Ceph - Bug #10787

mon: OSDMonitor::map_cache is buggy, send_incremental is not conservative

02/06/2015 10:54 PM - Sage Weil

Status: Resolved
Priority: Urgent
Assignee: Sage Weil
Category:
Target version:
Source: Q/A
Backport: giant, firefly
Regression:
Severity: 3 - minor

Start date: 02/06/2015
Due date:
% Done: 0%
Estimated time: 0.00 hour
Spent time: 0.00 hour
Reviewed:
Affected Versions:

Description

2015-02-06 22:29:31.387753 7fed116df700 10 mon.peon5752@0(leader).osd e598595 preprocess_pgtemp osd_pgtemp(e598595)

2015-02-06 22:29:31.387773 7fed116df700 7 mon.peon5752@0(leader).osd e598595 preprocess_pgtemp e598591 no changes from osd.807 [2607f298:4:2243:5522]:6828/3810

2015-02-06 22:29:31.387779 7fed116df700 7 mon.peon5752@0(leader).osd e598595 _reply_map 598591 from osd.807 [2607f298:4:2243:5522]:6828/3810

2015-02-06 22:29:31.387783 7fed116df700 5 mon.peon5752@0(leader).osd e598595 send_latest to osd.807 [2607f298:4:2243:5522]:6828/3810 start 598591


2015-02-06 22:29:31.387792 7fed116df700 10 mon.peon5752@0(leader).osd e598595 osd.807 should have epoch <b>594144</b>/</b>

2015-02-06 22:29:31.387794 7fed116df700 10 mon.peon5752@0(leader).osd e598595 build_incremental [594145..594245]

the passed-in first is > (much >!) the cached map value.

we should maybe only use the cached value here if it is larger, update the map_cache while we are here so that we do have the best value for this osd...

Associated revisions

Revision 86d38351 - 02/07/2015 08:37 PM - Sage Weil

mon/OSDMonitor: do not trust small values in osd epoch cache

If the epoch cache says the osd has epoch 100 and the osd is asking for epoch 200+, do not send it 100+.

Fixes: #10787
Backport: giant, firefly
Signed-off-by: Sage Weil <sage@redhat.com>
(cherry picked from commit 19fda32f4e414d3b7d449e5a1d9e465c1fd3a3b)
mon/OSDMonitor: do not trust small values in osd epoch cache

If the epoch cache says the osd has epoch 100 and the osd is asking for epoch 200+, do not send it 100+.

Fixes: #10787
Backport: giant, firefly
Signed-off-by: Sage Weil <sage@redhat.com>

(cherry picked from commit a5759e9b97107488a8508f36ad9ca1aba3fae07)

Revision 92c352d0 - 03/17/2015 05:08 PM - Sage Weil
mon/OSDMonitor: do not trust small values in osd epoch cache

If the epoch cache says the osd has epoch 100 and the osd is asking for epoch 200+, do not send it 100+.

Fixes: #10787
Backport: giant, firefly
Signed-off-by: Sage Weil <sage@redhat.com>
(cherry picked from commit a5759e9b97107488a8508f36ad9ca1aba3fae07)

History

#1 - 02/07/2015 02:17 PM - Sage Weil
  - Status changed from New to Need Review

#2 - 02/21/2015 06:09 PM - Sage Weil
  - Status changed from Need Review to Pending Backport

#3 - 03/17/2015 05:09 PM - Loic Dachary
  - giant backport https://github.com/ceph/ceph/pull/4049
#4 - 03/19/2015 12:21 AM - Loic Dachary
- Backport changed from giant,firefly to giant,firefly

#5 - 03/26/2015 02:14 PM - Loic Dachary
92c352d mon/OSDMonitor: do not trust small values in osd epoch cache (in giant), 27dbbb3 mon/OSDMonitor: do not trust small values in osd epoch cache (in firefly).

#6 - 03/26/2015 02:14 PM - Loic Dachary
- Status changed from Pending Backport to Resolved