

知 CAS修改将共享文件系统分布式锁改为硬件辅助锁的方法

张昊 2020-09-30 发表

组网及说明

CAS E0535H08版本及以上

问题描述

CAS E0535H05版本后开始支持修改磁盘锁的类型，E0535H08版本完善后，可以考虑将磁盘锁类型手动修改，规避管理网异常可能导致的集群fence等问题。

过程分析

修改磁盘锁类型需注意存储支持CAW，才可以进行修改。检查方法如下：

cvknode# sg_vpd -p bl /dev/dm-xx //cvk上执行

Block limits VPD page (SBC):

Write same non-zero (WSNZ): 0
Maximum compare and write length: 1 blocks //如果是0表示不支持CAW
Optimal transfer length granularity: 32 blocks
Maximum transfer length: 32768 blocks
Optimal transfer length: 32768 blocks
Maximum prefetch length: 0 blocks
Maximum unmap LBA count: 65536
Maximum unmap block descriptor count: 10
Optimal unmap granularity: 32
Unmap granularity alignment valid: 1
Unmap granularity alignment: 2147483648
Maximum write same length: 0x0 blocks
Maximum atomic transfer length: 0
Atomic alignment: 0
Atomic transfer length granularity: 0

解决方法

修改磁盘锁的方法有两种，首先在存储卷没有数据时，可以直接格式化的方法，进行修改；如果存储卷内有数据，则需要手动进行修改。

手动修改方法如下：

- 1、确保主机的管理网络、存储网络全部处于健康状态。
- 2、检查存储池剩余空间是否满足转换条件数据存储可用空间必须大于1%，同时可用空间需大于1G。
- 3、关闭或迁移存储池上的所有虚拟机，在所有节点上暂停需要修改的存储池，并确认所有节点对应的存储池已经全部暂停。
- 4、在其中任何一个节点执行fsck.ocfs2 -f /dev/dm-xx, 检查对应存储池是否有错，如有则用fsck命令进行修复。
- 5、备份文件系统元数据，以备不时之需，命令：o2image <device> <image file>，备份文件需要4G+的空间，备份文件要存放到非转换的lun上。

例：root@wxbcvknode96:/# o2image /dev/dm-4 /home/dm-4.bak

- 6、当以上条件都满足后，执行命令进行硬件辅助锁格式转换

virsh pool-list --all查看对应的pool name

fsmcli setlocktype --name <pool name> dlock 进行转换

- 7、转换完成，重新启动存储池。在前台界面 主机池->共享文件系统页面查看存储池锁类型已变成硬件辅助锁。

注意事项：

- 1、转换过程中，不允许进行任何与存储、虚拟机相关的业务操作。
- 2、转换过程中，若主机重启、掉电等异常导致转换没有完成，可用步骤5中备份的元数据进行恢复，恢复完成后再次执行转换。

恢复方法：o2image -l <device> <image file>

例：o2image -l /dev/dm-4 /home/dm-4.bak

- 3、转换完成，确保集群一切正常后，方可手动删除备份文件。