



H3C S5500-HI-D 系列以太网交换机



IRF 配置指导

杭州华三通信技术有限公司
<http://www.h3c.com.cn>

资料版本: 6W101-20141025
产品版本: Release 52xx 系列

Copyright © 2013-2014 杭州华三通信技术有限公司及其许可者 版权所有，保留一切权利。

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

H3C、**H3C**、H3CS、H3CIE、H3CNE、Aolynk、、H³Care、、IRF、NetPilot、Netflow、SecEngine、SecPath、SecCenter、SecBlade、Comware、ITCMM、HUASAN、华三均为杭州华三通信技术有限公司的商标。对于本手册中出现的其它公司的商标、产品标识及商品名称，由各自权利人拥有。

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

前言

H3C S5500-HI-D 系列以太网交换机配置指导共分为十本手册，介绍了 S5500-HI-D 系列以太网交换机 Release52xx 系列软件版本各软件特性的原理及其配置方法，包含原理简介、配置任务描述和配置举例。《IRF 配置指导》主要介绍如何使用多台 S5500-HI-D 交换机组建基于 IRF 技术的虚拟化设备，包括规划 IRF 中设备的角色、IRF 链路连接、以及 IRF 形成后的检测和维护等内容。

前言部分包含如下内容：

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

读者对象

本手册主要适用于如下工程师：

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

新增及修改特性说明

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

- Release5206 与 Release5203 版本相比，无新增、修改特性。
- Release5203 与 Release5101 版本相比，无新增、修改特性。

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。






格 式	意 义
[x y ...]*	表示从多个选项中选择的一个、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参考，并不代表设备上实际具有此编号的端口，实际使用中请以设备上存在的端口编号为准。

产品配套资料

H3C S5500-HI-D 系列以太网交换机的配套资料包括如下部分：

大类	资料名称	内容介绍
硬件介绍及安装	安全手册	帮助您了解安装和日常维护本设备需要注意的安全规范
	H3C 设备防雷安装指导手册	帮助您了解防雷接地设计和工程安装方法，以保证交换机具有良好的抗雷击性能
	快速安装指南	指导您对设备进行初始安装，通常针对最常用的情况，减少您的检索时间
	安装指导	帮助您详细了解设备硬件规格和安装方法，指导您对设备进行安装
	风扇安装手册	帮助您了解产品支持的可插拔风扇模块的外观、功能、规格、安装及拆卸方法
	电源手册	帮助您了解产品支持的可插拔电源模块的外观、功能、规格、安装及拆卸方法
	RPS电源用户手册	帮助您了解产品支持的RPS电源的外观、功能、规格
	接口模块扩展卡用户手册	帮助您了解该接口模块扩展卡的外观、规格、安装及拆卸方法
	H3C光模块手册	帮助您了解产品支持的可插拔模块类型、外观和规格
业务配置	H3C可插拔SFP[SFP+][XFP]模块安装指南	帮助您掌握SFP/SFP+/XFP模块的正确安装方法，避免因操作不当而造成器件损坏
	配置指导	帮助您掌握设备软件功能的配置方法及配置步骤
运行维护	命令参考	详细介绍设备的命令，相当于命令字典，方便您查阅各个命令的功能
	版本说明书	帮助您了解产品版本的相关信息（包括：版本配套说明、兼容性说明、特性变更说明、技术支持信息）及软件升级方法

资料获取方式

您可以通过H3C网站（www.h3c.com.cn）获取最新的产品资料：

H3C 网站与产品资料相关的主要栏目介绍如下：

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

技术支持

用户支持邮箱：service@h3c.com

技术支持热线电话：400-810-0504（手机、固话均可拨打）

网址：<http://www.h3c.com.cn>

资料意见反馈

如果您在使用过程中发现产品资料的任何问题，可以通过以下方式反馈：

E-mail：info@h3c.com

感谢您的反馈，让我们做得更好！

目 录

1 IRF配置	1-1
1.1 IRF简介	1-1
1.1.1 IRF概述	1-1
1.1.2 IRF的优点	1-1
1.1.3 IRF的应用	1-1
1.2 IRF基本概念	1-2
1.3 IRF工作原理	1-3
1.3.1 物理连接	1-3
1.3.2 拓扑收集	1-5
1.3.3 角色选举	1-6
1.3.4 IRF的管理与维护	1-6
1.3.5 多IRF冲突检测（MAD功能）	1-8
1.4 IRF配置任务简介	1-9
1.5 IRF配置	1-10
1.5.1 配置IRF域编号	1-11
1.5.2 配置成员编号	1-12
1.5.3 配置IRF端口	1-13
1.5.4 配置成员优先级	1-14
1.5.5 配置成员设备的描述信息	1-15
1.5.6 配置IRF链路的负载分担类型	1-15
1.5.7 配置IRF的桥MAC保留时间	1-16
1.5.8 使能IRF系统启动文件的自动加载功能	1-17
1.5.9 配置IRF链路down延迟上报功能	1-18
1.6 MAD配置	1-18
1.6.1 LACP MAD检测	1-19
1.6.2 BFD MAD检测	1-21
1.6.3 ARP MAD检测	1-24
1.6.4 配置保留接口	1-26
1.6.5 MAD故障恢复	1-26
1.7 访问IRF	1-28
1.7.1 访问Master	1-28
1.7.2 访问Slave	1-28
1.8 IRF显示和维护	1-29

1.9 IRF典型配置举例	1-29
1.9.1 IRF典型配置举例（LACP MAD检测方式）	1-29
1.9.2 IRF典型配置举例（BFD MAD检测方式）	1-32
1.9.3 IRF典型配置举例（ARP MAD检测方式）	1-35

1 IRF配置



说明

- 本章中提到的三层以太网端口是指被配置为三层模式以太网端口,有关以太网端口模式切换的操作,请参见“二层技术-以太网交换配置指导”中的“以太网端口”
- S5500-HI-D 系列交换机按设备型号可以分为两组: 第一组包括 S5500-28SC-HI-D 和 S5500-52SC-HI-D; 其它款型属于第二组。同一组的设备之间可以建立 IRF, 不同组的设备间无法建立 IRF。

1.1 IRF简介

1.1.1 IRF概述

IRF (Intelligent Resilient Framework, 智能弹性架构) 是 H3C 自主研发的软件虚拟化技术。它的核心思想是将多台设备通过 IRF 物理端口连接在一起, 进行必要的配置后, 虚拟化成一台“分布式设备”。使用这种虚拟化技术可以集合多台设备的硬件资源和软件处理能力, 实现多台设备的协同工作、统一管理和不间断维护。

为了便于描述, 这个“虚拟设备”也称为 IRF。所以, 本文中的 IRF 有两层意思, 一个是指 IRF 技术, 一个是指 IRF 设备。

1.1.2 IRF的优点

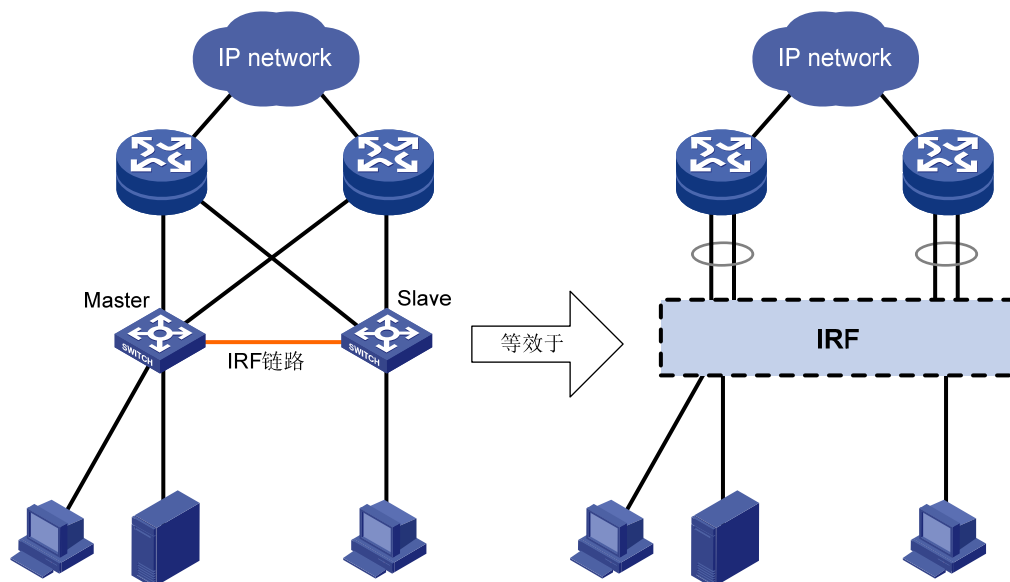
IRF 主要具有以下优点:

- 简化管理。IRF 形成之后, 用户通过任意成员设备的任意端口都可以登录 IRF 系统, 对 IRF 内所有成员设备进行统一管理。
- 高可靠性。IRF 的高可靠性体现在多个方面, 例如: IRF 由多台成员设备组成, Master 设备负责 IRF 的运行、管理和维护, Slave 设备在作为备份的同时也可以处理业务。一旦 Master 设备故障, 系统会迅速自动选举新的 Master, 以保证业务不中断, 从而实现了设备的 1:N 备份; 此外, 成员设备之间的 IRF 链路支持聚合功能, IRF 和上、下层设备之间的物理链路也支持聚合功能, 多条链路之间可以互为备份也可以进行负载分担, 从而进一步提高了 IRF 的可靠性。
- 强大的网络扩展能力。通过增加成员设备, 可以轻松自如的扩展 IRF 的端口数、带宽。因为各成员设备都有 CPU, 能够独立处理协议报文、进行报文转发, 所以 IRF 还能够轻松自如的扩展处理能力。

1.1.3 IRF的应用

如 [图 1-1](#) 所示, Master和Slave组成IRF, 对上、下层设备来说, 它们就是一台设备——IRF。

图1-1 IRF 组网应用示意图



1.2 IRF基本概念

IRF 虚拟化技术涉及如下基本概念：

1. 角色

IRF 中每台设备都称为成员设备。成员设备按照功能不同，分为两种角色：

- **Master:** 负责管理整个 IRF。
- **Slave:** 作为 Master 的备份设备运行。当 Master 故障时，系统会自动从 Slave 中选举一个新的 Master 接替原 Master 工作。

Master和Slave均由角色选举产生。一个IRF中同时只能存在一台Master，其它成员设备都是Slave。关于设备角色选举过程的详细介绍请参见 [1.3.3 角色选举](#)。

2. IRF端口

一种专用于 IRF 的逻辑接口，分为 IRF-Port1 和 IRF-Port2。它需要和物理端口绑定之后才能生效。

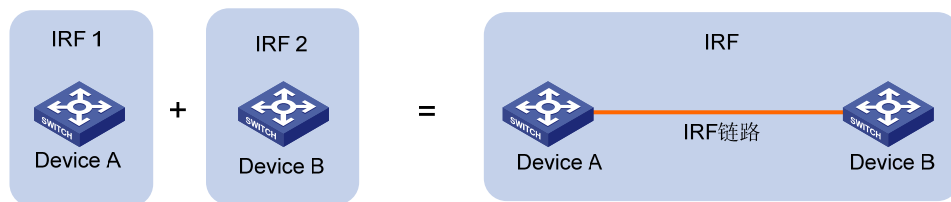
3. IRF物理端口

设备上可以用于 IRF 连接的物理端口。S5500-HI-D 系列交换机可以使用前面板上的 SFP+口或接口模块扩展卡上的万兆口作为 IRF 物理端口。通常情况下，接口用于传输业务报文，当它们与 IRF 端口绑定后就作为 IRF 物理端口，用于成员设备之间转发报文。可转发的报文包括 IRF 相关协商报文以及需要跨成员设备转发的业务报文。

4. IRF合并

如 [图 1-2](#) 所示，两个IRF各自已经稳定运行，通过物理连接和必要的配置，形成一个IRF，这个过程称为IRF合并（merge）。

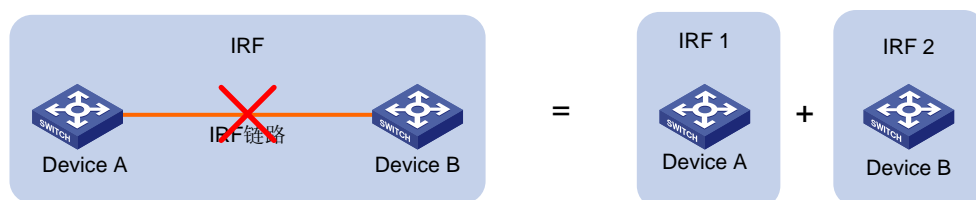
图1-2 IRF 合并示意图



5. IRF分裂

如 [图 1-3](#) 所示，一个IRF形成后，由于IRF链路故障，导致IRF中两相邻成员设备物理上不连通，一个IRF变成两个IRF，这个过程称为IRF分裂（split）。

