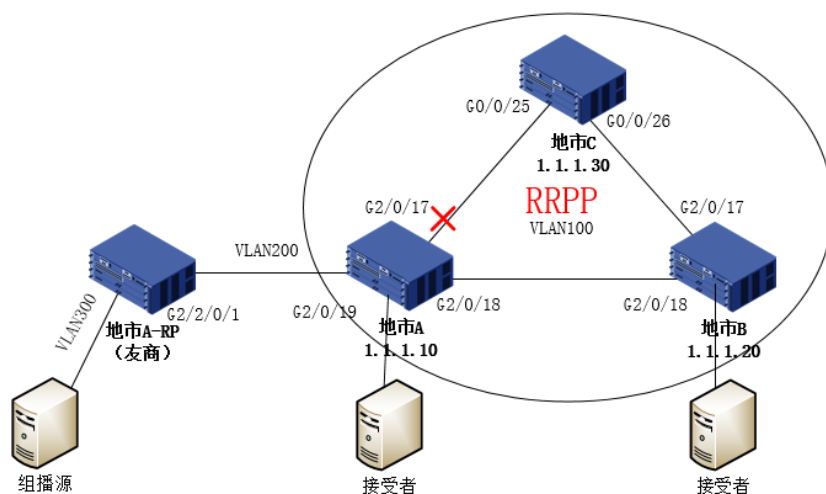


某局点RRPP环网中组播业务无法正常接收案例分析

PIM 刘凯 2017-03-15 发表



1、客户组网如上图：

地市A连接RP第三方厂商设备，RP设备连接组播源，发出2路组播信号分别为

(239.20.1.6和239.20.1.8)，在地市A及地市B都有相应的组播接受者。各个地市之间通过RRPP环网相互连接，地市A为主环主节点，保护VLAN为VLAN100，地市A的G2/0/18为主端口、G2/0/17为副端口（阻塞端口），地市B及地市C均为传输节点。

2、在地市A、B、C都与RP三层互通的情况下：

问题1：地市A、B在VLAN100内都无法接收到相应的组播数据，有（*, G）和（S, G）的表项，但是没有下行接口，除VLAN100外，在地市A上新建VLAN400即可接收到相应的组播数据。

问题2：在非VLAN100的其他VLAN内，地市B只有在地市A接收到相应组播组后，才可以正常接收组播。例如，在地市B上新建VLAN500请求组播，无法接收，此时在地市A上同样创建相应新的VLAN 400，请求组播组239.20.1.6的组播流量成功，则地市B的VLAN500内才可以收到相应的组播流量，而239.20.1.8组播组同样如此。

1、问题一：地市A、B在VLAN100内都无法接收到相应的组播数据分析，在地市A内VLAN100存在组播接收者的故障时刻，我们查看其pim routing-table表项，发现有（*, G）和（S, G）表项，但没有相应的出接口：

```
<DS-A>dis pim routing-table
```

```
Total 2 (*, G) entries; 2 (S, G) entries
```

```
(*, 239.20.1.6)
```

```
RP: 192.168.200.201
```

```
Protocol: pim-sm, Flag: WC
```

```
UpTime: 00:06:41
```

```
Upstream interface: Vlan-interface200
```

```
Upstream neighbor: 192.168.200.201
```

```
RPF prime neighbor: 192.168.200.201
```

```
Downstream interface(s) information: None
```

```
(1.1.1.2, 239.20.1.6)
```

```
RP: 192.168.200.201
```

```
Protocol: pim-sm, Flag: SPT
```

```
UpTime: 00:00:04
```

```
Upstream interface: Vlan-interface100
```

```
Upstream neighbor: 10.255.51.15
```

```
RPF prime neighbor: 10.255.51.15
```

```
Downstream interface(s) information: None
```

由于出接口由IGMP协议生成，检查三个地市中VLAN100内IGMP配置如下：

地市A：

```
interface Vlan-interface100
```

```
description manager
```

```
ip address 1.1.1.10 255.255.255.0
```

```
igmp enable
```

```
igmp version 3
```

```
pim sm
```

地市B:

```
interface Vlan-interface100
```

```
description manager
```

```
ip address 1.1.1.20 255.255.255.0
```

```
pim sm
```

地市C:

```
interface Vlan-interface100
```

```
description manager
```

```
ip address 1.1.1.30 255.255.255.0
```

```
pim sm
```

