

H3C SecPath M9000 系列 多业务安全网关

三层技术-IP 业务配置指导(V7)

Copyright © 2021-2024 新华三技术有限公司及其许可者 版权所有，保留一切权利。

未经本公司书面许可，任何单位和个人不得擅自摘抄、复制本书内容的部分或全部，并不得以任何形式传播。

除新华三技术有限公司的商标外，本手册中出现的其它公司的商标、产品标识及商品名称，由各自权利人拥有。

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

前言

本配置指导主要介绍 IP 业务相关技术的相关功能原理及配置。

前言部分包含如下内容：

- [读者对象](#)
- [本书约定](#)
- [资料意见反馈](#)

读者对象

本手册主要适用于如下工程师：

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

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。
<i>斜体</i>	命令行参数（命令中必须由实际值进行替代的部分）采用 <i>斜体</i> 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。
	该图标及其相关描述文字代表无线控制器、无线控制器业务板和有线无线一体化交换机的无线控制引擎设备。
	该图标及其相关描述文字代表无线接入点设备。
	该图标及其相关描述文字代表无线终结单元。
	该图标及其相关描述文字代表无线终结者。
	该图标及其相关描述文字代表无线Mesh设备。
	该图标代表发散的无线射频信号。
	该图标代表点到点的无线射频信号。
	该图标及其相关描述文字代表防火墙、UTM、多业务安全网关、负载均衡等安全设备。
	该图标及其相关描述文字代表防火墙插卡、负载均衡插卡、NetStream插卡、SSL VPN插卡、IPS插卡、ACG插卡等安全插卡。

5. 示例约定

由于设备型号不同、配置不同、版本升级等原因，可能造成本手册中的内容与用户使用的设备显示信息不一致。实际使用中请以设备显示的内容为准。

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。

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	1-1
1.1 ARP 简介	1-1
1.1.1 ARP 报文结构	1-1
1.1.2 ARP 地址解析过程	1-1
1.1.3 ARP 表项类型	1-2
1.2 vSystem 相关说明	1-3
1.3 ARP 配置任务简介	1-3
1.4 手工添加静态 ARP 表项	1-4
1.4.1 手工添加短静态 ARP 表项	1-4
1.4.2 手工添加长静态 ARP 表项	1-4
1.5 配置动态 ARP 表项的相关功能	1-5
1.5.1 配置设备学习动态 ARP 表项的最大数目	1-5
1.5.2 配置接口学习动态 ARP 表项的最大数目	1-5
1.5.3 配置动态 ARP 表项的老化时间	1-6
1.5.4 开启动态 ARP 表项的检查功能	1-6
1.6 开启在地址借用的接口学习不同网段 ARP 表项的功能	1-6
1.7 开启 ARP 日志信息功能	1-7
1.8 ARP 显示和维护	1-8
2 免费 ARP	2-1
2.1 免费 ARP 简介	2-1
2.1.1 IP 地址冲突检测	2-1
2.1.2 免费 ARP 报文学习	2-1
2.1.3 定时发送免费 ARP	2-1
2.2 免费 ARP 配置任务简介	2-2
2.3 开启源 IP 地址冲突提示功能	2-2
2.4 开启免费 ARP 报文学习功能	2-2
2.5 开启定时发送免费 ARP 功能	2-3
2.6 开启设备收到非同一网段 ARP 请求时发送免费 ARP 报文功能	2-3
2.7 配置当接口 MAC 地址变化时，该接口重新发送免费 ARP 报文的次数和时间间隔	2-3
3 代理 ARP	3-1
3.1 代理 ARP 简介	3-1
3.2 开启普通代理 ARP 功能	3-1

3.3 开启本地代理 ARP 功能	3-1
3.4 代理 ARP 显示和维护	3-2
4 ARP Snooping	4-1
4.1 ARP Snooping 简介	4-1
4.1.1 ARP Snooping 表项建立机制	4-1
4.1.2 ARP Snooping 表项老化机制	4-1
4.1.3 ARP Snooping 表项冲突处理机制	4-1
4.2 开启 ARP Snooping 功能	4-1
4.3 ARP Snooping 显示和维护	4-2

1 ARP

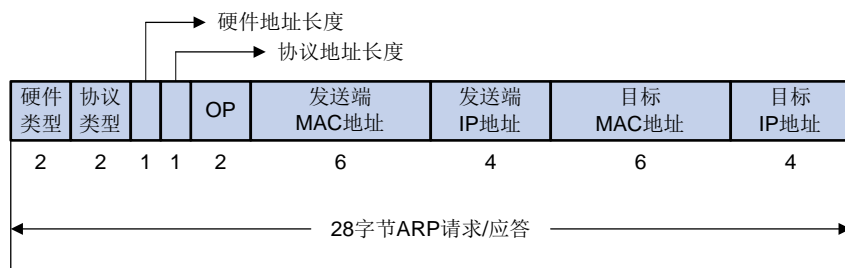
1.1 ARP简介

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。在网络中，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），由于 IP 数据报必须封装成帧才能通过物理网络发送，因此还需要知道对方的物理地址，所以设备上需要存在一个从 IP 地址到物理地址的映射关系。ARP 就是实现这个功能的协议。

1.1.1 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。

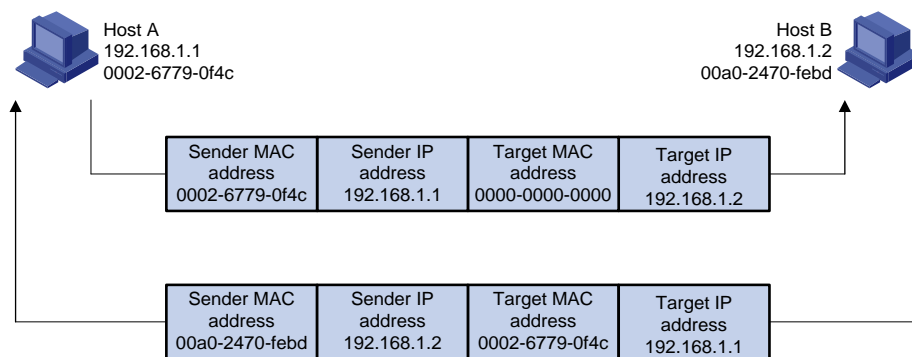
1.1.2 ARP 地址解析过程

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 数据报进行帧封装，并将 IP 数据报发送给主机 B。

- (2) 如果主机 A 在 ARP 表中找不到对应的 MAC 地址，则将缓存该 IP 数据报，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。
- (3) 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。
- (4) 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 数据报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，将报文封装并发送给网关。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。

1.1.3 ARP 表项类型

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 地址和 MAC 地址映射关系的表项，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、静态 ARP 表项和 Rule 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、接口状态 down 时，系统会删除相应的动态 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 表项分为短静态 ARP 表项、长静态 ARP 表项。

- 长静态 ARP 表项可以直接用于报文转发，除了包括 IP 地址和 MAC 地址外，还需要包括以下两种表项内容之一：
 - 该 ARP 表项所在 VLAN 和出接口；
 - 该 ARP 表项的入接口和出接口对应关系。
- 短静态 ARP 表项只包括 IP 地址和 MAC 地址。

如果出接口是三层以太网接口，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，需要对表项进行解析：当要发送 IP 数据报时，设备先发送 ARP 请求报文，如果收到的响应报文中的发送端 IP 地址和发送端 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，此时，该短静态 ARP 表项由未解析状态变为解析状态，之后就可以用于报文转发。

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。

3. Rule ARP 表项

Rule ARP 表项不会被老化，不能通过 ARP 报文更新，可以被静态 ARP 表项覆盖，可以直接用于转发报文。Rule ARP 表项可由如下特性添加：

- IPoE，IPoE 的详细介绍请参见“安全配置指导”中的“IPoE”。
- Portal，Portal 的详细介绍请参见“安全配置指导”中的“Portal”。
- VXLAN，VXLAN 的详细介绍请参见“VXLAN 配置指导”中的“VXLAN”。

1.2 vSystem 相关说明

非缺省 vSystem 支持本特性部分功能，包括手工添加静态 ARP 表项。



说明

非缺省 vSystem 对具体命令的支持情况，请见本特性的命令参考。有关 vSystem 的详细介绍请参见“虚拟化技术配置指导”中的“vSystem”。

1.3 ARP 配置任务简介

本节中的所有配置均为可选，请根据实际情况选择配置。

- [手工添加静态 ARP 表项](#)
 - [手工添加短静态 ARP 表项](#)
 - [手工添加长静态 ARP 表项](#)
- [配置动态 ARP 表项的相关功能](#)
 - [配置设备学习动态 ARP 表项的最大数目](#)
 - [配置接口学习动态 ARP 表项的最大数目](#)

- [配置动态 ARP 表项的老化时间](#)
- [开启动态 ARP 表项的检查功能](#)
- [开启在地址借用的接口学习不同网段 ARP 表项的功能](#)
- [开启 ARP 日志信息功能](#)

1.4 手工添加静态ARP表项

静态 ARP 表项在设备正常工作期间一直有效。

1.4.1 手工添加短静态 ARP 表项

1. 配置限制和指导

对于已经解析的短静态 ARP 表项，会由于外部事件，比如解析到的出接口状态 down 或设备的 ARP 表项所对应的 VLAN 或 VLAN 接口被删除等原因，恢复到未解析状态。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 手工添加短静态 ARP 表项。

```
arp static ip-address mac-address [ vpn-instance vpn-instance-name ]  
[ description text ]
```

1.4.2 手工添加长静态 ARP 表项

1. 功能简介

长静态 ARP 表项根据设备的当前状态可能处于有效或无效两种状态。处于无效状态的原因可能是该 ARP 表项中的 IP 地址与本地 IP 地址冲突或设备上没有与该 ARP 表项中的 IP 地址在同一网段的接口地址等原因。处于无效状态的长静态 ARP 表项不能指导报文转发。当长静态 ARP 表项所对应的 VLAN 或 VLAN 接口被删除时，该 ARP 表项会被删除。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 手工添加长静态 ARP 表项。

```
arp static ip-address mac-address [ vlan-id interface-type  
interface-number | interface-type interface-number interface-type  
interface-number ] [ vpn-instance vpn-instance-name ] [ description  
text ]
```

1.5 配置动态ARP表项的相关功能

1.5.1 配置设备学习动态 ARP 表项的最大数目

1. 功能简介

设备可以通过 ARP 协议自动生成动态 ARP 表项。为了防止用户占用过多的 ARP 资源，可以通过设置设备学习动态 ARP 表项的最大数目来进行限制。当设备学习动态 ARP 表项的数目达到所设置的值时，该设备上将不再学习动态 ARP 表项。

当本命令配置的动态 ARP 表项的最大数目小于设备当前已经学到的动态 ARP 表项数目，已学到的动态 ARP 表项不会被直接删除，用户可以通过执行 **reset arp dynamic** 命令直接清除动态 ARP 表项。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置设备允许学习动态 ARP 表项的最大数目。

（独立运行模式）

```
arp max-learning-number max-number slot slot-number [ cpu cpu-number ]
```

（IRF 模式）

```
arp max-learning-number max-number chassis chassis-number slot  
slot-number [ cpu cpu-number ]
```

缺省情况下，设备允许学习动态 ARP 表项的最大数目为 16384。

当配置设备允许学习动态 ARP 表项的最大数目为 0 时，表示禁止本设备学习动态 ARP 表项。

1.5.2 配置接口学习动态 ARP 表项的最大数目

1. 功能简介

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 表项的数目达到所设置的值时，该接口将不再学习动态 ARP 表项。

如果二层接口及其所属的 VLAN 接口都配置了允许学习动态 ARP 表项的最大数目，则只有二层接口及 VLAN 接口上的动态 ARP 表项数目都没有超过各自配置的最大值时，才会学习 ARP 表项。

设备各接口学习的动态 ARP 表项之和不会超过该设备学习动态 ARP 表项的最大数目。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口允许学习动态 ARP 表项的最大数目。

```
arp max-learning-num max-number
```

缺省情况下，设备允许学习动态 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。

1.5.3 配置动态 ARP 表项的老化时间

1. 功能简介

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生存周期，到达生存周期仍得不到刷新的记录将从 ARP 表中删除，这个生存周期被称作老化时间。如果在到达老化时间前记录被刷新，则重新计算老化时间。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置动态 ARP 表项的老化时间。

```
arp timer aging aging-time
```

缺省情况下，动态 ARP 表项的老化时间为 20 分钟。

1.5.4 开启动态 ARP 表项的检查功能

1. 功能简介

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

- 开启 ARP 表项的检查功能后，设备上不能学习 ARP 报文中发送端 MAC 地址为组播 MAC 的动态 ARP 表项，也不能手工添加 MAC 地址为组播 MAC 的静态 ARP 表项。
- 关闭 ARP 表项的检查功能后，设备可以学习以太网源 MAC 地址为单播 MAC 且 ARP 报文中发送端 MAC 地址为组播 MAC 的动态 ARP 表项，也可以手工添加 MAC 地址为组播 MAC 的静态 ARP 表项。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启动态 ARP 表项的检查功能。

```
arp check enable
```

缺省情况下，动态 ARP 表项的检查功能处于开启状态。

1.6 开启在地址借用的接口学习不同网段 ARP 表项的功能

1. 功能简介

在某些组网环境中，当配置某个接口借用指定接口 IP 地址后，由于借到的地址所在的网段和对端接口的地址可能处于不同网段，导致接口收到不同网段的 ARP 报文时无法学习到 ARP 表项。配置本功能后，可以使接口收到不在同一网段的 ARP 报文后学习对应的 ARP 表项，保证该接口和对端可以通信。

关闭本功能后，配置了地址借用功能的接口不再学习不同网段的 ARP 表项，已经学到的 ARP 表项老化后删除。

2. 配置限制和指导

请不要在非地址借用的接口执行 **arp ip-unnumbered learning enable** 命令，否则可能会出现学习 ARP 表项异常的问题。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置本接口借用指定接口的 IP 地址。

```
ip address unnumbered interface interface-type interface-number
```

缺省情况下，本接口未借用其它接口的 IP 地址。

- (4) 开启在地址借用的接口学习不同网段 ARP 表项的功能。

```
arp ip-unnumbered learning enable
```

缺省情况下，在地址借用的接口学习不同网段 ARP 表项的功能处于关闭状态。

1.7 开启ARP日志信息功能

1. 功能简介

ARP 日志可以方便管理员定位问题和解决问题，对处理 ARP 报文的信息进行的记录。例如，ARP 日志可以记录如下事件：

- 设备未使能 ARP 代理功能时收到目的 IP 不是设备接口 IP 地址、VRRP 备份组中的虚拟 IP 地址或 NAT 转换的外部网络地址；
- 收到的 ARP 报文中源地址和接收接口 IP 地址、VRRP 备份组中的虚拟 IP 地址或 NAT 转换的外部网络地址冲突，且此报文不是 ARP 请求报文等。

设备生成的 ARP 日志信息会交给信息中心模块处理，信息中心模块的配置将决定日志信息的发送规则和发送方向。关于信息中心的详细描述请参见“网络管理和监控配置指导”中的“信息中心”。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 ARP 日志信息功能。

```
arp check log enable
```

缺省情况下，ARP 日志信息功能处于关闭状态。

1.8 ARP显示和维护



注意

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

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 表项。



说明

非缺省 vSystem 不支持部分显示和维护命令，具体情况请参见本特性的命令参考。

表1-1 ARP 显示和维护

操作	命令
显示ARP表项	(独立运行模式) display arp [[all dynamic static] [slot <i>slot-number</i> [cpu <i>cpu-number</i>]] vlan <i>vlan-id</i> interface <i>interface-type</i> <i>interface-number</i>] [count verbose] (IRF模式) display arp [[all dynamic static] [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>]] vlan <i>vlan-id</i> interface <i>interface-type</i> <i>interface-number</i>] [count verbose]
显示指定IP地址的ARP表项	(独立运行模式) display arp ip-address [slot <i>slot-number</i> [cpu <i>cpu-number</i>]] [verbose] (IRF模式) display arp ip-address [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>]] [verbose]
显示动态ARP表项的老化时间	display arp timer aging
显示指定VPN实例的ARP表项	display arp vpn-instance <i>vpn-instance-name</i> [count verbose]
清除ARP表项	(独立运行模式) reset arp { all dynamic interface <i>interface-type</i> <i>interface-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>] static } (IRF模式) reset arp { all chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>] dynamic interface <i>interface-type</i> <i>interface-number</i> static }

2 免费 ARP

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 地址都是本机的 IP 地址。设备通过对外发送免费 ARP 报文来确定其他设备的 IP 地址是否与本机的 IP 地址冲突，并实现在设备硬件地址改变时通知其它设备更新 ARP 表项。

2.1.1 IP 地址冲突检测

设备接口获取到 IP 地址时可以在接口所在局域网内广播发送免费 ARP 报文。如果设备收到 ARP 应答报文，表示局域网中存在与该设备 IP 地址相同的设备，则设备不会使用此 IP 地址，并打印日志提示管理员修改该 IP 地址。如果设备未收到 ARP 应答报文，表示局域网中不存在与该设备 IP 地址相同的设备，则设备可以正常使用 IP 地址。

2.1.2 免费 ARP 报文学习

开启了免费 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（发送端 IP 地址、发送端 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 报文中的发送端 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。

2.1.3 定时发送免费 ARP

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：

- 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。为了降低这种仿冒网关的 ARP 攻击所带来的影响，可以在网关的接口上开启定时发送免费 ARP 功能。开启该功能后，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。
- 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开启定时发送免费 ARP 功能。启用该功能后，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。

- 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 Master 路由器 VRRP 虚拟 IP 地址相同的设备。免费 ARP 报文中的发送端 MAC 为 VRRP 虚拟路由器对应的虚拟 MAC 地址。

2.2 免费ARP配置任务简介

本节中的所有配置均为可选，请根据实际情况选择配置。当以下功能均未开启时，免费 ARP 的冲突地址检测功能仍然生效。

- [开启源 IP 地址冲突提示功能](#)
- [开启免费 ARP 报文学习功能](#)
- [开启定时发送免费 ARP 功能](#)
- [开启设备收到非同一网段 ARP 请求时发送免费 ARP 报文功能](#)
- [配置当接口 MAC 地址变化时，该接口重新发送免费 ARP 报文的次数和时间间隔](#)

2.3 开启源IP地址冲突提示功能

1. 功能简介

设备接收到其它设备发送的 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，该设备会根据当前源 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时，设备发送一个免费 ARP 报文确认是否冲突，只有收到对应的 ARP 应答后才提示存在 IP 地址冲突。
- 如果源 IP 地址冲突提示功能处于开启状态时，设备立刻提示存在 IP 地址冲突。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启源 IP 地址冲突提示功能。

```
arp ip-conflict log prompt
```

缺省情况下，源 IP 地址冲突提示功能处于关闭状态。

2.4 开启免费ARP报文学习功能

- (1) 进入系统视图。

```
system-view
```

- (2) 开启免费 ARP 报文学习功能。

gratuitous-arp-learning enable

缺省情况下，免费 ARP 报文的学习功能处于开启状态。

2.5 开启定时发送免费ARP功能

1. 配置限制和指导

- 设备最多允许同时在 1024 个接口上开启定时发送免费 ARP 功能。
- 开启定时发送免费 ARP 功能后，只有当接口链路状态 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 报文的发送时间间隔，则在下一个发送时间间隔才能生效。
- 如果同时在很多接口下开启定时发送免费 ARP 功能，或者每个接口有大量的从 IP 地址，又或者是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 报文的实际发送时间间隔可能会远远高于用户设定的时间间隔。

2. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 进入接口视图。

interface *interface-type* *interface-number*

- (3) 开启定时发送免费 ARP 功能。

arp send-gratuitous-arp [*interval interval*]

缺省情况下，定时发送免费 ARP 功能处于关闭状态。

2.6 开启设备收到非同一网段ARP请求时发送免费ARP报文功能

- (1) 进入系统视图。

system-view

- (2) 开启设备收到非同一网段 ARP 请求时发送免费 ARP 报文功能。

gratuitous-arp-sending enable

缺省情况下，设备收到非同一网段的 ARP 请求时发送免费 ARP 报文功能处于关闭状态。

2.7 配置当接口MAC地址变化时，该接口重新发送免费ARP报文的次数和时间间隔

1. 功能简介

当设备的 MAC 地址发生变化后，设备会通过免费 ARP 报文将修改后的 MAC 地址通告给其他设备。由于目前免费 ARP 报文没有重传机制，其他设备可能无法收到免费 ARP 报文。为了解决这个问题，用户可以配置当接口 MAC 地址变化时，该接口重新发送免费 ARP 报文的次数和时间间隔，保证其他设备可以收到该免费 ARP 报文。

2. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 配置当接口 MAC 地址变化时，重新发送免费 ARP 报文的次数和时间间隔

gratuitous-arp mac-change retransmit *times* interval *seconds*

缺省情况下，当设备的接口 MAC 地址变化时，该接口只会发送一次免费 ARP 报文。

3 代理 ARP

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 的应用场景为：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 的应用场景为：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。

3.2 开启普通代理ARP功能

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

普通代理 ARP 功能可在 VLAN 接口视图/三层以太网接口视图/三层以太网子接口视图/三层聚合接口视图/三层聚合子接口视图下进行配置。

- (3) 开启普通代理 ARP 功能。

```
proxy-arp enable
```

缺省情况下，普通代理 ARP 功能处于关闭状态。

3.3 开启本地代理ARP功能

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

本地代理 ARP 功能可在 VLAN 接口视图/三层以太网接口视图/三层以太网子接口视图/三层聚合接口视图/三层聚合子接口视图下进行配置。

- (3) 开启本地代理 ARP 功能。

```
local-proxy-arp enable [ ip-range start-ip-address to end-ip-address ]
```

缺省情况下，本地代理 ARP 功能处于关闭状态。

3.4 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3-1 代理 ARP 显示和维护

操作	命令
显示本地代理ARP的状态	display local-proxy-arp [interface <i>interface-type</i> <i>interface-number</i>]
显示普通代理ARP的状态	display proxy-arp [interface <i>interface-type</i> <i>interface-number</i>]

4 ARP Snooping

4.1 ARP Snooping简介

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。

4.1.1 ARP Snooping 表项建立机制

设备上在一个 VLAN 中启用 ARP Snooping 后，该 VLAN 内接收的 ARP 报文都会被上送到 CPU。CPU 对上送的 ARP 报文进行分析，获取 ARP 报文的发送端 IP 地址、发送端 MAC 地址、VLAN 和入端口信息，建立记录用户信息的 ARP Snooping 表项。

4.1.2 ARP Snooping 表项老化机制

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 12 分钟内没有收到 ARP 更新报文，设备会向外主动发送一个 ARP 请求进行探测；若 ARP Snooping 表项自最后一次更新后 15 分钟时，还没有收到 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。当收到发送端 IP 地址和发送端 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。

4.1.3 ARP Snooping 表项冲突处理机制

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 1 分钟后删除此表项。

4.2 开启ARP Snooping功能

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 VLAN 视图。

```
vlan vlan-id
```

- (3) 开启 ARP Snooping 功能。

```
arp snooping enable
```

缺省情况下，ARP Snooping 功能处于关闭状态。

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4-1 ARP Snooping 显示和维护

操作	命令
显示ARP Snooping表项	<p>(独立运行模式)</p> <pre>display arp snooping [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [cpu <i>cpu-number</i>] [count]</pre> <pre>display arp snooping ip <i>ip-address</i> [slot <i>slot-number</i>] [cpu <i>cpu-number</i>]]</pre> <p>(IRF模式)</p> <pre>display arp snooping [vlan <i>vlan-id</i>] [chassis <i>chassis-number</i>] [slot <i>slot-number</i>] [cpu <i>cpu-number</i>]] [count]</pre> <pre>display arp snooping ip <i>ip-address</i> [chassis <i>chassis-number</i>] [slot <i>slot-number</i>] [cpu <i>cpu-number</i>]]</pre>
清除ARP Snooping表项	<pre>reset arp snooping [ip <i>ip-address</i> vlan <i>vlan-id</i>]</pre>

目 录

1 IP 地址	1-1
1.1 IP 地址简介	1-1
1.1.1 IP 地址的表示和分类	1-1
1.1.2 特殊的 IP 地址	1-2
1.1.3 子网和掩码	1-2
1.1.4 IP 地址的获取方式	1-2
1.2 vSystem 相关说明	1-3
1.3 手工指定接口的 IP 地址	1-3
1.4 配置接口借用 IP 地址	1-4
1.5 IP 地址显示和维护	1-4

1 IP 地址

1.1 IP地址简介

若非特别指明，本文所指的 IP 地址均为 IPv4 地址。

1.1.1 IP 地址的表示和分类

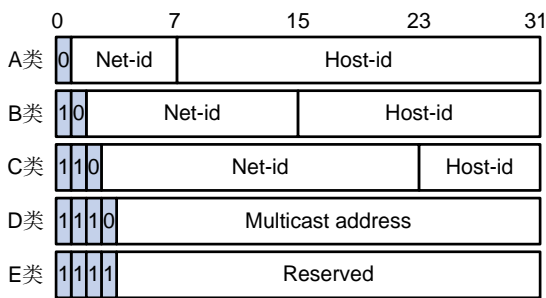
连接到 IPv4 网络上的设备通过 IP 地址标识。IP 地址长度为 32 比特，通常采用点分十进制方式表示，即每个 IP 地址被表示为以小数点隔开的 4 个十进制整数，每个整数对应一个字节，如 10.1.1.1。

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
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，其它地址保留今后使用

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。

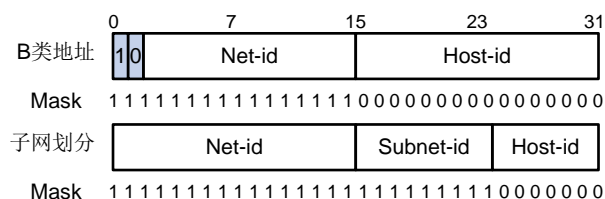
1.1.3 子网和掩码

随着 Internet 的快速发展，IP 地址已近枯竭。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 $512 \times 126 = 64512$ 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。

1.1.4 IP 地址的获取方式

接口获取 IP 地址有以下几种方式：

- 通过手动指定 IP 地址，本手册只介绍通过手动指定 IP 地址的方式。
- 通过 BOOTP 分配得到 IP 地址，通过 BOOTP 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“BOOTP 客户端”。
- 通过 DHCP 分配得到 IP 地址，通过 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 客户端”。
- 通过 PPP 协商获得 IP 地址，通过 PPP 协商获得 IP 地址方式的介绍请参见“二层技术-广域网接入配置指导”中的“PPP”。

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,首先通过手动指定了 IP 地址,然后使用 DHCP 协议申请 IP 地址,那么手动指定的 IP 地址会被删除,接口的 IP 地址是通过 DHCP 协议分配的。

1.2 vSystem相关说明

vSystem 支持本特性的所有功能。有关 vSystem 的详细介绍请参见“虚拟化技术配置指导”中的“vSystem”。

1.3 手工指定接口的IP地址

1. 功能简介

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。

在 RBM 组网下可以为 IP 地址赋予浮动属性,浮动 IP 地址可以简化 HA (High Availability, 高可靠性) 功能的配置。将浮动 IP 地址配置在 HA 主设备接入下行租户的接口,该地址会自动同步到备设备,不需要在主/备设备的业务接口上配置 VRRP (Virtual Router Redundancy Protocol, 虚拟路由器冗余协议) 虚拟地址。有关 HA 和 RBM 的详细介绍请参见“可靠性”中的“高可靠性”,有关 VRRP 的详细介绍请参见“可靠性”中的“VRRP”。

2. 配置限制和指导

- 一个接口只能有一个主 IP 地址,在同一个接口上多次执行本命令,最后一次执行的命令生效。
- 当接口被配置为通过 BOOTP、DHCP、PPP 方式获取 IP 地址或借用其它接口的 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、主接口及其子接口之间、同一主接口下不同子接口之间的 IP 地址不可以在同一网段。
- 设备支持在不同接口上配置掩码不同但最短掩码对应网络位相同的地址,比如地址 1.1.1.1/16 和 1.1.2.1/24,这两个地址的最短掩码 16 对应的网络位都是 1.1.0.0。缺省连接这两个接口上的用户不能互通,如需互通,需要配置普通代理 ARP 功能,关于普通代理 ARP 的描述,请参见“三层技术-IP 业务配置指导”中的“ARP”。
- 浮动 IP 地址仅在主备模式的主设备上配置,不支持双主模式,不支持在备设备上配置、修改或删除。
- 通过 **ip address** 命令配置同样的 IP 地址,但不携带 **float** 参数,可以取消该地址的浮动属性。

3. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 进入接口视图。

```
interface interface-type interface-number
```

(3) 配置接口的 IP 地址。

```
ip address ip-address { mask-length | mask } [ sub ] [ float ]
```

缺省情况下，未配置接口 IP 地址。

1.4 配置接口借用IP地址

1. 功能简介

IP 地址借用是指一个接口上未配置 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 地址借用的使用场景如下：

- 在 IP 地址资源比较匮乏的环境下，为了节约 IP 地址资源，可以配置某个接口借用其它接口的 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它接口的 IP 地址，而不必让其一直占用一个单独的 IP 地址。

2. 配置限制和指导

- Loopback 接口的 IP 地址可被其它接口借用，但本身不能借用其它接口的地址。
- 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手动配置的 IP 地址，则只有手动配置的主 IP 地址能被借用。
- 由于借用方接口本身没有 IP 地址，无法在此接口上启用动态路由协议。所以必须手动配置一条到对端网段的静态路由，才能实现设备间的连通。
- 三层以太网子接口或三层聚合子接口通过地址借用功能借用到 IP 地址后，无法学习到与借用地址对应的 ARP 表项，故不建议在以太网子接口或三层聚合子接口配置地址借用功能。

3. 配置准备

被借用接口的 IP 地址已经配置，配置方法可以为手动指定、通过 BOOTP 或 DHCP 动态获取或通过 PPP 协商分配。

4. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置本接口借用指定接口的 IP 地址。

```
ip address unnumbered interface interface-type interface-number
```

缺省情况下，本接口未借用其它接口的 IP 地址。

1.5 IP地址显示和维护

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1-2 IP 地址的显示和维护

操作	命令
显示IPv4浮动地址详细信息	<p>(独立运行模式)</p> <p>display ip address float [interface <i>interface-type</i> <i>interface-number</i> vpn-instance <i>vpn-instance-name</i> [<i>ip-address</i>]] [slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</p> <p>(IRF模式)</p> <p>display ip address float [interface <i>interface-type</i> <i>interface-number</i> vpn-instance <i>vpn-instance-name</i> [<i>ip-address</i>]] [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</p>
显示三层接口与IP相关的简要信息	display ip interface [<i>interface-type</i> [<i>interface-number</i>]] brief [description]
显示三层接口与IP相关的配置和统计信息	display ip interface [<i>interface-type</i> [<i>interface-number</i>]]

目 录

1 DHCP 概述	1-1
1.1 DHCP 组网模型	1-1
1.2 DHCP 的 IP 地址分配	1-1
1.2.1 IP 地址分配策略	1-1
1.2.2 IP 地址获取过程	1-2
1.2.3 IP 地址的租约更新	1-2
1.3 DHCP 报文格式	1-3
1.4 DHCP 选项介绍	1-4
1.5 DHCP 常用选项	1-4
1.6 自定义 DHCP 选项	1-5
1.6.1 厂商特定信息选项 (Option 43)	1-5
1.6.2 中继代理信息选项 (Option 82)	1-6
1.7 协议规范	1-7
2 DHCP 服务器	2-1
2.1 DHCP 服务器简介	2-1
2.1.1 地址池的地址管理方式	2-1
2.1.2 地址池的选取原则	2-2
2.1.3 DHCP 服务器分配 IP 地址的优先次序	2-2
2.2 DHCP 服务器配置任务简介	2-3
2.3 创建 DHCP 用户类	2-4
2.4 配置 DHCP 服务器的地址池	2-4
2.4.1 DHCP 服务器地址池配置任务简介	2-4
2.4.2 创建 DHCP 地址池	2-5
2.4.3 配置一个主网段多个地址范围的动态地址管理方式	2-5
2.4.4 配置一个主网段多个从网段的动态地址管理方式	2-6
2.4.5 配置静态地址绑定	2-8
2.4.6 配置 DHCP 客户端使用的网关地址	2-9
2.4.7 配置 DHCP 客户端使用的域名后缀	2-9
2.4.8 配置 DHCP 客户端使用的 DNS 服务器地址	2-10
2.4.9 配置 DHCP 客户端使用的 WINS 服务器地址和 NetBIOS 节点类型	2-10
2.4.10 配置 DHCP 客户端使用的 BIMS 服务器信息	2-11
2.4.11 配置 DHCP 客户端使用的远程启动文件信息	2-12

2.4.12 配置 DHCP 客户端使用的下一个提供服务的服务器 IP 地址.....	2-12
2.4.13 自定义 DHCP 选项.....	2-13
2.4.14 为 DHCP 服务器上的地址池绑定 VPN 实例	2-14
2.4.15 配置 DHCP 用户类白名单功能	2-15
2.4.16 配置 DHCP 服务器辅助网关信息	2-16
2.4.17 配置 DHCP 服务器辅助路由信息	2-16
2.5 配置接口引用地址池	2-17
2.6 配置 DHCP 策略动态分配地址和其他参数	2-18
2.7 开启 DHCP 服务	2-18
2.8 配置接口工作在 DHCP 服务器模式	2-19
2.9 配置 IP 地址冲突检测功能	2-19
2.10 配置 Option 82 的处理方式.....	2-20
2.11 配置 DHCP 服务器兼容性	2-20
2.11.1 配置 DHCP 服务器始终以广播方式回复请求报文.....	2-20
2.11.2 配置 DHCP 服务器忽略 BOOTP 请求报文	2-21
2.11.3 配置 DHCP 服务器以 RFC 1048 规定的格式发送 BOOTP 应答报文.....	2-21
2.11.4 配置 DHCP 服务器发送 DHCP 应答报文不携带 Option 60 选项	2-22
2.12 配置 DHCP 服务器发送 DHCP 报文的 DSCP 优先级	2-22
2.13 配置 DHCP 服务器租约固化功能	2-22
2.14 开启 DHCP 服务器的用户下线探测功能.....	2-23
2.15 配置 DHCP 告警功能.....	2-23
2.16 开启 DHCP 服务器日志信息功能	2-24
2.17 DHCP 服务器显示和维护	2-24
2.18 DHCP 服务器常见故障处理.....	2-25
2.18.1 DHCP 客户端获取到冲突的 IP 地址.....	2-25
3 DHCP 中继.....	3-1
3.1 DHCP 中继简介	3-1
3.1.2 DHCP 中继的基本原理	3-1
3.1.3 DHCP 中继支持 Option 82 功能.....	3-2
3.2 DHCP 中继配置任务简介	3-2
3.3 开启 DHCP 服务	3-3
3.4 配置接口工作在 DHCP 中继模式.....	3-3
3.5 指定 DHCP 服务器的地址	3-3
3.5.1 指定 DHCP 中继对应的 DHCP 服务器地址	3-3
3.5.2 指定中继地址池对应的 DHCP 服务器地址	3-4
3.6 指定 DHCP 客户端对应的 DHCP 中继地址池	3-5

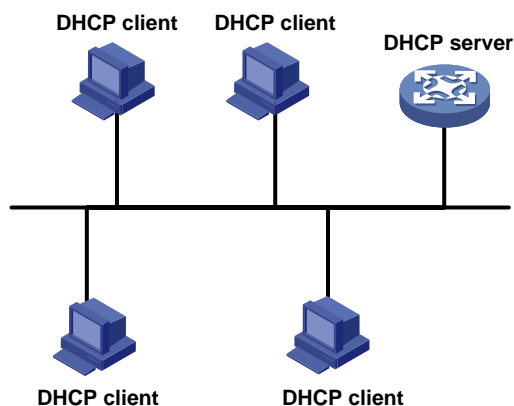
3.7 配置 DHCP 中继的安全功能	3-5
3.7.1 配置 DHCP 中继用户地址表项记录功能	3-5
3.7.2 配置 DHCP 中继动态用户地址表项定时刷新功能	3-6
3.7.3 配置防止 DHCP 饿死攻击	3-6
3.7.4 配置 DHCP 中继支持代理功能	3-7
3.7.5 配置 DHCP 中继的用户下线探测功能	3-8
3.8 配置通过 DHCP 中继释放客户端的 IP 地址	3-8
3.9 配置 DHCP 中继支持 Option 82 功能	3-9
3.10 配置 DHCP 中继发送 DHCP 报文的 DSCP 优先级	3-9
3.11 配置 DHCP 中继在 DHCP 报文中填充的中继地址	3-10
3.11.1 手工指定在 DHCP 报文中填充的中继地址	3-10
3.11.2 通过 smart-relay 功能指定 DHCP 报文中填充的中继地址	3-10
3.12 指定 DHCP 中继向 DHCP 服务器转发报文的源地址	3-11
3.13 配置 DHCP 中继通过 Option82 信息转发 DHCP 应答报文	3-11
3.14 DHCP 中继显示和维护	3-12
3.15 DHCP 中继常见故障处理	3-13
3.15.1 DHCP 客户端无法通过 DHCP 中继获取配置信息	3-13
4 DHCP 客户端	4-1
4.1 DHCP 客户端简介	4-1
4.2 DHCP 客户端配置限制和指导	4-1
4.3 DHCP 客户端配置任务简介	4-1
4.4 配置接口通过 DHCP 协议获取 IP 地址	4-1
4.5 配置接口使用的 DHCP 客户端 ID	4-2
4.6 开启地址冲突检查功能	4-2
4.7 配置 DHCP 客户端发送 DHCP 报文的 DSCP 优先级	4-3
4.8 DHCP 客户端显示和维护	4-3
5 BOOTP 客户端	5-1
5.1 BOOTP 客户端简介	5-1
5.1.1 BOOTP 客户端的应用环境	5-1
5.1.2 IP 地址动态获取过程	5-1
5.1.3 协议规范	5-1
5.2 配置接口通过 BOOTP 协议获取 IP 地址	5-1
5.3 BOOTP 客户端显示和维护	5-2

1 DHCP 概述

1.1 DHCP组网模型

DHCP（Dynamic Host Configuration Protocol，动态主机配置协议）采用客户端/服务器模式，由服务器为网络设备动态地分配 IP 地址等网络配置参数。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](#)”。

图1-1 同网段 DHCP 组网应用



1.2 DHCP的IP地址分配

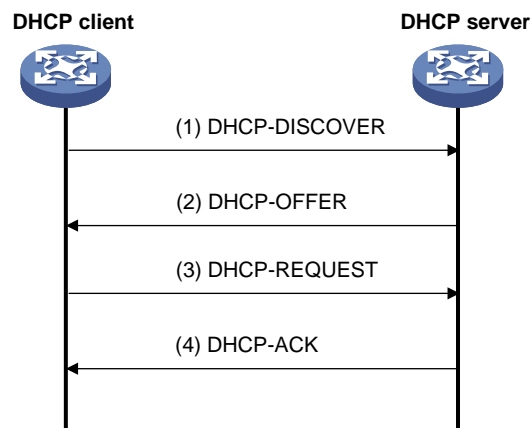
1.2.1 IP 地址分配策略

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 地址分配给客户端。
- 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。
- (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。

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未收到回应，并且客户端上不存在与该地址同网段的其他地址时，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

DHCP 服务器分配给客户端的 IP 地址具有一定的租借期限（除自动分配的 IP 地址），该租借期限称为租约。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，则 DHCP 客户端需要申请延长 IP 地址租约。

在 DHCP 客户端的 IP 地址租约期限达到一半左右时间时，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。

如果在租约的一半左右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，每种报文的格式都相同，只是某些字段的取值不同。DHCP 的报文格式如图 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

0	7	15	23	31
op (1)	htype (1)	hlen (1)	hops (1)	
xid (4)				
secs (2)		flags (2)		
ciaddr (4)				
yiaddr (4)				
siaddr (4)				
giaddr (4)				
chaddr (16)				
sname (64)				
file (128)				
options (variable)				

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 options 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。如果客户端有合法和可用的 IP 地址，则将其添加到此字段，否则字段设置为 0。此字段不用于客户端申请某个特定的 IP 地址。
- yiaddr: DHCP 服务器分配给客户端的 IP 地址。
- siaddr: DHCP 客户端获取启动配置信息的服务器 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 客户端获取启动配置信息的服务器名称。
- 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。

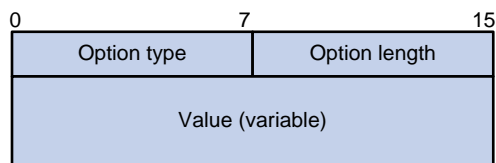
1.4 DHCP选项介绍

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 (Options) 字段。DHCP 在 BOOTP 基础上增加的功能, 通过 Options 字段来实现。

DHCP 利用 Options 字段传递控制信息和网络配置参数, 实现地址动态分配的同时, 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



1.5 DHCP常用选项

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。如果 Option 3 和 Option 121 同时存在, 则忽略 Option 3。
- 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 (即目的网络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（即目的网络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 3、Option 33 和 Option 121 同时存在，则忽略 Option 3 和 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。