图1-3 IRF 分裂示意图



6. 成员优先级

成员优先级是成员设备的一个属性，主要用于角色选举过程中确定成员设备的角色。优先级越高当选为 **Master** 的可能性越大。

设备的缺省优先级均为 1，如果想让某台设备当选为 **Master**，则在组建 IRF 前，可以通过命令行手工提高该设备的成员优先级。

1.3 IRF工作原理

IRF系统将经历 [物理连接](#)、[拓扑收集](#)、[角色选举](#)、[IRF的管理与维护](#)四个阶段。成员设备之间需要先建立IRF物理连接，然后会自动进行拓扑收集和角色选举，完成IRF的建立，此后进入IRF管理和维护阶段。

1.3.1 物理连接

1. 连接介质

要形成一个 IRF，需要先连接成员设备的 IRF 物理端口。S5500-HI-D 系列交换机使用前面板上的 SFP+口或接口模块扩展卡上的万兆口作为 IRF 物理端口。可用于提供 IRF 物理端口的接口模块扩展卡包括：

- 短距双端口 10GE CX4 接口模块扩展卡（LSPM1CX2P）
- 双端口 10GE SFP+接口模块扩展卡（LSPM2SP2P）
- 双端口 10GE XFP 接口模块扩展卡（LSPM1XP2P）
- 单端口 10GE XFP 接口模块扩展卡（LSPM1XP1P）
- 双端口 10GE 以太网接口模块扩展卡（LSPM1XGT2P）

根据您使用的接口类型，您可以选择 CX4 电缆、SFP+电缆、双绞线、或 SFP+/XFP 光模块与光纤在多台成员设备之间进行物理连接。

SFP+/XFP 模块与光纤的搭配适用于在距离很远的设备间进行 IRF 连接，使得应用更加灵活；而双绞线/CX4 电缆/SFP+电缆适用于机房内部短距离的 IRF 连接。

说明

- S5500-HI-D 支持的 SFP+/XFP 模块以及 SFP+/CX4 电缆请参见《H3C S5500-HI-D 系列以太网交换机 安装指导》。
- 有关各接口模块的详细介绍，请参见《H3C 低端系列以太网交换机 可插拔模块手册》。
- H3C SFP+模块和 SFP+电缆的种类随着时间变化有更新的可能性，所以，若您需要准确的模块种类信息，请咨询 H3C 公司市场人员或技术支持人员。

2. 连接要求

IRF 端口的连接是基于 IRF 物理端口的连接而建立的，因此需要将 IRF 端口和 IRF 物理端口对应起来，即将 IRF 端口和 IRF 物理端口进行绑定。一个 IRF 端口可以对应一个 IRF 物理端口，也可以通过对应多个 IRF 物理端口来实现聚合 IRF 端口，以达到链路备份和扩展带宽的效果。

在 S5500-HI-D 系列交换机上进行 IRF 端口和 IRF 物理端口的绑定时，需要注意不同机型对于绑定关系的限制。

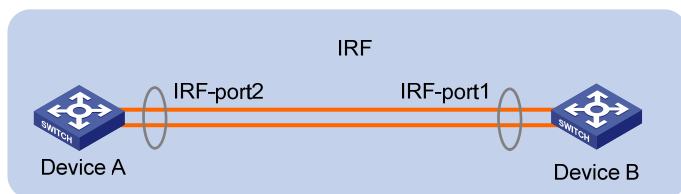
表1-1 S5500-HI-D 系列交换机上 IRF 端口与 IRF 物理端口的对应方式

设备型号	IRF 端口与 IRF 物理端口的对应方式
S5500-28SC-HI-D S5500-34C-HI-D	前面板上的2个SFP+口以及后面板上通过插入接口模块扩展卡得到的IRF物理端口可以对应到任意IRF端口
S5500-58C-HI-D	当配置聚合IRF端口时： <ul style="list-style-type: none">● 同一块接口模块扩展卡上的 IRF 物理端口必须对应到相同的 IRF 端口● 前面板编号为 53 的 SFP+口只能与槽位 2 中的接口模块扩展卡上的 IRF 物理端口对应到相同的 IRF 端口● 前面板编号为 54 的 SFP+口只能与槽位 1 中的接口模块扩展卡上的 IRF 物理端口对应到相同的 IRF 端口
S5500-52SC-HI-D	当配置聚合IRF端口时： <ul style="list-style-type: none">● 如果您选用双端口的接口模块扩展卡，则前面板编号 51 的 SFP+口只能与接口模块扩展卡的 1 号端口对应到相同的 IRF 端口，前面板编号 52 的 SFP+口只能与接口模块扩展卡的 2 号端口对应到相同的 IRF 端口。● 如果您选用单端口的接口模块扩展卡，只能将前面板编号 51 的 SFP+口与扩展卡上的端口对应到相同的 IRF 端口。

在 S5500-HI-D 系列交换机中，S5500-52SC-HI-D 和 S5500-28SC-HI-D 交换机最多支持将 2 个 IRF 物理端口与同一个 IRF 端口进行绑定，其它机型最多支持将 3 个 IRF 物理端口与同一个 IRF 端口进行绑定。

在进行成员设备的连接时,本设备上与IRF-Port1 绑定的IRF物理端口只能和邻居成员设备IRF-Port2 口上绑定的IRF物理端口相连,本设备上与IRF-Port2 口绑定的IRF物理端口只能和邻居成员设备IRF-Port1 口上绑定的IRF物理端口相连,如 图 1-4 所示。否则,不能形成IRF。

图1-4 IRF 物理连接示意图

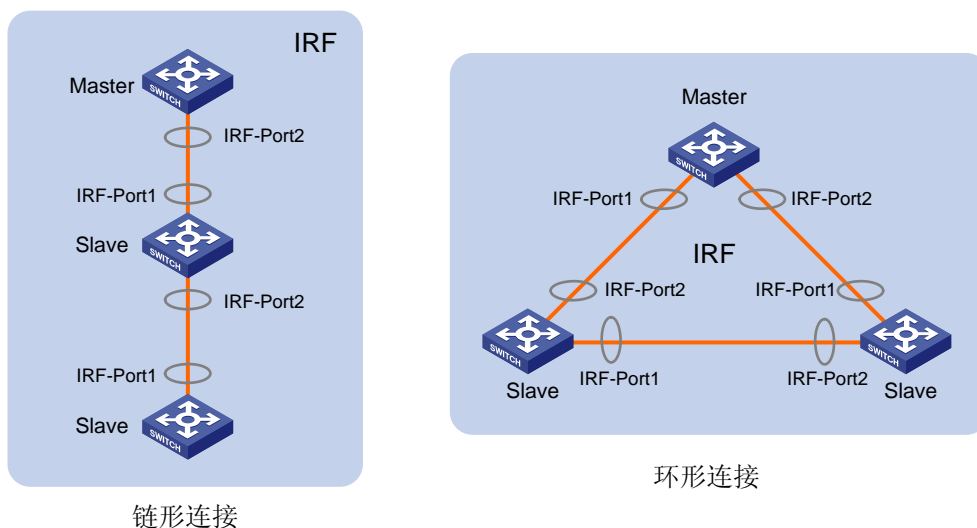


3. 连接拓扑

IRF的连接拓扑有两种：链形连接和环形连接,如 图 1-5 所示。

- 相比环形连接,链形连接对成员设备的物理位置要求更低,主要用于成员设备物理位置分散的组网。
- 环形连接比链形连接更可靠。因为当链形连接中出现链路故障时,会引起 IRF 分裂;而环形连接中某条链路故障时,会形成链形连接,IRF 的业务不会受到影响。

图1-5 IRF 连接拓扑示意图



1.3.2 拓扑收集

每个成员设备和邻居成员设备通过交互 IRF Hello 报文来收集整个 IRF 的拓扑。IRF Hello 报文会携带拓扑信息,具体包括 IRF 端口连接关系、成员设备编号、成员设备优先级、成员设备的桥 MAC 等内容。

每个成员设备在本地记录自己已知的拓扑信息。设备刚启动时只记录了自身的拓扑信息。当 IRF 端口状态变为 up 后,设备会将已知的拓扑信息周期性的从 up 状态的 IRF 端口发送出去;直接邻居收到该信息后,会更新本地记录的拓扑信息;如此往复,经过一段时间的收集,所有成员设备都会收集到完整的拓扑信息(称为拓扑收敛)。

此时会进入角色选举阶段。

1.3.3 角色选举

确定成员设备角色为 **Master** 或 **Slave** 的过程称为角色选举。

角色选举会在拓扑变更的情况下产生，比如 **IRF** 建立、新设备加入、**Master** 设备离开或者故障、两个 **IRF** 合并等。角色选举规则如下：

- (1) 当前 **Master** 优先(**IRF** 系统形成时, 没有 **Master** 设备, 所有加入的设备都认为自己是 **Master**, 会跳转到第二条规则继续比较)；
- (2) 成员优先级大的优先；
- (3) 系统运行时间长的优先；
- (4) 桥 **MAC** 地址小的优先。

从第一条开始判断，如果判断的结果是多个最优，则继续判断下一条，直到找到唯一最优的成员设备才停止比较。此最优成员设备即为 **Master**，其它成员设备则均为 **Slave**。

在角色选举完成后，**IRF** 形成，进入 **IRF** 管理与维护阶段。



说明

- **IRF** 合并的情况下，两个 **IRF** 会进行 **IRF** 竞选，竞选仍然遵循角色选举的规则，竞选完成后，失败一方将自动重新启动，然后全部作为 **Slave** 设备加入竞选获胜的一方，最终合并为一个 **IRF**。
 - 不管设备与其它设备一起形成 **IRF**，还是加入已有 **IRF**，如果该设备被当选为 **Slave**，则该设备会使用 **Master** 的配置重新初始化和启动，以保证和 **Master** 上的配置一致，而不管该设备在重新初始化之前有哪些配置、是否保存了当前配置。
-

1.3.4 **IRF**的管理与维护

角色选举完成之后，**IRF** 形成，所有的成员设备组成一台虚拟设备存在于网络中，所有成员设备上的资源归该虚拟设备拥有并由 **Master** 统一管理。

1. 成员编号

在运行过程中，**IRF** 系统使用成员编号 (**Member ID**) 来标志和管理成员设备，并在端口编号和文件系统中引入成员编号的标识信息。该编号关系到整个 **IRF** 的管理和运行，因此，需要用户在设备加入 **IRF** 前统一规划、配置设备的成员编号，以保证 **IRF** 中成员编号的唯一性。

2. 接口命名规则

对于单独运行的设备(即没有加入任何 **IRF**)，接口编号采用设备编号/子槽位编号/接口序号的格式，其中：

- 缺省情况下，设备编号为 1。
- 如果设备曾经加入过 **IRF**，则在退出 **IRF** 后，仍然会使用在 **IRF** 中时的成员编号作为自身的设备编号。

- 子槽位编号：接口所在子槽位的编号。对于 S5500-HI-D 系列交换机，前面板固定端口的子槽位编号为 0；对于后面板具有单扩展插槽的设备，扩展槽位的子槽位编号为 1；对于后面板具有双扩展插槽的设备，扩展槽位的子槽位编号分别为 1 和 2。
- 接口序号与各型号交换机支持的接口数量相关，请查看设备接口板上的丝印。

比如，要将单独运行的设备 Sysname 的接口 GigabitEthernet1/0/1 的接口链路类型设置为 Trunk，可参照以下步骤：

```
<Sysname> system-view
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] port link-type trunk
```

对于 IRF 中的成员设备，接口编号仍然采用成员设备编号/子槽位编号/接口序号的格式，其中：

- 成员设备编号用来标志不同成员设备上的接口。
- 子槽位编号和接口序号的含义和取值与单独运行时的一样。

比如，要将 IRF 中的成员设备 Slave3（成员编号为 3）子槽位 0 上第一个端口的链路类型设置为 Trunk，可参照以下步骤：

```
<Sysname> system-view
[Sysname] interface gigabitethernet 3/0/1
[Sysname-GigabitEthernet3/0/1] port link-type trunk
```

