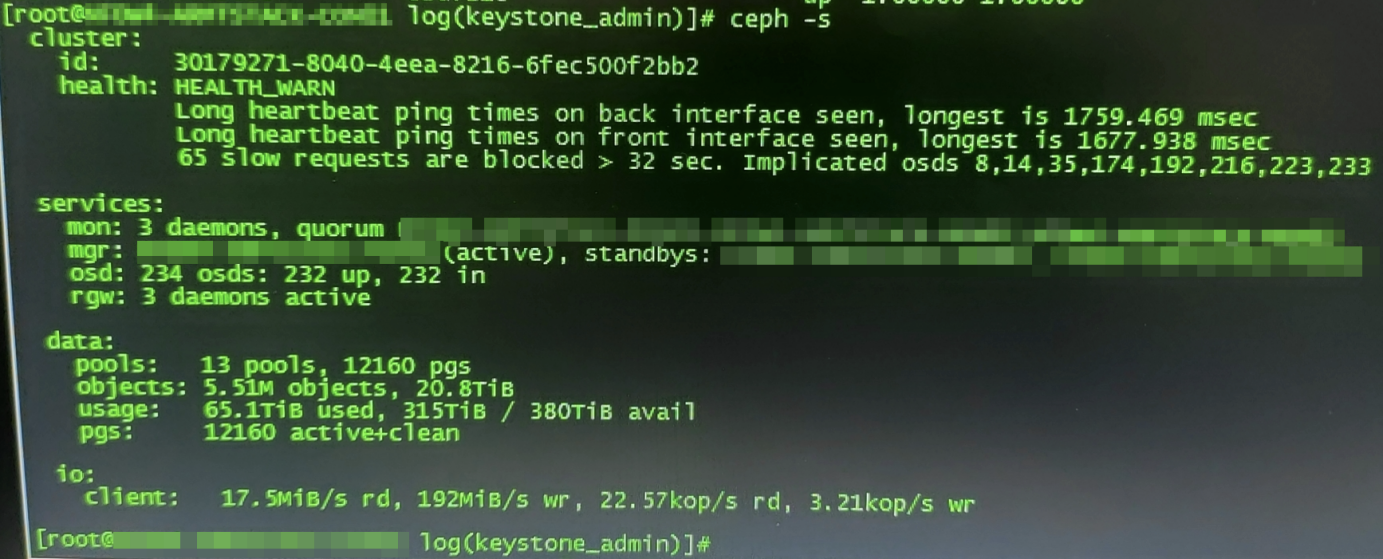
arm环境ceph慢请求问题分析

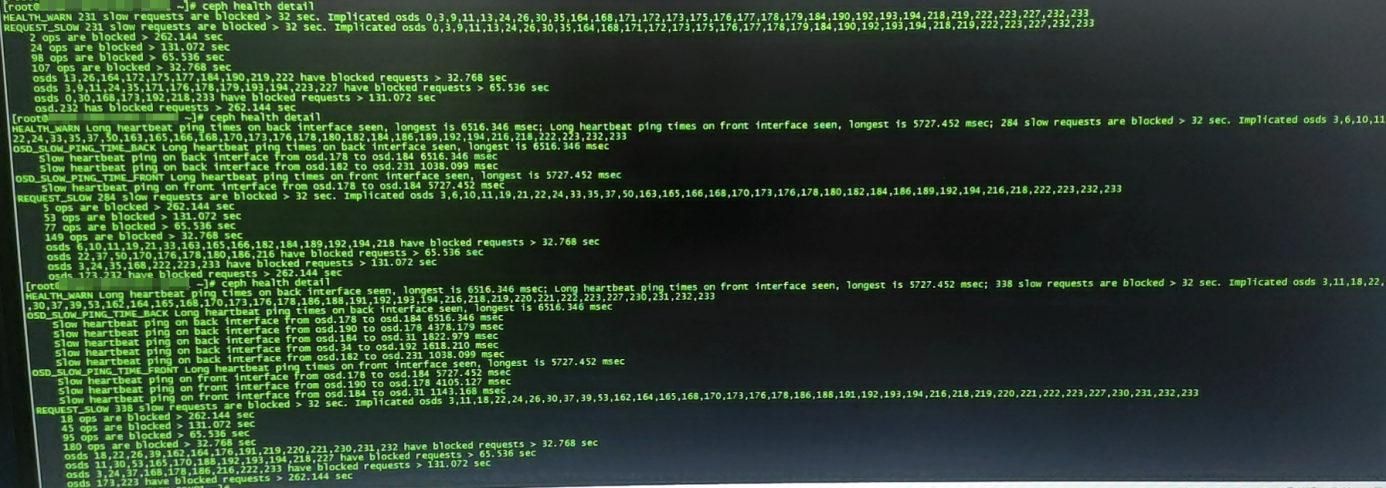
# 问题现象

一线反馈业务io不高，集群持续出现slow request，从0930到1001持续了1整天时间

已知信息

1. 存在部分节点时钟相差2分钟，通过步进方式慢慢调整后已同步
2. mon一直有心跳检测时间过长告警，网络组排查网络正常





# 原因分析

从主mon01日志找到如下错误信息

grep -in "reported failed" MON01-ceph.log

101035:2021-10-01 11:43:07.464745 mon.xxxxxxx mon.0 x.x.x.x:6789/0 758912 : cluster [DBG] osd.189 x.x.x.x:6818/34208 reported failed by osd.172 x.x.x.x:6820/34783

103789:2021-10-01 12:00:31.831697 mon.xxxxxxx mon.0 x.x.x.x:6789/0 759222 : cluster [DBG] osd.192 x.x.x.x:6824/34572 reported failed by osd.28 x.x.x.x:6822/16008

104777:2021-10-01 12:06:57.566970 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 759341 : cluster [DBG] osd.188 x.x.x.x:6816/34095 reported failed by osd.175 x.x.x.x:6826/35123

106207:2021-10-01 12:09:42.805258 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 759394 : cluster [DBG] osd.187 x.x.x.x:6814/33982 reported failed by osd.231 x.x.x.x:6830/30178

110773:2021-10-01 12:52:51.723832 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760054 : cluster [DBG] osd.3 x.x.x.x:6803/80872 reported failed by osd.184 x.x.x.x:6808/33642

110961:2021-10-01 12:53:36.666903 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760074 : cluster [DBG] osd.197 x.x.x.x:6834/35138 reported failed by osd.167 x.x.x.x:6810/34190

111597:2021-10-01 12:55:05.584627 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760111 : cluster [DBG] osd.189 x.x.x.x:6818/34208 reported failed by osd.165 x.x.x.x:6806/33963

