From HPC and GRID infrastructures to digital ecosystems

Stefano Cozzini

Belgrade, June 12 2026

Introduction & Objectives

A Rapid Journey Through Two Decades

Goal: Trace the transformation of scientific computing infrastructure.

Starting Point: The Landscape of 2002 (Grid & HPC era).

Ending Point: Today Landscape (AI, Exascale, Cloud-Native).

Focus:

- Technological shifts (Hardware/Software).
- Human roles

AREA SCIENCE PARK

Why 2002 ?

SCHOOL IN

High Performance Computing on Linux Clusters

4 - 15 FEBRUARY 2002
Miramare, Trieste, Italy

A two-week School in "High Performance Computing on Linux Clusters" will be held at the Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy, from 4 to 15 February 2002, with the co-sponsorship of the Istituto Nazionale per la Fisica della Materia (INFM), Italy. The School will be directed by Stefano Cozzini (INFM, udr SISSA, Italy), with Alvise Nobile (ICTP) as the local organizer.

PURPOSE AND NATURE

The availability of high-speed networks and increasingly powerful commodity microprocessors are making the usage of clusters (or networks of computers) an appealing vehicle for cost effective parallel computing. Clusters, built using commodity-of-the-shelf (COTS) hardware components, as well as free, or commonly used, software, are playing a major role in redefining the concept of high performance computing. The School aims at providing the skills needed to benefit from this generation of HPC solutions, giving a basic knowledge of programming, administering and tuning, as well as building Linux-based clusters.

Provisionally, lectures will be on:

- Designing and building linux clusters
- Network solutions for linux clusters (characteristics and performances)
- Parallel programming techniques
- Optimization and profiling techniques for PC-based clusters
- System Management of a linux cluster
- Advanced operating system issues: parallel file systems and I/O

These lectures will be complemented by hands-on lab sessions where participants, grouped in small teams, will build their own cluster from scratch. During the first week of the school these clusters will be assembled, configured and tested. In the second part of the course, some representative parallel programmes in the areas of computational condensed matter, engineering and weather forecasts will be distributed to the participants. These parallel codes will be presented in a series of tutorials illustrating their usage and the parallelisation strategies they adopt. During the lab sessions, the participants will be asked to install, analyse and profile some of these codes on the previously built clusters.

