Ceph - Bug #16351

jewel : 60-ceph-partuuid-workaround-rules still needed by debian jessie (udev 215-17)

06/16/2016 03:42 PM - alexandre derumier

Status:	Resolved		**Start date:**	06/16/2016
Priority:	Urgent		**Due date:**	
Assignee:	Loic Dachary		**% Done:**	0%
Category:			**Estimated time:**	0.00 hour
Target version:			**Spent time:**	0.00 hour
Source:	other		**Reviewed:**	
Tags:			**Affected Versions:**	v10.2.0, v10.2.1, v10.2.2
Backport:	jewel		**ceph-qa-suite:**	
Regression:	Yes		**Pull request ID:**	
Severity:	3 - minor			

Description

associated mail thread

Hi,
since this commit
https://github.com/ceph/ceph/commit/9f77244b8e0782921663e52005b725cca58a8753

osd disk are not mounted anymore on debian jessie ((udev 215-17) at boot

I have looked at udev log, at the 95-ceph-osd.rules is never triggered

if I add the old 60-ceph-partuuid-workaround-rules, it's working fine

Related issues:

Copied to Ceph - Backport #16577: jewel : 60-ceph-partuuid-workaround-rules s...	**Resolved**

History

#1 - 06/16/2016 04:13 PM - Loic Dachary

- Assignee set to Loic Dachary

- Priority changed from Normal to Urgent

#2 - 06/20/2016 02:39 PM - Loic Dachary

- Description updated

- Status changed from New to In Progress

#3 - 06/21/2016 09:18 AM - Daniel Swarbrick

I ran into this bug also, and can confirm that re-adding the old 60-ceph-partuuid-workaround-rules fixes the problem.

#4 - 06/21/2016 01:54 PM - Loic Dachary

On Jessie 8.5, the rules work as expected when removing/adding a disk

```
udevadm monitor
...
KERNEL[399.124705] add        /devices/pci0000:00/0000:00:07.0 (pci)
KERNEL[399.128475] add        /devices/pci0000:00/0000:00:07.0/virtio3 (virtio)
UDEV  [399.131617] add        /devices/pci0000:00/0000:00:07.0 (pci)
KERNEL[399.132247] add        /devices/virtual/bdi/254:16 (bdi)
UDEV  [399.132499] add        /devices/pci0000:00/0000:00:07.0/virtio3 (virtio)
UDEV  [399.133016] add        /devices/virtual/bdi/254:16 (bdi)
KERNEL[399.142008] add        /devices/pci0000:00/0000:00:07.0/virtio3/block/vdb (block)
KERNEL[399.142033] add        /devices/pci0000:00/0000:00:07.0/virtio3/block/vdb/vdb1 (block)
```

```
KERNEL[399.142045] add        /devices/pci0000:00/0000:00:07.0/virtio3/block/vdb/vdb2 (block)
UDEV  [399.526162] add        /devices/pci0000:00/0000:00:07.0/virtio3/block/vdb (block)
UDEV  [403.304562] add        /devices/pci0000:00/0000:00:07.0/virtio3/block/vdb/vdb2 (block)
UDEV  [403.395816] add        /devices/pci0000:00/0000:00:07.0/virtio3/block/vdb/vdb1 (block)
```

and

```
ceph       938  6.4  0.4 831352 32052 ?        Ssl  13:50   0:00 /usr/bin/ceph-osd -f --cluster ceph --id 0 --
setuser ceph --setgroup ceph
```

But it does not work when rebooting. Trying to figure out why. The fact that adding 60-ceph-partuuid-workaround-rules fixes the issue is helpful.

#5 - 06/21/2016 01:56 PM - Loic Dachary

For the record, it works on CentOS 7.2 but a user reported it failed in the same way on CentOS 7 (maybe an older version) ?

#6 - 06/21/2016 03:02 PM - Loic Dachary

Maybe the problem is that [ADD udev events are not fired at boot time](#)

#7 - 06/22/2016 03:37 PM - Loic Dachary

The CentOS failure happens when the root file system is on LVM. When the root file system is on a regular partition, the OSD shows up as expected.

