

Ceph - Bug #16474

OSD scrubs same PG over and over

06/24/2016 09:47 PM - Paul Emmerich

| | | | |
|------------------------|------------------|---------------------------|-----------------|
| Status: | Resolved | Start date: | 06/24/2016 |
| Priority: | Normal | Due date: | |
| Assignee: | David Zafman | % Done: | 0% |
| Category: | OSD | Estimated time: | 0.00 hour |
| Target version: | | Spent time: | 0.00 hour |
| Source: | Community (user) | Reviewed: | |
| Tags: | | Affected Versions: | v10.2.2, v9.2.2 |
| Backport: | jewel | ceph-qa-suite: | |
| Regression: | No | Pull request ID: | |
| Severity: | 3 - minor | | |

Description

Hi,

we've encountered a weird issue where some OSDs start scrubbing the same PG over and over while the other PGs on the OSD are simply never scrubbed, for example:

```
2016-06-24 23:27:47.503326 osd.73 [INF] 46.1a4 scrub ok
2016-06-24 23:27:50.476240 osd.73 [INF] 46.1a4 scrub starts
2016-06-24 23:27:50.511307 osd.73 [INF] 46.1a4 scrub ok
2016-06-24 23:27:51.475678 osd.73 [INF] 46.1a4 scrub starts
2016-06-24 23:27:51.513781 osd.73 [INF] 46.1a4 scrub ok
2016-06-24 23:27:52.476481 osd.73 [INF] 46.1a4 deep-scrub starts
2016-06-24 23:27:53.109623 osd.73 [INF] 46.1a4 deep-scrub ok
2016-06-24 23:27:54.476875 osd.73 [INF] 46.1a4 scrub starts
2016-06-24 23:27:54.507103 osd.73 [INF] 46.1a4 scrub ok
2016-06-24 23:27:59.476587 osd.73 [INF] 46.1a4 scrub starts
2016-06-24 23:27:59.510383 osd.73 [INF] 46.1a4 scrub ok
2016-06-24 23:28:06.476843 osd.73 [INF] 46.1a4 deep-scrub starts
2016-06-24 23:28:07.190836 osd.73 [INF] 46.1a4 deep-scrub ok
2016-06-24 23:28:07.477200 osd.73 [INF] 46.1a4 scrub starts
2016-06-24 23:28:07.508380 osd.73 [INF] 46.1a4 scrub ok
2016-06-24 23:28:08.477347 osd.73 [INF] 46.1a4 scrub starts
2016-06-24 23:28:08.512599 osd.73 [INF] 46.1a4 scrub ok
2016-06-24 23:28:17.478138 osd.73 [INF] 46.1a4 scrub starts
2016-06-24 23:28:17.511248 osd.73 [INF] 46.1a4 scrub ok
2016-06-24 23:28:20.478949 osd.73 [INF] 46.1a4 scrub starts
2016-06-24 23:28:20.511489 osd.73 [INF] 46.1a4 scrub ok
```

We can't see anything unusual with ceph pg query, the scrub timestamps are all current, the output of ceph pg query is attached at the end of this post

Other PGs on the same OSD also have a significantly older timestamp (beyond the max time between scrubs), yet they are completely ignored.

Restarting the affected OSDs fixes the problem.

Unfortunately, it is not reproducible and this only happened twice in three months. It seems to be related to the creation of new pools/PGs as it appeared twice within 24 hours of the creation of a new pool. The PGs being scrubbed all belong to the new pool.

We first encountered it with 9.2.2 a few month ago and today with 10.2.2.

