ComputeOps project - Conseil Scientifique IN2P3 2022

Table of Content

Abstract

ComputeOps is a transverse project of the IN2P3 resulting from the Decalog masterproject, aiming at studying software containerization technologies on heterogeneous computing hardware (CPUs with vector instructions, GPUs, FPGAs,...).

The Decalog project was initiated to face the new challenges of high energy physics, in particular the increasing heterogeneity of computing means and resources. In this context, the members of ComputeOps are studying the possibilities offered by software containerization technologies with a view to portability and reproducibility of codes.

Moreover, IT practices are evolving very quickly, and the "DevOps" movement has appeared in recent years. This movement leads more and more developers to deliver their codes accompanied by a container image, or a file allowing to rebuild such a container (light virtual machines). Thus, the application can be deployed much more easily, and in a similar execution context to the developer's, without requiring the intervention of an administrator to install this or that application library. Docker is the emblematic tool of this movement.

This concept is now spreading at high speed in the world of cloud and supercomputers; in the latter case, rather around the Singularity/Apptainer tool. To keep the possibility of using these computing resources, and to benefit from the flexibility brought by containers, it is vital to prepare the use of these containers in our disciplines and in our computing grids.

Scientific challenges

"DevOps" software development practices (continuous development and integration) as well as the global offer of French and international computing centers are pushing the education and research community to integrate these new methods of deployment and service orchestration. Today, all "mésocentres" offer containerization tools and large scientific collaborations are adapting to these new challenges. Large space projects, for example, have all integrated an IT development organization with continuous integration and deployment according to the precepts of the DevOps work organization. In this context, the credibility of the experiments in which our laboratories participate depends on knowing and mastering these methods.

Within IN2P3, the Gitlab service has met a growing success, even beyond our institute. This tool is particularly appreciated for its ability to rely on "DevOps" tools and containers. However, this tool, as well as the underlying technologies, are evolving very quickly and many features are unknown and little used. All the professions linked to infrastructures and IT development can benefit from this common tool to build a stable and performing application base, at the state of the art.

Moreover, the use of containerization and continuous integration tools allows the reproducibility of scientific calculations and algorithms (provided that good practices are respected), which is essential for scientific publications and more broadly in the context of Open Science.

Project

The main objective of the project is to study the advantages of containers for HPC applications. The sub-objectives follow from this premise: organize technology watch, training as well as dissemination of knowledge and know-how to the whole community (research as well as technical staff).

Since 2018, the project is funded by master project DecaLog IN2P3 up to 7 k€ for missions and the organization of training or technical seminars.

The project participants are staff from the institute's laboratories (APC, IJCLAB, LPC, LPNHE, LLR, IPHC, CC), external partners (IAS, INRAE, CEA, Universities, INS2I, CNES) as well as the private sector (Sylabs, INTEL).

Expected outputs

- Publications, tutorials and workshops
- Comparison of container technologies for IN2P3 computing
- Services for the community (container sharing platforms, orchestration services...)
- Containerization of pilot applications
- Prototype of a multi-container and multi-resource submission tool

Goals

- Compare different container technologies (docker, rocket, lxd, udocker, singularity, shifter): effect on performance, vectoring, access to compute boosters, security, ease of administration and use. Is there a decisive advantage to disdaining the dominant tool (Docker) and turning to "HPC" alternatives (Singularity, Shifter....)?
- Study the interoperability of technologies. Especially images and image reconstruction files. For example, could a user develop on his workstation in Docker, then deploy on the computing center in Singularity?
- Validate the compatibility of containers with the grid.
- Prototype a job submission tool that can deploy a containerized application on a personal machine, the grid, clouds, supercomputers, etc., i.e., allow a physicist, via a single interface, to run his or her application on his or her workstation or on any computing resource at his or her disposal.
- When running containers in production on a cluster, it is necessary to use a container orchestrator that distributes the tasks on the available resources and on demand. Several solutions exist (Docker Swarm, Kubernetes, Mesos....), so it is necessary to follow and compare technologies.
- Facilitate the dissemination and use of ready-to-use codes for the research community in the various computing infrastructures ("mésocentres", grids, clouds, etc.)
- To sensitize the research communities on the good practices of software reproducibility within the framework of scientific publications or large experiments

