

Storage Management in INDIGO

Wednesday, 11 November 2015 17:00 (20 minutes)

The INDIGO DataCloud project is set out to develop a data/computing platform targeting scientific communities, to enhance usefulness of existing e- infrastructures [1]. The developments in the storage are focus on two levels. On the IaaS level, QoS in storage will be addressed, by implementing a standardised extension to the CDMI standard, which enables management of storage quality e.g. access-latency, retention-policy, migration-strategy or data-lifecycle. This is closely related with intelligent identity management to harmonise access via different protocols, such as gridFTP, sftp and CDMI. This will allow to use CDMI to manage QoS of data that is accessible via gridFTP or sftp.

On the PaaS level, INDIGO DataCloud will provide flexible data federation functionality, enabling users to transparently store and access their data between heterogeneous infrastructures. DataCloud will provide unified API's for data management based on state of the art standards, allowing both users and application developers to easily integrate DataCloud high level data management functionality into their use cases. One of the key features of this solution will be optimization of data access in various scenarios, which will include automatic pre-staging, maximum bandwidth usage via parallel transfers and enabling instant access to remote data through streaming. Furthermore, the layer will provide information to Cloud orchestration services allowing placing the computations in the sites where the data is already staged, or where it can be delivered efficiently.

[1]<https://www.indigo-datacloud.eu/>

Primary authors: Dr DONVITO, Giacinto (INFN); Dr DUTKA, Lukasz (CYFRONET); HARDT, Marcus (KIT-G); FUHRMANN, Patrick (DESY)

Presenters: HARDT, Marcus (KIT-G); FUHRMANN, Patrick (DESY)

Session Classification: Open Data workshop