1.6 自定义DHCP选项

有些选项的内容，RFC 2132 中没有统一规定，例如 Option 43、Option 82 和 Option 184。下面将介绍设备上定义的几种选项。

1.6.1 厂商特定信息选项（Option 43）

1. Option 43 的作用

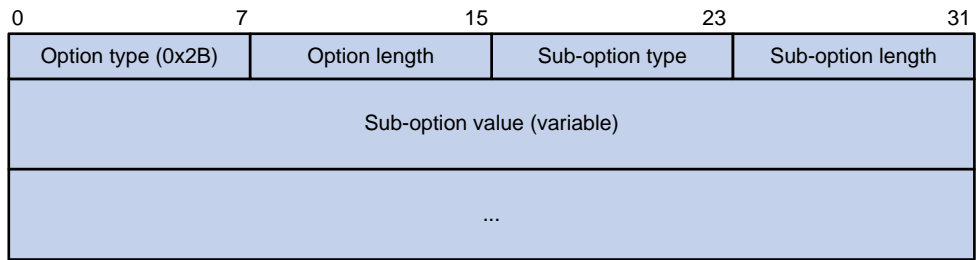
Option 43 称为厂商特定信息选项。DHCP 服务器和 DHCP 客户端通过 Option 43 交换厂商特定的信息。

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。

2. 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。

3. 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 20）隔开。

图1-6 ACS 参数子选项取值字段的格式

URL of ACS (variable)	20
User name of ACS (variable)	20
Password of ACS (variable)	

- 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取值字段的格式

0	7	15
PXE server type (0x0000)		
Server number		
Server IP addresses (variable)		

1.6.2 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。目前设备只支持两个子选项：sub-option 1（Circuit ID，电路 ID 子选项）、sub-option 2（Remote ID，远程 ID 子选项）。

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，Circuit ID 的填充模式有以下几种：

- 采用 string 模式填充：sub-option 1 的内容是用户配置的字符串。
- 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。
- 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，接收到 DHCP 客户端请求报文的接口类型、接口编号和接口所属的 VLAN ID。

Remote ID 的填充模式有以下几种：

- 采用 **string** 模式填充：sub-option 2 的内容是用户配置的字符串。
- 采用 **normal** 模式填充：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）。
- 采用 **sysname** 模式填充：sub-option 2 的内容是设备的系统名称。设备的系统名称可以通过系统视图下的 **sysname** 命令配置。

1.7 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

2.1 DHCP服务器简介

DHCP 服务器通过地址池保存 IP 地址和网络参数，从地址池中选择 IP 地址和网络参数分配给客户端。

2.1.1 地址池的地址管理方式

地址池的地址管理方式有以下几种：静态绑定 IP 地址，即通过将客户端的 MAC 地址或客户端 ID 与 IP 地址绑定的方式，实现为特定的客户端分配特定的 IP 地址；动态选择 IP 地址，即在地址池中指定可供分配的 IP 地址范围，当收到客户端的 IP 地址申请时，从该地址范围中动态选择 IP 地址，分配给该客户端。

在地址池中指定可供分配的 IP 地址范围，有以下几种方法：

1. 为地址池指定一个主网段，并将该网段划分为多个地址范围。

多个地址范围是指一个地址池动态分配的 IP 地址范围（公共地址范围）和多个为 DHCP 用户类分配的 IP 地址范围。

DHCP 服务器通过定义 DHCP 用户类，实现为满足特定条件的客户端分配特定地址范围的 IP 地址。DHCP 服务器根据客户端发送的请求报文，判断 DHCP 客户端所属的用户类。每个用户类可以配置多个匹配条件，只要客户端发送的 DHCP 请求报文满足任意一个匹配条件，就认为该客户端属于该用户类。在地址池下，可以为不同的用户类指定不同的地址范围。如果 DHCP 客户端属于某个用户类，则从该用户类的地址范围内选择地址分配给该客户端。

采用这种地址管理方式时，地址选择过程为：

- (1) 按照地址池下用户类地址范围的配置顺序，将 DHCP 客户端和用户类进行匹配。
- (2) 如果 DHCP 客户端属于某个用户类，则从该用户类的地址范围中选择地址分配给客户端。
- (3) 如果该用户类中没有可供分配的地址，则继续匹配下一个用户类。如果所有匹配上的用户类地址范围都没有可供分配的地址，则从公共地址范围中选择地址分配给客户端。
- (4) 如果 DHCP 客户端不属于任何一个 DHCP 用户类，则会从地址池动态分配的 IP 地址范围（通过 **address range** 命令配置）中选择地址分配给 DHCP 客户端。
- (5) 如果动态分配的 IP 地址范围内也没有空闲地址，或者未配置动态分配的 IP 地址范围，则地址分配失败，即 DHCP 服务器无法为 DHCP 客户端分配地址。



说明

每个地址范围内的地址都必须属于指定的主网段，否则无法分配该范围内的地址。

2. 为地址池指定一个主网段，并指定多个从网段。

采用此种地址分配方式时，地址选择的过程是：首先从地址池主网段中查找可供分配的 IP 地址。如果主网段中没有可供分配的 IP 地址，则按照该地址池下从网段的配置顺序，依次查找可供分配的 IP 地址。

2.1.2 地址池的选取原则

DHCP 服务器为客户端分配 IP 地址时，按照如下顺序选择地址池：

- (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 地址和其他网络参数分配给客户端。
- (2) 如果配置了 DHCP 策略，则 DHCP 客户端匹配某个 DHCP 用户类时，DHCP 服务器选择与该 DHCP 用户类关联的 DHCP 地址池；DHCP 客户端未匹配到 DHCP 用户类时，若配置了默认 DHCP 地址池，则选择该 DHCP 地址池；若未配置默认 DHCP 地址池或 DHCP 默认地址池不存在可供分配的 IP 地址时，IP 地址或其他参数分配失败。
- (3) 如果接收到 DHCP 请求报文的接口引用了某个地址池，则选择该地址池，从该地址池中选取 IP 地址和其他网络参数分配给客户端。
- (4) 如果上述条件均不满足，则使用以下方法选择 DHCP 地址池：
 - 如果客户端与服务器在同一网段，则将 DHCP 请求报文接收接口的 IP 地址与所有地址池配置的主网段进行匹配，并选择最长匹配的主网段所对应的地址池。如果未匹配到主网段，则将 DHCP 请求报文接收接口的 IP 地址与所有地址池配置的从网段进行匹配，并选择最长匹配的网段所对应的地址池。
 - 如果客户端与服务器不在同一网段，即客户端通过 DHCP 中继获取 IP 地址，则将 DHCP 请求报文中 giaddr 字段指定的 IP 地址与所有地址池配置的主网段进行匹配，并选择最长匹配的网段所对应的地址池。如果未匹配到主网段，则将 DHCP 请求报文中 giaddr 字段指定的 IP 地址与所有地址池配置的从网段进行匹配，并选择最长匹配的网段所对应的地址池。

例如，DHCP 服务器上配置了两个地址池，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，且未引用地址池，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，请尽量保证其与 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。
- 建议合理规划 DHCP 服务器上各地址池中主网段的配置，尽量避免客户端匹配不到主网段、直接匹配从网段的情况发生。

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。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。
- (4) 按照“[2.1.1 地址池的地址管理方式](#)”和“[2.1.2 地址池的选取原则](#)”中所述的动态分配地址选择原则，顺序查找可供分配的 IP 地址，选择最先找到的 IP 地址。
- (5) 如果未找到可用的 IP 地址，则从当前匹配地址池中依次查询租约过期、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

- 如果客户端所在的网段发生变化，服务器不会为客户端分配曾经分配给它的 IP 地址，而是从匹配新网段的地址池中重新选择 IP 地址。
 - 使用曾经发生过冲突的 IP 地址时，只有冲突状态超过一小时的 IP 地址才能够被服务器分配给新的 DHCP 客户端。
-

2.2 DHCP服务器配置任务简介

DHCP 服务器配置任务如下：

- (1) （可选）[创建 DHCP 用户类](#)
- (2) [配置 DHCP 服务器的地址池](#)
- (3) （可选）修改 DHCP 服务器的地址池选择方式
 - [配置接口引用地址池](#)
 - [配置 DHCP 策略动态分配地址和其他参数](#)
- (4) [开启 DHCP 服务](#)
- (5) [配置接口工作在 DHCP 服务器模式](#)
- (6) （可选）配置高级功能
 - [配置 IP 地址冲突检测功能](#)
 - [配置 Option 82 的处理方式](#)
 - [配置 DHCP 服务器兼容性](#)
 - [配置 DHCP 服务器发送 DHCP 报文的 DSCP 优先级](#)
 - [配置 DHCP 服务器租约固化功能](#)
 - [开启 DHCP 服务器的用户下线探测功能](#)
- (7) （可选）配置告警及日志功能
 - [配置 DHCP 告警功能](#)
 - [开启 DHCP 服务器日志信息功能](#)

2.3 创建DHCP用户类

1. 功能简介

DHCP 用户类通过 DHCP 请求报文中的硬件地址、Option 信息或 Giaddr 字段来匹配一组特定的 DHCP 客户端，以实现为特定的 DHCP 客户端分配特定的 IP 地址和其他参数。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 创建 DHCP 用户类，并进入 DHCP 用户类视图。

```
dhcp class class-name
```

- (3) 配置 DHCP 用户类的匹配规则。

```
if-match rule rule-number { hardware-address hardware-address mask  
hardware-address-mask | option option-code [ ascii ascii-string [ offset  
offset | partial ] | hex hex-string [ mask mask | offset offset length  
length | partial ] ] | relay-agent gateway-address }
```

缺省情况下，未配置 DHCP 用户类的匹配规则。

2.4 配置DHCP服务器的地址池

2.4.1 DHCP 服务器地址池配置任务简介

DHCP 服务器地址池配置任务如下：

- (1) [创建 DHCP 地址池](#)

- (2) 配置为 DHCP 客户端分配地址

同一个地址池中不能同时配置两种动态地址管理方式，但可以同时配置动态地址管理方式和静态地址绑定。

- [配置一个主网段多个地址范围的动态地址管理方式](#)
- [配置一个主网段多个从网段的动态地址管理方式](#)
- [配置静态地址绑定](#)

- (3) 配置为 DHCP 客户端分配其他参数

- [配置 DHCP 客户端使用的网关地址](#)
- [配置 DHCP 客户端使用的域名后缀](#)
- [配置 DHCP 客户端使用的 DNS 服务器地址](#)
- [配置 DHCP 客户端使用的 WINS 服务器地址和 NetBIOS 节点类型](#)
- [配置 DHCP 客户端使用的 BIMS 服务器信息](#)
- [配置 DHCP 客户端使用的远程启动文件信息](#)
- [配置 DHCP 客户端使用的下一个提供服务的服务器 IP 地址](#)
- [自定义 DHCP 选项](#)

- (4) （可选）[为 DHCP 服务器上的地址池绑定 VPN 实例](#)

- (5) （可选）[配置 DHCP 用户类白名单功能](#)

- (6) (可选) [配置 DHCP 服务器辅助网关信息](#)
- (7) (可选) [配置 DHCP 服务器辅助路由信息](#)

2.4.2 创建 DHCP 地址池

- (1) 进入系统视图。
system-view
- (2) 创建 DHCP 地址池，并进入 DHCP 地址池视图。
dhcp server ip-pool pool-name

2.4.3 配置一个主网段多个地址范围的动态地址管理方式

1. 功能简介

在某些组网应用中，需要将一个网段下的不同客户端，按照一定的规则划分到不同的地址范围中。此时，可以按照客户端划分规则创建对应的 DHCP 用户类，并在地址池内为不同的用户类配置不同的地址范围，从而实现为特定的客户端分配特定范围的地址。在这种情况下，还可以配置一个公共地址范围，为不匹配任何用户类的客户端分配给该范围的地址。如果不配置公共地址范围，则不匹配任何用户类的客户端将无法获取到 IP 地址。

如果不需要对客户端进行分类，而仅需要限制网段内可分配的动态地址范围，则可以只配置公共地址范围，而不配置用户类的地址范围。

2. 配置限制和指导

配置为客户端分配的 IP 地址时，需要注意：

- 在同一个 DHCP 地址池中，如果多次执行 **network** 或 **address range** 命令，新的配置会覆盖已有配置；如果多次执行 **class** 命令，则可以为多个用户类指定不同的地址范围；多次执行 **forbidden-ip** 命令或 **forbidden-ip-range** 命令，可以配置多个不参与自动分配的 IP 地址或 IP 地址段。
- 在 DHCP 地址池视图下通过 **forbidden-ip** 命令或 **forbidden-ip-range** 命令配置不参与自动分配的 IP 地址或 IP 地址段后，只有当前的地址池不能分配这些 IP 地址或 IP 地址段，其他地址池仍然可以分配这些 IP 地址或 IP 地址段；通过 **dhcp server forbidden-ip** 命令指定不参与自动分配的 IP 地址后，所有地址池都不能分配这些 IP 地址。
- 当用户配置 **class range** 命令修改已存在的为 DHCP 用户类动态分配的 IP 地址范围，且新的 IP 地址范围包括之前 IP 地址范围中已分配的地址租约时，如果 DHCP 服务器收到该地址租约的续约需求，DHCP 服务器会给该 DHCP 客户端分配新的 IP 地址租约，已分配的地址租约会继续老化等待超期释放。如果需要已分配的地址租约立即释放，则需配置 **reset dhcp server ip-in-use** 命令进行清除地址租约操作。

3. 配置步骤

- (1) 进入系统视图。
system-view
- (2) 进入 DHCP 地址池视图。
dhcp server ip-pool pool-name
- (3) 配置 DHCP 地址池动态分配的主网段。

network *network-address* [*mask-length* | **mask** *mask*]

缺省情况下，未配置主网段。

- (4) （可选）配置地址池动态分配的 IP 地址范围，即公共地址范围。

address range *start-ip-address end-ip-address*

缺省情况下，未配置动态分配的 IP 地址范围。

- (5) （可选）配置 DHCP 地址池为指定 DHCP 用户类动态分配的 IP 地址范围。

class *class-name range start-ip-address end-ip-address*

缺省情况下，未配置为指定 DHCP 用户类动态分配的 IP 地址范围。

只有先通过 **dhcp class** 命令创建 DHCP 用户类，再通过本命令指定该用户类，才能为该用户类分配指定范围的地址。

- (6) （可选）配置动态分配的 IP 地址的租约有效期限。

expired { **day** *day* [**hour** *hour* [**minute** *minute* [**second** *second*]]] | **unlimited** }

缺省情况下，IP 地址租约有效期限为 1 天。

- (7) （可选）配置 DHCP 地址池中不参与自动分配的 IP 地址。

forbidden-ip *ip-address*&<1-8>

缺省情况下，未配置 DHCP 地址池中不参与自动分配的 IP 地址。

- (8) （可选）配置 DHCP 地址池中不参与自动分配的 IP 地址段。

forbidden-ip-range *start-ip-address* [*end-ip-address*]

缺省情况下，未配置 DHCP 地址池中不参与自动分配的 IP 地址段。

- (9) （可选）在系统视图配置全局不参与自动分配的 IP 地址。

- a. 退回系统视图。

quit

- b. 配置全局不参与自动分配的 IP 地址。

dhcp server forbidden-ip *start-ip-address* [*end-ip-address*]
[**vpn-instance** *vpn-instance-name*]

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。

2.4.4 配置一个主网段多个从网段的动态地址管理方式

1. 功能简介

在配置了一个主网段和多个从网段的地址池中，从网段的作用是对主网段地址空间的补充。当主网段中没有空闲地址分配给客户端时，服务器会从该地址池中的从网段获取地址分配给客户端。

2. 配置限制和指导

在 DHCP 地址池视图下通过 **forbidden-ip** 命令或 **forbidden-ip-range** 命令配置不参与自动分配的 IP 地址或 IP 地址段后，只有当前的地址池不能分配这些 IP 地址或 IP 地址段，其他地址池仍然可以分配这些 IP 地址或 IP 地址段；通过 **dhcp server forbidden-ip** 命令指定不参与自动分配的 IP 地址后，所有地址池都不能分配这些 IP 地址。

3. 在地址池中配置一个主网段和多个从网段

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置 DHCP 地址池动态分配的主网段。

```
network network-address [ mask-length | mask mask ]
```

缺省情况下，未配置主网段。

每个 DHCP 地址池中只能配置一个主网段，如果多次执行 **network** 命令配置主网段，则新的配置会覆盖已有配置。

- (4) （可选）配置 DHCP 地址池动态分配的从网段。

```
network network-address [ mask-length | mask mask ] secondary
```

缺省情况下，未配置从网段。

每个 DHCP 地址池中，最多可以配置 32 个从网段。

- (5) （可选）退回地址池视图。

```
quit
```

4. 在地址池中配置动态分配的 IP 地址的租约有效期限

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置动态分配的 IP 地址的租约有效期限。

```
expired { day day [ hour hour [ minute minute [ second second ] ] ] | unlimited }
```

缺省情况下，IP 地址租约有效期限为 1 天。

5. 配置不参与自动分配的地址

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置 DHCP 地址池中不参与自动分配的 IP 地址。

```
forbidden-ip ip-address&<1-8>
```

缺省情况下，未配置地址池中不参与自动分配的 IP 地址。

多次执行 **forbidden-ip** 命令，可以配置多个不参与自动分配的 IP 地址段。

- (4) 配置地址池中不参与自动分配的 IP 地址段。

```
forbidden-ip-range start-ip-address [ end-ip-address ]
```

缺省情况下，未配置地址池中不参与自动分配的 IP 地址段。

多次执行 **forbidden-ip-range** 命令，可以配置多个不参与自动分配的 IP 地址段。

(5) (可选) 在系统视图下配置全局不参与自动分配的 IP 地址。

a. 退回系统视图。

quit

b. 配置全局不参与自动分配的 IP 地址。

```
dhcp server forbidden-ip start-ip-address [ end-ip-address ]  
[ vpn-instance vpn-instance-name ]
```

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。

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。

2.4.5 配置静态地址绑定

1. 功能简介

某些客户端（如 Web 服务器等）需要固定的 IP 地址，通过以下几种方式可以实现为特定的客户端分配特定的 IP 地址：

- 将客户端的硬件地址与 IP 地址绑定：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。
- 将客户端 ID 与 IP 地址绑定：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。

2. 配置限制和指导

- 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获取到 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。

3. 配置步骤

(1) 进入系统视图。

system-view

(2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

(3) 配置静态地址绑定。

```
static-bind ip-address ip-address [ mask-length | mask mask ]  
{ client-identifier client-identifier | hardware-address  
hardware-address [ ethernet | token-ring ] } [ description  
description-text ]
```

缺省情况下，未配置静态地址绑定。

同一地址只能绑定给一个客户端。不允许通过重复执行 **static-bind ip-address** 命令的方式修改 IP 地址与客户端的绑定关系。只有删除了某个地址的绑定关系，才能将该地址与其他客户端绑定。

- (4) （可选）配置静态绑定 IP 地址的租约有效期限。

```
expired { day day [ hour hour [ minute minute [ second second ] ] ] |  
unlimited }
```

缺省情况下，IP 地址租约有效期限为 1 天。

2.4.6 配置 DHCP 客户端使用的网关地址

1. 功能简介

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 服务器可以为客户端指定网关的地址。

2. 配置限制和指导

- 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。目前，每个 DHCP 地址池视图下、每个从网段视图下最多可以配置 64 个网关地址。
- DHCP 地址池视图下执行 **gateway-list** 命令，配置的是为地址池中所有 DHCP 客户端分配的网关地址。如果用户需要为地址池下某个从网段的 DHCP 客户端分配其它的网关地址，可以在地址池的从网段视图下执行 **gateway-list** 命令。如果在地址池视图和从网段视图下都配置了网关地址，则优先将从网段视图下配置的网关地址分配给从网段的 DHCP 客户端。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置为 DHCP 客户端分配的网关地址。

```
gateway-list ip-address&<1-64>
```

缺省情况下，未配置为 DHCP 客户端分配的网关地址。

- (4) （可选）在从网段视图中配置为 DHCP 客户端分配的网关地址。

- a. 进入从网段视图。

```
network network-address [ mask-length | mask mask ] secondary
```

- b. 配置为 DHCP 客户端分配的网关地址。

```
gateway-list ip-address&<1-64>
```

缺省情况下，未配置为 DHCP 客户端分配的网关地址。

2.4.7 配置 DHCP 客户端使用的域名后缀

1. 功能简介

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从 DHCP 服务器获得的域名后缀进行解析。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“域名解析”。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置为 DHCP 客户端分配的域名后缀。

```
domain-name domain-name
```

缺省情况下，未配置为 DHCP 客户端分配的域名后缀。

2.4.8 配置 DHCP 客户端使用的 DNS 服务器地址

1. 功能简介

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指定 DNS（Domain Name System，域名系统）服务器地址。目前，每个 DHCP 地址池视图下最多可以配置 8 个 DNS 服务器地址。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置为 DHCP 客户端分配的 DNS 服务器地址。

```
dns-list ip-address&<1-8>
```

缺省情况下，未配置为 DHCP 客户端分配的 DNS 服务器地址。

2.4.9 配置 DHCP 客户端使用的 WINS 服务器地址和 NetBIOS 节点类型

1. 功能简介

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 服务器应该为客户端指定 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。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置为 DHCP 客户端分配的 WINS 服务器地址。

```
nbns-list ip-address <1-8>
```

缺省情况下，未配置为 DHCP 客户端分配的 WINS 服务器地址。

对于 b 类节点，为可选；其他情况下，为必选。每个 DHCP 地址池视图下最多可以配置 8 个 WINS 服务器地址。

- (4) 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 }
```

缺省情况下，未配置为 DHCP 客户端分配的 NetBIOS 节点类型。

2.4.10 配置 DHCP 客户端使用的 BIMS 服务器信息

1. 功能简介

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 服务器需要将 BIMS 服务器的 IP 地址、端口号以及加密的共享密钥等信息发给 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。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 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 } string
```

缺省情况下，未配置为 DHCP 客户端分配的 BIMS 服务器信息。

2.4.11 配置 DHCP 客户端使用的远程启动文件信息

1. 功能简介

服务器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或远程启动文件的 HTTP 形式 URL 等。

2. 配置 DHCP 客户端使用的 TFTP 服务器地址及启动文件名

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置 DHCP 客户端使用的 TFTP 服务器信息。请选择其中至少一项进行配置。

- 配置 DHCP 客户端使用的 TFTP 服务器地址。

```
tftp-server ip-address ip-address
```

缺省情况下,未配置 DHCP 客户端使用的 TFTP 服务器地址。

- 配置 DHCP 客户端使用的 TFTP 服务器名。

```
tftp-server domain-name domain-name
```

缺省情况下,未配置 DHCP 客户端使用的 TFTP 服务器名。

- (4) 配置 DHCP 客户端使用的启动文件名。

```
bootfile-name bootfile-name
```

缺省情况下,未配置 DHCP 客户端使用的启动文件名。

3. 配置 DHCP 客户端使用的远程启动文件的 HTTP 形式 URL

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置 DHCP 客户端使用的远程启动文件的 HTTP 形式 URL。

```
bootfile-name url
```

缺省情况下,未配置 DHCP 客户端使用的远程启动文件的 HTTP 形式 URL。

2.4.12 配置 DHCP 客户端使用的下一个提供服务的服务器 IP 地址

1. 功能简介

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。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 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
```

缺省情况下，未配置 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。

2.4.13 自定义 DHCP 选项

1. 自定义 DHCP 选项应用场景

本配置为 DHCP 服务器提供了灵活的选项配置方式，使得 DHCP 服务器可以为 DHCP 客户端提供更加丰富的选项内容。在以下情况下，可以使用本命令自定义 DHCP 选项：

- 随着 DHCP 的不断发展，新的 DHCP 选项会陆续出现。通过自定义 DHCP 选项，可以方便地添加新的 DHCP 选项。
- 有些选项的内容，RFC 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。通过自定义 DHCP 选项，可以为 DHCP 客户端提供厂商指定的信息。
- 设备上只提供了有限的选项配置命令（如 **gateway-list**、**dns-list** 命令），对于没有专门命令来配置的 DHCP 选项，可以通过 **option** 命令配置选项内容。例如，可以通过 **option 4 ip-address 1.1.1.1** 命令指定为 DHCP 客户端分配的时间服务器地址为 1.1.1.1。
- 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），可以通过自定义 DHCP 选项的方式进行扩展。

2. 常用 Option 选项

[表 2-1](#) 中列出了常用的 DHCP 选项名称、对应的配置命令和推荐的 Option 命令参数信息。

表2-1 常用 Option 选项信息

选项编号	选项名称	对应的配置命令	推荐的 option 命令参数
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

3. 配置限制和指导

- 自定义 DHCP 选项时，取值的获取比较复杂，配置错误可能会对 DHCP 的工作过程造成影响，请谨慎使用该功能。

- 用户可在 DHCP 地址池中自定义选项信息。
- 用户可在 DHCP 选项组中自定义选项信息,并在 DHCP 地址池中配置 DHCP 用户类和 DHCP 选项组关联,为 DHCP 客户端分配选项信息。

4. 自定义 DHCP 地址池选项

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 自定义 DHCP 地址池选项。

```
option code { ascii ascii-string | hex hex-string | ip-address ip-address&<1-8> }
```

缺省情况下,未自定义 DHCP 地址池选项。

DHCP 服务器在应答 DHCP 客户端报文时,如果 DHCP 选项组的选项编号和 DHCP 地址池选项编号相同且匹配用户类时,以 DHCP 选项组的选项为准。

5. 自定义 DHCP 选项组选项

- (1) 进入系统视图。

```
system-view
```

- (2) 创建 DHCP 选项组,并进入 DHCP 选项组视图。

```
dhcp option-group option-group-number
```

- (3) 自定义 DHCP 选项组选项。

```
option code { ascii ascii-string | hex hex-string | ip-address ip-address&<1-8> }
```

缺省情况下,未定义 DHCP 选项组的选项。

DHCP 服务器在应答客户端报文时,如果多个 DHCP 选项组的选项编号相同时,以最先匹配的 DHCP 用户类对应的 DHCP 选项组的选项为准。

- (4) 返回系统视图。

```
quit
```

- (5) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (6) 配置 DHCP 用户类与 DHCP 选项组的关联。

```
class class-name option-group option-group-number
```

缺省情况下,未配置指定 DHCP 用户类与 DHCP 选项组的关联。

2.4.14 为 DHCP 服务器上的地址池绑定 VPN 实例

1. 功能简介

当地址池绑定了 VPN 实例后,DHCP 服务器可以将网络划分成公网和 VPN 私网。未配置 VPN 属性的地址池被划分到公网,配置了 VPN 属性的地址池被划分到相应的 VPN 私网,这样,对于处于

公网或 VPN 私网中的客户端，服务器都能够选择合适的地址池来为客户端分配租约并且记录该客户端的状态信息。

DHCP 服务器可以通过如下方式判断 DHCP 客户端所属的 VPN 实例：

- 认证模块，用户接入时在 AAA 服务器处授权获得 VPN 实例信息；
- DHCP 服务器接收报文的接口绑定的 VPN 实例即为该客户端所属的 VPN 实例。

如果以上两种方式都可获取到 DHCP 客户端所属的 VPN 实例，则以认证模块为准。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 为 DHCP 服务器上的地址池绑定 VPN 实例。

```
vpn-instance vpn-instance-name
```

缺省情况下，DHCP 服务器上的地址池未绑定 VPN 实例。

2.4.15 配置 DHCP 用户类白名单功能

1. 功能简介

配置 DHCP 用户类白名单功能后：

- 当没有 DHCP 用户在线时，DHCP 服务器仅会处理属于用户类白名单用户发送的请求报文和非用户类白名单用户的续约请求报文。对于非用户类白名单用户发送的续约请求报文，DHCP 服务器将回复 DHCP-NAK 报文。
- 当有 DHCP 用户在线时，DHCP 服务器只有收到属于用户类白名单的用户发送的请求报文，才会进行处理。

2. 配置限制和指导

如果 DHCP 客户端请求的是静态绑定租约，则 DHCP 服务器不进行白名单检查直接进行处理。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 开启 DHCP 用户类白名单功能。

```
verify class
```

缺省情况下，DHCP 用户类白名单功能处于关闭状态。

- (4) 配置 DHCP 用户类白名单包括的用户类名。

```
valid class class-name<1-8>
```

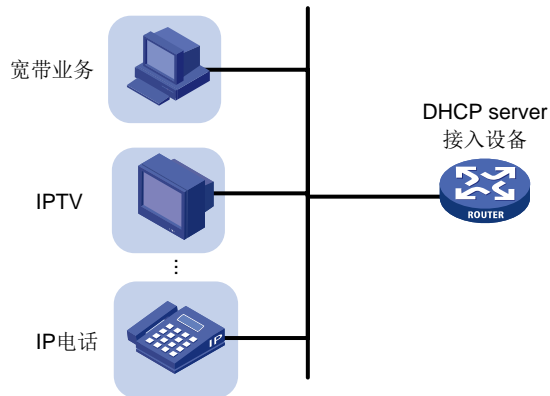
缺省情况下，未配置 DHCP 用户类白名单包括的用户类名。

2.4.16 配置 DHCP 服务器辅助网关信息

1. 功能简介

在某些接入组网类型中，如图 2-1 所示，接入设备上除了配置接入特性还需要配置 DHCP 服务器功能。由于接入设备需要接入多种业务的客户端（如 IPTV、IP 电话和宽带业务等），而不同业务的设备需要获取不同网段的 IP 地址，所以接入设备的下行口一般不能配置 IP 地址。此时可以通过在接入设备的 DHCP 地址池中配置辅助网关功能使不同类型的业务流量能够正常转发。本特性使用辅助网关的 IP 地址和 MAC 地址信息应答客户端的 ARP 请求，即可实现对不同类型的业务流量的引导。

图2-1 DHCP 服务器辅助网关组网图



2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置 DHCP 服务器辅助网关信息。

```
gateway-list ip-address&<1-64> export-route
```

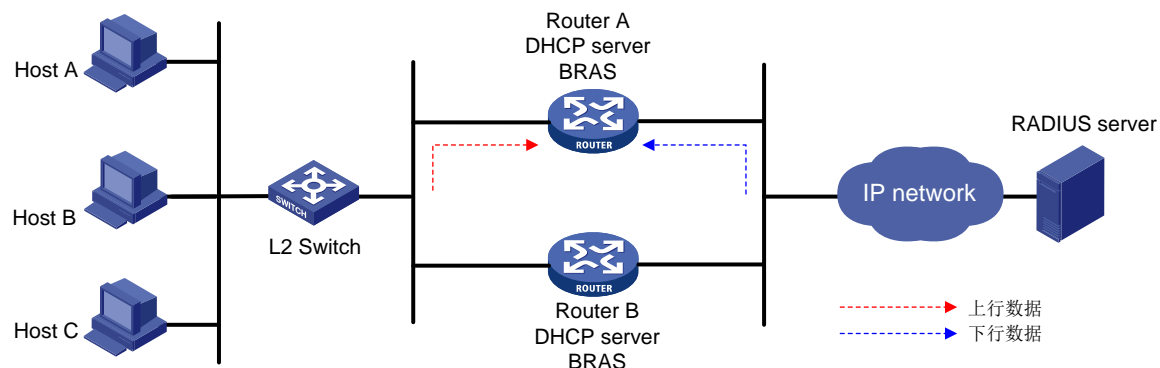
缺省情况下，未配置 DHCP 服务器辅助网关信息。

2.4.17 配置 DHCP 服务器辅助路由信息

1. 功能简介

在某些特定的业务模型（如 BRAS 组网）下，BRAS 设备需要实时监测网络流量，并将统计数据发送到 RADIUS 服务器。该统计数据为用户上线以来产生的所有上下行流量数据，而不能是设备在某个时间段内发生的上下行流量数据。由于 RADIUS 服务器刷新计数的方法是覆盖以前数据而不是进行累加，所以当一台设备的上下行流量分别从两台 BRAS 设备上通过时，在 RADIUS 服务器上记录的数据就会相互覆盖，这时 RADIUS 服务器得到的统计数据是不准确的。为了提高准确性，需保证一台设备的上下行流量经过同一台 BRAS 设备。通过配置辅助路由信息，并对外发布此网段路由，引导指定网段的下行数据流量来保证上下行流量从一台 BRAS 设备经过。

图2-2 DHCP 服务器辅助路由组网图



2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 配置 DHCP 服务器辅助路由信息。

```
network network-address [ mask-length | mask mask ] [ secondary ]  
export-route
```

缺省情况下，未配置 DHCP 服务器辅助路由信息。

2.5 配置接口引用地址池

1. 功能简介

创建地址池，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，则从引用的地址池中选择 IP 地址分配给客户端。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口引用地址池。

```
dhcp server apply ip-pool pool-name
```

缺省情况下，接口未引用地址池。

如果接口引用的地址池不存在，将导致无法动态分配地址。

2.6 配置DHCP策略动态分配地址和其他参数

1. 功能简介

创建 DHCP 策略，并在接口引用该策略后，该接口接收到 DHCP 请求报文时，则根据配置顺序逐个匹配 DHCP 策略中通过 **class ip-pool** 命令指定的 DHCP 用户类。匹配情况如下：

- 若匹配 DHCP 用户类成功，当该 DHCP 用户类关联的 DHCP 地址池中存在可供分配的地址信息时，则从该 DHCP 地址池中分配 IP 地址和其他参数；当该 DHCP 用户类关联的 DHCP 地址池中不存在可供分配的地址信息时，IP 地址和其他参数分配失败。
- 若匹配 DHCP 策略中的所有 DHCP 用户类失败，当配置了默认 DHCP 地址池时，则从该地址池中分配 IP 地址和其他参数；当未配置默认 DHCP 地址池或默认 DHCP 地址池中不存在可供分配的地址信息时，IP 地址和其他参数分配失败。

若接收 DHCP 请求报文的接口引用的 DHCP 策略不存在或匹配的 DHCP 用户类关联的 DHCP 地址池不存在时，IP 地址和其他参数分配失败。

2. 配置限制和指导

DHCP 策略需要在接口上引用才生效。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 创建 DHCP 策略，并进入 DHCP 策略视图。

```
dhcp policy policy-name
```

- (3) 指定 DHCP 用户类关联的 DHCP 地址池。

```
class class-name ip-pool pool-name
```

缺省情况下，未指定 DHCP 用户类关联的 DHCP 地址池。

- (4) 指定默认 DHCP 地址池。

```
default ip-pool pool-name
```

缺省情况下，未指定默认 DHCP 地址池。

- (5) 退回系统视图。

```
quit
```

- (6) 进入接口视图。

```
interface interface-type interface-number
```

- (7) 指定接口引用的 DHCP 策略。

```
dhcp apply-policy policy-name
```

缺省情况下，接口未引用 DHCP 策略。

2.7 开启DHCP服务

1. 配置限制和指导

只有开启 DHCP 服务后，其它相关的 DHCP 服务器配置才能生效。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DHCP 服务。

```
dhcp enable
```

缺省情况下，DHCP 服务处于关闭状态。

2.8 配置接口工作在DHCP服务器模式

1. 功能简介

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等参数。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口工作在 DHCP 服务器模式。

```
dhcp select server
```

缺省情况下，接口工作在 DHCP 服务器模式。

2.9 配置IP地址冲突检测功能

1. 功能简介

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检测。

DHCP 服务器的地址探测是通过 ping 功能实现的，通过检测是否能在指定时间内得到 ping 响应来判断是否存在地址冲突。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 ICMP 回显响应报文，则认为存在地址冲突。DHCP 服务器从地址池中选择新的 IP 地址，并重复上述操作。如果在指定时间内未收到 ICMP 回显响应报文，则继续发送 ICMP 回显请求报文，直到发送的 ICMP 回显显示报文数目达到最大值。如果仍然未收到 ICMP 回显响应报文，则将地址分配给客户端，从而确保客户端获得的 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。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) （可选）配置 DHCP 服务器发送 ICMP 回显请求报文的最大数目。

```
dhcp server ping packets number
```

缺省情况下，DHCP 服务器发送 ICMP 回显请求报文的最大数目为 1。

0 表示 DHCP 服务器将 IP 地址分配给 DHCP 客户端之前，不会通过 ping 操作探测该地址是否冲突。

- (3) （可选）配置 DHCP 服务器等待 ICMP 回显响应报文的超时时间。

```
dhcp server ping timeout milliseconds
```

缺省情况下，DHCP 服务器等待 ICMP 回显响应报文的超时时间为 500 毫秒。

0 表示 DHCP 服务器将 IP 地址分配给 DHCP 客户端之前，不会通过 ping 操作探测该地址是否冲突。

2.10 配置 Option 82 的处理方式

1. 功能简介

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。

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 中继支持 Option 82 功能的相关配置请参见“[3.9 配置 DHCP 中继支持 Option 82 功能](#)”。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 DHCP 服务器处理 Option 82。

```
dhcp server relay information enable
```

缺省情况下，DHCP 服务器处理 Option 82。

2.11 配置 DHCP 服务器兼容性

当 DHCP 客户端的行为不符合 RFC 协议规定时，为了与之兼容，需要配置 DHCP 服务器兼容性功能。

2.11.1 配置 DHCP 服务器始终以广播方式回复请求报文

1. 功能简介

一般情况下，只有 DHCP 请求报文的广播标志位为 1 的时候，DHCP 服务器才会以广播的方式发送应答报文。如果 DHCP 客户端发送的请求报文中广播标志位为 0，且该客户端不支持接收单播的应答报文，则可以配置 DHCP 服务器忽略请求报文的广播标志位，始终以广播方式发送应答报文。

当已经存在 IP 地址的客户端发出请求报文（即报文中 ciaddr 字段不为 0）时，无论是否开启 DHCP 服务器的广播回应报文功能，DHCP 服务器都会以单播形式将回应报文发送给 DHCP 客户端（即目的地址为 ciaddr）。

当请求报文通过 DHCP 中继转发到 DHCP 服务器（即报文中 giaddr 字段不为 0）时，无论是否开启 DHCP 服务器的广播回应报文功能，DHCP 服务器都会以单播形式将回应报文发送给 DHCP 中继（即目的地址为 giaddr）。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DHCP 服务器的广播回应报文功能。

```
dhcp server always-broadcast
```

缺省情况下，DHCP 服务器的广播回应报文功能处于关闭状态。DHCP 服务器根据请求报文中的广播标志位来决定以广播还是单播的形式发送应答报文。

2.11.2 配置 DHCP 服务器忽略 BOOTP 请求报文

1. 功能简介

BOOTP 客户端申请到的地址租约是无限期的。在某些组网环境中，可能不希望出现无限期的地址租约。此时，可以通过配置 DHCP 服务器忽略 BOOTP 请求报文，避免分配无限期的地址租约。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 DHCP 服务器忽略 BOOTP 请求报文。

```
dhcp server bootp ignore
```

缺省情况下，DHCP 服务器不会忽略 BOOTP 请求报文。

2.11.3 配置 DHCP 服务器以 RFC 1048 规定的格式发送 BOOTP 应答报文

1. 功能简介

有些 BOOTP 客户端发送的请求报文中，vend 字段的格式不符合 RFC 1048 的要求。对于这种报文，DHCP 服务器的缺省处理方法是不解析 vend 字段内容，将报文中 vend 字段的内容拷贝到回复报文中的 vend 字段回应给 BOOTP 客户端。

开启 DHCP 服务器的回应 RFC 1048 格式报文功能后，对于这种格式不符合 RFC 1048 要求的报文，DHCP 服务器会将需要回应的选项以符合 RFC 1048 要求的格式，封装到回复报文的 vend 字段，并回应给 BOOTP 客户端。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DHCP 服务器回应 RFC 1048 格式报文功能。

```
dhcp server bootp reply-rfc-1048
```

缺省情况下，DHCP 服务器回应 RFC 1048 格式报文功能处于关闭状态。

2.11.4 配置 DHCP 服务器发送 DHCP 应答报文不携带 Option 60 选项

1. 功能简介

如果网络中存在不支持解析 Option 60 的 DHCP 客户端，DHCP 服务器需要配置 DHCP 服务器发送 DHCP 应答报文时不携带 Option 60 选项功能。配置该功能后，DHCP 服务器无论收到的 DHCP 报文中是否携带 Option 60 选项，也无论 DHCP 地址池中是否已经配置了 Option 60 选项内容，DHCP 服务器应答的 DHCP 报文中都不携带 Option 60 选项。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 DHCP 服务器发送 DHCP 应答报文不携带 Option 60 选项。

```
dhcp server reply-exclude-option60
```

缺省情况下，DHCP 服务器发送 DHCP 应答报文时可以携带 Option 60 选项。

2.12 配置 DHCP 服务器发送 DHCP 报文的 DSCP 优先级

1. 功能简介

DSCP 优先级用来体现报文自身的优先等级，决定报文传输的优先程度。通过本配置可以指定 DHCP 服务器发送的 DHCP 报文的 DSCP 优先级。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 DHCP 服务器发送 DHCP 报文的 DSCP 优先级。

