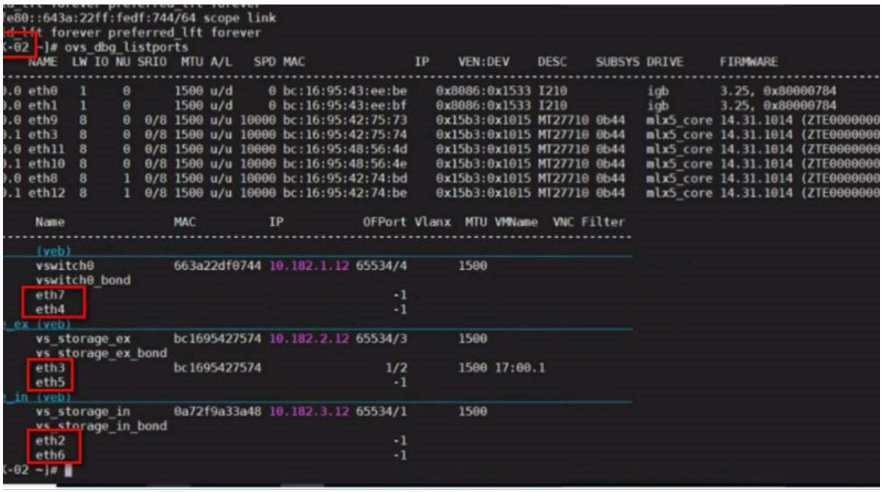
Workspace E1011P11超融合环境主机eth口自动改变

芮恒 2023-03-30 发表

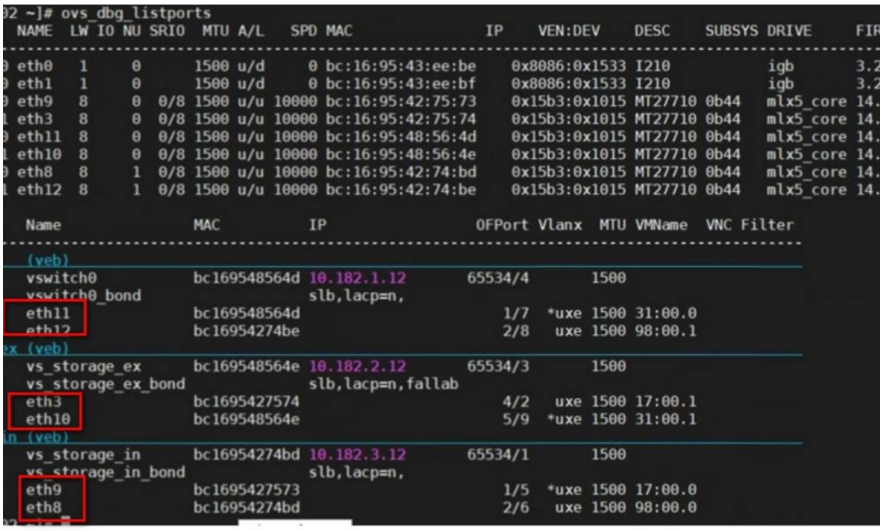
问题描述

现场发现一台主机的eth口自动改变，导致其和集群其他主机之间网络不通。主机原eth口分别为eth2、3、4、5、6、7，改变后的eth口分别为eth3、8、9、10、11、12



```
fe80::643a:22ff:fedf:744/64 scope link
    lft forever preferred_lft forever
-02 ~]# ovs_dbg_listports
  NAME   LW IO NU SRIO  MTU A/L   SPD MAC                IP    VEN:DEV     DESC        SUBSYS DRIVE     FIRMWARE
-------------------------------------------------------------------------------------------------------------------------
).0 eth0    1     0       1500 u/d     0 bc:16:95:43:ee:be     0x8086:0x1533 I210                igb       3.25, 0x80000784
).0 eth1    1     0       1500 u/d     0 bc:16:95:43:ee:bf     0x8086:0x1533 I210                igb       3.25, 0x80000784
).0 eth9    8     0  0/8 1500 u/u 10000 bc:16:95:42:75:73     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.31.1014 (ZTE0000000
).1 eth3    8     0  0/8 1500 u/u 10000 bc:16:95:42:75:74     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.31.1014 (ZTE0000000
).0 eth11   8     0  0/8 1500 u/u 10000 bc:16:95:48:56:4d     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.31.1014 (ZTE0000000
).1 eth10   8     0  0/8 1500 u/u 10000 bc:16:95:48:56:4e     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.31.1014 (ZTE0000000
).0 eth8    8     1  0/8 1500 u/u 10000 bc:16:95:42:74:bd     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.31.1014 (ZTE0000000
).1 eth12   8     1  0/8 1500 u/u 10000 bc:16:95:42:74:be     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.31.1014 (ZTE0000000

    Name              MAC          IP          OFPort Vlanx  MTU VMName  VNC Filter
-----------------------------------------------------------------------------------
    (veb)
    vswitch0          663a22df0744 10.182.1.12 65534/4        1500
    vswitch0_bond
    eth7                                            -1
    eth4                                            -1
_ex (veb)
    vs_storage_ex     bc1695427574 10.182.2.12 65534/3        1500
    vs_storage_ex_bond
    eth3              bc1695427574                 1/2        1500 17:00.1
    eth5                                            -1
_in (veb)
    vs_storage_in     0a72f9a33a48 10.182.3.12 65534/1        1500
    vs_storage_in_bond
