RADOS - Bug #38345

mon: segv in MonOpRequest::~MonOpRequest OpHistory::cleanup

02/15/2019 06:31 PM - Sage Weil

Status:	Verified	**Start date:**	02/15/2019
Priority:	High	**Due date:**	
Assignee:		**% Done:**	0%
Category:	Correctness/Safety	**Estimated time:**	0.00 hour
Target version:		**Spent time:**	0.00 hour
Source:		**Affected Versions:**	
Tags:		**ceph-qa-suite:**	
Backport:		**Component(RADOS):**	Monitor
Regression:	No	**Pull request ID:**	
Severity:	3 - minor	**Crash signature:**	
Reviewed:			

Description

```
2019-02-15 09:48:45.381 7f30cf9fd700  1 -- v1:172.21.15.148:6789/0 <== mon.1 v1:172.21.15.102:6789
/0 2 ==== election(8ec471df-d03c-4d8b-9fa6-53fd073b70cc propose 11) v7 ==== 435+0+0 (539273810 0 0
) 0x3cd5e40 con 0x437cd80
...
2019-02-15 09:48:45.389 7f30cf9fd700 20 -- v1:172.21.15.148:6789/0 done calling dispatch on 0x3cd5
e40
...
2019-02-15 09:49:57.145 7f30d3204700 -1 *** Caught signal (Segmentation fault) **
 in thread 7f30d3204700 thread_name:OpHistorySvc

 ceph version 14.0.1-3738-gca5ec13 (ca5ec139e4b98becd397a459018a659a332bc291) nautilus (dev)
 1: (()+0x11390) [0x7f30db1ad390]
 2: (RefCountedObject::put() const+0x44) [0x6d6014]
 3: (MonOpRequest::~MonOpRequest()+0x43) [0x6d7663]
 4: (std::_Rb_tree<std::pair<double, boost::intrusive_ptr<TrackedOp> >, std::pair<double, boost::i
ntrusive_ptr<TrackedOp> >, std::_Identity<std::pair<double, boost::intrusive_ptr<TrackedOp> > >, s
td::less<std::pair<double, boost::intrusive_ptr<TrackedOp> > >, std::allocator<std::pair<double, b
oost::intrusive_ptr<TrackedOp> > > >::_M_erase_aux(std::_Rb_tree_const_iterator<std::pair<double,
boost::intrusive_ptr<TrackedOp> > >)+0xa6) [0x7f30dc4421d6]
 5: (OpHistory::cleanup(utime_t)+0x31d) [0x7f30dc43c66d]
 6: (OpHistory::_insert_delayed(utime_t const&, boost::intrusive_ptr<TrackedOp>)+0x276) [0x7f30dc4
3d8e6]
 7: (OpHistoryServiceThread::entry()+0xf9) [0x7f30dc43df49]
 8: (()+0x76ba) [0x7f30db1a36ba]
 9: (clone()+0x6d) [0x7f30da9cc41d]


(gdb) bt
#0  0x00007f30db1ad269 in raise (sig=sig@entry=11) at ../sysdeps/unix/sysv/linux/pt-raise.c:35
#1  0x00000000008b7473 in reraise_fatal (signum=11) at /build/ceph-14.0.1-3738-gca5ec13/src/global
/signal_handler.cc:81
#2  handle_fatal_signal (signum=11) at /build/ceph-14.0.1-3738-gca5ec13/src/global/signal_handler.
cc:298
#3  <signal handler called>
#4  RefCountedObject::put (this=0x428eb40) at /build/ceph-14.0.1-3738-gca5ec13/src/common/RefCount
edObj.h:58
#5  0x00000000006d7663 in intrusive_ptr_release (p=<optimized out>) at /build/ceph-14.0.1-3738-gca
5ec13/src/common/RefCountedObj.h:174
#6  boost::intrusive_ptr<RefCountedObject>::~intrusive_ptr (this=0x2e88740, __in_chrg=<optimized o
ut>) at /build/ceph-14.0.1-3738-gca5ec13/obj-x86_64-linux-gnu/boost/include/boost/smart_ptr/intrus
ive_ptr.hpp:98
```

```
#7  MonOpRequest::~MonOpRequest (this=0x2e88670, __in_chrg=<optimized out>) at /build/ceph-14.0.1-
3738-gca5ec13/src/mon/MonOpRequest.h:127
#8  MonOpRequest::~MonOpRequest (this=0x2e88670, __in_chrg=<optimized out>) at /build/ceph-14.0.1-
3738-gca5ec13/src/mon/MonOpRequest.h:129
#9  0x00007f30dc4421d6 in std::_Rb_tree<std::pair<double, boost::intrusive_ptr<TrackedOp> >, std::
pair<double, boost::intrusive_ptr<TrackedOp> >, std::_Identity<std::pair<double, boost::intrusive_
ptr<TrackedOp> > >, std::less<std::pair<double, boost::intrusive_ptr<TrackedOp> > >, std::allocato
r<std::pair<double, boost::intrusive_ptr<TrackedOp> > > >::_M_erase_aux(std::_Rb_tree_const_iterat
or<std::pair<double, boost::intrusive_ptr<TrackedOp> > >) () from /usr/lib/ceph/libceph-common.so.
1
#10 0x00007f30dc43c66d in OpHistory::cleanup(utime_t) () from /usr/lib/ceph/libceph-common.so.1
#11 0x00007f30dc43d8e6 in OpHistory::_insert_delayed(utime_t const&, boost::intrusive_ptr<TrackedO
p>) () from /usr/lib/ceph/libceph-common.so.1
#12 0x00007f30dc43df49 in OpHistoryServiceThread::entry() () from /usr/lib/ceph/libceph-common.so.
1
#13 0x00007f30db1a36ba in start_thread (arg=0x7f30d3204700) at pthread_create.c:333
#14 0x00007f30da9cc41d in clone () at ../sysdeps/unix/sysv/linux/x86_64/clone.S:109
```