the
abdus salam
international
centre
for theoretical
physics

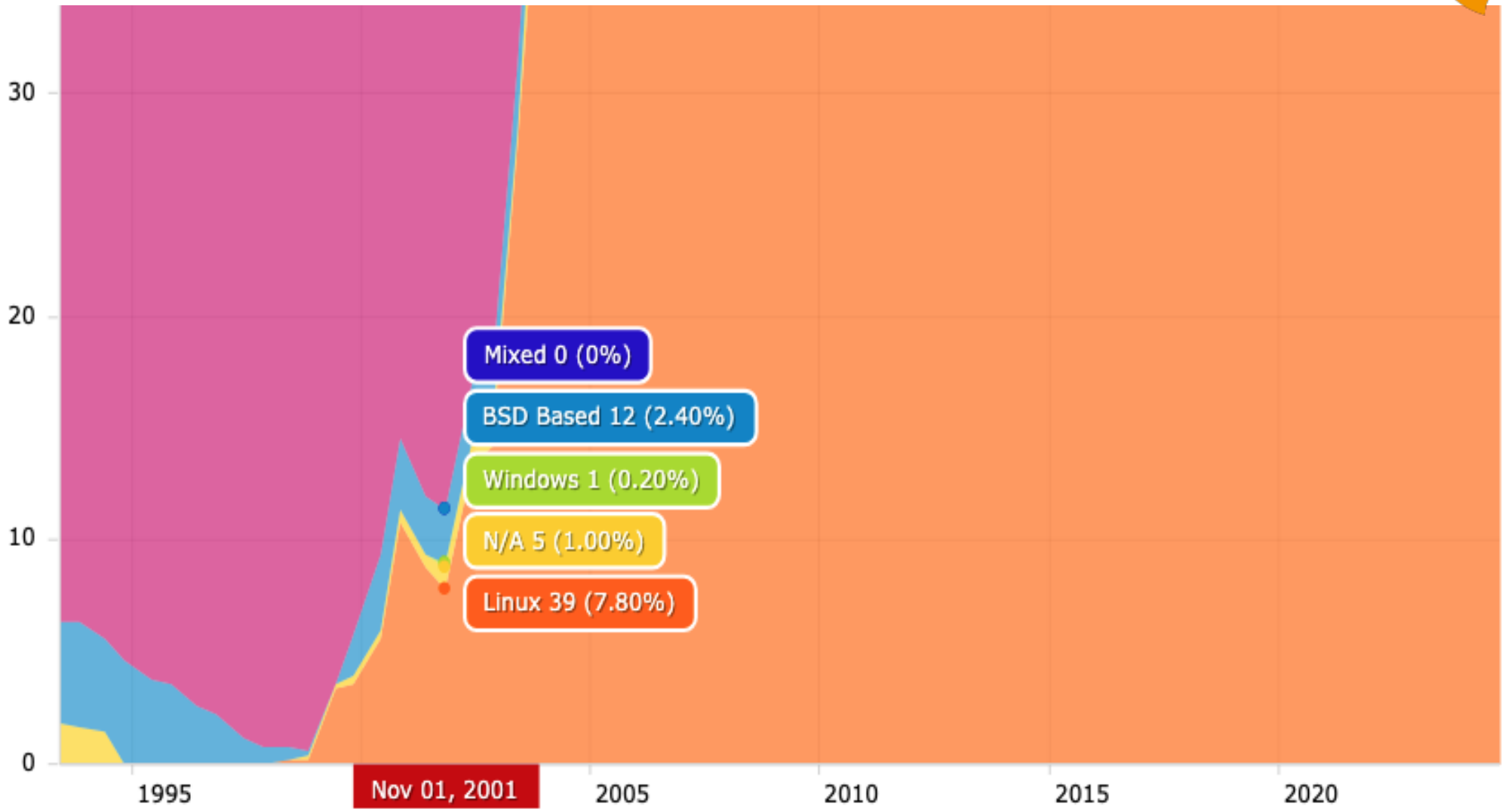
CO-SPONSORED BY
I.N.F.M.
(Istituto Nazionale per la Fisica della Materia)

DIRECTOR
S. Cozzini
(INFM, Italy)

LOCAL ORGANIZER
A. Nobile
(ICTP)

AREA SCIENCE PARK

Top500: november 2001



Contribution

Student presentation: "Some Improvements on Monte Carlo Calculations in QFT".

ICTP/INFM School in High Performance Computing on Linux Clusters
Trieste, Italy, 31 January – 15 February 2002

Some Improvements on Monte Carlo Calculations in QFT

Antun Balaž

E-mail: antun@phy.bg.ac.yu
Institute of Physics, Belgrade, Yugoslavia
http://www.phy.bg.ac.yu/

2002 HPC Landscape

The Era of Dedicated Clusters

Dominant Paradigm: High-Performance Computing (HPC) & Early Grid.

Primary Users: Large scientific communities (Physics, Astronomy, Climate).

Infrastructure:

- Physical clusters located in university and national labs
- hundreds of cpu
- Dedicated network links (myrinet)

Access: Batch processing, long queue times, remote login..

GRID computing era (2000-2010)

Compute as a Utility: The Electrical Grid

The Vision: Just as you plug a device into a wall socket without knowing which power plant generated the electricity, researchers should plug into the Grid without knowing which supercomputer runs their job.

The great GRID success: WLHC

Worldwide LHC Computing Grid

A key tool for physics

The most sophisticated data-taking & analysis system ever built for science, providing near real-time access to LHC data.