3. 文件系统命名规则

对于单独运行的设备，直接使用存储介质的名称就可以访问设备的文件系统了（存储介质的命名请参见“基础配置指导”中的“文件系统管理配置”）。

对于 IRF 中的成员设备，直接使用存储介质的名称可以访问 Master 设备的文件系统，使用“slotMember-ID#存储介质的名称”才可以访问 Slave 设备的文件系统。

比如：

(1) 创建并访问 IRF 中 Master 设备存储介质 Flash 根目录下的 test 文件夹，可参照以下步骤：

```
<Master> mkdir test
...
%Created dir flash:/test.
<Master> dir
Directory of flash:/
   0   -rw-   10105088   Apr 26 2000 13:44:57   test.bin
   1   -rw-     2445     Apr 26 2000 15:18:19   config.cfg
   2   drw-     -       Jul 14 2008 15:20:35   test
515712 KB total (505812 KB free)
```

(2) 创建并访问 IRF 中 Slave 设备（成员编号为 3）存储介质 Flash 根目录下的 test 文件夹，可参照以下步骤：

```
<Master> mkdir slot3#flash:/test
%Created dir slot3#flash:/test.
<Master> cd slot3#flash:/test
<Master> pwd
slot3#flash:/test
```

或者：

```
<Master> cd slot3#flash:/
<Master> mkdir test
```



```
%Created dir slot3#flash:/test.
```

(3) 将 Master 的 test.bin 文件拷贝到 Slave3 Flash 的根目录下，可参照以下步骤：

```
<Master> pwd
slot3#flash:
// 以上显示信息表明，当前的工作路径是 Slave3 Flash 的根目录
<Master> cd flash:/
<Master> pwd
flash:
// 以上操作表明，当前的工作路径已经回到了 Master Flash 的根目录
<Master> copy test.bin slot3#flash:/
Copy flash:/test.bin to slot3#flash:/test.bin?[Y/N]:y
%Copy file flash:/test.bin to slot3#flash:/test.bin...Done.
```

4. 配置文件的同步

IRF 技术使用了严格的配置文件同步机制，来保证 IRF 中的多台设备能够像一台设备一样在网络中工作，并且在 Master 设备出现故障之后，其余设备仍能够正常执行各项功能。

- IRF 中的 Slave 设备在启动时，会自动寻找 Master 设备，并将 Master 设备的当前配置文件同步到本地并执行；如果 IRF 中的所有设备同时启动，则 Slave 设备会将 Master 设备的起始配置文件同步至本地并执行。
- 在 IRF 正常工作后，用户所进行的任何配置，都会记录到 Master 设备的当前配置文件中，并同步到 IRF 中的各个设备执行；用户在执行 **save** 命令时，如果开启了配置文件同步保存功能（缺省为开启），Master 设备的当前配置文件将被同步保存到 IRF 的所有成员设备上，作为起始配置文件，以便使 IRF 中所有设备的起始配置文件保持一致；如果未开启配置文件同步保存功能，当前配置文件将仅在 Master 设备上保存。

通过即时的同步，IRF 中所有设备均保存有相同的配置文件，即使 Master 设备出现故障，其它设备仍能够按照相同的配置文件执行各项功能。

5. IRF 拓扑维护

如果某成员设备 A down 或者 IRF 链路 down，其邻居设备会立即将“成员设备 A 离开”的信息广播通知给 IRF 中的其它设备。获取到离开消息的成员设备会根据本地维护的 IRF 拓扑信息表来判断离开的是 Master 还是 Slave，如果离开的是 Master，则触发新的角色选举，再更新本地的 IRF 拓扑；如果离开的是 Slave，则直接更新本地的 IRF 拓扑，以保证 IRF 拓扑能迅速收敛。



说明

IRF 端口的状态由与它绑定的 IRF 物理端口的状态决定。与 IRF 端口绑定的所有 IRF 物理端口状态均为 down 时，IRF 端口的状态才会变成 down。

1.3.5 多 IRF 冲突检测（MAD 功能）

IRF 链路故障会导致一个 IRF 变成两个新的 IRF。这两个 IRF 拥有相同的 IP 地址等三层配置，会引起地址冲突，导致故障在网络中扩大。为了提高系统的可用性，当 IRF 分裂时我们就需要一种机制，能够检测出网络中同时存在多个 IRF，并进行相应的处理尽量降低 IRF 分裂对业务的影响。MAD（Multi-Active Detection，多 Active 检测）就是这样一种检测和处理机制。它主要提供以下功能：

(1) 分裂检测

通过 LACP(Link Aggregation Control Protocol, 链路聚合控制协议)、BFD(Bidirectional Forwarding Detection, 双向转发检测) 或者免费 ARP (Gratuitous Address Resolution Protocol) 来检测网络中是否存在多个 IRF。

(2) 冲突处理

IRF 分裂后, 通过分裂检测机制 IRF 会检测到网络中存在其它处于 Active 状态 (表示 IRF 处于正常工作状态) 的 IRF。冲突处理会让 Master 成员编号最小的 IRF 继续正常工作 (维持 Active 状态), 其它 IRF 会迁移到 Recovery 状态 (表示 IRF 处于禁用状态), 并关闭 Recovery 状态 IRF 中所有成员设备上除保留端口以外的其它所有物理端口 (通常为业务接口), 以保证该 IRF 不能再转发业务报文。(缺省情况下, 只有 IRF 物理端口是保留端口, 如果要将其它端口, 比如用于远程登录的端口, 也作为保留端口, 需要使用命令行进行手工配置。)

(3) MAD 故障恢复

IRF 链路故障导致 IRF 分裂, 从而引起多 Active 冲突。因此修复故障的 IRF 链路, 让冲突的 IRF 重新合并为一个 IRF, 就能恢复 MAD 故障。如果在 MAD 故障恢复前, 处于 Recovery 状态的 IRF 也出现了故障, 则需要将故障 IRF 和故障链路都修复后, 才能让冲突的 IRF 重新合并为一个 IRF, 恢复 MAD 故障; 如果在 MAD 故障恢复前, 故障的是 Active 状态的 IRF, 则可以通过命令行先启用 Recovery 状态的 IRF, 让它接替原 IRF 工作, 以便保证业务尽量少受影响, 再恢复 MAD 故障。



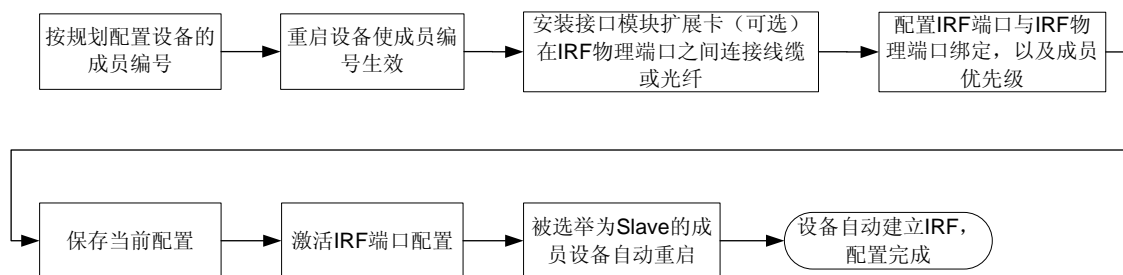
说明

- IRF 分裂后, 竞选失败的 IRF 会自动关闭所有成员设备上的部分端口 (等效于在接口下执行 **shutdown** 命令), 但有些端口不会被自动关闭, 这些端口称为保留端口。缺省情况下, 只有 IRF 物理端口是保留端口, 如果要将其它端口 (比如用于远程登录的端口) 也作为保留端口, 需要使用命令行进行手工配置。
- 关于 LACP 的详细介绍请参见“二层技术-以太网交换配置指导”中的“以太网链路聚合”; 关于 BFD 的详细介绍请参见“可靠性配置指导”中的“BFD”; 关于免费 ARP 的详细介绍请参见“三层技术-IP 业务配置指导”中的“免费 ARP”。

1.4 IRF配置任务简介

用户配置 IRF 前, 要做好前期规划工作, 需要明确 IRF 内各成员设备的角色和功能。因为有些参数的配置需要重启设备才能生效, 所以建议用户按照下面的流程进行配置。

图1-6 IRF 配置流程图



用户也可以在激活 IRF 端口配置后再进行 IRF 物理连接，当设备检测到 IRF 端口正常连接后，将立刻开始角色选举，选举为 **Slave** 的设备将自动重启。

在 IRF 形成后，用户通过 IRF 中的任意一台设备进行登录，均可以对 IRF 系统进行配置和管理。

表1-2 IRF 配置任务简介

配置任务	说明	详细配置	
配置IRF域编号	可选	1.5.1	
配置成员编号	必选	1.5.2	
配置IRF端口	必选	1.5.3	
配置成员优先级	可选	1.5.4	
配置成员设备的描述信息	可选	1.5.5	
配置IRF链路的负载分担类型	可选	1.5.6	
配置IRF的桥MAC保留时间	可选	1.5.6	
使能IRF系统启动文件的自动加载功能	可选	1.5.8	
配置IRF链路down延迟上报功能	可选	1.5.9	
连接IRF物理端口，确保IRF链路是连通的（推荐使用环形连接）			
MAD配置	LACP MAD检测	选择其中一种方式，也可以多选 建议用户在IRF建立完成后配置MAD检测功能	1.6.1
	BFD MAD检测		1.6.2
	ARP MAD检测		1.6.3
	配置保留接口	可选	1.6.4
	MAD故障恢复	可选	1.6.5
访问IRF	访问Master	必选	1.7.1
	访问Slave	可选	1.7.2

1.5 IRF配置



说明

在组建 IRF 前，请确认各成员设备 FIPS 功能的开启/关闭状态一致。关于 FIPS 功能的介绍，请参见“安全配置指导”中的“FIPS”。

1.5.1 配置IRF域编号

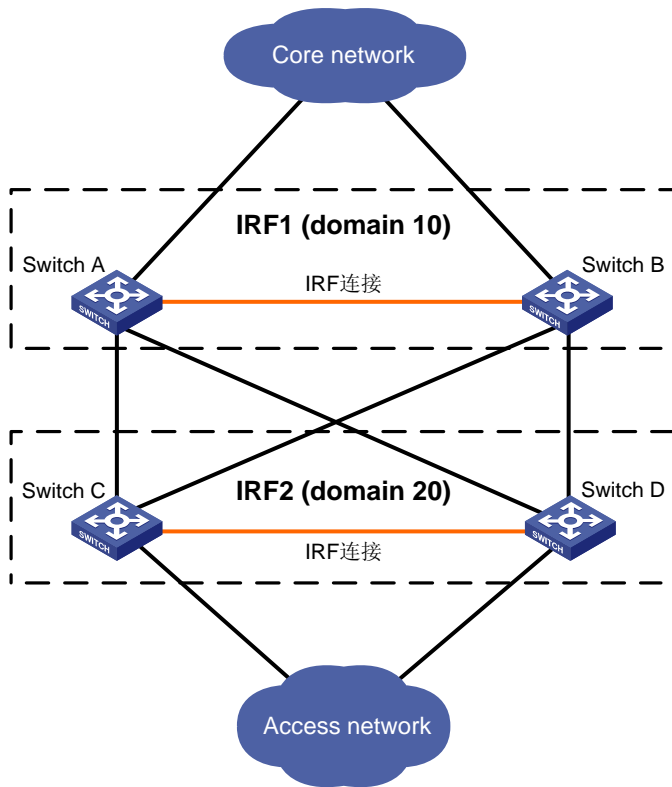
1. IRF域简介

域是一个逻辑概念，设备通过 IRF 链路连接在一起就组成一个 IRF，这些成员设备的集合就是一个 IRF 域。

为了适应各种组网应用，同一个网络里可以部署多个IRF，IRF之间使用域编号（DomainID）来以示区别。如 图 1-7 所示，Switch A和Switch B组成IRF1，Switch C和Switch D组成IRF2。如果IRF1和IRF2 之间有LACP MAD检测链路，则IRF1 和IRF2 会通过检测链路互相发送MAD检测报文，从而彼此影响IRF系统的状态和运行。这种情况下，可以给两个IRF配置不同的域编号，以保证两个IRF互不干扰。

配置 IRF 域编号后，成员设备发出的扩展 LACP 报文中将携带 IRF 域信息，用以区分不同 IRF 的 LACP 检测报文，避免与其它 IRF 产生混淆。

图1-7 多 IRF 域示意图



2. 配置IRF域编号

表1-3 配置 IRF 域编号

操作	命令	说明
进入系统视图	system-view	-
配置IRF域编号	irf domain domain-id	对于LACP MAD和ARP MAD检测方式该步骤必选 对于BFD MAD检测方式该步骤可选 缺省情况下，IRF的域编号为0



