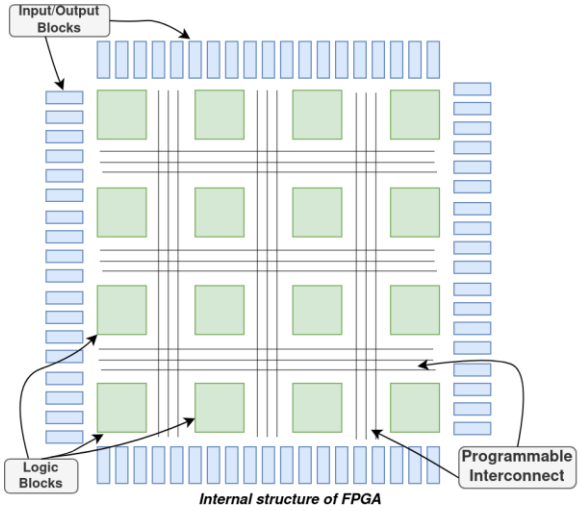
Chair for Embedded Systems

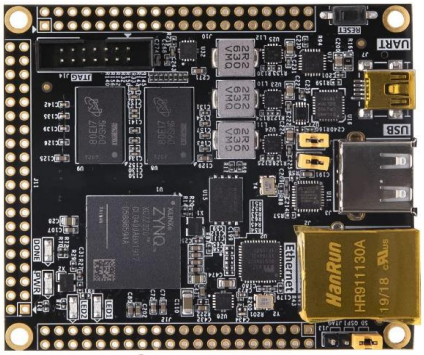
Prof. Dr. J. Henkel

Multiple Bachelor/Master Theses

Hardware Architectures for DNN Accelerators



Internal structure of FPGA


Xlinx Zynq-7000


AMD Versal AI Core

As the demand for efficient Deep Neural Network (**DNN**) implementations continues to rise, hardware accelerators play a crucial role. **Field Programmable Gate Array (FPGA)** platforms offer flexibility and performance advantages, making them an attractive choice for deploying DNN models.

The offered theses focus on different kinds of mappings of DNNs to FPGA architectures and cloud-based FPGA platforms. By exploring various mapping strategies and utilizing cloud FPGA resources, we aim to improve performance, resource utilization, and scalability in embedded systems and machine learning applications.

Depending on your interest and depending on whether it is a BA or MA, the tasks could include (but are not limited to):

- Deployment of DNNs onto FPGA platforms
  - using weight- or input stationary,
  - using row-stationary etc.
- Utilization of Cloud-based FPGA platforms, e.g. via our access to the very recent Versal AI accelerators at ETH Zurich's Heterogeneous Accelerated Compute Clusters (HACC)
- Evaluation of mapping strategies and performance analysis

Skills required/beneficial for the thesis

- Programming skills (C++, Python)
- For hardware implementations on FPGAs, background in VHDL/Verilog or FPGAs is beneficial

Skills acquired within the thesis

- Apply your programming experience to research
- Gain practical insight in hardware accelerators for DNNs and data-locality challenges when mapping DNNs to accelerators

Language

- The collaboration with the colleagues can be in English.

Contact

- Zeynep Gülbeyaz Demirdag, zeynep.demirdag@kit.edu, webpage
- Lars Bauer, lars.bauer@kit.edu, webpage