【不推荐Reth口对Reth】某局点安全设备IRF组网标准拆堆叠升级导致业务中断

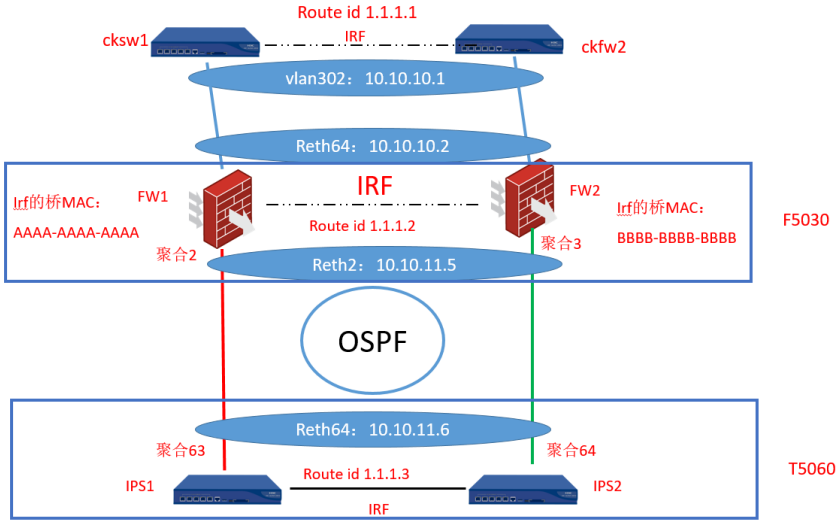
IRF　孔梦龙　2021-12-20 发表

组网及说明

组网如下图（隐去关键信息）：

FW主备标准组网，track上下行物理口；IPS主备标准组网，track上下行物理口。两者之间reth口对接reth口，起OSPF建立邻居发布路由。

FW拆堆叠以后，FW1和FW2上的桥MAC分别如下如所示。



配置步骤

（为保护客户隐私，本案例只描述现象，不展示具体的抓包文件和配置信息，shutdown和手工拔线效果等同）

目标是升级中间的防火墙F5030设备，操作过程是标准的拆堆叠升级；

（1）现场FW标准的IRF主备配置，IPS标准的IRF主备配置；

（2）拔掉FW的上行的所有业务口，业务切换至备框，IPS同步切换，OSPF正常建立；断掉堆叠线；

（3）升级主框的FW，确定主框状态正常；

（4）shutdown IPS备框上的所有与FW备框链接的口，插上FW上行的业务口；

（5）业务中断，OSPF邻居建立失败。OSPF停留的状态是：**ExStart。**

（6）流量再次切换至备框，业务恢复。

配置关键点

现象分析：

（1）业务中断的原因是OSPF邻居建立失败；故障的时候，在FW上和IPS同时抓包看OSPF报文的交互的状态，有如下的结果：

1）FW的主设备已经发了DB报文给IPS，源MAC是主防火墙的MAC，实例中的AAAA，目的MAC是IPS的reth64口的MAC，假设是DDDD。这个　　报文一直在重传，相当于IPS不回包。

