

IRF切换DRNI组网操作指南

H3C交换机技术专题篇



我看了上期的DRNI升级，GIR升级好方便好丝滑呀，可惜咱这儿还都是IRF，升级需要以迅雷不及掩耳之势插拔堆线，有一点麻烦，也有一点风险，难过(づ￣)￣(づ)

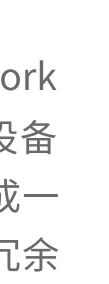


宝爸心里话

真慌真慌，全体工程师，IRF切换DRNI组网操作指南，来咯来咯~



拿捏了



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。

DRNI (Distributed Resilient Network Interconnect，分布式弹性网络互连)是一种跨设备链路聚合技术，将两台物理设备在聚合层面虚拟成一台设备来实现跨设备链路聚合，从而提供设备级冗余保护和流量负载均衡。

如下图表格，IRF和DRNI对比，组网可靠性要求高、升级过程要求业务中断时间短的场景推荐使用DRNI。

项目	IRF	DRNI
控制面	所有成员设备控制面统一，集中管理 所有成员设备需要同步所有表项	两台独立设备，控制平面解耦 主要同步MAC/ARP表项
设备面	紧耦合 硬件要求：芯片架构相同，一般要求同系列 软件要求：必须相同版本	松耦合 硬件要求：支持不同型号 软件要求：支持不同版本（由于DRNI的特性支持情况还在快速发展阶段，现阶段部分产品要求同版本）
版本升级	需要成员设备同步升级，或者主设备、从设备分开升级但操作较复杂 升级时业务中断时间2s左右	可独立升级，升级时业务中断时间小于1s 对于支持Graceful Transition and Renewal，平滑插入和移除的设备，可以做到不中断，关于使用GIR进行DRNI系统升级的更多介绍，请参见“DRNI升级指南”
配置管理类	统一配置、统一管理、操作简单 耦合度高，和控制器配合存在单点故障可能	独立配置，DR系统会进行配置一致性检查，具体业务配置需基于工保证 独立管理，耦合度低，和控制器配合使用不存在单点故障，可靠性更高

那如果现网运行的是IRF，可以直接切换到DRNI吗？当然可以啦~今天我们就一起来看下IRF切换DRNI组网的操作流程和注意事项，主要讲解underlay IRF组网，设备三层转发场景下切换DRNI组网的步骤。（二层转发或overlay EVPN组网切换步骤类似，只是具体的三层配置有所差异，这里不作展开介绍，有需要详见《H3C+IRF组网切换DRNI组网指导书》）

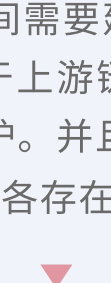


组网介绍

如下图所示，原本SWA/B和SWC/D分别是由两台成员设备组成的IRF，上下行均是跨IRF设备的动态聚合，SWA/B设备为网关节三层转发设备。为了提高SWA/B的可靠性和方便升级，将SWA/B的IRF系统切换为DR系统，SWA/B的上下行的聚合链路切换为DR接口。



下滑切换DRNI组网



切换注意事项

#1 对于框式设备，配置DRNI之前，需要先执行undo chassis convert mode命令将设备由IRF模式切换为独立运行模式。

设备进行模式切换后接口索引会发生变化（包括VLAN接口、聚合接口、VSI接口等）。ADNET解决方案等涉及接口索引的应用请注意处理。对于盒式设备，删除IRF配置、进行DRNI配置的过程中，如果没有删除过对应接口，接口索引不会变。

#2 检查业务是否存在单挂接入，如果业务对中断不敏感，可以先中断业务，将设备切换成DRNI后，再恢复业务；如果业务对中断敏感，需要先将单归接入改造成双归接入，再按照本文描述的方式切换。

#3 设备配置DRNI时，会占用一定数量的ACL资源。切换DRNI前请计算设备剩余ACL资源的数量是否足够DRNI使用，如果ACL资源不足，需要释放部分ACL资源。

#4 检查IRF设备桥MAC地址配置，进行IRF切换DRNI前，配置undo irf mac-address persistent命令和undo irf mac-address命令（若有），使IRF分裂后桥MAC立即变化为当前设备的桥MAC地址。