111598:2021-10-01 12:55:05.880083 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760112 : cluster [DBG] osd.190 x.x.x.x:6820/34346 reported failed by osd.46 x.x.x.x:6831/96552

112471:2021-10-01 12:56:05.417190 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760137 : cluster [DBG] osd.216 x.x.x.x:6800/28417 reported failed by osd.190 x.x.x.x:6820/34346

112660:2021-10-01 12:56:10.889688 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760141 : cluster [DBG] osd.189 x.x.x.x:6818/34208 reported failed by osd.41 xxxxxxxx:6824/23308

116785:2021-10-01 13:07:33.376493 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760391 : cluster [DBG] osd.184 x.x.x.x:6808/33642 reported failed by osd.178 x.x.x.x:6832/35478

116897:2021-10-01 13:07:39.216323 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760394 : cluster [DBG] osd.178 x.x.x.x:6832/35478 reported failed by osd.190 x.x.x.x:6820/34346

119003:2021-10-01 13:11:10.913973 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760480 : cluster [DBG] osd.182 x.x.x.x:6804/33417 reported failed by osd.13 x.x.x.x:6808/82358

121426:2021-10-01 13:25:54.023468 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760801 : cluster [DBG] osd.233 x.x.x.x:6834/30408 reported failed by osd.197 x.x.x.x:6834/35138

121466:2021-10-01 13:25:58.477214 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760803 : cluster [DBG] osd.184 x.x.x.x:6808/33642 reported failed by osd.230 x.x.x.x:6828/30060

122621:2021-10-01 13:28:13.389350 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760854 : cluster [DBG] osd.185 x.x.x.x:6810/33756 reported failed by osd.32 xxxxxxxx:6814/14762