    eth2                                            -1
    eth6                                            -1
-02 ~]#
```

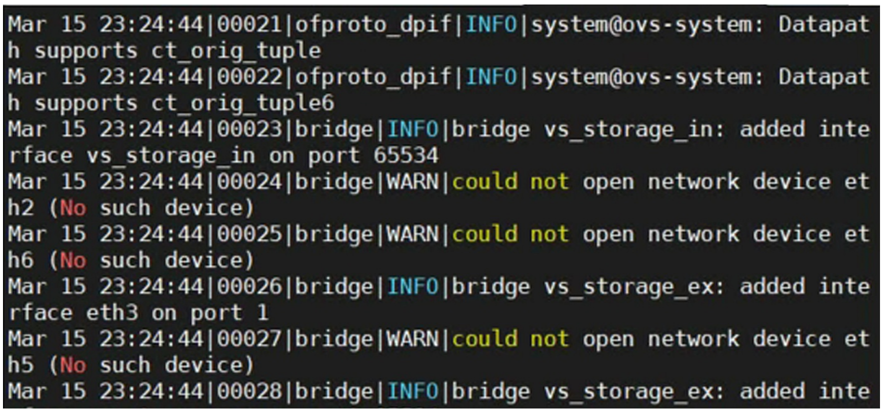


```
02 ~]# ovs_dbg_listports
 NAME   LW IO NU SRIO  MTU A/L   SPD MAC                IP    VEN:DEV     DESC        SUBSYS DRIVE     FIR
-----------------------------------------------------------------------------------------------------------
9 eth0    1     0       1500 u/d     0 bc:16:95:43:ee:be     0x8086:0x1533 I210                igb       3.2
9 eth1    1     0       1500 u/d     0 bc:16:95:43:ee:bf     0x8086:0x1533 I210                igb       3.2
9 eth9    8     0  0/8 1500 u/u 10000 bc:16:95:42:75:73     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.
1 eth3    8     0  0/8 1500 u/u 10000 bc:16:95:42:75:74     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.
9 eth11   8     0  0/8 1500 u/u 10000 bc:16:95:48:56:4d     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.
1 eth10   8     0  0/8 1500 u/u 10000 bc:16:95:48:56:4e     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.
9 eth8    8     1  0/8 1500 u/u 10000 bc:16:95:42:74:bd     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.
1 eth12   8     1  0/8 1500 u/u 10000 bc:16:95:42:74:be     0x15b3:0x1015 MT27710 0b44        mlx5_core 14.

  Name              MAC          IP                   OFPort Vlanx  MTU VMName  VNC Filter
-------------------------------------------------------------------------------------------
  (veb)
  vswitch0          bc169548564d 10.182.1.12          65534/4        1500
  vswitch0_bond                  slb,lacp=n,
  eth11             bc169548564d                          1/7  *uxe 1500 31:00.0
  eth12             bc16954274be                          2/8   uxe 1500 98:00.1
x (veb)
  vs_storage_ex     bc169548564e 10.182.2.12          65534/3        1500
  vs_storage_ex_bond             slb,lacp=n,fallab
  eth3              bc1695427574                          4/2   uxe 1500 17:00.1
  eth10             bc169548564e                          5/9  *uxe 1500 31:00.1
in (veb)
  vs_storage_in     bc16954274bd 10.182.3.12          65534/1        1500
  vs_storage_in_bond             slb,lacp=n,
  eth9              bc1695427573                          1/5  *uxe 1500 17:00.0
  eth8              bc16954274bd                          2/6   uxe 1500 98:00.0
```

