

物理网络转发丢包导致虚拟机跨CVK访问丢包

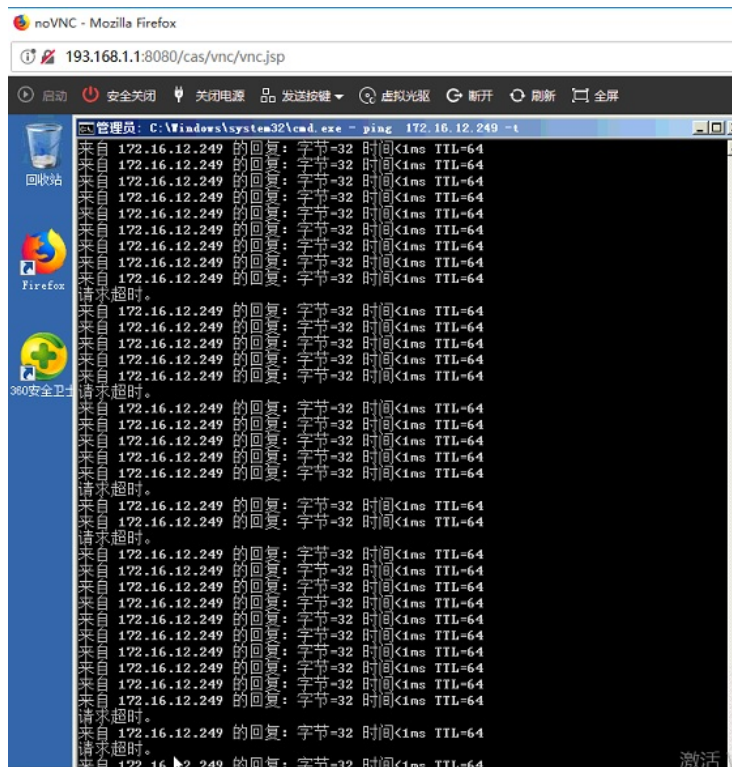
厉梦如 2018-09-11 发表

组网及说明

主机tljyj上有一台linux虚拟机“linux test”，地址172.16.12.249，另外一台主机tljyj-2上有一台虚拟机“windows 2008R2 test”，地址172.16.12.254。

问题描述

两台虚拟机测试ping发现丢包。



过程分析

查看虚拟机所在vswitch为专门的业务vswitch“yw”。管理网CVM和CVK之间走的管理网尝试ping不丢包。主机和虚拟机状态都是正常的。

云资源 / 主机池: 物理 / 主机: 物理

增加虚拟机 导入虚拟机 进入维护模式 退出维护模式 ... 更多操作

三 概要 硬件监控 性能监控 虚拟机 存储 虚拟交换机 物理网卡 存储适配器 高级设置 任务

虚拟交换机

名称	网络类型	物理接口	转发模式	VLAN ID	状态	IP地址	子网掩码	网关
vswitch0	管理网络	eth4	VEB		活动	193.168.1.1	255.255.255.0	193.168.1.254
yw	业务网络	eth5	VEB	1016	活动			

虚拟端口流量

虚拟机	MAC地址	接收报文数	接收字节数
windows 2008R2 test 2	0c:da:41:1d:04:f0	39811	3.33MB
linux test	0c:da:41:1d:a4:e4	27761	1.94MB
test	0c:da:41:1d:9b:93	6857	825.21KB

虚拟交换机

名称	网络类型	物理接口	转发模式	VLAN ID	状态	IP地址	子网掩码	网关
vswitch0	管理网络	eth4	VEB		● 活动	193.168.1.3	255.255.255.0	193.168.1.254
yw	业务网络	eth5	VEB	1016	● 活动			

<

虚拟端口流量

虚拟机	MAC地址	接收报文数	接收字节数
郑好东-2-29	0c:da:41:1d:15:9a	68	5.33KB
windows 2008R2 test	0c:da:41:1d:e0:5e	4690821	3.84GB
名博-7-26	0c:da:41:1d:94:54	276	21.76KB
名博-5-24	0c:da:41:1d:ef:00	8561	1.12MB



尝试从虚拟机"windows 2008R2 test"(172.16.12.254)向虚拟机"linux test" (172.16.12.249) 持续ping , 然后在对应CVK和vnet口进行抓包。涉及4个端口抓包详细说明见下:

