

知 多路径链路震荡导致主机fence重启

H3Cloud CAS CVK 重启 主机 虚拟机 汪普深 2020-09-28 发表

组网及说明

无

问题描述

CAS平台出现部分虚拟机业务异常，虚拟机显示蓝色状态，主机触发fence重启，后台执行共享存储相关命令会挂住。

过程分析

在CVK主机后台执行共享存储检测脚本，会有共享存储阻塞的提示：

```
root@L01-E6-ZWW-CVK-R6900-03:~# virsh pool-list --all
^C
root@L01-E6-ZWW-CVK-R6900-03:~# ocf2_kernel_stat_collect.sh
Process have been hung up!
Find busy locks on disk /dev/dm-2.
Maybe power off nodes( ) could unblock ocf2-cluster, then power on it/them until ocf2-cluster recovers to normal.
root@L01-E6-ZWW-CVK-R6900-03:~#
```

查看主机kern.log日志，可以看到有共享存储多路径链路震荡，以其中一台主机日志为例，日志内容如下，多路径链路不停断开再恢复：

Jul 23 17:42:52 L01-E3-ZWW-CVK-R6900-01 kernel: [10958703.208382] (o2hb-B14C005525,10951,8):o2hb_thread:1447 do disk heartbeat used 3048 msecs on device(dm-1), ret = 0.//出现共享存储心跳超时

Jul 23 17:42:52 L01-E3-ZWW-CVK-R6900-01 kernel: [10958703.208574] (o2hb-774E8026D9,10750,10):o2hb_thread:1447 do disk heartbeat used 2632 msecs on device(dm-0), ret = 0.

Jul 23 17:42:52 L01-E3-ZWW-CVK-R6900-01 kernel: [10958703.213676] (o2hb-DF5A141CCF,11150,89):o2hb_thread:1447 do disk heartbeat used 2189 msecs on device(dm-2), ret = 0.

Jul 23 17:43:11 L01-E3-ZWW-CVK-R6900-01 kernel: [10958722.124639] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:0 -- 0 2002.//多路径链路异常

Jul 23 17:47:17 L01-E3-ZWW-CVK-R6900-01 kernel: [10958968.401071] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:1 -- 0 2002.

Jul 23 17:52:33 L01-E3-ZWW-CVK-R6900-01 kernel: [10959284.686307] (o2hb-69E6B455D9,11460,15):o2hb_thread:1447 do disk heartbeat used 3075 msecs on device(dm-3), ret = 0.

Jul 23 18:03:48 L01-E3-ZWW-CVK-R6900-01 kernel: [10959959.238115] sd 1:0:3:0: [sdo] tag# 0 FAILED Result: hostbyte=DID_ERROR driverbyte=DRIVER_OK

Jul 23 18:03:48 L01-E3-ZWW-CVK-R6900-01 kernel: [10959959.238125] sd 1:0:3:0: [sdo] tag# 0 CDB: Read(16) 88 00 00 00 00 02 d3 e7 ac 00 00 00 02 00 00 00

Jul 23 18:03:48 L01-E3-ZWW-CVK-R6900-01 kernel: [10959959.238128] print_req_error: I/O error, dev sdo, sector 12145110016//主机访问存储I/O报错

Jul 23 18:03:48 L01-E3-ZWW-CVK-R6900-01 kernel: [10959959.238211] device-mapper: mult ipath: Failing path 8:224.//多路径链路断开

Jul 23 18:03:52 L01-E3-ZWW-CVK-R6900-01 kernel: [10959963.847899] device-mapper: mult ipath: Reinstating path 8:224.//多路径链路恢复

Jul 23 18:03:52 L01-E3-ZWW-CVK-R6900-01 kernel: [10959963.859228] sd 1:0:3:0: alua: port group 01 state A preferred supports tolnsA

Jul 23 18:03:52 L01-E3-ZWW-CVK-R6900-01 kernel: [10959963.859327] sd 1:0:3:0: alua: port group 01 state A preferred supports tolnsA

Jul 23 18:13:19 L01-E3-ZWW-CVK-R6900-01 kernel: [10960530.029822] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:1 -- 0 2002.

