



CloudOS nexus容器crash问题

汪普深 2021-03-31 发表

组网及说明

无

问题描述

CloudOS 平台nexus容器出现crash无法正常启动问题。
查看nexus容器处于CrashLoopBackoff状态：

```
[root@xsq-omz-cloudos-mgser-02 ~]# pod |grep -v Running
NAMESPACE   NAME          READY   STATUS    RESTARTS   AGE   IP            NODE           NOMINATED NODE
k8s-form    nexus-jdlde   0/1     CrashLoopBackOff   20      3h    10.101.67.7   10.72.195.4   <none>
```

过程分析

通过查看crash容器的日志有如下报错，怀疑是共享存储卷文件系统损坏导致I/O读写异常，引起容器crash:

```
2021-02-26 02:26:42.276+0000 ERROR [RelixStartLevel] *SYSTEM org.sonatype.nexus.extender.NexusContextListener - Failed to start nexus
com.orienttechnologies.orient.core.exception.OSTorageException: Cannot open local storage "/nexus-data/db/accesslog" with mode=rw
at com.orienttechnologies.orient.core.storage.impl.local.OAbstractPaginatedStorage.open(OAbstractPaginatedStorage.java:323)
at com.orienttechnologies.orient.core.db.document.ODatabaseDocumentTx.open(ODatabaseDocumentTx.java:259)
at org.sonatype.nexus.orient.DatabaseManagerSupport.connect(DatabaseManagerSupport.java:143)
at org.sonatype.nexus.orient.DatabaseInstanceImpl.doStart(DatabaseInstanceImpl.java:36)
at org.sonatype.goodies.lifecycle.LifecycleSupport.start(LifecycleSupport.java:104)
at org.sonatype.goodies.lifecycle.Lifecycle.start(Lifecycle.java:44)
at org.sonatype.nexus.orient.DatabaseManagerSupport.createInstance(DatabaseManagerSupport.java:252)
at java.util.concurrent.ConcurrentHashMap.computeIfAbsent(ConcurrentHashMap.java:1660)
at org.sonatype.nexus.orient.DatabaseManagerSupport.getInstance(DatabaseManagerSupport.java:231)
at java.util.stream.ForEachOps$ForEachOp$OfRef.accept(ForEachOps.java:184)
at java.util.Spliterators$ArraySpliterator.forEachRemaining(Spliterators.java:948)
at java.util.stream.AbstractPipeline.copyInto(AbstractPipeline.java:481)
at java.util.stream.ForEachOps$ForEachTask.compute(ForEachOps.java:261)
at java.util.concurrent.CountedCompleter.exec(CountedCompleter.java:731)
at java.util.concurrent.ForkJoinTask.doExec(ForkJoinTask.java:289)
at java.util.concurrent.ForkJoinPool$WorkQueue.runTask(ForkJoinPool.java:1056)
at java.util.concurrent.ForkJoinPool.runWorker(ForkJoinPool.java:1692)
at java.util.concurrent.ForkJoinWorkerThread.run(ForkJoinWorkerThread.java:157)
Caused by: java.lang.NegativeArraySizeException: null
at com.orienttechnologies.orient.core.storage.cache.local.OWOWCache.readNextNameEntry(OWOWCache.java:1438)
at com.orienttechnologies.orient.core.storage.cache.local.OWOWCache.readNameMap(OWOWCache.java:1356)
at com.orienttechnologies.orient.core.storage.cache.local.OWOWCache.initNameMapping(OWOWCache.java:1333)
at com.orienttechnologies.orient.core.storage.cache.local.OWOWCache.loadRegisteredFiles(OWOWCache.java:1237)
at com.orienttechnologies.orient.core.storage.impl.local.OAbstractPaginatedStorage.initWalAndDiskCache(OAbstractPaginatedStorage.java:606)
at com.orienttechnologies.orient.core.storage.impl.local.OAbstractPaginatedStorage.open(OAbstractPaginatedStorage.java:285)
... 17 common frames omitted
```

