PhD in Information Technology and Electrical Engineering
Università degli Studi di Napoli Federico II

PhD Student: Manuel Maddaluno

Cycle: XXXIX

Training and Research Activities Report

Year: First

Tutor: Prof. Alessandro Cilardo

Firmato digitalmente da
Alessandro CILARDO
CN = CILARDO Alessandro
O = Università degli Studi di Napoli Federico II

Date: December 06, 2024

1. Information:

- **PhD student: Manuel Maddaluno**
- **DR number: DR997203**
- **Date of birth: 05/07/1999**
- **Master Science degree: Computer Engineering**
- **University: Università degli Studi di Napoli Federico II**
- **Doctoral Cycle: XXXIX**
- **Scholarship type: PNRR – Centro Nazionale CN1 - National Centre for High-Performance Computing, Big Data and Quantum Computing – Spoke: Future HPC**
- **Tutor: Prof. Alessandro Cilardo**

2. Study and training activities:

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
Ensuring Electronic Reliability Against CERN's Radiation Environment	Seminar	2	0.4	01/12/2023	Prof. Francesco Fienga	Y
Energy-Efficient Data Science	Seminar	1	0.2	13/12/2023	Prof. Elio Masciari	Y
Quantum Technologies: state of the art and perspectives	Seminar	2	0.4	08/02/2024	Prof. Fabio Villone	Y
Hominis	Seminar	5	1	21/02/2024	Prof. Carlo Sansone, Eng. Stefano Marrone	Y
Analytic center selection of optimization-based controllers for robot ecology	Seminar	1	0.2	09/04/2024	Prof. Bruno Siciliano	Y
IEEE Authorship and Open Access Symposium: Tips and best Practices to Get Published from IEEE Editors	Seminar	1.5	0.4	07/05/2024	IEEE	Y
Sustainable IT: Strategies and best practices for a green engineering future	Seminar	5	1	27/05/2024	5G Academy	Y
Generative AI for software engineering:	Seminar	5	1	29/05/2024	5G Academy	Y

strategies, impacts, and practical applications						
Real-time resource management for adaptive embedded systems and applications	Seminar	1	0.2	26/06/2024	Prof. Marcello Cinque	Y
On the single allocation hub location problems: new formulations and solving methods	Seminar	1	0.2	26/06/2024	Prof. Claudio Sterle, Prof. Maurizio Boccia, Prof. Adriano Masone	Y
Using support vector machines for feature selection and outlier detection	Seminar	1	0.2	26/06/2024	Prof. Claudio Sterle, Prof. Maurizio Boccia, Prof. Adriano Masone	Y
Resource management and orchestration for mixed-criticality cloud/distributed systems	Seminar	1	0.2	27/06/2024	Prof. Marcello Cinque	Y
Including elastic demand in the hub line location problem	Seminar	1	0.2	28/06/2024	Prof. Claudio Sterle, Prof. Maurizio Boccia, Prof. Adriano Masone	Y
The maximal covering location problem with edge downgrades	Seminar	1	0.2	28/06/2024	Prof. Claudio Sterle, Prof. Maurizio Boccia, Prof. Adriano Masone	Y
IEEE Photonic Society Seminars	Seminar	4	0.8	18/11/2024	Prof. Giovanni Breglio	Y

Strategic Orientation for STEM Research & Writing	Course	20	5	Dec. 07, 15 2023 - Jan. 12, 19 - Feb. 09, 23 2024	Dr. Chie Shin Fraser	Y
Virtualization Technologie and their application	Course	20	5	Jan. 08, 10, 15, 19, 24, 25, 2, 31 – Feb. 07, 26 2024	Prof. Luigi De Simone	Y
IoT Data Analysis	Course	12	4	Feb. 15, 19, 21, 23, 27, 29 2024	Prof. Raffaele Della Corte	Y
Statistical data analysis for science and engineering research	Course	12	4	Feb. 15, 19, 21, 23, 27, 29 2024	Prof. Roberto Pietrantuono	Y

1) Courses, Seminar, Doctoral School, Research, Tutorship
2) Choose: Y or N

2.1. Study and training activities - credits earned

	Courses	Seminars	Research	Tutorship	Total
Bimonth 1	10	2	1	0	13
Bimonth 2	8	0.2	1	0	9.2
Bimonth 3	0	3.6	6	0	9.6
Bimonth 4	0	0	10	0	10
Bimonth 5	0	0	10	0	10
Bimonth 6	0	0.8	10	0	10.8
Total	18	6.6	38	0	62.6
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	

