

## Achievements of the eXtreme-DataCloud project

*Friday, 28 August 2020 10:00 (30 minutes)*

The eXtreme-DataCloud project (XDC) is a software development initiative, funded by the European Commission under the H2020 framework program, aimed at developing and implementing data management scalable services. The project addresses high level topics like: policy driven data management based on storage Quality-of-Services, Data Life-cycle management, storage federations creation, smart placement of data with caching mechanisms, metadata 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 service catalogue is based on a Toolbox made of already existing, production quality components and services that the project enriched with new features and functionalities as requested by the user communities represented in the Consortium and belonging to a wide range of scientific domains: High Energy Physics (WLCG), Astronomy (CTA and LSST), Photon and Life Science (XFEL and LifeWatch), Medical research (ECRIN). The list of services include dCache, Onedata, EOS, FTS, Indigo-Orchestrator, Indigo-CDMI server, RUCIO and Dynafed. All of them have been organized in a coherent architecture that can be easily plugged into the current distributed e-Infrastructures. This is possible thanks to the adoption of established standard for protocols and authentication methods.

XDC started in November 2017 and is now close to its end. Two releases have been produced: XDC-1, code-named Pulsar, at the beginning of 2019 and XDC-2, codenamed Quasar, released at the end of 2019. This work presents the new features introduced with these two releases, how they impacted on the tools composing the XDC-Toolbox and, in general, the main achievements of the project.

Some of the developments that will be presented are: the OpenIDConnect support for all the tools, Data Caching methods based on xrootd and http protocols to support geographic deployment of distributed caches and the inclusion of diskless sites in the e-infrastructures, bulk storage QoS transitions support, new Onedata features for metadata management, sensitive data handling and QoS change, storage events notifications to support complex workflows.

**Primary author:** CESINI, Daniele (INFN-CNAF)

**Co-authors:** Dr COSTANTINI, Alessandro (INFN-CNAF); Dr OHMANN, Christian (ECRIN); DUMA, Doina Cristina (INFN - CNAF); Dr LUKASZ, Dutka (AGH/Cyfronet); Dr AGUILAR, Fernando (UC/IFCA); Mr DONVITO, Giacinto (INFN); Mr DELL'AGNELLO, Luca (INFN); VILJOEN, Matthew (EGI Foundation); Dr KEEBLE, Oliver (CERN); Dr FUHRMANN, Patrick (DESY/dCache.org); Dr LEMRANI, Rachid (CNRS); Dr BATTAGLIA, Serena (ECRIN); Dr POIREAU, Vincent (CNRS)

**Presenter:** CESINI, Daniele (INFN-CNAF)

**Session Classification:** Data Management Session

**Track Classification:** Data Management & Big Data