

知 H3C S10500 VRRP配置相同VID导致业务中断问题

VRRP 彭绍穗 2014-10-29 发表

组网如图所示：设备2015及设备2016均是S10504，G7/0/1、G7/0/7、G7/0/47、G7/0/42均配置成route模式。两台设备之间建立VRRP。

10504作为网关设备，底下业务出现中断，且中断时候slave设备ping vrrp的虚地址不通，实地址没有问题。且将端口执行shutdown/undo shutdown后业务正常。

在实验室进行复现，将软件版本设置与现场相同，并将配置清空。并采用逐渐增加配置的方法分四步来复现问题。

为排除影响，首先我们分别对2015和2016上的G7/0/7和G7/0/42作shutdown处理，仅启用G7/0/1和G7/0/47两个端口，对这两个端口作如下配置，在每一端口上开启2个VRRP。

```
2015上端口G7/0/1配置
#
interface GigabitEthernet7/0/1
port link-mode route
flow-interval 30
ip address 211.151.131.34 255.255.255.240
ip address 211.151.131.50 255.255.255.240 sub
vrrp vrid 23 virtual-ip 211.151.131.33
vrrp vrid 23 priority 120
vrrp vrid 24 virtual-ip 211.151.131.49
vrrp vrid 24 priority 120
#
2016上端口G7/0/47配置
#
interface GigabitEthernet7/0/47
port link-mode route
flow-interval 30
ip address 211.151.131.35 255.255.255.240
ip address 211.151.131.51 255.255.255.240 sub
vrrp vrid 23 virtual-ip 211.151.131.33
vrrp vrid 23 priority 110
vrrp vrid 24 virtual-ip 211.151.131.49
vrrp vrid 24 priority 110
#
```

配置完成后查看VRRP状态如下：

```
[2015]dis vrrp ver
IPv4 Standby Information:
Run Mode : Standard
Run Method : Virtual MAC
Total number of virtual routers : 2
Interface GigabitEthernet7/0/1
VRID : 23 Adver Timer : 1
Admin Status : Up State : Master
Config Pri : 120 Running Pri : 120
Preempt Mode : Yes Delay Time : 0
Auth Type : None
Virtual IP : 211.151.131.33
Virtual MAC : 0000-5e00-0117
Master IP : 211.151.131.34
```

```
Interface GigabitEthernet7/0/1
VRID : 24 Adver Timer : 1
Admin Status : Up State : Master
Config Pri : 120 Running Pri : 120
Preempt Mode : Yes Delay Time : 0
Auth Type : None
Virtual IP : 211.151.131.49
Virtual MAC : 0000-5e00-0118
Master IP : 211.151.131.34
```

可见，VRRP工作正常。这里需要说明的是，对于105设备，虚拟mac地址的分配规则是0000-5e00-0100 + vrid，如vrid=24，则虚拟MAC地址为0000-5e00-0118。所以，如果在三层接口下为VRRP业务指定了相同的vrid，则分配的虚拟MAC地址是相同的。这是导致之后网络不通问题的条件之一。

此时，查看设备2015底层表项如下：

```
[2015-diagnose]bcm 7 0 l2/show
mac=00:23:89:d5:e7:ce vlan=4095 modid=18 port=0/cpu0 Static DHit L3 Group=L3Intf
mac=00:0f:e2:da:cb:c0 vlan=1 modid=18 port=0/cpu0 Static COS(src=7,dst=7) CPU ReplacePriority Group=Reserve
mac=00:0f:e2:da:cb:c2 vlan=1 modid=18 port=1/ge48 Static Group=Learnt
mac=00:00:5e:00:01:17 vlan=4095 modid=18 port=0/cpu0 Static DHit L3 Group=L3Intf
mac=00:00:5e:00:01:18 vlan=4095 modid=18 port=0/cpu0 Static DHit L3 Group=L3Intf
mac=00:23:89:d5:e7:ce vlan=0 modid=18 port=0/cpu0 Static L3 Group=L3Intf
```

可见，当VRRP生效时，对应于两个备份组（23和24），系统向芯片下发了两条表项，表示当设备收到目的MAC为0000-5e00-0117（备份组23）和0000-5e00-0118（备份组24）的报文时，走三层