Jul 23 18:14:17 L01-E3-ZWW-CVK-R6900-01 kernel: [10960588.910855] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:0 -- 0 2002.

Jul 23 18:17:16 L01-E3-ZWW-CVK-R6900-01 kernel: [10960767.090128] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:0 -- 0 2002.

Jul 23 18:17:16 L01-E3-ZWW-CVK-R6900-01 kernel: [10960767.108009] (o2hb-774E8026D9,10750,4):o2hb_thread:1447 do disk heartbeat used 3166 msecs on device(dm-0), ret = 0.

Jul 23 18:17:39 L01-E3-ZWW-CVK-R6900-01 kernel: [10960790.185856] sd 1:0:3:0: [sdo] tag# 0 FAILED Result: hostbyte=DID_ERROR driverbyte=DRIVER_OK

Jul 23 18:17:39 L01-E3-ZWW-CVK-R6900-01 kernel: [10960790.185865] sd 1:0:3:0: [sdo] tag# 0 CDB: Read(16) 88 00 00 00 00 02 d3 e7 ac 00 00 00 02 00 00 00

Jul 23 18:17:39 L01-E3-ZWW-CVK-R6900-01 kernel: [10960790.185867] print_req_error: I/O error, dev sdo, sector 12145110016

Jul 23 18:17:39 L01-E3-ZWW-CVK-R6900-01 kernel: [10960790.185936] device-mapper: mult ipath: Failing path 8:224.

Jul 23 18:17:43 L01-E3-ZWW-CVK-R6900-01 kernel: [10960794.503803] device-mapper: mult ipath: Reinstating path 8:224.

Jul 23 18:17:43 L01-E3-ZWW-CVK-R6900-01 kernel: [10960794.514581] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:17:43 L01-E3-ZWW-CVK-R6900-01 kernel: [10960794.514721] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:18:46 L01-E3-ZWW-CVK-R6900-01 kernel: [10960857.715695] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:1 -- 0 2002.

Jul 23 18:19:18 L01-E3-ZWW-CVK-R6900-01 kernel: [10960889.204218] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:0 -- 0 2002.

Jul 23 18:19:23 L01-E3-ZWW-CVK-R6900-01 kernel: [10960894.853971] (o2hb-DF5A141CCF,11150,4):o2hb_thread:1447 do disk heartbeat used 2173 msecs on device(dm-2), ret = 0.

Jul 23 18:19:23 L01-E3-ZWW-CVK-R6900-01 kernel: [10960894.860052] (o2hb-B14C005525,10951,8):o2hb_thread:1447 do disk heartbeat used 2243 msecs on device(dm-1), ret = 0.

Jul 23 18:19:54 L01-E3-ZWW-CVK-R6900-01 kernel: [10960925.813000] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:0 -- 0 2002.

Jul 23 18:19:59 L01-E3-ZWW-CVK-R6900-01 kernel: [10960930.933122] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:0 -- 0 2002.

Jul 23 18:19:59 L01-E3-ZWW-CVK-R6900-01 kernel: [10960930.953332] (o2hb-774E8026D9,10750,4):o2hb_thread:1447 do disk heartbeat used 8608 msecs on device(dm-0), ret = 0.

Jul 23 18:20:54 L01-E3-ZWW-CVK-R6900-01 kernel: [10960985.462076] qla2xxx [0000:ad:00.0]-801c:1: Abort command issued nexus=1:3:0 -- 0 2002.

Jul 23 18:20:54 L01-E3-ZWW-CVK-R6900-01 kernel: [10960985.478316] (o2hb-774E8026D9,10750,4):o2hb_thread:1447 do disk heartbeat used 6140 msecs on device(dm-0), ret = 0.

