Contribution ID: **6** Type: **Poster presentation**

An SQL-based approach to Physics Analysis

Monday, 14 October 2013 15:00 (45 minutes)

As part of the CERN Openlab collaboration, an investigation has been made into the use of an SQL-based approach for physics analysis with various up-to-date software and hardware options.
Currently physics analysis is done using data stored in customised root-ntuples that contain only the variables needed for a specific analysis. Production of these ntuples is mainly done by accessing the centrally produced analysis data through the LHC computing grid and can take several days to complete.
We'll present an alternative approach to physics data analysis where analysis data is stored in a database, removing the need for customized ntuple production, and allowing calculations that are part of the analysis to be done on the database side. An example implementation of such a database will be shown, demonstrating how physics analysis in a database can be done via ROOT.
The use of an Oracle database in this setup is compared to use of a Hadoop-based data structure. The advantages and drawbacks of the different database setups are presented, with a detailed analysis of the CPU and I/O usage, in particular for the case when many users need to access the database at the same time.

Summary

Primary author: LIMPER, Maaike (CERN)

Presenter: LIMPER, Maaike (CERN)

Session Classification: Poster presentations

Track Classification: Event Processing, Simulation and Analysis