

EGU22-5931, updated on 06 Jul 2022

<https://doi.org/10.5194/egusphere-egu22-5931>

EGU General Assembly 2022

© Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.



C-SCALE: A new Data and Compute Federation for Earth Observation

Christian Briese¹, **Charis Chatzikyriakou**¹, Diego Scardaci², Zdeněk Šustr³, Enol Fernández², Björn Backeberg⁴, and Elonora Testa²

¹EODC Earth Observation Data Centre for Water Resources Monitoring GmbH, Vienna, Austria

²EGI Foundation, Amsterdam, The Netherlands

³CESNET, z. s. p. o., Prague, Czech Republic

⁴Deltares Foundation, Delft, The Netherlands

Through the provision of massive streams of high-resolution Earth Observation (EO) data, the EU Copernicus programme has established itself globally as the predominant spatial data provider. These data are widely used by research communities to monitor and address global challenges, such as environmental monitoring and climate change, supporting European policy initiatives, such as the Green Deal and others. To date, there is no single European data sharing and processing infrastructure that serves all datasets of interest, and Europe is falling behind international developments in Big Data analytics and computing.

The C-SCALE (Copernicus - eoSC AnaLytics Engine, <https://c-scale.eu>) project federates European EO infrastructure services, such as ESA's Sentinel Collaborative Ground Segment, the Copernicus DIASes (Data and Information Access Services under the EC), independent nationally-funded EO service providers, and European Open Science Cloud (EOSC) e-infrastructure providers. It capitalises on EOSC's capacity and capabilities to support Copernicus research and operations with large and easily accessible European computing environments. The project will implement and publish the C-SCALE Federation in the EOSC Portal as a suite of complementary services that can be easily exploited. It will consist of a Data Federation, a service providing access to a large EO data archive, a Compute Federation, and analytics tools.

The C-SCALE Data Federation aims at making EO data providers under EOSC findable, their metadata databases searchable, and their product storage accessible. While a centralised, monolithic, complete Copernicus data archive may not be feasible, some organisations maintain various archives for limited areas of their interest. C-SCALE, therefore, integrates these heterogeneous resources into a "system of systems" that will offer the users an interface that, in most cases, provides similar functionality and quality of service as a centralised, monolithic data archive would. The federation is built on existing technologies, avoiding redundancy and replication of functions and not disrupting existing usage patterns at participating sites, instead only adding a simple layer for improved discovery and seamless access.

At the same time, the C-SCALE Compute Federation provides access to a wide range of computing providers (IaaS VMs, container orchestration platforms, HPC and HTC systems) to enable the

analysis of Copernicus and EO data under EOSC. The design of the federation allows users to deploy their applications using federated authentication mechanisms, find their software under a common catalogue, and have access to data using C-SCALE Data Federation tools. The federation relies on existing tools and services already compliant with EOSC, thus facilitating the integration into the larger EOSC ecosystem.

By making such a scalable Big Copernicus Data Analytics federated services available through EOSC and its Portal and linking the problems and results with experience from other research disciplines, C-SCALE helps to support the EO sector in its development. By abstracting the set-up of computing and storage resources from the end-users, it enables the deployment of custom workflows to generate meaningful results quickly and easily. Furthermore, the project will deliver a blueprint, setting up an interaction model between service providers to facilitate interoperability between commercial and public cloud infrastructures.