

Cloud Computing Needs for Earth Observation Data Analysis: EGI and the European Open Science Cloud

Björn Backeberg¹, Yin Chen¹, *Tiziana Ferrari¹, Pedro Gonçalves⁴, Paolo Mazzetti³, Anabela Oliveira², Diego Scardaci¹, Gergely Sipos¹

1. EGI Foundation, 2. LNEC, 3. CNR, 4. Terradue

Over recent years, the vision of Open Science has emerged as a new paradigm of transparent, data-driven science capable of accelerating competitiveness and innovation. The embodiment of this vision in Europe is the European Open Science Cloud (EOSC), first proposed by the European Commission in April 2016 as part of the Communication on the ‘European Cloud Initiative’, one of the pillars of the Digital Single Market Strategy. The EOSC puts into practice the European vision for Open Innovation, Open Science and Open to the World by bringing together services and research products such as computing, storage, data, publications, software and workflows from national and international research infrastructures, research performing organisations, collaborations and projects.

Within EOSC and the EOSC-hub H2020 project, the European GEOSS community of the EuroGEOSS initiative, aims at demonstrating the benefit of Earth Observation for the society. Earth Observation data is being used to create the new knowledge needed to support informed decision-making for sustainable development and policy making. In addition, the GEOSS is contributing to EOSC by providing mediation and harmonisation services for accessing information and knowledge from in-situ and Earth Observation data, opening access to models for the generation of indicators.

EOSC will be instrumental in providing the large computation and storage capabilities needed to run scientific models at local, national, regional and potentially global scale. Within EOSC, EGI is the European digital infrastructure responsible for providing a distributed platform for scientific computing. The EGI Federation is the publicly funded e-infrastructure that brings together hundreds of data centres worldwide and also includes the largest community of research clouds in Europe with more than 20 cloud providers across 12 European countries offering IaaS cloud and storage services. The EGI cloud infrastructure is based on open standards allowing a seamless user experience over the cloud sites belonging to the federation through standardised interfaces. Advanced services are available to manage computing services and datasets in a distributed environment hiding the complexity of the geographical distribution to the developers of scientific applications and services.

The EGI Federated Cloud will give scientists the possibility to access and run analytic models in a transparent way, without the need of a local infrastructure hosting data and computing facilities. In this presentation, we will introduce the EGI technical solutions which provide the following capabilities: (1) the federation of services, data and resources; (2) the provision of geographically distributed large capacity computing resources and specialized computing facilities; (3) Jupyter notebooks data analysis capabilities integrated with federated Authentication and Authorization and the EGI Cloud; and (4) the support of added-value services, like the the OpenCoastS - for the production on-demand circulation forecast systems, the GEOSS Web Portal - for discovering and accessing GEOSS data from than 170 data systems globally, the Geohazards Thematic Exploitation Platform GEP - providing access to a set of on-demand terrain motion services supporting interferogram generation, co-seismic displacement mapping, landslide rapid mapping and landslide displacement field monitoring with Copernicus Sentinel-1 and Sentinel-2 data.

Keywords: Open Science, EOSC, EGI, Earth observation data analysis