

bluestore - Bug #36268

Unable to recover from ENOSPC in BlueFS

10/01/2018 01:58 PM - Igor Fedotov

Status: Resolved	Start date: 10/01/2018
Priority: Normal	Due date:
Assignee: Igor Fedotov	% Done: 0%
Category:	Estimated time: 0.00 hour
Target version:	
Source:	Reviewed:
Tags:	Affected Versions:
Backport:	ceph-qa-suite:
Regression: No	Pull request ID:
Severity: 3 - minor	Crash signature:
Description Under heavy load and full DB volume BlueStore might fall into the state where it lacks additional space for BlueFS even if the space is still available at block device. This is caused by the "lazy" behavior of free space rebalancing - it happens periodically in background rather than on demand. On the first allocation failure OSD asserts and then is unable to restart since log replay during BlueFS open needs the space as well but rebalance is still not executed. Then assertion again and hence getting a sort of unrecoverable deadlock for OSD.	
Related issues: Copied to bluestore - Backport #36640: luminous: Unable to recover from ENOSP... New Copied to bluestore - Backport #36641: mimic: Unable to recover from ENOSPC i... New	

History

#1 - 10/05/2018 09:19 PM - Yuri Weinstein

<https://github.com/ceph/ceph/pull/24352> merged

#2 - 10/30/2018 02:44 PM - Sage Weil

- Status changed from New to Pending Backport

- Backport set to mimic,luminous

also <https://github.com/ceph/ceph/pull/23103>

#3 - 10/30/2018 05:15 PM - Patrick Donnelly

- Copied to Backport #36640: luminous: Unable to recover from ENOSPC in BlueFS added

#4 - 10/30/2018 05:15 PM - Patrick Donnelly

- Copied to Backport #36641: mimic: Unable to recover from ENOSPC in BlueFS added

#5 - 11/13/2018 10:36 AM - Igor Fedotov

- Status changed from Pending Backport to In Progress

In fact previously mentioned PR is just a workaround to be able to manually fix the issue.
Working on the actual solution to fix BlueFS allocation strategy.

#6 - 11/29/2018 03:54 PM - Igor Fedotov

- Status changed from In Progress to Need Review

<https://github.com/ceph/ceph/pull/25132>

#7 - 02/28/2019 03:30 PM - Sage Weil

- Status changed from *Need Review* to *Resolved*

#8 - 03/01/2019 03:09 PM - Nathan Cutler

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

Sage, did you mean to cancel the mimic and luminous backports when you changed the status to Resolved?

#9 - 03/04/2019 03:23 PM - Sage Weil

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

- Backport deleted (*mimic,luminous*)

Alternative fix for mimic and luminous: <https://github.com/ceph/ceph/pull/26735>

#10 - 05/15/2019 09:25 AM -

Sage Weil wrote:

Alternative fix for mimic and luminous: <https://github.com/ceph/ceph/pull/26735>

hello sage weil , i have meet the same issue before in Luminous and i have merged the new patch you mentioned, but it is unuseful. restart the osd have the same assert.Is there any other way to restore OSD such as clean up bluefs size or expand bluefs size