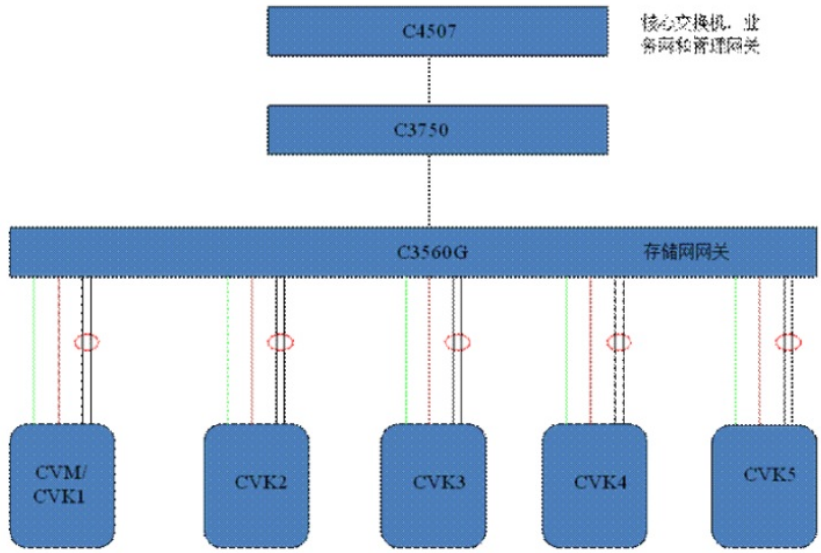
CVK主机中病毒导致集群主机反复重启问题

蒋立明 2016-04-19 发表

某局点集群异常重启，且异常重启后服务器处于反复重启的状态。现场远程业务恢复后，该局点集群几周后再次发生异常重启现象。如下图所示，是该局点组网图，现场一共五台CVK主机，其中一台CVK主机即是CVM管理节点又是CVK计算节点，且这五台CVK都是零存储节点。



远程进行问题定位后发现处于异常重启的四台服务器的有一个关于零存储的系统配置文件已损坏，修复配置文件后再次重启，服务器正常启动，再次重启也可以正常启动。确认服务器可以正常启后，CAS平台上的业务也恢复了正常。

2.1集群第一次异常重启原因分析

2.1.1分析“ocfs2_fence_restart”日志

发现系统重启是因为CAS平台的fence保护机制重启了服务器。

```
o2quo_fence_self: Heartbeat or quorum error, restart at 1452161689.790394 (Beijing Time: 2016-1-7 18:14:49), please have a check.
```

2.1.2分析syslog日志

从syslog日志中可以看出fence重启的原因是CAS平台上的共享存储文件无法访问。对应的日志如下：

```
Jan  7 18:14:12 CQYT-SXB-CAS-4 kernel: [21010409.259873] connection1:0: detected conn error (1020)
Jan  7 18:14:13 CQYT-SXB-CAS-4 iscsid: Kernel reported iSCSI connection 1:0 error (1020 - ISCSI_ERR_TCP_CONN_CLOSE: TCP connection closed) state (3)
Jan  7 18:14:15 CQYT-SXB-CAS-4 iscsid: connect to 127.0.0.1:3260 failed (Connection refused)
Jan  7 18:14:30  iscsid: last message repeated 3 times
Jan 7 18:14:30 CQYT-SXB-CAS-4 kernel: [21010426.859231] o2net: Connection to node CQYT-SXB-CAS-3 (num 3) at 10.78.246.215:7100 has been idle for 90.208 secs.
Jan  7 18:14:30 CQYT-SXB-CAS-4 kernel: [21010426.859250] o2net_idle_timer 1628: Local and remote node is heartbeating, and try connect
Jan  7 18:14:30 CQYT-SXB-CAS-4 iscsid: connect to 127.0.0.1:3260 failed (Connection refused)
Jan  7 18:14:43  iscsid: last message repeated 3 times
```