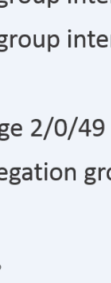
除了上述注意事项之外，建议进行如下配置，以便加快IRF切换DRNI组网过程中网络的收敛时间。

● 将Keepalive口、DR接口所属VLAN的VLAN接口、两台DR设备之间的三层互联接口配置为DRNI保留接口

● 一般情况下建议将DRNI延迟恢复时间（drni restore-delay）配置为300秒，如果ARP/MAC表项较大，可以再酌情调大。S12500X-AF的ARP表项接近48K规格时，延迟时间需要配置为900秒

● 将MAC地址老化时间调整为最大值（mac-address timer aging 2400）

● 先将IRF从设备切换为DR设备，再将IRF主设备切换为DR设备，减少一次IRF设备倒换



切换准备及流程

切换前请先准备好如下配置，以便切换过程中快速下发配置。

● DRNI系统配置，通过配置角色优先级确定设备主、从关系。

● IRF物理链路修改为IPL链路，IPL链路允许所有VLAN通过。

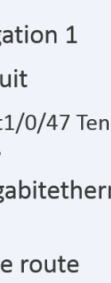
● 跨IRF成员设备的聚合接口修改为DR接口，DR接口关闭STP功能。

● 两台DR设备间配置Keepalive链路。

● 由于IRF切换为DRNI后，组网中的逻辑设备由一台变成两台，因此涉及为DR口上的业务做三层转发的接口均需要配置VRRP。VRRP配置不抢占模式，VRRP主从和DRNI主从保持一致。由于VRRP配置至少需要3个地址（主设备三层接口地址、从设备三层接口地址、VRRP虚拟IP地址），所以可能需要调整网段和IP地址的设计。建议VRRP虚拟IP地址配置为切换前三层接口的MAC。

● IRF设备配置了动态路由时，IRF切换DRNI后，组成IRF的两台设备成为独立设备，切换前IRF设备上配置的路由ID出现冲突，重新配置路由由协议Router ID，并重启路由进程。

● 使用动态路由实现DR设备与上行设备的路由互通时，DR主、从设备之间需要建立三层路由，打通主、从设备的三层互联，用于上游链路故障的业务保护，以及DRNI故障时的业务保护。并且需要确保上行设备到DR系统的两台设备之间至少各存在一条等价路由。



具体切换过程

A 关闭从设备业务物理接口

● 关闭从设备所有业务物理接口（除了IRF物理接口），流量切到主设备。

```
[SwitchB] interface range ten-gigabitethernet 2/0/1 ten-gigabitethernet 2/0/2 ten-gigabitethernet 2/0/45 ten-gigabitethernet 2/0/47 ten-gigabitethernet 2/0/48
[SwitchB-if-range] shutdown
[SwitchB] undo interface bridge-aggregation 1
[SwitchB-if-range] quit
```

● 主设备关闭IRF物理接口，IRF分裂。

```
[SwitchA] interface range fortygige 1/0/49 fortygige 1/0/50
[SwitchA-if-range] shutdown
[SwitchA-if-range] quit
```

B 从设备配置DRNI

● 配置IRF设备桥MAC恢复缺省情况，配置IRF桥MAC不保留，避免分裂后的两台设备具有相同MAC导致网络故障。

```
# 配置IRF桥MAC不保留。
[SwitchB] undo irf mac-address persistent
# 配置IRF设备桥MAC恢复缺省情况。
[SwitchB] undo irf mac-address
```

● DRNI系统配置

```
# 配置延迟恢复定时器超时时间为300秒。
[SwitchB] drni restore-delay 300
# 配置DR设备的角色优先级为100，使本设备作为DR系统主设备。
[SwitchB] drni role priority 100
# 配置DR系统MAC地址（同一DR系统的DR设备配置相同的系统MAC）。
[SwitchB] drni system-mac 0001-0001-0001
# 配置DR系统编号（同一DR系统的DR设备配置不同的系统编号）。
[SwitchB] drni system-number 1
# 配置DR系统优先级（同一DR系统的DR设备配置相同的系统优先级）。
[SwitchB] drni system-priority 123
# 配置Keepalive接口的目的IPv4地址为1.1.1.1，源IPv4地址为1.1.1.1。
[SwitchB] drni keepalive ip destination 1.1.1.1 source 1.1.1.1
```

