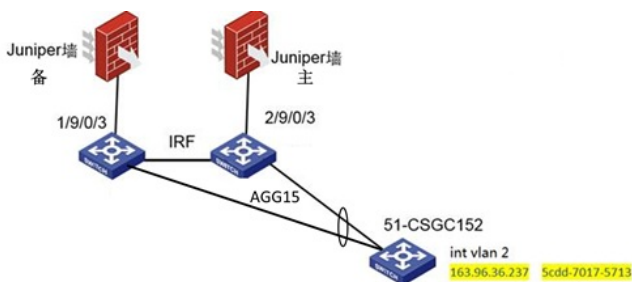


某局点S12508X-AF FX单板配置入方向学习命令后业务转发异常问题案例

二层转发 镜像 张文宁 2020-04-13 发表

组网及说明

拓扑：如下图所示，两台S12509X-AF堆叠作为网关，通过AGG15与下行的接入设备互联。



问题描述

问题：现场配置入方向mac学习命令后，导致下挂业务全部异常，与接入设备直连ping也时通时不通，现场取消命令后并重启后恢复。

过程分析

分析：查看故障时候表项，发现mac和arp的表项出端口在频繁变化，表项紊乱：

```
<GD002H3SW12-CSGE021>dis arp | i 163.96.36.237
163.96.36.237 5cdd-7017-5713 2 BAGG15 1077 D
<GD002H3SW12-CSGE021>dis arp | i 163.96.36.237
163.96.36.237 5cdd-7017-5713 2 GE2/9/0/3 1159 D
<GD002H3SW12-CSGE021>dis arp | i 163.96.36.237
163.96.36.237 5cdd-7017-5713 2 GE2/4/0/45 1159 D
```

```
<GD002H3SW12-CSGE021>dis mac-address | i 5713
5cdd-7017-5713 2 Learned GE2/9/0/3 Y
<GD002H3SW12-CSGE021>dis mac-address | i 5713
5cdd-7017-5713 2 Learned BAGG15 Y
<GD002H3SW12-CSGE021>dis mac-address | i 5713
5cdd-7017-5713 2 Learned GE2/4/0/45 Y
```

排查配置发现现场存在不规范的镜像配置（F系列板卡芯片限制，不支持跨框镜像）：

```
mirroring-group 1 local
#
防火墙互联口的配置
interface GigabitEthernet1/9/0/3
port link-mode bridge
description %To_GD000JF55-CS01 E0/0.35
port link-type trunk
port trunk permit vlan all
mirroring-group 1 mirroring-port both

interface GigabitEthernet2/9/0/3
port link-mode bridge
description %To_GD000JF55-CS02 E0/0.35
port link-type trunk
port trunk permit vlan all
mirroring-group 1 mirroring-port both
```

下连端口及镜像源配置

```
Aggregate Interface: Bridge-Aggregation15
Aggregation Mode: Dynamic
Loadsharing Type: Shar
Management VLANs: None
System ID: 0x8000, ac74-09c2-9e00
```

Local:

Port	Status	Priority	Index	Oper-Key	Flag
XGE1/4/0/43	S	32768	11	4	{ACDEF}
XGE2/4/0/43	S	32768	24	4	{ACDEF}

Remote:

Actor	Priority	Index	Oper-Key	SystemID	Flag
XGE1/4/0/43(R)	32768	50	1	0x8000, 5cdd-7017-5712	{ACDEF}
XGE2/4/0/43	32768	52	1	0x8000, 5cdd-7017-5712	{ACDEF}

```
interface Ten-GigabitEthernet1/4/0/43
port link-mode bridge
description %GD002H3SW51-CSGC152_G1/0/50
port link-type trunk
port trunk permit vlan all
mirroring-group 1 mirroring-port both
port link-aggregation group 15
```

```
interface Ten-GigabitEthernet2/4/0/43
port link-mode bridge
description %GD002H3SW51-CSGC152_G1/0/52
port link-type trunk
port trunk permit vlan all
mirroring-group 1 mirroring-port both
port link-aggregation group 15
```

镜像目的地配置

```
interface Ten-GigabitEthernet1/3/0/9
port link-mode bridge
flow-interval 5
mirroring-group 1 monitor-port
```

