



Contribution ID: 218

Type: oral presentation

On Distributed Database Deployment for the LHC Experiments

Monday, 27 September 2004 15:00 (20 minutes)

While there are differences among the LHC experiments in their views of the role of databases and their deployment, there is relatively widespread agreement on a number of principles:

1. Physics codes will need access to database-resident data. The need for database access is not confined to middleware and services: physics-related data will reside in databases.
2. Database-resident data will be distributed, and replicated. A single, centralized database, at CERN or elsewhere, does not suffice.
3. Distributed deployment infrastructure should be open to the use of different technologies as appropriate at the various Tier N sites.

A variety of approaches to distributed deployment have been explored in the context of individual experiments; indeed, a degree of distributed deployment has been integral to the computing model tests of some experiments (cf. ATLAS) in their 2004 data challenges. Approaches to replication have also been investigated in the context of specific databases, often with vendor-specific replication tools (e.g., Oracle Replication via Streams for the LCG File Catalog and the Oracle instantiation of the LCG conditions database; MySQL tools for replication in the MySQL instantiation of the LCG conditions database). XML exchange mechanisms have also been discussed. Distributed database deployment, though, is more than a middleware and applications software issue—a successful strategy must involve those who will be responsible for systems deployment and administration at LHC grid sites.

We describe the status of ongoing work in this area, and discuss the prospects for components of a common approach to distributed deployment in the time frame of the 2005 LHC data challenges.

Primary authors: VANIACHINE, A. (ANL); DUELLMANN, D. (CERN); MALON, D. (ANL); BIRD, I. (CERN); SHIERS, J. (CERN); GIRONE, M. (CERN); LAMANNA, M. (CERN)

Presenter: DUELLMANN, Dirk

Session Classification: Distributed Computing Services

Track Classification: Track 4 - Distributed Computing Services