Foundations of High Performance Computing

Lecture 1: INTRODUCTION TO HPC

"Foundation of HPC" course

DATA SCIENCE & SCIENTIFIC COMPUTING

2020-2021 Stefano Cozzini

First Week

- 07.10 Introduction to HPC (part1)
- 09.10
  - Morning:
    - Introduction to HPC (part2)
  - Afternoon:
    - Complete presentation (if needed)
    - Q&A section
    - Tutorial 0: accessing the HPC system available for this course

Second Week

- 14.10 Introduction to parallel world:
  - Parallel concepts
  - Parallel Hardware
- 16.10
  - Morning:
    - Introduction to HPC software stacks
  - Afternoon
    - Tutorial 1 :Using an HPC system
    - Assignment 1: Compute scalability on a toy program

Some more information

- Slides and materials of the course available here:

https://github.com/Foundations-of-HPC/Foundations_of_HPC_2020

- For any slot of lecture a directory will be created and informations and materials will be loaded there: i.e for today:
- https://github.com/Foundations-of-HPC/Foundations_of_HPC_2020/tree/master/day1

Day/README.md

Day 1: introduction to course and introduction to HPC

date: Wednesday 07/10/2020

lectures

- Luca Tornatore : Introduction to the course
- Stefano Cozzini : introduction to HPC

references

- from reference 1:
  - section 1.2 for more example on HPC
  - section 1.4 for SLOW model
- FLOPS definition from wikipedia
- HPC from European perspective

additional materials

- a short article on the exaflop day
- what can we do with an exascale machine
- more to come

Before starting: HPC prefix..

Factor	Name	Symbol
10^{24}	yotta	Y
10^{21}	zetta	Z
10^{18}	exa	E
10^{15}	peta	P
10^{12}	tera	T
10^{9}	giga	G
10^{6}	mega	M
10^{3}	kilo	k

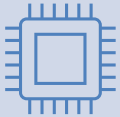
- How many data are produced daily?
- How large is your HD on your laptop ?
- How large is your RAM?
- How powerful is your CPU in your laptop ?
- How large is the L1 cache of your CPU ?

Agenda

- Prologue: why and where HPC ?
- What is HPC ?
- Performance and metrics ?
- Supercomputers and TOP500
- Parallel Computers

Where and Why HPC ?

Traditionally HPC system (a.k.a supercomputers) were confined in research and academic lab…



Today they are everywhere: HPC is now an enabler not just for science but also for business

Today HPC does not necessarily means supercomputers

HPC not easy to define..

High performance computing (HPC), also known as supercomputing, refers to computing systems with extremely high computational power that are able to solve hugely complex and demanding problems.

[Taken from https://ec.europa.eu/digital-single-market/en/high-performance-computing]

Complex problem:

Weather forecast..

Recipe:

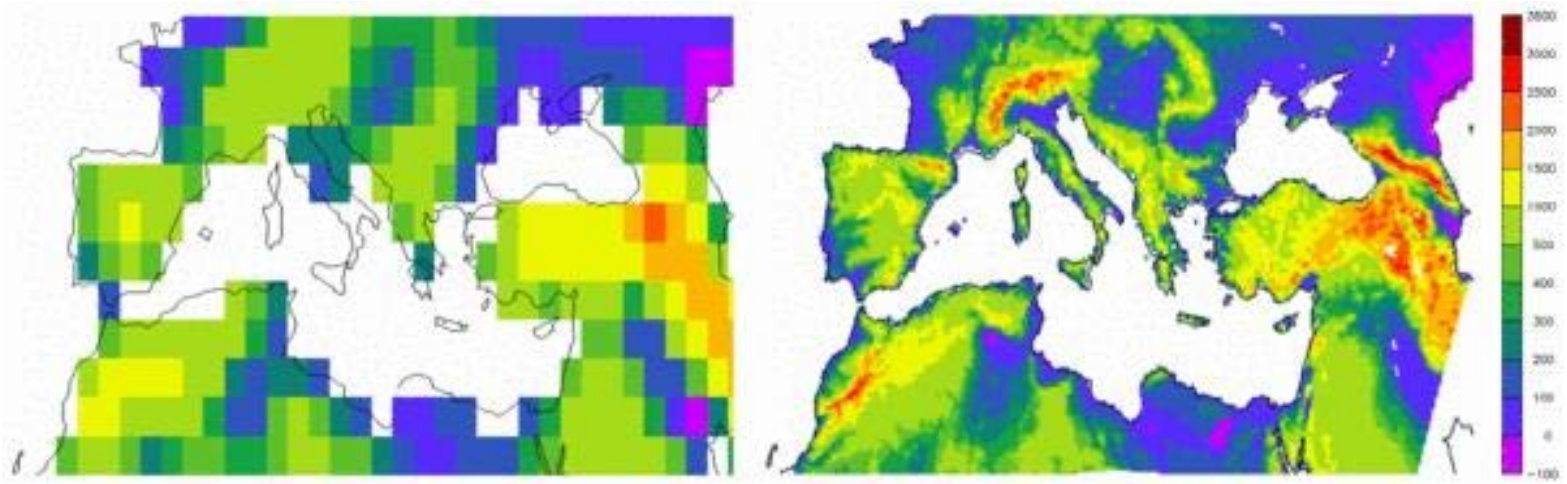
- Define a mathematical model to describe the problem
- Solve it computationally
  - Discretization over a 3d grid
  - Integrate equations
  - Check results..