Jul 23 18:21:27 L01-E3-ZWW-CVK-R6900-01 kernel: [10961018.071958] sd 1:0:3:0: [sdo] tag# 0 FAILED Result: hostbyte=DID_ERROR driverbyte=DRIVER_OK

Jul 23 18:21:27 L01-E3-ZWW-CVK-R6900-01 kernel: [10961018.071978] sd 1:0:3:0: [sdo] tag#

0 CDB: Read(16) 88 00 00 00 00 03 be d6 54 00 00 00 02 00 00 00

Jul 23 18:21:27 L01-E3-ZWW-CVK-R6900-01 kernel: [10961018.071980] print_req_error: I/O error, dev sdo, sector 16086619136

Jul 23 18:21:27 L01-E3-ZWW-CVK-R6900-01 kernel: [10961018.072049] device-mapper: multipath: Failing path 8:224.

Jul 23 18:21:31 L01-E3-ZWW-CVK-R6900-01 kernel: [10961022.686528] device-mapper: multipath: Reinstating path 8:224.

Jul 23 18:21:31 L01-E3-ZWW-CVK-R6900-01 kernel: [10961022.694713] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:21:31 L01-E3-ZWW-CVK-R6900-01 kernel: [10961022.694809] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:22:16 L01-E3-ZWW-CVK-R6900-01 kernel: [10961067.087595] sd 1:0:0:1: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:22:16 L01-E3-ZWW-CVK-R6900-01 kernel: [10961067.087697] sd 1:0:0:1: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:22:55 L01-E3-ZWW-CVK-R6900-01 kernel: [10961106.428510] (o2hb-774E8026D9,10750,3):o2hb_thread:1447 do disk heartbeat used 2000 msecs on device(dm-0), ret = 0.

Jul 23 18:23:25 L01-E3-ZWW-CVK-R6900-01 kernel: [10961136.464537] sd 1:0:3:0: [sdo] tag#0 FAILED Result: hostbyte=DID_ERROR driverbyte=DRIVER_OK

Jul 23 18:23:25 L01-E3-ZWW-CVK-R6900-01 kernel: [10961136.464556] sd 1:0:3:0: [sdo] tag#0 CDB: Read(16) 88 00 00 00 00 04 a0 0f 70 00 00 00 02 00 00 00

Jul 23 18:23:25 L01-E3-ZWW-CVK-R6900-01 kernel: [10961136.464558] print_req_error: I/O error, dev sdo, sector 19865235456

Jul 23 18:23:25 L01-E3-ZWW-CVK-R6900-01 kernel: [10961136.464621] device-mapper: multipath: Failing path 8:224.

Jul 23 18:23:27 L01-E3-ZWW-CVK-R6900-01 kernel: [10961138.128479] device-mapper: multipath: Reinstating path 8:224.

Jul 23 18:23:27 L01-E3-ZWW-CVK-R6900-01 kernel: [10961138.136838] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:23:27 L01-E3-ZWW-CVK-R6900-01 kernel: [10961138.136939] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:23:38 L01-E3-ZWW-CVK-R6900-01 kernel: [10961149.050969] (o2hb-774E8026D9,10750,5):o2hb_thread:1447 do disk heartbeat used 2286 msecs on device(dm-0), ret = 0.

Jul 23 18:23:39 L01-E3-ZWW-CVK-R6900-01 kernel: [10961150.763518] sd 1:0:3:0: [sdo] tag#0 FAILED Result: hostbyte=DID_ERROR driverbyte=DRIVER_OK

Jul 23 18:23:39 L01-E3-ZWW-CVK-R6900-01 kernel: [10961150.763538] sd 1:0:3:0: [sdo] tag#0 CDB: Read(16) 88 00 00 00 00 00 63 05 8a 00 00 00 02 00 00 00

Jul 23 18:23:39 L01-E3-ZWW-CVK-R6900-01 kernel: [10961150.763540] print_req_error: I/O error, dev sdo, sector 1661307392

Jul 23 18:23:39 L01-E3-ZWW-CVK-R6900-01 kernel: [10961150.763604] device-mapper: multipath: Failing path 8:224.