转发流程。需要注意的是，两条表项vlan均为4095。这里需要说明，对于105v5设备，为节省vlan接口资源，路由口的设计实际是将所有路由口均加入到vlan 4095中，并采取端口隔离的措施。

此时，设备2015和2016无论实地址还是虚地址均是可通的。

在两个路由口上启用VRRP，配置相同的vrid。分别开启2015和2016上的G7/0/7和G7/0/42端口，并分别配置如下：

设备2015配置

```
#
interface GigabitEthernet7/0/7
port link-mode route
flow-interval 30
ip address 211.151.130.130 255.255.255.224
vrrp vrid 24 virtual-ip 211.151.130.129
vrrp vrid 24 priority 120
#
```

设备2016

```
#
interface GigabitEthernet7/0/42
port link-mode route
ip address 211.151.130.131 255.255.255.224
vrrp vrid 24 virtual-ip 211.151.130.129
vrrp vrid 24 priority 120
#
```

注意这里现场再次指定了vrid 24，这个vrid已经在之前指定过了。

此时，我们查看两台设备的VRRP状态：这里需要注意Interface GigabitEthernet7/0/1和Interface GigabitEthernet7/0/7下的虚拟mac是相同的，原因指定了相同的vrid 24，所以分配了相同的虚拟地址。

设备2015

```
[2015]dis vrrp ver
IPv4 Standby Information:
Run Mode : Standard
Run Method : Virtual MAC
Total number of virtual routers : 3
Interface GigabitEthernet7/0/1
VRID : 23 Adver Timer : 1
Admin Status : Up State : Master
Config Pri : 120 Running Pri : 120
Preempt Mode : Yes Delay Time : 0
Auth Type : None
Virtual IP : 211.151.131.33
Virtual MAC : 0000-5e00-0117
Master IP : 211.151.131.34
```

Interface GigabitEthernet7/0/1

```
VRID : 24 Adver Timer : 1
Admin Status : Up State : Master
Config Pri : 120 Running Pri : 120
Preempt Mode : Yes Delay Time : 0
Auth Type : None
Virtual IP : 211.151.131.49
Virtual MAC : 0000-5e00-0118
Master IP : 211.151.131.34
```

Interface GigabitEthernet7/0/7

```
VRID : 24 Adver Timer : 1
Admin Status : Up State : Master
Config Pri : 120 Running Pri : 120
Preempt Mode : Yes Delay Time : 0
Auth Type : None
Virtual IP : 211.151.130.129
Virtual MAC : 0000-5e00-0118
Master IP : 211.151.130.130
```

设备2016

```
[2016]dis vrrp ver
IPv4 Standby Information:
Run Mode : Standard
Run Method : Virtual MAC
Total number of virtual routers : 3
Interface GigabitEthernet7/0/42
VRID : 24 Adver Timer : 1
Admin Status : Up State : Backup
Config Pri : 120 Running Pri : 120
Preempt Mode : Yes Delay Time : 0
Become Master : 2650ms left
Auth Type : None
Virtual IP : 211.151.130.129
Master IP : 211.151.130.130
```

Interface GigabitEthernet7/0/47

```
VRID : 23 Adver Timer : 1
Admin Status : Up State : Backup
Config Pri : 110 Running Pri : 110
Preempt Mode : Yes Delay Time : 0
Become Master : 2600ms left
Auth Type : None
Virtual IP : 211.151.131.33
Master IP : 211.151.131.34
```

Interface GigabitEthernet7/0/47

```
VRID : 24 Adver Timer : 1
Admin Status : Up State : Backup
Config Pri : 110 Running Pri : 110
Preempt Mode : Yes Delay Time : 0
Become Master : 2500ms left
Auth Type : None
Virtual IP : 211.151.131.49
```

