H3C S5500-HI 系列以太网交换机

三层技术-IP 路由配置指导

杭州华三通信技术有限公司
http://www.h3c.com.cn

资料版本：6W102-20131220
产品版本：Release 52xx 系列

Copyright © 2013 杭州华三通信技术有限公司及其许可者 版权所有，保留一切权利。
未经本公司书面许可，任何单位和个人不得擅自摘抄、复制本书内容的部分或全部，并不得以任何形式传播。

H3C、H3C、H3CS、H3CIE、H3CNE、Aolynk、Aolynk、H^{3}Care、H3care、IRF、NetPilot、Netflow、SecEngine、SecPath、SecCenter、SecBlade、Comware、ITCMM、HUASAN、华三均为杭州华三通信技术有限公司的商标。对于本手册中出现的其它公司的商标、产品标识及商品名称，由各自权利人拥有。

由于产品版本升级或其他原因，本手册内容有可能变更。H3C 保留在没有任何通知或者提示的情况下对本手册的内容进行修改的权利。本手册仅作为使用指导，H3C 尽全力在本手册中提供准确的信息，但是 H3C 并不确保手册内容完全没有错误，本手册中的所有陈述、信息和建议也不构成任何明示或暗示的担保。

前 言

H3C S5500-HI 系列以太网交换机配置指导共分为十一本手册，介绍了 S5500-HI 系列以太网交换机 Release52xx 系列软件版本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：

- 读者对象
- 新增及修改特性说明
- 本书约定
- 产品配套资料
- 资料获取方式
- 技术支持
- 资料意见反馈

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

新增及修改特性说明

本手册对应 S5500-HI 系列以太网交换机的 Release52xx 系列软件版本，各版本间特性差异如下：

- Release5206 与 Release5203 版本相比，无新增、修改特性。
- Release5203 与Release5101 版本相比，新增、修改了部分特性，具体请参见 表 1。

表1 Release5203 与 Release5101 版本间特性差异

配置指导	新增及修改特性
IP路由基础	无
静态路由	无
RIP	新增特性：配置RIP报文的DSCP优先级
OSPF	新增特性： • 配置 OSPF 报文的 DSCP 优先级 • 配置 ABR 向 Stub 区域内发布缺省路由的 Type-3 LSA 时不检查骨干区域是否存在 FULL 状态的邻居
IS-IS	无

配置指导	新增及修改特性
BGP	新增特性：配置BGP报文的DSCP优先级
IPv6静态路由	无
RIPng	无
OSPFv3	无
IPv6 IS-IS	新增特性：配置IS-IS支持IPv6单播拓扑
IPv6 BGP	新增特性：配置IPv6 BGP报文的DSCP优先级
路由策略	无
策略路由	无

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. 图形界面格式约定

格　式	意　义
< >	带尖括号“< >”表示按钮名，如“单击<确定>按钮”。
[ ]	带方括号“[ ]”表示窗口名、菜单名和数据表，如“弹出[新建用户]窗口”。
/	多级菜单用“/”隔开。如[文件/新建/文件夹]多级菜单表示[文件]菜单下的[新建]子菜单下的[文件夹]菜单项。

3. 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。

4. 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。

5. 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 S5500-HI 系列以太网交换机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	产品彩页	帮助您了解产品的主要规格参数及亮点
	技术白皮书	帮助您了解产品和特性功能，对于特色及复杂技术从细节上进行介绍
硬件介绍及安装	安全兼容性手册	列出产品的兼容性声明，并对兼容性和安全的细节进行说明
	H3C 设备防雷安装指导手册	帮助您了解防雷接地设计和工程安装方法，以保证交换机具有良好的抗雷击性能
	快速安装指南	指导您对设备进行初始安装，通常针对最常用的情况，减少您的检索时间
	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
	风扇安装手册	帮助您了解产品支持的可插拔风扇模块的外观、功能、规格、安装及拆卸方法

大类	资料名称	内容介绍
	电源手册	帮助您了解产品支持的可插拔电源模块的外观、功能、规格、安装及拆卸方法
	RPS电源用户手册	帮助您了解产品支持的RPS电源的外观、功能、规格
	H3C低端系列以太网交换机RPS电源选购指南	帮助您了解各种RPS电源适用的交换机产品型号及RPS电源配套电缆的相关规格
	接口模块扩展卡用户手册	帮助您了解该接口模块扩展卡的外观、规格、安装及拆卸方法
	H3C低端系列以太网交换机可插拔模块手册	帮助您了解产品支持的可插拔模块类型、外观和规格
	H3C可插拔SFP[SFP+][XFP]模块安装指南	帮助您掌握SFP/SFP+/XFP模块的正确安装方法，避免因操作不当而造成器件损坏
业务配置	配置指导	帮助您掌握设备软件功能的配置方法及配置步骤
	命令参考	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	故障处理手册	指导您快速定位并处理软件故障
	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）

010-62982107

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路由基础配置 ····· 1-1
1.1 IP路由简介 ····· 1-1
1.1.1 路由表 ····· 1-1
1.1.2 路由协议分类 ····· 1-2
1.1.3 路由优先级 ····· 1-3
1.1.4 负载分担 ····· 1-4
1.1.5 路由备份 ····· 1-4
1.1.6 路由迭代 ····· 1-4
1.1.7 路由信息共享 ····· 1-4
1.2 路由表显示和维护 ····· 1-4

1 IP路由基础配置

说明

- 本文所指的路由器代表运行了路由协议的三层设备。
- 除配置举例外，本文组网图中涉及的接口类型是为了方便描述而进行的示意，并不代表实际设备具有该类型的接口。
- 路由功能中所指的“接口”为三层口，包括 VLAN 接口、三层以太网端口等。三层以太网端口是指被配置为三层模式的以太网端口，有关以太网端口模式切换的操作，请参见“二层技术-以太网交换配置指导”中的“以太网端口配置”。

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。路径中最后的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

路由器通过路由表选择路由，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。每个路由器中都至少保存着一张路由表和一张 FIB 表。

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：动态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。

说明

FIB 表的具体内容，请参见“三层技术-IP 业务配置指导”中的“IP 性能优化”。

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
Routing Tables: Public
         Destinations : 7        Routes : 7

Destination/Mask    Proto   Pre  Cost        NextHop         Interface

1.1.1.0/24          Direct  0    0           1.1.1.1         Vlan11
2.2.2.0/24          Static  60   0           12.2.2.2        Vlan12
80.1.1.0/24         OSPF    10   2           80.1.1.1        Vlan13
......（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。
- 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

- 内部网关协议（Interior Gateway Protocol，简称 IGP）：在一个自治系统内部运行，常见的 IGP 协议包括 RIP、OSPF 和 IS-IS。
- 外部网关协议（Exterior Gateway Protocol，简称 EGP）：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）是拥有同一选路策略，并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。

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”。

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当前路由。

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
OSPF	10
IS-IS	15
STATIC（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，即等价路由。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。

说明

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，该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（仅配置了下一跳、未配置出接口的情况）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。

1.1.7 路由信息共享

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。

1.2 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1-2 路由表显示和维护

操作	命令
显示路由表的信息	**display ip routing-table** [**vpn-instance** *vpn-instance-name*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示通过指定基本访问控制列表过滤的路由	**display ip routing-table** [**vpn-instance** *vpn-instance-name*] **acl** *acl-number* [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定目的地址的路由	**display ip routing-table** [**vpn-instance** *vpn-instance-name*] *ip-address* [*mask* | *mask-length*] [**longer-match**] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定目的地址范围内的路由	**display ip routing-table** [**vpn-instance** *vpn-instance-name*] *ip-address1* { *mask* | *mask-length* } *ip-address2* { *mask* | *mask-length* } [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示通过指定前缀列表过滤的路由	**display ip routing-table** [**vpn-instance** *vpn-instance-name*] **ip-prefix** *ip-prefix-name* [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定协议发现的路由	**display ip routing-table** [**vpn-instance** *vpn-instance-name*] **protocol** *protocol* [**inactive** | **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示路由表中的路由统计信息	**display ip routing-table** [**vpn-instance** *vpn-instance-name*] **statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除路由表中的综合路由统计信息	**reset ip routing-table statistics protocol** [**vpn-instance** *vpn-instance-name*] { *protocol* | **all** }
显示IPv6路由表的信息	**display ipv6 routing-table** [**vpn-instance** *vpn-instance-name*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示经过指定的基本IPv6 ACL（访问控制列表）过滤的路由	**display ipv6 routing-table** [**vpn-instance** *vpn-instance-name*] **acl** *acl6-number* [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定目的地址的IPv6路由信息	**display ipv6 routing-table** [**vpn-instance** *vpn-instance-name*] *ipv6-address prefix-length* [**longer-match**] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定地址范围内的IPv6路由信息	**display ipv6 routing-table** [**vpn-instance** *vpn-instance-name*] *ipv6-address1 prefix-length1 ipv6-address2 prefix-length2* [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示经过指定IPv6前缀列表过滤的路由信息	**display ipv6 routing-table** [**vpn-instance** *vpn-instance-name*] **ipv6-prefix** *ipv6-prefix-name* [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定协议发现的IPv6路由信息	**display ipv6 routing-table** [**vpn-instance** *vpn-instance-name*] **protocol** *protocol* [**inactive** | **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6路由表中的综合路由统计信息	**display ipv6 routing-table** [**vpn-instance** *vpn-instance-name*] **statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IPv6路由表中的综合路由统计信息	**reset ipv6 routing-table statistics protocol** [**vpn-instance** *vpn-instance-name*] { *protocol* | **all** }

目 录

1 静态路由配置 ····· 1-1
1.1 简介 ····· 1-1
1.1.1 静态路由 ····· 1-1
1.1.2 缺省路由 ····· 1-1
1.1.3 静态路由应用 ····· 1-1
1.2 配置静态路由 ····· 1-2
1.2.1 配置准备 ····· 1-2
1.2.2 配置静态路由 ····· 1-2
1.3 配置静态路由与BFD联动 ····· 1-3
1.3.1 双向检测 ····· 1-3
1.3.2 单跳检测 ····· 1-4
1.4 配置静态路由快速重路由功能 ····· 1-5
1.5 静态路由显示和维护 ····· 1-6
1.6 静态路由典型配置举例 ····· 1-7
1.6.1 静态路由基本功能配置举例 ····· 1-7
1.6.2 静态路由快速重路由配置举例 ····· 1-9
1.6.3 配置静态路由与BFD联动（直连） ····· 1-11
1.6.4 配置静态路由与BFD联动（非直连） ····· 1-13

1 静态路由配置

说明

本文所指的路由器代表运行了路由协议的三层设备。

1.1 简介

1.1.1 静态路由

静态路由是一种特殊的路由，由管理员手工配置。配置静态路由后，去往指定目的地的数据报文将按照管理员指定的路径进行转发。

在组网结构比较简单的网络中，只需配置静态路由就可以实现网络互通。恰当地设置和使用静态路由可以改善网络的性能，并可为重要的网络应用保证带宽。

静态路由的缺点在于：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，可能会出现路由不可达，导致网络中断，此时必须由网络管理员手工修改静态路由的配置。

1.1.2 缺省路由

如果到达某个指定网络的数据报文在路由器的路由表里找不到对应的表项，那么该报文将被路由器丢弃。

通过给当前路由器配置一条缺省路由，那些在路由表里找不到匹配路由表入口项的数据报文将会转发给另外一台路由器（如果这台路由器的路由能力比较强，包括到达大部分所有网络的路由信息），由另外一台路由器进行报文的转发。

缺省路由是在路由器没有找到匹配的路由表入口项时才使用的路由：

- 如果报文的目的地址不能与路由表的任何入口项相匹配，那么该报文将选取缺省路由。
- 如果没有缺省路由且报文的目的地不在路由表中，那么该报文将被丢弃。

缺省路由有两种生成方式：

- 第一种是通过网络管理员在路由器上配置到网络 0.0.0.0（掩码也为 0.0.0.0）的静态路由，对于一个到来的数据报文，如果在当前路由器里找不到匹配的路由表项，将会把报文发给在配置的静态路由里指定的下一跳路由器。
- 第二种是通过动态路由协议生成（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。

1.1.3 静态路由应用

配置静态路由时，需要了解以下内容：

(1) 目的地址与掩码

在 **ip route-static** 命令中，IPv4 地址为点分十进制格式，掩码可以用点分十进制表示，也可用掩码长度（即掩码中连续‘1’的位数）表示。

(2) 出接口和下一跳地址

在配置静态路由时，可指定出接口，也可指定下一跳地址。指定出接口还是指定下一跳地址要视具体情况而定，下一跳地址不能为本地接口 IP 地址，否则路由不会生效。

实际上，所有的路由项都必须明确下一跳地址。在发送报文时，首先根据报文的目的地址寻找路由表中与之匹配的路由。只有指定了下一跳地址，链路层才能找到对应的链路层地址，并转发报文。

指定出接口时需要注意：

- 对于 Null0 接口，配置了出接口就不再配置下一跳地址。
- 在配置静态路由时，如果配置广播接口（如以太网接口、VLAN 接口等）为出接口，则必须同时指定其对应的下一跳地址。

(3) 其它属性

对于不同的静态路由，可以为它们配置不同的优先级，从而更灵活地应用路由管理策略。例如：配置到达相同目的地的多条路由，如果指定相同优先级，则可实现负载分担，如果指定不同优先级，则可实现路由备份。

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* | *mask-length* } { *next-hop-address* [**track** *track-entry-number*] | *interface-type interface-number* [*next-hop-address*] | **vpn-instance** *d-vpn-instance-name next-hop-address* [**track** *track-entry-number*] } [**preference** *preference-value*] [**tag** *tag-value*] [**permanent**] [**description** *description-text*]	二者必选其一 缺省情况下，静态路由的优先级**preference**为60，静态路由**tag**值为0，未配置描述信息

操作	命令	说明
	ip route-static vpn-instance *s-vpn-instance-name*&<1-6> *dest-address* { *mask* | *mask-length* } { *next-hop-address* [**public**] [**track** *track-entry-number*] | *interface-type interface-number* [*next-hop-address*] | **vpn-instance** *d-vpn-instance-name next-hop-address* [**track** *track-entry-number*] } [**preference** *preference-value*] [**tag** *tag-value*] [**permanent**] [**description** *description-text*]	参数**permanent**不能和**track**一起进行配置。即使在出接口down时，配置的永久静态路由仍然保持active状态
配置静态路由的缺省优先级	**ip route-static default-preference** *default-preference-value*	可选 缺省情况下，静态路由的缺省优先级为60

说明

- 在配置静态路由时，如果先指定下一跳地址，然后将该地址配置为本地接口（如以太网接口、VLAN 接口等）的 IP 地址，静态路由不会生效。
- 如果在配置静态路由时没有指定优先级，就会使用缺省优先级。重新设置缺省优先级后，新设置的缺省优先级仅对新增的静态路由有效。
- 设置静态路由的 Tag 值，可以在路由策略中根据 Tag 值对路由进行灵活的控制。
- 在使用 **ip route-static** 配置静态路由时，如果将目的地址与掩码配置为全零(0.0.0.0 0.0.0.0)，则表示配置的是缺省路由。
- 关于 Track 的详细介绍，请参见“可靠性配置指导”中的“Track”。

1.3 配置静态路由与BFD联动

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，可以为各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”。

BFD 需要检测两端建立会话。对于动态路由协议来说是有邻居概念的，因此在检测两端动态路由协议都会通知 BFD 会话邻居信息，从而检测两端的 BFD 任务可以通过向邻居发送 BFD 控制报文来建立会话。由于静态路由没有什么邻居的概念一般使用以下方法解决。

1.3.1 双向检测

静态路由使用控制报文方式 BFD 功能时，对端也必须存在对应的 BFD 会话。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。

控制报文方式支持直连下一跳和非直连下一跳。当配置出接口和下一跳时，BFD 建立直连会话；当配置非直连下一跳和指定 BFD 源 IP 地址时，BFD 建立非直连会话。

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* | *mask-length* } *interface-type interface-number next-hop-address* **bfd control-packet** [**preference** *preference-value*] [**tag** *tag-value*] [**description** *description-text*]	二者必选其一
	ip route-static vpn-instance *s-vpn-instance-name*&<1-6> *dest-address* { *mask* | *mask-length* } *interface-type interface-number next-hop-address* **bfd control-packet** [**preference** *preference-value*] [**tag** *tag-value*] [**description** *description-text*]	

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* | *mask-length* } { *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*]	二者必选其一
	ip route-static vpn-instance *s-vpn-instance-name*&<1-6> *dest-address* { *mask* | *mask-length* } { *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

利用 BFD echo 报文，通过报文建立会话，echo 报文的目的地址为本设备接口地址，发送给下一跳设备后会直接转发回本设备而不经过 BFD 任务处理。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报文源地址	**bfd echo-source-ip** *ip-address*	必选 缺省情况下，没有配置echo报文源地址 本命令的详细情况请参见“可靠性命令参考”中的“BFD”

操作	命令	说明
配置静态路由运行BFD	**ip route-static** *dest-address* { *mask* | *mask-length* } *interface-type interface-number next-hop-address* **bfd echo-packet** [**preference** *preference-value*] [**tag** *tag-value*] [**description** *description-text*]	二者必选其一
	ip route-static vpn-instance *s-vpn-instance-name&<1-6> dest-address* { *mask* | *mask-length* } *interface-type interface-number next-hop-address* **bfd echo-packet** [**preference** *preference-value*] [**tag** *tag-value*] [**description** *description-text*]	

注意

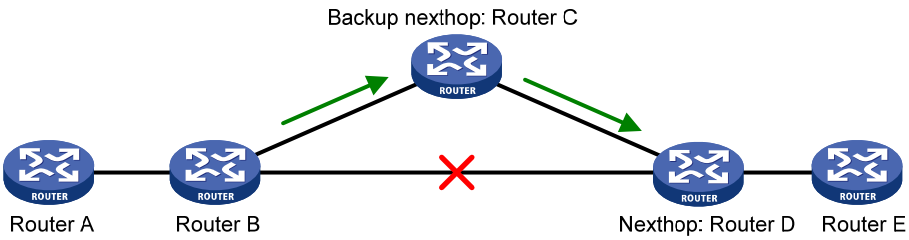
- 路由振荡时，使能 BFD 检测功能可能会加剧振荡，需谨慎使用。
- 当 BFD 会话工作于 echo 报文方式时，必须配置 echo 报文源地址。另外，仅在一端建立 BFD 会话。
- 配置单跳检测时，必须同时指定出接口和路由的下一跳 IP 地址。
- 静态路由如果出接口为含 SPOOFING 属性的接口，不能使用 BFD 进行检测。

1.4 配置静态路由快速重路由功能

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，当路由器通过BFD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. 配置准备

要为静态路由配置快速重路由功能，需要先配置路由策略，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”。

2. 配置过程

表1-5 配置静态路由快速重路由功能

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*	必选 缺省情况下，没有配置BFD echo报文源地址 本命令的详细情况请参见“可靠性命令参考”中的“BFD”
配置静态路由支持快速重路由功能	**ip route-static** [**vpn-instance** *vpn-instance-name*] **fast-reroute route-policy** *route-policy-name*	必选 缺省情况下，没有配置静态路由支持快速重路由功能

说明

- 静态路由支持快速重路由功能仅对在配置中同时指定了出接口和下一跳的静态路由有效。
- 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。

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。在系统视图下执行 **delete** 命令可以删除配置的所有静态路由。

表1-6 静态路由显示和维护

操作	命令
查看静态路由表信息	**display ip routing-table protocol static** [**inactive** | **verbose**] [| { **begin** | **exclude** | **include** } *regular-expression*]
删除所有静态路由	**delete** [**vpn-instance** *vpn-instance-name*] **static-routes all**

说明

命令 **display ip routing-table protocol static** [**inactive** | **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]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”。

1.6 静态路由典型配置举例

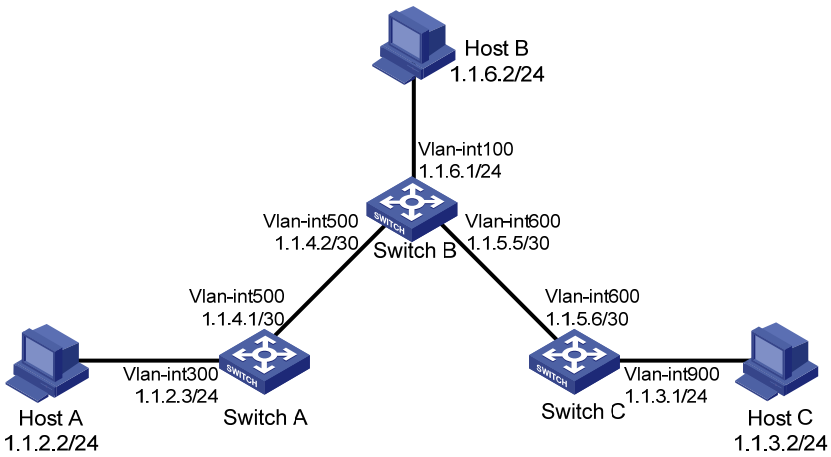
1.6.1 静态路由基本功能配置举例

1. 组网需求

交换机各接口及主机的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 Switch A 上配置缺省路由。

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

在 Switch B 上配置两条静态路由。

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

在 Switch C 上配置缺省路由。

```
<SwitchC> system-view
[SwitchC] 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) 查看配置结果

显示 Switch A 的 IP 路由表。

```
[SwitchA] display ip routing-table
Routing Tables: Public
         Destinations : 7       Routes : 7

Destination/Mask    Proto  Pre  Cost         NextHop         Interface

0.0.0.0/0           Static 60   0            1.1.4.2         Vlan500
1.1.2.0/24          Direct 0    0            1.1.2.3         Vlan300
1.1.2.3/32          Direct 0    0            127.0.0.1       InLoop0
1.1.4.0/30          Direct 0    0            1.1.4.1         Vlan500
1.1.4.1/32          Direct 0    0            127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0            127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
```

显示 Switch B 的 IP 路由表。

```
[SwitchB] display ip routing-table
Routing Tables: Public
         Destinations : 10       Routes : 10

Destination/Mask    Proto  Pre  Cost         NextHop         Interface

1.1.2.0/24          Static 60   0            1.1.4.1         Vlan500
1.1.3.0/24          Static 60   0            1.1.5.6         Vlan600
1.1.4.0/30          Direct 0    0            1.1.4.2         Vlan500
1.1.4.2/32          Direct 0    0            127.0.0.1       InLoop0
1.1.5.0/30          Direct 0    0            1.1.5.5         Vlan600
1.1.5.5/32          Direct 0    0            127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0            127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
1.1.6.0/24          Direct 0    0            1.1.6.1         Vlan100
1.1.6.1/32          Direct 0    0            127.0.0.1       InLoop0
```

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=255
Reply from 1.1.2.2: bytes=32 time=1ms TTL=255
Reply from 1.1.2.2: bytes=32 time=1ms TTL=255
Reply from 1.1.2.2: bytes=32 time=1ms TTL=255

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.
```

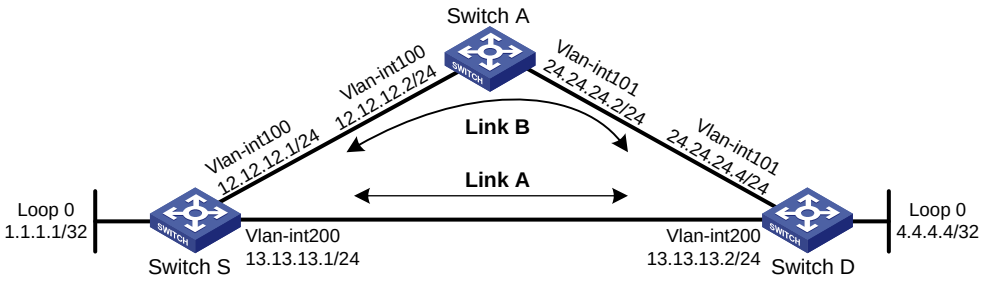
1.6.2 静态路由快速重路由配置举例

1. 组网需求

如 图 1-3 所示，Switch S、Switch A和Switch D通过静态路由实现网络互连。要求当Switch S和Switch 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1-3 静态路由快速重路由配置举例



3. 配置步骤

(1)　配置各交换机接口的 IP 地址与静态路由

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。

在 Switch S、Switch A 和 Switch D 上配置静态路由使 Switch S 可以到达 Switch D 的 Loopback0 口，Switch D 可以到达 Switch S 的 Loopback0 口。

具体配置如下：

在 Switch S 上配置静态路由。

```
<SwitchS> system-view
[SwitchS] ip route-static 4.4.4.4 32 vlan-interface 200 13.13.13.2
[SwitchS] ip route-static 4.4.4.4 32 vlan-interface 100 12.12.12.2 preference 65
```

在 Switch D 上配置静态路由。

```
<SwitchD> system-view
[SwitchD] ip route-static 1.1.1.1 32 vlan-interface 200 13.13.13.1
[SwitchD] ip route-static 1.1.1.1 32 vlan-interface 101 24.24.24.2 preference 65
```

在 Switch A 上配置静态路由。

```
<SwitchA> system-view
[SwitchA] ip route-static 4.4.4.4 32 vlan-interface 101 24.24.24.4
[SwitchA] ip route-static 1.1.1.1 32 vlan-interface 100 12.12.12.1
```

(2) 配置静态路由快速重路由

配置 Switch S。

```
[SwitchS] bfd echo-source-ip 2.2.2.2
[SwitchS] ip ip-prefix abc index 10 permit 4.4.4.4 32
[SwitchS] route-policy frr permit node 10
[SwitchS-route-policy] if-match ip-prefix abc
[SwitchS-route-policy] apply fast-reroute backup-interface vlan-interface 100
backup-nexthop 12.12.12.2
[SwitchS-route-policy] quit
[SwitchS] ip route-static fast-reroute route-policy frr
```

配置 Switch D。

```
[SwitchD] bfd echo-source-ip 3.3.3.3
[SwitchD] ip ip-prefix abc index 10 permit 1.1.1.1 32
[SwitchD] route-policy frr permit node 10
[SwitchD-route-policy] if-match ip-prefix abc
[SwitchD-route-policy] apply fast-reroute backup-interface vlan-interface 101
backup-nexthop 24.24.24.2
[SwitchD-route-policy] quit
[SwitchD] ip route-static fast-reroute route-policy frr
```

(3) 检验配置效果

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

```
[SwitchS] display ip routing-table 4.4.4.4 verbose
Routing Table : Public
Summary Count : 1

  Destination: 4.4.4.4/32
     Protocol: Static          Process ID: 0
   Preference: 60                    Cost: 0
 IpPrecedence:                    QosLcId:
      NextHop: 13.13.13.2       Interface: vlan 200
    BkNextHop: 12.12.12.2     BkInterface: vlan 100
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv             Age: 00h01m27s
          Tag: 0
```

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

```
[SwitchD] display ip routing-table 1.1.1.1 verbose
Routing Table : Public
Summary Count : 1

  Destination: 1.1.1.1/32
     Protocol: Static          Process ID: 0
   Preference: 60                    Cost: 0
```

```
IpPrecedence:                    QosLcId:
     NextHop: 13.13.13.1       Interface: vlan 200
   BkNextHop: 24.24.24.2     BkInterface: vlan 101
 RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
   Tunnel ID: 0x0                 Label: NULL
 BKTunnel ID: 0x0               BKLabel: NULL
       State: Active Adv            Age: 00h01m27s
         Tag: 0
```

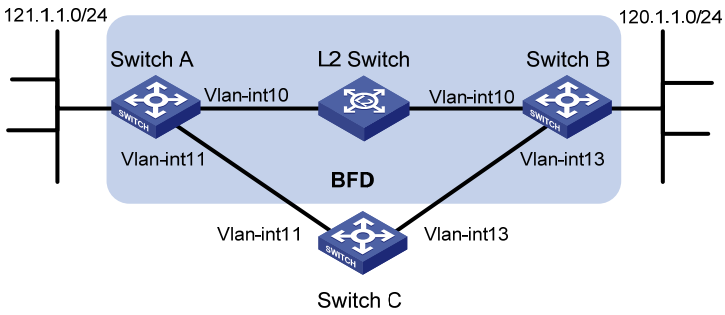
1.6.3 配置静态路由与BFD联动（直连）

1. 组网需求

- 在 Switch A 上配置静态路由可以到达 120.1.1.0/24 网段路由，在 Switch B 上配置静态路由可以到达 121.1.1.0/24 网段路由，并都使能 BFD 检测功能。
- 在 Switch 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 Switch A 和 Switch 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 Switch C 进行通信。

2. 组网图

图1-4 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
Switch A	Vlan-int10	12.1.1.1/24	Switch B	Vlan-int10	12.1.1.2/24
	Vlan-int11	10.1.1.102/24		Vlan-int13	13.1.1.1/24
Switch C	Vlan-int11	10.1.1.100/24			
	Vlan-int13	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 Switch A 上配置静态路由，并使能 BFD 检测功能，通过 BFD 控制报文方式实现 BFD 功能。

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 10
[SwitchA-vlan-interface10] bfd min-transmit-interval 500
[SwitchA-vlan-interface10] bfd min-receive-interval 500
[SwitchA-vlan-interface10] bfd detect-multiplier 9
[SwitchA-vlan-interface10] quit
[SwitchA] ip route-static 120.1.1.0 24 vlan-interface 10 12.1.1.2 bfd control-packet
```

```
[SwitchA] ip route-static 120.1.1.0 24 vlan-interface 11 10.1.1.100 preference 65
[SwitchA] quit
```

在 Switch B 上配置静态路由，并使能 BFD 检测功能，通过 BFD 控制报文方式实现 BFD 功能。

```
<SwitchB> system-view
[SwitchB] interface vlan-interface 10
[SwitchB-vlan-interface10] bfd min-transmit-interval 500
[SwitchB-vlan-interface10] bfd min-receive-interval 500
[SwitchB-vlan-interface10] bfd detect-multiplier 9
[SwitchB-vlan-interface10] quit
[SwitchB] ip route-static 121.1.1.0 24 vlan-interface 10 12.1.1.1 bfd control-packet
[SwitchB] ip route-static 121.1.1.0 24 vlan-interface 13 13.1.1.2 preference 65
[SwitchB] quit
```

在 Switch C 上配置静态路由。

```
<SwitchC> system-view
[SwitchC] ip route-static 120.1.1.0 24 vlan-interface 13 13.1.1.1
[SwitchC] ip route-static 121.1.1.0 24 vlan-interface 11 10.1.1.102
```

(3) 检查配置结果

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

查看 BFD 会话，可以看到 BFD 会话已经创建。

```
<SwitchA> display bfd session

 Total Session Num: 1            Init Mode: Active

 Session Working Under Ctrl Mode:

 LD/RD         SourceAddr      DestAddr        State Holdtime Interface
 4/7           12.1.1.1        12.1.1.2        Up    2000ms   Vlan10
```

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

```
<SwitchA> display ip routing-table protocol static
Public Routing Table : Static
Summary Count : 2

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        Vlan10

Direct Routing table Status : <Inactive>
Summary Count : 1
Destination/Mask    Proto  Pre  Cost         NextHop         Interface
120.1.1.0/24        Static 65   0            10.1.1.100      Vlan11
```

当 Switch A 和二层交换机之间链路发生故障时， Switch A 能够快速感知 Switch B 的变化。

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

```
<SwitchA> display ip routing-table protocol static
Public Routing Table : 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      Vlan11
Static Routing table Status : <Inactive>
Summary Count : 0
```

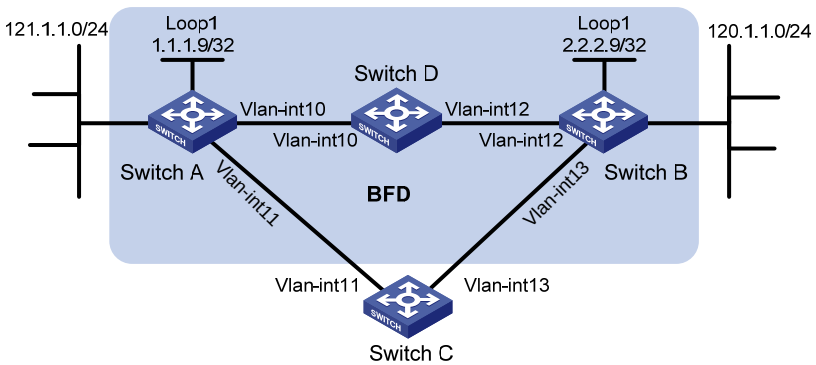
1.6.4 配置静态路由与BFD联动（非直连）

1. 组网需求

- 在 Switch A 上配置静态路由可以到达 120.1.1.0/24 网段路由，在 Switch B 上配置静态路由可以到达 121.1.1.0/24 网段路由，并都使能 BFD 检测功能。
- 在 Switch C 和 Switch 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。
- Switch A 存在到 Switch B 的接口 Loopback1（2.2.2.9/32）的路由，出接口为 Vlan-interface10；Switch B 存在到 Switch A 的接口 Loopback1（1.1.1.9/32）的路由，出接口为 Vlan-interface12；Switch D 存在到 1.1.1.9/32 的路由，出接口为 Vlan-interface10，存在到 2.2.2.9/32 的路由，出接口为 Vlan-interface12。
- 当 Switch A 和 Switch B 通过 Switch D 通信的链路出现故障时，BFD 能够快速感知，并且切换到 Switch C 进行通信。

2. 组网图

图1-5 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
Switch A	Vlan-int10	12.1.1.1/24	Switch B	Vlan-int12	11.1.1.1/24
	Vlan-int11	10.1.1.102/24		Vlan-int13	13.1.1.1/24
	Loop1	1.1.1.9/32		Loop1	2.2.2.9/32
Switch C	Vlan-int11	10.1.1.100/24	Switch D	Vlan-int10	12.1.1.2/24
	Vlan-int13	13.1.1.2/24		Vlan-int12	11.1.1.2/24

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 Switch A 上配置静态路由，并使能 BFD 检测功能，通过 BFD 控制报文方式实现 BFD 功能。

```
<SwitchA> system-view
[SwitchA] interface loopback 1
[SwitchA-LoopBack1] bfd min-transmit-interval 500
```

```
[SwitchA-LoopBack1] bfd min-receive-interval 500
[SwitchA-LoopBack1] bfd detect-multiplier 9
[SwitchA-LoopBack1] quit
[SwitchA] ip route-static 120.1.1.0 24 2.2.2.9 bfd control-packet bfd-source 1.1.1.9
[SwitchA] ip route-static 120.1.1.0 24 vlan-interface 11 10.1.1.100 preference 65
[SwitchA] quit
```

在 Switch B 上配置静态路由，并使能 BFD 检测功能，通过 BFD 控制报文方式实现 BFD 功能。

```
<SwitchB> system-view
[SwitchB] interface loopback 1
[SwitchB-LoopBack1] bfd min-transmit-interval 500
[SwitchB-LoopBack1] bfd min-receive-interval 500
[SwitchB-LoopBack1] bfd detect-multiplier 9
[SwitchB-LoopBack1] quit
[SwitchB] ip route-static 121.1.1.0 24 1.1.1.9 bfd control-packet bfd-source 2.2.2.9
[SwitchB] ip route-static 121.1.1.0 24 vlan-interface 13 13.1.1.2 preference 65
[SwitchB] quit
```

在 Switch C 上配置静态路由。

```
<SwitchC> system-view
[SwitchC] ip route-static 120.1.1.0 24 vlan-interface 13 13.1.1.1
[SwitchC] ip route-static 121.1.1.0 24 vlan-interface 11 10.1.1.102
```

在 Switch D 上配置静态路由。

```
<SwitchD> system-view
[SwitchD] ip route-static 120.1.1.0 24 vlan-interface 12 11.1.1.1
[SwitchD] ip route-static 121.1.1.0 24 vlan-interface 10 12.1.1.1
```

(3) 检查配置结果

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

查看 BFD 会话，可以看到 BFD 会话已经创建。

```
<SwitchA> display bfd session

 Total Session Num: 1            Init Mode: Active

 Session Working Under Ctrl Mode:

 LD/RD         SourceAddr      DestAddr        State Holdtime Interface
 4/7           1.1.1.9         2.2.2.9         Up    2000ms   Loop1
```

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

```
<SwitchA> display ip routing-table protocol static
Public Routing Table : Static
Summary Count : 2

Static Routing table Status : <Active>
Summary Count : 1
Destination/Mask    Proto  Pre  Cost         NextHop         Interface
120.1.1.0/24        Static 60   0            2.2.2.9         Vlan10

Static Routing table Status : <Inactive>
Summary Count : 1
```

```
Destination/Mask    Proto  Pre  Cost         NextHop         Interface
120.1.1.0/24        Static 65   0            10.1.1.100      Vlan11
```

当 Switch A 和 Switch D 之间链路发生故障时， Switch A 能够快速感知 Switch B 的变化。

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

```
<SwitchA> display ip routing-table protocol static
Public Routing Table : 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      Vlan11
Static Routing table Status : <Inactive>
Summary Count : 0
```

目 录

1 RIP配置 ····· 1-1

1.1 简介 ····· 1-1

1.1.1 RIP工作机制 ····· 1-1

1.1.2 RIP的启动和运行过程 ····· 1-2

1.1.3 RIP的版本 ····· 1-3

1.1.4 RIP的报文格式 ····· 1-3

1.1.5 支持的RIP特性 ····· 1-5

1.1.6 协议规范 ····· 1-5

1.2 RIP配置任务简介 ····· 1-5

1.3 配置RIP的基本功能 ····· 1-6

1.3.1 配置准备 ····· 1-6

1.3.2 配置RIP的基本功能 ····· 1-6

1.4 配置RIP的路由信息控制 ····· 1-9

1.4.1 配置接口附加度量值 ····· 1-9

1.4.2 配置RIP-2 路由聚合 ····· 1-9

1.4.3 禁止RIP接收主机路由 ····· 1-11

1.4.4 配置RIP发布缺省路由 ····· 1-11

1.4.5 配置RIP对接收/发布的路由进行过滤 ····· 1-12

1.4.6 配置RIP协议优先级 ····· 1-13

1.4.7 配置RIP引入外部路由 ····· 1-13

1.5 调整和优化RIP网络 ····· 1-13

1.5.1 配置RIP定时器 ····· 1-14

1.5.2 配置水平分割和毒性逆转 ····· 1-14

1.5.3 配置最大等价路由条数 ····· 1-15

1.5.4 配置RIP-1 报文的零域检查 ····· 1-15

1.5.5 配置源地址检查 ····· 1-16

1.5.6 配置RIP-2 报文的认证方式 ····· 1-16

1.5.7 配置RIP邻居 ····· 1-17

1.5.8 配置RIP和MIB绑定 ····· 1-17

1.5.9 配置RIP报文的发送速率 ····· 1-17

1.6 配置RIP快速重路由功能 ····· 1-18

1.7 配置RIP与BFD联动 ····· 1-19

1.7.1 echo报文单跳检测 ····· 1-19

1.7.2 control报文双向检测 ····· 1-20
1.8 RIP显示和维护 ····· 1-20
1.9 RIP典型配置举例 ····· 1-21
1.9.1 配置RIP的版本 ····· 1-21
1.9.2 配置RIP引入外部路由 ····· 1-22
1.9.3 配置RIP接口附加度量值 ····· 1-25
1.9.4 配置RIP发布聚合路由 ····· 1-27
1.9.5 配置RIP快速重路由 ····· 1-29
1.9.6 配置RIP运行BFD（echo报文单跳检测） ····· 1-31
1.9.7 配置RIP运行BFD（control报文双向检测） ····· 1-34
1.10 常见配置错误举例 ····· 1-37
1.10.1 收不到邻居的RIP更新报文 ····· 1-37
1.10.2 RIP网络发生路由振荡 ····· 1-38

1 RIP配置

说明

本文所指的路由器代表运行了路由协议的三层设备。

1.1 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工作机制

1. RIP的基本概念

RIP 是一种基于距离矢量（Distance-Vector，D-V）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定时器

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。

4. 防止路由环路

RIP 是一种基于 D-V 算法的路由协议，由于它向邻居通告的是自己的路由表，存在发生路由环路的可能性。

RIP 通过以下机制来避免路由环路的产生：

- 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。
- 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。
- 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。

1.1.2 RIP的启动和运行过程

RIP 启动和运行的整个过程可描述如下：

- 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）。
- 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。
- 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。

1.1.3 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，并提供简单验证和 MD5 验证两种方式，增强安全性。

说明

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。

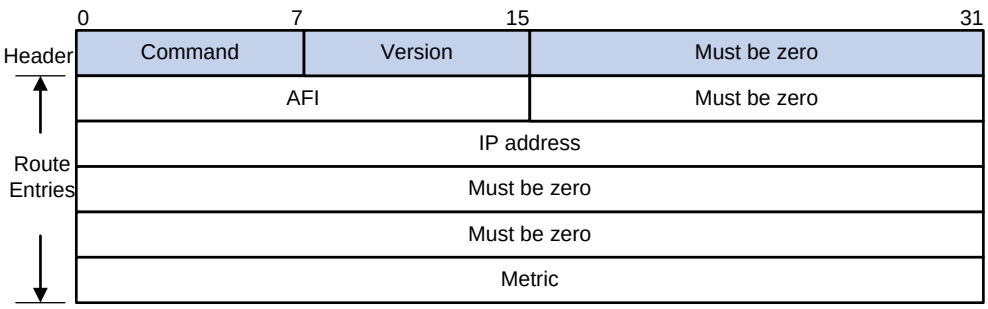
1.1.4 RIP的报文格式

RIP 报文由头部（Header）和多个路由表项（Route Entries）部分组成。在一个 RIP 报文中，最多可以有 25 个路由表项（如果是 RIP-2 的验证报文，由于第一个路由表项作为验证项，所以最多可以有 24 个路由表项）。

1. RIP-1 的报文格式

RIP-1 的报文格式如 图 1-1 所示。

图1-1 RIP-1 的报文格式



各字段的解释如下：

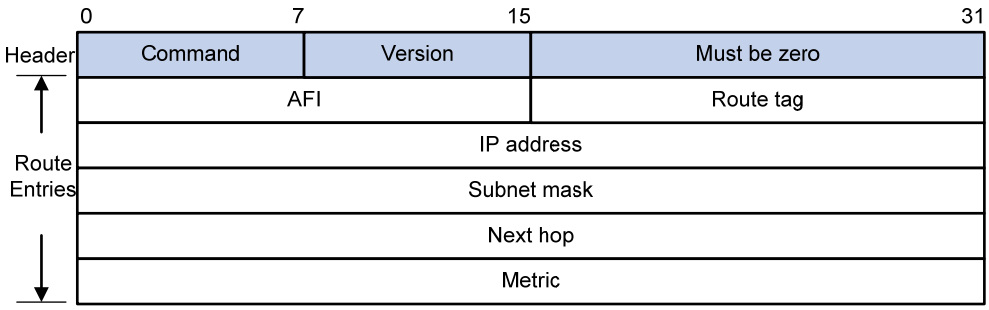
- Command：标识报文的类型。值为 1 时表示 Request 报文，向邻居请求全部或部分路由信息；值为 2 表示 Response 报文，发送全部或部分路由信息，1 个 Response 报文中最多包含 25 个路由表项。

- Version：RIP 的版本号。对于 RIP-1 来说其值为 0x01。
- Must be zero：必须为 0 字段。
- AFI（Address Family Identifier）：地址族标识，其值为 2 时表示 IP 协议。
- IP address：该路由的目的 IP 地址，可以是自然网段地址、子网地址或主机地址。
- Metric：路由的度量值。

2. RIP-2 的报文格式

RIP-2 的报文格式与RIP-1 类似，如 图 1-2 所示。

图1-2 RIP-2 的报文格式



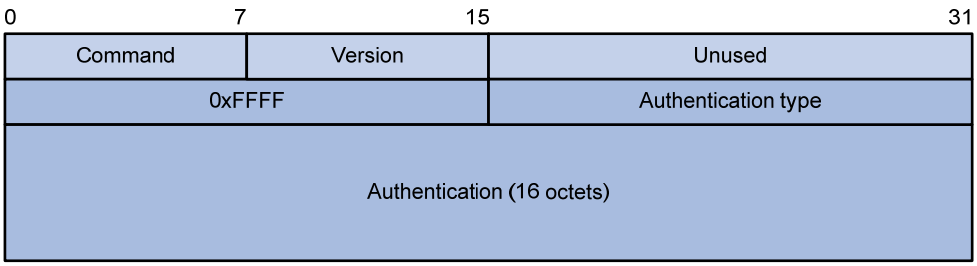
其中，与 RIP-1 不同的字段有：

- Version：RIP 的版本号。对于 RIP-2 来说其值为 0x02。
- Route Tag：路由标记。
- IP address：该路由的目的 IP 地址，可以是自然网段地址、子网地址或主机地址。
- Subnet Mask：目的地址的掩码。
- Next Hop：如果为 0.0.0.0，则表示发布此条路由信息的路由器地址就是最优下一跳地址，否则表示提供了一个比发布此条路由信息的路由器地址更优的下一跳地址。

3. RIP-2 的验证报文格式

RIP-2 为了支持报文验证，使用第一个路由表项（Route Entry）作为验证项，并将AFI字段的值设为 0xFFFF标识报文携带认证信息，如 图 1-3 所示。

图1-3 RIP-2 的验证报文格式



各字段的解释如下：

- Authentication Type：验证类型。值为 2 时表示简单验证，值为 3 时表示 MD5 验证。

- Authentication：验证字。当使用简单验证时包含了密码信息；当使用 MD5 验证时包含了 Key ID、MD5 验证数据长度和序列号的信息。

说明

- RFC 1723 中只定义了简单验证方式，关于 MD5 验证的详细信息，请参见 RFC 2453“RIP Version 2”。
- 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。

1.1.5 支持的RIP特性

目前设备支持以下 RIP 特性：

- 支持 RIP-1 和 RIP-2
- 支持 RIP 多实例

它可以作为 VPN 内部路由协议，在 BGP/MPLS VPN 网络中的 CE 和 PE 之间运行，相关概念及应用请参见“MPLS 配置指导”中的“MPLS L3VPN”。

- RIP 快速重路由
- 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。

1.1.6 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 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		必选	1.3
配置RIP路由特性	配置接口附加度量值	可选	1.4.1

配置任务		说明	详细配置
	配置RIP-2路由聚合	可选	1.4.2
	禁止RIP接收主机路由	可选	1.4.3
	配置RIP发布缺省路由	可选	1.4.4
	配置RIP对接收/发布的路由进行过滤	可选	1.4.5
	配置RIP协议优先级	可选	1.4.6
	配置RIP引入外部路由	可选	1.4.7
调整和优化RIP网络	配置RIP定时器	可选	1.5.1
	配置水平分割和毒性逆转	可选	1.5.2
	配置最大等价路由条数	可选	1.5.3
	配置RIP-1报文的零域检查	可选	1.5.4
	配置源地址检查	可选	1.5.5
	配置RIP-2报文的认证方式	可选	1.5.6
	配置RIP邻居	可选	1.5.7
	配置RIP和MIB绑定	可选	1.5.8
	配置RIP报文的发送速率	可选	1.5.9
配置RIP快速重路由功能		可选	1.6
配置RIP与BFD联动	echo报文单跳检测	可选	1.7.1
	control报文双向检测	可选	1.7.2

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 配置RIP的基本功能

1. 启动RIP，配置指定网段范围内的接口运行RIP

表1-2 启动 RIP，配置指定的接口运行 RIP

操作	命令	说明
进入系统视图	**system-view**	-
创建RIP进程并进入RIP视图	**rip** [*process-id*] [**vpn-instance** *vpn-instance-name*]	必选 缺省情况下，RIP进程处于关闭状态

操作	命令	说明
在指定网段接口上使能RIP	**network** *network-address*	必选 缺省情况下，接口上的RIP功能处于关闭状态

- 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 只在指定网段的接口上运行；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。
- **network** 0.0.0.0 命令用来在所有接口上使能 RIP。
- 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。

2. 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，即接口只接收路由更新报文而不发送路由更新报文
- 配置接口接收 RIP 报文
- 配置接口发送 RIP 报文

表1-3 配置接口的工作状态

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*]	-
配置接口工作在抑制状态	**silent-interface** { *interface-type interface-number* | **all** }	可选 缺省情况下，所有使能RIP的接口发送路由更新报文
退回系统视图	**quit**	-
进入接口视图	**interface** *interface-type interface-number*	-
允许接口接收RIP报文	**rip input**	可选 缺省情况下，所有使能RIP的接口接收RIP报文
允许接口发送RIP报文	**rip output**	可选 缺省情况下，所有使能RIP的接口发送RIP报文

3. 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，如果希望接口配置的 RIP 版本与全局配置的不一样，则进入接口视图配置接口运行的 RIP 版本。
- 当接口运行的 RIP 版本为 RIP-1 时，发送 RIP-1 广播报文，可以接收 RIP-1 广播/单播报文。
- 当接口运行的 RIP 版本为 RIP-2 且工作在组播方式时，发送 RIP-2 组播报文，可以接收 RIP-2 广播/组播/单播报文。
- 当接口运行的 RIP 版本为 RIP-2 且工作在广播方式时，发送 RIP-2 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。

表1-4 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** }	可选 缺省情况下，如果接口配置了RIP版本，以接口配置的为准，如果接口也没有配置，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**] }	可选 缺省情况下，如果没有配置接口运行的RIP版本，接口运行的RIP版本以全局配置的为准；如果也没有进行全局RIP版本的配置，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

4. 配置RIP报文的DSCP优先级

在 IPv4 报文头中，包含一个 8bit 的 ToS 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 RIP 报文的 DSCP 优先级进行配置。

表1-5 配置 RIP 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
创建RIP进程并进入RIP视图	**rip** [*process-id*] [**vpn-instance** *vpn-instance-name*]	必选 缺省情况下，RIP进程处于关闭状态
配置RIP报文的DSCP优先级	**dscp** *dscp-value*	可选 缺省情况下，RIP报文的DSCP优先级为48

1.4 配置RIP的路由信息控制

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 基本功能

1.4.1 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，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*	可选 缺省情况下，接口接收RIP路由时的附加度量值为0
配置接口发送RIP路由时的附加度量值	**rip metricout** [**route-policy** *route-policy-name*] *value*	可选 缺省情况下，接口发送RIP路由时的附加路由度量值为1

1.4.2 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。

在大型网络中，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**	可选 缺省情况下，RIP-2自动路由聚合功能处于使能状态 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 Vlan-interface11 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**	必选 缺省情况下，RIP-2自动路由聚合功能处于使能状态

操作	命令	说明
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* | *mask-length* }	必选

1.4.3 禁止RIP接收主机路由

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**	必选 缺省情况下，允许RIP接收主机路由

说明

禁止接收主机路由仅对 RIPv2 有效，对 RIPv1 无效。

1.4.4 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*]	-

操作	命令	说明
配置RIP发布缺省路由	**default-route** { **only** | **originate** } [**cost** *cost*]	可选 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** }	可选 缺省情况下，RIP接口是否发布缺省路由以RIP进程配置为准

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。

1.4.5 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，通过指定访问控制列表和地址前缀列表，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，还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** *ip-prefix-name* | **ip-prefix** *ip-prefix-name* [**gateway** *ip-prefix-name*] } **import** [*interface-type interface-number*]	必选 缺省情况下，RIP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** { *acl-number* | **ip-prefix** *ip-prefix-name* } **export** [*protocol* [*process-id*] | *interface-type interface-number*]	必选 缺省情况下，RIP不对发布的路由信息进行过滤

说明

- **filter-policy import** 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。
- **filter-policy export** 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。

1.4.6 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*	可选 缺省情况下，RIP路由的优先级为100

1.4.7 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、ISIS、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*]	-
配置引入路由的缺省度量值	**default cost** *value*	可选 缺省情况下，引入路由的缺省度量值为0
引入外部路由	**import-route** *protocol* [*process-id* | **all-processes** | **allow-ibgp**] [**cost** *cost* | **route-policy** *route-policy-name* | **tag** *tag*] *	必选 缺省情况下，RIP不引入其它路由

说明

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。

1.5 调整和优化RIP网络

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 基本功能

1.5.1 配置RIP定时器

通过调整 RIP 定时器可以改变 RIP 网络的收敛速度。

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* } *	可选 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

说明

在配置 RIP 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 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。

1.5.2 配置水平分割和毒性逆转

说明

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*	-

操作	命令	说明
使能水平分割功能	**rip split-horizon**	可选 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**	必选 缺省情况下，毒性逆转功能处于关闭状态

1.5.3 配置最大等价路由条数

通过配置最大等价路由跳数，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*	可选 缺省值为8

1.5.4 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**	可选 缺省情况下，RIP-1报文的零域检查功能处于使能状态

1.5.5 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。

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**	可选 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

说明

当存在 RIP 非直连的邻居时，应该取消源地址检查。

1.5.6 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**] *key-string key-id* | **rfc2453** [**cipher**] *key-string* } | **simple** [**cipher**] *password* }	必选

说明

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。

1.5.7 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*	必选
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**	必选 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

说明

- 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。
- 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。

1.5.8 配置RIP和MIB绑定

通过将 MIB 操作绑定在指定的 RIP 进程上，可以让指定的 RIP 进程接收 SNMP 请求。

表1-22 配置 RIP 和 MIB 绑定

操作	命令	说明
进入系统视图	**system-view**	-
配置RIP和MIB绑定	**rip mib-binding** *process-id*	可选 缺省情况下，MIB操作绑定在RIP进程1上

1.5.9 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**	-
启动RIP并进入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*	可选 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

1.6 配置RIP快速重路由功能

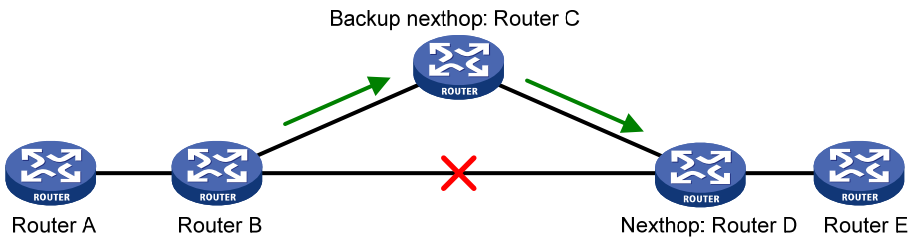
说明

- RIP 支持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 支持快速重路由功能不能与 RIP 支持 BFD 监测同时使用，否则可能导致快速重路由功能失效。

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1-4 RIP 快速重路由功能示意图



如 图 1-4 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探测到网络故障时，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. 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”。

2. 配置限制和指导

该功能在主链路三层口 UP、主链路状态由双通变为单通或者不通的情况下，可以实现下一跳的快速切换（单通现象是指一条链路上的两端，有且只有一端可收到另一端发来的报文，此时链路称为单向链路）；在主链路三层口 Down 的情况下，本功能不可用。

3. 配置过程

表1-24 配置 RIP 快速重路由功能

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*	必选 缺省情况下，没有配置BFD Echo报文源地址
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*	必选 缺省情况下，RIP快速重路由功能处于关闭状态

1.7 配置RIP与BFD联动

说明

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”。

RIP 支持 BFD 提供了两种检测方式：

- 直连邻居采用 echo 报文单跳检测方式，在对端有路由发送时才能建立 BFD 会话。
- 非直连邻居采用 control 报文双向检测方式，当两端互有路由发送时，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1-25 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**	-

操作	命令	说明
配置echo报文源地址	**bfd echo-source-ip** *ip-address*	必选 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BFD功能	**rip bfd enable**	必选 缺省情况下，BFD功能处于关闭状态

1.7.2 control报文双向检测

表1-26 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**	-
创建RIP进程并进入RIP视图	**rip** [*process-id*] [**vpn-instance** *vpn-instance-name*]	必选 缺省情况下，RIP进程处于关闭状态
配置RIP邻居	**peer** *ip-address*	必选 缺省情况下，RIP不向任何定点地址发送更新报文
进入接口视图	**interface** *interface-type interface-number*	-
使能BFD功能	**rip bfd enable**	必选 缺省情况下，BFD功能处于关闭状态

说明

- 当检测 RIP 的会话两端在直连网段（即 IP 报文的一跳），适合采用 BFD 的 echo 单向检测方式，但是，经过多跳到达邻居时 echo 方式则会失效。
- 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。

1.8 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1-27 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* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示RIP发布数据库的所有激活路由	**display rip** *process-id* **database** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示RIP的接口信息	**display rip** *process-id* **interface** [*interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定RIP进程的路由信息	**display rip** *process-id* **route** [*ip-address* { *mask* | *mask-length* } | **peer** *ip-address* | **statistics**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
重启RIP进程	**reset rip** *process-id* **process**
清除指定RIP进程的统计信息	**reset rip** *process-id* **statistics**

1.9 RIP典型配置举例

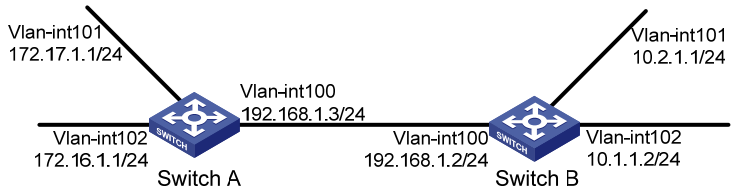
1.9.1 配置RIP的版本

1. 组网需求

如 图 1-5 所示，要求在Switch A和Switch B的所有接口上使能RIP，并使用RIP-2 进行网络互连。

2. 组网图

图1-5 配置 RIP 的版本



3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 RIP 基本功能

配置 Switch A。

```
[SwitchA] rip
[SwitchA-rip-1] network 192.168.1.0
[SwitchA-rip-1] network 172.16.0.0
[SwitchA-rip-1] network 172.17.0.0
[SwitchA-rip-1] quit
```

配置 Switch B。

```
[SwitchB] rip
[SwitchB-rip-1] network 192.168.1.0
[SwitchB-rip-1] network 10.0.0.0
```

```
[SwitchB-rip-1] quit
```

查看 Switch A 的 RIP 路由表。

```
[SwitchA] display rip 1 route
 Route Flags: R - RIP, T - TRIP
              P - Permanent, A - Aging, S - Suppressed, G - Garbage-collect
 ------------------------------------------------------------------------------
 Peer 192.168.1.2  on Vlan-interface100
      Destination/Mask        Nexthop     Cost    Tag    Flags    Sec
         10.0.0.0/8        192.168.1.2      1       0     RA       11
```

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 Switch A 上配置 RIP-2。

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

在 Switch B 上配置 RIP-2。

```
[SwitchB] rip
[SwitchB-rip-1] version 2
[SwitchB-rip-1] undo summary
```

查看 Switch A 的 RIP 路由表。

```
[SwitchA] display rip 1 route
 Route Flags: R - RIP, T - TRIP
              P - Permanent, A - Aging, S - Suppressed, G - Garbage-collect
 ------------------------------------------------------------------------------

 Peer 192.168.1.2  on Vlan-interface100
      Destination/Mask        Nexthop     Cost    Tag    Flags    Sec
         10.0.0.0/8        192.168.1.2      1       0     RA       50
         10.2.1.0/24       192.168.1.2      1       0     RA       16
         10.1.1.0/24       192.168.1.2      1       0     RA       16
```

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，路由表中可能还会存在 RIP-1 的路由信息。

1.9.2 配置RIP引入外部路由

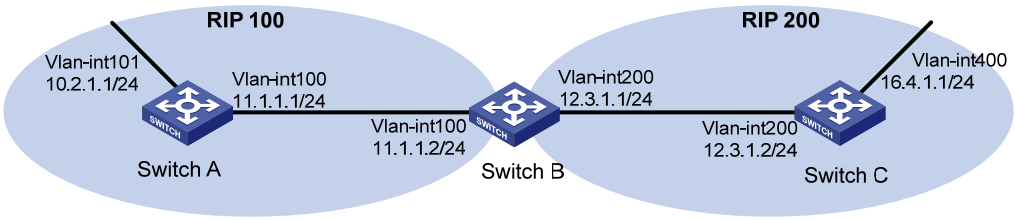
1. 组网需求

- Switch B 上运行两个 RIP 进程：RIP 100 和 RIP 200。Switch B 通过 RIP 100 和 Switch A 交换路由信息，通过 RIP 200 和 Switch C 交换路由信息。

- 在 Switch 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 Switch C 能够学习到达 10.2.1.0/24 和 11.1.1.0/24 的路由，但 Switch A 不能学习到达 12.3.1.0/24 和 16.4.1.0/24 的路由。
- 在 Switch B 配置过滤策略，对引入的 RIP 100 的一条路由（10.2.1.1/24）进行过滤，使其不发布给 Switch C。

2. 组网图

图1-6 配置 RIP 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 RIP 基本功能

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

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

在 Switch 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。

```
<SwitchB> system-view
[SwitchB] rip 100
[SwitchB-rip-100] network 11.0.0.0
[SwitchB-rip-100] version 2
[SwitchB-rip-100] undo summary
[SwitchB-rip-100] quit
[SwitchB] rip 200
[SwitchB-rip-200] network 12.0.0.0
[SwitchB-rip-200] version 2
[SwitchB-rip-200] undo summary
[SwitchB-rip-200] quit
```

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

```
<SwitchC> system-view
[SwitchC] rip 200
[SwitchC-rip-200] network 12.0.0.0
[SwitchC-rip-200] network 16.0.0.0
[SwitchC-rip-200] version 2
```

```
[SwitchC-rip-200] undo summary
[SwitchC-rip-200] quit
```

查看 Switch C 的路由表信息。

```
[SwitchC] display ip routing-table
Routing Tables: Public
         Destinations : 6        Routes : 6
Destination/Mask    Proto  Pre  Cost         NextHop         Interface
12.3.1.0/24         Direct 0    0            12.3.1.2        Vlan200
12.3.1.2/32         Direct 0    0            127.0.0.1       InLoop0
16.4.1.0/24         Direct 0    0            16.4.1.1        Vlan400
16.4.1.1/32         Direct 0    0            127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0            127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
```

(3) 配置 RIP 引入外部路由

在 Switch 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。

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

查看路由引入后 Switch C 的路由表信息。

```
[SwitchC] display ip routing-table
Routing Tables: Public
         Destinations : 8        Routes : 8
Destination/Mask    Proto  Pre  Cost         NextHop         Interface
10.2.1.0/24         RIP    100  1            12.3.1.1        Vlan200
11.1.1.0/24         RIP    100  1            12.3.1.1        Vlan200
12.3.1.0/24         Direct 0    0            12.3.1.2        Vlan200
12.3.1.2/32         Direct 0    0            127.0.0.1       InLoop0
16.4.1.0/24         Direct 0    0            16.4.1.1        Vlan400
16.4.1.1/32         Direct 0    0            127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0            127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
```

(4) 配置 RIP 对引入的路由进行过滤

在 Switch B 上配置 ACL，并对引入的 RIP 进程 100 的路由进行过滤，使其不发布给 Switch C。

```
[SwitchB] acl number 2000
[SwitchB-acl-basic-2000] rule deny source 10.2.1.1 0.0.0.255
[SwitchB-acl-basic-2000] rule permit
[SwitchB-acl-basic-2000] quit
[SwitchB] rip 200
[SwitchB-rip-200] filter-policy 2000 export rip 100
```

查看过滤后 Switch C 的路由表。

```
[SwitchC] display ip routing-table
Routing Tables: Public
         Destinations : 7        Routes : 7

Destination/Mask    Proto  Pre  Cost         NextHop         Interface
```

```
11.1.1.0/24         RIP    100  1           12.3.1.1        Vlan200
12.3.1.0/24         Direct 0    0           12.3.1.2        Vlan200
12.3.1.2/32         Direct 0    0           127.0.0.1       InLoop0
16.4.1.0/24         Direct 0    0           16.4.1.1        Vlan400
16.4.1.1/32         Direct 0    0           127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0           127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0           127.0.0.1       InLoop0
```

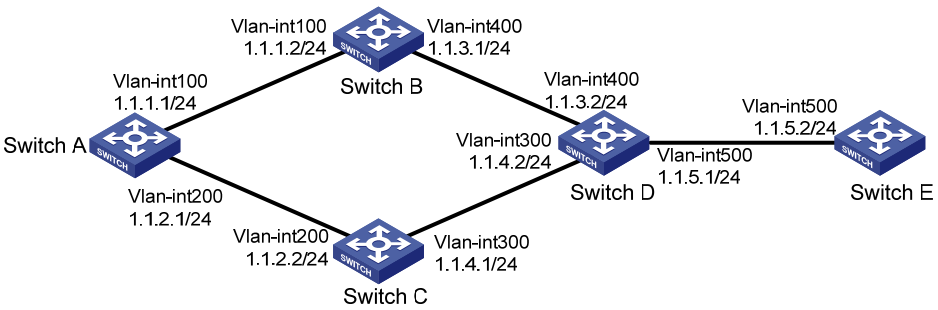
1.9.3 配置RIP接口附加度量值

1. 组网需求

- 在 Switch A、Switch B、Switch C、Switch D 和 Switch E 的所有接口上使能 RIP，并使用 RIP-2 进行网络互连。
- Switch A 有两条链路可以到达 Switch D，其中，通过 Switch B 到达 Switch D 的链路比通过 Switch C 到达 Switch D 的链路更加稳定。通过在 Switch A 的 Vlan-interface200 上配置接口接收 RIP 路由的附加度量值，使得 Switch A 优选从 Switch B 学到的 1.1.5.0/24 网段的路由。

2. 组网图

图1-7 配置 RIP 接口附加度量值



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 Switch A。

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

配置 Switch B。

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

配置 Switch C。

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

配置 Switch D。

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

配置 Switch E。

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

查看 Switch A 的 IP 路由表。

```
[SwitchA] display rip 1 database
   1.0.0.0/8, cost 0, ClassfulSumm
       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 Switch B（IP 地址为 1.1.1.2）和 Switch C（IP 地址为 1.1.2.2），cost 值都是 2；到达网段 1.1.4.0/24 的下一跳是 Switch B，cost 为 1。

(3) 配置 RIP 接口附加度量值

在 Switch A 上配置接口 Vlan-interface200 的接口附加度量值为 3。

```
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] rip metricin 3
[SwitchA-Vlan-interface200] display rip 1 database
   1.0.0.0/8, cost 0, ClassfulSumm
       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 Switch B（IP 地址为 1.1.1.2），cost 值为 2。

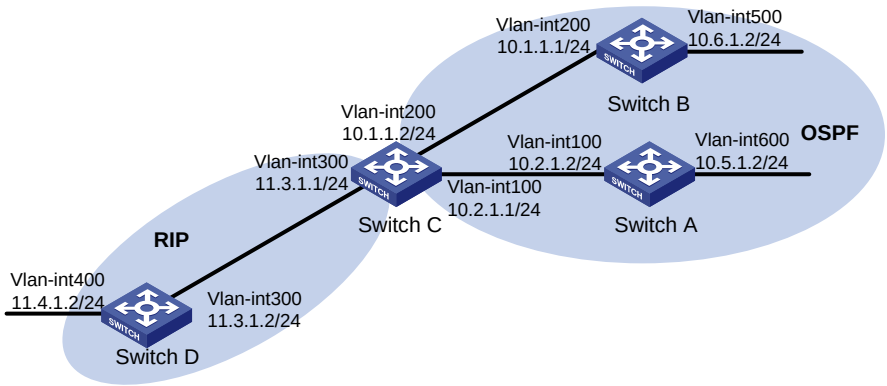
1.9.4 配置RIP发布聚合路由

1. 组网需求

- Switch A、Switch B 运行 OSPF，Switch D 运行 RIP，Switch C 同时运行 OSPF 和 RIP。
- 在 Switch C 上配置 RIP 进程引入 OSPF 路由，使 Switch 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 Switch D 的路由表规模，在 Switch 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1-8 配置 RIP 发布聚合路由



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 Switch A。

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

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network 10.6.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf-1] area 0
```

```
[SwitchC-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

(3) 配置 RIP 基本功能

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] rip 1
[SwitchC-rip-1] network 11.3.1.0
[SwitchC-rip-1] version 2
[SwitchC-rip-1] undo summary
```

配置 Switch D。

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

在 Switch 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。

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

查看 Switch D 的路由表信息。

```
[SwitchD] display ip routing-table
Routing Tables: Public
         Destinations : 10       Routes : 10

Destination/Mask    Proto  Pre  Cost         NextHop         Interface

10.1.1.0/24         RIP    100  1            11.3.1.1        Vlan300
10.2.1.0/24         RIP    100  1            11.3.1.1        Vlan300
10.5.1.0/24         RIP    100  1            11.3.1.1        Vlan300
10.6.1.0/24         RIP    100  1            11.3.1.1        Vlan300
11.3.1.0/24         Direct 0    0            11.3.1.2        Vlan300
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        Vlan400
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
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
```

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

```
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] rip summary-address 10.0.0.0 8
```

查看 Switch D 的路由表信息。

```
[SwitchD] display ip routing-table
Routing Tables: Public
         Destinations : 7        Routes : 7
```

```
Destination/Mask    Proto   Pre  Cost        NextHop         Interface

10.0.0.0/8          RIP     100  1           11.3.1.1        Vlan300
11.3.1.0/24         Direct  0    0           11.3.1.2        Vlan300
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        Vlan400
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
127.0.0.1/32        Direct  0    0           127.0.0.1       InLoop0
```

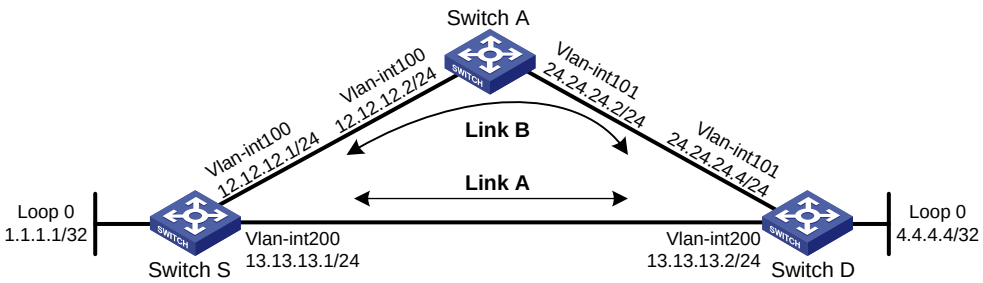
1.9.5 配置RIP快速重路由

1. 组网需求

如 图 1-9所示，Switch S 、Switch A和Switch D通过RIPv2 协议实现网络互连。要求当Switch S和Switch 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 配置 RIP 快速重路由



3. 配置步骤

(1) 配置各交换机接口的 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 Switch S、Switch A 和 Switch 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 Switch S。

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 2.2.2.2
[SwitchS] ip ip-prefix abc index 10 permit 4.4.4.4 32
[SwitchS] route-policy frr permit node 10
[SwitchS-route-policy] if-match ip-prefix abc
[SwitchS-route-policy] apply fast-reroute backup-interface vlan-interface 100
backup-nexthop 12.12.12.2
[SwitchS-route-policy] quit
[SwitchS] rip 1
[SwitchS-rip-1] fast-reroute route-policy frr
```

```
[SwitchS-rip-1] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] bfd echo-source-ip 3.3.3.3
[SwitchD] ip ip-prefix abc index 10 permit 1.1.1.1 32
[SwitchD] route-policy frr permit node 10
[SwitchD-route-policy] if-match ip-prefix abc
[SwitchD-route-policy]  apply  fast-reroute  backup-interface  vlan-interface  101
backup-nexthop 24.24.24.2
[SwitchD-route-policy] quit
[SwitchD] rip 1
[SwitchD-rip-1] fast-reroute route-policy frr
[SwitchD-rip-1] quit
```

(3) 检验配置效果

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

```
[SwitchS] display ip routing-table 4.4.4.4 verbose
Routing Table : Public
Summary Count : 1

  Destination: 4.4.4.4/32
     Protocol: RIP             Process ID: 1
   Preference: 100                   Cost: 1
 IpPrecedence:                    QosLcId:
      NextHop: 13.13.13.2       Interface: vlan200
    BkNextHop: 12.12.12.2     BkInterface: vlan100
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv             Age: 00h01m27s
          Tag: 0
```

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

```
[SwitchD] display ip routing-table 1.1.1.1 verbose
Routing Table : Public
Summary Count : 1

  Destination: 1.1.1.1/32
     Protocol: RIP             Process ID: 1
   Preference: 100                   Cost: 1
 IpPrecedence:                    QosLcId:
      NextHop: 13.13.13.1       Interface: vlan200
    BkNextHop: 24.24.24.2     BkInterface: vlan101
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv             Age: 00h01m27s
          Tag: 0
```

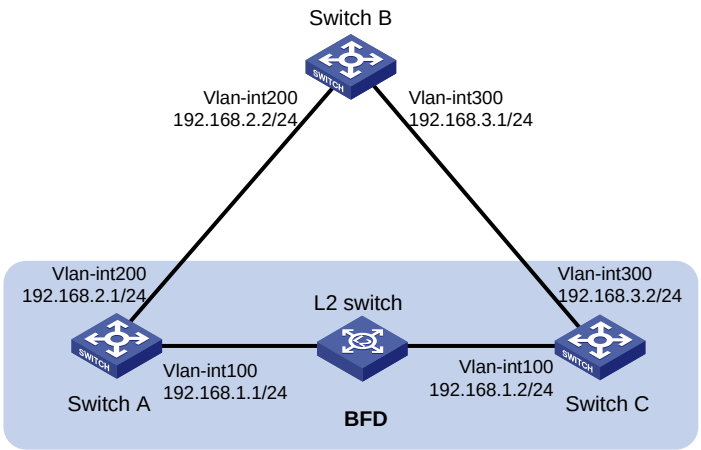
1.9.6 配置RIP运行BFD（echo报文单跳检测）

1. 组网需求

- Switch A、Switch C通过二层交换机互连，它们的接口Vlan-interface100都运行RIP进程1。并且 Switch A 的接口 Vlan-interface100 上还使能了 BFD 检测功能。
- Switch A 通过 Switch B 与 Switch C 互连，Switch A 的接口 Vlan-interface200 运行 RIP 进程 2。Switch C 的接口 Vlan-interface300、Switch B 的接口 Vlan-interface200 和 Vlan-interface300 上都运行 RIP 进程 1。
- Switch C 上配置静态路由，并且将静态路由引入在 RIP 进程中，使 Switch C 有路由发送至 Switch A。Switch A 上学习到 Switch C 发送的静态路由，出接口为与二层交换机相连的接口。
- 在 Switch 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 Switch C 的邻居，并删除从 Switch C 学习的路由。Switch A 上学习到 Switch C 上发送的静态路由，出接口为与 Switch B 相连的接口。

2. 组网图

图1-10 配置 RIP 运行 BFD 特性组网图（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

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] rip 1
[SwitchA-rip-1] version 2
[SwitchA-rip-1] undo summary
[SwitchA-rip-1] network 192.168.1.0
[SwitchA-rip-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] rip bfd enable
[SwitchA-Vlan-interface100] quit
[SwitchA] rip 2
```

```
[SwitchA-rip-2] version 2
[SwitchA-rip-2] undo summary
[SwitchA-rip-2] network 192.168.2.0
[SwitchA-rip-2] quit
```

配置 Switch B。

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

配置 Switch C。

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

(3) 配置 BFD 参数

配置 Switch A。

```
[SwitchA] bfd session init-mode active
[SwitchA] bfd echo-source-ip 11.11.11.11
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] bfd min-transmit-interval 500
[SwitchA-Vlan-interface100] bfd min-receive-interval 500
[SwitchA-Vlan-interface100] bfd detect-multiplier 7
[SwitchA-Vlan-interface100] quit
[SwitchA] quit
```

(4) Switch C 配置静态路由

```
[SwitchC] ip route-static 100.1.1.1 24 null 0
```

(5) 检查配置结果

显示 Switch A 的 BFD 信息。

```
<SwitchA> display bfd session
 Total Session Num: 1            Init Mode: Active
 Session Working Under Echo Mode:
 LD          SourceAddr      DestAddr        State   Holdtime    Interface
 5           192.168.1.1     192.168.1.2     Up      2000ms      Vlan100
```

显示 Switch A 上学到的路由 100.1.1.0/24。

```
<SwitchA> display ip routing-table 100.1.1.0 24 verbose
Routing Table : Public
Summary Count : 2
  Destination: 100.1.1.0/24
     Protocol: RIP             Process ID: 1
```

```
   Preference: 100                  Cost: 1
 IpPrecedence:                    QosLcId:
      NextHop: 192.168.1.2      Interface: vlan-interface 100
    BkNextHop: 0.0.0.0         BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 192.168.1.2
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv             Age: 00h00m47s
          Tag: 0
  Destination: 100.1.1.0/24
     Protocol: RIP             Process ID: 2
   Preference: 100                  Cost: 2
 IpPrecedence:                    QosLcId:
      NextHop: 192.168.2.2      Interface: vlan-interface 200
    BkNextHop: 0.0.0.0         BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 192.168.2.2
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Inactive Adv           Age: 00h12m50s
        Tag: 0
```

当 Switch C 和二层交换机之间的链路发生故障时， Switch A 能够快速感知 Switch C 的变化。

#查看 Switch A 的 BFD 信息，Switch A 已取消与 Switch C 的邻居关系，没有任何输出信息。

```
<SwitchA> display bfd session
```

显示 Switch A 的 RIP 进程 1 的路由信息。

查看 Switch A 的 RIP 进程 1 的路由信息，没有从 Switch C 学到的 RIP 路由。

```
<SwitchA> display rip 1 route
 Route Flags: R - RIP, T - TRIP
              P - Permanent, A - Aging, S - Suppressed, G - Garbage-collect
 ------------------------------------------------------------------------------
```

显示 Switch A 上学到的路由 100.1.1.0/24。

```
<SwitchA> display ip routing-table 100.1.1.0 24 verbose
Routing Table : Public
Summary Count : 1

  Destination: 100.1.1.0/24
     Protocol: RIP             Process ID: 2
   Preference: 100                  Cost: 2
 IpPrecedence:                    QosLcId:
      NextHop: 192.168.2.2      Interface: vlan-interface 200
    BkNextHop: 0.0.0.0         BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 192.168.2.2
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv             Age: 00h18m40s
          Tag: 0
```

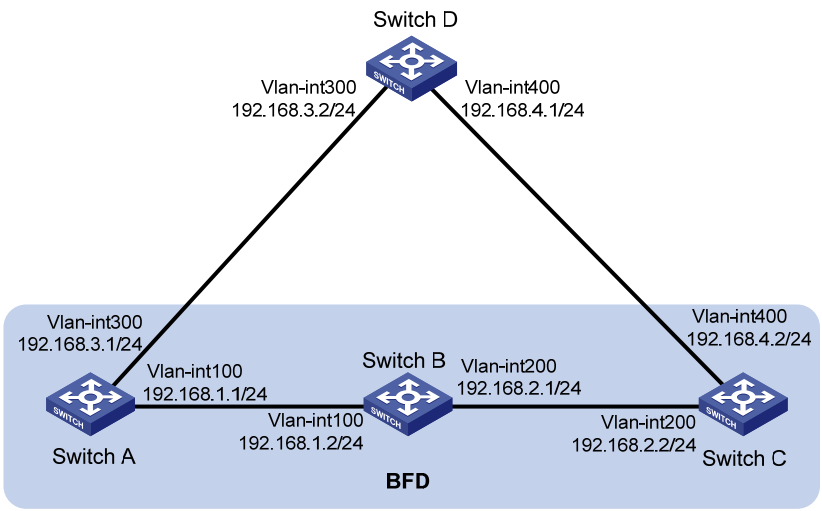
1.9.7 配置RIP运行BFD（control报文双向检测）

1. 组网需求

- Switch A、Switch B 与 Switch C 互连。Switch A 的接口 Vlan-interface100、Switch C 的接口 Vlan-interface200、Switch B 的接口 Vlan-interface100 和 Vlan-interface200 上都运行 RIP 进程 1。
- 分别在 Switch A 和 Switch C 上配置到达对端的静态路由，并在 Switch A 的接口 Vlan-interface100 和 Switch C 的接口 Vlan-interface200 上使能 BFD 检测功能。
- Switch A 通过 Switch D 与 Switch C 互连。Switch A 的接口 Vlan-interface300 运行 RIP 进程 2。Switch C 的接口 Vlan-interface400、Switch D 的接口 Vlan-interface300 和 Vlan-interface400 上运行 RIP 进程 1。
- 为使 Switch A 与 Switch C 互有路由发送，在 Switch A 与 Switch C 上将到达对端的静态路由引入 RIP 协议中。Switch A 与 Switch C 之间建立 BFD 会话。Switch A 上学习到 Switch C 发送的静态路由，出接口为与 Switch B 连接的接口。
- 在 Switch B 与 Switch 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 Switch C 的邻居，并删除从 Switch C 学习的路由。Switch A 上学习到 Switch C 发送的静态路由，出接口为与 Switch D 连接的接口。

2. 组网图

图1-11 配置 RIP 运行 BFD 特性组网图（control 报文双向检测）



3. 配置步骤

(1) 配置各接口的 IP 地址（略）

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

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] rip 1
[SwitchA-rip-1] version 2
[SwitchA-rip-1] undo summary
```

```
[SwitchA-rip-1] network 192.168.1.0
[SwitchA-rip-1] peer 192.168.2.2
[SwitchA-rip-1] undo validate-source-address
[SwitchA-rip-1] import-route static
[SwitchA-rip-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] rip bfd enable
[SwitchA-Vlan-interface100] quit
[SwitchA] rip 2
[SwitchA-rip-2] version 2
[SwitchA-rip-2] undo summary
[SwitchA-rip-2] network 192.168.3.0
[SwitchA-rip-2] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] rip 1
[SwitchC-rip-1] version 2
[SwitchC-rip-1] undo summary

