

The eXtreme-DataCloud project: advanced data management services for distributed e-infrastructures

Thursday, 4 April 2019 14:30 (30 minutes)

The eXtreme-DataCloud project (XDC) is a software development initiative aimed at implementing data management scalable services to address the following high level topics: policy driven data management based on Quality-of-Service, Data Life-cycle management, storage federations creation, smart placement of data with caching mechanisms, meta-data with no predefined schema handling, execution of pre-processing applications during ingestion, data management and protection of sensitive data in distributed e-infrastructures. The project is driven by user communities belonging to different scientific domains: High Energy Physics (WLCG), Astronomy (CTA and LSST), Photon and Life Science (XFEL and LifeWatch), Medical research (ECRIN). XDC is funded by the European Commission under the Horizon 2020 framework program, it started in November 2017 and its first major release was launched at the end of 2018. This release is based on a toolbox composed by well known, production quality services that have been enriched with new functionalities. The list of services include dCache, ONEDATA, EOS, FTS, Indigo-Orchestrator, Indigo-CDMI server and Dynafed. All the newly implemented functionalities can be easily plugged into the existing e-infrastructures but, given the use of standard protocols and authorization mechanisms, can be used also as building blocks of the new generation scientific infrastructures. This contribution will introduce the project, its overall architecture and the first release main features. Some use cases addressed by the project developments will also be presented.

Primary authors: COSTANTINI, Alessandro (INFN-CNAF); OHMANN, Christian (ECRIN); CESINI, Daniele (INFN-CNAF); DUMA, Doina Cristina (INFN - CNAF); AGUILAR GOMEZ, Fernando (IFCA); DONVITO, Giacinto (INFN); DELL'AGNELLO, Luca (INFN); DUTKA, Lukasz (AGH/CYFRONET); VILJOLEN, Matthew (EGI Foundation); KEEBLE, Oliver (CERN); FUHRMANN, Patrick (DESY/dCache.org); LEMRANI, Rachid (IN2P3); BATTAGLIA, Serena (ECRIN); POIREAU, Vincent (IN2P3-LAPP)

Presenter: CESINI, Daniele (INFN-CNAF)

Session Classification: Data Management & Big Data

Track Classification: Data Management & Big Data