

# Gianluca Sabella

## Innovative benchmark for evaluating heterogeneous computing resources and High-Performance clusters deployed via Kubernetes

Tutor: Prof. Carlo Sansone

co-Tutor: Prof. Elvira Rossi

Cycle: XXXVIII

Year: First

# My background

- MSc degree in **Computer Science** from University of Naples “Federico II”
- Research groups: PICUS, ICSC - Spoke2 – WP2 and WP5, INFN Sezione di Napoli
- PhD start date: 01/01/2023
- Scholarship type: no funded scholarship

# Research field of interest

- ***High-Performance Computing (HPC) heterogeneous resources and benchmarking:***  
Delving into the realm of supercomputers and computing architectures that combine CPU and GPU resources, evaluating and leveraging benchmarks to optimize scientific applications
  
- ***High Performance Resource Containerization:***  
Exploring lightweight, easily deployable container technology for efficient application delivery and efficient use of high-performance computing resources

# Summary of study activities

## Ad hoc PhD courses

- IoT Data Analysis
- Percorso per il rafforzamento delle competenze sulla progettazione europea

## Courses

- First course about the porting on GPUs of code and algorithms
- Ethics and AI

## School

- *The Fifth International School on Open Science Cloud (SOSC 2023)*

## Events attended

- ATLAS Italia Computing annual meeting, Genova, Italy, 27/11/2023 to 29/11/2023, presenting author - Le attività del T2 di Napoli con Kubernetes: il caso d'uso delle analysis facilities
- SPOKE 2 annual meeting, Bologna, Italy, 18/12/2023 to 20/12/2023, presenting author

Courses	Seminars	Research	Tutorship	Total
16,5	5,4	39	0	60,9

# Research activity (1): Overview

## ❑ **Problem:**

Insufficient benchmarks exist to accurately gauge the performance of heterogeneous computing resources (combining GPU + CPU) in today's scientific and technological sphere.

## ❑ **Objective:**

Defining a new benchmark to precisely and realistically assess the performance of diverse computing resources under real workloads. This initiative aims to bridge the current gap in scientific computing by offering a more comprehensive evaluation.

## ❑ **Methodology:**

- *Identifying and analyzing* requirements from real-world workloads.
- Defining, implementing, and evolving a benchmark algorithm based on principles such as the Schur Complement.
- Testing heterogeneous systems to assess their scalability.
- Comparing the obtained results with existing benchmark solutions

# Research activity (2): Overview

## ❑ **Problem:**

The need to expedite application development and deployment in high-performance computing contexts. These procedures can often become lengthy and complicated due to the complexities of establishing suitable infrastructures and configuring specific development environments. Addressing the need for quick setups is crucial to adapt to dynamic market demands and efficiently meet end-user needs.

## ❑ **Objective:**

To create a versatile computing platform that effortlessly hosts diverse applications and services, efficiently managing various workloads.

## ❑ **Methodology:**

- Implementing a container cluster leveraging Kubernetes as orchestration system
- Studying the best configuration and optimizing resources to ensure flexibility and adaptability across various applications and services
- Employing diverse use cases as stress tests to showcase our cluster's ability to adapt to a wide array of applications and computing resources. This approach aims to highlight the versatility of our infrastructure in successfully managing multiple workloads while ensuring efficiency and scalability.

# Next year

## □ Problem:

1. How does the advantage of a quick and user-friendly application deployment impact performance?
2. *What is the behavior of a high-performance system in the case of virtualized resources in terms of overhead?*

## □ Objectives:

To answer all these questions, the need for a benchmarking system designed for virtual resources is crucial.

The goal will be to converge the two problems studied this year, and to integrate GPUs into the infrastructure, to test new types of use cases (e.g., deep learning use cases)

# Products

[P1]

Carracciuolo, L, Mele, V, Sabella, G. Toward a new linpack-like benchmark for heterogeneous computing resources. *Concurrency Computat Pract Exper.* 2023;e7962. doi: 10.1002/cpe.7962

Published