Ceph - Bug #2866

osd: pg stuck with unfound

07/27/2012 04:29 PM - Sage Weil

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

Description

on congress, observed pg stuck with unfound objects. kicking peering (marking primary down) resolved it.

in testing #2860 fix, observed:

```
- osd A's pg peers, finds missing on osd B
 - osd B goes down.  now becomes unfound
 - osd B comes back up, still unfound.
```

log attached (pg 1.6). osd.2 comes back in epoch 31, but pg 1.6 doesn't notice. it checks for sources going down, but not down sources coming up.

Associated revisions

Revision 7b9d37c6 - 07/28/2012 05:05 PM - Sage Weil

osd: set STRAY on pg load when non-primary

The STRAY bit indicates that we should annouce ourselves to the primary,
but it is only set in start_peering_interval(). We also need to set it
initially, so that a PG that is loaded but whose role does not change
(e.g., the stray replica stays a stray) will notify the primary.

Observed:
- osd starts up
- mapping does not change, STRAY not set
- does not announce to primary
- primary does not re-check must_have_unfound, objects appear unfound

Fix this by initializing STRAY when pg is loaded or created whenever we
are not the primary.

Fixes: #2866
Signed-off-by: Sage Weil <sage@inktank.com>

Revision 9e5d4e61 - 07/30/2012 05:49 PM - Sage Weil

osd: initialize send_notify on pg load

When the PG is loaded, we need to set send_notify if we are not the
primary. Otherwise, if the PG does not go through
start_peering_interval() or experience a role change, we will not set
the flag and tell the primary that we exist. This can cause problems
for example if we have unfound objects that the primary needs, although
I'm sure there are other bad implications as well.

Fixes: #2866
Signed-off-by: Sage Weil <sage@inktank.com>

History

#1 - 07/27/2012 04:30 PM - Sage Weil

- File osd.1.log.gz added

#2 - 07/27/2012 05:18 PM - Sage Weil

- Status changed from 12 to Fix Under Review

#3 - 07/30/2012 10:52 AM - Sage Weil

- Status changed from Fix Under Review to Resolved

- Backport set to argonaut

9e5d4e61a73343397e67e918e87f1e6dcb8ec72d and 7b9d37c662313929b52011ddae47cc8abab99095

Files

osd.1.log.gz	15.6 MB	07/27/2012		Sage Weil