Ceph - Bug #11493
mon: adding existing pool as tier with --force-nonempty clobbers removed_snaps
04/28/2015 08:18 PM - Sage Weil

### Status: Resolved
### Start date: 04/28/2015
### Priority: Urgent
### Due date:
### Assignee: Sage Weil
### % Done: 0%
### Category:
### Estimated time: 0.00 hour
### Target version:
### Spent time: 0.00 hour
### Source: Community (user)
### Reviewed:
### Tags: firefly, hammer
### Affected Versions:
### Backport: ceph qa-suite:
### Regression: No
### Release: 
### Severity: 3 - minor

#### Description

1. images had snapshots (and removed_snaps)

2. images was added as a tier of img, and img's removed_snaps was copied to images, clobbering the removed_snaps value (see OSDMap::Incremental::propagate_snaps_to_tiers)

3. tiering relation was undone, but removed_snaps was still gone

4. on OSD startup, when we load the PG, removed_snaps is initialized with the older map. later, in PGPool::update(), we assume that removed_snaps always grows (never shrinks) and we trigger an assert.

To fix this I think we need to do 2 things:

1. make the OSD forgiving out removed_snaps getting smaller. This is probably a good thing anyway: once we know snaps are removed on all OSDs we can prune the interval_set in the OSDMap. Maybe.

2. Fix the mon to prevent this from happening, even when --force-nonempty is specified. (This is the root cause.)

#### Related issues:

- Related to Bug #7915: ./include/interval_set.h: 385: FAILED assert(_size >= 0) Duplicate 03/31/2014
- Copied to Backport #12312: mon: adding existing pool as tier with --force-none... Resolved 04/28/2015
- Copied to Backport #12314: mon: adding existing pool as tier with --force-none... Resolved 04/28/2015

#### Associated revisions

**Revision bbec53ed - 05/15/2015 05:39 PM - Sage Weil**

mon: prevent pool with snapshot state from being used as a tier

If we add a pool with snap state as a tier the snap state gets clobbered by OSDMap::Incremental::propagate_snaps_to_tiers(), and may prevent OSDs from starting. Disallow this.

Include a test.

Fixes: #11493

Backport: hammer, giant, firefly

Signed-off-by: Sage Weil <sage@redhat.com>
Revision 13c8d58d - 06/15/2015 07:25 PM - Sage Weil

mon: prevent pool with snapshot state from being used as a tier

If we add a pool with snapshot state as a tier the snap state gets clobbered by OSDMap::Incremental::propogate_snaps_to_tiers(), and may prevent OSDs from starting. Disallow this.

Include a test.

Fixes: #11493
Backport: hammer, giant, firefly
Signed-off-by: Sage Weil <sage@redhat.com>
(cherry picked from commit bbec53edf9e585af4e20bbca9057d6fda342)

Revision efe943ed - 07/07/2015 06:07 PM - Sage Weil

mon: prevent pool with snapshot state from being used as a tier

If we add a pool with snapshot state as a tier the snap state gets clobbered by OSDMap::Incremental::propogate_snaps_to_tiers(), and may prevent OSDs from starting. Disallow this.

Include a test.

Fixes: #11493
Backport: hammer, giant, firefly
Signed-off-by: Sage Weil <sage@redhat.com>
(cherry picked from commit bbec53edf9e585af4e20bbca9057d6fda342)

Revision d7f53442 - 10/05/2015 01:14 PM - Sage Weil

mon: prevent pool with snapshot state from being used as a tier

If we add a pool with snapshot state as a tier the snap state gets clobbered by OSDMap::Incremental::propogate_snaps_to_tiers(), and may prevent OSDs from starting. Disallow this.

Include a test.

Fixes: #11493
Backport: hammer, giant, firefly
Signed-off-by: Sage Weil <sage@redhat.com>
(cherry picked from commit bbec53edf9e585af4e20bbca9057d6fda342)

Conflicts:
qa/workunits/cephpool/test.sh
properly co-exist with "# make sure we can't create an ec pool tier"
src/moun/OSDMonitor.cc
properly co-exist with preceding "if (tp->ec_pool())"
(The changes to both files would have applied cleanly if
https://github.com/ceph/ceph/pull/5389 had not been merged first.)
History

#1 - 04/28/2015 08:19 PM - Sage Weil
- Description updated

#2 - 04/28/2015 09:37 PM - Sage Weil
- Status changed from New to Need Review

https://github.com/ceph/ceph/pull/4485

#3 - 05/29/2015 09:24 PM - Samuel Just
- Status changed from Need Review to Resolved
- Regression set to No

#4 - 07/13/2015 06:06 PM - Nathan Cutler
- Status changed from Resolved to Pending Backport
- Backport set to firefly, hammer

#5 - 07/13/2015 06:07 PM - Nathan Cutler
master commit: https://github.com/ceph/ceph/commit/bbec53edf9e585af4e20bbca9057b810fda342

#6 - 10/20/2015 07:31 PM - Loic Dachary
- Status changed from Pending Backport to Resolved