Workflow Management for Complex HEP Analyses

Thursday, 13 October 2016 16:30 (15 minutes)

We present the novel Analysis Workflow Management (AWM) that provides users with the tools and competences of professional large scale workflow systems. The approach presents a paradigm shift from executing parts of the analysis to defining the analysis.

Within AWM an analysis consists of steps. For example, a step defines to run a certain executable for multiple files of an input data collection. Each call to the executable for one of those input files can be submitted to the desired run location, which could be the local computer or a remote batch system. An integrated software manager enables automated user installation of dependencies in the working directory at the run location. Each execution of a step item creates one report for bookkeeping purposes containing error codes and output data or file references. Required files, e.g. created by previous steps, are retrieved automatically. Since data storage and run locations are exchangeable from the step's perspective, computing resources can be used opportunistically. A visualization of the workflow as a graph of the steps in the web browser provides a high-level view on the analysis. The workflow system is developed and tested alongzaʻside of a ttbb cross section measurement where, for instance, the event selection is represented by one step and a Bayesian statistical inference is performed by another.

The clear interface and dependencies between steps enables a make-like execution of the whole analysis.

Tertiary Keyword (Optional)

Analysi tools and techniques

Secondary Keyword (Optional)

Distributed workload management

Primary Keyword (Mandatory)

Data processing workflows and frameworks/pipelines

Primary author: FISCHER, Robert (Rheinisch-Westfaelische Tech. Hoch. (DE))

Co-authors: RIEGER, Marcel (Rheinisch-Westfaelische Tech. Hoch. (DE)); ERDMANN, Martin (Rheinisch-Westfaelische Tech. Hoch. (DE)); VON CUBE, Ralf Florian (Rheinisch-Westfaelische Tech. Hoch.(DE))

Presenter: FISCHER, Robert (Rheinisch-Westfaelische Tech. Hoch. (DE))

Session Classification: Posters B / Break

Track Classification: Track 7: Middleware, Monitoring and Accounting