

Prediction of Event Processing Time at the CMS Experiment of the Large Hadron Collider

The physics event reconstruction in LHC/CMS is one of the biggest challenges for computing.

Among the different tasks that computing systems perform, the reconstruction takes most of the CPU resources that are available. The reconstruction time of a single event varies according to the event complexity. Measurements were done in order to find precisely this correlation, creating means to predict it based on the physics conditions of the input data.

Currently the data processing system do not account that when splitting a task in chunks(jobs), this can cause a considerable variation in the job length, thus a considerable increase into the workflow Estimated Time of Arrival.

The goal is to use this prediction to more efficiently split the work in chunks, considering the CPU time needed for each chunk and due to this, lowering the standard deviation of the job length distribution in a workflow.