



Contribution ID: 44

Type: **Presentation short (15 min)**

A HPC as a Service solution for climate data analytics on the EGI infrastructure

Thursday, 21 October 2021 12:25 (15 minutes)

The deluge of data started at the beginning of this century has caused a profound transformation in the way scientific discovery is carried out. In several domains, such as climate science, scientific advances now rely on technologies and software solutions from both the HPC and Big Data landscapes. However, being able to efficiently exploit HPC infrastructures for running scientific data analysis is not easy. A unified model that also allows the deployment on HPC of the same services already exploited in the cloud can pave the way for a wider range of opportunities in the scientific community, further fueling the adoption of the HPC as a Service (HPCaaS) paradigm.

In this respect, software containers are good candidates for supporting portability and deployment of data analytics frameworks over multiple platforms. Thanks to the recent development of HPC-friendly container technologies (e.g., udocker, Singularity, Sarus), it has now become possible for scientists to exploit the benefits of this model also on HPC infrastructures.

Containers can allow encapsulation of the application, together with its dependencies, into a single and portable image file. Nevertheless, several issues must be addressed to effectively exploit these technologies for operational scientific services and applications. Specifically, some of the main issues concern possible performance degradation and the integration with services, including those distributed across different infrastructures.

In the European Open Science Cloud (EOSC) context, the ENES Climate Analytics Service (ECAS) is a central component of the ENES Data Space set up in the EGI-ACE project, which aims to provide an open and cloud-enabled data science environment for climate scientists. In this environment, the Ophidia HPDA framework represents the core computing engine of the ECAS service and it can greatly benefit from the exploitation of HPC resources for running data analysis applications and workflows.

This work presents the container-based approach investigated in the EGI-ACE project, for transparent and portable deployment of ECAS on top of the HPC resources made available in the EGI infrastructure.

“Fabrizio Antonio was awarded a Master’s Degree with first-class honors in Computer Engineering from the University of Salento, Faculty of Engineering, in April 2016, with a thesis on High Performance Computing and Big Data. In May 2016, he joined the Data Science and Learning Research Team within the Advanced Scientific Computing (ASC) Division of CMCC.

His research activities focus on distributed data management and high-performance data analytics and mining for eScience in the context of climate change.”

Most suitable track

Innovating services together

By submitting my abstract, I agree that my personal data is being stored in accordance to conference Privacy Policy

Primary authors: Mr ELIA, Donatello (CMCC Foundation, Lecce, Italy; University of Salento, Lecce, Italy); Mr ANTONIO, Fabrizio (CMCC Foundation, Lecce, Italy); Mr GIANNOTTA, Andrea (CMCC Foundation, Lecce, Italy); Mr PALAZZO, Cosimo (CMCC Foundation, Lecce, Italy); Mr D'ANCA, Alessandro (CMCC Foundation, Lecce, Italy); Prof. FIORE, Sandro (University of Trento, Trento, Italy); Prof. ALOISIO, Giovanni (CMCC Foundation, Lecce, Italy; University of Salento, Lecce, Italy)

Presenter: Mr ANTONIO, Fabrizio (CMCC Foundation, Lecce, Italy)

Session Classification: Innovating Services Together - Workshop