Contribution ID: **534**   Type: **Parallel**

Evaluation of benefits of a three tier data model for WLCG analysis

Tuesday 22 May 2012 17:00 (25 minutes)

One of the most crucial requirement for online storage is the fast and efficient access to data.

Although smart client side caching often compensates for discomforts like latencies and server disk congestion, spinning disks, with their limited ability to serve multi stream random access patterns, seem to be the cause of most of the observed inefficiencies.

With the appearance of the different variants of solid state disks (SSD), this deficiency could be overcome, however, replacing the entire experiment data repositories by SSDs is not feasible in the foreseeable future.

Moreover, spinning disks are still appropriate media for controlled streaming applications.

Assuming a deployment of a mixture of media, like spinning disks, SSDs and tape, at a site, the authors argue for the introduction of a three tier media structure within a single storage system with automatic transitions, based on usage patterns, in contrast to interlinking and maintaining different mediatypes in different systems with external procedures taking care of proper data placement.

The feasibility of the suggested approach is studied, using the analysis of access logs of the DESY WLCG Tier II storage elements, hosting the largest part of the data
to be analyzed by the CMS and ATLAS Collaborations.

Finally we will report on a prototype implementing of the three tier media structure into dCache, a storage technology widely used in WLCG.

Authors: OZEROV, Dmitry (Deutsches Elektronen-Synchrotron (DE)); Dr FUHRMANN, Patrick (DESY)

Presenters: OZEROV, Dmitry (Deutsches Elektronen-Synchrotron (DE)); Dr FUHRMANN, Patrick (DESY)

Session Classification: Distributed Processing and Analysis on Grids and Clouds

Track Classification: Distributed Processing and Analysis on Grids and Clouds (track 3)