History and Timeline

History

- Fall 2017: identification of pilot sites and administrators; identification of representative applications and pilot users, with vectorization, GPU, OpenMP, MPI.
- 2018: technologies to be tested are distributed among pilot administrators and users; preparation of sites, containerization of applications; testing campaign; following the tests, problematic technologies are discarded. Organization of a training session "ANF École IN2P3 Informatique on containers in production".
- 2019: installation of the selected technologies (at least docker and an alternative) on all the pilot sites and cross-testing campaign; interoperability is studied; a white paper summarizes the comparison.
- 2020: Implementation of a CSAN PoC, dropping of sregistry, comparison between Nomad and Kubernetes.
- 2023: Organization of a training session "ANF École IN2P3 Informatique on container orchestration".

Publications

- Alexandre Dehne, Rémy Dernat, Jerome Pansanel, Tovo Rabemanantsoa, Richard Randriatoamanana, et al.. CSAN : a Comprehensive Software Archive Network. Journées réseaux de l'enseignement et de la recherche (JRES) 2022, May 2022, Marseille, France. ⟨hal-03673448⟩
- CSAN, a Comprehensive Software Archive Network - JCAD 2021
- ComputeOps: Container for High Performance Computing,
  Cécile Cavet, Martin Souchal, Sébastien Gadrat, Gilles Grasseau, Andrea Satirana, Aurélien Bailly-Reyre, Olivier Dadoun, Victor Mendoza, David Chamont, Gérard Marchal-Duval,

Emmanuel Medernach and Jérôme Pansanel,
EPJ Web Conf., 245 (2020) 07006
DOI: https://doi.org/10.1051/epjconf/202024507006

- ComputeOps: containers for High Performance Computing,
  Cécile Cavet, Aurélien Bailly-Reyre, David Chamont, Olivier Dadoun, Alexandre Dehne Garcia, Pierre-Emmanuel Guérin, Pascale Hennion, Oleg Lodygensky, Gérard Marchal-Duval, Emmanuel Medernach, Victor Mendoza, Jérôme Pansanel, Richard Randriatoamanana, Andrea Sartirana, Martin Souchal and Julien Tugler,
  EPJ Web Conf., 214 (2019) 07004
  DOI: https://doi.org/10.1051/epjconf/201921407004
- ComputeOps: Container for High Performance Computing - CHEP 2018
- Transporter ses applications parallèles avec les containers LXD et Singularity - JDEV 2017

Talks

- Conteneurs pour le calcul - AI DevTalks INRIA (13 juillet 2021)
- Conteneurs pour le calcul - Formation UST4HPC (20 janvier 2021)
- Kubernetes pour le calcul - Journée Kubernetes du CC IN2P3 (19 février 2020)
- ComputeOps : conteneurs pour le HPC - JI (9 Octobre 2018)
- ComputeOps - Journée Projets IN2P3 R&D transverse Calcul&Données (29 Janvier 2018)

Tutorials

- Piscine Singularity aux JI (oct. 2018).
- Singularity à SBAC-PAD (sept. 2018).
- Docker/Singularity au FACe (janvier 2018).
- Atelier Docker à l'APC (juillet-aout 2020).

