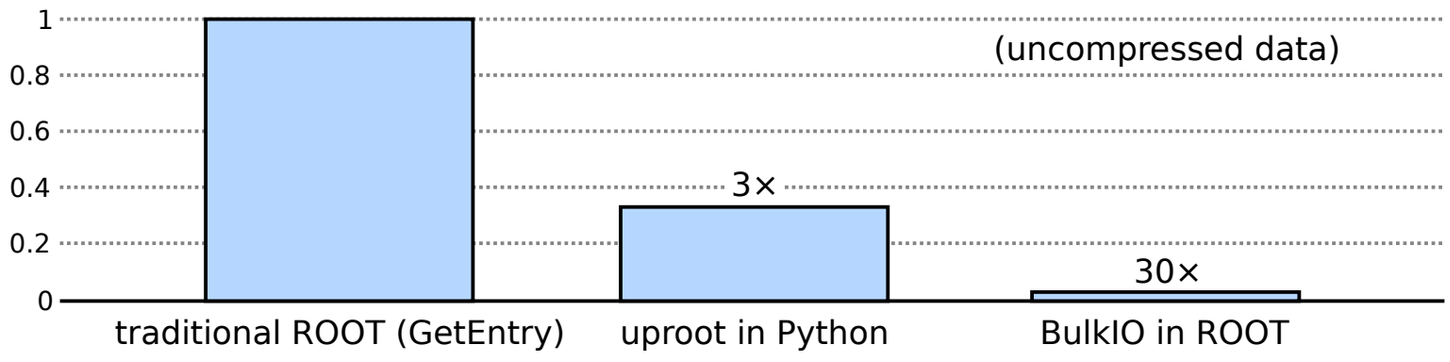
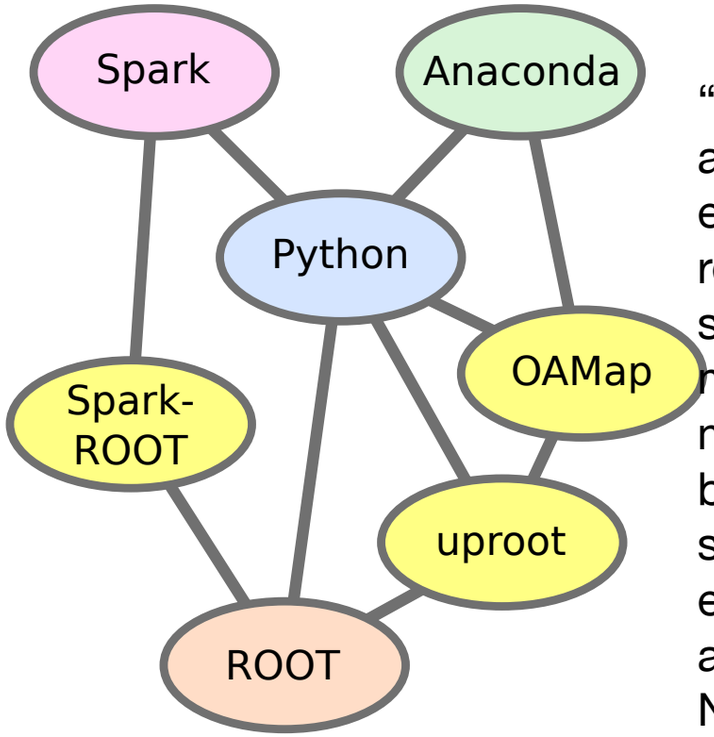
Advancing Analysis for HEP



Improved Performance

To reduce the time to scientific discovery and to enable more in-depth analyses, we are increasing the rate of access to ROOT data files. This includes streamlined access to simpler data types (uproot and BulkIO) and faster compression algorithms (LZ4 and ZSTD). These efforts have already provided factors-of-several improvements.



Bridging to Big Data

“Big Data” software in industry, such as the Spark and scientific Python ecosystems, both complement and reproduce functionality of HEP software developed. To provide more options and reduce maintenance burdens, DIANA is building bridges between HEP software and the Big Data ecosystems: Spark-ROOT to Spark and uproot/OAMap to Numpy, Numba, and Dask.

Statistical Techniques

We are developing tools and methods for statistical analysis in HEP, including research for simulator-based inference (Carl), machine learning for particle physics (Scikit-Optimize), and software for efficient numerical computations.

High Level Tools

We are therefore striving to present HEP analysis with higher-level interfaces. Scikit-HEP incorporates HEP techniques in Pythonic idioms, uproot provides access to ROOT data as Numpy and Pandas abstractions, and OAMap compiles object-centric user code into fast array operations.

Scikit-HEP

uproot

NYU

diana hep