#8 - 06/23/2016 04:22 AM - alexandre derumier

Loic Dachary wrote:

> The CentOS failure happens when the root file system is on LVM. When the root file system is on a regular partition, the OSD shows up as
> expected.

on my debian jessie setup, it is without lvm

/etc/fstab
UUID=a9309e01-0f80-475b-95e1-33a9ddb4bdff /   xfs   defaults   0   1

#9 - 06/23/2016 07:38 AM - Loic Dachary

- Status changed from In Progress to Need Review

#10 - 06/23/2016 07:39 AM - Loic Dachary

- Backport set to jewel

#11 - 07/03/2016 11:33 AM - Loic Dachary

- Status changed from Need Review to Pending Backport

#12 - 07/03/2016 11:34 AM - Loic Dachary

- Copied to Backport #16577: jewel : 60-ceph-partuuid-workaround-rules still needed by debian jessie (udev 215-17) added

#13 - 08/05/2016 03:20 PM - Denis Pudov

- File ceph-disk.strace.sde1.4787.gz added

- File ceph-disk.strace.sde1.4545.gz added

alexandre derumier wrote:

> [associated mail thread](#)
>
> Hi,
> since this commit
> https://github.com/ceph/ceph/commit/9f77244b8e0782921663e52005b725cca58a8753
>
> osd disk are not mounted anymore on debian jessie ((udev 215-17) at boot
>
> I have looked at udev log, at the 95-ceph-osd.rules is never triggered

On my Jessie (8.5) with 4.6.0 backports kernel its triggered.
I have added strace to ceph-disk at the 95-ceph-osd.rules:

```
RUN+="/usr/bin/strace -tt -s 1024 -o /var/tmp/ceph-disk.strace.$name -ff /usr/sbin/ceph-disk --log-stdout -v t
rigger /dev/$name"
```

and got four strace output per device.

At the stage of calling trigger (/bin/systemctl --no-block restart ceph-disk@dev/sde1.service) get an error from systemd (full strace at attach):

```
org.freedesktop.DBus.Error.InvalidArgs.*sysinit.target is queued, ignoring restart request for unit ceph-disk@
dev-sde1.service
```


> if I add the old 60-ceph-partuuid-workaround-rules, it's working fine

For me it's not fine... systemd tries many times to run ceph-osd@.service and eventually get:

```
авг 05 18:00:18 storage-ceph-01y systemd[1]: ceph-osd@58.service: main process exited, code=exited, status=1/F
AILURE
авг 05 18:00:18 storage-ceph-01y systemd[1]: Unit ceph-osd@58.service entered failed state.
авг 05 18:00:18 storage-ceph-01y systemd[1]: ceph-osd@58.service start request repeated too quickly, refusing
to start.
авг 05 18:00:18 storage-ceph-01y systemd[1]: Failed to start Ceph object storage daemon.
```

#14 - 08/05/2016 03:24 PM - alexandre derumier

@denis:

do you have tried the fix from loic ?
https://github.com/ceph/ceph/pull/9885/commits/35004a628b2969d8b2f1c02155bb235165a1d809

#15 - 08/05/2016 03:42 PM - Denis Pudov

Yes, i have added 60-ceph-partuuid-workaround-rules to /lib/udev/rules.d/ and got mounts but have problems with start ceph-osd@.service.

#16 - 08/05/2016 04:41 PM - alexandre derumier

The fix is
60-ceph-by-parttypeuuid.rules
https://github.com/dachary/ceph/blob/35004a628b2969d8b2f1c02155bb235165a1d809/udev/60-ceph-by-parttypeuuid.rules

not old
60-ceph-partuuid-workaround-rules

#17 - 08/05/2016 07:40 PM - Denis Pudov

Alexandre, thank you!
60-ceph-by-parttypeuuid.rules works!

#18 - 08/09/2016 06:46 AM - Loic Dachary

Additional pull request, for the record: https://github.com/ceph/ceph/pull/10115

#19 - 08/25/2016 07:23 AM - Loic Dachary

- Status changed from Pending Backport to Resolved

#20 - 09/23/2016 02:08 AM - joke lee

it's still not mount osd after reboot when the osd disk is raid0

Files

ceph-disk.strace.sde1.4787.gz	8.77 KB	08/05/2016	Denis Pudov
ceph-disk.strace.sde1.4545.gz	142 KB	08/05/2016	Denis Pudov