[SwitchC-rip-1] network 192.168.2.0
[SwitchC-rip-1] network 192.168.4.0
[SwitchC-rip-1] peer 192.168.1.1
[SwitchC-rip-1] undo validate-source-address
[SwitchC-rip-1] import-route static
[SwitchC-rip-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] rip bfd enable
[SwitchC-Vlan-interface200] quit
```

配置 Switch D。

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

(3) 配置 BFD 参数

配置 Switch A。

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] bfd min-transmit-interval 500
[SwitchA-Vlan-interface100] bfd min-receive-interval 500
[SwitchA-Vlan-interface100] bfd detect-multiplier 7
[SwitchA-Vlan-interface100] quit
```

配置 Switch C。

```
[SwitchC] bfd session init-mode active
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] bfd min-transmit-interval 500
```

```
[SwitchC-Vlan-interface200] bfd min-receive-interval 500
[SwitchC-Vlan-interface200] bfd detect-multiplier 7
[SwitchC-Vlan-interface200] quit
```

(4) 配置静态路由

配置 Switch A。

```
[SwitchA] ip route-static 192.168.2.0 24 vlan-interface 100 192.168.1.2
[SwitchA] ip route-static 101.1.1.0 24 null 0
[SwitchA] quit
```

配置 Switch C。

```
[SwitchC] ip route-static 192.168.1.0 24 vlan-interface 200 192.168.2.1
[SwitchC] ip route-static 100.1.1.0 24 null 0
```

注意

出接口为 Null0 的静态路由，目的 IP 地址不能为直连网段的 IP 地址。

(5) 检查配置结果

显示 Switch A 的 BFD 信息。

```
<SwitchA> display bfd session
Total Session Num: 1            Init Mode: Active
 Session Working Under Ctrl Mode:
 LD/RD         SourceAddr      DestAddr        State Holdtime Interface
 6/3           192.168.1.1     192.168.2.2     Up    1700ms   vlan100
```

显示 Switch A 上学到的路由 100.1.1.0/24。

```
<SwitchA> display ip routing-table 100.1.1.0 24 verbose
Routing Table : Public
Summary Count : 2
  Destination: 100.1.1.0/24
     Protocol: RIP             Process ID: 1
   Preference: 100                   Cost: 1
 IpPrecedence:                   QosLcId:
      NextHop: 192.168.2.2      Interface: vlan-interface 100
    BkNextHop: 0.0.0.0         BkInterface:
  RelyNextHop: 192.168.1.2       Neighbor : 192.168.2.2
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv GotQ        Age: 00h00m47s
          Tag: 0
  Destination: 100.1.1.0/24
     Protocol: RIP             Process ID: 2
   Preference: 100                   Cost: 2
 IpPrecedence:                   QosLcId:
      NextHop: 192.168.3.2      Interface: vlan-interface 300
    BkNextHop: 0.0.0.0        BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 192.168.3.2
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
```

```
        State: Inactive Adv           Age: 00h12m50s
        Tag: 0
```

当 Switch B 和 Switch C 之间的链路发生故障时 Switch A 能够快速检测到链路的变化。

查看 Switch A 的 BFD 信息，Switch A 已取消与 Switch C 的邻居关系，没有任何输出信息。

```
<SwitchA> display bfd session
```

显示 Switch A 的 RIP 进程 1 的路由信息。

查看 Switch A 的 RIP 进程 1 的路由信息，没有从 Switch C 学到的 RIP 路由。

```
<SwitchA> display rip 1 route
 Route Flags: R - RIP, T - TRIP
              P - Permanent, A - Aging, S - Suppressed, G - Garbage-collect
 ----------------------------------------------------------------------------
```

显示 Switch A 上学到的路由 100.1.1.0/24。

```
<SwitchA> display ip routing-table 100.1.1.0 24 verbose
Routing Table : Public
Summary Count : 1

  Destination: 100.1.1.0/24
     Protocol: RIP             Process ID: 2
   Preference: 100                   Cost: 2
 IpPrecedence:                   QosLcId:
      NextHop: 192.168.3.2      Interface: vlan-interface 300
    BkNextHop: 0.0.0.0        BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 192.168.3.2
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv             Age: 00h18m40s
          Tag: 0
```

1.10 常见配置错误举例

1.10.1 收不到邻居的RIP更新报文

1. 故障现象

在链路正常的情况下收不到邻居的 RIP 更新报文。

2. 故障分析

RIP 启动后，必须使用 **network** 命令使能相应的接口。如果对接口的工作状态单独进行了配置，应确认没有抑制相关接口或禁止其收发 RIP 报文。

如果在对端路由器上配置的是组播方式发送 RIP 报文，在本地路由器上也应该配置为组播方式。

3. 故障排除

(1) 执行 **display current-configuration** 命令，检查 RIP 的配置。

(2) 执行 **display rip** 命令，检查是否抑制了相关 RIP 接口。

1.10.2 RIP网络发生路由振荡

1. 故障现象

在链路正常的情况下，运行 RIP 的网络发生路由振荡，查看路由表时发现部分路由时有时无。

2. 故障分析

在 RIP 网络中，应确保全网定时器的配置一致，且各定时器之间的关系合理，如 Timeout 定时器的值应大于 Update 定时器的值。

3. 故障排除

(1) 使用 **display rip** 命令查看 RIP 定时器的配置。

(2) 使用 **timers** 将全网的定时器配置一致。

目 录

1 OSPF配置 1-1

1.1 OSPF简介 1-1

1.1.1 OSPF的基本概念 1-2

1.1.2 OSPF区域 1-3

1.1.3 路由器的类型 1-7

1.1.4 OSPF的网络类型 1-8

1.1.5 DR/BDR 1-9

1.1.6 OSPF的协议报文 1-10

1.1.7 系统支持的OSPF特性 1-18

1.1.8 协议规范 1-20

1.2 OSPF配置任务简介 1-20

1.3 使能OSPF功能 1-22

1.3.1 配置准备 1-22

1.3.2 使能OSPF功能 1-22

1.4 配置OSPF区域 1-24

1.4.1 配置准备 1-24

1.4.2 配置Stub区域 1-24

1.4.3 配置NSSA区域 1-25

1.4.4 配置虚连接 1-25

1.5 配置OSPF的网络类型 1-26

1.5.1 配置准备 1-26

1.5.2 配置OSPF接口网络类型为广播 1-26

1.5.3 配置OSPF接口网络类型为NBMA 1-27

1.5.4 配置OSPF接口网络类型为P2MP 1-28

1.5.5 配置OSPF接口网络类型为P2P 1-28

1.6 配置OSPF的路由信息控制 1-29

1.6.1 配置准备 1-29

1.6.2 配置OSPF路由聚合 1-29

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 1-30

1.6.4 配置过滤Type-3 LSA 1-31

1.6.5 配置OSPF接口的开销值 1-31

1.6.6 配置OSPF支持的路由最大数目 1-32

1.6.7 配置OSPF最大等价路由条数 1-32

1.6.8 配置OSPF协议的优先级 ····· 1-33
1.6.9 配置OSPF引入外部路由 ····· 1-33
1.6.10 配置发布一条主机路由 ····· 1-35
1.7 调整和优化OSPF网络 ····· 1-35
1.7.1 配置准备 ····· 1-35
1.7.2 配置OSPF报文定时器 ····· 1-36
1.7.3 配置接口传送LSA的延迟时间 ····· 1-37
1.7.4 配置SPF计算时间间隔 ····· 1-37
1.7.5 配置LSA重复到达的最小时间间隔 ····· 1-37
1.7.6 配置LSA重新生成的时间间隔 ····· 1-38
1.7.7 禁止接口收发OSPF报文 ····· 1-38
1.7.8 配置Stub路由器 ····· 1-39
1.7.9 配置OSPF验证 ····· 1-40
1.7.10 配置DD报文中的MTU ····· 1-40
1.7.11 配置LSDB中External LSA的最大数量 ····· 1-41
1.7.12 配置兼容RFC 1583 的外部路由选择规则 ····· 1-41
1.7.13 配置邻居状态变化的输出开关 ····· 1-42
1.7.14 配置OSPF网管功能 ····· 1-42
1.7.15 使能日志功能 ····· 1-43
1.7.16 使能Opaque LSA发布接收能力 ····· 1-43
1.7.17 配置OSPF优先接收并处理Hello报文 ····· 1-43
1.7.18 配置LSU报文的发送速率 ····· 1-44
1.7.19 配置OSPF报文的DSCP优先级 ····· 1-44
1.7.20 配置OSPF ISPF ····· 1-44
1.8 配置OSPF快速重路由 ····· 1-45
1.9 配置OSPF GR ····· 1-47
1.9.1 配置GR Restarter ····· 1-47
1.9.2 配置GR Helper ····· 1-48
1.9.3 以GR方式重启OSPF进程 ····· 1-49
1.10 配置OSPF与BFD联动 ····· 1-49
1.10.1 control报文双向检测 ····· 1-49
1.10.2 echo报文单跳检测 ····· 1-50
1.11 OSPF显示和维护 ····· 1-50
1.12 典型配置举例 ····· 1-51
1.12.1 配置OSPF基本功能 ····· 1-51
1.12.2 配置OSPF引入自治系统外部路由 ····· 1-55

1.12.3 配置OSPF发布聚合路由 ····· 1-56
1.12.4 配置OSPF的Stub区域 ····· 1-59
1.12.5 配置OSPF的NSSA区域 ····· 1-62
1.12.6 配置OSPF的DR选择 ····· 1-64
1.12.7 配置OSPF虚连接 ····· 1-68
1.12.8 OSPF GR配置举例 ····· 1-70
1.12.9 配置路由过滤 ····· 1-72
1.12.10 OSPF快速重路由配置举例 ····· 1-75
1.12.11 配置OSPF与BFD联动 ····· 1-77
1.13 常见配置错误举例 ····· 1-80
1.13.1 OSPF邻居无法建立 ····· 1-80
1.13.2 OSPF路由信息不正确 ····· 1-80

1 OSPF配置

说明

本文所指的路由器代表运行了路由协议的三层设备。

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（RFC 2328）。

说明

本章若没有特别说明，下文中所提到的 OSPF 均指 OSPFv2。

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，区域间传送的路由信息被进一步抽象，从而减少了占用的网络带宽。

- 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。

1.1.1 OSPF的基本概念

1. 自治系统（Autonomous System）

一组使用相同路由协议交换路由信息的路由器，缩写为 AS。

2. OSPF路由的计算过程

同一个区域内，OSPF 协议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（Link State Advertisement，链路状态通告），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（Link State Database，链路状态数据库）。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。

3. 路由器ID号

一台运行 OSPF 协议路由器，每一个 OSPF 进程必须存在自己的 Router ID（路由器 ID）。Router ID 是一个 32 比特无符号整数，可以在一个自治系统中唯一的标识一台路由器。

4. OSPF的协议报文

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。内容包括一些定时器的数值、DR（Designated Router，指定路由器）、BDR（Backup Designated Router，备份指定路由器）以及自己已知的邻居。
- DD（Database Description，数据库描述）报文：描述了本地 LSDB 中每一条 LSA 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。内容包括所需要的 LSA 的摘要。
- 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。内容是需要确认的 LSA 的 Header（一个报文可对多个 LSA 进行确认）。

5. LSA的类型

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（Type1）：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（Type2）：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（Type3）：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（Type4）：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（Type5）：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（Type7）：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：是一个被提议的 LSA 类别，由标准的 LSA 头部后面跟随特殊应用的信息组成，可以直接由 OSPF 协议使用，或者由其它应用分发信息到整个 OSPF 域间接使用。Opaque LSA 分为 Type9、Type10、Type11 三种类型，泛洪区域不同；其中，Type9 的 Opaque LSA 仅在本地链路范围进行泛洪，Type10 的 Opaque LSA 仅在本地区域范围进行泛洪，Type11 的 LSA 可以在一个自治系统范围进行泛洪。

6. 邻居和邻接

在 OSPF 中，邻居（Neighbor）和邻接（Adjacency）是两个不同的概念。

OSPF 路由器启动后，便会通过 OSPF 接口向外发送 Hello 报文。收到 Hello 报文的 OSPF 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。

形成邻居关系的双方不一定都能形成邻接关系，这要根据网络类型而定。只有当双方成功交换 DD 报文，交换 LSA 并达到 LSDB 的同步之后，才形成真正意义上的邻接关系。

1.1.2 OSPF区域

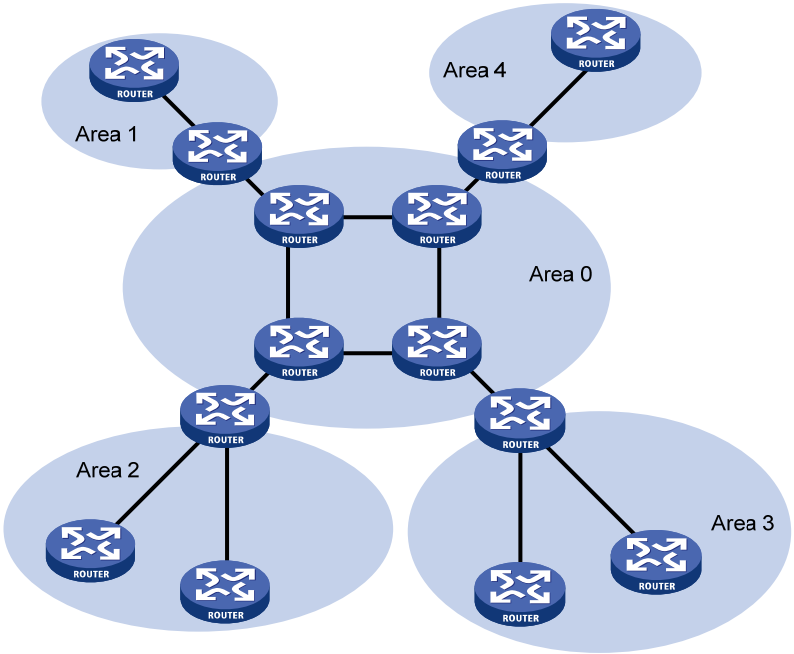
1. 区域划分

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 路由协议时，路由器数量的增多会导致 LSDB 非常庞大，占用大量的存储空间，并使得运行 SPF 算法的复杂度增加，导致 CPU 负担很重。

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（Area）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（Area ID）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（Area ID）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。

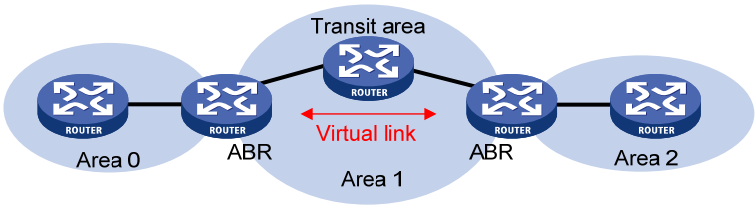
但在实际应用中，可能会因为各方面条件的限制，无法满足这个要求。这时可以通过配置 OSPF 虚连接（Virtual Link）予以解决。

(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）。

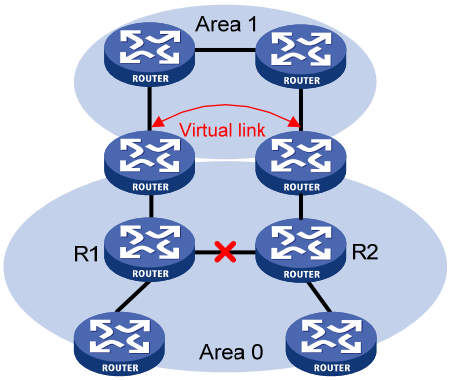
在 图 1-2 中，Area2 与骨干区域之间没有直接相连的物理链路，但可以在ABR上配置虚连接，使Area2通过一条逻辑链路与骨干区域保持连通。

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

Stub 区域是一些特定的区域，Stub 区域不允许注入 Type5 LSA，在这些区域中路由器的路由表规模以及路由信息传递的数量都会大大减少。

为了进一步减少 Stub 区域中路由器的路由表规模以及路由信息传递的数量，可以将该区域配置为 Totally Stub（完全 Stub）区域，该区域的 ABR 不会将区域间的路由信息和外部路由信息传递到本区域。

（Totally）Stub 区域是一种可选的配置属性，但并不是每个区域都符合配置的条件。通常来说，（Totally）Stub 区域位于自治系统的边界。

为保证到本自治系统的其他区域或者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，并发布给本区域中的其他非 ABR 路由器。

配置（Totally）Stub 区域时需要注意下列几点：

- 骨干区域不能配置成（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。

- （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。
- 虚连接不能穿过（Totally）Stub 区域。

4. NSSA区域

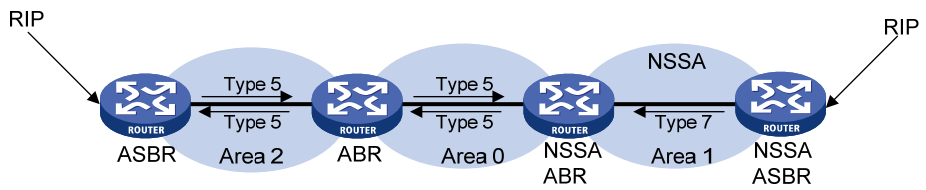
NSSA（Not-So-Stubby Area）区域是 Stub 区域的变形，与 Stub 区域有许多相似的地方。NSSA 区域也不允许 Type5 LSA 注入，但可以允许 Type7 LSA 注入。Type7 LSA 由 NSSA 区域的 ASBR 产生，在 NSSA 区域内传播。当 Type7 LSA 到达 NSSA 的 ABR 时，由 ABR 将 Type7 LSA 转换成 Type5 LSA，传播到其他区域。

如 图 1-4 所示，运行OSPF协议的自治系统包括 3 个区域：区域 1、区域 2 和区域 0，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Type7 LSA在区域 1 内传播，当Type7 LSA到达NSSA ABR后，转换成Type5 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 Type5 LSA 在 OSPF 自治系统中传播。但由于区域 1 是 NSSA 区域，所以 Type5 LSA 不会到达区域 1。

与 Stub 区域一样，虚连接也不能穿过 NSSA 区域。

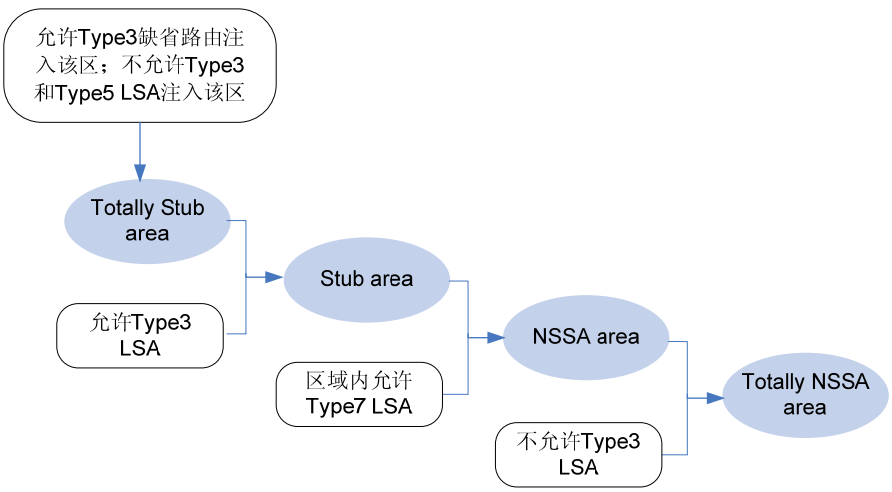
图1-4 NSSA 区域



5. 各区域特性

如 图 1-5 所示，列出了各个区域之间的区别。

图1-5 区域特性示意图



各区域特性：

- Totally Stub 区域：允许 ABR 发布的 Type3 缺省路由，不允许自治系统外部路由和区域间的路由；

- Stub 区域：和 Totally Stub 区域不同在于该区域允许区域间路由；
- NSSA 区域：和 Stub 区域不同在于该区域允许自治系统外部路由的引入，由 ASBR 发布 Type7 LSA 通告给本区域；
- Totally NSSA 区域：和 NSSA 区域不同在于该区域不允许区域间路由。

1.1.3 路由器的类型

1. 路由器的分类

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（Area Border Router）

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（骨干区域的介绍请参见下一小节）。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。

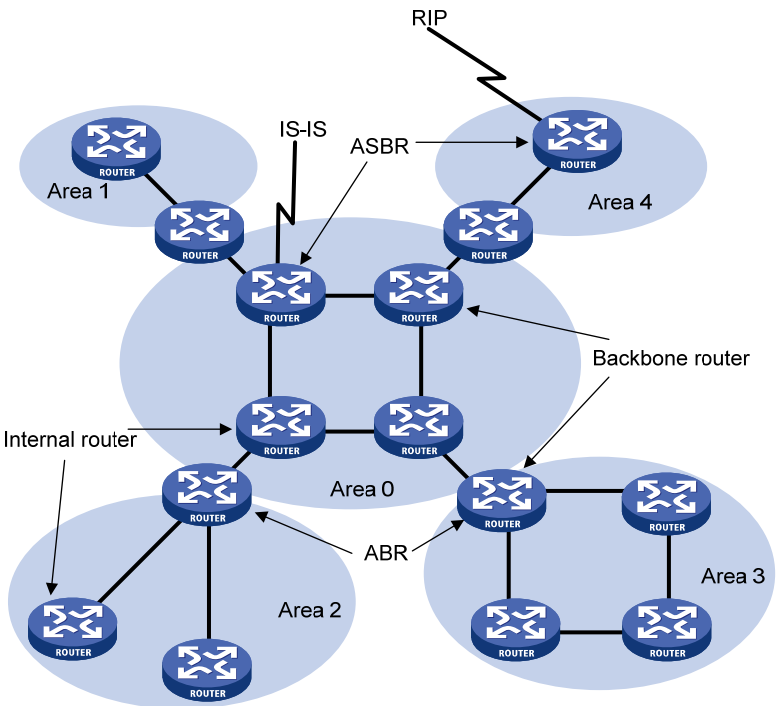
(3) 骨干路由器（Backbone Router）

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1-6 OSPF 路由器的类型



2. 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）
- 第二类外部路由（Type2 External）

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。OSPF 将引入的 AS 外部路由分为两类：Type1 和 Type2。

第一类外部路由：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：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。

1.1.4 OSPF的网络类型

1. OSPF的 4 种网络类型

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，OSPF 缺省认为网络类型是 Broadcast。在该类型的网络中，通常以组播形式（224.0.0.5：含义是 OSPF 路由器的预留 IP 组播地址；224.0.0.6：含义是 OSPF DR 的预留 IP 组播地址）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，OSPF 缺省认为网络类型是 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。点到多点必须是由其他的网络类型强制更改的。常用做法是将 NBMA 改为点到多点的网络。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，OSPF 缺省认为网络类型是 P2P。在该类型的网络中，以组播形式（224.0.0.5）发送协议报文。

2. NBMA网络的配置原则

NBMA 网络是指非广播、多点可达的网络，比较典型的有 ATM 和帧中继网络。

对于接口的网络类型为 NBMA 的网络需要进行一些特殊的配置。由于无法通过广播 Hello 报文的形式发现相邻路由器，必须手工为该接口指定相邻路由器的 IP 地址，以及该相邻路由器是否有 DR 选举权等。

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。如果部分路由器之间没有直接可达的链路时，应将接口配置成 P2MP 类型。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 类型。

NBMA 与 P2MP 网络之间的区别如下：

- NBMA 网络是指那些全连通的、非广播、多点可达网络。而 P2MP 网络，则并不需要一定是全连通的。
- 在 NBMA 网络中需要选举 DR 与 BDR，而在 P2MP 网络中没有 DR 与 BDR。
- NBMA 是一种缺省的网络类型，而 P2MP 网络必须是由其它的网络强制更改的。最常见的做法是将 NBMA 网络改为 P2MP 网络。
- 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。

1.1.5 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 协议定义了指定路由器 DR（Designated Router），所有路由器都只将信息发送给 DR，由 DR 将网络链路状态发送出去。

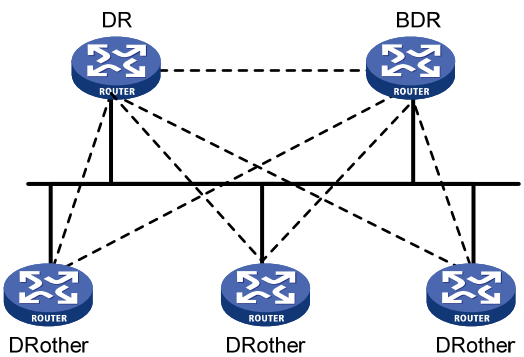
如果 DR 由于某种故障而失效，则网络中的路由器必须重新选举 DR，再与新的 DR 同步。这需要较长的时间，在这段时间内，路由的计算是不正确的。为了能够缩短这个过程，OSPF 提出了 BDR（Backup Designated Router，备份指定路由器）的概念。

BDR 是对 DR 的一个备份，在选举 DR 的同时也选举出 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 DR。由于不需要重新选举，并且邻接关系之前已建立，所以这个过程是非常短暂的。当然这时还需要再重新选举出一个新的 BDR，虽然需要较长的时间，但并不会影响路由的计算。

运行 OSPF 进程的网络中，既不是 DR 也不是 BDR 的路由器为 DR Other。DR Other 仅与 DR 和 BDR 之间建立邻接关系，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 1-7 所示，用实线代表以太网物理连接，虚线代表建立的邻接关系。可以看到，采用DR/BDR 机制后，5 台路由器之间只需要建立 7 个邻接关系就可以了。

图1-7 DR 和 BDR 示意图



2. DR/BDR选举过程

DR 和 BDR 是由同一网段中所有的路由器根据路由器优先级、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，发给网段上的每台运行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。如果一台路由器的优先级为0，则它不会被选举为DR或BDR。

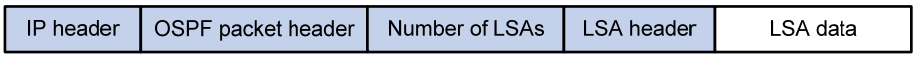
需要注意的是：

- 只有在广播或NBMA类型接口才会选举DR，在点到点或点到多点类型的接口上不需要选举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。
- 路由器的优先级可以影响DR/BDR的选举过程，但是当DR/BDR已经选举完毕，就算一台具有更高优先级的路由器变为有效，也不会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。

1.1.6 OSPF的协议报文

OSPF报文直接封装为IP报文协议报文，协议号为89。一个比较完整的OSPF报文（以LSU报文为例）结构如 图 1-8 所示。

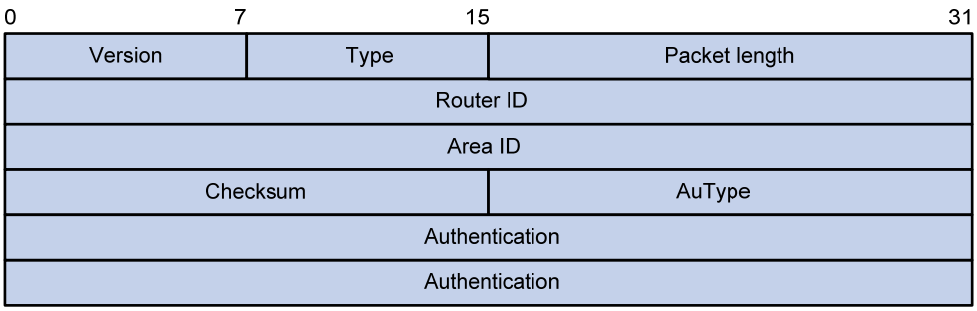
图1-8 OSPF 报文结构



1. OSPF报文头

OSPF有五种报文类型，它们有相同的报文头。如 图 1-9 所示。

图1-9 OSPF 报文头格式



主要字段的解释如下：

- Version：OSPF的版本号。对于OSPFv2来说，其值为2。
- Type：OSPF报文的类型。数值从1到5，分别对应Hello报文、DD报文、LSR报文、LSU报文和LSAck报文。
- Packet length：OSPF报文的总长度，包括报文头在内，单位为字节。

- Router ID：始发该 LSA 的路由器的 ID。
- Area ID：始发 LSA 的路由器所在的区域 ID。
- Checksum：对整个报文的校验和。
- AuType：验证类型。可分为不验证、简单（明文）口令验证和 MD5 验证，其值分别为 0、1、2。
- Authentication：其数值根据验证类型而定。当验证类型为 0 时未作定义，为 1 时此字段为密码信息，类型为 2 时此字段包括 Key ID、MD5 验证数据长度和序列号的信息。

说明

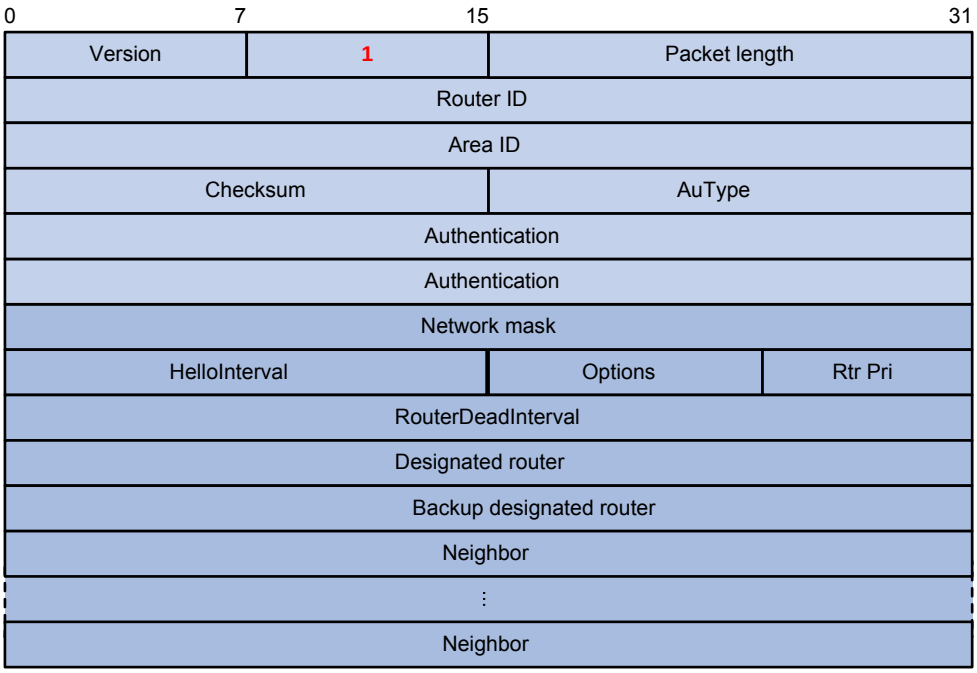
MD5 验证数据添加在 OSPF 报文后面，不包含在 Authenticaiton 字段中。

---

2. Hello报文（Hello Packet）

最常用的一种报文，周期性的发送给邻居路由器用来维持邻居关系以及DR/BDR的选举，内容包括一些定时器的数值、DR、BDR以及自己已知的邻居。Hello报文格式如 图 1-10 所示。

图1-10 Hello 报文格式



主要字段解释如下：

- Network mask：发送 Hello 报文的接口所在网络的掩码，如果相邻两台路由器的网络掩码不同，则不能建立邻居关系。
- HelloInterval：发送 Hello 报文的时间间隔。如果相邻两台路由器的 Hello 间隔时间不同，则不能建立邻居关系。
- Rtr Pri：路由器优先级。如果设置为 0，则该路由器接口不能成为 DR/BDR。
- RouterDeadInterval：失效时间。如果在此时间内未收到邻居发来的 Hello 报文，则认为邻居失效。如果相邻两台路由器的失效时间不同，则不能建立邻居关系。

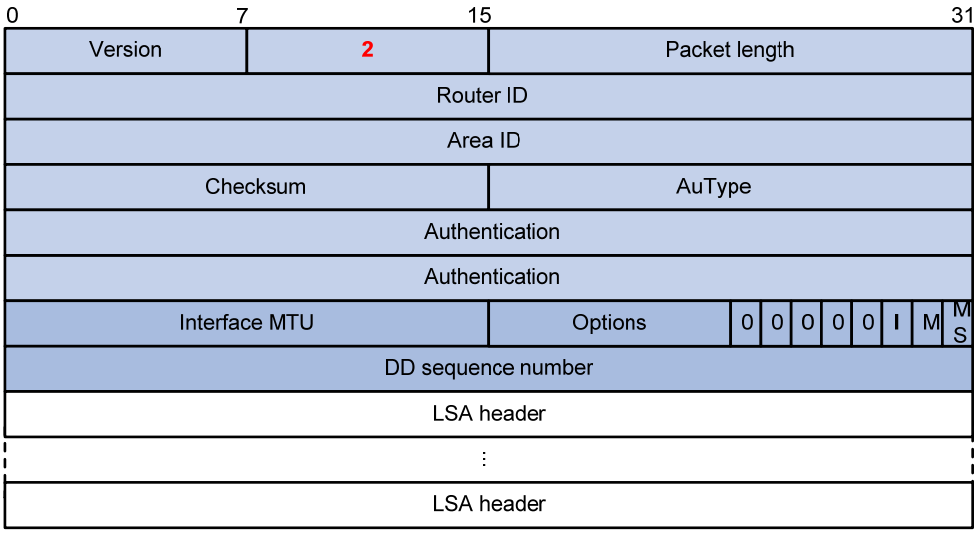
- Designated router：指定路由器的接口的 IP 地址。
- Backup designated router：备份指定路由器的接口的 IP 地址。
- Neighbor：邻居路由器的 Router ID。

3. DD报文（Database Description Packet）

两台路由器进行数据库同步时，用 DD 报文来描述自己的 LSDB，内容包括 LSDB 中每一条 LSA 的摘要（摘要是指 LSA 的 Header，通过该 Header 可以唯一标识一条 LSA）。这样做是为了减少路由器之间传递信息的量，因为 LSA 的 Header 只占一条 LSA 的整个数据量的一小部分，根据 Header，对端路由器就可以判断出是否已有这条 LSA。

DD报文格式如 图 1-11 所示。

图1-11 DD 报文格式



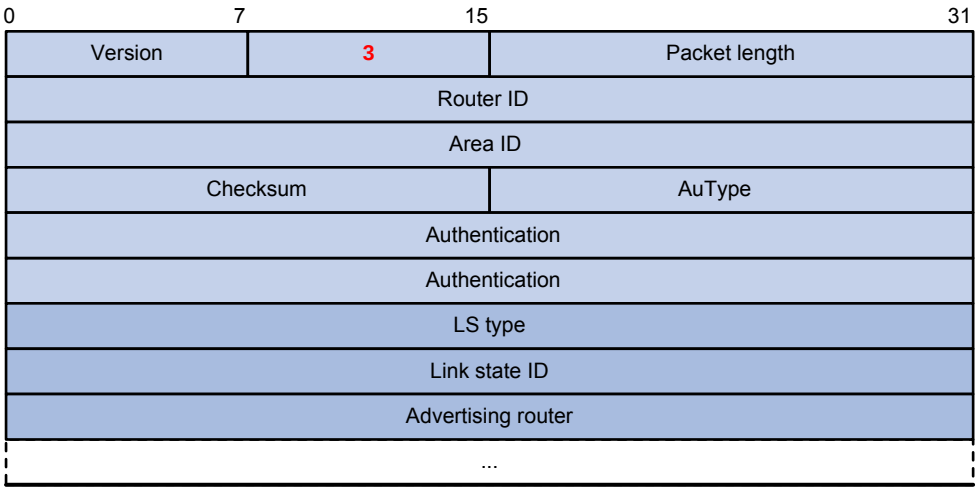
主要字段的解释如下：

- Interface MTU：在不分片的情况下，此接口最大可发出的 IP 报文长度。
- I（Initial）：当发送连续多个 DD 报文时，如果这是第一个 DD 报文，则置为 1，否则置为 0。
- M（More）：当连续发送多个 DD 报文时，如果这是最后一个 DD 报文，则置为 0。否则置为 1，表示后面还有其他的 DD 报文。
- MS（Master/Slave）：当两台 OSPF 路由器交换 DD 报文时，首先需要确定双方的主（Master）从（Slave）关系，Router ID 大的一方会成为 Master。当值为 1 时表示发送方为 Master。
- DD Sequence Number：DD 报文序列号，由 Master 方规定起始序列号，每发送一个 DD 报文序列号加 1，Slave 方使用 Master 的序列号作为确认。主从双方利用序列号来保证 DD 报文传输的可靠性和完整性。

4. LSR报文（Link State Request Packet）

两台路由器互相交换过DD报文之后，知道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。内容包括所需要的LSA的摘要。LSR报文格式如 图 1-12 所示。

图1-12 LSR 报文格式



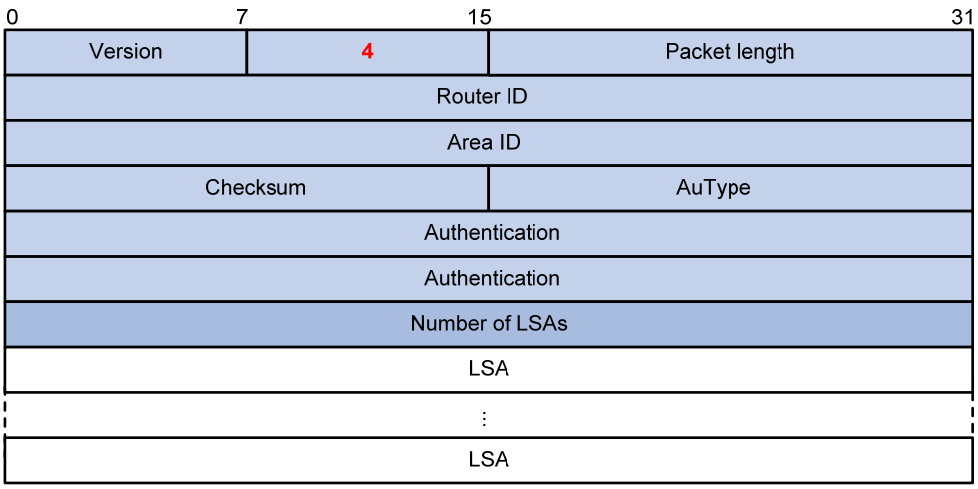
主要字段解释如下：

- LS type：LSA 的类型号。例如 Type1 表示 Router LSA。
- Link State ID：链路状态标识，根据 LSA 的类型而定。
- Advertising Router：产生此 LSA 的路由器的 Router ID。

5. LSU报文（Link State Update Packet）

LSU报文用来向对端路由器发送它所需要的LSA，内容是多条LSA（全部内容）的集合。LSU报文格式如 图 1-13 所示。

图1-13 LSU 报文格式



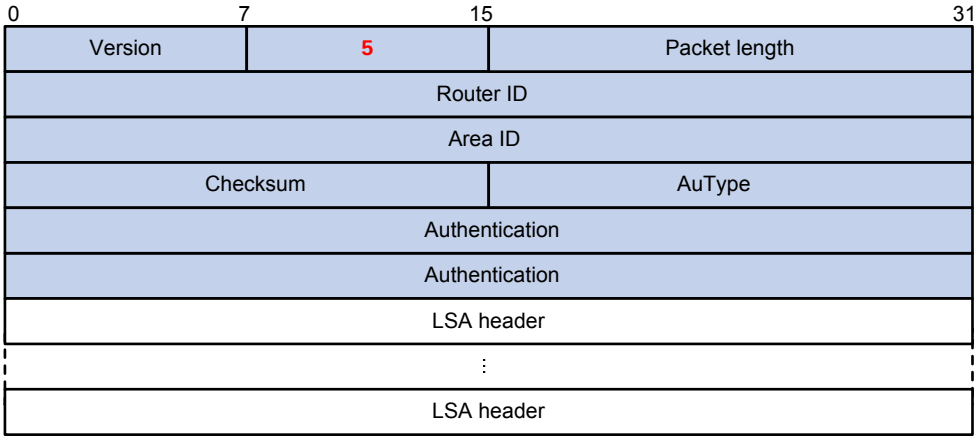
主要字段解释如下：

- Number of LSAs：该报文包含的 LSA 的数量。
- LSA：该报文包含的 LSA。

6. LSAck报文（Link State Acknowledgment Packet）

LSAck报文用来对接收到的LSU报文进行确认，内容是需要确认的LSA的Header。一个LSAck报文可对多个LSA进行确认。报文格式如 图 1-14 所示。

图1-14 LSAck 报文格式



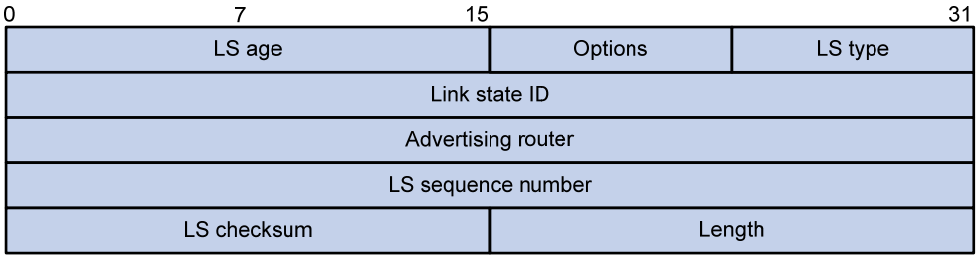
主要字段解释如下：

LSA Headers：该报文包含的 LSA 头部。

7. LSA头格式

所有的LSA都有相同的报文头，其格式如 图 1-15 所示。

图1-15 LSA 的头格式



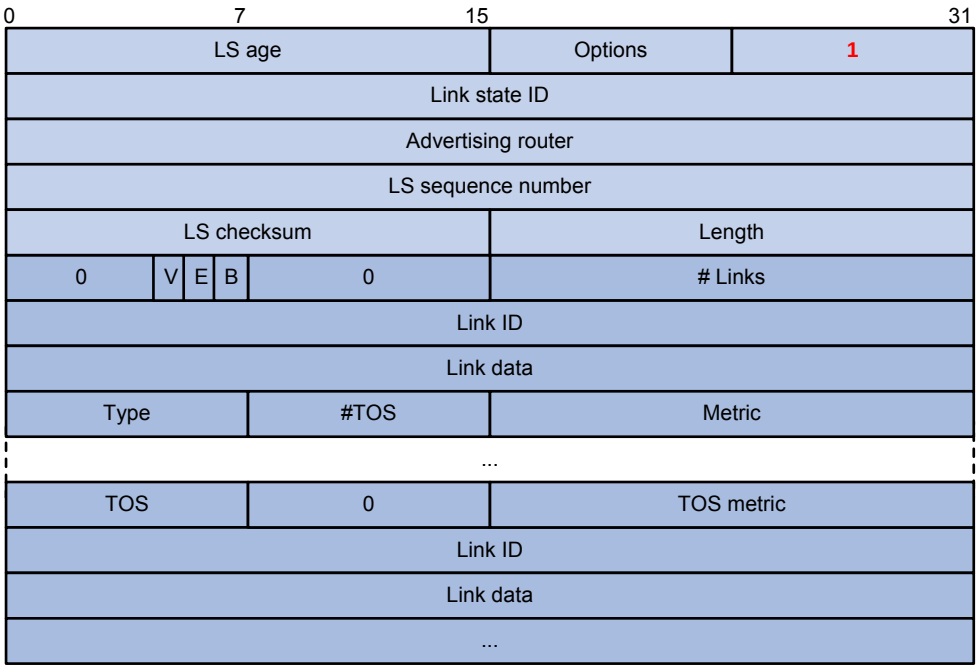
主要字段的解释如下：

- LS age：LSA 产生后所经过的时间，以秒为单位。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。
- LS type：LSA 的类型。
- Link State ID：具体数值根据 LSA 的类型而定。
- Advertising Router：始发 LSA 的路由器的 ID。
- LS sequence number：LSA 的序列号，其他路由器根据这个值可以判断哪个 LSA 是最新的。
- LS checksum：除了 LS age 字段外，关于 LSA 的全部信息的校验和。
- length：LSA 的总长度，包括 LSA Header，以字节为单位。

8. LSA类型

(1) Router LSA

图1-16 Router LSA 格式



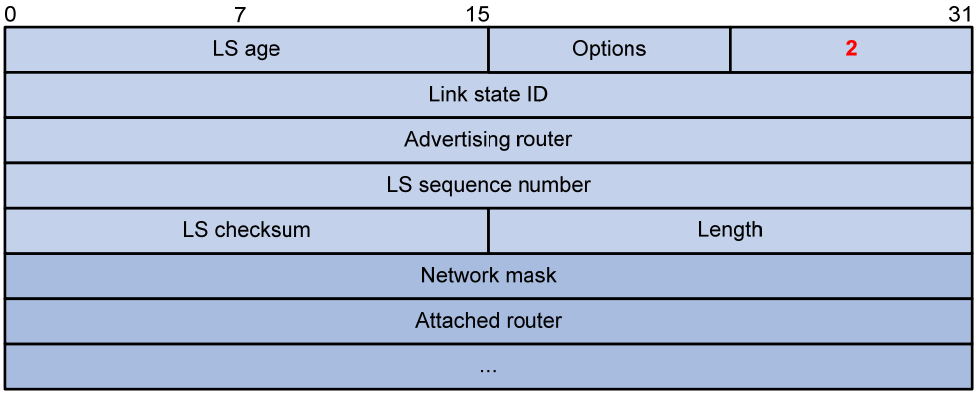
主要字段的解释如下：

- Link State ID：产生此 LSA 的路由器的 Router ID。
- V（Virtual Link）：如果产生此 LSA 的路由器是虚连接的端点，则置为 1。
- E（External）：如果产生此 LSA 的路由器是 ASBR，则置为 1。
- B（Border）：如果产生此 LSA 的路由器是 ABR，则置为 1。
- # Links：LSA 中所描述的链路信息的数量，包括路由器上处于某区域中的所有链路和接口。
- Link ID：链路标识，具体的数值根据链路类型而定。
- Link data：链路数据，具体的数值根据链路类型而定。
- Type：链路类型，取值为 1 表示通过点对点链路与另一路由器相连，取值为 2 表示连接到传送网络，取值为 3 表示连接到 Stub 网络，取值为 4 表示虚连接。
- #TOS：描述链路不同 TOS 开销的数量。如果没有 TOS 开销和当前链路关联，那么这个字段就设置为 0。RFC 2328 已经不再支持 TOS，保留该字段主要是为了兼容早期的 OSPF 版本。
- Metric：链路的开销。
- TOS：服务类型。
- TOS metric：指定服务类型的链路的开销。

(2) Network LSA

Network LSA 由广播网或 NBMA 网络中的 DR 发出，LSA 中记录了这一网段上所有路由器的 Router ID。

图1-17 Network LSA 格式



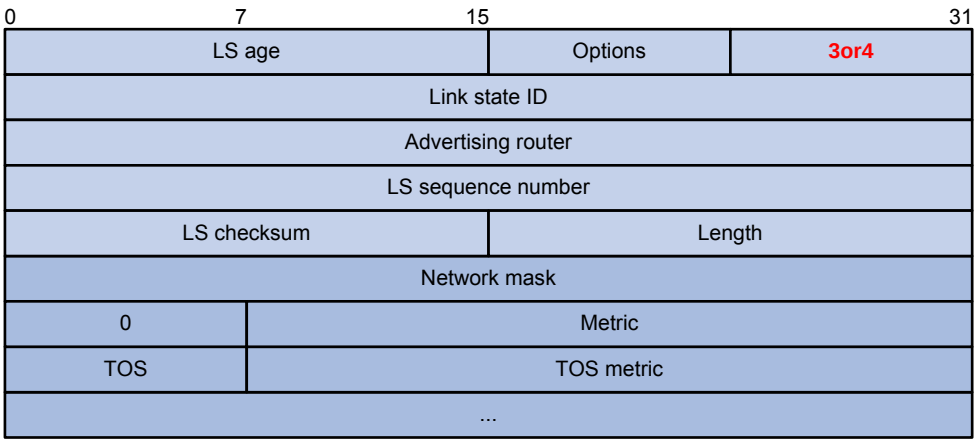
主要字段的解释如下：

- Link State ID：DR 的 IP 地址。
- Network mask：广播网或 NBMA 网络地址的掩码。
- Attached router：连接在同一个网段上的所有与 DR 形成了完全邻接关系的路由器的 Router ID，也包括 DR 自身的 Router ID。

(3) Summary LSA

Network Summary LSA（Type3 LSA）和 ASBR Summary LSA（Type4 LSA）除 Link State ID 字段有所不同外，有着相同的格式，它们都是由 ABR 产生。

图1-18 Summary LSA 格式



主要字段的解释如下：

- Link State ID：对于 Type3 LSA 来说，它是所通告的区域外的网络地址；对于 Type4 来说，它是所通告区域外的 ASBR 的 Router ID。
- Network mask：Type3 LSA 的网络地址掩码。对于 Type4 LSA 来说没有意义，设置为 0.0.0.0。
- Metric：到目的地址的路由开销。

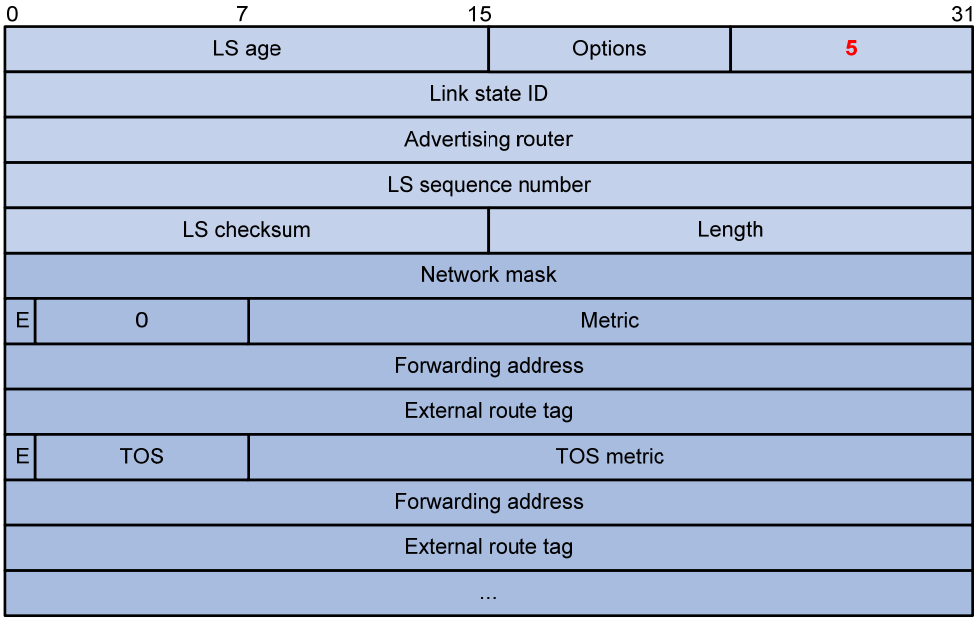
说明

Type3 的 LSA 可以用来通告缺省路由，此时 Link State ID 和 Network Mask 都设置为 0.0.0.0。

(4) AS External LSA

由 ASBR 产生，描述到 AS 外部的路由信息。

图1-19 AS External LSA 格式



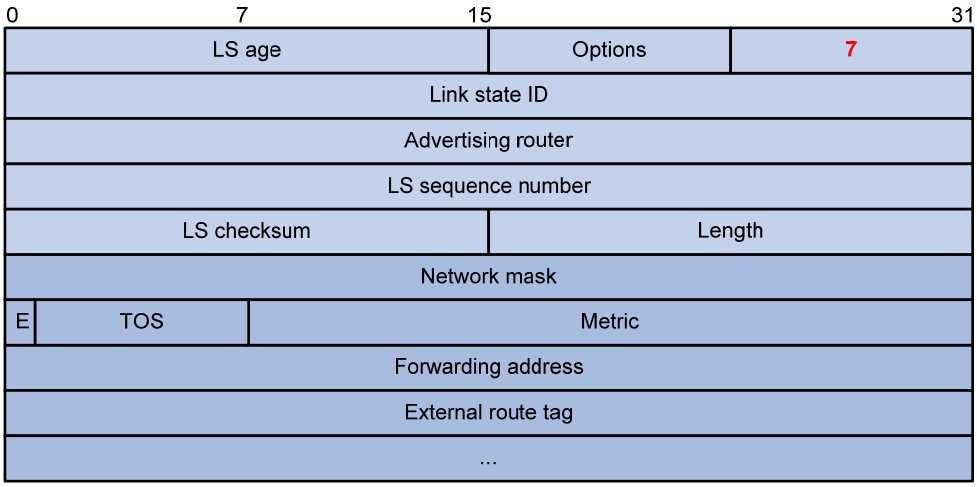
主要字段的解释如下：

- Link State ID：所要通告的其他外部 AS 的目的地址，如果通告的是一条缺省路由，那么链路状态 ID（Link State ID）和网络掩码（Network Mask）字段都将设置为 0.0.0.0。
- Network mask：所通告的目的地址的掩码。
- E（External Metric）：外部度量值的类型。如果是第 2 类外部路由就设置为 1，如果是第 1 类外部路由则设置为 0。关于外部路由类型的详细描述请参见 1.1.3 2. 路由类型部分。
- Metirc：路由开销。
- Forwarding Address：到所通告的目的地址的报文将被转发到的地址。
- External Route Tag：添加到外部路由上的标记。OSPF 本身并不使用这个字段，它可以用来对外部路由进行管理。

(5) NSSA External LSA

由NSSA区域内的ASBR产生，且只能在NSSA区域内传播。其格式与AS External LSA相同，如 图 1-20 所示。

图1-20 NSSA External LSA 格式



1.1.7 系统支持的OSPF特性

1. 多进程

OSPF 支持多进程，在同一台路由器上可以运行多个不同的 OSPF 进程，它们之间互不影响，彼此独立。不同 OSPF 进程之间的路由交互相当于不同路由协议之间的路由交互。支持多个 OSPF 进程共用一个 RID。

路由器的一个接口只能属于某一个 OSPF 进程。

2. 验证

OSPF 支持报文验证功能，只有通过验证的 OSPF 报文才能接收，否则将不能正常建立邻居关系。一个区域中所有的接口的验证类型（不支持验证、支持明文验证或者支持 MD5 密文验证）必须一致。一个网段中所有接口的验证字口令必须一致。

3. OSPF GR

GR 是一种在协议重启或主备切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。

基于 OSPF 的 GR Restarter 进行协议重启后，为了与其 OSPF 邻居重新同步链路状态数据库，它必须完成下列两项任务：

- 在不改变邻接关系的前提下，重新获取网络中的有效 OSPF 邻居信息；
- 重新获取网络链路状态数据库的内容。

在 OSPF 协议重启前，GR Restarter 与 GR Helper 协商 GR 能力，如果 GR Restarter 发生重启，在重启过程中，GR Helper 宣告与 GR Restarter 的邻接状态不变。

OSPF 协议重启完毕后，GR Restarter 会立即向邻接的 GR Helper 发送一个 OSPF GR 信号。这样，OSPF 邻居就不会复位与 GR Restarter 的邻居关系。在收到 OSPF 邻居的响应后，GR Restarter 会重新恢复与 GR Helper 的邻居关系列表。

邻居关系重新建立后，GR Restarter 与所有具备 GR 能力的 OSPF 邻居之间同步数据库，并交换路由信息。交换完成后，GR Restarter 根据新的路由转发信息更新路由表和转发表，删除失效的路由，完成 OSPF 协议收敛。

4. TE与DS-TE

OSPF TE（Traffic Engineering，流量工程）支持建立和维护 TE 的 LSP（Label Switched Path，标签交换路径）。

MPLS（Multiprotocol Label Switching，多协议标签交换）在构建 CR LSP（Constraint-based Routed LSP，基于约束路由的 LSP）时，需要了解本区域中所有链路的流量属性信息。它通过 OSPF 来获取链路的流量工程信息。

OSPF 支持 Opaque LSA，可以用来承载流量工程信息。

DS-TE（DiffServ Aware TE）主要用于优化和分配网络传输资源、对流进行分类、指定每条流在链路带宽中所占的比例。流量工程是基于划分后的类（细粒度聚合类）实现的，而不是聚合的类（粗粒度聚合类）。这一点改善了性能和带宽的利用率。

为了支持 MPLS 中 DS-TE 的应用，OSPF 支持子 TLV（Local Overbooking Multiplier TLV）和 BC（Bandwidth Constraint，带宽约束）TLV。

说明

OSPF TE 的详细配置请参见“MPLS 配置指导”中的“MPLS TE”。

---

5. IGP Shortcut和Forwarding Adjacency

OSPF 支持 IGP Shortcut 和 Forwarding Adjacency 特性，这两个特性允许 OSPF 使用 LSP 作为到达某个目的地址的出接口。否则，即使存在到达某个目的地址的 LSP，OSPF 也不能使用它作为出接口。

IGP Shortcut 和 Forwarding Adjacency 的区别在于：

- 如果仅使能了 Forwarding Adjacency 特性，OSPF 也可以使用 LSP 到达目的地址；
- 如果仅使能了 IGP Shortcut，则只有使能此特性的路由器才可以在路由中使用 LSP。

说明

此特性的详细配置请参见“MPLS 配置指导”中的“MPLS TE”。

---

6. OSPF支持多实例

OSPF 支持多实例，可以运行在 VPN 网络中的 PE 上。

在 BGP MPLS VPN 网络中，属于同一个 VPN 的多个 Site 可以使用 OSPF 作为内部路由协议。然而，它们会被看作属于不同的自治系统来处理。这样，在一个节点学到的 OSPF 路由，将被作为外部路由传送给另一节点。这种处理方式导致了比较高的 OSPF 路由协议流量，并带来了一些原本可以避免的网络管理问题。

目前，可以在 PE 上通过配置域 ID 来区分 Site 所在的 VPN。属于同一个 VPN 的不同 Site 之间彼此看作是直接相连的。这样，PE 路由器之间交换 OSPF 路由信息时就好像是通过一条专线相连，改善了网络管理并使 OSPF 的应用更为有效。

说明

此特性的相关配置请参见“MPLS 配置指导”中的“MPLS L3VPN”。

7. OSPF伪连接

OSPF 伪连接（sham link）是 MPLS VPN 骨干网上两个 PE 路由器之间的点到点链路。

通常情况下，BGP 对等体之间通过 BGP 扩展团体属性在 MPLS VPN 骨干网上承载路由信息。另一端 PE 上运行的 OSPF 可利用这些信息来生成 PE 到 CE 的 Type3 LSA，这些路由是区域间路由。但是，如果路由器和它同一区域内的 PE 路由器相连，且建立到达特定目的地址的内部区域路由（后门路由），那么 VPN 流量就将总是穿越这条后门路由，而不是骨干路由。这是因为在路由表中建立的 OSPF 内部区域路由的优先级较高。为了避免这一异常现象，可以在 PE 路由器之间配置一条 unnumbered 的点到点伪连接。这样，就可以通过一条低开销的内部区域路由到达 PE 路由器。

说明

OSPF 伪连接的相关配置请参见“MPLS 配置指导”中的“MPLS L3VPN”。

8. OSPF与BFD联动

说明

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”。

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。

1.1.8 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 3630：Traffic Engineering Extensions to OSPF Version 2
- 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

在一个运行 OSPF 的路由域里，有各种类型的路由器，区域内路由器、ABR、ASBR 等。

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；使能 OSPF 功能后，路由器可以使用一些缺省配置，如 OSPF 各种协议报文的发送时间间隔、LSA 延迟时间、SPF 计算时间间隔等，也可以根据需要修改缺省配置。

在 OSPF 路由器上进行各项配置的时候都应该先做好网络规划，在配置同一区域内的路由器时，配置时应该以区域为基础来统一考虑，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。

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
	配置对Type-3 LSA进行过滤	可选	1.6.4
	配置OSPF接口的开销值	可选	1.6.5
	配置OSPF支持的路由最大数目	可选	1.6.6
	配置OSPF最大等价路由条数	可选	1.6.7
	配置OSPF协议的优先级	可选	1.6.8
	配置OSPF引入外部路由	可选	1.6.9
	配置发布一条主机路由	可选	1.6.10
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SPF计算时间间隔	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DD报文中的MTU	可选	1.7.10

配置任务		说明	详细配置
	配置LSDB中External LSA的最大数量	可选	1.7.11
	配置兼容RFC 1583的外部路由选择规则	可选	1.7.12
	配置邻接状态输出	可选	1.7.13
	配置OSPF网管功能	可选	1.7.14
	使能日志功能	可选	1.7.15
	使能Opaque LSA发布接收能力	可选	1.7.16
	配置OSPF优先处理Hello报文功能	可选	1.7.17
	配置OSPF发送更新报文的速度	可选	1.7.18
	配置OSPF报文的DSCP优先级	可选	1.7.19
	配置OSPF ISPF	可选	1.7.20
配置OSPF快速重路由		可选	1.8
配置OSPF GR	配置GR Restarter	可选	1.9.1
	配置GR Helper	可选	1.9.2
	重启OSPF GR进程	可选	1.9.3
配置OSPF与BFD联动		可选	1.10

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，其它功能特性的配置才能生效。

1.3.1 配置准备

在配置 OSPF 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的标识一台路由器，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。
- 可以配置 OSPF 在 VPN 实例中运行。如果指定了 VPN 实例，那么此 OSPF 进程属于指定的 VPN 实例。

表1-2 使能 OSPF 功能

操作	命令	说明
进入系统视图	**system-view**	-
配置全局Router ID	**router id** *router-id*	可选 缺省情况下，未配置全局Router ID 如果没有配置全局Router ID，则按照下面的规则进行选择： (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 (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）
创建OSPF，进入OSPF视图	**ospf** [*process-id* | **router-id** *router-id* | **vpn-instance** *vpn-instance-name*] *	必选 缺省情况下，系统不运行OSPF
配置OSPF进程描述	**description** *description*	可选 缺省情况下，没有配置进程描述
配置OSPF区域，进入OSPF区域视图	**area** *area-id*	必选 缺省情况下，没有配置OSPF区域
配置区域描述	**description** *description*	可选 缺省情况下，没有配置区域描述
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*	必选 缺省情况下，接口不属于任何区域且OSPF功能处于关闭状态

说明

- 一个网段只能属于一个区域。
- 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。
- 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。

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

1.4.2 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，Type5 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（完全 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。

表1-3 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**] *	必选 缺省情况下，没有区域被设置为Stub区域
配置ABR发送到Stub区域缺省路由的开销	**default-cost** *cost*	可选 缺省情况下，ABR发送到Stub区域缺省路由的开销为1

说明

- Stub 区域内的所有路由器必须使用 **stub** 命令将该区域配置成 Stub 属性。
- **default-cost** 命令只有在 Stub 区域的 ABR 上配置才能生效。
- 骨干区域不能配置成（Totally）Stub 区域。
- （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。
- 虚连接不能穿过（Totally）Stub 区域。

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。

表1-4 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** | **no-import-route** | **no-summary** | **translate-always** | **translator-stability-interval** *value*] *	必选 缺省情况下，没有区域被设置为NSSA区域
配置发送到NSSA区域缺省路由的开销	**default-cost** *cost*	可选 缺省情况下，发送到NSSA区域的缺省路由的开销为1

说明

- NSSA 区域内的所有路由器必须使用 **nssa** 命令将该区域配置成 NSSA 属性。
- **default-cost** 命令只有在 NSSA 区域的 ABR/ASBR 上配置才能生效。

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。

表1-5 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* [**hello** *seconds* | **retransmit** *seconds* | **trans-delay** *seconds* | **dead** *seconds* | **simple** [**plain** | **cipher**] *password* | { **md5** | **hmac-md5** } *key-id* [**plain** | **cipher**] *password*] *	必选 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值也不同：

- 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；
- 点对点：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点对点。

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，那么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点对多点，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1-6 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**	-

操作	命令	说明
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**	必选 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
配置OSPF接口的路由器优先级	**ospf dr-priority** *priority*	可选 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1-7 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**	必选 缺省情况下，接口的网络类型根据物理接口而定
配置OSPF接口的路由器优先级	**ospf dr-priority** *priority*	可选 缺省情况下，接口的路由器优先级为1
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* [**cost** *value* | **dr-priority** *dr-priority*]	必选

说明

使用 **peer** 命令和使用 **ospf dr-priority** 命令设置的优先级具有不同的用途：

- **peer** 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，以建立邻居关系。

- **ospf dr-priority** 命令设置的优先级用于实际的 DR 选举。

1.5.4 配置OSPF接口网络类型为P2MP

表1-8 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**]	必选 缺省情况下，接口的网络类型根据物理接口而定 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*] *	-
配置P2MP单播网络的邻居	**peer** *ip-address* [**cost** *value* | **dr-priority** *dr-priority*]	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1-9 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**	必选 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 Type3 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1-10 配置 ABR 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*] *	-
进入OSPF区域视图	**area** *area-id*	-
配置OSPF的ABR路由聚合	**abr-summary** *ip-address* { *mask* | *mask-length* } [**advertise** | **not-advertise**] [**cost** *cost*]	必选 此命令只有在ABR上配置才会有效 缺省情况下，ABR不会对路由进行聚合

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 Type5 LSA 进行聚合。

如果 ASBR 在 NSSA 区域里面，将对聚合地址范围内的 Type7 LSA 进行聚合，当本地路由器同时是 ASBR 和 ABR 时，将对由 Type7 LSA 转化成的 Type5 LSA 进行聚合处理。

表1-11 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* | *mask-length* } [**tag** *tag* | **not-advertise** | **cost** *cost*] *	必选 此命令只有在ASBR上配置才会有效 缺省情况下，ASBR不会对引入的路由进行聚合

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

说明

- 关于地址前缀列表的详细介绍，请参见“三层技术-IP 路由配置指导”中的“路由策略”。
- 关于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”。

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1-12 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*] *	-

操作	命令	说明
配置对接收的路由进行过滤	**filter-policy** { *acl-number* [**gateway** *ip-prefix-name*] | **gateway** *ip-prefix-name* | **ip-prefix** *ip-prefix-name* [**gateway** *ip-prefix-name*] | **route-policy** *route-policy-name* } **import**	必选 缺省情况下，不对接收到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1-13 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* | **ip-prefix** *ip-prefix-name* } { **import** | **export** }	必选 缺省情况下，没有对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（100Mbps）÷接口带宽（Mbps）。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。

如果没有在接口下配置接口的开销值，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1-14 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*	-

操作	命令	说明
设置OSPF接口的开销值	**ospf cost** *value*	可选 缺省情况下，VLAN接口的缺省值为1；Loopback接口的缺省值为0；其他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

2. 配置带宽参考值

表1-15 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** *vpn-instance-name*] *	-
配置带宽参考值	**bandwidth-reference** *value*	可选 缺省情况下，带宽参考值为100Mbps

1.6.6 配置OSPF支持的路由最大数目

表1-16 配置 OSPF 支持的路由最大数目

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支持的路由最大数目	**maximum-routes** { **external** | **inter** | **intra** } *number*	可选 本命令的缺省值与设备的型号有关，请以设备的实际情况为准

1.6.7 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*] *	-

操作	命令	说明
配置OSPF最大等价路由条数	**maximum load-balancing** *maximum*	可选 缺省值为8

1.6.8 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*	可选 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

1.6.9 配置OSPF引入外部路由

1. 配置OSPF引入其它协议的路由

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**] [**cost** *cost* | **type** *type* | **tag** *tag* | **route-policy** *route-policy-name*] *	必选 缺省情况下，没有引入其他协议的路由信息

操作	命令	说明
配置对引入的路由进行过滤	**filter-policy** { *acl-number* | **ip-prefix** *ip-prefix-name* } **export** [*protocol* [*process-id*]]	可选 缺省情况下，没有对引入的路由信息进行过滤

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。

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*]	必选 缺省情况下，没有引入缺省路由

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、路由上限、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*] *	-

操作	命令	说明
配置引入外部路由时的参数缺省值（开销、路由上限、标记、类型）	**default** { **cost** *cost* | **limit** *limit* | **tag** *tag* | **type** *type* } *	可选 缺省情况下，OSPF引入外部路由的缺省值如下： • 路由度量值为 1 • 单位时间内引入外部路由数量的上限为 1000 • 外部路由标记值为 1 • 引入的外部路由类型为 Type2

1.6.10 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*	必选 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。
- OSPF 同时支持网管功能，可以配置 OSPF MIB 与某一进程绑定，以及发送 Trap 消息和日志功能。

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*	可选 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
配置Poll定时器	**ospf timer poll** *seconds*	可选 缺省情况下，发送轮询Hello报文的时间间隔为120秒
配置邻居失效时间	**ospf timer dead** *seconds*	可选 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
配置接口重传LSA的时间间隔	**ospf timer retransmit** *interval*	可选 缺省情况下，时间间隔为5秒

说明

- 修改了网络类型后，Hello 定时器与邻居失效时间都将恢复缺省值。
- 在同一接口上邻居失效时间应至少为 Hello 时间间隔的 4 倍。
- 轮询 Hello 报文的时间间隔至少应为 Hello 时间间隔的 4 倍。
- 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*	可选 缺省情况下，接口传送LSA的延迟时间为1秒

1.7.4 配置SPF计算时间间隔

当 OSPF 的链路状态数据库（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。

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*] *	-
配置SPF计算时间间隔	**spf-schedule-interval** *maximum-interval* [*minimum-interval* [*incremental-interval*]]	可选 缺省情况下，SPF计算的时间间隔为5秒

说明

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*。

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*	可选 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。

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* [*initial-interval* [*incremental-interval*]]	可选 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5000毫秒

说明

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*。

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。

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** }	可选 缺省情况下，允许接口收发OSPF报文

说明

- 不同的进程可以对同一接口禁止收发 OSPF 报文，但 **silent-interface** 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。
- 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。

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 分别表示通过点对点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。

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Stub路由器	**stub-router**	必选 缺省情况下，没有路由器被配置为Stub路由器

说明

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对 OSPF 路由器进行恶意攻击，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。

要配置 OSPF 报文验证，同一个区域的所有路由器上都需要配置区域验证模式，且配置的验证模式必须相同，同一个网段内的路由器需要配置相同的接口验证模式和口令。

表1-30 配置 OSPF 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** { **md5** | **simple** }	必选 缺省情况下，没有配置区域验证模式
退回OSPF视图	**quit**	-
退回系统视图	**quit**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置OSPF接口的验证模式（简单验证）	**ospf authentication-mode simple** [**cipher** | **plain**] *password*	二者必选其一 缺省情况下，接口不对OSPF报文进行验证
配置OSPF接口的验证模式（MD5验证）	**ospf authentication-mode** { **hmac-md5** | **md5** } *key-id* [**cipher** | **plain**] *password*	

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1-31 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**	可选 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LSDB中External LSA的最大数量

表1-32 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*	可选 缺省情况下，不对LSDB中External LSA的最大条目数进行限制

1.7.12 配置兼容RFC 1583 的外部路由选择规则

当通过多条 LSA 计算出同一条外部路由时，在 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 中定义的方式。当 RFC 2328 兼容 RFC 1583 时，优选骨干区的区域内路由；当 RFC 2328 不兼容 RFC 1583 时，优选非骨干区的区域内路由，这样做的目的是尽量减少骨干区的负担。

表1-33 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**	可选 缺省情况下，使能兼容RFC 1583的选路规则

说明

为了避免路由环路，对于是否兼容 RFC 1583 的外部路由选择规则，同一路由域内的路由器建议配置相同，即要么配置所有路由器都兼容 RFC 1583 的外部路由选择规则，要么配置所有路由器都不兼容 RFC 1583 的外部路由选择规则。

1.7.13 配置邻居状态变化的输出开关

表1-34 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*] *	-
使能OSPF日志记录功能，记录邻居关系连接以及断开事件的日志信息	**log-peer-change**	可选 缺省情况下，使能OSPF日志记录功能

1.7.14 配置OSPF网管功能

开启 OSPF 模块的 Trap 功能后，该模块会生成 Trap 报文，用于报告该模块的重要事件。Trap 报文的级别如下：

- 错误级别 Trap 报文，级别为 level3；
- 警告级别 Trap 报文，级别为 level4；
- 正常出现但是重要信息级别 Trap 报文，级别为 level5；
- 通知信息级别 Trap 报文，级别为 level6。

生成的 Trap 报文将被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 Trap 报文的输出规则（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1-35 配置 OSPF 网管功能

操作	命令	说明
进入系统视图	**system-view**	-
配置OSPF MIB绑定	**ospf mib-binding** *process-id*	可选 缺省情况下，MIB绑定在进程号最小的OSPF进程上
使能OSPF的TRAP功能	**snmp-agent trap enable ospf** [*process-id*] [**ifauthfail** | **ifcfgerror** | **ifrxbadpkt** | **ifstatechange** | **iftxretransmit** | **lsdbapproachoverflow** | **lsdboverflow** | **maxagelsa** | **nbrstatechange** | **originatelsa** | **vifcfgerror** | **virifauthfail** | **virifrxbadpkt** | **virifstatechange** | **viriftxretransmit** | **virnbrstatechange**] *	可选 缺省情况下，OSPF的TRAP功能处于使能状态

1.7.15 使能日志功能

表1-36 使能日志功能

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*] *	-
使能日志功能	**enable log** [**config** | **error** | **state**]	必选 缺省情况下，日志功能处于关闭状态

1.7.16 使能Opaque LSA发布接收能力

通过使能 Opaque LSA 发布接收能力，OSPF 可以接收和发布 Type9、Type10 和 Type11 的 Opaque LSA。

表1-37 使能 Opaque LSA 发布接收能力

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Opaque LSA发布接收能力	**opaque-capability enable**	可选 缺省情况下，Opaque LSA发布接收能力处于关闭状态

1.7.17 配置OSPF优先接收并处理Hello报文

为了保证协议的正常运行，路由器要同时接收和处理 Hello 报文和其它类型的协议报文，当路由器与多个邻居路由器同时建立邻居关系并且路由表中的路由条数比较多时，需要路由器接收和处理的报文数量会很大，可以通过配置 OSPF 优先接收和处理 Hello 报文，来确保邻居关系的稳定性。

表1-38 配置 OSPF 优先接收和处理 Hello 报文

操作	命令	说明
进入系统视图	**system-view**	-
配置OSPF优先接收和处理Hello报文	**ospf packet-process prioritized-treatment**	必选 缺省情况下，OSPF不优先接收和处理Hello报文

1.7.18 配置LSU报文的发送速率

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 LSU 报文的发送速率

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LSU报文的发送速率	**transmit-pacing interval** *interval* **count** *count*	可选 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

1.7.19 配置OSPF报文的DSCP优先级

在 IPv4 报文头中，包含一个 8bit 的 ToS 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 OSPF 报文的 DSCP 优先级进行配置。

表1-40 配置 OSPF 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
创建OSPF，进入OSPF视图	**ospf** [*process-id* | **router-id** *router-id* | **vpn-instance** *vpn-instance-name*] *	必选 缺省情况下，系统不运行OSPF
配置OSPF报文的DSCP优先级	**dscp** *dscp-value*	可选 缺省情况下，OSPF报文的DSCP优先级为48

1.7.20 配置OSPF ISPF

ISPF（Incremental Shortest Path First，增量最短路径优先）计算是对 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，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 OSPF ISPF

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*] *	-
使能OSPF ISPF功能，即增量SPF计算功能	**ispf enable**	可选 缺省情况下，OSPF ISPF功能处于关闭状态

1.8 配置OSPF快速重路由

说明

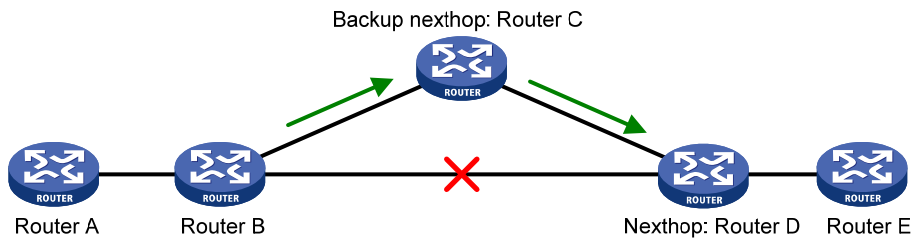
- OSPF 支持快速重路由功能不能与 OSPF 支持 BFD 监测同时使用，否则可能导致快速重路由功能失效。
- OSPF 支持快速重路由功能（自动计算备份下一跳）不能与 **vlink-peer**、**sham-link**（MPLS 命令参考/MPLS L3VPN）、**enable traffic-adjustment**（MPLS 命令参考/MPLS TE）、**enable traffic-adjustment advertise**（MPLS 命令参考/MPLS TE）命令同时使用。

1. 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1-21 OSPF 快速重路由功能示意图



如 图 1-21 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探测到网络故障时，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。

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

3. 配置OSPF支持快速重路由功能（自动计算备份下一跳）

表1-42 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**	-
配置BFD Echo报文源地址	**bfd echo-source-ip** *ip-address*	必选 缺省情况下，没有配置BFD Echo报文源地址
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（自动计算备份下一跳）	**fast-reroute auto** [**abr-only**]	必选 缺省情况下，没有配置OSPF快速重路由功能 **abr-only**表示仅选取到ABR设备的路由作为备份下一跳

4. 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1-43 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**	-
配置BFD Echo报文源地址	**bfd echo-source-ip** *ip-address*	必选 缺省情况下，没有配置BFD Echo报文源地址
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*	必选 缺省情况下，没有配置OSPF快速重路由功能

.

