



UNIS 入侵防御系统产品

可靠性配置指导

北京紫光恒越网络科技有限公司
<http://www.unis-hy.com>

资料版本：5PW100-20160929

Copyright © 2016 北京紫光恒越网络科技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。

UNIS 为北京紫光恒越网络科技有限公司的商标。对于本手册中出现的其它公司的商标、产品标识及商品名称，由各自权利人拥有。

由于产品版本升级或其他原因，本手册内容有可能变更。紫光恒越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，紫光恒越尽全力在本手册中提供准确的信息，但是紫光恒越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

UNIS 入侵防御系统产品配置指导介绍了入侵防御系统产品各软件特性的原理及其配置方法，包含原理简介、配置任务描述和配置举例。《可靠性配置指导》主要介绍备份组、冗余备份、BFD、Track 和进程分布优化相关的特性。

前言部分包含如下内容：

- [适用款型](#)
- [读者对象](#)
- [本书约定](#)
- [技术支持](#)
- [资料意见反馈](#)

适用款型

入侵防御系统产品款型较多，形态丰富，本手册所描述的内容适用于如下产品款型：

表1 手册适用的产品款型

款型	形态
UNIS T5000-M06	分布式设备，可以运行在： <ul style="list-style-type: none">• 独立运行模式• IRF 模式
UNIS T5000-G20	集中式IRF设备
UNIS T1000-G20/G50/G60/G80	集中式IRF设备

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 ...]	表示从多个选项中选择一个或者不选。
{x y ...}*	表示从多个选项中至少选取一个。
[x y ...]*	表示从多个选项中选择一个、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. 端口编号示例约定

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

技术支持

用户支持邮箱：zgsm_service@thunis.com

技术支持热线电话：400-910-9998（手机、固话均可拨打）

网址：<http://www.unis-hy.com>

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：zgsm_info@thunis.com

感谢您的反馈，让我们做得更好！

目 录

1 备份组.....	1-1
1.1 备份组简介.....	1-1
1.1.1 备份组分类.....	1-1
1.1.2 备份组的应用.....	1-2
1.2 配置备份组.....	1-2
1.3 备份组显示和维护.....	1-2

1 备份组

设备各款型对于本节所描述的特性支持情况有所不同，详细差异信息如下：

型号	特性	描述
T5000-M06	备份组	支持
T5000-G20		不支持
T1000-G20/G50/G60/G80		不支持



说明

- 本文中的节点指的是 CPU。为了保证业务在主、备节点迁移后，仍能正常运行，建议将不同单板上的性能相当的两个 CPU 互为备份。
- 对于本节命令中的 CPU 参数，仅 T5000-M06 产品支持。

1.1 备份组简介

备份组用于实现特定业务在指定 CPU 之间的数据备份，为特定业务的高可靠性运行提供保障。

备份组由主节点和备节点组成。业务模块引用备份组后，主节点处理业务流量并将业务数据备份到备节点，备节点处于备份状态，不处理业务。当主节点故障时，流量自动切换到备节点。当主节点再次恢复后，流量会切回到主节点，实现业务的备份和可靠运行。

1.1.1 备份组分类

备份组有两种：自动备份组和手动备份组。

1. 自动备份组

为了满足特定业务的备份需求，系统会自动为每块引擎板创建一个备份组，这种备份组称为自动备份组，这块引擎板即为自动备份组的主节点。

缺省情况下自动备份组没有备节点。当引擎板被拔出时，系统会自动选择同一安全引擎组内的其它引擎板作为备节点，并将流量自动切换到备节点。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”。

2. 手动备份组

用户通过命令行创建的备份组。

手动备份组的主节点和备节点可通过命令行配置。

1.1.2 备份组的应用

使用备份组可以实现 1:1 备份：创建一个备份组，加入主节点和备节点。由主节点处理流量，并将业务数据备份到备节点。

1.2 配置备份组

表1-1 配置备份组

操作	命令	说明
进入系统视图	system-view	-
创建备份组，并进入备份组视图。如果指定的备份组已存在，则直接进入该备份组的视图	failover group <i>group-name</i>	缺省情况下，存在自动备份组（名字以 AutoBackup为前缀），不存在手动备份组
节点加入备份组（分布式设备—独立运行模式）	bind slot <i>slot-number</i> cpu <i>cpu-number</i> { primary secondary }	只有指定的单板支持多CPU时，才能继续配置 cpu 参数。如果指定的单板为单CPU单板，则不能指定 cpu 参数 缺省情况下，备份组内不存在节点 不同备份组的主节点不能相同，同一备份组的主节点和备节点不能相同
节点加入备份组（分布式设备—IRF模式）	bind chassis <i>chassis-number</i> slot <i>slot-number</i> cpu <i>cpu-number</i> { primary secondary }	只有指定的单板支持多CPU时，才能继续配置该参数。如果指定的单板为单CPU单板，则不能指定该参数 缺省情况下，备份组内不存在节点 不同备份组的主节点不能相同，同一备份组的主节点和备节点不能相同

1.3 备份组显示和维护

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备份组的运行情况，通过查看显示信息验证配置的效果。

表1-2 备份组显示和维护

操作	命令
显示备份组的信息	display failover group [<i>group-name</i>]

目 录

1 以太网冗余接口	1-1
1.1 以太网冗余接口简介.....	1-1
1.2 配置以太网冗余接口.....	1-2
1.3 以太网冗余接口显示和维护.....	1-3
1.4 以太网冗余接口典型配置举例.....	1-4
2 冗余组	2-1
2.1 冗余组简介.....	2-1
2.2 冗余组原理.....	2-2
2.2.1 冗余组节点.....	2-2
2.2.2 冗余组成员.....	2-5
2.2.3 冗余组定时器.....	2-7
2.2.4 冗余组的倒换机制.....	2-7
2.3 配置冗余组.....	2-7
2.3.1 冗余组配置任务简介.....	2-7
2.3.2 创建冗余组.....	2-8
2.3.3 配置冗余组节点.....	2-8
2.3.4 配置冗余组的成员接口.....	2-9
2.3.5 配置冗余组成员备份组.....	2-9
2.3.6 配置冗余组定时器.....	2-10
2.3.7 手工触发冗余组倒换.....	2-10
2.3.8 开启冗余组告警功能.....	2-10
2.4 冗余组显示和维护.....	2-11
2.5 冗余组典型配置举例（集中式IRF设备）.....	2-11
2.5.1 工作在三层，上下行分别连接两台路由器，两台路由器接口不在同一网段.....	2-11
2.5.2 工作在三层，上下行分别连接一台路由器.....	2-15
2.6 冗余组典型配置举例（分布式设备—IRF模式）.....	2-18
2.6.1 工作在三层，上下行分别连接两台路由器，两台路由器接口不在同一网段.....	2-18
2.6.2 工作在三层，上下行分别连接一台路由器.....	2-22

1 以太网冗余接口

1.1 以太网冗余接口简介

以太网冗余接口（Redundant Ethernet, Reth）是一种三层虚拟接口。一个以太网冗余接口中包含两个成员接口，这两个成员接口的类型可以为：三层以太网接口、三层 GigabitEthernet 接口、三层 Ten-GigabitEthernet 接口、三层 TwentyGigE 接口、三层 FortyGigE 接口、三层 HundredGigE 接口、三层聚合接口、EFM 接口及上述接口的子接口。使用以太网冗余接口不但可以实现两个接口之间的备份，还可以实现两个聚合接口之间的备份。

以太网冗余接口主要用途是作为冗余组的成员接口，根据冗余组的指令完成以太网冗余接口内成员接口的状态切换，从而配合实现冗余组的整体主备倒换需求。以太网冗余接口也可以单独组网使用。本章节仅描述以太网冗余接口单独组网使用的情况，以太网冗余组相关描述，以及以太网冗余接口和冗余组配合使用的情况，请参见“[2 冗余组](#)”。

如 [图 1-1](#) 所示，在 Device A 和 Device B 组成的 IRF 中，使用以太网冗余接口技术，可以实现 GE1/0/1 和 GE2/0/1 互为备份，GE1/0/2 和 GE2/0/2 互为备份。正常情况下，报文通过 GE1/0/2 和 GE1/0/1 转发，GE2/0/1 和 GE2/0/2 处于协议关闭状态。当 GE1/0/1 故障，系统会自动启用 GE2/0/1 接口转发报文。（集中式 IRF 设备）

如 [图 1-2](#) 所示，在 Device A 和 Device B 组成的 IRF 中，使用以太网冗余接口技术，可以实现 GE1/1/0/1 和 GE2/1/0/1 互为备份，GE1/1/0/2 和 GE2/1/0/2 互为备份。正常情况下，报文通过 GE1/1/0/2 和 GE1/1/0/1 转发，GE2/1/0/1 和 GE2/1/0/2 处于协议关闭状态。当 GE1/1/0/1 故障，系统会自动启用 GE2/1/0/1 接口转发报文。（分布式设备—IRF 模式）

图1-1 以太网冗余接口典型应用组网图（集中式 IRF 设备）

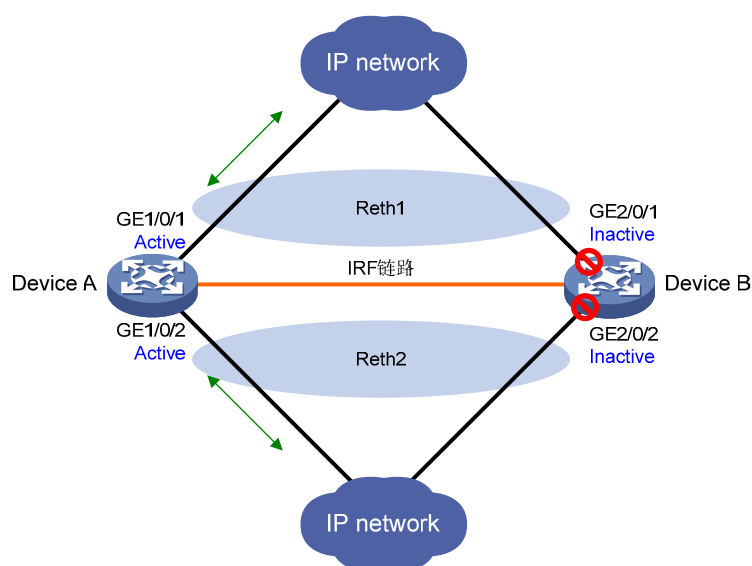
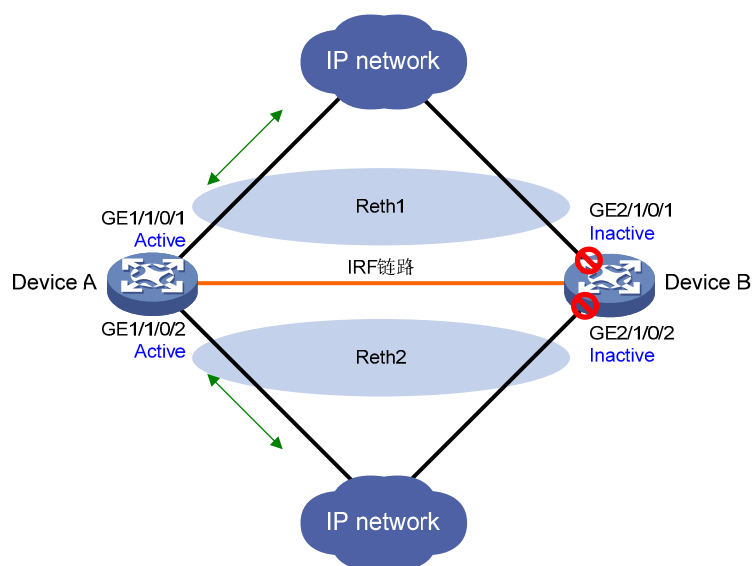


图1-2 以太网冗余接口典型应用组网图（分布式设备—IRF 模式）



2. 以太网冗余接口的成员接口的状态

以太网冗余接口的成员接口有两种状态：

- 激活状态：能够收发报文。
- 非激活状态：不能收发报文。

成员接口的激活/非激活状态由以下因素决定：

- 当两个成员接口的物理状态均为 **up** 时：
 - 优先级较高的成员接口处于激活状态，优先级较低成员接口处于非激活状态。优先级可通过命令行配置。
 - 如果以太网冗余接口加入了冗余组，由冗余组决定哪个成员接口处于激活状态。具体描述请参见“[2 冗余组](#)”。
- 当激活接口物理状态变为 **down** 时，系统会自动激活另外一个成员接口。
- 当两个成员接口的物理状态均为 **down** 时，两个接口均为非激活状态。

3. 以太网冗余接口的工作原理

任意时刻，同一个以太网冗余接口内只有一个成员接口处于激活状态。当激活接口的链路变为 **Down** 时，处于非激活状态的接口会自动激活，用来接替原激活接口收发报文，实现接口间的备份。

以太网冗余接口使用设备的桥 **MAC** 地址作为 **MAC** 地址，可配置 **IP** 地址和安全特性（如 **NAT**、**IP Sec** 等）。设备使用以太网冗余接口和上、下行设备通信，成员接口收到报文后，交给以太网冗余接口处理，以太网冗余接口处理完后交给成员接口发送。在上、下行设备看来，与其连接的是以太网冗余接口。即便以太网冗余接口的单个成员接口的物理状态发生变化，也不会影响上、下行设备。

1.2 配置以太网冗余接口

用户可以手动创建和删除以太网冗余接口。当以太网冗余接口中还有成员接口时，请将成员接口从以太网冗余接口中删除后，再删除以太网冗余接口。

添加成员接口有如下限制：

- 最多可添加两个成员接口，且两个成员接口的优先级不能相同。
- 同一以太网冗余接口的成员接口的类型和速率最好相同，例如均为 100M 三层以太网接口。从而能够保证成员接口切换后不因带宽不同，影响正常的流量转发。
- 一个接口/子接口加入一个以太网冗余接口后，不能加入其它以太网冗余接口。
- 两个成员接口如果都是子接口，则不能是同一主接口的两个子接口，并且其 VLAN 终结配置必须一致。关于 VLAN 终结的详细介绍请参见“二层技术-以太网交换”中的“VLAN 终结”。
- 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 静态邻居表项的出接口。关于 IPv6 静态邻居表项的详细描述请参见“三层技术-IP 业务配置指导”中的“Ipv6 基础”。

表1-1 配置以太网冗余接口

操作	命令	说明
进入系统视图	system-view	-
创建以太网冗余接口，并进入该接口视图	interface reth interface-number	缺省情况下，不存在以太网冗余接口 如果以太网冗余接口已创建，执行该命令，则直接进入该以太网冗余接口的视图
给以太网冗余接口添加成员接口	member interface interface-type interface-number priority priority	缺省情况下，以太网冗余接口下不存在成员接口
（可选）配置以太网冗余接口的期望带宽	bandwidth bandwidth-value	缺省情况下，接口的期望带宽为 10000kbit/s
（可选）配置以太网冗余接口的描述信息	description text	缺省情况下，接口的描述信息为“该接口的接口名 Interface”，比如：Reth-1 Interface
（可选）配置以太网冗余接口的 MTU（Maximum Transmission Unit，最大传输单元）值	mtu size	缺省情况下，以太网冗余接口的 MTU 值为 1500 字节
打开以太网冗余接口	undo shutdown	缺省情况下，以太网冗余接口处于开启状态
（可选）恢复接口的缺省配置	default	-

1.3 以太网冗余接口显示和维护

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以太网冗余接口的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令可以清除以太网冗余接口的统计信息。

表1-2 以太网冗余显示和维护

操作	命令
显示以太网冗余接口的成员接口的信息	display reth interface interface-type interface-number

操作	命令
显示以太网冗余接口的相关信息	display interface [reth [<i>interface-number</i>]] [brief [description down]]
清除以太网冗余接口的统计信息	reset counters interface <i>interface-type</i> <i>interface-number</i>

1.4 以太网冗余接口典型配置举例

1. 组网需求

Device 通过三层聚合接口 RAGG1 和 RAGG2 连接到 Internet, 使用以太网冗余接口技术让 RAGG1 和 RAGG2 互为备份, 正常情况下, 使用 RAGG1 转发报文。

2. 组网图

图1-3 配置以太网冗余接口组网图（集中式 IRF 设备）

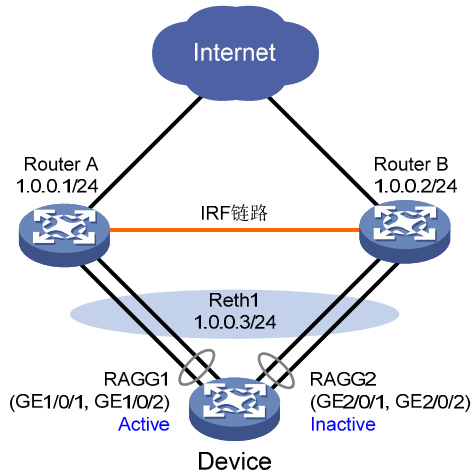
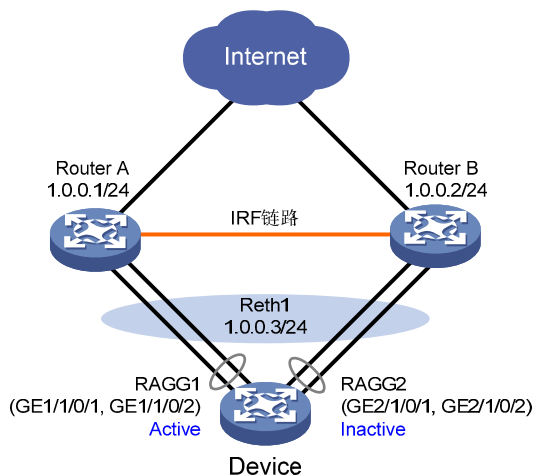


图1-4 配置以太网冗余接口组网图（分布式设备—IRF 模式）



3. 配置步骤（集中式IRF设备）

(1) 配置两条聚合链路

创建三层聚合接口 1。

```
<Sysname> system-view
[Sysname] interface route-aggregation 1
[Sysname-Route-Aggregation1] quit
```

分别将接口 GigabitEthernet1/0/1 和 GigabitEthernet1/0/2 加入到聚合组 1 中。

```
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] port link-aggregation group 1
[Sysname-GigabitEthernet1/0/1] quit
[Sysname] interface gigabitethernet 1/0/2
[Sysname-GigabitEthernet1/0/2] port link-aggregation group 1
[Sysname-GigabitEthernet1/0/2] quit
```

创建三层聚合接口 2。

```
[Sysname] interface route-aggregation 2
[Sysname-Route-Aggregation2] quit
```

分别将接口 GigabitEthernet2/0/1 和 GigabitEthernet2/0/2 加入到聚合组 2 中。

```
[Sysname] interface gigabitethernet 2/0/1
[Sysname-GigabitEthernet2/0/1] port link-aggregation group 2
[Sysname-GigabitEthernet2/0/1] quit
[Sysname] interface gigabitethernet 2/0/2
[Sysname-GigabitEthernet2/0/2] port link-aggregation group 2
[Sysname-GigabitEthernet2/0/2] quit
```

