H3C S5500-HI 系列以太网交换机

三层技术-IP 业务配置指导

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

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 业务配置指导》主要介绍了 IP 业务相关技术的原理及具体配置方法。通过这些技术您可以完成 IP 地址的配置，进行 IP 参数的调整，将 IP 地址解析为以太网 MAC 地址，进行域名与 IP 地址之间的转换，对指定 UDP 端口的 IP 广播报文进行中继转发，实现 IPv4 网络和 IPv6 网络间的互通。

前言部分包含如下内容：

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，新增、修改了部分特性，具体请参见 表 1。
- Release5203 与Release5101 版本相比，新增、修改了部分特性，具体请参见 表 2。

表1 Release5206 与 Release5203 版本间特性差异

配置指导	新增及修改特性
ARP	无
IP地址	变更特性：显示三层接口的IP基本配置信息时，通过增加**description**参数用来显示三层接口的完整描述信息。
DHCP	无
域名解析	无
IRDP	无

配置指导	新增及修改特性
IP性能优化	无
UDP Helper	无
IPv6基础	新增特性：组播ND
DHCPv6	无
IPv6域名解析	无
隧道	无
GRE	无

表2 Release5203 与 Release5101 版本间特性差异

配置指导	新增及修改特性
ARP	新增特性：开启源IP地址冲突提示功能
IP地址	新增特性：支持IP地址的子网掩码长度为31
DHCP	新增特性： • 配置 DHCP 服务器的用户下线检测功能 • 配置 DHCP 服务器、DHCP 中继和 DHCP 客户端发送的 DHCP 报文的 DSCP 优先级
域名解析	新增特性： • 配置 DNS 报文的 DSCP 优先级 • 配置 DNS 报文的源接口
IRDP	无
IP性能优化	无
UDP Helper	新增特性：配置中继转发的私网中的目的服务器
IPv6基础	无
DHCPv6	新增特性： • 配置 DHCPv6 服务器、DHCPv6 中继和 DHCPv6 客户端发送的 DHCPv6 报文的 DSCP 优先级 • 配置 DHCPv6 Snooping 支持 Option 18 和 Option 37
IPv6域名解析	新增特性：配置IPv6 DNS报文的DSCP优先级
隧道	无
GRE	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
	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

大类	资料名称	内容介绍
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 ARP配置

说明

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 ARP简介

1.1.1 ARP作用

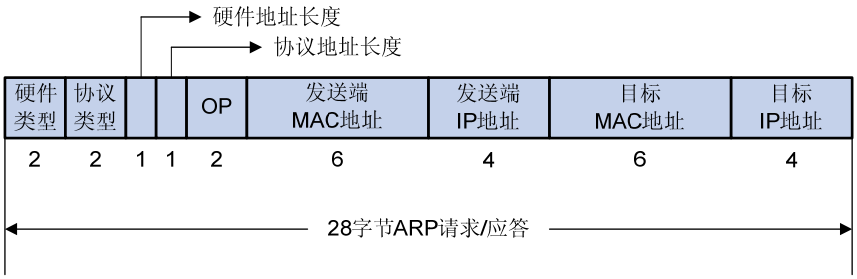
ARP（Address Resolution Protocol，地址解析协议）是将 IP 地址解析为以太网 MAC 地址（或称物理地址）的协议。

在局域网中，当主机或其它网络设备有数据要发送给另一个主机或设备时，它必须知道对方的网络层地址（即 IP 地址）。但是仅仅有 IP 地址是不够的，因为 IP 数据报文必须封装成帧才能通过物理网络发送，因此发送站还必须有接收站的物理地址，所以需要一个从 IP 地址到物理地址的映射。ARP 就是实现这个功能的协议。

1.1.2 ARP报文结构

ARP报文分为ARP请求和ARP应答报文，报文格式如 图 1-1 所示。

图1-1 ARP 报文结构



- 硬件类型：表示硬件地址的类型。它的值为 1 表示以太网地址；
- 协议类型：表示要映射的协议地址类型。它的值为 0x0800 即表示 IP 地址；
- 硬件地址长度和协议地址长度分别指出硬件地址和协议地址的长度，以字节为单位。对于以太网上 IP 地址的 ARP 请求或应答来说，它们的值分别为 6 和 4；
- 操作类型（OP）：1 表示 ARP 请求，2 表示 ARP 应答；
- 发送端 MAC 地址：发送方设备的硬件地址；
- 发送端 IP 地址：发送方设备的 IP 地址；
- 目标 MAC 地址：接收方设备的硬件地址。

- 目标 IP 地址：接收方设备的 IP 地址。

1.1.3 ARP地址解析过程

假设主机A和B在同一个网段，主机A要向主机B发送信息。如 图 1-2 所示，具体的地址解析过程如下：

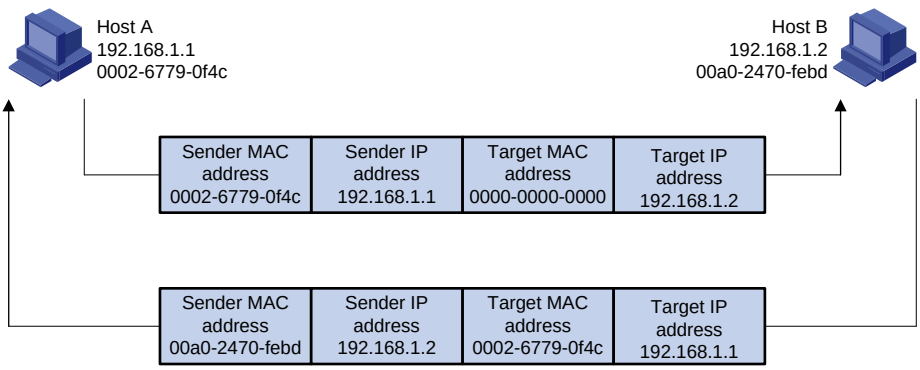
(1) 主机 A 首先查看自己的 ARP 表，确定其中是否包含有主机 B 对应的 ARP 表项。如果找到了对应的 MAC 地址，则主机 A 直接利用 ARP 表中的 MAC 地址，对 IP 数据包进行帧封装，并将数据包发送给主机 B。

(1) 如果主机 A 在 ARP 表中找不到对应的 MAC 地址，则将缓存该数据报文，然后以广播方式发送一个 ARP 请求报文。ARP 请求报文中的发送端 IP 地址和发送端 MAC 地址为主机 A 的 IP 地址和 MAC 地址，目标 IP 地址和目标 MAC 地址为主机 B 的 IP 地址和全 0 的 MAC 地址。由于 ARP 请求报文以广播方式发送，该网段上的所有主机都可以接收到该请求，但只有被请求的主机（即主机 B）会对该请求进行处理。

(2) 主机 B 比较自己的 IP 地址和 ARP 请求报文中的目标 IP 地址，当两者相同时进行如下处理：将 ARP 请求报文中的发送端（即主机 A）的 IP 地址和 MAC 地址存入自己的 ARP 表中。之后以单播方式发送 ARP 响应报文给主机 A，其中包含了自己的 MAC 地址。

(3) 主机 A 收到 ARP 响应报文后，将主机 B 的 MAC 地址加入到自己的 ARP 表中以用于后续报文的转发，同时将 IP 数据包进行封装后发送出去。

图1-2 ARP 地址解析过程



当主机 A 和主机 B 不在同一网段时，主机 A 就会先向网关发出 ARP 请求，ARP 请求报文中的目标 IP 地址为网关的 IP 地址。当主机 A 从收到的响应报文中获得网关的 MAC 地址后，将报文封装并发给网关。如果网关没有主机 B 的 ARP 表项，网关会广播 ARP 请求，目标 IP 地址为主机 B 的 IP 地址，当网关从收到的响应报文中获得主机 B 的 MAC 地址后，就可以将报文发给主机 B；如果网关已经有主机 B 的 ARP 表项，网关直接把报文发给主机 B。

1.1.4 ARP表

设备通过 ARP 解析到目的 MAC 地址后，将会在自己的 ARP 表中增加 IP 地址到 MAC 地址的映射表项，以用于后续到同一目的地报文的转发。

ARP 表项分为动态 ARP 表项和静态 ARP 表项。

1. 动态ARP表项

动态 ARP 表项由 ARP 协议通过 ARP 报文自动生成和维护，可以被老化，可以被新的 ARP 报文更新，可以被静态 ARP 表项覆盖。当到达老化时间、接口 down 时会删除相应的动态 ARP 表项。

2. 静态ARP表项

静态 ARP 表项通过手工配置和维护，不会被老化，不会被动态 ARP 表项覆盖。

配置静态 ARP 表项可以增加通信的安全性。静态 ARP 表项可以限制和指定 IP 地址的设备通信时只使用指定的 MAC 地址，此时攻击报文无法修改此表项的 IP 地址和 MAC 地址的映射关系，从而保护了本设备和指定设备间的正常通信。

静态 ARP 表项分为长静态 ARP 表项、短静态 ARP 表项和多端口 ARP 表项。

- 在配置长静态 ARP 表项时，除了配置 IP 地址和 MAC 地址项外，还必须配置该 ARP 表项所在 VLAN 和出接口。长静态 ARP 表项可以直接用于报文转发。
- 在配置短静态 ARP 表项时，只需要配置 IP 地址和 MAC 地址项。如果出接口是三层以太网端口，短静态 ARP 表项可以直接用于报文转发；如果出接口是 VLAN 接口，短静态 ARP 表项不能直接用于报文转发，当要发送 IP 数据包时，先发送 ARP 请求报文，如果收到的响应报文中的源 IP 地址和源 MAC 地址与所配置的 IP 地址和 MAC 地址相同，则将接收 ARP 响应报文的接口加入该静态 ARP 表项中，之后就可以用于 IP 数据包的转发。
- 多端口 ARP 表项通过配置短静态 ARP 表项和组播 MAC 地址表项形成，当短静态 ARP 表项中的 MAC 地址与组播 MAC 地址表项中的 MAC 地址相同时，则生成多端口 ARP 表项。当设备要发送 IP 数据包时，多端口 ARP 表项将指导 IP 数据包从多个出端口发送。

说明

- 一般情况下，ARP 动态执行并自动寻求 IP 地址到以太网 MAC 地址的解析，无需管理员的介入。
- 当希望设备和指定用户只能使用某个固定的 IP 地址和 MAC 地址通信时，可以配置短静态 ARP 表项，当进一步希望限定这个用户只在某 VLAN 内的某个特定接口上连接时就可以配置长静态 ARP 表项。

1.2 配置ARP

1.2.1 手工添加静态ARP表项

静态 ARP 表项在设备正常工作时间一直有效，当设备的 ARP 表项所对应的 VLAN 或 VLAN 接口被删除时，如果是长静态 ARP 表项则被删除，如果是已经解析的短静态 ARP 表项则重新变为未解析状态。

表1-1 手工添加静态 ARP 表项

操作		命令	说明
进入系统视图		**system-view**	-
手工添加静态ARP表	手工添加长静态ARP表项	**arp static** *ip-address mac-address vlan-id interface-type interface-number* [**vpn-instance** *vpn-instance-name*]	两者必选其一

操作		命令	说明
项	手工添加短静态ARP表项	**arp static** *ip-address mac-address* [**vpn-instance** *vpn-instance-name*]	

注意

- 参数 *vlan-id* 用于指定 ARP 表项所对应的 VLAN，*vlan-id* 必须是用户已经创建好的 VLAN 的 ID，且 *vlan-id* 参数后面指定的以太网端口必须属于这个 VLAN。VLAN 对应的 VLAN 接口必须已经创建。
- 指定参数 *vlan-id* 和 *ip-address* 的情况下，参数 *vlan-id* 对应的 VLAN 接口的 IP 地址必须和参数 *ip-address* 指定的 IP 地址属于同一网段。

1.2.2 手工添加多端口ARP表项

多端口 ARP 表项由组播 MAC 地址表项指定 VLAN 和出端口，由短静态 ARP 表项指定 VPN 和 IP 地址。多端口 ARP 表项不会被其它动态、短静态或长静态 ARP 表项覆盖。

表1-2 手工添加多端口 ARP 表项

操作	命令	说明
进入系统视图	**system-view**	-
配置组播MAC地址表项	**mac-address multicast** *mac-address* **interface** *interface-list* **vlan** *vlan-id*	必选
手工添加短静态ARP表项	**arp static** *ip-address mac-address* [**vpn-instance** *vpn-instance-name*]	必选 其中 *mac-address* 应该和组播 MAC 中的 *mac-address* 一致

注意

- 当对应的 VLAN 接口未创建、VLAN 接口 down 或者 VLAN 接口和 VPN 不匹配的时候，该多端口 ARP 表项不能正常指导转发，匹配该多端口 ARP 表项的报文被丢弃；当满足 VLAN 接口 up 且 VLAN 接口和 VPN 匹配时，该多端口 ARP 表项能正常指导转发。
- 短静态 ARP 按照 MAC 地址查找对应的组播 MAC 地址表项，当同样的 MAC 在多个 VLAN 中存在的时候，短静态 ARP 关联其中 VLAN ID 最小的 MAC 生成多端口 ARP 表项。
- 当达到多端口 ARP 表项的最大数目时，即使短静态 ARP 存在匹配的组播 MAC 地址表项，仍然会解析失败。如果后续其它多端口 ARP 表项被删掉，短静态 ARP 表项不能自动解析生成多端口 ARP 表项，需要手动重新配置。
- 组播 MAC 的相关内容，请参见“IP 组播命令参考/组播路由与转发”中的命令 **mac-address multicast**。

1.2.3 配置接口学习动态ARP表项的最大数目

设备可以通过 ARP 协议自动生成动态 ARP 表项。为了防止部分接口下的用户占用过多的 ARP 资源，可以通过设置接口学习动态 ARP 表项的最大数目来进行限制。当接口学习动态 ARP 表项的最大数目达到所设置的值时，该接口将不再学习动态 ARP 表项。

表1-3 配置接口学习动态 ARP 表项的最大数目

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*	-
配置接口允许学习动态ARP表项的最大数目	**arp max-learning-num** *number*	可选 缺省情况下，二层接口不对允许学习动态ARP表项的最大数目进行限制，对于S5500-28SC-HI和S5500-52SC-HI交换机的三层接口允许学习动态ARP表项的最大数目为8192；对于本系列交换机其它机型的三层接口允许学习动态ARP表项的最大数目为16384 当配置接口允许学习动态ARP表项的最大数目为0时，表示禁止接口学习动态ARP表项

说明

如果二层接口及其所属的 VLAN 三层接口都配置了允许学习动态 ARP 表项的最大数目，则以数目小的进行限制。

1.2.4 配置动态ARP表项的老化时间

为适应网络的变化，ARP 表需要不断更新。ARP 表中的动态 ARP 表项并非永远有效，每一条记录都有一个老化时间，到达老化时间仍得不到刷新的记录将被从 ARP 表中删除。如果在到达老化时间前记录被刷新，则重新计算老化时间。用户可以根据网络实际情况调整老化时间。

表1-4 配置动态 ARP 的老化时间

操作	命令	说明
进入系统视图	**system-view**	-
配置动态ARP的老化时间	**arp timer aging** *aging-time*	可选 缺省情况下，动态ARP的老化时间为20分钟

1.2.5 使能动态ARP表项的检查功能

动态 ARP 表项检查功能可以控制设备上是否可以学习 MAC 地址为组播 MAC 的动态 ARP 表项。

- 使能 ARP 表项的检查功能后，设备上不能学习 MAC 地址为组播 MAC 的动态 ARP 表项。
- 关闭 ARP 表项的检查功能后，设备上可以学习 MAC 地址为组播 MAC 的动态 ARP 表项。

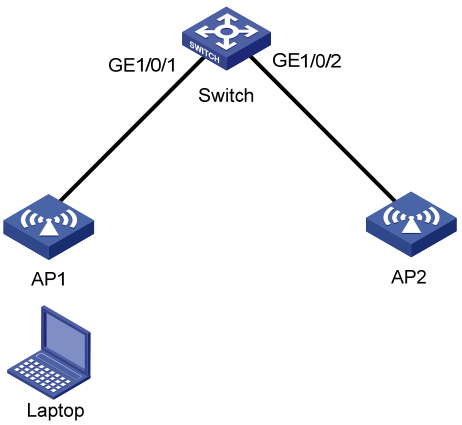
表1-5 使能动态 ARP 表项的检查功能

操作	命令	说明
进入系统视图	**system-view**	-
使能动态ARP表项的检查功能	**arp check enable**	可选 缺省情况下，使能动态ARP表项的检查功能

1.2.6 配置快速更新ARP表项

如 图 1-3 所示，Laptop经常在无线站点AP1 和AP2 之间漫游，导致Switch上记录的Laptop的MAC地址与出端口的对应关系经常发生改变，但是Switch上的ARP表项不会立即更新，影响到数据业务的正常转发。

图1-3 快速更新 ARP 表项应用环境



使能快速更新 ARP 表项功能后，如果交换机上记录的 MAC 地址与出端口的对应关系发生改变，系统会立刻更新 ARP 表项，保证了数据业务的不间断转发。

表1-6 配置快速更新 ARP 表项

操作	命令	说明
进入系统视图	**system-view**	-
使能快速更新ARP表项功能	**mac-address station-move quick-notify enable**	可选 缺省情况下，没有使能快速更新ARP表项功能

说明

快速更新 ARP 表项功能通常用于无线组网环境中，在其它组网环境中建议用户不要使用该功能。

1.2.7 配置组播ARP

微软的网络负载均衡（NLB，Network Load Balancing）功能，是其在 Windows Server 上开发的一个多服务器集群负载均衡特性。

NLB 支持集群内服务器之间的负载分担以及冗余备份，当发生服务器故障时可以支持数据快速切换。为了保证快速切换，NLB 要求交换机将业务流量转发至集群内的所有服务器或指定服务器，然后由各服务器将该服务器不期望的流量过滤掉，因此对于那些使用 Windows Server 作为服务器操作系统的中、小型数据中心来说，交换机与 NLB 的协同工作非常重要。

为了让业务流量能够被转发到所有服务器或指定服务器，微软 NLB 采取了一些处理机制，包括单播模式处理机制、组播模式处理机制和 IGMP 组播模式处理机制。

- 单播模式：在单播模式下，NLB 重新为每个 NLB 节点分配一个共同的 MAC 地址（该 MAC 地址为集群 MAC 地址），并且在发送时修改数据报文的源 MAC 地址，从而使交换机不能将集群 MAC 地址学习到 MAC 地址表中，这样目的地址为集群 MAC 地址的数据报文将作为未知单播报文在交换机的所有端口上进行转发。
- 组播模式：在组播模式下，NLB 使用一个组播地址（该 MAC 地址为虚拟 MAC）用于 NLB 的通信，例如使用 0300-5e11-1111 来作为 NLB 节点的虚 MAC 地址。
- IGMP 组播模式：IGMP 组播模式和组播模式的区别在于：IGMP 组播模式可以通过 IGMP 协议使交换机只将数据报文发送到连接 NLB 节点的端口，而不是所有端口。

说明

- 目前，本系列交换机支持和 Windows Server 的 NLB 协同工作。
- 组播 ARP 功能仅适用于 NLB 采用组播模式处理机制的情况。
- 关于 NLB 的详细介绍请参见 Windows Server 的相关文档。

表1-7 配置组播 ARP

操作	说明	详细配置
关闭ARP表项的检查功能	**undo arp check enable**	必选
手工添加静态ARP表项	**arp static** *ip-address mac-address vlan-id interface-type interface-number* [**vpn-instance** *vpn-instance-name*]	可选
配置静态组播MAC地址表项	**mac-address multicast** *mac-address* **interface** *interface-list* **vlan** *vlan-id*	必选

说明

mac-address multicast 命令的详细介绍请参见“IP 组播命令参考/IGMP Snooping 配置命令”。

1.3 ARP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 ARP 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** 命令清除 ARP 表中的 ARP 表项。

表1-8 ARP 显示和维护

操作	命令
显示ARP表项	**display arp** [[**all** | **dynamic** | **static**] [**slot** *slot-number*] | **vlan** *vlan-id* | **interface** *interface-type interface-number*] [**count** | **verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定IP地址的ARP表项	**display arp** *ip-address* [**slot** *slot-number*] [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定VPN实例的ARP表项	**display arp vpn-instance** *vpn-instance-name* [**count**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示动态ARP表项的老化时间	**display arp timer aging** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除ARP表项	**reset arp** { **all** | **dynamic** | **static** | **slot** *slot-number* | **interface** *interface-type interface-number* }

说明

清除 ARP 表项，将取消 IP 地址和 MAC 地址的映射关系，可能导致设备无法正常通信。清除前请务必仔细确认。

1.4 ARP典型配置举例

1.4.1 静态ARP表项配置举例

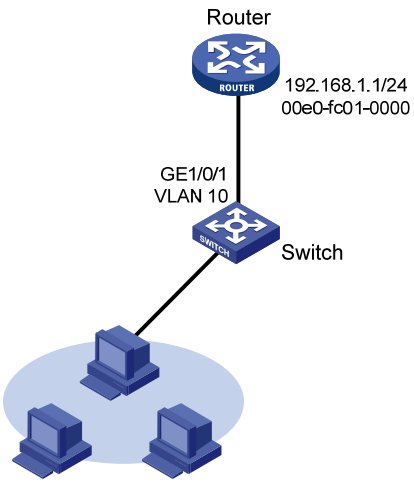
1. 组网需求

- Switch 连接主机，通过接口 GigabitEthernet1/0/1 连接 Router。接口 GigabitEthernet1/0/1 属于 VLAN 10。
- Router 的 IP 地址为 192.168.1.1/24，MAC 地址为 00e0-fc01-0000。

网络管理员希望通过某种方法来防止恶意用户对 Switch 进行 ARP 攻击，增加 Switch 和 Router 通信的安全性。如果 Router 的 IP 地址和 MAC 地址是固定的，则可以通过在 Switch 上配置静态 ARP 表项的方法，防止恶意用户进行 ARP 攻击。

2. 组网图

图1-4 配置静态 ARP 表项组网图



3. 配置步骤

在 Switch 上进行下列配置。

创建 VLAN 10。

```
<Switch> system-view
[Switch] vlan 10
[Switch-vlan10] quit
```

将接口 GigabitEthernet1/0/1 加入到 VLAN 10 中。

```
[Switch] interface GigabitEthernet 1/0/1
[Switch-GigabitEthernet1/0/1] port link-type trunk
[Switch-GigabitEthernet1/0/1] port trunk permit vlan 10
[Switch-GigabitEthernet1/0/1] quit
```

创建接口 Vlan-interface10，并配置 IP 地址。

```
[Switch] interface vlan-interface 10
[Switch-vlan-interface10] ip address 192.168.1.2 24
[Switch-vlan-interface10] quit
```

配置一条静态 ARP 表项，IP 地址为 192.168.1.1，对应的 MAC 地址为 00e0-fc01-0000，此条 ARP 表项对应的出接口为属于 VLAN 10 的接口 GigabitEthernet1/0/1。

```
[Switch] arp static 192.168.1.1 00e0-fc01-0000 10 GigabitEthernet 1/0/1
```

查看静态 ARP 表项信息。

```
[Switch] display arp static
                Type: S-Static    D-Dynamic    A-Authorized
IP Address       MAC Address     VLAN ID  Interface             Aging Type
192.168.1.1      00e0-fc01-0000  10       GE1/0/1                N/A   S
```

1.4.2 组播ARP配置举例

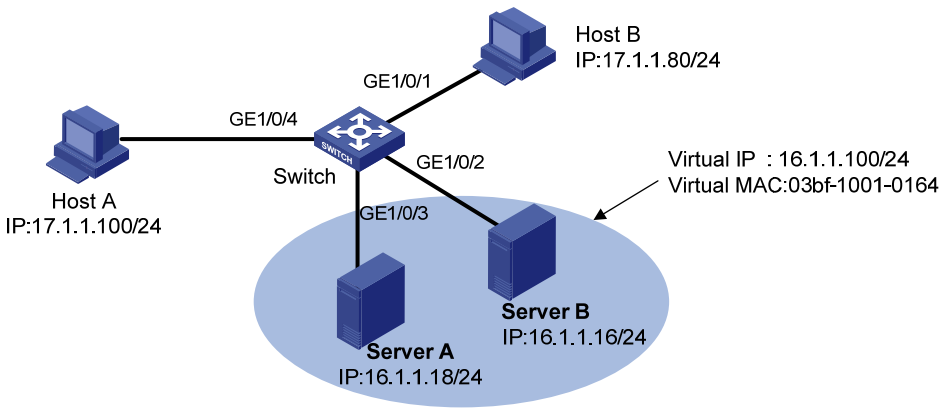
1. 组网需求

在一个小型数据中心里，采用微软的 NLB（Network Loadbalancing）功能的组播模式处理机制，为了使交换机能够和 NLB 协同工作，在交换机上进行如下配置：

- 端口 GigabitEthernet1/0/2 和 GigabitEthernet1/0/3 属于 Vlan 1，Vlan-interface1 的 IP 地址为 16.1.1.30/24；
- 端口 GigabitEthernet1/0/1 和 GigabitEthernet1/0/4 属于 Vlan 2，Vlan-interface2 的 IP 地址为 17.1.1.1/24；
- Host A 和 Host B 的网关的 IP 地址为 17.1.1.1/24；
- Server A 和 Srever B 的网关的 IP 地址为 16.1.1.30/24。
- 关闭 ARP 表项检查功能。关闭 ARP 表项的检查功能后，可以学习 MAC 地址为组播 MAC 的动态 ARP 表项；
- 为了保证除 Switch 的 GigabitEthernet1/0/2 和 GigabitEthernet1/0/3 以外的其他端口不会接收到相关报文，在 Switch 上手工添加静态组播 MAC 表项。

2. 组网图

图1-5 组播 ARP 配置举例



3. 配置步骤

说明

- 本配置举例仅给出交换机上的配置，有关服务器上 NLB 的配置请参见 Windows Server 的相关手册。
- 该配置举例仅适用于 NLB 的组播模式，假设 Server 的虚 IP 地址为 16.1.1.100/24，虚 MAC 地址为 03bf-1001-0164。

- 配置 Switch

配置接口 Vlan-interface2 的 IP 地址。

```
<Switch> system-view
[Switch] vlan 2
[Switch-vlan2] port GigabitEthernet 1/0/4
[Switch-vlan2] port GigabitEthernet 1/0/1
[Switch-vlan2] quit
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ip address 17.1.1.1 255.255.255.0
[Switch-Vlan-interface2] quit
```

配置接口 Vlan-interface1 的 IP 地址。

```
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 16.1.1.30 255.255.255.0
[Switch-Vlan-interface1] quit
```

#关闭 ARP 表项检查功能。

```
[Switch] undo arp check enable
```

手工添加静态组播 MAC 表项。

```
[Switch] mac-address multicast 03bf-1001-0164 interface GigabitEthernet 1/0/2 Gigabi
tEthernet 1/0/3 vlan 1
```

4. 验证结果

- NLB 负载分担功能测试：分别在 Server A 和 Server B 上启动 FTP Server，两台 host 向虚拟 IP 发起登录请求，登录在不同的 Server 上；
- NLB 冗余备份功能测试：禁用 Server A 的网卡，两台 host 向虚拟 IP 发起登录请求，应均登录到 Server B 的 FTP Server 上。

2 免费ARP配置

2.1 免费ARP简介

免费 ARP 报文是一种特殊的 ARP 报文，该报文中携带的发送端 IP 地址和目标 IP 地址都是本机 IP 地址，报文源 MAC 地址是本机 MAC 地址，报文的目的 MAC 地址是广播地址。

设备通过对外发送免费 ARP 报文来实现以下功能：

- 确定其它设备的 IP 地址是否与本机的 IP 地址冲突。当其它设备收到免费 ARP 报文后，如果发现报文中的 IP 地址和自己的 IP 地址相同，则给发送免费 ARP 报文的设备返回一个 ARP 应答，告知该设备 IP 地址冲突。
- 设备改变了硬件地址，通过发送免费 ARP 报文通知其它设备更新 ARP 表项。

1. 免费ARP报文学习功能的作用

使能了免费 ARP 报文学习功能后，设备会根据收到的免费 ARP 报文中携带的信息（源 IP 地址、源 MAC 地址）对自身维护的 ARP 表进行修改。设备先判断 ARP 表中是否存在与此免费 ARP 报文源 IP 地址对应的 ARP 表项：

- 如果没有对应的 ARP 表项，设备会根据该免费 ARP 报文中携带的信息新建 ARP 表项；
- 如果存在对应的 ARP 表项，设备会根据该免费 ARP 报文中携带的信息更新对应的 ARP 表项。

关闭免费 ARP 报文学习功能后，设备不会根据收到的免费 ARP 报文来新建 ARP 表项，但是会更新已存在的对应 ARP 表项。如果用户不希望通过免费 ARP 报文来新建 ARP 表项，可以关闭免费 ARP 报文学习功能，以节省 ARP 表项资源。

2. 定时发送免费ARP功能的作用

定时发送免费 ARP 功能可以及时通知下行设备更新 ARP 表项或者 MAC 地址表项，主要应用场景如下：

(1) 防止仿冒网关的 ARP 攻击

如果攻击者仿冒网关发送免费 ARP 报文，就可以欺骗同网段内的其它主机，使得被欺骗的主机访问网关的流量，被重定向到一个错误的 MAC 地址，导致其它主机用户无法正常访问网络。

为了尽量避免这种仿冒网关的 ARP 攻击，可以在网关的接口上使能定时发送免费 ARP 功能。使能该功能后，网关接口上将按照配置的时间间隔周期性发送接口主 IP 地址和手工配置的从 IP 地址的免费 ARP 报文。这样，每台主机都可以学习到正确的网关，从而正常访问网络。

(2) 防止主机 ARP 表项老化

在实际环境中，当网络负载较大或接收端主机的 CPU 占用率较高时，可能存在 ARP 报文被丢弃或主机无法及时处理接收到的 ARP 报文等现象。这种情况下，接收端主机的动态 ARP 表项会因超时而老化，在其重新学习到发送设备的 ARP 表项之前，二者之间的流量就会发生中断。

为了解决上述问题，可以在网关的接口上使能定时发送免费 ARP 功能。使能该功能后，网关接口上将按照配置的时间间隔周期性发送接口主 IP 地址和手工配置的从 IP 地址的免费 ARP 报文。这样，接收端主机可以及时更新 ARP 映射表，从而防止了上述流量中断现象。

(3) 防止 VRRP 虚拟 IP 地址冲突

当网络中存在 VRRP 备份组时，需要由 VRRP 备份组的 Master 路由器周期性的向网络内的主机发送免费 ARP 报文，使主机更新本地 ARP 地址表，从而确保网络中不会存在 IP 地址与 VRRP 虚拟 IP 地址相同的设备。

由于用户可以设定 VRRP 虚拟 IP 地址和 MAC 地址对应关系，因此有以下两种情况：

- 如果当前 VRRP 虚拟 IP 地址和虚拟 MAC 地址对应，则免费 ARP 报文中的源 MAC 地址为 VRRP 虚拟路由器对应的虚拟 MAC 地址。
- 如果当前 VRRP 虚拟 IP 地址和实际 MAC 地址对应，则免费 ARP 报文中的源 MAC 地址为 VRRP 备份组中 Master 路由器接口的 MAC 地址。

说明

关于 VRRP 的详细介绍，请参见“可靠性配置指导”中的“VRRP”。

2.2 配置免费ARP

表2-1 配置免费 ARP

操作	命令	说明
进入系统视图	**system-view**	-
使能免费ARP报文学习功能	**gratuitous-arp-learning enable**	可选 缺省情况下，设备免费ARP报文的学习功能处于开启状态
使能收到非同一网段ARP请求时发送免费ARP报文功能	**gratuitous-arp-sending enable**	可选 缺省情况下，设备收到非同一网段的ARP请求时不发送免费ARP报文
进入接口视图	**interface** *interface-type interface-number*	-
使能定时发送免费ARP功能，并设置发送免费ARP报文的周期	**arp send-gratuitous-arp** [**interval** *milliseconds*]	可选 缺省情况下，定时发送免费ARP功能处于关闭状态

说明

- 设备最多允许同时在 1024 个接口上使能定时发送免费 ARP 功能。
- 配置定时发送免费 ARP 功能后，只有当接口链路 up 并且配置 IP 地址后，此功能才真正生效。
- 如果修改了免费 ARP 报文的发送周期，则在下一个发送周期才能生效。
- 如果同时在很多接口下使能定时发送免费 ARP 功能，或者每个接口有大量的从 IP 地址，或者两种情况共存的同时又配置很小的发送时间间隔，那么免费 ARP 报文的发送频率可能会远远低于用户的预期。

2.3 开启源IP地址冲突提示功能

交换机接收到其它设备发送的免费 ARP 报文后，如果发现报文中的源 IP 地址和自己的 IP 地址相同，交换机会根据当前源 IP 地址冲突提示功能的状态，进行如下处理：

- 如果源 IP 地址冲突提示功能处于关闭状态，则发送一个免费 ARP 报文，在收到对应的 ARP 应答后才提示存在 IP 地址冲突，并在提示存在 IP 地址冲突的 5 秒钟后，再次进行该项操作（发送免费 ARP 报文，在收到应答报文后提示存在 IP 地址冲突），这个过程会反复进行，直到冲突结束为止。
- 如果源 IP 地址冲突提示功能处于开启状态，则不会发送免费 ARP 报文，而是直接提示存在 IP 地址冲突。如果持续收到冲突的免费 ARP 报文，交换机将会每隔 30 秒钟，提示一次存在 IP 地址冲突，直到冲突结束为止。

表2-2 开启源 IP 地址冲突提示功能

操作	命令	说明
进入系统视图	**system-view**	-
开启源IP地址冲突提示功能	**arp ip-conflict prompt**	可选 缺省情况下，源IP地址冲突提示功能处于关闭状态

3 代理ARP配置

3.1 代理ARP简介

如果 ARP 请求是从一个网络的主机发往同一网段却不在同一物理网络上的另一台主机，那么连接它们的具有代理 ARP 功能的设备就可以回答该请求，这个过程称作代理 ARP（Proxy ARP）。

代理 ARP 功能屏蔽了分离的物理网络这一事实，使用户使用起来，好像在同一个物理网络上。

代理 ARP 分为普通代理 ARP 和本地代理 ARP，二者的应用场景有所区别：

- 普通代理 ARP 的应用环境为：想要互通的主机分别连接到设备的不同三层接口上，且这些主机不在同一个广播域中。
- 本地代理 ARP 的应用环境为：想要互通的主机连接到设备的同一个三层接口上，且这些主机不在同一个广播域中。

说明

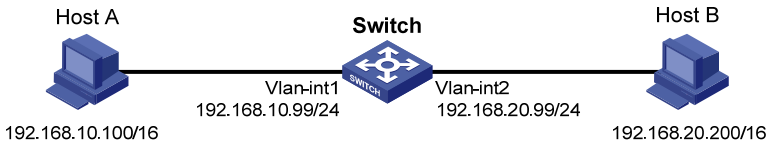
如无特殊说明，本章后续描述中的代理 ARP 均指普通代理 ARP。

3.1.1 代理ARP

处于同一网段内的主机，当连接到设备的不同三层接口时，可以利用设备的代理 ARP 功能，通过三层转发实现互通。

代理ARP的典型应用环境如 图 3-1 所示。设备Switch通过两个VLAN接口Vlan-interface1 和Vlan-interface2 连接两个网络，两个VLAN接口的IP地址不在同一个网段，接口地址分别为192.168.10.99/24、192.168.20.99/24。但是两个网络内的主机Host A和Host B的地址通过掩码的控制，既与相连设备的接口地址在同一网段，同时二者也处于同一个网段。

图3-1 代理 ARP 的应用环境



在这种组网情况下，当 Host A 需要与 Host B 通信时，由于目的 IP 地址与本机的 IP 地址为同一网段，因此 Host A 会广播发送 ARP 请求报文，请求 Host B 的 MAC 地址。但是，此时的两台主机处于不同的广播域中，Host B 无法收到 Host A 的 ARP 请求报文，当然也就无法应答。

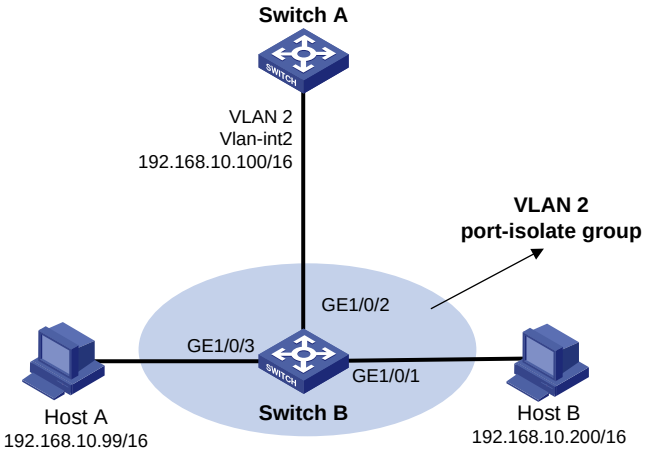
通过在 Switch 上启用代理 ARP 功能，可以解决此问题。启用代理 ARP 后，Switch 可以应答 Host A 的 ARP 请求。同时，Switch 相当于 Host B 的代理，把从其他主机发送过来的报文转发给它。

代理 ARP 的优点是，它可以只被应用在一个设备上（此时该设备的作用相当于网关），不会影响到网络中其他设备的路由表。代理 ARP 功能可以在主机没有配置缺省网关或者主机没有任何路由能力的情况下使用。

3.1.2 本地代理ARP

本地代理ARP的应用场景如 图 3-2 所示。Host A和Host B属于同一个VLAN 2，但它们分别连接到被二层隔离的端口GigabitEthernet1/0/3 和GigabitEthernet1/0/1 上，通过在SwitchA上启用本地代理ARP功能，可以实现Host A和Host B的三层互通。

图3-2 本地代理 ARP 的应用环境



本地代理 ARP 可以在下列三种情况下实现主机之间的三层互通：

- 想要互通的主机分别连接到同一个 VLAN 中的不同二层隔离端口下；
- 使能 Super VLAN 功能后，想要互通的主机属于不同的 Sub VLAN；
- 使能 Isolate-user-vlan 功能后，想要互通的主机属于不同的 Secondary VLAN。

3.2 配置代理ARP功能

代理 ARP 和本地代理 ARP 功能均可在 VLAN 接口视图/三层以太网端口视图/三层聚合接口视图下进行配置。

表3-1 配置代理 ARP 功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
开启代理ARP功能	**proxy-arp enable**	必选 缺省情况下，关闭代理ARP功能

表3-2 配置本地代理 ARP 功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-

操作	命令	说明
开启本地代理ARP功能	**local-proxy-arp enable** [**ip-range** *startIP* **to** *endIP*]	必选 缺省情况下，关闭本地代理ARP功能

3.3 代理ARP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代理 ARP 的运行情况，查看显示信息验证配置的效果。

表3-3 代理 ARP 显示和维护

操作	命令
显示代理ARP的状态	**display proxy-arp** [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示本地代理ARP的状态	**display local-proxy-arp** [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

3.4 代理ARP典型配置举例

3.4.1 代理ARP配置举例

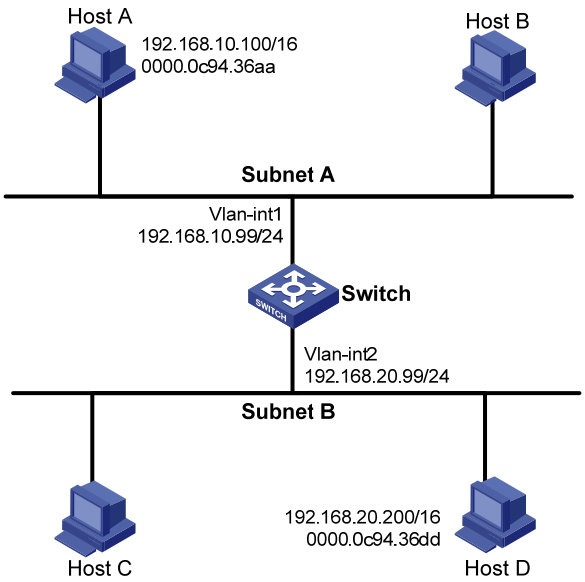
1. 组网需求

Host A 和 Host D 为同一网段的主机（Host A 的 IP 地址是 192.168.10.100/16，Host D 的 IP 地址是 192.168.20.200/16），但却被设备 Switch 分在两个不同的子网（Host A 属于 VLAN 1，Host D 属于 VLAN 2）。

当 Host A 需要与 Host D 通信时，由于目的 IP 地址与本机的 IP 地址属于同一网段，Host A 会直接发出请求 Host D 硬件地址的 ARP 请求。但是，因为两台主机处于不同的广播域中，Host D 无法收到 Host A 的 ARP 请求报文，当然也就无法应答。通过在 Switch 上启用代理 ARP 功能，可以使得处在两个子网的 Host A 和 Host D 互通。

2. 组网图

图3-3 配置代理 ARP 组网图



3. 配置步骤

创建 VLAN 2。

```
<Switch> system-view
[Switch] vlan 2
[Switch-vlan2] quit
```

配置接口 Vlan-interface1 的 IP 地址。

```
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 192.168.10.99 255.255.255.0
```

开启接口 Vlan-interface1 的代理 ARP 功能。

```
[Switch-Vlan-interface1] proxy-arp enable
[Switch-Vlan-interface1] quit
```

配置接口 Vlan-interface2 的 IP 地址。

```
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ip address 192.168.20.99 255.255.255.0
```

开启接口 Vlan-interface2 的代理 ARP 功能。

```
[Switch-Vlan-interface2] proxy-arp enable
```

配置完成后，Host A 和 Host D 可以互相 ping 通。

3.4.2 端口隔离时的本地代理ARP配置举例

1. 组网需求

- Host A 和 Host B 属于同一个 VLAN，分别与设备 Switch B 的端口 GigabitEthernet1/0/3 和 GigabitEthernet1/0/1 相连；
- 设备 Switch B 通过端口 GigabitEthernet1/0/2 端口与 Switch A 相连。

在此网络中要求：

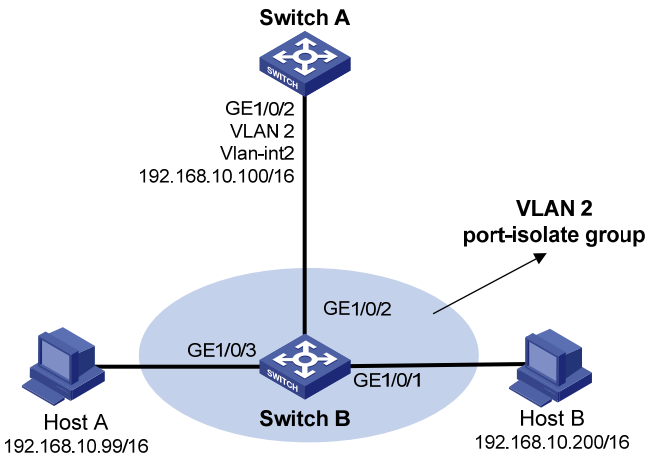
- Host A 和 Host B 之间不能二层互通。
- Host A 和 Host B 之间可以进行三层通信。

通过如下方法，可以满足上述需求：

- 配置端口隔离实现 Host A 和 Host B 之间不能二层互通。
- 配置本地代理 ARP 功能实现 Host A 和 Host B 之间三层互通。

2. 组网图

图3-4 配置端口隔离时的本地代理 ARP 组网图



3. 配置步骤

(1) 配置 Switch B

在 Switch B 上端口 GigabitEthernet1/0/3、GigabitEthernet1/0/1、GigabitEthernet1/0/2 属于同一 VLAN 2；Host A 和 Host B 彼此之间二层报文不能互通。

```
<SwitchB> system-view
[SwitchB] vlan 2
[SwitchB-vlan2] port GigabitEthernet 1/0/3
[SwitchB-vlan2] port GigabitEthernet 1/0/1
[SwitchB-vlan2] port GigabitEthernet 1/0/2
[SwitchB-vlan2] quit
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] port-isolate enable
[SwitchB-GigabitEthernet1/0/3] quit
[SwitchB] interface GigabitEthernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port-isolate enable
[SwitchB-GigabitEthernet1/0/1] quit
```

(2) 配置 SwitchA

创建 VLAN 2，添加端口 GigabitEthernet1/0/2 到 VLAN 2。

```
<SwitchA> system-view
[SwitchA] vlan 2
[SwitchA-vlan2] port GigabitEthernet 1/0/2
[SwitchA-vlan2] quit
[SwitchA] interface vlan-interface 2
```

```
[SwitchA-Vlan-interface2] ip address 192.168.10.100 255.255.0.0
```

从 Host A 上 ping 不通 Host B，表明 Host A 和 Host B 二层隔离。

配置本地代理 ARP，实现 Host A 和 Host B 之间的三层互通。

```
[SwitchA-Vlan-interface2] local-proxy-arp enable
```

从 Host A 上可以 ping 通 Host B，表明 Host A 和 Host B 三层互通了。

3.4.3 Super VLAN中的本地代理ARP配置举例

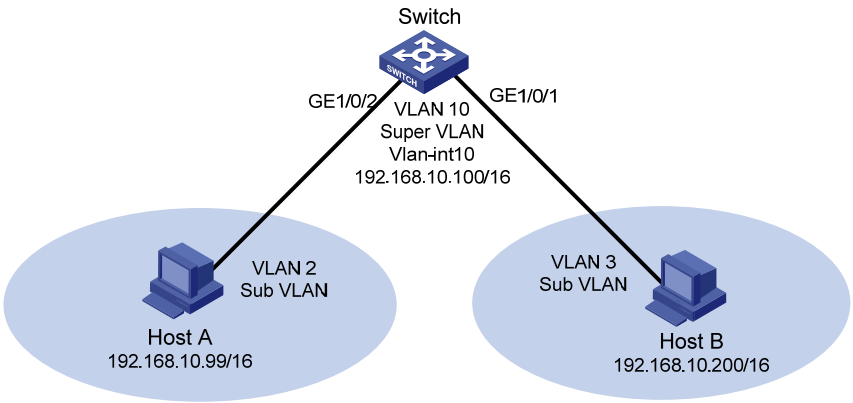
1. 组网需求

- 创建 Super VLAN：VLAN 10，VLAN 接口的 IP 地址为 192.168.10.100/16。
- 创建 Sub VLAN：VLAN 2、VLAN 3。
- 端口 GigabitEthernet1/0/2 属于 VLAN 2，端口 GigabitEthernet1/0/1 属于 VLAN 3。
- Host A 属于 VLAN 2，与 Switch 的端口 GigabitEthernet1/0/2 相连；Host B 属于 VLAN 3，与 Switch 的端口 GigabitEthernet1/0/1 相连。

由于 Host A 和 Host B 属于不同的 Sub VLAN，Host A 和 Host B 之间二层隔离。通过在 Switch 上配置本地代理 ARP 功能，可以实现 Host A 和 Host B 之间的三层互通。

2. 组网图

图3-5 配置 Super VLAN 中的本地代理 ARP 组网图



3. 配置步骤

创建 Super VLAN，Sub VLAN；添加端口 GigabitEthernet1/0/2 到 VLAN 2，端口 GigabitEthernet1/0/1 到 VLAN 3；配置 VLAN 10 接口的 IP 地址为 192.168.10.100/16。

```
<Switch> system-view
[Switch] vlan 2
[Switch-vlan2] port GigabitEthernet 1/0/2
[Switch-vlan2] quit
[Switch] vlan 3
[Switch-vlan3] port GigabitEthernet 1/0/1
[Switch-vlan3] quit
[Switch] vlan 10
[Switch-vlan10] supervlan
[Switch-vlan10] subvlan 2 3
```

```
[Switch-vlan10] quit
[Switch] interface vlan-interface 10
[Switch-Vlan-interface10] ip address 192.168.10.100 255.255.0.0
```

从 Host A 上 ping 不通 Host B，表明 Host A 和 Host B 二层隔离。

配置本地代理 ARP，实现 Sub VLAN 之间的三层互通。

```
[Switch-Vlan-interface10] local-proxy-arp enable
```

从 Host A 上可以 ping 通 Host B，表明 Host A 和 Host B 三层互通了。

3.4.4 Isolate-user-vlan中的本地代理ARP配置举例

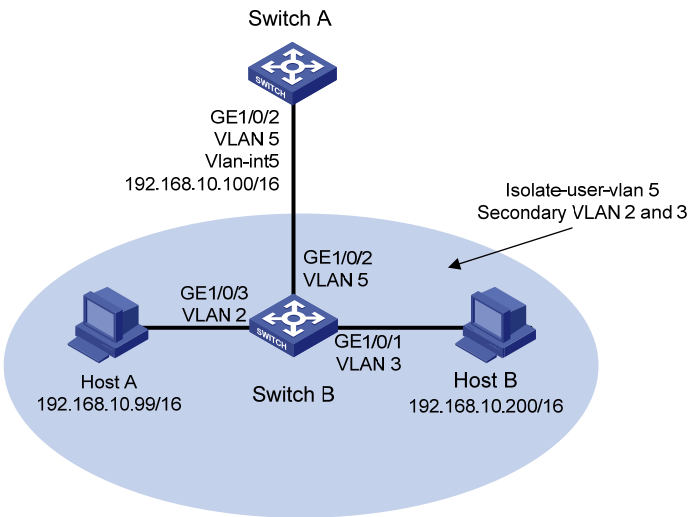
1. 组网需求

- 设备 Switch A 下接设备 Switch B。
- 设备 Switch B 上的 VLAN 5 为 Isolate-user-vlan，包含上行端口 GigabitEthernet1/0/2 和两个 Secondary VLAN（VLAN 2 和 VLAN 3），VLAN 2 包含端口 GigabitEthernet1/0/3，VLAN 3 包含端口 GigabitEthernet1/0/1。
- Host A 属于 VLAN 2，与 Switch B 的端口 GigabitEthernet1/0/3 相连；Host B 属于 VLAN 3，与 Switch B 的端口 GigabitEthernet1/0/1 相连。

由于 Host A 和 Host B 属于不同的 Secondary VLAN，Host A 和 Host B 之间二层隔离。通过在 Switch A 上启用本地代理 ARP 功能，可以实现 Host A 和 Host B 之间的三层互通。

2. 组网图

图3-6 配置 Isolate-user-vlan 中的本地代理 ARP 组网图



3. 配置步骤

(1) 配置 Switch B

在设备 Switch B 上创建 VLAN 2、VLAN 3 和 VLAN 5；添加端口 GigabitEthernet1/0/3 到 VLAN 2，端口 GigabitEthernet1/0/1 到 VLAN 3，端口 GigabitEthernet1/0/2 到 VLAN 5；配置 VLAN 5 为 Isolate-user-vlan，VLAN 2 和 VLAN 3 为 Secondary VLAN；配置 Isolate-user-vlan 和 Secondary VLAN 间的映射关系。

```
<SwitchB> system-view
[SwitchB] vlan 2
[SwitchB-vlan2] port GigabitEthernet 1/0/3
[SwitchB-vlan2] quit
[SwitchB] vlan 3
[SwitchB-vlan3] port GigabitEthernet 1/0/1
[SwitchB-vlan3] quit
[SwitchB] vlan 5
[SwitchB-vlan5] port GigabitEthernet 1/0/2
[SwitchB-vlan5] isolate-user-vlan enable
[SwitchB-vlan5] quit
[SwitchB] interface GigabitEthernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] port isolate-user-vlan 5 promiscuous
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface GigabitEthernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port isolate-user-vlan host
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] port isolate-user-vlan host
[SwitchB-GigabitEthernet1/0/3] quit
[SwitchB] isolate-user-vlan 5 secondary 2 3
```

(2) 配置 Switch A

创建 VLAN 5，添加端口 GigabitEthernet1/0/2 到 VLAN 5。

```
<SwitchA> system-view
[SwitchA] vlan 5
[SwitchA-vlan5] port GigabitEthernet 1/0/2
[SwitchA-vlan5] quit
[SwitchA] interface vlan-interface 5
[SwitchA-Vlan-interface5] ip address 192.168.10.100 255.255.0.0
```

从 Host A 上 ping 不通 Host B，表明 Host A 和 Host B 二层隔离。

配置本地代理 ARP，实现 Secondary VLAN 之间的三层互通。

```
[SwitchA-Vlan-interface5] local-proxy-arp enable
```

从 Host A 上可以 ping 通 Host B，表明 Host A 和 Host B 三层互通了。

4 ARP Snooping配置

4.1 ARP Snooping简介

4.1.1 作用

ARP Snooping 功能是一个用于二层交换网络环境的特性，通过侦听 ARP 报文建立 ARP Snooping 表项。

4.1.2 工作机制

设备上的一个 VLAN 使能 ARP Snooping 后，该 VLAN 内所有端口接收的 ARP 报文会被重定向到 CPU。CPU 对重定向上送的 ARP 报文进行分析，获取 ARP 报文的源 IP 地址、源 MAC 地址、VLAN 和入端口信息，建立记录用户信息的 ARP Snooping 表项。

ARP Snooping 表项的老化时间为 25 分钟，有效时间为 15 分钟。如果一个 ARP Snooping 表项自最后一次更新后 15 分钟内没有收到 ARP 更新报文，则此表项开始进入失效状态，不再对外提供服务，其他特性查找此表项将会失败。当收到源 IP 地址和源 MAC 与已存在的 ARP Snooping 表项 IP 地址和 MAC 均相同的 ARP 报文时，此 ARP Snooping 表项进行更新，重新开始生效，并重新老化计时。当 ARP Snooping 表项达到老化时间后，则将此 ARP Snooping 表项删除。

如果 ARP Snooping 收到 ARP 报文时检查到相同 IP 的 ARP Snooping 表项已经存在，但是 MAC 地址发生了变化，则认为发生了攻击，此时 ARP Snooping 表项处于冲突状态，表项失效，不再对外提供服务，并在 25 分钟后删除此表项。

4.2 配置ARP Snooping

表4-1 配置 ARP Snooping

操作	命令	说明
进入系统视图	**system-view**	-
进入VLAN视图	**vlan** *vlan-id*	-
使能ARP Snooping功能	**arp-snooping enable**	必选 缺省情况下，关闭ARP Snooping功能

4.3 ARP Snooping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 ARP Snooping 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** 命令清除 ARP Snooping 表中的表项。

表4-2 ARP Snooping 显示和维护

操作	命令
显示ARP Snooping表项	**display arp-snooping** [**ip** *ip-address* | **vlan** *vlan-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除ARP Snooping表项	**reset arp-snooping** [**ip** *ip-address* | **vlan** *vlan-id*]

目 录

1 IP地址配置

说明

本文中所指的“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地址简介

1.1.1 IP地址的分类和表示

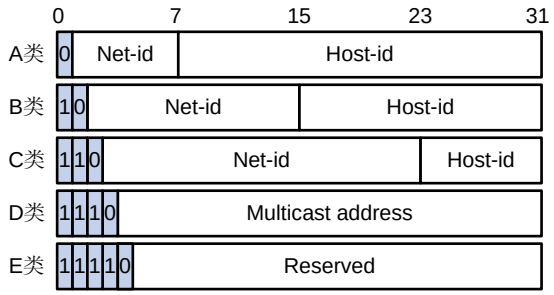
连接到 Internet 上的设备接口必须有一个全球唯一的 IP 地址。IP 地址长度为 32 比特，通常采用点分十进制方式表示，即每个 IP 地址被表示为以小数点隔开的 4 个十进制整数，每个整数对应一个字节，如 10.1.1.1。

IP 地址由两部分组成：

- 网络号码字段（Net-id）：用于区分不同的网络。网络号码字段的前几位称为类别字段（又称为类别比特），用来区分 IP 地址的类型。
- 主机号码字段（Host-id）：用于区分一个网络内的不同主机。

为了方便管理及组网，IP地址分成五类，如 图 1-1 所示，其中蓝色部分为类别字段。

图1-1 五类 IP 地址



上述五类IP地址的地址范围如 表 1-1 所示。目前大量使用的IP地址属于A、B、C三类。

表1-1 IP 地址分类及范围

地址类型	地址范围	说明
A	0.0.0.0～127.255.255.255	IP地址0.0.0.0仅用于主机在系统启动时进行临时通信，并且永远不是有效目的地址 127.0.0.0网段的地址都保留作环回测试，发送到这个地址的分组不会输出到链路上，它们被当作输入分组在内部进行处理

地址类型	地址范围	说明
B	128.0.0.0～191.255.255.255	-
C	192.0.0.0～223.255.255.255	-
D	224.0.0.0～239.255.255.255	组播地址
E	240.0.0.0～255.255.255.255	255.255.255.255用于广播地址，其他地址保留今后使用

1.1.2 特殊的IP地址

下列 IP 地址具有特殊的用途，不能作为主机的 IP 地址。

- Net-id 为全 0 的地址：表示本网络内的主机。例如，0.0.0.16 表示本网络内 Host-id 为 16 的主机。
- Host-id 为全 0 的地址：网络地址，用于标识一个网络。
- Host-id 为全 1 的地址：网络广播地址。例如，目的地址为 192.168.1.255 的报文，将转发给 192.168.1.0 网络内所有的主机。

1.1.3 子网和掩码

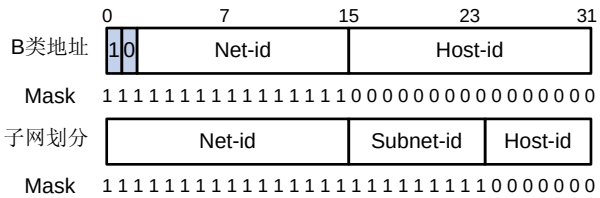
随着 Internet 的快速发展，IP 地址已近枯竭。如果将一个可以容纳 65534（2^{16}-2，去掉主机号码字段全 1 的广播地址和主机号码字段全 0 的网段地址）个主机的 B 类地址分配给一个只有几百台主机的网络，就会浪费大量的 IP 地址。

为了充分利用已有的 IP 地址，可以使用子网掩码将网络划分为更小的部分（即子网）。通过从主机号码字段部分划出一些比特位作为子网号码字段，能够将一个网络划分为多个子网。子网号码字段的长度由子网掩码确定。

子网掩码是一个长度为 32 比特的数字，由一串连续的“1”和一串连续的“0”组成。“1”对应于网络号码字段和子网号码字段，而“0”对应于主机号码字段。

图 1-2 所示是一个B类地址划分子网的情况。

图1-2 IP 地址子网划分



多划分出一个子网号码字段会浪费一些 IP 地址。例如，一个 B 类地址可以容纳 65534（2^{16}-2，去掉主机号码字段全 1 的广播地址和主机号码字段全 0 的网段地址）个主机号码。但划分出 9 比特长的子网字段后，最多可有 512（2^9）个子网，每个子网有 7 比特的主机号码，即每个子网最多可有 126（2^7-2，去掉主机号码字段全 1 的广播地址和主机号码字段全 0 的网段地址）个主机号码。因此主机号码的总数是 512*126=64512 个，比不划分子网时要少 1022 个。

若不进行子网划分，则子网掩码为默认值，此时子网掩码中“1”的长度就是网络号码的长度，即 A、B、C 类 IP 地址对应的子网掩码默认值分别为 255.0.0.0、255.255.0.0 和 255.255.255.0。

1.2 配置IP地址

接口有了 IP 地址后就可以与其他主机进行 IP 通信。接口获取 IP 地址有以下几种方式：

- 通过手工指定 IP 地址
- 通过 BOOTP 分配得到 IP 地址
- 通过 DHCP 分配得到 IP 地址

这几种方式是互斥的，通过新的配置方式获取的 IP 地址会覆盖通过原有方式获取的 IP 地址。例如，首先通过手工指定了 IP 地址，然后使用 DHCP 协议申请 IP 地址，那么手工指定的 IP 地址会被删除，接口的 IP 地址是通过 DHCP 协议分配的。

说明

本节只介绍通过手工指定 IP 地址的方式。通过 BOOTP 和 DHCP 分配得到 IP 地址方式的介绍请参见“三层技术-IP 业务配置指导”中的“DHCP”。

1.2.1 配置接口的IP地址

设备的每个接口可以配置多个 IP 地址，其中一个为主 IP 地址，其余为从 IP 地址。

一般情况下，一个接口只需配置一个主 IP 地址，但在有些特殊情况下需要配置从 IP 地址。比如，一台设备通过一个接口连接了一个局域网，但该局域网中的计算机分别属于 2 个不同的子网，为了使设备与局域网中的所有计算机通信，就需要在该接口上配置一个主 IP 地址和一个从 IP 地址。

表1-2 配置接口的 IP 地址

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IP地址	**ip address** *ip-address* { *mask-length* | *mask* } [**sub**]	必选 缺省情况下，没有为接口配置IP地址

注意

- 一个接口只能有一个主 IP 地址。新配置的主 IP 地址将覆盖原有主 IP 地址。
- 当接口被配置为通过 BOOTP、DHCP 方式获取 IP 地址后，则不能再给该接口配置从 IP 地址。
- 同一接口的主、从 IP 地址可以在同一网段，但不同接口之间的 IP 地址不可以在同一网段。
- 为满足点对点连接通信中节约 IP 地址的使用需求，交换机支持配置 IP 地址的子网掩码长度为 31（或子网掩码为 255.255.255.254）。

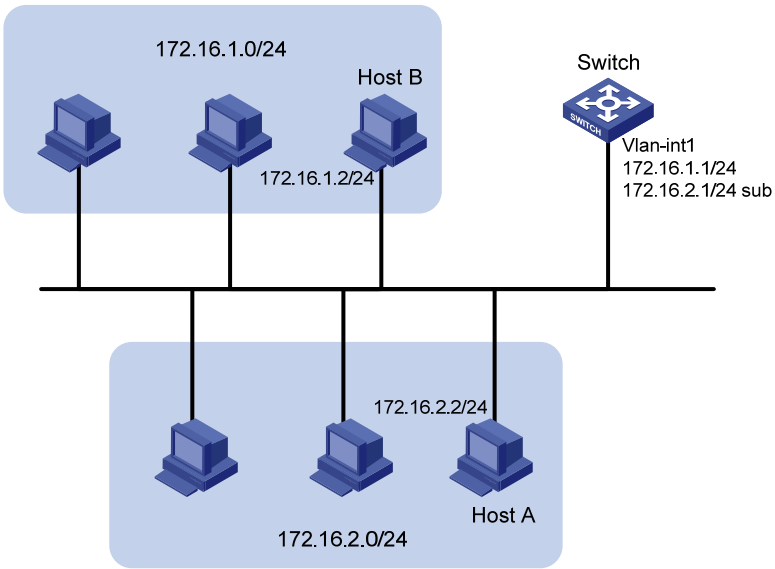
1.2.2 IP地址配置举例

1. 组网需求

Switch 的端口（属于 VLAN 1）连接一个局域网，局域网中的计算机分别属于 2 个网段：172.16.1.0/24 和 172.16.2.0/24。要求这两个网段的主机都可以通过 Switch 与外部网络通信，且这两个网段中的主机能够互通。

2. 组网图

图1-3 IP 地址配置组网图



3. 配置步骤

针对上述的需求，如果在 Switch 的 VLAN 接口 1 上只配置一个 IP 地址，则只有一部分主机能够通过 Switch 与外部网络通信。为了使局域网内的所有主机都能够通过 Switch 访问外部网络，需要配置 VLAN 接口 1 的从 IP 地址。为了使两个网段中的主机能够互通，两个网段中的主机需要分别将 Switch 上 VLAN 接口 1 的主 IP 地址和从 IP 地址设置为网关。

配置 VLAN 接口 1 的主 IP 地址和从 IP 地址。

```
<Switch> system-view
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 172.16.1.1 255.255.255.0
[Switch-Vlan-interface1] ip address 172.16.2.1 255.255.255.0 sub
```

在 172.16.1.0/24 网段中的主机上配置网关为 172.16.1.1；在 172.16.2.0/24 网段中的主机上配置网关为 172.16.2.1。

使用 **ping** 命令检测 Switch 与网络 172.16.1.0/24 内主机的连通性。

```
<Switch> ping 172.16.1.2
  PING 172.16.1.2: 56  data bytes, press CTRL_C to break
    Reply from 172.16.1.2: bytes=56 Sequence=1 ttl=255 time=25 ms
    Reply from 172.16.1.2: bytes=56 Sequence=2 ttl=255 time=27 ms
    Reply from 172.16.1.2: bytes=56 Sequence=3 ttl=255 time=26 ms
    Reply from 172.16.1.2: bytes=56 Sequence=4 ttl=255 time=26 ms
```

```
    Reply from 172.16.1.2: bytes=56 Sequence=5 ttl=255 time=26 ms

  --- 172.16.1.2 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 25/26/27 ms
```

