

Ceph - Bug #6003

journal Unable to read past sequence 406 ...

08/15/2013 09:14 PM - Sage Weil

Status:	Resolved	% Done:	0%
Priority:	Urgent	Spent time:	0.00 hour
Assignee:	Samuel Just		
Category:	OSD		
Target version:			
Source:	Q/A	Affected Versions:	
Tags:		ceph-qa-suite:	
Backport:	dumpling, firefly, giant	Pull request ID:	
Regression:	No	Crash signature (v1):	
Severity:	2 - major	Crash signature (v2):	
Reviewed:			

Description

- fix : [bae1f3eaa09c4747b8bfc6fb5dc673aa6989b695](#)

```
2013-08-15 12:44:50.877768 7f51d4230780 -1 journal Unable to read past sequence 406 but header indicates the journal has committed up through 416, journal is corrupt
```

```
2013-08-15 12:44:50.901410 7f51d4230780 -1 os/FileJournal.cc: In function 'bool FileJournal::read_entry(ceph::bufferlist&, uint64_t&, bool*)' thread 7f51d4230780 time 2013-08-15 12:44:50.877826 os/FileJournal.cc: 1697: FAILED assert(0)
```

```
ceph version 0.67-125-g4422f21 (4422f21a6586467a63ce6841552d0f60aa849cf1)
```

```
1: (FileJournal::read_entry(ceph::buffer::list&, unsigned long&, bool*)+0x7d6) [0x77e966]
```

```
2: (FileJournal::open(unsigned long)+0x733) [0x77bf73]
```

```
3: (JournalingObjectStore::journal_replay(unsigned long)+0x271) [0x7cad61]
```

```
4: (FileStore::mount()+0x34f3) [0x7a1503]
```

```
5: (OSD::do_convertfs(ObjectStore*)+0x1a) [0x66dcba]
```

```
6: (OSD::convertfs(std::string const&, std::string const&)+0x47) [0x66e737]
```

```
7: (main()+0x1c13) [0x5d0e13]
```

```
8: (__libc_start_main()+0xed) [0x7f51d1ccc76d]
```

```
9: ceph-osd() [0x5d479d]
```

```
NOTE: a copy of the executable, or `objdump -rDS <executable>` is needed to interpret this.
```

job was

```
ubuntu@teuthology:/a/teuthology-2013-08-15_01:00:05-rados-next-testing-basic-plana/107661$ cat orig.config.yaml
```

```
kernel:
```

```
  kdb: true
```

```
  sha1: 68d04c9a036bea12d76fdd51e926aef771e8e041
```

```
machine_type: plana
```

```
nuke-on-error: true
```

```
os_type: ubuntu
```

```
overrides:
```

```
  admin_socket:
```

```
    branch: next
```

```
  ceph:
```

```
    conf:
```

```
      global:
```

```
        ms inject socket failures: 5000
```

```
      mon:
```

```
        debug mon: 20
```

```

    debug ms: 1
    debug paxos: 20
fs: xfs
log-whitelist:
- slow request
shal: 4422f21a6586467a63ce6841552d0f60aa849cf1
ceph-deploy:
  branch:
    dev: next
  conf:
    client:
      log file: /var/log/ceph/ceph-$name.$pid.log
    mon:
      debug mon: 1
      debug ms: 20
      debug paxos: 20
install:
  ceph:
    shal: 4422f21a6586467a63ce6841552d0f60aa849cf1
s3tests:
  branch: next
workunit:
  shal: 4422f21a6586467a63ce6841552d0f60aa849cf1
roles:
- - mon.a
- - mon.c
- - osd.0
- - osd.1
- - osd.2
- - mon.b
- - mds.a
- - osd.3
- - osd.4
- - osd.5
- - client.0
tasks:
- chef: null
- clock.check: null
- install: null
- ceph:
  log-whitelist:
  - wrongly marked me down
  - objects unfound and apparently lost
- thrashosds:
  chance_pgnum_grow: 2
  chance_pgnum_fix: 1
  timeout: 1200
- radosbench:
  clients:
  - client.0
  time: 1800
teuthology_branch: next

```

Related issues:

Related to Ceph - Bug #9073: OSD with device/partition journals down after fr...	Resolved	08/11/2014
Related to Ceph - Bug #10632: ceph-objectstore-tool: crash in FileJournal.cc	Duplicate	01/24/2015
Duplicated by Ceph - Bug #9682: "os/FileJournal.cc: 1677: FAILED assert(0)" i...	Duplicate	10/07/2014
Duplicated by Ceph - Bug #10609: "os/FileJournal.cc: 1693: FAILED assert(0)" ...	Duplicate	01/22/2015
Duplicated by Ceph - Bug #10629: osd/PG.cc: 2775: FAILED assert(values.size())...	Duplicate	01/24/2015

Associated revisions

Revision **cd30e5fb** - 11/01/2013 11:12 PM - Samuel Just

common/buffer.cc: fix rebuild_page_aligned typo

Introduced: 66a9fbe2c7ba59b7cd034c17865adce3432cd2cb
Fixes: #6003
Signed-off-by: Samuel Just <sam.just@inktank.com>

Revision d6a53ea0 - 09/30/2014 04:43 PM - Samuel Just

common/buffer.cc: fix rebuild_page_aligned typo

Introduced: 66a9fbe2c7ba59b7cd034c17865adce3432cd2cb
Fixes: #6003
Signed-off-by: Samuel Just <sam.just@inktank.com>
(cherry picked from commit cd30e5fbb1d2d29b4fc0c2b13fa3b9b4493c24f0)

History

#1 - 08/16/2013 09:32 AM - Ian Colle

- Assignee set to Samuel Just

#2 - 08/17/2013 09:09 AM - Sage Weil

- Status changed from New to Need More Info

#3 - 08/19/2013 09:38 AM - Ian Colle

Run this tasks repeatedly with logging

#4 - 08/19/2013 09:39 AM - Sage Weil

- Assignee changed from Samuel Just to Sage Weil

#5 - 08/19/2013 03:42 PM - Sage Weil

200 runs later and no luck reproducing this with logs.