1.9 配置OSPF GR

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。

1.9.1 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1-44 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**	必选 缺省情况下，Opaque LSA发布接收能力处于关闭状态
使能OSPF协议的IETF标准GR能力	**graceful-restart ietf**	必选 缺省情况下，OSPF协议的IETF标准GR能力处于关闭状态
配置OSPF协议的GR重启间隔时间	**graceful-restart interval** *timer*	可选 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1-45 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**	必选 缺省情况下，OSPF本地链路信令能力处于关闭状态

操作	命令	说明
使能OSPF带外同步能力	**enable out-of-band-resynchronization**	必选 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**]	必选 缺省情况下，OSPF协议的非IETF标准GR能力处于关闭状态
配置OSPF协议的GR重启间隔时间	**graceful-restart interval** *timer*	可选 缺省情况下，OSPF协议的GR重启间隔时间为120秒

1.9.2 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1-46 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**	必选 缺省情况下，Opaque LSA发布接收能力处于关闭状态
配置允许设备可以作哪些OSPF邻居的GR Helper	**graceful-restart helper** { *acl-number* | **prefix** *prefix-list* }	可选 缺省情况下，设备可以作任一OSPF邻居的GR Helper

2. 配置非IETF标准GR Helper

表1-47 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**	必选 缺省情况下，OSPF本地链路信令能力处于关闭状态

操作	命令	说明
使能OSPF带外同步能力	**enable out-of-band-resynchronization**	必选 缺省情况下，OSPF带外同步能力处于关闭状态
配置允许设备可以作哪些OSPF邻居的GR Helper	**graceful-restart help** { *acl-number* | **prefix** *prefix-list* }	可选 缺省情况下，设备可以作任一OSPF邻居的GR Helper

1.9.3 以GR方式重启OSPF进程

发生 Master/Slave 切换或者路由器上进行如下操作均可以以 GR 方式重启 OSPF 进程。

表1-48 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**	必选 请在用户视图下执行该命令

1.10 配置OSPF与BFD联动

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1-49 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*	-
指定接口上使能BFD	**ospf bfd enable**	必选 缺省情况下，运行OSPF的接口未使能BFD提供的链路检测功能

说明

- 一个网段只能属于同一个区域，并且必须为每个运行 OSPF 协议的接口指明属于某一个特定的区域。
- 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1-50 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**	-
配置echo报文源地址	**bfd echo-source-ip** *ip-address*	必选 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*	-
指定接口上使能BFD功能	**ospf bfd enable echo**	必选 缺省情况下，运行OSPF的接口未使能BFD提供的链路检测功能

1.11 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** 命令可以复位 OSPF 计数器或连接。

表1-51 OSPF 显示和维护

操作	命令
显示OSPF的概要信息	**display ospf** [*process-id*] **brief** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF的统计信息	**display ospf** [*process-id*] **cumulative** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF的LSDB信息	**display ospf** [*process-id*] **lsdb** [**brief** | [{ **ase** | **router** | **network** | **summary** | **asbr** | **nssa** | **opaque**-link | **opaque-area** | **opaque**-as } [*link-state-id*]] [**originate-router** *advertising-router-id* | **self-originate**]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF邻居的信息	**display ospf** [*process-id*] **peer** [**verbose**] [*interface-type interface-number*] [*neighbor-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF各区域邻居的统计信息	**display ospf** [*process-id*] **peer statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
显示OSPF下一跳信息	**display ospf** [*process-id*] **nexthop** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF路由表的信息	**display ospf** [*process-id*] **routing** [**interface** *interface-type interface-number*] [**nexthop** *nexthop-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF虚连接信息	**display ospf** [*process-id*] **vlink** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF请求列表	**display ospf** [*process-id*] **request-queue** [*interface-type interface-number*] [*neighbor-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF重传列表	**display ospf** [*process-id*] **retrans-queue** [*interface-type interface-number*] [*neighbor-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF ABR及ASBR信息	**display ospf** [*process-id*] **abr-asbr** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF接口信息	**display ospf** [process-id] **interface** [**all** | *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF错误信息	**display ospf** [*process-id*] **error** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPF ASBR聚合信息	**display ospf** [*process-id*] **asbr-summary** [*ip-address* { *mask* | *mask-length* }] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示全局Router ID	**display router id** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除OSPF计数器	**reset ospf** [process-id] **counters** [**neighbor** [*interface-type interface-number*] [*router-id*]]
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**

1.12 典型配置举例

注意

在配置举例中，只列出了与 OSPF 配置相关的命令。

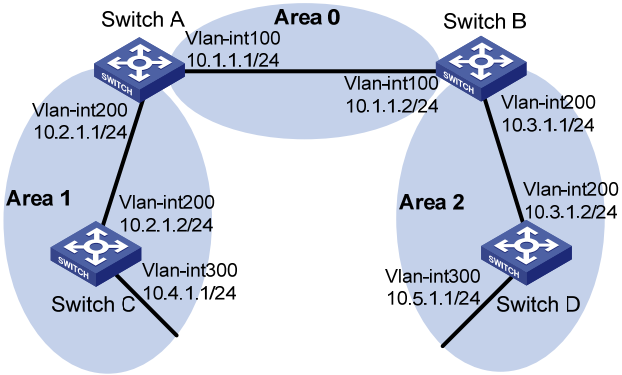
1.12.1 配置OSPF基本功能

1. 组网需求

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

2. 组网图

图1-22 OSPF 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 OSPF 基本配置

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] area 1
[SwitchA-ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.1] quit
[SwitchA-ospf-1] quit
```

配置 Switch B。

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

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf-1] area 1
[SwitchC-ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.1] network 10.4.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.1] quit
[SwitchC-ospf-1] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] ospf
[SwitchD-ospf-1] area 2
[SwitchD-ospf-1-area-0.0.0.2] network 10.3.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.2] network 10.5.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.2] quit
[SwitchD-ospf-1] quit
```

(3) 检验配置结果

查看 Switch A 的 OSPF 邻居。

```
[SwitchA] 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(Vlan-interface100)'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
   Dead timer due in 37  sec
   Neighbor is up for 06:03:59
   Authentication Sequence: [ 0 ]
   Neighbor state change count: 5

                  Neighbors

 Area 0.0.0.1 interface 10.2.1.1(Vlan-interface200)'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
   Dead timer due in 32  sec
   Neighbor is up for 06:03:12
   Authentication Sequence: [ 0 ]
   Neighbor state change count: 5
```

显示 Switch A 的 OSPF 路由信息。

```
[SwitchA] 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        10       Transit 10.2.1.1        10.2.1.1        0.0.0.1
 10.3.1.0/24        4        Inter   10.1.1.2        10.3.1.1        0.0.0.0
 10.4.1.0/24        13       Stub    10.2.1.2        10.4.1.1        0.0.0.1
 10.5.1.0/24        14       Inter   10.1.1.2        10.3.1.1        0.0.0.0
 10.1.1.0/24        2        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
```

显示 Switch A 的 LSDB。

```
[SwitchA] display ospf lsdb

          OSPF Process 1 with Router ID 10.2.1.1
                  Link State Database

                          Area: 0.0.0.0
 Type      LinkState ID    AdvRouter          Age  Len   Sequence   Metric
 Router    10.2.1.1        10.2.1.1          1069  36    80000012       0
 Router    10.3.1.1        10.3.1.1           780  36    80000011       0
 Network   10.1.1.1        10.2.1.1          1069  32    80000010       0
 Sum-Net   10.5.1.0        10.3.1.1           780  28    80000003      12
 Sum-Net   10.2.1.0        10.2.1.1          1069  28    8000000F      10
 Sum-Net   10.3.1.0        10.3.1.1           780  28    80000014       2
 Sum-Net   10.4.1.0        10.2.1.1           769  28    8000000F      13
                          Area: 0.0.0.1
 Type      LinkState ID    AdvRouter          Age  Len   Sequence   Metric
 Router    10.2.1.1        10.2.1.1           769  36    80000012       0
 Router    10.4.1.1        10.4.1.1          1663  48    80000012       0
 Network   10.2.1.1        10.2.1.1           769  32    80000010       0
 Sum-Net   10.5.1.0        10.2.1.1           769  28    80000003      14
 Sum-Net   10.3.1.0        10.2.1.1          1069  28    8000000F       4
 Sum-Net   10.1.1.0        10.2.1.1          1069  28    8000000F       2
 Sum-Asbr  10.3.1.1        10.2.1.1          1069  28    8000000F       2
```

查看 Switch D 的路由表。

```
[SwitchD] 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

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

在 Switch D 上使用 Ping 进行测试连通性。

```
[SwitchD] ping 10.4.1.1
  PING 10.4.1.1: 56  data bytes, press CTRL_C to break
    Reply from 10.4.1.1: bytes=56 Sequence=2 ttl=253 time=2 ms
    Reply from 10.4.1.1: bytes=56 Sequence=2 ttl=253 time=1 ms
    Reply from 10.4.1.1: bytes=56 Sequence=3 ttl=253 time=1 ms
    Reply from 10.4.1.1: bytes=56 Sequence=4 ttl=253 time=1 ms
```

```
  Reply from 10.4.1.1: bytes=56 Sequence=5 ttl=253 time=1 ms

--- 10.4.1.1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/1/2 ms
```

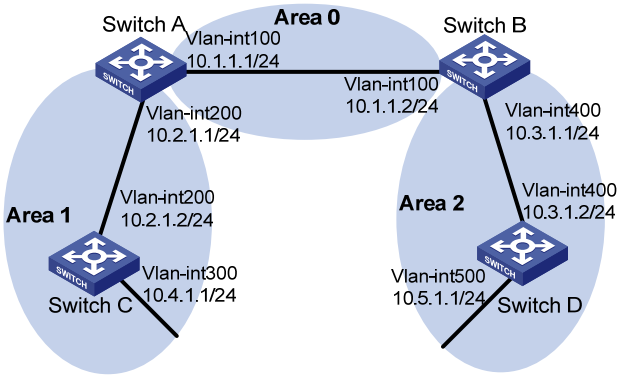
1.12.2 配置OSPF引入自治系统外部路由

1. 组网需求

- 所有的交换机都运行 OSPF，整个自治系统划分为 3 个区域。
- 其中 Switch A 和 Switch B 作为 ABR 来转发区域之间的路由。
- 在 Switch 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1-23 配置 OSPF 引入自治系统外部路由



3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置OSPF基本功能（同前例 1.12.1 ）

(3) 配置引入自治系统外部路由

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

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

在 Switch C 上配置 OSPF 引入静态路由。

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

(4) 检验配置效果

查看 Switch D 的 ABR/ASBR 信息。

```
<SwitchD> 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 Switch D 的 OSPF 路由表。

```
<SwitchD> 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
```

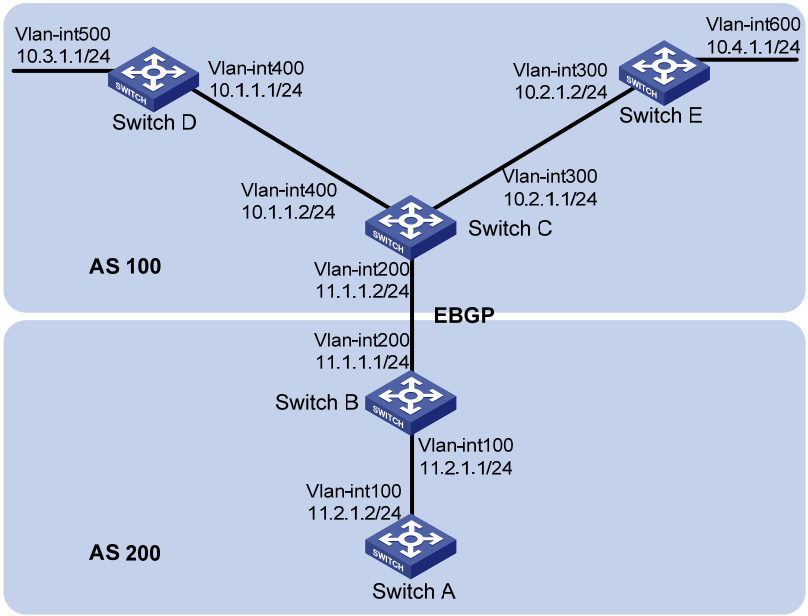
1.12.3 配置OSPF发布聚合路由

1. 组网需求

- Switch A 和 Switch B 位于 AS 200 内，AS 200 内使用 OSPF 作为 IGP 协议。
- Switch C、Switch D 和 Switch E 位于 AS 100 内，AS 100 内使用 OSPF 作为 IGP 协议。
- Switch B 和 Switch 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 Switch A 的路由表规模，在 Switch 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1-24 配置 OSPF 发布聚合路由



3. 配置步骤

(1) 配置接口的 IP 地址（略）

(2) 配置 OSPF

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 11.2.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 11.2.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

配置 Switch C。

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

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] ospf
[SwitchD-ospf-1] area 0
[SwitchD-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.0] network 10.3.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.0] quit
```

配置 Switch E。

```
<SwitchE> system-view
[SwitchE] ospf
[SwitchE-ospf-1] area 0
[SwitchE-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[SwitchE-ospf-1-area-0.0.0.0] network 10.4.1.0 0.0.0.255
[SwitchE-ospf-1-area-0.0.0.0] quit
[SwitchE-ospf-1] quit
```

(3) 配置 BGP，引入 OSPF 和直连路由

配置 Switch B。

```
[SwitchB] bgp 200
[SwitchB-bgp] peer 11.1.1.2 as 100
[SwitchB-bgp] import-route ospf
[SwitchB-bgp] import-route direct
[SwitchB-bgp] quit
```

配置 Switch C。

```
[SwitchC] bgp 100
[SwitchC-bgp] peer 11.1.1.1 as 200
[SwitchC-bgp] import-route ospf
[SwitchC-bgp] import-route direct
[SwitchC-bgp] quit
```

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

在 Switch B 上配置 OSPF 引入 BGP 路由。

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

在 Switch C 上配置 OSPF 引入 BGP 路由。

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

查看 SwitchA 的路由表信息。

```
[SwitchA] display ip routing-table
Routing Tables: Public
         Destinations : 8        Routes : 8

Destination/Mask    Proto  Pre  Cost         NextHop         Interface

10.1.1.0/24         O_ASE  150  1            11.2.1.1        Vlan100
10.2.1.0/24         O_ASE  150  1            11.2.1.1        Vlan100
10.3.1.0/24         O_ASE  150  1            11.2.1.1        Vlan100
10.4.1.0/24         O_ASE  150  1            11.2.1.1        Vlan100
```

```
11.2.1.0/24         Direct 0    0           11.2.1.2        Vlan100
11.2.1.2/32         Direct 0    0           127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0           127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0           127.0.0.1       InLoop0
```

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

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

查看 Switch A 的路由表信息。

```
[SwitchA] display ip routing-table
Routing Tables: Public
         Destinations : 5        Routes : 5

Destination/Mask    Proto   Pre  Cost         NextHop         Interface

10.0.0.0/8          O_ASE   150  2            11.2.1.1        Vlan100
11.2.1.0/24         Direct 0    0           11.2.1.2        Vlan100
11.2.1.2/32         Direct 0    0           127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0           127.0.0.1       InLoop0
127.0.0.1/32        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。

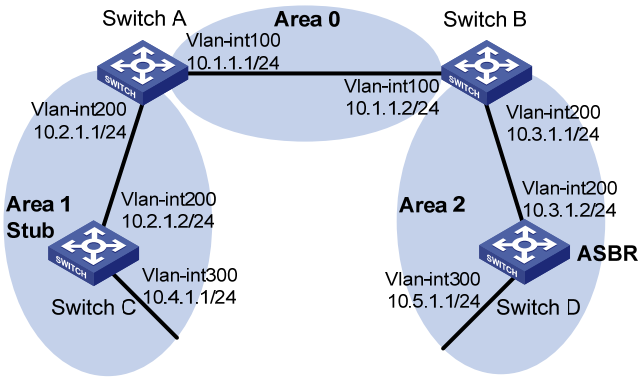
1.12.4 配置OSPF的Stub区域

1. 组网需求

- 所有的交换机都运行 OSPF，整个自治系统划分为 3 个区域。
- 其中 Switch A 和 Switch B 作为 ABR 来转发区域之间的路由，Switch 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1-25 配置 OSPF Stub 区域组网图



3. 配置步骤

(1) 配置接口的 IP 地址（略）

(2) 配置OSPF基本功能（同前例 1.12.1 ）

(3) 配置 Switch D 引入静态路由

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

查看 Switch C 的 ABR/ASBR 信息。

```
<SwitchC> 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       10.3.1.1        0.0.0.1         5     10.2.1.1        ABR
 Inter       10.5.1.1        0.0.0.1         7     10.2.1.1        ASBR
```

查看 Switch C 的 OSPF 路由表。

```
<SwitchC> 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 Switch 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 Switch A。

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

```
[SwitchA-ospf-1] quit
```

配置 Switch C。

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

显示 Switch C 的 OSPF 路由表。

```
[SwitchC] 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 Switch 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。

配置禁止向 Stub 区域通告 Type3 LSA。

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

查看 Switch C 的 OSPF 路由表。

```
[SwitchC] 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。

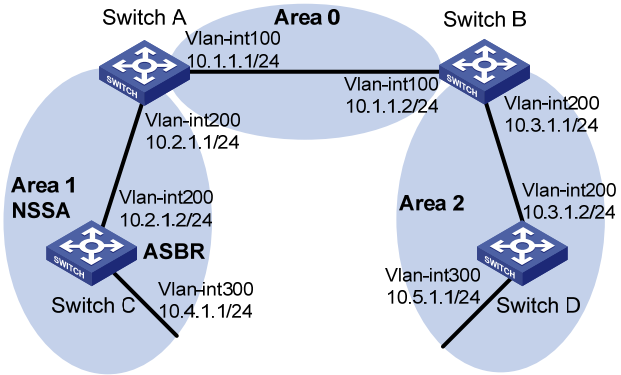
1.12.5 配置OSPF的NSSA区域

1. 组网需求

- 所有的交换机都运行 OSPF，整个自治系统划分为 3 个区域。
- 其中 Switch A 和 Switch 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 Switch 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1-26 配置 OSPF NSSA 区域组网图



3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置OSPF基本功能（同前例 1.12.1 ）

(3) 配置 Area1 区域为 NSSA 区域

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 1
[SwitchA-ospf-1-area-0.0.0.1] nssa default-route-advertise no-summary
[SwitchA-ospf-1-area-0.0.0.1] quit
[SwitchA-ospf-1] quit
```

配置 Switch C。

```
<SwitchC> system-view
```

```
[SwitchC] ospf
[SwitchC-ospf-1] area 1
[SwitchC-ospf-1-area-0.0.0.1] nssa
[SwitchC-ospf-1-area-0.0.0.1] quit
[SwitchC-ospf-1] quit
```

- 如果 NSSA 区域内路由器（Switch C）需要获取通往 AS 内其他区域的路由，ABR（Switch A）上必须配置 **default-route-advertise** 参数，这样 Switch C 才可以获取到缺省路由。
- 建议在 ABR(Switch 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 Switch C 的 OSPF 路由表。

```
[SwitchC] 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          65536    Inter   10.2.1.1        10.2.1.1        0.0.0.1
 10.2.1.0/24        65535    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
```

(4) 配置 Switch C 引入静态路由

```
[SwitchC] ip route-static 3.1.3.1 24 10.4.1.2
[SwitchC] ospf
[SwitchC-ospf-1] import-route static
[SwitchC-ospf-1] quit
```

查看 Switch D 的 OSPF 路由表。

```
<SwitchD> 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
3.1.3.0/24         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 Switch D 上可以看到 NSSA 区域引入的一条 AS 外部的路由。

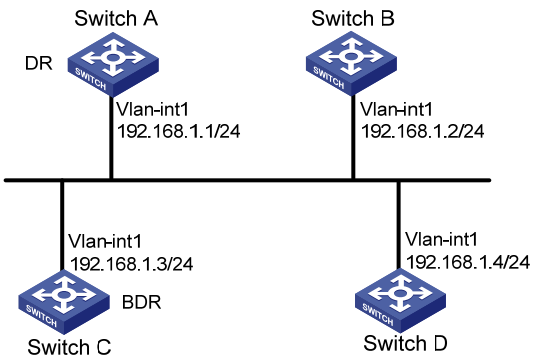
1.12.6 配置OSPF的DR选择

1. 组网需求

Switch A、Switch B、Switch C、Switch D 在同一网段，运行 OSPF 协议；

配置 Switch A 为 DR，Switch C 为 BDR。

图1-27 配置 OSPF 优先级的 DR 选择组网图



2. 配置思路

(1) 配置各接口的 IP 地址

(2) 配置 OSPF 基本功能

(3) 改变交换机接口的路由器优先级使 Switch A 成为 DR，Switch 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 Switch A。

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

配置 Switch B。

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

配置 Switch C。

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

配置 Switch D。

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

查看 Switch A 的邻居信息。

```
[SwitchA] 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(Vlan-interface1)'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
   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
   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
  Dead timer due in 31  sec
  Neighbor is up for 00:01:28
  Authentication Sequence: [ 0 ]
```

可以看到 Switch D 为 DR，Switch C 为 BDR。

(3) 配置接口上的路由器优先级

配置 Switch A。

```
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] ospf dr-priority 100
[SwitchA-Vlan-interface1] quit
```

配置 Switch B。

```
[SwitchB] interface vlan-interface 1
[SwitchB-Vlan-interface1] ospf dr-priority 0
[SwitchB-Vlan-interface1] quit
```

配置 Switch C。

```
[SwitchC] interface vlan-interface 1
[SwitchC-Vlan-interface1] ospf dr-priority 2
[SwitchC-Vlan-interface1] quit
```

查看 Switch D 的邻居信息。

```
<SwitchD> 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(Vlan-interface1)'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
   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
   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
   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 Switch D 的进程。

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

查看 Switch D 的邻居信息。

```
<SwitchD> 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(Vlan-interface1)'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
   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
   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
   Dead timer due in 39  sec
   Neighbor is up for 00:01:41
   Authentication Sequence: [ 0 ]
```

可以看到 Switch A 成为 DR，Switch C 为 BDR。

说明

- 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。

```
[SwitchA] 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

[SwitchB] 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。

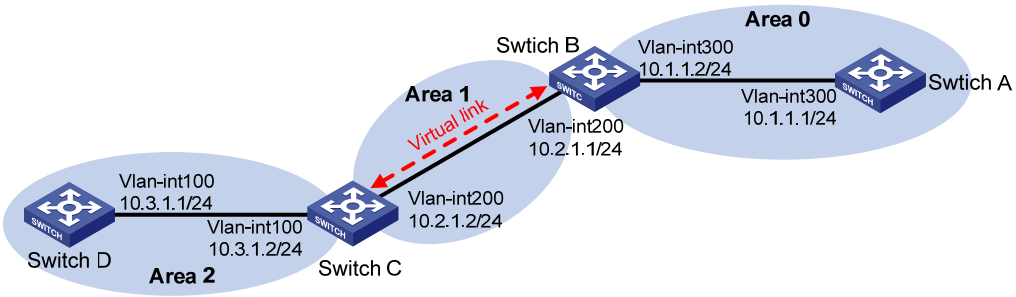
1.12.7 配置OSPF虚连接

1. 组网需求

- Area 2 与 Area 0 没有直接相连。Area 1 被用作传输区域（Transit Area）来连接 Area 2 和 Area 0。Switch B 和 Switch C 之间配置一条虚连接。
- 配置完成后，Switch B 能够学到 Area 2 中的路由。

2. 组网图

图1-28 配置 OSPF 虚链路组网图



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 Switch A。

```
<SwitchA> system-view
[SwitchA] ospf 1 router-id 1.1.1.1
```

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

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ospf 1 router-id 2.2.2.2
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] area 1
[SwitchB–ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[SwitchB–ospf-1-area-0.0.0.1] quit
[SwitchB-ospf-1] quit
```

配置 Switch C。

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

配置 Switch D。

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

查看 Switch B 的 OSPF 路由表。

```
[SwitchB] 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 Area 0 没有与 Area 2 直接相连，所以 Switch B 的路由表中没有 Area 2 的路由。

(3) 配置虚连接

配置 Switch B。

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

配置 Switch C。

```
[SwitchC] ospf 1
[SwitchC-ospf-1] area 1
[SwitchC-ospf-1-area-0.0.0.1] vlink-peer 2.2.2.2
[SwitchC-ospf-1-area-0.0.0.1] quit
```

查看 Switch B 的 OSPF 路由表。

```
[SwitchB] 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，Switch B 已经学到了 Area 2 的路由 10.3.1.0/24。

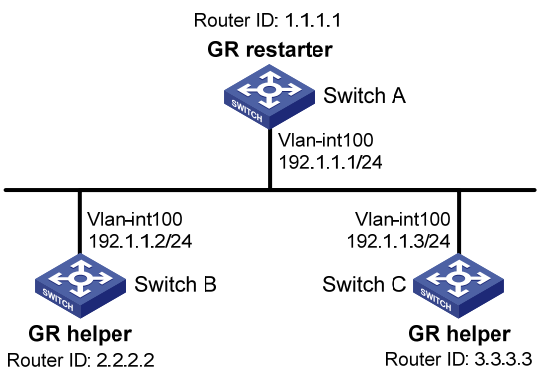
1.12.8 OSPF GR配置举例

1. 组网需求

- Switch A、Switch B 和 Switch 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。
- Switch A 作为非 IETF 标准 GR Restarter，Switch B 和 Switch C 作为 GR Helper 并且通过 GR 机制与 Switch A 保持带外同步。

2. 组网图

图1-29 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 Switch A。

```
<SwitchA> system-view
[SwitchA] router id 1.1.1.1
[SwitchA] ospf 100
[SwitchA-ospf-100] area 0
[SwitchA-ospf-100-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[SwitchA-ospf-100-area-0.0.0.0] quit
```

配置 Switch B。

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

配置 Switch C。

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

(3) 配置 OSPF GR

配置 Switch 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。

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

配置 Switch 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。

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

配置 Switch 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。

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

(4) 检验配置效果

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

```
<SwitchA> debugging ospf event graceful-restart
<SwitchA> terminal monitor
```

```
<SwitchA> terminal debugging
<SwitchA> reset ospf 100 process graceful-restart
Warning : Reset OSPF process? [Y/N]:y
%Dec 12 09:36:12:500 2006 SwitchA RM/3/RMLOG:OSPF-NBRCHANGE: Process 100, Neighbour
192.1.1.1(Vlan100) from Full to Down
OSPF 100: Intf 192.1.1.1 Rcv InterfaceDown State BackupDR -> Down.
OSPF 100 nonstandard GR Started for OSPF Router
OSPF 100 notify RM that OSPF process will enter GR.
OSPF 100 created GR wait timer, timeout interval is 40(s).
OSPF 100 created GR Interval timer,timeout interval is 120(s).
OSPF 100: Intf 192.1.1.1 Rcv InterfaceUp State Down -> Waiting.
OSPF 100: Intf 192.1.1.1 Rcv BackupSeen State Waiting -> BackupDR.
OSPF 100 created OOB Progress timer for neighbor 192.1.1.2.
OSPF 100 restarted OOB Progress timer for neighbor 192.1.1.2.
OSPF 100 restarted OOB Progress timer for neighbor 192.1.1.2.
%Oct 22 09:36:12:566 2008 SwitchA RM/3/RMLOG:OSPF-NBRCHANGE: Process 100, Neighbour
192.1.1.2(Vlan100) from Loading to Full
OSPF 100 restarted OOB Progress timer for neighbor 192.1.1.2.
OSPF 100 deleted OOB Progress timer for neighbor 192.1.1.2.
OSPF 100 Gr Wait Timeout timer fired.
OSPF 100 deleted GR wait timer.
OSPF 100 deleted GR Interval timer.
OSPF 100 GR Completed for OSPF Router
OSPF 100 notified RM that OSPF process left GR.
RM notified that all protocol left GR.
OSPF 100 started flushing STALE LSA after all protocol left GR.
OSPF 100: Flush Stale Area LSAs
OSPF 100: Start Flush Stale ASE + NSSA LSAs
OSPF 100: End Flush Stale ASE + NSSA LSAs
```

从上面的信息可以看出 Switch A 在 Switch B 的协助下完成了 GR。

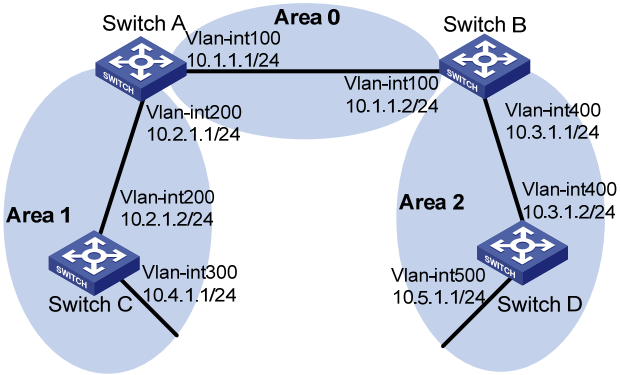
1.12.9 配置路由过滤

1. 组网需求

- 所有的交换机都运行 OSPF，整个自治系统划分为 3 个区域。
- 其中 Switch A 和 Switch B 作为 ABR 来转发区域之间的路由。
- 在 Switch C 上配置为 ASBR 引入外部路由（静态路由），并在 Switch C 上配置过滤策略，对引入的一条路由（3.1.3.0/24）进行过滤。
- 在 Switch A 上配置路由策略，对路由（10.5.1.0/24）进行过滤。

2. 组网图

图1-30 配置路由过滤



3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置OSPF（同前例 1.12.1 ）

(3) 配置引入自治系统外部路由

在 Switch C 上配置一条到目的网段 3.1.1.0/24 的静态路由。

```
<SwitchC> system-view
[SwitchC] ip route-static 3.1.1.0 24 10.4.1.2
```

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

```
[SwitchC] ip route-static 3.1.2.0 24 10.4.1.2
```

在 Switch C 上配置一条到目的网段 3.1.3.0/24 的静态路由。

```
[SwitchC] ip route-static 3.1.3.0 24 10.4.1.2
```

在 Switch C 上配置 OSPF 引入静态路由。

```
[SwitchC] ospf 1
[SwitchC-ospf-1] import-route static
[SwitchC-ospf-1] quit
```

在 Switch A 上查看路由信息。

```
<SwitchA> display ip routing-table
Routing Tables: Public
         Destinations : 12       Routes : 12

Destination/Mask    Proto   Pre  Cost         NextHop         Interface

3.1.1.0/24          O_ASE   150  1            10.2.1.2        Vlan200
3.1.2.0/24          O_ASE   150  1            10.2.1.2        Vlan200
3.1.3.0/24          O_ASE   150  1            10.2.1.2        Vlan200
10.1.1.0/24         Direct  0    0            10.1.1.1        Vlan200
10.1.1.1/32         Direct  0    0            127.0.0.1       InLoop0
10.2.1.0/24         Direct  0    0            10.2.1.1        Vlan200
10.2.1.1/32         Direct  0    0            127.0.0.1       InLoop0
10.3.1.0/24         OSPF    10   4            10.1.1.2        Vlan100
10.4.1.0/24         OSPF    10   13           10.2.1.2        Vlan200
```

```
10.5.1.0/24         OSPF   10   14           10.1.1.2        Vlan100
127.0.0.0/8         Direct 0    0            127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
```

(4) 在 Switch C 配置对路由 3.1.3.0/24 进行过滤

配置 IPv4 地址前缀列表。

```
[SwitchC] ip ip-prefix prefix1 index 1 deny 3.1.3.0 24
[SwitchC] ip ip-prefix prefix1 index 2 permit 3.1.1.0 24
[SwitchC] ip ip-prefix prefix1 index 3 permit 3.1.2.0 24
```

配置对引入的静态路由信息进行过滤，过滤掉路由 3.1.3.0/24。

```
[SwitchC] ospf 1
[SwitchC-ospf-1] filter-policy ip-prefix prefix1 export static
```

在 Switch A 上查看路由信息。

```
<SwitchA> display ip routing-table
Routing Tables: Public
         Destinations : 11       Routes : 11

Destination/Mask    Proto  Pre  Cost         NextHop         Interface

3.1.1.0/24          O_ASE  150  1            10.2.1.2        Vlan200
3.1.2.0/24          O_ASE  150  1            10.2.1.2        Vlan200
10.1.1.0/24         Direct 0    0            10.1.1.1        Vlan100
10.1.1.1/32         Direct 0    0            127.0.0.1       InLoop0
10.2.1.0/24         Direct 0    0            10.2.1.1        Vlan200
10.2.1.1/32         Direct 0    0            127.0.0.1       InLoop0
10.3.1.0/24         OSPF   10   4            10.1.1.2        Vlan100
10.4.1.0/24         OSPF   10   13           10.2.1.2        Vlan200
10.5.1.0/24         OSPF   10   14           10.1.1.2        Vlan100
127.0.0.0/8         Direct 0    0            127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0            127.0.0.1       InLoop0
```

可以看到，到目的网段 3.1.3.0/24 的路由被过滤掉了。

(5) 在 Switch A 上配置对路由 10.5.1.0/24 进行过滤

在 Switch A 上配置访问控制列表。

```
<SwitchA> system-veiw
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule 0 deny source 10.5.1.0 0.0.0.255
[SwitchA-acl-basic-2000] rule 1 permit source any
[SwitchA-acl-basic-2000] quit
```

配置对通过 LSA 计算出来的路由信息 10.5.1.0/24 进行过滤。

```
[SwitchA] ospf 1
[SwitchA-ospf-1] filter-policy 2000 import
[SwitchA-ospf-1] quit
```

在 Switch A 上查看路由信息。

```
[SwitchA] display ip routing-table
Routing Tables: Public
         Destinations : 10       Routes : 10
```

```
Destination/Mask    Proto   Pre  Cost        NextHop         Interface

3.1.1.0/24          O_ASE   150  1           10.2.1.2        Vlan200
3.1.2.0/24          O_ASE   150  1           10.2.1.2        Vlan200
10.1.1.0/24         Direct  0    0           10.1.1.1        Vlan100
10.1.1.1/32         Direct  0    0           127.0.0.1       InLoop0
10.2.1.0/24         Direct  0    0           10.2.1.1        Vlan200
10.2.1.1/32         Direct  0    0           127.0.0.1       InLoop0
10.3.1.0/24         OSPF    10   4           10.1.1.2        Vlan100
10.4.1.0/24         OSPF    10   13          10.2.1.2        Vlan200
127.0.0.0/8         Direct  0    0           127.0.0.1       InLoop0
127.0.0.1/32        Direct  0    0           127.0.0.1       InLoop0
```

可以看到，到 10.5.1.0/24 的路由被过滤掉了。

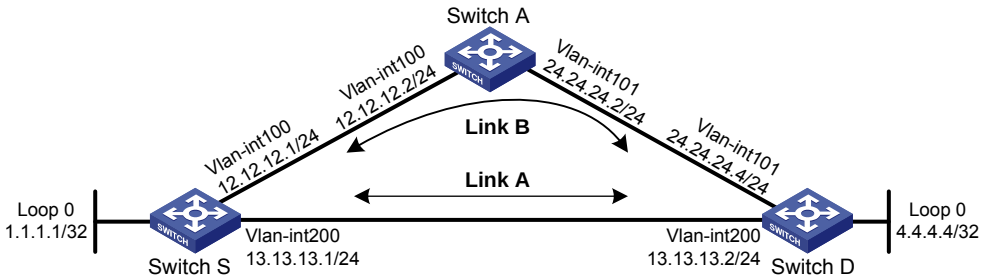
1.12.10 OSPF快速重路由配置举例

1. 组网需求

如 图 1-31 所示，Switch S 、Switch A和Switch D属于同一OSPF区域，通过OSPF协议实现网络互连。要求当Switch S和Switch 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1-31 OSPF 快速重路由配置举例



3. 配置步骤

(1) 配置各交换机接口的 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。

配置各交换机之间采用 OSPF 协议进行互连，确保 Switch S、Switch A 和 Switch D 之间能够在网络层互通，并且各交换机之间能够借助 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，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 Switch S 和 Switch D 的 OSPF 协议的自动计算快速重路由能力

配置 Switch S。

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 2.2.2.2
```

```
[SwitchS] ospf 1
[SwitchS-ospf-1] fast-reroute auto
[SwitchS-ospf-1] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] bfd echo-source-ip 3.3.3.3
[SwitchD] ospf 1
[SwitchD-ospf-1] fast-reroute auto
[SwitchD-ospf-1] quit
```

方法二：使能 Switch S 的 OSPF 协议的指定路由策略快速重路由能力

配置 Switch S。

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 1.1.1.1
[SwitchS] ip ip-prefix abc index 10 permit 4.4.4.4 32
[SwitchS] route-policy frr permit node 10
[SwitchS-route-policy] if-match ip-prefix abc
[SwitchS-route-policy] apply fast-reroute backup-interface vlan-interface 100
backup-nexthop 12.12.12.2
[SwitchS-route-policy] quit
[SwitchS] ospf 1
[SwitchS-ospf-1] fast-reroute route-policy frr
[SwitchS-ospf-1] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] bfd echo-source-ip 4.4.4.4
[SwitchD] ip ip-prefix abc index 10 permit 1.1.1.1 32
[SwitchD] route-policy frr permit node 10
[SwitchD-route-policy] if-match ip-prefix abc
[SwitchD-route-policy] apply fast-reroute backup-interface vlan-interface 101
backup-nexthop 24.24.24.2
[SwitchD-route-policy] quit
[SwitchD] ospf 1
[SwitchD-ospf-1] fast-reroute route-policy frr
[SwitchD-ospf-1] quit
```

(3) 检验配置效果

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

```
[SwitchS] display ip routing-table 4.4.4.4 verbose
Routing Table : Public
Summary Count : 1

  Destination: 4.4.4.4/32
     Protocol: OSPF            Process ID: 1
   Preference: 10                    Cost: 1
 IpPrecedence:                    QosLcId:
      NextHop: 13.13.13.2       Interface: Vlan-interface200
    BkNextHop: 12.12.12.2     BkInterface: Vlan-interface100
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
```

```
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv            Age: 00h01m27s
          Tag: 0
```

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

```
[SwitchD] display ip routing-table 1.1.1.1 verbose
Routing Table : Public
Summary Count : 1

  Destination: 1.1.1.1/32
     Protocol: OSPF            Process ID: 1
   Preference: 10                    Cost: 1
 IpPrecedence:                   QosLcId:
      NextHop: 13.13.13.1       Interface: Vlan-interface200
    BkNextHop: 24.24.24.2     BkInterface: Vlan-interface101
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv            Age: 00h01m27s
          Tag: 0
```

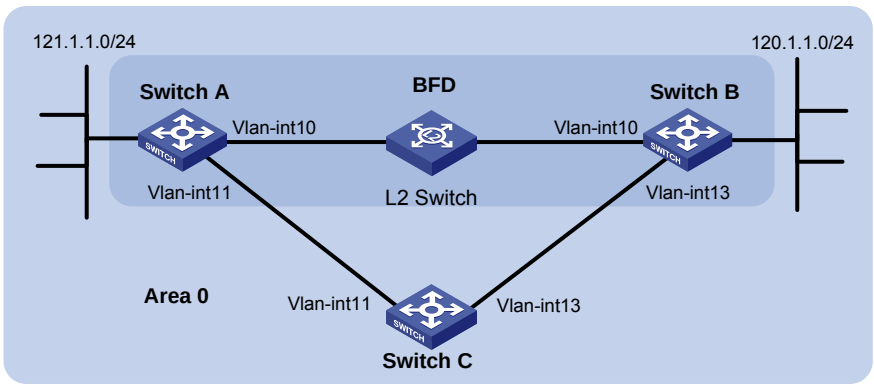
1.12.11 配置OSPF与BFD联动

1. 组网需求

- Switch A、Switch B 和 Switch C 上运行 OSPF，网络层相互可达。
- 当 Switch A 和 Switch 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 Switch C 进行通信。

2. 组网图

图1-32 OSPF 与 BFD 联动配置组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int10	10.1.0.102/24	Switch B	Vlan-int10	10.1.0.100/24
	Vlan-int11	11.1.1.1/24		Vlan-int13	13.1.1.1/24
Switch C	Vlan-int11	11.1.1.2/24			
	Vlan-int13	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 Switch A。

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 10.1.0.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network 11.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network 121.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
[SwitchA] interface vlan 11
[SwitchA-Vlan-interface11] ospf cost 2
[SwitchA-Vlan-interface11] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 10.1.0.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network 13.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network 120.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
[SwitchB] interface vlan-interface 13
[SwitchB-Vlan-interface13] ospf cost 2
[SwitchA-Vlan-interface13] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 11.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 13.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

(3) 配置 BFD 功能

在 Switch A 上使能 BFD 检测功能，并配置 BFD 参数。

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] ospf bfd enable
[SwitchA-Vlan-interface10] bfd min-transmit-interval 500
[SwitchA-Vlan-interface10] bfd min-receive-interval 500
[SwitchA-Vlan-interface10] bfd detect-multiplier 7
[SwitchA-Vlan-interface10] quit
[SwitchA] quit
```

在 Switch B 上使能 BFD 检测功能，并配置 BFD 参数。

```
[SwitchB] bfd session init-mode active
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ospf bfd enable
[SwitchB-Vlan-interface10] bfd min-transmit-interval 500
[SwitchB-Vlan-interface10] bfd min-receive-interval 500
[SwitchB-Vlan-interface10] bfd detect-multiplier 6
```

(4) 检查配置结果

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

显示 Switch A 的 BFD 信息。

```
<SwitchA> display bfd session
 Total Session Num: 1            Init Mode: Active
 Session Working Under Ctrl Mode:
 LD/RD         SourceAddr      DestAddr        State Holdtime Interface
 3/1           10.1.0.102      10.1.0.100      Up    1700ms   vlan10
```

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

```
<SwitchA> display ip routing-table 120.1.1.0 verbose
Routing Table : Public
Summary Count : 1
  Destination: 120.1.1.0/24
     Protocol: OSPF            Process ID: 0
   Preference: 0                     Cost: 2
 IpPrecedence:                   QosLcId:
      NextHop: 10.1.0.100    Interface: Vlan-interface10
    BkNextHop: 0.0.0.0        BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                 Label: NULL
  BKTunnel ID: 0x0               BKLabel: NULL
        State: Active Adv            Age: 00h58m10s
          Tag: 0
```

当 Switch A 和 Switch B 通过 L2 Switch 通信的链路出现故障时， Switch A 能够快速感知 Switch B 的变化。

查看 Switch A 的 BFD 信息。此时 Switch A 和 Switch B 的 BFD 会话已经被删除，没有任何输出信息。

```
<SwitchA> display bfd session
```

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

```
<SwitchA> display ip routing-table 120.1.1.0 verbose
Routing Table : Public
Summary Count : 1
  Destination: 120.1.1.0/24
     Protocol: OSPF            Process ID: 1
   Preference: 10                    Cost: 4
 IpPrecedence:                   QosLcId:
      NextHop: 11.1.1.2       Interface: Vlan-interface11
    BkNextHop: 0.0.0.0        BkInterface:
```

```
RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
  Tunnel ID: 0x0                  Label: NULL
BKTunnel ID: 0x0                BKLabel: NULL
      State: Active Adv             Age: 00h58m10s
        Tag: 0
```

1.13 常见配置错误举例

1.13.1 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。

1.13.2 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配置 ····· 1-1
1.1 IS-IS简介 ····· 1-1
1.1.1 基本概念 ····· 1-1
1.1.2 IS-IS区域 ····· 1-3
1.1.3 IS-IS的网络类型 ····· 1-5
1.1.4 IS-IS的PDU格式 ····· 1-6
1.1.5 支持的IS-IS特性 ····· 1-13
1.1.6 协议规范 ····· 1-15
1.2 IS-IS配置任务简介 ····· 1-16
1.3 配置IS-IS基本功能 ····· 1-17
1.3.1 配置准备 ····· 1-17
1.3.2 使能IS-IS功能 ····· 1-17
1.3.3 配置路由器类型和接口的链路邻接关系类型 ····· 1-17
1.3.4 配置接口网络类型 ····· 1-18
1.4 配置IS-IS路由信息控制 ····· 1-19
1.4.1 配置准备 ····· 1-19
1.4.2 配置IS-IS链路开销值 ····· 1-19
1.4.3 配置IS-IS路由优先级 ····· 1-20
1.4.4 配置IS-IS最大等价路由条数 ····· 1-21
1.4.5 配置IS-IS路由聚合 ····· 1-21
1.4.6 配置IS-IS发布缺省路由 ····· 1-22
1.4.7 配置IS-IS引入外部路由 ····· 1-22
1.4.8 配置IS-IS路由过滤 ····· 1-23
1.4.9 配置IS-IS路由渗透 ····· 1-24
1.5 调整和优化IS-IS网络 ····· 1-24
1.5.1 配置准备 ····· 1-24
1.5.2 配置Hello/CSNP报文发送时间间隔 ····· 1-25
1.5.3 配置Hello报文失效数目 ····· 1-25
1.5.4 配置DIS优先级 ····· 1-26
1.5.5 禁止接口发送和接收IS-IS报文 ····· 1-26
1.5.6 配置接口发送小型Hello报文 ····· 1-26
1.5.7 配置LSP参数 ····· 1-27
1.5.8 配置SPF参数 ····· 1-30

1.5.9 配置优先级参数 ····· 1-30
1.5.10 配置LSDB过载标志位 ····· 1-31
1.5.11 配置IS-IS主机名映射 ····· 1-31
1.5.12 配置邻接状态输出 ····· 1-32
1.6 提高IS-IS网络的安全性 ····· 1-33
1.6.1 配置准备 ····· 1-33
1.6.2 配置邻居关系验证 ····· 1-33
1.6.3 配置区域验证 ····· 1-34
1.6.4 配置路由域验证 ····· 1-34
1.7 配置IS-IS GR ····· 1-35
1.8 配置IS-IS NSR ····· 1-36
1.9 配置IS-IS快速重路由 ····· 1-36
1.10 使能IS-IS的TRAP功能 ····· 1-38
1.11 配置IS-IS和MIB绑定 ····· 1-38
1.12 配置IS-IS与BFD联动 ····· 1-39
1.13 IS-IS显示和维护 ····· 1-39
1.14 IS-IS典型配置案例 ····· 1-40
1.14.1 IS-IS基本配置 ····· 1-40
1.14.2 配置IS-IS的DIS选择 ····· 1-44
1.14.3 配置IS-IS引入外部路由 ····· 1-49
1.14.4 IS-IS GR配置举例 ····· 1-52
1.14.5 IS-IS NSR配置举例 ····· 1-54
1.14.6 IS-IS快速重路由配置举例 ····· 1-56
1.14.7 IS-IS验证配置举例 ····· 1-59
1.14.8 配置IS-IS与BFD联动 ····· 1-61

1 IS-IS配置

说明

在以下路由协议的介绍中所指的路由器及路由器图标，代表了一般意义下的路由器以及运行了路由协议的以太网交换机。

1.1 IS-IS简介

IS-IS（Intermediate System-to-Intermediate System，中间系统到中间系统）最初是国际标准化组织（International Organization for Standardization，ISO）为它的无连接网络协议（ConnectionLess Network Protocol，CLNP）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 属于内部网关协议（Interior Gateway Protocol，IGP），用于自治系统内部。IS-IS 是一种链路状态协议，使用最短路径优先（Shortest Path First，SPF）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

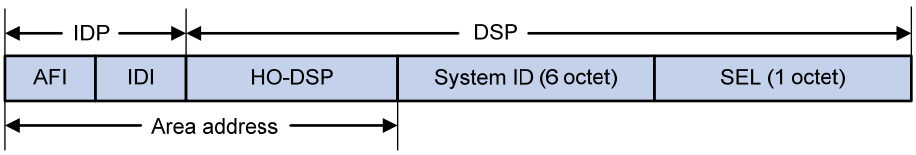
如 图 1-1 所示，NSAP由IDP（Initial Domain Part）和DSP（Domain Specific Part）组成。IDP相当于IP地址中的主网络号，DSP相当于IP地址中的子网号和主机地址。

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，因此，它们一起被称为区域地址（Area Address）。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（NSAP Selector，有时也写成 N-SEL）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。

(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。

(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。

说明

- 属于不同区域的 Level-1 路由器不能形成邻居关系。
- Level-2 路由器是否形成邻居关系则与区域无关。

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

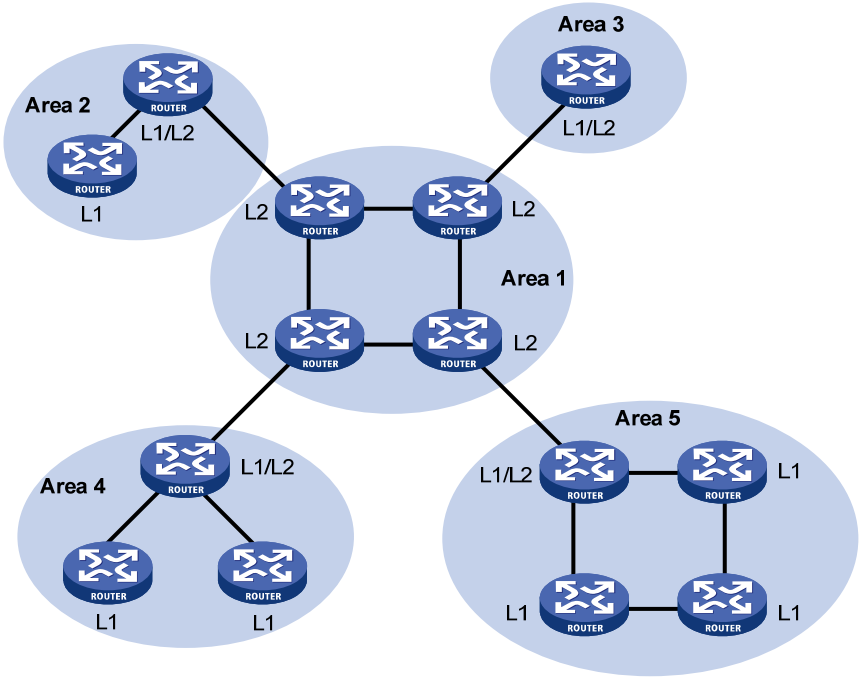
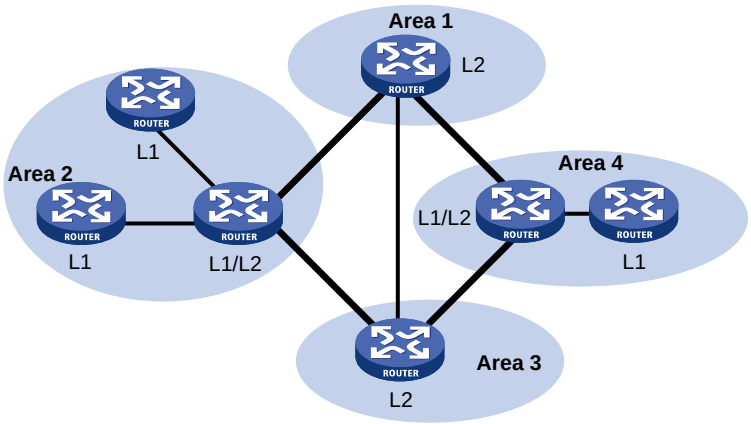


图 1-3 是IS-IS的另外一种拓扑结构图。其中Level-1-2 路由器不仅仅用来连接Level-1 和Level-2 路由器，而且还与其它Level-2 路由器一起构成了IS-IS的骨干网。在这个拓扑中，并没有规定哪个区域是骨干区域。所有Level-2 路由器和Level-1-2 路由器构成了IS-IS的骨干网，它们可以属于不同的区域，但必须是物理连续的。

图1-3 IS-IS 拓扑结构图之二



说明

IS-IS 的骨干网（Backbone）指的不是一个特定的区域。

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-1-2 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。

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。

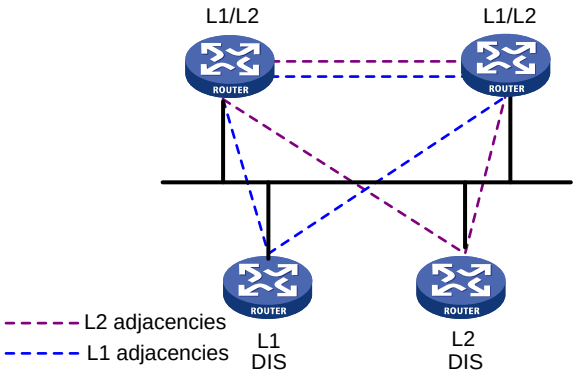
Level-1 和 Level-2 的 DIS 是分别选举的，用户可以为不同级别的 DIS 选举设置不同的优先级。DIS 优先级数值越高，被选中的可能性就越大。如果优先级最高的路由器有多台，则其中 SNPA（Subnetwork Point of Attachment，子网连接点）地址（广播网络中的 SNPA 地址是 MAC 地址）最大的路由器会被选中。不同级别的 DIS 可以是同一台路由器，也可以是不同的路由器。

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的PDU格式

1. PDU头格式

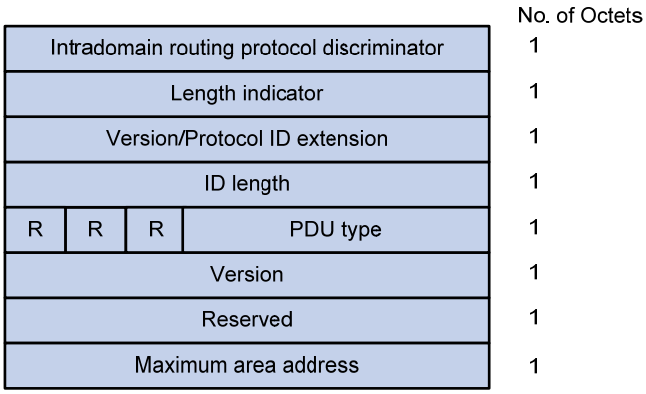
IS-IS报文是直接封装在数据链路层的帧结构中的。PDU（Protocol Data Unit，协议数据单元）可以分为两个部分，报文头和变长字段部分。其中报文头又可分为通用报头和专用报头。对于所有PDU来说，通用报头都是相同的，但专用报头根据PDU类型不同而有所差别，如 图 1-5 所示。

图1-5 PDU 格式

2. 通用报头格式

所有的PDU都有相同的通用报头格式，如 图 1-6 所示。

图1-6 PDU 头格式



主要字段的解释如下：

- Intradomain Routing Protocol Discriminator（域内路由协议鉴别符）：设置为 0x83。
- Length Indicator（长度标识符）：PDU 头部的长度（包括通用报头和专用报头），以字节为单位。
- Version/Protocol ID Extension（版本/协议标识扩展）：设置为 1（0x01）。
- ID Length（标识长度）：NSAP 地址和 NET 的 ID 长度。
- R（Reserved，保留）：设置为 0。
- PDU Type（PDU类型）：详细信息请参考 表 1-1。
- Version（版本）：设置为 1（0x01）。
- Maximum Area Address（最大区域地址数）：支持的最大区域个数。

表1-1 PDU 类型对应关系表

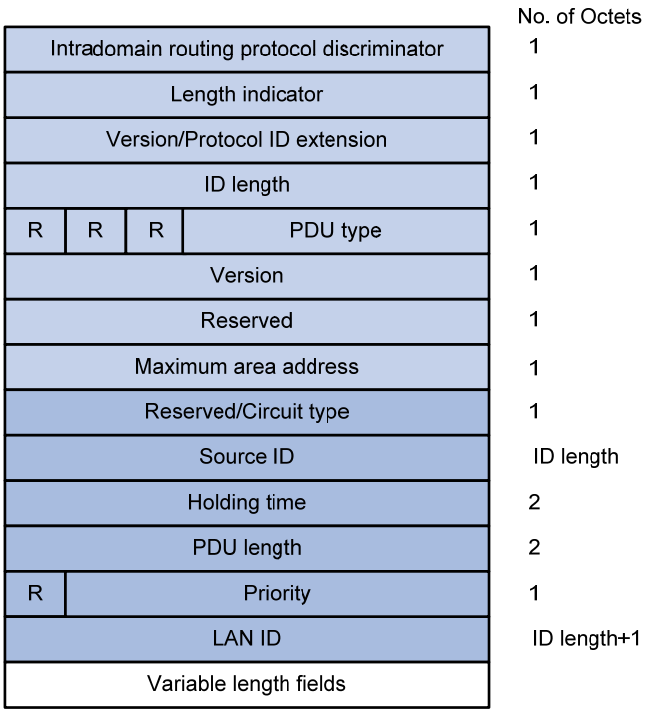
类型值	PDU 类型	简称
15	Level-1 LAN IS-IS Hello PDU	L1 LAN IIH
16	Level-2 LAN IS-IS Hello PDU	L2 LAN IIH
17	Point-to-Point IS-IS Hello PDU	P2P IIH
18	Level-1 Link State PDU	L1 LSP
20	Level-2 Link State PDU	L2 LSP
24	Level-1 Complete Sequence Numbers PDU	L1 CSNP
25	Level-2 Complete Sequence Numbers PDU	L2 CSNP
26	Level-1 Partial Sequence Numbers PDU	L1 PSNP
27	Level-2 Partial Sequence Numbers PDU	L2 PSNP

3. Hello

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。

它们的报文格式有所不同。广播网中的Hello报文格式如 图 1-7 所示（浅蓝色部分是通用报文头）。

图1-7 L1/L2 LAN IIH 格式

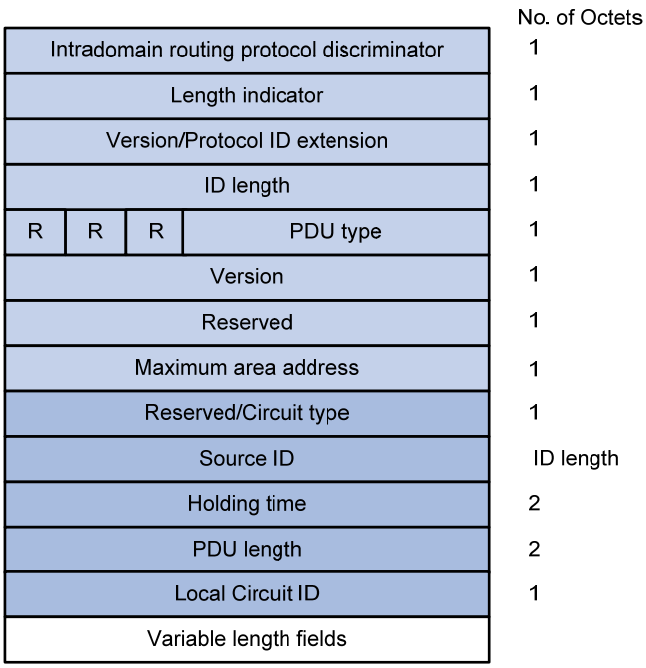


主要字段的解释如下：

- Reserved/Circuit Type：高位的 6 比特保留，值为 0。低位的 2 比特表示路由器的类型（00 保留，01 表示 L1，10 表示 L2，11 表示 L1/2）。
- Source ID：发送 Hello 报文的路由器的 System ID。
- Holding Time：保持时间。在此时间内如果没有收到邻居发来的 Hello 报文，则中止已建立的邻居关系。
- PDU Length：PDU 的总长度，以字节为单位。
- Priority：选举 DIS 的优先级。
- LAN ID：包括 System ID 和一字节的伪节点 ID。

点到点网络中的Hello报文格式如 图 1-8 所示（浅蓝色部分是通用报文头）。

图1-8 P2P IIH 格式



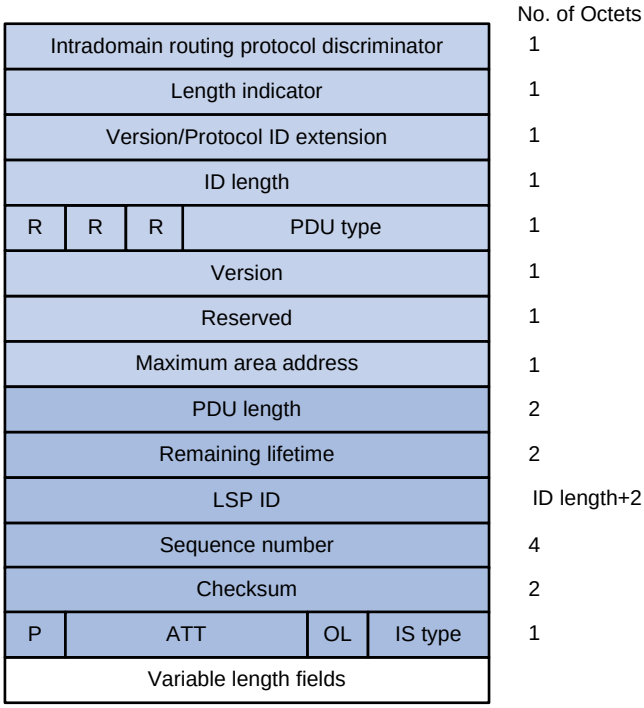
从图中可以看出，P2P IIH 中的多数字段与 LAN IIH 相同。不同的是没有 Priority 和 LAN ID 字段，而多了一个 Local Circuit ID 字段，表示本地链路 ID。

4. LSP报文格式

LSP 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。

两类LSP有相同的报文格式，如 图 1-9 所示（浅蓝色部分是通用报文头）。

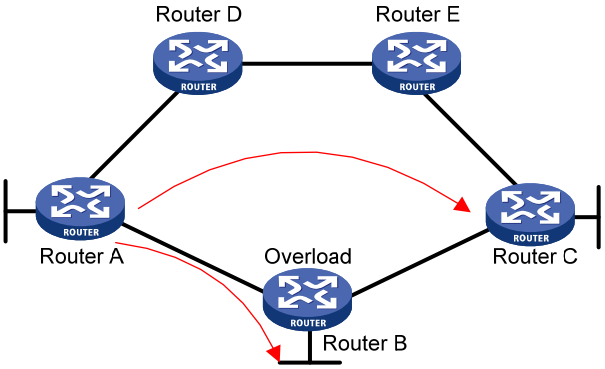
图1-9 L1/L2 LSP 格式



主要字段的解释如下：

- PDU Length：PDU 的总长度，以字节为单位。
- Remaining Lifetime：LSP 的存活时间，以秒为单位。
- LSP ID：由三部分组成，System ID、伪节点 ID（一字节）和 LSP 的分片号（一字节）。
- Sequence Number：LSP 的序列号。
- Checksum：LSP 的校验和。
- P（Partition Repair）：分区位。仅与 L2 LSP 有关，表示路由器是否支持自动修复区域分割。
- ATT（Attachment）：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。
- OL（LSDB Overload）：过载标志位。表示本路由器因内存不足而导致LSDB不完整。其它路由器在得知这一信息后，就不会再利用这台路由器转发需要经过它传送的数据流，但到此路由器直连地址的报文仍然可以被转发。如 图 1-10 所示，假设正常情况下Router A到Router C的报文都是经过Router B转发，但如果Router B的OL位置 1，则Router A会认为Router B的路由不完整，从而将报文通过Router D、Router E转发给Router C，但到Router B直连地址的报文不受影响。

图1-10 LSDB Overload 示意图



- IS Type：生成 LSP 的路由器的类型。

5. SNP格式

时序报文 SNP（Sequence Number PDUs）用于确认邻居之间最新接收的 LSP，作用类似于确认（Acknowledge）报文，但更有效。

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 的摘要信息，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 在邻居关系建立时会发送，只发送一次。

CSNP的报文格式如 图 1-11 所示（浅蓝色部分是通用报文头）。

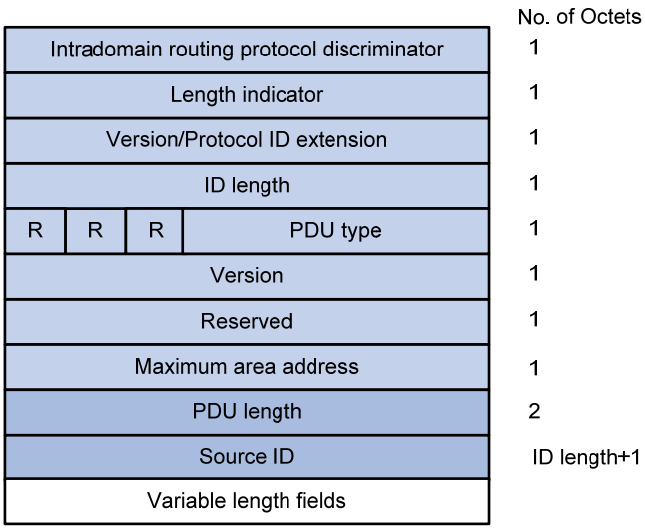
图1-11 L1/L2 CSNP 格式

	No. of Octets
Intradomain routing protocol discriminator	1
Length indicator	1
Version/Protocol ID extension	1
ID length	1
R R R PDU type	1
Version	1
Reserved	1
Maximum area address	1
PDU length	2
Source ID	ID length+1
Start LSP ID	ID length+2
End LSP ID	ID length+2
Variable length fields	

PSNP 只列举最近收到的一个或多个 LSP 的序号，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。

PSNP的报文格式如 图 1-12 所示：

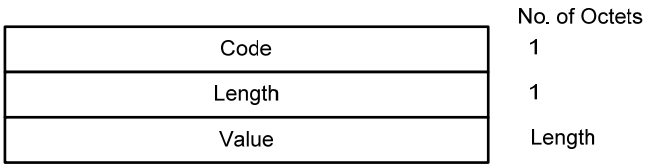
图1-12 L1/L2 PSNP 格式



6. 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 1-13 所示：

图1-13 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
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

CLV Code	名称	所应用的 PDU 类型
130	IP External Reachability Information	L2 LSP
131	Inter-Domain Routing Protocol Information	L2 LSP
132	IP Interface Address	IIH、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），其它几种 CLV 在 RFC 1195 中定义。

1.1.5 支持的IS-IS特性

1. 多实例和多进程

为了方便管理，提高控制效率，IS-IS 支持多进程和多实例特性。多进程允许为一个指定的 IS-IS 进程关联一组接口，从而保证该进程进行的所有协议操作都仅限于这一组接口。这样，就可以实现一台路由器有多个 IS-IS 协议进程，每个进程负责唯一的一组接口。

对于支持 VPN 的路由器，每个 IS-IS 进程都与一个指定的 VPN 实例相关联。这样，所有附加到该进程的接口都应与该进程相关联的 VPN 实例相关联。

2. IS-IS GR

GR（Graceful Restart，平滑重启）是一种在协议重启或主备切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。

基于 IS-IS 的 GR Restarter 进行协议重启后，为了与其 IS-IS 邻居重新同步链路状态数据库，它必须完成下列两项任务：

- 在不改变邻接关系的前提下，重新获取网络中的有效 IS-IS 邻居信息；
- 重新获取网络链路状态数据库的内容。

IS-IS 协议重启完毕后，GR Restarter 会立即向邻接的 GR Helper 发送一个 IS-IS GR 信号。这样，IS-IS 邻居就不会复位与其的邻居关系。在收到其 IS-IS 邻居的响应后，GR Restarter 会重新恢复与其的邻居关系列表。

邻居关系重新建立后，GR Restarter 与其所有具备 GR 感知能力的 IS-IS 邻居之间同步数据库，并交换路由信息。交换完成后，GR Restarter 根据新的路由转发信息更新路由表和转发表，删除失效的路由，完成 IS-IS 协议收敛。

3. IS-IS NSR

NSR（Nonstop Routing，不间断路由）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 Master 设备备份到 Slave 设备，在 Master/Slave 切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。

4. IS-IS TE

IS-IS TE（Traffic Engineering，流量工程）支持建立和维护 TE 的 LSP（Label Switched Path，标签交换路径）。

MPLS 在构建 CR LSP（Constraint-based Routed LSP，基于约束路由的 LSP）时，需要了解本区域中所有链路的流量属性信息。它可以通过 IS-IS 来获取链路的流量工程信息。

说明

IS-IS TE 的详细配置请参见“MPLS 配置指导”中的“MPLS TE”。

5. 管理标记

管理标记简化了对路由信息的管理，该标记用来携带关于 IP 地址前缀的管理信息，利用它可以控制对各种外部路由协议的引入，以及承载 BGP 的标准或扩展团体属性。

6. LSP分片扩展

IS-IS 通过泛洪 LSP 来宣告链路状态信息，由于一个 LSP 能够承载的信息量有限，IS-IS 将对 LSP 进行分片。每个 LSP 分片由产生该 LSP 的结点或伪结点的 SystemID、PseudonodeID（普通 LSP 为 0，Pseudonode LSP 为非 0）、LSPNumber（LSP 分片号）组合起来唯一标识，由于 LSPNumber 字段的长度是 1 字节，因此，IS-IS 路由器可产生的分片数最大为 256，限制了 IS-IS 路由器可以发布的链路信息量。

IS-IS LSP 分片扩展特性可使 IS-IS 路由器生成更多的 LSP 分片，通过为路由器配置附加的虚拟系统、每个虚拟系统都可生成 256 个 LSP 分片（最多可配置 50 个虚拟系统），使得 IS-IS 路由器可最多生成 13056 个 LSP 分片。

(1) 常用术语

- 初始系统（Originating System）：实际运行 IS-IS 协议的路由器。使能 IS-IS LSP 分片扩展功能后，可以为路由器配置附加的虚拟系统，“Originating System”指的是原来运行的、“真正”的 IS-IS 进程。
- 系统 ID（System ID）：初始系统的系统 ID。
- 附加系统 ID（Additional System ID）：使能 IS-IS LSP 分片扩展功能后，为 IS-IS 路由器配置的附加的虚拟 SystemID。每个附加系统 ID 都可以生成 256 个额外扩展的 LSP 分片。附加系统 ID 和系统 ID 一样，在整个路由域中必须唯一。
- 虚拟系统（Virtual System）：由附加系统 ID 标识的系统，用来生成扩展 LSP 分片。
- 原始 LSP（Original LSP）：由初始系统生成的 LSP，在其 LSP ID 中指定的 System ID 为初始系统的系统 ID。
- 扩展 LSP（Extended LSP）：由虚拟系统生成的 LSP，在其 LSP ID 中指定的 System ID 为附加系统 ID。

通过配置附加系统 ID，IS-IS 路由器可以在扩展分片 LSP 中宣告更多的链路状态信息，可以把虚拟系统看作一个虚拟的路由器，扩展分片 LSP 就是由附加系统 ID 标识的虚拟系统发布的。

(2) 操作模式

IS-IS 路由器运行 LSP 分片扩展特性的操作模式有两种：

- Mode-1：当网络中的部分路由器不支持 LSP 分片扩展特性时，使用工作模式 1。在该模式下，初始系统与每个虚拟系统都建立邻居关系，且初始系统到虚拟系统的链路开销为 0，因此，虚拟系统就好像网络中与初始系统相连的路由器，但只能通过初始系统才能到达虚拟系统，因此，不支持分片扩展特性的 IS-IS 路由器收到扩展分片 LSP 时无需做任何修改也能够正常工作，但是对通过虚拟系统发布的扩展 LSP 分片中的链路状态信息有所限制。

- Mode-2：当网络中所有路由器都支持 LSP 分片扩展特性时，推荐使用工作模式 2。在该模式下，网络中所有 IS-IS 路由器都知道虚拟系统生成的 LSP 属于的初始系统，对虚拟系统发布的 LSP 扩展分片的链路状态信息没有限制。

LSP 分片扩展特性的操作模式基于区域和路由层次配置，mode-1 可以向前兼容，支持分片扩展特性的路由器能够与不支持分片扩展特性的路由器协同工作；mode-2 要求处在同一区域且 Level 相同的路由器都支持分片扩展特性。

7. 动态主机名交换机制

动态主机名交换机制是为了方便对 IS-IS 网络的维护和管理而引入的，它为 IS-IS 路由器提供了一种从主机名到 System ID 映射的服务。这个动态的主机名信息在 LSP 中以一个动态主机名 CLV 的形式发布。

这个机制同时还提供将主机名与广播网中的 DIS 相关联的服务，并将此信息通过伪节点的 LSP 以动态主机名 CLV 的形式发布出去。

在维护和管理中，使用主机名比使用 System ID 会更直观，也更容易记忆。配置此功能后，当在路由器上使用 **display** 命令显示 IS-IS 相关信息时，看到的是路由器的主机名，而不再是 System ID。

8. IS-IS与BFD联动

说明

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”。

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。

1.1.6 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 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 3784：IS-IS extensions for Traffic Engineering
- RFC 3847：Restart signaling for IS-IS

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
	配置路由器类型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链路开销值	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/CSNP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DIS优先级	可选	1.5.4
	禁止接口发送和接收IS-IS报文	可选	1.5.5
	配置接口发送小型Hello报文	可选	1.5.6
	配置LSP参数	可选	1.5.7
	配置SPF参数	可选	1.5.8
	配置优先级参数	可选	1.5.9
	配置LSDB过载标志位	可选	1.5.10
	配置IS-IS主机名映射	可选	1.5.11
	配置邻接状态输出	可选	1.5.12
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快速重路由		可选	1.9

配置任务	说明	详细配置
使能IS-IS的Trap功能	可选	1.10
配置IS-IS和MIB绑定	可选	1.11
配置IS-IS与BFD联动	可选	1.12

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**	-
启动IS-IS路由进程，进入IS-IS视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	必选 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*	必选 缺省情况下，没有配置NET
退回至系统视图	**quit**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置指定接口上使能IS-IS路由进程	**isis enable** [*process-id*]	必选 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 配置路由器类型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 配置路由器类型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**	-
启动IS-IS路由进程，进入IS-IS视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
配置路由器类型	**is-level** { **level-1** | **level-1-2** | **level-2** }	可选 缺省情况下，IS-IS路由器的类型为**level-1-2**
退回至系统视图	**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**]	可选 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**	可选 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

说明

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链路开销值

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=（参考值÷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。

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*]	-
配置IS-IS开销值的类型	**cost-style** { **narrow** | **wide** | **wide-compatible** | { **compatible** | **narrow-compatible** } [**relax-spf-limit**] }	可选 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	**interface** *interface-type interface-number*	-

操作	命令	说明
配置IS-IS接口的链路开销值	**isis cost** *value* [**level-1** | **level-2**]	必选 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	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
配置IS-IS开销值的类型	**cost-style** { **narrow** | **wide** | **wide-compatible** | { **compatible** | **narrow-compatible** } [**relax-spf-limit**] }	可选 缺省情况下，IS-IS开销值的类型为narrow
全局配置IS-IS的链路开销值	**circuit-cost** *value* [**level-1** | **level-2**]	必选 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	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
配置IS-IS开销值的类型	**cost-style** { **wide** | **wide-compatible** }	必选 缺省情况下，IS-IS开销值的类型为narrow
使能自动计算接口链路开销值功能	**auto-cost enable**	必选 缺省情况下，自动计算接口链路开销值功能处于关闭状态
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*	可选 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视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
配置IS-IS路由优先级	**preference** { **route-policy** *route-policy-name* | *preference* } *	必选 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视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
配置在负载分担方式下IS-IS最大等价路由条数	**maximum load-balancing** *number*	必选 缺省值为8

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。

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视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
配置聚合路由	**summary** *ip-address* { *mask* | *mask-length* } [**avoid-feedback** | **generate_null0_route** | **tag** *tag* | [**level-1** | **level-1-2** | **level-2**]] *	必选 缺省情况下，没有对路由进行聚合

说明

- 聚合后路由的开销值取所有被聚合路由中最小的开销值。
- 路由器只对本地生成的 LSP 中的路由进行聚合。

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视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
配置IS-IS发布缺省路由	**default-route-advertise** [**route-policy** *route-policy-name* | [**level-1** | **level-1-2** | **level-2**]] *	必选 缺省情况下，此功能关闭

说明

该配置产生的缺省路由只被发布到同级别的路由器。通过使用路由策略，可以强制 IS-IS 只在路由表中有匹配的路由项时才生成缺省路由。

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视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-

操作	命令	说明
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**] [**cost** *cost* | **cost-type** { **external** | **internal** } | [**level-1** | **level-1-2** | **level-2**] | **route-policy** *route-policy-name* | **tag** *tag*] *	必选 缺省情况下，IS-IS不引入其它协议的路由信息 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
配置引入Level1/Level2的IPv4路由最大条数	**import-route limit** *number*	可选 本命令的缺省值与设备的型号有关，请以设备的实际情况为准

说明

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。

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，包括过滤通过接收到的 LSP 计算出来的路由信息和过滤从其它路由协议引入的路由信息。

1. 过滤通过接收到的LSP计算出来的路由信息

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，并使用 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。

通过 ACL、IP 地址前缀列表或路由策略可以对计算出来的将要加入到 IS-IS 路由表中的路由进行过滤，满足条件则加入到 IS-IS 路由表中，否则将不能加入到 IS-IS 路由表中。

表1-15 过滤通过接收到的 LSP 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IS-IS视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
过滤通过接收到的LSP计算出来的路由信息	**filter-policy** { *acl-number* | **ip-prefix** *ip-prefix-name* | **route-policy** *route-policy-name* } **import**	必选 缺省情况下，IS-IS不对通过接收到的LSP 计算出来的路由信息进行过滤

2. 过滤引入路由信息

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 过滤引入路由信息

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*]	-
过滤引入路由信息	**filter-policy** { *acl-number* | **ip-prefix** *ip-prefix-name* | **route-policy** *route-policy-name* } **export** [*protocol* [*process-id*]]	必选 缺省情况下，IS-IS不对引入的路由信息进行过滤

1.4.9 配置IS-IS路由渗透

通过 IS-IS 路由渗透功能，Level-1-2 路由器可以将它所知道的其它 Level-1 区域路由信息和 Level-2 区域路由信息发布到 Level-1 路由器。

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视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
使能IS-IS路由渗透	**import-route isis level-2 into level-1** [**filter-policy** { *acl-number* | **ip-prefix** *ip-prefix-name* | **route-policy** *route-policy-name* } | **tag** *tag*] *	必选 缺省情况下，Level-2区域的路由信息不向Level-1区域发布

说明

将 Level-2 区域的路由信息引入到 Level-1 区域时，需要注意：

- 如果指定了过滤策略，则只有通过过滤的路由才能够被发布到 Level-1 区域中。
- 如果要通过路由策略对从 Level-2 区域引入到 Level-1 区域的路由信息进行过滤，必须在 **import-route isis level-2 into level-1** 命令中同时指定要应用的路由策略，否则路由过滤将不会生效；其它路由策略，如在接收或引入路由时指定的路由策略对路由渗透无效。

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/CSNP报文发送时间间隔

表1-18 配置 Hello/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Hello报文的发送时间间隔	**isis timer hello** *seconds* [**level-1** | **level-2**]	可选 缺省情况下，Hello报文的发送时间间隔为10秒
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**]	可选 缺省情况下，CSNP报文的发送时间间隔为10秒

说明

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

当路由器在抑制时间内没有收到来自邻居的 Hello 报文时将宣告邻居失效，抑制时间等于 Hello 报文失效数目与 Hello 报文发送时间间隔的乘积，通过设置 Hello 报文失效数目和 Hello 报文的发送时间间隔，可以调整抑制时间，即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1-19 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**]	可选 缺省情况下，Hello报文失效数目为3

说明

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。

1.5.4 配置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1-20 配置 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**]	可选 缺省情况下，接口的DIS优先级为64

1.5.5 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1-21 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**	必选 缺省情况下，接口既发送也接收IS-IS报文

1.5.6 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1-22 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**	-

操作	命令	说明
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**	必选 缺省情况下，接口发送标准Hello报文

1.5.7 配置LSP参数

1. 配置LSP时间参数

(1) 配置 LSP 最大生存时间

每一个 LSP 都包含一个最大生存时间，当 LSP 驻留在 LSDB 中时，随着时间的推移，每一个 LSP 的最大生存时间将逐渐递减至 0，当 LSP 的最大生存时间为 0 时，IS-IS 将删掉这个过期的 LSP。用户可根据网络规模对 LSP 的最大生存时间进行调整。

表1-23 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*	可选 缺省情况下，LSP最大生存时间为1200秒

(2) 配置 LSP 生成时间间隔

每一个 LSP 都有一个最大生存时间，每个 LSP 都会随着时间的推移而被逐渐老化，因此每个路由器必须定时刷新自己生成的 LSP，以防止 LSP 的最大生存时间减小至 0；并通过定时刷新 LSP 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。

表1-24 配置 LSP 生成时间间隔

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*	可选 缺省情况下，LSP刷新周期为900秒

操作	命令	说明
配置LSP生成时间间隔	**timer lsp-generation** *maximum-interval* [*initial-interval* [*second-wait-interval*]] [**level-1** | **level-2**]	可选 缺省情况下，LSP生成时间间隔为2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。

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1-25 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可以最多发送的LSP的数目	**isis timer lsp** *time* [**count** *count*]	可选 缺省情况下，LSP的发送最小时间间隔为33毫秒，一次可以最多发送的LSP的数目为5
配置LSP在在点到点链路上的重传时间间隔	**isis timer retransmit** *seconds*	可选 缺省情况下，LSP报文在在点到点链路上的重传间隔为5秒

说明

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。

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1-26 配置 LSP 报文长度

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*]	-
配置生成的Level-1 LSP和Level-2 LSP的最大长度	**lsp-length originate** *size* [**level-1** | **level-2**]	可选 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*	可选 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1-27 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**]] *	必选 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1-28 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，并配置LSP分片扩展特性的工作模式	**lsp-fragments-extend** [[**level-1** | **level-1-2** | **level-2**] | [**mode-1** | **mode-2**]] *	必选 缺省情况下，LSP分片扩展功能处于关闭状态
配置IS-IS进程的虚拟系统ID	**virtual-system** *virtual-system-id*	必选 缺省情况下，没有配置IS-IS进程的虚拟系统ID

说明

使能 LSP 分片扩展功能时需要注意：

- IS-IS 进程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。
- 为了使路由器生成扩展 LSP 分片，应至少配置一个虚拟 System ID。一个 IS-IS 进程最多可配置 50 个虚拟 System ID。

1.5.8 配置SPF参数

IS-IS 协议中，当 LSDB 发生变化时需要进行路由计算。频繁的路由计算会占用大量的系统资源，导致系统性能下降，而周期性地计算 SPF 可以在一定程度上提高效率，计算 SPF 的时间间隔可以由用户根据需要进行配置。

表1-29 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	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
配置IS-IS路由计算时间间隔	**timer spf** *maximum-interval* [*initial-interval* [*second-wait-interval*]]	可选 缺省情况下，IS-IS路由计算的时间间隔为10秒

1.5.9 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。路由收敛的优先级可以由用户根据需要进行配置。

表1-30 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*]	-
配置指定IS-IS路由为高优先级	**priority high** { **ip-prefix** *prefix-name* | **tag** *tag-value* }	可选 缺省情况下，没有配置IS-IS路由为高优先级，即不能快速收敛 IS-IS主机路由的收敛优先级低于高优先级，而高于其他IS-IS路由的收敛优先级

1.5.10 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1-31 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	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
设置过载标志位	**set-overload** [**on-startup** [[**start-from-nbr** *system-id* [*timeout1* [*nbr-timeout*]]] | *timeout2*] [**allow** { **external** | **interlevel** } *]	必选 缺省情况下，不设置过载标志位

1.5.11 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，点分十进制表示的 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，在路由器上使用 **display isis lsdb** 命令显示 IS-IS 链路状态数据库时，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1-32 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*	必选 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1-33 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*	必选 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*	可选 缺省情况下，没有配置本地局域网名称 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

1.5.12 配置邻接状态输出

表1-34 配置邻接状态输出

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*]	-
打开邻接状态输出开关	**log-peer-change**	必选 缺省情况下，邻接状态输出开关处于打开状态

说明

当打开邻接状态输出开关后，IS-IS 邻接状态的变化会输出到配置终端上，直至邻接状态输出开关被关闭。

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。

表1-35 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** } [**cipher**] *password* [**level-1** | **level-2**] [**ip** | **osi**]	必选 缺省情况下，接口没有配置邻居关系验证，既不会验证收到的Hello报文，也不会把验证密码插入到Hello报文中

说明

- 必须先使用 **isis enable** 命令使能该接口才能进行参数 **level-1** 和 **level-2** 的配置。
- 如果没有指定 **level-1** 或 **level-2** 参数，将同时为 **level-1** 和 **level-2** 的 Hello 报文配置验证方式及验证密码。
- 如果没有指定 **ip** 或 **osi** 参数，将检查 Hello 报文中 OSI 的相应字段的配置内容。

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。

表1-36 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*]	-
配置区域验证方式和验证密码	**area-authentication-mode** { **md5** | **simple** } [**cipher**] *password* [**ip** | **osi**]	必选 缺省情况下，系统没有配置区域验证，既不会验证收到的Level-1报文，也不会把验证密码插入到Level-1报文中

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。

表1-37 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** } [**cipher**] *password* [**ip** | **osi**]	必选 缺省情况下，系统没有配置路由域验证，既不会验证收到的Level-2报文，也不会把验证密码插入到Level-2报文中

1.7 配置IS-IS GR

说明

IS-IS GR 特性与 IS-IS NSR 特性互斥，不能同时配置。

在普通的路由协议重启的情况下，路由器与其它邻居的邻接关系被拆除。此时会引起网络暂时的中断。

基于 IS-IS 的 GR 可以解决这个问题：重启路由器将重启状态通知给邻居，允许邻居重新建立邻接关系而不终止连接。

基于 IS-IS 的 GR 特性有以下几个特点：

- 发生了协议重启时，应用了 GR 特性的设备可重新向邻居发送连接请求，而不会终止邻接关系。
- 在路由协议重启完毕后，邻居协助当前设备完成路由信息同步，在尽量短的时间内使该设备的各种路由信息恢复到重启前的状态。通过 GR，在整个协议重启过程中，网络路由和转发保持高度稳定。
- 对于第一次启动的路由器，在 LSP 报文中设置 OL（过载标记位）直到数据库同步，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。

分别在作为 GR Restarter 和 GR Helper 的设备上进行以下配置。

表1-38 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*]	必选 缺省情况下，系统没有运行IS-IS
使能IS-IS协议的GR能力	**graceful-restart**	必选 缺省情况下，IS-IS协议的GR能力处于关闭状态
配置IS-IS协议的GR重启间隔时间	**graceful-restart interval** *timer*	可选 缺省情况下，IS-IS协议的GR重启间隔时间为300秒 重启间隔时间用来指定路由器的重启时间。重启间隔时间在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
配置重启时抑制SA位	**graceful-restart suppress-sa**	可选 缺省情况下，重启时不抑制SA位 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

1.8 配置IS-IS NSR

说明

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 Master/Slave 切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 Master/Slave 切换后新 Master 设备上的 IS-IS 进程重新学习所有的路由，如果在 Master/Slave 切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 Master 设备备份到 Slave 设备，在发生 Master/Slave 切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1-39 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**	必选 缺省情况下，IS-IS协议的NSR功能处于关闭状态
配置NSR的时间间隔	**non-stop-routing interval** *interval-value*	可选 缺省情况下，NSR的时间间隔为0秒，即没有配置NSR的时间间隔

1.9 配置IS-IS快速重路由

说明

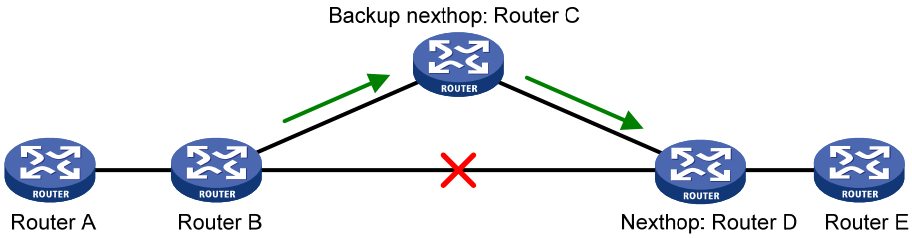
- IS-IS 支持快速重路由功能不能与 IS-IS 支持 BFD 监测同时使用，否则可能导致快速重路由功能失效。
- IS-IS 支持快速重路由自动计算备份下一跳功能与 IS-IS TE 特性互斥。

1. 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1-14 IS-IS 快速重路由功能示意图



如 图 1-14 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探测到网络故障时，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。

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

3. 配置IS-IS支持快速重路由功能（自动计算备份下一跳）

表1-40 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**	-
配置BFD Echo报文源地址	**bfd echo-source-ip** *ip-address*	必选 缺省情况下，没有配置BFD Echo报文源地址
进入IS-IS视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
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**	必选 缺省情况下，IS-IS支持快速重路由功能处于关闭状态

4. 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1-41 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**	-
配置BFD Echo报文源地址	**bfd echo-source-ip** *ip-address*	必选 缺省情况下，没有配置BFD Echo报文源地址
进入IS-IS视图	**isis** [*process-id*] [**vpn-instance** *vpn-instance-name*]	-
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*	必选 缺省情况下，IS-IS支持快速重路由功能处于关闭状态

1.10 使能IS-IS的TRAP功能

开启 IS-IS 模块的 Trap 功能后，该模块会生成 Trap 报文，用于报告该模块的重要事件。生成的 Trap 报文将被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 Trap 报文的输出规则（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1-42 使能 IS-IS 的 TRAP 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 的 TRAP 功能	**is-snmp-traps enable**	必选 缺省情况下，IS-IS的TRAP功能处于使能状态

1.11 配置IS-IS和MIB绑定

通过将 MIB 操作绑定在指定的 IS-IS 进程上，可以让指定的 IS-IS 进程接收 SNMP 请求。有关 MIB 的相关内容请参见“网络管理和监控配置指导”中的“SNMP”。

表1-43 配置 IS-IS 和 MIB 绑定

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和MIB绑定	**isis mib-binding** *process-id*	必选 缺省情况下，MIB操作绑定在IS-IS进程1上

1.12 配置IS-IS与BFD联动

表1-44 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*]	必选 缺省情况下，接口上不使能IS-IS路由进程
指定接口上使能BFD	**isis bfd enable**	必选 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

1.13 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 的数据库信息，或者复位特定邻居的数据信息。

表1-45 IS-IS 显示和维护

操作	命令
显示IS-IS的摘要信息	**display isis brief** [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS调试开关的状态	**display isis debug-switches** { *process-id* | **vpn-instance** *vpn-instance-name* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS协议的GR状态	**display isis graceful-restart status** [**level-1** | **level-2**] [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示使能IS-IS功能接口的信息	**display isis interface** [**statistics** | [*interface-type interface-number*] [**verbose**]] [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS的链路状态数据库	**display isis lsdb** [[**l1** | **l2** | **level-1** | **level-2**] | [**lsp-id** *lspid* | **lsp-name** *lspname*] | **local** | **verbose**] * [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示系统ID到主机名称的映射关系表	**display isis name-table** [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS的邻居信息	**display isis peer** [**statistics | verbose**] [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS的IPv4路由信息	**display isis route** [**ipv4**] [[**level-1** | **level-2**] | **verbose**] * [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
显示IS-IS进行SPF计算的日志信息	**display isis spf-log** [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS统计信息	**display isis statistics** [**level-1** | **level-1-2** | **level-2**] [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除所有IS-IS的数据结构信息	**reset isis all** [*process-id* | **vpn-instance** *vpn-instance-name*]
清除IS-IS特定邻居的数据信息	**reset isis peer** *system-id* [*process-id* | **vpn-instance** *vpn-instance-name*]

1.14 IS-IS典型配置案例

1.14.1 IS-IS基本配置

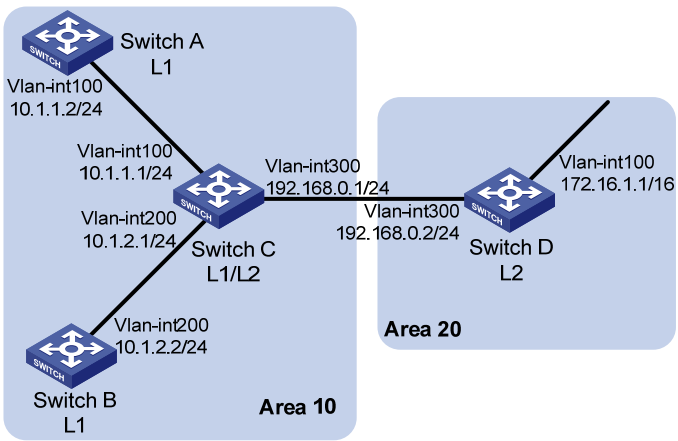
1. 组网需求

如 图 1-15 所示，Switch A、Switch B、Switch C和Switch 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。

其中，Switch A 和 Switch B 为 Level-1 交换机，Switch D 为 Level-2 交换机，Switch C 作为 Level-1-2 交换机将两个区域相连。Switch A、Switch B 和 Switch C 的区域号为 10，Switch D 的区域号为 20。

2. 组网图

图1-15 IS-IS 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 IS-IS

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] is-level level-1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
```

```
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis enable 1
[SwitchA-Vlan-interface100] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] isis 1
[SwitchB-isis-1] is-level level-1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis enable 1
[SwitchB-Vlan-interface200] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] isis enable 1
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis enable 1
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis enable 1
[SwitchC-Vlan-interface300] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] isis 1
[SwitchD-isis-1] is-level level-2
[SwitchD-isis-1] network-entity 20.0000.0000.0004.00
[SwitchD-isis-1] quit
[SwitchD] interface vlan-interface 100
[SwitchD-Vlan-interface100] isis enable 1
[SwitchD-Vlan-interface100] quit
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] isis enable 1
[SwitchD-Vlan-interface300] quit
```

