1st FPGA Developers' Forum (FDF) meeting

Contribution ID: **8**

Type: **not specified**

Tree Tensor Network inference on FPGA

Wednesday 12 June 2024 13:45 (20 minutes)

TTNs are hierarchical tensor structures commonly used for representing many-body quantum systems but can also be applied to ML tasks such as classification or optimization. The algorithmic nature of TTNs makes them easily deployable on FPGAs, which are naturally suitable for concurrent tasks like matrix multiplications. Moreover, the hardware resource limitation can be optimally tuned exploiting the intrinsic properties of said networks. We study the deployment of TTNs in high-frequency real-time applications, showing different classifier implementations on FPGA, and performing inference on synthetic ML datasets for benchmarking. A projection of the needed resources for the HW implementation of a classifier will also be provided by comparing how different degrees of parallelism affect physical resources and latency. The full firmware has been developed in VHDL, exploiting Xilinx IPs for explicit DSP declaration and AXI Stream and AXI Lite communication protocols.

Talk's Q&A

During the talk

Talk duration

15'+7'

Will you be able to present in person?

Yes

Primary authors: BORELLA, Lorenzo (Universita e INFN, Padova (IT)); Mr COPPI, Alberto (University of Padua); PAZZINI, Jacopo (Università e INFN, Padova (IT)); Dr STANCO, Andrea (University of Padua); TRIOSSI, Andrea (Universita e INFN, Padova (IT)); ZANETTI, Marco (Universita e INFN, Padova (IT))

Presenter: BORELLA, Lorenzo (Universita e INFN, Padova (IT))

Session Classification: Algorithm implementation

Track Classification: Algorithm implementation in HDL and HLS