

Studies on track finding algorithms based on machine learning with CPU, GPU and FPGA

Friday 19 July 2024 17:19 (17 minutes)

In this presentation we describe the performance obtained running machine learning models studied for the ATLAS Muon High Level Trigger. These models are designed for hit position reconstruction and track pattern recognition with a tracking detector, on different models of commercially available Xilinx FPGA cards: Alveo U50, Alveo U250, and Versal VCK5000. We compare the inference times obtained on a CPU, on a GPU and on the FPGA cards. These tests are done using TensorFlow libraries as well as the TensorRT framework, and software frameworks for AI-based applications acceleration. The inference times and other performance benchmark parameters are compared to the needs of present and future experiments at LHC.

Alternate track

I read the instructions above

Yes

Authors: ZHU, Junjie (University of Michigan (US)); CARNESALE, Maria (Sapienza Universita e INFN, Roma I (IT))

Presenter: CARNESALE, Maria (Sapienza Universita e INFN, Roma I (IT))

Session Classification: Computing and Data handling

Track Classification: 14. Computing, AI and Data Handling