

RADOS - Bug #40622

PG stuck in active+clean+remapped

07/02/2019 02:46 PM - Mike Almateia

Status:	Resolved	Start date:	07/02/2019
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:	Backfill/Recovery	Estimated time:	0.00 hour
Target version:	v12.2.12	Spent time:	0.00 hour
Source:		Affected Versions:	
Tags:		ceph-qa-suite:	
Backport:		Component(RADOS):	
Regression:	No	Pull request ID:	
Severity:	3 - minor	Crash signature:	
Reviewed:			

Description

A cluster have 6 servers, in 3 racks, 2 servers per a rack.

A replication rule distributes replicas to the 3 racks: one replica per a rack

We start removing one server in each rack: all replicas must move to the remaining server in each rack.

In the second rack, the second server has two DSOs less than the one being removed from the cluster.

When moving data from server in a second rack, 1 PG stuck at active+clean+remapped status: apparently can not find the desired OSD for moving inside second rack.

I'm try use:

- ceph osd out 21
- ceph crush reweight osd.21 0

but the same PG (id 5.783) stuck in active+clean+remapped status.

I have mon_max_pg_per_osd=400 set up, it's can be a barrier.

History

#1 - 07/03/2019 09:30 PM - Sage Weil

- Status changed from New to Resolved

This looks like crush is just failing to find a good replica because 50% of the osds in a rack are down. Try using the optimal tunables.. if you already are using those, try increasing choose_tries to something larger than 50 (70?) and the pgs will probably go active.

Files

ceph_osd_df_tree	7.5 KB	07/02/2019	Mike Almateia
replicated_rule	417 Bytes	07/02/2019	Mike Almateia
crush_dump	35.1 KB	07/02/2019	Mike Almateia