123396:2021-10-01 13:29:19.203899 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760879 : cluster [DBG] osd.221 x.x.x.x:6810/28989 reported failed by osd.188 x.x.x.x:6816/34095

123510:2021-10-01 13:29:28.665162 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760884 : cluster [DBG] osd.197 x.x.x.x:6834/35138 reported failed by osd.177 x.x.x.x:6830/35355

124254:2021-10-01 13:30:40.044623 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760909 : cluster [DBG] osd.228 x.x.x.x:6824/29829 reported failed by osd.188 x.x.x.x:6816/34095

上述日志说明osd.180 ~ osd.197都有出现过心跳失败问题，这几个osd都在09节点上，查看09节点osd日志，心跳包能够正常收到，但心跳包返回时间超出了deadline时间

grep -in "no reply from" ceph-osd\*

ceph-osd.180.log:1674:2021-10-01 02:39:53.718755 fffd57e5d940 -1 osd.180 25810 heartbeat\_check: no reply from x.x.x.x:6807 osd.165 since back 2021-10-01 02:39:26.950935 front 2021-10-01 02:39:51.899042 (oldest deadline 2021-10-01 02:39:52.845892)

ceph-osd.181.log:2972:2021-10-01 02:36:51.875047 fffeeb13d940 -1 osd.181 25810 heartbeat\_check: no reply from x.x.x.x:6817 osd.224 since back 2021-10-01 02:36:28.369302 front 2021-10-01 02:36:51.302426 (oldest deadline 2021-10-01 02:36:51.656757)

......

ceph-osd.194.log:7815:2021-10-01 11:40:36.879396 fffde5fed940 -1 osd.194 25814 heartbeat\_check: no reply from x.x.x.x:6805 osd.8 since back 2021-10-01 11:40:05.057912 front 2021-10-01 11:40:32.967935

(oldest deadline 2021-10-01 11:40:29.722534)

ceph-osd.197.log:7578:2021-10-01 13:26:30.484815 fffe9171d940 -1 osd.197 25819 heartbeat\_check: no reply from x.x.x.x:6835 osd.233 since back 2021-10-01 13:26:29.050046 front 2021-10-01 13:26:28.886862 (oldest deadline 2021-10-01 13:26:26.747012)

部分osd出现过down,并不是所有osd都down，原因是osd一直在反复重连，而down的条件之一需要满足mon\_osd\_min\_down\_reporters(默认为2)个osd同时上报给mon，从日志中看到每次重连都能成功，因此down的概率很小，down之后进程还在过一会又会up起来，这种反复重连或处于up和down之间来回震荡的osd对读写性能影响较大

grep -in "DOWN" MON01-ceph.log

48192:2021-10-01 04:20:18.638729 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 752816 : cluster [WRN] Health check failed: 1 osds down (OSD\_DOWN)

48229:2021-10-01 04:20:20.721892 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 752820 : cluster [INF] Health check cleared: OSD\_DOWN (was: 1 osds down)

48361:2021-10-01 04:20:20.048420 osd.180 osd.180 x.x.x.x:6800/33189 6181 : cluster [WRN] Monitor daemon marked osd.180 down, but it is still running

48362:2021-10-01 04:20:20.048430 osd.180 osd.180 x.x.x.x:6800/33189 6182 : cluster [DBG] map e25812 wrongly marked me down at e25811

120000:2021-10-01 13:21:16.181527 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760685 : cluster [INF] osd.190 marked itself down

120001:2021-10-01 13:21:16.715928 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760686 : cluster [WRN] Health check failed: 1 osds down (OSD\_DOWN)

120011:2021-10-01 13:21:22.787714 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760693 : cluster [INF] Health check cleared: OSD\_DOWN (was: 1 osds down)

以其中某一个时间点出现的slow request为例

125003 2021-10-01 13:32:43.153212 mon.xxxxxxxx mon.0 x.x.x.x:6789/0 760956 : cluster [WRN] Health check update: 93 slow requests are blocked > 32 sec. Implicated osds 6,14,20,26,34,39,162 ,173,179,182,187,219,224,226,228,229 (REQUEST\_SLOW)

