EOS workshop



Contribution ID: 9

Type: not specified

EOS as storage back-end for JRC data science

Monday 3 February 2020 10:45 (15 minutes)

The Joint Research Centre (JRC) of the European Commission has set up the JRC Earth Observation Data and Processing Platform (JEODPP) as an infrastructure to enable the JRC projects to process and analyze big data, extracting knowledge and insights in support of EU policy making. The main focus is related to geospatial data, but has been extended to other data domains. EOS is the main storage component of the platform, operationally used since mid 2016, maintained and extended with support by the CERN EOS team.

The JEODPP is actively used by more than 40 JRC projects as platform for data science, covering a wide range of data analysis activities. In order to serve the growing needs for data storage and processing capacity by the JRC projects, the platform has been extended in 2019. It currently consists of the EOS system as storage back-end with a total gross capacity of 15.5 PB, and service nodes with a total of 2000 CPU cores.

As main changes in 2019, the EOS service has been migrated to the QuarkDB namespace, and half of the service nodes have been migrated to FUSEX client. The presentation will give an overview about the implemented platform, the current status, experiences made, and issues identified with EOS as main storage back-end of the JRC data science platform.

Author: Mr BURGER, Armin (European Commission - Joint Research Centre)

Co-authors: Mr EYRAUD, Franck (Contractor of European Commission); Mr TOGNOLI, Pier Valerio (European Commission - Joint Research Centre); Mr SCAVAZZON, Marco (European Commission - Joint Research Centre)

Presenter: Mr BURGER, Armin (European Commission - Joint Research Centre)

Session Classification: EOS Production Services