

RADOS - Bug #24368

osd: should not restart on permanent failures

05/31/2018 11:17 PM - Greg Farnum

Status:	Resolved	% Done:	0%
Priority:	Normal	Spent time:	0.00 hour
Assignee:	Greg Farnum		
Category:	Administration/Usability		
Target version:			
Source:	Development	Affected Versions:	
Tags:		ceph-qa-suite:	
Backport:		Component(RADOS):	OSD
Regression:	No	Pull request ID:	
Severity:	3 - minor	Crash signature (v1):	
Reviewed:		Crash signature (v2):	

Description

Last week at OpenStack I heard a few users report OSDs were not failing hard and fast as they should be on disk issues. For some of them, there were definitely multiple causes. But one of the easy ones is that systemd (especially as we configure it) tries to keep services running, so when an OSD crashes it gets restarted and tries to rejoin the cluster.

There are two different approaches to take here:

- 1) Modify how frequently systemd can restart the service. (Changing the StartLimitInterval and StartLimitBurst values)
- 2) Modify in what cases systemd restarts the service. It turns out you can configure varying combinations of the ways for a process to exit to behave differently (in systemd: "Clean exit code or signal", "Unclean exit code", "Unclean signal", "Timeout", "Watchdog" are handled differently in the 6 options for when to restart on exist), AND you can specify that the service shouldn't restart on specific return values or signals. I'm not sure if our exit statuses are distinct enough for that to be useful right now, but we can definitely get there!

History

#1 - 05/31/2018 11:25 PM - Greg Farnum

See <https://www.freedesktop.org/software/systemd/man/systemd.service.html#Restart=> for the details on Restart options.

#2 - 05/31/2018 11:34 PM - Greg Farnum

- Status changed from New to In Progress
- Assignee set to Greg Farnum

<https://github.com/ceph/ceph/pull/22349> has the simple restart interval change. Will investigate the options for conditionally limiting restarts.

#3 - 06/01/2018 06:01 PM - Nathan Cutler

- Backport set to mimic, luminous

Sounds like something that would be useful in our stable releases - Greg, do you agree?

#4 - 06/04/2018 06:32 PM - Greg Farnum

It would, but the previous settings were there for a reason so I'm not sure if it's feasible to backport this for ceph-disk users, or if they'd hit the startup race discussed in that PR commit.

Although maybe hitting a startup race is better than OSDs taking forever to get kicked out of the cluster, anyway?
Planning and need to investigate it more. :)

#5 - 07/09/2018 08:12 AM - guotao Yao

I've had a similar problem recently. One OSD crash and exit, and the OSD process starts up quickly by systemd. It causes the OSD to flapping up and down. Many PGs are always in the peering state. The client I/O drops sharply.

I found that this was caused by the systemd service. and I found the options `StartLimitInterval` and `StartLimitBurst` limit the number of service restarts. I counted the time between service startup and crash exit from the OSD log. It restart twice a minute in my scenario, and I set the `StartLimitInterval` is 2 minutes, and the `StartLimitBurst` is 2.

I also want to know how the Ceph community solves this issue.

#6 - 07/09/2018 08:42 AM - guotao Yao

guotao Yao wrote:

I've had a similar problem recently. One OSD crash and exit, and the OSD process starts up quickly by systemd. It causes the OSD to flapping up and down. Many PGs are always in the peering state. The client I/O drops sharply.

I found that this was caused by the systemd service. and I found the options `StartLimitInterval` and `StartLimitBurst` limit the number of service restarts. I counted the time between service startup and crash exit from the OSD log. It restart twice a minute in my scenario, and I set the `StartLimitInterval` is 2 minutes, and the `StartLimitBurst` is 2.

I also want to know how the Ceph community solves this issue.

In addition, I didn't change the `RestartSec` parameter, but I kept it for 20 seconds.

#7 - 07/09/2018 04:09 PM - Greg Farnum

I don't think the issue has moved beyond the PR linked above to change the systemd settings. I sent this out to one or two large users and was hoping to get some reports back on how it worked before doing any backports.

I looked briefly at if we could somewhat easily tell systemd how it should behave by tweaking our exit codes or signals, but it didn't seem like a short project, so I let it fall. :(

#8 - 10/18/2018 04:02 PM - Greg Farnum

From a user:

There is some class of OSD out there (all filestore, IIRC) that are ultra slow to start at boot time in Luminous.

So on those machines I've been applying this change to `ceph-osd@.service`:

- `RestartSec=20s`
+ `RestartSec=1s`

And it fixes everything :)

#9 - 10/19/2018 08:01 PM - Greg Farnum

- *Status changed from In Progress to Resolved*

Okay, after discussing with CERN I've merged the PR to master so this isn't an issue going forward.

But unfortunately I think we're just going to have to live with it on existing installs, as the ceph-disk races remain common and require so many restarts as to render this tuning pretty broken. :/

#10 - 10/19/2018 08:18 PM - Nathan Cutler

- *Backport deleted (mimic, luminous)*

Clearing backport field on the assumption that's what was intended by the previous edit.