显示信息表示 Switch 与网络 172.16.1.0/24 内的主机可以互通。

使用 **ping** 命令检测 Switch 与网络 172.16.2.0/24 内主机的连通性。

```
<Switch> ping 172.16.2.2
  PING 172.16.2.2: 56  data bytes, press CTRL_C to break
    Reply from 172.16.2.2: bytes=56 Sequence=1 ttl=255 time=25 ms
    Reply from 172.16.2.2: bytes=56 Sequence=2 ttl=255 time=26 ms
    Reply from 172.16.2.2: bytes=56 Sequence=3 ttl=255 time=26 ms
    Reply from 172.16.2.2: bytes=56 Sequence=4 ttl=255 time=26 ms
    Reply from 172.16.2.2: bytes=56 Sequence=5 ttl=255 time=26 ms

  --- 172.16.2.2 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 25/25/26 ms
```

显示信息表示 Switch 与网络 172.16.2.0/24 内的主机可以互通。

使用 **ping** 命令检测网络 172.16.1.0/24 和网络 172.16.2.0/24 内主机的连通性。在 Host A 上可以 ping 通 Host B。

1.3 配置接口借用IP地址

所谓“IP 地址借用”，是指一个接口上没有配置 IP 地址，但为了使该接口能正常使用，就向同一设备上其它有 IP 地址的接口借用一个 IP 地址。IP 地址借用使用的场合如下：

- 在 IP 地址资源比较缺乏的环境下，为了节约 IP 地址资源，可以配置某个接口借用其他接口的 IP 地址。
- 如果某个接口只是偶尔使用，可以配置该接口借用其他接口的 IP 地址，而不必让其一直占用一个单独的 IP 地址。

1.3.1 配置准备

被借用接口的主 IP 地址已经配置，配置方法可以为手工指定、通过 BOOTP 或 DHCP 动态获取。

1.3.2 配置接口借用IP地址

表1-3 配置接口借用 IP 地址

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
进入Tunnel接口视图	**interface tunnel** *number*	-
配置本接口借用指定接口的IP地址	**ip address unnumbered interface** *interface-type interface-number*	必选 缺省情况下，不借用其它接口的IP地址

注意

- 被借用接口的地址本身不能为借用地址。
- 一个接口的地址可以借给多个接口。
- 如果被借用接口有多个 IP 地址，则只有主 IP 地址能被借用。
- 借用接口的 IP 地址始终与被借用接口的 IP 地址保持一致，并且随着被借用接口 IP 地址的变化而变化。即如果被借用接口已经配置 IP 地址，则借用接口的 IP 地址与被借用接口相同；如果被借用接口没有配置 IP 地址，则借用接口的 IP 地址也处于未配置状态。

1.4 IP地址的显示和维护

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 IP 地址的运行情况，通过查看显示信息验证配置的效果。

表1-4 IP 地址的显示和维护

操作	命令
显示三层接口的相关信息	**display ip interface** [*interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示三层接口的IP基本配置信息	**display ip interface** [*interface-type* [*interface-number*]] **brief** [**description**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

说明

显示三层接口的 IP 基本配置信息中的 **description** 参数仅 R5206 及以上版本支持。

目 录

1 DHCP概述

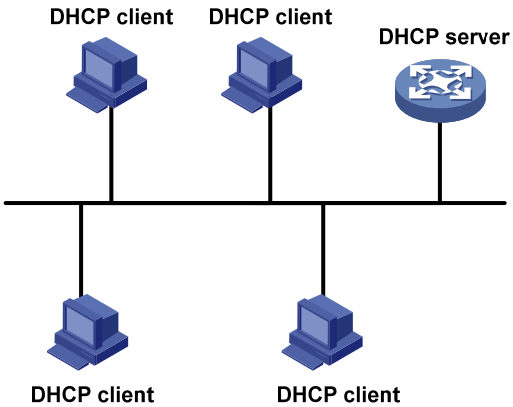
1.1 DHCP简介

DHCP（Dynamic Host Configuration Protocol，动态主机配置协议）用来为网络设备动态地分配 IP 地址等网络配置参数。

DHCP 采用客户端/服务器通信模式，由客户端向服务器提出配置申请，服务器返回为客户端分配的 IP 地址等相应的配置信息，以实现 IP 地址等信息的动态配置。

在DHCP的典型应用中，一般包含一台DHCP服务器和多台客户端（如PC和便携机），如 图 1-1 所示。

图1-1 DHCP 典型应用



说明

DHCP客户端和DHCP服务器处于不同物理网段时，客户端可以通过DHCP中继与服务器通信，获取IP地址及其他配置信息。DHCP中继的详细介绍，请参见“3.1 DHCP中继简介”。

1.2 DHCP的IP地址分配

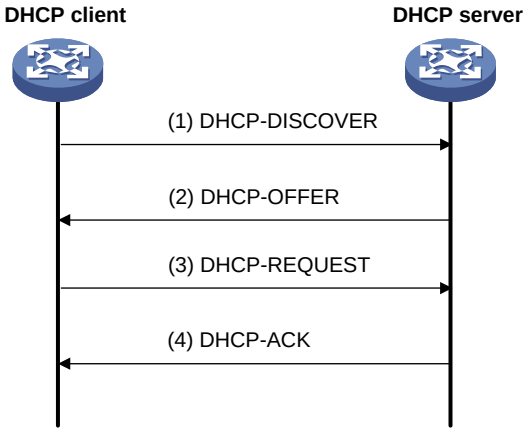
1.2.1 IP地址分配策略

针对客户端的不同需求，DHCP 提供三种 IP 地址分配策略：

- 手工分配地址：由管理员为少数特定客户端（如 WWW 服务器等）静态绑定固定的 IP 地址。通过 DHCP 将配置的固定 IP 地址发给客户端。
- 自动分配地址：DHCP 为客户端分配租期为无限长的 IP 地址。
- 动态分配地址：DHCP 为客户端分配具有一定有效期限的 IP 地址，到达使用期限后，客户端需要重新申请地址。绝大多数客户端得到的都是这种动态分配的地址。

1.2.2 IP地址动态获取过程

图1-2 IP 地址动态获取过程



如 图 1-2 所示，DHCP客户端从DHCP服务器动态获取IP地址，主要通过四个阶段进行：

(1) 发现阶段，即 DHCP 客户端寻找 DHCP 服务器的阶段。客户端以广播方式发送 DHCP-DISCOVER 报文。

(2) 提供阶段，即DHCP服务器提供IP地址的阶段。DHCP服务器接收到客户端的DHCP-DISCOVER报文后，根据IP地址分配的优先次序选出一个IP地址，与其他参数一起通过DHCP-OFFER报文发送给客户端。DHCP-OFFER报文的发送方式由DHCP-DISCOVER报文中的flag字段决定，具体请参见“1.3 DHCP报文格式”的介绍。

(3) 选择阶段，即 DHCP 客户端选择 IP 地址的阶段。如果有多台 DHCP 服务器向该客户端发来 DHCP-OFFER 报文，客户端只接受第一个收到的 DHCP-OFFER 报文，然后以广播方式发送 DHCP-REQUEST 报文，该报文中包含 DHCP 服务器在 DHCP-OFFER 报文中分配的 IP 地址。

(4) 确认阶段，即 DHCP 服务器确认 IP 地址的阶段。DHCP 服务器收到 DHCP 客户端发来的 DHCP-REQUEST 报文后，只有 DHCP 客户端选择的服务器会进行如下操作：如果确认将地址分配给该客户端，则返回 DHCP-ACK 报文；否则返回 DHCP-NAK 报文，表明地址不能分配给该客户端。

说明

- 客户端收到服务器返回的 DHCP-ACK 确认报文后，会以广播的方式发送免费 ARP 报文，探测是否有主机使用服务器分配的 IP 地址，如果在规定的时间内没有收到回应，客户端才使用此地址。否则，客户端会发送 DHCP-DECLINE 报文给 DHCP 服务器，并重新申请 IP 地址。
- 如果网络中存在多个 DHCP 服务器，除 DHCP 客户端选中的服务器外，其它 DHCP 服务器中本次未分配出的 IP 地址仍可分配给其他客户端。

1.2.3 IP地址的租约更新

如果采用动态地址分配策略，则 DHCP 服务器分配给客户端的 IP 地址有一定的租借期限，当租借期满后服务器会收回该 IP 地址。如果 DHCP 客户端希望继续使用该地址，需要更新 IP 地址租约。

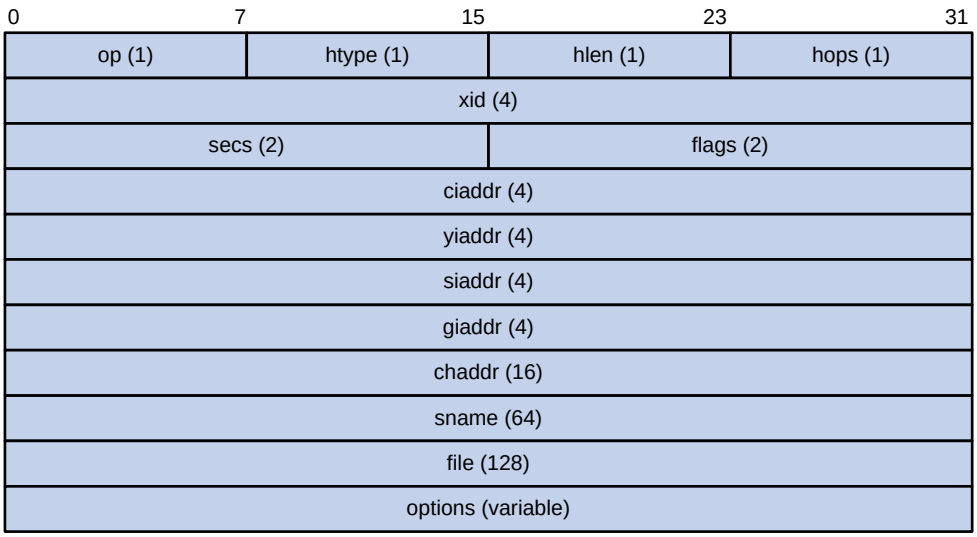
在 DHCP 客户端的 IP 地址租约期限达到一半时间时，DHCP 客户端会向为它分配 IP 地址的 DHCP 服务器单播发送 DHCP-REQUEST 报文，以进行 IP 租约的更新。如果客户端可以继续使用此 IP 地址，则 DHCP 服务器回应 DHCP-ACK 报文，通知 DHCP 客户端已经获得新 IP 租约；如果此 IP 地址不可以再分配给该客户端，则 DHCP 服务器回应 DHCP-NAK 报文，通知 DHCP 客户端不能获得新的租约。

如果在租约的一半时间进行的续约操作失败，DHCP 客户端会在租约期限达到 7/8 时，广播发送 DHCP-REQUEST 报文进行续约。DHCP 服务器的处理方式同上，不再赘述。

1.3 DHCP报文格式

DHCP有 8 种类型的报文，每种报文的格式相同，只是某些字段的取值不同。DHCP报文格式基于BOOTP的报文格式，具体格式如 图 1-3 所示（括号中的数字表示该字段所占的字节）。

图1-3 DHCP 报文格式



各字段的解释如下：

- op：报文的操作类型，分为请求报文和响应报文，1 为请求报文；2 为响应报文。具体的报文类型在 option 字段中标识。
- htype、hlen：DHCP 客户端的硬件地址类型及长度。
- hops：DHCP 报文经过的 DHCP 中继的数目。DHCP 请求报文每经过一个 DHCP 中继，该字段就会增加 1。
- xid：客户端发起一次请求时选择的随机数，用来标识一次地址请求过程。
- secs：DHCP 客户端开始 DHCP 请求后所经过的时间。目前没有使用，固定为 0。
- flags：第一个比特为广播响应标识位，用来标识 DHCP 服务器响应报文是采用单播还是广播方式发送，0 表示采用单播方式，1 表示采用广播方式。其余比特保留不用。

- ciaddr：DHCP 客户端的 IP 地址。
- yiaddr：DHCP 服务器分配给客户端的 IP 地址。
- siaddr：DHCP 客户端获取 IP 地址等信息的服务器 IP 地址。
- giaddr：DHCP 客户端发出请求报文后经过的第一个 DHCP 中继的 IP 地址。
- chaddr：DHCP 客户端的硬件地址。
- sname：DHCP 客户端获取 IP 地址等信息的服务器名称。
- file：DHCP 服务器为 DHCP 客户端指定的启动配置文件名称及路径信息。
- options：可选变长选项字段，包含报文的类型、有效租期、DNS 服务器的 IP 地址、WINS 服务器的 IP 地址等配置信息。

1.4 DHCP选项

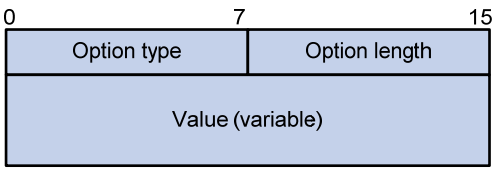
1.4.1 DHCP选项简介

为了与 BOOTP（Bootstrap Protocol，自举协议）兼容，DHCP 保留了 BOOTP 的消息格式。DHCP 和 BOOTP 消息的不同主要体现在选项（Option）字段。DHCP 在 BOOTP 基础上增加的功能，通过 Option 字段来实现。

DHCP 利用 Option 字段传递控制信息和网络配置参数，实现地址的动态分配，为客户端提供更加丰富的网络配置信息。

DHCP选项的格式如 图 1-4 所示。

图1-4 DHCP 选项格式



1.4.2 DHCP常用选项介绍

常见的 DHCP 选项有：

- Option 3：路由器选项，用来指定为客户端分配的网关地址。
- Option 6：DNS 服务器选项，用来指定为客户端分配的 DNS 服务器地址。
- Option 33：静态路由选项。该选项中包含一组有分类静态路由（即目的地址的掩码固定为自然掩码，不能划分子网），客户端收到该选项后，将在路由表中添加这些静态路由。如果 Option 33 和 Option 121 同时存在，则忽略 Option 33。
- Option 51：IP 地址租约选项。
- Option 53：DHCP 消息类型选项，标识 DHCP 消息的类型。
- Option 55：请求参数列表选项。客户端利用该选项指明需要从服务器获取哪些网络配置参数。该选项内容为客户端请求的参数对应的选项值。

- Option 60：厂商标识选项。客户端利用该选项标识自己所属的厂商；DHCP 服务器可以根据该选项区分客户端所属的厂商，并为其分配特定范围的 IP 地址。
- Option 66：TFTP 服务器名选项，用来指定为客户端分配的 TFTP 服务器的域名。
- Option 67：启动文件名选项，用来指定为客户端分配的启动文件名。
- Option 121：无分类路由选项。该选项中包含一组无分类静态路由（即目的地址的掩码为任意值，可以通过掩码来划分子网），客户端收到该选项后，将在路由表中添加这些静态路由。如果 Option 33 和 Option 121 同时存在，则忽略 Option 33。
- Option 150：TFTP 服务器地址选项，用来指定为客户端分配的 TFTP 服务器的地址。

更多 DHCP 选项的介绍，请参见 RFC 2132 和 RFC 3442。

1.4.3 自定义的选项格式

有些选项的内容，RFC 2132 中没有统一规定，例如 Option 43。下面将介绍设备上定义的几种选项格式。

1. 厂商特定信息选项（Option 43）

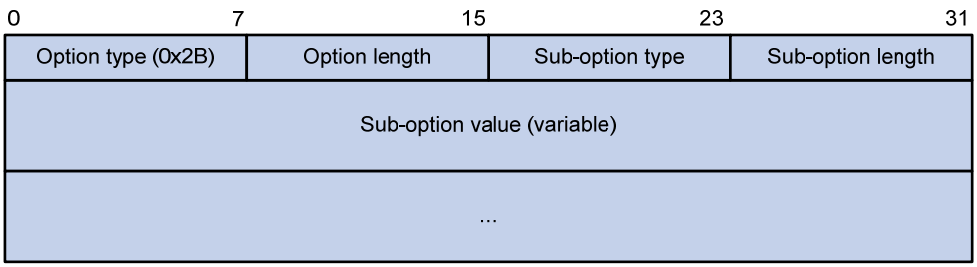
Option 43 称为厂商特定信息选项。DHCP 服务器和 DHCP 客户端通过 Option 43 交换厂商特定的信息。

设备作为 DHCP 客户端时，可以通过 Option 43 获取：

- ACS（Auto-Configuration Server，自动配置服务器）的参数，包括 URL 地址、用户名和密码。
- 服务提供商标识，CPE（Customer Premises Equipment，用户侧设备）从 DHCP 服务器获取该信息后，将该信息通告给 ACS，以便 ACS 选择服务提供商特有的配置和参数等。
- PXE（Preboot eXecution Environment，预启动执行环境）引导服务器地址，以便客户端从 PXE 引导服务器获取启动文件或其他控制信息。

(1) Option 43 格式

图1-5 Option 43 格式



为了提供可扩展性，通过Option 43 为客户端分配更多的信息，Option 43 采用子选项的形式，通过不同的子选项为用户分配不同的网络配置参数，如 图 1-5 所示。子选项中各字段的含义为：

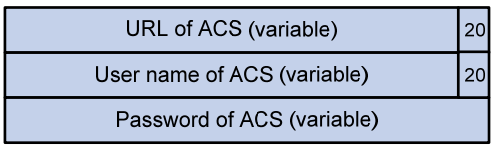
- Sub-option type：子选项类型。目前，子选项类型值可以为 0x01 表示 ACS 参数子选项，0x02 表示服务提供商标识子选项，0x80 表示 PXE 引导服务器地址子选项。
- Sub-option length：子选项的长度，不包括子选项类型和子选项长度字段。

- Sub-option value：子选项的取值。不同类型的子选项，取值格式有所不同，详细介绍请参见下文。

(2) Option 43 子选项取值字段的格式

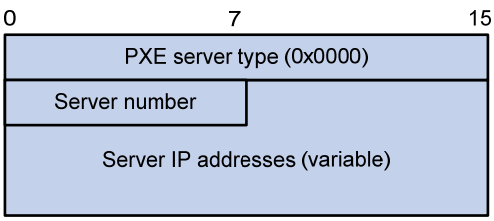
- ACS参数子选项的取值字段格式如 图 1-6 所示。ACS的URL地址、用户名和密码长度可变，每个参数之间用空格（十六进制数为 0x20）隔开。

图1-6 ACS 参数子选项格式



- 服务提供商标识子选项的取值字段内容为服务提供商的标识。
- PXE引导服务器地址子选项的取值字段格式如 图 1-7 所示。其中，PXE服务器类型目前取值只能为 0；Server number为子选项中包含的PXE服务器地址的数目；Server IP addresses为PXE服务器的IP地址。

图1-7 PXE 引导服务器地址子选项格式



2. 中继代理信息选项（Option 82）

Option 82 称为中继代理信息选项，该选项记录了 DHCP 客户端的位置信息。DHCP 中继或 DHCP Snooping 设备接收到DHCP客户端发送给DHCP服务器的请求报文后，在该报文中添加Option 82，并转发给 DHCP 服务器。

管理员可以从 Option 82 中获得 DHCP 客户端的位置信息，以便定位 DHCP 客户端，实现对客户端的安全和计费等控制。支持 Option 82 的服务器还可以根据该选项的信息制定 IP 地址和其他参数的分配策略，提供更加灵活的地址分配方案。

Option 82 最多可以包含 255 个子选项。若定义了 Option 82，则至少要定义一个子选项。目前，DHCP 中继支持两个子选项：sub-option 1（Circuit ID，电路 ID 子选项）和 sub-option 2（Remote ID，远程 ID 子选项）；DHCP Snooping 设备支持三个子选项：sub-option 1（Circuit ID，电路 ID 子选项）、sub-option 2（Remote ID，远程 ID 子选项）和 sub-option 9（子选项 9）。

由于 Option 82 的内容没有统一规定，不同厂商通常根据需要进行填充。

设备上，可以通过两种方式配置 Option 82 的内容：

- 用户自定义方式：用户手工指定 Option 82 的内容；
- 非用户自定义方式：采用默认的 normal 模式、verbose 模式、private 模式或者 standard 模式填充 Option 82。

说明

目前，只有 DHCP Snooping 设备支持填充 sub-option 9、支持 private 和 standard 填充模式。

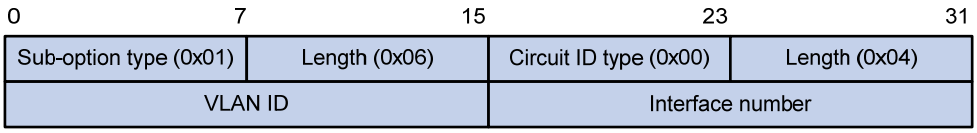
normal 和 verbose 填充模式中，子选项内容的填充格式可以是 ASCII 格式和 HEX 格式。

采用不同的填充模式时，各种子选项的内容如下：

(1) 采用 normal 模式填充

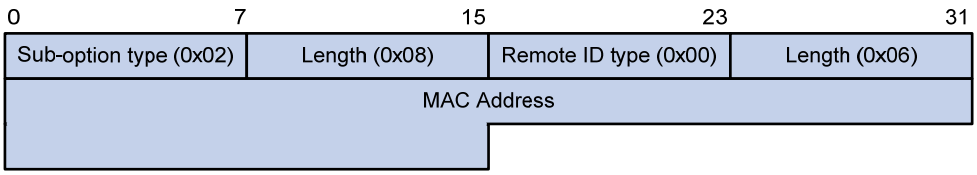
- sub-option 1 的内容是接收到DHCP客户端请求报文的接口所属的VLAN ID以及接口编号。如图 1-8 所示，子选项类型值为 1，电路ID类型值为 0。

图1-8 normal 模式填充的 sub-option 1



- sub-option 2 的内容是接收到DHCP客户端请求报文的接口MAC地址（DHCP中继）或设备的桥MAC地址（DHCP Snooping设备）。如 图 1-9 所示，子选项类型值为 2，远程ID类型值为 0。

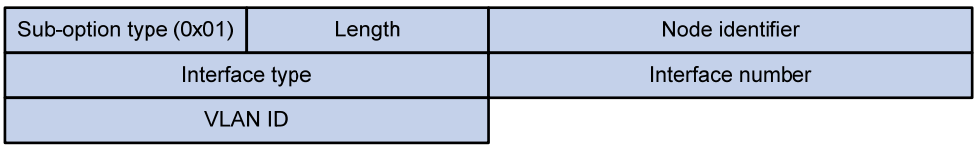
图1-9 normal 模式填充的 sub-option 2



(2) 采用 verbose 模式填充

- sub-option 1 的内容包括用户配置的接入节点标识（在报文中添加Option 82 的设备的标识）、接收到DHCP客户端请求报文的接口类型、接口编号和VLAN ID，如 图 1-10 所示。

图1-10 verbose 模式填充的 sub-option 1



说明

图 1-10 中除VLAN ID固定为两字节外，其他sub-option 1 的填充内容均为可变长度。

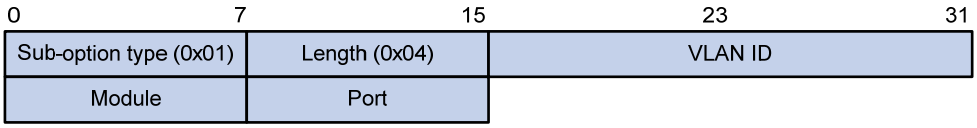
- sub-option 2 的内容是接收到DHCP客户端请求报文的接口MAC地址（DHCP中继）或设备的桥MAC地址（DHCP Snooping设备）。verbose和normal填充模式的sub-option 2 内容相同，如 图 1-9 所示。

private 填充模式和 standard 填充模式中，子选项的内容如下：

(3) 采用 private 模式填充

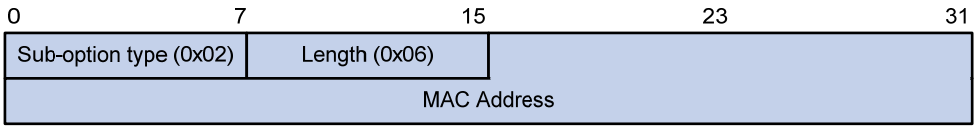
- sub-option 1 的内容是接收到DHCP客户端请求报文的接口所属的VLAN ID、Module（接口的子卡号）和Port（该接口的编号）。如 图 1-11 所示，子选项类型值为 1。

图1-11 private 模式填充的 sub-option 1



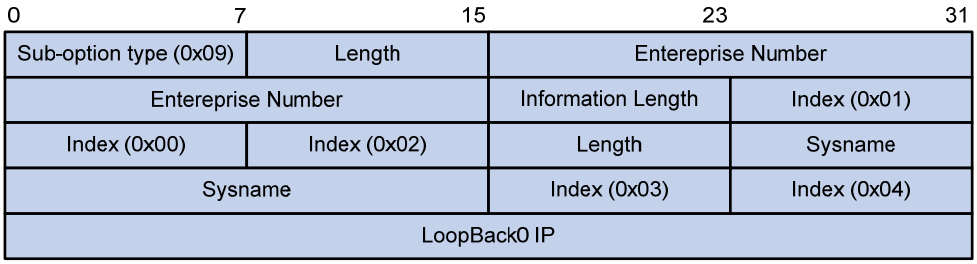
- sub-option 2 的内容是接收到DHCP客户端请求报文的DHCP Snooping设备的桥MAC地址。如 图 1-12 所示，子选项类型值为 2。

图1-12 private 模式填充的 sub-option 2



- sub-option 9 的内容是Sysname和Loopback0 接口的主IP地址。如 图 1-13 所示，子选项类型值为 9。

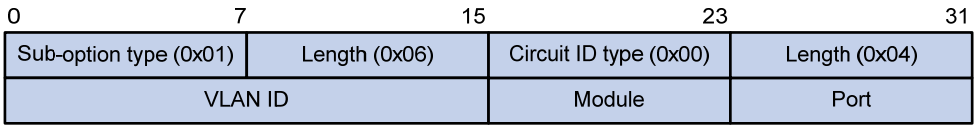
图1-13 private 模式填充的 sub-option 9



(4) 采用 standard 模式填充

- sub-option 1 的内容是接收到DHCP客户端请求报文的接口所属的VLAN ID、Module（接口的子卡号）和Port（该接口的编号）。如 图 1-14 所示，子选项类型值为 1，电路ID类型值为 0。

图1-14 standard 模式填充的 sub-option 1



- sub-option 2 的内容是接收到DHCP客户端请求报文的DHCP Snooping设备的桥MAC地址。子选项类型值为 2，远程ID类型值为 0，格式跟normal模式的sub-option 2 相同，如 图 1-9 所示。

3. Option 184

Option 184 是 RFC 中规定的保留选项，用户可以自定义该选项中携带的信息。设备上，Option 184 携带了语音呼叫所需的信息。通过 Option 184，可以实现在为具有语音功能的 DHCP 客户端分配 IP 地址的同时，为客户端提供语音呼叫相关信息。

目前 Option 184 支持四个子选项，承载的内容如下：

- sub-option 1，网络呼叫处理器的 IP 地址：用来标识作为网络呼叫控制源及应用程序下载的服务器。
- sub-option 2，备用服务器的 IP 地址：当 sub-option 1 中携带的网络呼叫处理器不可达或不合法时，DHCP 客户端使用该选项指定的备用服务器作为网络呼叫处理器。
- sub-option 3，语音 VLAN 信息：指定语音 VLAN 的 ID 及 DHCP 客户端是否会将所指定的 VLAN ID 作为语音 VLAN。
- sub-option 4，自动故障转移呼叫路由：指定故障转移呼叫路由的 IP 地址及其关联的拨号串，即 SIP（Session Initiation Protocol，会话初始协议）用户之间互相通信时对端的 IP 地址和呼叫号码。当网络呼叫处理器和备用服务器均不可达时，SIP 用户可以使用对端 IP 地址及呼叫号码直接与对端 SIP 用户建立连接并通信。

说明

只有定义了 sub-option 1（网络呼叫处理器的 IP 地址子选项），其他子选项才能生效。

1.5 协议规范

与 DHCP 相关的协议规范有：

- RFC 2131：Dynamic Host Configuration Protocol
- RFC 2132：DHCP Options and BOOTP Vendor Extensions
- RFC 1542：Clarifications and Extensions for the Bootstrap Protocol
- RFC 3046：DHCP Relay Agent Information Option
- RFC 3442：The Classless Static Route Option for Dynamic Host Configuration Protocol (DHCP) version 4

2 DHCP服务器配置

说明

DHCP 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”。

2.1 DHCP服务器简介

2.1.1 DHCP服务器的应用环境

在以下场合通常利用 DHCP 服务器来完成 IP 地址分配：

- 网络规模较大，手工配置需要很大的工作量，并难以对整个网络进行集中管理。
- 网络中主机数目大于该网络支持的 IP 地址数量，无法给每个主机分配一个固定的 IP 地址。例如，Internet 接入服务提供商限制同时接入网络的用户数目，大量用户必须动态获得自己的 IP 地址。
- 网络中只有少数主机需要固定的 IP 地址，大多数主机没有固定的 IP 地址需求。

说明

设备作为 MCE（Multi-VPN-instance Customer Edge，多实例用户网络边界设备）时，在设备上配置 DHCP 服务器功能，不仅可以为公网上的 DHCP 客户端分配 IP 地址，还可以实现为私网内的 DHCP 客户端分配 IP 地址，但是公网和私网之间、不同私网之间的 IP 地址空间不能重叠。MCE 的详细介绍，请参见“MPLS 配置指导”中的“MPLS L3VPN”。

2.1.2 DHCP地址池

1. 地址池分类

DHCP 地址池分为两类：

- 普通模式的地址池：支持静态绑定和动态分配两种地址分配方式。
- 扩展模式的地址池：只支持地址的动态分配方式。

2. 普通模式地址池结构

DHCP 服务器从地址池中为客户端选择并分配 IP 地址及其他相关参数。

DHCP 服务器的普通模式地址池采用树状结构：树根是自然网段的地址池，分支是该网段的子网地址池，叶节点是手工绑定的客户端地址。同一级别地址池的顺序由配置的先后决定。这种树状结构实现了配置的继承性，即子网配置继承自然网段的配置，客户端的配置继承子网的配置。这样，对

于一些通用参数（如 DNS 服务器地址），只需要在自然网段或者子网上配置即可。具体的继承情况如下：

(1) 在父子关系建立时，子地址池将会继承父地址池的已有配置。

(2) 在父子关系建立后，对父地址池进行的配置，子地址池是否会继承，则有下面两种情况：

- 如果子地址池没有该项配置，则继承父地址池的配置；
- 如果子地址池已有该项配置，则不会继承父地址池的配置。

说明

- 设备独立存储每个扩展模式地址池。扩展地址池之间不存在父子关系和继承关系。
- IP 地址的租用有效期限不具有继承关系。

3. 地址池的选取原则

DHCP 服务器为客户端分配 IP 地址时，地址池的选择原则如下：

(1) 如果存在将客户端MAC地址或客户端ID与IP地址静态绑定的地址池，则选择该地址池，并将静态绑定的IP地址分配给客户端。该地址池的配置方法请参见“2.3.3 1. 配置采用静态绑定方式进行地址分配”。

(2) 如果接收到DHCP请求报文的接口引用了扩展模式的地址池，则选择该地址池，从该地址池中选取IP地址分配给客户端。如果该地址池中没有可供分配的IP地址，则服务器无法为客户端分配IP地址。该地址池的配置方法请参见“2.3.4 配置扩展模式地址池采用动态分配方式进行地址分配”。

(3) 如果不存在静态绑定的地址池，且接收到DHCP请求报文的接口没有引用扩展模式的地址池，则选择包含DHCP请求报文接收接口的IP地址（客户端与服务器在同一网段时）或DHCP请求报文中giaddr字段指定的IP地址（客户端与服务器不在同一网段，客户端通过DHCP中继获取IP地址时）、网段最小的普通模式地址池。如果该地址池中没有可供分配的IP地址，则服务器无法为客户端分配地址，服务器不会将父地址池中的IP地址分配给客户端。该地址池的配置方法请参见“2.3.3 2. 配置采用动态分配方式进行地址分配”。

例如，DHCP 服务器上配置了两个普通模式的地址池，动态分配的网段分别是 1.1.1.0/24 和 1.1.1.0/25，如果接收 DHCP 请求报文的接口 IP 地址为 1.1.1.1/25，服务器将从 1.1.1.0/25 地址池中选择 IP 地址分配给客户端，1.1.1.0/25 地址池中如果没有可供分配的 IP 地址，则服务器无法为客户端分配地址；如果接收 DHCP 请求报文的接口 IP 地址为 1.1.1.130/25，服务器将从 1.1.1.0/24 地址池中选择 IP 地址分配给客户端。

说明

配置地址池动态分配的网段和 IP 地址范围时，请尽量保证与 DHCP 服务器接口或 DHCP 中继接口地址的网段一致，以免分配错误的 IP 地址。

2.1.3 DHCP服务器分配IP地址的优先次序

DHCP 服务器为客户端分配 IP 地址的优先次序如下：

(1) 与客户端 MAC 地址或客户端 ID 静态绑定的 IP 地址。

(2) DHCP 服务器记录的曾经分配给客户端的 IP 地址。

(3) 客户端发送的 DHCP-DISCOVER 报文中 Option 50 字段指定的 IP 地址。

(4) 在扩展或普通模式的动态分配地址池中，顺序查找可供分配的 IP 地址，最先找到的 IP 地址。

(5) 如果未找到可用的 IP 地址，则依次查询租约过期、曾经发生过冲突的 IP 地址，如果找到则进行分配，否则将不予处理。

说明

Option 50 为客户端请求的 IP 地址选项（Requested IP Address），客户端通过在 DHCP-DISCOVER 报文中添加该选项来指明客户端希望获取的 IP 地址。该选项的内容由客户端决定。

2.2 DHCP服务器配置任务简介

表2-1 DHCP 服务器配置任务简介

操作	说明	详细配置
配置DHCP服务器的地址池	必选	2.3
使能DHCP服务	必选	2.4
配置接口工作在DHCP服务器模式	必选	2.5
配置接口引用扩展模式的地址池	使用扩展模式的地址池时，为必选；使用普通模式的地址池时，不需要执行该配置任务	2.6
配置DHCP服务器的安全功能	可选	2.7
配置DHCP服务器的用户下线检测功能	可选	2.8
配置Option 82的处理方式	可选	2.9
配置DHCP服务器发送Trap消息	可选	2.10
配置DHCP报文的DSCP优先级	可选	2.11

2.3 配置DHCP服务器的地址池

2.3.1 DHCP服务器地址池配置任务简介

表2-2 DHCP 服务器地址池配置任务简介

操作		说明	详细配置
创建DHCP地址池		必选	2.3.2
配置普通模式地址池的地址分配方式	配置采用静态绑定方式进行地址分配	对于普通模式地址池，二者必选其一，且对同一个地址池只能选一种方式	2.3.3
	配置采用动态分配方式进行地址分配		

操作	说明	详细配置
配置扩展模式地址池采用动态分配方式进行地址分配	对于扩展模式地址池，为必选	2.3.4
配置DHCP客户端的域名后缀	可选	2.3.5
配置DHCP客户端的DNS服务器地址	可选	2.3.6
配置DHCP客户端的WINS服务器地址和NetBIOS节点类型	可选	2.3.7
配置DHCP客户端的BIMS服务器信息	可选	2.3.8
配置DHCP客户端的网关地址	可选	2.3.9
配置DHCP客户端的Option 184参数	可选	2.3.10
配置DHCP客户端的TFTP服务器地址及启动文件名	可选	2.3.11
配置DHCP客户端的下一个提供服务的服务器IP地址	可选	2.3.12
配置DHCP自定义选项	可选	2.3.13

2.3.2 创建DHCP地址池

创建 DHCP 地址池时，需要指定地址池的类型：普通模式地址池或扩展模式地址池。

表2-3 创建 DHCP 地址池

操作	命令	说明
进入系统视图	**system-view**	-
创建DHCP地址池并进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	必选 缺省情况下，没有创建DHCP地址池

说明

普通模式地址池和扩展模式地址池中地址分配方式的配置方法不同，其余网络参数（如域名后缀、DNS 服务器地址等）的配置方法相同。

2.3.3 配置普通模式地址池的地址分配方式

注意

根据客户端的实际需要，可以将地址池配置为采用静态绑定或动态分配方式进行地址分配，但对一个 DHCP 地址池不能同时配置这两种方式。

动态地址分配需要指定用于分配的网段，而静态地址绑定则可以看作是只包含一个地址的特殊的DHCP 地址池。

1. 配置采用静态绑定方式进行地址分配

某些客户端（如 WWW 服务器等）需要固定的 IP 地址，可以通过将客户端的 MAC 地址与 IP 地址绑定的方式实现。当具有此 MAC 地址的客户端申请 IP 地址时，DHCP 服务器将根据客户端的 MAC 地址查找到对应的 IP 地址，并分配给客户端。

某些客户端在向 DHCP 服务器发送 DHCP-DISCOVER 报文申请 IP 地址时，会构建客户端 ID 并添加到报文中一起发送。如果在 DHCP 服务器上将客户端 ID 与 IP 地址绑定，则当该客户端申请 IP 地址时，DHCP 服务器将根据客户端 ID 查找到对应的 IP 地址并分配给客户端。

目前一个 DHCP 地址池中只能配置一个静态绑定，可以是 IP 地址与 MAC 地址的绑定，也可以是 IP 地址与客户端 ID 的绑定。

表2-4 配置采用静态绑定方式进行地址分配

操作		命令	说明
进入系统视图		**system-view**	-
进入普通模式地址池视图		**dhcp server ip-pool** *pool-name*	-
配置静态绑定的IP地址		**static-bind ip-address** *ip-address* [*mask-length* | **mask** *mask*]	必选 缺省情况下，没有配置静态绑定的IP地址
配置静态绑定的MAC地址或客户端ID	配置静态绑定的客户端MAC地址	**static-bind mac-address** *mac-address*	二者必选其一 缺省情况下，没有配置静态绑定的MAC地址和客户端ID
	配置静态绑定的客户端ID	**static-bind client-identifier** *client-identifier*	
配置静态绑定IP地址的租用有效期限		**expired** { **day** *day* [**hour** *hour* [**minute** *minute* [**second** *second*]]] | **unlimited** }	可选 缺省情况下，IP地址租用有效期限为**unlimited**，即租约期限为无限长

说明

- **static-bind ip-address** 和 **static-bind mac-address** 或 **static-bind client-identifier** 命令必须配合使用。
- 在同一个 DHCP 地址池中，如果配置了 **static-bind mac-address** 命令后，再配置 **static-bind client-identifier** 命令，则后面的配置将会覆盖前面的配置，反之亦然。
- 在同一个 DHCP 地址池中，如果多次执行 **static-bind ip-address** 或 **static-bind mac-address** 或 **static-bind client-identifier** 命令，新的配置会覆盖已有配置。
- 静态绑定的 IP 地址不能是 DHCP 服务器的接口 IP 地址，否则会导致 IP 地址冲突，被绑定的客户端将无法正常获取到 IP 地址。
- 静态绑定的客户端 ID，要与在待绑定客户端通过 **display dhcp client verbose** 命令显示的客户端 ID 一致。否则，客户端无法成功获取 IP 地址。
- 静态绑定方式的 DHCP 地址池可以配置租期，且租期会生效。但通过命令 **display dhcp server ip-in-use all** 查看时，显示的租期还是会是 Unlimited，不会随配置改变。
- 设备作为 DHCP 客户端时，需要在 DHCP 服务器上配置 IP 地址与客户端 ID 静态绑定，设备作为 BOOTP 客户端时，需要配置 IP 地址与客户端 MAC 静态绑定，否则设备无法申请到静态绑定的 IP 地址。
- 如果作为 DHCP 客户端的设备，接口的 MAC 地址相同，则为了区分不同接口，采用静态绑定方式进行地址分配时，需要在服务器上配置静态绑定的客户端 ID，而不能配置静态绑定的客户端 MAC 地址，否则可能导致客户端无法成功获取 IP 地址。

2. 配置采用动态分配方式进行地址分配

对于采用动态地址分配方式的地址池，需要配置该地址池可分配的网段。目前，一个地址池中只能配置一个地址段。

DHCP 服务器在分配地址时，需要排除已经被占用的 IP 地址（如网关、FTP 服务器等）。否则，同一地址分配给两个客户端会造成 IP 地址冲突。

DHCP 服务器可以为不同的地址池指定不同的地址租用期限，但同一 DHCP 地址池中的所有地址都具有相同的期限。地址租用有效期限不具有继承关系。

表2-5 配置采用动态分配方式进行地址分配

操作	命令	说明
进入系统视图	**system-view**	-
进入普通模式地址池视图	**dhcp server ip-pool** *pool-name*	-
配置动态分配的网段	**network** *network-address* [*mask-length* | **mask** *mask*]	必选 缺省情况下，没有配置动态分配的网段，即没有可供分配的IP地址

操作	命令	说明
在网段的基础上进一步限制动态分配的IP地址范围	**network ip range** *min-address max-address*	可选 缺省情况下，没有配置动态分配的IP地址范围
配置动态分配的IP地址的租用有效期限	**expired** { **day** *day* [**hour** *hour* [**minute** *minute* [**second** *second*]]] | **unlimited** }	可选 缺省情况下，IP地址租用有效期限为1天
退回系统视图	**quit**	-
配置DHCP地址池中不参与自动分配的IP地址	**dhcp server forbidden-ip** *low-ip-address* [*high-ip-address*]	可选 缺省情况下，除DHCP服务器接口的IP地址外，DHCP地址池中的所有IP地址都参与自动分配

说明

- 在同一个 DHCP 地址池中，如果多次执行 **network** 或 **network ip range** 命令，新的配置会覆盖已有配置。
- 通过 **dhcp server forbidden-ip** 命令指定不参与自动分配的 IP 地址后，所有动态分配方式的地址池（包括普通模式地址池和扩展模式地址池）都不能分配这些 IP 地址。
- 多次执行 **dhcp server forbidden-ip** 命令，可以配置多个不参与自动分配的 IP 地址段。

2.3.4 配置扩展模式地址池采用动态分配方式进行地址分配

扩展模式地址池只支持动态方式分配地址，不支持静态绑定方式。

配置扩展模式地址池的地址分配方式时，需要指定：

- 可分配的 IP 地址范围：扩展模式地址池中采用最小 IP 地址、最大 IP 地址的方式指定地址范围。
- 分配的 IP 地址的掩码。

只有同时配置了可分配的 IP 地址范围和掩码，该地址池才会生效。

表2-6 配置扩展模式地址池采用动态分配方式进行地址分配

操作	命令	说明
进入系统视图	**system-view**	-
进入扩展模式地址池视图	**dhcp server ip-pool** *pool-name* **extended**	-
配置动态分配的IP地址范围	**network ip range** *min-address max-address*	必选 缺省情况下，没有配置动态分配的IP地址范围，即没有可供分配的IP地址

操作	命令	说明
配置动态分配的IP地址掩码	**network mask** *mask*	必选 缺省情况下，没有配置动态分配的IP地址掩码
配置为指定厂商的DHCP客户端动态分配的IP地址范围	**vendor-class-identifier** *hex-string*&<1-255> **ip range** *min-address max-address*	可选 缺省情况下，没有配置DHCP扩展地址池为指定厂商的客户端动态分配的IP地址范围
配置动态分配的IP地址的租用有效期限	**expired** { **day** *day* [**hour** *hour* [**minute** *minute* [**second** *second*]]] | **unlimited** }	可选 缺省情况下，IP地址租用有效期限为1天
配置该地址池中不参与自动分配的IP地址	**forbidden-ip** *ip-address*&<1-8>	可选 缺省情况下，除DHCP服务器接口的IP地址外，扩展模式地址池中的所有IP地址都参与自动分配

说明

在扩展模式的 DHCP 地址池视图下通过 **forbidden-ip** 命令配置不参与自动分配的 IP 地址后，只有当前扩展模式的地址池不能分配这些 IP 地址，其他地址池仍然可以分配这些 IP 地址。

2.3.5 配置DHCP客户端的域名后缀

在 DHCP 服务器上，可以为每个地址池指定客户端使用的域名后缀。在给客户端分配 IP 地址的同时，也将域名后缀发送给客户端。

在客户端进行域名解析时，用户只需要输入域名的部分字段，客户端会自动将输入的域名加上域名后缀进行解析。有关域名后缀的详细介绍，请参见“三层技术-IP 业务配置指导”中的“IPv4 域名解析”。

表2-7 配置 DHCP 客户端的域名后缀

操作	命令	说明
进入系统视图	**system-view**	-
进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	-
配置为DHCP客户端分配的域名后缀	**domain-name** *domain-name*	必选 缺省情况下，没有配置为DHCP客户端分配的域名后缀

2.3.6 配置DHCP客户端的DNS服务器地址

通过域名访问 Internet 上的主机时，需要将域名解析为 IP 地址，这是通过 DNS（Domain Name System，域名系统）实现的。为了使 DHCP 客户端能够通过域名访问 Internet 上的主机，DHCP 服务器应在为客户端分配 IP 地址的同时指定 DNS 服务器地址。目前，每个 DHCP 地址池最多可以配置 8 个 DNS 服务器地址。

表2-8 配置 DHCP 客户端的 DNS 服务器地址

操作	命令	说明
进入系统视图	**system-view**	-
进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	-
配置为DHCP客户端分配的DNS服务器地址	**dns-list** *ip-address*&<1-8>	必选 缺省情况下，没有配置为DHCP客户端分配的DNS服务器地址

2.3.7 配置DHCP客户端的WINS服务器地址和NetBIOS节点类型

对于使用 Microsoft Windows 操作系统的客户端，由 WINS（Windows Internet Naming Service，Windows Internet 名称服务）服务器为通过 NetBIOS 协议通信的主机提供主机名到 IP 地址的解析。所以，大部分 Windows 网络客户端需要进行 WINS 的设置。

为了使 DHCP 客户端实现主机名到 IP 地址的解析，DHCP 服务器应在为客户端分配 IP 地址的同时指定 WINS 服务器地址。目前，每个 DHCP 地址池最多可以配置 8 个 WINS 服务器地址。

DHCP 客户端在网络上使用 NetBIOS 协议通信时，需要在主机名和 IP 地址之间建立映射关系。根据获取映射关系方式的不同，NetBIOS 节点分为四种：

- b 类节点（b-node）：“b”代表广播（broadcast），即此类节点采用广播方式获取映射关系。源节点通过发送带有目的节点主机名的广播报文来获取目的节点的 IP 地址，目的节点收到广播报文后，就将自己的 IP 地址返回给源节点。
- p 类节点（p-node）：“p”代表端到端（peer-to-peer），即此类节点采用发送单播报文与 WINS 服务器通信的方式获取映射关系。源节点给 WINS 服务器发送单播报文，WINS 服务器收到单播报文后，返回源节点请求的目的节点名所对应的 IP 地址。
- m 类节点（m-node）：“m”代表混合（mixed），是具有部分广播特性的 p 类节点。即此类节点首先发送广播报文来获取映射关系，如果没有获取到，则再发送单播报文与 WINS 服务器通信来获取映射关系。
- h 类节点（h-node）：“h”代表混合（hybrid），是具备“端到端”通信机制的 b 类节点。即此类节点首先发送单播报文与 WINS 服务器通信来获取映射关系，如果没有获取到，再发送广播报文来获取映射关系。

表2-9 配置 DHCP 客户端的 WINS 服务器地址和 NetBIOS 节点类型

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	-
配置为DHCP客户端分配的WINS服务器地址	**nbns-list** *ip-address*&<1-8>	必选 对于b类节点，为可选 缺省情况下，没有配置为DHCP客户端分配的WINS服务器地址
配置为DHCP客户端分配的NetBIOS节点类型	**netbios-type** { **b-node** | **h-node** | **m-node** | **p-node** }	必选 缺省情况下，没有配置为DHCP客户端分配的NetBIOS节点类型

说明

如果配置 DHCP 客户端为 b 类 NetBIOS 节点，则不需要配置 WINS 服务器地址。

2.3.8 配置DHCP客户端的BIMS服务器信息

为了使 DHCP 客户端通过 BIMS（Branch Intelligent Management System，分支网点智能管理系统）服务器进行软件的备份和升级等操作，DHCP 服务器需要在为 DHCP 客户端分配 IP 地址的同时，将 BIMS 服务器的 IP 地址、端口号以及加密的共享密钥等信息也发给 DHCP 客户端。之后，DHCP 客户端就可以定期向 BIMS 服务器发送连接请求，从 BIMS 服务器上获取配置文件，进行软件的备份和升级等操作。

表2-10 配置 DHCP 客户端的 BIMS 服务器信息

操作	命令	说明
进入系统视图	**system-view**	-
进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	-
配置为DHCP客户端分配的BIMS服务器的IP地址、端口及共享密钥信息	**bims-server ip** *ip-address* [**port** *port-number*] **sharekey** [**cipher** | **simple**] *key*	必选 缺省情况下，没有配置为DHCP客户端分配的BIMS服务器信息

2.3.9 配置DHCP客户端的网关地址

DHCP 客户端访问本网段以外的服务器或主机时，数据必须通过网关进行转发。DHCP 服务器可以在为客户端分配 IP 地址的同时指定网关的地址。

在 DHCP 服务器上，可以为每个地址池分别指定客户端对应的网关地址。在给客户端分配 IP 地址的同时，也将网关地址发送给客户端。目前，每个 DHCP 地址池最多可以配置 8 个网关地址。

表2-11 配置 DHCP 客户端的网关地址

操作	命令	说明
进入系统视图	**system-view**	-
进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	-
配置为DHCP客户端分配的网关地址	**gateway-list** *ip-address*&<1-8>	必选 缺省情况下，没有配置为DHCP客户端分配的网关地址

2.3.10 配置DHCP客户端的Option 184 参数

为了使具有语音功能的DHCP客户端能够在通过DHCP获取IP地址的同时，获取到语音呼叫所需的相关信息，需要在DHCP服务器上配置Option 184。Option 184 内容的详细介绍，请参见“1.4.3 3. Option 184”。

表2-12 配置 DHCP 客户端的 Option 184 参数

操作	命令	说明
进入系统视图	**system-view**	-
进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	-
配置网络呼叫处理器的地址	**voice-config ncp-ip** *ip-address*	必选 缺省情况下，没有配置网络呼叫处理器的地址
配置备用服务器的地址	**voice-config as-ip** *ip-address*	可选 缺省情况下，没有配置备用服务器的地址
配置语音VLAN	**voice-config voice-vlan** *vlan-id* { **disable** | **enable** }	可选 缺省情况下，没有配置语音VLAN
配置自动故障转移呼叫路由	**voice-config fail-over** *ip-address dialer-string*	可选 缺省情况下，没有配置自动故障转移呼叫路由

说明

只有配置了网络呼叫处理器的地址，其他配置才能生效。

2.3.11 配置DHCP客户端的TFTP服务器地址及启动文件名

设备在空配置启动时自动获取并执行配置文件的功能，被称为自动配置。具体过程如下：

(1) 设备在空配置启动时，系统会自动将处于 up 状态的接口（如缺省 VLAN 对应的虚接口）设置为 DHCP 客户端，并向 DHCP 服务器获取 IP 地址及后续获取配置文件所需要的信息（例如：TFTP 服务器的 IP 地址、TFTP 服务器名、启动文件名等）。

(2) 如果获取到相关信息，则 DHCP 客户端就可发起 TFTP 请求，从指定的 TFTP 服务器获取配置文件，之后设备就使用获取到的配置文件进行设备初始化工作。如果没有获取到相关信息，则设备在空配置的情况下正常启动。

自动配置功能在空配置启动的设备上不需要进行任何配置，但需要在 DHCP 服务器上配置一些必需的参数，包括 TFTP 服务器地址、TFTP 服务器名和启动文件名。

表2-13 配置 DHCP 客户端的 TFTP 服务器地址及启动文件名

操作	命令	说明
进入系统视图	**system-view**	-
进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	-
配置为DHCP客户端分配的TFTP服务器地址	**tftp-server ip-address** *ip-address*	二者至少选择其一 缺省情况下，没有配置为DHCP客户端分配的TFTP服务器地址和TFTP服务器名
配置为DHCP客户端分配的TFTP服务器名	**tftp-server domain-name** *domain-name*	
配置为DHCP客户端分配的启动文件名	**bootfile-name** *bootfile-name*	必选 缺省情况下，没有配置为DHCP客户端分配的启动文件名

2.3.12 配置DHCP客户端的下一个提供服务的服务器IP地址

设备在启动后，可能需要访问某些服务器获取设备运行需要的信息，例如从 TFTP 服务器上获取配置文件。通过本配置可以指定 DHCP 服务器为 DHCP 客户端分配的下一个提供服务的服务器 IP 地址，以便客户端启动后访问该服务器，获取必要的信息。

表2-14 配置 DHCP 客户端的下一个提供服务的服务器 IP 地址

操作	命令	说明
进入系统视图	**system-view**	-
进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	-
配置DHCP地址池为DHCP客户端分配的下一个提供服务的服务器IP地址	**next-server** *ip-address*	必选 缺省情况下，没有配置DHCP地址池为DHCP客户端分配的下一个提供服务的服务器IP地址

2.3.13 配置DHCP自定义选项

通过配置 DHCP 自定义选项，用户可以：

- 定义新的 DHCP 选项。随着 DHCP 的不断发展，新的可选项会陆续出现，为了支持这些新的选项，可以通过手工定义的方式将新选项添加到 DHCP 服务器的属性列表中。
- 定义已有选项的内容。有些选项的内容，RFC 2132 中没有统一规定。厂商可以根据需要定义选项的内容，如 Option 43。通过配置 DHCP 自定义选项，可以为 DHCP 客户端提供厂商指定的信息。
- 扩展已有的 DHCP 选项。当前已提供的方式无法满足用户需求时（比如通过 **dns-list** 命令最多只能配置 8 个 DNS 服务器地址，如果用户需要配置的 DNS 服务器地址数目大于 8，则该命令无法满足需求），可以通过配置 DHCP 自定义选项的方式进行扩展。

表2-15 配置 DHCP 自定义选项

操作	命令	说明
进入系统视图	**system-view**	-
进入DHCP地址池视图	**dhcp server ip-pool** *pool-name* [**extended**]	-
配置DHCP自定义选项	**option** *code* { **ascii** *ascii-string* | **hex** *hex-string*&<1-16> | **ip-address** *ip-address*&<1-8> }	必选 缺省情况下，没有配置DHCP自定义选项

表2-16 常用 Option 配置说明

选项代码	选项名称	对应的配置命令	option 命令参数选择
3	Router Option	**gateway-list**	**ip-address**
6	Domain Name Server Option	**dns-list**	**ip-address**
15	Domain Name	**domain-name**	**ascii**
44	NetBIOS over TCP/IP Name Server Option	**nbns-list**	**ip-address**
46	NetBIOS over TCP/IP Node Type Option	**netbios-type**	**hex**
66	TFTP server name	**tftp-server**	**ascii**
67	Bootfile name	**bootfile-name**	**ascii**
43	Vendor Specific Information	-	**hex**

注意

配置自定义选项可能会对 DHCP 的工作过程造成影响，请谨慎使用 DHCP 自定义选项。

2.4 使能DHCP服务

只有使能 DHCP 服务后，其它相关的 DHCP 服务器配置才能生效。

表2-17 使能 DHCP 服务

操作	命令	说明
进入系统视图	**system-view**	-
使能DHCP服务	**dhcp enable**	必选 缺省情况下，DHCP服务处于禁止状态

2.5 配置接口工作在DHCP服务器模式

配置接口工作在 DHCP 服务器模式后，当接口收到 DHCP 客户端发来的 DHCP 报文时，将从 DHCP 服务器的地址池中分配地址。

表2-18 配置接口工作在 DHCP 服务器模式

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*	-
配置接口工作在DHCP服务器模式	**dhcp select server global-pool** [**subaddress**]	可选 缺省情况下，接口工作在DHCP服务器模式

说明

subaddress 关键字只在客户端与 DHCP 服务器之间通信不需要通过三层路由转发时有效。如果客户端与服务器之间通过 DHCP 中继相连，则不管是否输入 **subaddress** 关键字，当服务器为客户端分配 IP 地址时，都是从与 DHCP 中继接口（与客户端相连的接口）的主 IP 地址在同一网段的地址池中选择地址分配给客户端。

说明

当客户端与 DHCP 服务器之间通信不需要通过三层路由转发时，根据是否输入 **subaddress** 关键字，DHCP 服务器的处理不同：

- 如果输入 **subaddress** 关键字，当服务器为客户端分配 IP 地址时，将优先从与服务器接口（与客户端相连的接口）的主 IP 地址在同一网段的地址池中选择地址分配给客户端，如果该地址池中没有可供分配的 IP 地址，则从与服务器接口的从 IP 地址在同一网段的地址池中选择地址分配给客户端。如果接口有多个从 IP 地址，则从第一个从 IP 地址依次匹配。
- 如果不输入 **subaddress** 关键字，当服务器为客户端分配 IP 地址时，将只能从与服务器接口的主 IP 地址在同一网段的地址池中选择地址分配给客户端。

2.6 配置接口引用扩展模式的地址池

创建扩展模式的地址池，并在接口引用该地址池后，接口接收到 DHCP 请求，将优先为客户端分配静态绑定的 IP 地址；如果不存在静态绑定的 IP 地址，则从引用的扩展模式地址池中选择 IP 地址分配给客户端。引用的地址池中不存在可供分配的 IP 地址时，设备无法为客户端分配 IP 地址，设备不会将其他地址池中的地址分配给客户端。

表2-19 配置接口引用扩展模式的地址池

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*	-
配置接口引用扩展模式的地址池	**dhcp server apply ip-pool** *pool-name*	可选 缺省情况下，接口没有引用任何扩展模式的地址池，接口接收到DHCP请求后，将从普通模式地址池中选择IP地址分配给客户端

说明

接口只能引用扩展模式的地址池，且引用的地址池必须已经存在。

2.7 配置DHCP服务的安全功能

在配置 DHCP 服务器后，为了提高 DHCP 服务的安全性，需要配置 DHCP 服务的安全功能。

2.7.1 配置准备

在配置 DHCP 服务的安全功能之前，需完成 DHCP 服务器的必配任务：

- 使能 DHCP 服务；
- 配置 DHCP 服务器的地址池。

2.7.2 配置伪DHCP服务器检测功能

如果网络中有私自架设的 DHCP 服务器，当客户端申请 IP 地址时，这台 DHCP 服务器就会与 DHCP 客户端进行交互，导致客户端获得错误的 IP 地址。这种私设的 DHCP 服务器称为伪 DHCP 服务器。

使能伪 DHCP 服务器检测功能后，DHCP 服务器会检查接收到的 DHCP 报文中是否携带 Option 54（Server Identifier Option，服务器标识选项）。如果携带该选项，则 DHCP 服务器记录此选项中的 IP 地址，即给客户端分配 IP 地址的服务器 IP 地址，并记录接收到 DHCP 报文的接口信息，以便管理员及时发现并处理伪 DHCP 服务器。

表2-20 配置伪 DHCP 服务器检测功能

操作	命令	说明
进入系统视图	**system-view**	-
使能伪DHCP服务器检测功能	**dhcp server detect**	必选 缺省情况下，禁止伪DHCP服务器检测功能

说明

- 使能伪 DHCP 服务器检测功能后，对所有 DHCP 服务器都会进行记录，包括合法的 DHCP 服务器。管理员需要从日志信息中查找伪 DHCP 服务器。
- 使能伪 DHCP 服务器检测功能后，对每个 DHCP 服务器只记录一次。

2.7.3 配置IP地址重复分配检测功能

为防止 IP 地址重复分配导致地址冲突，DHCP 服务器为客户端分配地址前，需要先对该地址进行探测。

地址探测是通过 ping 功能实现的，通过检测是否能在指定时间内得到 ping 响应来判断是否有地址冲突。DHCP 服务器发送目的地址为待分配地址的 ICMP 回显请求报文，如果在指定时间内收到回显响应报文，DHCP 服务器从地址池中选择新的 IP 地址，并重复上述操作；如果在指定时间内没有收到回显响应报文，则继续发送 ICMP 回显请求报文，直到发送的回显请求报文达到最大值，如果仍然没有收到回显响应报文，则将地址分配给客户端，从而确保客户端被分得的 IP 地址唯一。

DHCP 服务器通过 ping 操作来检测是否发生地址冲突，而 DHCP 客户端则通过发送免费 ARP 报文检测是否发生地址冲突。

表2-21 配置 IP 地址重复分配检测功能

操作	命令	说明
进入系统视图	**system-view**	-
配置DHCP服务器发送回显请求报文的最大数目	**dhcp server ping packets** *number*	可选 缺省情况下，DHCP服务器发送回显请求报文的最大数目为1 0表示不进行ping操作
配置DHCP服务器等待回显响应报文的超时时间	**dhcp server ping timeout** *milliseconds*	可选 缺省情况下，DHCP服务器等待回显响应报文的超时时间为500毫秒 0表示不进行ping操作

2.8 配置DHCP服务器的用户下线检测功能

DHCP 服务器的用户下线检测功能以 ARP 表项老化功能为基础，当 ARP 表项老化时认为该表项对应的用户已经下线。

如果在接口上使能了 DHCP 服务器的用户下线检测功能，则当 ARP 表项老化时，系统会删除该表项对应用户的地址绑定信息。

表2-22 配置 DHCP 服务器的用户下线检测功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能DHCP服务器的用户下线检测功能	**dhcp server client-detect enable**	必选 缺省情况下，DHCP服务器的用户下线检测功能处于关闭状态

说明

手工删除 ARP 表项，不会触发删除对应的地址绑定信息。

2.9 配置Option 82的处理方式

如果配置 DHCP 服务器处理 Option 82，则当 DHCP 服务器收到带有 Option 82 的报文后，会在响应报文中携带 Option 82，并为客户端分配 IP 地址等信息。

如果配置 DHCP 服务器忽略 Option 82，则当 DHCP 服务器收到带有 Option 82 的报文后，不会在响应报文中携带 Option 82，只为客户端分配 IP 地址等信息。

1. 配置准备

在配置 Option 82 的处理方式之前，需完成 DHCP 服务器的必配任务：

- 使能 DHCP 服务；
- 配置 DHCP 服务器的地址池。

2. 配置Option 82 的处理方式

表2-23 配置 Option 82 的处理方式

操作	命令	说明
进入系统视图	**system-view**	-
配置DHCP服务器处理Option 82	**dhcp server relay information enable**	可选 缺省情况下，DHCP服务器处理Option 82

说明

为使Option 82 功能正常使用，需要在DHCP服务器和DHCP中继（或DHCP Snooping设备）上都进行相应配置。DHCP中继支持Option 82 功能的相关配置请参见“3.3.7 配置DHCP中继支持Option 82 功能”；DHCP Snooping支持Option 82 功能的相关配置请参见“5.4 配置DHCP Snooping支持Option 82 功能”。

2.10 配置DHCP服务器发送Trap消息的门限值

2.10.1 配置准备

在配置 DHCP 服务器发送 Trap 消息之前，需要先通过 **snmp-agent target-host** 命令配置 Trap 消息的目的地址。**snmp-agent target-host** 命令的详细介绍，请参见“网络管理和监控命令参考”中的“SNMP”。

2.10.2 配置DHCP服务器发送Trap消息的门限值

在 DHCP 服务器上，可以配置成功分配 IP 地址的比率、地址池的平均 IP 利用率、地址池的最大 IP 利用率达到门限值时，向网管服务器发送 Trap 消息，以便网络管理员及时了解 DHCP 服务器的使用情况。

表2-24 配置 DHCP 服务器发送 Trap 消息的门限值

操作	命令	说明
进入系统视图	**system-view**	-
配置DHCP服务器在指定条件下，向网管服务器发送Trap消息	**dhcp server threshold** { **allocated-ip** *threshold-value* | **average-ip-use** *threshold-value* | **max-ip-use** *threshold-value* }	可选 缺省情况下，DHCP服务器不会向网管服务器发送Trap消息

2.11 配置DHCP服务器发送的DHCP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 DHCP 服务器发送的 DHCP 报文的 DSCP 优先级进行配置。

表2-25 配置 DHCP 服务器发送的 DHCP 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
配置DHCP服务器发送的DHCP报文的DSCP优先级	**dhcp dscp** *dscp-value*	可选 缺省情况下，DHCP服务器发送的DHCP报文的DSCP优先级为56

