

Orchestrator - Bug #45595

qa/tasks/cephadm: No filesystem is configured and MDS daemon gets deployed repeatedly

05/19/2020 08:04 AM - Varsha Rao

Status:	Need More Info	% Done:	0%
Priority:	Low		
Assignee:			
Category:	teuthology		
Target version:			
Source:		Affected Versions:	
Tags:		ceph-qa-suite:	
Backport:		Pull request ID:	
Regression:	No	Crash signature (v1):	
Severity:	3 - minor	Crash signature (v2):	
Reviewed:			

Description

On adding mds to roles

```
2020-05-07T11:21:19.160 ERROR:teuthology.run_tasks:Saw exception from tasks.
Traceback (most recent call last):
  File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/run_tasks.py", line 90
, in run_tasks
    manager.__enter__()
  File "/usr/lib/python3.6/contextlib.py", line 81, in __enter__
    return next(self.gen)
  File "/home/teuthworker/src/github.com_varshar16_ceph_wip-varsha-testing/qa/tasks/cephadm.py", l
ine 1106, in task
    healthy(ctx=ctx, config=config)
  File "/home/teuthworker/src/github.com_varshar16_ceph_wip-varsha-testing/qa/tasks/ceph.py", line
1428, in healthy
    ceph_fs.wait_for_daemons(timeout=300)
  File "/home/teuthworker/src/github.com_varshar16_ceph_wip-varsha-testing/qa/tasks/cephfs/filesys
tem.py", line 908, in wait_for_daemons
    if self.are_daemons_healthy(status=status, skip_max_mds_check=skip_max_mds_check):
  File "/home/teuthworker/src/github.com_varshar16_ceph_wip-varsha-testing/qa/tasks/cephfs/filesys
tem.py", line 758, in are_daemons_healthy
    mds_map = self.get_mds_map(status=status)
  File "/home/teuthworker/src/github.com_varshar16_ceph_wip-varsha-testing/qa/tasks/cephfs/filesys
tem.py", line 659, in get_mds_map
    return status.get_fsmap(self.id) ['mdsmap']
  File "/home/teuthworker/src/github.com_varshar16_ceph_wip-varsha-testing/qa/tasks/cephfs/filesys
tem.py", line 111, in get_fsmap
    raise RuntimeError("FSCID {0} not in map".format(fscid))
RuntimeError: FSCID None not in map
```

http://qa-proxy.ceph.com/teuthology/varsha-2020-05-07_09:45:36-rados-wip-varsha-testing-distro-basic-smithi/5030231/

Filesystem was not configured. So, I have added setup_cephfs(). But still the test fails with repeated deployment of MDS daemon.
<https://github.com/varshar16/ceph/commit/5d4c3bb634a87dceb004562bd186d3e3a6b3bbda#diff-8264ee52f589f4c0191aa94f87aa1aeb>

```
2020-05-12T10:06:16.729 ERROR:teuthology.run_tasks:Saw exception from tasks.
Traceback (most recent call last):
  File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/run_tasks.py", line 90
, in run_tasks
```

```
manager.__enter__()
File "/usr/lib/python3.6/contextlib.py", line 81, in __enter__
    return next(self.gen)
File "/home/teuthworker/src/github.com_varsharl6_ceph_wip-varsha-testing-3/qa/tasks/cephadm.py",
line 1107, in task
    healthy(ctx=ctx, config=config)
File "/home/teuthworker/src/github.com_varsharl6_ceph_wip-varsha-testing-3/qa/tasks/ceph.py", li
ne 1423, in healthy
    manager.wait_until_healthy(timeout=300)
File "/home/teuthworker/src/github.com_varsharl6_ceph_wip-varsha-testing-3/qa/tasks/ceph_manager
.py", line 2894, in wait_until_healthy
    'timeout expired in wait_until_healthy'
AssertionError: timeout expired in wait_until_healthy
```

http://qa-proxy.ceph.com/teuthology/varsha-2020-05-12_09:46:35-rados-wip-varsha-testing-distro-basic-smithi/5048961

Related issues:

Related to Orchestrator - Feature #46265: test cephadm MDS deployment	Duplicate
Related to Orchestrator - Bug #48916: "File system None does not exist in the..."	Duplicate

History

#1 - 05/19/2020 10:55 AM - Sebastian Wagner

the logs of a failed MDS:

```
starting mds.a at
debug 2020-05-12T10:00:53.931+0000 7f6d6f39b700 -1 mds.0.openfiles _load_finish got (2) No such file or direct
ory
cluster 2020-05-12T10:00:53.626686+0000 mgr.a (mgr.14142) 74 : cluster [DBG] pgmap v72: 17 pgs: 17 active+clea
n; 2.2 KiB data, 812 KiB used, 265 GiB / 268 GiB avail; 1023 B/s rd, 1 op/s
cluster 2020-05-12T10:00:53.923579+0000 mon.a (mon.0) 425 : cluster [INF] Health check cleared: FS_WITH_FAILED
_MDS (was: 1 filesystem has a failed mds daemon)
cluster 2020-05-12T10:00:53.926056+0000 mon.a (mon.0) 426 : cluster [DBG] mds.? [v2:172.21.15.31:6826/98158823
2,v1:172.21.15.31:6827/981588232] up:boot
cluster 2020-05-12T10:00:53.926118+0000 mon.a (mon.0) 427 : cluster [INF] Standby daemon mds.a assigned to fil
esystem cephfs as rank 0
cluster 2020-05-12T10:00:53.926216+0000 mon.a (mon.0) 428 : cluster [INF] Health check cleared: MDS_ALL_DOWN (
was: 1 filesystem is offline)
cluster 2020-05-12T10:00:53.926428+0000 mon.a (mon.0) 429 : cluster [DBG] fsmap cephfs:0/1 1 up:standby, 1 fai
led
audit 2020-05-12T10:00:53.926571+0000 mon.a (mon.0) 430 : audit [DBG] from='mgr.14142 172.21.15.31:0/350760470
1' entity='mgr.a' cmd=[{"prefix": "mds metadata", "who": "a"}]: dispatch
cluster 2020-05-12T10:00:53.928851+0000 mon.a (mon.0) 431 : cluster [DBG] fsmap cephfs:1/1 {0=a=up:replay}
audit 2020-05-12T10:00:54.457020+0000 mon.a (mon.0) 433 : audit [INF] from='mgr.14142 172.21.15.31:0/350760470
1' entity='mgr.a' cmd=[{"prefix": "config-key set", "key": "mgr/cephadm/host.smithi031", "val": {"\daemons\": {\
mon.a\": {\\"hostname\": \"smithi031\", \"container_id\": \"105bf202ec76\", \"container_image_id\": \"ba1862563
a7ec5bee9a9a7b56b0087f68457fcc4dec68b196c2f0023b5d5822f\", \"container_image_name\": \"quay.io/ceph-ci/ceph:5d
96f0c9612029b065cea7c34cd161b174878f8c\", \"daemon_id\": \"a\", \"daemon_type\": \"mon\", \"version\": \"16.0.
0-1393-g5d96f0c9612\", \"status\": 1, \"status_desc\": \"running\", \"last_refresh\": \"2020-05-12T10:00:54.45
3939\", \"created\": \"2020-05-12T09:58:29.620917\", \"started\": \"2020-05-12T09:58:34.349610\"}, \"mgr.a\":
{\\"hostname\": \"smithi031\", \"container_id\": \"ae5a1b713eee\", \"container_image_id\": \"ba1862563a7ec5bee9
a9a7b56b0087f68457fcc4dec68b196c2f0023b5d5822f\", \"container_image_name\": \"quay.io/ceph-ci/ceph:5d96f0c9612
029b065cea7c34cd161b174878f8c\", \"daemon_id\": \"a\", \"daemon_type\": \"mgr\", \"version\": \"16.0.0-1393-g5
d96f0c9612\", \"status\": 1, \"status_desc\": \"running\", \"last_refresh\": \"2020-05-12T10:00:54.454013\", \
\"created\": \"2020-05-12T09:58:35.932808\", \"started\": \"2020-05-12T09:58:35.989411\"}, \"osd.0\": {\\"hostna
me\": \"smithi031\", \"container_id\": \"1f2d6fcd37e\", \"container_image_id\": \"ba1862563a7ec5bee9a9a7b56b0
087f68457fcc4dec68b196c2f0023b5d5822f\", \"container_image_name\": \"quay.io/ceph-ci/ceph:5d96f0c9612029b065ce
a7c34cd161b174878f8c\", \"daemon_id\": \"0\", \"daemon_type\": \"osd\", \"version\": \"16.0.0-1393-g5d96f0c961