(2) 配置以太网冗余接口

创建以太网冗余接口 1，并为该接口配置 IP 地址和子网掩码。

```
[Sysname] interface reth 1
[Sysname-Reth-1] ip address 1.0.0.3 24
```

将三层聚合接口 1 加入以太网冗余接口 1，并指定优先级为 100；将三层聚合接口 2 加入以太网冗余接口 1，并指定优先级为 80。

```
[Sysname-Reth1] member interface route-aggregation 1 priority 100
[Sysname-Reth1] member interface route-aggregation 2 priority 80
```

4. 验证配置

显示以太网冗余接口 1 的当前状态。由于三层聚合接口 1 的优先级较高，应为激活接口，三层聚合接口 2 应为非激活接口。

```
[Sysname-Reth1] display reth interface reth 1
```

Reth1 :

```
Redundancy group   : N/A
Member             Physical status   Forwarding status   Presence status
Route-aggregation 1 UP                 Active              Normal
Route-aggregation 2 UP                 Inactive            Normal
```

手动关闭三层聚合接口 1，再次显示以太网冗余接口 1 的当前状态。应看到三层聚合接口 2 为激活接口，三层聚合接口 1 为非激活接口。

```
[Sysname-Reth1] quit
[Sysname] interface route-aggregation 1
```

```
[Sysname-Route-Aggregation1] shutdown
```

```
[Sysname-Route-Aggregation1] display reth interface reth 1
```

```
Reth1 :
```

```
Redundancy group : N/A
```

Member	Physical status	Forwarding status	Presence status
RAGG 1	DOWN	Inactive	Normal
RAGG 2	UP	Active	Normal

2 冗余组

2.1 冗余组简介

在如 [图 2-1](#)[图 2-2](#) 所示的组网中，Device A 和 Device B 组成 IRF，双上行到 IP 网络，双下行到 IP 网络。为了便于业务的处理和统计：

- 使用以太网冗余接口技术可以使得报文进入某成员设备，经过处理后，从该成员设备的接口转发出去。
- 使用冗余组技术能够提供监控机制，快速检测上、下行链路是否故障，如果故障，则通知组内所有成员（包括以太网冗余接口，如图中的 Reth1 和 Reth2）整体进行倒换，以便保证倒换后，报文的出接口和入接口仍然在同一台设备上。

图2-1 冗余组典型应用组网图（集中式 IRF 设备）

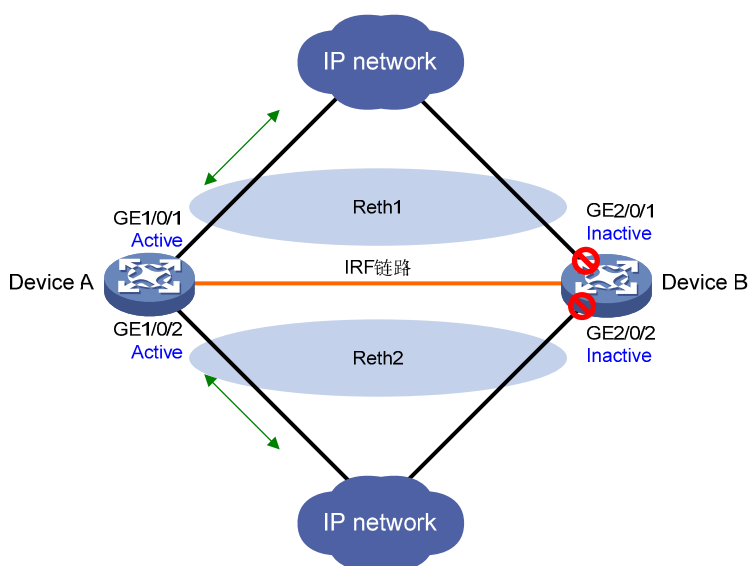
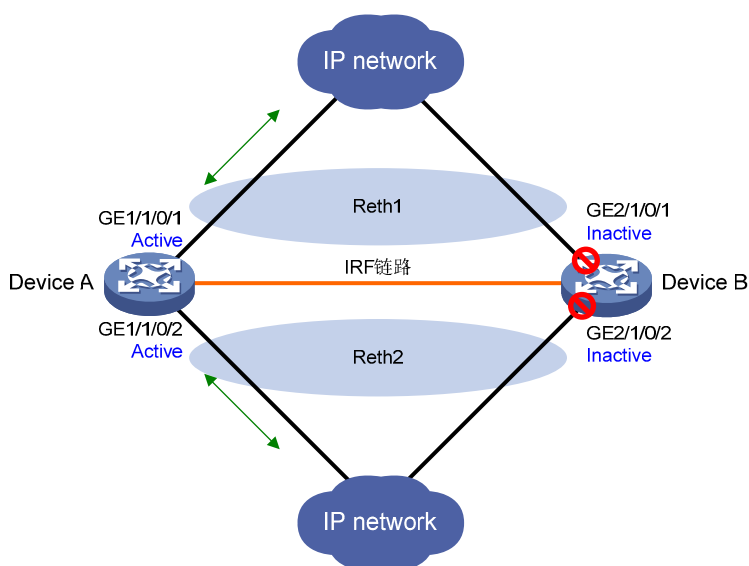


图2-2 冗余组典型应用组网图（分布式设备—IRF 模式）



一个冗余组包含两个冗余组节点以及若干冗余组成员，冗余组成员可以是以太网冗余接口和（或）备份组。一个冗余组节点和一台成员设备绑定，冗余组成员则部署在冗余组节点绑定的成员设备上，冗余组成员的主备状态必须和冗余组节点的主备状态保持一致。

当 IRF 中存在大于两个成员设备时，可以配置多个冗余组，多个冗余组相互独立，本冗余组内的主备状态以及主备倒换不会影响其它冗余组。

2.2 冗余组原理

2.2.1 冗余组节点

冗余组节点和成员设备绑定，冗余组节点通过状态机来决定哪台成员设备处理业务，冗余组节点通过权重值来决定什么时候发生倒换。

1. 冗余组节点的状态

冗余组节点有两个状态：主（Primary）和备（Secondary）。和主节点绑定的成员设备处理业务，转发报文。

冗余组节点的主备状态由以下因素决定：

- 当冗余组节点只绑定了一个成员设备或者绑定的两个成员设备中有一个不能正常工作时，则与能正常工作的成员设备绑定的节点成为主节点。节点能否正常工作由监控机制决定。
- 当冗余组节点绑定的成员设备均能正常工作时：
 - 优先级高的为主节点。优先级可通过命令行配置。
 - 当两个节点的优先级相等时，节点编号小的为主节点。节点编号可通过命令行配置。

2. 冗余组节点的成员接口

冗余组节点下可以配置成员接口，这些接口必须是冗余组节点绑定的成员设备上的接口。两个节点分别和两个成员设备绑定，这样，这两个节点下的成员接口可以互为备份。

如 图 2-1图 2-2 所示，当上下行设备运行了动态路由协议时，可将上下行接口均配置为冗余组节点的成员接口，来实现报文的出接口和入接口在同一台设备上的需求。

当优先级高的节点状态为主时，两个节点下成员接口都会保持UP状态，能正常收发报文，如 图 2-3 图 2-4 所示；当冗余组主备倒换，优先级低的节点状态变为主时，则原主节点的成员接口会被冗余组模块关闭（使用 **display interface** 命令查看时，接口的状态显示为Down（**redundancy down**）），以便上下行设备能感知链路故障并切换路由，从而实现上下行流量从当前主节点的成员接口收发，如 图 2-5图 2-6 所示。

图2-3 两节点均正常时，冗余组节点的成员接口状态示意图（集中式 IRF 设备）

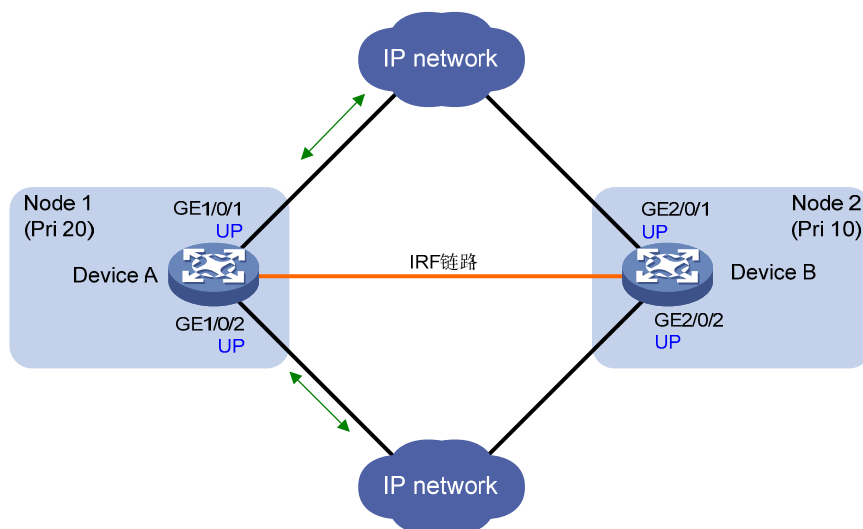


图2-4 两节点均正常时，冗余组节点的成员接口状态示意图（分布式设备—IRF 模式）

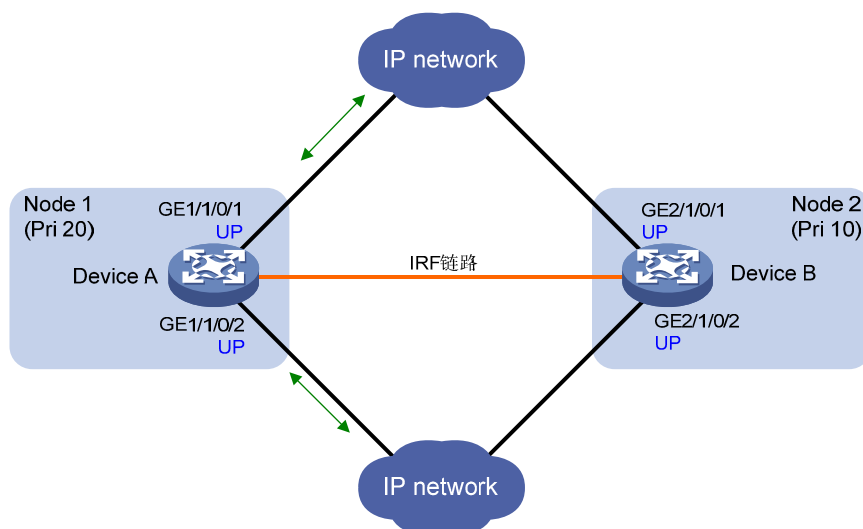


图2-5 优先级高的节点故障时，冗余组节点的成员接口状态示意图（集中式 IRF 设备）

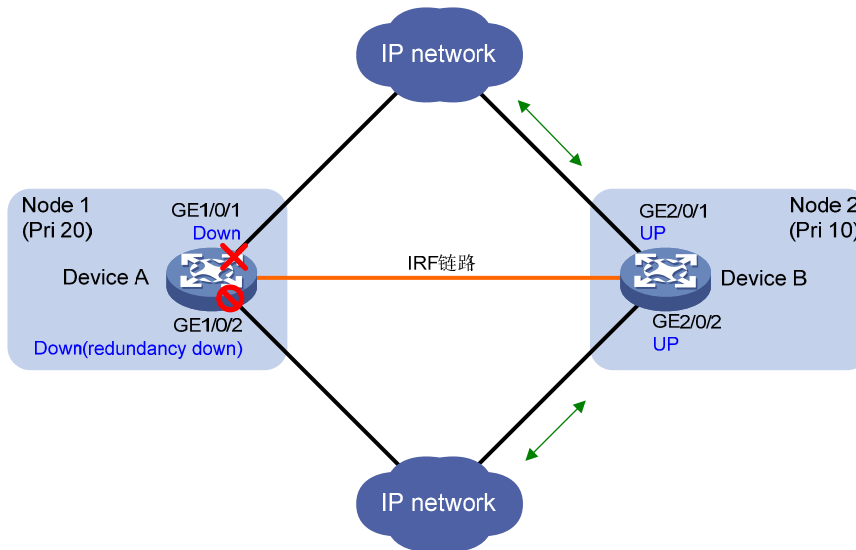
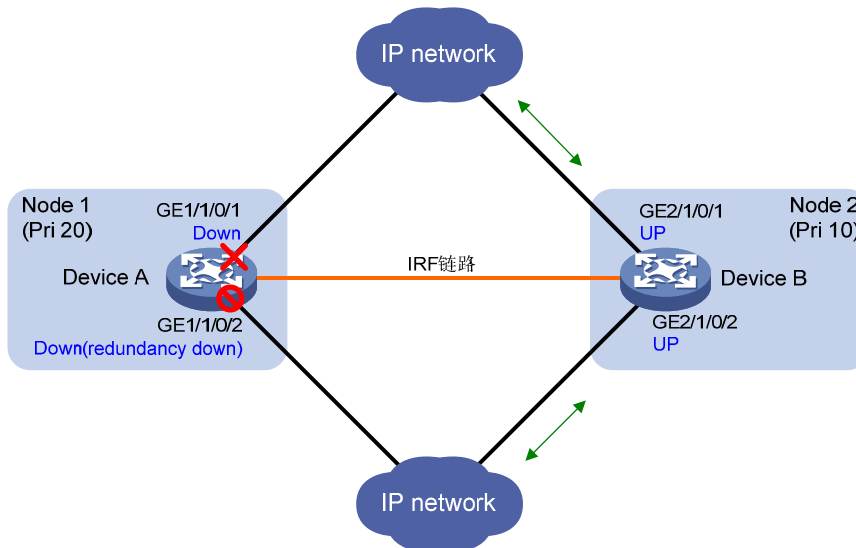


图2-6 优先级高的节点故障时，冗余组节点的成员接口状态示意图（分布式设备—IRF 模式）



3. 冗余组节点的监控机制

冗余组节点通过和 Track 联动来快速检测上、下行链路、接口的状态，从而控制冗余组节点的状态。每个冗余组节点都有权重，缺省值为 255，每个冗余组节点必须关联至少一个 Track 项并配置权重增量。当 Track 项变为 NotReady 或 Negative 状态时，冗余组节点用当前权重减去对应的权重增量获得新的当前权重。当 Track 项变为 Positive 时，冗余组节点用当前权重加上对应的权重增量获得新的当前权重。当新的当前权重小于或等于 0 时，则认为该节点故障，无法正常工作。若 Track 模块尚未启动，则节点绑定的 Track 项状态始终为 Positive。关于 Track 的详细介绍请参见“可靠性配置指导”中的“Track”。

一个冗余组节点可以关联多个 Track 项，用户关联 Track 项时，可以指定一个关联接口。当 Track 项监控的接口为以太网冗余接口的成员接口或是冗余组节点的成员接口时，请将监控接口配置为关联接口。当主节点因为关联接口故障，导致主备倒换时，除了关联接口，原主节点的其他成员接口均会被冗余组模块关闭。关联接口用于触发自动整体倒回（即业务重新回到原主节点处理），原理如下：

在优先级高的节点中，首个状态变为 NotReady 或 Negative 的 Track 项的关联接口会被标识为故障口。当故障口所在的 Track 项恢复为 Positive 状态时，会触发冗余组进行整体倒回。人工触发倒回，或倒回时间配置为 0 时，不标识故障口。

2.2.2 冗余组成员

1. 以太网冗余接口

冗余组下可以配置以太网冗余接口作为成员接口，以太网冗余接口拥有两个成员接口。以太网冗余接口和冗余组节点之间的联动关系为：

- 优先级高的冗余组节点状态为主时，以太网冗余接口下优先级高的成员接口处于激活状态。
- 优先级高的冗余组节点状态为备时，以太网冗余接口下优先级高的成员接口处于非激活状态。

所以，用户需要通过配置保证：

- 以太网冗余接口的成员接口分别处于冗余组节点绑定的两个成员设备上。
- 处于和高优先级绑定的成员设备上的成员接口的优先级更高。

如 [图 2-1图 2-2](#) 所示，当上下行设备没有运行动态路由协议时，比如使用 VRRP 技术，可将上行接口加入一个以太网冗余接口，将下行接口加入另一个以太网冗余接口，并将这两个以太网冗余接口加入一个冗余组，来实现报文的出接口和入接口在同一台设备上的需求。推荐配置和状态变化如 [表 2-1](#) 和 [表 2-2](#) 所示，当 Node 1 为主节点时，和 Node 1 绑定的成员设备上的接口 GE1/1/0/1 和 GE1/1/0/2 处于激活状态；当 Node 2 为主节点时，和 Node 2 绑定的成员设备上的接口 GE2/1/0/1 和 GE2/1/0/2 处于激活状态。

表2-1 冗余组与冗余接口的联动

冗余组	主节点	冗余组节点	以太网冗余接口			
			名称	成员接口	名称	成员接口
Redundancy Group 1	Node 1	Node 1(Pri 20) 绑定成员设备1	Reth1	GE1/1/0/1(Pri 20) (激活状态)	Reth2	GE1/1/0/2(Pri 20) (激活状态)
		Node 2(Pri 10) 绑定成员设备2		GE2/1/0/1(Pri 10) (非激活状态)		GE2/1/0/2(Pri 10) (非激活状态)

表2-2 冗余组与冗余接口的联动（假设 GE1/1/0/1 故障）

冗余组	主节点	冗余组节点	以太网冗余接口			
			名称	成员接口	名称	成员接口
Redundancy Group 1	Node 1	Node 1(Pri 20) 绑定成员设备1	Reth1	GE1/1/0/1(Pri 20) (非激活状态)	Reth2	GE1/1/0/2(Pri 20) (非激活状态)

冗余组	主节点	冗余组节点	以太网冗余接口			
			名称	成员接口	名称	成员接口
		Node 2(Pri 10) 绑定成员设备2		GE2/1/0/1(Pri 10) (激活状态)		GE2/1/0/2(Pri 10) (激活状态)

2. 备份组

冗余组下可以配置备份组作为成员，备份组拥有两个 CPU 成员，用于实现这两个 CPU 上安全业务的备份。备份组和冗余组节点之间的联动关系为：

- 优先级高的冗余组节点状态为主时，备份组里的主成员处理业务。如果备份组里的主成员故障，备成员能正常工作，则使用备成员处理业务。
- 优先级高的冗余组节点状态为备时，备份组里优先级高的成员不处理业务。如果备份组里的备成员故障，主成员能正常工作，则使用主成员处理业务。

所以，用户需要通过配置保证：

- 备份组的两个成员分别处于冗余组节点绑定的两个成员设备上。
- 处于和高优先级绑定的成员设备上的备份组成员属性配置为 Primary。

如 [图 2-1](#) [图 2-2](#) 所示，当需要将安全业务集中在 Device A 上处理，当 Device A 故障，需要迁移到 Device B，为了保证业务处理的连续性，请配置冗余组和备份组。推荐配置和状态变化如 [表 2-3](#) 和 [表 2-4](#) 所示，当 Node 1 为主节点时，和 Node 1 绑定的成员设备上的安全引擎处理业务；当 Node 2 为主节点时，和 Node 2 绑定的成员设备上的安全引擎处理业务。

