D 的 BGP 路由表。

```
[RouterD] display bgp routing-table ipv4

 Total number of routes: 1

 BGP local router ID is 4.4.4.4
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >i 9.1.1.0/24         10.1.3.1        0          100        0       100i
```

```
[RouterD] display bgp routing-table ipv4 9.1.1.0

 BGP local router ID: 4.4.4.4
 Local AS number: 65001

 Paths:   1 available, 1 best

 BGP routing table information of 9.1.1.0/24:
 From            : 10.1.3.1 (1.1.1.1)
 Rely nexthop    : 10.1.3.1
 Original nexthop: 10.1.3.1
 OutLabel        : NULL
 AS-path         : 100
 Origin          : igp
 Attribute value : MED 0, localpref 100, pref-val 0, pre 255
 State           : valid, internal-confed, best,
 IP precedence   : N/A
 QoS local ID    : N/A
```

通过以上显示信息可以看出：

- Router F 只需要和 Router A 建立 EBGP 连接，而不需要和 Router B、Router C 建立连接，同样可以通过联盟将路由信息传递给 Router B 和 Router C。
- Router B 和 Router D 在同一个联盟里，但是属于不同的子自治系统，它们都是通过 Router A 来获取外部路由信息，生成的 BGP 路由表项也是一致的，等效于在同一个自治系统内，但是又不需要物理上全连接。

## 1.17.8 BGP路径选择配置

### 1. 组网需求

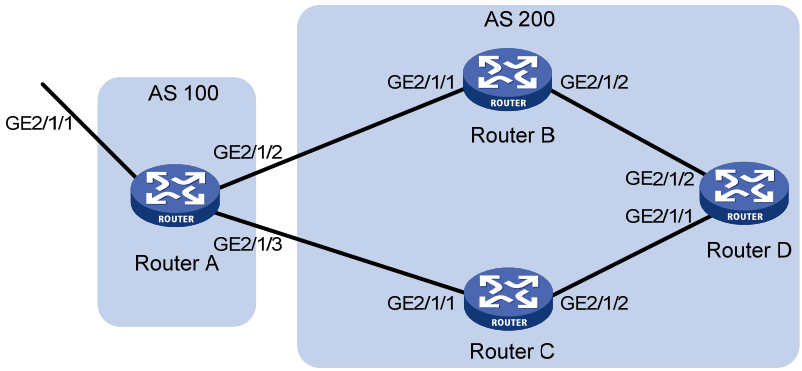
所有路由器都运行 BGP 协议。Router A 与 Router B 和 Router C 之间运行 EBGP；Router D 与 Router B 和 Router C 之间运行 IBGP。

AS 200 中运行 OSPF 协议。

配置路由策略，使得 Router D 优选从 Router C 学到的 1.0.0.0/8 路由。

### 2. 组网图

图1-23 配置 BGP 路径选择的组网图




| 设备 | 接口 | IP地址 | 设备 | 接口 | IP地址 |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 1.0.0.1/8 | Router D | GE2/1/1 | 195.1.1.1/24 |
| | GE2/1/2 | 192.1.1.1/24 | | GE2/1/2 | 194.1.1.1/24 |
| | GE2/1/3 | 193.1.1.1/24 | Router C | GE2/1/1 | 193.1.1.2/24 |
| Router B | GE2/1/1 | 192.1.1.2/24 | | GE2/1/2 | 195.1.1.2/24 |
| | GE2/1/2 | 194.1.1.2/24 | | | |

### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 Router B、Router C 和 Router D 之间运行 OSPF 协议

# 配置 Router B。

```
<RouterB> system-view
[RouterB] ospf
[RouterB-ospf] area 0
[RouterB-ospf-1-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
[RouterB-ospf-1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] ospf
[RouterC-ospf] area 0
[RouterC-ospf-1-area-0.0.0.0] network 193.1.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[RouterC-ospf-1-area-0.0.0.0] quit
[RouterC-ospf-1] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] ospf
[RouterD-ospf] area 0
[RouterD-ospf-1-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[RouterD-ospf-1-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[RouterD-ospf-1-area-0.0.0.0] quit
[RouterD-ospf-1] quit
```

(3) 配置 BGP 连接

# 配置 Router A。

```
<RouterA> system-view
[RouterA] bgp 100
[RouterA-bgp] peer 192.1.1.2 as-number 200
[RouterA-bgp] peer 193.1.1.2 as-number 200
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 192.1.1.2 enable
[RouterA-bgp-ipv4] peer 193.1.1.2 enable
```

# 将 1.0.0.0/8 网段通告到 Router A 的 BGP 路由表中。

```
[RouterA-bgp-ipv4] network 1.0.0.0 8
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
[RouterB] bgp 200
[RouterB-bgp] peer 192.1.1.1 as-number 100
[RouterB-bgp] peer 194.1.1.1 as-number 200
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] peer 192.1.1.1 enable
[RouterB-bgp-ipv4] peer 194.1.1.1 enable
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
```

# 配置 Router C。

```
[RouterC] bgp 200
[RouterC-bgp] peer 193.1.1.1 as-number 100
[RouterC-bgp] peer 195.1.1.1 as-number 200
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 193.1.1.1 enable
[RouterC-bgp-ipv4] peer 195.1.1.1 enable
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
```

# 配置 Router D。

```
[RouterD] bgp 200
[RouterD-bgp] peer 194.1.1.2 as-number 200
[RouterD-bgp] peer 195.1.1.2 as-number 200
[RouterD-bgp] address-family ipv4 unicast
[RouterD-bgp-ipv4] peer 194.1.1.2 enable
[RouterD-bgp-ipv4] peer 195.1.1.2 enable
[RouterD-bgp-ipv4] quit
[RouterD-bgp] quit
```

(4) 通过配置 1.0.0.0/8 路由的不同属性值，使得 Router D 优选从 Router C 学到的路由。

- 方法一：在 Router A 上对发布给对等体 192.1.1.2 的 1.0.0.0/8 路由配置较高的 MED 属性值，使得 Router D 优选从 Router C 学到的路由。

# 定义编号为 2000 的 ACL，允许路由 1.0.0.0/8 通过。

```
[RouterA] acl number 2000
[RouterA-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[RouterA-acl-basic-2000] quit
```

# 定义两个 Route-policy，一个名为 apply_med_50，为路由 1.0.0.0/8 设置 MED 属性值为 50；另一个名为 apply_med_100，为路由 1.0.0.0/8 设置 MED 属性值为 100。

```
[RouterA] route-policy apply_med_50 permit node 10
[RouterA-route-policy-apply_med_50-10] if-match ip address acl 2000
[RouterA-route-policy-apply_med_50-10] apply cost 50
[RouterA-route-policy-apply_med_50-10] quit
[RouterA] route-policy apply_med_100 permit node 10
[RouterA-route-policy-apply_med_100-10] if-match ip address acl 2000
[RouterA-route-policy-apply_med_100-10] apply cost 100
[RouterA-route-policy-apply_med_100-10] quit
```

# 对发布给对等体 193.1.1.2（Router C）的路由应用名为 apply_med_50 的 Route-policy，对发布给对等体 192.1.1.2（Router B）的路由应用名为 apply_med_100 的 Route-policy。

```
[RouterA] bgp 100
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 193.1.1.2 route-policy apply_med_50 export
[RouterA-bgp-ipv4] peer 192.1.1.2 route-policy apply_med_100 export
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 查看 Router D 的 BGP 路由表。

```
[RouterD] display bgp routing-table ipv4

 Total number of routes: 2

 BGP local router ID is 195.1.1.1
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >i 1.0.0.0            193.1.1.1       50         100        0       100i
*  i                    192.1.1.1       100        100        0       100i
```

可以看到，Router D 从 Router C 学到 1.0.0.0/8 的路由是最优的。

- 方法二：在 Router B 和 Router C 上分别对 1.0.0.0/8 路由配置不同的本地优先级，使得 Router D 优选从 Router C 学到的路由。

# 在 Router C 上定义编号为 2000 的 ACL，允许 1.0.0.0/8 路由通过。

```
[RouterC] acl number 2000
[RouterC-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[RouterC-acl-basic-2000] quit
```

# 在 Router C 上定义名为 localpref 的 Route-policy，设置路由 1.0.0.0/8 的本地优先级为 200（缺省的本地优先级为 100）。

```
[RouterC] route-policy localpref permit node 10
[RouterC-route-policy-localpref-10] if-match ip address acl 2000
[RouterC-route-policy-localpref-10] apply local-preference 200
[RouterC-route-policy-localpref-10] quit
```

# 为从 BGP 对等体 193.1.1.1 的路由应用名为 localpref 的 Route-policy。

```
[RouterC] bgp 200
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 193.1.1.1 route-policy localpref import
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
```

# 查看 Router D 的 BGP 路由表。

```
[RouterD] display bgp routing-table ipv4

 Total number of routes: 2

 BGP local router ID is 195.1.1.1
 Status codes: * - valid, > - best, d - dampened, h - history,
```

```
                s - suppressed, S - stale, i - internal, e - external
                Origin: i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

* >i 1.0.0.0            193.1.1.1                  200        0       100i
*  i                    192.1.1.1                  100        0       100i
```

可以看到，Router D 从 Router C 学到 1.0.0.0/8 的路由是最优的。

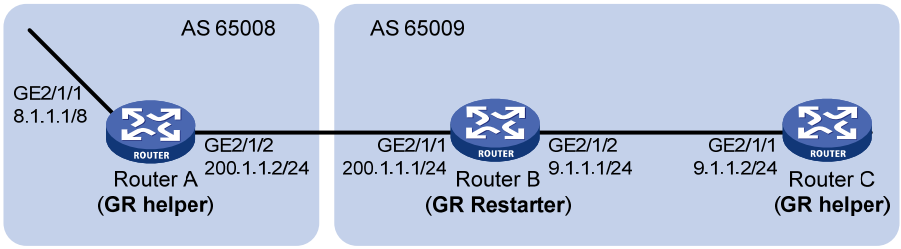
## 1.17.9 BGP GR配置

### 1. 组网需求

如 图 1-24 所示，所有路由器均运行BGP协议，Router A和Router B之间建立EBGP连接，Router B和Router C之间建立IBGP连接。现要求实现即便Router B发生主备倒换，也不会影响Router A和Router C之间正在进行的数据传输。

### 2. 组网图

图1-24 BGP GR 配置组网图




### 3. 配置步骤

(1) Router A 的配置

# 配置各接口的 IP 地址（略）。

# 配置 Router A 与 Router B 的 EBGP 连接。

```
<RouterA> system-view
[RouterA] bgp 65008
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 200.1.1.1 as-number 65009
```

# 使能 BGP GR 功能。

```
[RouterA-bgp] graceful-restart
```

# 将 8.0.0.0/8 网段路由通告到 IPv4 BGP 路由表中。

```
[RouterA-bgp] address-family ipv4
[RouterA-bgp-ipv4] network 8.0.0.0
```

# 使能与 Router B 交换 BGP IPv4 单播路由的能力。

```
[RouterA-bgp-ipv4] peer 200.1.1.1 enable
```

(2) Router B 的配置

# 配置各接口的 IP 地址（略）。

# 配置 Router B 与 Router A 的 EBGP 连接。

```
<RouterB> system-view
[RouterB] bgp 65009
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 200.1.1.2 as-number 65008
```

# 配置 Router B 与 Router C 的 IBGP 连接。

```
[RouterB-bgp] peer 9.1.1.2 as-number 65009
```

# 使能 BGP GR 功能。

```
[RouterB-bgp] graceful-restart
```

# 将 200.1.1.0/24 和 9.1.1.0/24 网段路由通告到 IPv4 BGP 路由表中。

```
[RouterB-bgp] address-family ipv4
[RouterB-bgp-ipv4] network 200.1.1.0 24
[RouterB-bgp-ipv4] network 9.1.1.0 24
```

# 使能与 Router A、Router C 交换 BGP IPv4 单播路由的能力。

```
[RouterB-bgp-ipv4] peer 200.1.1.2 enable
[RouterB-bgp-ipv4] peer 9.1.1.2 enable
```

(3) Router C 的配置

# 配置各接口的 IP 地址（略）。

# 配置 Router C 与 Router B 的 IBGP 连接。

```
<RouterC> system-view
[RouterC] bgp 65009
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 9.1.1.1 as-number 65009
```

# 使能 BGP GR 功能。

```
[RouterC-bgp] graceful-restart
```

# 使能与 Router B 交换 BGP IPv4 单播路由的能力。

```
[RouterC-bgp] address-family ipv4
[RouterC-bgp-ipv4] peer 9.1.1.1 enable
```

#### 4. 验证配置

在 Router A 上 ping Router C，同时在 Router B 上触发主备倒换，可以发现在整个倒换过程中 Router A 都可以 ping 通 Router C。

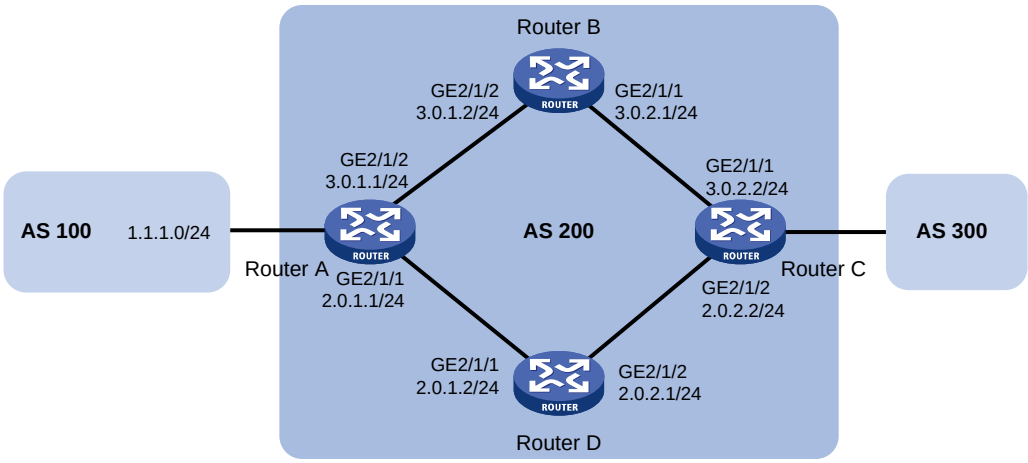
### 1.17.10 BGP与BFD联动配置

#### 1. 组网需求

- 在 AS 200 内使用 OSPF 作为 IGP 协议，实现 AS 内的互通。
- Router A 与 Router C 之间建立两条 IBGP 连接。当 Router A 与 Router C 之间的两条路径均连通时，Router C 与 1.1.1.0/24 之间的报文使用 Router A<－>Router B<－>Router C 这条路径转发；当 Router A<－>Router B<－>Router C 这条路径发生故障时，BFD 能够快速检测并通告 BGP 协议，使得 Router A<－>Router D<－>Router C 这条路径能够迅速生效。

## 2. 组网图

图1-25 配置 BGP 与 BFD 联动组网图




## 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 OSPF，保证 Router A 和 Router C 之间路由可达（略）

(3) Router A 上的 BGP 配置

# 配置 Router A 和 Router C 建立两条 IBGP 连接。

```
<RouterA> system-view
[RouterA] bgp 200
[RouterA-bgp] peer 3.0.2.2 as-number 200
[RouterA-bgp] peer 2.0.2.2 as-number 200
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 3.0.2.2 enable
[RouterA-bgp-ipv4] peer 2.0.2.2 enable
[RouterA-bgp-ipv4] quit
```

# 配置当 Router A 与 Router C 之间的两条路径均连通时，Router C 与 1.1.1.0/24 之间的报文使用 Router A<—>Router B<—>Router C 这条路径转发。（在 Router A 上对发布给对等体 2.0.2.2 的 1.1.1.0/24 路由配置较高的 MED 属性值）

- 定义编号为 2000 的 ACL，允许路由 1.1.1.0/24 通过。

```
[RouterA] acl number 2000
[RouterA-acl-basic-2000] rule permit source 1.1.1.0 0.0.0.255
[RouterA-acl-basic-2000] quit
```

- 定义两个 Route-policy，一个名为 apply_med_50，为路由 1.1.1.0/24 设置 MED 属性值为 50；另一个名为 apply_med_100，为路由 1.1.1.0/24 设置 MED 属性值为 100。

```
[RouterA] route-policy apply_med_50 permit node 10
[RouterA-route-policy-apply_med_50-10] if-match ip address acl 2000
[RouterA-route-policy-apply_med_50-10] apply cost 50
[RouterA-route-policy-apply_med_50-10] quit
[RouterA] route-policy apply_med_100 permit node 10
[RouterA-route-policy-apply_med_100-10] if-match ip address acl 2000
[RouterA-route-policy-apply_med_100-10] apply cost 100
```

```
[RouterA-route-policy-apply_med_100-10] quit
```

- 对发布给对等体 3.0.2.2 的路由应用名为 apply_med_50 的 Route-policy，对发布给对等体 2.0.2.2 的路由应用名为 apply_med_100 的 Route-policy。

```
[RouterA] bgp 200
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 3.0.2.2 route-policy apply_med_50 export
[RouterA-bgp-ipv4] peer 2.0.2.2 route-policy apply_med_100 export
[RouterA-bgp-ipv4] quit
```

# 配置当Router A<—>Router B<—>Router C这条路径发生故障时，BFD能够快速检测并通告BGP协议，使得 Router A<—>Router D<—>Router C 这条路径能够迅速生效。

```
[RouterA-bgp] peer 3.0.2.2 bfd
[RouterA-bgp] quit
```

(4) Router C 上的 BGP 配置。

# 配置 Router A 和 Router C 建立两条 IBGP 连接。

```
<RouterC> system-view
[RouterC] bgp 200
[RouterC-bgp] peer 3.0.1.1 as-number 200
[RouterC-bgp] peer 2.0.1.1 as-number 200
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 3.0.1.1 enable
[RouterC-bgp-ipv4] peer 2.0.1.1 enable
[RouterA-bgp-ipv4] quit
```

# 配置当Router A<—>Router B<—>Router C这条路径发生故障时，BFD能够快速检测并通告BGP协议，使得 Router A<—>Router D<—>Router C 这条路径能够迅速生效。

```
[RouterC-bgp] peer 3.0.1.1 bfd
[RouterC-bgp] quit
[RouterC] quit
```

### 4. 验证配置

下面以 Router C 为例，Router A 和 Router C 类似，不再赘述。

# 显示 Router C 的 BFD 信息。

```
<RouterC> display bfd session verbose

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv4 Session Working Under Ctrl Mode:

       Local Discr: 513                 Remote Discr: 513
         Source IP: 3.0.2.2          Destination IP: 3.0.1.1
     Session State: Up                     Interface: N/A
      Min Tx Inter: 500ms                Act Tx Inter: 500ms
      Min Rx Inter: 500ms                Detect Inter: 2500ms
          Rx Count: 135                      Tx Count: 135
      Connect Type: Indirect           Running Up for: 00:00:58
         Hold Time: 2457ms                   Auth mode: None
       Detect Mode: Async                         Slot: 0
          Protocol: BGP
```

```
        Diag Info: No Diagnostic
```

以上显示信息表明：Router A 和 Router C 之间已经建立了 BFD 连接，而且 BFD 协议运行正常。

# 在 Router C 上查看 BGP 邻居信息，可以看出 Router A 和 Router C 之间建立两条 BGP 连接，且均处于 Established 状态。

```
<RouterC> display bgp peer ipv4

 BGP local router ID: 3.3.3.3
 Local AS number: 200
 Total number of peers: 2                 Peers in established state: 2

  Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State

  2.0.1.1                200        4        5    0       0 00:01:55 Established
  3.0.1.1                200        4        5    0       0 00:01:52 Established
```

# 在 Router C 上查看 1.1.1.0/24 的路由信息，可以看出 Router C 通过 Router A<—>Router B<—>Router C 这条路径与 1.1.1.0/24 网段通信。

```
<RouterC> display ip routing-table 1.1.1.0 24 verbose

Summary Count : 1

Destination: 1.1.1.0/24
   Protocol: BGP             Process ID: 0
  SubProtID: 0x1                    Age: 00h00m09s
       Cost: 50             Preference: 255
        Tag: 0                    State: Active Adv
  OrigTblID: 0x1               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
      NBRID: 0x15000001         LastAs: 0
     AttrID: 0x1               Neighbor: 3.0.1.1
      Flags: 0x10060        OrigNextHop: 3.0.1.1
      Label: NULL           RealNextHop: 3.0.2.1
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid        BkInterface: N/A
```

# Router A 和 Router B 之间的链路发生故障后，在 Router C 上查看 1.1.1.0/24 的路由信息，可以看出 Router C 通过 Router A<—>Router D<—>Router C 这条路径与 1.1.1.0/24 网段通信。

```
<RouterC> display ip routing-table 1.1.1.0 24 verbose

Summary Count : 1

Destination: 1.1.1.0/24
   Protocol: BGP             Process ID: 0
  SubProtID: 0x1                    Age: 00h03m08s
       Cost: 100            Preference: 255
        Tag: 0                    State: Active Adv
  OrigTblID: 0x1               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 0
```

```
      NBRID: 0x15000000          LastAs: 0
    AttrID: 0x0                Neighbor: 2.0.1.1
     Flags: 0x10060          OrigNextHop: 2.0.1.1
     Label: NULL             RealNextHop: 2.0.2.1
   BkLabel: NULL               BkNextHop: N/A
 Tunnel ID: Invalid            Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid         BkInterface: N/A
```

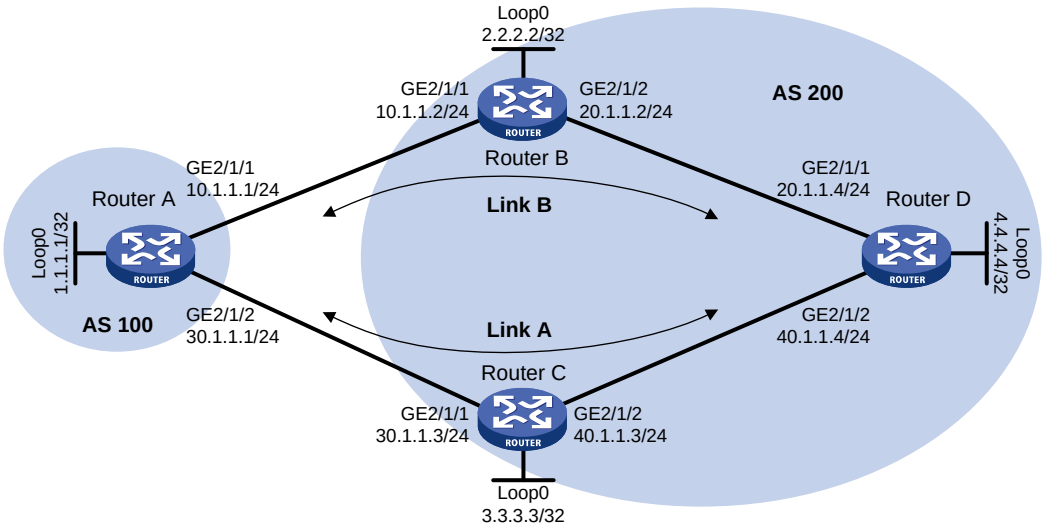
## 1.17.11 BGP快速重路由配置

### 1. 组网需求

如 图 1-26 所示，Router A、Router B、Router C和Router D通过BGP协议实现网络互连。要求链路B正常时，Router A和Router D之间的流量通过链路B转发；链路B出现故障时，流量可以快速切换到链路A上。

### 2. 组网图

图1-26 配置 BGP 快速重路由组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 在 AS 200 内配置 OSPF，发布接口地址所在网段的路由(包括 Loopback 接口)，确保 Router B、Router C 和 Router D 之间路由可达（略）

(3) 配置 BGP 连接

# 配置 Router A 分别与 Router B 和 Router C 建立 EBGP 会话，并配置通过 BGP 发布路由 1.1.1.1/32。

```
<RouterA> system-view
[RouterA] bgp 100
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 10.1.1.2 as-number 200
[RouterA-bgp] peer 30.1.1.3 as-number 200
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] peer 10.1.1.2 enable
```

```
[RouterA-bgp-ipv4] peer 30.1.1.3 enable
[RouterA-bgp-ipv4] network 1.1.1.1 32
```

# 配置 Router B 与 Router A 建立 EBGP 会话，与 Router D 建立 IBGP 会话。

```
<RouterB> system-view
[RouterB] bgp 200
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 10.1.1.1 as-number 100
[RouterB-bgp] peer 4.4.4.4 as-number 200
[RouterB-bgp] peer 4.4.4.4 connect-interface loopback 0
[RouterB-bgp] address-family ipv4 unicast
[RouterB-bgp-ipv4] peer 10.1.1.1 enable
[RouterB-bgp-ipv4] peer 4.4.4.4 enable
[RouterB-bgp-ipv4] peer 4.4.4.4 next-hop-local
[RouterB-bgp-ipv4] quit
[RouterB-bgp] quit
```

# 配置 Router C 与 Router A 建立 EBGP 会话，与 Router D 建立 IBGP 会话。

```
<RouterC> system-view
[RouterC] bgp 200
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 30.1.1.1 as-number 100
[RouterC-bgp] peer 4.4.4.4 as-number 200
[RouterC-bgp] peer 4.4.4.4 connect-interface loopback 0
[RouterC-bgp] address-family ipv4 unicast
[RouterC-bgp-ipv4] peer 30.1.1.1 enable
[RouterC-bgp-ipv4] peer 4.4.4.4 enable
[RouterC-bgp-ipv4] peer 4.4.4.4 next-hop-local
[RouterC-bgp-ipv4] quit
[RouterC-bgp] quit
```

# 配置 Router D 分别与 Router B 和 Router C 建立 IBGP 会话，并配置 BGP 发布路由 4.4.4.4/32。

```
<RouterD> system-view
[RouterD] bgp 200
[RouterD-bgp] router-id 4.4.4.4
[RouterD-bgp] peer 2.2.2.2 as-number 200
[RouterD-bgp] peer 2.2.2.2 connect-interface loopback 0
[RouterD-bgp] peer 3.3.3.3 as-number 200
[RouterD-bgp] peer 3.3.3.3 connect-interface loopback 0
[RouterD-bgp] address-family ipv4 unicast
[RouterD-bgp-ipv4] peer 2.2.2.2 enable
[RouterD-bgp-ipv4] peer 3.3.3.3 enable
[RouterD-bgp-ipv4] network 4.4.4.4 32
```

(4)　修改路由的首选值，使得 Router A 和 Router D 之间的流量优先通过链路 B 转发

# 在 Router A 上配置从 Router B 接收到的路由的首选值为 100。

```
[RouterA-bgp-ipv4] peer 10.1.1.2 preferred-value 100
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 在 Router D 上配置从 Router B 接收到的路由的首选值为 100。

```
[RouterD-bgp-ipv4] peer 2.2.2.2 preferred-value 100
```

```
[RouterD-bgp-ipv4] quit
[RouterD-bgp] quit
```

(5) 配置 BGP 快速重路由

# 配置 Router A：配置通过 Echo 方式的 BFD 会话检测主路由的下一跳是否可达，并配置 BFD echo 报文的源 IP 地址为 11.1.1.1；创建路由策略 frr，为路由 4.4.4.4/32 指定备份下一跳的地址为 30.1.1.3（对等体 Router C 的地址）；在 BGP IPv4 单播地址族下应用该路由策略。

```
[RouterA] bfd echo-source-ip 11.1.1.1
[RouterA] ip prefix-list abc index 10 permit 4.4.4.4 32
[RouterA] route-policy frr permit node 10
[RouterA-route-policy] if-match ip address prefix-list abc
[RouterA-route-policy] apply fast-reroute backup-nexthop 30.1.1.3
[RouterA-route-policy] quit
[RouterA] bgp 100
[RouterA-bgp] primary-path-detect bfd echo
[RouterA-bgp] address-family ipv4 unicast
[RouterA-bgp-ipv4] fast-reroute route-policy frr
[RouterA-bgp-ipv4] quit
[RouterA-bgp] quit
```

# 配置 Router D：配置通过 Echo 方式的 BFD 会话检测主路由的下一跳是否可达，并配置 BFD echo 报文的源 IP 地址为 44.1.1.1；创建路由策略 frr，为路由 1.1.1.1/32 指定备份下一跳的地址为 3.3.3.3（对等体 Router C 的地址）；在 BGP IPv4 单播地址族下应用该路由策略。

```
[RouterD] bfd echo-source-ip 44.1.1.1
[RouterD] ip prefix-list abc index 10 permit 1.1.1.1 32
[RouterD] route-policy frr permit node 10
[RouterD-route-policy] if-match ip address prefix-list abc
[RouterD-route-policy] apply fast-reroute backup-nexthop 3.3.3.3
[RouterD-route-policy] quit
[RouterD] bgp 200
[RouterD-bgp] primary-path-detect bfd echo
[RouterD-bgp] address-family ipv4 unicast
[RouterD-bgp-ipv4] fast-reroute route-policy frr
[RouterD-bgp-ipv4] quit
[RouterD-bgp] quit
```

## 4. 验证配置

# 在 Router A 上查看 4.4.4.4/32 路由，可以看到备份下一跳信息。

```
[RouterA] display ip routing-table 4.4.4.4 32 verbose

Summary Count : 1

Destination: 4.4.4.4/32
   Protocol: BGP             Process ID: 0
  SubProtID: 0x2                    Age: 00h01m52s
       Cost: 0              Preference: 255
      IpPre: N/A            QosLocalID: N/A
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
```

```
    TableID: 0x2                OrigAs: 200
      NibID: 0x15000003         LastAs: 200
     AttrID: 0x5              Neighbor: 10.1.1.2
      Flags: 0x10060       OrigNextHop: 10.1.1.2
      Label: NULL          RealNextHop: 10.1.1.2
    BkLabel: NULL            BkNextHop: 30.1.1.3
  Tunnel ID: Invalid         Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid       BkInterface: GigabitEthernet2/1/2
   FtnIndex: 0x0
```

# 在 Router D 上查看 1.1.1.1/32 路由，可以看到备份下一跳信息。

```
[RouterD] display ip routing-table 1.1.1.1 32 verbose

Summary Count : 1

Destination: 1.1.1.1/32
   Protocol: BGP             Process ID: 0
  SubProtID: 0x1                    Age: 00h00m36s
       Cost: 0               Preference: 255
      IpPre: N/A             QosLocalID: N/A
        Tag: 0                    State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0x2                OrigAs: 100
      NibID: 0x15000003         LastAs: 100
     AttrID: 0x1              Neighbor: 2.2.2.2
      Flags: 0x10060       OrigNextHop: 2.2.2.2
      Label: NULL          RealNextHop: 20.1.1.2
    BkLabel: NULL            BkNextHop: 40.1.1.3
  Tunnel ID: Invalid         Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid       BkInterface: GigabitEthernet2/1/2
   FtnIndex: 0x0
```

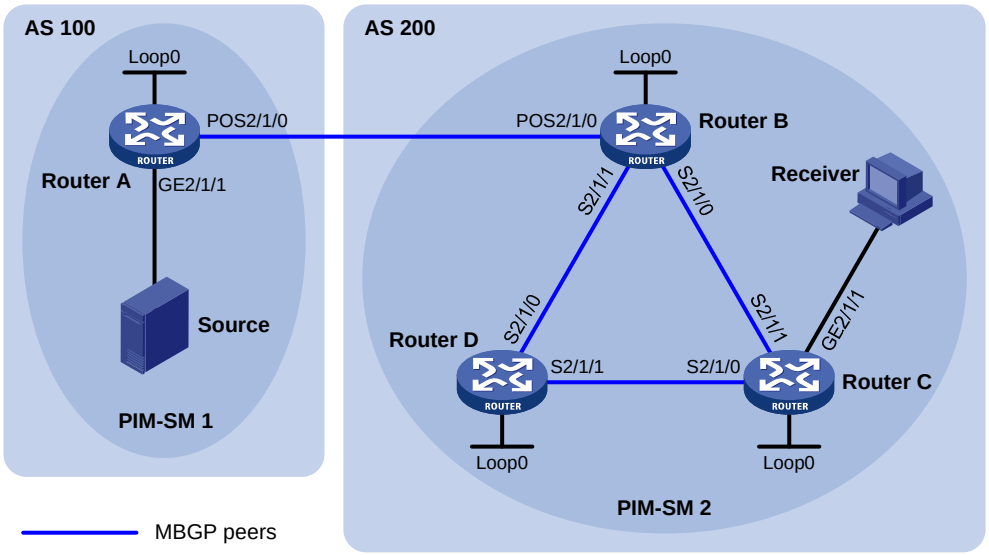
## 1.17.12 MBGP配置

### 1. 组网需求

- 网络中存在两个自治系统：PIM-SM 1 属于 AS 100，PIM-SM 2 属于 AS 200。各 AS 内部采用 OSPF 交换路由信息，AS 之间采用 MBGP 交换用于 RPF 检查的 IPv4 单播路由信息。
- 组播源属于 AS 100 内的 PIM-SM 1，接收者则属于 AS 200 内的 PIM-SM 2。
- 将 Router A 和 Router B 各自的 Loopback0 接口分别配置为各自 PIM-SM 域的 C-BSR 和 C-RP。
- 在 Router A 与 Router B 之间通过 MBGP 建立 MSDP 对等体关系。

## 2. 组网图

图1-27 MBGP 配置组网图




| 设备 | 接口 | IP地址 | 设备 | 接口 | IP地址 |
|---|---|---|---|---|---|
| Source | - | 10.110.1.100/24 | Router C | GE2/1/1 | 10.110.2.1/24 |
| Router A | GE2/1/1 | 10.110.1.1/24 | | S2/1/0 | 192.168.4.1/24 |
| | POS2/1/0 | 192.168.1.1/24 | | S2/1/1 | 192.168.2.2/24 |
| | Loop0 | 1.1.1.1/32 | | Loop0 | 3.3.3.3/32 |
| Router B | POS2/1/0 | 192.168.1.2/24 | Router D | S2/1/0 | 192.168.3.2/24 |
| | S2/1/0 | 192.168.2.1/24 | | S2/1/1 | 192.168.4.2/24 |
| | S2/1/1 | 192.168.3.1/24 | | Loop0 | 4.4.4.4/32 |
| | Loop0 | 2.2.2.2/32 | | | |

## 3. 配置步骤

(1) 配置各路由器接口的 IP 地址和单播路由协议

- 请按照 图 1-27 配置各接口的IP地址和子网掩码，具体配置过程略。
- 配置 AS 200 内的各路由器之间采用 OSPF 路由协议交换路由信息（AS 内各路由器使用的 OSPF 进程号为 1），确保各 AS 内部在网络层互通，能学到彼此 Loopback 接口的路由，具体配置过程略。

(2) 使能 IP 组播路由，使能 PIM-SM 和 IGMP，并配置 BSR 的服务边界

# 在 Router A 上使能 IP 组播路由，在各接口上使能 PIM-SM。

```
<RouterA> system-view
[RouterA] multicast routing
[RouterA-mrib] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] pim sm
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface pos 2/1/0
[RouterA-Pos2/1/0] pim sm
[RouterA-Pos2/1/0] quit
```

Router B 和 Router D 上的配置与 Router A 相似，配置过程略。

# 在 Router C 上使能 IP 组播路由，在各接口上使能 PIM-SM，并在主机侧接口 GigabitEthernet2/1/1 上使能 IGMP。

```
<RouterC> system-view
[RouterC] multicast routing
[RouterC-mrib] quit
[RouterC] interface serial 2/1/0
[RouterC-Serial2/1/0] pim sm
[RouterC-Serial2/1/0] quit
[RouterC] interface serial 2/1/1
[RouterC-Serial2/1/1] pim sm
[RouterC-Serial2/1/1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] pim sm
[RouterC-GigabitEthernet2/1/1] igmp enable
[RouterC-GigabitEthernet2/1/1] quit
```

# 在 Router A 上配置 BSR 的服务边界。

```
[RouterA] interface pos 2/1/0
[RouterA-Pos2/1/0] pim bsr-boundary
[RouterA-Pos2/1/0] quit
```

# 在 Router B 上配置 BSR 的服务边界。

```
[RouterB] interface pos 2/1/0
[RouterB-Pos2/1/0] pim bsr-boundary
[RouterB-Pos2/1/0] quit
```

(3) 配置 Loopback0 接口和 C-BSR、C-RP 的位置

# 在 Router A 上配置 Loopback0 接口和 C-BSR、C-RP 的位置。

```
[RouterA] interface loopback 0
[RouterA-LoopBack0] ip address 1.1.1.1 32
[RouterA-LoopBack0] pim sm
[RouterA-LoopBack0] quit
[RouterA] pim
[RouterA-pim] c-bsr 1.1.1.1
[RouterA-pim] c-rp 1.1.1.1
[RouterA-pim] quit
```

# 在 Router B 上配置 Loopback0 接口和 C-BSR、C-RP 的位置。

```
[RouterB] interface loopback 0
[RouterB-LoopBack0] ip address 2.2.2.2 32
[RouterB-LoopBack0] pim sm
[RouterB-LoopBack0] quit
[RouterB] pim
[RouterB-pim] c-bsr 2.2.2.2
[RouterB-pim] c-rp 2.2.2.2
[RouterB-pim] quit
```

(4) 配置 BGP 协议，建立 BGP IPv4 组播对等体，并引入路由

# 在 Router A 上配置其与 Router B 建立 EBGP 会话，使能 Router A 与 Router B 交换用于 RPF 检查的 IPv4 单播路由的能力，并引入直连路由。

```
[RouterA] bgp 100
```

```
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 192.168.1.2 as-number 200
[RouterA-bgp] address-family ipv4 multicast
[RouterA-bgp-mul-ipv4] peer 192.168.1.2 enable
[RouterA-bgp-mul-ipv4] import-route direct
[RouterA-bgp-mul-ipv4] quit
[RouterA-bgp] quit
```

# 在 Router B 上配置其与 Router A 建立 EBGP 会话，使能 Router A 与 Router B 交换用于 RPF 检查的 IPv4 单播路由的能力，并引入 OSPF 路由。

```
[RouterB] bgp 200
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 192.168.1.1 as-number 100
[RouterB-bgp] address-family ipv4 multicast
[RouterB-bgp-mul-ipv4] peer 192.168.1.1 enable
[RouterB-bgp-mul-ipv4] import-route ospf 1
[RouterB-bgp-mul-ipv4] quit
[RouterB-bgp] quit
```

(5) 配置 MSDP 对等体

# 在 Router A 上配置 MSDP 对等体。

```
[RouterA] msdp
[RouterA-msdp] peer 192.168.1.2 connect-interface pos 2/1/0
[RouterA-msdp] quit
```

# 在 Router B 上配置 MSDP 对等体。

```
[RouterB] msdp
[RouterB-msdp] peer 192.168.1.1 connect-interface pos 2/1/0
[RouterB-msdp] quit
```

### 4. 验证配置

# 执行 **display bgp peer ipv4 multicast** 命令查看 BGP IPv4 组播对等体。以 Router B 为例：

```
[RouterB] display bgp peer ipv4 multicast

 BGP local router ID : 2.2.2.2
 Local AS number : 200
 Total number of peers : 1                 Peers in established state : 1

  Peer            AS  MsgRcvd  MsgSent  OutQ PrefRcv Up/Down  State

  192.168.1.1    100       56       56     0       0 00:40:54 Established
```

# 执行 **display msdp brief** 命令查看路由器之间 MSDP 对等体的建立情况。以 Router B 为例：

```
[RouterB] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  1            1            0            0            0            0
  Peer's Address    State    Up/Down time    AS     SA Count   Reset Count
  192.168.1.1       Up       00:07:17        100    1          0
```

# 1.18 IPv6 BGP典型配置举例

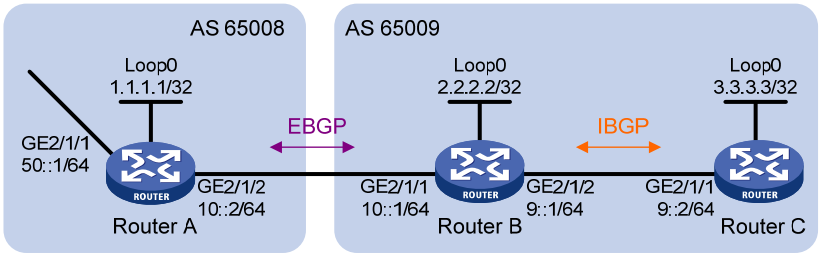
## 1.18.1 IPv6 BGP基本配置

### 1. 组网需求

如 图 1-28 所示，所有路由器均运行IPv6 BGP协议。Router A位于AS 65008；Router B和Router C位于AS 65009。要求Router A和Router B之间建立EBGP连接，Router B和Router C之间建立IBGP连接，使得Router C能够访问Router A直连的 50::/64 网段。

### 2. 组网图

图1-28 IPv6 BGP 基本配置组网图




### 3. 配置步骤

(1) 配置各接口的 IPv6 地址及 Loopback 接口的 IPv4 地址（略）

(2) 配置 IBGP 连接

# 配置 Router B。

```
<RouterB> system-view
[RouterB] bgp 65009
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 9::2 as-number 65009
[RouterB-bgp] address-family ipv6
[RouterB-bgp-ipv6] peer 9::2 enable
[RouterB-bgp-ipv6] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] bgp 65009
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 9::1 as-number 65009
[RouterC-bgp] address-family ipv6
[RouterC-bgp-ipv6] peer 9::1 enable
```

(3) 配置 EBGP 连接

# 配置 Router A。

```
<RouterA> system-view
[RouterA] bgp 65008
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 10::1 as-number 65009
[RouterA-bgp] address-family ipv6
```

```
[RouterA-bgp-ipv6] peer 10::1 enable
```

# 配置 Router B。

```
[RouterB-bgp] peer 10::2 as-number 65008
[RouterB-bgp] address-family ipv6
[RouterB-bgp-ipv6] peer 10::2 enable
```

(4) 配置通过 IPv6 BGP 发布的网段路由

# 配置 Router A。

```
[RouterA-bgp-ipv6] network 10:: 64
[RouterA-bgp-ipv6] network 50:: 64
[RouterA-bgp-ipv6] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
[RouterB-bgp-ipv6] network 10:: 64
[RouterB-bgp-ipv6] network 9:: 64
[RouterB-bgp-ipv6] quit
[RouterB-bgp] quit
```

# 配置 Router C。

```
[RouterC-bgp-ipv6] network 9:: 64
[RouterC-bgp-ipv6] quit
[RouterC-bgp] quit
```

### 4. 验证配置

# 在 Router B 上查看 IPv6 BGP 对等体的信息。可以看出，Router A 和 Router B 之间建立了 EBGP 连接，Router B 和 Router C 之间建立了 IBGP 连接。

```
[RouterB] display bgp peer ipv6

 BGP local router ID: 2.2.2.2
 Local AS number: 65009
 Total number of peers: 2                 Peers in established state: 2

  Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State

  9::2                 65009       41       43    0       1 00:29:00 Established
  10::2                65008       38       38    0       2 00:27:20 Established
```

# 在 Router A 上查看 IPv6 BGP 路由表信息。可以看出，Router A 学习到了 AS 65009 内的路由信息。

```
[RouterA] display bgp routing-table ipv6

 Total number of routes: 4

 BGP local router ID is 1.1.1.1
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

* >e Network : 9::                                    PrefixLen : 64
     NextHop : 10::1                                  LocPrf    :
```

```
     PrefVal : 0                                          OutLabel  : NULL
     MED     : 0
     Path/Ogn: 65009i

* >  Network : 10::                                       PrefixLen : 64
     NextHop : ::                                         LocPrf    :
     PrefVal : 32768                                      OutLabel  : NULL
     MED     : 0
     Path/Ogn: i

*  e Network : 10::                                       PrefixLen : 64
     NextHop : 10::1                                      LocPrf    :
     PrefVal : 0                                          OutLabel  : NULL
     MED     : 0
     Path/Ogn: 65009i

* >  Network : 50::                                       PrefixLen : 64
     NextHop : ::                                         LocPrf    :
     PrefVal : 32768                                      OutLabel  : NULL
     MED     : 0
     Path/Ogn: i
```

# 在 Router C 上查看 IPv6 BGP 路由表信息。可以看出，Router C 学习到了到达 50::/64 网段的路由。

```
[RouterC] display bgp routing-table ipv6

 Total number of routes: 4

 BGP local router ID is 3.3.3.3
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

* >  Network : 9::                                        PrefixLen : 64
     NextHop : ::                                         LocPrf    :
     PrefVal : 32768                                      OutLabel  : NULL
     MED     : 0
     Path/Ogn: i

*  i Network : 9::                                        PrefixLen : 64
     NextHop : 9::1                                       LocPrf    : 100
     PrefVal : 0                                          OutLabel  : NULL
     MED     : 0
     Path/Ogn: i

* >i Network : 10::                                       PrefixLen : 64
     NextHop : 9::1                                       LocPrf    : 100
     PrefVal : 0                                          OutLabel  : NULL
     MED     : 0
```

```
     Path/Ogn: i

* >i Network : 50::                                        PrefixLen : 64
     NextHop : 10::2                                      LocPrf    : 100
     PrefVal : 0                                          OutLabel  : NULL
     MED     : 0
     Path/Ogn: 65008i
```

# 在 Router C 上可以 ping 通 50::/64 网段内的主机。（略）

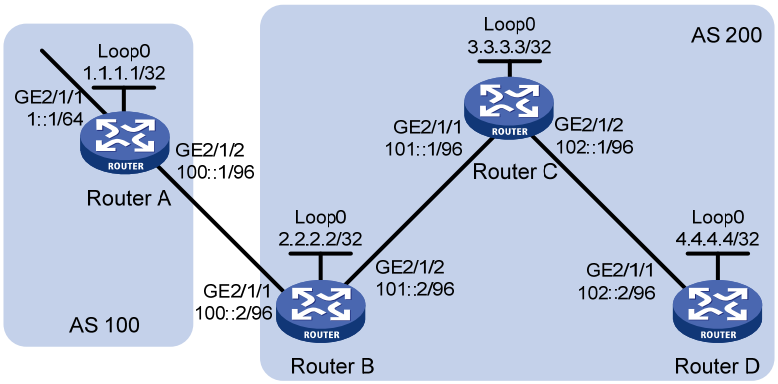
## 1.18.2 IPv6 BGP路由反射器配置

### 1. 组网需求

所有路由器均运行 IPv6 BGP 协议。Router A 与 Router B 之间建立 EBGP 连接。配置 Router C 作为路由反射器，Router B 和 Router D 为 Router C 的客户机，以避免在 AS 200 内建立 IBGP 全连接。即，只需在 Router C 与 Router B、Router C 与 Router D 之间建立 IBGP 连接，Router B 和 Router D 之间无需建立 IBGP 连接，即可实现 BGP 路由信息在整个 AS 200 内传递。

