

某局点 7800XP-G 虚拟机访问设备loopback口地址不通

IPv4 IP转发 刘贝 2023-12-01 发表

组网及说明

主机-leaf-spine-border, 主机和border建立Vxlan隧道

告警信息

不涉及

问题描述

现场反馈一组UNIS 7800XP-G组成的M-LAG设备作为border-leaf设备 (L3GW-leaf), 和OVS建立vxlan, 现场发现虚拟机去访问7800XP-G设备上的loopback地址不通

过程分析

1. 经过流统确认虚拟机ping的报文进到L3GW-leaf, 且已经去除vxlan封装, 理论上访问的设备上的地址, 应该三层转发上cpu处理, 但是在设备上debug ip icmp无报文打印, 将报文镜像到cpu后发现流量可以通, 说明问题在于访问设备的报文未正常上送cpu。

2. 现场流量进到设备后需要从vpn-vpcc跨到vpn-L3GW, 是通过华为云下发的静态路由实现, 通过排查排除了跨VPN 下发硬件参数错误的可能, 跟正常跨VPN报文diag抓包对比发现, 有问题的流量缺少EPE Header Adjust 过程, 即没有进行环回口egress封装, 通过diag/trig/pkt/mod/net/watch-key/src-port/85抓到报文, 即通过环回口可以抓到报文, 但是发现报文在跨vpn 过环回口的时候, 报文被丢弃, 进一步通过show/qos/stat/port/84/drop-pkt分析发现是因为环回口队列被占满, 导致报文在环回口被丢弃。

```
[H3C-1-probe]sdk slot 1 show/qos/stat/port/84/drop-pkt
show qos stat port 84 drop-pkt
Gport Queue id Deq(Pkts) Deq(Bytes) Drop(Pkts) Drop(Bytes)
-----
0x54 0 0 0 0 0
1 0 0 0 0 0
2 0 0 400966 0
3 0 0 0 0 0
4 0 0 0 0 0
5 0 0 0 0 0
6 0 0 0 0 0
7 0 0 0 0 0
-----
```

3. 通过内部环回口抓到的ip地址11.191.8.182的报文, 检查设备上配置如以下2条跨vpn的静态路由, 该路由会导致访问这个地址的报文在环回口形成环路, 设备跨vpn转发的报文经过环回口进行环回, 此时报文跨vpn转发不会做TTL-1的操作, 所以报文会一直环回, 导致环回口队列占满, 现场ping的流量也需要跨vpn转发, 也需要到环回口转发, 由于队列已经满了, 所以无法被丢弃无法转发。

```
ip route-static vpn-instance vpna 0.0.0.0 0 vpn-instance vpnb
ip route-static vpn-instance vpnb 11.X.8.0 24 vpn-instance vpna
```

```
[H3C-1-probe]sdk slot 1 show/diag/pkt/de
show diag pkt de
Show Diag Packet Trace:
-----
Dest Type :Network
Dest Port :0x0056
Dest Channel :114
Interface Id :1
Loop Flags :iLoop eLoop
Current Position :EPE
-----
Detail Path Info:
-----
1st Parsing
mac-da :dc2d.XXXX.b58a
mac-sa :dc2d.XXXX.d44a
ether-type :0x00000800
layer2-ext-type :0
layer3-type :L3TYPE_IPV4
ip-protocol :0x00000001
ip-da :11.X.8.182
ip-sa :200.X.6.9
```

```
layer3-ext-type :0
layer4-type :L4TYPE_ICMP
I2-decoder-index :1
I3-decoder-index :2
I4-decoder-index :4
```

4. 现场通过配置24位网段的黑洞路由，将这样异常环回的报文匹配黑洞路由丢弃后故障恢复

```
ip route-static vpn-instance vpna 11.x.8.0 24 NULL 0
```

解决方法

规划静态路由来规避