



H3C S12500-X & S12500X-AF 系列交换机



虚拟化技术配置指导

新华三技术有限公司
<http://www.h3c.com>

资料版本：6W102-20171230
产品版本：Release 1150 及以上

Copyright © 2016-2017 新华三技术有限公司及其许可者 版权所有，保留一切权利。

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

H3C、**H3C**、H3CS、H3CIE、H3CNE、Aolynk、、H³Care、、IRF、NetPilot、Netflow、SecEngine、SecPath、SecCenter、SecBlade、Comware、ITCMM、HUASAN、华三均为新华三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

本配置指导主要介绍虚拟化技术 IRF 和 MDC。通过 IRF，可以将多台设备虚拟为一台设备在网络中运行，提高管理效率的同时，还能够提供 1:N 的设备级备份，并简化网络拓扑。通过 MDC，可以将一台物理设备划分成多台逻辑设备，每台逻辑设备都拥有自己专属的软硬件资源，独立运行，独立转发。

前言部分包含如下内容：

- [读者对象](#)
- [本书约定](#)
- [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

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. 示例约定

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

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

资料获取方式

您可以通过H3C网站（<http://www.h3c.com/>）获取最新的产品资料：

- 获取安装类、配置类或维护类产品资料
http://www.h3c.com/cn/Technical_Documents
- 获取版本说明书等与软件版本配套的资料
http://www.h3c.com/cn/Software_Download

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/>

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	1-1
1.1 IRF简介	1-1
1.1.1 IRF的优点	1-1
1.1.2 IRF的应用	1-1
1.1.3 IRF基本概念	1-2
1.2 IRF工作原理	1-6
1.2.1 物理连接	1-6
1.2.2 拓扑收集	1-7
1.2.3 角色选举	1-7
1.2.4 IRF的管理与维护	1-8
1.2.5 MAD功能	1-9
1.3 配置限制和指导	1-10
1.3.1 组建IRF时的注意事项	1-10
1.3.2 选择IRF物理端口时的注意事项	1-11
1.3.3 IRF形成后的配置限制和指导	1-11
1.4 IRF配置任务简介	1-11
1.5 独立运行模式下预配置IRF	1-13
1.5.1 配置成员编号	1-13
1.5.2 配置成员优先级	1-13
1.5.3 配置IRF端口	1-13
1.6 将当前配置保存到设备的下次启动配置文件	1-14
1.7 切换到IRF模式	1-14
1.8 访问IRF	1-15
1.9 IRF模式下配置IRF	1-15
1.9.1 配置成员编号	1-15
1.9.2 配置成员优先级	1-16
1.9.3 配置IRF端口	1-16
1.9.4 开启IRF合并自动重启功能	1-17
1.9.5 配置成员设备的描述信息	1-18
1.9.6 配置IRF的桥MAC地址	1-18
1.9.7 配置IRF的桥MAC保留时间	1-19
1.9.8 开启启动文件的自动加载功能	1-19
1.9.9 配置IRF链路down延迟上报功能	1-20

1.9.10 MAD配置	1-21
1.10 IRF显示和维护	1-28
1.11 IRF典型配置举例	1-29
1.11.1 IRF典型配置举例（LACP MAD检测方式）	1-29
1.11.2 IRF典型配置举例（BFD MAD检测方式）	1-31
1.11.3 将成员设备从IRF模式恢复到独立运行模式配置举例	1-34

1 IRF

1.1 IRF简介

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连接在一起, 进行必要的配置后, 虚拟化成一台设备。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。

1.1.1 IRF的优点

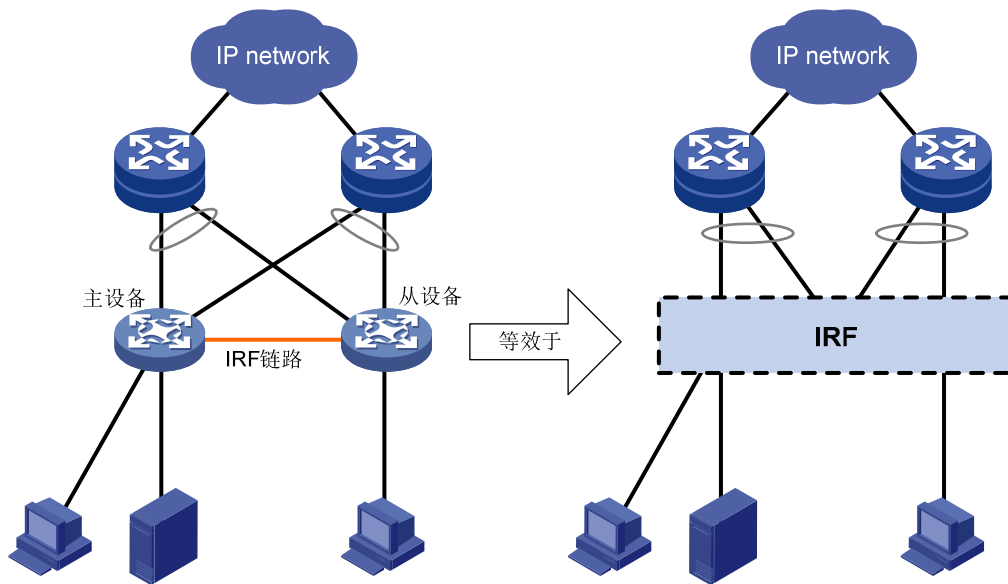
IRF 主要具有以下优点:

- 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, 主设备负责 IRF 的运行、管理和维护, 从设备在作为备份的同时也可以处理业务。一旦主设备故障, 系统会迅速自动选举新的主设备, 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 还能轻松自如的扩展处理能力。

1.1.2 IRF的应用

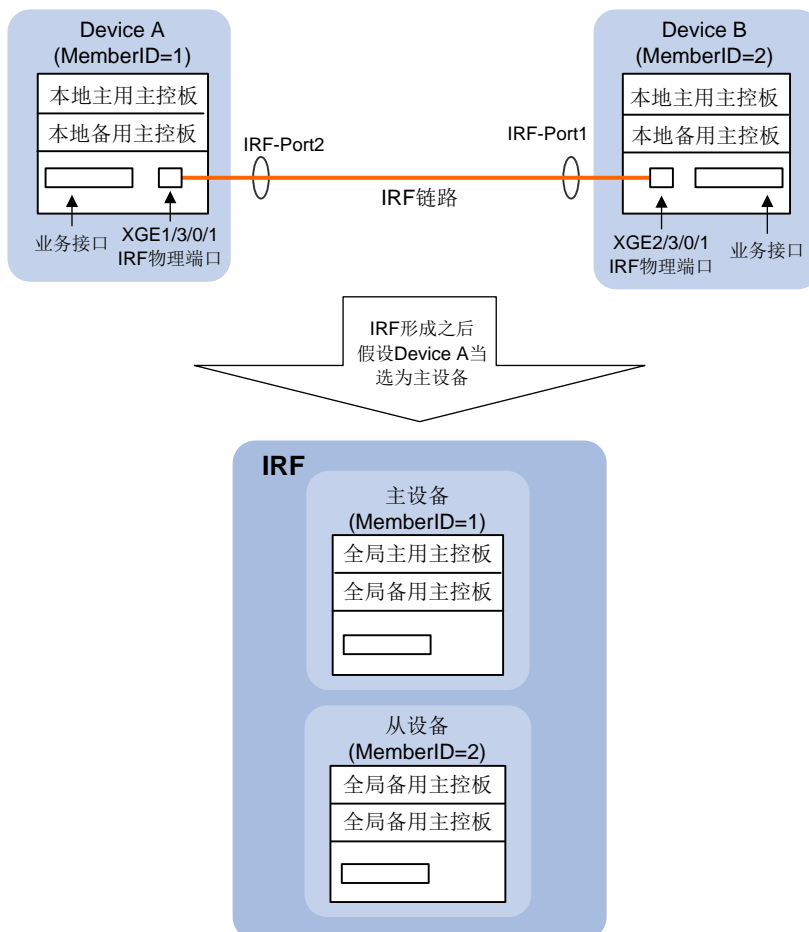
如 [图 1-1](#) 所示, 主设备和从设备组成 IRF, 对上、下层设备来说, 它们就是一台设备——IRF。

图1-1 IRF 组网应用示意图



1.1.3 IRF基本概念

图1-2 IRF 虚拟化示意图



如 [图 1-2](#) 所示，将 Device A 和 Device B 物理连线，进行必要的配置后，就能形成 IRF。IRF 拥有四块主控板（一块主用主控板，三块备用主控板），两块接口板。IRF 统一管理 Device A 和 Device B 的物理资源和软件资源。



注意

在 IRF 环境中，从设备上的主控板不能被全部拔出，每台设备上都必须至少存在一块主控板才能正常运行。

IRF 虚拟化技术涉及如下基本概念：

1. 运行模式

设备支持两种运行模式：

- 独立运行模式：处于该模式下的设备只能单机运行，不能与别的设备形成 IRF。
- IRF 模式：处于该模式下的设备可以与其它设备互连形成 IRF。

两种模式之间通过命令行进行切换。

2. 成员设备的角色

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：

- 主用设备（简称为主设备）：负责管理整个 IRF。
- 从属设备（简称为从设备）：作为主设备的备份设备运行。当主设备故障时，系统会自动从从设备中选举一个新的主设备接替原主设备工作。

主设备和从设备均由角色选举产生。一个 IRF 中同时只能存在一台主设备，其它成员设备都是从设备。关于设备角色选举过程的详细介绍请参见“[1.2.3 角色选举](#)”。

3. 主控板的角色

设备加入 IRF 后，设备上的主控板就具有两重身份（身份不同责任不同）：

- 本地身份：负责管理本设备的事宜，比如主用主控板和备用主控板间的同步、协议报文的处理、路由表项的生成维护等。
- 全局身份：负责处理 IRF 相关事宜，比如角色选举、拓扑收集等。

表1-1 主控板的角色

主控板角色	描述
本地主控板	成员设备的主控板，负责管理本台设备，是成员设备的必备硬件
全局主用主控板	IRF的主用主控板，负责管理整个IRF，就是主设备的本地主控板
全局备用主控板	IRF的备用主控板，是全局主用主控板的备份。除了全局主用主控板，IRF中所有成员设备的主控板均为全局备用主控板

4. IRF端口

一种专用于 IRF 成员设备之间进行连接的逻辑接口，每台成员设备上可以配置两个 IRF 端口，分别为 IRF-Port1 和 IRF-Port2。它需要和物理端口绑定之后才能生效。

- 在独立运行模式下，IRF 端口采用一维编号，分为 IRF-Port1 和 IRF-Port2；
- 在 IRF 模式下，IRF 端口采用二维编号，分为 IRF-Portn/1 和 IRF-Portn/2，其中 n 为设备的成员编号。为简洁起见，本文描述时统一使用 IRF-Port1 和 IRF-Port2。

5. IRF物理端口

与 IRF 端口绑定，用于 IRF 成员设备之间进行连接的物理接口。在本系列交换机上，可以将任意 10GE、40GE、或 100GE CXP/CFP2 端口配置为 IRF 物理端口。

通常情况下，本系列交换机上的 10GE/40GE 端口以及 100GE CXP/CFP2 端口作为普通的业务端口，负责向网络中转发业务报文，将它们与 IRF 端口绑定后就作为 IRF 物理端口，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。

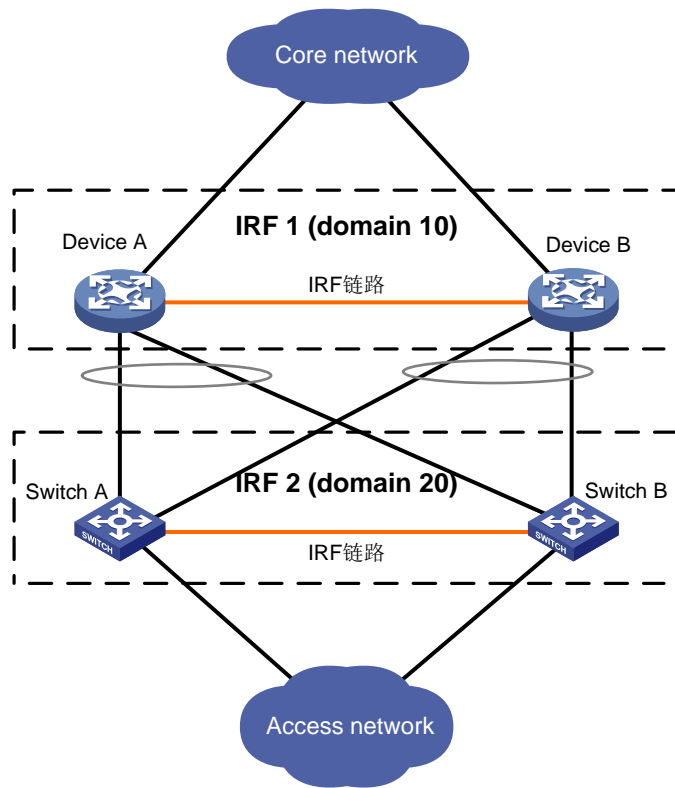
由于 IRF 物理端口上不能开启 STP 或其它环路控制协议，IRF 成员设备需要根据接收和发送报文的端口以及 IRF 的当前拓扑，来判断报文在发送后是否会产生环路。如果判断结果为会产生环路，设备将在位于环路路径上的发送端口处将报文丢弃。该方式会造成大量广播报文在 IRF 物理端口上被丢弃，此为正常现象。在使用 SNMP 工具监测设备端口的收发报文记录时，取消对 IRF 物理端口的监测，可以避免收到大量丢弃报文的告警信息。

6. IRF域

域是一个逻辑概念，一个 IRF 对应一个 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 1-3](#) 所示，Device A 和 Device B 组成 IRF 1，Switch A 和 Switch B 组成 IRF 2。如果 IRF 1 和 IRF 2 之间有 MAD 检测链路，则两个 IRF 各自的成员设备间发送的 MAD 检测报文会被另外的 IRF 接收到，从而对两个 IRF 的 MAD 检测造成影响。这种情况下，需要给两个 IRF 配置不同的域编号，以保证两个 IRF 互不干扰。

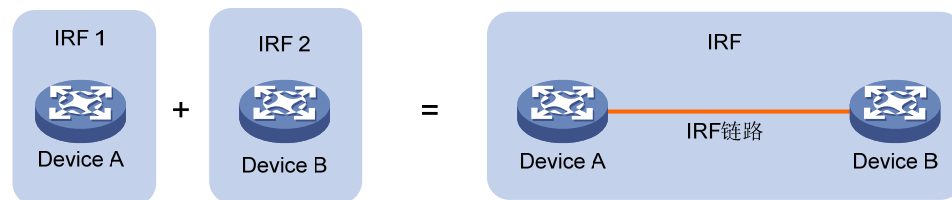
图1-3 多 IRF 域示意图



7. IRF合并

如 图 1-4 所示，两个（或多个）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。

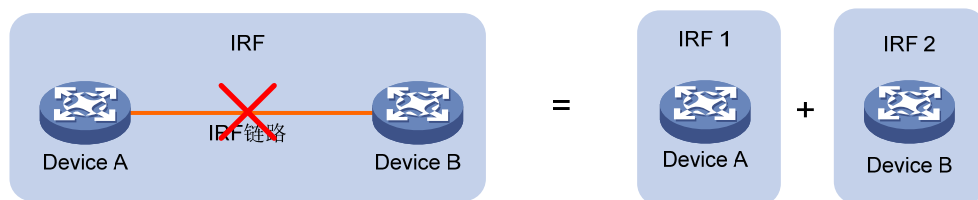
图1-4 IRF 合并示意图



8. IRF分裂

如 图 1-5 所示，一个 IRF 形成后，由于 IRF 链路故障，导致 IRF 中两相邻成员设备不连通，一个 IRF 变成两个 IRF，这个过程称为 IRF 分裂。

图1-5 IRF 分裂示意图



9. 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。优先级越高当选为主设备的可能性越大。

设备的缺省优先级均为 1，如果想让某台设备当选为主设备，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。

