Containers Created at Runtime for Batch Jobs

Didrik A. Rognstad
Master's Student at Bergen University College
Supervisors:
Bjarte Kileng and Kristin Fanebust Hetland

March 31, 2016

Agenda

- Why run jobs in containers?
- Brief introduction to containers.
- My implementation **cbatch**.

Why run jobs in containers?

- Homogeneous environments for batch jobs.
- Job environments can be updated easily.
- Job isolation.
- Minimal impact on performance.

What are containers?

A container is one or more processes contained using features present in the Linux Kernel.

- Namespaces
- CGroups
- Kernel Capabilities

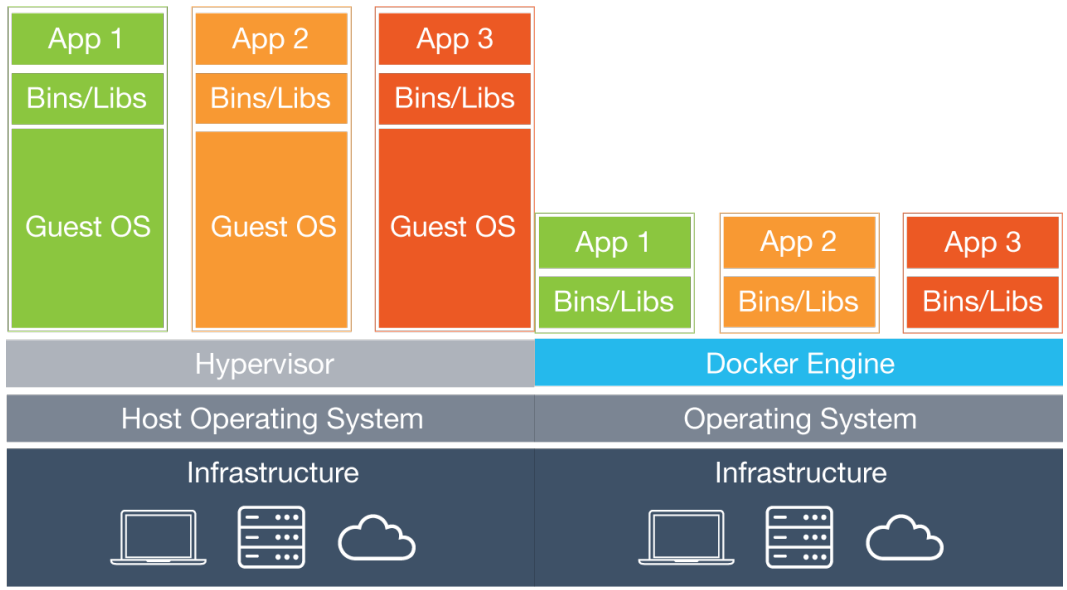

Figure: VM vs Container

Container implementations

- **Docker**
- LXC
- rkt
- runC

Goals

- Provide homogeneous environments on heterogeneous infrastructure.
- Isolation, prevent jobs from disturbing each other.
- Security, a malicious job should not cause harm to or be able to take control over the host system.
- Efficient, minimal impact on host performance.

Implementation

My implementation: **cbatch**

- Use Docker as container implementation.
- Installed on each worker node
- Docker must be installed on each worker node.
- Receives jobs from Torque.
- Container image must be present.

Docker

- Application containers.
- Data written to running containers are not persisted.
- Images consists of read only layers. Each container each have their own RW layer. Storing images and creating the RW container layer is handled by a storage driver.
- Kernel namespaces (user namespaces in v1.10+), cgroups, capabilites.

