排除HSRP感知PIM故障

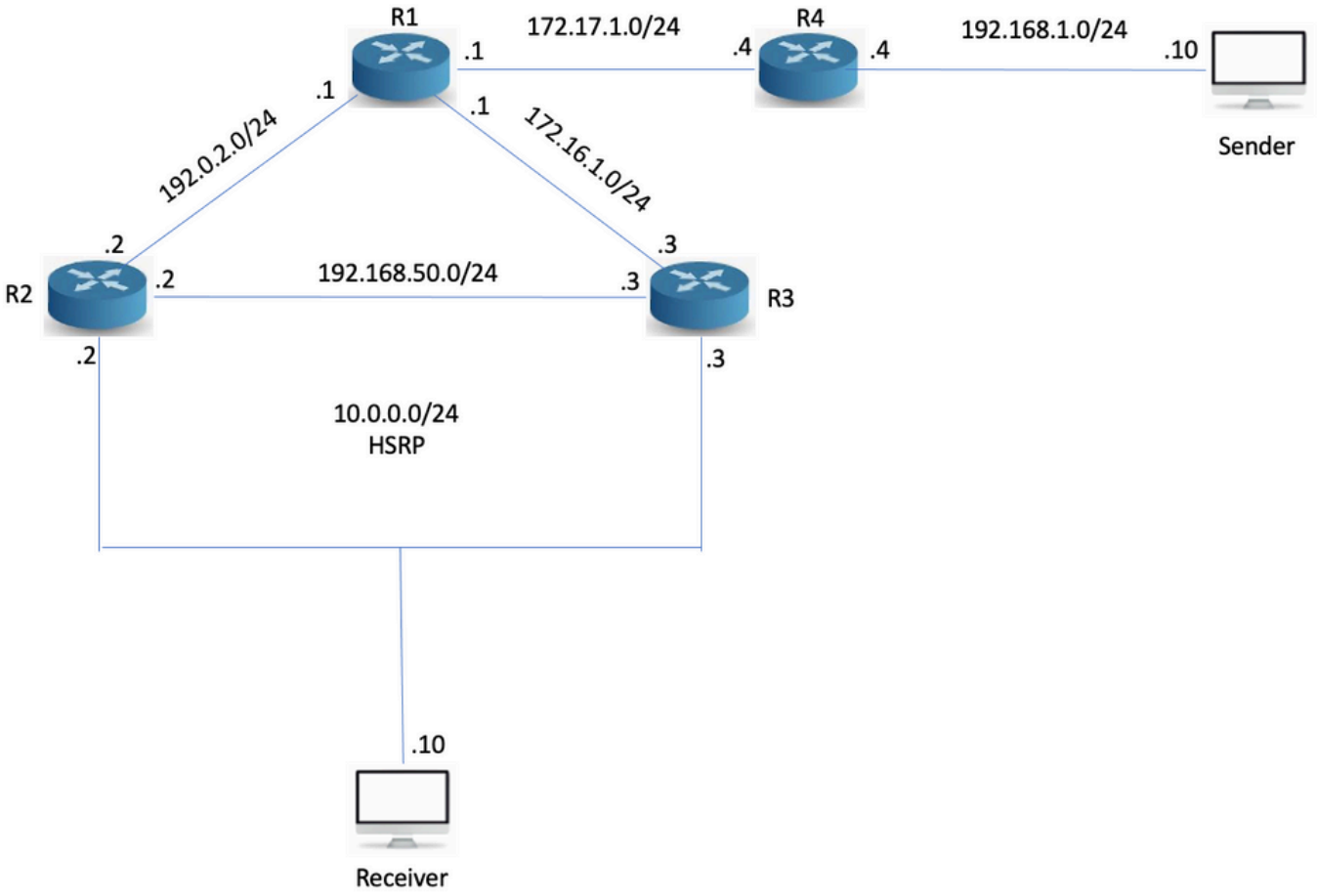
目录

简介

本文档介绍如何对热备份路由器协议(HSRP)感知协议独立组播(PIM)功能以及可以使用该功能的场景进行故障排除。

解释

在需要冗余的环境中，HSRP正常运行。HSRP是一种经过验证的协议，它有效，但是当您有需要组播的客户端时，您如何处理？当活动路由器(AR)关闭时，什么触发组播以收敛？在本例中，使用拓扑1:



拓扑 1

这里需要注意的一点是，R3是PIM指定路由器(DR)，即使R2是HSRP AR。网络已使用开放最短路径优先(OSPF)建立，PIM和R1是IP地址为10.1.1.1的交汇点(RP)。R2和R3均收到互联网组管理协议(IGMP)报告，但只有R3发送PIM加入，因为它是PIM DR。R3向RP构建“*,G”：

```
R3#sh ip mroute 239.0.0.1 IP Multicast Routing Table Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected, L - Local, P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet, X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement, U - URD, I - Received Source Specific Host Report, Z - Multicast Tunnel, z - MDT-data group sender, Y - Joined MDT-data group, y - Sending to MDT-data group, G - Received BGP C-Mroute, g - Sent BGP C-Mroute, N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed, Q - Received BGP S-A Route, q - Sent BGP S-A Route, V - RD & Vector, v - Vector, p - PIM Joins on route Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join Timers: Uptime/Expires Interface state: Interface, Next-Hop or VCD, State/Mode (*, 239.0.0.1), 02:54:15/00:02:20, RP 10.1.1.1, flags: SJC Incoming interface: Ethernet0/0, RPF nbr 172.16.1.1 Outgoing interface list: Ethernet0/2, Forward/Sparse, 00:25:59/00:02:20
```

然后，从组播源ping 239.0.0.1，以构建S，G:

```
Sender#ping 239.0.0.1 re 3 Type escape sequence to abort. Sending 3, 100-byte ICMP Echos to 239.0.0.1, timeout is 2 seconds: Reply to request 0 from 10.0.0.10, 35 ms Reply to request 1 from 10.0.0.10, 1 ms Reply to request 2 from 10.0.0.10, 2 ms
```

S，G已经建立：

```
R3#sh ip mroute 239.0.0.1 IP Multicast Routing Table Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected, L - Local, P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet, X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement, U - URD, I - Received Source Specific Host Report, Z - Multicast Tunnel, z - MDT-data group sender, Y - Joined MDT-data group, y - Sending to MDT-data group, G - Received BGP C-Mroute, g - Sent BGP C-Mroute, N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed, Q - Received BGP S-A Route, q - Sent BGP S-A Route, V - RD & Vector, v - Vector, p - PIM Joins on route Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join Timers: Uptime/Expires Interface state: Interface, Next-Hop or VCD, State/Mode (*, 239.0.0.1), 02:57:14/stopped, RP 10.1.1.1, flags: SJC Incoming interface: Ethernet0/0, RPF nbr 172.16.1.1 Outgoing interface list: Ethernet0/2, Forward/Sparse, 00:28:58/00:02:50 (192.168.1.10, 239.0.0.1), 00:02:03/00:00:56, flags: JT Incoming interface: Ethernet0/0, RPF nbr 172.16.1.1 Outgoing interface list: Ethernet0/2, Forward/Sparse, 00:02:03/00:02:50
```

单播和组播拓扑当前不一致。这可能重要，也可能不重要。当R3发生故障时会发生什么情况？

```
R3(config)#int e0/2 R3(config-if)#sh R3(config-if)#
```

在R2上的PIM检测到R3已经消失并接管DR角色之前，不会收到对ping的应答。使用默认计时器时，需要60至90秒。

```
Sender#ping 239.0.0.1 re 100 ti 1 Type escape sequence to abort. Sending 100, 100-byte ICMP Echos to 239.0.0.1, timeout is 1 seconds: Reply to request 0 from 10.0.0.10, 18 ms Reply to request 1 from 10.0.0.10, 2
ms............................................................... ....... Reply to request 77 from 10.0.0.10, 10 ms Reply to request 78 from 10.0.0.10, 1 ms Reply to request 79 from 10.0.0.10, 1 ms Reply to request 80 from 10.0.0.10, 1 ms
```