- 1、在主机tljy-2的虚拟机"windows 2008R2 test"对应的vnet口抓包
root@tljy-2:~# tcpdump -lni vnet1 host 172.16.12.249
- 2、在主机tljy-2的虚拟机"windows 2008R2 test"所在vswitch口对应eth口上进行抓包
root@tljy-2:~# tcpdump -lni eth5 host 172.16.12.249
- 3、在主机tljy的虚拟机"linux test" (172.16.12.249) 对应的vnet口抓包
root@tljy:~# tcpdump -lni vnet2 host 172.16.12.254
- 4、在主机tljy的虚拟机"linux test" (172.16.12.249) 对应的vnet口抓包
root@tljy:~# tcpdump -lni eth5 host 172.16.12.254

抓包分析说明:

在主机tljy-2的虚拟机"windows 2008R2 test"对应的vnet口抓包。正常ping包应该是1个request包, 1个reply包, 对应seq相同。发现seq13001和13002只有request包, 没有reply包。

```
14:42:48.223030 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13000, length 40
14:42:48.223168 IP 172.16.12.249 > 172.16.12.254: ICMP echo reply, id 1, seq 13000, length 40
14:42:49.223043 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13001, length 40
14:42:54.223121 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13002, length 40
14:42:59.223099 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13003, length 40
14:42:59.223744 IP 172.16.12.249 > 172.16.12.254: ICMP echo reply, id 1, seq 13003, length 40
```

对应主机tljy-2的eth5物理口查看也是如此, 说明虚拟机"windows 2008R2 test"把ping request包从主机上发出去了, 但是没有收到reply报文。

```
14:42:48.223042 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13000, length 40
14:42:48.223163 IP 172.16.12.249 > 172.16.12.254: ICMP echo reply, id 1, seq 13000, length 40
14:42:49.223051 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13001, length 40
14:42:54.223133 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13002, length 40
14:42:59.223112 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13003, length 40
14:42:59.223605 IP 172.16.12.249 > 172.16.12.254: ICMP echo reply, id 1, seq 13003, length 40
```

查看对端CVM主机的虚拟机"linux test"对应vnet2收发情况: 对应seq13001和13002的request报文并没有收到。说明虚拟机没有收到ping request报文

```
14:42:55.528810 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13000, length 40
14:42:55.528905 IP 172.16.12.249 > 172.16.12.254: ICMP echo reply, id 1, seq 13000, length 40
14:43:06.528993 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13003, length 40
14:43:06.529425 IP 172.16.12.249 > 172.16.12.254: ICMP echo reply, id 1, seq 13003, length 40
```

对应CVM主机的eth5收发情况: 对应seq13001和13002的request报文也没有收到, 说明CVM主机物理端口并没有收到ping request报文, 即交换机没有转发虚拟机"windows 2008R2 test"的ping request包给CVM主机, 即报文丢在物理网络上。

```
14:42:55.528816 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13000, length 40
14:42:55.528895 IP 172.16.12.249 > 172.16.12.254: ICMP echo reply, id 1, seq 13000, length 40
14:43:06.529162 IP 172.16.12.254 > 172.16.12.249: ICMP echo request, id 1, seq 13003, length 40
14:43:06.529362 IP 172.16.12.249 > 172.16.12.254: ICMP echo reply, id 1, seq 13003, length 40
```

