Contribution ID: **83**   Type: **Oral presentation to parallel session**

Disk storage at CERN: handling LHC data and beyond

Monday 14 October 2013 14:16 (20 minutes)

Data Storage and Services (DSS) group at CERN stores and provides access to the data coming from the LHC and other physics experiments. We implement specialized storage services to provide tools for an optimal data management, based on the evolution of data volumes, the available technologies and the observed experiment and users usage patterns. Our current solutions are CASTOR for highly-reliable tape-backed storage for heavy-duty Tier-0 workflows and EOS for disk-only storage for full-scale analysis activities.

CASTOR has been the main physics storage system at CERN since 2001 and successfully caters for the LHC experiments' needs, storing 90 PB of data and more than 350 M files. During the last LHC run CASTOR was routinely storing 1 PB/week of data to tape. CASTOR is currently evolving towards a simplified disk layer in front of the tape robotics, focusing on recording the primary data from the detectors.

EOS is now a well established storage service used intensively by the four big LHC experiments, holding over 15 PB of data and more than 130M files (30 PB usable disk space expected at the end of the year). Its conceptual design based on multi-replica and in-memory namespace make it the perfect system for data intensive workflows and its usage will expand via a shared instance for non-LHC experiments. In the short term EOS usage will absorb most of the newly installed capacity at CERN and expand via a shared instance for the non-LHC experiments. An additional challenge will be to run this service across two geographically different sites (CERN Geneva and Budapest).

LHC-Long Shutdown 1 presents a window of opportunity to shape up both of our storage services and validate against the ongoing analysis activity in order to successfully face the new LHC data taking period in 2015. Besides summarizing the current state and foreseen evolutions, the talk will focus on the detailed analysis of the operational experience of both systems, in particular service efficiency, performance and reliability.

Authors: CHAN, Belinda (CERN); IVEN, Jan (CERN); MASCETTI, Luca (CERN); LAMANNA, Massimo (CERN); ESPINAL CURULL, Xavier (CERN)

Presenter: ESPINAL CURULL, Xavier (CERN)

Session Classification: Data Stores, Data Bases, and Storage Systems

Track Classification: Data Stores, Data Bases, and Storage Systems