

Introduction to GPUs and Parallel Computing

Prof. Jacopo Pantaleoni¹

¹ NVIDIA Research
Email: jpantaleoni@nvidia.com

Timetable: 12 hrs. Lectures on June 15-16, 2015. First lecture on June 15, 2015, (schedule to be confirmed) (dates already fixed, see the calendar), Torre Archimede, LabTA on 2nd floor.

Course requirements: C/C++ programming basic elements of processor architecture

S.S.D.: INF/01

Course contents:

With the exponential growth in parallelism of contemporary processor architectures, parallel computing is becoming a pervasive reality used across all fields of science and entertainment, from climate modeling, to combustion engine simulation, from cancer research to drug discovery, from visual effects to social networks.

This course will first provide a brief introduction to the fundamental limits of technology scaling that led to the development of massively parallel processors, as well as giving a rough overview of a modern many-core architecture, and later focus on the implications for algorithm development.

Particularly, this course will focus on the paradigm change required by parallel computing, analyze a suite of parallel constructs that form the building blocks of most parallel programs, and give an overview of the basic instruments of GPU programming, from the lowest level programming languages (CUDA), to directive-based C++ extensions (OpenACC/OpenMP), to higher level template libraries (Thrust/CUB).

The course will be accompanied by a series of practical exercises of growing complexity.