2\", \"status\": 1, \"status_desc\": \"running\", \"last_refresh\": \"2020-05-12T10:00:54.454047\", \"created\
\": \"2020-05-12T09:59:29.312886\", \"started\": \"2020-05-12T09:59:30.872638\"}, \"osd.1\": {\\"hostname\": \"s
mithi031\", \"container_id\": \"50d07b96ccd8\", \"container_image_id\": \"ba1862563a7ec5bee9a9a7b56b0087f68457
fcc4dec68b196c2f0023b5d5822f\", \"container_image_name\": \"quay.io/ceph-ci/ceph:5d96f0c9612029b065cea7c34cd16
1b174878f8c\", \"daemon_id\": \"1\", \"daemon_type\": \"osd\", \"version\": \"16.0.0-1393-g5d96f0c9612\", \"st
atus\": 1, \"status_desc\": \"running\", \"last_refresh\": \"2020-05-12T10:00:54.454140\", \"created\": \"2020
-05-12T09:59:44.896618\", \"started\": \"2020-05-12T09:59:46.443553\"}, \"osd.2\": {\\"hostname\": \"smithi031\
\", \"container_id\": \"30629d30788e\", \"container_image_id\": \"ba1862563a7ec5bee9a9a7b56b0087f68457fcc4dec68
b196c2f0023b5d5822f\", \"container_image_name\": \"quay.io/ceph-ci/ceph:5d96f0c9612029b065cea7c34cd161b174878f
8c\", \"daemon_id\": \"2\", \"daemon_type\": \"osd\", \"version\": \"16.0.0-1393-g5d96f0c9612\", \"status\": 1
, \"status_desc\": \"running\", \"last_refresh\": \"2020-05-12T10:00:54.454204\", \"created\": \"2020-05-12T10
:00:00.193354\", \"started\": \"2020-05-12T10:00:01.739960\"}, \"mds.a\": {\\"hostname\": \"smithi031\", \"cont
ainer_id\": \"a053eff517d8\", \"container_image_id\": \"ba1862563a7ec5bee9a9a7b56b0087f68457fcc4dec68b196c2f00
```

```
23b5d5822f\", \"container_image_name\": \"quay.io/ceph-ci/ceph:5d96f0c9612029b065cea7c34cd161b174878f8c\", \"d
aemon_id\": \"a\", \"daemon_type\": \"mds\", \"version\": \"16.0.0-1393-g5d96f0c9612\", \"status\": 1, \"statu
s_desc\": \"running\", \"last_refresh\": \"2020-05-12T10:00:54.454265\", \"created\": \"2020-05-12T10:00:07.08
3235\", \"started\": \"2020-05-12T10:00:52.996854\"}}, \"devices\": [{\"rejected_reasons\": [\"Insufficient sp
ace (<5GB) on vgs\", \"LVM detected\", \"locked\"], \"available\": false, \"path\": \"/dev/nvme0n1\", \"sys_ap
i\": {\"removable\": \"0\", \"ro\": \"0\", \"vendor\": \"\", \"model\": \"INTEL SSDPEDMD400G4\", \"rev\": \"\",
\"sas_address\": \"\", \"sas_device_handle\": \"\", \"support_discard\": \"512\", \"rotational\": \"0\", \"n
r_requests\": \"1023\", \"scheduler_mode\": \"none\", \"partitions\": {}, \"sectors\": 0, \"sectorsize\": \"51
2\", \"size\": 400088457216.0, \"human_readable_size\": \"372.61 GB\", \"path\": \"/dev/nvme0n1\", \"locked\":
1}, \"lvs\": [{\"name\": \"lv_1\", \"comment\": \"not used by ceph\"}, {\"name\": \"lv_2\", \"osd_id\": \"2\"
, \"cluster_name\": \"ceph\", \"type\": \"block\", \"osd_fsid\": \"7cef8f05-5030-46eb-9733-7cc96b2329a6\", \"c
luster_fsid\": \"09bf22bc-9437-11ea-a069-001a4aab830c\", \"osdspec_affinity\": \"\", \"block_uuid\": \"D2fXzg-
9b6C-x9Qk-PvTG-cFEN-pcyi-zGTgCr\"}, {\"name\": \"lv_3\", \"osd_id\": \"1\", \"cluster_name\": \"ceph\", \"type
\": \"block\", \"osd_fsid\": \"a856c3cd-7121-42c5-bcce-804b31cf7c33\", \"cluster_fsid\": \"09bf22bc-9437-11ea-
a069-001a4aab830c\", \"osdspec_affinity\": \"\", \"block_uuid\": \"hKW7JS-J8Vp-qMu5-RemP-V1eP-EMd-bVZ2ap\"},
