



Computing Infrastructures

 POLITECNICO DI MILANO



Course Introduction

Prof. Gianluca Palermo



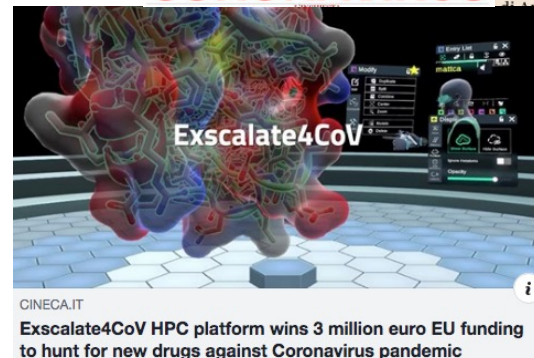
- **Email:**
 - gianluca.palermo@polimi.it
 - SUBJECT: CI25 - <actualSubject>
- **Phone:**
 - (+39 02 2399) 3552
- **Research:**
 - Embedded and HPC Systems
 - Accelerating SW for Molecular Docking in Drug Discovery applications on HPC architectures
 - GPU programming
- **Teaching Assistant:**
 - Roberto Rocco
 - Fault tolerance in MPI-based HPC applications
 - HPC-Quantum Integration



Virus, supercalcolo a caccia della terapia. «Testeremo 500 miliardi di molecole»

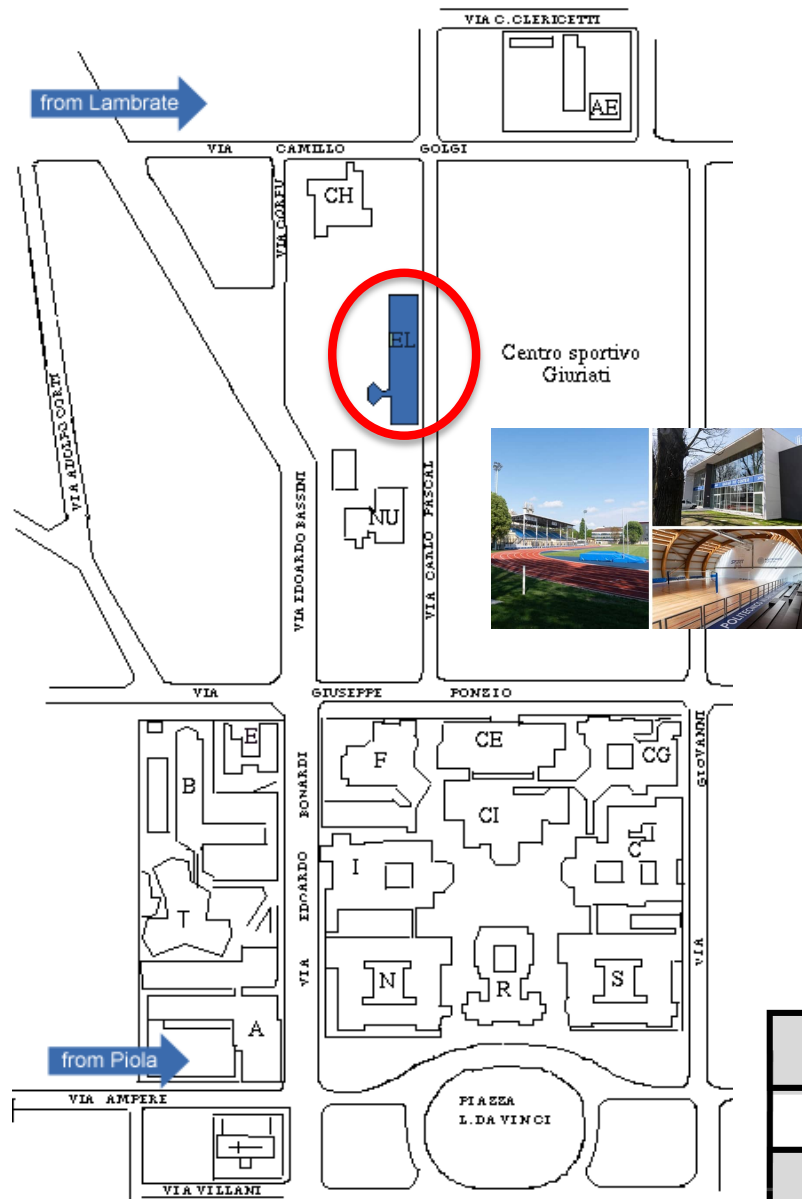
Progetto guidato da Dompé. Già in corso lo screening su 10mila farmaci esistenti. Al Politecnico di Milano il compito di accelerare il processo computazionale