/var/lib/teuthworker/archive/sage-bug-6003-a
200 passes

#6 - 08/19/2013 04:11 PM - Sage Weil

- Priority changed from Urgent to High

#7 - 08/22/2013 09:42 AM - Sage Weil

- Priority changed from High to Urgent

ubuntu@teuthology:/var/lib/teuthworker/archive/teuthology-2013-08-22_01:00:13-rados-next-testing-basic-plana/804

#8 - 08/24/2013 01:36 PM - Sage Weil

teuthology-2013-08-23_01:00:10-rados-master-testing-basic-plana/1275

#9 - 09/04/2013 09:04 PM - Sage Weil

ubuntu@teuthology:/a/teuthology-2013-09-04_17:24:00-rados-master-testing-basic-plana/21102

#10 - 09/05/2013 08:26 AM - Sage Weil

ubuntu@teuthology:/a/teuthology-2013-09-04_17:24:00-rados-master-testing-basic-plana/21102

#11 - 10/13/2013 10:23 AM - Samuel Just

ubuntu@teuthology:/a/teuthology-2013-10-10_22:04:00-rados-master-testing-basic-plana/44493/remote

#12 - 10/15/2013 01:35 PM - Sage Weil

- Assignee deleted (Sage Weil)

#13 - 10/24/2013 03:11 PM - Sage Weil

ubuntu@teuthology:/a/teuthology-2013-10-23_19:00:21-rados-dumpling-testing-basic-plana/65408

#14 - 10/28/2013 04:15 PM - Sage Weil

ubuntu@teuthology:/a/teuthology-2013-10-27_19:00:21-rados-dumpling-testing-basic-plana/71126

#15 - 10/30/2013 09:03 PM - David Zafman

- File *ceph-osd.3.log* added

I hit this same problem in recovery testing. Attached is the complete log.

#16 - 10/31/2013 08:28 AM - Greg Farnum

/a/teuthology-2013-10-28_23:01:45-kcephfs-master-testing-basic-plana/73927

#17 - 10/31/2013 10:07 AM - Samuel Just

slider:~samuelj/bug_logs/13-10-31-04:05:43/remote

With journal logging! I seem to be able to reproduce this pretty much constantly. ext4. David's was ext4 as well, iirc.

#18 - 10/31/2013 10:09 AM - Samuel Just

Kernel: 3.8.0-21-generic

#19 - 10/31/2013 10:21 AM - Sage Weil

- Status changed from *Need More Info* to *12*

#20 - 10/31/2013 05:50 PM - David Zafman

This looks like the same problem seen in a unit test.

```
2013-10-31T15:23:41.910 INFO:teuthology.orchestra.run.err:[10.214.132.32]: 2013-10-31 15:23:41.909749 7f87dfb7d780 1 journal _open
/tmp/ceph_test_filejournal.tmp.558768921 fd 8: 209715200 bytes, block size 4096 bytes, directio = 1, aio = 0
2013-10-31T15:23:41.986 INFO:teuthology.orchestra.run.out:[10.214.132.32]: test/test_filejournal.cc:199: Failure
2013-10-31T15:23:41.986 INFO:teuthology.orchestra.run.out:[10.214.132.32]: Value of: j.read_entry(inbl, seq)
2013-10-31T15:23:41.986 INFO:teuthology.orchestra.run.out:[10.214.132.32]: Actual: false
2013-10-31T15:23:41.986 INFO:teuthology.orchestra.run.out:[10.214.132.32]: Expected: true
2013-10-31T15:23:41.986 INFO:teuthology.orchestra.run.err:[10.214.132.32]: 2013-10-31 15:23:41.985155 7f87dfb7d780 -1 journal Unable to read
past sequence 0 but header indicates the journal has committed up through 1, journal is corrupt
2013-10-31T15:23:41.989 INFO:teuthology.orchestra.run.out:[10.214.132.32]: [ FAILED ] TestFileJournal.WriteManyVecs (355 ms)
```

#21 - 11/01/2013 10:41 AM - Sage Weil

- Assignee set to *Samuel Just*

#22 - 11/01/2013 02:04 PM - Samuel Just

- Status changed from 12 to Need More Info

af0c5b415ad5a863940add7c2e8f3c1ac8040ef2 should fix the most recent iteration, but it doesn't explain the instances prior to the typo.

#23 - 11/06/2013 02:36 PM - Samuel Just

- Assignee deleted (Samuel Just)

#24 - 11/15/2013 03:09 PM - Samuel Just

- Status changed from Need More Info to Can't reproduce

#25 - 12/29/2013 12:20 PM - Sage Weil

- Status changed from Can't reproduce to 12

ubuntu@teuthology:/a/teuthology-2013-12-28_23:00:01-rados-next-distro-basic-plana/18074

```
-4> 2013-12-29 08:27:34.549948 7f602df89780 2 journal open /var/lib/ceph/osd/ceph-0/journal fsid dad1a3cb-5fce-44d2-ae19-1df8090c97fd fs_op_seq 1285
-3> 2013-12-29 08:27:34.549978 7f602df89780 -1 journal FileJournal::_open: disabling aio for non-block journal. Use journal_force_aio to force use of aio anyway
-2> 2013-12-29 08:27:34.549998 7f602df89780 1 journal _open /var/lib/ceph/osd/ceph-0/journal fd 20: 104857600 bytes, block size 4096 bytes, directio = 1, aio = 0
-1> 2013-12-29 08:27:34.571046 7f602df89780 -1 journal Unable to read past sequence 1274 but header indicates the journal has committed up through 1280, journal is corrupt
0> 2013-12-29 08:27:34.573614 7f602df89780 -1 os/FileJournal.cc: In function 'bool FileJournal::read_entry(ceph::bufferlist&, uint64_t&, bool*)' thread 7f602df89780 time 2013-12-29 08:27:34.571062 os/FileJournal.cc: 1697: FAILED assert(0)
```