```
dhcp dscp dscp-value
```

缺省情况下，DHCP 服务器发送 DHCP 报文的 DSCP 优先级为 56。

2.13 配置 DHCP 服务器租约固化功能

1. 功能简介

DHCP 服务器重启后，设备上记录的租约信息将丢失，会影响 DHCP 服务器的正常业务。

DHCP 服务器租约固化功能将 DHCP 服务器的在用地址租约和冲突表项保存到指定的文件中，DHCP 服务器设备重启后，自动根据该文件恢复 DHCP 服务器的租约信息，从而保证 DHCP 服务器的租约信息不会丢失。

当 DHCP 服务器设备重启后，自动根据该文件恢复 DHCP 服务器的租约信息，租约恢复的过程中，DHCP 服务器不能提供 DHCP 业务。所以当恢复过程出现问题导致恢复过程无法结束时，用户可配置 **dhcp server database update stop** 命令终止当前的 DHCP 服务器表项恢复操作，以便 DHCP 服务器能及时提供 DHCP 服务。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 指定存储 DHCP 服务器表项的文件名称。

```
dhcp server database filename { filename | url url [ username username [ password { cipher | simple } string ] ] }
```

缺省情况下，未指定存储 DHCP 服务器表项的文件名称。

执行本命令后，会立即触发一次表项备份。

- (3) （可选）将当前的 DHCP 服务器表项保存到用户指定的文件中。

```
dhcp server database update now
```

本命令只用来触发一次 DHCP 服务器表项的备份。

- (4) （可选）配置刷新 DHCP 服务器表项存储文件的延迟时间。

```
dhcp server database update interval interval
```

缺省情况下，若 DHCP 服务器表项不变化，则不刷新存储文件；若 DHCP 服务器表项发生变化，默认在 300 秒之后刷新存储文件。

- (5) （可选）终止当前的 DHCP 服务器表项恢复操作。

```
dhcp server database update stop
```

本命令只用来触发一次终止 DHCP 服务器表项信息的恢复。

2.14 开启DHCP服务器的用户下线探测功能

1. 功能简介

DHCP 服务器的用户下线探测功能以 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下线。

如果在接口上开启了 DHCP 服务器的用户下线探测功能，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。

2. 配置限制和指导

手工删除 ARP 表项，不会触发 DHCP 服务器删除对应用户的地址绑定信息。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 开启 DHCP 服务器的用户下线探测功能。

```
dhcp client-detect
```

缺省情况下，DHCP 服务器的用户下线探测功能处于关闭状态。

2.15 配置DHCP告警功能

1. 功能简介

为了避免地址池地址耗尽，导致用户无法上线，用户可以设置地址池使用率的告警阈值，当地址池中地址使用率超过阈值时，系统会生成告警信息，并将告警信息发送到设备的 SNMP 模块，通过

设置 SNMP 中告警信息的发送参数，来决定告警信息输出的相关属性。有关告警信息的详细介绍，请参见“网络管理和监控配置指导”中的“SNMP”。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCP 地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) （可选）设置地址池使用率告警门限阈值。

```
ip-in-use threshold threshold-value
```

缺省情况下，地址池使用率告警门限阈值为 100%。

2.16 开启DHCP服务器日志信息功能

1. 功能简介

DHCP 服务器日志可以方便管理员定位问题和解决问题。设备生成 DHCP 日志信息会交给信息中心模块处理，信息中心模块的配置将决定日志信息的发送规则和发送方向。关于信息中心的详细描述请参见“网络管理和监控配置指导”中的“信息中心”。

2. 配置限制和指导

比如大量 DHCP 客户端发生上下线操作时，DHCP 服务器会输出大量日志信息，这可能会降低设备性能，影响 DHCP 服务器分配 IP 地址的速度。为了避免该情况的发生，用户可以关闭 DHCP 服务器日志信息功能，使得 DHCP 服务器不再输出日志信息。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DHCP 服务器日志信息功能。

```
dhcp log enable
```

缺省情况下，DHCP 服务器日志信息功能处于关闭状态。

2.17 DHCP服务器显示和维护



提示

DHCP 服务器重启或使用 **reset dhcp server ip-in-use** 命令清除租约后，DHCP 服务器上不存在租约信息。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。

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2-2 DHCP 服务器显示和维护

操作	命令
显示DHCP的地址冲突信息	<code>display dhcp server conflict [ip ip-address] [vpn-instance vpn-instance-name]</code>
显示DHCP服务器的表项备份信息	<code>display dhcp server database</code>
显示租约过期的地址绑定信息	<code>display dhcp server expired [[ip ip-address] [vpn-instance vpn-instance-name] pool pool-name]</code>
显示DHCP地址池的空闲地址信息	<code>display dhcp server free-ip [pool pool-name vpn-instance vpn-instance-name]</code>
显示DHCP地址绑定信息	<code>display dhcp server ip-in-use [[ip ip-address] [vpn-instance vpn-instance-name] pool pool-name]</code>
显示DHCP地址池的信息	<code>display dhcp server pool [pool-name vpn-instance vpn-instance-name]</code>
显示DHCP服务器的统计信息	<code>display dhcp server statistics [pool pool-name vpn-instance vpn-instance-name]</code>
清除DHCP的地址冲突信息	<code>reset dhcp server conflict [ip ip-address] [vpn-instance vpn-instance-name]</code>
清除租约过期的地址绑定信息	<code>reset dhcp server expired [[ip ip-address] [vpn-instance vpn-instance-name] pool pool-name]</code>
清除DHCP的正式绑定和临时绑定信息	<code>reset dhcp server ip-in-use [[ip ip-address] [vpn-instance vpn-instance-name] pool pool-name]</code>
清除DHCP服务器的统计信息	<code>reset dhcp server statistics [vpn-instance vpn-instance-name]</code>

2.18 DHCP服务器常见故障处理

2.18.1 DHCP 客户端获取到冲突的 IP 地址

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. 故障处理

- (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 环境下运行 `cmd` 进入 DOS 环境，使用 `ipconfig /release` 命令释放 IP 地址，之后使用 `ipconfig /renew` 重新获取 IP 地址。

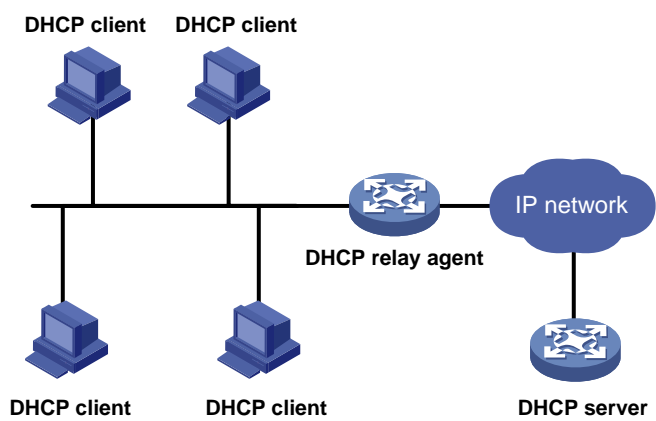
3 DHCP 中继

3.1 DHCP中继简介

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 DHCP 服务器通信，获取 IP 地址及其他配置信息。

[图 3-1](#) 是 DHCP 中继的典型应用示意图。

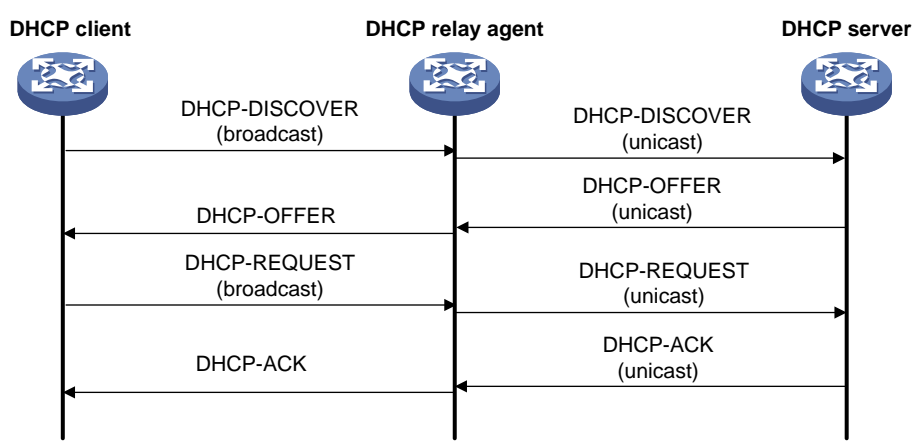
图3-1 DHCP 中继的典型组网应用



3.1.2 DHCP 中继的基本原理

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，并根据配置将报文中广播转发给指定的 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，实现根据 Option 82 为客户端分配特定范围的地址、对客户端进行安全和计费控制。Option 82 的详细介绍请参见“[1.6.2 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	DHCP 中继对报文的处理
收到的报文中带有 Option 82	Drop	丢弃报文
	Keep	保持报文中的 Option 82 不变并进行转发
	Replace	根据 DHCP 中继上配置的填充模式、内容、格式等填充 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	-	根据 DHCP 中继上配置的填充模式、内容、格式等填充 Option 82，添加到报文中并进行转发

3.2 DHCP 中继配置任务简介

DHCP 中继配置任务如下：

- (1) [开启 DHCP 服务](#)
- (2) [配置接口工作在 DHCP 中继模式](#)
- (3) [指定 DHCP 服务器的地址](#)
- (4) [（可选）指定 DHCP 客户端对应的 DHCP 中继地址池](#)
- (5) [（可选）配置 DHCP 中继的安全功能](#)
- (6) [（可选）配置通过 DHCP 中继释放客户端的 IP 地址](#)
- (7) [（可选）配置 DHCP 中继支持 Option 82 功能](#)
- (8) [（可选）配置 DHCP 中继发送 DHCP 报文的 DSCP 优先级](#)
- (9) [（可选）配置 DHCP 中继在 DHCP 报文中填充的中继地址](#)
- (10) [（可选）指定 DHCP 中继向 DHCP 服务器转发报文的源地址](#)

(11) [\(可选\)配置 DHCP 中继通过 Option82 信息转发 DHCP 应答报文](#)

3.3 开启DHCP服务

1. 配置限制和指导

只有开启 DHCP 服务后，其它相关的 DHCP 中继配置才能生效。

2. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 开启 DHCP 服务。

```
dhcp enable
```

缺省情况下，DHCP 服务处于关闭状态。

3.4 配置接口工作在DHCP中继模式

1. 功能简介

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。

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 IP 地址所在网段（网络号和掩码）匹配的地址池，否则会导致 DHCP 客户端无法获得正确的 IP 地址。

2. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 进入接口视图。

```
interface interface-type interface-number
```

(3) 配置接口工作在 DHCP 中继模式。

```
dhcp select relay
```

缺省情况下，开启 DHCP 服务后，接口工作在 DHCP 服务器模式。

3.5 指定DHCP服务器的地址

3.5.1 指定 DHCP 中继对应的 DHCP 服务器地址

1. 功能简介

为了提高可靠性，可以在一个网络中设置多个 DHCP 服务器。DHCP 中继上配置多个 DHCP 服务器后，DHCP 中继会将客户端发来的 DHCP 报文转发给所有的服务器。

2. 配置限制和指导

指定的 DHCP 服务器的 IP 地址不能与 DHCP 中继的接口 IP 地址在同一网段。否则，可能导致客户端无法获得 IP 地址。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 指定 DHCP 中继对应的 DHCP 服务器地址。

```
dhcp relay server-address ip-address
```

缺省情况下，未指定 DHCP 服务器的地址。

通过多次执行 **dhcp relay server-address** 命令可以指定多个 DHCP 服务器，一个接口下最多可以指定 8 个 DHCP 服务器。

3.5.2 指定中继地址池对应的 DHCP 服务器地址

1. 功能简介

对于某些特定的用户接入方式，基于用户接入位置信息的不同，网络中存在大量不同类型的用户。为了使相同类型的用户可以从指定的 DHCP 服务器申请 IP 地址等网络参数，接入模块根据用户注册信息，使不同的用户选择不同的 DHCP 中继地址池，并从中继地址池下配置的 DHCP 服务器获取 IP 地址等网络参数。

为了提高可靠性，一个 DHCP 中继地址池下配置多个 DHCP 服务器地址，当 DHCP 客户端匹配该中继地址池后，DHCP 中继会将 DHCP 客户端发来的 DHCP 报文转发给该地址池对应所有的 DHCP 服务器。

一台 DHCP 中继的一个接口下可能连接不同类型的用户，当 DHCP 中继转发 DHCP 客户端请求报文给 DHCP 服务器时，不能再以中继接口的 IP 地址作为选择地址池的依据。为了解决这个问题，需要使用 **gateway-list** 命令指定某个类型用户所在的网段，并将该地址添加到转发给 DHCP 服务器的报文字段中，为 DHCP 服务器选择地址池提供依据。

2. 配置限制和指导

配置本功能需要注意：

- 当 PPPoE 用户下线时，DHCP 中继需要查询中继用户地址表项，若存在对应表项，则会向 DHCP 服务器发送 Release 报文，通知 DHCP 服务器释放该地址租约。这就需要在 DHCP 中继上使用 **dhcp relay client-information record** 命令开启 DHCP 中继用户地址表项记录功能。
- 和 PPPoE 配合使用时，如果设备的地址池中配置了 **remote-server** 命令，则可以认定该设备一定是 DHCP 中继设备，所以不需要在接口视图下执行 **dhcp select relay** 命令。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 创建 DHCP 中继地址池，并进入中继地址池视图。

```
dhcp server ip-pool pool-name
```

- (3) 指定匹配该地址池的 DHCPv4 客户端所在的网段地址。

```
gateway-list ip-address&<1-64> [ export-route ]
```

缺省情况下，未指定匹配该地址池的 DHCP 客户端所在的网段地址。

- (4) 指定中继地址池对应的 DHCP 服务器地址。

remote-server *ip-address*&<1-8>

缺省情况下，未指定中继地址池对应的 DHCP 服务器的地址。

通过执行 **remote-server** 命令一次最多可以指定 8 个 DHCP 服务器的地址信息。

3.6 指定DHCP客户端对应的DHCP中继地址池

1. 功能简介

在 DHCP 中继上可能存在多个 DHCP 中继地址池，可以在接口下直接指定 DHCP 客户端对应的地址池。如果希望更进一步根据 DHCP 客户端报文的 Option 信息来选择对应的 DHCP 中继地址池，则可以指定 Option 参数。

2. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 进入 DHCP 中继地址池视图。

dhcp server ip-pool *pool-name*

- (3) 指定中继地址池对应的 DHCP 服务器地址。

remote-server *ip-address*&<1-8>

缺省情况下，未指定中继地址池的 DHCP 服务器的地址。

- (4) 指定匹配该地址池的 DHCP 客户端所在的网段的地址。

gateway-list *ip-address*&<1-64>

缺省情况下，未指定匹配该地址池的 DHCP 客户端所在的网段地址。

3.7 配置DHCP中继的安全功能

3.7.1 配置 DHCP 中继用户地址表项记录功能

1. 功能简介

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 中继用户地址表项记录功能。

开启该功能后，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 地址与硬件地址的绑定关系，生成 DHCP 中继的用户地址表项。

本功能与其他 IP 地址安全功能（如 ARP 地址检查、授权 ARP）配合，可以实现只允许匹配用户地址表项中绑定关系的报文通过 DHCP 中继。从而，保证非法主机不能通过 DHCP 中继与外部网络通信。

2. 配置限制和指导

同异步串口作为 DHCP 客户端申请 IP 地址时，DHCP 中继不会记录该客户端对应的用户地址表项。

3. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 开启 DHCP 中继的用户地址表项记录功能。

dhcp relay client-information record

缺省情况下，DHCP 中继用户地址表项记录功能处于关闭状态。

3.7.2 配置 DHCP 中继动态用户地址表项定时刷新功能

1. 功能简介

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 中继每隔指定时间采用客户端获取到的 IP 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未接收到 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 中继会删除动态用户地址表中对应的表项。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 中继不会删除该 IP 地址对应的表项。

2. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 开启 DHCP 中继动态用户地址表项定时刷新功能。

dhcp relay client-information refresh enable

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。

- (3) （可选）配置 DHCP 中继动态用户地址表项的定时刷新周期。

dhcp relay client-information refresh { auto | interval interval }

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。

3.7.3 配置防止 DHCP 饿死攻击

1. 功能简介

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源 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,需要开启 DHCP 中继的 MAC 地址检查功能。开启该功能后,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。

因为 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 中继设备上开启 MAC 地址检查功能。

设备支持配置 DHCP 中继的 MAC 地址检查表项老化时间,当老化时间到达以后,该表项信息会被老化掉,DHCP 中继收到该 MAC 地址对应的 DHCP 请求报文后重新进行合法性检查。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 DHCP 中继的 MAC 地址检查表项的老化时间。

```
dhcp relay check mac-address aging-time time
```

缺省情况下,DHCP 中继的 MAC 地址检查表项的老化时间为 30 秒。

如果未通过 `dhcp relay check mac-address` 命令开启 DHCP 中继的 MAC 地址检查功能,则本命令的配置不会生效。

- (3) 进入接口视图。

```
interface interface-type interface-number
```

- (4) 开启 DHCP 中继的 MAC 地址检查功能。

```
dhcp relay check mac-address
```

缺省情况下,DHCP 中继的 MAC 地址检查功能处于关闭状态。

3.7.4 配置 DHCP 中继支持代理功能

1. 功能简介

设备可以通过配置 DHCP 中继支持代理功能,来防止非法用户攻击 DHCP 服务器。

开启该功能后,DHCP 中继收到 DHCP 服务器的应答报文,会把报文中的 DHCP 服务器地址修改为中继的接口地址,并转发给 DHCP 客户端。当 DHCP 客户端通过 DHCP 中继从 DHCP 服务器获取到 IP 地址等网络参数后,DHCP 客户端会把 DHCP 中继当做自己的服务器,来进行后续的 DHCP 功能的报文交互。从而达到了把真正的 DHCP 服务器和 DHCP 客户端隔离开,保护 DHCP 服务器的目的。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置 DHCP 中继支持代理功能。

```
dhcp select relay proxy
```

缺省情况下,开启 DHCP 服务后,接口工作在 DHCP 服务器模式。

3.7.5 配置 DHCP 中继的用户下线探测功能

1. 功能简介

DHCP 中继的用户下线探测功能以 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。

如果在接口上配置了 DHCP 中继的用户下线检测功能，则当 ARP 表项老化时，DHCP 中继认为该表项对应的用户已经下线，删除对应的用户地址表项，并通过发送 Release 报文通知 DHCP 服务器删除下线用户的 IP 地址租约。

2. 配置限制和指导

手工删除 ARP 表项，不会触发 DHCP 中继删除对应的用户地址表项。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DHCP 中继的用户地址表项记录功能。

```
dhcp relay client-information record
```

缺省情况下，DHCP 中继用户地址表项记录功能处于关闭状态。

用户需要开启 DHCP 中继用户地址表项记录功能，否则用户下线探测功能无法完全生效。

- (3) 进入接口视图。

```
interface interface-type interface-number
```

- (4) 配置接口工作在 DHCP 中继模式。

```
dhcp select relay
```

缺省情况下，开启 DHCP 服务后，接口工作在 DHCP 服务器模式。

- (5) 开启 DHCP 中继的用户下线探测功能。

```
dhcp client-detect
```

缺省情况下，DHCP 中继的用户下线探测功能处于关闭状态。

3.8 配置通过 DHCP 中继释放客户端的 IP 地址

1. 功能简介

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。

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。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 向 DHCP 服务器请求释放客户端申请到的 IP 地址。

```
dhcp relay release ip ip-address [ vpn-instance vpn-instance-name ]
```

3.9 配置DHCP中继支持Option 82功能

1. 配置限制和指导

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“[2.10 配置 Option 82 的处理方式](#)”。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 开启 DHCP 中继支持 Option 82 功能。

```
dhcp relay information enable
```

缺省情况下，DHCP 中继支持 Option 82 功能处于关闭状态。

- (4) （可选）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 }
```

缺省情况下，处理策略为 **replace**。

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 的填充模式和填充格式；处理策略为 **keep** 或 **drop** 时，不需要配置 Option 82 的填充模式和填充格式。

- (5) （可选）配置 Circuit ID 子选项的填充模式和填充格式。

```
dhcp relay information circuit-id { bas [ sub-interface-vlan ] | string  
circuit-id | { normal | verbose [ node-identifier { mac | sysname |  
user-defined node-identifier } ] [ interface ] } [ sub-interface-vlan ]  
[ format { ascii | hex } ] }
```

缺省情况下，Circuit ID 子选项的填充模式为 Normal，填充格式为 hex。

如果以设备的系统名称（**sysname**）作为节点标识填充 DHCP 报文的 Option 82，则系统名称中不能包含空格；否则，DHCP 中继添加或替换 Option 82 失败。

- (6) （可选）配置 Remote ID 子项的填充模式和填充格式。

```
dhcp relay information remote-id { normal [ format { ascii | hex } ] |  
string remote-id | sysname }
```

缺省情况下，Remote ID 子选项的填充模式为 Normal；填充格式为 hex。

3.10 配置DHCP中继发送DHCP报文的DSCP优先级

1. 功能简介

DSCP 优先级用来体现报文自身的优先等级，决定报文传输的优先程度。通过本配置可以指定 DHCP 中继发送的 DHCP 报文的 DSCP 优先级。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 DHCP 中继发送 DHCP 报文的 DSCP 优先级。

```
dhcp dscp dscp-value
```

缺省情况下，DHCP 中继发送的 DHCP 报文的 DSCP 优先级为 56。

3.11 配置DHCP中继在DHCP报文中填充的中继地址

3.11.1 手工指定在 DHCP 报文中填充的中继地址

1. 功能简介

当未开启该功能时，DHCP 中继收到 DHCP 客户端的请求报文后，只能将接口的主 IP 地址添加到报文中，然后转发给 DHCP 服务器。对于某些特定需求，DHCP 中继需要添加指定的地址到报文中，这时就需要配置此功能。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 手工指定在 DHCP 报文中填充的中继地址。

```
dhcp relay gateway ip-address
```

缺省情况下，DHCP 中继填充的中继地址是接口下的主 IP 地址。

3.11.2 通过 smart-relay 功能指定 DHCP 报文中填充的中继地址

1. 功能简介

当 DHCP 中继收到 DHCP 客户端发来的请求报文时，会使用中继接口的主 IP 地址填充请求报文的 giaddr 字段，然后转发给 DHCP 服务器，DHCP 服务器根据 giaddr 字段中的地址选择合适的地址池为客户端分配 IP 地址。当 DHCP 服务器中该网段地址分配完毕后，不管 DHCP 服务器上是否存在其他网段的地址，都不会再为该 DHCP 中继下的其他 DHCP 客户端分配 IP 地址。

DHCP 中继通过 smart-relay 解决上述问题，开启该功能后，DHCP 中继可以使用除中继接口主地址外的其他 IP 地址来填充 giaddr 字段，从而使 DHCP 客户端可以获取到其他网段的 IP 地址。

DHCP 中继转发 3 次 DHCP-DISCOVER 报文后，若还未收到 DHCP 服务器的应答报文，DHCP 中继将使用下一个可用 IP 地址来填充 giaddr 字段。DHCP 中继使用所有配置的 IP 地址填充 giaddr 字段之后，将重新选择第一个配置的 IP 地址进入下一个循环。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DHCP 中继支持 smart-relay 功能。

```
dhcp smart-relay enable
```

缺省情况下，DHCP 中继支持 smart-relay 功能处于关闭状态。

3.12 指定DHCP中继向DHCP服务器转发报文的源地址

1. 功能简介

在某些组网中，多个 DHCP 中继接口 IP 地址相同或者中继接口 IP 到服务器没有可达路由，用户需要配置本命令指定一个 IP 地址或选择中继设备上的另一个接口（一般选择的是 Loopback 口）的 IP 地址填充到发送到 DHCP 服务器的 DHCP 请求报文中的源地址字段和 Giaddr 中。

当多个 DHCP 中继接口 IP 地址相同时，导致 DHCP 中继转发 DHCP 应答报文时无法根据目的 IP 地址找到唯一的出接口。配置本功能时需要先开启 DHCP 中继支持 Option 82 功能，DHCP 中继收到 DHCP 请求报文时在 Option 82 中的子选项 sub-option5 填充正确的子网网段，服务器可以根据中继填充的 sub-option5 来分配地址，之后 DHCP 中继处理 DHCP 应答报文时通过 MAC 地址表中的接口信息转发 DHCP 报文。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 指定 DHCP 中继向 DHCP 服务器转发报文的源地址。

```
dhcp relay source-address { ip-address | gateway | relay-interface }
```

缺省情况下，当 DHCP 中继连接服务器的接口和 DHCP 服务器属于同一个 VPN 时，DHCP 中继向 DHCP 服务器转发报文出接口的地址作为报文源地址；当 DHCP 中继连接服务器的接口和 DHCP 服务器不属于同一个 VPN 时，DHCP 中继会选择一个中继上和 DHCP 服务器同 VPN 的最小地址作为报文源地址。

3.13 配置DHCP中继通过Option82信息转发DHCP应答报文

1. 功能简介

对于某些组网环境，DHCP 中继需要通过 DHCP 服务器的 DHCP 应答报文中携带的 Option 82 信息（VLAN ID）把 DHCP 应答报文转发给 DHCP 客户端。如在 IPRAN 组网主备 PW 网关收敛 N:1 方式中，主备网关设备上配置 L3VE 口作为中继接口，配置多个 L2VE 子接口（L2VE 子接口与 PW 是一一对应的）接收报文。只有主设备会收到 DHCP 请求报文，但是主备设备都可能收到 DHCP 应答报文，主设备可通过记录的用户信息将报文从对应的 PW 发送，但备设备没有记录用户信息，所以备设备会向所有的 PW 发送 DHCP 应答报文。

为了解决这个问题，需要配置：

- 在主设备上配置 `dhcp relay information enable` 和 `dhcp relay information circuit-id`（指定 `sub-interface-vlan` 关键字）命令。当主设备收到 DHCP 请求报文时，会添加 Option 82 选项，并记录 L2VE 子接口所在的 VLAN 编号。
- 在备设备上配置 `dhcp relay information enable`、`dhcp relay information circuit-id`（指定 `sub-interface-vlan` 关键字）和 `dhcp relay forward reply`

by-option82 命令。这样，当设备收到 DHCP 应答报文时，可根据报文中 Option 82 选项记录的 L2VE 子接口所在的 VLAN 编号直接向对应的 PW 转发 DHCP 应答报文。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 开启 DHCP 中继支持 Option 82 功能。

```
dhcp relay information enable
```

缺省情况下，DHCP 中继支持 Option 82 功能处于关闭状态。

- (4) 配置 Circuit ID 子选项的填充模式和填充格式。

```
dhcp relay information circuit-id { bas [ sub-interface-vlan ] | string circuit-id | { normal | verbose [ node-identifier { mac | sysname | user-defined node-identifier } ] [ interface ] } [ sub-interface-vlan ] [ format { ascii | hex } ] }
```

缺省情况下，Circuit ID 子选项的填充模式为 Normal，填充格式为 hex。

如果以设备的系统名称（**sysname**）作为节点标识填充 DHCP 报文的 Option 82，则系统名称中不能包含空格；否则，DHCP 中继添加或替换 Option 82 失败。

- (5) 配置 DHCP 中继通过 Option82 信息转发 DHCP 应答报文。

```
dhcp relay forward reply by-option82
```

缺省情况下，DHCP 中继不通过 Option82 信息转发 DHCP 应答报文。

3.14 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3-2 DHCP 中继显示和维护

操作	命令
显示DHCP中继的MAC地址检查表项	display dhcp relay check mac-address
显示DHCP中继的用户地址表项信息	display dhcp relay client-information [interface <i>interface-type</i> <i>interface-number</i> ip <i>ip-address</i> [vpn-instance <i>vpn-instance-name</i>]]
显示DHCP中继上的Option 82配置信息	display dhcp relay information [interface <i>interface-type interface-number</i>]
显示接口上指定的DHCP服务器地址信息	display dhcp relay server-address [interface <i>interface-type</i> <i>interface-number</i>]
显示DHCP中继的相关报文统计信息	display dhcp relay statistics [interface <i>interface-type interface-number</i>]

操作	命令
清除DHCP中继的用户地址表项信息	reset dhcp relay client-information [interface <i>interface-type</i> <i>interface-number</i> ip <i>ip-address</i> [vpn-instance <i>vpn-instance-name</i>]]
清除DHCP中继的相关报文统计信息	reset dhcp relay statistics [interface <i>interface-type</i> <i>interface-number</i>]

3.15 DHCP中继常见故障处理

3.15.1 DHCP 客户端无法通过 DHCP 中继获取配置信息

1. 故障现象

DHCP 客户端无法通过 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. 故障处理

- 检查 DHCP 服务器和 DHCP 中继是否开启了 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指定了正确的 DHCP 服务器地址。

4 DHCP 客户端

4.1 DHCP客户端简介

为了方便用户配置和集中管理，可以指定设备的接口作为 DHCP 客户端，使用 DHCP 协议从 DHCP 服务器动态获得 IP 地址等参数。

4.2 DHCP客户端配置限制和指导

DHCP 客户端中对于接口的相关配置，目前只能在三层以太网接口（包括子接口）、VLAN 接口和三层聚合接口上进行。

4.3 DHCP客户端配置任务简介

DHCP 客户端配置任务如下：

- (1) [配置接口通过 DHCP 协议获取 IP 地址](#)
- (2) [配置接口使用的 DHCP 客户端 ID](#)

DHCP 客户端使用客户端 ID 从 DHCP 服务器获取特定地址时配置。

- (3) （可选）[开启地址冲突检查功能](#)
- (4) （可选）[配置 DHCP 客户端发送 DHCP 报文的 DSCP 优先级](#)

4.4 配置接口通过DHCP协议获取IP地址

1. 配置限制和指导

配置接口通过 DHCP 协议获取 IP 地址，需要注意：

- 接口作为 DHCP 客户端多次申请 IP 地址失败后，将停止申请，并为接口配置缺省 IP 地址。
- 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，且会再向 DHCP 服务器重新申请 IP 地址。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口通过 DHCP 协议获取 IP 地址。

```
ip address dhcp-alloc
```

缺省情况下，接口不通过 DHCP 协议获取 IP 地址。

4.5 配置接口使用的DHCP客户端ID

1. 功能简介

DHCP 客户端 ID 用来填充 DHCP 报文 Option 61，作为识别 DHCP 客户端的唯一标识。DHCP 服务器可以根据客户端 ID 为特定的客户端分配特定的 IP 地址。DHCP 客户端 ID 包括类型和取值两部分，用户可以通过 ASCII 字符串、十六进制数和指定接口的 MAC 地址来指定 DHCP 客户端 ID：

- 当客户端 ID 的取值为 ASCII 字符串时，对应的类型值为 00；
- 当客户端 ID 的取值为十六进制数时，对应的类型值为该十六进制数的前两个字符；
- 当客户端 ID 使用指定接口的 MAC 地址时，对应的类型值为 01。

DHCP 客户端 ID 类型值可通过命令 **display dhcp server ip-in-use** 或 **display dhcp client** 进行查看。

2. 配置限制和指导

用户在指定客户端 ID 时，需要确保不同客户端的客户端 ID 不能相同。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口使用的 DHCP 客户端 ID。

```
dhcp client identifier { ascii ascii-string | hex hex-string | mac  
interface-type interface-number }
```

缺省情况下，根据本接口 MAC 地址生成 DHCP 客户端 ID，如果本接口没有 MAC 地址，则获取设备第一个以太接口的 MAC 地址生成 DHCP 客户端 ID。

4.6 开启地址冲突检查功能

1. 功能简介

通常情况下，DHCP 客户端上开启地址冲突检查功能，通过发送和接收 ARP 报文，对 DHCP 服务器分配的 IP 地址进行地址冲突检测。

如果攻击者仿冒地址拥有者进行 ARP 应答，就可以欺骗 DHCP 客户端，导致 DHCP 客户端无法正常使用分配到的 IP 地址。在网络中存在上述攻击者时，建议在客户端上关闭地址冲突检查功能。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启地址冲突检查功能。

```
dhcp client dad enable
```

缺省情况下，地址冲突检查功能处于开启状态。

4.7 配置DHCP客户端发送DHCP报文的DSCP优先级

1. 功能简介

DSCP 优先级用来体现报文自身的优先等级，决定报文传输的优先程度。通过本配置可以指定 DHCP 客户端发送的 DHCP 报文的 DSCP 优先级。

2. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 配置 DHCP 客户端发送 DHCP 报文的 DSCP 优先级。

```
dhcp client dscp dscp-value
```

缺省情况下，DHCP 客户端发送的 DHCP 报文的 DSCP 优先级为 56。

4.8 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4-1 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 <i>interface-type interface-number</i>]

5 BOOTP 客户端

5.1 BOOTP客户端简介

5.1.1 BOOTP 客户端的应用环境

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，该接口可以通过 BOOTP 协议从 BOOTP 服务器获取 IP 地址等信息，从而方便用户配置。

使用 BOOTP 协议时，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 BOOTP 协议需要在 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，推荐采用 DHCP 协议。

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。DHCP 服务器会按照服务器分配地址的优先次序为 BOOTP 客户端分配 IP 地址。

5.1.2 IP 地址动态获取过程

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。

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。

5.1.3 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

5.2 配置接口通过BOOTP协议获取IP地址

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

BOOTP 客户端中对于接口的相关配置，目前只能在三层以太网接口（包括子接口）、三层聚合接口和 VLAN 接口上进行。

- (3) 配置接口通过 BOOTP 协议获取 IP 地址。

ip address bootp-alloc

缺省情况下，接口不通过 BOOTP 协议获取 IP 地址。

5.3 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5-1 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 <i>interface-type</i> <i>interface-number</i>]

目 录

1 域名解析	1-1
1.1 域名解析简介	1-1
1.1.1 域名解析类型	1-1
1.1.2 静态域名解析	1-1
1.1.3 服务器域名解析	1-1
1.1.4 DNS 代理	1-2
1.1.5 DNS spoofing	1-4
1.2 域名解析配置任务简介	1-5
1.3 配置 DNS 客户端	1-5
1.3.1 功能简介	1-5
1.3.2 配置静态域名解析	1-5
1.3.3 配置服务器域名解析	1-6
1.4 配置 DNS proxy	1-7
1.4.1 开启 DNS proxy 功能	1-7
1.4.2 配置域名服务器的地址	1-7
1.5 配置 DNS 源地址透明代理	1-8
1.6 配置 DNS spoofing	1-10
1.7 配置 DNS Snooping 功能	1-11
1.8 开启 DNS Snooping 的日志功能	1-12
1.9 开启 DNS Snooping 的报文限速功能	1-13
1.10 配置 DNS 报文的源接口	1-14
1.11 配置 DNS 信任接口	1-15
1.12 指定 DNS 报文的 DSCP 优先级	1-15
1.13 配置 DNS 过滤功能	1-16
1.14 域名解析显示和维护	1-17
1.15 域名解析常见故障处理	1-17
1.15.1 无法解析到正确的 IP 地址	1-17
1.15.2 无法解析到正确的 IPv6 地址	1-18

1 域名解析

1.1 域名解析简介

DNS（Domain Name System，域名系统）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。

1.1.1 域名解析类型

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。由于动态域名解析需要域名服务器（DNS server）的配合，会花费一定的时间，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。

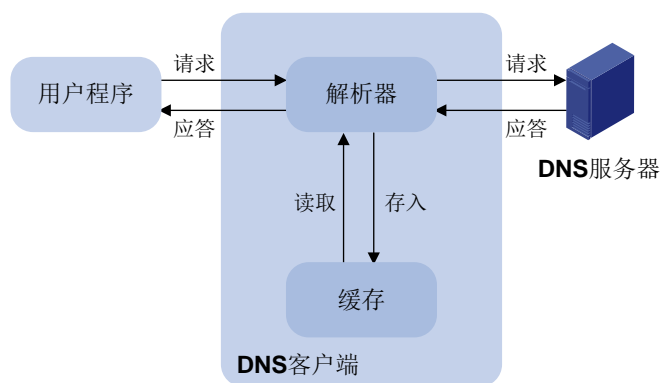
1.1.2 静态域名解析

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。

1.1.3 服务器域名解析

1. 体系结构

图1-1 服务器域名解析



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 客户端和域名服务器一般分布在两台设备上。目前，设备只能作为 DNS 客户端，不能作为 DNS 服务器。



说明

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

2. 解析过程

服务器域名解析通过向域名服务器查询域名和 IP 地址之间的对应关系来实现将域名解析为 IP 地址。服务器域名解析过程如下：

- (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；如果未发现所要查找的映射项，就向域名服务器发送查询请求。
- (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，就将请求交给其他域名服务器处理，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，将解析结果返回用户程序。

3. 缓存功能

服务器域名解析支持缓存功能。每次解析成功的域名与 IP 地址的映射均存放在 DNS 客户端的动态域名缓存区中，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老化而被删除，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 客户端从域名服务器的应答报文中获得老化时间。

4. 域名后缀列表功能

服务器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。例如，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结果。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）。

1.1.4 DNS 代理

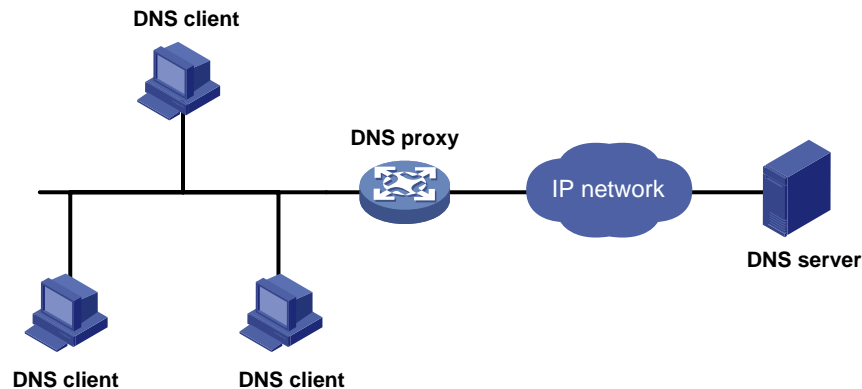
DNS 代理（DNS proxy）用来在 DNS client 和 DNS server 之间转发 DNS 请求和应答报文。局域网内的 DNS client 把 DNS proxy 当作 DNS server，将 DNS 请求报文发送给 DNS proxy。DNS proxy

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



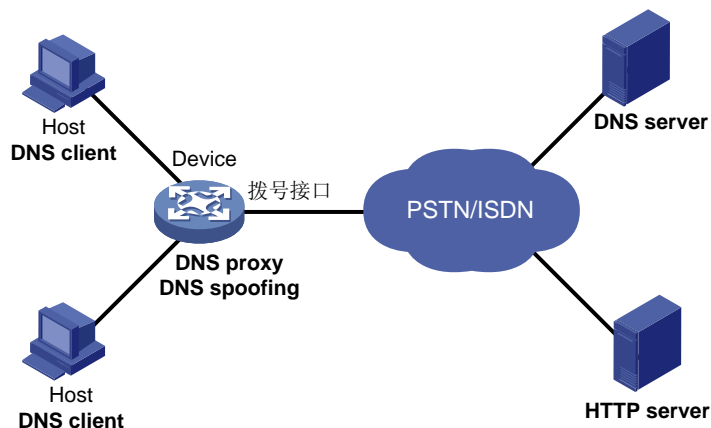
DNS 代理的工作过程如下：

- (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缓存表，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 将报文转发给域名服务器进行解析：
 - a. 当 DNS proxy 上配置了域名服务器组时，则 DNS proxy 根据域名匹配规则选择对应的域名服务器组，并将请求转发给服务器组中服务器；
 - b. 如果不存在匹配的域名服务器组，且配置了域名服务器地址时，DNS proxy 将报文转发给 DNS 服务器，通过 DNS 服务器进行域名解析。
- (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。

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。

1.1.5 DNS spoofing

图1-3 DNS spoofing 典型应用场景



DNS spoofing（DNS 欺骗）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 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 上未开启 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 DNS server 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，如果拨号接口尚未建立连接，Device 上不存在 DNS server 地址，或者设备上配置的 DNS server 地址均不可达，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“1.1.4 DNS 代理”。
- (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 域名解析配置任务简介

域名解析配置任务如下：

(1) [配置 DNS 客户端](#)

请至少选择其中一项进行配置。

- [配置静态域名解析](#)
- [配置服务器域名解析](#)

(2) （可选）配置 DNS 代理功能

- [配置 DNS proxy](#)
- [配置 DNS 源地址透明代理](#)

(3) （可选）[配置 DNS spoofing](#)

本功能用于拨号网络。

(4) （可选）配置 DNS 安全功能

- [配置 DNS Snooping 功能](#)
- [开启 DNS Snooping 的日志功能](#)
- [开启 DNS Snooping 的报文限速功能](#)

(5) （可选）[配置 DNS 报文的源接口](#)

(6) （可选）[配置 DNS 信任接口](#)

(7) （可选）[指定 DNS 报文的 DSCP 优先级](#)

(8) （可选）[配置 DNS 过滤功能](#)

