

目 录

1 IRF3.1	1-1
1.1 IRF3.1 简介	1-1
1.1.1 IRF3.1 典型组网应用	1-1
1.1.2 IRF3.1 的优点	1-1
1.1.3 IRF3.1 物理连接拓扑	1-2
1.1.4 IRF3.1 基本概念	1-5
1.1.5 PEX 设备加入 IRF3.1 系统	1-7
1.1.6 PEX 设备上的接口命名规则	1-7
1.1.7 配置管理	1-7
1.1.8 数据转发	1-8
1.1.9 协议规范	1-8
1.2 IRF3.1 配置限制和指导	1-9
1.2.1 硬件兼容性限制	1-9
1.2.2 软件兼容性限制	1-9
1.2.3 MDC 配置兼容性限制	1-9
1.2.4 上行接口的物理端口选择	1-9
1.2.5 环路避免	1-10
1.2.6 PEX 堆叠配置限制和指导	1-10
1.2.7 PEX 二层聚合接口配置限制和指导	1-10
1.2.8 PEX 设备配置管理限制和指导	1-10
1.2.9 MAD 配置限制和指导	1-10
1.2.10 配置回滚限制	1-11
1.3 IRF3.1 配置任务简介	1-11
1.3.1 手工配置 IRF3.1 系统	1-11
1.3.2 自动配置 IRF3.1 系统	1-12
1.4 IRF3.1 配置准备	1-12
1.5 搭建 IRF 作为 IRF3.1 父设备	1-12
1.6 创建 PEX 组	1-12
1.7 配置连接一级 PEX 的级联接口	1-13
1.8 配置连接二级 PEX 的级联接口	1-14
1.9 分配虚拟框号	1-16
1.10 开启 IRF3.1 系统自动配置功能	1-16

1.11 将 PEX 连接到上级设备	1-17
1.12 开启 PEX 设备本地转发功能	1-18
1.13 配置 PEX 设备下线后的报文转发功能	1-18
1.14 从父设备上登录到 PEX 设备	1-19
1.15 在父设备上保存 PEX 独立支持特性的配置文件.....	1-20
1.16 删除空闲的级联接口	1-21
1.17 IRF3.1 显示和维护.....	1-21
1.18 IRF3.1 典型配置举例.....	1-22
1.18.1 两级 PEX IRF3.1 典型配置举例	1-22

1 IRF3.1

1.1 IRF3.1简介

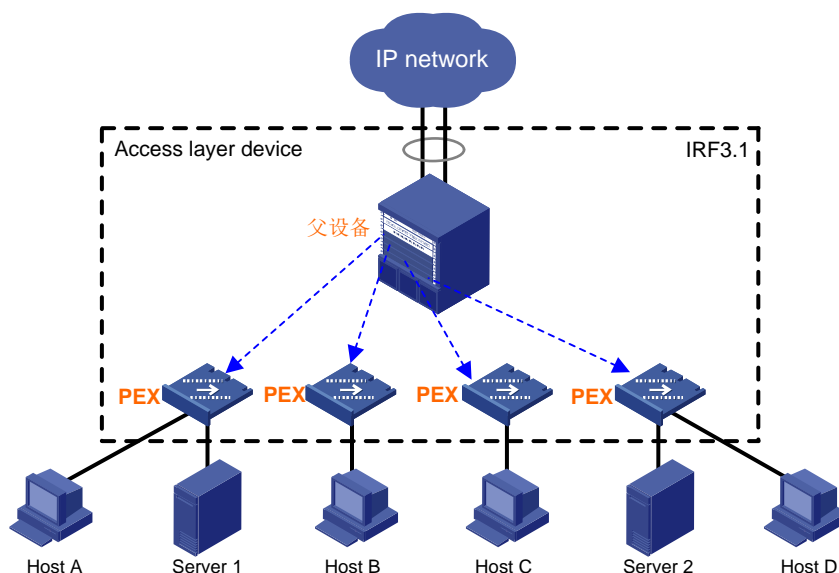
IRF3.1 是一种能够提高网络接入层的接入能力和管理效率的纵向网络整合虚拟化技术，采用 IEEE 802.1BR 标准协议实现。

IRF3.1 将多台 PEX 设备（Bridge Port Extender）连接到父设备（Parent device）上，PEX 中的每台物理设备相当于父设备的一块远程业务板，由父设备统一管理。使用这种虚拟化技术可以以较低的成本，来提高父设备的接口密度，简化网络拓扑，降低网络维护成本。

1.1.1 IRF3.1 典型组网应用

IRF3.1 技术通常和 IRF 配合使用，用于数据中心和大型企业网络的接入层。如图 1-1 所示，父设备是由两台设备组成的 IRF，PEX 设备用于接入终端、服务器，最终实现将所有接入层设备虚拟成一台分布式接入设备（IRF3.1）。

图1-1 IRF3.1 组网应用示意图



1.1.2 IRF3.1 的优点

IRF3.1 主要具有以下优点：

- 简化管理

使用 IRF3.1 技术，可以将每台 PEX 设备虚拟成父设备上的一块远程业务板。对于上下层设备和网管来说，网络中只存在一台设备（一个管理点），用户只需在父设备上进行操作即可实现对父设备和 PEX 设备的统一管理。

- 强大的网络扩展能力
通过增加 PEX 设备，可以轻松自如地扩展父设备的端口数。
- 降低成本
在 IRF3.1 中，父设备承担协议的运算、表项的生成和下发，PEX 设备仅负责数据转发，所以对 PEX 设备的性能、规格要求不高。只需要使用较低配置的款型，就能达到网络扩容的效果，有利于成本控制；另外，PEX 设备自动同步父设备的配置，降低了网络管理成本。
- 高可靠性
PEX 设备上存在多个高速率的物理接口，可以用于连接父设备。在这些接口上建立的 PEX 链路可以进行负载分担并互为备份，从而确保了 PEX 设备和父设备之间报文转发的高可靠性。
- 父设备和 PEX 设备软件版本兼容性高
父设备和 PEX 设备采用 IEEE 802.1BR 标准协议交互信息。PEX 设备运行复杂度低，和父设备软件版本松耦合。父设备和 PEX 设备的软件版本可以独立更新，更新父设备或 PEX 设备的软件版本时，通常不需要对端配合更新软件版本，双方可以保持长时间的版本兼容。

1.1.3 IRF3.1 物理连接拓扑

1. 单层接入方式

如图 1-2 所示，每台 PEX 设备通过动态类型二层聚合链路（需要为动态类型）与父设备直接相连。父设备只连接一级 PEX 设备。PEX 设备之间不能有物理连接。

