Docker for space missions

Cécile Cavet

François Arago Centre (FACe), APC, Univ. Paris Diderot, CNRS/IN2P3, CEA/Irfu, Obs. de Paris, Sorbonne Paris Cité,

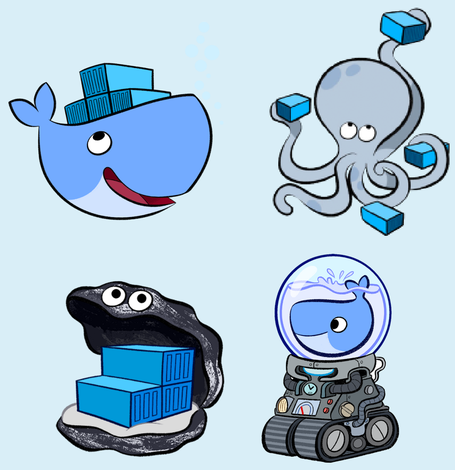
13 rue Watt, 75013, Paris, France

ccavet at apc.in2p3.fr

Abstract

Docker is a recent technology which offers multi-infrastructure environnement. Based on Linux containers, the Docker solution provides an ecosystem and a Hub for sharing images. Containers run lighter and quicker than virtual machines (VMs) but they are complementary technologies: on the IaaS (Infrastructure-as-a-Service) cloud infrastructure, the Docker plugging is not always integrated and on non Linux machines, the Docker daemon run inside a light VM. For space missions, scientists and engineers need a flexible environment in order to develop and run pipelines in production. Docker allows to develop the code locally, push it on a registry and run it on the scalable cloud in a multi-infrastructure approach. Furthermore, in the case of load peaks of the data analysis, the complex workflow can be managed by a hybrid cluster/cloud infrastructure including a Docker orchestrator.

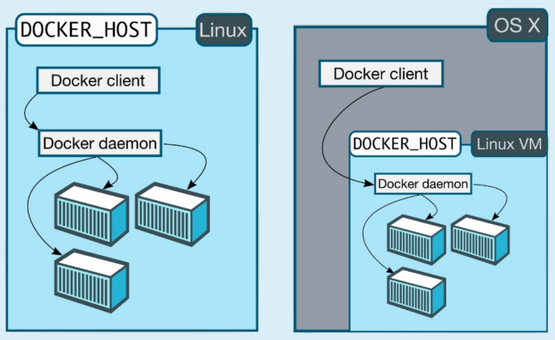
Docker ecosystem



- ▶ **Engine**: daemon and CLI client.
- ▶ **Compose**: multi-container application.
- ▶ **Registry**: secure private registry.
- ▶ **Machine**: local and cloud VMs.

User mode

- ▶ **Non Linux local machine**: **Engine** in a light VM (HyperKit virtualisation for macOS).
- ▶ **IaaS cloud**:
  - Linux VM + **Engine**.
  - **Machine** + Linux VM.
  - **OpenStack Magnum** plugging.



Linux vs macOS local machines.

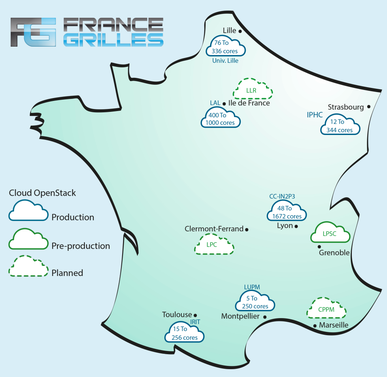
OpenStack cloud

OpenStack is an open-source solution used on academic public IaaS cloud since 2012. Virtual resources such as VMs are:

- instantiated by the **KVM** hypervisor.
- managed by a stack of services (**Nova**, **Neutron**, **Cinder**...).
- run from OS disk images provided by the **Glance** catalog.
- managed by the **Horizon Dashboard** and the CLI client.

French cloud Federation

The French actors of the academic cloud computing are federated by France Grilles [1] in order to offer cloud services.



French cloud infrastructures [2].

Docker history

The **Docker** solution has been developed since 2013. The **Docker** technology is based on LXC (Linux Containers) and allows the isolation of processes. In the **DevOps** approach, each application is isolated within its specific environnement and shared with other users.

