Ceph - Feature #13505

scrub/repair: persist scrub results.

10/16/2015 07:19 AM - Kefu Chai

**Status**: In Progress  
**Priority**: Normal  
**Assignee**: Kefu Chai  
**Category**:  
**Target version**: v10.0.1  
**Source**: other  
**Tags**: Affected Versions:  
**Backport**:  
**Reviewed**:  
**Start date**: 10/16/2015  
**Due date**:  
**% Done**: 0%  
**Estimated time**: 0.00 hour  
**Spent time**: 0.00 hour  
**User Impact**:  
**Target version**:  
**Spent time**:  
**Release**:  

**Description**

- write out temp object as scrub goes. with key of object name, value will present what's wrong with the object, object name => whats_wrong: inconsistency_t  
  inconsistency_t:  
  - most recent log version, prior version  
  - osd_id => shard_info_t  
  shard_info_t  
  - exists  
  - omap_sha1  
  - data_sha1  
  - size  
  - xattrs -> useronly  
  - missing on clone -> snapshot  
  - object_info_t  
  - data error?  
  - metadata error?  
- use pagination when querying the scrub result.  
- should always pass epic of the begin of the interval in the scrub APIs. if the epoch passes, EAGAIN is returned.

1. dump above metadata related to scrub/repair in the form of temp object, (they are already in the scrub map)  
2. add simple pg command to dump it  
3. add teuthology test accordingly

**Related issues:**

- Related to Feature #13506: scrub/repair: add librados APIs  
  - New  
  - 10/16/2015  
- Related to Feature #13507: scrub APIs to read replica  
  - New  
  - 10/16/2015  
- Related to Feature #13508: scrub/repair: repair corrupted/missing objects  
  - New  
  - 10/16/2015

**History**

#1 - 10/16/2015 07:20 AM - Kefu Chai

- Description updated

#2 - 10/16/2015 07:21 AM - Kefu Chai

- Description updated

#3 - 10/16/2015 07:25 AM - Kefu Chai

- Description updated
There are some scrub errors which are not related to a specific object or involve multiple objects.

1. The pg_stat_t (object_stat_sum_t) contains stats for the pg as a whole. Needs to be fixed last.
2. A missing SnapSet in a head object requires rebuilding the SnapSet or removing all clones. Are the clones in error or the head object?
3. A corruption of the clone_overlap requires clone_size to be repaired first. We could use a hierarchy of inconsistencies.

For the first stage of this change, we should worry about object data and omap inconsistencies keeping in mind some of these more complex error types will be handled later. For pg_stat_t we could just have repair run after the last object is repaired.

2. A missing SnapSet in a head object requires rebuilding the SnapSet or removing all clones. Are the clones in error or the head object?

they will be in the error.

- Target version set to v10.0.1