{\"name\": \"lv_4\", \"osd_id\": \"0\", \"cluster_name\": \"ceph\", \"type\": \"block\", \"osd_fsid\": \"02984
fcd-ea6e-4bb9-a8ce-7b4165ff17ff\", \"cluster_fsid\": \"09bf22bc-9437-11ea-a069-001a4aab830c\", \"osdspec_affin
ity\": \"\", \"block_uuid\": \"ltCJoY-aYDA-BgUq-SD2p-2SU0-s8Mo-EVPpxd\"}, {\"name\": \"lv_5\", \"comment\": \"
not used by ceph\"}], \"human_readable_type\": \"ssd\", \"device_id\": \"INTEL SSDPEDMD400G4_CVFT53310008400BG
N\"}, {\"rejected_reasons\": [\"locked\"], \"available\": false, \"path\": \"/dev/sda\", \"sys_api\": {\"remov
able\": \"0\", \"ro\": \"0\", \"vendor\": \"ATA\", \"model\": \"ST1000NM0033-9ZM\", \"rev\": \"SN04\", \"sas_a
ddress\": \"\", \"sas_device_handle\": \"\", \"support_discard\": \"0\", \"rotational\": \"1\", \"nr_requests\"
: \"64\", \"scheduler_mode\": \"mq-deadline\", \"partitions\": {\"sdal\": {\"start\": \"2048\", \"sectors\":
\"1953522688\", \"sectorsize\": 512, \"size\": 1000203616256.0, \"human_readable_size\": \"931.51 GB\", \"hold
ers\": []}, \"sectors\": 0, \"sectorsize\": \"512\", \"size\": 1000204886016.0, \"human_readable_size\": \"93
1.51 GB\", \"path\": \"/dev/sda\", \"locked\": 1}, \"lvs\": [], \"human_readable_type\": \"hdd\", \"device_id\"
: \"ST1000NM0033-9ZM173_Z1W4HQEW\"}], \"daemon_config_deps\": {\"osd.0\": {\"deps\": [], \"last_config\": \"2
020-05-12T09:59:27.963479\"}, \"osd.1\": {\"deps\": [], \"last_config\": \"2020-05-12T09:59:43.514126\"}, \"os
d.2\": {\"deps\": [], \"last_config\": \"2020-05-12T09:59:58.767400\"}, \"mds.a\": {\"deps\": [], \"last_confi
g\": \"2020-05-12T10:00:48.936080\"}}, \"last_daemon_update\": \"2020-05-12T10:00:54.454344\", \"last_device_u
pdate\": \"2020-05-12T10:00:05.549895\", \"networks\": {\"networks\": {\"172.21.0.0/20\": [\"172.21.15.31\"], \"last_host_ch
eck\": \"2020-05-12T09:58:55.496480\"}}}' : finished
audit 2020-05-12T10:00:54.457810+0000 mon.a (mon.0) 434 : audit [INF] from='mgr.14142 172.21.15.31:0/350760470
1' entity='mgr.a' cmd=[{\"prefix\": \"config set\", \"who\": \"mds.all\", \"name\": \"mds_join_fs\", \"value\": \"all\"}]: dis
patch
audit 2020-05-12T10:00:54.458428+0000 mon.a (mon.0) 435 : audit [INF] from='mgr.14142 172.21.15.31:0/350760470
1' entity='mgr.a' cmd=[{\"prefix\": \"auth get-or-create\", \"entity\": \"mds.a\", \"caps\": [\"mon\", \"profile mds\", \"osd
\", \"allow rwx\", \"mds\", \"allow\"]}]: dispatch
audit 2020-05-12T10:00:54.459033+0000 mon.a (mon.0) 436 : audit [DBG] from='mgr.14142 172.21.15.31:0/350760470
1' entity='mgr.a' cmd=[{\"prefix\": \"config generate-minimal-conf\"}]: dispatch
audit 2020-05-12T10:00:54.459757+0000 mon.a (mon.0) 437 : audit [DBG] from='mgr.14142 172.21.15.31:0/350760470
1' entity='mgr.a' cmd=[{\"prefix\": \"config get\", \"who\": \"mds.a\", \"key\": \"container_image\"}]: dispatch
topping Ceph mds.a for 09bf22bc-9437-11ea-a069-001a4aab830c...
debug 2020-05-12T10:00:55.919+0000 7f6d763a9700 -1 received signal: Terminated from Kernel ( Could be generat
ed by pthread_kill(), raise(), abort(), alarm() ) UID: 0
debug 2020-05-12T10:00:55.919+0000 7f6d763a9700 -1 mds.a *** got signal Terminated ***
cephadm 2020-05-12T10:00:54.459456+0000 mgr.a (mgr.14142) 75 : cephadm [INF] Deploying daemon mds.a on smithi0
31
```

