

## bluestore - Bug #21068

### ceph-disk deploy bluestore fails to create correct block symlink for multipath devices

08/23/2017 12:50 AM - Justin Mammarella

<b>Status:</b>	Won't Fix	<b>Start date:</b>	08/23/2017
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>		<b>% Done:</b>	0%
<b>Category:</b>		<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	v11.2.1	<b>Reviewed:</b>	
<b>Source:</b>		<b>Affected Versions:</b>	v12.1.0
<b>Tags:</b>	ceph-disk multipath	<b>ceph-qa-suite:</b>	ceph-disk
<b>Backport:</b>	luminous	<b>Pull request ID:</b>	
<b>Regression:</b>	No	<b>Crash signature:</b>	
<b>Severity:</b>	3 - minor		

#### Description

Instead of pointing to the aggregated multipath device /dev/dm-XX, the block symlink would point to a single path.

/dev/part-byuuid only lists 1 device per part-uuid, ignoring any subsequent multipath devices.

Also, ceph-osd fails to start due to permission errors? Once we corrected the link, the errors went away.

#### Pull request:

<https://github.com/ceph/ceph/pull/17132>

#### Supporting Logs:

```
ceph-common.x86_64      2:12.1.4-0.el7 @Ceph
```

```
[root@flash1-ceph2-qh2 mapper]# DEV=4;ceph-disk prepare --bluestore --crush-device-class sandisk --block.db
/dev/disk/by-partlabel/osd-dm${DEV}-db --block.wal /dev/disk/by-partlabel/osd-dm${DEV}-wal /dev/dm-{$DEV}
Creating new GPT entries.
Warning: The kernel is still using the old partition table.
The new table will be used at the next reboot.
The operation has completed successfully.
prepare_device: OSD will not be hot-swappable if block.db is not the same device as the osd data
prepare_device: Block.db /dev/disk/by-partlabel/osd-dm4-db was not prepared with ceph-disk. Symlinking directly.
prepare_device: OSD will not be hot-swappable if block.wal is not the same device as the osd data
prepare_device: Block.wal /dev/disk/by-partlabel/osd-dm4-wal was not prepared with ceph-disk. Symlinking directly.
Warning: The kernel is still using the old partition table.
The new table will be used at the next reboot.
The operation has completed successfully.
Warning: The kernel is still using the old partition table.
The new table will be used at the next reboot.
The operation has completed successfully.
specified blocksize 4096 is less than device physical sector size 16384
switching to logical sector size 4096
meta-data=/dev/dm-14      isize=2048  agcount=4, agsize=6400 blks   =           sectsz=4096  attr=2, projid32bit=1
=           crc=1      finobt=0, sparse=0
data   =           bsize=4096  blocks=25600, imaxpct=25   =           sunit=0    swidth=0 blks
naming  =version 2      bsize=4096  ascii-ci=0  ftype=1
log     =internal log   bsize=4096  blocks=1608, version=2    =           sectsz=4096  sunit=1 blks, lazy-count=1
realtime =none         extsz=4096  blocks=0, rtextents=0
Warning: The kernel is still using the old partition table.
The new table will be used at the next reboot.
The operation has completed successfully.

[root@flash1-ceph2-qh2 ~]# less /var/log/ceph/ceph-osd.99.log
```

```

2017-08-23 10:35:21.978590 7fc79ea17d00 0 set uid:gid to 167:167 (ceph:ceph)
2017-08-23 10:35:21.978605 7fc79ea17d00 0 ceph version 12.1.4 (a5f84b37668fc8e03165aaf5cbb380c78e4deba4) luminous (rc),
process (unknown), pid 36240
2017-08-23 10:35:21.995301 7fc79ea17d00 1 bluestore(/var/lib/ceph/tmp/mnt.C1J3Ur) mkfs path /var/lib/ceph/tmp/mnt.C1J3Ur
2017-08-23 10:35:21.995361 7fc79ea17d00 -1 bluestore(/var/lib/ceph/tmp/mnt.C1J3Ur) _setup_block_symlink_or_file failed to open
block file: (13) Permission denied
2017-08-23 10:35:21.995385 7fc79ea17d00 -1 bluestore(/var/lib/ceph/tmp/mnt.C1J3Ur) mkfs failed, (13) Permission denied
2017-08-23 10:35:21.995391 7fc79ea17d00 -1 OSD::mkfs: ObjectStore::mkfs failed with error (13) Permission denied