2.12 DHCP服务器显示和维护

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 DHCP 服务器的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令清除 DHCP 服务器的相关信息。

表2-26 DHCP 服务器显示和维护

操作	命令
显示DHCP的地址冲突统计信息	**display dhcp server conflict** { **all** | **ip** *ip-address* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP地址池中的租约超期信息	**display dhcp server expired** { **all** | **ip** *ip-address* | **pool** [*pool-name*] } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP地址池的可用地址信息	**display dhcp server free-ip** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP地址池中不参与自动分配的IP地址	**display dhcp server forbidden-ip** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP地址池中的地址绑定信息	**display dhcp server ip-in-use** { **all** | **ip** *ip-address* | **pool** [*pool-name*] } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP服务器的统计信息	**display dhcp server statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP地址池的树状结构信息	**display dhcp server tree** { **all** | **pool** [*pool-name*] } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除DHCP地址冲突的统计信息	**reset dhcp server conflict** { **all** | **ip** *ip-address* }
清除DHCP动态地址绑定信息	**reset dhcp server ip-in-use** { **all** | **ip** *ip-address* | **pool** [*pool-name*] }
清除DHCP服务器的统计信息	**reset dhcp server statistics**

说明

执行 **save** 命令时不会保存 DHCP 服务器的租约信息，故当系统重新启动或使用 **reset dhcp server ip-in-use** 命令清除租约后，配置文件中将没有任何租约信息。此时客户端如果发出续约请求将会被拒绝，客户端需要重新申请 IP 地址。

2.13 DHCP服务器典型配置举例

常见的 DHCP 组网方式可分为两类：一种是 DHCP 服务器和客户端在同一个子网内，直接进行 DHCP 报文的交互；第二种是 DHCP 服务器和客户端处于不同的子网中，必须通过 DHCP 中继代理实现 IP 地址的分配。无论哪种情况下，DHCP 的配置都是相同的。

2.13.1 静态绑定地址典型配置举例

1. 组网需求

Switch B 和 Switch C 分别作为 DHCP 客户端和 BOOTP 客户端，从 DHCP 服务器 Switch A 获取静态绑定的 IP 地址、域名服务器、网关地址等信息。

其中：

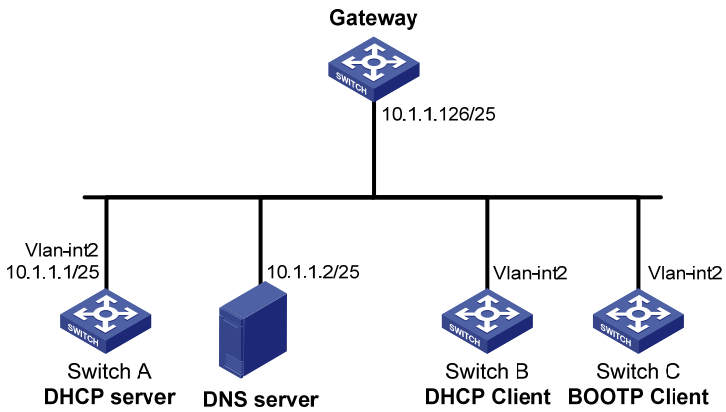
- Switch B 上 VLAN 接口 2 的客户端 ID 为：

3030-3066-2e65-3234-392e-3830-3530-2d56-6c61-6e2d-696e-7465-7266-6163-6532；

- Switch C 上 VLAN 接口 2 的 MAC 地址为：000f-e249-8050。

2. 组网图

图2-1 静态绑定地址组网图



3. 配置步骤

(1) 配置接口的 IP 地址

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ip address 10.1.1.1 25
[SwitchA-Vlan-interface2] quit
```

(2) 配置 DHCP 服务

使能 DHCP 服务。

```
[SwitchA] dhcp enable
```

配置 VLAN 接口 2 工作在 DHCP 服务器模式。

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] dhcp select server global-pool
[SwitchA-Vlan-interface2] quit
```

配置 DHCP 地址池 0，采用静态绑定方式为 Switch B 分配 IP 地址。

```
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] static-bind ip-address 10.1.1.5 25
[SwitchA-dhcp-pool-0]                    static-bind                    client-identifier
3030-3066-2e65-3234-392e-3830-3530-2d56-6c61-6e2d-696e-7465-7266-6163-6532
```

```
[SwitchA-dhcp-pool-0] dns-list 10.1.1.2
[SwitchA-dhcp-pool-0] gateway-list 10.1.1.126
[SwitchA-dhcp-pool-0] quit
```

配置 DHCP 地址池 1，采用静态绑定方式为 Switch C 分配 IP 地址。

```
[SwitchA] dhcp server ip-pool 1
[SwitchA-dhcp-pool-1] static-bind ip-address 10.1.1.6 25
[SwitchA-dhcp-pool-1] static-bind mac-address 000f-e249-8050
[SwitchA-dhcp-pool-1] dns-list 10.1.1.2
[SwitchA-dhcp-pool-1] gateway-list 10.1.1.126
```

(3) 验证配置结果

配置完成后，Switch B 和 Switch C 可以从 DHCP 服务器 Switch A 分别申请到 IP 地址 10.1.1.5 和 10.1.1.6，并获取相关网络配置参数。通过 **display dhcp server ip-in-use** 命令可以查看 DHCP 服务器为客户端分配的 IP 地址。

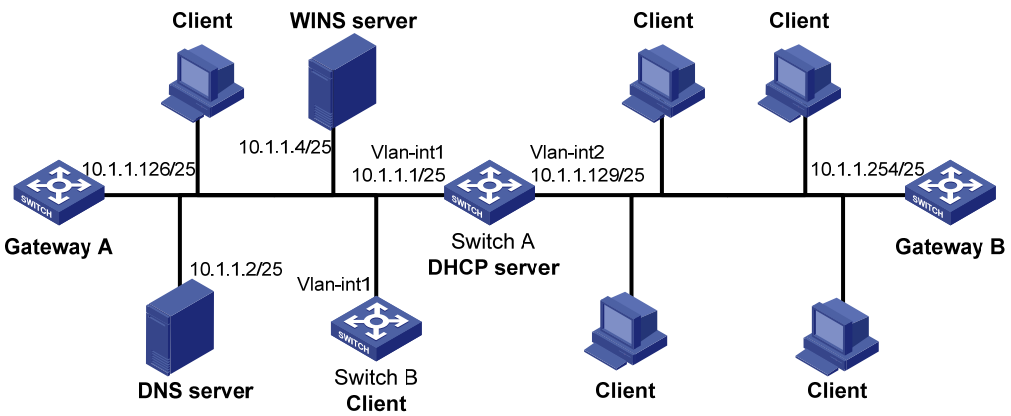
2.13.2 动态分配地址典型配置举例

1. 组网需求

- 作为 DHCP 服务器的 Switch A 为网段 10.1.1.0/24 中的客户端动态分配 IP 地址，该地址池网段分为两个子网网段：10.1.1.0/25 和 10.1.1.128/25；
- Switch A 的两个 VLAN 接口，VLAN 接口 1 和 VLAN 接口 2 的地址分别为 10.1.1.1/25 和 10.1.1.129/25；
- 10.1.1.0/25 网段内的地址租用期限为 10 天 12 小时，域名后缀为 aabbcc.com，DNS 服务器地址为 10.1.1.2/25，WINS 服务器地址为 10.1.1.4/25，网关的地址为 10.1.1.126/25；
- 10.1.1.128/25 网段内的地址租用期限为 5 天，域名后缀为 aabbcc.com，DNS 服务器地址为 10.1.1.2/25，无 WINS 服务器地址，网关的地址为 10.1.1.254/25。
- 10.1.1.0/25 网段与 10.1.1.128/25 网段的域名后缀、DNS 服务器地址相同，可以只配置 10.1.1.0/24 网段的域名后缀和 DNS 服务器地址，10.1.1.0/25 网段与 10.1.1.128/25 网段继承 10.1.1.0/24 网段的配置。

2. 组网图

图2-2 DHCP 组网图



3. 配置步骤

(1) 配置端口所属 VLAN 及对应 VLAN 接口的 IP 地址（略）

(2) 配置 DHCP 服务

使能 DHCP 服务。

```
<SwitchA> system-view
[SwitchA] dhcp enable
```

配置 VLAN 接口 1 和 VLAN 接口 2 工作在 DHCP 服务器模式。

```
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] dhcp select server global-pool
[SwitchA-Vlan-interface1] quit
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] dhcp select server global-pool
[SwitchA-Vlan-interface2] quit
```

配置不参与自动分配的 IP 地址（DNS 服务器、WINS 服务器和网关地址）。

```
[SwitchA] dhcp server forbidden-ip 10.1.1.2
[SwitchA] dhcp server forbidden-ip 10.1.1.4
[SwitchA] dhcp server forbidden-ip 10.1.1.126
[SwitchA] dhcp server forbidden-ip 10.1.1.254
```

配置 DHCP 地址池 0 的共有属性（网段、客户端域名后缀、DNS 服务器地址）。

```
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] network 10.1.1.0 mask 255.255.255.0
[SwitchA-dhcp-pool-0] domain-name aabbcc.com
[SwitchA-dhcp-pool-0] dns-list 10.1.1.2
[SwitchA-dhcp-pool-0] quit
```

配置 DHCP 地址池 1 的属性（网段、网关、地址租用期限、WINS 服务器地址）。

```
[SwitchA] dhcp server ip-pool 1
[SwitchA-dhcp-pool-1] network 10.1.1.0 mask 255.255.255.128
[SwitchA-dhcp-pool-1] gateway-list 10.1.1.126
[SwitchA-dhcp-pool-1] expired day 10 hour 12
[SwitchA-dhcp-pool-1] nbns-list 10.1.1.4
[SwitchA-dhcp-pool-1] quit
```

配置 DHCP 地址池 2 的属性（网段、地址租用期限、网关）。

```
[SwitchA] dhcp server ip-pool 2
[SwitchA-dhcp-pool-2] network 10.1.1.128 mask 255.255.255.128
[SwitchA-dhcp-pool-2] expired day 5
[SwitchA-dhcp-pool-2] gateway-list 10.1.1.254
```

(3) 验证配置结果

配置完成后，10.1.1.0/25 和 10.1.1.128/25 网段的客户端可以从 DHCP 服务器 Switch A 申请到相应网段的 IP 地址和网络配置参数。通过 **display dhcp server ip-in-use** 命令可以查看 DHCP 服务器为客户端分配的 IP 地址。

2.13.3 自定义选项典型配置举例

1. 组网需求

DHCP 客户端 Switch B 从 DHCP 服务器 Switch A 获取 IP 地址和 PXE 引导服务器地址信息：

- IP 地址所在网段为 10.1.1.0/24；
- PXE 引导服务器地址为 1.2.3.4 和 2.2.2.2。

DHCP服务器需要通过自定义选项的方式配置Option 43 的内容，从而实现为客户端分配PXE引导服务器地址。Option 43 和PXE服务器地址列表的格式分别如 图 1-5 和 图 1-7。DHCP服务器上配置的Option 43 选项内容为 80 0B 00 00 02 01 02 03 04 02 02 02 02，其中 80 为子选项类型（Sub-option type），0B为子选项长度（Sub-option length），00 00 为PXE服务器类型（PXE server type），02 为服务器数目（Server number），01 02 03 04 02 02 02 02 为服务器的IP地址 1.2.3.4 和 2.2.2.2。

2. 组网图

图2-3 自定义选项典型配置举例

3. 配置步骤

(1) 配置各接口的 IP 地址（略）

(2) 配置 DHCP 服务

使能 DHCP 服务。

```
<SwitchA> system-view
[SwitchA] dhcp enable
```

配置 VLAN 接口 2 工作在 DHCP 服务器模式。

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] dhcp select server global-pool
[SwitchA-Vlan-interface2] quit
```

配置 DHCP 地址池 0。

```
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] network 10.1.1.0 mask 255.255.255.0
[SwitchA-dhcp-pool-0] option 43 hex 80 0B 00 00 02 01 02 03 04 02 02 02 02
```

(3) 验证配置结果

配置完成后，Switch B 可以从 DHCP 服务器 Switch A 获取到 10.1.1.0/24 网段的 IP 地址和 PXE 引导服务器地址。通过 **display dhcp server ip-in-use** 命令可以查看 DHCP 服务器为客户端分配的 IP 地址。

2.14 DHCP服务器常见配置错误举例

1. 故障现象

客户端从 DHCP 服务器动态获得的 IP 地址与其他主机 IP 地址冲突。

2. 故障分析

可能是网络上有主机私自配置了 IP 地址，导致冲突。

3. 处理过程

(1) 禁用客户端的网卡或断开其网线，从另外一台主机执行 ping 操作，检查网络中是否已经存在该 IP 地址的主机。

(2) 如果能够收到 ping 操作的响应消息，则说明该 IP 地址已由用户静态配置。在 DHCP 服务器上执行 **dhcp server forbidden-ip** 命令，禁止该 IP 地址参与动态地址分配。

(3) 重新启用客户端的网卡或连接好其网线，在客户端释放并重新获取 IP 地址。以 Windows XP 为例，在 Windows 环境下，选择[开始]/[运行]，在“运行”对话框中输入 cmd，点击[确认]按钮，进入命令行界面，使用 **ipconfig/release** 命令释放 IP 地址，之后使用 **ipconfig/renew** 重新获取 IP 地址。

3 DHCP中继配置

说明

DHCP 中继中对于接口的相关配置，目前只能在三层以太网端口、三层聚合接口和 VLAN 接口上进行。

3.1 DHCP中继简介

3.1.1 DHCP中继的应用环境

由于在 IP 地址动态获取过程中采用广播方式发送请求报文，因此 DHCP 只适用于 DHCP 客户端和服务器处于同一个子网内的情况。为进行动态主机配置，需要在所有网段上都设置一个 DHCP 服务器，这显然是很不经济的。

DHCP 中继功能的引入解决了这一难题：客户端可以通过 DHCP 中继与其他网段的 DHCP 服务器通信，最终获取到 IP 地址。这样，多个网络上的 DHCP 客户端可以使用同一个 DHCP 服务器，既节省了成本，又便于进行集中管理。

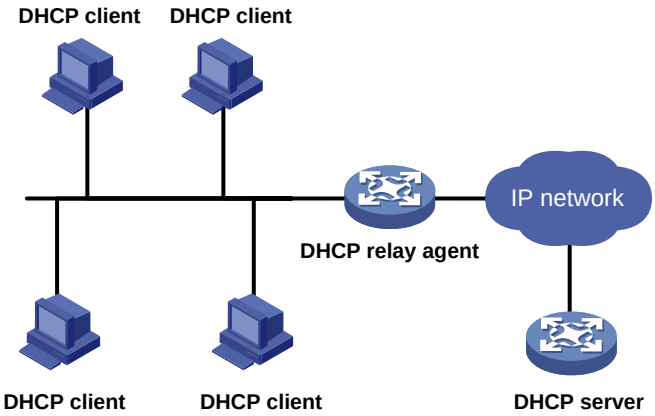
说明

设备作为 MCE（Multi-VPN-instance Customer Edge，多实例用户网络边界设备）时，在设备上配置 DHCP 中继功能，不仅可以为公网上的 DHCP 服务器和 DHCP 客户端转发 DHCP 报文，还可以实现为私网内的 DHCP 服务器和 DHCP 客户端转发 DHCP 报文，但是公网和私网之间、不同私网之间的 IP 地址空间不能重叠。MCE 的详细介绍，请参见“MPLS 配置指导”中的“MPLS L3VPN”。

3.1.2 DHCP中继的基本原理

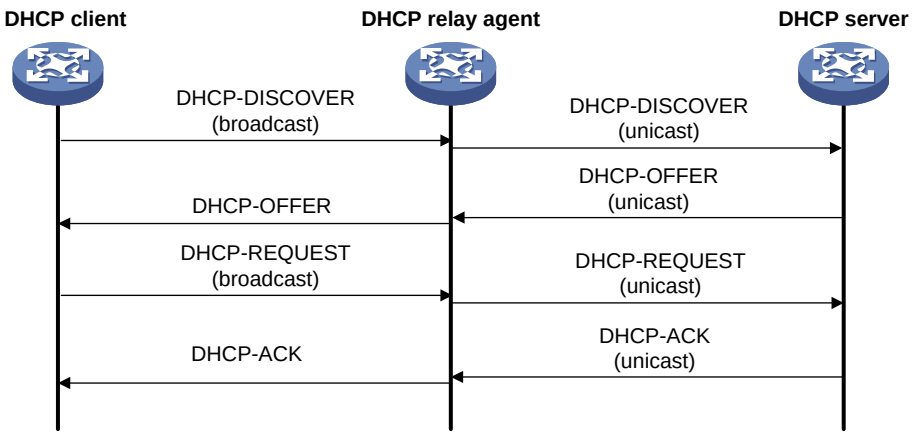
图 3-1 是DHCP中继的典型应用示意图。

图3-1 DHCP 中继的典型组网应用



通过DHCP中继完成动态配置的过程中，DHCP客户端与DHCP服务器的处理方式与不通过DHCP中继时的处理方式基本相同。下面只说明DHCP中继的转发过程，报文的具体交互过程请参见“1.2.2 IP地址动态获取过程”。

图3-2 DHCP 中继的工作过程



如 图 3-2 所示，DHCP中继的工作过程为：

(1) 具有 DHCP 中继功能的网络设备收到 DHCP 客户端以广播方式发送的 DHCP-DISCOVER 或 DHCP-REQUEST 报文后，将报文中的 giaddr 字段填充为 DHCP 中继的 IP 地址，并根据配置将报文单播转发给指定的 DHCP 服务器。

(2) DHCP 服务器根据 giaddr 字段为客户端分配 IP 地址等参数，并通过 DHCP 中继将配置信息转发给客户端，完成对客户端的动态配置。

3.1.3 DHCP中继支持Option 82 功能

Option 82 记录了DHCP客户端的位置信息。管理员可以利用该选项定位DHCP客户端，实现对客户端的安全和计费等控制。Option 82 的详细介绍请参见“1.4.3 2. 中继代理信息选项（Option 82）”。

如果DHCP中继支持Option 82 功能，则当DHCP中继接收到DHCP请求报文后，将根据报文中是否包含Option 82 以及用户配置的处理策略及填充模式等对报文进行相应的处理，并将处理后的报文转发给DHCP服务器。具体的处理方式见 表 3-1。

如果 DHCP 中继收到的应答报文中带有 Option 82，则会将 Option 82 删除后再转发给 DHCP 客户端。

表3-1 DHCP 中继支持 Option 82 的处理方式

收到 DHCP 请求报文	处理策略	填充模式	DHCP 中继对报文的处理
收到的报文中带有Option 82	Drop	任意	丢弃报文
	Keep	任意	保持报文中的Option 82不变并进行转发
	Replace	normal	采用normal模式填充Option 82，替换报文中原有的Option 82并进行转发
		verbose	采用verbose模式填充Option 82，替换报文中原有的Option 82并进行转发
		用户自定义	采用用户自定义的内容填充Option 82，替换报文中原有的Option 82并进行转发
收到的报文中不带有Option 82	-	normal	采用normal模式填充Option 82并进行转发
		verbose	采用verbose模式填充Option 82并进行转发
		用户自定义	采用用户自定义的内容填充Option 82并进行转发

3.2 DHCP中继配置任务简介

表3-2 DHCP 中继配置任务简介

配置任务	说明	详细配置
使能DHCP服务	必选	3.3.1
配置接口工作在DHCP中继模式	必选	3.3.2
配置DHCP服务器组	必选	3.3.3
配置DHCP中继的安全功能	可选	3.3.4
配置DHCP中继的用户下线检测功能	可选	3.3.5
配置通过DHCP中继释放客户端的IP地址	可选	3.3.6
配置DHCP中继支持Option 82功能	可选	3.3.7
配置DHCP报文的DSCP优先级	可选	3.3.8

3.3 配置DHCP中继

3.3.1 使能DHCP服务

只有使能 DHCP 服务后，其它相关的 DHCP 中继配置才能生效。

表3-3 使能 DHCP 服务

操作	命令	说明
进入系统视图	**system-view**	-
使能DHCP服务	**dhcp enable**	必选 缺省情况下，DHCP服务处于禁止状态

3.3.2 配置接口工作在DHCP中继模式

配置接口工作在中继模式后，当接口收到 DHCP 客户端发来的 DHCP 报文时，会将报文转发给 DHCP 服务器，由服务器分配地址。

表3-4 配置接口工作在 DHCP 中继模式

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*	-
配置接口工作在DHCP中继模式	**dhcp select relay**	必选 缺省情况下，使能DHCP服务后，接口工作在DHCP服务器模式

说明

DHCP 客户端通过 DHCP 中继获取 IP 地址时，DHCP 服务器上需要配置与 DHCP 中继连接 DHCP 客户端的接口的 IP 地址所在网段（网络号和掩码）完全相同的地址池，否则会导致 DHCP 客户端无法获得正确的 IP 地址。

3.3.3 配置DHCP服务器组

为了提高可靠性，可以在一个网络中设置多个 DHCP 服务器。多个 DHCP 服务器构成一个 DHCP 服务器组。当接口与 DHCP 服务器组建立归属关系后，会将客户端发来的 DHCP 报文转发给服务器组中的所有服务器。

表3-5 配置 DHCP 服务器组

操作	命令	说明
进入系统视图	**system-view**	-
配置DHCP服务器组中DHCP服务器的IP地址	**dhcp relay server-group** *group-id* **ip** *ip-address*	必选 缺省情况下，没有配置DHCP服务器组中服务器的IP地址
进入接口视图	**interface** *interface-type interface-number*	-

操作	命令	说明
配置接口与DHCP服务器组的归属关系	**dhcp relay server-select** *group-id*	必选 缺省情况下，接口没有与任何一个DHCP服务器组建立归属关系

说明

- 设备上最多可以配置 20 个 DHCP 服务器组。
- 通过重复执行 **dhcp relay server-group** 命令，可以在一个 DHCP 服务器组中配置多个 DHCP 服务器的 IP 地址。每个 DHCP 服务器组中最多可以配置 8 个 DHCP 服务器地址。
- DHCP 服务器组中服务器的 IP 地址不能与 DHCP 中继连接客户端的接口 IP 地址在同一网段。否则，可能导致客户端无法获得 IP 地址。
- 每个 DHCP 服务器组可以对应多个接口。但每个接口只能对应一个 DHCP 服务器组。在同一接口下多次执行 **dhcp relay server-select** 命令，新的配置会覆盖已有配置。但是，如果新指定的 DHCP 服务器组不存在时，新的归属关系配置不成功，接口还是维持与上一次配置的 DHCP 服务器组的归属关系。
- **dhcp relay server-select** 命令中所指定的组号，需事先通过 **dhcp relay server-group** 命令进行配置。

3.3.4 配置DHCP中继的安全功能

1. DHCP中继的地址匹配检查功能

为了防止非法主机静态配置一个 IP 地址并访问外部网络，设备支持 DHCP 中继的地址匹配检查功能。

接口上使能该功能后，当客户端通过 DHCP 中继从 DHCP 服务器获取到 IP 地址时，DHCP 中继可以自动记录客户端 IP 地址与 MAC 地址的绑定关系，生成 DHCP 中继的动态用户地址表项。同时，为满足用户采用合法固定 IP 地址访问外部网络的需求，DHCP 中继也支持静态配置用户地址表项，即在 DHCP 中继上手工配置 IP 地址与 MAC 地址的绑定关系。

DHCP 中继接收到主机发送的报文后，如果在用户地址表中（包括 DHCP 中继动态记录的表项以及手工配置的用户地址表项）没有与报文源 IP 地址和源 MAC 地址匹配的表项，则不学习该主机的 ARP 表项。DHCP 中继收到应答给该主机的报文后，无法将应答报文发送给该主机。从而，保证非法主机不能通过 DHCP 中继与外部网络通信。

表3-6 配置 DHCP 中继的地址匹配检查功能

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
配置DHCP中继的静态用户地址表项	**dhcp relay security static** *ip-address mac-address* [**interface** *interface-type interface-number*]	可选 缺省情况下，没有配置DHCP中继的静态用户地址表项
进入接口视图	**interface** *interface-type interface-number*	-
使能DHCP中继的地址匹配检查功能	**dhcp relay address-check enable**	必选 缺省情况，禁止DHCP中继的地址匹配检查功能

说明

- 目前，只能在三层以太网端口和 VLAN 接口上执行 **dhcp relay address-check enable** 命令。
- 在接口上使能 DHCP 中继的地址匹配检查功能之前，需要先使能 DHCP 服务、并配置该接口工作在 DHCP 中继模式，否则地址匹配检查功能不会生效。
- 执行 **dhcp relay address-check enable** 命令后将只检查 IP 和 MAC 地址，不检查接口。
- 通过 **dhcp relay security static** 命令配置静态用户地址表项时，如果静态用户地址表项与接口绑定，配置的接口必须工作在 DHCP 中继模式，否则可能引起地址表项冲突。

2. DHCP中继动态用户地址表项定时刷新功能

DHCP 客户端释放动态获取的 IP 地址时，会向 DHCP 服务器单播发送 DHCP-RELEASE 报文，DHCP 中继不会处理该报文的内容。如果此时 DHCP 中继上记录了该 IP 地址与 MAC 地址的绑定关系，则会造成 DHCP 中继的用户地址表项无法实时刷新。为了解决这个问题，DHCP 中继支持动态用户地址表项的定时刷新功能。

DHCP 中继动态用户地址表项定时刷新功能开启时，DHCP 中继每隔指定时间以客户端分配到的 IP 地址和 DHCP 中继接口的 MAC 地址向 DHCP 服务器发送 DHCP-REQUEST 报文：

- 如果 DHCP 中继接收到 DHCP 服务器响应的 DHCP-ACK 报文或在指定时间内没有接收到 DHCP 服务器的响应报文，则表明这个 IP 地址已经可以进行分配，DHCP 中继会将动态用户地址表中对应的表项老化掉。为了避免地址浪费，DHCP 中继收到 DHCP-ACK 报文后，会发送 DHCP-RELEASE 报文释放申请到的 IP 地址。
- 如果 DHCP 中继接收到 DHCP 服务器响应的 DHCP-NAK 报文，则表示该 IP 地址的租约仍然存在，DHCP 中继不会老化该 IP 地址对应的表项。

表3-7 配置 DHCP 中继动态用户地址表项定时刷新周期

操作	命令	说明
进入系统视图	**system-view**	-
开启DHCP中继动态用户地址表项定时刷新功能	**dhcp relay security refresh enable**	可选 缺省情况下，DHCP中继动态用户地址表项定时刷新功能处于开启状态

操作	命令	说明
配置DHCP中继动态用户地址表项的定时刷新周期	**dhcp relay security tracker** { *interval* | **auto** }	可选 缺省情况下，定时刷新周期为**auto**，即根据表项的数目自动计算刷新时间间隔

3. 配置伪DHCP服务器检测功能

如果网络中有私自架设的 DHCP 服务器，当客户端申请 IP 地址时，这台 DHCP 服务器就会与 DHCP 客户端进行交互，导致客户端获得错误的 IP 地址，这种私设的 DHCP 服务器称为伪 DHCP 服务器。

使能伪 DHCP 服务器检测功能后，DHCP 中继会检查接收到的 DHCP 报文中是否携带 Option 54（Server Identifier Option，服务器标识选项）。如果携带该选项，则 DHCP 中继记录此选项中的 IP 地址，即给客户端分配 IP 地址的服务器 IP 地址，并记录接收到报文的接口信息，以便管理员及时发现并处理伪 DHCP 服务器。

表3-8 配置伪 DHCP 服务器检测功能

操作	命令	说明
进入系统视图	**system-view**	-
使能伪DHCP服务器检测功能	**dhcp relay server-detect**	必选 缺省情况下，禁止伪DHCP服务器检测功能

说明

- 使能伪 DHCP 服务器检测功能后，对所有 DHCP 服务器都会进行记录，包括合法的 DHCP 服务器，管理员需要从日志信息中查找伪 DHCP 服务器。
- 使能伪 DHCP 服务器检测功能后，对每个 DHCP 服务器只记录一次。

4. 配置防止DHCP饿死攻击

DHCP 饿死攻击是指攻击者伪造 chaddr 字段各不相同的 DHCP 请求报文，向 DHCP 服务器申请大量的 IP 地址，导致 DHCP 服务器地址池中的地址耗尽，无法为合法的 DHCP 客户端分配 IP 地址，或导致 DHCP 服务器消耗过多的系统资源，无法处理正常业务。

如果封装 DHCP 请求报文的数据帧的源 MAC 地址各不相同，则限制三层接口上可以学习到的 ARP 表项数，或限制二层端口上可以学习到的 MAC 地址数，并配置学习到的 MAC 地址数达到最大值时，丢弃源 MAC 地址不在 MAC 地址表里的报文，能够避免攻击者申请过多的 IP 地址，在一定程度上缓解 DHCP 饿死攻击。

如果封装 DHCP 请求报文的数据帧的 MAC 地址都相同，则通过上述方法无法防止 DHCP 饿死攻击。在这种情况下，需要使能 DHCP 中继的 MAC 地址检查功能。使能该功能后，DHCP 中继检查接收到的 DHCP 请求报文中的 chaddr 字段和数据帧的源 MAC 地址字段是否一致。如果一致，则认为该报文合法，将其转发给 DHCP 服务器；如果不一致，则丢弃该报文。

表3-9 使能 DHCP 中继的 MAC 地址检查功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能DHCP中继的MAC地址检查功能	**dhcp relay check mac-address**	必选 缺省情况下，DHCP中继的MAC地址检查功能处于关闭状态

说明

由于 DHCP 中继转发 DHCP 报文时会修改报文的源 MAC 地址，所以只能在靠近 DHCP 客户端的第一跳 DHCP 中继设备上使能 MAC 地址检查功能。在非第一跳 DHCP 中继设备上使能 MAC 地址检查功能，会使 DHCP 中继设备错误的丢弃报文，导致客户端地址申请不成功。

3.3.5 配置DHCP中继的用户下线检测功能

DHCP 中继的用户下线检测功能以 ARP 表项老化功能为基础，ARP 表项老化时认为该表项对应的用户已经下线。

如果接口上使能了 DHCP 中继用户下线检测功能，则 ARP 表项老化时，会删除对应的用户地址表项。同时，DHCP 中继还会向 DHCP 服务器发送 DHCP-RELEASE 报文，释放下线用户的 IP 地址租约。

表3-10 配置 DHCP 中继的用户下线检测功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能DHCP中继的用户下线检测功能	**dhcp relay client-detect enable**	必选 缺省情况下，DHCP中继的用户下线检测功能处于关闭状态

说明

手工删除 ARP 表项，不会触发删除对应的用户地址表项。该 ARP 表项对应的用户下线时，需要通过 **undo dhcp relay security** 命令手工删除对应的用户地址表项。

3.3.6 配置通过DHCP中继释放客户端的IP地址

在某些情况下，可能需要通过 DHCP 中继手工释放客户端申请到的 IP 地址。如果 DHCP 中继上存在客户端 IP 地址对应的动态用户地址表项，则配置通过 DHCP 中继释放该客户端 IP 地址后，DHCP 中继会主动向 DHCP 服务器发送 DHCP-RELEASE 报文。DHCP 服务器收到该报文后，将会释放指定 IP 地址的租约。DHCP 中继也会删除该动态用户地址表项。

表3-11 配置通过 DHCP 中继释放客户端的 IP 地址

操作	命令	说明
进入系统视图	**system-view**	-
向DHCP服务器请求释放客户端申请到的IP地址	**dhcp relay release ip** *client-ip*	必选

说明

- 释放的客户端 IP 地址必须是动态用户地址表项中存在的 IP 地址，否则 DHCP 中继无法释放该 IP 地址。
- 只有使能 DHCP 中继的地址匹配检查功能或 IP Source Guard 功能后，才会自动生成 DHCP 中继的动态用户地址表项。有关 IP Source Guard 的详细介绍，请参见“安全配置指导”中的“IP Source Guard”。

3.3.7 配置DHCP中继支持Option 82 功能

1. 配置准备

在配置 DHCP 中继支持 Option82 功能之前，需完成 DHCP 中继的必配任务，即：

- 使能 DHCP 服务
- 配置接口工作在 DHCP 中继模式
- 配置 DHCP 服务器组

2. 配置DHCP中继支持Option 82 功能

表3-12 配置 DHCP 中继支持 Option 82 功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能DHCP中继支持Option 82功能	**dhcp relay information enable**	必选 缺省情况下，禁止DHCP中继支持Option 82功能
配置DHCP中继对包含Option 82的请求报文的处理策略	**dhcp relay information strategy** { **drop** | **keep** | **replace** }	可选 缺省情况下，处理策略为**replace**

操作		命令	说明
配置非用户自定义的Option 82	配置Option 82的填充模式	**dhcp relay information format** { **normal** | **verbose** [**node-identifier** { **mac** | **sysname** | **user-defined** *node-identifier* }] }	可选 缺省情况下，Option 82的填充模式为**normal**
	配置Circuit ID子选项的填充格式	**dhcp relay information circuit-id format-type** { **ascii** | **hex** }	可选 缺省情况下，Circuit ID子选项的填充格式由Option 82的填充模式决定，每个字段的填充格式不同 配置的填充格式只对非用户自定义的填充内容有效
	配置Remote ID子选项的填充格式	**dhcp relay information remote-id format-type** { **ascii** | **hex** }	可选 缺省情况下，采用HEX格式填充Remote ID子选项 配置的填充格式只对非用户自定义的填充内容有效
配置用户自定义的Option 82	配置Circuit ID子选项的内容	**dhcp relay information circuit-id string** *circuit-id*	可选 缺省情况下，Circuit ID子选项的内容由Option 82的填充模式决定
	配置Remote ID子选项的内容	**dhcp relay information remote-id string** { *remote-id* | **sysname** }	可选 缺省情况下，Remote ID子选项的内容由Option 82的填充模式决定

说明

- 为使Option 82 功能正常使用，需要在DHCP服务器和DHCP中继上都进行相应配置。DHCP服务器的相关配置请参见“2.9 配置Option 82 的处理方式”。
- DHCP 中继对包含 Option 82 请求报文的处理策略为 **replace** 时，需要配置 Option 82 的填充格式；处理策略为 **keep** 或 **drop** 时，不需要配置 Option 82 的填充格式。
- 如果以节点的设备名称（**sysname**）作为节点标识填充 DHCP 报文的 Option 82，则设备名称中不能包含空格；否则，DHCP 中继将丢弃该报文。

3.3.8 配置DHCP中继发送的DHCP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 DHCP 中继发送的 DHCP 报文的 DSCP 优先级进行配置。

表3-13 配置 DHCP 中继发送的 DHCP 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
配置DHCP中继发送的DHCP报文的DSCP优先级	**dhcp dscp** *dscp-value*	可选 缺省情况下，DHCP中继发送的DHCP报文的DSCP优先级为56

3.4 DHCP中继显示和维护

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 DHCP 中继的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令清除 DHCP 中继的统计信息。

表3-14 DHCP 中继显示和维护

操作	命令
显示接口对应的DHCP服务器组的信息	**display dhcp relay** { **all** | **interface** *interface-type interface-number* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP中继上Option 82的配置信息	**display dhcp relay information** { **all** | **interface** *interface-type interface-number* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP中继的用户地址表项信息	**display dhcp relay security** [*ip-address* | **dynamic** | **static**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP中继用户地址表项的统计信息	**display dhcp relay security statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP中继动态用户地址表项的定时刷新周期	**display dhcp relay security tracker** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP服务器组中服务器的IP地址。	**display dhcp relay server-group** { *group-id* | **all** } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP中继的相关报文统计信息	**display dhcp relay statistics** [**server-group** { *group-id* | **all** }] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除DHCP中继的统计信息	**reset dhcp relay statistics** [**server-group** *group-id*]

3.5 DHCP中继典型配置举例

3.5.1 DHCP中继配置举例

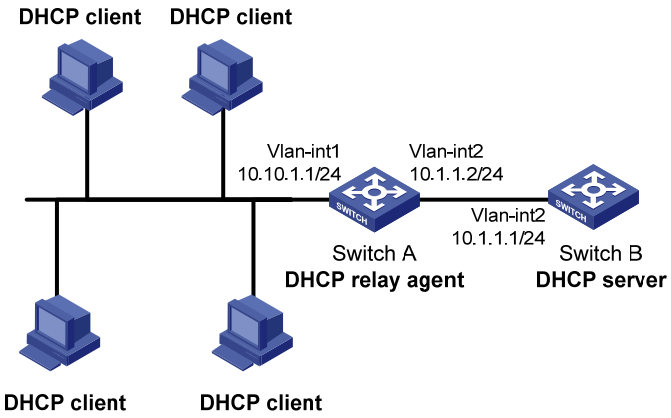
1. 组网需求

- DHCP 客户端所在网段为 10.10.1.0/24，DHCP 服务器的 IP 地址为 10.1.1.1/24；
- 由于 DHCP 客户端和 DHCP 服务器不在同一网段，因此，需要在客户端所在网段设置 DHCP 中继设备，以便客户端可以从 DHCP 服务器申请到 10.10.1.0/24 网段的 IP 地址及相关配置信息；

- Switch A 作为 DHCP 中继通过端口（属于 VLAN1）连接到 DHCP 客户端所在的网络，交换机 VLAN 接口 1 的 IP 地址为 10.10.1.1/24，VLAN 接口 2 的 IP 地址为 10.1.1.2/24。

2. 组网图

图3-3 DHCP 中继组网示意图



3. 配置步骤

配置各接口的 IP 地址（略）。

使能 DHCP 服务。

```
<SwitchA> system-view
[SwitchA] dhcp enable
```

配置 DHCP 服务器的地址。

```
[SwitchA] dhcp relay server-group 1 ip 10.1.1.1
```

配置 VLAN 接口 1 工作在 DHCP 中继模式。

```
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] dhcp select relay
```

配置 VLAN 接口 1 对应 DHCP 服务器组 1。

```
[SwitchA-Vlan-interface1] dhcp relay server-select 1
```

配置完成后，DHCP 客户端可以通过 DHCP 中继从 DHCP 服务器获取 IP 地址及相关配置信息。通过 **display dhcp relay statistics** 命令可以显示 DHCP 中继转发的 DHCP 报文统计信息；如果在 DHCP 中继上通过 **dhcp relay address-check enable** 命令使能了 DHCP 中继的地址匹配检查功能，则可以通过 **display dhcp relay security** 命令可以显示通过 DHCP 中继获取 IP 地址的客户端信息。

说明

- 由于 DHCP 中继连接客户端的接口 IP 地址与 DHCP 服务器的 IP 地址不在同一网段，因此需要在 DHCP 服务器上通过静态路由或动态路由协议保证两者之间路由可达。
- 为了使DHCP客户端能从DHCP服务器获得IP地址，还需要在DHCP服务器上进行一些配置。DHCP服务器的配置方法，请参见“2.13 DHCP服务器典型配置举例”。

3.5.2 DHCP中继支持Option 82 配置举例

1. 组网需求

- 在 DHCP 中继 Switch A 上使能 Option 82 功能；
- 对包含 Option 82 的请求报文的处理策略为 **replace**；
- Ciruict ID 填充内容为 company001，Remote ID 填充内容为 device001；
- Switch A 将添加 Option 82 的 DHCP 请求报文转发给 DHCP 服务器 Switch B，使得 DHCP 客户端可以获取到 IP 地址。

2. 组网图

如 图 3-3 所示。

3. 配置步骤

配置各接口的 IP 地址（略）。

使能 DHCP 服务。

```
<SwitchA> system-view
[SwitchA] dhcp enable
```

配置 DHCP 服务器的地址。

```
[SwitchA] dhcp relay server-group 1 ip 10.1.1.1
```

配置 VLAN 接口 1 工作在 DHCP 中继模式。

```
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] dhcp select relay
```

配置 VLAN 接口 1 对应 DHCP 服务器组 1。

```
[SwitchA-Vlan-interface1] dhcp relay server-select 1
```

配置 Option 82 的处理策略和填充内容。

```
[SwitchA-Vlan-interface1] dhcp relay information enable
[SwitchA-Vlan-interface1] dhcp relay information strategy replace
[SwitchA-Vlan-interface1] dhcp relay information circuit-id string company001
[SwitchA-Vlan-interface1] dhcp relay information remote-id string device001
```

说明

为使 Option 82 功能正常使用，DHCP 服务器也需要进行相应配置。

3.6 DHCP中继常见配置错误举例

1. 故障现象

客户端不能通过 DHCP 中继获得配置信息。

2. 故障分析

DHCP 中继或 DHCP 服务器的配置可能有问题。可以打开调试开关显示调试信息，并通过执行 **display** 命令显示接口状态信息的方法来分析定位。

3. 处理过程

- 检查 DHCP 服务器和 DHCP 中继是否使能了 DHCP 服务。
- 检查 DHCP 服务器是否配置有 DHCP 客户端所在网段的地址池。
- 检查具有 DHCP 中继功能的网络设备和 DHCP 服务器是否配置有相互可达的路由。
- 检查具有 DHCP 中继功能的网络设备是否在连接 DHCP 客户端所在网段的接口上配置有正确的 DHCP 服务器组，且 DHCP 服务器组的 IP 地址配置正确。

4 DHCP客户端配置

说明

- DHCP 客户端中对于接口的相关配置，目前只能在三层以太网端口、三层聚合接口和 VLAN 接口上进行。
- 多个具有相同 MAC 地址的 VLAN 接口通过中继以 DHCP 方式申请 IP 地址时，不能用 Windows 2000 Server 和 Windows 2003 Server 作为 DHCP 服务器。
- 加入聚合组的端口不能配置为 DHCP 客户端。

4.1 DHCP客户端简介

指定设备的接口作为 DHCP 客户端后，可以使用 DHCP 协议从 DHCP 服务器动态获得 IP 地址等参数，方便用户配置，也便于集中管理。

4.2 配置接口使用DHCP方式获取IP地址

表4-1 配置接口使用 DHCP 方式获取 IP 地址

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置接口使用DHCP方式获取IP地址	**ip address dhcp-alloc** [**client-identifier mac** *interface-type interface-number*]	必选 • 空配置启动时，使用软件功能缺省值，接口不使用 DHCP 方式获取 IP 地址 • 缺省配置启动时，使用软件功能出厂值，指定接口的 MAC 地址作为客户端 ID 来获取 IP 地址 关于空配置启动和缺省配置启动，请参见“基础配置指导”中的“配置文件管理”

说明：

- 接口可以采用多种方式获得 IP 地址，新的配置方式会覆盖原有的配置方式。
- 当接口被配置为通过 DHCP 动态获取 IP 地址后，不能再给该接口配置从 IP 地址。
- 如果 DHCP 服务器为接口分配的 IP 地址与设备上其他接口的 IP 地址在同一网段，则该接口不会使用该 IP 地址，且不再向 DHCP 服务器申请 IP 地址，除非手动删除冲突接口的 IP 地址，并重新使能接口（先后执行 **shutdown** 和 **undo shutdown** 命令）或重新配置接口使用 DHCP 方式获取 IP 地址（先后执行 **undo ip address dhcp-alloc** 和 **ip address dhcp-alloc** 命令）。

4.3 配置DHCP客户端发送的DHCP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 DHCP 客户端发送的 DHCP 报文的 DSCP 优先级进行配置。

表4-2 配置 DHCP 客户端发送的 DHCP 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
配置DHCP客户端发送的DHCP报文的DSCP优先级	**dhcp client dscp** *dscp-value*	可选 缺省情况下，DHCP客户端发送的DHCP报文的DSCP优先级为56

4.4 DHCP客户端显示和维护

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 DHCP 客户端的信息，通过查看显示信息验证配置的效果。

表4-3 DHCP 客户端显示和维护

操作	命令
显示DHCP客户端的相关信息	**display dhcp client** [**verbose**] [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

4.5 DHCP客户端典型配置举例

1. 组网需求

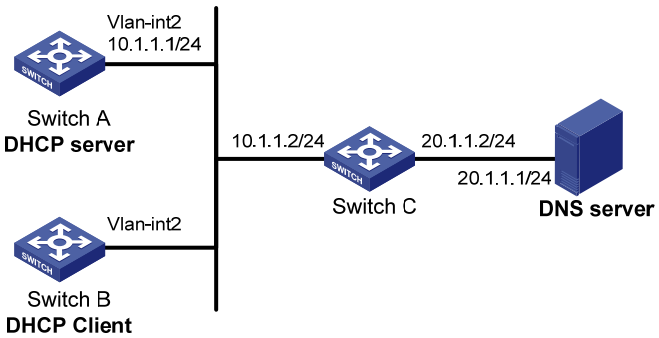
Switch B 的端口（属于 VLAN2）接入局域网，VLAN 接口 2 通过 DHCP 协议从 DHCP 服务器获取 IP 地址、DNS 服务器地址和静态路由信息：

- DHCP 客户端的 IP 地址所在网段为 10.1.1.0/24；
- DNS 服务器地址为 20.1.1.1；
- 静态路由信息为到达 20.1.1.0/24 网段的下一跳地址是 10.1.1.2。

DHCP 服务器需要通过自定义选项的方式配置 Option 121 的内容，以便为客户端分配静态路由信息。其中，目的描述符由子网掩码长度和目的网络地址两部分组成。在本例中，目的描述符字段取值为 18 14 01 01（十六进制数值，表示子网掩码长度为 24，目的网络地址为 20.1.1.0）；下一跳地址字段取值为 0A 01 01 02（十六进制数值，表示下一跳地址为 10.1.1.2）。

2. 组网图

图4-1 DHCP 客户端配置举例组网图



3. 配置步骤

(1) 配置 DHCP 服务器 Switch A

配置接口的 IP 地址。

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ip address 10.1.1.1 24
[SwitchA-Vlan-interface2] quit
```

使能 DHCP 服务。

```
[SwitchA] dhcp enable
```

配置不参与自动分配的 IP 地址。

```
[SwitchA] dhcp server forbidden-ip 10.1.1.2
```

配置 DHCP 地址池 0，采用动态绑定方式分配 IP 地址。可分配的网段为 10.1.1.0/24，租约有效期限为 10 天，DNS 服务器地址为 20.1.1.1，到达 20.1.1.0/24 网段的下一跳地址是 10.1.1.2。

```
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] network 10.1.1.0 mask 255.255.255.0
[SwitchA-dhcp-pool-0] expired day 10
[SwitchA-dhcp-pool-0] dns-list 20.1.1.1
[SwitchA-dhcp-pool-0] option 121 hex 18 14 01 01 0A 01 01 02
```

(2) 配置 DHCP 客户端 Switch B

配置 VLAN 接口 2 通过 DHCP 动态获取地址。

```
<SwitchB> system-view
```

```
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ip address dhcp-alloc
```

(3) 验证配置结果

通过 **display dhcp client** 命令可以查看 Switch B 申请到的 IP 地址和网络配置参数。

```
[SwitchB-Vlan-interface2] display dhcp client verbose
Vlan-interface2 DHCP client information:
 Current machine state: BOUND
 Allocated IP: 10.1.1.3 255.255.255.0
 Allocated lease: 864000 seconds, T1: 432000 seconds, T2: 756000 seconds
 Lease from 2009.02.20 11:06:35   to   2009.03.02 11:06:35
 DHCP server: 10.1.1.1
 Transaction ID: 0x410090f0
 Classless static route:
   Destination: 20.1.1.0, Mask: 255.255.255.0, NextHop: 10.1.1.2
 DNS server: 20.1.1.1
 Client ID: 3030-3066-2e65-3230-
            302e-3030-3032-2d45-
            7468-6572-6e65-7430-
            2f30
 T1 will timeout in 4 days 23 hours 59 minutes 50 seconds.
```

通过 **display ip routing-table** 命令可以查看 Switch B 的路由表中添加了到达 20.1.1.0/24 网络的静态路由。

```
[SwitchB-Vlan-interface2] display ip routing-table
Routing Tables: Public
         Destinations : 5        Routes : 5

Destination/Mask    Proto  Pre  Cost         NextHop         Interface

10.1.1.0/24         Direct 0    0            10.1.1.3        Vlan2
10.1.1.3/32         Direct 0    0            127.0.0.1       InLoop0
20.1.1.0/24         Static 70   0            10.1.1.2        Vlan2
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 DHCP Snooping配置

说明

设备只有位于 DHCP 客户端与 DHCP 服务器之间，或 DHCP 客户端与 DHCP 中继之间时，DHCP Snooping 功能配置后才能正常工作;设备位于 DHCP 服务器与 DHCP 中继之间时，DHCP Snooping 功能配置后不能正常工作。

5.1 DHCP Snooping简介

5.1.1 DHCP Snooping作用

DHCP Snooping 是 DHCP 的一种安全特性，具有如下功能：

(1) 保证客户端从合法的服务器获取 IP 地址。

(2) 记录 DHCP 客户端 IP 地址与 MAC 地址的对应关系。

1. 保证客户端从合法的服务器获取IP地址

网络中如果存在私自架设的伪 DHCP 服务器，则可能导致 DHCP 客户端获取错误的 IP 地址和网络配置参数，无法正常通信。为了使 DHCP 客户端能通过合法的 DHCP 服务器获取 IP 地址，DHCP Snooping 安全机制允许将端口设置为信任端口和不信任端口：

- 信任端口正常转发接收到的 DHCP 报文。
- 不信任端口接收到 DHCP 服务器响应的 DHCP-ACK 和 DHCP-OFFER 报文后，丢弃该报文。

连接 DHCP 服务器和其他 DHCP Snooping 设备的端口需要设置为信任端口，其他端口设置为不信任端口，从而保证 DHCP 客户端只能从合法的 DHCP 服务器获取 IP 地址，私自架设的伪 DHCP 服务器无法为 DHCP 客户端分配 IP 地址。

2. 记录DHCP客户端IP地址与MAC地址的对应关系

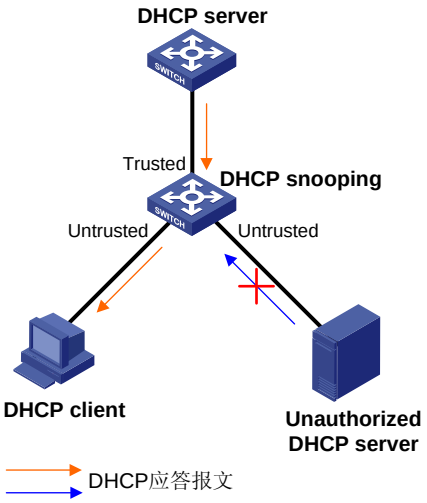
DHCP Snooping 通过监听 DHCP-REQUEST 报文和信任端口收到的 DHCP-ACK 报文，记录 DHCP Snooping 表项，其中包括客户端的 MAC 地址、获取到的 IP 地址、与 DHCP 客户端连接的端口及该端口所属的 VLAN 等信息。利用这些信息可以实现：

- ARP Detection：根据 DHCP Snooping 表项来判断发送 ARP 报文的用户是否合法，从而防止非法用户的 ARP 攻击。ARP Detection 的详细介绍请参见“安全配置指导”中的“ARP 攻击防御”。
- IP Source Guard：通过动态获取 DHCP Snooping 表项对端口转发的报文进行过滤，防止非法报文通过该端口。IP Source Guard 的详细介绍请参见“安全配置指导”中的“IP Source Guard”。
- VLAN 映射：发送给用户的报文通过查找映射 VLAN（SVLAN）对应的 DHCP Snooping 表项中的 DHCP 客户端 IP 地址、MAC 地址和原始 VLAN（CVLAN）的信息，将报文的 SVLAN 修改为原来的 CVLAN。VLAN 映射的详细介绍请参见“二层技术-以太网交换配置指导”中的“VLAN 映射”。

5.1.2 信任端口的典型应用环境

1. 连接DHCP服务器

图5-1 信任端口和非信任端口



如 图 5-1 所示，连接DHCP服务器的端口需要配置为信任端口，以便DHCP Snooping设备正常转发DHCP服务器的应答报文，保证DHCP客户端能够从合法的DHCP服务器获取IP地址。

2. DHCP Snooping级联网络

在多个 DHCP Snooping 设备级联的网络中，为了节省系统资源，不需要每台 DHCP Snooping 设备都记录所有 DHCP 客户端的 IP 地址和 MAC 地址绑定，只需在与客户端直接相连的 DHCP Snooping 设备上记录绑定信息。通过将间接与 DHCP 客户端相连的端口配置为不记录 IP 地址和 MAC 地址绑定的信任端口，可以实现该功能。如果 DHCP 客户端发送的请求报文从此类信任端口到达 DHCP Snooping 设备，DHCP Snooping 设备不会记录客户端 IP 地址和 MAC 地址的绑定。

图5-2 DHCP Snooping 级联组网图

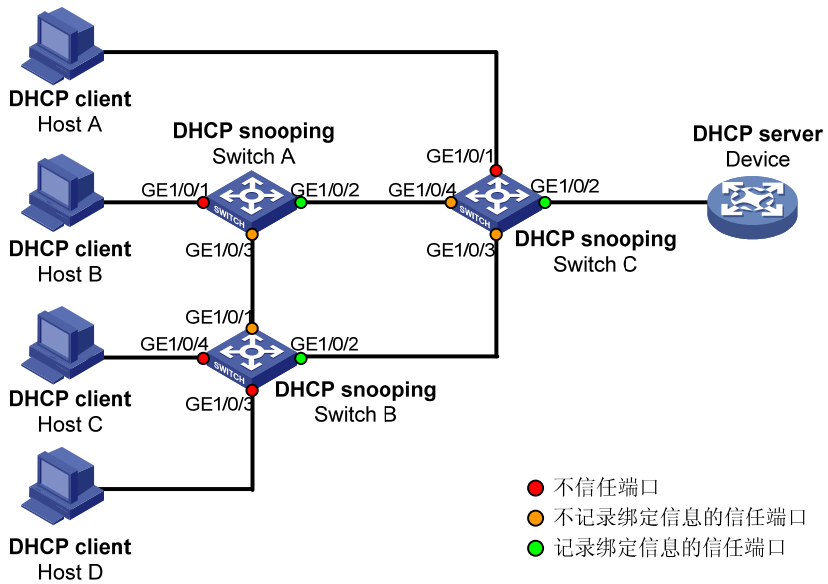


图 5-2 中设备各端口的角色如 表 5-1 所示。

表5-1 端口的角色

设备	不信任端口	不记录绑定信息的信任端口	记录绑定信息的信任端口
Switch A	GigabitEthernet1/0/1	GigabitEthernet1/0/3	GigabitEthernet1/0/2
Switch B	GigabitEthernet1/0/3和 GigabitEthernet1/0/4	GigabitEthernet1/0/1	GigabitEthernet1/0/2
Switch C	GigabitEthernet1/0/1	GigabitEthernet1/0/3和 GigabitEthernet1/0/4	GigabitEthernet1/0/2

5.1.3 DHCP Snooping支持Option 82 功能

Option 82 记录了DHCP客户端的位置信息。管理员可以利用该选项定位DHCP客户端，实现对客户端的安全和计费等控制。Option 82 的详细介绍请参见“1.4.3 2. 中继代理信息选项（Option 82）”。

如果DHCP Snooping支持Option 82 功能，则当设备接收到DHCP请求报文后，将根据报文中是否包含Option 82 以及用户配置的处理策略及填充模式等对报文进行相应的处理，并将处理后的报文转发给DHCP服务器。具体的处理方式见 表 5-2。

当设备接收到 DHCP 服务器的响应报文时，如果报文中含有 Option 82，则删除 Option 82，并转发给 DHCP 客户端；如果报文中不含有 Option 82，则直接转发。

表5-2 DHCP Snooping 支持 Option 82 的处理方式

收到 DHCP 请求报文	处理策略	填充模式	DHCP Snooping 对报文的处理
收到的报文中带有Option 82	Drop	-	丢弃报文
	Keep	任意	保持报文中的Option 82不变并进行转发
	Replace	normal	采用normal模式填充Option 82，替换报文中原有的Option 82并进行转发
		verbose	采用verbose模式填充Option 82，替换报文中原有的Option 82并进行转发
		用户自定义	采用用户自定义的内容填充Option 82，替换报文中原有的Option 82并进行转发
	Append	normal	保持报文中的Option 82不变并进行转发
		verbose	保持报文中的Option 82不变并进行转发
		private	在报文中原有Option 82的基础上追加子选项9的内容，如果原有Option 82中携带子选项9，在原有子选项9的基础上追加内容，然后转发报文
		standard	保持报文中的Option 82不变并进行转发
		用户自定义	保持报文中的Option 82不变并进行转发
收到的报文中不带有Option 82	-	normal	采用normal模式填充Option 82并进行转发
		private	采用private填充Option 82并进行转发
		standard	采用standard模式填充Option 82并进行转发

收到 DHCP 请求报文	处理策略	填充模式	DHCP Snooping 对报文的处理
		verbose	采用verbose模式填充Option 82并进行转发
		用户自定义	采用用户自定义的内容填充Option 82并进行转发

说明

DHCP Snooping 对 Option 82 的处理策略、填充模式与 DHCP 中继相同。

5.2 DHCP Snooping配置任务简介

表5-3 DHCP Snooping 配置任务简介

配置任务	说明	详细配置
配置DHCP Snooping基本功能	必选	5.3
配置DHCP Snooping支持Option 82功能	可选	5.4
配置DHCP Snooping表项备份功能	可选	5.5
配置防止DHCP饿死攻击	可选	5.6
配置防止伪造DHCP续约报文攻击	可选	5.7
配置DHCP Snooping报文限速功能	可选	5.8

5.3 配置DHCP Snooping基本功能

表5-4 配置 DHCP Snooping 基本功能

操作	命令	说明
进入系统视图	**system-view**	-
使能DHCP Snooping功能	**dhcp-snooping**	必选 缺省情况下，DHCP Snooping功能处于关闭状态
进入接口视图	**interface** *interface-type interface-number*	- 此接口为连接DHCP服务器的接口
配置端口为信任端口，并记录客户端IP地址和MAC地址的绑定关系	**dhcp-snooping trust**	必选 缺省情况下，在使能DHCP Snooping功能后，设备的所有端口均为不信任端口
退回系统视图	**quit**	-

操作	命令	说明
进入接口视图	**interface** *interface-type interface-number*	- 此接口为间接与DHCP客户端相连的接口
配置端口为不记录IP地址和MAC地址绑定的信任端口	**dhcp-snooping trust no-user-binding**	可选 缺省情况下，在使能DHCP Snooping功能后，设备的所有端口均为不信任端口

说明

- 为了使 DHCP 客户端能从合法的 DHCP 服务器获取 IP 地址，必须将与合法 DHCP 服务器相连的端口设置为信任端口,设置的信任端口和与 DHCP 客户端相连的端口必须在同一个 VLAN 内。
- 目前,可以配置为 DHCP Snooping 信任端口的接口类型包括: 二层以太网端口、二层聚合接口。关于聚合接口的详细介绍，请参见“二层技术-以太网交换配置指导”中的“以太网链路聚合”。
- 如果二层以太网端口加入聚合组，则在该接口上进行的 DHCP Snooping 相关配置不会生效；该接口退出聚合组后，之前的配置才会生效。
- DHCP Snooping 功能可以与基本 QinQ、灵活 QinQ 功能同时使用：接收到 DHCP 客户端发送给 DHCP 服务器的报文后，如果报文不带有 VLAN Tag，则为其添加一层 VLAN Tag；如果报文只带一层 VLAN Tag，则在报文前面再添加一层 VLAN Tag，(添加的 VLAN Tag 为第一层 VLAN Tag，原来的 VLAN Tag 为第二层 VLAN Tag)，并通过 DHCP Snooping 表项记录两层 VLAN Tag；如果报文带有两层 VLAN Tag，则不添加 VLAN Tag，直接转发给 DHCP 服务器。
- 在添加 VLAN Tag 的同时，还需要修改原有的 VLAN Tag 时，DHCP Snooping 功能不能与灵活 QinQ 功能同时使用。

5.4 配置DHCP Snooping支持Option 82功能

表5-5 配置 DHCP Snooping 支持 Option 82 功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能DHCP Snooping支持Option 82功能	**dhcp-snooping information enable**	必选 缺省情况下，禁止DHCP Snooping支持Option 82功能
配置DHCP Snooping对包含Option 82的请求报文的处理策略	**dhcp-snooping information strategy** { **append** | **drop** | **keep** | **replace** }	可选 缺省情况下，处理策略为**replace**

操作		命令	说明
配置非用户自定义的Option 82	配置Option 82的填充模式	**dhcp-snooping information format** { **normal** | **private** *private* | **standard | verbose** [**node-identifier** { **mac** | **sysname** | **user-defined** *node-identifier* }] }	可选 缺省情况下，Option 82的填充模式为**normal**
	配置Circuit ID子选项的填充格式	**dhcp-snooping information circuit-id format-type** { **ascii** | **hex** }	可选 缺省情况下，Circuit ID子选项的填充格式由Option 82的填充模式决定，每个字段的填充格式不同 配置的填充格式只对非用户自定义的填充内容有效 Private填充模式下，仅配置为hex格式有效
	配置Remote ID子选项的填充格式	**dhcp-snooping information remote-id format-type** { **ascii** | **hex** }	可选 缺省情况下，采用HEX格式填充Remote ID子选项 配置的填充格式只对非用户自定义的填充内容有效 Private填充模式下，仅配置为hex格式有效
	配置子选项9使能	**dhcp-snooping information** [**vlan** *vlan-id*] **sub-option** *sub-option-code*	可选 缺省情况下，子选项9没有使能 配置append策略情况下，填充sysname和Loopback0接口的主IP，其它策略情况下仅填充sysname
配置用户自定义的Option 82	配置Circuit ID子选项的内容	**dhcp-snooping information** [**vlan** *vlan-id*] **circuit-id string** *circuit-id*	可选 缺省情况下，Circuit ID子选项的内容由Option 82的填充模式决定
	配置Remote ID子选项的内容	**dhcp-snooping information** [**vlan** *vlan-id*] **remote-id string** { *remote-id* | **sysname** }	可选 缺省情况下，Remote ID子选项的内容由Option 82的填充模式决定
	配置子选项9的内容	**dhcp-snooping information** [**vlan** *vlan-id*] **sub-option** *sub-option-code* [**string** *user-string*&<1-8>]	可选 缺省情况下，子选项9不填充

说明

- 只能在二层以太网端口、二层聚合接口上配置 DHCP Snooping 支持 Option 82 功能。
- 如果二层以太网端口加入聚合组，则该接口上进行的 DHCP Snooping 支持 Option 82 功能的配置不会生效；该接口退出聚合组后，之前的配置才会生效。
- 为使Option 82 功能正常使用，需要在DHCP服务器和DHCP Snooping上都进行相应配置。DHCP服务器的相关配置请参见“2.9 配置Option 82 的处理方式”。
- DHCP Snooping 对包含 Option 82 请求报文的处理策略为 **replace** 时，需要配置 Option 82 的填充格式；处理策略为 **keep** 或 **drop** 时，不需要配置 Option 82 的填充格式。
- 如果以设备名称（**sysname**）作为节点标识填充 DHCP 报文的 Option 82，则设备名称中不能包含空格；否则，DHCP Snooping 将丢弃该报文。用户可以通过 **sysname** 命令配置设备名称，该命令的详细介绍请参见“基本配置命令参考”中的“设备管理”。
- DHCP Snooping 功能和 QinQ 功能同时使用，或 DHCP Snooping 设备接收到的 DHCP 报文带有两层 VLAN Tag 时，如果采用 normal 或 verbose 模式填充 Option 82，则 sub-option 1 中 VLAN ID 字段的格式为“第一层 VLAN Tag.第二层 VLAN Tag”。例如，第一层 VLAN Tag 为 10（十六进制值为 a），第二层 VLAN Tag 为 20（十六进制值为 14），则 VLAN ID 字段的内容为“000a.0014”。

5.5 配置DHCP Snooping表项备份功能

DHCP Snooping 设备重启后，设备上记录的 DHCP Snooping 表项将丢失。如果 DHCP Snooping 与安全模块（如 IP Source Guard）配合使用，则表项丢失会导致安全模块无法通过 DHCP Snooping 获取到相应的表项，进而导致 DHCP 客户端不能顺利通过安全检查、正常访问网络。

DHCP Snooping 表项备份功能将 DHCP Snooping 表项保存到指定的文件中，DHCP Snooping 设备重启后，自动根据该文件恢复 DHCP Snooping 表项，从而保证 DHCP Snooping 表项不会丢失。

表5-6 配置 DHCP Snooping 表项备份功能

操作	命令	说明
进入系统视图	**system-view**	-
指定存储DHCP Snooping表项的文件名称	**dhcp-snooping binding database filename** *filename*	必选 缺省情况下，未指定存储文件名称 执行本命令后，会立即触发一次表项备份。之后，按照**dhcp-snooping binding database update interval**命令设置的刷新时间间隔定时更新表项文件
将当前的DHCP Snooping表项保存到用户指定的文件中	**dhcp-snooping binding database update now**	可选 本命令只用来触发一次DHCP Snooping表项的备份

操作	命令	说明
配置DHCP Snooping表项存储文件的刷新时间间隔	**dhcp-snooping binding database update interval** *minutes*	可选 缺省情况下，不会定期刷新DHCP Snooping表项存储文件

说明

执行 **undo dhcp-snooping** 命令关闭 DHCP Snooping 功能后，设备会删除所有 DHCP Snooping 表项，文件中存储的 DHCP Snooping 表项也将被删除。

5.6 配置防止DHCP饿死攻击

DHCP 饿死攻击是指攻击者伪造 chaddr 字段各不相同的 DHCP 请求报文，向 DHCP 服务器申请大量的 IP 地址，导致 DHCP 服务器地址池中的地址耗尽，无法为合法的 DHCP 客户端分配 IP 地址，或导致 DHCP 服务器消耗过多的系统资源，无法处理正常业务。

如果封装 DHCP 请求报文的数据帧的源 MAC 地址各不相同，则通过 **mac-address max-mac-count** 命令限制端口可以学习到的 MAC 地址数，并配置学习到的 MAC 地址数达到最大值时，丢弃源 MAC 地址不在 MAC 地址表里的报文，能够避免攻击者申请过多的 IP 地址，在一定程度上缓解 DHCP 饿死攻击。此时，不存在 DHCP 饿死攻击的端口下的 DHCP 客户端可以正常获取 IP 地址，但存在 DHCP 饿死攻击的端口下的 DHCP 客户端仍可能无法获取 IP 地址。

如果封装 DHCP 请求报文的数据帧的 MAC 地址都相同，则通过 **mac-address max-mac-count** 命令无法防止 DHCP 饿死攻击。在这种情况下，需要使能 DHCP Snooping 的 MAC 地址检查功能。使能该功能后，DHCP Snooping 设备检查接收到的 DHCP 请求报文中的 chaddr 字段和数据帧的源 MAC 地址字段是否一致。如果一致，则认为该报文合法，将其转发给 DHCP 服务器；如果不一致，则丢弃该报文。

表5-7 使能 DHCP Snooping 的 MAC 地址检查功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能DHCP Snooping的MAC地址检查功能	**dhcp-snooping check mac-address**	必选 缺省情况下，DHCP Snooping的MAC地址检查功能处于关闭状态

说明

只能在二层以太网端口、二层聚合接口上使能 DHCP Snooping 的 MAC 地址检查功能。

5.7 配置防止伪造DHCP续约报文攻击

伪造DHCP续约报文攻击是指攻击者冒充合法的DHCP客户端，向DHCP服务器发送伪造的DHCP续约报文，导致DHCP服务器和DHCP客户端无法按照自己的意愿及时释放IP地址租约。如果攻击者冒充不同的DHCP客户端发送大量伪造的DHCP续约报文，则会导致大量IP地址被长时间占用，DHCP服务器没有足够的地址分配给新的DHCP客户端。

在DHCP Snooping设备上使能DHCP Request报文检查功能，可以有效地防止伪造DHCP续约报文攻击。如果使能了该功能，则DHCP Snooping设备接收到DHCP Request报文后，检查本地是否存在与DHCP Request报文匹配的DHCP Snooping表项。若存在，则DHCP Request报文信息与DHCP Snooping表项信息一致时，认为该报文为合法的续约报文，将其转发给DHCP服务器；不一致时，认为该报文为伪造的续约报文，将其丢弃。若不存在，则认为该报文合法，将其转发给DHCP服务器。

表5-8 使能DHCP Snooping的DHCP Request报文检查功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能DHCP Snooping的DHCP Request报文检查功能	**dhcp-snooping check request-message**	必选 缺省情况下，DHCP Snooping的DHCP Request报文检查功能处于关闭状态

说明

只能在二层以太网端口、二层聚合接口上使能DHCP Snooping的DHCP Request报文检查功能。

5.8 配置DHCP Snooping报文限速功能

为了避免非法用户发送大量的DHCP报文，对网络造成攻击，DHCP Snooping支持报文限速功能，限制接口接收DHCP报文的速率。当接口接收的DHCP报文速率超过限制的最高速率时，DHCP Snooping设备将丢弃超过速率限制的报文。

表5-9 配置DHCP Snooping报文限速功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
使能DHCP Snooping的报文限速功能	**dhcp-snooping rate-limit** *rate*	必选 缺省情况下，关闭DHCP Snooping的报文限速功能，即不限制接口接收DHCP报文的速率

说明

- 只能在二层以太网端口和二层聚合接口上使能 DHCP Snooping 的报文限速功能。
- 如果二层以太网端口加入了聚合组，则该接口采用对应二层聚合接口下的 DHCP Snooping 报文限速配置。

5.9 DHCP Snooping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 DHCP Snooping 的配置情况，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 DHCP Snooping 的统计信息。

表5-10 DHCP Snooping 显示和维护

操作	命令
显示DHCP Snooping表项信息	**display dhcp-snooping** [**ip** *ip-address*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP Snooping上Option 82的配置信息	**display dhcp-snooping information** { **all** | **interface** *interface-type interface-number* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP Snooping设备上的DHCP报文统计信息	**display dhcp-snooping packet statistics** [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示信任端口信息	**display dhcp-snooping trust** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCP Snooping表项备份信息	**display dhcp-snooping binding database** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除DHCP Snooping表项	**reset dhcp-snooping** { **all** | **ip** *ip-address* }
清除DHCP Snooping设备上的DHCP报文统计信息	**reset dhcp-snooping packet statistics** [**slot** *slot-number*]

5.10 DHCP Snooping典型配置举例

5.10.1 DHCP Snooping配置举例

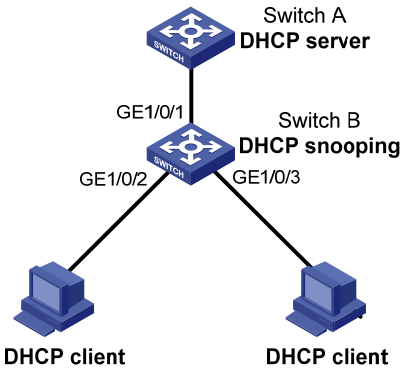
1. 组网需求

Switch B 通过以太网端口 GigabitEthernet1/0/1 连接到 DHCP 服务器，通过以太网端口 GigabitEthernet1/0/2、GigabitEthernet1/0/3 连接到 DHCP 客户端。要求：

- 与 DHCP 服务器相连的端口可以转发 DHCP 服务器的响应报文，而其他端口不转发 DHCP 服务器的响应报文。
- 记录 DHCP-REQUEST 报文和信任端口收到的 DHCP-ACK 报文中 DHCP 客户端 IP 地址及 MAC 地址的绑定关系。

2. 组网图

图5-3 DHCP Snooping 组网示意图



3. 配置步骤

使能 DHCP Snooping 功能。