(3) 验证配置结果

显示各交换机的 IS-IS LSDB 信息，查看 LSP 是否完整。

```
[SwitchA] display isis lsdb

                      Database information for ISIS(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
[SwitchB] display isis lsdb

                        Database information for ISIS(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
[SwitchC] display isis lsdb

                        Database information for ISIS(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
[SwitchD] display isis lsdb

                        Database information for ISIS(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
```

显示各交换机的 IS-IS 路由信息。Level-1 交换机的路由表中应该有一条缺省路由，且下一跳为 Level-1-2 交换机，Level-2 交换机的路由表中应该有所有 Level-1 和 Level-2 的路由。

```
[SwitchA] display isis route

                        Route information for ISIS(1)
                        ------------------------------

                     ISIS(1) IPv4 Level-1 Forwarding Table
                     -------------------------------------

 IPV4 Destination     IntCost   ExtCost ExitInterface   NextHop      Flags
-------------------------------------------------------------------------------
 10.1.1.0/24          10        NULL    Vlan100         Direct       D/L/-
 10.1.2.0/24          20        NULL    Vlan100         10.1.1.1     R/-/-
 192.168.0.0/24       20        NULL    Vlan100         10.1.1.1     R/-/-
 0.0.0.0/0            10        NULL    Vlan100         10.1.1.1     R/-/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
[SwitchC] display isis route

                        Route information for ISIS(1)
                        ------------------------------

                     ISIS(1) IPv4 Level-1 Forwarding Table
                     -------------------------------------

 IPV4 Destination     IntCost    ExtCost ExitInterface  NextHop      Flags
-------------------------------------------------------------------------------
 192.168.0.0/24       10         NULL    Vlan300        Direct       D/L/-
```

```
 10.1.1.0/24          10         NULL    Vlan100        Direct       D/L/-
 10.1.2.0/24          10         NULL    Vlan200        Direct       D/L/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set


                     ISIS(1) IPv4 Level-2 Forwarding Table
                     -------------------------------------

 IPV4 Destination     IntCost   ExtCost ExitInterface   NextHop      Flags
------------------------------------------------------------------------------
 192.168.0.0/24       10        NULL    Vlan300         Direct       D/L/-
 10.1.1.0/24          10        NULL    Vlan100         Direct       D/L/-
 10.1.2.0/24          10        NULL    Vlan200         Direct       D/L/-
 172.16.0.0/16        20        NULL    Vlan300         192.168.0.2  R/-/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
[SwitchD] display isis route

                        Route information for ISIS(1)
                        -----------------------------

                     ISIS(1) IPv4 Level-2 Forwarding Table
                     -------------------------------------

 IPV4 Destination     IntCost   ExtCost ExitInterface   NextHop      Flags
------------------------------------------------------------------------------
 192.168.0.0/24       10        NULL    Vlan300         Direct       D/L/-
 10.1.1.0/24          20        NULL    Vlan300         192.168.0.1  R/-/-
 10.1.2.0/24          20        NULL    Vlan300         192.168.0.1  R/-/-
 172.16.0.0/16        10        NULL    Vlan100         Direct       D/L/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

1.14.2 配置IS-IS的DIS选择

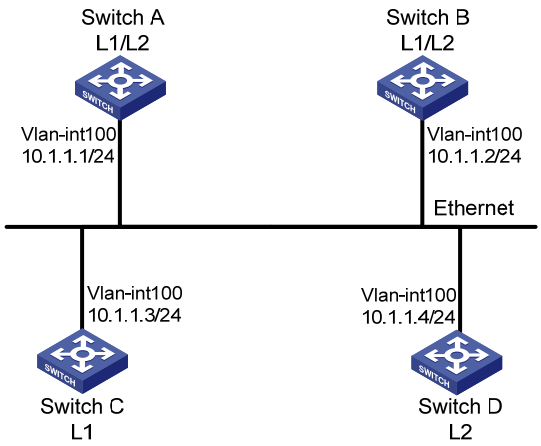
1. 组网需求

如 图 1-16 所示，Switch A、Switch B、Switch C和Switch 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）。其中Switch A和Switch B是Level-1-2 交换机，Switch C为Level-1 交换机，Switch D为Level-2 交换机。

要求通过改变接口的 DIS 优先级，将 Switch A 配置为 Level-1-2 的 DIS。

2. 组网图

图1-16 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 Switch A。

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis enable 1
[SwitchA-Vlan-interface100] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] isis 1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] isis enable 1
[SwitchB-Vlan-interface100] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] is-level level-1
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] isis enable 1
[SwitchC-Vlan-interface100] quit
```

配置 Switch D。

```
<SwitchD> system-view
```

```
[SwitchD] isis 1
[SwitchD-isis-1] network-entity 10.0000.0000.0004.00
[SwitchD-isis-1] is-level level-2
[SwitchD-isis-1] quit
[SwitchD] interface vlan-interface 100
[SwitchD-Vlan-interface100] isis enable 1
[SwitchD-Vlan-interface100] quit
```

查看 Switch A 的 IS-IS 邻居信息。

```
[SwitchA] display isis peer

                        Peer information for ISIS(1)
                        ----------------------------
  System Id: 0000.0000.0002
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0003.01
  State: Up     HoldTime: 21s        Type: L1(L1L2)     PRI: 64

  System Id: 0000.0000.0003
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0003.01
  State: Up     HoldTime: 27s        Type: L1           PRI: 64

  System Id: 0000.0000.0002
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0004.01
  State: Up     HoldTime: 28s        Type: L2(L1L2)     PRI: 64

  System Id: 0000.0000.0004
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0004.01
  State: Up     HoldTime: 30s        Type: L2          PRI: 64
```

显示 Switch A 的 IS-IS 接口信息。

```
[SwitchA] display isis interface

                     Interface information for ISIS(1)
                     ---------------------------------
Interface: Vlan-interface100
Id      IPV4.State      IPV6.State    MTU    Type    DIS
001        Up             Down       1497    L1/L2   No/No
```

显示 Switch C 的 IS-IS 接口信息。

```
[SwitchC] display isis interface

                     Interface information for ISIS(1)
                     ---------------------------------
Interface: Vlan-interface100
Id      IPV4.State      IPV6.State    MTU    Type    DIS
001         Up            Down        1497   L1/L2   Yes/No
```

显示 Switch D 的 IS-IS 接口信息。

```
[SwitchD] display isis interface

                     Interface information for ISIS(1)
```

```
                    --------------------------------
Interface: Vlan-interface100
Id      IPV4.State       IPV6.State   MTU     Type     DIS
001         Up             Down       1497    L1/L2    No/Yes
```

说明

从接口信息中可以看到，在使用缺省 DIS 优先级的情况下，Switch C 为 Level-1 的 DIS，Switch 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。

(3) 配置 Switch A 的 DIS 优先级

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis dis-priority 100
[SwitchA-Vlan-interface100] quit
```

查看 Switch A 的 IS-IS 邻居信息。

```
[SwitchA] display isis peer

                         Peer information for ISIS(1)
                         ----------------------------

  System Id: 0000.0000.0002
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 21s        Type: L1(L1L2)     PRI: 64

  System Id: 0000.0000.0003
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 27s        Type: L1           PRI: 64


  System Id: 0000.0000.0002
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 28s        Type: L2(L1L2)     PRI: 64

  System Id: 0000.0000.0004
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 30s        Type: L2           PRI: 64
```

查看 Switch A 的 IS-IS 接口信息。

```
[SwitchA] display isis interface

                      Interface information for ISIS(1)
                      ---------------------------------
Interface: Vlan-interface100
Id      IPV4.State      IPV6.State      MTU    Type   DIS
001         Up             Down         1497   L1/L2  Yes/Yes
```

说明

从上述信息中可以看到，在改变 IS-IS 接口的 DIS 优先级后，Switch A 立即成为 Level-1-2 的 DR（DIS），且伪节点是 0000.0000.0001.01。

显示 Switch C 的 IS-IS 邻居和接口信息。

```
[SwitchC] display isis peer

                        Peer information for ISIS(1)
                        ----------------------------
  System Id: 0000.0000.0002
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 25s        Type: L1           PRI: 64

  System Id: 0000.0000.0001
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 7s         Type: L1           PRI: 100
[SwitchC] display isis interface

                     Interface information for ISIS(1)
                     ---------------------------------
Interface: Vlan-interface100
Id      IPV4.State          IPV6.State    MTU    Type   DIS
001         Up                 Down       1497   L1/L2  No/No
```

显示 Switch D 的 IS-IS 邻居和接口信息。

```
[SwitchD] display isis peer

                        Peer information for ISIS(1)
                        ----------------------------
  System Id: 0000.0000.0001
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 9s         Type: L2           PRI: 100

  System Id: 0000.0000.0002
  Interface: Vlan-interface100       Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 28s        Type: L2           PRI: 64
[SwitchD] display isis interface

                     Interface information for ISIS(1)
                     ---------------------------------
Interface: Vlan-interface100
Id      IPV4.State          IPV6.State    MTU   Type    DIS
001         Up                 Down       1497  L1/L2   No/No
```

1.14.3 配置IS-IS引入外部路由

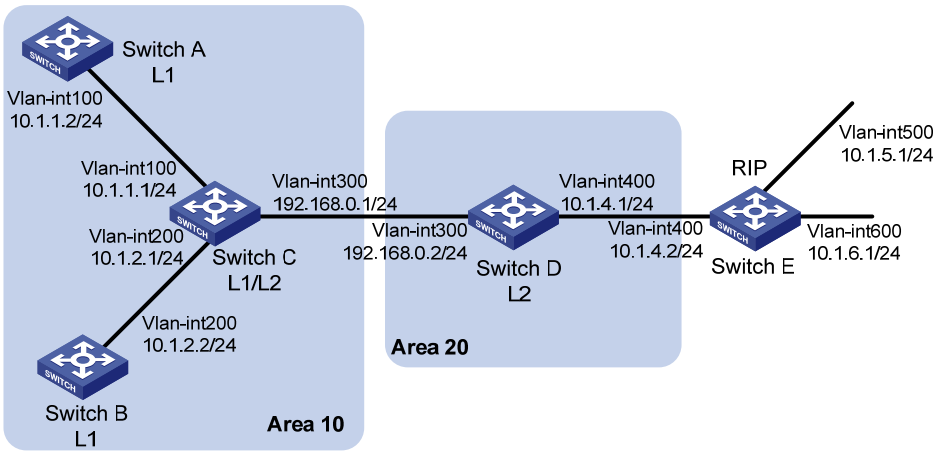
1. 组网需求

如 图 1-17 所示：

- Switch A、Switch B、Switch C 和 Switch 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。
- Switch A 和 Switch B 为 Level-1 路由器，Switch D 为 Level-2 路由器，Switch C 作为 Level-1-2 路由器将两个区域相连。Switch A、Switch B 和 Switch C 的区域号为 10，Switch D 的区域号为 20。
- 在 Switch D 的 IS-IS 进程中引入 RIP 路由。

2. 组网图

图1-17 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 Switch A。

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] is-level level-1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis enable 1
[SwitchA-Vlan-interface100] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] isis 1
[SwitchB-isis-1] is-level level-1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
```

```
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis enable 1
[SwitchB-Vlan-interface200] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis enable 1
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] isis enable 1
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis enable 1
[SwitchC-Vlan-interface300] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] isis 1
[SwitchD-isis-1] is-level level-2
[SwitchD-isis-1] network-entity 20.0000.0000.0004.00
[SwitchD-isis-1] quit
[SwitchD] interface interface vlan-interface 300
[SwitchD-Vlan-interface300] isis enable 1
[SwitchD-Vlan-interface300] quit
```

显示各路由器的 IS-IS 路由信息。

```
[SwitchA] display isis route

                      Route information for ISIS(1)
                      ------------------------------

                    ISIS(1) IPv4 Level-1 Forwarding Table
                    -------------------------------------

 IPV4 Destination     IntCost  ExtCost ExitInterface  NextHop       Flags
--------------------------------------------------------------------------------
 0.0.0.0/0            10         NULL     VLAN100           10.1.1.1      R/-/-
 10.1.1.0/24          10         NULL     VLAN100           Direct        D/L/-
 10.1.2.0/24          20         NULL     VLAN100           10.1.1.1      R/-/-
 192.168.0.0/24       20         NULL     VLAN100           10.1.1.1      R/-/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
[SwitchC] display isis route

                      Route information for ISIS(1)
                      ------------------------------
```

```
                    ISIS(1) IPv4 Level-1 Forwarding Table
                    ------------------------------------

 IPV4 Destination     IntCost   ExtCost ExitInterface   NextHop     Flags
--------------------------------------------------------------------------
 10.1.1.0/24          10           NULL     VLAN100            Direct      D/L/-
 10.1.2.0/24          10           NULL     VLAN200            Direct      D/L/-
 192.168.0.0/24       10           NULL     VLAN300            Direct      D/L/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set


                    ISIS(1) IPv4 Level-2 Forwarding Table
                    ------------------------------------

 IPV4 Destination     IntCost   ExtCost ExitInterface  NextHop      Flags
--------------------------------------------------------------------------
 10.1.1.0/24          10          NULL     VLAN100           Direct       D/L/-
 10.1.2.0/24          10          NULL     VLAN200           Direct       D/L/-
 192.168.0.0/24       10          NULL     VLAN300           Direct       D/L/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
[SwitchD] display isis route

                       Route information for ISIS(1)
                       -----------------------------

                    ISIS(1) IPv4 Level-2 Forwarding Table
                    ------------------------------------

 IPV4 Destination     IntCost   ExtCost ExitInterface  NextHop      Flags
--------------------------------------------------------------------------
 10.1.1.0/24           20          NULL     VLAN300           192.168.0.1  R/-/-
 10.1.2.0/24           20          NULL     VLAN300           192.168.0.1  R/-/-
 192.168.0.0/24        10          NULL     VLAN300           Direct       D/L/-


      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

(3) 在 Switch D 和 Switch E 之间运行 RIPv2，在 Switch D 上配置 IS-IS 进程引入 RIP 路由。

在 Switch D 上配置 RIPv2。

```
[SwitchD] rip 1
[SwitchD-rip-1] network 10.0.0.0
[SwitchD-rip-1] version 2
[SwitchD-rip-1] undo summary
```

在 Switch E 上配置 RIPv2。

```
[SwitchE] rip 1
[SwitchE-rip-1] network 10.0.0.0
[SwitchE-rip-1] version 2
```

```
[SwitchE-rip-1] undo summary
```

在 Switch D 上配置 IS-IS 进程引入 RIP 进程的路由。

```
[SwitchD-rip-1] quit
[SwitchD] isis 1
[SwitchD–isis-1] import-route rip level-2
```

显示 Switch C 的 IS-IS 路由信息。

```
[SwitchC] display isis route

                        Route information for ISIS(1)
                        ------------------------------

                    ISIS(1) IPv4 Level-1 Forwarding Table
                    -------------------------------------

 IPV4 Destination     IntCost   ExtCost ExitInterface   NextHop      Flags
--------------------------------------------------------------------------------
 10.1.1.0/24          10         NULL    VLAN100            Direct      D/L/-
 10.1.2.0/24          10         NULL    VLAN200            Direct      D/L/-
 192.168.0.0/24       10         NULL    VLAN300            Direct      D/L/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set


                    ISIS(1) IPv4 Level-2 Forwarding Table
                    -------------------------------------

 IPV4 Destination     IntCost   ExtCost ExitInterface  NextHop      Flags
--------------------------------------------------------------------------------
 10.1.1.0/24          10          NULL    VLAN100           Direct       D/L/-
 10.1.2.0/24          10         NULL    VLAN200           Direct       D/L/-
 10.1.4.0/24          20         NULL    VLAN300           192.168.0.2  R/L/-
 10.1.5.0/24          10         0       VLAN300           192.168.0.2  R/L/-
 10.1.6.0/24          10         0       VLAN300           192.168.0.2  R/L/-
 192.168.0.0/24       10         NULL    VLAN300           Direct       D/L/-


      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

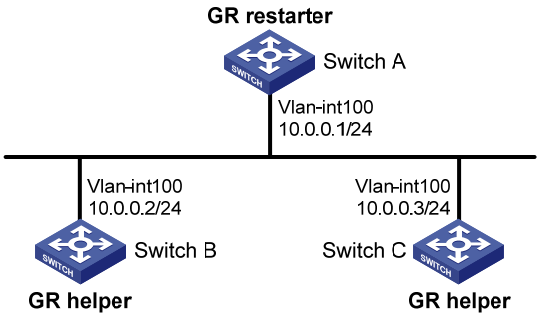
1.14.4 IS-IS GR配置举例

1. 组网需求

如 图 1-18 所示，Switch A、Switch B和Switch C属于同一域。这三台交换机都运行IS-IS协议以实现路由互连。

2. 组网图

图1-18 IS-IS GR 配置组网图



3. 配置步骤

(1) 配置各交换机接口的 IP 地址和 IS-IS 协议

请按照 图 1-18 配置各接口的IP地址和子网掩码，具体配置过程略。

配置各交换机之间采用 IS-IS 协议进行互连，确保 Switch A、Switch B 和 Switch C 之间能够在网络层互通，并且各交换机之间能够借助 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 Switch A 的 IS-IS 协议的 GR 能力，并配置重启间隔时间。

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] graceful-restart
[SwitchA-isis-1] graceful-restart interval 150
[SwitchA-isis-1] return
```

Switch B 和 Switch C 的配置与 Switch A 相似，配置过程略。

(3) 检验配置效果

Switch A 分别与 Switch B 和 Switch C 建立邻接关系后，三台交换机开始交换路由信息。Switch 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 Switch A 上 IS-IS 协议的 GR 状态。

重启 Switch A 的 IS-IS 进程。

```
<SwitchA> reset isis all 1
Warning : Reset ISIS process? [Y/N]:y
```

查看 Switch A 上 IS-IS 协议的 GR 状态。

```
<SwitchA> display isis graceful-restart status
                Restart information for IS-IS(1)
--------------------------------------------------------------------------
IS-IS(1) Level-1 Restart Status
Restart Interval: 150
SA Bit Supported
  Total Number of Interfaces = 1
  Restart Status: RESTARTING
  Number of LSPs Awaited: 3
  T3 Timer Status:
```

```
    Remaining Time: 140
  T2 Timer Status:
    Remaining Time: 59

IS-IS(1) Level-2 Restart Status
Restart Interval: 150
SA Bit Supported
  Total Number of Interfaces = 1
  Restart Status: RESTARTING
  Number of LSPs Awaited: 3
  T3 Timer Status:
    Remaining Time: 140
  T2 Timer Status:
Remaining Time: 59
```

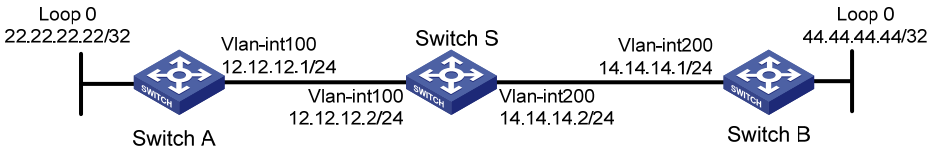
1.14.5 IS-IS NSR配置举例

1. 组网需求

如 图 1-19 所示，Switch S、Switch A、Switch B属于同一IS-IS区域，通过IS-IS协议实现网络互连，其中Switch S为形成了IRF的设备，与Switch A、Switch B连接的链路为跨设备聚合链路。要求对Switch S进行Master/Slave切换时，Switch A和Switch B到Switch S的邻居没有中断，Switch A到Switch B的流量没有中断。

2. 组网图

图1-19 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 1-19 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 Switch S、Switch A 和 Switch 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 Switch A 的 IS-IS NSR 功能。

```
<SwitchS> system-view
[SwitchS] isis 1
[SwitchS-isis-1] non-stop-routing
[SwitchS-isis-1] non-stop-routing interval 30
[SwitchS-isis-1] return
```

(3) 检验配置效果

Switch S 分别与 Switch A 和 Switch 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，Switch S 进行 Master/Slave 切换。在 Switch S Master/Slave 切换期间，使用 **display isis peer** 命令查看 Switch A 和 Switch B 上到 Switch S 的邻居是否发生任何变化；使用 **display isis route** 命令查看 Switch A 上是否有 Switch B 上 Loopback 接口的路由，Switch B 上是否有 Switch A 上 Loopback 接口的路由。

在 Switch S 进行 Master/Slave 切换时，查看 Switch A 上 IS-IS 协议的邻居和路由。

```
<SwitchA> display isis peer

      Peer information for ISIS(1)
                          ----------------------------

  System Id: 0000.0000.0001
  Interface: vlan100                Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 25s       Type: L1(L1L2)     PRI: 64

  System Id: 0000.0000.0001
  Interface: vlan100                Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 27s       Type: L2(L1L2)     PRI: 64
<SwitchA> display isis route
 Route information for ISIS(1)
                           ----------------------------
                     ISIS(1) IPv4 Level-1 Forwarding Table
                     ------------------------------------
IPV4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 12.12.12.0/24       10         NULL    vlan100         Direct          D/L/-
 22.22.22.22/32      10         NULL    Loop0           Direct          D/-/-
 14.14.14.0/32       10         NULL    vlan100         12.12.12.2      R/L/-
 44.44.44.44/32      10         NULL    vlan100         12.12.12.2      R/L/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set


                     ISIS(1) IPv4 Level-2 Forwarding Table
                     ------------------------------------

 IPV4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 12.12.12.0/24        10         NULL    vlan100         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 RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

查看 Switch B 上 IS-IS 协议的邻居和路由。

```
<SwitchB> display isis peer
```

```
      Peer information for ISIS(1)
                        ----------------------------

  System Id: 0000.0000.0001
  Interface: vlan200                Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 25s       Type: L1(L1L2)     PRI: 64

  System Id: 0000.0000.0001
  Interface: vlan200                Circuit Id: 0000.0000.0001.01
  State: Up     HoldTime: 27s       Type: L2(L1L2)     PRI: 64
<SwitchB> display isis route
 Route information for ISIS(1)
                          ----------------------------
                     ISIS(1) IPv4 Level-1 Forwarding Table
                     ------------------------------------
IPV4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 14.14.14.0/24       10         NULL    vlan200         Direct          D/L/-
 44.44.44.44/32      10         NULL    Loop0           Direct          D/-/-
 12.12.12.0/32       10         NULL    vlan200         14.14.14.4      R/L/-
 22.22.22.22/32      10         NULL    vlan200         14.14.14.4      R/L/-

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set


                     ISIS(1) IPv4 Level-2 Forwarding Table
                     ------------------------------------

 IPV4 Destination     IntCost    ExtCost ExitInterface   NextHop         Flags
--------------------------------------------------------------------------------
 14.14.14.0/24        10         NULL    vlan200         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 RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

通过上面信息可以看出 Switch A 和 Switch 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 Switch S 的 Master/Slave 切换。

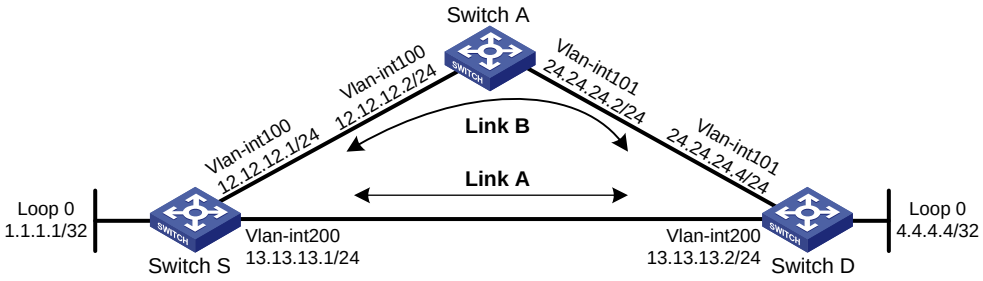
1.14.6 IS-IS快速重路由配置举例

1. 组网需求

如 图 1-20 所示，Switch S、Switch A和Switch D属于同一IS-IS路由域，通过IS-IS协议实现网络互连。要求当Switch S和Switch 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1-20 IS-IS 快速重路由配置组网图



3. 配置步骤

(1) 配置各交换机接口的 IP 地址和 IS-IS 协议

请按照 图 1-18 配置各接口的IP地址和子网掩码，具体配置过程略。

配置各交换机之间采用 IS-IS 协议进行互连，确保 Switch A、Switch D 和 Switch S 之间能够在网络层互通，并且各交换机之间能够借助 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 Switch S 和 Switch D 的 IS-IS 协议的自动计算快速重路由能力

配置 Switch S。

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 1.1.1.1
[SwitchS] isis 1
[SwitchS-isis-1] fast-reroute auto
[SwitchS-isis-1] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] bfd echo-source-ip 4.4.4.4
[SwitchD] isis 1
[SwitchD-isis-1] fast-reroute auto
[SwitchD-isis-1] quit
```

方法二：使能 Switch S 和 Switch D 的 IS-IS 协议的指定路由策略快速重路由能力

配置 Switch S。

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 1.1.1.1
[SwitchS] ip ip-prefix abc index 10 permit 4.4.4.4 32
[SwitchS] route-policy frr permit node 10
[SwitchS-route-policy] if-match ip-prefix abc
[SwitchS-route-policy] apply fast-reroute backup-interface vlan-interface 100
backup-nexthop 12.12.12.2
[SwitchS-route-policy] quit
[SwitchS] isis 1
```

```
[SwitchS-isis-1] fast-reroute route-policy frr
[SwitchS-isis-1] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] bfd echo-source-ip 4.4.4.4
[SwitchD] ip ip-prefix abc index 10 permit 1.1.1.1 32
[SwitchD] route-policy frr permit node 10
[SwitchD-route-policy] if-match ip-prefix abc
[SwitchD-route-policy] apply fast-reroute backup-interface vlan-interface 101
backup-nexthop 24.24.24.2
[SwitchD-route-policy] quit
[SwitchD] isis 1
[SwitchD-isis-1] fast-reroute route-policy frr
[SwitchD-isis-1] quit
```

(3) 检验配置效果

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

```
[SwitchS] display ip routing-table 4.4.4.4 verbose
Routing Table : Public
Summary Count : 1

  Destination: 4.4.4.4/32
     Protocol: ISIS            Process ID: 1
   Preference: 10                    Cost: 10
 IpPrecedence:                   QosLcId:
      NextHop: 13.13.13.2       Interface: Vlan-interface200
    BkNextHop: 12.12.12.2     BkInterface: Vlan-interface100
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                 Label: NULL
  BKTunnel ID: 0x0               BKLabel: NULL
        State: Active Adv            Age: 00h01m27s
          Tag: 0
```

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

```
[SwitchD] display ip routing-table 1.1.1.1 verbose
Routing Table : Public
Summary Count : 1

  Destination: 1.1.1.1/32
     Protocol: ISIS            Process ID: 1
   Preference: 10                    Cost: 10
 IpPrecedence:                   QosLcId:
      NextHop: 13.13.13.1       Interface: Vlan-interface200
    BkNextHop: 24.24.24.2     BkInterface: Vlan-interface101
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                 Label: NULL
  BKTunnel ID: 0x0               BKLabel: NULL
        State: Active Adv            Age: 00h01m27s
          Tag: 0
```

1.14.7 IS-IS验证配置举例

1. 组网需求

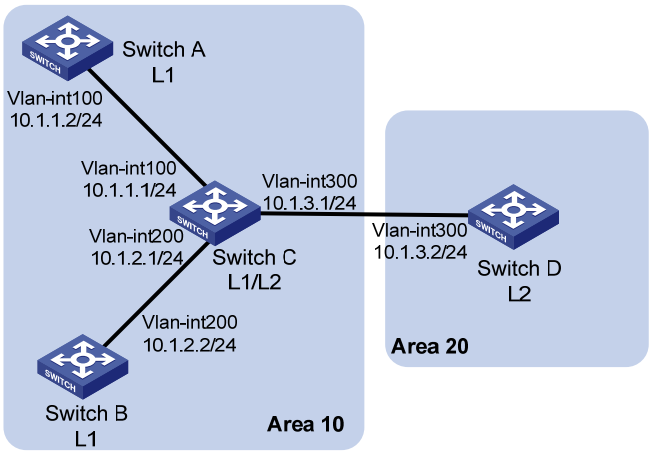
如 图 1-21 所示，Switch A、Switch B、Switch C和Switch D属于同一路由域，要求它们之间通过IS-IS协议达到IP网络互连的目的。

其中，Switch A、Switch B 和 Switch C 属于同一个区域，区域号为 10，Switch 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 Switch C 和 Switch 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 Switch A、Switch B、Switch C 和 Switch D 上配置邻居关系验证。

2. 组网图

图1-21 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 Switch A。

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis enable 1
[SwitchA-Vlan-interface100] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] isis 1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis enable 1
```

```
[RouterB--Vlan-interface200] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis enable 1
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis enable 1
[SwitchC-Vlan-interface300] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis enable 1
[SwitchC-Vlan-interface300] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] isis 1
[SwitchD-isis-1] network-entity 20.0000.0000.0001.00
[SwitchD-isis-1] quit
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] isis enable 1
[SwitchD-Vlan-interface300] quit
```

(3) 在 Switch A、Switch B、Switch C 和 Switch D 之间建立邻居关系验证

分别在 Switch A 的 Vlan-interface100、Switch C 的 Vlan-interface100 配置邻居关系验证，验证方式为 MD5 密文，验证密码为“eRg”。

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis authentication-mode md5 eRg
[SwitchA-Vlan-interface100] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] isis authentication-mode md5 eRg
[SwitchC-Vlan-interface100] quit
```

分别在 Switch B 的 Vlan-interface200、Switch C 的 Vlan-interface200 上配置邻居关系验证，验证方式为 MD5 密文，验证密码为“t5Hr”。

```
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis authentication-mode md5 t5Hr
[SwitchB-Vlan-interface200] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis authentication-mode md5 t5Hr
[SwitchC-Vlan-interface200] quit
```

分别在 Switch C 的 Vlan-interface300、Switch D 的 Vlan-interface300 配置邻居关系验证，验证方式为 MD5 密文，验证密码为“hSec”。

```
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis authentication-mode md5 hSec
[SwitchC-Vlan-interface300] quit
[SwitchD] interface vlan-interface 300
```

```
[SwitchD-Vlan-interface300] isis authentication-mode md5 hSec
[SwitchD-Vlan-interface300] quit
```

(4) 在 Switch A、Switch B 和 Switch C 上配置区域验证，验证方式为 MD5 密文验证，验证密码为“10Sec”。

```
[SwitchA] isis 1
[SwitchA-isis-1] area-authentication-mode md5 10Sec
[SwitchA-isis-1] quit
[SwitchB] isis 1
[SwitchB-isis-1] area-authentication-mode md5 10Sec
[SwitchB-isis-1] quit
[SwitchC] isis 1
[SwitchC-isis-1] area-authentication-mode md5 10Sec
[SwitchC-isis-1] quit
```

(5) 在 Switch C 和 Switch D 上配置路由域验证，验证方式为 MD5 密文验证，验证密码为“1020Sec”。

```
[SwitchC] isis 1
[SwitchC-isis-1] domain-authentication-mode md5 1020Sec
[SwitchC-isis-1] quit
[SwitchD] isis 1
[SwitchD-isis-1] domain-authentication-mode md5 1020Sec
```

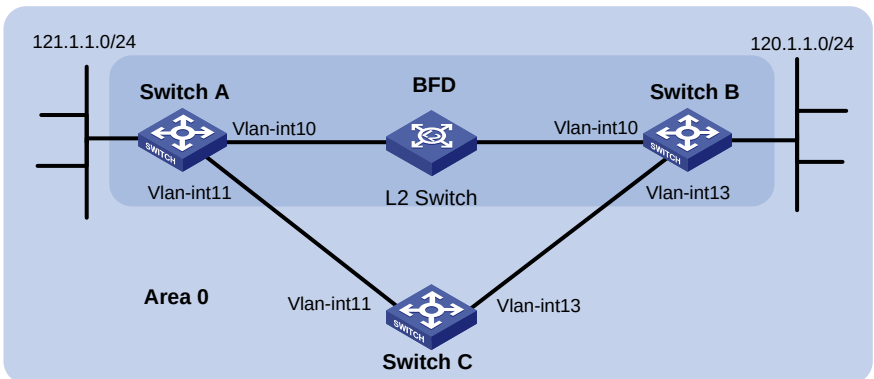
1.14.8 配置IS-IS与BFD联动

1. 组网需求

- Switch A、Switch B 和 Switch C 上运行 IS-IS，网络层相互可达。
- 当 Switch A 和 Switch 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 Switch C 进行通信。

2. 组网图

图1-22 IS-IS 与 BFD 联动配置组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int10	10.1.0.102/24	Switch B	Vlan-int10	10.1.0.100/24
	Vlan-int11	11.1.1.1/24		Vlan-int13	13.1.1.1/24
Switch C	Vlan-int11	11.1.1.2/24			
	Vlan-int13	13.1.1.2/24			

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 Switch A。

```
<SwitchA> system-view
[SwitchA] isis
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] isis enable
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] isis enable
[SwitchA-Vlan-interface11] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] isis
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] isis enable
[SwitchB-Vlan-interface10] quit
[SwitchB] interface vlan-interface 13
[SwitchB-Vlan-interface13] isis enable
[SwitchB-Vlan-interface13] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] isis
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 11
[SwitchC-Vlan-interface11] isis enable
[SwitchC-Vlan-interface11] quit
[SwitchC] interface vlan-interface 13
[SwitchC-Vlan-interface13] isis enable
[SwitchC-Vlan-interface13] quit
```

(3) 配置 BFD 功能

在 Switch A 上使能 BFD 检测功能，并配置 BFD 参数。

```
[SwitchA] bfd session init-mode passive
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] isis bfd enable
[SwitchA-Vlan-interface10] bfd min-receive-interval 500
[SwitchA-Vlan-interface10] bfd min-transmit-interval 500
[SwitchA-Vlan-interface10] bfd detect-multiplier 7
```

在 Switch B 上使能 BFD 检测功能，并配置 BFD 参数。

```
[SwitchB] bfd session init-mode active
```

```
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] isis bfd enable
[SwitchB-Vlan-interface10] bfd min-receive-interval 500
[SwitchB-Vlan-interface10] bfd min-transmit-interval 500
[SwitchB-Vlan-interface10] bfd detect-multiplier 8
[SwitchB-Vlan-interface10] return
```

(4) 检查配置结果

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

显示 Switch A 的 BFD 信息。

```
<SwitchA> display bfd session
 Total Session Num: 1            Init Mode: Active
 Session Working Under Ctrl Mode:
 LD/RD         SourceAddr      DestAddr        State Holdtime Interface
 3/1           10.1.0.102      10.1.0.100      Up    1700ms   vlan10
```

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

```
<SwitchA> display ip routing-table 120.1.1.0 verbose
Routing Table : Public
Summary Count : 1
  Destination: 120.1.1.0/24
     Protocol: ISIS            Process ID: 1
   Preference: 0                     Cost: 2
 IpPrecedence:                    QosLcId:
      NextHop: 192.168.0.100    Interface: Vlan-interface10
    BkNextHop: 0.0.0.0        BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Active Adv             Age: 00h58m10s
          Tag: 0
```

当 Switch B 和二层交换机之间的链路发生故障时 Switch A 能够快速检测 Switch B 的变化。

查看 Switch A 的 BFD 信息，Switch A 已取消与 Switch B 的邻居关系，没有任何输出信息。

```
<SwitchA> display bfd session
```

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

```
<SwitchA> display ip routing-table 120.1.1.0 verbose
Routing Table : Public
Summary Count : 1
  Destination: 120.1.1.0/24
     Protocol: ISIS            Process ID: 1
   Preference: 10                    Cost: 4
 IpPrecedence:                    QosLcId:
      NextHop: 10.1.1.100       Interface: Vlan-interface11
    BkNextHop: 0.0.0.0        BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                  Label: NULL
```

```
BKTunnel ID: 0x0                BKLabel: NULL
      State: Active Adv             Age: 00h58m10s
        Tag: 0
```

目 录

1 BGP配置 ····· 1-1
1.1 BGP简介 ····· 1-1
1.1.1 BGP概述 ····· 1-1
1.1.2 BGP的消息类型 ····· 1-2
1.1.3 BGP的路由属性 ····· 1-4
1.1.4 BGP的选路规则 ····· 1-9
1.1.5 IBGP和IGP同步 ····· 1-11
1.1.6 大规模BGP网络所遇到的问题 ····· 1-11
1.1.7 BGP GR ····· 1-15
1.1.8 MP-BGP ····· 1-16
1.1.9 协议规范 ····· 1-16
1.2 BGP配置任务简介 ····· 1-17
1.3 配置BGP连接 ····· 1-18
1.3.1 配置准备 ····· 1-18
1.3.2 创建BGP连接 ····· 1-18
1.3.3 配置建立TCP连接使用的源接口 ····· 1-20
1.3.4 配置允许同非直连邻居建立EBGP连接 ····· 1-20
1.4 控制路由信息的生成 ····· 1-21
1.4.1 配置准备 ····· 1-21
1.4.2 配置BGP发布本地路由 ····· 1-21
1.4.3 配置BGP引入其它路由 ····· 1-21
1.4.4 配置BGP引入其它协议缺省路由 ····· 1-22
1.5 控制路由信息的发布与接收 ····· 1-22
1.5.1 配置准备 ····· 1-22
1.5.2 配置BGP路由聚合 ····· 1-23
1.5.3 配置向对等体/对等体组发送缺省路由 ····· 1-23
1.5.4 配置BGP路由信息的发布/接收策略 ····· 1-24
1.5.5 配置BGP与IGP路由同步 ····· 1-25
1.5.6 限制从BGP对等体/对等体组接收的路由前缀数量 ····· 1-26
1.5.7 配置BGP路由衰减 ····· 1-26
1.5.8 配置shortcut路由 ····· 1-26
1.6 配置BGP的路由属性 ····· 1-27
1.6.1 配置准备 ····· 1-27

1.6.2 为接收路由分配首选值 ····· 1-27
1.6.3 配置BGP的路由优先级 ····· 1-27
1.6.4 配置Local-Preference缺省值 ····· 1-28
1.6.5 配置MED属性 ····· 1-28
1.6.6 配置下一跳属性 ····· 1-31
1.6.7 配置AS_PATH属性 ····· 1-32
1.7 调整和优化BGP网络 ····· 1-34
1.7.1 配置准备 ····· 1-34
1.7.2 配置BGP存活时间间隔与保持时间 ····· 1-34
1.7.3 配置发送路由更新报文的时间间隔 ····· 1-35
1.7.4 配置BGP软复位 ····· 1-35
1.7.5 配置BGP ORF能力 ····· 1-37
1.7.6 使能 4 字节AS号抑制功能 ····· 1-38
1.7.7 配置BGP报文的DSCP优先级 ····· 1-38
1.7.8 使能EBGP连接快速复位功能 ····· 1-39
1.7.9 配置BGP建立TCP连接时进行MD5 认证 ····· 1-39
1.7.10 配置BGP负载分担 ····· 1-39
1.7.11 禁止与对等体/对等体组建立会话 ····· 1-40
1.8 配置大型BGP网络 ····· 1-40
1.8.1 配置准备 ····· 1-40
1.8.2 配置BGP对等体组 ····· 1-40
1.8.3 配置BGP团体 ····· 1-42
1.8.4 配置BGP路由反射器 ····· 1-43
1.8.5 配置BGP联盟 ····· 1-44
1.9 配置BGP GR ····· 1-45
1.10 开启Trap功能 ····· 1-46
1.11 使能BGP日志功能 ····· 1-46
1.12 配置BGP与BFD联动 ····· 1-47
1.13 BGP显示和维护 ····· 1-47
1.13.1 显示BGP ····· 1-47
1.13.2 复位BGP连接 ····· 1-48
1.13.3 清除BGP信息 ····· 1-49
1.14 BGP典型配置举例 ····· 1-49
1.14.1 BGP基本配置 ····· 1-49
1.14.2 BGP与IGP交互配置 ····· 1-54
1.14.3 BGP负载分担配置 ····· 1-56

1.14.4 BGP团体配置……1-59
1.14.5 BGP路由反射器配置……1-61
1.14.6 BGP联盟配置……1-63
1.14.7 BGP路径选择配置 ……1-67
1.14.8 BGP GR配置 ……1-70
1.14.9 BGP与BFD联动配置……1-71
1.15 BGP常见错误配置举例……1-76
1.15.1 BGP对等体关系不能建立 ……1-76

1 BGP配置

说明

在以下路由协议的介绍中所指的路由器及路由器图标，代表了一般意义下的路由器以及运行了路由协议的以太网交换机。

1.1 BGP简介

1.1.1 BGP概述

BGP（Border Gateway Protocol，边界网关协议）是一种用于 AS（Autonomous System，自治系统）之间的动态路由协议。AS 是拥有同一选路策略，在同一技术管理部门下运行的一组路由器。早期发布的三个版本分别是 BGP-1（RFC 1105）、BGP-2（RFC 1163）和 BGP-3（RFC 1267），当前使用的版本是 BGP-4（RFC 1771，已更新至 RFC 4271）。BGP-4 作为事实上的 Internet 外部路由协议标准，被广泛应用于 ISP（Internet Service Provider，互联网服务提供商）之间。

说明

下文中若不做特殊说明，所指的 BGP 均为 BGP-4。

BGP 特性描述如下：

- BGP 是一种外部网关协议（Exterior Gateway Protocol，EGP），与 OSPF、RIP 等内部网关协议（Interior Gateway Protocol，IGP）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 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。

发送 BGP 消息的路由器称为 BGP 发言者（BGP Speaker），它接收或产生新的路由信息，并发布（Advertise）给其它 BGP 发言者。当 BGP 发言者收到来自其它自治系统的新路由时，如果该路由比当前已知路由更优、或者当前还没有该路由，它就把这条路由发布给自治系统内所有其它 BGP 发言者。

相互交换消息的 BGP 发言者之间互称对等体（Peer），若干相关的对等体可以构成对等体组（Peer group）。

BGP 在路由器上以下列两种方式运行：

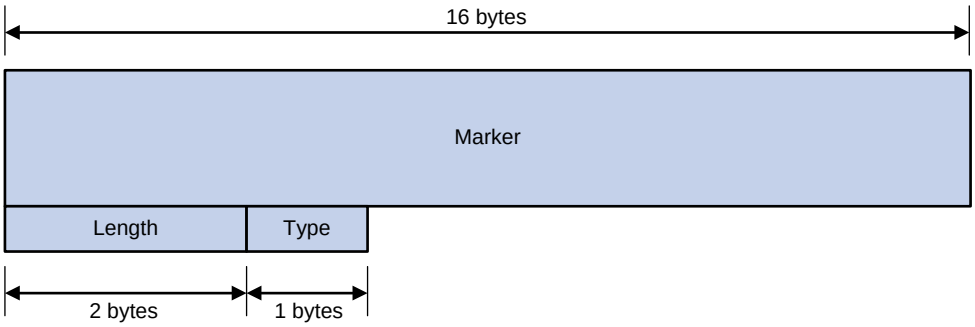
- IBGP（Internal BGP）：当 BGP 运行于同一自治系统内部时，被称为 IBGP；
- EBGP（External BGP）：当 BGP 运行于不同自治系统之间时，称为 EBGP。

1.1.2 BGP的消息类型

1. 消息头格式

BGP有 5 种消息类型：Open、Update、Notification、Keepalive和Route-refresh。这些消息有相同的报文头，其格式如 图 1-1 所示。

图1-1 BGP 消息的报文头格式



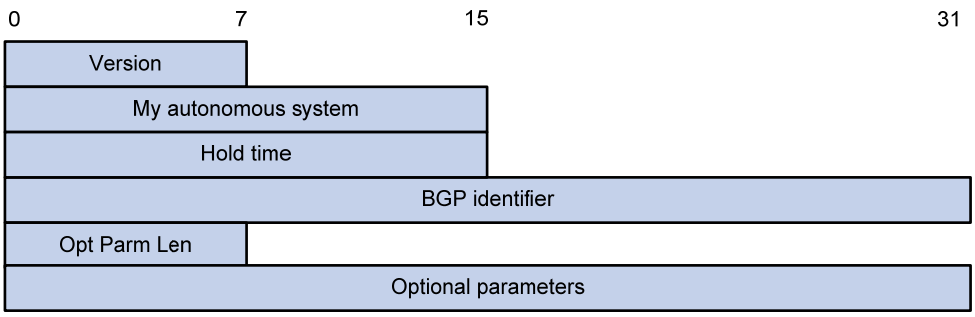
主要字段的解释如下：

- Marker：16 字节，用于标明 BGP 报文边界，所有比特均为“1”。
- Length：2 字节，BGP 消息总长度（包括报文头在内），以字节为单位。
- Type：1 字节，BGP 消息的类型。其取值从 1 到 5，分别表示 Open、Update、Notification、Keepalive 和 Route-refresh 消息。其中，前四种消息是在 RFC 1771 中定义，而 Type 为 5 的消息则是在 RFC 2918 中定义的。

2. Open

Open消息是TCP连接建立后发送的第一个消息，用于建立BGP对等体之间的连接关系。其消息格式如 图 1-2 所示。

图1-2 BGP Open 消息格式



主要字段的解释如下：

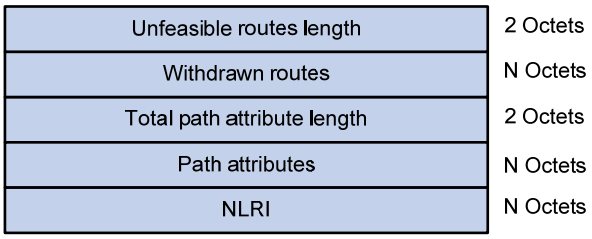
- Version：BGP 的版本号。对于 BGP-4 来说，其值为 4。

- My autonomous system：本地 AS 号。通过比较两端的 AS 号可以确定是 EBGP 连接还是 IBGP 连接。
- Hold time：保持时间。在建立对等体关系时两端要协商 Hold Time，并保持一致。如果在这个时间内未收到对端发来的 Keepalive 消息或 Update 消息，则认为 BGP 连接中断。
- BGP identifier：BGP 标识符。以 IP 地址的形式表示，用来识别 BGP 路由器。
- Opt Parm Len（Optional Parameters Length）：可选参数的长度。如果为 0 则没有可选参数。
- Optional parameters：可选参数。用于多协议扩展（Multiprotocol Extensions）等功能。

3. Update

Update消息用于在对等体之间交换路由信息。它既可以发布可达路由信息，也可以撤销不可达路由信息。其消息格式如 图 1-3 所示。

图1-3 BGP Update 消息格式



一条 Update 报文可以通告一类具有相同路径属性的可达路由，这些路由放在 NLRI（Network Layer Reachability Information，网络层可达性信息）字段中，Path Attributes 字段携带了这些路由的属性，BGP 根据这些属性进行路由的选择；同时 Update 报文还可以携带多条不可达路由，被撤销的路由放在 Withdrawn Routes 字段中。

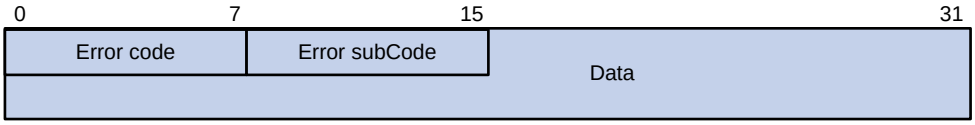
主要字段的解释如下：

- Unfeasible routes length：不可达路由字段的长度，以字节为单位。如果为 0 则说明没有 Withdrawn Routes 字段。
- Withdrawn routes：不可达路由的列表。
- Total path attribute length：路径属性字段的长度，以字节为单位。如果为 0 则说明没有 Path Attributes 字段。
- Path attributes：与 NLRI 相关的所有路径属性列表，每个路径属性由一个 TLV（Type-Length-Value）三元组构成。BGP 正是根据这些属性值来避免环路，进行选路，协议扩展等。
- NLRI：可达路由的前缀和前缀长度二元组。

4. Notification

当BGP检测到错误状态时，就向对等体发出Notification消息，之后BGP连接会立即中断。其消息格式如 图 1-4 所示。

图1-4 BGP Notification 消息格式



主要字段的解释如下：

- Error code：差错码，指定错误类型。
- Error subcode：差错子码，错误类型的详细信息。
- Data：用于辅助发现错误的原因，它的内容依赖于具体的差错码和差错子码，记录的是出错部分的数据，长度不固定。

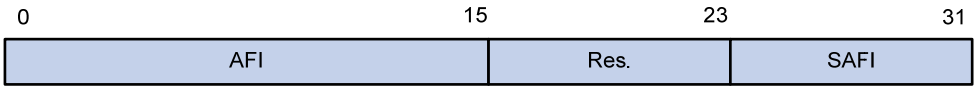
5. Keepalive

BGP 会周期性地向对等体发出 Keepalive 消息，用来保持连接的有效性。其消息格式中只包含报文头，没有附加其他任何字段。

6. Route-refresh

Route-refresh消息用来要求对等体重新发送指定地址族的路由信息。其消息格式如 图 1-5 所示。

图1-5 BGP Route-refresh 消息格式



主要的字段解释如下：

- AFI：Address Family Identifier，地址族标识。
- Res.：保留，必须置 0。
- SAFI：Subsequent Address Family Identifier，子地址族标识。

1.1.3 BGP的路由属性

1. 路由属性的分类

BGP 路由属性是跟随路由一起发送出去的一组参数，封装在 Update 报文的 Path attributes 字段中。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。事实上，所有的 BGP 路由属性都可以分为以下四类：

- 公认必须遵循（Well-known mandatory）：所有 BGP 路由器都必须能够识别这种属性，且必须存在于 Update 消息中。如果缺少这种属性，路由信息就会出错。
- 公认可选（Well-known discretionary）：所有 BGP 路由器都可以识别，但不要求必须存在于 Update 消息中，可以根据具体情况来选择。
- 可选过渡（Optional transitive）：在 AS 之间具有可传递性的属性。BGP 路由器可以不支持此属性，但它仍然会接收带有此属性的路由，并通告给其他对等体。

- 可选非过渡（Optional non-transitive）：如果 BGP 路由器不支持此属性，该属性被忽略，且不会通告给其他对等体。

BGP路由几种基本属性和对应的类别如 表 1-1 所示。

表1-1 路由属性和类别

属性名称	类别
ORIGIN	公认必须遵循
AS_PATH	公认必须遵循
NEXT_HOP	公认必须遵循
LOCAL_PREF	公认可选
ATOMIC_AGGREGATE	公认可选
AGGREGATOR	可选过渡
COMMUNITY	可选过渡
MULTI_EXIT_DISC (MED)	可选非过渡
ORIGINATOR_ID	可选非过渡
CLUSTER_LIST	可选非过渡

2. 几种主要的路由属性

(1) 源（ORIGIN）属性

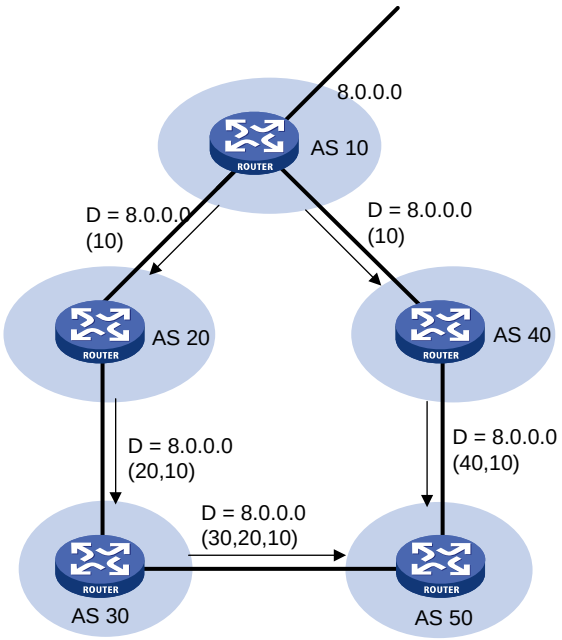
ORIGIN 属性定义路由信息的来源，标记一条路由是怎么成为 BGP 路由的。它有以下三种类型：

- IGP：优先级最高，说明路由产生于本 AS 内。
- EGP：优先级次之，说明路由通过 EGP 学到。
- incomplete：优先级最低，它并不是说明路由不可达，而是表示路由的来源无法确定。例如，引入的其它路由协议的路由信息。

(2) AS 路径（AS_PATH）属性

AS_PATH属性按一定次序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将一条路由通告到其他AS时，便会把本地AS号添加在AS_PATH列表的最前面。收到此路由的BGP路由器根据AS_PATH属性就可以知道去目的地址所要经过的AS。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。如 图 1-6 所示。

图1-6 AS_PATH 属性



通常情况下，BGP 不会接受 AS_PATH 中已包含本地 AS 号的路由，从而避免了形成路由环路的可能。

说明

在目前的实现中，根据某些特殊的具体应用需求，可以通过配置 **peer allow-as-loop** 命令允许 AS 号重复。

同时，AS_PATH 属性也可用于路由的选择和过滤。在其他因素相同的情况下，BGP 会优先选择路径较短的路由。比如在上图中，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。

在某些应用中，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。

通过配置 AS 路径过滤列表，还可以针对 AS_PATH 属性中所包含的 AS 号来对路由进行过滤。

(3) 下一跳（NEXT_HOP）属性

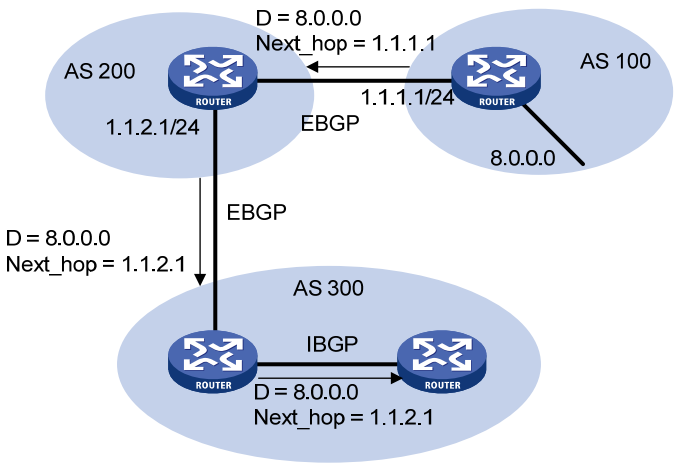
BGP 的下一跳属性和 IGP 的有所不同，不一定就是邻居路由器的 IP 地址。

下一跳属性取值情况分为三种，如 图 1-7 所示。

- BGP 发言者把自己产生的路由发给所有邻居时，将把该路由信息的下一跳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，将把该路由信息的下一跳属性设置为本地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下一跳属性。如果配置了负载分担，等价路由被发给IBGP邻居时则会修改下一跳属性。关于“负载分担”的概念请参见“1.1.4 BGP的选路规则”。

图1-7 下一跳属性

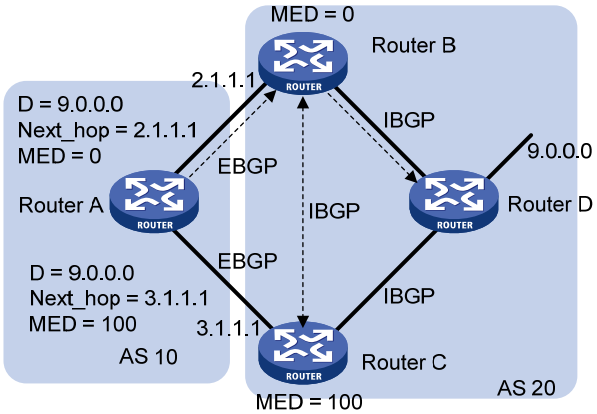


(4) MED（MULTI_EXIT_DISC）

MED 属性仅在相邻两个 AS 之间交换，收到此属性的 AS 一方不会再将其通告给任何其他第三方 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。当一个运行BGP的路由器通过不同的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 1-8 所示，从AS 10 到AS 20 的流量将选择Router B作为入口。

图1-8 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。

说明

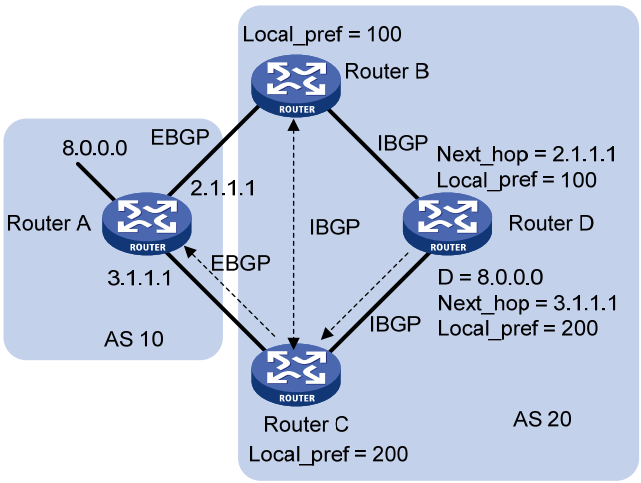
在目前的实现中，通过配置 **compare-different-as-med** 命令，可以强制 BGP 比较来自不同 AS 的路由的 MED 属性值。

(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的路由器通过不同的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1-9所示，从AS 20 到AS 10 的流量将选择Router C作为出口。

图1-9 LOCAL_PREF 属性



(6) 团体（COMMUNITY）属性

团体属性也是跟随路由一起发送出去的一组特殊数据。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就可以根据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），从而能够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“1.1.6 大规模BGP网络所遇到的问题”）。
- 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。

1.1.4 BGP的选路规则

1. BGP选择路由的策略

在目前的实现中，BGP 选择路由时采取如下策略：

(1) 首先丢弃下一跳（NEXT_HOP）不可达的路由；

(2) 优选 Preferred-value 值最大的路由；

(3) 优选本地优先级（LOCAL_PREF）最高的路由；

(4) 优选聚合路由；

(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、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，反射器将自己的 CLUSTER_ID 加入 CLUSTER_LIST 中，若反射器收到路由中 CLUSTER_LIST 中包含有自己的 CLUSTER_ID，则丢弃该路由，从而避免群内环路的发生。
- 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。

---

2. 应用BGP负载分担时的选路

在 BGP 中，由于协议本身的特殊性，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（recursion）。

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 也会生成相同数量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。

在实现方法上，BGP 的负载分担与 IGP 的负载分担有所不同：

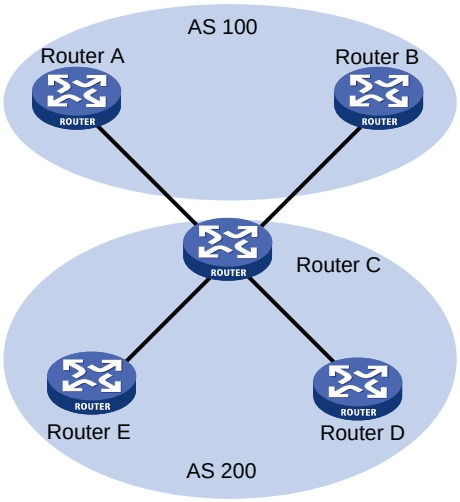
- IGP 是通过协议定义的路由算法，对到达同一目的地址的不同路由，根据计算结果，将度量值（metric）相等的（如 RIP、OSPF）路由进行负载分担，选择的标准很明确（按 metric）。

- 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。

说明

- BGP 只对 AS_PATH 属性、ORIGIN 属性、LOCAL_PREF 和 MED 值完全相同的路由进行负载分担。
- BGP 负载分担特性适用于 EBGP、IBGP 以及联盟之间。
- 如果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。

图1-10 BGP 负载分担示意图



在 图 1-10 中，Router D和Router E是Router C的IBGP对等体。当Router A和Router B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配置了负载分担（如balance 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和MED值时，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 D和Router E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。其它的BGP过渡属性将按最佳路由的属性传递。

3. BGP发布路由的策略

在目前的实现中，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；
- BGP 发言者只把自己使用的路由发布给对等体；
- BGP 发言者从 EBGP 获得的路由会向它所有 BGP 对等体发布（包括 EBGP 对等体和 IBGP 对等体）；
- BGP 发言者从 IBGP 获得的路由不向它的 IBGP 对等体发布；

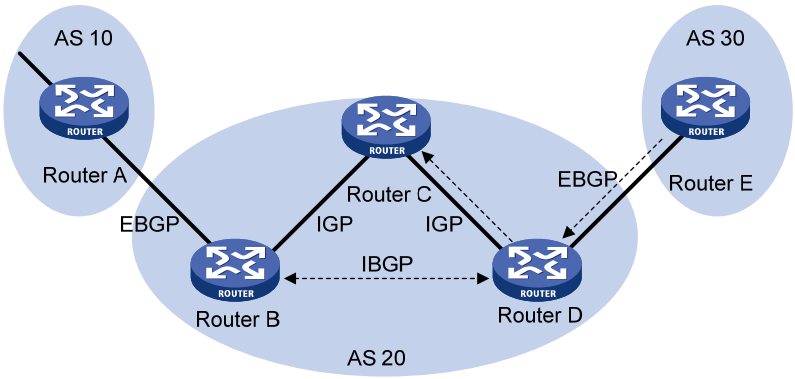
- BGP 发言者从 IBGP 获得的路由发布给它的 EBGP 对等体(关闭 BGP 与 IGP 同步的情况下，IBGP 路由被直接发布；开启 BGP 与 IGP 同步的情况下，该 IBGP 路由只有在 IGP 也发布了这条路由时才会被同步并发布给 EBGP 对等体)；
- 连接一旦建立，BGP 发言者将把自己所有的 BGP 路由发布给新对等体。

1.1.5 IBGP和IGP同步

同步是指 IBGP 和 IGP 之间的同步，其目的是为了避免出现误导外部 AS 路由器的现象发生。

如果一个AS中有非BGP路由器提供转发服务，经该AS转发的IP报文将可能因为目的地址不可达而被丢弃。如 图 1-11 所示，Router E通过BGP从Router D可以学到Router A的一条路由 8.0.0.0/8，于是将到这个目的地址的报文转发给Router D，Router D查询路由表，发现下一跳是Router B（通过**peer next-hop-local**命令手动设置）。由于Router D从IGP学到了到Router B的路由，所以通过路由迭代，Router D将报文转发给Router C。但Router C并不知道去 8.0.0.0/8 的路由，于是将报文丢弃。

图1-11 IBGP 和 IGP 同步



如果设置了同步特性，在 IBGP 路由加入路由表并发布给 EBGP 对等体之前，会先检查 IGP 路由表。只有在 IGP 也知道这条 IBGP 路由时，它才会被发布给 EBGP 对等体。

在下面的情况中，可以关闭同步特性。

- 本 AS 不是过渡 AS（上图中的 AS 20 就属于一个过渡 AS）
- 本 AS 内所有路由器建立 IBGP 全连接

1.1.6 大规模BGP网络所遇到的问题

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. BGP路由衰减

BGP 路由衰减（Route Dampening）用来解决路由不稳定的问题。路由不稳定的主要表现形式是路由震荡（Route flaps），即路由表中的某条路由反复消失和重现。

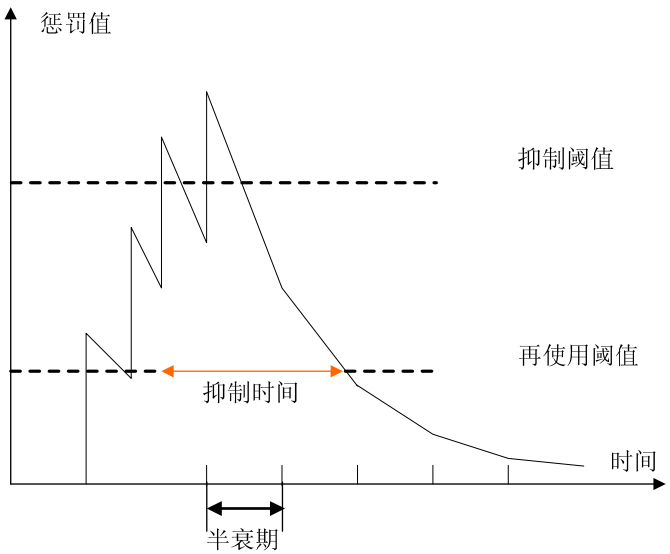
发生路由震荡时，路由协议就会向邻居发布路由更新，收到更新报文的路由器需要重新计算路由并修改路由表。所以频繁的路由震荡会消耗大量的带宽资源和 CPU 资源，严重时会影响到网络的正常工作。

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震荡带来的不利影响，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，惩罚值越高则说明路由越不稳定。路由每发生一次震荡（路由从激活状态变为未激活状态，称为一次路由震荡），BGP 便会给此路由增加一定的惩罚值（1000，此数值为系统固定，不可修改）。当惩罚值超过抑制阈值时，此路由被抑制，不加入到路由表中，也不再向其他 BGP 对等体发布更新报文。

被抑制的路由每经过一段时间，惩罚值便会减少一半，这个时间称为半衰期（Half-life）。当惩罚值降到再使用阈值时，此路由变为可用并被加入到路由表中，同时向其他 BGP 对等体发布更新报文。

图1-12 BGP 衰减示意图



3. 对等体组

对等体组（Peer Group）是一些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中时，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。

在大型 BGP 网络中，对等体的数量会很多，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，利用对等体组在很多情况下可以简化配置。

将对等体加入对等体组中，对等体与对等体组具有相同的路由更新策略，提高了路由发布效率。

注意

如果对等体和对等体组都对某个选项做了配置，配置以最后一次的修改为准。

4. 团体

对等体组可以使一组对等体共享相同的策略，而利用团体可以使多个 AS 中的一组 BGP 路由器共享相同的策略。团体是一个路由属性，在 BGP 对等体之间传播，它并不受到 AS 范围的限制。

BGP 路由器在将带有团体属性的路由发布给其它对等体之前，可以改变此路由原有的团体属性。

除了使用公认的团体属性外，用户还可以使用团体属性列表自定义扩展团体属性，以便更为灵活地控制路由策略。

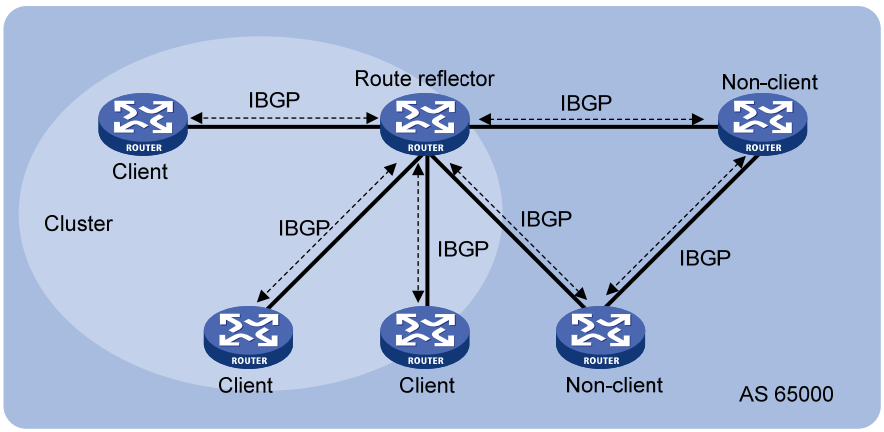
5. 路由反射器

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路由反射器 RR（Route Reflector），其它路由器作为客户机（Client）与路由反射器之间建立 IBGP 连接。路由反射器在客户机之间传递（反射）路由信息，而客户机之间不需要建立 BGP 连接。

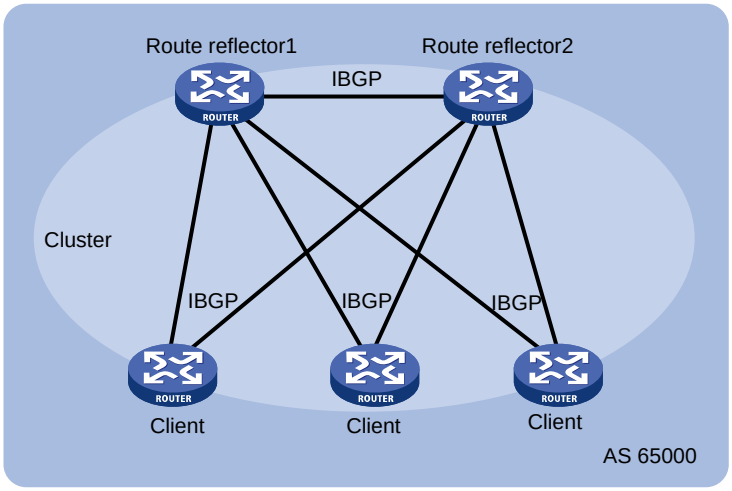
既不是反射器也不是客户机的BGP路由器被称为非客户机（Non-Client）。非客户机与路由反射器之间，以及所有的非客户机之间仍然必须建立全连接关系。其示意图如 图 1-13 所示。

图1-13 路由反射器示意图



路由反射器和它的客户机组成了一个集群（Cluster）。某些情况下，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，位于相同集群中的每个路由反射器都要配置相同的Cluster_ID，以避免路由循环。如 图 1-14 所示。

图1-14 多路由反射器



在某些网络中，路由反射器的客户机之间已经建立了全连接，它们可以直接交换路由信息，此时客户机到客户机之间的路由反射是没有必要的，而且还占用带宽资源。目前，系统支持配置相关命令来禁止在客户机之间反射路由。

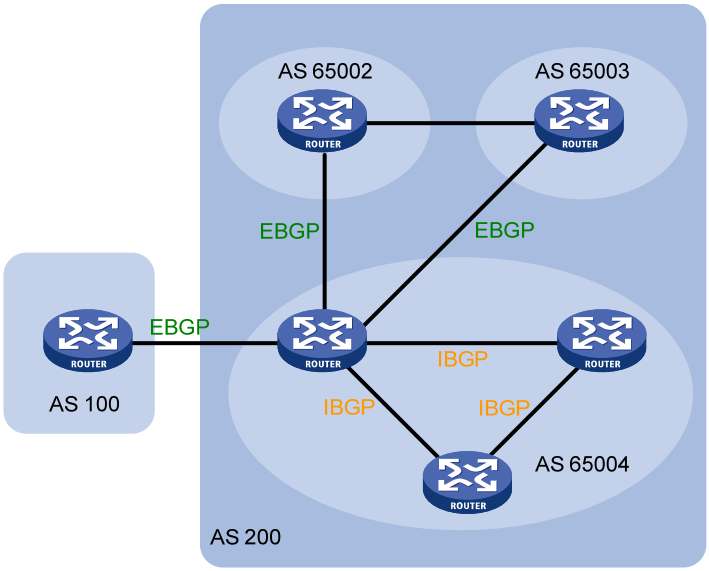
说明

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联盟内部EBGP连接关系。其示意图如 图 1-15 所示。

图1-15 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。

1.1.7 BGP GR

GR 是一种在协议重启或 Master/Slave 切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。

基于 BGP 的 GR Restarter 为了与 BGP 对等体建立一个 BGP 会话连接，首先要发送一个包含了 GR 能力的 OPEN 消息到对端，BGP 对等体收到该消息后，得知发送方已具有 GR 能力。这样，通过 OPEN 消息交互 GR 能力，GR Restarter 与其 BGP 对等体之间协商建立起 GR Session 连接。如果双方都没有交换 GR 能力的信息，建立起的会话也就不具备 GR 能力。

对于 IRF 设备，当进行 Master/Slave 切换时，会话项将丢失，此时具备 GR 感知能力的 BGP 对等体会将所有与该 GR Restarter 有关的路由进行失效标记。但在 GR Time 内仍按照这些路由进行报文转发，这样确保了在从 BGP 对等体重新收集路由信息的过程中没有报文丢失。

对于 IRF 设备，当进行 Master/Slave 切换时，GR Restarter 会重新与 BGP 对等体建立 GR Session 连接，同时发送新的 GR 消息以宣告其重启完毕。此时两个 BGP 对等体间进行路由信息交换。交换完成后，GR Restarter 根据新的路由转发信息更新路由表和转发表，删除失效的路由，完成 BGP 协议收敛。

1.1.8 MP-BGP

1. MP-BGP概述

传统的 BGP-4 只能管理 IPv4 单播的路由信息，对于使用其它网络层协议（如 IPv6 等）的应用，在跨自治系统传播时就受到一定限制。

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，目前的 MP-BGP 标准是 RFC 4760（Multiprotocol Extensions for BGP-4，BGP-4 的多协议扩展）。

支持 BGP 扩展的路由器与不支持 BGP 扩展的路由器可以互通。

2. MP-BGP的扩展属性

BGP-4 使用的报文中，与 IPv4 地址格式相关的三条信息都由 Update 报文携带，这三条信息分别是：NLRI、路径属性中的 NEXT_HOP、路径属性中的 AGGREGATOR（该属性中包含形成聚合路由的 BGP 发言者的 IP 地址）。

为实现对多种网络层协议的支持，BGP-4 需要将网络层协议的信息反映到 NLRI 及 NEXT_HOP。MP-BGP 中引入了两个新的路径属性：

- MP_REACH_NLRI：Multiprotocol Reachable NLRI，多协议可达 NLRI。用于发布可达路由及下一跳信息。
- MP_UNREACH_NLRI：Multiprotocol Unreachable NLRI，多协议不可达 NLRI。用于撤销不可达路由。

这两种属性都是可选非过渡（Optional non-transitive）的，因此，不提供多协议能力的 BGP 发言者将忽略这两个属性的信息，不把它们传递给其它邻居。

3. 地址族

MP-BGP 采用地址族（Address Family）来区分不同的网络层协议，关于地址族的一些取值可以参考 RFC 1700（Assigned Numbers）。目前，系统实现了多种 MP-BGP 扩展应用，包括对 VPN 的扩展、对 IPv6 的扩展等，不同的扩展应用在各自的地址族视图下配置。

说明

- 有关 VPN 的扩展应用，请参见“MPLS 配置指导”中的“MPLS L3VPN”。
- 有关 IPv6 的扩展应用，请参见“三层技术-IP 路由配置指导”中的“IPv6 BGP”。
- 本章不对 MP-BGP 地址族视图下的、与特定应用相关的命令作详细介绍。

1.1.9 协议规范

与 BGP 相关的协议规范有：

- RFC 1771：A Border Gateway Protocol 4 (BGP-4)
- RFC 2858：Multiprotocol Extensions for BGP-4
- RFC 3392：Capabilities Advertisement with BGP-4
- RFC 2918：Route Refresh Capability for BGP-4
- RFC 2439：BGP Route Flap Damping

- RFC 1997：BGP Communities Attribute
- RFC 2796：BGP Route Reflection
- RFC 3065：Autonomous System Confederations for BGP
- RFC 4271，A Border Gateway Protocol 4 (BGP-4)
- RFC 5291，Outbound Route Filtering Capability for BGP-4
- RFC 5292，Address-Prefix-Based Outbound Route Filter for BGP-4
- draft-ietf-idr-restart-08：Graceful Restart Mechanism for BGP

1.2 BGP配置任务简介

表1-2 BGP 配置任务简介

配置任务		说明	详细配置
配置BGP连接	创建BGP连接	必选	1.3.2
	配置建立TCP连接使用的源接口	可选	1.3.3
	配置允许同非直连邻居建立EBGP连接	可选	1.3.4
控制路由信息的生成	配置BGP发布本地路由	二者至少选其一	1.4.2
	配置BGP引入其它路由		1.4.3
	配置BGP引入其它协议缺省路由	可选	1.4.4
控制路由信息的发布与接收	配置BGP路由聚合	可选	1.5.2
	配置向对等体/对等体组发送缺省路由		1.5.3
	配置BGP路由信息的发布/接收策略		1.5.4
	配置BGP与IGP路由同步		1.5.5
	限制从BGP对等体/对等体组接收的路由前缀数量		1.5.6
	配置BGP路由衰减		1.5.7
	配置shortcut路由		1.5.8
配置BGP的路由属性	为接收路由分配首选值	可选	1.6.2
	配置BGP的路由优先级		1.6.3
	配置Local-Preference缺省值		1.6.4
	配置MED属性		1.6.5
	配置下一跳属性		1.6.6
	配置AS_PATH属性		1.6.7
调整和优化BGP网络	配置BGP存活时间间隔与保持时间	可选	1.7.2
	配置发送路由更新报文的时间间隔		1.7.3
	配置BGP软复位		1.7.4
	配置BGP ORF能力		1.7.5

配置任务		说明	详细配置
	使能4字节AS号抑制功能		1.7.6
	配置BGP报文的DSCP优先级		1.7.7
	使能EBGP连接快速复位功能		1.7.8
	配置BGP建立TCP连接时进行MD5认证		1.7.9
	配置BGP负载分担		1.7.10
	禁止与对等体/对等体组建立会话		1.7.11
配置大型BGP网络	配置BGP对等体组	可选	1.8.2
	配置BGP团体		1.8.3
	配置BGP路由反射器		1.8.4
	配置BGP联盟		1.8.5
配置BGP GR		可选	1.9
开启Trap功能		可选	1.10
使能BGP日志功能		可选	1.11
配置BGP与BFD联动		可选	1.12

1.3 配置BGP连接

本节讲述最基本的 BGP 网络配置过程。

说明

在本节中，不对 BGP 和 MP-BGP 进行严格的区分，命令的适用情况请参考所在的视图。

1.3.1 配置准备

在配置 BGP 连接之前，需保证相邻节点在网络层互通。

1.3.2 创建BGP连接

要在路由器上创建 BGP 连接，必须配置该路由器所在的 AS 编号即启动 BGP、配置对等体（对等体组）的 AS 编号。

Router ID 用来在一个自治系统中唯一的标识一台路由器，一台路由器如果要运行 BGP 协议，则必须存在 Router ID。

- 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，则缺省使用全局 Router ID。

- 如果 BGP 使用的是全局 Router ID，当 Router ID 所在接口被删除时会重新选择路由器的 Router ID。
- 如果是在 BGP 视图下手工配置的 Router ID，只有使用 **undo router-id** 命令才能重新选择路由器的 Router ID。

表1-3 创建 BGP 连接

操作	命令	说明
进入系统视图	**system-view**	-
配置全局Router ID	**router id** *router-id*	可选 缺省情况下，未配置全局Router ID 如果没有配置全局Router ID，则按照下面的规则进行选择： (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 (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BGP，进入BGP视图	**bgp** *as-number*	必选 缺省情况下，系统没有运行BGP
配置Router ID	**router-id** *router-id*	可选 缺省情况下，与全局Router ID相同
指定对等体/对等体组及其AS号	**peer** { *group-name* | *ip-address* } **as-number** *as-number*	必选 缺省情况下，没有指定对等体/对等体组AS号
配置使用**peer as-number**命令创建邻居时默认使能IPv4单播地址族	**default ipv4-unicast**	可选 缺省情况下，IPv4单播地址族处于使能状态
激活指定对等体	**peer** *ip-address* **enable**	可选 缺省情况下，BGP对等体是激活的
配置对等体/对等体组的描述信息	**peer** { *group-name* | *ip-address* } **description** *description-text*	可选 缺省情况下，对等体/对等体组没有描述信息

注意

- 一个路由器只能位于一个 AS 内，因此一台路由器只能运行一个 BGP 进程。
- 必须首先创建对等体组，才能进行对等体组的相关配置。关于对等体组的创建可以参见“1.8.2 配置BGP对等体组”。

1.3.3 配置建立TCP连接使用的源接口

BGP 使用 TCP 作为其传输层协议，缺省情况下，BGP 使用到达对等体最佳路由的出接口作为与对等体/对等体组建立 TCP 连接的源接口。

当建立 BGP 连接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 连接的源接口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，建议网络管理员使用 Loopback 接口建立 BGP 连接，并将建立 BGP 连接所使用的源接口配置为 Loopback 接口，来提高 BGP 连接的可靠性和稳定性。

表1-4 配置建立 TCP 连接使用的源接口

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
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*	必选 缺省情况下，BGP使用到达BGP对等体的最佳路由的出接口作为与对等体/对等体组创建BGP会话时建立TCP连接的源接口

注意

当两个设备之间建立多条 BGP 连接时，如果没有为每个连接明确指定建立 TCP 连接的源接口，可能会由于无法根据到达 BGP 对等体的最优路由确定 TCP 连接源接口从而导致无法建立 TCP 连接，在此情况下建议用户配置 BGP 对等体时明确配置 BGP 会话建立 TCP 连接的源接口。

1.3.4 配置允许同非直连邻居建立EBGP连接

当前路由器要与另外一个路由器建立 EBGP 连接，它们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** 命令允许它们之间经过多跳建立 TCP 连接。

表1-5 配置允许同非直连邻居建立 EBGP 连接

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置允许同非直接相连网络上的邻居建立EBGP连接	**peer** { *group-name* | *ip-address* } **ebgp-max-hop** [*hop-count*]	必选 缺省情况下，不允许同非直接相连网络上的邻居建立EBGP连接

说明

直连 EBGP 对等体不需要配置 **peer ebgp-max-hop** 命令。

1.4 控制路由信息的生成

BGP 与 IGP 不同，其着眼点主要在于控制路由的传播和选择最佳路由。

要生成 BGP 路由，主要有以下两种途径：

- 配置 BGP 发布本地路由
- 引入其它路由

网络管理员也可以根据需要配置引入其它协议的缺省路由。

1.4.1 配置准备

在控制 BGP 路由信息的生成之前，需要先创建 BGP 连接。

1.4.2 配置BGP发布本地路由

用户可以在 BGP 视图下配置发布某个网段的路由，从而将该路由发布给对等体，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。需要注意的是，要发布的本地路由必须存在于 IP 路由表中。

表1-6 配置 BGP 发布本地路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
将网段路由发布到BGP路由表中	**network** *ip-address* [*mask* | *mask-length*] **route-policy** *route-policy-name*	必选 缺省情况下，BGP不发布任何网段路由

1.4.3 配置BGP引入其它路由

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 的路由信息引入到 BGP 路由表中，并发布给对等体。在引入 IGP 路由时，可以针对不同的路由协议来对路由信息进行过滤。

通过引入方式发布的路由的 Origin 属性为 incomplete。

表1-7 配置 BGP 引入其它路由

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	**bgp** *as-number*	-
引入其它协议路由信息并通告	**import-route** *protocol* [{ *process-id* | **all-processes** } [**allow-direct** | **med** *med-value* | **route-policy** *route-policy-name*] *]	必选 缺省情况下，BGP不引入其它协议的路由 目前只有OSPF协议支持配置**allow-direct**参数

说明

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。

1.4.4 配置BGP引入其它协议缺省路由

BGP 不能通过 **import-route** 命令从其它协议引入缺省路由，如果要引入其它协议缺省路由，必须要使用下面命令进行配置。

表1-8 配置 BGP 引入其它协议缺省路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
引入其它协议路由信息并通告	**import-route** *protocol* [{ *process-id* | **all-processes** } [**allow-direct** | **med** *med-value* | **route-policy** *route-policy-name*] *]	必选 缺省情况下，BGP不引入其它协议的路由 目前只有OSPF协议支持配置**allow-direct**参数
允许将缺省路由引入到BGP路由表中	**default-route imported**	必选 缺省情况下，BGP不允许将缺省路由引入到BGP路由表中

1.5 控制路由信息的发布与接收

1.5.1 配置准备

在控制 BGP 路由信息的发布与接收之前，需要先配置 BGP 连接。

1.5.2 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，需要配置路由聚合，减小对等体路由表中的路由数量。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。

1. 配置路由自动聚合

配置自动聚合功能后，BGP 将对引入的 IGP 子网路由进行聚合，不再发布子网路由，而是发布聚合后的自然网段的路由。

需要注意的是，用 **network** 命令发布的路由不能进行自动聚合。

表1-9 配置路由自动聚合

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置对引入的IGP子网路由进行自动聚合	**summary automatic**	必选 缺省情况下，不对引入的IGP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 IGP 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1-10 配置路由手动聚合

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置路由手动聚合	**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*] *	必选 缺省情况下，不进行路由聚合

1.5.3 配置向对等体/对等体组发送缺省路由

通过配置向对等体/对等体组发送缺省路由，不论本地路由表中是否存在缺省路由，都将向指定对等体/对等体组发布一条下一跳地址为本地地址的缺省路由。

表1-11 配置向对等体/对等体组发送缺省路由

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
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*]	必选 缺省情况下，不向对等体/对等体组发送缺省路由

1.5.4 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
- IP 地址前缀列表
- 路由策略
- AS 路径过滤列表

关于访问控制列表的详细配置过程，请参考“ACL 和 QoS 配置指导”中的“ACL”。

关于 IP 地址前缀列表、路由策略和 AS 路径过滤列表的详细配置过程，请参考“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

用户可以使用路由策略、访问控制列表、AS 路径过滤列表或 IP 地址前缀列表对发布给对等体的路由信息进行过滤，可以根据需求选择过滤策略。

表1-12 配置 BGP 路由信息的发布策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置引入的路由在发布时进行过滤	**filter-policy** { *acl-number* | **ip-prefix** *ip-prefix-name* } **export** [**direct** | **isis** *process-id* | **ospf** *process-id* | **rip** *process-id* | | **static**]	至少选其一 缺省情况下，不对发布的路由信息进行过滤 可以根据需求选择过滤策略 同时配置几种过滤策略时，按照如下顺序执行： • **filter-policy export** • **peer filter-policy export** • **peer as-path-acl export** • **peer ip-prefix 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
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**	
为对等体/对等体组设置基于ACL的过滤策略	**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BGP路由过滤策略	**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IP前缀列表的路由过滤策略	**peer** { *group-name* | *ip-address* } **ip-prefix** *ip-prefix-name* **export**	

3. 配置BGP路由信息的接收策略

用户可以使用路由策略、访问控制列表、AS 路径过滤列表或 IP 地址前缀列表对从等体接收的路由信息进行过滤，只有满足某些条件的路由才能被 BGP 接收，并加到路由表中。

网络管理员可以根据需求选择过滤策略，可以对从所有对等体接收的路由信息进行过滤，也可以从指定对等体接收的路由信息进行过滤。

需要注意的是，对等体组的成员可以与所在的组使用不同的入方向路由策略，即接收路由时，各对等体可以选择自己的策略。

表1-13 配置 BGP 路由信息的接收策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
对接收的路由信息进行过滤	**filter-policy** { *acl-number* | **ip-prefix** *ip-prefix-name* } **import**	至少选其一 缺省情况下，不对接收的路由信息进行过滤 可以根据需求选择过滤策略 同时配置几种过滤策略时，按照如下顺序执行： • **filter-policy import** • **peer filter-policy import** • **peer as-path-acl import** • **peer ip-prefix 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
对来自对等体/对等体组的路由指定路由策略	**peer** { *group-name* | *ip-address* } **route-policy** *route-policy-name* **import**	
为对等体/对等体组设置基于ACL的过滤策略	**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BGP路由过滤策略	**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IP前缀列表的路由过滤策略	**peer** { *group-name* | *ip-address* } **ip-prefix** *ip-prefix-name* **import**	

1.5.5 配置BGP与IGP路由同步

BGP 路由器收到一条 IBGP 路由，缺省只检查该路由的下一跳是否可达。如果可达， BGP 路由器就将这条 IBGP 路由发布给 EBGP 对等体。如果配置了同步特性，则需要同时满足以下条件，BGP 路由器才会将该 IBGP 路由发布给 EBGP 对等体：

- 该路由的下一跳可达；
- 在 IGP 路由表中存在一条 active 的路由，与该 IBGP 路由的目的网段完全相同。（IGP 路由是否为 active 状态可以通过 **display ip routing-table protocol** 命令来查看）

表1-14 配置 BGP 与 IGP 路由同步

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
配置BGP与IGP路由同步	**synchronization**	必选 缺省情况下，BGP和IGP路由不同步

1.5.6 限制从BGP对等体/对等体组接收的路由前缀数量

表1-15 限制从 BGP 对等体/对等体组接收的路由前缀数量

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置当从指定对等体/对等体组收到的路由前缀数量超过最大值时，路由器中断BGP连接	**peer** { *group-name* | *ip-address* } **route-limit** *prefix-number* [*percentage-value*]	三者必选其一 缺省情况下，没有限制从BGP对等体/对等体组接收的路由前缀数量
配置当从指定对等体/对等体组收到的路由前缀数量超过最大值时，仅打印告警信息，路由器保持与指定对等体/对等体组的连接	**peer** { *group-name* | *ip-address* } **route-limit** *prefix-number* **alert-only** [*percentage-value*]	
配置当从指定对等体/对等体组收到的路由前缀数量超过允许的最大值时，路由器中断BGP连接，但会自动与对等体重建连接	**peer** { *group-name* | *ip-address* } **route-limit** *prefix-number* **reconnect** *reconnect-time* [*percentage-value*]	

1.5.7 配置BGP路由衰减

通过配置 BGP 衰减，可以抑制不稳定的路由信息，不将这类路由加入到路由表中，也不将这类路由向其他 BGP 对等体发布。

表1-16 配置 BGP 路由衰减

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置BGP路由衰减	**dampening** [*half-life-reachable half-life-unreachable reuse suppress ceiling* | **route-policy** *route-policy-name*] *	必选 缺省情况下，没有配置BGP路由衰减

1.5.8 配置shortcut路由

缺省情况下，接收到的 EBGP 路由的路由管理优先级为 255，比本地生成的路由的管理优先级低。通过将一条 EBGP 路由被配置成 short-cut，那么这条路由在加入到路由表时将使用本地生成路由的路由管理值，而不是 EBGP 路由的路由管理值，此时，这条路由成为最佳路由的可能性被提高了很多。

表1-17 配置 shortcut 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**short-cut**路由	**network** *ip-address* [*mask* | *mask-length*] **short-cut**	可选 缺省情况下，接收到的EBGP路由的路由管理优先级为255

1.6 配置BGP的路由属性

BGP 具有很多路由属性，利用这些属性可以改变 BGP 的选路策略。

1.6.1 配置准备

在配置 BGP 的路由属性之前，需要先配置 BGP 连接。

1.6.2 为接收路由分配首选值

BGP 选择路由时将首先丢弃下一跳不可达的路由，其次再优选 Preferred-value 值最大的路由。

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1-18 为接收路由分配首选值

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
为从对等体/对等体组接收的路由分配首选值	**peer** { *group-name* | *ip-address* } **preferred-value** *value*	可选 缺省情况下，从对等体/对等体组接收的路由的首选值为0

1.6.3 配置BGP的路由优先级

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。

用户可以修改 EBGP 路由、IBGP 路由以及用本地产生的 BGP 路由的路由优先级；还可以应用路由策略为匹配过滤条件的特定路由配置优先级，对于那些没有匹配的路由，使用缺省优先级。

表1-19 配置 BGP 的路由优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置BGP路由的管理优先级	**preference** { *external-preference internal-preference local-preference* | **route-policy** *route-policy-name* }	可选 缺省情况下，EBGP路由的管理优先级为255，IBGP路由的管理优先级为255，本地产生的BGP路由的管理优先级为130

1.6.4 配置Local-Preference缺省值

Local-Preference 用来判断流量离开 AS 时的最佳路由。当 BGP 的路由器通过不同的 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-Preference 值较高的路由。

用户可以通过 **default local-preference** 命令改变 BGP 路由器向 IBGP 对等体发送的路由 Local-Preference 的缺省值。

表1-20 配置 Local-Preference 缺省值

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置本地优先级的缺省值	**default local-preference** *value*	可选 缺省情况下，本地优先级的缺省值为100

1.6.5 配置MED属性

MED 用来判断流量进入 AS 时的最佳路由。当一个运行 BGP 的路由器通过不同的 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1-21 配置 MED 缺省值

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置系统MED的缺省值	**default med** *med-value*	可选 缺省情况下，MED的缺省值为0

2. 配置允许比较来自不同AS邻居路由的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1-22 配置允许比较来自不同 AS 邻居路由的 MED 属性值

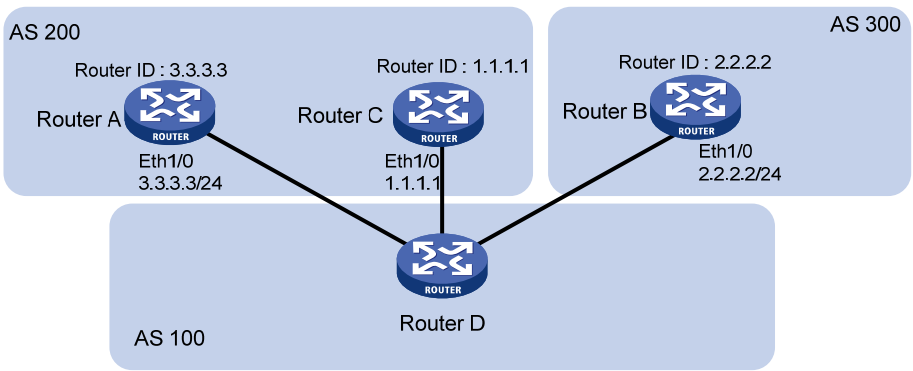
操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置允许比较来自不同AS邻居的路由路径的MED属性值	**compare-different-as-med**	必选 缺省情况下，不允许比较来自不同AS邻居的路由路径的MED属性值

3. 配置对来自同一AS的路由进行MED排序优选

缺省情况下，系统不会对来自同一 AS 的路由进行 MED 排序优选。

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1-16 MED 排序优选示意图



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，那么由于来自同一 AS 200 且 MED 值更大，应该视为无效路由。

在 Router D 上配置 **bestroute compare-med** 命令，Router D 对于学习到新的路由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，选出来 MED 值最小的路由，然后再对来自不同 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  10.0.0.0          3.3.3.3         50                     0       200e
 * i                    1.1.1.1         60                     0       200e
```