#26 - 02/08/2014 09:15 PM - Sage Weil

ubuntu@teuthology:/a/sage-2014-02-08_15:50:23-rados:thrash-wip-agent-testing-basic-plana/72854

#27 - 02/21/2014 02:14 PM - Sage Weil

- Status changed from 12 to Need More Info

#28 - 03/03/2014 03:56 PM - David Zafman

- Status changed from Need More Info to 12

/a/dzafman-2014-02-28_12:09:58-rados:thrash-wip-7458-testing-basic-plana/111738

```
2014-03-01 02:09:40.455986 7f2b0c373780 10 filestore(/var/lib/ceph/osd/ceph-2) open_journal at /var/lib/ceph/osd/ceph-2/journal
2014-03-01 02:09:40.456008 7f2b0c373780 0 filestore(/var/lib/ceph/osd/ceph-2) mount: enabling WRITEAHEAD journal mode: checkpoint is not enabled
2014-03-01 02:09:40.456014 7f2b0c373780 10 filestore(/var/lib/ceph/osd/ceph-2) list_collections
```

```
2014-03-01 02:09:40.456462 7f2b0c373780 -1 journal FileJournal::_open: disabling aio for non-block journal. Use journal_force_aio to force use of aio anyway
2014-03-01 02:09:40.456498 7f2b0c373780 1 journal _open /var/lib/ceph/osd/ceph-2/journal fd 20: 104857600 bytes, block size 4096 bytes, directio = 1, aio = 0
2014-03-01 02:09:40.474869 7f2b0c373780 -1 journal Unable to read past sequence 958 but header indicates the journal has committed up through 967, journal is corrupt
2014-03-01 02:09:40.489832 7f2b0c373780 -1 os/FileJournal.cc: In function 'bool FileJournal::read_entry(ceph::bufferlist&, uint64_t&, bool*)' thread 7f2b0c373780 time 2014-03-01 02:09:40.474886
os/FileJournal.cc: 1673: FAILED assert(0)
```

```
ceph version 0.77-618-g123ff9e (123ff9e18a27c6517a291aable0613b27d70d711)
1: (FileJournal::read_entry(ceph::buffer::list&, unsigned long&, bool*)+0x86f) [0x9244af]
2: (FileJournal::open(unsigned long)+0x6ad) [0x921a7d]
3: (JournalingObjectStore::journal_replay(unsigned long)+0x14f) [0x779b6f]
4: (FileStore::mount()+0x30c2) [0x74e812]
5: (OSD::do_convertfs(ObjectStore*)+0x1a) [0x6243ba]
6: (main()+0x1d88) [0x60e7b8]
7: (__libc_start_main()+0xed) [0x7f2b09fd076d]
8: ceph-osd() [0x6129b9]
```

#29 - 03/13/2014 09:18 PM - Sage Weil

- Severity changed from 3 - minor to 2 - major

#30 - 03/14/2014 09:34 AM - Sage Weil

- Status changed from 12 to Need More Info

I think we need a log for this.. hopefully we will hit it now that things are cranked up on the nightlies?

#31 - 03/25/2014 01:20 PM - Samuel Just

- Status changed from Need More Info to 12

#32 - 04/08/2014 01:14 PM - Sage Weil

- Status changed from 12 to Need More Info

#33 - 05/20/2014 01:16 PM - Sage Weil

- Priority changed from Urgent to High

#34 - 06/16/2014 01:35 PM - Sage Weil

ubuntu@teuthology:/a/teuthology-2014-06-15_02:30:07-rados-next-testing-basic-plana/308538 (btrfs)

the problem in this case looks like premature trimming. the filestore start seq is X and the first journal entry is X+20 or something. appears that perhaps the btrfs snap wait method returned before it was really durable?

#35 - 07/27/2014 09:47 PM - Sage Weil

ubuntu@teuthology:/var/lib/teuthworker/archive/teuthology-2014-07-27_02:30:01-rados-next-testing-basic-plana/380261
rados/thrash/{clusters/fixed-2.yaml fs/ext4.yaml msgr-failures/osd-delay.yaml thrashers/default.yaml workloads/radosbench.yaml}

this time on ext4, not btrfs.

The header prints out valid info:

```
-6> 2014-07-27 15:46:20.583861 7f92aa089780 10 journal read_header
-5> 2014-07-27 15:46:20.597188 7f92aa089780 10 journal header: block_size 4096 alignment 4096 max_size 104
857600
-4> 2014-07-27 15:46:20.597210 7f92aa089780 10 journal header: start 100716544
-3> 2014-07-27 15:46:20.597213 7f92aa089780 10 journal write_pos 4096
-2> 2014-07-27 15:46:20.597220 7f92aa089780 10 journal open header.fsid = 0d6e2a30-a9e5-4c1d-ba4c-fd203e3c
c64f
```

but a few lines down when we start reading entries we fail. printing the header shows bad info:

```
$27 = {flags = 41735368, fsid = {uuid = "\000\214{\002\000\000\000\000\000\300S\231\002\000\000\000"}, block_size
= 15938360, alignment = 0,
max_size = 140736923900312, start = 140267867036691, committed_up_to = 0, start_seq = 41651200}
```