1.2 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的管理和维护阶段。

1.2.1 物理连接

1. 连接介质

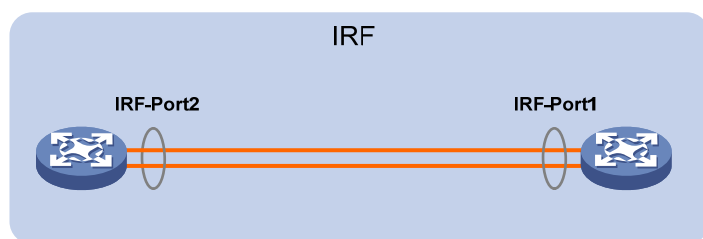
要形成一个 IRF，需要先连接成员设备的 IRF 物理端口。本系列交换机支持使用 10GE/40GE/100GE 端口作为 IRF 物理端口，使用 SFP+/QSFP+/CXP 线缆能够为成员设备间报文的传输提供很高的可靠性和性能，使用 SFP+/QSFP+/CXP/CFP2 模块和光纤可以将距离很远的物理设备连接成为一台虚拟设备，使得应用更加灵活。

2. 连接要求

本设备上与IRF-Port1 口绑定的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 1-6](#) 所示。否则，不能形成IRF。

一个 IRF 端口可以与一个或多个 IRF 物理端口绑定，以提高 IRF 链路的带宽以及可靠性。

图1-6 IRF 物理连接示意图





说明

在进行 IRF 物理连接时，必须保证成员设备间物理连接形式为直连，即 IRF 连接的物理路径上不能存在其他网络设备。

3. 连接拓扑

目前本系列设备支持两台设备实现链型拓扑的 IRF，如 [图 1-6](#) 所示。

1.2.2 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。

每个成员设备由本地主控板进行管理，在本地记录自己已知的拓扑信息。设备刚启动时，本地主控板只记录了自身的拓扑信息。当 IRF 端口状态变为 up 后，本地主控板会进行以下操作：

- (1) 将已知的拓扑信息周期性的从 up 状态的 IRF 端口发送出去；
- (2) 在收到邻居的拓扑信息后，更新本地记录的拓扑信息；
- (3) 如果成员设备上配备了备用主控板，则本地主用主控板会将自己记录的拓扑信息同步到本地备用主控板上，以便保持两块主控板上拓扑信息的一致。

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。此时会进入角色选举阶段。

1.2.3 角色选举

确定成员设备角色为主设备或从设备的过程称为角色选举。角色选举会在以下情况下进行：IRF 建立、主设备离开或者故障、IRF 合并等。角色选举规则如下：

- (1) 当前主设备优先，IRF 不会因为新的成员设备加入而重新选举主设备。不过，当 IRF 形成时，因为没有主设备，所有加入的设备都认为自己是主设备，则继续下一条规则的比较。
- (2) 成员优先级大的优先。如果优先级相同，则继续下一条规则的比较。
- (3) 系统运行时间长的优先。在 IRF 中，成员设备启动时间间隔精度为 10 分钟，即 10 分钟之内启动的设备，则认为它们是同时启动的，则继续下一条规则的比较。
- (4) 成员编号小的优先。

通过以上规则选出的最优成员设备即为主设备，其它成员设备则均为从设备。

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，每个IRF的主设备间会进行IRF竞选，竞选仍然遵循角色选举的规则，竞选失败方的所有成员设备重启后均以从设备的角色加入获胜方，最终合并为一个IRF。合并过程中的重启是设备自动完成还是需要用户手工完成与用户的配置有关，请参见[开启IRF合并自动重启功能](#)。
 - 不管设备与其它设备一起形成 IRF，还是加入已有 IRF，如果该设备被选为从设备，则该设备会使用主设备的配置重新启动，以保证和主设备上的配置一致，本设备上的配置文件还在，但不再生效，除非设备恢复到独立运行模式。
-

1.2.4 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主设备统一管理。

1. 成员编号

在运行过程中，IRF 使用成员编号来标识成员设备，以便对其进行管理。例如，IRF 中接口的编号会加入成员编号信息：当设备处于独立运行模式时，接口编号采用三维格式（如 `FortyGigE3/0/1`）；加入 IRF 后，接口编号会变为四维，第一维表示成员编号（如 `FortyGigE2/3/0/1`）。成员编号还被引入到文件系统管理中：当设备处于独立运行模式时，某文件的路径为 `slot1#flash:/test.cfg`；加入 IRF 后，该文件路径前需要添加“`chassisA#`”信息，变为 `chassis1#slot1#flash:/test.cfg`，用来表明文件位于成员设备 1 的 1 号单板上。所以，在 IRF 中必须保证所有设备成员编号的唯一性。

如果建立 IRF 时成员设备的编号不唯一（即存在编号相同的成员设备），则不能建立 IRF；如果新设备加入 IRF，但是该设备与已有成员设备的编号冲突，则该设备不能加入 IRF。因此，在建立 IRF 前，请统一规划各成员设备的编号，并逐一进行手工配置，以保证各设备成员编号的唯一性。



说明

成员设备编号和优先级的配置是以设备为单位的，配置后，先保存在本地主用主控板，再同步给本地备用主控板。如果某成员设备上本地主用主控板和本地备用主控板保存的成员编号不一致，则以本地主用主控板的配置为准。

2. IRF拓扑维护

如果某成员设备 A 故障或者 IRF 链路故障，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主设备还是从设备，如果离开的是主设备，则触发新的角色选举，再更新本地的 IRF 拓扑；如果离开的是从设备，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。

3. IRF 端口连接状态维护

当一个端口与 IRF 端口绑定后，将只能用于 IRF 成员设备之间的连接，如果将其连接到一个普通业务端口上，可能会导致 IRF 成员间流量和协议运算受到影响，甚至出现成员设备无法正常工作的现象，造成 IRF 分裂。

为了在连接错误时及时提示用户，当 IRF 物理端口与普通业务端口相连时，在 IRF 物理端口侧的设备上，将输出日志信息，形式为“The port *port* can't receive irf pkt and has been changed to inactive status, please check.”。其中 *port* 为出现连接错误的 IRF 物理端口编号。

如果您在设备上看到上述日志信息，请根据信息中提供的端口编号检查 IRF 物理连接和 IRF 端口的配置是否一致。

需要说明的是，在 IRF 处于非稳定状态时（包括但不限于：建立 IRF 的过程、IRF 重启的过程、用于 IRF 连接的单板的重启过程，以及修改用于 IRF 连接的端口时），出现此提示信息为正常现象，待 IRF 稳定工作后将不再提示该信息。

1.2.5 MAD 功能

IRF 链路故障会导致一个 IRF 变成多个新的 IRF。这些 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）或 BFD（Bidirectional Forwarding Detection，双向转发检测）来检测网络中是否存在多个 IRF。详细信息请参考“[1.9.10 MAD 配置](#)”。

(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（即正常工作状态）的 IRF。

- 对于 LACP MAD 检测，冲突处理会先比较两个 IRF 中成员设备的数量：数量多的 IRF 处于 Active 状态，继续正常工作；数量少的迁移到 Recovery 状态（即禁用状态）；如果成员数量相等，则主设备成员编号小的 IRF 处于 Active 状态，继续正常工作，其它 IRF 迁移到 Recovery 状态。
- 对于 BFD MAD 检测，冲突处理会直接让主设备成员编号小的 IRF 处于 Active 状态，继续正常工作；其它 IRF 迁移到 Recovery 状态。

IRF 迁移到 Recovery 状态后会关闭该 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，可通过 **mad exclude interface** 命令配置。

(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。

- 如果出现故障的是 Active 状态的 IRF，则在进行 MAD 故障恢复前，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。
- 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。

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”。

1.3 配置限制和指导



对于本文列出的命令，缺省 MDC 均支持，非缺省 MDC 只支持 **display irf link**、**mad enable** 以及 **mad exclude interface** 命令。关于 MDC 的详细介绍请参见“虚拟化技术配置指导”中的“MDC”。

1.3.1 组建 IRF 时的注意事项

- 本系列交换机最多支持使用两台设备建立 IRF。
- 在本系列交换机中，只有相同型号的机型之间才能建立 IRF。
- 使用本系列交换机组建 IRF 时，成员编号只能使用 1 或 2。
- IRF 中所有成员设备的软件版本必须相同，如果有软件版本不同的设备要加入 IRF，请确保 IRF 的启动文件同步加载功能处于开启状态。
- 如果两台设备的桥 MAC 地址相同，请修改其中一台设备的桥 MAC 地址，否则，它们不能合并为一个 IRF（这里指设备自身的桥 MAC，使用 **irf mac-address mac-address** 命令配置的 IRF 虚拟设备桥 MAC 不受此影响）。
- 在多台设备形成 IRF 之前，请确保各成员设备上以下配置保持一致：
 - 工作模式（通过 **system-working-mode** 命令配置）。关于工作模式的介绍，请参见“基础配置指导”中的“设备管理”。
 - 设备的聚合能力（通过 **link-aggregation capability** 命令配置）。关于设备聚合能力的介绍，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”。
 - 等价路由增强模式（通过 **ecmp mode enhanced** 命令配置）。关于等价路由增强模式的介绍，请参见“三层技术-IP 路由配置指导”中的“IP 路由基础”。
- 同时配置 MDC 和 IRF MAD 检测功能的情况下，请将 IRF 物理端口和 MAD 检测 VLAN 都配置在缺省 MDC 中，并请先为 MDC 分配物理接口再配置 MAD 功能。
- 在 IRF 分裂后，以及再次合并前，请确保各成员设备上 MDC 的相关配置以及 IRF 的相关配置和分裂前保持一致，避免再次合并后出现错误。

1.3.2 选择IRF物理端口时的注意事项

- 建议为 IRF 端口绑定多个 IRF 物理端口，并选择不同业务板上的端口作为 IRF 物理端口，以提高 IRF 链路的可靠性。
- 设备工作在 IRF 模式下时，在将组内某个端口与 IRF 端口进行绑定或取消绑定之前，必须先对该端口执行 **shutdown** 操作，在完成绑定或取消绑定操作后，再对该端口执行 **undo shutdown** 操作。
- 当选用具有 40GE QSFP+端口的单板进行 IRF 连接时，由于 40GE QSFP+端口具备拆分成 4 个 10GE 端口的功能，当执行拆分或合并操作时需要重启该单板，因此请提前规划好 40GE QSFP+端口的使用方式，避免 IRF 形成后重启单板可能对 IRF 拓扑产生的影响。
- 在独立运行模式下将 IRF 端口和 IRF 物理端口绑定，并不会影响 IRF 物理端口的当前业务。当设备切换到 IRF 模式后，IRF 物理端口的配置将恢复到缺省状态（即原有的业务配置会被删除）。

1.3.3 IRF形成后的配置限制和指导

- IRF 物理端口下只能配置以下命令：
 - 接口基本属性命令，包括 **shutdown**、**description** 和 **flow-interval**。关于这些命令的介绍，请参见“二层技术-以太网交换配置指导”中的“以太网接口”。
 - LLDP 功能命令，包括 **lldp admin-status**、**lldp check-change-interval**、**lldp enable**、**lldp encapsulation snap**、**lldp notification remote-change enable** 和 **lldp tlv-enable**。关于这些命令的介绍，请参见“二层技术-以太网交换配置指导”中的“LLDP”。
- 因为 LACP MAD 和 BFD MAD 的冲突处理原则不同，请不要同时配置。
- 在 LACP MAD 和 BFD 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。
- IRF 域编号是一个全局变量，IRF 中的所有成员设备、所有 MDC 都共用这个 IRF 域编号。在缺省 MDC 上通过 **irf domain** 命令，或者在任意 MDC 上通过 **mad enable** 命令均可修改全局 IRF 域编号。因此，请按照网络规划来修改 IRF 域编号，不要随意修改。
- IRF 迁移到 **Recovery** 状态后会关闭该 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），保留端口可通过 **mad exclude interface** 命令配置。
- 如果接口因为多 **Active** 冲突被关闭，则只能等 IRF 恢复到 **Active** 状态后，接口才能自动被激活，不能通过 **undo shutdown** 命令来激活。

1.4 IRF配置任务简介

成员编号、成员优先级、IRF 端口是形成 IRF 的基本参数，这三个参数的配置方式有两种：

- 设备处于独立运行模式时预配置，使用该方式最终组成 IRF 只需要一次重启。该方式是在独立运行的设备上配置这三个参数，这些配置不会影响本设备的运行，只有设备切换到 IRF 模式下才会生效。在组建 IRF 前，通常使用该方式配置。成员编号必须在独立运行模式时预配置，设备才能切换到 IRF 模式，与别的设备组成 IRF；将成员优先级配置为较大值，当多台设

备初次形成 IRF 时，该设备就能在角色选举中获胜，成为主设备；配置 IRF 端口，以便将运行模式切换到 IRF 模式后，就能直接和别的设备形成 IRF。

- 设备切换到 IRF 模式后再配置。该方式是在一个已经运行在 IRF 模式的设备上配置这三个参数。该配置方式通常用于修改当前配置。比如，将某个成员设备的编号修改为指定值（需要注意的是修改成员编号可能导致原编号相关的部分配置失效）；修改成员设备的优先级，让该设备在下次 IRF 竞选时成为主设备；修改 IRF 端口的已有绑定关系（删除某个绑定或者添加新的绑定），IRF 端口的配置可能会影响本设备的运行（比如引起 IRF 分裂、IRF 合并）。

如上所述，成员编号、成员优先级、IRF 端口配置方式不同，时效不同。建议用户使用以下步骤来建立 IRF：

- (1) 进行网络规划，明确使用哪台设备作为主设备、各成员设备的编号以及成员设备之间的物理连接；
- (2) 在独立运行模式下预配置 IRF，包括配置成员编号、成员优先级、IRF 端口；
- (3) 将当前配置保存到设备的下次启动配置文件，以便设备重启后，IRF 配置能够继续生效；
- (4) 连接 IRF 线缆，确保 IRF 物理端口之间是连通的；
- (5) 将设备的运行模式切换到 IRF 模式（执行该步骤设备会自动重启），形成 IRF；
- (6) 访问 IRF；
- (7) 根据需要，在 IRF 模式下配置 IRF，比如原 IRF 物理端口故障需要绑定其它 IRF 物理端口等。

表1-2 IRF 配置任务简介

配置任务		说明	详细配置
独立运行模式下预配置IRF	配置成员编号	必须先配置成员编号,设备才能从独立运行模式切换到IRF模式	1.5.1
	配置成员优先级	成员优先级、IRF端口在IRF模式下也可以配置,但为了切换到IRF模式后这些配置能够直接生效,建议采用该方式配置	1.5.2
	配置IRF端口		1.5.3
将当前配置保存到设备的下次启动配置文件		必选	1.6
切换IRF模式		必选	1.7
访问IRF		必选	1.8
IRF模式下配置IRF	配置成员编号	必选	1.9.1
	配置成员优先级	可选	1.9.2
	配置IRF端口	如果在独立运行模式下已经配置了IRF端口,则该步骤可选,否则必选	1.9.3
	开启IRF合并自动重启功能	可选	1.9.4
	配置成员设备的描述信息	可选	1.9.5
	配置IRF的桥MAC地址	可选	1.9.6
	配置IRF的桥MAC保留时间	可选	1.9.7
	开启IRF系统启动文件的自动加载功能	可选	1.9.8
	配置IRF链路down延迟上报功能	可选	1.9.9
MAD配置	必须选择一种MAD检测方式进行配置	1.9.10	

1.5 独立运行模式下预配置IRF

1.5.1 配置成员编号

出厂时，设备处于独立运行模式，没有成员编号。必须配置成员编号后，才能将设备从独立运行模式切换到 IRF 模式。用户可以使用 **display irf configuration** 命令查看成员编号，如果“MemberID”字段显示为“--”则表示当前没有配置成员编号。

请确认 IRF 中的成员设备编号唯一。如果存在相同的成员编号，则不能建立 IRF。如果新设备加入 IRF，但是该设备与已有成员设备的编号冲突，则该设备不能加入 IRF。

表1-3 配置成员编号

操作	命令	说明
进入系统视图	system-view	-
在独立运行模式下配置设备的成员编号	irf member member-id	缺省情况下，没有配置成员编号

1.5.2 配置成员优先级

表1-4 配置成员优先级

操作	命令	说明
进入系统视图	system-view	-
在独立运行模式下配置设备的成员优先级	irf priority priority	缺省情况下，设备的成员优先级为1

