Contribution ID: **55**   Type: **not specified**

Large-scale data services for science: present and future challenges

Monday, 28 September 2015 14:10 (30 minutes)

CERN IT operates the main storage resources for data taking and physics analysis mainly via three system: AFS, CASTOR and EOS. Managed disk-storage amounts to about 100 PB (with relative ratios 1:10:30). EOS deploys disk resources evenly across the two CERN computer centres (Meyrin and Wigner). The physics data archive (CASTOR) contains about 100 PB so far. We are also providing sizeable resources for general IT services most notably OpenStack and NFS clients. This is implemented with a Ceph infrastructure for a total capacity of ~1 PB (which we scaled up for testing by a factor of 10). Recently a new service, CERNBOX, has been added to provide file synchronisation and sharing functionality (more than 2000 users).

We will describe the operational experience and plans for the future
- Data services for LHC data taking (new roles of CASTOR and EOS)
- Experience in deploying EOS across multiple sites
- Experience in coupling commodity and home-grown solution (e.g. Ceph disk pools for AFS, CASTOR and NFS)
- Future evolution of these systems in the WLCG realm and beyond, especially with the promising field of cloud synchronisation systems

Primary author: Dr LAMANNA, Massimo (CERN)

Presenter: Dr LAMANNA, Massimo (CERN)