



Contribution ID: 334

Type: oral presentation

Production mode Data-Replication framework in STAR using the HRM Grid

Monday, 27 September 2004 16:50 (20 minutes)

The STAR experiment utilizes two major computing facilities for its data processing needs - the RCF at Brookhaven and the PDSF at LBNL/NERSC. The sharing of data between these facilities utilizes data grid services for file replication, and the deployment of these services was accomplished in conjunction with the Particle Physics Data Grid (PPDG). For STAR's 2004 run it will be necessary to replicate ~100 TB. The file replication is based on Hierarchical Resource Managers (HRMs) along with Globus tools for security (GSI) and data transport (GridFTP). HRMs are grid middleware developed by the Scientific Data Management group at LBNL, and STAR file replication consists of an HRM interfaced to HPSS at each site with GridFTP transfers between the HRMs. Each site also has its own installation of the STAR file and metadata catalog, which is implemented in MySQL. Queries to the catalogs are used to generate file transfer requests. Single requests typically consist of many thousands of files with a volume of hundreds of GBs. The HRMs implement a plugin to a Replica Registration Service (or RRS) which is utilized for automatic registration of new files as they are successfully transferred across sites. This allows STAR users immediate use of the distributed data. Data transfer statistics and system architecture will be presented.

Primary authors: SHOSHANI, A. (Lawrence Berkeley Laboratory); SIM, A. (Lawrence Berkeley Laboratory); OLSON, D. (Lawrence Berkeley Laboratory); HJORT, E. (LAWRENCE BERKELEY LABORATORY); LAURET, Jerome (BROOKHAVEN NATIONAL LABORATORY)

Presenter: HJORT, E. (LAWRENCE BERKELEY LABORATORY)

Session Classification: Distributed Computing Services

Track Classification: Track 4 - Distributed Computing Services