RADOS - Bug #40620

Explicitly requested repair of an inconsistent PG cannot be scheduled timely on a OSD with ongoing recovery

07/02/2019 01:33 PM - Jeegn Chen

Status:	Resolved		**% Done:**	0%
Priority:	Normal		**Spent time:**	0.00 hour
Assignee:				
Category:				
Target version:				
Source:			**Affected Versions:**	v12.2.8
Tags:			**ceph-qa-suite:**	
Backport:	nautilus		**Component(RADOS):**	OSD
Regression:	No		**Pull request ID:**	28839
Severity:	3 - minor		**Crash signature (v1):**	
Reviewed:			**Crash signature (v2):**	

Description

Since osd_scrub_during_recovery=false is used as default, when a OSD has some recovering PG, it will not schedule any new scrub including explicitly request repair. Thus inconsistent data cannot be fixed in time, which is not good for data safety.

The proposal is that we introduce a new config option osd_repair_during_recovery, whose default value is false:

- When osd_scrub_during_recovery is true, ignore osd_repair_during_recovery (no behavior change)
- When osd_scrub_during_recovery is false and osd_repair_during_recovery is false, no behavior change
- When osd_scrub_during_recovery is false and osd_repair_during_recovery is true, we would allow `OSD::sched_scrub()` to schedule explicitly request repair (scrubber.must_repair=true)

Related issues:

Copied to RADOS - Backport #40840: nautilus: Explicitly requested repair of a...	**Resolved**

History

#1 - 07/03/2019 01:45 AM - Jeegn Chen

PR: https://github.com/ceph/ceph/pull/28839

#2 - 07/03/2019 09:29 PM - Neha Ojha

- Status changed from New to Fix Under Review

#3 - 07/03/2019 11:19 PM - David Zafman

- Pull request ID set to 28839

#4 - 07/16/2019 03:08 PM - Sage Weil

- Status changed from Fix Under Review to Pending Backport

- Backport set to nautilus

#5 - 07/22/2019 08:20 AM - Nathan Cutler

- Copied to Backport #40840: nautilus: Explicitly requested repair of an inconsistent PG cannot be scheduled timely on a OSD with ongoing recovery added

#6 - 11/14/2019 03:51 PM - Nathan Cutler

- Status changed from Pending Backport to Resolved

While running with --resolve-parent, the script "backport-create-issue" noticed that all backports of this issue are in status "Resolved" or "Rejected".