表2-3 冗余组与备份组的联动

冗余组	主节点	冗余组节点	备份组	
			名称	成员
Redundancy Group 1	Node 1	Node 1(Pri 20) 绑定成员设备1	Failover Group 1	Chassis 1 slot 4 cpu 1 (处理业务)
		Node 2(Pri 10) 绑定成员设备2		Chassis 2 slot 4 cpu 1 (不处理业务)

表2-4 冗余组与备份组的联动（假设 Node 1 无法正常工作）

冗余组	主节点	冗余组节点	备份组	
			名称	成员
Redundancy Group 1	Node 1	Node 1(Pri 20) 绑定成员设备1	Failover Group 1	Chassis 1 slot 4 cpu 1 (不处理业务)
		Node 2(Pri 10) 绑定成员设备2		Chassis 2 slot 4 cpu 1 (处理业务)

2.2.3 冗余组定时器

1. 保持定时器

当网络不稳定，监测接口/链路状态频繁改变，会导致 **Track** 项状态在短时间内频繁改变，连带导致冗余组需要不断的响应主备倒换事件，使用保持定时器可以避免这种情况的发生。当节点完成主备倒换后，系统启动保持定时器。在保持时间内，不允许再次发生主备倒换。

2. 倒回定时器

当冗余组内优先级高的节点倒回条件就绪时（譬如故障恢复），会触发倒回事件，并启动倒回定时器。由于需要整体倒回，在冗余组倒回的过程中会同时触发很多事件（比如接口状态变化等），这些事件的处理需要时间。倒回定时器能够为冗余组提供一段时间，让节点准备完毕后，再将业务从优先级低的节点倒换到优先级高的节点。

2.2.4 冗余组的倒换机制

1. 自动倒换机制

Track 项的状态变化会引起冗余组节点权重的变化，冗余组根据冗余组节点的权重值来判断节点是否能够正常工作。当权重减为 0 或负值时，冗余组认为节点无法正常工作并触发冗余组的整体倒换。

- 如果是将业务从优先级高的节点倒换到优先级低的节点，则系统收到整体倒换请求后，等到保持定时器超时后，进行整体主备倒换。
- 如果是优先级高的节点故障恢复，需要将业务从优先级低的节点倒回，则系统收到整体倒换请求后，等到保持定时器超时后，认为倒回条件就绪，并等到倒回定时器超时后，再进行整体倒回。

2. 手工倒换机制

如果两个节点均能正常工作，但用户需要更换主节点上的硬件，此时，可手工触发倒换，让业务迁移到优先级低的节点。如果两个节点均能正常工作，但用户未配置 **Track** 项关联接口或者将倒回时延配置为 0，系统不能自动倒回时，可手工触发倒回，让业务迁移到优先级高的节点。

自动倒换是权重值触发的，手工倒换是命令行触发的。

2.3 配置冗余组

2.3.1 冗余组配置任务简介

表2-5 冗余组配置任务简介表

配置任务	说明
创建冗余组	必选
配置冗余组节点	必选
配置冗余组的成员接口	可选
配置冗余组成员备份组	可选
配置冗余组定时器	可选
手工触发冗余组倒换	可选

配置任务	说明
开启冗余组告警功能	可选

2.3.2 创建冗余组

表2-6 创建冗余组

操作	命令	说明
进入系统视图	system-view	-
创建冗余组，并进入该冗余组视图	redundancy group <i>group-name</i>	缺省情况下，不存在冗余组 如果冗余组已创建，则执行该命令，直接进入冗余组视图

2.3.3 配置冗余组节点



注意

冗余组节点关联的 Track 项必须是监控接口物理状态的 Track 项，否则，冗余组将不能正常工作。

配置冗余组节点时，请遵循以下要求：

- 一个冗余组下最多可创建两个冗余组节点。不同冗余组下冗余组节点的编号可以相同。
- 冗余组节点必须和成员设备绑定，一个冗余组节点只能和一个成员设备绑定，一台成员设备只能和一个冗余组节点绑定。当冗余组节点下存在成员接口或 Track 项时，用户不能取消绑定。
- 用户可根据组网需求，给冗余组节点配置成员接口。请先绑定成员设备再添加成员接口，并确保各节点的成员接口是本节点绑定的成员设备上的物理接口，不能是聚合接口和子接口，不能是以太网冗余接口的成员接口。物理接口只能属于一个冗余组节点，不能同时加入其它节点。
- 同一个 Track 项不能与同一冗余组下的两个冗余组节点都关联。当已将某物理接口配置为某冗余组内高优先级冗余组节点的成员接口，或者为某冗余组内以太网冗余接口的高优先级成员接口时，请不要将该物理接口的子接口配置为该冗余组内高优先级冗余组节点的 Track 项关联接口。因为物理接口被协议关闭时，会导致其子接口状态为 Down，该子接口将无法触发自动倒回，此时，需要手工倒回。

表2-7 创建冗余组节点

操作	命令	说明
进入系统视图	system-view	-
进入冗余组视图	redundancy group <i>group-name</i>	-
创建冗余组节点，并进入冗余组节点视图	node <i>node-id</i>	缺省情况下，不存在冗余组节点

操作	命令	说明
配置冗余组节点的优先级	priority <i>priority</i>	缺省情况下，冗余组节点的优先级为1
将冗余组节点和IRF成员设备绑定（分布式设备—IRF模式）	bind chassis <i>chassis-number</i>	缺省情况下，冗余组节点未绑定成员设备
将冗余组节点和IRF成员设备绑定（集中式IRF设备）	bind slot <i>slot-number</i>	缺省情况下，冗余组节点未绑定成员设备
（可选）为冗余组节点添加成员接口	node-member interface <i>interface-type interface-number</i>	缺省情况下，冗余组节点下不存在成员接口
关联Track项	track <i>track-entry-number</i> [reduced weight-reduced] [interface <i>interface-type interface-number</i>]	缺省情况下，冗余组节点下未关联Track项 建议先创建Track项，再将该Track项和冗余组关联。否则，可能会导致冗余组没有有效的Track项而触发倒换

2.3.4 配置冗余组的成员接口

表2-8 配置冗余组成员接口

操作	命令	说明
进入系统视图	system-view	-
进入冗余组视图	redundancy group <i>group-name</i>	-
将以太网冗余接口加入冗余组	member interface reth <i>interface-number</i>	缺省情况下，冗余组下不存在以太网冗余接口

2.3.5 配置冗余组成员备份组

设备各款型对于本节所描述的特性支持情况有所不同，详细差异信息如下：

型号	特性	描述
T5000-M06	配置冗余组成员备份组	支持
T5000-G20		不支持
T1000-G20/G50/G60/G80		不支持

表2-9 配置冗余组备份组

操作	命令	说明
进入系统视图	system-view	-
进入冗余组视图	redundancy group <i>group-name</i>	-

操作	命令	说明
将备份组加入冗余组	member failover group <i>group-name</i>	缺省情况下，冗余组下不存在备份组添加的备份组必须是已经创建的。否则，本命令将执行失败

2.3.6 配置冗余组定时器

表2-10 配置冗余组定时器

操作	命令	说明
进入系统视图	system-view	-
进入冗余组视图	redundancy group <i>group-name</i>	-
配置冗余组节点状态的保持时间，这段时间内不能发生主备倒换	hold-down-interval <i>second</i>	缺省情况下，冗余组节点状态的保持时间为1秒
配置冗余组节点的倒回延时	preempt-delay <i>min</i>	缺省情况下，冗余组节点的倒回延时为1分钟 如果将倒回时间配置为0，则表示不倒回

2.3.7 手工触发冗余组倒换

表2-11 手工触发冗余组倒换

操作	命令	说明
进入系统视图	system-view	-
进入冗余组视图	redundancy group <i>group-name</i>	-
手工触发冗余组进行主备倒换，让冗余组工作在优先级低的节点	switchover request	二者选其一
手工触发一次冗余组倒回，让冗余组工作在优先级高的节点	switchover reset	

2.3.8 开启冗余组告警功能

开启冗余组告警功能后，在冗余组人工倒换、故障接口恢复、故障接口生成时，会生成告警信息，并将该信息发送到设备的 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特性。

有关告警信息的详细描述，请参见“网络管理和监控配置指导”中的“SNMP”。

表2-12 开启冗余组告警功能

操作	命令	说明
进入系统视图	system-view	-

操作	命令	说明
开启冗余组告警功能	snmp-agent trap enable rddc	缺省情况下，冗余组告警功能处于开启状态

2.4 冗余组显示和维护

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冗余组的运行情况，通过查看显示信息验证配置的效果。

表2-13 冗余组显示和维护

操作	命令
显示冗余组的相关信息	display redundancy group [group-name]

2.5 冗余组典型配置举例（集中式IRF设备）

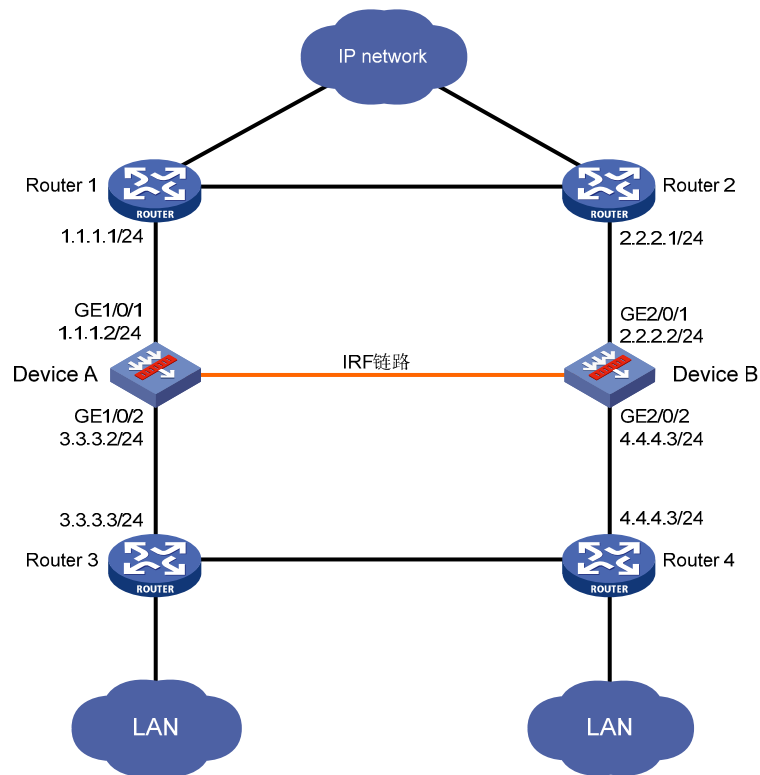
2.5.1 工作在三层，上下行分别连接两台路由器，两台路由器接口不在同一网段

1. 组网需求

- 如 [图 2-7](#) 所示，Device A和Device B组成IRF，Router 1 和IRF相连的接口与Router 2 和IRF相连的接口不在同一网段，Router 3 和IRF相连的接口与Router 4 和IRF相连的接口不在同一网段。
- 正常情况下，流量走 Router 1—Device A—Router 3；当这条通道上的任一链路或者设备故障时，流量切换到 Router 2—Device B—Router 4。正常通道故障恢复时，流量再切回。

2. 组网图

图2-7 配置组网图



3. 配置步骤



说明

请参见“虚拟化配置指导”中的“IRF”，将 Device A 和 Device B 组成 IRF。本文只描述 IRF 组成后，在 IRF 上配置冗余组。

(1) 配置三层接口及其 IP 地址

请按照图 2-7 所示，配置各接口的 IP 地址和子网掩码，具体配置过程略。

(2) 配置 Track，监测上、下行接口的物理状态。

```
<Sysname> system-view
[Sysname] track 1 interface gigabitethernet 1/0/1 physical
[Sysname] track 2 interface gigabitethernet 1/0/2 physical
[Sysname] track 3 interface gigabitethernet 2/0/1 physical
[Sysname] track 4 interface gigabitethernet 2/0/2 physical
```

(3) 配置冗余组。

创建 Node 1，Node 1 和 Device A 绑定，为主节点，成员接口为 GE1/0/1 和 GE1/0/2。关联的 Track 项为 1 和 2。

```
[Sysname] redundancy group aaa
[Sysname] redundancy-group-aaa] node 1
```

```
[Sysname-redundancy-group-aaa-node1] bind slot 1
[Sysname-redundancy-group-aaa-node1] priority 100
[Sysname-redundancy-group-aaa-node1] node-member interface gigabitethernet 1/0/1
[Sysname-redundancy-group-aaa-node1] node-member interface gigabitethernet 1/0/2
[Sysname-redundancy-group-aaa-node1] track 1 interface gigabitethernet 1/0/1
[Sysname-redundancy-group-aaa-node1] track 2 interface gigabitethernet 1/0/2
[Sysname-redundancy-group-aaa-node1] quit
```

创建 Node 2，Node 2 和 Device B 绑定，为备节点，成员接口为 GE2/0/1 和 GE2/0/2。关联的 Track 项为 3 和 4。

```
[Sysname]redundancy-group-aaa] node 2
[Sysname-redundancy-group-aaa-node2] bind slot 2
[Sysname-redundancy-group-aaa-node2] priority 50
[Sysname-redundancy-group-aaa-node2] node-member interface gigabitethernet 2/0/1
[Sysname-redundancy-group-aaa-node2] node-member interface gigabitethernet 2/0/2
[Sysname-redundancy-group-aaa-node2] track 3 interface gigabitethernet 2/0/1
[Sysname-redundancy-group-aaa-node2] track 4 interface gigabitethernet 2/0/2
[Sysname-redundancy-group-aaa-node2] quit
```

4. 验证配置

(1) 缺省情况下的显示信息

显示冗余组信息。可以看到优先级高的 Node 1 为主节点，Node 1 和 Node 2 下面的成员接口都处于 UP 状态。

```
[Sysname-redundancy-group-aaa] display redundancy group aaa
```

```
Redundancy group aaa (ID 1):
```

Node ID	Slot	Priority	Status	Track weight
1	Slot1	100	Primary	255
2	Slot2	50	Secondary	255

```
Preempt delay time remained      : 0    min
Preempt delay timer setting      : 1    min
Remaining hold-down time         : 0    sec
Hold-down timer setting          : 1    sec
Manual switchover requeset       : No
```

```
Member interfaces:
```

```
Member failover groups:
```

```
Node 1:
```

Node member	Physical status
GE1/0/1	UP
GE1/0/2	UP

```
Track info:
```

Track	Status	Reduced weight	Interface
1	Positive	255	GE1/0/1
2	Positive	255	GE1/0/2

```
Node 2:
```

Node member	Physical status
GE2/0/1	UP

```

GE2/0/2  UP
Track info:
  Track   Status      Reduced weight  Interface
  3       Positive    255             GE2/0/1
  4       Positive    255             GE2/0/2

```

(2) 冗余组内主备倒换后的显示信息

手工关闭接口 GE1/0/2，显示冗余组信息。可以看到优先级低的 Node 2 为主节点，Node 1 的成员接口 GE1/0/2 故障 (DOWN)，GE1/0/1 被协议关闭 (DOWN(redundancy down))，Node 2 的成员接口转发报文。

```

[Sysname-redundancy-group-aaa] quit
[Sysname] interface gigabitethernet 1/0/2
[Sysname-GigabitEthernet1/0/2] shutdown
[Sysname-GigabitEthernet1/0/2] display redundancy group aaa
Redundancy group aaa (ID 1):
  Node ID   Slot      Priority  Status      Track weight
  1         Slot1     100      Secondary   -255
  2         Slot2     50       Primary     255

```

```

Preempt delay time remained      : 0    min
Preempt delay timer setting      : 1    min
Remaining hold-down time         : 0    sec
Hold-down timer setting          : 1    sec
Manual switchover request       : No

```

```

Member interfaces:
Member failover groups:

```

```

Node 1:
  Node member   Physical status
  GE1/0/1      DOWN(redundancy down)
  GE1/0/2      DOWN
Track info:
  Track   Status      Reduced weight  Interface
  1       Negative    255             GE1/0/1
  2       Negative    255             GE1/0/2 (Fault)

```

```

Node 2:
  Node member   Physical status
  GE2/0/1      UP
  GE2/0/2      UP
Track info:
  Track   Status      Reduced weight  Interface
  3       Positive    255             GE2/0/1
  4       Positive    255             GE2/0/2

```

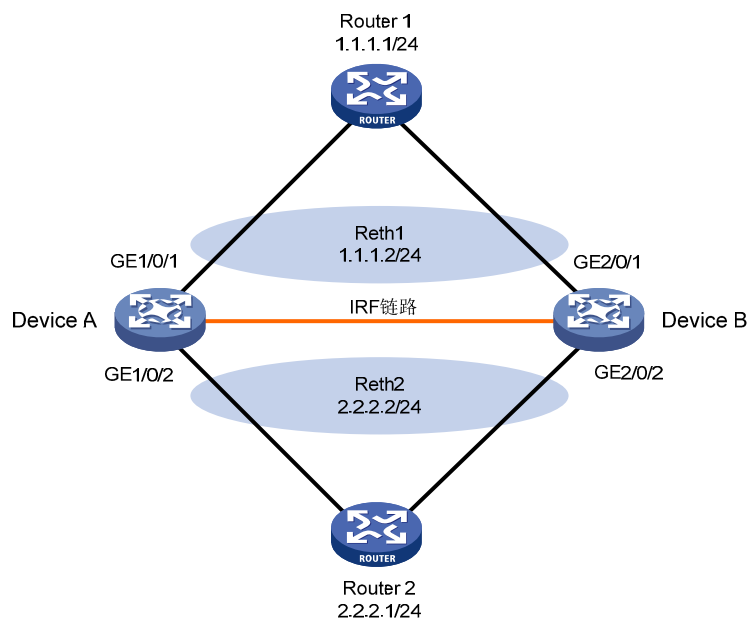
2.5.2 工作在三层，上下行分别连接一台路由器

1. 组网需求

- 如 图 2-8 所示，Device A和Device B组成IRF，Device A和Device B分别用一个接口连接上行设备Router 1，再分别用一个接口连接下行设备Router 2。
- Device A 和 Device B 分别使用三层聚合接口连接 Router 1 和 Router 2。
- 正常情况下，流量走 Router 1——Device A——Router 2；当这条通道上的任一链路或者设备故障时，流量切换到 Router 1——Device B——Router 2。正常通道故障恢复时，流量再切回。