pg info

```

{
  "state": "active+clean",
  "snap_trimq": "[]",
  "epoch": 151167,
  "up": [
    73,
    23,
    138
  ],
  "acting": [
    73,
    23,
    138
  ],
  "actingbackfill": [
    "23",
    "73",
    "138"
  ],
  "info": {
    "pgid": "46.1a4",
    "last_update": "151143'51",
    "last_complete": "151143'51",
    "log_tail": "0'0",
    "last_user_version": 51,
    "last_backfill": "MAX",
    "last_backfill_bitwise": 0,
    "purged_snaps": "[1~3]",
    "history": {
      "epoch_created": 150853,
      "last_epoch_started": 150955,
      "last_epoch_clean": 150955,
      "last_epoch_split": 0,
      "last_epoch_marked_full": 0,
      "same_up_since": 150854,
      "same_interval_since": 150954,
      "same_primary_since": 150828,
      "last_scrub": "151143'51",
      "last_scrub_stamp": "2016-06-24 23:39:02.606989",
      "last_deep_scrub": "151143'51",
      "last_deep_scrub_stamp": "2016-06-24 23:38:45.246075",
      "last_clean_scrub_stamp": "2016-06-24 23:39:02.606989"
    },
    "stats": {
      "version": "151143'51",
      "reported_seq": "22519",
      "reported_epoch": "151167",
      "state": "active+clean",
      "last_fresh": "2016-06-24 23:39:02.607169",
      "last_change": "2016-06-24 23:39:02.607169",
      "last_active": "2016-06-24 23:39:02.607169",
      "last_peered": "2016-06-24 23:39:02.607169",
      "last_clean": "2016-06-24 23:39:02.607169",
      "last_became_active": "2016-06-23 06:26:51.096878",
      "last_became_peered": "2016-06-23 06:26:51.096878",
      "last_unstale": "2016-06-24 23:39:02.607169",
      "last_undegraded": "2016-06-24 23:39:02.607169",
      "last_fullsized": "2016-06-24 23:39:02.607169",
      "mapping_epoch": 150853,
      "log_start": "0'0",
      "ondisk_log_start": "0'0",
      "created": 150853,
      "last_epoch_clean": 150955,
      "parent": "0.0",
      "parent_split_bits": 9,
    }
  }
}

```

```

"last_scrub": "151143'51",
"last_scrub_stamp": "2016-06-24 23:39:02.606989",
"last_deep_scrub": "151143'51",
"last_deep_scrub_stamp": "2016-06-24 23:38:45.246075",
"last_clean_scrub_stamp": "2016-06-24 23:39:02.606989",
"log_size": 51,
"ondisk_log_size": 51,
"stats_invalid": false,
"dirty_stats_invalid": false,
"omap_stats_invalid": false,
"hitset_stats_invalid": false,
"hitset_bytes_stats_invalid": false,
"pin_stats_invalid": false,
"stat_sum": {
  "num_bytes": 167772160,
  "num_objects": 40,
  "num_object_clones": 0,
  "num_object_copies": 120,
  "num_objects_missing_on_primary": 0,
  "num_objects_missing": 0,
  "num_objects_degraded": 0,
  "num_objects_misplaced": 0,
  "num_objects_unfound": 0,
  "num_objects_dirty": 40,
  "num_whiteouts": 0,
  "num_read": 203,
  "num_read_kb": 11872,
  "num_write": 0,
  "num_write_kb": 0,
  "num_scrub_errors": 0,
  "num_shallow_scrub_errors": 0,
  "num_deep_scrub_errors": 0,
  "num_objects_recovered": 0,
  "num_bytes_recovered": 0,
  "num_keys_recovered": 0,
  "num_objects_omap": 0,
  "num_objects_hit_set_archive": 0,
  "num_bytes_hit_set_archive": 0,
  "num_flush": 0,
  "num_flush_kb": 0,
  "num_evict": 0,
  "num_evict_kb": 0,
  "num_promote": 0,
  "num_flush_mode_high": 0,
  "num_flush_mode_low": 0,
  "num_evict_mode_some": 0,
  "num_evict_mode_full": 0,
  "num_objects_pinned": 0
},
"up": [
  73,
  23,
  138
],
"acting": [
  73,
  23,
  138
],
"blocked_by": [],
"up_primary": 73,
"acting_primary": 73
},
"empty": 0,
"dne": 0,
"incomplete": 0,
"last_epoch_started": 150955,