三台设备只有地市A配置了igmp enable，但地市A却同样接收不到组播报文。查看配置发现地市C的VLAN100内IP地址大于地市A，那么在VLAN100内地市C就会被PIM选举称为DR（Designated Router，指定路由器），下面我们来回顾下DR的作用：

无论是与组播源相连的网络，还是与接收者相连的网络，都需要选举DR（Designated Router，指定路由器）。接收者侧的DR负责向RP发送加入报文（Join Message）；组播源侧的DR负责向RP发送注册报文（Register Message），而DR的选举过程如下：

（1）共享网络上的各路由器相互之间发送Hello报文（携带有竞选DR优先级的参数），拥有最高优先级的路由器将成为DR；

（2）如果优先级相同，或者网络中至少有一台路由器不支持在Hello报文中携带竞选DR优先级的参数，则根据各路路由器的IP地址大小来竞选DR，IP地址最大的路由器将成为DR。

通过上面DR的作用我们可以了解到，在充当接收者侧DR的设备上必须使能IGMP，否则连接在该DR上的接收者将无法通过该DR加入组播组。在现网中，VLAN100是环网上的互联VLAN，并且使能了三层接口；在VLAN100内要接受组播流，必须每个节点的VLAN100上的都要使能IGMP，同网段内必须DR设备才能向RP去发送(*,G)加入，在单播路由都是正常的情况下，如果DR设备没使能IGMP，没有处理组播组加入，也就不能向RP去发送加入，请求组播流下来。

2、问题二：在非VLAN 100的其他VLAN内，地市B只有在地市A接收到相应组播组后，才可以正常接收组播分析，首选查看地市A上VLAN 400的配置和地市B上VLAN 500的配置。

地市A

```
interface Vlan-interface400
```

```
ip address 2.2.2.1 255.255.255.0
```

```
pim sm
```

```
igmp enable
```

```
igmp version 3
```

地市B

```
interface Vlan-interface500
```

```
ip address 3.3.3.1 255.255.255.0
```

```
pim sm
```

```
igmp enable
```

```
igmp version 3
```

发现在地市A和地市B上都启用的是IGMP V3，IGMP V3除了兼容IGMP V2和V1外，而且还增加了对特定源组查询的支持。IGMP分别有如下机制查询报文：

- （1）普遍组查询报文中，既不携带组地址，也不携带源地址；
- （2）特定组查询报文中，携带组地址，但不携带源地址；
- （3）特定源组查询报文中，既携带组地址，还携带一个或多个源地址。

为了方便解决故障，与客户协商进行问题复现，在组播源侧新增加一路组播源为239.20.1.9，同时在地市B加入组播接受者。同时在地市A和地市B上分别查看IGMP表项及pim routing-table表项如下：

地市A:

```
[DS-A]dis igmp group
```

```
Total 2 IGMP Group(s).
```

```
Interface group report information of VPN-Instance: public net
```

```
Vlan-interface100(10.255.51.19):
```

```
Total 2 IGMP Groups reported
```

Group Address	Last Reporter	Uptime	Expires
---------------	---------------	--------	---------

239.20.1.9	10.255.51.100	00:00:01	00:02:09
------------	---------------	----------	----------

239.255.255.250	10.255.51.100	00:00:01	00:02:09
-----------------	---------------	----------	----------

```
[DSA]dis pim routing-table
```

```
Total 1 (*, G) entries; 0 (S, G) entry
```

```
(*, 239.20.1.9)
```

```
RP: 192.168.200.201
```

```
Protocol: pim-sm, Flag: WC
```

```
UpTime: 00:00:49
```

```
Upstream interface: Vlan-interface100
Upstream neighbor: 1.1.1.30
RPF prime neighbor: 1.1.1.30
Downstream interface(s) information:
Total number of downstreams: 1
1: Vlan-interface100
Protocol: igmp, UpTime: 00:00:49, Expires: -
```

地市B

```
[DSA]dis igmp group
```

```
Total 2 IGMP Group(s).
```

```
Interface group report information of VPN-Instance: public net
```

```
Vlan-interface500(10.255.51.19):
```

```
Total 2 IGMP Groups reported
```