1.3 配置DNS客户端

1.3.1 功能简介

如果 DNS 客户端上同时配置了静态域名解析和动态域名解析，则收到域名请求后，匹配顺序如下：

- (1) 静态域名解析；
- (2) 使用域名服务器进行域名解析。

如果以上方式均无法解析成功，则域名解析过程失败。

1.3.2 配置静态域名解析

1. 配置限制和指导

在公网或单个 VPN 实例内，一个主机名只能对应一个 IPv4 地址和 IPv6 地址。

公网或单个 VPN 实例内最多可以配置 2048 个主机名和地址的对应关系。可同时在公网和 VPN 实例内配置主机名和地址的对应关系。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置主机名和对应的地址。

(IPv4 网络)

```
ip host host-name ip-address [ vpn-instance vpn-instance-name ]
```

(IPv6 网络)

```
ipv6 host host-name ipv6-address [ vpn-instance vpn-instance-name ]
```

1.3.3 配置服务器域名解析

1. 配置限制和指导

- 设备上允许配置的域名服务器数目限制为：
 - 系统视图下，公网或单个 VPN 实例内最多可以配置 6 个域名服务器的 IPv4 地址。可同时在公网和 VPN 实例内配置域名服务器的 IPv4 地址。
 - 系统视图下，公网或单个 VPN 实例内最多可以配置 6 个域名服务器的 IPv6 地址。可同时在公网和 VPN 实例内配置域名服务器的 IPv6 地址。
 - 接口视图下，公网或单个 VPN 实例内最多可以配置 6 个域名服务器的 IPv4 地址。可同时在公网和 VPN 实例内配置域名服务器的 IPv4 地址。
- 如果同时配置域名服务器的 IPv4 地址和 IPv6 地址，DNS 客户端向域名服务器发送请求的处理方式如下：
 - 查询主机名对应的 IPv4 地址时，优先向域名服务器的 IPv4 地址发送查询请求。如果查询失败，则再向域名服务器的 IPv6 地址发送查询请求；
 - 查询主机名对应的 IPv6 地址时，优先向域名服务器的 IPv6 地址发送查询请求。如果查询失败，则再向域名服务器的 IPv4 地址发送查询请求。
- 域名服务器的优先级顺序为：
 - 系统视图下配置的域名服务器优先级高于接口视图下配置的域名服务器。
 - 配置多个域名服务器的 IPv4 地址时，IPv4 地址越小的域名服务器，其优先级越高。
 - 配置多个域名服务器的 IPv6 地址时，IPv6 地址越小的域名服务器，其优先级越高。
 - 设备上手工配置的域名服务器优先级高于通过 DHCP 等方式动态获取的域名服务器。

设备首先向优先级最高的域名服务器发送查询请求，失败后再根据优先级从高到低的次序向其他域名服务器发送查询请求。
- 配置域名解析后缀时，需要注意：
 - 公网或单个 VPN 实例内最多可以配置 16 个域名后缀。可同时在公网和 VPN 实例内配置域名后缀。
 - 添加域名后缀的优先级顺序为：先配置的域名后缀优先级高于后配置的域名后缀；设备上手工配置的域名后缀优先级高于通过 DHCP 等方式动态获取的域名后缀。设备首先添加优先级最高的域名后缀，查询失败后再根据优先级从高到低的次序添加其他域名后缀。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) (可选) 配置域名解析表项的有效时间。

```
dns cache ttl { maximum max-value | minimum min-value } *
```

缺省情况下，域名解析表项的有效时间为 DNS 响应报文中的 TTL。

- (3) (可选) 配置域名后缀。

```
dns domain domain-name [ vpn-instance vpn-instance-name ]
```

缺省情况下，未配置域名后缀，即只根据用户输入的域名信息进行解析。

- (4) 配置域名服务器的地址。

- 系统视图下配置域名服务器的地址。

(IPv4 网络)

```
dns server ip-address [ vpn-instance vpn-instance-name ]
```

(IPv6 网络)

```
ipv6 dns server ipv6-address [ interface-type interface-number ]  
[ vpn-instance vpn-instance-name ]
```

- 请依次执行以下命令在接口视图下配置域名服务器的 IPv4 地址。

```
interface interface-type interface-number
```

```
dns server ip-address [ vpn-instance vpn-instance-name ]
```

缺省情况下，未配置域名服务器的地址。

1.4 配置 DNS proxy

1.4.1 开启 DNS proxy 功能

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DNS proxy 功能。

```
dns proxy enable
```

缺省情况下，DNS proxy 功能处于关闭状态。

1.4.2 配置域名服务器的地址

1. 配置限制和指导

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

DNS proxy 可同时配置域名服务器的 IPv4 地址和 IPv6 地址。无论 DNS proxy 接收到的查询请求是来自 IPv4 客户端还是来自 IPv6 客户端，DNS proxy 都会按照优先级顺序向域名服务器的 IPv4 地址和 IPv6 地址转发查询请求。如果查询请求是 IPv4 报文，则优先向域名服务器的 IPv4 地址转发查询请求。如果查询请求是 IPv6 报文，则优先向域名服务器的 IPv6 地址转发查询请求。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置域名服务器的地址。

- o 系统视图下配置域名服务器的地址。

(IPv4 网络)

```
dns server ip-address [ vpn-instance vpn-instance-name ]
```

(IPv6 网络)

```
ipv6 dns server ipv6-address [ interface-type interface-number ]  
[ vpn-instance vpn-instance-name ]
```

- o 请依次执行以下命令在接口视图下配置域名服务器的 IPv4 地址。

```
interface interface-type interface-number
```

```
dns server ip-address [ vpn-instance vpn-instance-name ]
```

缺省情况下，未配置域名服务器的地址。

1.5 配置DNS源地址透明代理

1. 功能简介

DNS 透明代理指的是，DNS 客户端感知不到代理服务存在的代理模式。DNS 源地址透明代理不仅能够提供代理服务，还会修改 DNS 请求报文的源地址，从而实现发送 DNS 请求报文的设备能够接收到相应的 DNS 响应报文，适用于需要基于域名做策略的场景（如安全策略、带宽策略等）。

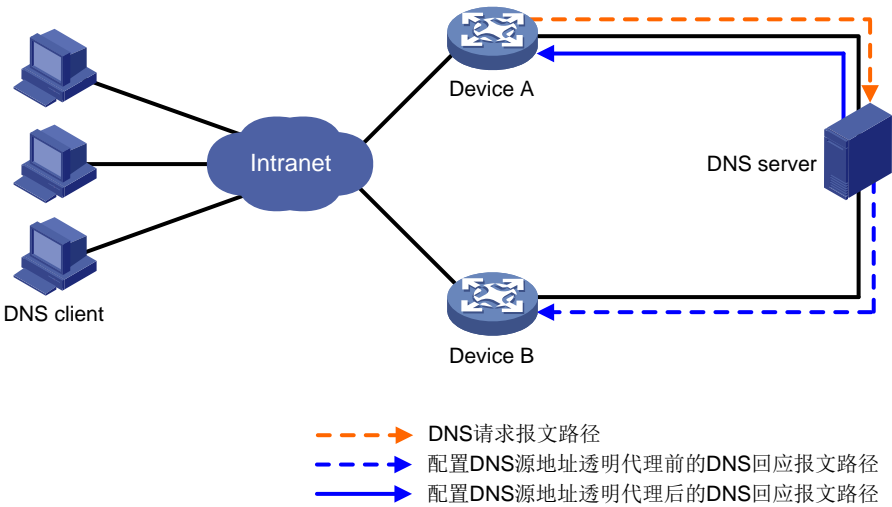
与 DNS 代理功能不同的是，DNS 客户端不需要将 DNS 服务器的地址指定为 DNS 源地址透明代理设备的地址，简化了客户端的配置。在一些负载均衡场景中，为了保证基于域名的策略能够成功对报文进行控制，建议使用 DNS 源地址透明代理功能。

开启 DNS 源地址透明代理功能后，设备开始监听过路的 DNS 请求报文和 DNS 应答报文，并记录域名解析信息，以便提供代理功能。具体工作机制如下：

- (1) DNS 源地址透明代理设备监听所有过路 DNS 报文，当收到 DNS 请求报文时，设备根据一定的规则从能够到达 DNS 服务器的本地 IP 地址选取一个，并将其作为修改后的 DNS 请求报文的源 IP 地址，使得 DNS 应答报文能够返回本设备。
- (2) DNS 源地址透明代理设备收到 DNS 服务器的应答报文后，记录域名解析的结果，并将报文转发给 DNS 客户端。DNS 客户端利用域名解析的结果进行相应的处理。后续 DNS 透明代理设备收到 DNS 请求报文后，首先查找 DNS 透明代理记录的表项，如果存在请求的信息，则 DNS 透明代理设备直接通过 DNS 应答报文将域名解析结果返回给 DNS client。如果不存在请求的信息，则 DNS 透明代理设备将报文转发给域名服务器进行解析。

如图 1-4 所示，在 Device A 上开启了 DNS 源地址透明代理功能。当 Device A 收到 DNS 请求报文后，将报文的源地址修改为本设备的地址，然后转发给 DNS 服务器。Device A 收到相应的 DNS 应答报文后，记录域名和 IP 地址的关系，并将 DNS 应答报文转发给 DNS 客户端。

图1-4 DNS 源地址透明代理原理图



2. 硬件适配关系

本特性的支持情况与设备型号有关，请以设备的实际情况为准。

设备型号	业务板类型	说明
M9006 M9010 M9014	Blade IV防火墙业务板	支持
	Blade V防火墙业务板	支持
	NAT业务板	支持
M9010-GM	加密业务板	支持
M9016-V	Blade V防火墙业务板	支持
M9008-S M9012-S	Blade IV防火墙业务板	支持
	入侵防御业务板	支持
	视频网关业务板	支持
M9008-S-V	Blade IV防火墙业务板	支持
M9000-AI-E4	Blade V防火墙业务板	不支持
M9000-AI-E8 M9000-AI-E16	Blade V防火墙业务板	支持
M9000-AK001	Blade V防火墙业务板	不支持
M9000-X06 M9000-X06-B M9000-X06-B-G M9000-X06-G M9000-X10	Blade VI防火墙业务板	支持
M9000-AI-X06 M9000-AI-X10	Blade VI防火墙业务板	支持

3. 配置限制和指导

DNS 源地址透明代理功能和 DNS Snooping 功能不能同时使用。

DNS 源地址透明代理功能不支持跨 VPN 使用,即设备上 DNS 报文的出入接口必须属于同一个 VPN。

4. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) (可选) 配置域名解析表项的有效时间。

```
dns cache ttl { maximum max-value | minimum min-value } *
```

缺省情况下, 域名解析表项的有效时间为 DNS 响应报文中的 TTL。

- (3) 开启 DNS 源地址透明代理功能。

```
dns transparent-proxy enable
```

缺省情况下, DNS 源地址透明代理功能处于关闭状态。

1.6 配置DNS spoofing

1. 配置限制和指导

公网或单个 VPN 实例内只能配置 1 个 DNS spoofing 应答的 IPv4 地址和 1 个 DNS spoofing 应答的 IPv6 地址。重复配置时, 新的配置会覆盖原有配置。

可同时在公网和 VPN 实例内配置 DNS spoofing 功能。

DNS spoofing 功能生效时, 即使设备上配置了静态域名解析, 也会使用 DNS spoofing 指定的 IP 地址来应答 DNS 请求。

2. 配置准备

设备上启用了 DNS proxy 功能。

设备上未指定域名服务器地址或不存在到达域名服务器的路由。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 启用 DNS proxy 功能。

```
dns proxy enable
```

缺省情况下, DNS proxy 功能处于关闭状态。

- (3) 开启 DNS spoofing 功能, 并指定 DNS spoofing 应答地址。

(IPv4 网络)

```
dns spoofing ip-address [ vpn-instance vpn-instance-name ]
```

(IPv6 网络)

```
ipv6 dns spoofing ipv6-address [ vpn-instance vpn-instance-name ]
```

缺省情况下, 未开启 DNS spoofing 功能。

1.7 配置DNS Snooping功能

1. 功能简介

DNS Snooping 功能适用于基于域名做策略的场景（如安全策略、带宽策略等）。设备使用基于域名的策略过滤用户流量时，需要获取域名对应的 IP 地址才能真正实现流量过滤。开启 DNS Snooping 功能后，设备会监听过路的 DNS 请求报文和 DNS 应答报文，如果 DNS 请求报文中的域名与策略中的域名相同，设备会在收到该域名的响应报文时记录域名解析结果，并上报给策略，使得策略可以基于此域名对应的 IP 地址实现流量过滤。如果 DNS 请求报文中的域名与过滤规则中的域名不同，设备不会记录域名解析结果。

2. 硬件适配关系

本特性的支持情况与设备型号有关，请以设备的实际情况为准。

设备型号	业务板类型	说明
M9006	Blade IV防火墙业务板	支持
M9010	Blade V防火墙业务板	支持
M9014	NAT业务板	支持
M9010-GM	加密业务板	支持
M9016-V	Blade V防火墙业务板	支持
M9008-S	Blade IV防火墙业务板	支持
M9012-S	入侵防御业务板	支持
	视频网关业务板	支持
M9008-S-V	Blade IV防火墙业务板	支持
M9000-AI-E4	Blade V防火墙业务板	不支持
M9000-AI-E8	Blade V防火墙业务板	支持
M9000-AI-E16		
M9000-AK001	Blade V防火墙业务板	不支持
M9000-X06	Blade VI防火墙业务板	支持
M9000-X06-B		
M9000-X06-B-G		
M9000-X06-G		
M9000-X10		
M9000-AI-X06	Blade VI防火墙业务板	支持
M9000-AI-X10		

3. 配置限制和指导

开启 DNS Snooping 功能时，需要注意：

- DNS Snooping 设备只有位于 DNS 客户端与 DNS 服务器之间，或 DNS 客户端与 DNS 代理设备之间时，DNS Snooping 功能配置后才能正常工作。
- DNS Snooping 功能和 DNS 源地址透明代理功能不能同时使用。

- DNS Snooping 功能不支持跨 VPN 使用，即设备上 DNS 报文的出入接口必须属于同一个 VPN。

4. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) （可选）配置域名解析表项的有效时间。

```
dns cache ttl { maximum max-value | minimum min-value } *
```

缺省情况下，域名解析表项的有效时间为 DNS 响应报文中的 TTL。

- (3) 开启 DNS Snooping 功能。

```
dns snooping enable
```

缺省情况下，DNS Snooping 功能处于关闭状态。

1.8 开启DNS Snooping的日志功能

1. 功能简介

网络环境中，DNS proxy 设备收到请求报文后，需要查询自己是否记录了请求域名对应的地址，如果存在，则直接应答需求；如果不存在，则需要向 DNS 服务器转发请求。如果网络中存在攻击源或有大量客户端同一时间发送大量 DNS 请求，则会增加网络负载并影响 DNS proxy 设备或 DNS 服务器的性能。为了防止上述问题的产生，可以在 DNS 客户端和 DNS proxy 或 DNS 服务器之间的设备配置 DNS Snooping 日志功能。

开启 DNS Snooping 的日志功能后，DNS 收到一个请求和对应的应答报文后会进行记录，并生成日志发给快速日志模块。管理员通过查询日志信息确认并解决问题。关于快速日志模块的详细介绍，请参见“网络管理和监控配置指导”中的“快速日志输出”。

2. 硬件适配关系

本特性的支持情况与设备型号有关，请以设备的实际情况为准。

设备型号	业务板类型	说明
M9006 M9010 M9014	Blade IV 防火墙业务板	支持
	Blade V 防火墙业务板	支持
	NAT 业务板	支持
M9010-GM	加密业务板	支持
M9016-V	Blade V 防火墙业务板	支持
M9008-S M9012-S	Blade IV 防火墙业务板	支持
	入侵防御业务板	支持
	视频网关业务板	支持
M9008-S-V	Blade IV 防火墙业务板	支持
M9000-AI-E4	Blade V 防火墙业务板	不支持
M9000-AI-E8	Blade V 防火墙业务板	支持

M9000-AI-E16		
M9000-AK001	Blade V防火墙业务板	不支持
M9000-X06 M9000-X06-B M9000-X06-B-G M9000-X06-G M9000-X10	Blade VI防火墙业务板	支持
M9000-AI-X06 M9000-AI-X10	Blade VI防火墙业务板	支持

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DNS Snooping 的日志功能。

```
dns snooping log enable
```

缺省情况下，DNS Snooping 的日志功能处于关闭状态。

1.9 开启DNS Snooping的报文限速功能

1. 功能简介

开启 DNS Snooping 的报文限速功能后，当接口上收到的 DNS 报文速率超过用户设定的限速值时，丢弃超过速率限制的 DNS 报文。

2. 硬件适配关系

本特性的支持情况与设备型号有关，请以设备的实际情况为准。

设备型号	业务板类型	说明
M9006	Blade IV防火墙业务板	支持
M9010	Blade V防火墙业务板	支持
M9014	NAT业务板	支持
M9010-GM	加密业务板	支持
M9016-V	Blade V防火墙业务板	支持
M9008-S	Blade IV防火墙业务板	支持
M9012-S	入侵防御业务板	支持
	视频网关业务板	支持
M9008-S-V	Blade IV防火墙业务板	支持
M9000-AI-E4	Blade V防火墙业务板	不支持
M9000-AI-E8 M9000-AI-E16	Blade V防火墙业务板	支持

M9000-AK001	Blade V 防火墙业务板	不支持
M9000-X06 M9000-X06-B M9000-X06-B-G M9000-X06-G M9000-X10	Blade VI 防火墙业务板	支持
M9000-AI-X06 M9000-AI-X10	Blade VI 防火墙业务板	支持

3. 配置限制和指导

只有开启了 DN 源地址透明代理功能或 DNS Snooping 的日志功能，本功能才会生效。

4. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DNS Snooping 的报文限速功能。

```
dns snooping rate-limit rate
```

缺省情况下，DNS Snooping 的报文限速功能处于关闭状态。

1.10 配置DNS报文的源接口

1. 功能简介

缺省情况下，设备根据域名服务器的地址，通过路由表查找请求报文的出接口，并将该出接口的主 IP 地址作为发送到该服务器的 DNS 请求报文的源地址。根据域名服务器的地址不同，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，域名服务器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。

2. 配置限制和指导

发送 IPv4 DNS 报文时，将使用源接口的主 IPv4 地址作为 DNS 报文的源地址。发送 IPv6 DNS 报文时，将根据 RFC 3484 中定义的规则从源接口上选择 IPv6 地址作为 DNS 报文的源地址。如果源接口上未配置对应的地址，则将导致报文发送失败。

公网或单个 VPN 实例内只能配置 1 个源接口。重复配置时，新的配置会覆盖原有配置。可同时在公网和 VPN 实例内配置源接口。

无论配置的源接口是否属于指定的 VPN，该配置都会生效。不建议为某个 VPN 配置一个不属于该 VPN 的源接口。否则，设备会使用不属于该 VPN 的地址作为 DNS 报文源地址，导致无法收到 DNS 应答。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 指定 DNS 报文的源接口。

```
dns source-interface interface-type interface-number [ vpn-instance
vpn-instance-name ]
```

缺省情况下，未指定 DNS 报文的源接口。

1.11 配置DNS信任接口

1. 功能简介

缺省情况下，任意接口通过 DHCP 等协议动态获得的域名后缀和域名服务器信息都将作为有效信息，用于域名解析。如果网络攻击者通过 DHCP 服务器为设备分配错误的域名后缀和域名服务器地址，则会导致设备域名解析失败，或解析到错误的结果。通过本配置指定信任接口后，域名解析时只采用信任接口动态获得的域名后缀和域名服务器信息，非信任接口获得的信息不能用于域名解析，从而在一定程度上避免这类攻击。

2. 配置限制和指导

设备最多可以配置 128 个 DNS 信任接口。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 指定 DNS 信任接口。

```
dns trust-interface interface-type interface-number
```

缺省情况下，未指定任何接口为信任接口。

1.12 指定DNS报文的DSCP优先级

1. 功能简介

DSCP 优先级用来体现报文自身的优先等级，决定报文传输的优先程度。通过本配置可以指定设备发送的 DNS 报文的 DSCP 优先级。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 指定 DNS 客户端或 DNS proxy 发出的 DNS 报文的 DSCP 优先级。

(IPv4 网络)

```
dns dscp dscp-value
```

缺省情况下，DNS 报文的 DSCP 优先级为 0。

(IPv6 网络)

```
ipv6 dns dscp dscp-value
```

缺省情况下，IPv6 DNS 报文的 DSCP 优先级为 0。

1.13 配置DNS过滤功能

1. 功能简介

通过在 DNS 代理上开启 DNS 过滤功能，可以实现对用户通过域名进行的业务访问进行控制。开启 DNS 过滤功能后，DNS 代理将会提取 DNS 客户端发送的 DNS 请求报文中的域名与本功能配置的白名单或黑名单进行匹配，根据匹配结果对 DNS 请求报文执行放行或丢弃动作。

DNS 过滤功能的机制如下：

- 如果 DNS 代理收到的 DNS 请求报文中的域名命中白名单，则 DNS 代理放行该 DNS 请求报文，并在收到 DNS 响应报文后记录域名解析的结果，然后将 DNS 响应报文转发给 DNS 客户端。如果 DNS 代理收到的 DNS 请求报文中的域名未命中白名单，则 DNS 代理丢弃该 DNS 请求报文。
- 如果 DNS 代理收到的 DNS 请求报文中的域名未命中黑名单，则 DNS 代理放行该 DNS 请求报文，并在收到 DNS 响应报文后记录域名解析的结果，然后将 DNS 响应报文转发给 DNS 客户端。如果 DNS 代理收到的 DNS 请求报文中的域名命中黑名单，则 DNS 代理丢弃该 DNS 请求报文。

如果希望实现严格的访问控制，建议使用白名单进行 DNS 过滤。如果希望实现宽松的访问控制，建议使用黑名单进行 DNS 过滤。

2. 硬件适配关系

本特性的支持情况与设备型号有关，请以设备的实际情况为准。

设备型号	业务板类型	说明
M9006 M9010 M9014	Blade IV 防火墙业务板	支持
	Blade V 防火墙业务板	支持
	NAT 业务板	支持
M9010-GM	加密业务板	支持
M9016-V	Blade V 防火墙业务板	支持
M9008-S M9012-S	Blade IV 防火墙业务板	支持
	入侵防御业务板	支持
	视频网关业务板	支持
M9008-S-V	Blade IV 防火墙业务板	支持
M9000-AI-E4	Blade V 防火墙业务板	不支持
M9000-AI-E8 M9000-AI-E16	Blade V 防火墙业务板	支持
M9000-AK001	Blade V 防火墙业务板	不支持
M9000-X06 M9000-X06-B M9000-X06-B-G M9000-X06-G M9000-X10	Blade VI 防火墙业务板	支持

设备型号	业务板类型	说明
M9000-AI-X06 M9000-AI-X10	Blade VI防火墙业务板	支持

3. 配置限制和指导

可以配置多个白名单或多个黑名单，但不允许同时配置白名单和黑名单。

4. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 DNS 过滤功能，并配置黑/白名单。

```
dns filter { allowlist | denylist } hostname
```

缺省情况下，DNS 过滤功能处于关闭状态。

1.14 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1-1 域名解析显示和维护

操作	命令
显示域名后缀信息	display dns domain [dynamic] [vpn-instance vpn-instance-name]
显示域名解析表信息	display dns host [ip ipv6] [vpn-instance vpn-instance-name]
显示域名服务器的IPv4地址信息	display dns server [dynamic] [vpn-instance vpn-instance-name]
显示域名服务器的IPv6地址信息	display ipv6 dns server [dynamic] [vpn-instance vpn-instance-name]
显示DNS Snooping记录的域名解析信息	display dns snooping host [ip ipv6] [vpn-instance vpn-instance-name]
清除动态域名解析缓存信息	reset dns host [ip ipv6] [vpn-instance vpn-instance-name]

1.15 域名解析常见故障处理

1.15.1 无法解析到正确的 IP 地址

1. 故障现象

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. 处理过程

- 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。
- 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。

1.15.2 无法解析到正确的 IPv6 地址

1. 故障现象

配置了动态域名解析，但不能根据域名解析到正确的 IPv6 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 IPv6 地址。

3. 处理过程

- 执行命令 **display dns host ipv6**，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。
- 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 IPv6 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 IP 转发基础	1-1
1.1 IP 转发表简介	1-1
1.2 vSystem 相关说明	1-1
1.3 开启保持上一跳功能	1-2
1.4 开启备份上一跳功能	1-3
1.5 IP 转发表显和维护	1-3
2 负载分担	2-1
2.1 负载分担简介	2-1
2.2 配置负载分担方式	2-1
2.3 开启 IPv4 基于带宽的负载分担功能	2-1
2.4 开启等价路由负载分担本地优先功能	2-2

1 IP 转发基础

1.1 IP转发表简介

FIB（Forwarding Information Base，转发信息库）表用来指导 IP 报文转发。

路由器通过路由表选择路由，把优选路由下发到 FIB 表中，通过 FIB 表指导 IP 报文转发。FIB 表中每条转发表项都指明了要到达某子网或某主机的报文的下一跳 IP 地址以及出接口。

关于路由表的详细介绍，请参见“三层技术-IP 路由配置指导”中的“IP 路由基础”。

通过命令 **display fib** 可以查看 FIB 表的信息，例如：

```
<Sysname> display fib
```

```
Destination count: 4 FIB entry count: 4
```

Flag:

```
U:Useable    G:Gateway    H:Host      B:Blackhole  D:Dynamic    S:Static
R:Relay      F:FRR
```

Destination/Mask	Nexthop	Flag	OutInterface/Token	Label
10.2.0.0/16	10.2.1.1	U	GE1/0/1	Null
10.2.1.1/32	127.0.0.1	UH	InLoop0	Null
127.0.0.0/8	127.0.0.1	U	InLoop0	Null
127.0.0.1/32	127.0.0.1	UH	InLoop0	Null

FIB 表中包含了下列关键项：

- **Destination:** 目的地址。用来标识 IP 报文的目的地址或目的网络。
- **Mask:** 网络掩码。与目的地址一起来标识目的主机或路由器所在的网段的地址。将目的地址和网络掩码“逻辑与”后可得到目的主机或路由器所在网段的地址。例如：目的地址为 192.168.1.40、掩码为 255.255.255.0 的主机或路由器所在网段的地址为 192.168.1.0。掩码由若干个连续“1”构成，既可以用点分十进制法表示，也可以用掩码中连续“1”的个数来表示。
- **NextHop:** 转发的下一跳地址。
- **Flag:** 路由的标志。
- **OutInterface:** 转发接口。指明 IP 报文将从哪个接口转发。
- **Token:** LSP（Label Switched Path，标签交换路径）索引号。
- **Label:** 内层标签值。

1.2 vSystem相关说明

非缺省 vSystem 不支持本特性部分功能，包括开启基于 VPN Peer 的业务功能和开启 IPv4 基于带宽的负载分担功能。



说明

非缺省 vSystem 对具体命令的支持情况，请见本特性的命令参考。有关 vSystem 的详细介绍请参见“虚拟化技术配置指导”中的“vSystem”。

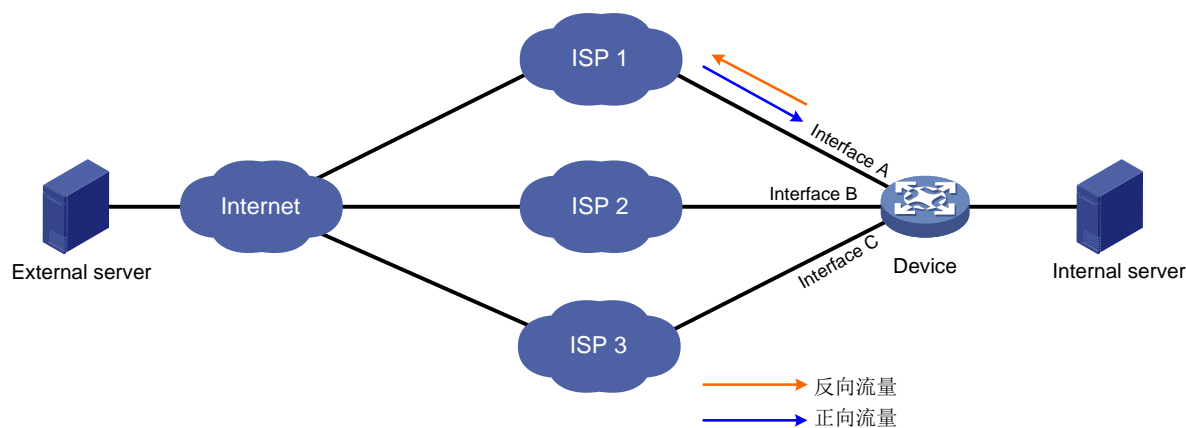
1.3 开启保持上一跳功能

1. 功能简介

接口上开启保持上一跳功能后，当该接口接收到正向流量的第一个 IP 报文，设备会根据流量特征以及上一跳信息，建立相反方向的快速转发表项，当反向流量报文到达设备进行转发时，可以直接通过该快速转发表项指导报文进行转发，使对端到本端的正向流量和本端到对端的反向流量走的是相同的路径，从而保证同一会话的流量能够进行相同的业务处理。

如图 1-1 所示，外网服务器向内网服务器发起业务请求，请求报文通过 ISP1 到达 Device 设备，访问内网服务器。用户希望相同会话或连接的正向流量与反向流量保持相同的转发路径，即回应报文到达 Device 设备后，通过接收请求报文的 Interface A 接口转发出去，经过 ISP1 到达外网服务器。未开启保持上一跳功能时，Device 设备会选择最佳链路进行报文的转发，这样就无法保证正反向流量路径一致。这种情况下，用户可以在接收正向流量的 Interface A 接口上开启保持上一跳功能。

图1-1 保持上一跳功能组网应用



2. 配置限制和指导

- 保持上一跳功能依赖于快速转发表项的建立，对于以太网类型的链路，如果上一跳的 MAC 地址发生变化，对应的快速转发表项需要重建才能使保持上一跳功能正常工作。
- 本特性不适用于 IRF 组网中，跨成员设备转发的业务报文。
- 本特性不适用于 RBM 双机热备组网中非对称流量的场景。有关 RBM 双机热备的详细介绍，请参考“高可靠性配置指导”中的“双机热备（RBM）”。
- 在支持部署多安全业务板的设备上，本特性对外部主动访问设备的流量不生效。

3. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 进入接口视图。

```
interface interface-type { interface-number |  
interface-number.subnumber }
```

- (3) 开启保持上一跳功能。

```
ip last-hop hold
```

缺省情况下，转发保持上一跳功能处于关闭状态。

1.4 开启备份上一跳功能

1. 功能简介

在 IRF 环境下，为了使对端设备到本端设备的正向流量和本端设备到对端设备的反向流量走相同的路径，可以在主设备的接口上开启转发保持上一跳功能，并在全局开启备份上一跳功能和会话热备功能（配置 **session synchronization enable** 命令）后，当该接口接收到正向流量的第一个 IP 报文，会保存上一跳信息，同时将该上一跳信息备份到从设备，当反向流量报文到达从设备上时可以直接通过该上一跳信息进行转发。关于 **session synchronization enable** 命令的详细解释请参见“安全命令参考”中的“会话管理”。

当设备上存在多个业务板，且板间有业务备份时，在接口上开启保持上一跳功能，并在全局开启备份上一跳功能和会话引流功能（配置 **session flow-redirect enable** 命令）后，当该接口接收到正向流量的第一个 IP 报文后，接收到该报文的业务板会保存上一跳信息，同时将该上一跳信息备份到所有业务板上，当反向流量报文到达本业务板或其他业务板上时可以直接通过该上一跳信息进行转发。关于 **session flow-redirect enable** 命令的详细解释请参见“安全命令参考”中的“会话管理”。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启备份上一跳功能。

```
last-hop backup enable
```

缺省情况下，备份上一跳功能处于开启状态。

1.5 IP转发表显和维护

查看转发表的信息是定位转发问题的基本方法。在任意视图下执行 **display** 命令可以显示转发表信息。

表1-1 IP 转发表显示和维护

操作	命令
显示FIB表项的信息	<div>(独立运行模式)</div> <div>display fib [topology <i>topology-name</i> vpn-instance <i>vpn-instance-name</i>] [<i>ip-address</i> [<i>mask</i> <i>mask-length</i>]] [slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</div> <div>(IRF模式)</div> <div>display fib [topology <i>topology-name</i> vpn-instance <i>vpn-instance-name</i>] [<i>ip-address</i> [<i>mask</i> <i>mask-length</i>]] [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</div>

2 负载分担

2.1 负载分担简介

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

2.2 配置负载分担方式

1. 功能简介

配置负载分担的内容包括：

- 配置负载分担方式：设备上存在多条等价路由时，可以根据报文中的信息（源 IP 地址、目的 IP 地址、源端口、目的端口和 IP 协议号）配置逐流进行负载分担，或者根据报文进行逐包负载分担。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置负载分担方式。

（独立运行模式）

```
ip load-sharing mode { per-flow [ dest-ip | dest-port | ip-pro | src-ip |  
src-port ] * | per-packet } { global | slot slot-number [ cpu cpu-number ] }
```

（IRF 模式）

```
ip load-sharing mode { per-flow [ dest-ip | dest-port | ip-pro | src-ip  
| src-port ] * | per-packet } { chassis chassis-number slot slot-number  
[ cpu cpu-number ] | global }
```

缺省情况下，设备基于报文逐流进行负载分担。

2.3 开启IPv4基于带宽的负载分担功能

1. 功能简介

开启 IPv4 基于带宽的负载分担功能情况下，如果转发时查到多个出接口/下一跳，则按照接口的带宽值计算出各个接口应该分配的报文比例，然后按照带宽比例对报文进行转发。

支持负载分担的协议（如 LISP）的设备，无论是否配置负载分担命令，负载分担比例以协议定义的负载分担比例为准。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 IPv4 基于带宽的负载分担功能。

bandwidth-based-sharing

缺省情况下，IPv4 基于带宽的负载分担功能处于关闭状态。

- (3) （可选）配置接口的期望带宽值。

- a. 进入接口视图。

interface *interface-type* *interface-number*

- b. 配置接口的期望带宽值。

bandwidth *bandwidth*

缺省情况下，接口期望带宽为接口的物理带宽。

2.4 开启等价路由负载分担本地优先功能

1. 功能简介

当 IRF 设备转发报文时，如果查询到的是等价路由且出接口在不同成员设备上，可能会将报文透传到某个成员设备再发送，这会使报文转发效率变低，也会影响成员设备间的数据处理能力。当配置了等价路由负载分担本地优先的功能以后，如果在处理报文的成员设备上存在等价路由的出接口，就只从当前设备发送报文，而不会再透传到其他成员设备发送。

2. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 开启等价路由负载分担本地优先功能。

ip load-sharing local-first enable

缺省情况下，等价路由负载分担本地优先功能处于关闭状态。

目 录

1 快速转发	1-1
1.1 快速转发简介.....	1-1
1.2 vSystem 相关说明	1-1
1.3 快速转发配置限制和指导.....	1-1
1.4 配置快速转发表项的老化时间	1-1
1.5 配置快速转发负载分担	1-2
1.6 配置硬件快速转发功能	1-2
1.7 配置对于 GRE 和 VXLAN 报文支持基于 DSCP 进行快速转发	1-3
1.8 配置硬件快速转发芯片出口将增量校验和封装到报文中	1-3
1.9 开启硬件快速转发芯片出口对报文的篡改检测功能.....	1-3
1.10 配置设备检测到报文被篡改时执行的动作	1-4
1.11 配置上行业务报文的处理模式为单硬件芯片转发模式	1-4
1.12 配置安全业务板的报文处理模式.....	1-5
1.13 开启硬件快速转发忽略接口序列号匹配功能	1-5
1.14 配置会话申请锁功能.....	1-6
1.15 配置逻辑聚合选路模式为 CRC Hash 模式	1-6
1.16 开启获取硬件快速转发会话状态功能	1-7
1.17 开启逻辑畸形报文检测功能	1-7
1.18 配置 SIP=DIP 畸形报文的丢弃功能.....	1-7
1.19 快速转发显示和维护	1-8

1 快速转发

1.1 快速转发简介

报文转发效率是衡量设备性能的一项关键指标。按照常规流程，设备收到一个报文后，根据报文的地址寻找路由表中与之匹配的路由，然后确定一条最佳的路径，同时还将报文按照数据链路层上使用的协议进行封装，最后进行报文转发。

快速转发是采用高速缓存来处理报文，采用了基于数据流的技术。

快速转发根据报文中的信息（比如源 IP 地址、目的 IP 地址、源端口、目的端口、IP 协议号等）来标识一条数据流。当一条数据流的第一个报文通过查找路由表转发后，在高速缓存中生成相应的转发信息，该数据流后续报文的转发就可以通过直接查找快速转发表进行转发。这样便大大缩减了 IP 报文的排队流程，减少报文的转发时间，提高 IP 报文的转发速率。

1.2 vSystem 相关说明

非缺省 vSystem 不支持本特性部分功能，包括配置硬件快速转发和配置快速转发功能。



说明

非缺省 vSystem 对具体命令的支持情况，请见本特性的命令参考。有关 vSystem 的详细介绍请参见“虚拟化技术配置指导”中的“vSystem”。

1.3 快速转发配置限制和指导

快速转发能处理已经分片的 IP 报文，但不支持对 IP 报文的再分片。

根据处理方式不同，快速转发分为软件快速转发和硬件快速转发。除非特别指明，否则下文中的快速转发均指软件快速转发。

1.4 配置快速转发表项的老化时间

1. 功能简介

快速转发表中的表项并非永远有效，每一条记录都有一个生存周期，到达生存周期仍得不到刷新的记录将从快速转发表中删除，这个生存周期被称作老化时间。如果在到达老化时间前纪录被刷新，则重新计算老化时间。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置快速转发表项的老化时间。

```
ip fast-forwarding aging-time aging-time
```

缺省情况下，快速转发表项的老化时间为 30 秒。

1.5 配置快速转发负载分担

1. 功能简介

缺省情况下，快速转发负载分担功能处于开启状态，快速转发根据报文中的信息来标识一条数据流；关闭快速转发负载分担功能后，快速转发根据报文中的信息和入接口来标识一条数据流。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置快速转发负载分担功能。请选择其中一项进行配置。

- 开启快速转发负载分担功能。

```
ip fast-forwarding load-sharing
```

- 关闭快速转发负载分担功能。

```
undo ip fast-forwarding load-sharing
```

缺省情况下，快速转发负载分担功能处于开启状态。

1.6 配置硬件快速转发功能

1. 功能简介

硬件快速转发功能可以在系统建立快转的时候存储会话信息，以便后续流量可以通过匹配会话表项加速报文的转发。当需要定位转发芯片是否存在故障时，可以关闭硬件快速转发功能。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置硬件快速转发功能。请选择其中一项进行配置。

- 开启硬件快速转发功能。

（独立运行模式）

```
hardware fast-forwarding enable [ slot slot-number [ cpu cpu-number ] ]
```

（IRF 模式）

```
hardware fast-forwarding enable [ chassis chassis-number slot  
slot-number [ cpu cpu-number ] ]
```

- 关闭硬件快速转发功能。

（独立运行模式）

```
undo hardware fast-forwarding enable [ slot slot-number [ cpu  
cpu-number ] ]
```

（IRF 模式）

```
undo hardware fast-forwarding enable [ chassis chassis-number slot  
slot-number [ cpu cpu-number ] ]
```


缺省情况下，硬件快速转发功能处于开启状态。

1.7 配置对于GRE和VXLAN报文支持基于DSCP进行快速转发

1. 功能简介

配置本功能后，对于 GRE 和 VXLAN 报文，使用外层报文的 DSCP 替代源端口号标识数据流。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置对于 GRE 和 VXLAN 报文支持基于 DSCP 进行快速转发。

```
ip fast-forwarding dscp
```

缺省情况下，GRE 和 VXLAN 报文不支持基于 DSCP 进行快速转发。

- (3) （可选）配置用于识别 VXLAN 报文的目的 UDP 端口号。

```
ip fast-forwarding vxlan-port port-number
```

缺省情况下，用于识别 VXLAN 报文的目的 UDP 端口号为 4789。

1.8 配置硬件快速转发芯片出口将增量校验和封装到报文中

- (1) 进入系统视图。

```
system-view
```

- (2) 配置硬件快速转发芯片出口将增量校验和封装到报文中。

（独立运行模式）

```
hardware fast-forwarding checksum encap incremental [ slot slot-number  
cpu cpu-number ]
```

（IRF 模式）

```
hardware fast-forwarding checksum encap incremental [ chassis  
chassis-number slot slot-number cpu cpu-number ]
```

缺省情况下，硬件快速转发芯片出口将增量校验和封装到报文中。

1.9 开启硬件快速转发芯片出口对报文的篡改检测功能

- (1) 进入系统视图。

```
system-view
```

- (2) 开启硬件快速转发芯片出口对报文的篡改检测功能。

（独立运行模式）

```
hardware fast-forwarding checksum inspect [ 13 | 14 [ tcp | udp ] ] enable  
[ slot slot-number cpu cpu-number ]
```

（IRF 模式）

```
hardware fast-forwarding checksum inspect [ 13 | 14 [ tcp | udp ] ] enable  
[ chassis chassis-number slot slot-number cpu cpu-number ]
```