1.5.3 配置IRF端口

IRF 端口是一个逻辑概念，创建 IRF 端口并与物理端口绑定后，物理端口才可以作为 IRF 物理端口与邻居设备建立 IRF 连接。

表1-5 配置 IRF 端口

操作	命令	说明
进入系统视图	system-view	-
在独立运行模式下创建IRF端口并进入IRF端口视图（如果该IRF端口已经创建，则直接进入IRF端口视图）	irf-port port-number	缺省情况下，设备上没有创建IRF端口

操作	命令	说明
将IRF端口和IRF物理端口绑定	port group [mdc mdc-id] interface interface-type interface-number [mode enhanced]	缺省情况下，IRF端口没有和任何IRF物理端口绑定。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。在本系列交换机上，一个IRF端口最多可以与8个IRF物理端口绑定，但不允许有超过4个IRF物理端口位于同一个芯片上（在Probe视图执行 debug port mapping 命令查看显示信息Unit列可以查看端口和芯片的对应关系）

说明

- IRF 物理端口必须工作在二层模式下，才能与 IRF 端口进行绑定。关于端口工作模式的介绍，请参见“二层技术-以太网交换配置指导”中的“以太网接口配置”。
- 缺省情况下，IRF 物理端口的工作模式为 **enhanced**，且只能工作在 **enhanced** 模式下。

1.6 将当前配置保存到设备的下次启动配置文件

表1-6 将当前配置保存到设备的下次启动配置文件

操作	命令	说明
将当前配置保存到存储介质的根目录下，并将该文件设置为下次启动配置文件	save [safely] [backup main] [force]	该命令可在任意视图下执行

说明

设备建立 IRF 之后，如果进行了其它配置，也需要及时保存；如在未保存配置时发生了主备倒换，则未保存的 IRF 相关配置会丢失。

1.7 切换到IRF模式

注意

- 在切换到 IRF 模式前，必须先配置成员编号。用户可以使用 **display irf configuration** 命令查看成员编号，如果“MemberID”字段显示为“-”则表示当前没有配置成员编号。
- 切换运行模式，设备会自动重启，使新的运行模式生效。

设备缺省处于独立运行模式。要使设备加入 IRF 或使设备的 IRF 配置生效，必须将设备运行模式切换到 IRF 模式。修改运行模式后，设备会自动重启使新的模式生效。为了解决模式切换后配置不可用的问题，在用户执行模式切换操作时，系统会提示用户是否需要自动转换下次启动配置文件。如

果用户选择了<Y>，则设备会自动将下次启动配置文件中槽位和接口的相关配置进行转换并保存，以便当前的配置在模式切换后能够尽可能多的继续生效。比如自动实现将 **slot slot-number** 与 **chassis chassis-number slot slot-number** 的转换、接口编号的转换等。

表1-7 切换到 IRF 模式

操作	命令	说明
进入系统视图	system-view	-
将设备的运行模式切换到IRF模式	chassis convert mode irf	缺省情况下，设备处于独立运行模式 因为管理和维护IRF需要耗费一定的系统资源。如果当前组网中设备不需要和别的设备组成IRF时，请执行 undo chassis convert mode ，将IRF模式切换到独立运行模式

1.8 访问IRF

IRF 模式切换，设备重启后，可通过如下方式登录 IRF：

- 本地登录：通过任意成员设备的 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、SNMP 等方式进行远程登录。

不管使用哪种方式登录 IRF，实际上登录的都是全局主用主控板。全局主用主控板是 IRF 系统的配置和控制中心，在全局主用主控板上配置后，全局主用主控板会将相关配置同步给全局备用主控板，以便保证全局主用主控板和全局备用主控板配置的一致性。

1.9 IRF模式下配置IRF

1.9.1 配置成员编号



注意

在 IRF 中以成员编号标识设备，IRF 端口和成员优先级的配置也和成员编号紧密相关。所以，修改设备成员编号可能导致配置发生变化或者失效，请慎重使用。

请确认 IRF 中的成员设备编号唯一。如果存在相同的成员编号，则不能建立 IRF。如果新设备加入 IRF，但是该设备与已有成员设备的编号冲突，则该设备不能加入 IRF。

- 修改成员编号后，但是没有重启本设备，则原编号继续生效，各物理资源仍然使用原编号来标识。
- 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1-8 配置成员编号

操作	命令	说明
进入系统视图	system-view	-
配置IRF中指定成员设备的成员编号	irf member <i>member-id</i> renumber <i>new-member-id</i>	IRF模式下，设备使用的是独立运行模式下预配置的成员编号



注意

- 需要重启成员编号为 *member-id* 的设备，新成员编号 *new-member-id* 才能生效。
- 当 IRF 中的多个成员设备需要对换编号时，请先将相关的成员设备切换为独立运行模式，修改成员编号后再切换为 IRF 模式，重新加入 IRF。

1.9.2 配置成员优先级

表1-9 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 <i>member-id</i> priority <i>priority</i>	缺省情况下，设备的成员优先级均为1

1.9.3 配置IRF端口

在 IRF 已经建立的情况下，用户可以使用下面的配置增加成员设备间连接所使用的物理端口。



说明

- 在 IRF 已经建立的情况下增加 IRF 物理端口时，需要先将各成员设备上所有的待增加的 IRF 物理端口全部关闭后，再与各 IRF 端口进行绑定。
- 在IRF模式下配置IRF端口时，所需注意的事项与独立运行模式下相同，请参见 [配置限制和指导](#) 以及 [配置IRF端口](#) 部分的介绍。

在配置IRF端口前，请确保IRF合并自动重启功能处于关闭状态，该功能的详细介绍请参见 [开启IRF合并自动重启功能](#)。

表1-10 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>	-

操作	命令	说明
关闭接口	shutdown	-
退回系统视图	quit	-
进入IRF端口视图	irf-port member-id/port-number	-
将IRF端口和IRF物理端口绑定	port group [mdc mdc-id] interface interface-type interface-number [mode enhanced]	缺省情况下, IRF端口没有和任何IRF物理端口绑定 多次执行该命令, 可以将IRF端口与多个IRF物理端口绑定, 以实现IRF链路的备份或负载分担, 从而提高IRF链路的带宽和可靠性。在本系列交换机上, 一个IRF端口最多可以与8个IRF物理端口绑定, 但不允许有超过4个IRF物理端口位于同一个芯片上(在Probe视图执行 debug port mapping 命令查看显示信息Unit列可以查看端口和芯片的对应关系)
退回到系统视图	quit	-
进入IRF物理端口视图	interface interface-type interface-number	-
激活接口	undo shutdown	-
退回系统视图	quit	-
保存当前配置	save	激活IRF端口会引起IRF合并, 进而设备需要重启。为了避免重启后配置丢失, 请在激活IRF端口前先将当前配置保存到下次启动配置文件
激活IRF端口下的配置	irf-port-configuration active	IRF物理线缆连接好, 并将IRF物理端口添加到IRF端口后, 必须通过该命令手工激活IRF端口的配置才能形成IRF

1.9.4 开启IRF合并自动重启功能

IRF合并时, 两台IRF会遵照角色选举的规则进行竞选, 竞选失败方IRF的所有成员设备需要重启才能加入获胜方IRF。其中:

- 如果没有开启IRF合并自动重启功能, 则合并过程中的重启需要用户根据系统提示手工完成。
- 如果开启IRF合并自动重启功能, 则合并过程中的重启由系统自动完成。

表1-11 开启IRF合并自动重启功能

操作	命令	说明
进入系统视图	system-view	-
开启IRF合并自动重启功能	irf auto-merge enable	缺省情况下, IRF合并自动重启功能处于开启状态, 即两台IRF合并时, 竞选失败方将自动重启完成合并 要使IRF合并自动重启功能正常运行, 请在即将合并的两台IRF上都开启IRF合并自动重启功能

1.9.5 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1-12 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 <i>member-id</i> description <i>text</i>	缺省情况下，成员设备没有描述信息

1.9.6 配置IRF的桥MAC地址



说明

- 桥 MAC 变化可能导致流量短时间中断，请谨慎配置。
- 如果两台设备自身的桥 MAC 相同，则它们不能合并为一个 IRF。本功能配置的桥 MAC 地址为 IRF 虚拟设备的桥 MAC。IRF 虚拟设备的桥 MAC 与其它设备 MAC 地址相同时允许合并，只要所有成员设备自身桥 MAC 唯一即可。
- 两台 IRF 合并后，IRF 设备的桥 MAC 为优先级高的一方的桥 MAC。
- 配置了桥 MAC 的 IRF 设备分裂后，分裂出的 IRF 设备的桥 MAC 都为配置的桥 MAC。

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。请确保 IRF 设备桥 MAC 地址在所处二层网络中唯一。如果同一二层网络中存在桥 MAC 相同的设备，则会引起桥 MAC 冲突，从而导致通信故障。

表1-13 配置 IRF 的桥 MAC 地址

操作	命令	说明
进入系统视图	system-view	-
配置IRF的桥MAC地址	irf mac-address <i>mac-address</i>	缺省情况下，IRF设备的桥MAC地址是Master设备的桥MAC地址

1.9.7 配置IRF的桥MAC保留时间



注意

- 桥 MAC 变化可能导致流量短时间中断，请谨慎配置。
- 如果两台设备自身的桥 MAC 相同，则它们不能合并为一个 IRF。
- 如果使用 **irf mac-address mac-address** 命令配置了 IRF 设备桥 MAC 地址，则 IRF 的桥 MAC 始终为该命令配置的桥 MAC 地址，不受本配置影响。

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。未手工配置 IRF 的桥 MAC 时，IRF 会选用某台成员设备的桥 MAC 作为 IRF 的桥 MAC，这台成员设备被称为 IRF 桥 MAC 拥有者。通常情况下，IRF 使用主设备的桥 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：

- 如果配置了 IRF 桥 MAC 保留时间为永久，则不管 IRF 桥 MAC 拥有者是否离开 IRF，IRF 桥 MAC 始终保持不变。
- 如果配置了 IRF 桥 MAC 不保留，则当 IRF 桥 MAC 拥有者离开 IRF 时，系统会立即使用 IRF 中当前主设备的桥 MAC 做 IRF 桥 MAC。

表1-14 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IRF的桥MAC会永久保留	irf mac-address persistent always	二者选其一 缺省情况下，IRF的桥MAC 地址保留时间为永久保留
配置IRF的桥MAC不保留，会立即变化	undo irf mac-address persistent	

1.9.8 开启启动文件的自动加载功能



说明

在独立运行模式下，用户可使用“使能备用主控板启动软件包自动加载功能”来自动保证备用主控板和主用主控板启动软件包版本的一致性；在 IRF 模式下，用户可使用本特性来自动保证全局备用主控板和全局主用主控板启动软件包版本的一致性。关于“使能备用主控板启动软件包自动加载功能”的详细介绍请参见“基础配置指导”中的“软件升级”。

 注意

- 仅当新加入从设备/备用主控板的现有软件版本和主用主控板的软件版本之间支持 ISSU 兼容升级时才能支持自动加载功能。版本间是否支持 ISSU 兼容升级请查看版本说明书中“ISSU 版本兼容列表”。
- 加载启动软件包需要一定时间,在加载期间,请不要插拔或者手工重启处于加载状态的主控板,否则,会导致该主控板加载启动软件包失败而不能启动。用户可打开日志信息显示开关,并根据日志信息的内容来判断加载过程是否开始以及是否结束。

如果新主控板加入 IRF, 并且新主控板的软件版本和全局主用主控板的软件版本不一致, 则新加入的主控板不能正常启动。此时:

- 如果没有开启启动文件的自动加载功能, 则需要用户手工升级新主控板后, 再将新主控板加入 IRF。或者在主设备上开启启动文件的自动加载功能, 断电重启新设备, 让新主控板重新加入 IRF。
- 如果已经开启了启动文件的自动加载功能, 则新主控板加入 IRF 时, 会与全局主用主控板的软件版本号进行比较, 如果不一致, 则自动从全局主用主控板下载启动文件, 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主控板上原有启动文件文件名重名, 则原有启动文件会被覆盖。

为了能够自动加载成功, 请确保从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。如果从设备存储介质上空闲空间不足, 系统会自动删除从设备的当前启动文件来完成加载。如果删除从设备的当前启动文件后空间仍然不足, 从设备将无法进行自动加载。此时, 需要管理员重启从设备并进入从设备的 Boot ROM 菜单, 删除一些不重要的文件后, 再让从设备重新加入 IRF。

表1-15 开启 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	-
开启IRF系统启动文件的自动加载功能	irf auto-update enable	缺省情况下, IRF系统启动文件自动加载功能处于开启状态

1.9.9 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。经过一定的时间间隔后, 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1-16 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	缺省情况下，IRF链路down延迟上报时间为4秒 当IRF链路down延迟时间为缺省值时，如果IRF链路的不稳定状态持续时间不超过4秒，则不会导致IRF分裂；但如果某些协议配置的超时时间小于4秒（例如VRRP、OSPF等），该协议将超时。此时请适当调整IRF链路down的延迟上报时间或者该协议的超时时间，使IRF链路down的延迟上报时间小于协议超时时间，保证协议状态不会发生不必要的切换 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

1.9.10 MAD配置

设备支持的 MAD 检测方式有：LACP MAD 检测和 BFD MAD 检测。两种 MAD 检测机制各有特点，用户可以根据现有组网情况进行选择。因为 LACP MAD 和 BFD MAD 冲突处理的原则不同，请不要同时配置。

表1-17 MAD 检测机制的比较

MAD 检测方式	优势	限制
LACP MAD	检测速度快，利用现有聚合组网即可实现，无需占用额外接口，利用聚合链路同时传输普通业务报文和 MAD 检测报文（扩展 LACP 报文）	组网中需要使用 H3C 设备作为中间设备，每个成员设备都需要连接到中间设备
BFD MAD	检测速度较快，组网形式灵活，对其它设备没有要求	配置专用三层接口，这些接口不能再传输普通业务流量 <ul style="list-style-type: none"> 如果不使用中间设备，则要求成员设备间是全链接，即每个成员设备都必须和其它所有成员设备相连。该链路专用于 MAD 检测，不能再传输普通业务流量。该方式适用于成员设备少，并且物理距离比较近的组网环境 如果使用中间设备，组网时每个成员设备都需要连接到中间设备，这些 BFD 链路专用于 MAD 检测

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（等于主设备的成员编号）。

开启 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。

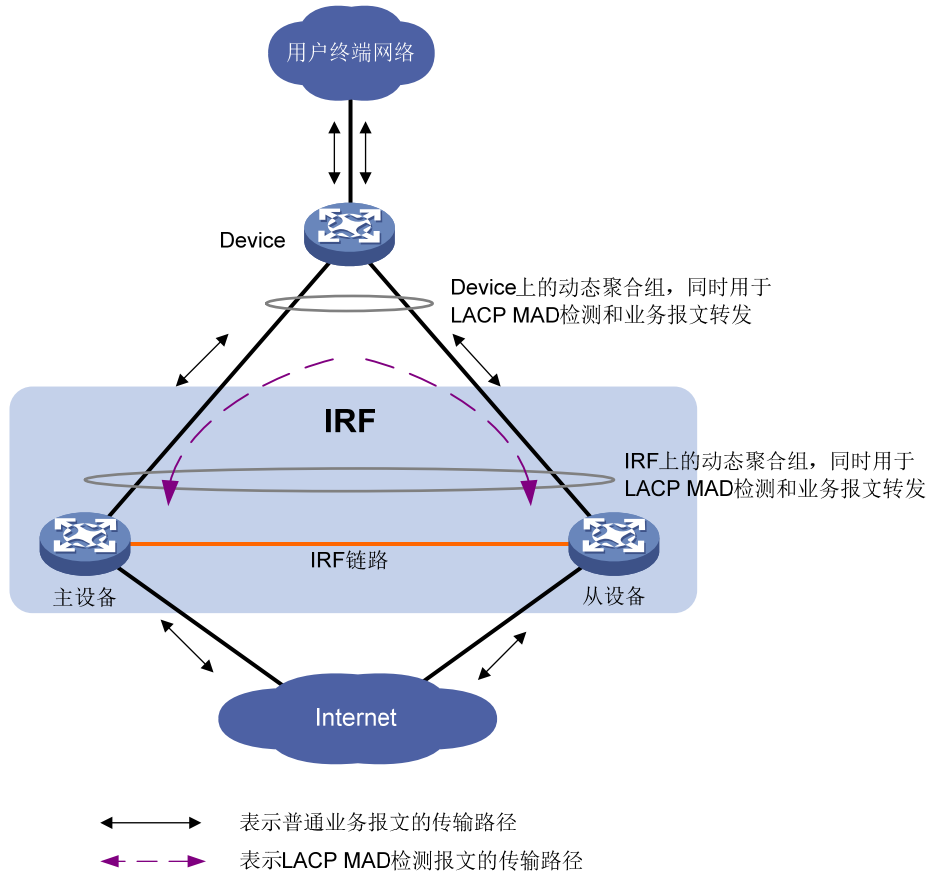
- 当成员设备收到 LACP 协议报文后，先比较 DomainID。如果 DomainID 相同，再比较 ActiveID；如果 DomainID 不同，则认为报文来自不同 IRF，不再进行 MAD 处理。