### 2. 组网图

图1-29 IPv6 BGP 路由反射器配置组网图




### 3. 配置步骤

(1) 配置各接口的 IPv6 地址及 Loopback 接口的 IPv4 地址（略）

(2) 配置 IBGP 和 EBGP 连接，并配置通过 IPv6 BGP 发布的网段路由

# 配置 Router A。

```
<RouterA> system-view
[RouterA] bgp 100
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 100::2 as-number 200
[RouterA-bgp] address-family ipv6
[RouterA-bgp-ipv6] peer 100::2 enable
[RouterA-bgp-ipv6] network 1:: 64
[RouterA-bgp-ipv6] network 100:: 96
[RouterA-bgp-ipv6] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] bgp 200
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 100::1 as-number 100
[RouterB-bgp] peer 101::1 as-number 200
[RouterB-bgp] address-family ipv6
[RouterB-bgp-ipv6] peer 100::1 enable
[RouterB-bgp-ipv6] peer 101::1 enable
[RouterB-bgp-ipv6] peer 101::1 next-hop-local
[RouterB-bgp-ipv6] network 100:: 96
[RouterB-bgp-ipv6] network 101:: 96
[RouterB-bgp-ipv6] quit
[RouterB-bgp] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] bgp 200
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 101::2 as-number 200
[RouterC-bgp] peer 102::2 as-number 200
[RouterC-bgp] address-family ipv6
[RouterC-bgp-ipv6] peer 101::2 enable
[RouterC-bgp-ipv6] peer 102::2 enable
[RouterC-bgp-ipv6] network 101:: 96
[RouterC-bgp-ipv6] network 102:: 96
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] bgp 200
[RouterD-bgp] router-id 4.4.4.4
[RouterD-bgp] peer 102::1 as-number 200
[RouterD-bgp] address-family ipv6
[RouterD-bgp-ipv6] peer 102::1 enable
[RouterD-bgp-ipv6] network 102:: 96
```

(3)　配置路由反射器

# 配置 Router C 作为路由反射器，Router B 和 Router D 是它的两个客户机。

```
[RouterC-bgp-ipv6] peer 101::2 reflect-client
[RouterC-bgp-ipv6] peer 102::2 reflect-client
[RouterC-bgp-ipv6] quit
[RouterC-bgp] quit
```

### 4. 验证配置

# 在 Router D 上执行 **display bgp routing-table ipv6** 命令，可以看到 Router D 通过 Router C 反射路由学习到了到达 1::/64 网段的路由。

```
[RouterD] display bgp routing-table ipv6

 Total number of routes: 5
```

```
 BGP local router ID is 4.4.4.4
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

* >i Network : 1::                                      PrefixLen : 64
     NextHop : 101::2                                   LocPrf    : 100
     PrefVal : 0                                        OutLabel  : NULL
     MED     : 0
     Path/Ogn: 100i

* >i Network : 100::                                    PrefixLen : 96
     NextHop : 101::2                                   LocPrf    : 100
     PrefVal : 0                                        OutLabel  : NULL
     MED     : 0
     Path/Ogn: i

* >i Network : 101::                                    PrefixLen : 96
     NextHop : 102::1                                   LocPrf    : 100
     PrefVal : 0                                        OutLabel  : NULL
     MED     : 0
     Path/Ogn: i

* >  Network : 102::                                    PrefixLen : 96
     NextHop : ::                                       LocPrf    :
     PrefVal : 32768                                    OutLabel  : NULL
     MED     : 0
     Path/Ogn: i

*  i Network : 102::                                    PrefixLen : 96
     NextHop : 102::1                                   LocPrf    : 100
     PrefVal : 0                                        OutLabel  : NULL
     MED     : 0
     Path/Ogn: i
```

### 1.18.3 6PE配置

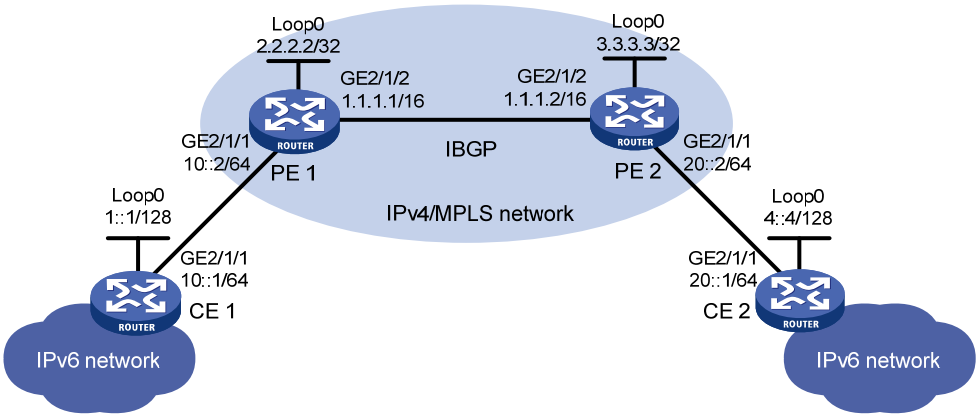
#### 1. 组网需求

通过配置 6PE 实现利用 MPLS 技术跨越运营商的 IPv4 网络连接隔离的两个 IPv6 用户网络。其中：

- 运营商网络内部采用 OSPF 作为 IGP 路由协议。
- PE 1 和 PE 2 为运营商网络的边缘设备，PE 1 和 PE 2 之间建立 IPv4 IBGP 连接。
- CE 1 和 CE 2 为 IPv6 用户网络的边缘设备，用户网络通过该设备接入运营商网络。
- CE 与 PE 之间配置 IPv6 静态路由，以指导 IPv6 报文的转发。

## 2. 组网图

图1-30 6PE 配置组网图




## 3. 配置过程

(1) 配置各接口的 IPv6 地址及 IPv4 地址（略）

(2) 配置 PE 1

# 全局使能 LDP 能力，并配置 LSP 触发策略。

```
<PE1> system-view
[PE1] mpls lsr-id 2.2.2.2
[PE1] mpls ldp
[PE1-ldp] lsp-trigger all
[PE1-ldp] quit
```

# 在接口 GigabitEthernet2/1/2 上使能 MPLS 和 LDP 能力。

```
[PE1] interface gigabitethernet 2/1/2
[PE1-GigabitEthernet2/1/2] mpls enable
[PE1-GigabitEthernet2/1/2] mpls ldp enable
[PE1-GigabitEthernet2/1/2] quit
```

# 配置 IBGP，使能对等体的 6PE 能力，并引入 IPv6 的直连和静态路由。

```
[PE1] bgp 65100
[PE1-bgp] router-id 2.2.2.2
[PE1-bgp] peer 3.3.3.3 as-number 65100
[PE1-bgp] peer 3.3.3.3 connect-interface loopback 0
[PE1-bgp] address-family ipv6
[PE1-bgp-ipv6] import-route direct
[PE1-bgp-ipv6] import-route static
[PE1-bgp-ipv6] peer 3.3.3.3 enable
[PE1-bgp-ipv6] peer 3.3.3.3 label-route-capability
[PE1-bgp-ipv6] quit
[PE1-bgp] quit
```

# 配置到 CE 1 的静态路由。

```
[PE1] ipv6 route-static 1::1 128 10::1
```

# 配置 OSPF，实现运营商网络内部互通。

```
[PE1] ospf
```

```
[PE1-ospf-1] area 0
[PE1-ospf-1-area-0.0.0.0] network 2.2.2.2 0.0.0.0
[PE1-ospf-1-area-0.0.0.0] network 1.1.0.0 0.0.255.255
[PE1-ospf-1-area-0.0.0.0] quit
[PE1-ospf-1] quit
```

(3) 配置 PE 2

# 全局使能 LDP 能力，并配置 LSP 触发策略。

```
<PE2> system-view
[PE2] mpls lsr-id 3.3.3.3
[PE2] mpls ldp
[PE2-mpls-ldp] lsp-trigger all
[PE2-mpls-ldp] quit
```

# 在接口 GigabitEthernet2/1/2 上使能 MPLS 和 LDP 能力。

```
[PE2] interface gigabitethernet 2/1/2
[PE2-GigabitEthernet2/1/2] mpls enable
[PE2-GigabitEthernet2/1/2] mpls ldp enable
[PE2-GigabitEthernet2/1/2] quit
```

# 配置 IBGP，使能对等体的 6PE 能力，并引入 IPv6 的直连和静态路由。

```
[PE2] bgp 65100
[PE2-bgp] router-id 3.3.3.3
[PE2-bgp] peer 2.2.2.2 as-number 65100
[PE2-bgp] peer 2.2.2.2 connect-interface loopback 0
[PE2-bgp] address-family ipv6
[PE2-bgp-ipv6] import-route direct
[PE2-bgp-ipv6] import-route static
[PE2-bgp-ipv6] peer 2.2.2.2 enable
[PE2-bgp-ipv6] peer 2.2.2.2 label-route-capability
[PE2-bgp-ipv6] quit
[PE2-bgp] quit
```

# 配置到 CE 2 的静态路由。

```
[PE2] ipv6 route-static 4::4 128 20::1
```

# 配置 OSPF，实现运营商内部互通。

```
[PE2] ospf
[PE2-ospf-1] area 0
[PE2-ospf-1-area-0.0.0.0] network 3.3.3.3 0.0.0.0
[PE2-ospf-1-area-0.0.0.0] network 1.1.0.0 0.0.255.255
[PE2-ospf-1-area-0.0.0.0] quit
[PE2-ospf-1] quit
```

(4) 配置 CE 1

# 配置静态路由，缺省下一跳为 PE 1。

```
<CE1> system-view
[CE1] ipv6 route-static :: 0 10::2
```

(5) 配置 CE 2

# 配置静态路由，缺省下一跳为 PE 2。

```
<CE2> system-view
[CE2] ipv6 route-static :: 0 20::2
```

#### 4. 验证配置

# 显示 PE 1 上的 IPv6 BGP 路由信息。可以看到 PE 1 上存在到达两个 IPv6 用户网络的路由。

```
[PE1] display bgp routing-table ipv6

 Total number of routes: 5

 BGP local router ID is 2.2.2.2
 Status codes: * - valid, > - best, d - dampened, h - history,
               s - suppressed, S - stale, i - internal, e - external
               Origin: i - IGP, e - EGP, ? - incomplete

* >  Network : 1::1                                    PrefixLen : 128
     NextHop : 10::1                                   LocPrf    :
     PrefVal : 32768                                   OutLabel  : NULL
     MED     : 0
     Path/Ogn: ?

* >i Network : 4::4                                    PrefixLen : 128
     NextHop : ::FFFF:3.3.3.3                          LocPrf    : 100
     PrefVal : 0                                       OutLabel  : 1279
     MED     : 0
     Path/Ogn: ?

* >  Network : 10::                                    PrefixLen : 64
     NextHop : ::                                      LocPrf    :
     PrefVal : 32768                                   OutLabel  : NULL
     MED     : 0
     Path/Ogn: ?

* >  Network : 10::2                                   PrefixLen : 128
     NextHop : ::1                                     LocPrf    :
     PrefVal : 32768                                   OutLabel  : NULL
     MED     : 0
     Path/Ogn: ?

* >i Network : 20::                                    PrefixLen : 64
     NextHop : ::FFFF:3.3.3.3                          LocPrf    : 100
     PrefVal : 0                                       OutLabel  : 1278
     MED     : 0
     Path/Ogn: ?
```

# CE 1 可以 ping 通 CE 2 的 IPv6 地址（Loopback 接口地址 4::4）。

### 1.18.4 IPv6 BGP与BFD联动配置

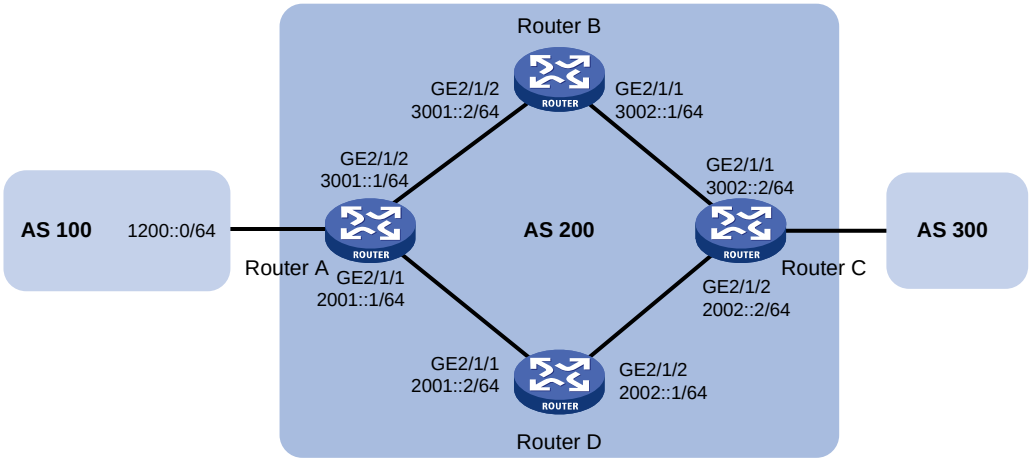
#### 1. 组网需求

- 在 AS 200 内使用 OSPFv3 作为 IGP 协议，实现 AS 内的互通。

- Router A 与 Router C 之间建立两条 IBGP 连接。当 Router A 与 Router C 之间的两条路径均连通时，Router C 与 1200::0/64 之间的报文使用 Router A<—>Router B<—>Router C 这条路径转发；当 Router A<—>Router B<—>Router C 这条路径发生故障时，BFD 能够快速检测并通告 IPv6 BGP 协议，使得 Router A<—>Router D<—>Router C 这条路径能够迅速生效。

## 2. 组网图

图1-31 IPv6 BGP 与 BFD 联动配置组网图




## 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 OSPFv3，保证 Router A 和 Router C 之间路由可达（略）

(3) Router A 上的 IPv6 BGP 配置

# 配置 Router A 和 Router C 建立两条 IBGP 连接。

```
<RouterA> system-view
[RouterA] bgp 200
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 2002::2 as-number 200
[RouterA-bgp] peer 3002::2 as-number 200
[RouterA-bgp] address-family ipv6
[RouterA-bgp-ipv6] peer 2002::2 enable
[RouterA-bgp-ipv6] peer 3002::2 enable
[RouterA-bgp-ipv6] quit
```

# 配置当 Router A 与 Router C 之间的两条路径均连通时，Router C 与 1200::0/64 之间的报文使用 Router A<—>Router B<—>Router C 这条路径转发。（在 Router A 上对发布给对等体 2002::2 的 1200::0/64 路由配置较高的 MED 属性值）

- 定义编号为 2000 的 IPv6 ACL，允许路由 1200::0/64 通过。

```
[RouterA] acl ipv6 number 2000
[RouterA-acl6-basic-2000] rule permit source 1200:: 64
[RouterA-acl6-basic-2000] quit
```

- 定义两个 Route-policy，一个名为 apply_med_50，为路由 1200::0/64 设置 MED 属性值为 50；另一个名为 apply_med_100，为路由 1200::0/64 设置 MED 属性值为 100。

```
[RouterA] route-policy apply_med_50 permit node 10
```

```
[RouterA-route-policy-apply_med_50-10] if-match ipv6 address acl 2000
[RouterA-route-policy-apply_med_50-10] apply cost 50
[RouterA-route-policy-apply_med_50-10] quit
[RouterA] route-policy apply_med_100 permit node 10
[RouterA-route-policy-apply_med_100-10] if-match ipv6 address acl 2000
[RouterA-route-policy-apply_med_100-10] apply cost 100
[RouterA-route-policy-apply_med_100-10] quit
```

- 对发布给对等体 3002::2 的路由应用名为 apply_med_50 的 Route-policy，对发布给对等体 2002::2 的路由应用名为 apply_med_100 的 Route-policy。

```
[RouterA] bgp 200
[RouterA-bgp] address-family ipv6 unicast
[RouterA-bgp-ipv6] peer 3002::2 route-policy apply_med_50 export
[RouterA-bgp-ipv6] peer 2002::2 route-policy apply_med_100 export
[RouterA-bgp-ipv6] quit
```

# 配置通过 BFD 检测 Router A<－>Router B<－>Router C 这条路径，当该路径出现故障时，BFD 能够快速检测到并通告 IPv6 BGP 协议，使得 Router A<－>Router D<－>Router C 这条路径能够迅速生效。

```
[RouterA-bgp] peer 3002::2 bfd
[RouterA-bgp] quit
```

(4) Router C 上的 IPv6 BGP 配置。

# 配置 Router A 和 Router C 建立两条 IBGP 连接。

```
<RouterC> system-view
[RouterC] bgp 200
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] peer 3001::1 as-number 200
[RouterC-bgp] peer 2001::1 as-number 200
[RouterC-bgp] address-family ipv6
[RouterC-bgp-ipv6] peer 3001::1 enable
[RouterC-bgp-ipv6] peer 2001::1 enable
[RouterC-bgp-ipv6] quit
```

# 配置通过 BFD 检测 Router A<－>Router B<－>Router C 这条路径，当该路径出现故障时，BFD 能够快速检测到，并通告 IPv6 BGP 协议，使得 Router A<－>Router D<－>Router C 这条路径能够迅速生效。

```
[RouterC-bgp] peer 3001::1 bfd
[RouterC-bgp] quit
[RouterC] quit
```

### 4. 验证配置

下面以 Router C 为例，Router A 与此类似，不再赘述。

# 显示 Router C 的 BFD 信息。可以看出，Router A 和 Router C 之间已经建立了 BFD 会话，而且 BFD 协议运行正常。

```
<RouterC> display bfd session verbose
 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv6 Session Working Under Ctrl Mode:
```

```
       Local Discr: 513                 Remote Discr: 513
         Source IP: 3002::2
    Destination IP: 3001::1
     Session State: Up                     Interface: N/A
      Min Tx Inter: 500ms                Act Tx Inter: 500ms
      Min Rx Inter: 500ms                Detect Inter: 2500ms
          Rx Count: 13                       Tx Count: 14
      Connect Type: Indirect           Running Up for: 00:00:05
         Hold Time: 2243ms                  Auth mode: None
       Detect Mode: Async                       Slot: 0
          Protocol: BGP6
         Diag Info: No Diagnostic
```

# 在 Router C 上查看 BGP 邻居信息。可以看出，Router A 和 Router C 之间建立两条 BGP 连接，且均处于 Established 状态。

```
<RouterC> display bgp peer ipv6

 BGP local router ID: 3.3.3.3
 Local AS number: 200
 Total number of peers: 2                 Peers in established state: 2

  Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State

  2001::1                200        8        8    0       0 00:04:45 Established
  3001::1                200        5        4    0       0 00:01:53 Established
```

# 在 Router C 上查看 1200::0/64 的路由信息，可以看出 Router C 通过 Router A<－>Router B<－>Router C 这条路径与 1200::0/64 网段通信。

```
<RouterC> display ipv6 routing-table 1200::0 64 verbose

Summary Count : 1

Destination: 1200::/64
   Protocol: BGP4+           Process ID: 0
  SubProtID: 0x1                    Age: 00h01m07s
       Cost: 50              Preference: 255
        Tag: 0                    State: Active Adv
  OrigTblID: 0x1                OrigVrf: default-vrf
    TableID: 0xa                 OrigAs: 0
      NBRID: 0x25000001          LastAs: 0
     AttrID: 0x1                Neighbor: 3001::1
      Flags: 0x10060        OrigNextHop: 3001::1
      Label: NULL           RealNextHop: FE80::20C:29FF:FE4A:3873
    BkLabel: NULL             BkNextHop: N/A
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid        BkInterface: N/A
```

# Router A<－>Router B<－>Router C 这条路径发生故障后，在 Router C 上查看 1200::0/64 的路由信息，可以看出 Router C 通过 Router A<－>Router D<－>Router C 这条路径转发报文。

```
<RouterC> display ipv6 routing-table 1200::0 64 verbose
```

```
Summary Count : 1

Destination: 1200::/64
   Protocol: BGP4+           Process ID: 0
  SubProtID: 0x1                    Age: 00h00m57s
       Cost: 100             Preference: 255
        Tag: 0                    State: Active Adv
  OrigTblID: 0x1               OrigVrf: default-vrf
    TableID: 0xa                OrigAs: 0
      NBRID: 0x25000000         LastAs: 0
     AttrID: 0x0                Neighbor: 2001::1
      Flags: 0x10060          OrigNextHop: 2001::1
      Label: NULL             RealNextHop: FE80::20C:29FF:FE40:715
    BkLabel: NULL               BkNextHop: N/A
  Tunnel ID: Invalid            Interface: GigabitEthernet2/1/2
BkTunnel ID: Invalid          BkInterface: N/A
```

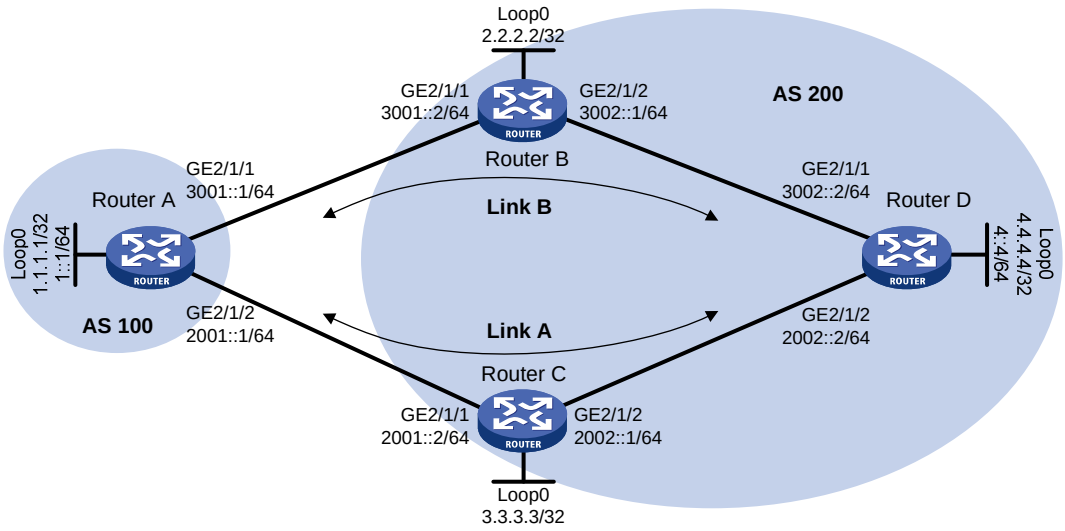
## 1.18.5 配置BGP快速重路由

### 1. 组网需求

如 图 1-32 所示，Router A、Router B、Router C和Router D通过BGP协议实现网络互连。要求链路B正常时，Router A和Router D之间的流量通过链路B转发；链路B出现故障时，流量可以快速切换到链路A上。

### 2. 组网图

图1-32 配置 BGP 快速重路由




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 在 AS 200 内配置 OSPFv3，发布接口地址所在网段的路由（包括 Loopback 接口），确保 Router B、Router C 和 Router D 之间 IPv6 路由可达（略）

(3) 配置 BGP 连接

# 配置 Router A 分别与 Router B 和 Router C 建立 EBGP 会话，并配置通过 BGP 发布路由 1::/64。

```
<RouterA> system-view
[RouterA] bgp 100
[RouterA] router-id 1.1.1.1
[RouterA-bgp] peer 3001::2 as-number 200
[RouterA-bgp] peer 2001::2 as-number 200
[RouterA-bgp] address-family ipv6 unicast
[RouterA-bgp-ipv6] peer 3001::2 enable
[RouterA-bgp-ipv6] peer 2001::2 enable
[RouterA-bgp-ipv6] network 1:: 64
[RouterA-bgp-ipv6] quit
[RouterA-bgp] quit
```

# 配置 Router B 与 Router A 建立 EBGP 会话，与 Router D 建立 IBGP 会话。

```
<RouterB> system-view
[RouterB] bgp 200
[RouterB] router-id 2.2.2.2
[RouterB-bgp] peer 3001::1 as-number 100
[RouterB-bgp] peer 3002::2 as-number 200
[RouterB-bgp] address-family ipv6 unicast
[RouterB-bgp-ipv6] peer 3001::1 enable
[RouterB-bgp-ipv6] peer 3002::2 enable
[RouterB-bgp-ipv6] peer 3002::2 next-hop-local
[RouterB-bgp-ipv6] quit
[RouterB-bgp] quit
```

# 配置 Router C 与 Router A 建立 EBGP 会话，与 Router D 建立 IBGP 会话。

```
<RouterC> system-view
[RouterC] bgp 200
[RouterC] router-id 3.3.3.3
[RouterC-bgp] peer 2001::1 as-number 100
[RouterC-bgp] peer 2002::2 as-number 200
[RouterC-bgp] address-family ipv6 unicast
[RouterC-bgp-ipv6] peer 2001::1 enable
[RouterC-bgp-ipv6] peer 2002::2 enable
[RouterC-bgp-ipv6] peer 2002::2 next-hop-local
[RouterC-bgp-ipv6] quit
[RouterC-bgp] quit
```

# 配置 Router D 分别与 Router B 和 Router C 建立 IBGP 会话，并配置 BGP 发布路由 4::/64。

```
<RouterD> system-view
[RouterD] bgp 200
[RouterD-bgp] peer 3002::1 as-number 200
[RouterD-bgp] peer 2002::1 as-number 200
[RouterD-bgp] address-family ipv6 unicast
[RouterD-bgp-ipv6] peer 3002::1 enable
[RouterD-bgp-ipv6] peer 2002::1 enable
[RouterD-bgp-ipv6] network 4:: 64
[RouterD-bgp-ipv6] quit
```

```
[RouterD-bgp] quit
```

(4) 修改路由的首选值，使得 Router A 和 Router D 之间的流量优先通过链路 B 转发

# 在 Router A 上配置从 Router B 接收到的路由的首选值为 100。

```
[RouterA-bgp-ipv6] peer 3001::2 preferred-value 100
[RouterA-bgp-ipv6] quit
[RouterA-bgp] quit
```

# 在 Router D 上配置从 Router B 接收到的路由的首选值为 100。

```
[RouterD-bgp-ipv6] peer 3002::1 preferred-value 100
[RouterD-bgp-ipv6] quit
[RouterD-bgp] quit
```

(5) 配置 BGP 快速重路由

# 配置 Router A：创建路由策略 frr，为路由 4::/64 指定备份下一跳的地址为 2001::2（对等体 Router C 的地址）；在 BGP IPv6 单播地址族下应用该路由策略。

```
<RouterA> system-view
[RouterA] ipv6 prefix-list abc index 10 permit 4:: 64
[RouterA] route-policy frr permit node 10
[RouterA-route-policy] if-match ipv6 address prefix-list abc
[RouterA-route-policy] apply ipv6 fast-reroute backup-nexthop 2001::2
[RouterA-route-policy] quit
[RouterA] bgp 100
[RouterA-bgp] address-family ipv6 unicast
[RouterA-bgp-ipv6] fast-reroute route-policy frr
[RouterA-bgp-ipv6] quit
[RouterA-bgp] quit
```

# 配置 Router D：创建路由策略 frr，为路由 1::/64 指定备份下一跳的地址为 2002::1（对等体 Router C 的地址）；在 BGP IPv6 单播地址族下应用该路由策略。

```
<RouterD> system-view
[RouterD] ipv6 prefix-list abc index 10 permit 1:: 64
[RouterD] route-policy frr permit node 10
[RouterD-route-policy] if-match ipv6 address prefix-list abc
[RouterD-route-policy] apply ipv6 fast-reroute backup-nexthop 2002::1
[RouterD-route-policy] quit
[RouterD] bgp 200
[RouterD-bgp] address-family ipv6 unicast
[RouterD-bgp-ipv6] fast-reroute route-policy frr
[RouterD-bgp-ipv6] quit
[RouterD-bgp] quit
```

### 4. 验证配置

# 在 Router A 上查看 4::/64 路由，可以看到备份下一跳信息。

```
[RouterA] display ipv6 routing-table 4:: 64 verbose

Summary Count : 1

Destination: 4::/64
   Protocol: BGP4+           Process ID: 0
  SubProtID: 0x2                    Age: 00h00m58s
```

```
       Cost: 0               Preference: 255
     IpPre: N/A             QosLocalID: N/A
        Tag: 0                   State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0xa                OrigAs: 200
      NibID: 0x25000003         LastAs: 200
     AttrID: 0x3              Neighbor: 3001::2
      Flags: 0x10060       OrigNextHop: 3001::2
      Label: NULL          RealNextHop: 3001::2
    BkLabel: NULL            BkNextHop: 2001::2
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid       BkInterface: GigabitEthernet2/1/2
   FtnIndex: 0x0
```

# 在 Router D 上查看 1::/64 路由，可以看到备份下一跳信息。

```
[RouterD] display ipv6 routing-table 1:: 64 verbose

Summary Count : 1

Destination: 1::/64
   Protocol: BGP4+           Process ID: 0
  SubProtID: 0x1                    Age: 00h03m24s
       Cost: 0               Preference: 255
     IpPre: N/A             QosLocalID: N/A
        Tag: 0                   State: Active Adv
  OrigTblID: 0x0               OrigVrf: default-vrf
    TableID: 0xa                OrigAs: 100
      NibID: 0x25000003         LastAs: 100
     AttrID: 0x4              Neighbor: 3002::1
      Flags: 0x10060       OrigNextHop: 3002::1
      Label: NULL          RealNextHop: 3002::1
    BkLabel: NULL            BkNextHop: 2002::1
  Tunnel ID: Invalid          Interface: GigabitEthernet2/1/1
BkTunnel ID: Invalid       BkInterface: GigabitEthernet2/1/2
   FtnIndex: 0x0
```

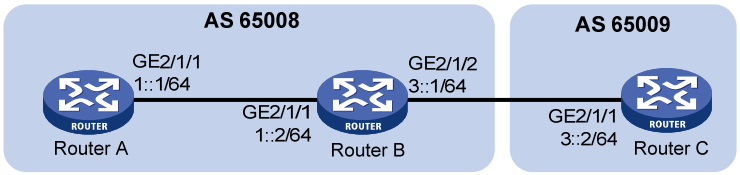
## 1.18.6 通过IPsec保护IPv6 BGP报文配置

### 1. 组网需求

- Router A、Router B 和 Router C 三台路由器之间运行 IPv6 BGP 交互路由信息。Router A 和 Router B 之间建立 IBGP 连接，Router B 和 Router C 之间建立 EBGP 连接。
- 为了提高安全性，配置通过 IPsec 对 IPv6 BGP 报文进行保护。

## 2. 组网图

图1-33 通过 IPsec 保护 IPv6 BGP 报文组网图




## 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 IBGP 连接

# 配置 Router A。

```
<RouterA> system-view
[RouterA] bgp 65008
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] group ibgp internal
[RouterA-bgp] peer 1::2 group ibgp
[RouterA-bgp] address-family ipv6 unicast
[RouterA-bgp-ipv6] peer ibgp enable
[RouterA-bgp-ipv6] quit
[RouterA-bgp] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] bgp 65008
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] group ibgp internal
[RouterB-bgp] peer 1::1 group ibgp
[RouterB-bgp] address-family ipv6 unicast
[RouterB-bgp-ipv6] peer ibgp enable
[RouterB-bgp-ipv6] quit
```

(3) 配置 EBGP 连接

# 配置 Router C。

```
<RouterC> system-view
[RouterC] bgp 65009
[RouterC-bgp] router-id 3.3.3.3
[RouterC-bgp] group ebgp external
[RouterC-bgp] peer 3::1 as-number 65008
[RouterC-bgp] peer 3::1 group ebgp
[RouterC-bgp] address-family ipv6 unicast
[RouterC-bgp-ipv6] peer ebgp enable
[RouterC-bgp-ipv6] quit
[RouterC-bgp] quit
```

# 配置 Router B。

```
[RouterB-bgp] group ebgp external
```

```
[RouterB-bgp] peer 3::2 as-number 65009
[RouterB-bgp] peer 3::2 group ebgp
[RouterB-bgp] address-family ipv6 unicast
[RouterB-bgp-ipv6] peer ebgp enable
[RouterB-bgp-ipv6] quit
[RouterB-bgp] quit
```

(4) 配置 IPsec 安全提议和安全框架

# 配置 Router A。创建名为 tran1 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议。创建手工方式的安全框架 policy001，配置 SPI 和密钥。

```
[RouterA] ipsec transform-set tran1
[RouterA-ipsec-transform-set-tran1] encapsulation-mode transport
[RouterA-ipsec-transform-set-tran1] esp encryption-algorithm des
[RouterA-ipsec-transform-set-tran1] esp authentication-algorithm sha1
[RouterA-ipsec-transform-set-tran1] quit
[RouterA] ipsec profile policy001 manual
[RouterA-ipsec-profile-policy001-manual] transform-set tran1
[RouterA-ipsec-profile-policy001-manual] sa spi outbound esp 12345
[RouterA-ipsec-profile-policy001-manual] sa spi inbound esp 12345
[RouterA-ipsec-profile-policy001-manual] sa string-key outbound esp simple abcdefg
[RouterA-ipsec-profile-policy001-manual] sa string-key inbound esp simple abcdefg
[RouterA-ipsec-profile-policy001-manual] quit
```

# 配置 Router B。创建名为 tran1 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议；创建手工方式的安全框架 policy001，配置 SPI 和密钥。创建名为 tran2 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议；创建手工方式的安全框架 policy002，配置 SPI 和密钥。

```
[RouterB] ipsec transform-set tran1
[RouterB-ipsec-transform-set-tran1] encapsulation-mode transport
[RouterB-ipsec-transform-set-tran1] esp encryption-algorithm des
[RouterB-ipsec-transform-set-tran1] esp authentication-algorithm sha1
[RouterB-ipsec-transform-set-tran1] quit
[RouterB] ipsec profile policy001 manual
[RouterB-ipsec-profile-policy001-manual] transform-set tran1
[RouterB-ipsec-profile-policy001-manual] sa spi outbound esp 12345
[RouterB-ipsec-profile-policy001-manual] sa spi inbound esp 12345
[RouterB-ipsec-profile-policy001-manual] sa string-key outbound esp simple abcdefg
[RouterB-ipsec-profile-policy001-manual] sa string-key inbound esp simple abcdefg
[RouterB-ipsec-profile-policy001-manual] quit
[RouterB] ipsec transform-set tran2
[RouterB-ipsec-transform-set-tran2] encapsulation-mode transport
[RouterB-ipsec-transform-set-tran2] esp encryption-algorithm des
[RouterB-ipsec-transform-set-tran2] esp authentication-algorithm sha1
[RouterB-ipsec-transform-set-tran2] quit
[RouterB] ipsec profile policy002 manual
[RouterB-ipsec-profile-policy002-manual] transform-set tran2
[RouterB-ipsec-profile-policy002-manual] sa spi outbound esp 54321
[RouterB-ipsec-profile-policy002-manual] sa spi inbound esp 54321
[RouterB-ipsec-profile-policy002-manual] sa string-key outbound esp simple gfedcba
```

```
[RouterB-ipsec-profile-policy002-manual] sa string-key inbound esp simple gfedcba
[RouterB-ipsec-profile-policy002-manual] quit
```

# 配置 Router C。创建名为 tran2 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议。创建手工方式的安全框架 policy002，配置 SPI 和密钥。

```
[RouterC] ipsec transform-set tran2
[RouterC-ipsec-transform-set-tran2] encapsulation-mode transport
[RouterC-ipsec-transform-set-tran2] esp encryption-algorithm des
[RouterC-ipsec-transform-set-tran2] esp authentication-algorithm sha1
[RouterC-ipsec-transform-set-tran2] quit
[RouterC] ipsec profile policy002 manual
[RouterC-ipsec-profile-policy002-manual] transform-set tran2
[RouterC-ipsec-profile-policy002-manual] sa spi outbound esp 54321
[RouterC-ipsec-profile-policy002-manual] sa spi inbound esp 54321
[RouterC-ipsec-profile-policy002-manual] sa string-key outbound esp simple gfedcba
[RouterC-ipsec-profile-policy002-manual] sa string-key inbound esp simple gfedcba
[RouterC-ipsec-profile-policy002-manual] quit
```

(5) 配置通过 IPsec 保护 Router A 和 Router B 之间的 IPv6 BGP 报文

# 配置 Router A。

```
[RouterA] bgp 65008
[RouterA-bgp] peer 1::2 ipsec-profile policy001
[RouterA-bgp] quit
```

# 配置 Router B。

```
[RouterB] bgp 65008
[RouterB-bgp] peer 1::1 ipsec-profile policy001
[RouterB-bgp] quit
```

(6) 配置通过 IPsec 保护 Router B 和 Router C 之间的 IPv6 BGP 报文

# 配置 Router C。

```
[RouterC] bgp 65009
[RouterC-bgp] peer ebgp ipsec-profile policy002
[RouterC-bgp] quit
```

# 配置 Router B。

```
[RouterB] bgp 65008
[RouterB-bgp] peer ebgp ipsec-profile policy002
[RouterB-bgp] quit
```

### 4. 验证配置

# 在 Router B 上显示 IPv6 BGP 对等体的详细信息。可以看出完成上述配置后 IBGP、EBGP 对等体能够正常建立，且发送和接收的 IPv6 BGP 报文都经过加密。

```
[RouterB] display bgp peer ipv6 verbose

          Peer: 1::1      Local: 2.2.2.2
          Type: IBGP link
          BGP version 4, remote router ID 1.1.1.1
          BGP current state: Established, Up for 00h05m54s
          BGP current event: KATimerExpired
          BGP last state: OpenConfirm
```

```
        Port:  Local - 24896    Remote - 179
        Configured: Active Hold Time: 180 sec   Keepalive Time: 60 sec
        Received  : Active Hold Time: 180 sec
        Negotiated: Active Hold Time: 180 sec   Keepalive Time: 60 sec
        Peer optional capabilities:
        Peer support BGP multi-protocol extended
        Peer support BGP route refresh capability
        Peer support BGP route AS4 capability
        Address family IPv6 Unicast: advertised and received

InQ updates: 0, OutQ updates: 0
NLRI statistics:
        Rcvd:   UnReach NLRI          0,       Reach NLRI          0
        Sent:   UnReach NLRI          0,       Reach NLRI          3

Message statistics:
Msg type     Last rcvd time/      Current rcvd count/      History rcvd count/
             Last sent time       Current sent count       History sent count
Open         18:59:15-2013.4.24   1                        1
             18:59:15-2013.4.24   1                        2
Update       -                    0                        0
             18:59:16-2013.4.24   1                        1
Notification -                    0                        0
             18:59:15-2013.4.24   0                        1
Keepalive    18:59:15-2013.4.24   1                        1
             18:59:15-2013.4.24   1                        1
RouteRefresh -                    0                        0
             -                    0                        0
Total        -                    2                        2
             -                    3                        5

Maximum allowed prefix number: 4294967295
Threshold: 75%
Minimum time between advertisements is 15 seconds
Optional capabilities:
 Multi-protocol extended capability has been enabled
 Route refresh capability has been enabled
Peer preferred value: 0
IPsec profile name: policy001

Routing policy configured:
No routing policy is configured

        Peer: 3::2      Local: 2.2.2.2
        Type: EBGP link
        BGP version 4, remote router ID 3.3.3.3
        BGP current state: Established, Up for 00h05m00s
        BGP current event: KATimerExpired
```

```
        BGP last state: OpenConfirm
        Port:  Local - 24897    Remote - 179
        Configured: Active Hold Time: 180 sec   Keepalive Time: 60 sec
        Received  : Active Hold Time: 180 sec
        Negotiated: Active Hold Time: 180 sec   Keepalive Time: 60 sec
        Peer optional capabilities:
        Peer support BGP multi-protocol extended
        Peer support BGP route refresh capability
        Peer support BGP route AS4 capability
        Address family IPv6 Unicast: advertised and received

Received: Total 8 messages, Update messages 1
Sent: Total 8 messages, Update messages 1
Maximum allowed prefix number: 4294967295
Threshold: 75%
Minimum time between advertisements is 30 seconds
Optional capabilities:
 Multi-protocol extended capability has been enabled
 Route refresh capability has been enabled
Peer preferred value: 0
IPsec profile name: policy002

Routing policy configured:
No routing policy is configured
```

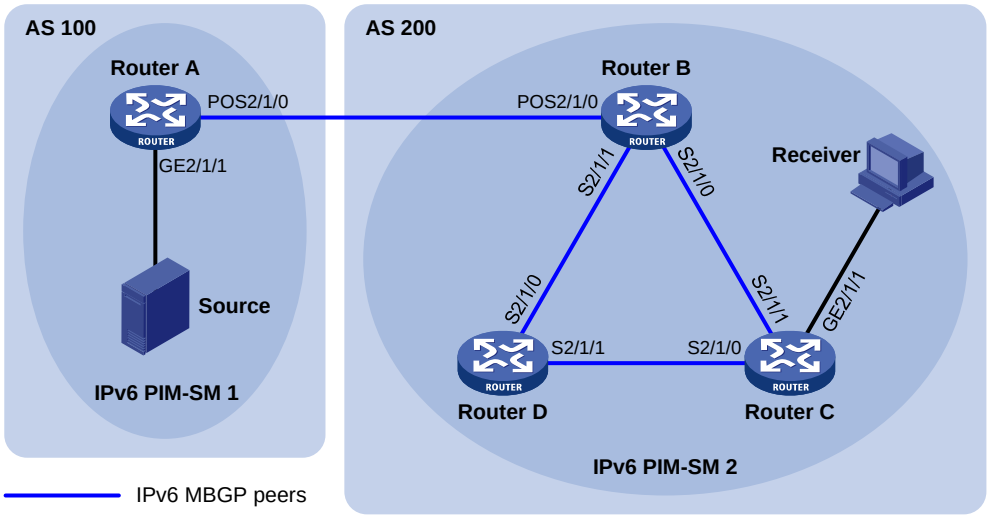
## 1.18.7 IPv6 MBGP配置

### 1. 组网需求

- 网络中存在两个自治系统：IPv6 PIM-SM 1 属于 AS 100，IPv6 PIM-SM 2 属于 AS 200。各 AS 内部采用 OSPFv3 交换路由信息，AS 之间采用 IPv6 MBGP 交换用于 RPF 检查的 IPv6 单播路由信息。
- IPv6 组播源属于 AS 100 内的 IPv6 PIM-SM 1，接收者则属于 AS 200 内的 IPv6 PIM-SM 2。
- 在 IPv6 PIM 域中的所有路由器上都使能嵌入式 RP 功能。

## 2. 组网图

图1-34 IPv6 MBGP 配置组网图




| 设备 | 接口 | IP地址 | 设备 | 接口 | IP地址 |
|---|---|---|---|---|---|
| Source | - | 1002::100/64 | Router C | GE2/1/1 | 3002::1/64 |
| Router A | GE2/1/1 | 1002::1/64 | | S2/1/0 | 3001::1/64 |
| | POS2/1/0 | 1001::1/64 | | S2/1/1 | 2001::2/64 |
| Router B | POS2/1/0 | 1001::2/64 | Router D | S2/1/0 | 2002::2/64 |
| | S2/1/0 | 2001::1/64 | | S2/1/1 | 3001::2/64 |
| | S2/1/1 | 2002::1/64 | | | |

## 3. 配置步骤

(1) 配置 IPv6 地址和 IPv6 单播路由协议

- 按照 图 1-34 配置各接口的IPv6 地址和前缀长度，具体配置过程略。
- 配置 AS 200 内的各路由器之间采用 OSPFv3 路由协议交换路由信息（AS 内各路由器使用的 OSPFv3 进程号为 1），确保各 AS 内部在网络层互通，具体配置过程略。

(2) 使能 IPv6 组播路由，使能 IPv6 PIM-SM 和 MLD，并配置 BSR 的服务边界

# 在 Router A 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM。

```
<RouterA> system-view
[RouterA] ipv6 multicast routing
[RouterA-mrib6] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ipv6 pim sm
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface pos 2/1/0
[RouterA-Pos2/1/0] ipv6 pim sm
[RouterA-Pos2/1/0] quit
```

Router B 和 Router D 上的配置与 Router A 相似，配置过程略。

# 在 Router C 上使能 IPv6 组播路由，在各接口上使能 IPv6 PIM-SM，并在主机侧接口 GigabitEthernet2/1/1 上使能 MLD。

```
<RouterC> system-view
[RouterC] ipv6 multicast routing
```

```
[RouterC-mrib6] quit
[RouterC] interface serial 2/1/0
[RouterC-Serial2/1/0] ipv6 pim sm
[RouterC-Serial2/1/0] quit
[RouterC] interface serial 2/1/1
[RouterC-Serial2/1/1] ipv6 pim sm
[RouterC-Serial2/1/1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ipv6 pim sm
[RouterC-GigabitEthernet2/1/1] mld enable
[RouterC-GigabitEthernet2/1/1] quit
```

# 在 Router A 上配置 BSR 的服务边界。

```
[RouterA] interface pos 2/1/0
[RouterA-Pos2/1/0] ipv6 pim bsr-boundary
[RouterA-Pos2/1/0] quit
```

# 在 Router B 上配置 BSR 的服务边界。

```
[RouterB] interface pos 2/1/0
[RouterB-Pos2/1/0] ipv6 pim bsr-boundary
[RouterB-Pos2/1/0] quit
```