LISA DPC

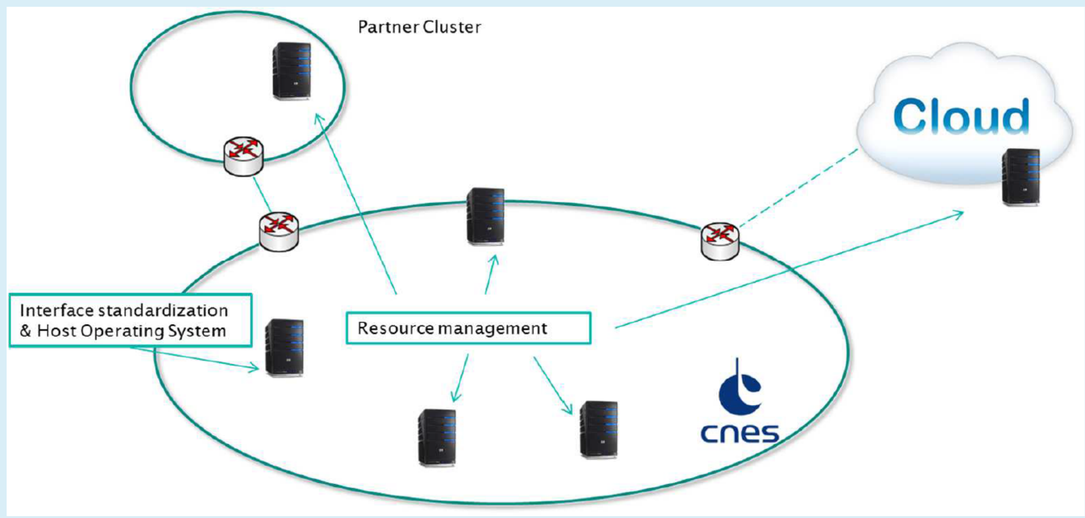
▶ The ESA L3 **LISA** space mission has the goal to study gravitational waves.

▶ The proto-**Data Processing Center** (DPC) for LISA [4] aimed to:

- efficiently manage computing.
- offer a development infrastructure.

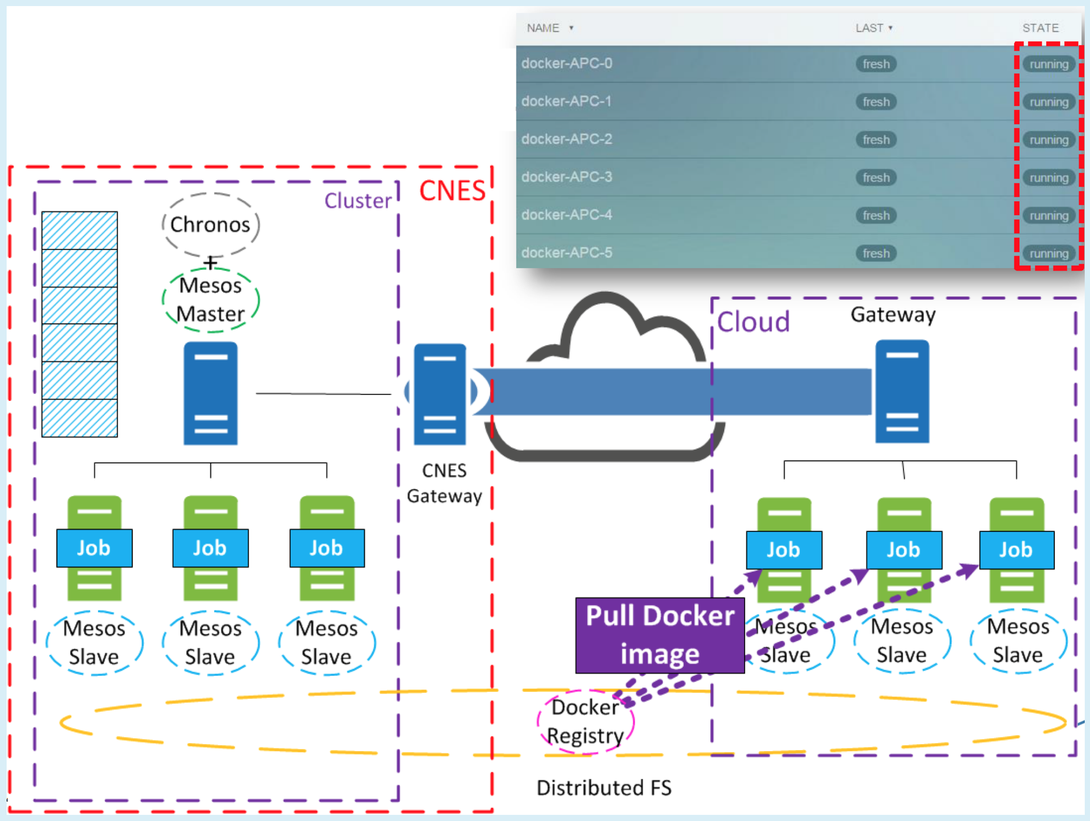
Hybrid cluster/cloud infrastructure

Complexe workflows such as load peak management can be supported by a hybrid infrastructure. For the LISA data analysis, the main processing will take place on a cluster of physical servers and the on-demand processing will run on the cloud VMs.