```

```

    "hit_set_history": {
      "current_last_update": "0'0",
      "history": []
    }
  },
  "peer_info": [
    {
      "peer": "23",
      "pgid": "46.1a4",
      "last_update": "151143'51",
      "last_complete": "151143'51",
      "log_tail": "0'0",
      "last_user_version": 0,
      "last_backfill": "MAX",
      "last_backfill_bitwise": 0,
      "purged_snaps": "[]",
      "history": {
        "epoch_created": 150853,
        "last_epoch_started": 150955,
        "last_epoch_clean": 150955,
        "last_epoch_split": 0,
        "last_epoch_marked_full": 0,
        "same_up_since": 150854,
        "same_interval_since": 150954,
        "same_primary_since": 150828,
        "last_scrub": "151143'51",
        "last_scrub_stamp": "2016-06-24 23:39:02.606989",
        "last_deep_scrub": "151143'51",
        "last_deep_scrub_stamp": "2016-06-24 23:38:45.246075",
        "last_clean_scrub_stamp": "2016-06-24 23:39:02.606989"
      },
      "stats": {
        "version": "0'0",
        "reported_seq": "16",
        "reported_epoch": "150906",
        "state": "peering",
        "last_fresh": "2016-06-23 06:22:48.093937",
        "last_change": "2016-06-23 06:22:48.088003",
        "last_active": "2016-06-23 06:18:33.658154",
        "last_peered": "2016-06-23 06:18:33.658154",
        "last_clean": "2016-06-23 06:18:33.658154",
        "last_became_active": "2016-06-23 06:18:33.658063",
        "last_became_peered": "2016-06-23 06:18:33.658063",
        "last_unstale": "2016-06-23 06:22:48.093937",
        "last_undegraded": "2016-06-23 06:22:48.093937",
        "last_fullsized": "2016-06-23 06:22:48.093937",
        "mapping_epoch": 150853,
        "log_start": "0'0",
        "ondisk_log_start": "0'0",
        "created": 150853,
        "last_epoch_clean": 150855,
        "parent": "0.0",
        "parent_split_bits": 9,
        "last_scrub": "0'0",
        "last_scrub_stamp": "2016-06-23 00:28:36.323398",
        "last_deep_scrub": "0'0",
        "last_deep_scrub_stamp": "2016-06-23 00:28:36.323398",
        "last_clean_scrub_stamp": "0.000000",
        "log_size": 0,
        "ondisk_log_size": 0,
        "stats_invalid": true,
        "dirty_stats_invalid": false,
        "omap_stats_invalid": false,
        "hitset_stats_invalid": false,
        "hitset_bytes_stats_invalid": false,
        "pin_stats_invalid": false,
        "stat_sum": {

```

```

        "num_bytes": 0,
        "num_objects": 0,
        "num_object_clones": 0,
        "num_object_copies": 0,
        "num_objects_missing_on_primary": 0,
        "num_objects_missing": 0,
        "num_objects_degraded": 0,
        "num_objects_misplaced": 0,
        "num_objects_unfound": 0,
        "num_objects_dirty": 0,
        "num_whiteouts": 0,
        "num_read": 0,
        "num_read_kb": 0,
        "num_write": 0,
        "num_write_kb": 0,
        "num_scrub_errors": 0,
        "num_shallow_scrub_errors": 0,
        "num_deep_scrub_errors": 0,
        "num_objects_recovered": 0,
        "num_bytes_recovered": 0,
        "num_keys_recovered": 0,
        "num_objects_omap": 0,
        "num_objects_hit_set_archive": 0,
        "num_bytes_hit_set_archive": 0,
        "num_flush": 0,
        "num_flush_kb": 0,
        "num_evict": 0,
        "num_evict_kb": 0,
        "num_promote": 0,
        "num_flush_mode_high": 0,
        "num_flush_mode_low": 0,
        "num_evict_mode_some": 0,
        "num_evict_mode_full": 0,
        "num_objects_pinned": 0
    },
    "up": [
        73,
        23,
        138
    ],
    "acting": [
        73,
        23,
        138
    ],
    "blocked_by": [],
    "up_primary": 73,
    "acting_primary": 73
},
"empty": 0,
"dne": 0,
"incomplete": 0,
"last_epoch_started": 150955,
"hit_set_history": {
    "current_last_update": "0'0",
    "history": []
}
},
{
    "peer": "138",
    "pgid": "46.1a4",
    "last_update": "151143'51",
    "last_complete": "151143'51",
    "log_tail": "0'0",
    "last_user_version": 0,
    "last_backfill": "MAX",
    "last_backfill_bitwise": 0,

