



UNIS R9900 系列路由器

虚拟化配置指导

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

资料版本: 6W100-20160510
产品版本: R9900-CMW710-R7353P08

Copyright © 2016 北京紫光恒越网络科技有限公司及其许可者版权所有，保留一切权利。

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

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

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

前言

本配置指导主要介绍了虚拟化技术-IRF。通过 IRF，可以将多台设备虚拟为一台设备在网络中运行，不仅能提高管理效率，还能实现设备间备份，并简化网络拓扑。

前言部分包含如下内容：

- [读者对象](#)
- [本书约定](#)
- [产品配套资料](#)
- [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”。

格 式	意 义
/	多级菜单用“/”隔开。如[文件/新建/文件夹]多级菜单表示[文件]菜单下的[新建]子菜单下的[文件夹]菜单项。

3. 各类标志

本书还采用各种醒目标志来表示在操作过程中应该特别注意的地方，这些标志的意义如下：

 警告	该标志后的注释需给予格外关注，不当的操作可能会对人身造成伤害。
 注意	提醒操作中应注意的事项，不当的操作可能会导致数据丢失或者设备损坏。
 提示	为确保设备配置成功或者正常工作而需要特别关注的操作或信息。
 说明	对操作内容的描述进行必要的补充和说明。
 窍门	配置、操作、或使用设备的技巧、小窍门。

4. 图标约定

本书使用的图标及其含义如下：

	该图标及其相关描述文字代表一般网络设备，如路由器、交换机、防火墙等。
	该图标及其相关描述文字代表一般意义下的路由器，以及其他运行了路由协议的设备。
	该图标及其相关描述文字代表二、三层以太网交换机，以及运行了二层协议的设备。
	该图标及其相关描述文字代表无线控制器、无线控制器业务板和有线无线一体化交换机的无线控制引擎设备。
	该图标及其相关描述文字代表无线接入点设备。
	该图标及其相关描述文字代表无线终结单元。
	该图标及其相关描述文字代表无线终结者。
	该图标及其相关描述文字代表无线Mesh设备。
	该图标代表发散的无线射频信号。
	该图标代表点到点的无线射频信号。
	该图标及其相关描述文字代表防火墙、UTM、多业务安全网关、负载均衡等安全设备。



该图标及其相关描述文字代表防火墙插卡、负载均衡插卡、NetStream插卡、SSL VPN插卡、IPS插卡、ACG插卡等安全插卡。

5. 端口编号示例约定

本手册中出现的端口编号仅作参考，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

紫光恒越 R9900 系列路由器的配套资料包括如下部分：

大类	资料名称	内容介绍
硬件描述与安装	安全兼容性手册	列出产品的兼容性声明，并对兼容性和安全的细节进行说明
	快速入门	指导您对设备进行初始安装、配置，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
业务配置	配置指导	帮助您掌握设备软件功能的配置方法及配置步骤
	命令参考	详细介绍设备的命令，相当于命令字典，方便您查阅各个命令的功能
运行维护	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

技术支持

用户支持邮箱：zgsm_service@thunis.com

技术支持热线电话：400-910-9998（手机、固话均可拨打）

网址：<http://www.unis-hy.com>

资料意见反馈

如果您在使用过程中发现产品资料的任何问题，可以通过以下方式反馈：

E-mail：zgsm_info@thunis.com

感谢您的反馈，让我们做得更好！

目 录

1 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	1-3
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
1.3 配置限制和指导	1-10
1.4 IRF配置任务简介	1-10
1.5 独立运行模式下预配置IRF	1-12
1.5.1 配置成员编号	1-12
1.5.2 配置成员优先级	1-12
1.5.3 配置IRF端口	1-12
1.6 将当前配置保存到设备的下次启动配置文件	1-13
1.7 切换到IRF模式	1-13
1.8 访问IRF	1-14
1.9 IRF模式下配置IRF	1-14
1.9.1 配置成员编号	1-14
1.9.2 配置成员优先级	1-15
1.9.3 配置IRF端口	1-15
1.9.4 使能IRF合并自动重启功能	1-16
1.9.5 配置成员设备的描述信息	1-16
1.9.6 配置IRF链路的负载分担类型	1-17
1.9.7 配置IRF的桥MAC保留时间	1-18
1.9.8 配置IRF链路down延迟上报功能	1-18
1.9.9 MAD配置	1-19
1.10 快速恢复IRF配置	1-25
1.10.1 配置准备	1-25
1.10.2 配置步骤	1-25
1.11 IRF显示和维护	1-26

1.12 IRF典型配置举例.....	1-27
1.12.1 在R9900-08 或R9900-12 上组建IRF典型配置举例（LACP MAD检测方式）	1-27
1.12.2 组建IRF典型配置举例（BFD MAD检测方式）	1-30
1.12.3 将成员设备从IRF模式恢复到独立运行模式配置举例.....	1-32

1 IRF

1.1 IRF简介

IRF（Intelligent Resilient Framework，智能弹性架构）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