Jul 23 18:23:42 L01-E3-ZWW-CVK-R6900-01 kernel: [10961153.141681] device-mapper: multipath: Reinstating path 8:224.

Jul 23 18:23:42 L01-E3-ZWW-CVK-R6900-01 kernel: [10961153.153194] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:23:42 L01-E3-ZWW-CVK-R6900-01 kernel: [10961153.153293] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:25:12 L01-E3-ZWW-CVK-R6900-01 kernel: [10961243.258856] qla2xxx [0000:ad:00:0]-801c:1: Abort command issued nexus=1:3:0 -- 0 2002.

Jul 23 18:25:12 L01-E3-ZWW-CVK-R6900-01 kernel: [10961243.275033] (o2hb-774E8026D9,10750,7):o2hb_thread:1447 do disk heartbeat used 25660 msecs on device(dm-0), ret = 0.

Jul 23 18:25:21 L01-E3-ZWW-CVK-R6900-01 kernel: [10961252.086171] (o2hb-DF5A141CCF,11150,4):o2hb_thread:1447 do disk heartbeat used 2407 msecs on device(dm-2), ret = 0.

Jul 23 18:25:21 L01-E3-ZWW-CVK-R6900-01 kernel: [10961252.086384] (o2hb-774E8026D9,10750,7):o2hb_thread:1447 do disk heartbeat used 2759 msecs on device(dm-0), ret = 0.

Jul 23 18:26:01 L01-E3-ZWW-CVK-R6900-01 kernel: [10961292.371735] sd 1:0:0:4: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:26:01 L01-E3-ZWW-CVK-R6900-01 kernel: [10961292.371835] sd 1:0:0:4: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:26:20 L01-E3-ZWW-CVK-R6900-01 kernel: [10961311.053520] sd 1:0:3:0: [sdo] tag# 1 FAILED Result: hostbyte=DID_ERROR driverbyte=DRIVER_OK

Jul 23 18:26:20 L01-E3-ZWW-CVK-R6900-01 kernel: [10961311.053523] sd 1:0:3:0: [sdo] tag# 1 CDB: Read(16) 88 00 00 00 00 02 9b bf b8 00 00 00 02 00 00 00

Jul 23 18:26:20 L01-E3-ZWW-CVK-R6900-01 kernel: [10961311.053526] print_req_error: I/O error, dev sdo, sector 11202967552

Jul 23 18:26:20 L01-E3-ZWW-CVK-R6900-01 kernel: [10961311.053543] device-mapper: mult ipath: Failing path 8:224.

Jul 23 18:26:24 L01-E3-ZWW-CVK-R6900-01 kernel: [10961315.374468] device-mapper: mult ipath: Reinstating path 8:224.

Jul 23 18:26:24 L01-E3-ZWW-CVK-R6900-01 kernel: [10961315.384142] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:26:24 L01-E3-ZWW-CVK-R6900-01 kernel: [10961315.384238] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:26:58 L01-E3-ZWW-CVK-R6900-01 kernel: [10961349.366801] (o2hb-69E6B455D9,11460,16):o2hb_thread:1447 do disk heartbeat used 2790 msecs on device(dm-3), ret = 0.