缺省情况下，硬件快速转发芯片出口对报文的各种篡改检测功能均处于开启状态。

1.10 配置设备检测到报文被篡改时执行的动作

- (1) 进入系统视图。

system-view

- (2) 配置设备检测到报文被篡改时执行的动作。

（独立运行模式）

```
hardware fast-forwarding checksum inspect action { drop-err | log }  
[ slot slot-number cpu cpu-number ]
```

（IRF 模式）

```
hardware fast-forwarding checksum inspect action { drop-err | log }  
[ chassis chassis-number slot slot-number cpu cpu-number ]
```

缺省情况下，设备检测到报文被篡改时执行的动作为转发报文和生成日志。

1.11 配置上行业务报文的处理模式为单硬件芯片转发模式

1. 功能简介

对于包含双硬件转发芯片的业务板，当只需要单个硬件转发芯片处理业务报文时，使用本命令将上行业务报文处理模式配置为单硬件芯片转发模式。模式切换后，上行报文由单硬件转发芯片处理，下行报文由双硬件转发芯片处理。

2. 配置限制和指导

配置本功能后，只有重启相应业务板，模式切换才能生效。如果安全引擎组内有多个安全引擎（即多业务板），需要针对安全引擎组内所有业务板执行本命令进行模式切换（即保证安全引擎组内所有业务板都处于同一转发模式下）并重启安全引擎组内所有业务板后，模式切换才能生效。有关安全引擎组的详细介绍请参见“虚拟化技术配置指导”中的“Context”。

对于仅包含单硬件转发芯片的业务板，本功能不生效。

3. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 配置上行业务报文的处理模式为单硬件芯片转发模式。

（独立运行模式）

```
hardware fast-forwarding standalone [ slot slot-number [ cpu  
cpu-number ] ]
```

（IRF 模式）

```
hardware fast-forwarding standalone [ chassis chassis-number slot  
slot-number [ cpu cpu-number ] ]
```

缺省情况下，上行业务报文的处理模式为双硬件芯片转发模式。

1.12 配置安全业务板的报文处理模式

1. 功能简介

当安全业务板上的 CPU 出现单核流量过大的情况时，可以将业务板的报文处理模式配置为抗攻击模式，以避免攻击流量影响其他 CPU 核心的正常处理。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 将安全业务板的报文处理模式配置为抗攻击模式。

```
hardware processing-mode attack-resistance
```

缺省情况下，安全业务板的报文处理模式为 CPU 模式，该模式下业务板使用 CPU 进行报文处理。

抗攻击模式下，业务板优先使用硬件快速转发芯片进行报文处理。

1.13 开启硬件快速转发忽略接口序列号匹配功能

1. 功能简介

在等价双出口组网环境中，一条流量的回程报文可能经由不同的入接口送至设备。这样会导致设备将这些入接口不同的报文视作不同的流量，进而无法对其进行硬件快速转发。

开启本功能后，硬件快速转发功能将不再要求报文的入接口序列号相同，同一条流量的后续报文可以正常进行硬件快速转发。

2. 配置限制和指导

只有在开启硬件快速转发功能的情况下，本功能才能生效。

如发现转发异常情况，可关闭本功能进行调试。

在等价双出口组网环境中，开启本功能后会影设备性能，管理员可根据当下网络现况，判断是否关闭本功能，以提高设备性能。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启硬件快速转发忽略接口序列号匹配功能。

（独立运行模式）

```
hardware fast-forwarding ifsn match enable [ slot slot-number cpu  
cpu-number ]
```

（IRF 模式）

```
hardware fast-forwarding ifsn match enable [ chassis chassis-number  
slot slot-number cpu cpu-number ]
```

1.14 配置会话申请锁功能

1. 配置限制和指导

此命令仅在 Blade IV 安全业务板和 Blade V 安全业务板上支持。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置会话申请锁功能。请选择其中一项进行配置。

- 关闭会话申请锁功能。

（独立运行模式）

```
hardware fast-forwarding session-lock disable slot slot-number cpu  
cpu-number
```

（IRF 模式）

```
hardware fast-forwarding session-lock disable chassis chassis-number  
slot slot-number cpu cpu-number
```

- 开启会话申请锁功能。

（独立运行模式）

```
undo hardware fast-forwarding session-lock disable slot slot-number cpu  
cpu-number
```

（IRF 模式）

```
undo hardware fast-forwarding session-lock disable chassis  
chassis-number slot slot-number cpu cpu-number
```

缺省情况下，会话申请锁功能处于开启状态。

1.15 配置逻辑聚合选路模式为CRC Hash模式

1. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置逻辑聚合选路模式为 CRC Hash 模式。

```
hardware fast-forwarding link-aggregation hash-mode crc
```

缺省情况下，逻辑聚合选路模式为异或模式

- (3) （可选）配置参与逻辑聚合选路 CRC 运算的 IPv6 地址的偏移量。

```
hardware fast-forwarding link-aggregation hash-mode crc ip-offset  
offset-vlaue
```

缺省情况下，参与逻辑聚合选路 CRC 运算的 IPv6 地址的偏移量为 0。

当逻辑聚合选路模式为 CRC HASH 模式时，执行本命令可以配置参与逻辑聚合选路 CRC 运算的 IPv6 地址的偏移量。从 *offset-vlaue* 参数所示位置开始（包含此比特），选取 32 比特进行逻辑聚合选路 CRC 运算。

1.16 开启获取硬件快速转发会话状态功能

1. 功能简介

开启本功能后，通过执行 **display session table ipv4 verbose/display session table ipv6** 命令可以查看到获取的硬件快速转发会话的状态。

目前仅 Blade4、Blade5 板卡支持配置本命令。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启获取硬件快速转发会话状态功能。

```
hardware fast-forwarding session-state enable
```

缺省情况下，获取硬件快速转发会话状态功能处于开启状态。

1.17 开启逻辑畸形报文检测功能

- (1) 进入系统视图。

```
system-view
```

- (2) 开启逻辑畸形报文检测功能。

（独立运行模式）

```
hardware fast-forwarding malpkt-filter enable [ slot slot-number cpu  
cpu-number ]
```

```
undo hardware fast-forwarding malpkt-filter enable [ slot slot-number  
cpu cpu-number ]
```

（IRF 模式）

```
hardware fast-forwarding malpkt-filter enable [ chassis chassis-number  
slot slot-number cpu cpu-number ]
```

```
undo hardware fast-forwarding malpkt-filter enable [ chassis  
chassis-number slot slot-number cpu cpu-number ]
```

缺省情况下，逻辑畸形报文检测功能处于开启状态。

1.18 配置SIP=DIP畸形报文的丢弃功能

1. 功能简介

在某些攻击情况下，攻击者会通过 SIP=DIP（即请求报文的源 IP 地址和目的 IP 地址相同）的方式进行网络欺骗。通过开启本命令，当检测到 SIP=DIP 畸形报文时，硬件转发芯片直接对报文进行丢弃处理，以增强网络的安全性和稳健性。

配置本命令前，请先执行 **hardware fast-forwarding malpkt-filter enable** 命令开启设备对畸形报文的检测功能。

2. 配置限制与指导

此命令仅在 Blade IV 安全业务板上配置生效。

3. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 配置 SIP=DIP 畸形报文的丢弃功能。

(独立运行模式)

```
hardware fast-forwarding malpkt-filter sip_dip discard [ slot
slot-number cpu cpu-number ]
undo hardware fast-forwarding malpkt-filter sip_dip discard [ slot
slot-number cpu cpu-number ]
```

(IRF 模式)

```
hardware fast-forwarding malpkt-filter sip_dip discard [ chassis
chassis-number slot slot-number cpu cpu-number ]
undo hardware fast-forwarding malpkt-filter sip_dip discard [ chassis
chassis-number slot slot-number cpu cpu-number ]
```

缺省情况下，硬件转发芯片对检测到的 SIP=DIP 畸形报文不进行丢弃，直接上送到 CPU 处理。

1.19 快速转发显示和维护

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

在用户视图下执行 **reset** 命令可以清除快速转发表中的内容。

表1-1 快速转发显示和维护

操作	命令
显示快速转发表项的老化时间	display ip fast-forwarding aging-time
显示快速转发表信息	(独立运行模式) display ip fast-forwarding cache [ip-address] [slot slot-number [cpu cpu-number]] (IRF模式) display ip fast-forwarding cache [ip-address] [chassis chassis-number slot slot-number [cpu cpu-number]]
显示虚拟分片报文快速转发表信息	(独立运行模式) display ip fast-forwarding fragcache [ip-address] [slot slot-number [cpu cpu-number]] (IRF模式) display ip fast-forwarding fragcache [ip-address] [chassis chassis-number slot slot-number [cpu cpu-number]]

操作	命令
清除快速转发表信息	<p>(独立运行模式)</p> <pre>reset ip fast-forwarding cache [slot slot-number [cpu cpu-number]]</pre> <p>(IRF模式)</p> <pre>reset ip fast-forwarding cache [chassis chassis-number slot slot-number [cpu cpu-number]]</pre>

目 录

1 多 CPU 报文负载分担	1-1
1.1 多 CPU 报文负载分担简介	1-1
1.2 多 CPU 报文负载分担配置限制和指导	1-1
1.3 配置报文负载分担策略	1-1

1 多 CPU 报文负载分担

1.1 多CPU报文负载分担简介

多核设备上，报文在 CPU 之间进行负载分担的策略包括：

- 基于流处理：用五元组（源 IP 地址、源端口号、目的 IP 地址、目的端口号、协议号）来区分和划定一条流，同一条流被分配到同一个或多个 CPU 进行处理，处理过程保证先进先出。
- 基于报文处理：将报文依次发送到不同的 CPU 进行处理，不保证报文的处理顺序。

1.2 多CPU报文负载分担配置限制和指导

当负载分担策略配置为基于报文处理时，同一条流的报文会被发送到不同的 CPU 进行处理，这样无法保证报文的处理顺序。当启动某些业务时，该业务可能无法处理在不同的 CPU 的同一条流的报文而将报文丢弃。

1.3 配置报文负载分担策略

- (1) 进入系统视图。

```
system-view
```

- (2) 配置报文负载分担策略。

```
forwarding policy { per-flow [ mode { destination-ip | destination-port  
| source-ip | source-port } | three-tuple ] | per-packet }
```

缺省情况下，采用基于流处理的报文负载分担策略。

目 录

1 邻接表	1-1
1.1 邻接表简介	1-1
1.2 vSystem 相关说明	1-1
1.3 邻接表显示和维护	1-1

1 邻接表

1.1 邻接表简介

邻接表用于管理各种链路层协议（如 PPP）的邻居信息。此处的邻居，是指 IP 层面的邻居，即对于三层转发来说一跳可达，不需要经过中间设备进行三层转发。

各种链路层协议通过协商（如 PPP 动态协商）或配置生成邻居信息后，将其下发给邻接表，生成邻接表表项。邻接表中记录了邻居的网络层地址（下一跳）、路由出接口、链路层协议类型、链路层地址等信息。邻接表表项的更新、删除也由各链路层协议模块通知完成。

IP/IPv6 转发时，设备通过查找 FIB（Forwarding Information Base，转发信息库）/IPv6 FIB 表项得到报文的出接口和下一跳信息，再以此出接口和下一跳为索引查找邻接表，获取到该下一跳的链路层转发信息，如链路层协议（PPP 等）及介质类型（P2P、NBMA）、封装报文的链路层头信息等，然后根据此信息对报文进行封装后转发。



说明

以太网类型邻居信息和非以太网类型邻居信息统一存储和管理，本文所描述的邻接表特指管理非以太网类型的邻居信息。

1.2 vSystem相关说明

vSystem 支持本特性的所有功能。有关 vSystem 的详细介绍请参见“虚拟化技术配置指导”中的“vSystem”。

1.3 邻接表显示和维护

在任意视图下执行 **display** 命令可以显示邻接表项的信息。

表1-1 邻接表显示和维护

操作	命令
显示IPv4邻接表项	<p>（独立运行模式）</p> <pre>display adjacent-table { all physical-interface <i>interface-type interface-number</i> routing-interface <i>interface-type interface-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>] } [count verbose]</pre> <p>（IRF模式）</p> <pre>display adjacent-table { all physical-interface <i>interface-type interface-number</i> routing-interface <i>interface-type interface-number</i> chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>] } [count verbose]</pre>

操作	命令
显示IPv6邻接表项	<p>(独立运行模式)</p> <pre>display ipv6 adjacent-table { all physical-interface interface-type interface-number routing-interface interface-type interface-number slot slot-number [cpu cpu-number] } [count verbose]</pre> <p>(IRF模式)</p> <pre>display ipv6 adjacent-table { all physical-interface interface-type interface-number routing-interface interface-type interface-number chassis chassis-number slot slot-number [cpu cpu-number] } [count verbose]</pre>

目 录

1 IP 性能优化	1-1
1.1 IP 性能优化配置任务简介	1-1
1.2 配置允许接口接收和转发直连网段的定向广播报文	1-1
1.2.1 功能简介	1-1
1.2.2 配置步骤	1-2
1.3 配置接口发送 IPv4 报文的 MTU	1-2
1.4 开启三层报文统计功能	1-2
1.5 开启 IP 分片报文本地重组功能	1-3
1.6 开启 IPv4 虚拟分片透传功能	1-3
1.7 强制关闭 IPv4 虚拟分片重组功能	1-4
1.8 开启 IPv6 虚拟分片透传功能	1-4
1.9 强制关闭 IPv6 虚拟分片重组功能	1-5
1.10 配置 ICMP 差错报文发送功能	1-5
1.10.1 功能简介	1-5
1.10.2 开启 ICMP 重定向报文发送功能	1-6
1.10.3 开启 ICMP 超时报文发送功能	1-6
1.10.4 开启 ICMP 目的不可达报文发送功能	1-7
1.11 配置发送 ICMP 差错报文对应的令牌刷新周期和令牌桶容量	1-7
1.12 指定 ICMP 报文源地址	1-8
1.13 配置接口的 TCP 最大报文段长度	1-8
1.14 配置缺省 TCP 最大报文段长度	1-9
1.15 配置 TCP 连接的 Path MTU 探测功能	1-9
1.16 开启 SYN Cookie 功能	1-10
1.17 配置 TCP 连接的缓冲区大小	1-11
1.18 配置 TCP 定时器	1-11
1.19 配置发送 TCP 报文时添加 TCP 时间戳选项信息	1-11
1.20 IP 性能优化显示和维护	1-12

1 IP 性能优化

1.1 IP性能优化配置任务简介

如下所有配置均为可选，请根据实际情况选择配置。

- 配置 IP 报文功能
 - [配置允许接口接收和转发直连网段的定向广播报文](#)
 - [配置接口发送 IPv4 报文的 MTU](#)
 - [开启三层报文统计功能](#)
 - [开启 IP 分片报文本地重组功能](#)
本功能适用于 IRF 组网环境。
 - [开启 IPv4 虚拟分片透传功能](#)
 - [强制关闭 IPv4 虚拟分片重组功能](#)
 - [开启 IPv6 虚拟分片透传功能](#)
 - [强制关闭 IPv6 虚拟分片重组功能](#)
- 配置 ICMP 报文功能
 - [配置 ICMP 差错报文发送功能](#)
 - [配置发送 ICMP 差错报文对应的令牌刷新周期和令牌桶容量](#)
 - [指定 ICMP 报文源地址](#)
- 配置 TCP 报文功能
 - [配置接口的 TCP 最大报文段长度](#)
 - [配置缺省 TCP 最大报文段长度](#)
 - [配置 TCP 连接的 Path MTU 探测功能](#)
 - [开启 SYN Cookie 功能](#)
 - [配置 TCP 连接的缓冲区大小](#)
 - [配置 TCP 定时器](#)
 - [配置发送 TCP 报文时添加 TCP 时间戳选项信息](#)

1.2 配置允许接口接收和转发直连网段的定向广播报文

1.2.1 功能简介

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。

接口接收和转发直连网段的定向广播报文包括以下几种情况：

- 在接收定向广播报文的情况下，如果在接口上配置了此命令，设备允许接收此接口直连网段的定向广播报文。

- 在转发定向广播报文的情况下，如果在接口上配置了此命令，设备从其他接口接收到目的地址为此接口直连网段的定向广播报文时，会从此接口转发此类报文。

1.2.2 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置允许接口接收和转发面向直连网段的定向广播报文。

```
ip forward-broadcast
```

缺省情况下，设备禁止转发直连网段的定向广播报文；设备允许接收直连网段的定向广播报文。

1.3 配置接口发送IPv4报文的MTU

1. 功能简介

当设备使用某个接口发送报文时，发现报文长度大于该接口的发送 IPv4 报文的 MTU 值，则进行下列处理：

- 如果报文不允许分片，则将报文丢弃；
- 如果报文允许分片，则将报文进行分片转发。

为了减轻转发设备在传输过程中的分片和重组数据包的压力，更高效的利用网络资源，请根据实际组网环境设置合适的接口 MTU 值，以减少分片的发生。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置发送 IPv4 报文的 MTU。

```
ip mtu mtu-size
```

缺省情况下，未配置接口发送 IPv4 报文的 MTU。

1.4 开启三层报文统计功能

1. 功能简介

开启本功能后，设备会统计接口接收或发送的 IP 报文的数量，该统计信息可通过 **display ip intreface** 命令查看；也会统计接口的 IP 报文统计速率，统计速率可通过命令行 **display interface** 查看。接口报文流量过大时，开启本功能会造成设备 CPU 占用率高，影响转发性能。因此，当用户不需要统计接口接收的 IP 报文数量时，建议关闭本功能。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 开启三层报文统计功能。

```
statistics l3-packet enable [ inbound | outbound ]
```

缺省情况下，三层报文统计功能处于关闭状态。

1.5 开启IP分片报文本地重组功能

1. 功能简介

当某单板收到目的为本设备的 IP 分片报文时，需要把分片报文送到主用主控板进行重组，这样会导致报文重组性能较低的问题。当开启 IP 分片报文本地重组功能后，分片报文会在该单板直接进行报文重组，这样就能提高报文的重组性能。开启 IP 分片报文本地重组功能后，如果分片报文是从设备上不同的单板进入的，会导致 IP 分片报文本地无法重组成功。

多台设备组成的 IRF 环境下，当某成员设备收到目的为本 IRF 设备的 IP 分片报文时，需要把分片报文送到主设备进行重组，这样会导致报文重组性能较低的问题。当开启 IP 分片报文本地重组功能后，分片报文会在该成员设备上直接进行报文重组，这样就能提高分片报文的重组性能。开启 IP 分片报文本地重组功能后，如果分片报文是从设备上不同的成员设备进入的，会导致 IP 分片报文本地无法重组成功。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 IP 分片报文本地重组功能。

```
ip reassemble local enable
```

缺省情况下，IP 分片报文本地重组功能处于关闭状态。

1.6 开启IPv4虚拟分片透传功能

1. 功能简介

在高可靠性组网环境下，如果设备上开启了 IPv4 虚拟分片重组功能，当出现同一条流的分片报文分别到达了两个设备的情况时，每个设备上收到的分片报文都不完整，IPv4 虚拟分片重组功能无法对不完整的分片报文进行重组，因此会对分片报文进行丢弃。配置本命令后，未收到首个分片的设备，会将接收到的分片报文全部透传给收到同一条流首片报文所在的设备，实现同一条流的分片报文在一个设备上集中处理，即 IPv4 虚拟分片重组。

2. 配置限制和指导

在非高可靠性组网环境中，或者设备上未开启 IPv4 虚拟分片重组功能，不建议配置本功能。

有关高可靠性组网，请参见“高可靠性配置指导”中的“双机热备（RBM）”。

本功能在三层转发加 **INLINE** 转发的场景下生效，且不支持隧道功能。关于 **INLINE** 转发的详情请参考“二层技术-以太网交换配置指导”中的“二层转发”。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 IPv4 虚拟分片透传功能。

```
ip virtual-reassembly centralize
```

缺省情况下，IPv4 虚拟分片透传功能处于关闭状态。

1.7 强制关闭IPv4虚拟分片重组功能

1. 功能简介

IPv4 虚拟分片重组功能可以提前对到达设备的分片报文进行检验、排序和缓存。开启 IPv4 虚拟分片重组功能，可以保证后续的报文重组功能处理的都是顺序正确的分片报文。

在高可靠性组网环境中，如果出现同一条流的分片报文分别到达了两个设备的情况，则每个设备上收到的分片报文都不完整，此时设备的 IPv4 虚拟分片重组功能无法对不完整的分片报文进行重组，因此会对其进行丢弃。

如果需要设备对接收到不完整的分片报文进行放行，则可以强行关闭设备的 IPv4 虚拟分片重组功能。

对于已配置 ASPF、连接数限制业务的设备，配置本功能后，ASPF、连接数限制业务将不会对分片报文进行校验，而是直接放行。

2. 配置限制和指导

本命令仅对三层转发的安全策略业务生效，关于安全策略的详情请参见“安全配置指导”中的“安全策略”。请明确实际组网对分片报文重组功能的需求后，谨慎配置本功能。

有关高可靠性组网，请参见“高可靠性配置指导”中的“双机热备（RBM）”。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 强制关闭 IPv4 虚拟分片重组功能。

```
ip virtual-reassembly suppress
```

缺省情况下，强制关闭 IP 虚拟分片重组功能处于关闭状态。

1.8 开启IPv6虚拟分片透传功能

1. 功能简介

在高可靠性组网环境下，如果设备上开启了 IPv6 虚拟分片重组功能，当出现同一条流的分片报文分别到达了两个设备的情况时，每个设备上收到的分片报文都不完整，IPv6 虚拟分片重组功能无法对不完整的分片报文进行重组，因此会对分片报文进行丢弃。配置本命令后，未收到首个分片的设备，会将接收到的分片报文全部透传给收到同一条流首片报文所在的设备，实现同一条流的分片报文在一个设备上集中处理，即 IPv6 虚拟分片重组。

2. 配置限制和指导

在非高可靠性组网环境中，或者设备上未开启 IPv6 虚拟分片重组功能，不建议配置本功能。
有关高可靠性组网，请参见“高可靠性配置指导”中的“双机热备（RBM）”。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 IPv6 虚拟分片透传功能。

```
ipv6 virtual-reassembly centralize
```

缺省情况下，IPv6 虚拟分片透传功能处于关闭状态。

1.9 强制关闭IPv6虚拟分片重组功能

1. 功能简介

IPv6 虚拟分片重组功能可以提前对到达设备的分片报文进行检验、排序和缓存。开启 IPv6 虚拟分片重组功能，可以保证后续的报文重组功能处理的都是顺序正确的分片报文。

在高可靠性组网环境中，如果出现同一条流的分片报文分别到达了两个设备的情况，则每个设备上收到的 IPv6 分片报文都不完整，此时设备的 IPv6 虚拟分片重组功能无法对不完整的分片报文进行重组，因此会对其进行丢弃。

如果需要设备对接收到不完整的分片报文进行放行，则可以强行关闭设备的 IPv6 虚拟分片重组功能。

对于已配置 ASPF、连接数限制业务的设备，配置本命令后，ASPF、连接数限制业务将不会对分片报文进行校验，而是直接放行。

2. 配置限制和指导

请明确实际组网对分片报文重组功能的需求后，谨慎配置本功能。

有关高可靠性组网，请参见“高可靠性配置指导”中的“双机热备（RBM）”。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 强制关闭 IPv6 虚拟分片重组功能。

```
ipv6 virtual-reassembly suppress
```

缺省情况下，强制关闭 IPv6 虚拟分片重组功能处于关闭状态。

1.10 配置ICMP差错报文发送功能

1.10.1 功能简介

ICMP 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。ICMP 差错报文的发送虽然方便了网络的控制管理，但是也存在缺陷：发送大量的 ICMP 报文，增大网络流量；如果有用户发送 ICMP 差错报文进行恶意攻击，会导致设备性能下降或影响正常工

作。为了避免上述现象发生，缺省情况下，ICMP 差错报文发送功能处于关闭状态，用户可以根据需要开启 ICMP 差错报文发送功能。

ICMP 差错报文包括重定向报文、超时报文和目的不可达报文。

1.10.2 开启 ICMP 重定向报文发送功能

1. 功能简介

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。

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。

同时满足下列条件时，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；
- 报文的源 IP 地址和报文接收接口的 IP 地址在同一个网段；
- 数据报文中没有源路由选项。

2. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 开启 ICMP 重定向报文发送功能。

```
ip redirects enable
```

缺省情况下，ICMP 重定向报文发送功能处于关闭状态。

1.10.3 开启 ICMP 超时报文发送功能

1. 功能简介

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。

2. 配置限制和指导

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。

3. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 开启 ICMP 超时报文发送功能。

```
ip ttl-expires enable
```

缺省情况下，ICMP 超时报文发送功能处于关闭状态。

1.10.4 开启 ICMP 目的不可达报文发送功能

1. 功能简介

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，如果在路由表中未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分片比特”ICMP 差错报文。

2. 配置限制和指导

设备开启 DHCP 服务后，在未发送 ICMP 回显请求(ECHO-REQUEST)报文情况下，收到非法 ICMP 回显应答(ECHO-REPLY)报文，此时设备不会回应“协议不可达”ICMP 差错报文。执行 **dhcp enable** 命令可以开启 DHCP 服务，关于 **dhcp enable** 的详细介绍，请参见“三层技术-IP 业务命令参考”中的“DHCP”。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 ICMP 目的不可达报文发送功能。

```
ip unreachable enable
```

缺省情况下，ICMP 目的不可达报文发送功能处于关闭状态。

1.11 配置发送ICMP差错报文对应的令牌刷新周期和令牌桶容量

1. 功能简介

如果网络中短时间内发送的 ICMP 差错报文过多，将可能导致网络拥塞。为了避免这种情况，用户可以控制设备在指定时间内发送 ICMP 差错报文的数目，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，即每隔多长时间发放一个令牌到令牌桶中，直到令牌桶中的令牌数达到配置的容量。一个令牌表示允许发送一个 ICMP 差错报文，每当发送一个 ICMP 差错报文，则令牌桶中减少一个令牌。如果连续发送的 ICMP 差错报文超过了令牌桶的容量，则后续的 ICMP 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。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置发送 ICMP 差错报文对应的令牌刷新周期和令牌桶容量。

```
ip icmp error-interval interval [ bucket-size ]
```

缺省情况下，令牌刷新周期为 100 毫秒，令牌桶容量为 10。

刷新周期为 0 时，表示不限制 ICMP 差错报文的发送。

1.12 指定ICMP报文源地址

1. 功能简介

在网络中 IP 地址配置较多的情况下，收到 ICMP 报文时，用户很难根据报文的源 IP 地址判断报文来自哪台设备。为了简化这一判断过程，可以指定 ICMP 报文源地址。用户配置特定地址（如环回口地址）为 ICMP 报文的源地址，可以简化判断。

设备发送 ICMP 差错报文（TTL 超时、端口不可达和参数错误等）和 ping echo request 报文时，都可以通过上述命令指定报文的源地址。

2. 配置限制和指导

用户发送 ping echo request 报文时，如果 ping 命令中已经指定源地址，则使用该源地址，否则使用 `ip icmp source` 配置的源地址。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 指定 ICMP 报文源地址。

```
ip icmp source [ vpn-instance vpn-instance-name ] ip-address
```

缺省情况下，未指定 ICMP 报文源地址。

发送 ICMP 差错报文（TTL 超时、端口不可达和参数错误等）时，设备使用触发 ICMP 差错报文的原始报文的入接口 IP 地址作为 ICMP 报文源地址。

发送 ICMP echo request 报文时，设备使用出接口 IP 地址作为 ICMP 报文源地址。

发送 ICMP echo reply 报文时，设备使用 ICMP echo request 报文的目的地址作为 ICMP 报文源地址。

1.13 配置接口的TCP最大报文段长度

1. 功能简介

TCP 最大报文段长度（Maximum Segment Size, MSS）表示 TCP 连接的对端发往本端的最大 TCP 报文段的长度，目前作为 TCP 连接建立时的一个选项来协商：当一个 TCP 连接建立时，连接的双方要将 MSS 作为 TCP 报文的一个选项通告给对端，对端会记录下这个 MSS 值，后续在发送 TCP 报文时，会限制 TCP 报文的大小不超过该 MSS 值。当对端发送的 TCP 报文的长度小于本端的 TCP 最大报文段长度时，TCP 报文不需要分段；否则，对端需要对 TCP 报文按照最大报文段长度进行分段处理后再发给本端。

2. 配置限制和指导

- 用户可以通过下面的命令配置接口的 TCP 最大报文段长度，配置后该接口接收和发送的 TCP 报文的大小都不能超过该值。
- 该配置仅对新建的 TCP 连接生效，对于配置前已建立的 TCP 连接不生效。
- 该配置仅对 IP 报文生效，当接口上配置了 MPLS 功能后，不建议再配置本功能。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口的 TCP 最大报文段长度。

```
tcp mss value
```

缺省情况下，未配置接口的 TCP 最大报文段长度。

1.14 配置缺省TCP最大报文段长度

1. 功能简介

TCP 建立连接后，TCP 连接会根据 TCP 最大报文段长度对报文进行分片后再传输。一般情况下，TCP 最大报文段长度等于出接口的 IP MTU 减 40。目前，缺省 TCP 最大报文段长度较小，这会导致报文分片过多，且会存在接收连接方应答的 TCP SYN 报文的 TCP 校验和错误等问题。为了解决这个问题，管理员可以通过本命令手工指定缺省 TCP 最大报文段长度。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置缺省 TCP 最大报文段长度。

```
tcp default-mss mss-value
```

缺省情况下，缺省 TCP 最大报文段长度为 512。

1.15 配置TCP连接的Path MTU探测功能

1. 功能简介

通过开启 TCP 连接的 Path MTU 探测功能，用户可确定 TCP 路径上从源端到目的端的最小 MTU（Path MTU），按照 Path MTU 组织 TCP 分段长度，避免 IP 分片的发生。为了在 Path MTU 增大时，减少资源浪费，可以开启 Path MTU 老化定时器，保证设备尽量按照 TCP 路径允许的最大报文长度发送数据。

RFC 1191 中规定的 TCP 连接的 Path MTU 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， $MSS = \text{最小 MTU 值} - \text{IP 头部长度} - \text{TCP 头部长度}$ 。

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）。

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 分钟内未收到 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。

2. 配置准备

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 unreachable enable`)，以确保 ICMP 差错报文可以发送到 TCP 源端。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 TCP 连接的 Path MTU 探测功能。

```
tcp path-mtu-discovery [ aging age-time | no-aging ]
```

缺省情况下，TCP 连接的 Path MTU 探测功能处于关闭状态。

1.16 开启 SYN Cookie 功能

1. 功能简介

SYN Cookie 功能用来防止 SYN Flood 攻击。SYN Flood 攻击中，攻击者向设备发送大量请求建立 TCP 连接的 SYN 报文，而不回应设备的 SYN ACK 报文，导致设备上建立了大量的 TCP 半连接。从而，达到耗费设备资源，使设备无法处理正常业务的目的。配置 SYN Cookie 功能后，当设备收到 TCP 连接请求时，不建立 TCP 半连接，而直接向发起者回复 SYN ACK 报文。设备接收到发起者回应的 ACK 报文后，建立连接，并进入 ESTABLISHED 状态。通过这种方式，可以避免在设备上建立大量的 TCP 半连接，防止设备受到 SYN Flood 攻击。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 SYN Cookie 功能。

tcp syn-cookie enable

缺省情况下，SYN Cookie 功能处于关闭状态。

1.17 配置TCP连接的缓冲区大小

- (1) 进入系统视图。

system-view

- (2) 配置 TCP 连接的接收和发送缓冲区的大小。

tcp window window-size

缺省情况下，TCP 连接的接收和发送缓冲区大小为 63KB。

1.18 配置TCP定时器

1. TCP 定时器简介

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 定时器和重传 SYN 报文定时器，当 synwait 定时器超时且 SYN 报文重传未达到最大次数时，如果设备未收到回应报文，则 TCP 连接建立不成功；当 synwait 定时器未超时但是 SYN 报文重传达到最大次数时，如果设备未收到回应报文，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未收到报文，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。

2. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 配置 TCP 的 synwait 定时器超时时间。

tcp timer syn-timeout time-value

缺省情况下，synwait 定时器超时时间为 75 秒。

- (3) 配置 TCP 的 finwait 定时器超时时间。

tcp timer fin-timeout time-value

缺省情况下，finwait 定时器超时时间为 675 秒。

1.19 配置发送TCP报文时添加TCP时间戳选项信息

1. 功能简介

TCP 报文中携带 TCP 时间戳选项信息时，建立 TCP 连接的两台设备通过 TCP 报文中的时间戳字段就可计算出 RTT（Round Trip Time，往返时间）值。在某些组网中，由于安全隐患，需要防止 TCP 连接上的中间设备获取到 TCP 时间戳信息，可以在建立 TCP 连接任意一端关闭发送 TCP 报文时添加时间戳选项信息功能。

2. 配置步骤

- (1) 进入系统视图。
`system-view`
- (2) 配置发送 TCP 报文时添加 TCP 时间戳选项信息。
`tcp timestamps enable`
缺省情况下，发送 TCP 报文时会添加 TCP 时间戳选项信息。

1.20 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 性能优化功能后的运行情况，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1-1 IP 性能优化显示和维护

操作	命令
显示ICMP流量统计信息	(独立运行模式) <code>display icmp statistics [slot slot-number [cpu cpu-number]]</code> (IRF模式) <code>display icmp statistics [chassis chassis-number slot slot-number [cpu cpu-number]]</code>
显示IP报文统计信息	(独立运行模式) <code>display ip statistics [slot slot-number [cpu cpu-number]]</code> (IRF模式) <code>display ip statistics [chassis chassis-number slot slot-number [cpu cpu-number]]</code>
显示RawIP连接摘要信息	(独立运行模式) <code>display rawip [slot slot-number [cpu cpu-number]]</code> (IRF模式) <code>display rawip [chassis chassis-number slot slot-number [cpu cpu-number]]</code>
显示RawIP连接详细信息	(独立运行模式) <code>display rawip verbose [slot slot-number [cpu cpu-number] [pcb pcb-index]]</code> (IRF模式) <code>display rawip verbose [chassis chassis-number slot slot-number [cpu cpu-number] [pcb pcb-index]]</code>

操作	命令
显示TCP连接摘要信息	(独立运行模式) display tcp [slot slot-number [cpu cpu-number]] (IRF模式) display tcp [chassis chassis-number slot slot-number [cpu cpu-number]]
显示TCP连接的流量统计信息	(独立运行模式) display tcp statistics [slot slot-number [cpu cpu-number]] (IRF模式) display tcp statistics [chassis chassis-number slot slot-number [cpu cpu-number]]
显示TCP连接详细信息	(独立运行模式) display tcp verbose [slot slot-number [cpu cpu-number] [pcb pcb-index]] (IRF模式) display tcp verbose [chassis chassis-number slot slot-number [cpu cpu-number] [pcb pcb-index]]
显示TCP代理连接的简要信息	(独立运行模式) display tcp-proxy slot slot-number [cpu cpu-number] (IRF模式) display tcp-proxy chassis chassis-number slot slot-number [cpu cpu-number]
显示TCP代理非保留端口的使用信息	(独立运行模式) display tcp-proxy port-info slot slot-number [cpu cpu-number] (IRF模式) display tcp-proxy port-info chassis chassis-number slot slot-number [cpu cpu-number]
显示UDP连接摘要信息	(独立运行模式) display udp [slot slot-number [cpu cpu-number]] (IRF模式) display udp [chassis chassis-number slot slot-number [cpu cpu-number]]
显示UDP流量统计信息	(独立运行模式) display udp statistics [slot slot-number [cpu cpu-number]] (IRF模式) display udp statistics [chassis chassis-number slot slot-number [cpu cpu-number]]

操作	命令
显示UDP连接详细信息	(独立运行模式) display udp verbose [slot <i>slot-number</i> [cpu <i>cpu-number</i>] [pcb <i>pcb-index</i>]] (IRF模式) display udp verbose [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>] [pcb <i>pcb-index</i>]]
显示当前开启的INET业务的信息	(独立运行模式) display inet open-service [slot <i>slot-number</i> [cpu <i>cpu-number</i>]] (IRF模式) display inet open-service [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>]]
清除IP报文统计信息	(独立运行模式) reset ip statistics [slot <i>slot-number</i> [cpu <i>cpu-number</i>]] (IRF模式) reset ip statistics [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>]]
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 IPv6 基础.....	1-1
1.1 IPv6 简介	1-1
1.1.1 IPv6 协议特点	1-1
1.1.2 IPv6 地址介绍	1-2
1.1.3 IPv6 PMTU 发现	1-5
1.1.4 IPv6 过渡技术介绍	1-5
1.1.5 协议规范	1-6
1.2 vSystem 相关说明	1-7
1.3 IPv6 基础配置任务简介	1-7
1.4 配置 IPv6 全球单播地址	1-8
1.4.1 功能简介	1-8
1.4.2 采用 EUI-64 格式形成 IPv6 地址	1-8
1.4.3 手工指定 IPv6 地址	1-8
1.4.4 无状态自动配置 IPv6 地址	1-9
1.4.5 引用前缀生成接口上的 IPv6 地址，并将此前缀分配给终端设备	1-10
1.5 配置 IPv6 链路本地地址	1-10
1.5.1 功能简介	1-10
1.5.2 配置限制和指导	1-11
1.5.3 配置自动生成链路本地地址	1-11
1.5.4 手工指定接口的链路本地地址	1-11
1.6 配置 IPv6 任播地址	1-11
1.7 配置 PMTU 发现	1-12
1.7.1 配置接口 MTU	1-12
1.7.2 配置指定地址的静态 PMTU	1-12
1.7.3 配置 PMTU 老化时间	1-13
1.8 配置 ICMPv6 报文发送功能	1-13
1.8.1 配置发送 ICMPv6 差错报文对应的令牌桶容量和令牌刷新周期	1-13
1.8.2 配置允许回复组播形式的 Echo request 报文	1-13
1.8.3 配置 ICMPv6 目的不可达差错报文发送功能	1-14
1.8.4 配置 ICMPv6 超时差错报文发送功能	1-14
1.8.5 配置 ICMPv6 重定向报文发送功能	1-15
1.8.6 配置 ICMPv6 报文指定源地址功能	1-16
1.9 开启 IPv6 分片报文本本地重组功能	1-16

1.10 开启 IPv6 报文扩展头丢弃功能.....	1-17
1.11 IPv6 基础显示和维护	1-17
2 IPv6 邻居发现	2-1
2.1 IPv6 邻居发现简介.....	2-1
2.1.1 IPv6 邻居发现使用的 ICMPv6 消息	2-1
2.1.1 地址解析	2-1
2.1.2 验证邻居是否可达	2-2
2.1.3 重复地址检测	2-2
2.1.4 路由器发现/前缀发现及地址无状态自动配置	2-2
2.1.5 重定向功能.....	2-3
2.1.6 协议规范	2-3
2.2 vSystem 相关说明	2-3
2.3 IPv6 邻居发现配置任务简介	2-4
2.4 配置静态邻居表项	2-4
2.5 配置接口上允许动态学习的邻居的最大个数.....	2-5
2.6 开启接口从未经请求的 NA 报文中学习邻居信息的功能	2-5
2.7 配置 STALE 状态 ND 表项的老化时间	2-6
2.8 配置链路本地 ND 表项资源占用最小化	2-6
2.9 配置设备的跳数限制	2-7
2.10 配置允许发布 RA 消息及相关参数	2-7
2.10.1 RA 消息及相关参数介绍	2-7
2.10.2 配置限制和指导	2-8
2.10.3 配置允许发布 RA 消息	2-8
2.10.4 配置 RA 消息中的相关参数.....	2-8
2.10.5 配置 RA 消息中的 DNS 服务器信息.....	2-10
2.10.6 配置 RA 消息中的 DNS 域名后缀信息	2-10
2.10.7 开启 RA 消息中的 DNS 信息抑制功能	2-11
2.11 配置重复地址检测时发送邻居请求消息的次数	2-12
2.12 配置 IPv6 地址冲突自恢复功能.....	2-12
2.13 配置 ND Proxy 功能.....	2-13
2.13.1 功能简介	2-13
2.13.2 配置普通 ND Proxy 功能	2-14
2.13.3 配置本地 ND Proxy 功能.....	2-14
2.14 IPv6 邻居发现显示和维护.....	2-15