(3) 使能嵌入式 RP 功能

# 在 Router A 上使能嵌入式 RP 功能。

```
[RouterA] ipv6 pim
[RouterA-pim6] embedded-rp
[RouterA-pim6] quit
```

Router B、Router C 和 Router D 上的配置与 Router A 相似，配置过程略。

(4) 配置 BGP 协议，建立 BGP IPv6 组播对等体，并引入路由

# 在 Router A 上配置其与 Router B 建立 EBGP 会话，使能 Router A 与 Router B 交换用于 RPF 检查的 IPv6 单播路由的能力，并引入直连路由。

```
[RouterA] bgp 100
[RouterA-bgp] router-id 1.1.1.1
[RouterA-bgp] peer 1001::2 as-number 200
[RouterA-bgp] address-family ipv6 multicast
[RouterA-bgp-mul-ipv6] peer 1001::2 enable
[RouterA-bgp-mul-ipv6] import-route direct
[RouterA-bgp-mul-ipv6] quit
[RouterA-bgp] quit
```

# 在 Router B 上配置其与 Router A 建立 EBGP 会话，使能 Router A 与 Router B 交换用于 RPF 检查的 IPv6 单播路由的能力，并引入 OSPFv3 路由。

```
[RouterB] bgp 200
[RouterB-bgp] router-id 2.2.2.2
[RouterB-bgp] peer 1001::1 as-number 100
[RouterB-bgp] address-family ipv6 multicast
[RouterB-bgp-mul-ipv6] peer 1001::1 enable
[RouterB-bgp-mul-ipv6] import-route ospfv3 1
[RouterB-bgp-mul-ipv6] quit
[RouterB-bgp] quit
```

#### 4. 验证配置

# 执行 **display bgp peer ipv6 multicast** 命令查看 BGP IPv6 组播对等体。以 Router B 为例：

```
[RouterB] display bgp peer ipv6 multicast

 BGP local router ID : 2.2.2.2
 Local AS number : 200
 Total number of peers : 3                 Peers in established state : 3

  Peer        AS  MsgRcvd  MsgSent  OutQ PrefRcv Up/Down  State

  1001::1     100       56       56      0        0 00:40:54 Established
```

## 1.19 BGP常见错误配置举例

### 1.19.1 故障现象

使用 **display bgp peer ipv4 unicast** 命令或 **display bgp peer ipv6 unicast** 命令查看 BGP 对等体的信息，发现与对端的连接无法进入 Established 状态。

### 1.19.2 故障分析

BGP 邻居的建立需要能够使用 179 端口建立 TCP 会话，以及能够正确交换 Open 消息。

### 1.19.3 故障处理

(1) 执行 **display current-configuration** 命令查看当前配置，检查邻居的 AS 号配置是否正确。

(2) 执行 **display bgp peer ipv4 unicast** 命令或 **display bgp peer ipv6 unicast** 命令检查邻居的 IP 地址/IPv6 地址是否正确。

(3) 如果使用 Loopback 接口，检查是否配置了 **peer connect-interface** 命令。

(4) 如果是物理上非直连的 EBGP 邻居，检查是否配置了 **peer ebgp-max-hop** 命令。

(5) 如果配置了 **peer ttl-security hops** 命令，请检查对端是否也配置了该命令，且保证双方配置的 *hop-count* 不小于两台设备实际需要经过的跳数。

(6) 检查路由表中是否存在到邻居的可用路由。

(7) 使用 **ping** 命令检查链路是否畅通。

(8) 使用 **display tcp verbose** 命令或 **display ipv6 tcp verbose** 命令检查 TCP 连接是否正常。

(9) 检查是否配置了禁止 TCP 端口 179 的 ACL。

# 目 录

# 1 策略路由

## 1.1 策略路由简介

与单纯依照 IP 报文的目的地址查找路由表进行转发不同，策略路由是一种依据用户制定的策略进行路由转发的机制。策略路由可以对于满足一定条件（ACL 规则、报文长度等）的报文，执行指定的操作（设置报文的下一跳、出接口、缺省下一跳和缺省出接口等）。

报文到达后，其后续的转发流程如下：

- 首先根据配置的策略路由转发。
- 若找不到匹配的节点或虽然找到了匹配的节点，但指导报文转发失败时，再根据路由表中除缺省路由之外的路由来转发报文。
- 若转发失败，则根据策略路由中配置的缺省下一跳和缺省出接口指导报文转发。
- 若转发失败，则再根据缺省路由来转发报文。

根据作用对象的不同，策略路由可分为本地策略路由和转发策略路由：

- 本地策略路由：对设备本身产生的报文（比如本地发出的 ping 报文）起作用，指导其发送。
- 转发策略路由：对接口接收的报文起作用，指导其转发。

### 1.1.1 策略简介

策略用来定义报文的匹配规则，以及对报文执行的操作。策略由节点组成。

一个策略可以包含一个或者多个节点。节点的构成如下：

- 每个节点由节点编号来标识。节点编号越小节点的优先级越高，优先级高的节点优先被执行。
- 每个节点的具体内容由 **if-match** 子句和 **apply** 子句来指定。**if-match** 子句定义该节点的匹配规则，**apply** 子句定义该节点的动作。
- 每个节点对报文的处理方式由匹配模式决定。匹配模式分为 **permit**（允许）和 **deny**（拒绝）两种。

应用策略后，系统将根据策略中定义的匹配规则和操作，对报文进行处理：系统按照优先级从高到低的顺序依次匹配各节点，如果报文满足这个节点的匹配规则，就执行该节点的动作；如果报文不满足这个节点的匹配规则，就继续匹配下一个节点；如果报文不能满足策略中任何一个节点的匹配规则，则根据路由表来转发报文。

#### 1. if-match子句

目前，策略路由提供了两种 **if-match** 子句，作用如下：

- **if-match acl**：设置 ACL 匹配规则。
- **if-match packet-length**：设置 IP 报文长度匹配规则。

在一个节点中可以配置多条 **if-match** 子句，同一类型的 **if-match** 子句最多只能有一条。

同一个节点中的各 **if-match** 子句之间是“与”的关系，即报文必须满足该节点的所有 **if-match** 子句才算满足这个节点的匹配规则。

### 2. apply子句

策略路由提供了十二种**apply**子句，同一个节点中可以配置多条**apply**子句，但配置的多条**apply**子句不一定都会执行。影响报文转发路径的**apply**子句有五条，优先级顺序是：**apply access-vpn vpn-instance**、**apply next-hop**、**apply output-interface**、**apply default-next-hop**和**apply default-output-interface**。**apply**子句的含义以及执行优先情况等说明如 表 1-1 所示。

表1-1 **apply** 子句的含义以及执行优先情况等说明

| 子句 | 含义 | 执行优先情况/详细说明 |
|---|---|---|
| **apply precedence** | 设置IP报文的IP优先级 | 只要配置了该子句，该子句就一定会执行 |
| **apply ip-df** *df-value* | 设置IP报文的DF（Don’t Fragment，不分片）标志 | 只要配置了该子句，该子句就一定会执行 |
| **apply loadshare { next-hop \| output-interface \| default-next-hop \| default-output-interface }** | 设置指导报文转发的下一跳、出接口、缺省下一跳和缺省出接口的工作模式为负载分担模式 | 下一跳、出接口、缺省下一跳和缺省出接口的工作模式有两种：主备模式、负载分担模式<br>• 主备模式：按照配置顺序，以第一个配置（下一跳、出接口、缺省下一跳或缺省出接口）作为主用，指导报文转发。当主用失效时，按配置顺序选择后续的第一个有效配置指导报文转发<br>• 负载分担模式：按照配置顺序，逐包轮流选择有效的下一跳、出接口、缺省下一跳或缺省出接口指导报文转发<br>缺省情况下，工作模式为主备模式 |
| **apply access-vpn vpn-instance** | 设置报文在指定VPN实例中进行转发 | 报文如果匹配了其中一个VPN实例下的转发表，报文将在该VPN实例中进行转发 |
| **apply next-hop**和**apply output-interface** | 设置报文的下一跳、出接口 | 当两条子句同时配置并且都有效时，系统只会执行**apply next-hop**子句 |
| **apply default-next-hop**和**apply default-output-interface** | 设置报文的缺省下一跳、缺省出接口 | 当两条子句同时配置并且都有效时，系统只会执行**apply default-next-hop**子句<br>执行缺省下一跳和出接口的前提是：在策略中没有配置下一跳或者出接口，或者配置的下一跳和出接口无效，并且在路由表中没有找到与报文目的IP地址匹配的路由表项 |
| **apply continue** | 设置匹配成功的当前节点转发失败后继续进行后续节点的处理 | 如果当前节点中没有配置影响报文转发路径的五个**apply**子句，或者配置了这五个子句中的一个或多个，但配置的子句都失效（下一跳不可达、出接口down或者报文在指定VPN内转发失败）时，会进行下一节点的处理 |

### 3. 节点的匹配模式与节点的**if-match**子句、**apply**子句的关系

一个节点的匹配模式与这个节点的**if-match**子句、**apply**子句的关系如 表 1-2 所示。

表1-2 节点的匹配模式、**if-match** 子句、**apply** 子句三者之间的关系

| 是否满足所有 **if-match** 子句 | 节点匹配模式 | |
|---|---|---|
| | **permit**（允许模式） | **deny**（拒绝模式） |

| 是否满足所有 **if-match** 子句 | 节点匹配模式 | |
|---|---|---|
| | **permit**（允许模式） | **deny**（拒绝模式） |
| 是 | • 如果节点配置了 **apply** 子句，则执行此节点 **apply** 子句<br>  ○ 如果节点指导报文转发成功，则不再匹配下一节点<br>  ○ 如果节点指导报文转发失败且没有配置 **apply continue** 子句，则不再匹配下一节点<br>  ○ 如果节点指导报文转发失败且配置了 **apply continue** 子句，则继续匹配下一节点<br>• 如果节点没有配置 **apply** 子句，则不会执行任何动作，且不再匹配下一节点，报文将根据路由表来进行转发 | 不执行此节点**apply**子句，不再匹配下一节点，报文将根据路由表来进行转发 |
| 否 | 不执行此节点**apply**子句，继续匹配下一节点 | 不执行此节点**apply**子句，继续匹配下一节点 |

说明

如果一个节点中没有配置任何 **if-match** 子句，则认为所有报文都满足该节点的匹配规则，按照“报文满足所有 **if-match** 子句”的情况进行后续处理。

### 1.1.2 策略路由与Track联动

策略路由通过与 Track 联动，增强了应用的灵活性和对网络环境变化的动态感知能力。

策略路由可以在配置报文的下一跳、出接口、缺省下一跳、缺省出接口时与 Track 项关联，根据 Track 项的状态来动态地决定策略的可用性。策略路由配置仅在关联的 Track 项状态为 Positive 或 NotReady 时生效。关于策略路由与 Track 联动的的详细介绍和相关配置，请参见“可靠性配置指导”中的“Track”。

## 1.2 策略路由配置任务简介

**表1-3 策略路由配置任务简介**

| 配置任务 | | 说明 | 详细配置 |
|---|---|---|---|
| 配置策略 | 创建策略节点 | 必选 | 1.3.1 |
| | 配置策略节点的匹配规则 | | 1.3.2 |
| | 配置策略节点的动作 | | 1.3.3 |
| 应用策略 | 对本地报文应用策略 | 必选<br>用户可根据实际情况进行选择 | 1.4.1 |
| | 对接口转发的报文应用策略 | | 1.4.2 |

## 1.3 配置策略

### 1.3.1 创建策略节点

表1-4 创建策略节点

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 创建策略节点，并进入策略节点视图 | **policy-based-route** *policy-name* [ **deny** \| **permit** ] **node** *node-number* | 缺省情况下，没有创建策略节点 |

### 1.3.2 配置策略节点的匹配规则

表1-5 配置策略节点的匹配规则

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入策略节点视图 | **policy-based-route** *policy-name* [ **deny** \| **permit** ] **node** *node-number* | - |
| 设置ACL匹配规则 | **if-match acl** { *acl-number* \| **name** *acl-name* } | 缺省情况下，未设置ACL匹配规则 |
| 设置IP报文长度匹配规则 | **if-match packet-length** *min-len max-len* | 缺省情况下，未设置IP报文长度匹配规则 |

说明

**if-match** 子句中使用 ACL 时，对于 ACL 规则的 **permit/deny** 动作以及 **time-range** 指定的规则生效时间段等的处理机制与设备的型号有关，请以设备的实际情况为准。

### 1.3.3 配置策略节点的动作

表1-6 配置策略节点的动作

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入策略节点视图 | **policy-based-route** *policy-name* [ **deny** \| **permit** ] **node** *node-number* | - |
| 设置IP报文的IP优先级 | **apply precedence** { *type* \| *value* } | 缺省情况下，不对IP报文的优先级进行设置 |
| 设置IP报文头中的DF标志 | **apply ip-df** *df-value* | 缺省情况下，不对IP报文头中的DF标志进行设置 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 设置报文在指定VPN实例中进行转发 | **apply access-vpn vpn-instance** *vpn-instance-name*&<1-*n*> | 缺省情况下，未设置报文在指定VPN实例中进行转发<br>每个节点最多可以配置*m*个VPN实例。当满足匹配规则后，将根据第一个可用的VPN实例转发表进行转发。*m*的实际取值为6 |
| 设置报文转发的下一跳 | **apply next-hop** [ **vpn-instance** *vpn-instance-name* \| **inbound-vpn** ] { *ip-address* [ **direct** ] [ **track** *track-entry-number* ] }&<1-*n*> | 缺省情况下，未设置报文转发的下一跳<br>用户可以同时配置多个下一跳（通过一次或多次配置本命令实现），起到主备或负载分担的作用<br>每个节点最多可以配置*m*个下一跳。*m*的实际取值为16 |
| 设置指导报文转发的多个下一跳工作在负载分担模式 | **apply loadshare next-hop** | 缺省情况下，多个下一跳工作在主备模式 |
| 设置指导报文转发的出接口 | **apply output-interface** { *interface-type interface-number* [ **track** *track-entry-number* ] }&<1-*n*> | 缺省情况下，未设置指导报文转发的出接口<br>用户可以同时配置多个出接口（通过一次或多次配置本命令实现），起到主备或负载分担的作用<br>每个节点最多可以配置*m*个出接口。*m*的实际取值为16 |
| 设置指导报文转发的多个出接口工作在负载分担模式 | **apply loadshare output-interface** | 缺省情况下，多个出接口工作在主备模式 |
| 设置指导报文转发的缺省下一跳 | **apply default-next-hop** [ **vpn-instance** *vpn-instance-name* \| **inbound-vpn** ] { *ip-address* [ **direct** ] [ **track** *track-entry-number* ] }&<1-*n*> | 缺省情况下，未设置指导报文转发的缺省下一跳<br>用户可以同时配置多个缺省下一跳（通过一次或多次配置本命令实现），起到主备或负载分担的作用<br>每个节点最多可以配置*m*个缺省下一跳。*m*的实际取值为16 |
| 设置指导报文转发的多个缺省下一跳工作在负载分担模式 | **apply loadshare default-next-hop** | 缺省情况下，多个缺省下一跳工作在主备模式 |
| 设置指导报文转发的缺省出接口 | **apply default-output-interface** { *interface-type interface-number* [ **track** *track-entry-number* ] }&<1-*n*> | 缺省情况下，未设置指导报文转发的缺省出接口<br>用户可以同时配置多个缺省出接口（通过一次或多次配置本命令实现），起到主备或负载分担的作用<br>每个节点最多可以配置*m*个缺省出接口。*m*的实际取值为16 |
| 设置指导报文转发的多个缺省出接口工作在负载分担模式 | **apply loadshare default-output-interface** | 缺省情况下，多个缺省出接口工作在主备模式 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 设置匹配成功的当前节点指定转发路径失败后继续进行后续节点的处理 | **apply continue** | 缺省情况下，匹配成功的当前节点指定转发路径失败后不再进行下一节点的匹配<br>本命令仅在策略节点的匹配模式为**permit**时生效 |

# 1.4 应用策略

## 1.4.1 对本地报文应用策略

通过本配置，可以将已经配置的策略应用到本地，指导设备本身产生报文的发送。应用策略时，该策略必须已经存在，否则配置将失败。

对本地报文只能应用一个策略。应用新的策略前必须删除本地原来已经应用的策略。

若无特殊需求，建议用户不要对本地报文应用策略。

**表1-7 对本地报文应用策略**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 对本地报文应用策略 | **ip local policy-based-route** *policy-name* | 缺省情况下，对本地报文没有应用策略 |

## 1.4.2 对接口转发的报文应用策略

通过本配置，可以将已经配置的策略应用到接口，指导接口接收的所有报文的转发。应用策略时，该策略必须已经存在，否则配置将失败。

对接口转发的报文应用策略时，一个接口只能应用一个策略。应用新的策略前必须删除接口上原来已经应用的策略。

一个策略可以同时被多个接口应用。

**表1-8 对接口转发的报文应用策略**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 对接口转发的报文应用策略 | **ip policy-based-route** *policy-name* | 缺省情况下，对接口转发的报文没有应用策略 |

# 1.5 策略路由显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置策略路由后的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除策略路由的统计信息。

表1-9 策略路由显示和维护

| 操作 | 命令 |
| --- | --- |
| 显示已经配置的策略 | **display ip policy-based-route** [ **policy** *policy-name* ] |
| 显示已经应用的策略路由信息 | **display ip policy-based-route setup** |
| 显示本地策略路由的配置信息和统计信息（MSR 2600/MSR 3600） | **display ip policy-based-route local** |
| 显示本地策略路由的配置信息和统计信息（MSR 5600） | **display ip policy-based-route local** [ **slot** *slot-number* ] |
| 显示接口下转发策略路由的配置信息和统计信息（MSR 2600/MSR 3600） | **display ip policy-based-route interface** *interface-type interface-number* |
| 显示接口下转发策略路由的配置信息和统计信息（MSR 5600） | **display ip policy-based-route interface** *interface-type interface-number* [ **slot** *slot-number* ] |
| 清除策略路由的统计信息 | **reset ip policy-based-route statistics** [ **policy** *policy-name* ] |

# 1.6 策略路由典型配置举例

## 1.6.1 基于报文协议类型的本地策略路由配置举例

### 1. 组网需求

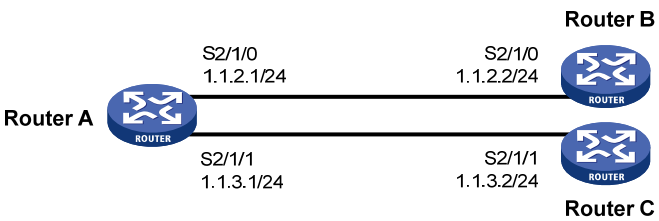
通过策略路由控制 Router A 产生的报文：

- 指定所有 TCP 报文的下一跳为 1.1.2.2；
- 其它报文仍然按照查找路由表的方式进行转发。

其中，Router A 分别与 Router B 和 Router C 直连。

### 2. 组网图

图1-1 基于报文协议类型的本地策略路由的配置举例组网图




### 3. 配置步骤

(1) 配置 Router A

# 配置 Serial 接口的 IP 地址。

```
<RouterA> system-view
[RouterA] interface serial 2/1/0
```

```
[RouterA-Serial2/1/0] ip address 1.1.2.1 24
[RouterA-Serial2/1/0] quit
[RouterA] interface serial 2/1/1
[RouterA-Serial2/1/1] ip address 1.1.3.1 24
[RouterA-Serial2/1/1] quit
```

# 定义访问控制列表 ACL 3101，用来匹配 TCP 报文。

```
[RouterA] acl number 3101
[RouterA-acl-adv-3101] rule permit tcp
[RouterA-acl-adv-3101] quit
```

# 定义 5 号节点，指定所有 TCP 报文的下一跳为 1.1.2.2。

```
[RouterA] policy-based-route aaa permit node 5
[RouterA-pbr-aaa-5] if-match acl 3101
[RouterA-pbr-aaa-5] apply next-hop 1.1.2.2
[RouterA-pbr-aaa-5] quit
```

# 在 Router A 上应用本地策略路由。

```
[RouterA] ip local policy-based-route aaa
```

(2)　配置 Router B

# 配置 Serial 接口的 IP 地址。

```
<RouterB> system-view
[RouterB] interface serial 2/1/0
[RouterB-Serial2/1/0] ip address 1.1.2.2 24
```

(3)　配置 Router C

# 配置 Serial 接口的 IP 地址。

```
<RouterC> system-view
[RouterC] interface serial 2/1/1
[RouterC-Serial2/1/1] ip address 1.1.3.2 24
```

#### 4. 验证配置

从 Router A 上通过 Telnet 方式登录 Router B（1.1.2.2/24），结果成功。

从 Router A 上通过 Telnet 方式登录 Router C（1.1.3.2/24），结果失败。

从 Router A 上 ping Router C（1.1.3.2/24），结果成功。

由于 Telnet 使用的是 TCP 协议，ping 使用的是 ICMP 协议，所以由以上结果可证明：Router A 产生的 TCP 报文的下一跳为 1.1.2.2，串口 Serial2/1/1 不发送 TCP 报文，但可以发送非 TCP 报文，策略路由设置成功。

### 1.6.2 基于报文协议类型的转发策略路由配置举例

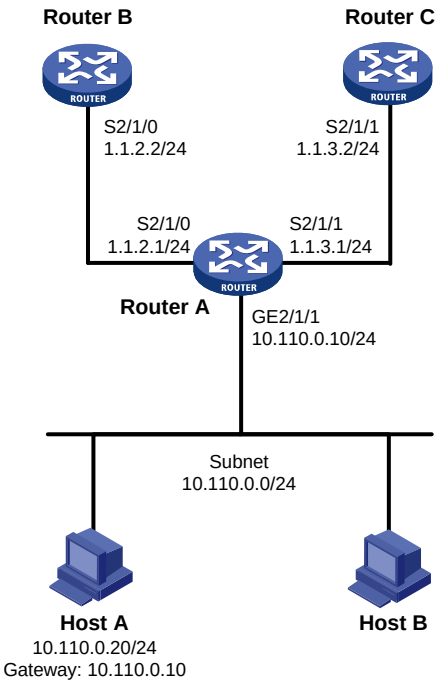
#### 1. 组网需求

通过策略路由控制从 Router A 的以太网接口 GigabitEthernet2/1/1 接收的报文：

- 指定所有 TCP 报文的下一跳为 1.1.2.2；
- 其它报文仍然按照查找路由表的方式进行转发。

## 2. 组网图

图1-2 基于报文协议类型的转发策略路由的配置举例组网图




## 3. 配置步骤

(1) 配置 Router A

# 配置 Serial 接口的 IP 地址。

```
<RouterA> system-view
[RouterA] interface serial 2/1/0
[RouterA-Serial2/1/0] ip address 1.1.2.1 24
[RouterA-Serial2/1/0] quit
[RouterA] interface serial 2/1/1
[RouterA-Serial2/1/1] ip address 1.1.3.1 24
[RouterA-Serial2/1/1] quit
```

# 定义访问控制列表 ACL 3101，用来匹配 TCP 报文。

```
[RouterA] acl number 3101
[RouterA-acl-adv-3101] rule permit tcp
[RouterA-acl-adv-3101] quit
```

# 定义 5 号节点，指定所有 TCP 报文的下一跳为 1.1.2.2。

```
[RouterA] policy-based-route aaa permit node 5
[RouterA-pbr-aaa-5] if-match acl 3101
[RouterA-pbr-aaa-5] apply next-hop 1.1.2.2
[RouterA-pbr-aaa-5] quit
```

# 在以太网接口 GigabitEthernet2/1/1 上应用转发策略路由，处理此接口接收的报文。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ip address 10.110.0.10 24
[RouterA-GigabitEthernet2/1/1] ip policy-based-route aaa
```

```
[RouterA-GigabitEthernet2/1/1] quit
```

(2) 配置 Router B

# 配置 Serial 接口的 IP 地址。

```
<RouterB> system-view
[RouterB] interface serial 2/1/0
[RouterB-Serial2/1/0] ip address 1.1.2.2 24
[RouterB-Serial2/1/0] quit
```

# 配置到网段 10.110.0.0/24 的静态路由。

```
[RouterB] ip route-static 10.110.0.0 24 1.1.2.1
```

(3) 配置 Router C

# 配置 Serial 接口的 IP 地址。

```
<RouterC> system-view
[RouterC] interface serial 2/1/1
[RouterC-Serial2/1/1] ip address 1.1.3.2 24
[RouterC-Serial2/1/1] quit
```

# 配置到网段 10.110.0.0/24 的静态路由。

```
[RouterC] ip route-static 10.110.0.0 24 1.1.3.1
```

#### 4. 验证配置

将 Host A 的 IP 地址配置为 10.110.0.20/24，网关地址配置为 10.110.0.10。

从 Host A 上通过 Telnet 方式登录 Router B，结果成功。

从 Host A 上通过 Telnet 方式登录 Router C，结果失败。

从 Host A 上 ping Router C，结果成功。

由于 Telnet 使用的是 TCP 协议，ping 使用的是 ICMP 协议，所以由以上结果可证明：从 Router A 的以太网接口 GigabitEthernet2/1/1 接收的 TCP 报文的下一跳为 1.1.2.2，串口 Serial2/1/1 不转发 TCP 报文，但可以转发非 TCP 报文，策略路由设置成功。

### 1.6.3 基于报文长度的转发策略路由配置举例

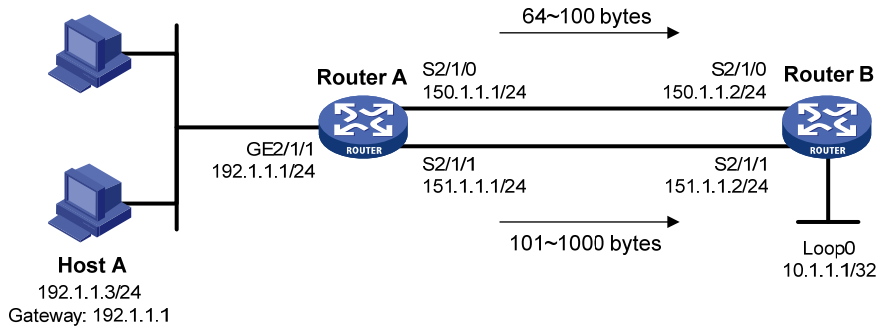
#### 1. 组网需求

通过策略路由控制从 Router A 的以太网接口 GigabitEthernet2/1/1 接收的报文：

- 长度为 64～100 字节的报文以 150.1.1.2/24 作为下一跳 IP 地址；
- 长度为 101～1000 字节的报文以 151.1.1.2/24 作为下一跳 IP 地址；
- 所有其它长度的报文都按照查找路由表的方式转发。

## 2. 组网图

图1-3 基于报文长度的转发策略路由的配置举例组网图




## 3. 配置步骤

(1) 配置 Router A

# 配置 Serial 接口的 IP 地址。

```
<RouterA> system-view
[RouterA] interface serial 2/1/0
[RouterA-Serial2/1/0] ip address 150.1.1.1 24
[RouterA-Serial2/1/0] quit
[RouterA] interface serial 2/1/1
[RouterA-Serial2/1/1] ip address 151.1.1.1 24
[RouterA-Serial2/1/1] quit
```

# 配置动态路由协议 RIP。

```
[RouterA] rip
[RouterA-rip-1] network 192.1.1.0
[RouterA-rip-1] network 150.1.0.0
[RouterA-rip-1] network 151.1.0.0
[RouterA-rip-1] quit
```

# 配置策略 lab1，将长度为 64～100 字节的报文转发到下一跳 150.1.1.2，而将长度为 101～1000 字节的报文转发到下一跳 151.1.1.2。

```
[RouterA] policy-based-route lab1 permit node 10
[RouterA-pbr-lab1-10] if-match packet-length 64 100
[RouterA-pbr-lab1-10] apply next-hop 150.1.1.2
[RouterA-pbr-lab1-10] quit
[RouterA] policy-based-route lab1 permit node 20
[RouterA-pbr-lab1-20] if-match packet-length 101 1000
[RouterA-pbr-lab1-20] apply next-hop 151.1.1.2
[RouterA-pbr-lab1-20] quit
```

# 在以太网接口 GigabitEthernet2/1/1 上应用定义的策略 lab1，处理此接口接收的报文。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ip address 192.1.1.1 24
[RouterA-GigabitEthernet2/1/1] ip policy-based-route lab1
[RouterA-GigabitEthernet2/1/1] quit
```

(2) 配置 Router B

# 配置 Serial 接口的 IP 地址。

```
<RouterB> system-view
[RouterB] interface serial 2/1/0
[RouterB-Serial2/1/0] ip address 150.1.1.2 24
[RouterB-Serial2/1/0] quit
[RouterB] interface serial 2/1/1
[RouterB-Serial2/1/1] ip address 151.1.1.2 24
[RouterB-Serial2/1/1] quit
```

# 配置 Loopback 接口的 IP 地址。

```
[RouterB] interface loopback 0
[RouterB-LoopBack0] ip address 10.1.1.1 32
[RouterB-LoopBack0] quit
```

# 配置动态路由协议 RIP。

```
[RouterB] rip
[RouterB-rip-1] network 10.0.0.0
[RouterB-rip-1] network 150.1.0.0
[RouterB-rip-1] network 151.1.0.0
[RouterB-rip-1] quit
```

### 4. 验证配置

# 在 Router A 上使用 **debugging ip policy-based-route** 命令监视策略路由。

```
<RouterA> debugging ip policy-based-route
<RouterA> terminal logging level 7
<RouterA> terminal monitor
```

# 从 Host A 上 Ping Router B 的 Loopback0，并将报文数据字段长度设为 64 字节。

```
C:\>ping –n 1 -l 64 10.1.1.1

Pinging 10.1.1.1 with 64 bytes of data:

Reply from 10.1.1.1: bytes=64 time=1ms TTL=64

Ping statistics for 10.1.1.1:
    Packets: Sent = 1, Received = 1, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 1ms, Maximum = 1ms, Average = 1ms
```

从 Router A 上显示的策略路由调试信息如下：

```
<RouterA>
*Jun  26 12:04:33:519 2012 RouterA PBR4/7/PBR Forward Info: -MDC=1; Policy:lab1, Node:
 10,match Succeeded.
*Jun  26 12:04:33:519 2012 RouterA PBR4/7/PBR Forward Info: -MDC=1; apply next-hop 150
.1.1.2.
```

以上策略路由信息显示，Router A 在接收到报文后，根据策略路由确定的下一跳为 150.1.1.2，也就是说将报文从接口 Serial2/1/0 转发出去。

# 从 Host A 上 Ping Router B 的 Loopback0，并将报文数据字段长度设为 200 字节。

```
C:\> ping –n 1 -l 200 10.1.1.1

Pinging 10.1.1.1 with 200 bytes of data:
```

```
Reply from 10.1.1.1: bytes=200 time=1ms TTL=64

Ping statistics for 10.1.1.1:
    Packets: Sent = 1, Received = 1, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 1ms, Maximum = 1ms, Average = 1ms
```

从 Router A 上显示的策略路由调试信息如下：

```
<RouterA>
*Jun  26 12:20:33:610 2012 RouterA PBR4/7/PBR Forward Info: -MDC=1; Policy:lab1, Node:
 20,match Succeeded.
*Jun  26 12:20:33:610 2012 RouterA PBR4/7/PBR Forward Info: -MDC=1; apply next-hop 151
.1.1.2.
```

以上策略路由信息显示，Router A 在接收到报文后，根据策略路由确定的下一跳为 151.1.1.2，也就是说将报文从接口 Serial2/1/1 转发出去。

## 1.6.4 基于报文源地址的转发策略路由配置举例

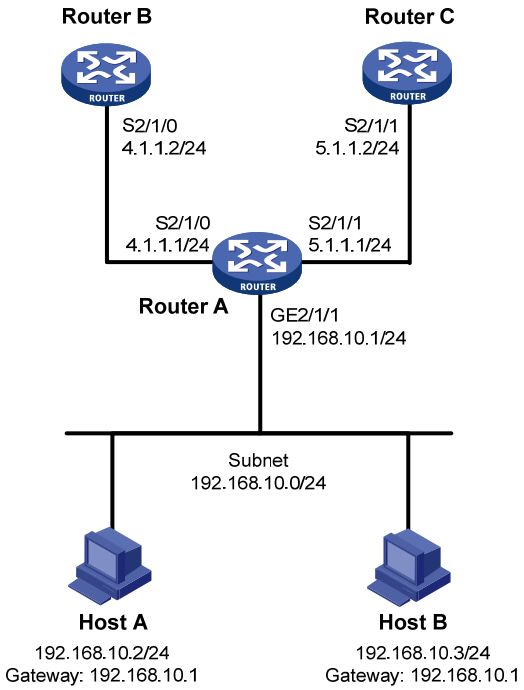
### 1. 组网需求

通过策略路由控制从 Router A 的以太网接口 GigabitEthernet2/1/1 接收的报文：

- 源地址为 192.168.10.2 的报文以 4.1.1.2/24 作为下一跳 IP 地址；
- 其它源地址的报文以 5.1.1.2/24 作为下一跳 IP 地址。

### 2. 组网图

**图1-4 基于报文源地址的转发策略路由的配置举例组网图**



### 3. 配置步骤

(1) 配置 Router A

# 配置 Serial 接口的 IP 地址。

```
<RouterA> system-view
[RouterA] interface serial 2/1/0
[RouterA-Serial2/1/0] ip address 4.1.1.1 24
[RouterA-Serial2/1/0] quit
[RouterA] interface serial 2/1/1
[RouterA-Serial2/1/1] ip address 5.1.1.1 24
[RouterA-Serial2/1/1] quit
```

# 定义访问控制列表 ACL 3000，用来匹配源地址为 192.168.10.2 的报文。

```
[RouterA] acl number 2000
[RouterA-acl-basic-2000] rule 10 permit source 192.168.10.2 0
[RouterA-acl-basic-2000] quit
```

# 定义 0 号节点，指定所有源地址为 192.168.10.2 的报文的下一跳为 4.1.1.2。

```
[RouterA] policy-based-route aaa permit node 0
[RouterA-pbr-aaa-0] if-match acl 2000
[RouterA-pbr-aaa-0] apply next-hop 4.1.1.2
[RouterA-pbr-aaa-0] quit
[RouterA] policy-based-route aaa permit node 1
[RouterA-pbr-aaa-1] apply next-hop 5.1.1.2
[RouterA-pbr-aaa-1] quit
```

# 在以太网接口 GigabitEthernet2/1/1 上应用转发策略路由，处理此接口接收的报文。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ip address 192.168.10.1 24
[RouterA-GigabitEthernet2/1/1] ip policy-based-route aaa
[RouterA-GigabitEthernet2/1/1] quit
```

(2) 配置 Router B

# 配置 Serial 接口的 IP 地址。

```
<RouterB> system-view
[RouterB] interface serial 2/1/0
[RouterB-Serial2/1/0] ip address 4.1.1.2 24
[RouterB-Serial2/1/0] quit
```

# 配置到网段 192.168.10.0/24 的静态路由。

```
[RouterB] ip route-static 192.168.10.0 24 4.1.1.1
```

(3) 配置 Router C

# 配置 Serial 接口的 IP 地址。

```
<RouterC> system-view
[RouterC] interface serial 2/1/1
[RouterC-Serial2/1/1] ip address 5.1.1.2 24
[RouterC-Serial2/1/1] quit
```

# 配置到网段 192.168.10.0/24 的静态路由。

```
[RouterC] ip route-static 192.168.10.0 24 5.1.1.1
```

### 4. 验证配置

将 Host A 的 IP 地址配置为 192.168.10.2/24，网关地址配置为 192.168.10.1；将 Host B 的 IP 地址配置为 192.168.10.3/24，网关地址配置为 192.168.10.1。

从 Host A 上 ping Router B，结果成功。

从 Host B 上 ping Router B，结果失败。

从 Host A 上 ping Router C，结果失败。

从 Host B 上 ping Router C，结果成功。

以上结果可证明：从 Router A 的以太网接口 GigabitEthernet2/1/1 接收的源地址为 192.168.10.2 的报文的下一跳为 4.1.1.2，所以 Host A 能 ping 通 Router B，源地址为 192.168.10.3 的下一跳 5.1.1.2，所以 Host B 能 ping 通 Router C，由此表明策略路由设置成功。

# 目 录

# 1 IPv6 静态路由

## 1.1 IPv6静态路由简介

静态路由是一种特殊的路由，由管理员手工配置。当网络结构比较简单时，只需配置静态路由就可以使网络正常工作。

静态路由不能自动适应网络拓扑结构的变化。当网络发生故障或者拓扑发生变化后，必须由网络管理员手工修改配置。

IPv6 静态路由与 IPv4 静态路由类似，适合于一些结构比较简单的 IPv6 网络。

## 1.2 配置IPv6静态路由

### 1.2.1 配置准备

在配置 IPv6 静态路由之前，需完成以下任务：

- 配置相关接口的物理参数
- 配置相关接口的链路层属性
- 相邻节点网络层（IPv6）可达

### 1.2.2 配置IPv6 静态路由

表1-1 配置 IPv6 静态路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置IPv6静态路由 | **ipv6 route-static** *ipv6-address prefix-length* { *interface-type interface-number* [ *next-hop-address* ] \| *next-hop-address* \| **vpn-instance** *d-vpn-instance-name nexthop-address* } [ **permanent** ] [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 二者选其一<br>缺省情况下，没有配置IPv6静态路由 |
| | **ipv6 route-static vpn-instance** *s-vpn-instance-name ipv6-address prefix-length* { *interface-type interface-number* [ *next-hop-address* ] \| *nexthop-address* [ **public** ] \| **vpn-instance** *d-vpn-instance-name nexthop-address* } [ **permanent** ] [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | |
| （可选）配置IPv6静态路由的缺省优先级 | **ipv6 route-static default-preference** *default-preference-value* | 缺省情况下，IPv6静态路由的缺省优先级为60 |
| （可选）删除所有IPv6静态路由 | **delete ipv6** [ **vpn-instance** *vpn-instance-name* ] **static-routes all** | - |

说明

- 使用 **undo ipv6 route-static** 命令可以删除一条 IPv6 静态路由，而使用 **delete ipv6 static-routes all** 命令可以删除包括缺省路由在内的所有 IPv6 静态路由。

## 1.3 配置IPv6静态路由与BFD联动

注意

路由振荡时，使能 BFD 功能可能会加剧振荡，请谨慎使用。

BFD（Bidirectional Forwarding Detection，双向转发检测）提供了一个通用的、标准化的、介质无关、协议无关的快速故障检测机制，可以为上层协议（如路由协议、MPLS 等）统一地快速检测两台路由器间双向转发路径的故障。

关于 BFD 的详细介绍，请参见“可靠性配置指导”中的“BFD”。

### 1.3.1 双向检测

双向检测，即本端和对端需要同时进行配置，通过控制报文检测两个方向上的链路状态，实现毫秒级别的链路故障检测。

双向检测支持直连下一跳和非直连下一跳。

#### 1. 直连下一跳

直连下一跳是指下一跳和本端是直连的，配置时必须指定出接口和下一跳。

表1-2 配置 IPv6 静态路由与 BFD 联动（双向检测—直连）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置静态路由与BFD联动 | **ipv6 route-static** *ipv6-address prefix-length interface-type interface-number next-hop-address* **bfd control-packet** [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 二者选其一<br>缺省情况下，没有配置IPv6静态路由与BFD联动 |
| | **ipv6 route-static vpn-instance** *s-vpn-instance-name ipv6-address prefix-length interface-type interface-number next-hop-address* **bfd control-packet** [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | |

#### 2. 非直连下一跳

非直连下一跳是指下一跳和本端不是直连的，中间还有其它设备。配置时必须指定下一跳和 BFD 源 IPv6 地址。

表1-3 配置 IPv6 静态路由与 BFD 联动（双向检测—非直连）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置静态路由与BFD联动 | **ipv6 route-static** *ipv6-address prefix-length* { *next-hop-address* **bfd control-packet bfd-source** *ipv6-address* \| **vpn-instance** *d-vpn-instance-name next-hop-address* **bfd control-packet bfd-source** *ipv6-address* } [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 二者选其一<br>缺省情况下，没有配置IPv6静态路由与BFD联动 |
| | **ipv6 route-static vpn-instance** *s-vpn-instance-name ipv6-address prefix-length* { *next-hop-address* **bfd control-packet bfd-source** *ipv6-address* \| **vpn-instance** *d-vpn-instance-name next-hop-address* **bfd control-packet bfd-source** *ipv6-address* } [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | |

## 1.3.2 单跳检测

单跳检测，即只需要本端进行配置，通过 echo 报文检测链路的状态。echo 报文的目的地址为本端接口地址，发送给下一跳设备后会直接转发回本端。这里所说的“单跳”是 IPv6 的一跳。

表1-4 配置静态路由与 BFD 联动（单跳检测）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置echo报文的源IPv6地址 | **bfd echo-source-ipv6** *ipv6-address* | 缺省情况下，没有配置echo报文的源IPv6地址<br>需要注意的是，echo报文源IPv6地址仅支持全球单播地址<br>本命令的详细情况请参见“可靠性命令参考”中的“BFD” |
| 配置静态路由与BFD联动 | **ipv6 route-static** *ipv6-address prefix-length interface-type interface-number next-hop-address* **bfd echo-packet** [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | 二者选其一<br>缺省情况下，没有配置IPv6静态路由与BFD联动<br>需要注意的是，下一跳IPv6地址必须为全球单播地址 |
| | **ipv6 route-static vpn-instance** *s-vpn-instance-name ipv6-address prefix-length interface-type interface-number next-hop-address* **bfd echo-packet** [ **preference** *preference-value* ] [ **tag** *tag-value* ] [ **description** *description-text* ] | |

提示

IPv6 静态路由的出接口为处于 SPOOFING 状态时，不能使用 BFD 进行检测。

## 1.4 IPv6静态路由显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令查看 IPv6 静态路由配置的运行情况并检验配置结果。

表1-5 IPv6 静态路由显示和维护

| 操作 | 命令 |
| --- | --- |
| 查看IPv6静态路由表信息（本命令的详细情况请参见“三层技术-IP路由命令参考”中的“IP路由基础”） | **display ipv6 routing-table protocol static** [ **inactive** \| **verbose** ] |
| 显示IPv6静态路由下一跳信息 | **display ipv6 route-static nib** [ *nib-id* ] [ **verbose** ] |
| 显示IPv6静态路由表信息 | **display ipv6 route-static routing-table** [ **vpn-instance** *vpn-instance-name* ] [ *ipv6-address prefix-length* ] |

## 1.5 IPv6静态路由典型配置举例

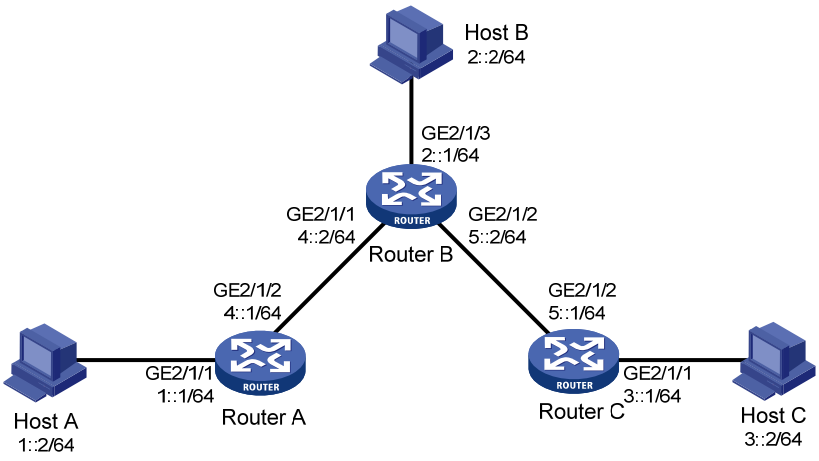
### 1.5.1 IPv6 静态路由基本功能配置举例

#### 1. 组网要求

要求各路由器之间配置 IPv6 静态路由后，可以使所有主机和路由器之间互通。

#### 2. 组网图

图1-1 IPv6 静态路由基本功能配置组网图



### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 IPv6 静态路由

# 在 Router A 上配置 IPv6 缺省路由。

```
<RouterA> system-view
[RouterA] ipv6 route-static :: 0 4::2
```

# 在 Router B 上配置两条 IPv6 静态路由。

```
<RouterB> system-view
[RouterB] ipv6 route-static 1:: 64 4::1
[RouterB] ipv6 route-static 3:: 64 5::1
```

# 在 Router C 上配置 IPv6 缺省路由。

```
<RouterC> system-view
[RouterC] ipv6 route-static :: 0 5::2
```

(3) 配置主机地址和网关

根据组网图配置好各主机的 IPv6 地址，并将 Host A 的缺省网关配置为 1::1，Host B 的缺省网关配置为 2::1，Host C 的缺省网关配置为 3::1。

### 4. 验证配置

# 查看 Router A 的 IPv6 静态路由信息。

```
[RouterA] display ipv6 routing-table protocol static

Summary Count : 1

Static Routing table Status : <Active>
Summary Count : 1

Destination: ::                                         Protocol  : Static
NextHop    : 4::2                                       Preference: 60
Interface  : GE2/1/2                                    Cost      : 0

Static Routing table Status : <Inactive>
Summary Count : 0
```

# 查看 Router B 的 IPv6 静态路由信息。

```
[RouterB] display ipv6 routing-table protocol static

Summary Count : 2

Static Routing table Status : <Active>
Summary Count : 2

Destination: 1::/64                                     Protocol  : Static
NextHop    : 4::1                                       Preference: 60
Interface  : GE2/1/1                                    Cost      : 0

Destination: 3::/64                                     Protocol  : Static
NextHop    : 5::1                                       Preference: 60
```

```
Interface   : GE2/1/2                                    Cost      : 0

Static Routing table Status : <Inactive>
Summary Count : 0
```

# 使用 Ping 进行验证。

```
[RouterA] ping ipv6 3::1
Ping6(104=40+8+56 bytes) 4::1 --> 3::1, press CTRL_C to break
56 bytes from 3::1, icmp_seq=0 hlim=62 time=0.700 ms
56 bytes from 3::1, icmp_seq=1 hlim=62 time=0.351 ms
56 bytes from 3::1, icmp_seq=2 hlim=62 time=0.338 ms
56 bytes from 3::1, icmp_seq=3 hlim=62 time=0.373 ms
56 bytes from 3::1, icmp_seq=4 hlim=62 time=0.316 ms

