Contribution ID: **390** Type: **poster**

Techniques for high-throughput, reliable transfer systems: break-down of PhEDEx design

Wednesday, 15 February 2006 09:00 (20 minutes)

Distributed data management at LHC scales is a staggering task, accompanied by equally challenging practical management issues with storage systems and wide-area networks. CMS data transfer management system, PhEDEx, is designed to handle this task with minimum operator effort, automating the workflows from large scale distribution of HEP experiment datasets down to reliable and scalable transfers of individual files over frequently unreliable infrastructure. PhEDEx has been designed and proven to scale beyond the current CMS needs. Few of the techniques we have used are novel, but rarely documented in HEP. We describe many of the techniques we have used to make the system robust and able to deliver high performance. On schema and data organisation we describe our use of hierarchical data organisation, separation of active and inactive data, and tuning the database for the data and access patterns. Regarding monitoring we describe our use of optimised queries, moving queries away from hot tables, and using multi-level performance histograms to precalculate partial aggregated results. Robustness applies to both detecting and recovering from local errors, and robustness in the distributed environment. We describe the coding patterns we use for error-resilient and selfhealing agents for the former, and the breakdown of handshakes in file transfer, routing files to destinations, and in managing site presence for the latter.

Primary authors: TUURA, Lassi (Northeastern University); BARRASS, Timothy Adam (University of Bristol)

Co-authors: BONACORSI, Daniele (INFN-CNAF); REHN, Jens (CERN); HERNANDEZ, Jose (CIEMAT); WU, Yujun (FNAL)

Presenter: BARRASS, Timothy Adam (University of Bristol)

Session Classification: Poster

Track Classification: Distributed Event production and processing