

Distributed Analysis in ATLAS using GANGA

Tuesday, 24 March 2009 17:50 (20 minutes)

The distributed data analysis using Grid resources is one of the fundamental applications in high energy physics to be addressed and realized before the start of LHC data taking. The needs to manage the resources are very high. In every experiment up to a thousand physicist will be submitting analysis jobs into the Grid. Appropriate user interfaces and helper applications have to be made available to assure that all users can use the Grid without expertise in Grid technology. These tools enlarge the number of grid users from a few production administrators to potentially all participating physicists.

The GANGA job management system (<http://cern.ch/ganga>), developed as a common project between the ATLAS and LHCb experiments provides and integrates these kind of tools.

GANGA provides a simple and consistent way of preparing, organizing and executing analysis tasks within the experiment analysis framework, implemented through a plug-in system. It allows trivial switching between running test jobs on a local batch system and running large-scale analyzes on the Grid, hiding Grid technicalities.

We will be reporting on the plug-ins and our experiences of distributed data analysis using GANGA within the ATLAS experiment. Support for all grids presently used by ATLAS, namely the LCG/EGEE, NDGF/NorduGrid, and OSG/PanDA is provided. The integration and interaction with the ATLAS data management system DQ2 into GANGA is a key functionality. An intelligent job brokering is setup by using the job splitting mechanism together with dataset and file location knowledge. The brokering is aided by an automated system that regularly processes test analysis jobs at all ATLAS DQ2 supported sites. Large amounts of analysis jobs can be sent to the locations of data following the ATLAS computing model.

GANGA supports amongst other things tasks of user analysis with reconstructed data and small scale production of Monte Carlo data.

Primary author: ELMSHEUSER, Johannes (Ludwig-Maximilians-Universität München)

Co-authors: SOROKO, Alexander (University of Oxford); MAIER, Andrew (CERN); GAIDIOZ, Benjamin (CERN); SAMSET, Bjorn (University of Oslo); VAN DER STER, Daniel (CERN); BROCHU, Frederic (University of Cambridge); COWAN, Greig (University of Edinburgh); LEE, Hurng-Chun (NIKHEF); MOSCICKI, Jakub (CERN); PAJCHEL, Katarina (University of Oslo); SLATER, Mark (University of Birmingham); WILLIAMS, Michael (Imperial College London); EGEDE, Ulrik (Imperial College London); REECE, Will (Imperial College London)

Presenter: ELMSHEUSER, Johannes (Ludwig-Maximilians-Universität München)

Session Classification: Distributed Processing and Analysis

Track Classification: Distributed Processing and Analysis