经确认，在这种错误的配置下，如果配置入方向mac学习，是会导致mac和arp学习出错的，具体如下：

关于mac学习：

报文首次从端口进来时会触发一个mac学习帧，包含了报文入端口等信息，会同步给芯片cpu，芯片cpu会同步给主控CPU。

入方向mac学习：报文进来的时候立刻从mac学习帧提取信息，学习到mac表。

出方向mac学习：报文出方向走到EPNI环节，说明该报文要从具体端口出去了，此时提取源mac并从mac学习帧中提取入端口等更多信息，学习到mac表。

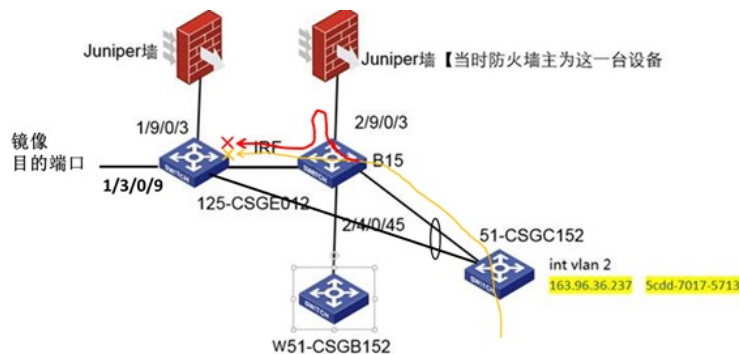
镜像转发原理：

正常非跨框本地镜像，走的是单播转发表。

配置跨框镜像后，芯片上会起用一个组播组，对镜像到的流量进行转发。流量在本框和转发到对框后找出端口，若找不到出端口报文会丢弃。

跨框镜像限制：

镜像报文跨框转发时要经过堆叠口，但芯片限制此时无法携带镜像标识和出端口信息，所以报文到目的端口所在的机框时查不到对应的出端口，报文被组播丢弃。



MAC漂移的产生过程：

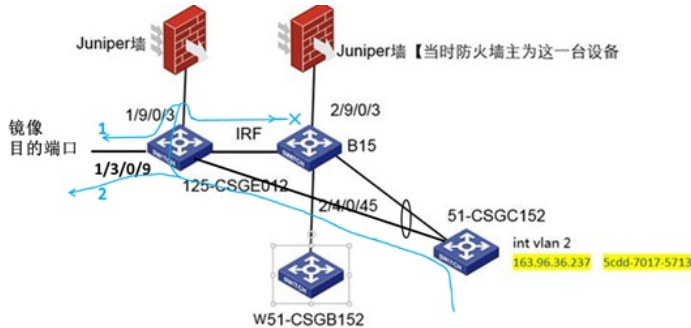
如上图现场的S12500X-AF配置了跨框镜像，源端口均为both方向，源端口包括与右下角的51-CSGC152互联的AGG15口、ten2/9/0/3、ten2/4/0/45口，目的端口在1框的ten1/3/0/9。

以右下角5cdd-7017-5713的mac报文为例子：

从橙色方向进来源mac为5cdd-7017-5713的报文，开启入端口mac学习后，此时mac正常学习在agg 15口。

由于AGG15配置了跨框镜像，设备会形成组播转发表来转发，其中一份从堆叠口转发到1框，另外有一份报文会去到2框所有的芯片上（如图该镜像报文会去到ten2/9/0/3所在芯片），不过因为2框没有具体目的出端口，报文只能转发到芯片的EGQ环节，无法到EPNI环节送出去。但是因为ten2/9/0/3配置了出方向镜像，到了EGQ环节的报文会被镜像送了回来（图中红色部分），于是这部分报文也要送往1框设备。这两份报文因为镜像标识和出端口信息的丢失，报文被丢弃，但是现场开启了入方向mac学习，红色部分报文导致设备学习该mac在ten2/9/0/3口。因此出现了该mac一会学习在ten2/9/0/3、一会学习在agg 15。同理，还会学习在其它一些配置了出方向镜像的端口入2/4/0/45口。

同样原因，该mac还会漂移到1框：



如图源mac流量为5cdd-7017-5713的蓝色流量从1框进，其中一份直接根据目的端口转发到ten1/3/0/9，还会复制一份到其它单板芯片，但是同样因为没有出端口而且配置了outbound的镜像，报文被重新转回来，其中一份转发给ten1/3/0/9，还有一份转发到2框，2框对这部分流量因为查表无出端口直接丢弃，同时因为开启了MAC入方向学习，导致该mac会从1/9/0/3学习到。

如果芯片上有大量的MAC和迁移速度较快时，软件存在感知不到的情况，这样的MAC迁移软件会存在记录不到的情况。

ARP表项学习紊乱过程：

现场的ARP报文学习错误原因同上，是由于ARP报文被错误的镜像导致学习端口出错，并且在mac学习紊乱的情况了，很容易出现MAC出端口和arp出端口不一致的情况，设备CPU会定时检查这种情况，如果发现就会认为当前的ARP错误了，会重新发ARP请求更新ARP，但由于现场错误的配置，更新正确后很快又会更新到错误的端口，因此出现时通时不通的现象。

定位结论：

1、现场存在不规范的镜像配置，F系列单板不支持跨框镜像，配合入方向mac学习后出现的异常，建议规范配置解决。

解决方法

解决方法：

去掉跨框镜像的配置。