Group Address	Last Reporter	Uptime	Expires
239.20.1.9	10.255.51.100	00:00:01	00:02:09
239.255.255.250	10.255.51.100	00:00:01	00:02:09

```
[DS-B]dis pim routing-table
```

```
Total 1 (*, G) entries; 0 (S, G) entry
```

```
(*, 239.20.1.9)
```

```
RP: 192.168.200.201
```

```
Protocol: pim-sm, Flag: WC
```

```
UpTime: 00:00:49
```

```
Upstream interface: Vlan-interface100
```

```
Upstream neighbor: 1.1.1.30
```

```
RPF prime neighbor: 1.1.1.30
```

```
Downstream interface(s) information:
```

```
Total number of downstreams: 1
```

```
1: Vlan-interface500
```

```
Protocol: igmp, UpTime: 00:00:49, Expires: -
```

从地市A中设备表项查询结果得出，在地市A只是建立了针对的239.20.1.9组播组的（*，G）表项从地市B中设备表项查询结果得出，同样只建立了239.20.1.9组播组的（*，G）表项。怀疑RP没有回应（*，G）表项的请求，此时查看友商RP设备配置如下：

```
ip pim ssm range acl_SSM-IPmc-range
ip access-list standard acl_SSM-IPmc-range
permit 232.0.0.0 0.255.255.255
permit 239.0.0.0 0.0.0.255
```

发现RP设备上将239段组播设置为PIM-SSM,我们可以先来回顾下PIM-SSM是如何运作的：

Any-Source Multicast，任意信源组播模型包括PIM-DM和PIM-SM两种模式，SSM（Source-Specific Multicast，指定信源组播）模型能够借助PIM-SM的部分技术来实现，也称为PIM-SSM。SSM模型为指定源组播提供了解决方案，通过IGMPv3来维护主机与路由器之间的关系。在实际应用中，通常采用IGMPv3以及PIM-SM的一部分技术来实现SSM模型。

组播信息的接收者（Receiver），由其借助IGMPv3的报告报文向DR报告自己对来自组播源S、发往组播组G的信息感兴趣，如果在SSM组地址范围内，则构建PIM-SSM，并向组播源S逐跳发送通道（Channel）的订阅报文（Subscribe Message）。沿途所有路由器上都创建（S，G）表项，从而在网络内构建了一棵以组播源S为根、以接收者为叶子的SPT，该SPT就是PIM-SSM中的传输通道，也就是说PIM-SSM必须结合IGMP V3中的指定组播源信息，才能建立完整的（*，G）和（S，G）表项。如果通过前面的配置查看，我们已经得出地市A和地市B都配置了IGMP V3，理论都可以接受组播信息，但故障现象是只有地市A主动接受组播信息，建立（S，G）表项后，地市B才会接受到组播信息。针对这个问题我们分别在2个地市设备上进行debug igmp查看pim表项的建立过程。

地市A:

```
*Mar 1 20:58:03:130 2017 FJQZ-ZhiBo-S7604X IGMP/7/REPORT: -MDC=1; Received IGMPv3 report on interface Vlan-interface400(2.2.2.1) from 2.2.2.2 (G161283)
*Mar 1 20:58:03:130 2017 FJQZ-ZhiBo-S7604X IGMP/7/REPORT: -MDC=1; Process IS_EX packet for EXCLUDE group(239.20.1.9) on interface Vlan-interface400(2.2.2.1) (G115513)
*Mar 1 20:58:03:130 2017 FJQZ-ZhiBo-S7604X IGMP/7/TIMER: -MDC=1; Set group aging timer for group(239.20.1.9) on interface Vlan-interface400(2.2.2.1) to 260s (G112963)
*Mar 1 20:58:23:303 2017 FJQZ-ZhiBo-S7604X IGMP/7/REPORT: -MDC=1; Received IGMPv3 report on interface Vlan-interface400(2.2.2.1) from 2.2.2.2 (G161283)
*Mar 1 20:58:23:304 2017 FJQZ-ZhiBo-S7604X IGMP/7/REPORT: -MDC=1; Process TO_IN packet for EXCLUDE group(239.20.1.9) on interface Vlan-interface400(2.2.2.1) (G115513)
*Mar 1 20:58:23:304 2017 FJQZ-ZhiBo-S7604X IGMP/7/EVENT: -MDC=1; Send JOIN for (1.1.1.2, 239.20.1.9) on interface Vlan-interface400(2.2.2.1) to MRIB (G11109)
```

地市B

```
*Mar 3 09:52:14:496 2017 FJ-ZhiBo-S7608XV-2 IGMP/7/REPORT: -MDC=1; Received IGMPv2 re
```

port for group 239.20.1.9 on interface Vlan-interface500(10.255.31.129) (G16992)