说明

- IRF 域编号的配置必须在开启 LACP MAD 和 ARP MAD 检测功能之前进行。
- 建议用户为同一 IRF 中的成员设备配置统一的 IRF 域编号，否则会影响 LACP MAD 和 ARP MAD 检测功能的正常运行。
- 在完成上述配置后，在任意视图下执行 **display irf** 命令可以显示 IRF 域编号的配置情况，通过查看显示信息验证配置的效果。

1.5.2 配置成员编号

IRF 通过成员编号唯一的识别各成员设备，设备上的许多信息、配置与成员编号相关，比如接口（包括物理接口和逻辑接口）的编号以及接口下的配置、成员优先级的配置等。

- 修改成员编号后，如果没有重启本设备，则原编号继续生效，各物理资源仍然使用原编号来标识；配置文件中，只有 IRF 端口的编号以及 IRF 端口下的配置、成员优先级的配置会跟着改变，其它配置均不会跟着改变。
- 修改成员编号后，如果保存当前配置，重启本设备，则新的成员编号生效，需要用新编号来标识物理资源；配置文件中，只有 IRF 端口的编号以及 IRF 端口下的配置、成员优先级会继续生效，其它与成员编号相关的配置（比如普通物理接口的配置等）不再生效，需要重新配置。

表1-4 配置成员编号

操作	命令	说明
进入系统视图	system-view	-
配置成员编号	irf member <i>member-id</i> renumber <i>new-member-id</i>	可选 缺省情况下，设备的成员编号均为1



注意

- 该配置需要重启 *member-id* 标志的设备才能生效；
- 在 IRF 中以成员编号标志设备，修改设备成员编号可能导致设备配置发生变化或者丢失。例如，IRF 中有三台设备（编号为 1、2、3），假定设备型号一样，每台设备都有若干端口，将设备 2 的成员编号改为 3，将设备 3 的成员编号改为 2，然后将设备 2 和 3 重启，再次加入 IRF 中，此时设备 2 将会使用先前设备 3 的接口配置，而设备 3 则使用先前设备 2 的端口配置。因此，请在建立 IRF 前规划好所有设备的成员编号，在 IRF 形成后尽量减少不必要的成员编号修改操作。

1.5.3 配置IRF端口

只有配置了 IRF 端口与 IRF 物理端口的绑定关系，并激活了 IRF 端口的配置之后，设备的 IRF 功能才能开始运行。

在完成成员设备间的物理连接，并将 IRF 物理端口与状态为 DIS 或 DOWN（可以使用 **display irf topology** 命令来查看）的 IRF 端口进行绑定后，还需要通过 **irf-port-configuration active** 命令激活 IRF 端口的配置。在执行激活操作后，当 IRF 端口状态变为 UP 时，设备间会进行 Master 竞选，竞选失败的设备会自动重启并以 Slave 身份加入 IRF。

表1-5 配置 IRF 端口

操作	命令	说明
进入系统视图	system-view	-
进入IRF物理端口视图	interface <i>interface-type</i> <i>interface-number</i>	-
关闭接口	shutdown	必选
退回系统视图	quit	-
进入IRF端口视图	irf-port <i>member-id/port-number</i>	-
将IRF端口和IRF物理端口绑定	port group interface <i>interface-type</i> <i>interface-number</i> [mode { enhanced normal }]	必选 缺省情况下，IRF端口没有和任何IRF物理端口绑定 S5500-52SC-HI-D 和 S5500-28SC-HI-D 交换机不支持 enhanced 参数
退回到系统视图	quit	-
进入IRF物理端口视图	interface <i>interface-type</i> <i>interface-number</i>	-
激活接口	undo shutdown	必选
退回系统视图	quit	-
保存当前配置	save	必选
激活IRF端口下的配置	irf-port-configuration active	必选



说明

- 多次执行 **port group interface**，可以将 IRF 端口与多个 IRF 物理端口绑定，以实现 IRF 链路的备份/负载分担，从而提高 IRF 链路的带宽和可靠性。
- 在 S5500-HI-D 系列交换机中，S5500-52SC-HI-D 和 S5500-28SC-HI-D 交换机最多支持将 2 个 IRF 物理端口与同一个 IRF 端口进行绑定，其它机型最多支持将 3 个 IRF 物理端口与同一个 IRF 端口进行绑定。
- 将 IRF 物理端口与 IRF 端口进行绑定或解除绑定前，必须先将涉及到的 IRF 物理端口手工关闭（即在端口上执行 **shutdown** 命令）；执行添加或者删除操作后，再将该 IRF 物理端口手工激活（即在端口上执行 **undo shutdown** 命令）。
- 在对 IRF 物理端口进行 **shutdown** 操作时，需要首先在 Master 设备或是距离 Master 设备较近（跳数较少）的设备上对 IRF 物理端口进行操作。
- 如果用户在将 IRF 端口与 IRF 物理端口绑定时使用 **mode** 参数配置了绑定模式，则 IRF 链路两端的 IRF 端口绑定模式需要配置为一致。
- 如果在 IRF 建立后，用户需要拔出 IRF 物理端口所在的接口模块扩展卡，请先拔掉用于 IRF 连接的线缆，或者先在 IRF 物理端口视图下执行 **shutdown** 命令关闭该端口后，再进行拔出接口模块扩展卡的操作。
- 以太网端口作为 IRF 物理端口与 IRF 端口绑定后，只支持 **bfd**、**default**、**shutdown**、**description** 和 **flow-interval** 命令。有关以上几条命令的详细介绍，请参见“二层技术-以太网交换命令参考”中的“以太网端口”。

1.5.4 配置成员优先级

在 Master 选举过程中，优先级数值大的成员设备将优先被选举成为 Master 设备。

表1-6 配置成员优先级

操作	命令	说明
进入系统视图	system-view	-
配置IRF中指定成员设备的优先级	irf member member-id priority priority	可选 缺省情况下，设备的成员优先级均为1



说明

优先级设置后立即生效，无需重启设备。

1.5.5 配置成员设备的描述信息

当网络中存在多个 IRF 或者同一 IRF 中存在多台成员设备且物理位置比较分散（比如在不同楼层甚至不同建筑）时，为了确认成员设备的物理位置，在组建 IRF 时可以将物理位置设置为成员设备的描述信息，以便后期维护。

表1-7 配置成员设备的描述信息

操作	命令	说明
进入系统视图	system-view	-
配置IRF中指定成员设备的描述信息	irf member member-id description text	可选 缺省情况下，成员设备没有描述信息

1.5.6 配置IRF链路的负载分担类型

当 IRF 端口与多个 IRF 物理端口（大于等于两个）绑定时，IRF 设备之间就会存在多条 IRF 链路。通过改变 IRF 链路负载分担的类型，可以灵活地实现成员设备间流量的负载分担。用户可以指定系统按照报文携带的源/目的 IP 地址、源/目的 MAC 地址、或它们之间的组合来选择所采用的负载分担类型。

用户可以通过全局配置（系统视图下）和端口下（IRF 端口视图下）的配置方式设置 IRF 链路的负载分担模式：

- 在系统视图下执行该命令，则该配置对所有 IRF 端口生效；
- 在 IRF 端口视图下执行该命令，则该配置只对当前 IRF 端口下的 IRF 链路生效；
- IRF 端口会优先采用端口下的配置。如果端口下没有配置，则采用全局配置。



说明

- 在同一视图下多次配置 **irf-port load-sharing mode** 命令，以最新的配置为准。
- 对于设备不支持的负载分担模式，系统将提示用户不支持。
- 在配置负载分担模式前，请先将 IRF 端口和 IRF 物理端口绑定。否则，负载分担模式将配置失败。

1. 全局配置IRF链路的负载分担类型

表1-8 全局配置 IRF 链路的负载分担类型

操作	命令	说明
进入系统视图	system-view	-
配置IRF链路的负载分担模式	irf-port load-sharing mode { destination-ip destination-mac source-ip source-mac } *	可选 缺省情况下，本系列交换机在处理报文时通过报文类型来进行负载分担

2. 端口下配置IRF链路的负载分担类型

表1-9 端口下配置 IRF 链路的负载分担类型

操作	命令	说明
进入系统视图	system-view	-
进入IRF端口视图	irf-port member-id/port-number	-
配置IRF链路的负载分担模式	irf-port load-sharing mode { destination-ip destination-mac source-ip source-mac }*	可选 缺省情况下,本系列交换机在处理报文时通过报文类型来进行负载分担

1.5.7 配置IRF的桥MAC保留时间

桥 MAC 是设备作为网桥与外界通信时使用的 MAC 地址。一些二层协议（例如 LACP）会使用桥 MAC 标识不同设备，所以网络上的桥设备必须具有唯一的桥 MAC。如果网络中存在两台桥 MAC 相同的设备，则会引起桥 MAC 冲突，从而导致通信故障。

IRF 作为一台虚拟设备与外界通信，也具有唯一的桥 MAC，称为 IRF 桥 MAC。通常情况下使用 Master 设备的桥 MAC 作为 IRF 桥 MAC。

因为桥 MAC 冲突会引起通信故障，桥 MAC 的切换又会导致流量中断。因此，用户需要根据网络实际情况配置 IRF 桥 MAC 的保留时间：

- 配置 IRF 桥 MAC 地址保留时间为 6 分钟。即当 Master 离开 IRF 时，IRF 桥 MAC 地址 6 分钟内保持不变；如果 6 分钟后 Master 没有回到 IRF，则使用新选举的 Master 的桥 MAC 作为 IRF 桥 MAC。该配置适用于 Master 设备短时间内离开又回到 IRF 的情况（比如 Master 重启或者链路临时故障等），可以减少不必要的桥 MAC 切换导致的流量中断。
- 如果配置了 IRF 桥 MAC 地址保留时间为永久，则不管 Master 设备是否离开 IRF，IRF 桥 MAC 始终保持不变。
- 如果配置了 IRF 桥 MAC 地址不保留，则当 Master 设备离开 IRF 时，系统立即会使用新选举的 Master 设备的桥 MAC 做 IRF 桥 MAC。

表1-10 配置 IRF 的桥 MAC 保留时间

操作	命令	说明
进入系统视图	system-view	-
配置当Master设备离开IRF时，IRF的桥MAC地址会永久保留	irf mac-address persistent always	可选 缺省情况下，IRF的桥MAC保留时间为6分钟
配置当Master设备离开IRF时，IRF的桥MAC地址的保留时间为6分钟	irf mac-address persistent timer	
配置当Master设备离开IRF时，IRF的桥MAC地址不保留，会立即变化	undo irf mac-address persistent	



注意

- 桥 MAC 变化可能导致流量短时间中断。
 - 如果两个 IRF 的桥 MAC 相同，则它们不能合并为一个 IRF。
 - 当使用 ARP MAD + 生成树的组网时，需要将 IRF 配置为 MAC 地址立即改变，即配置 **undo irf mac-address persistent** 命令。
-

1.5.8 使能IRF系统启动文件的自动加载功能

- 如果没有使能自动加载功能，当参与 IRF 的设备软件版本与 Master 设备的不一致时，则新加入或者优先级低的设备不能正常启动。此时需要用户手工升级设备的软件版本后，再将设备加入 IRF。
- 使能自动加载功能后，成员设备加入 IRF 时，会与 Master 设备的软件版本号进行比较，如果不一致，则自动从 Master 设备下载启动文件，然后使用新的系统启动文件重启，重新加入 IRF。如果新下载的启动文件的文件名与设备上原有启动文件的文件名重名，则原有启动文件会被覆盖。

表1-11 使能 IRF 系统启动文件的自动加载功能

操作	命令	说明
进入系统视图	system-view	-
使能IRF系统启动文件的自动加载功能	irf auto-update enable	可选 缺省情况下，IRF系统启动文件的自动加载功能处于使能状态



注意

- Slave 设备自动加载 Master 的启动文件后，会将该文件设置为 Slave 设备的下次启动文件，并使用该文件重启本设备。
 - 为了能够自动加载成功，请确保 Slave 设备存储介质上有足够的空闲空间用于存放新的启动文件。
-



