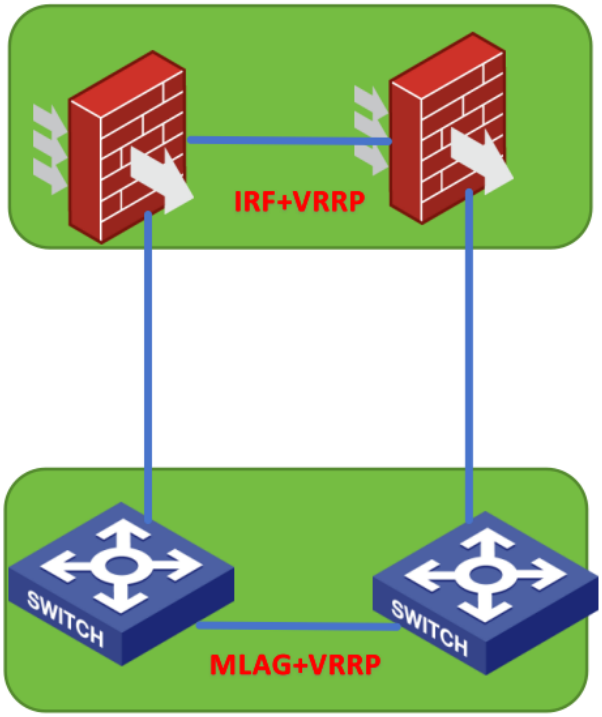
交换机 m-lag+vrrp直连irf防火墙不通问题

M-LAG/DRNI VRRP 陈泽勇 2024-05-06 发表

组网及说明

组网拓扑如下;



两台交换机M-LAG+VRRP单挂分别接入第三方防火墙IRF+VRRP。

问题描述

第三方防火墙做测试，IRF备防火墙用VRRP虚地址去ping交换机M-LAG备接口实地址发现无法ping通，备交换机主动去ping防火墙vrrp虚地址和接口实地址都无问题。

过程分析

1、让现场复现故障也就是备防火墙指导源地址为vrrp虚地址主动去ping交换机备接口实地址，查看arp防火墙那边都是稳定存在的。

2、在备交换机上debug ip packet和ip info、icmp等信息可以看到备交换机收到了对应的icmp报文，但是没有回包。

```
<FJFZ-MS-RADIUS-SW32>*Apr 28 20:48:32:870 2024 FJFZ-MS-RADIUS-SW32 IPFW/7/IPFW_
PACKET:
Receiving, interface = Vlan-interface538, version = 4, headlen = 20, tos = 0,
pktlen = 84, pktid = 1052, offset = 0, ttl = 255, protocol = 1,
checksum = 37286, s = 172.23.102.217, d = 172.23.102.222
channelID = 0, vpn-InstanceIn = 0, vpn-InstanceOut = 0.
prompt: Receiving IP packet.

*Apr 28 20:48:32:870 2024 FJFZ-MS-RADIUS-SW32 IPFW/7/IPFW_PACKET:
Delivering, interface = Vlan-interface538, version = 4, headlen = 20, tos = 0,
pktlen = 84, pktid = 1052, offset = 0, ttl = 255, protocol = 1,
checksum = 37286, s = 172.23.102.217, d = 172.23.102.222
channelID = 0, vpn-InstanceIn = 0, vpn-InstanceOut = 0.
prompt: IP packet is delivering up.

*Apr 28 20:48:32:870 2024 FJFZ-MS-RADIUS-SW32 SOCKET/7/ICMP:
ICMP Input:
ICMP Packet: src = 172.23.102.217, dst = 172.23.102.222
```

type = 8, code = 0 (echo)

*Apr 28 20:48:34:840 2024 FJFZ-MS-RADIUS-SW32 IPFW/7/IPFW_PACKET:
Delivering, interface = Vlan-interface538, version = 4, headlen = 20, tos = 0,
pktlen = 84, pktid = 1081, offset = 0, ttl = 255, protocol = 1,
checksum = 37257, s = 172.23.102.217, d = 172.23.102.222
channelID = 0, vpn-InstanceIn = 0, vpn-InstanceOut = 0.
prompt: IP packet is delivering up.