--- Ping6 statistics for 3::1 ---
5 packet(s) transmitted, 5 packet(s) received, 0.0% packet loss
round-trip min/avg/max/std-dev = 0.316/0.416/0.700/0.143 ms
```

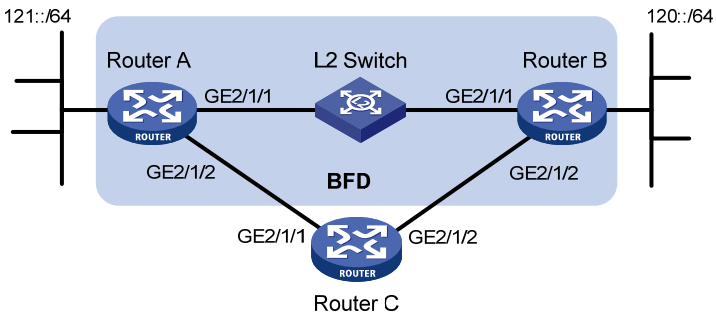
## 1.5.2 配置IPv6 静态路由与BFD联动（直连）

### 1. 组网需求

- 在 Router A 上配置 IPv6 静态路由可以到达 120::/64 网段，在 Router B 上配置 IPv6 静态路由可以到达 121::/64 网段，并使能 BFD 检测功能。
- 在 Router C 上配置 IPv6 静态路由可以到达 120::/64 网段和 121::/64 网段。
- 当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时，BFD 能够快速感知，并且切换到 Router C 进行通信。

### 2. 组网图

图1-2 IPv6 静态路由与 BFD 联动（直连）配置组网图




| 设备 | 接口 | IPv6地址 | 设备 | 接口 | IPv6地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 12::1/64 | Router B | GE2/1/1 | 12::2/64 |
| | GE2/1/2 | 10::102/64 | | GE2/1/2 | 13::1/64 |
| Router C | GE2/1/1 | 10::100/64 | | | |
| | GE2/1/2 | 13::2/64 | | | |

### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 IPv6 静态路由和 BFD

# 在 Router A 上配置 IPv6 静态路由，并使能 BFD 检测功能，使用双向检测方式。

```
<RouterA> system-view
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterA-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterA-GigabitEthernet2/1/1] bfd detect-multiplier 9
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] ipv6 route-static 120:: 64 gigabitethernet 2/1/1 FE80::2E0:FCFF:FE58:123E bfd
control-packet
[RouterA] ipv6 route-static 120:: 64 10::100 preference 65
[RouterA] quit
```

# 在 Router B 上配置静态路由，并使能 BFD 检测功能，使用双向检测方式。

```
<RouterB> system-view
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterB-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterB-GigabitEthernet2/1/1] bfd detect-multiplier 9
[RouterB-GigabitEthernet2/1/1] quit
[RouterB] ipv6 route-static 121:: 64 gigabitethernet 2/1/1 FE80::2A0:FCFF:FE00:580A bfd
control-packet
[RouterB] ipv6 route-static 121:: 64 13::2 preference 65
[RouterB] quit
```

# 在 Router C 上配置静态路由。

```
<RouterC> system-view
[RouterC] ipv6 route-static 120:: 64 13::1
[RouterC] ipv6 route-static 121:: 64 10::102
```

### 4. 验证配置

下面以 Router A 为例，Router B 和 Router A 类似，不再赘述。

# 查看 BFD 会话，可以看到 BFD 会话已经创建。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv6 Session Working Under Ctrl Mode:

       Local Discr: 513                 Remote Discr: 33
         Source IP: FE80::2A0:FCFF:FE00:580A（Router A 接口 GigabitEthernet2/1/1 的链路本地地
址）
    Destination IP: FE80::2E0:FCFF:FE58:123E（Router B 接口 GigabitEthernet2/1/1 的链路本地地
址）
     Session State: Up                      Interface: GE2/1/1
         Hold Time: 2012ms
```

# 查看 IPv6 静态路由，可以看到 Router A 经过 L2 Switch 到达 Router B。

```
<RouterA> display ipv6 routing-table protocol static

Summary Count : 1
```

```
Static Routing table Status : <Active>
Summary Count : 1

Destination: 120::/64                                   Protocol  : Static
NextHop    : 12::2                                      Preference: 60
Interface  : GE2/1/1                                    Cost      : 0

Static Routing table Status : <Inactive>
Summary Count : 0
```

当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时：

# 查看静态路由，可以看到 Router A 经过 Router C 到达 Router B。

```
<RouterA> display ipv6 routing-table protocol static

Summary Count : 1

Static Routing table Status : <Active>
Summary Count : 1

Destination: 120::/64                                   Protocol  : Static
NextHop    : 10::100                                    Preference: 65
Interface  : GE2/1/2                                    Cost      : 0

Static Routing table Status : <Inactive>
Summary Count : 0
```

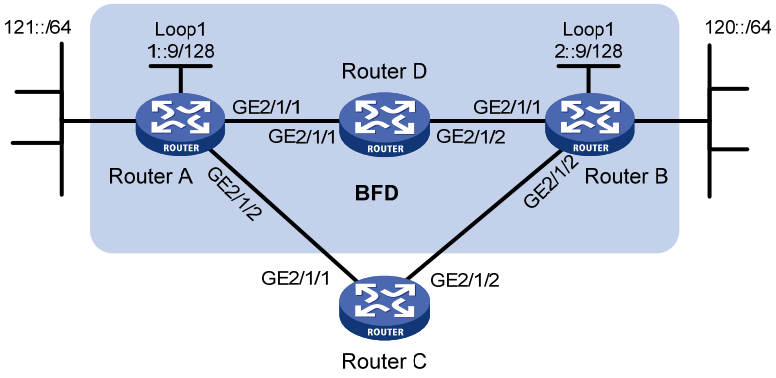
### 1.5.3 配置IPv6 静态路由与BFD联动（非直连）

#### 1. 组网需求

- 在 Router A 上配置 IPv6 静态路由可以到达 120::/64 网段，在 Router B 上 IPv6 配置静态路由可以到达 121::/64 网段，并使能 BFD 检测功能。
- 在 Router C 和 Router D 上配置 IPv6 静态路由可以到达 120::/64 网段和 121::/64 网段。
- Router A 存在到 Router B 的接口 Loopback1（2::9/128）的路由，出接口为 GigabitEthernet2/1/1；Router B 存在到 Router A 的接口 Loopback1（1::9/128）的路由，出接口为 GigabitEthernet2/1/1；Router D 存在到 1::9/128 的路由，出接口为 GigabitEthernet2/1/1，存在到 2::9/128 的路由，出接口为 GigabitEthernet2/1/2。
- 当 Router A 和 Router B 通过 Router D 通信的链路出现故障时，BFD 能够快速感知，并且切换到 Router C 进行通信。

## 2. 组网图

图1-3 IPv6 静态路由与 BFD 联动（非直连）配置组网图




| 设备 | 接口 | IPv6地址 | 设备 | 接口 | IPv6地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 12::1/64 | Router B | GE2/1/1 | 11::2/64 |
| | GE2/1/2 | 10::102/64 | | GE2/1/2 | 13::1/64 |
| | Loop1 | 1::9/128 | | Loop1 | 2::9/128 |
| Router C | GE2/1/1 | 10::100/64 | Router D | GE2/1/1 | 12::2/64 |
| | GE2/1/2 | 13::2/64 | | GE2/1/2 | 11::1/64 |

## 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 IPv6 静态路由和 BFD

# 在 Router A 上配置 IPv6 静态路由，并使能 BFD 检测功能，使用双向检测方式。

```
<RouterA> system-view
[RouterA] bfd multi-hop min-transmit-interval 500
[RouterA] bfd multi-hop min-receive-interval 500
[RouterA] bfd multi-hop detect-multiplier 9
[RouterA] ipv6 route-static 120:: 64 2::9 bfd control-packet bfd-source 1::9
[RouterA] ipv6 route-static 120:: 64 10::100 preference 65
[RouterA] quit
```

# 在 Router B 上配置 IPv6 静态路由，并使能 BFD 检测功能，使用双向检测方式。

```
<RouterB> system-view
[RouterB] bfd multi-hop min-transmit-interval 500
[RouterB] bfd multi-hop min-receive-interval 500
[RouterB] bfd multi-hop detect-multiplier 9
[RouterB] ipv6 route-static 121:: 64 1::9 bfd control-packet bfd-source 2::9
[RouterB] ipv6 route-static 121:: 64 13::2 preference 65
[RouterB] quit
```

# 在 Router C 上配置静态路由。

```
<RouterC> system-view
[RouterC] ipv6 route-static 120:: 64 13::1
[RouterC] ipv6 route-static 121:: 64 10::102
```

# 在 Router D 上配置静态路由。

```
<RouterC> system-view
[RouterC] ipv6 route-static 120:: 64 11::2
[RouterC] ipv6 route-static 121:: 64 12::1
```

### 4. 验证配置

下面以 Router A 为例，Router B 和 Router A 类似，不再赘述。

# 查看 BFD 会话，可以看到 BFD 会话已经创建。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv6 Session Working Under Ctrl Mode:

       Local Discr: 513                 Remote Discr: 33
         Source IP: FE80::1:1B49（Router A 接口 Loopback1 的链路本地地址）
    Destination IP: FE80::1:1B49（Router B 接口 Loopback1 的链路本地地址）
     Session State: Up                      Interface: N/A
         Hold Time: 2012ms
```

# 查看 IPv6 静态路由，可以看到 Router A 经过 Router D 到达 Router B。

```
<RouterA> display ipv6 routing-table protocol static

Summary Count : 1

Static Routing table Status : <Active>
Summary Count : 1

Destination: 120::/64                                   Protocol  : Static
NextHop    : 2::9                                       Preference: 60
Interface  : GE2/1/1                                    Cost      : 0

Static Routing table Status : <Inactive>
Summary Count : 0
```

当 Router A 和 Router B 通过 Router D 通信的链路出现故障时：

# 查看 IPv6 静态路由，可以看到 Router A 经过 Router C 到达 Router B。

```
<RouterA> display ipv6 routing-table protocol static

Summary Count : 1

Static Routing table Status : <Active>
Summary Count : 1

Destination: 120::/64                                   Protocol  : Static
NextHop    : 10::100                                    Preference: 65
Interface  : GE2/1/2                                    Cost      : 0

Static Routing table Status : <Inactive>
Summary Count : 0
```

# 2 IPv6 缺省路由

## 2.1 IPv6缺省路由简介

IPv6 缺省路由是在路由器没有找到匹配的 IPv6 路由表项时使用的路由。

IPv6 缺省路由有两种生成方式：

- 第一种是网络管理员手工配置。配置请参见 表 1-1，指定的目的地址为::/0（前缀长度为 0）。
- 第二种是动态路由协议生成（如 OSPFv3、IPv6 IS-IS 和 RIPng），由路由能力比较强的路由器将 IPv6 缺省路由发布给其它路由器，其它路由器在自己的路由表里生成指向那台路由器的缺省路由。配置请参见各个路由协议手册。

# 目 录

# 1 RIPng

## 1.1 简介

RIPng（RIP next generation，下一代 RIP 协议）是对原来的 IPv4 网络中 RIP-2 协议的扩展。大多数 RIP 的概念都可以用于 RIPng。

为了在 IPv6 网络中应用，RIPng 对原有的 RIP 协议进行了如下修改：

- UDP 端口号：使用 UDP 的 521 端口发送和接收路由信息。
- 组播地址：使用 FF02::9 作为链路本地范围内的 RIPng 路由器组播地址。
- 前缀长度：目的地址使用 128 比特的前缀长度。
- 下一跳地址：使用 128 比特的 IPv6 地址。
- 源地址：使用链路本地地址 FE80::/10 作为源地址发送 RIPng 路由信息更新报文。

### 1.1.1 RIPng工作机制

RIPng 协议是基于距离矢量（Distance-Vector）算法的协议。它通过 UDP 报文交换路由信息，使用的端口号为 521。

RIPng 使用跳数来衡量到达目的地址的距离（也称为度量值或开销）。在 RIPng 中，从一个路由器到其直连网络的跳数为 0，通过与其相连的路由器到达另一个网络的跳数为 1，其余以此类推。当跳数大于或等于 16 时，目的网络或主机就被定义为不可达。

RIPng 每 30 秒发送一次路由更新报文。如果在 180 秒内没有收到网络邻居的路由更新报文，RIPng 将从邻居学到的所有路由标识为不可达。如果再过 120 秒内仍没有收到邻居的路由更新报文，RIPng 将从路由表中删除这些路由。

为了提高性能并避免形成路由环路，RIPng 既支持水平分割也支持毒性逆转。此外，RIPng 还可以从其它的路由协议引入路由。

每个运行 RIPng 的路由器都管理一个路由数据库，该路由数据库包含了到所有可达目的地的路由项，这些路由项包含下列信息：

- 目的地址：主机或网络的 IPv6 地址。
- 下一跳地址：为到达目的地，需要经过的相邻路由器的接口 IPv6 地址。
- 出接口：转发 IPv6 报文通过的出接口。
- 度量值：本路由器到达目的地的开销。
- 路由时间：从路由项最后一次被更新到现在所经过的时间，路由项每次被更新时，路由时间重置为 0。
- 路由标记（Route Tag）：用于标识外部路由，以便在路由策略中根据 Tag 对路由进行灵活的控制。关于路由策略的详细信息，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

### 1.1.2 RIPng报文

RIPng 有两种报文：Request 报文和 Response 报文。

当 RIPng 路由器启动后或者需要更新部分路由表项时，便会发出 Request 报文，向邻居请求需要的路由信息。通常情况下以组播方式发送 Request 报文。

Response 报文包含本地路由表的信息，一般在下列情况下产生：

- 对某个 Request 报文进行响应
- 作为更新报文周期性地发出
- 在路由发生变化时触发更新

收到 Request 报文的 RIPng 路由器会以 Response 报文形式发回给请求路由器。

收到 Response 报文的路由器会更新自己的 RIPng 路由表。为了保证路由的准确性，RIPng 路由器会对收到的 Response 报文进行有效性检查，比如源 IPv6 地址是否是链路本地地址，端口号是否正确等，没有通过检查的报文会被忽略。

### 1.1.3 协议规范

与 RIPng 相关的规范有：

- RFC 2080：RIPng for IPv6
- RFC 2081：RIPng Protocol Applicability Statement

## 1.2 RIPng配置任务简介

表1-1 RIPng 配置任务简介

| 配置任务 | | 说明 | 详细配置 |
|---|---|---|---|
| 配置RIPng的基本功能 | | 必选 | 1.3 |
| 配置RIPng路由特性 | 配置接口附加度量值 | 可选 | 1.4.1 |
| | 配置RIPng路由聚合 | 可选 | 1.4.2 |
| | 配置RIPng发布缺省路由 | 可选 | 1.4.3 |
| | 配置RIPng对接收/发布的路由进行过滤 | 可选 | 1.4.4 |
| | 配置RIPng协议优先级 | 可选 | 1.4.5 |
| | 配置RIPng引入外部路由 | 可选 | 1.4.6 |
| 调整和优化RIPng网络 | 配置RIPng定时器 | 可选 | 1.5.1 |
| | 配置水平分割和毒性逆转 | 可选 | 1.5.2 |
| | 配置RIPng报文的零域检查 | 可选 | 1.5.3 |
| | 配置最大等价路由条数 | 可选 | 1.5.4 |
| 配置RIPng GR | | 可选 | 1.6 |
| 配置RIPng IPsec安全框架 | | 可选 | 1.7 |

## 1.3 配置RIPng的基本功能

在配置 RIPng 基本功能之前，需要配置接口的网络层地址，使相邻节点的网络层可达。

**表1-2 配置 RIPng 的基本功能**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 创建RIPng进程，并进入RIPng视图 | **ripng** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 缺省情况下，没有RIPng进程在运行 |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 在接口上使能RIPng路由协议 | **ripng** *process-id* **enable** | 缺省情况下，接口禁用RIPng路由协议<br>如果接口没有使能RIPng，那么RIPng进程在该接口上既不发送也不接收RIPng路由 |

## 1.4 配置RIPng路由特性

在配置 RIPng 的路由特性之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点的网络层可达
- 配置 RIPng 的基本功能

### 1.4.1 配置接口附加度量值

附加度量值是在 RIPng 路由原来度量值的基础上所增加的度量值（跳数），包括发送附加度量值和接收附加度量值。

- 发送附加度量值：不会改变路由表中的路由度量值，仅当接口发送 RIPng 路由信息时才会添加到发送路由上。
- 接收附加度量值：会影响接收到的路由度量值，接口接收到一条合法的 RIPng 路由时，在将其加入路由表前会把附加度量值加到该路由上。

**表1-3 配置接口附加度量值**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 设置接口接收RIPng路由时的附加度量值 | **ripng metricin** *value* | 缺省情况下，接口接收RIPng路由时的附加度量值为0 |
| 设置接口发送RIPng路由时的附加度量值 | **ripng metricout** *value* | 缺省情况下，接口发送RIPng路由时的附加度量值为1 |

### 1.4.2 配置RIPng路由聚合

RIPng 的路由聚合是在接口上实现的，在接口上配置路由聚合，此时可以将 RIPng 要在这个接口上发布出去的路由按最长匹配原则聚合后发布出去。

RIPng 路由聚合可提高网络的可扩展性和效率，缩减路由表。

RIPng 将多条路由聚合成一条路由时，聚合路由的 Metric 值将取所有路由 Metric 的最小值。

例如，RIPng 从接口发布出去的路由有两条：11:11:11::24 Metric=2 和 11:11:12::34 Metric=3，在此接口上配置的聚合路由为 11::0/16，则最终发布出去的路由为 11::0/16 Metric=2。

表1-4 配置 RIPng 路由聚合

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置RIPng在接口发布聚合的IPv6地址，并指定被聚合的路由的IPv6前缀 | **ripng summary-address** *ipv6-address prefix-length* | 缺省情况下，没有配置RIPng在接口发布聚合的IPv6地址 |

### 1.4.3 配置RIPng发布缺省路由

表1-5 配置 RIPng 发布缺省路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置RIPng发布缺省路由 | **ripng default-route** { **only** \| **originate** } [ **cost** *cost* ] | 缺省情况下，RIPng进程不发布缺省路由<br>缺省路由将被强制通过指定接口的路由更新报文发布出去，该路由的发布不考虑其是否已经存在于本设备的IPv6路由表中 |

### 1.4.4 配置RIPng对接收/发布的路由进行过滤

用户可通过使用 IPv6 ACL 和 IPv6 前缀列表对接收到的路由信息进行过滤，只有通过过滤的路由才能被加入到 RIPng 路由表；此外，还可对本机所有要发布的路由进行过滤，包括从其它路由协议引入的路由和从邻居学到的 RIPng 路由，只有通过过滤的路由才能被发布给 RIPng 邻居。

表1-6 配置 RIPng 对接收/发布的路由进行过滤

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIPng视图 | **ripng** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 对接收的路由信息进行过滤 | **filter-policy** { *acl6-number* \| **prefix-list** *prefix-list-name* } **import** | 缺省情况下，RIPng不对接收的路由信息进行过滤 |
| 对发布的路由信息进行过滤 | **filter-policy** { *acl6-number* \| **prefix-list** *prefix-list-name* } **export** [ *protocol* [ *process-id* ] ] | 缺省情况下，RIPng不对发布的路由信息进行过滤 |

### 1.4.5 配置RIPng协议优先级

任何路由协议都具备特有的协议优先级，在设备进行路由选择时能够在不同的协议中选择最佳路由。可以手工设置 RIPng 协议的优先级，设置的值越小，其优先级越高。

表1-7 配置 RIPng 协议优先级

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIPng视图 | **ripng** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIPng路由的优先级 | **preference** [ **route-policy** *route-policy-name* ] *value* | 缺省情况下，RIPng路由的优先级为100 |

### 1.4.6 配置RIPng引入外部路由

表1-8 配置 RIPng 引入外部路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIPng视图 | **ripng** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 引入外部路由 | **import-route** *protocol* [ *process-id* ] [ **allow-ibgp** ] [ **allow-direct** \| **cost** *cost* \| **route-policy** *route-policy-name* ] * | 缺省情况下，RIPng不引入其它路由 |
| （可选）配置引入路由的缺省度量值 | **default cost** *cost* | 缺省情况下，引入路由的缺省度量值为0 |

## 1.5 调整和优化RIPng网络

本节将介绍如何调整和优化 RIPng 网络的性能，以及在特殊网络环境中某些 RIPng 特性的应用，在调整和优化 RIPng 网络之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点的网络层可达
- 配置 RIPng 的基本功能

### 1.5.1 配置RIPng定时器

用户可通过调节 RIPng 定时器来调整 RIPng 路由协议的性能，以满足网络需要。

在配置 RIPng 定时器时需要注意，定时器值的调整应考虑网络的性能，并在所有运行 RIPng 的路由器上进行统一配置，避免增加不必要的网络流量。

**表1-9 配置 RIPng 定时器**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 进入RIPng视图 | **ripng** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIPng定时器的值 | **timers** { **garbage-collect** *garbage-collect-value* \| **suppress** *suppress-value* \| **timeout** *timeout-value* \| **update** *update-value* } * | 缺省情况下，Update定时器的值为30秒，Timeout定时器的值为180秒，Suppress定时器的值为120秒，Garbage-collect定时器的值为120秒 |

### 1.5.2 配置水平分割和毒性逆转

提示

如果同时配置了水平分割和毒性逆转，则只有毒性逆转功能生效。

#### 1. 配置水平分割

提示

通常情况下，为了防止路由环路的出现，水平分割都是必要的，因此，建议不要关闭水平分割。

配置水平分割可以使得从一个接口学到的路由不能通过此接口向外发布，用于避免相邻路由器间的路由环路。

**表1-10 配置水平分割**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能水平分割功能 | **ripng split-horizon** | 缺省情况下，水平分割功能处于使能状态 |

### 2. 配置毒性逆转

配置毒性逆转可以使得从一个接口学到的路由还可以从这个接口向外发布，但此时这些路由的度量值已设置为 16，即不可达。

表1-11 配置毒性逆转

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能毒性逆转功能 | **ripng poison-reverse** | 缺省情况下，毒性逆转功能处于关闭状态 |

## 1.5.3 配置RIPng报文的零域检查

RIPng 报文头部中的一些字段必须配置为 0，也称为零域。使能 RIPng 报文的零域检查功能后，如果报文头部零域中的值不为零，这些报文将被丢弃，不做处理。如果能确保所有报文都是可信任的，则不需要进行该项检查，以节省 CPU 处理时间。

表1-12 配置 RIPng-1 报文的零域检查

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 启动 RIPng 并进入 RIPng 视图 | **ripng** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能对RIPng报文头部的零域检查功能 | **checkzero** | 缺省情况下，RIPng进行零域检查操作 |

## 1.5.4 配置最大等价路由条数

表1-13 配置最大等价路由条数

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 启动 RIPng 并进入 RIPng 视图 | **ripng** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIPng最大等价路由条数 | **maximum load-balancing** *number* | 缺省情况下，RIPng支持的最大等价路由条数为32。 |

## 1.6 配置RIPng GR

GR（Graceful Restart，平滑重启）是一种在协议重启或主备倒换时 RIPng 进行平滑重启，保证转发业务不中断的机制。

GR 有两个角色：

- GR Restarter：发生协议重启或主备倒换事件且具有 GR 能力的设备。
- GR Helper：和 GR Restarter 具有邻居关系，协助完成 GR 流程的设备。

在普通的路由协议重启的情况下，路由器需要重新学习 RIPng 路由，并更新 FIB 表，此时会引起网络暂时的中断，基于 RIPng 的 GR 可以解决这个问题。

应用了 GR 特性的设备向外发送 RIPng 全部路由表请求报文，重新从邻居处学习 RIPng 路由，在此期间 FIB 表不变化。在路由协议重启完毕后，设备将重新学到的 RIPng 路由下刷给 FIB 表，使该设备的路由信息恢复到重启前的状态。

在作为 GR Restarter 的设备上进行以下配置。启动了 RIPng 的设备缺省就是 GR Helper。

表1-14 配置 RIPng GR

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 启动 RIPng 并进入 RIPng 视图 | **ripng** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 使能 RIPng 协议的 GR 能力 | **graceful-restart** | 缺省情况下，RIPng 协议的 GR 能力处于关闭状态 |

## 1.7 配置RIPng IPsec安全框架

在安全性要求较高的网络环境中，可以通过配置基于 IPsec 安全框架的认证方式来对 RIPng 报文进行有效性检查和验证。IPsec 安全框架的具体情况请参见“安全配置指导”中的“IPsec”。

设备在发送的报文中会携带配置好的 IPsec 安全框架的 SPI（Security Parameter Index，安全参数索引）值，接收报文时通过 SPI 值进行 IPsec 安全框架匹配：只有安全框架匹配的报文才能接收；否则将不会接收报文，从而不能正常建立邻居和学习路由。

RIPng 支持在进程和接口下配置 IPsec 安全框架。进程下配置的 IPsec 安全框架对该进程下的所有报文有效，接口下的 IPsec 安全框架只对接口下的报文有效。当接口和接口所在进程均配置了 IPsec 安全框架时，接口下的配置生效。

表1-15 配置 RIPng IPsec 安全框架（RIPng 进程）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入RIPng视图 | **ripng** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置RIPng进程应用IPsec安全框架 | **enable ipsec-profile** *profile-name* | 缺省情况下，RIPng进程没有应用IPsec安全框架 |

表1-16 配置 RIPng IPsec 安全框架（接口）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置使能了RIPng的接口上应用IPsec安全框架 | **ripng ipsec-profile** *profile-name* | 缺省情况下，RIPng接口没有应用IPsec安全框架 |

## 1.8 RIPng显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 RIPng 的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以重启 RIPng 进程或清除指定 RIPng 进程的统计信息。

表1-17 RIPng 显示和维护

| 操作 | 命令 |
| --- | --- |
| 显示RIPng进程的配置信息 | **display ripng** [ *process-id*] |
| 显示RIPng发布数据库中的路由 | **display ripng** *process-id* **database** [ *ipv6-address prefix-length* ] |
| 显示指定RIPng进程的路由信息 | **display ripng** *process-id* **route** [ *ipv6-address prefix-length* [ **verbose** ] \| **peer** *ipv6-address* \| **statistics** ] |
| 显示RIPng的接口信息 | **display ripng** *process-id* **interface** [ *interface-type interface-number* ] |
| 重启指定RIPng进程 | **reset ripng** *process-id* **process** |
| 清除RIPng进程的统计信息 | **reset ripng** *process-id* **statistics** |

## 1.9 RIPng典型配置举例

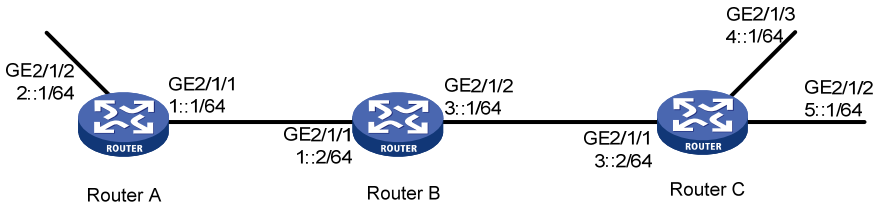
### 1.9.1 配置RIPng基本功能

#### 1. 组网需求

- Router A、Router B 和 Router C 相连并通过 RIPng 来学习网络中的 IPv6 路由信息。
- 在 Router B 上对接收的 Router A 的路由（2::/64）进行过滤，使其不加入到 Router B 的 RIPng 进程的路由表中，发布给 Router A 的路由只有（4::/64）。

## 2. 组网图

图1-1 RIPng 基本功能配置组网图




## 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 RIPng 的基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] ripng 1
[RouterA-ripng-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ripng 1 enable
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] ripng 1 enable
[RouterA-GigabitEthernet2/1/2] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] ripng 1
[RouterB-ripng-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ripng 1 enable
[RouterB-GigabitEthernet2/1/1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] ripng 1 enable
[RouterB-GigabitEthernet2/1/2] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] ripng 1
[RouterC-ripng-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ripng 1 enable
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] ripng 1 enable
[RouterC-GigabitEthernet2/1/2] quit
[RouterC] interface gigabitethernet 2/1/3
[RouterC-GigabitEthernet2/1/3] ripng 1 enable
```

```
[RouterC-GigabitEthernet2/1/3] quit
```

# 查看 Router B 的 RIPng 路由表。

```
[RouterB] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect, D – Direct
                O - Optimal, F - Flush to RIB
 ------------------------------------------------------------------

 Peer FE80::20F:E2FF:FE23:82F5 on GigabitEthernet2/1/1
 Destination 1::/64,
     via FE80::20F:E2FF:FE23:82F5, cost 1, tag 0, AOF, 6 secs
 Destination 2::/64,
     via FE80::20F:E2FF:FE23:82F5, cost 1, tag 0, AOF, 6 secs

 Peer FE80::20F:E2FF:FE00:100  on GigabitEthernet2/1/2
 Destination 3::/64,
     via FE80::20F:E2FF:FE00:100, cost 1, tag 0, AOF, 11 secs
 Destination 4::/64,
     via FE80::20F:E2FF:FE00:100, cost 1, tag 0, AOF, 11 secs
 Destination 5::/64,
     via FE80::20F:E2FF:FE00:100, cost 1, tag 0, AOF, 11 secs
```

# 查看 Router A 的 RIPng 路由表。

```
[RouterA] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect, D – Direct
                O - Optimal, F - Flush to RIB
 ------------------------------------------------------------------

 Peer FE80::200:2FF:FE64:8904 on GigabitEthernet2/1/1
 Destination 1::/64,
     via FE80::200:2FF:FE64:8904, cost 1, tag 0, AOF, 31 secs
 Destination 3::/64,
     via FE80::200:2FF:FE64:8904, cost 1, tag 0, AOF, 31 secs
 Destination 4::/64,
     via FE80::200:2FF:FE64:8904, cost 2, tag 0, AOF, 31 secs
 Destination 5::/64,
     via FE80::200:2FF:FE64:8904, cost 2, tag 0, AOF, 31 secs
```

(3) 配置 Router B 对接收和发布的路由进行过滤

```
[RouterB] ipv6 prefix-list aaa permit 4:: 64
[RouterB] ipv6 prefix-list bbb deny 2:: 64
[RouterB] ipv6 prefix-list bbb permit :: 0 less-equal 128
[RouterB] ripng 1
[RouterB-ripng-1] filter-policy prefix-list aaa export
[RouterB-ripng-1] filter-policy prefix-list bbb import
[RouterB-ripng-1] quit
```

# 查看 Router B 和 Router A 的 RIPng 路由表。

```
[RouterB] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect, D – Direct
                O - Optimal, F - Flush to RIB
```

```
 ----------------------------------------------------------------

 Peer FE80::1:1 on GigabitEthernet2/1/1
 Destination 1::/64,
     via FE80::2:1, cost 1, tag 0, AOF, 6 secs

 Peer FE80::3:1 on GigabitEthernet2/1/2
 Destination 3::/64,
     via FE80::2:2, cost 1, tag 0, AOF, 11 secs
 Destination 4::/64,
     via FE80::2:2, cost 1, tag 0, AOF, 11 secs
 Destination 5::/64,
     via FE80::2:2, cost 1, tag 0, AOF, 11 secs
[RouterA] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect, D – Direct
                O - Optimal, F - Flush to RIB
 ----------------------------------------------------------------

 Peer FE80::2:1 on GigabitEthernet2/1/1
 Destination 4::/64,
     via FE80::1:1, cost 2, tag 0, AOF, 2 secs
```

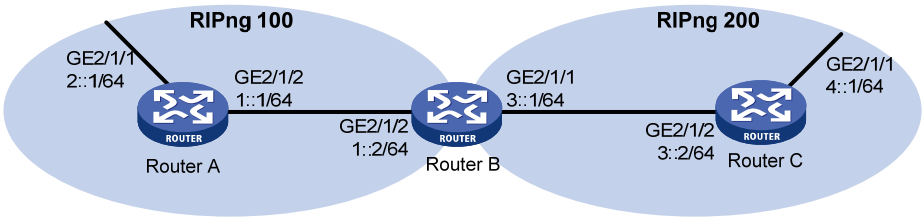
## 1.9.2 配置RIPng引入外部路由

### 1. 组网需求

- Router B 上运行两个 RIPng 进程：RIPng100 和 RIPng200。Router B 通过 RIPng100 和 Router A 交换路由信息，通过 RIPng200 和 Router C 交换路由信息。
- 要求在 Router B 上配置路由引入，将两个不同进程的 RIPng 路由相互引入到对方的 RIPng 进程中。

### 2. 组网图

图1-2 RIPng 引入外部路由配置组网图




### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 RIPng

# 在 Router A 上启动 RIPng 进程 100。

```
<RouterA> system-view
[RouterA] ripng 100
```

```
[RouterA-ripng-100] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ripng 100 enable
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] ripng 100 enable
```

# 在 Router B 上启动两个 RIPng 进程，进程号分别为 100 和 200。

```
<RouterB> system-view
[RouterB] ripng 100
[RouterB-ripng-100] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] ripng 100 enable
[RouterB-GigabitEthernet2/1/2] quit
[RouterB] ripng 200
[RouterB-ripng-200] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ripng 200 enable
```

# 在 Router C 上启动 RIPng 进程 200。

```
<RouterC> system-view
[RouterC] ripng 200
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ripng 200 enable
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] ripng 200 enable
[RouterC-GigabitEthernet2/1/2] quit
```

# 查看 Router A 的路由表信息。

```
[RouterA] display ipv6 routing-table

Destinations : 7 Routes : 7

Destination: ::1/128                                Protocol  : Direct
NextHop    : ::1                                    Preference: 0
Interface  : InLoop0                                Cost      : 0

Destination: 1::/64                                 Protocol  : Direct
NextHop    : 1::1                                   Preference: 0
Interface  : GE2/1/2                                Cost      : 0

Destination: 1::1/128                               Protocol  : Direct
NextHop    : ::1                                    Preference: 0
Interface  : InLoop0                                Cost      : 0

Destination: 2::/64                                 Protocol  : Direct
NextHop    : 2::1                                   Preference: 0
Interface  : GE2/1/1                                Cost      : 0

Destination: 2::1/128                               Protocol  : Direct
```

```
NextHop    : ::1                                    Preference: 0
Interface  : InLoop0                                Cost      : 0

Destination: FE80::/10                              Protocol  : Direct
NextHop    : ::                                     Preference: 0
Interface  : NULL0                                  Cost      : 0

Destination: FF00::/8                               Protocol  : Direct
NextHop    : ::                                     Preference: 0
Interface  : NULL0                                  Cost      : 0
```

(3) 配置 RIPng 引入外部路由

# 在 Router B 上将两个不同 RIPng 进程的路由相互引入到对方的路由表中。

```
[RouterB] ripng 100
[RouterB-ripng-100] import-route ripng 200
[RouterB-ripng-100] quit
[RouterB] ripng 200
[RouterB-ripng-200] import-route ripng 100
[RouterB-ripng-200] quit
```

# 查看路由引入后 Router A 的路由表信息。

```
[RouterA] display ipv6 routing-table

Destinations : 8 Routes : 8

Destination: ::1/128                                Protocol  : Direct
NextHop    : ::1                                    Preference: 0
Interface  : InLoop0                                Cost      : 0

Destination: 1::/64                                 Protocol  : Direct
NextHop    : 1::1                                   Preference: 0
Interface  : GE2/1/2                                Cost      : 0

Destination: 1::1/128                               Protocol  : Direct
NextHop    : ::1                                    Preference: 0
Interface  : InLoop0                                Cost      : 0

Destination: 2::/64                                 Protocol  : Direct
NextHop    : 2::1                                   Preference: 0
Interface  : GE2/1/1                                Cost      : 0

Destination: 2::1/128                               Protocol  : Direct
NextHop    : ::1                                    Preference: 0
Interface  : InLoop0                                Cost      : 0

Destination: 4::/64                                 Protocol  : RIPng
NextHop    : FE80::200:BFF:FE01:1C02                Preference: 100
Interface  : GE2/1/2                                Cost      : 1

Destination: FE80::/10                              Protocol  : Direct
```

```
NextHop    : ::                                    Preference: 0
Interface  : NULL0                                 Cost      : 0

Destination: FF00::/8                              Protocol  : Direct
NextHop    : ::                                    Preference: 0
Interface  : NULL0                                 Cost      : 0
```

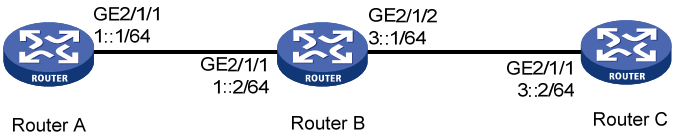
### 1.9.3 配置RIPng IPsec安全框架

#### 1. 组网需求

- Router A、Router B 和 Router C 相连并通过 RIPng 来学习网络中的 IPv6 路由信息。
- 要求配置 IPsec 安全框架对 Router A、Router B 和 Router C 之间的 RIPng 报文进行有效性检查和验证。

#### 2. 组网图

图1-3 RIPng IPsec 安全框架配置组网图




#### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 RIPng 基本功能

# 配置 Router A。

```
<RouterA> system-view
[RouterA] ripng 1
[RouterA-ripng-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ripng 1 enable
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] ripng 1
[RouterB-ripng-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ripng 1 enable
[RouterB-GigabitEthernet2/1/1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] ripng 1 enable
[RouterB-GigabitEthernet2/1/2] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] ripng 1
[RouterC-ripng-1] quit
```

```
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ripng 1 enable
[RouterC-GigabitEthernet2/1/1] quit
```

(3) 配置 RIPng IPsec 安全框架

# 配置 Router A。创建名为 protrf1 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议。创建一条安全框架 profile001，协商方式为 manual，配置 SPI 和密钥。

```
[RouterA] ipsec transform-set protrf1
[RouterA-ipsec-transform-set-protrf1] esp encryption-algorithm 3des-cbc
[RouterA-ipsec-transform-set-protrf1] esp authentication-algorithm md5
[RouterA-ipsec-transform-set-protrf1] encapsulation-mode transport
[RouterA-ipsec-transform-set-protrf1] quit
[RouterA] ipsec profile profile001 manual
[RouterA-ipsec-profile-profile001-manual] transform-set protrf1
[RouterA-ipsec-profile-profile001-manual] sa spi inbound esp 256
[RouterA-ipsec-profile-profile001-manual] sa spi outbound esp 256
[RouterA-ipsec-profile-profile001-manual] sa string-key inbound esp simple abc
[RouterA-ipsec-profile-profile001-manual] sa string-key outbound esp simple abc
[RouterA-ipsec-profile-profile001-manual] quit
```

# 配置 Router B。创建名为 protrf1 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议。创建一条安全框架 profile001，协商方式为 manual，配置 SPI 和密钥。

```
[RouterB] ipsec transform-set protrf1
[RouterB-ipsec-transform-set-protrf1] esp encryption-algorithm 3des-cbc
[RouterB-ipsec-transform-set-protrf1] esp authentication-algorithm md5
[RouterB-ipsec-transform-set-protrf1] encapsulation-mode transport
[RouterB-ipsec-transform-set-protrf1] quit
[RouterB] ipsec profile profile001 manual
[RouterB-ipsec-profile-profile001-manual] transform-set protrf1
[RouterB-ipsec-profile-profile001-manual] sa spi inbound esp 256
[RouterB-ipsec-profile-profile001-manual] sa spi outbound esp 256
[RouterB-ipsec-profile-profile001-manual] sa string-key inbound esp simple abc
[RouterB-ipsec-profile-profile001-manual] sa string-key outbound esp simple abc
[RouterB-ipsec-profile-profile001-manual] quit
```

# 配置 Router C。创建名为 protrf1 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议。创建一条安全框架 profile001，协商方式为 manual，配置 SPI 和密钥。

```
[RouterC] ipsec transform-set protrf1
[RouterC-ipsec-transform-set-protrf1] esp encryption-algorithm 3des-cbc
[RouterC-ipsec-transform-set-protrf1] esp authentication-algorithm md5
[RouterC-ipsec-transform-set-protrf1] encapsulation-mode transport
[RouterC-ipsec-transform-set-protrf1] quit
[RouterC] ipsec profile profile001 manual
[RouterC-ipsec-profile-profile001-manual] transform-set protrf1
[RouterC-ipsec-profile-profile001-manual] sa spi inbound esp 256
[RouterC-ipsec-profile-profile001-manual] sa spi outbound esp 256
[RouterC-ipsec-profile-profile001-manual] sa string-key inbound esp simple abc
[RouterC-ipsec-profile-profile001-manual] sa string-key outbound esp simple abc
[RouterC-ipsec-profile-profile001-manual] quit
```

(4) RIPng 进程上应用 IPsec 安全框架

# 配置 Router A。

```
[RouterA] ripng 1
[RouterA-ripng-1] enable ipsec-profile profile001
[RouterA-ripng-1] quit
```

# 配置 Router B。

```
[RouterB] ripng 1
[RouterB-ripng-1] enable ipsec-profile profile001
[RouterB-ripng-1] quit
```

# 配置 Router C。

```
[RouterC] ripng 1
[RouterC-ripng-1] enable ipsec-profile profile001
[RouterC-ripng-1] quit
```

### 4. 验证配置

以上配置完成后，Router A、Router B 和 Router C 之间的 RIPng 报文将被加密传输。

# 目 录

# 1 OSPFv3

## 1.1 OSPFv3简介

### 1.1.1 OSPFv3 概述

OSPFv3 是 OSPF（Open Shortest Path First，开放最短路径优先）版本 3 的简称，主要提供对 IPv6 的支持，遵循的标准为 RFC 5340。关于 OSPFv2 的介绍，请参见“三层技术-IP 路由”中的“OSPF”。

OSPFv3 和 OSPFv2 在很多方面是相同的：

- Router ID，Area ID 仍然是 32 位的。
- 相同类型的报文：Hello 报文，DD（Database Description，数据库描述）报文，LSR（Link State Request，链路状态请求）报文，LSU（Link State Update，链路状态更新）报文和 LSAck（Link State Acknowledgment，链路状态确认）报文。
- 相同的邻居发现机制和邻接形成机制。
- 相同的 LSA 扩散机制和老化机制。

OSPFv3 和 OSPFv2 的不同主要有：

- OSPFv3 是基于链路运行；OSPFv2 是基于网段运行。在配置 OSPFv3 时，不需要考虑是否配置在同一网段，只要在同一链路，就可以直接建立联系。
- OSPFv3 在同一条链路上可以运行多个实例，即一个接口可以使能多个 OSPFv3 进程（使用不同的实例）。
- OSPFv3 是通过 Router ID 来标识邻居；OSPFv2 则是通过 IPv4 地址来标识邻居。

### 1.1.2 OSPFv3 的协议报文

和 OSPFv2 一样，OSPFv3 也有五种报文类型，如下：

- Hello 报文：周期性发送，用来发现和维持 OSPFv3 邻居关系，以及进行 DR（Designated Router，指定路由器）/BDR（Backup Designated Router，备份指定路由器）的选举。
- DD（Database Description，数据库描述）报文：描述了本地 LSDB（Link State DataBase，链路状态数据库）中每一条 LSA（Link State Advertisement，链路状态通告）的摘要信息，用于两台路由器进行数据库同步。
- LSR（Link State Request，链路状态请求）报文：向对方请求所需的 LSA。两台路由器互相交换 DD 报文之后，得知对端的路由器有哪些 LSA 是本地的 LSDB 所缺少的，这时需要发送 LSR 报文向对方请求所需的 LSA。
- LSU（Link State Update，链路状态更新）报文：向对方发送其所需要的 LSA。
- LSAck（Link State Acknowledgment，链路状态确认）报文：用来对收到的 LSA 进行确认。

### 1.1.3 OSPFv3 的LSA类型

LSA（Link State Advertisement，链路状态通告）是 OSPFv3 协议计算和维护路由信息的主要来源，常用的 LSA 有以下几种类型：

- Router LSA（Type-1）：由每个路由器生成，描述本路由器的链路状态和开销，只在路由器所处区域内传播。
- Network LSA（Type-2）：由广播网络和 NBMA（Non-Broadcast Multi-Access，非广播多路访问）网络的 DR（Designated Router，指定路由器）生成，描述本网段接口的链路状态，只在 DR 所处区域内传播。
- Inter-Area-Prefix LSA（Type-3）：由 ABR（Area Border Router，区域边界路由器）生成，在与该 LSA 相关的区域内传播，描述一条到达本自治系统内其他区域的 IPv6 地址前缀的路由。
- Inter-Area-Router LSA（Type-4）：由 ABR 生成，在与该 LSA 相关的区域内传播，描述一条到达本自治系统内的 ASBR（Autonomous System Boundary Router，自治系统边界路由器）的路由。
- AS External LSA（Type-5）：由 ASBR 生成，描述到达其它 AS（Autonomous System，自治系统）的路由，传播到整个 AS（Stub 区域和 NSSA 区域除外）。缺省路由也可以用 AS External LSA 来描述。
- NSSA LSA（Type-7）：由 NSSA 区域内的 ASBR 生成，描述到 AS 外部的路由，仅在 NSSA 区域内传播。
- Link LSA（Type-8）：路由器为每一条链路生成一个 Link-LSA，在本地链路范围内传播，描述该链路上所连接的 IPv6 地址前缀及路由器的 Link-local 地址。
- Intra-Area-Prefix LSA（Type-9）：包含路由器上的 IPv6 前缀信息，Stub 区域信息或穿越区域（Transit Area）的网段信息，该 LSA 在区域内传播。由于 Router LSA 和 Network LSA 不再包含地址信息，导致了 Intra-Area-Prefix LSA 的引入。
- Grace LSA（Type-11）：由 Restarter 在重启时候生成的，在本地链路范围内传播。这个 LSA 描述了重启设备的重启原因和重启时间间隔，目的是通知邻居本设备将进入 GR（Graceful Restart，平滑重启）。