di Antonio Larizza





Office Hours

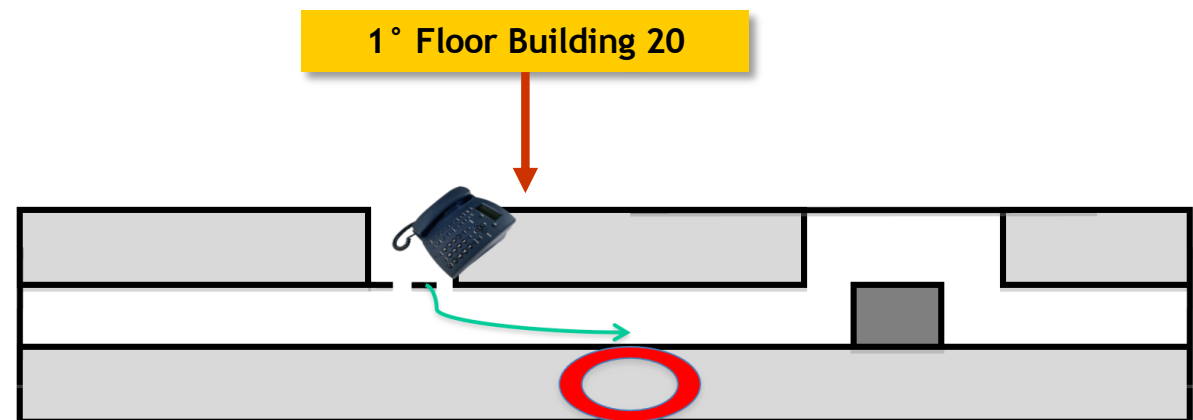


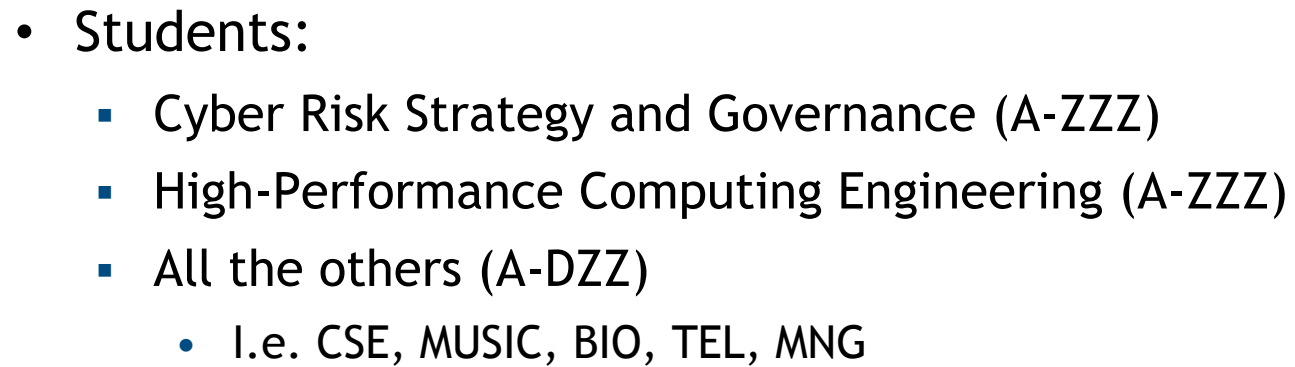
• Office

- DEIB - Dipartimento di Elettronica, Informazione e Bioingegneria
- Building 20 - 1st Floor - Room 124
- On request

• Microsoft Teams:

- Asynchronous and Remote Interactions with students





	TIME	ROOM
Monday	14.15-16.15	5.0.1
Tuesday	12.15-14.15	T.2.1



NOT USED UNLESS THERE ARE SPECIFIC NEEDS



- Central Point Of Synchronization for the course

095898 - COMPUTING INFRASTRUCTURES (PALERMO GIANLUCA)

Introduzione
Official website for the Computing Infrastructure course (Prof. Palermo)

Course Schedule

Notice board
Check course announcements. Interact with the professor, tutors and your fellow students via the forum.

Materials
Check the materials provided by the professor.

Recordings
Rewatch the recorded course lectures.