需要注意的是，BGP 将无法进行负载分担，因为 BGP 要求进行负载分担的路由必须 AS PATH 相同。

表1-23 配置对来自同一 AS 的路由进行 MED 排序优选

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置根据路由来自的AS进行分组，并根据MED排序优选	**bestroute compare-med**	必选 缺省情况下，不根据路由来自的AS进行分组对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1-24 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

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
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**	必选 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

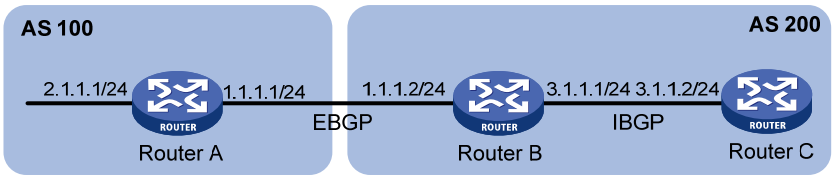
说明

只有 AS_PATH 里不包含联盟体外的子自治系统编号时，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，例如有三条路由，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子自治系统编号，因此在选择最优路由时第一条路由将成为最优路由。

1.6.6 配置下一跳属性

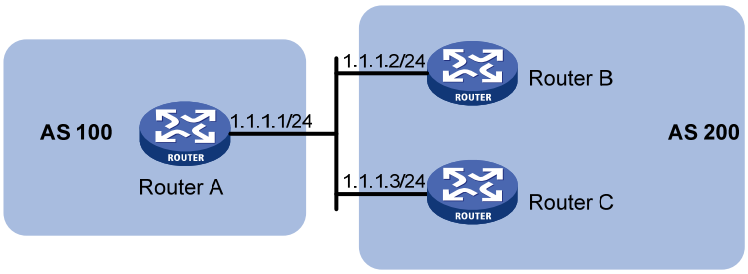
缺省情况下，路由器向 IBGP 对等体/对等体组发布路由时，不将自身地址作为下一跳，但有的时候为了保证 IBGP 邻居能够找到下一跳，可以配置将自身地址作为下一跳。以下图为例，Router A 与 Router B 建立 EBGP 邻居关系，Router B 与 Router C 建立 IBGP 邻居关系，Router B 在向 Router C 发布从 Router A 学到的 BGP 路由时，如果 Router C 上没有到达 1.1.1.1/24 的路由，可以在 Router B 上配置 **peer next-hop-local** 命令将 3.1.1.1/24 作为下一跳，这样，Router C 就能找到下一跳。

图1-17 配置 BGP 下一跳属性应用组网图一



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1-18 配置 BGP 下一跳属性应用组网图二



需要注意的是，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。

表1-25 配置 BGP 的下一跳属性

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置发布路由时将自身地址作为下一跳	**peer** { *group-name* | *ip-address* } **next-hop-local**	可选 缺省情况下，向EBGP对等体/对等体组发布路由时，将自身地址作为下一跳；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

1.6.7 配置AS_PATH属性

1. 配置允许本地AS号重复出现的次数

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。

通过配置**peer allow-as-loop**命令，可以允许本地AS号在所接收的路由的AS_PATH属性中出现，并可同时配置允许重复的次数。

表1-26 配置允许本地 AS 号重复出现的次数

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置允许本地AS号重复出现的次数	**peer** { *group-name* | *ip-address* } **allow-as-loop** [*number*]	必选 缺省情况下，不允许本地AS重复出现

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1-27 禁止路由器将 AS_PATH 当作选路算法中的一个因素

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
禁止路由器将AS_PATH当作选路算法中的一个因素	**bestroute as-path-neglect**	必选 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/对等体组定制一个虚拟的自治系统号

当网络管理员需要进行系统移植时，例如，Router A 原来位于 AS 2，现在需要将它移植到 AS 3 里，可以在 Router A 上为已经创建的 EBGP 对等体（组）配置一个虚拟的本地自治系统号 2，将本地真实的 AS 号 3 隐藏起来；由于位于其他 AS 内的 EBGP 对等体只能看到这个虚拟 AS 号 2，因此它们不需要改变现有配置，从而在系统升级调试过程中可以不中断现有的 BGP 服务。

表1-28 为对等体/对等体组定制一个虚拟的自治系统号

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
为对等体/对等体组定制一个虚拟的自治系统号	**peer** { *group-name* | *ip-address* } **fake-as** *as-number*	必选 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

说明

peer fake-as 命令只适用于 EBGP 对等体/对等体组。

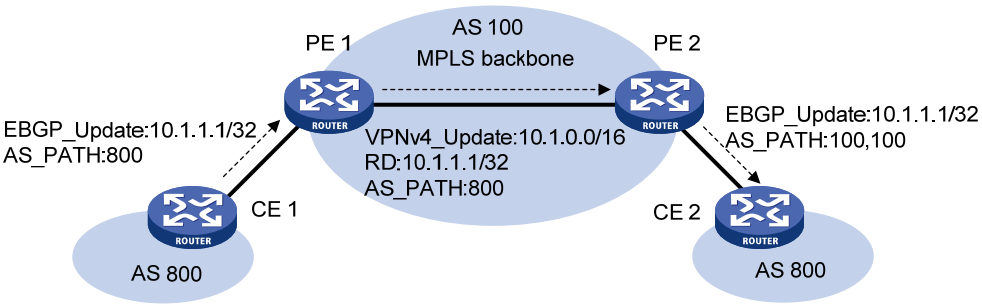
4. 配置AS号替换功能

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。

如果物理分散的 CE 复用相同的 AS 号，就应该在 PE 上配置 BGP 的 AS 号替换功能。此功能是 BGP 的出口策略，在发布路由时有效。

使能了 BGP 的 AS 号替换功能后，当 PE 向指定对等体中的 CE 发布路由时，如果路由的 AS_PATH 中有与 CE 相同的 AS 号，将被替换成 PE 的 AS 号后再发布。

图1-19 BGP AS 号替换应用示意图



如 图 1-19 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，这需要注意的是，如果需要完全的连接性，PE 1 上也需要做类似的配置。

表1-29 配置 AS 号替换功能

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
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**	必选 缺省情况下，没有用本地AS号替换AS_PATH属性中指定对等体/对等体组的AS号

注意

替换 AS_PATH 属性中的 AS 号命令仅在特定组网环境下使用，错误的配置会引起路由环路。

5. 配置发送BGP更新报文时AS_PATH属性中不携带私有AS号

表1-30 配置发送 BGP 更新报文时 AS_PATH 属性中不携带私有 AS 号

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置发送BGP更新报文时AS_PATH属性中不携带私有AS号	**peer** { *group-name* | *ip-address* } **public-as-only**	必选 缺省情况下，发送BGP更新报文时，携带私有自治系统号

6. 配置不检测EBGP路由的第一个AS号

通常情况下，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1-31 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**	必选 缺省情况下，系统收到EBGP路由后，会检测路由的第一个AS号

1.7 调整和优化BGP网络

1.7.1 配置准备

在调整和优化 BGP 网络之前，需要先配置 BGP 连接。

1.7.2 配置BGP存活时间间隔与保持时间

当对等体间建立了 BGP 连接后，它们定时向对端发送 Keepalive 消息，以防止路由器认为 BGP 连接已中断。若路由器在设定的连接保持时间（Holdtime）内未收到对端的 Keepalive 消息或 Update 报文，则认为此 BGP 连接已中断，从而断开此 BGP 连接。

用户可以配置当前路由的 BGP 存活时间间隔与保持时间，也可以配置与指定对等体/对等体组建立连接的存活时间间隔和保持时间：

- 如果当前路由器的保持时间与对等体/对等体组上配置的不一致，则数值较小者作为协商后的保持时间。
- 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发送 keepalive 消息。

- 存活时间间隔不为 0 时，以协商的保持时间的三分之一再与配置的存活时间间隔比较取最小值作为存活时间间隔发送 keepalive 消息。

表1-32 配置 BGP 存活时间间隔与保持时间

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置BGP的存活时间间隔与保持时间	**timer keepalive** *keepalive* **hold** *holdtime*	可选 缺省情况下，存活时间间隔为60秒，保持时间为180秒
配置指定对等体/对等体组的存活时间间隔和保持时间	**peer** { *group-name* | *ip-address* } **timer keepalive** *keepalive* **hold** *holdtime*	

注意

- 合理的最大 Keepalive 消息发送间隔为保持时间的三分之一，且该发送间隔不能小于 1 秒，因此，保持时间如果不为 0，则最小为 3 秒。
- 使用 **timer** 命令配置的定时器比使用 **peer timer** 命令配置的定时器优先级要低。
- 如果与对等体的邻居关系已经建立，必须要复位 BGP 连接才能使新配置的存活时间间隔与保持时间生效。
- 配置 **timer** 命令后，不会影响已建立的 BGP 连接，只对新建立的连接生效。
- 配置 **peer timer** 命令后，如果新配置的值与之前的配置值不同，会影响已建立的 BGP 连接：BGP 连接会直接断开，再使用新参数重新建立连接。

1.7.3 配置发送路由更新报文的时间间隔

表1-33 配置发送路由更新报文的时间间隔

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置对等体/对等体组发送同一路由更新报文的时间间隔	**peer** { *group-name* | *ip-address* } **route-update-interval** *interval*	可选 缺省情况下，向IBGP对等体发送同一路由更新报文的时间间隔为15秒，向EBGP对等体发送同一路由更新报文的时间间隔为30秒

1.7.4 配置BGP软复位

BGP 的选路策略改变后，为了使新的策略生效，必须复位 BGP 连接，但这样会造成短暂的 BGP 连接中断。

如果当前路由器和对等体都支持 Route-Refresh 功能，通过使能 Route-Refresh 功能，在策略改变时，系统可以在不中断 BGP 连接的情况下，自动对 BGP 路由表进行动态刷新。

如果对等体不支持 Route-Refresh 功能，则需要将从对等体接收的所有路由更新保存在本地，当选路策略发生改变后，在不中断连接的情况下重新刷新 BGP 路由表，并应用新的策略。

1. 通过使能Route-Refresh实现BGP软复位

在对等体支持并使能 Route-Refresh 功能的情况下，如果 BGP 的路由策略发生了变化，本地路由器会向 BGP 对等体发布 Router-Refresh 消息，收到此消息的对等体会将其路由信息重新发给本地路由器。这样，在不中断 BGP 连接的情况下，就可以对 BGP 路由表进行动态更新，并应用新的选路策略。

表1-34 通过使能 Route-Refresh 实现 BGP 软复位

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路由刷新功能	**peer** { *group-name* | *ip-address* } **capability-advertise route-refresh**	可选 缺省情况下，BGP路由刷新功能处于使能状态

2. 通过将所有路由更新保存在本地实现BGP软复位

当对等体不支持 Route-Refresh 功能时，可通过配置 **peer keep-all-routes** 命令将从对等体接收的所有路由更新保存在本地用户，然后再执行 **refresh bgp** 命令对保存在本地的所有路由使用路由策略重新过滤一遍。

表1-35 通过将所有路由更新保存在本地实现 BGP 软复位

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
禁止BGP路由刷新和多协议扩展功能	**peer** { *group-name* | *ip-address* } **capability-advertise conventional**	必选 缺省情况下，BGP路由刷新和多协议扩展功能处于使能状态
保存所有来自对等体/对等体组的原始路由信息	**peer** { *group-name* | *ip-address* } **keep-all-routes**	必选 缺省情况下，不保存所有来自对等体/对等体组的原始路由信息
退回用户视图	**return**	-
手工对BGP连接进行软复位	**refresh bgp** { **all** | *ip-address* | **group** *group-name* | **external** | **internal** } { **export** | **import** }	必选

注意

如果对端邻居不支持 Route-Refresh 功能，同时本端又未配置 **peer keep-all-routes** 命令，那么在修改 BGP 路由信息的接收策略后，需要根据策略的影响来决定是否手工断开邻居，重新学习路由。

1.7.5 配置BGP ORF能力

BGP ORF（Outbound Route Filtering，输出路由过滤）功能是将本地入口策略通过 Route-refresh 报文发送给邻居，当邻居需要向 BGP 对等体发送 Update 更新报文时，通过本地的路由策略后还需要进行 ORF 策略的过滤，只有通过 ORF 策略的路由信息才会发给 BGP 对等体，以达到减少 BGP 邻居间 Update 更新报文的交互，节省网络资源的目的。

使能 BGP ORF 能力后，本地和 BGP 对等体会通过 Open 报文协商 ORF 能力（即收发的报文里是否允许携带 ORF 信息，如果允许携带，是否可以携带非标准的 ORF 信息），当协商完毕并成功建立邻居关系后，可以通过特殊的 Route-refresh 报文交互 ORF 信息。

ORF能力协商成功需要两端的配置来保证，关于两端参数的选择请参见 表 1-37。

表1-36 配置 BGP ORF 能力

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路由刷新功能	**peer** { *group-name* | *ip-address* } **capability-advertise route-refresh**	必选 缺省情况下，BGP路由刷新功能处于使能状态
使能BGP邻居的非标准ORF能力	**peer** { *group-name* | *ip-address* } **capability-advertise orf non-standard**	可选 缺省情况下，BGP ORF能力支持RFC5291、RFC5292的标准能力 如果对端只支持非标准值的ORF报文，则必须配置该命令
使能BGP邻居的ORF能力	**peer** { *group-name* | *ip-address* } **capability-advertise orf ip-prefix** { **both** | **receive** | **send** }	必选 缺省情况下，BGP对等体/对等体组不使能邻居的ORF能力

表1-37 both、send、receive 参数选择以及配置效果描述表

本地选择参数	对端选择参数	协商成功后
send	**receive**	本端的ORF发送能力，对端的ORF接收能力
	both	
receive	**send**	本端的ORF接收能力，对端的ORF发送能力
	both	

本地选择参数	对端选择参数	协商成功后
both	**both**	双向的ORF发送和接收能力

1.7.6 使能 4 字节AS号抑制功能

通常情况下，设备在与对端设备建立 BGP 对等体连接关系时会发送 BGP Open 消息，在消息中的 Optional parameters 字段携带信息：说明 AS 号取值占用 4 字节（即取值范围为 1～4294967295）。此时，如果对端设备不支持 AS 号取值为 4 字节（比如只支持 2 字节），则连接关系无法建立。

为了解决上述问题，您可以使能 4 字节 AS 号抑制功能：当对端设备不支持 AS 号取值为 4 字节时，设备发出的 BGP Open 消息仍然可以被对端设备正常识别，从而 BGP 对等体连接关系可以成功建立。

表1-38 使能 4 字节 AS 号抑制功能

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
使能4字节AS号抑制功能	**peer** { *group-name* | *ip-address* } **capability-advertise suppress-4-byte-as**	必选 缺省情况下，设备没有使能4字节AS号抑制功能

注意

如果对端设备支持 4 字节 AS 号能力时，请不要使能该功能，否则会导致对等体无法建立。

1.7.7 配置BGP报文的DSCP优先级

在 IPv4 报文头中，包含一个 8bit 的 ToS 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 BGP 报文的 DSCP 优先级进行配置。

表1-39 配置向 BGP 邻居发送报文的 DSCP 优先级

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*	二者必选其一
	进入BGP-VPN实例视图	**bgp** *as-number*	
		ipv4-family vpn-instance *vpn-instance-name*	

操作	命令	说明
配置向指定对等体/对等体组发送BGP报文的DSCP优先级	**peer** { *group-name* | *ip-address* } **dscp** dscp-value	可选 缺省情况下，BGP报文使用的DSCP优先级为48

1.7.8 使能EBGP连接快速复位功能

路由器在设定的连接保持时间（Holdtime）内未收到对端的 Keepalive 消息或 Update 报文，则认为此 BGP 连接已中断，从而断开此 BGP 连接。

通过使能 EBGP 连接快速复位功能，如果当前路由器与直连的 EBGP 邻居链路变为 down 状态，系统将重新建立与该邻居的 EBGP 会话。

表1-40 使能 EBGP 连接快速复位功能

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
使能EBGP连接快速复位功能	**ebgp-interface-sensitive**	可选 缺省情况下，使能EBGP连接快速复位功能处于关闭状态

1.7.9 配置BGP建立TCP连接时进行MD5 认证

BGP 使用 TCP 作为传输层协议，为提高 BGP 的安全性，可以配置 BGP 在建立 TCP 连接时进行 MD5 认证，即两台路由器必须配置相同的密码，才能建立 TCP 连接。

BGP 的 MD5 认证并不能对 BGP 报文认证，它只是为 TCP 连接设置 MD5 认证密码，由 TCP 完成认证。如果认证失败，则不建立 TCP 连接。

表1-41 配置 BGP 建立 TCP 连接时进行 MD5 认证

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置BGP建立TCP连接时进行MD5认证	**peer** { *group-name* | *ip-address* } **password** { **cipher** | **simple** } *password*	可选 缺省情况下，BGP在建立TCP连接时不进行MD5认证

1.7.10 配置BGP负载分担

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 BGP 协议的最大等价路由条数。

表1-42 配置 BGP 负载分担

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置进行BGP负载分担的路由条数	**balance** *number*	可选 缺省情况下，BGP不进行路由负载分担

1.7.11 禁止与对等体/对等体组建立会话

表1-43 禁止与对等体/对等体组建立会话

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
禁止与对等体/对等体组建立会话	**peer** { *group-name* | *ip-address* } **ignore**	可选 缺省情况下，允许与BGP对等体/对等体组建立会话

1.8 配置大型BGP网络

在大型 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。

1.8.1 配置准备

在配置大型 BGP 网络之前，需完成以下任务：相邻节点的网络层互通。

1.8.2 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。

需要注意的是，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。

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1-44 配置 IBGP 对等体组

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	- 缺省情况下，系统没有运行BGP
创建IBGP对等体组	**group** *group-name* [**internal**]	必选
向对等体组中加入对等体	**peer** *ip-address* **group** *group-name*	必选

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 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；
- 第二种方式是创建对等体组完毕后，先配置对等体的 AS 号，再将对等体加入对等体组中。
- 第三种方式是创建对等体组完毕后，将对等体加入对等组的同时指定 AS 号。

表1-45 配置 EBGP 对等体组（方式一）

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**	必选
设置对等体组的AS号	**peer** *group-name* **as-number** *as-number*	必选
向对等体组中加入对等体	**peer** *ip-address* **group** *group-name*	必选

说明

- 该方式下，对等体加入对等体之前不能配置 AS 号；
- 该方式下，对等体组中所有对等体的 AS 号等于对等体组的 AS 号。

表1-46 配置 EBGP 对等体组（方式二）

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**	必选
设置对等体的AS号	**peer** *ip-address* **as-number** *as-number*	必选
向对等体组中加入对等体	**peer** *ip-address* **group** *group-name* [**as-number** *as-number*]	必选 **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

说明

该方式下，对等体组中对等体的 AS 号可以相同也可以不同。

表1-47 配置 EBGP 对等体组（方式三）

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**	必选
向对等体组中加入对等体	**peer** *ip-address* **group** *group-name* **as-number** *as-number*	必选

说明

- 该方式下，对等体加入对等体之前不能配置 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。

1.8.3 配置BGP团体

团体是一组有相同特征的目的地址的集合，没有物理上的边界，与其所在的 AS 无关。

用户可以先通过路由策略定义一组目的地址属于某个指定团体，并配置向对等体/对等体组发布团体属性：

- 在向对等体/对等体组发布路由时可以指定发布策略，在发布策略中通过团体属性对发布给对等体/对等体组的路由信息进行过滤；

- 对等体/对等体组在接收路由信息时也可以指定接收策略，在接收策略中通过团体属性对从等体/对等体组接收的路由信息进行过滤。

在路由策略中通过团体来进行路由过滤，可以提高路由策略配置的灵活度，并简化路由策略的管理，从而降低维护管理的难度。

关于路由策略的详细配置，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

表1-48 配置 BGP 团体

操作		命令	说明
进入系统视图		**system-view**	-
进入BGP视图		**bgp** *as-number*	-
配置向对等体/对等体组发布团体属性	配置向对等体/对等体组发布团体属性	**peer** { *group-name* | *ip-address* } **advertise-community**	必选 缺省情况下，不将团体属性和扩展团体属性发布给任何对等体/对等体组
	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**	
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**	必选 缺省情况下，不指定对等体/对等体组的路由策略

1.8.4 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反射器，其它路由器作为客户机。

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，以避免路由循环。

表1-49 配置 BGP 路由反射器

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
配置将本机作为路由反射器，并将对等体/对等体组作为路由反射器的客户	**peer** { *group-name* | *ip-address* } **reflect-client**	必选 缺省情况下，没有配置路由反射器及其客户 BGP视图和BGP-VPNv4子地址族视图下均支持**peer reflect-client**命令。当在BGP视图下配置该命令时，表示本机可以反射公网路由；当在BGP-VPNv4子地址族视图下配置该命令时，表示本机可以反射私网路由。（在BGP视图下使用**ipv4-family vpnv4**命令可以进入BGP-VPNv4子地址族视图，**ipv4-family vpnv4**命令的详细介绍请参见“MPLS命令参考”中的“MPLS L3VPN”）

操作	命令	说明
配置允许客户到客户的路由反射	**reflect between-clients**	可选 缺省情况下，允许客户到客户的路由反射
配置路由反射器的集群ID	**reflector cluster-id** *cluster-id*	可选 缺省情况下，每个路由反射器是使用自己的Router ID作为集群ID

注意

- 通常情况下，路由反射器的客户之间不要求是全连接的，路由缺省通过反射器从一个客户反射到其它客户；如果客户之间是全连接的，可以禁止客户间的反射，以便减少开销。
- 通常，一个集群里只有一个路由反射器。此时是由反射器的路由器 ID 来识别该集群的。设置多个路由反射器可提高网络的稳定性。如果一个集群中配有多个路由反射器，请使用 **reflector cluster-id** 命令为所有的路由反射器配置同样的集群 ID，以避免路由循环。

1.8.5 配置BGP联盟

联盟（Confederation）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) 首先，启动 BGP，并指定该路由器所属的子自治系统号；

(2) 其次，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1-50 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ID	**confederation id** *as-number*	必选 缺省情况下，未配置联盟的ID

操作	命令	说明
指定一个联盟体中包含了哪些子自治系统	**confederation peer-as** *as-number-list*	必选 缺省情况下，未指定一个联盟体中包含了哪些子自治系统

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1-51 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*	-
配置与未采用RFC 3065配置的AS联盟兼容	**confederation nonstandard**	可选 缺省情况下，配置的联盟与RFC 3065一致

1.9 配置BGP GR

分别在作为 GR Restarter 和 GR Helper 的设备上进行以下配置。

说明

一台设备可以同时具备 GR Restarter 和 GR Helper 能力。

表1-52 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**	必选 缺省情况下，BGP协议的GR能力处于关闭状态
配置对端重建BGP会话连接的最大时间	**graceful-restart timer restart** *timer*	可选 缺省情况下，对端重建BGP会话连接的最大时间为150秒
配置等待End-Of-RIB标记的时间	**graceful-restart timer wait-for-rib** *timer*	可选 缺省情况下，等待End-Of-RIB标记的时间为180秒

说明

- 通常情况下，对端重建会话连接的最大时间应小于 OPEN 消息中的 Holdtime 时间。接收侧（Receiving Speaker）检测到会话中断后，将等待指定时间以便重新建立会话连接。
- End-Of-RIB（End of Routing-Information-Base，路由信息库结束）标记用来标定路由更新发送的结束。

1.10 开启Trap功能

开启 BGP 模块的 Trap 功能后，该模块会生成级别为 level4 的警告 Trap 报文，用于报告该模块的重要事件。生成的 Trap 报文将被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 Trap 报文的输出规则（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1-53 开启 Trap 功能

操作	命令	说明
进入系统视图	**system-view**	-
开启BGP模块的Trap功能	**snmp-agent trap enable bgp**	可选 缺省情况下，BGP模块的Trap功能处于开启状态

1.11 使能BGP日志功能

表1-54 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*	- 缺省情况下，系统没有运行BGP
全局使能BGP日志功能	**log-peer-change**	可选 缺省情况下，全局BGP日志功能处于使能状态
记录指定对等体/对等体组的会话状态和事件信息	**peer** { *group-name* | *ip-address* } **log-change**	可选 记录对等体/对等体组的会话状态和事件信息

1.12 配置BGP与BFD联动

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，使得 BFD 能够为 BGP 邻居之间的链路提供更快速的检测。当邻居之间的链路出现故障时，能够加快 BGP 协议的收敛速度。

表1-55 配置 BGP 与 BFD 联动

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
在指定邻居上使能BFD	**peer** *ip-address* **bfd**	必选 缺省情况下，BGP的所有对等体邻居上都没有使能BFD链路检测功能

说明

- 配置 BFD 之前，需要使能 BGP 功能。
- 在 BGP 使能 GR 能力后，请慎用 BFD。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。
- 有关 BFD 的介绍和基本功能配置，请参见“可靠性配置指导”中的“BFD”。

1.13 BGP显示和维护

1.13.1 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1-56 BGP 配置显示

操作	命令
显示对等体组信息	**display bgp group** [*group-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示BGP发布的路由信息	**display bgp network** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示AS路径信息	**display bgp paths** [*as-regular-expression* | **|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
显示BGP对等体/对等体组的信息	**display bgp peer** [*ip-address* { **log-info** | **verbose** } | *group-name* **log-info** | **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示从指定邻居收到的ORF报文中的前缀信息	**display bgp peer** *ip-address* **received ip-prefix** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示BGP路由信息	**display bgp routing-table** [*ip-address* [{ *mask* | *mask-length* } [**longer-prefixes**]]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配指定AS路径过滤列表的BGP路由信息	**display bgp routing-table as-path-acl** *as-path-acl-number* [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示CIDR的BGP路由信息	**display bgp routing-table cidr** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配指定BGP团体的BGP路由信息	**display bgp routing-table community** [*aa:nn*&<1-13>] [**no-advertise** | **no-export** | **no-export-subconfed**] * [**whole-match**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配指定BGP团体列表的BGP路由	**display bgp routing-table community-list** { { *basic-community-list-number* | *comm-list-name* } [**whole-match**] | *adv-community-list-number* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示BGP衰减的BGP路由信息	**display bgp routing-table dampened** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示BGP衰减的配置参数	**display bgp routing-table dampening parameter** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示源AS不一致的路由	**display bgp routing-table different-origin-as** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示BGP路由震荡统计信息	**display bgp routing-table flap-info** [**regular-expression** *as-regular-expression* | [**as-path-acl** *as-path-acl-number* | *ip-address* [{ *mask* | *mask-length* } [**longer-match**]]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]]
显示BGP标签路由信息	**display bgp routing-table label** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示向指定的BGP对等体发送或者从BGP对等体收到的路由信息	**display bgp routing-table peer** *ip-address* { **advertised-routes** **|** **received-routes** } [*network-address* [*mask* | *mask-length*] | **statistic**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配AS正则表达式的路由信息	**display bgp routing-table regular-expression** *as-regular-expression*
显示BGP的路由统计信息	**display bgp routing-table statistic** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示全局路由器ID	**display router id** [**|** { **begin** | **exclude** | **include** } *regular-expression*]

1.13.2 复位BGP连接

当 BGP 路由策略或协议发生变化后，如果需要通过复位 BGP 连接使新的配置生效，请在用户视图下进行下列配置。

表1-57 复位 BGP 连接

操作	命令
复位所有BGP连接	**reset bgp all**
复位与指定AS之间的BGP连接	**reset bgp** *as-number*
复位与指定对等体的BGP连接	**reset bgp** *ip-address* [**flap-info**]
复位所有EBGP连接	**reset bgp external**
复位与指定对等体组的BGP连接	**reset bgp group** *group-name*
复位所有IBGP连接	**reset bgp internal**
复位IPv4单播地址族下的所有BGP连接	**reset bgp ipv4 all**

1.13.3 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1-58 清除 BGP 信息

操作	命令
清除路由的衰减信息并释放被抑制的路由	**reset bgp dampening** [*ip-address* [*mask* | *mask-length*]]
清除路由的震荡统计信息	**reset bgp flap-info** [*ip-address* [*mask-length* | *mask*] | **as-path-acl** *as-path-acl-number* | **regexp** *as-path-regular-expression*] **reset bgp** *peer-ip-address* **flap-info**

1.14 BGP典型配置举例

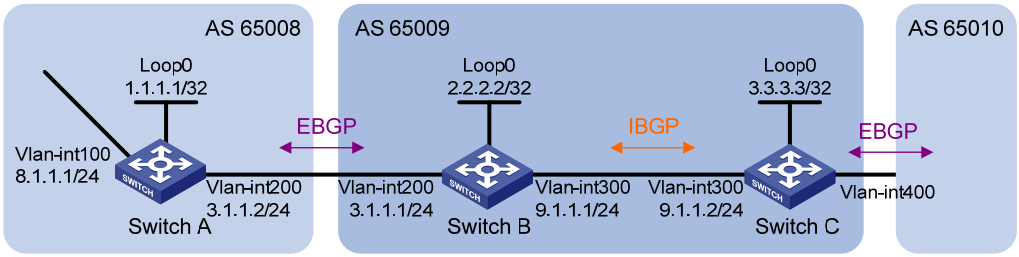
1.14.1 BGP基本配置

1. 组网需求

如 图 1-20 所示，所有设备均运行BGP协议，Switch A和Switch B之间建立EBGP连接，Switch B和Switch C之间建立IBGP连接。要求Switch C能够访问Switch A直连的 8.1.1.0/24 网段。

2. 组网图

图1-20 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 Switch B 到 Switch C 的 Loopback 接口路由可达，Switch C 到 Switch B 的 Loopback 接口路由可达。

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 3.3.3.3 as-number 65009
[SwitchB-bgp] peer 3.3.3.3 connect-interface loopback 0
[SwitchB-bgp] quit
[SwitchB] ospf 1
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 2.2.2.2 0.0.0.0
[SwitchB-ospf-1-area-0.0.0.0] network 9.1.1.1 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 2.2.2.2 as-number 65009
[SwitchC-bgp] peer 2.2.2.2 connect-interface loopback 0
[SwitchC-bgp] quit
[SwitchC] ospf 1
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 3.3.3.3 0.0.0.0
[SwitchC-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

```
[SwitchC] display bgp peer

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        2        2    0       0 00:00:13 Established
```

以上显示信息表明 Switch B 和 Switch 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 BGP 邻居。
- 因为要求 Switch C 能够访问 Switch 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 Switch A。

```
<SwitchA> system-view
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 3.1.1.1 as-number 65009
[SwitchA-bgp] network 8.1.1.1 24
[SwitchA-bgp] quit
```

配置 Switch B。

```
[SwitchB] bgp 65009
[SwitchB-bgp] peer 3.1.1.2 as-number 65008
[SwitchB-bgp] quit
```

查看 Switch B 的 BGP 对等体的连接状态。

```
[SwitchB] display bgp peer

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        4        4    0       0 00:02:49 Established
  3.1.1.2              65008        2        2    0       0 00:00:05 Established
```

可以看出，Switch B 与 Switch C、Switch B 与 Switch A 之间的 BGP 连接均已建立。

查看 Switch A 的 BGP 路由表。

```
[SwitchA] display bgp routing-table

 Total Number of Routes: 1

 BGP Local router ID is 1.1.1.1
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
```

```
              Origin : i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

 *>  8.1.1.0/24         0.0.0.0         0                     0       i
```

显示 Switch B 的 BGP 路由表。

```
[SwitchB] display bgp routing-table

 Total Number of Routes: 1

 BGP Local router ID is 2.2.2.2
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

 *>  8.1.1.0/24         3.1.1.2         0                     0       65008i
```

显示 Switch C 的 BGP 路由表。

```
[SwitchC] display bgp routing-table

 Total Number of Routes: 1

 BGP Local router ID is 3.3.3.3
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               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
```

说明

从路由表可以看出，Switch A 没有学到 AS 65009 内部的任何路由，Switch 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 Switch B 上配置 BGP 引入直连路由，以便 Switch A 能够获取到网段 9.1.1.0/24 的路由，Switch C 能够获取到网段 3.1.1.0/24 的路由。

配置 Switch B。

```
[SwitchB] bgp 65009
[SwitchB-bgp] import-route direct
```

显示 Switch A 的 BGP 路由表。

```
[SwitchA] display bgp routing-table

 Total Number of Routes: 4
```

```
 BGP Local router ID is 1.1.1.1
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? – incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

 *>  2.2.2.2/32         3.1.1.1         0                     0       65009?
 *   3.1.1.0/24         3.1.1.1         0                     0       65009?
 *>  8.1.1.0/24         0.0.0.0         0                     0       i
 *>  9.1.1.0/24         3.1.1.1         0                     0       65009?
```

以上显示信息表明，在 Switch B 上引入直连路由后，Switch A 新增了到 2.2.2.2/32 和 9.1.1.0/24 两条路由。

显示 Switch C 的 BGP 路由表。

```
[SwitchC] display bgp routing-table

 Total Number of Routes: 4

 BGP Local router ID is 3.3.3.3
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               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 Switch A 的地址。

(5) 结果验证

使用 Ping 进行验证。

```
[SwitchC] ping 8.1.1.1
  PING 8.1.1.1: 56  data bytes, press CTRL_C to break
    Reply from 8.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
    Reply from 8.1.1.1: bytes=56 Sequence=2 ttl=254 time=2 ms
    Reply from 8.1.1.1: bytes=56 Sequence=3 ttl=254 time=2 ms
    Reply from 8.1.1.1: bytes=56 Sequence=4 ttl=254 time=2 ms
    Reply from 8.1.1.1: bytes=56 Sequence=5 ttl=254 time=2 ms
  --- 8.1.1.1 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 2/2/2 ms
```

1.14.2 BGP与IGP交互配置

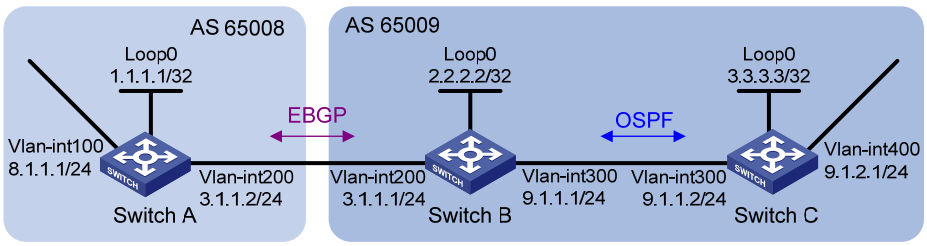
1. 组网需求

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 Switch A 和 Switch B 相连。

现要求实现 Switch A 能够访问 AS 65009 内的网段 9.1.2.0/24，Switch C 能够访问 AS 65008 内的网段 8.1.1.0/24。

2. 组网图

图1-21 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 Switch B 能获取到到 9.1.2.0/24 网段的路由。

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ospf 1
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 2.2.2.2 0.0.0.0
[SwitchB-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ospf 1
[SwitchC-ospf-1] import-route direct
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

(3) 配置 EBGP 连接

配置 EBGP 连接，并在 Switch A 上将 8.1.1.0/24 网段通告到 BGP 路由表中，以便 Switch B 获取到网段 8.1.1.0/24 的路由。

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] bgp 65008
```

```
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 3.1.1.1 as-number 65009
[SwitchA-bgp] network 8.1.1.0 24
[SwitchA-bgp] quit
```

配置 Switch B。

```
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 3.1.1.2 as-number 65008
```

(4) 配置 BGP 与 IGP 交互

- 在 Switch B 上配置 BGP 引入 OSPF 路由，以便 Switch A 能够获取到到 9.1.2.0/24 网段的路由。
- 在 Switch B 上配置 OSPF 引入 BGP 路由，以便 Switch C 能够获取到到 8.1.1.0/24 网段的路由。

在 Switch B 上配置 BGP 引入 OSPF 路由。

```
[SwitchB-bgp] import-route ospf 1
[SwitchB-bgp] quit
[SwitchB] ospf 1
[SwitchB-ospf-1] import-route bgp
[SwitchB-ospf-1] quit
```

查看 Switch A 的 BGP 路由表。

```
[SwitchA] display bgp routing-table

Total Number of Routes: 3

 BGP Local router ID is 1.1.1.1
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

 *>  3.3.3.3/32         3.1.1.1         1                     0       65009?
 *>  8.1.1.0/24         0.0.0.0         0                     0       i
 *>  9.1.2.0/24         3.1.1.1         1                     0       65009?
```

查看 SwitchC 的路由表。

```
[SwitchC] display ip routing-table
Routing Tables: Public
         Destinations : 9        Routes : 9

Destination/Mask    Proto  Pre  Cost         NextHop         Interface

2.2.2.2/32          OSPF   10   1            9.1.1.1         Vlan300
3.3.3.3/32          Direct 0    0            127.0.0.1       InLoop0
8.1.1.0/24          O_ASE  150  1            9.1.1.1         Vlan300
9.1.1.0/24          Direct 0    0            9.1.1.2         Vlan300
9.1.1.2/32          Direct 0    0            127.0.0.1       InLoop0
```

```
9.1.2.0/24          Direct 0    0           9.1.2.1         Vlan400
9.1.2.1/32          Direct 0    0           127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0           127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0           127.0.0.1       InLoop0
```

(5) 结果验证

使用 Ping 进行验证。

```
[SwitchA] ping -a 8.1.1.1 9.1.2.1
  PING 9.1.2.1: 56  data bytes, press CTRL_C to break
    Reply from 9.1.2.1: bytes=56 Sequence=1 ttl=254 time=15 ms
    Reply from 9.1.2.1: bytes=56 Sequence=2 ttl=254 time=31 ms
    Reply from 9.1.2.1: bytes=56 Sequence=3 ttl=254 time=47 ms
    Reply from 9.1.2.1: bytes=56 Sequence=4 ttl=254 time=46 ms
    Reply from 9.1.2.1: bytes=56 Sequence=5 ttl=254 time=47 ms

  --- 9.1.2.1 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 15/37/47 ms
[SwitchC] ping -a 9.1.2.1 8.1.1.1
  PING 8.1.1.1: 56  data bytes, press CTRL_C to break
    Reply from 8.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
    Reply from 8.1.1.1: bytes=56 Sequence=2 ttl=254 time=2 ms
    Reply from 8.1.1.1: bytes=56 Sequence=3 ttl=254 time=2 ms
    Reply from 8.1.1.1: bytes=56 Sequence=4 ttl=254 time=2 ms
    Reply from 8.1.1.1: bytes=56 Sequence=5 ttl=254 time=2 ms

  --- 8.1.1.1 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 2/2/2 ms
```

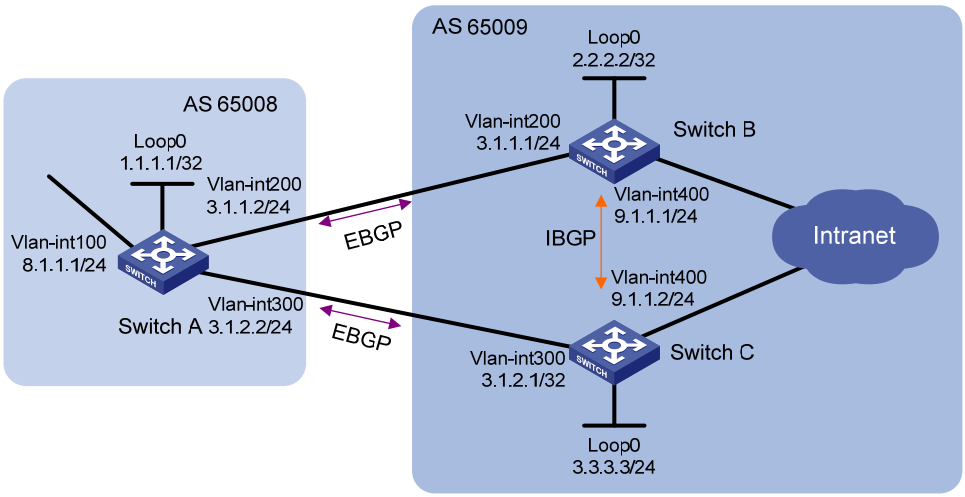
1.14.3 BGP负载分担配置

1. 组网需求

所有交换机都配置 BGP，Switch A 在 AS 65008 中，Switch B 和 Switch C 在 AS 65009 中。Switch A 与 Switch B、Switch C 之间运行 EBGP，Switch B 和 Switch C 之间运行 IBGP。在 Switch 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1-22 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 Switch A 上与 Switch B、Switch C 分别建立 EBGP 连接，并将 8.1.1.0/24 网段的路由通告给 Switch B 和 Switch C，以便 Switch B 和 Switch C 能够访问 Switch A 的内部网络。
- 在 Switch B 上与 Switch A 建立 EBGP 连接，与 Switch C 建立 IBGP 连接，并将 9.1.1.0/24 网段的路由通告给 Switch A，以便 Switch A 能够通过 Switch B 访问内部网络。同时，在 Switch B 上配置一条到 Switch 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 Switch C 上与 Switch A 建立 EBGP 连接，与 Switch B 建立 IBGP 连接，并将 9.1.1.0/24 网段的路由通告给 Switch A，以便 Switch A 能够通过 Switch C 访问内部网络。同时，在 Switch C 上配置一条到 Switch 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 Switch A。

```
<SwitchA> system-view
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 3.1.1.1 as-number 65009
[SwitchA-bgp] peer 3.1.2.1 as-number 65009
[SwitchA-bgp] network 8.1.1.1 24
[SwitchA-bgp] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 3.1.1.2 as-number 65008
[SwitchB-bgp] peer 3.3.3.3 as-number 65009
```

```
[SwitchB-bgp] peer 3.3.3.3 connect-interface loopback 0
[SwitchB-bgp] network 9.1.1.0 255.255.255.0
[SwitchB-bgp] quit
[SwitchB] ip route-static 3.3.3.3 32 9.1.1.2
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 3.1.2.2 as-number 65008
[SwitchC-bgp] peer 2.2.2.2 as-number 65009
[SwitchC-bgp] peer 2.2.2.2 connect-interface loopback 0
[SwitchC-bgp] network 9.1.1.0 255.255.255.0
[SwitchC-bgp] quit
[SwitchC] ip route-static 2.2.2.2 32 9.1.1.1
```

查看 Switch A 的路由表。

```
[SwitchA] display bgp routing-table

 Total Number of Routes: 3

 BGP Local router ID is 1.1.1.1
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? – incomplete

     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

 *>  8.1.1.0/24         0.0.0.0         0                     0       i
 *>  9.1.1.0/24         3.1.1.1         0                     0       65009i
 *                      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 Switch B 的路由器 ID 要小一些）；而下一跳为 3.1.2.1 的路由前只有标志 “*”，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 Vlan200。

(3) 配置负载分担

因为 Switch A 有两条路径到达 AS 65009 的内部网络，所以，在 Switch A 配置负载分担的路由条数为 2，以提高链路利用率。

配置 Switch A。

```
[SwitchA] bgp 65008
[SwitchA-bgp] balance 2
[SwitchA-bgp] quit
```

(4) 结果验证

查看 Switch A 的路由表。

```
[SwitchA] display bgp routing-table
```

```
Total Number of Routes: 3

BGP Local router ID is 1.1.1.1
Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete
     Network            NextHop        MED        LocPrf    PrefVal Path/Ogn

 *>  8.1.1.0/24         0.0.0.0         0                     0       i
 *>   9.1.1.0/24         3.1.1.1         0                   0       65009i
 *>                      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 Vlan200；另一条的下一跳地址为 3.1.2.1，出接口为 Vlan300。

1.14.4 BGP团体配置

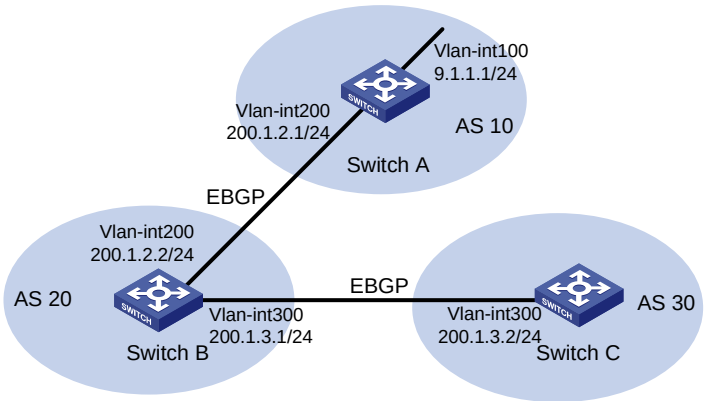
1. 组网需求

Switch B 分别与 Switch A、Switch C 之间建立 EBGP 连接。

通过在 Switch A 上配置 NO_EXPORT 团体属性，使得 AS 10 发布到 AS 20 中的路由，不再被 AS 20 向其他 AS 发布。

2. 组网图

图1-23 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 Switch A。

```
<SwitchA> system-view
[SwitchA] bgp 10
[SwitchA-bgp] router-id 1.1.1.1
```

```
[SwitchA-bgp] peer 200.1.2.2 as-number 20
[SwitchA-bgp] network 9.1.1.0 255.255.255.0
[SwitchA-bgp] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] bgp 20
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 200.1.2.1 as-number 10
[SwitchB-bgp] peer 200.1.3.2 as-number 30
[SwitchB-bgp] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] bgp 30
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 200.1.3.1 as-number 20
[SwitchC-bgp] quit
```

查看 Switch B 的路由表。

```
[SwitchB] display bgp routing-table 9.1.1.0

 BGP local router ID : 2.2.2.2
 Local AS number : 20
 Paths:   1 available, 1 best

BGP routing table entry information of 9.1.1.0/24:
 From            : 200.1.2.1 (1.1.1.1)
 Original nexthop: 200.1.2.1
 AS-path         : 10
 Origin          : igp
 Attribute value : MED 0, pref-val 0, pre 255
 State           : valid, external, best,
 Advertised to such 1 peers:
    200.1.3.2
```

可以看出，Switch B 把收到的路由发布给了位于 AS 30 内的 Switch C。

查看 Switch C 的路由表。

```
[SwitchC] display bgp routing-table

 Total Number of Routes: 1

 BGP Local router ID is 3.3.3.3
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
      Network            NextHop        MED        LocPrf    PrefVal Path/Ogn

 *>   9.1.1.0/24         200.1.3.1      0                    0       20 10i
```

从路由表可以确认，Switch C 从 Switch 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。

```
[SwitchA] route-policy comm_policy permit node 0
[SwitchA-route-policy] apply community no-export
[SwitchA-route-policy] quit
```

应用路由策略。

```
[SwitchA] bgp 10
[SwitchA-bgp] peer 200.1.2.2 route-policy comm_policy export
[SwitchA-bgp] peer 200.1.2.2 advertise-community
```

查看 Switch B 的路由表。

```
[SwitchB] display bgp routing-table 9.1.1.0
 BGP local router ID : 2.2.2.2
 Local AS number : 20
 Paths:   1 available, 1 best

 BGP routing table entry information of 9.1.1.0/24:
 From            : 200.1.2.1 (1.1.1.1)
 Original nexthop: 200.1.2.1
 Community       : No-Export
 AS-path         : 10
 Origin          : igp
 Attribute value : MED 0, pref-val 0, pre 255
 State           : valid, external, best,
 Not advertised to any peers yet
```

在 Switch B 的 BGP 路由表中可以看到配置的团体属性。此时在 Switch C 的 BGP 路由表中已经没有到目的地址 9.1.1.0/24 的路由。

1.14.5 BGP路由反射器配置

1. 组网需求

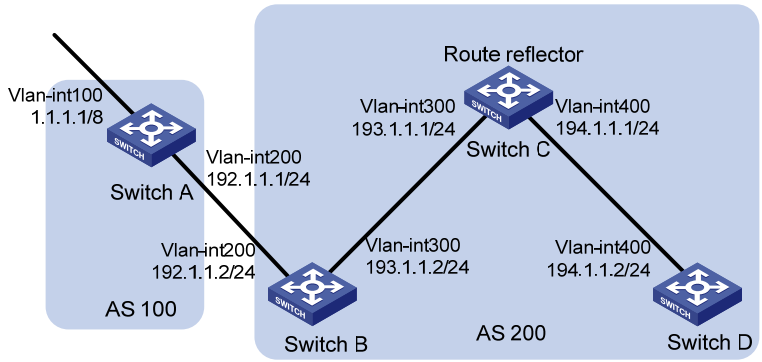
所有交换机运行 BGP 协议，Switch A 与 Switch B 建立 EBGP 连接，Switch C 与 Switch B 和 Switch D 之间建立 IBGP 连接。

Switch C 作为路由反射器，Switch B 和 Switch D 为 Switch C 的客户机。

Switch D 能够通过 Switch C 学到路由 1.0.0.0/8。

2. 组网图

图1-24 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 Switch A。

```
<SwitchA> system-view
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 192.1.1.2 as-number 200
```

通告 1.0.0.0/8 网段路由到 BGP 路由表中。

```
[SwitchA-bgp] network 1.0.0.0
[SwitchA-bgp] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 192.1.1.1 as-number 100
[SwitchB-bgp] peer 193.1.1.1 as-number 200
[SwitchB-bgp] peer 193.1.1.1 next-hop-local
[SwitchB-bgp] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] bgp 200
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 193.1.1.2 as-number 200
[SwitchC-bgp] peer 194.1.1.2 as-number 200
[SwitchC-bgp] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] bgp 200
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] peer 194.1.1.1 as-number 200
[SwitchD-bgp] quit
```

(3) 配置路由反射器

配置 Switch C。

```
[SwitchC] bgp 200
[SwitchC-bgp] peer 193.1.1.2 reflect-client
[SwitchC-bgp] peer 194.1.1.2 reflect-client
[SwitchC-bgp] quit
```

(4) 验证配置效果

查看 Switch B 的 BGP 路由表。

```
[SwitchB] display bgp routing-table

 Total Number of Routes: 1

 BGP Local router ID is 200.1.2.2
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
     Network        NextHop         MED        LocPrf     PrefVal Path/Ogn

 *>  1.0.0.0        192.1.1.1       0                     0       100i
```

查看 Switch D 的 BGP 路由表。

```
[SwitchD] display bgp routing-table

 Total Number of Routes: 1

 BGP Local router ID is 200.1.2.1
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
     Network        NextHop         MED        LocPrf     PrefVal Path/Ogn

   i 1.0.0.0        193.1.1.2       0          100        0       100i
```

可以看出，Switch D 从 Switch C 已经学到了 1.0.0.0/8 路由。

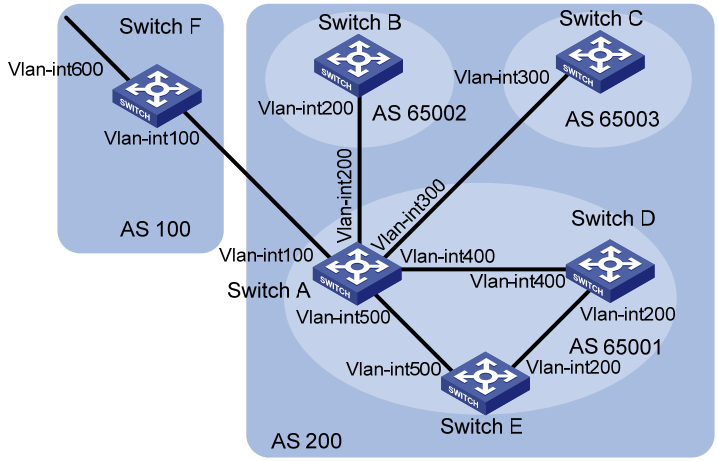
1.14.6 BGP联盟配置

1. 组网需求

AS 200 中有多台 BGP 交换机，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 内的三台交换机建立 IBGP 全连接。

2. 组网图

图1-25 配置联盟组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int100	200.1.1.1/24	Switch D	Vlan-int200	10.1.5.1/24
	Vlan-int200	10.1.1.1/24		Vlan-int400	10.1.3.2/24
	Vlan-int300	10.1.2.1/24	Switch E	Vlan-int200	10.1.5.2/24
	Vlan-int400	10.1.3.1/24		Vlan-int500	10.1.4.2/24
	Vlan-int500	10.1.4.1/24	Switch F	Vlan-int100	200.1.1.2/24
Switch B	Vlan-int200	10.1.1.2/24		Vlan-int600	9.1.1.1/24
Switch C	Vlan-int300	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 Switch A。

```
<SwitchA> system-view
[SwitchA] bgp 65001
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] confederation id 200
[SwitchA-bgp] confederation peer-as 65002 65003
[SwitchA-bgp] peer 10.1.1.2 as-number 65002
[SwitchA-bgp] peer 10.1.1.2 next-hop-local
[SwitchA-bgp] peer 10.1.2.2 as-number 65003
[SwitchA-bgp] peer 10.1.2.2 next-hop-local
[SwitchA-bgp] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] bgp 65002
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] confederation id 200
[SwitchB-bgp] confederation peer-as 65001 65003
[SwitchB-bgp] peer 10.1.1.1 as-number 65001
[SwitchB-bgp] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] bgp 65003
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] confederation id 200
[SwitchC-bgp] confederation peer-as 65001 65002
[SwitchC-bgp] peer 10.1.2.1 as-number 65001
[SwitchC-bgp] quit
```

(3) 配置 AS 65001 内的 IBGP 连接

配置 Switch A。

```
[SwitchA] bgp 65001
[SwitchA-bgp] peer 10.1.3.2 as-number 65001
[SwitchA-bgp] peer 10.1.3.2 next-hop-local
[SwitchA-bgp] peer 10.1.4.2 as-number 65001
[SwitchA-bgp] peer 10.1.4.2 next-hop-local
[SwitchA-bgp] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] bgp 65001
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] confederation id 200
[SwitchD-bgp] peer 10.1.3.1 as-number 65001
[SwitchD-bgp] peer 10.1.5.2 as-number 65001
[SwitchD-bgp] quit
```

配置 Switch E。

```
<SwitchE> system-view
[SwitchE] bgp 65001
[SwitchE-bgp] router-id 5.5.5.5
[SwitchE-bgp] confederation id 200
[SwitchE-bgp] peer 10.1.4.1 as-number 65001
[SwitchE-bgp] peer 10.1.5.1 as-number 65001
[SwitchE-bgp] quit
```

(4) 配置 AS 100 和 AS 200 之间的 EBGP 连接

配置 Switch A。

```
[SwitchA] bgp 65001
[SwitchA-bgp] peer 200.1.1.2 as-number 100
[SwitchA-bgp] quit
```

配置 Switch F。

```
<SwitchF> system-view
[SwitchF] bgp 100
[SwitchF-bgp] router-id 6.6.6.6
[SwitchF-bgp] peer 200.1.1.1 as-number 200
[SwitchF-bgp] network 9.1.1.0 255.255.255.0
[SwitchF-bgp] quit
```

(5) 验证配置结果

查看 Switch B 的 BGP 路由表。

```
[SwitchB] display bgp routing-table
```

```
 Total Number of Routes: 1

 BGP Local router ID is 2.2.2.2
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               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
[SwitchB] display bgp routing-table 9.1.1.0

 BGP local router ID : 2.2.2.2
 Local AS number : 65002
 Paths:   1 available, 1 best

 BGP routing table entry information of 9.1.1.0/24:
 From            : 10.1.1.1 (1.1.1.1)
 Relay Nexthop   : 0.0.0.0
 Original nexthop: 10.1.1.1
 AS-path         : (65001) 100
 Origin          : igp
 Attribute value : MED 0, localpref 100, pref-val 0, pre 255
 State           : valid, external-confed, best,
 Not advertised to any peers yet
```

查看 Switch D 的 BGP 路由表。

```
[SwitchD] display bgp routing-table

 Total Number of Routes: 1

 BGP Local router ID is 4.4.4.4
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               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
[SwitchD] display bgp routing-table 9.1.1.0

 BGP local router ID : 4.4.4.4
 Local AS number : 65001
 Paths:   1 available, 1 best

 BGP routing table entry information of 9.1.1.0/24:
 From            : 10.1.3.1 (1.1.1.1)
 Relay Nexthop   : 0.0.0.0
 Original nexthop: 10.1.3.1
 AS-path         : 100
```

```
Origin          : igp
Attribute value : MED 0, localpref 100, pref-val 0, pre 255
State           : valid, internal, best,
Not advertised to any peers yet
```

通过以上显示信息可以看出：

- Switch F 只需要和 Switch A 建立 EBGP 连接，而不需要和 Switch B、Switch C 建立连接，同样可以通过联盟将路由信息传递给 Switch B 和 Switch C。
- Switch B 和 Switch D 在同一个联盟里，但是属于不同的子自治域，它们都是通过 Switch 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域内，但是又不需要物理上全连接。

1.14.7 BGP路径选择配置

1. 组网需求

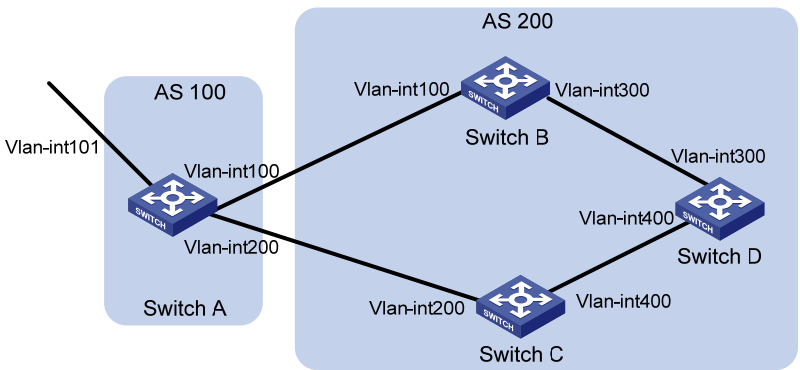
所有交换机都运行 BGP 协议。Switch A 与 Switch B 和 Switch C 之间运行 EBGP；Switch D 与 Switch B 和 Switch C 之间运行 IBGP。

AS 200 中运行 OSPF 协议。

配置不同的路由策略，使得 Switch D 优选 Switch C 学到的 1.0.0.0/8 路由。

2. 组网图

图1-26 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
Switch A	Vlan-int101	1.0.0.0/8	Switch D	Vlan-int400	195.1.1.1/24
	Vlan-int100	192.1.1.1/24		Vlan-int300	194.1.1.1/24
	Vlan-int200	193.1.1.1/24	Switch C	Vlan-int400	195.1.1.2/24
Switch B	Vlan-int100	192.1.1.2/24		Vlan-int200	193.1.1.2/24
	Vlan-int300	194.1.1.2/24			

3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 Switch B、Switch C 和 Switch D 之间运行 OSPF 协议

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ospf
[SwitchB-ospf] area 0
```

```
[SwitchB-ospf-1-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 193.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] ospf
[SwitchD-ospf] area 0
[SwitchD-ospf-1-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.0] quit
[SwitchD-ospf-1] quit
```

(3) 配置 BGP 连接

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] bgp 100
[SwitchA-bgp] peer 192.1.1.2 as-number 200
[SwitchA-bgp] peer 193.1.1.2 as-number 200
```

将 1.0.0.0/8 网段通告到 Switch A 的 BGP 路由表中。

```
[SwitchA-bgp] network 1.0.0.0 8
[SwitchA-bgp] quit
```

配置 Switch B。

```
[SwitchB] bgp 200
[SwitchB-bgp] peer 192.1.1.1 as-number 100
[SwitchB-bgp] peer 194.1.1.1 as-number 200
[SwitchB-bgp] quit
```

配置 Switch C。

```
[SwitchC] bgp 200
[SwitchC-bgp] peer 193.1.1.1 as-number 100
[SwitchC-bgp] peer 195.1.1.1 as-number 200
[SwitchC-bgp] quit
```

配置 Switch D。

```
[SwitchD] bgp 200
[SwitchD-bgp] peer 194.1.1.2 as-number 200
[SwitchD-bgp] peer 195.1.1.2 as-number 200
[SwitchD-bgp] quit
```

(4) 通过配置 1.0.0.0/8 路由的不同属性值，使得 Switch D 优选 Switch C 学到的路由。

- 方法一：在 Switch 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 Switch D 优选 Switch 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。

```
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[SwitchA-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。

```
[SwitchA] route-policy apply_med_50 permit node 10
[SwitchA-route-policy] if-match acl 2000
[SwitchA-route-policy] apply cost 50
[SwitchA-route-policy] quit
[SwitchA] route-policy apply_med_100 permit node 10
[SwitchA-route-policy] if-match acl 2000
[SwitchA-route-policy] apply cost 100
[SwitchA-route-policy] quit
```

对发布给对等体 193.1.1.2（SwitchC）的路由应用名为 apply_med_50 的 Route-policy，对发布给对等体 192.1.1.2（SwitchB）的路由应用名为 apply_med_100 的 Route-policy。

```
[SwitchA] bgp 100
[SwitchA-bgp] peer 193.1.1.2 route-policy apply_med_50 export
[SwitchA-bgp] peer 192.1.1.2 route-policy apply_med_100 export
[SwitchA-bgp] quit
```

查看 Switch D 的 BGP 路由表。

```
[SwitchD] display bgp routing-table

 Total Number of Routes: 2

 BGP Local router ID is 194.1.1.1
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               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，Switch D 从 Switch C 学到 1.0.0.0/8 的路由是最优的。

- 方法二：在 Switch B 和 Switch C 上分别对 1.0.0.0/8 路由配置不同的本地优先级，使得 Switch D 优选 Switch C 学到的路由。

在 Switch C 上定义编号为 2000 的 ACL，允许 1.0.0.0/8 路由通过。

```
[SwitchC] acl number 2000
[SwitchC-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[SwitchC-acl-basic-2000] quit
```