including a corrupt stack from

```
#11 0x00000000009815ef in FileJournal::read_entry (this=0x53b, bl=..., next_seq=<optimized out>, corrupt=0x0)
at os/FileJournal.cc:1677
1677     in os/FileJournal.cc
```

some of FileJournal fields are okay...

```
(gdb) p write_pos
$21 = 10726185
```

others come from a bad header

```
(gdb) p read_pos
$29 = 140736923900312
(gdb) p header.start
$30 = 140267867036691
```

and other are also corrupt. fn follows header:

```
(gdb) p fn
$31 = {static npos = <optimized out>,
_M_dataplus = {<std::allocator<char>> = {<__gnu_cxx::new_allocator<char>> = {<No data fields>}, <No data fie
lds>}},
_M_p = 0x27cda80 "\275P\274"}}
```

and these precede it by a bunch

```
(gdb) p journaled_seq
$33 = 140267894734912
(gdb) p /x journaled_seq
$34 = 0x7f92aa0a0040
(gdb) p finisher_lock
$35 = {name = 0x0, id = -564455648, recursive = 255, lockdep = 127, backtrace = false, _m = {__data = {__lock
= 1315, __count = 0,
```

```
__owner = 1698063408, __users = 808673586, __kind = 1698259245, __spins = 1664363829, __list = {__prev = 0x2d633461622d6431, __next = 0x6333653330326466}}, __size = "#\005\000\000\000\000\000\000\000\060d6e2a30-a9e5-4c1d-ba4c-fd203e3c", __align = 1315}, nlock = 1714697827, locked_by = 15961913567585126656, cct = 0x27bb230, logger = 0x27cd1d8}
```

and immediately prior is

```
(gdb) p completions_lock
$39 = {name = 0x8 <Address 0x8 out of bounds>, id = -564455576, recursive = 255, lockdep = 127, backtrace = false, _m = {__data = {__lock = -1467469952, __count = 32658, __owner = 0, __users = 0, __kind = 41812256, __spins = 0, __list = {__prev = 0x27d76f0, __next = 0x7f92a8883200}}, __size = "\200\210\250\222\177\000\000\000\000\000\000\000\000 \001~\002\000\000\000\000\360v}\002\000\000\000\000\000\062\210\250\222\177\000", __align = 140267869450112}, nlock = -1467468512, locked_by = 140267869451536, cct = 0x27ef480, logger = 0x27cd4c0}
(gdb) p completions
$40 = {<std::_Deque_base<FileJournal::completion_item, std::allocator<FileJournal::completion_item> >> = {_M_impl = {<std::allocator<FileJournal::completion_item>> = {<__gnu_cxx::new_allocator<FileJournal::completion_item>> = {<No data fields>}, <No data fields>}, _M_map = 0x27cd4c0, _M_map_size = 9215206, _M_start = {_M_cur = 0x0, _M_first = 0x27cd420, _M_last = 0x27cd420, _M_node = 0x27cd4c0}, _M_finish = {_M_cur = 0x0, _M_first = 0x0, _M_last = 0x2995408, _M_node = 0x8c7c34}}}, <No data fields>}
```

so.. looks like something scribbled all over memory... ? :/

#36 - 08/02/2014 08:07 AM - Sage Weil

- Priority changed from High to Urgent

ubuntu@teuthology:/var/lib/teuthworker/archive/teuthology-2014-08-01_02:32:01-rados-master-testing-basic-plana/392342

description: rados/thrash/{clusters/fixed-2.yaml fs/xfs.yaml msgr-failures/few.yaml thrashers/default.yaml workloads/radosbench.yaml}

this time, the read_entry frame is partly corrupt, too: this is bad, although the other fields look ok:

```
#11 0x0000000000984eff in FileJournal::read_entry (this=0x59e, bl=..., next_seq=<optimized out>, corrupt=0x0)
at os/FileJournal.cc:1678
1678 os/FileJournal.cc: No such file or directory.
(gdb) p this
$13 = (FileJournal * const) 0x59e
(gdb) p pos
```



```
$14 = 44053728
(gdb) p next_pos
$15 = 97988608
(gdb) p seq
$16 = 1413
(gdb) up
#12 0x0000000009824cd in FileJournal::open (this=0x7fff89c59f30, fs_op_seq=1444) at os/FileJournal.cc:514
514      in os/FileJournal.cc
(gdb) p header
$17 = {flags = 43775176, fsid = {uuid = "\000\\\232\002\000\000\000\000\3000\246\002\000\000\000"}, block_size
= 15946952, alignment = 0, max_size = 140735504818856,
start = 140553230486547, committed_up_to = 0, start_seq = 43670528}
```

*this contents still screwed up, though!

#37 - 09/28/2014 11:02 PM - Shambhu Rajak

- File Log79set_5.zip added

I saw similar issue hitting my 3 node cluster:

```
node1 - mon.0
        osd.0
        osd.1
        osd.2
node2 - mon.1
        osd.3
        osd.4
        osd.5
node3 - mon.2
        client.0
```

Platform: Ubuntu 14.04 Trusty

Snippet from teuthology Log:

While reviving osd.4 saw error os/FileJournal.cc: 1693: FAILED assert(0)

