

## Ceph - Bug #6528

### btrfs osd gets kicked out while removing large pg

10/11/2013 09:26 PM - Alexandre Oliva

<b>Status:</b>	Resolved	<b>% Done:</b>	0%
<b>Priority:</b>	Urgent	<b>Spent time:</b>	0.00 hour
<b>Assignee:</b>	Samuel Just		
<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>	dumping	<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>			
<b>Description</b>			
<p>btrfs isn't exactly blazingly fast for file removal; so much is well known to me. However, there are two issues at hand here that don't make sense to me.</p> <p>Context is a pg that finished being moved elsewhere, and the osd proceeds to remove its local copy. After a while, it starts showing OSD::disk_t heartbeat timeouts in its log. After a longer while, other osds kick it out of the cluster.</p> <p>The osd, however, does not die because of the timeout. It notifies that cluster that it was incorrectly marked as down, and keeps on removing the pg. However, it will only reconnect to the cluster after removal is complete. This is the first oddity; it should ideally succeed in reconnecting before the pg removal is complete (it can take a while to remove 100k files), but suicide would be fine too, so that a pid-based monitor can bring it back up.</p> <p>The other oddity is that, if I restart the osd manually when it's in this timed-out disk_t heartbeat state, it will often quickly rejoin the cluster and keep on removing the pg files, but without hitting the heartbeat timeouts any more! What gives? Couldn't it do whatever it does to remove the pg files after a restart without the restart, so as to avoid being kicked out in the first place?</p>			

#### Associated revisions

##### Revision c658258d - 10/15/2013 10:16 PM - Samuel Just

OSD: ping tphandle during pg removal

Fixes: #6528

Signed-off-by: Samuel Just <[sam.just@inktank.com](mailto:sam.just@inktank.com)>

Reviewed-by: Sage Weil <[sage@inktank.com](mailto:sage@inktank.com)>

##### Revision 24711cd4 - 03/04/2014 12:35 AM - Samuel Just

OSD: ping tphandle during pg removal

Fixes: #6528

Signed-off-by: Samuel Just <[sam.just@inktank.com](mailto:sam.just@inktank.com)>

Reviewed-by: Sage Weil <[sage@inktank.com](mailto:sage@inktank.com)>

(cherry picked from commit c658258d9e2f590054a30c0dee14a579a51bda8c)

Conflicts:

src/osd/OSD.cc

## History

---

### #1 - 10/15/2013 09:38 AM - Sage Weil

- Priority changed from Normal to High

### #2 - 10/15/2013 01:07 PM - Samuel Just

Ah, we don't ping the tphandle during that process, creating patch!

### #3 - 10/15/2013 01:07 PM - Samuel Just

- Status changed from New to In Progress

- Assignee set to Samuel Just

- Priority changed from High to Urgent

### #4 - 10/15/2013 01:09 PM - Samuel Just

- Backport set to dumpling

### #5 - 10/15/2013 01:33 PM - Sage Weil

- Status changed from In Progress to Fix Under Review

### #6 - 10/15/2013 04:00 PM - Sage Weil

- Status changed from Fix Under Review to Resolved