



Contribution ID: 82

Type: **Lightning Talk 8 mins**

Bridging Cloud and HPC towards High Performance Data Analytics for climate science

Wednesday, 21 September 2022 12:25 (8 minutes)

D. Elia¹, F. Antonio¹, C. Palazzo¹, A. D'Anca¹, S. Fiore² and G. Aloisio^{1,3}

¹ Euro-Mediterranean Center on Climate Change (CMCC) Foundation, Lecce, Italy

² University of Trento, Trento, Italy

³ University of Salento, Lecce, Italy

The Big Data revolution started at the beginning of this century has been propelled also by the advent of cloud computing solutions, which provided an efficient and cost-effective model for accessing resources on-demand according to the application workload and functional requirements. These new technologies have been gradually exploited in several scientific domains to address the issues associated with large data volumes, besides the more traditional use of High Performance Computing (HPC), which is still required for several compute-intensive applications. The next natural step in this evolution concerns the integration of the Big Data (cloud-based) and HPC software ecosystems for supporting High Performance Data Analytics (HPDA) scientific scenarios at extreme scale. However, the two software ecosystems rely on very different service usage models

and target different application requirements, making their mixed usage complicated. Software containers can represent the layer for supporting software portability and transparent deployment of scientific HPDA solutions over multiple platforms, allowing developers to bundle the application and all its dependencies (including data dependencies) into a single software image. In this regard, the recent emergence of HPC-friendly container technologies (e.g., udocker, Singularity, Sarus) can actually enable the use of this model also on HPC infrastructures, thus providing a bridge between Cloud and HPC-based solutions and enabling new paradigms such as HPC as a Service (HPCaaS). In the context of the EGI-ACE project, a HPC pilot concerning the use of data science, management and HPDA solutions for climate science applications is being developed. The pilot is aimed at understanding how containerization technologies can support the integration of cloud and HPC infrastructures to support large-scale data analytics and management. This contribution presents the container-based solutions explored and implemented in the context of the HPC pilot towards transparent and portable deployment of HPDA solutions for climate science on top of the resources made available in the EGI infrastructure.

Any relevant links

Topic

A Federated Compute Continuum

Primary authors: ELIA, Donatello (CMCC Foundation); ANTONIO, Fabrizio (CMCC Foundation); PALAZZO, Cosimo (CMCC Foundation); D'ANCA, Alessandro (CMCC Foundation); FIORE, Sandro (University of Trento, Trento, Italy); Prof. ALOISIO, Giovanni (Advanced Scientific Computing Division, Centro Euro-Mediterraneo sui Cambiamenti Climatici)

Presenters: ELIA, Donatello (CMCC Foundation); ANTONIO, Fabrizio (CMCC Foundation); PALAZZO, Cosimo (CMCC Foundation); D'ANCA, Alessandro (CMCC Foundation); FIORE, Sandro (University of Trento, Trento, Italy); Prof. ALOISIO, Giovanni (Advanced Scientific Computing Division, Centro Euro-Mediterraneo sui Cambiamenti Climatici); FERNANDEZ, Enol (EGL.eu)

Session Classification: EGI-ACE Lightning Talks: Compute continuum use cases

Track Classification: A Federated Compute Continuum