Snippet from teuthology log(complete Teuthology attached Log79set_5.zip):

```
2014-09-15T15:02:46.135 INFO:teuthology.task.thrashosds.thrasher:Reviving osd 4
2014-09-15T15:02:46.135 INFO:teuthology.task.ceph.osd.4:Restarting
2014-09-15T15:02:46.135 DEBUG:teuthology.orchestra.run:Running [10.15.16.60]: 'sudo adjust-ulimits ceph-covera
ge /home/ubuntu/ceph/test/archive/coverage daemon-helper kill ceph-osd -f -i 4'
2014-09-15T15:02:46.693 INFO:teuthology.task.ceph.osd.4:Started
2014-09-15T15:02:46.694 DEBUG:teuthology.orchestra.run:Running [10.15.16.60]: 'sudo adjust-ulimits ceph-covera
ge /home/ubuntu/ceph/test/archive/coverage ceph --admin-daemon /var/run/ceph/ceph-osd.4.asok dump_ops_in_flight
'
2014-09-15T15:02:46.758 INFO:teuthology.task.ceph.osd.4.out:[10.15.16.60]: starting osd.4 at :/0 osd_data /var
/lib/ceph/osd/ceph-4 /var/lib/ceph/osd/ceph-4/journal
2014-09-15T15:02:46.831 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2014-09-15 09:32:46.701026 7f3be6f9
6900 -1 journal FileJournal::_open: disabling aio for non-block journal. Use journal_force_aio to force use o
f aio anyway
2014-09-15T15:02:46.831 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2014-09-15 09:32:46.701614 7f3be6f9
6900 -1 journal Unable to read past sequence 1738 but header indicates the journal has committed up through 17
74, journal is corrupt
2014-09-15T15:02:46.866 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: os/FileJournal.cc: In function 'boo
l FileJournal::read_entry(ceph::bufferlist&, uint64_t&, bool*)' thread 7f3be6f96900 time 2014-09-15 09:32:46.7
01625
2014-09-15T15:02:46.866 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: os/FileJournal.cc: 1693: FAILED ass
```

```

ert(0)
2014-09-15T15:02:46.874 INFO:teuthology.task.radosbench.radosbench.0.out:[10.15.16.59]: 336 16 17
90 1774 21.1158 0 - 2.98315
2014-09-15T15:02:46.883 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: ceph version 0.84-sd-sprint4 (3215
c520e1306f50d0094b5646636c02456c9df4)
2014-09-15T15:02:46.884 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 1: (ceph::__ceph_assert_fail(char
const*, char const*, int, char const*)+0x8b) [0xb776db]
2014-09-15T15:02:46.884 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2: (FileJournal::read_entry(ceph::
buffer::list&, unsigned long&, bool*)+0xaa4) [0xa3ec34]
2014-09-15T15:02:46.884 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 3: (FileJournal::open(unsigned lon
g)+0x66d) [0xa4055d]
2014-09-15T15:02:46.884 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 4: (JournalingObjectStore::journal
_replay(unsigned long)+0x103) [0x92e9b3]
2014-09-15T15:02:46.885 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 5: (FileStore::mount()+0x3eaa) [0x
902e3a]
2014-09-15T15:02:46.885 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 6: (OSD::do_convertfs(ObjectStore*
)+0x1a) [0x66b9ba]
2014-09-15T15:02:46.885 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 7: (main()+0x23ee) [0x64379e]
2014-09-15T15:02:46.885 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 8: (__libc_start_main()+0xf5) [0x7
f3be3f4dec5]
2014-09-15T15:02:46.885 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 9: ceph-osd() [0x65bc07]
2014-09-15T15:02:46.885 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: NOTE: a copy of the executable, or
`objdump -rds <executable>` is needed to interpret this.
2014-09-15T15:02:46.886 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2014-09-15 09:32:46.741006 7f3be6f9
6900 -1 os/FileJournal.cc: In function 'bool FileJournal::read_entry(ceph::bufferlist&, uint64_t&, bool*)' thr
ead 7f3be6f96900 time 2014-09-15 09:32:46.701625
2014-09-15T15:02:46.886 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: os/FileJournal.cc: 1693: FAILED ass
ert(0)
2014-09-15T15:02:46.886 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]:
2014-09-15T15:02:46.886 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: ceph version 0.84-sd-sprint4 (3215
c520e1306f50d0094b5646636c02456c9df4)
2014-09-15T15:02:46.887 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 1: (ceph::__ceph_assert_fail(char
const*, char const*, int, char const*)+0x8b) [0xb776db]
2014-09-15T15:02:46.887 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2: (FileJournal::read_entry(ceph::
buffer::list&, unsigned long&, bool*)+0xaa4) [0xa3ec34]
2014-09-15T15:02:46.887 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 3: (FileJournal::open(unsigned lon
g)+0x66d) [0xa4055d]
2014-09-15T15:02:46.887 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 4: (JournalingObjectStore::journal
_replay(unsigned long)+0x103) [0x92e9b3]
2014-09-15T15:02:46.887 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 5: (FileStore::mount()+0x3eaa) [0x
902e3a]
2014-09-15T15:02:46.888 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 6: (OSD::do_convertfs(ObjectStore*
)+0x1a) [0x66b9ba]
2014-09-15T15:02:46.888 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 7: (main()+0x23ee) [0x64379e]
2014-09-15T15:02:46.888 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 8: (__libc_start_main()+0xf5) [0x7
f3be3f4dec5]
2014-09-15T15:02:46.888 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 9: ceph-osd() [0x65bc07]
2014-09-15T15:02:46.888 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: NOTE: a copy of the executable, or
`objdump -rds <executable>` is needed to interpret this.
2014-09-15T15:02:46.888 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]:
2014-09-15T15:02:47.058 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: -3> 2014-09-15 09:32:46.701026
7f3be6f96900 -1 journal FileJournal::_open: disabling aio for non-block journal. Use journal_force_aio to for
ce use of aio anyway
2014-09-15T15:02:47.058 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: -1> 2014-09-15 09:32:46.701614
7f3be6f96900 -1 journal Unable to read past sequence 1738 but header indicates the journal has committed up th
rough 1774, journal is corrupt
2014-09-15T15:02:47.059 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 0> 2014-09-15 09:32:46.741006
7f3be6f96900 -1 os/FileJournal.cc: In function 'bool FileJournal::read_entry(ceph::bufferlist&, uint64_t&, boo
l*)' thread 7f3be6f96900 time 2014-09-15 09:32:46.701625
2014-09-15T15:02:47.059 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: os/FileJournal.cc: 1693: FAILED ass
ert(0)
2014-09-15T15:02:47.059 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]:
2014-09-15T15:02:47.060 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: ceph version 0.84-sd-sprint4 (3215
c520e1306f50d0094b5646636c02456c9df4)
2014-09-15T15:02:47.060 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 1: (ceph::__ceph_assert_fail(char
const*, char const*, int, char const*)+0x8b) [0xb776db]
2014-09-15T15:02:47.060 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2: (FileJournal::read_entry(ceph::
buffer::list&, unsigned long&, bool*)+0xaa4) [0xa3ec34]
2014-09-15T15:02:47.060 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 3: (FileJournal::open(unsigned lon
g)+0x66d) [0xa4055d]
2014-09-15T15:02:47.060 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 4: (JournalingObjectStore::journal
_replay(unsigned long)+0x103) [0x92e9b3]
2014-09-15T15:02:47.061 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 5: (FileStore::mount()+0x3eaa) [0x
902e3a]
2014-09-15T15:02:47.061 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 6: (OSD::do_convertfs(ObjectStore*