- 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H3C设备作为中间设备。通常采用如 [图 1-7](#) 所示的组网，成员设备之间通过中间设备（Device）交互LACP扩展报文。

图1-7 LACP MAD 检测组网示意图



⚠ 注意

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。

(3) 配置 LACP MAD 检测

LACP MAD 检测的配置步骤为：

- 创建聚合接口；（中间设备上也需要进行该项配置）
- 将聚合接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）
- 在动态聚合接口下开启 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1-18 配置 LACP MAD 检测

操作	命令	说明
进入系统视图	system-view	-
(可选) 配置IRF域编号	irf domain <i>domain-id</i>	缺省情况下, IRF的域编号为0
创建并进入二层聚合接口视图	interface bridge-aggregation <i>interface-number</i>	-
配置聚合组工作在动态聚合模式下	link-aggregation mode dynamic	缺省情况下, 聚合组工作在静态聚合模式下
开启LACP MAD检测功能	mad enable	缺省情况下, LACP MAD检测未开启
退回系统视图	quit	-
进入以太网接口视图	interface <i>interface-type</i> <i>interface-number</i>	-
将以太网接口加入聚合组	port link-aggregation group <i>number</i>	-

2. BFD MAD检测



说明

仅缺省 MDC 支持 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, 除在三层接口下开启 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, 且所有成员设备的 MAD IP 必须属于同一网段。

- 当 IRF 正常运行时, 只有主设备上配置的 MAD IP 地址生效, 从设备上配置的 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 分裂形成多个 IRF 时, 不同 IRF 中主设备上配置的 MAD IP 地址均会生效, BFD 会话被激活, 此时会检测到多 Active 冲突。

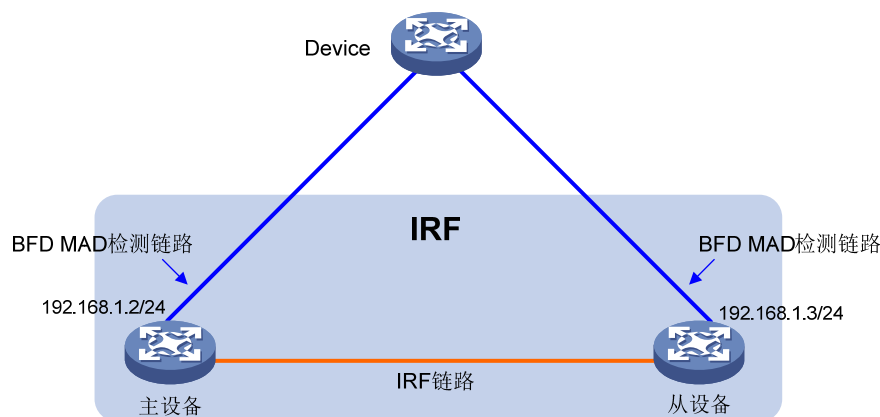
(2) BFD MAD 检测组网要求

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。在使用中间设备时, 可以使用以太网端口或管理以太网口来实现 BFD MAD 检测, 如 [图 1-8](#) 所示。

用于 BFD MAD 检测的以太网端口需要属于同一 VLAN, 在该 VLAN 接口视图下为不同成员设备配置同一网段内的不同 MAD IP 地址。

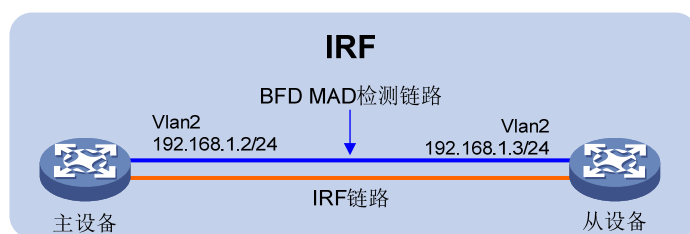
如果使用管理以太网口实现 BFD MAD 检测, 只需要为每台成员设备配置同一网段内的不同 MAD IP 地址即可。

图1-8 使用中间设备实现 BFD MAD 检测组网示意图



在没有中间设备时，需要采用如 [图 1-9](#) 所示的组网方式：每台成员设备必须和其它所有成员设备之间使用以太网端口建立BFD MAD检测链路（即成员设备之间是全连接组网）。这些链路连接的接口必须属于同一VLAN，在该VLAN接口视图下为不同成员设备配置同一网段下的不同IP地址。

图1-9 不使用中间设备实现 BFD MAD 检测组网示意图



 提示

除管理用以太网口外，开启 BFD MAD 检测功能的三层接口只能专用于 BFD MAD 检测，该接口下建议只配置 **mad bfd enable** 和 **mad ip address** 命令。如果用户配置了其它命令，可能会影响该业务以及 BFD MAD 检测功能的运行。

 注意

在 BFD 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。

(3) 配置 BFD MAD 检测

配置 BFD MAD 检测时，请遵循以下要求：

- 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。
- 开启 BFD MAD 检测功能的 VLAN 接口以及对应 VLAN 仅用于 BFD MAD 功能，请勿配置任何其它功能。

- BFD MAD 检测 VLAN 中仅加入用于 BFD MAD 检测的端口。
- 假设开启 BFD MAD 检测功能的 VLAN 接口为 D，请不要创建编号为 D 的三层以太网子接口或三层聚合子接口（三层以太网子接口的编号规则为 **interface type A/B/C.D**，三层聚合子接口的编号规则为 **interface type A.D**，D 表示子接口编号），否则会导致该 VLAN 不可用。
- 不允许在 Vlan-interface1 接口上开启 BFD MAD 检测功能。
- BFD MAD 检测功能与生成树功能互斥，在开启了 BFD MAD 检测功能的三层接口对应 VLAN 内的端口上，请不要开启生成树协议。
- 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。
- 为保证 MAD 检测功能正常运行，请不要将 MAD IP 地址配置为设备上已经使用的 IP 地址。

使用以太网端口进行 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；（对于使用中间设备的组网，中间设备上也需要进行该项配置）
- 确定哪些物理端口用于 BFD MAD 检测，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，在接口下开启 BFD MAD 检测功能，并配置 MAD IP 地址。

表1-19 配置使用以太网端口进行 BFD MAD 检测

操作		命令	说明
进入系统视图		system-view	-
（可选）配置IRF域编号		irf domain domain-id	缺省情况下，IRF的域编号为0
创建一个新VLAN专用于BFD MAD检测		vlan vlan-id	缺省情况下，设备上只存在VLAN 1
退回系统视图		quit	-
进入以太网接口视图		interface interface-type interface-number	-
端口加入 BFD MAD 检测专用 VLAN	Access端口	port access vlan vlan-id	请根据端口的当前链路类型选择对应的配置命令 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
	Trunk端口	port trunk permit vlan vlan-id	
	Hybrid端口	port hybrid vlan vlan-id { tagged untagged }	
退回系统视图		quit	-
进入VLAN接口视图		interface vlan-interface interface-number	-
开启BFD MAD检测功能		mad bfd enable	缺省情况下，没有开启BFD MAD检测功能
给指定成员设备配置MAD IP地址		mad ip address ip-address { mask mask-length } member member-id	缺省情况下，没有为接口配置MAD IP地址

使用管理用以太网口进行 BFD MAD 检测功能的配置顺序为：

- 将 IRF 中所有成员设备的管理用以太网口连接到同一台中间设备。如果成员设备安装了两块主控板，请将每块主控板的管理用以太网口都连接到中间设备，避免主备倒换后 BFD MAD 检测失效。
- 将中间设备上与 IRF 成员设备相连的端口配置在一个 VLAN 内。
- 在主设备的管理用以太网口下开启 BFD MAD 检测功能，并为每台成员设备配置 MAD IP 地址。

表1-20 配置使用管理用以太网口进行 BFD MAD 检测

操作	命令	说明
进入系统视图	system-view	-
(可选) 配置IRF域编号	irf domain <i>domain-id</i>	缺省情况下，IRF的域编号为0
进入主设备管理用以太网口的接口视图	interface M-GigabitEthernet <i>interface-number</i>	-
开启BFD MAD检测功能	mad bfd enable	缺省情况下，没有开启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>	缺省情况下，没有为接口配置MAD IP地址

3. 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 IRF 中的所有业务接口。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。

使用 VLAN 接口进行远程登录时，需要将该 VLAN 接口及其对应的以太网端口都配置为保留接口。但如果在正常工作状态的 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。

表1-21 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>	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 IRF 物理端口和 Console 口自动作为保留接口，不需要配置

4. MAD故障恢复

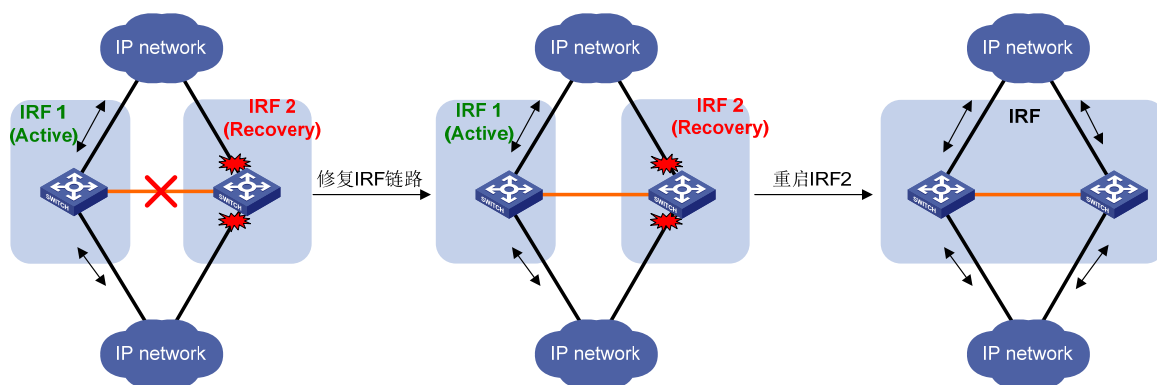
IRF 链路故障将一个 IRF 分裂为两个 IRF，从而导致多 Active 冲突。当系统检测到多 Active 冲突后，两个冲突的 IRF 会进行竞选，主设备成员编号小的获胜，继续正常运行，失败的 IRF 会转入 Recovery 状态，暂时不能转发业务报文。此时通过修复 IRF 链路可以恢复 IRF 系统（设备会尝试自动修复 IRF 链路，如果修复失败的话，则需要用户手工修复）。IRF 链路修复后，系统会自动重启或者给出提示信息要求用户手工重启处于 Recovery 状态的 IRF。重启后，原 Recovery 状态 IRF 中所有成员设备以

从设备身份加入原Active状态的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，整个IRF系统恢复，如 图 1-10 所示。

 注意

- 系统是否会自动重启或者给出提示信息要求用户手工重启处于 Recovery 状态的 IRF，与用户是否配置了 **irf auto-merge enable** 命令有关。
- 请根据提示重启处于 Recovery 状态的 IRF，如果错误的重启了 Active 状态的 IRF，会导致合并后的 IRF 仍然处于 Recovery 状态，所有成员设备的业务接口都会被关闭。此时，需要执行 **mad restore** 命令让整个 IRF 系统恢复。

图1-10 MAD 故障恢复（IRF 链路故障）



如果MAD故障还没来得及修复而处于Active的IRF也故障了（原因可能是设备故障或者上下行线路故障），如 图 1-11 所示。此时可以在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，先接替IRF 1工作。然后再修复IRF 1和IRF链路，修复后，两个IRF发生合并，整个IRF系统恢复。

图1-11 MAD 故障恢复（IRF 链路故障+Active 状态的 IRF 故障）

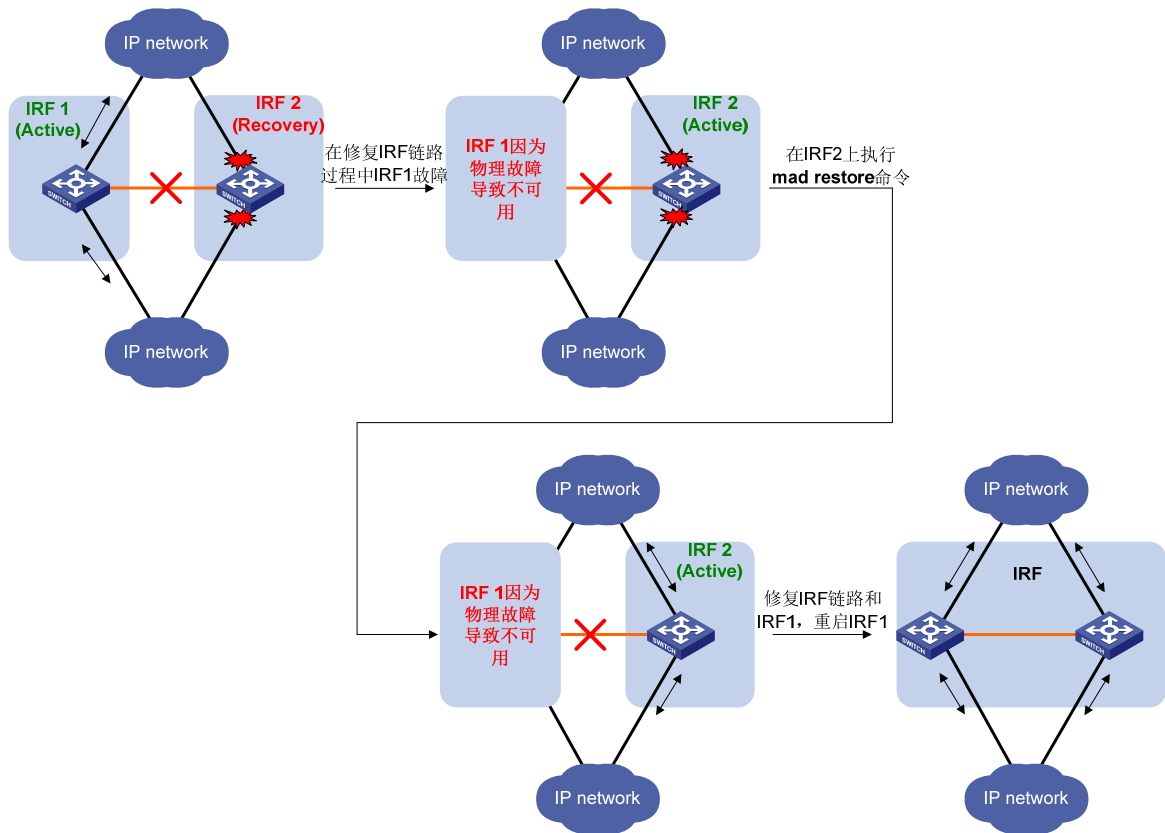


表1-22 手动恢复处于 Recovery 状态的设备

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	-

说明

对于因 MAD 检测冲突而转入 Recovery 状态的设备，如果需要开启被关闭的端口，建议使用 **mad restore** 命令将设备恢复至 Active 状态，而不要在端口上执行 **undo shutdown** 命令进行手工恢复。

1.10 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1-23 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
显示IRF的拓扑信息	display irf topology
显示IRF链路信息	display irf link
显示所有成员设备上重启以后生效的IRF配置	display irf configuration
显示MAD配置信息	display mad [verbose]