您可以增加R2的DR优先级，使其成为DR。

```
R2(config-if)#ip pim dr-priority 50 *May 30 12:42:45.900: %PIM-5-DRCHG: DR change from neighbor 10.0.0.3 to 10.0.0.2 on interface Ethernet0/2
```

HSRP感知PIM是使HSRP成为PIM DR的功能。它还会从虚拟IP发送PIM消息，在路由器具有通向虚拟IP(VIP)的静态路由时，此消息非常有用。 以下是思科对功能的描述：

HSRP感知PIM使组播流量能够通过HSRP AR转发，允许PIM利用HSRP冗余，避免潜在的重复流量，并启用故障转移，这取决于设备中的HSRP状态。PIM-DR与HSRP AR在同一网关上运行并维护mroute状态。

在拓扑1中，HSRP向客户端运行，因此，尽管此功能听起来很适合，但在组播融合中却无济于事。在R2上配置此功能：

```
R2(config-if)#ip pim redundancy HSRP1 hsrp dr-priority 100 R2(config-if)# *May 30 12:48:20.024: %PIM-5-DRCHG: DR change from neighbor 10.0.0.3 to 10.0.0.2 on interface Ethernet0/2
```

R2现在是PIM DR，R3现在在接口E0/2上看到两个PIM邻居：

```
R3#sh ip pim nei e0/2 PIM Neighbor Table Mode: B - Bidir Capable, DR - Designated Router, N - Default DR Priority, P - Proxy Capable, S - State Refresh Capable, G - GenID Capable Neighbor Interface Uptime/Expires Ver DR Address Prio/Mode 10.0.0.1 Ethernet0/2 00:00:51/00:01:23 v2 0 / S P G 10.0.0.2 Ethernet0/2 00:07:24/00:01:23 v2 100/ DR S P G
```

R2现在拥有S、G，您可以看到它是断言赢家，因为R3以前是LAN网段的组播转发器。

```
R2#sh ip mroute 239.0.0.1 IP Multicast Routing Table Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected, L - Local, P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet, X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement, U - URD, I - Received Source Specific Host Report, Z - Multicast Tunnel, z - MDT-data group sender, Y - Joined MDT-data group, y - Sending to MDT-data group, G - Received BGP C-Mroute, g - Sent BGP C-Mroute, N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed, Q - Received BGP S-A Route, q - Sent BGP S-A Route, V - RD & Vector, v - Vector, p - PIM Joins on route Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join Timers: Uptime/Expires Interface state: Interface, Next-Hop or VCD, State/Mode (*, 239.0.0.1), 00:20:31/stopped, RP 10.1.1.1, flags: SJC Incoming interface: Ethernet0/0, RPF nbr 192.0.2.1 Outgoing interface list: Ethernet0/2, Forward/Sparse, 00:16:21/00:02:35 (192.168.1.10, 239.0.0.1), 00:00:19/00:02:40, flags: JT Incoming interface: Ethernet0/0, RPF nbr 192.0.2.1 Outgoing interface list: Ethernet0/2, Forward/Sparse, 00:00:19/00:02:40, A
```

当R2的LAN接口关闭时会发生什么情况？R3能否成为DR?它能以多快的速度收敛？

```
R2(config)#int e0/2 R2(config-if)#sh
```