```
<SwitchB> system-view
[SwitchB] dhcp-snooping
```

配置 GigabitEthernet1/0/1 端口为信任端口。

```
[SwitchB] interface GigabitEthernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] dhcp-snooping trust
[SwitchB-GigabitEthernet1/0/1] quit
```

5.10.2 DHCP Snooping支持Option 82 配置举例

1. 组网需求

- Switch B 上使能 DHCP Snooping 功能，并支持 Option 82 功能；
- 对包含 Option 82 的请求报文的处理策略为 **replace**；
- 在 GigabitEthernet1/0/2 上配置 Ciruict ID 填充内容为 company001，Remote ID 填充内容为 device001；
- 在 GigabitEthernet1/0/3 上配置以 verbose 模式填充 Option 82，接入节点标识为 **sysname**，填充格式为 ASCII 格式。
- Switch B 将添加 Option 82 的 DHCP 请求报文转发给 DHCP 服务器 Switch A，使得 DHCP 客户端可以获取到 IP 地址。

2. 组网图

如 图 5-3 所示。

3. 配置步骤

使能 DHCP Snooping 功能。

```
<SwitchB> system-view
[SwitchB] dhcp-snooping
```

配置 GigabitEthernet1/0/1 端口为信任端口。

```
[SwitchB] interface GigabitEthernet 1/0/1
```

```
[SwitchB-GigabitEthernet1/0/1] dhcp-snooping trust
[SwitchB-GigabitEthernet1/0/1] quit
```

在 GigabitEthernet1/0/2 上配置 DHCP Snooping 支持 Option 82 功能。

```
[SwitchB] interface GigabitEthernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] dhcp-snooping information enable
[SwitchB-GigabitEthernet1/0/2] dhcp-snooping information strategy replace
[SwitchB-GigabitEthernet1/0/2] dhcp-snooping information circuit-id string company001
[SwitchB-GigabitEthernet1/0/2] dhcp-snooping information remote-id string device001
[SwitchB-GigabitEthernet1/0/2] quit
```

在端口 GigabitEthernet1/0/3 上配置 DHCP Snooping 支持 Option 82 功能。

```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information enable
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information strategy replace
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information format verbose node-identifier
sysname
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information circuit-id format-type ascii
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information remote-id format-type ascii
```

6 BOOTP客户端配置

说明

- BOOTP 客户端中对于接口的相关配置，目前只能在三层以太网端口、三层聚合接口和 VLAN 接口上进行。
- 多个具有相同 MAC 地址的 VLAN 接口通过中继以 BOOTP 方式申请 IP 地址时，不能用 Windows 2000 Server 和 Windows 2003 Server 作为 BOOTP 服务器。
- 加入聚合组的接口不能配置为 BOOTP 客户端。

6.1 BOOTP客户端简介

6.1.1 BOOTP应用

BOOTP 是 Bootstrap Protocol（自举协议）的简称。指定设备的接口作为 BOOTP 客户端后，该接口可以使用 BOOTP 协议从 BOOTP 服务器获得 IP 地址等信息，从而方便用户配置。

使用 BOOTP 协议，管理员需要在 BOOTP 服务器上为每个 BOOTP 客户端配置 BOOTP 参数文件，该文件包括 BOOTP 客户端的 MAC 地址及其对应的 IP 地址等信息。当 BOOTP 客户端向 BOOTP 服务器发起请求时，服务器会查找 BOOTP 参数文件，并返回相应的配置信息。

由于需要在 BOOTP 服务器上为每个客户端事先配置参数文件，BOOTP 一般运行在相对稳定的环境中。当网络变化频繁时，可以采用 DHCP 协议。

说明

由于 DHCP 服务器可以与 BOOTP 客户端进行交互，因此用户可以不配置 BOOTP 服务器，而使用 DHCP 服务器为 BOOTP 客户端分配 IP 地址。

6.1.2 IP地址动态获取过程

说明

在下面的 IP 地址动态获取过程中，BOOTP 服务器的功能可以用 DHCP 服务器替代。

BOOTP 客户端从 BOOTP 服务器动态获取 IP 地址的具体过程如下：

(1) BOOTP 客户端以广播方式发送 BOOTP 请求报文，其中包含了 BOOTP 客户端的 MAC 地址；

(2) BOOTP 服务器接收到请求报文后，根据报文中的 BOOTP 客户端 MAC 地址，从配置文件数据库中查找对应的 IP 地址等信息，并向客户端返回包含这些信息的 BOOTP 响应报文；

(3) BOOTP 客户端从接收到的响应报文中即可获得 IP 地址等信息。

6.1.3 协议规范

与 BOOTP 相关的协议规范有：

- RFC 951：Bootstrap Protocol (BOOTP)
- RFC 2132：DHCP Options and BOOTP Vendor Extensions
- RFC 1542：Clarifications and Extensions for the Bootstrap Protocol

6.2 配置接口通过BOOTP协议获取IP地址

表6-1 配置接口通过 BOOTP 协议获取 IP 地址

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*	-
配置接口通过BOOTP协议获取IP地址	**ip address bootp-alloc**	必选 缺省情况下，接口不通过BOOTP协议获取IP地址

6.3 BOOTP客户端显示和维护

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 BOOTP 客户端的运行情况，通过查看显示信息验证配置的效果。

表6-2 BOOTP 客户端显示和维护

操作	命令
显示BOOTP客户端的相关信息	**display bootp client** [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

6.4 BOOTP客户端典型配置举例

1. 组网需求

Switch B 的端口（属于 VLAN1）接入局域网，VLAN 接口 1 通过 BOOTP 协议从 DHCP 服务器获取 IP 地址。

2. 组网图

如 图 2-2 所示。

3. 配置步骤

下面只列出 图 2-2 中，作为客户端的Switch B的配置。

配置 VLAN 接口 1 通过 BOOTP 动态获取地址。

```
<SwitchB> system-view
[SwitchB] interface vlan-interface 1
[SwitchB-Vlan-interface1] ip address bootp-alloc
```

通过 **display bootp client** 命令可以查看 BOOTP 客户端申请到的 IP 地址。

说明

为了使BOOTP客户端能从DHCP服务器获得IP地址，还需要在DHCP服务器上进行一些配置，具体内容请参见“2.13 DHCP服务器典型配置举例”。

目 录

1 IPv4 域名解析配置

1.1 域名解析简介

域名系统（DNS，Domain Name System）是一种用于 TCP/IP 应用程序的分布式数据库，提供域名与 IP 地址之间的转换。通过域名系统，用户进行某些应用时，可以直接使用便于记忆的、有意义的域名，而由网络中的域名解析服务器将域名解析为正确的 IP 地址。

域名解析分为静态域名解析和动态域名解析，二者可以配合使用。在解析域名时，首先采用静态域名解析（查找静态域名解析表），如果静态域名解析不成功，再采用动态域名解析。由于动态域名解析可能会花费一定的时间，且需要域名服务器的配合，因而可以将一些常用的域名放入静态域名解析表中，这样可以大大提高域名解析效率。

1.1.1 静态域名解析

静态域名解析就是手工建立域名和 IP 地址之间的对应关系。当用户使用域名进行某些应用（如 telnet 应用）时，系统查找静态域名解析表，从中获取指定域名对应的 IP 地址。

1.1.2 动态域名解析

1. 解析过程

动态域名解析是通过对域名服务器的查询完成的。解析过程如下：

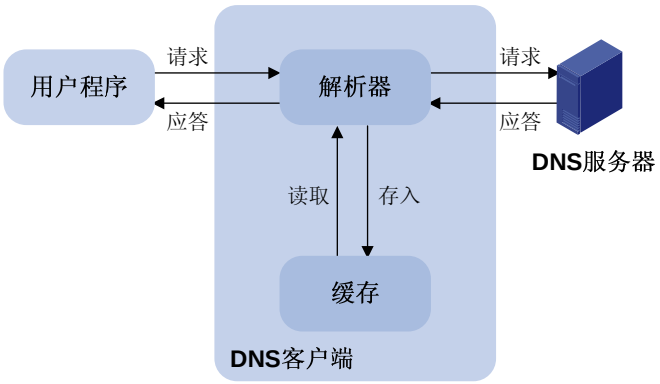
(1) 当用户使用域名进行某些应用时，用户程序首先向 DNS 客户端中的解析器发出请求。

(2) DNS 客户端收到请求后，首先查询本地的域名缓存。如果存在已解析成功的映射项，就将域名对应的 IP 地址返回给用户程序；如果没有发现所要查找的映射项，就向域名服务器（DNS Server）发送查询请求。

(3) 域名服务器首先从自己的数据库中查找域名对应的 IP 地址。如果判断该域名不属于本域范围之内，就将请求交给上一级的域名解析服务器处理，直到完成解析，并将解析的结果返回给 DNS 客户端。

(4) DNS 客户端收到域名服务器的响应报文后，将解析结果返回给应用程序。

图1-1 动态 DNS



用户程序、DNS客户端及域名服务器的关系如 图 1-1 所示，其中解析器和缓存构成DNS客户端。用户程序、DNS客户端在同一台设备上，而DNS客户端和服务器一般分布在两台设备上。

动态域名解析支持缓存功能。每次动态解析成功的域名与 IP 地址的映射均存放在动态域名缓存区中，当下一次查询相同域名的时候，就可以直接从缓存区中读取，不用再向域名服务器进行请求。缓存区中的映射在一段时间后会被老化删除，以保证及时从域名服务器得到最新的内容。老化时间由域名服务器设置，DNS 客户端从协议报文中获得老化时间。

2. 域名后缀列表功能

动态域名解析支持域名后缀列表功能。用户可以预先设置一些域名后缀，在域名解析的时候，用户只需要输入域名的部分字段，系统会自动将输入的域名加上不同的后缀进行解析。举例说明，用户想查询域名 aabbcc.com，那么可以先在后缀列表中配置 com，然后输入 aabbcc 进行查询，系统会自动将输入的域名与后缀连接成 aabbcc.com 进行查询。

使用域名后缀的时候，根据用户输入域名方式的不同，查询方式分成以下几种情况：

- 如果用户输入的域名中没有“.”，比如 aabbcc，系统认为这是一个主机名，会首先加上域名后缀进行查询，如果所有加后缀的域名查询都失败，将使用最初输入的域名（如 aabbcc）进行查询。
- 如果用户输入的域名中间有“.”，比如 www.aabbcc，系统直接用它进行查询，如果查询失败，再依次加上各个域名后缀进行查询。
- 如果用户输入的域名最后有“.”，比如 aabbcc.com.，表示不需要进行域名后缀添加，系统直接用输入的域名进行查询，不论成功与否都直接返回。就是说，如果用户输入的字符中最后一个字符为“.”，就只根据用户输入的字符进行查找，而不会去匹配用户预先设置的域名后缀，因此最后这个“.”，也被称为查找终止符。带有查询终止符的域名，称为 FQDN（Fully Qualified Domain Name，完全合格域名）。

目前，设备支持静态域名解析和动态域名解析的客户端功能。

说明

如果域名服务器上配置了域名的别名，设备可以通过别名来解析主机的 IP 地址。

1.1.3 DNS代理

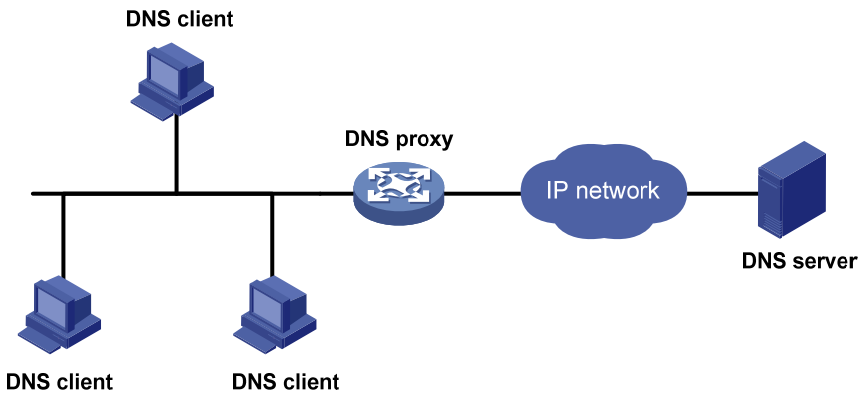
1. DNS代理简介

DNS 代理（DNS proxy）用来在 DNS client 和 DNS server 之间转发 DNS 请求和应答报文。局域网内的 DNS client 把 DNS proxy 当作 DNS server，将 DNS 请求报文发送给 DNS proxy。DNS proxy 将该请求报文转发到真正的 DNS server，并将 DNS server 的应答报文返回给 DNS client，从而实现域名解析。

使用 DNS proxy 功能后，当 DNS server 的地址发生变化时，只需改变 DNS proxy 上的配置，无需改变局域网内每个 DNS client 的配置，从而简化了网络管理。

DNS proxy的典型应用环境如 图 1-2 所示。

图1-2 DNS 代理典型组网应用



2. DNS代理的工作机制

DNS 代理的工作过程如下：

(1) DNS client 把 DNS proxy 当作 DNS server，将 DNS 请求报文发送给 DNS proxy，即请求报文的目的地址为 DNS proxy 的 IP 地址。

(2) DNS proxy 收到请求报文后，首先查找本地的静态域名解析表和动态域名解析表，如果存在请求的信息，则 DNS proxy 直接通过 DNS 应答报文将域名解析结果返回给 DNS client。

(3) 如果不存在请求的信息，则 DNS proxy 将报文转发给 DNS server，通过 DNS server 进行域名解析。

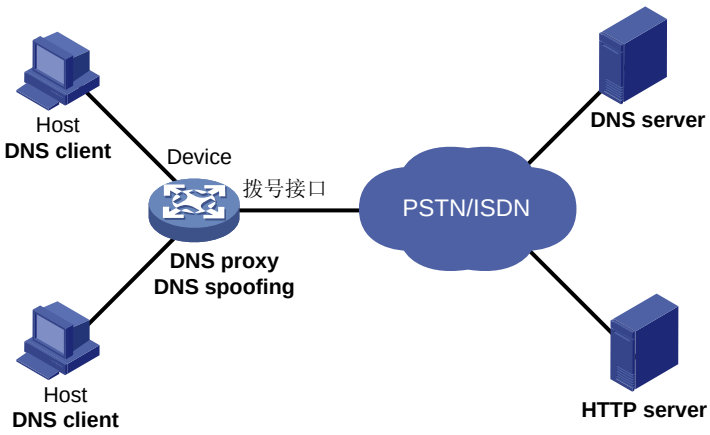
(4) DNS proxy 收到 DNS server 的应答报文后，记录域名解析的结果，并将报文转发给 DNS client。DNS client 利用域名解析的结果进行相应的处理。

说明

只有 DNS proxy 上存在域名服务器地址，并存在到达域名服务器的路由，DNS proxy 才会向 DNS server 发送域名解析请求。否则，DNS proxy 不会向 DNS server 发送域名解析请求，也不会应答 DNS client 的请求。

1.1.4 DNS spoofing

图1-3 DNS spoofing 典型应用场景



DNS spoofing主要应用于 图 1-3 所示的拨号网络。在该网络中：

- Device 通过拨号接口连接到 PSTN/ISDN 等拨号网络。只有存在通过拨号接口转发的报文时，才会触发拨号接口建立连接。
- Device 作为 DNS proxy。在 Host 上将 Device 指定为 DNS 服务器；拨号接口建立连接后，Device 通过 DHCP 等方式动态获取 DNS 服务器地址。

Device 上没有使能 DNS spoofing 功能时，Device 接收到 Host 发送的域名解析请求报文后，如果不存在对应的静态域名解析表项和动态域名解析表项，则需要向 DNS server 发送域名解析请求。但是，由于此时拨号接口尚未建立连接，Device 上不存在 DNS server 地址，Device 不会向 DNS server 发送域名解析请求，也不会应答 DNS client 的请求。从而，导致域名解析失败，且没有流量触发拨号接口建立连接。

DNS spoofing 功能可以解决上述问题。使能 DNS spoofing 功能后，即便 Device 上不存在域名服务器地址或到达 DNS server 的路由，Device 也会利用指定的 IP 地址作为域名解析结果，应答 DNS client 的域名解析请求。DNS client 后续发送的报文可以用来触发拨号接口建立连接。

图 1-3 所示网络中，Host访问HTTP server的报文处理流程为：

(1) Host 通过域名访问 HTTP server 时，首先向 Device 发送域名解析请求，将 HTTP server 的域名解析为 IP 地址。

(2) Device 接收到域名解析请求后，在本地没有找到对应的静态域名解析表项和动态域名解析表项，且拨号接口尚未建立连接，Device 上不存在 DNS server 地址，则 Device 利用 DNS spoofing 中指定的 IP 地址作为域名解析结果，应答 DNS client 的域名解析请求。该域名解析应答的老化时间为 0。并且，应答的 IP 地址满足如下条件：Device 上存在到达该 IP 地址的路由，且路由的出接口为拨号接口。

(3) Host 接收到 Device 的应答报文后，向应答的 IP 地址发送 HTTP 请求。

(4) Device 通过拨号接口转发 HTTP 请求时，触发拨号接口建立连接，并通过 DHCP 等方式动态获取 DNS server 的地址。

(5) 域名解析应答老化后，Host 再次发送域名解析请求。

(6) 之后，Device的处理过程与DNS proxy工作过程相同，请参见“1.1.3 2. DNS代理的工作机制”。

(7) Host 获取到正确的 HTTP server 地址后，可以正常访问 HTTP server。

说明

由于 DNS spoofing 功能指定的 IP 地址并不是待解析域名对应的 IP 地址，为了防止 DNS client 上保存错误的域名解析表项，该 IP 地址对应域名解析应答的老化时间为 0。

1.2 配置IPv4 DNS client

1.2.1 配置静态域名解析

配置静态域名解析就是通过配置使主机名与 IPv4 地址相互对应。当使用 Telnet 等应用时，可以直接使用主机名，由系统解析为 IPv4 地址。

表1-1 配置静态域名解析

操作	命令	说明
进入系统视图	**system-view**	-
配置主机名和对应的IPv4地址	**ip host** *hostname ip-address*	必选 缺省情况下，静态域名解析表中没有主机名及IPv4地址的对应关系

说明

- 每个主机名只能对应一个 IPv4 地址，当对同一主机名进行多次配置时，最后配置的 IPv4 地址有效。
- 最多可配置 50 条 IPv4 静态域名解析信息。

1.2.2 配置动态域名解析

如果用户需要使用动态域名解析功能，可以使用下面的命令使能动态域名解析功能，并配置域名服务器，这样才能将查询请求报文发送到正确的服务器进行解析。

用户还可以配置域名后缀，以便实现只输入域名的部分字段，而由系统自动加上预先设置的后缀进行解析。

表1-2 配置动态域名解析

操作	命令	说明
进入系统视图	**system-view**	-

操作		命令	说明
开启动态域名解析功能		**dns resolve**	必选 缺省情况下，动态域名解析功能处于关闭状态
配置域名服务器的IPv4地址	系统视图下	**dns server** *ip-address*	二者至少选择其一 缺省情况下，没有配置域名服务器的IPv4地址
	接口视图下	**interface** *interface-type interface-number*	
		dns server *ip-address*	
		quit	
配置域名后缀		**dns domain** *domain-name*	可选 缺省情况下，没有配置域名后缀，即只根据用户输入的域名信息进行解析

说明

- 包括 IPv6 域名服务器在内，系统视图下最多可配置 6 个域名服务器；所有接口下配置的域名服务器总数不能超过 6。
- DNS server 的优先级顺序为：系统视图下配置的 DNS server 优先级高于接口视图下配置的 DNS server；在同一视图下先配置的 DNS server 优先级高于后配置的 DNS server；设备上手工配置的 DNS server 优先级高于通过 DHCP 等方式动态获取的 DNS server。设备首先向优先级最高的 DNS server 发送查询请求，失败后再依次向其他 DNS server 发送查询请求。
- 设备上最多可以配置 10 个域名后缀。

1.3 配置DNS proxy

表1-3 配置 DNS proxy

操作		命令	说明
进入系统视图		**system-view**	-
开启DNS proxy功能		**dns proxy enable**	必选 缺省情况下，DNS proxy功能处于关闭状态
配置域名服务器的IPv4地址	系统视图下	**dns server** *ip-address*	二者至少选择其一 缺省情况下，没有配置域名服务器的IPv4地址
	接口视图下	**interface** *interface-type interface-number*	
		dns server *ip-address*	

说明

通过 **dns server** 命令可以指定多个 DNS server。DNS proxy 接收到客户端的查询请求后，首先向优先级最高的 DNS server 转发查询请求，失败后再依次向其他 DNS server 转发查询请求。

1.4 配置DNS spoofing

1.4.1 配置准备

只有在以下条件均满足的情况下，DNS spoofing 功能才会生效：

- 设备上使能了 DNS proxy 功能
- 设备上没有指定域名服务器地址或不存在到达域名服务器的路由

因此，配置 DNS spoofing 前，需要先使能 DNS proxy 功能。

1.4.2 配置DNS spoofing

表1-4 配置 DNS spoofing

操作	命令	说明
进入系统视图	**system-view**	-
开启DNS spoofing功能，并指定应答的IP地址	**dns spoofing** *ip-address*	必选 缺省情况下，DNS spoofing功能处于关闭状态

1.5 配置DNS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 DNS 报文的 DSCP 优先级进行配置。

表1-5 配置 DNS 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
配置发送的DNS报文的DSCP优先级	**dns dscp** *dscp-value*	可选 缺省情况下，发送的DNS报文的DSCP优先级为0

1.6 配置DNS报文的源接口

缺省情况下，设备根据 DNS server 的地址，通过路由表查找报文的出接口，并将该出接口的主 IP 地址作为发送到该服务器的 DNS 查询报文的源地址。根据 DNS server 的地址不同，发送报文的源地址可能会发生变化。在某些特殊的组网环境中，DNS server 只应答来自特定源地址的 DNS 请求报文。这种情况下，必须指定 DNS 报文的源接口。如果为设备配置了 DNS 报文的源接口，则设备在发送 DNS 报文时，将固定使用该接口的主 IP 地址作为报文的源地址。

表1-6 配置 DNS 报文的源接口

操作	命令	说明
进入系统视图	**system-view**	-
配置DNS报文的源接口	**dns source-interface** *interface-type interface-number*	必选 缺省情况下，未指定DNS报文的源接口，设备根据DNS server的地址，通过路由表查找报文的出接口，并将该出接口的主IP地址作为发送到该服务器的DNS查询报文的源地址。

1.7 IPv4域名解析显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示 IPv4 域名解析配置后的运行情况，通过查看显示信息验证配置的效果。

在用户视图下，执行 **reset** 命令可以清除动态域名缓存信息。

表1-7 域名解析显示和维护

操作	命令
显示IPv4静态域名解析表	**display ip host** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv4域名服务器信息	**display dns server** [**dynamic**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示域名后缀列表信息	**display dns domain** [**dynamic**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv4动态域名缓存信息	**display dns host ip** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IPv4动态域名缓存信息	**reset dns host ip**

1.8 IPv4域名解析典型配置举例

1.8.1 静态域名解析配置举例

1. 组网需求

为了避免记忆复杂的 IP 地址，Device 希望通过便于记忆的主机名访问某一主机。在 Device 上手工配置 IP 地址对应的主机名，利用静态域名解析功能，就可以实现通过主机名访问该主机。

在本例中，Device 访问的主机 IP 地址为 10.1.1.2，主机名为 host.com。

2. 组网图

图1-4 静态域名解析配置组网图

3. 配置步骤

配置主机名 host.com 对应的 IP 地址为 10.1.1.2。

```
<Sysname> system-view
[Sysname] ip host host.com 10.1.1.2
```

执行 ping host.com 命令，Device 通过静态域名解析可以解析到 host.com 对应的 IP 地址为 10.1.1.2。

```
[Sysname] ping host.com
  PING host.com (10.1.1.2):
  56  data bytes, press CTRL_C to break
    Reply from 10.1.1.2: bytes=56 Sequence=1 ttl=128 time=1 ms
    Reply from 10.1.1.2: bytes=56 Sequence=2 ttl=128 time=4 ms
    Reply from 10.1.1.2: bytes=56 Sequence=3 ttl=128 time=3 ms
    Reply from 10.1.1.2: bytes=56 Sequence=4 ttl=128 time=2 ms
    Reply from 10.1.1.2: bytes=56 Sequence=5 ttl=128 time=3 ms

  --- host.com ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 1/2/4 ms
```

1.8.2 动态域名解析配置举例

1. 组网需求

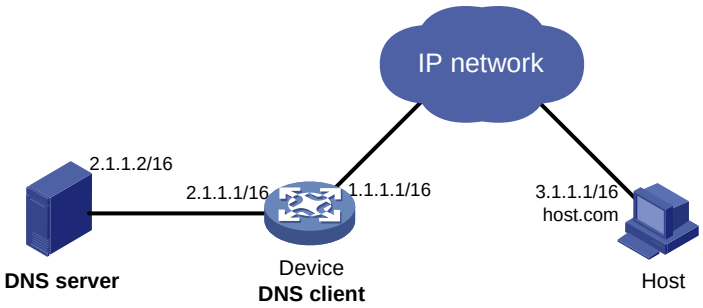
为了避免记忆复杂的 IP 地址，Device 希望通过便于记忆的域名访问某一主机。如果网络中存在域名服务器，则可以利用动态域名解析功能，实现通过域名访问主机。

在本例中：

- 域名服务器的 IP 地址是 2.1.1.2/16，域名服务器上存在 com 域，且 com 域中包含域名“host”和 IP 地址 3.1.1.1/16 的对应关系。
- Device 作为 DNS 客户端，使用动态域名解析功能，将域名解析为 IP 地址。
- Device 上配置域名后缀 com，以便简化访问主机时输入的域名，例如通过输入 host 即可访问域名为 host.com、IP 地址为 3.1.1.1/16 的主机 Host。

2. 组网图

图1-5 动态域名解析组网图



3. 配置步骤

说明

- 在开始下面的配置之前，假设设备与主机之间的路由可达，设备和主机都已经配置完毕，接口IP地址如 图 1-5 所示。
- 不同域名服务器的配置方法不同，下面仅以 Windows Server 2000 为例，说明域名服务器的配置方法。

(1) 配置域名服务器

进入域名服务器配置界面。

在开始菜单中，选择[程序/管理工具/DNS]。

创建区域 com。

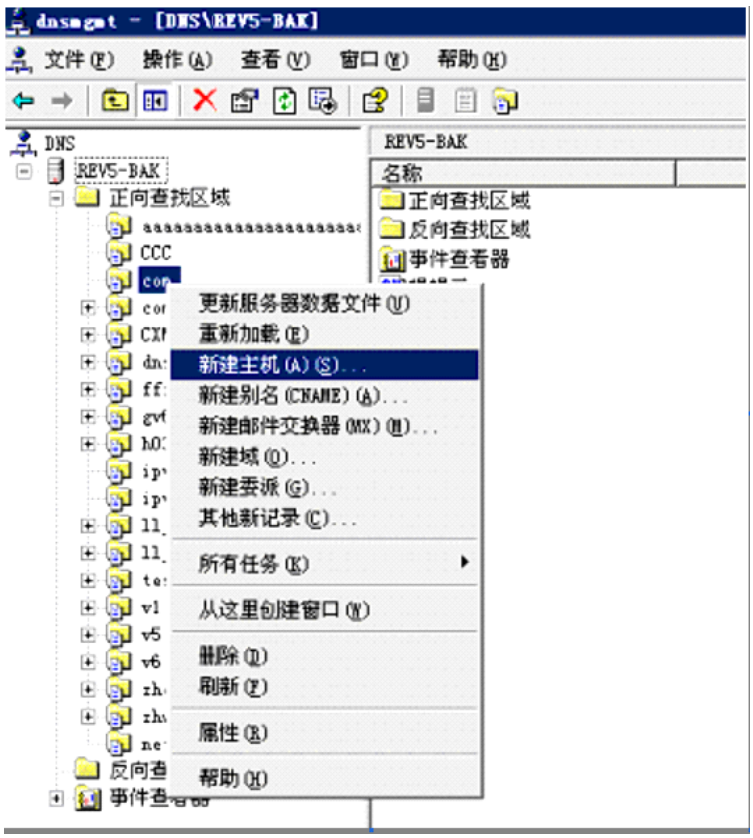
如 图 1-6 所示，右键点击[正向查找区域]，选择[新建区域]，按照提示创建新的区域com。

图1-6 创建区域

添加域名和 IP 地址的映射。

如 图 1-7 所示，右键点击区域com。

图1-7 新建主机



选择[新建主机]，弹出如 图 1-8 的对话框。按照 图 1-8 输入域名host和IP地址 3.1.1.1。

图1-8 添加域名和 IP 地址的映射

(2) 配置 DNS 客户端 Device

开启动态域名解析功能。

```
<Sysname> system-view
[Sysname] dns resolve
```

配置域名服务器的 IP 地址为 2.1.1.2。

```
[Sysname] dns server 2.1.1.2
```

配置域名后缀 com。

```
[Sysname] dns domain com
```

(3) 验证配置结果

在设备上执行 ping host 命令，可以 ping 通主机，且对应的目的地址为 3.1.1.1。

```
[Sysname] ping host
 Trying DNS resolve, press CTRL_C to break
 Trying DNS server (2.1.1.2)
  PING host.com (3.1.1.1):
  56  data bytes, press CTRL_C to break
    Reply from 3.1.1.1: bytes=56 Sequence=1 ttl=126 time=3 ms
    Reply from 3.1.1.1: bytes=56 Sequence=2 ttl=126 time=1 ms
    Reply from 3.1.1.1: bytes=56 Sequence=3 ttl=126 time=1 ms
    Reply from 3.1.1.1: bytes=56 Sequence=4 ttl=126 time=1 ms
    Reply from 3.1.1.1: bytes=56 Sequence=5 ttl=126 time=1 ms

  --- host.com ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 1/1/3 ms
```

1.8.3 DNS proxy典型配置举例

1. 组网需求

某局域网内拥有多台设备，每台设备上都指定了域名服务器的 IP 地址，以便直接通过域名访问外部网络。当域名服务器的 IP 地址发生变化时，网络管理员需要更改局域网内所有设备上配置的域名服务器 IP 地址，工作量将会非常巨大。

通过 DNS proxy 功能，可以大大减少网络管理员的工作量。当域名服务器 IP 地址改变时，只需更改 DNS proxy 上的配置，即可实现局域网内设备通过新的域名服务器解析域名。

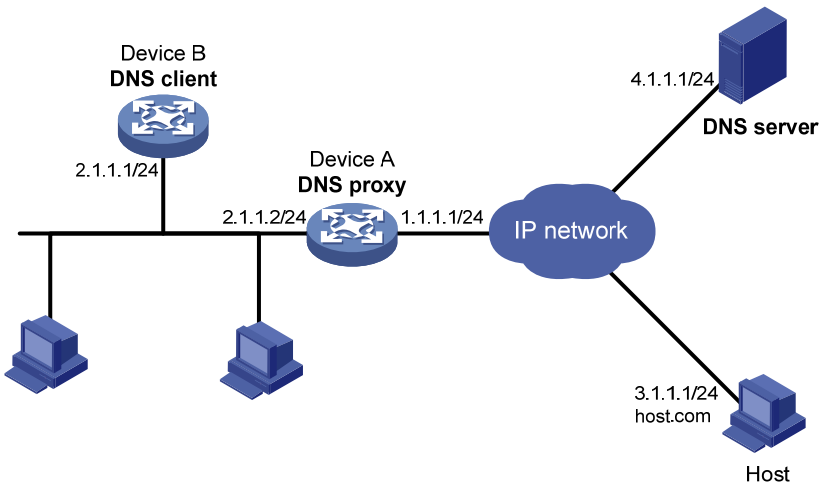
在本例中，具体配置步骤为：

(1) 局域网中的某台设备 Device A 配置为 DNS proxy，DNS proxy 上指定域名服务器 IP 地址为真正的域名服务器的地址 4.1.1.1

(2) 局域网中的其他设备（如 Device B）上，域名服务器的 IP 地址配置为 DNS proxy 的地址，域名解析报文将通过 DNS proxy 转发给真正的域名服务器。

2. 组网图

图1-9 DNS proxy 组网图



3. 配置步骤

说明

在开始下面的配置之前，假设设备与域名服务器、主机之间的路由可达，并已按照 图 1-9 配置各接口的IP地址。

(1) 配置域名服务器

不同的域名服务器的配置方法不同。Windows Server 2000 作为域名服务器时，配置方法请参见“1.8.2 动态域名解析配置举例”。

(2) 配置 DNS 代理 Device A

配置域名服务器的 IP 地址为 4.1.1.1。

```
<DeviceA> system-view
```

```
[DeviceA] dns server 4.1.1.1
```

开启 DNS proxy 功能。

```
[DeviceA] dns proxy enable
```

(3) 配置 DNS 客户端 Device B

开启动态域名解析功能。

```
<DeviceB> system-view
[DeviceB] dns resolve
```

配置域名服务器的 IP 地址为 2.1.1.2。

```
[DeviceB] dns server 2.1.1.2
```

(4) 验证配置结果

在 Device B 上执行 ping host.com 命令，可以 ping 通主机，且对应的目的地址为 3.1.1.1。

```
[DeviceB] ping host.com
Trying DNS resolve, press CTRL_C to break
 Trying DNS server (2.1.1.2)
  PING host.com (3.1.1.1):
  56  data bytes, press CTRL_C to break
    Reply from 3.1.1.1: bytes=56 Sequence=1 ttl=126 time=3 ms
    Reply from 3.1.1.1: bytes=56 Sequence=2 ttl=126 time=1 ms
    Reply from 3.1.1.1: bytes=56 Sequence=3 ttl=126 time=1 ms
    Reply from 3.1.1.1: bytes=56 Sequence=4 ttl=126 time=1 ms
    Reply from 3.1.1.1: bytes=56 Sequence=5 ttl=126 time=1 ms

  --- host.com ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 1/1/3 ms
```

1.9 IPv4域名解析常见配置错误举例

1. 现象描述

配置了动态域名解析，但不能根据域名解析到正确的 IP 地址。

2. 故障分析

DNS 客户端需要和域名服务器配合使用，才能根据域名解析到正确的 IP 地址。

3. 故障排除

- 执行命令 **display dns host ip**，检查动态域名缓存信息是否存在指定域名。
- 如果不存在要解析的域名，检查 DNS 客户端是否和域名服务器通信正常，域名服务器是否工作正常，动态域名解析功能是否已经开启。
- 如果存在要解析的域名，但地址不对，则检查 DNS 客户端所配置的域名服务器的 IP 地址是否正确。
- 检查域名服务器所设置的域名和地址映射表是否正确。

目 录

1 IRDP配置

说明

本文提到的主机表示支持 IRDP 功能的主机。

1.1 IRDP简介

IRDP（ICMP Router Discovery Protocol，ICMP 路由器发现协议）是 ICMP 协议的一个扩展，它使得主机能够动态地发现本地网络中路由器的 IP 地址，并设置自己的缺省路由。

1.1.1 IRDP的产生背景

一个网络中的主机如果要发送报文到网络外部，它至少需要获取本网络内的一个路由器的 IP 地址，由路由器来把报文转发出去。主机通常有两种方式获取路由器的 IP 地址：一是在主机上配置默认网关，二是让主机侦听网络内的路由协议报文，从报文中获取路由器的 IP 地址。

两种方法都有缺点。前一种要求静态配置，必须要手工维护，而且不能适应网络的动态变化；后一种方法要求主机能够识别各种路由协议的报文，这对于一个主机来说要求太高了，而且有时路由器上不运行动态路由协议，此时主机便无法侦听到路由协议报文。

为了解决上述问题，出现了 IRDP 协议，该协议采用两种新的 ICMP 消息类型来实现主机对路由器的发现。IRDP 可以动态适应网络的变化，也不用手工维护大量的配置，并且不依赖于任何一种具体的路由协议，可以很好的解决上面存在的问题。

1.1.2 IRDP的工作机制

IRDP 中用到两种 ICMP 消息：

- 路由公告消息 RA（Router Advertisements）：由路由器发送，用于公告路由器的 IP 地址、优先级等信息。
- 路由请求消息 RS（Router Solicitations）：由主机发送，用于主动向网络中的路由器请求路由器的 IP 地址。

IRDP 的工作机制如下：

- 路由器周期性的从接口发送 RA，公告该接口的 IP 地址（包括接口的主 IP 地址和手工配置的从 IP 地址）。主机接收到 RA 后，就会获取到网络中路由器的 IP 地址。
- 当一台主机刚刚连接到网络上，它可以主动发送 RS 来请求路由器的 IP 地址，而不是被动等待 RA。如果主机发送的 RS 没有回应，它可以重传几次 RS。如果主机通过上述主动方式不能获取路由器的IP地址，那么还可以通过后续路由器周期性公告RA来获取路由器的IP地址。
- 主机接收到 RA 后，将根据 RA 中包含的 IP 地址，添加本地路由。如果主机希望缺省路由也从 RA 中获取，那么主机将从收到的所有 RA 包含的 IP 地址中选择一个优先级最高的 IP 地址作为本机缺省路由。

IRDP 功能只能让主机知道路由器的存在，而并不知道到达某个地址，哪个路由器实际上是最优的。如果主机选中了一个到达某目的地非最优的路由器作为报文转发路由器，一旦报文转发到那台路由器后，主机会收到从那个路由器发来的一个 ICMP 重定向报文，让主机发往这个目的地的报文重定向到一个更优的路由器上。

1.1.3 IRDP概念介绍

下面将介绍 IRDP 协议中用到的一些重要概念。

1. IP地址的优先级

RA 中每一个被公告的 IP 地址都对应一个“优先级”值，这个优先级是主机选择缺省路由的依据。当主机希望从 RA 中获取缺省路由时，它将从收到的所有 RA 包含的 IP 地址中选择一个优先级最高的 IP 地址作为本机缺省路由。

用户可以对路由器公告的 IP 地址的优先级进行配置，以控制主机优先选择哪个 IP 地址作为缺省路由。

优先级值越大表示优先级越高。最小的优先级值（-2147483648）表示主机不要使用这个地址作为缺省路由。

2. IP地址的生命周期

生命周期表示路由器公告的 IP 地址可以在主机上存在的时间。如果后续主机没有收到包含该 IP 地址的新的 RA，那么在这个时间过后，该 IP 地址将被删除。

通过同一个接口公告出去的所有 IP 地址具有相同的生命周期。

3. 周期性发送RA的时间间隔

使能 IRDP 功能后，路由器会周期性发送 RA。发送 RA 不是完全周期性的，每两次发送 RA 的时间间隔是在最小时间间隔和最大时间间隔之间的一个随机值，从而避免同一链路上多个路由器同时发送 RA 对网络性能的影响。

在丢包严重的链路上，建议缩短 RA 的发送周期。

4. RA消息的目的地址

RA 消息的目的 IP 地址可以有两种：广播地址 255.255.255.255、组播地址 224.0.0.1（本地链路所有主机）。

缺省情况下，RA 消息的目的 IP 地址采用广播地址。如果发送 RA 的接口支持组播报文，那么建议使用组播地址 224.0.0.1 作为 RA 消息的目的 IP 地址。

5. 代理公告IP地址

缺省情况下，接口仅向外公告接口的主 IP 地址和手工配置的从 IP 地址，如果用户希望接口公告其他 IP 地址，可以通过命令行手工配置该接口代理公告的 IP 地址。

1.1.4 协议规范

与 IRDP 相关的协议规范有：

- RFC 1256：ICMP Router Discovery Messages

1.2 配置IRDP

表1-1 配置 IRDP

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	- 可以是三层以太网端口、VLAN接口
使能接口的IRDP功能	**ip irdp**	必选 缺省情况下，接口的IRDP功能处于关闭状态
配置接口公告的接口IP地址的优先级	**ip irdp preference** *preference-value*	可选 缺省情况下，接口公告的接口IP地址的优先级为0 本配置对接口公告出去的所有接口IP地址（包括接口的主IP地址和手工配置的从IP地址）有效
配置接口公告的IP地址的生命周期	**ip irdp lifetime** *life-number*	可选 缺省情况下，接口公告的IP地址的生命周期为1800秒 本配置对接口公告出去的所有IP地址（包括接口IP地址、代理公告的IP地址）有效
配置接口发送周期性RA的最小时间间隔	**ip irdp minadvinterval** *min-value*	可选 缺省情况下，接口发送周期性RA的最小时间间隔为450秒
配置接口发送周期性RA的最大时间间隔	**ip irdp maxadvinterval** *max-value*	可选 缺省情况下，接口发送周期性RA的最大时间间隔为600秒
配置接口发送的RA消息的目的IP地址为组播地址224.0.0.1	**ip irdp multicast**	可选 缺省情况下，接口发送的RA消息的目的IP地址为广播地址255.255.255.255
配置接口代理公告的IP地址	**ip irdp address** *ip-address preference*	可选

说明

只有使能接口的 IRDP 功能，其他 IRDP 相关配置才生效。

1.3 IRDP典型配置举例

1. 组网需求

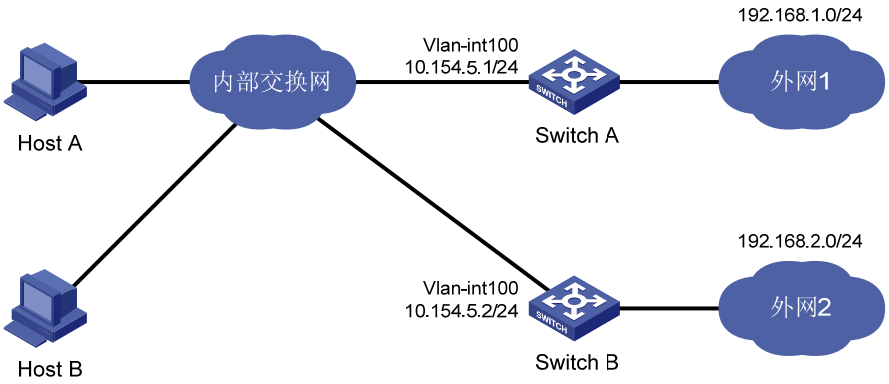
- 公司内部网络中有两台 Linux 系统的主机 Host A 和 Host B，支持 IRDP 功能。

- 内部交换网连接了两台出口交换机 Switch A 和 Switch B，分别到达外网 192.168.1.0/24 和 192.168.2.0/24。

用户希望两台主机使用 Switch A 作为缺省网关，并且到两个外网的报文能正确路由。

2. 组网图

图1-1 配置 IRDP 组网图



3. 配置步骤

(1) 配置 Switch A

配置接口 Vlan-interface100 的 IP 地址。

```
<SwitchA> system-view
[SwitchA] interface Vlan-interface 100
[SwitchA-Vlan-interface100] ip address 10.154.5.1 24
```

使能接口 Vlan-interface100 的 IRDP 功能。

```
[SwitchA-Vlan-interface100] ip irdp
```

配置接口 Vlan-interface100 公告的接口 IP 地址的优先级为 1000。

```
[SwitchA-Vlan-interface100] ip irdp preference 1000
```

配置接口 Vlan-interface100 发送的 RA 消息的目的 IP 地址为组播地址。

```
[SwitchA-Vlan-interface100] ip irdp multicast
```

配置接口 Vlan-interface100 代理公告 IP 地址 192.168.1.0。

```
[SwitchA-Vlan-interface100] ip irdp address 192.168.1.0 400
```

(2) 配置 Switch B

配置接口 Vlan-interface100 的 IP 地址。

```
<SwitchB> system-view
[SwitchB] interface Vlan-interface 100
[SwitchB-Vlan-interface100] ip address 10.154.5.2 24
```

使能接口 Vlan-interface100 的 IRDP 功能。

```
[SwitchB-Vlan-interface100] ip irdp
```

配置接口 Vlan-interface100 公告的接口 IP 地址的优先级为 500。

```
[SwitchB-Vlan-interface100] ip irdp preference 500
```

配置接口 Vlan-interface100 发送的 RA 消息的目的 IP 地址为组播地址。

```
[SwitchB-Vlan-interface100] ip irdp multicast
```

配置接口 Vlan-interface100 代理公告 IP 地址 192.168.2.0。

```
[SwitchB-Vlan-interface100] ip irdp address 192.168.2.0 400
```

(3) 验证配置结果

Host A 和 Host B 打开 IRDP 功能后，查看主机的路由表（以 Host A 为例）。

```
[HostA@localhost ~]$ netstat -rne
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
10.154.5.0      0.0.0.0         255.255.255.0   U     0      0        0 eth1
192.168.1.0     0.0.0.0         255.255.255.0   U     0      0        0 eth1
192.168.2.0     0.0.0.0         255.255.255.0   U     0      0        0 eth1
0.0.0.0         10.154.5.1      0.0.0.0         UG    0      0        0 eth1
```

从上面的信息可以看出，Host A 的缺省路由是 10.154.5.1，并且有到达外网 192.168.1.0/24、192.168.2.0/24 的路由。

目 录

1 IP性能优化配置

说明

IP 性能优化功能中所指的“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性能优化简介

在一些特定的网络环境里，可以通过调整 IP 的参数，以使网络性能达到最佳。IP 性能的优化配置包括：

- 配置允许接收和发送定向广播报文
- 配置 TCP 连接的 Path MTU 探测功能
- 配置 TCP 连接的接收和发送缓冲区的大小
- 配置 TCP 定时器
- 配置 ICMP 差错报文发送功能
- 配置 ICMP 携带扩展信息功能

1.2 配置允许接收和发送定向广播报文

定向广播报文是指发送给特定网络的广播报文。该报文的目的 IP 地址中网络号码字段为特定网络的网络号，主机号码字段为全 1。

如果允许设备接收并转发目的地址为接口所在网络的定向广播报文，黑客就可以利用这样的报文来攻击网络系统，给网络的安全带来了很大的隐患。但在某些应用环境下，设备又需要转发定向广播报文，例如使用 UDP Helper 功能，将广播报文转换为单播报文发送给指定的服务器。

在上述情况下，用户可以通过命令配置设备允许接收和转发定向广播报文。

1.2.1 配置允许接收定向广播报文

如果允许设备接收定向广播报文，则由接口上的配置决定是否转发该报文；否则，直接丢弃定向广播报文。

表1-1 配置允许接收定向广播报文

操作	命令	说明
进入系统视图	**system-view**	-
配置允许设备接收定向广播报文	**ip forward-broadcast**	必选 缺省情况下，禁止设备接收定向广播报文

1.2.2 配置允许转发定向广播报文

表1-2 配置允许转发定向广播报文

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置允许接口转发定向广播报文	**ip forward-broadcast** [**acl** *acl-number*]	必选 缺省情况下，禁止接口转发定向广播报文

说明

- 允许接口转发定向的广播报文时，如果配置了 ACL 规则，则在转发广播报文的同时还需要对报文进行过滤，不符合 ACL 规则的报文将被丢弃，只转发符合 ACL 规则的报文。
- 如果在同一接口下重复执行 **ip forward-broadcast acl** 命令，则后面配置的 ACL 会覆盖以前配置的 ACL；如果后配置的命令不带 **acl** *acl-number*，则以前配置中的 ACL 规则将被取消。

1.2.3 允许接收和发送定向广播报文配置举例

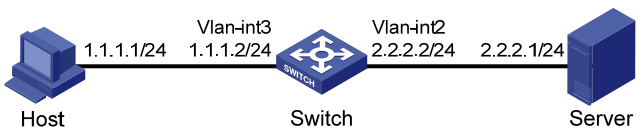
1. 组网需求

如 图 1-1 所示，Host的接口和Switch的VLAN接口 3 处于同一个网段（1.1.1.0/24），Switch的VLAN接口 2 和Server处于另外一个网段（2.2.2.0/24）。Host上配置默认网关为Switch的VLAN接口 3 的地址（1.1.1.2/24）。

要求通过配置使得 Server 可以收到 Host 发送的定向广播报文。

2. 组网图

图1-1 配置收发定向广播报文组网图



3. 配置步骤

配置允许 Switch 接收定向的广播报文。

```
<Switch> system-view
[Switch] ip forward-broadcast
```

配置 VLAN 接口 3 和 VLAN 接口 2 的 IP 地址。

```
[Switch] interface vlan-interface 3
```

```
[Switch-Vlan-interface3] ip address 1.1.1.2 24
[Switch-Vlan-interface3] quit
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ip address 2.2.2.2 24
```

配置允许 VLAN 接口 2 转发定向广播报文。

```
[Switch-Vlan-interface2] ip forward-broadcast
```

1.3 配置TCP属性

1.3.1 配置TCP连接的Path MTU探测功能

RFC 1191 中规定的 TCP 连接的 Path MTU 探测功能，可以探测 TCP 路径上从源端到目的端的最小 MTU，其探测机制如下：

(1) TCP 源端将发送的 TCP 数据段的外层 IP 报文设置 DF（不可分片）标记。

(2) 如果 TCP 路径上某路由器的出接口 MTU 值小于该 IP 报文长度，则会丢弃报文，并给 TCP 源端发送 ICMP 差错报文，报文中会携带该出接口 MTU 值。

(3) TCP 源端通过解析该 ICMP 差错报文，可知 TCP 路径上当前最小的单向 MTU 值。

(4) 后续 TCP 源端发送数据段的长度不超过 MSS。其中，MSS＝最小 MTU 值－IP 头部长度－TCP 头部长度。

说明

- 当 MSS 已经达到系统规定的最小的 32 字节后，如果再次收到减少 MSS 的 ICMP 差错报文，系统将允许该 TCP 连接发送的报文进行分片。
- 产生 ICMP 差错报文的路由器可能不支持 RFC 1191，其产生的 ICMP 差错报文中的出接口 MTU 字段值为 0，对于这种报文，TCP 源端将按照 RFC 1191 中规定的 MTU 表获取比当前路径 MTU 更小的值作为计算 TCP MSS 的基础。MTU 表的内容为（单位为字节）：68、296、508、1006、1280、1492、2002、4352、8166、17914、32000、65535（由于系统规定的 TCP 最小 MSS 为 32，所以对应最小的 MTU 实际为 72 字节）。

用户通过命令行开启 TCP 连接的 Path MTU 探测功能后，新建的 TCP 连接均会携带 Path MTU 探测属性，可以通过上述探测机制确定 Path MTU，按照数据路径上的最小 MTU 组织 TCP 分段长度，最大限度利用网络资源，避免 IP 分片的发生。

Path MTU 值可以老化，这样当 Path MTU 增大时可以充分利用网络资源，尽量按照转发路径可以容忍的最大报文长度发送数据。Path MTU 的老化机制如下：

- 当 TCP 源端收到 ICMP 差错报文后，除了减小 Path MTU 值，同时会为该 Path MTU 值启动老化定时器。
- 当该定时器超时后，系统将按照 RFC 1191 规定的 MTU 表依次递增 TCP 的 MSS 值。
- 如果增加一次 MSS 之后的 2 分钟内没有收到 ICMP 差错报文，则继续递增，直到 MSS 增长到对端在 TCP 三次握手阶段通告的 MSS 值。

表1-3 配置 TCP 连接的 Path MTU 探测功能

操作	命令	说明
进入系统视图	**system-view**	-
开启TCP连接的Path MTU探测功能	**tcp path-mtu-discovery** [**aging** *minutes* | **no-aging**]	必选 缺省情况下，TCP连接的Path MTU探测功能处于关闭状态

注意

TCP 连接的 Path MTU 探测功能依赖 IP 报文的 DF 标记位设置后触发 ICMP 差错报文，因此需要 TCP 路径上的所有设备打开 ICMP 差错报文发送功能（**ip unreachables enable**），以确保 ICMP 差错报文可以发送到 TCP 源端。

1.3.2 配置TCP连接的接收和发送缓冲区大小

表1-4 配置 TCP 连接的接收和发送缓冲区大小

操作	命令	说明
进入系统视图	**system-view**	-
配置TCP连接的接收和发送缓冲区的大小	**tcp window** *window-size*	可选 缺省情况下，TCP连接的接收和发送缓冲区大小为8KB

1.3.3 配置TCP定时器

可以配置的 TCP 定时器包括：

- synwait 定时器：当发送 SYN 报文时，TCP 启动 synwait 定时器，如果 synwait 超时前未收到回应报文，则 TCP 连接建立不成功。
- finwait 定时器：当 TCP 的连接状态为 FIN_WAIT_2 时，启动 finwait 定时器，如果在定时器超时前没有收到报文，则 TCP 连接终止；如果收到 FIN 报文，则 TCP 连接状态变为 TIME_WAIT 状态；如果收到非 FIN 报文，则从收到的最后一个非 FIN 报文开始重新计时，在超时后中止连接。

表1-5 配置 TCP 定时器

操作	命令	说明
进入系统视图	**system-view**	-
配置TCP的synwait定时器超时时间	**tcp timer syn-timeout** *time-value*	可选 缺省情况下，synwait定时器超时时间为75秒

操作	命令	说明
配置TCP的finwait定时器超时时间	**tcp timer fin-timeout** *time-value*	可选 缺省情况下，finwait定时器超时时间为675秒

注意

finwait 定时器的实际超时时间由如下公式决定：finwait 定时器的实际超时时间 =（配置的 finwait 定时器超时时间 – 75）+ 配置的 synwait 定时器超时时间。

1.4 配置ICMP差错报文发送功能

1.4.1 ICMP差错报文发送功能简介

发送差错报文是 ICMP（Internet Control Message Protocol，互联网控制消息协议）的主要功能之一。差错报文通常被网络层或传输层协议用来在异常情况发生时通知相应设备，从而便于进行控制管理。

1. ICMP差错报文发送功能的作用

重定向报文、超时报文、目的不可达报文是 ICMP 差错报文中的三种。下面分别介绍这三种差错报文发送的条件及作用。

(1) ICMP 重定向报文发送功能

主机启动时，它的路由表中可能只有一条到缺省网关的缺省路由。当满足一定的条件时，缺省网关会向源主机发送 ICMP 重定向报文，通知主机重新选择正确的下一跳进行后续报文的发送。

满足下列条件时，设备会发送 ICMP 重定向报文：

- 接收和转发数据报文的接口是同一接口；
- 被选择的路由本身没有被 ICMP 重定向报文创建或修改过；
- 被选择的路由不是设备的默认路由；
- 数据报文中没有源路由选项。

ICMP 重定向报文发送功能可以简化主机的管理，使具有很少选路信息的主机逐渐建立较完善的路由表，从而找到最佳路由。

(2) ICMP 超时报文发送功能

ICMP 超时报文发送功能是在设备收到 IP 数据报文后，如果发生超时差错，则将报文丢弃并给源端发送 ICMP 超时差错报文。

设备在满足下列条件时会发送 ICMP 超时报文：

- 设备收到 IP 数据报文后，如果报文的目的地不是本地且报文的 TTL 字段是 1，则发送“TTL 超时”ICMP 差错报文；
- 设备收到目的地址为本地的 IP 数据报文的第一个分片后，启动定时器，如果所有分片报文到达之前定时器超时，则会发送“重组超时”ICMP 差错报文。

(3) ICMP 目的不可达报文发送功能

ICMP 目的不可达报文发送功能是在设备收到 IP 数据报文后，如果发生目的不可达的差错，则将报文丢弃并给源端发送 ICMP 目的不可达差错报文。

设备在满足下列条件时会发送目的不可达报文：

- 设备在转发报文时，如果在路由表中没有找到对应的转发路由，且路由表中没有缺省路由，则给源端发送“网络不可达”ICMP 差错报文；
- 设备收到目的地址为本地的数据报文时，如果设备不支持数据报文采用的传输层协议，则给源端发送“协议不可达”ICMP 差错报文；
- 设备收到目的地址为本地、传输层协议为 UDP 的数据报文时，如果报文的端口号与正在使用的进程不匹配，则给源端发送“端口不可达”ICMP 差错报文；
- 源端如果采用“严格的源路由选择”发送报文，当中间设备发现源路由所指定的下一个设备不在其直接连接的网络上，则给源端发送“源站路由失败”的 ICMP 差错报文；
- 设备在转发报文时，如果转发接口的 MTU 小于报文的长度，但报文被设置了不可分片，则给源端发送“需要进行分片但设置了不可分片比特”ICMP 差错报文。

2. ICMP差错报文发送功能的弊端

ICMP 差错报文的发送虽然方便了网络的控制管理，但也存在一定的弊端：

- 由于发送大量的 ICMP 报文，增大了网络流量。
- 如果设备接收到大量需要发送 ICMP 差错报文的恶意攻击报文，设备会因为处理大量该类报文而导致性能降低。
- 由于重定向功能会在主机的路由表中增加主机路由，当增加的主机路由很多时，会降低主机性能。
- 由于 ICMP 目的不可达报文传递给用户进程的信息为不可达信息，如果有用户恶意攻击，可能会影响终端用户的正常使用。

为了避免上述现象发生，可以关闭设备的 ICMP 差错报文发送功能，从而减少网络流量、防止遭到恶意攻击。

1.4.2 配置ICMP差错报文发送功能

表1-6 配置 ICMP 差错报文发送功能

操作	命令	说明
进入系统视图	**system-view**	-
开启ICMP重定向报文发送功能	**ip redirects enable**	必选 缺省情况下，ICMP重定向报文发送功能处于关闭状态
开启ICMP超时报文发送功能	**ip ttl-expires enable**	必选 缺省情况下，ICMP超时报文发送功能处于关闭状态
开启ICMP目的不可达报文发送功能	**ip unreachables enable**	必选 缺省情况下，ICMP目的不可达报文发送功能处于关闭状态

说明

关闭 ICMP 超时报文发送功能后，设备不会再发送“TTL 超时”ICMP 差错报文，但“重组超时”ICMP 差错报文仍会正常发送。

1.5 配置ICMP携带扩展信息功能

1.5.1 ICMP携带扩展信息功能简介

传统的 ICMP 报文格式是固定的，不能携带扩展信息。在使能 ICMP 携带扩展信息功能后，设备可以在需要的时候在 ICMP 报文的后面加上扩展信息字段。目前，设备仅支持在 ICMP 报文中扩展携带 MPLS 标签信息。

1. ICMP携带扩展信息的作用

在 MPLS 网络中，当报文在转发过程中出现 TTL 超时后，MPLS 会剥离掉 MPLS 头，构造一个 ICMP TTL 超时报文，送到 MPLS 隧道出口，隧道出口再重新把报文回送给隧道的源端。传统 ICMP 报文里面不能携带标签等信息，而这些信息对于源端来说却非常重要。使能 ICMP 携带扩展信息功能后，设备可以在 TTL 超时后，把当时的标签附加到 TTL 超时报文的后面，回送给源端。

ICMP 携带扩展信息功能通常在 MPLS 网络中进行 Tracert 时使用，通过在 ICMP 报文中携带标签，就可以打印出 MPLS 转发过程每一跳的标签信息。

2. ICMP携带扩展信息功能原理

ICMP 报文可以分为三类：

- 传统 ICMP 报文：报文中不携带扩展信息。
- 携带长度字段的扩展 ICMP 报文：报文中携带扩展信息，并且携带长度字段，长度字段的值为 ICMP 头后面的原始数据长度（不包含扩展信息的长度）。此类报文符合 RFC 4884 的要求。
- 不携带长度字段的扩展 ICMP 报文：报文中携带扩展信息，但是不携带长度字段。此类报文不符合 RFC 4884 的要求。

根据设备对三类ICMP报文的处理机制，可将设备分为三种模式：传统模式、兼容模式和非兼容模式。三种模式设备对三类ICMP报文的处理机制如 表 1-7 所示。

表1-7 三种模式设备对三类 ICMP 报文的处理机制

设备模式	可以发送的 ICMP 报文	可以正确接收的 ICMP 报文	备注
传统模式	传统ICMP报文	传统ICMP报文	如果收到扩展ICMP报文，不会处理报文中的扩展信息
兼容模式	传统ICMP报文 携带长度字段的扩展ICMP报文	传统ICMP报文 携带长度字段的扩展ICMP报文	如果收到不携带长度字段的扩展ICMP报文，则认为报文不携带扩展信息，按传统ICMP报文进行处理

设备模式	可以发送的 ICMP 报文	可以正确接收的 ICMP 报文	备注
非兼容模式	传统ICMP报文 不携带长度字段的扩展ICMP报文	传统ICMP报文 携带长度字段的扩展ICMP报文 不携带长度字段的扩展ICMP报文	-

说明

IPv4 的重定向报文、超时报文、目的不可达报文，以及 IPv6 的目的不可达、超时报文可以携带扩展信息，其余类型的 ICMP/ICMPv6 报文不能携带扩展信息。

1.5.2 配置ICMP携带扩展信息功能

表1-8 配置 ICMP 携带扩展信息功能

操作	命令	说明
进入系统视图	**system-view**	-
使能ICMP携带扩展信息功能，采用兼容模式	**ip icmp-extensions compliant**	可选 缺省情况下，ICMP不携带扩展信息
使能ICMP携带扩展信息功能，采用非兼容模式	**ip icmp-extensions non-compliant**	可选 缺省情况下，ICMP不携带扩展信息

说明

关闭 ICMP 携带扩展信息功能后，设备发送的 ICMP 报文都不携带扩展信息。

1.6 IP性能优化显示和维护

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 IP 性能后的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令清除 IP、TCP 和 UDP 的流量统计信息。

表1-9 IP 性能优化显示和维护

操作	命令
显示TCP连接的流量统计信息	**display tcp statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示UDP流量统计信息	**display udp statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IP报文统计信息	**display ip statistics** [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
显示ICMP流量统计信息	**display icmp statistics** [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示套接口信息	**display ip socket** [**socktype** *sock-type*] [*task-id socket-id*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示FIB信息	**display fib** [**vpn-instance** *vpn-instance-name*] [**acl** *acl-number* | **ip-prefix** *ip-prefix-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示与指定目的IP地址匹配的FIB信息	**display fib** [**vpn-instance** *vpn-instance-name*] *ip-address* [*mask* | *mask-length*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IP报文统计信息	**reset ip statistics** [**slot** *slot-number*]
清除TCP连接的流量统计信息	**reset tcp statistics**
清除UDP流量统计信息	**reset udp statistics**

目 录

1 UDP Helper配置

说明

UDP Helper 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 UDP Helper简介

网络中的主机有时需要通过发送广播报文，来获得网络配置或查询网络中其他设备的名称。但是，当主机与服务器或待查询的设备不在同一个广播域时，主机就无法获得所需要的信息。

为解决上述问题，设备提供了 UDP Helper 功能。通过该功能可以实现对指定 UDP 端口的 IP 广播报文进行中继转发，即将指定 UDP 端口的广播报文转换为单播报文发送给指定的目的服务器，起到中继的作用。

使能 UDP Helper 功能后，如果设备接收到广播报文，将根据报文的 UDP 目的端口号来判断是否要对其中继转发，并进行相应的处理：

- 如果报文的 UDP 目的端口号与配置的需要中继转发的 UDP 端口号匹配，则修改 IP 报文头的目的 IP 地址，将报文发给指定的目的服务器；
- 如果报文的 UDP 目的端口号与配置的需要中继转发的 UDP 端口号不匹配，则直接将报文送给上层协议处理。

1.2 配置UDP Helper

表1-1 配置 UDP Helper

操作	命令	说明
进入系统视图	**system-view**	-
使能UDP Helper功能	**udp-helper enable**	必选 缺省情况下，UDP Helper 功能处于关闭状态
配置需要中继转发的UDP端口	**udp-helper port** { *port-number* | **dns** | **netbios-ds** | **netbios-ns** | **tacacs** | **tftp** | **time** }	必选 缺省情况下，没有配置中继转发的UDP端口
进入接口视图	**interface** *interface-type interface-number*	-
配置中继转发的目的服务器	**udp-helper server** [**vpn-instance** *vpn-instance-name*] *ip-address*	必选 缺省情况下，没有配置中继转发的目的服务器

注意

- 本系列交换机缺省情况下禁止接收所有的定向广播报文。因此，只有在系统视图下配置 **ip forward-broadcast** 命令后，UDP Helper 功能才可用。定向广播报文抑制功能的详细介绍请参见“三层技术-IP 业务配置指导”中的“IP 性能优化”。
- UDP Helper 功能不能中继转发 DHCP 广播报文，即中继转发的 UDP 端口不能配置为 67 和 68。
- 需要中继转发的 UDP 端口有两种配置方法：指定端口号配置和指定参数配置。例如：**udp-helper port** 53 和 **udp-helper port dns** 的效果是一样的。
- 关闭 UDP Helper 功能后，所有已配置的 UDP 端口都被取消。
- 设备上最多可以配置 256 个需要中继转发的 UDP 端口。
- 一个接口上最多可以配置 20 个中继转发的目的服务器。

1.3 UDP Helper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 UDP 中继转发的目的服务器信息以及中继转发到目的服务器的报文数目，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 UDP 中继转发的报文统计数目。

表1-2 UDP Helper 显示和维护

操作	命令
显示UDP中继转发的相关信息	**display udp-helper server** [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除UDP中继转发的报文统计数目	**reset udp-helper packet**

1.4 UDP Helper典型配置举例

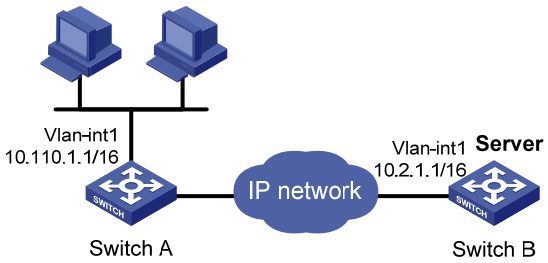
1.4.1 UDP Helper配置举例

1. 组网需求

Switch A 的 VLAN 接口 1 的 IP 地址为 10.110.1.1/16，连接到网段 10.110.0.0/16。配置将目的 UDP 端口号为 55、目的 IP 地址为 255.255.255.255 和此网段子网广播地址 10.110.255.255 的广播报文，中继转发到公网中的目的服务器 10.2.1.1/16。

2. 组网图

图1-1 UDP Helper 配置举例组网图



3. 配置步骤

说明

用户需保证 Switch A 到网段 10.2.0.0/16 路由可达。

配置允许 Switch A 接收定向的广播报文。