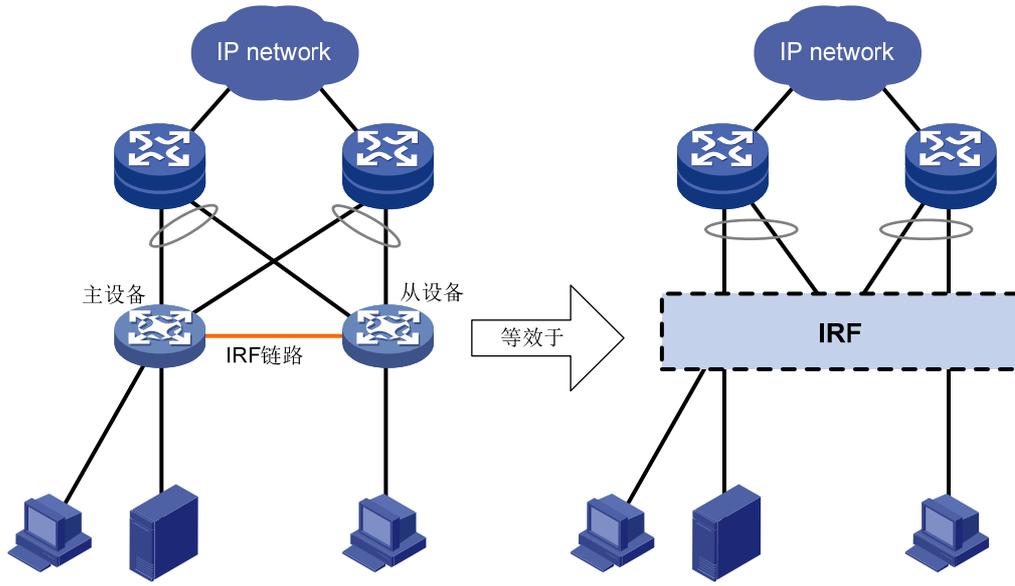
IRF 主要具有以下优点：

- 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

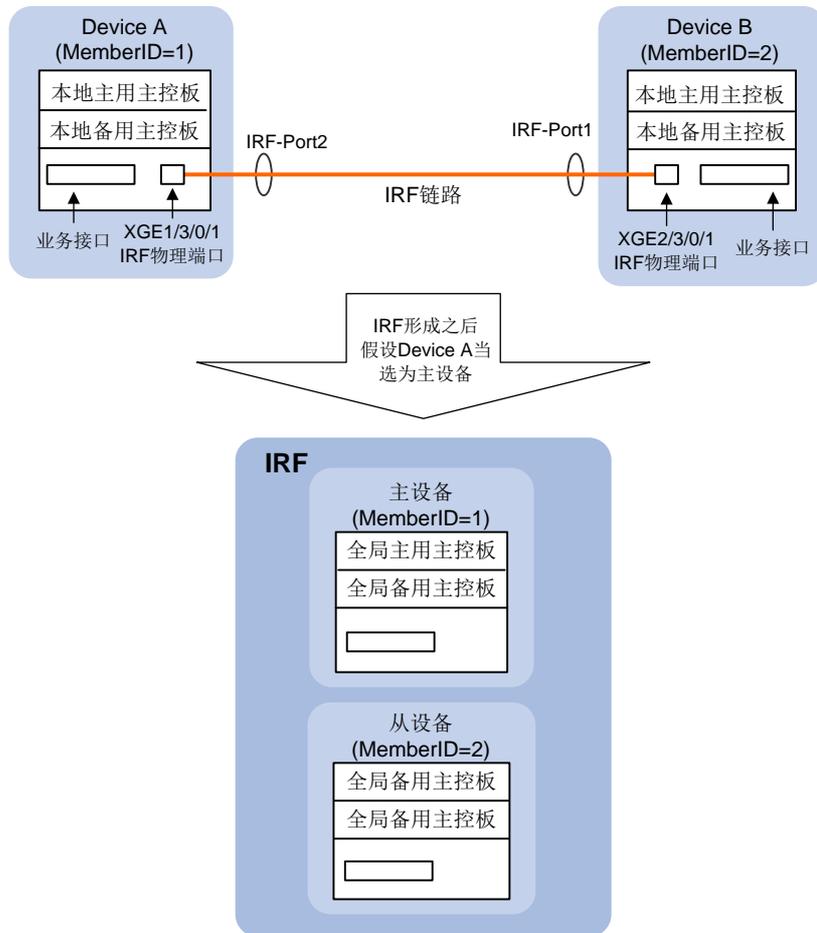
如 [图 1-1](#) 所示，主设备和从设备组成 IRF，对上、下层设备来说，它们就是一台设备——IRF。

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。

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
本地主用主控板	成员设备的主用主控板，负责管理本台设备，是成员设备的必备硬件
本地备用主控板	成员设备的备用主控板，是本地主用主控板的备份，是成员设备的可选硬件
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用主控板
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。

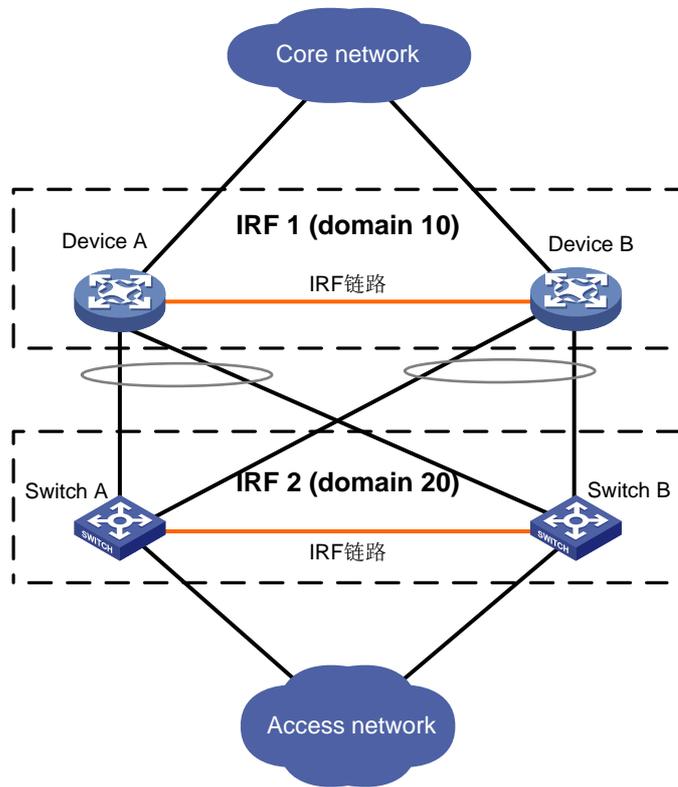
通常情况下，以太网光接口负责向网络中转发业务报文，将它们与 IRF 端口绑定后就作为 IRF 物理端口，可转发的报文包括 IRF 相关协商报文以及需要跨成员设备转发的业务报文。

