Contribution ID: **129** Type: **Poster**

Optimizing ROOT I/O for Analysis

Thursday, 24 August 2017 16:25 (20 minutes)

The ROOT I/O (RIO) subsystem is foundational to most HEP experiments - it provides a file format, a set of APIs/semantics, and a reference implementation in C++. It is often found at the base of an experiment's framework and is used to serialize the experiment's data; in the case of an LHC experiment, this may be hundreds of petabytes of files! Individual physicists will further use RIO to perform their end-stage analysis, reading from intermediate files they generate from experiment data.

RIO is thus incredibly flexible: it must serve as a file format for archival (optimized for space) and for working data (optimized for read speed). To date, most of the technical work has focused on improving the former use case. We present work designed to help improve RIO for analysis. We analyze the real-world impact of LZ4 to decrease decompression times (and the corresponding cost in disk space). We introduce new APIs that read RIO data in bulk, removing the per-event overhead of a C++ function call. We compare the performance with the existing RIO APIs for simple structure data and show how this can be complimentary with efforts to improve the parallelism of the RIO stack.

Primary authors: BOCKELMAN, Brian Paul (University of Nebraska-Lincoln (US)); ZHANG, Zhe

Co-author: PIVARSKI, Jim (Princeton University)

Presenter: BOCKELMAN, Brian Paul (University of Nebraska-Lincoln (US))

Session Classification: Poster Session

Track Classification: Track 1: Computing Technology for Physics Research