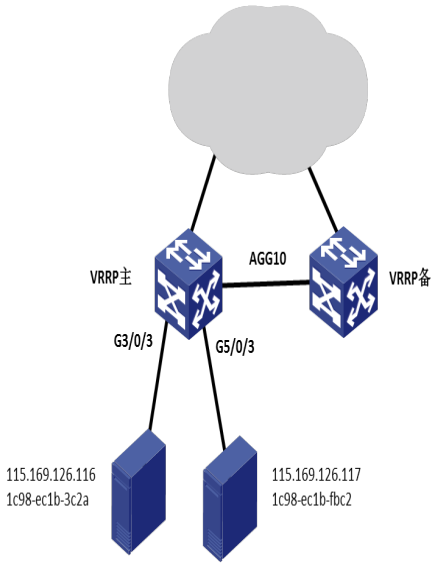
S12508F-AF VRRP组网流量异常泛洪问题案例

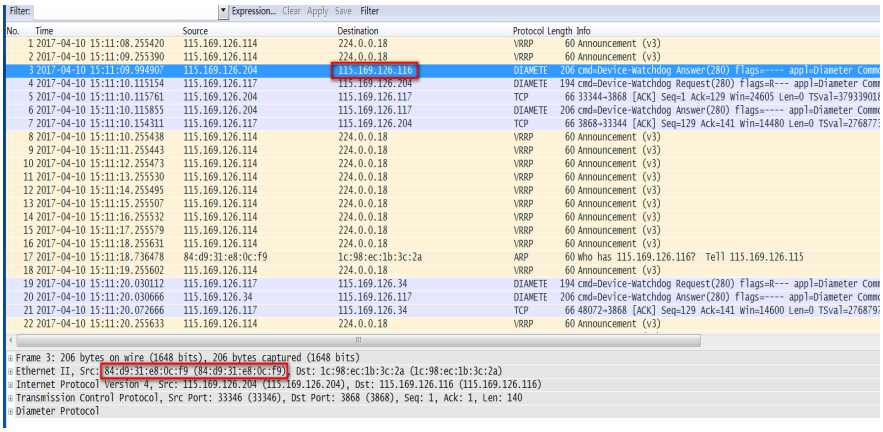
VRRP 陆世叶 2017-04-22 发表

组网拓扑：



如图，某局点两台S12508F-AF作为VRRP的主备，同时作为下行服务器的网关，两台服务器同接在VRRP主上，上行等价链路分别连接到VRRP主备。
客户表示之前在服务器侧抓包的时候，抓到了发往同网段的其他服务器的报文。故而查设备上的mac表和arp表的情况，发现有ARP表项，但是没有MAC表项，理论上来说这没关系，因为S12508F-AF是作为服务器的网关的，直接查ARP表项就可以。

1、开始是没有mac表的，第一次ping通后，查看有服务器的mac表，等待一个mac表老化时间后，再查看的时候，mac表消失，arp表正常，重新ping之后，mac表未生成。这个现象是正常的，这个S125X/S125F的一个机制：默认MAC地址学习是出接口学习，不是传统的在入接口学习源MAC。但从客户那边反馈的抓包来看，在115.169.126.117侧的服务器确实抓到了去往服务器115.169.126.116的报文，抓包截图如下：