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。

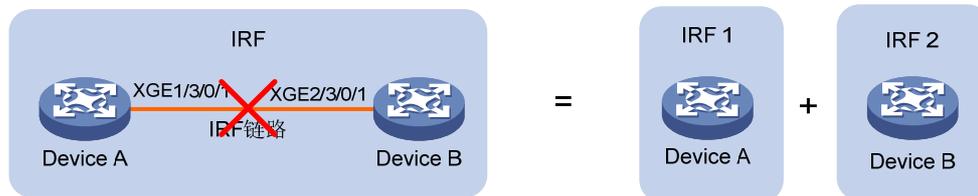
图1-4 IRF 合并示意图



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

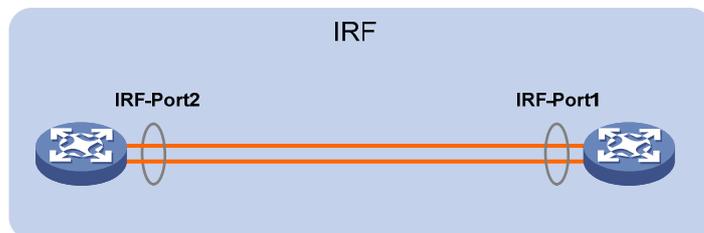
要形成一个 IRF，需要先连接成员设备的 IRF 物理端口。设备支持使用 10GE 二层以太网光接口作为 IRF 物理端口，不同类型的光接口则需要使用不同类型的光纤或电缆进行连接，可以将距离很远的物理设备连接组成 IRF，使得应用更加灵活。

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



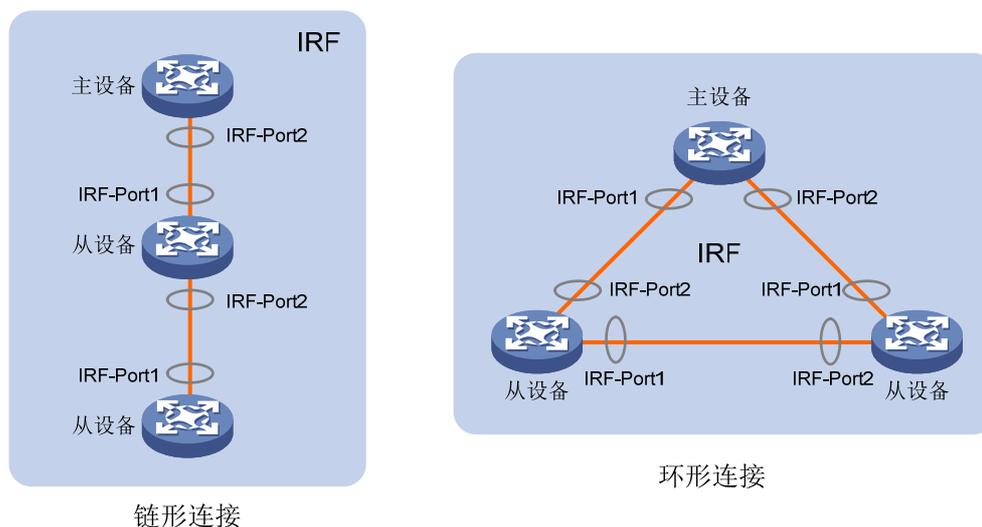
3. 连接拓扑

IRF的连接拓扑有两种：链形连接和环形连接，如 [图 1-7](#)所示。

- 链形连接对成员设备的物理位置要求比环形连接低，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。

目前，路由器仅支持链型连接。

图1-7 IRF 连接拓扑示意图



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用主控板进行管理，在本地记录自己已知的拓扑信息。设备刚启动时，本地主用主控板只记录了自身的拓扑信息。当 IRF 端口状态变为 up 后，本地主用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 不会因为有了新的成员设备/主控板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 [1.9.4 使能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 GigabitEthernet3/0/1）；加入 IRF 后，接口编号会变为四维，第一维表示成员编号（如 GigabitEthernet2/3/0/1）。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。比如设备上只有一块主用主控板，配置的成员编号为 2，此时插入一块成员编号是 3 的备用主控板，则该设备的成员编号仍然为 2，并会将备用主控板上保存的成员编号同步为 2。

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. 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。

(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正常工作状态的 IRF。

- 对于 LACP MAD 检测，冲突处理机制会直接让主设备成员编号小的 IRF 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 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。

- 如果出现故障的是继续正常工作的 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

1. 组建IRF时的注意事项

- R9900 路由器之间可以建立 IRF，但不能与其他路由器产品建立 IRF，且 IRF 中的成员设备最多为 2 台。
- 路由器上组建 IRF 必须使用 SPC 单板上的 10GE 以太网光接口作为 IRF 物理端口。
- 建立 IRF 前，请确保组成 IRF 的 R9900 使用的主控板丝印完全相同，否则不能形成 IRF。
- 建立 IRF 前，请确保组成 IRF 的 R9900 使用的是 B 类交换网板或 D 类交换网板。

2. IRF形成后的配置限制和指导

- 以太网接口作为 IRF 物理端口与 IRF 端口绑定后，只支持 **shutdown** 和 **description** 命令，这些命令的详细介绍，请参见“接口管理命令参考”中的“以太网接口”。
- 因为 LACP MAD 和 BFD MAD 冲突处理的原则不同，请不要同时配置。
- 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。在 BFD MAD 检测组网中，IRF 域编号为可选配置。
- 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 恢复到正常工作状态后，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
	使能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链路的负载分担类型	可选	1.9.6
	配置IRF的桥MAC保留时间	可选	1.9.7
	配置IRF链路down延迟上报功能	可选	1.9.8
MAD配置		可选	1.9.9
快速恢复IRF配置		可选	1.10

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 <i>member-id</i>	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 <i>priority</i>	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。

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），并且 IRF 物理端口下只能配置 **shutdown** 和 **description** 命令，有关这些命令的详细介绍，请参见“接口管理命令参考”中的“以太网接口”。

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>	-
切换以太网接口工作模式为二层模式	port link-mode bridge	缺省情况下，以太网接口工作在三层模式
退回系统视图	quit	-