/a/sage-2019-02-15_04:19:44-upgrade:mimic-x-wip-mimic-upgrade-distro-basic-smithi/3592782

Related issues:

Related to RADOS - Bug #38306: ceph-mon: "*** Caught signal (Segmentation fau...	**New**	**02/13/2019**
Related to RADOS - Bug #24664: osd: crash in OpTracker::unregister_inflight_o...	**Pending Backport**	**06/26/2018**

History

#1 - 02/15/2019 10:06 PM - Sage Weil

- Subject changed from mon: segv in to mon: segv in MonOpRequest::~MonOpRequest OpHistory::cleanup

- Priority changed from High to Urgent

/a/sage-2019-02-15_20:30:53-upgrade:mimic-x-wip-mimic-upgrade-distro-basic-smithi/3595562

#2 - 02/20/2019 10:43 PM - Greg Farnum

*- Related to Bug #38306: ceph-mon: "*** Caught signal (Segmentation fault) **" in upgrade:luminous-x-mimic added*

#3 - 03/11/2019 02:09 PM - Joao Eduardo Luis

- Category set to Correctness/Safety

- Assignee set to Joao Eduardo Luis

- Component(RADOS) Monitor added

#4 - 03/13/2019 10:55 PM - Sage Weil

```
Missing separate debuginfos, use: debuginfo-install bzip2-libs-1.0.6-13.el7.x86_64 elfutils-libelf-0.170-4.el7
.x86_64 elfutils-libs-0.170-4.el7.x86_64 fuse-libs-2.9.2-10.el7.x86_64 glibc-2.17-222.el7.x86_64 gperftools-li
bs-2.6.1-1.el7.x86_64 leveldb-1.12.0-11.el7.x86_64 libaio-0.3.109-13.el7.x86_64 libattr-2.4.46-13.el7.x86_64 l
ibblkid-2.23.2-52.el7.x86_64 libcap-2.22-9.el7.x86_64 libgcc-4.8.5-36.el7.x86_64 libibverbs-15-7.el7_5.x86_64
libnl3-3.2.28-4.el7.x86_64 librdmacm-15-7.el7_5.x86_64 libstdc++-4.8.5-36.el7.x86_64 libuuid-2.23.2-52.el7.x86
_64 lttng-ust-2.4.1-4.el7.x86_64 lz4-1.7.5-2.el7.x86_64 nspr-4.19.0-1.el7_5.x86_64 nss-3.36.0-5.el7_5.x86_64 n
ss-softokn-3.36.0-5.el7_5.x86_64 nss-softokn-freebl-3.36.0-5.el7_5.x86_64 nss-util-3.36.0-1.el7_5.x86_64 opens
sl-libs-1.0.2k-12.el7.x86_64 snappy-1.1.0-3.el7.x86_64 sqlite-3.7.17-8.el7.x86_64 systemd-libs-219-57.el7.x86_
64 userspace-rcu-0.7.16-1.el7.x86_64 xz-libs-5.2.2-1.el7.x86_64 zlib-1.2.7-17.el7.x86_64
(gdb) bt
#0  0x00007fa5323ae59b in raise () from /lib64/libpthread.so.0
#1  0x000056198d2e04e5 in reraise_fatal (signum=11) at /usr/src/debug/ceph-14.1.0-638-ged9e549/src/global/sign
al_handler.cc:81
#2  handle_fatal_signal (signum=11) at /usr/src/debug/ceph-14.1.0-638-ged9e549/src/global/signal_handler.cc:29
8
#3  <signal handler called>
#4  0x00007fa5323a8c80 in pthread_mutex_lock () from /lib64/libpthread.so.0
#5  0x00007fa535ade441 in __gthread_mutex_lock (__mutex=0x10000000100038) at /opt/rh/devtoolset-7/root/usr/inc
lude/c++/7/x86_64-redhat-linux/bits/gthr-default.h:748
#6  lock (this=0x10000000100038) at /opt/rh/devtoolset-7/root/usr/include/c++/7/bits/std_mutex.h:103
#7  lock (this=0x7fa52b007410) at /opt/rh/devtoolset-7/root/usr/include/c++/7/bits/std_mutex.h:267
```

```
#8  unique_lock (__m=..., this=0x7fa52b007410) at /opt/rh/devtoolset-7/root/usr/include/c++/7/bits/std_mutex.h
:197
#9  ceph::logging::Log::submit_entry(ceph::logging::Entry&&) (this=0x10000000100000, e=e@entry=<unknown type i
n /usr/lib/debug/usr/lib64/ceph/libceph-common.so.0.debug, CU 0x4530b47, DIE 0x459b6ad>) at /usr/src/debug/cep
h-14.1.0-638-ged9e549/src/log/Log.cc:177