```

```

)+0x1a) [0x66b9ba]
2014-09-15T15:02:47.061 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 7: (main()+0x23ee) [0x64379e]
2014-09-15T15:02:47.061 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 8: (__libc_start_main()+0xf5) [0x7f3be3f4dec5]
2014-09-15T15:02:47.061 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 9: ceph-osd() [0x65bc07]
2014-09-15T15:02:47.061 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: NOTE: a copy of the executable, or
`objdump -rdS <executable>` is needed to interpret this.
2014-09-15T15:02:47.062 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]:
2014-09-15T15:02:47.062 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: terminate called after throwing an
instance of 'ceph::FailedAssertion'
2014-09-15T15:02:47.062 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: *** Caught signal (Aborted) **
2014-09-15T15:02:47.062 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: in thread 7f3be6f96900
2014-09-15T15:02:47.062 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: ceph version 0.84-sd-sprint4 (3215
c520e1306f50d0094b5646636c02456c9df4)
2014-09-15T15:02:47.063 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 1: ceph-osd() [0xa8c7ff]
2014-09-15T15:02:47.063 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2: (()+0x10340) [0x7f3be5c49340]
2014-09-15T15:02:47.063 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 3: (gsignal()+0x39) [0x7f3be3f62bb
9]
2014-09-15T15:02:47.063 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 4: (abort()+0x148) [0x7f3be3f65fc8
]
2014-09-15T15:02:47.063 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 5: (__gnu_cxx::__verbose_terminate
_handler()+0x155) [0x7f3be486e6b5]
2014-09-15T15:02:47.063 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 6: (()+0x5e836) [0x7f3be486c836]
2014-09-15T15:02:47.064 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 7: (()+0x5e863) [0x7f3be486c863]
2014-09-15T15:02:47.064 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 8: (()+0x5eaa2) [0x7f3be486caa2]
2014-09-15T15:02:47.064 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 9: (ceph::__ceph_assert_fail(char
const*, char const*, int, char const*)+0x278) [0xb778c8]
2014-09-15T15:02:47.064 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 10: (FileJournal::read_entry(ceph:
:buffer::list&, unsigned long&, bool*)+0xaa4) [0xa3ec34]
2014-09-15T15:02:47.064 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 11: (FileJournal::open(unsigned lo
ng)+0x66d) [0xa4055d]
2014-09-15T15:02:47.064 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 12: (JournalingObjectStore::journa
l_replay(unsigned long)+0x103) [0x92e9b3]
2014-09-15T15:02:47.065 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 13: (FileStore::mount()+0x3eaa) [0
x902e3a]
2014-09-15T15:02:47.065 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 14: (OSD::do_convertfs(ObjectStore
*)+0x1a) [0x66b9ba]
2014-09-15T15:02:47.065 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 15: (main()+0x23ee) [0x64379e]
2014-09-15T15:02:47.065 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 16: (__libc_start_main()+0xf5) [0x
7f3be3f4dec5]
2014-09-15T15:02:47.065 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 17: ceph-osd() [0x65bc07]
2014-09-15T15:02:47.104 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2014-09-15 09:32:46.767054 7f3be6f9
6900 -1 *** Caught signal (Aborted) **
2014-09-15T15:02:47.104 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: in thread 7f3be6f96900
2014-09-15T15:02:47.104 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]:
2014-09-15T15:02:47.105 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: ceph version 0.84-sd-sprint4 (3215
c520e1306f50d0094b5646636c02456c9df4)
2014-09-15T15:02:47.105 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 1: ceph-osd() [0xa8c7ff]
2014-09-15T15:02:47.105 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2: (()+0x10340) [0x7f3be5c49340]
2014-09-15T15:02:47.105 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 3: (gsignal()+0x39) [0x7f3be3f62bb
9]
2014-09-15T15:02:47.105 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 4: (abort()+0x148) [0x7f3be3f65fc8
]
2014-09-15T15:02:47.106 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 5: (__gnu_cxx::__verbose_terminate
_handler()+0x155) [0x7f3be486e6b5]
2014-09-15T15:02:47.106 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 6: (()+0x5e836) [0x7f3be486c836]
2014-09-15T15:02:47.106 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 7: (()+0x5e863) [0x7f3be486c863]
2014-09-15T15:02:47.106 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 8: (()+0x5eaa2) [0x7f3be486caa2]
2014-09-15T15:02:47.106 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 9: (ceph::__ceph_assert_fail(char
const*, char const*, int, char const*)+0x278) [0xb778c8]
2014-09-15T15:02:47.106 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 10: (FileJournal::read_entry(ceph:
:buffer::list&, unsigned long&, bool*)+0xaa4) [0xa3ec34]
2014-09-15T15:02:47.107 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 11: (FileJournal::open(unsigned lo
ng)+0x66d) [0xa4055d]
2014-09-15T15:02:47.107 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 12: (JournalingObjectStore::journa
l_replay(unsigned long)+0x103) [0x92e9b3]
2014-09-15T15:02:47.107 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 13: (FileStore::mount()+0x3eaa) [0
x902e3a]
2014-09-15T15:02:47.107 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 14: (OSD::do_convertfs(ObjectStore
*)+0x1a) [0x66b9ba]
2014-09-15T15:02:47.107 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 15: (main()+0x23ee) [0x64379e]
2014-09-15T15:02:47.108 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 16: (__libc_start_main()+0xf5) [0x
7f3be3f4dec5]
2014-09-15T15:02:47.108 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 17: ceph-osd() [0x65bc07]
2014-09-15T15:02:47.108 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: NOTE: a copy of the executable, or