HSRP在R3上更改为活动，但PIM DR角色在PIM查询间隔过期(3x hello)之前不会收敛。

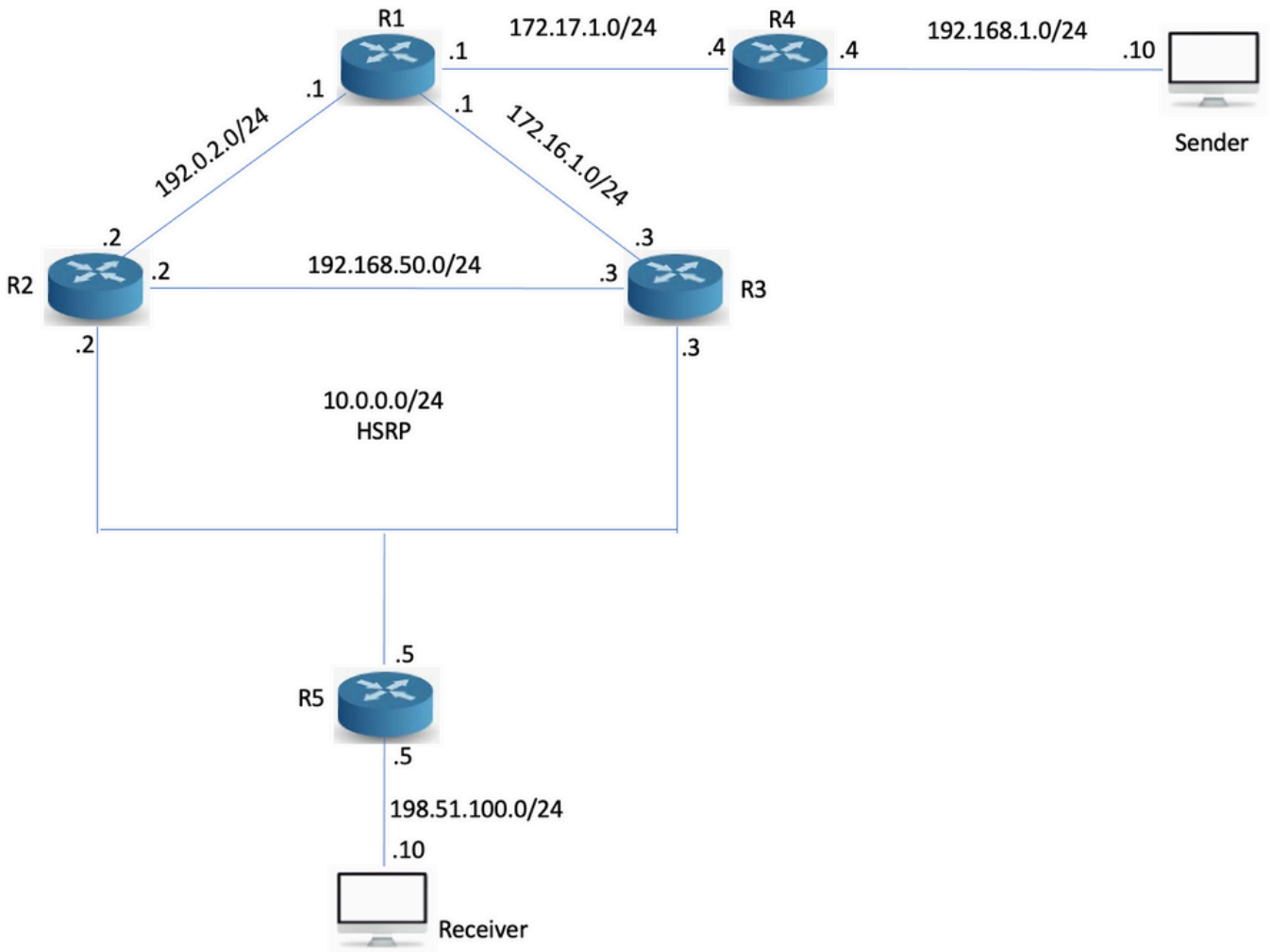
```
*May 30 12:51:44.204: HSRP: Et0/2 Grp 1 Redundancy "hsrp-Et0/2-1" state Standby -> Active R3#sh ip pim nei e0/2 PIM Neighbor Table Mode: B - Bidir Capable, DR - Designated Router, N - Default DR Priority, P - Proxy Capable, S - State Refresh Capable, G - GenID Capable Neighbor Interface Uptime/Expires Ver DR Address Prio/Mode 10.0.0.1 Ethernet0/2 00:04:05/00:00:36 v2 0 / S P G 10.0.0.2 Ethernet0/2 00:10:39/00:00:36 v2 100/ DR S P G R3# *May 30 12:53:02.013: %PIM-5-NBRCHG: neighbor 10.0.0.2 DOWN on interface Ethernet0/2 DR *May 30 12:53:02.013: %PIM-5-DRCHG: DR change from neighbor 10.0.0.2 to 10.0.0.3 on interface Ethernet0/2 *May 30 12:53:02.013: %PIM-5-NBRCHG: neighbor 10.0.0.1 DOWN on interface Ethernet0/2 non DR
```

