# ceph 重启osd io block问题

## 环境信息

操作系统：CentOS 7.1.1503

内核版本：3.10.0-229.el7.x86\_64

ceph版本：10.2.7

总共33个节点，其中3个节点部署3个mon，30个节点部署269个osd

2块万兆网卡组bond，划分2个子接口，分别用于ceph public和cluster网络， 3个mon节点mtu 1500，30个osd节点mtu 9000，使用4台交换机二层通信，交换机支持大包

## 问题现象

最近遇到3次slow request问题

第一次 起始时间 2021-05-25 13:54:45 终止时间 2021-05-25 15:05:50

第二次 起始时间2021-05-29 00:05:48 终止时间 2021-05-29 21:47:28

第三次 起始时间2021-06-03 09:45:53 终止时间 2021-06-03 10:00:38

其中第1次和第3次持续时间较短是手动将osd.331踢出集群

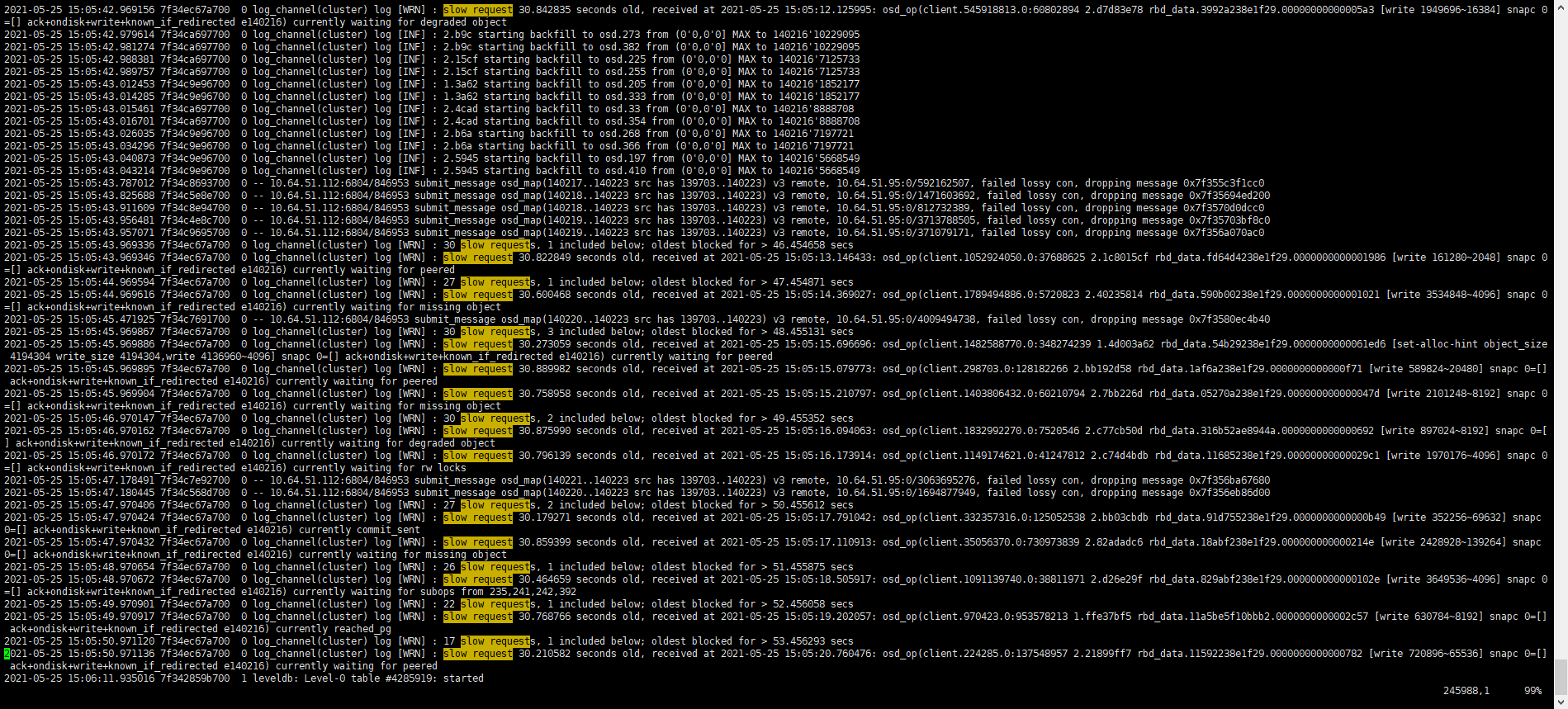
第2次是osd.328 down了之后被mon踢出集群，mon\_osd\_down\_out\_interval = 1800

