ceph-volume - Bug #24152

Starting all OSDs created by ceph-volume via systemd fails

05/16/2018 10:08 PM - Andras Pataki

Status:	Resolved	**% Done:**	0%
Priority:	Normal		
Assignee:			
Category:			
Target version:			
Source:	Community (user)	**Affected Versions:**	v12.2.5
Tags:		**ceph-qa-suite:**	
Backport:		**Pull request ID:**	
Regression:	No	**Crash signature (v1):**	
Severity:	3 - minor	**Crash signature (v2):**	
Reviewed:			

Description

Currently when ceph-volume creates a new OSD, it creates a systemd unit instance in /etc/systemd/system/multi-user.target.wants. This makes the OSD start at boot, but it is inconvenient to start all OSDs on the same host (for upgrades without reboots for example).

Perhaps it would be preferable to move the ceph-volume@.service to live under ceph-osd.target instead. That way systemctl start ceph-osd.target would start all OSDs (which currently does not work), and systemctl stop ceph-osd.target stops them (which currently does work).

When that is done, however, a different problem rears its head. When running systemctl start 'ceph-osd.target', it executes the unit(s) ceph-volume@OSDID-OSDUUID.service, which in turn run ceph-volume lvm trigger OSDID-OSDUUID. That command internally runs activate, which, among other things runs something like
systemctl enable ceph-volume@OSDID-OSDUUID
(and then starts the OSD). This means that the unit that is being run is trying to activate itself. In CentOS, this causes multiple invocations of the systemctl start ceph-osd.target command to hang and not return (even though the OSDs start correctly).

It would be preferable not to test the edge cases of systemd, and not have the unit that runs try to activate itself.

For example, in systemd.py replacing

def enable(unit):
process.run(['systemctl', 'enable', unit])

with

def enable(unit):
stdout, stderr, retcode = process.call(['systemctl', 'is-enabled',
unit], show_command=True)
if retcode != 0:
process.run(['systemctl', 'enable', unit])

is one way to resolve this.

Or perhaps some flag to activate to distinguish the cases when it is called for creating an osd (ceph-volume activate) vs. by systemd (ceph-volume trigger).

History

#1 - 07/30/2018 01:12 PM - Dan van der Ster

Does this cover your use-case? https://github.com/ceph/ceph/pull/23321

#2 - 08/03/2018 12:08 PM - Alfredo Deza

- Status changed from New to Resolved

Mimic PR https://github.com/ceph/ceph/pull/23393
Lumimnous PR https://github.com/ceph/ceph/pull/23394