当PIM收敛发生时，您会丢失大量数据包：

```
Sender#ping 239.0.0.1 re 100 time 1 Type escape sequence to abort. Sending 100, 100-byte ICMP Echos to 239.0.0.1, timeout is 1 seconds: Reply to request 0 from 10.0.0.10, 5 ms Reply to request 0 from 10.0.0.10, 14
```

```
ms..... Reply to request 68 from 10.0.0.10, 10 ms Reply to request 69 from 10.0.0.10, 2 ms Reply to request 70 from 10.0.0.10, 1 ms
```

HSRP知道PIM在这方面没有真正帮助。如果改为使用拓扑2，则此命令非常有用：



拓扑 2

路由器R5已添加，而接收方位于R5后面。R5不与R2和R3运行路由，只与RP和组播源处的静态路由点运行：

```
R5(config)#ip route 10.1.1.1 255.255.255.255 10.0.0.1 R5(config)#ip route 192.168.1.0 255.255.255.0 10.0.0.1
```

如果没有HSRP感知PIM，反向路径转发(RPF)检查将失败，因为具有物理地址的PIM对等体，但R5在网段上看到三个邻居，其中一个是VIP:

```
R5#sh ip pim nei PIM Neighbor Table Mode: B - Bidir Capable, DR - Designated Router, N - Default DR Priority, P - Proxy Capable, S - State Refresh Capable, G - GenID Capable Neighbor Interface Uptime/Expires Ver DR Address Prio/Mode 10.0.0.2 Ethernet0/0 00:03:00/00:01:41 v2 100/ DR S P G 10.0.0.1 Ethernet0/0 00:03:00/00:01:41 v2 0 / S P G 10.0.0.3 Ethernet0/0 00:03:00/00:01:41 v2 1 / S P G
```

R2是在正常情况下转发组播的路由器，因为它是通过活动路由器的HSRP状态转发PIM DR:

```
R2#sh ip mroute 239.0.0.1 IP Multicast Routing Table Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected, L - Local, P - Pruned, R - RP-bit set, F - Register flag, T
```

```
- SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet, X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement, U - URD, I - Received Source Specific Host Report, Z - Multicast Tunnel, z - MDT-data group sender, Y - Joined MDT-data group, y - Sending to MDT-data group, G - Received BGP C-Mroute, g - Sent BGP C-Mroute, N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed, Q - Received BGP S-A Route, q - Sent BGP S-A Route, V - RD & Vector, v - Vector, p - PIM Joins on route Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join Timers: Uptime/Expires Interface state: Interface, Next-Hop or VCD, State/Mode (*, 239.0.0.1), 00:02:12/00:02:39, RP 10.1.1.1, flags: S Incoming interface: Ethernet0/0, RPF nbr 192.0.2.1 Outgoing interface list: Ethernet0/2, Forward/Sparse, 00:02:12/00:02:39
```