操作	命令	说明
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 <i>port-number</i>	缺省情况下,设备上没有创建IRF端口
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 }]	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。最多可以绑定的8个IRF物理端口,当绑定的物理端口数达到上限时,该命令将执行失败

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

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** 的转换等。

表1-7 切换到IRF模式

操作	命令	说明
进入系统视图	system-view	-

操作	命令	说明
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 AUX 或者 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 member-id renumber new-member-id	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

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

两台设备切换到 IRF 模式后，创建各自的 IRF 端口，并将 IRF 端口与各自的物理端口绑定，最后用 IRF 线缆分别连接到两台设备的 IRF 物理端口，设备的 IRF 功能才能生效。一台成员设备上的 IRF-Port1 端口只能和另一台成员设备 IRF-Port2 端口相连。

需要注意的是：以太网接口作为 IRF 物理端口与 IRF 端口绑定后，只支持 **shutdown** 和 **description** 命令，这些命令的详细介绍，请参见“接口管理命令参考”中的“以太网接口”。

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>	-
切换以太网接口工作模式为二层模式	port link-mode bridge	缺省情况下，以太网接口工作在三层模式
关闭接口	shutdown	如果允许关闭当前端口，则直接在该接口视图下执行 shutdown 命令即可；如果不能关闭该端口，请根据系统提示信息关闭该端口直连的邻居设备上的端口
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>	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。最多可以绑定8个IRF物理端口，当绑定的物理端口数达到上限时，该命令将执行失败
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	-
退回系统视图	quit	-

操作	命令	说明
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 使能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：

- 如果没有使能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。
- 如果使能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 使能IRF合并自动重启功能

操作	命令	说明
进入系统视图	system-view	-
使能IRF合并自动重启功能	irf auto-merge enable	缺省情况下，没有使能IRF合并自动重启功能，即两台IRF合并时，竞选失败方不会自动重启，需要用户手工重启后才能完成合并 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上都使能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时可配置成员设备的描述信息进行标识。例如当成员设备的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 member-id description text	缺省情况下，成员设备没有描述信息

1.9.6 配置IRF链路的负载分担类型



在配置负载分担模式前，IRF 端口必须至少和一个 IRF 物理端口绑定。否则，负载分担模式将配置失败。

当 IRF 端口与多个 IRF 物理端口绑定时，成员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。用户既可以指定系统按照报文携带的 IP 地址、MAC 地址等信息之一或其组合来选择所采用的负载分担类型，也可以指定系统按照报文类型（如二层、IPv4、IPv6 等）自动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）配置的方式设置 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。

表1-13 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 global load-sharing mode { destination-ip destination-mac ingress-port source-ip source-mac } *	缺省情况下，按报文的入端口实现负载分担 多次执行该命令配置不同负载分担模式时，以最新的配置为准

表1-14 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 <i>member-id/port-number</i>	-
配置IRF链路的负载分担模式	irf-port load-sharing mode { destination-ip destination-mac ingress-port source-ip source-mac } *	缺省情况下，按报文的入端口实现负载分担 多次执行该命令配置不同负载分担模式时，以最新的配置为准

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。
- 如果两个 IRF 的桥 MAC 相同，则它们不能合并为一个 IRF。

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。通常情况下使用主设备的桥 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 6 分钟，则当主设备离开 IRF 时，IRF 桥 MAC 在 6 分钟内保持不变；如果 6 分钟后主设备没有回到 IRF，则使用新选举的主设备的桥 MAC 作为 IRF 桥 MAC。该配置适用于主设备短时间内离开又回到 IRF 的情况（比如主设备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 保留时间为永久，则不管主设备是否离开 IRF，IRF 桥 MAC 始终保持不变。
- 如果配置了 IRF 桥 MAC 不保留，则当主设备离开 IRF 时，系统会立即使用新选举的主设备的桥 MAC 做 IRF 桥 MAC。
- 如果在 IRF 中需要配置 PPPoE 功能，请先修改 IRF 的桥 MAC 地址保留时间为永久保留，否则，当 IRF 的桥 MAC 发生变化时，已在线 PPPoE 用户会直接下线。关于 PPPoE 的详细介绍，请参见“用户接入配置指导”中的“PPPoE”。

表1-15 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	三者选其一 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的保留时间为6分钟	irf mac-address persistent timer	
配置IRF的桥MAC不保留，会立即变化	undo irf mac-address persistent	

1.9.8 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1秒 在IRF环境中使用BFD功能时，请保证IRF链路down延迟上报时间小于BFD的超时时间，关于BFD功能的介绍，请参见“可靠性配置指导”中的“BFD”

1.9.9 MAD配置

设备支持两种 MAD 检测方式：LACP MAD 检测和 BFD MAD 检测。这两种 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UNIS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连接，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. 配置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）。

