

知 UIS虚拟机定时发生故障重启

王向杰 2023-03-29 发表

问题描述

UISE0720, 有一台虚拟机每天五点半左右发生故障重启。



完成时间	登录地址	操作分类	操作对象	操作描述	执行
2023-03-03 05:35:10	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-03-02 05:52:07	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-03-01 09:19:18	53.208.153.154	虚拟机动作	基础组件服务器	修改虚拟机基础组件服务器。	成功
2023-03-01 08:44:03	53.208.153.154	虚拟机动作	基础组件服务器	修改虚拟机基础组件服务器cpu个数: 40 -->...	成功
2023-03-01 05:43:04	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-28 05:43:01	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-27 05:54:58	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-26 05:51:55	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-25 05:50:22	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-24 17:47:18	10.129.64.246	虚拟机动作	基础组件服务器	修改虚拟机基础组件服务器cpu个数: 20 -->...	成功
2023-02-24 06:21:20	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-23 05:38:17	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-22 05:39:44	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-21 06:57:11	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-20 08:34:08	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-19 06:42:05	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-17 06:47:59	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-16 06:01:56	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-15 06:05:23	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功
2023-02-14 05:33:50	127.0.0.1	虚拟机动作	基础组件服务器	虚拟机基础组件服务器故障重启成功。	成功

过程分析

1、确认同主机其他虚拟机未发生相同现象，检查/var/log/libvirt/qemu下的虚拟机日志发现虚拟机被异常shutdown，原因为crash。

```
2023-03-03 21:36:37.544+0000: Domain id=62 is tainted: custom-ga-command
2023-03-05 05:25:18.226+0000: shutting down, reason=crashed
2023-03-05 05:26:46.125+0000: starting up Libvirt version: 4.0.0, package: 1.e17 (Unknown, 2021-02-24-14:20:17, build-e0705), qemu version: 2.12.0qemu-2.12-1.e17, hostname: UIS-CVK08
```

2、检查对应主机日志发现同一时间点，虚拟机进程使用内存触发cgroup阈值，进程被异常kill。

```
inactive_anon:10KB active_anon:117089376KB inactive_file:729KB active_file:0KB unevictable:0KB
Mar 5 05:25:16 UIS-CVK08 kernel: [ pid ] uid tgid total_vm rss nr_ptes nr_pmds swappents oom_score_adj name
Mar 5 05:25:16 UIS-CVK08 kernel: [ 40256 ] 0 40256 4506927 3984261 5049 20 0 0 kvm
Mar 5 05:25:16 UIS-CVK08 kernel: [ 7031 ] 0 7031 17394096 16764621 33062 69 0 0 kvm
Mar 5 05:25:16 UIS-CVK08 kernel: [ 18270 ] 0 18270 2403932 2002029 4330 12 0 0 kvm
Mar 5 05:25:16 UIS-CVK08 kernel: [ 44249 ] 0 44249 32444433 29504399 57993 121 0 0 kvm
Mar 5 05:25:16 UIS-CVK08 kernel: Memory cgroup out of memory: Kill process 44249 (kvm) score 546 or sacrifice child
Mar 5 05:25:16 UIS-CVK08 kernel: Killed process 44249 (kvm) total_vm:1297773248, anon-rss:117089376KB, file-rss:2902048, shmem-rss:2096
Mar 5 05:25:17 UIS-CVK08 kernel: device vnet1 left promiscuous mode
Mar 5 05:25:17 UIS-CVK08 kvm: 3 guests now active
Mar 5 05:25:17 UIS-CVK08 systemd-machined: Machine qemu-02-ljchuzujm-36 terminated.
Mar 5 05:25:18 UIS-CVK08 ovs-vsctl: ovs|00001|vsctl|INFO|called as ovs-vsctl --timeout=5 ... --if-exists del-port vnet1
Mar 5 05:25:19 UIS-CVK08 ntpd[65000]: Deleting interface #70 vnet1, fe80::fc16:3eff:fedd:a501e123, interface stats: received=0, sent=0, dropped=0, active_time=85743 secs
Mar 5 05:26:46 UIS-CVK08 ovs-vsctl: ovs|00001|vsctl|INFO|called as ovs-vsctl --timeout=5 ... --if-exists del-port vnet1 --add-port vs_business vnet1 tag=1003 -- set Interface vnet1
Interface vnet1 "external_id:attached-mac='f4:16:3e:0d:95:011'" -- set Interface vnet1 "external_id:iface_id='3966dafa-6554-470a-f5dd-914d2d7059d3'" -- set Interface vnet1
"external_id:vs-id='ae5da9eb-d644-446e-ab42-57d9f2219bc6'" -- set Interface vnet1 external_id:iface-status=active -- set Interface vnet1 other_config:tes=0 -- set Interface vnet1
```

3、确认该节点为超融合节点，检查/etc/cgconfig.conf文件。

```
[root@UIS-CVK08 ~]# cat /etc/cgconfig.conf
group onestor
{
    memory
    {
        memory.swappiness=2;
        memory.oom_control=1;
    }
}
group system.slice
{
    memory
    {
        memory.swappiness=2;
        memory.oom_control=1;
    }
}
group machine.slice
{
    memory
    {
        memory.limit_in_bytes=200g;
        memory.swappiness=20;
    }
}
[root@UIS-CVK08 ~]#
```

经确认，现场在该时间段虚拟机内有一批任务执行会导致虚拟机使用的内存上升，现场该主机上所有虚拟机内存总和为208G，超出cgroup阈值。Cgroup触发是根据虚拟机kvm进程实际使用的物理内存来计算的。由于kvm进程本身也需要消耗资源，所以一般虚拟机kvm进程会比虚拟机内部显示使用内存大。所以在配置内存时候还要考虑预留一部分内存，不然容易导致虚拟机被随机kill掉。

解决方法

- 1、给该主机扩充内存;
- 2、将该内存大的虚拟机迁移到其他内存充足的主机上。

