

## Ceph - Bug #1370

### pgmon: PGMap::creating\_pgs is not checked on leader election

08/05/2011 05:09 PM - Greg Farnum

<b>Status:</b>	Resolved	<b>Start date:</b>	08/05/2011
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	Greg Farnum	<b>% Done:</b>	0%
<b>Category:</b>	Monitor	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	v0.34	<b>Spent time:</b>	3.50 hours
<b>Source:</b>		<b>Reviewed:</b>	
<b>Tags:</b>		<b>Affected Versions:</b>	
<b>Backport:</b>		<b>ceph-qa-suite:</b>	
<b>Regression:</b>	No	<b>Pull request ID:</b>	
<b>Severity:</b>	3 - minor		
<b>Description</b>			
So, state is lost when the leader changes. This needs a little bit of thought to handle leaders updating their maps properly.			

#### History

##### #1 - 08/05/2011 05:12 PM - Greg Farnum

- Status changed from New to In Progress

##### #2 - 08/08/2011 09:18 AM - Sage Weil

- Target version changed from v0.33 to v0.34

##### #3 - 08/08/2011 09:23 AM - Sage Weil

- translation missing: en.field\_position set to 31

##### #4 - 08/08/2011 01:53 PM - Greg Farnum

Ah, not quite what I thought it was. creating\_pgs is actually maintained via stat\_pg\_[add|sub], and it's done correctly. The problem is that creating\_pgs is only dealt with in register\_new\_pgs, which is only called via check\_osd\_map, which is only called by OSDMonitor::update\_from\_paxos. Which is called frequently, but backs out unless there's a new map it didn't have before. And if your cluster is stable, you can keep the same OSD map for a loooooong time.

Which meant that if you got a new leader election at the wrong time, you could have PG creations that are indefinitely delayed. To handle this, use the virtual on\_active() method and call check\_osd\_map (if you're the leader).

##### #5 - 08/08/2011 02:36 PM - Greg Farnum

- Subject changed from pgmon: PGMap::creating\_pgs is not serialized to disk to pgmon: PGMap::creating\_pgs is not checked on leader election

- Status changed from In Progress to Testing

##### #6 - 08/10/2011 04:53 PM - Greg Farnum

- Status changed from Testing to Resolved

Well, I haven't seen any more errors from here, so I guess I misread something or fixed it elsewhere. Merged all the pg create bits into master in [97a77c1cbc70af7bf266afe820f2fa1627684806](#)