Seamless access

Computing resources which include data storage capacity, processing power, sensors, visualization tools and more.

Global collaboration

40+ countries
170 computing centres
Over 1 million computer cores
2 exabytes of storage

Enabling discovery

WLCG computing enabled physicists to announce the discovery of the Higgs Boson on 4 July 2012.

The Worldwide LHC Computing Grid (WLCG) is a global computing infrastructure whose mission is to provide computing resources to store, distribute and analyse the data generated by the Large Hadron Collider (LHC), making the data equally available to all partners, regardless of their physical location.

WLCG is the world's largest computing grid. It is supported by many associated national and international grids across the world, such as European Grid Initiative (Europe-based) and Open Science Grid (US-based), as well as many other regional grids.

WLCG is co-ordinated by CERN. It is managed and operated by a worldwide collaboration between the experiments (ALICE, ATLAS, CMS and LHCb) and the participating computer centres. It is reviewed by a board of delegates from partner country funding agencies, and scientifically reviewed by the LHC Experiments Committee.

GRID computing era (2000-2010)

Compute as a Utility: The Electrical Grid

The Reality

- **Complexity:** Unlike a wall socket, the "plug" required complex certificates, manual configuration, and job scripts.
- **Reliability:** Power is stable; Grid jobs often failed due to network glitches or site maintenance.
- **Usability:** Required specialized training (SysAdmin level) rather than simple access.

Our challenge: use it outside HE and Europe.

Intercontinental cooperation - Iran

PURPOSE AND

Grid computing is
science in the next y
world, among whic
opportunity to learn
areas, providing an
constraints that the g
applications present

MAIN OBJECTIV

The Workshop aim
introduction to ho
computational scie
experience of the current middleware and GRID infrastructures available in the Indian subcontinent, namely Swelanka and EUINDIAGRID, focusing on a few case studies where scientific applications were ported and successfully executed on

PROGRAMME

The School will consist of theoretical lectures given in the morning, and laboratory sessions in the afternoon. The first week of the school will focus on the needs of beginner and intermediate level participants whereas the second will be dedicated to more advanced topics and giving researchers the opportunity to apply the techniques learned in the first week to their own projects. Previous

