

Ceph - Bug #10787

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

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

Status:	Resolved	Start date:	02/06/2015
Priority:	Urgent	Due date:	
Assignee:	Sage Weil	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:		Spent time:	0.00 hour
Source:	Q/A	Reviewed:	
Tags:		Affected Versions:	
Backport:	giant,firefly	ceph-qa-suite:	
Regression:	No	Pull request ID:	
Severity:	3 - minor		

Description

```
2015-02-06 22:29:31.387753 7fed116df700 10 mon.peon5752@0(leader).osd e598595 preprocess_pgtemp os
d_pgtemp(e598591 {5.7e99=[37,536,882],5.98a4=[482,738,869],5.f15b=[180,778,585],11.7e93=[37,536,88
2],11.989e=[482,738,869],11.f155=[180,778,585]} v598591) v1
2015-02-06 22:29:31.387773 7fed116df700 7 mon.peon5752@0(leader).osd e598595 preprocess_pgtemp e5
98591 no changes from osd.807 [2607:f298:4:2243::5522]:6828/3810
2015-02-06 22:29:31.387779 7fed116df700 7 mon.peon5752@0(leader).osd e598595 _reply_map 598591 fr
om osd.807 [2607:f298:4:2243::5522]:6828/3810
2015-02-06 22:29:31.387783 7fed116df700 5 mon.peon5752@0(leader).osd e598595 send_latest to osd.8
07 [2607:f298:4:2243::5522]:6828/3810 start 598591
2015-02-06 22:29:31.387787 7fed116df700 5 mon.peon5752@0(leader).osd e598595 send_incremental [59
8591..598595] to osd.807 [2607:f298:4:2243::5522]:6828/3810
2015-02-06 22:29:31.387792 7fed116df700 10 mon.peon5752@0(leader).osd e598595 osd.807 should have
epoch <b>594144</b>
2015-02-06 22:29:31.387794 7fed116df700 10 mon.peon5752@0(leader).osd e598595 build_incremental [5
94145..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 a5759e9b - 02/12/2015 09:49 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>

Revision 27dbbb3c - 02/27/2015 03:36 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 a5759e9b97107488a8508f36adf9ca1aba3fae07)

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 a5759e9b97107488a8508f36adf9ca1aba3fae07)

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, firefy* 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