```
<SwitchA> system-view
[SwitchA] ip forward-broadcast
```

使能 UDP Helper 功能。

```
[SwitchA] udp-helper enable
```

配置将目的 UDP 端口号为 55 的广播报文进行中继转发。

```
[SwitchA] udp-helper port 55
```

配置公网中的中继转发的目的服务器为 10.2.1.1。

```
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] ip address 10.110.1.1 16
[SwitchA-Vlan-interface1] udp-helper server 10.2.1.1
```

目 录

1 IPv6 基础配置

说明

本文中所指的“接口”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 IPv6简介

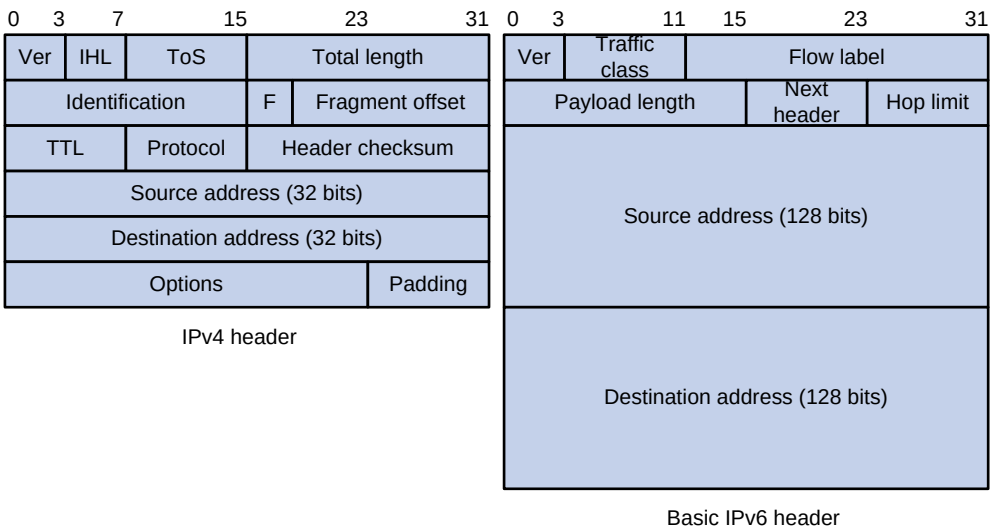
IPv6（Internet Protocol Version 6，因特网协议版本 6）是网络层协议的第二代标准协议，也被称为 IPng（IP Next Generation，下一代因特网），它是 IETF（Internet Engineering Task Force，互联网工程任务组）设计的一套规范，是 IPv4 的升级版本。IPv6 和 IPv4 之间最显著的区别为：IP 地址的长度从 32 比特增加到 128 比特。

1.1.1 IPv6 协议特点

1. 简化的报文头格式

通过将 IPv4 报文头中的某些字段裁减或移入到扩展报文头，减小了 IPv6 基本报文头的长度。IPv6 使用固定长度的基本报文头，从而简化了转发设备对 IPv6 报文的处理，提高了转发效率。尽管 IPv6 地址长度是 IPv4 地址长度的四倍，但 IPv6 基本报文头的长度只有 40 字节，为 IPv4 报文头长度（不包括选项字段）的两倍。

图1-1 IPv4 报文头和 IPv6 基本报文头格式比较



2. 充足的地址空间

IPv6 的源地址与目的地址长度都是 128 比特（16 字节）。它可以提供超过 3.4×10^{38} 种可能的地址空间，完全可以满足多层次的地址划分需要，以及公有网络和机构内部私有网络的地址分配。

3. 层次化的地址结构

IPv6 的地址空间采用了层次化的地址结构，有利于路由快速查找，同时可以借助路由聚合，有效减少 IPv6 路由表占用的系统资源。

4. 地址自动配置

为了简化主机配置，IPv6 支持有状态地址配置和无状态地址配置：

- 有状态地址配置是指从服务器（如 DHCP 服务器）获取 IPv6 地址及相关信息；
- 无状态地址配置是指主机根据自己的链路层地址及路由器发布的前缀信息自动配置 IPv6 地址及相关信息。

同时，主机也可根据自己的链路层地址及默认前缀（FE80::/10）形成链路本地地址，实现与本链路上其他主机的通信。

5. 内置安全性

IPv6 将 IPSec 作为它的标准扩展头，可以提供端到端的安全特性。这一特性也为解决网络安全问题提供了标准，并提高了不同 IPv6 应用之间的互操作性。

6. 支持QoS

IPv6 报文头的流标签（Flow Label）字段实现流量的标识，允许设备对某一流中的报文进行识别并提供特殊处理。

7. 增强的邻居发现机制

IPv6 的邻居发现协议是通过一组 ICMPv6（Internet Control Message Protocol for IPv6，IPv6 互联网控制消息协议）消息实现的，管理着邻居节点间（即同一链路上的节点）信息的交互。它代替了 ARP（Address Resolution Protocol，地址解析协议）、ICMPv4 路由器发现和 ICMPv4 重定向消息，并提供了一系列其他功能。

8. 灵活的扩展报文头

IPv6 取消了 IPv4 报文头中的选项字段，并引入了多种扩展报文头，在提高处理效率的同时还大大增强了 IPv6 的灵活性，为 IP 协议提供了良好的扩展能力。IPv4 报文头中的选项字段最多只有 40 字节，而 IPv6 扩展报文头的大小只受到 IPv6 报文大小的限制。

1.1.2 IPv6 地址介绍

1. IPv6 地址表示方式

IPv6 地址被表示为以冒号（:）分隔的一连串 16 比特的十六进制数。每个 IPv6 地址被分为 8 组，每组的 16 比特用 4 个十六进制数来表示，组和组之间用冒号隔开，比如：2001:0000:130F:0000:0000:09C0:876A:130B。

为了简化 IPv6 地址的表示，对于 IPv6 地址中的“0”可以有下面的处理方式：

- 每组中的前导“0”可以省略，即上述地址可写为 2001:0:130F:0:0:9C0:876A:130B。
- 如果地址中包含连续两个或多个均为 0 的组，则可以用双冒号“::”来代替，即上述地址可写为 2001:0:130F::9C0:876A:130B。

注意

在一个 IPv6 地址中只能使用一次双冒号“::”，否则当设备将“::”转变为 0 以恢复 128 位地址时，将无法确定“::”所代表的 0 的个数。

IPv6 地址由两部分组成：地址前缀与接口标识。其中，地址前缀相当于 IPv4 地址中的网络号码字段部分，接口标识相当于 IPv4 地址中的主机号码部分。

地址前缀的表示方式为：IPv6 地址/前缀长度。其中，IPv6 地址是前面所列出的任一形式，而前缀长度是一个十进制数，表示 IPv6 地址最左边多少位为地址前缀。

2. IPv6 的地址分类

IPv6 主要有三种类型的地址：单播地址、组播地址和任播地址。

- 单播地址：用来唯一标识一个接口，类似于 IPv4 的单播地址。发送到单播地址的数据报文将被传送给此地址所标识的接口。
- 组播地址：用来标识一组接口（通常这组接口属于不同的节点），类似于 IPv4 的组播地址。发送到组播地址的数据报文被传送给此地址所标识的所有接口。
- 任播地址：用来标识一组接口（通常这组接口属于不同的节点）。发送到任播地址的数据报文被传送给此地址所标识的一组接口中距离源节点最近（根据使用的路由协议进行度量）的一个接口。

说明

IPv6 中没有广播地址，广播地址的功能通过组播地址来实现。

IPv6 地址类型是由地址前面几位（称为格式前缀）来指定的，主要地址类型与格式前缀的对应关系如 表 1-1 所示。

表1-1 地址类型与格式前缀的对应关系

地址类型		格式前缀（二进制）	IPv6 前缀标识
单播地址	未指定地址	00...0 (128 bits)	::/128
	环回地址	00...1 (128 bits)	::1/128
	链路本地地址	1111111010	FE80::/10
	站点本地地址	1111111011	FEC0::/10
	全球单播地址	其他形式	-
组播地址		11111111	FF00::/8
任播地址		从单播地址空间中进行分配，使用单播地址的格式	

3. 单播地址的类型

IPv6 单播地址的类型可有多种，包括全球单播地址、链路本地地址和站点本地地址等。

- 全球单播地址等同于 IPv4 公网地址，提供给网络服务提供商。这种类型的地址允许路由前缀的聚合，从而限制了全球路由表项的数量。
- 链路本地地址用于邻居发现协议和无状态自动配置中链路本地上节点之间的通信。使用链路本地地址作为源或目的地址的数据报文不会被转发到其他链路上。
- 站点本地地址与 IPv4 中的私有地址类似。使用站点本地地址作为源或目的地址的数据报文不会被转发到本站点（相当于一个私有网络）外的其它站点。
- 环回地址：单播地址 0:0:0:0:0:0:0:1（简化表示为::1）称为环回地址，不能分配给任何物理接口。它的作用与在 IPv4 中的环回地址相同，即节点用来给自己发送 IPv6 报文。
- 未指定地址：地址“::”称为未指定地址，不能分配给任何节点。在节点获得有效的 IPv6 地址之前，可在发送的IPv6报文的源地址字段填入该地址，但不能作为IPv6报文中的目的地址。

4. 组播地址

表 1-2 所示的组播地址，是预留的特殊用途的组播地址。

表1-2 预留的 IPv6 组播地址列表

地址	应用
FF01::1	表示节点本地范围所有节点的组播地址
FF02::1	表示链路本地范围所有节点的组播地址
FF01::2	表示节点本地范围所有路由器的组播地址
FF02::2	表示链路本地范围所有路由器的组播地址

另外，还有一类组播地址：被请求节点（Solicited-Node）地址。该地址主要用于获取同一链路上邻居节点的链路层地址及实现重复地址检测。每一个单播或任播 IPv6 地址都有一个对应的被请求节点地址。其格式为：

FF02:0:0:0:0:1:FFXX:XXXX

其中，FF02:0:0:0:0:1:FF 为 104 位固定格式；XX:XXXX 为单播或任播 IPv6 地址的后 24 位。

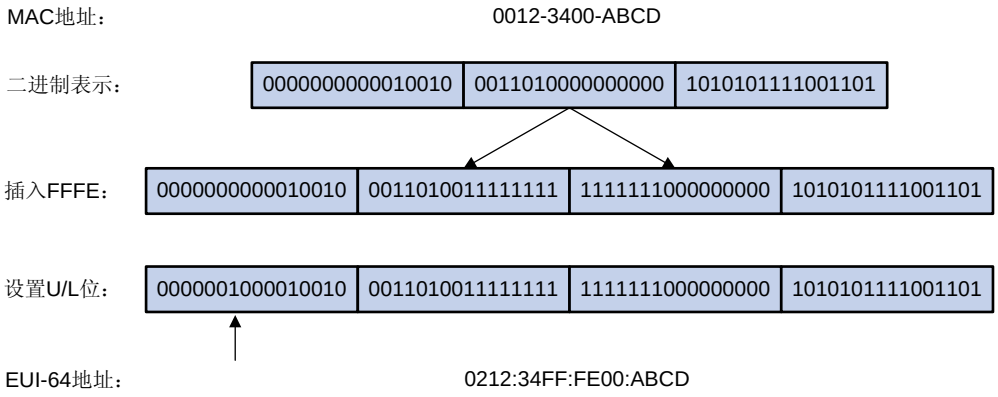
5. IEEE EUI-64 格式的接口标识符

IPv6 单播地址中的接口标识符用来标识链路上的一个唯一的接口。目前 IPv6 单播地址基本上都要求接口标识符为 64 位。

不同接口的 IEEE EUI-64（64-bit Extended Unique Identifier，64 位扩展唯一标识符）格式的接口标识符的生成方法不同，分别介绍如下：

- 所有 IEEE 802 接口类型（例如，VLAN 接口）：IEEE EUI-64 格式的接口标识符是从接口的链路层地址（MAC 地址）变化而来的。IPv6 地址中的接口标识符是 64 位，而 MAC 地址是 48 位，因此需要在 MAC 地址的中间位置（从高位开始的第 24 位后）插入十六进制数 FFFE（1111111111111110）。为了表示这个从 MAC 地址得到的接口标识符是全球唯一的，还要将 Universal/Local (U/L)位（从高位开始的第 7 位）设置为“1”。最后得到的这组数就作为 EUI-64 格式的接口标识符。

图1-2 MAC 地址到 EUI-64 格式接口标识符的转换过程



- Tunnel 接口：IEEE EUI-64 格式的接口标识符的低 32 位为 Tunnel 接口的源 IPv4 地址，ISATAP 隧道的接口标识符的高 32 位为 0000:5EFE，其他隧道的接口标识符的高 32 位为全 0。关于各种隧道的介绍，请参见“三层技术-IP 业务配置指导”中的“隧道”。
- 其他接口类型：IEEE EUI-64 格式的接口标识符由设备随机生成。

1.1.3 IPv6 邻居发现协议介绍

IPv6 邻居发现（Neighbor Discovery，ND）协议使用五种类型的 ICMPv6 消息，实现下面一些功能：地址解析、验证邻居是否可达、重复地址检测、路由器发现/前缀发现、地址自动配置和重定向等功能。

邻居发现协议使用的ICMPv6 消息的类型及作用如 表 1-3 所示。

表1-3 邻居发现协议使用的 ICMPv6 消息类型及作用

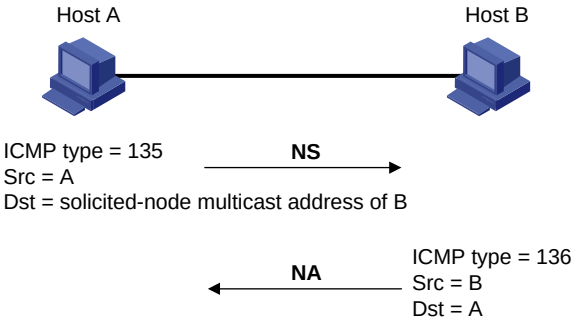
ICMPv6 消息	类型号	作用
邻居请求消息NS（Neighbor Solicitation）	135	获取邻居的链路层地址
		验证邻居是否可达
		进行重复地址检测
邻居通告消息NA（Neighbor Advertisement）	136	对NS消息进行响应
		节点在链路层变化时主动发送NA消息，向邻居节点通告本节点的变化信息
路由器请求消息RS（Router Solicitation）	133	节点启动后，通过RS消息向路由器发出请求，请求前缀和其他配置信息，用于节点的自动配置
路由器通告消息RA（Router Advertisement）	134	对RS消息进行响应
		在没有抑制RA消息发布的条件下，路由器会周期性地发布RA消息，其中包括前缀信息选项和一些标志位的信息
重定向消息（Redirect）	137	当满足一定的条件时，缺省网关通过向源主机发送重定向消息，使主机重新选择正确的下一跳地址进行后续报文的发送

邻居发现协议提供的主要功能如下：

1. 地址解析

获取同一链路上邻居节点的链路层地址（与IPv4 的ARP功能相同），通过邻居请求消息NS和邻居通告消息NA实现。如 图 1-3 所示，节点A要获取节点B的链路层地址。

图1-3 地址解析示意图



(1) 节点 A 以组播方式发送 NS 消息。NS 消息的源地址是节点 A 的接口 IPv6 地址，目的地址是节点 B 的被请求节点组播地址，消息内容中包含了节点 A 的链路层地址。

(2) 节点 B 收到 NS 消息后，判断报文的目的地址是否为自己的 IPv6 地址对应的被请求节点组播地址。如果是，则节点 B 可以学习到节点 A 的链路层地址，并以单播方式返回 NA 消息，其中包含了自己的链路层地址。

(3) 节点 A 从收到的 NA 消息中就可获取到节点 B 的链路层地址。

2. 验证邻居是否可达

在获取到邻居节点的链路层地址后，通过邻居请求消息 NS 和邻居通告消息 NA 可以验证邻居节点是否可达。

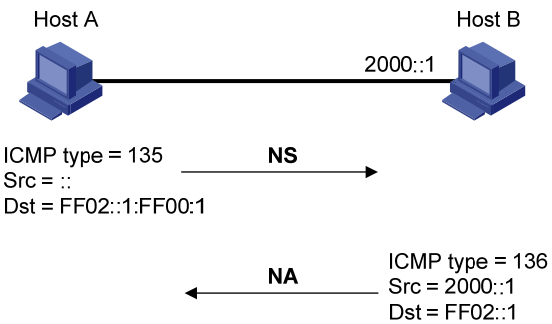
(1) 节点发送 NS 消息，其中目的地址是邻居节点的 IPv6 地址。

(2) 如果收到邻居节点的确认报文，则认为邻居可达；否则，认为邻居不可达。

3. 重复地址检测

当节点获取到一个IPv6 地址后，需要使用重复地址检测功能确定该地址是否已被其他节点使用（与IPv4 的免费ARP功能相似）。通过NS和NA可以实现重复地址检测，如 图 1-4 所示。

图1-4 重复地址检测示意图



(1) 节点 A 发送 NS 消息，NS 消息的源地址是未指定地址::，目的地址是待检测的 IPv6 地址对应的被请求节点组播地址，消息内容中包含了待检测的 IPv6 地址。

(2) 如果节点 B 已经使用这个 IPv6 地址，则会返回 NA 消息。其中包含了自己的 IPv6 地址。

(3) 节点 A 收到节点 B 发来的 NA 消息，就知道该 IPv6 地址已被使用。反之，则说明该地址未被使用，节点 A 就可使用此 IPv6 地址。

4. 路由器发现/前缀发现及地址自动配置

路由器发现/前缀发现是指节点从收到的 RA 消息中获取邻居路由器及所在网络的前缀，以及其他配置参数。

地址无状态自动配置是指节点根据路由器发现/前缀发现所获取的信息，自动配置 IPv6 地址。

路由器发现/前缀发现通过路由器请求消息 RS 和路由器通告消息 RA 来实现，具体过程如下：

(1) 节点启动时，通过 RS 消息向路由器发出请求，请求前缀和其他配置信息，以便用于节点的配置。

(2) 路由器返回 RA 消息，其中包括前缀信息选项（路由器也会周期性地发布 RA 消息）。

(3) 节点利用路由器返回的 RA 消息中的地址前缀及其他配置参数，自动配置接口的 IPv6 地址及其他信息。

说明

- 前缀信息选项中不仅包括地址前缀的信息，还包括该地址前缀的首选生命期（preferred lifetime）和有效生命期（valid lifetime）。节点收到周期性发送的 RA 消息后，会根据该消息更新前缀的首选生命期和有效生命期。
- 在有效生命期内，自动生成的地址可以正常使用；有效生命期过期后，自动生成的地址将被删除。

---

5. 重定向功能

当主机启动时，它的路由表中可能只有一条到缺省网关的缺省路由。当满足一定的条件时，缺省网关会向源主机发送 ICMPv6 重定向消息，通知主机选择更好的下一跳进行后续报文的发送（与 IPv4 的 ICMP 重定向消息的功能相同）。

设备在满足下列条件时会发送对主机重定向的 ICMPv6 重定向报文：

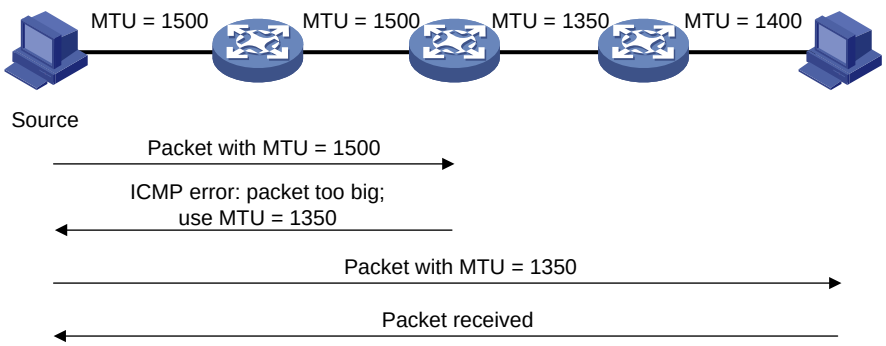
- 接收和转发数据报文的接口是同一接口；
- 被选择的路由本身没有被 ICMPv6 重定向报文创建或修改过；
- 被选择的路由不是缺省路由。

1.1.4 IPv6 PMTU发现

报文从源端到目的端的传输路径中所经过的链路可能具有不同的 MTU。在 IPv6 中，当报文的长度大于链路的 MTU 时，报文的分片将在源端进行，从而减轻中间转发设备的处理压力，合理利用网络资源。

PMTU（Path MTU，路径MTU）发现机制的目的就是要找到从源端到目的端的路径上最小的MTU。PMTU的工作过程如 图 1-5 所示。

图1-5 PMTU 发现工作过程



(1) 源端主机按照自己的 MTU 对报文进行分片，之后向目的主机发送报文。

(2) 中间转发设备接收到该报文进行转发时，如果发现转发报文的接口支持的 MTU 值小于报文长度，则会丢弃报文，并给源端返回一个 ICMPv6 差错报文，其中包含了转发失败的接口的 MTU。

(3) 源主机收到该差错报文后，将按照报文中所携带的 MTU 重新对报文进行分片并发送。

(4) 如此反复，直到目的端主机收到这个报文，从而确定报文从源端到目的端路径中的最小 MTU。

1.1.5 IPv6 过渡技术介绍

在 IPv6 成为主流协议之前，首先使用 IPv6 协议栈的网络希望能与当前仍被 IPv4 支撑着的 Internet 进行正常通信，因此必须开发出 IPv4 和 IPv6 互通技术以保证 IPv4 能够平稳过渡到 IPv6。互通技术应该对信息传递做到高效无缝。目前已经出现了多种过渡技术，这些技术各有特点，用于解决不同过渡时期、不同环境的通信问题。

目前解决过渡问题的基本技术主要有 2 种：双协议栈（RFC 2893）和隧道技术（RFC 2893）。

1. 双协议栈

双协议栈是一种最简单直接的过渡机制。同时支持 IPv4 协议和 IPv6 协议的网络节点称为双协议栈节点。当双协议栈节点配置 IPv4 地址和 IPv6 地址后，就可以在相应接口上转发 IPv4 和 IPv6 报文。当一个上层应用同时支持 IPv4 和 IPv6 协议时，根据协议要求可以选用 TCP 或 UDP 作为传输层的协议，但在选择网络层协议时，它会优先选择 IPv6 协议栈。双协议栈技术适合 IPv4 网络节点之间或者 IPv6 网络节点之间通信，是所有过渡技术的基础。但是，这种技术要求运行双协议栈的节点有一个全球唯一的地址，实际上没有解决 IPv4 地址资源匮乏的问题。

2. 隧道技术

隧道是一种封装技术，它利用一种网络协议来传输另一种网络协议，即利用一种网络传输协议，将其他协议产生的数据报文封装在它自己的报文中，然后在网络中传输。关于隧道技术的详细介绍，请参见“三层技术-IP 业务配置指导”中的“隧道”。

1.1.6 协议规范

与 IPv6 基础相关的协议规范有：

- RFC 1881：IPv6 Address Allocation Management
- RFC 1887：An Architecture for IPv6 Unicast Address Allocation

- RFC 1981：Path MTU Discovery for IP version 6
- RFC 2375：IPv6 Multicast Address Assignments
- RFC 2460：Internet Protocol, Version 6 (IPv6) Specification
- RFC 2464：Transmission of IPv6 Packets over Ethernet Networks
- RFC 2526：Reserved IPv6 Subnet Anycast Addresses
- RFC 2894：Router Renumbering for IPv6
- RFC 3307：Allocation Guidelines for IPv6 Multicast Addresses
- RFC 3513：Internet Protocol Version 6 (IPv6) Addressing Architecture
- RFC 4191：Default Router Preferences and More-Specific Routes
- RFC 4443：Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification.
- RFC 4861：Neighbor Discovery for IP Version 6 (IPv6)
- RFC 4862：IPv6 Stateless Address Autoconfiguration

1.2 IPv6基础配置任务简介

表1-4 IPv6 基础配置任务简介

配置任务		说明	详细配置
配置IPv6基本功能	使能IPv6报文转发功能	必选	1.3.1
	配置IPv6全球单播地址	三者必选其一	1.3.2
	配置IPv6链路本地地址		1.3.3
	配置IPv6任播地址		1.3.4
配置IPv6邻居发现协议	配置静态邻居表项	可选	1.4.1
	配置接口上允许动态学习的邻居的最大个数	可选	1.4.2
	配置STALE状态ND表项的老化时间	可选	1.4.3
	配置RA消息的相关参数	可选	1.4.4
	配置重复地址检测时发送邻居请求消息的次数	可选	1.4.5
	配置ND Snooping功能	可选	1.4.6
	配置ND Proxy功能	可选	1.4.7
配置PMTU发现	配置指定地址的静态PMTU	可选	1.5.1
	配置PMTU老化时间	可选	1.5.2
配置TCP6		可选	1.6
配置ICMPv6报文发送	配置指定时间内发送ICMPv6差错报文的最大个数	可选	1.7.1
	配置允许回复组播形式的Echo request报文	可选	1.7.2
	配置ICMPv6超时差错报文发送功能	可选	1.7.3

配置任务		说明	详细配置
	配置ICMPv6目的不可达差错报文发送功能	可选	1.7.4
配置组播ND		可选	1.8

1.3 配置IPv6基本功能

1.3.1 使能IPv6 报文转发功能

在进行 IPv6 的相关配置以前，必须先使能 IPv6 报文转发功能。否则即使在接口上配置了 IPv6 地址，仍无法转发 IPv6 的报文，造成 IPv6 网络无法互通。

表1-5 使能 IPv6 报文转发功能

操作	命令	说明
进入系统视图	**system-view**	-
使能IPv6报文转发功能	**ipv6**	必选 缺省情况下，IPv6报文转发功能处于关闭状态

1.3.2 配置IPv6 全球单播地址

IPv6 全球单播地址可以通过下面三种方式配置：

- 采用 EUI-64 格式形成：当配置采用 EUI-64 格式形成 IPv6 地址时，接口的 IPv6 地址的前缀需要手工配置，而接口标识符则由接口自动生成。
- 手工配置：用户手工配置 IPv6 全球单播地址。
- 无状态自动配置：根据接收到的 RA 报文中携带的地址前缀信息，自动生成 IPv6 全球单播地址。

说明

- 每个接口可以有多个前缀不同的全球单播地址。
- 手工配置的全球单播地址的优先级高于自动生成的全球单播地址。如果在接口已经自动生成全球单播地址的情况下，手工配置前缀相同的全球单播地址，自动生成的地址将被覆盖。此后，即使删除手工配置的全球单播地址，已被覆盖的自动生成的全球单播地址也不会恢复。再次接收到 RA 报文后，设备根据报文携带的地址前缀信息，重新生成全球单播地址。

1. 采用EUI-64 格式形成IPv6 地址

表1-6 采用 EUI-64 格式形成 IPv6 地址

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*	-
采用EUI-64格式形成IPv6地址	**ipv6 address** *ipv6-address/prefix-length* **eui-64**	必选 缺省情况下，接口上没有配置全球单播地址

2. 手工指定IPv6 地址

表1-7 手工指定 IPv6 地址

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*	-
手工指定IPv6地址	**ipv6 address** { *ipv6-address prefix-length* | *ipv6-address/prefix-length* }	必选 缺省情况下，接口上没有配置全球单播地址

3. 无状态自动配置IPv6 地址

表1-8 无状态自动配置 IPv6 地址

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*	-
无状态自动配置IPv6地址	**ipv6 address auto**	必选 缺省情况下，接口上没有配置全球单播地址

说明

在接口上执行 **undo ipv6 address auto** 命令，将删除该接口上所有自动生成的全球单播地址。

在配置了无状态自动配置 IPv6 地址功能后，接口会根据接收到的 RA 报文中携带的地址前缀信息和接口 ID，自动生成 IPv6 全球单播地址。如果接口是 IEEE 802 类型的接口（例如，VLAN 接口），其接口 ID 是由 MAC 地址根据一定的规则生成，此接口 ID 具有全球唯一性。对于不同的前缀，接口 ID 部分始终不变，攻击者通过接口 ID 可以很方便的识别出通信流量是由哪台设备产生的，并分析其规律，从而窥探到设备和谁进行通信、在什么时间进行通信，会造成一定的安全隐患。

如果在地址无状态自动配置时，自动生成接口 ID 不断变化的 IPv6 地址，就可以加大攻击的难度，从而保护网络。为此，设备提供了临时地址功能，使得系统可以生成临时地址，并优先选择临时地

址作为接口发送报文的源地址。配置该功能后，通过地址无状态自动配置，IEEE 802 类型的接口可以同时生成两类地址：

- 公共地址：地址前缀采用 RA 报文携带的前缀，接口 ID 由 MAC 地址产生。接口 ID 始终不变。
- 临时地址：地址前缀采用 RA 报文携带的前缀，接口 ID 由系统根据 MD5 算法计算产生。接口 ID 不断变化。

发送报文时，系统将优先选择临时地址作为报文的源地址。当临时地址的有效生命期过期后，这个临时地址将被删除，同时，系统会通过 MD5 算法重新生成一个接口 ID 不同的临时地址。所以，该接口发送报文的源地址的接口 ID 总是在不停变化。如果生成的临时地址因为 DAD 冲突不可用，就采用公共地址作为报文的源地址。

临时地址的首选生命期和有效生命期的确定原则如下：

- 首选生命期是如下两个值之中的较小者：RA 前缀中的首选生命期和（配置的临时地址首选生命期减去 DESYNC_FACTOR）。DESYNC_FACTOR 是一个 0～600 秒的随机值。
- 有效生命期是如下两个值之中的较小者：RA 前缀中的有效生命期和配置的临时地址有效生命期。

表1-9 配置系统生成临时地址，并优先选择临时地址作为报文的源地址

操作	命令	说明
进入系统视图	**system-view**	-
配置系统生成临时地址，并优先选择临时地址作为报文的源地址	**ipv6 prefer temporary-address** [*valid-lifetime preferred-lifetime*]	必选 缺省情况下，系统不生成临时地址，也就不会用临时地址作为接口发送报文的源地址

注意

- 设备的接口必须启用地址无状态自动配置功能才能生成临时地址，而且临时地址不会覆盖公共地址，因此会出现一个接口下有多个前缀相同但是接口 ID 不同的地址。
- 如果公共地址生成失败，例如前缀冲突，则不会生成临时地址。

1.3.3 配置IPv6 链路本地地址

IPv6 的链路本地地址可以通过两种方式获得：

- 自动生成：设备根据链路本地地址前缀（FE80::/10）及接口的链路层地址，自动为接口生成链路本地地址；
- 手工指定：用户手工配置 IPv6 链路本地地址。

说明

- 每个接口只能有一个链路本地地址，为了避免链路本地地址冲突，推荐使用链路本地地址的自动生成方式。
- 配置链路本地地址时，手工指定方式的优先级高于自动生成方式。即如果先采用自动生成方式，之后手工指定，则手工指定的地址会覆盖自动生成的地址；如果先手工指定，之后采用自动生成的方式，则自动配置不生效，接口的链路本地地址仍是手工指定的。此时，如果删除手工指定的地址，则自动生成的链路本地地址会生效。

表1-10 配置自动生成链路本地地址

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*	-
配置自动生成链路本地地址	**ipv6 address auto link-local**	可选 缺省情况下，接口上没有链路本地地址。当接口配置了IPv6全球单播地址后，会自动生成链路本地地址

表1-11 手工指定接口的链路本地地址

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*	-
手工指定接口的链路本地地址	**ipv6 address** *ipv6-address* **link-local**	可选 缺省情况下，接口上没有链路本地地址。当接口配置了IPv6全球单播地址后，会自动生成链路本地地址

说明

- 当接口配置了 IPv6 全球单播地址后，同时会自动生成链路本地地址。且与采用 **ipv6 address auto link-local** 命令生成的链路本地地址相同。此时如果手工指定接口的链路本地地址，则手工指定的有效。如果删除手工指定的链路本地地址，则接口的链路本地地址恢复为系统自动生成的地址。
- **undo ipv6 address auto link-local** 命令只能删除使用 **ipv6 address auto link-local** 命令生成的链路本地地址。即如果此时已经配置了 IPv6 全球单播地址，由于系统会自动生成链路本地地址，则接口仍有链路本地地址；如果此时没有配置 IPv6 全球单播地址，则接口没有链路本地地址。

1.3.4 配置IPv6 任播地址

用户需要手工配置接口的 IPv6 任播地址。

表1-12 配置 IPv6 任播地址

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IPv6任播地址	**ipv6 address** *ipv6-address/prefix-length* **anycast**	必选 缺省情况下，接口上没有配置任播地址

1.4 配置IPv6邻居发现协议

1.4.1 配置静态邻居表项

将邻居节点的 IPv6 地址解析为链路层地址，可以通过邻居请求消息 NS 及邻居通告消息 NA 来动态实现，也可以通过手工配置静态邻居表项来实现。

设备根据邻居节点的 IPv6 地址和与此邻居节点相连的三层接口号来唯一标识一个静态邻居表项。目前，静态邻居表项有两种配置方式：

- 配置本节点的三层接口对应的邻居节点的 IPv6 地址、链路层地址；
- 配置本节点 VLAN 中的端口对应的邻居节点的 IPv6 地址、链路层地址。

表1-13 配置静态邻居表项

操作	命令	说明
进入系统视图	**system-view**	-
配置静态邻居表项	**ipv6 neighbor** *ipv6-address mac-address* { *vlan-id port-type port-number* | **interface** *interface-type interface-number* } [**vpn-instance** *vpn-instance-name*]	必选

注意

对于 VLAN 接口，可以采用上述两种方式来配置静态邻居表项：

- 采用第一种方式配置静态邻居表项后，设备还需要解析 VLAN 对应的二层端口信息。
- 采用第二种方式配置静态邻居表项，需要保证 VLAN 所对应的 VLAN 接口已经存在，且 *port-type port-number* 指定的二层端口属于 *vlan-id* 指定的 VLAN。在配置后，设备会将 VLAN 所对应的 VLAN 接口与 IPv6 地址相对应来唯一标识一个静态邻居表项。

1.4.2 配置接口上允许动态学习的邻居的最大个数

设备可以通过 NS 消息和 NA 消息来动态获取邻居节点的链路层地址，并将其加入到邻居表中。如果动态获取的邻居表过大，将可能导致设备的转发性能下降。为此，可以通过设置接口上允许动态学习的邻居的最大个数来进行限制。当接口上动态学习的邻居个数达到所设置的最大值时，该接口将不再学习邻居信息。

表1-14 配置接口上允许动态学习的邻居的最大个数

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*	-
配置接口上允许动态学习的邻居的最大个数	**ipv6 neighbors max-learning-num** *number*	可选 缺省情况下，二层接口不对允许动态学习的邻居的最大个数进行限制，对于S5500-28SC-HI和S5500-52SC-HI交换机的三层接口允许动态学习的邻居的最大个数为4096；对于本系列交换机其它机型的三层接口允许动态学习的邻居的最大个数为8192

1.4.3 配置STALE状态ND表项的老化时间

为适应网络的变化，ND 表需要不断更新。ND 表中的 STALE 状态 ND 表项并非永远有效，每一条记录都有一个老化时间。到达老化时间的 STALE 状态 ND 表项将迁移到 DELAY 状态。5 秒钟后 DELAY 状态超时，ND 表项将迁移到 PROBE 状态，并发送 3 次 NS 报文进行可达性探测。若邻居已经下线，则收不到回应的 NA 报文，此时会将该 ND 表项删除。

缺省情况下，STALE 状态 ND 表项的老化时间为 4 小时。用户可以根据网络实际情况调整老化时间。

表1-15 配置 STALE 状态 ND 表项的老化时间

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
配置STALE状态ND表项的老化时间	**ipv6 neighbor stale-aging** *aging-time*	可选 缺省情况下，STALE状态ND表项的老化时间为4小时

1.4.4 配置RA消息的相关参数

用户可以根据实际情况，配置接口是否发送RA消息及发送RA消息的时间间隔，同时可以配置RA消息中的相关参数以通告给主机。当主机接收到RA消息后，就可以采用这些参数进行相应操作。可以配置的RA消息中的参数及含义如 表 1-16 所示。

表1-16 RA 消息中的参数及描述

参数	描述
跳数限制（Cur Hop Limit）	主机在发送IPv6报文时，将使用该参数值填充IPv6报文头中的Hop Limit字段。同时该参数值也作为设备应答报文中的Hop Limit字段值。
前缀信息（Prefix Information）	在同一链路上的主机收到设备发布的前缀信息后，可以进行无状态自动配置等操作。
MTU	发布链路的MTU，可以用于确保同一链路上的所有节点采用相同的MTU值。
被管理地址配置标志位（M flag）	用于确定主机是否采用有状态自动配置获取IPv6地址。 如果设置该标志位为1，主机将通过有状态自动配置（例如DHCP服务器）来获取IPv6地址；否则，将通过无状态自动配置获取IPv6地址，即根据自己的链路层地址及路由器发布的前缀信息生成IPv6地址。
其他信息配置标志位（O flag）	用于确定主机是否采用有状态自动配置获取除IPv6地址外的其他信息。 如果设置其他信息配置标志位为1，主机将通过有状态自动配置（例如DHCP服务器）来获取除IPv6地址外的其他信息；否则，将通过无状态自动配置获取其他信息。
路由器生存时间（Router Lifetime）	用于设置发布RA消息的路由器作为主机的默认路由器的时间。主机根据接收到的RA消息中的路由器生存时间参数值，就可以确定是否将发布该RA消息的路由器作为默认路由器。
邻居请求消息重传时间间隔（Retrans Timer）	设备发送NS消息后，如果未在指定的时间间隔内收到响应，则会重新发送NS消息。
保持邻居可达状态的时间（Reachable Time）	当通过邻居可达性检测确认邻居可达后，在所设置的可达时间内，设备认为邻居可达；超过设置的时间后，如果需要向邻居发送报文，会重新确认邻居是否可达。

表1-17 配置允许发布 RA 消息

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
取消对RA消息发布的抑制	**undo ipv6 nd ra halt**	必选 缺省情况下，抑制发布RA消息
配置RA消息发布的最大时间间隔和最小时间间隔	**ipv6 nd ra interval** *max-interval-value min-interval-value*	可选 缺省情况下，RA消息发布的最大间隔时间为600秒，最小时间间隔为200秒 RA消息周期性发布时，相邻两次的时间间隔是在最大时间间隔与最小时间间隔之间随机选取一个值作为周期性发布RA消息的时间间隔 配置的最小时间间隔应该小于等于最大时间间隔的0.75倍

表1-18 配置 RA 消息中的相关参数

操作	命令	说明
进入系统视图	**system-view**	-
配置跳数限制	**ipv6 nd hop-limit** *value*	可选 缺省情况下，路由器发布的跳数限制为64跳
进入接口视图	**interface** *interface-type interface-number*	-
配置RA消息中的前缀信息	**ipv6 nd ra prefix** { *ipv6-prefix prefix-length* | *ipv6-prefixIprefix-length* } *valid-lifetime preferred-lifetime* [**no-autoconfig** | **off-link**] *	可选 缺省情况下，没有配置RA消息中的前缀信息，此时将使用发送RA消息的接口IPv6地址作为RA消息中的前缀信息，其有效生命期是2592000秒（30天），首选生命期是604800（7天）
配置RA消息中不携带MTU选项	**ipv6 nd ra no-advlinkmtu**	可选 缺省情况下，RA消息中携带MTU选项
设置被管理地址配置标志位为1	**ipv6 nd autoconfig managed-address-flag**	可选 缺省情况下，被管理地址标志位为0，即主机通过无状态自动配置获取IPv6地址
设置其他配置标志位为1	**ipv6 nd autoconfig other-flag**	可选 缺省情况下，其他配置标志位为0，即主机通过无状态自动配置获取其他信息
配置RA消息中路由器的生存时间	**ipv6 nd ra router-lifetime** *value*	可选 缺省情况下，RA消息中路由器的生存时间为1800秒
配置邻居请求消息重传时间间隔	**ipv6 nd ns retrans-timer** *value*	可选 缺省情况下，接口发送NS消息的时间间隔为1000毫秒；接口发布的RA消息中Retrans Timer字段的值为0，即不对主机进行指定

操作	命令	说明
配置保持邻居可达状态的时间	**ipv6 nd nud reachable-time** *value*	可选 缺省情况下，接口保持邻居可达状态的时间为30000毫秒；接口发布的RA消息中Reachable Timer字段的值为0，即不对主机进行指定

说明

- RA 消息发布的最大间隔时间应该小于或等于 RA 消息中路由器的生存时间，以保证在路由器失效之前得到更新的 RA 消息。
- 在接口上配置的邻居请求消息重传时间间隔及保持邻居可达状态的时间，既可作为 RA 消息中的信息发布给主机，也可作为本接口发送邻居请求消息的时间间隔及保持邻居可达状态的时间。

1.4.5 配置重复地址检测时发送邻居请求消息的次数

接口获得 IPv6 地址后，将发送邻居请求消息进行重复地址检测，如果在指定的时间内（通过 **ipv6 nd ns retrans-timer** 命令配置）没有收到响应，则继续发送邻居请求消息，当发送的次数达到所设置的次数后，仍未收到响应，则认为该地址可用。

表1-19 配置重复地址检测时发送邻居请求消息的次数

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*	-
配置重复地址检测时发送邻居请求消息的次数	**ipv6 nd dad attempts** *value*	可选 缺省情况下，重复地址检测时发送邻居请求报文的次数为1，当*value*值为0时，表示禁止重复地址检测

1.4.6 配置ND Snooping功能

1. ND Snooping功能简介

ND Snooping 功能用于二层交换网络环境，通过侦听 DAD（Duplicate Address Detection，重复地址检测） NS 消息来建立 ND Snooping 表项，表项内容包括报文的源 IPv6 地址、源 MAC 地址、所属 VLAN、入端口等信息。

目前，ND Snooping 表项有以下用途：

- 与 ND Detection 功能配合使用。关于 ND Detection 的详细介绍，请参见“安全配置指导”中的“ND 攻击防御”。
- 与 IP Source Guard 功能配合使用。关于 IP Source Guard 的详细介绍，请参见“安全配置指导”中的“IP Source Guard”。

- 在 SAVI 各场景中使用。关于 SAVI 的详细介绍，请参见“安全配置指导”中的“SAVI”。

当一个 VLAN 使能 ND Snooping 后，该 VLAN 内所有端口接收的 ND 报文均会被重定向到 CPU。全局使能 ND Snooping 后，CPU 会对这些 ND 报文进行分析，获取报文的源 IPv6 地址、源 MAC 地址、源 VLAN 和入端口信息，并根据这些信息来新建或更新 ND Snooping 表项。

下面将具体介绍 ND Snooping 表项的新建、更新和老化机制。

(1) 新建表项机制

设备只会根据收到的 DAD NS 消息来新建 ND Snooping 表项。

(2) 更新表项机制

设备收到一个 ND 报文后，首先找到与该报文源 IPv6 地址对应的表项。如果表项的刷新时间没有超过 1 秒，则设备不会对表项进行更新。如果表项的刷新时间已经超过 1 秒，设备要判断收到的 ND 报文的 MAC 地址、入端口信息与现有该表项的 MAC 地址、入端口信息是否一致。

- 如果一致，则更新该表项的老化时间；
- 如果不一致，但收到的是 DAD NS 消息，则忽略该消息；
- 如果不一致，并且收到的是除 DAD NS 以外的 ND 报文，则进行主动确认。

主动确认过程如下：

- 首先，设备将探测现有该表项的正确性。设备对外发送 DAD NS 消息（该消息仅发送一次），消息中的待检测 IPv6 地址就是 ND Snooping 表项中的 IPv6 地址。如果在发送过程中收到对应的 NA 消息（消息的源 IPv6 地址、源 MAC 地址、入端口、源 VLAN 信息均和现有表项信息一致），则更新该表项的老化时间；如果在发送 DAD NS 消息后的 1 秒内都没有收到对应的 NA 消息，则开始探测新收到报文的真实性。
- 探测新收到报文（报文 A）真实性的过程如下：设备对外发送 DAD NS 消息（该消息仅发送一次），消息中的待检测 IPv6 地址就是报文 A 的源 IPv6 地址。如果在发送过程中收到对应的 NA 消息（消息的源 IPv6 地址、源 MAC 地址、入端口、源 VLAN 信息均和报文 A 一致），则更新该表项的老化时间；如果在发送 DAD NS 消息后的 1 秒内都没有收到对应的 NA 消息，则该表项不会被更新。

(3) 老化表项机制

ND Snooping 表项的老化时间为 25 分钟。如果一个 ND Snooping 表项自最后一次更新后 15 分钟内没有收到 ND 更新报文，则开始进行主动确认：设备对外发送 DAD NS 消息（该消息仅发送一次），消息中的待检测 IPv6 地址就是 ND Snooping 表项中的 IPv6 地址。

- 如果在发送过程中收到对应的 NA 消息（消息的源 IPv6 地址、源 MAC 地址、入端口、源 VLAN 信息均和现有表项信息一致），则更新表项的老化时间；
- 如果 1 秒内没有收到对应的 NA 消息，则等到表项的老化时间过后删除该表项。

2. 配置ND Snooping功能

表1-20 配置 ND Snooping 功能

操作	命令	说明
进入系统视图	**system-view**	-
使能全球单播类型地址的ND Snooping功能，根据IPv6地址为全球单播地址的DAD NS报文建立ND Snooping表项	**ipv6 nd snooping enable global**	二者至少选其一 缺省情况下，全球单播类型地址和链路本地类型地址的ND Snooping功

操作	命令	说明
使能链路本地类型地址的ND Snooping功能，根据IPv6地址为链路本地地址的DAD NS报文建立ND Snooping表项	**ipv6 nd snooping enable link-local**	能均处于关闭状态
进入VLAN视图	**vlan** *vlan-id*	-
使能ND Snooping功能	**ipv6 nd snooping enable**	必选 缺省情况下，ND Snooping功能处于关闭状态
退回系统视图	**quit**	-
进入二层以太网端口视图/二层聚合接口视图	**interface** *interface-type interface-number*	-
配置接口允许学习ND Snooping表项的最大个数	**ipv6 nd snooping max-learning-num** *number*	可选 缺省情况下，不对接口允许学习ND Snooping表项的最大个数进行限制
配置接口为上行口，禁止接口学习ND Snooping表项	**ipv6 nd snooping uplink**	可选 缺省情况下，ND Snooping功能使能后，接口上允许学习ND Snooping表项

1.4.7 配置ND Proxy功能

说明

目前只支持对 NS 及 NA 报文的 ND Proxy 功能。

1. ND Proxy功能简介

如果 NS 请求是从一个网络的主机发往同一网段却不在同一物理网络上的另一台主机，那么连接它们的具有 ND Proxy 功能的设备就可以通过三层转发将这个 NS 请求发往被请求的另一台主机，该主机回应的 NA 报文也会通过连接它们的设备经过三层转发转给发起请求的主机，实现两台主机的互通，这个过程称作 ND Proxy。

ND Proxy 功能根据应用场景不同分为普通 ND Proxy 和本地 ND Proxy。

说明

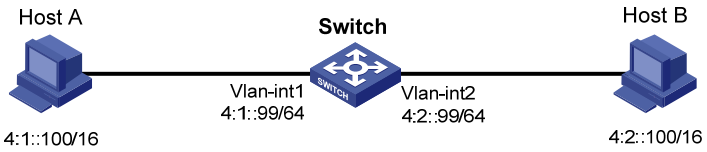
如无特殊说明，本章后续描述中的 ND Proxy 均指普通 ND Proxy。

(1) ND Proxy

ND Proxy的典型应用环境如 图 1-6 所示。设备Switch通过两个三层接口Vlan-interface1 和Vlan-interface2 连接两个网络，两个三层接口的IPv6 地址不在同一个网段，接口地址分别为

4:1::99/64、4:2::99/64。但是两个网络内的主机Host A和Host B的地址通过掩码的控制，既与相连设备的接口地址在同一网段，同时二者也处于同一个网段。

图1-6 ND Proxy 的应用环境



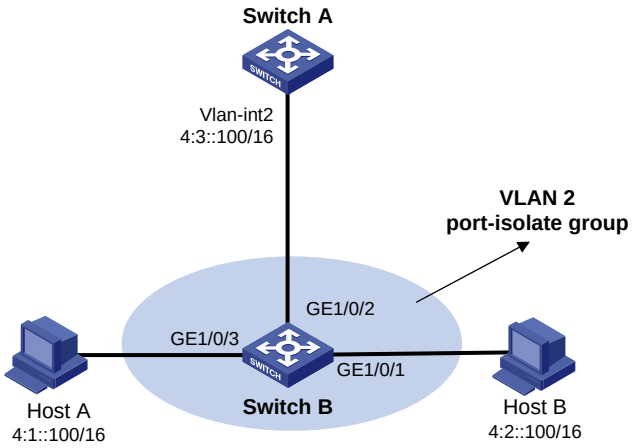
在这种组网情况下，当 Host A 需要与 Host B 通信时，由于目的 IPv6 地址与本机的 IPv6 地址为同一网段，因此 Host A 会直接发出请求 Host B 硬件地址的 NS 请求。但是，此时的两台主机处于不同的广播域中，Host B 无法收到 Host A 的 NS 请求报文，当然也就无法应答。

通过在 Switch 上启用 ND Proxy 功能，可以解决此问题。在接口 Vlan-interface1 和 Vlan-interface2 上启用 ND Proxy 后，Switch 可以根据 NS 请求的 IPv6 地址查找转发表项获得出接口，并将 NS 请求报文转发出去，这样 Host B 可以收到请求自己的 NS 报文，并回应 NA；NA 报文也会通过启用了 ND Proxy 的设备将 NA 报文转发给 Host A；这样，实现 Host A 与 Host B 之间的通信。

(2) 本地 ND Proxy

本地ND Proxy的应用场景如 图 1-7 所示。Host A和Host B属于同一个VLAN 2，但它们分别连接到被二层隔离的端口GigabitEthernet1/0/3 和GigabitEthernet1/0/1 上。

图1-7 本地 ND Proxy 的应用环境



在这种组网情况下，当 Host A 需要与 Host B 通信时，由于目的 IPv6 地址与本机的 IPv6 地址为同一网段，因此 Host A 会直接发出请求 Host B 硬件地址的 NS 请求。但是，因为连接两台主机的端口处于端口隔离状态，Host B 无法收到 Host A 的 NS 请求报文。

通过在 Switch A 上启用本地 ND Proxy 功能，可以解决此问题。在接口 Vlan-interface2 上启用本地 ND Proxy 后，Switch A 会代替 Host B 回应 NA，Host A 发给 Host B 的报文就会通过 Switch A 进行转发，从而实现 Host A 与 Host B 之间的通信。

本地 ND Proxy 可以在下列三种情况下实现主机之间的三层互通：

- 想要互通的主机分别连接到同一个 VLAN 中的不同二层隔离端口下；
- 使能 Super VLAN 功能后，想要互通的主机属于不同的 Sub VLAN；

- 使能 Isolate-user-vlan 功能后，想要互通的主机属于不同的 Secondary VLAN。

2. 配置ND Proxy功能

ND Proxy 和本地 ND Proxy 功能均可在 VLAN 接口视图/三层以太网端口下进行配置。

表1-21 配置 ND Proxy 功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
开启ND Proxy功能	**proxy-nd enable**	必选 缺省情况下，ND Proxy功能处于关闭状态

表1-22 配置本地 ND Proxy 功能

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
开启本地ND Proxy功能	**local-proxy-nd enable**	必选 缺省情况下，本地ND Proxy功能处于关闭状态

1.5 配置PMTU发现

1.5.1 配置指定地址的静态PMTU

用户可以为指定的目的 IPv6 地址配置静态的 PMTU 值。当源端主机从接口发送报文时，将比较该接口的 MTU 与指定目的 IPv6 地址的静态 PMTU，如果报文长度大于二者中的最小值，则采用此最小值对报文进行分片。

表1-23 配置指定地址的静态 PMTU

操作	命令	说明
进入系统视图	**system-view**	-
配置指定IPv6地址对应的静态PMTU值	**ipv6 pathmtu** [**vpn-instance** *vpn-instance-name*] *ipv6-address* [*value*]	必选 缺省情况下，没有配置静态PMTU值

1.5.2 配置PMTU老化时间

通过“1.1.4 IPv6 PMTU发现”中的方法动态确定源端主机到目的端主机的PMTU后，源端主机将使用这个MTU值发送后续报文到目的端主机。当PMTU老化时间超时后，动态确定的PMTU值将会被删除，源端主机会通过PMTU机制重新确定发送报文的MTU值。

该配置对静态 PMTU 不起作用。

表1-24 配置 PMTU 老化时间

操作	命令	说明
进入系统视图	**system-view**	-
配置PMTU老化时间	**ipv6 pathmtu age** *age-time*	可选 缺省情况下，PMTU的老化时间是10分钟

1.6 配置TCP6

可以配置的 TCP6 属性包括：

- synwait 定时器：当发送 SYN 报文时，TCP6 启动 synwait 定时器，如果 synwait 定时器超时前未收到回应报文，则 TCP6 连接建立不成功。
- finwait 定时器：当 TCP6 的连接状态为 FIN_WAIT_2 时，启动 finwait 定时器，如果在定时器超时前没有收到报文，则 TCP6 连接终止；如果收到 FIN 报文，则 TCP6 连接状态变为 TIME_WAIT 状态；如果收到非 FIN 报文，则从收到的最后一个非 FIN 报文开始重新计时，在超时后中止连接。
- TCP6 的接收和发送缓冲区的大小。

表1-25 配置 TCP6

操作	命令	说明
进入系统视图	**system-view**	-
配置TCP6的synwait定时器	**tcp ipv6 timer syn-timeout** *wait-time*	可选 缺省情况下，synwait定时器的值75秒
配置TCP6的finwait定时器	**tcp ipv6 timer fin-timeout** *wait-time*	可选 缺省情况下，finwait定时器的值为675秒
配置TCP6的接收和发送缓冲区大小	**tcp ipv6 window** *size*	可选 缺省情况下，TCP6的接收和发送缓冲区大小均为8KB

1.7 配置ICMPv6报文发送

1.7.1 配置指定时间内发送ICMPv6 差错报文的最大个数

如果网络中短时间内发送的 ICMPv6 差错报文过多，将可能导致网络拥塞。为了避免这种情况，用户可以控制在指定时间内发送 ICMPv6 差错报文的最大个数，目前采用令牌桶算法来实现。

用户可以设置令牌桶的容量，即令牌桶中可以同时容纳的令牌数；同时可以设置令牌桶的刷新周期，即每隔多长时间将令牌桶内的令牌个数刷新为所配置的容量。一个令牌表示允许发送一个 ICMPv6 差错报文，每当发送一个 ICMPv6 差错报文，则令牌桶中减少一个令牌。如果连续发送的 ICMPv6

差错报文超过了令牌桶的容量，则后续的 ICMPv6 差错报文将不能被发送出去，直到按照所设置的刷新频率将新的令牌放入令牌桶中。

表1-26 配置指定时间内发送 ICMPv6 差错报文的最大个数

操作	命令	说明
进入系统视图	**system-view**	-
配置控制ICMPv6差错报文发送的令牌桶容量和刷新周期	**ipv6 icmp-error** { **bucket** *bucket-size* | **ratelimit** *interval* } *	可选 缺省情况下，令牌桶容量为10，令牌桶的刷新周期为100毫秒，即每一个刷新周期内最多可以发送10个ICMPv6差错报文 刷新周期为0时，表示不限制ICMPv6差错报文的发送

1.7.2 配置允许回复组播形式的Echo request报文

如果允许主机回复组播形式的 Echo request 报文，则主机 A 可以构造目的地址为组播地址、源地址为主机 B 的 Echo request 报文，使该组播组中所有的主机都向主机 B 发送 Echo reply 报文，从而达到攻击主机 B 的目的。因此，为了避免主机利用设备达到攻击的目的，缺省情况下，不允许设备回复组播形式的 Echo request 报文。

在某些应用场景下，可能需要使用组播形式的 Echo request 报文来获取信息，此时可以通过下面的命令，配置允许设备回复组播形式的 Echo request 报文。

表1-27 配置允许回复组播形式的 Echo request 报文

操作	命令	说明
进入系统视图	**system-view**	-
配置允许回复组播形式的Echo request报文	**ipv6 icmpv6 multicast-echo-reply enable**	必选 缺省情况下，不允许回复组播形式的Echo request报文

1.7.3 配置ICMPv6 超时差错报文发送功能

ICMPv6 超时报文发送功能是在设备收到 IPv6 数据报文后，如果发生超时差错，则将报文丢弃并给源端发送 ICMPv6 超时差错报文。

设备在满足下列条件时会发送 ICMPv6 超时报文：

- 设备收到 IPv6 数据报文后，如果报文的目的地不是本地且报文的 Hop limit 字段是 1，则发送“Hop limit 超时”ICMPv6 差错报文；
- 设备收到目的地址为本地的 IPv6 数据报文的第一个分片后，启动定时器，如果所有分片报文到达之前定时器超时，则会发送“重组超时”ICMPv6 差错报文。

设备接收到大量需要发送 ICMPv6 差错报文的恶意攻击报文，设备会因为处理大量该类报文而导致性能降低。

为了避免上述现象发生，可以关闭设备的 ICMPv6 超时报文发送功能，从而减少网络流量、防止遭到恶意攻击。

表1-28 配置 ICMPv6 超时差错报文发送功能

操作	命令	说明
进入系统视图	**system-view**	-
开启设备的ICMPv6超时报文的发送功能	**ipv6 hoplimit-expires enable**	可选 缺省情况下，ICMPv6超时报文发送功能处于开启状态

1.7.4 配置ICMPv6 目的不可达差错报文发送功能

ICMPv6 目的不可达报文发送功能是在设备收到 IPv6 数据报文后，如果发生目的不可达的差错，则将报文丢弃并给源端发送 ICMPv6 目的不可达差错报文。

设备在满足下列条件时会发送目的不可达报文：

- 设备在转发报文时，如果在路由表中没有找到对应的转发路由，且路由表中没有缺省路由，则给源端发送“没有到达目的地址的路由”ICMPv6 差错报文；
- 设备在转发报文时，如果是因为管理策略（例如防火墙过滤、ACL 等）导致无法发送报文时，则给源端发送“与目的地址的通信被管理策略禁止”ICMPv6 差错报文；
- 设备在转发报文时，如果报文的目的 IPv6 地址超出源 IPv6 地址的范围（例如，报文的源 IPv6 地址为链路本地地址，报文的目的 IPv6 地址为全球单播地址），会导致报文无法到达目的端，此时要给源端发送“超出源地址范围”ICMPv6 差错报文；
- 设备在转发报文时，如果不能解析目的 IPv6 地址对应的链路层地址，则给源端发送“地址不可达”ICMPv6 差错报文；
- 设备收到目的地址为本地、传输层协议为 UDP 的数据报文时，如果报文的目的端口号与正在使用的进程不匹配，则给源端发送“端口不可达”ICMPv6 差错报文。

由于 ICMPv6 目的不可达报文传递给用户进程的信息为不可达信息，如果有用户恶意攻击，可能会影响终端用户的正常使用。为了避免上述现象发生，可以关闭设备的 ICMPv6 目的不可达报文发送功能，从而减少网络流量、防止遭到恶意攻击。

表1-29 配置 ICMPv6 目的不可达报文发送功能

操作	命令	说明
进入系统视图	**system-view**	-
开启设备的ICMPv6目的不可达报文的发送功能	**ipv6 unreachables enable**	必选 缺省情况下，ICMPv6目的不可达报文发送功能处于关闭状态

1.8 配置组播ND

微软的网络负载均衡（NLB，Network Load Balancing）功能，是其在 Windows Server 上开发的一个多服务器集群负载均衡特性。

NLB 支持集群内服务器之间的负载分担以及冗余备份，当发生服务器故障时可以支持数据快速切换。为了保证快速切换，NLB 要求交换机将业务流量转发至集群内的所有服务器或指定服务器，然后由各服务器将该服务器不期望的流量过滤掉，因此对于那些使用 Windows Server 作为服务器操作系统的中、小型数据中心来说，交换机与 NLB 的协同工作非常重要。

为了让业务流量能够被转发到所有服务器或指定服务器，微软 NLB 采取了一些处理机制，包括单播模式处理机制、组播模式处理机制和 IGMP 组播模式处理机制。

- 单播模式：在单播模式下，NLB 重新为每个 NLB 节点分配一个共同的 MAC 地址（该 MAC 地址为集群 MAC 地址），并且在发送时修改数据报文的源 MAC 地址，从而使交换机不能将集群 MAC 地址学习到 MAC 地址表中，这样目的地址为集群 MAC 地址的数据报文将作为未知单播报文在交换机的所有端口上进行转发。
- 组播模式：在组播模式下，NLB 使用一个组播 MAC 地址（该 MAC 地址为虚拟 MAC）用于 NLB 的通信，例如使用 0300-5e11-1111 来作为 NLB 节点的虚 MAC 地址。
- IGMP 组播模式：IGMP 组播模式和组播模式的区别在于：IGMP 组播模式可以通过 IGMP 协议使交换机只将数据报文发送到连接 NLB 节点的端口，而不是所有端口。

说明

- 仅 R5206 及以上版本支持组播 ND。
- 组播 ND 功能仅适用于 NLB 采用组播模式处理机制的情况。
- 关于 NLB 的详细介绍请参见 Windows Server 的相关文档。

表1-30 配置组播 ND

操作	说明	详细配置
进入系统视图	**system-view**	-
手工添加静态ND表项	**ipv6 neighbor** *ipv6-address mac-address vlan-id port-type port-number* [**vpn-instance** *vpn-instance-name*]	可选
配置静态组播MAC地址表项	**mac-address multicast** *mac-address* **interface** *interface-list* **vlan** *vlan-id*	必选

说明

- **ipv6 neighbor** 命令的详细介绍请参考“三层技术-IP 业务命令参考”内的“IPv6 基础命令”。
- **mac-address multicast** 命令的详细介绍请参考 “IP 组播命令参考”内的“IGMP Snooping 命令”。

1.9 IPv6基础显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示 IPv6 配置后的运行情况，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** 命令可以清除相应的统计信息。

表1-31 IPv6 基础显示和维护

操作	命令
显示IPv6 FIB转发信息表项	**display ipv6 fib** [**vpn-instance** *vpn-instance-name*] [**acl6** *acl6-number* | **ipv6-prefix** *ipv6-prefix-name*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定目的IPv6地址的IPv6 FIB转发信息表项	**display ipv6 fib** [**vpn-instance** *vpn-instance-name*] *ipv6-address* [*prefix-length*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示接口的IPv6信息	**display ipv6 interface** [*interface-type* [*interface-number*]] [**brief**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示邻居信息	**display ipv6 neighbors** { { *ipv6-address* | **all** | **dynamic** | **static** } [**slot** *slot-number*] | **interface** *interface-type interface-number* | **vlan** *vlan-id* } [**verbose**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示符合指定条件的邻居表项的总个数	**display ipv6 neighbors** { { **all** | **dynamic** | **static** } [**slot** *slot-number*] | **interface** *interface-type interface-number* | **vlan** *vlan-id* } **count** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定VPN实例的邻居信息	**display ipv6 neighbors vpn-instance** *vpn-instance-name* [**count**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6的PMTU信息	**display ipv6 pathmtu** [**vpn-instance** *vpn-instance-name*] { *ipv6-address* | **all** | **dynamic** | **static** } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示指定套接字的相关信息	**display ipv6 socket** [**socktype** *socket-type*] [*task-id socket-id*] [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6报文及ICMPv6报文的统计信息	**display ipv6 statistics** [**slot** *slot-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示TCP6连接的统计信息	**display tcp ipv6 statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示TCP6连接的状态信息	**display tcp ipv6 status** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示UDP6的统计信息	**display udp ipv6 statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示ND Snooping表项	**display ipv6 nd snooping** [*ipv6-address* | **vlan** *vlan-id*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IPv6邻居信息	**reset ipv6 neighbors** { **all** | **dynamic** | **interface** *interface-type interface-number* | **slot** *slot-number* | **static** }
清除PMTU值	**reset ipv6 pathmtu** { **all** | **static** | **dynamic**}
清除IPv6报文及ICMPv6报文的统计信息	**reset ipv6 statistics** [**slot** *slot-number*]
清除所有TCP6连接的统计信息	**reset tcp ipv6 statistics**

操作	命令
清除所有UDP6统计信息	**reset udp ipv6 statistics**
清除ND Snooping表项	**reset ipv6 nd snooping** [*ipv6-address* | **vlan** *vlan-id*]

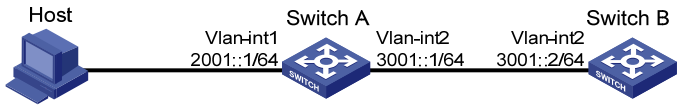
1.10 IPv6基础典型配置举例

1. 组网需求

- 如 图 1-8 所示，Host、Switch A和Switch B之间通过以太网端口相连，将以太网端口分别加入相应的VLAN里，在VLAN接口上配置IPv6 地址，验证它们之间的互通性。
- Switch A 的 VLAN 接口 1 的全球单播地址为 2001::1/64，VLAN 接口 2 的全球单播地址为 3001::1/64。
- Switch B 的 VLAN 接口 2 的全球单播地址为 3001::2/64，有可以到 Host 的路由。
- Host 上安装了 IPv6，根据 IPv6 邻居发现协议自动配置 IPv6 地址，有可以到 Switch B 的路由。

2. 组网图

图1-8 IPv6 地址配置组网图



说明

交换机上已经创建相应的 VLAN 接口。

3. 配置步骤

(1) 配置 Switch A

使能交换机的 IPv6 转发功能。