PEX 可以为单台物理设备，也可以为多台物理设备组成的 IRF 设备，如图 1-3 所示。

图1-2 单层 PEX 组网应用示意图（PEX 为单台物理设备）

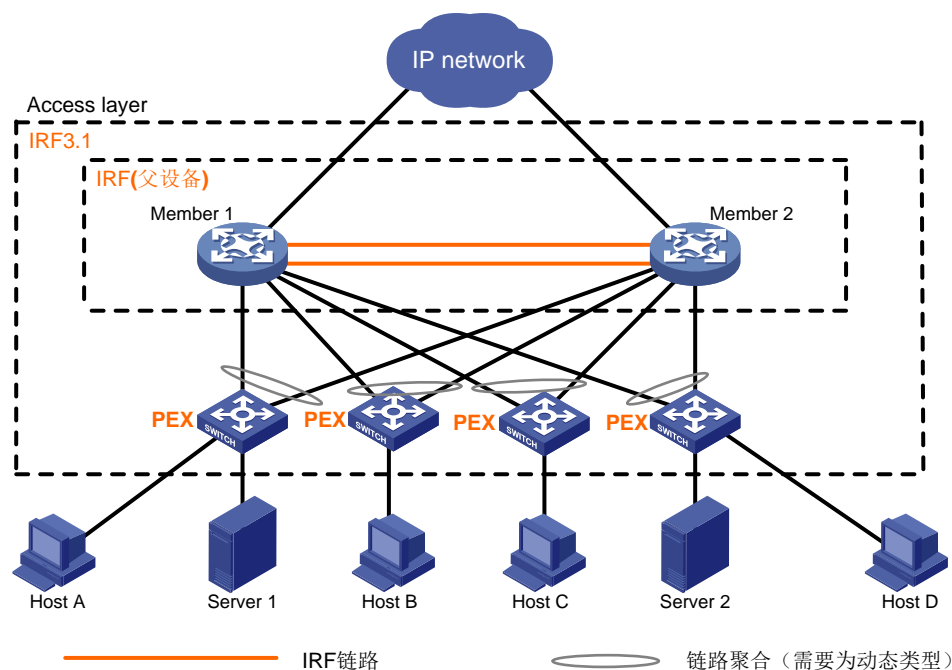
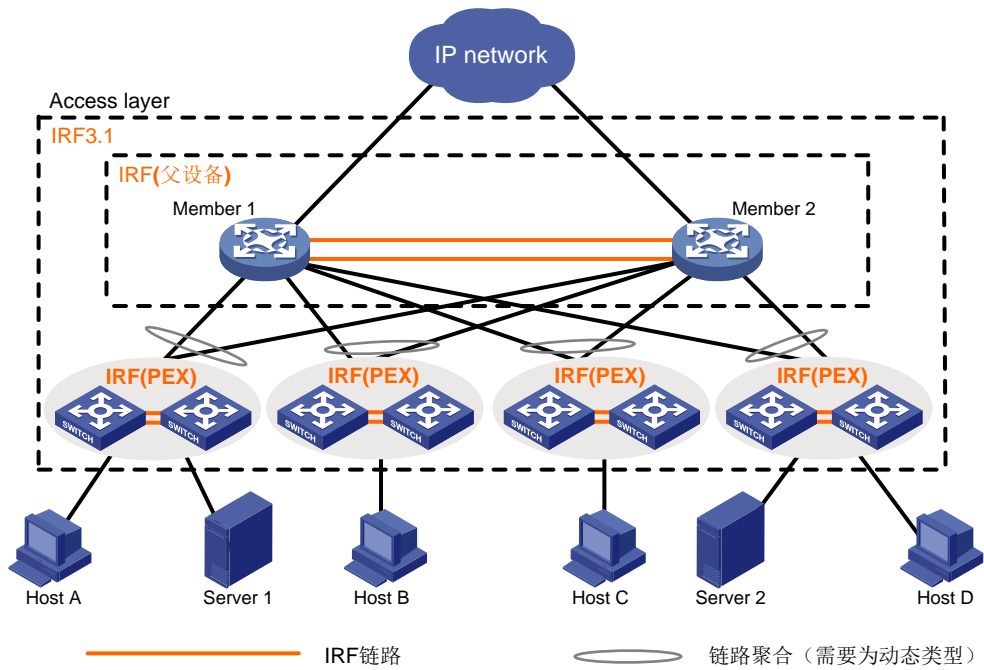


图1-3 单层 PEX 组网应用示意图 (PEX 为 IRF 设备)



2. 级联接入方式

如[图 1-4](#)所示，每台 PEX 设备通过动态类型二层聚合链路与父设备或上级 PEX 相连。我们将直接连接到父设备的 PEX 称为一级 PEX，直接连接到一级 PEX 的 PEX 称为二级 PEX，以此类推。每个二级 PEX 只能连接一个一级 PEX。PEX 设备之间不能有级联链路之外的物理连接。设备最多支持二层 PEX 级联。

一级 PEX 可以为单台物理设备，也可以为多台物理设备组成的 IRF 设备，如[图 1-5](#)所示。二级 PEX 只能为单台物理设备。

图1-4 级联 PEX 组网应用示意图 (所有 PEX 为单台物理设备)

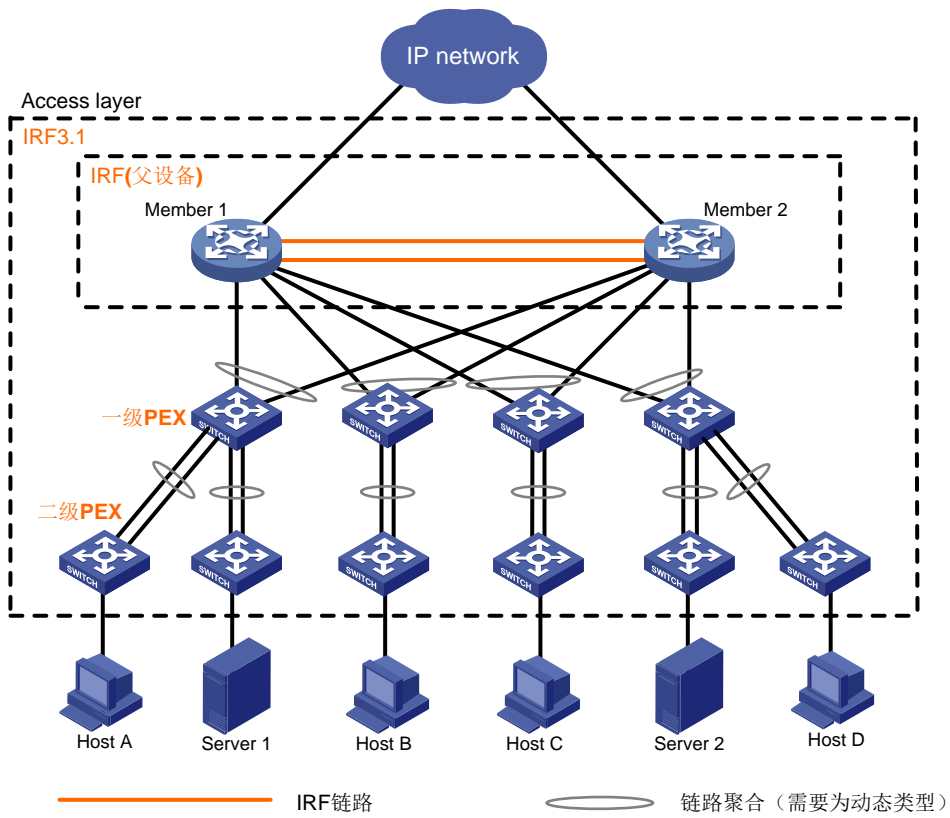
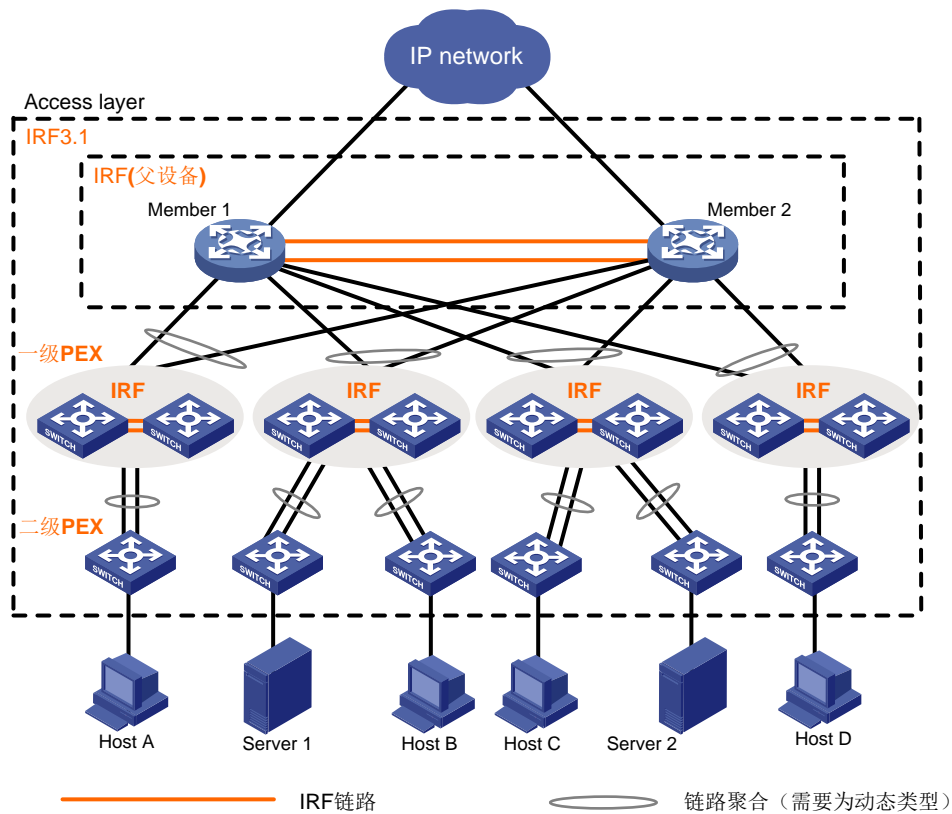