2. 组网图

图2-8 配置组网图



3. 配置步骤



说明

请参见“虚拟化配置指导”中的“IRF”，将 Device A 和 Device B 组成 IRF。本文只描述 IRF 组成后，在 IRF 上配置冗余组。

(1) 配置以太网冗余接口

创建 Reth1，IP 地址为 1.1.1.2/24，成员接口为 GE1/0/1 和 GE2/0/1，其中 GE1/0/1 的优先级为 255，GE2/0/1 的优先级为 50。

```
<Sysname> system-view
[Sysname] interface reth 1
[Sysname-Reth1] ip address 1.1.1.2 24
[Sysname-Reth1] member interface gigabitethernet 1/0/1 priority 255
[Sysname-Reth1] member interface gigabitethernet 2/0/1 priority 50
```

```
[Sysname-Reth1] quit
# 创建 Reth2, IP 地址为 2.2.2.2/24, 成员接口为 GE1/0/2 和 GE2/0/2, 其中 GE1/0/2 的优先级为 255, GE2/0/2 的优先级为 50。
```

```
[Sysname] interface reth 2
[Sysname-Reth2] ip address 2.2.2.2 24
[Sysname-Reth2] member interface gigabitethernet 1/0/2 priority 255
[Sysname-Reth2] member interface gigabitethernet 2/0/2 priority 50
[Sysname-Reth2] quit
```

(2) 配置 Track, 监测上、下行接口的物理状态。

```
[Sysname] track 1 interface gigabitethernet 1/0/1 physical
[Sysname] track 2 interface gigabitethernet 1/0/2 physical
[Sysname] track 3 interface gigabitethernet 2/0/1 physical
[Sysname] track 4 interface gigabitethernet 2/0/2 physical
```

(3) 配置冗余组

创建 Node 1, Node 1 和 Device A 绑定, 为主节点。关联的 Track 项为 1 和 2。

```
[Sysname] redundancy group aaa
[Sysname-redundancy-group-aaa] node 1
[Sysname-redundancy-group-aaa-node1] bind slot 1
[Sysname-redundancy-group-aaa-node1] priority 100
[Sysname-redundancy-group-aaa-node1] track 1 interface gigabitethernet 1/0/1
[Sysname-redundancy-group-aaa-node1] track 2 interface gigabitethernet 1/0/2
[Sysname-redundancy-group-aaa-node1] quit
```

创建 Node 2, Node 2 和 Device B 绑定, 为备节点。关联的 Track 项为 3 和 4。

```
[Sysname-redundancy-group-aaa] node 2
[Sysname-redundancy-group-aaa-node2] bind slot 2
[Sysname-redundancy-group-aaa-node2] priority 50
[Sysname-redundancy-group-aaa-node2] track 3 interface gigabitethernet 2/0/1
[Sysname-redundancy-group-aaa-node2] track 4 interface gigabitethernet 2/0/2
[Sysname-redundancy-group-aaa-node2] quit
```

将 Reth1 和 Reth2 添加到冗余组中。

```
[Sysname-redundancy-group-aaa] member interface reth 1
[Sysname-redundancy-group-aaa] member interface reth 2
[Sysname-redundancy-group-aaa] quit
```

4. 验证配置

(1) 缺省情况下的显示信息

显示冗余组信息。可以看到优先级高的 Node 1 为主节点, Reth1 和 Reth2 中优先级高的成员接口处于激活状态。

```
[Sysname] display redundancy group aaa
Redundancy group aaa (ID 1):
   Node ID   Slot      Priority  Status      Track weight
-----
   Node1     Slot1     100      Primary     255
   Node2     Slot2     50       Secondary   255

Preempt delay time remained : 0 min
Preempt delay timer setting  : 1 min
```

```

Remaining hold-down time      : 0    sec
Hold-down timer setting      : 300  sec
Manual switchover request    : No

```

Member interfaces:

```

    Reth1          Reth2

```

Member failover groups:

Node 1:

Track info:

Track	Status	Reduced weight	Interface
1	Positive	255	GE1/0/1
2	Positive	255	GE1/0/2

Node 2:

Track info:

Track	Status	Reduced weight	Interface
3	Positive	255	GE2/0/1
4	Positive	255	GE2/0/2

显示 Reth 信息。可以看到 Reth1 和 Reth2 中优先级高的成员接口处于激活状态。

```
[Sysname] display reth interface reth 1
```

Reth1 :

Redundancy group : aaa

Member	Physical status	Forwarding status	Presence status
GE1/0/1	UP	Active	Normal
GE2/0/1	UP	Inactive	Normal

```
[Sysname] display reth interface reth 2
```

Reth2 :

Redundancy group : aaa

Member	Physical status	Forwarding status	Presence status
GE1/0/2	UP	Active	Normal
GE2/0/2	UP	Inactive	Normal

(2) 冗余组内主备倒换后的显示信息

手工关闭接口 GE1/0/2，显示冗余组信息。可以看到优先级低的 Node 2 为主节点。

```
[Sysname] interface gigabitethernet 1/0/2
```

```
[Sysname-GigabitEthernet1/0/2] shutdown
```

```
[Sysname-GigabitEthernet1/0/2] display redundancy group aaa
```

Redundancy group aaa (ID 1):

Node ID	Slot	Priority	Status	Track weight
Node1	Slot1	100	Secondary	-255
Node2	Slot2	50	Primary	255

```
Preempt delay time remained : 0    min
```

```
Preempt delay timer setting : 1    min
```

```
Remaining hold-down time   : 0    sec
```

```
Hold-down timer setting    : 300  sec
```

```
Manual switchover request  : No
```

Member interfaces:


```
Reth1          Reth2
Member failover groups:
```

```
Node 1:
```

```
Track info:
```

Track	Status	Reduced weight	Interface
1	Negative	255	GE1/0/1
2	Negative	255	GE1/0/2(Fault)

```
Node 2:
```

```
Track info:
```

Track	Status	Reduced weight	Interface
3	Positive	255	GE2/0/1
4	Positive	255	GE2/0/2

显示 Reth 的信息。Reth2 下的接口 GE1/0/2 故障 (DOWN)，Reth1 下的接口 GE1/0/1 被协议关闭 (DOWN(redundancy down))。GE2/0/1 和 GE2/0/2 激活。

```
[Sysname-GigabitEthernet1/0/2] display reth interface reth 1
```

```
Reth1 :
```

```
Redundancy group : aaa
```

Member	Physical status	Forwarding status	Presence status
GE1/0/1	DOWN(redundancy down)	Inactive	Normal
GE2/0/1	UP	Active	Normal

```
[Sysname-GigabitEthernet1/0/2] display reth interface reth 2
```

```
Reth2 :
```

```
Redundancy group : aaa
```

Member	Physical status	Forwarding status	Presence status
GE1/0/2	DOWN	Inactive	Normal
GE2/0/2	UP	Active	Normal

2.6 冗余组典型配置举例（分布式设备—IRF模式）

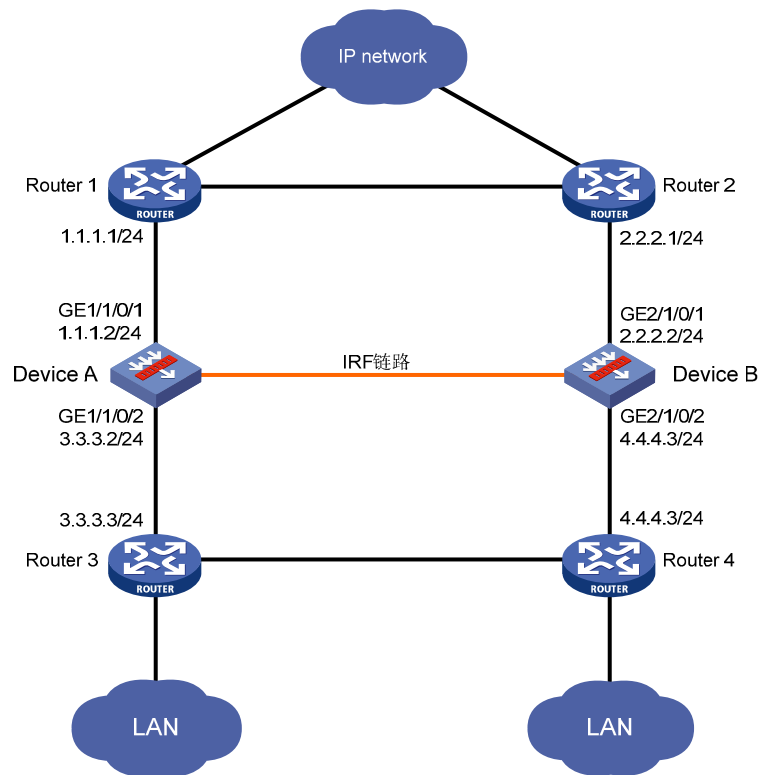
2.6.1 工作在三层，上下行分别连接两台路由器，两台路由器接口不在同一网段

1. 组网需求

- 如 [图 2-9](#) 所示，Device A 和 Device B 组成 IRF，Router 1 和 IRF 相连的接口与 Router 2 和 IRF 相连的接口不在同一网段，Router 3 和 IRF 相连的接口与 Router 4 和 IRF 相连的接口不在同一网段。
- 正常情况下，流量走 Router 1—Device A—Router 3；当这条通道上的任一链路或者设备故障时，流量切换到 Router 2—Device B—Router 4。正常通道故障恢复时，流量再切回。

2. 组网图

图2-9 配置组网图



3. 配置步骤



说明

请参见“虚拟化配置指导”中的“IRF”，将 Device A 和 Device B 组成 IRF。本文只描述 IRF 组成后，在 IRF 上配置冗余组。

(1) 配置三层接口及其 IP 地址

请按照图 2-9 所示，配置各接口的 IP 地址和子网掩码，具体配置过程略。

(2) 配置 Track，监测上、下行接口的以及安全引擎的状态。

```
<Sysname> system-view
[Sysname] track 1 interface gigabitethernet 1/1/0/1 physical
[Sysname] track 2 interface gigabitethernet 1/1/0/2 physical
[Sysname] track 3 interface gigabitethernet 2/1/0/1 physical
[Sysname] track 4 interface gigabitethernet 2/1/0/2 physical
[Sysname] track 5 interface blade 1/4/0/1 physical
[Sysname] track 6 interface blade 2/4/0/1 physical
```

(3) 配置备份组，并指定 Device A 上的安全引擎为主，Device B 上的安全引擎为备。

```
[Sysname] failover group group1
[Sysname-failover-group-group1] bind chassis 1 slot 4 cpu 1 primary
[Sysname-failover-group-group1] bind chassis 2 slot 4 cpu 1 secondary
```

```
[Sysname-failover-group-group1] quit
```

(4) 配置冗余组。

创建 Node 1，Node 1 和 Device A 绑定，为主节点，成员接口为 GE1/1/0/1 和 GE1/1/0/2。关联的 Track 项为 1、2 和 5。

```
[Sysname] redundancy group aaa
[Sysname]redundancy-group-aaa] node 1
[Sysname-redundancy-group-aaa-node1] bind chassis 1
[Sysname-redundancy-group-aaa-node1] priority 100
[Sysname-redundancy-group-aaa-node1] node-member interface gigabitethernet 1/1/0/1
[Sysname-redundancy-group-aaa-node1] node-member interface gigabitethernet 1/1/0/2
[Sysname-redundancy-group-aaa-node1] track 1 interface gigabitethernet 1/1/0/1
[Sysname-redundancy-group-aaa-node1] track 2 interface gigabitethernet 1/1/0/2
[Sysname-redundancy-group-aaa-node1] track 5 interface blade 1/4/0/1
[Sysname-redundancy-group-aaa-node1] quit
```

创建 Node 2，Node 2 和 Device B 绑定，为备节点，成员接口为 GE2/1/0/1 和 GE2/1/0/2。关联的 Track 项为 3、4 和 6。

```
[Sysname]redundancy-group-aaa] node 2
[Sysname-redundancy-group-aaa-node2] bind chassis 2
[Sysname-redundancy-group-aaa-node2] priority 50
[Sysname-redundancy-group-aaa-node2] node-member interface gigabitethernet 2/1/0/1
[Sysname-redundancy-group-aaa-node2] node-member interface gigabitethernet 2/1/0/2
[Sysname-redundancy-group-aaa-node2] track 3 interface gigabitethernet 2/1/0/1
[Sysname-redundancy-group-aaa-node2] track 4 interface gigabitethernet 2/1/0/2
[Sysname-redundancy-group-aaa-node2] track 6 interface blade 2/4/0/1
[Sysname-redundancy-group-aaa-node2] quit
```

将备份组 1 添加到冗余组中。

```
[Sysname-redundancy-group-aaa] member failover group group1
[Sysname-redundancy-group-aaa] quit
```

4. 验证配置

(1) 缺省情况下的显示信息

显示冗余组信息。可以看到优先级高的 Node 1 为主节点，Node 1 和 Node 2 下面的成员接口都处于 UP 状态。

```
[Sysname-redundancy-group-aaa] display redundancy group aaa
```

```
Redundancy group aaa (ID 1):
```

Node ID	Chassis	Priority	Status	Track weight
1	Chassis1	100	Primary	255
2	Chassis4	50	Secondary	255

```
Preempt delay time remained      : 0 min
```

```
Preempt delay timer setting      : 1 min
```

```
Remaining hold-down time        : 0 sec
```

```
Hold-down timer setting         : 1 sec
```

```
Manual switchover request       : No
```

```
Member interfaces:
```

```
Member failover groups:
```

group1

Node 1:

```
Node member      Physical status
-----
GE1/1/0/1      UP
GE1/1/0/2      UP
Track info:
Track   Status      Reduced weight  Interface
-----
1       Positive    255             GE1/1/0/1
2       Positive    255             GE1/1/0/2
5       Positive    255             Blade1/4/0/1
```

Node 2:

```
Node member      Physical status
-----
GE2/1/0/1      UP
GE2/1/0/2      UP
Track info:
Track   Status      Reduced weight  Interface
-----
3       Positive    255             GE2/1/0/1
4       Positive    255             GE2/1/0/2
6       Positive    255             Blade2/4/0/1
```

显示备份组信息。可以看到备份组中配置为 **Primary** 的安全引擎处理业务。

```
[Sysname-redundancy-group-aaa] display failover group group1
Stateful failover group information:
```

ID	Name	Primary	Secondary	Active Status
255	group1	1/4.1	2/4.1	Primary

(2) 冗余组内主备倒换后的显示信息

手工关闭接口 **GE1/1/0/2**，显示冗余组信息。可以看到优先级低的 **Node 2** 为主节点，**Node 1** 的成员接口 **GE1/1/0/2** 故障 (**DOWN**)，**GE1/1/0/1** 被协议关闭 (**DOWN(redundancy down)**)，**Node 2** 的成员接口转发报文。

```
[Sysname-redundancy-group-aaa] quit
[Sysname] interface gigabitethernet 1/1/0/2
[Sysname-GigabitEthernet1/1/0/2] shutdown
[Sysname-GigabitEthernet1/1/0/2] display redundancy group aaa
Redundancy group aaa (ID 1):
```

Node ID	Chassis	Priority	Status	Track weight
1	Chassis1	100	Secondary	-255
2	Chassis4	50	Primary	255

```
Preempt delay time remained      : 0 min
Preempt delay timer setting      : 1 min
Remaining hold-down time         : 0 sec
Hold-down timer setting          : 1 sec
Manual switchover requeset       : No
```

Member interfaces:

Member failover groups:

group1

Node 1:

Node member	Physical status
GE1/1/0/1	DOWN(redundancy down)
GE1/1/0/2	DOWN

Track info:

Track	Status	Reduced weight	Interface
1	Negative	255	GE1/1/0/1
2	Negative	255	GE1/1/0/2 (Fault)
5	Positive	255	Blade1/4/0/1

Node 2:

Node member	Physical status
GE2/1/0/1	UP
GE2/1/0/2	UP

Track info:

Track	Status	Reduced weight	Interface
3	Positive	255	GE2/1/0/1
4	Positive	255	GE2/1/0/2
6	Positive	255	Blade2/4/0/1

显示备份组信息。可以看到备份组中配置为 **Secondary** 的安全引擎处理业务。

```
[Sysname-GigabitEthernet1/1/0/2] display failover group group1
```

Stateful failover group information:

ID	Name	Primary	Secondary	Active Status
255	group1	1/4.1	2/4.1	Secondary

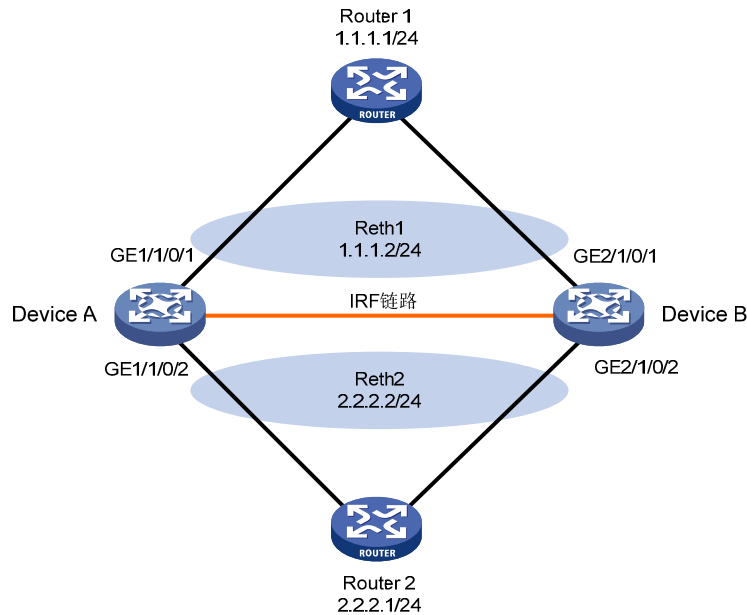
2.6.2 工作在三层，上下行分别连接一台路由器

1. 组网需求

- 如 [图 2-10](#) 所示，Device A 和 Device B 组成 IRF，Device A 和 Device B 分别用一个接口连接上行设备 Router 1，再分别用一个接口连接下行设备 Router 2。
- Router 1 和 Router 2 分别使用三层聚合接口连接 Device A 和 Device B。
- 正常情况下，流量走 Router 1—Device A—Router 2；当这条通道上的任一链路或者设备故障时，流量切换到 Router 1—Device B—Router 2。正常通道故障恢复时，流量再切回。

2. 组网图

图2-10 配置组网图



3. 配置步骤



说明

请参见“虚拟化配置指导”中的“IRF”，将 Device A 和 Device B 组成 IRF。本文只描述 IRF 组成后，在 IRF 上配置冗余组。

(1) 配置以太网冗余接口

创建 Reth1，IP 地址为 1.1.1.2/24，成员接口为 GE1/1/0/1 和 GE2/1/0/1，其中 GE1/1/0/1 的优先级为 255，GE2/1/0/1 的优先级为 50。

```
<Sysname> system-view
[Sysname] interface reth 1
[Sysname-Reth1] ip address 1.1.1.2 24
[Sysname-Reth1] member interface gigabitethernet 1/1/0/1 priority 255
[Sysname-Reth1] member interface gigabitethernet 2/1/0/1 priority 50
[Sysname-Reth1] quit
```

