

# Ceph - Bug #8753

## PG::activate assert failed when recover finished

07/07/2014 12:54 AM - Haomai Wang

<b>Status:</b>	Resolved	<b>% Done:</b>	0%
<b>Priority:</b>	Urgent	<b>Spent time:</b>	0.00 hour
<b>Assignee:</b>			
<b>Category:</b>			
<b>Target version:</b>			
<b>Source:</b>	other	<b>Affected Versions:</b>	
<b>Tags:</b>		<b>ceph-qa-suite:</b>	
<b>Backport:</b>	firefly,hammer	<b>Pull request ID:</b>	
<b>Regression:</b>	No	<b>Crash signature (v1):</b>	
<b>Severity:</b>	3 - minor	<b>Crash signature (v2):</b>	
<b>Reviewed:</b>			

### Description

```
60> 2014-06-28 14:19:50.353700 7f9a354d8700-0 10.1.16.72:6805/35913 >> 10.1.16.83:6812/1038272 pipe(0x3d389e00 sd=19
:50460 s=1 pgs=89400 cs=144 l=0 c=0x2c989080).fault
59> 2014-06-28 14:19:50.353727 7f9a196e9700-0 10.1.16.72:6805/35913 >> 10.1.16.71:6801/10910 pipe(0x2bc1400 sd=405
:39701 s=2 pgs=242288 cs=2061 l=0 c=0x2131ce0).fault, initiating reconnect
58> 2014-06-28 14:19:50.354165 7f9a19ef1700-0 10.1.16.72:6805/35913 >> 10.1.16.76:6801/9800 pipe(0x2bc3700 sd=82
:33045 s=2 pgs=467742 cs=995 l=0 c=0x2133020).fault, initiating reconnect
57> 2014-06-28 14:19:50.354190 7f9a06411700-0 10.1.16.72:6805/35913 >> 10.1.16.85:6801/29271 pipe(0x2c9ecb00 sd=86
:42948 s=2 pgs=21909 cs=123 l=0 c=0x39a86e0).fault, initiating reconnect
56> 2014-06-28 14:19:50.354260 7f9a037db700-0 10.1.16.72:6805/35913 >> 10.1.16.78:6801/4277 pipe(0x2bc2080 sd=96
:43756 s=2 pgs=224617 cs=1953 l=0 c=0x21323c0).fault, initiating reconnect
55> 2014-06-28 14:19:50.354362 7f9a19ff2700-0 10.1.16.72:6805/35913 >> 10.1.16.77:6801/4420 pipe(0x2bc1b80 sd=40 :40859
s=2 pgs=221587 cs=1961 l=0 c=0x2132100).fault, initiating reconnect
54> 2014-06-28 14:19:50.354553 7f9a182dc700-0 10.1.16.72:6805/35913 >> 10.1.16.74:6801/9452 pipe(0x2bc3980 sd=36
:37865 s=2 pgs=318367 cs=1769 l=0 c=0x2133180).fault, initiating reconnect
53> 2014-06-28 14:19:50.354667 7f9a19aed700-0 10.1.16.72:6805/35913 >> 10.1.16.85:6809/30200 pipe(0x360f9e00 sd=423
:37434 s=2 pgs=109303 cs=203 l=0 c=0x33756040).fault, initiating reconnect
52> 2014-06-28 14:19:50.354688 7f99eb65f700-0 10.1.16.72:6805/1035913 >> 10.1.16.85:6809/30200 pipe(0x360f9e00 sd=423
:37434 s=1 pgs=109303 cs=204 l=0 c=0x33756040).fault
51> 2014-06-28 14:19:50.354726 7f99eb0eb700-0 10.1.16.72:6808/1035913 >> 10.1.16.74:6805/20411 pipe(0x2bc1900 sd=31
:42298 s=2 pgs=251791 cs=1919 l=0 c=0x2131fa0).fault, initiating reconnect
50> 2014-06-28 14:19:50.356623 7f9a037db700-0 10.1.16.72:6808/1035913 >> 10.1.16.78:6801/4277 pipe(0x2bc2080 sd=96
:43764 s=2 pgs=224618 cs=1955 l=0 c=0x21323c0).fault, initiating reconnect
49> 2014-06-28 14:19:50.357017 7f99eb0eb700-0 10.1.16.72:6808/1035913 >> 10.1.16.74:6805/20411 pipe(0x2bc1900 sd=31
:42314 s=2 pgs=251792 cs=1921 l=0 c=0x2131fa0).fault, initiating reconnect
48> 2014-06-28 14:19:50.357924 7f9a037db700-0 10.1.16.72:6808/1035913 >> 10.1.16.78:6801/4277 pipe(0x2bc2080 sd=96
:43769 s=2 pgs=224619 cs=1957 l=0 c=0x21323c0).fault, initiating reconnect
47> 2014-06-28 14:19:50.358551 7f99eb0eb700-0 10.1.16.72:6808/1035913 >> 10.1.16.74:6805/20411 pipe(0x2bc1900 sd=31
:42316 s=2 pgs=251793 cs=1923 l=0 c=0x2131fa0).fault, initiating reconnect
46> 2014-06-28 14:19:50.358608 7f9a037db700-0 10.1.16.72:6808/1035913 >> 10.1.16.78:6801/4277 pipe(0x2bc2080 sd=96
:43771 s=2 pgs=224620 cs=1959 l=0 c=0x21323c0).fault, initiating reconnect
45> 2014-06-28 14:19:50.359146 7f9a037db700-0 10.1.16.72:6808/1035913 >> 10.1.16.78:6801/4277 pipe(0x2bc2080 sd=96
:43773 s=2 pgs=224621 cs=1961 l=0 c=0x21323c0).fault, initiating reconnect
44> 2014-06-28 14:19:50.359457 7f99eb0eb700-0 10.1.16.72:6808/1035913 >> 10.1.16.74:6805/20411 pipe(0x2bc1900 sd=31
:42318 s=2 pgs=251794 cs=1925 l=0 c=0x2131fa0).fault, initiating reconnect

37> 2014-06-28 14:19:55.592409 7f9a42514700-0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(ab40a209/rbd_data
.3633a55b527db.000000000000006fb/head/3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, d
ropping message 0x7c555a00

36> 2014-06-28 14:19:55.607802 7f9a42514700-0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(7664a209/rbd_data
.130186704ef0cd.000000000000005f/head/3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con,
```

