



CAS 虚拟机迁移存储停止任务后磁盘IO延迟高

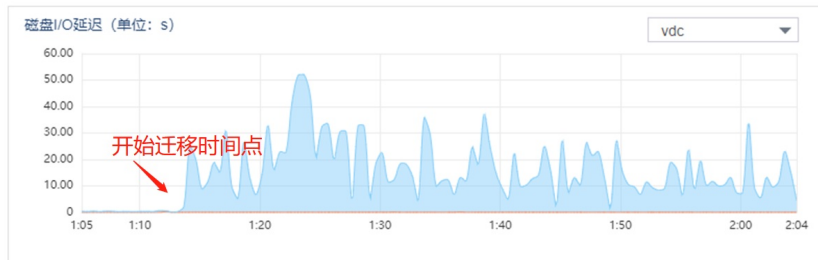
薛金涛 2021-12-13 发表

组网及说明

不涉及

问题描述

1. 在线迁移虚拟机存储，迁移时磁盘IO达到50s左右，严重影响业务



2. 后台kill掉相关磁盘cp进程后，IO延迟仍然居高不下

```
[root@cvk18-GPU03 ~]# lsof /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
COMMAND PID USER  FD  TYPE DEVICE          SIZE/OFF      NODE NAME
kvm      806  root   40r  REG  253,1 1610603560960 602409630 /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
cp       70158 root    3r  REG  253,1 1610603560960 602409630 /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
[root@cvk18-GPU03 ~]#
```

```
[root@cvk18-GPU03 ~]# kill -9 70158
[root@cvk18-GPU03 ~]# lsof /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
COMMAND PID USER  FD  TYPE DEVICE          SIZE/OFF      NODE NAME
kvm      806  root   40r  REG  253,1 1610603560960 602409630 /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
[root@cvk18-GPU03 ~]# lsof /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
COMMAND PID USER  FD  TYPE DEVICE          SIZE/OFF      NODE NAME
kvm      806  root   40r  REG  253,1 1610603560960 602409630 /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
```

过程分析

1. 查看CVK主机上对应共享存储卷IO状态正常



2. 后台查看对应磁盘镜像, 发现该磁盘镜像有临时镜像, 后台执行在线commit命令 `virsh blockcommit $vmName $incImageName --active --shallow --persist --pivot --wait`, 合并镜像后, 问题依旧

```
root@cvk18-GPU03 log]# qemu-img info /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd_inc_1.1639329165632_2_807 --backing-chain
image: /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd_inc_1.1639329165632_2_807
file format: qcow2
virtual size: 1.5T (1610612736000 bytes)
disk size: 26G
cluster_size: 262144
backing file: /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd_inc_1
backing file format: qcow2
format specific information:
  compat: 1.1
  lazy refcounts: false
  refcount bits: 16
  corrupt: false

image: /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd_inc_1
file format: qcow2
virtual size: 1.5T (1610612736000 bytes)
disk size: 535G
cluster_size: 262144
backing file: /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
backing file format: qcow2
format specific information:
  compat: 1.1
  lazy refcounts: false
  refcount bits: 16
  corrupt: false

image: /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
file format: qcow2
virtual size: 1.5T (1610612736000 bytes)
disk size: 1.5T
cluster_size: 262144
```

```
[root@cvk18-GPU03 ~]# qemu-img info /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd_inc_1 --backing-chain
image: /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd_inc_1
file format: qcow2
virtual size: 1.5T (1610612736000 bytes)
disk size: 564G
cluster_size: 262144
backing file: /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
backing file format: qcow2
format specific information:
  compat: 1.1
  lazy refcounts: false
  refcount bits: 16
  corrupt: false

image: /vms/hdd_50T/volume-ce5ca99b-aa51-4129-88a3-b2a8bd8c95cd
file format: qcow2
virtual size: 1.5T (1610612736000 bytes)
disk size: 1.5T
cluster_size: 262144
format specific information:
  compat: 1.1
  lazy refcounts: false
  refcount bits: 16
  corrupt: false
```

3. 查看对应磁盘是否限速: `virsh blkdeviotune <VMName> --device vdc`, 回显如下

```
[root@cvk18-GPU03 ~]# virsh blkdeviotune huanghe1-3-140 --device vdc
total_bytes_sec: 0
read_bytes_sec: 0
write_bytes_sec: 1048576
total_iops_sec: 0
read_iops_sec: 0
write_iops_sec: 0
total_bytes_sec_max: 0
read_bytes_sec_max: 0
write_bytes_sec_max: 0
total_iops_sec_max: 0
read_iops_sec_max: 0
write_iops_sec_max: 0
size_iops_sec: 0
group_name: drive-virtio-disk2
total_bytes_sec_max_length: 0
read_bytes_sec_max_length: 0
write_bytes_sec_max_length: 0
total_iops_sec_max_length: 0
```

read_iops_sec_max_length: 0

write_iops_sec_max_length: 0

解决方法

可以看到有对应限速在，可知，只kill掉cp命令未能解除迁移速度限制，需要手动解除限制，使用命令

:

```
virsh blkdeviotune <VMName> --device vda --write_bytes_sec 0
```