## 原因分析

### 第一次问题分析

由于第一次问题持续时间较短，重启osd后业务已恢复，没有深入分析

当时也是osd.331导致的问题



### 第二次问题分析

1. 第2次slow request持续了一整天，也就说osd.328一直都是up状态，直到2021-05-29 21:47:28之后处于down状态集群才恢复正常

osd.328进程并没有挂掉，通过ceph osd tree命令看到状态为down，down的原因从日志能看到是由于心跳上报有问题，示例

1. 2021-05-31 10:20:01.910993 7f7e676bd700 -1 osd.328 143458 heartbeat\_check: no reply from x.x.x.x:6807 osd.315 since back 2021-05-29 04:43:10.520257 front 2021-05-29 04:43:10.520257 (cutoff 2021-05-31 10:19:41.910811)

通过检查osd.328跟其它osd心跳使用的ip和端口，例如osd.328和osd.315的public网卡心跳使用端口日志打印为6807，实际查看进程监听到的端口为6806和6816

[root@nodex ~]# ps -ef | grep 315

root      32526      1 37  2019 ?        221-02:45:52 /usr/bin/ceph-osd -f --cluster ceph --id 315 --setuser root --setgroup root

root     246006 245670  0 10:26 pts/0    00:00:00 grep --color=auto 315

[root@nodex ~]# netstat -anp | grep -i listen | grep 32526

tcp        0      0 x.x.x.x:6817    0.0.0.0:\*               LISTEN      32526/ceph-osd

tcp        0      0 x.x.x.x:6806       0.0.0.0:\*               LISTEN      32526/ceph-osd

tcp        0      0 x.x.x.x:6816       0.0.0.0:\*               LISTEN      32526/ceph-osd

tcp        0      0 x.x.x.x:6816    0.0.0.0:\*               LISTEN      32526/ceph-osd

telnet检查日志中使用的端口确实不通，更换正确的端口6806后能够正常telnet

[root@nodex ceph]# telnet x.x.x.x 6807

Trying x.x.x.x...

telnet: connect to address x.x.x.x: Connection refused

[root@nodex osd\_328]# telnet x.x.x.x 6806

Trying x.x.x.x...

Connected to x.x.x.x.

Escape character is '^]'.

ceph v027

@3ny

@3p^]

telnet> quit

Connection closed.

1. [root@nodex osd\_328]#

查看mon最新的osdmap，看到osd.315上报的端口为6806而不是6807，说明osd.328进程使用了错误的endpoint去上报心跳

[root@nodex ~]# ceph osd dump | grep osd.315

1. osd.315 up   in  weight 1 up\_from 143481 up\_thru 143923 down\_at 143469 last\_clean\_interval [26328,143477) x.x.x.x:6806/32526 x.x.x.x:6816/1032526 x.x.x.x:6817/1032526 x.x.x.x:6816/1032526 exists,up x-x-x-x-x

查看osd.328当前的osdmap版本和集群最新版本是一致的都是144891

[root@nodex osd\_328]# ceph daemon osd.328 status

{

    "cluster\_fsid": "x-x-x-x-x",

    "osd\_fsid": "x-x-x-x-x",

    "whoami": 328,

    "state": "active",

    "oldest\_map": 144391,

    "newest\_map": 144891,

    "num\_pgs": 827

}

[root@nodex osd\_328]# ceph -s

    cluster x-x-x-x-x

     health HEALTH\_OK

     monmap e1: 3 mons at {ceph-01=x.x.x.x:6789/0,ceph-02=x.x.x.x:6789/0,ceph-03=x.x.x.x:6789/0}