*Apr 28 20:48:34:840 2024 FJFZ-MS-RADIUS-SW32 SOCKET/7/ICMP:
ICMP Input:
ICMP Packet: src = 172.23.102.217, dst = 172.23.102.222
type = 8, code = 0 (echo)

3、此时查看备交换机的arp表项可以看到，m-lag备是通过peerlink口学到的防火墙vrrp虚地址arp，m-lag主是通过直连口1/0/15学到的vrrp虚地址，所以转发情况就是；防火墙备用虚地址去ping交换机m-lag备，交换机备收到了报文之后通过peerlink口转发给了m-lag主，m-lag主再转发给了防火墙的主。这样的话，防火墙就会认为来回路径不一致从而导致丢包。

m-lag备聚合1是peerlink

```
interface Bridge-Aggregation1
 description TO-[FJFZ-MS-RADIUS-SW31]-BAGG1
 port link-type trunk
 undo port trunk permit vlan 1
 port trunk permit vlan 528 538 568 588 608 628 648 738 748 758
 port trunk permit vlan 768
 link-aggregation mode dynamic
 port m-lag peer-link 1
 undo mac-address static source-check enable
#
```

Arp表项

```
----------------display arp all----------------
  Type: S-Static   D-Dynamic   O-Openflow   R-Rule   M-Multiport  I-Invalid
IP address      MAC address    VLAN/VSI name Interface               Aging Type
1.1.1.1         7c7a-3c5d-ebf4 --            RAGG1                   88    D
172.23.103.129  0000-5e00-0114 588           BAGG1                   432   D
172.23.103.130  7c7a-3c5d-ec0c 588           BAGG1                   584   D
172.23.103.132  30b9-3013-4446 588           BAGG1                   367   D
172.23.103.133  30b9-3013-444e 588           BAGG1                   637   D
172.23.103.134  30b9-3013-445e 588           BAGG1                   319   D
172.23.103.146  30b9-3013-4456 588           BAGG1                   81    D
172.23.103.147  30b9-3013-4466 588           BAGG1                   872   D
172.23.103.150  f4e9-7595-4c52 588           BAGG1                   874   D
172.23.103.152  10a4-da6f-e900 588           BAGG1                   793   D
172.23.104.130  7c7a-3c5d-ec08 648           BAGG1                   355   D
172.23.102.217  0000-5e00-0101 538           BAGG1                   1111  D
```

m-lag主arp是通过直连口学到的

```
  ================display arp all================
  Type: S-Static   D-Dynamic   O-Openflow   R-Rule   M-Multiport  I-Invalid
IP address      MAC address    VLAN/VSI name Interface               Aging Type
1.1.1.2         7c7a-3c5e-23f4 --            RAGG1                   325   D
172.23.103.131  7c7a-3c5e-240c 588           BAGG1                   822   D
172.23.103.151  f4e9-7595-46e2 588           BAGG1                   906   D
172.23.103.153  10a4-da6f-ebe0 588           BAGG1                   1057  D
172.23.104.131  7c7a-3c5e-2408 648           BAGG1                   592   D
172.23.104.138  30b9-3008-daa2 648           BAGG1                   450   D
172.23.104.139  30b9-3008-dab6 648           BAGG1                   667   D
172.23.104.140  30b9-3008-dab6 648           BAGG1                   107   D
172.23.102.219  10a4-da6f-ec0d 538           BAGG1                   1191  D
172.23.102.222  7c7a-3c5e-241a 538           BAGG1                   501   D
172.23.103.150  f4e9-7595-4c52 588           XGE1/0/21               1111  D
172.23.103.146  30b9-3013-4456 588           XGE1/0/8                638   D
172.23.103.147  30b9-3013-4466 588           XGE1/0/11               1151  D
172.23.103.134  30b9-3013-445e 588           XGE1/0/6                556   D
172.23.103.132  30b9-3013-4446 588           XGE1/0/2                604   D
172.23.103.133  30b9-3013-444e 588           XGE1/0/4                875   D
172.23.103.152  10a4-da6f-e900 588           XGE1/0/14               1061  D
192.168.18.2    7c7a-3c5e-19f4 --            XGE1/0/22               1105  D
172.23.102.217  0000-5e00-0101 538           XGE1/0/15               1137  D
```

4、 至于为什么m-lag备通过peerlink学到的因为vrrp主才响应arp报文，所以m-lag交换机arp请求防火墙vrrp虚地址广播arp报文的时候，仅有m-lag主直连的防火墙主会响应arp报文，虽然m-lag备也会收到arp广播报文并泛洪给防火墙的备，但是vrrp备不响应arp报文，后续m-lag主再将arp响应报文通过rlink封装发给m-lag备，所以m-lag备显示arp通过peerlink学到的。

解决方法

正常情况，无需关注。m-lag备收到报文之后，通过m-lag主回包了，防火墙认为来回路径不一致导致丢包。