```

```

`objdump -rdS <executable>` is needed to interpret this.
2014-09-15T15:02:47.108 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]:
2014-09-15T15:02:47.110 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 0> 2014-09-15 09:32:46.767054
7f3be6f96900 -1 *** Caught signal (Aborted) **
2014-09-15T15:02:47.110 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: in thread 7f3be6f96900
2014-09-15T15:02:47.110 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]:
2014-09-15T15:02:47.110 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: ceph version 0.84-sd-sprint4 (3215
c520e1306f50d0094b5646636c02456c9df4)
2014-09-15T15:02:47.110 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 1: ceph-osd() [0xa8c7ff]
2014-09-15T15:02:47.111 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 2: (()+0x10340) [0x7f3be5c49340]
2014-09-15T15:02:47.111 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 3: (gsignal()+0x39) [0x7f3be3f62bb
9]
2014-09-15T15:02:47.111 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 4: (abort()+0x148) [0x7f3be3f65fc8
]
2014-09-15T15:02:47.111 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 5: (__gnu_cxx::__verbose_terminate
_handler()+0x155) [0x7f3be486e6b5]
2014-09-15T15:02:47.111 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 6: (()+0x5e836) [0x7f3be486c836]
2014-09-15T15:02:47.111 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 7: (()+0x5e863) [0x7f3be486c863]
2014-09-15T15:02:47.112 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 8: (()+0x5eaa2) [0x7f3be486caa2]
2014-09-15T15:02:47.112 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 9: (ceph::__ceph_assert_fail(char
const*, char const*, int, char const*)+0x278) [0xb778c8]
2014-09-15T15:02:47.112 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 10: (FileJournal::read_entry(ceph:
:buffer::list&, unsigned long&, bool*)+0xaa4) [0xa3ec34]
2014-09-15T15:02:47.112 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 11: (FileJournal::open(unsigned lo
ng)+0x66d) [0xa4055d]
2014-09-15T15:02:47.112 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 12: (JournalingObjectStore::journa
l_replay(unsigned long)+0x103) [0x92e9b3]
2014-09-15T15:02:47.112 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 13: (FileStore::mount()+0x3eaa) [0
x902e3a]
2014-09-15T15:02:47.113 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 14: (OSD::do_convertfs(ObjectStore
*)+0x1a) [0x66b9ba]
2014-09-15T15:02:47.113 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 15: (main()+0x23ee) [0x64379e]
2014-09-15T15:02:47.113 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 16: (__libc_start_main()+0xf5) [0x
7f3be3f4dec5]
2014-09-15T15:02:47.113 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: 17: ceph-osd() [0x65bc07]
2014-09-15T15:02:47.113 INFO:teuthology.task.ceph.osd.4.err:[10.15.16.60]: NOTE: a copy of the executable, or
`objdump -rdS <executable>` is needed to interpret this.

```

#38 - 10/07/2014 06:37 AM - Sage Weil

```
-1> 2014-10-06 16:29:02.286272 7f2368425780 -1 journal Unable to read past sequence 39581 but header indicates the journal has committed up through 39647, journal is corrupt
0> 2014-10-06 16:29:02.297264 7f2368425780 -1 os/FileJournal.cc: In function 'bool FileJournal::read_entry(ceph::bufferlist&, uint64_t&, bool*)' thread 7f2368425780 time 2014-10-06 16:29:02.286293
os/FileJournal.cc: 1677: FAILED assert(0)
```

```
ceph version 0.80.6-48-g884f7c4 (884f7c40c4a28d519847d3995c8d98e5837ceaf0)
1: (FileJournal::read_entry(ceph::bufferlist&, unsigned long&, bool*)+0x86f) [0x97918f]
2: (FileJournal::open(unsigned long)+0x6b5) [0x9766f5]
3: (JournalingObjectStore::journal_replay(unsigned long)+0x26d) [0x8bd39d]
4: (FileStore::mount()+0x30c2) [0x890792]
5: (OSD::do_convertfs(ObjectStore*)+0x1a) [0x61fd8a]
6: (main()+0x1d88) [0x6079b8]
7: (__libc_start_main()+0xed) [0x7f236622776d]
8: ceph-osd() [0x60bca9]
NOTE: a copy of the executable, or `objdump -rdS <executable>` is needed to interpret this.
```

ubuntu@teuthology:/a/teuthology-2014-10-05_10:00:04-upgrade:firefly-firefly-distro-basic-multi/528734