2）在IPS抓包显示，收到了主防火墙源MAC是AAAA，目的MAC是DDDD的报文；IPS的回包源MAC是DDDD，目的MAC是BBBB

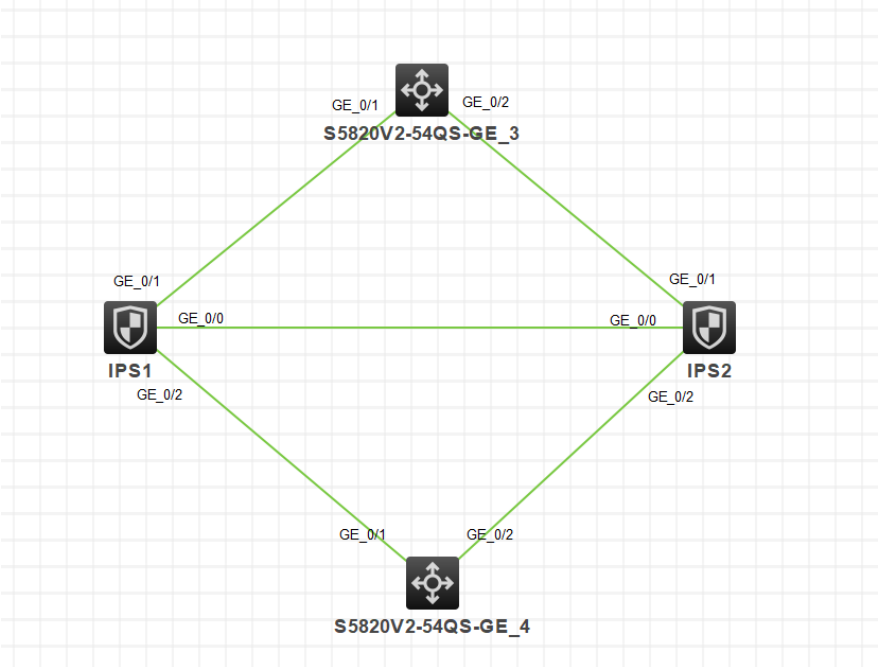
相当于IPS把报文回复给了备防火墙，导致OSPF邻居建立失败。

（2）分析IPS回复给备防火墙的原因是，目前IPS学习到的10.10.11.5的ARP仍然是BBBB。

原因分析：

（1）实验室模拟测试发现，FW拆堆叠以后，流量由备框切换至主框的时候，现场先shutdown IPS的备框的所有的上行口，此时IPS备框的冗余不发生切换，按现场的操作，此时再在主FW上插上业务口，此时发现IPS上的冗余倒回延时开始计时，默认60s，60s以后，IPS的冗余切换至IPS主框，OSPF正常建立，业务正常。

（2）实验室按现场情况简化如下的组网，上行的交换机模拟拆了堆叠以后的FW，下行的F1060模拟现场的IPS，就操作看冗余的状态看流量的走向。（reth 1的成员口是1/0/1和2/0/1，主框优先级高 ）



1）shutdown交换机的1/0/1口，此时IPS的冗余发生切换至备框，模拟流量在备框运行；

2）开始流量回切，此时先shutdown交换机上的1/0/2口；

此时在IPS查看冗余的状态（业务中断）：

```
<H3C>dis redundancy group
Redundancy group a (ID 1):
Preempt delay time remained     : 0    min
Preempt delay timer setting     : 1    min
Remaining hold-down time        : 0    sec
Hold-down timer setting         : 1    sec
Manual switchover request       : No

Member interfaces:

Redundancy group q (ID 2):
  Node ID      Slot          Priority   Status        Track weight
  1            Slot1         30         Secondary     -255
  2            Slot2         20         Primary       0
```

3）此时在SW上undo shutdown 1/0/1，目的是把流量引回到主框，发现IPS上的冗余节点的权重值增加到了255，冗余开始向主设备切换

```
<H3C>
<H3C>dis redundancy group
Redundancy group a (ID 1):
Preempt delay time remained     : 0    min
Preempt delay timer setting     : 1    min
Remaining hold-down time        : 0    sec
Hold-down timer setting         : 1    sec
Manual switchover request       : No

Member interfaces:

Redundancy group q (ID 2):
  Node ID      Slot          Priority   Status        Track weight
  1            Slot1         30         Secondary     255
  2            Slot2         20         Primary       0

Preempt delay time remained     : 1    min
Preempt delay timer setting     : 1    min
Remaining hold-down time        : 0    sec
Hold-down timer setting         : 1    sec
Manual switchover request       : No
```

此时冗余加口的情况，物理上up，但是状态上是inactive，也就是不响应报文。

```
<H3C>
<H3C>dis reth int reth
<H3C>dis reth int Reth 1
Reth1 :
  Redundancy group  : q
  Member           Physical status         Forwarding status   Presence status
  GE1/0/1          UP                      Inactive            Normal
  GE2/0/1          DOWN                    Active              Normal
<H3C>
```

60s以后，冗余组切换完成，reth口状态正常，业务正常

```
<H3C>dis reth int Reth 1
Reth1 :
  Redundancy group  : q
  Member           Physical status         Forwarding status   Presence status
  GE1/0/1          UP                      Inactive            Normal
  GE2/0/1          DOWN                    Active              Normal
<H3C>
<H3C>
<H3C>%Dec 20 17:35:12:097 2021 H3C RDDC/5/RDDC_ACTIVENODE_CHANGE: -Context=1; Redundancy group q active node changed to node 1 (slot 1), because of node's weight changed.

<H3C>dis reth int Reth 1
Reth1 :
  Redundancy group  : q
  Member           Physical status         Forwarding status   Presence status
  GE1/0/1          UP                      Active              Normal
  GE2/0/1          DOWN                    Inactive            Normal
```

综上来看，现场OSPF中断的时候，IPS上因为冗余正在切换，ARP没有刷新，即使主防火墙发了ARP报文到IPS得主设备，但是IPS得reth口此时不响应，导致邻居建立失败，业务中断。

规避措施：

（1）不建议Reth口对Reth口，升级时候不注意，会导致业务异常；

（2）如果遇到上述得问题，直接在IPS的响应的接口上清理一下ARP（现场清理的reth64）；

（3）如果是升级前，可以在IPS的冗余组下配置倒回时间延时，不使用默认的60s，IPS切换至主框时间短一点。待业务正常以后，再改回默认的值。 命令是：preempt-delay seconds XX （单位：s）。