

Ceph - Backport #13335

hammer: OSD crashed when reached pool's max_bytes quota

10/02/2015 08:04 AM - Loic Dachary

Status: Resolved	Spent time: 0.00 hour
Priority: Normal	
Assignee: Alexey Sheplyakov	
Target version: v0.94.6	
Release: hammer	Crash signature:
Description https://github.com/ceph/ceph/pull/6918	
Related issues:	
Related to Ceph - Bug #15019: hammer: fs test fails with log [ERR] : OSD full...	Duplicate
Copied from Ceph - Bug #13098: OSD crashed when reached pool's max_bytes quota	Resolved 09/15/2015

Associated revisions

Revision 2817ffcf - 01/30/2016 09:13 AM - Alexey Sheplyakov

Check for full before changing the cached obc

ReplicatedPG::prepare_transaction(): check if the pool is full before updating the cached ObjectContext to avoid the discrepancy between the cached and the actual object size (and other metadata). While at it improve the check itself: consider cluster full flag, not just the pool full flag, also consider object count changes too, not just bytes.

Based on commit a1eb380c3d5254f9f1fe34b4629e51d77fe010c1

Fixes: #13335

Signed-off-by: Alexey Sheplyakov <asheplyakov@mirantis.com>

History

#1 - 01/11/2016 09:40 AM - Loic Dachary

- Description updated
- Status changed from New to In Progress
- Assignee set to Loic Dachary

#2 - 01/11/2016 10:20 AM - Alexey Sheplyakov

There's a fix here: <https://github.com/ceph/ceph/pull/6918>

#3 - 01/11/2016 01:25 PM - Loic Dachary

- Description updated
- Status changed from In Progress to New

- Assignee deleted (Loic Dachary)

This needs to be adapted to hammer because a few things are different (no flag in messages, CEPH_OSD_FLAG_FULL_FORCE not implemented).

#4 - 01/11/2016 02:03 PM - Alexey Sheplyakov

Basically I've moved the check for a full pool to the right place (before updating the cached ObjectContext) without changing the check itself (well, almost).

#5 - 01/29/2016 04:03 AM - Loic Dachary

- Status changed from New to In Progress

- Assignee set to Loic Dachary

#6 - 01/29/2016 04:03 AM - Loic Dachary

- Description updated

#7 - 02/03/2016 06:01 AM - Loic Dachary

- Status changed from In Progress to Resolved

- Target version set to v0.94.6

#8 - 03/09/2016 03:36 AM - Loic Dachary

- Status changed from Resolved to New

- Assignee deleted (Loic Dachary)

- Target version deleted (v0.94.6)

#9 - 03/09/2016 03:37 AM - Loic Dachary

The commit introduces a regression and is reverted by <http://tracker.ceph.com/issues/15019>

#10 - 03/09/2016 03:38 AM - Loic Dachary

- Related to Bug #15019: hammer: fs test fails with log [ERR] : OSD full dropping all updates 100% full added

#11 - 03/09/2016 08:38 AM - Alexey Sheplyakov

The commit introduces a regression

The commit exposes a bug in the test which assumes it's possible to write more data than the storage capacity is. I believe that OSD should reject such writes to prevent further damage (ENOSPC handling in filesystems' code is not 100% fool proof), and it does so in Infernalis and Jewel.

and is reverted by <http://tracker.ceph.com/issues/15019>

I don't think reverting it is a good idea, the test case itself should be fixed instead. Even if we want to pretend that it's possible to write 144 MB of data to a 100 MB drive the check should be slightly modified, that is, <https://github.com/ceph/ceph/blob/hammer/src/osd/ReplicatedPG.cc#L5693> should be removed, instead of reintroducing the obc corruption. However I think checking for a full OSD is actually correct.

#12 - 03/11/2016 04:53 AM - Loic Dachary

- Status changed from New to In Progress
- Assignee set to Alexey Sheplyakov

#13 - 03/11/2016 04:54 AM - Loic Dachary

- Subject changed from OSD crashed when reached pool's max_bytes quota to hammer: OSD crashed when reached pool's max_bytes quota

#14 - 04/06/2016 03:39 PM - Loic Dachary

- Status changed from In Progress to Resolved

#15 - 04/06/2016 03:40 PM - Loic Dachary

- Target version set to v0.94.6