#2 - 05/20/2020 01:54 AM - Michael Fritch

We attempted to configure an MDS with file system `all` using an explicit daemon id of `a`?

```
ceph orch mds all --placement '1;smithi031=a'
```

mgr log:

```
2020-05-12T10:00:03.729+0000 7fce8d3a1700 1 -- [v2:172.21.15.31:6800/3029948604,v1:172.21.15.31:6801/3029948604] <== client.14194 172.21.15.31:0/31310088 1 ==== mgr_command(tid 0: {"prefix": "orch apply mds", "fs_name": "all", "placement": "1;smithi031=a", "target": ["mon-mgr", ""]}) v1 ==== 127+0+0 (secure 0 0 0) 0x55a5e88ae9a0 con 0x55a5e8f5d800
```

mon log:

```
2020-05-12T10:00:16.657+0000 7fb04d6e1700 5 mon.a@0(leader).mds e10 prepare_beacon mds.0 up:active -> down:down
2020-05-12T10:00:16.657+0000 7fb04d6e1700 1 mon.a@0(leader).mds e10 fail_mds_gid 14222 mds.a role 0
2020-05-12T10:00:16.657+0000 7fb04d6e1700 10 mon.a@0(leader).osd e24 blacklist [v2:172.21.15.31:6826/790498861,v1:172.21.15.31:6827/790498861] until 2020-05-13T10:00:16.658143+0000
```

From the logs it would appear the mds is being deployed in a loop:

- 1) cephadm deploys daemon `mds.a`
- 2) mds is killed after `rejoin_done`
- 3) cephadm does not see a running daemon (goto 1)

Maybe we need to validate the daemon_id starts with the fsname?

#3 - 05/20/2020 02:08 AM - Michael Fritch

Also, any ideas why the fsname is `all` ??

#4 - 05/22/2020 09:17 AM - Varsha Rao

Michael Fritch wrote:

Also, any ideas why the fsname is `all` ??

