Ceph - Bug #11211

osd/ReplicatedPG.cc: 2141: FAILED assert(op == prdop->op)

03/23/2015 06:07 PM - Sage Weil

<table>
<thead>
<tr>
<th>Status:</th>
<th>Resolved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>Urgent</td>
</tr>
<tr>
<td>Assignee:</td>
<td></td>
</tr>
<tr>
<td>Category:</td>
<td></td>
</tr>
<tr>
<td>Start date:</td>
<td>03/23/2015</td>
</tr>
<tr>
<td>Due date:</td>
<td></td>
</tr>
<tr>
<td>% Done:</td>
<td>0%</td>
</tr>
<tr>
<td>Estimated time:</td>
<td>0.00 hour</td>
</tr>
<tr>
<td>Spent time:</td>
<td>0.00 hour</td>
</tr>
<tr>
<td>Source:</td>
<td>Q/A</td>
</tr>
<tr>
<td>Tags:</td>
<td></td>
</tr>
<tr>
<td>Backport:</td>
<td>hammer</td>
</tr>
<tr>
<td>Regression:</td>
<td>No</td>
</tr>
<tr>
<td>Severity:</td>
<td>2 - major</td>
</tr>
</tbody>
</table>

Description

0> 2015-03-21 07:04:19.095762 7f4c41ed4700 -1 osd/ReplicatedPG.cc: In function 'void ReplicatedPG::finish_proxy_read(hobject_t, ceph_tid_t, int)' thread 7f4c41ed4700 time 2015-03-21 07:04:19.094195 osd/ReplicatedPG.cc: 2141: FAILED assert(op == prdop->op)

ceph version 0.93-164-g120300f (120300fbebdc34c23fc24633edfd8af65e93c)
1: (ceph::__ceph_assert_fail(char const*, char const*, int, char const*)+0x8b) [0xbc6f6b]
2: (ReplicatedPG::finish_proxy_read(hobject_t, unsigned long, int)+0x109d) [0x847f0d]
3: (C_ProxyRead::finish(int)+0xbd) [0x8d930d]
4: (Context::complete(int)+0x9) [0x6ca639]
5: (Finisher::finisher_thread_entry()+0x158) [0xaf38d8]
6: (clone()+0x6d) [0x7f4c62ceb38d]

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

ubuntu@teuthology:/a/sage-2015-03-20_17:42:17-rgw-hammer-distro-basic-multi/814306
description: rgw/multifs/{clusters/fixed-2.yaml frontend/apache.yaml fs/ext4.yaml overrides.yaml rgw_pool_type/ec-cache.yaml tasks/rgw_readwrite.yaml}

Related issues:
Duplicated by Bug #11375: Proxy_Read BUG in Cachetier when using ecpool as co...

Duplicate 04/13/2015

Associated revisions

Revision 560a5839 - 03/25/2015 06:47 AM - Zhiqiang Wang

osd/ReplicatedPG: don't check order in finish_proxy_read

Read doesn't need to be ordered. So when proxy read comes back from base tier, it's not necessarily at the front of the in progress list.

Fixes: #11211

Signed-off-by: Zhiqiang Wang <zhiqiang.wang@intel.com>

Revision 0ee022b1 - 04/24/2015 09:07 PM - Zhiqiang Wang

11/25/2015 1/5
osd/ReplicatedPG: don't check order in finish_proxy_read

Read doesn't need to be ordered. So when proxy read comes back from base tier, it's not necessarily at the front of the in progress list.

Fixes: #11211

Signed-off-by: Zhiqiang Wang <zhiqiang.wang@intel.com>
(cherry picked from commit 560a5839c0d1852b5816937b845b60390777636c)

---

**History**

**#1 - 03/24/2015 07:10 AM - Zhiqiang Wang**

This is the problem we talked before. Read is not necessarily ordered. So when the proxy read comes back from the base tier, it may be not the first element in the list. I'm going to write a patch to fix this.

**#2 - 03/24/2015 08:11 AM - Zhiqiang Wang**

A fix for this is in https://github.com/ceph/ceph/pull/4155. Pls help review, thx!