### 1.1.4 协议规范

与 OSPFv3 相关的协议规范有：

- RFC 5340：OSPF for IPv6
- RFC 2328：OSPF Version 2
- RFC 3101：OSPF Not-So-Stubby Area (NSSA) Option
- RFC 5187：OSPFv3 Graceful Restart

## 1.2 OSPFv3配置任务简介

表1-1 OSPFv3 配置任务简介

| 配置任务 | | 说明 | 详细配置 |
|---|---|---|---|
| 使能OSPFv3功能 | | 必选 | 1.3 |
| 配置OSPFv3的区域属性 | 配置OSPFv3的Stub区域 | 可选 | 1.4.2 |
| | 配置OSPFv3的NSSA区域 | 可选 | 1.4.3 |

| 配置任务 | | 说明 | 详细配置 |
| --- | --- | --- | --- |
| | 配置OSPFv3的虚连接 | 可选 | 1.4.4 |
| 配置OSPFv3网络类型 | 配置OSPFv3接口的网络类型 | 可选 | 1.5.2 |
| | 配置NBMA或者P2MP网络的邻居 | 可选 | 1.5.3 |
| 配置OSPFv3的路由信息控制 | 配置配置OSPFv3的路由聚合 | 可选 | 1.6.2 |
| | 配置过滤通过接收到的LSA计算出来的路由信息 | 可选 | 1.6.3 |
| | 配置过滤Inter-Area-Prefix-LSA | 可选 | 1.6.4 |
| | 配置OSPFv3接口的开销值 | 可选 | 1.6.5 |
| | 配置OSPFv3最大等价路由条数 | 可选 | 1.6.6 |
| | 配置OSPFv3协议的优先级 | 可选 | 1.6.7 |
| | 配置OSPFv3引入外部路由 | 可选 | 1.6.8 |
| 调整和优化OSPFv3网络 | 配置OSPFv3定时器 | 可选 | 1.7.2 |
| | 配置接口的LSA传输延迟时间 | 可选 | 1.7.3 |
| | 配置SPF计算时间间隔 | 可选 | 1.7.4 |
| | 配置LSA重新生成的时间间隔 | 可选 | 1.7.5 |
| | 配置接口的DR优先级 | 可选 | 1.7.6 |
| | 忽略DD报文中的MTU检查 | 可选 | 1.7.7 |
| | 禁止接口收发OSPFv3报文 | 可选 | 1.7.8 |
| | 配置邻居状态变化的输出开关 | 可选 | 1.7.9 |
| | 配置接口发送LSU报文的时间间隔和一次发送LSU报文的最大个数 | 可选 | 1.7.10 |
| 配置OSPFv3 GR | 配置GR Restarter | 可选 | 1.8.1 |
| | 使能GR Helper | 可选 | 1.8.2 |
| 配置OSPFv3 NSR | | 可选 | 1.9 |
| 配置OSPFv3与BFD联动 | | 可选 | 1.10 |
| 配置OSPFv3 IPsec安全框架 | | 可选 | 1.11 |

# 1.3 使能OSPFv3功能

## 1.3.1 配置准备

配置接口的网络层地址，使各相邻节点的网络层可达。

### 1.3.2 使能OSPFv3 功能

要在路由器上使能 OSPFv3 功能，必须先创建 OSPFv3 进程、指定该进程的 Router ID 以及在接口上使能 OSPFv3 功能。

Router ID 用来在一个自治系统中唯一的标识一台路由器。在 OSPFv3 中，用户必须手工配置一个 Router ID，而且必须保证自治系统中任意两台路由器的 Router ID 都不相同。因此，为了保证 OSPFv3 运行的稳定性，在进行网络规划时，应确定路由器 ID 的划分并手工配置。需要注意的是，如果在同一台路由器上运行了多个 OSPFv3 进程，必须为不同的进程指定不同的 Router ID。

在一台路由器上可以创建多个 OSPFv3 进程，OSPFv3 进程是本地概念。不同的路由器之间，即使进程不同也可以进行报文交换。

表1-2 使能 OSPFv3 功能

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 创建OSPFv3进程，进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | 缺省情况下，系统不运行OSPFv3进程 |
| 配置路由器的Router ID | **router-id** *router-id* | 缺省情况下，运行OSPFv3协议的路由器没有Router ID |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 在接口上使能OSPFv3 | **ospfv3** *process-id* **area** *area-id* [ **instance** *instance-id* ] | 缺省情况下，接口上没有使能OSPFv3 |

## 1.4 配置OSPFv3的区域属性

OSPFv3 支持 Stub 区域、NSSA 区域和虚连接的配置，其原理及应用环境与 OSPFv2 相同。

OSPFv3 划分区域后，可以减少网络中 LSA 的数量，OSPFv3 的扩展性也得以增强。对于位于 AS 边缘的一些非骨干区域，为了更多的缩减其路由表规模和降低 LSA 的数量，可以将它们配置为 Stub 区域。

Stub 区域不能引入外部路由，为了在允许将自治系统外部路由通告到 OSPFv3 路由域内部的同时，保持其余部分的 Stub 区域的特征，网络管理员可以将区域配置为 NSSA 区域。NSSA 区域也是位于 AS 边缘的非骨干区域。

在划分区域之后，非骨干区域之间的 OSPFv3 路由更新是通过骨干区域来交换完成的。对此，OSPFv3 要求所有非骨干区域必须与骨干区域保持连通，并且骨干区域自身也要保持连通。但在实际应用中，可能会因为各方面条件的限制，无法满足这个要求。这时可以通过配置 OSPFv3 虚连接予以解决。

### 1.4.1 配置准备

在配置 OSPFv3 的区域属性之前，需完成以下任务：使能 OSPFv3 功能。

### 1.4.2 配置OSPFv3 的Stub区域

对于位于 Stub 区域中的所有路由器都必须配置 **stub** 命令，参数 **no-summary** 只能在 ABR 上配置。如果 ABR 使用了 **stub** 命令中的参数 **no-summary**，则此 ABR 只向区域内发布一条描述缺省路由的 Inter-Area-Prefix-LSA。这种既没有 AS-external-LSA，也没有其它 Inter-Area-Prefix-LSA、Inter-Area-Router-LSA 的 Stub 区域，又称为 Totally Stub 区域。

表1-3 配置 OSPFv3 的 Stub 区域

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPFv3区域视图 | **area** *area-id* | - |
| 配置一个区域为Stub区域 | **stub** [ **default-route-advertise-always** \| **no-summary** ] * | 缺省情况下，没有区域被配置为Stub区域 |
| （可选）配置发送到Stub区域的缺省路由的开销值 | **default-cost** *value* | 缺省情况下，发送到Stub区域的缺省路由的开销值为1 |

### 1.4.3 配置OSPFv3 的NSSA区域

Stub 区域不能引入外部路由，为了在允许将自治系统外部路由通告到 OSPFv3 路由域内部的同时，保持其余部分的 Stub 区域的特征，网络管理员可以将区域配置为 NSSA 区域。NSSA 区域也是位于 AS 边缘的非骨干区域。

对于位于 NSSA 区域中的所有路由器都必须配置 **nssa** 命令。配置 **nssa** 命令时指定 **no-summary** 参数可以将该区域配置为 Totally NSSA 区域，该区域的 ABR 不会将区域间的路由信息传递到本区域。

表1-4 配置 OSPFv3 的 NSSA 区域

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPFv3区域视图 | **area** *area-id* | - |
| 配置一个区域为NSSA区域 | **nssa** [ **default-route-advertise** [ **cost** *cost* \| **nssa-only** \| **route-policy** *route-policy-name* \| **tag** *tag* \| **type** *type* ] * \| **no-import-route** \| **no-summary** \| [ **translate-always** \| **translate-never** ] \| **suppress-fa** \| **translator-stability-interval** *value* ] * | 缺省情况下，没有区域被配置为NSSA区域 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| （可选）配置发送到NSSA区域的缺省路由的开销值 | **default-cost** *cost* | 缺省情况下，发送到NSSA区域的缺省路由的开销值为1<br>本命令只有在NSSA区域和Totally NSSA区域的ABR/ASBR上配置才能生效 |

### 1.4.4 配置OSPFv3 的虚连接

提示

- 虚连接的两端必须是 ABR，而且必须在两端同时配置才可生效。
- 虚连接不能在使能了 GR 能力的进程下的区域进行配置。

对于没有和骨干区域直接相连的非骨干区域，或者不连续的骨干区域，可以使用该配置建立逻辑上的连通性。

表1-5 配置 OSPFv3 的虚连接

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPFv3区域视图 | **area** *area-id* | - |
| 创建并配置虚连接 | **vlink-peer** *router-id* [ **dead** *seconds* \| **hello** *seconds* \| **instance** *instance-id* \| **ipsec-profile** *profile-name* \| **retransmit** *seconds* \| **trans-delay** *seconds* ] * | 缺省情况下，没有虚连接 |

## 1.5 配置OSPFv3网络类型

OSPFv3 根据链路层协议类型将网络分为四种不同的类型：广播、NBMA、P2MP 和 P2P。

缺省情况下，当接口封装的链路层协议不同时，OSPFv3 接口网络类型的缺省值也不同：

- 当接口封装的链路层协议是 Ethernet、FDDI 时，OSPFv3 接口网络类型的缺省值为广播类型；
- 当接口封装的链路层协议是 ATM、帧中继或 X.25 时，OSPFv3 接口网络类型的缺省值为 NBMA；
- 当接口封装的链路层协议是 PPP、LAPB、HDLC 或 POS 时，OSPFv3 接口网络类型的缺省值为 P2P。

用户可以根据需要配置 OSPFv3 接口的网络类型：

- 由于 NBMA 网络必须是全连通的，即网络中任意两台路由器之间都必须有一条虚电路直接可达。但在很多情况下，这个要求无法满足，这时就需要通过命令强制改变网络的类型。

- 对于 NBMA 网络，如果部分路由器之间没有直接可达的链路时，应将接口的网络类型配置为 P2MP。如果路由器在 NBMA 网络中只有一个对端，也可将接口类型配置为 P2P。

### 1.5.1 配置准备

在配置 OSPFv3 的网络类型之前，需完成以下任务：使能 OSPFv3 功能。

### 1.5.2 配置OSPFv3 接口的网络类型

表1-6 配置 OSPFv3 接口的网络类型

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置OSPFv3接口的网络类型 | **ospfv3 network-type** { **broadcast** \| **nbma** \| **p2mp** [ **unicast** ] \| **p2p** } [ **instance** *instance-id* ] | 缺省情况下，接口的网络类型根据物理接口而定 |

### 1.5.3 配置NBMA或者P2MP网络的邻居

当路由器的接口类型为如下网络类型时，需要为其指定相邻路由器 IP 地址：

- NBMA 网络
- P2MP 网络（仅当接口选择单播形式发送报文时，需要此配置）

由于无法通过广播 Hello 报文的形式发现相邻路由器，必须手工指定相邻路由器的本地链路地址。对于 NBMA 网络，可以指定该相邻路由器是否有 DR 选举权等。

表1-7 配置 NBMA 或者 P2MP 网络的邻居

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置NBMA或者P2MP（单播）网络的邻居 | **ospfv3 peer** *ipv6-address* [ **cost** *value* \| **dr-priority** *dr-priority* ] [ **instance** *instance-id* ] | 缺省情况下，没有指定邻居接口的链路本地地址 |

## 1.6 配置OSPFv3的路由信息控制

### 1.6.1 配置准备

在配置 OSPFv3 的路由信息控制之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点的网络层可达
- 使能 OSPFv3 功能

### 1.6.2 配置OSPFv3 的路由聚合

如果该区域中存在多个连续的网段，则可以在 ABR 上配置 **abr-summary** 命令将它们聚合成一个网段，ABR 只发送一条聚合后的 LSA，所有落入本命令指定的聚合网段范围的 LSA 将不再会被单独发送出去，这样可减小其它区域中 LSDB 的规模。

表1-8 配置 OSPFv3 的路由聚合

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPFv3区域视图 | **area** *area-id* | - |
| 配置OSPFv3区域路由聚合 | **abr-summary** *ipv6-address prefix-length* [ **not-advertise** ] [ **cost** value ] | 缺省情况下，ABR没有对路由进行聚合<br>路由聚合只有在ABR上配置才会有效 |

### 1.6.3 过滤通过接收到的LSA计算出来的路由信息

OSPFv3 接收到 LSA 后，可以根据一定的过滤条件来决定是否将计算后得到的路由信息加入到本地路由表中。

表1-9 过滤通过接收到的 LSA 计算出来的路由信息

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 过滤通过接收到的LSA计算出来的路由信息 | **filter-policy** { *acl6-number* [ **gateway** *prefix-list-name* ] \| **prefix-list** *prefix-list-name* [ **gateway** *prefix-list-name* ] \| **gateway** *prefix-list-name* \| **route-policy** *route-policy-name* } **import** | 缺省情况下，不对通过接收到的LSA计算出来的路由信息进行过滤<br>本命令只对OSPFv3计算出来的路由进行过滤，没有通过过滤的路由将不被加入到本地路由表中，从而不能用于转发报文 |

### 1.6.4 配置过滤Inter-Area-Prefix-LSA

此命令只在 ABR 路由器上有效，对区域内部路由器无效。

表1-10 配置过滤 Inter-Area-Prefix-LSA

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入OSPFv3区域视图 | **area** *area-id* | - |
| 配置对Inter-Area-Prefix-LSA进行过滤 | **filter** { *acl6-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } { **export** \| **import** } | 缺省情况下，没有配置对Inter-Area-Prefix-LSA进行过滤 |

### 1.6.5 配置OSPFv3 接口的开销值

OSPFv3 有两种方式来配置接口的开销值：

- 第一种方法是在接口视图下直接配置开销值；
- 第二种方法是配置接口的带宽参考值，OSPFv3 根据带宽参考值自动计算接口的开销值，计算公式为：接口开销＝带宽参考值（100Mbps）÷接口带宽（Mbps），当计算出来的开销值大于 65535，开销取最大值 65535；当计算出来的开销值小于 1 时，开销取最小值 1。

如果没有在接口视图下配置接口的开销值，OSPFv3 会根据该接口的带宽自动计算开销值。

#### 1. 配置接口的开销值

表1-11 配置 OSPFv3 接口的开销值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置OSPFv3接口的开销值 | **ospfv3 cost** *value* [ **instance** *instance-id* ] | 缺省情况下，OSPFv3根据接口的带宽自动计算链路开销，对于VLAN接口，缺省值为1；对于Loopback接口，缺省取值为0 |

#### 2. 配置带宽参考值

表1-12 配置带宽参考值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置带宽参考值 | **bandwidth-reference** *value* | 缺省情况下，带宽参考值为100Mbps |

### 1.6.6 配置OSPFv3 最大等价路由条数

如果到一个目的地有几条开销相同的路径，可以通过等价路由负载分担来提高链路利用率。

表1-13 配置 OSPFv3 最大等价路由条数

| 操作 | 命令 | 说明 |
| --- | --- | --- |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPFv3最大等价路由条数 | **maximum load-balancing** *maximum* | 缺省情况下，OSPFv3支持的最大等价路由条数为32。 |

## 1.6.7 配置OSPFv3 协议的优先级

由于路由器上可能同时运行多个动态路由协议，就存在各个路由协议之间路由信息共享和选择的问题。系统为每一种路由协议设置一个优先级，在不同协议发现同一条路由时，优先级高的路由将被优选。

表1-14 配置 OSPFv3 协议的优先级

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置OSPFv3协议的优先级 | **preference** [ **ase** ] [ **route-policy** *route-policy-name* ] *preference* | 缺省情况下，OSPFv3内部路由的优先级为10，OSPFv3 外部路由的优先级为150 |

## 1.6.8 配置OSPFv3 引入外部路由

由于 OSPFv3 是基于链路状态的路由协议，不能直接对发布的 LSA 进行过滤，所以只能在 OSPFv3 引入路由时进行过滤，只有符合条件的路由才能转换成 LSA 发布出去。

表1-15 配置 OSPFv3 引入外部路由

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| （可选）配置外部路由的缺省开销值 | **default cost** *value* | 缺省情况下，缺省开销值为1 |
| 引入外部路由信息 | **import-route** *protocol* [ *process-id* \| **all-processes** \| **allow-ibgp** ] [ **allow-direct** \| **cost** *cost* \| **nssa-only** \| **route-policy** *route-policy-name* \| **tag** *tag* \| **type** *type* ] * | 缺省情况下，没有引入外部路由信息 |
| （可选）配置OSPFv3引入缺省路由 | **default-route-advertise** [ [ **always** \| **permit-calculate-other** ] \| **cost** *cost* \| **route-policy** *route-policy-name* \| **type** *type* ] * | 缺省情况下，没有引入缺省路由<br>只能通过本命令引入并发布缺省路由 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| （可选）对引入的外部路由信息进行过滤 | **filter-policy** { *acl6-number* \| **prefix-list** *prefix-list-name* } **export** [ *protocol* [ *process-id* ] ] | 缺省情况下，没有对引入的路由信息过滤<br>本命令只对本设备使用**import-route**引入的路由起作用。如果没有配置**import-route**命令来引入其它外部路由（包括不同进程的OSPFv3路由），则本命令失效 |

说明

在 OSPFv3 路由器上配置 **import-route** 或 **default-route-advertise** 命令后，这台 OSPFv3 路由器就成为 ASBR。

# 1.7 调整和优化OSPFv3网络

本节主要介绍配置 OSPFv3 定时器、配置接口的 DR 优先级和邻居状态变化的输出开关。

## 1.7.1 配置准备

在调整和优化 OSPFv3 网络之前，需完成以下任务：

- 配置接口的网络层地址，使相邻节点的网络层可达
- 使能 OSPFv3 功能

## 1.7.2 配置OSPFv3 定时器

表1-16 配置 OSPFv3 定时器

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口发送hello报文的时间间隔 | **ospfv3 timer hello** *seconds* [ **instance** *instance-id* ] | 缺省情况下，P2P、Broadcast网络类型接口发送Hello报文的时间间隔的值为10秒 |
| 配置相邻路由器间失效时间 | **ospfv3 timer dead** *seconds* [ **instance** *instance-id* ] | 缺省情况下，P2P、Broadcast网络类型接口的OSPFv3邻居失效时间为40秒<br>相邻路由器间失效时间的值不要设置得太小，否则邻居很容易失效 |
| 配置Poll定时器 | **ospfv3 timer poll** *seconds* [ **instance** *instance-id* ] | 缺省情况下，发送轮询Hello报文的时间间隔为120秒 |
| 配置相邻路由器重传LSA的时间间隔 | **ospfv3 timer retransmit** *interval* [ **instance** *instance-id* ] | 缺省情况下，LSA的重传时间间隔为5秒<br>相邻路由器重传LSA时间间隔的值不要设置得太小，否则将会引起不必要的重传 |

### 1.7.3 配置接口的LSA传输延迟时间

LSA 在本路由器的链路状态数据库（LSDB）中会随时间老化（每秒钟加 1），但在网络的传输过程中却不会，所以有必要在发送之前将 LSA 的老化时间增加一定的延迟时间。此配置对低速率的网络尤其重要。

表1-17 配置接口的 LSA 传输延迟时间

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的LSA传输延迟时间 | **ospfv3 trans-delay** *seconds* [ **instance** *instance-id* ] | 缺省情况下，接口的LSA传输延迟时间为1秒 |

### 1.7.4 配置SPF计算时间间隔

当 OSPFv3 的 LSDB 发生改变时，需要重新计算最短路径。如果网络频繁变化，且每次变化都立即计算最短路径，将会占用大量系统资源，并影响路由器的效率。通过调节 SPF 计算时间间隔，可以抑制由于网络频繁变化带来的影响。

本命令在网络变化不频繁的情况下将连续路由计算的时间间隔缩小到 *minimum-interval*，而在网络变化频繁的情况下可以进行相应惩罚，增加 *incremental-interval*×$2^{n-2}$（n 为连续触发路由计算的次数），将等待时间按照配置的惩罚增量延长，最大不超过 *maximum-interval*。

表1-18 配置 SPF 计算时间间隔

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置SPF计算时间间隔 | **spf-schedule-interval** *maximum-interval* [ *minimum-interval* [ *incremental-interval* ] ] | 缺省情况下，SPF计算的最大时间间隔为5秒，最小时间间隔为50毫秒，时间间隔惩罚增量为200毫秒 |

### 1.7.5 配置LSA重新生成的时间间隔

通过调节 LSA 重新生成的时间间隔，可以抑制网络频繁变化可能导致的占用过多带宽资源和路由器资源。

本命令在网络变化不频繁的情况下将 LSA 重新生成时间间隔缩小到 *minimum-interval*，而在网络变化频繁的情况下可以进行相应惩罚，增加 *incremental-interval*×$2^{n-2}$（n 为连续触发路由计算的次数），将等待时间按照配置的惩罚增量延长，最大不超过 *maximum-interval*。

表1-19 配置 LSA 发送间隔

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置LSA重新生成的时间间隔 | **lsa-generation-interval** *maximum-interval* [ *minimum-interval* [ *incremental-interval* ] ] | 缺省情况下，最大时间间隔为5秒，最小时间间隔为0毫秒，惩罚增量为0毫秒 |

## 1.7.6 配置接口的DR优先级

路由器接口的 DR 优先级将影响接口在选举 DR 时所具有的资格，优先级为 0 的路由器不会被选举为 DR 或 BDR。

表1-20 配置接口的 DR 优先级

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的DR优先级 | **ospfv3 dr-priority** *priority* [ **instance** *instance-id* ] | 缺省情况下，接口的DR优先级为1 |

## 1.7.7 忽略DD报文中的MTU检查

在 LSA 数量不多的情况下，没有必要去检查 MTU 大小，可以设置忽略 DD 报文中的 MTU 检查，从而提高性能。

表1-21 忽略 DD 报文中的 MTU 检查

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 忽略DD报文中的MTU检查 | **ospfv3 mtu-ignore** [ **instance** *instance-id* ] | 缺省情况下，接口在进行DD报文交换时执行MTU检查 |

## 1.7.8 禁止接口收发OSPFv3 报文

当运行 OSPFv3 协议的接口被配置为 Silent 状态后，该接口的直连路由仍可以由同一路由器的其他接口通过 Intra-Area-Prefix-LSA 发布，但 OSPFv3 报文将被阻塞，接口上不会建立 OSPFv3 邻居关系。这一特性可以增强 OSPFv3 的组网适应能力。

表1-22 禁止接口收发 OSPFv3 报文

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 禁止接口收发OSPFv3报文 | **silent-interface** { *interface-type interface-number* \| **all** } | 缺省情况下，允许接口收发OSPFv3报文<br>不同的进程可以对同一接口禁止收发OSPFv3报文，但本命令只对本进程已经使能的OSPFv3接口起作用，不对其它进程的接口起作用 |

## 1.7.9 配置邻居状态变化的输出开关

打开邻居状态变化的输出开关后，OSPFv3 邻居状态变化时会生成日志信息发送到设备的信息中心，通过设置信息中心的参数，最终决定日志信息的输出规则（即是否允许输出以及输出方向）。（有关信息中心参数的配置请参见“网络管理和监控配置指导”中的“信息中心”。）

表1-23 配置邻居状态变化的输出开关

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置邻居状态变化的输出开关 | **log-peer-change** | 缺省情况下，邻居状态变化的输出开关处于打开状态 |

## 1.7.10 配置接口发送LSU报文的时间间隔和一次发送LSU报文的最大个数

如果路由器路由表里的路由条目很多，在与邻居进行 LSDB 同步时，可能需要发送大量 LSU，有可能会对当前设备和网络带宽带来影响；因此，路由器将 LSU 报文分为多个批次进行发送，并且对 OSPFv3 接口每次允许发送的 LSU 报文的最大个数做出限制。

用户可根据需要配置 OSPFv3 接口发送 LSU 报文的时间间隔以及接口一次发送 LSU 报文的最大个数。

表1-24 配置接口发送 LSU 报文的时间间隔和一次发送 LSU 报文的最大个数

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置接口发送LSU报文的时间间隔和一次发送LSU报文的最大个数 | **transmit-pacing interval** *interval* **count** *count* | 缺省情况下，OSPFv3接口发送LSU报文的时间间隔为20毫秒，一次最多发送3个LSU报文 |

## 1.8 配置OSPFv3 GR

GR（Graceful Restart，平滑重启）是一种在协议重启或主备倒换时保证转发业务不中断的机制。GR 有两个角色：

- GR Restarter：发生协议重启或主备倒换事件且具有 GR 能力的设备。
- GR Helper：和 GR Restarter 具有邻居关系，协助完成 GR 流程的设备。

支持 OSPFv3 的 GR Restarter 能力的设备主备倒换后，为了实现设备转发业务的不中断，它必须完成下列两项任务：

- 重启过程 GR Restarter 转发表项保持稳定；
- 重启流程结束后重建所有邻居关系，重新获取完整的网络拓扑信息。

设备（GR Restarter）主备倒换后，首先向邻居发送 Grace LSA 通告邻居本设备进入 GR；邻居收到 Grace-LSA 后，如果支持 GR Helper 能力则进入 Helper 模式（此时该邻居称为 GR Helper）。GR Restarter 重新建立邻居，GR Helper 帮助 GR Restarter 进行 LSDB 的同步。同步完成之后，GR 流程结束，进入正常的 OSPFv3 流程。这样就能实现设备在主备倒换时转发业务正常进行。

### 1.8.1 配置GR Restarter

可以在 GR Restarter 设备上配置 GR Restarter 能力。

**表1-25 配置 GR Restarter**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能GR能力 | **graceful-restart enable** | 缺省情况下，OSPFv3协议的GR Restarter能力处于关闭状态 |
| （可选）配置GR重启时间间隔 | **graceful-restart interval** *interval-value* | 缺省情况下，OSPFv3协议的GR重启间隔时间为120秒 |

### 1.8.2 配置GR Helper

可以在 GR Helper 设备上配置 GR Helper 能力。

**表1-26 配置 GR Helper**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能helper能力 | **graceful-restart helper enable** | 缺省情况下，OSPFv3的GR Helper能力处于打开状态 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 使能LSA严格检查能力 | **graceful-restart helper strict-lsa-checking** | 缺省情况下，OSPFv3协议的GR Helper严格LSA检查能力处于关闭状态 |

## 1.9 配置OSPFv3 NSR

MSR 系列路由器各款型对于本节所描述的特性的支持情况有所不同，详细差异信息如下：

| 型号 | 特性 | 描述 |
| --- | --- | --- |
| MSR 2600 | 配置OSPFv3 NSR | 不支持 |
| MSR 3600 | | 不支持 |
| MSR 5600 | | 支持 |

说明

设备配置了 OSPFv3 NSR 功能后不能再充当 GR Restarter。

NSR（Nonstop Routing，不间断路由）通过将 OSPFv3 链路状态信息从主进程备份到备进程，使设备在发生主备倒换时可以自行完成链路状态的恢复和路由的重新生成，邻接关系不会发生中断，从而避免了主备倒换对转发业务的影响。

GR 特性需要周边设备配合才能完成路由信息的恢复，在网络应用中有一定的限制。NSR 特性不需要周边设备的配合，网络应用更加广泛。

表1-27 配置 NSR

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 使能NSR能力 | **non-stop-routing** | 缺省情况下，OSPFv3协议的NSR能力处于关闭状态 |

## 1.10 配置OSPFv3与BFD联动

BFD（Bidirectional Forwarding Detection，双向转发检测）能够为 OSPFv3 邻居之间的链路提供快速检测功能。当邻居之间的链路出现故障时，加快 OSPFv3 协议的收敛速度。关于 BFD 的介绍和基本功能配置，请参见“可靠性配置指导”中的“BFD”。

OSPFv3 使用 BFD 来进行快速故障检测时，可以通过 Hello 报文动态发现邻居，将邻居地址通知 BFD 就开始建立会话。BFD 会话建立前处于 down 状态，此时 BFD 控制报文以不小于 1 秒的时间间隔周期发送以减少控制报文流量，直到会话建立以后才会以协商的时间间隔发送以实现快速检测。

进行配置 BFD 之前，需要配置 OSPFv3 功能。

表1-28 配置 OSPFv3 与 BFD 联动

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 配置路由器的ID | **router-id** *router-id* | - |
| 退出OSPFv3视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 在接口上使能OSPFv3 | **ospfv3** *process-id* **area** *area-id* [ **instance** *instance-id* ] | - |
| 在指定接口上使能OSPFv3 BFD | **ospfv3 bfd enable** [ **instance** *instance-id* ] | 缺省情况下，运行OSPFv3的接口未使能BFD功能 |

## 1.11 配置OSPFv3 IPsec安全框架

从安全性角度来考虑，为了避免路由信息外泄或者对设备进行恶意攻击，OSPFv3 提供基于 IPsec 的报文验证功能。IPsec 安全框架的具体情况请参见“安全配置指导”中的“IPsec”。

设备在发送的报文中会携带配置好的 IPsec 安全框架的 SPI（Security Parameter Index，安全参数索引）值，接收报文时通过 SPI 值进行 IPsec 安全框架匹配：只有能够匹配的报文才能接收；否则将不会接收报文，从而不能正常建立邻居和学习路由。

OSPFv3 支持在区域、接口和虚连接下配置 IPsec 安全框架。

- 当需要保护区域内的所有报文时，可以在区域下配置 IPsec 安全框架，此时区域内所有路由器都需要配置相同的 IPsec 安全框架。
- 当需要保护区域下某些接口的报文时，可以在接口下配置 IPsec 安全框架，此时直连邻居接口需要配置相同的 IPsec 安全框架。
- 当需要保护虚连接的报文时，可以配置虚连接应用 IPsec 安全框架，此时虚连接上的两个邻居需要配置相同的 IPsec 安全框架。

当接口和接口所在区域均配置了 IPsec 安全框架时，接口下的生效；当虚连接和区域 0 均配置了 IPsec 安全框架时，虚连接的生效。

表1-29 配置 OSPFv3 IPsec 安全框架（区域）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPFv3区域视图 | **area** *area-id* | - |
| 配置OSPFv3区域应用IPsec安全框架 | **enable ipsec-profile** *profile-name* | 缺省情况下，OSPFv3区域没有应用IPsec安全框架 |

表1-30 配置 OSPFv3 IPsec 安全框架（接口）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置使能了OSPFv3的接口上应用IPsec安全框架 | **ospfv3 ipsec-profile** *profile-name* | 缺省情况下，OSPFv3接口没有应用IPsec安全框架 |

表1-31 配置 OSPFv3 IPsec 安全框架（虚连接）

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入OSPFv3视图 | **ospfv3** [ *process-id* \| **vpn-instance** *vpn-instance-name* ] * | - |
| 进入OSPFv3区域视图 | **area** *area-id* | - |
| 配置OSPFv3虚连接应用IPsec安全框架 | **vlink-peer** *router-id* [ **dead** *seconds* \| **hello** *seconds* \| **instance** *instance-id* \| **retransmit** *seconds* \| **trans-delay** *seconds* \| **ipsec-profile** *profile-name* ] * | 缺省情况下，OSPFv3虚连接没有应用IPsec安全框架 |

## 1.12 OSPFv3显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 OSPFv3 的运行情况，通过查看显示信息验证配置的效果。

表1-32 OSPFv3 显示和维护

| 操作 | 命令 |
| --- | --- |
| 显示到OSPFv3的区域边界路由器和自治系统边界路由器的路由信息 | **display ospfv3** [ *process-id* ] **abr-asbr** |
| 显示OSPFv3的ABR聚合信息 | **display ospfv3** [ *process-id* ] [ **area** *area-id* ] **abr-summary** [ *ipv6-address prefix-length* ] [ **verbose** ] |
| 显示OSPFv3的进程信息 | **display ospfv3** [ *process-id* ] [ **verbose** ] |
| 显示OSPFv3进程的GR状态信息 | **display ospfv3** [ *process-id* ] **graceful-restart** |
| 显示OSPFv3的接口信息 | **display ospfv3** [ *process-id* ] **interface** [ *interface-type interface-number* \| **verbose** ] |
| 显示OSPFv3的链路状态数据库信息 | **display ospfv3** [ *process-id* ] **lsdb** [ { **external** \| **grace** \| **inter-prefix** \| **inter-router** \| **intra-prefix** \| **link** \| **network** \| **nssa** \| **router** \| **unknown** [ *type* ] } [ *link-state-id* ] [ **originate-router** *router-id* \| **self-originate** ] \| **statistics** \| **total** \| **verbose** ] |
| 显示OSPFv3的路由下一跳信息 | **display ospfv3** [ *process-id* ] **nexthop** |

| 操作 | 命令 |
| --- | --- |
| 显示OSPFv3邻居信息 | **display ospfv3** [ *process-id* ] [ **area** *area-id* ] **peer** [ [ *interface-type interface-number* ] [ **verbose** ] \| *peer-router-id* \| **statistics** ] |
| 显示OSPFv3请求列表的信息 | **display ospfv3** [ *process-id* ] [ **area** *area-id* ] **request-queue** [ *interface-type interface-number* ] [ *neighbor-id* ] |
| 显示OSPFv3重传列表的信息 | **display ospfv3** [ *process-id* ] [ **area** *area-id* ] **retrans-queue** [ *interface-type interface-number* ] [ *neighbor-id* ] |
| 显示OSPFv3路由表信息 | **display ospfv3** [ *process-id* ] **routing** [ *ipv6-address prefix-length* ] |
| 显示OSPFv3区域的拓扑信息 | **display ospfv3** [ *process-id* ] [ **area** *area-id* ] **spf-tree** [ **verbose** ] |
| 显示OSPFv3的报文统计信息 | **display ospfv3** [ *process-id* ] **statistics** [ **error** ] |
| 显示OSPFv3的虚连接信息 | **display ospfv3** [ *process-id* ] **vlink** |

# 1.13 OSPFv3配置举例

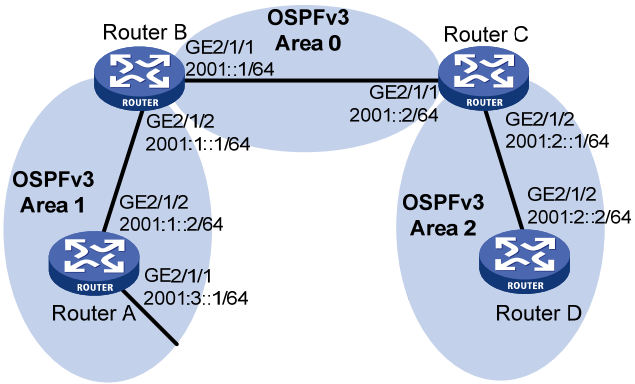
## 1.13.1 配置OSPFv3 的Stub区域

### 1. 组网需求

- 所有的路由器都运行 OSPFv3，整个自治系统划分为 3 个区域。其中 Router B 和 Router C 作为 ABR 来转发区域之间的路由。
- 要求将 Area 2 配置为 Stub 区域，减少通告到此区域内的 LSA 数量，但不影响路由的可达性。

### 2. 组网图

图1-1 OSPFv3 的 Stub 区域配置组网图




### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 OSPFv3 基本功能

# 配置 Router A，启动 OSPFv3，并配置其 Router ID 为 1.1.1.1。

```
<RouterA> system-view
[RouterA] ospfv3 1
```

```
[RouterA-ospfv3-1] router-id 1.1.1.1
[RouterA-ospfv3-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospfv3 1 area 1
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] ospfv3 1 area 1
[RouterA-GigabitEthernet2/1/2] quit
```

# 配置 Router B，启动 OSPFv3，并配置其 Router ID 为 2.2.2.2。

```
<RouterB> system-view
[RouterB] ospfv3 1
[RouterB-ospfv3-1] router-id 2.2.2.2
[RouterB-ospfv3-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterB-GigabitEthernet2/1/1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] ospfv3 1 area 1
[RouterB-GigabitEthernet2/1/2] quit
```

# 配置 Router C，启动 OSPFv3，并配置其 Router ID 为 3.3.3.3。

```
<RouterC> system-view
[RouterC] ospfv3 1
[RouterC-ospfv3-1] router-id 3.3.3.3
[RouterC-ospfv3-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] ospfv3 1 area 2
[RouterC-GigabitEthernet2/1/2] quit
```

# 配置 Router D，启动 OSPFv3，并配置其 Router ID 为 4.4.4.4。

```
<RouterD> system-view
[RouterD] ospfv3 1
[RouterD-ospfv3-1] router-id 4.4.4.4
[RouterD-ospfv3-1] quit
[RouterD] interface gigabitethernet 2/1/2
[RouterD-GigabitEthernet2/1/2] ospfv3 1 area 2
[RouterD-GigabitEthernet2/1/2] quit
```

# 查看 Router B 的 OSPFv3 邻居状态。

```
[RouterB] display ospfv3 peer

              OSPFv3 Process 1 with Router ID 2.2.2.2

 Area: 0.0.0.0
--------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 3.3.3.3         1   Full/BDR          00:00:40  0      GE2/1/1
```

```
 Area: 0.0.0.1
--------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 1.1.1.1         1   Full/DR           00:00:40  0      GE2/1/2
```

# 查看 Router C 的 OSPFv3 邻居状态。

```
[RouterC] display ospfv3 peer

              OSPFv3 Process 1 with Router ID 3.3.3.3

 Area: 0.0.0.0
--------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 2.2.2.2         1   Full/DR           00:00:40  0      GE2/1/1

 Area: 0.0.0.2
--------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 4.4.4.4         1   Full/BDR          00:00:40  0      GE2/1/2
```

# 查看 Router D 的 OSPFv3 路由表信息。

```
[RouterD] display ospfv3 routing

              OSPFv3 Process 1 with Router ID 4.4.4.4
--------------------------------------------------------------------------
 I  - Intra area route,  E1 - Type 1 external route,  N1 - Type 1 NSSA route
 IA - Inter area route,  E2 - Type 2 external route,  N2 - Type 2 NSSA route
 *  - Selected route

 *Destination: 2001::/64
  Type       : IA                                    Cost     : 2
  NextHop    : FE80::F40D:0:93D0:1                   Interface: GE2/1/2
  AdvRouter  : 3.3.3.3                               Area     : 0.0.0.0
  Preference : 10

 *Destination: 2001:1::/64
  Type       : IA                                    Cost     : 3
  NextHop    : FE80::F40D:0:93D0:1                   Interface: GE2/1/2
  AdvRouter  : 3.3.3.3                               Area     : 0.0.0.0
  Preference : 10

 *Destination: 2001:2::/64
  Type       : I                                     Cost     : 1
  Nexthop    : ::                                    Interface: GE2/1/2
  AdvRouter  : 4.4.4.4                               Area     : 0.0.0.2
  Preference : 10

 *Destination: 2001:3::/64
  Type       : IA                                    Cost     : 4
```

```
 NextHop    : FE80::F40D:0:93D0:1                      Interface: GE2/1/2
 AdvRouter  : 3.3.3.3                                  Area     : 0.0.0.0
 Preference : 10

Total: 4
Intra area: 1         Inter area: 3         ASE: 0         NSSA: 0
```

(3) 配置 Stub 区域

# 配置 Router D 的 Stub 区域。

```
[RouterD] ospfv3
[RouterD-ospfv3-1] area 2
[RouterD-ospfv3-1-area-0.0.0.2] stub
[RouterD-ospfv3-1-area-0.0.0.2] quit
[RouterD-ospfv3-1] quit
```

# 配置 Router C 的 Stub 区域，设置发送到 Stub 区域的缺省路由的开销为 10。

```
[RouterC] ospfv3
[RouterC-ospfv3-1] area 2
[RouterC-ospfv3-1-area-0.0.0.2] stub
[RouterC-ospfv3-1-area-0.0.0.2] default-cost 10
```

# 查看 Router D 的 OSPFv3 路由表信息，可以看到路由表中多了一条缺省路由，它的开销值为直连路由的开销和所配置的开销值之和。

```
[RouterD] display ospfv3 routing

              OSPFv3 Process 1 with Router ID 4.4.4.4
------------------------------------------------------------------------
 I  - Intra area route,  E1 - Type 1 external route,  N1 - Type 1 NSSA route
 IA - Inter area route,  E2 - Type 2 external route,  N2 - Type 2 NSSA route
 *  - Selected route

 *Destination: ::/0
  Type       : IA                                     Cost     : 11
  NextHop    : FE80::F40D:0:93D0:1                    Interface: GE2/1/2
  AdvRouter  : 4.4.4.4                                Area     : 0.0.0.2
  Preference : 10

 *Destination: 2001::/64
  Type       : IA                                     Cost     : 2
  NextHop    : FE80::F40D:0:93D0:1                    Interface: GE2/1/2
  AdvRouter  : 3.3.3.3                                Area     : 0.0.0.0
  Preference : 10

 *Destination: 2001:1::/64
  Type       : IA                                     Cost     : 3
  NextHop    : FE80::F40D:0:93D0:1                    Interface: GE2/1/2
  AdvRouter  : 3.3.3.3                                Area     : 0.0.0.0
  Preference : 10

 *Destination: 2001:2::/64
```

```
 Type       : I                                     Cost     : 1
 Nexthop    : ::                                    Interface: GE2/1/2
 AdvRouter  : 4.4.4.4                               Area     : 0.0.0.2
 Preference : 10

*Destination: 2001:3::/64
 Type       : IA                                    Cost     : 4
 NextHop    : FE80::F40D:0:93D0:1                   Interface: GE2/1/2
 AdvRouter  : 3.3.3.3                               Area     : 0.0.0.0
 Preference : 10

Total: 5
Intra area: 1         Inter area: 4         ASE: 0         NSSA: 0
```

(4) 进一步减少 Stub 区域路由表规模，将 Area 2 配置为 Totally Stub 区域

# 配置 Router C，设置 Area 2 为 Totally Stub 区域。

```
[RouterC-ospfv3-1-area-0.0.0.2] stub no-summary
```

# 查看 Router D 的 OSPFv3 路由表，可以发现路由表项数目减少了，其他非直连路由都被抑制，只有缺省路由被保留。

```
[RouterD] display ospfv3 routing

              OSPFv3 Process 1 with Router ID 4.4.4.4
------------------------------------------------------------------------
 I  - Intra area route,  E1 - Type 1 external route,  N1 - Type 1 NSSA route
 IA - Inter area route,  E2 - Type 2 external route,  N2 - Type 2 NSSA route
 *  - Selected route

 *Destination: ::/0
  Type       : IA                                    Cost     : 11
  NextHop    : FE80::F40D:0:93D0:1                   Interface: GE2/1/2
  AdvRouter  : 4.4.4.4                               Area     : 0.0.0.2
  Preference : 10

 *Destination: 2001:2::/64
  Type       : I                                     Cost     : 1
  Nexthop    : ::                                    Interface: GE2/1/2
  AdvRouter  : 4.4.4.4                               Area     : 0.0.0.2
  Preference : 10

 Total: 2
 Intra area: 1         Inter area: 1         ASE: 0         NSSA: 0
```

## 1.13.2 配置OSPFv3 的NSSA区域

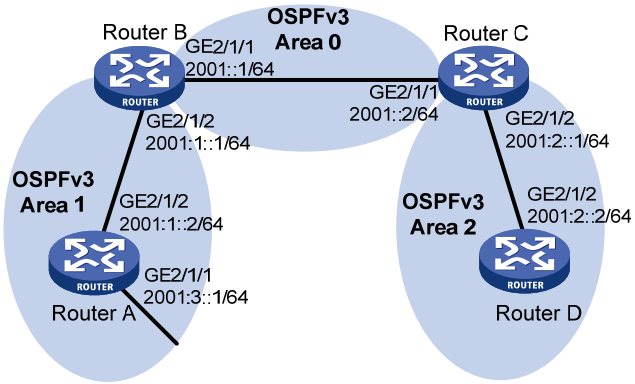
### 1. 组网需求

- 所有的路由器都运行 OSPFv3，整个自治系统划分为 3 个区域。其中 Router B 和 Router C 作为 ABR 来转发区域之间的路由。

- 要求将 Area 1 配置为 NSSA 区域，同时将 Router A 配置为 ASBR 引入外部路由（静态路由），且路由信息可正确的在 AS 内传播。

## 2. 组网图

图1-2 OSPFv3 的 NSSA 区域配置组网图




## 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置OSPFv3 基本功能（同前例 1.13.1 ）

(3) 配置 Area 1 为 NSSA 区域

# 配置 Router A 的 NSSA 区域。

```
<RouterA> system-view
[RouterA] ospfv3
[RouterA-ospfv3-1] area 1
[RouterA-ospfv3-1-area-0.0.0.2] nssa
[RouterA-ospfv3-1-area-0.0.0.2] quit
[RouterA-ospfv3-1] quit
```

# 配置 Router B 的 NSSA 区域。

```
<RouterB> system-view
[RouterB] ospfv3
[RouterB-ospfv3-1] area 1
[RouterB-ospfv3-1-area-0.0.0.2] nssa
[RouterB-ospfv3-1-area-0.0.0.2] quit
[RouterB-ospfv3-1] quit
```

说明

- 如果 NSSA 区域内路由器（Router A）需要获取通往 AS 内其他区域的路由，ABR（Router B）上必须配置 **default-route-advertise** 参数，这样 Router A 才可以获取到缺省路由。
- 建议在 ABR( Router B )上配置 **no-summary** 参数，这样可以减少 NSSA 路由器的路由表数量。其他 NSSA 路由器只需配置 **nssa** 命令就可以。

# 查看 Router A 的 OSPFv3 路由表信息。

```
[RouterA] display ospfv3 1 routing

              OSPFv3 Process 1 with Router ID 1.1.1.1
-------------------------------------------------------------------------
 I  - Intra area route,  E1 - Type 1 external route,  N1 - Type 1 NSSA route
 IA - Inter area route,  E2 - Type 2 external route,  N2 - Type 2 NSSA route
 *  - Selected route

 *Destination: 2001::/64
  Type       : IA                                    Cost     : 2
  NextHop    : FE80::20C:29FF:FE74:59C6              Interface: GE2/1/1
  AdvRouter  : 2.2.2.2                               Area     : 0.0.0.1
  Preference : 10

 *Destination: 2001:1::/64
  Type       : I                                     Cost     : 1
  Nexthop    : ::                                    Interface: GE2/1/1
  AdvRouter  : 1.1.1.1                               Area     : 0.0.0.2
  Preference : 10

 *Destination: 2001:2::/64
  Type       : IA                                    Cost     : 3
  NextHop    : FE80::20C:29FF:FE74:59C6              Interface: GE2/1/1
  AdvRouter  : 2.2.2.2                               Area     : 0.0.0.1
  Preference : 10

 Total: 3
 Intra area: 1         Inter area: 2         ASE: 0         NSSA: 0
```