1 IPv6 基础

1.1 IPv6简介

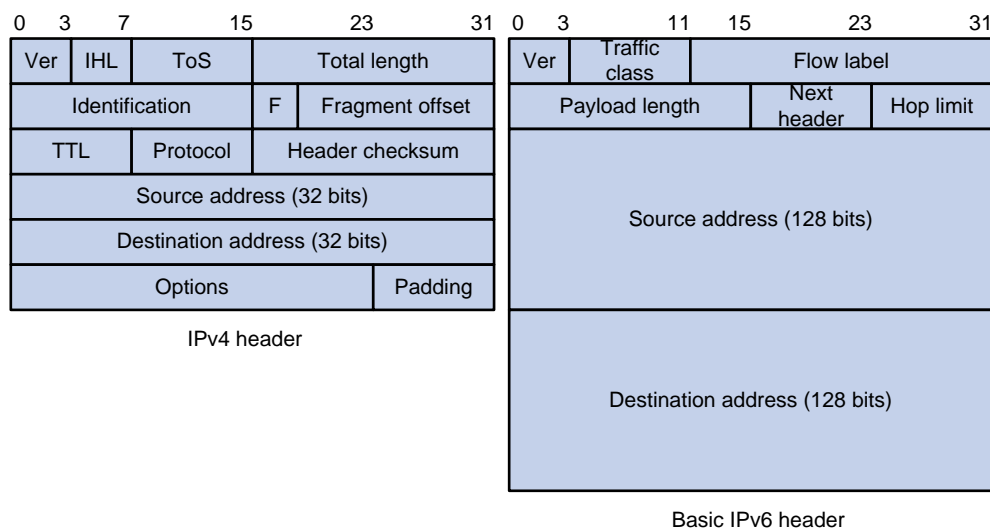
IPv6（Internet Protocol Version 6，互联网协议版本 6）是网络层协议的第二代标准协议，也被称为 IPng（IP Next Generation，下一代互联网协议），它是 IETF（Internet Engineering Task Force，互联网工程任务组）设计的一套规范，是 IPv4 的升级版本。IPv6 和 IPv4 之间最显著的区别为：IP 地址的长度从 32 比特增加到 128 比特。

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 DHCPv6 服务器）获取 IPv6 地址及相关信息，详细介绍请参见“三层技术-IP 业务配置指导”中的“DHCPv6 服务器”；
- 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。
- 如果地址中包含一组或连续多组均为 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。



说明

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，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。

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
	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。
- 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：

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

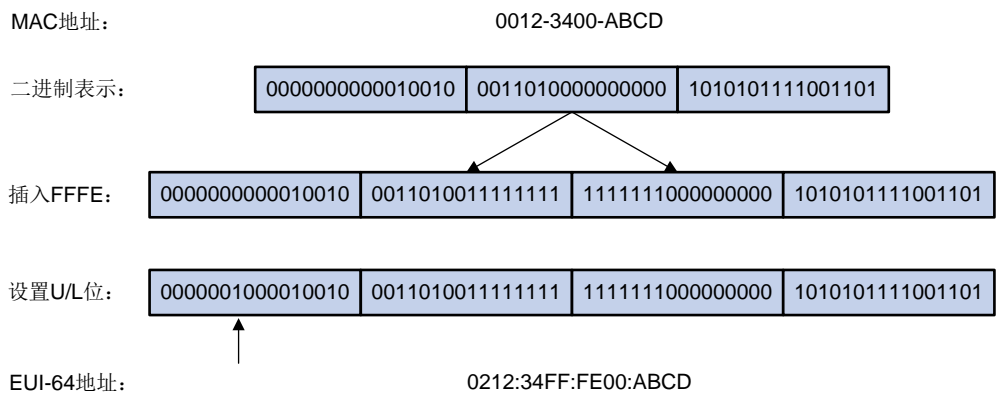
其中，FF02:0:0:0:1:FF 为 104 位固定格式；XX:XXXX 为单播或任播 IPv6 地址的后 24 位。

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唯一标识链路上的一个接口。目前 IPv6 单播地址基本上都要求接口标识符为 64 位。

对于所有 IEEE 802 接口类型（例如，VLAN 接口）的接口，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（111111111111110）。为了使接口标识符的作用范围与原 MAC 地址一致，还要将 Universal/Local (U/L)位（从高位开始的第 7 位）进行取反操作。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



对于 Tunnel 类型的接口，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“VPN 配置指导”中的“隧道”。

对于其他接口类型的接口，IEEE EUI-64 格式的接口标识符由设备随机生成。

1.1.3 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。

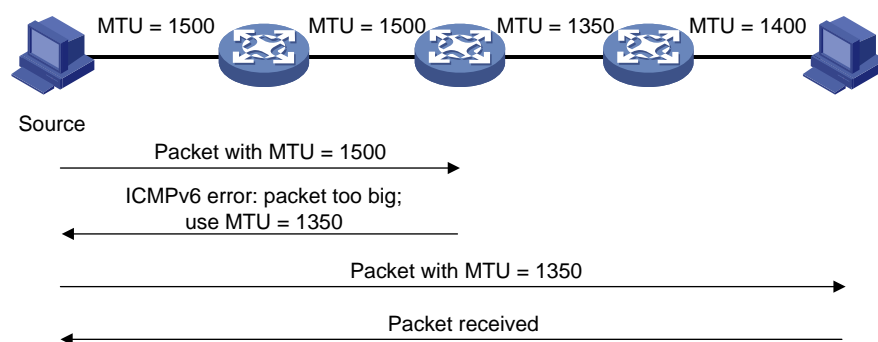
PMTU (Path MTU, 路径 MTU) 发现机制的目的就是要找到从源端到目的端的路径上最小的 MTU。

如图 1-3 所示，PMTU 的工作过程为：

- (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。

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。

图1-3 PMTU 发现工作过程



1.1.4 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 支撑着的互联网进行正常通信，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。

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“VPN 配置指导”中的“隧道”。

3. AFT

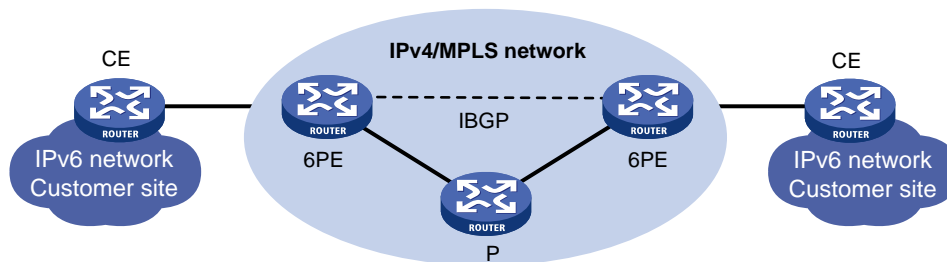
AFT（Address Family Translation，地址族转换）提供了 IPv4 和 IPv6 地址之间的相互转换功能，使 IPv4 网络和 IPv6 网络可以直接通信。AFT 作用于 IPv4 和 IPv6 网络边缘设备上，所有的地址转换过程都在该设备上实现，对 IPv4 和 IPv6 网络内的用户来说是透明的，即用户不必改变目前网络中主机的配置就可实现 IPv6 网络与 IPv4 网络的通信。有关 AFT 的详细介绍，请参见“三层技术-IP 业务配置指导”中的“AFT”。

4. 6PE

6PE 是一种过渡技术，ISP 可以利用已有的 IPv4 骨干网为分散用户的 IPv6 网络提供接入能力。

6PE 的主要思想是：6PE（IPv6 Provider Edge，IPv6 供应商边缘）路由器将用户的 IPv6 路由信息转换为带有标签的 IPv6 路由信息，并且通过 IBGP（Internal Border Gateway Protocol，内部边界网关协议）会话扩散到 ISP 的 IPv4 骨干网中。6PE 路由器转发 IPv6 报文时，首先会将进入骨干网隧道的数据流打上标签。隧道可以是 GRE 隧道或者 MPLS LSP 等。有关 6PE 的详细介绍及配置请参见“三层技术-IP 路由配置指导”中的“BGP”。

图1-4 6PE 组网图



当 ISP 想利用自己原有的 IPv4 网络，使其通过 MPLS 具有 IPv6 流量交换能力时，只需要升级 PE 路由器就可以了。所以对于运营商来说，使用 6PE 技术作为 IPv6 过渡机制无疑是一个高效的解决方案，其操作风险也会小得多。

1.1.5 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 3307: Allocation Guidelines for IPv6 Multicast Addresses
- RFC 4191: Default Router Preferences and More-Specific Routes
- RFC 4291: IP Version 6 Addressing Architecture

- RFC 4443: Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification
- RFC 4862: IPv6 Stateless Address Autoconfiguration

1.2 vSystem相关说明

非缺省 vSystem 支持本特性的部分功能，具体包括：

- 配置 IPv6 全球单播地址
- 配置 IPv6 链路本地地址
- 配置 IPv6 任播地址



说明

非缺省 vSystem 对具体命令的支持情况，请见本特性的命令参考。有关 vSystem 的详细介绍请参见“虚拟化技术配置指导”中的“vSystem”。

1.3 IPv6基础配置任务简介

IPv6 基础配置任务如下：

(1) 配置 IPv6 地址

请选择以下至少一项任务进行配置：

- [配置 IPv6 全球单播地址](#)
- [配置 IPv6 链路本地地址](#)
- [配置 IPv6 任播地址](#)

(2) （可选）[配置 PMTU 发现](#)

- [配置接口 MTU](#)
- [配置指定地址的静态 PMTU](#)
- [配置 PMTU 老化时间](#)

(3) （可选）[配置 ICMPv6 报文发送功能](#)

- [配置发送 ICMPv6 差错报文对应的令牌桶容量和令牌刷新周期](#)
- [配置允许回复组播形式的 Echo request 报文](#)
- [配置 ICMPv6 目的不可达差错报文发送功能](#)
- [配置 ICMPv6 超时差错报文发送功能](#)
- [配置 ICMPv6 重定向报文发送功能](#)
- [配置 ICMPv6 报文指定源地址功能](#)

(4) （可选）[开启 IPv6 分片报文本本地重组功能](#)

(5) （可选）[开启 IPv6 报文扩展头丢弃功能](#)

1.4 配置IPv6全球单播地址

1.4.1 功能简介

IPv6 全球单播地址可以通过下面几种方式配置：

- 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 ID 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。
- 引用前缀生成 IPv6 地址：引用前缀生成 IPv6 地址时，接口的 IPv6 地址的前缀可以通过手工配置或 DHCPv6 动态获取，同时该前缀还会分配给终端设备。
- 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。

每个接口可以有多个全球单播地址。

手工配置的全局单播地址（包括采用 EUI-64 格式形成的全局单播地址）的优先级高于自动生成的全局单播地址。如果在接口已经自动生成全局单播地址的情况下，手工配置前缀相同的全局单播地址，不会覆盖之前自动生成的全局单播地址。如果删除手工配置的全局单播地址，设备还可以使用自动生成的全局单播地址进行通信。

在 RBM 组网下可以为 IPv6 地址赋予浮动属性，IPv6 浮动地址可以简化 HA（High Availability，高可靠性）功能的配置。将 IPv6 浮动地址配置在 HA 主设备接入下行租户的接口，该地址会自动同步到备设备，不需要在主/备设备的业务接口上配置 VRRP（Virtual Router Redundancy Protocol，虚拟路由器冗余协议）虚拟地址。IPv6 浮动地址仅在主备模式的主设备上配置，不支持双主模式，不支持在备设备上配置、修改或删除。有关 HA 和 RBM 的详细介绍请参见“可靠性”中的“高可靠性”，有关 VRRP 的详细介绍请参见“可靠性”中的“VRRP”。

通过 `ipv6 address` 命令配置同样的 IPv6 地址，但不携带 `float` 参数，可以取消该地址的浮动属性。

1.4.2 采用 EUI-64 格式形成 IPv6 地址

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 采用 EUI-64 格式形成 IPv6 地址。

```
ipv6 address { ipv6-address prefix-length | ipv6-address/prefix-length }  
eui-64
```

缺省情况下，接口上未配置 IPv6 全球单播地址。

1.4.3 手工指定 IPv6 地址

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

(3) 手工指定 IPv6 地址。

```
ipv6 address { ipv6-address prefix-length | ipv6-address/prefix-length }  
[ float ]
```

缺省情况下，接口上未配置 IPv6 全球单播地址。

1.4.4 无状态自动配置 IPv6 地址

1. 功能简介

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，以太网接口、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，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。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。

在配置了优先选择临时地址功能前提下发送报文，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”。

2. 配置限制和指导

如果 RA 报文携带的前缀长度不是 64 位，则该接口自动生成 IPv6 全球单播地址失败。

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。

在接口上开启无状态地址自动配置功能后，接口通过无状态自动配置方式生成全球单播地址。如果通过 **undo ipv6 address auto** 命令关闭该功能，将删除该接口上所有自动生成的全球单播地址和链路本地地址。

3. 开启无状态自动配置 IPv6 地址功能

(1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 开启无状态地址自动配置功能，使接口通过无状态自动配置方式生成全球单播地址。

```
ipv6 address auto
```

缺省情况下，接口上无状态地址自动配置功能处于关闭状态。

4. 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

- (1) 进入系统视图。

```
system-view
```

- (2) 配置系统生成临时地址。

```
ipv6 temporary-address [ valid-lifetime preferred-lifetime ]
```

缺省情况下，系统不生成临时地址。

- (3) 优先选择临时地址作为报文的源地址。

```
ipv6 prefer temporary-address
```

缺省情况下，不会用临时地址作为接口发送报文的源地址。

1.4.5 引用前缀生成接口上的 IPv6 地址，并将此前缀分配给终端设备

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 IPv6 前缀。请选择其中一项进行配置。

- 手工配置静态的 IPv6 前缀。

```
ipv6 prefix prefix-number ipv6-prefix/prefix-length
```

缺省情况下，未配置静态 IPv6 前缀。

- 配置设备作为 DHCPv6 客户端动态获取 IPv6 前缀，并生成指定编号的 IPv6 前缀。

配置方法请参见“三层技术-IP 业务配置指导”中的“DHCPv6 客户端”。

- (3) 进入接口视图。

```
interface interface-type interface-number
```

- (4) 引用前缀生成接口上的 IPv6 地址，并将此前缀分配给终端设备。

```
ipv6 address prefix-number sub-prefix/prefix-length
```

缺省情况下，接口上未引用前缀，也不会向终端设备分配该前缀。

1.5 配置IPv6链路本地地址

1.5.1 功能简介

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。

1.5.2 配置限制和指导

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。

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。

1.5.3 配置自动生成链路本地地址

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置自动生成链路本地地址。

```
ipv6 address auto link-local
```

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。

1.5.4 手工指定接口的链路本地地址

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 手工指定接口的链路本地地址。

```
ipv6 address ipv6-address link-local
```

缺省情况下，未指定接口的链路本地地址。

1.6 配置IPv6任播地址

- (1) 进入系统视图。

```
system-view
```


- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置 IPv6 任播地址。

```
ipv6 address { ipv6-address prefix-length | ipv6-address/prefix-length }  
anycast
```

缺省情况下，接口上未配置任播地址。

1.7 配置PMTU发现

1.7.1 配置接口 MTU

1. 功能简介

由于 IPv6 路由器不支持对报文进行分片，当路由器接口收到一个报文后，如果发现报文长度比转发接口的 MTU 值大，则会将其丢弃；同时将转发接口的 MTU 值通过 ICMPv6 报文的“Packet Too Big”消息发给源端主机，源端主机以该值重新发送 IPv6 报文。为减少报文被丢弃带来的额外流量开销，需要根据实际组网环境设置合适的接口 MTU 值。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口 MTU。

```
ipv6 mtu size
```

缺省情况下，未配置接口上发送 IPv6 报文的 MTU。

1.7.2 配置指定地址的静态 PMTU

1. 功能简介

用户可以为指定的目的 IPv6 地址配置静态的 PMTU 值。当设备作为源端从接口发送报文时，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发送。发送过程中再通过“[1.1.3 IPv6 PMTU 发现](#)”中的方法动态确定设备作为源端到目的端主机的 PMTU 值。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置指定 IPv6 地址对应的静态 PMTU 值。

```
ipv6 pathmtu [ vpn-instance vpn-instance-name ] ipv6-address value
```

缺省情况下，未配置静态 PMTU 值。

1.7.3 配置 PMTU 老化时间

1. 功能简介

通过“[1.1.3 IPv6 PMTU 发现](#)”中的方法动态确定设备作为源端到目的端主机的 PMTU 后，设备将使用这个 MTU 值发送后续报文到目的端主机。当 PMTU 老化时间超时后，源端主机将通过 PMTU 机制重新确定发送报文的 MTU 值。

2. 配置限制与指导

该配置对静态 PMTU 不起作用。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 PMTU 老化时间。

```
ipv6 pathmtu age age-time
```

缺省情况下，PMTU 的老化时间是 10 分钟。

1.8 配置 ICMPv6 报文发送功能

1.8.1 配置发送 ICMPv6 差错报文对应的令牌桶容量和令牌刷新周期

1. 功能简介

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，即每隔多长时间发放一个令牌到令牌桶中，直到令牌桶中的令牌数达到配置的容量。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。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置发送 ICMPv6 差错报文对应的令牌桶容量和令牌刷新周期。

```
ipv6 icmpv6 error-interval interval [ bucketsize ]
```

缺省情况下，令牌桶容量为 10，令牌刷新周期为 100 毫秒。

刷新周期为 0 时，表示不限制 ICMPv6 差错报文的发送。

1.8.2 配置允许回复组播形式的 Echo request 报文

- (1) 进入系统视图。

```
system-view
```

- (2) 配置设备允许回复组播形式的 Echo request 报文。

```
ipv6 icmpv6 multicast-echo-reply enable
```

缺省情况下，不允许设备回复组播形式的 Echo request 报文。

1.8.3 配置 ICMPv6 目的不可达差错报文发送功能

1. 功能简介

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。

设备在满足下列任一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。

2. 配置限制和指导

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。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启设备的 ICMPv6 目的不可达报文的发送功能。

```
ipv6 unreachable enable
```

缺省情况下，ICMPv6 目的不可达报文发送功能处于关闭状态。

1.8.4 配置 ICMPv6 超时差错报文发送功能

1. 功能简介

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。

设备在满足下列任一条件时会发送 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。

如果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。

2. 配置限制和指导

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。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启设备的 ICMPv6 超时报文的发送功能。

```
ipv6 hoplimit-expires enable
```

缺省情况下，ICMPv6 超时报文发送功能处于开启状态。

1.8.5 配置 ICMPv6 重定向报文发送功能

1. 功能简介

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 重定向报文，通知主机重新选择更好的下一跳进行后续报文的发送。

同时满足下列条件时，设备会发送 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；
- 被选择的路由不是设备的缺省路由；
- 被转发的 IPv6 数据报文中不包含路由扩展头。

ICMPv6 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。但是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。因此缺省情况下设备的 ICMPv6 重定向报文发送功能处于关闭状态。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启设备的 ICMPv6 重定向报文发送功能。

```
ipv6 redirects enable
```

缺省情况下，ICMPv6 重定向报文发送功能处于关闭状态。

1.8.6 配置 ICMPv6 报文指定源地址功能

1. 功能简介

在网络中 IPv6 地址配置较多的情况下，收到 ICMPv6 报文时，用户很难根据报文的源 IPv6 地址判断报文来自哪台设备。为了简化这一判断过程，可以配置 ICMPv6 报文指定源地址功能。用可配置特定地址（如环回口地址）为 ICMPv6 报文的源地址，可以简化判断。

设备发送 ICMPv6 差错报文（TTL 超时、报文过大、端口不可达和参数错误等）和 ping echo request 报文时，都可以通过上述命令指定报文的源地址。

2. 配置限制与指导

用户发送 ping echo request 报文时，如果 ping 命令中已经指定源地址，则使用该源地址，否则使用 `ipv6 icmpv6 source` 配置的源地址。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 ICMPv6 报文指定源地址功能。

```
ipv6 icmpv6 source [ vpn-instance vpn-instance-name ] ipv6-address
```

缺省情况下，ICMPv6 报文指定源地址功能处于关闭状态。

1.9 开启 IPv6 分片报文本地重组功能

1. 功能简介

多台设备组成的 IRF 环境下，当某成员设备收到目的为本 IRF 设备的 IPv6 分片报文时，需要把分片报文送到主设备进行重组，这样会导致报文重组性能较低的问题。

当开启 IPv6 分片报文本地重组功能后，分片报文会在该成员设备上直接进行报文重组，这样就能提高分片报文的重组性能。

2. 配置限制与指导

开启 IPv6 分片报文本地重组功能后，如果分片报文是从 IRF 系统中不同的成员设备进入的，会导致 IPv6 分片报文本地无法重组成功。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启设备的 IPv6 分片报文本地重组功能。

```
ipv6 reassemble local enable
```

缺省情况下，IPv6 分片报文本地重组功能处于关闭状态。

1.10 开启IPv6报文扩展头丢弃功能

1. 功能简介

IPv6 协议引入了多种扩展报文头，开启 IPv6 扩展报文丢弃功能后，如果接收到无法处理的 IPv6 扩展头的报文，设备将直接丢弃。

2. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 开启 IPv6 报文扩展头丢弃功能。

```
ipv6 extension-header drop enable
```

缺省情况下，IPv6 报文扩展头丢弃功能处于关闭状态。

1.11 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。

display tcp statistics、**display udp statistics**、**reset tcp statistics** 和 **reset udp statistics** 命令的详细介绍请参见“三层技术-IP 业务命令参考”中的“IP 性能优化”。



说明

非缺省 vSystem 不支持部分显示和维护命令，具体情况请参见本特性的命令参考。

表1-3 IPv6 基础显示和维护

操作	命令
显示IPv6浮动地址详细信息	<pre>display ipv6 address float [interface interface-type interface-number [vpn-instance vpn-instance-name] [ipv6-address]] （独立运行模式） display ipv6 address float [interface interface-type interface-number [vpn-instance vpn-instance-name] [ipv6-address]] [slot slot-number [cpu cpu-number]] （IRF模式） display ipv6 address float [interface interface-type interface-number [vpn-instance vpn-instance-name] [ipv6-address]] [chassis chassis-number slot slot-number [cpu cpu-number]]</pre>
显示IPv6 FIB信息	<pre>（独立运行模式） display ipv6 fib [vpn-instance vpn-instance-name] [ipv6-address [prefix-length]] [slot slot-number [cpu cpu-number]]</pre>

	(IRF模式) display ipv6 fib [vpn-instance vpn-instance-name] [ipv6-address [prefix-length]] [chassis chassis-number slot slot-number [cpu cpu-number]]
显示IPv6 ICMP流量统计信息	(独立运行模式) display ipv6 icmp statistics [slot slot-number [cpu cpu-number]] (IRF模式) display ipv6 icmp statistics [chassis chassis-number slot slot-number [cpu cpu-number]]
显示接口的IPv6信息	display ipv6 interface [interface-type [interface-number]] [brief]
显示接口的IPv6前缀信息	display ipv6 interface interface-type interface-number prefix
显示IPv6的PMTU信息	display ipv6 pathmtu [vpn-instance vpn-instance-name] { ipv6-address { all dynamic static } [count] }
显示IPv6前缀信息	display ipv6 prefix [prefix-number]
显示IPv6 RawIP连接摘要信息	(独立运行模式) display ipv6 rawip [slot slot-number [cpu cpu-number]] (IRF模式) display ipv6 rawip [chassis chassis-number slot slot-number [cpu cpu-number]]
显示IPv6 RawIP连接详细信息	(独立运行模式) display ipv6 rawip verbose [slot slot-number [cpu cpu-number] [pcb pcb-index]] (IRF模式) display ipv6 rawip verbose [chassis chassis-number slot slot-number [cpu cpu-number] [pcb pcb-index]]
显示IPv6报文及ICMPv6报文的统计信息	(独立运行模式) display ipv6 statistics [slot slot-number [cpu cpu-number]] (IRF模式) display ipv6 statistics [chassis chassis-number slot slot-number [cpu cpu-number]]
显示IPv6 TCP连接摘要信息	(独立运行模式) display ipv6 tcp [slot slot-number [cpu cpu-number]] (IRF模式) display ipv6 tcp [chassis chassis-number slot slot-number [cpu cpu-number]]
显示IPv6 TCP连接详细信息	(独立运行模式) display ipv6 tcp verbose [slot slot-number [cpu cpu-number] [pcb pcb-index]] (IRF模式) display ipv6 tcp verbose [chassis chassis-number slot slot-number [cpu cpu-number] [pcb

	<code>pcb-index]]</code>
显示IPv6 TCP代理连接的简要信息	<p>(独立运行模式)</p> <p>display ipv6 tcp-proxy slot slot-number [cpu cpu-number]</p> <p>(IRF模式)</p> <p>display ipv6 tcp-proxy chassis chassis-number slot slot-number [cpu cpu-number]</p>
显示IPv6 TCP代理非保留端口的使用信息	<p>(独立运行模式)</p> <p>display ipv6 tcp-proxy port-info slot slot-number [cpu cpu-number]</p> <p>(IRF模式)</p> <p>display ipv6 tcp-proxy port-info chassis chassis-number slot slot-number [cpu cpu-number]</p>
显示IPv6 UDP连接摘要信息	<p>(独立运行模式)</p> <p>display ipv6 udp [slot slot-number [cpu cpu-number]]</p> <p>(IRF模式)</p> <p>display ipv6 udp [chassis chassis-number slot slot-number [cpu cpu-number]]</p>
显示IPv6 UDP连接详细信息	<p>(独立运行模式)</p> <p>display ipv6 udp verbose [slot slot-number [cpu cpu-number] [pcb pcb-index]]</p> <p>(IRF模式)</p> <p>display ipv6 udp verbose [chassis chassis-number slot slot-number [cpu cpu-number] [pcb pcb-index]]</p>
显示IPv6 TCP连接的流量统计信息	<p>(独立运行模式)</p> <p>display tcp statistics [slot slot-number [cpu cpu-number]]</p> <p>(IRF模式)</p> <p>display tcp statistics [chassis chassis-number slot slot-number [cpu cpu-number]]</p>
显示IPv6 UDP流量统计信息	<p>(独立运行模式)</p> <p>display udp statistics [slot slot-number [cpu cpu-number]]</p> <p>(IRF模式)</p> <p>display udp statistics [chassis chassis-number slot slot-number [cpu cpu-number]]</p>
清除PMTU值	reset ipv6 pathmtu { all dynamic static }
清除IPv6报文及ICMPv6报文的统计信息	<p>(独立运行模式)</p> <p>reset ipv6 statistics [slot slot-number [cpu cpu-number]]</p> <p>(IRF模式)</p> <p>reset ipv6 statistics [chassis chassis-number slot slot-number [cpu cpu-number]]</p>
清除IPv6 TCP连接的流量统计信息	reset tcp statistics
清除IPv6 UDP流量统计信息	reset udp statistics

2 IPv6 邻居发现

2.1 IPv6邻居发现简介

2.1.1 IPv6 邻居发现使用的 ICMPv6 消息

IPv6 ND（IPv6 Neighbor Discovery，IPv6 邻居发现）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。

邻居发现协议使用的 ICMPv6 消息的类型及作用如[表 2-1](#)所示。

表2-1 邻居发现协议使用的 ICMPv6 消息类型及作用

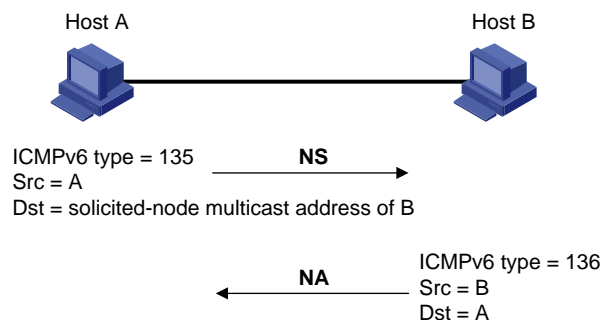
ICMPv6 消息	类型号	作用
邻居请求消息NS（Neighbor Solicitation）	135	获取邻居的链路层地址
		验证邻居是否可达
		进行重复地址检测
邻居通告消息NA（Neighbor Advertisement）	136	对NS消息进行响应
		节点在链路层变化时主动发送NA消息，向邻居节点通告本节点的变化信息
路由器请求消息RS（Router Solicitation）	133	节点启动后，通过RS消息向路由器发出请求，请求前缀和其他配置信息，用于节点的自动配置
路由器通告消息RA（Router Advertisement）	134	对RS消息进行响应
		在没有抑制RA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

2.1.1 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 2-1](#)所示，节点 A 要获取节点 B 的链路层地址的过程为：

- (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 的链路层地址和请求的目标地址。
- (2) 节点 B 收到 NS 消息后，判断报文的目标地址是否为自己的 IPv6 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。

图2-1 地址解析示意图



2.1.2 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。

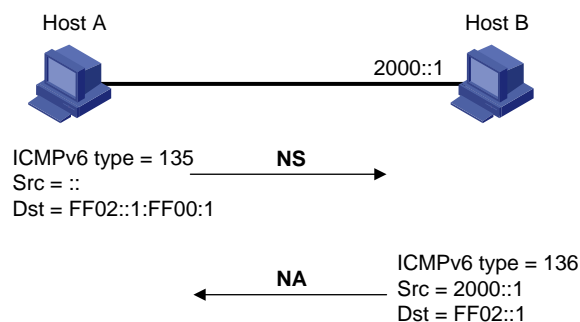
- (1) 节点发送 **NS** 消息，其中目的地址是邻居节点的 **IPv6** 地址。
- (2) 如果收到邻居节点的确认报文，则认为邻居可达；否则，认为邻居不可达。

2.1.3 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）。如图 2-2 所示，通过 **NS** 和 **NA** 实现重复地址检测的过程为：

- (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。

图2-2 重复地址检测示意图



2.1.4 路由器发现/前缀发现及地址无状态自动配置

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。

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。

- 有效生命期：表示前缀有效期。在有效生命期内，通过该前缀自动生成的地址可以正常使用；有效生命期过期后，通过该前缀自动生成的地址变为无效，将被删除。
- 首选生命期：表示首选通过该前缀无状态自动配置地址的时间。首选生命期过期后，节点通过该前缀自动配置的地址将被废止。节点不能使用被废止的地址建立新的连接，但是仍可以接收目的地址为被废止地址的报文。首选生命期必须小于或等于有效生命期。

2.1.5 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）。

同时满足下列条件时，设备会发送 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；
- 被选择的路由不是设备的缺省路由；
- 被转发的 IPv6 数据报文中不包含路由扩展头。

2.1.6 协议规范

与 IPv6 邻居发现相关的协议规范有：

- RFC 4861: Neighbor Discovery for IP Version 6 (IPv6)
- RFC 8106: IPv6 Router Advertisement Options for DNS Configuration

2.2 vSystem相关说明

非缺省 vSystem 支持本特性的部分功能，具体包括：

- 配置静态邻居表项
- 配置 RA 消息中的前缀信息
- 配置通过 RA 消息发布的前缀使用的缺省参数



说明

非缺省 vSystem 对具体命令的支持情况，请见本特性的命令参考。有关 vSystem 的详细介绍请参见“虚拟化技术配置指导”中的“vSystem”。

2.3 IPv6邻居发现配置任务简介

本节中的所有配置均为可选，请根据实际情况选择配置。

- [配置静态邻居表项](#)
- [配置接口上允许动态学习的邻居的最大个数](#)
- [开启接口从未经请求的 NA 报文中学习邻居信息的功能](#)
- [配置 STALE 状态 ND 表项的老化时间](#)
- [配置链路本地 ND 表项资源占用最小化](#)
- [配置设备的跳数限制](#)
- [配置允许发布 RA 消息及相关参数](#)
- [配置 RA 消息中的 DNS 服务器信息](#)
- [配置 RA 消息中的 DNS 域名后缀信息](#)
- [开启 RA 消息中的 DNS 信息抑制功能](#)
- [配置重复地址检测时发送邻居请求消息的次数](#)
- [配置 IPv6 地址冲突自恢复功能](#)
- [配置 ND Proxy 功能](#)

2.4 配置静态邻居表项

1. 功能简介

邻居表项保存的是设备在链路范围内的邻居信息，设备邻居表项可以通过邻居请求消息 NS 及邻居通告消息 NA 来动态创建，也可以通过手工配置来静态创建。

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相连的邻居节点的 IPv6 地址和链路层地址；
- 配置本节点 VLAN 中的二层端口相连的邻居节点的 IPv6 地址和链路层地址。

2. 配置限制与指导

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，设备还需要解析该 VLAN 下的二层端口信息。
- 采用第二种方式配置静态邻居表项后，需要保证 *port-type port-number* 指定的二层端口属于 *vlan-id* 指定的 VLAN，且该 VLAN 已经创建了 VLAN 接口。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。

当以太网冗余接口的成员接口包含子接口时，不能指定该以太网冗余接口为 IPv6 静态邻居表项所对应的接口。关于以太网冗余接口的详细介绍，请参见“虚拟化配置指导”中的“冗余备份”。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 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 ]
```

缺省情况下，未配置静态邻居表项。

2.5 配置接口上允许动态学习的邻居的最大个数

1. 功能简介

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。为了防止部分接口下的用户占用过多的资源，可以通过设置接口学习动态邻居表项的最大个数来进行限制。当接口学习到的动态邻居表项的个数达到所设置的最大值时，该接口将不再学习动态邻居表项。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口上允许学习的动态邻居表项的最大个数。

```
ipv6 neighbors max-learning-num max-number
```

缺省情况下,接口上允许学习的动态邻居表项的最大个数请参考命令手册中对应的描述。

2.6 开启接口从未经请求的NA报文中学习邻居信息的功能

1. 功能简介

在一些组网中，为了保证网络的冗余链路备份，服务器会发送 NA 报文给两台对端设备，此时因为不关注对端的地址，所以发送的是组播 NA，设备无法从组播 NA 中学习 ND 表项。所以在没有服务器对端到服务器的 IPv6 流量，且服务器没有逐个发送单播给对端设备的情况下，对端设备无法学习到服务器的 ND 表项。

2. 配置限制和指导

配置本功能后，接口可以从未经请求而收到的 NA 报文中学习邻居信息，学习到的表项状态为 STALE，因此建议本功能仅在信任网络内开启。

为防止设备上学习过多的 ND 表项，占用系统资源，建议开启本功能前，通过 **ipv6 neighbor stale-aging** 命令配置较小的 STALE 老化时间，加速老化无效的 ND 表项。

本功能仅支持在三层接口视图下开启。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入三层接口视图。

```
interface interface-type interface-number
```

- (3) 开启接口从未经请求的 NA 报文中学习邻居信息的功能。

```
ipv6 nd unsolicited-na-learning enable
```

缺省情况下，接口不学习未经请求的 NA 报文中的邻居信息到 ND 表项中。

2.7 配置STALE状态ND表项的老化时间

1. 功能简介

为适应网络的变化，ND 表需要不断更新。在 ND 表中，处于 STALE 状态的 ND 表项并非永远有效，而是有一个老化时间。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，并且设备会发送 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，此时设备会将该 ND 表项删除。用户可以根据网络实际情况调整老化时间。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 STALE 状态 ND 表项的老化时间。

```
ipv6 neighbor stale-aging aging-time
```

缺省情况下，STALE 状态 ND 表项的老化时间为 240 分钟。

2.8 配置链路本地ND表项资源占用最小化

1. 功能简介

本功能可以对链路本地 ND 表项（该 ND 表项的 IPv6 地址为链路本地地址）占用的资源进行优化。缺省情况下，所有 ND 表项均会下发硬件表项。配置本功能后，新学习的、未被引用的链路本地 ND 表项（该 ND 表项的链路本地地址不是某条路由的下一跳）不下发硬件表项，以节省资源。本功能只对后续新学习的 ND 表项生效，已经存在的 ND 表项不受影响。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置链路本地 ND 表项资源占用最小化。

```
ipv6 neighbor link-local minimize
```

缺省情况下，所有 ND 表项均会下发硬件表项。

2.9 配置设备的跳数限制

1. 功能简介

本功能可以对设备发送的 IPv6 数据报文的跳数（即 IPv6 数据报文的 Hop Limit 字段的值）进行配置。

2. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 配置设备的跳数限制。

```
ipv6 hop-limit value
```

缺省情况下，设备的跳数限制为 64 跳。

2.10 配置允许发布RA消息及相关参数

2.10.1 RA 消息及相关参数介绍

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 2-2 所示。

表2-2 RA 消息中的参数及描述

参数	描述
跳数限制（Hop Limit）	在RA消息中发布本设备的跳数限制，收到该RA消息之后，主机在发送IPv6报文时，将使用该跳数值填充IPv6报文头中的Hop Limit字段。
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DHCPv6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DHCPv6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。发布RA消息中路由器生存时间为0的路由器不能作为默认路由器。
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。

参数	描述
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。
配置路由优先级 (Router Preference)	用于设置发布RA消息的路由器的路由器优先级，主机根据接收到的RA消息中的路由器优先级，可以选择优先级最高的路由器作为默认网关。在路由器的优先级相同的情况下，遵循“先来先用”的原则，优先选择先接收到的RA消息对应的发送路由器作为默认网关。
DNS服务器信息 (DNS Server)	用于设置RA消息中的DNS服务器信息。主机通过接收到的RA消息便能够获取DNS服务器信息，不需再通过DHCPv6方式获取
DNS域名后缀信息 (DNS Search List)	用于设置RA消息中的DNS域名后缀信息。主机通过接收到的RA消息便能够获取DNS域名后缀信息，不需再通过DHCPv6方式获取

2.10.2 配置限制和指导

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。

2.10.3 配置允许发布 RA 消息

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 取消对 RA 消息发布的抑制。

```
undo ipv6 nd ra halt
```

缺省情况下，抑制发布 RA 消息。

- (4) 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
```

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。

接口将在最大时间间隔与最小时间间隔之间随机选取一个值来发布 RA 消息。

配置的最小时间间隔应该小于或等于最大时间间隔的 0.75 倍。

2.10.4 配置 RA 消息中的相关参数

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```


- (3) 配置 RA 消息中的前缀信息。

```
ipv6 nd ra prefix { ipv6-prefix prefix-length |  
ipv6-prefix/prefix-length } [ valid-lifetime preferred-lifetime  
[ no-autoconfig | off-link ] * | no-advertise ]
```

缺省情况下，未配置 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，其手工配置地址的有效生命期是 2592000 秒（30 天），首选生命期是 604800（7 天）；其他自动分配地址（如 DHCPv6 分配地址）的有效生命期和首选生命期与地址本身的生命期相同。

- (4) 配置通过 RA 消息发布的前缀使用的缺省参数。

```
ipv6 nd ra prefix default [ valid-lifetime preferred-lifetime  
[ no-autoconfig | off-link ] * | no-advertise ]
```

缺省情况下，未配置通过 RA 消息发布的前缀使用的缺省参数。

- (5) 配置 RA 消息中不携带 MTU 选项。

```
ipv6 nd ra no-advlinkmtu
```

缺省情况下，RA 消息中携带 MTU 选项。

- (6) 配置 RA 消息中不指定跳数限制。

```
ipv6 nd ra hop-limit unspecified
```

缺省情况下，RA 消息中发布本设备的跳数限制，本设备的跳数限制默认为 64 跳。

- (7) 设置被管理地址配置标志位为 1。

```
ipv6 nd autoconfig managed-address-flag
```

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。

- (8) 设置其他配置标志位为 1。

```
ipv6 nd autoconfig other-flag
```

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。

- (9) 配置 RA 消息中路由器的生存时间。

```
ipv6 nd ra router-lifetime time
```

缺省情况下，RA 消息中路由器的生存时间为 RA 消息发布的最大时间间隔的 3 倍。

- (10) 配置邻居请求消息重传时间间隔。

```
ipv6 nd ns retrans-timer value
```

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。

- (11) 配置 RA 消息中路由器的优先级。

```
ipv6 nd router-preference { high | low | medium }
```

缺省情况下，RA 消息中路由器的优先级为 medium。

- (12) 配置保持邻居可达状态的时间。

```
ipv6 nd nud reachable-time time
```

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。

2.10.5 配置 RA 消息中的 DNS 服务器信息

1. 功能简介

RA 消息中携带 DNS 服务器选项，可为主机提供 DNS 服务器信息。当主机通过无状态地址自动配置获取 IPv6 地址的同时，能够获取 DNS 服务器信息，不需再通过 DHCPv6 方式获取。

一条 DNS 服务器信息对应 RA 消息中的一个 DNS 服务器选项。选项按照配置序列号进行排列，序列号小的排在前面。

执行 **ipv6 nd ra dns server** 命令后，立即发送 RA 报文，报文中携带新配置的以及之前配置的 DNS 服务器信息。执行 **undo ipv6 nd ra dns server** 命令后，立即发送两个 RA 报文，第一个报文携带所有 DNS 服务器信息，包括被删除的 DNS 服务器，其生存时间为 0；第二个报文携带剩余的 DNS 服务器信息。发送 RA 报文后，会立即刷新接口下 RA 消息发布的时间间隔。

2. 配置限制和指导

同一个接口下最多可配置 8 条 DNS 服务器信息。

如果没有配置 DNS 服务器的生存时间，也未指定 DNS 服务器的生存时间为无限期，则 DNS 服务器的生存时间为 3*RA 消息发布的最大时间间隔，RA 消息发布的最大时间间隔可以通过 **ipv6 nd ra interval** 命令配置。

对于 PPP ND RA 应用场景以及 IPoE 中的 IPv6 ND RS 接入用户，DNS 服务器地址可以通过 AAA 授权获取，此时，也是通过 RA 报文将 IPv6 DNS 服务器地址分配给终端用户。若 AAA 为终端用户授权了 IPv6 DNS 服务器地址，也使用本命令配置了 DNS 服务器地址，则在 RA 报文中 AAA 授权的 IPv6 DNS 服务器选项排在最前面，和本命令配置的 DNS 服务器选项一起通过 RA 报文发送出去；如果 AAA 授权的 DNS 服务器地址和本命令配置的 DNS 服务器地址冲突，则以 AAA 授权为准。

