RADOS - Feature #40955

Extend the scrub sleep time when the period is outside [osd_scrub_begin_hour, osd_scrub_end_hour)

07/25/2019 12:29 PM - Jeegn Chen

<table>
<thead>
<tr>
<th>Status:</th>
<th>Resolved</th>
<th>% Done:</th>
<th>0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>Normal</td>
<td>Spent time:</td>
<td>0.00 hour</td>
</tr>
<tr>
<td>Assignee:</td>
<td>Jeegn Chen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target version:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affected Versions:</td>
<td>v12.2.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Component(RADOS):</td>
<td>OSD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backport:</td>
<td></td>
<td>Pull request ID:</td>
<td>29342</td>
</tr>
<tr>
<td>Reviewed:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Description

We already have osd_scrub_begin_week_day, osd_scrub_end_week_day, osd_scrub_begin_hour and osd_scrub_end_hour to tell OSD what period is proper to schedule scrubbing.

However, some scrub may happen to be scheduled just a short time before osd_scrub_begin_hour and cannot complete before osd_scrub_begin_hour, thus crosses the boundary of [osd_scrub_begin_hour, osd_scrub_end_hour).

In order to mitigate such impact of the scrubs, we may add a new option osd_scrub_extended_sleep (default is 0) and the logic is as follows:

1. when scrub_time_permit() return true, the existing osd_scrub_sleep is used
2. when scrub_time_permit() return false, there may be 2 scenarios
   2.1 if osd_scrub_extended_sleep <= osd_scrub_sleep, let's take osd_scrub_sleep
   2.2 otherwise, let's take osd_scrub_extended_sleep

So we can set a relatively large value (such as 10) to osd_scrub_extended_sleep. Thus the scrubs across the boundary of [osd_scrub_begin_hour, osd_scrub_end_hour) may be suppressed and the impact on client IO will be mitigated.

History

#1 - 07/26/2019 04:03 AM - Jeegn Chen
PR: https://github.com/ceph/ceph/pull/29342

#2 - 07/26/2019 06:43 AM - Kefu Chai
- Status changed from New to Fix Under Review
- Assignee set to Jeegn Chen
- Pull request ID set to 29342

#3 - 07/31/2019 03:00 AM - Jeegn Chen
One more scencario dzafman's comment:
@Jeegn-Chen Another way a scrub could happen even with scrub_time_permit() returns false would be when the scrub deadline is reached (scrub_max_interval). So a scrub could even start when it is not normally permitted if there hasn't been enough time to scrub even once during the scrub_min_interval due to load or time restrictions.

06/05/2021 1/2
Updated logic:

1. always take osd_scrub_sleep for manually initiated scrubs
2. when scrub_time_permit() return true for scheduled ones, the existing osd_scrub_sleep is used
3. when scrub_time_permit() return false for scheduled ones, there may be 2 scenarios
   3.1 if osd_scrub_extended_sleep <= osd_scrub_sleep, let's take osd_scrub_sleep
   3.2 otherwise, let's take osd_scrub_extended_sleep

#5 - 10/05/2019 03:14 AM - David Zafman

- Status changed from Fix Under Review to Resolved