(4) 配置 Router A 引入静态路由

# 配置 Router A 上的静态路由，并配置 OSPFv3 引入静态路由。

```
[RouterA] ipv6 route-static 1234:: 64 null 0
[RouterA] ospfv3 1
[RouterA-ospfv3-1] import-route static
[RouterA-ospfv3-1] quit
```

# 查看 Router D 的 OSPFv3 路由表，可以看到 NSSA 区域引入的一条 AS 外部的路由。

```
[RouterD] display ospfv3 1 routing

              OSPFv3 Process 1 with Router ID 4.4.4.4
-------------------------------------------------------------------------
 I  - Intra area route,  E1 - Type 1 external route,  N1 - Type 1 NSSA route
 IA - Inter area route,  E2 - Type 2 external route,  N2 - Type 2 NSSA route
 *  - Selected route

 *Destination: 2001::/64
  Type       : IA                                    Cost     : 2
  NextHop    : FE80::20C:29FF:FEB9:F2EF              Interface: GE2/1/2
  AdvRouter  : 3.3.3.3                               Area     : 0.0.0.2
  Preference : 10
```

```
*Destination: 2001:1::/64
 Type       : IA                                    Cost     : 3
 NextHop    : FE80::20C:29FF:FEB9:F2EF              Interface: GE2/1/2
 AdvRouter  : 3.3.3.3                               Area     : 0.0.0.2
 Preference : 10

*Destination: 2001:2::/64
 Type       : I                                     Cost     : 1
 NextHop    : ::                                    Interface: GE2/1/2
 AdvRouter  : 4.4.4.4                               Area     : 0.0.0.2
 Preference : 10

*Destination: 1234::/64
 Type       : E2                                    Cost     : 1
 NextHop    : FE80::20C:29FF:FEB9:F2EF              Interface: GE2/1/2
 AdvRouter  : 2.2.2.2                               Tag      : 0
 Preference : 10

Total: 4
Intra area: 1         Inter area: 2         ASE: 1         NSSA: 0
```

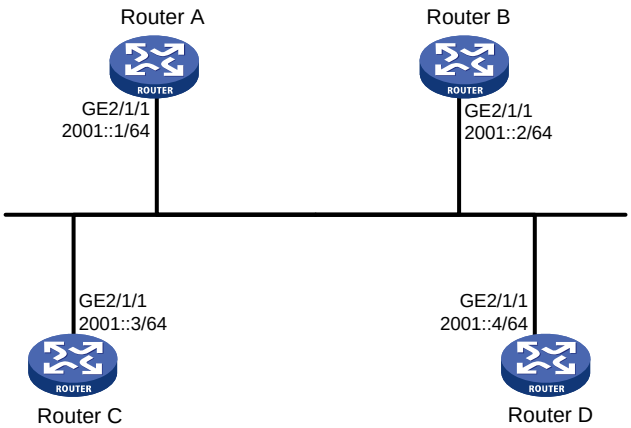
## 1.13.3 配置OSPFv3 的DR选择

### 1. 组网需求

- Router A 的优先级配置为 100，它是网络上的最高优先级，所以 Router A 被选为 DR；
- Router C 的优先级配置为 2，它是优先级次高的，被选为 BDR；
- Router B 的优先级配置为 0，这意味着它将无法成为 DR；
- Router D 没有配置优先级，取缺省值 1。

### 2. 组网图

图1-3 OSPFv3 的 DR 选择配置组网图



### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 OSPFv3 基本功能

# 配置 Router A，启动 OSPFv3，并配置其 Router ID 为 1.1.1.1。

```
<RouterA> system-view
[RouterA] ospfv3
[RouterA-ospfv3-1] router-id 1.1.1.1
[RouterA-ospfv3-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B，启动 OSPFv3，并配置其 Router ID 为 2.2.2.2。

```
<RouterB> system-view
[RouterB] ospfv3
[RouterB-ospfv3-1] router-id 2.2.2.2
[RouterB-ospfv3-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterB-GigabitEthernet2/1/1] quit
```

# 配置 Router C，启动 OSPFv3，并配置其 Router ID 为 3.3.3.3。

```
<RouterC> system-view
[RouterC] ospfv3
[RouterC-ospfv3-1] router-id 3.3.3.3
[RouterC-ospfv3-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterC-GigabitEthernet2/1/1] quit
```

# 配置 Router D，启动 OSPFv3，并配置其 Router ID 为 4.4.4.4。

```
<RouterD> system-view
[RouterD] ospfv3
[RouterD-ospfv3-1] router-id 4.4.4.4
[RouterD-ospfv3-1] quit
[RouterD] interface gigabitethernet 2/1/1
[RouterD-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterD-GigabitEthernet2/1/1] quit
```

# 查看 Router A 的邻居信息，可以看到 DR 优先级（缺省为 1）以及邻居状态。此时优先级相等，Router ID 大者被选为 DR，可以看到 Router D 为 DR，Router C 为 BDR。

```
[RouterA] display ospfv3 peer

              OSPFv3 Process 1 with Router ID 1.1.1.1

 Area: 0.0.0.0
--------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 2.2.2.2         1   2-Way/DROther     00:00:36  0      GE2/1/1
 3.3.3.3         1   Full/BDR          00:00:35  0      GE2/1/1
```

```
 4.4.4.4         1   Full/DR           00:00:33  0      GE2/1/1
```

# 查看 Router D 的邻居信息，可以看到 Router D 和其他邻居之间的邻居状态都为 Full。

```
[RouterD] display ospfv3 peer

              OSPFv3 Process 1 with Router ID 4.4.4.4

 Area: 0.0.0.0
------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 1.1.1.1         1   Full/DROther      00:00:30  0      GE2/1/1
 2.2.2.2         1   Full/DROther      00:00:37  0      GE2/1/1
 3.3.3.3         1   Full/BDR          00:00:31  0      GE2/1/1
```

(3) 配置接口的 DR 优先级

# 配置 Router A 的接口 GigabitEthernet2/1/1 的 DR 优先级为 100。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospfv3 dr-priority 100
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B 的接口 GigabitEthernet2/1/1 的 DR 优先级为 0。

```
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospfv3 dr-priority 0
[RouterB-GigabitEthernet2/1/1] quit
```

# 配置 Router C 的接口 GigabitEthernet2/1/1 的 DR 优先级为 2。

```
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ospfv3 dr-priority 2
[RouterC-GigabitEthernet2/1/1] quit
```

# 显示 Router A 的邻居信息，可以看到 DR 优先级已经更新，但 DR/BDR 并未改变。

```
[RouterA] display ospfv3 peer

              OSPFv3 Process 1 with Router ID 1.1.1.1

 Area: 0.0.0.0
------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 2.2.2.2         1   2-Way/DROther     00:00:36  0      GE2/1/1
 3.3.3.3         1   Full/BDR          00:00:35  0      GE2/1/1
 4.4.4.4         1   Full/DR           00:00:33  0      GE2/1/1
```

# 显示 Router D 的邻居信息，可以看到 Router D 仍然为 DR。

```
[RouterD] display ospfv3 peer

              OSPFv3 Process 1 with Router ID 4.4.4.4

 Area: 0.0.0.0
------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 1.1.1.1         1   Full/DROther      00:00:30  0      GE2/1/1
 2.2.2.2         1   Full/DROther      00:00:37  0      GE2/1/1
 3.3.3.3         1   Full/BDR          00:00:31  0      GE2/1/1
```

(4) 重新进行 DR/BDR 选择

# 将所有接口进行一次**shutdown**和**undo shutdown**，使OSPFv3进行DR/BDR的重新选举（略）。

# 查看 Router A 的邻居信息，可以看到 Router C 为 BDR。

```
[RouterA] display ospfv3 peer

               OSPFv3 Process 1 with Router ID 1.1.1.1

 Area: 0.0.0.0
------------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 2.2.2.2         0   Full/DROther      00:00:36  0      GE2/1/1
 3.3.3.3         2   Full/BDR          00:00:35  0      GE2/1/1
 4.4.4.4         1   Full/DROther      00:00:33  0      GE2/1/1
```

# 查看 Router D 的邻居信息，可以看到 Router A 为 DR。

```
[RouterD] display ospfv3 peer

               OSPFv3 Process 1 with Router ID 4.4.4.4

 Area: 0.0.0.0
------------------------------------------------------------------------------
 Router ID       Pri State             Dead-Time InstID Interface
 1.1.1.1         100 Full/DR           00:00:30  0      GE2/1/1
 2.2.2.2         0   2-Way/DROther     00:00:37  0      GE2/1/1
 3.3.3.3         2   Full/BDR          00:00:31  0      GE2/1/1
```

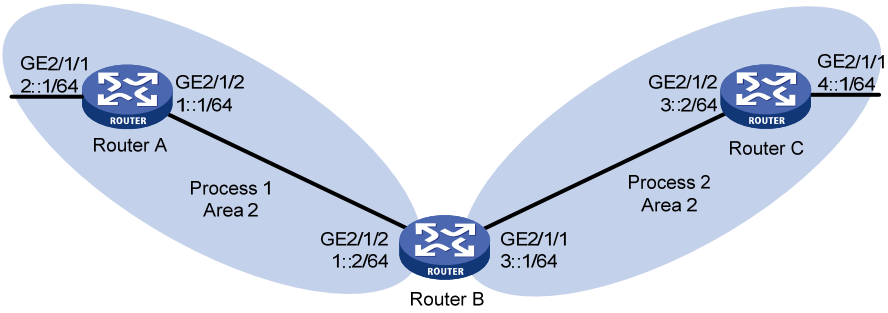
## 1.13.4 配置OSPFv3 引入外部路由

### 1. 组网需求

- Router A、Router B 和 Router C 位于 Area 2 内；
- Router B 上运行两个 OSPFv3 进程：OSPFv3 1 和 OSPFv3 2。Router B 通过 OSPFv3 1 和 Router A 交换路由信息，通过 OSPFv3 2 和 Router C 交换路由信息；
- 在 Router B 上配置 OSPFv3 进程 2 引入外部路由，引入直连路由和 OSPFv3 进程 1 的路由，并将引入的外部路由的缺省度量值设置为 3，使得 Router C 能够学习到达 1::0/64 和 2::0/64 的路由，但 Router A 不能学习到达 3::0/64 和 4::0/64 的路由。

## 2. 组网图

图1-4 配置 OSPFv3 引入外部路由组网图




## 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 OSPFv3

# 在 Router A 上启动 OSPFv3 进程 1。

```
<RouterA> system-view
[RouterA] ospfv3 1
[RouterA-ospfv3-1] router-id 1.1.1.1
[RouterA-ospfv3-1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] ospfv3 1 area 2
[RouterA-GigabitEthernet2/1/2] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospfv3 1 area 2
[RouterA-GigabitEthernet2/1/1] quit
```

# 在 Router B 上启动两个 OSPFv3 进程，进程号分别为 1 和 2。

```
<RouterB> system-view
[RouterB] ospfv3 1
[RouterB-ospfv3-1] router-id 2.2.2.2
[RouterB-ospfv3-1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] ospfv3 1 area 2
[RouterB-GigabitEthernet2/1/2] quit
[RouterB] ospfv3 2
[RouterB-ospfv3-2] router-id 3.3.3.3
[RouterB-ospfv3-2] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospfv3 2 area 2
[RouterB-GigabitEthernet2/1/1] quit
```

# 在 Router C 上启动 OSPFv3 进程 2。

```
<RouterC> system-view
[RouterC] ospfv3 2
[RouterC-ospfv3-2] router-id 4.4.4.4
[RouterC-ospfv3-2] quit
```

```
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] ospfv3 2 area 2
[RouterC-GigabitEthernet2/1/2] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ospfv3 2 area 2
[RouterC-GigabitEthernet2/1/1] quit
```

# 查看 Router C 的路由表信息。

```
[RouterC] display ipv6 routing-table

Destinations : 7 Routes : 7

Destination: ::1/128                                Protocol  : Direct
NextHop    : ::1                                    Preference: 0
Interface  : InLoop0                                Cost      : 0

Destination: 3::/64                                 Protocol  : Direct
NextHop    : 3::2                                   Preference: 0
Interface  : GE2/1/2                                Cost      : 0

Destination: 3::2/128                               Protocol  : Direct
NextHop    : ::1                                    Preference: 0
Interface  : InLoop0                                Cost      : 0

Destination: 4::/64                                 Protocol  : Direct
NextHop    : 4::1                                   Preference: 0
Interface  : GE2/1/1                                Cost      : 0

Destination: 4::1/128                               Protocol  : Direct
NextHop    : ::1                                    Preference: 0
Interface  : InLoop0                                Cost      : 0

Destination: FE80::/10                              Protocol  : Direct
NextHop    : ::                                     Preference: 0
Interface  : NULL0                                  Cost      : 0

Destination: FF00::/8                               Protocol  : Direct
NextHop    : ::                                     Preference: 0
Interface  : NULL0
```

(3) 配置 OSPFv3 引入外部路由

# 在 Router B 上配置 OSPFv3 进程 2 引入外部路由,，引入直连路由和 OSPFv3 进程 1 的路由。

```
[RouterB] ospfv3 2
[RouterB-ospfv3-2] default cost 3
[RouterB-ospfv3-2] import-route ospfv3 1
[RouterB-ospfv3-2] import-route direct
[RouterB-ospfv3-2] quit
```

# 查看路由引入后 Router C 的路由表信息。

```
[RouterC] display ipv6 routing-table
```

```
Destinations : 9 Routes : 9

Destination: ::1/128                                    Protocol  : Direct
NextHop    : ::1                                        Preference: 0
Interface  : InLoop0                                    Cost      : 0

Destination: 1::/64                                     Protocol  : OSPFv3
NextHop    : FE80::200:CFF:FE01:1C03                    Preference: 150
Interface  : GE2/1/2                                    Cost      : 3

Destination: 2::/64                                     Protocol  : OSPFv3
NextHop    : FE80::200:CFF:FE01:1C03                    Preference: 150
Interface  : GE2/1/2                                    Cost      : 3

Destination: 3::/64                                     Protocol  : Direct
NextHop    : 3::2                                       Preference: 0
Interface  : GE2/1/2                                    Cost      : 0

Destination: 3::2/128                                   Protocol  : Direct
NextHop    : ::1                                        Preference: 0
Interface  : InLoop0                                    Cost      : 0

Destination: 4::/64                                     Protocol  : Direct
NextHop    : 4::1                                       Preference: 0
Interface  : GE2/1/1                                    Cost      : 0

Destination: 4::1/128                                   Protocol  : Direct
NextHop    : ::1                                        Preference: 0
Interface  : InLoop0                                    Cost      : 0

Destination: FE80::/10                                  Protocol  : Direct
NextHop    : ::                                         Preference: 0
Interface  : NULL0                                      Cost      : 0

Destination: FF00::/8                                   Protocol  : Direct
NextHop    : ::                                         Preference: 0
Interface  : NULL0
```

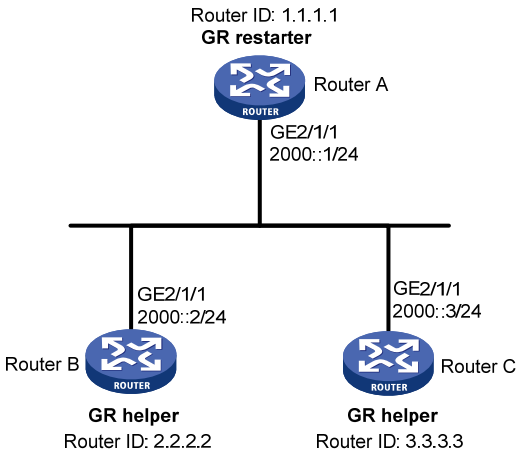
### 1.13.5 配置OSPFv3 GR

#### 1. 组网需求

- Router A、Router B 和 Router C 既属于同一自治系统，也属于同一 OSPFv3 区域，通过 OSPFv3 协议实现网络互连，并提供 GR 机制；
- Router A 作为 GR Restarter，Router B 和 Router C 作为 GR Helper 并且通过 GR 机制与 Router A 保持带外同步。

## 2. 组网图

图1-5 配置 OSPFv3 的 GR 组网图




## 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 OSPFv3 基本功能

# 配置 Router A，启动 OSPFv3，并设置其 Router ID 为 1.1.1.1。

```
<RouterA> system-view
[RouterA] ospfv3 1
[RouterA-ospfv3-1] router-id 1.1.1.1
[RouterA-ospfv3-1] graceful-restart enable
[RouterA-ospfv3-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospfv3 1 area 1
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B，启动 OSPFv3，并设置其 Router ID 为 2.2.2.2。缺省情况下，Router B 的 GR helper 能力处于开启状态。

```
<RouterB> system-view
[RouterB] ospfv3 1
[RouterB-ospfv3-1] router-id 2.2.2.2
[RouterB-ospfv3-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospfv3 1 area 1
[RouterB-GigabitEthernet2/1/1] quit
```

# 在 Router C，启动 OSPFv3，并设置其 Router ID 为 3.3.3.3。缺省情况下，Router C 的 GR helper 能力处于开启状态。

```
<RouterC> system-view
[RouterC] ospfv3 1
[RouterC-ospfv3-1] router-id 3.3.3.3
[RouterC-ospfv3-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ospfv3 1 area 1
[RouterC-GigabitEthernet2/1/1] quit
```

#### 4. 验证配置

运行稳定后，在 Router A 上主备倒换进入 OSPFv3 协议的 GR 进程。

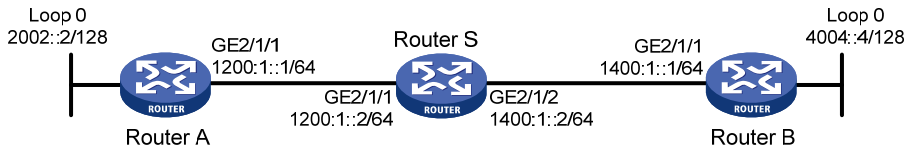
### 1.13.6 配置OSPFv3 NSR

#### 1. 组网需求

- Router A、Router B 和 Router S 既属于同一自治系统，也属于同一 OSPFv3 区域，通过 OSPFv3 协议实现网络互连。Router S 为分布式设备，提供 NSR 机制；
- 当 Router S 进行主备倒换时，Router A 和 Router B 与 Router S 的邻居没有中断，Router A 到 Router B 的流量没有中断。

#### 2. 组网图

图1-6 配置 OSPFv3 的 NSR 组网图




#### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 OSPFv3 基本功能

# 配置 Router A，启动 OSPFv3，并设置其 Router ID 为 1.1.1.1。

```
<RouterA> system-view
[RouterA] ospfv3 1
[RouterA-ospfv3-1] router-id 1.1.1.1
[RouterA-ospfv3-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospfv3 1 area 1
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B，启动 OSPFv3，并设置其 Router ID 为 2.2.2.2。

```
<RouterB> system-view
[RouterB] ospfv3 1
[RouterB-ospfv3-1] router-id 2.2.2.2
[RouterB-ospfv3-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospfv3 1 area 1
[RouterB-GigabitEthernet2/1/1] quit
```

# 配置 Router S，启动 OSPFv3，并设置其 Router ID 为 3.3.3.3。使能 NSR 能力。

```
<RouterS> system-view
[RouterS] ospfv3 1
[RouterS-ospfv3-1] router-id 3.3.3.3
[RouterS-ospfv3-1] non-stop-routing
[RouterS-ospfv3-1] quit
[RouterS] interface gigabitethernet 2/1/1
```

```
[RouterS-GigabitEthernet2/1/1] ospfv3 1 area 1
[RouterS-GigabitEthernet2/1/1] quit
[RouterS] interface gigabitethernet 2/1/2
[RouterS-GigabitEthernet2/1/2] ospfv3 1 area 1
[RouterS-GigabitEthernet2/1/2] quit
```

#### 4. 验证配置

运行稳定后，在 Router S 上主备倒换进入 OSPFv3 协议的 NSR 阶段，保证倒换期间不断流，倒换后平滑升级。

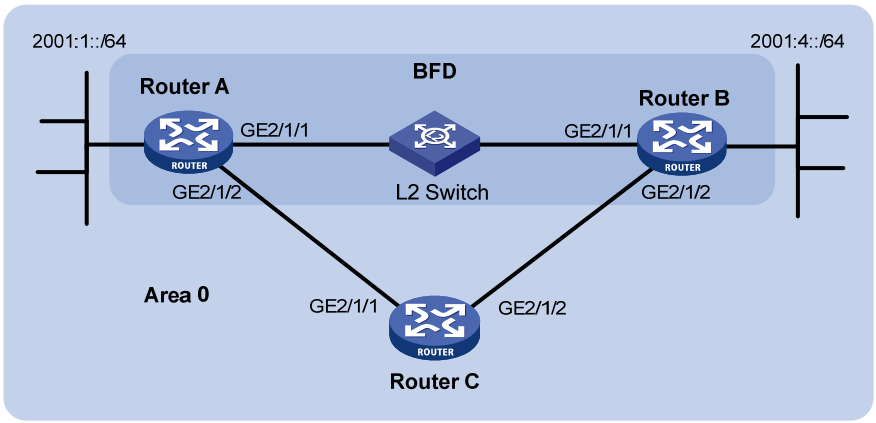
### 1.13.7 配置OSPFv3 与BFD联动

#### 1. 组网需求

- Router A、Router B 和 Router C 上运行 OSPFv3，网络层相互可达。
- 当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时 BFD 能够快速感知通告 OSPFv3 协议，并且切换到 Router C 进行通信。

#### 2. 组网图

图1-7 配置 OSPFv3 与 BFD 联动组网图




| 设备 | 接口 | IPv6地址 | 设备 | 接口 | IPv6地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 2001::1/64 | Router B | GE2/1/1 | 2001::2/64 |
| | GE2/1/2 | 2001:2::1/64 | | GE2/1/2 | 2001:3::2/64 |
| Router C | GE2/1/1 | 2001:2::2/64 | | | |
| | GE2/1/2 | 2001:3::1/64 | | | |

#### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 OSPFv3 基本功能

# 配置 Router A，启动 OSPFv3，并设置其 Router ID 为 1.1.1.1。

```
<RouterA> system-view
[RouterA] ospfv3 1
[RouterA-ospfv3-1] router-id 1.1.1.1
[RouterA-ospfv3-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospfv3 1 area 0
```

```
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] ospfv3 1 area 0
[RouterA-GigabitEthernet2/1/2] quit
```

# 配置 Router B，启动 OSPFv3，并设置其 Router ID 为 2.2.2.2。

```
<RouterB> system-view
[RouterB] ospfv3 1
[RouterB-ospfv3-1] router-id 2.2.2.2
[RouterB-ospfv3-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterB-GigabitEthernet2/1/1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] ospfv3 1 area 0
[RouterB-GigabitEthernet2/1/2] quit
```

# 配置 Router C，启动 OSPFv3，并设置其 Router ID 为 3.3.3.3。

```
<RouterC> system-view
[RouterC] ospfv3 1
[RouterC-ospfv3-1] router-id 3.3.3.3
[RouterC-ospfv3-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] ospfv3 1 area 0
[RouterC-GigabitEthernet2/1/2] quit
```

(3) 配置 BFD 功能

# 在 Router A 上使能 BFD 检测功能，并配置 BFD 参数。

```
[RouterA] bfd session init-mode active
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ospfv3 bfd enable
[RouterA-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterA-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterA-GigabitEthernet2/1/1] bfd detect-multiplier 7
[RouterA-GigabitEthernet2/1/1] return
```

# 在 Router B 上使能 BFD 检测功能，并配置 BFD 参数。

```
[RouterB] bfd session init-mode active
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospfv3 bfd enable
[RouterB-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterB-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterB-GigabitEthernet2/1/1] bfd detect-multiplier 6
```

### 4. 验证配置

下面以 Router A 为例，Router B 和 Router A 类似，不再赘述。

# 显示 Router A 的 BFD 信息。

```
<RouterA> display bfd session
```

```
 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv6 Session Working Under Ctrl Mode:

       Local Discr: 1441                 Remote Discr: 1450
         Source IP: FE80::20F:FF:FE00:1202（Router A 接口 GigabitEthernet2/1/1 的链路本地地址）
    Destination IP: FE80::20F:FF:FE00:1200（Router B 接口 GigabitEthernet2/1/1 的链路本地地址）
     Session State: Up                      Interface: GE2/1/1
         Hold Time: 2319ms
```

# 在 Router A 上查看 2001:4::0/64 的路由信息，可以看出 Router A 和 Router B 是通过 L2 Switch 进行通信的。

```
<RouterA> display ipv6 routing-table 2001:4::0 64

Summary Count : 1

Destination: 2001:4::/64                              Protocol  : OSPFv3
NextHop    : FE80::20F:FF:FE00:1200                   Preference: 10
Interface  : GE2/1/1                                  Cost      : 1
```

当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时：

# 在 Router A 上查看 2001:4::0/64 的路由信息，可以看出 Router A 和 Router B 已经切换到 Router C 进行通信。

```
<RouterA> display ipv6 routing-table 2001:4::0 64

Summary Count : 1

Destination: 2001:4::/64                              Protocol  : OSPFv3
NextHop    : FE80::BAAF:67FF:FE27:DCD0                Preference: 10
Interface  : GE2/1/2                                  Cost      : 2
```

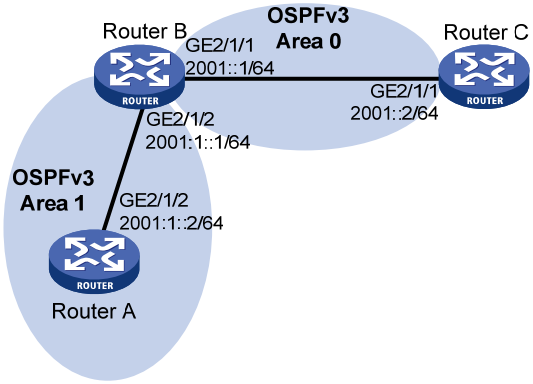
## 1.13.8 配置OSPFv3 IPsec安全框架

### 1. 组网需求

- 所有的路由器都运行 OSPFv3，整个自治系统划分为 2 个区域。
- 要求配置 IPsec 安全框架对 Router A、Router B 和 Router C 之间的 OSPFv3 报文进行有效性检查和验证。

## 2. 组网图

图1-8 配置 OSPFv3 IPsec 安全框架组网图




## 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 OSPFv3 基本功能

# 配置 Router A，启动 OSPFv3，并设置其 Router ID 为 1.1.1.1。

```
<RouterA> system-view
[RouterA] ospfv3 1
[RouterA-ospfv3-1] router-id 1.1.1.1
[RouterA-ospfv3-1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] ospfv3 1 area 1
[RouterA-GigabitEthernet2/1/2] quit
```

# 配置 Router B，启动 OSPFv3，并设置其 Router ID 为 2.2.2.2。

```
<RouterB> system-view
[RouterB] ospfv3 1
[RouterB-ospfv3-1] router-id 2.2.2.2
[RouterB-ospfv3-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterB-GigabitEthernet2/1/1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] ospfv3 1 area 1
[RouterB-GigabitEthernet2/1/2] quit
```

# 配置 Router C，启动 OSPFv3，并设置其 Router ID 为 3.3.3.3。

```
<RouterC> system-view
[RouterC] ospfv3 1
[RouterC-ospfv3-1] router-id 3.3.3.3
[RouterC-ospfv3-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] ospfv3 1 area 0
[RouterC-GigabitEthernet2/1/1] quit
```

(3) 配置 OSPFv3 IPsec 安全框架

# 配置 Router A。创建名为 trans 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议。创建一条安全框架 profile001，协商方式为 manual，配置 SPI 和密钥。

```
[RouterA] ipsec transform-set trans
[RouterA-ipsec-transform-set-trans] encapsulation-mode transport
[RouterA-ipsec-transform-set-trans] esp encryption-algorithm 3des-cbc
[RouterA-ipsec-transform-set-trans] esp authentication-algorithm md5
[RouterA-ipsec-transform-set-trans] ah authentication-algorithm md5
[RouterA-ipsec-transform-set-trans] quit
[RouterA] ipsec profile profile001 manual
[RouterA-ipsec-profile-profile001-manual] transform-set trans
[RouterA-ipsec-profile-profile001-manual] sa spi inbound ah 100000
[RouterA-ipsec-profile-profile001-manual] sa spi outbound ah 100000
[RouterA-ipsec-profile-profile001-manual] sa spi inbound esp 200000
[RouterA-ipsec-profile-profile001-manual] sa spi outbound esp 200000
[RouterA-ipsec-profile-profile001-manual] sa string-key inbound ah simple abc
[RouterA-ipsec-profile-profile001-manual] sa string-key outbound ah simple abc
[RouterA-ipsec-profile-profile001-manual] sa string-key inbound esp simple 123
[RouterA-ipsec-profile-profile001-manual] sa string-key outbound esp simple 123
[RouterA-ipsec-profile-profile001-manual] quit
```

# 配置 Router B。创建名为 trans 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议。创建一条安全框架 profile001，协商方式为 manual，配置 SPI 和密钥。创建一条安全框架 profile002，协商方式为 manual，配置 SPI 和密钥。

```
[RouterB] ipsec transform-set trans
[RouterB-ipsec-transform-set-trans] encapsulation-mode transport
[RouterB-ipsec-transform-set-trans] esp encryption-algorithm 3des-cbc
[RouterB-ipsec-transform-set-trans] esp authentication-algorithm md5
[RouterB-ipsec-transform-set-trans] ah authentication-algorithm md5
[RouterB-ipsec-transform-set-trans] quit
[RouterB] ipsec profile profile001 manual
[RouterB-ipsec-profile-profile001-manual] transform-set trans
[RouterB-ipsec-profile-profile001-manual] sa spi inbound ah 100000
[RouterB-ipsec-profile-profile001-manual] sa spi outbound ah 100000
[RouterB-ipsec-profile-profile001-manual] sa spi inbound esp 200000
[RouterB-ipsec-profile-profile001-manual] sa spi outbound esp 200000
[RouterB-ipsec-profile-profile001-manual] sa string-key inbound ah simple abc
[RouterB-ipsec-profile-profile001-manual] sa string-key outbound ah simple abc
[RouterB-ipsec-profile-profile001-manual] sa string-key inbound esp simple 123
[RouterB-ipsec-profile-profile001-manual] sa string-key outbound esp simple 123
[RouterB-ipsec-profile-profile001-manual] quit
[RouterB] ipsec profile profile002 manual
[RouterB-ipsec-profile-profile002-manual] transform-set trans
[RouterB-ipsec-profile-profile002-manual] sa spi inbound ah 400000
[RouterB-ipsec-profile-profile002-manual] sa spi outbound ah 400000
[RouterB-ipsec-profile-profile002-manual] sa spi inbound esp 256
[RouterB-ipsec-profile-profile002-manual] sa spi outbound esp 256
[RouterB-ipsec-profile-profile002-manual] sa string-key inbound ah simple hello
[RouterB-ipsec-profile-profile002-manual] sa string-key outbound ah simple hello
```

```
[RouterB-ipsec-profile-profile002-manual] sa string-key inbound esp simple byebye
[RouterB-ipsec-profile-profile002-manual] sa string-key outbound esp simple byebye
[RouterB-ipsec-profile-profile002-manual] quit
```

# 配置 Router C。创建名为 trans 的安全提议，报文封装形式采用传输模式，安全协议采用 ESP 协议。创建一条安全框架 profile002，协商方式为 manual，配置 SPI 和密钥。

```
[RouterC] ipsec transform-set trans
[RouterC-ipsec-transform-set-trans] encapsulation-mode transport
[RouterC-ipsec-transform-set-trans] esp encryption-algorithm 3des-cbc
[RouterC-ipsec-transform-set-trans] esp authentication-algorithm md5
[RouterC-ipsec-transform-set-trans] ah authentication-algorithm md5
[RouterC-ipsec-transform-set-trans] quit
[RouterC] ipsec profile profile002 manual
[RouterC-ipsec-profile-profile002-manual] transform-set trans
[RouterC-ipsec-profile-profile002-manual] sa spi inbound ah 400000
[RouterC-ipsec-profile-profile002-manual] sa spi outbound ah 400000
[RouterC-ipsec-profile-profile002-manual] sa spi inbound esp 256
[RouterC-ipsec-profile-profile002-manual] sa spi outbound esp 256
[RouterC-ipsec-profile-profile002-manual] sa string-key inbound ah simple hello
[RouterC-ipsec-profile-profile002-manual] sa string-key outbound ah simple hello
[RouterC-ipsec-profile-profile002-manual] sa string-key inbound esp simple byebye
[RouterC-ipsec-profile-profile002-manual] sa string-key outbound esp simple byebye
[RouterC-ipsec-profile-profile002-manual] quit
```

(4) 配置 OSPFv3 区域上应用 IPsec 安全框架

# 配置 Router A。

```
[RouterA] ospfv3 1
[RouterA-ospfv3-1] area 1
[RouterA-ospfv3-1-area-0.0.0.1] enable ipsec-profile profile001
[RouterA-ospfv3-1-area-0.0.0.1] quit
[RouterA-ospfv3-1] quit
```

# 配置 Router B。

```
[RouterB] ospfv3 1
[RouterB-ospfv3-1] area 0
[RouterB-ospfv3-1-area-0.0.0.0] enable ipsec-profile profile002
[RouterB-ospfv3-1-area-0.0.0.0] quit
[RouterB-ospfv3-1] area 1
[RouterB-ospfv3-1-area-0.0.0.1] enable ipsec-profile profile001
[RouterB-ospfv3-1-area-0.0.0.1] quit
[RouterB-ospfv3-1] quit
```

# 配置 Router C。

```
[RouterC] ospfv3 1
[RouterC-ospfv3-1] area 0
[RouterC-ospfv3-1-area-0.0.0.0] enable ipsec-profile profile002
[RouterC-ospfv3-1-area-0.0.0.0] quit
[RouterC-ospfv3-1] quit
```

### 4. 验证配置

以上配置完成后，Router A、Router B 和 Router C 之间的 OSPFv3 报文将被加密传输。

# 目 录

# 1 IPv6 IS-IS

IPv6 IS-IS 实现了 IPv4 IS-IS 的所有功能，与 IPv4 IS-IS 的区别在于发布的是 IPv6 路由信息，本章只列出了 IPv6 IS-IS 专有的配置任务，其他相关配置任务请参见“三层技术-IP 路由配置指导”中的“IS-IS”。

## 1.1 IPv6 IS-IS简介

IS-IS（Intermediate System-to-Intermediate System，中间系统到中间系统）支持多种网络层协议，其中包括 IPv6 协议，支持 IPv6 协议的 IS-IS 路由协议又称为 IPv6 IS-IS 动态路由协议。IETF 的 draft-ietf-isis-ipv6-05 中规定了 IS-IS 为支持 IPv6 所新增的内容，主要是新添加的支持 IPv6 协议的两个 TLV（Type-Length-Values）和一个新的 NLPID（Network Layer Protocol Identifier，网络层协议标识符）。

TLV 是 LSP（Link State PDU，链路状态协议数据单元）中的一个可变长字段值。新增的两个 TLV 分别是：

- IPv6 Reachability：类型值为 236（0xEC），通过定义路由信息前缀、度量值等信息来说明网络的可达性。
- IPv6 Interface Address：类型值为 232（0xE8），它对应于 IPv4 中的“IP Interface Address” TLV，只不过把原来的 32 比特的 IPv4 地址改为 128 比特的 IPv6 地址。

NLPID 是标识网络层协议报文的一个 8 比特字段，IPv6 的 NLPID 值为 142（0x8E）。如果 IS-IS 路由器支持 IPv6，那么它必须以这个 NLPID 值向外发布路由信息。

## 1.2 配置IPv6 IS-IS的基本特性

在 IPv6 网络环境中，可以通过配置 IPv6 IS-IS 路由协议来实现 IPv6 网络的互连。

### 1.2.1 配置准备

在配置之前，需完成以下任务：

- 配置接口的网络层地址，使各相邻节点网络层可达
- 启动 IS-IS

### 1.2.2 配置IS-IS的IPv6 基本特性

表1-1 配置 IS-IS 的 IPv6 基本特性

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 启动IS-IS路由进程，进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | 缺省情况下，系统没有运行IS-IS |
| 配置网络实体名称（NET） | **network-entity** *net* | 缺省情况下，没有配置NET |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 创建并进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | 缺省情况下，没有创建IS-IS IPv6地址族视图 |
| 退回到IS-IS视图 | **quit** | - |
| 退回到系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能接口IS-IS路由进程的IPv6能力并指定要关联的IS-IS进程号 | **isis ipv6 enable** [ *process-id* ] | 缺省情况下，接口上没有使能IS-IS路由进程的IPv6能力 |

## 1.3 配置IPv6 IS-IS的路由信息控制

### 1.3.1 配置准备

在进行 IPv6 IS-IS 的路由特性配置之前，需完成 IPv6 IS-IS 基本配置。

### 1.3.2 配置IPv6 IS-IS的路由信息控制

表1-2 配置 IPv6 IS-IS 的路由信息控制

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 进入IS-IS IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 配置IPv6 IS-IS路由优先级 | **preference** { **route-policy** *route-policy-name* \| *preference* } * | 缺省情况下，IPv6 IS-IS路由优先级为15 |
| 配置IPv6 IS-IS聚合路由 | **summary** *ipv6-prefix prefix-length* [ **avoid-feedback** \| **generate_null0_route** \| [ **level-1** \| **level-1-2** \| **level-2** ] \| **tag** *tag* ] * | 缺省情况下，没有配置IPv6聚合路由 |
| 配置生成IPv6 IS-IS缺省路由 | **default-route-advertise** [ **avoid-learning** \| [ **level-1** \| **level-1-2** \| **level-2** ] \| **route-policy** *route-policy-name* \| **tag** *tag* ] * | 缺省情况下，不生成IPv6 IS-IS缺省路由 |
| 配置IPv6 IS-IS对引入的路由进行过滤 | **filter-policy** { *acl6-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } **export** [ *protocol* [ *process-id* ] ] | 缺省情况下，IPv6 IS-IS不对引入的路由信息进行过滤<br>本命令一般和**import-route**命令结合使用 |
| 配置IPv6 IS-IS对接收的路由进行过滤 | **filter-policy** { *acl6-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } **import** | 缺省情况下，IPv6 IS-IS不对接收的路由进行过滤 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置IPv6 IS-IS引入其他协议的路由信息 | **import-route** *protocol* [ *process-id* ] [ **allow-ibgp** ] [ **allow-direct** \| **cost** *cost* \| [ **level-1** \| **level-1-2** \| **level-2** ] \| **route-policy** *route-policy-name* \| **tag** *tag* ] * | 缺省情况下，IPv6 IS-IS不引入其他协议的路由信息 |
| 配置引入Level1/Level2的IPv6路由最大条数 | **import-route limit** *number* | （number 的缺省情况如下）<br>MSR 2600：200000<br>MSR 36-10、MSR 3600-28/ MSR 3600-51：200000<br>其他MSR 3600：500000<br>MSR 5600：1000000 |
| 配置从Level-2向Level-1进行路由渗透 | **import-route isisv6 level-2 into level-1** [ **filter-policy** { *acl6-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } \| **tag** *tag* ] * | 缺省情况下，不从Level-2向Level-1进行路由渗透 |
| 配置从Level-1向Level-2进行路由渗透 | **import-route isisv6 level-1 into level-2** [ **filter-policy** { *acl6-number* \| **prefix-list** *prefix-list-name* \| **route-policy** *route-policy-name* } \| **tag** *tag* ] * | 缺省情况下，从Level-1向Level-2进行路由渗透 |
| 配置在负载分担方式下IPv6 IS-IS等价路由的最大数量 | **maximum load-balancing** *number* | 缺省情况下，IPv6 IS-IS支持的最大等价路由条数为32。 |

## 1.3.3 配置IS-IS IPv6 链路开销值

### 1. 配置接口的IPv6 链路开销值

表1-3 配置接口的 IPv6 链路开销值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS开销值的类型 | **cost-style** { **wide** \| **wide-compatible** \| **compatible** } | 缺省情况下，IS-IS只收发采用narrow方式的报文 |
| 进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 配置IS-IS IPv6单播拓扑 | **multi-topology** [ **compatible** ] | 缺省情况下，没有配置支持IPv6拓扑 |
| 退回到IS-IS视图 | **quit** | - |
| 退回到系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能接口IS-IS路由进程的IPv6能力并指定要关联的IS-IS进程号 | **isis ipv6 enable** [ *process-id* ] | 缺省情况下，接口在指定拓扑上没有配置IPv6链路开销值 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置接口的IPV6链路开销值 | **isis ipv6 cost** | 缺省情况下，接口没有配置ipv6 cost |

### 2. 全局配置IPv6 IS-IS的链路开销值

表1-4 全局配置 IPv6 IS-IS 链路开销值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS开销值的类型 | **cost-style** { **wide** \| **wide-compatible** \| **compatible** } | 缺省情况下，IS-IS只收发采用narrow方式的报文 |
| 进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 配置IS-IS IPv6单播拓扑 | **multi-topology** [ **compatible** ] | 缺省情况下，没有配置支持IPv6拓扑 |
| 全局配置IPv6 IS-IS的链路开销值 | **circuit-cost** *value* [ **level-1** \| **level-2** ] | 缺省情况下，没有全局配置IPv6 IS-IS的链路开销值 |

### 3. 配置IS-IS自动计算链路开销值

表1-5 配置 IS-IS 自动计算链路开销值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS开销值的类型 | **cost-style** { **wide** \| **wide-compatible** } | 缺省情况下，IS-IS只收发采用narrow方式的报文 |
| 进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 配置IS-IS IPv6单播拓扑 | **multi-topology** [ **compatible** ] | 缺省情况下，没有配置支持IPv6拓扑 |
| 使能自动计算接口链路开销值功能 | **auto-cost enable** | 缺省情况下，自动计算接口链路开销值功能处于关闭状态 |
| （可选）配置IPv6 IS-IS自动计算链路开销值时依据的带宽参考值 | **bandwidth-reference** *value* | 缺省情况下，带宽参考值为100Mbps |

## 1.4 调整和优化IPv6 IS-IS网络

### 1.4.1 配置准备

在进行 IPv6 IS-IS 的路由特性配置之前，需完成 IPv6 IS-IS 基本配置。

### 1.4.2 配置优先级参数

IS-IS 协议中，当网络拓扑发生变化时，路由要重新收敛。IPv6 IS-IS 路由收敛的优先级由高到低包括：

- critical：最高优先级。
- high：高优先级。
- medium：中优先级。
- 低优先级：缺省优先级。需要注意的是，IPv6 IS-IS 主机路由的缺省优先级为中优先级。

IPv6 IS-IS 路由的优先级越高收敛的速度越快。

表1-6 配置优先级参数

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS开销值的类型 | **cost-style** { **wide** \| **wide-compatible** \| **compatible** } | 必选<br>缺省情况下，IS-IS只收发采用narrow方式的报文 |
| 进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | |
| 配置IS-IS IPv6单播拓扑 | **multi-topology** [ **compatible** ] | 缺省情况下，没有配置支持IPv6拓扑 |
| 配置指定IPv6 IS-IS路由收敛的优先级 | **prefix-priority** { **critical** \| **high** \| **medium** } { **prefix-list** *prefix-list-name* \| **tag** *tag-value* }<br>**prefix-priority route-policy** *route-policy-name* | 缺省情况下，IPv6 IS-IS路由收敛的优先级为低优先级 |

### 1.4.3 配置IPv6 拓扑的LSDB过载标志位

表1-7 配置 IPv6 拓扑的 LSDB 过载标志位

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS开销值的类型 | **cost-style** { **wide** \| **wide-compatible** \| **compatible** } | 缺省情况下，IS-IS只收发采用narrow方式的报文 |
| 进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 配置IS-IS IPv6单播拓扑 | **multi-topology** [ **compatible** ] | 缺省情况下，没有配置支持IPv6拓扑 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置IPv6拓扑的LSDB过载标志位 | **set-overload** [ **on-startup** [ [ **start-from-nbr** *system-id* [ *timeout1* [ *nbr-timeout* ] ] ] \| *timeout2* ] [ **allow** { **external** \| **interlevel** } * ] | 缺省情况下，不设置过载标志位 |

### 1.4.4 配置接口的Tag值

当 cost-sytle 为 wide、wide-compatible 或 compatible 时，如果发布可达的 IP 地址前缀具有 tag 属性，IS-IS 会将 tag 加入到该前缀的 IP 可达信息 TLV 中。

表1-8 配置接口的 Tag 值

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的Tag值 | **isis ipv6 tag** *tag* | 缺省情况下，没有配置接口的Tag值 |

### 1.4.5 配置IPv6 IS-IS的路由计算的时间间隔

表1-9 配置 IPv6 IS-IS 路由计算的时间间隔

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS开销值的类型 | **cost-style** { **wide** \| **wide-compatible \| compatible** } | 缺省情况下，IS-IS只收发采用narrow方式的报文 |
| 进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 配置IS-IS IPv6单播拓扑 | **multi-topology** [ **compatible** ] | 缺省情况下，没有配置支持IPv6拓扑 |
| 配置IPv6 IS-IS路由计算的时间间隔 | **timer spf** *maximum-interval* [ *minimum-interval* [ *incremental-interval* ] ] | 缺省情况下，IS-IS路由计算的最大时间间隔为5秒，最小时间间隔为50毫秒，时间间隔惩罚增量为200毫秒 |

### 1.4.6 配置IPv6 IS-IS 的ISPF

表1-10 配置 IPv6 IS-IS 的 ISPF

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS开销值的类型 | **cost-style** { **wide** \| **wide-compatible** \| **compatible** } | 缺省情况下，IS-IS只收发采用narrow方式的报文 |
| 进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 配置IS-IS IPv6单播拓扑 | **multi-topology** [ **compatible** ] | 缺省情况下，没有配置支持IPv6拓扑 |
| 使能IPV6 IS-IS ISPF功能，即增量SPF计算功能 | **ispf enable** | 缺省情况下，使能IPv6 IS-IS ISPF功能 |

### 1.4.7 配置前缀抑制

接口上配置本功能后，禁止此接口的前缀在 LSP 中携带，屏蔽内部节点被发布，提高安全性，加快路由收敛。

表1-11 配置前缀抑制

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 配置接口的前缀抑制功能 | **isis ipv6 prefix-suppression** | 缺省情况下，未配置接口的前缀抑制功能 |

## 1.5 配置IPv6 IS-IS与BFD联动

BFD（Bidirectional Forwarding Detection，双向转发检测）能够为 IPv6 IS-IS 邻居之间的链路提供快速检测功能。当邻居之间的链路出现故障时，加快 IPv6 IS-IS 协议的收敛速度。关于 BFD 的介绍和基本功能配置，请参见“可靠性配置指导”中的“BFD”。

表1-12 配置 IPv6 IS-IS 与 BFD 联动

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 启动IS-IS路由进程，进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置网络实体名称（NET） | **network-entity** *net* | 缺省情况下，没有配置NET |
| 进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 使能接口IS-IS路由进程的IPv6能力并指定要关联的IS-IS进程号 | **isis ipv6 enable** [ *process-id* ] | 缺省情况下，接口上没有使能IS-IS路由进程的IPv6能力 |
| 在指定接口上使能IPv6 IS-IS BFD | **isis ipv6 bfd enable** | 缺省情况下，IPv6 IS-IS的BFD功能处于关闭状态 |

# 1.6 配置IS-IS支持IPv6单播拓扑

说明

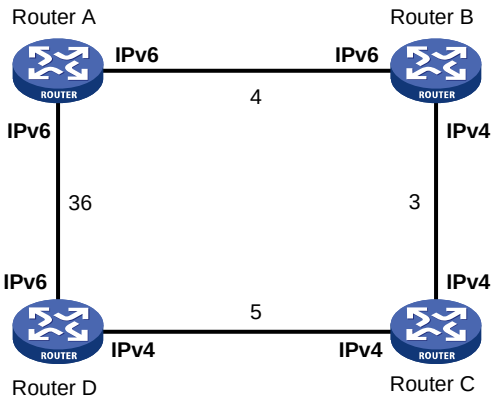
- 关于 MTR 的介绍，请参见“三层技术-IP 路由配置指导”中的“MTR”。
- IS-IS 支持 IPv4 单播拓扑的相关内容，请参见“三层技术-IP 路由配置指导”中的“IS-IS”。

## 1.6.1 简介

IPv6 IS-IS 和 IPv4 IS-IS 使用同样的最短路径进行路由计算，IPv4 和 IPv6 的混合拓扑被看成是一个集成的拓扑，这就要求所有 IPv4 和 IPv6 的拓扑信息必须一致。但是 IPv4 和 IPv6 协议在网络中的部署可能不一致，IPv4 和 IPv6 的拓扑信息可能不同。当一些路由器和链路不支持 IPv6 协议时，支持双协议栈的路由器因为无法感知到这些路由器和链路不支持 IPv6，仍然会把 IPv6 报文转发给它们，这就导致 IPv6 报文由于无法转发而被丢弃。

IS-IS MTR（Multi-Topology Routing，多拓扑路由）的功能之一就是实现 IS-IS 支持 IPv6 单播拓扑，即 IPv4 和 IPv6 分拓扑计算，从而解决上面的问题。

图1-1 IS-IS 支持 IPv6 单播拓扑功能示意图




如 图 1-1 所示，图中的数值表示对应链路上的开销值；Router A、Router B和Router D支持IPv4 和 IPv6 双协议栈；Router C只支持IPv4 协议，不能转发IPv6 报文。

在 Router A、Router B、Router C、Router D 上都配置 IS-IS 支持 IPv6 单播拓扑，所有的路由器对于 IPv4、IPv6 都分为两个拓扑进行计算，则 Router A 能够感知到 Router B 和 Router C 之间，Router C 和 Router D 之间的链路不支持 IPv6，即不会将到达 Router D 的 IPv6 报文转发给 Router B 而造成报文丢弃。

### 1.6.2 配置IS-IS支持IPv6 单播拓扑

#### 1. 配置准备

在配置 IS-IS 支持 IPv6 单播拓扑功能之前，需完成以下任务：配置 IS-IS IPv4 和 IPv6 基本功能，网络建立 IS-IS 邻居，有基本拓扑。

#### 2. 配置IS-IS支持IPv6 单播拓扑

表1-13 配置 IS-IS 支持 IPv6 单播拓扑

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IS-IS视图 | **isis** [ *process-id* ] [ **vpn-instance** *vpn-instance-name* ] | - |
| 配置IS-IS开销值的类型 | **cost-style** { **wide** \| **wide-compatible** \| **compatible** } | 缺省情况下，IS-IS只收发采用narrow方式的报文 |
| 进入IPv6地址族视图 | **address-family ipv6** [ **unicast** ] | - |
| 配置IS-IS IPv6单播拓扑 | **multi-topology** [ **compatible** ] | 缺省情况下，没有配置支持IPv6拓扑 |

## 1.7 IPv6 IS-IS显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 IPv6 IS-IS 的运行情况，用户可以通过查看显示信息验证配置的效果。其他相关的显示和维护请参见“三层技术-IP 路由配置指导”中的“IS-IS”。

表1-14 IPv6 IS-IS 显示和维护

| 操作 | 命令 |
| --- | --- |
| 显示IPv6 IS-IS引入路由信息 | **display isis redistribute ipv6** [ *ipv6-address mask-length* ] [ **level-1** \| **level-2** ] [ *process-id* ] |
| 显示IPv6 IS-IS路由信息 | **display isis route ipv6** [ *ipv6-address* ] [ [ **level-1** \| **level-2** ] \| **verbose** ] * [ *process-id* ] |
| 显示IPv6 IS-IS拓扑信息 | **display isis spf-tree ipv6** [ [ **level-1** \| **level-2** ] \| **verbose** ] * [ *process-id* ] |

# 1.8 IPv6 IS-IS典型配置举例

## 1.8.1 IPv6 IS-IS基本配置

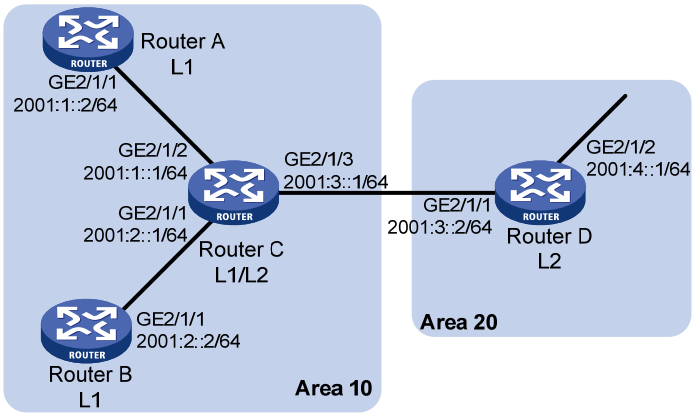
### 1. 组网需求

如下图所示，Router A、Router B、Router C 和 Router D 属于同一自治系统，所有路由器已使能了 IPv6 能力，要求它们之间通过 IPv6 IS-IS 协议达到 IPv6 网络互连的目的。

其中 Router A 和 Router B 是 Level-1 路由器，Router D 是 Level-2 路由器，Router C 是 Level-1-2 路由器。Router A、Router B 和 Router C 属于区域 10，而 Router D 属于区域 20。

### 2. 组网图

图1-2 IPv6 IS-IS 基本配置组网图




| 设备 | 接口 | IPv6地址 | 设备 | 接口 | IPv6地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 2001:1::2/64 | Router B | GE2/1/1 | 2001:2::2/64 |
| Router C | GE2/1/1 | 2001:2::1/64 | Router D | GE2/1/1 | 2001:3::2/64 |
| | GE2/1/2 | 2001:1::1/64 | | GE2/1/2 | 2001:4::1/64 |
| | GE2/1/3 | 2001:3::1/64 | | | |

### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 IPv6 IS-IS

# 配置 Router A。

