

Orchestration of optimized GPU containers using a cloud management platform in a hybrid environment

Nowadays, with the advent of container technologies like Docker or Singularity there is no need to have applications installed on any scientific resource; instead, they can be encapsulated inside a container to run in any operating system, CPU architecture, interconnect, and even leverage specific hardware accelerators like GPUs.

With a proper container definition, hardware or GNU/Linux distro-related incompatibilities can be seamlessly avoided, while retaining native or near-native performance depending on the host and the container configuration.

With ATRIO Composable Cloud (ACC) we provide a hardware and resource manager agnostic computing platform that is able to orchestrate any workload and use the bare-metal clusters in CETA-Ciemat (Slurm, OpenStack Nova or Ironic deployments), public cloud service providers (AWS, Azure, Google Cloud, etc.) and potentially any compute resource, like the European Open Science Cloud.

In this poster, we show how ACC can orchestrate optimized containers using a variety of computing resources and deployments provided by CETA-Ciemat and some public cloud providers. We also demonstrate how simple is to deploy a machine learning workflow using Tensorflow with MPI, CUDA and low-latency interconnect (Infiniband) capabilities across multiple heterogeneous clusters.

Primary authors: Mr GÓMEZ-MARTÍN, César (Atrio); Mr ALONSO-GÓMEZ, Javier (Atrio); Mr DÍAZ, Guillermo (CETA-Ciemat); Mr PARDO, Alfonso (CETA-Ciemat); Mr FRANCO, José M. (CETA-Ciemat)

Presenter: Mr DÍAZ, Guillermo (CETA-Ciemat)

Session Classification: IBERGRID Contributions

Track Classification: R&D for computing services, networking, and data-driven science.