● 配置DR口，配置连接下层设备的聚合200和连接DCGW的聚合接口201加入分布式聚合组，关闭DR口的STP功能。

```
[SwitchB] interface Bridge-Aggregation 200
[SwitchB-Bridge-Aggregation200] port drni group 200
[SwitchB-Bridge-Aggregation200] undo stp enable
[SwitchB] interface Bridge-Aggregation 201
[SwitchB-Bridge-Aggregation201] port drni group 201
[SwitchB-Bridge-Aggregation201] undo stp enable
[SwitchB-Bridge-Aggregation201] quit
```

● 配置VRRP，DR接口所属VLAN的VLAN接口11、40需要配置VRRP功能。将VLAN接口的原IP地址指定为VRRP备份组的虚拟IP，为VLAN接口指定同一网段的另一IP地址。配置设备在VRRP备份组中工作在非抢占模式。

```
[SwitchB] interface Vlan-interface 11
[SwitchB-Vlan-interface11] ip address 19.1.128.253 255.255.255.0
[SwitchB-Vlan-interface11] vrrp vrid 11 virtual-ip 19.1.128.254
[SwitchB-Vlan-interface11] undo vrrp vrid 11 preempt-mode
[SwitchB-Vlan-interface11] quit
[SwitchB] interface Vlan-interface 40
[SwitchB-Vlan-interface40] ip address 19.1.32.43 255.255.255.248
[SwitchB-Vlan-interface40] vrrp vrid 40 virtual-ip 19.1.32.44
[SwitchB-Vlan-interface40] undo vrrp vrid 40 preempt-mode
[SwitchB-Vlan-interface40] quit
```

● 配置IPL链路

```
# 删除切换前IRF设备上用于MAD检测的二层聚合接口1和2，重新创建二层聚合接口1。
[SwitchB] undo interface bridge-aggregation 1
[SwitchB] undo interface bridge-aggregation 2
[SwitchB] interface bridge-aggregation 1
[SwitchB-Bridge-Aggregation1] quit
# 配置IRF物理接口FortyGigE1/0/49和FortyGigE2/0/50取消与IRF接口绑定，并加入聚合组1作为IPL链路物理接口。
[SwitchB] irf-port 2/2
[SwitchB-irf-port2/2] undo port group interface fortygige 2/0/49
[SwitchB-irf-port2/2] undo port group interface fortygige 2/0/50
[SwitchB-irf-port2/2] quit
[SwitchB] interface range fortygige 2/0/49 fortygige 2/0/50
[SwitchB-if-range] port link-aggregation group 1
[SwitchB-if-range] quit
# 配置二层聚合接口1作为IPL接口。
[SwitchA] interface Bridge-Aggregation 1
[SwitchA-Bridge-Aggregation1] port link-type trunk
[SwitchA-Bridge-Aggregation1] port trunk permit vlan all
[SwitchA-Bridge-Aggregation1] link-aggregation mode dynamic
[SwitchA-Bridge-Aggregation1] port drni intra-portal-port 1
[SwitchA-Bridge-Aggregation1] quit
```

● 配置Keepalive链路接口

```
# 创建三层聚合接口1。
[SwitchA] undo interface vlan-interface 4094
[SwitchA] interface route-aggregation 1
[SwitchA-Route-Aggregation1] quit
# 配置物理接口Ten-GigabitEthernet1/0/47 Ten-GigabitEthernet2/0/48加入三层聚合组1作为Keepalive链路物理接口。
[SwitchA] interface range ten-gigabitethernet 1/0/47 ten-gigabitethernet 2/0/48
[SwitchA-if-range] port link-mode route
[SwitchA-if-range] port link-aggregation group 1
[SwitchA-if-range] quit
# 配置三层聚合接口1作为Keepalive接口，并配置IP地址。
[SwitchA] interface route-aggregation1
[SwitchA-Route-Aggregation1] link-aggregation mode dynamic
[SwitchA-Route-Aggregation1] ip address 1.1.1.1 255.255.255.0
[SwitchA-Route-Aggregation1] quit
```