2.1.3分析零存储日志

检查零存储日志可以看到如下的错误：

```
2016-01-07 18:13:11[1452161591]:../net/net_table.c:1941 : netable_closesock( ): 2994 :140402006136576: INFO: 1_v1422339237 closed name:(172.26.254.1/0), closed by peer
2016-01-07 18:13:11[1452161591]:../diskpool/disktable.c:645 : __disktable_reset_handler( ): 2994 :140402006136576: INFO: disktable 172.26.254.1/0 offline
2016-01-07 18:13:11[1452161591]:../diskpool/diskmap.c:993 : diskmap_del( ): 2994 :140402006136576: INFO: del 172.26.254.1/0
2016-01-07 18:13:11[1452161591]:../diskpool/diskmap.c:650 : __diskmap_del_entry( ): 2994 :140402006136576: INFO: del disk 0, parent 0  list 4
2016-01-07 18:13:11[1452161591]:../diskpool/diskmap.c:650 : __diskmap_del_entry( ): 2994 :14040
```

```
2006136576: INFO: del node CQYT-SXB-CAS-1, parent 4  list 4
2016-01-07 18:13:11[1452161591]:../msgqueue/mq_master.c:156 : __mq_master_remove( ): 2994 :
140402006136576: INFO: close peer left 3
2016-01-07 18:13:11[1452161591]:../net/net_table.c:1434 : __netable_close_nolock( ): 2994
:140402006136576: INFO: net 172.26.254.1/0 closed
2016-01-07 18:13:11[1452161591]:../net/net_table.c:353 : __netable_stm_hb( ): 2994 :14040111800
4992: INFO: conn 172.26.254.3/0/172.26.254.3 lost hb ack 1
2016-01-07 18:13:12[1452161592]:../node/paxos_srv.c:492 : paxos_srv_getvote( ): 2994
:140401357534976: INFO: get voteid from 172.26.254.1/0 2 --> 2
2016-01-07 18:13:13[1452161593]:../node/paxos_srv.c:492 : paxos_srv_getvote( ): 2994
:140401357534976: INFO: get voteid from 172.26.254.1/0 2 --> 2
2016-01-07 18:13:13[1452161593]:../net/net_table.c:2056 : __netable_worker( ): 2994
:140401118004992: INFO: netable worker
2016-01-07 18:13:14[1452161594]:../net/net_table.c:341 : __netable_stm_hb( ): 2994 :14040113872
6656: INFO: conn 172.26.254.3/0/172.26.254.3 timeout
2016-01-07 18:13:14[1452161594]:../net/net_table.c:1941 : netable_closesock( ): 2994 :1404011387
26656: INFO: 4_v1422339237 closed name:(172.26.254.3/0), timeout at hb
```