关于 PPP 支持 IPv6 的详细介绍请参见“PPP 和 PPPoE”中的“PPP”。

关于 IPoE 中的 IPv6 ND RS 接入用户的详细介绍请参见“安全配置指导”中的“IPoE”。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置 RA 消息中的 DNS 服务器信息。

```
ipv6 nd ra dns server ipv6-address [ seconds | infinite ] sequence seqno
```

缺省情况下，未配置 RA 消息中的 DNS 服务器信息，RA 消息中不带 DNS 服务器信息。

2.10.6 配置 RA 消息中的 DNS 域名后缀信息

1. 功能简介

RA 消息中携带 DNSSL（DNS Search List，DNS 查询列表）选项，可为主机提供域名后缀信息。当主机通过无状态地址自动配置获取 IPv6 地址的同时，能够获取域名后缀信息，不需再通过 DHCPv6 方式获取。

一条 DNS 域名后缀信息对应 RA 消息中的一个 DNSSL 选项。选项按照配置序列号进行排列，序列号小的排在前面。

执行 **ipv6 nd ra dns search-list** 命令后，立即发送 RA 报文，报文中携带新配置的以及之前配置的 DNS 域名后缀信息。执行 **undo ipv6 nd ra dns search-list** 命令后，立即发送两个 RA 报文，第一个报文携带所有域名后缀信息，包括被删除的 DNS 域名后缀，其生存时间为 0；第二个报文携带剩余的域名后缀信息。发送 RA 报文后，会立即刷新接口下 RA 消息发布的时间间隔。

2. 配置限制和指导

同一个接口下最多可配置 8 条 DNS 域名后缀信息。

如果没有配置 DNS 域名后缀的生存时间，也未指定 DNS 域名后缀的生存时间为无限期，则 DNS 域名后缀的生存时间为 3*RA 消息发布的最大时间间隔，RA 消息发布的最大时间间隔可以通过 **ipv6 nd ra interval** 命令配置。

3. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 进入接口视图。

```
interface interface-type interface-number
```

(3) 配置 RA 消息中的 DNS 域名后缀信息。

```
ipv6 nd ra dns search-list domain-name [ seconds | infinite ] sequence  
seqno
```

缺省情况下，未配置 RA 消息中的 DNS 域名后缀信息，RA 消息中不带 DNS 域名后缀信息。

2.10.7 开启 RA 消息中的 DNS 信息抑制功能

1. 功能简介

配置 RA 消息中的 DNS 信息（包括 DNS 服务器信息和 DNS 域名后缀信息）后，如果希望 RA 消息不发布 DNS 信息，可以开启相应的抑制功能。

开启 RA 消息中的 DNS 服务器信息抑制功能后：

- 若接口下已经配置了 DNS 服务器信息或者已经通过 AAA 授权了 DNS 服务器地址，则立即发送两个 RA 报文，第一个报文携带 DNS 服务器的生存时间为 0，第二个报文不携带 DNS 服务器信息。
- 若接口下未配置 DNS 服务器信息且未通过 AAA 授权 DNS 服务器地址，则不发送 RA 报文。
- 若接口下配置了新的 DNS 服务器信息或者删除了一条 DNS 服务器信息，则立即发送 RA 报文，但是不携带任何 DNS 服务器信息。

关闭 RA 消息中的 DNS 服务器信息抑制功能后：

- 若接口下已经配置了 DNS 服务器信息或者已经通过 AAA 授权了 DNS 服务器地址，则立即发送 RA 报文，携带 DNS 服务器信息。
- 若接口下未配置 DNS 服务器信息且未通过 AAA 授权 DNS 服务器地址，则不发送 RA 报文。

发送 RA 报文后，会立即刷新接口下 RA 消息发布的时间间隔。

开启、关闭 RA 消息中的 DNS 域名后缀信息抑制功能后，RA 报文的发送逻辑和上述开启、关闭 RA 消息中的 DNS 服务器信息抑制功能后的 RA 报文发送逻辑相同，此处不再赘述。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 开启 RA 消息中的 DNS 服务器信息抑制功能。

```
ipv6 nd ra dns server suppress
```

缺省情况下，RA 消息中的 DNS 服务器信息抑制功能处于关闭状态。

- (4) 开启 RA 消息中的 DNS 域名后缀信息抑制功能。

```
ipv6 nd ra dns search-list suppress
```

缺省情况下，RA 消息中的 DNS 域名后缀信息抑制功能处于关闭状态。

2.11 配置重复地址检测时发送邻居请求消息的次数

1. 功能简介

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，当发送的次数达到使用 **ipv6 nd dad attempts** 命令所设置的次数后，仍未收到响应，则认为该地址可用。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置重复地址检测时发送邻居请求消息的次数。

```
ipv6 nd dad attempts times
```

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 *times* 值为 0 时，表示禁止重复地址检测。

2.12 配置IPv6地址冲突自恢复功能

1. 功能简介

当接口获取到一个 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。此接口会通过 ND 协议向被检测节点发送 NS 消息，地址冲突的节点会向此接口返回 NA 消息，接口收到 NA 消息后认为此 IPv6 地址是重复的。此 IPv6 地址在这个接口上被标识为 **duplicate** 状态，无法被用于通信。

接口不会自动为被标识为 **duplicate** 状态的 IPv6 地址使用重复地址检测功能，因此即便地址不再冲突，被标识为 **duplicate** 状态的 IPv6 地址也不会自动恢复到正常状态。若开启地址冲突自恢复功能，接口在收到被检测节点的 NA 消息后，每隔一段时间会恢复重复地址检测功能，继续向被检测节点发送 NS 消息，直到不再收到 NA 消息或地址冲突自恢复功能被关闭。有关具体探测过程的详细介绍，请参见“[2.1.3 重复地址检测](#)”。

若开启了地址冲突自恢复功能且配置地址冲突自恢复的最大时间间隔为 *interval*，在接口收到被检测节点的 NA 消息后系统会等待 1~*interval* 秒之间的一个随机时间间隔，然后恢复重复地址检测功能。每次接口收到 NA 消息系统生成的随机时间间隔都可能不同，这样可以降低因同时发送 NS 消息而造成报文拥塞的几率。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 开启 IPv6 地址冲突自恢复功能。

```
ipv6 address duplicate-detect enable
```

缺省情况下，IPv6 地址冲突自恢复功能处于关闭状态。

- (3) （可选）配置 IPv6 地址冲突自恢复的最大时间间隔。

```
ipv6 address duplicate-detect interval interval
```

缺省情况下，IPv6 地址冲突自恢复的最大时间间隔为 5 秒。

2.13 配置 ND Proxy 功能

2.13.1 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代理功能的设备就可以代答该请求，回应 NA 报文，这个过程称作 ND 代理（ND Proxy）。

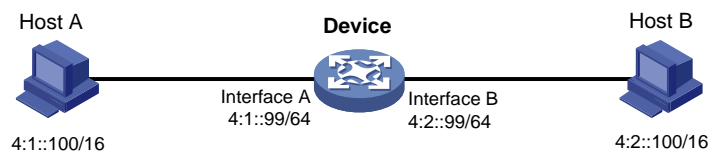
ND Proxy 功能屏蔽了分离的物理网络这一事实，使用户使用起来，好像在同一个物理网络上。

ND Proxy 功能根据应用场景不同分为普通 ND Proxy 和本地 ND Proxy。

1. 普通 ND Proxy

普通 ND Proxy 的典型应用环境如[图 2-3](#)所示。设备 Device 通过两个三层接口 Interface A 和 Interface B 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2-3 普通 ND Proxy 的应用环境



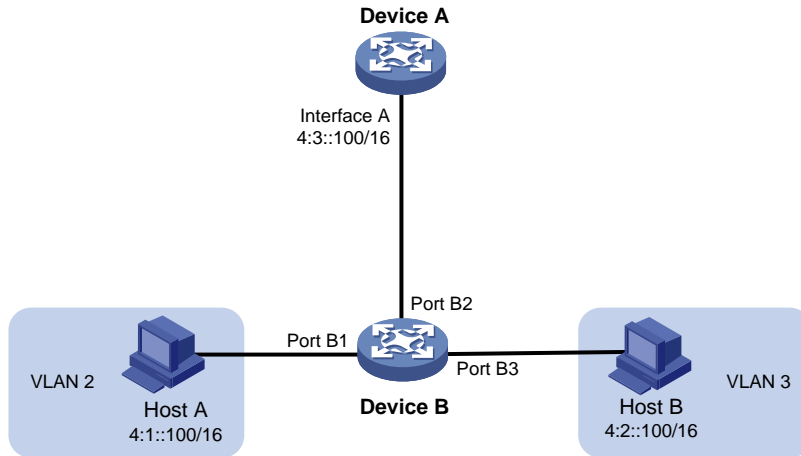
在这种组网情况下，当 Host A 需要与 Host B 通信时，由于目的 IPv6 地址与本机的 IPv6 地址为同一网段，因此 Host A 会直接发出请求 Host B 硬件地址的 NS 请求。但是，此时的两台主机处于不同的广播域中，Host B 无法收到 Host A 的 NS 请求报文，当然也就无法应答。

通过在 Device 上启用 ND Proxy 功能，可以解决此问题。在接口 Interface A 和 Interface B 上启用 ND Proxy 后，Device 可以应答 Host A 的 NS 请求。同时，Device 作为 Host B 的代理，把其它主机发送过来的报文转发给 Host B。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 2-4 所示。Host A 属于 VLAN 2，Host B 属于 VLAN 3。但它们分别连接到端口 Port B1 和 Port B3 上。

图2-4 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，因为连接两台主机属于不同的 VLAN 中，Host B 无法收到 Host A 的 NS 请求报文。

通过在 Device A 上启用本地 ND Proxy 功能，可以解决此问题。在接口 Interface A 上启用本地 ND Proxy 后，Device A 会代替 Host B 回应 NA，Host A 发给 Host B 的报文就会通过 Device A 进行转发，从而实现 Host A 与 Host B 之间的通信。

本地 ND Proxy 可以在下列四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台设备的不同 VLAN 中的端口下；
- 想要互通的主机分别连接到同一个 VLAN 中的同一个隔离组内的不同二层隔离端口下。

2.13.2 配置普通 ND Proxy 功能

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 开启普通 ND Proxy 功能。

```
proxy-nd enable
```

缺省情况下，ND Proxy 功能处于关闭状态。

2.13.3 配置本地 ND Proxy 功能

- (1) 进入系统视图。

```
system-view
```

(2) 进入接口视图。

```
interface interface-type interface-number
```

(3) 开启本地 ND Proxy 功能。

```
local-proxy-nd enable
```

缺省情况下，本地 ND Proxy 功能处于关闭状态。

2.14 IPv6邻居发现显示和维护

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。



说明

非缺省 vSystem 不支持部分显示和维护命令，具体情况请参见本特性的命令参考。

表2-3 IPv6 邻居发现显示和维护

操作	命令
显示邻居表项的个数	<p>(独立运行模式)</p> <pre>display ipv6 neighbors { { all dynamic static } [slot slot-number [cpu cpu-number]] interface interface-type interface-number vlan vlan-id } count</pre> <p>(IRF模式)</p> <pre>display ipv6 neighbors { { all dynamic static } [chassis chassis-number slot slot-number [cpu cpu-number]] interface interface-type interface-number vlan vlan-id } count</pre>
显示邻居信息	<p>(独立运行模式)</p> <pre>display ipv6 neighbors { { ipv6-address all dynamic static } [slot slot-number [cpu cpu-number]] interface interface-type interface-number vlan vlan-id } [verbose]</pre> <p>(IRF模式)</p> <pre>display ipv6 neighbors { { ipv6-address all dynamic static } [chassis chassis-number slot slot-number [cpu cpu-number]] interface interface-type interface-number vlan vlan-id } [verbose]</pre>
显示指定VPN实例的邻居信息	<pre>display ipv6 neighbors vpn-instance vpn-instance-name [count]</pre>

操作	命令
清除IPv6邻居信息	<p>(独立运行模式)</p> <pre>reset ipv6 neighbors { all dynamic interface <i>interface-type interface-number</i> slot slot-number [cpu cpu-number] static }</pre> <p>(IRF模式)</p> <pre>reset ipv6 neighbors { all dynamic interface <i>interface-type interface-number</i> chassis <i>chassis-number slot slot-number</i> [cpu cpu-number] static }</pre>

目 录

1 DHCPv6 概述	1-1
1.1 DHCPv6 的优点	1-1
1.2 DHCPv6 地址/前缀分配过程	1-1
1.2.1 交互两个消息的快速分配过程	1-1
1.2.2 交互四个消息的分配过程	1-1
1.3 地址/前缀租约更新过程	1-2
1.4 DHCPv6 无状态配置	1-3
1.5 DHCPv6 选项介绍	1-4
1.5.1 Option 18	1-4
1.5.2 Option 37	1-4
1.6 协议规范	1-5
2 DHCPv6 服务器	2-1
2.1 DHCPv6 服务器简介	2-1
2.1.1 DHCPv6 服务器应用环境	2-1
2.1.2 基本概念	2-2
2.1.3 DHCPv6 地址池	2-3
2.1.4 地址/前缀的选择优先次序	2-4
2.2 DHCPv6 服务器配置任务简介	2-4
2.3 配置为 DHCPv6 客户端分配 IPv6 前缀	2-5
2.4 配置为 DHCPv6 客户端分配 IPv6 地址	2-6
2.5 配置为 DHCPv6 客户端分配网络参数	2-8
2.5.1 功能简介	2-8
2.5.2 直接在 DHCPv6 地址池中配置网络参数	2-8
2.5.3 通过 DHCPv6 选项组配置网络参数	2-9
2.6 配置接口工作在 DHCPv6 服务器模式，并配置地址/前缀分配方式	2-10
2.7 配置 DHCPv6 策略动态分配 IPv6 地址、前缀和其他参数	2-11
2.8 配置 DHCPv6 服务器发送 DHCPv6 报文的 DSCP 优先级	2-12
2.9 配置 DHCPv6 服务器租约固化功能	2-12
2.10 配置 DHCPv6 服务器辅助路由信息	2-13
2.11 指定 DHCPv6 服务器上的地址池所属的 VPN 实例	2-14
2.12 开启 DHCPv6 服务器的日志信息功能	2-14
2.13 DHCPv6 服务器显示和维护	2-15

3 DHCPv6 中继	3-1
3.1 DHCPv6 中继简介	3-1
3.1.1 应用环境	3-1
3.1.2 DHCPv6 中继的工作过程	3-1
3.2 DHCPv6 中继配置任务简介	3-2
3.3 配置接口工作在 DHCPv6 中继模式	3-2
3.4 指定 DHCPv6 服务器的地址	3-3
3.4.1 指定 DHCPv6 中继对应的 DHCPv6 服务器地址	3-3
3.4.2 指定中继地址池上对应的 DHCPv6 服务器地址	3-3
3.5 配置 DHCPv6 中继为 DHCPv6 客户端分配的网关地址	3-4
3.6 配置 DHCPv6 中继发送 DHCPv6 报文的 DSCP 优先级	3-4
3.7 配置 DHCPv6 中继支持的 Interface ID 选项填充模式	3-5
3.8 DHCPv6 中继显示和维护	3-5
4 DHCPv6 客户端	4-1
4.1 DHCPv6 客户端简介	4-1
4.2 DHCPv6 客户端配置限制和指导	4-1
4.3 DHCPv6 客户端配置任务简介	4-1
4.4 配置接口使用的 DHCPv6 客户端 DUID	4-2
4.5 配置 DHCPv6 客户端获取 IPv6 地址和网络配置参数	4-2
4.6 配置 DHCPv6 客户端获取 IPv6 前缀和网络配置参数	4-2
4.7 配置 DHCPv6 客户端同时获取 IPv6 地址、IPv6 前缀和网络配置参数	4-3
4.8 配置 DHCPv6 客户端获取除地址/前缀外的其他网络配置参数	4-3
4.9 配置 DHCPv6 客户端发送 DHCPv6 报文的 DSCP 优先级	4-4
4.10 DHCPv6 客户端显示和维护	4-4

1 DHCPv6 概述

DHCPv6（Dynamic Host Configuration Protocol for IPv6，支持 IPv6 的动态主机配置协议）针对 IPv6 编址方案设计，用来为主机分配 IPv6 前缀、IPv6 地址和其他网络配置参数。

1.1 DHCPv6的优点

与其他 IPv6 地址分配方式（包括手工配置、通过路由器公告消息中的网络前缀无状态自动配置等，关于这两种形式的配置，请参见“三层技术-IP 业务配置指导”中的“IPv6 基础”）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。
- 为客户端分配前缀，以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、域名后缀等网络配置参数。

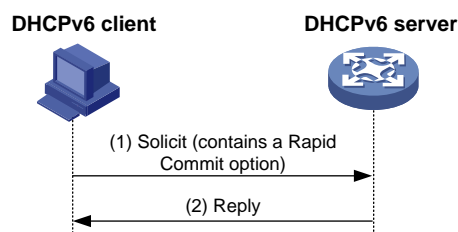
1.2 DHCPv6地址/前缀分配过程

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



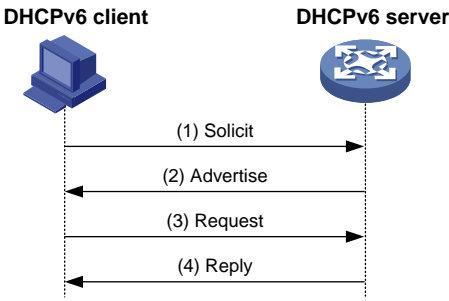
如图 1-1 所示，地址/前缀快速分配过程为：

- (1) DHCPv6 客户端在向 DHCPv6 服务器发送的 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/前缀和其他网络配置参数。
- (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](#)。

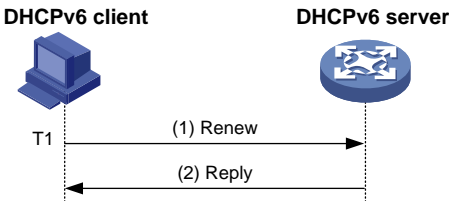
表1-1 交互四个消息的分配过程

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，该租借期限称为租约。租借期限由有效生命期决定。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，则需要申请延长地址/前缀租约。

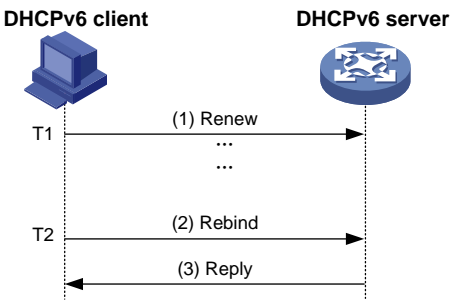
图1-3 通过 Renew 更新地址/前缀租约



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的一半）时，DHCPv6 客户端会向为它分配地址/前缀的 DHCPv6 服务器发送 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



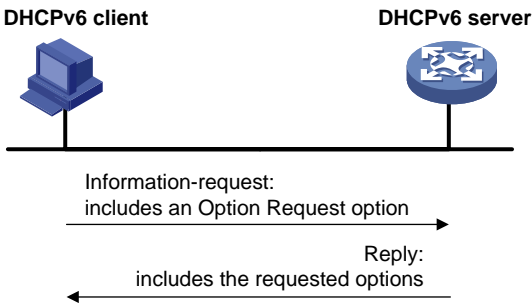
如图 1-4 所示，如果在 T1 时发送 Renew 请求更新租约，但是未收到 DHCPv6 服务器的回应报文，则 DHCPv6 客户端会在 T2（推荐值为首选生命期的 0.8 倍）时，向所有 DHCPv6 服务器组播发送 Rebind 报文请求更新租约。如果客户端可以继续使用该地址/前缀，则 DHCPv6 服务器回应续约成功的 Reply 报文，通知 DHCPv6 客户端已经成功更新地址/前缀租约；如果该地址/前缀不可以再分配给该客户端，则 DHCPv6 服务器回应续约失败的 Reply 报文，通知客户端不能获得新的租约；如果 DHCPv6 客户端未收到服务器的应答报文，则到达有效生命期后，客户端停止使用该地址/前缀。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

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，即 DHCPv6 客户端根据路由器发现/前缀发现所获取的信息自动配置 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）取值为 0、O 标志位（Other stateful configuration flag，其他配置标志位）取值为 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。

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：

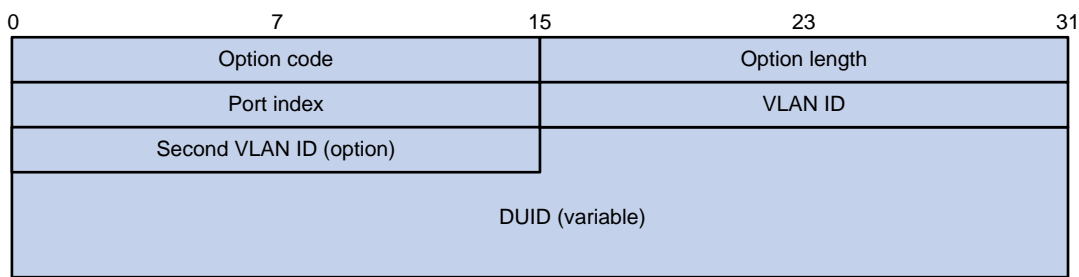
- (1) 客户端以组播的方式向 DHCPv6 服务器发送 Information-request 报文，该报文中携带 Option Request 选项，指定客户端需要从服务器获取的配置参数。
- (2) 服务器收到 Information-request 报文后，为客户端分配网络配置参数，并单播发送 Reply 报文将网络配置参数返回给客户端。
- (3) 客户端检查 Reply 报文中提供的信息，如果与 Information-request 报文中请求的配置参数相符，则按照 Reply 报文中提供的参数进行网络配置；否则，忽略该参数。如果接收到多个与请求相符的 Reply 报文，客户端将选择最先收到的 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 DHCPv6选项介绍

1.5.1 Option 18

Option 18 称为接口 ID 选项（Interface ID），设备接收到 DHCPv6 客户端发送的 DHCPv6 请求报文后，在该报文中添加 Option 18 选项，并转发给 DHCPv6 服务器。服务器可根据 Option 18 选项中的客户端信息选择合适的地址池为 DHCPv6 客户端分配 IPv6 地址。[图 1-6](#) 为 Option 18 选项格式。

图1-6 Option 18 选项格式



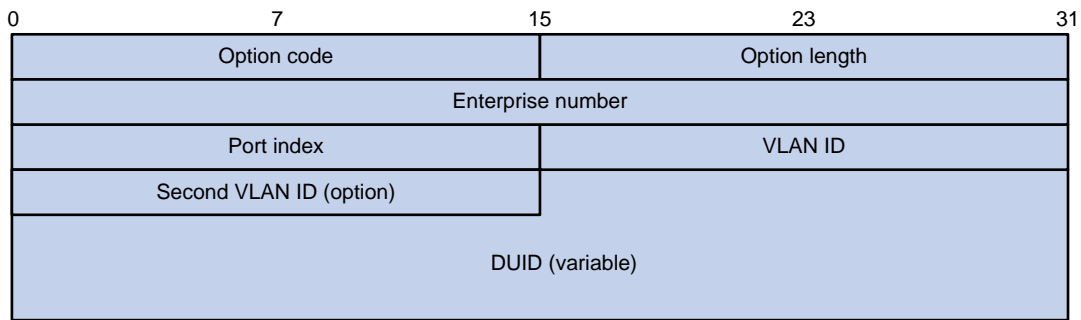
各字段的解释如下：

- Option code: Option 编号，取值为 18。
- Option length: Option 字段长度。
- Port index: DHCPv6 设备收到客户端请求报文的端口索引。
- VLAN ID: 第一层 VLAN 信息。
- Second VLAN ID: 第二层 VLAN 信息。选项格式中的 Second VLAN ID 字段为可选，如果 DHCPv6 报文中不含有 Second VLAN，则 Option 18 中也不包含 Second VLAN ID 内容。
- DUID: DHCPv6 客户端的 DUID 信息。

1.5.2 Option 37

Option 37 称为远程 ID 选项（Remote ID），设备接收到 DHCPv6 客户端发送的 DHCPv6 请求报文后，在该报文中添加 Option 37 选项，并转发给 DHCPv6 服务器。服务器可根据 Option 37 选项中的信息对 DHCPv6 客户端定位，为分配 IPv6 地址提供帮助。[图 1-7](#) 为 Option 37 选项格式。

图1-7 Option 37 选项格式



各字段的解释如下：

- Option code: Option 编号，取值为 37。
- Option length: Option 字段长度。
- Enterprise number: 企业编号。
- Port index: DHCPv6 设备收到客户端请求报文的端口索引。
- VLAN ID: 第一层 VLAN 信息。
- Second VLAN ID: 第二层 VLAN 信息。选项格式中的 Second VLAN ID 字段为可选，如果 DHCPv6 报文中不含有 Second VLAN，则 Option 37 中也不包含 Second VLAN ID 内容。
- DUID: DHCPv6 客户端的 DUID 信息。

1.6 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

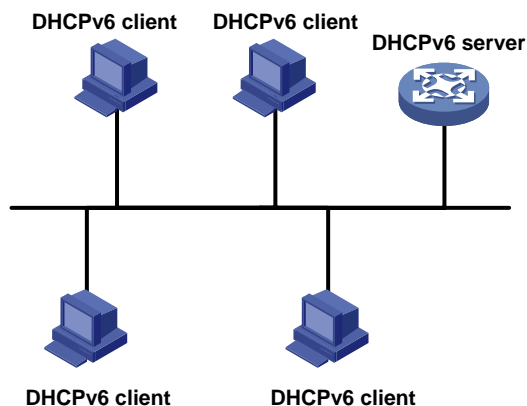
2.1 DHCPv6服务器简介

2.1.1 DHCPv6 服务器应用环境

DHCPv6 服务器可以为客户端分配 IPv6 地址/前缀和其他网络配置参数。

1. DHCPv6 服务器为客户端分配 IPv6 地址和其他网络配置参数

图2-1 DHCPv6 服务器为客户端分配 IPv6 地址和其他网络配置参数应用环境



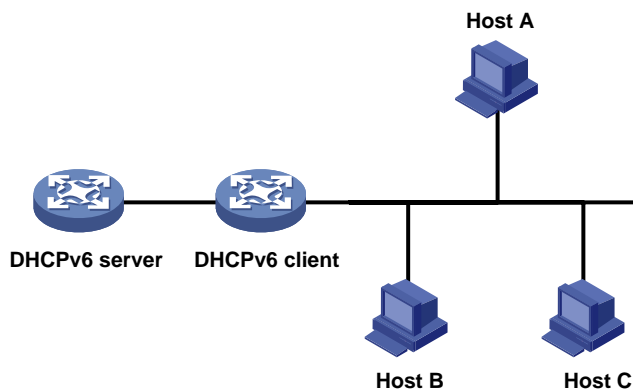
如图 2-1 所示，为了便于集中管理 IPv6 地址，简化网络配置，DHCPv6 服务器可以用来为 DHCPv6 客户端提供诸如 IPv6 地址、域名后缀、DNS 服务器地址等网络配置参数。DHCPv6 客户端根据服务器分配的参数来实现主机的配置。

DHCPv6 服务器为客户端分配的 IPv6 地址分为以下两类：

- 临时 IPv6 地址：在短期内经常变化且不用续约的地址；
- 非临时 IPv6 地址：正常使用，可以进行续约的地址。

2. DHCPv6 服务器为客户端分配 IPv6 前缀

图2-2 DHCPv6 服务器前缀分配应用组网图



如图 2-2 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，向所在网络组播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。

2.1.2 基本概念

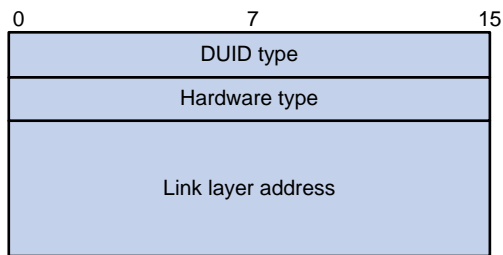
1. DHCPv6 采用的组播地址

DHCPv6 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）是一台 DHCPv6 设备（包括客户端、服务器和中继）的唯一标识。在 DHCPv6 报文交互过程中，DHCPv6 客户端、服务器和中继通过在报文中添加 DUID 来标识自己。

图2-3 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 2-3 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，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前缀绑定信息，前缀绑定信息中记录了 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 DHCPv6 地址池

每个 DHCPv6 地址池都拥有一组可供分配的 IPv6 地址、IPv6 前缀和网络配置参数。DHCPv6 服务器从地址池中为客户端选择并分配 IPv6 地址、IPv6 前缀及其他参数。

1. DHCPv6 地址池的地址管理方式

DHCPv6 地址池的地址管理方式有以下几种：

- 静态绑定 IPv6 地址：通过将客户端 DUID 和 IAID 与 IPv6 地址绑定的方式，实现为特定的客户端分配特定的 IPv6 地址；
- 动态选择 IPv6 地址：在地址池中指定可供分配的 IPv6 地址范围，当收到客户端的 IPv6 地址申请时，从该地址范围中动态选择 IPv6 地址，分配给该客户端。

在 DHCPv6 地址池中指定可供分配的 IPv6 地址范围时，需要：

- (1) 指定动态分配的 IPv6 地址网段。
- (2) 将该网段划分为非临时地址范围和临时地址范围。每个地址范围内的地址必须属于该网段，否则无法分配。

采用动态选择 IPv6 地址方式时，如果接收到客户端的地址申请，则 DHCPv6 服务器选择一个合适的地址池，并按照客户端申请的地址类型（非临时地址或临时地址），从该地址池对应的地址范围（非临时地址范围或临时地址范围）中选择合适的 IPv6 地址分配给客户端。

2. DHCPv6 地址池的前缀管理方式

DHCPv6 地址池的前缀管理方式有以下几种：

- 静态绑定 IPv6 前缀：通过将客户端 DUID 和 IAID 与 IPv6 前缀绑定的方式，实现为特定的客户端分配特定的 IPv6 前缀；
- 动态选择 IPv6 前缀：在地址池中指定可供分配的 IPv6 前缀范围，当收到客户端的 IPv6 前缀申请时，从该前缀范围中动态选择 IPv6 前缀，分配给该客户端。

在 DHCPv6 地址池中指定可供分配的 IPv6 前缀范围时，需要：

- (1) 创建前缀池，指定前缀池中包括的 IPv6 前缀范围。
- (2) 在地址池中指定动态分配的 IPv6 地址网段。
- (3) 在地址池中引用前缀池。

3. 地址池的选取原则

DHCPv6 服务器为客户端分配 IPv6 地址或前缀时，按照如下顺序选择地址池：

- (1) 如果存在将客户端 DUID、IAID 与 IPv6 地址或前缀静态绑定的地址池，则选择该地址池，并将静态绑定的 IPv6 地址或前缀、及该地址池中的网络参数分配给客户端。
- (2) 如果配置了 DHCPv6 策略，则 DHCPv6 客户端匹配某个 DHCPv6 用户类时，DHCPv6 服务器选择与该 DHCPv6 用户类关联的 DHCPv6 地址池；DHCPv6 客户端未匹配到 DHCPv6 用户类时，若配置了默认 DHCPv6 地址池，则选择该 DHCPv6 地址池；若未配置默认 DHCPv6 地址池或 DHCPv6 默认地址池不存在可供分配的 IPv6 地址或前缀时，IPv6 地址、前缀或其他参数分配失败。
- (3) 如果接收到 DHCPv6 请求报文的接口引用了某个地址池，则选择该地址池，从该地址池中选取 IPv6 地址或前缀、及网络配置参数分配给客户端。
- (4) 如果上述条件均不满足，则使用以下方法选择 DHCPv6 地址池：

- 如果客户端与服务器在同一网段，则将接收到 DHCPv6 请求报文的接口的 IPv6 地址与所有地址池配置的网段进行匹配，并选择最长匹配的网段所对应的地址池。
- 如果客户端与服务器不在同一网段，即客户端通过 DHCPv6 中继获取 IPv6 地址或前缀，则将离 DHCPv6 客户端最近的 DHCPv6 中继接口的 IPv6 地址与所有地址池配置的网段进行匹配，并选择最长匹配的网段所对应的地址池。

配置地址池动态分配的网段和 IPv6 地址范围时，请尽量保证与 DHCPv6 服务器接口或 DHCPv6 中继接口的 IPv6 地址所在的网段一致，以免分配错误的 IPv6 地址。

2.1.4 地址/前缀的选择优先次序

DHCPv6 服务器为客户端分配 IPv6 地址/前缀的优先次序如下：

- (1) DUID、IAID 与客户端 DUID、IAID 匹配，且与客户端期望地址/前缀匹配的静态绑定地址/前缀；
- (2) DUID、IAID 与客户端 DUID、IAID 匹配的静态绑定地址/前缀；
- (3) DUID 与客户端的 DUID 匹配，且与客户端期望地址/前缀匹配的静态绑定地址/前缀，该地址/前缀中未指定客户端的 IAID；
- (4) DUID 与客户端 DUID 匹配的静态绑定地址/前缀，该地址/前缀中未指定客户端的 IAID；
- (5) 地址池/前缀池中与客户端期望地址/前缀匹配的空闲地址/前缀；
- (6) 服务器记录的曾经分配给客户端的地址/前缀；
- (7) 地址池/前缀池中的其他空闲地址/前缀；
- (8) 如果未找到可用的地址/前缀，则依次查询租约过期地址/前缀、曾经发生过冲突的地址，如果找到则进行分配，否则将不予处理。

如果客户端的网段发生变化，服务器不会为客户端分配曾经分配给它的地址/前缀，而是从匹配新网段的地址池中重新选择地址/前缀等信息。



说明

使用曾经发生过冲突的 IPv6 地址时，只有冲突状态超过一小时的地址租约才能够被服务器分配给新的 DHCPv6 客户端。

2.2 DHCPv6服务器配置任务简介

DHCPv6 服务器配置任务如下：

- (1) 配置为 DHCPv6 客户端分配 IPv6 前缀、IPv6 地址和其他网络参数
请至少选择以下一项任务进行配置：
 - [配置为 DHCPv6 客户端分配 IPv6 前缀](#)
 - [配置为 DHCPv6 客户端分配 IPv6 地址](#)
 - [配置为 DHCPv6 客户端分配网络参数](#)
- (2) 修改 DHCPv6 服务器的地址池选择方式
请至少选择以下一项任务进行配置：

- [配置接口工作在 DHCPv6 服务器模式，并配置地址/前缀分配方式](#)
- [配置 DHCPv6 策略动态分配 IPv6 地址、前缀和其他参数](#)
- (3) (可选) [配置 DHCPv6 服务器发送 DHCPv6 报文的 DSCP 优先级](#)
- (4) (可选) [配置 DHCPv6 服务器租约固化功能](#)
- (5) (可选) [配置 DHCPv6 服务器辅助路由信息](#)
- (6) (可选) [指定 DHCPv6 服务器上的地址池所属的 VPN 实例](#)
- (7) (可选) [开启 DHCPv6 服务器的日志信息功能](#)

2.3 配置为DHCPv6客户端分配IPv6前缀

1. 功能简介

可以通过以下两种方式配置 DHCPv6 服务器为 DHCPv6 客户端分配 IPv6 前缀：

- 在地址池中配置静态绑定前缀：指定 DUID、IAID 及前缀的静态绑定关系后，如果 DHCPv6 请求报文中的 DUID、IAID 与静态绑定的 DUID、IAID 都相同，则将静态绑定的前缀分配给此 DHCPv6 客户端。如果只指定了 DUID 和前缀的绑定关系，未指定静态绑定的 IAID，则只要请求报文中的 DUID 与静态绑定的 DUID 相同，就将静态绑定的前缀分配给此 DHCPv6 客户端。
- 在地址池中引用包含一定前缀范围的前缀池：接收到 DHCPv6 客户端的前缀分配请求后，DHCPv6 服务器从前缀范围中动态选择可用前缀，分配给客户端。

在实际组网中，某些前缀是保留前缀，不应该动态分配给客户端。通过配置不参与自动分配的前缀，可以避免 DHCPv6 服务器分配这些前缀。

2. 配置限制和指导

配置为 DHCPv6 客户端分配 IPv6 前缀时，需要注意：

- 一个 IPv6 前缀只能与一个客户端绑定。不允许通过重复执行 **static-bind prefix** 命令的方式修改 IPv6 前缀与客户端的绑定关系、前缀的首选生命期和有效生命期。只有删除该 IPv6 前缀的静态绑定配置后，才能将该 IPv6 前缀与其他客户端绑定，或修改前缀的首选生命期和有效生命期。
- 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。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) (可选) 配置不参与自动分配的 IPv6 前缀。

```
ipv6 dhcp server forbidden-prefix start-prefix/prefix-len  
[ end-prefix/prefix-len ] [ vpn-instance vpn-instance-name ]
```

缺省情况下，DHCPv6 前缀池中的所有 IPv6 前缀都参与自动分配。

如果通过 **ipv6 dhcp server forbidden-prefix** 命令将已经静态绑定的 IPv6 前缀配置为不参与自动分配的前缀，则该前缀仍然可以分配给静态绑定的用户。

(3) 创建前缀池。

```
ipv6 dhcp prefix-pool prefix-pool-number prefix { prefix-number |  
prefix/prefix-len } assign-len assign-len [ vpn-instance  
vpn-instance-name ]
```

仅在 DHCPv6 服务器为 DHCPv6 客户端动态分配 IPv6 前缀时需要进行本配置。

当配置前缀池引用前缀编号时，必须保证指定前缀号有对应的生效前缀，否则配置不生效。

(4) 创建 DHCPv6 地址池，并进入 DHCPv6 地址池视图。

```
ipv6 dhcp pool pool-name
```

(5) 配置动态分配的 IPv6 地址网段。

```
network { prefix/prefix-length | prefix prefix-number  
[ sub-prefix/sub-prefix-length ] } [ preferred-lifetime  
preferred-lifetime valid-lifetime valid-lifetime ]
```

缺省情况下，未配置动态分配的 IPv6 地址网段。

不能在不同地址池下使用 **network** 命令配置相同的地址网段。

不能在不同地址池下引用完全一致的前缀编号、子前缀和子前缀长度。

(6) 配置地址池引用前缀信息。请至少选择其中一项进行配置。

○ 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 ] [ description description-text ]
```

缺省情况下，未配置地址池的静态绑定前缀。

重复执行 **static-bind prefix** 命令，可以配置多个静态绑定的 IPv6 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 ]
```

缺省情况下，未配置可动态分配的前缀。

2.4 配置为DHCPv6客户端分配IPv6地址

1. 功能简介

可以通过以下两种方式配置 DHCPv6 服务器为 DHCPv6 客户端分配 IPv6 地址：

- 在地址池中配置静态绑定地址：指定 DUID、IAID 及地址的静态绑定关系后，如果 DHCPv6 请求报文中的 DUID、IAID 与静态绑定的 DUID、IAID 都相同，则将静态绑定的地址分配给此 DHCPv6 客户端。如果只指定了 DUID 和地址的绑定关系，未指定静态绑定的 IAID，则只要请求报文中的 DUID 与静态绑定的 DUID 相同，就将静态绑定的地址分配给此 DHCPv6 客户端。
- 在地址池中配置动态分配的地址网段和地址范围：

- 在进行非临时地址分配时，如果未在地址池下通过 **address range** 命令配置动态分配的 IPv6 非临时地址范围，则 **network** 命令指定的网段内的单播地址都可以分配给 DHCPv6 客户端。如果配置了 **address range** 命令，则只会从该地址范围内分配 IPv6 非临时地址，即使该范围内的地址分配完毕，也不会从 **network** 命令指定的地址范围内分配 IPv6 非临时地址。
- 在进行临时地址分配时，如果未在地址池下通过 **temporary address range** 命令配置动态分配的 IPv6 临时地址范围，则地址池无法分配临时地址。如果配置了 **temporary address range** 命令，则只会从该地址范围内分配 IPv6 临时地址，不会从 **network** 或者 **address range** 命令配置的地址范围内分配临时地址。

在实际组网中，某些地址是服务器的地址或者是保留地址，不应该动态分配给客户端。通过配置不参与自动分配的地址，可以避免 DHCPv6 服务器分配这些地址。

2. 配置限制和指导

配置为 DHCPv6 客户端分配 IPv6 地址，需要注意：

- 一个地址池下只能配置一个 IPv6 非临时地址范围和一个 IPv6 临时地址范围。
- **address range** 命令和 **temporary address range** 命令配置的地址范围应该在 **network** 命令配置的网段内，否则地址不能被分配。
- 一个 IPv6 地址只能与一个客户端绑定。不允许通过重复执行 **static-bind address** 命令的方式修改 IPv6 地址与客户端的绑定关系、地址的首选生命期和有效生命期。只有删除该 IPv6 地址的静态绑定配置后，才能通过重新配置将该 IPv6 地址与其他客户端绑定，或修改地址的首选生命期和有效生命期。
- 每个 DHCPv6 地址池只能配置一个网段，在相同地址池中重复执行 **network** 命令，新的配置会覆盖已有配置。如果相邻两次 **network** 命令配置的地址网段相同而首选生命期和有效生命期不同，则新配置的首选生命期和有效生命期只能在新生成的绑定信息中生效，原有绑定信息不受影响。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) （可选）配置不参与自动分配的 IPv6 地址。

```
ipv6 dhcp server forbidden-address start-ipv6-address
[ end-ipv6-address ] [ vpn-instance vpn-instance-name ]
```