Co-sponsored by

e-Lab, joint SISSA/DEMOCRITOS Laboratory for e-Science

IPM-Grid

DIRECTORS

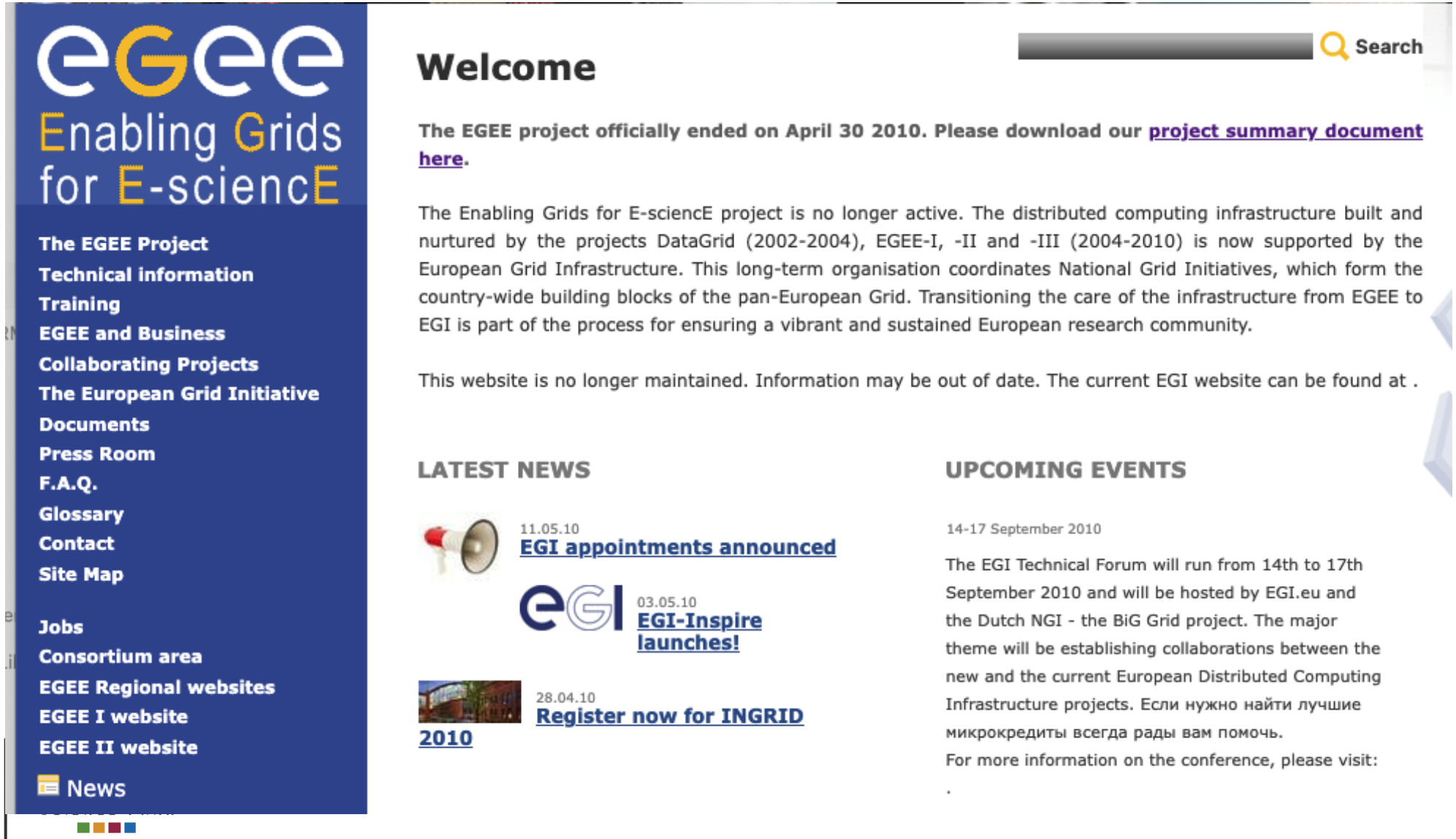
S. Cozzini *
CNR-IOM/Democritos and e-Lab/SISSA, Italy

R. Asgari and S. Rouhani **
Institute for Studies in Theoretical Physics and Mathematics, Tehran, The Islamic Republic of Iran

A. Balaz
Scientific Computing Laboratory, Institute of Physics, Pregrevica Belgrade, Serbia

SCIENCE PARK

EGEE comes to an end (2010)



Legacy of the EGEE

Success: Proved distributed computing at scale works.

Failure:

- Required too much expertise (SysAdmin-level skills).
- Few scientific communities on board
- HPC and GRID remains quite separate

Lesson: Infrastructure must be **invisible** to the end-user.

