Contribution ID: **383**　　　　　　　　　　　　　　　　　　　　　　　Type: **Poster**

# Speeding HEP Analysis with ROOT Bulk I/O

Distinct HEP workflows have distinct I/O needs; while ROOT I/O excels at serializing complex C++ objects common to reconstruction, analysis workflows typically have simpler objects and can sustain higher event rates. To meet these workflows, we have developed a "bulk I/O"interface, allowing multiple events'data to be returned per library call. This reduces ROOT-related overheads and increases event rates – orders-of-magnitude improvements are shown in microbenchmarks.

Unfortunately, this bulk interface is difficult to use as it requires users to identify when it is applicable and they still "think"in terms of events, not arrays of data. We have integrated the bulk I/O interface into the new RDataFrame analysis framework inside ROOT. As RDataFrame's interface can provide improved type information, the framework itself can determine what data is readable via the bulk IO and automatically switch between interfaces. We demonstrate how this can improve event rates when reading analysis data formats, such as CMS's NanoAOD.

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

**Presenter:** SHADURA, Oksana (University of Nebraska Lincoln (US))

**Session Classification:** Poster Session

**Track Classification:** Track 1: Computing Technology for Physics Research