注意

如果待升级的软件版本要求您升级 Bootrom（请参见版本说明书中的要求），请按如下步骤操作：

- 1、请将待升级的软件版本文件上传至 Master 设备，使用 **bootrom update** 命令升级 Master 设备的 Bootrom。
- 2、使用 **boot-loader** 命令指定待升级的软件版本文件为下次启动文件，并通过使用 **slot** 参数使该更改应用于 Master 设备。
- 3、重启 IRF 系统，完成软件升级。

1.5.9 配置IRF链路down延迟上报功能

配置 IRF 链路 down 延迟上报功能后，

- 如果 IRF 链路状态从 up 变为 down，端口不会立即向系统报告链路状态变化。经过配置的时间间隔后，如果 IRF 链路仍然处于 down 状态，端口才向系统报告链路状态的变化，系统再作出相应的处理；
- 如果 IRF 链路状态从 down 变为 up，链路层会立即向系统报告。

该功能用于避免因端口链路层状态在短时间内频繁改变，导致 IRF 分裂/合并的频繁发生。

表1-12 配置 IRF 链路 down 延迟上报功能

操作	命令	说明
进入系统视图	system-view	-
配置IRF链路down延迟上报时间	irf link-delay interval	可选 缺省情况下，IRF链路down延迟上报时间为4秒



注意

在对主备倒换速度和 IRF 链路切换速度要求较高，或部署了 BFD、GR 功能的环境中，建议将 IRF 链路 down 延迟上报时间配置为 0。

1.6 MAD配置

本系列交换机支持的 MAD 检测方式有：LACP MAD 检测、BFD MAD 检测和 ARP MAD 检测。三种检测方式虽然原理不同但是功能效果相同，能够满足不同组网需求：

- LACP MAD 检测用于基于 LACP 的组网检测需求；
- BFD MAD 检测用于基于 BFD 的组网检测需求；
- ARP MAD 检测用于基于非聚合场合的 Resilient ARP 的组网检测需求。

这三种方式独立工作，彼此之间互不干扰。因此，同一 IRF 内可以配置多种 MAD 检测方式。

1.6.1 LACP MAD检测

(1) LACP MAD 检测原理

LACP MAD 检测是通过扩展 LACP 协议报文内容实现的，即在 LACP 协议报文的扩展字段内定义一个新的 TLV（Type/Length/Value，类型/长度/值）数据域——用于交互 IRF 的 DomainID（域编号）和 ActiveID。当网络中同时存在多个 IRF 时（比如 IRF 级联的组网情况），DomainID 用于区别不同的 IRF。当某个 IRF 分裂时，ActiveID 用于 MAD 检测，用 IRF 中 Master 设备的成员编号来表示。

使能 LACP MAD 检测后，成员设备通过 LACP 协议报文和其它成员设备交互 DomainID 和 ActiveID 信息。

- 当成员设备收到 LACP 协议报文后，先比较 DomainID。如果 DomainID 相同，再比较 ActiveID；如果 DomainID 不同，则认为报文来自不同 IRF，不再进行 MAD 处理。
- 如果 ActiveID 相同，则表示 IRF 正常运行，没有发生多 Active 冲突；如果 ActiveID 值不同，则表示 IRF 分裂，检测到多 Active 冲突。

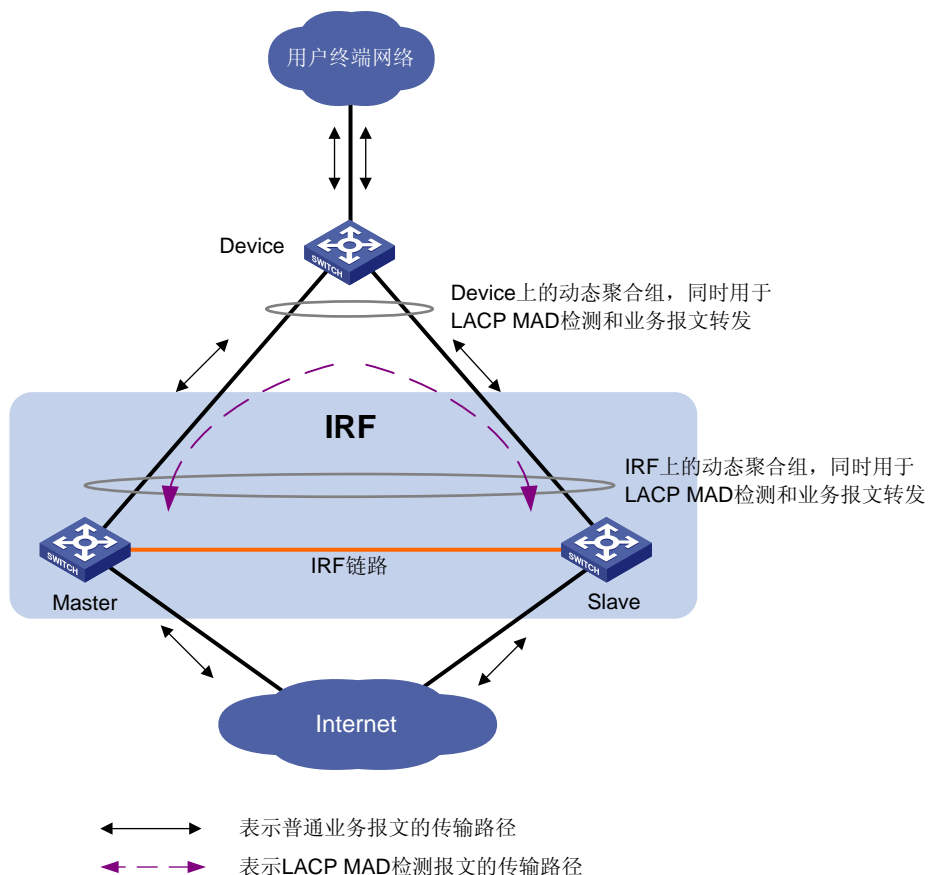
(2) LACP MAD 检测组网要求

LACP MAD检测方式组网中需要使用中间设备，支持LACP协议扩展功能的H3C设备都能作为中间设备（H3C设备是否支持LACP协议扩展功能请参见该设备操作手册中“LACP协议”部分的相关描述）。通常采用如 [图 1-8](#) 所示的组网：成员设备之间通过Device交互LACP扩展报文。



在 LACP MAD 检测组网中，如果中间设备本身也是一个 IRF 系统，则必须通过配置确保其 IRF 域编号与被检测的 IRF 系统不同，否则可能造成检测异常，甚至导致业务中断。

图1-8 LACP MAD 检测组网示意图



(3) 配置 LACP MAD 检测

LACP MAD 检测的配置步骤为：

- 配置 IRF 域编号
- 创建二层聚合端口/三层聚合接口；（中间设备上也需要进行该项配置）
- 将聚合端口的工作模式配置为动态聚合模式；（中间设备上也需要进行该项配置）
- 在动态聚合端口下使能 LACP MAD 检测功能；
- 给聚合组添加成员端口。（中间设备上也需要进行该项配置）

表1-13 配置 LACP MAD 检测

操作	命令	说明
进入系统视图	system-view	-
配置IRF域编号	irf domain <i>domain-id</i>	必选 缺省情况下，IRF的域编号为0
创建并进入聚合接口视图	interface bridge-aggregation <i>interface-number</i>	二者必选其一
	interface route-aggregation <i>interface-number</i>	

操作	命令	说明
配置聚合组工作在动态聚合模式下	link-aggregation mode dynamic	必选 缺省情况下，聚合组工作在静态聚合模式下
使能LACP MAD检测功能	mad enable	必选 缺省情况下，LACP MAD检测未使能 该命令可以在动态或静态聚合口下配置，但由于LACP MAD检测依赖于LACP协议，因此只在动态聚合端口下生效
退回系统视图	quit	-
进入二层以太网端口视图或三层以太网端口视图	interface <i>interface-type</i> <i>interface-number</i>	-
将以太网端口加入聚合组	port link-aggregation group <i>number</i>	必选

1.6.2 BFD MAD检测

(1) BFD MAD 检测原理

BFD MAD 检测是通过 BFD 协议来实现的。要使 BFD MAD 检测功能正常运行，除在三层接口下使能 BFD MAD 检测功能外，还需要在该接口上配置 MAD IP 地址。MAD IP 地址与普通 IP 地址不同的地方在于 MAD IP 地址与成员设备是绑定的，IRF 中的每个成员设备上都需要配置，且必须属于同一网段。

- 当 IRF 正常运行时，只有 Master 上配置的 MAD IP 地址生效，Slave 设备上配置的 MAD IP 地址不生效，BFD 会话处于 down 状态；（使用 **display bfd session** 命令查看 BFD 会话的状态。如果 Session State 显示为 Up，则表示激活状态；如果显示为 Down，则表示处于 down 状态）
- 当 IRF 分裂后会形成多个 IRF，不同 IRF 中 Master 上配置的 MAD IP 地址均会生效，BFD 会话被激活，此时会检测到多 Active 冲突。

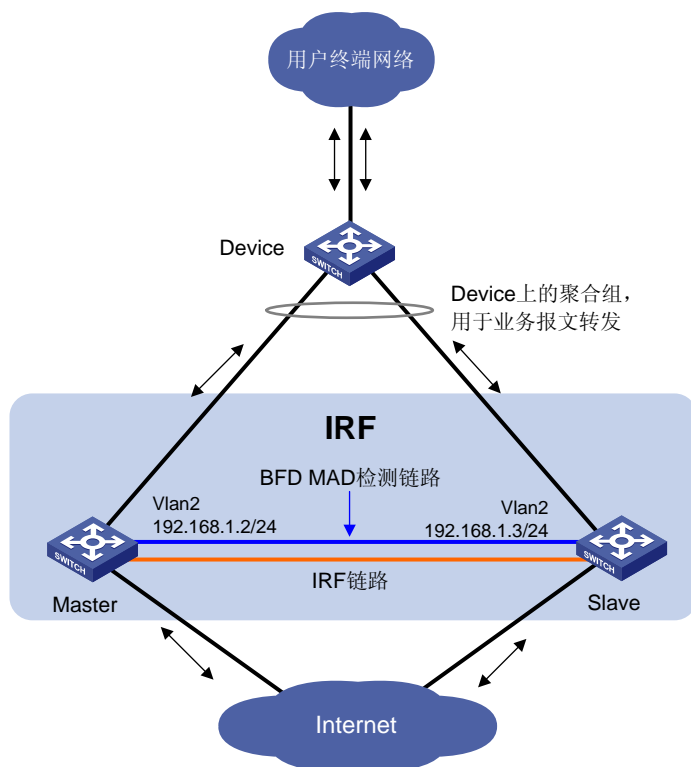
(2) BFD MAD 检测组网要求

BFD MAD检测方式可以使用中间设备来进行连接，也可以不使用中间设备。在没有中间设备的组网情况下，需要IRF中的成员设备每两台之间都要建立BFD MAD检测链路，即所有成员设备使用全连接（Full Mesh）的方式建立BFD MAD检测链路。这些链路上的端口必须属于同一VLAN，然后在该VLAN接口视图下给不同成员设备配置同一网段下的不同IP地址。常见组网如 [图 1-9](#) 所示。

 注意

- 使能 BFD MAD 检测功能的三层接口只能专用于 BFD MAD 检测，不允许运行其它业务。如果用户配置了其它业务，可能会影响该业务以及 BFD MAD 检测功能的运行。
- 用于 BFD MAD 检测的三层接口对应的 VLAN 中只能包含 BFD MAD 检测链路上的端口，请不要将其它端口加入该 VLAN。当某个业务端口需要使用 **port trunk permit vlan all** 命令允许所有 VLAN 通过时，请使用 **undo port trunk permit** 命令将用于 BFD MAD 的 VLAN 排除。

图1-9 BFD MAD 检测组网示意图



(3) 配置 BFD MAD 检测

BFD MAD 检测功能的配置顺序为：

- 创建一个新 VLAN，专用于 BFD MAD 检测；（如果用到中间设备组网，中间设备上也需要进行该项配置）
- 确定使用哪些物理端口用作 BFD MAD 检测（每台成员设备上至少一个），并将这些端口都添加到 BFD MAD 检测专用 VLAN 中；（如果用到中间设备组网，中间设备上也需要进行该项配置）
- 为 BFD MAD 检测专用 VLAN 创建 VLAN 接口，在接口下使能 BFD MAD 检测功能，并配置 MAD IP 地址。