图1-5 级联 PEX 组网应用示意图 (一级 PEX 为 IRF 设备)



1.1.4 IRF3.1 基本概念

1. 角色

IRF3.1 组网中设备按照功能不同，分为两种角色：

- 父设备：作为所有 PEX 设备的主控设备，负责配置和管理整个 IRF3.1 系统。父设备为 IRF 技术构建的虚拟设备，也被称为 CB（Controlling Bridge，控制设备）。
- PEX 设备：作为父设备的远程业务板，负责将用户发送的报文传送到父设备处理。设备作为 PEX 设备运行时，只允许通过父设备进行配置。通过 PEX 设备能够按需灵活地增加父设备上可用端口的数量。PEX 设备有两种工作状态，如表 1-1 所示：

表1-1 PEX 设备状态描述表

状态	说明
Offline	下线状态，表示PEX设备与父设备之间没有建立PE CSP连接，不能通过父设备管理
Online	上线状态，表示父设备已经通过LLDP发现PEX邻居设备，并完成PE CSP（Port Extender Control and Status Protocol，端口扩展设备控制和状态协议）协商，PEX设备成功上线

2. 级联接口

动态类型的二层聚合接口开启连接 PEX 的能力后称为级联接口，包括：

- 父设备上与一级 PEX 设备连接的接口。
- 上级 PEX 上与下级 PEX 连接的接口。

3. 上行接口

下级设备上与级联接口相连的接口，包括：

- 一级 PEX 设备上连接父设备的接口。
- 下级 PEX 上连接上级 PEX 的接口。

PEX 设备上不需要手动创建与上级设备连接的二层聚合接口。设备作为 PEX 设备启动时自动创建动态类型的二层聚合接口并将连接上级设备的物理端口加入聚合组。

4. PEX 组

PEX 组为一组 PEX 的集合。当级联接口配置在同一个 PEX 组中时，我们称这些 PEX 为同一个组中的 PEX。下级 PEX 设备只能与其直连的上级 PEX 设备配置在同一 PEX 组。

5. 扩展端口

IRF3.1 系统中位于 PEX 上的物理端口称为扩展端口，除了上行接口的物理端口。

6. PEX 二层聚合接口

成员端口为扩展端口的聚合组称为 PEX 二层聚合组，对应的聚合接口称为 PEX 二层聚合接口。PEX 二层聚合接口可以作为上级 PEX 连接下级 PEX 的级联接口或者普通业务接口使用。

有关 PEX 二层聚合接口的详细介绍，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”。

7. PEX 堆叠

当 PEX 设备为多台成员设备组成的 IRF 时，我们将其称为 PEX 堆叠。本文中如无特殊说明，PEX 设备的描述也适用于 PEX 堆叠。

8. 虚拟框号

用来在父设备上标识并管理 PEX 设备。一台 PEX 设备对应一个虚拟框。物理设备相当于安插在虚拟框上的业务板，对应的槽位号为设备的 IRF 成员编号。

1.1.5 PEX 设备加入 IRF3.1 系统

1. 邻居设备发现

在父设备上完成 PEX 配置且父设备与 PEX 设备互连的物理链路状态为 UP 后，父设备和 PEX 设备互相发送 LLDP 报文发现邻居设备。

对于二级 PEX，其上级 PEX 会透传它和父设备之间的 LLDP 报文，以使父设备完成邻居发现。

2. PE CSP 连接建立

建立 LLDP 邻居关系后，父设备和 PEX 设备分别发送 PE CSP Open 协议报文，如果双方均能在发送请求后 60 秒内接收到对端回复的 PE CSP Open 报文，则父设备和 PEX 之间的连接建立成功。对于二级 PEX，其上级 PEX 会透传它和父设备之间的 PE CSP 报文，以使父设备建立 PE CSP 连接。

3. PEX 注册

父设备与 PEX 设备建立 PE CSP 连接后，进行下面的操作：

- (1) PEX 设备向父设备注册。父设备用级联接口上配置的虚拟框号标识 PEX 设备。
- (2) PEX 设备发送接口创建请求。父设备收到请求后，创建扩展端口并为端口分配 ECID（E-channel Identifier，E 通道标识）。ECID 在 PEX 组内唯一。同时，PEX 设备上接口的状态、双工、速率属性会同步到父设备创建的扩展端口上。

4. PEX 链路维护

父设备和 PEX 设备采用动态聚合接口连接，动态聚合接口通过 LACP 协议和 PE CSP 协议检测和维持链路状态。当聚合接口状态变为 down 或父设备/PEX 设备在发送请求报文后 60 秒未收到对端回应时，PEX 设备下线。父设备与 PEX 设备将会重新互发 Open 报文，以使 PEX 重试上线。

1.1.6 PEX 设备上的接口命名规则

PEX 设备完成向父设备的注册后，PEX 设备上接口编号将由三维变成四维：第一维为虚拟框号，第二维为 PEX 设备的 IRF 成员编号，三维和四维为原接口编号的二维和三维。例如，PEX 设备在加入 IRF3.1 之前，IRF 成员编号为 1，某接口编号为 1/0/1，父设备为该 PEX 设备分配的虚拟框号为 100，则该接口的编号将变为 100/1/0/1。

1.1.7 配置管理

PEX 设备虚拟化为父设备的业务板。请登录父设备，来完成 PEX 设备的配置，比如配置 PEX 设备上端口所属的 VLAN 等。用户在父设备执行 **save** 命令保存当前配置时，也会将 PEX 设备对应配置保存到父设备上。当系统重启或者更换新的 PEX 设备时，父设备会将 PEX 对应配置下发给该槽位的 PEX 设备。