尝试从源执行ping操作：

```
Sender#ping 239.0.0.1 re 3 Type escape sequence to abort. Sending 3, 100-byte ICMP Echos to 239.0.0.1, timeout is 2 seconds: Reply to request 0 from 198.51.100.10, 1 ms Reply to request 1 from 198.51.100.10, 2 ms Reply to request 2 from 198.51.100.10, 2 ms
```

ping工作，R2具有S、G:

```
R2#sh ip mroute 239.0.0.1 IP Multicast Routing Table Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected, L - Local, P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet, X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement, U - URD, I - Received Source Specific Host Report, Z - Multicast Tunnel, z - MDT-data group sender, Y - Joined MDT-data group, y - Sending to MDT-data group, G - Received BGP C-Mroute, g - Sent BGP C-Mroute, N - Received BGP Shared-Tree Prune, n - BGP C-Mroute suppressed, Q - Received BGP S-A Route, q - Sent BGP S-A Route, V - RD & Vector, v - Vector, p - PIM Joins on route Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join Timers: Uptime/Expires Interface state: Interface, Next-Hop or VCD, State/Mode (*, 239.0.0.1), 00:04:18/00:03:29, RP 10.1.1.1, flags: S Incoming interface: Ethernet0/0, RPF nbr 192.0.2.1 Outgoing interface list: Ethernet0/2, Forward/Sparse, 00:04:18/00:03:29 (192.168.1.10, 239.0.0.1), 00:01:35/00:01:24, flags: T Incoming interface: Ethernet0/0, RPF nbr 192.0.2.1 Outgoing interface list: Ethernet0/2, Forward/Sparse, 00:01:35/00:03:29
```

当R2发生故障时会发生什么情况？

```
R2#conf t Enter configuration commands, one per line. End with CNTL/Z. R2(config)#int e0/2 R2(config-if)#sh R2(config-if)#

Sender#ping 239.0.0.1 re 200 ti 1 Type escape sequence to abort. Sending 200, 100-byte ICMP Echos to 239.0.0.1, timeout is 1 seconds: Reply to request 0 from 198.51.100.10, 9 ms Reply to request 1 from 198.51.100.10, 2 ms Reply to request 1 from 198.51.100.10, 11 ms.....
.....
.....
```

ping超时，因为当R5的PIM加入进入时，R3不会意识到它必须处理加入。

```
*May 30 13:20:13.236: PIM(0): Received v2 Join/Prune on Ethernet0/2 from 10.0.0.5, not to us
*May 30 13:20:32.183: PIM(0): Generation ID changed from neighbor 10.0.0.2
```

结果是，PIM冗余命令也必须在辅助路由器上配置，以便处理PIM加入到VIP。

```
R3(config-if)#ip pim redundancy HSRP1 hsrp dr-priority 10
```

配置完后，将处理传入的加入。R3触发R5发送新的加入，因为GenID在PIM hello中设置为新值。

```
*May 30 13:59:19.333: PIM(0): Matched redundancy group VIP 10.0.0.1 on Ethernet0/2 Active, processing the Join/Prune, to us *May 30 13:40:34.043: PIM(0): Generation ID changed from
```

```
neighbor 10.0.0.1
```

在此配置后，PIM DR角色的收敛速度与HSRP允许的速度一样快。此场景中使用双向转发检测(BFD)。

结论

了解HSRP感知PIM的关键概念是：

- 最初，AR上的PIM冗余配置使其成为DR。
- PIM冗余也必须在辅助路由器上配置，否则，它无法处理PIM加入到VIP。
- PIM DR角色在PIM呼叫超时之前不会收敛。辅助路由器处理Joins，因此组播会收敛。

要点

当HSRP LAN上有接收器时，此功能不起作用，因为DR角色在PIM邻接过期之前不会移动。

相关信息

- https://www.cisco.com/en/US/docs/ios-xml/ios/ipmulti_pim/configuration/15-2s/imc_hsrp_aware.html#GUID-1294B212-466A-4D8D-AB20-D8DE0B3645CD
- 技术支持和文档 - Cisco Systems