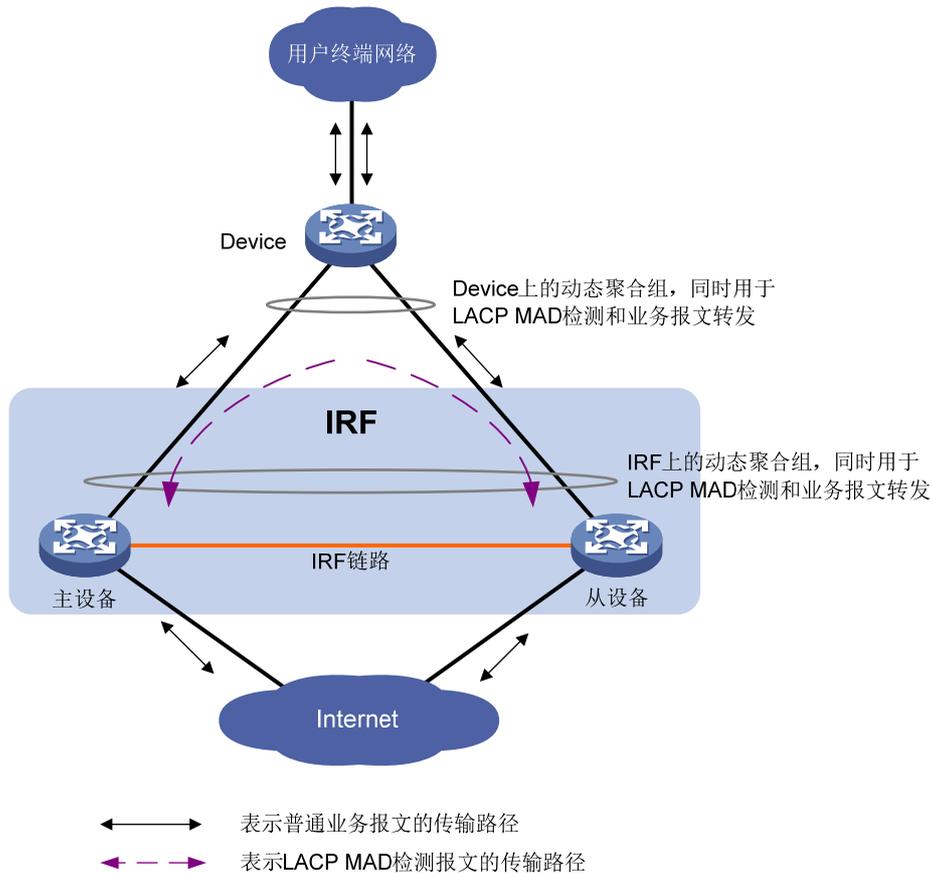
使能 LACP MAD 检测后，成员设备通过 LACP 协议报文和其它成员设备交互 DomainID 和 ActiveID 信息。

- 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UNIS设备作为中间设备。通常采用如 图 1-8 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1-8 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:

- 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1-18 配置 LACP MAD 检测

操作	命令	说明
进入系统视图	system-view	-

操作		命令	说明
配置IRF域编号		irf domain <i>domain-id</i>	缺省情况下，IRF的域编号为0
创建并进入聚合接口视图	进入二层聚合接口视图	interface bridge-aggregation <i>interface-number</i>	二者选其一
	进入三层聚合接口视图	interface route-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
使能LACP MAD检测功能		mad enable	缺省情况下，LACP MAD检测未使能
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

(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

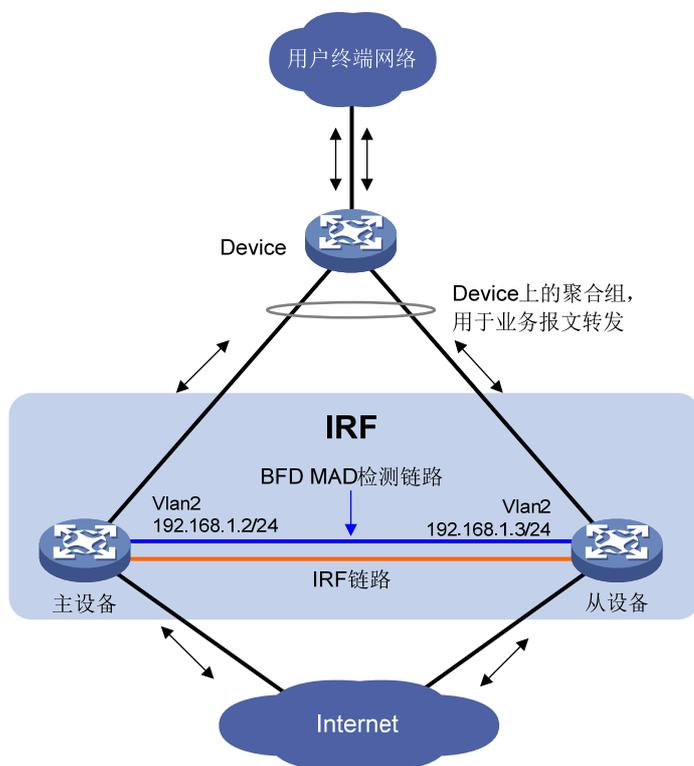
BFD MAD检测方式可以使用中间设备来进行连接，也可以不使用中间设备。通常采用如 [图 1-9](#) 所示的组网方式：每台成员设备必须和其它所有成员设备之间有一条BFD MAD检测链路（即成员设备之间是全连接组网）。这些链路连接的接口必须属于同一VLAN，在该VLAN接口视图下给不同成员设备配置同一网段下的不同IP地址。



提示

使能 BFD MAD 检测功能的三层接口只能专用于 BFD MAD 检测，这些接口下建议只配置 **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。

图1-9 BFD MAD 检测组网示意图



(3) 配置 BFD MAD 检测

配置 BFD MAD 检测时，请遵循以下要求：

- 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 VLAN 接口绑定的二层以太网接口上，请关闭生成树协议。
- 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。
- 配置 BFD MAD 检测功能之前，建议修改 IRF 链路 down 延迟上报时间为 0，以避免 BFD MAD 状态振荡。

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，在接口下使能 BFD MAD 检测功能，并配置 MAD IP 地址。

表1-19 配置 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
创建一个新VLAN专用于BFD MAD检测		vlan <i>vlan-id</i>	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 <i>interface-type</i> <i>interface-number</i>	-
端口加入 BFD MAD 检测专用 VLAN	Access端口	port access 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 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 <i>vlan-id</i>	
	Hybrid端口	port hybrid vlan <i>vlan-id</i> { tagged untagged }	
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	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>	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。