The 2010x Transition

Virtualization & The Cloud Wave

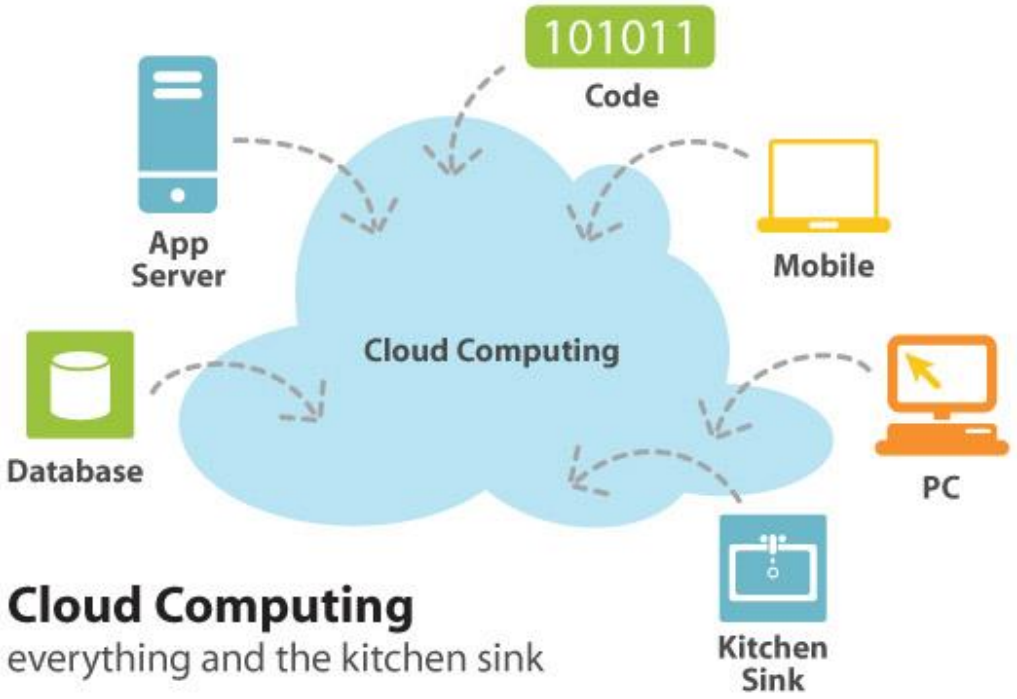
Shift: From Bare Metal to Virtual Machines (VMs

Drivers:

- Cost efficiency (multi-tenancy).
- Flexibility in resource allocation.

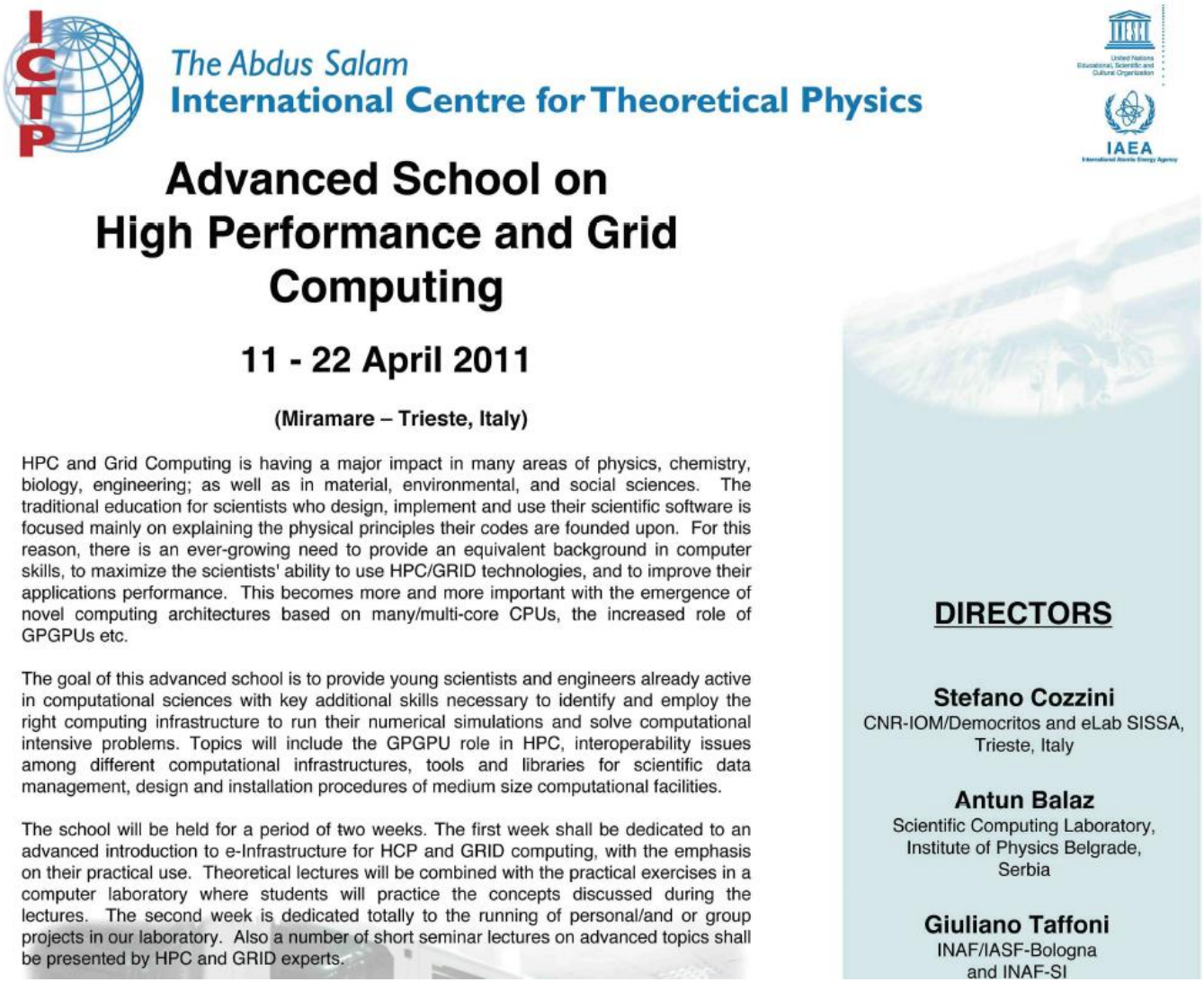
Emerging Tech:

- OpenStack for private clouds.
- AWS/Azure for burst computing.



Impact: Decoupling software from specific hardware.

But we still believed in GRID :-)



The Abdus Salam
International Centre for Theoretical Physics

IAEA

Advanced School on High Performance and Grid Computing

11 - 22 April 2011

(Miramare – Trieste, Italy)

HPC and Grid Computing is having a major impact in many areas of physics, chemistry, biology, engineering; as well as in material, environmental, and social sciences. The traditional education for scientists who design, implement and use their scientific software is focused mainly on explaining the physical principles their codes are founded upon. For this reason, there is an ever-growing need to provide an equivalent background in computer skills, to maximize the scientists' ability to use HPC/GRID technologies, and to improve their applications performance. This becomes more and more important with the emergence of novel computing architectures based on many/multi-core CPUs, the increased role of GPGPUs etc.

The goal of this advanced school is to provide young scientists and engineers already active in computational sciences with key additional skills necessary to identify and employ the right computing infrastructure to run their numerical simulations and solve computational intensive problems. Topics will include the GPGPU role in HPC, interoperability issues among different computational infrastructures, tools and libraries for scientific data management, design and installation procedures of medium size computational facilities.

The school will be held for a period of two weeks. The first week shall be dedicated to an advanced introduction to e-Infrastructure for HCP and GRID computing, with the emphasis on their practical use. Theoretical lectures will be combined with the practical exercises in a computer laboratory where students will practice the concepts discussed during the lectures. The second week is dedicated totally to the running of personal/and or group projects in our laboratory. Also a number of short seminar lectures on advanced topics shall be presented by HPC and GRID experts.

DIRECTORS

Stefano Cozzini
CNR-IOM/Democritos and eLab SISSA, Trieste, Italy

Antun Balaz
Scientific Computing Laboratory, Institute of Physics Belgrade, Serbia