从所有关联osd日志中找对应时间点日志，写head起始时间2021-10-01 13:32:49.261742，结束返回时间2021-10-01 13:36:47.393110，由于osd重连导致写一个object响应延时达到4分钟

7356 2021-10-01 13:32:49.261742 fffd91bfd940 0 log\_channel(cluster) log [WRN] : slow request 30.903496 seconds old, received at 2021-10-01 13:32:18.358171: osd\_op(client.xxxxxxxxxxxxxxx:head [stat,write 4149248~4096] snapc 0=[] ondisk+write+known\_if\_redirected e25819) currently op\_applied

7357 2021-10-01 13:34:00.915348 fffd950ed940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6814/13337 conn(0xaaad0e3ff800 :-1 s=STATE\_OPEN pgs=1135 cs=299 l=0).fault initiating reconnect

7358 2021-10-01 13:35:23.156494 fffd950ed940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6832/1033189 conn(0xaaad03082000 :-1 s=STATE\_OPEN pgs=38971 cs=73 l=0).fault initiating reconnect

7359 2021-10-01 13:35:43.517490 fffd958fd940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6830/1034459 conn(0xaaad3e423000 :-1 s=STATE\_OPEN pgs=62548 cs=299 l=0).fault initiating reconnect

7360 2021-10-01 13:35:43.526008 fffd958fd940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6804/1034208 conn(0xaaad16f4a800 :-1 s=STATE\_OPEN pgs=70458 cs=299 l=0).fault initiating reconnect

7361 2021-10-01 13:35:43.526062 fffd958fd940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6834/1033756 conn(0xaaacfe656000 :-1 s=STATE\_OPEN pgs=85376 cs=299 l=0).fault initiating reconnect

7362 2021-10-01 13:35:43.527021 fffd950ed940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6820/2034095 conn(0xaaad12bd8800 :-1 s=STATE\_OPEN pgs=48740 cs=299 l=0).fault initiating reconnect

7363 2021-10-01 13:35:57.609197 fffd958fd940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6808/20895 conn(0xaaad1a146800 :-1 s=STATE\_OPEN pgs=35067 cs=299 l=0).fault initiating reconnect

7364 2021-10-01 13:35:57.696976 fffd958fd940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6812/21805 conn(0xaaad0a76e000 :-1 s=STATE\_OPEN pgs=39963 cs=299 l=0).fault initiating reconnect

7365 2021-10-01 13:35:57.717841 fffd958fd940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6834/27885 conn(0xaaacfe870000 :-1 s=STATE\_OPEN pgs=34728 cs=299 l=0).fault initiating reconnect

7366 2021-10-01 13:35:58.952893 fffd958fd940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6822/91844 conn(0xaaad049d6000 :-1 s=STATE\_OPEN pgs=33479 cs=299 l=0).fault initiating reconnect

7367 2021-10-01 13:35:59.280305 fffd958fd940 0 -- x.x.x.x:6822/2033417 >> x.x.x.x:6818/15070 conn(0xaaad269ec000 :-1 s=STATE\_OPEN pgs=29097 cs=299 l=0).fault initiating reconnect

7368 2021-10-01 13:36:47.393087 fffd91bfd940 0 log\_channel(cluster) log [WRN] : 1 slow requests, 1 included below; oldest blocked for > 30.299554 secs

7369 2021-10-01 13:36:47.393110 fffd91bfd940 0 log\_channel(cluster) log [WRN] : slow request 30.299554 seconds old, received at 2021-10-01 13:36:17.093478: osd\_op(client.xxxxxxxxxxx:head [stat,write 4149248~4096] snapc 0=[] ondisk+write+known\_if\_redirected e25819) currently sub\_op\_commit\_rec from 101

排查了其它slow request，基本都能找到对应时间点osd重连时间过长的日志，且全部集中在09节点，其它节点没有看到异常日志。现场操作过程：重启完09节点上所有osd之后没有出现心跳丢失和重连现象，slow request消失

# 结论

osd心跳包异常可能原因有

1. 网络故障
2. 时间不同步

检查过09节点网络都正常，且重启该节点上所有osd后问题已解决，间接说明网络没问题。

目前没有证据能证明09节点发生过时间跳变，初步判断根因属于第2种情况