2017-08-23 10:35:21.995452 7fc79ea17d00 -1 ESC[0;31m * ERROR: error creating empty object store in
/var/lib/ceph/tmp/mnt.C1J3Ur: (13) Permission deniedESC[0m
2017-08-23 10:35:22.518523 7fd4a81d4d00 0 set uid:gid to 167:167 (ceph:ceph)
2017-08-23 10:35:22.518556 7fd4a81d4d00 0 ceph version 12.1.4 (a5f84b37668fc8e03165aaf5cbb380c78e4deba4) luminous
(rc), process (unknown), pid 36315
2017-08-23 10:35:22.534819 7fd4a81d4d00 1 bluestore(/var/lib/ceph/tmp/mnt.1UY9IA) mkfs path
/var/lib/ceph/tmp/mnt.1UY9IA
2017-08-23 10:35:22.534864 7fd4a81d4d00 -1 bluestore(/var/lib/ceph/tmp/mnt.1UY9IA) _setup_block_symlink_or_file failed
to open block file: (13) Permission denied
2017-08-23 10:35:22.534878 7fd4a81d4d00 -1 bluestore(/var/lib/ceph/tmp/mnt.1UY9IA) mkfs failed, (13) Permission denied
2017-08-23 10:35:22.534881 7fd4a81d4d00 -1 OSD::mkfs: ObjectStore::mkfs failed with error (13) Permission denied
2017-08-23 10:35:22.534953 7fd4a81d4d00 -1 ESC[0;31m * ERROR: error creating empty object store in
/var/lib/ceph/tmp/mnt.1UY9IA: (13) Permission deniedESC[0m

[root@flash1-ceph2-qh2 ~]# mount -t xfs /dev/dm-14 /tmp/temp_mount
[root@flash1-ceph2-qh2 ~]# ls -la /tmp/temp_mount
total 48
drwxr-xr-x 2 ceph ceph 245 Aug 23 10:35 .
drwxrwxrwt 21 root root 4096 Aug 23 03:09 ..
-rw-r--r- 1 root root 474 Aug 23 10:35 activate.monmap
lrwxrwxrwx 1 ceph ceph 58 Aug 23 10:35 block > /dev/disk/by-partuuid/4eb20735-e84d-458e-b83e-1c73654e2c48 <- BAD SYMLINK
lrwxrwxrwx 1 ceph ceph 33 Aug 23 10:35 block.db > /dev/disk/by-partlabel/osd-dm4-db
-rw-r--r- 1 ceph ceph 37 Aug 23 10:35 block.db_uuid
-rw-r--r- 1 ceph ceph 37 Aug 23 10:35 block_uuid
lrwxrwxrwx 1 ceph ceph 34 Aug 23 10:35 block.wal > /dev/disk/by-partlabel/osd-dm4-wal
-rw-r--r- 1 ceph ceph 37 Aug 23 10:35 block.wal_uuid
-rw-r--r- 1 ceph ceph 37 Aug 23 10:35 ceph_fsid
-rw-r--r- 1 ceph ceph 8 Aug 23 10:35 crush_device_class
-rw-r--r- 1 ceph ceph 37 Aug 23 10:35 fsid
-rw----- 1 ceph ceph 57 Aug 23 10:35 keyring
-rw-r--r- 1 ceph ceph 21 Aug 23 10:35 magic
-rw-r--r- 1 ceph ceph 10 Aug 23 10:35 type
-rw-r--r- 1 ceph ceph 3 Aug 23 10:35 whoami

[root@flash1-ceph2-qh2 ~]# ls -la /dev/disk/by-partuuid/4eb20735-e84d-458e-b83e-1c73654e2c48
lrwxrwxrwx 1 root root 10 Aug 23 10:37 4eb20735-e84d-458e-b83e-1c73654e2c48 -> ../sdo2 <- SYMLINK IS SINGLE PATH

[root@flash1-ceph2-qh2 ~]# multipath -ll | grep dm-4 -A 4

mpatho (35001173100dd65ac) dm-4 SANDISK ,SDIFC10-0720801
size=7.0T features='1 queue_if_no_path' hwhandler='0' wp=rw
`-+ policy='round-robin 0' prio=1 status=active | - 4:0:8:0 sdg 8:96 active ready running
`- 4:0:17:0 sdo 8:224 active ready running

Correct block symlink should be:

[root@flash1-ceph2-qh2 ~]# ls -la /dev/disk/by-parttypeuid/ | grep 4eb20735
lrwxrwxrwx 1 root root 11 Aug 23 10:37 cafe-cafe-8ae0-4982-bf9d-5a8d867af560.4eb20735-e84d-458e-b83e-1c73654e2c48 ->
../dm-15

```

## History

#1 - 08/23/2017 12:59 AM - Justin Mammarella

Title of bug should be: 'ceph-disk prepare --bluestore fails to create correct block simlink for multipath devices.'

Apologies for the lack of formatting in the logs, if someone with edit ability could correct it, that would be appreciated.

**#2 - 08/23/2017 01:01 AM - Justin Mammarella**

Also target version should be luminous, next release.

**#3 - 08/23/2017 11:16 AM - Nathan Cutler**

- Subject changed from *ceph-disk deploy bluestore fails to create correct block symlink for multipath devices.* to *ceph-disk deploy bluestore fails to create correct block symlink for multipath devices*

- Backport set to *luminous*

**#4 - 11/29/2017 11:07 PM - Sage Weil**

- Project changed from *Ceph* to *bluestore*

- Category deleted (*ceph cli*)

**#5 - 02/20/2018 02:28 PM - Sage Weil**

- Status changed from *New* to *Won't Fix*

focusing on *ceph-volume* instead of *ceph-disk* for *bluestore* support.