Giuliano Taffoni
INAF/IASF-Bologna and INAF-SI

AREA SCIENCE PARK

Top500 Evolution

2008: a new metric added

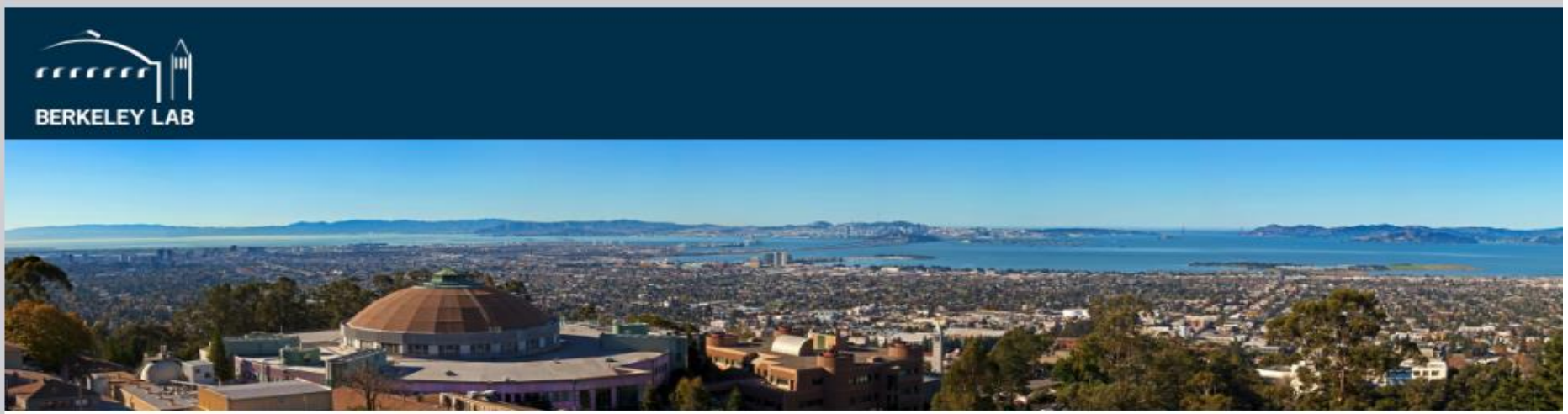
Power consumption (Mwatt)

2013: The first Green500

Metric Shift: Energy efficiency (FLOPS/Watt) begins to matter.

Green500 Rank	MFlops/ watts	Site	System	Total Power(kW)
1	3208.8	CINECA	Eurotech Aurora HPC 10-20, Xeon E5-2687W 8C 3.100GHz, Infiniband QDR, NVIDIA K20	30.7

Exascale expectation



Why we need Exascale and why we won't get there by 2020

Horst Simon
Lawrence Berkeley National Laboratory

AASTCS 2: Exascale Radio Astronomy Meeting
Monterey, California
March 31, 2014

Why Was Exascale not reached in time ?

The Power Wall

The Memory/I/O Bottleneck

System Reliability

Software Rewrite Complexity

Exascale at the end...

JUNE 2022

The 59th edition of the TOP500 revealed the Frontier system to be the first true exascale machine with an HPL score of 1.102 Exaflop/s.

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	8,730,112	1,102.00	1,685.65	21,100

AREA SCIENCE PARK

AI & HPC Convergence

The New Paradigm: Simulation + Data + Generative AI

The Convergence:

- **Traditional HPC:** Physics-based simulation (CFD, Climate).
- **AI Workloads:** Data-driven models (LLMs, Foundation Models).

HPC-AI: Hybrid workflows (AI for Science, Digital Twins, Surrogate Models).

Hardware Shift:

- **GPU Dominance:** Standard compute nodes now include NVIDIA Hx00/Bx00 or AMD MI300.
- **Specialized Accelerators:** NPUs and TPUs integrated into HPC clusters.
- **Liquid cooling system**

AREA SCIENCE PARK

SC25: More Pumps than Processors

From: https://www.hpcwire.com/2025/11/26/sc25-observations-more-pumps-than-processors/

From infrastructure to digital ecosystem

DATA is at the core of the scientific process and its appropriate FAIR management links physical space to the digital dimension

AREA SCIENCE PARK

Our modest approach...

