FaDER

Fast Data and Event Reduction:
lowering data volumes in high-intensity experiments

Carlo Schiavi on behalf of the FaDER Team

04 10 2024

Università di Genova | DIFI DIPARTIMENTO DI FISICA

INFN GENOVA

Project summary - the Jefferson Lab Path (BDX)

Focus on real-time fast data reduction (BDX experiment study case)

Test of low-level algorithms (NN autoencoders, classifiers, cluster-finders) for efficient data reduction for real-time applications (e.g. streaming readout DAQ).

- Real-time low-level data reduction: JLab fADC 250
  → M1.1 (M6): definition of NN autoencoder algorithm to compress waveforms (BDX simulated data)
  → M1.2 (M12): deployment and test of NN autoencoder algorithm on CPU and test (BDX prototype data)
  → M1.3 (M16): deployment of the NN autoencoder algorithm on FPGA and NN autoencoder performance profiling

- Real-time high-level data reduction: BDX experiment ‘events’
  → M2.1 (M8): definition of a NN classifier to select a soft trigger condition for BDX experiment (simulations)
  → M2.1 (M16): deployment and test of a NN clustering algorithm to select BDX ‘events’ and NNs performance profiling (M2.2)

Project summary - the LHC Path (ATLAS)

Focus on improving track-based triggers

Test multiple **tracking detector hit pre-selection / clustering techniques**, based on ML, aiding the track finding in the dense LHC/HL-LHC environment (Mxx indicates estimated milestone completion month):

- **reject hits from pile-up or low-pT tracks** before track reconstruction, to reduce combinatorics
- **estimate the position of the primary vertex** along the beam line starting from detector hits
  → M1.1 (M8): implementation using CNNs; possibly test on FPGA
  → M1.2 (M12): implementation using Transformers and symbolic approx.; possibly test on FPGA
  → M1.3 (M16): compare and optimize the results from M1.2 and M1.3 on physics samples

- study the **impact on b-jet and hadronic tau selections** and optimize a possible **unified tagger**
  → M2.1 (M16): test and optimize algs from M1.1 and M1.2 on final states containing b-jets and taus
  → possibly measure the impact on a unified b-jet and tau trigger tagger

Project computational requests

The Jefferson Lab path

Real-time streaming readout DAQ require fast data-reduction algorithms running on fast and heterogeneous hardware (CPU/GPU/FPGA).

- AI-supported algorithms development and validation: (local) CPU/GPU/FPGA
- Deployment on fast hw and profiling*: (remote) GPUs and FPGA (O(10) GPU and FPGA days)
- Storage: local/remote O(1TB)

* for profiling tests remote resources should be exclusively allocated

The LHC path

Interested in exploiting ICSC for larger trainings, while development ones can be efficiently done locally

Apart from FPGAs for implementation tests, some quantitative request, estimated with some safety margin:

GPUs

- Computing resources for large trainings: 5-10 x O(10) days of allocation of a single GPU
- Supporting storage resources needed: O(100GB/1TB)

CPUs

- Production of large toy model sample: O(10) days of CPU usage, large memory preferred

BONUS: NP-Twins2024 workshop - Genoa, Dec 16-18, 2024 → collab with Alessandro Pilloni (UniME, GeV-AI-PI)