Ceph - Bug #345

OSD crash: PG::read_state

08/11/2010 12:24 AM - Wido den Hollander

Status:	Resolved	**% Done:**	0%
Priority:	Normal	**Spent time:**	0.00 hour
Assignee:			
Category:	OSD		
Target version:			
Source:		**Affected Versions:**	
Tags:		**ceph-qa-suite:**	
Backport:		**Pull request ID:**	
Regression:	No	**Crash signature (v1):**	
Severity:	3 - minor	**Crash signature (v2):**	
Reviewed:			

Description

This might be a duplicate of #279 but i'm not sure.

This morning i saw that 4 of my 12 OSD's were down (most of them killed by the OOM killer while i'm using tcmalloc).

Tried to start them again, but then osd5 crashed:

```
10.08.11_09:12:35.468138 7f17fb9d3720 osd5 2559 pg[0.18d( v 889'2876 lc 0'0 (889'2874,889'2876]+ba
cklog n=2758 ec=2 les=2497 2552/2552/879) [] r=0 (info mismatch, log(0'0,0'0]) mlcod 0'0 inactive]
 read_log 0~250978
10.08.11_09:12:35.468157 7f17fb9d3720 filestore(/srv/ceph/osd5) read /srv/ceph/osd5/current/meta/p
glog_0.18d_0 0~250978
10.08.11_09:12:35.468218 7f17fb9d3720 filestore(/srv/ceph/osd5) read couldn't open /srv/ceph/osd5/
current/meta/pglog_0.18d_0 errno 2 No such file or directory
10.08.11_09:12:35.468228 7f17fb9d3720 filestore(/srv/ceph/osd5) read /srv/ceph/osd5/current/meta/p
glog_0.18d_0 0~250978 = -2
10.08.11_09:12:35.468237 7f17fb9d3720 osd5 2559 pg[0.18d( v 889'2876 lc 0'0 (889'2874,889'2876]+ba
cklog n=2758 ec=2 les=2497 2552/2552/879) [] r=0 (info mismatch, log(889'2874,0'0]+backlog) (log b
ound mismatch, empty) mlcod 0'0 inactive] read_log got 0 bytes, expected 250978-0=250978
osd/PG.cc: In function 'void PG::read_log(ObjectStore*)':
osd/PG.cc:2168: FAILED assert(0)
 1: (PG::read_state(ObjectStore*)+0x846) [0x532746]
 2: (OSD::load_pgs()+0x145) [0x4e6b25]
 3: (OSD::init()+0x4b8) [0x4e7508]
 4: (main()+0x1d72) [0x458022]
 5: (__libc_start_main()+0xfd) [0x7f17fa295c4d]
 6: /usr/bin/cosd() [0x456099]
 NOTE: a copy of the executable, or `objdump -rdS <executable>` is needed to interpret this.
```

Like the log says, **/srv/ceph/osd5/current/meta/pglog_0.18d_0** is missing from the disk (fs is not full).

This looks like Christian Brunner's post on the ML (cosd dying after start), but i am not using the **rbd** branch on my client, i'm running the latest unstable (0eb6cd49f6e3ec523787d09cf08d3179be270db4).

Like mentioned on the ML, i tried a scrub, but that fails:

```
root@node14:~# ceph osd scrub 5
10.08.11_09:24:22.453954 mon <- [osd,scrub,5]
10.08.11_09:24:22.455788 mon0 -> 'unknown command scrub' (-22)
root@node14:~#
```

I've uploaded the core, logfile and binary to **logger.ceph.widodh.nl** in the directory **/srv/ceph/issues/cosd_crash_pg_read_state**

History

#1 - 08/11/2010 01:48 AM - Wido den Hollander

I just had the same crash on another osd. This OSD had some troubles with cephx, so i restarted it, then it crashed with the same message:

```
10.08.11_10:45:23.178767 7f4947d2e720 osd11 2820 pg[0.137( v 889'2760 lc 0'0 (889'2759,889'2760] n=2643 ec=2 l
es=2731 2764/2764/2764) [] r=0 (info mismatch, log(0'0,0'0]) mlcod 0'0 inactive] read_log 237510~3003
10.08.11_10:45:23.178790 7f4947d2e720 filestore(/srv/ceph/osd11) read /srv/ceph/osd11/current/meta/pglog_0.137
_0 237510~3003
10.08.11_10:45:23.178866 7f4947d2e720 filestore(/srv/ceph/osd11) read couldn't open /srv/ceph/osd11/current/me
ta/pglog_0.137_0 errno 2 No such file or directory
10.08.11_10:45:23.178879 7f4947d2e720 filestore(/srv/ceph/osd11) read /srv/ceph/osd11/current/meta/pglog_0.137
_0 237510~3003 = -2
10.08.11_10:45:23.178891 7f4947d2e720 osd11 2820 pg[0.137( v 889'2760 lc 0'0 (889'2759,889'2760] n=2643 ec=2 l
es=2731 2764/2764/2764) [] r=0 (info mismatch, log(889'2759,0'0]) (log bound mismatch, empty) mlcod 0'0 inacti
ve] read_log got 0 bytes, expected 240513-237510=3003
osd/PG.cc: In function 'void PG::read_log(ObjectStore*)':
osd/PG.cc:2168: FAILED assert(0)
 1: (PG::read_state(ObjectStore*)+0x846) [0x532746]
 2: (OSD::load_pgs()+0x145) [0x4e6b25]
 3: (OSD::init()+0x4b8) [0x4e7508]
 4: (main()+0x1d72) [0x458022]
 5: (__libc_start_main()+0xfd) [0x7f49465f0c4d]
 6: /usr/bin/cosd() [0x456099]
 NOTE: a copy of the executable, or `objdump -rdS <executable>` is needed to interpret this.
```

I added the log and coredump of osd11 to the same directory at **logger.ceph.widodh.nl**

#2 - 08/11/2010 01:57 AM - Wido den Hollander

Checked out the code, it seems you have to specify the full OSD name or * to scrub:

```
ceph osd scrub osd11
```

Or

```
ceph osd scrub '*'
```

Now my OSD's are scrubbing and i'll check if i can start the OSD's again.

#3 - 08/11/2010 10:06 AM - Sage Weil

fixed by [fd080d538e9594ed6203b20e2c65a91f5aaae2d4](#)

for any of these that aren't starting, just do 'rmdir /srv/ceph/osd$num/current/$badpgid', where $badpgid in the above case is 0.137, and start cosd again.

#4 - 08/11/2010 10:06 AM - Sage Weil

- Status changed from New to Resolved