**#3 - 03/24/2015 09:04 PM - Yuri Weinstein**

Jobs: ['817403', '817431', '817459', '817515', '817543']

Assertion: osd/ReplicatedPG.cc: 2141: FAILED assert(op == prdop->op)
ceph version 0.93-167-gc60ba6a (c60ba6af8cf45ce8fcf2df04aco1e76edf4a96)
  1: (ceph::__ceph_assert_fail(char const*, char const*, int, char const*)+0x85) [0xbc93d5]
  2: (ReplicatedPG::finish_proxy_read(hobject_t, unsigned long, int)+0x109a) [0x83704a]
  3: (C_ProxyRead::finish(int)+0xed) [0x8c7dad]
  4: (Context::complete(int)+0x9) [0x6be449]
  5: (Finisher::finisher_thread_entry()+0x168) [0xaee748]
  6: (;+0x7d3f) [0x7fe43c93c93c93]
  7: (clone()+0x6d) [0x7fe43b41f3dd]

**#4 - 03/26/2015 04:35 AM - Sage Weil**

- Backport set to hammer

this should go to hammer, too; the rgw test suite is triggering it:

ubuntu@teuthology:/a/sage-2015-03-24_11:15:21-rgw-hammer-distro-basic-multi/819683
ubuntu@teuthology:/a/sage-2015-03-24_11:15:21-rgw-hammer-distro-basic-multi/819604

11/25/2015
Assertion: osd/ReplicatedPG.cc: 2063: FAILED assert(op == prdop->op)
ceph version 0.93.203-g59143d1 (59143d141ccbd3e8bede87be0a13e8a56e52c32)
1: (ceph::__ceph_assert_fail(char const*, char const*, int, char const*)+0x7f) [0xaea18f]
2: (ReplicatedPG::finish_proxy_read(hobject_t, unsigned long, int)+0xf35) [0x806dd5]
3: (C_ProxyRead::finish(int)+0xc7) [0x8781b7]
4: (Context::complete(int)+0xa9) [0x68a849]
5: (Finisher::finisher_thread_entry()+0x160) [0xa16420]
6: (()+0x7e9a) [0x7f9115c5ce9a]
7: (clone()+0x6d) [0x7f911440631d]

---

**#6 - 03/29/2015 11:22 PM - Yuri Weinstein**

Jobs: ['826416', '826417', '826418', '826419']

Jobs: ['826416', '826417', '826418', '826419']

---

**#7 - 04/01/2015 06:55 AM - Kefu Chai**

the fix has not landed on hammer yet.

---

**#8 - 04/16/2015 03:51 PM - Loic Dachary**

- Severity changed from 3 - minor to 2 - major

---

**#9 - 04/21/2015 08:13 PM - Sage Weil**

- Status changed from Need Review to Pending Backport

---

**#10 - 04/24/2015 09:08 PM - Sage Weil**

- hammer backport [https://github.com/ceph/ceph/pull/4462](https://github.com/ceph/ceph/pull/4462)

---

**#11 - 04/29/2015 03:28 PM - Yuri Weinstein**

Jobs: ['866219', '866275']

---

**#12 - 05/06/2015 08:51 AM - Loic Dachary**

- Status changed from Pending Backport to Resolved
- Regression set to No

---

**#13 - 05/08/2015 06:45 PM - Yuri Weinstein**

See it in next
Job: ['878207']

Assertion: osd/ReplicatedPG.cc: 2072: FAILED assert(op == prdop->op)
#14 - 05/08/2015 06:47 PM - Yuri Weinstein
- Status changed from Resolved to New

#15 - 05/18/2015 09:00 AM - Loic Dachary
- Status changed from New to Verified


2015-05-17 16:09:19.260294 7f53284f9700 10 write_log with: dirty_to: 0'0, dirty_from: 4294967295'18446744073709551615, dirty_divergent_priors: 0, writeout_from: 28'10937, trimmed: 28'10936


ceph version 0.94.1-7-g293affe (293affe9921186e604f65030b263a794ca624)
1: (ceph::__ceph_assert_fail(char const*, char const*, int, char const*)+0x8b) [0xaebf4f]
2: (ReplicatedPG::finish_proxy_read(hobject_t, unsigned long, int)+0x3f5) [0x808d15]
3: (C_ProxyRead::finish(int)+0x9) [0x879997]
4: (Context::complete(int)+0x9) [0x86f151]
5: (Finisher::finisher_thread_entry()+0x158) [0xb5d938]
6: (clone()+0x6d) [0x7f5344a3ce9a]
7: (context()+0x6d) [0x7f53431e818d]

NOTE: a copy of the executable, or 'objdump -rdS <executable>' is needed to interpret this.
the backport is 0ee022b1ae832c70a80e9d2cdf32403039f3f125 which was not present in that failed run.