1.11 IRF典型配置举例

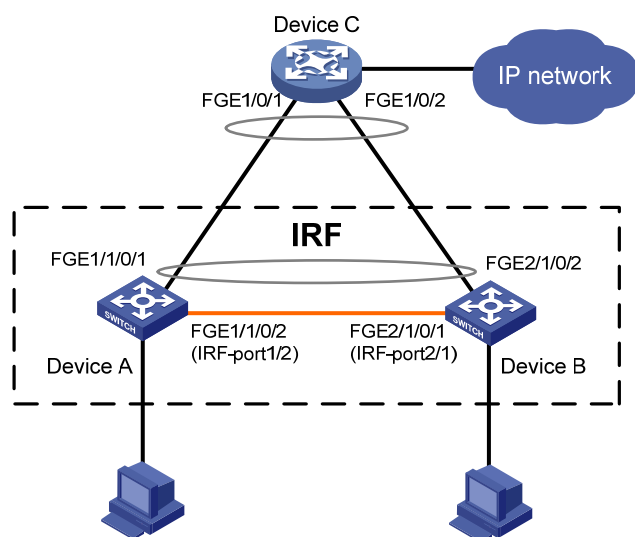
1.11.1 IRF典型配置举例（LACP MAD检测方式）

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1-12 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

(1) 配置 Device A

配置 Device A 的成员编号为 1，创建 IRF 端口 2，并将它与物理端口 FortyGigE1/0/2 绑定。

```
<Sysname> system-view
[Sysname] irf member 1
[Sysname] irf-port 2
[Sysname-irf-port2] port group interface FortyGigE 1/0/2
[Sysname-irf-port2] quit
```

将当前配置保存到下次启动配置文件。

```
[Sysname] quit
<Sysname> save
```

将设备的运行模式切换到 IRF 模式。

```
<Sysname> system-view
[Sysname] chassis convert mode irf
The device will switch to IRF mode and reboot.
You are recommended to save the current running configuration and specify the configuration
file for the next startup. Continue? [Y/N]:y
Do you want to convert the content of the next startup configuration file flash:/startup.cfg
to make it available in IRF mode? [Y/N]:y
Now rebooting, please wait...
```

设备重启后 Device A 组成了只有一台成员设备的 IRF。

(2) 配置 Device B

配置 Device B 的成员编号为 2，创建 IRF 端口 1，并将它与物理端口 FortyGigE1/0/1 绑定。

```
<Sysname> system-view
[Sysname] irf member 2
[Sysname] irf-port 1
[Sysname-irf-port1] port group interface FortyGigE 1/0/1
[Sysname-irf-port1] quit
```

将当前配置保存到下次启动配置文件。

```
[Sysname] quit
<Sysname> save
```

参照 [图 1-12](#) 进行物理连线。

将设备的运行模式切换到 IRF 模式。

```
<Sysname> system-view
[Sysname] chassis convert mode irf
The device will switch to IRF mode and reboot.
You are recommended to save the current running configuration and specify the configuration
file for the next startup. Continue? [Y/N]:y
Do you want to convert the content of the next startup configuration file flash:/startup.cfg
to make it available in IRF mode? [Y/N]:y
Now rebooting, please wait...
```

设备 B 重启后与设备 A 形成 IRF。

(3) 配置 LACP MAD 检测

设置 IRF 域编号为 1。

```
<Sysname> system-view
[Sysname] irf domain 1
```

创建一个动态聚合接口，并开启 LACP MAD 检测功能。

```
[Sysname] interface bridge-aggregation 2
[Sysname-Bridge-Aggregation2] link-aggregation mode dynamic
[Sysname-Bridge-Aggregation2] mad enable
You need to assign a domain ID (range: 0-4294967295)
[Current domain is: 0]: 1
The assigned domain ID is: 1
```

MAD LACP only enable on dynamic aggregation interface.

```
[Sysname-Bridge-Aggregation2] quit
```

在聚合接口中添加成员端口 FortyGigE1/1/0/1 和 FortyGigE2/1/0/2，专用于 Device A 和 Device B 实现 LACP MAD 检测。

```
[Sysname] interface FortyGigE 1/1/0/1
[Sysname-FortyGigE1/1/0/1] port link-aggregation group 2
[Sysname-FortyGigE1/1/0/1] quit
[Sysname] interface FortyGigE 2/1/0/2
[Sysname-FortyGigE2/1/0/2] port link-aggregation group 2
```

(4) 配置中间设备 Device C



提示

如果中间设备是一个 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。

Device C 作为中间设备来转发、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 协议扩展功能的交换机即可。

创建一个动态聚合接口。

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

在聚合接口中添加成员端口 FortyGigE1/0/1 和 FortyGigE1/0/2，用于帮助 LACP MAD 检测。

```
[Sysname] interface FortyGigE 1/0/1
[Sysname-FortyGigE1/0/1] port link-aggregation group 2
[Sysname-FortyGigE1/0/1] quit
[Sysname] interface FortyGigE 1/0/2
[Sysname-FortyGigE1/0/2] port link-aggregation group 2
```

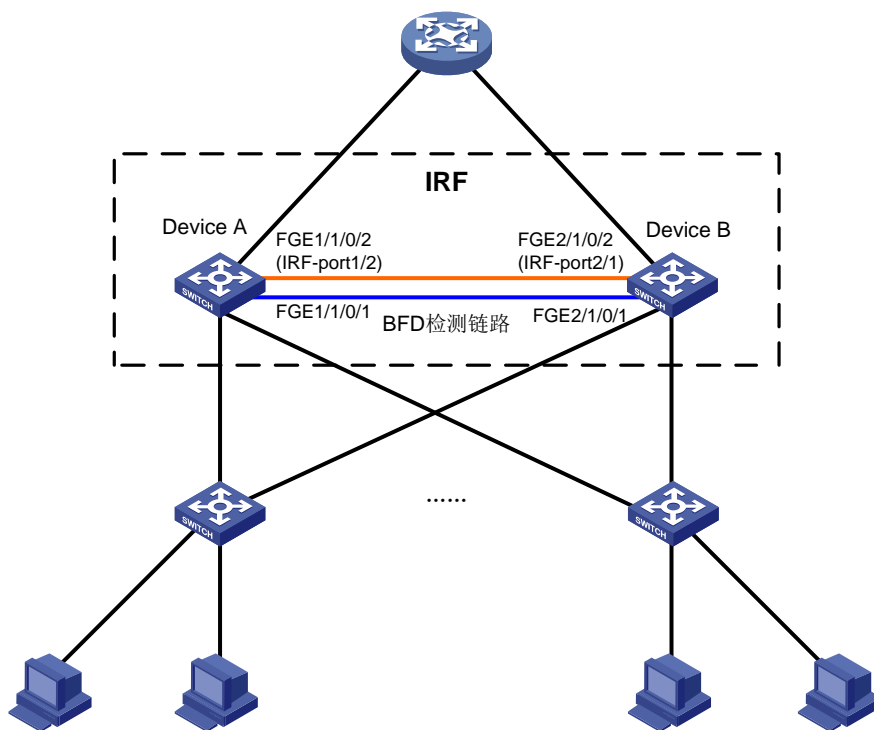
1.11.2 IRF 典型配置举例（BFD MAD 检测方式）

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1-13 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 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

(1) 配置 Device A

设置 Device A 的成员编号为 1。

```
<Sysname> system-view  
[Sysname] irf member 1
```

将设备的运行模式切换到 IRF 模式。

```
<Sysname> system-view  
[Sysname] chassis convert mode irf
```

The device will switch to IRF mode and reboot.

You are recommended to save the current running configuration and specify the configuration file for the next startup. Continue? [Y/N]:y

Do you want to convert the content of the next startup configuration file flash:/startup.cfg to make it available in IRF mode? [Y/N]:y

Now rebooting, please wait...

设备重启后 **Device A** 形成了只有一台成员设备的 IRF。

关闭 **FortyGigE1/1/0/2** 端口。

```
[Sysname] interface FortyGigE1/1/0/2
[Sysname-FortyGigE1/1/0/2] shutdown
[Sysname-FortyGigE1/1/0/2] quit
```

创建 IRF 端口 **1/2**，并将它与物理端口 **FortyGigE1/1/0/2** 绑定。

```
[Sysname] irf-port 1/2
[Sysname-irf-port1/2] port group interface FortyGigE 1/1/0/2
[Sysname-irf-port1/2] quit
```

开启 **FortyGigE1/1/0/2** 端口。

```
[Sysname] interface FortyGigE1/1/0/2
[Sysname-FortyGigE1/1/0/2] undo shutdown
[Sysname-FortyGigE1/1/0/2] quit
```

将当前配置保存到下次启动配置文件。

```
[Sysname] quit
<Sysname> save
```

(2) 配置 Device B

配置 **Device B** 的成员编号为 **2**。

```
<Sysname> system-view
[Sysname] irf member 2
```

将设备的运行模式切换到 IRF 模式。

```
[Sysname] chassis convert mode irf
```

The device will switch to IRF mode and reboot.

You are recommended to save the current running configuration and specify the configuration file for the next startup. Continue? [Y/N]:y

Do you want to convert the content of the next startup configuration file flash:/startup.cfg to make it available in IRF mode? [Y/N]:y

Now rebooting, please wait...

设备重启后 **Device B** 形成了只有一台成员设备的 IRF。

关闭 **FortyGigE2/1/0/2** 端口。

```
[Sysname] interface FortyGigE2/1/0/2
[Sysname-FortyGigE2/1/0/2] shutdown
[Sysname-FortyGigE2/1/0/2] quit
```

创建 IRF 端口 **2/1**，并将它与物理端口 **FortyGigE2/1/0/2** 绑定。

```
[Sysname] irf-port 2/1
[Sysname-irf-port2/1] port group interface FortyGigE 2/1/0/2
[Sysname-irf-port2/1] quit
```

开启 **FortyGigE2/1/0/2** 端口。

```
[Sysname] interface FortyGigE2/1/0/2
[Sysname-FortyGigE2/1/0/2] undo shutdown
[Sysname-FortyGigE2/1/0/2] quit
```

将当前配置保存到下次启动配置文件。

```
[Sysname] quit
<Sysname> save
```


(3) 激活 IRF 端口的配置

参照 [图 1-13](#) 进行物理连线。

激活 Device A 设备的 IRF 端口配置。

```
[Sysname] irf-port-configuration active
```

激活 Device B 设备的 IRF 端口配置，系统将提示发生 IRF 合并，用户需要手工重启其中一台设备，加入以另一台设备为主设备的 IRF。

```
[Sysname] irf-port-configuration active
```

```
%Jul 9 09:04:48:279 2013 STM/4/STM_MERGE_NEED_REBOOT: -MDC=1; IRF merge occurs and the IRF system needs a reboot.
```

```
%Jul 9 14:03:06:855 2013 STM/5/STM_MERGE: -MDC=1; IRF merge occurs and the IRF system does not need to reboot.
```

在本例中，使用 Device A 作为主设备，因此需要重启 Device B。

```
[Sysname] quit
```

```
<Sysname> reboot
```

重启完成后，Device A 和 Device B 合并成为一个 IRF，Device A 为主设备。

(4) 配置 BFD MAD 检测

创建 VLAN 3，并将 Device A（成员编号为 1）上的端口 FortyGigE1/1/0/1 和 Device B（成员编号为 2）上的端口 FortyGigE2/1/0/1 加入 VLAN 中。

```
<Sysname> system-view
```

```
[Sysname] vlan 3
```

```
[Sysname-vlan3] port FortyGigE 1/1/0/1 FortyGigE 2/1/0/1
```

```
[Sysname-vlan3] quit
```

创建 VLAN 接口 3，并配置 MAD IP 地址。

```
[Sysname] interface vlan-interface 3
```

```
[Sysname-Vlan-interface3] mad bfd enable
```

```
[Sysname-Vlan-interface3] mad ip address 192.168.2.1 24 member 1
```

```
[Sysname-Vlan-interface3] mad ip address 192.168.2.2 24 member 2
```

```
[Sysname-Vlan-interface3] quit
```

因为 BFD MAD 和生成树功能互斥，所以在 FortyGigE1/1/0/1 和 FortyGigE2/1/0/1 上关闭生成树协议。

```
[Sysname] interface FortyGigE 1/1/0/1
```

```
[Sysname-FortyGigE-1/1/0/1] undo stp enable
```

```
[Sysname-FortyGigE-1/1/0/1] quit
```

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

```
[Sysname-FortyGigE-2/1/0/1] undo stp enable
```

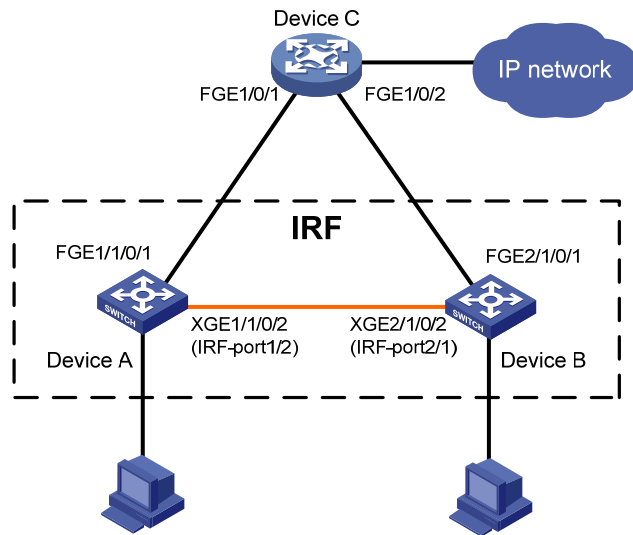
1.11.3 将成员设备从IRF模式恢复到独立运行模式配置举例

1. 组网需求

如 [图 1-14](#) 所示，IRF 已经稳定运行，Device A 和 Device B 是 IRF 的成员设备。现因网络调整，需要将 Device A 和 Device B 从 IRF 模式下恢复到独立运行模式待用。

2. 组网图

图1-14 将成员设备从 IRF 模式恢复到独立运行模式组网图



3. 配置思路

- (1) 断开 IRF 连接。可以直接将 IRF 物理连接线缆拔出也可以使用命令行关闭主设备上所有的 IRF 物理端口。本举例采用命令行关闭的方式。
- (2) IRF 分裂后，分别将两台成员设备从 IRF 模式切换到独立运行模式。

4. 配置步骤

- (1) 确定主设备。

```
<IRF> display irf
```

MemberID	Slot	Role	Priority	CPU-Mac	Description
*+1	16	Master	1	00e0-fc0a-15e0	DeviceA
1	17	Standby	1	00e0-fc0f-8c02	DeviceA
2	16	Standby	1	00e0-fc0f-15e1	DeviceB
2	17	Standby	1	00e0-fc0f-15e2	DeviceB

* indicates the device is the master.

+ indicates the device through which the user logs in.

The Bridge MAC of the IRF is: 000f-e26a-58ed

Auto upgrade : no

Mac persistent : always

Domain ID : 0

通过以上显示信息可以看出，Device A 是主设备。

- (2) 断开 IRF 连接：手工关闭主设备（Device A）的 IRF 物理端口 FortyGigE1/1/0/2。（本举例中只有一条 IRF 物理链路，如果有多条，则需要手工关闭所有的 IRF 物理端口）

```
<IRF> system-view
```

```
[IRF] interface FortyGigE 1/1/0/2
```

```
[IRF-FortyGigE1/1/0/2] shutdown
```

```
[IRF-FortyGigE1/1/0/2] quit
```

(3) 将 Device A 的运行模式切换到独立运行模式。

```
[IRF] undo chassis convert mode
```

The device will switch to stand-alone mode and reboot.

You are recommended to save the current running configuration and specify the configuration file for the next startup. Continue? [Y/N]:y

Do you want to convert the content of the next startup configuration file flash:/vrpcfg.cfg to make it available in stand-alone mode? [Y/N]:y

Now rebooting, please wait...

Device A 自动重启来完成模式的切换。

(4) 登录 Device B 后，将 Device B 的运行模式切换到独立运行模式。

```
<IRF> system-view
```

```
[IRF] undo chassis convert mode
```

The device will switch to stand-alone mode and reboot.

You are recommended to save the current running configuration and specify the configuration file for the next startup. Continue? [Y/N]:y

Do you want to convert the content of the next startup configuration file flash:/vrpcfg.cfg to make it available in stand-alone mode? [Y/N]:y

Now rebooting, please wait...