It is from here in ceph_mdss()

```
_shell(ctx, cluster_name, remote, [
    'ceph', 'orch', 'apply', 'mds',
    'all',
    str(len(nodes)) + ';' + ';' .join(nodes)]
)
```

<https://github.com/ceph/ceph/blob/master/qa/tasks/cephadm.py#L645>

mgr log

```
2020-05-12T10:00:08.245+0000 7fcea993f700 0 [cephadm DEBUG root] Applying service mds.all spec
2020-05-12T10:00:08.245+0000 7fcea993f700 0 [cephadm DEBUG cephadm.module] place 0 over provided host list: [
HostPlacementSpec(hostname='smithi031', network='', name='a')]
2020-05-12T10:00:08.245+0000 7fcea993f700 0 [cephadm DEBUG cephadm.module] Combine hosts with existing daemon
s [] + new hosts [HostPlacementSpec(hostname='smithi031', network='', name='a')]
2020-05-12T10:00:08.245+0000 7fcea993f700 0 [cephadm DEBUG root] hosts with daemons: set()
2020-05-12T10:00:08.245+0000 7fcea993f700 1 -- 172.21.15.31:0/3507604701 --> [v2:172.21.15.31:3300/0,v1:172.2
1.15.31:6789/0] -- mon_command({"prefix": "config set", "who": "mds.all", "name": "mds_join_fs", "value": "all
"} v 0) v1 -- 0x55a5e266a180 con 0x55a5e8608400
2020-05-12T10:00:08.246+0000 7fcea993f700 1 --2- 172.21.15.31:0/3507604701 >> [v2:172.21.15.31:3300/0,v1:172.
21.15.31:6789/0] conn(0x55a5e8608400 0x55a5e860e480 secure :-1 s=THROTTLE_DONE pgs=53 cs=0 l=1 rx=0x55a5e26620
00 tx=0x55a5e8664e60).handle_read_frame_epilogue_main read frame epilogue bytes=32
2020-05-12T10:00:08.246+0000 7fcea993f700 1 -- 172.21.15.31:0/3507604701 <== mon.0 v2:172.21.15.31:3300/0 246
==== mon_command_ack({"prefix": "config set", "who": "mds.all", "name": "mds_join_fs", "value": "all"})=0 v
6) v1 ==== 115+0+0 (secure 0 0 0) 0x55a5e88f0000 con 0x55a5e8608400
```

It goes in loop after orch apply. Why cephadm does not detect already deployed mds?

I got this patch verified by Patrick to fix fscid error. It is correct fix for it.

```
diff --git a/qa/tasks/cephadm.py b/qa/tasks/cephadm.py
index 1ffd02553d..2cb92cae06 100644
--- a/qa/tasks/cephadm.py
+++ b/qa/tasks/cephadm.py
@@ -21,7 +21,7 @@ from teuthology.orchestra.daemon import DaemonGroup
    from teuthology.config import config as teuth_config

    # these items we use from ceph.py should probably eventually move elsewhere
-from tasks.ceph import get_mons, healthy
+from tasks.ceph import get_mons, healthy, cephfs_setup

    CEPH_ROLE_TYPES = ['mon', 'mgr', 'osd', 'mds', 'rgw', 'prometheus']

@@ -1086,6 +1086,7 @@ def task(ctx, config):
    lambda: ceph_mgrs(ctx=ctx, config=config),
    lambda: ceph_osds(ctx=ctx, config=config),
    lambda: ceph_mdss(ctx=ctx, config=config),
+    lambda: cephfs_setup(ctx=ctx, config=config),
    lambda: ceph_rgw(ctx=ctx, config=config),
    lambda: ceph_monitoring('prometheus', ctx=ctx, config=config),
    lambda: ceph_monitoring('node-exporter', ctx=ctx, config=config),
```

#5 - 06/30/2020 08:06 AM - Sebastian Wagner

- Related to Feature #46265: test cephadm MDS deployment added

#6 - 11/19/2020 11:27 AM - Sebastian Wagner

- Status changed from New to Need More Info

#7 - 02/03/2021 10:56 AM - Sebastian Wagner

- Related to Bug #48916: "File system None does not exist in the map" in upgrade:octopus-x:parallel-master added

#8 - 02/17/2021 12:13 PM - Juan Miguel Olmo Martínez

- Assignee set to Sebastian Wagner

#9 - 05/27/2021 11:18 AM - Sebastian Wagner

- Priority changed from Normal to Low

#10 - 07/06/2021 04:01 PM - Sebastian Wagner

- Assignee deleted (Sebastian Wagner)