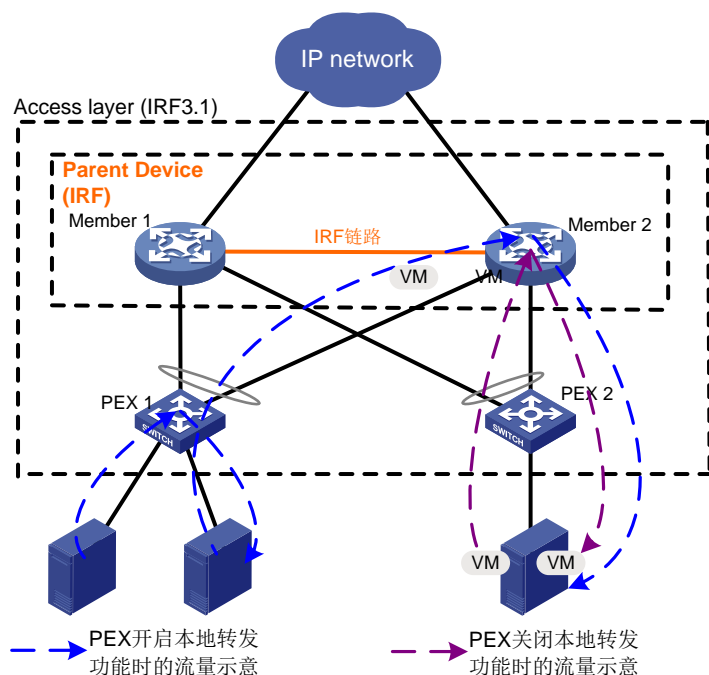
部分 PEX 独立支持的功能，需要使用 `switchto pex` 命令登录到 PEX 设备上配置。完成配置后，请返回到父设备执行 `save pex-configuration` 命令进行 PEX 独立支持特性的配置保存。PEX 上线或重启时会使用父设备上保存的配置文件进行 PEX 独立支持特性的配置恢复。

1.1.8 数据转发

PEX 设备上的数据转发处理机制与 PEX 本地转发功能的开启状态有关，具体如下：

- PEX 本地转发功能处于开启状态时，PEX 设备对二层已知单播报文进行本地转发，其它报文转发给父设备处理。
- PEX 本地转发功能处于关闭状态时，PEX 设备收到的所有报文均转发给父设备处理，父设备进行转发决策，再转发给出接口，如 [图 1-6](#) 所示。PEX 设备的接口接收报文时，为报文打上 E-TAG，E-TAG 中携带接口 ECID。IRF3.1 系统根据 ECID 信息完成报文在系统内部的转发。当报文离开 IRF3.1 系统时，设备去掉 E-TAG 转发报文。

图1-6 IRF3.1 数据转发示意图



1.1.9 协议规范

与 IRF 3.1 相关的协议规范有：IEEE 802.1BR: Virtual Bridged Local Area Networks—Bridge Port Extension。

1.2 IRF3.1配置限制和指导

1.2.1 硬件兼容性限制

父设备上仅下列单板上的接口可以作为级联接口，与 PEX 进行连接：

- 主控板：LSQM1CTGS24QSFD0
- FD 系列接口板
- FE 系列接口板
- SG 系列接口板

本设备作为父设备时，支持以下设备作 PEX 设备：

- S5130S-HI 系列交换机。
- S3100V3-EI 系列交换机。
- S5130S-EI 系列交换机。
- S5560X-EI 系列交换机。

不同型号不同版本的 PEX 设备配置不同，请参见具体 PEX 设备对应版本的配置指导。

在 IRF3.1 组网时，在设备上请不要使用以下单板：

- EB 系列接口板
- 下列 SA 系列接口板：LSQM1TGS16GPSA0
- 下列 SC 系列接口板：LSQ1QGS4SC0
- SD 系列接口板
- SF 系列接口板

1.2.2 软件兼容性限制

设备的融合 AC（Access Controller，无线控制器）特性安装后，仅本设备支持使用该特性，PEX 设备不支持。有关融合 AC 特性的详细介绍，请参见《融合 AC 用户手册》。

1.2.3 MDC 配置兼容性限制

在 MDC 环境中部署 IRF3.1 时，需要按照如下步骤配置：

- 搭建 IRF 父设备
- 完成 MDC 的划分。
- 将 PEX 的使用权限分配给 MDC。
- 在 MDC 中完成 PEX 的配置。配置方式与普通物理设备上的配置方式相同。

更多信息请参考“虚拟化配置指导”中的“MDC”。

1.2.4 上行接口的物理端口选择

PEX 设备上只有特定的高速率端口可以作为上行接口的物理端口，我们称这些端口为上行接口的候选成员端口。可通过如下方式来确定上行接口的候选成员端口：

- 设备未切换为 PEX 模式时，请查看 PEX 设备的配置手册确认上行接口的候选成员端口。
- 当设备处于 PEX 模式时，可以在系统视图下执行 **probe** 命令进入 Probe 视图，在 Probe 视图执行 **display system internal pex upstreamport** 命令查看上行接口的候选成员端口。

上行接口的候选成员端口中，我们将同一类型的端口视为一组。例如某设备的 Ten-GigabitEthernet1/0/43~Ten-GigabitEthernet1/0/48 和 FortyGigE1/0/49~FortyGigE1/0/54 可以作为上行接口的物理端口。则 Ten-GigabitEthernet1/0/43~Ten-GigabitEthernet1/0/48 为一组，FortyGigE1/0/49~FortyGigE1/0/54 为一组。当组中有端口成为上行接口的物理端口时，同组中的其它端口也被系统保留，不创建为扩展端口，在 IRF3.1 系统中不可见。如果需要修改 PEX 使用的上行物理端口组，需要重启 PEX 后重新配置。

IRF3.1 系统组建完成后，如需将 PEX 上行接口的物理接口更换成不同速率的接口，请配置完成后重启该 PEX 设备。

1.2.5 环路避免

缺省情况下，为了减少系统开销，PEX 设备上端口的生成树协议处于关闭状态。为了避免环路，请选择其中一种方式实施：

- 确保 IRF3.1 组网中不存在环路（同一台 PEX 的两个端口之间、两台 PEX 设备之间、PEX 设备和主设备之间不存在级联链路之外的连接）。
- 开启 PEX 设备上端口的生成树协议。

1.2.6 PEX 堆叠配置限制和指导

PEX 设备可组建 IRF，但是必须在加入 IRF3.1 前组建，组建时，请确保 PEX 设备的 IRF 成员编号范围为 1~5。PEX 设备组建 IRF 的方法请参见 PEX 设备的 IRF 配置指导。

1.2.7 PEX 二层聚合接口配置限制和指导

只有同一个 PEX 组中同一级 PEX 设备上的扩展端口才允许加入同一个 PEX 二层聚合组，且请使用同系列的 PEX 设备进行链路聚合。

仅同一级 PEX 设备的接口可以加入同一个聚合组，且请使用相同系列的 PEX 设备进行链路聚合。

1.2.8 PEX 设备配置管理限制和指导

配置 PEX 设备时：

- 对于非 PEX 独立支持特性，只允许通过父设备进行配置。
- 对于 PEX 独立支持特性，请通过“[1.15 在父设备上保存 PEX 独立支持特性](#)”中所示方法进行。

查看 PEX 设备运行信息，请通过“[1.14 从父设备上登录到 PEX 设备](#)”中所示方法进行。

1.2.9 MAD 配置限制和指导

为了使 PEX 设备能够迅速感知到作为父设备的 IRF 是否发生分裂，请在父设备上进行如下配置：

- 使用 **lACP system-priority** 命令配置系统的 LACP 优先级值为小于缺省值 32768。
- 使用 **undo irf mac-address persistent** 命令将 IRF 配置为桥 MAC 地址立即改变，同时请不要使用 **irf mac-address mac-address** 命令配置 IRF 的桥 MAC 为指定 MAC 地址。

仅父设备上的接口支持作 MAD 检测接口，PEX 设备上的接口不支持。

在 PEX 堆叠接入组网中，系统会自动在 PEX 上行接口上下发 LACP MAD 配置。当 PEX 堆叠发生分裂时，PEX 堆叠中进入 Recovery 状态的 IRF 关闭业务接口，从而使 PEX 从 IRF3.1 系统中下线，不影响 IRF3.1 系统的稳定性。