            election epoch 154, quorum 0,1,2 ceph-01,ceph-02,ceph-03

     osdmap e144891: 269 osds: 268 up, 268 in

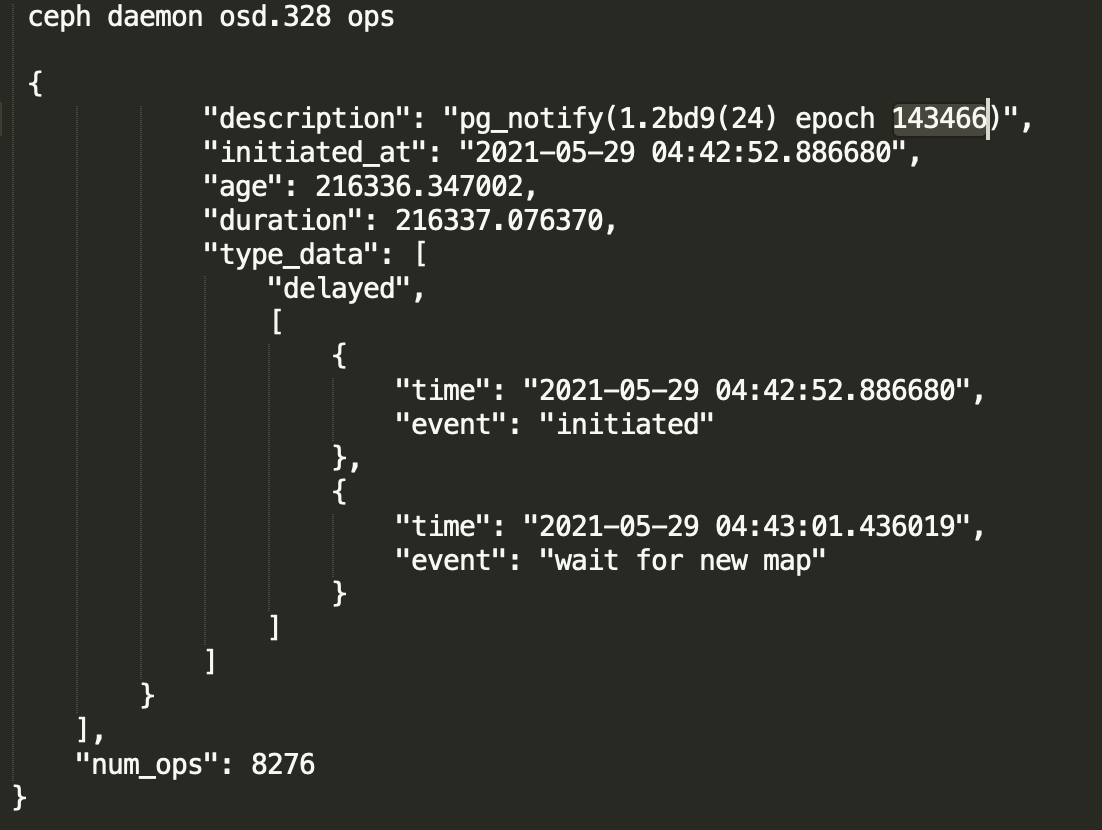
            flags sortbitwise,require\_jewel\_osds

      pgmap v49419952: 66240 pgs, 5 pools, 210 TB data, 54068 kobjects

            628 TB used, 1128 TB / 1757 TB avail

               66240 active+clean

1. client io 236 MB/s rd, 259 MB/s wr, 3923 op/s rd, 7669 op/s wr
2. 分析osd.328当前的op卡在wait for new map阶段，在epoch 143466版本时卡住，且当时osd.315确实有down过但osd.315进程id一直没变过，分析osd.315日志，可能是触发了rebind流程导致osd端口号发生变化