Device B 自动重启来完成模式的切换。



如果 IRF 上创建了 VLAN 接口、配置了 IP 地址，并且 Device A 和 Device B 上都存在该 VLAN 的成员端口（即配置了端口加入 VLAN）。此时，Device A 和 Device B 恢复到独立运行模式后，会产生 IP 地址冲突，请登录其中一台设备，修改该 VLAN 接口的 IP 地址。

目 录

1 MDC	1-1
1.1 MDC简介	1-1
1.1.1 MDC的应用	1-1
1.2 缺省MDC和非缺省MDC	1-2
1.3 MDC配置任务简介	1-2
1.4 创建MDC	1-3
1.5 为MDC分配硬件资源	1-3
1.5.1 为MDC分配物理接口和业务板	1-4
1.5.2 设置MDC的CPU权重	1-6
1.5.3 为MDC分配内存空间	1-7
1.6 启动MDC	1-7
1.7 访问和管理MDC	1-7
1.8 MDC显示和维护	1-8
1.9 MDC典型配置举例（独立运行模式）	1-9
1.10 MDC典型配置举例（IRF模式）	1-12

1 MDC



MDC 功能需要安装 License 才能使用。当 License 到期或被卸载后，所有已创建的非缺省 MDC 都会被停止，且不允许再创建或启动 MDC，请重新安装有效的 License。关于 License 的详细介绍请参见“基础配置指导”中的“License 管理”。

1.1 MDC简介

通过虚拟化技术将一台物理设备划分成多台逻辑设备，每台逻辑设备就称为一台 MDC (multitenant device contexts, 多租户设备环境)。

每台 MDC 拥有自己专属的软硬件资源，独立运行，独立转发，独立提供业务。创建、启动、重启、删除一台 MDC，不会影响其它 MDC 的运行。

对于用户来说，每台 MDC 就是一台独立的物理设备。MDC 之间相互隔离，不能直接通信，具有很高的安全性。

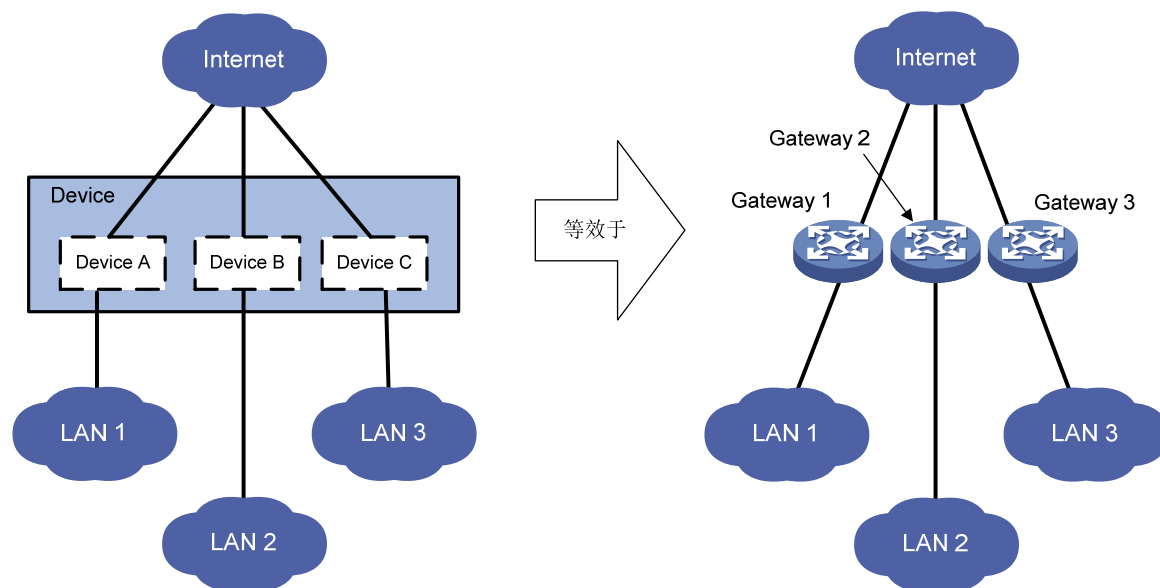
对于管理者来说，当有新的分支机构加入时，可通过划分 MDC，来节省购置新网络设备和网络设备硬件升级的开销，提高现有网络资源利用率。同时，多台 MDC 集成在一台物理设备上，又有效的减少了管理和维护成本。

1.1.1 MDC的应用

MDC 的应用十分广泛，比如提供设备出租和业务托管，应用于实验教学等。

如 [图 1-1](#) 所示，LAN 1、LAN 2 和 LAN 3 是三个不同公司的局域网，它们通过同一台物理设备 Device 连接到外网。通过虚拟化技术，能让一台设备当三台设备使用。具体做法是，在 Device 上创建三台 MDC：Device A、Device B 和 Device C，分别负责 LAN 1、LAN 2、LAN 3 的网络接入和控制。LAN 1、LAN 2、LAN 3 的网络管理员可以（也只能）分别登录到自己的接入设备进行配置、保存、重启等操作，不会影响其它网络的使用，其效果等同于 LAN 1、LAN 2 和 LAN 3 分别通过各自的网关 Gateway 1、Gateway 2、Gateway 3 接入 Internet。

图1-1 MDC 组网应用示意图



1.2 缺省MDC和非缺省MDC

- 设备支持MDC功能后，整台物理设备就是一个MDC，称为缺省MDC，如 [图 1-1](#) 中的Device。当用户登录物理设备时，实际登录的就是缺省MDC。用户在物理设备上的配置实质就是对缺省MDC的配置。缺省MDC的名称为Admin，编号为 1。缺省MDC不需要创建，不能删除。
- 与缺省MDC相对应的是非缺省MDC，如 [图 1-1](#) 中的Device A、Device B、Device C。
- 缺省 MDC 拥有对整台物理设备的所有权限，它可以使用和管理设备所有的硬件资源。缺省 MDC 下可以创建/删除非缺省 MDC，给非缺省 MDC 分配接口、CPU 资源、内存空间，没有分配的接口、CPU 资源、内存空间由缺省 MDC 使用和管理。
- 非缺省 MDC 下不可再创建/删除非缺省 MDC，它只能使用缺省 MDC 分配给自己的硬件资源，并在缺省 MDC 指定的硬件资源限制范围内工作，不能抢占其他 MDC 或者系统剩余的硬件资源。

说明

- 如无特殊说明，下文中的 MDC 均指非缺省 MDC。
- 如无特殊说明，本文所列操作均在缺省 MDC 下执行。

1.3 MDC配置任务简介

为 MDC 分配物理接口、设置 MDC 的 CPU 权重、为 MDC 分配内存空间可以在启动 MDC 前配置也可以在启动 MDC 后配置。建议先做好规划，在启动 MDC 前配置。

请先组建或加入 IRF 再创建和配置 MDC。因为设备组建或加入 IRF 后，会以 IRF 中主设备的配置重启，除设备上的 IRF 端口配置外，之前的其它配置将不再生效。

表1-1 MDC 配置任务简介

配置任务		说明	详细配置
创建MDC		必选	1.4
为MDC分配硬件资源	为MDC分配物理接口和业务板	必选	1.5.1
	设置MDC的CPU权重	可选	1.5.2
	为MDC分配内存空间	可选	1.5.3
启动MDC		必选	1.6
访问和管理MDC		必选	1.7

1.4 创建MDC

创建 MDC 相当于构造了一台新的设备。但此时 MDC 还没有初始化，没有启动，不能使用。

表1-2 创建 MDC

操作	命令	说明
进入系统视图	system-view	-
创建MDC	mdc mdc-name [id mdc-id]	缺省情况下，设备上存在缺省MDC，该MDC的名称为Admin，编号为1。缺省MDC不需要创建，不能删除 最多可创建8个非缺省MDC

删除 MDC 时，请按照以下推荐步骤执行，勿直接删除 MDC，避免涉及的业务板后续被分配给其他 MDC 时出现问题：

- (1) 进入 MDC 的视图。
- (2) 使用 **display this** 命令查看该 MDC 下生效的配置。
- (3) 执行 **undo mdc start** 命令停止运行当前 MDC。
- (4) 执行 **undo location** 命令，分别取消该 MDC 对所有业务板的使用权。
- (5) 执行 **undo allocate interface** 命令，将接口从该 MDC 中删除。
- (6) 执行 **undo mdc** 命令，删除 MDC。

1.5 为MDC分配硬件资源

缺省 MDC 上可以给非缺省 MDC 分配的硬件资源包括业务接口、CPU 资源和内存空间。MDC 创建后，缺省 MDC 的管理员：

- 必须为非缺省 MDC 分配物理接口和业务板，否则非缺省 MDC 上将没有业务接口，不能转发业务报文。
- 可以根据需要在各 MDC 间进行 CPU 资源和内存空间的合理、灵活分配，因为系统在创建 MDC 时已经自动预留了一定的 CPU 资源和内存空间以保证 MDC 的正常运行。

物理设备的 Console 口不能分配。Console 口属于缺省 MDC，非缺省 MDC 上没有 Console 口。

物理设备的管理以太网接口不能也不需要分配。物理管理以太网接口属于缺省 MDC，系统在创建非缺省 MDC 时会在该 MDC 下创建相同数量的虚拟管理以太网接口。

- 所有 MDC 的管理以太网接口名称和编号相同，相同编号的管理以太网口共用同一个物理接口和物理链路。
 - 不同 MDC 的管理以太网接口下可以配置相同网段或者不同网段的 IP 地址，以便不同 MDC 的管理员登录自己的 MDC。
 - 只有在缺省 MDC 下可以对管理以太网接口执行 **shutdown** 命令，非缺省 MDC 下不可以。
- 下文文中为 MDC 分配的物理接口均指业务接口。

1.5.1 为MDC分配物理接口和业务板

1. 配置指导及注意事项

(1) 独立运行模式和 IRF 模式下的通用注意事项

- MDC 必须同时拥有接口和接口所属业务板的使用权限，才能使用接口。分配顺序为先分配接口，再分配业务板使用权限。
- 业务板上的接口是按组划分的。为MDC分配接口或将接口从MDC中删除时，需要按组分配或删除，而不能只分配或删除这组接口中的部分接口。接口的分组方式如 [表 1-3](#) 所示：

表1-3 业务板接口分组

业务板类型	接口分组
FC、FX	从接口编号1开始，按编号由小到大的顺序： <ul style="list-style-type: none"> • 1G SFP、10G SFP+、10/100/1000Base-T、10GBase-T 接口业务板：每 24 个接口为一组 • 40G QSFP+接口业务板：每 6 个接口为一组 • 100G CXP、100G CFP2 接口业务板：每 2 个接口为一组
FE	同一单板上的所有接口为一组

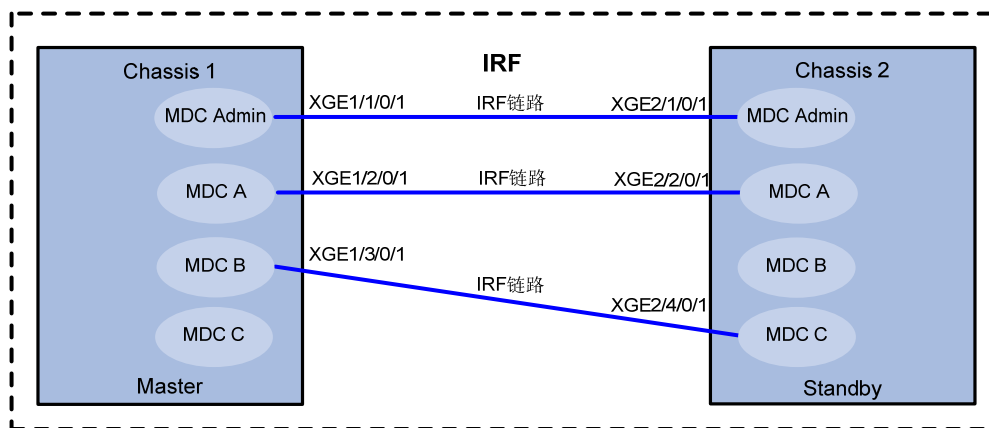
- 一个物理接口只能属于一台 MDC。
- 同一业务板上的接口可以按组分配给不同的 MDC，此时，需要将该业务板也分配给这些 MDC。
- 接口只能从缺省 MDC 分配到非缺省 MDC。如果待分配接口已被分配给非缺省 MDC，则需要先将接口从该 MDC 中删除。

(2) IRF 模式下的其它注意事项

- 每个机框上至少需要配置两块支持 IRF 物理端口的业务板，并建议在不同业务板上建立至少两条 IRF 链路。MDC 配置过程中需要注意确保成员设备间至少有一条 IRF 链路正常通信，避免 IRF 分裂。
- 如果需要将非缺省 MDC 中 IRF 物理端口所在的单板拔出，请在拔出单板前，先取消该单板上的 IRF 物理端口配置，并且保存当前配置后再拔出单板。
- IRF 链路可以属于缺省 MDC（名称为 Admin），也可以属于非缺省 MDC。为了增强链路的可靠性，建议在不同接口板上建立至少两条 IRF 链路。

- 一条 IRF 链路可以属于一个 MDC，也可以属于不同 MDC。即连接一条 IRF 链路的两个 IRF 物理端口可以属于同一个 MDC 也可以属于不同 MDC。

图1-2 MDC 中 IRF 链路连接示意图



- 将 IRF 物理端口分配给其它 MDC 或者从当前 MDC 中删除时，需要按以下步骤：
 - 将端口 **shutdown**
 - 执行 **undo port group interface** 命令取消 IRF 端口和 IRF 物理端口的绑定关系。有关 **undo port group interface** 命令的详细使用，请参见“虚拟化技术命令参考”中的“IRF”
 - 执行分配或者删除操作
 - 执行 **save** 命令保存当前配置文件
 - 将 IRF 物理端口分配给其它 MDC 或者从当前 MDC 中删除的过程会导致该 IRF 物理端口上的 IRF 配置丢失、IRF 链路断开。因此请确保进行此操作时成员设备间还有其它 IRF 链路正常通信，否则会导致 IRF 分裂。
- (3) 给 MDC 分配接口后的注意事项
- 将物理接口分配给 MDC 或者从 MDC 中删除时，该接口下的所有配置都会恢复到缺省情况。
 - 请确保不同用户对同一个接口的操作时序，在一个用户分配或删除接口时及时通知其他用户，让其停止配置该接口，否则可能导致接口达不到非缺省 MDC 用户预期的配置效果。
 - 物理接口分配给 MDC 后，需要登录该 MDC 后，才能对接口下的参数进行配置。
 - 将物理接口从 MDC 中删除时，需要先执行 **undo location** 命令取消该 MDC 以及缺省 MDC 对接口所属业务板的使用权限。

2. 配置步骤

请参照以下步骤为 MDC 分配物理接口和业务板：

- (1) 为每个 MDC 规划物理接口和业务板，确认接口数量、接口编号以及接口所在业务板的槽位号。
- (2) 在缺省 MDC 视图下执行 **display this** 命令，如果显示信息中没有该业务板相关的 **undo location** 命令，则需要执行 **undo location** 命令取消缺省 MDC 对该业务板的使用权限。
- (3) 在当前 MDC 视图下执行 **display this** 命令，如果显示信息中有该业务板相关的 **location** 命令，则需要执行 **undo location** 命令取消当前 MDC 对该业务板的使用权限。
- (4) 将接口分配给 MDC。

(5) 将接口所属业务板的使用权限分配给该 MDC。

表1-4 为 MDC 分配物理接口和业务板

操作	命令	说明
进入系统视图	system-view	-
进入MDC视图	mdc mdc-name [id mdc-id]	-
为MDC分配物理接口	allocate interface interface-list	缺省情况下，物理设备上的所有物理接口都属于缺省MDC，非缺省MDC下没有接口 执行本命令可以给同一MDC分配多个接口
将业务板的使用权限分配给MDC（独立运行模式）	location slot slot-number	缺省情况下，缺省MDC可以使用物理设备上的所有业务板，非缺省MDC只有主控板没有业务板 如果没有给MDC分配某业务板上的接口，请不要给这个MDC分配该业务板
将业务板的使用权限分配给MDC（IRF模式）	location chassis chassis-number slot slot-number	缺省情况下，缺省MDC可以使用物理设备上的所有业务板，非缺省MDC只有主控板没有业务板 如果没有给MDC分配某业务板上的接口，请不要给这个MDC分配该业务板