1.2.10 配置回滚限制

因为 PEX 组的配置、接口连接 PEX 能力的配置、PEX 设备虚拟槽位号的配置不支持回滚。如果当前配置文件和回滚配置文件之间存在 IRF3.1 相关配置的差异，则不建议进行配置回滚。否则 IRF3.1 系统的配置恢复结果可能与预期不同。有关配置回滚的详细介绍，请参见“基础配置指导”中的“配置文件”。

1.3 IRF3.1 配置任务简介

IRF3.1 系统支持两种配置方式：手工配置和自动配置。

当自动配置功能处于开启状态时，不建议进行手工配置；否则手工配置和自动配置可能会互相影响，造成 PEX 无法正常上线或配置效果与预期不同。

1.3.1 手工配置 IRF3.1 系统

手工配置 IRF3.1 任务如下：

- (1) [搭建 IRF 作为 IRF3.1 父设备](#)
- (2) 在父设备上进行一次 PEX 配置
 - a. [创建 PEX 组](#)
 - b. [配置连接一级 PEX 的级联接口](#)
 - c. [分配虚拟框号](#)
 - d. [将 PEX 连接到上级设备](#)
- (3) 在父设备上进行一次二级 PEX 配置
 - a. [配置连接二级 PEX 的级联接口](#)
 - b. [分配虚拟框号](#)
 - c. [将 PEX 连接到上级设备](#)

PEX 级联接入时必须。先完成上级 PEX 的配置并使上级 PEX 设备上线，才能进行下级 PEX 设备的配置。

- (4) （可选）在父设备上进行一次 PEX 的调整和维护
 - o [开启 PEX 设备本地转发功能](#)
 - o [配置 PEX 设备下线后的报文转发功能](#)
 - o [从父设备上登录到 PEX 设备](#)

- [在父设备上保存 PEX 独立支持特性](#)
- [删除空闲的级联接口](#)

1.3.2 自动配置 IRF3.1 系统

- (1) [搭建 IRF 作为 IRF3.1 父设备](#)
- (2) [自动配置 IRF3.1 系统](#)
 - a. [开启 IRF3.1 系统自动配置功能](#)
 - b. [将 PEX 连接到上级设备](#)
- (3) (可选) 在父设备上进行 PEX 的调整和维护
 - [开启 PEX 设备本地转发功能](#)
 - [配置 PEX 设备下线后的报文转发功能](#)
 - [从父设备上登录到 PEX 设备](#)
 - [在父设备上保存 PEX 独立支持特性](#)
 - [删除空闲的级联接口](#)

1.4 IRF3.1配置准备

进行网络规划，确定以下事项：

- IRF3.1 系统中父设备成员设备数量、PEX 设备数量
- 父设备和 PEX 设备、PEX 设备间的层级关系
- PEX 设备虚拟框号分配方案
- 级联接口、上行接口和连线方案

1.5 搭建IRF作为IRF3.1父设备

在 IRF3.1 系统中，父设备必须为 IRF 设备，可以是单台成员设备组成的 IRF。

1.6 创建PEX组

1. 配置限制和指导

创建 PEX 组时，设备会自动开启 NTP 服务，并配置本地时钟作为参考时钟，层数为 2，用于父设备和 PEX 间进行时间同步。

删除 PEX 组会导致组中所有 PEX 下线，请谨慎配置。PEX 设备下线时，父设备下发到 PEX 设备的所有配置将被清除（PEX 独立支持特性的配置不会被清除）。

2. 配置步骤

表1-2 创建 PEX 组

操作	命令	说明
进入系统视图	system-view	-
创建PEX组	pex group <i>group-id</i>	-
(可选) 为PEX组配置描述信息	description <i>text</i>	缺省情况下, PEX组的描述信息为“PEX group <i>group-id</i> ”

1.7 配置连接一级PEX的级联接口

1. 功能简介

父设备需要使用非 PEX 聚合的动态类型二层聚合接口作为与一级 PEX 连接的级联接口。一个级联接口对应一台 PEX 设备。

开启接口连接 PEX 的能力后, 接口会自动配置为动态聚合模式和生成树边缘端口。级联接口自动配置为生成树边缘端口可以使 PEX 设备尽快完成协议报文交互并上线。

2. 配置限制和指导

- 为保证 IRF3.1 协议正常运行和 PEX 正常上线, 配置如下端口时请注意:
 - 级联接口上除了 IRF3.1 相关配置命令、以及 **shutdown**、**description** 命令外, 请勿配置其它命令。
 - 聚合成员端口除了 **lldp enable** 外, 请勿配置其它命令。
- 如需关闭级联轴接口的 PEX 连接能力或修改其所属的 PEX 组, 必须保证其对应的 PEX 设备上连接二级 PEX 的接口已经关闭连接 PEX 的能力。
- 关闭接口连接 PEX 的能力会导致对应 PEX 设备下线。修改级联接口所属的 PEX 组, PEX 设备会先下线, 然后重新上线。PEX 设备下线时, 父设备下发到 PEX 设备的所有配置被清除 (PEX 独立支持特性的配置不会被清除)。如果该 PEX 设备连接了下级 PEX, 且下级 PEX 需要重新上线, 则需要重新配置级联接口。

3. 配置步骤

表1-3 配置连接一级 PEX 的级联接口

操作	命令	说明
进入系统视图	system-view	-

操作		命令	说明
全局开启LLDP功能		lldp global enable	缺省情况下，全局LLDP功能处于关闭状态
创建二层聚合接口，并进入二层聚合接口视图		interface bridge-aggregation interface-number	创建二层聚合接口后，系统将自动生成同编号的二层聚合组，且该聚合组缺省工作在静态聚合模式下
开启接口连接PEX的能力，并将接口加入PEX组		pex-capability enable group group-id	缺省情况下，接口不支持连接PEX
退回系统视图		quit	-
进入连接一级PEX的二层以太网接口视图	进入二层以太网接口视图	interface interface-type interface-number	在对多个接口进行相同配置时，通过接口批量配置视图可以更快速完成操作
	进入一组接口的批量配置视图	interface range { interface-type interface-number [to interface-type interface-number] } &<1-24>	
将二层以太网接口加入聚合组		port link-aggregation group group-id	同一个聚合接口的成员端口需要连接到同一台PEX设备上
开启接口的LLDP功能		lldp enable	缺省情况下，LLDP功能在接口上处于开启状态

1.8 配置连接二级PEX的级联接口

1. 功能简介

父设备需要使用动态类型 **PEX** 二层聚合接口作为连接二级 **PEX** 的级联接口。聚合成员端口必须全部是上级 **PEX** 上的物理端口。

开启接口连接 **PEX** 的能力后，接口会自动配置为动态聚合模式和生成树边缘端口。级联接口自动配置为生成树边缘端口可以使 **PEX** 设备尽快完成协议报文交互并上线。

2. 配置限制和指导

- 必须先向 **PEX** 二层聚合组添加成员端口，然后才允许开启接口连接 **PEX** 的能力。否则由于设备无法获取到上级 **PEX** 所属的 **PEX** 组，而无法完成开启接口连接 **PEX** 能力的配置。
- 建议选择设备上的高速率端口作为级联轴口的成员端口。
- 为保证 **IRF3.1** 协议正常运行和 **PEX** 正常上线，配置如下端口时请注意：
 - 级联轴口上除了 **IRF3.1** 相关配置命令、以及 **shutdown**、**description** 命令外，请勿配置其它命令。
 - 聚合成员端口除了 **lldp enable** 外，请勿配置其它命令。
- 二级 **PEX** 只能与其上级 **PEX** 加入同一个 **PEX** 组。如需修改二级 **PEX** 所属的 **PEX** 组，需要先关闭级联轴口的 **PEX** 连接能力，然后修改一级 **PEX** 所属的 **PEX** 组。

- 关闭接口连接 PEX 的能力会导致对应 PEX 设备下线。PEX 设备下线时，父设备下发到 PEX 设备的所有配置被清除（PEX 独立支持特性的配置不会被清除）。