200 km

25 km

Complex problem: climate change over the Mediterranean sea

- What are the requirements in term of RAM to have decent results ?

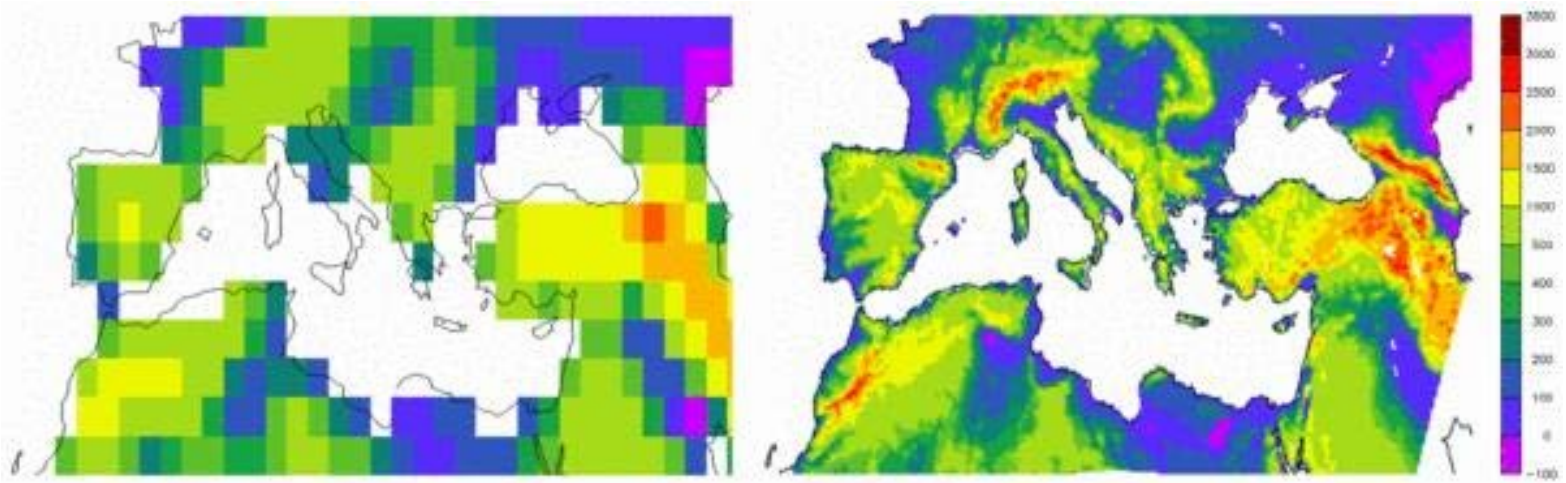


200 km

25 km

Complex problem: climate change over the Mediterranean sea

- Resolution:
  - 200km -> ~ 1GB 2km -> ? GB



200 km 25 km

Complex problems solved by simulations

- Simulation has become the way to research and develop new scientific and engineering solutions.
- Used nowadays in leading science domains like aerospace industry, astrophysics, etc.
- Challenges related to the complexity, scalability and data production of the simulators arise.
- Impact on the relaying IT infrastructure.