*Mar 3 09:52:14:497 2017 FJ-ZhiBo-S7608XV-2 IGMP/7/REPORT: -MDC=1; Process IS_EX packet for EXCLUDE group(239.20.1.9) on interface Vlan-interface500(10.255.31.129) (G115528)

*Mar 3 09:52:18:851 2017 FJ-ZhiBo-S7608XV-2 IGMP/7/REPORT: -MDC=1; Received IGMPv2 report for group 239.20.1.3 on interface Vlan-interface500(10.255.31.129) (G16992)

由以上debug可以看出，地市A终端所发出的IGMP请求为IGMP V3，携带组播源地址，而地市B终端所发出的IGMP请求为IGMP V2，没有携带组播源地址。结合之前右商RP配置得出，地市B无法接收到组播的原因因为接收者IGMP版本不正确导致。那么为何在地市A接受到组播流量后，地市B也可以接受到呢，通过理论分析，可以得出如下结论：

地市B设备收到组播接收者请求，掌握了组播接收者的相关信息后，会向该组播组的RP（Rendezvous Point）设备逐跳发送（*，G）加入报文，以构建RP到组播接收者之间用于转发组播流量的共享树。此时，若在组播接收者到RP之间的路径上存在运行PIM SSM协议的网络设备，或被指定为RP设备的友商设备本身就是PIM SSM设备，则会在组播树无法构建的现象。因为PIM SSM本身就是为指定源组播设计的，当友商设备收到下游设备地市A发送而来的未指定组播源的（*，G）加入报文时，会自动丢弃该报文，并不会创建任何组播表项。因此友商设备不认为地市A的方向会有该组播组接收者，也就无法将该组播组的组播流量引至地市A及下游设备处，最终导致接收者无法正常接收组播流，当地市A直连的接受者发送IGMP V3请求，在地市A设备上建立完善的（*，G）和（S,G）表项后，地市B的（*，G）请求到达地市A设备，地市A设备会直接回应该（*，G）表项，无需RP回应。这样地市B设备也会建立完整的（*，G）和（S,G）表项，这样地市B直连的组播接收者也可以接收到组播数据。

问题一：

1、在所有RRPP环网中的VLAN 100内开启IGMP功能，这样RP设备可以正常回应组播组加入请求，建立相应转发表项

2、如果只需要地市A的VLAN 100内的组播接收者接收组播报文，可以将地市A内VLAN 100的IP地址改为最大，保证接收者接在DR设备下即可

问题二：

1、全网统一切换为PIM-SM

PIM SM协议和IGMPv2协议均遵循ASM（Any-Source Multicast，任意信源组播）模型，无需指定组播源。在无法要求终端统一使用IGMPv3的环境下，建议将全网组播路由设备修改为PIM SM协议，避免出现上文所描述的PIM SM与PIM SSM的兼容问题。

2、旁挂IGMP V3主机

此方案仅作为在无法将PIM SSM设备修改为PIM SM的情况下的临时规避方案。在组播接收者和RP之间路径上的网络设备（例如图中的地市B）旁挂支持IGMPv3的主机终端，IGMPv3终端将发送携带组播源地址列表的IGMPv3成员报告消息。当地市A设备收到IGMPv3成员报告消息后，会创建（S，G）表项，并向上游SSM友商设备邻居发送（S，G）加入报文。友商设备收到（S，G）加入报文后会构建SPT，最终将组播流量引至地市B及下游PIM SM设备。

3、启用IGMP SSM Mapping

IGMP SSM Mapping通过在组播接收者的网关设备上配置SSM静态映射规则，从而为运行IGMPv2的组播接收者提供对SSM模型的支持。SSM模型要求在组播接收者所在的网段，网关设备能够了解组播接收者加入组播组时所指定的组播源。如果组播接收者只能运行IGMPv2，则可以通过在网关上配置IGMP SSM Mapping规则，将IGMPv2报告报文中所包含的（*，G）信息映射为（G，INCLUDE，（S1，S2...））信息。网关上有了这些信息后，就可以上游设备发送（S，G）加入报文。运行PIM SSM的设备（友商设备）收到（S，G）加入报文后会构建SPT，最终将组播流量引至地市B及下游PIM SM设备。此方案配置相对复杂，建议在需要考虑组播源安全性，并且组播接收端无法运行IGMPv3的情况下部署。

针对于组播类问题，建议大家优先查看组播表项的完整性，根据表项结合组播转发原理，配合相应的debug等排查手段，进行问题一步步的排查。