#10 0x000056198d105fa2 in RefCountedObject::put (this=0x56198f7b7880) at /usr/src/debug/ceph-14.1.0-638-ged9e5
49/src/common/RefCountedObj.h:60
#11 0x000056198d107399 in intrusive_ptr_release (p=<optimized out>) at /usr/src/debug/ceph-14.1.0-638-ged9e549
/src/common/RefCountedObj.h:174
#12 ~intrusive_ptr (this=0x561990dc7ec8, __in_chrg=<optimized out>) at /usr/src/debug/ceph-14.1.0-638-ged9e549
/build/boost/include/boost/smart_ptr/intrusive_ptr.hpp:98
#13 ~MonOpRequest (this=0x561990dc7e10, __in_chrg=<optimized out>) at /usr/src/debug/ceph-14.1.0-638-ged9e549/
src/mon/MonOpRequest.h:127
#14 MonOpRequest::~MonOpRequest (this=0x561990dc7e10, __in_chrg=<optimized out>) at /usr/src/debug/ceph-14.1.0
-638-ged9e549/src/mon/MonOpRequest.h:129
#15 0x00007fa535841ab6 in put (this=<optimized out>) at /usr/src/debug/ceph-14.1.0-638-ged9e549/src/common/Tra
ckedOp.h:308
#16 intrusive_ptr_release (o=<optimized out>) at /usr/src/debug/ceph-14.1.0-638-ged9e549/src/common/TrackedOp.
h:394
#17 ~intrusive_ptr (this=<optimized out>, __in_chrg=<optimized out>) at /usr/src/debug/ceph-14.1.0-638-ged9e54
9/build/boost/include/boost/smart_ptr/intrusive_ptr.hpp:98
#18 ~pair (this=<optimized out>, __in_chrg=<optimized out>) at /opt/rh/devtoolset-7/root/usr/include/c++/7/bit
s/stl_pair.h:198
#19 destroy<std::pair<double, boost::intrusive_ptr<TrackedOp> > > (this=<optimized out>, __p=<optimized out>)
 at /opt/rh/devtoolset-7/root/usr/include/c++/7/ext/new_allocator.h:140
#20 destroy<std::pair<double, boost::intrusive_ptr<TrackedOp> > > (__a=..., __p=<optimized out>) at /opt/rh/de
vtoolset-7/root/usr/include/c++/7/bits/alloc_traits.h:487
#21 _M_destroy_node (this=0x561990bf3268, __p=0x561990b28ab0) at /opt/rh/devtoolset-7/root/usr/include/c++/7/b
its/stl_tree.h:650
#22 _M_drop_node (this=0x561990bf3268, __p=0x561990b28ab0) at /opt/rh/devtoolset-7/root/usr/include/c++/7/bits
/stl_tree.h:658
#23 std::_Rb_tree<std::pair<double, boost::intrusive_ptr<TrackedOp> >, std::pair<double, boost::intrusive_ptr<
TrackedOp> >, std::_Identity<std::pair<double, boost::intrusive_ptr<TrackedOp> > >, std::less<std::pair<double
, boost::intrusive_ptr<TrackedOp> > >, std::allocator<std::pair<double, boost::intrusive_ptr<TrackedOp> > > >:
:_M_erase_aux (this=this@entry=0x561990bf3268, __position=...) at /opt/rh/devtoolset-7/root/usr/include/c++/7/
bits/stl_tree.h:2477
#24 0x00007fa53583b944 in erase (__position=..., this=0x561990bf3268) at /opt/rh/devtoolset-7/root/usr/include
/c++/7/bits/stl_tree.h:1113
#25 erase (__position=..., this=0x561990bf3268) at /opt/rh/devtoolset-7/root/usr/include/c++/7/bits/stl_set.h:
645
#26 OpHistory::cleanup (this=this@entry=0x561990bf3238, now=...) at /usr/src/debug/ceph-14.1.0-638-ged9e549/sr
c/common/TrackedOp.cc:101
#27 0x00007fa53583cd9d in OpHistory::_insert_delayed (this=0x561990bf3238, now=..., op=...) at /usr/src/debug/
ceph-14.1.0-638-ged9e549/src/common/TrackedOp.cc:83
#28 0x00007fa53583d378 in OpHistoryServiceThread::entry (this=0x561990bf3308) at /usr/src/debug/ceph-14.1.0-63
8-ged9e549/src/common/TrackedOp.cc:54
#29 0x00007fa5323a6e25 in start_thread () from /lib64/libpthread.so.0
#30 0x00007fa53126fbad in clone () from /lib64/libc.so.6
```