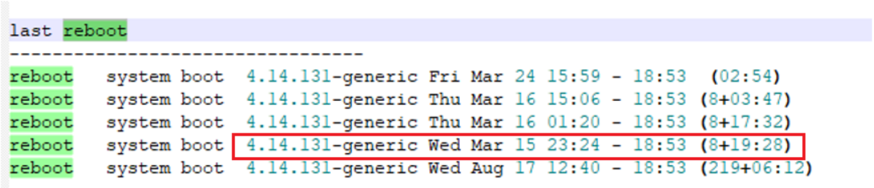
过程分析

1、查看/var/log/openvswitch/ovs-vswitchd.log日志，可以看到现场最开始报原eth2口查询异常是在3月15日23：24左右。



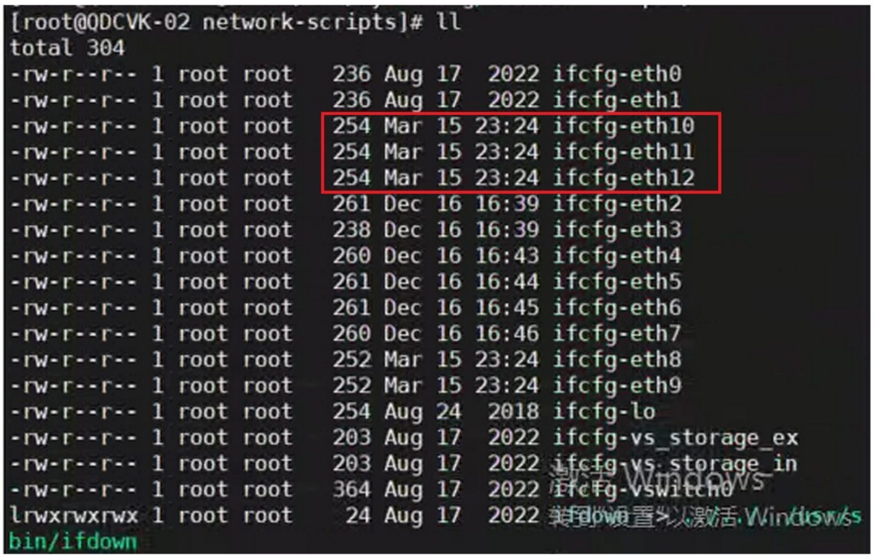
```
Mar 15 23:24:44|00021|ofproto_dpif|INFO|system@ovs-system: Datapat
h supports ct_orig_tuple
Mar 15 23:24:44|00022|ofproto_dpif|INFO|system@ovs-system: Datapat
h supports ct_orig_tuple6
Mar 15 23:24:44|00023|bridge|INFO|bridge vs_storage_in: added inte
rface vs_storage_in on port 65534
Mar 15 23:24:44|00024|bridge|WARN|could not open network device et
h2 (No such device)
Mar 15 23:24:44|00025|bridge|WARN|could not open network device et
h6 (No such device)
Mar 15 23:24:44|00026|bridge|INFO|bridge vs_storage_ex: added inte
rface eth3 on port 1
Mar 15 23:24:44|00027|bridge|WARN|could not open network device et
h5 (No such device)
Mar 15 23:24:44|00028|bridge|INFO|bridge vs_storage_ex: added inte
```