在 Switch 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）。

```
[SwitchC] route-policy localpref permit node 10
```

```
[SwitchC-route-policy] if-match acl 2000
[SwitchC-route-policy] apply local-preference 200
[SwitchC-route-policy] quit
```

为从 BGP 对等体 193.1.1.1 的路由应用名为 localpref 的 Router-policy。

```
[SwitchC] bgp 200
[SwitchC-bgp] peer 193.1.1.1 route-policy localpref import
[SwitchC-bgp] quit
```

查看 Switch D 的 BGP 路由表。

```
[SwitchD] display bgp routing-table

 Total Number of Routes: 2

 BGP Local router ID is 194.1.1.1
 Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
     Network          NextHop       MED        LocPrf     PrefVal Path/Ogn

 *>i 1.0.0.0          193.1.1.1     0          200        0       100i
 * i                  192.1.1.1     0          100        0       100i
```

可以看到，Switch D 从 Switch C 学到 1.0.0.0/8 的路由是最优的。

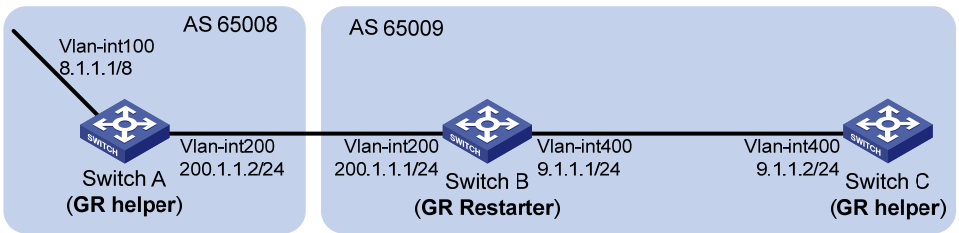
1.14.8 BGP GR配置

1. 组网需求

如 图 1-27 所示，所有路由器均运行BGP协议，Switch A和Switch B之间建立EBGP连接，Switch B和Switch C之间建立IBGP全连接。现要求实现即便Switch B发生主备倒换，也不会影响Switch A和Switch C之间正在进行的数据传输。

2. 组网图

图1-27 BGP GR 配置组网图



3. 配置步骤

(1) Switch A 的配置

配置各接口的 IP 地址（略）。

配置 Switch A 与 Switch B 的 EBGP 连接。

```
<SwitchA> system-view
[SwitchA] bgp 65008
```

```
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 200.1.1.1 as-number 65009
```
将 8.0.0.0/8 网段路由通告到 BGP 路由表中。
```
[SwitchA-bgp] network 8.0.0.0
```
使能 BGP GR 功能。
```
[SwitchA-bgp] graceful-restart
```
(2) Switch B 的配置

配置各接口的 IP 地址（略）。

配置 Switch B 与 Switch A 的 EBGP 连接。
```
<SwitchB> system-view
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 200.1.1.2 as-number 65008
```
配置 Switch B 与 Switch C 的 IBGP 连接。
```
[SwitchB-bgp] peer 9.1.1.2 as-number 65009
```
将 200.1.1.0/24 和 9.1.1.0/24 网段路由通告到 BGP 路由表中。
```
[SwitchB-bgp] network 200.1.1.0 24
[SwitchB-bgp] network 9.1.1.0 24
```
使能 BGP GR 功能。
```
[SwitchB-bgp] graceful-restart
```
(3) Switch C 的配置

配置各接口的 IP 地址（略）。

配置 Switch C 与 Switch B 的 IBGP 连接。
```
<SwitchC> system-view
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 9.1.1.1 as-number 65009
```
使能 BGP GR 功能。
```
[SwitchC-bgp] graceful-restart
```
4. 功能验证

在 Switch A 上 ping Switch C，同时在 Switch B 上触发主备倒换，可以发现在整个倒换过程中 Switch A 都可以 ping 通 Switch C。

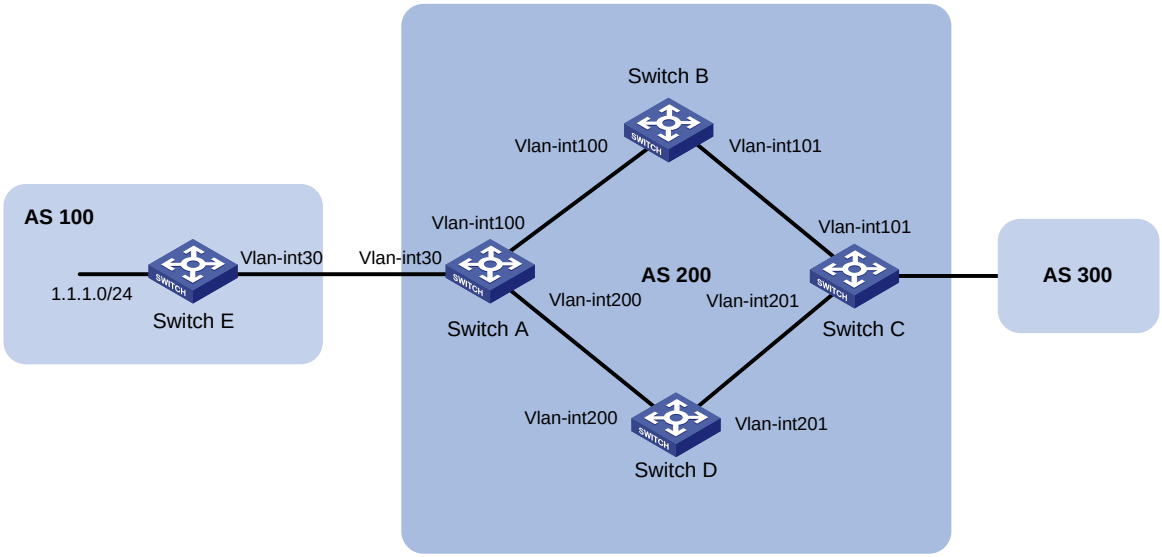
1.14.9 BGP与BFD联动配置

1. 组网需求

- 在 AS 200 内使用 OSPF 作为 IGP 协议；
- Switch A 与 Switch C 建立两条 IBGP 连接。当 Switch A 与 Switch C 之间的两条链路均连通时，Switch C 与 1.1.1.0/24 之间的报文使用 Switch A<－>Switch B<－>Switch C 这条路径；当 Switch A<－>Switch B<－>Switch C 链路故障时，BFD 能够快速检测并通告 BGP 协议，使得 Switch A<－>Switch D<－>Switch C 这条路径能够迅速生效。

2. 组网图

图1-28 配置 BGP 与 BFD 联动组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int100	3.0.1.1/24	Switch C	Vlan-int101	3.0.2.2/24
	Vlan-int200	2.0.1.1/24		Vlan-int201	2.0.2.2/24
	Vlan-int30	30.1.1.1/24	Switch D	Vlan-int200	2.0.1.2/24
Switch B	Vlan-int100	3.0.1.2/24		Vlan-int201	2.0.2.1/24
	Vlan-int101	3.0.2.1/24	Switch E	Vlan-int30	30.1.1.2/24

3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 在 AS 200 内配置 OSPF（略），保证 Switch A 和 Switch C 之间路由可达，并在 Switch A 上引入 BGP 路由（略）

(3) Switch A 上的 BGP 配置

配置 Switch A 和 Switch C 建立两条 IBGP 连接，并配置发布路由时将自身地址作为下一跳。

```
<SwitchA> system-view
[SwitchA] bgp 100
[SwitchA-bgp] peer 3.0.2.2 as-number 200
[SwitchA-bgp] peer 2.0.2.2 as-number 200
[SwitchA-bgp] peer 3.0.2.2 next-hop-local
[SwitchA-bgp] peer 2.0.2.2 next-hop-local
```

引入 OSPF 路由。

```
[SwitchA-bgp] import-route ospf
[SwitchA-bgp] quit
```

配置当 Switch A 与 Switch C 之间的两条链路均连通时，Switch C 与 1.1.1.0/24 之间的报文使用 Switch A<－>Switch B<－>Switch C 这条路径。（在 Switch 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。

```
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule permit source 1.1.1.0 24
```

```
[SwitchA-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。

```
[SwitchA] route-policy apply_med_50 permit node 10
[SwitchA-route-policy] if-match acl 2000
[SwitchA-route-policy] apply cost 50
[SwitchA-route-policy] quit
[SwitchA] route-policy apply_med_100 permit node 10
[SwitchA-route-policy] if-match acl 2000
[SwitchA-route-policy] apply cost 100
[SwitchA-route-policy] 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。

```
[SwitchA] bgp 100
[SwitchA-bgp] peer 3.0.2.2 route-policy apply_med_50 export
[SwitchA-bgp] peer 2.0.2.2 route-policy apply_med_100 export
```

配置当 Switch A<－>Switch B<－>Switch C 链路故障时，BFD 能够快速检测并通告 BGP 协议，使得 Switch A<－>Switch D<－>Switch C 这条路径能够迅速生效。

```
[SwitchA-bgp] peer 3.0.2.2 bfd
```

配置 Switch A 和 Switch E 建立 EBGP 连接。

```
[SwitchA-bgp] peer 30.1.1.2 as-number 100
[SwitchA-bgp] quit
```

(4) Switch C 上的 BGP 配置。

配置 Switch A 和 Switch C 建立两条 IBGP 连接。

```
<SwitchC> system-view
[SwitchC] bgp 100
[SwitchC-bgp] peer 3.0.1.1 as-number 200
[SwitchC-bgp] peer 2.0.1.1 as-number 200
```

配置当 Switch A<－>Switch B<－>Switch C 链路故障时，BFD 能够快速检测并通告 BGP 协议，使得 Switch A<－>Switch D<－>Switch C 这条路径能够迅速生效。

```
[SwitchC-bgp] peer 3.0.1.1 bfd
```

配置 Switch C 学习到的 IBGP 路由的优先级为 140。

```
[SwitchC-bgp] preference 255 140 130
[SwitchC-bgp] quit
```

(5) Switch E 上的 BGP 配置

配置 Switch A 和 Switch E 建立 EBGP 连接。

```
<SwitchE> system-view
[SwitchE] bgp 100
[SwitchE-bgp] peer 30.1.1.1 as-number 200
```

配置 Switch E 通过 BGP 发布网段路由 1.1.1.0/24。

```
[SwitchE-bgp] network 1.1.1.0 24
[SwitchE-bgp] quit
```

(6) 配置各路由器接口及 BFD 参数（该步骤可省略，使用 BFD 参数的缺省值即可）

配置 Switch A。

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 100
```

- 配置发送 BFD 控制报文的最小时间间隔为 500 毫秒。

```
[SwitchA-Vlan-interface100] bfd min-transmit-interval 500
```

- 配置接收 BFD 控制报文的最小时间间隔为 500 毫秒。

```
[SwitchA-Vlan-interface100] bfd min-receive-interval 500
```

- 配置 BFD 报文最大失效的个数为 7。

```
[SwitchA-Vlan-interface100] bfd detect-multiplier 7
```

- 配置接口建立 BFD 会话时需要认证，采用明文认证方式方式，密码为 ibgpbfd。

```
[SwitchA-Vlan-interface100] bfd authentication-mode simple 1 ibgpbfd
[SwitchA-Vlan-interface100] quit
```

配置 Switch C。

```
[SwitchC] bfd session init-mode active
[SwitchC] interface vlan-interface 101
[SwitchC-Vlan-interface101] bfd min-transmit-interval 500
[SwitchC-Vlan-interface101] bfd min-receive-interval 500
[SwitchC-Vlan-interface101] bfd detect-multiplier 7
[SwitchC-Vlan-interface101] bfd authentication-mode simple 1 ibgpbfd
[SwitchC-Vlan-interface101] return
```

(7) 结果验证

下面以 Switch C 为例，Switch A 和 Switch C 类似，不再赘述。

显示 Switch C 的 BFD 信息。

```
<SwitchC> display bfd session verbose

 Total Session Num: 1            Init Mode: Active

 IP Session Working Under Ctrl Mode:

     Local Discr: 17                 Remote Discr: 13
      Source IP: 3.0.2.2           Destination IP:  3.0.1.1
   Session State: Up                     Interface: Vlan-interface101
 Min Trans Inter: 500ms           Act Trans Inter: 500ms
  Min Recv Inter: 500ms          Act Detect Inter: 3000ms
  Running Up for: 00:00:06              Auth mode: None
    Connect Type: Indirect              Board Num: 0
        Protocol: BGP
       Diag Info: No Diagnostic
```

以上显示信息表明：Switch A 的 VLAN 接口 100 和 Switch C 的 VLAN 接口 101 之间已经建立了 BFD 连接，而且 BFD 协议运行正常。

在 Switch C 上查看 BGP 邻居信息，可以看出 Switch A 和 Switch C 之间 BGP 邻居处于 Established 状态。

```
<SwitchC> display bgp peer

 BGP local router ID : 1.1.1.1
 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        7       10    0       0 00:01:05 Established
  3.0.1.1                200        7       10    0       0 00:01:34 Established
```

在 Switch C 上查看 1.1.1.0/24 的路由信息，可以看出 Switch A 和 Switch C 是通过 Switch B 进行通信的。

```
<SwitchC> display ip routing-table 1.1.1.0 24 verbose
Routing Table : Public
Summary Count : 4

  Destination: 1.1.1.0/24
     Protocol: BGP             Process ID: 0
   Preference: 0                     Cost: 50
      NextHop: 3.0.1.1          Interface: Vlan-interface101
    BkNextHop: 2.0.2.1        BkInterface: Vlan-interface201
  RelyNextHop: 3.0.2.1          Neighbor : 3.0.1.1
    Tunnel ID: 0x0                  Label: NULL
        State: Active Adv GotQ            Age: 00h08m54s
          Tag: 0

  Destination: 1.1.1.0/24
     Protocol: BGP             Process ID: 0
   Preference: 0                     Cost: 100
      NextHop: 2.0.1.1          Interface: Vlan-interface201
    BkNextHop: 0.0.0.0        BkInterface:
  RelyNextHop: 2.0.2.1          Neighbor : 2.0.1.1
    Tunnel ID: 0x0                  Label: NULL
        State: Invalid Adv GotQ            Age: 00h08m54s
          Tag: 0

  Destination: 1.1.1.0/24
     Protocol: O_ASE           Process ID: 1
   Preference: 150                   Cost: 1
IpPrecedence:                    QosLcId:
      NextHop: 3.0.2.1          Interface: Vlan-interface101
    BkNextHop: 0.0.0.0        BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
    Tunnel ID: 0x0                  Label: NULL
  BKTunnel ID: 0x0                BKLabel: NULL
        State: Inactive Adv           Age: 00h14m10s
          Tag: 1

  Destination: 1.1.1.0/24
     Protocol: O_ASE           Process ID: 1
   Preference: 150                   Cost: 1
IpPrecedence:                    QosLcId:
```

```
      NextHop: 2.0.2.1           Interface: Vlan-interface201
    BkNextHop: 0.0.0.0         BkInterface:
  RelyNextHop: 0.0.0.0           Neighbor : 0.0.0.0
    Tunnel ID: 0x0                   Label: NULL
  BKTunnel ID: 0x0                 BKLabel: NULL
        State: Inactive Adv            Age: 00h14m10s
          Tag: 1
```

以上显示信息表明：Switch C 到达 1.1.1.0/24 网段有两条路径：当前生效的是 Switch A<－>Switch B<－>Switch C，Switch A<－>Switch D<－>Switch C 这条路径没有生效。

当 Switch A 和 Switch B 之间的链路发生故障时 Switch C 能够快速检测到链路的变化。

在 Switch C 上查看 1.1.1.0/24 的路由信息，可以看出 Switch A 和 Switch C 是通过 Switch D 进行通信的。

```
<SwitchC> display ip routing-table 1.1.1.0 24 verbose
Routing Table : Public
Summary Count : 1

  Destination: 1.1.1.0/24
     Protocol: BGP             Process ID: 0
   Preference: 0                     Cost: 100
      NextHop: 2.0.1.1           Interface: Vlan-interface201
    BkNextHop: 0.0.0.0         BkInterface:
  RelyNextHop: 2.0.2.1           Neighbor : 2.0.1.1
    Tunnel ID: 0x0                   Label: NULL
        State: Active Adv              Age: 00h09m54s
          Tag: 0
```

以上显示信息表明：Switch C 到达 1.1.1.0/24 网段有一条路径：Switch A<－>Switch D<－>Switch C。

1.15 BGP常见错误配置举例

1.15.1 BGP对等体关系不能建立

1. 故障现象

使用 **display bgp peer** 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。

2. 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。

3. 处理过程

(1) 用 **display current-configuration** 命令检查邻居的 AS 号配置是否正确。

(2) 用 **display bgp peer** 命令检查邻居的 IP 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) 检查路由表中是否存在到邻居的可用路由。

(6) 使用 **ping** 命令检查链路是否畅通。

(7) 使用 **display tcp status** 命令检查 TCP 连接是否正常。

(8) 检查是否配置了禁止 TCP 端口 179 的 ACL。

目 录

1 IPv6 静态路由配置 ····· 1-1
1.1 IPv6 静态路由简介 ····· 1-1
1.1.1 IPv6 静态路由属性及功能 ····· 1-1
1.1.2 IPv6 缺省路由 ····· 1-1
1.2 配置IPv6 静态路由 ····· 1-1
1.2.1 配置准备 ····· 1-1
1.2.2 配置IPv6 静态路由 ····· 1-2
1.3 IPv6 静态路由显示和维护 ····· 1-2
1.4 IPv6 静态路由典型配置举例 ····· 1-3

1 IPv6 静态路由配置

说明

本文所指的路由器代表运行了路由协议的三层设备。

1.1 IPv6静态路由简介

静态路由是一种特殊的路由，它由管理员手工配置。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。恰当地设置和使用静态路由可以改进网络的性能，并可为重要的应用保证带宽。

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。

1.1.1 IPv6 静态路由属性及功能

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。

它们之间的主要区别是目的地址和下一跳地址有所不同，IPv6 静态路由使用的是 IPv6 地址，而 IPv4 静态路由使用 IPv4 地址。

1.1.2 IPv6 缺省路由

在配置 IPv6 静态路由时，如果指定的目的地址为::/0（前缀长度为 0），则表示配置了一条 IPv6 缺省路由。如果报文的目的地址无法匹配路由表中的任何一项，路由器将选择 IPv6 缺省路由来转发 IPv6 报文。

1.2 配置IPv6静态路由

在小型 IPv6 网络中，可以通过配置 IPv6 静态路由达到网络互连的目的。相对使用动态路由来说，可以节省带宽。

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
- 使能 IPv6 报文转发能力
- 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* } [**preference** *preference-value*]	二者必选其一 缺省情况下，IPv6静态路由的优先级为60
	ipv6 route-static vpn-instance *s-vpn-instance-name*&<1-6> *ipv6-address prefix-length* { *interface-type interface-number* [*next-hop-address*] | *nexthop-address* [**public**] | **vpn-instance** *d-vpn-instance-name nexthop-address* } [**preference** *preference-value*]	

说明

在配置静态路由时，如果出接口类型为广播（如以太网接口、VLAN 接口等）类型，必须指定下一跳地址。

1.3 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。

在系统视图下执行 **delete** 命令可以删除所有 IPv6 静态路由。

表1-2 IPv6 静态路由显示和维护

操作	命令
查看IPv6静态路由表信息	**display ipv6 routing-table protocol static** [**inactive** | **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
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。
- 命令 **display ipv6 routing-table protocol static** [**inactive** \| **verbose**] [**\|** { **begin** \| **exclude** \| **include** } *regular-expression*]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”。

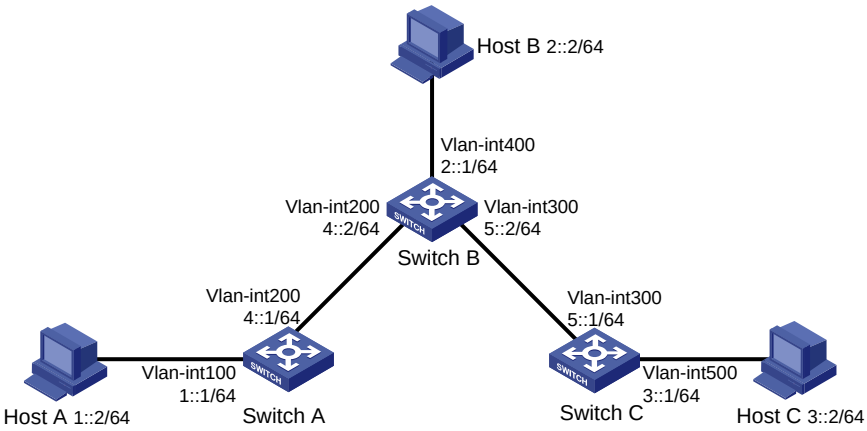
1.4 IPv6静态路由典型配置举例

1. 组网要求

要求各交换机之间配置 IPv6 静态路由协议后，可以使所有主机和交换机之间互通。

2. 组网图

图1-1 IPv6 静态路由配置举例组网图



3. 配置步骤

(1) 配置各 VLAN 虚接口的 IPv6 地址（略）

(2) 配置 IPv6 静态路由

在 Switch A 上使能 IPv6 报文转发功能，并配置 IPv6 缺省路由。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ipv6 route-static :: 0 4::2
```

在 Switch B 上使能 IPv6 报文转发功能，并配置两条 IPv6 静态路由。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ipv6 route-static 1:: 64 4::1
[SwitchB] ipv6 route-static 3:: 64 5::1
```

在 Switch C 上使能 IPv6 报文转发功能，并配置 IPv6 缺省路由。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] 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) 查看配置结果

查看 Switch A 的 IPv6 路由表。

```
[SwitchA] display ipv6 routing-table
Routing Table :
```

```
        Destinations : 5        Routes : 5

 Destination  : ::                                  Protocol     : Static
 NextHop      : 4::2                                 Preference   : 60
 Interface    : Vlan200                   Cost         : 0

 Destination  : ::1/128                              Protocol     : Direct
 NextHop      : ::1                                  Preference   : 0
 Interface    : InLoop0                              Cost         : 0

 Destination  : 1::/64                               Protocol     : Direct
 NextHop      : 1::1                                 Preference   : 0
 Interface    : Vlan100                   Cost         : 0

 Destination  : 1::1/128                             Protocol     : Direct
 NextHop      : ::1                                  Preference   : 0
 Interface    : InLoop0                              Cost         : 0

 Destination  : FE80::/10                            Protocol     : Direct
 NextHop      : ::                                   Preference   : 0
 Interface    : NULL0                                Cost         : 0
```

使用 Ping 进行验证。

```
[SwitchA] ping ipv6 3::1
  PING 3::1 : 56  data bytes, press CTRL_C to break
    Reply from 3::1
    bytes=56 Sequence=1 hop limit=254  time = 63 ms
    Reply from 3::1
    bytes=56 Sequence=2 hop limit=254  time = 62 ms
    Reply from 3::1
    bytes=56 Sequence=3 hop limit=254  time = 62 ms
    Reply from 3::1
    bytes=56 Sequence=4 hop limit=254  time = 63 ms
    Reply from 3::1
    bytes=56 Sequence=5 hop limit=254  time = 63 ms

  --- 3::1 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 62/62/63 ms
```

目 录

1 RIPng配置 ····· 1-1

1.1 简介 ····· 1-1

1.1.1 RIPng工作机制 ····· 1-1

1.1.2 RIPng的报文格式 ····· 1-2

1.1.3 RIPng报文处理过程 ····· 1-3

1.1.4 协议规范 ····· 1-3

1.2 RIPng配置任务简介 ····· 1-4

1.3 配置RIPng的基本功能 ····· 1-4

1.3.1 配置准备 ····· 1-4

1.3.2 配置RIPng的基本功能 ····· 1-4

1.4 配置RIPng路由特性 ····· 1-5

1.4.1 配置接口附加度量值 ····· 1-5

1.4.2 配置RIPng路由聚合 ····· 1-6

1.4.3 配置RIPng发布缺省路由 ····· 1-6

1.4.4 配置RIPng对接收/发布的路由进行过滤 ····· 1-6

1.4.5 配置RIPng协议优先级 ····· 1-7

1.4.6 配置RIPng引入外部路由 ····· 1-7

1.5 调整和优化RIPng网络 ····· 1-7

1.5.1 配置RIPng定时器 ····· 1-8

1.5.2 配置水平分割和毒性逆转 ····· 1-8

1.5.3 配置RIPng报文的零域检查 ····· 1-9

1.5.4 配置最大等价路由条数 ····· 1-9

1.6 配置RIPng IPsec安全策略 ····· 1-10

1.7 RIPng显示和维护 ····· 1-11

1.8 RIPng典型配置举例 ····· 1-11

1.8.1 配置RIPng基本功能 ····· 1-11

1.8.2 配置RIPng引入外部路由 ····· 1-14

1.8.3 配置RIPng IPsec安全策略 ····· 1-16

1 RIPng配置

说明

本文所指的路由器代表运行了路由协议的三层设备。

1.1 简介

RIPng 又称为下一代 RIP 协议（RIP next generation），它是对原来的 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”。

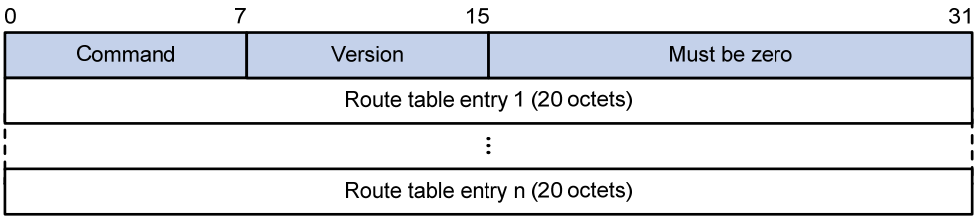
1.1.2 RIPng的报文格式

1. 基本格式

RIPng 报文由头部（Header）和多个路由表项（RTEs）组成。在同一个 RIPng 报文中，RTE 的最大条数与发送接口设置的 IPv6 MTU 有关。

RIPng 报文基本格式如下图所示。

图1-1 RIPng 报文基本格式



各字段的含义如下：

- Command：定义报文的类型。0x01 表示 Request 报文，0x02 表示 Response 报文。
- Version：RIPng 的版本，目前其值只能为 0x01。
- RTE（Route Table Entry）：路由表项，每项的长度为 20 字节。

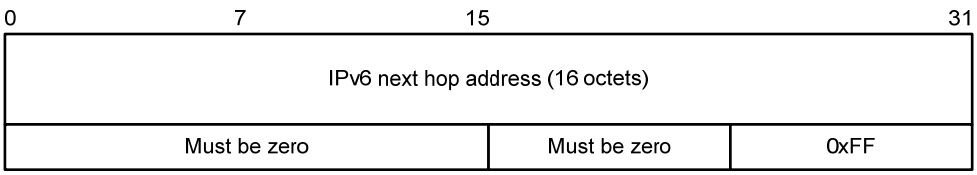
2. RTE的格式

在 RIPng 里有两类 RTE，分别是：

- 下一跳 RTE：位于一组具有相同下一跳的“IPv6 前缀 RTE”的前面，它定义了下一跳的 IPv6 地址。
- IPv6 前缀 RTE：位于某个“下一跳 RTE”的后面。同一个“下一跳 RTE”的后面可以有多个不同的“IPv6 前缀 RTE”。它描述了 RIPng 路由表中的目的 IPv6 地址、路由标记、前缀长度以及度量值。

下一跳 RTE 的格式如下图所示。

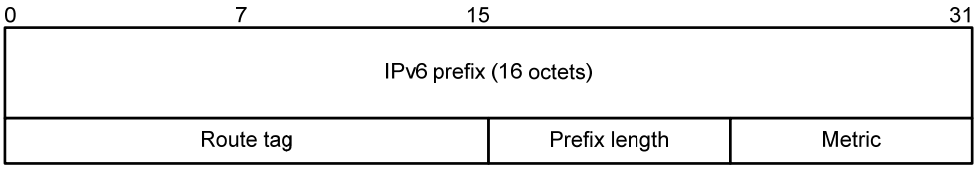
图1-2 下一跳 RTE 格式



其中，IPv6 next hop address 表示下一跳的 IPv6 地址。

IPv6 前缀 RTE 的格式如下图所示。

图1-3 IPv6 前缀 RTE 格式



各字段的解释如下：

- IPv6 prefix：目的 IPv6 地址的前缀。
- route tag：路由标记。
- prefix len：IPv6 地址的前缀长度。
- metric：路由的度量值。

1.1.3 RIPng报文处理过程

1. Request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。

收到Request报文的RIPng路由器会对其中的RTE进行处理。如果Request报文中只有一项RTE，且 IPv6 前缀和前缀长度都为 0，度量值为 16，则表示请求邻居发送全部路由信息，被请求路由器收到后会把当前路由表中的全部路由信息，以 Response 报文形式发回给请求路由器。如果 Request 报文中有多项 RTE，被请求路由器将对 RTE 逐项处理，更新每条路由的度量值，最后以 Response 报文形式返回给请求路由器。

2. Response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 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。

1.1.4 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
使能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 IPsec安全策略		可选	1.6

1.3 配置RIPng的基本功能

RIPng 基本功能的配置任务包括 RIPng 特性配置中涉及的基本配置，完成本节的配置就可以使用 RIPng 特性。

在 RIPng 的配置中，应该最先使能 RIPng。但是与 RIPng 功能特性相关的接口配置不受是否启动了 RIPng 的限制，比如给接口配置一个 IPv6 地址。

1.3.1 配置准备

在配置 RIPng 基本功能之前，需完成以下任务：

- 启动 IPv6 报文转发功能
- 配置接口的网络层地址，使相邻节点的网络层可达

1.3.2 配置RIPng的基本功能

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*]	必选 缺省情况下，没有RIPng进程在运行

操作	命令	说明
退回系统视图	**quit**	-
进入接口视图	**interface** *interface-type interface-number*	-
在指定的网络接口上使能RIPng	**ripng** *process-id* **enable**	必选 缺省情况下，接口禁用RIPng

说明

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
- 配置 RIPng 基本功能
- 如果使用 IPv6 ACL 进行路由过滤，需要定义 IPv6 ACL，相关配置请参见“ACL 和 QoS 配置指导”中的“ACL”。
- 如果使用 IPv6 地址前缀列表进行路由过滤，需要定义 IPv6 地址前缀列表，相关配置请参见“三层技术-IP 路由配置指导”中的“路由策略”。

1.4.1 配置接口附加度量值

附加度量值是附加在 RIPng 路由上的输入输出度量值，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*	可选 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*	可选 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

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前缀	**ripng** *ipv6-address prefix-length* **summary-address**	必选

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*]	必选 缺省情况下，RIPng进程不发布缺省路由

说明

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* | **ipv6-prefix** *ipv6-prefix-name* } **import**	必选 缺省情况下，RIPng不对接收的路由信息进行过滤

操作	命令	说明
对发布的路由信息进行过滤	**filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* } **export** [*protocol* [*process-id*]]	必选 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*] *preference*	可选 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*]	-
配置引入路由的缺省度量值	**default cost** *cost*	可选 缺省情况下，引入路由的缺省度量值为0
引入外部路由	**import-route** *protocol* [*process-id*] [**allow-ibgp**] [**cost** *cost* | **route-policy** *route-policy-name*] *	必选 缺省情况下，RIPng不引入其它路由

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
- 配置 RIPng 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

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* } *	可选 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

说明

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。

1.5.2 配置水平分割和毒性逆转

说明

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

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**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能水平分割功能	**ripng split-horizon**	可选 缺省情况下，水平分割功能处于使能状态

说明

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。

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**	必选 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 报文的零域检查

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*]	-
使能对RIPng报文头部的零域检查操作	**checkzero**	可选 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**	-
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*	可选 缺省为8

1.6 配置RIPng IPsec安全策略

在安全性要求较高的网络环境中，可以通过配置基于 IPsec 安全策略的认证方式来对 RIPng 报文进行有效性检查和验证。

路由器在发送的报文中会携带配置好的 IPsec 安全策略的 SPI（Security Parameter Index，安全参数索引）值，接收报文时通过 SPI 值进行 IPsec 安全策略匹配：只有策略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 安全策略。进程下配置的 IPsec 安全策略对该进程下的所有报文有效，接口下的 IPsec 安全策略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 安全策略时，接口下的策略生效。

1. 配置准备

配置 RIPng IPsec 安全策略之前，需要满足下面配置：

- 安全提议已经创建。
- IPsec 安全策略已经创建。

关于安全提议和 IPsec 安全策略的详细配置过程，请参见“安全配置指导”中的“IPsec”。

2. 配置RIPng IPsec安全策略

表1-14 配置 RIPng IPsec 安全策略（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安全策略	**enable ipsec-policy** *policy-name*	必选 缺省情况下，RIPng进程没有应用IPsec安全策略

表1-15 配置 RIPng IPsec 安全策略（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安全策略	**ripng ipsec-policy** *policy-name*	必选 缺省情况下，使能了RIPng的接口上没有应用IPsec安全策略

说明

RIPng IPsec 安全策略只支持手工配置方式，具体情况请参见“安全配置指导”中的“IPsec”。

1.7 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1-16 RIPng 显示和维护

操作	命令
显示RIPng进程的配置信息	**display ripng** [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示RIPng发布数据库中的路由	**display ripng** *process-id* **database** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定RIPng进程的路由信息	**display ripng** *process-id* **route** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示RIPng的接口信息	**display ripng** *process-id* **interface** [*interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
重启RIPng进程	**reset ripng** *process-id* **process**
清除RIPng进程的统计信息	**reset ripng** *process-id* **statistics**

1.8 RIPng典型配置举例

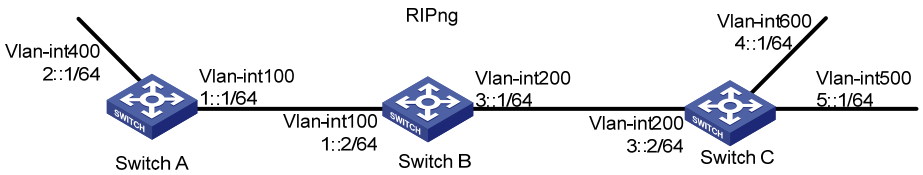
1.8.1 配置RIPng基本功能

1. 组网需求

- Switch A、Switch B 和 Switch C 相连并通过 RIPng 来学习网络中的 IPv6 路由信息；
- 在 Switch B 上对接收的 Switch C 的路由（3::/64）进行过滤，使其不加入到 Switch B 的 RIPng 进程的路由表中，也不发布给 Switch A。

2. 组网图

图1-4 RIPng 配置举例组网图



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 Switch A。

```
<SwitchA> system-view
```

```
[SwitchA] ripng 1
[SwitchA-ripng-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ripng 1 enable
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 400
[SwitchA-Vlan-interface400] ripng 1 enable
[SwitchA-Vlan-interface400] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ripng 1
[SwitchB-ripng-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ripng 1 enable
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ripng 1 enable
[SwitchB-Vlan-interface100] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ripng 1
[SwitchC-ripng-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] ripng 1 enable
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 500
[SwitchC-Vlan-interface500] ripng 1 enable
[SwitchC-Vlan-interface500] quit
[SwitchC] interface vlan-interface 600
[SwitchC-Vlan-interface600] ripng 1 enable
[SwitchC-Vlan-interface600] quit
```

查看 Switch B 的 RIPng 路由表。

```
[SwitchB] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
 ----------------------------------------------------------------

 Peer FE80::20F:E2FF:FE23:82F5  on Vlan-interface100
 Dest 1::/64,
     via FE80::20F:E2FF:FE23:82F5, cost  1, tag 0, A, 6 Sec
 Dest 2::/64,
     via FE80::20F:E2FF:FE23:82F5, cost  1, tag 0, A, 6 Sec

 Peer FE80::20F:E2FF:FE00:100  on Vlan-interface200
 Dest 3::/64,
     via FE80::20F:E2FF:FE00:100, cost  1, tag 0, A, 11 Sec
 Dest 4::/64,
     via FE80::20F:E2FF:FE00:100, cost  1, tag 0, A, 11 Sec
 Dest 5::/64,
```

```
    via FE80::20F:E2FF:FE00:100, cost  1, tag 0, A, 11 Sec
```

查看 Switch A 的 RIPng 路由表。

```
[SwitchA] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
 ----------------------------------------------------------------

 Peer FE80::200:2FF:FE64:8904  on Vlan-interface100
 Dest 1::/64,
     via FE80::200:2FF:FE64:8904, cost  1, tag 0, A, 31 Sec
 Dest 4::/64,
     via FE80::200:2FF:FE64:8904, cost  2, tag 0, A, 31 Sec
 Dest 5::/64,
     via FE80::200:2FF:FE64:8904, cost  2, tag 0, A, 31 Sec
 Dest 3::/64,
     via FE80::200:2FF:FE64:8904, cost  1, tag 0, A, 31 Sec
```

(3) 配置 Switch B 对接收和发布的路由进行过滤

```
[SwitchB] acl ipv6 number 2000
[SwitchB-acl6-basic-2000] rule deny source 3::/64
[SwitchB-acl6-basic-2000] rule permit
[SwitchB-acl6-basic-2000] quit
[SwitchB] ripng 1
[SwitchB-ripng-1] filter-policy 2000 import
[SwitchB-ripng-1] filter-policy 2000 export
```

查看 Switch B 和 Switch A 的 RIPng 路由表。

```
[SwitchB] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
 ----------------------------------------------------------------

 Peer FE80::20F:E2FF:FE23:82F5  on Vlan-interface100
 Dest 1::/64,
     via FE80::20F:E2FF:FE23:82F5, cost  1, tag 0, A, 2 Sec
 Dest 2::/64,
     via FE80::20F:E2FF:FE23:82F5, cost  1, tag 0, A, 2 Sec

 Peer FE80::20F:E2FF:FE00:100  on Vlan-interface200
 Dest 4::/64,
     via FE80::20F:E2FF:FE00:100, cost  1, tag 0, A, 5 Sec
 Dest 5::/64,
     via FE80::20F:E2FF:FE00:100, cost  1, tag 0, A, 5 Sec
[SwitchA] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
 ----------------------------------------------------------------

 Peer FE80::20F:E2FF:FE00:1235  on Vlan-interface100
 Dest 1::/64,
     via FE80::20F:E2FF:FE00:1235, cost  1, tag 0, A, 2 Sec
 Dest 4::/64,
     via FE80::20F:E2FF:FE00:1235, cost  2, tag 0, A, 2 Sec
```

```
Dest 5::/64,
    via FE80::20F:E2FF:FE00:1235, cost  2, tag 0, A, 2 Sec
```

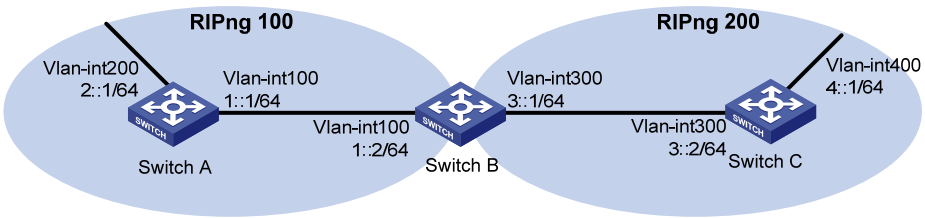
1.8.2 配置RIPng引入外部路由

1. 组网需求

- Switch B 上运行两个 RIPng 进程：RIPng100 和 RIPng200。Switch B 通过 RIPng100 和 Switch A 交换路由信息，通过 RIPng200 和 Switch C 交换路由信息。
- 要求在 Switch 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，将引入的 RIPng200 的路由缺省度量值设为 3。

2. 组网图

图1-5 配置 RIPng 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 RIPng

在 Switch A 上启动 RIPng 进程 100。

```
<SwitchA> system-view
[SwitchA] ripng 100
[SwitchA-ripng-100] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ripng 100 enable
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ripng 100 enable
[SwitchA-Vlan-interface200] quit
```

在 Switch B 上启动两个 RIPng 进程，进程号分别为 100 和 200。

```
<SwitchB> system-view
[SwitchB] ripng 100
[SwitchB-ripng-100] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ripng 100 enable
[SwitchB-Vlan-interface100] quit
[SwitchB] ripng 200
[SwitchB-ripng-200] quit
[SwitchB] interface vlan-interface 300
[SwitchB-Vlan-interface300] ripng 200 enable
[SwitchB-Vlan-interface300] quit
```

在 Switch C 上启动 RIPng 进程 200。

```
<SwitchC> system-view
[SwitchC] ripng 200
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] ripng 200 enable
[SwitchC-Vlan-interface300] quit
[SwitchC] interface vlan-interface 400
[SwitchC-Vlan-interface400] ripng 200 enable
[SwitchC-Vlan-interface400] quit
```

查看 Switch A 的路由表信息。

```
[SwitchA] display ipv6 routing-table
Routing Table :
         Destinations : 6        Routes : 6

Destination: ::1/128                                    Protocol  : Direct
NextHop    : ::1                                        Preference: 0
Interface  : InLoop0                                    Cost      : 0

Destination: 1::/64                                     Protocol  : Direct
NextHop    : 1::1                                       Preference: 0
Interface  : Vlan100                                    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  : Vlan200                                    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
```

(3) 配置 RIPng 引入外部路由

在 Switch 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。

```
[SwitchB] ripng 100
[SwitchB-ripng-100] default cost 3
[SwitchB-ripng-100] import-route ripng 200
[SwitchB-ripng-100] quit
[SwitchB] ripng 200
[SwitchB-ripng-200] import-route ripng 100
[SwitchB-ripng-200] quit
```

查看路由引入后 Switch A 的路由表信息。

```
[SwitchA] display ipv6 routing-table
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  : Vlan100                                    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  : Vlan200                                    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  : Vlan100                                    Cost      : 4

Destination: FE80::/10                                  Protocol  : Direct
NextHop    : ::                                         Preference: 0
Interface  : NULL0                                      Cost      : 0d
```

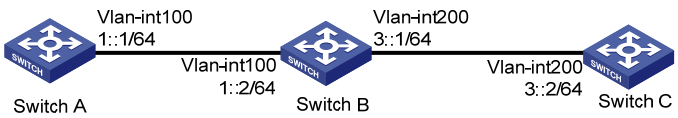
1.8.3 配置RIPng IPsec安全策略

1. 组网需求

- Switch A、Switch B 和 Switch C 相连并通过 RIPng 来学习网络中的 IPv6 路由信息。
- 要求配置 IPsec 安全策略对 Switch A、Switch B 和 Switch C 之间的 RIPng 报文进行有效性检查和验证。

2. 组网图

图1-6 RIPng IPsec 安全策略组网图



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 Switch A。

```
<SwitchA> system-view
[SwitchA] ripng 1
[SwitchA-ripng-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ripng 1 enable
[SwitchA-Vlan-interface100] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ripng 1
[SwitchB-ripng-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ripng 1 enable
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ripng 1 enable
[SwitchB-Vlan-interface100] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ripng 1
[SwitchC-ripng-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] ripng 1 enable
[SwitchC-Vlan-interface200] quit
```

(3) 配置 RIPng IPsec 安全策略

配置 Switch 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。创建一条安全策略 policy001，协商方式为 manual，配置 SPI 和密钥。

```
[SwitchA] ipsec proposal tran1
[SwitchA-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchA-ipsec-proposal-tran1] transform esp
[SwitchA-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchA-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchA-ipsec-proposal-tran1] quit
[SwitchA] ipsec policy policy001 10 manual
[SwitchA-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchA-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] quit
```

配置 Switch 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。创建一条安全策略 policy001，协商方式为 manual，配置 SPI 和密钥。

```
[SwitchB] ipsec proposal tran1
```

```
[SwitchB-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran1] transform esp
[SwitchB-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran1] quit
[SwitchB] ipsec policy policy001 10 manual
[SwitchB-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchB-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] quit
```

配置 Switch C。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。创建一条安全策略 policy001，协商方式为 manual，配置 SPI 和密钥。

```
[SwitchC] ipsec proposal tran1
[SwitchC-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchC-ipsec-proposal-tran1] transform esp
[SwitchC-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchC-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchC-ipsec-proposal-tran1] quit
[SwitchC] ipsec policy policy001 10 manual
[SwitchC-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchC-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchC-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchC-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchC-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchC-ipsec-policy-manual-policy001-10] quit
```

(4) RIPng 进程上应用 IPsec 安全策略

配置 Switch A。

```
[SwitchA] ripng 1
[SwitchA-ripng-1] enable ipsec-policy policy001
[SwitchA-ripng-1] quit
```

配置 Switch B。

```
[SwitchB] ripng 1
[SwitchB-ripng-1] enable ipsec-policy policy001
[SwitchB-ripng-1] quit
```

配置 Switch C。

```
[SwitchC] ripng 1
[SwitchC-ripng-1] enable ipsec-policy policy001
[SwitchC-ripng-1] quit
```

4. 检验配置结果

以上配置完成后，Switch A、Switch B 和 Switch C 之间的 RIPng 报文将被加密传输。

目 录

1 OSPFv3 配置 ····· 1-1
1.1 OSPFv3 简介 ····· 1-1
1.1.1 OSPFv3 概述 ····· 1-1
1.1.2 OSPFv3 的协议报文 ····· 1-1
1.1.3 OSPFv3 的LSA类型 ····· 1-2
1.1.4 OSPFv3 的定时器 ····· 1-3
1.1.5 系统支持的OSPFv3 特性 ····· 1-3
1.1.6 协议规范 ····· 1-4
1.2 OSPFv3 配置任务简介 ····· 1-4
1.3 使能OSPFv3 功能 ····· 1-5
1.3.1 配置准备 ····· 1-5
1.3.2 使能OSPFv3 功能 ····· 1-5
1.4 配置OSPFv3 的区域属性 ····· 1-5
1.4.1 配置准备 ····· 1-6
1.4.2 配置OSPFv3 的Stub区域 ····· 1-6
1.4.3 配置OSPFv3 的虚连接 ····· 1-7
1.5 配置OSPFv3 网络类型 ····· 1-7
1.5.1 配置准备 ····· 1-7
1.5.2 配置OSPFv3 接口的网络类型 ····· 1-8
1.5.3 配置NBMA或者P2MP网络的邻居 ····· 1-8
1.6 配置OSPFv3 的路由信息控制 ····· 1-8
1.6.1 配置准备 ····· 1-8
1.6.2 配置OSPFv3 的路由聚合 ····· 1-9
1.6.3 配置OSPFv3 对通过接收到的LSA计算出来的路由信息进行过滤 ····· 1-9
1.6.4 配置OSPFv3 接口的开销值 ····· 1-10
1.6.5 配置OSPFv3 最大等价路由条数 ····· 1-10
1.6.6 配置OSPFv3 协议的优先级 ····· 1-11
1.6.7 配置OSPFv3 引入外部路由 ····· 1-11
1.7 调整和优化OSPFv3 网络 ····· 1-12
1.7.1 配置准备 ····· 1-12
1.7.2 配置OSPFv3 定时器 ····· 1-12
1.7.3 配置接口的DR优先级 ····· 1-13
1.7.4 忽略DD报文中的MTU检查 ····· 1-14

1.7.5 禁止接口收发OSPFv3 报文 ····· 1-14
1.7.6 配置邻接状态变化输出开关 ····· 1-15
1.8 配置OSPFv3 GR ····· 1-15
1.8.1 配置GR Restarter ····· 1-16
1.8.2 配置GR Helper ····· 1-16
1.9 配置OSPFv3 与BFD联动 ····· 1-16
1.10 配置OSPFv3 IPsec安全策略 ····· 1-17
1.11 OSPFv3 显示和维护 ····· 1-19
1.12 OSPFv3 配置举例 ····· 1-20
1.12.1 配置OSPFv3 区域 ····· 1-20
1.12.2 配置OSPFv3 的DR选择 ····· 1-23
1.12.3 配置OSPFv3 引入外部路由 ····· 1-26
1.12.4 配置OSPFv3 GR ····· 1-29
1.12.5 配置OSPFv3 与BFD联动 ····· 1-31
1.12.6 配置OSPFv3 IPsec安全策略 ····· 1-33
1.13 常见配置错误举例 ····· 1-37
1.13.1 OSPFv3 邻居无法建立 ····· 1-37
1.13.2 OSPFv3 路由信息不正确 ····· 1-37

1 OSPFv3 配置

说明

本文所指的路由器代表运行了路由协议的三层设备。

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 2740（OSPF for IPv6）。

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（Link）运行，OSPFv2 是基于网段（Network）运行。
- OSPFv3 在同一条链路上可以运行多个实例。
- OSPFv3 是通过 Router ID 来标识邻接的邻居。OSPFv2 则是通过 IP 地址来标识邻接的邻居。

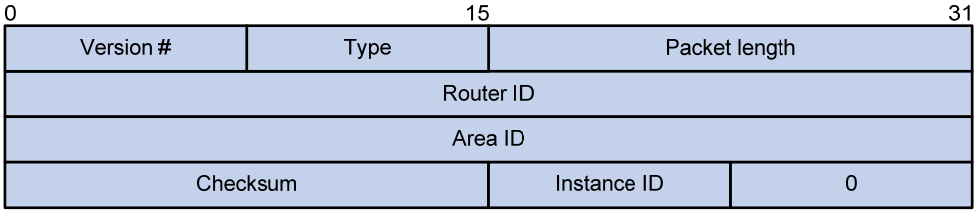
1.1.2 OSPFv3 的协议报文

和 OSPFv2 一样，OSPFv3 也有五种报文类型，分别是 Hello 报文、DD 报文、LSR 报文、LSU 报文和 LSAck 报文。

这五种报文有相同的报文头，但是它和 OSPFv2 的报文头有一些区别，其长度只有 16 字节，且没有认证字段。另外就是多了一个 Instance ID 字段，用来支持在同一条链路上运行多个实例。

OSPFv3 的报文头如 图 1-1 所示。

图1-1 OSPFv3 报文头示意图



主要字段的解释如下：

- Version #：OSPF 的版本号。对于 OSPFv3 来说，其值为 3。
- Type：OSPF 报文的类型。数值从 1 到 5，分别对应 Hello 报文、DD 报文、LSR 报文、LSU 报文和 LSAck 报文。
- Packet Length：OSPF 报文的总长度，包括报文头在内，单位为字节。
- Instance ID：同一条链路上的实例标识。
- 0：保留位，必须为 0。

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。在 RFC 2740 中定义了七类 LSA，描述如下：

- Router-LSA：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：由广播网络和 NBMA（Non-Broadcast Multi-Access）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：和 OSPFv2 中的 Type-3 LSA 类似，该 LSA 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。每一条 Inter-Area-Prefix-LSA 描述了一条到达本自治系统内其他区域的 IPv6 地址前缀（IPv6 Address Prefix）的路由。
- Inter-Area-Router-LSA：和 OSPFv2 中的 Type-4 LSA 类似，该 LSA 由 ABR 生成，在与该 LSA 相关的区域内传播。每一条 Inter-Area-Router-LSA 描述了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：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 区域除外）。缺省路由也可以用 AS-external-LSA 来描述。
- Link-LSA：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。每一个 Link-LSA 描述了该链路上所连接的 IPv6 地址前缀及路由器的 Link-local 地址。
- Intra-Area-Prefix-LSA：每个 Intra-Area-Prefix-LSA 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。

在 RFC 5187 中定义了第十一类 LSA，Grace-LSA。Grace-LSA 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 OSPFv3 的定时器

OSPFv3 的定时器包括：

- OSPFv3 的报文定时器
- LSA 的延迟时间
- SPF 定时器
- GR 定时器

1. OSPFv3 报文定时器

Hello 报文周期性地被发送至邻居路由器，用于发现与维持邻居关系、选举 DR 与 BDR。需要注意的是，网络邻居间的 Hello 时间间隔必须一致，并且 Hello 时钟的值与路由收敛速度、网络负荷大小成反比。

在一定时间间隔内，如果路由器未收到对方的 Hello 报文，则认为对端路由器失效，这个时间间隔被称为相邻路由器间的失效时间。

当一台路由器向它的邻接发送一条 LSA 后，需要等到对方的确认报文。若在设定的重传间隔时间内没有收到对方的确认报文，就会向邻接重传这条 LSA。重传间隔的值必须大于一个报文在两台路由器之间传送一个来回的时间。

2. LSA的延迟时间

由于 LSA 在本路由器的 LSDB（Link State Database，链路状态数据库）中会随时间老化（每秒加 1），但在网络的传输过程中却不会随时间老化，所以有必要在发送之前就将 LSA 的老化时间增加上传送延迟时间。对于低速网络，该项配置尤为重要。

3. SPF定时器

当 OSPFv3 的 LSDB 发生改变时，需要重新计算最短路径，如果每次改变都立即计算最短路径，将占用大量资源，并会影响路由器的效率，通过调节 SPF（Shortest Path First，最短路径优先）的计算延迟时间和间隔时间，可以避免在网络频繁变化时过多的占用资源。

4. GR定时器

在进行 GR 时，可能出现重建邻居失败的情况，这将导致设备长期处于 GR 过程中。GR 定时器即 GR 的最长时间，当它超时时则退出 GR，从而避免设备长期处于 GR 过程中。

1.1.5 系统支持的OSPFv3 特性

在目前的实现中，支持以下 OSPFv3 特性：

- RFC 2740 规定的基本特性
- OSPFv3 Stub 区域
- OSPFv3 多进程（Multi-Process）
- OSPFv3 支持多实例
- OSPFv3 GR
- OSPFv3 与 BFD 联动

1.1.6 协议规范

与 OSPFv3 相关的协议规范有：

- RFC 2740：OSPF for IPv6
- RFC 2328：OSPF Version 2
- 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的虚连接	可选	1.4.3
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OSPFv3的路由聚合	可选	1.6.2
	配置OSPFv3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OSPFv3接口的开销值	可选	1.6.4
	配置OSPFv3最大等价路由条数	可选	1.6.5
	配置OSPFv3协议的优先级	可选	1.6.6
	配置OSPFv3引入外部路由	可选	1.6.7
调整和优化OSPFv3网络	配置OSPFv3定时器	可选	1.7.2
	配置接口的DR优先级	可选	1.7.3
	忽略DD报文中的MTU检查	可选	1.7.4
	禁止接口收发OSPFv3报文	可选	1.7.5
	配置邻接状态变化输出开关	可选	1.7.6
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与BFD联动		可选	1.9
配置OSPFv3 IPsec安全策略		可选	1.10

1.3 使能OSPFv3功能

1.3.1 配置准备

在使能 OSPFv3 功能之前，需完成以下任务：

- 各相邻节点网络层可达
- 使能 IPv6 报文转发功能

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。

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*]	必选 缺省情况下，系统不运行OSPFv3进程
配置路由器的ID	**router-id** *router-id*	必选
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*]	必选 缺省情况下，接口上没有使能OSPFv3

说明

有关多实例的介绍请参见“MPLS 配置指导”中的“IPv6 MPLS L3VPN”。

1.4 配置OSPFv3的区域属性

OSPFv3 支持 Stub 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。

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 IPv6 报文转发功能
- 配置 OSPFv3 基本功能

1.4.2 配置OSPFv3 的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*]	-
进入OSPFv3区域视图	**area** *area-id*	-
配置一个区域为Stub区域	**stub** [**no-summary**]	必选 缺省情况下，没有区域被配置为Stub区域
配置发送到Stub区域的缺省路由的开销值	**default-cost** *value*	可选 缺省情况下，发送到Stub区域的缺省路由的开销值为1

说明

- OSPFv3 的区域不能直接删除，只有区域视图下的所有配置都删除后，且该区域中相关接口的状态都为 down，此区域才会自动删除。
- 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。

1.4.3 配置OSPFv3 的虚连接

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1-4 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*]	-
进入OSPFv3区域视图	**area** *area-id*	-
创建并配置虚连接	**vlink-peer** *router-id* [**hello** *seconds* | **retransmit** *seconds* | **trans-delay** *seconds* | **dead** *seconds* | **instance** *instance-id*] *	必选

说明

- 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。

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。

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，例如：

- 当接口封装的链路层协议是 PPP 时，OSPFv3 接口网络类型的缺省值为点对点。
- 当接口封装的链路层协议是 Ethernet 时，OSPFv3 接口网络类型的缺省值为广播类型。

用户可以根据需要配置 OSPFv3 接口的网络类型，例如：

- 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 IPv6 报文转发功能
- 配置 OSPFv3 基本功能

1.5.2 配置OSPFv3 接口的网络类型

表1-5 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** [**non-broadcast**] | **p2p** } [**instance** *instance-id*]	可选 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选举权等。

表1-6 配置 NBMA 或者 P2MP(单播)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* [**dr-priority** *dr-priority*] [**instance** *instance-id*]	必选

1.6 配置OSPFv3的路由信息控制

通过这部分的配置，可以控制 OSPFv3 的路由信息的发布与接收，并引入外部路由。

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：

- 使能 IPv6 报文转发功能
- 配置 OSPFv3 基本功能

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1-7 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*]	-
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**]	必选 缺省情况下，ABR没有对路由进行聚合

说明

路由聚合只有在 ABR 上配置才会有效。

1.6.3 配置OSPFv3 对通过接收到的LSA计算出来的路由信息进行过滤

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1-8 配置 OSPFv3 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OSPFv3视图	**ospfv3** [*process-id*]	-
配置对接收的路由信息进行过滤	**filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* } **import**	必选 缺省情况下，没有对接收的路由信息过滤

说明

filter-policy import 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。

如果没有在接口下配置接口开销值，OSPFv3 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1-9 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*	-
设置OSPFv3接口的开销值	**ospfv3 cost** *value* [**instance** *instance-id*]	可选 缺省情况下，VLAN接口的缺省值为1；Loopback接口的缺省值为0；其他接口按照当前的带宽自动计算接口运行OSPFv3协议所需的开销

2. 配置带宽参考值

表1-10 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*]	-
配置带宽参考值	**bandwidth-reference** *value*	可选 缺省情况下，带宽参考值为100Mbits/s

1.6.5 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1-11 配置 OSPFv3 最大等价路由条数

操作	命令	说明
进入系统视图	**system-view**	-
进入OSPFv3视图	**ospfv3** [*process-id*]	-
配置OSPFv3最大等价路由条数	**maximum load-balancing** *maximum*	可选 缺省值为8

1.6.6 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1-12 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* |	-
配置OSPFv3协议的优先级	**preference** [**ase**] [**route-policy** *route-policy-name*] *preference*	可选 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

1.6.7 配置OSPFv3 引入外部路由

表1-13 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*]	-
配置外部路由的缺省开销值	**default cost** *value*	可选 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* | **allow-ibgp**] [**cost** *value* | **route-policy** *route-policy-name* | **type** *type*] *	必选 缺省情况下，没有引入外部路由信息
配置OSPFv3引入缺省路由	**default-route-advertise** [**always** | **cost** *value* | **type** *type* | **route-policy** *route-policy-name*] *	可选 缺省情况下，没有引入缺省路由
对引入的外部路由信息进行过滤	**filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* } **export** [**isisv6** *process-id* | **ospfv3** *process-id* | **ripng** *process-id* | **bgp4+** | **direct** | **static**]	可选 缺省情况下，没有对引入的路由信息过滤

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。
- 只能通过 **default-route-advertise** 命令引入并发布缺省路由。
- 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。
- **filter-policy export** 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 **filter-policy export** 命令失效。

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、忽略 DD 报文中的 MTU 检查和禁止接口收发 OSPFv3 报文。

OSPFv3 包括如下定时器：

- OSPFv3 的报文定时器，可以调整 OSPFv3 网络的收敛速度以及协议报文带来的网络负荷。
- LSA 的延迟时间：在一些低速链路上，需要考虑接口传送 LSA 的延迟时间。
- SPF 定时器：通过调整 SPF 计算的延迟时间和抑制间隔时间，可以抑制由于网络频繁变化带来的资源消耗问题。

对于广播网，通过配置接口的 DR 优先级来影响 DR/BDR 的选择。

如果要使 OSPFv3 路由信息不被某一网络中的路由器获得，可禁止在与该网络相连的接口上发送 OSPFv3 报文。

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：

- 使能 IPv6 报文转发功能
- 配置 OSPFv3 基本功能

1.7.2 配置OSPFv3 定时器

表1-14 配置 OSPF 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*	-

操作	命令	说明
配置接口发送hello报文的时间间隔	**ospfv3 timer hello** *seconds* [**instance** *instance-id*]	可选 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Poll定时器	**ospfv3 timer poll** *seconds* [**instance** *instance-id*]	可选 缺省情况下，发送轮询Hello报文的时间间隔为120秒
配置相邻路由器间失效时间	**ospfv3 timer dead** *seconds* [**instance** *instance-id*]	可选 缺省情况下，P2P、Broadcast网络类型接口的OSPFv3邻居失效时间为40秒
配置相邻路由器重传LSA的时间间隔	**ospfv3 timer retransmit** *interval* [**instance** *instance-id*]	可选 缺省情况下，LSA的重传时间间隔为5秒
配置接口的LSA传输延迟时间	**ospfv3 trans-delay** *seconds* [**instance** *instance-id*]	可选 缺省情况下，接口的LSA传输延迟时间为1秒
退回系统视图	**quit**	-
进入OSPFv3视图	**ospfv3** [*process-id*]	-
设置SPF定时器	**spf timers** *delay-interval hold-interval*	可选 缺省情况下，delay-interval值为5秒，hold-interval值为10秒 当延迟时间和间隔时间都配置为0秒时，表示立即触发SPF计算，可以加快收敛速度

说明

- 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。
- 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DR优先级

表1-15 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*	-

操作	命令	说明
配置接口的DR优先级	**ospfv3 dr-priority** *priority* [**instance** *instance-id*]	可选 缺省情况下，接口的DR优先级为1

说明

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。

1.7.4 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1-16 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*]	必选 缺省情况下，不忽略MTU检查

1.7.5 禁止接口收发OSPFv3 报文

表1-17 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*]	-
禁止接口收发OSPFv3报文	**silent-interface** { *interface-type interface-number* | **all** }	必选 缺省情况下，允许接口收发OSPFv3报文

说明

- 不同的进程可以对同一接口禁止收发 OSPFv3 报文，但 **silent-interface** 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。
- 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。

1.7.6 配置邻接状态变化输出开关

表1-18 配置邻接状态变化输出开关

操作	命令	说明
进入系统视图	**system-view**	-
进入OSPFv3视图	**ospfv3** [*process-id*]	-
配置邻接状态变化输出开关	**log-peer-change**	必选 缺省情况下，邻接状态变化输出开关处于打开状态

1.8 配置OSPFv3 GR

说明

OSPFv3 GR 特性和 OSPFv3 虚连接特性不同时支持，在配置了虚连接后将不能再配置 GR 特性。

GR（Graceful Restart，平滑重启）是一种在协议重启或主备切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1-19 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*]	-
使能GR能力	**graceful-restart enable**	必选 缺省情况下，OSPFv3协议的GR Restarter能力处于关闭状态
配置GR重启时间间隔	**graceful-restart interval** *interval-value*	可选 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1-20 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*]	-
使能helper能力	**graceful-restart helper enable**	可选 缺省情况下，OSPFv3的GR Helper能力处于打开状态
使能LSA严格检查能力	**graceful-restart helper strict-lsa-checking**	可选 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。

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1-21 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*]	-
配置路由器的ID	**router-id** *router-id*	必选
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*]	必选 缺省情况下，接口上没有使能OSPFv3
在指定接口上使能OSPFv3 BFD	**ospfv3 bfd enable** [**instance** *instance-id*]	必选 缺省情况下，运行OSPFv3的接口未使能BFD功能

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”。

1.10 配置OSPFv3 IPsec安全策略

从安全性角度来考虑，为了避免路由信息外泄或者对路由器进行恶意攻击，OSPFv3 提供基于 IPsec 的报文验证功能。

路由器在发送的报文中会携带配置好的 IPsec 安全策略的 SPI（Security Parameter Index，安全参数索引）值，接收报文时通过 SPI 值进行 IPsec 安全策略匹配：只有策略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 安全策略。

- 当需要保护区域内的所有报文时，可以在区域下配置 IPsec 安全策略，此时区域内所有路由器都需要配置相同的 IPsec 安全策略。
- 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 安全策略，此时直连邻居接口需要配置相同的 IPsec 安全策略。
- 当需要保护虚连接的报文时，可以配置虚连接应用 IPsec 安全策略，此时虚连接上的两个邻居需要配置相同的 IPsec 安全策略。

当接口和接口所在区域均配置了 IPsec 安全策略时，接口下的策略生效；当虚连接和区域 0 均配置了 IPsec 安全策略时，虚连接的策略有效。

1. 配置准备

配置 OSPFv3 IPsec 安全策略之前，需要满足下面配置：

- 安全提议已经创建。
- IPsec 安全策略已经创建。

关于安全提议和 IPsec 安全策略的详细配置过程，请参见“安全配置指导”中的“IPsec”。

2. 配置OSPFv3 IPsec安全策略

表1-22 配置 OSPFv3 IPsec 安全策略（区域）

操作	命令	说明
进入系统视图	**system-view**	-
进入OSPFv3视图	**ospfv3** [*process-id*]	-
进入OSPF区域视图	**area** *area-id*	-
配置OSPFv3区域应用IPsec安全策略	**enable ipsec-policy** *policy-name*	必选 缺省情况下，OSPFv3区域没有应用IPsec安全策略

表1-23 配置 IPsec 安全策略（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安全策略	**ospfv3 ipsec-policy** *policy-name* [**instance** *instance-id*]	必选 缺省情况下，使能了OSPFv3的接口上没有应用IPsec安全策略

表1-24 配置 IPsec 安全策略（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*]	-
进入OSPF区域视图	**area** *area-id*	-
配置虚连接应用IPsec安全策略	**vlink-peer** *router-id* [**hello** *seconds* | **retransmit** *seconds* | **trans-delay** *seconds* | **dead** *seconds* | **instance** *instance-id* | **ipsec-policy** *policy-name*] *	必选 缺省情况下，虚连接没有应用IPsec安全策略

说明

OSPFv3 IPsec 安全策略只支持手工配置方式，具体情况请参见“安全配置指导”中的“IPsec”。

1.11 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1-25 OSPFv3 显示和维护

操作	命令
显示OSPFv3进程的概要信息	**display ospfv3** [*process-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3接口信息	**display ospfv3 interface** [*interface-type interface-number* | **statistic**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3的LSDB信息	**display ospfv3** [*process-id*] **lsdb** [[**external** | **inter-prefix** | **inter-router** | **intra-prefix** | **link** | **network** | **router** | **grace**] [*link-state-id*] [**originate-router** *router-id*] | **total**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3 LSDB中LSA的统计信息	**display ospfv3 lsdb statistic** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3邻居信息	**display ospfv3** [*process-id*] [**area** *area-id*] **peer** [[*interface-type interface-number*] [**verbose**] | *peer-router-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3邻居的统计信息	**display ospfv3 peer statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3路由表信息	**display ospfv3** [*process-id*] **routing** [*ipv6-address prefix-length* **|** *ipv6-address/prefix-length* | **abr-routes** | **asbr-routes** | **all | statistics**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3区域拓扑信息	**display ospfv3** [*process-id*] **topology** [**area** *area-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3虚连接信息	**display ospfv3** [*process-id*] **vlink** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3下一跳信息	**display ospfv3** [*process-id*] **next-hop** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3链路状态请求列表信息	**display ospfv3** [*process-id*] **request-list** [{ **external** | **inter-prefix** | **inter-router** | **intra-prefix** | **link** | **network** | **router** | **grace** } [*link-state-id*] [**originate-router** *ip-address*] | **statistics**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3链路状态重传列表信息	**display ospfv3** [*process-id*] **retrans-list** [{ **external** | **inter-prefix** | **inter-router** | **intra-prefix** | **link** | **network** | **router** | **grace** } [*link-state-id*] [**originate-router** *ip-address*] | **statistics**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示接口收发OSPFv3报文的统计信息	**display ospfv3 statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示OSPFv3进程的GR状态信息	**display ospfv3** [*process-id*] **graceful-restart status** [**|** { **begin** | **exclude** | **include** } *regular-expression*]

1.12 OSPFv3配置举例

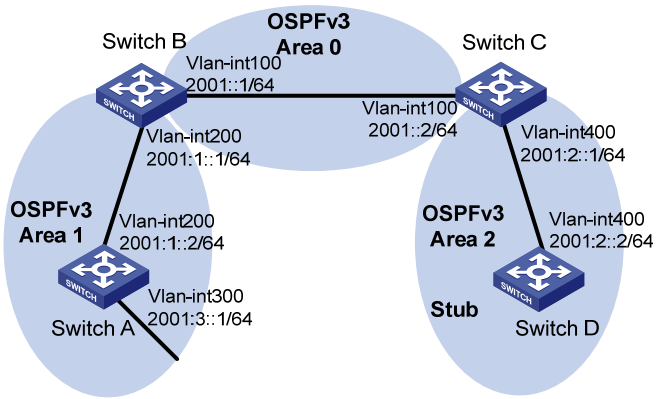
1.12.1 配置OSPFv3 区域

1. 组网需求

- 所有的交换机都运行 OSPFv3，整个自治系统划分为 3 个区域。其中 Switch B 和 Switch C 作为 ABR 来转发区域之间的路由。
- 要求将 Area2 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1-2 配置 OSPFv3 区域组网图



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 Switch A。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 300
[SwitchA-Vlan-interface300] ospfv3 1 area 1
[SwitchA-Vlan-interface300] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ospfv3 1 area 1
[SwitchA-Vlan-interface200] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 100
```

```
[SwitchB-Vlan-interface100] ospfv3 1 area 0
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ospfv3 1 area 1
[SwitchB-Vlan-interface200] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] ospfv3 1 area 0
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 400
[SwitchC-Vlan-interface400] ospfv3 1 area 2
[SwitchC-Vlan-interface400] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] ipv6
[SwitchD] ospfv3
[SwitchD-ospfv3-1] router-id 4.4.4.4
[SwitchD-ospfv3-1] quit
[SwitchD] interface vlan-interface 400
[SwitchD-Vlan-interface400] ospfv3 1 area 2
[SwitchD-Vlan-interface400] quit
```

查看 Switch B 的 OSPFv3 邻居状态。

```
[SwitchB] display ospfv3 peer

            OSPFv3 Area ID 0.0.0.0 (Process 1)
 ------------------------------------------------------------------------
Neighbor ID    Pri   State          Dead Time   Interface      Instance ID
3.3.3.3        1     Full/DR        00:00:39    Vlan100        0

            OSPFv3 Area ID 0.0.0.1 (Process 1)
 ------------------------------------------------------------------------
Neighbor ID     Pri   State         Dead Time   Interface      Instance ID
1.1.1.1         1     Full/Backup   00:00:38    Vlan200        0
```

查看 Switch C 的 OSPFv3 邻居状态。

```
[SwitchC] display ospfv3 peer
            OSPFv3 Area ID 0.0.0.0 (Process 1)
 ------------------------------------------------------------------------
Neighbor ID    Pri   State          Dead Time   Interface      Instance ID
2.2.2.2        1     Full/Backup    00:00:39    Vlan100        0

            OSPFv3 Area ID 0.0.0.2 (Process 1)
 ------------------------------------------------------------------------
```

```
Neighbor ID     Pri   State         Dead Time   Interface      Instance ID
4.4.4.4         1     Full/DR       00:00:38    Vlan400        0
```

查看 Switch D 的 OSPFv3 路由表信息。

```
[SwitchD] display ospfv3 routing

E1 - Type 1 external route,    IA - Inter area route,    I  - Intra area route
E2 - Type 2 external route,    *  - Seleted route

            OSPFv3 Router with ID (4.4.4.4) (Process 1)
------------------------------------------------------------------------
 *Destination: 2001::/64
  Type       : IA                                    Cost     : 2
  NextHop    : FE80::F40D:0:93D0:1                   Interface: Vlan400

 *Destination: 2001:1::/64
  Type       : IA                                    Cost     : 3
  NextHop    : FE80::F40D:0:93D0:1                   Interface: Vlan400

 *Destination: 2001:2::/64
  Type       : I                                     Cost     : 1
  NextHop    : directly-connected                    Interface: Vlan400

 *Destination: 2001:3::/64
  Type       : IA                                    Cost     : 4
  NextHop    : FE80::F40D:0:93D0:1                   Interface: Vlan400
```

(3) 配置 Stub 区域

配置 Switch D 的 Stub 区域。

```
[SwitchD] ospfv3
[SwitchD-ospfv3-1] area 2
[SwitchD-ospfv3-1-area-0.0.0.2] stub
```

配置 Switch C 的 Stub 区域，设置发送到 Stub 区域的缺省路由的开销为 10。

```
[SwitchC] ospfv3
[SwitchC-ospfv3-1] area 2
[SwitchC-ospfv3-1-area-0.0.0.2] stub
[SwitchC-ospfv3-1-area-0.0.0.2] default-cost 10
```

查看 Switch 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。

```
[SwitchD] display ospfv3 routing
E1 - Type 1 external route,    IA - Inter area route,    I  - Intra area route
E2 - Type 2 external route,    *  - Seleted route

            OSPFv3 Router with ID (4.4.4.4) (Process 1)
 ------------------------------------------------------------------------
 *Destination: ::/0
  Type       : IA                                    Cost     : 11
  NextHop    : FE80::F40D:0:93D0:1                   Interface: Vlan400
```

```
 *Destination: 2001::/64
  Type       : IA                                  Cost     : 2
  NextHop    : FE80::F40D:0:93D0:1                 Interface: Vlan400

 *Destination: 2001:1::/64
  Type       : IA                                  Cost     : 3
  NextHop    : FE80::F40D:0:93D0:1                 Interface: Vlan400

 *Destination: 2001:2::/64
  Type       : I                                   Cost     : 1
  NextHop    : directly-connected                  Interface: Vlan400

 *Destination: 2001:3::/64
  Type       : IA                                  Cost     : 4
  NextHop    : FE80::F40D:0:93D0:1                 Interface: Vlan400
```

(4) 配置 Totally Stub 区域

配置 Switch C，设置 Area2 为 Totally Stub 区域。

```
[SwitchC-ospfv3-1-area-0.0.0.2] stub no-summary
```

查看 Switch 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。

```
[SwitchD] display ospfv3 routing
E1 - Type 1 external route,    IA - Inter area route,    I  - Intra area route
E2 - Type 2 external route,    *  - Seleted route

            OSPFv3 Router with ID (4.4.4.4) (Process 1)
 ------------------------------------------------------------------------
 *Destination: ::/0
  Type       : IA                                  Cost     : 11
  NextHop    : FE80::F40D:0:93D0:1                 Interface: Vlan400

 *Destination: 2001:2::/64
  Type       : I                                   Cost     : 1
  NextHop    : directly-connected                  Interface: Vlan400
```

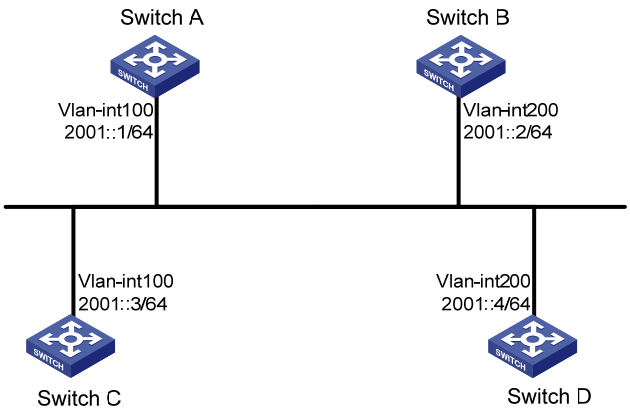
1.12.2 配置OSPFv3 的DR选择

1. 组网需求

- Switch A 的优先级为 100，它是网络上的最高优先级，所以 Switch A 被选为 DR；
- Switch C 的优先级为 2，它是优先级次高的，被选为 BDR；
- Switch B 的优先级为 0，这意味着它将无法成为 DR；
- Switch D 没有配置优先级，取缺省值 1。

2. 组网图

图1-3 配置 OSPFv3 的 DR 选择组网图



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 Switch A，启动 OSPFv3，并设置其 Router ID 为 1.1.1.1。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ospfv3 1 area 0
[SwitchA-Vlan-interface100] quit
```

配置 Switch B，启动 OSPFv3，并设置其 Router ID 为 2.2.2.2。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ospfv3 1 area 0
[SwitchB-Vlan-interface200] quit
```

配置 Switch C，启动 OSPFv3，并设置其 Router ID 为 3.3.3.3。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] ospfv3 1 area 0
[SwitchC-Vlan-interface100] quit
```

配置 Switch D，启动 OSPFv3，并设置其 Router ID 为 4.4.4.4。

```
<SwitchD> system-view
[SwitchD] ipv6
[SwitchD] ospfv3
[SwitchD-ospfv3-1] router-id 4.4.4.4
[SwitchD-ospfv3-1] quit
[SwitchD] interface vlan-interface 200
[SwitchD-Vlan-interface200] ospfv3 1 area 0
[SwitchD-Vlan-interface200] quit
```

查看 Switch 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 Switch D 为 DR，Switch C 为 BDR。

```
[SwitchA] display ospfv3 peer
            OSPFv3 Area ID 0.0.0.0 (Process 1)
 ------------------------------------------------------------------------
Neighbor ID     Pri   State            Dead Time   Interface      Instance ID
2.2.2.2         1     2-Way/DROther    00:00:36    Vlan200        0
3.3.3.3         1     Full/Backup      00:00:35    Vlan100        0
4.4.4.4         1     Full/DR          00:00:33    Vlan200        0
```

查看 Switch D 的邻居信息，可以看到 Switch D 和其他邻居之间的邻居状态都为 Full。

```
[SwitchD] display ospfv3 peer
            OSPFv3 Area ID 0.0.0.0 (Process 1)
 ------------------------------------------------------------------------
Neighbor ID     Pri   State            Dead Time   Interface      Instance ID
1.1.1.1         1     Full/DROther     00:00:30    Vlan100        0
2.2.2.2         1     Full/DROther     00:00:37    Vlan200        0
3.3.3.3         1     Full/Backup      00:00:31    Vlan100        0
```

(3) 配置接口的 DR 优先级

配置 Switch A 的接口 Vlan-interface100 的 DR 优先级为 100。

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ospfv3 dr-priority 100
[SwitchA-Vlan-interface100] quit
```

配置 Switch B 的接口 Vlan-interface200 的 DR 优先级为 0。

```
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ospfv3 dr-priority 0
[SwitchB-Vlan-interface200] quit
```

配置 Switch C 的接口 Vlan-interface100 的 DR 优先级为 2。

```
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] ospfv3 dr-priority 2
[SwitchC-Vlan-interface100] quit
```

显示 Switch 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。

```
[SwitchA] display ospfv3 peer
            OSPFv3 Area ID 0.0.0.0 (Process 1)
 ------------------------------------------------------------------------
Neighbor ID     Pri   State            Dead Time   Interface      Instance ID
2.2.2.2         0     2-Way/DROther    00:00:38    Vlan200        0
3.3.3.3         2     Full/Backup      00:00:32    Vlan100        0
4.4.4.4         1     Full/DR          00:00:36    Vlan200        0
```

显示 Switch D 的邻居信息，可以看到 Switch D 仍然为 DR。

```
[SwitchD] display ospfv3 peer
            OSPFv3 Area ID 0.0.0.0 (Process 1)
 ------------------------------------------------------------------------
Neighbor ID     Pri   State            Dead Time   Interface      Instance ID
1.1.1.1         100   Full/DROther     00:00:33    Vlan100        0
2.2.2.2         0     Full/DROther     00:00:36    Vlan200        0
3.3.3.3         2     Full/Backup      00:00:40    Vlan100        0
```

(4) 重新进行 DR/BDR 选择

将所有接口一次 **shutdown** 和 **undo shutdown**，使 OSPFv3 进行 DR/BDR 的重新选举。

查看 Switch A 的邻居信息，可以看到 Switch C 为 BDR。

```
[SwitchA] display ospfv3 peer
            OSPFv3 Area ID 0.0.0.0 (Process 1)
 ------------------------------------------------------------------------
Neighbor ID     Pri   State            Dead Time   Interface      Instance ID
2.2.2.2         0     Full/DROther     00:00:31    Vlan200        0
3.3.3.3         2     Full/Backup      00:00:39    Vlan100        0
4.4.4.4         1     Full/DROther     00:00:37    Vlan200        0
```

查看 Switch D 的邻居信息，可以看到 Switch A 为 DR。

```
[SwitchD] display ospfv3 peer
            OSPFv3 Area ID 0.0.0.0 (Process 1)
 ------------------------------------------------------------------------
Neighbor ID     Pri   State            Dead Time   Interface      Instance ID
1.1.1.1         100   Full/DR          00:00:34    Vlan100        0
2.2.2.2         0     2-Way/DROther    00:00:34    Vlan200        0
3.3.3.3         2     Full/Backup      00:00:32    Vlan100        0
```

