Contribution ID: **328** Type: **Poster**

Big data log mining: the key to efficiency

Tuesday 22 May 2012 13:30 (4h 45m)

In addition to the physics data generated each day from the CMS detector, the experiment also generates vast quantities of supplementary log data. From reprocessing logs to transfer logs this data could shed light on operational issues and assist with reducing inefficiencies and eliminating errors if properly stored, aggregated and analyzed. The term "big data" has recently taken the spotlight with organizations worldwide using tools such as CouchDB, Hadoop and Hive. In this paper we present a way of evaluating the capture and storage of log data from various experiment components to provide analytics and visualization in near real time.

Author: ROSSMAN, Paul (Fermi National Accelerator Laboratory (FNAL))

Presenter: ROSSMAN, Paul (Fermi National Accelerator Laboratory (FNAL))

Session Classification: Poster Session

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