

知 某局点跨Leaf虚拟机之间二层不通问题排查

ADDC解决方案 孙秀丽 2020-08-28 发表

组网及说明

ADDC 3.0组网，EVPN分布式网关网络overlay

问题描述

如题，跨Leaf之间虚拟机二层不通

过程分析

EVPN分布式网关在部署时可采用ARP代答和ARP代理两种方式。采用ARP代答时，Leaf会通过EVPN将远端Leaf下主机的ARP信息学到ARP抑制表，相同VXLAN的主机进行二层通信时，由Leaf对请求远端主机的ARP报文进行代答，此时依然走二层转发流程。采用ARP代理时，Leaf会直接用自己的网关MAC代理响应对远端主机的ARP请求，此时相同VXLAN的主机互访会走三层转发流程。

查看leaf是交换机转发模式为arp代理还是arp代答：dis cu | in arp, arp suppression enable为arp代答（L2转发模式）。local-proxy-arp enable为arp代理模式（L3转发模式）。EVPN代答模式中，交换机上本端虚拟机的MAC地址表，arp抑制表项都是由收到的虚拟机ARP信息生成。本端虚拟机MAC地址表项和ARP抑制表项查询命令：dis l2vpn mac、dis arp suppression vsi。

ARP代答模式，查看两侧的leaf交换机上是否学习到对端虚拟机的MAC表项和arp抑制表，本端虚拟机访问对端虚拟机时ARP信息由ARP抑制表代答，之后参考L2VPN MAC表项转发，这两种表项都是由EVPN BGP2类路由生成。检查leaf对端虚拟机的mac地址表项是否学习到，使用dis l2vpn mac-address查看leaf交换机上是否存在对端虚拟机的Mac地址表项。对端的mac地址对应的Link ID接口是vxlan隧道名称，类型是EVPN表示由BGP EVPN消息生成。dis arp suppression vsi查看对端虚拟机的arp抑制表项是否学习到。

如果没有查询到相关表项信息，进一步排查两个vtep设备隧道是否正常建立（dis interface tunnel），BGP邻居关系是否正常（dis bgp peer l2vpn evpn），如果隧道未建立排查vtep设备之间连通性，BGP关系没有建立起来邻居，需要排查underlay链路，其次排查路由是否可达。

此案例中两个Leaf设备之间隧道没有建立，leaf之间连通性正常。最后定位两台leaf设备的router-id一致导致无法正常建立隧道。

```
[YZ-F4B2-1-2/3-DXY-Leaf-SC-IRF]bgp 6001
[YZ-F4B2-1-2/3-DXY-Leaf-SC-IRF-bgp-default]dis this
#
bgp 6001
non-stop-routing
router-id 11.107.129.6
peer 11.107.129.4 as-number 6001
peer 11.107.129.4 connect-interface LoopBack0
peer 11.107.129.5 as-number 6001
peer 11.107.129.5 connect-interface LoopBack0
#
address-family l2vpn vpn
undo policy vpn-target
peer 11.107.129.4 enable
peer 11.107.129.5 enable
#
ip vpn-instance 21gnk42hmb9jeouge1i016a90j
#
address-family ipv4 unicast
balance 4
network 10.0.0.0 255.255.255.0
network 10.0.0.1 255.255.255.255
network 30.0.0.0 255.255.255.0
network 30.0.0.1 255.255.255.255
#
--- More ---
```

解决方法

修改其中一台设备的router-id之后，两台leaf设备之间vxlan隧道正常建立后，虚拟机二层互通恢复正常可达。