```
<SwitchA> system-view
[SwitchA] ipv6
```

手工指定 VLAN 接口 2 的全球单播地址。

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address 3001::1/64
[SwitchA-Vlan-interface2] quit
```

手工指定 VLAN 接口 1 的全球单播地址，并允许其发布 RA 消息。（缺省情况下，所有的接口不会发布 RA 消息）

```
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] ipv6 address 2001::1/64
[SwitchA-Vlan-interface1] undo ipv6 nd ra halt
[SwitchA-Vlan-interface1] quit
```

(2) 配置 Switch B

使能交换机的 IPv6 转发功能。

```
<SwitchB> system-view
[SwitchB] ipv6
```

配置 VLAN 接口 2 的全球单播地址。

```
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ipv6 address 3001::2/64
[SwitchB-Vlan-interface2] quit
```

配置 IPv6 静态路由，该路由的目的地址为 2001::/64，下一跳地址为 3001::1。

```
[SwitchB] ipv6 route-static 2001:: 64 3001::1
```

(3) 配置 Host

Host 上安装 IPv6，根据 IPv6 邻居发现协议自动配置 IPv6 地址。

从 Switch A 上使用 **ping ipv6** 命令检查 Switch B 是否可达。

```
[SwitchA] ping ipv6 3001::1
  PING 3001::1 : 56  data bytes, press CTRL_C to break
    Reply from 3001::1
    bytes=56 Sequence=0 hop limit=64  time = 3 ms
    Reply from 3001::1
    bytes=56 Sequence=1 hop limit=64  time = 2 ms
    Reply from 3001::1
    bytes=56 Sequence=2 hop limit=64  time = 2 ms
    Reply from 3001::1
    bytes=56 Sequence=3 hop limit=64  time = 3 ms
    Reply from 3001::1
    bytes=56 Sequence=4 hop limit=64  time = 9 ms

  --- 3001::1 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
```

round-trip min/avg/max = 2/3/9 ms

从 Switch A 上查看端口 GigabitEthernet1/0/2 的邻居信息。

```
[SwitchA] display ipv6 neighbors interface GigabitEthernet 1/0/2
               Type: S-Static    D-Dynamic
IPv6 Address                Link-layer      VID  Interface    State T Age
FE80::215:E9FF:FEA6:7D14     0015-e9a6-7d14  1     GE1/0/2         STALE D 1238
2001::15B:E0EA:3524:E791     0015-e9a6-7d14  1     GE1/0/2         STALE D 1248
```

通过上面的信息可以知道 Host 上获得的 IPv6 全球单播地址为 2001::15B:E0EA:3524:E791。

4. 验证配置结果

显示 Switch A 的接口信息，可以看到各接口配置的 IPv6 全球单播地址。

```
[SwitchA] display ipv6 interface vlan-interface 2
Vlan-interface2 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::20F:E2FF:FE00:2
  Global unicast address(es):
```

```
    3001::1, subnet is 3001::/64
  Joined group address(es):
    FF02::1:FF00:0
    FF02::1:FF00:1
    FF02::1:FF00:2
    FF02::2
    FF02::1
  MTU is 1500 bytes
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  ND retransmit interval is 1000 milliseconds
  Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
  InReceives:                   25829
  InTooShorts:                  0
  InTruncatedPkts:              0
  InHopLimitExceeds:            0
  InBadHeaders:                 0
  InBadOptions:                 0
  ReasmReqds:                   0
  ReasmOKs:                     0
  InFragDrops:                  0
  InFragTimeouts:               0
  OutFragFails:                 0
  InUnknownProtos:              0
  InDelivers:                   47
  OutRequests:                  89
  OutForwDatagrams:             48
  InNoRoutes:                   0
  InTooBigErrors:               0
  OutFragOKs:                   0
  OutFragCreates:               0
  InMcastPkts:                  6
  InMcastNotMembers:            25747
  OutMcastPkts:                 48
  InAddrErrors:                 0
  InDiscards:                   0
  OutDiscards:                  0

[SwitchA] display ipv6 interface vlan-interface 1
Vlan-interface1 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::20F:E2FF:FE00:1C0
  Global unicast address(es):
    2001::1, subnet is 2001::/64
  Joined group address(es):
    FF02::1:FF00:0
    FF02::1:FF00:1
```

```
    FF02::1:FF00:1C0
    FF02::2
    FF02::1
  MTU is 1500 bytes
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  ND retransmit interval is 1000 milliseconds
  ND advertised reachable time is 0 milliseconds
  ND advertised retransmit interval is 0 milliseconds
  ND router advertisements are sent every 600 seconds
  ND router advertisements live for 1800 seconds
  Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
  InReceives:                    272
  InTooShorts:                   0
  InTruncatedPkts:               0
  InHopLimitExceeds:             0
  InBadHeaders:                  0
  InBadOptions:                  0
  ReasmReqds:                    0
  ReasmOKs:                      0
  InFragDrops:                   0
  InFragTimeouts:                0
  OutFragFails:                  0
  InUnknownProtos:               0
  InDelivers:                    159
  OutRequests:                   1012
  OutForwDatagrams:              35
  InNoRoutes:                    0
  InTooBigErrors:                0
  OutFragOKs:                    0
  OutFragCreates:                0
  InMcastPkts:                   79
  InMcastNotMembers:             65
  OutMcastPkts:                  938
  InAddrErrors:                  0
  InDiscards:                    0
  OutDiscards:                   0
```

显示 Switch B 的接口信息，可以看到接口配置的 IPv6 全球单播地址。

```
[SwitchB] display ipv6 interface vlan-interface 2
Vlan-interface2 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::20F:E2FF:FE00:1234
  Global unicast address(es):
    3001::2, subnet is 3001::/64
  Joined group address(es):
    FF02::1:FF00:0
```

```
    FF02::1:FF00:2
    FF02::1:FF00:1234
    FF02::2
    FF02::1
  MTU is 1500 bytes
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  ND retransmit interval is 1000 milliseconds
  Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
  InReceives:                   117
  InTooShorts:                  0
  InTruncatedPkts:              0
  InHopLimitExceeds:            0
  InBadHeaders:                 0
  InBadOptions:                 0
  ReasmReqds:                   0
  ReasmOKs:                     0
  InFragDrops:                  0
  InFragTimeouts:               0
  OutFragFails:                 0
  InUnknownProtos:              0
  InDelivers:                   117
  OutRequests:                  83
  OutForwDatagrams:             0
  InNoRoutes:                   0
  InTooBigErrors:               0
  OutFragOKs:                   0
  OutFragCreates:               0
  InMcastPkts:                  28
  InMcastNotMembers:            0
  OutMcastPkts:                 7
  InAddrErrors:                 0
  InDiscards:                   0
  OutDiscards:                  0
```

在 Host 上使用 Ping 测试和 Switch A 及 Switch B 的互通性；在 Switch B 上使用 Ping 测试和 Switch A 及 Host 的互通性。

注意

在 Ping 链路本地地址时，需要使用**-i** 参数来指定链路本地地址的接口。

```
[SwitchB] ping ipv6 -c 1 3001::1
  PING 3001::1 : 56  data bytes, press CTRL_C to break
    Reply from 3001::1
    bytes=56 Sequence=1 hop limit=64  time = 2 ms
```

```
  --- 3001::1 ping statistics ---
    1 packet(s) transmitted
    1 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 2/2/2 ms
[SwitchB-Vlan-interface2] ping ipv6 -c 1 2001::15B:E0EA:3524:E791
  PING 2001::15B:E0EA:3524:E791 : 56  data bytes, press CTRL_C to break
    Reply from 2001::15B:E0EA:3524:E791
    bytes=56 Sequence=1 hop limit=63  time = 3 ms

  --- 2001::15B:E0EA:3524:E791 ping statistics ---
    1 packet(s) transmitted
    1 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 3/3/3 ms
```

从 Host 上也可以 ping 通 Switch B 和 Switch A，证明它们是互通的。

1.11 常见配置错误举例

1. 故障现象

无法 Ping 通对端的 IPv6 地址。

2. 故障排除

- 在任意视图使用 **display current-configuration** 命令或在系统视图下使用 **display this** 命令检查是否使能了 IPv6 报文转发功能。
- 在任意视图下使用 **display ipv6 interface** 命令检查接口配置的 IPv6 地址是否正确，接口状态是否为 up。
- 在用户视图下使用 **debugging ipv6 packet** 命令打开 IPv6 报文调试开关，根据调试信息进行判断。

目 录

1 DHCPv6 简介

1.1 DHCPv6概述

DHCPv6（Dynamic Host Configuration Protocol for IPv6，支持 IPv6 的动态主机配置协议）是针对 IPv6 编址方案设计的，为主机分配 IPv6 前缀、IPv6 地址和其他网络配置参数的协议。

与其他 IPv6 地址分配方式（手工配置、通过路由器公告消息中的网络前缀无状态自动配置等）相比，DHCPv6 具有以下优点：

- 更好地控制地址的分配。通过 DHCPv6 不仅可以记录为主机分配的地址，还可以为特定主机分配特定的地址，以便于网络管理。
- 为设备分配前缀，便于全网络的自动配置和管理。
- 除了 IPv6 前缀、IPv6 地址外，还可以为主机分配 DNS 服务器、域名等网络配置参数。

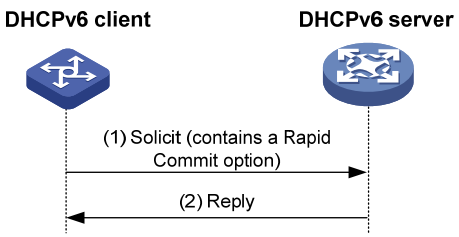
1.2 DHCPv6地址/前缀分配过程

DHCPv6 服务器为客户端分配地址/前缀的过程分为两类：

- 交互两个消息的快速分配过程
- 交互四个消息的分配过程

1.2.1 交互两个消息的快速分配过程

图1-1 地址/前缀快速分配过程



如 图 1-1 所示，地址/前缀快速分配过程为：

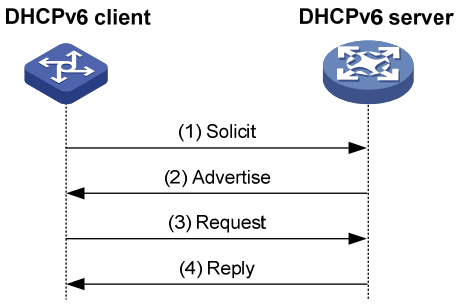
(1) DHCPv6 客户端在发送的 Solicit 消息中携带 Rapid Commit 选项，标识客户端希望服务器能够快速为其分配地址/前缀和网络配置参数；

(2) 如果DHCPv6 服务器支持快速分配过程，则直接返回Reply消息，为客户端分配IPv6 地址/前缀和其他网络配置参数。如果DHCPv6 服务器不支持快速分配过程，则采用“1.2.2 交互四个消息的分配过程”为客户端分配IPv6 地址/前缀和其他网络配置参数。

1.2.2 交互四个消息的分配过程

交互四个消息的分配过程如 图 1-2 所示。

图1-2 交互四个消息的分配过程



交互四个消息分配过程的简述如 表 1-1。

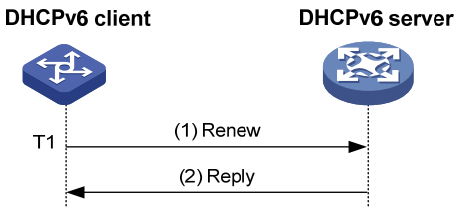
表1-1 交互四个消息的分配过程

步骤	发送的消息	说明
(1)	Solicit	DHCPv6客户端发送该消息，请求DHCPv6服务器为其分配IPv6地址/前缀和网络配置参数
(2)	Advertise	如果Solicit消息中没有携带Rapid Commit选项，或Solicit消息中携带Rapid Commit选项，但服务器不支持快速分配过程，则DHCPv6服务器回复该消息，通知客户端可以为其分配的地址/前缀和网络配置参数
(3)	Request	如果DHCPv6客户端接收到多个服务器回复的Advertise消息，则根据消息接收的先后顺序、服务器优先级等，选择其中一台服务器，并向该服务器发送Request消息，请求服务器确认为其分配地址/前缀和网络配置参数
(4)	Reply	DHCPv6服务器回复该消息，确认将地址/前缀和网络配置参数分配给客户端使用

1.3 地址/前缀租约更新过程

DHCPv6 服务器分配给客户端的 IPv6 地址/前缀具有一定的租借期限。租借期限由有效生命期（Valid Lifetime）决定。地址/前缀的租借时间到达有效生命期后，DHCPv6 客户端不能再使用该地址/前缀。在有效生命期到达之前，如果 DHCPv6 客户端希望继续使用该地址/前缀，则需要更新地址/前缀租约。

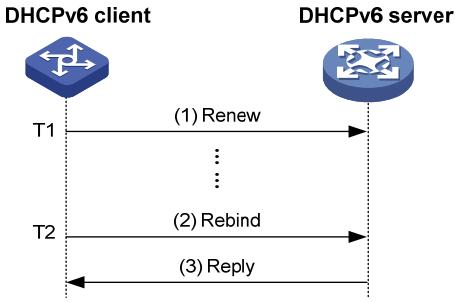
图1-3 通过 Renew 更新地址/前缀租约



如 图 1-3 所示，地址/前缀租借时间到达时间T1（推荐值为首选生命期Preferred Lifetime的一半）时，DHCPv6 客户端会向为它分配地址/前缀的DHCPv6 服务器单播发送Renew报文，以进行地址/前缀租约的更新。如果客户端可以继续使用该地址/前缀，则DHCPv6 服务器回应续约成功的Reply

报文，通知DHCPv6 客户端已经成功更新地址/前缀租约；如果该地址/前缀不可以再分配给该客户端，则DHCPv6 服务器回应续约失败的Reply报文，通知客户端不能获得新的租约。

图1-4 通过 Rebind 更新地址/前缀租约



如 图 1-4 所示，如果在T1 时发送Renew请求更新租约，但是没有收到DHCPv6 服务器的回应报文，则DHCPv6 客户端会在T2（推荐值为首选生命期的 0.8 倍）时，向所有DHCPv6 服务器组播发送Rebind报文请求更新租约。如果客户端可以继续使用该地址/前缀，则DHCPv6 服务器回应续约成功的Reply报文，通知DHCPv6 客户端已经成功更新地址/前缀租约；如果该地址/前缀不可以再分配给该客户端，则DHCPv6 服务器回应续约失败的Reply报文，通知客户端不能获得新的租约；如果DHCPv6 客户端没有收到服务器的应答报文，则到达有效生命期后，客户端停止使用该地址/前缀。

说明

有效生命期和首选生命期的详细介绍请参见“三层技术-IP 业务配置指导”中的“IPv6 基础”。

1.4 DHCPv6无状态配置

1.4.1 DHCPv6 无状态配置简介

DHCPv6 服务器可以为已经具有 IPv6 地址/前缀的客户端分配其他网络配置参数，该过程称为DHCPv6 无状态配置。

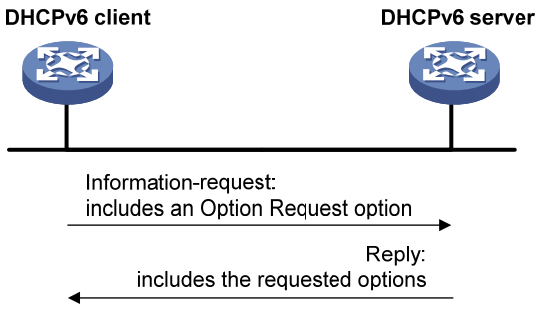
DHCPv6 客户端通过地址无状态自动配置功能成功获取 IPv6 地址后，如果接收到的 RA（Router Advertisement，路由器通告）报文中 M 标志位（Managed address configuration flag，被管理地址配置标志位）为 0、O 标志位（Other stateful configuration flag，其他配置标志位）为 1，则 DHCPv6 客户端会自动启动 DHCPv6 无状态配置功能，以获取除地址/前缀外的其他网络配置参数。

说明

地址无状态自动配置是指节点根据路由器发现/前缀发现所获取的信息，自动配置 IPv6 地址。详细介绍请参见“三层技术-IP 业务配置指导”的“IPv6 基础”。

1.4.2 DHCPv6 无状态配置过程

图1-5 DHCPv6 无状态配置工作过程



如 图 1-5 所示，DHCPv6 无状态配置的具体过程为：

(1) 客户端以组播的方式向 DHCPv6 服务器发送 Information-request 报文，该报文中携带 Option Request 选项，指定客户端需要从服务器获取的配置参数。

(2) 服务器收到 Information-request 报文后，为客户端分配网络配置参数，并单播发送 Reply 报文将网络配置参数返回给客户端。

(3) 客户端检查 Reply 报文中提供的信息，如果与 Information-request 报文中请求的配置参数相符，则按照 Reply 报文中提供的参数进行网络配置；否则，忽略该参数。如果接收到多个 Reply 报文，客户端将选择最先收到的 Reply 报文，并根据该报文中提供的参数完成客户端无状态配置。

1.5 协议规范

与 DHCPv6 相关的协议规范有：

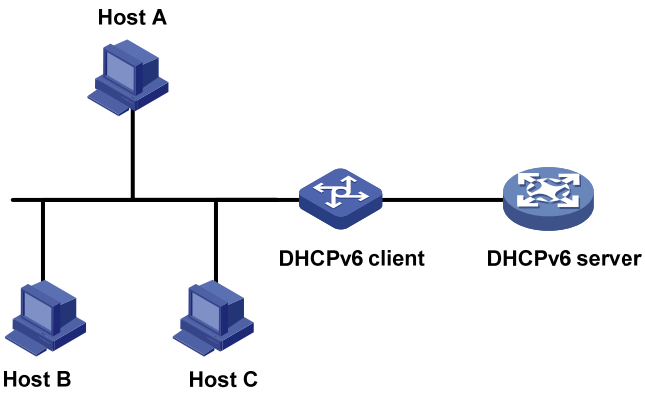
- RFC 3736：Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6
- RFC 3315：Dynamic Host Configuration Protocol for IPv6 (DHCPv6)
- RFC 2462：IPv6 Stateless Address Autoconfiguration
- RFC 3633：IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6

2 DHCPv6 服务器配置

2.1 DHCPv6服务器简介

2.1.1 DHCPv6 服务器应用环境

图2-1 DHCPv6 服务器应用组网图



如 图 2-1 所示，为了便于集中管理IPv6 地址，简化网络配置，DHCPv6 服务器可以用来为DHCPv6 客户端分配IPv6 前缀。DHCPv6 客户端获取到IPv6 前缀后，向所在网络发送包含该前缀信息的RA 消息，以便网络内的主机根据该前缀自动配置IPv6 地址。

说明

目前，设备作为 DHCPv6 服务器时，只能为 DHCPv6 客户端分配前缀，不能分配地址；且支持 DHCPv6 无状态配置，为 DHCPv6 客户端分配除 IPv6 地址/前缀外的其他网络配置参数。

2.1.2 基本概念

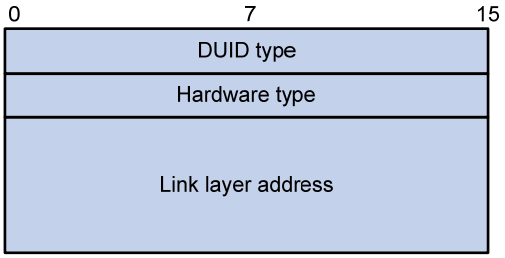
1. DHCPv6 采用的组播地址

DHCPv6 中采用组播地址 FF05::1:3 来表示站点本地范围内所有的 DHCPv6 服务器；采用组播地址 FF02::1:2 来表示链路本地范围内所有的 DHCPv6 服务器和中继。

2. DUID

DUID（DHCP Unique Identifier，DHCP 唯一标识符）用来标识一台 DHCPv6 设备（包括客户端、服务器和中继）。

图2-2 DUID-LL 结构



目前，设备采用RFC 3315 规定的DUID-LL（DUID Based on Link-layer Address，基于链路层地址的DUID）作为DHCPv6 设备的标识。DUID-LL的结构如 图 2-2 所示：

- DUID type：DUID 类型。设备支持的 DUID 类型为 DUID-LL，取值为 0x0003。
- Hardware type：硬件类型。设备支持的硬件类型为以太网，取值为 0x0001。
- Link layer address：链路层地址。取值为设备的桥 MAC 地址。

3. IA

IA（Identity Association，标识联盟）用于管理分配给客户端的一组地址和前缀等信息，通过 IAID 标识。一个客户端可以有多个 IA，如客户端的每个接口拥有一个 IA，用来管理该接口获取的地址和前缀等信息。

4. IAID

IAID 是 IA 的标识符，由客户端选择。在一个客户端上不同 IA 的 IAID 不能相同。

5. PD

PD（Prefix Delegation，前缀授权）是 DHCPv6 服务器为分配的前缀创建的租约，前缀租约中记录了 IPv6 前缀、客户端 DUID、IAID、有效时间、首选时间、租约过期时间、申请前缀的客户端的 IPv6 地址等信息。

2.1.3 前缀的选择过程

DHCPv6 服务器从接口接收到客户端的请求后，从该接口应用的地址池中选择前缀和其他网络配置参数，分配给客户端。地址池既可以包含静态绑定前缀，用于为特定客户端分配固定前缀；也可以引用包含一定前缀范围的前缀池，用于动态选择可用前缀分配给客户端。

DHCPv6 服务器按照如下顺序从地址池中为客户端选择匹配的前缀：

(1) DUID、IAID 与客户端 DUID、IAID 匹配，且与客户端期望前缀匹配的静态绑定前缀；

(2) DUID、IAID 与客户端 DUID、IAID 匹配的静态绑定前缀；

(3) DUID 与客户端的 DUID 匹配，且与客户端期望前缀匹配的静态绑定前缀，该前缀中未指定客户端的 IAID；

(4) DUID 与客户端 DUID 匹配的静态绑定前缀，该前缀中未指定客户端的 IAID；

(5) 前缀池中与客户端期望前缀匹配的空闲前缀；

(6) 前缀池中的其他空闲前缀。

2.2 配置DHCPv6服务器

2.2.1 DHCPv6 服务器配置任务简介

表2-1 DHCPv6 服务器配置任务简介

配置任务	说明	详细配置
使能DHCPv6服务器功能	必选	2.2.3
创建前缀池	必选	2.2.4
配置DHCPv6地址池	必选	2.2.5
配置接口引用地址池	必选	2.2.6
配置DHCPv6服务器发送的DHCPv6报文的DSCP优先级	可选	2.2.7

2.2.2 配置准备

配置 DHCPv6 服务器功能之前，需要先执行 **ipv6** 命令使能 IPv6 报文收发功能。**ipv6** 命令的详细介绍请参见“三层技术-IP 业务命令参考”中的“IPv6 基础”。

2.2.3 使能DHCPv6 服务器功能

表2-2 使能 DHCPv6 服务器功能

操作	命令	说明
进入系统视图	**system-view**	-
使能DHCPv6服务器功能	**ipv6 dhcp server enable**	必选 缺省情况下，DHCPv6服务器功能处于关闭状态

2.2.4 创建前缀池

前缀池用来定义一个前缀范围。

表2-3 创建前缀池

操作	命令	说明
进入系统视图	**system-view**	-
创建前缀池	**ipv6 dhcp prefix-pool** *prefix-pool-number* **prefix** *prefix/prefix-len* **assign-len** *assign-len*	必选 缺省情况下，设备上不存在任何前缀池

2.2.5 配置DHCPv6 地址池

DHCPv6 地址池中包括了供分配的前缀和网络配置参数等信息，如 DNS 服务器地址、域名、SIP 服务器地址、SIP 服务器域名和 AFTR（Address Family Transition Router，地址族转换路由器）地址。DHCPv6 服务器从地址池中选择前缀和其他网络配置参数，分配给客户端。

表2-4 配置 DHCPv6 地址池

操作	命令	说明
进入系统视图	**system-view**	-
创建DHCPv6地址池，并进入DHCPv6地址池视图	**ipv6 dhcp pool** *pool-number*	必选 缺省情况下，设备上不存在任何DHCPv6地址池
配置静态绑定前缀	**static-bind prefix** *prefix/prefix-len* **duid** *duid* [**iaid** *iaid*] [**preferred-lifetime** *preferred-lifetime* **valid-lifetime** *valid-lifetime*]	二者至少选择其一 缺省情况下，没有指定地址池分配的前缀
配置地址池引用前缀池	**prefix-pool** *prefix-pool-number* [**preferred-lifetime** *preferred-lifetime* **valid-lifetime** *valid-lifetime*]	
配置为客户端分配的DNS服务器地址	**dns-server** *ipv6-address*	可选 缺省情况下，没有指定为客户端分配的DNS服务器地址
配置为客户端分配的域名	**domain-name** *domain-name*	可选 缺省情况下，没有指定为客户端分配的域名
配置为客户端分配的SIP服务器地址或域名	**sip-server** { **address** *ipv6-address* | **domain-name** *domain-name* }	可选 缺省情况下，没有指定为客户端分配的SIP服务器地址或域名
配置为客户端分配的AFTR地址	**ds-lite address** *ipv6-address*	可选 缺省情况下，没有指定为客户端分配的AFTR地址

说明

- 一个地址池最多可以引用一个前缀池。
- 地址池可以引用并不存在的前缀池，但是，此时设备无法从该地址池中动态选择前缀分配给客户端。只有创建该前缀池后，才能支持前缀的动态选择。
- 不允许通过重复执行 **prefix-pool** 命令的方式修改地址池引用的前缀池、前缀的首选生命期和有效生命期。只有取消当前地址池引用的前缀池后，才能引用其他的前缀池，或修改首选生命期和有效生命期。
- 一个地址池下最多可以配置 8 个 DNS 服务器地址、1 个域名、8 个 SIP 服务器地址和 8 个 SIP 服务器域名。

2.2.6 配置接口引用地址池

在接口上引用地址池后，该接口将具有 DHCPv6 服务器功能，即从该接口接收到客户端的请求后，设备将从该接口应用的地址池中选择前缀和其他网络配置参数，分配给客户端。

表2-5 配置接口引用地址池

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*	-
配置接口引用DHCPv6地址池	**ipv6 dhcp server apply pool** *pool-number* [**allow-hint** | **preference** *preference-value* | **rapid-commit**] *	必选 缺省情况下，接口没有引用DHCPv6地址池

说明

- 一个接口不能同时作为 DHCPv6 服务器和 DHCPv6 中继。
- 建议不要在一个接口上同时配置 DHCPv6 服务器和 DHCPv6 客户端功能。
- 一个接口上最多只能引用一个地址池。
- 接口可以引用并不存在的地址池，但是，此时该接口无法为客户端分配前缀等信息。只有创建该地址池后，才能为客户端分配前缀等信息。
- 不允许通过重复执行 **ipv6 dhcp server apply pool** 命令的方式修改接口引用的地址池和服务器优先级等参数。只有取消当前接口引用的地址池后，才能引用其他的地址池，或修改服务器优先级等参数。

2.2.7 配置DHCPv6 服务器发送的DHCPv6 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 DHCPv6 服务器发送的 DHCPv6 报文的 DSCP 优先级进行配置。

表2-6 配置 DHCPv6 服务器发送的 DHCPv6 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
配置DHCPv6服务器发送的DHCPv6报文的DSCP优先级	**ipv6 dhcp dscp** *dscp-value*	可选 缺省情况下，DHCPv6服务器发送的DHCPv6报文的DSCP优先级为56

2.3 DHCPv6服务器显示和维护

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 DHCPv6 服务器的运行情况，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 DHCPv6 服务器的统计信息。

表2-7 DHCPv6 服务器显示和维护

操作	命令
显示本设备DUID	**display ipv6 dhcp duid** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCPv6地址池的信息	**display ipv6 dhcp pool** [*pool-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示前缀池的信息	**display ipv6 dhcp prefix-pool** [*prefix-pool-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCPv6服务器的配置信息	**display ipv6 dhcp server** [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示前缀租约信息	**display ipv6 dhcp server pd-in-use** { **all** | **pool** *pool-number* | **prefix** *prefix/prefix-len* | **prefix-pool** *prefix-pool-number* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCPv6服务器的报文统计信息	**display ipv6 dhcp server statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除DHCPv6服务器的前缀租约信息	**reset ipv6 dhcp server pd-in-use** { **all** | **pool** *pool-number* | **prefix** *prefix/prefix-len* }
清除DHCPv6服务器的报文统计信息	**reset ipv6 dhcp server statistics**

2.4 DHCPv6服务器典型配置举例

1. 组网需求

DHCPv6 客户端从 DHCPv6 服务器获取 IPv6 地址前缀，以及网络配置参数：DNS 服务器地址、域名、SIP 服务器地址和 SIP 服务器域名。其中：

- Switch 作为 DHCPv6 服务器，地址为 1::1/64。
- DHCPv6 服务器为 DUID 为 00030001CA0006A40000 的客户端固定分配前缀 2001:0410:0201::/48；为其他客户端分配 2001:0410::/48～2001:0410:FFFF::/48 之间除 2001:0410:0201::/48 外的前缀。
- DNS 服务器地址为 2:2::3。
- DHCPv6 客户端所属域的域名为 aaa.com。
- SIP 服务器地址为 2:2::4，域名为 bbb.com。

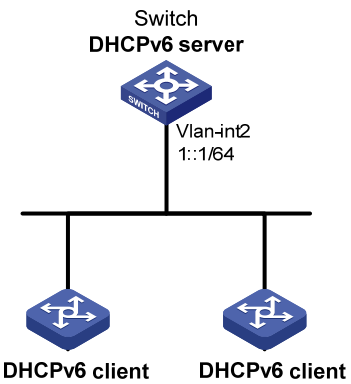
2. 配置思路

在 DHCPv6 服务器上需要进行如下配置：

- 使能 IPv6 报文转发功能和 DHCPv6 服务器功能。
- 创建前缀池。为了满足为客户端分配 2001:0410::/48～2001:0410:FFFF::/48 范围的前缀这一需求，需要配置前缀池包含的前缀为 2001:0410::/32，分配的前缀长度为 48。
- 创建地址池。在地址池配置静态绑定前缀，引用前缀池，并配置 DNS 服务器地址等参数。
- 在连接 DHCPv6 客户端的接口上引用地址池。

3. 组网图

图2-3 DHCPv6 服务器配置组网图



4. 配置步骤

(1) 配置 DHCPv6 服务器

使能 IPv6 报文转发功能及 DHCPv6 服务器功能。

```
<Switch> system-view
[Switch] ipv6
[Switch] ipv6 dhcp server enable
```

配置 VLAN 接口 2 的 IPv6 地址。

```
[Switch] interface vlan-interface 2
```

```
[Switch-Vlan-interface2] ipv6 address 1::1/64
[Switch-Vlan-interface2] quit
```
配置前缀池 1，包含的前缀为 2001:0410::/32，分配的前缀长度为 48。
```
[Switch] ipv6 dhcp prefix-pool 1 prefix 2001:0410::/32 assign-len 48
```
创建地址池 1。
```
[Switch] ipv6 dhcp pool 1
```
配置地址池 1 引用已存在的前缀池 1，并设置首选生命期为 1 天，有效生命期为 3 天。
```
[Switch-ipv6-dhcp-pool-1] prefix-pool 1 preferred-lifetime 86400 valid-lifetime 259200
```
在地址池 1 中配置静态绑定前缀：绑定的前缀为 2001:0410:0201::/48，绑定的客户端 DUID 为 00030001CA0006A40000，并设置首选生命期为 1 天，有效生命期为 3 天。
```
[Switch-ipv6-dhcp-pool-1] static-bind prefix 2001:0410:0201::/48 duid 00030001CA0006A40000
preferred-lifetime 86400 valid-lifetime 259200
```
配置为客户端分配的 DNS 服务器地址为 2:2::3。
```
[Switch-ipv6-dhcp-pool-1] dns-server 2:2::3
```
配置为客户端分配的域名为 aaa.com。
```
[Switch-ipv6-dhcp-pool-1] domain-name aaa.com
```
配置为客户端分配的 SIP 服务器地址为 2:2::4，域名为 bbb.com。
```
[Switch-ipv6-dhcp-pool-1] sip-server address 2:2::4
[Switch-ipv6-dhcp-pool-1] sip-server domain-name bbb.com
[Switch-ipv6-dhcp-pool-1] quit
```
在 VLAN 接口 2 上引用已存在的地址池 1，使能期望前缀分配和前缀快速分配功能，并将优先级设置为最高。
```
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ipv6 dhcp server apply pool 1 allow-hint preference 255 rapid-commit
```
(2) 验证配置结果

完成上述配置后，查看 VLAN 接口 2 上的 DHCPv6 服务器配置信息。
```
[Switch-Vlan-interface2] display ipv6 dhcp server interface vlan-interface 2
Using pool: 1
Preference value: 255
Allow-hint: Enabled
Rapid-commit: Enabled
```
显示地址池 1 的信息。
```
[Switch-Vlan-interface2] display ipv6 dhcp pool 1
DHCPv6 pool: 1
  Static bindings:
    DUID: 00030001CA0006A40000
    IAID: A1A1A1A1
    Prefix: 2001:410:201::/48
      preferred lifetime 86400, valid lifetime 2592000
  Prefix pool: 1
    preferred lifetime 86400, valid lifetime 2592000
  DNS server address:
    2:2::3
  Domain name: aaa.com
  SIP server address:
```

```
    2:2::4
  SIP server domain name:
    bbb.com
```

显示前缀池 1 的信息。

```
[Switch-Vlan-interface2] display ipv6 dhcp prefix-pool 1
Prefix: 2001:410::/32
Assigned length: 48
Total prefix number: 65536
Available: 65535
In-use: 0
Static: 1
```

DUID 为 00030001CA0006A40000 的客户端获取 IPv6 前缀后，显示前缀租约信息。

```
[Switch-Vlan-interface2] display ipv6 dhcp server pd-in-use all
Total number = 1
Prefix                                    Type      Pool Lease-expiration
2001:410:201::/48                         Static(C) 1    Jul 10 2009 19:45:01
```

其他客户端获取 IPv6 前缀后，显示前缀租约信息。

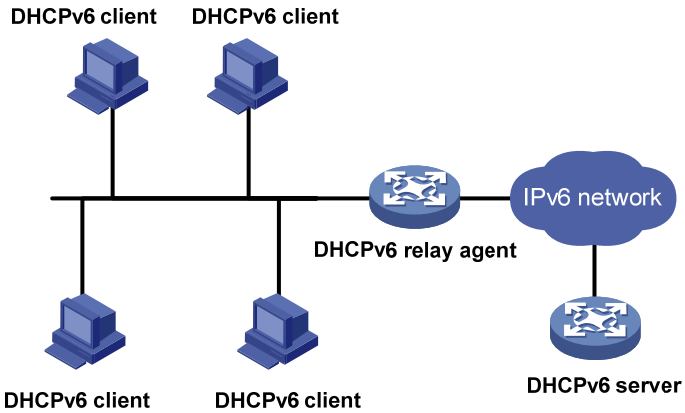
```
[Switch-Vlan-interface2] display ipv6 dhcp server pd-in-use all
Total number = 2
Prefix                                    Type      Pool Lease-expiration
2001:410:201::/48                         Static(C) 1    Jul 10 2009 19:45:01
2001:410::/48                             Auto(C)   1    Jul 10 2009 20:44:05
```

3 DHCPv6 中继配置

3.1 DHCPv6中继简介

3.1.1 应用环境

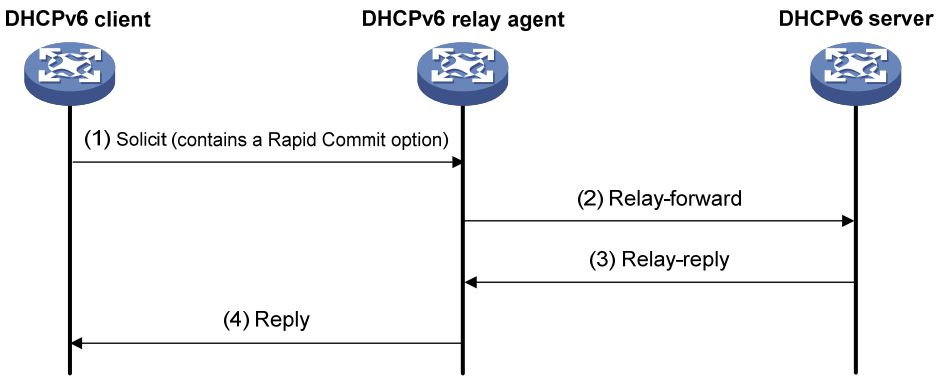
图3-1 DHCPv6 中继应用组网图



DHCPv6 客户端通常通过链路本地范围的组播地址与DHCPv6 服务器通信，以获取IPv6 地址和其他网络配置参数。如 图 3-1 所示，服务器和客户端不在同一个链路范围内时，服务器和客户端无法直接通信，需要通过DHCPv6 中继来转发报文。部署DHCPv6 中继可以避免在每个链路范围内都部署DHCPv6 服务器，既节省了成本，又便于进行集中管理。

3.1.2 DHCPv6 中继的工作过程

图3-2 DHCPv6 中继的工作过程



如 图 3-2 所示，以交互两个消息的快速分配过程为例，DHCPv6 客户端通过DHCPv6 中继，从DHCPv6 服务器获取IPv6 地址和其他网络配置参数的过程为：

(1) DHCPv6 客户端向所有 DHCPv6 服务器和中继的组播地址 FF02::1:2 发送携带 Rapid Commit 选项的 Solicit 消息；

(2) DHCPv6 中继接收到 Solicit 消息后，将其封装在 Relay-forward 报文的中继消息选项（Relay Message Option）中，并将 Relay-forward 报文发送给 DHCPv6 服务器；

(3) DHCPv6 服务器从 Relay-forward 报文中解析出客户端的 Solicit 消息，为客户端选取 IPv6 地址和其他参数，构造 Reply 消息，将 Reply 消息封装在 Relay-reply 报文的中继消息选项中，并将 Relay-reply 报文发送给 DHCPv6 中继；

(4) DHCPv6 中继从 Relay-reply 报文中解析出服务器的 Reply 消息，转发给 DHCPv6 客户端，以便 DHCPv6 客户端根据 DHCPv6 服务器分配的 IPv6 地址和其他参数进行网络配置。

3.2 配置DHCPv6中继

工作在 DHCPv6 中继模式的接口接收到 DHCPv6 客户端发来的报文后，将其封装在 Relay-forward 报文中，并发送给指定的 DHCPv6 服务器，由 DHCPv6 服务器为客户端分配 IPv6 地址和其他网络配置参数。

3.2.1 配置准备

配置 DHCPv6 中继之前，需要通过系统视图下的 **ipv6** 命令使能 IPv6 报文的转发功能。

3.2.2 配置步骤

表3-1 配置 DHCPv6 中继

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*	-
配置接口工作在DHCPv6中继模式，并指定DHCPv6服务器的地址	**ipv6 dhcp relay server-address** *ipv6-address* [**interface** *interface-type interface-number*]	必选 缺省情况下，接口未工作在DHCPv6中继模式，也未指定任何DHCPv6服务器

说明

- 通过多次执行 **ipv6 dhcp relay server-address** 命令可以指定多个 DHCPv6 服务器，一个接口下最多可以指定 8 个 DHCPv6 服务器。DHCPv6 中继接收到 DHCPv6 客户端报文后，将其转发给所有的 DHCPv6 服务器。
- 如果指定的 DHCPv6 服务器地址为链路本地地址或链路范围的组播地址，则必须通过 **ipv6 dhcp relay server-address** 命令的 **interface** 参数指定出接口，否则报文可能会无法到达服务器。
- 通过 **undo ipv6 dhcp relay server-address** 命令删除接口上指定的全部 DHCPv6 服务器后，该接口不再工作在 DHCPv6 中继模式。
- 一个接口不能同时作为 DHCPv6 中继和 DHCPv6 服务器。
- 建议不要在一个接口上同时配置 DHCPv6 中继和 DHCPv6 客户端功能。

3.3 配置DHCPv6中继发送的DHCPv6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 DHCPv6 中继发送的 DHCPv6 报文的 DSCP 优先级进行配置。

表3-2 配置 DHCPv6 中继发送的 DHCPv6 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
配置 DHCPv6 中继发送的 DHCPv6报文的DSCP优先级	**ipv6 dhcp dscp** *dscp-value*	可选 缺省情况下，DHCPv6 中继发送的 DHCPv6报文的DSCP优先级为56

3.4 DHCPv6中继显示和维护

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 DHCPv6 中继的运行情况，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 DHCPv6 中继的统计信息。

表3-3 DHCPv6 中继显示和维护

操作	命令
显示本设备DUID	**display ipv6 dhcp duid** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCPv6中继指定的DHCPv6服务器地址信息	**display ipv6 dhcp relay server-address** { **all** | **interface** *interface-type interface-number* } [**|** { **begin** | **exclude** | **include** } *regular-expression*]

操作	命令
显示DHCPv6中继的报文统计信息	**display ipv6 dhcp relay statistics** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除DHCPv6中继的报文统计信息	**reset ipv6 dhcp relay statistics**

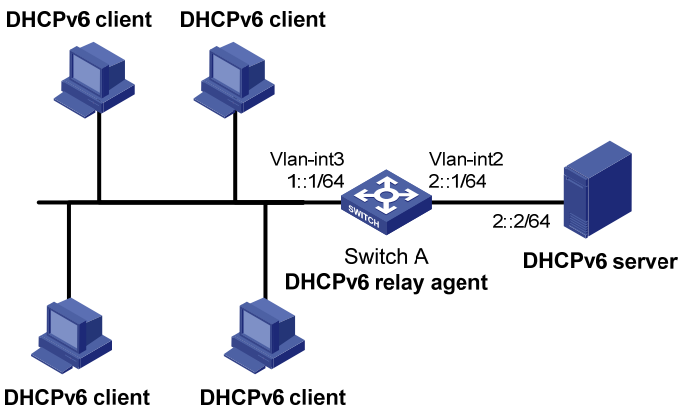
3.5 DHCPv6中继典型配置举例

1. 组网需求

- DHCPv6 客户端所在网络地址为 1::/64，DHCPv6 服务器的地址为 2::2/64。客户端和服务器不在同一个链路，需要通过 DHCPv6 中继转发报文。
- Switch A 作为 DHCPv6 中继，为客户端和服务器转发报文。
- Switch A 同时作为 1::/64 网络的网关设备，通过 RA 消息中的 M 标志位和 O 标志位指定该网络中的主机通过 DHCPv6 获取 IPv6 地址和其他网络配置参数。

2. 组网图

图3-3 DHCPv6 中继组网图



3. 配置步骤

(1) 配置 Switch A 作为 DHCPv6 中继

使能 IPv6 报文转发功能。

```
<SwitchA> system-view
[SwitchA] ipv6
```

配置 VLAN 接口 2 和 VLAN 接口 3 的 IPv6 地址。

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address 2::1 64
[SwitchA-Vlan-interface2] quit
[SwitchA] interface vlan-interface 3
[SwitchA-Vlan-interface3] ipv6 address 1::1 64
```

配置 VLAN 接口 3 工作在 DHCPv6 中继模式，并指定 DHCPv6 服务器地址。

```
[SwitchA-Vlan-interface3] ipv6 dhcp relay server-address 2::2
```

(2) 配置 Switch A 作为网关

配置发布 RA 消息，并配置 M 和 O 标志位。

```
[SwitchA-Vlan-interface3] undo ipv6 nd ra halt
[SwitchA-Vlan-interface3] ipv6 nd autoconfig managed-address-flag
[SwitchA-Vlan-interface3] ipv6 nd autoconfig other-flag
```

(3) 验证配置结果

完成上述配置后，查看 DHCPv6 中继指定的 DHCPv6 服务器的地址信息。

```
[SwitchA-Vlan-interface3] display ipv6 dhcp relay server-address all
Interface: Vlan3
Server address(es)                              Output Interface
2::2
```

查看 DHCPv6 中继转发报文的统计信息。

```
[SwitchA-Vlan-interface3] display ipv6 dhcp relay statistics
Packets dropped               :  0
    Error                     :  0
    Excess of rate limit      :  0
Packets received              :  14
    SOLICIT                   :  0
    REQUEST                   :  0
    CONFIRM                   :  0
    RENEW                     :  0
    REBIND                    :  0
    RELEASE                   :  0
    DECLINE                   :  0
    INFORMATION-REQUEST       :  7
    RELAY-FORWARD             :  0
    RELAY-REPLY               :  7
Packets sent                  :  14
    ADVERTISE                 :  0
    RECONFIGURE               :  0
    REPLY                     :  7
    RELAY-FORWARD             :  7
    RELAY-REPLY               :  0
```

4 DHCPv6 客户端配置

4.1 DHCPv6客户端简介

设备作为 DHCPv6 客户端时，只支持 DHCPv6 无状态配置，即只能通过 DHCPv6 获取除地址/前缀外的其他网络配置参数，不能获取 IPv6 地址和前缀。

DHCPv6 客户端通过地址无状态自动配置功能成功获取 IPv6 地址后，如果接收到的 RA 报文中 M 标志位为 0、O 标志位为 1，则设备会自动启动 DHCPv6 无状态配置功能，以获取除地址/前缀外的其他网络配置参数。

4.2 配置DHCPv6客户端

4.2.1 配置准备

为了使客户端能够通过 DHCPv6 无状态配置成功获取网络配置参数，需要确保 DHCPv6 服务器可用。

4.2.2 配置步骤

表4-1 配置 DHCPv6 客户端

操作	命令	说明
进入系统视图	**system-view**	-
使能IPv6报文转发功能	**ipv6**	必选
进入接口视图	**interface** *interface-type interface-number*	-
使能IPv6地址无状态自动配置功能	**ipv6 address auto**	必选

说明

- **ipv6 address auto** 命令的详细介绍请参见“三层技术-IP 业务命令参考”中的“IPv6 基础”。
- 建议不要在一个接口上同时配置 DHCPv6 客户端和 DHCPv6 服务器功能，也不要在一个接口上同时配置 DHCPv6 客户端和 DHCPv6 中继功能。

4.3 配置DHCPv6客户端发送的DHCPv6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 DHCPv6 客户端发送的 DHCPv6 报文的 DSCP 优先级进行配置。

表4-2 配置 DHCPv6 客户端发送的 DHCPv6 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
配置DHCPv6客户端发送的DHCPv6报文的DSCP优先级	**ipv6 dhcp client dscp** *dscp-value*	可选 缺省情况下，DHCPv6客户端发送的DHCPv6报文的DSCP优先级为56

4.4 DHCPv6客户端显示和维护

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 DHCPv6 客户端的运行情况，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 DHCPv6 客户端的统计信息。

表4-3 DHCPv6 客户端显示和维护

操作	命令
显示DHCPv6客户端的信息	**display ipv6 dhcp client** [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCPv6客户端的统计信息	**display ipv6 dhcp client statistics** [**interface** *interface-type interface-number*] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示本设备DUID	**display ipv6 dhcp duid** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除DHCPv6客户端的统计信息	**reset ipv6 dhcp client statistics** [**interface** *interface-type interface-number*]

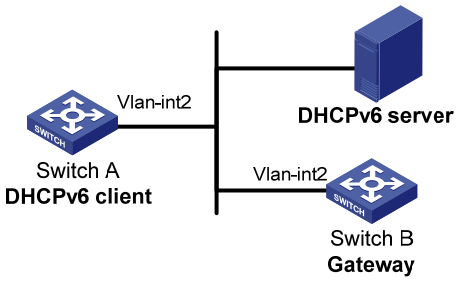
4.5 DHCPv6无状态配置典型配置举例

1. 组网需求

- Switch A 通过 DHCPv6 无状态配置获取域名服务器、域名等信息；
- Switch B 作为网关，周期性发布 RA 消息。

2. 组网图

图4-1 DHCPv6 无状态配置组网图



3. 配置步骤

(1) 配置网关 Switch B

使能 IPv6 报文转发功能。

```
<SwitchB> system-view
[SwitchB] ipv6
```

配置 VLAN 接口 2 的 IPv6 地址。

```
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ipv6 address 1::1 64
```

配置 RA 消息中 O 标志位为 1。

```
[SwitchB-Vlan-interface2] ipv6 nd autoconfig other-flag
```

配置允许发送 RA 消息。

```
[SwitchB-Vlan-interface2] undo ipv6 nd ra halt
```

(2) 配置 DHCPv6 客户端 Switch A

使能 IPv6 报文转发功能。

```
<SwitchA> system-view
[SwitchA] ipv6
```

在 VLAN 接口 2 上使能 IPv6 地址无状态自动配置功能。

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address auto
```

执行此命令后，如果 VLAN 接口 2 下没有配置地址，Switch A 会自动生成本地链路地址，并主动发送 RS（Router Solicitation，路由器请求）报文，请求网关 Switch B 立即回应 RA 报文。

4. 验证配置结果

如果收到的 RA 报文中 M 标志位为 0、O 标志位为 1，Switch A 就会启动 DHCPv6 客户端无状态配置。

可以通过 **display ipv6 dhcp client** 命令查看当前客户端的配置信息，如果从服务器成功获取了配置，将会有类似的显示信息。

```
[SwitchA-Vlan-interface2] display ipv6 dhcp client interface vlan-interface 2
Vlan-interface2 is in stateless DHCPv6 client mode
State is OPEN
Preferred Server:
    Reachable via address     :  FE80::213:7FFF:FEF6:C818
    DUID                      :  0003000100137ff6c818
```

```
    DNS servers              :  1:2:3::5
                                1:2:4::7
    Domain names             :  abc.com
                                Sysname.com
```

可以通过 **display ipv6 dhcp client statistics** 命令查看当前客户端的统计信息。

```
[SwitchA-Vlan-interface2] display ipv6 dhcp client statistics
Interface                    :  Vlan-interface2
Packets Received             :  1
        Reply                :  1
        Advertise            :  0
        Reconfigure          :  0
        Invalid              :  0
Packets Sent                 :  5
        Solicit              :  0
        Request              :  0
        Confirm              :  0
        Renew                :  0
        Rebind               :  0
        Information-request  :  5
        Release              :  0
        Decline              :  0
```

5 DHCPv6 Snooping配置

说明

- 设备只有位于 DHCPv6 客户端与 DHCPv6 服务器之间，或 DHCPv6 客户端与 DHCPv6 中继之间时，DHCPv6 Snooping 功能配置后才能正常工作；设备位于 DHCPv6 服务器与 DHCPv6 中继之间时，DHCPv6 Snooping 功能配置后不能正常工作。
- DHCPv6 Snooping 中对于接口的相关配置，目前只能在二层以太网端口或二层聚合接口上进行。关于聚合接口的详细介绍，请参见“二层技术-以太网交换配置指导”中的“以太网链路聚合”。

5.1 DHCPv6 Snooping简介

DHCPv6 Snooping 是 DHCPv6 的一种安全特性，具有如下功能：

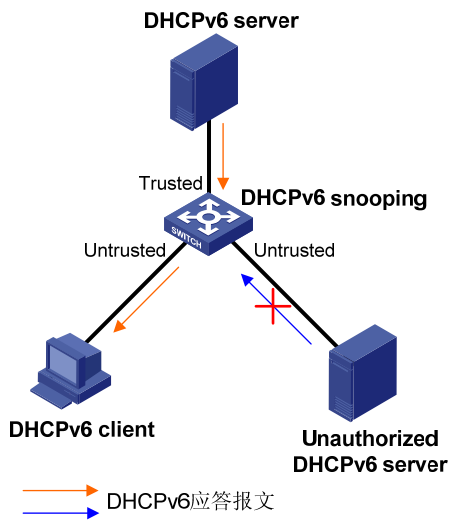
- 保证客户端从合法的服务器获取 IPv6 地址。
- 记录 DHCPv6 客户端 IPv6 地址与 MAC 地址的对应关系。

1. 保证客户端从合法的服务器获取IPv6 地址

网络中如果存在私自架设的伪 DHCPv6 服务器，则可能导致 DHCPv6 客户端获取错误的 IPv6 地址和网络配置参数，无法正常通信。为了使 DHCPv6 客户端能通过合法的 DHCPv6 服务器获取 IPv6 地址，DHCPv6 Snooping 安全机制允许将端口设置为信任端口（Trusted Port）和不信任端口（Untrusted Port）：

- 信任端口正常转发接收到的 DHCPv6 报文。
- 不信任端口接收到 DHCPv6 服务器发送的应答报文后，丢弃该报文。

图5-1 信任端口和非信任端口



连接DHCPv6 服务器、DHCPv6 中继或其他DHCPv6 Snooping设备的端口需要设置为信任端口，其他端口设置为不信任端口，从而保证DHCPv6 客户端只能从合法的DHCPv6 服务器获取地址，私自架设的伪DHCPv6 服务器无法为DHCPv6 客户端分配地址。如 图 5-1 中，将连接DHCPv6 服务器的端口设置为信任端口，其他端口设置为非信任端口。

2. 记录DHCPv6 客户端IPv6 地址与MAC地址的对应关系

DHCPv6 Snooping 通过监听 DHCPv6 报文，记录 DHCPv6 Snooping 表项，其中包括客户端的 MAC 地址、获取到的 IPv6 地址、与 DHCPv6 客户端连接的端口及该端口所属的 VLAN 等信息。网络管理员可以通过 **display ipv6 dhcp snooping user-binding** 命令查看客户端获取的 IPv6 地址信息，以便了解用户上网时所用的 IPv6 地址，并对其进行管理和监控。

5.2 使能DHCPv6 Snooping

使能 DHCPv6 Snooping 功能，必须首先在系统视图下全局使能 DHCPv6 Snooping 功能。全局使能 DHCPv6 Snooping 功能，并正确地配置信任端口和非信任端口后，可以保证客户端从合法的服务器获取 IPv6 地址。但是，此时不会记录 DHCPv6 Snooping 表项。

如果需要记录 DHCPv6 Snooping 表项，则需要在全局使能 DHCPv6 Snooping 功能的基础上，在 VLAN 视图下使能 VLAN 内的 DHCPv6 Snooping 功能。使能 VLAN 内的 DHCPv6 Snooping 功能，还可以实现 DHCPv6 Snooping 设备接收到该 VLAN 内客户端发送的请求报文后，只通过该 VLAN 内的信任端口转发该请求报文，不会通过其他非信任端口转发请求报文，以减轻网络负担。

表5-1 使能 DHCPv6 Snooping

操作	命令	说明
进入系统视图	**system-view**	-
全局使能DHCPv6 Snooping功能	**ipv6 dhcp snooping enable**	必选 缺省情况下，DHCPv6 Snooping功能处于关闭状态
进入VLAN视图	**vlan** *vlan-id*	-
在VLAN内使能DHCPv6 Snooping功能	**ipv6 dhcp snooping vlan enable**	可选 缺省情况下，VLAN内DHCPv6 Snooping功能处于关闭状态