3. 配置准备

完成上级 PEX 的配置，并在父设备上执行 **display pex interface brief** 命令确认上级 PEX 已经上线。

4. 配置步骤

表1-4 配置连接二级 PEX 的级联接口

操作		命令	说明
进入系统视图		system-view	-
创建PEX二层聚合组，并进入PEX二层聚合接口视图		interface bridge-aggregation interface-number pex	创建二层聚合接口后，系统将自动生成同编号的二层聚合组，且该聚合组缺省工作在静态聚合模式下
退回系统视图		quit	-
进入一级PEX上连接二级PEX的二层以太网接口视图	进入二层以太网接口视图	interface interface-type interface-number	在对多个接口进行相同配置时，通过接口批量配置视图可以更快速完成操作
	进入一组接口的批量配置视图	interface range { interface-type interface-number [to interface-type interface-number] } &<1-24>	
将二层以太网接口加入聚合组		port link-aggregation group group-id	-
开启接口的LLDP功能		lldp enable	缺省情况下，LLDP功能在接口上处于开启状态
退回系统视图		quit	-
进入二层聚合接口视图		interface bridge-aggregation interface-number	-
开启接口连接PEX的能力，并将接口加入PEX组		pex-capability enable group group-id	缺省情况下，接口不支持连接PEX

1.9 分配虚拟框号

1. 配置限制和指导

删除 PEX 设备的虚拟框号会导致对应 PEX 设备下线。修改 PEX 设备的虚拟框号，PEX 设备会先下线，然后使用更新的虚拟框号重新上线。并且，PEX 设备下线时，父设备下发到 PEX 设备的所有配置被清除（PEX 独立支持特性的配置不会被清除）。

在 PEX 级联环境中：

- 如果上级 PEX 因为删除/修改虚拟框号下线，所有下级 PEX 也会同时下线，但下级 PEX 的配置不会被清除。
- 上级 PEX 修改虚拟框号重新上线后，需要重新配置上级 PEX 物理端口加入下级 PEX 的级联接口，才能使下级 PEX 重新上线。

2. 配置步骤

表1-5 分配虚拟框号

操作	命令	说明
进入系统视图	system-view	-
进入连接PEX设备的二层聚合接口视图	interface bridge-aggregation interface-number	-
为PEX设备分配虚拟框号	pex associate chassis chassis-number	缺省情况下，没有为PEX设备分配虚拟框号

1.10 开启IRF3.1系统自动配置功能

1. 功能简介

父设备开启 IRF3.1 系统自动配置功能后，如果检测到对端发送的 LLDP 报文中携带 Port Extension TLV，则认为对端是 PEX 设备。在与对端建立 LLDP 邻居关系后，父设备自动下发 IRF3.1 相关配置，完成 IRF3.1 系统构建。

以一台一级 PEX 设备接入为例，自动配置的过程如下：

- (1) 父设备自动创建 PEX 组 1（通过自动配置功能上线的 PEX 都将加入 PEX 组 1）。
- (2) 父设备检测到一级 PEX 设备发送的 LLDP 报文，自动创建 1 个二层聚合接口用于连接一级 PEX，并将级联物理接口自动加入聚合组中（设备自动创建的聚合接口编号选取当前系统未被使用的接口编号）。
- (3) 开启二层聚合接口连接 PEX 的能力并将接口加入 PEX 组 1。
- (4) 父设备自动为一级 PEX 分配虚拟框号（虚拟框号选取当前系统中未被使用的编号）。

完成上述步骤后，一级 PEX 的配置完成，PEX 上线。

IRF3.1 系统自动配置功能下发的具体配置请参见系统输出的日志信息或查看 **display current-configuration** 命令显示信息。

2. 配置限制和指导

自动配置功能仅支持一级 PEX 组网，两级 PEX 组网请使用手工配置方式。

为保证 IRF3.1 协议正常运行和 PEX 正常上线，请注意：

- IRF3.1 相关配置命令会自动下发到级联接口上，请勿在级联接口上进行任何其他业务的配置。
- 级联轴接口的物理端口上除了 **lldp enable** 外，请勿配置其它命令。

关闭 IRF3.1 系统自动配置功能不会影响已经下发的配置。

3. 配置步骤

表1-6 开启 IRF3.1 系统自动配置功能

操作	命令	说明
进入系统视图	system-view	-
全局开启LLDP功能	lldp global enable	缺省情况下，全局LLDP功能处于关闭状态
进入连接一级 PEX的二层以太网接口视图	进入二层以太网接口视图 interface interface-type interface-number	在对多个接口进行相同配置时，通过接口批量配置视图可以更快完成操作
	进入一组接口的批量配置视图 interface range { interface-type interface-number [to interface-type interface-number] } &<1-24>	
开启接口的LLDP功能	lldp enable	缺省情况下，LLDP功能在接口上处于开启状态
退回系统视图	quit	-
开启IRF3.1系统自动配置功能	pex auto-config enable	缺省情况下，IRF3.1系统自动配置功能处于关闭状态

1.11 将PEX连接到上级设备

请按照拓扑规划将 PEX 上行接口的物理端口连接到上级设备（上级 PEX 或父设备）级联轴接口的物理端口。

PEX 上行接口的物理接口请根据“[1.2.4 上行接口的物理端口选择](#)”中所示方法选择。连接要求请参见“[1.1.3 IRF3.1 物理连接拓扑](#)”。

1.12 开启PEX设备本地转发功能

1. 功能简介

开启 PEX 本地转发功能后，PEX 设备对二层已知单播报文进行本地转发，其它报文转发给父设备处理。

PEX 本地转发功能处于关闭状态时，PEX 将收到的所有报文都转发给父设备，父设备进行转发处理。

开启/关闭 PEX 的本地转发功能后，PEX 及其下级 PEX 设备会先下线然后重新上线。

2. 配置步骤

表1-7 开启 PEX 设备本地转发功能

操作	命令	说明
进入系统视图	system-view	-
进入连接PEX设备的二层聚合接口视图	interface bridge-aggregation interface-number	-
开启PEX设备本地转发功能	pex local-forwarding	缺省情况下，PEX本地转发功能处于关闭状态

1.13 配置PEX设备下线后的报文转发功能

1. 功能简介

PEX 设备下线后的报文转发功能处于开启状态时：PEX 设备下线后，不清除运行数据，不关闭业务端口。PEX 可以继续对报文进行本地转发。

PEX 设备下线后的报文转发功能处于关闭状态时：PEX 设备下线后会清除运行数据，将所有业务端口设置为 DOWN 状态，不进行任何报文转发。但上行接口的物理端口不会被本功能设置为 DOWN 状态，如其仍在开启状态下，会继续发送 IRF3.1 协议报文尝试恢复与父设备的连接。

2. 配置限制和指导

仅当 PEX 本地转发功能处于开启状态时，本功能才能生效。

3. 配置步骤

表1-8 开启 PEX 设备本地转发功能

操作	命令	说明
进入系统视图	system-view	-
进入连接PEX设备的二层聚合接口视图	interface bridge-aggregation interface-number	-
开启PEX设备下线后的报文转发功能	pex persistent-forwarding	缺省情况下，PEX设备下线后的报文转发功能处于关闭状态

1.14 从父设备上登录到PEX设备

1. 功能简介

登录到 PEX 设备后，可以执行以下命令：

- **display** 命令。
- 文件系统管理相关命令（具体命令可以通过 **display role feature name filesystem** 命令查看）。有关文件系统管理相关命令的详细介绍，请参见“基础配置命令参考”中的“文件系统管理”。有关 **display role feature** 命令的详细介绍，请参见“基础配置命令参考”中的“RBAC”。
- PEX 独立支持特性。支持情况与 PEX 设备型号有关，请以设备的实际情况为准，例如支持 PoE 功能的 PEX 设备开启 PoE 功能。

