Ph.D. in Information Technology: Thesis Defenses

July 19th, 2019

Room Seminari – 11.00 am

Stefano CHERUBIN – XXXI Cycle

"Compiler-Assisted Dynamic Precision Tuning"

Advisor: Prof. **Giovanni Agosta**

Abstract:

Given the current technology, approximating real numbers with finite-precision is unavoidable. Determining which finite-precision representation to exploit for each variable in the program is a difficult task. To face this problem, several precision mix solutions have been proposed so far in the state-of-the-art. However, the best precision mix configuration may vary at runtime along with input data.
In this thesis we aim at suggesting two effective approaches to solve the precision tuning problem.
The first approach follows the static precision tuning paradigm, i.e. it generates a single mixed precision version from the original code, which is designed to be used in place of the original version.
We later allow the possibility of changing the input conditions that may affect the best precision mix configuration. To solve this problem we propose a novel approach and a new toolchain that automatizes a large portion of this process. We present each component of the toolchain, and we provide guidelines to use them properly. We refer to this second approach as dynamic precision tuning.
We evaluate the static and the dynamic precision tuning solutions on a set of high performance computing and approximate computing benchmarks. We show how the dynamic precision tuning toolchain can be used -- under certain conditions -- also for static precision tuning. Our second toolchain is capable of achieving good results in terms of performance gain while maintaining acceptable precision loss threshold.
In the future we aim at further improving this toolchain to extend its applicability to other use cases.
Additionally, we highlight which improvements on the current toolchain may provide greater benefits on the quality of the output.

Danilo FILGUEIRA MENDONÇA – XXX Cycle

"Self-Management of Geographically Distributed Infrastructures and Services"

Advisor: Prof. **Luciano BARESI**

Abstract:

The paradigm of edge computing re-emerged in the last years to fill the gap between cloud data centers and information prosumers at the network edge. Computing and storage resources are located in proximity to maximize network performance and enable latency-sensitive and data-intensive applications. The

decentralized nature of edge computing entails many challenges. Resources are limited and must be managed efficiently to allow more customers and services to be admitted into the system.
Operational aspects like deployment, placement, and scaling of geo-distributed services make automation and self-management first-class requirements. Typical cloud computing models and techniques are not suitable or need to be adapted.
As contributions, we first proposed a Serverless Architecture for Multi-Access Edge Computing. We then expanded our landscape with heterogeneous resources from mobile, edge, and cloud platforms ---which we refer to as the Mobile-Edge-Cloud Continuum. To tackle the life-cycle of serverless functions deployed to the Continuum, we proposed A3-E. We concluded our contributions with PAPS, a framework that tackles the efficient and effective placement and scaling of serverless functions onto densely distributed edge nodes through multi-level self-management. Our solutions were evaluated through comprehensive experiments. Results show the feasibility and benefits of the proposed serverless architecture and management frameworks.

PhD Committee:
Prof. **Giovanni Agosta**, DEIB
Prof. **Pietro Braione**, Universita' di Milano Bicocca
Prof. **Salvatore Distefano**, Universita' di Messina