LISA DPC infrastructure [3].

- ▶ **CNES cluster**: local cluster with **Engine** and **Mesos** master.
- ▶ **Private registry**: hosted by a cloud VM, hosting the LISA Pathfinder **Docker** image.
- ▶ **HXN ATOS cloud**: VMs configured with **Ansible** and **Engine**, on run/stop states.
- ▶ **Mesos/Chronos orchestrator**: management of cluster and cloud resources.



Orchestration of Docker containers [3].

Container orchestrators

- ▶ **Docker Swarm**: the native **Docker** solution, a cluster of **Engine**.
- ▶ **Kubernetes**: the Google open-source platform working with **Docker** and **Rkt**.
- ▶ **Mesos**: the Apache open-source solution managing various APIs (**Docker**, Spark...).

Container life cycle

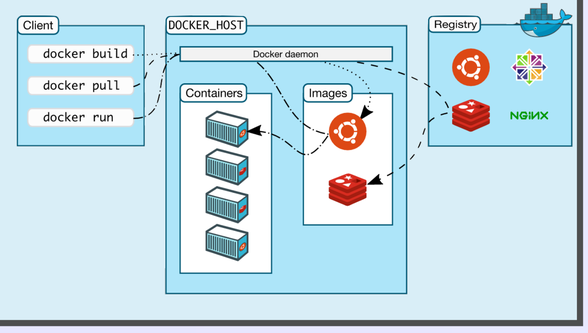
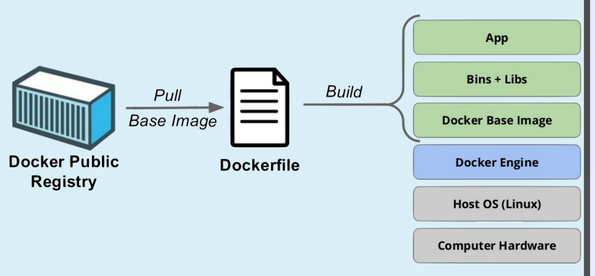


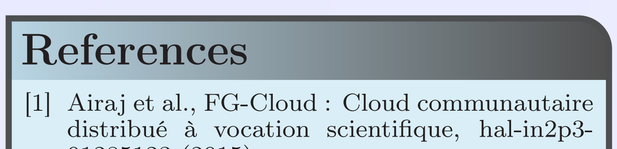
Image creation process



- ▶ **Docker Hub**: a marketplace for sharing images of various OS and applications.
- ▶ **Dockerfile**: a kind of shell script with specific instructions (RUN...).
- ▶ **Compose file**: a YAML file allowing to automatize the building of a multi-container application.

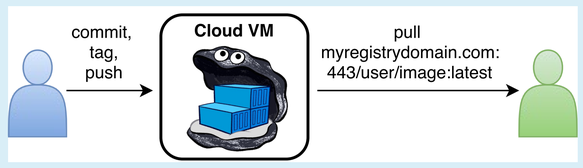
Public registry

- ▶ **Image sharing**: **Docker Hub**, a SaaS platform (cloud hosted service).
- ▶ **Automatic image build**:
  - Version control repository: GitHub or Bitbucket hosting **Dockerfile** and Readme.
  - Image build on new commit.
  - Branches on-demand (latest / develop).



Private registry

- ▶ **Image sharing**: **Registry**, a containerized service for secure image sharing (non public code, restricted users).
- ▶ **Manual image creation**.
- ▶ **SaaS service**:
  - Simple Web server (without Web interface) created with **Compose**: **Registry** and **Nginx** containers.
  - Security: TLS certificats.
  - Domaine name: **myregistrydomain.com** associated with IP adress.



References

[1] Airaj et al., FG-Cloud : Cloud communautaire distribué à vocation scientifique, hal-in2p3-01285123 (2015)

[2] FG-cloud: http://www.france-grilles.fr/services-catalogue/fg-cloud/

[3] Poncet et al., Private and cloud based clusters, BiDS'16 (2016)

[4] C. Cavet et al., A proto-DPC for LISA, 11th International LISA Symposium (2016)