Ceph - Bug #2275

osd: crash in FileJournal::wrap_read_bl

04/12/2012 09:35 PM - Sage Weil

Status:	Resolved	**% Done:**	0%
Priority:	Urgent	**Spent time:**	0.00 hour
Assignee:			
Category:			
Target version:	v0.47		
Source:	Support	**Reviewed:**	
Tags:		**Affected Versions:**	
Backport:		**ceph-qa-suite:**	
Regression:	No	**Pull request ID:**	
Severity:	3 - minor	**Crash signature:**	

Description

```
2012-04-13 04:30:16.604664 7fb526c98780 ceph version 0.45-3-gcfac4a9 (commit:cfac4a97160f7497a7a4c
782bb9e6fe0361ad812), process ceph-osd, pid 19357
2012-04-13 04:30:16.604791 7fb526c98780 -- [2607:f298:4:2243::7052]:6800/19357 accepter.bind my_in
st.addr is [2607:f298:4:2243::7052]:6800/19357 need_addr=0
2012-04-13 04:30:16.604812 7fb526c98780 -- [2607:f298:4:3243::7052]:6800/19357 accepter.bind my_in
st.addr is [2607:f298:4:3243::7052]:6800/19357 need_addr=0
2012-04-13 04:30:16.604825 7fb526c98780 -- [2607:f298:4:3243::7052]:6801/19357 accepter.bind my_in
st.addr is [2607:f298:4:3243::7052]:6801/19357 need_addr=0
2012-04-13 04:30:16.605624 7fb526c98780 filestore(/srv/ceph/osd/850) basedir /srv/ceph/osd/850 jou
rnal /srv/ceph/devices/osd.850.journal
2012-04-13 04:30:16.605939 7fb526c98780 filestore(/srv/ceph/osd/850) mount FIEMAP ioctl is NOT sup
ported
2012-04-13 04:30:16.605949 7fb526c98780 filestore(/srv/ceph/osd/850) mount did NOT detect btrfs
2012-04-13 04:30:16.605989 7fb526c98780 filestore(/srv/ceph/osd/850) mount fsid is 2f97b61b-26c2-4
8ea-81c8-2861d70ebe1a
2012-04-13 04:30:16.606022 7fb526c98780 filestore(/srv/ceph/osd/850) mount found snaps <>
2012-04-13 04:30:16.606036 7fb526c98780 filestore(/srv/ceph/osd/850) mount op_seq is 369
2012-04-13 04:30:16.664633 7fb526c98780 filestore (init)dbobjectmap: seq is 1
2012-04-13 04:30:16.664651 7fb526c98780 filestore(/srv/ceph/osd/850) open_journal at /srv/ceph/dev
ices/osd.850.journal
2012-04-13 04:30:16.664659 7fb526c98780 filestore(/srv/ceph/osd/850) mount: enabling WRITEAHEAD jo
urnal mode: btrfs not detected
2012-04-13 04:30:16.664666 7fb526c98780 filestore(/srv/ceph/osd/850) list_collections
2012-04-13 04:30:16.664794 7fb526c98780 journal journal_replay fs op_seq 369
2012-04-13 04:30:16.664805 7fb526c98780 journal open /srv/ceph/devices/osd.850.journal fsid 2f97b6
1b-26c2-48ea-81c8-2861d70ebe1a fs_op_seq 369
2012-04-13 04:30:16.664828 7fb523912700 filestore(/srv/ceph/osd/850) sync_entry waiting for max_in
terval 5.000000
2012-04-13 04:30:16.664848 7fb526c98780 journal _open_block_device: no journal size specified in c
onfiguration. We'll use the entire block device (size: 10736401408)
2012-04-13 04:30:16.665862 7fb526c98780 journal _check_disk_write_cache: fclose error: (61) No dat
a available
2012-04-13 04:30:16.665883 7fb526c98780 journal _open /srv/ceph/devices/osd.850.journal fd 21: 107
36398336 bytes, block size 4096 bytes, directio = 1, aio = 0
2012-04-13 04:30:16.665888 7fb526c98780 journal read_header
2012-04-13 04:30:16.666215 7fb526c98780 journal header: block_size 4096 alignment 4096 max_size 10
736398336
2012-04-13 04:30:16.666225 7fb526c98780 journal header: start 1746382848
2012-04-13 04:30:16.666227 7fb526c98780 journal  write_pos 4096
2012-04-13 04:30:16.666242 7fb526c98780 journal open header.fsid = 2f97b61b-26c2-48ea-81c8-2861d70
ebe1a
2012-04-13 04:30:16.680653 7fb526c98780 journal read_entry 1746382848 : seq 367 871 bytes