2. 配置步骤

表1-9 从父设备上登录到 PEX 设备

操作	命令	说明
进入系统视图	system-view	-
从父设备上登录到PEX设备	switchto pex chassis chassis-number	在PEX设备上执行 quit 可以退出登录

1.15 在父设备上保存PEX独立支持特性的配置文件

1. 功能简介

配置 PEX 独立支持特性需要使用 **switchto pex** 命令登录到 PEX 设备上配置。完成配置后，请返回到父设备执行 **save pex-configuration** 命令进行配置保存。父设备侧执行该命令后，下发保存配置文件的指令给 PEX，使 PEX 生成配置文件，并通过 FTP 服务获取该配置文件。父设备收到 PEX 的配置文件后，将配置文件同步保存到所有成员设备的所有主控板文件系统根目录的 **pex_images** 文件夹中。PEX 上线或重启时会使用该配置文件进行配置恢复。对于非 PEX 独立支持的功能，配置仍然保存在父设备的配置文件中。

PEX 独立支持的特性与产品型号有关，请以设备的实际情况为准。

2. 配置限制和指导

请您在完成 PEX 独立支持特性的配置或配置修改后立即返回到父设备上配置 PEX 独立支持特性的配置保存。否则，PEX 设备下线再重新上线后，PEX 独立支持特性的配置可能被父设备上当前保存的配置覆盖。这是由于 PEX 设备下线时不清除 PEX 独立支持特性的配置，上线时使用父设备上的配置文件进行配置恢复，包括 PEX 独立支持特性的配置。配置恢复时如果出现父设备下发的配置与本地配置冲突，父设备下发的配置将覆盖本地配置。此时，如果配置结果与您的预期不符，需要重新登录 PEX 设备进行配置修改。

3. 配置步骤

从父设备上登录到 PEX 设备后，可配置 PEX 独立支持特性，具体配置方法请参见 PEX 设备配置指导和命令参考对应特性的手册。配置完成后需要在父设备上保存 PEX 独立支持特性。

表1-10 在父设备上保存 PEX 独立支持特性的配置文件

操作	命令	说明
进入系统视图	system-view	-
从父设备上登录到PEX设备	switchto pex chassis <i>chassis-number</i>	-
配置PEX独立支持特性	具体配置方法请参见PEX设备配置指导和命令参考对应特性的手册	-
退回系统视图	quit	-
在父设备上保存PEX独立支持特性的配置文件	save pex-configuration	本命令仅保存PEX独立支持功能的配置，非PEX独立支持功能的配置仍然保存在父设备的配置文件中

1.16 删除空闲的级联接口

1. 功能简介

配置本功能后，如果当前有级联接口连接的 PEX 设备处于下线状态，系统会自动删除作为级联接口的聚合接口以释放资源。

2. 配置限制和指导

当被删除级联接口的 PEX 设备需要再次上线时，需要重新配置级联接口。如果 PEX 设备只是计划中的暂时下线，请勿配置此功能。

3. 配置步骤

表1-11 删除空闲的级联接口

操作	命令
进入系统视图	system-view
删除空闲的级联接口	pex idle-cascade delete

1.17 IRF3.1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 IRF3.1 的运行情况，通过查看显示信息验证配置的效果。

表1-12 IRF3.1 显示和维护

操作	命令
显示设备信息	display device
显示设备的电子标签信息	display device manuinfo
显示接口连接的PEX设备状态，以及接口 PE CSP协议状态信息。	display pex interface [<i>interface-name</i>] [brief]
显示PEX的拓扑信息	display pex topology
显示系统版本信息	display version



说明

有关 **display device**、**display device manuinfo**、**display version** 命令的详细介绍请参见“基础配置命令参考”中的“设备管理”。

1.18 IRF3.1典型配置举例

1.18.1 两级 PEX IRF3.1 典型配置举例

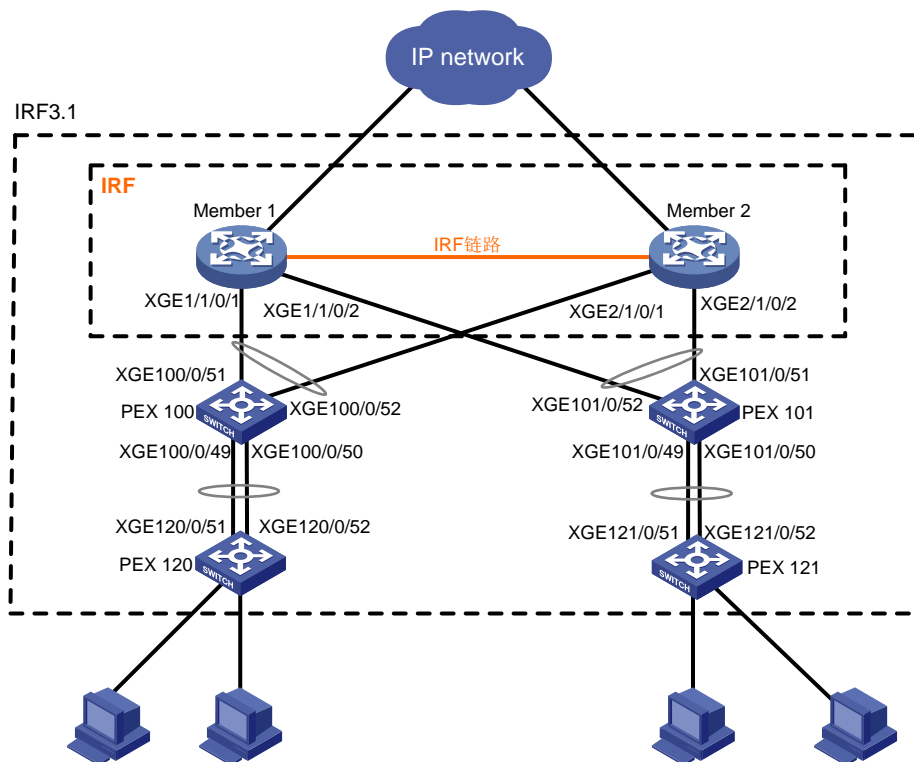
1. 组网需求

如[图 1-7](#)所示，构建如下 IRF3.1 系统：

- 两台设备组成 IRF。
- IRF 下连接两台 S5130S-52S-HI 交换机作为一级 PEX（PEX 100 和 PEX 101）。
- 两台一级 PEX 下再各自连接一台 S5130S-52S-HI 交换机作为二级 PEX（PEX 120 和 PEX 121）。

2. 组网图

图1-7 IRF3.1 典型配置组网图



3. 搭建 IRF 父设备

将 Member 1 和 Member 2 两台设备组成 IRF。关于 IRF 的配置请参见“虚拟化技术配置指导”中“IRF”的配置举例，此处步骤略。

4. 在父设备上配置连接一级 PEX 的级联接口

进入系统视图。

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

全局使能 LLDP。

```
[Sysname] lldp global enable
```

创建 PEX 组 1。

```
[Sysname] pex group 1
```

```
[Sysname-pex-group-1] quit
```

创建二层聚合接口 100 作为连接一级 PEX 100 的级联接口。（为方便记录对应关系，这里选择与 PEX 设备虚拟框号相同的聚合组编号）

```
[Sysname] interface bridge-aggregation 100
```