同样的方法查看到seq12931、seq12976等等都是相同情况，说明报文都是丢在物理网络上，请排查。

解决方法

问题定位在虚拟机发出的ping request报文物理网络并没有正常转发，建议物理交换机上去抓包看看，通过ping包对应连续序列号查看看看报文丢在哪里了。

附上：查看vnet口和虚拟机对应关系的命令ovs_dbg_listports

root@tljyj-2:~# ovs_dbg_listports

```
root@tljyj-2:~# ovs_dbg_listports
PCI NAME LW IO NU SRT0 MTU A/L SPEED MAC IP VEN:DEV DESC SUBSYS DRIVER FIRMWARE
04:00.0 8 0 0/16 -/- - 1425:5001 T520-CR 0000 cxgb4
04:00.1 8 0 0/16 -/- - 1425:5001 T520-CR 0000 cxgb4
04:00.2 8 0 0/16 -/- - 1425:5001 T520-CR 0000 cxgb4
04:00.3 8 0 0/16 -/- - 1425:5001 T520-CR 0000 cxgb4
04:00.4 eth5 8 0 1500 u/u 10000 38ad8e6ce29a 1425:5401 T520-CR 0000 cxgb4 1.15.37.0, TP 0.1.4.9
04:00.4 eth4 8 0 1500 u/u 10000 38ad8e6ce292 1425:5401 T520-CR 0000 cxgb4 1.15.37.0, TP 0.1.4.9
08:00.0 eth0 4 0 0/8 1500 u/d 0 38adbed814dd 8086:1521 I350 igb 1.63, 0x800009fa
08:00.1 eth1 4 0 0/8 1500 u/d 0 38adbed814de 8086:1521 I350 igb 1.63, 0x800009fa
08:00.2 eth2 4 0 0/8 1500 u/d 0 38adbed814df 8086:1521 I350 igb 1.63, 0x800009fa
08:00.3 eth3 4 0 0/8 1500 u/d 0 38adbed814e0 8086:1521 I350 igb 1.63, 0x800009fa

Type Name MAC IP OFFPort Vlanx VName VNC NwFilter
vswitch0 (vrb)
int vswitch0 38ad8e6ce292 193.168.1.3 65534/2
eth4 38ad8e6ce292 172.16.12.28 4/14 zh1-2-28 5909 04:00.4
yw (vrb)
int yw 38ad8e6ce29a 172.16.12.3 65534/4 1016x windows2008R2test 5901
vnet1 0cda4110c05e 172.16.12.254 3/6
vnet0 0cda4110b993 172.16.12.250 4/5
vnet5 0cda4110294a 172.16.12.23 21/10 mb-4-23 5905
vnet2 0cda41109454 172.16.12.26 25/7 mb-7-26 5902
vnet4 0cda41101523 172.16.12.22 20/9 mb-3-22 5904
vnet7 0cda4110431f 172.16.12.25 24/12 mb-6-25 5907
vnet11 0cda4110c03c 172.16.12.30 34/16 ds-1-30 5911
vnet8 0cda41102457 172.16.12.27 26/13 zh1-1-27 5908
vnet6 0cda4110def00 172.16.12.24 22/11 mb-5-24 5906
vnet0 0cda4110b988 172.16.12.20 17/5 mb-1-20 5900
vnet3 0cda4110e417 172.16.12.21 19/8 mb-2-21 5903
eth5 38ad8e6ce29a 1/3 04:00.4
root@tljyj-2:~#
```

root@tljyj:~# ovs_dbg_listports

```
root@tljyj:~# ovs_dbg_listports
PCI NAME LW IO NU SRT0 MTU A/L SPEED MAC IP VEN:DEV DESC SUBSYS DRIVER FIRMWARE
04:00.0 8 0 0/16 -/- - 1425:5001 T520-CR 0000 cxgb4
04:00.1 8 0 0/16 -/- - 1425:5001 T520-CR 0000 cxgb4
04:00.2 8 0 0/16 -/- - 1425:5001 T520-CR 0000 cxgb4
04:00.3 8 0 0/16 -/- - 1425:5001 T520-CR 0000 cxgb4
04:00.4 eth5 8 0 1500 u/u 10000 9ce89517b606 1425:5401 T520-CR 0000 cxgb4 1.15.37.0, TP 0.1.4.9
04:00.4 eth4 8 0 1500 u/u 10000 9ce89517b5fe 1425:5401 T520-CR 0000 cxgb4 1.15.37.0, TP 0.1.4.9
08:00.0 eth0 4 0 0/8 1500 u/d 0 38adbed814fb 8086:1521 I350 igb 1.63, 0x800009fa
08:00.1 eth1 4 0 0/8 1500 u/d 0 38adbed814fc 8086:1521 I350 igb 1.63, 0x800009fa
08:00.2 eth2 4 0 0/8 1500 u/d 0 38adbed814fd 8086:1521 I350 igb 1.63, 0x800009fa
08:00.3 eth3 4 0 0/8 1500 u/d 0 38adbed814fe 8086:1521 I350 igb 1.63, 0x800009fa

Type Name MAC IP OFFPort Vlanx VName VNC NwFilter
vswitch0 (vrb)
int vswitch0 9ce89517b5fe 193.168.1.1 65534/2
eth4 9ce89517b5fe 172.16.12.28 1/1 04:00.4
yw (vrb)
int yw 9ce89517b606 65534/4 1016x windows2008R2test2 5901
vnet1 0cda4110d0f0 172.16.12.252 3/6
eth5 9ce89517b606 1/3 04:00.4
vnet2 0cda4110b4e4 172.16.12.249 3/7 linuxtest 5902
vnet0 0cda4110b993 172.16.12.250 4/5 test 5900
root@tljyj:~#
```