dropping message 0x1c80c920  
~~35> 2014-06-28 14:19:55.725734 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(a025a209/rbd\_data.1afc06d2e45f7.0000000000000200/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x1dec8f00  
~~34> 2014-06-28 14:19:55.747289 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(3d88a209/rbd\_data.15950f7398a425.00000000000004020/head//3),PushReplyOp(142fa209/rbd\_data.14b4b934321f4c.000000000000020a/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x36d83a20  
~~33> 2014-06-28 14:19:55.773851 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.26b 21732 [PushReplyOp(40a3026b/rbd\_data.1a52c62d284e18.000000000000035f/head//3),PushReplyOp(de9a026b/rbd\_data.4651113639f46.0000000000000001/head//3)]) v1 remote, 10.1.16.85:6801/29271, failed lossy con, dropping message 0x34689680  
~~32> 2014-06-28 14:19:55.790342 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.26b 21732 [PushReplyOp(1b1b026b/rbd\_data.e70d318359ed.00000000000008cfd/head//3)]) v1 remote, 10.1.16.85:6801/29271, failed lossy con, dropping message 0x7c553c00  
~~31> 2014-06-28 14:19:55.790889 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(e0a6b209/rbd\_data.e70d318359ed.000000000000092a4/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x15013840  
~~30> 2014-06-28 14:19:55.805503 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(fda8b209/rbd\_data.16c8313d746e3a.0000000000000593/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x7c553de0  
~~29> 2014-06-28 14:19:55.805887 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.26b 21732 [PushReplyOp(d97f026b/rbd\_data.f1d966acb425a.000000000000018d/head//3)]) v1 remote, 10.1.16.85:6801/29271, failed lossy con, dropping message 0x1dec92c0  
~~28> 2014-06-28 14:19:55.820022 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(184bb209/rbd\_data.e70d318359ed.0000000000000af11/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x2f08ddc0  
~~27> 2014-06-28 14:19:55.822514 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.26b 21732 [PushReplyOp(f090126b/rbd\_data.185e8c6d010a57.00000000000001e6/head//3)]) v1 remote, 10.1.16.85:6801/29271, failed lossy con, dropping message 0x7c553a20  
~~26> 2014-06-28 14:19:55.833693 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(f35fb209/rbd\_data.e70d318359ed.0000000000000b4e4/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x2f088f00  
~~25> 2014-06-28 14:19:55.839213 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.26b 21732 [PushReplyOp(94a5126b/rbd\_data.c29316cd69c73.000000000000033b/head//3)]) v1 remote, 10.1.16.85:6801/29271, failed lossy con, dropping message 0x36d83fc0  
~~24> 2014-06-28 14:19:55.852903 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(bellc209/rbd\_data.19c14a493f9830.0000000000000152/head//3),PushReplyOp(4229c209/rbd\_data.16c82e7c0f15af.00000000000003dc/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x2f08ba20  
~~23> 2014-06-28 14:19:55.854557 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.26b 21732 [PushReplyOp(f19126b/rbd\_data.7ff757148d7aa.00000000000000c1/head//3)]) v1 remote, 10.1.16.85:6801/29271, failed lossy con, dropping message 0x15013660  
~~22> 2014-06-28 14:19:55.872031 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(b06ac209/rbd\_data.e70d318359ed.00000000000006175/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x2f08bde0  
~~21> 2014-06-28 14:19:55.890019 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(206dc209/rbd\_data.e70d318359ed.000000000000c716/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x2f08d280  
~~20> 2014-06-28 14:19:55.905227 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(4ea1d209/rbd\_data.7ff78248b1a8e.000000000000013b/head//3)]) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x2f08dfa0  
~~19> 2014-06-28 14:19:55.918277 7f9a42514700 0~~  
10.1.16.72:6808/1035913 submit\_message MOSDPGPushReply(3.209 21732 [PushReplyOp(82bbd209/rbd\_data

```
.167b1464ff7440.000000000000cb3/head//3)) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x2f089fe0
18> 2014-06-28 14:19:55.932649 7f9a42514700 0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(a471e209/rbd_data.138d436cf8a190.0000000000000469/head//3)) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x2f08e360
17> 2014-06-28 14:19:55.955793 7f9a42514700 0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(1f13e209/rbd_data.17d6132ae8944a.0000000000000323/head//3),PushReplyOp(ddc3e209/rbd_data.16c7da140bf204.00000000000009f5/head//3)) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x1deca00