Training Sessions

- ANF 2018 : Conteneurs en production (4 jours)
- ANF 2023+ : Conteneurs en production deuxième session

Workshops

- Sessions de formation sur Kubernetes par Fabrice Jammes pour les membres de ComputeOps, à distance, 2021 - 2022 : https://www.k8s-school.fr/formations-kubernetes/
- ComputeOps, APC, 19 Septembre 2019 : https://indico.in2p3.fr/event/19346/
- ComputeOps, LPNHE, 5 Mars 2019 : https://indico.in2p3.fr/event/18626/
- ComputeOps, APC, 13 Novembre 2018 : https://indico.in2p3.fr/event/18077/

Reports

- Contribution to the IN2P3 prospective document: GT09 - Computation, algorithms and data (2021)

State of the art

To our knowledge, there is no equivalent project at different scales. But container solutions are already used at CERN (WLCG community) and at the level of the European EGI grid so studies had to be done in this context.

Resources and means

Technical resources

The project itself does not require any particular technical resources. Meetings are held by videoconference and once a year in person. Minutes and collaborative work tools are those of the institute (gitlab, etherpad, Big Blue Button, etc...).

Pilot applications and proofs of concept are tested on IN2P3 infrastructures: mainly at the CC but also in laboratories (APC, LPNHE, IPHC...).

The services operated by ComputeOps are hosted on multi-institute infrastructures: cloud resources at IPHC and APC and storage at INRAE.

The training courses and the practical work supports are hosted on the CC-IN2P3 gitlab.

Human Resources

Laboratoire	Nom	ETP	Tutelle	Statut
APC	Cécile Cavet	0.1	IN2P3	IR
LPNHE	Aurélien Baily-Reyre	0.05	IN2P3	Ingénieur
APC	Martin Souchal	0.1	UdP	IE
APC	Pavel Zakharov	0.1	IN2P3	IE
LPC	Fabrice Jammes	0.1	IN2P3	Ingénieur
IJCLab	David Chamont	0.1	IN2P3	IR
IPHC	Jérôme Pansanel		IPHC	IR
IPHC	Yanis Govinda		IPHC	CDD

Laboratoire	Nom	ETP	Tutelle	Statut
IPHC	Emmanuel Medernach		IPHC	Ingénieur
LPNHE	Olivier Dadoun	0	IN2P3	Ingénieur
LPNHE	Victor Mendoza	0.05	IN2P3	Ingénieur
LLR	Michael Mellin	0	IN2P3	Ingénieur
CC	Sebastien Gadrat	0.05	IN2P3	Ingénieur
IJCLab	Gérard MARCHAL-DUVAL	0	IN2P3	Ingénieur
LS2N	Richard Randriatoamanana		INS2I	Ingénieur
LS2N	Pierre-Emmanuel Guerin		INS2I	Ingénieur
LS2N	Davide Rovelli		INS2I	Ingénieur
DipSO	Alexandre Dehne-Garcia		INRAE	IE
DNO/ISA/CID	Guillaume Eynard-Bontemps		CNES	Ingénieur
	Yann COSTES		Université Cergy	Ingénieur
ISEM	Remy Dernat		Université Montpellier	IR
IRFU	Andrea Formica		CEA	Chercheur-Ingénieur
IRFU	BOCQUIER Maxime		CEA	Chercheur-Ingénieur
IRFU	LOUVIN Henri		CEA	Chercheur-Ingénieur

Financial resources

The budget allocated to ComputeOps by IN2P3 is 7000 € renewable every year. This budget is exclusively used for the organization of seminars or the financing of missions.

Technical achievements

Pilot applications for compute containers

Within the ComputeOps group we have identified the following pilot applications. These are scientific applications used in the institute's computing centers and cover a wide variety of cases (AI, Machine Learning, HPC, HTC...):

- SMILEI: an open-source, particle-in-cell code.
- CMS MEM: Matrix Element Method for the analysis of the Higgs boson production
- HAhRD: machine-learning application based on Tensorflow.
- electron_capture: electron capture rates.
- TensorFlow: machine learning applications using TensorFlow library.
- LDC: simulation pipeline for scientific challenges of the LISA space mission.
- Geant4 simulation for direct Dark-Matter detection experiments (Xenon & DarkSide).

