Contribution ID: **3** Type: **Talk**

Accelerating Graph Neural Networks on FPGAs for Particle Track Reconstruction using OpenCL and hls4ml

Wednesday, 2 December 2020 15:30 (8 minutes)

Current charged particle tracking algorithms at the CERN Large Hadron Collider (LHC) scale quadratically or worse with increasing number of overlapping proton-proton collisions in an event (pileup). As the LHC moves into its high-luminosity phase, pileup is expected to increase to an average of 200 overlapping collisions, highlighting the need for new algorithmic strategies. Recent work has shown that graph neural networks (GNNs) are well-suited to classifying segments of tracks. The real-time data filter at the LHC (L1 trigger) requires submicrosecond latencies that can only be met by devices like field-programmable gate arrays (FPGAs). Accelerating neural networks on FPGAs facilitates energy efficient data-processing on large datasets with execution times that meet the L1 trigger latency requirements.

In this talk, we present two complementary FPGA implementations of an interaction network, a type of GNN, using OpenCL, an open-source framework for writing programs that execute across heterogenous acceleration platforms, and hls4ml, an open-source compiler of machine learning models into firmware. The OpenCL implementation adopts a CPU-plus-FPGA coprocessing approach where the CPU host program manages the application and all computational operations are accelerated using dedicated kernels deployed to the FPGA and take advantage of the FPGA hardware architecture to parallelize operations. The hls4ml implementation utilizes Xilinx high-level-synthesis tools to convert the GNN model to FPGA firmware making it suitable for both FPGA-only and co-processing applications. We will present comparisons of the two implementations in terms of their resource usage, latency, and tracking performance on the publicly-available TrackML benchmark dataset.

Primary authors: HEINTZ, Aneesh (Cornell University (US)); RAZAVIMALEKI, Vesal (Univ. of California San Diego (US)); DUARTE, Javier Mauricio (Univ. of California San Diego (US)); DEZOORT, Gage (Princeton University (US)); OJALVO, Isobel (Princeton University (US)); THAIS, Savannah Jennifer (Princeton University (US)); ATKINSON, Markus Julian (Univ. Illinois at Urbana Champaign (US)); NEUBAUER, Mark (Univ. Illinois at Urbana Champaign (US)); GRAY, Lindsey (Fermi National Accelerator Lab. (US))

Presenter: HEINTZ, Aneesh (Cornell University (US))