# X10000 2.0修复PG 不一致的步骤

告警监控 李志强 2020-06-29 发表

## 组网及说明

无

## 问题描述

集群中出现pg不一致告警

[root@node30 ~]# ceph -s cluster: id: d4a48f0e-490d-4d80-aef3-43e3c9c99298 health: HEALTH_ER R 2 scrub errors Possible data damage: 1 pg inconsistent services: mon: 3 daemons, quorum node30,node31,node32 mgr: node30(active), standbys: node31, node32 osd: 90 osds: 90 up, 90 in d ata: pools: 2 pools, 8192 pgs objects: 52231 objects, 202 GB usage: 2992 GB used, 326 TB / 329 TB avail pgs: 8191 active+clean 1 active+clean+inconsistent io: client: 115 MB/s wr, 0 op/s rd, 490 op/s w r

## 过程分析

无

## 解决方法

步骤1:提前关闭scrub和deep-scrub，任意节点执行：
ceph osd set noscrub
ceph osd set nodeep-scrub

步骤2：当没有pg在进行scrub和deep-scrub的时候，调整max_scrubs参数：
ceph tell osd.* injectargs --osd_max_scrubs=100
ceph tell mon.* injectargs --osd_max_scrubs=100
可以使用以下命令确认是否修改成功：
ceph daemon osd.X config show|grep max_scrubs

步骤3：根据ceph health detail 的输出找到需要修复的pg编号；

步骤4：执行修复命令：ceph pg repair PGID，且查看ceph –s有pg在进行repair

步骤5：等待pg修复完成，根据之前两个pg的修复情况可能会较长时间（2个小时以上）。
可以观察主osd的日志进行确认：
tailf /var/log/ceph/ceph-osd.X.log |grep fixed （当有fixed输出是表示修复成功）

步骤6：重复步骤3-步骤5修复其他pg，根据修复时间控制修复pg数量避免影响白天业务。

步骤7：结束修复之后将max_scrub值调回默认值：
ceph tell osd.* injectargs --osd_max_scrubs=1
ceph tell mon.* injectargs --osd_max_scrubs=1

可以使用以下命令确认是否修改成功：
ceph daemon osd.X config show|grep max_scrubs

步骤8：关开启scrub和deep-scrub，任意节点执行：
ceph osd unset noscrub
ceph osd unset nodeep-scrub