此后出现零存储进程重启的情况：

```
2016-01-07 18:13:35[1452161615]:../node/node_srv.c:471 : __node_admin_check( ): 2994
:140401349142272: WARNING: check peer, need 5, min 2, got 0
2016-01-07 18:13:35[1452161615]:../node/node_srv.c:667 : node_admin_start( ): 2994 :1404013491
42272: WARNING: cluster not ready, exist
2016-01-07 18:13:35[1452161615]:../node/env.c:423 : env_server_run( ): 2980 :140402129090368: I
NFO: worker require restart, restarting...
2016-01-07 18:13:35[1452161615]:tpool.c:46 : __tpool_new( ): 7002 :140402129090368: INFO:
tpool worker create new threads, total 1, left 0 idle 1
2016-01-07 18:13:35[1452161615]:../net/hosts.c:185 : __hosts_insert( ): 7002 :140402129090368: I
NFO: host:172.26.254.1 site:default rack:default node:CQYT-SXB-CAS-1
2016-01-07 18:13:35[1452161615]:../net/hosts.c:185 : __hosts_insert( ): 7002 :140402129090368: I
NFO: host:172.26.254.2 site:default rack:default node:CQYT-SXB-CAS-2
2016-01-07 18:13:35[1452161615]:../net/hosts.c:185 : __hosts_insert( ): 7002 :140402129090368: I
NFO: host:172.26.254.3 site:default rack:default node:CQYT-SXB-CAS-3
2016-01-07 18:13:35[1452161615]:../net/hosts.c:185 : __hosts_insert( ): 7002 :140402129090368: I
NFO: host:172.26.254.4 site:default rack:default node:CQYT-SXB-CAS-4
2016-01-07 18:13:35[1452161615]:../net/hosts.c:185 : __hosts_insert( ): 7002 :140402129090368: I
NFO: host:172.26.254.5 site:default rack:default node:CQYT-SXB-CAS-5
2016-01-07 18:13:35[1452161615]:tpool.c:46 : __tpool_new( ): 7002 :140402129090368: INFO:
tpool sdevent[0] create new threads, total 1, left 0 idle 1
2016-01-07 18:13:35[1452161615]:tpool.c:46 : __tpool_new( ): 7002 :140402129090368: INFO:
tpool sdevent[1] create new threads, total 1, left 0 idle 1
2016-01-07 18:13:35[1452161615]:tpool.c:46 : __tpool_new( ): 7002 :140402129090368: INFO:
tpool sdevent[2] create new threads, total 1, left 0 idle 1
```

至此可以确认第一次重启的原因为：零存储网络出现异常引发了零存储服务的重新启动，之后因CAS集群上使用了零存储作为共享文件系统，CAS集群发现共享文件系统无法访问后，触发了fence保护机制重启了服务器以尝试重启关联共享文件系统。

2.1.4分析服务器VMCORE文件的栈信息

为了进一行步确认出现存储网网络异常的原因，我们继续分析了服务器的vmcore日志。因为CVK服务器在重启过程中留下了vmcore文件，而vmcore文件是根据出问题时的服务器内存数据生成的。所以分析环境中的VMCORE文件可以进一步确认问题原因。

从VMCORE文件中分析确认存储网网络不通时，CVK服务器的OVS也出现了异常。但引发OVS异常的原因没有更多的日志进行分析。之后，我司研发在实验室根据以上的栈信息进行反复模拟测试后确认：OVS收到某种异常报文后，触发了OVS正常情况下处于关闭的内部接口被打开，同时该接口收到了IPV6报文，但该接口无法处理此报文，最终导致OVS也出现异常。

综上分析，引发第一次异常重启的原因为：服务器OVS收到网络侧发过来的异常报文后出现OVS工作异常，之后出现零存储存储内网通信异常，最终出现服务器异常重启。

2.2反复重启的原因分析

2.2.1查看系统日志

服务器反复重过程期间，通过分析相关信息和日志发现存在以下问题，分析反复重启时的syslog日志，发现存在以下错误信息：

```
root@cvknode3:~/jfl06974# zgrep -i "bad" ~/cas_0112/cas/*/logdir/var/log/syslog.*.gz
/root/cas_0112/cas/CQYT-SXB-CAS-1/logdir/var/log/syslog.4.gz:Jan  8 09:21:03 CQYT-SXB-CAS-1
cron[3015]: Error: bad minute; while reading /etc/crontab
/root/cas_0112/cas/CQYT-SXB-CAS-1/logdir/var/log/syslog.4.gz:Jan  8 09:22:01 CQYT-SXB-CAS-1
cron[3015]: Error: bad minute; while reading /etc/crontab
```