16> 2014-06-28 14:19:55.969277 7f9a42514700 0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(f1d5e209/rbd_data.e70d318359ed.000000000000d7aa/head//3)) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x1deca20
15> 2014-06-28 14:19:55.982543 7f9a42514700 0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(d65ae209/rbd_data.e70d318359ed.000000000000ef7c/head//3)) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x1decb0c0
14> 2014-06-28 14:19:55.991566 7f9a42514700 0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(350be209/rbd_data.134d1f34638c7.0000000000000458/head//3)) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x1decb480
13> 2014-06-28 14:19:56.004707 7f9a42514700 0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(57cee209/rbd_data.e70d318359ed.0000000000004515/head//3)) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x1decb2a0
12> 2014-06-28 14:19:56.017954 7f9a42514700 0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(dd75f209/rbd_data.e70d318359ed.00000000000042be/head//3)) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0x1decb840
11> 2014-06-28 14:19:56.031307 7f9a42514700 0
10.1.16.72:6808/1035913 submit_message MOSDPGPushReply(3.209 21732 [PushReplyOp(f25af209/rbd_data.318165dfd3d68.00000000000000f/203//3)) v1 remote, 10.1.16.83:6812/1038272, failed lossy con, dropping message 0xf00bc00
```

```
0> 2014-06-28 14:20:02.071914 7f9a1f20c700 -1 osd/PG.cc: In function 'void PG::activate(ObjectStore::Transaction&, epoch_t, std::list<Context*, std::allocator<Context*>&> &, std::map<int, std::map<pg_t, pg_query_t, std::less<pg_t>&>, std::allocator<std::pair<const pg_t, pg_query_t>&>> &>, std::less<int>&>, std::allocator<std::pair<const int, std::map<pg_t, pg_query_t, std::less<pg_t>&>, std::allocator<std::pair<const pg_t, pg_query_t>&>> &>>> &>, std::map<int, std::vector<std::pair<pg_notify_t, std::map<unsigned int, pg_interval_t, std::less<unsigned int>&>, std::allocator<std::pair<const unsigned int, pg_interval_t>&>> &>>, std::allocator<std::pair<pg_notify_t, std::map<unsigned int, pg_interval_t, std::less<unsigned int>&>, std::allocator<std::pair<const unsigned int, pg_interval_t>&>> &>>, std::allocator<std::pair<pg_notify_t, std::map<unsigned int, pg_interval_t, std::less<unsigned int>&>, std::allocator<std::pair<const unsigned int, pg_interval_t>&>> &>>>>>*)' thread 7f9a1f20c700 time 2014-06-28 14:20:02.026644
osd/PG.cc: 1091: FAILED assert(callbacks_for_degraded_object.empty())
```

```
ceph version 0.67.5 (a60ac9194718083a4b6a225fc17cad6096c69bd1)
1: (PG::activate(ObjectStore::Transaction&, unsigned int, std::list<Context*, std::allocator<Context*>&> &, std::map<int, std::map<pg_t, pg_query_t, std::less<pg_t>&>, std::allocator<std::pair<pg_t const, pg_query_t>&>> &>, std::less<int>&>, std::allocator<std::pair<int const, std::map<pg_t, pg_query_t, std::less<pg_t>&>, std::allocator<std::pair<pg_t const, pg_query_t>&>> &>>> &>, std::map<int, std::vector<std::pair<pg_notify_t, std::map<unsigned int, pg_interval_t, std::less<unsigned int>&>, std::allocator<std::pair<unsigned int const, pg_interval_t>&>> &>>, std::allocator<std::pair<pg_notify_t, std::map<unsigned int, pg_interval_t, std::less<unsigned int>&>, std::allocator<std::pair<unsigned int const, pg_interval_t>&>> &>>>>, std::less<int>&>, std::allocator<std::pair<int const, std::vector<std::pair<pg_notify_t, std::map<unsigned int, pg_interva
```

```

l_t, std::less<unsigned int>;, std::allocator<std::pair<unsigned int const, pg_interval_t>>; > > >, std::allocator<std::pair<pg_notify_t, std::map<unsigned int, pg_interval_t, std::less<unsigned int>;, std::allocator<std::pair<unsigned int const, pg_interval_t>>>>>+0x2b0f) [0x6748af]
2: (PG::RecoveryState::Active::Active(boost::statechart::state<PG::RecoveryState::Active, PG::RecoveryState::Primary, PG::RecoveryState::Activating, (boost::statechart::history_mode)0>;::my_context)+0x383) [0x6753c3]
3: /usr/bin/ceph-osd() [0x6a3852]
4: /usr/bin/ceph-osd() [0x6a3e51]
5: /usr/bin/ceph-osd() [0x6a2c01]
6: /usr/bin/ceph-osd() [0x69e89f]
7: (PG::handle_peering_event(std::tr1::shared_ptr<PG::CephPeeringEvt>;, PG::RecoveryCtx)+0x31e) [0x6684ce]
8: (OSD::process_peering_events(std::list<PG*, std::allocator<PG*>; > const&, ThreadPool::TPHandle&)+0x2fc) [0x5dab5c]
9: /usr/bin/ceph-osd() [0x612b06]
10: (ThreadPool::worker(ThreadPool::WorkThread*)+0x551) [0x7956f1]
11: /usr/bin/ceph-osd() [0x798720]
12: (()+0x79d1) [0x7f9a5065a9d1]
13: (clone()+0x6d) [0x7f9a503a7b6d]
NOTE: a copy of the executable, or `objdump -rdS &lt;executable>` is needed to interpret this.

