Contribution ID: **394**   Type: **Oral**

Merging multidimensional histograms via hypercube algorithm

Thursday, 18 April 2019 17:15 (15 minutes)

Scientists in high energy physics produce their output mostly in form of histograms. Set of histograms are saved in output
file for each grid job. As the next step is to merge these files/histograms to one file where scientist can produce final plots
for publication. Merging of these out files may be done sequentially as one job or do it in parallel via binary tree algorithm
as it is done by many users. Using histogram with low dimensions (1D or 2D) one can fit in memory with final merged
objects. On the other side, if dimensions or binning of histograms are increaced, sparse implementation of histogram has
to be used in analysis and final object might grow so much that user will not be able to merge or open final merged object
because it will not fit in memory at some point. Our task is merge these multidimensional histograms to N independed
objects to multiple files, where each file will contain uniqe part of merged object sorted by some axis in histogram
dimension. For optimalization reasons hypercube algorithm is used.

Primary author: VALA, Martin (JINR)

Co-authors: BULATOV, Andrey (State University Dubna, JINR); BUTENKO, Yuri (JINR)

Presenter: BULATOV, Andrey (State University Dubna, JINR)

Session Classification: Information Technologies

Track Classification: Information Technology