5.3 配置DHCPv6 Snooping信任端口

DHCPv6 Snooping 将端口分为两种：

- 信任端口：正常转发接收到的 DHCPv6 报文。
- 不信任端口：接收到 DHCPv6 服务器发送的应答报文后，丢弃该报文。

使能 VLAN 内的 DHCPv6 Snooping 功能，DHCPv6 Snooping 设备接收到该 VLAN 内客户端发送的请求报文后，只通过该 VLAN 内的信任端口转发该请求报文，不会通过其他非信任端口转发请求报文，以减轻网络负担。

表5-2 配置 DHCPv6 Snooping 信任端口

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*	-
配置端口为信任端口	**ipv6 dhcp snooping trust**	必选 缺省情况下，全局使能DHCPv6 Snooping功能后，设备的所有端口均为不信任端口

说明

- 为了使 DHCPv6 客户端能从合法的 DHCPv6 服务器获取 IPv6 地址，必须将与合法 DHCPv6 服务器相连的端口设置为信任端口，且设置的信任端口和与 DHCPv6 客户端相连的端口必须在同一个 VLAN 内。
- 如果二层以太网端口加入了聚合组，则加入聚合组之前和加入聚合组之后在该接口上进行的 DHCPv6 Snooping 相关配置不会生效；该接口退出聚合组后，DHCPv6 Snooping 的配置才会生效。

5.4 配置接口动态学习DHCPv6 Snooping表项的最大数目

通过本配置可以限制接口动态学习 DHCPv6 Snooping 表项的最大数目，以防止接口学习到大量 DHCPv6 Snooping 表项，占用过多地系统资源。

表5-3 配置接口动态学习 DHCPv6 Snooping 表项的最大数目

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*	-
配置接口动态学习DHCPv6 Snooping表项的最大数目	**ipv6 dhcp snooping max-learning-num** *number*	可选 缺省情况下，不限制接口动态学习DHCPv6 Snooping表项的数目

5.5 配置DHCPv6 Snooping支持Option 18和Option 37

Option 18 称为接口 ID 选项（Interface ID）、Option 37 称为远程 ID 选项（Remote ID），DHCPv6 Snooping 设备接收到 DHCPv6 客户端发送给 DHCPv6 服务器的请求报文后，在该报文中添加 Option 18 或 Option 37，并转发给 DHCPv6 服务器。

图5-2 Option 18 选项格式

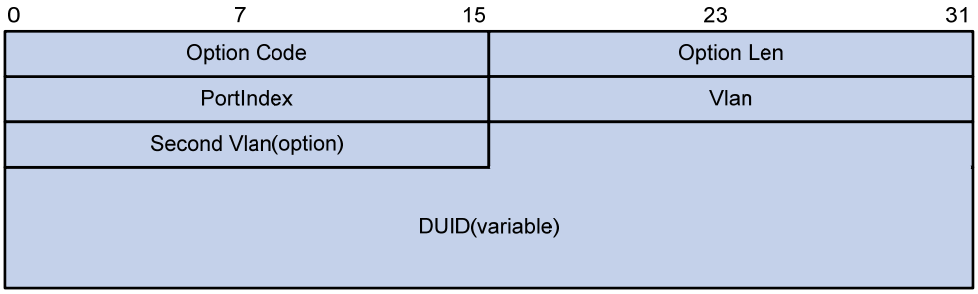
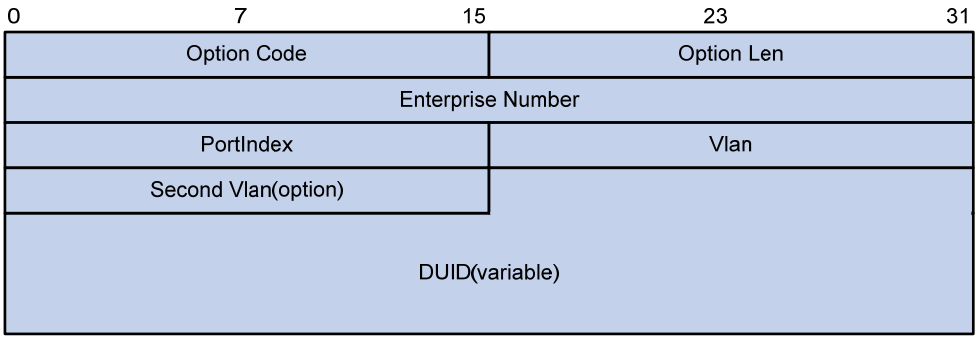


图5-3 Option 37 选项格式



说明

选项格式中的 Second Vlan 字段为可选，如果报文中不含有 Second Vlan，则 Option 18 或 Option 37 中也不包含 Second Vlan 内容。

表5-4 配置 DHCPv6 Snooping 支持 Option 18 和 Option 37

操作	命令	说明
进入系统视图	**system-view**	-
全局使能DHCPv6 Snooping功能	**ipv6 dhcp snooping enable**	必选 缺省情况下，DHCPv6 Snooping功能处于关闭状态
进入VLAN视图	**vlan** *vlan-id*	-
在VLAN内使能DHCPv6 Snooping功能	**ipv6 dhcp snooping vlan enable**	必选 缺省情况下，VLAN内DHCPv6 Snooping功能处于关闭状态
进入二层以太网端口视图或二层聚合接口视图	**interface** *interface-type interface-number*	-
使能DHCPv6 Snooping支持Option 18功能	**ipv6 dhcp snooping option interface-id enable**	必选 缺省情况下，禁止DHCPv6 Snooping支持Option 18功能

操作	命令	说明
配置Option 18选项中的DUID	**ipv6 dhcp snooping option interface-id string** *interface-id*	可选 缺省情况下，Option 18选项中的DUID为本设备的DUID
使能DHCPv6 Snooping支持Option 37功能	**ipv6 dhcp snooping option remote-id enable**	必选 缺省情况下，禁止DHCPv6 Snooping支持Option 37功能
配置Option 37选项中的DUID	**ipv6 dhcp snooping option remote-id string** *remote-id*	可选 缺省情况下，Option 37选项中的DUID为本设备的DUID

5.6 DHCPv6 Snooping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 DHCPv6 Snooping 的配置情况，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 DHCPv6 Snooping 表项信息。

表5-5 DHCPv6 Snooping 显示和维护

操作	命令
显示DHCPv6 Snooping信任端口信息	**display ipv6 dhcp snooping trust** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示DHCPv6 Snooping表项信息	**display ipv6 dhcp snooping user-binding** { *ipv6-address* | **dynamic** } [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除DHCPv6 Snooping表项	**reset ipv6 dhcp snooping user-binding** { *ipv6-address* | **dynamic** }

5.7 DHCPv6 Snooping典型配置举例

5.7.1 组网需求

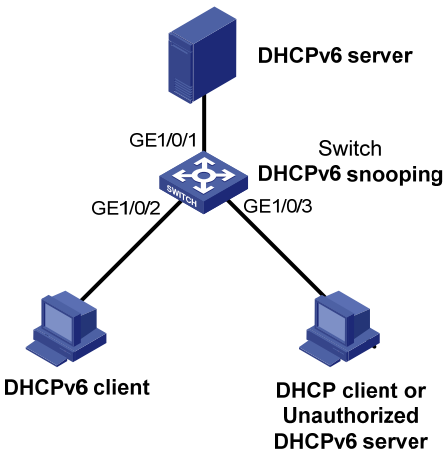
Switch 通过以太网端口 GigabitEthernet1/0/1 连接到 DHCPv6 服务器，通过以太网端口 GigabitEthernet1/0/2、GigabitEthernet1/0/3 连接到 DHCPv6 客户端。GigabitEthernet1/0/1、GigabitEthernet1/0/2 和 GigabitEthernet1/0/3 都属于 VLAN 2。

要求：

- 与 DHCPv6 服务器相连的端口可以转发 DHCPv6 服务器的响应报文，而其他端口不转发 DHCPv6 服务器的响应报文。
- 记录 DHCPv6 客户端 IPv6 地址及 MAC 地址的绑定关系。

5.7.2 组网图

图5-4 DHCPv6 Snooping 组网示意图



5.7.3 配置步骤

全局使能 DHCPv6 Snooping 功能。

```
<Switch> system-view
[Switch] ipv6 dhcp snooping enable
```

将端口 GigabitEthernet1/0/1、GigabitEthernet1/0/2 和 GigabitEthernet1/0/3 加入 VLAN 2。

```
[Switch] vlan 2
[Switch-vlan2] port GigabitEthernet 1/0/1 GigabitEthernet 1/0/2 GigabitEthernet 1/0/3
```

在 VLAN 2 内使能 DHCPv6 Snooping 功能。

```
[Switch-vlan2] ipv6 dhcp snooping vlan enable
[Switch] quit
```

配置 GigabitEthernet1/0/1 端口为信任端口。

```
[Switch] interface GigabitEthernet 1/0/1
[Switch-GigabitEthernet1/0/1] ipv6 dhcp snooping trust
```

验证配置结果。

配置完成后，通过 GigabitEthernet1/0/2 连接 DHCPv6 客户端、GigabitEthernet1/0/1 连接 DHCPv6 服务器，则可以发现 DHCPv6 客户端能够从 DHCPv6 服务器获取 IPv6 地址。通过 **display ipv6 dhcp snooping user-binding** 命令可以查看生成的 DHCPv6 Snooping 表项。如果 GigabitEthernet1/0/3 连接私自架设的伪 DHCPv6 服务器，则该服务器无法为 DHCPv6 客户端分配 IPv6 地址。

目 录

1 IPv6 域名解析配置

1.1 IPv6域名解析简介

IPv6 网络中，DNS 客户端通过 IPv6 域名解析功能实现域名与 IPv6 地址的转换。IPv6 DNS 与 IPv4 DNS 相同，分为静态域名解析和动态域名解析。两种域名解析的作用和实现方式也与 IPv4 DNS 相同。具体描述请参见“三层技术-IP 业务配置指导”中的“IPv4 域名解析”。IPv6 DNS 与 IPv4 DNS 的区别仅在于 IPv6 DNS 将域名转换为 IPv6 地址，而非 IPv4 地址。

1.2 配置IPv6 DNS client

1.2.1 配置静态域名解析

配置静态域名解析就是配置将主机名与 IPv6 地址相对应。当使用 Telnet 等应用时，可以直接使用主机名，由系统解析为 IPv6 地址。

表1-1 配置静态域名解析

操作	命令	说明
进入系统视图	**system-view**	-
配置主机名和对应的IPv6地址	**ipv6 host** *hostname ipv6-address*	必选 缺省情况下，静态域名解析表中没有主机名及IPv6地址的对应关系

说明

- 每个主机名只能对应一个 IPv6 地址，当对同一主机名进行多次配置时，最后配置的 IPv6 地址有效。
- 设备上最多可配置 50 条 IPv6 静态域名解析信息。

1.2.2 配置动态域名解析

如果用户需要使用动态域名解析功能，可以使用下面的命令使能动态域名解析功能，并配置域名服务器，这样才能将查询请求报文发送到正确的服务器进行解析。

用户还可以配置域名后缀，以便实现只输入域名的部分字段，而由系统自动加上预先设置的后缀进行解析。

表1-2 配置动态域名解析

操作	命令	说明
进入系统视图	**system-view**	-

操作	命令	说明
使能动态域名解析功能	**dns resolve**	必选 缺省情况下，动态域名解析功能处于关闭状态
配置域名服务器的IPv6地址	**dns server ipv6** *ipv6-address* [*interface-type interface-number*]	必选 缺省情况下，没有配置域名服务器的IPv6地址 当域名服务器的IPv6地址为链路本地地址时，需要指定参数*interface-type*和*interface-number*
配置域名后缀	**dns domain** *domain-name*	可选 缺省情况下，没有配置域名后缀，即只根据用户输入的域名信息进行解析

说明

- **dns resolve** 和 **dns domain** 命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“IPv4 域名解析”。
- 包括 IPv4 域名服务器在内，设备上最多可配置 6 个域名服务器。
- 设备上最多可以配置 10 个域名后缀。

1.3 配置IPv6 DNS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 DNS 报文的 DSCP 优先级进行配置。

表1-3 配置 IPv6 DNS 报文的 DSCP 优先级

操作	命令	说明
进入系统视图	**system-view**	-
配置发送的IPv6 DNS报文的DSCP优先级	**dns ipv6 dscp** *dscp-value*	可选 缺省情况下，发送的IPv6 DNS报文的DSCP优先级为0

1.4 IPv6域名解析显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示 IPv6 域名解析配置后的运行情况，通过查看显示信息验证配置的效果。

在用户视图下，执行 **reset** 命令可以清除动态域名缓存信息。

表1-4 域名解析显示和维护

操作	命令
显示IPv6静态域名解析表	**display ipv6 host** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6域名服务器信息	**display dns ipv6 server** [**dynamic**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示域名后缀列表信息	**display dns domain** [**dynamic**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示IPv6动态域名缓存信息	**display dns host ipv6** [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除IPv6动态域名缓存信息	**reset dns host ipv6**

说明

display dns domain、**display dns host ipv6** 和 **reset dns host ipv6** 命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“IPv4 域名解析”。

1.5 IPv6域名解析典型配置举例

1.5.1 静态域名解析配置举例

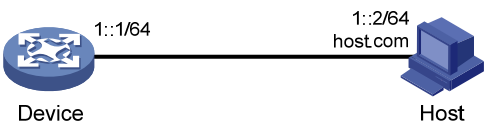
1. 组网需求

为了避免记忆复杂的 IPv6 地址，Device 希望通过便于记忆的主机名访问某一主机。在 Device 上手工配置 IPv6 地址对应的主机名，利用静态域名解析功能，就可以实现通过主机名访问该主机。

在本例中，Device 访问的主机 IPv6 地址为 1::2，主机名为 host.com。

2. 组网图

图1-1 静态域名解析配置组网图



3. 配置步骤

配置主机名 host.com 对应的 IPv6 地址为 1::2。

```
<Device> system-view
[Device] ipv6 host host.com 1::2
```

使能 IPv6 报文转发功能。

```
[Device] ipv6
```

执行 ping ipv6 host.com 命令，Device 通过静态域名解析可以解析到 host.com 对应的 IPv6 地址为 1::2。

```
[Device] ping ipv6 host.com
  PING host.com (1::2):
  56  data bytes, press CTRL_C to break
    Reply from 1::2
    bytes=56 Sequence=1 hop limit=64  time = 3 ms
    Reply from 1::2
    bytes=56 Sequence=2 hop limit=64  time = 1 ms
    Reply from 1::2
    bytes=56 Sequence=3 hop limit=64  time = 1 ms
    Reply from 1::2
    bytes=56 Sequence=4 hop limit=64  time = 2 ms
    Reply from 1::2
    bytes=56 Sequence=5 hop limit=64  time = 2 ms
  --- host.com ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 1/1/3 ms
```

1.5.2 动态域名解析配置举例

1. 组网需求

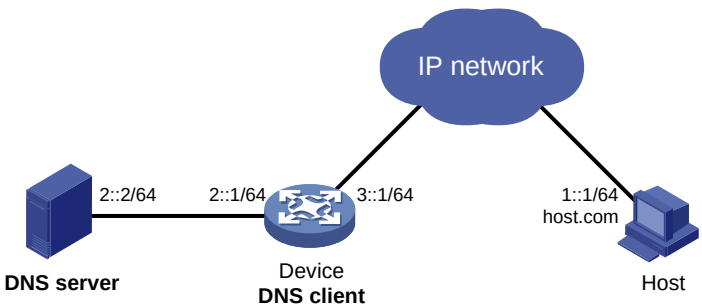
为了避免记忆复杂的 IPv6 地址，Device 希望通过便于记忆的域名访问某一主机。如果网络中存在域名服务器，则可以利用动态域名解析功能，实现通过域名访问主机。

在本例中：

- 域名服务器的 IPv6 地址是 2::2/64，域名服务器上存在 com 域，且 com 域中包含域名“host”和 IPv6 地址 1::1/64 的对应关系。
- Device 作为 DNS 客户端，使用动态域名解析功能，将域名解析为 IP 地址。
- Device 上配置域名后缀 com，以便简化访问主机时输入的域名，例如通过输入 host 即可访问域名为 host.com、IPv6 地址为 1::1/64 的主机 Host。

2. 组网图

图1-2 动态域名解析组网图



3. 配置步骤

说明

- 在开始下面的配置之前，假设设备与主机之间的路由可达，设备和主机都已经配置完毕，接口IPv6地址如 图 1-2 所示。
- 不同域名服务器的配置方法不同，下面仅以 Windows Server 2003 为例，说明域名服务器的配置方法。配置之前，需确保 DNS 服务器支持 IPv6 DNS 功能，以便处理 IPv6 域名解析报文；且 DNS 服务器的接口可以转发 IPv6 报文。

(1) 配置域名服务器

进入域名服务器配置界面。

在开始菜单中，选择[程序/管理工具/DNS]。

创建区域 com。

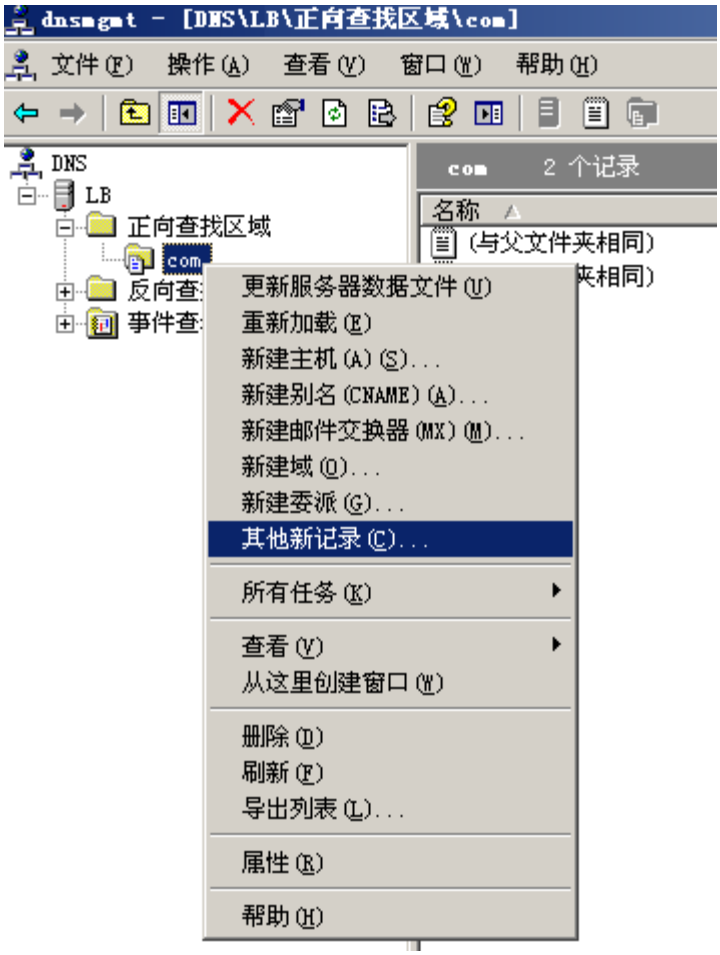
如 图 1-3 所示，右键点击[正向查找区域]，选择[新建区域]，按照提示创建新的区域com。

图1-3 创建区域

添加域名和 IPv6 地址的映射。

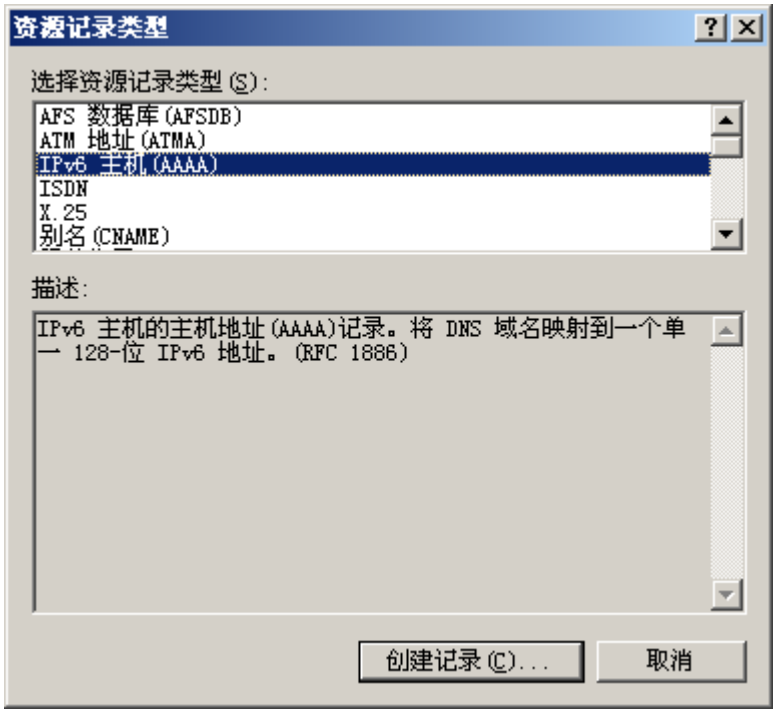
如 图 1-4 所示，右键点击区域com。

图1-4 创建记录



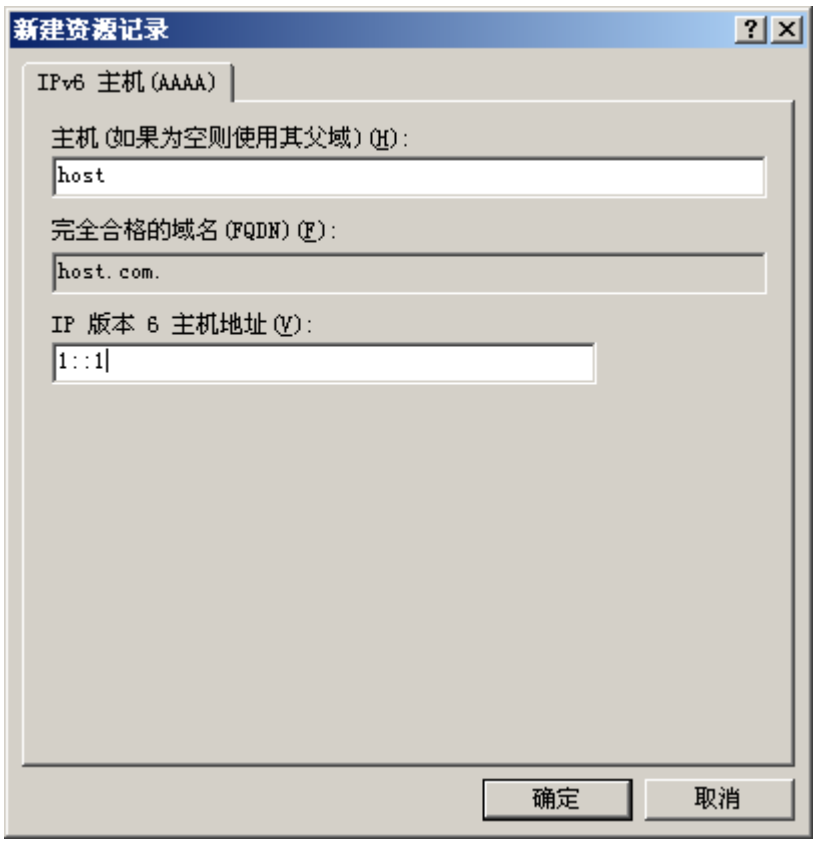
选择[其他新记录]，弹出如 图 1-5 的对话框，选择资源记录类型为“IPv6 主机（AAAA）”。

图1-5 选择资源记录类型



按照 图 1-6 输入域名host和IPv6 地址 1::1。点击<确定>按钮，添加域名和IPv6 地址的映射。

图1-6 添加域名和 IPv6 地址的映射



(2) 配置 DNS 客户端 Device

开启动态域名解析功能。

```
<Device> system-view
[Device] dns resolve
```

配置域名服务器的 IPv6 地址为 2::2。

```
[Device] dns server ipv6 2::2
```

配置域名后缀 com。

```
[Device] dns domain com
```

(3) 验证配置结果

在设备上执行 ping ipv6 host 命令，可以 ping 通主机，且对应的目的地址为 1::1。

```
[Device] ping ipv6 host
 Trying DNS resolve, press CTRL_C to break
 Trying DNS server (2::2)
  PING host.com (1::1):
  56  data bytes, press CTRL_C to break
    Reply from 1::1
    bytes=56 Sequence=1 hop limit=126  time = 2 ms
    Reply from 1::1
    bytes=56 Sequence=2 hop limit=126  time = 1 ms
    Reply from 1::1
    bytes=56 Sequence=3 hop limit=126  time = 1 ms
    Reply from 1::1
    bytes=56 Sequence=4 hop limit=126  time = 1 ms
    Reply from 1::1
    bytes=56 Sequence=5 hop limit=126  time = 1 ms

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

目 录

1 隧道配置

1.1 隧道概述

隧道技术是一种封装技术，它利用一种网络协议来传输另一种网络协议，即一种网络协议将其他网络协议的数据报文封装在自己的报文中，然后在网络中传输。封装后的数据报文在网络中传输的路径，称为隧道。隧道是一条虚拟的点对点连接，隧道的两端需要对数据报文进行封装及解封装。隧道技术就是指包括数据封装、传输和解封装在内的全过程。

隧道技术可以：

- 作为过渡技术，实现 IPv4 和 IPv6 网络互通，如 IPv6 over IPv4 隧道技术。
- 创建 VPN（Virtual Private Network，虚拟专用网络），保证通信的安全性，如 IPv4 over IPv4 隧道、IPv4/IPv6 over IPv6 隧道、GRE（Generic Routing Encapsulation，通用路由封装）、DVPN（Dynamic Virtual Private Network，动态虚拟专用网络）和 IPsec 隧道技术。
- 实现流量工程，避免由于负载不均衡导致网络拥塞，如 MPLS TE（Multiprotocol Label Switching Traffic Engineering，多协议标记交换流量工程）。

本文只介绍 IPv6 over IPv4 隧道、IPv4 over IPv4 隧道、IPv4 over IPv6 隧道和 IPv6 over IPv6 隧道。如无特殊说明，下文中的隧道技术均指此类隧道。

1.1.1 IPv6 over IPv4 隧道

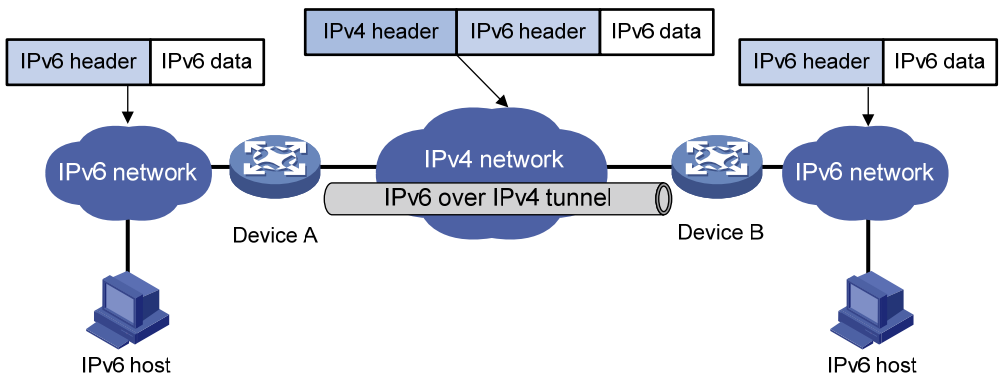
1. IPv6 over IPv4 隧道原理

IPv6 over IPv4 隧道是在IPv6 数据报文前封装上IPv4 的报文头，通过隧道使IPv6 报文穿越IPv4 网络，实现隔离的IPv6 网络的互通，如 图 1-1 所示。IPv6 over IPv4 隧道可以建立在主机—主机、主机—设备或设备—设备之间。隧道的终点可能是IPv6 报文的最终目的地，也可能需要进一步转发。

说明

IPv6 over IPv4 隧道两端的设备必须支持 IPv4/IPv6 双协议栈。

图1-1 IPv6 over IPv4 隧道原理图



IPv6 over IPv4 隧道对报文的处理过程如下：

(1) IPv6 网络中的设备发送 IPv6 报文，该报文到达隧道的源端设备 Device A。

(2) Device A 根据路由表判定该报文要通过隧道进行转发后，在 IPv6 报文前封装上 IPv4 的报文头，通过隧道的实际物理接口将报文转发出去。

(3) 封装报文通过隧道到达隧道目的端设备 Device B，Device B 判断该封装报文的目的地是本设备后，将对报文进行解封装。

(4) Device B 根据解封装后的 IPv6 报文的目的地址转发该 IPv6 报文。如果目的地就是本设备，则将 IPv6 报文转给上层协议处理。

2. IPv6 over IPv4 隧道模式

根据隧道终点的 IPv4 地址的获取方式不同，隧道分为“配置隧道”和“自动隧道”。

- 如果 IPv6 over IPv4 隧道的终点地址不能从 IPv6 报文的目的地址中自动获取，需要进行手工配置，这样的隧道称为“配置隧道”。
- 如果 IPv6 over IPv4 隧道的终点地址采用内嵌 IPv4 地址的特殊 IPv6 地址形式，则可以从 IPv6 报文的目的地址中自动获取隧道终点的 IPv4 地址，这样的隧道称为“自动隧道”。

如 表 1-1 所示，根据对IPv6 报文的封装方式的不同，IPv6 over IPv4 隧道分为以下几种模式。表 1-1 中还列举了各隧道模式的关键配置参数。

表1-1 IPv6 over IPv4 隧道模式

隧道类型	隧道模式	隧道源/目的地址	隧道接口地址
配置隧道	IPv6手动隧道	源/目的地址为手动配置的IPv4地址	IPv6地址
自动隧道	6to4隧道	源地址为手动配置的IPv4地址，目的地址不需配置	6to4地址，其格式为： 2002:IPv4-source-address::/48
	ISATAP（Intra-Site Automatic Tunnel Addressing Protocol，站点内自动隧道寻址协议）隧道	源地址为手动配置的IPv4地址，目的地址不需配置	ISATAP地址，其格式为： Prefix:0:5EFE:IPv4-source-address/64

(1) IPv6 手动隧道

手动隧道是点到点之间的链路，一条链路就是一个单独的隧道。主要用于边缘路由器—边缘路由器或主机—边缘路由器之间定期安全通信的稳定连接，可实现与远端 IPv6 网络的连接。

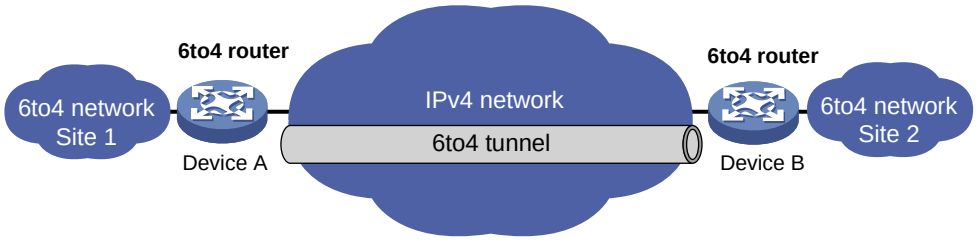
(2) 6to4 隧道

6to4 隧道是点到多点的自动隧道，主要建立在边缘路由器之间，主要用于将多个 IPv6 孤岛通过 IPv4 网络连接到 IPv6 网络。6to4 隧道通过在 IPv6 报文的目的地址中嵌入 IPv4 地址，来实现自动获取隧道终点的 IPv4 地址。

6to4 隧道采用特殊的 6to4 地址，其格式为：2002:abcd:efgh:子网号::接口 ID/64，其中 2002 表示固定的 IPv6 地址前缀，abcd:efgh 表示该 6to4 隧道对应的 32 位全球唯一的 IPv4 地址，用 16 进制表示（如 1.1.1.1 可以表示为 0101:0101）。2002:abcd:efgh 之后的部分唯一标识了一个主机在 6to4 网络内的位置。通过这个嵌入的 IPv4 地址可以自动确定隧道的终点，使隧道的建立非常方便。

由于 6to4 地址的 64 位地址前缀中的 16 位子网号可以由用户自定义，前缀中的前 48 位已由固定数值、隧道起点或终点设备的 IPv4 地址确定，使 IPv6 报文通过隧道进行转发成为可能。

图1-2 6to4 隧道原理图



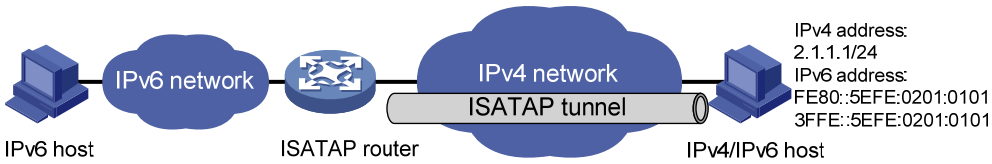
(3) ISATAP 隧道

随着 IPv6 技术的推广，现有的 IPv4 网络中将会出现越来越多的 IPv6 主机，ISATAP 隧道技术为这种应用提供了一个较好的解决方案。ISATAP 隧道是点到多点的自动隧道技术，通过在 IPv6 报文的目的地址中嵌入的 IPv4 地址，可以自动获取隧道的终点。

使用 ISATAP 隧道时，IPv6 报文的目的地址和隧道接口的 IPv6 地址都要采用特殊的 ISATAP 地址。ISATAP 地址格式为：Prefix(64bit):0:5EFE:abcd:efgh。其中，64 位的 Prefix 为任何合法的 IPv6 单播地址前缀，abcd:efgh 表示 32 位 IPv4 源地址，用 16 进制表示（如 1.1.1.1 可以表示为 0101:0101），该 IPv4 地址不要求全球唯一。通过这个嵌入的 IPv4 地址就可以自动建立隧道，完成 IPv6 报文的传送。

ISATAP 隧道主要用于在 IPv4 网络中 IPv6 路由器—IPv6 路由器、IPv6 主机—IPv6 路由器的连接。

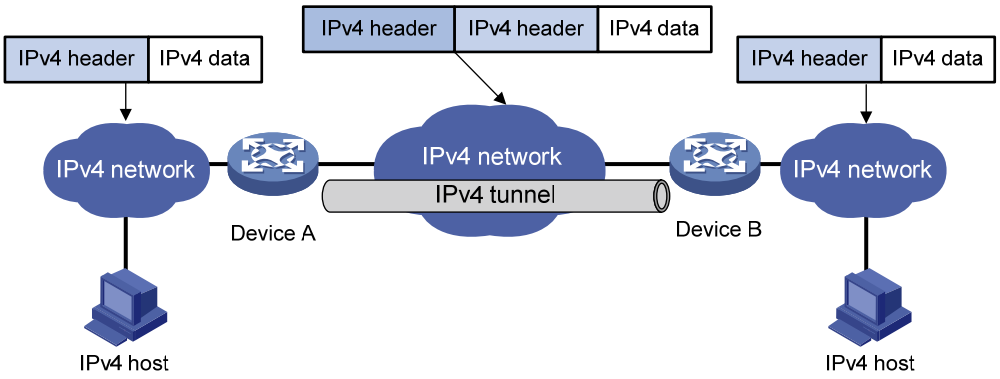
图1-3 ISATAP 隧道原理图



1.1.2 IPv4 over IPv4 隧道

IPv4 over IPv4 隧道（RFC 1853）是对 IP 数据报进行封装，使得一个 IPv4 网络的数据能够在另一个 IPv4 网络中传输。

图1-4 IPv4 over IPv4 隧道原理图



报文在隧道中传输经过封装与解封装两个过程，以上图为例说明这两个过程：

- 封装过程

Device A 连接 IPv4 主机所在子网的接口收到 IP 数据报后，首先交由 IP 协议栈处理。IP 协议栈根据 IP 报头中的目的地址来确定如何转发此包。如果报文的目的地址为与 Device B 相连的 IPv4 主机的地址，则将此报文发给 Device A 上连接 Device B 的 Tunnel 接口。

Tunnel 接口收到此包后，进行 IPv4 over IPv4 的封装，封装完成后重新交给 IP 协议栈处理，IP 协议栈根据添加的 IP 报头确定出接口。

- 解封装过程

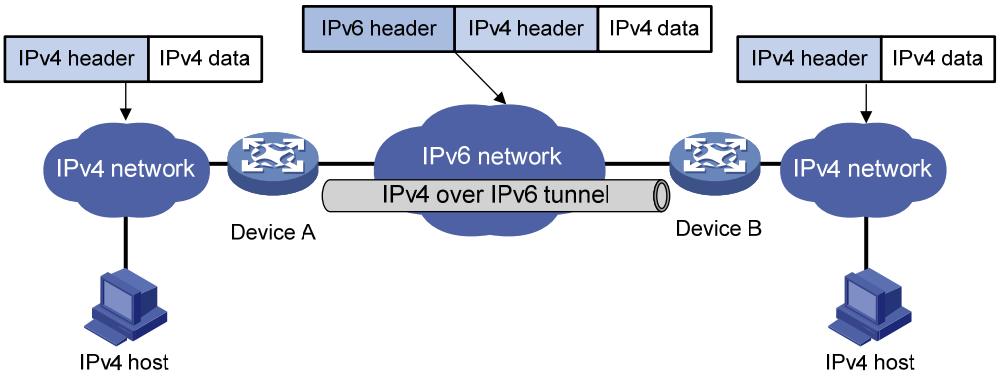
解封装过程和封装的过程相反。从网络接口收到的 IP 报文被送到 IP 协议栈。IP 协议栈检查接收到的 IP 报文头中的协议号。如果协议号为 4（表示封装的报文为 IPv4 报文），则将此 IP 数据包发送到隧道模块进行解封装处理。解封装之后的 IP 报文将重新被送到 IP 协议栈进行处理。

1.1.3 IPv4 over IPv6 隧道

随着 IPv6 网络的广泛部署，IPv6 网络将逐渐取代 IPv4 网络，占据主导地位。尚未被 IPv6 网络取代的 IPv4 网络将形成孤岛，需要通过 IPv6 网络互通。IPv4 over IPv6 隧道用来实现通过 IPv6 网络连接隔离的 IPv4 网络孤岛。

IPv4 over IPv6 隧道是在 IPv4 报文上封装 IPv6 的报文头，通过隧道使 IPv4 报文穿越 IPv6 网络，从而实现隔离的 IPv4 网络通过 IPv6 网络互通。

图1-5 IPv4 over IPv6 隧道原理图



IPv4 报文在隧道中传输经过封装与解封装两个过程，以 图 1-5 为例说明这两个过程：

- 封装过程

Device A 连接 IPv4 网络的接口收到 IPv4 报文后，首先交由 IPv4 协议栈处理。IPv4 协议栈根据 IPv4 报头中的目的地址来确定如何转发此包。如果报文的目的地址为与 Device B 相连的 IPv4 主机的地址，则将此报文发给 Device A 上连接 Device B 的 Tunnel 接口。

Tunnel 接口收到此报文后添加 IPv6 报文头，封装完成后交给 IPv6 模块处理。IPv6 协议模块根据 IPv6 报文头的目的地址重新确定如何转发此数据包。

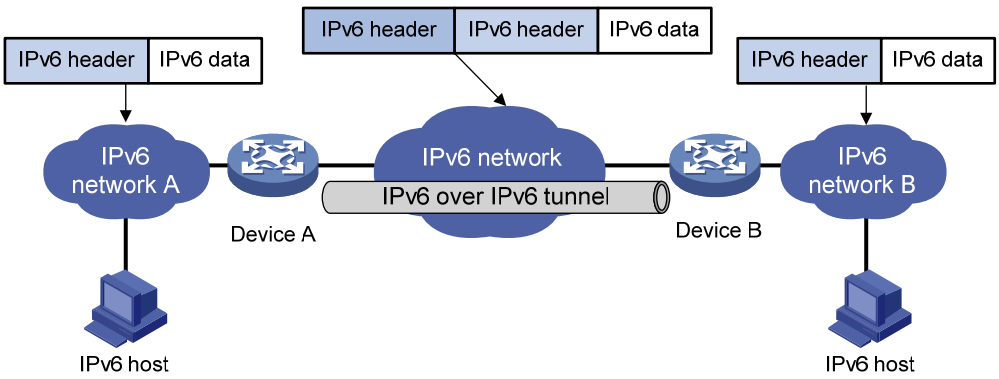
- 解封装流程

解封装过程和封装的过程相反。从连接 IPv6 网络的接口接收到 IPv6 报文后，将其送到 IPv6 协议模块。IPv6 协议模块检查 IPv6 报文封装的协议类型。若封装的协议为 IPv4，则报文进入隧道处理模块进行解封装处理。解封装之后的 IPv4 报文被送往 IPv4 协议模块进行二次路由处理。

1.1.4 IPv6 over IPv6 隧道

IPv6 over IPv6 隧道（RFC 2473）是对 IPv6 数据报进行封装，使这些被封装的数据报能够在另一个 IPv6 网络中传输，封装后的数据报文即 IPv6 隧道报文。

图1-6 IPv6 over IPv6 隧道原理图



IPv6 报文在隧道中传输经过封装与解封装两个过程，以 图 1-6 为例说明这两个过程：

- 封装过程

Device A 连接网络 A 的接口收到 IPv6 报文后，首先交由 IPv6 协议模块处理，并确定如何转发此报文。若此报文的目的地址为与 Device B 相连的主机的地址，则将此报文发给 Device A 上连接 Device B 的 Tunnel 接口。

Tunnel 口收到此报文后添加 IPv6 报文头，封装完成后交给 IPv6 模块处理。IPv6 协议模块根据 IPv6 报文头的目的地址重新确定如何转发此报文。

- 解封装流程

解封装过程和封装的过程相反。从 IPv6 网络接口接收的报文被送到 IPv6 协议模块。IPv6 协议模块检查 IPv6 报文封装的协议类型。若封装的协议为 IPv6，则报文进入隧道处理模块进行解封装处理；解封装之后的报文被送往相应的协议模块进行二次路由处理。

1.1.5 协议规范

与隧道技术相关的协议规范有：

- RFC 1853：IP in IP Tunneling
- RFC 2473：Generic Packet Tunneling in IPv6 Specification
- RFC 2893：Transition Mechanisms for IPv6 Hosts and Routers
- RFC 3056：Connection of IPv6 Domains via IPv4 Clouds
- RFC 4214：Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)

1.2 隧道配置任务简介

表1-2 隧道配置任务简介

配置任务		说明	详细配置
配置Tunnel接口		必选	1.3
配置IPv6 over IPv4隧道	配置IPv6手动隧道	根据组网情况，选择其一	1.4
	配置6to4隧道		1.5
	配置ISATAP隧道		1.6
配置IPv4 over IPv4隧道			1.7
配置IPv4 over IPv6隧道			1.8
配置IPv6 over IPv6隧道			1.9

1.3 配置Tunnel接口

隧道两端的设备上，需要创建虚拟的三层接口——Tunnel 接口，以便隧道两端的设备利用隧道发送报文、识别并处理来自隧道的报文。

1.3.1 配置准备

配置 Tunnel 接口前，需要先创建业务类型为 Tunnel 的业务环回组，并将设备上未使用的二层以太网端口加入该业务环回组。

表1-3 配置 Tunnel 接口

配置步骤	命令	说明
进入系统视图	**system-view**	-
创建Tunnel接口，并进入Tunnel接口视图	**interface tunnel** *number*	必选 缺省情况下，设备上无Tunnel接口
配置接口描述信息	**description** *text*	可选 缺省情况下，接口描述信息为“*该接口的接口名* Interface”
指定Tunnel接口引用的业务环回组ID	**service-loopback-group** *number*	必选 缺省情况下，隧道未引用任何业务环回组

配置步骤	命令	说明
配置MTU值	**mtu** *mtu-size*	可选 缺省情况下，Tunnel接口的MTU值和Tunnel接口的状态有关 • 如果 Tunnel 接口状态为 down，Tunnel 接口的 MTU 值显示为 64000 字节 • 如果 Tunnel 接口状态为 up，Tunnel 接口的 MTU 值由系统自动计算生成 需要注意的是： • 如果配置 Tunnel 接口的 MTU 值，请保证所有配置的 Tunnel 接口使用相同的 MTU 值 • 在任意一个 Tunnel 接口下配置的 MTU，对所有已创建的 Tunnel 接口都有效 • Tunnel 接口的 MTU 可多次配置，但是只有最后一次配置生效
配置Tunnel接口的带宽	**tunnel bandwidth** *bandwidth-value*	可选 缺省情况下，Tunnel接口的带宽为64kbps
恢复当前接口的缺省配置	**default**	可选
关闭Tunnel接口	**shutdown**	可选 缺省情况下，接口处于开启状态

说明

- 对于交换机，封装后的报文不能根据目的地址和路由表进行第二次三层转发，需要将封装后的报文发送给业务环回口，由业务环回口将报文回送给转发模块后，再进行三层转发。因此，需要指定隧道接口引用的业务环回组，以实现隧道报文的接收和发送。隧道接口引用的业务环回组必须已创建，否则隧道接口状态不会 up，隧道无法通讯。关于业务环回组的创建和配置，请参见“二层技术-以太网交换配置指导”中的“业务环回组”。
- 目前，通过 **tunnel bandwidth** 命令配置的 Tunnel 接口带宽只用于动态路由协议计算隧道所在路径的 cost 值，不会影响接口的实际带宽。建议根据报文实际出接口的带宽值设置 Tunnel 接口带宽。
- 在设备上配置对 IPv6 单播数据报文进行封装的隧道（如 IPv6 over IPv4 隧道、IPv6 over IPv6 隧道），请确保配置的 MTU 值不小于 1280 字节。
- 缺省情况下，ICMP 目的不可达报文发送功能处于关闭状态，为了使设备能够发送 ICMP 差错报文，用户需要通过 **ip unreachables enable** 命令开启 ICMP 目的不可达报文发送功能。

1.4 配置IPv6手动隧道

1.4.1 配置准备

设备上存在已经配置 IP 地址、能够进行正常通讯的接口（如 VLAN 接口，Loopback 接口等），该接口将作为 Tunnel 接口的源接口。

1.4.2 配置IPv6 手动隧道

表1-4 配置 IPv6 手动隧道

操作		命令	说明
进入系统视图		**system-view**	-
使能IPv6报文转发功能		**ipv6**	必选 缺省情况下，关闭IPv6报文转发功能
进入Tunnel接口视图		**interface tunnel** *number*	-
设置Tunnel接口的IPv6地址	配置IPv6全球单播地址或站点本地地址	**ipv6 address** { *ipv6-address prefix-length* | *ipv6-addressIprefix-length* }	二者必选其一 缺省情况下，Tunnel接口上没有设置IPv6全球单播地址或站点本地地址
		ipv6 address *ipv6-addressIprefix-length* **eui-64**	
	配置IPv6链路本地地址	**ipv6 address auto link-local**	可选 缺省情况下，当接口配置了IPv6全球单播地址或站点本地地址后，会自动生成链路本地地址
		ipv6 address *ipv6-address* **link-local**	
配置隧道模式为IPv6手动隧道		**tunnel-protocol ipv6-ipv4**	必选 缺省情况下，为GRE over IPv4隧道模式 在隧道的两端应配置相同的隧道模式，否则可能造成报文传输失败
设置Tunnel接口的源端地址或接口		**source** { *ip-address* | *interface-type interface-number* }	必选 缺省情况下，Tunnel接口上没有设置源端地址和接口
设置Tunnel接口的目的端地址		**destination** *ip-address*	必选 缺省情况下，Tunnel接口上没有设置目的端地址
退回系统视图		**quit**	-
配置丢弃含有IPv4兼容IPv6地址的IPv6报文		**tunnel discard ipv4-compatible-packet**	可选 缺省情况下，不会丢弃含有IPv4兼容IPv6地址的IPv6报文

说明

- 以上各项 Tunnel 接口下进行的功能特性配置，在删除 Tunnel 接口后，该接口上的所有配置也将被删除。
- Tunnel 接口上设置的源端和目的端必须是公网的地址或接口。
- 如果封装前 IPv6 报文的目的 IPv6 地址与 Tunnel 接口的 IPv6 地址不在同一个网段，则必须配置通过 Tunnel 接口到达目的 IPv6 地址的转发路由，以便需要进行封装的报文能正常转发。用户可以配置静态路由，指定到达目的 IPv6 地址的路由出接口为本端 Tunnel 接口或下一跳为对端 Tunnel 接口地址。用户也可以配置动态路由，在 Tunnel 接口使能动态路由协议。在隧道的两端都要进行此项配置，配置的详细情况请参见“三层技术-IP 路由配置指导”中的“IPv6 静态路由”或其他路由协议配置。

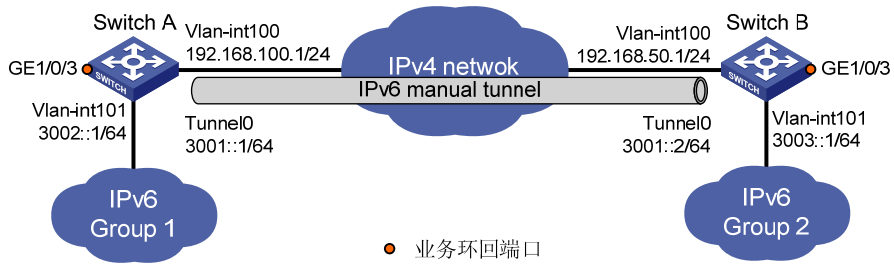
1.4.3 配置举例

1. 组网需求

如 图 1-7 所示，两个IPv6 网络分别通过Switch A和Switch B与IPv4 网络连接，要求在Switch A和Switch B之间建立IPv6 over IPv4 隧道，使两个IPv6 网络可以互通。如果隧道终点的IPv4 地址不能从IPv6 报文的目的地址中自动获取，则需要配置IPv6 手动隧道。

2. 组网图

图1-7 IPv6 手动隧道组网图



3. 配置步骤

说明

在开始下面的配置之前，请确保 Switch A 和 Switch B 上已经创建相应的 VLAN 接口，且两者之间 IPv4 报文路由可达。

(1) 配置 Switch A

使能 IPv6 转发功能。

```
<SwitchA> system-view
[SwitchA] ipv6
```

配置接口 Vlan-interface100 的地址。

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ip address 192.168.100.1 255.255.255.0
[SwitchA-Vlan-interface100] quit
```
配置接口 Vlan-interface101 的 IPv6 地址。
```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] ipv6 address 3002::1 64
[SwitchA-Vlan-interface101] quit
```
创建业务环回组 1，并配置服务类型为 tunnel。
```
[SwitchA] service-loopback group 1 type tunnel
```
将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。
```
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```
配置手动隧道。
```
[SwitchA] interface tunnel 0
[SwitchA-Tunnel0] ipv6 address 3001::1/64
[SwitchA-Tunnel0] source vlan-interface 100
[SwitchA-Tunnel0] destination 192.168.50.1
[SwitchA-Tunnel0] tunnel-protocol ipv6-ipv4
```
在 Tunnel 接口视图下指定隧道引用业务环回组 1。
```
[SwitchA-Tunnel0] service-loopback-group 1
[SwitchA-Tunnel0] quit
```
配置从 Switch A 经过 Tunnel0 接口到 Group 2 的静态路由。
```
[SwitchA] ipv6 route-static 3003:: 64 tunnel 0
```
(2)　配置 Switch B

使能 IPv6 转发功能。
```
<SwitchB> system-view
[SwitchB] ipv6
```
配置接口 Vlan-interface100 的地址。
```
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ip address 192.168.50.1 255.255.255.0
[SwitchB-Vlan-interface100] quit
```
配置接口 Vlan-interface101 的 IPv6 地址。
```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] ipv6 address 3003::1 64
[SwitchB-Vlan-interface101] quit
```
创建业务环回组 1，并配置服务类型为 tunnel。
```
[SwitchB] service-loopback group 1 type tunnel
```
将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。
```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] undo stp enable
[SwitchB-GigabitEthernet1/0/3] undo ndp enable
```

```
[SwitchB-GigabitEthernet1/0/3] undo lldp enable
[SwitchB-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchB-GigabitEthernet1/0/3] quit
```

配置手动隧道。

```
[SwitchB] interface tunnel 0
[SwitchB-Tunnel0] ipv6 address 3001::2/64
[SwitchB-Tunnel0] source vlan-interface 100
[SwitchB-Tunnel0] destination 192.168.100.1
[SwitchB-Tunnel0] tunnel-protocol ipv6-ipv4
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchB-Tunnel0] service-loopback-group 1
[SwitchB-Tunnel0] quit
```

配置从 Switch B 经过 Tunnel0 接口到 Group 1 的静态路由。

```
[SwitchB] ipv6 route-static 3002:: 64 tunnel 0
```

4. 验证配置结果

完成以上配置之后，分别查看 Switch A 和 Switch B 的 Tunnel 接口状态如下：

```
[SwitchA] display ipv6 interface tunnel 0
Tunnel0 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::C0A8:6401
  Global unicast address(es):
    3001::1, subnet is 3001::/64
  Joined group address(es):
    FF02::1:FF00:0
    FF02::1:FF00:1
    FF02::1:FFA8:6401
    FF02::2
    FF02::1
  MTU is 1480 bytes
  ND reachable time is 30000 milliseconds
  ND retransmit interval is 1000 milliseconds
  Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
  InReceives:                   55
……（略）
[SwitchB] display ipv6 interface tunnel 0
Tunnel0 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::C0A8:3201
  Global unicast address(es):
    3001::2, subnet is 3001::/64
  Joined group address(es):
    FF02::1:FF00:0
    FF02::1:FF00:1
    FF02::1:FFA8:3201
    FF02::2
    FF02::1
```

```
  MTU is 1480 bytes
  ND reachable time is 30000 milliseconds
  ND retransmit interval is 1000 milliseconds
  Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
  InReceives:                   55
……（略）
```

从 Switch A 上可以 Ping 通对端的 Vlan-int101 接口的 IPv6 地址：

```
[SwitchA] ping ipv6 3003::1
  PING 3003::1 : 56  data bytes, press CTRL_C to break
    Reply from 3003::1
    bytes=56 Sequence=1 hop limit=64  time = 1 ms
    Reply from 3003::1
    bytes=56 Sequence=2 hop limit=64  time = 1 ms
    Reply from 3003::1
    bytes=56 Sequence=3 hop limit=64  time = 1 ms
    Reply from 3003::1
    bytes=56 Sequence=4 hop limit=64  time = 1 ms
    Reply from 3003::1
    bytes=56 Sequence=5 hop limit=64  time = 1 ms

  --- 3003::1 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 1/1/1 ms
```

1.5 配置6to4隧道

1.5.1 配置准备

设备上存在已经配置 IP 地址、能够进行正常通讯的接口（如 VLAN 接口，Loopback 接口等），该接口将作为 Tunnel 接口的源接口。

1.5.2 配置 6to4 隧道

表1-5 配置 6to4 隧道

操作	命令	说明
进入系统视图	**system-view**	-
使能IPv6报文转发功能	**ipv6**	必选 缺省情况下，关闭IPv6报文转发功能
进入Tunnel接口视图	**interface tunnel** *number*	-

操作		命令	说明
设置Tunnel接口的IPv6地址	配置IPv6全球单播地址或站点本地地址	**ipv6 address** { *ipv6-address prefix-length* | *ipv6-addressIprefix-length* }	二者必选其一 缺省情况下，Tunnel接口上没有设置IPv6全球单播地址或站点本地地址
		ipv6 address *ipv6-addressIprefix-length* **eui-64**	
	配置IPv6链路本地地址	**ipv6 address auto link-local**	可选 缺省情况下，当接口配置了IPv6全球单播地址或站点本地地址后，会自动生成链路本地地址
		ipv6 address *ipv6-address* **link-local**	
配置隧道模式为6to4隧道		**tunnel-protocol ipv6-ipv4 6to4**	必选 缺省情况下，为GRE over IPv4隧道模式 在隧道的两端应配置相同的隧道模式，否则可能造成报文传输失败
设置Tunnel接口的源端地址或接口		**source** { *ip-address* | *interface-type interface-number* }	必选 缺省情况下，Tunnel接口上没有设置源端地址和接口
退回系统视图		**quit**	-
配置丢弃含有IPv4兼容IPv6地址的IPv6报文		**tunnel discard ipv4-compatible-packet**	可选 缺省情况下，不会丢弃含有IPv4兼容IPv6地址的IPv6报文

说明

- Tunnel 接口上设置的源端地址或接口必须是公网的地址或接口。
- 6to4 隧道不需要配置目的地址，因为隧道的目的地址可以通过 6to4 IPv6 地址中嵌入的 IPv4 地址自动获得。
- 如果封装前 IPv6 报文的目的 IPv6 地址与 Tunnel 接口的 IPv6 地址不在同一个网段，则必须配置通过 Tunnel 接口到达目的 IPv6 地址的转发路由，以便需要进行封装的报文能正常转发。对于自动隧道，用户只能配置静态路由，指定到达目的 IPv6 地址的路由出接口为本端 Tunnel 接口或下一跳为对端 Tunnel 接口地址，不支持动态路由。在隧道的两端都要进行转发路由的配置，配置的详细情况请参见“三层技术-IP 路由配置指导”中的“IPv6 静态路由”。
- 对于自动隧道，使用同种封装协议的 Tunnel 接口不能同时配置完全相同的源地址。

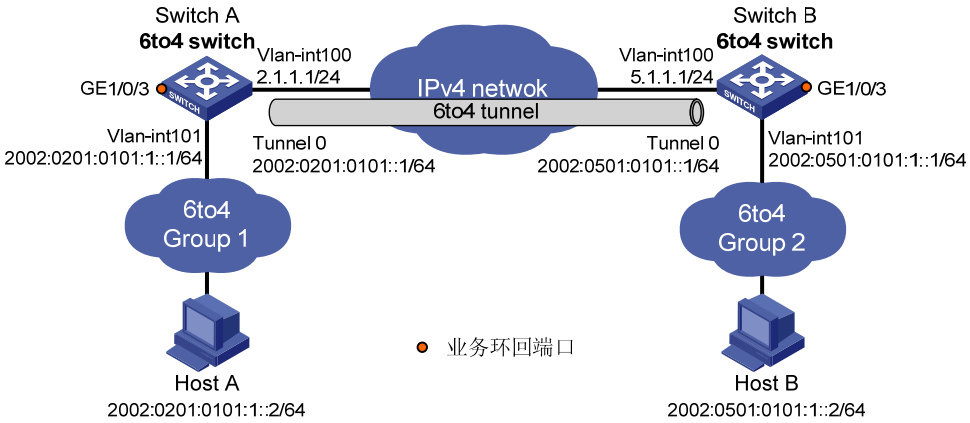
1.5.3 配置 6to4 隧道举例

1. 组网需求

如 图 1-8 所示，两个 6to4 网络通过网络边缘 6to4 switch（Switch A和Switch B）与IPv4 网络相连。在Switch A和Switch B之间建立 6to4 隧道，实现 6to4 网络中的主机Host A和Host B之间的互通。

2. 组网图

图1-8 6to4 隧道组网图



3. 配置思路

为了实现 6to4 网络之间的互通，除了配置 6to4 隧道外，还需要为 6to4 网络内的主机及 6to4 switch 配置 6to4 地址。

- Switch A 上接口 Vlan-int100 的 IPv4 地址为 2.1.1.1/24，转换成 IPv6 地址后使用 6to4 前缀 2002:0201:0101::/48。对此前缀进行子网划分，Tunnel0 使用 2002:0201:0101::/64 子网，Vlan-int101 使用 2002:0201:0101:1::/64 子网。
- Switch B 上接口 Vlan-int100 的 IPv4 地址为 5.1.1.1/24，转换成 IPv6 地址后使用 6to4 前缀 2002:0501:0101::/48。对此前缀进行子网划分，Tunnel0 使用 2002:0501:0101::/64 子网，Vlan-int101 使用 2002:0501:0101:1::/64 子网。

4. 配置步骤

说明

在开始下面的配置之前，请确保 Switch A 和 Switch B 上已经创建相应的 VLAN 接口，且两者之间 IPv4 报文路由可达。

(1) 配置 Switch A

使能 IPv6 转发功能。

```
<SwitchA> system-view
[SwitchA] ipv6
```

配置接口 Vlan-interface100 的地址。

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ip address 2.1.1.1 24
[SwitchA-Vlan-interface100] quit
```

配置接口 Vlan-interface101 的地址。

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] ipv6 address 2002:0201:0101:1::1/64
[SwitchA-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchA] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```

配置 6to4 隧道。

```
[SwitchA] interface tunnel 0
[SwitchA-Tunnel0] ipv6 address 2002:201:101::1/64
[SwitchA-Tunnel0] source vlan-interface 100
[SwitchA-Tunnel0] tunnel-protocol ipv6-ipv4 6to4
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchA-Tunnel0] service-loopback-group 1
[SwitchA-Tunnel0] quit
```

配置到目的地址 2002::/16，下一跳为 Tunnel 接口的静态路由。

```
[SwitchA] ipv6 route-static 2002:: 16 tunnel 0
```

(2) 配置 Switch B

使能 IPv6 转发功能。

```
<SwitchB> system-view
[SwitchB] ipv6
```

配置接口 Vlan-interface100 的地址。

```
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ip address 5.1.1.1 24
[SwitchB-Vlan-interface100] quit
```

配置接口 Vlan-interface101 的地址。

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] ipv6 address 2002:0501:0101:1::1/64
[SwitchB-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchB] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] undo stp enable
[SwitchB-GigabitEthernet1/0/3] undo ndp enable
[SwitchB-GigabitEthernet1/0/3] undo lldp enable
[SwitchB-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchB-GigabitEthernet1/0/3] quit
```

配置 6to4 隧道。

```
[SwitchB] interface tunnel 0
[SwitchB-Tunnel0] ipv6 address 2002:0501:0101::1/64
[SwitchB-Tunnel0] source vlan-interface 100
[SwitchB-Tunnel0] tunnel-protocol ipv6-ipv4 6to4
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchB-Tunnel0] service-loopback-group 1
[SwitchB-Tunnel0] quit
```

配置到目的地址 2002::/16，下一跳为 Tunnel 接口的静态路由。

```
[SwitchB] ipv6 route-static 2002:: 16 tunnel 0
```

5. 验证配置结果

完成以上配置之后，Host A 与 Host B 可以互相 Ping 通。

```
D:\>ping6 -s 2002:201:101:1::2 2002:501:101:1::2

Pinging 2002:501:101:1::2
from 2002:201:101:1::2 with 32 bytes of data:

Reply from 2002:501:101:1::2: bytes=32 time=13ms
Reply from 2002:501:101:1::2: bytes=32 time=1ms
Reply from 2002:501:101:1::2: bytes=32 time=1ms
Reply from 2002:501:101:1::2: bytes=32 time<1ms

Ping statistics for 2002:501:101:1::2:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 0ms, Maximum = 13ms, Average = 3ms
```

1.6 配置ISATAP隧道

1.6.1 配置准备

设备上存在已经配置 IP 地址、能够进行正常通讯的接口（如 VLAN 接口，Loopback 接口等），该接口将作为 Tunnel 接口的源接口。

1.6.2 配置ISATAP隧道

表1-6 配置 ISATAP 隧道

操作		命令	说明
进入系统视图		**system-view**	-
使能IPv6报文转发功能		**ipv6**	必选 缺省情况下，IPv6报文转发功能处于关闭状态
进入Tunnel接口视图		**interface tunnel** *number*	-
设置Tunnel接口的IPv6地址	配置IPv6全球单播地址或站点本地地址	**ipv6 address** { *ipv6-address prefix-length* | *ipv6-addressIprefix-length* }	二者必选其一 缺省情况下，Tunnel接口上没有设置IPv6全球单播地址或站点本地地址
		ipv6 address *ipv6-addressIprefix-length* **eui-64**	
	配置IPv6链路本地地址	**ipv6 address auto link-local**	可选

操作		命令	说明
		ipv6 address *ipv6-address* **link-local**	缺省情况下，当接口配置了IPv6全球单播地址或站点本地地址后，会自动生成链路本地地址
配置隧道模式为ISATAP隧道		**tunnel-protocol ipv6-ipv4 isatap**	必选 缺省情况下，为GRE over IPv4隧道模式 在隧道的两端应配置相同的隧道模式，否则可能造成报文传输失败
设置Tunnel接口的源端地址或接口		**source** { *ip-address* | *interface-type interface-number* }	必选 缺省情况下，Tunnel接口上没有设置源端地址和接口
退回系统视图		**quit**	-
配置丢弃含有IPv4兼容IPv6地址的IPv6报文		**tunnel discard ipv4-compatible-packet**	可选 缺省情况下，不会丢弃含有IPv4兼容IPv6地址的IPv6报文

说明

- Tunnel 接口上设置的源端地址或接口必须是公网的地址或接口。
- ISATAP 隧道不需要配置目的地址，因为隧道的目的地址可以通过 ISATAP 地址中嵌入的 IPv4 地址自动获得。
- 如果封装前 IPv6 报文的目的 IPv6 地址与 Tunnel 接口的 IPv6 地址不在同一个网段，则必须配置通过 Tunnel 接口到达目的 IPv6 地址的转发路由，以便需要进行封装的报文能正常转发。对于自动隧道，用户只能配置静态路由，指定到达目的 IPv6 地址的路由出接口为本端 Tunnel 接口或下一跳为对端 Tunnel 接口地址，不支持动态路由。在隧道的两端都要进行转发路由的配置，配置的详细情况请参见“三层技术-IP 路由配置指导”中的“IPv6 静态路由”。
- 对于自动隧道，使用同种封装协议的 Tunnel 接口不能同时配置完全相同的源地址。

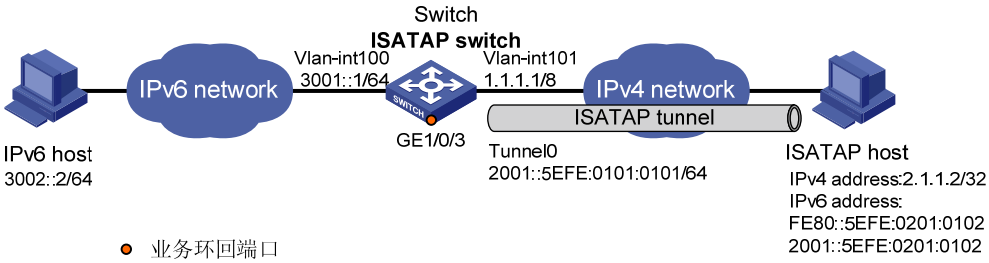
1.6.3 配置举例

1. 组网需求

如 图 1-9 所示，IPv6 网络和IPv4 网络通过ISATAP交换机相连，在IPv4 网络侧分布着一些IPv6 主机。要求将IPv4 网络中的IPv6 主机通过ISATAP隧道接入到IPv6 网络。

2. 组网图

图1-9 ISATAP 隧道组网图



3. 配置步骤

说明

在开始下面的配置之前，请确保 Switch 上已经创建相应的 VLAN 接口，且 Switch 的 Vlan-interface101 和 ISATAP host 之间 IPv4 报文路由可达。

(1) 配置 Switch

使能 IPv6 转发功能。

