ECASLab: a user-friendly, integrated environment for scientific data analytics and visualization in the European Open Science Cloud landscape

Sandro Fiore (1), Alessandro D'Anca (1), Donatello Elia (1), Cosimo Palazzo (1), Alessandra Nuzzo (1), Tobias Weigel (2), Sofiane Bendoukha (2), Giovanni Aloisio (1,3)

(1) Euro-Mediterranean Center on Climate Change (CMCC Foundation), Lecce, Italy (sandro.fiore@cmcc.it), (2) Deutsches Klimarechenzentrum (DKRZ), Hamburg, Germany, (3) University of Salento, Italy

ECASLab is a user-friendly scientific data analysis environment integrating data and analysis tools to support scientists in their daily research activities. The environment joins the features of the ENES Climate Analytice Service (ECAS) with a large set of Python libraries for running data manipulation, analysis, and visualization.

ECAS (an evolution of the Ophidia big data analytics framework in the EOSC landscape) is the main component of the ECASLab. It represents a complete software stack developed for the analysis of large multidimensional data ('datacubes') in several eScience domains (e.g. climate change).

The ECASLab environment integrates the following services:
- a multi-node ECAS framework instance with WPS-enabled interface accessible through the ECAS Terminal and any WPS-compliant client;
- a JupyterHub instance providing the user with a web-based system for creating, executing and sharing Jupyter notebooks (Python-based) supporting live-coding and visualization;
- a GUI for interactive workflow editing;
- a monitoring system based on Grafana.

Moreover, the environment provides access to a set of data (also through a THREDDS Data Server), a number of example Jupyter notebooks and real-world workflows describing indicators from several use cases. Experiment output can be exported in the user space or on the publicly accessible data server, whereas JupyterHub provides the features to update files and navigate the file system. The features of the ECAS framework can be directly exploited in the notebooks to run data analytics tasks on big datasets and plot the results on charts and maps using wellknown Python libraries in a user-friendly interface. PyOphidia - the Python bindings - allows to easily interact with ECAS and other Python-based modules (e.g. Matplotlib, NumPy, etc.).

As part of the H2020 EOSC-Hub project, the ECASLab environment will be running both at CMCC and DKRZ and it will be hosting two major instances of the ENES Climate Analytics Service to serve climate scientists as well as other research communities interested in the implementation of large-scale scientific data analytics workflows.