表1-20 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

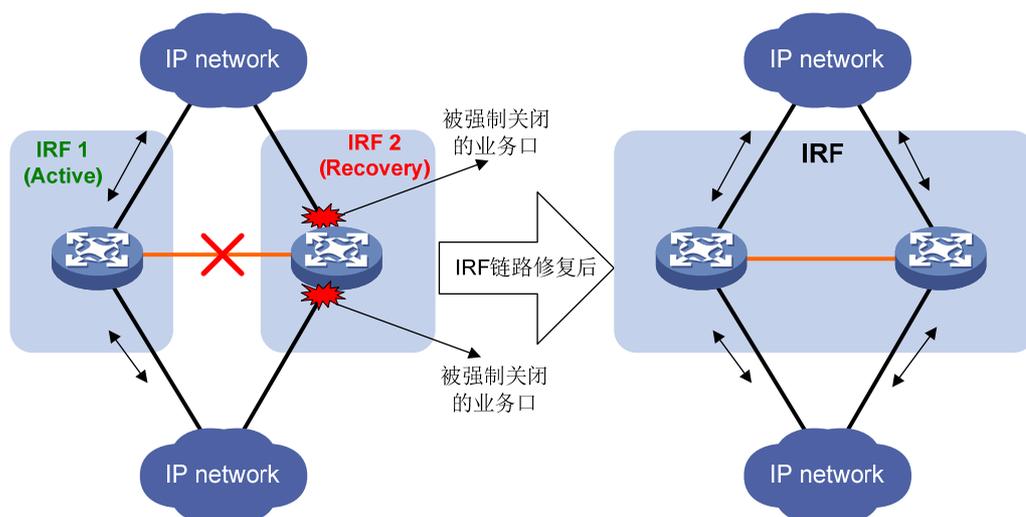
IRF链路故障将一个IRF分裂为两个IRF，从而导致多Active冲突。当系统检测到多Active冲突后，两个冲突的IRF会进行竞选，主设备成员编号小的获胜，继续正常运行，失败的IRF会转入Recovery

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正常工作状态的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，与设备是否支持以及用户是否配置了 **irf auto-merge enable** 命令有关。
- 请根据提示重启处于 Recovery 状态的 IRF，如果错误的重启了正常工作状态的 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正常工作状态的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+正常工作状态的 IRF 故障）

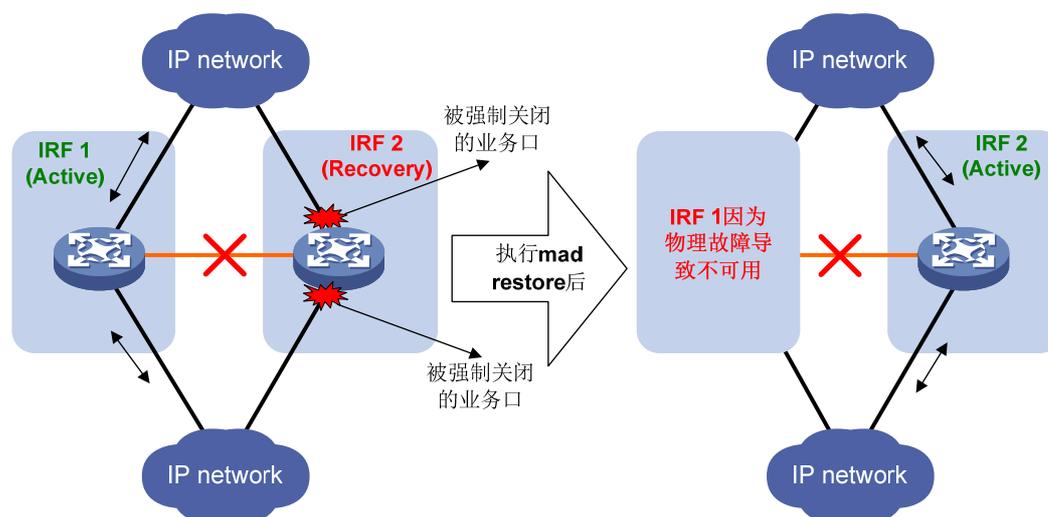


表1-21 手动恢复处于 Recovery 状态的设备

操作	命令	说明
进入系统视图	system-view	-
将IRF从Recovery状态恢复到正常工作状态	mad restore	-

1.10 快速恢复IRF配置

如果 IRF 中某台成员设备上只有一块主控板，此主控板一旦损坏，将引起 IRF 分裂。用户就需要重新进行 IRF 配置。本节介绍了一种快速恢复 IRF 配置的方法，可以大大减少重新配置 IRF 的工作量。

1.10.1 配置准备

- 推荐在 IRF 模式下进行本配置。
- 已经将 IRF 配置文件保存在 IRF 中的所有主控板上，假设名称为 a.cfg。

1.10.2 配置步骤

1. 情况 1：成员设备 1 上只有 1 块主控板，成员设备 2 上有 2 块主控板，成员设备 1 上的主控板损坏

- (1) 在成员设备 2 上，修改备用主控板的成员编号（配置命令请参见 [表 1-22](#)），使该备用主控板的成员编号与成员设备 1 的成员编号一致。
- (2) 拔出成员设备 1 上损坏的主控板，将成员设备 2 上的备用主控板插到成员设备 1 上。
- (3) 重新在成员设备 1 上配置 IRF 端口，与成员设备 2 进行连接。

快速恢复配置任务完成。

2. 情况 2: 2 台成员设备都只有 1 块主控板, 成员设备 1 上的主控板损坏

- (1) 在成员设备 2 上再插入一块主控板, 作为成员设备 2 的备用主控板。
- (2) 将成员设备 2 上主用主控板中的配置文件 `a.cfg` 拷贝到备用主控板中, 并将 `a.cfg` 设为下次启动的配置文件。
- (3) 修改成员设备 2 的备用主控板的成员编号 (配置命令请参见 [表 1-22](#)), 使该备用主控板成员编号与成员设备 1 的成员编号一致。
- (4) 拔出成员设备 1 上损坏的主控板, 将成员 2 设备上的备用主控板插到成员 1 设备上。
- (5) 重新在成员设备 1 上配置 IRF 端口, 与成员设备 2 进行连接。

快速恢复配置任务完成。

表1-22 修改备用主控板成员编号

操作	命令	说明
修改备用主控板的成员编号 (独立运行模式)	<code>irf slot slot-number member member-id</code>	此命令仅用于快速恢复IRF配置, 其它场合下请勿使用, 否则可能会发生未知错误 该命令在用户视图下执行
修改备用主控板的成员编号 (IRF 模式)	<code>irf chassis chassis-number slot slot-number member member-id</code>	