● 配置保留接口

```
[SwitchB] drni mad exclude interface route-aggregation 1
[SwitchB] drni mad exclude interface ten-gigabitethernet 1/0/47
[SwitchB] drni mad exclude interface ten-gigabitethernet 2/0/48
[SwitchB] drni mad exclude interface vlan-interface 11
[SwitchB] drni mad exclude interface vlan-interface 40
```

C 开启从设备接口，关闭主设备接口

● 从设备所有业务物理接口和逻辑接口执行undo shutdown命令开启接口。

```
[SwitchB] interface range ten-gigabitethernet 2/0/1 ten-gigabitethernet 2/0/2 ten-gigabitethernet 2/0/45 ten-gigabitethernet 2/0/47 ten-gigabitethernet 2/0/48
[SwitchB-if-range] undo shutdown
[SwitchB-if-range] quit
```

● 主设备业务物理接口执行shutdown命令关闭接口，流量切到从设备。

```
[SwitchA] interface range ten-gigabitethernet 1/0/1 ten-gigabitethernet 1/0/2 ten-gigabitethernet 1/0/45 ten-gigabitethernet 1/0/47 ten-gigabitethernet 1/0/50
[SwitchA-if-range] shutdown
[SwitchA-if-range] quit
```

D 主设备配置DRNI

● 配置IRF设备桥MAC恢复缺省情况，配置IRF桥MAC不保留，避免分裂后的两台设备具有相同MAC导致网络故障。

```
# 配置IRF桥MAC不保留。
[SwitchA] undo irf mac-address persistent
# 配置IRF设备桥MAC恢复缺省情况。
[SwitchA] undo irf mac-address
```

● DRNI系统配置

```
# 配置延迟恢复定时器超时时间为300秒。
[SwitchA] drni restore-delay 300
# 配置DR系统MAC地址（同一DR系统的两台DR设备需要配置相同的系统MAC）。
[SwitchA] drni system-mac 0001-0001-0001
# 配置DR系统编号（同一DR系统的两台DR设备需要配置不同的系统编号）。
[SwitchA] drni system-number 1
# 配置DR系统优先级（同一DR系统的两台DR设备需要配置相同的系统优先级）。
[SwitchA] drni system-priority 123
# 配置Keepalive接口的目的IPv4地址为1.1.1.1，源IPv4地址为1.1.1.1。
[SwitchA] drni keepalive ip destination 1.1.1.1 source 1.1.1.1
```

● 配置DR口，配置连接下层设备的聚合200和连接DCGW的聚合接口201加入分布式聚合组。关闭DR口的STP功能。

```
[SwitchA] interface bridge-aggregation 200
[SwitchA-Bridge-Aggregation200] port drni group 200
[SwitchA-Bridge-Aggregation200] undo stp enable
[SwitchA] interface bridge-aggregation 201
[SwitchA-Bridge-Aggregation201] port drni group 201
[SwitchA-Bridge-Aggregation201] undo stp enable
[SwitchA-Bridge-Aggregation201] quit
```

● 配置VRRP，DR接口所属VLAN的VLAN接口11、40需要配置VRRP功能。将VLAN接口的原IP地址指定为VRRP备份组的虚拟IP，为VLAN接口指定同一网段的另一IP地址。配置设备在VRRP备份组中工作在非抢占模式。

```
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] ip address 19.1.128.252 255.255.255.0
[SwitchA-Vlan-interface11] vrrp vrid 11 virtual-ip 19.1.128.254
[SwitchA-Vlan-interface11] undo vrrp vrid 11 preempt-mode
[SwitchA-Vlan-interface11] quit
[SwitchA] interface vlan-interface 40
[SwitchA-Vlan-interface40] ip address 19.1.32.42 255.255.255.248
[SwitchA-Vlan-interface40] vrrp vrid 40 virtual-ip 19.1.32.44
[SwitchA-Vlan-interface40] undo vrrp vrid 40 preempt-mode
[SwitchA-Vlan-interface40] quit
```