These pilot applications allowed us to test the ease of implementation in the form of containers, to validate the performance in operation as well as the portability and sharing. The results of this study have been published (see publications).

Sharing container

As we have seen, one of the major interests of containerization technologies is to be able to easily share a containerized application: indeed, the various container technologies offer native integration with application catalogs (the best known being the DockerHub for Docker) open to all. A portal makes it easy to find what you're looking for and an integrated API allows you to share and retrieve images quickly.

Within the ComputeOps group we started by identifying and analyzing existing tools and more specifically those used by the HPC community. Our study has shown that for computationally oriented technologies, these catalogs are not or hardly used, for several reasons, the main one being a lack of visibility of higher education and research projects on general-purpose platforms. On the other hand, large computing centers have developed internal solutions for the distribution of images, for example a catalog hosted on distributed file systems such as CVMFS (for CERN and CC IN2P3).

Following these observations, and in order to respond to user demand, the ComputeOps group proposed the hosting of a Singularity Hub for the ESR in the form of a proof of concept. The Singularity Hub was to allow all ComputeOps members, but also partners outside IN2P3, to share containers in order to perform tests and benchmarks on computing centers with different hardware. The service was subsequently opened to outside IN2P3 from 2018 to 2020. The CSAN project then took over from this project (see below).

CSAN, a Comprehensive Software Archive Network

The initiative of the CSAN (Comprehensive Software Archive Network) project is to propose to the ESRI community a catalog of open-source scientific applications ready to use and optimized for the national computing centers and "mésocentres". This proposal is based on the observation that scientific applications in the various "mésocentres" are often the same depending on the sector (bio-info, physics, etc.). The CSAN platform will allow software authors to deposit versions of their sources (or to point to versioned repositories). Then, the CSAN expert group will process this code and analyze its compilation and deployment configuration, via integration and continuous development methods, in order to make it effortlessly accessible and installable on different operating systems and to as many users and infrastructures as possible. This application

catalog will be accessible through a web portal and an open API for integration in the various "mésocentres".

A first version of the portal has been deployed in 2021 on an APC lab-specific infrastructure. This version served as a proof of concept, and a production version is being deployed on a multi-site infrastructure between the IPHC in Strasbourg and the Montpellier "mésocentre". The end of the deployment is planned for summer 2022, when the infrastructure test phase will be completed.

This project is led by a subgroup of the ComputeOps project composed of about ten people from different institutes (IN2P3, IP2I, INRAE and University). It was the subject of two publications during the JCAD 2021 and JRES 2022 (see publications).

Convergence of container orchestrators and job managers

The advent of container orchestrators has changed the traditional compute landscape where job schedulers (scheduler) installed on data centers interface with the user. Container orchestrators such as Kubernetes were designed to manage containerized workloads, primarily providing resource management for microservice containers such as web services. But the convergence of technologies in the area of High Performance Data Analysis (HPDA) suggests that container orchestrators could come to replace traditional solutions. Currently, container orchestrators lack task scheduling mechanisms (no management of potential constraints and links between tasks, nor of input and output data), and multi-user management.

Workflow managers manage data processing pipelines and allow interaction with different types of resources (HPC cluster, cloud, container). This brick upstream of the computation stage can communicate via operators with container orchestrators (API) and job schedulers (interface to be developed). In this context, a study with the workflow manager Airflow has been started in 2021. This study is carried out internally and the first complete prototype (multi-resource) was completed in 2022. This study is based on shared equipment, of which two components present at IN2P3 are used (HPC cluster and cloud). But it lacks the availability of a Kubernetes cluster in production (access planned for this year at CC-IN2P3).

SWOT self-analysis

- Strengths: very attractive R&D topic
- Weaknesses: several members have left and have not been renewed.
- Opportunities: the CSAN project is part of the Open Science approach with a direct connection to Software Heritage
- Threats: the volatility of container solutions (Singularity was bought by Sylabs and then there was a fork in the open source community with Apptainer) requires constant R&D monitoring