

RADOS - Bug #22413

can't delete object from pool when Ceph out of space

12/12/2017 08:57 PM - Ben England

| | | | |
|------------------------|--------------------|---------------------------|------------|
| Status: | Resolved | Start date: | 12/12/2017 |
| Priority: | High | Due date: | |
| Assignee: | Kefu Chai | % Done: | 0% |
| Category: | Correctness/Safety | Estimated time: | 0.00 hour |
| Target version: | v13.0.0 | Spent time: | 0.00 hour |
| Source: | Community (user) | Reviewed: | |
| Tags: | | Affected Versions: | v12.2.2 |
| Backport: | luminous | ceph-qa-suite: | |
| Regression: | No | Component(RADOS): | librados |
| Severity: | 2 - major | Pull request ID: | |

Description

I ran into a situation where python librados script would hang while trying to delete an object when Ceph storage was full. This is very bad, deleting is the very thing that you need to do in this situation. My workaround, deleting an entire pool, would not be valid in a production system. I haven't reproduced it yet. I'm running Ceph Luminous ceph-common-12.2.2-0.el7.x86_64 . ceph-mgr dashboard gui shows the state we are in below:

```
2017-12-12 20:32:51.120492 [WRN] Health check update: 3 nearfull osd(s) (OSD_NEARFULL)
2017-12-12 20:32:46.269235 [INF] Health check cleared: POOL_BACKFILLFULL (was: 1 pool(s) backfillfull)
2017-12-12 20:32:46.269200 [WRN] Health check failed: 1 pool(s) full (POOL_FULL)
2017-12-12 20:32:46.269084 [ERR] Health check failed: 1 full osd(s) (OSD_FULL)
2017-12-12 20:32:44.199469 [WRN] Health check update: 1 nearfull osd(s) (OSD_NEARFULL)
2017-12-12 20:32:42.100491 [INF] Health check cleared: POOL_NEARFULL (was: 1 pool(s) nearfull)
```

This was done in linode.com, no special ceph.conf params were used, jjust 1 mon, 8 OSDs, 4 clients and a ceph-mgr host.

```
[root@li1832-149 ~]# ceph -s
cluster:
id: 05d304a7-929a-4cc5-9a18-2b4676c8ab64
health: HEALTH_ERR
1 backfillfull osd(s)
1 full osd(s)
3 nearfull osd(s)
1 pool(s) full

services:
  mon: 1 daemons, quorum li469-35
  mgr: li1832-149(active)
  osd: 8 osds: 8 up, 8 in

data:
  pools: 1 pools, 128 pgs
  objects: 9245 objects, 36980 MB
  usage: 109 GB used, 17203 MB / 126 GB avail
  pgs: 128 active+clean

python script was:

#!/usr/bin/env python
import rados, sys

poolnm = sys.argv1
```

```
print('deleting all objs in pool %s' % poolnm)
cluster = rados.Rados(conffile='/etc/ceph/ceph.conf', conf = dict(keyring = '/etc/ceph/ceph.client.admin.keyring'))
cluster.connect()
ioctx = cluster.open_ioctx(poolnm)
iter = ioctx.list_objects()
cnt = 0
while True:
try:
o = iter.next()
#print(o)
ioctx.remove_object(o.key)
cnt += 1
except StopIteration:
break
ioctx.close()
print('%d objects removed from pool %s' % (cnt, poolnm))
```

