Evolution of the LHCb Online run bookkeeping and data consistency system

Tuesday, 11 October 2016 16:30 (15 minutes)

Since its original commissioning in 2008, the LHCb data acquisition system has seen several fundamental architectural changes. The original design had a single, continuous stream of data in mind, going from the read-out boards through a software trigger straight to a small set of parallelly written files. Over the years the enormous increase in available storage capacity has made it possible to reduce the amount of real-time computing at the experiment site and move to a more Offline like processing of data. A reduced, software based trigger is pre-processing the detector data which is then stored in a pool of storage elements in our computing farm. This data is then further processed at a later time, when computing resources are available.

Today, the LHCb Online System is keeping track of several thousand files and hundreds of runs being processed concurrently. The additional storage and parallel processing made it necessary to scale the run and file bookkeeping system far beyond its original design specifications. The additional complexity of the data flow also called for improved sanity checks and post processing, before the data can be shipped to the grid for analysis.

In this paper we are going to show the evolution of our scaled up system, with particular focus on handling several runs in parallel, output file merging for easier offline processing, data integrity checking and assurance that events are only sent offline once.

Tertiary Keyword (Optional)

Secondary Keyword (Optional)

Data processing workflows and frameworks/pipelines

Primary Keyword (Mandatory)

DAQ

Primary author: SCHWEMMER, Rainer (CERN)

Co-author: NEUFELD, Niko (CERN)

Presenter: SCHWEMMER, Rainer (CERN)

Session Classification: Posters A / Break

Track Classification: Track 1: Online Computing