● 配置IPL接口

```
# 删除切换前IRF设备上用于MAD检测的二层聚合接口1和2，重新创建二层聚合接口1。
[SwitchA] undo interface bridge-aggregation 1
[SwitchA] undo interface bridge-aggregation 2
[SwitchA] interface bridge-aggregation 1
[SwitchA-Bridge-Aggregation1] port link-type trunk
[SwitchA-Bridge-Aggregation1] port trunk permit vlan all
[SwitchA-Bridge-Aggregation1] link-aggregation mode dynamic
[SwitchA-Bridge-Aggregation1] port drni intra-portal-port 1
[SwitchA-Bridge-Aggregation1] quit
```

● 配置Keepalive链路

```
# 删除切换前IRF设备上用于MAD检测的VLAN接口4094，创建三层聚合接口1。
[SwitchA] undo interface vlan-interface 4094
[SwitchA] interface route-aggregation 1
[SwitchA-Route-Aggregation1] quit
# 配置物理接口Ten-GigabitEthernet1/0/47 Ten-GigabitEthernet2/0/48加入三层聚合组1作为Keepalive链路物理接口。
[SwitchA] interface range ten-gigabitethernet 1/0/47 ten-gigabitethernet 2/0/48
[SwitchA-if-range] port link-mode route
[SwitchA-if-range] port link-aggregation group 1
[SwitchA-if-range] quit
# 配置三层聚合接口1作为Keepalive接口，并配置IP地址。
[SwitchA] interface route-aggregation1
[SwitchA-Route-Aggregation1] link-aggregation mode dynamic
[SwitchA-Route-Aggregation1] ip address 1.1.1.1 255.255.255.0
[SwitchA-Route-Aggregation1] quit
```

● 配置保留接口

```
[SwitchA] drni mad exclude interface route-aggregation 1
[SwitchA] drni mad exclude interface ten-gigabitethernet 1/0/47
[SwitchA] drni mad exclude interface ten-gigabitethernet 2/0/48
[SwitchA] drni mad exclude interface vlan-interface 11
[SwitchA] drni mad exclude interface vlan-interface 40
```

E 开启主设备业务接口，完成DRNI系统切换

● 主设备IPL链路接口执行undo shutdown命令开启接口。

```
[SwitchA] interface range fortygige 1/0/49 fortygige 1/0/50 bridge-aggregation 1 [SwitchA-if-range] undo shutdown
[SwitchA-if-range] quit
```

● 主设备Keepalive链路接口执行undo shutdown命令开启接口。

```
[SwitchA] interface range ten-gigabitethernet 1/0/47 ten-gigabitethernet 1/0/48 route-aggregation 1
[SwitchA-if-range] undo shutdown
[SwitchA-if-range] quit
```

● 主设备所有业务物理接口和逻辑接口执行undo shutdown命令开启接口。

```
[SwitchA] interface range ten-gigabitethernet 1/0/1 ten-gigabitethernet 1/0/2 ten-gigabitethernet 1/0/45 ten-gigabitethernet 1/0/47 ten-gigabitethernet 1/0/50
[SwitchA-if-range] undo shutdown
[SwitchA-if-range] quit
```


到这里组网就切换成功啦，我们最后再一起来总结一下IRF切换DRNI组网的配置要点：

● 原IRF物理端口切换为IPL链路。

● 原MAD检测物理链路作为Keepalive链路。

● 连接上下行设备的跨IRF成员设备聚合接口切换为DR接口。

● DR口上需要使能三层接口的，均需要配置VRRP，VRRP配置不抢占模式。

● 使用动态路由实现DR设备与上行设备的路由互通时，DR主、从设备之间建立三层路由，打通主、从设备的三层互联，用于上游链路故障的业务保护，以及DRNI故障时的业务保护。

● IRF切换DRNI后，组成IRF的两台设备成为独立设备，切换前三层接口的IP地址出现冲突，需重新配置IP地址。

以上就是今天的全部内容，有问题或者建议记得联系我们哦~

—end—

扫码关注我们