

# Prototype of a Parallel Analysis System for CMS using PROOF

*Tuesday, 14 February 2006 16:20 (20 minutes)*

A typical HEP analysis in the LHC experiments involves the processing of data corresponding to several million events, terabytes of information, to be analysed in the last phases. Currently, processing one million events in a single modern workstation takes several hours, thus slowing the analysis cycle. The desirable computing model for a physicist would be closer to a High Performance Computing one where a large number of CPUs are required for short periods (of the order of several minutes). Where CPU farms are available, parallel computing is an obvious solution to this problem. Here we present the tests along this line using a tool for parallel physics analysis in CMS based on the PROOF libraries. Special attention has been paid in the development of this tool to modularity and easiness of usage to enable the possibility of sharing algorithms and simplifying software extensibility while hiding the details of the parallelisation.

The first tests performed using a medium size (90 nodes) cluster of dual processor machines on a typical CMS analysis dataset (corresponding to root files for one million top quark pairs producing fully leptonic final state events distributed uniformly among the computers) show quite promising results on scalability.

**Primary author:** Dr GONZALEZ CABALLERO, Isidro (Instituto de Fisica de Cantabria (CSIC-UC))

**Co-authors:** Mr CANO, Daniel (Instituto de Fisica de Cantabria (CSIC-UC)); Dr CUEVAS, Javier (Departamento de Fisica, Universidad de Oviedo); Mr MARCO, Rafael (Instituto de Fisica de Cantabria (CSIC-UC))

**Presenter:** Dr GONZALEZ CABALLERO, Isidro (Instituto de Fisica de Cantabria (CSIC-UC))

**Session Classification:** Distributed Data Analysis

**Track Classification:** Distributed Data Analysis