

Quality of Service in storage and the INDIGO-DataCloud project.

Tuesday, 15 March 2016 14:00 (30 minutes)

The pressure to provide cheap, reliable and unlimited cloud storage space in the commercial area has provided science with affordable storage hardware and open source storage solutions offering low maintenance costs and tuneable performance as well as durability properties, resulting in different cost models per storage unit. Those models, already introduced by WLCG a decade ago (disk vs tape) are now offered by large companies like Amazon (Bock Storage, S3 and Glacier) or Google (Standard, Durable Reduced, Cloud Storage Nearline). Industry appliances or software stacks (e.g. HPSS) offer similar storage properties for your locally installed computer centre storage space.

However, other than with SRM for WLCG, those offered storage quality properties don't follow a common description or specification and are hard to compare programatically. Moreover they aren't even close to a common way of been negotiated between the requesting client and the providing storage technology, which would be a prerequisite for federating different public and private storage services as planned by EGI and EUDAT.

To fill this gap, the INDIGO-DataCloud project is proposing a process to agree on common semantics in describing QoS attributes in storage in a consistent way, independently of the used API or protocol.

The process involves gathering uses-cases from scientific communities and creating working groups within international organisations, like RDA, OGF and SNIA to further discuss possible solutions with other interested parties like EUDAT and EGI. The activity has already been presented at the Paris RDA meeting and an RDA interest group is being prepared.

In a second step, based on feedback received, INDIGO will propose an implementation of the defined semantics to steer quality of service in storage. As a proof of concept INDIGO will implement the proposed solution in storage systems used within the INDIGO project, like dCache, StoRM, some typical industry products and a selected public storage cloud.

For within INDIGO-DataCloud a consistent QoS specification to uniquely describe the different distributed storage infrastructure components is essential to allow the INDIGO platform layer to broker the most appropriate endpoint for each individual data store, replication or access request.

We are presenting our work at ISGC, in order to receive feedback from interested communities in the Asian and Australian area as we believe that the approach is extremely useful beyond the activities within INDIGO-DataCloud.

This presentation will report on the goals achieved so far and on our next steps.

Primary authors: Dr ERTL, Benjamin (KIT); Dr DONVITO, Giacinto (INFN/BARI); Dr HARDT, Marcus (KIT); Dr FUHRMANN, Patrick (DESY/dCache.org); Dr MILLAR, Paul (DESY)

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

Session Classification: Data Management Session I

Track Classification: Data Management