1.5.2 设置MDC的CPU权重

如果将一块单板的使用权分配给了多个 MDC，那么这些 MDC 共享该单板的 CPU 资源。为了防止一个 MDC 过多的占用该单板的 CPU，而导致其他 MDC 无法运行，需要限制 MDC 对单板 CPU 的使用。

系统根据 MDC 的 CPU 权重占有所有 MDC CPU 权重总和的比率来确定该 MDC 的任务在一个 CPU 上占用时间的比率。比如当 3 个 MDC 的 CPU 权重分别为 10、10、5，则系统为第一个 MDC 分配的 CPU 时间和为第二个 MDC 分配的时间近似都是第三个 MDC 的 CPU 时间的 2 倍，此时和配置权重值分别为 2、2、1 效果一致。

给 MDC 配置的 CPU 权重对其所有具有使用权限的单板有效。“所有具有使用权限的单板”包括主控板和通过 **location** 命令分配的接口板。

表1-5 设置 MDC 的 CPU 权重

操作	命令	说明
进入系统视图	system-view	-
进入MDC视图	mdc mdc-name [id mdc-id]	-
设置MDC的CPU权重	limit-resource cpu weight weight-value	缺省情况下，缺省MDC在所有单板上的CPU权重均为10（不能修改）。非缺省MDC在所有具有使用权限的单板上的CPU权重均为10

1.5.3 为MDC分配内存空间

MDC 创建后，这些 MDC 将共享主控板的内存空间。为了防止一个 MDC 过多的占用内存，而导致其他 MDC 无法正常运行业务，需要限制 MDC 对内存的使用。

执行 **limit-resource memory** 命令前，请使用 **display mdc resource** 命令查看设备当前的内存分配。配置的内存值不宜过小，至少需要保证该 MDC 的正常启动并运行。

表1-6 为 MDC 分配内存空间

操作	命令	说明
进入系统视图	system-view	-
进入MDC视图	mdc mdc-name [id mdc-id]	-
配置MDC可使用的内存空间上限（独立运行模式）	limit-resource memory slot slot-number [cpu cpu-number] ratio limit-ratio	缺省情况下，所有MDC共享主控板的所有内存空间，每个MDC可使用的内存空间上限为空闲内存空间值
配置MDC可使用的内存空间上限（IRF模式）	limit-resource memory chassis chassis-number slot slot-number [cpu cpu-number] ratio limit-ratio	缺省情况下，所有MDC共享主控板的所有内存空间，每个MDC可使用的内存空间上限为空闲内存空间值

1.6 启动MDC

创建 MDC 相当于构造了一台新的物理设备。创建后必须要启动 MDC，MDC 才能正常运行，相当于给设备上电启动。

执行 **mdc start** 命令后，MDC 开始启动，并进入自动配置过程。请使用 **switchto mdc** 命令登录该 MDC，确保该 MDC 的自动配置过程已经结束。如果网络中没有部署自动配置环境，请根据设备提示信息，使用<Ctrl+D>快捷键结束自动配置过程。关于自动配置的详细介绍请参见“基础配置指导”中的“自动配置”。

表1-7 启动 MDC

操作	命令	说明
进入系统视图	system-view	-
进入MDC视图	mdc mdc-name [id mdc-id]	-
启动MDC	mdc start	-

1.7 访问和管理MDC

MDC 创建并启动成功后，就可以当作一台独立设备使用了。如果用户当前已经登录了缺省 MDC，可以通过以下方式来登录非缺省 MDC，借用用户和缺省 MDC 之间的物理连接和会话连接来访问 MDC。此时，命令行视图将从缺省 MDC 的系统视图切换到指定 MDC 的用户视图。

表1-8 登录 MDC

操作	命令	说明
进入系统视图	system-view	-
登录MDC	switchto mdc mdc-name	只有MDC处于active状态时,才允许使用该命令来登录MDC

通过以上方式登录 MDC 后,可以给 MDC 的管理以太网接口配置 IP 地址,或者在 MDC 上创建 VLAN 接口并配置 IP 地址,并确保访问终端和 MDC 之间路由可达。这样,访问终端就可以直接使用 Telnet 或 SSH 等方式登录 MDC 了。

用户登录 MDC 后,可以在 MDC 的用户视图执行 **switchback** 或 **quit** 命令来退出登录。此时,命令视图将从当前 MDC 的用户视图返回到缺省 MDC 的系统视图。

1.8 MDC显示和维护

在完成 MDC 相关配置后,在任意视图下执行 **display** 命令,均可以显示配置后 MDC 的运行情况,通过查看显示信息,来验证配置的效果。

[表 1-9](#)所列是缺省MDC上可以执行的MDC显示命令;[表 1-10](#)所列是非缺省MDC上可以执行的MDC显示命令。

表1-9 缺省 MDC 上的显示和维护

操作	命令
显示MDC的相关信息	display mdc [name mdc-name]
显示MDC的接口列表	display mdc [name mdc-name] interface
显示MDC对CPU/内存资源的使用情况 (独立运行模式)	display mdc [name mdc-name] resource [cpu memory] [slot slot-number [cpu cpu-number]]
显示MDC对CPU/内存资源的使用情况 (IRF模式)	display mdc [name mdc-name] resource [cpu memory] [chassis chassis-number slot slot-number [cpu cpu-number]]

表1-10 非缺省 MDC 上的显示和维护

操作	命令
显示MDC的相关信息	display mdc
显示MDC的接口列表	display mdc interface
显示MDC对CPU/内存资源的使用情况 (独立运行模式)	display mdc resource [cpu memory] [slot slot-number [cpu cpu-number]]
显示MDC对CPU/内存资源的使用情况 (IRF模式)	display mdc resource [cpu memory] [chassis chassis-number slot slot-number [cpu cpu-number]]

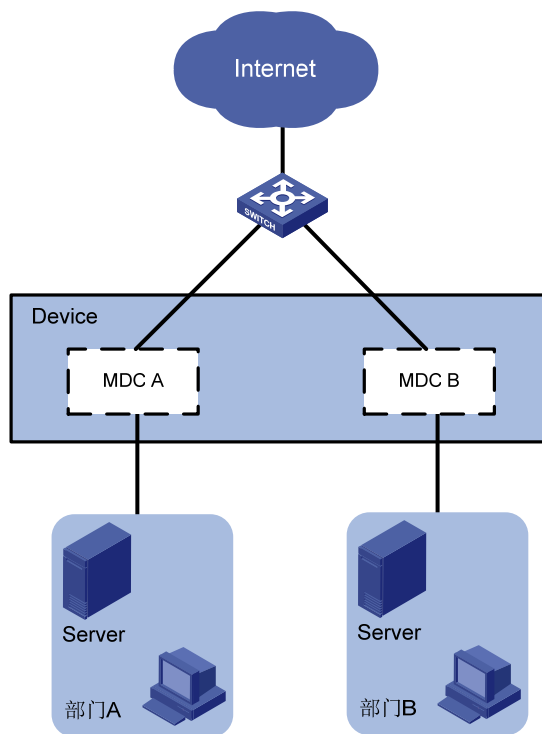
1.9 MDC典型配置举例（独立运行模式）

1. 组网需求

某公司有一台 12500-X 系列交换机，现要求将这台设备虚拟成两台独立的设备：MDC A 和 MDC B，分给该公司的 A、B 两个不同部门用作接入公司网络的网关。使得各部门可以独享设备的表项资源，并且可以根据各部门的业务需求分配设备的内存、CPU 资源，未分配的使用设备缺省值。

2. 组网图

图1-3 MDC 典型配置组网图



3. 配置步骤

(1) 创建 MDC A 和 MDC B

创建 MDC A，供部门 A 使用。

```
<Device> system-view
[Device] mdc MDCA
It will take some time to create MDC...
MDC created successfully.
[Device-mdc-2-MDCA] quit
```

创建 MDC B，供部门 B 使用。

```
[Device] mdc MDCB
It will take some time to create MDC...
MDC created successfully.
[Device-mdc-3-MDCB] quit
```

取消缺省 MDC 对 2 号业务板的使用权。

```
[Device] mdc Admin
```

```

[Device-mdc-1-Admin] undo location slot 2
The configuration associated with the specified slot of MDC will be lost. Continue? [Y/N]:y
[Device-mdc-1-Admin] quit
# 为 MDC A 分配接口 Ten-GigabitEthernet2/0/1~Ten-GigabitEthernet2/0/24。
[Device] mdc MDCA
[Device-mdc-2-MDCA] allocate interface ten-gigabitethernet 2/0/1 to ten-gigabitethernet
2/0/24
Configuration of the interfaces will be lost. Continue? [Y/N]:y
Execute the location slot command in this view to make the configuration take effect.
# 将 2 号业务板的使用权限分配给 MDC A。
[Device-mdc-2-MDCA] location slot 2
# 配置 MDC A 的 CPU 权重为 5。
[Device-mdc-2-MDCA] limit-resource cpu weight 5
# 启动 MDC A。
[Device-mdc-2-MDCA] mdc start
It will take some time to start MDC...
MDC started successfully.
[Device-mdc-2-MDCA] quit
# 为 MDC B 分配接口 Ten-GigabitEthernet2/0/25~Ten-GigabitEthernet2/0/48。
[Device] mdc MDCA
[Device-mdc-3-MDCB] allocate interface ten-gigabitethernet 2/0/25 to ten-gigabitethernet
2/0/48
Configuration of the interfaces will be lost. Continue? [Y/N]:y
Execute the location slot command in this view to make the configuration take effect.
# 将 2 号业务板的使用权限分配给 MDC B。
[Device-mdc-3-MDCB] location slot 2
# 配置 MDC B 的 CPU 权重为 5。
[Device-mdc-3-MDCB] limit-resource cpu weight 5
# 启动 MDC B。
[Device-mdc-3-MDCB] mdc start
It will take some time to start MDC...
MDC started successfully.
[Device-mdc-3-MDCB] quit
(2) 登录 MDC A，配置 MDC A 的管理以太网接口。
# 切换到 MDC A，并按照提示信息键入组合键 Ctrl+D 终止 MDC 的自动配置过程，进入 MDC A 的
命令行界面进行配置。
[Device] switcho mdc MDCA
*****
* Copyright (c) 2004-2017 New H3C Technologies Co., Ltd. All rights reserved.*
* Without the owner's prior written consent, *
* no decompiling or reverse-engineering shall be allowed. *
*****

Automatic configuration is running, press CTRL_D to break or press CTRL_B to
switch back to the default MDC.
<Device> system-view

```

将 MDC A 的名称修改为 MDCA，以便和其它 MDC 区别。

```
[Device] sysname MDCA
```

配置 MDC A 的网管口的 IP 地址、使能 Telnet 服务等基础配置，供用户远程登录 MDC A。

```
[MDCA] interface m-gigabitethernet 0/0/0
[MDCA-M-GigabitEthernet0/0/0] ip address 192.168.1.251 24
[MDCA-M-GigabitEthernet0/0/0] quit
[MDCA] telnet server enable
[MDCA] user-interface vty 0 63
[MDCA-line-vty0-63] authentication-mode none
[MDCA-line-vty0-63] user-role mdc-admin
```

从 MDC A 返回缺省 MDC。

```
[MDCA-line-vty0-63] return
<MDCA> switchback
```

```
[Device]
```

(3) 登录 MDC B，配置 MDC B 的管理以太网接口。

```
[Device] switchto mdc MDCB
```

```
*****
* Copyright (c) 2004-2017 New H3C Technologies Co., Ltd. All rights reserved.*
* Without the owner's prior written consent, *
* no decompiling or reverse-engineering shall be allowed. *
*****
```

Automatic configuration is running, press CTRL_D to break or press CTRL_B to switch back to the default MDC.

```
<Device> system-view
```

将 MDC B 的名称修改为 MDCB，以便和其它 MDC 区别。

```
[Device] sysname MDCB
```

配置 MDC B 的网管口的 IP 地址、使能 Telnet 服务等基础配置，供用户远程登录 MDC B。

```
[MDCB] interface m-gigabitethernet 0/0/0
[MDCB-M-GigabitEthernet0/0/0] ip address 192.168.2.252 24
[MDCB-M-GigabitEthernet0/0/0] quit
[MDCB] telnet server enable
[MDCB] user-interface vty 0 63
[MDCB-line-vty0-63] authentication-mode none
[MDCB-line-vty0-63] user-role mdc-admin
```

从 MDC B 返回缺省 MDC。

```
[MDCB-line-vty0-63] return
<MDCB> switchback
```

```
[Device]
```

4. 结果验证

(1) 查看 MDC 是否存在并且运转正常。（此时，Device 上应该有三台处于正常工作 active 状态的 MDC）

```
<Device> display mdc
```

ID	Name	Status
1	Admin	active
2	MDCA	active

3 MDCB active

(2) 模拟部门 A 的管理用户登录到 MDC A，可以查看本设备的当前配置。

```
C:\> telnet 192.168.1.251
*****
* Copyright (c) 2004-2017 New H3C Technologies Co., Ltd. All rights reserved.*
* Without the owner's prior written consent,                               *
* no decompiling or reverse-engineering shall be allowed.                   *
*****

<MDCA> display current-configuration
.....配置文件信息略.....
```

1.10 MDC典型配置举例（IRF模式）

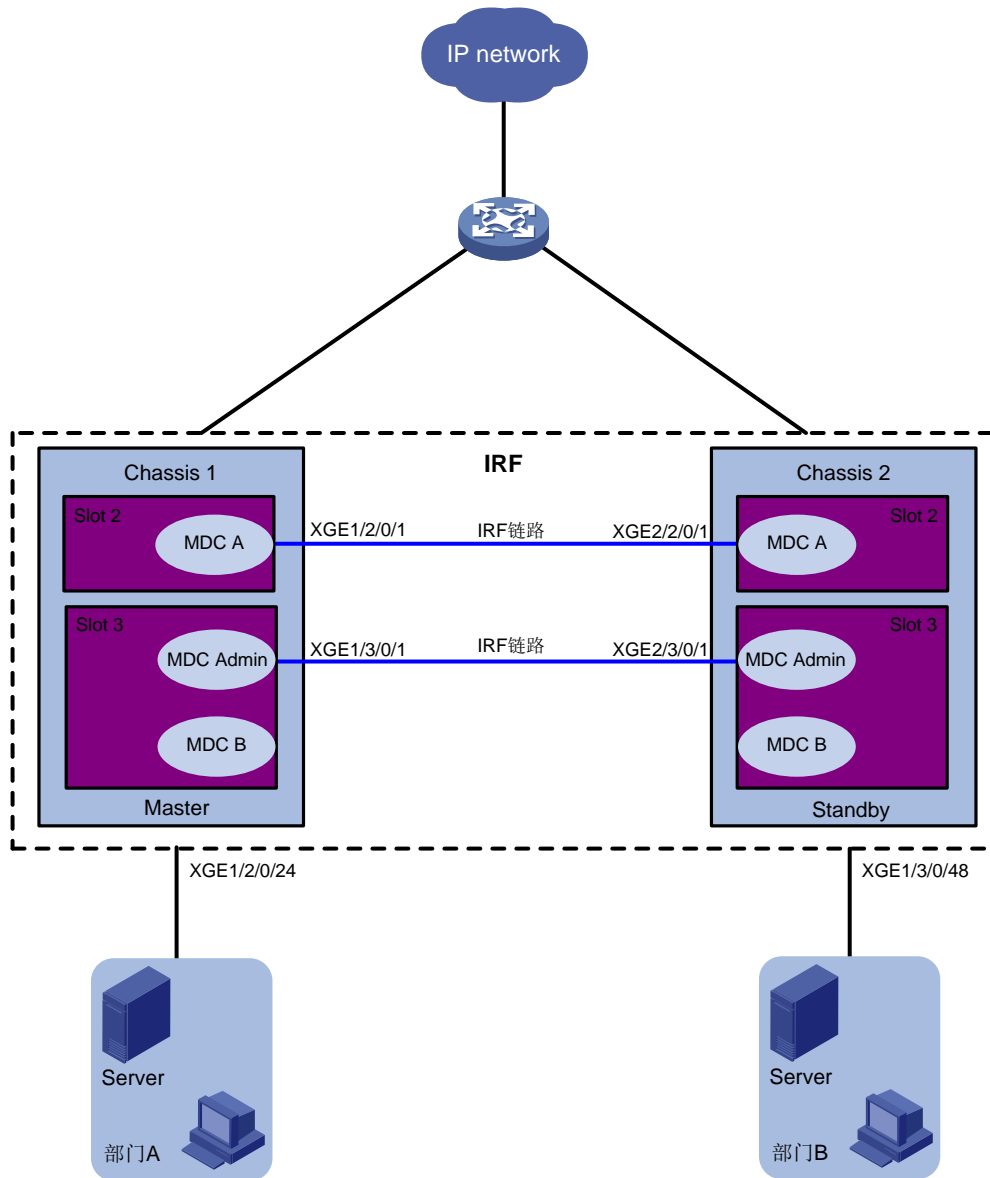
1. 组网需求

如 [图 1-4](#) 所示，某公司有两台 12500-X 系列交换机已经组成了 IRF。Master 和 Standby 设备上各有业务板 2 和 3，每个业务板有 48 个万兆以太网口接口。Master 使用 IRF 端口 1 和 Standby 的 IRF 端口 2 通过两条物理链路相连，Master 的 IRF 端口 1 绑定了物理接口 Ten-GigabitEthernet1/2/0/1 和 Ten-GigabitEthernet1/3/0/1，Standby 的 IRF 端口 2 绑定了物理接口 Ten-GigabitEthernet2/2/0/1 和 Ten-GigabitEthernet2/3/0/1。

