Embedded Systems Laboratory (ESL)
Director: Prof. David Atienza Alonso http://esl.epfl.ch

A bottom-up digital back-end design for the next generation of configurable accelerators for Edge-AI platforms

Contact Persons: Ms. Lara Orlandic (lara.orlandic@epfl.ch)

Dr. Davide Schiavone (davide.schiavone@epfl.ch)

Prof. David Atienza (david.atienza@epfl.ch)

Project Description

Edge Artificial Intelligence is a novel computing paradigm that has the potential to revolutionize the field of wearable medical devices. Instead of uploading sensitive biosignal data to remote servers, these devices perform all of their data processing on-board, thus preserving patients' privacy. However, to execute such complex edge-AI operations while maximizing battery lifetime, these devices must be equipped with high-performance, ultra-low power processors.

To this end, the Embedded Systems Laboratory (ESL) of EPFL has designed HEEP-Alive, an ultra-low power processor for biomedical edge-AI. It combines the open-source X-HEEP processor with the Very Wide Register Reconfigurable Array (VWR2A), an architecture that integrates high computational density and low-power memory structures to efficiently execute biomedical edge-AI kernels.

The project will involve performing hierarchical iterations of synthesis and place-and-route (PnR) of the VWR2A architecture at multiple layers of abstraction to optimize performance. As a baseline, the student will perform a traditional "top-down" synthesis and PnR of the whole design. Then, they will synthesize and PnR sub-modules of the design separately and place them on the final layout as black-box components. The goal of the project is to explore the effects of such a "bottom-up" design approach on the achieved power, performance, and area of the final design.

The expected outcomes of this project are:

- Performing a traditional "top-down" synthesis and PnR of the VWR2A
- Performing a "bottom-up" synthesis of various VWR2A sub-modules
- Modifying the back-end design flow to integrate the aforementioned sub-modules
- Investigating the effects of each "back-end" design iteration on PPA by performing:

- ○ Static timing analysis
- ○ (optional) Static power analysis
- Verifying the functionality of the design at every design iteration with post-synthesis and post-PnR netlist simulation
- Producing reproducible results using Tcl scripting, automating experiments, and regularly committing to the project's git repository

Throughout the project, the student will learn:
- Advanced back-end design methodologies
- How to assess the performance and power of the design and draw conclusions about the optimal design methodology based on these results
- How to verify the developed software and hardware.
- How to work with git repositories and in a team of people

The project will be carried out at the ESL at EPFL, one of the world's top-class universities. ESL is an active group (24 Ph.D. students among 45 members) involved in many research aspects. The student will be under the supervision of Ms. Lara Orlandic and Dr. Davide Schiavone. The student is expected to work 12 hours per week on the project throughout the semester.

Project objectives:
1. Understanding the VWR2A architecture, the function of its sub-modules, and how the interconnects between the sub-modules influence how they should be placed in the back-end design.
2. Comparing different iterative designs in terms of power, performance, and area
3. Validation of the proposed work with Verilog testbench simulations

Required knowledge and skills:

- Good knowledge of semi-custom design flows and tools:
  - ○ Synopsys Design Compiler or Cadence Genus for synthesis
  - ○ Cadence Innovus for PnR
- RTL design and in any HDL
- Good background in computer architecture
- Teamwork and git

Appreciated skills:
- Scientific curiosity
- Good communication skills
- Advanced English

Type of work: 20% theory analysis, 80% design and simulation