Life of a cbatch job

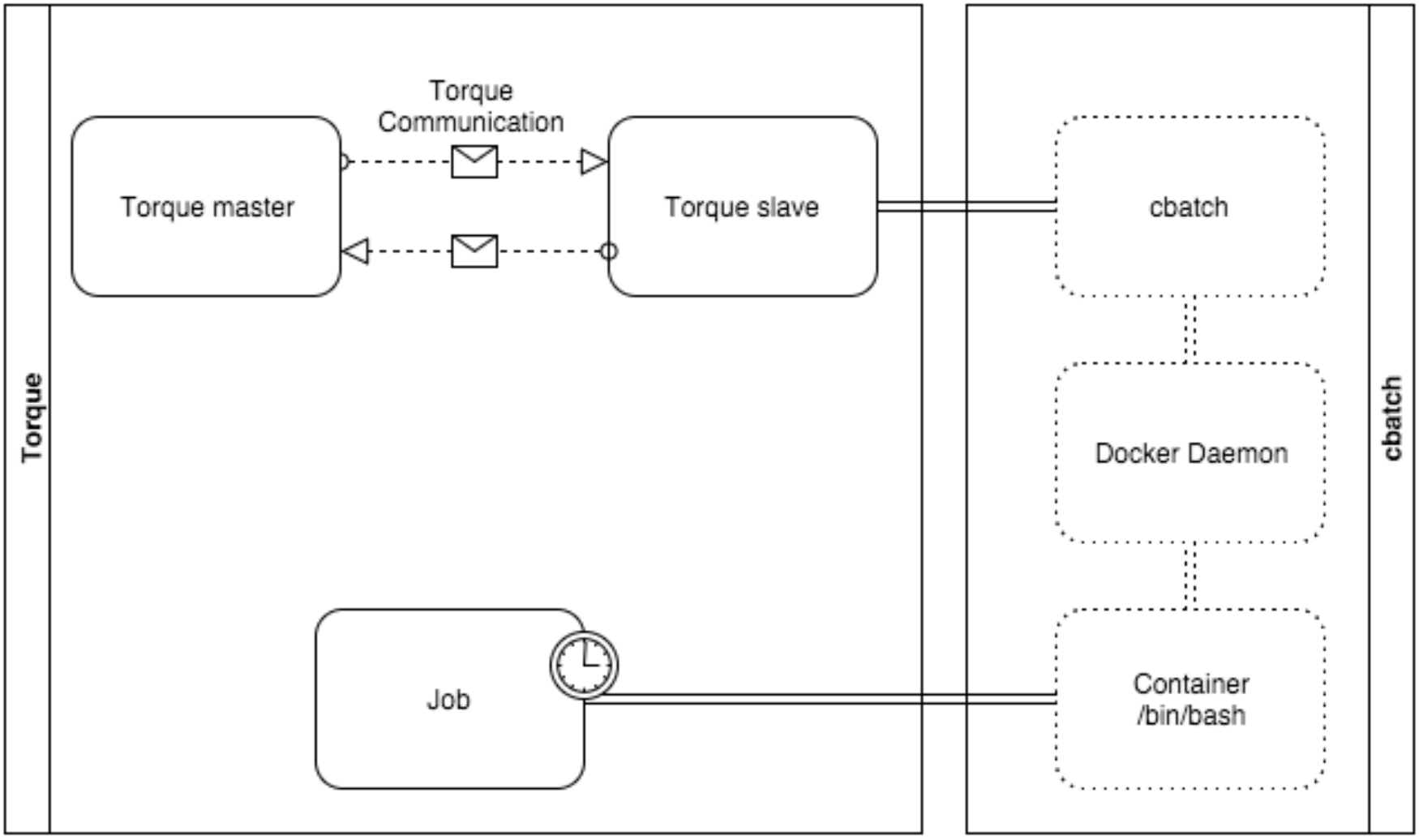

Figure: cbatch and Docker is transparent to Torque.

File access

The container needs access to several files to run the batch job from Torque:

- Access to the batch job is provided by mounting it from the host.
- Read and write to mounted resources operate at native host speed.
- Input and output files is accessed by mounting folders from the host as data volumes.

Challenges

- Run job as correct user in the container.
  - Solved by mounting */etc/passwd* and */etc/group* from the host.
- Jobs saving files to the container RW layer.
  - Dockers storage driver might be a bottleneck.

Outlook

- Use CernVM docker image.
- Make cbatch run jobs submitted through AliEn.
- Compare performance to native host and VMs.
- Evaluate security.
- Support other container implementations.

Questions?

Email: didrik.rognstad@gmail.com