创建 Reth2，IP 地址为 2.2.2.2/24，成员接口为 GE1/1/0/2 和 GE2/1/0/2，其中 GE1/1/0/2 的优先级为 255，GE2/1/0/2 的优先级为 50。

```
[Sysname] interface reth 2
[Sysname-Reth2] ip address 2.2.2.2 24
[Sysname-Reth2] member interface gigabitethernet 1/1/0/2 priority 255
[Sysname-Reth2] member interface gigabitethernet 2/1/0/2 priority 50
[Sysname-Reth2] quit
```

(2) 配置 Track，监测上、下行接口的状态以及安全引擎的状态。

```
[Sysname] track 1 interface gigabitethernet 1/1/0/1 physical
```

```
[Sysname] track 2 interface gigabitethernet 1/1/0/2 physical
[Sysname] track 3 interface gigabitethernet 2/1/0/1 physical
[Sysname] track 4 interface gigabitethernet 2/1/0/2 physical
[Sysname] track 5 interface blade 1/4/0/1 physical
[Sysname] track 6 interface blade 2/4/0/1 physical
```

(3) 配置备份组，并指定 Device A 上的安全引擎为主，Device B 上的安全引擎为备。

```
[Sysname] failover group group1
[Sysname-failover-group-group1] bind chassis 1 slot 4 cpu 1 primary
[Sysname-failover-group-group1] bind chassis 2 slot 4 cpu 1 secondary
[Sysname-failover-group-group1] quit
```

(4) 配置冗余组

创建 Node 1，Node 1 和 Device A 绑定，为主节点。关联的 Track 项为 1、2 和 5。

```
[Sysname] redundancy group aaa
[Sysname-redundancy-group-aaa] node 1
[Sysname-redundancy-group-aaa-node1] bind chassis 1
[Sysname-redundancy-group-aaa-node1] priority 100
[Sysname-redundancy-group-aaa-node1] track 1 interface gigabitethernet 1/1/0/1
[Sysname-redundancy-group-aaa-node1] track 2 interface gigabitethernet 1/1/0/2
[Sysname-redundancy-group-aaa-node1] track 5 interface blade 1/4/0/1
[Sysname-redundancy-group-aaa-node1] quit
```

创建 Node 2，Node 2 和 Device B 绑定，为备节点。关联的 Track 项为 3、4 和 6。

```
[Sysname-redundancy-group-aaa] node 2
[Sysname-redundancy-group-aaa-node2] bind chassis 2
[Sysname-redundancy-group-aaa-node2] priority 50
[Sysname-redundancy-group-aaa-node2] track 3 interface gigabitethernet 2/1/0/1
[Sysname-redundancy-group-aaa-node2] track 4 interface gigabitethernet 2/1/0/2
[Sysname-redundancy-group-aaa-node2] track 6 interface blade 2/4/0/1
[Sysname-redundancy-group-aaa-node2] quit
```

将 Reth1、Reth2 和备份组 1 添加到冗余组中。

```
[Sysname-redundancy-group-aaa] member interface reth 1
[Sysname-redundancy-group-aaa] member interface reth 2
[Sysname-redundancy-group-aaa] member failover group group1
[Sysname-redundancy-group-aaa] quit
```

4. 验证配置

(1) 缺省情况下的显示信息

显示冗余组信息。可以看到优先级高的 Node 1 为主节点，Reth1 和 Reth2 中优先级高的成员接口处于激活状态。

```
[Sysname] display redundancy group aaa
```

```
Redundancy group aaa (ID 1):
```

Node ID	Chassis	Priority	Status	Track weight
Node1	chassis1	100	Primary	255
Node2	chassis2	50	Secondary	255

```
Preempt delay time remained : 0 min
```

```
Preempt delay timer setting : 1 min
```

```
Remaining hold-down time : 0 sec
```

```
Hold-down timer setting      : 300 sec
Manual switchover request   : No
```

Member interfaces:

```
Reth1      Reth2
```

Member failover groups:

```
group1
```

Node 1:

Track info:

Track	Status	Reduced weight	Interface
1	Positive	255	GE1/1/0/1
2	Positive	255	GE1/1/0/2
5	Positive	255	Blade1/4/0/1

Node 2:

Track info:

Track	Status	Reduced weight	Interface
3	Positive	255	GE2/1/0/1
4	Positive	255	GE2/1/0/2
6	Positive	255	Blade2/4/0/1

显示 Reth 信息。可以看到 Reth1 和 Reth2 中优先级高的成员接口处于激活状态。

```
[Sysname] display reth interface reth 1
```

Reth1 :

```
Redundancy group : aaa
```

Member	Physical status	Forwarding status	Presence status
GE1/1/0/1	UP	Active	Normal
GE2/1/0/1	UP	Inactive	Normal

```
[Sysname] display reth interface reth 2
```

Reth2 :

```
Redundancy group : aaa
```

Member	Physical status	Forwarding status	Presence status
GE1/1/0/2	UP	Active	Normal
GE2/1/0/2	UP	Inactive	Normal

显示备份组信息。可以看到备份组中配置为 Primary 的安全引擎处理业务。

```
[Sysname] display failover group group1
```

Stateful failover group information:

ID	Name	Primary	Secondary	Active Status
255	group1	1/4.1	2/4.1	Primary

(2) 冗余组内主备倒换后的显示信息

手工关闭接口 GE1/1/0/2，显示冗余组信息。可以看到优先级低的 Node 2 为主节点。

```
[Sysname] interface gigabitethernet 1/1/0/2
```

```
[Sysname-GigabitEthernet1/1/0/2] shutdown
```

```
[Sysname-GigabitEthernet1/1/0/2] display redundancy group aaa
```

Redundancy group aaa (ID 1):

Node ID	Chassis	Priority	Status	Track weight
Node1	chassis1	100	Secondary	-255
Node2	chassis2	50	Primary	255


```

Preempt delay time remained : 0 min
Preempt delay timer setting : 1 min
Remaining hold-down time : 0 sec
Hold-down timer setting : 300 sec
Manual switchover request : No

```

Member interfaces:

```

Reth1 Reth2

```

Member failover groups:

```

group1

```

Node 1:

Track info:

Track	Status	Reduced weight	Interface
1	Negative	255	GE1/1/0/1
2	Negative	255	GE1/1/0/2(Fault)
5	Positive	255	Blade1/4/0/1

Node 2:

Track info:

Track	Status	Reduced weight	Interface
3	Positive	255	GE2/1/0/1
4	Positive	255	GE2/1/0/2
6	Positive	255	Blade2/4/0/1

显示 Reth 的信息。Reth2 下的接口 GE1/1/0/2 故障 (DOWN)，Reth1 下的接口 GE1/1/0/1 被协议关闭 (DOWN(redundancy down))。GE2/1/0/1 和 GE2/1/0/2 激活。

```

[Sysname-GigabitEthernet1/1/0/2] display reth interface reth 1

```

Reth1 :

Redundancy group : aaa

Member	Physical status	Forwarding status	Presence status
GE1/1/0/1	DOWN(redundancy down)	Inactive	Normal
GE2/1/0/1	UP	Active	Normal

```

[Sysname-GigabitEthernet1/1/0/2] display reth interface reth 2

```

Reth2 :

Redundancy group : aaa

Member	Physical status	Forwarding status	Presence status
GE1/1/0/2	DOWN	Inactive	Normal
GE2/1/0/2	UP	Active	Normal

显示备份组信息。可以看到备份组中配置为 Secondary 的安全引擎处理业务。

```

[Sysname-GigabitEthernet1/1/0/2] display failover group group1

```

Stateful failover group information:

ID	Name	Primary	Secondary	Active Status
255	group1	1/4.1	2/4.1	Secondary

目 录

1 BFD.....	1-1
1.1 BFD简介.....	1-1
1.1.1 BFD会话的建立与拆除	1-1
1.1.2 BFD会话的工作方式和检测模式	1-1
1.1.3 BFD支持的应用	1-2
1.1.4 协议规范	1-2
1.2 配置BFD.....	1-2
1.2.1 echo报文方式配置	1-3
1.2.2 控制报文方式配置	1-3
1.3 开启告警功能.....	1-5
1.4 BFD显示和维护.....	1-5

1 BFD

1.1 BFD简介

BFD (Bidirectional Forwarding Detection, 双向转发检测) 是一个通用的、标准化的、介质无关和协议无关的快速故障检测机制, 用于检测 IP 网络中链路的连通状况, 保证设备之间能够快速检测到通信故障, 以便能够及时采取措施, 保证业务持续运行。

BFD 可以为各种上层协议 (如路由协议) 快速检测两台设备间双向转发路径的故障。上层协议通常采用 Hello 报文机制检测故障, 所需时间为秒级, 而 BFD 可以提供毫秒级检测。

实际应用中, BFD 可以用来进行单跳和多跳检测:

- 单跳检测: 是指对两个直连设备进行 IP 连通性检测, 这里所说的“单跳”是 IP 的一跳。
- 多跳检测: BFD 可以检测两个设备间任意路径的链路情况, 这些路径可能跨越很多跳。

1.1.1 BFD会话的建立与拆除

BFD 本身并没有发现机制, 而是靠被服务的上层协议通知来建立会话, 具体过程如下:

- (1) 上层协议通过自己的 Hello 机制发现邻居并建立连接;
- (2) 上层协议在建立新的邻居关系后, 将邻居的参数及检测参数 (包括目的地址和源地址等) 通告给 BFD;
- (3) BFD 根据收到的参数建立 BFD 会话。

当网络出现故障时:

- (4) BFD 检测到链路故障后, 拆除 BFD 会话, 通知上层协议邻居不可达;
- (5) 上层协议中止邻居关系;
- (6) 如果网络中存在备用路径, 设备将选择备用路径进行通信。

1.1.2 BFD会话的工作方式和检测模式

BFD 会话通过下面两种报文来实现:

- echo 报文: 封装在 UDP 报文中传送, 其 UDP 目的端口号为 3785。
- 控制报文: 封装在 UDP 报文中传送, 对于单跳检测其 UDP 目的端口号为 3784, 对于多跳检测其 UDP 目的端口号为 4784。

1. echo报文方式

本端发送 echo 报文建立 BFD 会话, 对链路进行检测。对端不建立 BFD 会话, 只需把收到的 echo 报文转发回本端。

当 BFD 会话工作于 echo 报文方式时, 仅支持单跳检测, 并且不受检测模式的控制。

2. 控制报文方式

链路两端通过周期性发送控制报文建立 BFD 会话, 对链路进行检测。

BFD 会话建立前有两种模式: 主动模式和被动模式。

- 主动模式：在建立会话前不管是否收到对端发来的 BFD 控制报文，都会主动发送 BFD 控制报文；
 - 被动模式：在建立会话前不会主动发送 BFD 控制报文，直到收到对端发送来的控制报文。
- 通信双方至少要有一方运行在主动模式才能成功建立起 BFD 会话。

BFD 会话建立后有两种模式：异步模式和查询模式。

- 异步模式：设备周期性发送 BFD 控制报文，如果在检测时间内没有收到对端发送的 BFD 控制报文，则认为会话 down。
- 查询模式：设备周期性发送 BFD 控制报文，但是对端（缺省为异步模式）会停止周期性发送 BFD 控制报文。如果通信双方都是查询模式，则双方都停止周期性发送 BFD 控制报文。当需要验证连接性的时候，设备会以协商的周期连续发送几个 P 比特位置 1 的 BFD 控制报文。如果在检测时间内没有收到返回的报文，就认为会话 down；如果收到对方回应的 F 比特位置 1 的报文，就认为连通，停止发送报文，等待下一次触发查询。

另外，也可以链路两端通过发送控制报文建立和保持 BFD 会话，任意一端通过发送 echo 报文检测链路状态。

1.1.3 BFD支持的应用

- 静态路由与 BFD 联动：详细情况请参见“三层技术-IP 路由配置指导”中的“静态路由”。
- IPv6 静态路由与 BFD 联动：详细情况请参见“三层技术-IP 路由配置指导”中的“IPv6 静态路由”。
- RIP 与 BFD 联动：详细情况请参见“三层技术-IP 路由配置指导”中的“RIP”。
- OSPF 与 BFD 联动：详细情况请参见“三层技术-IP 路由配置指导”中的“OSPF”。
- OSPFv3 与 BFD 联动：详细情况请参见“三层技术-IP 路由配置指导”中的“OSPFv3”。
- Track 与 BFD 联动：详细情况请参见“可靠性配置指导”中的“Track”。
- IP 快速重路由：目前支持快速重路由的有 BGP、OSPF、RIP 和静态路由。详细情况请参见“三层技术-IP 路由配置指导”中的“BGP”、“OSPF”、“RIP”和“静态路由”。

1.1.4 协议规范

与 BFD 相关的协议规范有：

- RFC 5880: Bidirectional Forwarding Detection (BFD)
- RFC 5881: Bidirectional Forwarding Detection (BFD) for IPv4 and IPv6 (Single Hop)
- RFC 5882: Generic Application of Bidirectional Forwarding Detection (BFD)
- RFC 5883: Bidirectional Forwarding Detection (BFD) for Multihop Paths
- RFC 5885: Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)

1.2 配置BFD

在配置 BFD 之前，需配置接口的网络层地址，使相邻节点之间网络层可达。

BFD 会话建立后，可以动态协商 BFD 的相关参数（例如最小发送间隔、最小接收间隔、初始模式、报文认证等），两端协议通过发送相应的协商报文后采用新的参数，不影响会话的当前状态。



提示

对于建立在跨板聚合接口上的 BFD 会话，当负责收发 BFD 报文的单板被拔出或异常重启时，备用单板接替收发 BFD 报文的工作需要一定的时间，如果 BFD 会话检测时间较短或者会话数量较多，可能会出现 BFD 会话震荡的情况。

1.2.1 echo报文方式配置

表1-1 echo 报文方式配置

操作	命令	说明
进入系统视图	system-view	-
配置echo报文源IP地址	bfd echo-source-ip <i>ip-address</i>	二者选其一
	bfd echo-source-ipv6 <i>ipv6-address</i>	缺省情况下，未配置echo报文的源IP地址 为了避免对端发送大量的ICMP重定向报文造成网络拥塞，建议不要将echo报文的源IP地址配置为属于该设备任何一个接口所在网段 echo报文源IPv6地址仅支持全球单播地址
进入接口视图	interface <i>interface-type</i> <i>interface-number</i>	-
(可选) 配置接收echo报文的最小时间间隔	bfd min-echo-receive-interval <i>interval</i>	-
(可选) 配置单跳BFD检测时间倍数	bfd detect-multiplier <i>interval</i>	缺省情况下，单跳BFD检测时间倍数为5

1.2.2 控制报文方式配置

1. 单跳检测

表1-2 控制报文方式配置（单跳检测）

操作	命令	说明
进入系统视图	system-view	-
配置BFD会话建立前的运行模式	bfd session init-mode { active passive }	缺省情况下，BFD会话建立前的运行模式为主动模式
进入接口视图	interface <i>interface-type</i> <i>interface-number</i>	-
配置单跳BFD控制报文进行认证的方式	bfd authentication-mode { m-md5 m-sha1 md5 sha1 simple } <i>key-id</i> { cipher <i>cipher-string</i> plain <i>plain-string</i> }	缺省情况下，单跳BFD控制报文不进行认证

操作	命令	说明
配置BFD会话为查询模式	bfd demand enable	缺省情况下，BFD会话为异步模式
使能echo功能	bfd echo [receive send] enable	缺省情况下，echo功能处于关闭状态 本功能在发送控制报文的BFD会话时使用。使能echo功能并且会话up后，设备周期性发送echo报文检测链路连通性，同时降低控制报文的接收速率
配置发送单跳BFD控制报文的最小时间间隔	bfd min-transmit-interval interval	-
配置接收单跳BFD控制报文的最小时间间隔	bfd min-receive-interval interval	-
配置单跳BFD检测时间倍数	bfd detect-multiplier interval	缺省情况下，单跳BFD检测时间倍数为5
创建一个检测本接口状态的BFD会话	bfd detect-interface source-ip ip-address	缺省情况下，不存在检测本接口状态的BFD会话

2. 多跳检测

表1-3 控制报文方式配置（多跳检测）

操作	命令	说明
进入系统视图	system-view	-
配置BFD会话建立前的运行模式	bfd session init-mode { active passive }	缺省情况下，BFD会话建立前的运行模式为主动模式
配置多跳BFD控制报文进行认证的方式	bfd multi-hop authentication-mode { m-md5 m-sha1 md5 sha1 simple } key-id { cipher cipher-string plain plain-string }	缺省情况下，多跳BFD控制报文不进行认证
配置多跳BFD控制报文的端口号	bfd multi-hop destination-port port-number	缺省情况下，多跳BFD控制报文的端口号为4784
配置多跳BFD检测时间倍数	bfd multi-hop detect-multiplier value	缺省情况下，多跳BFD检测时间倍数为5
配置接收多跳BFD控制报文的最小时间间隔	bfd multi-hop min-receive-interval interval	-
配置发送多跳BFD控制报文的最小时间间隔	bfd multi-hop min-transmit-interval interval	-

3. 配置BFD模板

对于未指定出接口的会话，无法通过会话出接口配置 BFD 会话参数。使用 BFD 全局多跳可以配置，但是缺乏灵活性。通过 BFD 模板可以对参数进行灵活配置，LSP 以及 PW 的 BFD 检测关联到 BFD 模板即可指定会话参数。

表1-4 配置 BFD 模板

操作	命令	说明
进入系统视图	system-view	-
创建BFD模板，并进入BFD模板视图	bfd template <i>template-name</i>	缺省情况下，不存在BFD模板
配置BFD控制报文进行认证的方式	bfd authentication-mode { m-md5 m-sha1 md5 sha1 simple } <i>key-id</i> { cipher <i>cipher-string</i> plain <i>plain-string</i> }	缺省情况下，BFD控制报文不进行认证
配置BFD检测时间倍数	bfd detect-multiplier <i>value</i>	缺省情况下，BFD检测时间倍数为5
配置接收BFD控制报文的 minimum 时间间隔	bfd min-receive-interval <i>interval</i>	-
配置发送BFD控制报文的 minimum 时间间隔	bfd min-transmit-interval <i>interval</i>	-

1.3 开启告警功能

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

表1-5 开启告警功能

操作	命令	说明
进入系统视图	system-view	-
开启BFD的告警功能	snmp-agent trap enable bfd	缺省情况下，BFD的告警功能处于开启状态

1.4 BFD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 BFD 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 BFD 会话的统计信息。

表1-6 BFD 显示和维护

操作	命令
显示BFD会话信息	display bfd session [discriminator <i>value</i> verbose]
清除BFD会话统计信息	reset bfd session statistics