/a/sage-2019-02-03_18:58:17-rados-wip-sage2-testing-2019-02-03-1047-distro-basic-smithi/3545666
core and log in top level dir

#5 - 03/13/2019 11:03 PM - Sage Weil

```
(gdb) do
#10 0x000056198d105fa2 in RefCountedObject::put (this=0x56198f7b7880) at /usr/src/debug/ceph-14.1.0-638-ged9e5
49/src/common/RefCountedObj.h:60
60                                          << dendl;
(gdb) list
55          CephContext *local_cct = cct;
56          int v = --nref;
57          if (local_cct)
58            lsubdout(local_cct, refs, 1) << "RefCountedObject::put " << this << " "
59                                         << (v + 1) << " -> " << v
60                                         << dendl;
61          if (v == 0) {
62            ANNOTATE_HAPPENS_AFTER(&nref);
63            ANNOTATE_HAPPENS_BEFORE_FORGET_ALL(&nref);
64            delete this;
(gdb) p cct
$16 = (CephContext *) 0x56198fae0000
(gdb) p local_cct
$17 = (CephContext *) 0x5619a1e93540
(gdb) p nref
$18 = {
  <std::__atomic_base<unsigned long>> = {
    static _S_alignment = 8,
    _M_i = 4
  },
  members of std::atomic<unsigned long>:
  static is_always_lock_free = true
}
(gdb) p cct->_log
$19 = (ceph::logging::Log *) 0x56198f7b8780
(gdb) p local_cct->_log
$20 = (ceph::logging::Log *) 0x10000000100000
```

something mangled cct before the old object was destroyed?