1.11 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	<code>display irf</code>
显示IRF的拓扑信息	<code>display irf topology</code>
显示IRF链路信息	<code>display irf link</code>
显示所有成员设备上重启以后生效的IRF配置	<code>display irf configuration</code>
显示IRF链路的负载分担模式	<code>display irf-port load-sharing mode [irf-port [member-id/port-number]]</code>
显示MAD配置信息	<code>display mad [verbose]</code>

1.12 IRF典型配置举例

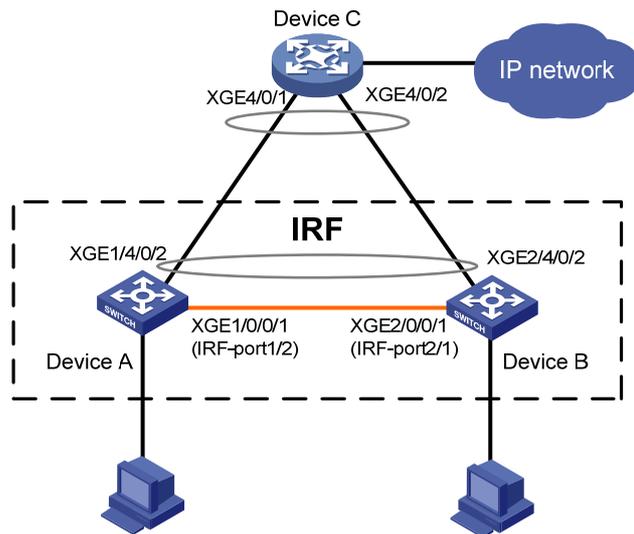
1.12.1 在R9900-08 或R9900-12 上组建IRF典型配置举例（LACP MAD检测方式）

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 组建 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 Ten-GigabitEthernet0/0/1 绑定。

```
<Sysname> system-view
[Sysname] irf member 1
[Sysname] interface ten-gigabitethernet 0/0/1
[Sysname-Ten-GigabitEthernet0/0/1] port link-mode bridge
[Sysname-Ten-GigabitEthernet0/0/1] quit
[Sysname] irf-port 2
```

```
[Sysname-irf-port2] port group interface ten-gigabitethernet 0/0/1
[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
Please wait...
Saving the converted configuration file to the main board succeeded.
Slot 4:
Saving the converted configuration file succeeded.
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 Ten-GigabitEthernet0/0/1 绑定。

```
<Sysname> system-view
[Sysname] irf member 2
[Sysname] interface ten-gigabitethernet 0/0/1
[Sysname-Ten-GigabitEthernet0/0/1] port link-mode bridge
[Sysname-Ten-GigabitEthernet0/0/1] quit
[Sysname] irf-port 1
[Sysname-irf-port1] port group interface ten-gigabitethernet 0/0/1
[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
Please wait...
Saving the converted configuration file to the main board succeeded.
Slot 4:
Saving the converted configuration file succeeded.
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，并使能 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 Ten-GigabitEthernet1/4/0/2 和 Ten-GigabitEthernet2/4/0/2，专用于 Device A 和 Device B 实现 LACP MAD 检测。

```
[Sysname] interface ten-gigabitethernet 1/4/0/2
```

```
[Sysname-Ten-GigabitEthernet1/4/0/2] port link-mode bridge
```

```
[Sysname-Ten-GigabitEthernet1/4/0/2] port link-aggregation group 2
```

```
[Sysname-Ten-GigabitEthernet1/4/0/2] quit
```

```
[Sysname] interface ten-gigabitethernet 2/4/0/2
```

```
[Sysname-Ten-GigabitEthernet2/4/0/2] port link-mode bridge
```

```
[Sysname-Ten-GigabitEthernet2/4/0/2] 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 Ten-GigabitEthernet4/0/1 和 Ten-GigabitEthernet4/0/2，用于帮助 LACP MAD 检测。

```
[Sysname] interface ten-gigabitethernet 4/0/1
```

```
[Sysname-Ten-GigabitEthernet4/0/1] port link-aggregation group 2
```

```
[Sysname-Ten-GigabitEthernet4/0/1] quit
```

```
[Sysname] interface ten-gigabitethernet 4/0/2
```

```
[Sysname-Ten-GigabitEthernet4/0/2] port link-aggregation group 2
```

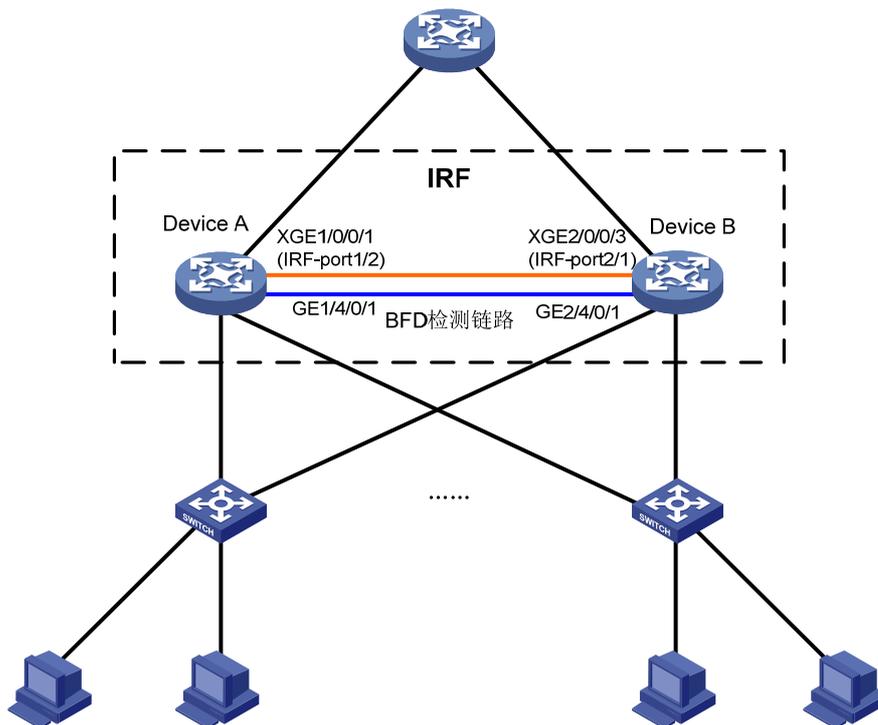
1.12.2 组建IRF典型配置举例（BFD MAD检测方式）

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 组建 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，创建 IRF 端口 2，并将它与物理端口 Ten-GigabitEthernet0/0/1 绑定。