```
<RouterA> system-view
[RouterA] isis 1
[RouterA-isis-1] is-level level-1
[RouterA-isis-1] network-entity 10.0000.0000.0001.00
[RouterA-isis-1] address-family ipv6
[RouterA-isis-1-ipv6] quit
[RouterA-isis-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis ipv6 enable 1
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置 Router B。

```
<RouterB> system-view
```

```
[RouterB] isis 1
[RouterB-isis-1] is-level level-1
[RouterB-isis-1] network-entity 10.0000.0000.0002.00
[RouterB-isis-1] address-family ipv6
[RouterB-isis-1-ipv6] quit
[RouterB-isis-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis ipv6 enable 1
[RouterB-GigabitEthernet2/1/1] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] isis 1
[RouterC-isis-1] network-entity 10.0000.0000.0003.00
[RouterC-isis-1] address-family ipv6
[RouterC-isis-1-ipv6] quit
[RouterC-isis-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] isis ipv6 enable 1
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] isis ipv6 enable 1
[RouterC-GigabitEthernet2/1/2] quit
[RouterC] interface gigabitethernet 2/1/3
[RouterC-GigabitEthernet2/1/3] isis ipv6 enable 1
[RouterC-GigabitEthernet2/1/3] quit
```

# 配置 Router D。

```
<RouterD> system-view
[RouterD] isis 1
[RouterD-isis-1] is-level level-2
[RouterD-isis-1] network-entity 20.0000.0000.0004.00
[RouterD-isis-1] address-family ipv6
[RouterD-isis-1-ipv6] quit
[RouterD-isis-1] quit
[RouterD] interface gigabitethernet 2/1/1
[RouterD-GigabitEthernet2/1/1] isis ipv6 enable 1
[RouterD-GigabitEthernet2/1/1] quit
[RouterD] interface gigabitethernet 2/1/2
[RouterD-GigabitEthernet2/1/2] isis ipv6 enable 1
[RouterD-GigabitEthernet2/1/2] quit
```

### 4. 验证配置

# 查看 Router A 的 IPv6 IS-IS 路由表。

```
[RouterA] display isis route ipv6

                      Route information for IS-IS(1)
                      ------------------------------

                      Level-1 IPv6 Forwarding Table
```

```
                        ----------------------------

 Destination : ::                                    PrefixLen: 0
 Flag        : R/-/-                                 Cost     : 10
 Next Hop    : FE80::200:FF:FE0F:4                   Interface: GE2/1/1

 Destination : 2001:1::                              PrefixLen: 64
 Flag        : D/L/-                                 Cost     : 10
 Next Hop    : Direct                                Interface: GE2/1/1

 Destination : 2001:2::                              PrefixLen: 64
 Flag        : R/-/-                                 Cost     : 20
 Next Hop    : FE80::200:FF:FE0F:4                   Interface: GE2/1/1

 Destination : 2001:3::                              PrefixLen: 64
 Flag        : R/-/-                                 Cost     : 20
 Next Hop    : FE80::200:FF:FE0F:4                   Interface: GE2/1/1

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
```

# 查看 Router B 的 IPv6 IS-IS 路由表。

```
[RouterB] display isis route ipv6


                        Route information for IS-IS(1)
                        ----------------------------

                        Level-1 IPv6 Forwarding Table
                        ----------------------------

 Destination : ::                                    PrefixLen: 0
 Flag        : R/-/-                                 Cost     : 10
 Next Hop    : FE80::200:FF:FE0F:4                   Interface: GE2/1/1

 Destination : 2001:1::                              PrefixLen: 64
 Flag        : D/L/-                                 Cost     : 10
 Next Hop    : FE80::200:FF:FE0F:4                   Interface: GE2/1/1

 Destination : 2001:2::                              PrefixLen: 64
 Flag        : R/-/-                                 Cost     : 20
 Next Hop    : Direct                                Interface: GE2/1/1

 Destination : 2001:3::                              PrefixLen: 64
 Flag        : R/-/-                                 Cost     : 20
 Next Hop    : FE80::200:FF:FE0F:4                   Interface: GE2/1/1

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
```

# 查看 Router C 的 IPv6 IS-IS 路由表。

```
[RouterC] display isis route ipv6
```

```
                        Route information for IS-IS(1)
                        ------------------------------

                        Level-1 IPv6 Forwarding Table
                        ------------------------------

 Destination : 2001:1::                              PrefixLen: 64
 Flag        : D/L/-                                 Cost     : 10
 Next Hop    : Direct                                Interface: GE2/1/2

 Destination : 2001:2::                              PrefixLen: 64
 Flag        : D/L/-                                 Cost     : 10
 Next Hop    : Direct                                Interface: GE2/1/1

 Destination : 2001:3::                              PrefixLen: 64
 Flag        : D/L/-                                 Cost     : 10
 Next Hop    : Direct                                Interface: GE2/1/3

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set

                        Level-2 IPv6 Forwarding Table
                        ------------------------------

 Destination : 2001:1::                              PrefixLen: 64
 Flag        : D/L/-                                 Cost     : 10
 Next Hop    : Direct                                Interface: GE2/1/2

 Destination : 2001:2::                              PrefixLen: 64
 Flag        : D/L/-                                 Cost     : 10
 Next Hop    : Direct                                Interface: GE2/1/1

 Destination : 2001:3::                              PrefixLen: 64
 Flag        : D/L/-                                 Cost     : 10
 Next Hop    : Direct                                Interface: GE2/1/3

 Destination : 2001:4::1                             PrefixLen: 128
 Flag        : R/-/-                                 Cost     : 10
 Next Hop    : FE80::20F:E2FF:FE3E:FA3D              Interface: GE2/1/3

      Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
```

# 查看 Router D 的 IPv6 IS-IS 路由表。

```
[RouterD] display isis route ipv6

                        Route information for IS-IS(1)
                        ------------------------------

                        Level-2 IPv6 Forwarding Table
```

```
                          ----------------------------
Destination : 2001:1::                          PrefixLen: 64
Flag        : R/-/-                             Cost     : 20
Next Hop    : FE80::200:FF:FE0F:4               Interface: GE2/1/1

Destination : 2001:2::                          PrefixLen: 64
Flag        : R/-/-                             Cost     : 20
Next Hop    : FE80::200:FF:FE0F:4               Interface: GE2/1/1

Destination : 2001:3::                          PrefixLen: 64
Flag        : D/L/-                             Cost     : 10
Next Hop    : Direct                            Interface: GE2/1/1

Destination : 2001:4::1                         PrefixLen: 128
Flag        : D/L/-                             Cost     : 0
Next Hop    : Direct                            Interface: GE2/1/2

     Flags: D-Direct, R-Added to Rib, L-Advertised in LSPs, U-Up/Down Bit Set
```

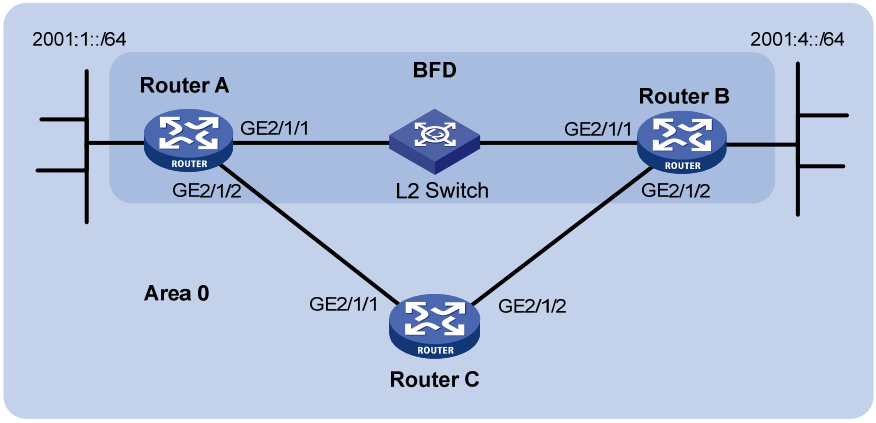
## 1.8.2 配置IPv6 IS-IS与BFD联动

### 1. 组网需求

- Router A、Router B 通过二层交换机互连，并且在双方接口上使能 BFD 应用，之间运行 IPv6 IS-IS，网络层相互可达。
- 当Router B 和二层交换机之间的链路发生故障后，BFD 能够快速检测并通告 IPv6 IS-IS 协议。

### 2. 组网图

图1-3 配置 IPv6 IS-IS 与 BFD 联动组网图




| 设备 | 接口 | IPv6地址 | 设备 | 接口 | IPv6地址 |
|---|---|---|---|---|---|
| Router A | GE2/1/1 | 2001::1/64 | Router B | GE2/1/1 | 2001::2/64 |
| | GE2/1/2 | 2001:2::1/64 | | GE2/1/2 | 2001:3::2/64 |
| Router C | GE2/1/1 | 2001:2::2/64 | | | |
| | GE2/1/2 | 2001:3::1/64 | | | |

### 3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 IPv6 IS-IS

# 配置 Router A。

```
<RouterA> system-view
[RouterA] isis 1
[RouterA-isis-1] is-level level-1
[RouterA-isis-1] network-entity 10.0000.0000.0001.00
[RouterA-isis-1] address-family ipv6
[RouterA-isis-1-ipv6] quit
[RouterA-isis-1] quit
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis ipv6 enable 1
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] isis ipv6 enable 1
[RouterA-GigabitEthernet2/1/2] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] isis 1
[RouterB-isis-1] is-level level-1
[RouterB-isis-1] network-entity 10.0000.0000.0002.00
[RouterB-isis-1] address-family ipv6
[RouterB-isis-1-ipv6] quit
[RouterB-isis-1] quit
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis ipv6 enable 1
[RouterB-GigabitEthernet2/1/1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] isis ipv6 enable 1
[RouterB-GigabitEthernet2/1/2] quit
```

# 配置 Router C。

```
<RouterC> system-view
[RouterC] isis 1
[RouterC-isis-1] network-entity 10.0000.0000.0003.00
[RouterC-isis-1] address-family ipv6
[RouterC-isis-1-ipv6] quit
[RouterC-isis-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] isis ipv6 enable 1
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] isis ipv6 enable 1
[RouterC-GigabitEthernet2/1/2] quit
```

(3) 配置 BFD 功能

# 在 Router A 上使能 IPv6 IS-IS BFD 功能，并配置 BFD 参数。

```
[RouterA] bfd session init-mode active
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] isis ipv6 bfd enable
```

```
[RouterA-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterA-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterA-GigabitEthernet2/1/1] bfd detect-multiplier 7
[RouterA-GigabitEthernet2/1/1] return
```

# 在 Router B 上使能 IPv6 IS-IS BFD 功能，并配置 BFD 参数。

```
[RouterB] bfd session init-mode active
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] isis ipv6 bfd enable
[RouterB-GigabitEthernet2/1/1] bfd min-transmit-interval 500
[RouterB-GigabitEthernet2/1/1] bfd min-receive-interval 500
[RouterB-GigabitEthernet2/1/1] bfd detect-multiplier 6
```

### 4. 验证配置

下面以 Router A 为例，Router B 和 Router A 类似，不再赘述。

# 显示 Router A 的 BFD 信息。

```
<RouterA> display bfd session

 Total Session Num: 1     Up Session Num: 1     Init Mode: Active

 IPv6 Session Working Under Ctrl Mode:

       Local Discr: 1441                 Remote Discr: 1450
        Source IP: FE80::20F:FF:FE00:1202（Router A 接口 GigabitEthernet2/1/1 的链路本地地址）
   Destination IP: FE80::20F:FF:FE00:1200（Router B 接口 GigabitEthernet2/1/1 的链路本地地址）
     Session State: Up                     Interface: GE2/1/1
        Hold Time: 2319ms
```

# 在 Router A 上查看 2001:4::0/64 的路由信息，可以看出 Router A 和 Router B 是通过 L2 Switch 进行通信的。

```
<RouterA> display ipv6 routing-table 2001:4::0 64

Summary Count : 2

Destination: 2001:4::/64                                Protocol  : ISISv6
NextHop    : FE80::20F:FF:FE00:1200                     Preference: 15
Interface  : GE2/1/1                                     Cost      : 10
```

当 Router A 和 Router B 通过 L2 Switch 通信的链路出现故障时：

# 在 Router A 上查看 2001:4::0/64 的路由信息，可以看出 Router A 和 Router B 已经切换到 Router C 进行通信。

```
<RouterA> display ipv6 routing-table 2001:4::0 64

Summary Count : 1

Destination: 2001:4::/64                                Protocol  : ISISv6
NextHop    : FE80::BAAF:67FF:FE27:DCD0                  Preference: 15
Interface  : GE2/1/2                                     Cost      : 20
```

# 目 录

# 1 IPv6 策略路由

## 1.1 IPv6策略路由简介

与单纯依照 IPv6 报文的目的地址查找路由表进行转发不同，策略路由是一种依据用户制定的策略进行路由转发的机制。策略路由可以对于满足一定条件（ACL 规则、报文长度）的报文，执行指定的操作（设置报文的下一跳、出接口、缺省下一跳和缺省出接口等）。

报文到达后，其后续的转发流程如下：

- 首先根据配置的策略路由转发。
- 若找不到匹配的节点或虽然找到了匹配的节点，但指导报文转发失败时，再根据路由表中除缺省路由之外的路由来转发报文。
- 若转发失败，则根据策略路由中配置的缺省下一跳和缺省出接口指导报文转发。
- 若转发失败，则再根据缺省路由来转发报文。

根据作用对象的不同，策略路由可分为本地策略路由和转发策略路由：

- 本地策略路由：对设备本身产生的报文（比如本地发出的 ping 报文）起作用，指导其发送。
- 转发策略路由：对接口接收的报文起作用，指导其转发。

### 1.1.1 IPv6 策略简介

IPv6 策略用来定义报文的匹配规则，以及对报文执行的操作。IPv6 策略由节点组成。

一个 IPv6 策略可以包含一个或者多个节点。节点的构成如下：

- 每个节点由节点编号来标识。节点编号越小节点的优先级越高，优先级高的节点优先被执行。
- 每个节点的具体内容由 **if-match** 子句和 **apply** 子句来指定。**if-match** 子句定义该节点的匹配规则，**apply** 子句定义该节点的动作。
- 每个节点对报文的处理方式由匹配模式决定。匹配模式分为 **permit**（允许）和 **deny**（拒绝）两种。

应用 IPv6 策略后，系统将根据 IPv6 策略中定义的匹配规则和操作，对报文进行处理：系统按照优先级从高到低的顺序依次匹配各节点，如果报文满足这个节点的匹配规则，就执行该节点的动作；如果报文不满足这个节点的匹配规则，就继续匹配下一个节点；如果报文不能满足 IPv6 策略中任何一个节点的匹配规则，则根据路由表来转发报文。

#### 1. if-match子句

目前，IPv6 策略路由提供了两种 **if-match** 子句，作用如下：

- **if-match acl**：设置 ACL 匹配规则。
- **if-match packet-length**：设置 IPv6 报文长度匹配规则。

在一个节点中可以配置多条 **if-match** 子句，同一类型的 **if-match** 子句最多只能有一条。

同一个节点中的各 **if-match** 子句之间是“与”的关系，即报文必须满足该节点的所有 **if-match** 子句才算满足这个节点的匹配规则。

### 2. apply子句

IPv6 策略路由提供了十一种**apply**子句，同一个节点中可以配置多条**apply**子句，但配置的多条**apply**子句不一定都会执行。影响报文转发路径的**apply**子句有五条，优先级顺序是：**apply access-vpn vpn-instance**、**apply next-hop**、**apply output-interface**、**apply default-next-hop**和**apply default-output-interface**。**apply**子句的含义以及执行优先情况等说明如 表 1-1 所示。

表1-1 **apply** 子句的含义以及执行优先情况等说明

| 子句 | 含义 | 执行优先情况/详细说明 |
|---|---|---|
| **apply precedence** | 设置IPv6报文的IP优先级 | 只要配置了该子句，该子句就一定会执行 |
| **apply loadshare { next-hop \| output-interface \| default-next-hop \| default-output-interface }** | 设置指导报文转发的下一跳、出接口、缺省下一跳和缺省出接口的工作模式为负载分担模式 | 下一跳、出接口、缺省下一跳和缺省出接口的工作模式有两种：主备模式、负载分担模式<br>• 主备模式：按照配置顺序，以第一个配置（下一跳、出接口、缺省下一跳或缺省出接口）作为主用，指导报文转发。当主用失效时，按配置顺序选择后续的第一个有效配置指导报文转发<br>• 负载分担模式：按照配置顺序，逐包轮流选择有效的下一跳、出接口、缺省下一跳或缺省出接口指导报文转发<br>缺省情况下，工作模式在主备模式 |
| **apply access-vpn vpn-instance** | 设置报文在指定VPN实例中进行转发 | 报文如果匹配了其中一个VPN实例下的转发表，报文将在该VPN实例中进行转发 |
| **apply next-hop**和**apply output-interface** | 设置报文的下一跳、出接口 | **apply next-hop**的优先级高于**apply output-interface**。当两条子句同时配置并且都有效时，系统只会执行**apply next-hop**子句 |
| **apply default-next-hop**和**apply default-output-interface** | 设置报文的缺省下一跳、缺省出接口 | **apply default-next-hop**的优先级高于**apply default-output-interface**。当两条子句同时配置并且都有效时，系统只会执行**apply default-next-hop**子句<br>执行缺省下一跳和出接口的前提是：在策略中没有配置下一跳或者出接口，或者配置的下一跳和出接口无效，并且在路由表中没有找到与报文目的IPv6地址匹配的路由表项 |
| **apply continue** | 设置匹配成功的当前节点转发失败后继续进行后续节点的处理 | 如果当前节点中没有配置影响报文转发路径的五个**apply**子句，或者配置了这五个子句中的一个或多个，但配置的子句都失效（下一跳不可达、出接口down或者报文在指定VPN内转发失败）时，会进行下一节点的处理 |

### 3. 节点的匹配模式与节点的**if-match**子句、**apply**子句的关系

一个节点的匹配模式与这个节点的**if-match**子句、**apply**子句的关系如 表 1-2 所示。

表1-2 节点的匹配模式、**if-match** 子句、**apply** 子句三者之间的关系

| 是否满足所有**if-match** 子句 | 节点匹配模式 | |
|---|---|---|
| | **permit**（允许模式） | **deny**（拒绝模式） |

| 是否满足所有 **if-match** 子句 | 节点匹配模式 | |
|---|---|---|
| | **permit**（允许模式） | **deny**（拒绝模式） |
| 是 | • 如果节点配置了 **apply** 子句，则执行此节点 **apply** 子句，不再匹配下一节点<br>  ○ 如果节点指导报文转发成功，则不再匹配下一节点<br>  ○ 如果节点指导报文转发失败且没有配置 **apply continue** 子句，则不再匹配下一节点<br>  ○ 如果节点指导报文转发失败且配置了 **apply continue** 子句，则继续匹配下一节点<br>• 如果节点没有配置 **apply** 子句，则不会执行任何动作，且不再匹配下一节点，报文将根据路由表来进行转发 | 不执行此节点**apply**子句，不再匹配下一节点，报文将根据路由表来进行转发 |
| 否 | 不执行此节点**apply**子句，继续匹配下一节点 | 不执行此节点**apply**子句，继续匹配下一节点 |

说明

如果一个节点中没有配置任何 **if-match** 子句，则认为所有报文都满足该节点的匹配规则，按照“报文满足所有 **if-match** 子句”的情况进行后续处理。

### 1.1.2 策略路由与Track联动

策略路由通过与 Track 联动，增强了应用的灵活性和对网络环境变化的动态感知能力。

策略路由可以在配置报文的下一跳、出接口、缺省下一跳、缺省出接口时与 Track 项关联，根据 Track 项的状态来动态地决定策略的可用性。策略路由配置仅在关联的 Track 项状态为 Positive 或 NotReady 时生效。关于策略路由与 Track 联动的的详细介绍和相关配置，请参见“可靠性配置指导”中的“Track”。

## 1.2 IPv6策略路由配置任务简介

表1-3 IPv6 策略路由配置任务简介

| 配置任务 | | 说明 | 详细配置 |
|---|---|---|---|
| 配置IPv6策略 | 创建IPv6策略节点 | 必选 | 1.3.1 |
| | 配置IPv6策略节点的匹配规则 | | 1.3.2 |
| | 配置IPv6策略节点的动作 | | 1.3.3 |
| 应用IPv6策略 | 对本地IPv6报文应用策略 | 必选<br>用户根据实际情况进行选择 | 1.4.1 |
| | 对接口转发的IPv6报文应用策略 | | 1.4.2 |

## 1.3 配置IPv6策略

### 1.3.1 创建IPv6 策略节点

表1-4 创建 IPv6 策略节点

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 创建IPv6策略节点，并进入IPv6策略节点视图 | **ipv6 policy-based-route** *policy-name* [ **deny** \| **permit** ] **node** *node-number* | 缺省情况下，没有创建IPv6策略节点 |

### 1.3.2 配置IPv6 策略节点的匹配规则

表1-5 配置 IPv6 策略节点的匹配规则

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IPv6策略节点视图 | **ipv6 policy-based-route** *policy-name* [ **deny** \| **permit** ] **node** *node-number* | - |
| 设置ACL匹配规则 | **if-match acl** { *acl6-number* \| **name** *acl6-name* } | 缺省情况下，未设置ACL匹配规则 |
| 设置IPv6报文长度匹配规则 | **if-match packet-length** *min-len max-len* | 缺省情况下，未设置IPv6报文长度匹配规则 |

说明

**if-match** 子句中使用 ACL 时，对于 ACL 规则的 **permit/deny** 动作以及 **time-range** 指定的规则生效时间段等的处理机制与设备的型号有关，请以设备的实际情况为准。

### 1.3.3 配置IPv6 策略节点的动作

表1-6 配置 IPv6 策略节点的动作

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入IPv6策略节点视图 | **ipv6 policy-based-route** *policy-name* [ **deny** \| **permit** ] **node** *node-number* | - |
| 设置IPv6报文的IP优先级 | **apply precedence** { *type* \| *value* } | 缺省情况下，不对IPv6报文的优先级进行设置 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 设置报文在指定VPN实例中进行转发 | **apply access-vpn vpn-instance** *vpn-instance-name*&<1-*n*> | 缺省情况下，未设置报文在指定VPN实例中进行转发<br>每个节点最多可以配置*m*个VPN实例。当满足匹配规则后，将根据第一个可用的VPN实例转发表进行转发。*m*的实际取值为6 |
| 设置报文转发的下一跳 | **apply next-hop** [ **vpn-instance** *vpn-instance-name* \| **inbound-vpn** ] { *ipv6-address* [ **direct** ] [ **track** *track-entry-number* ] }&<1-*n*> | 缺省情况下，未设置报文转发的下一跳<br>用户可以同时配置多个下一跳通过一次或多次配置本命令实现），起到主备或负载分担的作用<br>每个节点最多可以配置*m*个下一跳。*m*的实际取值为16 |
| 设置指导报文转发的多个下一跳工作在负载分担模式 | **apply loadshare next-hop** | 缺省情况下，多个下一跳工作在主备模式 |
| 设置指导报文转发的出接口 | **apply output-interface** { *interface-type interface-number* [ **track** *track-entry-number* ] }&<1-*n*> | 缺省情况下，未设置指导报文转发的出接口<br>用户可以同时配置多个出接口（通过一次或多次配置本命令实现），起到主备或负载分担的作用<br>每个节点最多可以配置*m*个出接口。*m*的实际取值为16 |
| 设置指导报文转发的多个出接口工作在负载分担模式 | **apply loadshare output-interface** | 缺省情况下，多个出接口工作在主备模式 |
| 设置指导报文转发的缺省下一跳 | **apply default-next-hop** [ **vpn-instance** *vpn-instance-name* \| **inbound-vpn** ] { *ipv6-address* [ **direct** ] [ **track** *track-entry-number* ] }&<1-*n*> | 缺省情况下，未设置指导报文转发的缺省下一跳<br>用户可以同时配置多个缺省下一跳（通过一次或多次配置本命令实现），起到主备或负载分担的作用<br>每个节点最多可以配置*m*个缺省下一跳。*m*的实际取值为16 |
| 设置指导报文转发的多个缺省下一跳工作在负载分担模式 | **apply loadshare default-next-hop** | 缺省情况下，多个缺省下一跳工作在主备模式 |
| 设置指导报文转发的缺省出接口 | **apply default-output-interface** { *interface-type interface-number* [ **track** *track-entry-number* ] }&<1-*n*> | 缺省情况下，未设置指导报文转发的缺省出接口<br>用户可以同时配置多个缺省出接口（通过一次或多次配置本命令实现），起到主备或负载分担的作用<br>每个节点最多可以配置*m*个缺省出接口。*m*的实际取值为16 |
| 设置指导报文转发的多个缺省出接口工作在负载分担模式 | **apply loadshare default-output-interface** | 缺省情况下，多个缺省出接口工作在主备模式 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 设置匹配成功的当前节点指定转发路径失败后继续进行后续节点的处理 | **apply continue** | 缺省情况下，匹配成功的当前节点指定转发路径失败后不再进行下一节点的匹配<br>本命令仅在策略节点的匹配模式为**permit**时生效 |

# 1.4 应用IPv6策略

## 1.4.1 对本地报文应用IPv6 策略

通过本配置，可以将已经配置的 IPv6 策略应用到本地，指导设备本身产生 IPv6 报文的发送。应用 IPv6 策略时，该 IPv6 策略必须已经存在，否则配置将失败。

对本地报文只能应用一个 IPv6 策略。应用新的 IPv6 策略前必须删除本地原来已经应用的 IPv6 策略。

若无特殊需求，建议用户不要对本地报文应用 IPv6 策略。

**表1-7 对本地报文应用 IPv6 策略**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 对本地报文应用IPv6策略 | **ipv6 local policy-based-route** *policy-name* | 缺省情况下，对本地报文没有应用IPv6策略 |

## 1.4.2 对接口转发的报文应用IPv6 策略

通过本配置，可以将已经配置的 IPv6 策略应用到接口，指导接口接收的所有 IPv6 报文的转发。应用 IPv6 策略时，该 IPv6 策略必须已经存在，否则配置将失败。

对接口转发的报文应用 IPv6 策略时，一个接口只能应用一个 IPv6 策略。应用新的 IPv6 策略前必须删除接口上原来已经应用的 IPv6 策略。

一个 IPv6 策略可以同时被多个接口应用。

**表1-8 对接口转发的报文应用 IPv6 策略**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 对接口转发的报文应用IPv6策略 | **ipv6 policy-based-route** *policy-name* | 缺省情况下，对接口转发的报文没有应用IPv6策略 |

## 1.5 IPv6策略路由显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示 IPv6 策略路由配置后的运行情况，通过查看显示信息验证配置的效果。

在用户视图下，用户可以执行 **reset** 命令可以清除 IPv6 策略路由的统计信息。

表1-9 IPv6 策略路由显示和维护

| 操作 | 命令 |
|---|---|
| 显示已经配置的IPv6策略 | **display ipv6 policy-based-route** [ **policy** *policy-name* ] |
| 显示已经应用的IPv6策略路由信息 | **display ipv6 policy-based-route setup** |
| 显示IPv6本地策略路由的配置信息和统计信息（MSR 2600/MSR 3600） | **display ipv6 policy-based-route local** |
| 显示IPv6本地策略路由的配置信息和统计信息（MSR 5600） | **display ipv6 policy-based-route local** [ **slot** *slot-number* ] |
| 显示接口下IPv6转发策略路由的配置信息和统计信息（MSR 2600/MSR 3600） | **display ipv6 policy-based-route interface** *interface-type interface-number* |
| 显示接口下IPv6转发策略路由的配置信息和统计信息（MSR 5600） | **display ipv6 policy-based-route interface** *interface-type interface-number* [ **slot** *slot-number* ] |
| 清除IPv6策略路由的统计信息 | **reset ipv6 policy-based-route statistics** [ **policy** *policy-name* ] |

## 1.6 IPv6策略路由典型配置举例

### 1.6.1 基于报文协议类型的IPv6 本地策略路由配置举例

#### 1. 组网需求

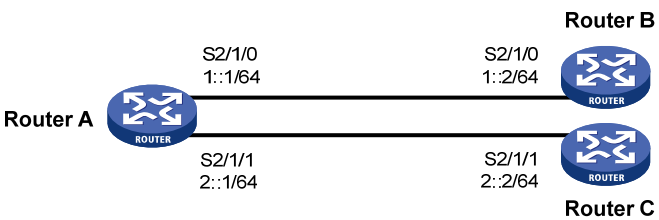
通过策略路由控制 Router A 产生的报文：

- 指定所有 TCP 报文的下一跳为 1::2；
- 其它 IPv6 报文仍然按照查找路由表的方式进行转发。

其中，Router A 分别与 Router B 和 Router C 直连。

#### 2. 组网图

图1-1 基于报文协议类型的策略路由的配置举例组网图



### 3. 配置步骤

(1) 配置 Router A

# 配置 Serial 接口的 IPv6 地址。

```
<RouterA> system-view
[RouterA] interface serial 2/1/0
[RouterA-Serial2/1/0] ipv6 address 1::1 64
[RouterA-Serial2/1/0] quit
[RouterA] interface serial 2/1/1
[RouterA-Serial2/1/1] ipv6 address 2::1 64
[RouterA-Serial2/1/1] quit
```

# 定义访问控制列表 ACL 3001，用来匹配 TCP 报文。

```
[RouterA] acl ipv6 number 3001
[RouterA-acl6-adv-3001] rule permit tcp
[RouterA-acl6-adv-3001] quit
```

# 定义 5 号节点，指定所有 TCP 报文的下一跳为 1::2。

```
[RouterA] ipv6 policy-based-route aaa permit node 5
[RouterA-pbr6-aaa-5] if-match acl 3001
[RouterA-pbr6-aaa-5] apply next-hop 1::2
[RouterA-pbr6-aaa-5] quit
```

# 在 Router A 上应用本地策略路由。

```
[RouterA] ipv6 local policy-based-route aaa
```

(2) 配置 Router B

# 配置 Serial 接口的 IPv6 地址。

```
<RouterB> system-view
[RouterB] interface serial 2/1/0
[RouterB-Serial2/1/0] ipv6 address 1::2 64
```

(3) 配置 Router C

# 配置 Serial 接口的 IPv6 地址。

```
<RouterC> system-view
[RouterC] interface serial 2/1/1
[RouterC-Serial2/1/1] ipv6 address 2::2 64
```

### 4. 验证配置

# 从 Router A 上通过 Telnet 方式登录 Router B（1::2/64），结果成功。

# 从 Router A 上通过 Telnet 方式登录 Router C（2::2/64），结果失败。

# 从 Router A 上 ping Router C（2::2/64），结果成功。

由于 Telnet 使用的是 TCP 协议，ping 使用的是 ICMP 协议，所以由以上结果可证明：Router A 产生的 TCP 报文的下一跳为 1::2，串口 Serial2/1/1 不发送 TCP 报文，但可以发送非 TCP 报文，策略路由设置成功。

## 1.6.2 基于报文协议类型的IPv6 转发策略路由配置举例

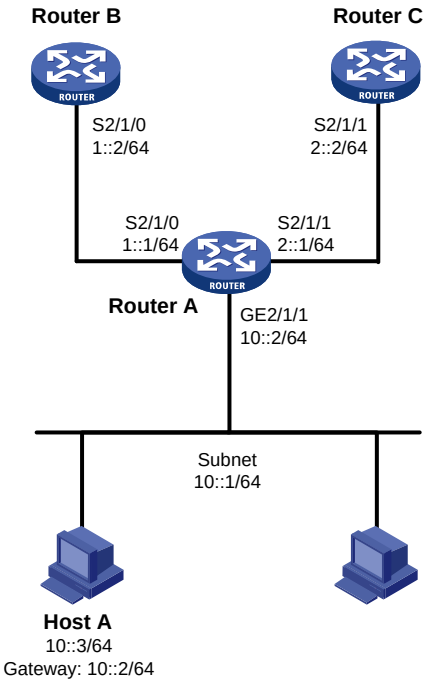
### 1. 组网需求

通过策略路由控制从 Router A 的以太网接口 GigabitEthernet2/1/1 接收的报文：

- 指定所有 TCP 报文的下一跳为 1::2；
- 其它 IPv6 报文仍然按照查找路由表的方式进行转发。

## 2. 组网图

图1-2 基于报文协议类型的 IPv6 转发策略路由配置举例组网图




## 3. 配置步骤

(1) 配置 Router A

# 配置动态路由协议 RIPng。

```
<RouterA> system-view
[RouterA] ripng 1
[RouterA-ripng-1] quit
[RouterA] interface serial 2/1/0
[RouterA-Serial2/1/0] ipv6 address 1::1 64
[RouterA-Serial2/1/0] ripng 1 enable
[RouterA-Serial2/1/0] quit
[RouterA] interface serial 2/1/1
[RouterA-Serial2/1/1] ipv6 address 2::1 64
[RouterA-Serial2/1/1] ripng 1 enable
[RouterA-Serial2/1/1] quit
```

# 定义访问控制列表 ACL 3001，用来匹配 TCP 报文。

```
[RouterA] acl ipv6 number 3001
[RouterA-acl6-adv-3001] rule permit tcp
[RouterA-acl6-adv-3001] quit
```

# 定义 5 号节点，指定所有 TCP 报文的下一跳为 1::2。

```
[RouterA] ipv6 policy-based-route aaa permit node 5
[RouterA-pbr6-aaa-5] if-match acl 3001
```

```
[RouterA-pbr6-aaa-5] apply next-hop 1::2
[RouterA-pbr6-aaa-5] quit
```

# 在以太网口 GigabitEthernet2/1/1 上应用转发策略路由，处理此接口接收的报文。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ipv6 address 10::2 64
[RouterA-GigabitEthernet2/1/1] undo ipv6 nd ra halt
[RouterA-GigabitEthernet2/1/1] ripng 1 enable
[RouterA-GigabitEthernet2/1/1] ipv6 policy-based-route aaa
[RouterA-GigabitEthernet2/1/1] quit
```

(2) 配置 Router B

# 配置动态路由协议 RIPng。

```
<RouterB> system-view
[RouterB] ripng 1
[RouterB-ripng-1] quit
[RouterB] interface serial 2/1/0
[RouterB-Serial2/1/0] ipv6 address 1::2 64
[RouterB-Serial2/1/0] ripng 1 enable
[RouterB-Serial2/1/0] quit
```

(3) 配置 Router C

# 配置动态路由协议 RIPng。

```
<RouterC> system-view
[RouterC] ripng 1
[RouterC-ripng-1] quit
[RouterC] interface serial 2/1/1
[RouterC-Serial2/1/1] ipv6 address 2::2 64
[RouterC-Serial2/1/1] ripng 1 enable
[RouterC-Serial2/1/1] quit
```

#### 4. 验证配置

在 Host A 上安装 IPv6 协议栈，并将 IPv6 地址配置为 10::3。

```
C:\>ipv6 install
Installing...
Succeeded.
C:\>ipv6 adu 4/10::3
```

从 Host A 上通过 Telnet 方式登录 Router B，结果成功。

从 Host A 上通过 Telnet 方式登录 Router C，结果失败。

从 Host A 上 ping Router C，结果成功。

由于 Telnet 使用的是 TCP 协议，ping 使用的是 ICMP 协议，所以由以上结果可证明：从 Router A 的以太网接口 GigabitEthernet2/1/1 接收的 TCP 报文的下一跳为 1::2，串口 Serial2/1/1 不转发 TCP 报文，但可以转发非 TCP 报文，策略路由设置成功。

### 1.6.3 基于报文长度的IPv6 转发策略路由配置举例

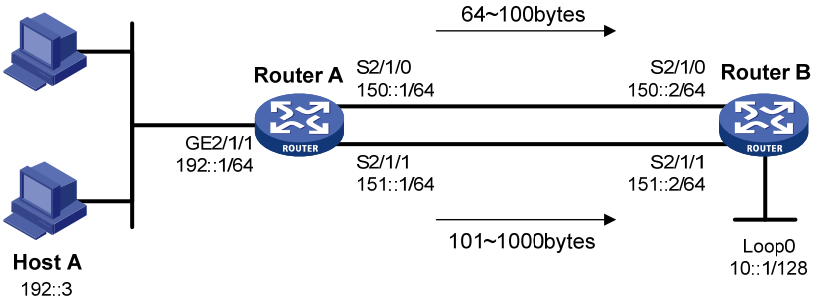
#### 1. 组网需求

通过策略路由控制从 Router A 的以太网接口 GigabitEthernet2/11 接收的报文：

- 长度为 64～100 字节的 IPv6 报文以 150::2/64 作为下一跳 IPv6 地址；
- 长度为 101～1000 字节的 IPv6 报文以 151::2/64 作为下一跳 IPv6 地址；
- 所有其它长度的 IPv6 报文都按照查找路由表的方式转发。

## 2. 组网图

图1-3 基于报文长度的 IPv6 转发策略路由配置举例组网图




## 3. 配置步骤

(1) 配置 Router A

# 配置动态路由协议 RIPng。

```
<RouterA> system-view
[RouterA] ripng 1
[RouterA-ripng-1] quit
[RouterA] interface serial 2/1/0
[RouterA-Serial2/1/0] ipv6 address 150::1 64
[RouterA-Serial2/1/0] ripng 1 enable
[RouterA-Serial2/1/0] quit
[RouterA] interface serial 2/1/1
[RouterA-Serial2/1/1] ipv6 address 151::1 64
[RouterA-Serial2/1/1] ripng 1 enable
[RouterA-Serial2/1/1] quit
```

# 配置策略 lab1，将长度为 64～100 字节的 IPv6 报文转发到下一跳 150::2/64，而将长度为 101～1000 字节的 IPv6 报文转发到下一跳 151::2/64。