(this->cct and nref are potentially meaningless since we may have freed the object and had someone else allocate underneath us?)

#6 - 03/14/2019 04:17 PM - Sage Weil

Forgot mention, the op appears to be an MForward.

#7 - 04/24/2019 06:48 PM - Greg Farnum

- Related to Bug #24664: osd: crash in OpTracker::unregister_inflight_op via OSD::get_health_metrics added

#8 - 04/24/2019 06:49 PM - Greg Farnum

We're also seeing Bus Errors instead of segfaults in the OpHistory cleanup at #24664 so these may be related...

#9 - 05/04/2019 10:43 PM - Kefu Chai

```
ceph version 13.2.5-268-g0d5c736 (0d5c736349288b66b3b12c1df49568d1ed29f90d) mimic (stable)
1: (()+0xf5d0) [0x7f11bd74d5d0]
2: (intrusive_ptr_release(RefCountedObject const*)+0x19f) [0x7f11be53ec2f]
3: (MonOpRequest::~MonOpRequest()+0x49) [0x557d810f8ba9]
4: (OpHistoryServiceThread::entry()+0x227) [0x7f11be41c0d7]
5: (()+0x7dd5) [0x7f11bd745dd5]
6: (clone()+0x6d) [0x7f11ba062ead]
```

/kchai-2019-05-04_11:38:20-rados-wip-ceph-release-distro-basic-smithi/3927368/

#10 - 05/10/2019 05:32 PM - Neha Ojha

/a/nojha-2019-05-10_00:33:57-upgrade-wip-parial-recovery-2019-05-09-distro-basic-smithi/3943156/

#11 - 08/20/2019 08:26 PM - Greg Farnum

- Assignee deleted (Joao Eduardo Luis)

- Priority changed from Urgent to High

#12 - 10/06/2019 02:05 PM - Sage Weil

```
2019-10-06T07:00:58.981 INFO:tasks.ceph.mon.c.smithi112.stderr:*** Caught signal (Segmentation fault) **
2019-10-06T07:00:58.981 INFO:tasks.ceph.mon.c.smithi112.stderr: in thread 7f5f2325b700 thread_name:OpHistorySv
c
2019-10-06T07:00:58.982 INFO:tasks.ceph.mon.c.smithi112.stderr: ceph version 13.2.6-466-g1720691 (172069129813
87416f4e0759804d403ab5e4cf52) mimic (stable)
2019-10-06T07:00:58.982 INFO:tasks.ceph.mon.c.smithi112.stderr: 1: (()+0xf630) [0x7f5f2c8fc630]
2019-10-06T07:00:58.983 INFO:tasks.ceph.mon.c.smithi112.stderr: 2: (intrusive_ptr_release(RefCountedObject con
st*)+0x19f) [0x7f5f2d6f1e6f]
2019-10-06T07:00:58.983 INFO:tasks.ceph.mon.c.smithi112.stderr: 3: (MonOpRequest::~MonOpRequest()+0x49) [0x564
4001b93a9]
2019-10-06T07:00:58.983 INFO:tasks.ceph.mon.c.smithi112.stderr: 4: (OpHistoryServiceThread::entry()+0x227) [0x
7f5f2d5ce837]
2019-10-06T07:00:58.983 INFO:tasks.ceph.mon.c.smithi112.stderr: 5: (()+0x7ea5) [0x7f5f2c8f4ea5]
2019-10-06T07:00:58.984 INFO:tasks.ceph.mon.c.smithi112.stderr: 6: (clone()+0x6d) [0x7f5f294158cd]
2019-10-06T07:00:58.984 INFO:tasks.ceph.mon.c.smithi112.stderr:2019-10-06 07:00:58.989 7f5f2325b700 -1 *** Cau
ght signal (Segmentation fault) **
```

/a/sage-2019-10-05_20:13:12-rados:upgrade-wip-sage2-testing-2019-10-04-1006-distro-basic-smithi/4362554
(mimic -> master(octopus))