注意

如果网络中存在多个 IRF，在配置 BFD MAD 时，各 IRF 必须使用不同的 VLAN 作为 BFD MAD 检测专用 VLAN。

表1-14 配置 BFD MAD 检测

操作		命令	说明
进入系统视图		system-view	-
创建一个新 VLAN 专用于 BFD MAD 检测		vlan <i>vlan-id</i>	必选 缺省情况下，设备上只存在 VLAN 1
退回系统视图		quit	-
进入二层以太网端口视图		interface <i>interface-type</i> <i>interface-number</i>	-
将端口加入 BFD MAD 检测专用 VLAN	Access 端口	port access vlan <i>vlan-id</i>	必选
	Trunk 端口	port trunk permit vlan <i>vlan-id</i>	请根据端口的当前链路类型选择对应的配置命令
	Hybrid 端口	port hybrid vlan <i>vlan-id</i> { tagged untagged }	BFD MAD 检测对检测端口的链路类型没有要求，不需要刻意修改端口的当前链路类型。缺省情况下，端口端的链路类型为 Access 端口
退回系统视图		quit	-
进入 VLAN 接口视图		interface vlan-interface <i>interface-number</i>	-
使能 BFD MAD 检测功能		mad bfd enable	必选 缺省情况下，没有使能 BFD MAD 检测功能
给指定成员设备配置 MAD IP 地址		mad ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } member <i>member-id</i>	必选 缺省情况下，没有为接口配置 MAD IP 地址



说明

- 用户需要将 BFD MAD 检测链路上的端口（包括 IRF 成员设备上的端口和中间设备上的端口）全部加入用于 BFD MAD 检测的 VLAN 接口对应的 VLAN。
- 使能了 BFD MAD 检测功能的 VLAN 接口以及对应 VLAN 内的端口上不支持包括 ARP 和 LACP 在内的所有的二层或三层协议应用。
- 不允许在 Vlan-interface1 接口上使能 BFD MAD 检测功能。
- BFD MAD 检测功能与 VPN 功能互斥，请不要将使能了 BFD MAD 检测功能的三层接口与 VPN 实例进行绑定。
- BFD MAD 检测功能与生成树功能互斥，在使能了 BFD MAD 检测功能的三层接口对应 VLAN 内的端口上，请不要使能生成树协议。
- 在用于 BFD MAD 检测的接口下必须使用 **mad ip address** 命令配置 MAD IP 地址，而不要配置其它 IP 地址（包括使用 **ip address** 命令配置的普通 IP 地址、VRRP 虚拟 IP 地址等），以免影响 MAD 检测功能。
- MAD IP 地址不能与各成员设备上已经配置的 IP 地址处于同一网段。
- 如果配置了 BFD MAD 检测功能的 IRF 发生分裂，则从原 IRF 分裂出去的其它 IRF 有可能因为仍保留着目的为原 Master 设备的转发表项而生成路由冲突信息（信息形式类似于%May 5 16:15:47:733 2010 H3C ARP/3/ROUTECONFLICT: Slot=5;Route conflict found, IP:192.168.2.1, VrfIndex:0），该信息对设备的转发没有任何影响，并将随着转发表项的老化而自动停止生成。

1.6.3 ARP MAD检测

(1) ARP MAD 检测原理

ARP MAD 检测是通过扩展免费 ARP 协议报文内容实现的，即使用免费 ARP 协议报文中未使用的字段来交互 IRF 的 DomainID 和 ActiveID。当网络中同时存在多个 IRF 时（比如 IRF 级联的组网情况），DomainID 用于区别不同的 IRF。当某个 IRF 分裂时，ActiveID 用于 MAD 检测，用 IRF 中 Master 设备的成员编号来表示。

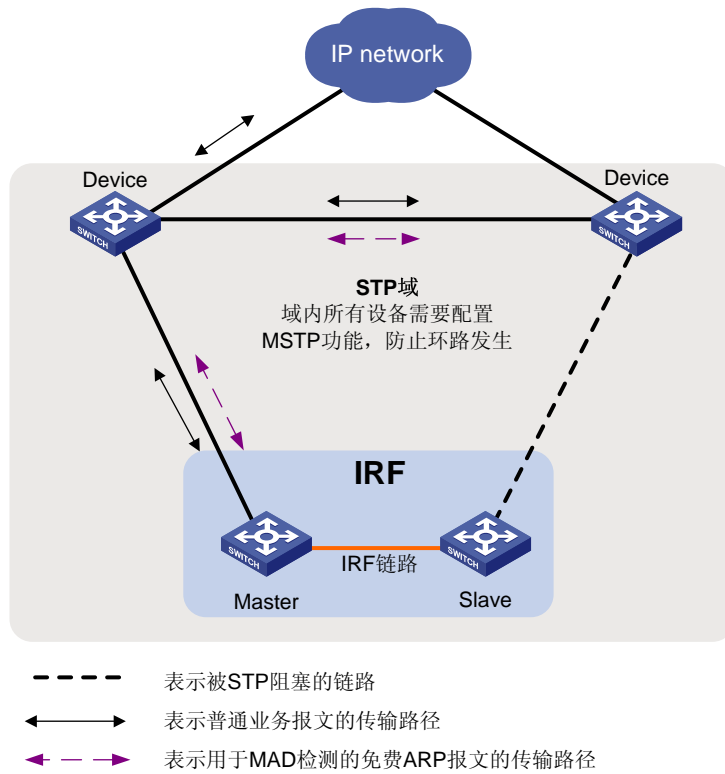
使能 ARP MAD 检测后，成员设备可以通过免费 ARP 协议报文和其它成员设备交互 DomainID 和 ActiveID 信息。

- 当成员设备收到免费 ARP 协议报文后，先比较 DomainID。如果 DomainID 相同，再比较 ActiveID；如果 DomainID 不同，则认为报文来自不同 IRF，不再进行 MAD 处理。
- 如果 ActiveID 相同，则表示 IRF 正常运行，没有发生多 Active 冲突；如果 ActiveID 值不同，则表示 IRF 分裂，检测到多 Active 冲突。

(2) ARP MAD 检测组网要求

ARP MAD检测方式可以使用中间设备来进行连接，也可以不使用中间设备。通常采用如 [图 1-10](#) 所示的组网：成员设备之间通过Device交互免费ARP报文，Device、Master和Slave上都要配置生成树功能，以防止形成环路。

图1-10 ARP MAD 检测组网示意图



(3) 配置 ARP MAD 检测

表1-15 配置 ARP MAD 检测

操作	命令	说明	
进入系统视图	system-view	-	
配置IRF域编号	irf domain <i>domain-id</i>	必选 缺省情况下，IRF的域编号为0	
创建一个新VLAN专用于ARP MAD检测	vlan <i>vlan-id</i>	必选 缺省情况下，设备上只存在VLAN 1	
退回系统视图	quit	-	
进入二层以太网端口视图	interface <i>interface-type</i> <i>interface-number</i>	-	
将端口加入 ARP MAD 检测专用VLAN	Access端口	port access vlan <i>vlan-id</i>	必选
	Trunk端口	port trunk permit vlan <i>vlan-id</i>	请根据端口的当前链路类型选择对应的配置命令
	Hybrid端口	port hybrid vlan <i>vlan-id</i> { tagged untagged }	ARP MAD检测对检测端口的链路类型没有要求，不需要刻意修改端口的当前链路类型。缺省情况下，端口端的链路类型为Access端口
退回系统视图	quit	-	

操作	命令	说明
进入VLAN接口视图	interface vlan-interface <i>interface-number</i>	-
配置IP地址	ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> }	必选 缺省情况下, 没有为接口配置IP地址
使能ARP MAD检测功能	mad arp enable	必选 缺省情况下, ARP MAD检测未使能

1.6.4 配置保留接口

IRF 系统在进行多 Active 处理的时候, 缺省情况下, 会关闭 Recovery 状态设备上的所有业务接口。如果接口有特殊用途需要保持 up 状态 (比如 Telnet 登录接口等), 则用户可以通过命令行将这些接口配置为保留接口。

表1-16 配置保留接口

操作	命令	说明
进入系统视图	system-view	-
配置保留接口, 当设备进入 Recovery 状态时, 该接口不会被关闭	mad exclude interface <i>interface-type</i> <i>interface-number</i>	必选 缺省情况下, 设备进入 Recovery 状态时会自动关闭本设备上所有的业务接口



说明

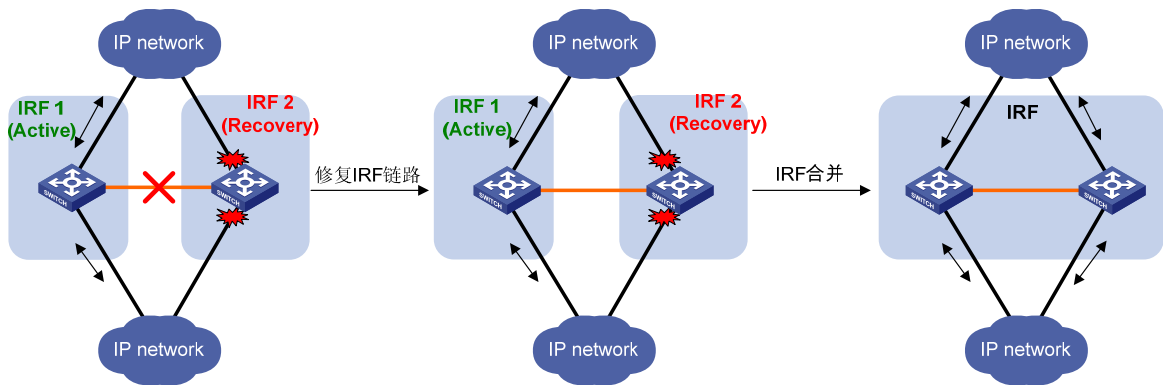
- IRF 物理端口自动作为保留接口, 不需要配置。
- 如果要求处于 Recovery 状态的 IRF 中的某个 VLAN 接口能够继续收发报文 (比如使用该 VLAN 接口进行远程登录), 则需要将该 VLAN 接口以及该 VLAN 接口对应的以太网端口都配置为保留接口。但如果在 Active 状态的 IRF 中该 VLAN 接口也处于 UP 状态, 则在网络中会产生 IP 地址冲突。

1.6.5 MAD故障恢复

IRF 链路故障将一个 IRF 分裂为两个 IRF, 从而导致多 Active 冲突。当系统检测到多 Active 冲突后, 两个冲突的 IRF 会进行竞选, Master 成员编号小的获胜, 继续正常运行, 失败的 IRF 会转入 Recovery 状态, 暂时不能转发业务报文。此时通过修复 IRF 链路可以恢复 IRF 系统 (设备会尝试自动修复 IRF 链路, 如果修复失败的话, 则需要用户手工修复)。

IRF 链路修复后, 处于 Recover 状态的 IRF 会自动重启, 从而与处于 Active 状态的 IRF 重新合并为一个 IRF, 原 Recovery 状态 IRF 中被强制关闭的业务接口会自动恢复到真实的物理状态, 如 [图 1-11](#) 所示。

图1-11 MAD 故障恢复（IRF 链路故障）



如果在IRF链路修复之前，处于Active的IRF也出现故障（原因可能是设备故障或者上下行线路故障），如 图 1-12 所示，可以在IRF 2（处于Recovery状态的IRF）上执行mad restore命令，让IRF 2恢复到正常状态（无需重启），先接替IRF 1工作，然后再修复IRF 1和IRF链路。完成修复工作后，两个IRF发生合并，此时双方将通过比较各自Master设备成员优先级的方式进行竞选，成员优先级高的Master所在IRF获胜，保持正常工作状态，竞选失败的Master所在IRF中的所有设备将自动重启并加入获胜方IRF，完成IRF合并过程，整个IRF系统恢复。