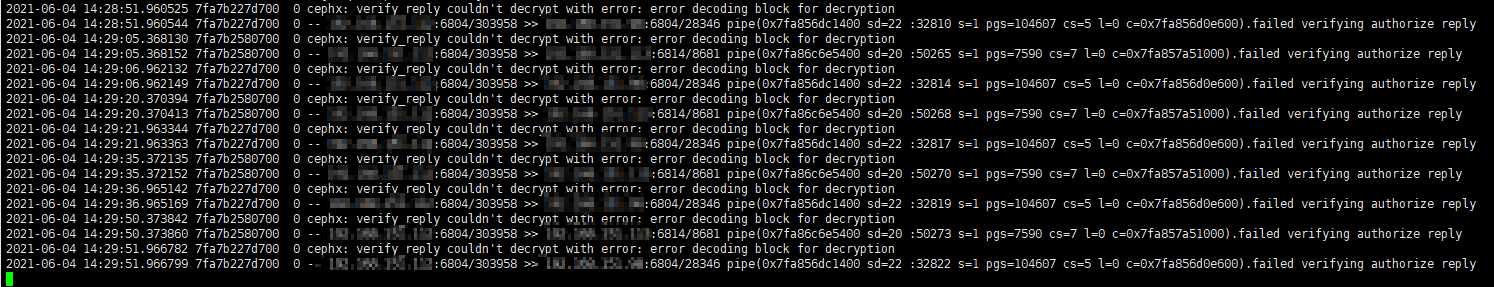


1. 怀疑过osd 和mon mtu 不一致可能会影响osd心跳，社区有个patch对心跳包做了填充（默认补充全0到2000可修改配置）是解决链路上如果存在不允许大包通过的情况由于心跳包很小可以正常收发，但是其它数据包过大可能会被交换机直接丢弃，填充超过1500之后如果由于交换机配置错误能够直接看到osd down不会出现业务io被block问题，因此这个patch解决不了当前环境问题https://github.com/ceph/ceph/pull/16059