Master IP : 211.151.131.34

可见，虽然2个不同路由口指定了相同的vrid，但VRRP交互正常。而且，2015与2016无论实地址与虚地址均能ping通。从现象来看，没有任何问题。然而，当我们查看2015的I2表项时，出现问题。

```
[2015-diagnose]bcm 7 0 I2/show
mac=00:23:89:d5:e7:ce vlan=4095 modid=18 port=0/cpu0 Static L3 Group=L3Intf
mac=00:0f:e2:da:cb:c0 vlan=1 modid=18 port=0/cpu0 Static COS(src=7,dst=7) CPU ReplacePriority
Group=Reserve
mac=00:0f:e2:da:cb:c2 vlan=1 modid=18 port=1/ge48 Static Group=Learnt
mac=00:00:5e:00:01:17 vlan=4095 modid=18 port=0/cpu0 Static L3 Group=L3Intf
mac=00:00:5e:00:01:18 vlan=4095 modid=18 port=0/cpu0 Static L3 Group=L3Intf
mac=00:23:89:d5:e7:ce vlan=0 modid=18 port=0/cpu0 Static L3 Group=L3Intf
```

--- 6 mac address(es) found ---

我们开启了3个备份组，但这里只有2条表项，进一步，查看local logbuffer 7中有如下提示信息，该信息同样存在于现场的diag信息日志中。

```
Sep 26 2014 18:07:38:0423:
LINE:5074, File platform_bcm/drv/intf/l3_intf/drv_l3intf_intf.c-TASK:IPCD-FUNC::
The VRRP-MAC already exist!ullntfId=4097,g_pstDrvIntf_VMac[1].ullntfId=4096,vlan=4095,Mac=00-00-5e-00-01-18.
```

由于在第一步中，已经通过开始vrid 24的VRRP下发了vlan=4095,Mac=00-00-5e-00-01-18这条I2表项。所以，当我们在第二步中，在路由口G7/0/7 (vlan 4095) 中启用vrid 24的VRRP时，系统发现本次要下发的表项vlan=4095,Mac=00-00-5e-00-01-18已经存在于I2表中了，所以下发失败。根本原因就在于使用了路由口vlan=4095以及指定了相同的vrid，造成分配的mac地址相同。

然而，此时2015与2016的互通不存在问题。因为虽然表项下发失败，但I2表中已经存在了相同的表项，所以转发不会出现问题。比如：当设备2015收到目的mac为0000-5e00-0118的报文时，无论该报文是从G7/0/47发来，还是G7/0/42发来，都能够匹配正确的I2表项，从而进入转发流程。

在G7/0/42端口下进行shutdown操作。而后分别查看2015和2016的VRRP状态。发现VRRP状态均无异常。

```
[2015]dis vrrp
IPv4 Standby Information:
Run Mode : Standard
Run Method : Virtual MAC
Total number of virtual routers : 3
Interface VRID State Run Adver Auth Virtual
Pri Timer Type IP
-----
GE7/0/1 23 Master 120 1 None 211.151.131.33
GE7/0/1 24 Master 120 1 None 211.151.131.49
GE7/0/7 24 Initialize 120 1 None 211.151.130.129
```

```
[2016]dis vrrp
IPv4 Standby Information:
Run Mode : Standard
Run Method : Virtual MAC
Total number of virtual routers : 3
Interface VRID State Run Adver Auth Virtual
Pri Timer Type IP
-----
GE7/0/42 24 Initialize 120 1 None 211.151.130.129
GE7/0/47 23 Backup 110 1 None 211.151.131.33
GE7/0/47 24 Backup 110 1 None 211.151.131.49
```

此时，再从2016 ping 2015，发现211.151.131.49这个地址ping不通，而实地址211.151.131.50却能ping通。查看2016上的arp表，无异常。该现象与现场一致。

```
<2016>dis arp
Type: S-Static D-Dynamic M-Multiport
IP Address MAC Address VLAN ID Interface Aging Type
211.151.131.33 0000-5e00-0117 N/A GE7/0/47 N/A D
211.151.131.34 0023-89d5-e7ce N/A GE7/0/47 N/A D
211.151.131.50 0023-89d5-e7ce N/A GE7/0/47 N/A D
211.151.131.49 0000-5e00-0118 N/A GE7/0/47 N/A D
```