```
<Sysname> system-view
[Sysname] irf member 1
[Sysname] interface ten-gigabitethernet 0/0/1
```

```
[Sysname-Ten-GigabitEthernet0/0/1] port link-mode bridge
[Sysname-Ten-GigabitEthernet0/0/1] quit
[Sysname] irf-port 2
[Sysname-irf-port2] port group interface ten-gigabitethernet 0/0/1
[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
Please wait...
Saving the converted configuration file to the main board succeeded.
Slot 4:
Saving the converted configuration file succeeded.
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 Ten-GigabitEthernet0/0/1 绑定。

```
<Sysname> system-view
[Sysname] irf member 2
[Sysname] interface ten-gigabitethernet 0/0/1
[Sysname-Ten-GigabitEthernet0/0/1] port link-mode bridge
[Sysname-Ten-GigabitEthernet0/0/1] quit
[Sysname] irf-port 1
[Sysname-irf-port1] port group interface ten-gigabitethernet 0/0/1
[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 1-13 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
Please wait...
Saving the converted configuration file to the main board succeeded.
Slot 4:
Saving the converted configuration file succeeded.
```

Now rebooting, please wait...

设备 B 重启后与设备 A 形成 IRF。

(3) 配置 BFD MAD 检测

创建 VLAN 3，并将 Device A（成员编号为 1）上的端口 1/4/0/1 和 Device B（成员编号为 2）上的端口 2/4/0/1 加入 VLAN 中。

```
<Sysname> system-view
[Sysname] interface range GigabitEthernet 1/4/0/1 GigabitEthernet 2/4/0/1
[Sysname-if-range] port link-mode bridge
[Sysname-if-range] quit
[Sysname] vlan 3
[Sysname-vlan3] port gigabitethernet 1/4/0/1 gigabitethernet 2/4/0/1
[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 GigabitEthernet1/4/0/1 和 GigabitEthernet2/4/0/1 上关闭生成树协议。

```
[Sysname] interface gigabitethernet 1/4/0/1
[Sysname-gigabitethernet-1/4/0/1] undo stp enable
[Sysname-gigabitethernet-1/4/0/1] quit
[Sysname] interface gigabitethernet 2/4/0/1
[Sysname-gigabitethernet-2/4/0/1] undo stp enable
```

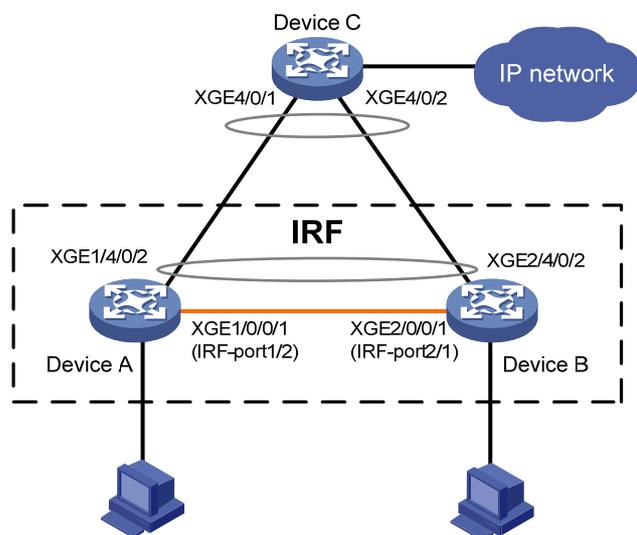
1.12.3 将成员设备从IRF模式恢复到独立运行模式配置举例

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	0	Master	1	00e0-fc0f-8c01	DeviceA
1	1	Standby	1	00e0-fc0f-8c02	DeviceA
2	0	Standby	1	00e0-fc0f-15e1	DeviceB
2	1	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: 0000-fc00-313e

```
Auto upgrade           : no
Mac persistent         : always
Domain ID              : 1
Auto merge             : no
```

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 Ten-Gigabitethernet 1/0/0/1。（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 ten-gigabitethernet 1/0/0/1
```

```
[IRF-Ten-Gigabitethernet1/0/0/1] shutdown
```

```
[IRF-Ten-GigabitEthernet1/0/0/1] quit
```

(3) 将 Device A 上 IRF 端口的配置删除。

```
[IRF] irf-port 1/2
```

```
[IRF-irf-port1/2] undo port group interface ten-gigabitEthernet1/0/0/1
```

```
[IRF-irf-port1/2] quit
```

(4) 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
```

```
Please wait.....
```

```
Saving the converted configuration file to main board succeeded.
```

```
Chassis 1 Slot 1:
```

```
Saving the converted configuration file succeeded.
```

```
Now rebooting, please wait...
```

Device A 自动重启来完成模式的切换。

(5) 登录 Device B 后，将 Device B 上 IRF 端口的配置删除。

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

```
[IRF] interface ten-gigabitEthernet 2/0/0/1
```

```
[IRF-Ten-GigabitEthernet2/0/0/1] shutdown
```

```
[IRF-Ten-GigabitEthernet2/0/0/1] quit
```

```
[IRF] irf-port 2/1
```

```
[IRF-irf-port2/1] undo port group interface ten-gigabitEthernet2/0/0/1
```

```
[IRF-irf-port2/1] quit
```

(6) 将 Device B 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
```

```
Please wait.....
```

```
Saving the converted configuration file to main board succeeded.
```

```
Chassis 2 Slot 1:
```

```
Saving the converted configuration file succeeded.
```

```
Now rebooting, please wait...
```

Device B 自动重启来完成模式的切换。



说明

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。