目 录

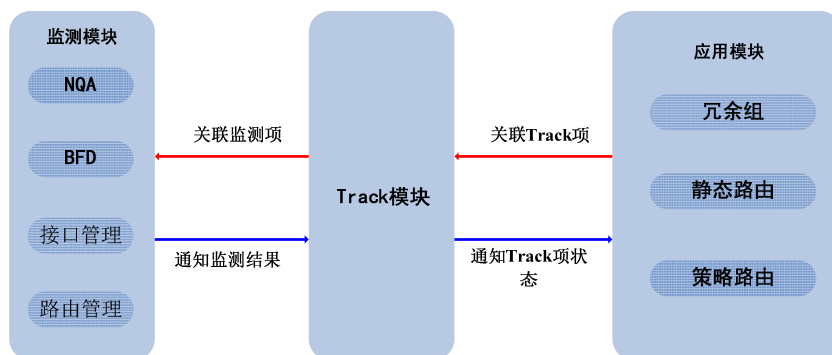
1 Track	1-1
1.1 Track简介	1-1
1.1.1 联动功能介绍.....	1-1
1.1.2 联动功能工作原理	1-1
1.1.3 联动功能应用举例	1-2
1.2 Track配置任务简介	1-2
1.3 配置Track与监测模块联动	1-3
1.3.1 配置Track与NQA联动	1-3
1.3.2 配置Track与BFD联动	1-3
1.3.3 配置Track与接口管理联动	1-4
1.3.4 配置Track与路由管理联动	1-5
1.4 配置Track与应用模块联动	1-5
1.4.1 配置Track与静态路由联动	1-5
1.4.2 配置Track与策略路由联动	1-6
1.4.3 配置Track与冗余组联动	1-8
1.5 Track显示和维护	1-8
1.6 Track典型配置举例	1-9
1.6.1 静态路由、Track与NQA联动配置举例	1-9
1.6.2 静态路由、Track与BFD联动配置举例.....	1-14

1 Track

1.1 Track简介

1.1.1 联动功能介绍

图1-1 联动功能实现示意图



Track的用途是实现联动功能。如 [图 1-1](#) 所示，联动功能通过在监测模块、Track模块和应用模块之间建立关联，实现这些模块之间的联合动作。联动功能利用监测模块对链路状态、网络性能等进行监测，并通过Track模块将监测结果及时通知给应用模块，以便应用模块进行相应的处理。例如，在静态路由、Track和NQA之间建立联动，利用NQA监测静态路由的下一跳地址是否可达。NQA监测到下一跳不可达时，通过Track通知静态路由模块该监测结果，以便静态路由模块将该条路由置为无效，确保报文不再通过该静态路由转发。

如果应用模块直接与监测模块关联，由于不同监测模块通知给应用模块的监测结果形式各不相同，应用模块需要分别处理不同形式的监测结果。联动功能在应用模块和监测模块之间增加了 Track 模块，通过 Track 模块屏蔽不同监测模块的差异，将监测结果以统一的形式通知给应用模块，从而简化应用模块的处理。

1.1.2 联动功能工作原理

联动功能的工作原理分为两部分：

- Track 模块与监测模块联动
- Track 模块与应用模块联动

1. Track模块与监测模块联动

Track 模块和监测模块之间通过 Track 项建立关联。监测模块负责对接口状态、链路状态等进行监测，并将监测结果通知给 Track 模块；Track 模块根据监测结果改变 Track 项的状态。

- 如果监测结果为监测对象工作正常（如接口处于 up 状态、网络可达），则对应 Track 项的状态为 Positive。
- 如果监测结果为监测对象出现异常（如接口处于 down 状态、网络不可达），则对应 Track 项的状态为 Negative。

- 如果监测结果无效（如 NQA 作为监测模块时，与 Track 项关联的 NQA 测试组不存在），则对应 Track 项的状态为 NotReady。

目前，可以与 Track 模块实现联动功能的监测模块包括：

- NQA（Network Quality Analyzer，网络质量分析）
- BFD（Bidirectional Forwarding Detection，双向转发检测）
- 接口管理
- 路由管理

2. Track模块与应用模块联动

Track 模块和应用模块之间通过 Track 项建立关联。Track 项的状态改变后，通知应用模块；应用模块根据 Track 项的状态，及时进行相应的处理，从而避免通信的中断或服务质量的降低。

目前，可以与 Track 模块实现联动功能的应用模块包括：

- 静态路由
- 策略路由
- 冗余组

在某些情况下，Track 项状态发生变化后，如果立即通知应用模块，则可能会由于路由无法及时恢复等原因，导致通信中断。

1.1.3 联动功能应用举例

下面以 NQA、Track 和静态路由联动为例，说明联动功能的工作原理。

用户在设备上配置了一条静态路由，下一跳地址为 192.168.0.88。如果 192.168.0.88 可达，则报文可以通过该静态路由转发，该静态路由有效；如果 192.168.0.88 不可达，则通过该静态路由转发报文会导致报文转发失败，此时，需要将该静态路由置为无效。通过在 NQA、Track 模块和静态路由之间建立联动，可以实现实时监测下一跳的可达性，以便及时判断静态路由是否有效。

在此例中联动功能的配置方法及其工作原理为：

- (1) 创建 NQA 测试组，通过 NQA 测试组监测目的地址 192.168.0.88 是否可达。
- (2) 创建和 NQA 测试组关联的 Track 项。192.168.0.88 可达时，NQA 会将监测结果通知给 Track 模块，Track 模块将该 Track 项的状态变为 Positive；192.168.0.88 不可达时，NQA 将监测结果通知给 Track 模块，Track 模块将该 Track 项的状态变为 Negative。
- (3) 配置这条静态路由和 Track 项关联。如果 Track 模块通知静态路由 Track 项的状态为 Positive，则静态路由模块将这条路由置为有效；如果 Track 模块通知静态路由 Track 项的状态为 Negative，则静态路由模块将这条路由置为无效。

1.2 Track配置任务简介

为了实现联动功能，需要在 Track 与监测模块、Track 与应用模块之间分别建立联动关系。

表1-1 Track 配置任务简介

配置任务		说明	详细配置
配置Track与监测模块联动	配置Track与NQA联动	四者必选其一	1.3.1
	配置Track与BFD联动		1.3.2
	配置Track与接口管理联动		1.3.3
	配置Track与路由管理联动		1.3.4
配置Track与应用模块联动	配置Track与静态路由联动	三者必选其一	1.4.1
	配置Track与策略路由联动		1.4.2
	配置Track与冗余组联动		1.4.3

1.3 配置Track与监测模块联动

1.3.1 配置Track与NQA联动

NQA 测试组周期性地探测某个目的地址是否可达、是否可以与某个目的服务器建立 TCP 连接等。如果在 Track 项和 NQA 测试组之间建立了关联，则当连续探测失败的次数达到指定的阈值时，NQA 将通知 Track 模块监测对象出现异常，Track 模块将与 NQA 测试组关联的 Track 项的状态置为 Negative；否则，NQA 通知 Track 模块监测对象正常工作，Track 模块将 Track 项的状态置为 Positive。NQA 的详细介绍，请参见“网络管理和监控配置指导”中的“NQA”。

表1-2 配置 Track 与 NQA 联动

操作	命令	说明
进入系统视图	system-view	-
创建与NQA测试组中指定联动项关联的Track项，并指定Track项状态变化时通知应用模块的延迟时间	track track-entry-number nqa entry admin-name operation-tag reaction item-number [delay { negative negative-time positive positive-time } *]	缺省情况下，不存在 Track项 配置Track项时，引用的NQA测试组或联动项可以不存在，此时该 Track项的状态为 NotReady

1.3.2 配置Track与BFD联动

如果在 Track 项和 BFD 会话之间建立了关联，则当 BFD 判断出对端不可达时，BFD 会通知 Track 模块将与 BFD 会话关联的 Track 项的状态置为 Negative；否则，通知 Track 模块将 Track 项的状态置为 Positive。

BFD 会话支持两种工作方式：Echo 报文方式和控制报文方式。Track 项只能与 Echo 报文方式的 BFD 会话建立关联，不能与控制报文方式的 BFD 会话建立联动。BFD 的详细介绍，请参见“可靠性配置指导”中的“BFD”。

1. 配置准备

配置 Track 与 BFD 联动前，需要配置 BFD echo 报文的源地址，配置方法请参见“可靠性配置指导”中的“BFD”。

2. 配置步骤

表1-3 配置 Track 与 BFD 联动

操作	命令	说明
进入系统视图	system-view	-
创建和BFD会话关联的Track项，并指定Track项状态变化时通知应用模块的延迟时间	track track-entry-number bfd echo interface interface-type interface-number remote ip remote-ip-address local ip local-ip-address [delay { negative negative-time positive positive-time } *]	缺省情况下，不存在Track项

1.3.3 配置Track与接口管理联动

接口管理用来监视接口的物理状态、链路状态和网络层协议状态。如果在 Track 项和接口之间建立了关联，则当接口的物理状态、链路状态或网络层协议状态为 up 时，接口管理通知 Track 模块将与接口关联的 Track 项的状态置为 Positive；接口的物理状态、链路状态或网络层协议状态为 down 时，接口管理通知 Track 模块将 Track 项的状态为 Negative。

表1-4 配置 Track 与接口管理联动

操作	命令	说明
进入系统视图	system-view	-
创建和接口管理关联的Track项，监视接口的链路状态，并指定Track项状态变化时通知应用模块的延迟时间	track track-entry-number interface interface-type interface-number [delay { negative negative-time positive positive-time } *]	三者选其一 缺省情况下，不存在Track项
创建和接口管理关联的Track项，监视接口的物理状态，并指定Track项状态变化时通知应用模块的延迟时间	track track-entry-number interface interface-type interface-number physical [delay { negative negative-time positive positive-time } *]	
创建和接口管理关联的Track项，监视接口的网络层协议状态，并指定Track项状态变化时通知应用模块的延迟时间	track track-entry-number interface interface-type interface-number protocol { ipv4 ipv6 } [delay { negative negative-time positive positive-time } *]	

1.3.4 配置Track与路由管理联动

如果在 Track 项和路由管理之间建立了关联，当对应的路由条目在路由表中存在时，路由管理通知 Track 模块将与之关联的 Track 项状态设置为 Positive；当对应的路由条目在路由表中被删除时，路由管理将通知 Track 模块将与之关联的 Track 项状态设置为 Negative。

表1-5 配置 Track 与路由管理联动

操作	命令	说明
进入系统视图	system-view	-
创建与路由管理关联的Track项，并指定Track项状态变化时通知应用模块的延迟时间	track track-entry-number ip route [vpn-instance vpn-instance-name] ip-address { mask-length mask } reachability [delay { negative negative-time positive positive-time } *]	缺省情况下，不存在Track项

1.4 配置Track与应用模块联动



说明

用户配置 Track 和应用模块联动时，需保证联动的 Track 项已被创建，否则应用模块可能会获取到错误的 Track 项状态信息。

1.4.1 配置Track与静态路由联动

静态路由是一种特殊的路由，由管理员手工配置。配置静态路由后，去往指定目的地的报文将按照管理员指定的路径进行转发。静态路由由配置的详细介绍，请参见“三层技术-IP 路由配置指导”中的“静态路由”。

静态路由的缺点在于：不能自动适应网络拓扑结构的变化，当网络发生故障或者拓扑发生变化时，可能会导致静态路由不可达，网络通信中断。

为了防止这种情况发生，可以配置其它路由和静态路由形成备份关系。静态路由可达时，根据静态路由转发报文，其它路由处于备份状态；静态路由不可达时，根据备份路由转发报文，从而避免通信中断，提高了网络可靠性。

通过在 Track 模块和静态路由之间建立联动，可以实现静态路由可达性的实时判断。

如果在配置静态路由时只指定了下一跳而未指定出接口，可以通过联动功能，利用监测模块监视静态路由下一跳的可达性，并根据 Track 项的状态来判断静态路由的可达性：

- 当 Track 项状态为 Positive 时，静态路由的下一跳可达，配置的静态路由将生效；
- 当 Track 项状态为 Negative 时，静态路由的下一跳不可达，配置的静态路由无效；
- 当 Track 项状态为 NotReady 时，无法判断静态路由的下一跳是否可达，此时配置的静态路由生效。

配置 Track 与静态路由联动时，需要注意的是：

- 如果 Track 模块通过 NQA 探测私网静态路由中下一跳的可达性，静态路由下一跳的 VPN 实例名与 NQA 测试组配置的实例名必须相同，才能进行正常的探测。
- 在静态路由进行迭代时，Track 项监测的应该是静态路由迭代后最终的下一跳地址，而不是配置中指定的下一跳地址。否则，可能导致错误地将有效路由判断为无效路由。

表1-6 配置 Track 与静态路由联动

操作	命令	说明
进入系统视图	system-view	-
配置通过Track与静态路由联动，检测静态路由下一跳是否可达	<pre> ip route-static { <i>dest-address</i> { <i>mask-length</i> <i>mask</i> } group <i>group-name</i> } { <i>next-hop-address</i> track <i>track-entry-number</i> vpn-instance <i>d-vpn-instance-name</i> <i>next-hop-address</i> track <i>track-entry-number</i> } [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] ip route-static vpn-instance <i>s-vpn-instance-name</i> { <i>dest-address</i> { <i>mask-length</i> <i>mask</i> } group <i>group-name</i> } { <i>next-hop-address</i> [public] track <i>track-entry-number</i> vpn-instance <i>d-vpn-instance-name</i> <i>next-hop-address</i> track <i>track-entry-number</i> } [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] </pre>	缺省情况下，未配置Track与静态路由联动

1.4.2 配置Track与策略路由联动

策略路由是一种依据用户制定的策略进行路由选择的机制。与单纯依照 IP 报文的目的地地址查找路由表进行转发不同，策略路由基于到达报文的源地址、长度等信息灵活地进行路由选择。对于满足一定条件（报文长度或 ACL 规则）的报文，将执行一定的操作（设置转发报文的 VPN 实例、设置报文的优先级、设置报文的出接口和下一跳、设置报文的缺省出接口和下一跳等），以指导报文的转发。策略路由配置的详细介绍，请参见“三层技术-IP 路由配置指导”中的“IP 策略路由”。

策略路由无法判断对报文执行的操作的可用性。当执行的操作不可用时，策略路由仍然对报文执行该操作，可能会导致报文转发失败。例如，策略路由中配置满足一定条件的报文，需要通过指定的出接口转发。当该出接口所在的链路出现故障时，策略路由无法感知链路故障，仍然通过该接口转发报文，导致报文转发失败。

通过联动功能，可以解决上述问题，增强了策略路由应用的灵活性，以及策略路由对网络环境的动态感知能力。配置策略路由执行的操作与 Track 项关联，利用监测模块监视链路的状态，通过 Track 项的状态来动态地决定策略路由操作的可用性：

- Track 项状态为 **Positive** 时，表示链路正常工作，与该 Track 项关联的策略路由操作生效，可以指导转发；
- Track 项状态为 **Negative** 时，表示链路出现故障，与该 Track 项关联的策略路由操作无效，转发时忽略该配置项；
- Track 项状态为 **NotReady** 时，与该 Track 项关联的策略路由操作生效，可以指导转发。

目前，支持与 Track 项关联的策略路由操作包括：

- 设置报文的出接口
- 设置报文的下一跳

- 设置报文的缺省出接口
- 设置报文的缺省下一跳

1. 配置准备

配置 Track 与策略路由联动前，需要先创建策略或一个策略节点，并配置匹配规则。

2. 配置步骤

表1-7 配置 Track 与 IPv4 策略路由联动

操作	命令	说明
进入系统视图	system-view	-
创建策略或一个策略节点，并进入该策略视图	policy-based-route <i>policy-name</i> [deny permit] node <i>node-number</i>	-
设置IP报文长度匹配条件	if-match packet-length <i>min-len max-len</i>	缺省情况下，所有报文都会通过该节点的过滤
设置ACL匹配条件	if-match acl { <i>acl-number</i> name <i>acl-name</i> }	
设置报文的出接口，并与Track项关联	apply output-interface { <i>interface-type interface-number</i> [track track-entry-number] }&<1-n>	四者至少选其一
设置报文的下一跳，并与Track项关联	apply next-hop [vpn-instance <i>vpn-instance-name</i> inbound-vpn] { <i>ip-address</i> [direct] [track track-entry-number] }&<1-n>	
设置报文缺省出接口，并与Track项关联	apply default-output-interface { <i>interface-type interface-number</i> [track track-entry-number] }&<1-n>	
设置报文缺省下一跳，并与Track项关联	apply default-next-hop [vpn-instance <i>vpn-instance-name</i> inbound-vpn] { <i>ip-address</i> [direct] [track track-entry-number] }&<1-n>	

表1-8 配置 Track 与 IPv6 策略路由联动

操作	命令	说明
进入系统视图	system-view	-
创建策略或一个策略节点，并进入该策略视图	ipv6 policy-based-route <i>policy-name</i> [deny permit] node <i>node-number</i>	-
设置IPv6报文长度匹配条件	if-match packet-length <i>min-len max-len</i>	缺省情况下，所有报文都会通过该节点的过滤
设置ACL匹配条件	if-match acl { <i>ipv6-acl-number</i> name <i>ipv6-acl-name</i> }	
设置报文的出接口，并与Track项关联	apply output-interface { <i>interface-type interface-number</i> [track track-entry-number] }&<1-n>	四者至少选其一
设置报文的下一跳，并与Track项关联	apply next-hop [vpn-instance <i>vpn-instance-name</i> inbound-vpn] { <i>ipv6-address</i> [direct] [track track-entry-number] }&<1-n>	

操作	命令	说明
设置报文缺省出接口，并与Track项关联	apply default-output-interface { <i>interface-type</i> <i>interface-number</i> [track <i>track-entry-number</i>] }<1-n>	
设置报文缺省下一跳，并与Track项关联	apply default-next-hop [vpn-instance <i>vpn-instance-name</i> / inbound-vpn] { <i>ipv6-address</i> [direct] [track <i>track-entry-number</i>] }<1-n>	

1.4.3 配置Track与冗余组联动

冗余组本身无法感知链路故障，所以需要与 Track 联动，来达到监控链路状态的目的。联动关系建立后，冗余组能够快速检测上、下行链路是否故障，如果故障，则通知组内所有成员（包括以太网冗余接口）整体进行倒换。

通过配置 Track 与冗余组联动，可以实时监控任意类型的故障，以触发冗余组进行整体倒换。

Track 模块根据监测模块的监测结果改变 Track 项的状态，并将 Track 项状态通知给冗余组模块；冗余组模块根据 Track 项状态进行相应处理：

- 如果 Track 项状态变为 **Positive**，则增加冗余组节点权重值，权重值为正值，则认为节点能够正常工作；
- 如果 Track 项状态变为 **Negative**，则减少冗余组节点权重值，如果权重值变为负值，则冗余组进行倒转；
- 如果 Track 项状态变为 **Notready**，处理方式和 **Negative** 一致，减少冗余组节点权重值，如果权重值变为负值，则冗余组进行倒转。