2、通过last reboot查询主机历史重启记录，可以看到在3月15日23：24左右发生过重启，怀疑是当时主机重启导致的eth口改变。



```
last reboot
-------------------------------------
reboot   system boot  4.14.131-generic Fri Mar 24 15:59 - 18:53  (02:54)
reboot   system boot  4.14.131-generic Thu Mar 16 15:06 - 18:53 (8+03:47)
reboot   system boot  4.14.131-generic Thu Mar 16 01:20 - 18:53 (8+17:32)
reboot   system boot  4.14.131-generic Wed Mar 15 23:24 - 18:53 (8+19:28)
reboot   system boot  4.14.131-generic Wed Aug 17 12:40 - 18:53 (219+06:12)
```

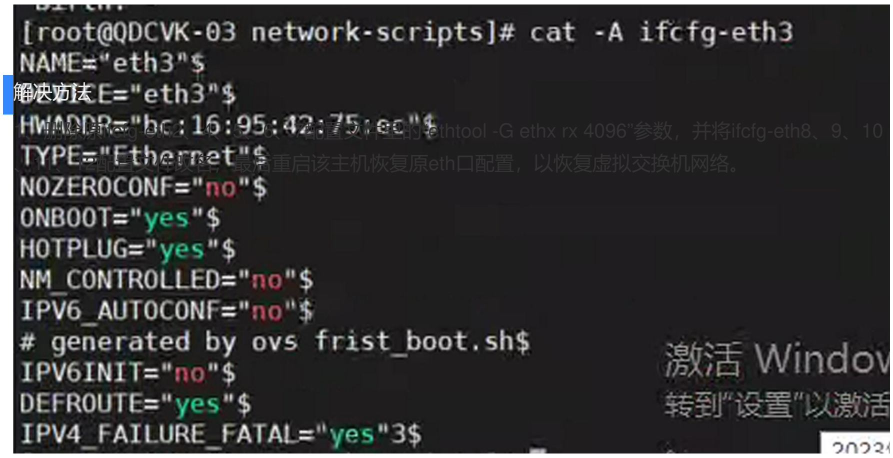
3、进一步查看该主机/etc/sysconfig/network-scripts目录下各eth网络配置文件的最新修改时间，可以看到eth8、9等eth文件确实是在主机重启时间点新生成。



