Hydra: A framework for data analysis in massively parallel platforms.

A. Augusto Alves Jr\textsuperscript{1}, M. D. Sokoloff\textsuperscript{1}

\textsuperscript{1} University of Cincinnati

What is it?

Hydra is a header only templated C++ library designed to perform common HEP data analyses on massively parallel platforms.

- It is implemented on top of the C++11 Standard Library and a variadic version of the Thrust library.
- It runs on Linux systems and uses OpenMP, CUDA and TBB back-ends.
- It is focused on portability, usability, performance and precision.

Design and features

- The library is structured using static polymorphism and the interfaces have a clean and concise semantics.
- There is absolutely no need to write explicit back-end oriented code.
- All supported back-ends can run concurrently in the same program using the suitable policies: hydra::omp::sys, hydra::cuda::sys, hydra::host::sys, hydra::device::sys, hydra::tbb::sys.
- The same source files written using Hydra and standard C++ compile for GPU, CPU or even both, just changing the extension from .cu to .cpp and one or two compiler flags.

Functionality

- Interface to ROOT::Minuit2 minimization package.
- Phase-space generation and integration.
- Multidimensional p.d.f. sampling.
- Parallel function evaluation on multidimensional datasets.
- Numerical integration: Monte Carlo and quadrature based.
- The VEGAS algorithm samples the integrand and adapts itself, so that the points are concentrated in the regions that make the largest contribution to the integral.
- No limit in the number of dimensions.
- Example: integrating a normalized Gaussian distribution in 10 dimensions.

Examples and performance

System configuration:

- GPU model: Tesla K40c
- CPU: Intel Xeon(R) CPU E5-2680 v3 @ 2.50GHz (one thread)

Vegas multidimensional numerical integration

The VEGAS algorithm samples the integrand and adapts itself, so that the points are concentrated in the regions that make the largest contribution to the integral.

- No limit in the number of dimensions.
- Example: integrating a normalized Gaussian distribution in 10 dimensions.

Phase-Space Monte Carlo

- No limitation on the number of particles in the final state.
- Support the generation of sequential decays and other features.

Results and performance:

- No limitation on the number of particles in the final state.
- Support the generation of sequential decays and other features.

Summary

Hydra’s development has been supported by the National Science Foundation under the grant number PHY-1414736.

- The project is hosted on GitHub: https://github.com/MultithreadCorner/Hydra
- The package includes a suite of examples.
- It is being used on the measurement of the Kaon mass at LHCb.
- A Google Summer of Code (GSoC) student has been working with the developers to add Python bindings. This is implemented for OpenMP and TBB back-ends.