现要求创建两台独立的虚拟设备：MDC A 和 MDC B，分给该公司的 A、B 两个部门用作接入公司网络的网关。使得各部门可以独享设备的表项资源，并且可以根据各部门的业务需求分配设备的内存、CPU 资源，Ten-GigabitEthernet1/2/0/1 ~ Ten-GigabitEthernet1/2/0/24 和 Ten-GigabitEthernet2/2/0/1 ~ Ten-GigabitEthernet2/2/0/24 分配给 MDC A，Ten-GigabitEthernet1/3/0/25 ~ Ten-GigabitEthernet1/3/0/48 和 Ten-GigabitEthernet2/3/0/25 ~ Ten-GigabitEthernet2/3/0/48 分配给 MDC B。

2. 组网图

图1-4 IRF+MDC 典型配置组网图



3. 配置步骤

(1) 分别取消 Ten-GigabitEthernet1/2/0/1 和 IRF 端口 1/1、Ten-GigabitEthernet2/2/0/1 和 IRF 端口 2/2 的绑定。

关闭接口 Ten-GigabitEthernet1/2/0/1 和 Ten-GigabitEthernet2/2/0/1。

```
<IRF> system-view
[IRF] interface range ten-gigabitethernet 1/2/0/1 ten-gigabitethernet 2/2/0/1
[IRF-if-range] shutdown
[IRF-if-range] quit
```

取消 Ten-GigabitEthernet1/2/0/1 和 IRF 端口 1/1 的绑定。

```
[IRF] irf-port 1/1
[IRF-irf-port1/1] undo port group interface Ten-GigabitEthernet1/2/0/1
```

```
[IRF-irf-port1/1] quit
# 取消 Ten-GigabitEthernet2/2/0/1 和 IRF 端口 2/2 的绑定。
[IRF] irf-port 2/2
[IRF-irf-port2/2] undo port group interface Ten-GigabitEthernet2/2/0/1
[IRF-irf-port2/2] quit
(2) 创建和配置 MDC A。
# 创建 MDC A，供部门 A 使用。
[IRF] mdc MDCA
It will take some time to create MDC...
MDC created successfully.
[IRF-mdc-2-MDCA] quit
# 取消缺省 MDC 对成员设备 1 的 2 号业务板和成员设备 2 的 2 号业务板的使用权。
[IRF] mdc Admin
[IRF-mdc-1-Admin] undo location chassis 1 slot 2
The configuration associated with the specified slot of MDC will be lost. Continue? [Y/N]:y
[IRF-mdc-1-Admin] undo location chassis 2 slot 2
The configuration associated with the specified slot of MDC will be lost. Continue? [Y/N]:y
[IRF-mdc-1-Admin] quit
# 为 MDC A 分配接口 Ten-GigabitEthernet1/2/0/1 ~ Ten-GigabitEthernet1/2/0/24 、
Ten-GigabitEthernet2/2/0/1~Ten-GigabitEthernet2/2/0/24。（该业务板上每 24 个接口为一组，这
24 个接口必须同时分配）
[IRF] mdc MDCA
[IRF-mdc-2-MDCA] allocate interface ten-gigabitethernet 1/2/0/1 to ten-gigabitethernet
1/2/0/24
Configuration of the interfaces will be lost. Continue? [Y/N]:y
Execute the location slot command in this view to make the configuration take effect.
[IRF-mdc-2-MDCA] allocate interface ten-gigabitethernet 2/2/0/1 to ten-gigabitethernet
2/2/0/24
Configuration of the interfaces will be lost. Continue? [Y/N]:y
Execute the location slot command in this view to make the configuration take effect.
[IRF-mdc-2-MDCA] quit
# 将成员设备 1 的 2 号业务板和成员设备 2 的 2 号业务板的使用权限分配给 MDC A。
[IRF-mdc-2-MDCA] location chassis 1 slot 2
[IRF-mdc-2-MDCA] location chassis 2 slot 2
# 配置 MDC A 的 CPU 权重为 5。
[IRF-mdc-2-MDCA] limit-resource cpu weight 5
# 启动 MDC A。
[IRF-mdc-2-MDCA] mdc start
It will take some time to start MDC...
MDC started successfully.
[IRF-mdc-2-MDCA] quit
(3) 将 MDC A 中的 Ten-GigabitEthernet1/2/0/1 和 Ten-GigabitEthernet2/2/0/1 配置为 IRF 物理端
口。
# 切换到 MDC A，并按照提示信息键入组合键 Ctrl+D 终止 MDC 的自动配置过程，进入 MDC A 的
命令行界面进行配置。
```

```
[IRF] switchto mdc MDCA
*****
* Copyright (c) 2004-2017 New H3C Technologies Co., Ltd. All rights reserved.*
* Without the owner's prior written consent, *
* no decompiling or reverse-engineering shall be allowed. *
*****
```

Automatic configuration is running, press CTRL_D to break or press CTRL_B to switch back to the default MDC.

```
<IRF> system-view
```

将 MDC A 的名称修改为 MDCA，以便和其它 MDC 区别。

```
[IRF] sysname MDCA
```

关闭接口 Ten-GigabitEthernet1/2/0/1 和 Ten-GigabitEthernet2/2/0/1。

```
[MDCA] interface range ten-gigabitethernet 1/2/0/1 ten-gigabitethernet 2/2/0/1
```

```
[MDCA-if-range] shutdown
```

```
[MDCA-if-range] quit
```

```
[MDCA] quit
```

从 MDC A 返回缺省 MDC。

```
<MDCA> switchback
```

```
[IRF]
```

显示 MDC A 的编号。

```
[IRF] display mdc
```

ID	Name	Status
1	Admin	active
2	MDCA	active

将 MDC A（编号为 2）中的 Ten-GigabitEthernet1/2/0/1 和 IRF 端口 1/1 绑定。

```
[IRF] irf-port 1/1
```

```
[IRF-irf-port1/1] port group mdc 2 interface ten-gigabitethernet 1/2/0/1
```

You must perform the following tasks for a successful IRF setup:

Save the configuration after completing IRF configuration.

Execute the "irf-port-configuration active" command to activate the IRF ports.

```
[IRF-irf-port1/1] quit
```

因为 IRF 已经组建，所以，不需要激活 IRF 端口配置，IRF 端口配置会自动生效。

将 MDC A（编号为 2）中的 Ten-GigabitEthernet2/2/0/1 和 IRF 端口 2/2 绑定。

```
[IRF] irf-port 2/2
```

```
[IRF-irf-port2/2] port group mdc 2 interface ten-gigabitethernet 2/2/0/1
```

You must perform the following tasks for a successful IRF setup:

Save the configuration after completing IRF configuration.

Execute the "irf-port-configuration active" command to activate the IRF ports.

```
[IRF-irf-port2/2] quit
```

因为 IRF 已经组建，所以，不需要激活 IRF 端口配置，IRF 端口配置会自动生效。

切换到 MDC A。

```
[IRF] switchto mdc MDCA
```

```
*****
* Copyright (c) 2004-2017 New H3C Technologies Co., Ltd. All rights reserved.*
```

```
* Without the owner's prior written consent, *
* no decompiling or reverse-engineering shall be allowed. *
*****
```

```
<MDCA> system-view
```

```
# 打开接口 Ten-GigabitEthernet1/2/0/1 和 Ten-GigabitEthernet2/2/0/1。
```

```
[MDCA] interface range ten-gigabitethernet 1/2/0/1 ten-gigabitethernet 2/2/0/1
[MDCA-if-range] undo shutdown
[MDCA-if-range] quit
```

```
# 配置 MDC A 的网管口的 IP 地址、使能 Telnet 服务等基础配置，供用户远程登录 MDC A。
```

```
[MDCA] display interface M-GigabitEthernet brief
Brief information on interfaces in route mode:
Link: ADM - administratively down; Stby - standby
Protocol: (s) - spoofing
Interface          Link Protocol Primary IP      Description
M-GE1/0/0/0        DOWN DOWN      --
M-GE1/0/0/1        DOWN DOWN      --
M-GE1/0/0/2        UP   UP         --
M-GE1/0/0/3        DOWN DOWN      --
```

```
[MDCA] interface m-gigabitethernet 1/0/0/2
[MDCA-M-GigabitEthernet1/0/0/2] ip address 192.168.1.251 24
[MDCA-M-GigabitEthernet1/0/0/2] quit
[MDCA] telnet server enable
[MDCA] user-interface vty 0 63
[MDCA-line-vty0-63] authentication-mode none
[MDCA-line-vty0-63] user-role mdc-admin
[MDCA-line-vty0-63] return
```

```
# 从 MDC A 返回缺省 MDC。
```

```
<MDCA> switchback
```

```
[IRF]
```

```
# 显示 IRF 链路信息，将看到两条 IRF 链路均处于 UP 状态。
```

```
<IRF> display irf link
```

```
Member 1
```

IRF Port	Interface	Status
1	Ten-GigabitEthernet1/2/0/1(MDC2)	UP
	Ten-GigabitEthernet1/3/0/1	UP
2	disable	--

```
Member 2
```

IRF Port	Interface	Status
1	disable	--
2	Ten-GigabitEthernet2/2/0/1(MDC2)	UP
	Ten-GigabitEthernet2/3/0/1	UP



说明

当需要将 IRF 物理端口分配给非缺省 MDC 或者需要在非缺省 MDC 中新建 IRF 链路时，请参照 MDC A 的步骤配置。下面将以 MDC B 为例，实现：IRF 物理端口仍归属缺省 MDC，将非 IRF 物理端口

分配给 MDC B 使用，这种方式配置相对简单，不需要修改 IRF 端口的配置，又能满足 MDC B 的要求。

(4) 创建和配置 MDC B。

创建 MDC B，供部门 B 使用。

```
[IRF] mdc MDCB
It will take some time to create MDC...
```

```
MDC created successfully.
```

```
[IRF-mdc-3-MDCB] quit
```

取消缺省 MDC 对成员设备 1 的 3 号业务板和成员设备 2 的 3 号业务板的使用权。

```
[IRF] mdc Admin
```

```
[IRF-mdc-1-Admin] undo location chassis 1 slot 3
```

```
The configuration associated with the specified slot of MDC will be lost. Continue? [Y/N]:y
```

```
[IRF-mdc-1-Admin] undo location chassis 2 slot 3
```

```
The configuration associated with the specified slot of MDC will be lost. Continue? [Y/N]:y
```

```
[IRF-mdc-1-Admin] quit
```

为 MDC B 分配接口 Ten-GigabitEthernet1/3/0/25 ~ Ten-GigabitEthernet1/3/0/48 和 Ten-GigabitEthernet2/3/0/25~Ten-GigabitEthernet2/3/0/48。

```
[IRF] mdc MDCB
```

```
[IRF-mdc-3-MDCB] allocate interface ten-gigabitethernet 1/3/0/25 to ten-gigabitethernet 1/3/0/48
```

```
Configuration of the interfaces will be lost. Continue? [Y/N]:y
```

```
Execute the location slot command in this view to make the configuration take effect.
```

```
[IRF-mdc-3-MDCB] allocate interface ten-gigabitethernet 2/3/0/25 to ten-gigabitethernet 2/3/0/48
```

```
Configuration of the interfaces will be lost. Continue? [Y/N]:y
```

```
Execute the location slot command in this view to make the configuration take effect.
```

将成员设备 1 的 3 号业务板和成员设备 2 的 3 号业务板的使用权限分配给 MDC B。

```
[IRF-mdc-3-MDCB] location chassis 1 slot 3
```

```
[IRF-mdc-3-MDCB] location chassis 2 slot 3
```

配置 MDC B 的 CPU 权重为 5。

```
[IRF-mdc-3-MDCB] limit-resource cpu weight 5
```

启动 MDC B。

```
[IRF-mdc-3-MDCB] mdc start
```

```
It will take some time to start MDC...
```

```
MDC started successfully.
```

```
[IRF-mdc-3-MDCB] quit
```

恢复缺省 MDC 对成员设备 1 的 3 号业务板和成员设备 2 的 3 号业务板的使用权限。

```
[IRF] mdc Admin
```

```
[IRF-mdc-2-Admin] location chassis 1 slot 3
```

```
[IRF-mdc-2-Admin] location chassis 2 slot 3
```

```
[IRF-mdc-2-Admin] quit
```

切换到 MDC B，并按照提示信息键入组合键 Ctrl+D 终止 MDC 的自动配置过程，进入 MDC B 的命令行界面进行配置。

```
[IRF] swichto mdc MDCB
```

```
*****
* Copyright (c) 2004-2017 New H3C Technologies Co., Ltd. All rights reserved.*
* Without the owner's prior written consent, *
* no decompiling or reverse-engineering shall be allowed. *
*****
```

Automatic configuration is running, press CTRL_D to break or press CTRL_B to switch back to the default MDC.

```
<IRF> system-view
```

将 MDC B 的名称修改为 MDCB，以便和其它 MDC 区别。

```
[IRF] sysname MDCB
```

配置 MDC B 的网管口的 IP 地址、使能 Telnet 服务等基础配置，供用户远程登录 MDC B。

```
[MDCB] display interface M-GigabitEthernet brief
```

```
Brief information on interfaces in route mode:
```

```
Link: ADM - administratively down; Stby - standby
```

```
Protocol: (s) - spoofing
```

Interface	Link	Protocol	Primary IP	Description
M-GE1/0/0/0	DOWN	DOWN	--	
M-GE1/0/0/1	DOWN	DOWN	--	
M-GE1/0/0/2	UP	UP	--	
M-GE1/0/0/3	DOWN	DOWN	--	

```
[MDCB] interface m-gigabitethernet 1/0/0/2
```

```
[MDCB-M-GigabitEthernet1/0/0/2] ip address 192.168.2.252 24
```

```
[MDCB-M-GigabitEthernet1/0/0/2] quit
```

```
[MDCB] telnet server enable
```

```
[MDCB] user-interface vty 0 63
```

```
[MDCB-line-vty0-63] authentication-mode none
```

```
[MDCB-line-vty0-63] user-role mdc-admin
```

从 MDC B 返回缺省 MDC。

```
[MDCB-line-vty0-63] return
```

```
<MDCB> switchback
```

```
[IRF]
```

4. 结果验证

(1) 查看 MDC 是否存在并且运转正常。（此时，IRF 上应该有三台处于正常工作 active 状态的 MDC）

```
<IRF> display mdc
```

ID	Name	Status
1	Admin	active
2	MDCA	active
3	MDCB	active

(2) 模拟部门 A 的管理用户登录到 MDC A，可以查看本设备的当前配置。

```
C:\> telnet 192.168.1.251
```

```
*****
* Copyright (c) 2004-2017 New H3C Technologies Co., Ltd. All rights reserved.*
* Without the owner's prior written consent, *
* no decompiling or reverse-engineering shall be allowed. *
*****
```

<MDCA> display current-configuration

.....配置文件信息略.....