No.	Time	Source	Destination	Protocol	Length	Info
1	2017-04-10 15:11:08.255420	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
2	2017-04-10 15:11:09.255390	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
3	2017-04-10 15:11:09.994907	115.169.126.204	115.169.126.116	DIAMETE	206	cmd=Device-Watchdog Answer(280) flags=---- appl=Diameter Comm
4	2017-04-10 15:11:10.115154	115.169.126.117	115.169.126.204	DIAMETE	194	cmd=Device-Watchdog Request(280) flags=R--- appl=Diameter Comm
5	2017-04-10 15:11:10.115761	115.169.126.204	115.169.126.117	TCP	66	33344→3868 [ACK] Seq=1 Ack=129 Win=24605 Len=0 TSval=37933901
6	2017-04-10 15:11:10.115855	115.169.126.204	115.169.126.117	DIAMETE	206	cmd=Device-Watchdog Answer(280) flags=---- appl=Diameter Comm
7	2017-04-10 15:11:10.154311	115.169.126.117	115.169.126.204	TCP	66	3868→33344 [ACK] Seq=129 Ack=141 Win=14480 Len=0 TSval=276877
8	2017-04-10 15:11:10.255438	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
9	2017-04-10 15:11:11.255443	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
10	2017-04-10 15:11:12.255473	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
11	2017-04-10 15:11:13.255530	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
12	2017-04-10 15:11:14.255495	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
13	2017-04-10 15:11:15.255507	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
14	2017-04-10 15:11:16.255532	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
15	2017-04-10 15:11:17.255579	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
16	2017-04-10 15:11:18.255631	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
17	2017-04-10 15:11:18.736478	84:d9:31:e8:0c:f9	1c:98:ec:1b:3c:2a	ARP	60	Who has 115.169.126.116? Tell 115.169.126.115
18	2017-04-10 15:11:19.255602	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)
19	2017-04-10 15:11:20.030112	115.169.126.117	115.169.126.34	DIAMETE	194	cmd=Device-Watchdog Request(280) flags=R--- appl=Diameter Comm
20	2017-04-10 15:11:20.030666	115.169.126.34	115.169.126.117	DIAMETE	206	cmd=Device-Watchdog Answer(280) flags=---- appl=Diameter Comm
21	2017-04-10 15:11:20.072666	115.169.126.117	115.169.126.34	TCP	66	48072→3868 [ACK] Seq=129 Ack=141 Win=14600 Len=0 TSval=276879
22	2017-04-10 15:11:20.255633	115.169.126.114	224.0.0.18	VRRP	60	Announcement (v3)

```
Frame 3: 206 bytes on wire (1648 bits), 206 bytes captured (1648 bits)
Ethernet II, Src: 84:d9:31:e8:0c:f9 (84:d9:31:e8:0c:f9), Dst: 1c:98:ec:1b:3c:2a (1c:98:ec:1b:3c:2a)
Internet Protocol Version 4, Src: 115.169.126.204 (115.169.126.204), Dst: 115.169.126.116 (115.169.126.116)
Transmission Control Protocol, Src Port: 33346 (33346), Dst Port: 3868 (3868), Seq: 1, Ack: 1, Len: 140
Diameter Protocol
```

2、从抓包来看，确实泛洪了，因为是单播流量，所有在VRRP主上这条流肯定是未知单播，否则不会泛洪。从上面的截图可以看到，泛洪报文的源MAC是84d9-31e8-0cf9，这个MAC是VRRP备对应vlan虚接口的MAC：
备上vlan 2016虚接口
Vlan-interface2016
Current state: UP
Line protocol state: UP
Description: pT:Media-DNS01AAA
Bandwidth: 10000000 kbps
Maximum transmission unit: 1500
Internet address: 115.169.126.115/28 (primary)
IP packet frame type: Ethernet II, hardware address:84d9-31e8-0cf9

IPv6 packet frame type: Ethernet II, hardware address:84d9-31e8-0cf9
Last clearing of counters: Never
去往115.169.126.116的流量从上行负载链路hash过来，从备进来的流量都广播了，而从主进来的流量都是单播转发出去。根据报文的源MAC就可以区别从主还是备进来的流量。

3、这里在回到组网上，两个服务器都是挂在VRRP主下面的，所有从VRRP备进来的流量查ARP表项发现出接口是互联的AGG10口，到VRRP主上在进行 二层转发，那因为没有MAC表项，所以出现泛洪的现象。

对于客户的这种组网，建议修改MAC地址学习方式传统入端口学习源MAC的方式，命令：mac-address mac-learning ingress。