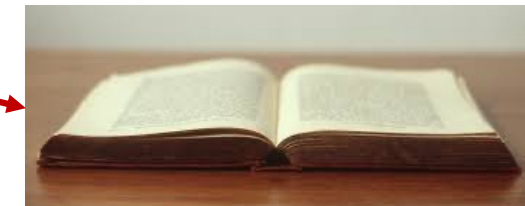


Dates, Topic, Link to Material



Announcements

Forum



Books

Slides

Exercises

Additional Readings

Exams

See Next Slide



095898 - COMPUTING INFRASTRUCTURES (PALERMO GIANLUCA) [2023-24]

2023-24



VIDEO OF THE LECTURES

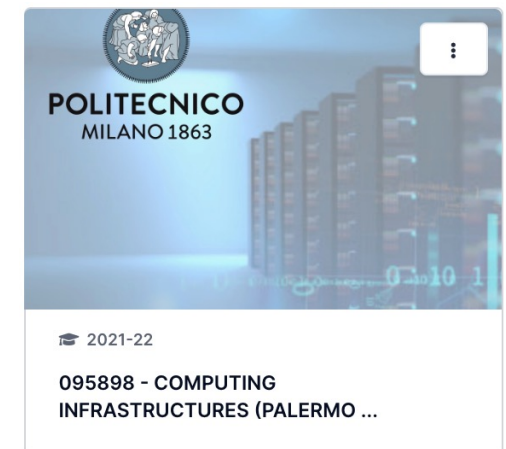


- Lectures will be recorded and made available
 - WEBEEP->Schedule
- **HOWEVER, the course is IN PRESENCE!**
 - **Videos are not thought for offline use**
- Videos on previous editions are also be available
 - AY 2022-2023 /2023-2024
 - Same content, a few updates on the slides & material
 - AY 2020-2021
 - Recorded during pandemic (with Prof. Roveri)



095898 - COMPUTING
INFRASTRUCTURES (PALERMO
GIANLUCA) [2023-24]

2023-24



POLITECNICO
MILANO 1863

2021-22

095898 - COMPUTING
INFRASTRUCTURES (PALERMO ...



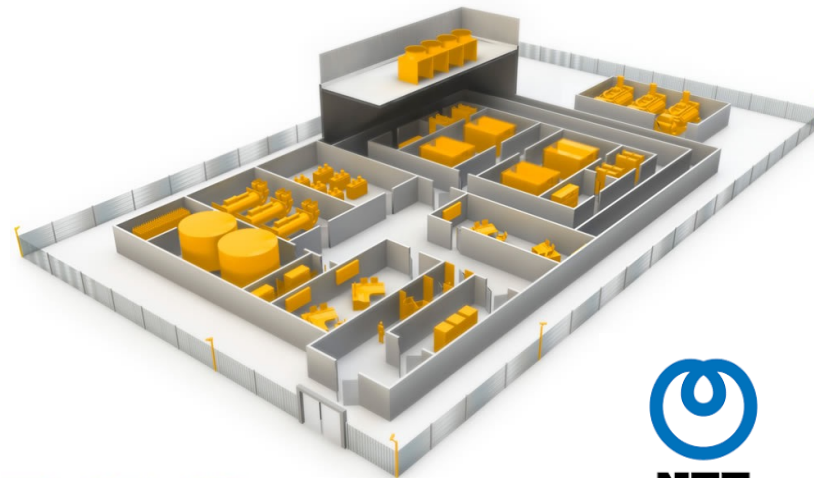
Introduction to the course



- Modern large-scale datacenters require the seamless integration of different components - applications, computation nodes, storage devices, and networks - into one computing infrastructure.



Google Cloud



- The course covers the basics of current datacenters architectures, ranging from the analysis of the single components to the global infrastructure.



Why should matter? (Some Examples)



Understand Infrastructure Complexity: Stress the interplay of networking, power, cooling, and hardware IT infrastructure.



Scalability and Resource Planning: In the context of expansion strategies, as computational needs grow (e.g., HPC clusters with more nodes or GPUs for AI workloads), the supporting infrastructure must grow with it.



Multidisciplinary Nature of the Engineering Elements involved: Show the role of mechanical/electrical engineering aspects (for cooling, power) and industrial designers (for layout, maintenance optimization, availability) in relation to CSE/HPC requirements.