此时怀疑是crontab文件可能存在问题，之后分析crontab文件的详细发现以下信息：

2.2.2查看/etc/crontab文件异常信息

```
# /etc/crontab: system-wide crontab
# Unlike any other crontab you don't have to run the `crontab'
# command to install the new version when you edit this file
# and files in /etc/cron.d. These files also have username fields,
# that none of the other crontabs do.
SHELL=/bin/sh
PATH=/usr/local/sbin:/usr/local/bin:/sbin:/bin:/usr/sbin:/usr/bin
# m h dom mon dow user  command
17 *    * * *   root    cd / && run-parts --report /etc/cron.hourly
25 6    * * *   root    test -x /usr/sbin/anacron || ( cd / && run-parts --report /etc/cron.daily )
47 6    * * 7   root    test -x /usr/sbin/anacron || ( cd / && run-parts --report /etc/cron.weekly )
52 6    1 * *   root    test -x /usr/sbin/anacron || ( cd / && run-parts --report /etc/cron.monthly )
#
* *    * * *   root    /opt/bin/cha_real_check_cvm.sh
* *    * * *   root    /opt/bin/cha_real_check_cvk.sh
* *    * * *   root    /opt/bin/libvirtd_check.sh
* *    * * *   root    /opt/bin/ocfs2_iscsi_conf_chg_timer.sh
*/10 *    * * *   root   python /opt/bin/ocfs2_cluster_config.pyc -s
1 2    * * *   root    /opt/bin/cas_clean_log.sh
* *    * * *   root    /opt/bin/novnc_check.sh
* *    * * *   root    /opt/bin/sys_alarmd_check.sh
* *    * * *   root    /opt/bin/tomcat_check.sh
^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^
@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^
@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^
@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^
@^@*/10
* * * * root /opt/mds/lich//bin/lich.node --recover >> /opt/mds/log/recover.log
1 */1 * * * root /opt/mds/lich//bin/lich.node --metabalance >> /opt/mds/log/metabalance.log
1 */8 * * * root /opt/mds/lich//bin/lich.node --chunkbalance --rand >> /opt/mds/log/chunkbalance.log
*/10 * * * * root /opt/mds/lich//bin/lich.node --disk_check --cache >> /opt/mds/log/disk_cache.log
*/3 * * * * root /etc/cron.hourly/cron.sh
```

正常情况下应该没有红色字体处的乱码信息，至此可以确认crontab文件已经损坏。由于零存储会在启动时或者正常退出时会修改crontab文件，怀疑服务器异常重启导致了crontab文件损坏。之后将crontab修复后再次重启服务器，服务器运行恢复正常。