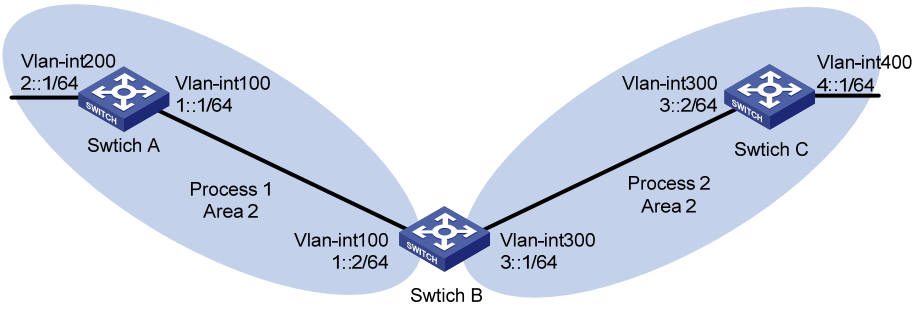
1.12.3 配置OSPFv3 引入外部路由

1. 组网需求

- Switch A、Switch B 和 Switch C 位于 Area 2 内；
- Switch B 上运行两个 OSPFv3 进程：OSPFv3 1 和 OSPFv3 2。Switch B 通过 OSPFv3 1 和 Switch A 交换路由信息，通过 OSPFv3 2 和 Switch C 交换路由信息；
- 在 Switch 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 Switch C 能够学习到达 1::0/64 和 2::0/64 的路由，但 Switch 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 Switch A 上启动 OSPFv3 进程 1。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3 1
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ospfv3 1 area 2
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ospfv3 1 area 2
[SwitchA-Vlan-interface200] quit
```

在 Switch B 上启动两个 OSPFv3 进程，进程号分别为 1 和 2。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3 1
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ospfv3 1 area 2
[SwitchB-Vlan-interface100] quit
[SwitchB] ospfv3 2
[SwitchB-ospfv3-2] router-id 3.3.3.3
[SwitchB-ospfv3-2] quit
[SwitchB] interface vlan-interface 300
[SwitchB-Vlan-interface300] ospfv3 2 area 2
[SwitchB-Vlan-interface300] quit
```

在 Switch C 上启动 OSPFv3 进程 2。

```
<SwitchC> system-view
[SwitchC] ipv6
```

```
[SwitchC] ospfv3 2
[SwitchC-ospfv3-2] router-id 4.4.4.4
[SwitchC-ospfv3-2] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] ospfv3 2 area 2
[SwitchC-Vlan-interface300] quit
[SwitchC] interface vlan-interface 400
[SwitchC-Vlan-interface400] ospfv3 2 area 2
[SwitchC-Vlan-interface400] quit
```

查看 Switch C 的路由表信息。

```
[SwitchC] display ipv6 routing-table
Routing Table :
         Destinations : 6        Routes : 6

Destination: ::1/128                                    Protocol  : Direct
NextHop    : ::1                                        Preference: 0
Interface  : InLoop0                                    Cost      : 0

Destination: 3::/64                                     Protocol  : Direct
NextHop    : 3::2                                       Preference: 0
Interface  : Vlan300                                    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  : Vlan400                                    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
```

(3)　配置 OSPFv3 引入外部路由

在 Switch B 上配置 OSPFv3 引入外部路由,，引入直连路由和 OSPFv3 进程 1 的路由。

```
[SwitchB] ospfv3 2
[SwitchB-ospfv3-2] default cost 3
[SwitchB-ospfv3-2] import-route ospfv3 1
[SwitchB-ospfv3-2] import-route direct
[SwitchB-ospfv3-2] quit
```

查看路由引入后 Switch C 的路由表信息。

```
[SwitchC] display ipv6 routing-table
Routing Table :
```

```
         Destinations : 8        Routes : 8

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  : Vlan300                                    Cost      : 3

Destination: 2::/64                                     Protocol  : OSPFv3
NextHop    : FE80::200:CFF:FE01:1C03                    Preference: 150
Interface  : Vlan300                                    Cost      : 3

Destination: 3::/64                                     Protocol  : Direct
NextHop    : 3::2                                       Preference: 0
Interface  : Vlan300                                    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  : Vlan400                                    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
```

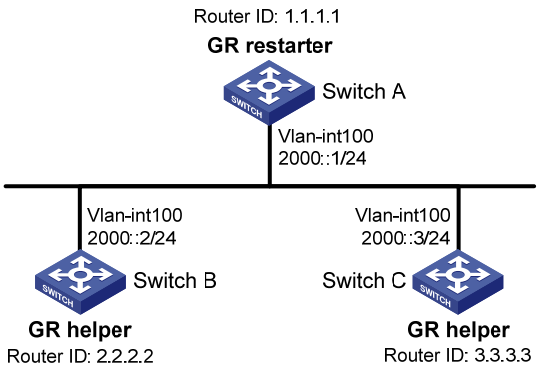
1.12.4 配置OSPFv3 GR

1. 组网需求

- Switch A、Switch B 和 Switch 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。
- Switch A 作为 GR Restarter，Switch B 和 Switch C 作为 GR Helper 并且通过 GR 机制与 Switch 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 Switch A，启动 OSPFv3，并设置其 Router ID 为 1.1.1.1。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3 1
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] graceful-restart enable
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ospfv3 1 area 1
[SwitchA-Vlan-interface100] quit
```

配置 Switch B，启动 OSPFv3，并设置其 Router ID 为 2.2.2.2。缺省情况下，Switch B 的 GR helper 能力处于开启状态。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3 1
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ospfv3 1 area 1
[SwitchB-Vlan-interface100] quit
```

配置 Switch C，启动 OSPFv3，并设置其 Router ID 为 3.3.3.3。缺省情况下，Switch C 的 GR helper 能力处于开启状态。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3 1
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] ospfv3 1 area 1
```

```
[SwitchC-Vlan-interface100] quit
```

(3) 检验配置效果

运行稳定后，在 Switch A 上主备倒换进入 OSPFv3 协议的 GR 进程。

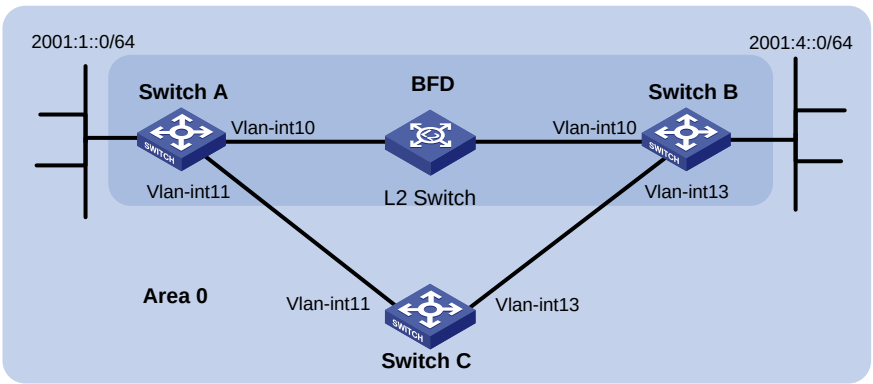
1.12.5 配置OSPFv3 与BFD联动

1. 组网需求

- Switch A、Switch B 和 Switch C 上运行 OSPFv3，网络层相互可达。
- 当 Switch A 和 Switch 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 Switch C 进行通信。

2. 组网图

图1-6 配置 OSPFv3 与 BFD 联动组网图



设备	接口	IPv6地址	设备	接口	IPv6地址
Switch A	Vlan-int10	2001::1/64	Switch B	Vlan-int10	2001::2/64
	Vlan-int11	2001:2::1/64		Vlan-int13	2001:3::2/64
Switch C	Vlan-int11	2001:2::2/64			
	Vlan-int13	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 Switch A，启动 OSPFv3，并设置其 Router ID 为 1.1.1.1。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] ospfv3 1 area 0
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] ospfv3 1 area 0
[SwitchA-Vlan-interface11] quit
```

配置 Switch B，启动 OSPFv3，并设置其 Router ID 为 2.2.2.2。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ospfv3 1 area 0
[SwitchB-Vlan-interface10] quit
[SwitchB] interface vlan-interface 13
[SwitchB-Vlan-interface13] ospfv3 1 area 0
[SwitchB-Vlan-interface13] quit
```

配置 Switch C，启动 OSPFv3，并设置其 Router ID 为 3.3.3.3。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 11
[SwitchC-Vlan-interface11] ospfv3 1 area 0
[SwitchC-Vlan-interface11] quit
[SwitchC] interface vlan-interface 13
[SwitchC-Vlan-interface13] ospfv3 1 area 0
[SwitchC-Vlan-interface13] quit
```

(3) 配置 BFD 功能

在 Switch A 上使能 BFD 检测功能，并配置 BFD 参数。

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] ospfv3 bfd enable
[SwitchA-Vlan-interface10] bfd min-transmit-interval 500
[SwitchA-Vlan-interface10] bfd min-receive-interval 500
[SwitchA-Vlan-interface10] bfd detect-multiplier 7
[SwitchA-Vlan-interface10] return
```

在 Switch B 上使能 BFD 检测功能，并配置 BFD 参数。

```
[SwitchB] bfd session init-mode active
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ospfv3 bfd enable
[SwitchB-Vlan-interface10] bfd min-transmit-interval 500
[SwitchB-Vlan-interface10] bfd min-receive-interval 500
[SwitchB-Vlan-interface10] bfd detect-multiplier 6
```

(4) 检查配置结果

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

显示 Switch A 的 BFD 信息。

```
<SwitchA> display bfd session
Total 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（Switch A 接口 Vlan-interface10 的链路本地地址）
  Destination IP: FE80::20F:FF:FE00:1200（Switch B 接口 Vlan-interface10 的链路本地地址）
   Session State: Up                    Interface: Vlan10
       Hold Time:    /
```

在 Switch A 上查看 2001:4::0/64 的路由信息，可以看出 Switch A 和 Switch B 是通过 L2 Switch 进行通信的。

```
<SwitchA> display ipv6 routing-table 2001:4::0 64 verbose
Routing Table :
Summary Count : 1

 Destination  : 2001:4::                          PrefixLength : 64
 NextHop      : 2001::2                          Preference   : 10
 IpPrecedence :                                  QosLcId      :
 RelayNextHop : ::                               Tag          : 0H
 Neighbor     : ::                               ProcessID    : 0
 Interface    : Vlan-interface10                 Protocol     : OSPFv3
 State        : Active Adv                       Cost         : 1
 Tunnel ID    : 0x0                              Label        : NULL
 Age          : 4538sec
```

当 Switch B 和二层交换机之间的链路发生故障时 Switch A 能够快速检测 Switch B 的变化。

查看 Switch A 的 BFD 信息，Switch A 已取消与 Switch B 的邻居关系，没有任何输出信息。

```
<SwitchA> display bfd session
```

在 Switch A 上查看 2001:4::0/64 的路由信息，可以看出 Switch A 和 Switch B 已经切换到 Switch C 进行通信。

```
<SwitchA> display ipv6 routing-table 2001:4::0 64 verbose
Routing Table :
Summary Count : 1

 Destination  : 2001:4::                          PrefixLength : 64
 NextHop      : 2001:2::2                        Preference   : 10
 IpPrecedence :                                  QosLcId      :
 RelayNextHop : ::                               Tag          : 0H
 Neighbor     : ::                               ProcessID    : 0
 Interface    : Vlan-interface11                 Protocol     : OSPFv3
 State        : Active  Adv                      Cost         : 2
 Tunnel ID    : 0x0                              Label        : NULL
 Age          : 4610sec
```

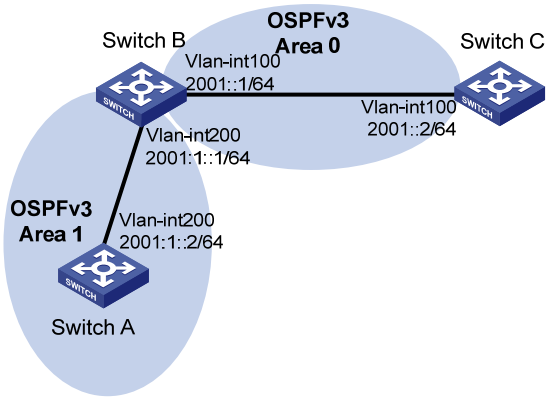
1.12.6 配置OSPFv3 IPsec安全策略

1. 组网需求

- 所有的交换机都运行 OSPFv3，整个自治系统划分为 2 个区域。
- 要求配置 IPsec 安全策略对 Switch A、Switch B 和 Switch C 之间的 OSPFv3 报文进行有效性检查和验证。

2. 组网图

图1-7 配置 OSPFv3 IPsec 策略组网图



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 Switch A，启动 OSPFv3，并设置其 Router ID 为 1.1.1.1。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3 1
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ospfv3 1 area 1
[SwitchA-Vlan-interface200] quit
```

配置 Switch B，启动 OSPFv3，并设置其 Router ID 为 2.2.2.2。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3 1
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ospfv3 1 area 0
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ospfv3 1 area 1
[SwitchB-Vlan-interface200] quit
```

在 Switch C，启动 OSPFv3，并设置其 Router ID 为 3.3.3.3。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3 1
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 100
```

```
[SwitchC-Vlan-interface100] ospfv3 1 area 0
[SwitchC-Vlan-interface100] quit
```

(3) 配置 OSPFv3 IPsec 安全策略

配置 Switch 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。创建一条安全策略 policy001，协商方式为 manual，配置 SPI 和密钥。

```
[SwitchA] ipsec proposal tran1
[SwitchA-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchA-ipsec-proposal-tran1] transform esp
[SwitchA-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchA-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchA-ipsec-proposal-tran1] quit
[SwitchA] ipsec policy policy001 10 manual
[SwitchA-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchA-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] quit
```

配置 Switch 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；创建一条安全策略 policy001，协商方式为 manual，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；创建一条安全策略 policy002，协商方式为 manual，配置 SPI 和密钥。

```
[SwitchB] ipsec proposal tran1
[SwitchB-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran1] transform esp
[SwitchB-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran1] quit
[SwitchB] ipsec policy policy001 10 manual
[SwitchB-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchB-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] quit
[SwitchB] ipsec proposal tran2
[SwitchB-ipsec-proposal-tran2] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran2] transform esp
[SwitchB-ipsec-proposal-tran2] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran2] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran2] quit
[SwitchB] ipsec policy policy002 10 manual
[SwitchB-ipsec-policy-manual-policy002-10] proposal tran2
[SwitchB-ipsec-policy-manual-policy002-10] sa spi outbound esp 54321
[SwitchB-ipsec-policy-manual-policy002-10] sa spi inbound esp 54321
[SwitchB-ipsec-policy-manual-policy002-10] sa string-key outbound esp gfedcba
[SwitchB-ipsec-policy-manual-policy002-10] sa string-key inbound esp gfedcba
```

```
[SwitchB-ipsec-policy-manual-policy002-10] quit
```

配置 Switch 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。创建一条安全策略 policy002，协商方式为 manual，配置 SPI 和密钥。

```
[SwitchC] ipsec proposal tran2
[SwitchC-ipsec-proposal-tran2] encapsulation-mode transport
[SwitchC-ipsec-proposal-tran2] transform esp
[SwitchC-ipsec-proposal-tran2] esp encryption-algorithm des
[SwitchC-ipsec-proposal-tran2] esp authentication-algorithm sha1
[SwitchC-ipsec-proposal-tran2] quit
[SwitchC] ipsec policy policy002 10 manual
[SwitchC-ipsec-policy-manual-policy002-10] proposal tran2
[SwitchC-ipsec-policy-manual-policy002-10] sa spi outbound esp 54321
[SwitchC-ipsec-policy-manual-policy002-10] sa spi inbound esp 54321
[SwitchC-ipsec-policy-manual-policy002-10] sa string-key outbound esp gfedcba
[SwitchC-ipsec-policy-manual-policy002-10] sa string-key inbound esp gfedcba
[SwitchC-ipsec-policy-manual-policy002-10] quit
```

(4) OSPFv3 区域上应用 IPsec 安全策略

配置 Switch A。

```
[SwitchA] ospfv3 1
[SwitchA-ospfv3-1] area 1
[SwitchA-ospfv3-1-area-0.0.0.1] enable ipsec-policy policy001
[SwitchA-ospfv3-1-area-0.0.0.1] quit
[SwitchA-ospfv3-1] quit
```

配置 Switch B。

```
[SwitchB] ospfv3 1
[SwitchB-ospfv3-1] area 0
[SwitchB-ospfv3-1-area-0.0.0.0] enable ipsec-policy policy002
[SwitchB-ospfv3-1-area-0.0.0.0] quit
[SwitchB-ospfv3-1] area 1
[SwitchB-ospfv3-1-area-0.0.0.1] enable ipsec-policy policy001
[SwitchB-ospfv3-1-area-0.0.0.1] quit
[SwitchB-ospfv3-1] quit
```

配置 Switch C。

```
[SwitchC] ospfv3 1
[SwitchC-ospfv3-1] area 0
[SwitchC-ospfv3-1-area-0.0.0.0] enable ipsec-policy policy002
[SwitchC-ospfv3-1-area-0.0.0.0] quit
[SwitchC-ospfv3-1] quit
```

(5) 检验配置结果

以上配置完成后，Switch A、Switch B 和 Switch C 之间的 OSPFv3 报文将被加密传输。

1.13 常见配置错误举例

1.13.1 OSPFv3 邻居无法建立

1. 故障现象

OSPFv3 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 OSPFv3 参数，必须保证与相邻路由器的参数一致，区域号相同。

相邻的两台路由器接口的网络类型必须一致。若网络类型为广播网，则至少有一个接口的 DR 优先级应大于零。

3. 处理过程

(1) 使用 **display ospfv3 peer** 命令查看 OSPFv3 邻居状态。

(2) 使用 **display ospfv3 interface** 命令查看 OSPFv3 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 OSPFv3 定时器，在同一接口上失效时间应至少为 Hello 间隔时间的 4 倍。

(5) 如果网络类型为广播网，则至少有一个接口的 DR 优先级大于零。

1.13.2 OSPFv3 路由信息不正确

1. 故障现象

OSPFv3 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 ospfv3 peer** 命令查看 OSPFv3 邻居状态。

(2) 使用 **display ospfv3 interface** 命令查看 OSPFv3 接口的信息。

(3) 使用 **display ospfv3 lsdb** 查看 LSDB 的信息是否完整。

(4) 使用 **display current-configuration configuration** 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。

(6) 如果配置了虚连接，使用 **display ospf vlink** 命令查看邻居状态是否正常。

目 录

1 IPv6 IS-IS配置 ····· 1-1
1.1 IPv6 IS-IS简介 ····· 1-1
1.2 配置IPv6 IS-IS的基本特性 ····· 1-1
1.2.1 配置准备 ····· 1-1
1.2.2 配置IS-IS的IPv6 基本特性 ····· 1-2
1.3 配置IPv6 IS-IS的路由信息控制 ····· 1-2
1.3.1 配置准备 ····· 1-2
1.3.2 配置IPv6 IS-IS的路由信息控制 ····· 1-2
1.4 配置IPv6 IS-IS与BFD联动 ····· 1-3
1.5 配置IS-IS支持IPv6 单播拓扑 ····· 1-4
1.5.1 简介 ····· 1-4
1.5.2 配置IS-IS支持IPv6 单播拓扑 ····· 1-5
1.6 IPv6 IS-IS显示和维护 ····· 1-6
1.7 IPv6 IS-IS典型配置举例 ····· 1-6
1.7.1 IPv6 IS-IS基本配置 ····· 1-6
1.7.2 配置IPv6 IS-IS与BFD联动 ····· 1-11
1.7.3 配置IS-IS支持IPv6 单播拓扑 ····· 1-13

1 IPv6 IS-IS配置

说明

- 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”。
- 本文所指的路由器代表运行了路由协议的三层设备。

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 或 Link State Packet）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。

IS-IS 的相关知识请参见“三层技术-IP 路由配置指导”中的“IS-IS”。

1.2 配置IPv6 IS-IS的基本特性

说明

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：

- 全局使能 IPv6 能力
- 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*]	必选 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*	必选 缺省情况下，没有配置NET
使能IS-IS进程的IPv6能力	**ipv6 enable**	必选 缺省情况下，没有使能IS-IS路由进程的IPv6能力
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*]	必选 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*]	-
配置IPv6 IS-IS路由优先级	**ipv6 preference** { **route-policy** *route-policy-name* | *preference* } *	可选 缺省情况下，IPv6 IS-IS路由优先级为15
配置IPv6 IS-IS聚合路由	**ipv6 summary** *ipv6-prefix prefix-length* [**avoid-feedback** | **generate_null0_route** | [**level-1** | **level-1-2** | **level-2**] | **tag** *tag*] *	可选 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	**ipv6 default-route-advertise** [[**level-1** | **level-1-2** | **level-2**] | **route-policy** *route-policy-name*] *	可选 缺省情况下，不生成IPv6 IS-IS缺省路由

操作	命令	说明
配置IPv6 IS-IS对接收的路由进行过滤	**ipv6 filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* | **route-policy** *route-policy-name* } **import**	可选 缺省情况下，IPv6 IS-IS不对接收的路由信息进行过滤
配置IPv6 IS-IS引入其他协议的路由信息	**ipv6 import-route** *protocol* [*process-id*] [**allow-ibgp**] [**cost** *cost* | [**level-1** | **level-1-2** | **level-2**] | **route-policy** *route-policy-name* | **tag** *tag*] *	可选 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	**ipv6 import-route limit** *number*	可选 本命令的缺省值与设备的型号有关，请以设备的实际情况为准
配置IPv6 IS-IS对引入的路由进行过滤	**ipv6 filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* | **route-policy** *route-policy-name* } **export** [*protocol* [*process-id*]]	可选 缺省情况下，IPv6 IS-IS不对引入的路由进行过滤
使能IPv6 IS-IS路由渗透功能	**ipv6 import-route isisv6 level-2 into level-1** [**filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* | **route-policy** *route-policy-name* } | **tag** *tag*] *	可选 缺省情况下，禁止IPv6 IS-IS路由渗透功能
配置在负载分担方式下IPv6 IS-IS等价路由的最大数量	**ipv6 maximum load-balancing** *number*	可选 缺省值为8

说明

- **ipv6 filter-policy export** 命令一般和 **ipv6 import-route** 命令结合使用，如果没有指定 *protocol* 参数，将对当前路由器从所有协议引入的路由进行过滤；如果指定了 *protocol* 参数，则只对当前路由器从特定协议引入的路由进行过滤。
- ACL 的相关知识请参见“ACL 和 QoS 配置指导”中的“ACL”。
- 路由策略、IPv6 地址前缀列表的相关知识请参见“三层技术-IP 路由配置指导”中的“路由策略”。

1.4 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。

表1-3 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*]	-

操作	命令	说明
配置网络实体名称（NET）	**network-entity** *net*	必选 缺省情况下，没有配置NET
使能IS-IS进程的IPv6能力	**ipv6 enable**	必选 缺省情况下，没有使能IS-IS路由进程的IPv6能力
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*]	必选 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**	必选 缺省情况下，运行IPv6 IS-IS的接口未使能BFD功能

说明

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”。

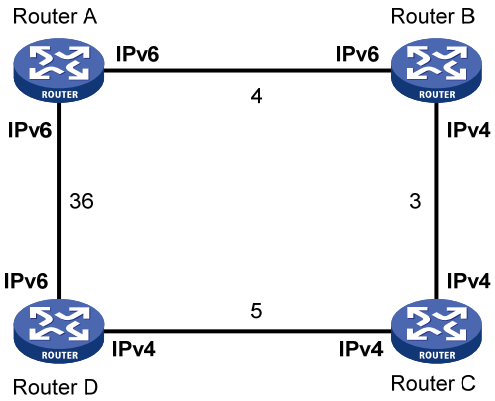
1.5 配置IS-IS支持IPv6单播拓扑

1.5.1 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。

1.5.2 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，需要配置 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1-4 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的Wide Metric属性	**cost-style** { **narrow** | **wide** | **wide-compatible** | { **compatible** | **narrow-compatible** } [**relax-spf-limit**] }	必选 缺省情况下，IS-IS只收发采用narrow方式的报文
配置IS-IS支持IPv6单播拓扑	**multiple-topology ipv6-unicast**	必选 缺省情况下，IS-IS不支持IPv6单播拓扑，即IS-IS不支持IPv4和IPv6分拓扑计算

1.6 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。

在用户视图下执行 **reset** 命令可以清除 IPv6 IS-IS 的数据库信息，或者复位特定邻居的数据信息。

表1-5 IPv6 IS-IS 显示和维护

操作	命令
显示IS-IS的摘要信息	**display isis brief** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS当前debug开关的状态	**display isis debug-switches** { *process-id* | **vpn-instance** *vpn-instance-name* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示使能了IS-IS的接口信息	**display isis interface** [**statistics** | [*interface-type interface-number*] [**verbose**]] [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS的链路状态数据库	**display isis lsdb** [[**l1** | **l2** | **level-1** | **level-2**] | [[**lsp-id** *lsp-id* | **lsp-name** *lspname* | **local**] | **verbose**] *] * [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS的mesh-group	**display isis mesh-group** [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示本地路由器名称到系统ID的映射关系表	**display isis name-table** [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS的邻居信息	**display isis peer** [**statistics | verbose**] [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 IS-IS路由信息	**display isis route ipv6** [[**level-1** | **level-2**] | **verbose**] * [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS的SPF计算日志记录	**display isis spf-log** [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IS-IS进程的统计信息	**display isis statistics** [**level-1** | **level-1-2** | **level-2**] [*process-id* | **vpn-instance** *vpn-instance-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除所有IS-IS的数据结构信息	**reset isis all** [*process-id* | **vpn-instance** *vpn-instance-name*]
清除IS-IS特定邻居的数据信息	**reset isis peer** *system-id* [*process-id* | **vpn** *vpn-instance-name*]

1.7 IPv6 IS-IS典型配置举例

1.7.1 IPv6 IS-IS基本配置

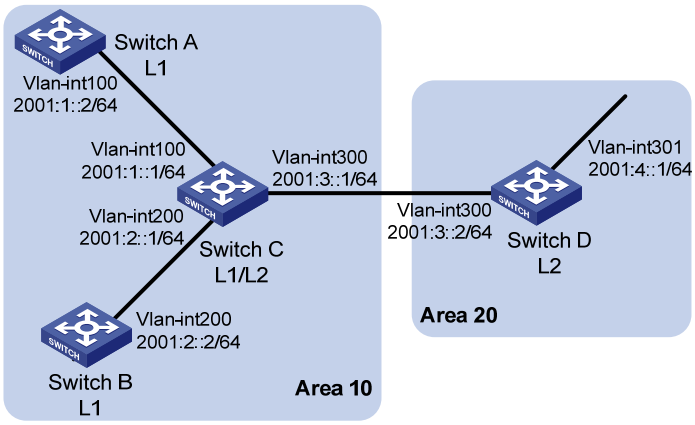
1. 组网需求

如下图所示，Switch A、Switch B、Switch C 和 Switch D 属于同一自治系统，所有交换机已使能了 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 Switch A 和 Switch B 是 Level-1 交换机，Switch D 是 Level-2 交换机，Switch C 是 Level-1-2 交换机。Switch A、Switch B 和 Switch C 属于区域 10，而 Switch D 属于区域 20。

2. 组网图

图1-2 IPv6 IS-IS 基本配置组网图



3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 IPv6 IS-IS

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] isis 1
[SwitchA-isis-1] is-level level-1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] ipv6 enable
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis ipv6 enable 1
[SwitchA-Vlan-interface100] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] isis 1
[SwitchB-isis-1] is-level level-1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] ipv6 enable
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis ipv6 enable 1
[SwitchB-Vlan-interface200] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
```

```
[SwitchC-isis-1] ipv6 enable
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] isis ipv6 enable 1
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis ipv6 enable 1
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis ipv6 enable 1
[SwitchC-Vlan-interface300] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] ipv6
[SwitchD] isis 1
[SwitchD-isis-1] is-level level-2
[SwitchD-isis-1] network-entity 20.0000.0000.0004.00
[SwitchD-isis-1] ipv6 enable
[SwitchD-isis-1] quit
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] isis ipv6 enable 1
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 301
[SwitchD-Vlan-interface301] isis ipv6 enable 1
[SwitchD-Vlan-interface301] quit
```

(3) 检查配置结果

查看 Switch A 的 IPv6 IS-IS 路由表。

```
[SwitchA] display isis route ipv6

                      Route information for ISIS(1)
                      -----------------------------

                   ISIS(1) IPv6 Level-1 Forwarding Table
                   -------------------------------------
 Destination: ::                                    PrefixLen: 0
 Flag       : R/-/-                                 Cost     : 10
 Next Hop   : FE80::200:FF:FE0F:4                   Interface: Vlan100

 Destination: 2001:1::                              PrefixLen: 64
 Flag       : D/L/-                                 Cost     : 10
 Next Hop   : Direct                                Interface: Vlan100

 Destination: 2001:2::                              PrefixLen: 64
 Flag       : R/-/-                                 Cost     : 20
 Next Hop   : FE80::200:FF:FE0F:4                   Interface: Vlan100

 Destination: 2001:3::                              PrefixLen: 64
 Flag       : R/-/-                                 Cost     : 20
```

```
 Next Hop   : FE80::200:FF:FE0F:4                    Interface: Vlan100

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

查看 Switch B 的 IPv6 IS-IS 路由表。

```
[SwitchB] display isis route ipv6

                        Route information for ISIS(1)
                        ------------------------------

                    ISIS(1) IPv6 Level-1 Forwarding Table
                    -------------------------------------
 Destination: ::                                    PrefixLen: 0
 Flag       : R/-/-                                 Cost     : 10
 Next Hop   : FE80::200:FF:FE0F:4                   Interface: Vlan200

 Destination: 2001:1::                              PrefixLen: 64
 Flag       : D/L/-                                 Cost     : 10
 Next Hop   : FE80::200:FF:FE0F:4                   Interface: Vlan200

 Destination: 2001:2::                              PrefixLen: 64
 Flag       : R/-/-                                 Cost     : 20
 Next Hop   : Direct                                Interface: Vlan200

 Destination: 2001:3::                              PrefixLen: 64
 Flag       : R/-/-                                 Cost     : 20
 Next Hop   : FE80::200:FF:FE0F:4                   Interface: Vlan200

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

查看 Switch C 的 IPv6 IS-IS 路由表。

```
[SwitchC] display isis route ipv6

                        Route information for ISIS(1)
                        ------------------------------

                    ISIS(1) IPv6 Level-1 Forwarding Table
                    -------------------------------------
 Destination: 2001:1::                              PrefixLen: 64
 Flag       : D/L/-                                 Cost     : 10
 Next Hop   : Direct                                Interface: Vlan100

 Destination: 2001:2::                              PrefixLen: 64
 Flag       : D/L/-                                 Cost     : 10
 Next Hop   : Direct                                Interface: Vlan200

 Destination: 2001:3::                              PrefixLen: 64
 Flag       : D/L/-                                 Cost     : 10
 Next Hop   : Direct                                Interface: Vlan300
```

```
      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set


                     ISIS(1) IPv6 Level-2 Forwarding Table
                     -------------------------------------
 Destination: 2001:1::                              PrefixLen: 64
 Flag       : D/L/-                                 Cost     : 10
 Next Hop   : Direct                                Interface: Vlan100

 Destination: 2001:2::                              PrefixLen: 64
 Flag       : D/L/-                                 Cost     : 10
 Next Hop   : Direct                                Interface: Vlan200

 Destination: 2001:3::                              PrefixLen: 64
 Flag       : D/L/-                                 Cost     : 10
 Next Hop   : Direct                                Interface: Vlan300

 Destination: 2001:4::1                             PrefixLen: 128
 Flag       : R/-/-                                 Cost     : 10
 Next Hop   : FE80::20F:E2FF:FE3E:FA3D              Interface: Vlan300

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

查看 Switch D 的 IPv6 IS-IS 路由表。

```
[SwitchD] display isis route ipv6

                         Route information for ISIS(1)
                         -----------------------------

                     ISIS(1) IPv6 Level-2 Forwarding Table
                     -------------------------------------
 Destination: 2001:1::                              PrefixLen: 64
 Flag       : R/-/-                                 Cost     : 20
 Next Hop   : FE80::200:FF:FE0F:4                   Interface: Vlan300

Destination: 2001:2::                              PrefixLen: 64
 Flag       : R/-/-                                 Cost     : 20
 Next Hop   : FE80::200:FF:FE0F:4                   Interface: Vlan300

 Destination: 2001:3::                              PrefixLen: 64
 Flag       : D/L/-                                 Cost     : 10
 Next Hop   : Direct                                Interface: Vlan300

 Destination: 2001:4::1                             PrefixLen: 128
 Flag       : D/L/-                                 Cost     : 0
 Next Hop   : Direct                                Interface: Loop1

      Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

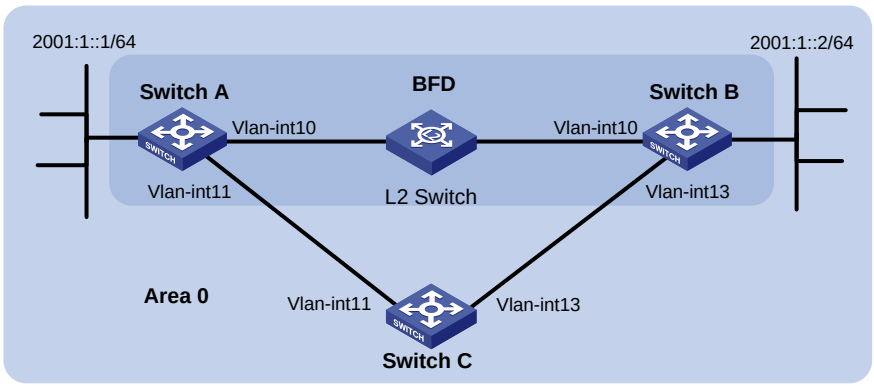
1.7.2 配置IPv6 IS-IS与BFD联动

1. 组网需求

- Switch A、Switch 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 Switch B 和二层交换机之间的链路出现故障后，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
Switch A	Vlan-int10	2001::1/64	Switch B	Vlan-int10	2001::2/64
	Vlan-int11	2001:2::1/64		Vlan-int13	2001:3::2/64
Switch C	Vlan-int11	2001:2::2/64			
	Vlan-int13	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 Switch A。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] isis 1
[SwitchA-isis-1] is-level level-1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] ipv6 enable
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] isis ipv6 enable 1
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] isis ipv6 enable 1
[SwitchA-Vlan-interface11] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ipv6
```

```
[SwitchB] isis 1
[SwitchB-isis-1] is-level level-1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] ipv6 enable
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] isis ipv6 enable 1
[SwitchB-Vlan-interface10] quit
[SwitchB] interface vlan-interface 13
[SwitchB-Vlan-interface13] isis ipv6 enable 1
[SwitchB-Vlan-interface13] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] ipv6 enable
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 11
[SwitchC-Vlan-interface11] isis ipv6 enable 1
[SwitchC-Vlan-interface11] quit
[SwitchC] interface vlan-interface 13
[SwitchC-Vlan-interface13] isis ipv6 enable 1
[SwitchC-Vlan-interface13] quit
```

(3) 配置 BFD 功能

在 Switch A 上使能 IPv6 IS-IS BFD 功能，并配置 BFD 参数。

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] isis ipv6 bfd enable
[SwitchA-Vlan-interface10] bfd min-transmit-interval 500
[SwitchA-Vlan-interface10] bfd min-receive-interval 500
[SwitchA-Vlan-interface10] bfd detect-multiplier 7
[SwitchA-Vlan-interface10] return
```

在 Switch B 上使能 IPv6 IS-IS BFD 功能，并配置 BFD 参数。

```
[SwitchB] bfd session init-mode active
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] isis ipv6 bfd enable
[SwitchB-Vlan-interface10] bfd min-transmit-interval 500
[SwitchB-Vlan-interface10] bfd min-receive-interval 500
[SwitchB-Vlan-interface10] bfd detect-multiplier 6
```

(4) 检查配置结果

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

显示 Switch A 的 BFD 信息。

```
<SwitchA> display bfd session
Total 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（Switch A 接口 Vlan-interface10 的链路本地地址）
  Destination IP: FE80::20F:FF:FE00:1200（Switch B 接口 Vlan-interface10 的链路本地地址）
   Session State: Up                     Interface: Vlan10
       Hold Time:    /
```

在 Switch A 上查看 2001:4::0/64 的路由信息，可以看出 Switch A 和 Switch B 是通过 L2 Switch 进行通信的。

```
<SwitchA> display ipv6 routing-table 2001:4::0 64 verbose
Routing Table :
Summary Count : 1

 Destination  : 2001:4::0                              PrefixLength : 64
 NextHop      : 2001::2                                Preference   : 15
 IpPrecedence :                                        QosLcId      :
 RelayNextHop : ::                                     Tag          : 0H
 Neighbor     : ::                                     ProcessID    : 0
 Interface    : Vlan-interface10                       Protocol     : ISISv6
 State        : Active Adv                             Cost         : 20
 Tunnel ID    : 0x0                                    Label        : NULL
 Age          : 4538sec
```

当 Switch B 和二层交换机之间的链路发生故障时 Switch A 能够快速检测 Switch B 的变化。

查看 Switch A 的 BFD 信息，Switch A 已取消与 Switch B 的邻居关系，没有任何输出信息。

```
<SwitchA> display bfd session
```

在 Switch A 上查看 2001:4::0/64 的路由信息，可以看出 Switch A 和 Switch B 已经切换到 Switch C 进行通信。

```
<SwitchA> display ipv6 routing-table 2001:4::0 64 verbose
Routing Table :
Summary Count : 1

 Destination  : 2001:4::0                              PrefixLength : 64
 NextHop      : 2001:2::2                              Preference   : 15
 IpPrecedence :                                        QosLcId      :
 RelayNextHop : ::                                     Tag          : 0H
 Neighbor     : ::                                     ProcessID    : 0
 Interface    : Vlan-interface11                       Protocol     : ISISv6
 State        : Active  Adv                            Cost         : 30
 Tunnel ID    : 0x0                                    Label        : NULL
 Age          : 4610sec
```

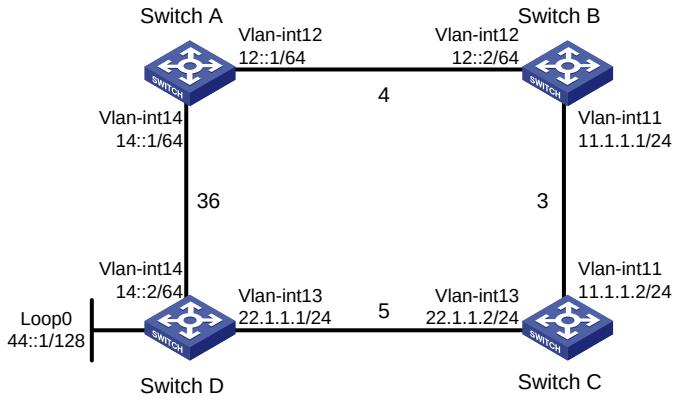
1.7.3 配置IS-IS支持IPv6 单播拓扑

1. 组网需求

如 图 1-4 所示，配置IS-IS支持IPv6 单播拓扑，即将IPv4 拓扑和IPv6 拓扑分开计算。

2. 组网图

图1-4 IS-IS 支持 IPv6 单播拓扑配置组网图



3. 配置步骤

(1) 配置各交换机接口的 IPv4/IPv6 地址和 IS-IS 协议

请按照 图 1-4 配置各接口的IPv4/IPv6 地址和子网掩码，具体配置过程略。

配置各交换机之间采用 IS-IS 协议进行互连，确保 Switch A、Switch B、Switch C 和 Switch D 之间能够在网络层互通，并且各交换机之间能够借助 IS-IS 协议实现动态路由更新。

请按照 图 1-4 配置各接口的链路开销值，过程略。

(2) 配置 IS-IS 支持 IPv6 单播拓扑

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] isis
[SwitchA-isis-1] cost-style wide
[SwitchA-isis-1] multiple-topology ipv6-unicast
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] isis
[SwitchB-isis-1] cost-style wide
[SwitchB-isis-1] multiple-topology ipv6-unicast
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] isis
[SwitchD-isis-1] cost-style wide
[SwitchD-isis-1] multiple-topology ipv6-unicast
```

(3) 检验配置效果

在 Switch A 上查看拓扑 IPv6 的 IS-IS 路由表。

```
[SwitchA-isis-1] display isis route ipv6

                        Route information for ISIS(1)
                        ------------------------------
```

```
                  ISIS(1) IPv6 Level-1 Forwarding Table
                  -------------------------------------
Destination: 12::                                 PrefixLen: 64
Flag       : D/L/-                                Cost     : 4
Next Hop   : Direct                               Interface: Vlan12

Destination: 44::1                                PrefixLen: 128
Flag       : R/L/-                                Cost     : 36
Next Hop   : FE80::200:5EFF:FE00:F11              Interface: Vlan14

Destination: 14::                                 PrefixLen: 64
Flag       : D/L/-                                Cost     : 36
Next Hop   : Direct                               Interface: Vlan14

     Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set


                  ISIS(1) IPv6 Level-2 Forwarding Table
                  -------------------------------------
Destination: 12::                                 PrefixLen: 64
Flag       : D/L/-                                Cost     : 4
Next Hop   : Direct                               Interface: Vlan12

Destination: 44::1                                PrefixLen: 128
Flag       : -/-/-                                Cost     : 36


Destination: 14::                                 PrefixLen: 64
Flag       : D/L/-                                Cost     : 36
Next Hop   : Direct                               Interface: Vlan14

     Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

Switch A 上查看拓扑到 Switch D 的 IPv6 路由 44::1/128 的下一跳为接口 Vlan-interface14。

目 录

1 IPv6 BGP配置 ····· 1-1

1.1 IPv6 BGP简介 ····· 1-1

1.2 配置任务简介 ····· 1-1

1.3 配置IPv6 BGP的基本功能 ····· 1-2

1.3.1 配置准备 ····· 1-2

1.3.2 配置IPv6 对等体 ····· 1-3

1.3.3 配置IPv6 BGP发布本地IPv6 路由 ····· 1-3

1.3.4 配置路由首选值 ····· 1-3

1.3.5 配置IPv6 BGP连接所使用的源接口 ····· 1-4

1.3.6 配置非直接相连的邻居建立EBGP连接 ····· 1-5

1.3.7 配置IPv6 对等体/IPv6 对等体组的描述信息 ····· 1-5

1.3.8 禁止与IPv6 对等体/IPv6 对等体组建立会话 ····· 1-6

1.3.9 记录指定IPv6 对等体/IPv6 对等体组的会话状态和事件信息 ····· 1-6

1.4 控制路由信息的发布与接收 ····· 1-6

1.4.1 配置准备 ····· 1-7

1.4.2 配置IPv6 BGP引入其他路由 ····· 1-7

1.4.3 配置IPv6 BGP路由聚合 ····· 1-7

1.4.4 配置向IPv6 对等体/IPv6 对等体组发送缺省路由 ····· 1-8

1.4.5 配置路由信息的发布策略 ····· 1-8

1.4.6 配置路由信息的接收策略 ····· 1-9

1.4.7 配置IPv6 BGP与IGP路由同步 ····· 1-10

1.4.8 配置路由衰减 ····· 1-10

1.5 配置IPv6 BGP的路由属性 ····· 1-11

1.5.1 配置准备 ····· 1-11

1.5.2 配置IPv6 BGP路由管理的优先级、缺省LOCAL_PREF及NEXT_HOP属性 ····· 1-11

1.5.3 配置MED属性 ····· 1-12

1.5.4 配置AS_PATH属性 ····· 1-12

1.6 调整和优化IPv6 BGP网络 ····· 1-13

1.6.1 配置准备 ····· 1-13

1.6.2 配置IPv6 BGP的时钟 ····· 1-14

1.6.3 配置IPv6 BGP软复位 ····· 1-14

1.6.4 使能IPv6 BGP ORF能力功能 ····· 1-15

1.6.5 使能 4 字节AS号抑制功能 ····· 1-16

1.6.6 配置IPv6 BGP报文的DSCP优先级 ····· 1-17
1.6.7 配置最大等价路由的条数 ····· 1-17
1.6.8 配置IPv6 BGP建立TCP连接时进行MD5 认证 ····· 1-17
1.6.9 配置IPv6 BGP IPsec安全策略 ····· 1-18
1.7 组建大型IPv6 BGP网络 ····· 1-19
1.7.1 配置准备 ····· 1-19
1.7.2 配置IPv6 BGP对等体组 ····· 1-19
1.7.3 配置IPv6 BGP团体 ····· 1-21
1.7.4 配置IPv6 BGP路由反射器 ····· 1-22
1.8 配置IPv6 BGP与BFD联动 ····· 1-22
1.9 IPv6 BGP显示和维护 ····· 1-23
1.9.1 IPv6 BGP显示 ····· 1-23
1.9.2 复位IPv6 BGP连接 ····· 1-24
1.9.3 清除IPv6 BGP信息 ····· 1-24
1.10 IPv6 BGP典型配置举例 ····· 1-25
1.10.1 IPv6 BGP基本配置 ····· 1-25
1.10.2 配置IPv6 BGP路由反射 ····· 1-27
1.10.3 配置IPv6 BGP IPsec安全策略 ····· 1-28
1.10.4 配置IPv6 BGP与BFD联动 ····· 1-33
1.11 IPv6 BGP常见错误配置举例 ····· 1-37
1.11.1 IPv6 BGP对等体关系不能建立 ····· 1-37

1 IPv6 BGP配置

说明

- 在以下路由协议的介绍中所指的路由器及路由器图标，代表了一般意义下的路由器以及运行了路由协议的以太网交换机。
- 本章只列出了 IPv6 BGP 专有的配置与操作，其他相关内容请参见“三层技术-IP 路由配置指导”中的 “BGP”。

1.1 IPv6 BGP简介

传统的 BGP-4 只能管理 IPv4 的路由信息，对于使用其它网络层协议（如 IPv6 等）的应用，在跨自治系统传播时就受到一定限制。

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 IPv6 BGP，目前的 IPv6 BGP 标准是 RFC 2858（Multiprotocol Extensions for BGP-4，BGP-4 多协议扩展）。

为了实现对 IPv6 协议的支持，IPv6 BGP 需要将 IPv6 网络层协议的信息反映到 NLRI（Network Layer Reachability Information，网络层可达性信息）及 NEXT_HOP 属性中。

IPv6 BGP 中引入的两个 NLRI 属性分别是：

- MP_REACH_NLRI：Multiprotocol Reachable NLRI，多协议可达 NLRI。用于发布可达路由及下一跳信息。
- MP_UNREACH_NLRI：Multiprotocol Unreachable NLRI，多协议不可达 NLRI。用于撤销不可达路由。

IPv6 BGP 中的 NEXT_HOP 属性用 IPv6 地址来表示，可以是 IPv6 全球单播地址或者链路本地地址。

IPv6 BGP 是利用 BGP 的多协议扩展属性，来达到在 IPv6 网络中应用的目的，BGP 协议原有的消息机制和路由机制并没有改变。

1.2 配置任务简介

表1-1 配置任务简介

配置任务		说明	详细配置
配置IPv6 BGP基本功能	配置IPv6对等体	必选	1.3.2
	配置IPv6 BGP发布本地IPv6路由	可选	1.3.3
	配置路由首选值	可选	1.3.4
	配置IPv6 BGP连接所使用的本地接口	可选	1.3.5
	配置EBGP连接的最大跳数	可选	1.3.6
	配置IPv6对等体/IPv6对等体组的描述信息	可选	1.3.7

配置任务		说明	详细配置
	禁止与IPv6对等体/IPv6对等体组建立会话	可选	1.3.8
	记录指定IPv6对等体/IPv6对等体组的会话状态和事件信息	可选	1.3.9
控制路由信息的发布与接收	配置IPv6 BGP引入其他路由	可选	1.4.2
	配置IPv6 BGP路由聚合	可选	1.4.3
	配置向IPv6对等体/IPv6对等体组发送缺省路由	可选	1.4.4
	配置路由信息的发布策略	可选	1.4.5
	配置路由信息的接收策略	可选	1.4.6
	配置IPv6 BGP与IGP路由同步	可选	1.4.7
	配置路由衰减	可选	1.4.8
配置IPv6 BGP的路由属性	配置IPv6 BGP路由管理的优先级、缺省LOCAL_PREF及NEXT_HOP属性	可选	1.5.2
	配置MED属性	可选	1.5.3
	配置AS_PATH属性	可选	1.5.4
调整和优化IPv6 BGP网络	配置IPv6 BGP的时钟	可选	1.6.2
	配置IPv6 BGP软复位	可选	1.6.3
	配置BGP ORF能力	可选	1.6.4
	使能4字节AS号抑制功能	可选	1.6.5
	配置IPv6 BGP报文的DSCP优先级	可选	1.6.6
	配置最大等价路由的条数	可选	1.6.7
	配置IPv6 BGP建立TCP连接时进行MD5认证	可选	1.6.8
	配置IPv6 BGP IPsec安全策略	可选	1.6.9
组建大型IPv6 BGP网络	配置IPv6 BGP对等体组	可选	1.7.2
	配置IPv6 BGP团体	可选	1.7.3
	配置IPv6 BGP路由反射器	可选	1.7.4
配置IPv6 BGP与BFD联动		可选	1.8

1.3 配置IPv6 BGP的基本功能

1.3.1 配置准备

在配置 IPv6 BGP 基本功能之前，需完成以下任务：

- 配置接口的网络层地址
- 在系统视图下通过 **ipv6** 命令使能 IPv6 报文转发功能

说明

必须首先创建对等体组，才能配置对等体组的基本功能。关于对等体组的创建请参见“1.7.2 配置IPv6 BGP对等体组”。

1.3.2 配置IPv6 对等体

表1-2 配置 IPv6 对等体

操作	命令	说明
进入系统视图	**system-view**	-
启动BGP，进入BGP视图	**bgp** *as-number*	必选 缺省情况下，系统没有运行BGP
为路由器指定ID	**router-id** *router-id*	可选 如果Loopback和其它接口没有配置IP地址，则该任务为必选
进入IPv6地址族视图	**ipv6-family**	-
配置对等体	**peer** *ipv6-address* **as-number** *as-number*	必选

1.3.3 配置IPv6 BGP发布本地IPv6 路由

表1-3 配置 IPv6 BGP 发布本地 IPv6 路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
将路由通告到IPv6 BGP路由表中	**network** *ipv6-address prefix-length* [**short-cut** | **route-policy** *route-policy-name*]	必选 缺省情况下，IPv6 BGP不通告任何路由

1.3.4 配置路由首选值

表1-4 配置路由首选值

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选

操作	命令	说明
进入IPv6地址族视图	**ipv6-family**	-
为从IPv6对等体/IPv6对等体组接收的路由分配首选值	**peer** { *ipv6-group-name* | *ipv6-address* } **preferred-value** *value*	可选 缺省情况下，从IPv6对等体/IPv6对等体组接收的路由的首选值为0

注意

当通过路由策略过滤 IPv6 路由时，路由的首选值将优先选取路由策略中设置的首选值。只有当路由策略里设置的首选值为 0 时，才选取 **peer** { *ipv6-group-name* | *ipv6-address* } **preferred-value** *value* 命令里设置的值。通过路由策略配置 BGP IPv6 路由信息首选值的相关配置可参考命令 **peer** { *group-name* | *ipv4-address* | *ipv6-address* } **route-policy** *route-policy-name* { **import** | **export** } 和“三层技术-IP 路由命令参考/路由策略”中的 **apply preferred-value** *preferred-value*。

1.3.5 配置IPv6 BGP连接所使用的源接口

IPv6 BGP 使用 TCP 作为其传输层协议，缺省情况下，IPv6 BGP 使用到达对等体最佳路由的出接口作为与对等体/对等体组建立 TCP 连接的源接口。

当建立 IPv6 BGP 连接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 连接的源接口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，建议网络管理员使用 Loopback 接口建立 IPv6 BGP 连接，并将建立 IPv6 BGP 连接所使用的源接口配置为 Loopback 接口，来提高 IPv6 BGP 连接的可靠性和稳定性。

表1-5 配置 IPv6 BGP 连接所使用的源接口

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
指定与IPv6对等体/IPv6对等体组创建BGP会话时建立TCP连接使用的源接口	**peer** { *ipv6-group-name* | *ipv6-address* } **connect-interface** *interface-type interface-number*	必选 缺省情况下，IPv6 BGP使用到达IPv6对等体的最佳路由的出接口作为与IPv6对等体/IPv6对等体组创建BGP会话时建立TCP连接的源接口

说明

当建立 BGP 连接时，如果没有明确指定建立 TCP 连接的源接口，可能会由于无法根据到达 BGP 对等体的最优路由确定 TCP 连接源接口从而导致无法建立 TCP 连接，因此建议用户在此情况下配置 BGP 对等体时明确配置 BGP 会话建立 TCP 连接的源接口为指定接口。

1.3.6 配置非直接相连的邻居建立EBGP连接

表1-6 配置非直接相连的邻居建立 EBGP 连接

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置允许同非直接相连网络上的邻居建立EBGP连接	**peer** { *ipv6-group-name* | *ipv6-address* } **ebgp-max-hop** [*hop-count*]	必选 缺省情况下，不允许同非直接相连网络上的邻居建立EBGP连接

注意

通常情况下，EBGP 对等体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** 命令允许它们之间经过多跳建立 TCP 连接。但是，对于直连 EBGP 使用 Loopback 接口建立邻居关系，不需要 **peer ebgp-max-hop** 命令的配置。

1.3.7 配置IPv6 对等体/IPv6 对等体组的描述信息

表1-7 配置 IPv6 对等体/IPv6 对等体组的描述信息

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置IPv6对等体/IPv6对等体组的描述信息	**peer** { *ipv6-group-name* | *ipv6-address* } **description** *description-text*	可选 缺省情况下，IPv6对等体/IPv6对等体组无描述信息

说明

如果配置对等体组的描述信息，需要先创建对等体组。

1.3.8 禁止与IPv6 对等体/IPv6 对等体组建立会话

表1-8 禁止与 IPv6 对等体/IPv6 对等体组建立会话

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
禁止与IPv6对等体/IPv6对等体组建立会话	**peer** { *ipv6-group-name* | *ipv6-address* } **ignore**	必选 缺省情况下，允许与IPv6对等体/IPv6对等体组建立会话

1.3.9 记录指定IPv6 对等体/IPv6 对等体组的会话状态和事件信息

表1-9 记录指定 IPv6 对等体/IPv6 对等体组的会话状态和事件信息

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
全局使能BGP日志记录功能	**log-peer-change**	可选 缺省情况下，使能BGP日志记录功能
进入IPv6地址族视图	**ipv6-family**	-
记录指定IPv6对等体/IPv6对等体组的会话状态和事件信息	**peer** { *ipv6-group-name* | *ipv6-address* } **log-change**	可选 缺省情况下，记录IPv6对等体/IPv6对等体组的会话状态和事件信息

说明

有关 **log-peer-change** 命令的详细信息请参见“三层技术-IP 路由命令参考”中的“BGP”。

1.4 控制路由信息的发布与接收

控制 IPv6 BGP 的路由信息的发布和接收，包括对路由信息进行过滤，应用路由策略，路由衰减。

1.4.1 配置准备

在控制路由信息的发布与接收之前，需完成以下任务：

- 使能 IPv6 功能
- 配置 IPv6 BGP 基本功能

1.4.2 配置IPv6 BGP引入其他路由

表1-10 配置 IPv6 BGP 引入和过滤外部路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
允许将缺省路由引入到IPv6 BGP路由表中	**default-route imported**	可选 缺省情况下，IPv6 BGP不允许将缺省路由引入到IPv6 BGP路由表中
配置引入其它协议路由信息并通告	**import-route** *protocol* [*process-id* [**med** *med-value* | **route-policy** *route-policy-name*] *]	必选 缺省情况下，IPv6 BGP不引入且不通告其它协议的路由

说明

如果没有配置 **default-route imported** 命令，则使用 **import-route** 命令引入 IGP 路由时，不能引入 IGP 的缺省路由。

1.4.3 配置IPv6 BGP路由聚合

在中型或大型 BGP 网络中，在向对等体发布 IPv6 路由信息时，可以配置路由聚合，减小对等体路由表中的路由数量。BGP 仅支持 IPv6 路由的手动聚合。

表1-11 配置 IPv6 路由聚合

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6地址族视图	**ipv6-family**	-
配置手动路由聚合	**aggregate** *ipv6-address prefix-length* [**as-set** | **attribute-policy** *route-policy-name* | **detail-suppressed** | **origin-policy** *route-policy-name* | **suppress-policy** *route-policy-name*] *	必选 缺省情况下，不进行路由聚合

1.4.4 配置向IPv6 对等体/IPv6 对等体组发送缺省路由

表1-12 配置向 IPv6 对等体/IPv6 对等体组发送缺省路由

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置向IPv6对等体/IPv6对等体组发送缺省路由	**peer** { *ipv6-group-name* | *ipv6-address* } **default-route-advertise** [**route-policy** *route-policy-name*]	必选 缺省情况下，不向IPv6对等体/IPv6对等体组发布缺省路由

说明

执行 **peer default-route-advertise** 命令后，不论本地路由表中是否存在缺省路由，都将向指定 IPv6 对等体/IPv6 对等体组发布一条下一跳地址为本地地址的缺省路由。

1.4.5 配置路由信息的发布策略

表1-13 配置路由信息的发布策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
对发布的路由进行过滤	**filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* } **export** [*protocol process-id*]	必选 缺省情况下，没有对发布的路由信息进行过滤
配置出方向路由策略	**peer** { *ipv6-group-name* | *ipv6-address* } **route-policy** *route-policy-name* **export**	必选 缺省情况下，没有指定IPv6对等体/IPv6对等体组的路由策略
配置基于IPv6 ACL的路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **filter-policy** *acl6-number* **export**	必选 缺省情况下，没有为IPv6对等体/IPv6对等体组配置基于IPv6 ACL的过滤策略
配置基于AS路径过滤列表的IPv6 BGP路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **as-path-acl** *as-path-acl-number* **export**	必选 缺省情况下，没有为IPv6对等体/IPv6对等体组配置基于AS路径过滤列表的IPv6 BGP路由过滤策略

操作	命令	说明
配置基于IPv6地址前缀列表的路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **ipv6-prefix** *ipv6-prefix-name* **export**	必选 缺省情况下，没有指定IPv6对等体/IPv6对等体组基于IPv6前缀列表的过滤策略

说明

IPv6 BGP 对引入的路由信息进行过滤后，会将符合条件的路由信息发布给 IPv6 BGP 对等体（Export）。指定 *protocol* 参数可以只对特定路由协议的信息进行过滤，如果没有指定此参数，则对所有要发布的 IPv6 BGP 路由信息进行过滤，包括引入的路由和使用 **network** 命令发布的路由。

1.4.6 配置路由信息的接收策略

表1-14 配置路由信息的接收策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
对接收的路由信息进行过滤	**filter-policy** { *acl6-number* | **ipv6-prefix** *ipv6-prefix-name* } **import**	必选 缺省情况下，不对接收的路由信息进行过滤
对接收的路由信息应用路由策略	**peer** { *ipv6-group-name* | *ipv6-address* } **route-policy** *route-policy-name* **import**	必选 缺省情况下，没有指定IPv6对等体/IPv6对等体组的路由策略
配置基于IPv6 ACL的路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **filter-policy** *acl6-number* **import**	必选 缺省情况下，没有为IPv6对等体/IPv6对等体组设置基于IPv6 ACL的路由过滤策略
配置基于AS路径过滤列表的IPv6 BGP路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **as-path-acl** *as-path-acl-number* **import**	必选 缺省情况下，没有为IPv6对等体/IPv6对等体组设置基于AS路径过滤列表的IPv6 BGP路由过滤策略
配置基于IPv6地址前缀列表的路由过滤策略	**peer** { *ipv6-group-name* | *ipv6-address* } **ipv6-prefix** *ipv6-prefix-name* **import**	必选 缺省情况下，没有为IPv6对等体/IPv6对等体组设置基于IPv6前缀列表的路由过滤策略
配置允许从IPv6对等体/IPv6对等体组接收的最大IPv6地址前缀数	**peer** { *ipv6-group-name* | *ipv6-address* } **route-limit** *limit* [*percentage*]	可选 缺省情况下，允许从IPv6对等体/IPv6对等体组接收的IPv6地址最大地址前缀数无限制

说明

- 对 IPv6 BGP 接收的路由进行过滤，只有满足某些条件的路由才能被 IPv6 BGP 接收，并加到路由表中。
- 对等体组的成员可以与所在的组使用不同的入方向路由策略，即接收路由时，各对等体可以选择自己的策略。

1.4.7 配置IPv6 BGP与IGP路由同步

IPv6 BGP 路由器收到一条 IBGP 路由，缺省只检查该路由的下一跳是否可达。如果可达，IPv6 BGP 路由器就将这条 IBGP 路由发布给 EBGP 对等体。如果配置了同步特性，则需要同时满足以下条件，IPv6 BGP 路由器才会将该 IBGP 路由发布给 EBGP 对等体：

- 该路由的下一跳可达；
- 在 IGP 路由表中存在一条 active 的路由，与该 IBGP 路由的目的网段完全相同。（IGP 路由是否为 active 状态可以通过 **display ipv6 routing-table protocol** 命令来查看）

表1-15 配置 IPv6 BGP 与 IGP 路由同步

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置IPv6 BGP与IGP路由同步	**synchronization**	必选 缺省情况下，IPv6 BGP和IGP路由不同步

1.4.8 配置路由衰减

表1-16 配置路由衰减

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置IPv6 BGP路由衰减	**dampening** [*half-life-reachable half-life-unreachable reuse suppress ceiling* | **route-policy** *route-policy-name*]*	可选 缺省情况下，没有配置路由衰减

1.5 配置IPv6 BGP的路由属性

本节主要介绍使用各类路由属性来改变 IPv6 BGP 的选路策略。包括如下属性：

- IPv6 BGP 协议的优先级
- 缺省 LOCAL_PREF 属性值
- MED 属性
- NEXT_HOP 属性
- AS_PATH 属性

1.5.1 配置准备

在配置 IPv6 BGP 的路由属性之前，需完成以下任务：

- 使能 IPv6 功能
- 配置 IPv6 BGP 基本功能

1.5.2 配置IPv6 BGP路由管理的优先级、缺省LOCAL_PREF及NEXT_HOP属性

表1-17 配置 IPv6 BGP 协议的优先级、缺省 LOCAL_PREF 及 NEXT_HOP 属性

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
设定IPv6 BGP路由管理的优先级	**preference** { *external-preference internal-preference local-preference* | **route-policy** *route-policy-name* }	可选 缺省情况下，EBGP路由的管理优先级为255，IBGP路由的管理优先级为255，本地产生的IPv6 BGP路由的管理优先级为130
配置本地优先级的缺省值	**default local-preference** *value*	可选 缺省情况下，本地优先级的缺省值为100
配置发布路由时将自身地址作为下一跳	**peer** { *ipv6-group-name* | *ipv6-address* } **next-hop-local**	可选 缺省情况下，向EBGP IPv6对等体/IPv6对等体组发布路由时，将自身地址作为下一跳；向IBGP IPv6对等体/IPv6对等体发布路由时，下一跳不变

说明

- 在某些组网环境中，为保证 IBGP 邻居能够找到正确的下一跳，可以配置在向 IBGP IPv6 对等体/IPv6 对等体组发布路由时，改变下一跳地址为自身地址。如果配置了 IPv6 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 IPv6 对等体/IPv6 对等体组发布路由时都先将下一跳地址改变为自身地址。

- 在第三方下一跳(即两个 BGP 连接在同一网段的广播网)这种特殊的组网环境中，缺省情况下，向 EBGP IPv6 对等体/IPv6 对等体组发布路由时，下一跳不变；只有配置了 **peer next-hop-local** 命令，才将自身地址作为下一跳。

1.5.3 配置MED属性

表1-18 配置 MED 属性

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置系统的缺省MED值	**default med** *med-value*	可选 缺省情况下，系统缺省的MED值为0
配置允许比较来自不同AS邻居的路由的MED属性值	**compare-different-as-med**	可选 缺省情况下，不比较来自不同AS邻居的路由的MED属性值
配置根据路由来自的AS进行分组对MED排序优选	**bestroute compare-med**	可选 缺省情况下，不根据路由来自的AS进行分组对MED排序优选
配置允许比较联盟对等体的路由按MED值进行优选	**bestroute med-confederation**	可选 缺省情况下，比较联盟对等体的路由时不考虑MED值

1.5.4 配置AS_PATH属性

表1-19 配置 AS_PATH 属性

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置允许本地AS号在所接收的路由的AS_PATH属性中出现，并可同时配置允许重复的次数	**peer** { *ipv6-group-name* | *ipv6-address* } **allow-as-loop** [*number*]	可选 缺省情况下，不允许本地AS号重复

操作	命令	说明
为IPv6对等体/IPv6对等体组配置一个虚拟的本地自治系统号	**peer** { *ipv6-group-name* | *ipv6-address* } **fake-as** *as-number*	可选 缺省情况下，没有为IPv6对等体/IPv6对等体组配置虚拟的本地自治系统号
禁止路由器将AS_PATH当作选路算法中的一个因素	**bestroute as-path-neglect**	可选 缺省情况下，路由器可以将AS_PATH当作选路算法中的一个因素
配置发送IPv6 BGP更新报文时不携带私有AS编号	**peer** { *ipv6-group-name* | *ipv6-address* } **public-as-only**	可选 缺省情况下，发送IPv6 BGP更新报文时携带私有自治系统号
配置用本地AS号替换AS_PATH属性中指定IPv6对等体/IPv6对等体组的AS号	**peer** { *ipv6-group-name* | *ipv6-address* } **substitute-as**	可选 缺省情况下，没有用本地AS号替换AS_PATH属性中指定IPv6对等体/IPv6对等体组的AS号

1.6 调整和优化IPv6 BGP网络

本节主要介绍配置 IPv6 BGP 时钟、软复位 IPv6 BGP 连接和配置最大等价路由的条数。

(1) IPv6 BGP 时钟

当对等体间建立了 IPv6 BGP 连接后，它们定时向对端发送 Keepalive 消息，以防止路由器认为 IPv6 BGP 连接已中断。若路由器在设定的连接保持时间（Holdtime）内未收到对端的 Keepalive 消息或任何其它类型的报文，则认为此 IPv6 BGP 连接已中断，从而退出此 IPv6 BGP 连接。

路由器在与对等体建立 IPv6 BGP 连接时，将比较双方保持时间，以数值较小者做为协商后的保持时间。如果协商结果为 0，则不发送 Keepalive 消息，且不检测 Holdtime 是否超时。

(2) 软复位 IPv6 BGP 连接

IPv6 BGP 的选路策略改变后，为了使新的策略生效，必须复位 IPv6 BGP 连接，但这样会造成短暂的 IPv6 BGP 连接中断。在目前的实现中，IPv6 BGP 支持 Route-refresh 功能。当策略改变后，系统可以在不中断 IPv6 BGP 连接的情况下，自动对 IPv6 BGP 路由表进行动态刷新。

在所有 IPv6 BGP 路由器使能 Route-refresh 功能的情况下，如果 IPv6 BGP 的路由策略发生了变化，本地路由器会向对等体发布 Route-refresh 消息，收到此消息的对等体会将其路由信息重新发给本地 IPv6 BGP 路由器。这样，在不中断 IPv6 BGP 连接的情况下，就可以对 IPv6 BGP 路由表进行动态更新，并应用新的策略。

如果网络中存在有不支持 Route-Refresh 的路由器，则需要配置 **peer keep-all-routes** 命令，将其所有路由更新保存在本地；当路由策略发生变化时，系统将对 IPv6 BGP 路由表进行动态更新，并应用新的策略。

1.6.1 配置准备

在调整 IPv6 BGP 的时钟之前，需完成以下任务：

- 使能 IPv6 能力

- 配置 IPv6 BGP 基本功能

1.6.2 配置IPv6 BGP的时钟

表1-20 配置 IPv6 BGP 的时钟

步骤		操作	命令
进入系统视图		**system-view**	-
进入BGP视图		**bgp** *as-number*	必选
进入IPv6地址族视图		**ipv6-family**	-
配置IPv6 BGP定时器	配置IPv6 BGP的存活时间间隔与保持时间	**timer keepalive** *keepalive* **hold** *holdtime*	可选 缺省情况下，存活时间间隔为60秒，保持时间为180秒
	配置指定IPv6对等体/IPv6对等体组的存活和保持时间	**peer** { *ipv6-group-name* | *ipv6-address* } **timer keepalive** *keepalive* **hold** *holdtime*	
配置IPv6对等体/IPv6对等体组的发送同一路由更新报文的时间间隔		**peer** { *ipv6-group-name* | *ipv6-address* } **route-update-interval** *interval*	可选 缺省情况下，向IBGP对等体发送同一路由更新的时间间隔为15秒，向EBGP对等体发送同一路由更新的时间间隔为30秒

说明

- 使用 **timer** 命令配置的定时器比使用 **peer timer** 命令配置的定时器优先级要低。
- 设置的保持时间应该至少为存活时间间隔的三倍。

1.6.3 配置IPv6 BGP软复位

1. 使能Route-refresh功能

表1-21 使能 Route-refresh 功能

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
使能IPv6 BGP路由刷新功能	**peer** { *ipv6-group-name* | *ipv6-address* } **capability-advertise route-refresh**	可选 缺省情况下，使能IPv6 BGP路由刷新功能

2. 手工软复位IPv6 BGP

表1-22 手工软复位 IPv6 BGP

步骤	操作	命令
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
保存所有来自IPv6对等体/IPv6对等体组的原始路由信息，即使这些路由没有通过已配置的入口策略	**peer** { *ipv6-group-name* | *ipv6-address* } **keep-all-routes**	可选 缺省情况下，不保存IPv6对等体/IPv6对等体组的原始路由信息
返回用户视图	**return**	-
手工对IPv6 BGP连接进行软复位	**refresh bgp ipv6** { **all** | *ipv6-address* | **group** *ipv6-group-name* | **external** | **internal** } { **export** | **import** }	必选

说明

配置命令 **peer keep-all-routes** 后，不论是否使用了过滤策略，都将保存指定对等体发来的所有路由更新。当软复位 IPv6 BGP 连接时，这些信息可以用来重新生成 IPv6 BGP 路由。

1.6.4 使能IPv6 BGP ORF能力功能

BGP ORF 是将本地入口策略通过 Route-refresh 报文推给邻居，当邻居需要向 BGP 对等体发送 Update 更新报文时，通过本地的路由策略后最后还需要进行 ORF 策略的过滤，只有通过 ORF 策略的路由信息才会发给 BGP 对等体，以达到减少 BGP 邻居间 Update 更新报文的交互，节省网络资源的目的。

使能 BGP ORF 能力后，本地和 BGP 对等体会通过 Open 报文协商 ORF 能力（即收发的报文里是否允许携带 ORF 信息，如果允许携带，是否可以携带非标准的 ORF 信息），当协商完毕并成功建立邻居关系后，可以通过特殊的 Route-refresh 报文交互 ORF 信息。

ORF能力协商成功需要两端的配置来保证，关于两端参数的选择请参见 表 1-24。

表1-23 配置 BGP ORF 能力

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
使能BGP路由刷新功能	**peer** { *group-name* | *ipv6-address* } **capability-advertise route-refresh**	必选 缺省情况下，BGP路由刷新功能处于使能状态

操作	命令	说明
使能BGP ORF非标准功能	**peer** { *group-name* | *ipv6-address* } **capability-advertise orf non-standard**	可选 缺省情况下，BGP ORF能力支持RFC5291,RFC5292的标准能力
使能BGP ORF地址前缀能力协商功能	**peer** { *group-name* | *ip-address* | *ipv6-address* } **capability-advertise orf ipv6-prefix** { **both** | **receive** | **send** }	必选 缺省情况下，BGP 不支持ORF 地址前缀的能力协商

表1-24 both、send、receive 参数选择以及配置效果描述表

本地选择参数	对端选择参数	协商成功后
send	**receive**	本端的ORF发送能力，对端的ORF接收能力
	both	
receive	**send**	本端的ORF接收能力，对端的ORF发送能力
	both	
both	**both**	双向的ORF发送和接收能力

1.6.5 使能 4 字节AS号抑制功能

通常情况下，设备在与对端设备建立 IPv6 BGP 对等体连接关系时会发送 IPv6 BGP Open 消息，在消息中的 Optional parameters 字段携带信息：说明 AS 号取值占用 4 字节（即取值范围为 1～4294967295）。此时，如果对端设备不支持 AS 号取值为 4 字节（比如只支持 2 字节），则连接关系无法建立。

为了解决上述问题，您可以使能 4 字节 AS 号抑制功能：当对端设备不支持 AS 号取值为 4 字节时，设备发出的 IPv6 BGP Open 消息仍然可以被对端设备正常识别，从而 IPv6 BGP 对等体连接关系可以成功建立。

表1-25 使能 4 字节 AS 号抑制功能

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6地址族视图	**ipv6-family**	-
使能4字节AS号抑制功能	**peer** { *group-name* | *ip-address* } **capability-advertise suppress-4-byte-as**	必选 缺省情况下，设备没有使能4字节AS号抑制功能

注意

如果对端设备支持 4 字节 AS 号能力时，请不要使能该功能，否则会导致对等体无法建立。

1.6.6 配置IPv6 BGP报文的DSCP优先级

在 IPv6 报文头中，包含一个 8bit 的 Traffic class 字段，用于标识 IP 报文的服务类型。RFC 2474 对这 8 个 bit 进行了定义，将前 6 个 bit 定义为 DSCP 优先级，最后 2 个 bit 作为保留位。在报文传输的过程中，DSCP 优先级可以被网络设备识别，并作为报文传输优先程度的参考。

用户可以对 IPv6 BGP 报文的 DSCP 优先级进行配置。

表1-26 配置向 IPv6 BGP 邻居发送报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** as-number	必选
进入IPv6地址族视图	**ipv6-family**	-
配置向指定IPv6对等体/ IPv6对等体组发送IPv6 BGP报文的DSCP优先级	**peer** { *ipv6-group-name* | *ipv6-address* } **dscp** *dscp-value*	可选 缺省情况下，IPv6 BGP报文使用的DSCP优先级为48

1.6.7 配置最大等价路由的条数

表1-27 配置最大等价路由的条数

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置进行IPv6 BGP负载分担的路由条数	**balance** *number*	必选 缺省情况下，不进行负载分担

1.6.8 配置IPv6 BGP建立TCP连接时进行MD5 认证

IPv6 BGP 使用 TCP 作为传输层协议，可靠性有一定保证。为了提高 IPv6 BGP 的安全性，可以配置 IPv6 BGP 在建立 TCP 连接时进行 MD5 认证，认证通过后才能建立 TCP 连接。如果认证失败，则不能建立 TCP 连接。

表1-28 配置 IPv6 BGP 建立 TCP 连接时进行 MD5 认证

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** as-number	必选
进入IPv6地址族视图	**ipv6-family**	-
配置IPv6 BGP建立TCP连接时进行MD5认证	**peer** { *ipv6-group-name* | *ipv6-address* } **password** { **cipher** | **simple** } *password*	可选 缺省情况下，IPv6 BGP建立TCP连接时不进行MD5认证

说明

- IPv6 BGP 的 MD5 认证并不能对 BGP 报文进行认证。
- 如果启用 MD5 认证，参与认证的双方必须配置完全一致的认证方式和密码，否则将因为无法通过认证而不能建立 TCP 连接。

1.6.9 配置IPv6 BGP IPsec安全策略

从安全性角度来考虑，为了避免路由信息外泄或者对路由器进行恶意攻击，IPv6 BGP 提供基于 IPsec 的报文验证功能。

路由器在发送的 IPv6 BGP 报文中会携带配置好的 IPsec 安全策略的 SPI（Security Parameter Index，安全参数索引）值，接收报文时通过 SPI 值进行 IPsec 安全策略匹配：只有策略匹配的报文才能接收；否则将不会接收报文，不能正常建立邻居和学习路由。

1. 配置准备

配置 IPv6 BGP IPsec 安全策略之前，需要满足下面配置：

- 安全提议已经创建；
- IPsec 安全策略已经创建。

关于安全提议和 IPsec 安全策略的详细配置过程，请参考“安全配置指导”中的“IPsec”。

2. 配置IPv6 BGP IPsec安全策略

表1-29 配置 IPv6 BGP IPsec 安全策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	-
进入IPv6地址族视图	**ipv6-family**	-
配置IPv6 BGP对等体/对等体组应用IPsec安全策略	**peer** { *group-name* | *ip-address* } **ipsec-policy** *policy-name*	必选 缺省情况下，IPv6 BGP对等体/对等体组没有应用IPsec安全策略

说明

IPv6 BGP IPsec 安全策略只支持手工配置方式，具体配置请参见“安全配置指导”中的“IPsec”。

1.7 组建大型IPv6 BGP网络

为方便管理，减少重复配置操作，管理员将一些享有相同更新策略的 IPv6 BGP 对等体划分到一个逻辑组织，这些组织就称为对等体组。只须对一个对等体组配置策略，该策略对组内所有成员都有效。

当对等体组的配置变化时，每个组员的配置也相应变化。对某些属性可以指定组员的 IPv6 地址进行配置，从而使指定组员在这些属性上不受对等体组配置的影响。

通常将同一自治系统的对等体配置为一个对等体组，当然，也可以将其它自治系统的对等体加入到该组。所有 IBGP 对等体可配置为另一个对等体组，也可以根据业务逻辑创建对等体组。

对等体组可以使一组对等体共享相同的策略，而利用团体可以使多个 AS 中的一组 IPv6 BGP 路由器共享相同的策略。团体是一个路由属性，在 IPv6 BGP 对等体之间传播，它并不受到 AS 范围的限制。

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。当 IBGP 对等体数目很多时，建立全连接网的开销很大。路由反射器和联盟是解决这个问题的两种方法。

IPv6 BGP 联盟的配置方式与 BGP 相同，这里不再赘述。以下主要介绍：

- 配置 IPv6 BGP 对等体组
- 配置 IPv6 BGP 团体
- 配置 IPv6 BGP 路由反射器

1.7.1 配置准备

在配置 IPv6 BGP 对等体组之前，需完成以下任务：

- 相邻节点网络层互通
- 启动 BGP，配置 Router ID

1.7.2 配置IPv6 BGP对等体组

1. 创建IBGP对等体组

表1-30 创建 IBGP 对等体组

操作	命令	说明
进入系统视图	**system-view**	-
启动BGP，进入BGP视图	**bgp** *as-number*	必选 缺省情况下，系统没有运行BGP

操作	命令	说明
进入IPv6地址族视图	**ipv6-family**	-
创建IBGP对等体组	**group** *ipv6-group-name* [**internal**]	必选
将对等体加入已存在的对等体组	**peer** *ipv6-address* **group** *ipv6-group-name* [**as-number** *as-number*]	必选 缺省情况下，对等体不属于任何对等体组

2. 创建纯EBGP对等体组

表1-31 创建纯 EBGP 对等体组

操作	命令	说明
进入系统视图	**system-view**	-
启动BGP，进入BGP视图	**bgp** *as-number*	必选 缺省情况下，系统没有运行BGP
进入IPv6地址族视图	**ipv6-family**	-
创建EBGP对等体组	**group** *ipv6-group-name* **external**	必选
配置对等体组的AS号	**peer** *ipv6-group-name* **as-number** *as-number*	必选 缺省情况下，对等体组无AS号
在已创建的对等体组中加入IPv6对等体	**peer** *ipv6-address* **group** *ipv6-group-name*	必选 缺省情况下，对等体不属于任何对等体组

说明

- 在创建纯 EBGP 对等体时，需要指定对等体组的自治系统号。
- 如果对等体组中已经加入了对等体，那么不能够为该对等体组指定自治系统号。

3. 创建混合EBGP对等体组

表1-32 创建混合 EBGP 对等体组

操作	命令	说明
进入系统视图	**system-view**	-
启动BGP，进入BGP视图	**bgp** *as-number*	必选 缺省情况下，系统没有运行BGP
进入IPv6地址族视图	**ipv6-family**	-
创建EBGP对等体组	**group** *ipv6-group-name* **external**	必选

操作	命令	说明
设置IPv6对等体的AS号	**peer** *ipv6-address* **as-number** *as-number*	必选 缺省情况下，IPv6对等体无AS号
加入已创建的IPv6对等体	**peer** *ipv6-address* **group** *ipv6-group-name*	必选 缺省情况下，对等体不属于任何对等体组

说明

在创建混合 EBGP 对等体组时，需要单独创建对等体，并可设置不同的自治系统号，但不能设置对等体组的自治系统号。

1.7.3 配置IPv6 BGP团体

1. 配置向IPv6 对等体/IPv6 对等体组发布团体属性

表1-33 配置向 IPv6 对等体/IPv6 对等体组发布团体属性

操作	命令	说明
进入系统视图	**system-view**	-
启动BGP，进入BGP视图	**bgp** *as-number*	必选 缺省情况下，系统没有运行BGP
进入IPv6地址族视图	**ipv6-family**	-
配置将团体属性传给IPv6对等体/IPv6对等体组	**peer** { *ipv6-group-name* | *ipv6-address* } **advertise-community**	必选 缺省情况下，不将团体属性发布给任何IPv6对等体/IPv6对等体组
配置将扩展团体属性传给IPv6对等体/IPv6对等体组	**peer** { *ipv6-group-name* | *ipv6-address* } **advertise-ext-community**	必选 缺省情况下，不将扩展团体属性发布给任何IPv6对等体/IPv6对等体组

2. 对发布的路由信息应用路由策略

表1-34 对发布的路由信息应用路由策略

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置出方向路由策略	**peer** { *ipv6-group-name* | *ipv6-address* } **route-policy** *route-policy-name* **export**	必选 缺省情况下，不指定IPv6对等体/IPv6对等体组的路由策略

说明

- 配置 IPv6 BGP 团体时，必须使用路由策略来定义具体的团体属性，然后在发布路由信息时应用此路由策略。
- 关于路由策略的配置，请参见“三层技术-IP 路由配置指导”中的“路由策略”。

1.7.4 配置IPv6 BGP路由反射器

表1-35 配置 IPv6 BGP 路由反射器

操作	命令	说明
进入系统视图	**system-view**	-
进入BGP视图	**bgp** *as-number*	必选
进入IPv6地址族视图	**ipv6-family**	-
配置将本机作为路由反射器，并将IPv6对等体/IPv6对等体组作为路由反射器的客户	**peer** { *ipv6-group-name* | *ipv6-address* } **reflect-client**	必选 缺省情况下，没有配置路由反射器及其客户
使能客户机之间的路由反射	**reflect between-clients**	可选 缺省情况下，允许客户到客户的路由反射
配置路由反射器的集群ID	**reflector cluster-id** *cluster-id*	可选 缺省情况下，每个路由反射器是使用自己的Router ID作为集群ID

说明

- 通常情况下，路由反射器的客户之间不要求是全连接的，路由缺省通过反射器从一个客户反射到其它客户；如果客户之间是全连接的，可以禁止客户间的反射，以便减少开销。
- 当一个集群里有多个路由反射器时，需要给所有位于同一个集群内的路由反射器配置相同的Cluster_ID，以避免路由循环。

1.8 配置IPv6 BGP与BFD联动

IPv6 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 IPv6 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 IPv6 BGP 与 BFD 联动，使得 BFD 能够为 IPv6 BGP 邻居之间的链路提供更快速的检测。当邻居之间的链路出现故障时，能够加快 IPv6 BGP 协议的收敛速度。

表1-36 配置 IPv6 BGP 与 BFD 联动

操作	命令	说明
进入系统视图	**system-view**	-
启动BGP，进入BGP视图	**bgp** *as-number*	必选 缺省情况下，系统不运行BGP
进入IPv6地址族视图	**ipv6-family**	-
在指定邻居上使能BFD	**peer** *ipv6-address* **bfd**	必选 缺省情况下，IPv6 BGP的所有对等体邻居上都没有使能BFD链路检测功能

说明

- 配置 BFD 之前，需要使能 BGP 功能。
- 在 BGP 使能 GR 能力后，请慎用 BFD。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。
- 有关 BFD 的介绍和基本功能配置，请参见“可靠性配置指导”中的“BFD”。

1.9 IPv6 BGP显示和维护

1.9.1 IPv6 BGP显示

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 BGP 的运行情况，通过查看显示信息验证配置的效果。

表1-37 显示 IPv6 BGP 的运行状态

操作	命令
显示对等体组信息	**display bgp ipv6 group** [*ipv6-group-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 BGP发布的IPv6路由信息	**display bgp ipv6 network** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示AS路径信息	**display bgp ipv6 paths** [*as-regular-expression* | **|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6对等体/IPv6对等体组信息	**display bgp ipv6 peer** [*group-name* **log-info** | *ipv4-address* **verbose** | *ipv6-address* { **log-info** | **verbose** } | **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示收到的邻居ORF信息中的前缀信息	**display bgp ipv6 peer** { *ip-address* | *ipv6-address* } **received ipv6-prefix** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 BGP路由信息	**display bgp ipv6 routing-table** [*ipv6-address prefix-length*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配指定AS路径过滤列表的路由	**display bgp ipv6 routing-table as-path-acl** *as-path-acl-number* [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
显示指定IPv6 BGP团体的路由信息	**display bgp ipv6 routing-table community** [*aa:nn*<1-13>] [**no-advertise** | **no-export** | **no-export-subconfed**]* [**whole-match**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配指定IPv6 BGP团体列表的路由信息	**display bgp ipv6 routing-table community-list** { { *basic-community-list-number* | *comm-list-name* } [**whole-match**] | *adv-community-list-number* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 BGP衰减的路由	**display bgp ipv6 routing-table dampened** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 BGP路由衰减参数	**display bgp ipv6 routing-table dampening parameter** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示来自不同自治系统的IPv6 BGP路由	**display bgp ipv6 routing-table different-origin-as** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6 BGP路由震荡统计信息	**display bgp ipv6 routing-table flap-info** [**regular-expression** *as-regular-expression* | [**as-path-acl** *as-path-acl-number* | *ipv6-address prefix-length* [**longer-match**]] [**|** { **begin** | **exclude** | **include** } *regular-expression*]]
显示向指定的IPv4/IPv6 BGP对等体发送或者从指定的IPv4/IPv6 BGP对等体收到的路由信息	**display bgp ipv6 routing-table peer** { *ipv4-address* | *ipv6-address* } { **advertised-routes** | **received-routes** } [*network-address prefix-length* | **statistic**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示匹配AS正则表达式的IPv6路由信息	**display bgp ipv6 routing-table regular-expression** *as-regular-expression*
显示IPv6 BGP的路由统计信息	**display bgp ipv6 routing-table statistic** [**|** { **begin** | **exclude** | **include** } *regular-expression*]

1.9.2 复位IPv6 BGP连接

当 BGP 路由策略或协议发生变化后，如果需要通过复位 IPv6 BGP 连接使新的配置生效，请在用户视图下进行下列配置。

表1-38 复位 IPv6 BGP 连接

操作	命令
手工软复位IPv6 BGP连接	**refresh bgp ipv6** { *ipv4-address* | *ipv6-address* | **all** | **external** | **group** *ipv6-group-name* | **internal** } { **export** | **import** }
复位所有IPv6 BGP连接	**reset bgp ipv6** { *as-number* | *ipv4-address* | *ipv6-address* | **all** | **external** | **group** *group-name* | **internal** }

1.9.3 清除IPv6 BGP信息

在用户视图下执行 **reset** 命令可清除 IPv6 BGP 路由的统计信息。

表1-39 清除 IPv6 BGP 信息

操作	命令
清除IPv6 BGP路由的衰减信息并释放被抑制的路由	**reset bgp ipv6 dampening** [*ipv6-address prefix-length*]
清除IPv6 BGP路由的震荡统计信息	**reset bgp ipv6 flap-info** [*ipv6-address/prefix-length* | **as-path-acl** *as-path-acl-number* | **regexp** *as-path-regexp*] **reset bgp ipv6** *peer-ipv6-address* **flap-info**

1.10 IPv6 BGP典型配置举例

说明

IPv6 BGP 的配置举例多数与 BGP4 的配置举例相似，本节不再一一列出。其他配置例请参见“三层技术-IP 路由配置指导”中的“BGP”。

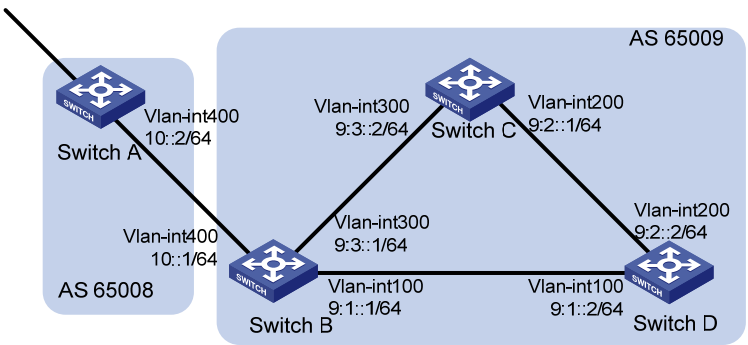
1.10.1 IPv6 BGP基本配置

1. 组网需求

如 图 1-1 所示，所有交换机运行IPv6 BGP协议，Switch A和Switch B之间建立EBGP连接，Switch B、Switch C和Switch D之间建立IBGP全连接。

2. 组网图

图1-1 IPv6 BGP 基本配置组网图



3. 配置步骤

(1) 配置各 VLAN 接口的 IPv6 地址（略）

(2) 配置 IBGP 连接

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ipv6
```

```
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer 9:1::2 as-number 65009
[SwitchB-bgp-af-ipv6] peer 9:3::2 as-number 65009
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] quit
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] peer 9:3::1 as-number 65009
[SwitchC-bgp-af-ipv6] peer 9:2::2 as-number 65009
[SwitchC-bgp-af-ipv6] quit
[SwitchC-bgp] quit
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] ipv6
[SwitchD] bgp 65009
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] ipv6-family
[SwitchD-bgp-af-ipv6] peer 9:1::1 as-number 65009
[SwitchD-bgp-af-ipv6] peer 9:2::1 as-number 65009
[SwitchD-bgp-af-ipv6] quit
[SwitchD-bgp] quit
```