2012-04-13 04:30:17.234685 7fb526c98780 journal read_entry 1746391040 : seq 368 79695435 bytes
```

```
2012-04-13 04:30:21.664984 7fb523912700 filestore(/srv/ceph/osd/850) sync_entry woke after 5.00016
0
2012-04-13 04:30:21.665018 7fb523912700 journal commit_start op_seq 369, applied_seq 369, committe
d_seq 369
2012-04-13 04:30:21.665022 7fb523912700 journal commit_start blocked, all open_ops have completed
2012-04-13 04:30:21.665025 7fb523912700 journal commit_start nothing to do
2012-04-13 04:30:21.665028 7fb523912700 journal commit_start
2012-04-13 04:30:21.665033 7fb523912700 filestore(/srv/ceph/osd/850) sync_entry waiting for max_in
terval 5.000000
2012-04-13 04:30:26.665113 7fb523912700 filestore(/srv/ceph/osd/850) sync_entry woke after 5.00007
0
2012-04-13 04:30:26.665136 7fb523912700 journal commit_start op_seq 369, applied_seq 369, committe
d_seq 369
2012-04-13 04:30:26.665140 7fb523912700 journal commit_start blocked, all open_ops have completed
2012-04-13 04:30:26.665142 7fb523912700 journal commit_start nothing to do
2012-04-13 04:30:26.665145 7fb523912700 journal commit_start
2012-04-13 04:30:26.665149 7fb523912700 filestore(/srv/ceph/osd/850) sync_entry waiting for max_in
terval 5.000000
2012-04-13 04:30:31.665266 7fb523912700 filestore(/srv/ceph/osd/850) sync_entry woke after 5.00010
8
2012-04-13 04:30:31.665289 7fb523912700 journal commit_start op_seq 369, applied_seq 369, committe
d_seq 369
2012-04-13 04:30:31.665292 7fb523912700 journal commit_start blocked, all open_ops have completed
2012-04-13 04:30:31.665304 7fb523912700 journal commit_start nothing to do
2012-04-13 04:30:31.665307 7fb523912700 journal commit_start
2012-04-13 04:30:31.665311 7fb523912700 filestore(/srv/ceph/osd/850) sync_entry waiting for max_in
terval 5.000000
2012-04-13 04:30:35.016606 7fb526c98780 journal FileJournal::wrap_read_bl: safe_read_exact 1826092
999~2897424950 returned -1397542346
os/FileJournal.cc: In function 'void FileJournal::wrap_read_bl(off64_t&, int64_t, ceph::bufferlist
&)' thread 7fb526c98780 time 2012-04-13 04:30:35.016646
os/FileJournal.cc: 1535: FAILED assert(0)
 ceph version 0.45-3-gcfac4a9 (commit:cfac4a97160f7497a7a4c782bb9e6fe0361ad812)
 1: (FileJournal::wrap_read_bl(long&, long, ceph::buffer::list&)+0x1f9) [0x7773d9]
 2: (FileJournal::read_entry(ceph::buffer::list&, unsigned long&)+0x35f) [0x78170f]
 3: (FileJournal::open(unsigned long)+0x834) [0x77bb74]
 4: (JournalingObjectStore::journal_replay(unsigned long)+0x1ae) [0x7bc4be]
 5: (FileStore::mount()+0x2533) [0x7a3693]
 6: ceph-osd() [0x5a0c1a]
 7: (OSD::convertfs(std::string const&, std::string const&)+0xbb) [0x5a1abb]
 8: (main()+0x218c) [0x5203cc]
 9: (__libc_start_main()+0xed) [0x7fb524b4f76d]
 10: ceph-osd() [0x52238d]
```

Associated revisions

Revision 3509b039 - 05/05/2012 05:01 PM - Sage Weil

safe_io: int -> ssize_t

int is 32-bit on 64-bit archs, but ssize_t is 64-bits. This fixes overflow
when reading large (>2GB) extends.

Fixes: #2275
Signed-off-by: Sage Weil <sage.weil@dreamhost.com>

History

#1 - 04/18/2012 03:02 PM - Sage Weil

- Priority changed from Immediate to Urgent

#2 - 05/05/2012 10:53 AM - Sage Weil

- Category deleted (OSD)

- Status changed from 12 to Resolved

- Target version set to v0.47

converted int -> ssize_t in common/safe_io.c, 3509b039a28d41c7ae1b3d482d67a27f8e5739e8