Jul 23 18:27:16 L01-E3-ZWW-CVK-R6900-01 kernel: [10961367.449795] sd 1:0:3:0: [sdo] tag# 0 FAILED Result: hostbyte=DID_ERROR driverbyte=DRIVER_OK

Jul 23 18:27:16 L01-E3-ZWW-CVK-R6900-01 kernel: [10961367.449814] sd 1:0:3:0: [sdo] tag# 0 CDB: Read(16) 88 00 00 00 00 00 63 05 86 00 00 00 02 00 00 00

Jul 23 18:27:16 L01-E3-ZWW-CVK-R6900-01 kernel: [10961367.449816] print_req_error: I/O error, dev sdo, sector 1661306368

Jul 23 18:27:16 L01-E3-ZWW-CVK-R6900-01 kernel: [10961367.449874] device-mapper: mult ipath: Failing path 8:224.

Jul 23 18:27:19 L01-E3-ZWW-CVK-R6900-01 kernel: [10961370.416732] device-mapper: mult ipath: Reinstating path 8:224.

Jul 23 18:27:19 L01-E3-ZWM-CVK-R6900-01 kernel: [10961370.425136] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

Jul 23 18:27:19 L01-E3-ZWM-CVK-R6900-01 kernel: [10961370.425232] sd 1:0:3:0: alua: port group 01 state A preferred supports tolusnA

从上面的日志可以看出，是多路径链路震荡，导致主机访问共享存储阻塞，引起集群锁。通过反复执行集群锁检测脚本可以看出，一会出现集群锁，一会又正常，与日志中多路径频繁断开再恢复能够对应上：

```
root@L01-E3-ZWM-CVK-R6900-01:/var/log# ocf2_kernel_stat_collect.sh
Process have been hung up!
Find busy locks on disk /dev/dm-3.
Find busy locks on disk /dev/dm-2.
Find busy locks on disk /dev/dm-0.
Find busy locks on disk /dev/dm-4.
Find busy locks on disk /dev/dm-1.
Maybe power off nodes( L01-E3-ZWM-CVK-R6900-01 ) could unblock ocf2-cluster, then power on it/them until ocf2-cluster recovers to normal.
root@L01-E3-ZWM-CVK-R6900-01:/var/log# ocf2_kernel_stat_collect.sh
Process have been hung up!
root@L01-E3-ZWM-CVK-R6900-01:/var/log# cat /etc/cas_cvk-version
V900000100000025005 V5.0 E0526005 Enterprise 48706
Build 2019-03-08 17:32:04. RELEASE SOP/TW/RE
root@L01-E3-ZWM-CVK-R6900-01:/var/log# ocf2_kernel_stat_collect.sh
root@L01-E3-ZWM-CVK-R6900-01:/var/log# ocf2_kernel_stat_collect.sh
```

当I/O阻塞严重时，触发主机fence重启，ocfs2_fence_restart.log日志记录如下：
Restarted,Thu Jul 23 21:59:00 2020,dm-0,774E8026D9E642519B7AA9A7FA35D908,DISK fault leading to being fenced
Restarted,Thu Jul 23 22:56:48 2020,dm-1,B14C005525B4442C85761DB0D6C69330,DISK fault leading to being fenced
Restarted,Thu Jul 23 23:42:16 2020,dm-2,DF5A141CCF3E44F38C99AF272A6D1291,DISK fault leading to being fenced
导致链路震荡的可能原因，包括，HBA卡/光模块/光纤线/FC交换机/存储等。
通过现场排查发现存储上一个端口有光衰的情况：

-----Port 1:0:2 DDM-----						
-Warning- --Alarm--						
--Type--	Units	Reading	Low	High	Low	High
Temp	C	35	0	70	-5	75
Voltage	mV	3350	3100	3500	3000	3600
TX Bias	mA	7	2	11	1	12
TX Power	uW	289	158	794	125	1000
RX Power	uW	410	15	1000	10	1258

通过disable此光衰端口后，主机访问共享存储恢复正常。

解决方法

通过目前的信息确认，共享存储阻塞问题，原因是链路震荡导致，链路震荡原因是存储端一个端口出现光衰导致。
对应此种情况，建议定期检查FC交换机及存储光模块光功率，发现问题尽快更换解决，避免引起严重后果。