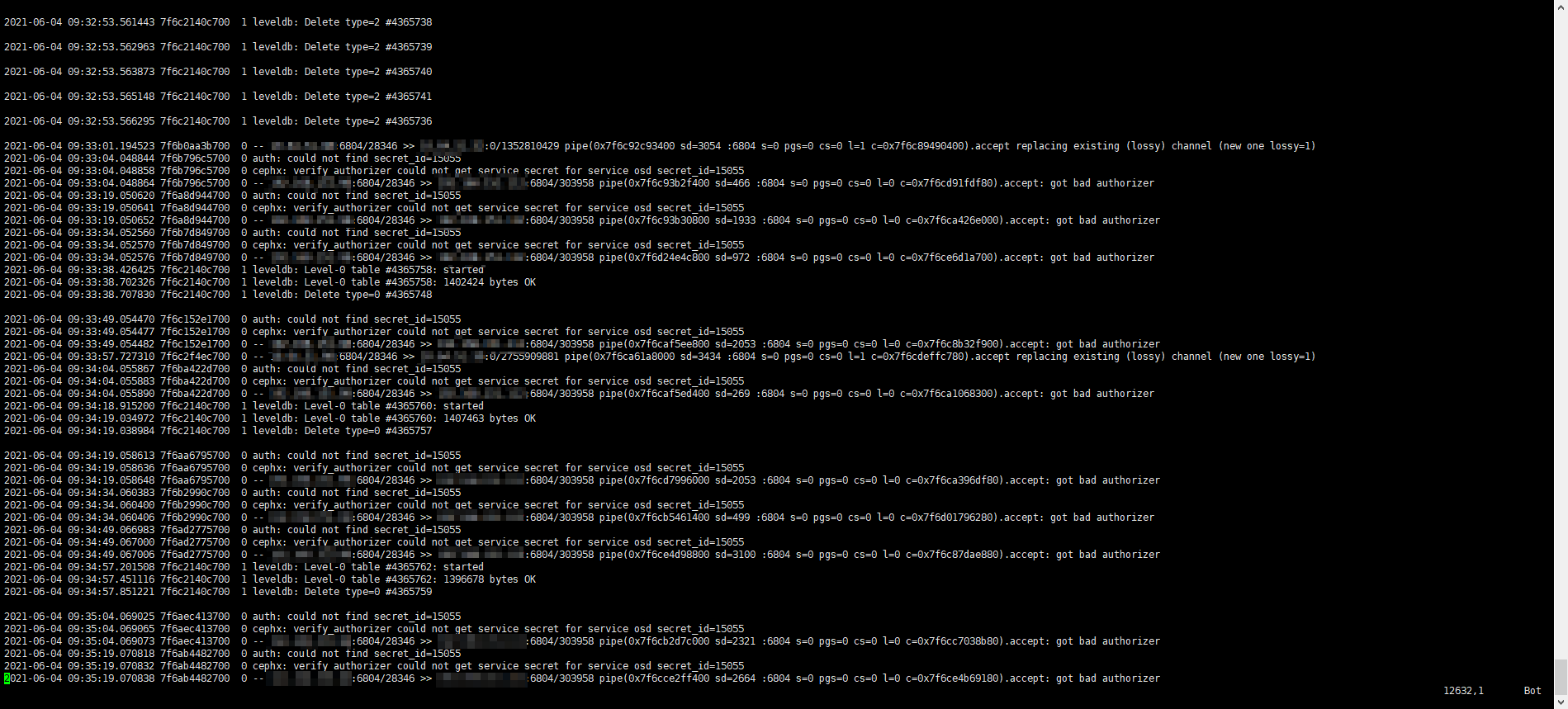
### 第三次问题分析

osd.328有人误操作被重启之后，持续了一晚上backfill+recovery，第二天上午开始出现slow request，发现osd.331再次down，手动out之后slow request消失