```
[RouterA] ipv6 policy-based-route lab1 permit node 10
[RouterA-pbr6-lab1-10] if-match packet-length 64 100
[RouterA-pbr6-lab1-10] apply next-hop 150::2
[RouterA-pbr6-lab1-10] quit
[RouterA] ipv6 policy-based-route lab1 permit node 20
[RouterA-pbr6-lab1-20] if-match packet-length 101 1000
[RouterA-pbr6-lab1-20] apply next-hop 151::2
[RouterA-pbr6-lab1-20] quit
```

# 在以太网接口 GigabitEthernet2/1/1 上应用定义的策略 lab1，处理此接口接收的报文。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ipv6 address 192::1 64
[RouterA-GigabitEthernet2/1/1] undo ipv6 nd ra halt
[RouterA-GigabitEthernet2/1/1] ripng 1 enable
```

```
[RouterA-GigabitEthernet2/1/1] ipv6 policy-based-route lab1
[RouterA-GigabitEthernet2/1/1] return
```

(2) 配置 Router B

# 配置动态路由协议 RIPng。

```
<RouterB> system-view
[RouterB] ripng 1
[RouterB-ripng-1] quit
[RouterB] interface serial 2/1/0
[RouterB-Serial2/1/0] ipv6 address 150::2 64
[RouterB-Serial2/1/0] ripng 1 enable
[RouterB-Serial2/1/0] quit
[RouterB] interface serial 2/1/1
[RouterB-Serial2/1/1] ipv6 address 151::2 64
[RouterB-Serial2/1/1] ripng 1 enable
[RouterB-Serial2/1/1] quit
[RouterB] interface loopback 0
[RouterB-LoopBack0] ipv6 address 10::1 128
[RouterB-LoopBack0] ripng 1 enable
```

## 4. 验证配置

# 在 Router A 上用 **debugging ipv6 policy-based-route** 命令监视策略路由。

```
<RouterA> debugging ipv6 policy-based-route
<RouterA> terminal logging level 7
<RouterA> terminal monitor
```

# 在 Host A 上安装 IPv6 协议栈，并将 IPv6 地址配置为 192::3。

```
C:\>ipv6 install
Installing...
Succeeded.
C:\>ipv6 adu 4/192::3
```

# 从 Host A 上 Ping Router B 的 Loopback0，并将报文数据字段长度设为 64 字节。

```
C:\>ping –n 1 -l 64 10::1

Pinging 10::1 with 64 bytes of data:

Reply from 10::1: time=1ms

Ping statistics for 10::1:
    Packets: Sent = 1, Received = 1, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 1ms, Maximum = 1ms, Average = 1ms
```

从 Router A 上显示的策略路由调试信息如下：

```
<RouterA>
*Jun  26 13:04:33:519 2012 RouterA PBR6/7/PBR Forward Info: -MDC=1; Policy:lab1, Node:
 10,match Succeeded.
*Jun  26 13:04:33:519 2012 RouterA PBR6/7/PBR Forward Info: -MDC=1; apply next-hop 150
::2.
```

以上策略路由信息显示，Router A 在接收到报文后，根据策略路由确定的下一跳为 150::2，也就是说将报文从接口 Serial2/1/0 转发出去。

# 从 Host A 上 Ping Router B 的 Loopback0，并将报文数据字段长度设为 200 字节。

```
C:\>ping –n 1 -l 200 10::1

Pinging 10::1 with 200 bytes of data:

Reply from 10::1: time=1ms

Ping statistics for 10::1:
    Packets: Sent = 1, Received = 1, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 1ms, Maximum = 1ms, Average = 1ms
```

从 Router A 上显示的策略路由调试信息如下：

```
<RouterA>
*Jun  26 13:20:33:619 2012 RouterA PBR6/7/PBR Forward Info: -MDC=1; Policy:lab1, Node:
 20,match Succeeded.
*Jun  26 13:20:33:619 2012 RouterA PBR6/7/PBR Forward Info: -MDC=1; apply next-hop 151
::2.
```

以上策略路由信息显示，Router A 在接收到报文后，根据策略路由确定的下一跳为 151::2，也就是说将报文从接口 Serial2/1/1 转发出去。

# 目 录

# 1 路由策略

本章所介绍的路由策略包括 IPv4 路由策略和 IPv6 路由策略，二者的配置基本一致，不同的部分在各节中另行说明。

## 1.1 路由策略简介

路由策略是为了改变网络流量所经过的途径而修改路由信息的技术，主要通过改变路由属性（包括可达性）来实现。

### 1.1.1 路由策略的应用

路由策略的应用灵活广泛，主要有下面几种方式：

- 控制路由的发布

路由协议在发布路由信息时，通过路由策略对路由信息进行过滤，只发布满足条件的路由信息。

- 控制路由的接收

路由协议在接收路由信息时，通过路由策略对路由信息进行过滤，只接收满足条件的路由信息，可以控制路由表项的数量，提高网络的安全性。

- 管理引入的路由

路由协议在引入其它路由协议发现的路由时，通过路由策略只引入满足条件的路由信息，并控制所引入的路由信息的某些属性，以使其满足本协议的要求。

- 设置路由的属性

对通过路由策略的路由设置相应的属性。

### 1.1.2 路由策略的实现

路由策略的实现步骤如下：

(1) 首先要定义将要实施路由策略的路由信息的特征，即定义一组匹配规则。可以用路由信息中的不同属性作为匹配依据进行设置，如目的地址、发布路由信息的路由器地址等。

(2) 然后再将匹配规则应用于路由的发布、接收和引入等过程的路由策略中。

可以灵活使用过滤器来定义各种匹配规则，过滤器的相关内容见下一节介绍。

### 1.1.3 过滤器

过滤器可以看作是路由策略过滤路由的工具，单独配置的过滤器没有任何过滤效果，只有在路由协议的相关命令中应用这些过滤器，才能够达到预期的过滤效果。

路由协议可以引用访问控制列表、地址前缀列表、AS 路径访问列表、团体属性列表、扩展团体属性列表、路由策略几种过滤器。下面对各种过滤器逐一进行介绍。

### 1. 访问控制列表

访问控制列表包括针对IPv4报文的ACL和针对IPv6报文的ACL。用户在定义ACL时可以指定IP(v6)地址和子网范围，用于匹配路由信息的目的网段地址或下一跳地址。

ACL 的相关内容请参见“ACL 和 QoS 配置指导”中的“ACL”。

### 2. 地址前缀列表

地址前缀列表包括 IPv4 地址前缀列表和 IPv6 地址前缀列表。

地址前缀列表的作用类似于 ACL，但比它更为灵活，且更易于用户理解。使用地址前缀列表过滤路由信息时，其匹配对象为路由信息的目的地址信息域；另外，用户可以指定 **gateway** 选项，指明只接收某些路由器发布的路由信息。关于 **gateway** 选项的设置请参见“三层技术-IP 路由命令参考”中的“RIP”和“OSPF”。

一个地址前缀列表由前缀列表名标识。每个前缀列表可以包含多个表项，每个表项可以独立指定一个网络前缀形式的匹配范围，并用一个索引号来标识，索引号指明了在地址前缀列表中进行匹配检查的顺序。

每个表项之间是“或”的关系，在匹配的过程中，路由器按升序依次检查由索引号标识的各个表项，只要有某一表项满足条件，就意味着通过该地址前缀列表的过滤（不再对下一个表项进行匹配）。

### 3. AS路径访问列表（as-path）

as-path 仅用于 BGP 路由的过滤。BGP 的路由信息中，包含有自治系统路径域。as-path 就是针对自治系统路径域指定匹配条件。

as-path 的相关内容请参见“三层技术-IP 路由配置指导”中的“BGP”。

### 4. 团体属性列表（community-list）

community-list 仅用于 BGP 路由的过滤。BGP 的路由信息包中，包含一个 community 属性域，用来标识一个团体。community-list 就是针对团体属性域指定匹配条件。

团体属性列表的相关内容请参见“三层技术-IP 路由配置指导”中的“BGP”。

### 5. 扩展团体属性列表（extcommunity-list）

extcommunity-list 仅用于 BGP 路由的过滤。BGP 扩展团体属性有两种，一种是用于 VPN 的 RT（Route Target，路由目标）扩展团体，另一种则是 SoO（Site of Origin，源站点）扩展团体。扩展团体属性列表就是针对这两种属性指定匹配条件。

扩展团体属性列表的相关内容请参见“MPLS 配置指导”中的“MPLS L3VPN”。

### 6. MAC地址列表（mac-list）

mac-list 仅用于 EVI IS-IS。EVI IS-IS 的表项中包含 MAC 地址信息，mac-list 就是针对 MAC 地址信息的匹配条件。

一个 MAC 地址列表由 MAC 地址列表名标识。每个 MAC 地址列表可以包含多个表项，每个表项可以独立指定一个 MAC 地址形式的匹配范围，并用一个索引号来标识，索引号指明了在 MAC 地址列表中进行匹配检查的顺序。

每个表项之间是“或”的关系，在匹配的过程中，路由器按升序依次检查由索引号标识的各个表项，只要有某一表项满足条件，就意味着通过该地址前缀列表的过滤（不再对下一个表项进行匹配）。

#### 7. 路由策略

路由策略是一种比较复杂的过滤器，它不仅可以匹配路由信息的某些属性，还可以在条件满足时改变路由信息的属性。路由策略可以使用前面几种过滤器定义自己的匹配规则。

一个路由策略可以由多个节点构成，每个节点是匹配检查的一个单元，在匹配过程中，系统按节点序号升序依次检查各个节点。不同节点间是“或”的关系，如果通过了其中一个节点，就意味着通过该路由策略，不再对其他节点进行匹配（配置了 **continue** 子句的情况除外）。

每个节点对路由信息的处理方式由匹配模式决定。匹配模式分为 **permit** 和 **deny** 两种。

- **permit**：指定节点的匹配模式为允许模式。当路由信息通过该节点的过滤后，将执行该节点的 **apply** 子句，不进入下一个节点的匹配（配置了 **continue** 子句的情况除外）；如果路由信息没有通过该节点过滤，将进入下一个节点继续匹配。
- **deny**：指定节点的匹配模式为拒绝模式（此模式下 **apply** 子句和 **continue** 子句不会被执行）。当路由信息通过该节点的过滤后，将被拒绝通过该节点，不进入下一个节点的匹配；如果路由信息没有通过该节点的过滤，将进入下一个节点继续匹配。

每个节点可以由一组 **if-match**、**apply** 和 **continue** 子句组成。

- **if-match** 子句：定义匹配规则，匹配对象是路由信息的一些属性。同一节点中的不同 **if-match** 子句是“与”的关系，只有满足节点内所有 **if-match** 子句指定的匹配条件，才能通过该节点的匹配。
- **apply** 子句：指定动作，也就是在通过节点的匹配后，对路由信息的一些属性进行设置。
- **continue** 子句：用来配置下一个执行节点。当路由成功匹配当前路由策略节点（必须是 **permit** 节点）时，可以指定路由继续匹配同一路由策略内的下一个节点，这样可以组合路由策略各个节点的 **if-match** 子句和 **apply** 子句，增强路由策略的灵活性。

**if-match**、**apply** 和 **continue** 子句可以根据应用进行设置，都是可选的。

- 如果只过滤路由，不设置路由的属性，则不需要使用 **apply** 子句。
- 如果某个 **permit** 节点没有配置任何 **if-match** 子句，则该节点匹配所有的路由。
- 通常在多个 **deny** 节点后设置一个不含 **if-match** 子句和 **apply** 子句的 **permit** 节点，用于允许其它的路由通过。

## 1.2 配置过滤列表

### 1.2.1 配置准备

在配置过滤列表之前，需要准备以下数据：

- 前缀列表名称
- 匹配的地址范围
- 扩展团体属性列表序号

### 1.2.2 配置地址前缀列表

#### 1. 配置IPv4 地址前缀列表

提示

如果所有表项都是 **deny** 模式，则任何路由都不能通过该过滤列表。这种情况下，需要在多条 **deny** 模式的表项后定义一条 **permit** 0.0.0.0 0 **less-equal** 32 表项，允许其它所有 IPv4 路由信息通过。

IPv4 地址前缀列表由列表名标识，每个前缀列表可以包含多个表项。各表项以网络前缀的形式，独立指定一个匹配范围，并使用索引号标识。

在匹配过程中，系统按索引号升序依次检查各个表项，只要路由信息满足一个表项，就认为通过该过滤列表，不再去匹配其他表项。

**表1-1 配置 IPv4 地址前缀列表**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 配置IPv4地址前缀列表 | **ip prefix-list** *prefix-list-name* [ **index** *index-number* ] { **deny** \| **permit** } *ip-address mask-length* [ **greater-equal** *min-mask-length* ] [ **less-equal** *max-mask-length* ] | 缺省情况下，没有配置IPv4地址前缀列表 |

#### 2. 配置IPv6 地址前缀列表

提示

如果所有表项都是 **deny** 模式，则任何路由都不能通过该过滤列表。这种情况下，需要在多条 **deny** 模式的表项后定义一条 **permit** :: 0 **less-equal** 128 表项，允许其它所有 IPv6 路由信息通过。

IPv6 地址前缀列表由列表名标识，每个前缀列表可以包含多个表项。各表项可以独立指定一个网络前缀形式的匹配范围，并使用索引号标识。

在匹配的过程中，系统按索引号升序依次检查各个表项，只要路由信息满足一个表项，就认为通过该过滤列表，不再去匹配其他表项。

**表1-2 配置 IPv6 地址前缀列表**

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 配置IPv6地址前缀列表 | **ipv6 prefix-list** *prefix-list-name* [ **index** *index-number* ] { **deny** \| **permit** } *ipv6-address prefix-length* [ **greater-equal** *min-prefix-length* ] [ **less-equal** *max-prefix-length* ] | 缺省情况下，没有配置IPv6地址前缀列表 |

### 1.2.3 配置AS路径过滤列表

一个 AS 路径过滤列表可以包含多个表项。在匹配过程中，各表项之间是“或”的关系，即只要路由信息通过该列表中的一条表项，就认为通过该 AS 路径过滤列表。

表1-3 配置 AS 路径过滤列表

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置AS路径过滤列表 | **ip as-path** *as-path-number* { **deny** \| **permit** } *regular-expression* | 缺省情况下，没有配置AS路径过滤列表 |

### 1.2.4 配置团体属性列表

一个团体属性列表可以定义多个表项。在匹配过程中，各表项之间是“或”的关系，即只要路由信息通过该列表中的一条表项，就认为通过该团体属性列表。

表1-4 配置团体属性列表

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 配置团体属性列表 | 配置基本团体属性列表 | **ip community-list** { *basic-comm-list-num* \| **basic** *basic-comm-list-name* } { **deny** \| **permit** } [ *community-number*&<1-32> \| *aa:nn*&<1-32> ] [ **internet** \| **no-advertise** \| **no-export** \| **no-export-subconfed** ] * | 二者选其一<br>缺省情况下，没有配置团体属性列表 |
| | 配置高级团体属性列表 | **ip community-list** { *adv-comm-list-num* \| **advanced** *adv-comm-list-name* } { **deny** \| **permit** } *regular-expression* | |

### 1.2.5 配置扩展团体属性列表

一个扩展团体属性列表可以定义多个表项。在匹配过程中，各表项之间是“或”的关系，即只要路由信息通过该列表中的一条表项，就认为通过该扩展团体属性列表。

表1-5 配置扩展团体属性列表

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置扩展团体属性列表 | **ip extcommunity-list** *ext-comm-list-number* { **deny** \| **permit** } { **rt** *route-target* \| **soo** *site-of-origin* }&<1-32> | 缺省情况下，没有配置扩展团体属性列表 |

### 1.2.6 配置MAC地址列表

提示

如果所有表项都是 **deny** 模式，则任何路由都不能通过该 MAC 地址列表。这种情况下，需要在多条 **deny** 模式的表项后定义一条 **permit** 0-0-0 0 表项，允许其它所有表项信息通过。

MAC 地址列表由列表名标识，每个 MAC 地址列表可以包含多个表项。各表项以 MAC 地址的形式，独立指定一个匹配范围，并使用索引号标识。

在匹配过程中，系统按索引号升序依次检查各个表项，只要满足一个表项，就认为通过该 MAC 地址列表，不再去匹配其他表项。

**表1-6 配置 MAC 地址列表**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 配置MAC地址列表 | **mac-list** *mac-list-name* [ **index** *index-number* ] { **deny** \| **permit** } *mac-address* [ *mask-length* ] | 缺省情况下，没有配置MAC地址列表 |

## 1.3 配置路由策略

路由策略用来根据路由信息的某些属性过滤路由信息，并改变与路由策略规则匹配的路由信息的属性。匹配条件可以使用前面几种过滤列表。

一个路由策略可由多个节点构成，每个节点又分为：

- **if-match** 子句：定义匹配规则，即路由信息通过当前路由策略所需满足的条件，匹配对象是路由信息的某些属性。
- **apply** 子句：指定动作，也就是在满足由 **if-match** 子句指定的过滤条件后所执行的一些配置命令，对路由的某些属性进行修改。
- **continue** 子句：用来配置下一个执行节点，当路由成功匹配当前路由策略节点时，可以指定路由继续匹配同一路由策略内的下一个节点。

### 1.3.1 配置准备

在配置路由策略之前，需完成以下任务：

- 配置过滤列表
- 配置路由协议

在配置之前，需要准备以下数据：

- 路由策略的名称、节点序号
- 匹配条件
- 要修改的路由属性值

### 1.3.2 创建一个路由策略

如果路由策略中定义了一个以上的节点，则各节点中至少应该有一个节点的匹配模式是 **permit**。如果路由策略的所有节点都是 **deny** 模式，则没有路由信息能通过该路由策略。

当路由策略用于路由信息过滤时，如果某路由信息没有通过任一节点，则认为该路由信息没有通过该路由策略。

**表1-7 创建一个路由策略**

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 创建路由策略，并进入该路由策略视图 | **route-policy** *route-policy-name* { **deny** \| **permit** } **node** *node-number* | 缺省情况下，没有配置路由策略 |

### 1.3.3 配置if-match子句

在一个节点中，可以没有 **if-match** 子句，也可以有多个 **if-match** 子句。当不指定 **if-match** 子句时，如果该节点的匹配模式为允许模式，则所有路由信息都会通过该节点的过滤；如果该节点的匹配模式为拒绝模式，则所有路由信息都会被拒绝。

**表1-8 配置 if-match 子句**

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 进入系统视图 | | **system-view** | - |
| 进入路由策略视图 | | **route-policy** *route-policy-name* { **deny** \| **permit** } **node** *node-number* | - |
| 配置路由的匹配条件 | 配置IPv4的路由信息的匹配条件 | **if-match ip** { **address** \| **next-hop** \| **route-source** } { **acl** *acl-number* \| **prefix-list** *prefix-list-name* } | 缺省情况下，没有配置IPv4的路由信息的匹配条件<br>如果**if-match**子句对应的ACL不存在或者ACL中没有任何规则时，则默认满足该匹配条件；如果**if-match**子句对应的ACL中没有匹配的ACL规则或者ACL规则处于非激活状态，则默认不满足该匹配条件<br>路由策略使用非VPN的ACL进行路由过滤 |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| | 配置IPv6的路由信息的匹配条件 | **if-match ipv6** { **address** \| **next-hop** \| **route-source** } { **acl** *acl6-number* \| **prefix-list** *prefix-list-name* } | 缺省情况下，没有配置IPv6的路由信息的匹配条件<br>如果**if-match**子句对应的ACL不存在或者ACL中没有任何规则时，则默认满足该匹配条件；如果**if-match**子句对应的ACL中没有匹配的ACL规则或者ACL规则处于非激活状态，则默认不满足该匹配条件<br>路由策略使用非VPN的ACL进行路由过滤 |
| 配置BGP路由信息的AS路径域的匹配条件 | | **if-match as-path** *as-path-number*&<1-32> | 缺省情况下，没有配置BGP路由信息的AS路径域的匹配条件 |
| 匹配BGP路由信息的团体属性的匹配条件 | | **if-match community** { { *basic-community-list-number* \| **name** *comm-list-name* } [ **whole-match** ] \| *adv-community-list-number* }&<1-32> | 缺省情况下，没有配置BGP路由信息的团体属性的匹配条件 |
| 配置路由信息的路由开销的匹配条件 | | **if-match cost** *value* | 缺省情况下，没有配置路由信息的路由开销的匹配条件 |
| 配置BGP扩展团体属性的匹配条件 | | **if-match extcommunity** *ext-comm-list-number*&<1-32> | 缺省情况下，没有配置BGP路由信息的扩展团体属性的匹配条件 |
| 配置路由信息的出接口的匹配条件 | | **if-match interface** { *interface-type interface-number* }&<1-16> | 缺省情况下，没有配置路由信息的出接口的匹配条件<br>将路由策略应用到BGP时，BGP协议不支持配置路由信息的出接口的匹配条件 |
| 配置BGP路由信息的本地优先级的匹配条件 | | **if-match local-preference** *preference* | 缺省情况下，没有配置BGP路由信息的本地优先级的匹配条件 |
| 配置MAC地址列表过滤的匹配条件 | | **if-match mac-list** *mac-list-name* | 缺省情况下，没有配置MAC地址列表的匹配条件<br>MAC地址列表的匹配条件仅用于EVI IS-IS |
| 配置路由信息的MPLS标签的匹配条件 | | **if-match mpls-label** | 缺省情况下，没有配置路由信息的MPLS标签的匹配条件 |
| 配置路由信息的类型的匹配条件 | | **if-match route-type** { **external-type1** \| **external-type1or2** \| **external-type2** \| **internal** \| **is-is-level-1** \| **is-is-level-2** \| **nssa-external-type1** \| **nssa-external-type1or2** \| **nssa-external-type2** } * | 缺省情况下，没有配置路由信息的类型的匹配条件 |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 配置RIP、OSPF、IS-IS路由信息的标记域的匹配条件 | **if-match tag** *value* | 缺省情况下，没有配置RIP、OSPF、IS-IS路由信息的标记域的匹配条件 |
| 配置匹配VLAN范围的匹配条件 | **if-match vlan** *vlan-list* | 缺省情况下，没有配置VLAN范围的匹配条件<br>VLAN ID范围的匹配条件仅用于EVI IS-IS |

说明

如果有 **if-match** 子句因超过命令行最大长度而出现多条相同类型的 **if-match** 子句时，这几条子句之间是“或”的关系，即满足一个匹配条件，就认为匹配该 **if-match** 语句，例如出现多条 **if-match community** 子句时，各个子句的团体属性之间是“或”的关系，即满足其中一个团体属性，就认为匹配 **if-match community** 子句。

### 1.3.4 配置apply子句

IPv4 路由策略和 IPv6 路由策略在配置 **apply** 子句时，不同之处在于设置路由信息的下一跳地址的命令不同。

表1-9 配置 apply 子句

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入系统视图 | **system-view** | - |
| 进入路由策略视图 | **route-policy** *route-policy-name* { **deny** \| **permit** } **node** *node-number* | - |
| 配置BGP路由信息的AS_PATH属性 | **apply as-path** *as-number*&<1-32> [ **replace** ] | 缺省情况下，没有配置BGP路由信息的AS_PATH属性 |
| 删除BGP路由信息的团体属性 | **apply comm-list** { *comm-list-number* \| *comm-list-name* } **delete** | 缺省情况下，没有删除BGP路由信息的团体属性 |
| 配置BGP路由信息的团体属性 | **apply community** { **none** \| **additive** \| { *community-number*&<1-32> \| *aa:nn*&<1-32> \| **internet** \| **no-advertise** \| **no-export** \| **no-export-subconfed** } * [ **additive** ] } | 缺省情况下，没有配置BGP路由信息的团体属性 |
| 配置路由信息的路由开销 | **apply cost** [ **+** \| **-** ] *value* | 缺省情况下，没有配置路由信息的路由开销 |
| 配置路由信息的开销类型 | **apply cost-type** { **external** \| **internal** \| **type-1** \| **type-2** } | 缺省情况下，没有配置路由开销类型 |
| 配置BGP路由信息的扩展团体属性 | **apply extcommunity** { **rt** *route-target* }&<1-32> [ **additive** ] | 缺省情况下，没有配置BGP路由信息的扩展团体属性 |

| 操作 | | 命令 | 说明 |
|---|---|---|---|
| 配置路由信息的下一跳地址 | 配置IPv4路由信息的下一跳地址 | **apply ip-address next-hop** *ip-address* [ **public** \| **vpn-instance** *vpn-instance-name* ] | 缺省情况下，没有配置IPv4路由信息的下一跳地址<br>对于引入的IPv4路由，使用本命令设置下一跳地址无效 |
| | 配置IPv6路由信息的下一跳地址 | **apply ipv6 next-hop** *ipv6-address* | 缺省情况下，没有配置IPv6路由信息的下一跳地址<br>对于引入的IPv6路由，使用本命令设置下一跳地址无效 |
| 配置路由的IP优先级 | | **apply ip-precedence** { *value* \| **clear** } | 缺省情况下，没有配置路由的IP优先级 |
| 配置引入路由到IS-IS某个级别的区域 | | **apply isis** { **level-1** \| **level-1-2** \| **level-2** } | 缺省情况下，没有配置引入路由到IS-IS某个级别的区域 |
| 配置BGP路由信息的本地优先级 | | **apply local-preference** *preference* | 缺省情况下，没有配置BGP路由信息的本地优先级 |
| 配置MPLS标签 | | **apply mpls-label** | 缺省情况下，没有配置MPLS标签 |
| 配置BGP路由信息的ORIGIN属性 | | **apply origin** { **egp** *as-number* \| **igp** \| **incomplete** } | 缺省情况下，没有配置BGP路由信息的ORIGIN属性 |
| 配置路由协议的优先级 | | **apply preference** *preference* | 缺省情况下，没有配置路由协议的优先级 |
| 配置BGP路由信息的首选值 | | **apply preferred-value** *preferred-value* | 缺省情况下，没有配置BGP路由信息的首选值 |
| 配置路由收敛优先级 | | **apply prefix-priority** { **critical** \| **high** \| **medium** } | 缺省情况下，没有配置路由收敛优先级<br>未配置时，路由的收敛优先级为低（Low） |
| 配置路由的QoS本地ID值 | | **apply qos-local-id** { *value* \| **clear** } | 缺省情况下，没有配置路由的QoS本地ID值 |
| 配置RIP、OSPF、IS-IS路由信息的标记域 | | **apply tag** *value* | 缺省情况下，没有配置RIP、OSPF、IS-IS路由信息的标记域 |
| 配置快速重路由 | | **apply fast-reroute** { **backup-interface** *interface-type interface-number* [ **backup-nexthop** *ip-address* ] \| **backup-nexthop** *ip-address* } | 缺省情况下，没有配置快速重路由 |
| | | **apply ipv6 fast-reroute backup-nexthop** *ipv6-address* | |

## 1.3.5 配置continue子句

表1-10 配置 continue 子句

| 操作 | 命令 | 说明 |
|---|---|---|
| **进入系**统视图 | **system-view** | - |

| 操作 | 命令 | 说明 |
| --- | --- | --- |
| 进入路由策略视图 | **route-policy** *route-policy-name* { **deny** \| **permit** } **node** *node-number* | - |
| 配置下一个执行节点 | **continue** [ *node-number* ] | 缺省情况下，没有配置下一个执行节点<br>需要注意的是，下一个执行节点序列号必须大于当前节点序列号 |

说明

- 当配置 **continue** 子句的多个节点配置相同的 **apply** 子句（没有叠加属性）只是子句的值不相同时，以最后一个 **apply** 子句为准；如果配置的是有叠加属性的 **apply** 子句（命令 **apply as-path** 不指定参数 **replace**/命令 **apply cost** 指定参数**+**或**-**/命令 **apply community** 指定参数 **additive**/命令 **apply extcommunity** 指定参数 **additive**），属性会全部叠加到路由上。
- 当配置**continue**子句的多个节点配置**apply community**子句时，使用命令行**apply comm-list delete** 不能删除前面节点中配置的团体属性。

## 1.4 路由策略显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后路由策略的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除路由策略的统计信息。

**表1-11 路由策略显示和维护**

| 操作 | 命令 |
| --- | --- |
| 显示BGP AS路径过滤列表信息 | **display ip as-path** [ *as-path-number* ] |
| 显示BGP团体属性列表信息 | **display ip community-list** [ *basic-community-list-number* \| *adv-community-list-number* \| **name** *comm-list-name* ] |
| 显示BGP扩展团体属性列表信息 | **display ip extcommunity-list** [ *ext-comm-list-number* ] |
| 显示IPv4地址前缀列表的统计信息 | **display ip prefix-list** [ **name** *prefix-list-name* ] |
| 显示IPv6地址前缀列表的统计信息 | **display ipv6 prefix-list** [ **name** *prefix-list-name* ] |
| 显示MAC地址列表的统计信息 | **display mac-list** [ **name** *mac-list-name* ] |
| 显示路由策略信息 | **display route-policy** [ **name** *route-policy-name* ] |
| 清除IPv4地址前缀列表的统计信息 | **reset ip prefix-list** [ *prefix-list-name* ] |
| 清除IPv6地址前缀列表的统计信息 | **reset ipv6 prefix-list** [ *prefix-list-name* ] |
| 清除MAC地址列表统计信息 | **reset mac-list** [ *mac-list-name* ] |

# 1.5 路由策略典型配置举例

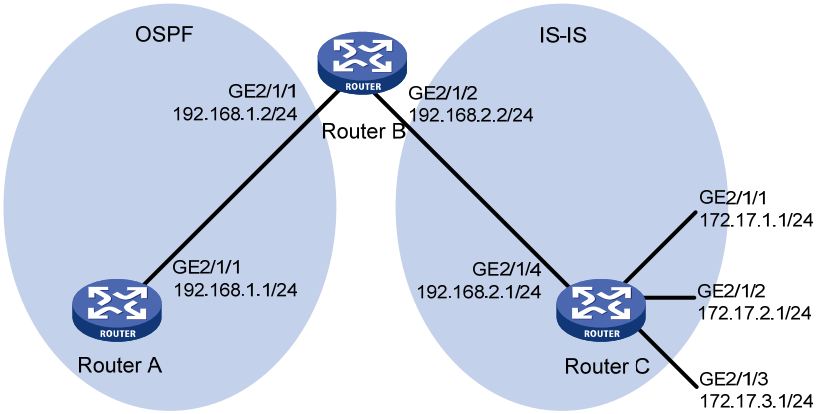
## 1.5.1 在IPv4 路由引入中应用路由策略

### 1. 组网需求

- Router B 与 Router A 之间通过 OSPF 协议交换路由信息，与 Router C 之间通过 IS-IS 协议交换路由信息。
- 要求在 Router B 上配置路由引入，将 IS-IS 路由引入到 OSPF 中去，并同时使用路由策略设置路由的属性。其中，设置 172.17.1.0/24 的路由的开销为 100，设置 172.17.2.0/24 的路由的 Tag 属性为 20。

### 2. 组网图

图1-1 在 IPv4 路由引入中应用路由策略配置组网图




### 3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 IS-IS 路由协议

# 配置 Router C。

```
<RouterC> system-view
[RouterC] isis
[RouterC-isis-1] is-level level-2
[RouterC-isis-1] network-entity 10.0000.0000.0001.00
[RouterC-isis-1] quit
[RouterC] interface gigabitethernet 2/1/1
[RouterC-GigabitEthernet2/1/1] isis enable
[RouterC-GigabitEthernet2/1/1] quit
[RouterC] interface gigabitethernet 2/1/2
[RouterC-GigabitEthernet2/1/2] isis enable
[RouterC-GigabitEthernet2/1/2] quit
[RouterC] interface gigabitethernet 2/1/3
[RouterC-GigabitEthernet2/1/3] isis enable
[RouterC-GigabitEthernet2/1/3] quit
[RouterC] interface gigabitethernet 2/1/4
[RouterC-GigabitEthernet2/1/4] isis enable
```

```
[RouterC-GigabitEthernet2/1/4] quit
```

# 配置 Router B。

```
<RouterB> system-view
[RouterB] isis
[RouterB-isis-1] is-level level-2
[RouterB-isis-1] network-entity 10.0000.0000.0002.00
[RouterB-isis-1] quit
[RouterB] interface gigabitethernet 2/1/2
[RouterB-GigabitEthernet2/1/2] isis enable
[RouterB-GigabitEthernet2/1/2] quit
```

(3) 配置 OSPF 路由协议及路由引入

# 配置 Router A，启动 OSPF。

```
<RouterA> system-view
[RouterA] ospf
[RouterA-ospf-1] area 0
[RouterA-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[RouterA-ospf-1-area-0.0.0.0] quit
[RouterA-ospf-1] quit
```

# 配置 RouterB，启动 OSPF，并引入 IS-IS 路由。

```
[RouterB] ospf
[RouterB-ospf-1] area 0
[RouterB-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[RouterB-ospf-1-area-0.0.0.0] quit
[RouterB-ospf-1] import-route isis 1
[RouterB-ospf-1] quit
```

# 查看 Router A 的 OSPF 路由表，可以看到引入的路由。

```
[RouterA] display ospf routing

          OSPF Process 1 with Router ID 192.168.1.1
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop        AdvRouter     Area
 192.168.1.0/24     1        Transit 192.168.1.1    192.168.1.1   0.0.0.0

 Routing for ASEs
 Destination        Cost     Type    Tag        NextHop        AdvRouter
 172.17.1.0/24      1        Type2   1          192.168.1.2    192.168.2.2
 172.17.2.0/24      1        Type2   1          192.168.1.2    192.168.2.2
 172.17.3.0/24      1        Type2   1          192.168.1.2    192.168.2.2

 Total Nets: 4
 Intra Area: 1  Inter Area: 0  ASE: 3  NSSA: 0
```

(4) 配置过滤列表

# 配置编号为 2002 的 ACL，允许 172.17.2.0/24 的路由通过。

```
[RouterB] acl number 2002
[RouterB-acl-basic-2002] rule permit source 172.17.2.0 0.0.0.255
```

```
[RouterB-acl-basic-2002] quit
```

# 配置名为 prefix-a 的地址前缀列表，允许 172.17.1.0/24 的路由通过。

```
[RouterB] ip prefix-list prefix-a index 10 permit 172.17.1.0 24
```

(5) 配置路由策略

```
[RouterB] route-policy isis2ospf permit node 10
[RouterB-route-policy-isis2ospf-10] if-match ip address prefix-list prefix-a
[RouterB-route-policy-isis2ospf-10] apply cost 100
[RouterB-route-policy-isis2ospf-10] quit
[RouterB] route-policy isis2ospf permit node 20
[RouterB-route-policy-isis2ospf-20] if-match ip address acl 2002
[RouterB-route-policy-isis2ospf-20] apply tag 20
[RouterB-route-policy-isis2ospf-20] quit
[RouterB] route-policy isis2ospf permit node 30
[RouterB-route-policy-isis2ospf-30] quit
```

(6) 在路由引入时应用路由策略

# 配置 Router B，设置在路由引入时应用路由策略。

```
[RouterB] ospf
[RouterB-ospf-1] import-route isis 1 route-policy isis2ospf
[RouterB-ospf-1] quit
```

# 查看 Router A 的 OSPF 路由表，可以看到目的地址为 172.17.1.0/24 的路由的开销为 100，目的地址为 172.17.2.0/24 的路由的标记域（Tag）为 20，而其他外部路由没有变化。

```
[RouterA] display ospf routing

          OSPF Process 1 with Router ID 192.168.1.1
                  Routing Tables

 Routing for Network
 Destination        Cost     Type    NextHop         AdvRouter     Area
 192.168.1.0/24     1        Transit 192.168.1.1     192.168.1.1   0.0.0.0

 Routing for ASEs
 Destination        Cost     Type    Tag         NextHop       AdvRouter
 172.17.1.0/24      100      Type2   1           192.168.1.2   192.168.2.2
 172.17.2.0/24      1        Type2   20          192.168.1.2   192.168.2.2
 172.17.3.0/24      1        Type2   1           192.168.1.2   192.168.2.2

 Total Nets: 4
 Intra Area: 1  Inter Area: 0  ASE: 3  NSSA: 0
```

## 1.5.2 在IPv6 路由引入中应用路由策略

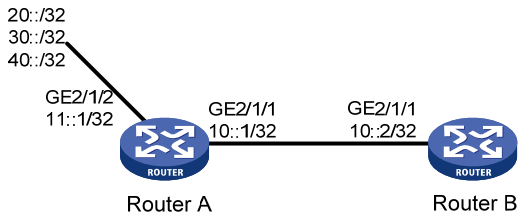
### 1. 组网需求

- Router A 与 Router B 通信，都运行 RIPng 协议。
- 使能 Router A 上的 RIPng 协议，配置三条静态路由。
- 设置在引入静态路由时应用路由策略，使三条静态路由部分引入、部分被屏蔽掉——20::/32 和 40::/32 网段的路由是可见的，30::/32 网段的路由则被屏蔽。

- 通过在 Router B 上查看 RIPng 路由表，验证路由策略是否生效。

## 2. 组网图

图1-2 在 IPv6 路由引入中应用路由策略配置组网图




## 3. 配置步骤

(1) 配置 Router A

# 配置接口 GigabitEthernet2/1/1 和 GigabitEthernet2/1/2 的 IPv6 地址。

```
<RouterA> system-view
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ipv6 address 10::1 32
[RouterA-GigabitEthernet2/1/1] quit
[RouterA] interface gigabitethernet 2/1/2
[RouterA-GigabitEthernet2/1/2] ipv6 address 11::1 32
[RouterA-GigabitEthernet2/1/2] quit
```

# 在接口 GigabitEthernet2/1/1 下使能 RIPng。

```
[RouterA] interface gigabitethernet 2/1/1
[RouterA-GigabitEthernet2/1/1] ripng 1 enable
[RouterA-GigabitEthernet2/1/1] quit
```

# 配置三条静态路由，其下一跳为 11::2，保证静态路由为 active 状态。

```
[RouterA] ipv6 route-static 20:: 32 11::2
[RouterA] ipv6 route-static 30:: 32 11::2
[RouterA] ipv6 route-static 40:: 32 11::2
```

# 配置路由策略。

```
[RouterA] ipv6 prefix-list a index 10 permit 30:: 32
[RouterA] route-policy static2ripng deny node 0
[RouterA-route-policy-static2ripng-0] if-match ipv6 address prefix-list a
[RouterA-route-policy-static2ripng-0] quit
[RouterA] route-policy static2ripng permit node 10
[RouterA-route-policy-static2ripng-10] quit
```

# 启动 RIPng 协议，同时应用路由策略 static2ripng 对引入的静态路由进行过滤。

```
[RouterA] ripng
[RouterA-ripng-1] import-route static route-policy static2ripng
```

(2) 配置 Router B

# 配置接口 GigabitEthernet2/1/1 的 IPv6 地址。

```
<RouterB> system-view
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ipv6 address 10::2 32
```

# 启动 RIPng 协议。

```
[RouterB] ripng
[RouterB-ripng-1] quit
```

# 在接口下使能 RIPng。

```
[RouterB] interface gigabitethernet 2/1/1
[RouterB-GigabitEthernet2/1/1] ripng 1 enable
[RouterB-GigabitEthernet2/1/1] quit
```

## 4. 验证配置

# 查看 Router B 的 RIPng 路由表。

```
[RouterB] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
 ----------------------------------------------------------------

 Peer FE80::7D58:0:CA03:1 on GigabitEthernet2/1/1
 Destination 10::/32,
     via FE80::7D58:0:CA03:1, cost 1, tag 0, A, 18 secs
 Destination 20::/32,
     via FE80::7D58:0:CA03:1, cost 1, tag 0, A, 8 secs
 Destination 40::/32,
     via FE80::7D58:0:CA03:1, cost 1, tag 0, A, 3 secs
```

# 目 录

# 1 MTR

## 1.1 MTR概述

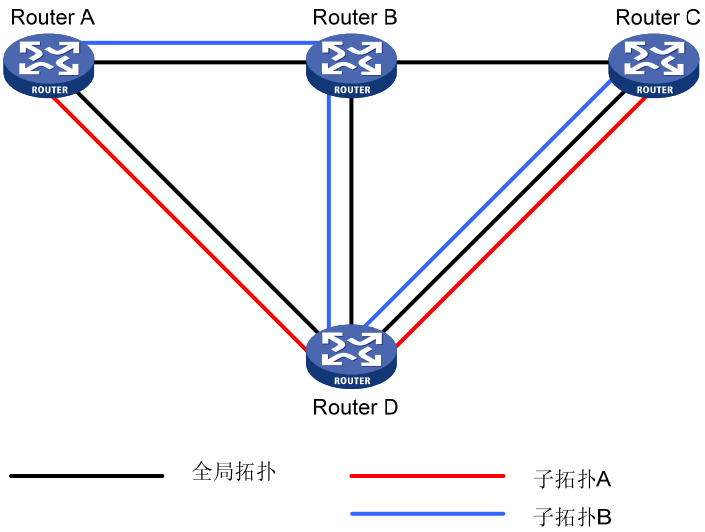
### 1.1.1 MTR简介

MTR（Multi-Topology Routing，多拓扑路由）是指将一个物理拓扑划分成多个逻辑的拓扑，这些逻辑的拓扑可能是交叉或者重叠的。不同拓扑运行各自的路由计算，实现网络的互通。

例如，IS-IS MTR 就是指在一个 IS-IS 自治域内运行多个独立的 IP 拓扑，例如 IPv4 拓扑和 IPv6 拓扑，而不是将它们视为一个集成的单一拓扑。这有利于 IS-IS 在路由计算中根据实际组网情况来单独考虑 IPv4 和 IPv6 网络。

### 1.1.2 MTR工作机制

图1-1 MTR 工作机制




如 图 1-1 所示，可以根据需要对全局拓扑进行划分，分为多个子拓扑，这样不同的流量就可以走不同的拓扑。例如，语音流可以走子拓扑A，视频流可以走子拓扑B。

对于子拓扑 A 而言，Router B 并不存在；而对于子拓扑 B 而言，它认为 Router A 和 Router D 没有直接相连，Router B 和 Router C 也没有直接相连。每一个单独的拓扑都根据路由协议计算出自己的路由，属于本拓扑的流量则根据本拓扑的路由表进行转发。

### 1.1.3 MTR支持的应用

- IS-IS MTR：详细情况请参见“三层技术-IP 路由配置指导”中的“IS-IS”和“IPv6 IS-IS”。
- 静态路由支持 MTR：详细情况请参见“三层技术-IP 路由配置指导”中的“静态路由”。

## 1.2 配置MTR

MTR 通过转发策略来对不同拓扑进行分类。转发策略使用 ACL、DSCP 优先级、IP 优先级定义自己的匹配规则。

一个转发策略可以由多个节点构成，每个节点是匹配检查的一个单元，在匹配过程中，系统按节点序号升序依次检查各个节点。不同节点间是“或”的关系，如果通过了其中一个节点，就意味着通过该转发策略，不再对其他节点进行匹配。

每个节点可以由一组 **if-match**、**apply** 子句组成。

- **if-match** 子句：定义匹配规则，匹配对象是报文信息的一些属性。同一节点中的不同 **if-match** 子句是“或”的关系，只要满足节点内任一 **if-match** 子句指定的匹配条件，就能通过该节点的匹配。
- **apply** 子句：指定动作，指定多拓扑转发策略节点应用的拓扑。

提示

节点中必须包含 **if-match** 和 **apply** 子句，否则子拓扑无法生效。

表1-1 配置 MTR

| 操作 | 命令 | 说明 |
|---|---|---|
| 进入系统视图 | **system-view** | - |
| 创建全局地址族视图 | **global-address-family ipv4** [ **unicast** ] | 缺省情况下，没有配置全局地址族视图 |
| 创建拓扑，并进入拓扑视图 | **topology** *topo-name* | 缺省情况下，不存在任何拓扑 |
| （可选）配置拓扑支持的最大激活路由前缀数 | **routing-table limit number** { *warn-threshold* \| **simply-alert** } | （*number* 的缺省情况如下）<br>MSR 2600：50000<br>MSR 36-10、MSR 3600-28/ MSR 3600-51：50000<br>其他MSR 3600：100000<br>MSR 5600：100000 |
| 退回系统视图 | **quit** | - |
| 进入接口视图 | **interface** *interface-type interface-number* | - |
| 创建并进入接口IPv4单播视图，将接口与指定拓扑进行关联 | **topology ipv4** [ **unicast** ] *topo-name* | 缺省情况下，接口没有关联到任何拓扑 |
| 退回系统视图 | **quit** | - |
| 创建多拓扑策略节点，并进入多拓扑策略节点视图 | **mtr-policy** *policy-name* **node** *node-value* | 缺省情况下，不存在多拓扑策略节点 |
| 配置做为该多拓扑转发策略节点应用的拓扑 | **apply topology** *topo-name* | 缺省情况下，没有配置多拓扑转发策略节点应用的拓扑 |

| 操作 | | 命令 | 说明 |
| --- | --- | --- | --- |
| 配置匹配条件 | 配置做为该多拓扑转发策略节点匹配条件的ACL | **if-match ip acl** *acl-number* | 三者至少选其一<br>缺省情况下，没有配置匹配条件 |
| | 配置DSCP的匹配条件 | **if-match ip dscp** *dscp-value* | |
| | 配置IP优先级的匹配条件 | **if-match ip precedence** *ip-prec-value* | |
| 退回系统视图 | | **quit** | - |
| 进入全局地址族视图 | | **global-address-family ipv4** [ **unicast** ] | - |
| 使能多拓扑转发策略 | | **topology-routing mtr-policy** *policy-name* | 缺省情况下，多拓扑转发策略处于关闭状态 |

## 1.3 MTR显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置后 MTR 的运行情况，用户可以通过查看显示信息验证配置的效果。

表1-2 MTR 显示和维护

| 操作 | 命令 |
| --- | --- |
| 显示拓扑信息 | **display topology** [ **name** *topo-name* ] |
| 显示多拓扑转发策略信息 | **display mtr-policy** [ **name** *mtr-policy-name* ] |