```

```

"purged_snaps": "[]",
"history": {
  "epoch_created": 150853,
  "last_epoch_started": 150955,
  "last_epoch_clean": 150955,
  "last_epoch_split": 0,
  "last_epoch_marked_full": 0,
  "same_up_since": 150854,
  "same_interval_since": 150954,
  "same_primary_since": 150828,
  "last_scrub": "151143'51",
  "last_scrub_stamp": "2016-06-24 23:39:02.606989",
  "last_deep_scrub": "151143'51",
  "last_deep_scrub_stamp": "2016-06-24 23:38:45.246075",
  "last_clean_scrub_stamp": "2016-06-24 23:39:02.606989"
},
"stats": {
  "version": "0'0",
  "reported_seq": "16",
  "reported_epoch": "150906",
  "state": "peering",
  "last_fresh": "2016-06-23 06:22:48.093937",
  "last_change": "2016-06-23 06:22:48.088003",
  "last_active": "2016-06-23 06:18:33.658154",
  "last_peered": "2016-06-23 06:18:33.658154",
  "last_clean": "2016-06-23 06:18:33.658154",
  "last_became_active": "2016-06-23 06:18:33.658063",
  "last_became_peered": "2016-06-23 06:18:33.658063",
  "last_unstale": "2016-06-23 06:22:48.093937",
  "last_undegraded": "2016-06-23 06:22:48.093937",
  "last_fullsized": "2016-06-23 06:22:48.093937",
  "mapping_epoch": 150853,
  "log_start": "0'0",
  "ondisk_log_start": "0'0",
  "created": 150853,
  "last_epoch_clean": 150855,
  "parent": "0.0",
  "parent_split_bits": 9,
  "last_scrub": "0'0",
  "last_scrub_stamp": "2016-06-23 00:28:36.323398",
  "last_deep_scrub": "0'0",
  "last_deep_scrub_stamp": "2016-06-23 00:28:36.323398",
  "last_clean_scrub_stamp": "0.000000",
  "log_size": 0,
  "ondisk_log_size": 0,
  "stats_invalid": true,
  "dirty_stats_invalid": false,
  "omap_stats_invalid": false,
  "hitset_stats_invalid": false,
  "hitset_bytes_stats_invalid": false,
  "pin_stats_invalid": false,
  "stat_sum": {
    "num_bytes": 0,
    "num_objects": 0,
    "num_object_clones": 0,
    "num_object_copies": 0,
    "num_objects_missing_on_primary": 0,
    "num_objects_missing": 0,
    "num_objects_degraded": 0,
    "num_objects_misplaced": 0,
    "num_objects_unfound": 0,
    "num_objects_dirty": 0,
    "num_whiteouts": 0,
    "num_read": 0,
    "num_read_kb": 0,
    "num_write": 0,
    "num_write_kb": 0,

```

```

        "num_scrub_errors": 0,
        "num_shallow_scrub_errors": 0,
        "num_deep_scrub_errors": 0,
        "num_objects_recovered": 0,
        "num_bytes_recovered": 0,
        "num_keys_recovered": 0,
        "num_objects_omap": 0,
        "num_objects_hit_set_archive": 0,
        "num_bytes_hit_set_archive": 0,
        "num_flush": 0,
        "num_flush_kb": 0,
        "num_evict": 0,
        "num_evict_kb": 0,
        "num_promote": 0,
        "num_flush_mode_high": 0,
        "num_flush_mode_low": 0,
        "num_evict_mode_some": 0,
        "num_evict_mode_full": 0,
        "num_objects_pinned": 0
    },
    "up": [
        73,
        23,
        138
    ],
    "acting": [
        73,
        23,
        138
    ],
    "blocked_by": [],
    "up_primary": 73,
    "acting_primary": 73
},
"empty": 0,
"dne": 0,
"incomplete": 0,
"last_epoch_started": 150955,
"hit_set_history": {
    "current_last_update": "0'0",
    "history": []
}
}
],
"recovery_state": [
    {
        "name": "Started\\Primary\\Active",
        "enter_time": "2016-06-23 06:26:50.813267",
        "might_have_unfound": [],
        "recovery_progress": {
            "backfill_targets": [],
            "waiting_on_backfill": [],
            "last_backfill_started": "MIN",
            "backfill_info": {
                "begin": "MIN",
                "end": "MIN",
                "objects": []
            },
            "peer_backfill_info": [],
            "backfills_in_flight": [],
            "recovering": [],
            "pg_backend": {
                "pull_from_peer": [],
                "pushing": []
            }
        }
    },
    {
        "scrub": {

```

```

        "scrubber.epoch_start": "150954",
        "scrubber.active": 0,
        "scrubber.state": "INACTIVE",
        "scrubber.start": "MIN",
        "scrubber.end": "MIN",
        "scrubber.subset_last_update": "0'0",
        "scrubber.deep": false,
        "scrubber.seed": 0,
        "scrubber.waiting_on": 0,
        "scrubber.waiting_on_whom": []
    }
},
{
    "name": "Started",
    "enter_time": "2016-06-23 06:26:49.718199"
}
],
"agent_state": {}
}

```

Related issues:

Duplicated by Ceph - Bug #17361: pg scrubbing far too often (with logs)

Duplicate

09/21/2016

Copied to Ceph - Backport #17666: jewel: OSD scrubs same PG over and over

Resolved

History

#1 - 06/24/2016 10:00 PM - Paul Emmerich

- File *ceph-osd-scrub-log.txt.gz* added

The attached log are a few seconds from an OSD with debug osd = 20/20 injected while the bug is present. There are several scrubs of PG 46.1d1 in the log file.