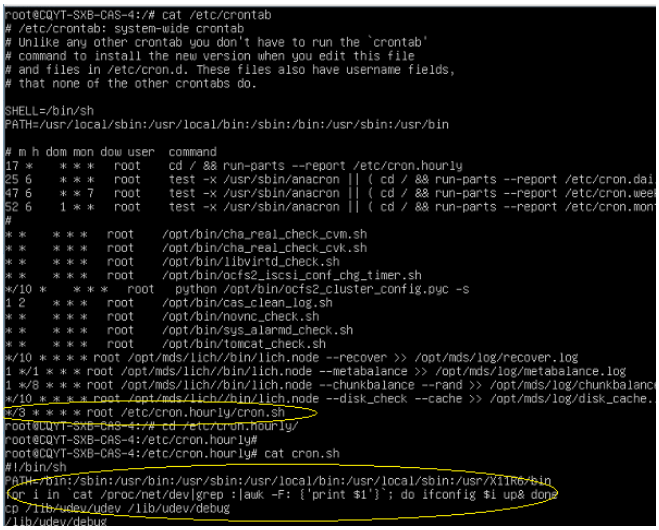
2.3几周后再次异常重启原因分析

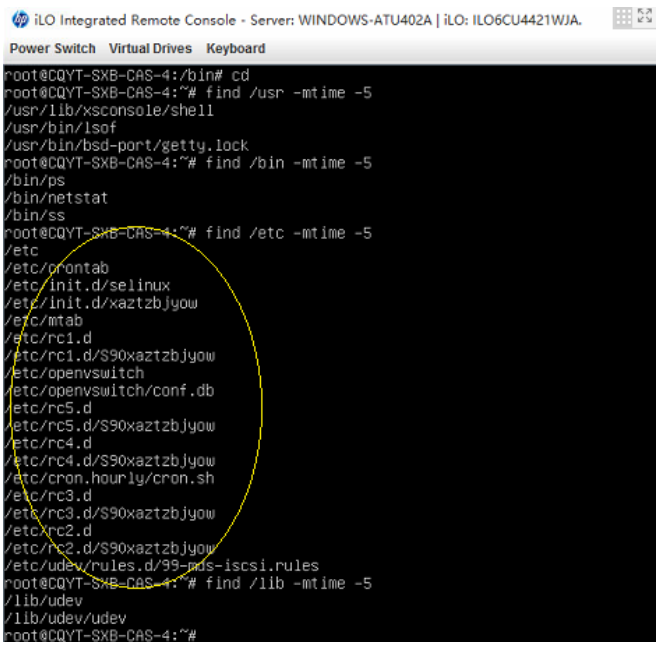
当时根据CVK系统奔溃时产生的vmcore文件分析，得知crontab文件被写坏，导致ovs内核内存异常，引发系统反复重启。因为急于恢复环境，没有保留太多现场信息，所以文件被写坏的根本原因没有机会定位。最后的解决方案是关闭CVK上ipv6功能先临时规避问题恢复CVK，后来提供CAS升级包针对该问题增加CAS内核的稳定性。

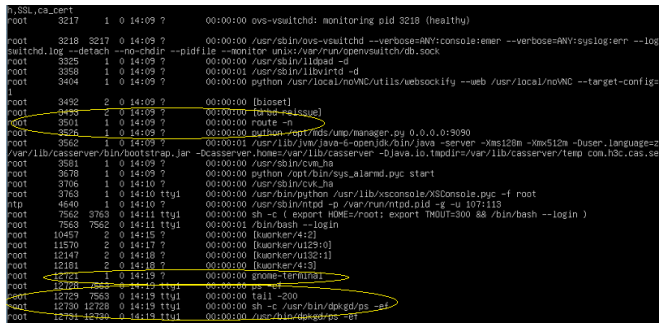
该问题复现后，发现内核出错点和上次是同样的现象。这次详细检查了出问题的CVK，发现它中了病毒，病毒也会不时向crontab写入任务，可能在某个时机和零存储写该文件时产生冲突。CVK中病毒体现在：

1. CVK定时任务中被加入异常任务（每3分钟执行一个可疑文件）
2. CVK开机启动项中被加入异常任务
3. CVK主机磁盘上存在一些非正常可疑文件
4. CVK内存中存在一些伪装的可疑进程
5. 根据以往经验，中毒CVK连入网络后，会导致用户网络拥塞以及向外发起一些非正常的连接。（用户可以根据之前出问题的环境验证这一点）

以下是关于上面现象的一些截图：







综上分析：CVK主机中病毒。该病毒会有一系列的动作，可能和CVK本身的一部分功能存在冲突，导致系统配置变化，文件破坏，甚至引发内核异常反复重启。CAS已经提供了升级包，增强系统稳定性，避免奔溃重启。

但是病毒可能引发的其他问题，比如占用系统资源，占用网络资源，执行破坏性工作。这些由病毒自身特性行为决定，CAS这边很难评估它的具体危害性。

1.不要使用简单的root密码

2.如非必要，不要把CVK主机映射到公网上。

3.中病毒后，建议重装CAS系统，杜绝病毒。建议安装新的CAS版本，它在稳定性和安全性方面都做了一些提高。

4.因急于恢复用户环境及保存用户数据，当时先手工临时清除了CVK上的病毒，包括删除可疑文件和

可疑任务，杀掉可疑进程等工作。这些措施临时恢复了环境，但不能保证还有潜伏的病毒会在以后爆发，所以还是建议后续重装系统。