

bluestore - Bug #38065

deep fsck fails on inspecting very large onodes

01/28/2019 02:21 PM - Igor Fedotov

| | |
|---|---------------------------|
| Status: Resolved | % Done: 0% |
| Priority: Normal | |
| Assignee: Igor Fedotov | |
| Category: | |
| Target version: | |
| Source: | Reviewed: |
| Tags: | Affected Versions: |
| Backport: mimic,luminous | ceph-qa-suite: |
| Regression: No | Pull request ID: |
| Severity: 3 - minor | Crash signature: |
| Description Steps to reproduce (100%): Put 3GB object to replicated pool via rados then do deep fsck either on mount or via ceph-bluestore-tool. Following output is provided: 2019-01-28 17:17:55.960 7f5d42af2b40 -1 bdev(0x557b43254700 dev/osd0/block) aio_submit retries 5423 2019-01-28 17:17:57.836 7f5d42af2b40 -1 bluestore(dev/osd0) fsck error: #4:238134c4:::tmp_3g:head# error during read: (1073741824) Unknown error 1073741824 | |
| Related issues: Copied to bluestore - Backport #38187: mimic: deep fsck fails on inspecting v... Resolved Copied to bluestore - Backport #38188: luminous: deep fsck fails on inspectin... Resolved | |

History

#1 - 01/28/2019 03:16 PM - Igor Fedotov

Looks like aio_queue_t::submit_batch timeouts due to long list of blocks to read. We need to cap this for fsck and, probably, introduce a sort of throttling for regular reads...

#2 - 01/28/2019 03:17 PM - Igor Fedotov

- Subject changed from deep fsck fails on inspecting objects >2Gb length to deep fsck fails on inspecting objects very large onodes

#3 - 01/28/2019 03:25 PM - Igor Fedotov

- Subject changed from deep fsck fails on inspecting objects very large onodes to deep fsck fails on inspecting very large onodes

#4 - 01/28/2019 03:27 PM - Igor Fedotov

<https://github.com/ceph/ceph/pull/26170>

#5 - 01/28/2019 03:28 PM - Igor Fedotov

- Status changed from New to Fix Under Review

#6 - 01/28/2019 04:08 PM - Sage Weil

- Backport set to mimic,luminous

#7 - 02/04/2019 08:51 PM - Yuri Weinstein

Igor Fedotov wrote:

<https://github.com/ceph/ceph/pull/26170>

merged

#8 - 02/04/2019 08:55 PM - Neha Ojha

- Status changed from *Fix Under Review* to *Pending Backport*

#9 - 02/05/2019 05:02 PM - Nathan Cutler

- Copied to Backport #38187: *mimic: deep fsck fails on inspecting very large onodes added*

#10 - 02/05/2019 05:02 PM - Nathan Cutler

- Copied to Backport #38188: *luminous: deep fsck fails on inspecting very large onodes added*

#11 - 04/04/2019 10:22 AM - Igor Fedotov

- Status changed from *Pending Backport* to *Resolved*