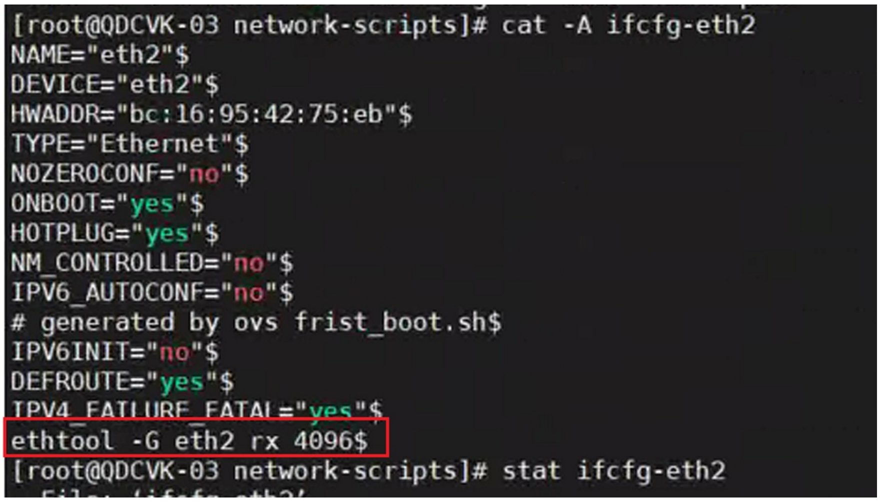
```
[root@QDCVK-02 network-scripts]# ll
total 304
-rw-r--r-- 1 root root   236 Aug 17  2022 ifcfg-eth0
-rw-r--r-- 1 root root   236 Aug 17  2022 ifcfg-eth1
-rw-r--r-- 1 root root   254 Mar 15 23:24 ifcfg-eth10
-rw-r--r-- 1 root root   254 Mar 15 23:24 ifcfg-eth11
-rw-r--r-- 1 root root   254 Mar 15 23:24 ifcfg-eth12
-rw-r--r-- 1 root root   261 Dec 16 16:39 ifcfg-eth2
-rw-r--r-- 1 root root   238 Dec 16 16:39 ifcfg-eth3
-rw-r--r-- 1 root root   260 Dec 16 16:43 ifcfg-eth4
-rw-r--r-- 1 root root   261 Dec 16 16:44 ifcfg-eth5
-rw-r--r-- 1 root root   261 Dec 16 16:45 ifcfg-eth6
-rw-r--r-- 1 root root   260 Dec 16 16:46 ifcfg-eth7
-rw-r--r-- 1 root root   252 Mar 15 23:24 ifcfg-eth8
-rw-r--r-- 1 root root   252 Mar 15 23:24 ifcfg-eth9
-rw-r--r-- 1 root root   254 Aug 24  2018 ifcfg-lo
-rw-r--r-- 1 root root   203 Aug 17  2022 ifcfg-vs_storage_ex
-rw-r--r-- 1 root root   203 Aug 17  2022 ifcfg-vs_storage_in
-rw-r--r-- 1 root root   364 Aug 17  2022 ifcfg-vswitch0
lrwxrwxrwx 1 root root    24 Aug 17  2022 ifdown [illegible]s
bin/ifdown
```

激活 Windows
转到设置以激活 Windows

4、对比前后eth口变化，发现存在原eth3口未改变，怀疑其ifcfg-eth配置文件和其他eth口不一样。对比后可看到，相较eth2、4、5、6、7，其ifcfg-eth配置文件中无“ethtool -G ethx rx 4096”参数，怀疑是该参数导致的现场eth口改变。



```
[root@QDCVK-03 network-scripts]# cat -A ifcfg-eth3
NAME="eth3"$
DEVICE="eth3"$
HWADDR="bc:16:95:42:75:ec"$
TYPE="Ethernet"$
NOZEROCONF="no"$
ONBOOT="yes"$
HOTPLUG="yes"$
NM_CONTROLLED="no"$
IPV6_AUTOCONF="no"$
# generated by ovs frist_boot.sh$
IPV6INIT="no"$
DEFROUTE="yes"$
IPV4_FAILURE_FATAL="yes"3$
```



```
[root@QDCVK-03 network-scripts]# cat -A ifcfg-eth2
NAME="eth2"$
DEVICE="eth2"$
HWADDR="bc:16:95:42:75:eb"$
TYPE="Ethernet"$
NOZEROCONF="no"$
ONBOOT="yes"$
HOTPLUG="yes"$
NM_CONTROLLED="no"$
IPV6_AUTOCONF="no"$
# generated by ovs frist_boot.sh$
IPV6INIT="no"$
DEFROUTE="yes"$
IPV4_FAILURE_FATAL="yes"$
ethtool -G eth2 rx 4096$
[root@QDCVK-03 network-scripts]# stat ifcfg-eth2
```

5、经确认，“ethtool -G ethx rx 4096”参数为现场手动添加，而Workspace主机系统禁止修改默认的ifcfg-eth网络配置文件参数。一旦ifcfg-eth文件内参数被改变，主机在重启时通过ovs读取文件内格式不一致，就会重新生成eth口。

解决方法

1、删除原ifcfg-eth2、3、4、5、6、7配置文件里的“ethtool -G ethx rx 4096”参数，并将ifcfg-eth8、9、10、11、12配置文件改名，最后重启该主机恢复原eth口配置，以恢复虚拟交换机网络。