查看2015的I2table，发现问题：
[2015-diagnose]bcm 7 0 I2/show
mac=00:23:89:d5:e7:ce vlan=4095 modid=18 port=0/cpu0 Static DHIt L3 Group=L3Intf
mac=00:0f:e2:da:cb:c0 vlan=1 modid=18 port=0/cpu0 Static COS(src=7,dst=7) CPU ReplacePriority Group=Reserve
mac=00:0f:e2:da:cb:c2 vlan=1 modid=18 port=1/ge48 Static Group=Learnt
mac=00:00:5e:00:01:17 vlan=4095 modid=18 port=0/cpu0 Static DHIt L3 Group=L3Intf
mac=00:23:89:d5:e7:ce vlan=0 modid=18 port=0/cpu0 Static L3 Group=L3Intf
--- 5 mac address(es) found ---

I2table中剩一条对应于vrid 23的表项，之前存在的对应于vrid 24的表项不存在了。回顾操作，可知，当shutdown 端口G7/0/42时，对端接口G7/0/7也down掉了，导致G7/0/7承载的VRRP业务状态变为initialize，这种状态下，系统会删除之前下发过的I2表项。所以，系统理所当然的将I2table中的对应表项mac=00:00:5e:00:01:17 vlan=4095删除掉了。然而，这一删除动作却导致了依赖于该表项的G7/0/1开启的VRRP (vrid=23) 业务受到影响。当从2016上ping 211.151.131.33时，2016发出报文目的mac为0000-5e00-0118，设备2015收到该报文后查找I2table，找不到对应的I2表项，对报文做丢弃处理。

查看show/c的rdbgc3计数，可以明确得知报文丢弃原因为“找不到出口”。

```
[2015-diagnose]bcm 7 0 show/c/ge1
RUC.ge1 : 5 +5
RDBGc3.ge1 : 5 +5
GR127.ge1 : 5 +5
GRPkt.ge1 : 5 +5
GRBYT.ge1 : 510 +510
GRUC.ge1 : 5 +5
GRPok.ge1 : 5 +5
GT64.ge1 : 56 +44 1/s
GTPkt.ge1 : 56 +44 1/s
GTMCA.ge1 : 56 +44 1/s
GTBYT.ge1 : 3,584 +2,816 124/s
```

GTPOK.ge1 : 56 +44 1/s

在这种情况下，对G7/0/47进行shutdown和undo shutdown操作，可以恢复业务。原因为：随着2015设备上端口的开关，端口承载的vrid 24的VRRP业务经历了由initialize到master的变化，而当VRRP状态变化master时，会再次下发vrid 23和vrid 24的I2表项，所以业务得到了恢复。此时从2016上又能ping通2015的虚地址了。

```
[2015-diagnose]bcm 7 0 I2/show
mac=00:23:89:d5:e7:ce vlan=4095 modid=18 port=0/cpu0 Static L3 Group=L3Intf
mac=00:0f:e2:da:cb:c0 vlan=1 modid=18 port=0/cpu0 Static COS(src=7,dst=7) CPU ReplacePriority Group=Reserve
mac=00:0f:e2:da:cb:c2 vlan=1 modid=18 port=1/ge48 Static Group=Learnt
mac=00:00:5e:00:01:17 vlan=4095 modid=18 port=0/cpu0 Static L3 Group=L3Intf
mac=00:00:5e:00:01:18 vlan=4095 modid=18 port=0/cpu0 Static L3 Group=L3Intf
mac=00:23:89:d5:e7:ce vlan=0 modid=18 port=0/cpu0 Static L3 Group=L3Intf
--- 6 mac address(es) found ---
```

综上，对于105设备，虚拟mac地址的分配规则是0000-5e00-0100 + vrid，如vrid=24，则虚拟MAC地址为0000-5e00-0118。所以，如果在三层接口下为VRRP业务指定了相同的vrid，则分配的虚拟MAC地址是相同的。当配置相同的vid时系统发现本次要下发的表项vlan=4095,Mac=00-00-5e-00-01-18已经存在于I2表中了，所以下发失败。当承载VRRP转发业务的端口shutdown后，表项被删除，报文找不到表项出现转发不通的问题。根本原因就在于使用了路由口vlan=4095以及指定了相同的vrid，造成分配的mac地址相同。

将不同路由口的vrid配置成不同即可。