通过fsck命令检查存储卷，确实有文件系统损坏，检查结果如下:

```
[root@xsq-omz-cloudos-mgser-03 accesslog]# fsck -fn /dev/mapper/nexus
fsck from util-linux 2.23.2
e2fsck 1.42.9 (28-Dec-2013)
warning! /dev/mapper/nexus is mounted.
Warning: skipping journal recovery because doing a read-only filesystem check.
Pass 1: Checking inodes, blocks, and sizes
Pass 2: Checking directory structure
Pass 3: Checking directory connectivity
Pass 4: Checking reference counts
Pass 5: Checking group summary information
Free blocks count wrong (50633924, counted=50633103).
Fix? no
Free inodes count wrong (13052674, counted=13052648).
Fix? no

/dev/mapper/nexus: 54526/13107200 files (3.3% non-contiguous), 1794876/52428800 blocks
[root@xsq-omz-cloudos-mgser-03 accesslog]# kubectl get rc
```

通过上面检查结果可以确认，共享文件存储卷文件系统损坏导致I/O异常，引起容器crash无法启动。

解决方法

1.通过如下命令停止nexus容器,并将存储卷取消挂载:

```
[root@xsq-omz-cloudos-mgser-03 ~]# kubectl scale --replicas=0 rc nexus -n terraform
replicationcontroller/nexus scaled
```

2.通过mkfs命令重新格式化此存储卷文件系统, fsck -fn /dev/mapper/nexus命令检查文件系统正常:

```
[root@xsq-omz-cloudos-mgser-03 log]# mkfs.ext4 /dev/mapper/nexus
mkfs: 1.42.9 (28-Dec-2013)
Discarding device blocks: done
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=4 blocks, stripe width=4096 blocks
13107200 inodes, 52428800 blocks
2621440 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=2199912448
1600 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000, 7962624, 11239424, 20480000, 23887872

Allocating group tables: done
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done

[root@xsq-omz-cloudos-mgser-03 log]# fsck -fn /dev/mapper/nexus
fsck from util-linux 2.23.2
```

3.再次使用命令kubectl scale --replicas=1 rc nexus -n terraform拉起容器, 并观察一段时间, 故障未复现:

```
[root@xsq-omz-cloudos-mgser-03 log]# kubectl scale --replicas=1 rc nexus -n terraform
replicationcontroller/nexus scaled
[root@xsq-omz-cloudos-mgser-03 log]# pod |grep nexus
terraform nexus-1xpb 1/1 Running 0 14s 10.101.67.7 10.71.195.4 <none>
terraform nexus-1xpb 1/1 Running 0 33s 10.101.67.7 10.71.195.4 <none>
terraform nexus-1xpb 1/1 Running 0 45s 10.101.67.7 10.71.195.4 <none>
terraform nexus-1xpb 1/1 Running 0 2m 10.101.67.7 10.71.195.4 <none>
terraform nexus-1xpb 1/1 Running 0 8m 10.101.67.7 10.71.195.4 <none>
terraform nexus-1xpb 1/1 Running 0 8m 10.101.67.7 10.71.195.4 <none>
terraform nexus-1xpb 1/1 Running 0 8m 10.101.67.7 10.71.195.4 <none>
terraform nexus-1xpb 1/1 Running 0 8m 10.101.67.7 10.71.195.4 <none>
terraform nexus-1xpb 1/1 Running 0 14m 10.101.67.7 10.71.195.4 <none>
terraform nexus-1xpb 1/1 Running 0 16m 10.101.67.7 10.71.195.4 <none>
```

后续建议:

共享存储卷文件系统损坏一般为存储链路不稳定存在链路抖动或存储端异常(如磁盘存在坏道等)导致, 建议排查存储链路光纤线质量, 光模块光衰等情况, 如果是磁盘坏道通过格式化可以修复。