3. Research activity:

Research topic:
The main macro topic of my research is High-Performance Computing (HPC) architectures. In more detail, my research work focuses on reducing and/or controlling the latency.
My first PhD year started with the study of new memory technologies, in particular, the main focus was on High Bandwidth Memory (HBM). This new Dynamic Random Access Memory (DRAM) interface was created with the aim of providing a much larger bandwidth than the common Double Data Rate (DDR). This is to accelerate massively parallel workloads like modern neural networks. Moreover, the

adoption of Artificial Intelligence (AI) models in automotive and autonomous driving is increasing. HBM memories, due to their small form factor, are integrated directly near the processing unit. This makes them ideal for use in domains such as automotive, avionics, railway, or IoT. However, they do not guarantee deterministic and predictable access to data that is necessary in these areas.
The goal of my first work is to fill this gap by building a HBM controller architecture that can provide deterministic access while maintaining high performance.To do this, a controller architecture model was first developed on which a detailed analysis of the latency times of each command to memory was carried out.

This is the starting point for another study aimed at reducing latency in latency-sensitive applications such as High-Frequency Trading (HFT). Architectures for low latency in-network computing used in Smart Network Interface Cards (SmartNICs) have been studied.
In this context, a rule-match architecture was designed to achieve very low latency.

In parallel, within the context of low latency, the field of quantum computing was explored to design architectures optimized for low-latency quantum error correction.

A challenge encountered during the development of various components, particularly the HBM controller, was the lack of a simulator capable of performing integration tests between individual components under development and the entire system.
This led to the proposal of future work: the development of a unified interface that bridges Electronic System Level (ESL) simulators with Hardware Description Language (HDL) simulators.

Methodology:
For the implementation of HBM controller and architecture components for SmartNIC, the following methodology was adopted and followed.

- **Architectural model:**
  - As a first step, an architectural model of the component was developed.

- **Analysis of the model:**
  - In this phase, the properties of interest on the architectural model were analysed.
  - In the case of HBM controllers, a detailed mathematical analysis was performed to calculate the worst-case latency (WCL) before each command to the memory (activate, precharge, read, write), then the whole request (load, store).
  - In the case of SmartNIC architecture, the response time of the component from sending a message to it has been estimated.

- **Development:**
  - In this phase, a prototype of the component was developed using the model produced in the previous phase as a reference.
  - The hardware description language SystemVerilog was used.
  - The AMD Vivado toolchain was used.

- **Validation:**

- In this phase, the developed component has been tested and validated against the model and the analysis of the starting point
- Mainly HDL simulators (QuestaSim, Vivado Simulator) were used.

- **Deploy and experiments:**
  - In this phase, the developed components were deployed in FPGA and experiments were carried out to evaluate the characteristics of interest (Throughput and Latency)
  - AMD Alveo U250/U280 FPGA accelerators were used.

The phases, especially those of development and verification, were carried out iteratively.

Results:
The research carried out in this first year yielded some significant results.

- **Predictable HBM controller:**
  - A predictable HBM controller in FPGA was developed, showing that using this type of memory interface in a safety critical environment is possible.
  - A peak throughput of 7 GB/s per memory channel, which is 10% higher than a recent state-of-the-art comparable controller design, was obtained.
  - An aggregate peak throughput of 112 GB/s under time predictability guarantees was obtained.
- **Low-latency SmartNIC architecture:**
  - A SmartNIC architecture capable of reacting to messages in less than 400 ns, 12% better than the state-of-the-art solutions in the HFT domain, was developed.

4. Research products:

- *Alessandro Cilardo, Manuel Maddaluno*
  Achieving deterministic High-Bandwidth Memory performance
  IEEE Transactions on Computers (SUBMITTED)

5. Conferences and seminars attended

- ITADATA2024: The 3rd Italian Conference on Big Data and Data Science - BigHPC2024: Special Track on Big Data and High-Performance Computing, Pisa, 17-19 September 2024. Website: https://www.itadata.it/2024/bighpc2024 - Extended abstract presented:

  Manuel Maddaluno, Alessandro Cilardo
  Extended Abstract: An FPGA platform for Latency-Sensitive HPC Applications

6. Activity abroad:

-

7. Activity in partner companies:

-

8. Tutorship

-