← NEWS

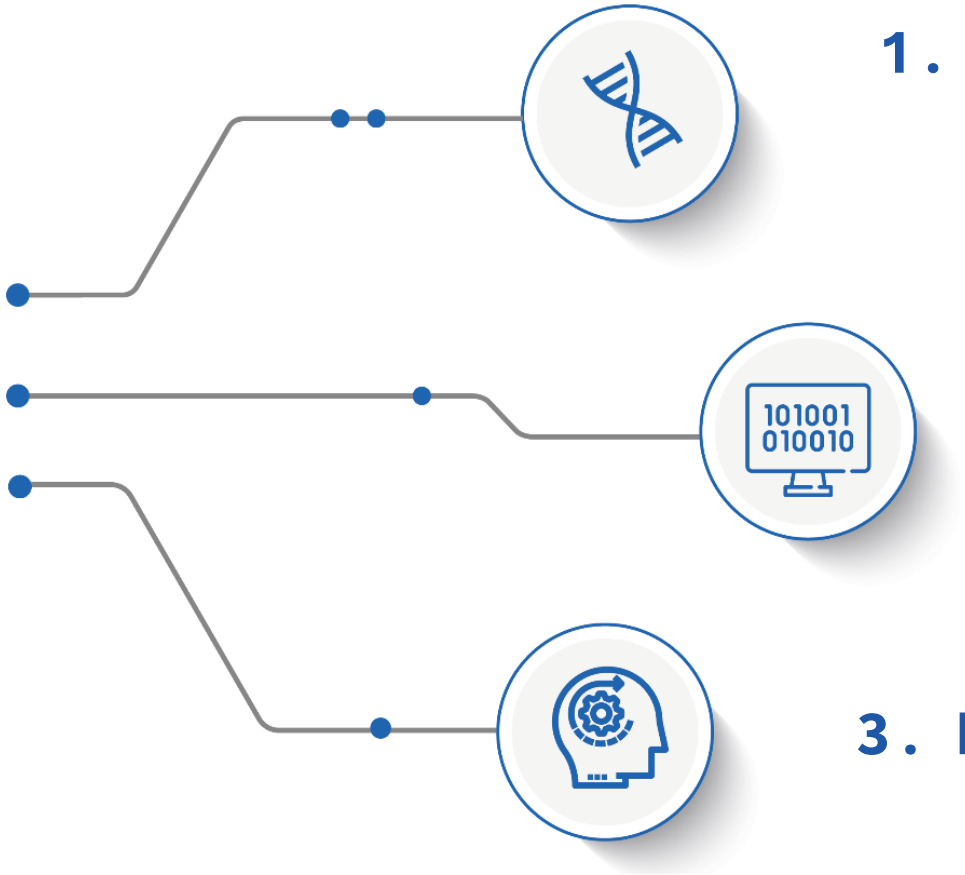
Più potenza e sostenibilità per ORFEO: cresce il Data Center di Area Science Park

22.10.2025

Con un investimento di 3 milioni di euro finanziato dal PNRR, ORFEO rafforza la capacità di calcolo a supporto di ricerca, formazione e trasformazione digitale delle imprese

AREA SCIENCE PARK

Digital Ecosystem components



1. Hardware
   - Orfeo infrastructure

2. Software
   - Tools and Codes
   - Data

3. Brainware
   - People

AREA SCIENCE PARK

Where are the human effort going ?

Training and education– and adapting it to AI tools – remains a focus.

AREA SCIENCE PARK

2014: www.mhpc.it

Master in High Performance Computing

About The Course

The Master in High Performance Computing (MHPC) is an innovative specialization program that prepares students for exciting careers in the fast-growing field of HPC

FOR ANY INQUIRIES, FEEL FREE TO CONTACT US AT INFO@MHPC.IT

2024: www.mdmc.it

SCIENCE PARK

Grazie per l’attenzione!

Area Science Park

Padriciano 99, Trieste-italia

www.areasciencepark.it