Ceph - Backport #17909

jewel: ReplicatedBackend::build_push_op: add a second config to limit omap entries/chunk independently of object data

11/15/2016 09:46 AM - Alexey Sheplyakov

Status:	Resolved	**Spent time:**	0.00 hour
Priority:	Normal		
Assignee:	Alexey Sheplyakov		
Target version:	v10.2.6		
Release:	jewel	**Crash signature (v2):**	
Crash signature (v1):			

Description

https://github.com/ceph/ceph/pull/11991

build_push_op assumes 8MB of omap entries is about as much work to read as 8MB of object data. This is probably false. Add a config (osd_recovery_max_omap_entries_per_chunk ?) with a sane default (50k?) and change build_push_op to use it.

Related issues:

Copied from Ceph - Feature #16128: ReplicatedBackend::build_push_op: add a se...	**Resolved**	**06/02/2016**

History

#1 - 11/15/2016 09:46 AM - Alexey Sheplyakov

- Copied from Feature #16128: ReplicatedBackend::build_push_op: add a second config to limit omap entries/chunk independently of object data added

#2 - 11/15/2016 09:53 AM - Alexey Sheplyakov

In ceph deployments with large numbers of objects (typically generated by use of radosgw for object storage) it's quite possible for OSDs recovering data to hit their suicide timeout and shutdown because of the number of objects each was trying to recover in a single chunk between heartbeats.
https://bugs.launchpad.net/ubuntu/+source/ceph/+bug/1628750/comments/0

#3 - 11/15/2016 10:01 AM - Alexey Sheplyakov

https://github.com/ceph/ceph/pull/11991

#4 - 11/15/2016 04:48 PM - Abhishek Varshney

- Description updated

- Status changed from New to In Progress

- Assignee set to Alexey Sheplyakov

#5 - 02/01/2017 11:26 AM - Nathan Cutler

- Status changed from In Progress to Resolved

- Target version set to v10.2.6