

Distributed Computing and Data Analysis for CMS in view of the LHC startup

Peter KREUZER
CERN, Switzerland
Peter.Kreuzer@cern.ch

The Large Hadron Collider (LHC) at CERN/Geneva is about to deliver first collision data to the CMS experiment. The CMS computing, software and analysis projects will need to meet the expected performances in terms of archiving of raw detector data, prompt calibration and primary reconstruction at the host laboratory, data distribution and handling at Tier-1 centers, data access and distributed physics analysis at Tier-2 centers. Hundreds of physicists located in nearly 200 institutions around the world will then expect to find the necessary infrastructure to easily access and process experimental data, including a large range of activities from low-level detector performance evaluation to involved discovery physics analysis.

In the past two years, CMS has conducted computing, software, and analysis challenges to demonstrate the functionality, scalability, and usability of the computing and software components. These challenges have been designed to validate the CMS distributed computing model by demonstrating the functionality of many components simultaneously. We will present the major boost demonstrated by CMS in event processing, data transfers and analysis processing during the CSA07 and on-going CCRC08 data challenges. In particular, we will describe relevant functional tests and the scale achieved from each CMS component and externally provided component. We will also summarize the main physics analysis lessons drawn from these challenges and the on-going tunings for an optimal begin of the experiment.