dCache, managing Quality of Service in Cloud Storage Infrastructures

Tuesday, 7 March 2017 14:40 (0:20)

Content

For the previous decade, high performance, high capacity Open Source storage systems have been designed and implemented, accommodating the demanding needs of the LHC experiments. However, with the general move away from the concept of local computer centers, supporting their associated communities, towards large infrastructures, providing Cloud-like solutions to a large variety of different scientific groups, storage systems needed to adjust their capabilities in many areas, as there are federated identities, non authenticated delegation to portals or platforms, modern sharing and user defined Quality of Storage.

This presentation will give an overview on how dCache is keeping up with modern Cloud storage requirements by partnering with EU projects, which provide the necessary contact to a large set of Scientific Communities. Regarding authentication, there is no strict relationship anymore between the individual scientist, the scientific community and the infrastructure, providing resources. Federated identity systems like SAML or "OpenID Connect" are growing into the method-of-choice for new scientific groups and are even sneaking their way into HEP. Therefor, under the umbrella of the INDIGO-DataCloud project, dCache is implementing those authentication mechanisms in addition to the already established ones, like username/password, Kerberos and X509 Certificates. To simplify the use of dCache as back-end of scientific portals, dCache is experimenting with new anonymous delegation methods, like "Macaroons", which the dCache team would like to introduce in order to start a discussion, targeting their broader acceptance in portals and at the level of service providers. As the separation between managing scientific mass data and scientific semi-private data, like publications, is no longer strict, large data management systems are supposed to provide a simple interface to easily share data among individuals or groups. While some systems are offering that feature through web portals only, dCache will show that this can be provided uniquely for all protocols the system supports, including NFS and GridFTP. Furthermore, in modern storage infrastructures, storage media, and consequently the quality and price of the request storage space are no longer negotiated with the responsible system administrators but dynamically selected by the end user or by automated computing platforms. The same is true for data migration between different qualities of storage. To accommodate this conceptual change, dCache is exposing it's entire data management interface through a RESTful service and a graphical user interface. The implemented mechanisms are following the recommendation of the corresponding working groups in RDA and SNIA and are agreed-upon with the INDIGO-DataCloud project to be compatible with similar functionalities of other INDIGO provided storage systems.

Summary

This presentation will give an overview on how dCache is keeping up with modern Cloud storage requirements by partnering with EU projects, especially INDIGO-DataCloud, providing the necessary contacts to a large set of Scientific Communities. We will elaborate on the integration of modern AAI mechanisms, like OpenID Connect and Macaroons, the close integration of dCache services using REST, an improved graphical interface for data access and manipulation and the first implementation of the new Quality of Service in storage interface, defined by the IaaS team of INDIGO.

Primary author(s) :   Dr. FUHRMANN, Patrick (DESY/dCache.org);   Dr. MILLAR, Paul (DESY);   Mrs. YASAR, Sibel (DESY);   Mr. MKRTCHYAN, Tigran (DESY);   Dr. ASHISH, Anupam (DESY);   Dr. ADEVEMI, Olufemi (DESY);   Dr. SAHAKYAN, Marina (DESY);   Dr. ROSSI, Albert (FNAL);   Dr. LITVINTSEV, Dmitry (FNAL);   Dr. BEHRMANN, Gerd (NDGF);   Mr. STAREK, Juergen (DESY)

Presenter(s) :   Dr. FUHRMANN, Patrick (DESY/dCache.org)

Session Classification :   Data Management & Big Data

Track Classification :   Data Management & Big Data