缺省情况下，除 DHCPv6 服务器接口的 IPv6 地址外，DHCPv6 地址池中的所有 IPv6 地址都参与自动分配。

如果通过 **ipv6 dhcp server forbidden-address** 命令将已经静态绑定的 IPv6 地址配置为不参与自动分配的地址，则该地址仍然可以分配给静态绑定的用户。

- (3) 进入 DHCPv6 地址池视图。

```
ipv6 dhcp pool pool-name
```

- (4) 配置动态分配的 IPv6 地址网段。

```
network { prefix/prefix-length | prefix prefix-number
[ sub-prefix/sub-prefix-length ] } [ preferred-lifetime
preferred-lifetime valid-lifetime valid-lifetime ]
```


缺省情况下，未配置动态分配的 IPv6 地址网段。

不能在不同地址池下使用 **network** 命令配置相同的地址网段。

不能在不同地址池下引用完全一致的前缀编号、子前缀和子前缀长度。

- (5) （可选）配置动态分配的 IPv6 非临时地址范围。

```
address range start-ipv6-address end-ipv6-address  
[ preferred-lifetime preferred-lifetime valid-lifetime  
valid-lifetime ]
```

缺省情况下，未配置地址池中动态分配的 IPv6 非临时地址范围，整个网段内的单播地址都可以作为非临时地址分配给客户端。

- (6) （可选）配置动态分配的 IPv6 临时地址范围。

```
temporary address range start-ipv6-address end-ipv6-address  
[ preferred-lifetime preferred-lifetime valid-lifetime  
valid-lifetime ]
```

缺省情况下，未配置动态分配的 IPv6 临时地址范围，不能分配 IPv6 临时地址。

- (7) （可选）配置静态绑定的 IPv6 地址。

```
static-bind address ipv6-address/addr-prefix-length duid duid [ iaid  
iaid ] [ preferred-lifetime preferred-lifetime valid-lifetime  
valid-lifetime ] [ description description-text ]
```

缺省情况下，不存在静态绑定的 IPv6 地址。

重复执行 **static-bind address** 命令，可以配置多个静态绑定的 IPv6 地址。

2.5 配置为DHCPv6客户端分配网络参数

2.5.1 功能简介

除了分配 IPv6 地址和 IPv6 前缀外，DHCPv6 地址池中还可以配置其他网络参数，如在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 服务器域名等。可以通过如下方式配置为 DHCPv6 客户端分配的网络参数：

- 直接在 DHCPv6 地址池视图下配置网络参数。
- 在 DHCPv6 选项组中配置网络参数，并在 DHCPv6 地址池视图下指定引用的 DHCPv6 选项组。

直接在 DHCPv6 地址池视图下配置的网络参数的优先级高于 DHCPv6 选项组中配置的网络参数。

2.5.2 直接在 DHCPv6 地址池中配置网络参数

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCPv6 地址池视图。

```
ipv6 dhcp pool pool-name
```

- (3) 配置动态分配的 IPv6 地址网段。

```
network { prefix/prefix-length | prefix prefix-number  
[ sub-prefix/sub-prefix-length ] } [ preferred-lifetime  
preferred-lifetime valid-lifetime valid-lifetime ]
```

缺省情况下，未配置动态分配的 IPv6 地址网段。

不能在不同地址池下使用 **network** 命令配置相同的地址网段。

不能在不同地址池下引用完全一致的前缀编号、子前缀和子前缀长度。

地址池引用前缀编号分配动态地址时必须保证前缀号有对应的生效前缀，否则配置不生效。

- (4) 配置为客户端分配的 DNS 服务器地址。

```
dns-server ipv6-address
```

缺省情况下，未指定为客户端分配的 DNS 服务器地址。

- (5) 配置为客户端分配的域名。

```
domain-name domain-name
```

缺省情况下，未指定为客户端分配的域名。

- (6) 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 }
```

缺省情况下，未指定为客户端分配的 SIP 服务器地址或域名。

- (7) 配置 DHCPv6 自定义选项。

```
option code hex hex-string
```

缺省情况下，未配置 DHCPv6 自定义选项。

2.5.3 通过 DHCPv6 选项组配置网络参数

1. 功能简介

DHCPv6 选项组的创建方法有以下几种：

- 通过 **ipv6 dhcp option-group** 命令手工创建静态 DHCPv6 选项组。
- 设备作为 DHCPv6 客户端获取 IPv6 地址、前缀和网络配置参数时，在 DHCPv6 客户端上根据获取的网络配置参数动态创建 DHCPv6 选项组。

手工创建的 DHCPv6 选项组优先级高于动态创建的 DHCPv6 选项组。本节只介绍手工创建静态 DHCPv6 选项组的方法，动态创建 DHCPv6 选项组的方法请参见“三层技术-IP 业务配置指导”中的“DHCPv6 客户端”。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 手工创建静态 DHCPv6 选项组，并进入 DHCPv6 选项组视图。

```
ipv6 dhcp option-group option-group-number
```

- (3) 配置为客户端分配的 DNS 服务器地址。

```
dns-server ipv6-address
```

缺省情况下，未指定为客户端分配的 DNS 服务器地址。

- (4) 配置为客户端分配的域名后缀。

domain-name *domain-name*

缺省情况下，未指定为客户端分配的域名后缀。

- (5) 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* }

缺省情况下，未指定为客户端分配的 SIP 服务器地址或域名。

- (6) 配置 DHCPv6 自定义选项。

option code **hex** *hex-string*

缺省情况下，未配置 DHCPv6 自定义选项。

- (7) 退回系统视图。

quit

- (8) 进入 DHCPv6 地址池视图。

ipv6 dhcp pool *pool-name*

- (9) 配置 DHCPv6 地址池引用选项组。

option-group *option-group-number*

缺省情况下，DHCPv6 地址池未引用选项组。

2.6 配置接口工作在DHCPv6服务器模式，并配置地址/前缀分配方式

1. 功能简介

配置接口工作在 DHCPv6 服务器模式后，当接口未引用地址池时，接口收到 DHCPv6 客户端发来的 DHCPv6 报文时，服务器根据该接口的地址或 DHCPv6 中继接口的地址选择最长匹配的 DHCPv6 地址池，并从该地址池中选择 IPv6 地址或前缀分配给客户端。当接口引用地址池时，则从引用的地址池中选择 IPv6 地址或前缀分配给客户端。如果引用的地址池中不存在可供分配的 IPv6 地址或前缀，则设备将无法为客户端分配 IPv6 地址或前缀。

2. 配置限制和指导

配置接口工作在 DHCPv6 服务器模式，并配置地址/前缀分配方式时，需要注意：

- 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。
- 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。

3. 配置步骤

- (1) 进入系统视图。

system-view

- (2) 进入接口视图。

interface *interface-type interface-number*

- (3) 配置接口工作在 DHCPv6 服务器模式。

ipv6 dhcp select server

缺省情况下，接口未工作在 DHCPv6 服务器模式，也未工作在 DHCPv6 中继模式，接口接收到 DHCPv6 客户端发来的 DHCPv6 报文后，丢弃该报文。

(4) 配置地址池选择方式。请选择其中一项进行配置。

- 配置全局查找地址池，并指定全局查找 DHCPv6 地址池时地址或前缀分配方式。

```
ipv6 dhcp server { allow-hint | preference preference-value |  
rapid-commit } *
```

缺省情况下，支持接口全局查找 DHCPv6 地址池，但不支持期望地址/前缀分配和快速分配功能，且未指定服务器优先级。

- 配置接口引用 DHCP 地址池。

```
ipv6 dhcp server apply pool pool-name [ allow-hint | preference  
preference-value | rapid-commit ] *
```

缺省情况下，接口未引用地址池。

2.7 配置DHCPv6策略动态分配IPv6地址、前缀和其他参数

1. 功能简介

创建 DHCPv6 策略，并在接口引用该策略后，该接口接收到 DHCPv6 请求报文时，则根据配置顺序逐个匹配 DHCPv6 策略中通过 **class pool** 命令指定的 DHCPv6 用户类。匹配情况如下：

- 若匹配 DHCPv6 用户类成功，当该 DHCPv6 用户类关联的 DHCPv6 地址池中存在可供分配的地址或前缀信息时，则从该 DHCPv6 地址池中分配 IPv6 地址、前缀或其他参数；当该 DHCPv6 用户类关联的 DHCPv6 地址池中不存在可供分配的地址或前缀信息时，IPv6 地址、前缀或其他参数分配失败。
- 若匹配 DHCPv6 策略中的所有 DHCPv6 用户类失败，当配置了默认 DHCPv6 地址池时，则从该 DHCPv6 地址池中分配 IPv6 地址、前缀或网络参数；当未配置默认 DHCPv6 地址池或 DHCPv6 默认地址池不存在可供分配的 IPv6 地址或前缀时，IPv6 地址、前缀或其他参数分配失败。
- 若接收 DHCPv6 请求报文的接口引用的 DHCPv6 策略不存在或匹配的 DHCPv6 用户类关联的 DHCPv6 地址池不存在，IPv6 地址、前缀或其他参数分配失败。
- 匹配规则中不支持匹配 DHCPv6 设备添加的选项，比如 Option 18 或 Option 37。

2. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 创建 DHCPv6 用户类，并进入 DHCPv6 用户类视图。

```
ipv6 dhcp class class-name
```

(3) 配置 DHCPv6 用户类的匹配规则。

```
if-match rule rule-number { option option-code [ ascii acsii-string  
[ offset offset | partial ] | hex hex-string [ mask mask | offset offset  
length length | partial ] | relay-agent gateway-ipv6-address }
```

缺省情况下，未配置 DHCPv6 用户类的匹配规则。

(4) 退回系统视图。

```
quit
```

(5) 创建 DHCPv6 策略，并进入 DHCPv6 策略视图。

```
ipv6 dhcp policy policy-name
```

DHCPv6 策略需要在接口上引用才生效。

- (6) 指定 DHCPv6 用户类关联的 DHCPv6 地址池。

```
class class-name pool pool-name
```

缺省情况下，未指定 DHCPv6 用户类关联的 DHCPv6 地址池。

- (7) （可选）指定默认 DHCPv6 地址池。

```
default pool pool-name
```

缺省情况下，未指定默认 DHCPv6 地址池。

- (8) 退回系统视图。

```
quit
```

- (9) 进入接口视图。

```
interface interface-type interface-number
```

- (10) 指定接口引用的 DHCPv6 策略。

```
ipv6 dhcp apply-policy policy-name
```

缺省情况下，接口未引用 DHCPv6 策略。

2.8 配置DHCPv6服务器发送DHCPv6报文的DSCP优先级

1. 功能简介

DSCP 优先级用来体现报文自身的优先等级，决定报文传输的优先程度。通过本配置可以指定 DHCPv6 服务器发送的 DHCPv6 报文的 DSCP 优先级。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 DHCPv6 服务器发送 DHCPv6 报文的 DSCP 优先级。

```
ipv6 dhcp dscp dscp-value
```

缺省情况下，DHCPv6 服务器发送的 DHCPv6 报文的 DSCP 优先级为 56。

2.9 配置DHCPv6服务器租约固化功能

1. 功能简介

DHCPv6 服务器重启后，设备上记录的租约信息将丢失，会影响 DHCP 服务器的正常业务。

DHCPv6 服务器租约固化功能将 DHCPv6 服务器的核心运行数据（在用地址租约、冲突表项）保存到指定的文件中，DHCPv6 服务器设备重启后，自动根据该文件恢复 DHCPv6 服务器的租约信息，从而保证 DHCPv6 服务器的租约信息不会丢失。

当 DHCPv6 服务器设备重启后，自动根据该文件恢复 DHCPv6 服务器的租约信息，租约恢复的过程中，DHCPv6 服务器不能提供 DHCPv6 业务。所以当恢复过程出现问题导致恢复过程无法结束时，用户可配置 **ipv6 dhcp server database update stop** 命令终止当前的 DHCPv6 服务器表项恢复操作，以便 DHCPv6 服务器能及时提供 DHCPv6 服务。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 指定存储 DHCPv6 服务器表项的文件名称。

```
ipv6 dhcp server database filename { filename | url url [ username  
username [ password { cipher | simple } string ] ] }
```

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。

- (3) （可选）将当前的 DHCPv6 服务器表项保存到用户指定的文件中。

```
ipv6 dhcp server database update now
```

本命令只用来触发一次 DHCPv6 服务器表项的备份。

- (4) （可选）配置刷新 DHCPv6 服务器表项存储文件的延迟时间。

```
ipv6 dhcp server database update interval interval
```

缺省情况下，若 DHCPv6 服务器表项不变化，则不刷新存储文件；若 DHCPv6 服务器表项发生变化，默认在 300 秒之后刷新存储文件。

- (5) （可选）终止当前的 DHCPv6 服务器表项恢复操作。

```
ipv6 dhcp server database update stop
```

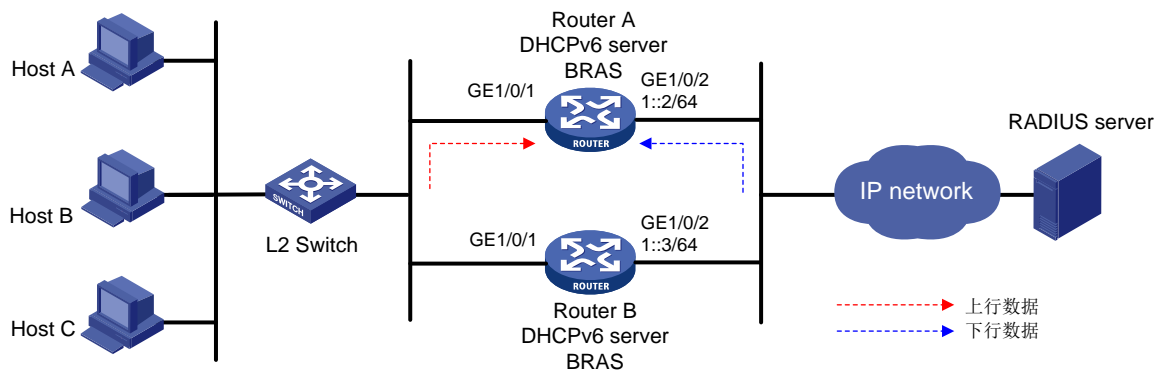
本命令只用来触发一次终止 DHCPv6 服务器表项信息的恢复。

2.10 配置DHCPv6服务器辅助路由信息

1. 功能简介

如图 2-4 所示，在某些特定的业务模型（如 BRAS 组网）下，BRAS 设备需要实时监测网络流量，并将统计数据发送到 RADIUS 服务器。该统计数据为用户上线以来产生的所有上下行流量数据，而不能是设备在某个时间段内发生的上下行流量数据。由于 RADIUS 服务器刷新计数的方法是覆盖以前数据而不是进行累加，所以当一台设备的上下行流量分别从两台 BRAS 设备上通过时，在 RADIUS 服务器上记录的数据就会相互覆盖，这时 RADIUS 服务器得到的统计数据是不准确的。为了提高准确性，需保证一台设备的上下行流量经过同一台 BRAS 设备。通过在 BRAS 设备上配置辅助路由信息，并对外发布此网段路由，引导指定网段的下行数据流量来保证上下行流量从一台 BRAS 设备经过。

图2-4 DHCPv6 服务器辅助路由组网图



2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCPv6 地址池视图。

```
ipv6 dhcp pool pool-name
```

- (3) 配置 DHCPv6 服务器辅助路由信息。

```
network { prefix/prefix-length | prefix prefix-number  
[ sub-prefix/sub-prefix-length ] } [ preferred-lifetime  
preferred-lifetime valid-lifetime valid-lifetime ] export-route
```

缺省情况下，未配置 DHCPv6 服务器辅助路由信息。

2.11 指定DHCPv6服务器上的地址池所属的VPN实例

1. 功能简介

当地址池绑定了 VPN 实例后，DHCPv6 服务器可以将网络划分成公网和 VPN 私网。未配置 VPN 属性的地址池被划分到公网，配置了 VPN 属性的地址池被划分到相应的 VPN 私网，这样，对于处于公网或 VPN 私网中的客户端，服务器都能够选择合适的地址池来为客户端分配租约并且记录该客户端的状态信息。

DHCPv6 服务器可以通过如下方式判断 DHCPv6 客户端所属的 VPN 实例：

- 认证模块用户接入时由 AAA 服务器授权 VPN 实例。
- DHCPv6 服务器接收报文的接口绑定的 VPN 实例即为该客户端所属的 VPN 实例。

如果以上两种方式都可获取到 DHCPv6 客户端所属的 VPN 实例，则以认证模块为准。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入 DHCPv6 地址池视图。

```
ipv6 dhcp pool pool-name
```

- (3) 指定 DHCPv6 服务器上的地址池所属的 VPN 实例。

```
vpn-instance vpn-instance-name
```

缺省情况下，未指定 DHCPv6 服务器上的地址池所属的 VPN 实例。

2.12 开启DHCPv6服务器的日志信息功能

1. 功能简介

DHCPv6 服务器日志可以方便管理员定位问题和解决问题。设备生成 DHCPv6 日志信息会交给信息中心模块处理，信息中心模块的配置将决定日志信息的发送规则和发送方向。关于信息中心的详细描述请参见“网络管理和监控配置指导”中的“信息中心”。

2. 配置限制和指导

大量 DHCPv6 客户端发生上下线操作时, DHCPv6 服务器需要输出大量日志信息, 这可能会降低设备性能, 影响 DHCPv6 服务器分配 IPv6 前缀或 IPv6 地址的速度。为了避免该情况的发生, 用户可以关闭 DHCPv6 服务器日志信息功能, 使得 DHCPv6 服务器不再输出日志信息。

3. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 开启 DHCPv6 服务器日志信息功能。

```
ipv6 dhcp log enable
```

缺省情况下, DHCPv6 服务器日志信息功能处于关闭状态。

2.13 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2-1 DHCPv6 服务器显示和维护

操作	命令
显示本设备的DUID	display ipv6 dhcp duuid
显示DHCPv6选项组信息	display ipv6 dhcp option-group [<i>option-group-number</i>]
显示DHCPv6地址池的信息	display ipv6 dhcp pool [<i>pool-name</i> vpn-instance <i>vpn-instance-name</i>]
显示前缀池的信息	display ipv6 dhcp prefix-pool [<i>prefix-pool-number</i>] [vpn-instance <i>vpn-instance-name</i>]
显示接口上的DHCPv6服务器信息	display ipv6 dhcp server [interface <i>interface-type interface-number</i>]
显示DHCPv6地址冲突信息	display ipv6 dhcp server conflict [address <i>ipv6-address</i>] [vpn-instance <i>vpn-instance-name</i>]
显示DHCPv6服务器表项备份信息	display ipv6 dhcp server database
显示租约过期的DHCPv6地址绑定信息	display ipv6 dhcp server expired [[address <i>ipv6-address</i>] [vpn-instance <i>vpn-instance-name</i>] pool <i>pool-name</i>]
显示DHCPv6地址绑定信息	display ipv6 dhcp server ip-in-use [[address <i>ipv6-address</i>] [vpn-instance <i>vpn-instance-name</i>] pool <i>pool-name</i>]
显示DHCPv6前缀绑定信息	display ipv6 dhcp server pd-in-use [pool <i>pool-name</i> [prefix <i>prefix/prefix-len</i>] [vpn-instance <i>vpn-instance-name</i>]]
显示DHCPv6服务器的报文统计信息	display ipv6 dhcp server statistics [pool <i>pool-name</i> vpn-instance <i>vpn-instance-name</i>]

操作	命令
清除DHCPv6地址冲突信息	reset ipv6 dhcp server conflict [address <i>ipv6-address</i>] [vpn-instance <i>vpn-instance-name</i>]
清除租约过期的DHCPv6地址绑定信息	reset ipv6 dhcp server expired [[address <i>ipv6-address</i>] [vpn-instance <i>vpn-instance-name</i>] pool <i>pool-name</i>]
清除DHCPv6的正式地址绑定和临时地址绑定信息	reset ipv6 dhcp server ip-in-use [[address <i>ipv6-address</i>] [vpn-instance <i>vpn-instance-name</i>] pool <i>pool-name</i>]
清除DHCPv6正式前缀绑定和临时前缀绑定信息	reset ipv6 dhcp server pd-in-use [pool <i>pool-name</i> [prefix <i>prefix/prefix-len</i>] [vpn-instance <i>vpn-instance-name</i>]]
清除DHCPv6服务器的报文统计信息	reset ipv6 dhcp server statistics [vpn-instance <i>vpn-instance-name</i>]

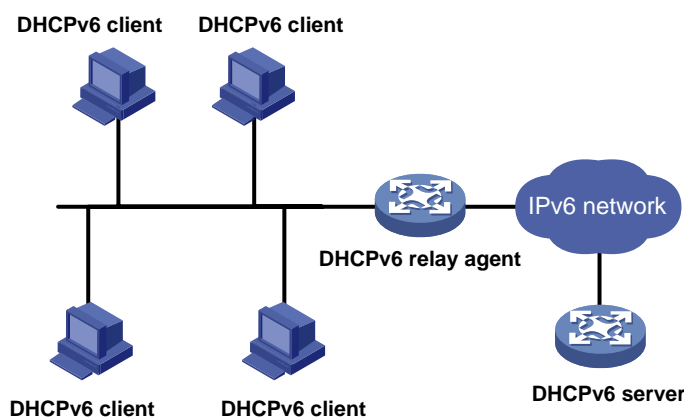
3 DHCPv6 中继

3.1 DHCPv6 中继简介

3.1.1 应用环境

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。

图3-1 DHCPv6 中继应用组网图

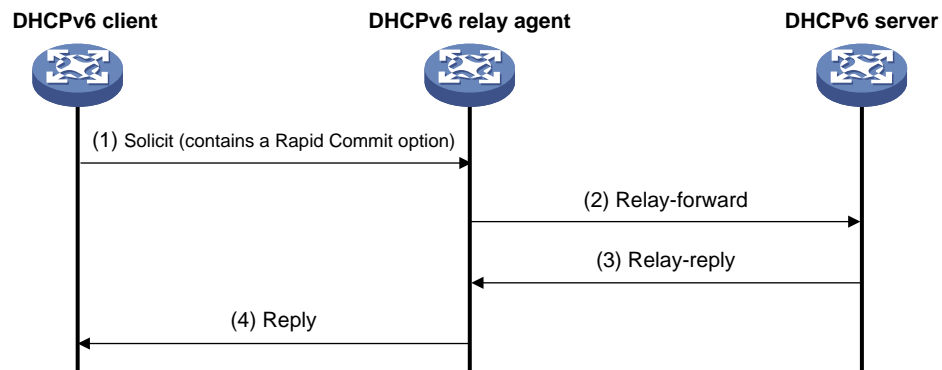


3.1.2 DHCPv6 中继的工作过程

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。

图3-2 DHCPv6 中继的工作过程



3.2 DHCPv6中继配置任务简介

DHCPv6 中继配置任务如下：

- (1) [配置接口工作在 DHCPv6 中继模式](#)
- (2) [指定 DHCPv6 服务器的地址](#)
- (3) （可选）[配置 DHCPv6 中继为 DHCPv6 客户端分配的网关地址](#)
- (4) （可选）[配置 DHCPv6 中继发送 DHCPv6 报文的 DSCP 优先级](#)
- (5) （可选）[配置 DHCPv6 中继支持的 Interface ID 选项填充模式](#)

3.3 配置接口工作在DHCPv6中继模式

1. 配置限制和指导

建议不要在一个接口上同时配置 DHCPv6 中继和 DHCPv6 客户端功能。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口工作在 DHCPv6 中继模式。

```
ipv6 dhcp select relay
```

缺省情况下，接口未工作在 DHCPv6 中继模式。

3.4 指定DHCPv6服务器的地址

3.4.1 指定 DHCPv6 中继对应的 DHCPv6 服务器地址

1. 功能简介

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、IPv6 前缀和其他网络配置参数。

2. 配置限制和指导

- 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。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 指定 DHCPv6 中继对应的 DHCPv6 服务器地址。

```
ipv6 dhcp relay server-address ipv6-address [ interface interface-type  
interface-number ]
```

缺省情况下，未指定 DHCPv6 中继对应的 DHCPv6 服务器地址。

3.4.2 指定中继地址池上对应的 DHCPv6 服务器地址

1. 功能简介

对于某些特定的用户接入方式，基于用户接入位置信息的不同，网络中存在大量不同类型的用户。为了使相同类型的用户可以从指定的 DHCPv6 服务器申请 IPv6 地址等网络参数，模块根据用户注册信息，使不同的用户选择不同的 DHCPv6 中继地址池，并从中继地址池下配置的 DHCPv6 服务器获取 IPv6 地址等网络参数。

一台 DHCPv6 中继的一个接口下可能连接不同类型的用户，当 DHCPv6 中继转发 DHCPv6 客户端请求报文给 DHCPv6 服务器时，不能再以中继接口的 IPv6 地址作为选择地址池的依据。为了解决这个问题，需要使用 **gateway-list** 命令指定某个类型用户所在的网段，并将该地址添加到转发给 DHCPv6 服务器的报文的 Link-address 字段中，为 DHCPv6 服务器选择地址池提供依据。

2. 配置限制和指导

- 为了提高可靠性，一个 DHCPv6 中继地址池下最多可以配置 8 个 DHCPv6 服务器地址，当 DHCPv6 客户端匹配该中继地址池后，DHCPv6 中继会将 DHCPv6 客户端发来的 DHCPv6 报文转发给该地址池对应所有的 DHCPv6 服务器。

- 当 PPPoE 用户下线时，DHCPv6 中继需要查询中继用户地址表项，若存在对应表项，则会向 DHCPv6 服务器发送 Release 报文，通知 DHCPv6 服务器释放该地址租约。这就需要在 DHCPv6 中继上使用 **ipv6 dhcp relay client-information record** 命令开启 DHCPv6 中继用户地址表项记录功能。
- 和 PPPoE 配合使用时，如果设备的地址池中配置了 **remote-server** 命令，则可以认定该设备一定是 DHCPv6 中继设备，所以不需要在接口视图下执行 **ipv6 dhcp select relay** 命令。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 创建 DHCPv6 中继地址池，并进入 DHCPv6 中继地址池视图。

```
ipv6 dhcp pool pool-name
```

- (3) 指定匹配该地址池的 DHCPv6 客户端所在的网段地址。

```
gateway-list ipv6-address<1-8>
```

缺省情况下，未指定匹配该地址池的 DHCPv6 客户端所在的网段地址。

- (4) 指定中继地址池对应的 DHCPv6 服务器地址。

```
remote-server ipv6-address [ interface interface-type  
interface-number ]
```

缺省情况下，未指定中继地址池对应的 DHCPv6 服务器的地址。

3.5 配置DHCPv6中继为DHCPv6客户端分配的网关地址

1. 功能简介

当未开启该功能时，DHCPv6 中继收到 DHCPv6 客户端的请求报文后，只能将接口的第一个 IPv6 添加到报文中，然后转发给 DHCPv6 服务器。对于某些特定需求，DHCPv6 中继需要添加指定的地址到报文中，这时就需要配置此功能。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置 DHCPv6 中继为 DHCPv6 客户端分配的网关地址。

```
ipv6 dhcp relay gateway ipv6-address
```

缺省情况下，DHCPv6 中继分配接口下的第一个 IPv6 地址作为 DHCPv6 客户端的网关地址。

3.6 配置DHCPv6中继发送DHCPv6报文的DSCP优先级

1. 功能简介

DSCP 优先级用来体现报文自身的优先等级，决定报文传输的优先程度。通过本配置可以指定 DHCPv6 中继发送的 DHCPv6 报文的 DSCP 优先级。

2. 配置步骤

- (1) 进入系统视图。
`system-view`
- (2) 配置 DHCPv6 中继发送 DHCPv6 报文的 DSCP 优先级。
`ipv6 dhcp dscp dscp-value`
缺省情况下，DHCPv6 中继发送的 DHCPv6 报文的 DSCP 优先级为 56。

3.7 配置DHCPv6中继支持的Interface ID选项填充模式

1. 功能简介

如果配置了 DHCPv6 中继支持的 Interface ID 选项填充模式，当 DHCPv6 中继接收到客户端发送的 DHCPv6 报文后，会以配置的填充方式将 DHCPv6 客户端的位置信息填充 Option 18 选项，并把填充好的报文转发给 DHCPv6 服务器。

2. 配置步骤

- (1) 进入系统视图。
`system-view`
- (2) 进入接口视图。
`interface interface-type interface-number`
- (3) 配置 DHCPv6 中继支持的 Interface ID 选项填充模式。
`ipv6 dhcp relay interface-id { bas | interface }`
缺省情况下，Interface ID 选项的填充模式为接口索引信息。

3.8 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3-1 DHCPv6 中继显示和维护

操作	命令
显示本设备DUID	<code>display ipv6 dhcp duid</code>
显示DHCPv6中继接口上指定的DHCPv6服务器地址信息	<code>display ipv6 dhcp relay server-address [interface interface-type interface-number]</code>
显示DHCPv6中继的相关报文统计信息	<code>display ipv6 dhcp relay statistics [interface interface-type interface-number]</code>
清除DHCPv6中继的相关报文统计信息	<code>reset ipv6 dhcp relay statistics [interface interface-type interface-number]</code>

4 DHCPv6 客户端

4.1 DHCPv6客户端简介

设备作为 DHCPv6 客户端时，可以具有如下功能：

- 通过 DHCPv6 获取 IPv6 地址和网络配置参数，IPv6 地址作为开启 DHCPv6 客户端功能的接口地址，当设备开启 DHCPv6 服务器功能后，获取的网络配置参数用来自动创建 DHCPv6 选项组。
- 通过 DHCPv6 获取 IPv6 前缀和网络配置参数，IPv6 前缀作为本地设备的 IPv6 前缀（本地设备根据该前缀生成 IPv6 地址）；当设备开启 DHCPv6 服务器功能后，获取的网络配置参数用来自动创建 DHCPv6 选项组。
- 通过 DHCPv6 同时获取 IPv6 地址、IPv6 前缀和网络配置参数，IPv6 地址作为开启 DHCPv6 客户端功能的接口地址，IPv6 前缀作为本地设备的 IPv6 前缀（本地设备根据该前缀生成 IPv6 地址）；当设备开启 DHCPv6 服务器功能后，获取的网络配置参数用来自动创建 DHCPv6 选项组。
- 通过 DHCPv6 无状态配置获取除 IPv6 地址/前缀外的其他网络配置参数。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 标志位的取值为 0、O 标志位的取值为 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。否则 DHCPv6 客户端不会开启无状态配置过程。

4.2 DHCPv6客户端配置限制和指导

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，否则会影响功能正常使用。

4.3 DHCPv6客户端配置任务简介

DHCPv6 客户端配置任务如下：

- (1) （可选）[配置接口使用的 DHCPv6 客户端 DUID](#)
- (2) 配置 DHCPv6 客户端获取 IPv6 地址、IPv6 前缀和网络配置参数
请至少选择以下一项任务进行配置：
 - [配置 DHCPv6 客户端获取 IPv6 地址和网络配置参数](#)
 - [配置 DHCPv6 客户端获取 IPv6 前缀和网络配置参数](#)
 - [配置 DHCPv6 客户端同时获取 IPv6 地址、IPv6 前缀和网络配置参数](#)
 - [配置 DHCPv6 客户端获取除地址/前缀外的其他网络配置参数](#)
- (3) （可选）[配置 DHCPv6 客户端发送 DHCPv6 报文的 DSCP 优先级](#)

4.4 配置接口使用的DHCPv6客户端DUID

1. 功能简介

DHCPv6 客户端 DUID 用来填充 DHCPv6 报文的 Option 1, 作为识别 DHCPv6 客户端的唯一标识。DHCPv6 服务器可以根据 DHCPv6 客户端 DUID 为特定的 DHCPv6 客户端分配特定的 IPv6 地址。用户可以通过三种方法指定 DHCPv6 客户端 DUID: ASCII 字符串、十六进制数或接口的 MAC 地址。

2. 配置限制和指导

用户在指定客户端 ID 时, 需要确保不同客户端的客户端 ID 不能相同。

3. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口使用的 DHCPv6 客户端 DUID。

```
ipv6 dhcp client duid { ascii ascii-string | hex hex-string | mac  
interface-type interface-number }
```

缺省情况下, 根据设备的桥 MAC 地址生成 DHCPv6 客户端 DUID。

4.5 配置DHCPv6客户端获取IPv6地址和网络配置参数

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口作为 DHCPv6 客户端, 通过 DHCPv6 方式获取 IPv6 地址和其他网络配置参数。

```
ipv6 address dhcp-alloc [ option-group group-number | rapid-commit ] *
```

缺省情况下, 接口不会作为 DHCPv6 客户端获取 IPv6 地址和网络配置参数。

4.6 配置DHCPv6客户端获取IPv6前缀和网络配置参数

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口作为 DHCPv6 客户端, 通过 DHCPv6 方式获取 IPv6 前缀和其他网络配置参数。

```
ipv6 dhcp client pd prefix-number [ option-group group-number |  
rapid-commit ] *
```

缺省情况下, 接口不会作为 DHCPv6 客户端获取 IPv6 前缀和网络配置参数。

4.7 配置DHCPv6客户端同时获取IPv6地址、IPv6前缀和网络配置参数

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 配置接口作为 DHCPv6 客户端，通过 DHCPv6 方式同时获取 IPv6 地址、IPv6 前缀和其他网络配置参数。

```
ipv6 dhcp client stateful prefix prefix-number [ option-group  
option-group-number | rapid-commit ] *
```

缺省情况下，接口不会作为 DHCPv6 客户端同时获取 IPv6 地址、IPv6 前缀和网络配置参数。

4.8 配置DHCPv6客户端获取除地址/前缀外的其他网络配置参数

1. 功能简介

DHCPv6 客户端可通过如下方式获取除地址/前缀外的其他网络参数：

- 如果接口上只配置了 **ipv6 address auto** 命令，则接口会通过无状态自动配置方式生成全球单播地址，同时自动生成链路本地地址。只有接收到的 RA 报文中 M 标志位的取值为 0、O 标志位的取值为 1 时，设备才会自动启动 DHCPv6 无状态配置功能。
- 如果接口只配置了 **ipv6 dhcp client stateless enable** 命令，则接口开启了 DHCPv6 客户端功能，并从 DHCPv6 服务器获取除地址/前缀外的其他网络配置参数。
- 如果接口上同时配置了 **ipv6 address auto** 命令和 **ipv6 dhcp client stateless enable** 命令，则接口通过无状态生成全球单播地址，同时自动生成链路本地地址，且直接从 DHCPv6 服务器获取除地址/前缀外的其他网络配置参数。

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”。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 进入接口视图。

```
interface interface-type interface-number
```

- (3) 开启无状态自动配置功能。请至少选择其中一项进行配置。

- 开启 IPv6 地址无状态自动配置功能。

```
ipv6 address auto
```

- 开启 DHCPv6 客户端无状态配置功能。

```
ipv6 dhcp client stateless enable
```

缺省情况下，接口不会作为 DHCPv6 客户端获取除地址/前缀外的其他网络配置参数。

4.9 配置DHCPv6客户端发送DHCPv6报文的DSCP优先级

1. 功能简介

DSCP 优先级用来体现报文自身的优先等级，决定报文传输的优先程度。通过本配置可以指定 DHCPv6 客户端发送的 DHCPv6 报文的 DSCP 优先级。

2. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 配置 DHCPv6 客户端发送的 DHCPv6 报文的 DSCP 优先级。

```
ipv6 dhcp client dscp dscp-value
```

缺省情况下，DHCPv6 客户端发送的 DHCPv6 报文的 DSCP 优先级为 56。

4.10 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4-1 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 <i>interface-type</i> <i>interface-number</i>]
显示DHCPv6客户端的统计信息	display ipv6 dhcp client statistics [interface <i>interface-type</i> <i>interface-number</i>]
清除DHCPv6客户端的统计信息	reset ipv6 dhcp client statistics [interface <i>interface-type</i> <i>interface-number</i>]

目 录

1 IPv6 快速转发	1-1
1.1 IPv6 快速转发简介	1-1
1.2 vSystem 相关说明	1-1
1.3 配置 IPv6 快速转发功能	1-1
1.4 配置 IPv6 快速转发表项的老化时间	1-2
1.5 配置 IPv6 快速转发负载分担	1-2
1.6 IPv6 快速转发显示和维护	1-2

1 IPv6 快速转发

1.1 IPv6快速转发简介

报文转发效率是衡量设备性能的一项关键指标。按照常规流程，设备收到一个报文后，根据报文的目地址寻找路由表中与之匹配的路由，然后确定一条最佳的路径，同时还将报文按照数据链路层上使用的协议进行封装，最后进行报文转发。

快速转发是采用高速缓存来处理报文，采用了基于数据流的技术。

IPv6 快速转发根据报文中的信息（比如源 IP 地址、目的 IP 地址、源端口、目的端口、IP 协议号等）来标识一条数据流。当一条数据流的第一个报文通过查找路由表转发后，相应的转发信息将被记录到高速缓存中的快速转发表中，该数据流后续报文就可以通过直接查找快速转发表进行转发。这样便大大缩减了 IPv6 报文的排队流程，减少报文的转发时间，提高 IPv6 报文的转发效率。

1.2 vSystem相关说明

非缺省 vSystem 不支持本特性部分功能，包括配置 IPv6 快速转发表项的老化时间和配置 IPv6 快速转发负载分担。



说明

非缺省 vSystem 对具体命令的支持情况，请见本特性的命令参考。有关 vSystem 的详细介绍请参见“虚拟化技术配置指导”中的“vSystem”。

1.3 配置IPv6快速转发功能

1. 功能简介

缺省情况下，IPv6 快速转发功能处于开启状态，IPv6 快速转发功能开启后，会生成快转表项。设备使用快转表项，能加快转发速度。当存在大量报文流生成大量快转表项时，可能会占用过多内存，导致其他业务申请内存失败无法继续工作。此时，可暂时关闭 IPv6 快速转发功能，通过表项老化释放内存空间。

2. 配置步骤

(1) 进入系统视图。

```
system-view
```

(2) 配置 IPv6 快速转发功能。请选择其中一项进行配置。

- 开启 IPv6 快速转发功能。

```
ipv6 fast-forwarding enable
```

- 关闭 IPv6 快速转发功能。

```
undo ipv6 fast-forwarding enable
```

缺省情况下，IPv6 快速转发功能处于开启状态。

1.4 配置IPv6快速转发表项的老化时间

1. 功能简介

IPv6 快速转发表中的表项并非永远有效，每一条记录都有一个生存周期，到达生存周期仍得不到刷新的记录将从 IPv6 快速转发表中删除，这个生存周期被称作老化时间。如果在到达老化时间前记录被刷新，则重新计算老化时间。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 IPv6 快速转发表项的老化时间。

```
ipv6 fast-forwarding aging-time aging-time
```

缺省情况下，IPv6 快速转发表项的老化时间为 30 秒。

1.5 配置IPv6快速转发负载分担

1. 功能简介

缺省情况下，IPv6 快速转发负载分担功能处于开启状态，IPv6 快速转发根据报文中的信息来标识一条数据流。关闭 IPv6 快速转发负载分担功能后，IPv6 快速转发根据报文中的信息和入接口来标识一条数据流。

2. 配置步骤

- (1) 进入系统视图。

```
system-view
```

- (2) 配置 IPv6 快速转发负载分担功能。请选择其中一项进行配置。

- 开启 IPv6 快速转发负载分担功能。

```
ipv6 fast-forwarding load-sharing
```

- 关闭 IPv6 快速转发负载分担功能。

```
undo ipv6 fast-forwarding load-sharing
```

缺省情况下，IPv6 快速转发负载分担功能处于开启状态。

1.6 IPv6快速转发显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示 IPv6 快速转发配置后的运行情况，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 IPv6 快速转发表中的内容。

表1-1 IPv6 快速转发显示和维护

操作	命令
显示IPv6快速转发表项的老化时间	display ipv6 fast-forwarding aging-time

操作	命令
显示IPv6快速转发表信息	<p>(独立运行模式)</p> <pre>display ipv6 fast-forwarding cache [<i>ipv6-address</i>] [slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</pre> <p>(IRF模式)</p> <pre>display ipv6 fast-forwarding cache [<i>ipv6-address</i>] [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</pre>
显示IPv6虚拟分片报文快转表信息	<p>(独立运行模式)</p> <pre>display ipv6 fast-forwarding fragcache [<i>ipv6-address</i>] [slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</pre> <p>(IRF模式)</p> <pre>display ipv6 fast-forwarding fragcache [<i>ipv6-address</i>] [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</pre>
清除IPv6快速转发表信息	<p>(独立运行模式)</p> <pre>reset ipv6 fast-forwarding cache [slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</pre> <p>(IRF模式)</p> <pre>reset ipv6 fast-forwarding cache [chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>]]</pre>