and stack trace of python script was:

```
(gdb) bt
#0 0x00007f4194432945 in pthread_cond_wait@@@GLIBC_2.3.2 () from /lib64/libpthread.so.0
#1 0x00007f418c0f2131 in librados::IoCtxImpl::operate(object_t const&, ObjectOperation*,
std::chrono::time_point<ceph::time_detail::real_clock, std::chrono::duration<unsigned long, std::ratio<1l, 1000000000l> > >*, int) ()
from /lib64/librados.so.2
#2 0x00007f418c0fcdd8 in librados::IoCtxImpl::remove(object_t const&) () from /lib64/librados.so.2
#3 0x00007f418c0bca28 in rados_remove () from /lib64/librados.so.2
#4 0x00007f418c3e9036 in __pyx_pw_5rados_5ioctx_57remove_object () from /usr/lib64/python2.7/site-packages/rados.so
#5 0x00007f419468e9a3 in PyObject_Call () from /lib64/libpython2.7.so.1.0
#6 0x00007f41947207b7 in PyEval_CallObjectWithKeywords () from /lib64/libpython2.7.so.1.0
#7 0x00007f41946a2da8 in methoddescr_call () from /lib64/libpython2.7.so.1.0
#8 0x00007f419468e9a3 in PyObject_Call () from /lib64/libpython2.7.so.1.0
#9 0x00007f418c40cdc9 in __pyx_pw_5rados_8requires_7wrapper_1validate_func () from
/usr/lib64/python2.7/site-packages/rados.so
#10 0x00007f419468e9a3 in PyObject_Call () from /lib64/libpython2.7.so.1.0
#11 0x00007f41947230f6 in PyEval_EvalFrameEx () from /lib64/libpython2.7.so.1.0
#12 0x00007f4194729efd in PyEval_EvalCodeEx () from /lib64/libpython2.7.so.1.0
#13 0x00007f419472a002 in PyEval_EvalCode () from /lib64/libpython2.7.so.1.0
#14 0x00007f419474343f in run_mod () from /lib64/libpython2.7.so.1.0
#15 0x00007f41947445fe in PyRun_FileExFlags () from /lib64/libpython2.7.so.1.0
#16 0x00007f419471dfec in builtin_execfile () from /lib64/libpython2.7.so.1.0
#17 0x00007f4194727bb0 in PyEval_EvalFrameEx () from /lib64/libpython2.7.so.1.0
#18 0x00007f4194729efd in PyEval_EvalCodeEx () from /lib64/libpython2.7.so.1.0
#19 0x00007f419472a002 in PyEval_EvalCode () from /lib64/libpython2.7.so.1.0
#20 0x00007f419474343f in run_mod () from /lib64/libpython2.7.so.1.0
#21 0x00007f41947442a5 in PyRun_StringFlags () from /lib64/libpython2.7.so.1.0
#22 0x00007f41947234f5 in PyEval_EvalFrameEx () from /lib64/libpython2.7.so.1.0
#23 0x00007f4194729efd in PyEval_EvalCodeEx () from /lib64/libpython2.7.so.1.0
#24 0x00007f41947273fc in PyEval_EvalFrameEx () from /lib64/libpython2.7.so.1.0
#25 0x00007f419472757d in PyEval_EvalFrameEx () from /lib64/libpython2.7.so.1.0
#26 0x00007f419472757d in PyEval_EvalFrameEx () from /lib64/libpython2.7.so.1.0
#27 0x00007f4194729efd in PyEval_EvalCodeEx () from /lib64/libpython2.7.so.1.0
#28 0x00007f419472a002 in PyEval_EvalCode () from /lib64/libpython2.7.so.1.0
#29 0x00007f41947262d3 in PyEval_EvalFrameEx () from /lib64/libpython2.7.so.1.0
#30 0x00007f4194729efd in PyEval_EvalCodeEx () from /lib64/libpython2.7.so.1.0
#31 0x00007f41947273fc in PyEval_EvalFrameEx () from /lib64/libpython2.7.so.1.0
#32 0x00007f4194729efd in PyEval_EvalCodeEx () from /lib64/libpython2.7.so.1.0
#33 0x00007f41946b3858 in function_call () from /lib64/libpython2.7.so.1.0
#34 0x00007f419468e9a3 in PyObject_Call () from /lib64/libpython2.7.so.1.0
#35 0x00007f4194755c30 in RunModule () from /lib64/libpython2.7.so.1.0
#36 0x00007f4194756414 in Py_Main () from /lib64/libpython2.7.so.1.0
#37 0x00007f419397cc05 in __libc_start_main () from /lib64/libc.so.6
#38 0x00000000040071e in _start ()
```

Related issues:

Copied to RADOS - Backport #23114: luminous: can't delete object from pool wh...

Resolved

History

#1 - 12/12/2017 10:35 PM - Ben England

forgot to mention I get errors like this when it fills up:

```
192.168.203.54: 2017-12-12 22:33:58.369563 7f2f1cb7ee40 0 client.20619.objecter FULL, paused modify 0x561c26fec240 tid 0
192.168.202.100: 2017-12-12 22:33:58.385279 7fa85329ce40 0 client.20621.objecter FULL, paused modify 0x55dd8986b890 tid 0
...
```

so the paused part makes sense, but the inability to delete objects does not.

#2 - 12/13/2017 12:18 AM - Josh Durgin

You can get around this by using `rados_write_op_operate` with the `'LIBRADOS_OPERATION_FULL_FORCE'` flag (128), like the `rados cli`, but I agree this should be the default for plain old `remove()`.

#3 - 01/03/2018 09:24 PM - Josh Durgin

- Priority changed from Normal to High

#4 - 02/22/2018 11:10 AM - Kefu Chai

- Status changed from New to Need Review

- Assignee set to Kefu Chai

- Component(RADOS) librados added

<https://github.com/ceph/ceph/pull/20534>

#5 - 02/24/2018 04:20 AM - Kefu Chai

- Status changed from Need Review to Pending Backport

- Backport set to luminous

#6 - 02/24/2018 11:22 AM - Nathan Cutler

- Copied to Backport #23114: luminous: can't delete object from pool when Ceph out of space added

#7 - 03/29/2018 03:33 PM - Nathan Cutler

- Status changed from Pending Backport to Resolved