



Contribution ID: 383

Type: **Parallel**

Using Functional Languages and Declarative Programming to Analyze Large Datasets: LINQToROOT

Tuesday 22 May 2012 14:45 (25 minutes)

Modern HEP analysis requires multiple passes over large datasets. For example, one has to first reweight the jet energy spectrum in Monte Carlo to match data before you can make plots of any other jet related variable. This requires a pass over the Monte Carlo and the Data to derive the reweighting, and then another pass over the Monte Carlo to plot the variables you are really interested in. With most modern ROOT based tools this requires separate analysis loops for each pass, and script files to glue to the two analysis loops together. A prototype framework has been developed that uses the functional and declarative features of C# and LINQ to specify the analysis. The framework uses language tools to convert the analysis into C++ and runs ROOT or PROOF as a backend to get the results. This gives the analyzer the full power of a object-oriented programming language to put together the analysis and at the same time the speed of C++ for the analysis loop. The tool allows one to incorporate C++ algorithms written for ROOT by others. The code is mature enough to have been used in ATLAS analyses. The package is open source and available on the open source site Codeplex.

Summary

A new approach to end-user analysis that tries to take advantage of modern (i.e. computer language research from the '70s) and the already exiting infrastructure in HEP (i.e. ROOT).

Primary author: WATTS, Gordon (University of Washington (US))

Presenter: WATTS, Gordon (University of Washington (US))

Session Classification: Event Processing

Track Classification: Event Processing (track 2)