



Contribution ID: 58

Type: oral presentation

Online data handling and storage at the CMS experiment

Tuesday, April 14, 2015 3:30 PM (15 minutes)

The CMS experiment at CERN is one of the two general-purpose detectors on the Large Hadron Collider (LHC) in the Geneva area, Switzerland. Its infrastructure has undergone massive upgrades during 2013 and 2014, which lead to major changes in the philosophy of its DAQ (Data Acquisition) system. One of the major components of this system is the Storage Manager, which is responsible for buffering the online data generated at the level of the readout units (RUs), which crosses the computing farm undergoing different processing and filtering stages all the way to the builder units (BUs). The Storage Manager at CMS is made up of three components: the distributed file system, the Merger service and the Transfer System. In the production DAQ system there will be around 50 BUs which will concurrently write their output data at an expected aggregated rate of 2 GB/s. A merger service has been put in place to aggregate this data. Counting the merger reading of the data provided by the BUs, its writing back of the merged data and the final reading for transfer to Tier0, an estimated bandwidth of 7GB/s in concurrent r/w mode is needed. Moreover, the Storage Manager has to be able to cope with being disconnected from Tier0 for 1 week, so an estimated of 250TB of total usable disk space is required. A unified name space file system (Lustre) has been chosen in order to cope with these requirements. Three different implementations of the merger service are proposed, each one providing different advantages: robustness, easy debugging, bandwidth requirements reduction. The merger is also providing the bookkeeping needed for establishing the handshake protocol between the Transfer System and the Tier0 facility at CERN, which ensures data consistency and integrity throughout the transfer process. Even if the handshake protocol itself is identical to the one from the Run1, some adjustments needed to be performed in the Transfer System in order to accommodate the new structures providing the required meta-information related to the acquired data. In addition to the nominal transfer of data to Tier0, the Transfer System needs to intelligently distribute the data, for a number of data streams need to be (also) stored locally in the CMS network for various consumers to process on site. In this article we present the various technological and implementation choices of the three components of the Storage Manager.

Primary authors: ANDRONIDIS, Anastasios (Aristotle Univ. of Thessaloniki (GR)); HOLZNER, Andre Georg; PETRUCCI, Andrea (CERN); FORREST, Andrew Kevin (University of Kent (GB)); Dr RACZ, Attila (CERN); DUPONT, Aymeric Arnaud (CERN); STIEGER, Benjamin (CERN); NUNEZ BARRANCO FERNANDEZ, Carlos (CERN); DELDICQUE, Christian (CERN); PAUS, Christoph (Massachusetts Inst. of Technology (US)); SCHWICK, Christoph (CERN); WAKEFIELD, Christopher Colin (Staffordshire University (GB)); GIGI, Dominique (CERN); MESCHI, Emilio (CERN); GLEGE, Frank (CERN); MEIJERS, Frans (C); DARLEA, Georgiana Lavinia (Massachusetts Inst. of Technology (US)); GOMEZ CEBALLOS RETUERTO, Guillermo (Massachusetts Inst. of Technology (US)); SAKULIN, Hannes (CERN); BRANSON, James Gordon (Univ. of California San Diego (US)); Mr VEVERKA, Jan (Massachusetts Inst. of Technology (US)); ANDRE, Jean-Marc Olivier (Fermi National Accelerator Lab. (US)); Dr HEGEMAN, Jeroen (CERN); SUMOROK, Konstanty (Massachusetts Inst. of Technology (US)); MASETTI, Lorenzo (CERN); ORSINI, Luciano (CERN); Dr DOBSON, Marc (CERN); PIERI, Marco (Univ. of California San Diego (US)); CHAZE, Olivier (CERN); ZEJDL, Petr (CERN); MOMMSEN, Remi (Fermi National Accelerator Lab. (US)); ERHAN, Samim (Univ. of California Los Angeles (US)); CITTOLIN, Sergio (Univ. of California San Diego (US)); MOROVIC, Srecko (CERN); BAWEJ, Tomasz Adrian (University of Wisconsin (US)); BEHRENS, Ulf (Deutsches Elektronen-Synchrotron (DE)); O'DELL, Vivian

(Fermi National Accelerator Laboratory (FNAL))

Presenter: DARLEA, Georgiana Lavinia (Massachusetts Inst. of Technology (US))

Session Classification: Track 1 Session

Track Classification: Track1: Online computing