图1-12 MAD 故障恢复（IRF 链路故障+Active 状态的 IRF 故障）

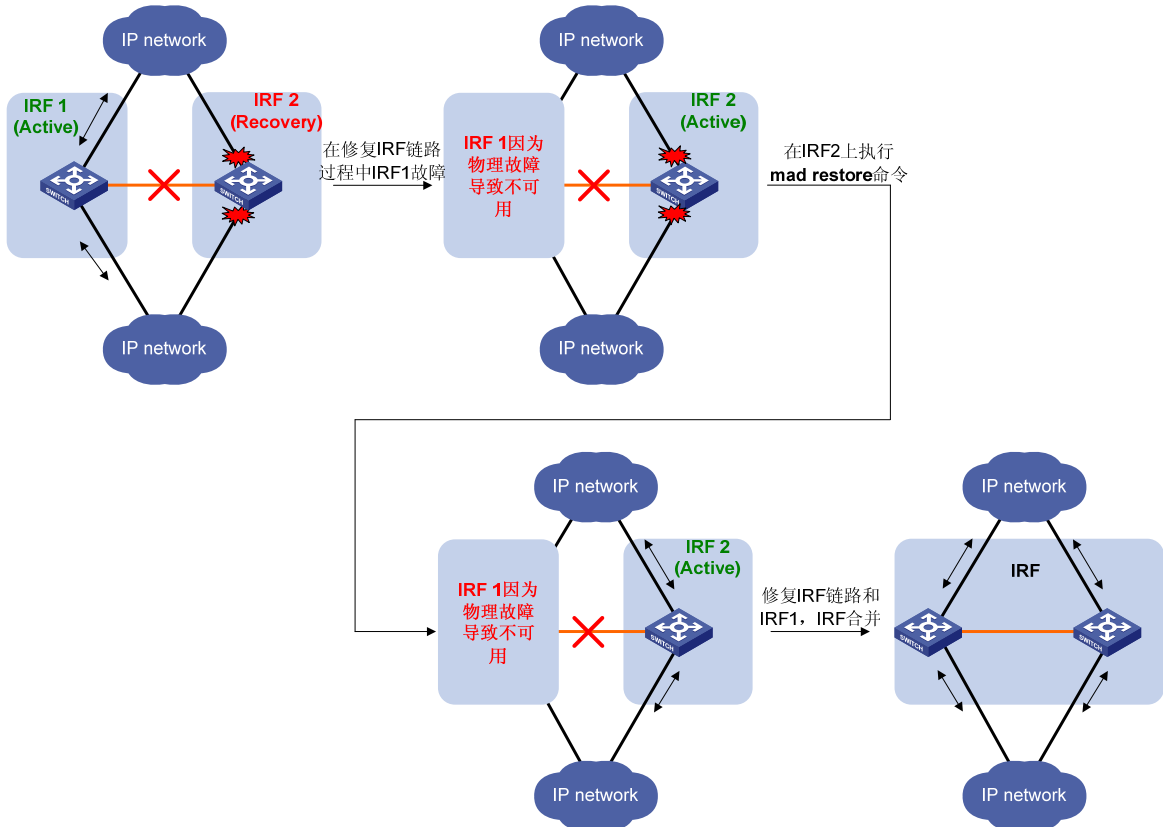


表1-17 手动恢复处于 Recovery 状态的设备

操作	命令	说明
进入系统视图	system-view	-
将IRF从Recovery状态恢复到Active状态	mad restore	必选

1.7 访问IRF

1.7.1 访问Master

IRF 的访问方式如下：

- 本地登录：通过任意成员设备 Console 口登录。
- 远程登录：给任意成员设备的任意三层接口配置 IP 地址，并且路由可达，就可以通过 Telnet、WEB、SNMP 等方式进行远程登录。

不管使用哪种方式登录 IRF，实际上登录的都是 Master。Master 是 IRF 系统的配置和控制中心，在 Master 上配置后，Master 会将相关配置同步给 Slave，以便保证 Master 和 Slave 配置的一致性。

1.7.2 访问Slave

用户访问 IRF 时，实际访问的是 IRF 中的 Master 设备，访问终端的操作界面显示的是 Master 设备的控制台。如果要打印 Slave 设备的日志、调试等信息，需要重定向到 Slave。重定向之后，用户访问终端的操作界面就会从 Master 的控制台切换到指定 Slave 的控制台，系统进入 Slave 的用户视图，“<系统名-Slave#X>”，其中“X”为成员设备编号，例如“<Sysname-Slave#2>”。用户从终端的输入指令都会转发给指定的 Slave，Master 不再进行处理。目前在 Slave 上只允许执行以下命令：

- **display**
- **quit**
- **return**
- **system-view**
- **debugging**
- **terminal debugging**
- **terminal trapping**
- **terminal logging**

用户可以使用 **quit** 命令退回到 Master 控制台，此时 Master 控制台重新激活，可以向外输出日志等信息。

表1-18 访问 Slave

操作	命令	说明
进入系统视图	system-view	-

操作	命令	说明
重定向到指定的 Slave 设备	irf switch-to member-id	必选 缺省情况下，用户访问 IRF 时，实际访问的是 Master



IRF 对同时登录设备的用户有数目限制，IRF 系统中最多允许 16 个 VTY 类型的用户同时登录，允许同时登录的 Console 用户数量与 IRF 中的成员设备数量相同。

1.8 IRF 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 IRF 的运行情况，通过查看显示信息验证配置的效果。

表1-19 IRF 显示和维护

操作	命令
显示 IRF 中所有成员设备的相关信息	display irf [{ begin exclude include } <i>regular-expression</i>]
查看 IRF 的拓扑信息	display irf topology [{ begin exclude include } <i>regular-expression</i>]
显示 IRF 中所有设备的配置信息	display irf configuration [{ begin exclude include } <i>regular-expression</i>]
显示 IRF 链路的负载分担模式	display irf-port load-sharing mode [irf-port [<i>member-id/port-number</i>]] [{ begin exclude include } <i>regular-expression</i>]
显示 IRF 设备的主备倒换状态	display switchover state [slot member-id] [{ begin exclude include } <i>regular-expression</i>]
显示 MAD 配置信息	display mad [verbose] [{ begin exclude include } <i>regular-expression</i>]

1.9 IRF 典型配置举例

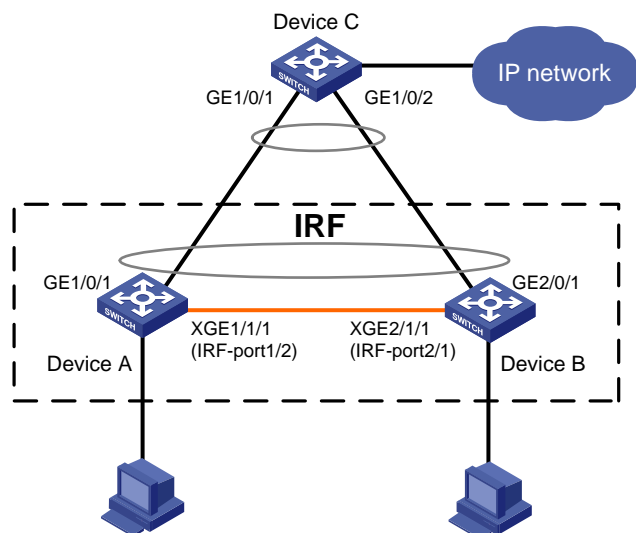
1.9.1 IRF 典型配置举例（LACP MAD 检测方式）

1. 组网需求

由于公司人员激增，接入层交换机提供的端口数目已经不能满足 PC 的接入需求。现需要在保护现有投资的基础上扩展端口接入数量，并要求网络易管理、易维护。

2. 组网图

图1-13 IRF 典型配置组网图（LACP MAD 检测方式）



3. 配置思路

- Device A 提供的接入端口数目已经不能满足网络需求，需要另外增加一台设备 Device B。（本文以两台设备组成 IRF 为例，在实际组网中可以根据需要，将多台设备组成 IRF，配置思路和配置步骤与本例类似）
- 鉴于第二代智能弹性架构 IRF 技术具有管理简便、网络扩展能力强、可靠性高等优点，所以本例使用 IRF 技术构建接入层（即在 Device A 和 Device B 上配置 IRF 功能）。
- 为了防止万一 IRF 链路故障导致 IRF 分裂、网络中存在两个配置冲突的 IRF，需要启用 MAD 检测功能。因为接入层设备较多，我们采用 LACP MAD 检测。

4. 配置步骤



说明

为便于区分，下文配置中假设 IRF 形成前 Device A 的系统名称为 DeviceA，Device B 的系统名称为 Device B；中间设备 Device C 的系统名称为 DeviceC。

(1) 配置设备编号

Device A 保留缺省编号为 1，不需要进行配置。

在 Device B 上将设备的成员编号修改为 2。

```
<DeviceB> system-view
[DeviceB] irf member 1 renumber 2
Warning: Renumbering the switch number may result in configuration change or loss. Continue?
[Y/N]:y
[DeviceB]
```

(2) 将两台设备断电后，按图 1-13 所示连接 IRF 链路，然后将两台设备上电。

在 Device A 上创建设备的 IRF 端口 2，与物理端口 Ten-GigabitEthernet1/1/1 绑定，并保存配置。

```

<DeviceA> system-view
[DeviceA] interface ten-gigabitethernet 1/1/1
[DeviceA-Ten-GigabitEthernet1/1/1] shutdown
[DeviceA-Ten-GigabitEthernet1/1/1] quit
[DeviceA] irf-port 1/2
[DeviceA-irf-port1/2] port group interface ten-gigabitethernet 1/1/1
[DeviceA-irf-port1/2] quit
[DeviceA] interface ten-gigabitethernet 1/1/1
[DeviceA-Ten-GigabitEthernet1/1/1] undo shutdown
[DeviceA-Ten-GigabitEthernet1/1/1] save

```

在 Device B 上创建设备的 IRF 端口 1，与物理端口 Ten-GigabitEthernet2/1/1 绑定，并保存配置。

```

<DeviceB> system-view
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] shutdown
[DeviceB-Ten-GigabitEthernet2/1/1] quit
[DeviceB] irf-port 2/1
[DeviceB-irf-port2/1] port group interface ten-gigabitethernet 2/1/1
[DeviceB-irf-port2/1] quit
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] undo shutdown
[DeviceB-Ten-GigabitEthernet2/1/1] save

```

激活 DeviceA 的 IRF 端口配置。

```

[DeviceA-Ten-GigabitEthernet1/1/1] quit
[DeviceA] irf-port-configuration active

```

激活 DeviceB 的 IRF 端口配置。

```

[DeviceB-Ten-GigabitEthernet2/1/1] quit
[DeviceB] irf-port-configuration active

```

(3) 两台设备间会进行 Master 竞选，竞选失败的一方将自动重启，重启完成后，IRF 形成，系统名称统一为 DeviceA。

(4) 配置 LACP MAD 检测

创建一个动态聚合端口，并使能 LACP MAD 检测功能，由于并不是在两个 IRF 之间配置 LACP MAD 检测，因此在系统提示输入 IRF 域 ID 时，可以保持为默认值 0。

```

<DeviceA> system-view
[DeviceA] interface bridge-aggregation 2
[DeviceA-Bridge-Aggregation2] link-aggregation mode dynamic
[DeviceA-Bridge-Aggregation2] mad enable
You need to assign a domain ID (range: 0-4294967295)
[Current domain is: 0]:
The assigned domain ID is: 0
Info: MAD LACP only enable on dynamic aggregation interface.
[DeviceA-Bridge-Aggregation2] quit

```

在聚合端口中添加成员端口 GigabitEthernet1/0/1 和 GigabitEthernet2/0/1，专用于两台 IRF 成员设备与中间设备进行 LACP MAD 检测。

```

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

```

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

(5) 中间设备 Device C 的配置

Device C 作为一台中间设备需要支持 LACP 功能，用来转发、处理 LACP 协议报文，协助 Device A 和 Device B 进行多 Active 检测。从节约成本的角度考虑，使用一台支持 LACP 功能的交换机即可。

创建一个动态聚合端口。

```
<DeviceC> system-view
[DeviceC] interface bridge-aggregation 2
[DeviceC-Bridge-Aggregation2] link-aggregation mode dynamic
[DeviceC-Bridge-Aggregation2] quit
```