```
<Switch> system-view
[Switch] ipv6
```

配置各接口地址。

```
[Switch] interface vlan-interface 100
[Switch-Vlan-interface100] ipv6 address 3001::1/64
[Switch-Vlan-interface100] quit
[Switch] interface vlan-interface 101
[Switch-Vlan-interface101] ip address 1.1.1.1 255.0.0.0
[Switch-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[Switch] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[Switch] interface GigabitEthernet 1/0/3
[Switch-GigabitEthernet1/0/3] undo stp enable
[Switch-GigabitEthernet1/0/3] undo ndp enable
[Switch-GigabitEthernet1/0/3] undo lldp enable
[Switch-GigabitEthernet1/0/3] port service-loopback group 1
[Switch-GigabitEthernet1/0/3] quit
```

配置 ISATAP 隧道。

```
[Switch] interface tunnel 0
[Switch-Tunnel0] ipv6 address 2001::5efe:0101:0101 64
[Switch-Tunnel0] source vlan-interface 101
[Switch-Tunnel0] tunnel-protocol ipv6-ipv4 isatap
```

取消对 RA 消息发布的抑制，使主机可以通过交换机发布的 RA 消息获取地址前缀等信息。

```
[Switch-Tunnel0] undo ipv6 nd ra halt
```
在 Tunnel 接口视图下指定隧道引用业务环回组 1。
```
[Switch-Tunnel0] service-loopback-group 1
[Switch-Tunnel0] quit
```
配置到 ISATAP 主机的静态路由。
```
[Switch] ipv6 route-static 2001:: 16 tunnel 0
```
(2) 配置 ISATAP 主机

ISATAP 主机上的具体配置与主机的操作系统有关，下面仅以 Windows XP 操作系统为例进行说明。

在主机上安装 IPv6 协议。
```
C:\>ipv6 install
```
在 Windows XP 上，ISATAP 接口通常为接口 2，只要在该接口上配置 ISATAP 交换机的 IPv4 地址即可完成主机侧的配置。先看看这个 ISATAP 接口的信息：
```
C:\>ipv6 if 2
Interface 2: Automatic Tunneling Pseudo-Interface
  Guid {48FCE3FC-EC30-E50E-F1A7-71172AEEE3AE}
  does not use Neighbor Discovery
  does not use Router Discovery
  routing preference 1
  EUI-64 embedded IPv4 address: 0.0.0.0
  router link-layer address: 0.0.0.0
    preferred link-local fe80::5efe:2.1.1.2, life infinite
  link MTU 1280 (true link MTU 65515)
  current hop limit 128
  reachable time 42500ms (base 30000ms)
  retransmission interval 1000ms
  DAD transmits 0
  default site prefix length 48
```
它自动生成了一个 ISATAP 格式的 link-local 地址（fe80::5efe:2.1.1.2）。我们需要设置这个接口上的 ISATAP 交换机的 IPv4 地址：
```
C:\>ipv6 rlu 2 1.1.1.1
```
只需要这么一个命令，这就完成了主机的配置，我们再来看看这个 ISATAP 接口的信息：
```
C:\>ipv6 if 2
Interface 2: Automatic Tunneling Pseudo-Interface
  Guid {48FCE3FC-EC30-E50E-F1A7-71172AEEE3AE}
  does not use Neighbor Discovery
  uses Router Discovery
  routing preference 1
  EUI-64 embedded IPv4 address: 2.1.1.2
  router link-layer address: 1.1.1.1
    preferred global 2001::5efe:2.1.1.2, life 29d23h59m46s/6d23h59m46s (public)
    preferred link-local fe80::5efe:2.1.1.2, life infinite
  link MTU 1500 (true link MTU 65515)
  current hop limit 255
  reachable time 42500ms (base 30000ms)
  retransmission interval 1000ms
  DAD transmits 0
```

```
  default site prefix length 48
```

对比前后的区别，我们可以看到主机获取了 2001::/64 的前缀，自动生成地址 2001::5efe:2.1.1.2，同时还会发现这么一行“uses Router Discovery”表明主机启用了路由器发现，这时 ping 一下交换机上隧道接口的 IPv6 地址，可以 ping 通，这时候表明 ISATAP 隧道已经成功建立。

```
C:\>ping 2001::5efe:1.1.1.1

Pinging 2001::5efe:1.1.1.1 with 32 bytes of data:

Reply from 2001::5efe:1.1.1.1: time=1ms
Reply from 2001::5efe:1.1.1.1: time=1ms
Reply from 2001::5efe:1.1.1.1: time=1ms
Reply from 2001::5efe:1.1.1.1: time=1ms

Ping statistics for 2001::5efe:1.1.1.1:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 1ms, Maximum = 1ms, Average = 1ms
```

4. 验证配置结果

完成以上配置之后，ISATAP 主机就可访问 IPv6 网络中的主机。

1.7 配置IPv4 over IPv4隧道

1.7.1 配置准备

设备上存在已经配置 IP 地址、能够进行正常通讯的接口（如 VLAN 接口，Loopback 接口等），该接口将作为 Tunnel 接口的源接口。

1.7.2 配置IPv4 over IPv4 隧道

表1-7 配置 IPv4 over IPv4 隧道

操作	命令	说明
进入系统视图	**system-view**	-
进入Tunnel接口视图	**interface tunnel** *number*	-
设置Tunnel接口的IPv4地址	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**]	必选 缺省情况下，Tunnel接口上没有设置IPv4地址
配置隧道模式为IPv4 over IPv4隧道	**tunnel-protocol ipv4-ipv4**	必选 缺省情况下，为GRE over IPv4隧道模式 在隧道的两端应配置相同的隧道模式，否则会造成报文传输失败
设置Tunnel接口的源端地址或接口	**source** { *ip-address* | *interface-type interface-number* }	必选 缺省情况下，Tunnel接口上没有设置源端地址和接口

操作	命令	说明
设置Tunnel接口的目的端地址	**destination** *ip-address*	必选 缺省情况下，Tunnel接口上没有设置目的端地址

说明

- Tunnel 接口上设置的源端和目的端必须是公网的地址或接口。
- 如果封装前 IPv4 报文的目的 IPv4 地址与 Tunnel 接口的 IPv4 地址不在同一个网段，则必须配置通过 Tunnel 接口到达目的 IPv4 地址的转发路由，以便需要进行封装的报文能正常转发。用户可以配置静态路由，指定到达目的 IPv4 地址的路由出接口为本端 Tunnel 接口或下一跳为对端 Tunnel 接口地址。用户也可以配置动态路由，在 Tunnel 接口使能动态路由协议。在隧道的两端都要进行转发路由的配置，配置的详细情况请参见“三层技术-IP 路由配置指导”中的“静态路由”或其他路由协议配置。
- 本端隧道接口的 IPv4 地址与隧道的目的地址不能在同一个网段内。
- 配置经过隧道接口的路由时，路由的目的地址不能与该隧道的目的地址在同一个网段内。
- 对两个或两个以上使用同种封装协议的 Tunnel 接口，不能同时配置完全相同的源地址和目的地址。
- 配置 Tunnel 接口的源端地址时，若采用配置源接口形式，则 Tunnel 的源地址取的是源接口的主 IP 地址。

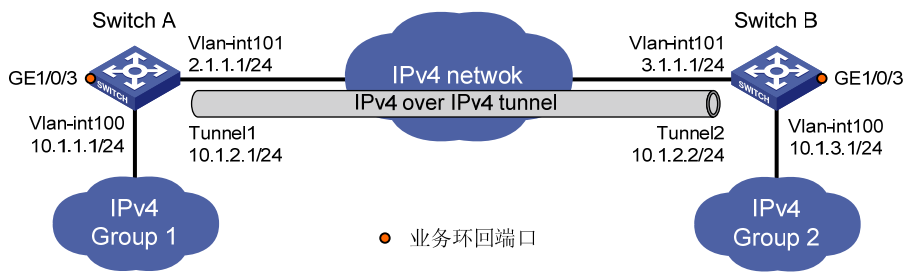
1.7.3 配置举例

1. 组网需求

运行 IP 协议的两个子网 Group 1 和 Group 2 位于不同的区域，这两个子网都使用私网地址。通过在交换机 Switch A 和交换机 Switch B 之间建立 IPv4 over IPv4 隧道，实现两个子网的互联。

2. 组网图

图1-10 IPv4 over IPv4 隧道组网图



3. 配置步骤

说明

在开始下面的配置之前，请确保 Switch A 和 Switch B 上已经创建相应的 VLAN 接口，且两者之间 IPv4 报文路由可达。

(1) 配置 Switch A

配置接口 Vlan-interface100。

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ip address 10.1.1.1 255.255.255.0
[SwitchA-Vlan-interface100] quit
```

配置接口 Vlan-interface101（隧道的实际物理接口）。

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] ip address 2.1.1.1 255.255.255.0
[SwitchA-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchA] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```

创建 Tunnel1 接口。

```
[SwitchA] interface tunnel 1
```

配置 Tunnel1 接口的 IP 地址。

```
[SwitchA-Tunnel1] ip address 10.1.2.1 255.255.255.0
```

配置 Tunnel 封装模式。

```
[SwitchA-Tunnel1] tunnel-protocol ipv4-ipv4
```

配置 Tunnel1 接口的源地址（Vlan-interface101 的 IP 地址）。

```
[SwitchA-Tunnel1] source 2.1.1.1
```

配置 Tunnel1 接口的目的地址（Switch B 的 Vlan-interface101 的 IP 地址）。

```
[SwitchA-Tunnel1] destination 3.1.1.1
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchA-Tunnel1] service-loopback-group 1
[SwitchA-Tunnel1] quit
```

配置从 Switch A 经过 Tunnel1 接口到 Group 2 的静态路由。

```
[SwitchA] ip route-static 10.1.3.0 255.255.255.0 tunnel 1
```

(2) 配置 Switch B

配置接口 Vlan-interface100。

```
<SwitchB> system-view
```

```
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ip address 10.1.3.1 255.255.255.0
[SwitchB-Vlan-interface100] quit
```

配置接口 Vlan-interface101（隧道的实际物理接口）。

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] ip address 3.1.1.1 255.255.255.0
[SwitchB-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchB] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] undo stp enable
[SwitchB-GigabitEthernet1/0/3] undo ndp enable
[SwitchB-GigabitEthernet1/0/3] undo lldp enable
[SwitchB-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchB-GigabitEthernet1/0/3] quit
```

创建 Tunnel2 接口。

```
[SwitchB] interface tunnel 2
```

配置 Tunnel2 接口的 IP 地址。

```
[SwitchB-Tunnel2] ip address 10.1.2.2 255.255.255.0
```

配置 Tunnel 封装模式。

```
[SwitchB-Tunnel2] tunnel-protocol ipv4-ipv4
```

配置 Tunnel2 接口的源地址（Vlan-interface101 的 IP 地址）。

```
[SwitchB-Tunnel2] source 3.1.1.1
```

配置 Tunnel2 接口的目的地址（SwitchA 的 Vlan-interface101 的 IP 地址）。

```
[SwitchB-Tunnel2] destination 2.1.1.1
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchB-Tunnel2] service-loopback-group 1
[SwitchB-Tunnel2] quit
```

配置从 Switch B 经过 Tunnel2 接口到 Group 1 的静态路由。

```
[SwitchB] ip route-static 10.1.1.0 255.255.255.0 tunnel 2
```

4. 验证配置结果

完成以上配置之后，分别察看 Switch A 和 Switch B 的 Tunnel 接口状态如下：

```
[SwitchA] display interface tunnel 1
Tunnel1 current state: UP
Line protocol current state: UP
Description: Tunnel1 Interface
The Maximum Transmit Unit is 1480
Internet Address is 10.1.2.1/24 Primary
Encapsulation is TUNNEL, service-loopback-group ID is 1.
Tunnel source 2.1.1.1（Vlan-interface101）, destination 3.1.1.1
Tunnel bandwidth 64 (kbps)
Tunnel protocol/transport IP/IP
Last clearing of counters:  Never
    Last 300 seconds input:  0 bytes/sec, 0 packets/sec
```

```
    Last 300 seconds output:  2 bytes/sec, 0 packets/sec
    4 packets input,  256 bytes
    0 input error
    12 packets output,  768 bytes
    0 output error

[SwitchB] display interface tunnel 2
Tunnel2 current state: UP
Line protocol current state: UP
Description: Tunnel2 Interface
The Maximum Transmit Unit is 1480
Internet Address is 10.1.2.2/24 Primary
Encapsulation is TUNNEL, service-loopback-group ID is 1.
Tunnel source 3.1.1.1（Vlan-interface101）, destination 2.1.1.1
Tunnel bandwidth 64 (kbps)
Tunnel protocol/transport IP/IP
Last clearing of counters:  Never
    Last 300 seconds input:  0 bytes/sec, 0 packets/sec
    Last 300 seconds output:  0 bytes/sec, 0 packets/sec
    5 packets input,  320 bytes
    0 input error
    9 packets output,  576 bytes
    0 output error
```

从 Switch A 可以 Ping 通对端的 Vlan-interface100 接口的 IPv4 地址：

```
[SwitchA] ping 10.1.3.1
  PING 10.1.3.1: 56  data bytes, press CTRL_C to break
    Reply from 10.1.3.1: bytes=56 Sequence=1 ttl=255 time=15 ms
    Reply from 10.1.3.1: bytes=56 Sequence=2 ttl=255 time=15 ms
    Reply from 10.1.3.1: bytes=56 Sequence=3 ttl=255 time=16 ms
    Reply from 10.1.3.1: bytes=56 Sequence=4 ttl=255 time=16 ms
    Reply from 10.1.3.1: bytes=56 Sequence=5 ttl=255 time=15 ms

  --- 10.1.3.1 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 15/15/16 ms
```

1.8 配置IPv4 over IPv6隧道

1.8.1 配置准备

设备上存在已经配置 IPv6 地址、能够进行正常通讯的接口（如 VLAN 接口，Loopback 接口等），该接口将作为 Tunnel 接口的源接口。

1.8.2 配置IPv4 over IPv6 隧道

表1-8 配置 IPv4 over IPv6 隧道

操作	命令	说明
进入系统视图	**system-view**	-
使能IPv6报文转发功能	**ipv6**	必选 缺省情况下，关闭IPv6报文转发功能
进入Tunnel接口视图	**interface tunnel** *number*	-
设置Tunnel接口的IPv4地址	**ip address** *ip-address* { *mask* | *mask-length* } [**sub**]	必选 缺省情况下，Tunnel接口上没有设置IPv4地址
配置隧道模式为IPv4 over IPv6隧道	**tunnel-protocol ipv4-ipv6**	必选 缺省情况下，为GRE over IPv4隧道模式 在隧道的两端应配置相同的隧道模式，否则可能造成报文传输失败
设置Tunnel接口的源端地址或接口	**source** { *ipv6-address* | *interface-type interface-number* }	必选 缺省情况下，Tunnel接口上没有设置源端地址和接口
设置Tunnel接口的目的端地址	**destination** *ipv6-address*	必选 缺省情况下，Tunnel接口上没有设置目的端地址

说明

- Tunnel 接口上设置的源端和目的端必须是公网的地址或接口。
- 如果封装前 IPv4 报文的目的 IPv4 地址与 Tunnel 接口的 IPv4 地址不在同一个网段，则必须配置通过 Tunnel 接口到达目的 IPv4 地址的转发路由，以便需要进行封装的报文能正常转发。用户可以配置静态路由，指定到达目的 IPv4 地址的路由出接口为本端 Tunnel 接口或下一跳为对端 Tunnel 接口地址。用户也可以配置动态路由，在 Tunnel 接口使能动态路由协议。在隧道的两端都要进行转发路由的配置，配置的详细情况请参见“三层技术-IP 路由配置指导”中的“静态路由”或其他路由协议配置。
- 对两个或两个以上使用同种封装协议的 Tunnel 接口，不能同时配置完全相同的源地址和目的地址。
- 配置 Tunnel 接口的源端地址时，若采用配置源接口形式，则 Tunnel 的源地址取的是源接口的主 IP 地址。

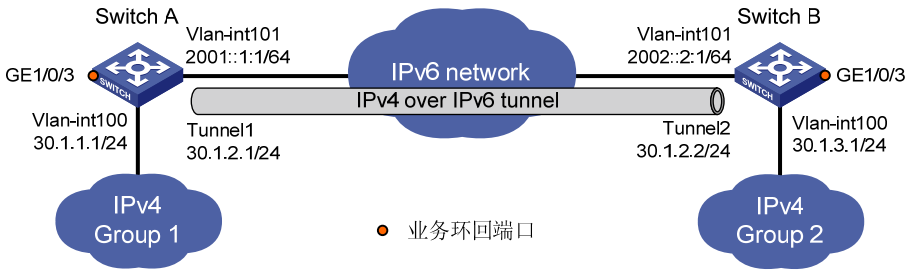
1.8.3 配置举例

1. 组网需求

运行 IP 协议的两个子网 Group 1 和 Group 2 通过 IPv6 网络相连。通过在交换机 Switch A 和交换机 Switch B 之间建立 IPv4 over IPv6 隧道，实现两个子网穿越 IPv6 网络互联。

2. 组网图

图1-11 IPv4 over IPv6 隧道组网图



3. 配置步骤

说明

在开始下面的配置之前，请确保 Switch A 和 Switch B 上已经创建相应的 VLAN 接口，且两者之间 IPv6 报文路由可达。

(1) 配置 Switch A

使能 IPv6 转发功能。

```
<SwitchA> system-view
[SwitchA] ipv6
```

配置接口 Vlan-interface100。

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ip address 30.1.1.1 255.255.255.0
[SwitchA-Vlan-interface100] quit
```

配置接口 Vlan-interface101（隧道的实际物理接口）。

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] ipv6 address 2001::1:1 64
[SwitchA-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchA] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```

创建 Tunnel1 接口。

```
[SwitchA] interface tunnel 1
```

配置 Tunnel1 接口的 IP 地址。

```
[SwitchA-Tunnel1] ip address 30.1.2.1 255.255.255.0
```

配置 Tunnel 封装模式。

```
[SwitchA-Tunnel1] tunnel-protocol ipv4-ipv6
```

配置 Tunnel1 接口的源地址（Vlan-interface101 的 IP 地址）。

```
[SwitchA-Tunnel1] source 2001::1:1
```

配置 Tunnel1 接口的目的地址（Switch B 的 Vlan-interface101 的 IP 地址）。

```
[SwitchA-Tunnel1] destination 2002::2:1
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchA-Tunnel1] service-loopback-group 1
[SwitchA-Tunnel1] quit
```

配置从 Switch A 经过 Tunnel1 接口到 Group 2 的静态路由。

```
[SwitchA] ip route-static 30.1.3.0 255.255.255.0 tunnel 1
```

(2) 配置 Switch B

使能 IPv6 转发功能。

```
<SwitchB> system-view
[SwitchB] ipv6
```

配置接口 Vlan-interface100。

```
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ip address 30.1.3.1 255.255.255.0
[SwitchB-Vlan-interface100] quit
```

配置接口 Vlan-interface101（隧道的实际物理接口）。

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] ipv6 address 2002::2:1 64
[SwitchB-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchB] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] undo stp enable
[SwitchB-GigabitEthernet1/0/3] undo ndp enable
[SwitchB-GigabitEthernet1/0/3] undo lldp enable
[SwitchB-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchB-GigabitEthernet1/0/3] quit
```

创建 Tunnel2 接口。

```
[SwitchB] interface tunnel 2
```

配置 Tunnel2 接口的 IP 地址。

```
[SwitchB-Tunnel2] ip address 30.1.2.2 255.255.255.0
```

配置 Tunnel 封装模式。

```
[SwitchB-Tunnel2] tunnel-protocol ipv4-ipv6
```

配置 Tunnel2 接口的源地址（Vlan-interface101 的 IP 地址）。

```
[SwitchB-Tunnel2] source 2002::2:1
```

配置 Tunnel2 接口的目的地址（Switch A 的 Vlan-interface101 的 IP 地址）。

```
[SwitchB-Tunnel2] destination 2001::1:1
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchB-Tunnel2] service-loopback-group 1
[SwitchB-Tunnel2] quit
```

配置从 Switch B 经过 Tunnel2 接口到 Group 1 的静态路由。

```
[SwitchB] ip route-static 30.1.1.0 255.255.255.0 tunnel 2
```

4. 验证配置结果

完成以上配置之后，分别查看 Switch A 和 Switch B 的 Tunnel 接口状态如下：

```
[SwitchA] display interface tunnel 1
Tunnel1 current state: UP
Line protocol current state: UP
Description: Tunnel1 Interface
The Maximum Transmit Unit is 1460
Internet Address is 30.1.2.1/24 Primary
Encapsulation is TUNNEL, service-loopback-group ID is 1.
Tunnel source 2002::0001:0001, destination 2002::0002:0001
Tunnel bandwidth 64 (kbps)
Tunnel protocol/transport IP/IPv6
Last clearing of counters:  Never
    Last 300 seconds input:  0 bytes/sec, 0 packets/sec
    Last 300 seconds output:  0 bytes/sec, 0 packets/sec
    152 packets input,  9728 bytes
    0 input error
    168 packets output,  10752 bytes
    0 output error

[SwitchB] display interface tunnel 2
Tunnel2 current state: UP
Line protocol current state: UP
Description: Tunnel2 Interface
The Maximum Transmit Unit is 1460
Internet Address is 30.1.2.2/24 Primary
Encapsulation is TUNNEL, service-loopback-group ID is 1.
Tunnel source 2002::0002:0001, destination 2002::0001:0001
Tunnel bandwidth 64 (kbps)
Tunnel protocol/transport IP/IPv6
    Last 300 seconds input:  1 bytes/sec, 0 packets/sec
    Last 300 seconds output:  1 bytes/sec, 0 packets/sec
    167 packets input,  10688 bytes
    0 input error
    170 packets output,  10880 bytes
    0 output error
```

从 Switch A 可以 Ping 通对端的 Vlan-interface100 接口的 IPv4 地址：

```
[SwitchA] ping 30.1.3.1
  PING 30.1.3.1: 56  data bytes, press CTRL_C to break
    Reply from 30.1.3.1: bytes=56 Sequence=1 ttl=255 time=46 ms
```

```
  Reply from 30.1.3.1: bytes=56 Sequence=2 ttl=255 time=15 ms
  Reply from 30.1.3.1: bytes=56 Sequence=3 ttl=255 time=16 ms
  Reply from 30.1.3.1: bytes=56 Sequence=4 ttl=255 time=15 ms
  Reply from 30.1.3.1: bytes=56 Sequence=5 ttl=255 time=16 ms

--- 30.1.3.1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 15/21/46 ms
```

1.9 配置IPv6 over IPv6隧道

1.9.1 配置准备

设备上存在已经配置 IPv6 地址、能够进行正常通讯的接口（如 VLAN 接口，Loopback 接口等），该接口将作为 Tunnel 接口的源接口。

1.9.2 配置IPv6 over IPv6 隧道

表1-9 配置 IPv6 over IPv6 隧道

操作		命令	说明
进入系统视图		**system-view**	-
使能IPv6报文转发功能		**ipv6**	必选 缺省情况下，关闭IPv6报文转发功能
进入Tunnel接口视图		**interface tunnel** *number*	-
设置Tunnel接口的IPv6地址	配置IPv6全球单播地址或站点本地地址	**ipv6 address** { *ipv6-address prefix-length* | *ipv6-addressIprefix-length* }	必选之一 缺省情况下，Tunnel接口上没有设置IPv6地址
		ipv6 address *ipv6-addressIprefix-length* **eui-64**	
	配置IPv6链路本地地址	**ipv6 address auto link-local**	
		ipv6 address *ipv6-address* **link-local**	
配置隧道模式为IPv6 over IPv6隧道		**tunnel-protocol ipv6-ipv6**	必选 缺省情况下，为GRE over IPv4隧道模式 在隧道的两端应配置相同的隧道模式，否则可能造成报文传输失败
设置Tunnel接口的源端地址或接口		**source** { *ipv6-address* | *interface-type interface-number* }	必选 缺省情况下，Tunnel接口上没有设置源端地址和接口

操作	命令	说明
设置Tunnel接口的目的端地址	**destination** *ipv6-address*	必选 缺省情况下，Tunnel接口上没有设置目的端地址
退回系统视图	**quit**	-
配置丢弃含有IPv4兼容IPv6地址的IPv6报文	**tunnel discard ipv4-compatible-packet**	可选 缺省情况下，不会丢弃含有IPv4兼容IPv6地址的IPv6报文

说明

- Tunnel 接口上设置的源端和目的端必须是公网的地址或接口。
- 如果封装前 IPv6 报文的目的 IPv6 地址与 Tunnel 接口的 IPv6 地址不在同一个网段，则必须配置通过 Tunnel 接口到达目的 IPv6 地址的转发路由，以便需要进行封装的报文能正常转发。用户可以配置静态路由，指定到达目的 IPv6 地址的路由出接口为本端 Tunnel 接口或下一跳为对端 Tunnel 接口地址。用户也可以配置动态路由，在 Tunnel 接口使能动态路由协议。在隧道的两端都要进行转发路由的配置，配置的详细情况请参见“三层技术-IP 路由配置指导”中的“IPv6 静态路由”或其他路由协议配置。
- 本端隧道接口的 IPv6 地址与隧道的目的地址不能在同一个网段内。
- 配置经过隧道接口的路由时，路由的目的地址不能与该隧道的目的地址在同一个网段内。
- 对两个或两个以上使用同种封装协议的 Tunnel 接口，不能同时配置完全相同的源地址和目的地址。
- 配置 Tunnel 接口的源端地址时，若采用配置源接口形式，则 Tunnel 的源地址取的是源接口的主 IP 地址。

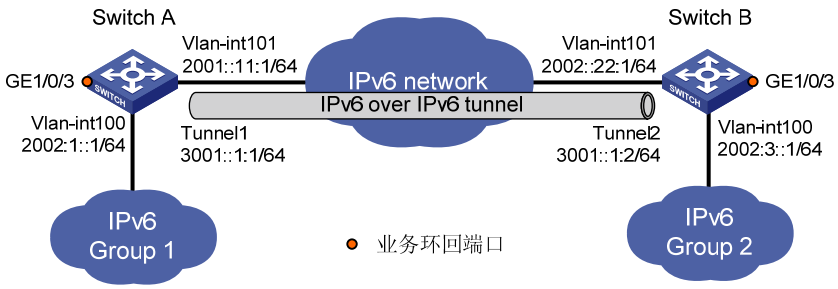
1.9.3 配置举例

1. 组网需求

运行 IPv6 协议的两个子网 Group 1 和 Group 2 的网络地址不希望泄露到 IPv6 网络中。网络管理员通过在交换机 Switch A 和交换机 Switch B 之间建立 IPv6 over IPv6 隧道，实现在 Group 1 和 Group 2 的网络地址不被泄露的情况下，确保 Group 1 和 Group 2 互通。

2. 组网图

图1-12 IPv6 over IPv6 隧道组网图



3. 配置步骤

说明

在开始下面的配置之前，请确保 Switch A 和 Switch B 上已经创建相应的 VLAN 接口，且两者之间 IPv6 报文路由可达。

(1) 配置 Switch A

使能 IPv6 转发功能。

```
<SwitchA> system-view
[SwitchA] ipv6
```

配置接口 Vlan-interface100。

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ipv6 address 2002:1::1 64
[SwitchA-Vlan-interface100] quit
```

配置接口 Vlan-interface101（隧道的实际物理接口）。

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] ipv6 address 2001::11:1 64
[SwitchA-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchA] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```

创建 Tunnel1 接口。

```
[SwitchA] interface tunnel 1
```

配置 Tunnel1 接口的 IP 地址。

```
[SwitchA-Tunnel1] ipv6 address 3001::1:1 64
```

配置 Tunnel 封装模式。

```
[SwitchA-Tunnel1] tunnel-protocol ipv6-ipv6
```
配置 Tunnel1 接口的源地址（Vlan-interface101 的 IP 地址）。
```
[SwitchA-Tunnel1] source 2001::11:1
```
配置 Tunnel1 接口的目的地址（Switch B 的 Vlan-interface101 的 IP 地址）。
```
[SwitchA-Tunnel1] destination 2002::22:1
```
在 Tunnel 接口视图下指定隧道引用业务环回组 1。
```
[SwitchA-Tunnel1] service-loopback-group 1
[SwitchA-Tunnel1] quit
```
配置从 Switch A 经过 Tunnel1 接口到 Group 2 的静态路由。
```
[SwitchA] ipv6 route-static 2002:3:: 64 tunnel 1
```
(2)　配置 Switch B

使能 IPv6 转发功能。
```
<SwitchB> system-view
[SwitchB] ipv6
```
配置接口 Vlan-interface100。
```
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ipv6 address 2002:3::1 64
[SwitchB-Vlan-interface100] quit
```
配置接口 Vlan-interface101（隧道的实际物理接口）。
```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] ipv6 address 2002::22:1 64
[SwitchB-Vlan-interface101] quit
```
创建业务环回组 1，并配置服务类型为 tunnel。
```
[SwitchB] service-loopback group 1 type tunnel
```
将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。
```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] undo stp enable
[SwitchB-GigabitEthernet1/0/3] undo ndp enable
[SwitchB-GigabitEthernet1/0/3] undo lldp enable
[SwitchB-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchB-GigabitEthernet1/0/3] quit
```
创建 Tunnel2 接口。
```
[SwitchB] interface tunnel 2
```
配置 Tunnel2 接口的 IP 地址。
```
[SwitchB-Tunnel2] ipv6 address 3001::1:2 64
```
配置 Tunnel 封装模式。
```
[SwitchB-Tunnel2] tunnel-protocol ipv6-ipv6
```
配置 Tunnel2 接口的源地址（Vlan-interface101 的 IP 地址）。
```
[SwitchB-Tunnel2] source 2002::22:1
```
配置 Tunnel2 接口的目的地址（Switch A 的 Vlan-interface101 的 IP 地址）。
```
[SwitchB-Tunnel2] destination 2001::11:1
```
在 Tunnel 接口视图下指定隧道引用业务环回组 1。
```
[SwitchB-Tunnel2] service-loopback-group 1
[SwitchB-Tunnel2] quit
```

配置从 Switch B 经过 Tunnel2 接口到 Group 1 的静态路由。

```
[SwitchB] ipv6 route-static 2002:1:: 64 tunnel 2
```

4. 验证配置结果

完成以上配置之后，分别查看 Switch A 和 Switch B 的 Tunnel 接口状态如下：

```
[SwitchA] display ipv6 interface tunnel 1
Tunnel1 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::2013:1
  Global unicast address(es):
    3001::1:1, subnet is 3001::/64
  Joined group address(es):
    FF02::1:FF13:1
    FF02::1:FF01:1
    FF02::1:FF00:0
    FF02::2
    FF02::1
  MTU is 1460 bytes
  ND reachable time is 30000 milliseconds
  ND retransmit interval is 1000 milliseconds
  Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
……（略）
[SwitchB] display ipv6 interface tunnel 2
Tunnel2 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::2024:1
  Global unicast address(es):
    3001::1:2, subnet is 3001::/64
  Joined group address(es):
    FF02::1:FF24:1
    FF02::1:FF01:2
    FF02::1:FF00:0
    FF02::2
    FF02::1
  MTU is 1460 bytes
  ND reachable time is 30000 milliseconds
  ND retransmit interval is 1000 milliseconds
  Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
……（略）
```

从 Switch A 可以 Ping 通对端的 Vlan-interface100 接口的 IPv6 地址：

```
[SwitchA] ping ipv6 2002:3::1
  PING 2002:3::1 : 56  data bytes, press CTRL_C to break
    Reply from 2002:3::1
    bytes=56 Sequence=1 hop limit=64  time = 31 ms
    Reply from 2002:3::1
    bytes=56 Sequence=2 hop limit=64  time = 1 ms
```

```
  Reply from 2002:3::1
  bytes=56 Sequence=3 hop limit=64  time = 16 ms
  Reply from 2002:3::1
  bytes=56 Sequence=4 hop limit=64  time = 16 ms
  Reply from 2002:3::1
  bytes=56 Sequence=5 hop limit=64  time = 31 ms

--- 2002:3::1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/19/31 ms
```

1.10 隧道显示和维护

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

在用户视图下执行 **reset** 命令可以清除 Tunnel 接口的统计信息。

表1-10 隧道显示和维护

操作	命令
显示Tunnel接口的相关信息	**display interface** [**tunnel**] [**brief** [**down**]] [**|** { **begin** | **exclude** | **include** } *regular-expression*] **display interface tunnel** *number* [**brief**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示Tunnel接口的IPv6相关信息	**display ipv6 interface tunnel** [*number*] [**brief**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
清除Tunnel接口的统计信息	**reset counters interface** [**tunnel** [*number*]]

1.11 常见配置错误举例

1. 故障现象

在 Tunnel 接口上配置了相关的参数后（例如隧道的起点、终点地址和隧道模式）仍未处于 up 状态。

2. 故障排除

- Tunnel 接口未处于 up 状态的最常见原因是隧道起点的物理接口没有处于 up 状态。使用 **display interface tunnel** 和 **display ipv6 interface tunnel** 命令查看隧道起点的物理接口状态为 up 还是 down。如果物理接口状态是 down 的，请检查网络连接。
- Tunnel 接口未处于 up 状态的另一个可能的原因是隧道的终点地址不可达。使用 **display ipv6 routing-table** 和 **display ip routing-table** 命令查看是否终点地址通过路由可达。如果路由表中没有保证隧道通讯的路由项，请配置相关路由。

目 录

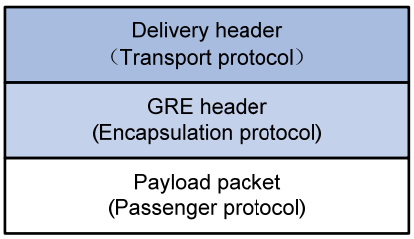
1 GRE配置

1.1 GRE简介

GRE（Generic Routing Encapsulation，通用路由封装）协议是对某些网络层协议（如 IP 和 IPX）的数据报文进行封装，使这些被封装的数据报文能够在另一个网络层协议（如 IP）中传输。封装后的数据报文在网络中传输的路径，称为 GRE 隧道。GRE 隧道是一个虚拟的点到点的连接，其两端的设备分别对数据报进行封装及解封装。

1.1.1 GRE封装后的报文格式

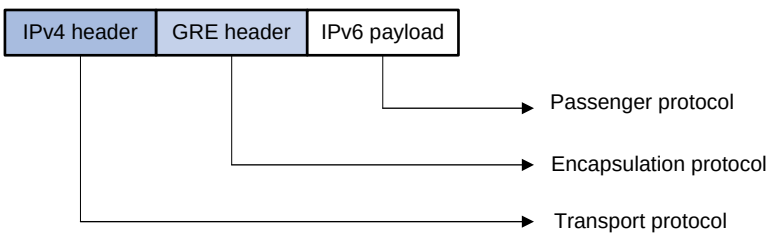
图1-1 GRE 封装后的报文格式



如 图 1-1 所示，GRE封装后的报文包括如下几个部分：

- 净荷数据（Payload packet）：需要封装和传输的数据报文。净荷数据的协议类型，称为乘客协议（Passenger Protocol）。
- GRE 头（GRE header）：系统收到净荷数据后，在净荷数据上添加 GRE 头，使其成为 GRE 报文。对净荷数据进行封装的 GRE 协议，称为封装协议（Encapsulation Protocol）。
- 传输协议的报文头（Delivery header）：负责转发封装后报文的网络协议，称为传输协议（Delivery Protocol 或者 Transport Protocol）。在 GRE 报文上需要增加传输协议的报文头，以便传输协议对封装后的报文进行转发处理。

图1-2 GRE 封装报文举例



IPv6 报文通过GRE隧道穿越IPv4 网络时，报文格式如 图 1-2 所示。其中，乘客协议为IPv6，封装协议为GRE，传输协议为IPv4。

根据传输协议的不同，GRE 隧道可以分为：

- GRE over IPv4：传输协议为 IPv4，乘客协议为任意网络层协议。

- GRE over IPv6：传输协议为 IPv6，乘客协议为任意网络层协议。

1.1.2 GRE加解封装过程

图1-3 X 协议网络通过 GRE 隧道互连

下面以 图 1-3 的网络为例说明X协议的报文穿越IP网络在GRE隧道中传输的过程：

1. 加封装过程

- Device A 连接 Group 1 的接口收到 X 协议报文后，首先交由 X 协议处理；
- X 协议检查报文头中的目的地址域来确定如何路由此包；
- 若报文的目的地址要经过 Tunnel 才能到达，则设备将此报文发给相应的 Tunnel 接口；
- Tunnel 接口收到此报文后进行 GRE 封装，再封装 IP 报文头后，设备根据此 IP 包的目的地址及路由表对报文进行转发，从相应的网络接口发送出去。

说明

对于交换机，封装后的报文不能根据目的地址和路由表进行第二次三层转发，需要将封装后的报文发送给业务环回口，由业务环回口将报文回送给转发模块后，再进行三层转发。

---

2. 解封装的过程

解封装过程和加封装的过程相反。

- Device B 从 Tunnel 接口收到 IP 报文，检查目的地址；
- 如果目的地是本路由器，且 IP 报文头中的协议号为 47（表示封装的报文为 GRE 报文），则 Device B 剥掉此报文的 IP 报头，交给 GRE 协议处理（进行检验密钥、检查校验和及报文的序列号等）；
- GRE 协议完成相应的处理后，剥掉 GRE 报头，再交由 X 协议对此数据报进行后续的转发处理。

说明

GRE 收发双方的加封装、解封装处理，以及由于封装造成的数据量增加，会导致使用 GRE 后设备的数据转发效率有一定程度的下降。

1.1.3 协议规范

与 GRE 相关的协议规范有：

- RFC 1701：Generic Routing Encapsulation (GRE)

- RFC 1702：Generic Routing Encapsulation over IPv4 networks
- RFC 2784：Generic Routing Encapsulation (GRE)

1.2 配置GRE over IPv4隧道

1.2.1 配置准备

- 设备上存在已经配置 IP 地址、能够进行正常通讯的接口（如 VLAN 接口，Loopback 接口等），该接口将作为 Tunnel 接口的源接口。
- 配置 GRE over IPv4 隧道前，需要先创建业务类型为 Tunnel 的业务环回组，并将设备上未使用的二层以太网接口加入该业务环回组。关于业务环回组的详细介绍，请参见“二层技术-以太网交换配置指导”中的“业务环回组”。

1.2.2 配置GRE over IPv4 隧道

表1-1 配置 GRE over IPv4 隧道

操作	命令	说明
进入系统视图	**system-view**	-
创建一个Tunnel接口，并进入该Tunnel接口视图	**interface tunnel** *interface-number*	必选 缺省情况下，设备上无Tunnel接口
设置Tunnel接口的IPv4地址	**ip address** *ip-address* { *mask* | *mask-length* }	必选 缺省情况下，Tunnel接口上没有设置IPv4地址
配置隧道模式为GRE over IPv4	**tunnel-protocol gre**	可选 缺省情况下，采用GRE over IPv4隧道模式 在隧道的两端应配置相同的隧道模式，否则可能造成报文传输失败
设置Tunnel接口的源端地址或接口	**source** { *ip-address* | *interface-type interface-number* }	必选 缺省情况下，Tunnel接口上没有设置源端地址和接口
设置Tunnel接口的目的端地址	**destination** *ip-address*	必选 缺省情况下，Tunnel接口上没有设置目的端地址
配置通过Tunnel转发报文的路由	配置的详细情况请参见“三层技术-IP路由配置指导”中的“静态路由”或其他路由协议配置	必选 在源端路由器和目的端路由器上都必须存在经过Tunnel转发报文的路由，这样需要进行GRE封装的报文才能正确转发。可以配置静态路由，也可以配置动态路由
退回系统视图	**quit**	-

操作	命令	说明
配置丢弃含有IPv4兼容IPv6地址的IPv6报文	**tunnel discard ipv4-compatible-packet**	可选 缺省情况下，不会丢弃含有IPv4兼容IPv6地址的IPv6报文

说明

- Tunnel 接口上设置的源端和目的端必须是公网的地址或接口。
- Tunnel 接口的详细介绍，及 Tunnel 接口下的更多配置命令，请参见“三层技术-IP 业务配置指导”中的“隧道”。
- **interface tunnel**、**tunnel-protocol**、**source**、**destination** 和 **tunnel discard ipv4-compatible-packet** 命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“隧道”。

说明

- Tunnel 的源端地址与目的端地址唯一标识了一个通道。Tunnel 两端必须配置源端地址与目的端地址，且两端地址互为源地址和目的地址。
- 两个或两个以上使用同种封装协议的 Tunnel 接口不能配置完全相同的源地址和目的地址。
- 配置 Tunnel 接口的源端地址时，若采用配置源接口形式，则 Tunnel 的源地址取的是源接口的主 IP 地址。
- 配置通过 Tunnel 转发的路由时，可以手工配置一条静态路由，目的地址是未进行 GRE 封装的报文的目的地址，下一跳是对端 Tunnel 接口的地址。也可以在 Tunnel 接口上和与私网相连的路由器接口上分别使能动态路由协议，由动态路由协议来建立通过 Tunnel 转发的路由表项。
- 在 Tunnel 接口配置的静态路由的目的地址不能与 Tunnel 接口的地址在同一网段。

1.3 配置GRE over IPv6隧道

1.3.1 配置准备

- 设备上存在已经配置 IP 地址、能够进行正常通讯的接口（如 VLAN 接口，Loopback 接口等），该接口将作为 Tunnel 接口的源接口。
- 配置 GRE over IPv6 隧道前，需要先创建业务类型为 Tunnel 的业务环回组，并将设备上未使用的二层以太网接口加入该业务环回组。关于业务环回组的详细介绍，请参见“二层技术-以太网交换配置指导”中的“业务环回组”。

1.3.2 配置GRE over IPv6 隧道

表1-2 配置 GRE over IPv6 隧道

操作	命令	说明
进入系统视图	**system-view**	-
使能IPv6报文转发功能	**ipv6**	必选 缺省情况下，关闭IPv6报文转发功能
创建一个Tunnel接口，并进入该Tunnel接口视图	**interface tunnel** *interface-number*	必选 缺省情况下，设备上无Tunnel接口
设置Tunnel接口的IPv4地址	**ip address** *ip-address* { *mask* | *mask-length* }	必选 缺省情况下，Tunnel接口上没有设置IPv4地址
配置隧道模式为GRE over IPv6	**tunnel-protocol gre ipv6**	必选 缺省情况下，采用GRE over IPv4隧道模式 在隧道的两端应配置相同的隧道模式，否则可能造成报文传输失败
设置Tunnel接口的源端地址或接口	**source** { *ipv6-address* | *interface-type interface-number* }	必选 缺省情况下，Tunnel接口上没有设置源端地址和接口
设置Tunnel接口的目的端地址	**destination** *ipv6-address*	必选 缺省情况下，Tunnel接口上没有设置目的端地址
退回系统视图	**quit**	-
配置丢弃含有IPv4兼容IPv6地址的IPv6报文	**tunnel discard ipv4-compatible-packet**	可选 缺省情况下，不会丢弃含有IPv4兼容IPv6地址的IPv6报文
配置通过Tunnel转发报文的路由	配置的详细情况请参见“三层技术-IP路由配置指导”中的“静态路由”或其他路由协议配置	必选 在源端路由器和目的端路由器上都必须存在经过Tunnel转发报文的路由，这样需要进行GRE封装的报文才能正确转发。可以配置静态路由，也可以配置动态路由

说明

- Tunnel 接口上设置的源端和目的端必须是公网的地址或接口。
- **interface tunnel**、**tunnel-protocol**、**source**、**destination** 和 **tunnel discard ipv4-compatible-packet** 命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“隧道”。
- Tunnel 接口的详细介绍，及 Tunnel 接口下的更多配置命令，请参见“三层技术-IP 业务配置指导”中的“隧道”。

说明

- 以上各项 Tunnel 接口下进行的功能特性配置，在删除 Tunnel 接口后，该接口上的所有配置也将被删除。
- Tunnel 的源端地址与目的端地址唯一标识了一个通道。Tunnel 两端必须配置源端地址与目的端地址，且两端地址互为源地址和目的地址。
- 两个或两个以上使用同种封装协议的 Tunnel 接口不能配置完全相同的源地址和目的地址。
- 配置 Tunnel 接口的源端地址时，若采用配置源接口形式，则 Tunnel 的源地址取的是源接口的主 IP 地址。
- 配置通过 Tunnel 转发的路由时，可以手工配置一条静态路由，目的地址是未进行 GRE 封装的报文的目的地址，下一跳是对端 Tunnel 接口的地址。也可以在 Tunnel 接口上和与私网相连的路由器接口上分别使能动态路由协议，由动态路由协议来建立通过 Tunnel 转发的路由表项。
- 在 Tunnel 接口配置的静态路由的目的地址不能与 Tunnel 接口的地址在同一网段。

1.4 GRE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 GRE 的运行情况，通过查看显示信息验证配置的效果。

表1-3 GRE 的显示和维护

操作	命令
显示Tunnel接口的相关信息	**display interface** [**tunnel**] [**brief** [**down**]] [**|** { **begin** | **exclude** | **include** } *regular-expression*] **display interface tunnel** *number* [**brief**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]
显示Tunnel接口的IPv6相关信息	**display ipv6 interface tunnel** [*number*] [**brief**] [**|** { **begin** | **exclude** | **include** } *regular-expression*]

说明

display interface tunnel 和 **display ipv6 interface tunnel** 命令的详细介绍，请参见“三层技术-IP 业务命令参考”中的“隧道”。

1.5 GRE over IPv4典型配置举例

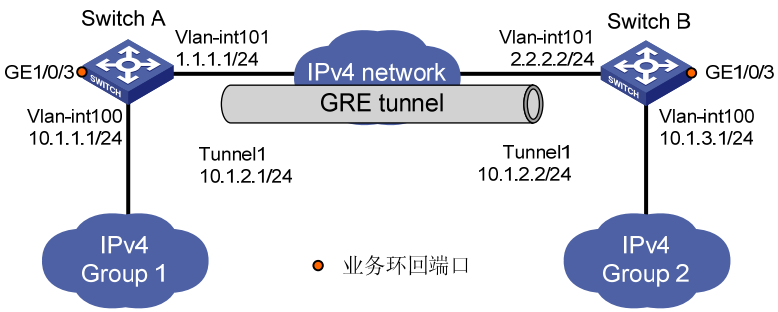
1.5.1 GRE over IPv4 典型配置举例

1. 组网需求

交换机 Switch A 和交换机 Switch B 之间通过 Internet 相连。运行 IP 协议的私有网络的两个子网 Group 1 和 Group 2，通过在两台交换机之间使用 GRE 建立隧道实现互联。

2. 组网图

图1-4 GRE over IPv4 应用组网图



3. 配置步骤

说明

在开始下面的配置之前，需确保 Switch A 和 Switch B 之间路由可达。

(1) 配置交换机 Switch A

配置接口 GigabitEthernet1/0/1。

```
<SwitchA> system-view
[SwitchA] vlan 100
[SwitchA-vlan100] port GigabitEthernet 1/0/1
[SwitchA-vlan100] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ip address 10.1.1.1 255.255.255.0
[SwitchA-Vlan-interface100] quit
```

配置接口 GigabitEthernet1/0/2（隧道的实际物理接口）。

```
[SwitchA] vlan 101
[SwitchA-vlan101] port GigabitEthernet 1/0/2
```

```
[SwitchA-vlan101] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] ip address 1.1.1.1 255.255.255.0
[SwitchA-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchA] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```

创建 Tunnel1 接口。

```
[SwitchA] interface tunnel 1
```

配置 Tunnel1 接口的 IP 地址。

```
[SwitchA-Tunnel1] ip address 10.1.2.1 255.255.255.0
```

配置 Tunnel 封装模式为 GRE over IPv4 隧道模式。

```
[SwitchA-Tunnel1] tunnel-protocol gre
```

配置 Tunnel1 接口的源地址（GigabitEthernet1/0/2 所属 VLAN 接口的 IP 地址）。

```
[SwitchA-Tunnel1] source vlan-interface 101
```

配置 Tunnel1 接口的目的地址（Switch B 的 GigabitEthernet1/0/2 所属 VLAN 接口的 IP 地址）。

```
[SwitchA-Tunnel1] destination 2.2.2.2
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchA-Tunnel1] service-loopback-group 1
[SwitchA-Tunnel1] quit
```

配置从 Switch A 经过 Tunnel1 接口到 Group 2 的静态路由。

```
[SwitchA] ip route-static 10.1.3.0 255.255.255.0 tunnel 1
```

(2)　配置交换机 Switch B

配置接口 GigabitEthernet1/0/1。

```
<SwitchB> system-view
[SwitchB] vlan 100
[SwitchB-vlan100] port GigabitEthernet 1/0/1
[SwitchB-vlan100] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ip address 10.1.3.1 255.255.255.0
[SwitchB-Vlan-interface100] quit
```

配置接口 GigabitEthernet1/0/2（隧道的实际物理接口）。

```
[SwitchB] vlan 101
[SwitchB-vlan101] port GigabitEthernet 1/0/2
[SwitchB-vlan101] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] ip address 2.2.2.2 255.255.255.0
[SwitchB-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchB] service-loopback group 1 type tunnel
```
将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。
```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] undo stp enable
[SwitchB-GigabitEthernet1/0/3] undo ndp enable
[SwitchB-GigabitEthernet1/0/3] undo lldp enable
[SwitchB-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchB-GigabitEthernet1/0/3] quit
```
创建 Tunnel1 接口。
```
[SwitchB] interface tunnel 1
```
配置 Tunnel1 接口的 IP 地址。
```
[SwitchB-Tunnel1] ip address 10.1.2.2 255.255.255.0
```
配置 Tunnel 封装模式为 GRE over IPv4 隧道模式。
```
[SwitchB-Tunnel1] tunnel-protocol gre
```
配置 Tunnel1 接口的源地址（GigabitEthernet1/0/2 所属 VLAN 接口的 IP 地址）。
```
[SwitchB-Tunnel1] source vlan-interface 101
```
配置 Tunnel1 接口的目的地址（Switch A 的 GigabitEthernet1/0/2 所属 VLAN 接口的 IP 地址）。
```
[SwitchB-Tunnel1] destination 1.1.1.1
```
在 Tunnel 接口视图下指定隧道引用业务环回组 1。
```
[SwitchB-Tunnel1] service-loopback-group 1
[SwitchB-Tunnel1] quit
```
配置从 Switch B 经过 Tunnel1 接口到 Group 1 的静态路由。
```
[SwitchB] ip route-static 10.1.1.0 255.255.255.0 Tunnel 1
```
(3) 验证配置结果

完成以上配置后，分别查看 Switch A 和 Switch B 的 Tunnel 接口状态。
```
[SwitchA] display interface tunnel 1
Tunnel1 current state: UP
Line protocol current state: UP
Description: Tunnel1 Interface
The Maximum Transmit Unit is 1476
Internet Address is 10.1.2.1/24 Primary
Encapsulation is TUNNEL, service-loopback-group ID is 1.
Tunnel source 1.1.1.1, destination 2.2.2.2
Tunnel bandwidth 64 (kbps)
Tunnel protocol/transport GRE/IP
    GRE key disabled
    Checksumming of GRE packets disabled
Last clearing of counters:  Never
    Last 300 seconds input:  0 bytes/sec, 0 packets/sec
    Last 300 seconds output:  0 bytes/sec, 0 packets/sec
    10 packets input,  840 bytes
    0 input error
    10 packets output,  840 bytes
    0 output error
[SwitchB] display interface tunnel 1
Tunnel1 current state: UP
```

```
Line protocol current state: UP
Description: Tunnel1 Interface
The Maximum Transmit Unit is 1476
Internet Address is 10.1.2.2/24 Primary
Encapsulation is TUNNEL, service-loopback-group ID is 1.
Tunnel source 2.2.2.2, destination 1.1.1.1
Tunnel bandwidth 64 (kbps)
Tunnel protocol/transport GRE/IP
    GRE key disabled
    Checksumming of GRE packets disabled
Last clearing of counters:  Never
    Last 300 seconds input:  2 bytes/sec, 0 packets/sec
    Last 300 seconds output:  2 bytes/sec, 0 packets/sec
    10 packets input,  840 bytes
    0 input error
    10 packets output,  840 bytes
    0 output error
```

从 Switch B 可以 Ping 通 Switch A 上 VLAN 接口 100 的地址。

```
[SwitchB] ping 10.1.1.1
  PING 10.1.1.1: 56  data bytes, press CTRL_C to break
    Reply from 10.1.1.1: bytes=56 Sequence=1 ttl=255 time=2 ms
    Reply from 10.1.1.1: bytes=56 Sequence=2 ttl=255 time=2 ms
    Reply from 10.1.1.1: bytes=56 Sequence=3 ttl=255 time=2 ms
    Reply from 10.1.1.1: bytes=56 Sequence=4 ttl=255 time=2 ms
    Reply from 10.1.1.1: bytes=56 Sequence=5 ttl=255 time=2 ms

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

1.6 GRE over IPv6典型配置举例

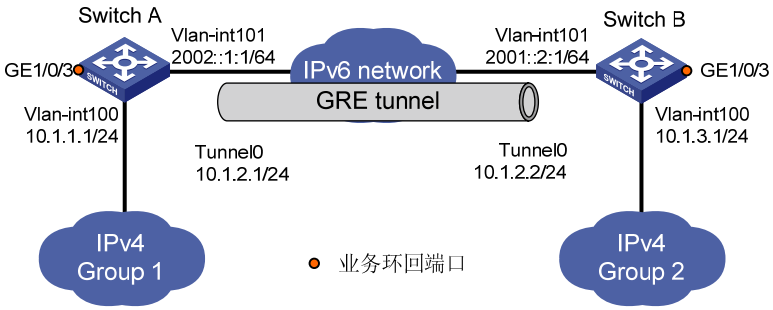
1.6.1 GRE over IPv6 典型配置举例

1. 组网需求

运行 IP 协议的两个子网 Group1 和 Group2 通过 IPv6 网络相连。通过在交换机 SwitchA 和交换机 SwitchB 之间建立 GRE over IPv6 隧道，实现两个子网穿越 IPv6 网络互联。

2. 组网图

图1-5 GRE over IPv6 应用组网图



3. 配置步骤

说明

在开始下面的配置之前，需确保 Switch A 和 Switch B 之间路由可达。

(1) 配置交换机 Switch A

```
<SwitchA> system-view
```

使能 IPv6。

```
[SwitchA] ipv6
```

配置接口 Vlan-interface100。

```
[SwitchA] vlan 100
[SwitchA-vlan100] port GigabitEthernet 1/0/1
[SwitchA-vlan100] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ip address 10.1.1.1 255.255.255.0
[SwitchA-Vlan-interface100] quit
```

配置接口 Vlan-interface101（隧道的实际物理接口）。

```
[SwitchA] vlan 101
[SwitchA-vlan101] port GigabitEthernet 1/0/2
[SwitchA-vlan101] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] ipv6 address 2002::1:1 64
[SwitchA-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchA] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```

创建 Tunnel0 接口。

```
[SwitchA] interface tunnel 0
```

配置 Tunnel0 接口的 IP 地址。

```
[SwitchA-Tunnel0] ip address 10.1.2.1 255.255.255.0
```

配置 Tunnel 封装模式为 GRE over IPv6 隧道模式。

```
[SwitchA-Tunnel0] tunnel-protocol gre ipv6
```

配置 Tunnel0 接口的源地址（Vlan-interface101 的 IP 地址）。

```
[SwitchA-Tunnel0] source 2002::1:1
```

配置 Tunnel0 接口的目的地址（Switch B 的 Vlan-interface101 的 IP 地址）。

```
[SwitchA-Tunnel0] destination 2001::2:1
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchA-Tunnel0] service-loopback-group 1
[SwitchA-Tunnel0] quit
```

配置从 Switch A 经过 Tunnel1 接口到 Group 2 的静态路由。

```
[SwitchA] ip route-static 10.1.3.0 255.255.255.0 tunnel 0
```

(2) 配置交换机 Switch B

```
<SwitchB> system-view
```

使能 IPv6。

```
[SwitchB] ipv6
```

配置接口 Vlan-interface100。

```
[SwitchB] vlan 100
[SwitchB-vlan100] port GigabitEthernet 1/0/1
[SwitchB-vlan100] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ip address 10.1.3.1 255.255.255.0
[SwitchB-Vlan-interface100] quit
```

配置接口 Vlan-interface101（隧道的实际物理接口）。

```
[SwitchB] vlan 101
[SwitchB-vlan101] port GigabitEthernet 1/0/2
[SwitchB-vlan101] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] ipv6 address 2001::2:1 64
[SwitchB-Vlan-interface101] quit
```

创建业务环回组 1，并配置服务类型为 tunnel。

```
[SwitchB] service-loopback group 1 type tunnel
```

将接口 GigabitEthernet1/0/3 加入业务环回组 1，并在该端口上关闭 STP、NDP 和 LLDP 功能。

```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] undo stp enable
[SwitchB-GigabitEthernet1/0/3] undo ndp enable
[SwitchB-GigabitEthernet1/0/3] undo lldp enable
[SwitchB-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchB-GigabitEthernet1/0/3] quit
```

创建 Tunnel0 接口。

```
[SwitchB] interface tunnel 0
```

配置 Tunnel0 接口的 IP 地址。

```
[SwitchB-Tunnel0] ip address 10.1.2.2 255.255.255.0
```

配置 Tunnel 封装模式为 GRE over IPv6 隧道模式。

```
[SwitchB-Tunnel0] tunnel-protocol gre ipv6
```

配置 Tunnel0 接口的源地址（Vlan-interface101 的 IP 地址）。

```
[SwitchB-Tunnel0] source 2001::2:1
```

配置 Tunnel0 接口的目的地址（Switch A 的 Vlan-interface101 的 IP 地址）。

```
[SwitchB-Tunnel0] destination 2002::1:1
```

在 Tunnel 接口视图下指定隧道引用业务环回组 1。

```
[SwitchB-Tunnel0] service-loopback-group 1
[SwitchB-Tunnel0] quit
```

配置从 Switch B 经过 Tunnel2 接口到 Group 1 的静态路由。

```
[SwitchB] ip route-static 10.1.1.0 255.255.255.0 tunnel 0
```

(3) 验证配置结果

完成以上配置后，分别查看 Switch A 和 Switch B 的 Tunnel 接口状态。

```
[SwitchA] display interface Tunnel 0
Tunnel0 current state: UP
Line protocol current state: UP
Description: Tunnel0 Interface
The Maximum Transmit Unit is 1456
Internet Address is 10.1.2.1/24 Primary
Encapsulation is TUNNEL, service-loopback-group ID is 1.
Tunnel source 2002::1:1, destination 2001::2:1
Tunnel bandwidth 64 (kbps)
Tunnel protocol/transport GRE/IPv6
    GRE key disabled
    Checksumming of GRE packets disabled
Last clearing of counters:  Never
    Last 300 seconds input:  0 bytes/sec, 0 packets/sec
    Last 300 seconds output:  0 bytes/sec, 0 packets/sec
    10 packets input,  840 bytes
    0 input error
    10 packets output,  840 bytes
    0 output error
[SwitchB] display interface Tunnel 0
Tunnel0 current state: UP
Line protocol current state: UP
Description: Tunnel0 Interface
The Maximum Transmit Unit is 1456
Internet Address is 10.1.2.2/24 Primary
Encapsulation is TUNNEL, service-loopback-group ID is 1.
Tunnel source 2001::2:1, destination 2002::1:1
Tunnel bandwidth 64 (kbps)
Tunnel protocol/transport GRE/IPv6
    GRE key disabled
    Checksumming of GRE packets disabled
```

```
Last clearing of counters:  Never
    Last 300 seconds input:  0 bytes/sec, 0 packets/sec
    Last 300 seconds output:  0 bytes/sec, 0 packets/sec
    10 packets input,  840 bytes
    0 input error
    10 packets output,  840 bytes
    0 output error
```

从 Switch B 可以 Ping 通 Switch A 上 VLAN 接口 100 的地址。

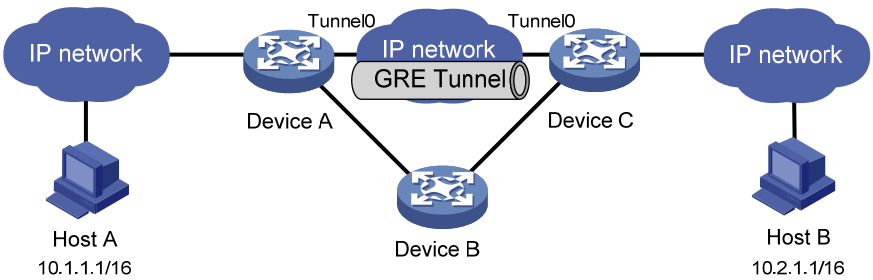
```
[SwitchB] ping 10.1.1.1
  PING 10.1.1.1: 56  data bytes, press CTRL_C to break
    Reply from 10.1.1.1: bytes=56 Sequence=1 ttl=255 time=3 ms
    Reply from 10.1.1.1: bytes=56 Sequence=2 ttl=255 time=2 ms
    Reply from 10.1.1.1: bytes=56 Sequence=3 ttl=255 time=2 ms
    Reply from 10.1.1.1: bytes=56 Sequence=4 ttl=255 time=2 ms
    Reply from 10.1.1.1: bytes=56 Sequence=5 ttl=255 time=3 ms

  --- 10.1.1.1 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 2/2/3 ms
```

1.7 常见配置错误举例

GRE的配置相对比较简单，但要注意配置的一致性，大部分的错误都可以通过使用调试命令**debugging gre**和**debugging tunnel**定位。这里仅就一种错误进行分析，如 图 1-6 所示。

图1-6 GRE 排错示例



1. 故障现象

Tunnel 两端接口配置正确且 Tunnel 两端可以 ping 通，但 Host A 和 Host B 之间却无法 ping 通。

2. 故障排除

- 在任意视图下，在 Device A 和 Device C 分别执行 **display ip routing-table** 命令，观察在 Device A 是否有经过 Tunnel0 接口到 10.2.0.0/16 的路由；在 Device C 是否有经过 Tunnel0 接口到 10.1.0.0/16 的路由。
- 如果在上一步的输出中发现缺少相应的静态路由，在系统视图下使用 **ip route-static** 命令添加。以 Device A 为例，配置如下：

```
[DeviceA] ip route-static 10.2.0.0 255.255.0.0 tunnel 0
```