

Implementation of a sub-array controller for in-SRAM computing architecture.

Contact Persons: Prof. David Atienza (david.atienza@epfl.ch)

Dr. Alexandre Levisse (alexandre.levisse@epfl.ch)

Project Description

With the incoming innovations in *Artificial Intelligence* (AI) such as self-driving cars or natural language recognition, *Convolutional Neural Networks* (CNN) have been gaining popularity. Their main asset lays in their natural capability to process data with spatial or temporal interrelationships. However, nowadays CNNs implementations are extremely energy hungry, forbidding their use in embedded systems.

In this context, the ESL laboratory proposed an innovative solution named BLADE. BLADE features *In-Memory Processing* (IMP) to execute CNNs inside an embedded system. It relies on a breakthrough trendy approach named *Bitline Computing* and consists in performing logic operations directly in the cache periphery. BLADE performances have been demonstrated through extensive architectural and circuit simulations. It shows breakthrough performance gains at both circuit and application level and opens the way for new application/architecture/circuit trade-offs.

This project aims at designing the sub-array controller which generates all the signals needed to achieve all the operations occurring inside the BLADE sub-array. The student will design a RTL circuit using VHDL or Verilog physical description language. This circuit generates *precharge, enable, addresses, operations selection* signals and manage signals timing. Ultimately, the controller will be synthesized, placed and routed and included inside a placed and routed memory macro to be tape-out-ed. Depending on the timing, the student may also design a custom decoding circuitry for in-SRAM computing (accessing more than one WordLine at a time).

The project will be carried out at the Embedded Systems Laboratory of EPFL under the supervision of Prof. David Atienza, Dr. Alexandre Levisse and he will be working closely with the two PhD students in charge of the BLADE architecture: William Simon and Marco Rios.

Several references concerning BLADE can be found here:

<https://infoscience.epfl.ch/record/265152?ln=en>

<https://infoscience.epfl.ch/record/264782?ln=en>

<https://infoscience.epfl.ch/record/267989?ln=en>

Project objectives:

1. Understanding of the memory operation and features.
2. Design of a RTL memory controller in 65nm TSMC technology
3. Place and route of the proposed controlled and integration inside the memory macro
4. Top simulation showing that the proposed controller works when tied to the memory in all the technology corners.

Required knowledge and skills:

- Good understanding of memory architectures
- Advanced knowledge on digital and analog circuit design
- Good analytical skills
- Good background on computer architecture

Appreciated skills:

- Scientific curiosity
- Good communication skills
- Advanced English
- Autonomous work ability
- Teamwork

Type of work: 20% theory analysis, 60% design and simulation, 20% top validation and simulation