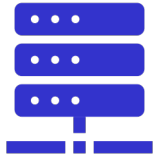
Engineering Constraints: Cooling and power design decisions affect hardware choices (e.g., density of racks, GPU-based computing). Importance to understand which innovation opportunity can help in moving such constraints.



Final Goal: Understanding datacenter infrastructure connects theory (like algorithms, distributed computing) with real-world challenges (power, space, cooling). This holistic view equips future engineers to create, operate, and optimize sustainable, high-performance computing systems. **OPEN THE BOX!!!**



The topics of the course



HW Infrastructures:

System-level: Computing Infrastructures and Data Center Architectures, Rack/Structure;

Node-level: Server (computation, HW accelerators), Storage (Type, technology), Networking (architecture and technology);

Building-level: Cooling systems, power supply, failure recovery



SW Infrastructures:

Virtualization: Process/System VM, Virtualization Mechanisms (Hypervisor, Para/Full virtualization)

Computing Architectures: Cloud Computing (types, characteristics), Edge/Fog Computing, X-as-a service



Methods:

Reliability and availability of datacenters (definition, fundamental laws, RBDs)

Disk performance (Type, Performance, RAID)

Scalability and performance of datacenters (definitions, fundamental laws, queuing network theory)



Bibliography

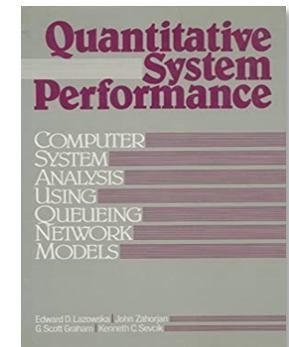
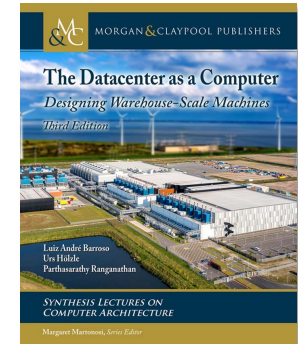


- Luiz André Barroso and Urs Hölzle, **The Datacenter as a Computer: Designing Warehouse-Scale Machines**. 3rd Edition (2018)
- Edward D. Lazowska, John Zahorjan, G. Scott Graham, Kenneth C. Sevcik, **Quantitative System Performance: Computer System Analysis Using Queueing Network Models**. (Ch. 1-6)
- *Additional Material provided in WEBEEP*



- **Further Readings:**

- Caesar Wu and Rajkumar Buyya, “Cloud Data Centers and Cost Modeling”
- James Smith and Ravi Nair, “Virtual Machines - Versatile Platforms for Systems and Processes”
- Massimo Lazzaroni Loredana Cristaldi Lorenzo Peretto Paola Rinaldi Marcantonio Catelani “Reliability Engineering: Basic Concepts and Applications in ICT”
- Giuseppe Serazzi, “Performance Engineering: Learning Through Applications Using JMT”



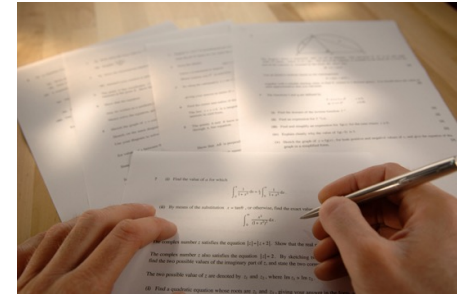
PDFs are
freely
available



Evaluation



- The course will have a **written exam (Closed book)**
- The exam will consist in:
 - a **set of exercises**, i.e., simple problems to solve (similar to the ones that will be covered during the classes)
 - **Open/Closed questions** dealing with general topics covered by the course
- *An oral examination is possible but **only** at the discretion of the professor*
- *Erasmus/Exchange Students*
 - *We follow the School directives -Oral examination in a different moment w.r.t. the exam*
- Exercises and some previous exams are available
 - (see next slide)





Exam Structure (Example)



10 T/F Questions

6 Exercises

2 Open Questions

Answer Sheets

Overall Grade
(10*1) + (6*2) + (5+5)



Exam Statistics from 2023-2024



- Data on June-July-September-January exams
 - February not yet finalized
- Some number
 - 231 students passed (358 overall “formally attending”)
 - 40 still under evaluation
 - 328 exams corrected
 - Average exam per student passed: 1.41
 - Rejected grade by the student: 19.5%
 - Average grade: 25.7
 - Distribution: See plot

