

## RADOS - Bug #21287

### 1 PG down, OSD fails with "FAILED assert(i->prior\_version == last || i->is\_error())"

09/07/2017 08:14 AM - Henrik Korkuc

<b>Status:</b>	Duplicate	<b>Start date:</b>	09/07/2017
<b>Priority:</b>	High	<b>Due date:</b>	
<b>Assignee:</b>		<b>% Done:</b>	0%
<b>Category:</b>	EC Pools	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>		<b>Spent time:</b>	0.00 hour
<b>Source:</b>		<b>Reviewed:</b>	
<b>Tags:</b>		<b>Affected Versions:</b>	v12.2.0
<b>Backport:</b>		<b>ceph-qa-suite:</b>	
<b>Regression:</b>	No	<b>Component(RADOS):</b>	OSD
<b>Severity:</b>	3 - minor	<b>Pull request ID:</b>	
<b>Description</b>			
<p>One PG went down for me during large rebalance (I added racks to OSD placement, almost all data had to be shuffled). RBD EC pool, min_size was set to k as some PGs went inactive after rebalance/OSD restarts.</p> <p>OSD is failing with FAILED assert(i-&gt;prior_version == last    i-&gt;is_error()) during peering. This does not seem to be OSD specific, as I moved PG to another OSD, but it experienced same issue after marking original OSD lost. And for example shutting down some other OSDs of that PG enables startup of previously failing OSD. E.g. OSD 133 fails to start. Shutdown OSD 65. OSD 133 starts successfully. Start OSD 65, OSD 133, 118 (it has a copy of that pg from OSD 133) and 65 crash. Shutdown OSD 381, OSD 65 can start.</p> <p>I poster cluster log, OSD 133 log from yesterday and will paste OSD 65 log with more debugging in next message.</p> <p>ceph-post-file for cluster log: 0d3a2d70-cb27-4ade-b0d7-6ee6f69ffbc9 ceph-post-file for OSD 133: b48cce81-5571-4575-a626-daa9f43725d7</p>			
<b>Related issues:</b>			
Duplicates RADOS - Bug #22916: OSD crashing in peering		<b>Duplicate</b>	<b>02/05/2018</b>

## History

### #1 - 09/07/2017 09:01 AM - Henrik Korkuc

btw down pg is 1.1735.

Starting OSD 381 crashes 65, 133 and 118. Stopping 65 enables to start remaining OSDs, starting it crash 118 and 133.

To decrease debug points I removed 1.1735s0 from OSD 118 (it was copied from 133 there).

added debug\_osd 20/20 to config, started OSD 133 and after it booted, started OSD 65. Both of them crashed, logs posted.

OSD 133: c396e54f-f96e-4312-9a18-77ea5a6ed5c2

OSD 65: a76c4057-8acf-46c2-aa02-39528b98dd27

### #2 - 09/19/2017 08:20 AM - Henrik Korkuc

I had to delete affected pool to reclaim occupied space so I am unable to verify any fixes

### #3 - 11/15/2017 10:27 AM - mingxin liu

we hit this bug too, RBD+EC



#5 - 11/15/2017 10:32 AM - mingxin liu

liu mingxin wrote:

```
@{
"op": "modify",
"object": "1:4ec1cf44:::rbd_data.100119495cff.0000000000002897:head",
"version": "1624'39707",
"prior_version": "1624'39706",
"reqid": "client.77068.0:18102704",
"extra_reqids": [],
"mtime": "2017-11-12 03:44:55.752943",
"return_code": 0,
"mod_desc": {
"object_mod_desc": {
"can_local_rollback": true,
"rollback_info_completed": false,
"ops": [ {
"code": "SETATTRS",
"attrs": [
" ",
"hinfo_key",
"snapset"
]
}, {
"code": "ROLLBACK_EXTENTS",
"gen": 39707,
"snaps": "[536576,4096]"
}
]
}
}, {
"op": "modify",
"object": "1:4ec1cf44:::rbd_data.100119495cff.0000000000002897:head",
"version": "1624'39709",
"prior_version": "1624'39708",
"reqid": "client.86630.0:19763476",
"extra_reqids": [],
"mtime": "2017-11-12 03:44:55.774969",
"return_code": 0,
"mod_desc": {
"object_mod_desc": {
"can_local_rollback": true,
"rollback_info_completed": false,
"ops": [ {
"code": "SETATTRS",
"attrs": [
" ",
"hinfo_key",
"snapset"
]
}, {
"code": "ROLLBACK_EXTENTS",
"gen": 39709,
"snaps": "[540672,4096]"
}
]
}
}
}
@
```

the last two updates, it's divergent entries, version 39708 disappear. i guess op has applied disorderly

**#6 - 11/28/2017 06:48 PM - Shinobu Kinjo**

- Assignee set to *Shinobu Kinjo*

**#7 - 11/29/2017 05:15 PM - Greg Farnum**

- Category set to *EC Pools*

- Priority changed from *Normal* to *High*

**#8 - 11/30/2017 04:42 AM - Shinobu Kinjo**

- Assignee deleted (*Shinobu Kinjo*)

**#9 - 02/07/2018 02:37 AM - lingjie kong**

we hit this bug too in ec pool 2+1, i find one peer did not receive one piece of op message sended from primary osd, but the former and latter had been handled and added to the log, which may cause this assert.

**#10 - 02/07/2018 03:52 AM - Chang Liu**

see <https://github.com/ceph/ceph/pull/16675>

**#11 - 02/07/2018 11:35 AM - Chang Liu**

- Duplicates Bug #22916: *OSD crashing in peering added*

**#12 - 02/07/2018 09:40 PM - Josh Durgin**

- Status changed from *New* to *Duplicate*