#39 - 11/21/2014 10:00 AM - Sage Weil

ubuntu@teuthology:/a/sage-2014-11-20_17:03:30-rados:thrash-wip-watch-notify-distro-basic-multi/611427

#40 - 12/16/2014 06:26 AM - Sage Weil

ubuntu@teuthology:/a/sage-2014-12-01_11:11:52-rados-next-distro-basic-multi/629379

#41 - 01/08/2015 11:26 AM - Loïc Dachary

http://qa-proxy.ceph.com/teuthology/sage-2015-01-06_09:44:19-rados-wip-sage-testing-firefly---basic-multi/688065/

```
-13> 2015-01-07 16:52:19.063303 7f6a93337780 10 filestore(/var/lib/ceph/osd/ceph-2) list_collections
-12> 2015-01-07 16:52:19.088037 7f6a93337780 10 journal journal_replay fs op_seq 2835
-11> 2015-01-07 16:52:19.088052 7f6a93337780 2 journal open /var/lib/ceph/osd/ceph-2/journal fsid 5af9380a-a5a6-4ca2-9310-311ee01ac6cc fs_op_seq 2835
-10> 2015-01-07 16:52:19.088074 7f6a93337780 -1 journal FileJournal::_open: disabling aio for non-block journal. Use journal_force_aio to force use of aio anyway
-9> 2015-01-07 16:52:19.088082 7f6a93337780 10 journal _open journal is not a block device, NOT checking disk write cache on '/var/lib/ceph/osd/ceph-2/journal'
-8> 2015-01-07 16:52:19.088084 7f6a93337780 1 journal _open /var/lib/ceph/osd/ceph-2/journal fd 20: 104857600 bytes, block size 4096 bytes, directio = 1, aio = 0
-7> 2015-01-07 16:52:19.088094 7f6a93337780 10 journal read_header
-6> 2015-01-07 16:52:19.088099 7f6a8dd70700 20 filestore(/var/lib/ceph/osd/ceph-2) sync_entry waiting for max_interval 5.000000
-5> 2015-01-07 16:52:19.092091 7f6a93337780 10 journal header: block_size 4096 alignment 4096 max_size 104857600
-4> 2015-01-07 16:52:19.092105 7f6a93337780 10 journal header: start 101089280
-3> 2015-01-07 16:52:19.092107 7f6a93337780 10 journal write_pos 4096
```

```
-2> 2015-01-07 16:52:19.092111 7f6a93337780 10 journal open header.fsid = 5af9380a-a5a6-4ca2-9310-311ee01a
c6cc
-1> 2015-01-07 16:52:19.096659 7f6a93337780 -1 journal Unable to read past sequence 2751 but header indica
tes the journal has committed up through 2778, journal is corrupt
0> 2015-01-07 16:52:19.111136 7f6a93337780 -1 os/FileJournal.cc: In function 'bool FileJournal::read_entr
y(ceph::bufferlist&, uint64_t&, bool*)' thread 7f6a93337780 time 2015-01-07 16:52:19.096681
os/FileJournal.cc: 1693: FAILED assert(0)
```

```
ceph version 0.80.7-185-gf43f147 (f43f14748d5efd382edc7f27104f29898e39b0d5)
1: (FileJournal::read_entry(ceph::buffer::list&, unsigned long&, bool*)+0x86f) [0x978c8f]
2: (FileJournal::open(unsigned long)+0x6b5) [0x9761f5]
3: (JournalingObjectStore::journal_replay(unsigned long)+0x26d) [0x8bca8d]
4: (FileStore::mount()+0x30c2) [0x88fe82]
5: (OSD::do_convertfs(ObjectStore*)+0x1a) [0x61f30a]
6: (main()+0x1d88) [0x606db8]
7: (__libc_start_main()+0xed) [0x7f6a9113a76d]
8: ceph-osd() [0x60b0a9]
NOTE: a copy of the executable, or `objdump -rdS <executable>` is needed to interpret this.
```

#42 - 02/06/2015 05:49 PM - Samuel Just

- Status changed from Need More Info to 7
- Assignee set to Samuel Just

The journalq is being populated with positions one sequence number off in do_read_entry. If we trim committed_thru to a sequence number within the journalq portion populated during replay, we end up with a header.start one forward of where it should be causing FileJournal::open() to assume that it's at the end of the journal. Subsequent calls into read_entry will be with the wrong offset/sequence number combo causing the assert to fail.

#43 - 02/06/2015 06:20 PM - Samuel Just

I introduced this in March 2013, the fix should be backported to dumpling, firefly, and giant.

#44 - 02/06/2015 06:21 PM - Samuel Just

- Backport set to dumpling, firefly, giant

#45 - 02/17/2015 06:49 PM - David Zafman

- Status changed from 7 to Pending Backport

#46 - 03/11/2015 11:01 AM - Loïc Dachary

- <https://github.com/ceph/ceph/pull/3691>

#47 - 03/11/2015 11:01 AM - Loïc Dachary

- Description updated

#48 - 03/11/2015 11:18 AM - Loïc Dachary

- firefly backport <https://github.com/ceph/ceph/pull/3960>

#49 - 03/19/2015 09:02 AM - Loïc Dachary

- giant backport <https://github.com/ceph/ceph/pull/4088>

#50 - 03/26/2015 12:03 PM - Loïc Dachary

- dumpling backport <https://github.com/ceph/ceph/pull/2611>

#51 - 03/26/2015 12:09 PM - Loïc Dachary

- Status changed from Pending Backport to Resolved

Files

ceph-osd.3.log	27.8 MB	10/30/2013	David Zafman
Log79set_5.zip	8.73 MB	09/28/2014	Shambhu Rajak