在聚合端口中添加成员端口 GigabitEthernet1/0/1 和 GigabitEthernet1/0/2，用于进行 LACP MAD 检测。

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

按 [图 1-13](#) 所示连接 LACP MAD 链路。

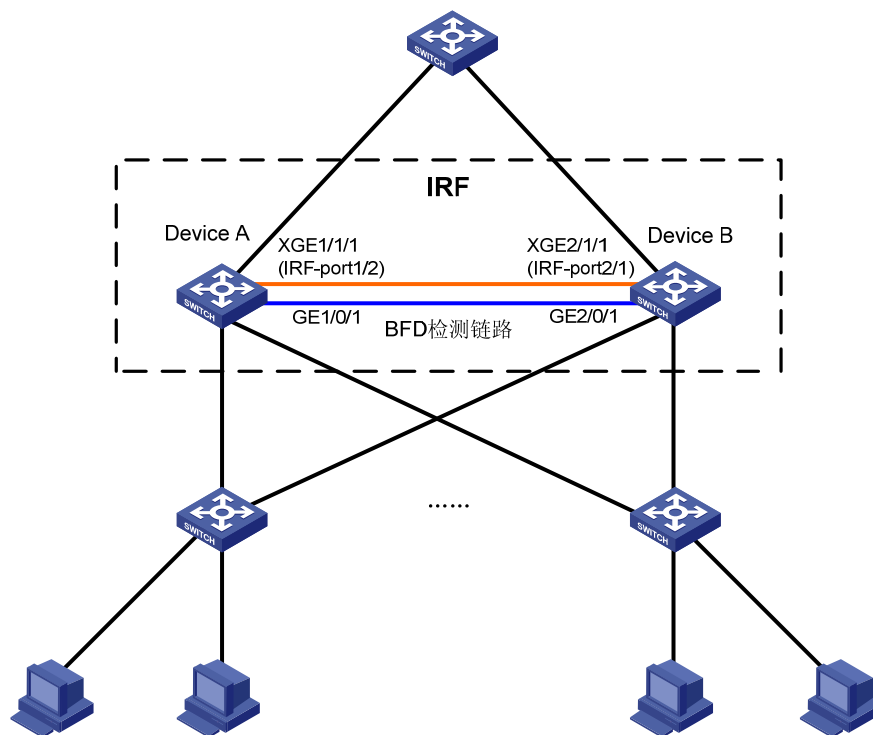
1.9.2 IRF 典型配置举例（BFD MAD 检测方式）

1. 组网需求

由于网络规模迅速扩大，当前中心交换机（Device A）转发能力已经不能满足需求，现需要在保护现有投资的基础上将网络转发能力提高一倍，并要求网络易管理、易维护。

2. 组网图

图1-14 IRF 典型配置组网图（BFD MAD 检测方式）



3. 配置思路

- Device A 处于局域网的汇聚层，为了将汇聚层的转发能力提高一倍，需要另外增加一台设备 Device B。
- 鉴于第二代智能弹性架构 IRF 技术具有管理简便、网络扩展能力强、可靠性高等优点，所以本例使用 IRF 技术构建网络汇聚层（即在 Device A 和 Device B 上配置 IRF 功能），接入层设备通过聚合双链路上行。
- 为了防止万一 IRF 链路故障导致 IRF 分裂、网络中存在两个配置冲突的 IRF，需要启用 MAD 检测功能。因为成员设备比较少，我们采用 BFD MAD 检测方式来监测 IRF 的状态。

4. 配置步骤



说明

为便于区分，下文配置中假设 IRF 形成前 Device A 的系统名称为 DeviceA，Device B 的系统名称为 Device B。

(1) 配置设备编号

Device A 保留缺省编号为 1，不需要进行配置。

在 Device B 上将设备的成员编号修改为 2。

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

```
[DeviceB] irf member 1 renumber 2
```

```
Warning: Renumbering the switch number may result in configuration change or loss. Continue?
```

```
[Y/N]:y
```

[DeviceB]

(2) 将两台设备断电后,按 [图 1-14](#) 所示连接IRF链路,然后将两台设备上电。

在 Device A 上创建设备的 IRF 端口 2,与物理端口 Ten-GigabitEthernet1/1/1 绑定,并保存配置。

```
<DeviceA> system-view
[DeviceA] interface ten-gigabitethernet 1/1/1
[DeviceA-Ten-GigabitEthernet1/1/1] shutdown
[DeviceA-Ten-GigabitEthernet1/1/1] quit
[DeviceA] irf-port 1/2
[DeviceA-irf-port1/2] port group interface ten-gigabitethernet 1/1/1
[DeviceA-irf-port1/2] quit
[DeviceA] interface ten-gigabitethernet 1/1/1
[DeviceA-Ten-GigabitEthernet1/1/1] undo shutdown
[DeviceA-Ten-GigabitEthernet1/1/1] save
```

在 Device B 上创建设备的 IRF 端口 1,与物理端口 Ten-GigabitEthernet2/1/1 绑定,并保存配置。

```
<DeviceB> system-view
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] shutdown
[DeviceB-Ten-GigabitEthernet2/1/1] quit
[DeviceB] irf-port 2/1
[DeviceB-irf-port2/1] port group interface ten-gigabitethernet 2/1/1
[DeviceB-irf-port2/1] quit
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] undo shutdown
[DeviceB-Ten-GigabitEthernet2/1/1] save
```

激活 DeviceA 的 IRF 端口配置。

```
[DeviceA-Ten-GigabitEthernet1/1/1] quit
[DeviceA] irf-port-configuration active
```

激活 DeviceB 的 IRF 端口配置。

```
[DeviceB-Ten-GigabitEthernet2/1/1] quit
[DeviceB] irf-port-configuration active
```

(3) 两台设备间将会进行 Master 竞选,竞选失败的一方将自动重启,重启完成后,IRF 形成,系统名称统一为 DeviceA。

(4) 配置 BFD MAD 检测

创建 VLAN 3,并将 Device A 上的端口 GigabitEthernet1/0/1 和 Device B 上的端口 GigabitEthernet2/0/1 加入 VLAN 中。

```
<DeviceA> system-view
[DeviceA] vlan 3
[DeviceA-vlan3] port gigabitethernet 1/0/1 gigabitethernet 2/0/1
[DeviceA-vlan3] quit
```

创建 VLAN 接口 3,并配置 MAD IP 地址。

```
[DeviceA] interface vlan-interface 3
[DeviceA-Vlan-interface3] mad bfd enable
[DeviceA-Vlan-interface3] mad ip address 192.168.2.1 24 member 1
[DeviceA-Vlan-interface3] mad ip address 192.168.2.2 24 member 2
[DeviceA-Vlan-interface3] quit
```

按 [图 1-14](#) 所示连接BFD MAD链路。

因为 BFD MAD 和生成树功能互斥，所以在 GigabitEthernet1/0/1 和 GigabitEthernet2/0/1 上关闭生成树协议。

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-Gigabitethernet1/0/1] undo stp enable
[DeviceA-Gigabitethernet1/0/1] quit
[DeviceA] interface gigabitethernet 2/0/1
[DeviceA-Gigabitethernet2/0/1] undo stp enable
```

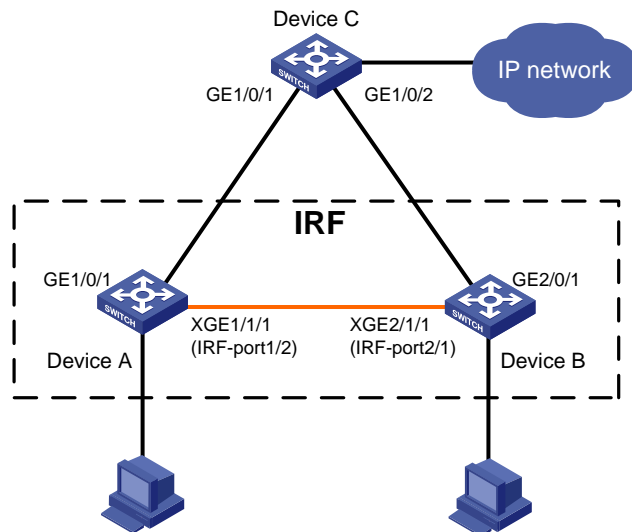
1.9.3 IRF典型配置举例（ARP MAD检测方式）

1. 组网需求

由于网络规模迅速扩大，当前中心交换机（Device A）转发能力已经不能满足需求，现需要在保护现有投资的基础上将网络转发能力提高一倍，并要求网络易管理、易维护。

2. 组网图

图1-15 IRF 典型配置组网图（ARP MAD 检测方式）



3. 配置思路

- Device A 处于局域网的汇聚层，为了将汇聚层的转发能力提高一倍，需要另外增加一台设备 Device B。
- 鉴于第二代智能弹性架构 IRF 技术具有管理简便、网络扩展能力强、可靠性高等优点，所以本例使用 IRF 技术构建网络接入层（即在 Device A 和 Device B 上配置 IRF 功能），IRF 通过双链路上行。
- 为了防止万一 IRF 链路故障导致 IRF 分裂、网络中存在两个配置冲突的 IRF，需要启用 MAD 检测功能。因为成员设备比较少，我们采用 ARP MAD 检测方式来监测 IRF 的状态，复用链路上行传递 ARP MAD 报文。为防止环路发生，在 IRF 和 Device C 上启用生成树功能。

4. 配置步骤



说明

为便于区分，下文配置中假设 IRF 形成前 Device A 的系统名称为 DeviceA，Device B 的系统名称为 Device B；中间设备 Device C 的系统名称为 DeviceC。

(1) 配置设备编号

Device A 保留缺省编号为 1，不需要进行配置。

在 Device B 上将设备的成员编号修改为 2。

```
<DeviceB> system-view
[DeviceB] irf member 1 renumber 2
Warning: Renumbering the switch number may result in configuration change or loss. Continue?
[Y/N]:y
[DeviceB]
```

(2) 将两台设备断电后，按 [图 1-14](#) 所示连接 IRF 链路，然后将两台设备上电。

在 Device A 上创建设备的 IRF 端口 2，与物理端口 Ten-GigabitEthernet1/1/1 绑定，并保存配置。

```
<DeviceA> system-view
[DeviceA] interface ten-gigabitethernet 1/1/1
[DeviceA-Ten-GigabitEthernet1/1/1] shutdown
[DeviceA-Ten-GigabitEthernet1/1/1] quit
[DeviceA] irf-port 1/2
[DeviceA-irf-port1/2] port group interface ten-gigabitethernet 1/1/1
[DeviceA-irf-port1/2] quit
[DeviceA] interface ten-gigabitethernet 1/1/1
[DeviceA-Ten-GigabitEthernet1/1/1] undo shutdown
[DeviceA-Ten-GigabitEthernet1/1/1] save
```

在 Device B 上创建设备的 IRF 端口 1，与物理端口 Ten-GigabitEthernet2/1/1 绑定，并保存配置。

```
<DeviceB> system-view
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] shutdown
[DeviceB-Ten-GigabitEthernet2/1/1] quit
[DeviceB] irf-port 2/1
[DeviceB-irf-port2/1] port group interface ten-gigabitethernet 2/1/1
[DeviceB-irf-port2/1] quit
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] undo shutdown
[DeviceB-Ten-GigabitEthernet2/1/1] save
```

激活 DeviceA 的 IRF 端口配置。

```
[DeviceA-Ten-GigabitEthernet1/1/1] quit
[DeviceA] irf-port-configuration active
```

激活 DeviceB 的 IRF 端口配置。

```
[DeviceB-Ten-GigabitEthernet2/1/1] quit
[DeviceB] irf-port-configuration active
```

(3) 两台设备间将会进行 Master 竞选，竞选失败的一方将自动重启，重启完成后，IRF 形成，系统名称统一为 DeviceA。

(4) 配置 ARP MAD

在 IRF 上全局使能 MSTP，以防止环路的发生。

```
<DeviceA> system-view
[DeviceA] stp enable
```

按 [图 1-15](#) 所示连接 ARP MAD 检测链路。

将 IRF 配置为 MAC 地址立即改变。

```
[DeviceA] undo irf mac-address persistent
```

创建 VLAN 3，并将 Device A（成员编号为 1）上的端口 GigabitEthernet1/0/1 和 Device B（成员编号为 2）上的端口 GigabitEthernet2/0/1 加入 VLAN 中。

```
[DeviceA] vlan 3
[DeviceA-vlan3] port gigabitethernet 1/0/1 gigabitethernet 2/0/1
[DeviceA-vlan3] quit
```

创建 VLAN-interface3，并配置 IP 地址，使能 ARP MAD 检测功能，由于并不是在两个 IRF 之间配置 ARP MAD 检测，因此在系统提示输入 IRF 域 ID 时，可以保持为默认值 0。

```
[DeviceA] interface vlan-interface 3
[DeviceA-Vlan-interface3] ip address 192.168.2.1 24
[DeviceA-Vlan-interface3] mad arp enable
You need to assign a domain ID (range: 0-4294967295)
[Current domain is: 0]:
The assigned domain ID is: 0
```

(5) 配置 Device C

在全局使能 MSTP，以防止环路的发生。

```
<DeviceC> system-view
[DeviceC] stp enable
```

创建 VLAN 3，并将端口 GigabitEthernet1/0/1 和 GigabitEthernet1/0/2 加入 VLAN 3 中，用于转发 ARP MAD 报文。

```
[DeviceC] vlan 3
[DeviceC-vlan3] port gigabitethernet 1/0/1 gigabitethernet 1/0/2
[DeviceC-vlan3] quit
```