Near-memory computing static profiling and offloading
Stefano Corda, s.corda@tue.nl, Eindhoven University of Technology, The Netherlands

“The conventional approach of moving data to the CPU for computation has become a significant performance bottleneck for emerging scale-out data-intensive applications due to their limited data reuse. At the same time, the advancement in 3D integration technologies has made the decade-old concept of coupling compute units close to the memory — called near-memory-computing (NMC) — more viable. Processing right at the “home” of data can significantly diminish the data movement problem of data-intensive applications” [1, 2]. This project will explore and compare static and dynamic application analysis approaches focusing on NMC offloading metrics.

Goal of this project and tasks:

- Application profiling using a static approach, e.g. [3] (see how a static cache model can be faster than a simulator).
- Performance comparison with dynamic analysis tools [4, 5, 6, 7] (also commercial like Intel VTune and Advisor). Benchmark could be MemBen [8].
- Offloading evaluation employing a simulator, e.g. Ramulator [9, 10].

Skills acquired in this project

- Hands-on experience on application characterization methodologies.
- Hands-on experience on simulation environments.
Pre-requisite:
- C/C++.
- Computer Architecture.

Helpful Skills
- Knowledge about Clang LLVM flow.
- Knowledge about performance counters (PAPI/perf).
- Work independently.

References