开启二层聚合接口 100 连接 PEX 的功能并将接口加入 PEX 组 1。

```

[Sysname-Bridge-Aggregation100] pex-capability enable group 1
# 为 PEX 分配虚拟框号 100。

[Sysname-Bridge-Aggregation100] pex associate chassis 100
[Sysname-Bridge-Aggregation100] quit
# 进入端口 Ten-GigabitEthernet1/1/0/1 和 Ten-GigabitEthernet2/1/0/1 的批量配置视图, 开启 LLDP
功能并将端口加入到聚合组 100 中。

[Sysname] interface range ten-gigabitethernet 1/1/0/1 ten-gigabitethernet 2/1/0/1
[Sysname-if-range] lldp enable
[Sysname-if-range] port link-aggregation group 100
[Sysname-if-range] quit
# 创建二层聚合接口 101 作为连接一级 PEX 101 的级联接口。

[Sysname] interface bridge-aggregation 101
# 开启二层聚合接口 101 连接 PEX 的功能并将接口加入 PEX 组 1。

[Sysname] interface bridge-aggregation 101
[Sysname-Bridge-Aggregation101] pex-capability enable group 1
# 为 PEX 分配虚拟框号 101。

[Sysname-Bridge-Aggregation101] pex associate chassis 101
[Sysname-Bridge-Aggregation101] quit
# 进入端口 Ten-GigabitEthernet1/1/0/2 和 Ten-GigabitEthernet2/1/0/2 的批量配置视图, 开启 LLDP
功能并将端口加入到聚合组 101 中。

[Sysname] interface range ten-gigabitethernet 1/1/0/2 ten-gigabitethernet 2/1/0/2
[Sysname-if-range] lldp enable
[Sysname-if-range] port link-aggregation group 101
[Sysname-if-range] quit

```

5. 配置一级 PEX 设备

对于 PEX 设备, 只需要配置 PEX 设备的工作模式为 `auto` 或 `pex`。下面以 PEX 100 为例给出 PEX 设备的配置, PEX 101 的配置与 PEX 100 完全一致, 此处不再赘述。

- 采用 `auto` 模式

将设备工作模式配置为 `auto` (设备缺省即处于 `auto` 模式, 如果未修改过 PEX 工作模式, 此步骤可省略)。

```

<Sysname> system-view
[Sysname] pex system-working-mode auto
# 保存当前配置。

[Sysname] save
The current configuration will be written to the device. Are you sure? [Y/N]:y
Please input the file name(*.cfg)[flash:/startup.cfg]
(To leave the existing filename unchanged, press the enter key):
flash:/startup.cfg exists, overwrite? [Y/N]:y
Validating file. Please wait...
Saved the current configuration to mainboard device successfully.

```

查看 PEX 设备配套资料确定上行接口的候选成员端口, 这里我们选择 XGE1/0/51 和 XGE1/0/52 作为上行接口的成员端口, 按照图 1-7 完成父设备与 PEX 100 设备之间的物理连线。

- 采用 **pex** 模式
 - # 将设备工作模式配置为 **pex**。

```
<Sysname> system-view
[Sysname] pex system-working-mode pex
```

 - # 保存当前配置。

```
[Sysname] save
[Sysname] quit
```

 - # 重新启动设备。

```
<Sysname> reboot
```

 - # 查看 **PEX** 设备上行接口的候选成员端口。

```
<Sysname> system-view
[Sysname] probe
[Sysname-probe] display system internal pex upstreamport
```

IfIndex	Interface	LldpPktRcvd	Priority	PeerPort-PE-CSP-Addr
51	XGE1/0/51	No	255	ffff-ffff-ffff
52	XGE1/0/52	No	255	ffff-ffff-ffff

 - # 本例中选择 **XGE1/0/51** 和 **XGE1/0/52** 作为上行接口的成员端口。请按照[图 1-7](#)完成父设备与 **PEX 1** 设备之间的物理连线。

6. 在父设备上配置连接二级 **PEX** 的级联接口

- # 创建 **PEX** 二层聚合接口 **120** 作为连接二级 **PEX 120** 的级联接口。

```
[Sysname] interface bridge-aggregation 120 pex
[Sysname-Bridge-Aggregation120] quit
```

- # 进入端口 **Ten-GigabitEthernet100/1/0/49** 和 **Ten-GigabitEthernet100/1/0/50** 的批量配置视图，开启 **LLDP** 功能并将端口加入到聚合组 **120** 中。

```
[Sysname] interface range ten-gigabitethernet 100/1/0/49 ten-gigabitethernet 100/1/0/50
[Sysname-if-range] lldp enable
[Sysname-if-range] port link-aggregation group 120
[Sysname-if-range] quit
```

- # 开启二层聚合接口 **120** 连接 **PEX** 的功能并将接口加入 **PEX** 组 **1**。

```
[Sysname] interface bridge-aggregation 120
[Sysname-Bridge-Aggregation120] pex-capability enable group 1
```

- # 为 **PEX** 分配虚拟框号 **120**。

```
[Sysname-Bridge-Aggregation120] pex associate chassis 120
[Sysname-Bridge-Aggregation120] quit
```

- # 创建 **PEX** 二层聚合接口 **121** 作为连接二级 **PEX 121** 的级联接口。

```
[Sysname] interface bridge-aggregation 121 pex
[Sysname-Bridge-Aggregation121] quit
```

- # 进入端口 **Ten-GigabitEthernet101/1/0/49** 和 **Ten-GigabitEthernet101/1/0/50** 的批量配置视图，开启 **LLDP** 功能并将端口加入到聚合组 **121** 中。

```
[Sysname] interface range ten-gigabitethernet 101/1/0/49 ten-gigabitethernet 101/1/0/50
[Sysname-if-range] lldp enable
[Sysname-if-range] port link-aggregation group 121
[Sysname-if-range] quit
```

开启二层聚合接口 121 连接 PEX 的功能并将接口加入 PEX 组 1。

```
[Sysname] interface bridge-aggregation 121
[Sysname-Bridge-Aggregation121] pex-capability enable group 1
```

为 PEX 分配虚拟框号 121。

```
[Sysname-Bridge-Aggregation121] pex associate chassis 121
[Sysname-Bridge-Aggregation121] quit
```

7. 配置二级 PEX 设备

与“[1.18.1 5. 配置一级 PEX 设备](#)”完全一致，此处不再赘述。

8. 验证配置

当 PEX 设备上线后，在父设备上显示设备信息，可以看到父设备和所有的 PEX 设备信息。

```
<Sysname> display device
```

Slot	Type	State	Subslot	Soft Ver	Patch Ver
1/0	LSQM3MPUB0	Master	0	S7500E-7536	None
1/1	NONE	Absent	0	NONE	None
1/2	LSQM2GT24PTSSCO	Normal	0	S7500E-7536	None
1/3	NONE	Absent	0	NONE	None
1/4	NONE	Fault	0	NONE	None
1/5	NONE	Absent	0	NONE	None
1/6	NONE	Absent	0	NONE	None
1/7	NONE	Absent	0	NONE	None

```
PEX in virtual chassis : 100
```

Slot	Type	State	Subslot	Soft Ver	Patch Ver
1	S5130S-52S-HI	Master	0	S5130S-HI6801L01	None

```
PEX in virtual chassis : 101
```

Slot	Type	State	Subslot	Soft Ver	Patch Ver
1	S5130S-52S-HI	Master	0	S5130S-HI6801L01	None

```
PEX in virtual slot : 120
```

Slot	Type	State	Subslot	Soft Ver	Patch Ver
1	S5130S-52S-HI	Master	0	S5130S-HI6801L01	None

```
PEX in virtual slot : 121
```

Slot	Type	State	Subslot	Soft Ver	Patch Ver
1	S5130S-52S-HI	Master	0	S5130S-HI6801L01	None

显示 PEX 拓扑信息。

```
<Sysname> display pex topology
```

```
Group 1
  Tier 1
    PEX 100 -----> Parent
    PEX 101 -----> Parent
  Tier 2
    PEX 120 -----> PEX 100
```

PEX 121 -----> PEX 101