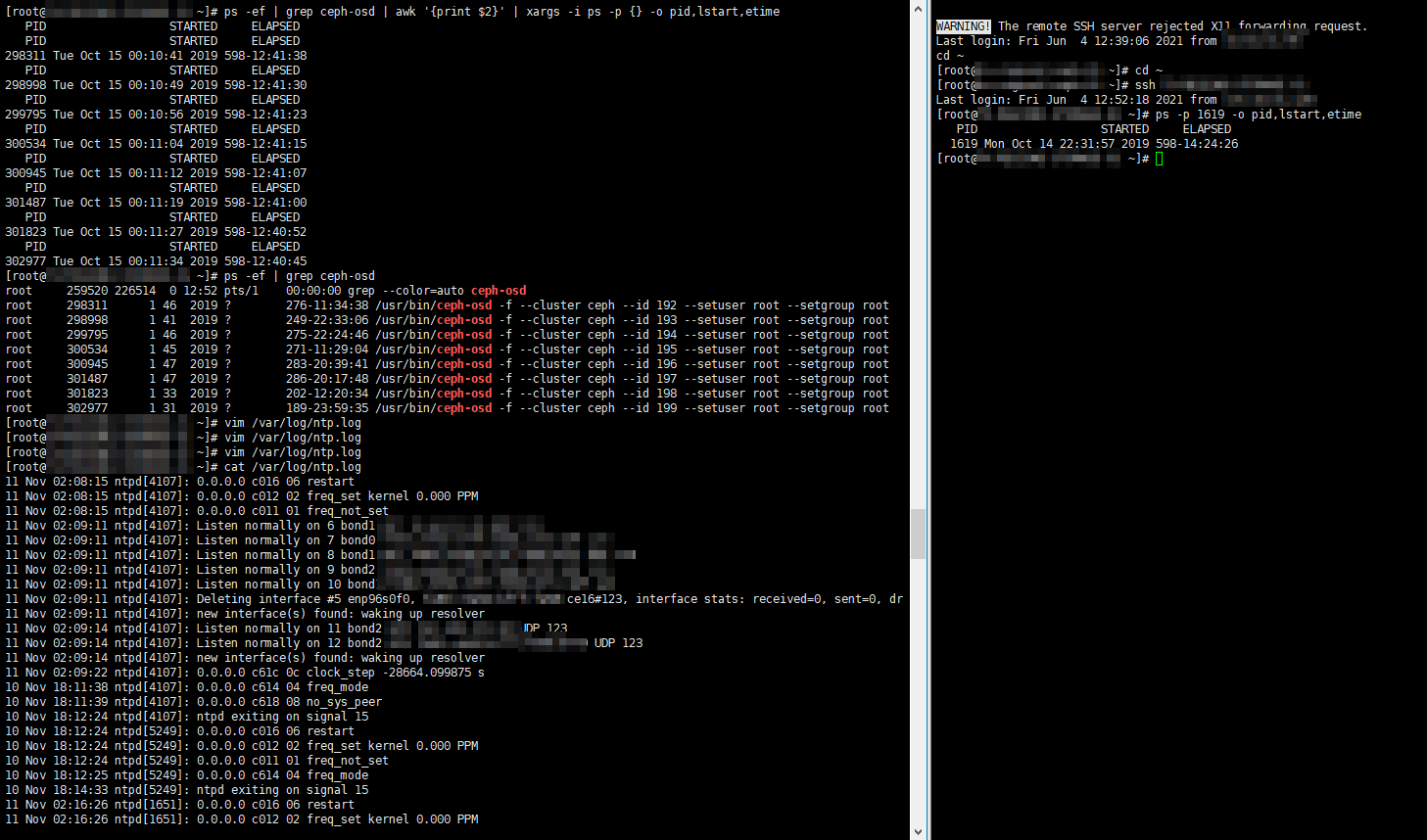
1. osd.331 一直打印和其它2个osd之间认证失败，见下图



查看 x.x.x.x:6814/8681和x.x.x.x:6804/28346这2个osd日志报错类似，怀疑和时间跳变有关系



1. 检查所有节点时间，没有不同步的现象
2. 3个mon启动时间在2019-10-14左右，osd最早的启动时间在2019-10-15，ntp时间看不出来是什么时候执行过同步



设置osd.218日志调高之后看到如下信息

ceph daemon osd.218 config set debug\_auth 10/10

2021-06-07 14:08:39.169514 7f6b2e3a6700 10 cephx: verify\_authorizer decrypted service osd secret\_id=15055

2021-06-07 14:08:39.169531 7f6b2e3a6700 0 auth: could not find secret\_id=15055

2021-06-07 14:08:39.169535 7f6b2e3a6700 10 auth: dump\_rotating:

2021-06-07 14:08:39.169538 7f6b2e3a6700 10 auth: id 15155 AQBbkL1gHwljLxAANA+ZpXVXf77ZPUWvMPtILA== expires 2021-06-07 13:19:55.794997

2021-06-07 14:08:39.169579 7f6b2e3a6700 10 auth: id 15156 AQBrnr1gHWaUERAA3CC+XFtJjqCvem638QThdQ== expires 2021-06-07 14:19:55.794997

2021-06-07 14:08:39.169591 7f6b2e3a6700 10 auth: id 15157 AQB/rL1g/BtcKBAAcXVaMcJBJ1BJPHH1JSjreg== expires 2021-06-07 15:19:59.677108

2021-06-07 14:08:39.169603 7f6b2e3a6700 0 cephx: verify\_authorizer could not get service secret for service osd secret\_id=15055

2021-06-07 14:08:39.169609 7f6b2e3a6700 0 -- x.x.x.x:6804/28346 >> x.x.x.x:6804/303958 pipe(0x7f6ce3329400 sd=267 :6804 s=0 pgs=0 cs=0 l=0 c=0x7f6ca7ab4d00).accept: got bad authorizer

设置osd.33日志级别调高之后看到日志类似

2021-06-07 14:15:27.875104 7fc432c3a700 10 cephx: verify\_authorizer decrypted service osd secret\_id=15055

2021-06-07 14:15:27.875121 7fc432c3a700 0 auth: could not find secret\_id=15055

2021-06-07 14:15:27.875125 7fc432c3a700 10 auth: dump\_rotating:

2021-06-07 14:15:27.875128 7fc432c3a700 10 auth: id 15155 AQBbkL1gHwljLxAANA+ZpXVXf77ZPUWvMPtILA== expires 2021-06-07 13:19:55.794997

2021-06-07 14:15:27.875163 7fc432c3a700 10 auth: id 15156 AQBrnr1gHWaUERAA3CC+XFtJjqCvem638QThdQ== expires 2021-06-07 14:19:55.794997

2021-06-07 14:15:27.875175 7fc432c3a700 10 auth: id 15157 AQB/rL1g/BtcKBAAcXVaMcJBJ1BJPHH1JSjreg== expires 2021-06-07 15:19:59.677108

2021-06-07 14:15:27.875186 7fc432c3a700 0 cephx: verify\_authorizer could not get service secret for service osd secret\_id=15055

2021-06-07 14:15:27.875192 7fc432c3a700 0 -- x.x.x.x:6814/8681 >> x.x.x.x:6804/303958 pipe(0x7fc64c689400 sd=2994 :6814 s=0 pgs=0 cs=0 l=0 c=0x7fc63e3a5380).accept: got bad authorizer

