

PhEDEx high-throughput data transfer management system

Thursday 16 February 2006 14:00 (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. Over the last year PhEDEx has matured to the point of handling virtually all CMS production data transfers. CMS pushes equally its own components to perform and the heavy investment into peer projects at all levels, from technical details to grid standards to world-wide projects, to ensure the end-to-end service is of sufficient quality. We present the throughput and service quality we have reached in the current daily 24/7 production work, the steps taken in LCG service challenges for the next generation transfer service, and the resulting changes in performance. We also report results from our scalability stress tests on PhEDEx alone. We offer an analysis of transfer-related problems we have encountered and how they have been affecting CMS data management.

Primary authors: Dr BONACORSI, Daniele - on behalf of the CMS Collaboration (INFN-CNAF, Bologna, Italy); TUURA, Lassi (Northeastern University); BARRASS, Tim (Bristol University)

Co-authors: PAPAGEORGIOU, Anastasios (Imperial College, London); NOWACK, Andreas (RWTH-Aachen); TRUNOV, Artem (CC IN2P3); BOCKELM, Brian (University of Nebraska-Licln); KUO, Chia-Ming (National Central University, Chung-li); FILIPPIDIS, Christos (NCSR Demokritos); PRESCOTT, Craig (University fo Florida); BONACORSI, Daniele (INFN-CNAF); FEICHTINGER, Derek (CERN); BOURILKOV, Dimitri (University of Florida); CALZOLARI, Federico (SNS Pisa); TEHRANI, Francesco (INFN-Roma); DONVITO, Giacinto (INFN-Bari); LETTS, James (University of California, San Diego); REHN, Jens (CERN); HERNANDEZ, Jose (CIEMAT); RABBERTZ, Klaus (University of Karlsruhe); MUENTEL, Mait (National Institute of Chemical Physics and Biophysics, Tallinn); KADASTIK, Mario (National Institute of Chemical Physics and Biophysics, Tallinn); BIASOTTO, Massimo (INFN-Legnaro); ERNST, Michael (DESY); KODOLOVA, Olga (Scobeltcyn Institute of Nuclear Physics); SMITH, Preston (Purdue University); DE WEIRDT, Stein (IIHE ULB-VUB); WAKEFIELD, Stuart (Imperial College, London); SINGH, Suresh (Caltech); WU, Yujun (FNAL)

Presenter: REHN, Jens (CERN)

Session Classification: Distributed Event production and Processing

Track Classification: Distributed Event production and processing