(3) 配置 EBGP 连接

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 10::1 as-number 65009
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] quit
```

配置 Switch B。

```
[SwitchB] bgp 65009
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer 10::2 as-number 65008
```

查看 Switch B 的对等体信息。

```
[SwitchB] display bgp ipv6 peer

 BGP local router ID : 2.2.2.2
 Local AS number : 65009
 Total number of peers : 3                 Peers in established state : 3
```

```
  Peer            AS  MsgRcvd  MsgSent  OutQ PrefRcv Up/Down  State

  10::2        65008        3        3     0       0 00:01:16 Established
  9:3::2       65009        2        3     0       0 00:00:40 Established
  9:1::2       65009        2        4     0       0 00:00:19 Established
```

查看 Switch C 的对等体信息。

```
[SwitchC] display bgp ipv6 peer

 BGP local router ID : 3.3.3.3
 Local AS number : 65009
 Total number of peers : 2                 Peers in established state : 2

  Peer            AS  MsgRcvd  MsgSent  OutQ PrefRcv Up/Down  State

  9:3::1       65009        4        4     0       0 00:02:18 Established
  9:2::2       65009        4        5     0       0 00:01:52 Established
```

可以看出，Switch A 和 Switch B 之间建立的 EBGP 连接，Switch B、Switch C 和 Switch D 之间建立了 IBGP 连接。

1.10.2 配置IPv6 BGP路由反射

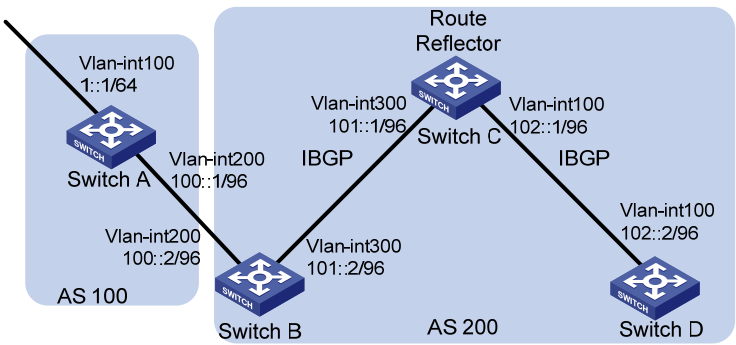
1. 组网需求

Switch B 接收了一个经过 EBGP 的更新报文并将之传给 Switch C。Switch C 被配置为路由反射器，它有两个客户机：Switch B 和 Switch D。

SwitchB 和 Switch D 间不需要建立 IBGP 连接，当 Switch C 接收了来自 Switch B 的路由更新时，它将此信息反射给 Switch D；反之，接收到的 Switch D 的路由信息也会被反射给 Switch B。

2. 组网图

图1-2 配置 IPv6 BGP 路由反射器组网图



3. 配置步骤

(1) 配置各 VLAN 接口的 IPv6 地址（略）

(2) 配置 IPv6 BGP 基本功能

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 100::2 as-number 200
[SwitchA-bgp-af-ipv6] network 1:: 64
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer 100::1 as-number 100
[SwitchB-bgp-af-ipv6] peer 101::1 as-number 200
[SwitchB-bgp-af-ipv6] peer 101::1 next-hop-local
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] bgp 200
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] peer 101::2 as-number 200
[SwitchC-bgp-af-ipv6] peer 102::2 as-number 200
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] ipv6
[SwitchD] bgp 200
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] ipv6-family
[SwitchD-bgp-af-ipv6] peer 102::1 as-number 200
```

(3) 配置路由反射器

配置 Switch C 作为路由反射器，Switch B 和 Switch D 是它的两个客户机。

```
[SwitchC-bgp-af-ipv6] peer 101::2 reflect-client
[SwitchC-bgp-af-ipv6] peer 102::2 reflect-client
```

在 Switch B 上用命令 **display bgp ipv6 routing-table**，可以看到 Switch B 已知道网络 1::/64 的存在。

在 Switch D 上用命令 **display bgp ipv6 routing-table**，可以看到 Switch D 也知道网络 1::/64 的存在。

1.10.3 配置IPv6 BGP IPsec安全策略

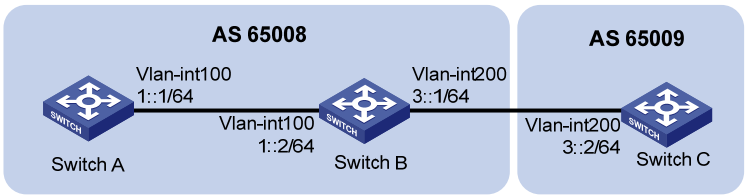
1. 组网需求

- Switch A、Switch B 和 Switch C 三台交换机互连，使能 IPv6 BGP 协议。Switch A 和 Switch B 之间建立 IBGP 连接，Switch B 和 Switch C 之间建立 EBGP 连接。

- 要求配置IPsec安全策略对Switch A和Switch B之间、Switch B和Switch C之间的IPv6 BGP报文进行验证和加密。

2. 组网图

图1-3 IPv6 BGP IPsec 安全策略组网图



3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 IBGP 连接

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] group ibgp internal
[SwitchA-bgp-af-ipv6] peer 1::2 group ibgp
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] bgp 65008
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] group ibgp internal
[SwitchB-bgp-af-ipv6] peer 1::1 group ibgp
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] quit
```

(3) 配置 EBGP 连接

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] group ebgp external
[SwitchC-bgp-af-ipv6] peer 3::1 as-number 65008
[SwitchC-bgp-af-ipv6] peer 3::1 group ebgp
[SwitchC-bgp-af-ipv6] quit
[SwitchC-bgp] quit
```

配置 Switch B。

```
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] group ebgp external
[SwitchB-bgp-af-ipv6] peer 3::2 as-number 65009
[SwitchB-bgp-af-ipv6] peer 3::2 group ebgp
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] quit
```

(4) 配置 IPv6 BGP IPsec 安全策略

配置 Switch 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。创建一条安全策略 policy001，协商方式为 manual，配置 SPI 和密钥。

```
[SwitchA] ipsec proposal tran1
[SwitchA-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchA-ipsec-proposal-tran1] transform esp
[SwitchA-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchA-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchA-ipsec-proposal-tran1] quit
[SwitchA] ipsec policy policy001 10 manual
[SwitchA-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchA-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] quit
```

配置 Switch 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；创建一条安全策略 policy001，协商方式为 manual，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；创建一条安全策略 policy002，协商方式为 manual，配置 SPI 和密钥。

```
[SwitchB] ipsec proposal tran1
[SwitchB-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran1] transform esp
[SwitchB-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran1] quit
[SwitchB] ipsec policy policy001 10 manual
[SwitchB-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchB-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] quit
[SwitchB] ipsec proposal tran2
[SwitchB-ipsec-proposal-tran2] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran2] transform esp
[SwitchB-ipsec-proposal-tran2] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran2] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran2] quit
[SwitchB] ipsec policy policy002 10 manual
```

```
[SwitchB-ipsec-policy-manual-policy002-10] proposal tran2
[SwitchB-ipsec-policy-manual-policy002-10] sa spi outbound esp 54321
[SwitchB-ipsec-policy-manual-policy002-10] sa spi inbound esp 54321
[SwitchB-ipsec-policy-manual-policy002-10] sa string-key outbound esp gfedcba
[SwitchB-ipsec-policy-manual-policy002-10] sa string-key inbound esp gfedcba
[SwitchB-ipsec-policy-manual-policy002-10] quit
```

配置 Switch 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。创建一条安全策略 policy002，协商方式为 manual，配置 SPI 和密钥。

```
[SwitchC] ipsec proposal tran2
[SwitchC-ipsec-proposal-tran2] encapsulation-mode transport
[SwitchC-ipsec-proposal-tran2] transform esp
[SwitchC-ipsec-proposal-tran2] esp encryption-algorithm des
[SwitchC-ipsec-proposal-tran2] esp authentication-algorithm sha1
[SwitchC-ipsec-proposal-tran2] quit
[SwitchC] ipsec policy policy002 10 manual
[SwitchC-ipsec-policy-manual-policy002-10] proposal tran2
[SwitchC-ipsec-policy-manual-policy002-10] sa spi outbound esp 54321
[SwitchC-ipsec-policy-manual-policy002-10] sa spi inbound esp 54321
[SwitchC-ipsec-policy-manual-policy002-10] sa string-key outbound esp gfedcba
[SwitchC-ipsec-policy-manual-policy002-10] sa string-key inbound esp gfedcba
[SwitchC-ipsec-policy-manual-policy002-10] quit
```

(5) 配置 IBGP 对等体 IPsec 安全策略

配置 Switch A。

```
[SwitchA] bgp 65008
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 1::2 ipsec-policy policy001
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] quit
```

配置 Switch B。

```
[SwitchB] bgp 65008
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer 1::1 ipsec-policy policy001
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] quit
```

(6) 配置 EBGP 对等体组 IPsec 安全策略

配置 Switch C。

```
[SwitchC] bgp 65009
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] peer ebgp ipsec-policy policy002
[SwitchC-bgp-af-ipv6] quit
[SwitchC-bgp] quit
```

配置 Switch B。

```
[SwitchB] bgp 65008
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer ebgp ipsec-policy policy002
[SwitchB-bgp-af-ipv6] quit
```

```
[SwitchB-bgp] quit
```

4. 验证配置结果

显示 Switch B 的对等体的详细信息。

```
[SwitchB] display bgp ipv6 peer verbose
         BGP Peer is 1::1,  remote AS 65008,
         Type: IBGP link
         BGP version 4, remote router ID 1.1.1.1
         BGP current state: Established, Up for 00h01m51s
         BGP current event: RecvKeepalive
         BGP last state: OpenConfirm
         Port:  Local – 1029     Remote - 179
         Configured: Active  Hold Time: 180 sec  Keepalive Time: 60 sec
         Received  : Active  Hold Time: 180 sec
         Negotiated: Active  Hold Time: 180 sec
         Peer optional capabilities:
         Peer support bgp multi-protocol extended
         Peer support bgp route refresh capability
         Address family IPv4 Unicast: advertised and received
 Received: Total 0 messages, Update messages 0
 Sent: Total 0 messages, Update messages 0
 Maximum allowed prefix number: 4294967295
 Threshold: 75%
 Minimum time between advertisement runs is 30 seconds
 Optional capabilities:
  Route refresh capability has been enabled
  ORF advertise capability based on prefix (type 64):
    Local: both
    Negotiated: send
 Peer Preferred Value: 0
 IPsec policy name: policy001, SPI :12345
 Routing policy configured:
 No routing policy is configured
         BGP Peer is 3::2,  remote AS 65009,
         Type: EBGP link
         BGP version 4, remote router ID 3.3.3.3
         BGP current state: Established, Up for 00h01m51s
         BGP current event: RecvKeepalive
         BGP last state: OpenConfirm
         Port:  Local – 1029     Remote - 179
         Configured: Active  Hold Time: 180 sec  Keepalive Time: 60 sec
         Received  : Active  Hold Time: 180 sec
         Negotiated: Active  Hold Time: 180 sec
         Peer optional capabilities:
         Peer support bgp multi-protocol extended
         Peer support bgp route refresh capability
         Address family IPv4 Unicast: advertised and received
 Received: Total 0 messages, Update messages 0
 Sent: Total 0 messages, Update messages 0
```

```
Maximum allowed prefix number: 4294967295
Threshold: 75%
Minimum time between advertisement runs is 30 seconds
Optional capabilities:
 Route refresh capability has been enabled
 ORF advertise capability based on prefix (type 64):
   Local: both
   Negotiated: send
Peer Preferred Value: 0
IPsec policy name: policy002, SPI :54321
Routing policy configured:
No routing policy is configured
```

完成上述配置后 IBGP、EBGP 邻居能正常建立，且报文都经过加密。

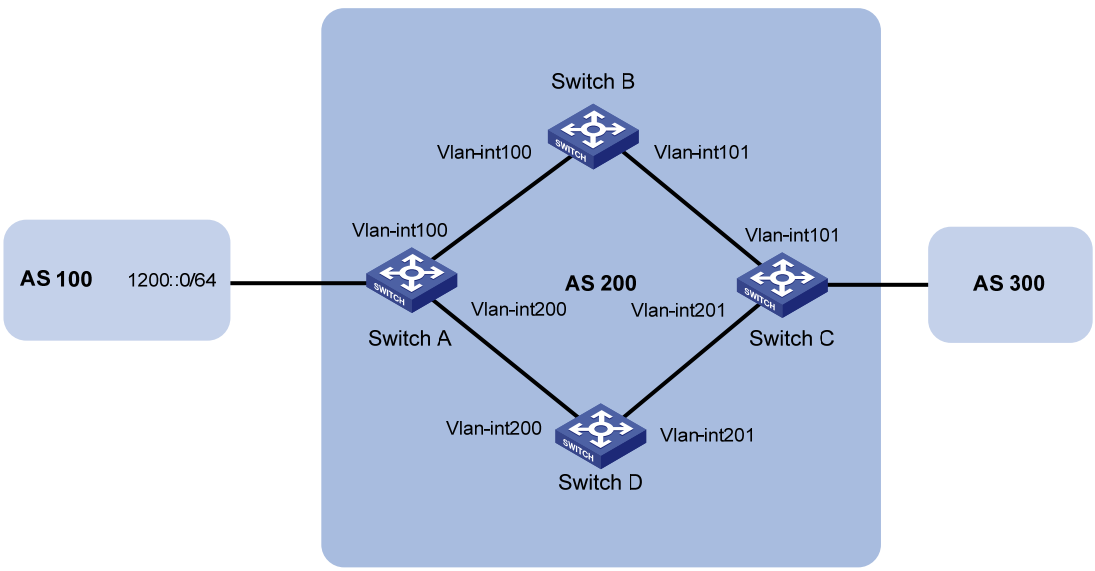
1.10.4 配置IPv6 BGP与BFD联动

1. 组网需求

- 在 AS 200 内使用 OSPFv3 作为 IGP 协议；
- Switch A 与 Switch C 建立两条 IBGP 连接。当 Switch A 与 Switch C 之间的两条链路均连通时，Switch C 与 1200::0/64 之间的报文使用 Switch A<—>Switch B<—>Switch C 这条路径；当 Switch A<—>Switch B<—>Switch C 链路故障时，BFD 能够快速检测并通告 IPv6 BGP 协议，使得 Switch A<—>Switch D<—>Switch C 这条路径能够迅速生效。

2. 组网图

图1-4 IPv6 BGP 与 BFD 联动配置组网图



设备	接口	IP地址	设备	接口	IP地址
Switch A	Vlan-int100	3000::1/64	Switch C	Vlan-int101	3001::3/64
	Vlan-int200	2000::1/64		Vlan-int201	2001::3/64
Switch B	Vlan-int100	3000::2/64	Switch D	Vlan-int200	2000::2/64
	Vlan-int101	3001::2/64		Vlan-int201	2001::2/64

3. 配置步骤

(1) 配置各接口的 IPv6 地址（略）

(2) 配置 OSPFv3（略），保证 Switch A 和 Switch C 之间路由可达

(3) Switch A 上的 IPv6 BGP 配置

配置 Switch A 和 Switch C 建立两条 IBGP 连接。

```
<SwitchA> system-view
[SwitchA] bgp 200
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 3001::3 as-number 200
[SwitchA-bgp-af-ipv6] peer 2001::3 as-number 200
[SwitchA-bgp-af-ipv6] quit
```

配置当 Switch A 与 Switch C 之间的两条链路均连通时，Switch C 与 1200::0/64 之间的报文使用 Switch A<－>Switch B<－>Switch C 这条路径。（在 Switch A 上对发布给对等体 2001::3 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。

```
[SwitchA] acl ipv6 number 2000
[SwitchA-acl6-basic-2000] rule permit source 1200::0 64
[SwitchA-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。

```
[SwitchA] route-policy apply_med_50 permit node 10
[SwitchA-route-policy] if-match ipv6 address acl 2000
[SwitchA-route-policy] apply cost 50
[SwitchA-route-policy] quit
[SwitchA] route-policy apply_med_100 permit node 10
[SwitchA-route-policy] if-match ipv6 address acl 2000
[SwitchA-route-policy] apply cost 100
[SwitchA-route-policy] quit
```

- 对发布给对等体 3001::3 的路由应用名为 apply_med_50 的 Route-policy，对发布给对等体 2001::3 的路由应用名为 apply_med_100 的 Route-policy。

```
[SwitchA] bgp 200
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] network 1200:: 64
[SwitchA-bgp-af-ipv6] peer 3001::3 route-policy apply_med_50 export
[SwitchA-bgp-af-ipv6] peer 2001::3 route-policy apply_med_100 export
```

配置当 Switch A<－>Switch B<－>Switch C 链路故障时，BFD 能够快速检测并通告 IPv6 BGP 协议，使得 Switch A<－>Switch D<－>Switch C 这条路径能够迅速生效。

```
[SwitchA-bgp-af-ipv6] peer 3001::3 bfd
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] quit
```

(4) Switch C 上的 IPv6 BGP 配置。

```
<SwitchC> system-view
[SwitchC] bgp 200
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] peer 3000::1 as-number 200
```

```
[SwitchC-bgp-af-ipv6] peer 3000::1 bfd
[SwitchC-bgp-af-ipv6] peer 2000::1 as-number 200
[SwitchC-bgp-af-ipv6] quit
[SwitchC-bgp] quit
```

(5) 配置各路由器接口及 BFD 参数（该步骤可省略，使用 BFD 参数的缺省值即可）

配置 Switch A。

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 100
```

- 配置发送 BFD 控制报文的最小时间间隔为 500 毫秒。

```
[SwitchA-Vlan-interface100] bfd min-transmit-interval 500
```

- 配置接收 BFD 控制报文的最小时间间隔为 500 毫秒。

```
[SwitchA-Vlan-interface100] bfd min-receive-interval 500
```

- 配置 BFD 报文最大失效的个数为 7。

```
[SwitchA-Vlan-interface100] bfd detect-multiplier 7
[SwitchA-Vlan-interface100] quit
```

配置 Switch C。

```
[SwitchC] bfd session init-mode active
[SwitchC] interface vlan-interface 101
```

- 配置发送 BFD 控制报文的最小时间间隔为 500 毫秒。

```
[SwitchC-Vlan-interface101] bfd min-transmit-interval 500
```

- 配置接收 BFD 控制报文的最小时间间隔为 500 毫秒。

```
[SwitchC-Vlan-interface101] bfd min-receive-interval 500
```

- 配置 BFD 报文最大失效的个数为 7。

```
[SwitchC-Vlan-interface101] bfd detect-multiplier 7
[SwitchC-Vlan-interface101] return
```

(6) 检查配置结果

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

显示 Switch C 的 BFD 信息。

```
<SwitchC> display bfd session verbose

 Total session number: 1   Up session number: 1   Init mode: Active

 IPv6 Session working under Ctrl mode:

     Local Discr: 17                 Remote Discr: 13
       Source IP: 3001::3
  Destination IP: 3000::1
   Session State: Up                    Interface: Vlan-interface101
 Min Trans Inter: 500ms            Act Trans Inter: 500ms
  Min Recv Inter: 500ms           Act Detect Inter: 3000ms
    Recv Pkt Num: 57                 Send Pkt Num: 53
       Hold Time: 2200ms             Connect Type: Direct
  Running Up for: 00:00:06              Auth mode: none
        Protocol: BGP6
       Diag Info: No Diagnostic
```

以上显示信息表明：Switch A 的 VLAN 接口 100 和 Switch C 的 VLAN 接口 101 之间已经建立了 BFD 连接，而且 BFD 协议运行正常。

在 Switch C 上查看 BGP 邻居信息，可以看出 Switch A 和 Switch C BGP 邻居处于 Established 状态。

```
<SwitchC> display bgp ipv6 peer
 BGP local router ID : 1.1.1.1
 Local AS number : 200
 Total number of peers : 2                 Peers in established state : 2
  Peer                    AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State
  2000::1                200        7       10    0       0 00:01:05 Established
  3000::1                200        7       10    0       0 00:01:34 Established
```

在 Switch C 上查看 1200::0/64 的路由信息，可以看出 Switch A 和 Switch C 是通过 Switch B 进行通信的。

```
<SwitchC> display ipv6 routing-table 1200::0 64 verbose
Routing Table :
Summary Count : 2

 Destination  : 1200::                                PrefixLength : 64
 NextHop      : 3000::1                               Preference   : 255
 RelayNextHop : 3001::2                               Tag          : 0H
 Neighbor     : 3000::1                               ProcessID    : 0
 Interface    : Vlan-interface101                    Protocol     : BGP4+
 State        : Active Adv                            Cost         : 50
 Tunnel ID    : 0x0                                   Label        : NULL
 Age          : 4538sec

 Destination  : 1200::                                PrefixLength : 64
 NextHop      : 2000::1                               Preference   : 255
 RelayNextHop : 2001::2                               Tag          : 0H
 Neighbor     : 2000::1                               ProcessID    : 0
 Interface    : Vlan-interface201                    Protocol     : BGP4+
 State        : Invalid Adv                           Cost         : 100
 Tunnel ID    : 0x0                                   Label        : NULL
 Age          : 4515sec
```

以上显示信息表明：Switch C 到达 1200::0/64 网段有两条路径：当前生效的是 Switch A<－>Switch B<－>Switch C，Switch A<－>Switch D<－>Switch C 这条路径没有生效。

当 Switch A 和 Switch B 之间的链路发生故障时 Switch C 能够快速检测 Switch B 的变化。

在 Switch C 上查看 1200::0/64 的路由信息，可以看出 Switch A 和 Switch C 是通过 Switch D 进行通信的。

```
<SwitchC> display ipv6 routing-table 1200::0 64 verbose
Routing Table :
Summary Count : 1
 Destination  : 1200::                                PrefixLength : 64
 NextHop       : 2000::1                              Preference   : 255
 RelayNextHop : 2001::2                               Tag          : 0H
 Neighbor     : 2000::1                               ProcessID    : 0
```

```
 Interface     : Vlan-interface201                    Protocol      : BGP4+
 State         : Active Adv                           Cost          : 100
 Tunnel ID     : 0x0                                  Label         : NULL
 Age           : 4635sec
```

以上显示信息表明：Switch C 到达 1200::0/64 网段有一条路径：Switch A<－>Switch D<－>Switch C。

1.11 IPv6 BGP常见错误配置举例

1.11.1 IPv6 BGP对等体关系不能建立

1. 错误现象

使用 **display bgp ipv6 peer** 命令查看 IPv6 BGP 对等体的信息，发现与对端的连接无法进入 Established 状态。

2. 分析

IPv6 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。

3. 处理过程

(1) 用 **display current-configuration configuration bgp** 命令检查邻居的 AS 号配置是否正确。

(2) 用 **display bgp ipv6 peer** 命令检查邻居的 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) 检查路由表中是否存在到邻居的可用路由。

(6) 使用 **ping** 命令检查链路是否畅通。

(7) 使用 **display tcp ipv6 status** 命令检查 TCP 连接是否正常。

(8) 检查是否配置了禁止 TCP 端口 179 的 ACL。

目 录

1 路由策略配置 ····· 1-1
1.1 路由策略简介 ····· 1-1
1.1.1 路由策略的应用 ····· 1-1
1.1.2 路由策略的实现 ····· 1-1
1.1.3 过滤器 ····· 1-1
1.2 配置过滤列表 ····· 1-3
1.2.1 配置准备 ····· 1-3
1.2.2 配置地址前缀列表 ····· 1-3
1.2.3 配置AS路径过滤列表 ····· 1-5
1.2.4 配置团体属性列表 ····· 1-5
1.2.5 配置扩展团体属性列表 ····· 1-5
1.3 配置路由策略 ····· 1-6
1.3.1 配置准备 ····· 1-6
1.3.2 创建一个路由策略 ····· 1-6
1.3.3 配置if-match子句 ····· 1-7
1.3.4 配置apply子句 ····· 1-9
1.3.5 配置continue子句 ····· 1-10
1.4 路由策略的显示和维护 ····· 1-11
1.5 路由策略典型配置举例 ····· 1-11
1.5.1 在IPv4 路由引入中应用路由策略 ····· 1-11
1.5.2 在IPv6 路由引入中应用路由策略 ····· 1-14
1.5.3 配置应用路由策略过滤接收的BGP路由 ····· 1-16
1.6 常见错误配置举例 ····· 1-18
1.6.1 无法实现IPv4 路由信息过滤 ····· 1-18
1.6.2 无法实现IPv6 路由信息过滤 ····· 1-19

1 路由策略配置

说明

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（Routing Policy）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（不再进入下一个表项的测试）。

3. AS路径访问列表（as-path）

as-path 仅用于 BGP。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。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。BGP 扩展团体属性有两种，一种是用于 VPN 的 Route-Target（路由目标）扩展团体，另一种则是 Source of Origin（源节点）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. 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（node）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。

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测试。
- **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

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 ip-prefix** *ip-prefix-name* [**index** *index-number*] { **deny** | **permit** } *ip-address mask-length* [**greater-equal** *min-mask-length*] [**less-equal** *max-mask-length*]	必选 缺省情况下，没有配置IPv4地址前缀列表

说明

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，建议在多条 **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。

例如，按如下配置可以保证仅过滤掉 10.1.0.0/16、10.2.0.0/16、10.3.0.0/16 三个网段的路由，而其它网段的路由信息可以通过。

```
<Sysname> system-view
[Sysname] ip ip-prefix abc index 10 deny 10.1.0.0 16
[Sysname] ip ip-prefix abc index 20 deny 10.2.0.0 16
[Sysname] ip ip-prefix abc index 30 deny 10.3.0.0 16
[Sysname] ip ip-prefix abc index 40 permit 0.0.0.0 0 less-equal 32
```

2. 配置IPv6 地址前缀列表

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	**ip ipv6-prefix** *ipv6-prefix-name* [**index** *index-number*] { **deny** | **permit** } *ipv6-address prefix-length* [**greater-equal** *min-prefix-length*] [**less-equal** *max-prefix-length*]	必选 缺省情况下，没有配置IPv6地址前缀列表

说明

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 的表项，以允许其它所有 IPv6 路由信息通过。

例如，按如下配置可以保证仅过滤掉 2000:1::/48、2000:2::/48、2000:3::/48 三个网段的路由，而其它网段的路由信息可以通过。

```
<Sysname> system-view
[Sysname] ip ipv6-prefix abc index 10 deny 2000:1:: 48
[Sysname] ip ipv6-prefix abc index 20 deny 2000:2:: 48
[Sysname] ip ipv6-prefix abc index 30 deny 2000:3:: 48
[Sysname] ip ipv6-prefix abc index 40 permit :: 0 less-equal 128
```

1.2.3 配置AS路径过滤列表

一个 AS 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*	必选 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** *comm-list-name* } { **deny** | **permit** } [*community-number-list*] [**internet** | **no-advertise** | **no-export** | **no-export-subconfed**] *	二者必选其一 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** *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* }&<1-16>	必选 缺省情况下，没有配置扩展团体属性列表

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

表1-6 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*	必选 缺省情况下，没有创建路由策略

说明

- **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测试；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测试。
- **deny** 指定节点的匹配模式为拒绝模式（此模式下 **apply** 子句不会被执行）。当路由项满足该节点的所有 **if-match** 子句时，将被拒绝通过该节点，不进入下一个节点的测试；如果路由项不满足该节点的 **if-match** 子句，将进入下一个节点继续测试。
- 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**。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。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。

1.3.3 配置if-match子句

表1-7 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*	必选
配置IPv4的路由信息的匹配条件	配置IPv4路由信息的目的地址范围的匹配条件	**if-match acl** *acl-number*	可选 缺省情况下，没有配置IPv4路由信息的目的地址范围的匹配条件
		if-match ip-prefix *ip-prefix-name*	
	配置IPv4的路由信息的下一跳地址或源地址的匹配条件	**if-match ip** { **next-hop** | **route-source** } { **acl** *acl-number* | **ip-prefix** *ip-prefix-name* }	可选 缺省情况下，没有配置IPv4的路由信息的下一跳地址或源地址的匹配条件
配置IPv6的路由信息的匹配条件		**if-match ipv6** { **address** | **next-hop** | **route-source** } { **acl** *acl-number* | **prefix-list** *ipv6-prefix-name* }	可选 缺省情况下，没有配置IPv6的路由信息的匹配条件
配置BGP路由信息的AS路径域的匹配条件		**if-match as-path** *as-path-number*&<1-16>	可选 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		**if-match community** { { *basic-community-list-number* | *comm-list-name* } [**whole-match**] | *adv-community-list-number* }&<1-16>	可选 缺省情况下，不匹配BGP路由信息的团体属性。
配置路由信息的路由开销的匹配条件		**if-match cost** *value*	可选 缺省情况下，没有配置路由信息的路由开销的匹配条件

操作	命令	说明
配置BGP扩展团体属性的匹配条件	**if-match extcommunity** *ext-comm-list-number*&<1-16>	可选 缺省情况下，没有配置BGP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>	可选 缺省情况下，没有配置路由信息的出接口的匹配条件
配置路由信息的MPLS标签的匹配条件	**if-match mpls-label**	可选 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** } *	可选 缺省情况下，没有配置路由信息的类型的匹配条件
配置RIP、OSPF、IS-IS路由信息的标记域的匹配条件	**if-match tag** *value*	可选 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

说明

- 对于同一个节点，在匹配的过程中，不同 **if-match** 子句间是“与”的关系，即路由信息必须同时满足所有匹配条件，才可以执行 **apply** 子句。如果有 **if-match** 子句因超过命令行最大长度而出现多条相同的 **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。
- 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。
- 如果 **if-match** 子句对应的 ACL 不存在，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。
- 路由策略应使用非 VPN 的 ACL 进行路由过滤。
- IPv4 路由策略和 IPv6 路由策略在配置 **if-match** 子句时，不同之处在于匹配路由信息的目的地址、下一跳和源地址的命令不同。
- 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。

1.3.4 配置apply子句

表1-8 配置 apply 子句

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*	必选 缺省情况下，不创建路由策略
配置BGP路由信息的AS_PATH属性		**apply as-path** *as-number*&<1-10> [**replace**]	可选 缺省情况下，没有配置BGP路由信息的AS_PATH属性
删除指定的BGP团体属性		**apply comm-list** { *comm-list-number* | *comm-list-name* } **delete**	可选 缺省情况下，不删除路由团体属性
配置BGP路由信息的团体属性		**apply community** { **none** | **additive** | { *community-number*&<1-16> | *aa:nn*&<1-16> | **internet** | **no-advertise** | **no-export** | **no-export-subconfed** } * [**additive**] }	可选 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*	可选 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** }	可选 缺省情况下，没有配置路由开销类型
配置BGP扩展团体属性		**apply extcommunity** { { **rt** *route-target* }&<1-16> [**additive**] | **soo** *site-of-origin* **additive** }	可选 缺省情况下，没有配置BGP扩展团体属性
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*	可选 缺省情况下，没有配置IPv4路由信息的下一跳地址
	配置IPv6路由信息的下一跳地址	**apply ipv6 next-hop** *ipv6-address*	可选 缺省情况下，没有配置IPv6路由信息的下一跳地址
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** }	可选 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*	可选 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**	可选 缺省情况下，没有配置MPLS标签

操作	命令	说明
配置BGP路由信息的ORIGIN属性	**apply origin** { **egp** *as-number* | **igp** | **incomplete** }	可选 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*	可选 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*	可选 缺省情况下，没有配置BGP路由信息的首选值
配置RIP、OSPF、IS-IS路由信息的标记域	**apply tag** *value*	可选 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*] }	可选 缺省情况下，没有配置快速重路由

说明

- 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。
- 对于引入的IPv4路由，使用**apply ip-address next-hop**命令设置下一跳地址无效；对于引入的IPv6路由，使用**apply ipv6 next-hop**命令设置下一跳地址无效。

1.3.5 配置continue子句

表1-9 配置 continue 子句

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*	必选 缺省情况下，不创建路由策略
配置下一个执行节点	**continue** [*node-number*]	可选 缺省情况下，没有配置下一个执行节点 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** 子句只是子句的值不相同时，以最后一个 **apply** 子句为准，包括有叠加属性的 **apply** 子句，当多个节点都配置了这样的子句，属性不会全部叠加到路由上，而是以最后一个子句为准。
- 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 路由策略的显示和维护

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1-10 路由策略的显示和维护

操作	命令
显示BGP AS路径过滤列表信息	**display ip as-path** [*as-path-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示BGP团体属性列表信息	**display ip community-list** [*basic-community-list-number* | *adv-community-list-number* | *comm-list-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示BGP扩展团体属性列表信息	**display ip extcommunity-list** [*ext-comm-list-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv4地址前缀列表的统计信息	**display ip ip-prefix** [*ip-prefix-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6地址前缀列表的统计信息	**display ip ipv6-prefix** [*ipv6-prefix-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示路由策略信息	**display route-policy** [*route-policy-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IPv4地址前缀列表的统计信息	**reset ip ip-prefix** [*ip-prefix-name*]
清除IPv6地址前缀列表的统计信息	**reset ip ipv6-prefix** [*ipv6-prefix-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

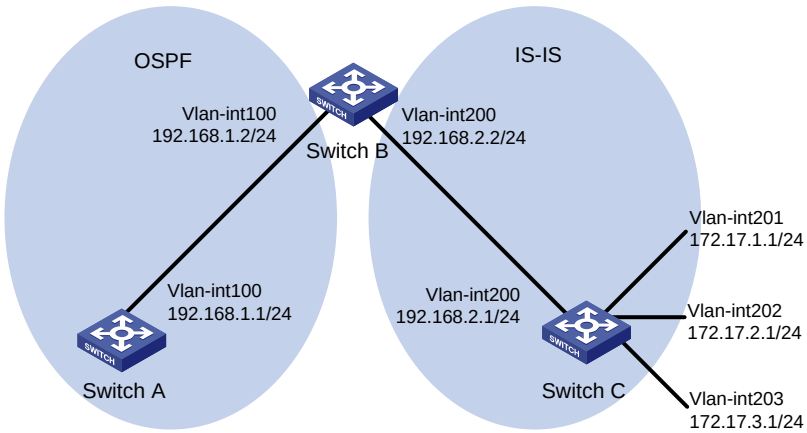
1. 组网需求

如下图所示，Switch B 与 Switch A 之间通过 OSPF 协议交换路由信息，与 Switch C 之间通过 IS-IS 协议交换路由信息。

要求在 Switch 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 Switch C。

```
<SwitchC> system-view
[SwitchC] isis
[SwitchC-isis-1] is-level level-2
[SwitchC-isis-1] network-entity 10.0000.0000.0001.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis enable
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 201
[SwitchC-Vlan-interface201] isis enable
[SwitchC-Vlan-interface201] quit
[SwitchC] interface vlan-interface 202
[SwitchC-Vlan-interface202] isis enable
[SwitchC-Vlan-interface202] quit
[SwitchC] interface vlan-interface 203
[SwitchC-Vlan-interface203] isis enable
[SwitchC-Vlan-interface203] quit
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] isis
[SwitchB-isis-1] is-level level-2
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
```

```
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis enable
[SwitchB-Vlan-interface200] quit
```

(3) 配置 OSPF 路由协议及路由引入

配置 Switch A，启动 OSPF。

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
```

配置 Switch B，启动 OSPF，并引入 IS-IS 路由。

```
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] import-route isis 1
[SwitchB-ospf-1] quit
```

查看 SwitchA 的 OSPF 路由表，可以看到引入的路由。

```
[SwitchA] 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     1562     Stub    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
 192.168.2.0/24     1        Type2   1           192.168.1.2     192.168.2.2

 Total Nets: 5
 Intra Area: 1  Inter Area: 0  ASE: 4  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。

```
[SwitchB] acl number 2002
[SwitchB-acl-basic-2002] rule permit source 172.17.2.0 0.0.0.255
[SwitchB-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。

```
[SwitchB] ip ip-prefix prefix-a index 10 permit 172.17.1.0 24
```

(5) 配置路由策略

```
[SwitchB] route-policy isis2ospf permit node 10
[SwitchB-route-policy] if-match ip-prefix prefix-a
[SwitchB-route-policy] apply cost 100
[SwitchB-route-policy] quit
[SwitchB] route-policy isis2ospf permit node 20
[SwitchB-route-policy] if-match acl 2002
[SwitchB-route-policy] apply tag 20
[SwitchB-route-policy] quit
[SwitchB] route-policy isis2ospf permit node 30
[SwitchB-route-policy] quit
```

(6) 在路由引入时应用路由策略

配置 Switch B，设置在路由引入时应用路由策略。

```
[SwitchB] ospf
[SwitchB-ospf-1] import-route isis 1 route-policy isis2ospf
[SwitchB-ospf-1] quit
```

查看 Switch 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。

```
[SwitchA] 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
 192.168.2.0/24     1        Type2   1           192.168.1.2   192.168.2.2

 Total Nets: 5
 Intra Area: 1  Inter Area: 0  ASE: 4  NSSA: 0
```

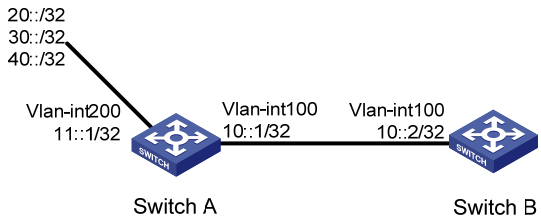
1.5.2 在IPv6 路由引入中应用路由策略

1. 组网需求

- 在 Switch A 和 Switch B 上使能 RIPng。
- 在 Switch A 上配置三条静态路由，并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 Switch 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组网图



3. 配置步骤

(1) 配置 Switch A

配置接口 Vlan-interface100 和 Vlan-interface200 的 IPv6 地址。

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ipv6 address 10::1 32
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ipv6 address 11::1 32
[SwitchA-Vlan-interface200] quit
```

在接口下使能 RIPng。

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ripng 1 enable
[SwitchA-Vlan-interface100] 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。

```
[SwitchA] ipv6 route-static 20:: 32 11::2
[SwitchA] ipv6 route-static 30:: 32 11::2
[SwitchA] ipv6 route-static 40:: 32 11::2
```

配置路由策略。

```
[SwitchA] ip ipv6-prefix a index 10 permit 30:: 32
[SwitchA] route-policy static2ripng deny node 0
[SwitchA-route-policy] if-match ipv6 address prefix-list a
[SwitchA-route-policy] quit
[SwitchA] route-policy static2ripng permit node 10
[SwitchA-route-policy] quit
```

启动 RIPng 协议并引入静态路由。

```
[SwitchA] ripng
[SwitchA-ripng-1] import-route static route-policy static2ripng
```

(2) 配置 Switch B

配置接口 Vlan-interface100 的 IPv6 地址。

```
[SwitchB] ipv6
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ipv6 address 10::2 32
```

在接口下使能 RIPng。

```
[SwitchB-Vlan-interface100] ripng 1 enable
[SwitchB-Vlan-interface100] quit
```

启动 RIPng 协议。

```
[SwitchB] ripng
```

查看 RIPng 路由表。

```
[SwitchB-ripng-1] display ripng 1 route
   Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
 ----------------------------------------------------------------

 Peer FE80::7D58:0:CA03:1  on Vlan-interface 100
 Dest 10::/32,
     via FE80::7D58:0:CA03:1, cost  1, tag 0, A, 18 Sec
 Dest 20::/32,
     via FE80::7D58:0:CA03:1, cost  1, tag 0, A, 8 Sec
 Dest 40::/32,
     via FE80::7D58:0:CA03:1, cost  1, tag 0, A, 3 Sec
```

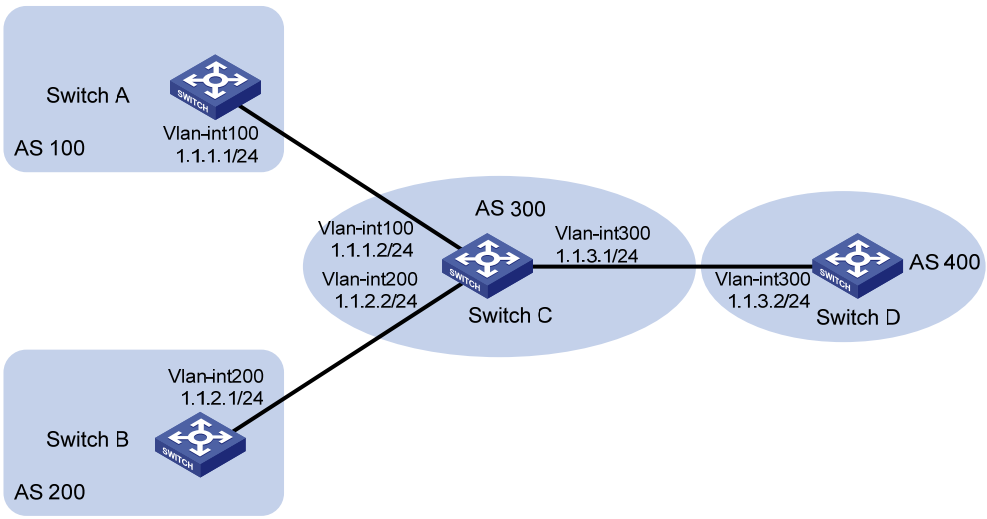
1.5.3 配置应用路由策略过滤接收的BGP路由

1. 组网需求

- 所有交换机均运行 BGP 协议，Switch A、Switch B 和 Switch C 之间建立 EBGP 连接，Switch C 和 Switch D 之间建立 EBGP 连接。
- 通过配置路由策略，使得 Switch D 拒绝接收来自 AS 200 的路由。

2. 组网图

图1-3 配置应用路由策略过滤接收的 BGP 路由



3. 配置步骤

(1) 配置接口 IP 地址（略）

(2) 配置 BGP 基本功能

配置 Switch A。

```
<SwitchA> system-view
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 1.1.1.2 as-number 300
```

配置 Switch B。

```
<SwitchB> system-view
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 1.1.2.2 as-number 300
```

配置 Switch C。

```
<SwitchC> system-view
[SwitchC] bgp 300
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 1.1.1.1 as-number 100
[SwitchC-bgp] peer 1.1.2.1 as-number 200
[SwitchC-bgp] peer 1.1.3.2 as-number 400
```

配置 Switch D。

```
<SwitchD> system-view
[SwitchD] bgp 400
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] peer 1.1.3.1 as-number 300
[SwitchD-bgp] quit
```

配置 Switch A，将网段路由 4.4.4.4/24、5.5.5.5/24、6.6.6.6/24 发布到 BGP 路由表中。

```
[SwitchA-bgp] network 4.4.4.4 24
[SwitchA-bgp] network 5.5.5.5 24
[SwitchA-bgp] network 6.6.6.6 24
```

配置 Switch B，将网段路由 7.7.7.7/24、8.8.8.8/24、9.9.9.9/24 发布到 BGP 路由表中。

```
[SwitchB-bgp] network 7.7.7.7 24
[SwitchB-bgp] network 8.8.8.8 24
[SwitchB-bgp] network 9.9.9.9 24
```

查看 Switch D 的 BGP 路由表。

```
[SwitchD-bgp] display bgp routing-table

 Total Number of Routes: 6

 BGP Local router ID is 4.4.4.4
 Status codes: * - valid, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

 *>  4.4.4.0/24        1.1.3.1                                0       300 100i
 *>  5.5.5.0/24        1.1.3.1                                0       300 100i
 *>  6.6.6.0/24        1.1.3.1                                0       300 100i
 *>  7.7.7.0/24        1.1.3.1                                0       300 200i
 *>  8.8.8.0/24        1.1.3.1                                0       300 200i
```

```
 *>  9.9.9.0/24          1.1.3.1                                0       300 200i
```

从路由表可以看出，Switch D 同时学到了来自 AS 100 的路由 4.4.4.0/24、5.5.5.0/24、6.6.6.0/24 和来自 AS 200 的路由 7.7.7.0/24、8.8.8.0/24、9.9.9.0/24。

(3) 配置 Switch D 拒绝接收来自 AS 200 的路由。

在 Switch D 上配置编号为 1 的 AS 路径过滤列表。

```
[SwitchD] ip as-path 1 permit .*200.*
```

在 Switch D 上配置名称为“rt1”的路由策略，其节点序列号为 1，匹配模式为 deny，即拒绝接收来自 AS 200 的路由。

```
[SwitchD] route-policy rt1 deny node 1
[SwitchD-route-policy] if-match as-path 1
[SwitchD-route-policy] quit
```

在 Switch D 上配置名称为“rt1”的路由策略，其节点序列号为 10，匹配模式为 permit，即可以接收来自其它 AS 的路由。

```
[SwitchD] route-policy rt1 permit node 10
[SwitchD-route-policy] quit
```

在 Switch D 上，配置对从对等体 1.1.3.1 接收的路由指定名称为“rt1”的路由策略。

```
[SwitchD] bgp 400
[SwitchD-bgp] peer 1.1.3.1 route-policy rt1 import
```

查看 Switch D 的 BGP 路由表。

```
[SwitchD-bgp] display bgp routing-table

 Total Number of Routes: 3

 BGP Local router ID is 4.4.4.4
 Status codes: * - valid, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
     Network            NextHop         MED        LocPrf     PrefVal Path/Ogn

 *>  4.4.4.0/24         1.1.3.1                                0       300 100i
 *>  5.5.5.0/24         1.1.3.1                                0       300 100i
 *>  6.6.6.0/24         1.1.3.1                                0       300 100i
```

从路由表可以看出，Switch D 只学到了来自 AS 100 的路由 4.4.4.0/24、5.5.5.0/24、6.6.6.0/24。

1.6 常见错误配置举例

1.6.1 无法实现IPv4 路由信息过滤

1. 错误现象

路由协议运行正常的情况下无法实现路由信息过滤。

2. 分析

地址前缀列表的各个表项中至少应该有一个表项的匹配模式是 **permit** 模式。Route-policy 的各个节点中至少应该有一个节点的匹配模式是 **permit** 模式。

3. 处理过程

(1) 使用 **display ip ip-prefix** 命令查看前缀列表的配置。

(2) 使用 **display route-policy** 命令查看路由策略的配置。

1.6.2 无法实现IPv6 路由信息过滤

1. 错误现象

路由协议运行正常的情况下无法实现路由信息过滤。

2. 分析

地址前缀列表的各个表项中至少应该有一个表项的匹配模式是 **permit** 模式。Route-policy 的各个节点中至少应该有一个节点的匹配模式是 **permit** 模式。

3. 处理过程

(1) 使用 **display ip ipv6-prefix** 命令查看前缀列表的配置。

(2) 使用 **display route-policy** 命令查看路由策略的配置。

目 录

1 策略路由配置 ····· 1-1
1.1 策略路由简介 ····· 1-1
1.1.1 PBR方式 ····· 1-1
1.1.2 QoS策略方式 ····· 1-3
1.2 配置PBR方式的策略路由 ····· 1-3
1.2.1 配置PBR策略 ····· 1-3
1.2.2 应用本地策略路由 ····· 1-4
1.2.3 应用接口策略路由 ····· 1-5
1.2.4 与Track联动 ····· 1-5
1.3 配置QoS策略方式的策略路由 ····· 1-5
1.3.1 配置QoS策略 ····· 1-6
1.3.2 应用QoS策略 ····· 1-6
1.4 策略路由显示和维护 ····· 1-7
1.4.1 PBR方式策略路由显示和维护 ····· 1-7
1.4.2 QoS策略方式策略路由显示和维护 ····· 1-8
1.5 策略路由典型配置举例 ····· 1-9
1.5.1 PBR方式本地策略路由典型配置举例 ····· 1-9
1.5.2 PBR方式接口策略路由典型配置举例 ····· 1-10
1.5.3 QoS策略方式IPv4 策略路由配置举例 ····· 1-12
1.5.4 QoS策略方式IPv6 策略路由配置举例 ····· 1-13

1 策略路由配置

1.1 策略路由简介

策略路由是一种依据用户制定的策略进行路由选择的机制与单纯依照 IP 报文的目的地址查找路由表进行转发不同，策略路由基于到达报文的源地址等信息灵活地进行路由选择。

一般来讲，策略路由的优先级要高于普通路由，即设备在转发报文时，首先将报文与策略路由的匹配规则进行比较，如果符合匹配条件，则按照策略路由进行转发。如果报文无法匹配策略路由的条件，再按照普通路由进行转发。

本系列交换机支持两种策略路由配置方式：

(1) PBR 方式（policy-based-route）

PBR 方式通过 ACL 制定匹配规则，支持对报文的下一跳，优先级及缺省下一跳进行设置，目前只支持 IPv4 单播策略路由。

PBR 方式支持动态修改 ACL 规则，以动态更新匹配规则，实现业务的灵活控制。

说明

- PBR 方式下，如果报文匹配了策略路由条件，但是配置的下一跳地址不存在，报文将继续进行普通路由的转发。
- 对于配置了缺省下一跳的情况，先进行普通路由的转发，如果无法匹配，再进行策略路由转发。

---

(2) QoS 策略方式

QoS 策略方式是通过 QoS 策略的流分类功能来制定细致的匹配规则，并使用流行为中的重定向动作将报文按指定的目的进行转发。QoS 策略方式能够实现 IPv4、IPv6 单播策略路由。

本系列交换机的策略路由是通过 QoS 策略方式实现，即通过流分类功能来制定细致的匹配规则，并使用流行为中的重定向到下一跳动作将报文按指定的目的进行转发。QoS 策略方式能够实现 IPv4、IPv6 单播策略路由。

下面分别对两种方式的实现规则进行介绍。

1.1.1 PBR方式

1. 简介

PBR 方式策略路由对于满足一定条件（ACL 规则）的报文，将执行指定的操作（设置报文的优先级、设置报文的下一跳、设置报文的缺省下一跳等），以指导报文的转发。

根据作用对象的不同，分为本地策略路由和接口策略路由：

- 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）进行策略路由，它只对设备本身产生的报文起作用，对转发的报文不起作用。
- 接口策略路由：对到达该接口的报文进行策略路由，它只对转发的报文起作用，对本地产生的报文不起作用。

对于一般转发和安全等方面的使用需求，大多数情况下只需使用接口策略路由。

说明

本文所指的“接口”为三层口，包括 VLAN 接口、三层以太网端口等。三层以太网端口是指被配置为三层模式的以太网端口，有关以太网端口模式切换的操作，请参见“二层技术-以太网交换配置指导”中的“以太网端口配置”。

2. 相关概念

一个策略用来引入一条路由，对 IP 报文转发进行路由选择。一个策略可以包含一个或多个节点。

(1) 节点

- 每个节点由 *node-number* 来指定。*node-number* 的值越小优先级越高，优先级高的节点优先被执行。
- 每个节点的具体内容由 **if-match** 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** 和 **deny** 两种。

(2) if-match 子句

本系列交换机支持的 **if-match** 子句为 **if-match acl**。

在一个节点中，同一类型的 **if-match** 子句最多只能有一条。

(3) apply 子句

本系列交换机支持三种 **apply** 子句，分别为：**apply ip-precedence**、**apply ip-address next-hop**、**apply ip-address default next-hop**。

说明

执行缺省下一跳命令（**apply ip-address default next-hop**）的前提是，在策略路由中报文没有配置下一跳，或者配置的下一跳无效，并且报文目的 IP 地址在路由表中没有查到相应的路由，这时才会使用策略路由配置的缺省下一跳。

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 1-1 所示。

表1-1 节点的匹配模式、if-match 子句、apply 子句三者之间的关系

节点匹配模式 是否满足 **if-match** 子句	**permit**（允许模式）	**deny**（拒绝模式）
报文满足此节点的所有**if-match**子句	执行此节点**apply**子句，不再匹配下一节点	不执行此节点**apply**子句，不再匹配下一节点，报文按正常转发流程处理
报文不满足此节点的**if-match**子句	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

同一个策略中的各节点之间是“或”的关系，即只要报文通过一个节点的过滤，就意味着通过该策略的过滤；如果报文不能通过一个策略中任何一个节点的过滤，则认为没有通过该策略的过滤，报文将按正常转发流程处理。

1.1.2 QoS策略方式

QoS 策略方式通过流分类来制定匹配规则，并使用“重定向到下一跳”的流行为动作，将满足流分类规则的报文向指定的路径进行转发，从而实现灵活的路由选择。

通过 QoS 策略方式实现的策略路由的优先级要高于普通路由，即报文先匹配流分类规则，如果匹配则按照策略路由进行转发，如果无法匹配所有流分类规则，再按照普通路由进行转发。

说明

关于 QoS 策略的详细介绍，请参见“ACL 和 QoS 配置指导”中的“QoS 配置”。

1.2 配置PBR方式的策略路由

1.2.1 配置PBR策略

表1-2 配置策略

操作	命令	说明
进入系统视图	**system-view**	-
创建策略或一个策略节点，并进入策略路由视图	**policy-based-route** *policy-name* [**deny** | **permit**] **node** *node-number*	必选
设置ACL匹配条件	**if-match acl** *acl-number*	可选
设置报文的优先级	**apply ip-precedence** *value*	可选
设置报文的下一跳	**apply ip-address next-hop** *ip-address* [**direct**] [**track** *track-entry-number*] [*ip-address* [**direct**] [**track** *track-entry-number*]]	可选 用户可以同时配置两个下一跳： • 应用于本地策略路由时，这两个下一跳同时有效，可以起到负载分担的作用 • 应用于接口策略路由时，第一个下一跳作为主路由，第二个下一跳作为备份路由，实现路由备份

操作	命令	说明
设置报文缺省下一跳	**apply ip-address default next-hop** *ip-address* [**track** *track-entry-number*] [*ip-address* [**track** *track-entry-number*]]	可选 用户可以同时配置两个下一跳： • 应用于本地策略路由时，这两个下一跳同时有效，可以起到负载分担的作用 • 应用于接口策略路由时，第一个下一跳作为主路由，第二个下一跳作为备份路由，实现路由备份

说明

- 如果 **if-match** 子句中使用了 ACL，将忽略 ACL 规则的 **permit/deny** 动作以及 **time-range** 指定的规则生效时间段，只使用 ACL 中的分类域来匹配报文。如果使用的 ACL 不存在，则不匹配任何报文。
- **apply ip-address next-hop** 和 **apply ip-address default next-hop** 命令可以配置两个下一跳，可以通过一次配置两个下一跳参数进行，也可以通过两条配置命令完成。当用户希望修改其中一个时，只需要继续执行 **apply ip-address next-hop** *ip-address* 命令，系统就会将最早配置的下一跳（一次配置两个下一跳参数的情况下为命令行输入的第二个参数值）替换；如果先前配置的两个下一跳都需要修改，则直接在 **apply ip-address next-hop** 命令后配置两个参数即可。

注意

- 如果某一节点不配置 **if-match** 子句，则所有报文都会通过该节点的过滤，根据 **permit/deny** 执行相应的操作，不再继续匹配下一节点。
- 如果某一 **permit** 模式的节点不配置 **apply** 子句，当报文满足此节点的所有 **if-match** 子句时，将不会执行任何动作，且不再继续匹配下一节点，报文按正常转发流程处理。
- 如果某一节点没有配置任何 **if-match** 子句和 **apply** 子句，则所有报文都会通过该节点的过滤，但不会执行任何动作，且不再继续匹配下一节点，报文按正常转发流程处理。

1.2.2 应用本地策略路由

在应用本地策略路由时，只能引用一个策略。

表1-3 配置本地策略路由

操作	命令	说明
进入系统视图	**system-view**	-
应用本地策略路由	**ip local policy-based-route** *policy-name*	必选 缺省情况下，没有应用本地策略路由

说明

如果配置时策略 *policy-name* 不存在，命令可以配置成功，但只有当策略 *policy-name* 创建后，本地策略路由才真正生效。

1.2.3 应用接口策略路由

在应用接口策略路由时，一个接口只能引用一个策略。

表1-4 配置接口策略路由

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*	-
应用接口策略路由	**ip policy-based-route** *policy-name*	必选 缺省情况下，没有应用接口策略路由

说明

如果配置时策略 *policy-name* 不存在，命令可以配置成功，但只有当策略 *policy-name* 创建后，本地策略路由才真正生效。

1.2.4 与Track联动

策略路由通过与 Track 联动，增强了应用的灵活性和策略路由对网络环境的动态感知能力。策略路由可以在配置报文的下一跳、缺省下一跳时与 Track 项关联，通过 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 Invalid 时生效。

说明

关于策略路由与 Track 联动的详细介绍和相关配置请参考“可靠性配置指导”中的“Track 配置”。

1.3 配置QoS策略方式的策略路由

配置策略路由功能需要完成以下两个配置步骤：

- 配置 QoS 策略：定义策略路由中匹配的报文规则以及重定向的路径
- 应用 QoS 策略：定义策略路由生效的范围

1.3.1 配置QoS策略

表1-5 配置 QoS 策略

操作	命令	说明
进入系统视图	**system-view**	-
创建流分类并进入流分类视图	**traffic classifier** *tcl-name* [**operator** { **and** | **or** }]	必选
定义匹配报文的规则	**if-match** *match-criteria*	必选
退回系统视图	**quit**	-
创建流行为并进入流行为视图	**traffic behavior** *behavior-name*	必选
配置重定向到下一跳的动作	**redirect next-hop** { *ipv4-add1* [*ipv4-add2*] | *ipv6-add1* [*interface-type interface-number*] [*ipv6-add2* [*interface-type interface-number*]] } [**fail-action** { **discard** | **forward** }]	必选
退回系统视图	**quit**	-
创建QoS策略并进入QoS策略视图	**qos policy** *policy-name*	必选
将流分类与流行为进行关联，即对匹配流分类规则的报文采用流行为中的动作	**classifier** *tcl-name* **behavior** *behavior-name*	必选

说明

如果不配置重定向下一跳失败的处理动作，默认的处理动作是转发。

1.3.2 应用QoS策略

在配置策略路由功能时，用户可以通过下列方式应用 QoS 策略：

- 基于全局应用 QoS 策略：简称为全局策略，QoS 策略对设备上的所有流量生效。
- 基于端口应用 QoS 策略：简称为端口策略，QoS 策略对端口接收的流量生效。
- 基于 VLAN 应用 QoS 策略：简称为 VLAN 策略，QoS 策略对 VLAN 内所有端口接收的流量生效。

说明

用于策略路由的 QoS 策略只能应用到端口/VLAN/全局的入方向。

表1-6 基于全局应用策略

操作	命令	说明
进入系统视图	**system-view**	-
应用全局策略	**qos apply policy** *policy-name* **global inbound**	必选

表1-7 基于端口应用策略

操作		命令	说明
进入系统视图		**system-view**	-
进入以太网端口视图或端口组视图	进入以太网端口视图	**interface** *interface-type interface-number*	二者必选其一 进入以太网端口视图后，下面进行的配置只在当前端口生效；进入端口组视图后，下面进行的配置将在端口组中的所有端口生效
	进入端口组视图	**port-group manual** *port-group-name*	
应用端口策略		**qos apply policy** *policy-name* **inbound**	必选

表1-8 基于 VLAN 应用策略

操作	命令	说明
进入系统视图	**system-view**	-
应用VLAN策略到指定的VLAN	**qos vlan-policy** *policy-name* **vlan** *vlan-id-list* **inbound**	必选

说明

VLAN 策略不能应用在动态 VLAN 上。例如，在运行 GVRP 协议的情况下，设备可能会动态创建 VLAN，相应的 VLAN 策略不能应用在该动态 VLAN 上。

1.4 策略路由显示和维护

1.4.1 PBR方式策略路由显示和维护

在完成 PBR 方式策略路由的配置后，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，用户可以执行 **reset** 命令重置策略路由的统计信息。

表1-9 PBR 方式策略路由显示和维护

操作	命令
显示系统和接口应用的所有策略路由的信息	**display ip policy-based-route** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示系统的策略路由的设置情况	**display ip policy-based-route setup** { **interface** *interface-type interface-number* | **local** | *policy-name* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示已经配置的策略路由	**display policy-based-route** [*policy-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

说明

- 如果只添加一个节点而没有配置任何 **if-match** 子句和 **apply** 子句，则所有报文都匹配，但不执行动作，不再继续往下匹配。策略路由的统计数字将改变。
- 如果添加一个节点，只配置 **if-match** 子句，没有配置 **apply** 子句，则进行匹配，但不执行动作，不再继续往下匹配。策略路由的统计数字将改变。
- 如果添加一个节点，没有配置 **if-match** 子句，只配置 **apply** 子句，则所有报文都匹配，根据 **permit/deny** 执行相应的操作，不再继续往下匹配。策略路由的统计数字将改变。
- 当配置节点的匹配模式为 **deny** 时，如果报文满足该节点的所有 **if-match** 子句，不执行节点的 **apply** 子句，也不再继续往下匹配，数据包将走正常路由表转发，因此没有 **deny** 的调试信息以及相应的统计信息。

1.4.2 QoS策略方式策略路由显示和维护

在完成 QoS 策略方式策略路由的配置后，在任意视图下执行 **display** 命令可以显示策略路由的运行情况，通过查看显示信息验证配置的效果。

表1-10 QoS 策略方式策略路由显示和维护

操作	命令
显示用户定义QoS策略的配置信息	**display traffic classifier user-defined** [*tcl-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定端口或所有端口上策略的配置信息和运行情况	**display qos policy interface** [*interface-type interface-number*] [**inbound** | **outbound**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示VLAN策略信息	**display qos vlan-policy** { **name** *policy-name* | **vlan** *vlan-id* } [**slot** *slot-number*] [**inbound** | **outbound**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示全局策略信息	**display qos policy global** [**slot** *slot-number*] [**inbound** | **outbound**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

1.5 策略路由典型配置举例

1.5.1 PBR方式本地策略路由典型配置举例

1. 组网需求

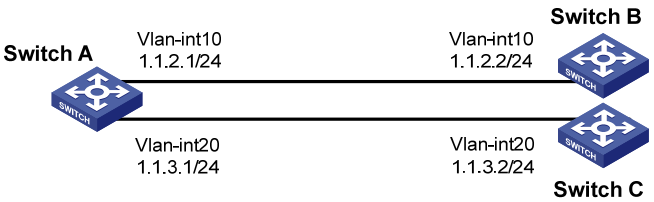
通过策略路由控制从 Switch A 发出的报文：

- 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，Switch A 分别与 Switch B 和 Switch C 直连。Switch B 与 Switch C 路由不可达。

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 Switch A

定义访问控制列表，ACL 3101 匹配 TCP 报文。

```
<SwitchA> system-view
[SwitchA] acl number 3101
[SwitchA-acl-adv-3101] rule permit tcp
[SwitchA-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。

```
[SwitchA] policy-based-route aaa permit node 5
[SwitchA-pbr-aaa-5] if-match acl 3101
[SwitchA-pbr-aaa-5] apply ip-address next-hop 1.1.2.2
[SwitchA-pbr-aaa-5] quit
```

在 Switch A 上应用本地策略路由。

```
[SwitchA] ip local policy-based-route aaa
```

配置接口 Vlan-interface10 和 Vlan-interface20 的 IP 地址。

```
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] ip address 1.1.2.1 255.255.255.0
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 20
[SwitchA-Vlan-interface20] ip address 1.1.3.1 255.255.255.0
```

(2) 配置 Switch B

配置接口 Vlan-interface10 的 IP 地址。

```
<SwitchB> system-view
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ip address 1.1.2.2 255.255.255.0
[SwitchB-Vlan-interface10] quit
```

(3) 配置 Switch C

配置接口 Vlan-interface20 的 IP 地址。

```
<SwitchC> system-view
[SwitchC] interface vlan-interface 20
[SwitchC-Vlan-interface20] ip address 1.1.3.2 255.255.255.0
[SwitchC-Vlan-interface20] quit
```

(4) 验证配置结果

从 Switch A 上 Telnet Switch B（1.1.2.2/24），结果成功。

从 Switch A 上 Telnet Switch C（1.1.3.2/24），结果失败。

```
<SwitchA> telnet 1.1.3.2
Trying 1.1.3.2 ...
Press CTRL+K to abort
Can't connect to the remote host!
```

从 Switch A 上 ping Switch C（1.1.3.2/24），结果成功。

```
<SwitchA> ping 1.1.3.2
  PING 1.1.3.2: 56  data bytes, press CTRL_C to break
    Reply from 1.1.3.2: bytes=56 Sequence=1 ttl=255 time=2 ms
    Reply from 1.1.3.2: bytes=56 Sequence=2 ttl=255 time=1 ms
    Reply from 1.1.3.2: bytes=56 Sequence=3 ttl=255 time=1 ms
    Reply from 1.1.3.2: bytes=56 Sequence=4 ttl=255 time=1 ms
    Reply from 1.1.3.2: bytes=56 Sequence=5 ttl=255 time=1 ms
  --- 1.1.3.2 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 1/1/2 ms
```

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：Switch A 发出的 TCP 报文的下一跳为 1.1.2.2，接口 Vlan-interface20 不发送 TCP 报文，但可以发送非 TCP 报文，策略路由设置成功。

1.5.2 PBR方式接口策略路由典型配置举例

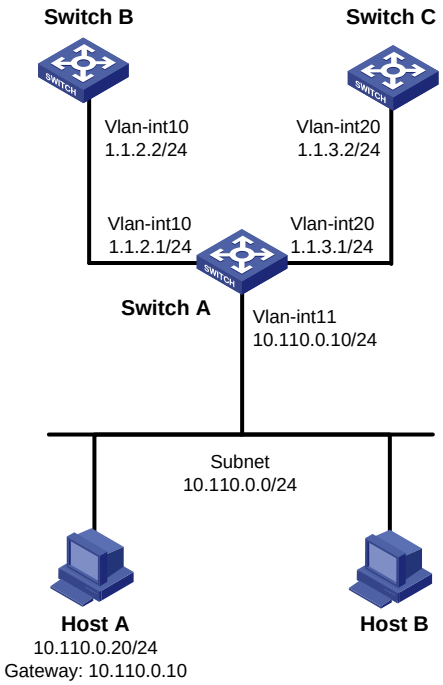
1. 组网需求

通过策略路由控制从 Switch A 的接口 Vlan-interface11 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接口策略路由的配置举例组网图



3. 配置步骤

说明

本例中采用静态路由保证各设备之间路由可达。

(1) 配置 Switch A

定义访问控制列表，ACL 3101 匹配 TCP 报文。

```
<SwitchA> system-view
[SwitchA] acl number 3101
[SwitchA-acl-adv-3101] rule permit tcp
[SwitchA-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。

```
[SwitchA] policy-based-route aaa permit node 5
[SwitchA-pbr-aaa-5] if-match acl 3101
[SwitchA-pbr-aaa-5] apply ip-address next-hop 1.1.2.2
[SwitchA-pbr-aaa-5] quit
```

在接口 Vlan-interface11 上应用接口策略路由，处理此接口接收的报文。

```
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] ip address 10.110.0.10 255.255.255.0
[SwitchA-Vlan-interface11] ip policy-based-route aaa
[SwitchA-Vlan-interface11] quit
```

配置接口 Vlan-interface10 和 Vlan-interface20 的 IP 地址。

```
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] ip address 1.1.2.1 255.255.255.0
```

```
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 20
[SwitchA-Vlan-interface20] ip address 1.1.3.1 255.255.255.0
```

(2) 配置 Switch B

配置到网段 10.110.0.0/24 的静态路由。

```
<SwitchB> system-view
[SwitchB] ip route-static 10.110.0.0 24 1.1.2.1
```

配置接口 Vlan-interface10 的 IP 地址。

```
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ip address 1.1.2.2 255.255.255.0
[SwitchB-Vlan-interface10] quit
```

(3) 配置 Switch C

配置到网段 10.110.0.0/24 的静态路由。

```
<SwitchC> system-view
[SwitchC] ip route-static 10.110.0.0 24 1.1.3.1
```

配置接口 Vlan-interface20 的 IP 地址。

```
[SwitchC] interface vlan-interface 20
[SwitchC-Vlan-interface20] ip address 1.1.3.2 255.255.255.0
[SwitchC-Vlan-interface20] quit
```

(4) 验证配置结果

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 Telnet 与 Switch A 直连的 Switch B（**telnet** 1.1.2.2），结果成功。

从 Host A 上 Telnet 与 Switch A 直连的 Switch C（**telnet** 1.1.3.2），结果失败。

从 Host A 上 ping 与 Switch A 直连的 Switch C（**ping** 1.1.3.2），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 Switch A 的接口 Vlan-interface11 接收的 TCP 报文的下一跳为 1.1.2.2，接口 Vlan-interface20 不转发 TCP 报文，但可以转发非 TCP 报文，策略路由设置成功。

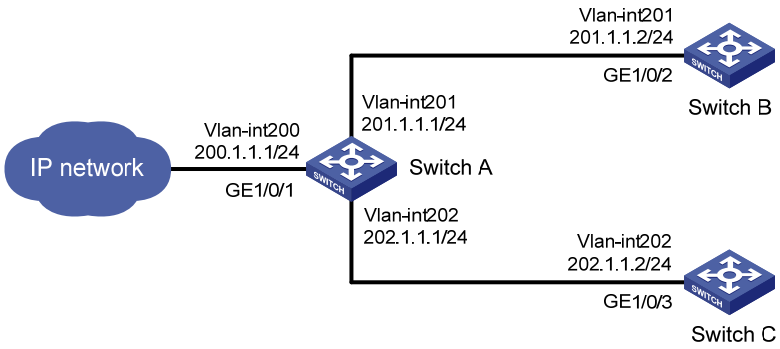
1.5.3 QoS策略方式IPv4 策略路由配置举例

1. 组网需求

控制从设备 Switch A 的 GigabitEthernet1/0/1 端口接收的所有报文，都转发到下一跳 202.1.1.2。

2. 组网图

图1-3 配置通过 QoS 策略实现 IPv4 策略路由组网图



3. 配置步骤

创建 IPv4 ACL，匹配规则为所有报文。

```
<SwitchA> system-view
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule 0 permit source any
[SwitchA-acl-basic-2000] quit
```

创建流分类，并配置流分类规则为匹配 IPv4 ACL 2000 的报文。

```
[SwitchA] traffic classifier a
[SwitchA-classifier-a] if-match acl 2000
[SwitchA-classifier-a] quit
```

创建流行为，并配置动作为将报文重定向到下一跳 202.1.1.2。

```
[SwitchA] traffic behavior a
[SwitchA-behavior-a] redirect next-hop 202.1.1.2
[SwitchA-behavior-a] quit
```

创建 QoS 策略，并将流分类与流行为进行关联。

```
[SwitchA] qos policy a
[SwitchA-qospolicy-a] classifier a behavior a
[SwitchA-qospolicy-a] quit
```

将 QoS 策略应用到端口 GigabitEthernet1/0/1。

```
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] qos apply policy a inbound
```

4. 验证结果

完成上面的配置后，当设备 Switch A 从 GigabitEthernet1/0/1 端口收到目的 IP 地址为 201.1.1.2 的报文时，此报文会被发送到 Switch C，而不会发送到 Switch B，说明策略路由应用成功。

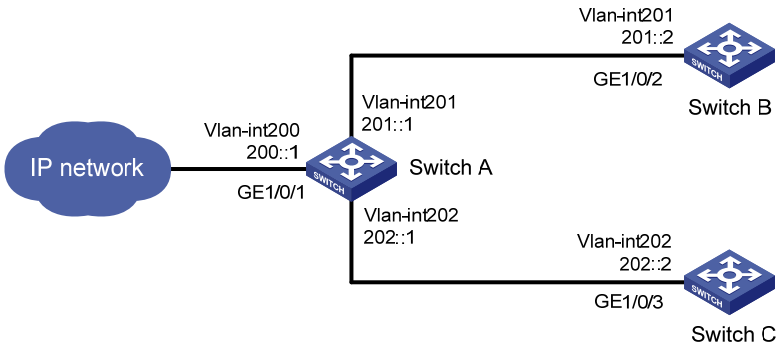
1.5.4 QoS策略方式IPv6 策略路由配置举例

1. 组网需求

控制从设备 Switch A 的 GigabitEthernet1/0/1 端口接收的所有报文，都转发到下一跳 202::2。

2. 组网图

图1-4 配置通过 QoS 策略实现 IPv6 策略路由组网图



3. 配置步骤

创建 IPv6 ACL，匹配规则为所有报文。

```
<SwitchA> system-view
[SwitchA] acl ipv6 number 2000
[SwitchA-acl6-basic-2000] rule 0 permit source any
[SwitchA-acl6-basic-2000] quit
```

#创建流分类，并配置流分类规则为匹配 IPv6 ACL 2000 的报文。

```
[SwitchA] traffic classifier a
[SwitchA-classifier-a] if-match acl ipv6 2000
[SwitchA-classifier-a] quit
```

创建流分类，并配置动作为将流量重定向到下一跳 202::2。

```
[SwitchA] traffic behavior a
[SwitchA-behavior-a] redirect next-hop 202::2
[SwitchA-behavior-a] quit
```

创建 QoS 策略，并将流分类与流行为进行关联。

```
[SwitchA] qos policy a
[SwitchA-qospolicy-a] classifier a behavior a
[SwitchA-qospolicy-a] quit
```

将 QoS 策略应用到端口 GigabitEthernet1/0/1。

```
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] qos apply policy a inbound
```

- 验证结果

完成上面的配置后，当设备 Switch A 从 GigabitEthernet1/0/1 端口收到目的 IP 地址为 201::2 的报文时，此报文会被发送到 Switch C，而不会发送到 Switch B，说明策略路由应用成功。