Contribution ID: **99** Type: **Presentation**

Next Generation Data Management Services: the eXtreme DataCloud project

Wednesday, 10 October 2018 14:45 (15 minutes)

The development of new scalable technologies for federating storage resources and managing data in the current and next generation e-Infrastructures deployed in Europe, such as the European Open Science Cloud (EOSC), the European Grid Infrastructure (EGI), the Worldwide LHC Computing Grid (WLCG) is the aim of the eXtreme-DataCloud (XDC) H2020 funded project.
The high-level objective of the project is the semi or fully automated placement of scientific data in the Exabyte region exploiting the resources made available by the modern, cloud based, e-Infrastructures.
XDC is focused on providing enriched high-level data management services to access heterogeneous storage resources and services. It enables scalable data processing on distributed infrastructures using established interfaces and allowing the use of legacy applications without the need for rewriting them from scratch.
The project will address high-level topics that include: i) federation of storage resources with standard protocols, ii) smart caching solutions to access transparently data stored in remote locations, iii) policy driven data management based on Quality of Service, iv) data lifecycle management, v) metadata handling and manipulation, vi) data preprocessing during ingestion, vii) optimized data management based on access patterns.
The solutions implemented by the XDC project are targeted to the real life use cases provided by different scientific communities represented within the project, such as: astrophysics (CTA), Photon Science (European X-FEL), High Energy Physics (WLCG), Life Science (LifeWatch) and Medical Science (ECRIN).
The XDC solutions are based on already well established data management components like dCache, FTS, EOS, the INDIGO PaaS Orchestrator and ONEDATA, just to mention some of them. These services will be enriched with new functionalities and organized in a coherent architecture to address the user requirements.
For a better understanding of the nature and the scope of the project, the high level architecture overview and related interfaces specification will be presented and described. Moreover, implementation examples on specific use cases will be presented.

Type of abstract

Presentation

Summary

We will present the H2020 funded eXtreme-DataCloud project. XDC is focused on providing enriched high-level data management services to access heterogeneous and distributed storage resources. It enables scalable data processing on the e-Infrastructures using established interfaces. XDC is a use case driven project and several communities are represented within its consortium: astrophysics (CTA), Photon Science (European X-FEL), High Energy Physics (WLCG), Life Science (LifeWatch) and Medical Science (ECRIN). Implementation examples on specific use cases will be presented.

Primary authors: COSTANTINI, Alessandro (INFN); CESINI, Daniele (INFN)

Co-authors: OHMANN, Christian (ECRIN); DUMA, Doina Cristina (INFN-CNAF); AGUILAR, Fernando (CSIC); Dr DONVITO, Giacinto (INFN); Dr DUTKA, Lukasz (CYFRONET); VILJOEN, Matthew (EGI.eu); KEEBLE, Oliver (CERN); FUHRMANN, Patrick (DESY); LEMRANI, Rachid (IN2P3-CC); BATTAGLIA, Serena (ECRIN (European Clinical Research Infrastructure Network)); POIREAU, Vincent (IN2P3-LAPP)

Presenters: COSTANTINI, Alessandro (INFN); CESINI, Daniele (INFN)

Session Classification: Data Management Services

Track Classification: Area 3. Computing and Virtual Research Environments