#2 - 08/11/2016 03:20 PM - Robert Klikics

We are facing the same issue after doubling the number of PGs. Restart of all OSDs stopped this for now.

Example from PG 0.25a:

```

2016-08-11 08:51:24.120671 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 08:51:44.393595 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:03:44.204498 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:04:09.605254 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:06:39.226237 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:07:00.564817 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:08:32.238131 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:08:51.314839 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:11:51.261601 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:12:12.390471 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:18:05.300804 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:18:26.596946 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:29:15.378694 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:29:37.197375 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:30:13.385262 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:30:34.708098 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:34:04.411574 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:34:27.984084 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:35:26.420798 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:35:41.591375 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:37:28.432331 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:37:53.735487 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:42:43.468118 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:43:06.689449 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:43:14.470394 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:43:30.385951 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:44:32.479196 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:44:53.231835 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 09:44:53.481041 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:45:13.230173 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok

```



```
2016-08-11 09:47:26.499868 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 09:47:44.678449 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 10:00:13.590341 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 10:00:49.909739 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
2016-08-11 10:07:26.640703 7f92c7520700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub starts
2016-08-11 10:07:45.944285 7f92c9d25700 0 log_channel(cluster) log [INF] : 0.25a deep-scrub ok
```

#3 - 09/21/2016 01:50 PM - Warren Wang

We are also seeing this problem.

```
2016-09-21 13:19:00.417766 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub ok
2016-09-21 13:19:02.415674 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:02.421542 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:05.416005 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:05.419691 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:06.417942 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:06.424146 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:13.429745 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:08.424397 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:09.420088 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub starts
2016-09-21 13:19:09.423914 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub ok
2016-09-21 13:19:11.432099 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:11.437787 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:13.429745 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:13.433968 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:20.441754 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:20.448617 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:23.437290 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub starts
2016-09-21 13:19:23.439749 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub ok
2016-09-21 13:19:26.441915 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub starts
2016-09-21 13:19:26.449701 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub ok
2016-09-21 13:19:28.448141 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:28.451956 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:31.449108 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:31.453150 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:34.461895 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:34.471187 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:40.458086 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub starts
2016-09-21 13:19:40.461601 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub ok
2016-09-21 13:19:43.463643 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:43.468535 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:45.465415 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:45.471669 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:56.479685 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:19:56.482410 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:19:58.483394 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub starts
2016-09-21 13:19:58.495058 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub ok
2016-09-21 13:20:02.484480 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:20:02.488747 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:20:03.487368 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub starts
2016-09-21 13:20:03.493172 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub ok
2016-09-21 13:20:05.492350 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
2016-09-21 13:20:05.495008 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 scrub ok
2016-09-21 13:20:06.500451 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub starts
2016-09-21 13:20:06.510892 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub ok
2016-09-21 13:20:10.500003 7f704106c700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub starts
2016-09-21 13:20:10.507958 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 deep-scrub ok
2016-09-21 13:20:12.505763 7f7043871700 0 log_channel(cluster) log [INF] : 4.c6 scrub starts
```

ceph version 10.2.2-508-g9bfc0cf (9bfc0cf178dc21b0fe33e0ce3b90a18858abaf1b)