```

This happened in production env, from the callback. I guess why "callbacks\_for\_degraded\_object" isn't empty is that the leak object is in "callbacks\_for\_degraded\_object" but not in "waiting\_for\_degraded\_object". The object in "callbacks\_for\_degraded\_object" will be cleaned when iterator "waiting\_for\_degraded\_object". But is it possible that the object get true from "is\_degraded\_object()" and not in "waiting\_for\_degraded\_object" vector. So when cleaning up degraded objects from "waiting\_for\_degraded\_object", it may happen some objects only exists in "callbacks\_for\_degraded\_object" not in "waiting\_for\_degraded\_object".

## History

### #1 - 07/15/2014 01:49 PM - Samuel Just

- Priority changed from Normal to High

### #2 - 07/29/2014 01:27 PM - Samuel Just

Has this happened since?

### #3 - 08/26/2014 02:08 PM - Sage Weil

- Status changed from New to Can't reproduce

### #4 - 02/09/2015 06:31 PM - Samuel Just

- Status changed from Can't reproduce to 7

- Backport set to *dumpling,firefly,hammer*

callbacks\_for\_degraded\_object can have an entry not reflected in waiting\_for\_degraded\_object, and therefore can be not cleaned up in on\_change.

### #5 - 02/09/2015 06:31 PM - Samuel Just

- Priority changed from High to Urgent

### #6 - 02/11/2015 06:33 PM - Samuel Just

- Status changed from 7 to Pending Backport

### #7 - 03/11/2015 12:15 AM - Loic Dachary

- fixed in hammer <https://github.com/ceph/ceph/pull/3692>

**#8 - 03/11/2015 12:17 AM - Loïc Dachary**

- firefly backport <https://github.com/ceph/ceph/pull/3940>

**#9 - 03/26/2015 12:16 PM - Loïc Dachary**

- Backport changed from *dumpling,firefly,hammer* to *firefly,hammer*

it is not in the last dumpling release

**#10 - 05/06/2015 08:52 AM - Loïc Dachary**

- Status changed from *Pending Backport* to *Resolved*

- Regression set to *No*