Arvados - Bug #17757

[a-d-c] the BootProbeCommand default leads to race conditions in our default/documented configuration

06/04/2021 03:00 PM - Ward Vandewege

Status:	Resolved	**Start date:**	06/04/2021
Priority:	Normal	**Due date:**	
Assigned To:	Ward Vandewege	**% Done:**	100%
Category:		**Estimated time:**	0.00 hour
Target version:	2021-06-09 sprint		

Description

The default BootProbeCommand is

```
    Containers:
      CloudVMs:
        BootProbeCommand: "docker ps -q"
```

Our installation instructions at https://doc.arvados.org/v2.2/install/crunch2-cloud/install-compute-node.html suggest to use our `build.sh` script to create a compute node image. That image will have this cloud-init script:

```
# Copyright (C) The Arvados Authors. All rights reserved.
#
# SPDX-License-Identifier: Apache-2.0
runcmd:
 - /bin/echo "cloudinit runcmd starting" | /usr/bin/logger
 - /usr/local/bin/ensure-encrypted-partitions.sh
 - /bin/echo "cloudinit runcmd ensure-encrypted-partitions.sh done" | /usr/bin/logger
 - /bin/echo "cloudinit runcmd finished" | /usr/bin/logger
 - /bin/touch /arvados-compute-node-boot.complete
```

The ensure-encrypted-partitions.sh script takes the attached scratch volumes, encrypts and mounts them, then stops docker and restarts it to make sure it uses the newly attached scratch volumes.

The build.sh script does not disable the docker service from starting.

There's a race condition here, particularly when the attached scratch space is large, and the docker image to be loaded for a workflow step is small. Here's an example of what can happen:

```
2021-06-04T13:38:37.720560086Z crunch-run 2.1.2 (go1.13.4) started
2021-06-04T13:38:37.720608151Z Executing container 'XXXXX-YYYYY-frdkk54nkyfvj6t'
2021-06-04T13:38:37.720620649Z Executing on host 'ip-172-31-8-234'
2021-06-04T13:38:37.871064526Z Fetching Docker image from collection 'fa39675198e1beb683f02cdc1aca75f2+261'
2021-06-04T13:38:37.897990135Z Using Docker image id 'sha256:361f55a39518e271a49368671e32ac2f17a447ddd72c56aef675a29aff7350bf'
2021-06-04T13:38:37.916400168Z Loading Docker image from keep
2021-06-04T13:38:37.871204475Z notice: reading stats from /sys/fs/cgroup/cpuacct/cgroup.procs
2021-06-04T13:38:37.871246154Z notice: reading stats from /sys/fs/cgroup/memory/memory.stat
2021-06-04T13:38:37.871462157Z mem 810102784 cache 0 swap 1815 pgmajfault 200908800 rss
2021-06-04T13:38:37.871498468Z notice: reading stats from /sys/fs/cgroup/cpuacct/cpuacct.stat
2021-06-04T13:38:37.871566828Z notice: reading stats from /sys/fs/cgroup/cpuset/cpuset.cpus
2021-06-04T13:38:37.871604624Z cpu 13.0700 user 12.5800 sys 2 cpus
2021-06-04T13:38:37.871804598Z net:docker0 0 tx 0 rx
2021-06-04T13:38:37.871811980Z net:ens5 295569 tx 95783318 rx
2021-06-04T13:38:44.096959603Z error in Run: While loading container image: While loading container image into Docker: error during connect: Post http://%2Fvar%2Frun%2Fdocker.sock/v1.21/images/load?quiet=1: EOF
2021-06-04T13:38:44.127642455Z error in CaptureOutput: error scanning files to copy to output: open .arvados#collection: no such file or directory
2021-06-04T13:38:44.127687005Z Cancelled
2021-06-04T13:38:44.740122288Z crunch-run finished
```

This failure is pretty harmless; the container will get retried.

But sometimes the effect is more serious: we've seen jobs where the BootProbeCommand says OK (docker is up and ready), while the `ensure-encrypted-partitions.sh` is still encrypting/mounting the scratch space. Meanwhile a (tiny) docker image is loaded and a container started. A few seconds in, `ensure-encrypted-partitions.sh` stops and restarts docker, which leads to a cryptic "Container exited with code: 137" (aka it was killed) and the container fails, no retries.

The workaround is simple, update the BootProbeCommand to

```
    Containers:
      CloudVMs:
        BootProbeCommand: ls /arvados-compute-node-boot.complete
```

The cleanest fix would probably be to update the packer script called from `build.sh` so that it disables the docker service. That way it would not be started on boot, and only be started when `ensure-encrypted-partitions.sh` is really done.

Subtasks:

Task # 17758: review 17757-bootprobe-race	**Resolved**

Associated revisions

Revision eba64f94 - 06/04/2021 09:31 PM - Ward Vandewege

Merge branch '17757-bootprobe-race'

closes #17757

Arvados-DCO-1.1-Signed-off-by: Ward Vandewege <ward@curii.com>

History

#1 - 06/04/2021 03:00 PM - Ward Vandewege

- Status changed from New to In Progress

#2 - 06/04/2021 03:10 PM - Ward Vandewege

- Subject changed from [config] the BootProbeCommand default leads to race conditions in our default/documented configuration to [a-d-c] the BootProbeCommand default leads to race conditions in our default/documented configuration

#3 - 06/04/2021 03:11 PM - Ward Vandewege

- Description updated

#5 - 06/04/2021 03:13 PM - Ward Vandewege

- Description updated

#6 - 06/04/2021 03:18 PM - Ward Vandewege

- Description updated

#7 - 06/04/2021 03:23 PM - Ward Vandewege

ready for review at b5178b3f99de705e2afcdc91f763f779481131cb on branch 17757-bootprobe-race

#8 - 06/04/2021 03:28 PM - Ward Vandewege

- Release set to 39

#9 - 06/04/2021 06:15 PM - Tom Clegg

LGTM, thanks! I suspect this is responsible for several occasions where I've seen containers terminate/retry and I didn't bother to follow up.

#10 - 06/04/2021 09:34 PM - Ward Vandewege

- % Done changed from 0 to 100

- Status changed from In Progress to Resolved

Applied in changeset arvados|eba64f9495cfcf65203eee41e90765cf5b6732d6.