#4 - 09/22/2016 04:02 AM - David Zafman

- Status changed from New to Verified

- Assignee set to David Zafman

Two logs show that OSD::sched_scrub() is entered and the PG returned by service.first_scrub_stamp() is always the same PG. It appears that the PG is never removed from the ScrubJob set because the sched_time isn't advancing which is impossible if it were removed and re-added.

```
2016-09-21 13:25:44.603314 7fc0e051b700 20 osd.12 550 sched_scrub load_is_low=0
2016-09-21 13:25:44.603333 7fc0e051b700 10 osd.12 550 sched_scrub scrubbing 25.3f at 2016-09-14 15:01:22.01344
7 deadline < now
... SCRUBBING HERE ...
2016-09-21 13:25:44.608794 7fc0ccb5e700 10 osd.12 pg_epoch: 550 pg[25.3f( v 550'50326 (538'47314,550'50326) lo
cal-les=534 n=22 ec=292 les/c/f 534/538/0 528/529/378) [12,39,1] r=0 lpr=529 crt=550'50324 lcod 550'50325 mlco
d 550'50325 active+clean+scrubbing] scrub got 22/22 objects, 0/0 clones, 22/22 dirty, 22/22 omap, 0/0 pinned,
0/0 hit_set_archive, 0/0 bytes, 0/0 hit_set_archive bytes.
...
2016-09-21 13:25:47.603858 7fc0e051b700 20 osd.12 550 sched_scrub load_is_low=0
2016-09-21 13:25:47.603871 7fc0e051b700 10 osd.12 550 sched_scrub scrubbing 25.3f at 2016-09-14 15:01:22.01344
7 deadline < now
```

As far as I can tell unreg_next_scrub() and then reg_next_scrub() must be called in scrub_finish(). I see the log message from _scrub_finish() that it calls.

One explanation that would cause this is for is_primary() to be returning false. This would cause unreg_next_scrub() and reg_next_scrub() to do nothing. However, PG::sched_scrub() won't proceed if !is_primary(), so that would be impossible.

#5 - 09/22/2016 03:00 PM - Samuel Just

- Duplicated by Bug #17361: pg scrubbing far too often (with logs) added

#6 - 09/22/2016 03:02 PM - Samuel Just

Another instance came up on the list (subject: '[ceph-users] Same pg scrubbed over and over (Jewel)') Log available at <http://filedump.ceph.com/samuely/ceph-osd.12.log.bz2>.

#7 - 09/22/2016 06:53 PM - David Zafman

I easily reproduced this and found two entries for 0.c which is the pg repeating the scrubbing:

```
(gdb) print osd->sched_scrub_pg
$7 = std::set with 3 elements = {[0] = {pgid = {pgid = {m_pool = 0, m_seed = 12, m_preferred = -1, static calc
_name_buf_size = 36 '$'}, shard = {id = -1 '\377', static NO_SHARD = {id = -1 '\377',
    static NO_SHARD = <same as static member of an already seen type>}}, static calc_name_buf_size = 40
'('), sched_time = {tv = {tv_sec = 1474569250, tv_nsec = 335163700}}, deadline = {tv = {
```

```
    tv_sec = 1475173891, tv_nsec = 355625841}}, [1] = {pgid = {pgid = {m_pool = 0, m_seed = 4, m_preferre
d = -1, static calc_name_buf_size = 36 '$'}, shard = {id = -1 '\377', static NO_SHARD = {
    id = -1 '\377', static NO_SHARD = <same as static member of an already seen type>}}, static calc_nam
e_buf_size = 40 '('), sched_time = {tv = {tv_sec = 1474569266, tv_nsec = 535410882}},
    deadline = {tv = {tv_sec = 1475173891, tv_nsec = 359957349}}, [2] = {pgid = {pgid = {m_pool = 0, m_seed =
    12, m_preferred = -1, static calc_name_buf_size = 36 '$'}, shard = {id = -1 '\377',
    static NO_SHARD = {id = -1 '\377', static NO_SHARD = <same as static member of an already seen type>}}
, static calc_name_buf_size = 40 '('), sched_time = {tv = {tv_sec = 1474569571,
    tv_nsec = 149387499}}, deadline = {tv = {tv_sec = 1475174242, tv_nsec = 818689957}}}}
(gdb) print this->scrubber.scrub_reg_stamp
$8 = {tv = {tv_sec = 1474569571, tv_nsec = 149387499}}
```

#8 - 09/22/2016 10:55 PM - David Zafman

- Status changed from Verified to Need Test
- Backport set to jewel

<https://github.com/ceph/ceph/pull/11206>

#9 - 10/22/2016 01:17 AM - David Zafman

- Copied to Backport #17666: jewel: OSD scrubs same PG over and over added

#10 - 03/10/2017 01:47 AM - David Zafman

- Status changed from Need Test to Resolved

Files

| | | | |
|---------------------------|--------|------------|---------------|
| ceph-osd-scrub-log.txt.gz | 211 KB | 06/24/2016 | Paul Emmerich |
|---------------------------|--------|------------|---------------|