在配置 Track 与冗余组联动时，需要配置对应的关联接口，当出现故障时，冗余组记录该接口为故障接口。如果冗余组功能支持倒回，需要配置该参数；否则可以不配置该参数。

关于“冗余组”的详细介绍，请参见“可靠性配置指导”中的“冗余备份”。

表1-9 配置 Track 与冗余组联动

操作	命令	说明
进入系统视图	system-view	-
创建冗余组并进入该冗余组视图	redundancy group <i>group-name</i>	缺省情况下，设备上不存在冗余组
创建冗余组节点，并进入冗余组节点视图	node <i>node-id</i>	缺省情况下，不存在冗余组节点
配置Track与冗余组关联	track <i>track-entry-number</i> [reduced <i>weight-reduced</i>] [interface <i>interface-type</i> <i>interface-number</i>]	缺省情况下，冗余组节点下未关联 Track项

1.5 Track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 Track 的运行情况，通过查看显示信息验证配置的效果。

表1-10 Track 显示和维护

操作	命令
显示Track项的信息	<code>display track { track-entry-number all [negative positive] } [brief]</code>

1.6 Track典型配置举例

1.6.1 静态路由、Track与NQA联动配置举例

1. 组网需求

Device A、Device B、Device C 和 Device D 连接了 20.1.1.0/24 和 30.1.1.0/24 两个网段，在路由器上配置静态路由以实现两个网段的互通，并配置路由备份以提高网络的可靠性。

Device A 作为 20.1.1.0/24 网段内主机的缺省网关，在 Device A 上存在两条到达 30.1.1.0/24 网段的静态路由，下一跳分别为 Device B 和 Device C。这两条静态路由形成备份，其中：

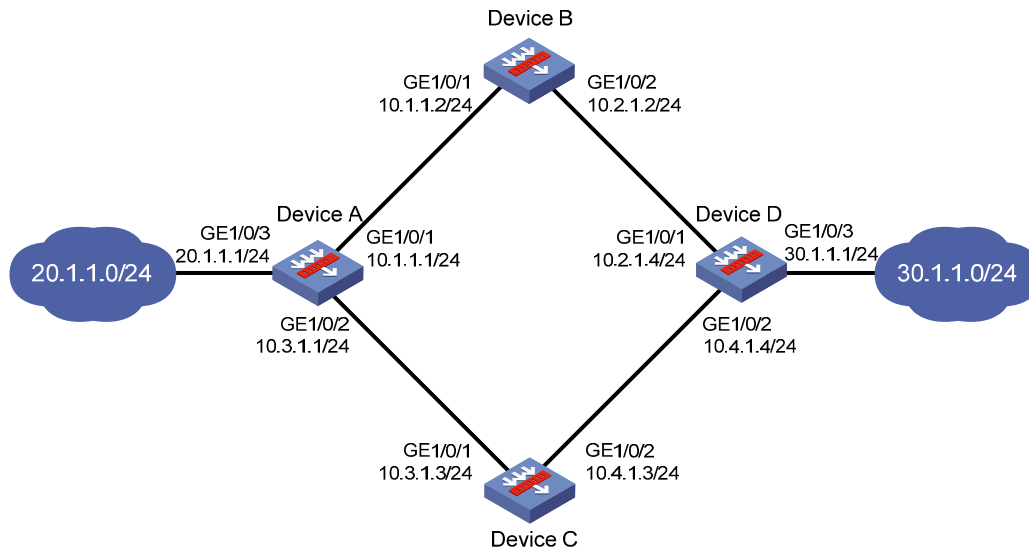
- 下一跳为 Device B 的静态路由优先级高，作为主路由。该路由可达时，Device A 通过 Device B 将报文转发到 30.1.1.0/24 网段。
- 下一跳为 Device C 的静态路由作为备份路由。
- 在 Device A 上通过静态路由、Track 与 NQA 联动，实时判断主路由是否可达。当主路由不可达时，备份路由生效，Device A 通过 Device C 将报文转发到 30.1.1.0/24 网段。

同样地，Device D 作为 30.1.1.0/24 网段内主机的缺省网关，在 Device D 上存在两条到达 20.1.1.0/24 网段的静态路由，下一跳分别为 Device B 和 Device C。这两条静态路由形成备份，其中：

- 下一跳为 Device B 的静态路由优先级高，作为主路由。该路由可达时，Device D 通过 Device B 将报文转发到 20.1.1.0/24 网段。
- 下一跳为 Device C 的静态路由作为备份路由。
- 在 Device D 上通过静态路由、Track 与 NQA 联动，实时判断主路由是否可达。当主路由不可达时，备份路由生效，Device D 通过 Device C 将报文转发到 20.1.1.0/24 网段。

2. 组网图

图1-2 静态路由、Track 与 NQA 联动配置组网图



3. 配置步骤

(1) 按照 [图 1-2](#) 配置各接口的 IP 地址，具体配置过程略。

(2) 配置 Device A

配置到达 30.1.1.0/24 网段的静态路由：下一跳地址为 10.1.1.2，优先级为缺省值 60，该路由与 Track 项 1 关联。

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

```
[DeviceA] ip route-static 30.1.1.0 24 10.1.1.2 track 1
```

配置到达 30.1.1.0/24 网段的静态路由：下一跳地址为 10.3.1.3，优先级为 80。

```
[DeviceA] ip route-static 30.1.1.0 24 10.3.1.3 preference 80
```

配置到达 10.2.1.4 的静态路由：下一跳地址为 10.1.1.2。

```
[DeviceA] ip route-static 10.2.1.4 24 10.1.1.2
```

创建管理员名为 admin、操作标签为 test 的 NQA 测试组。

```
[DeviceA] nqa entry admin test
```

配置测试类型为 ICMP-echo。

```
[DeviceA-nqa-admin-test] type icmp-echo
```

配置测试的目的地址为 10.2.1.4，下一跳地址为 10.1.1.2，以便通过 NQA 检测 Device A—Device B—Device D 这条路径的连通性。

```
[DeviceA-nqa-admin-test-icmp-echo] destination ip 10.2.1.4
```

```
[DeviceA-nqa-admin-test-icmp-echo] next-hop 10.1.1.2
```

配置测试频率为 100ms。

```
[DeviceA-nqa-admin-test-icmp-echo] frequency 100
```

配置联动项 1（连续失败 5 次触发联动）。

```
[DeviceA-nqa-admin-test-icmp-echo] reaction 1 checked-element probe-fail threshold-type consecutive 5 action-type trigger-only
```

```
[DeviceA-nqa-admin-test-icmp-echo] quit
```

启动探测。

```
[DeviceA] nqa schedule admin test start-time now lifetime forever
```

配置 Track 项 1，关联 NQA 测试组（管理员为 admin，操作标签为 test）的联动项 1。

```
[DeviceA] track 1 nqa entry admin test reaction 1
```

(3) 配置 Device B

配置到达 30.1.1.0/24 网段的静态路由：下一跳地址为 10.2.1.4。

```
<DeviceB> system-view
```

```
[DeviceB] ip route-static 30.1.1.0 24 10.2.1.4
```

配置到达 20.1.1.0/24 网段的静态路由：下一跳地址为 10.1.1.1。

```
[DeviceB] ip route-static 20.1.1.0 24 10.1.1.1
```

(4) 配置 Device C

配置到达 30.1.1.0/24 网段的静态路由：下一跳地址为 10.4.1.4。

```
<DeviceC> system-view
```

```
[DeviceC] ip route-static 30.1.1.0 24 10.4.1.4
```

配置到达 20.1.1.0/24 网段的静态路由：下一跳地址为 10.3.1.1。

```
[DeviceC] ip route-static 20.1.1.0 24 10.3.1.1
```

(5) 配置 Device D

配置到达 20.1.1.0/24 网段的静态路由：下一跳地址为 10.2.1.2，优先级为缺省值 60，该路由与 Track 项 1 关联。

```
<DeviceD> system-view
```

```
[DeviceD] ip route-static 20.1.1.0 24 10.2.1.2 track 1
```

配置到达 20.1.1.0/24 网段的静态路由：下一跳地址为 10.4.1.3，优先级为 80。

```
[DeviceD] ip route-static 20.1.1.0 24 10.4.1.3 preference 80
```

配置到达 10.1.1.1 的静态路由：下一跳地址为 10.2.1.2。

```
[DeviceD] ip route-static 10.1.1.1 24 10.2.1.2
```

创建管理员名为 admin、操作标签为 test 的 NQA 测试组。

```
[DeviceD] nqa entry admin test
```

配置测试类型为 ICMP-echo。

```
[DeviceD-nqa-admin-test] type icmp-echo
```

配置测试的目的地址为 10.1.1.1，下一跳地址为 10.2.1.2，以便通过 NQA 检测 Device D—Device B—Device A 这条路径的连通性。

```
[DeviceD-nqa-admin-test-icmp-echo] destination ip 10.1.1.1
```

```
[DeviceD-nqa-admin-test-icmp-echo] next-hop 10.2.1.2
```

配置测试频率为 100ms。

```
[DeviceD-nqa-admin-test-icmp-echo] frequency 100
```

配置联动项 1（连续失败 5 次触发联动）。

```
[DeviceD-nqa-admin-test-icmp-echo] reaction 1 checked-element probe-fail threshold-type  
consecutive 5 action-type trigger-only
```

```
[DeviceD-nqa-admin-test-icmp-echo] quit
```

启动探测。

```
[DeviceD] nqa schedule admin test start-time now lifetime forever
```

配置 Track 项 1，关联 NQA 测试组（管理员为 admin，操作标签为 test）的联动项 1。

```
[DeviceD] track 1 nqa entry admin test reaction 1
```

4. 验证配置

显示 Device A 上 Track 项的信息。

```
[DeviceA] display track all
Track ID: 1
  State: Positive
  Duration: 0 days 0 hours 0 minutes 32 seconds
  Notification delay: Positive 0, Negative 0 (in seconds)
  Tracked object:
    NQA entry: admin test
    Reaction: 1
    Remote IP/URL:--
    Local IP:--
    Interface:--
```

显示 Device A 的路由表。

```
[DeviceA] display ip routing-table

Destinations : 10          Routes : 10

Destination/Mask    Proto  Pre  Cost           NextHop          Interface
10.1.1.0/24         Direct  0    0             10.1.1.1          GE1/0/1
10.1.1.1/32         Direct  0    0             127.0.0.1         InLoop0
10.2.1.0/24         Static  60   0             10.1.1.2          GE1/0/1
10.3.1.0/24         Direct  0    0             10.3.1.1          GE1/0/2
10.3.1.1/32         Direct  0    0             127.0.0.1         InLoop0
20.1.1.0/24         Direct  0    0             20.1.1.1          GE1/0/3
20.1.1.1/32         Direct  0    0             127.0.0.1         InLoop0
30.1.1.0/24         Static  60   0             10.1.1.2          GE1/0/1
127.0.0.0/8         Direct  0    0             127.0.0.1         InLoop0
127.0.0.1/32        Direct  0    0             127.0.0.1         InLoop0
```

以上显示信息表示, NQA 测试的结果为主路由可达(Track 项状态为 Positive), Device A 通过 Device B 将报文转发到 30.1.1.0/24 网段。

在 Device B 上删除接口 GigabitEthernet1/0/1 的 IP 地址。

```
<DeviceB> system-view
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] undo ip address
```

显示 Device A 上 Track 项的信息。

```
[DeviceA] display track all
Track ID: 1
  State: Negative
  Duration: 0 days 0 hours 0 minutes 32 seconds
  Notification delay: Positive 0, Negative 0 (in seconds)
  Tracked object:
    NQA entry: admin test
    Reaction: 1
    Remote IP/URL:--
    Local IP:--
```

Interface:--

显示 Device A 的路由表。

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

```
Destinations : 10          Routes : 10
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
10.1.1.0/24	Direct	0	0	10.1.1.1	GE1/0/1
10.1.1.1/32	Direct	0	0	127.0.0.1	InLoop0
10.2.1.0/24	Static	60	0	10.1.1.2	GE1/0/1
10.3.1.0/24	Direct	0	0	10.3.1.1	GE1/0/2
10.3.1.1/32	Direct	0	0	127.0.0.1	InLoop0
20.1.1.0/24	Direct	0	0	20.1.1.1	GE1/0/3
20.1.1.1/32	Direct	0	0	127.0.0.1	InLoop0
30.1.1.0/24	Static	80	0	10.3.1.3	GE1/0/2
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

以上显示信息表示，NQA 测试的结果为主路由不可达（Track 项状态为 **Negative**），则备份路由生效，Device A 通过 Device C 将报文转发到 30.1.1.0/24 网段。

主路由出现故障后，20.1.1.0/24 网段内的主机仍然可以与 30.1.1.0/24 网段内的主机通信。

```
[DeviceA] ping -a 20.1.1.1 30.1.1.1
```

```
Ping 30.1.1.1: 56 data bytes, press CTRL_C to break
```

```
Reply from 30.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
```

```
Reply from 30.1.1.1: bytes=56 Sequence=2 ttl=254 time=1 ms
```

```
Reply from 30.1.1.1: bytes=56 Sequence=3 ttl=254 time=1 ms
```

```
Reply from 30.1.1.1: bytes=56 Sequence=4 ttl=254 time=2 ms
```

```
Reply from 30.1.1.1: bytes=56 Sequence=5 ttl=254 time=1 ms
```

```
--- Ping statistics for 30.1.1.1 ---
```

```
5 packet(s) transmitted, 5 packet(s) received, 0.00% packet loss
```

```
round-trip min/avg/max/std-dev = 1/1/2/1 ms
```

Device D 上的显示信息与 Device A 类似。主路由出现故障后，30.1.1.0/24 网段内的主机仍然可以与 20.1.1.0/24 网段内的主机通信。

```
[DeviceD] ping -a 30.1.1.1 20.1.1.1
```

```
Ping 20.1.1.1: 56 data bytes, press CTRL_C to break
```

```
Reply from 20.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
```

```
Reply from 20.1.1.1: bytes=56 Sequence=2 ttl=254 time=1 ms
```

```
Reply from 20.1.1.1: bytes=56 Sequence=3 ttl=254 time=1 ms
```

```
Reply from 20.1.1.1: bytes=56 Sequence=4 ttl=254 time=1 ms
```

```
Reply from 20.1.1.1: bytes=56 Sequence=5 ttl=254 time=1 ms
```

```
--- Ping statistics for 20.1.1.1 ---
```

```
5 packet(s) transmitted, 5 packet(s) received, 0.00% packet loss
```

```
round-trip min/avg/max/std-dev = 1/1/2/1 ms
```

1.6.2 静态路由、Track与BFD联动配置举例

1. 组网需求

Device A、Device B 和 Device C 连接了 20.1.1.0/24 和 30.1.1.0/24 两个网段，在路由器上配置静态路由以实现两个网段的互通，并配置路由备份以提高网络的可靠性。

Device A 作为 20.1.1.0/24 网段内主机的缺省网关，在 Route A 上存在两条到达 30.1.1.0/24 网段的静态路由，下一跳分别为 Device B 和 Device C。这两条静态路由形成备份，其中：

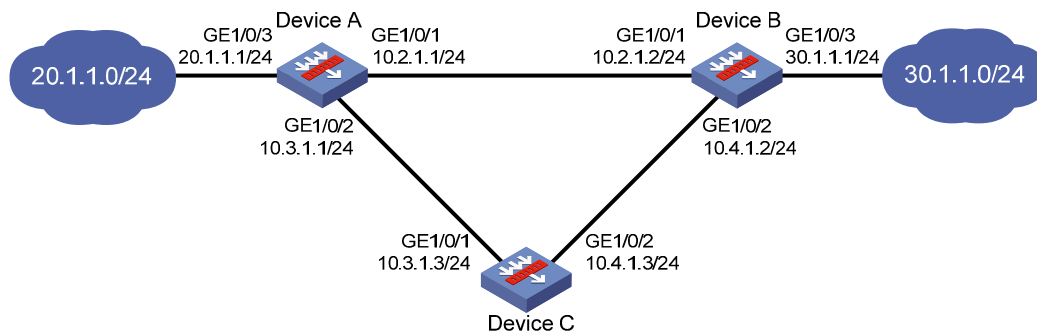
- 下一跳为 Device B 的静态路由由优先级高，作为主路由。该路由可达时，Device A 通过 Device B 将报文转发到 30.1.1.0/24 网段。
- 下一跳为 Device C 的静态路由作为备份路由。
- 在 Device A 上通过静态路由、Track 与 BFD 联动，实时判断主路由是否可达。当主路由不可达时，BFD 能够快速检测到路由故障，使得备份路由生效，Device A 通过 Device C 和 Device B 将报文转发到 30.1.1.0/24 网段。

同样地，Device B 作为 30.1.1.0/24 网段内主机的缺省网关，在 Route B 上存在两条到达 20.1.1.0/24 网段的静态路由，下一跳分别为 Device A 和 Device C。这两条静态路由形成备份，其中：

- 下一跳为 Device A 的静态路由由优先级高，作为主路由。该路由可达时，Device B 通过 Device A 将报文转发到 20.1.1.0/24 网段。
- 下一跳为 Device C 的静态路由作为备份路由。
- 在 Device B 上通过静态路由、Track 与 BFD 联动，实时判断主路由是否可达。当主路由不可达时，BFD 能够快速检测到路由故障，使得备份路由生效，Device B 通过 Device C 和 Device A 将报文转发到 20.1.1.0/24 网段。

2. 组网图

图1-3 静态路由、Track 与 BFD 联动配置组网图



3. 配置步骤

(1) 按照 图 1-3 配置各接口的IP地址，具体配置过程略。

(2) 配置 Device A

配置到达 30.1.1.0/24 网段的静态路由：下一跳地址为 10.2.1.2，优先级为缺省值 60，该路由与 Track 项 1 关联。

```
<DeviceA> system-view
[DeviceA] ip route-static 30.1.1.0 24 10.2.1.2 track 1
```

配置到达 30.1.1.0/24 网段的静态路由：下一跳地址为 10.3.1.3，优先级为 80。

```
[DeviceA] ip route-static 30.1.1.0 24 10.3.1.3 preference 80
```

配置 BFD echo 报文的源地址为 10.10.10.10。

```
[DeviceA] bfd echo-source-ip 10.10.10.10
```

创建和 BFD 会话关联的 Track 项 1，检测 Device A 是否可以与静态路由的下一跳 Device B 互通。

```
[DeviceA] track 1 bfd echo interface gigabitethernet 1/0/1 remote ip 10.2.1.2 local ip 10.2.1.1
```

(3) 配置 Device B

配置到达 20.1.1.0/24 网段的静态路由：下一跳地址为 10.2.1.1，优先级为缺省值 60，该路由与 Track 项 1 关联。

```
<DeviceB> system-view
```

```
[DeviceB] ip route-static 20.1.1.0 24 10.2.1.1 track 1
```

配置到达 20.1.1.0/24 网段的静态路由：下一跳地址为 10.4.1.3，优先级为 80。

```
[DeviceB] ip route-static 20.1.1.0 24 10.4.1.3 preference 80
```

配置 BFD echo 报文的源地址为 1.1.1.1。

```
[DeviceB] bfd echo-source-ip 1.1.1.1
```

创建和 BFD 会话关联的 Track 项 1，检测 Device B 是否可以与静态路由的下一跳 Device A 互通。

```
[DeviceB] track 1 bfd echo interface gigabitethernet 1/0/1 remote ip 10.2.1.1 local ip 10.2.1.2
```

(4) 配置 Device C

配置到达 30.1.1.0/24 网段的静态路由：下一跳地址为 10.4.1.2。

```
<DeviceC> system-view
```

```
[DeviceC] ip route-static 30.1.1.0 24 10.4.1.2
```

配置到达 20.1.1.0/24 网段的静态路由：下一跳地址为 10.3.1.1。

```
[DeviceB] ip route-static 20.1.1.0 24 10.3.1.1
```

4. 验证配置

显示 Device A 上 Track 项的信息。

```
[DeviceA] display track all
```

```
Track ID: 1
```

```
State: Positive
```

```
Duration: 0 days 0 hours 0 minutes 32 seconds
```

```
Notification delay: Positive 0, Negative 0 (in seconds)
```

```
Tracked object:
```

```
  BFD session mode: Echo
```

```
  Outgoing interface: GigabitEthernet1/0/1
```

```
  VPN instance name: -
```

```
  Remote IP: 10.2.1.2
```

```
  Local IP: 10.2.1.1
```

显示 Device A 的路由表。

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

```
Destinations : 9          Routes : 9
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
------------------	-------	-----	------	---------	-----------

```

10.2.1.0/24      Direct 0    0          10.2.1.1      GE1/0/1
10.2.1.1/32     Direct 0    0          127.0.0.1     InLoop0
10.3.1.0/24     Direct 0    0          10.3.1.1      GE1/0/2
10.3.1.1/32     Direct 0    0          127.0.0.1     InLoop0
20.1.1.0/24     Direct 0    0          20.1.1.1      GE1/0/3
20.1.1.1/32     Direct 0    0          127.0.0.1     InLoop0
30.1.1.0/24     Static 60   0          10.2.1.2      GE1/0/1
127.0.0.0/8     Direct 0    0          127.0.0.1     InLoop0
127.0.0.1/32    Direct 0    0          127.0.0.1     InLoop0

```

以上显示信息表示，BFD 检测的结果为下一跳地址 10.2.1.2 可达（Track 项状态为 Positive），主路由生效，Device A 通过 Device B 将报文转发到 30.1.1.0/24 网段。

在 Device B 上删除接口 GigabitEthernet1/0/1 的 IP 地址。

```

<DeviceB> system-view
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] undo ip address

```

显示 Device A 上 Track 项的信息。

```

[DeviceA] display track all
Track ID: 1
  State: Negative
  Duration: 0 days 0 hours 0 minutes 32 seconds
  Notification delay: Positive 0, Negative 0 (in seconds)
  Tracked object:
    BFD session mode: Echo
    Outgoing interface: GigabitEthernet1/0/1
    VPN instance name: -
Remote IP: 10.2.1.2
  Local IP: 10.2.1.1

```

显示 Device A 的路由表。

```

[DeviceA] display ip routing-table

```

```

Destinations : 9          Routes : 9

Destination/Mask    Proto  Pre  Cost           NextHop           Interface
10.2.1.0/24         Direct 0    0           10.2.1.1          GE1/0/1
10.2.1.1/32         Direct 0    0           127.0.0.1         InLoop0
10.3.1.0/24         Direct 0    0           10.3.1.1          GE1/0/2
10.3.1.1/32         Direct 0    0           127.0.0.1         InLoop0
20.1.1.0/24         Direct 0    0           20.1.1.1          GE1/0/3
20.1.1.1/32         Direct 0    0           127.0.0.1         InLoop0
30.1.1.0/24         Static 80   0           10.3.1.3          GE1/0/2
127.0.0.0/8         Direct 0    0           127.0.0.1         InLoop0
127.0.0.1/32        Direct 0    0           127.0.0.1         InLoop0

```

以上显示信息表示，BFD 检测的结果为下一跳地址 10.2.1.2 不可达（Track 项状态为 Negative），备份路由生效，Device A 通过 Device C 和 Device B 将报文转发到 30.1.1.0/24 网段。

主路由出现故障后，20.1.1.0/24 网段内的主机仍然可以与 30.1.1.0/24 网段内的主机通信。

```

[DeviceA] ping -a 20.1.1.1 30.1.1.1
Ping 30.1.1.1: 56 data bytes, press CTRL_C to break

```



```
Reply from 30.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
Reply from 30.1.1.1: bytes=56 Sequence=2 ttl=254 time=1 ms
Reply from 30.1.1.1: bytes=56 Sequence=3 ttl=254 time=1 ms
Reply from 30.1.1.1: bytes=56 Sequence=4 ttl=254 time=2 ms
Reply from 30.1.1.1: bytes=56 Sequence=5 ttl=254 time=1 ms
```

```
--- Ping statistics for 30.1.1.1 ---
```

```
5 packet(s) transmitted, 5 packet(s) received, 0.00% packet loss
round-trip min/avg/max/std-dev = 1/1/2/1 ms
```

Device B 上的显示信息与 Device A 类似。主路由出现故障后，30.1.1.0/24 网段内的主机仍然可以与 20.1.1.0/24 网段内的主机通信。

```
[DeviceB] ping -a 30.1.1.1 20.1.1.1
```

```
Ping 20.1.1.1: 56 data bytes, press CTRL_C to break
```

```
Reply from 20.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
Reply from 20.1.1.1: bytes=56 Sequence=2 ttl=254 time=1 ms
Reply from 20.1.1.1: bytes=56 Sequence=3 ttl=254 time=1 ms
Reply from 20.1.1.1: bytes=56 Sequence=4 ttl=254 time=1 ms
Reply from 20.1.1.1: bytes=56 Sequence=5 ttl=254 time=1 ms
```

```
--- Ping statistics for 20.1.1.1 ---
```

```
5 packet(s) transmitted, 5 packet(s) received, 0.00% packet loss
round-trip min/avg/max/std-dev = 1/1/2/1 ms
```

目 录

1 进程分布优化.....	1-1
1.1 进程分布优化简介.....	1-1
1.1.1 进程级HA.....	1-1
1.1.2 进程级 1:N备份	1-1
1.1.3 进程的分布策略与优化	1-2
1.2 进程分布优化配置任务简介.....	1-3
1.3 配置进程的分布策略.....	1-3
1.3.1 配置进程对节点位置的偏好	1-3
1.3.2 配置进程对节点类型的偏好	1-4
1.3.3 配置进程对其它进程的偏好	1-4
1.3.4 配置进程多实例之间的偏好	1-5
1.4 手工进行进程分布优化.....	1-5
1.5 进程分布优化的显示和维护.....	1-5

1 进程分布优化



说明

对于本节命令中的 CPU 参数，仅 T5000-M06 产品支持。

- 如果设备上只有一块主控板，则所有进程都运行在该主控板的主 CPU 上，此时，配置进程分布优化没有意义。
- 如果设备上有多个主控板，进程会进行 1:N 备份。但运行几个备份进程，运行在哪些 CPU 上，由功能模块决定，不过 N 肯定小于 CPU 的数量。当有备份进程运行时，用户配置进程分布策略，执行分布优化命令，可以让主控进程运行在用户期望的位置。通常情况下，进程在启动时，系统会给主控进程确定一个最佳运行位置，该位置不需要调整。如果确实需要调整，请在工程师的指导下进行，以免错误操作造成业务运行异常。
- 进程分布优化只涉及主控板，不涉及接口板。
- 在配置进程分布优化时，主控板上所有的 CPU 都是对等的，即便主 CPU 和辅助 CPU 物理上会有不同，分工会有不同，但一个 CPU 就是一个节点，可以配置进程在主 CPU 上运行，也可以配置主进程在辅助 CPU 上运行，最终结果可使用 **placement reoptimize** 命令了解。

1.1 进程分布优化简介

1.1.1 进程级 HA

进程是系统运行的基本单元，它能够给用户特定的服务。比如 aaa 进程能给用户 AAA 功能。

每个进程运行于受保护的内存空间，进程之间、进程与内核之间相互隔离。若某个进程出现问题不会影响整个系统。

1.1.2 进程级 1:N 备份

进程的 1:N 备份指的是：系统要运行某业务，会同时在设备的多个节点运行其对应进程。其中，一个为主控进程，其它 N 个为备份进程，主控进程处于工作状态，为用户提供服务，备份进程处于备用状态。

进程的 1:N 备份有以下优点：

- 保证了进程的高可用性。当主控进程故障时，系统能迅速从备份进程中选举一个新的主控进程，接替原主控进程继续为用户提供服务，不会影响到其它进程提供的业务。
- 提高了系统整体的稳定性，能适应设备的各种变化。如单板插拔、IRF 成员设备离开等，系统均可快速恢复稳定。
- 系统负载可在多节点间分担。不同业务的主进程分别指定到不同节点运行，能充分利用多个节点的 CPU 和内存资源，有效地分担系统的负载。

1.1.3 进程的分布策略与优化

1. 进程的分布策略

当设备启动或者用户配置/使能了某功能时，系统会自动同时为该业务运行一个主控进程和多个备份进程。

- 对于一些业务，其主控进程只能运行在主设备，这样的进程不支持进程分布优化配置（配置时会提示失败）。当主控进程异常时，系统会自动重启该主控进程。备份进程主要用于主备倒换和 ISSU 升级环境。另一些业务，其主控进程可以运行在主设备上，也可以运行在从设备上。当主控进程异常时，需要从备份进程中选举一个新的主控进程，从而保证业务不受影响。在众多的备份进程中到底选用哪个作为新的主控进程，由该进程的分布策略决定。（集中式 IRF 设备）
- 对于一些业务，其主控进程只能运行在主用主控板，这样的进程不支持进程分布优化配置（配置时会提示失败）。当主控进程异常时，系统会自动重启该主控进程。备份进程主要用于主备倒换和 ISSU 升级环境。另一些业务，其主控进程可以运行在主用主控板上，也可以运行在备用主控板上。当主控进程异常时，需要从备份进程中选举一个新的主控进程，从而保证业务不受影响。在众多的备份进程中到底选用哪个作为新的主控进程，由该进程的分布策略决定。（分布式设备—独立运行模式）
- 对于一些业务，其主控进程只能运行在全局主用主控板，这样的进程不支持进程分布优化配置（配置时会提示失败）。当主控进程异常时，系统会自动重启该主控进程。备份进程主要用于主备倒换和 ISSU 升级环境。另一些业务，其主控进程可以运行在全局主用主控板上，也可以运行在全局备用主控板上。当主控进程异常时，需要从备份进程中选举一个新的主控进程，从而保证业务不受影响。在众多的备份进程中到底选用哪个作为新的主控进程，由该进程的分布策略决定。（分布式设备—IRF 模式）

为了方便管理和维护进程，系统定义了缺省分布策略，该策略对所有进程以及进程的实例生效。设备出厂时，缺省分布策略如下：

- 主控进程运行在主用主控板的主 CPU 上，备份进程运行在备用主控板的主 CPU 上。（分布式设备—独立运行模式）
- 主控进程运行在主设备的主 CPU 上，备份进程运行在备设备的主 CPU 上。（集中式 IRF 设备）
- 主控进程运行在全局主用主控板的主 CPU 上，备份进程运行在全局备用主控板的主 CPU 上。（分布式设备—IRF 模式）
- 进程会继承上次运行的位置，启动和运行过程中都不会自动随意更改运行位置。
- 当有新节点加入时，对当前运行的进程没有影响，只对将要启动的进程有影响。即将启动的主控进程会选择一个 CPU 资源和内存资源更丰富的节点运行，其它节点上则运行该进程的备份进程。（可通过 **display cpu-usage** 和 **display memory** 命令了解设备当前 CPU 和内存的使用情况）

用户可以使用 **placement program default** 命令进入缺省分布策略视图，修改出厂缺省分布策略。也可以根据实际需要使使用 **placement program program-name [instance instance-name]** 命令进入进程的分布策略视图，为某个进程单独配置分布策略，对该进程的分布进行调整和优化，例如指定某些对资源有特殊要求的进程到特定的节点上运行，以取得更好的性能。和缺省分布策略相比，进程将优先应用单独配置的分布策略。进程的分布策略由进程的各种偏好（affinity）构成，包括：

- **location-set**: 进程对指定节点位置的偏好。

- **location-type**: 进程对指定节点类型的偏好。
- **program**: 进程对其它不同进程的偏好。
- **self**: 进程对同进程的其它实例的偏好。

以上四种偏好从不同角度表达了用户对进程在某个位置运行的期望，可以同时配置，并可以指定偏好程度。

偏好程度分为正向偏好程度和反向偏好程度，均用正整数表示，取值范围为 1~100000:

- 正向偏好程度表示希望运行在该位置，数值越大，表示期望越大，主进程运行在该位置的可能性越大。
- 反向偏好程度表示不希望运行在该位置，数值越大，主进程运行在该位置的可能性越小。

2. 进程的分布优化

配置了进程的分布策略之后，并不会立即改变进程的分布，用户还需要实施分布策略，才能对进程的分布进行优化。在优化进程分布的过程中，系统会综合进程分布策略、进程当前运行位置、设备节点是否在位和拓扑状态等因素，对优化后各进程理想的运行位置做出决策。如果为某进程决策出的位置不同于当前运行的位置，则对该进程执行主备倒换，即当前运行位置的主控进程降级为备份进程，预期位置的备份进程升级为主控进程。

1.2 进程分布优化配置任务简介

表1-1 进程分布优化配置任务简介

配置任务		说明	详细配置
配置进程的分布策略	配置进程对节点位置的偏好	请根据需要选择，可多选	1.3.1
	配置进程对节点类型的偏好		1.3.2
	配置进程对其它进程的偏好		1.3.3
	配置进程多实例之间的偏好		1.3.4
手工进行进程分布优化		必选	1.4

1.3 配置进程的分布策略

1.3.1 配置进程对节点位置的偏好

表1-2 配置进程对节点位置的偏好

操作	命令	说明
进入系统视图	system-view	-
进入进程的分布策略视图	placement program default	二者选其一 缺省进程分布策略视图下的配置对所有进程均生效；指定进程的分布策略视图下的配置只对指定进程生效
	placement program <i>program-name</i> [instance <i>instance-name</i>]	

操作	命令	说明
配置进程对节点位置的偏好 (集中式IRF设备/分布式设备-独立运行模式)	affinity location-set { slot slot-number [cpu cpu-number] } &<1-5> { attract strength default none repulse strength }	缺省情况下，系统未配置进程对节点位置的偏好
配置进程对节点位置的偏好 (分布式设备-IRF模式)	affinity location-set { chassis chassis-number slot slot-number [cpu cpu-number] } &<1-5> { attract strength default none repulse strength }	缺省情况下，系统未配置进程对节点位置的偏好

1.3.2 配置进程对节点类型的偏好

系统有 3 种节点类型：

- **current**: 主控进程当前运行的位置，可以通过 **display placement program** 命令查看。
- **paired**: 所有备份进程当前运行的位置。
- **primary**: 系统的主用主控板（分布式设备-独立运行模式）。
- **primary**: 主设备（集中式 IRF 设备）。
- **primary**: 全局主用主控板（分布式设备-IRF 模式）。

通过以下命令可以配置进程对某种节点类型的偏好。

表1-3 配置进程对节点类型的偏好

操作	命令	说明
进入系统视图	system-view	-
进入进程的分布策略视图	进入缺省进程分布策略视图 placement program default	二者选其一
	进入指定进程的分布策略视图 placement program program-name [instance instance-name]	缺省进程分布策略视图下的配置对所有进程均生效；指定进程的分布策略视图下的配置只对指定进程生效
配置进程对节点类型的偏好	affinity location-type { current paired primary } { attract strength repulse strength default none }	缺省情况下，系统未配置进程对节点类型的偏好

1.3.3 配置进程对其它进程的偏好

该配置以特定进程的运行位置为参照点，配置的是本进程和其它进程运行在同一位置的偏好。

表1-4 配置进程对其它进程的偏好

操作	命令	说明
进入系统视图	system-view	-
进入进程的分布策略视图	进入缺省进程分布策略视图 placement program default	二者选其一 缺省进程分布策略视图下的配

操作		命令	说明
图	进入指定进程的分布策略视图	placement program <i>program-name</i> [instance <i>instance-name</i>]	置对所有进程均生效；指定进程的分布策略视图下的配置只对指定进程生效
配置本进程和其它进程运行在同一位置的偏好		affinity program <i>program-name</i> { attract strength default none repulse strength }	缺省情况下，未配置进程和其它进程运行在同一位置的偏好

1.3.4 配置进程多实例之间的偏好

该配置用以决定一个进程的多个实例之间是否运行于同一个位置上，如果进程只有一个实例，则该配置不会产生作用。

表1-5 配置进程多实例之间的偏好

操作		命令	说明
进入系统视图		system-view	-
进入进程的分布策略视图	进入缺省进程分布策略视图	placement program default	二者选其一 缺省进程分布策略视图下的配置对所有进程均生效；指定进程的分布策略视图下的配置只对指定进程生效
	进入指定进程的分布策略视图	placement program <i>program-name</i> [instance <i>instance-name</i>]	
配置进程多实例之间的偏好		affinity self { attract strength repulse strength default none }	缺省情况下，系统未配置进程多实例之间的偏好

1.4 手工进行进程分布优化

表1-6 手工进行进程分布优化

操作	命令	说明
进入系统视图	system-view	-
手工进行进程分布优化	placement reoptimize	执行此命令时请保持系统的稳定性，不建议在执行此命令的过程中进行任务涉及进程重启的操作

1.5 进程分布优化的显示和维护

完成进程相关的配置后，在任意视图下执行 **display** 命令可以显示进程的分布策略配置、实际分布位置、分布优化的预测位置、服务器组的当前位置和状态等信息，通过查看显示信息可验证配置的效果。

表1-7 进程分布优化的显示和维护

操作	命令
显示进程的分布策略	display placement policy program { <i>program-name</i> all default }
显示进程的当前运行位置	display placement program { <i>program-name</i> all }
显示具体位置上正在运行的进程（集中式IRF设备/分布式设备—独立运行模式）	display placement location { slot <i>slot-number</i> [cpu <i>cpu-number</i>] all }
显示具体位置上正在运行的进程（分布式设备—IRF模式）	display placement location { chassis <i>chassis-number</i> slot <i>slot-number</i> [cpu <i>cpu-number</i>] all }
显示进程分布优化后的预测位置	display placement reoptimize program { <i>program-name</i> [instance <i>instance-name</i>] all }
显示服务组主控进程的位置和状态等信息	display ha service-group { <i>program-name</i> [instance <i>instance-name</i>] all }