200 km

25 km

Other example where to use HPC

HPC capabilities are used to solve and address scientific, industrial and societal challenges.

Health

- Development of personalised and precision medicine to provide individual and accurate patient treatment.
- Saving time and money on the development of new drugs from the initial idea phase to the final phase of reaching the market.

development of new drugs reduced from 17 years	HPC can help save hundreds of millions
↓	
to several months	up to billions of Euros

- Early detection of diseases and quicker diagnosis.

Climate change & weather forecast

- Europe paid severe weather damage costs between 1970 and 2012[1].

150.000 lives — €270 billion in economic damages

- With HPC technology, climate scientists will be able to predict the size and paths of storms and floods with accuracy.
- This will allow implementation of measures such as alerting or evacuating people, thus saving human lives.

Industry

- Reducing development time, minimising costs, optimising decision processes and producing higher quality goods and services.
- E.g. automotive industry will save time and money to develop new vehicle platforms, with improved environmental friendliness and passenger comfort and safety[2].

development reduced from 60 months	HPC can help save
↓	
to 24 months	up to € 40 billions

Cybersecurity

- Enabling complex encryption technologies and better reactions to cyberattacks.
- Combined with Artificial Intelligence, HPC detects:
  - strange systems behaviour,
  - insider threats and electronic fraud,
  - very early cyberattack patterns (few hours instead of a few days)

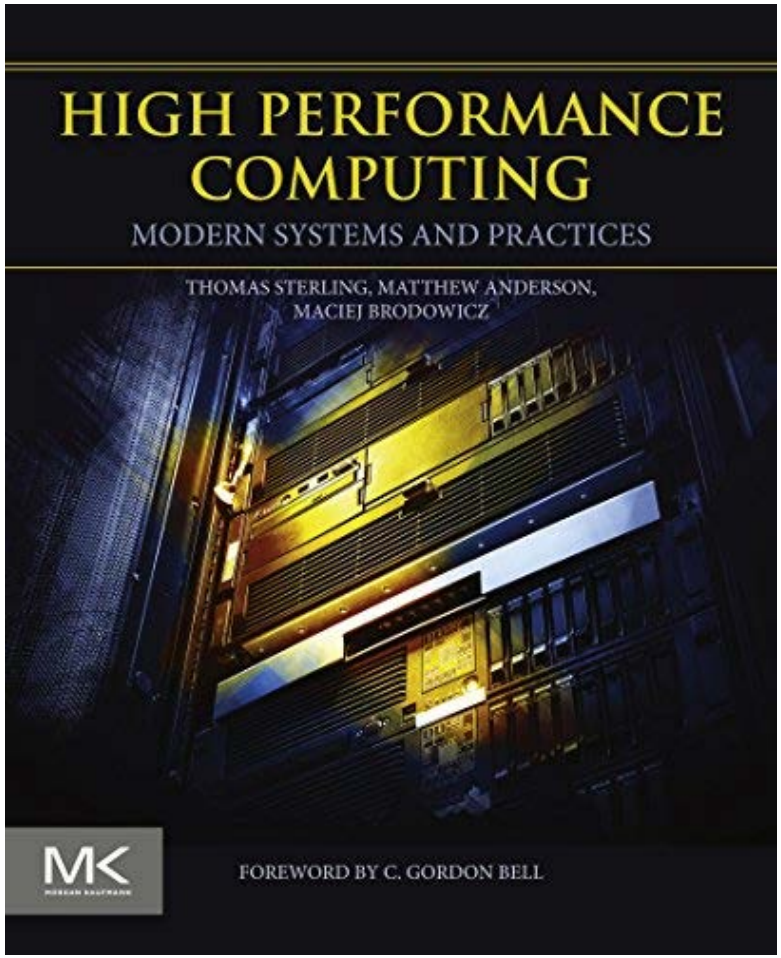
  This will allow automated and immediate actions even before a cyberattack happens.

Energy

- HPC provides critical tools for example in:
  - designing renewable energy parks
  - designing high-performance photovoltaic materials,
  - optimising turbines for electricity production.
- HPC expenditure in the energy sector is projected to grow by 5% in the next years.

Interested in more example ?

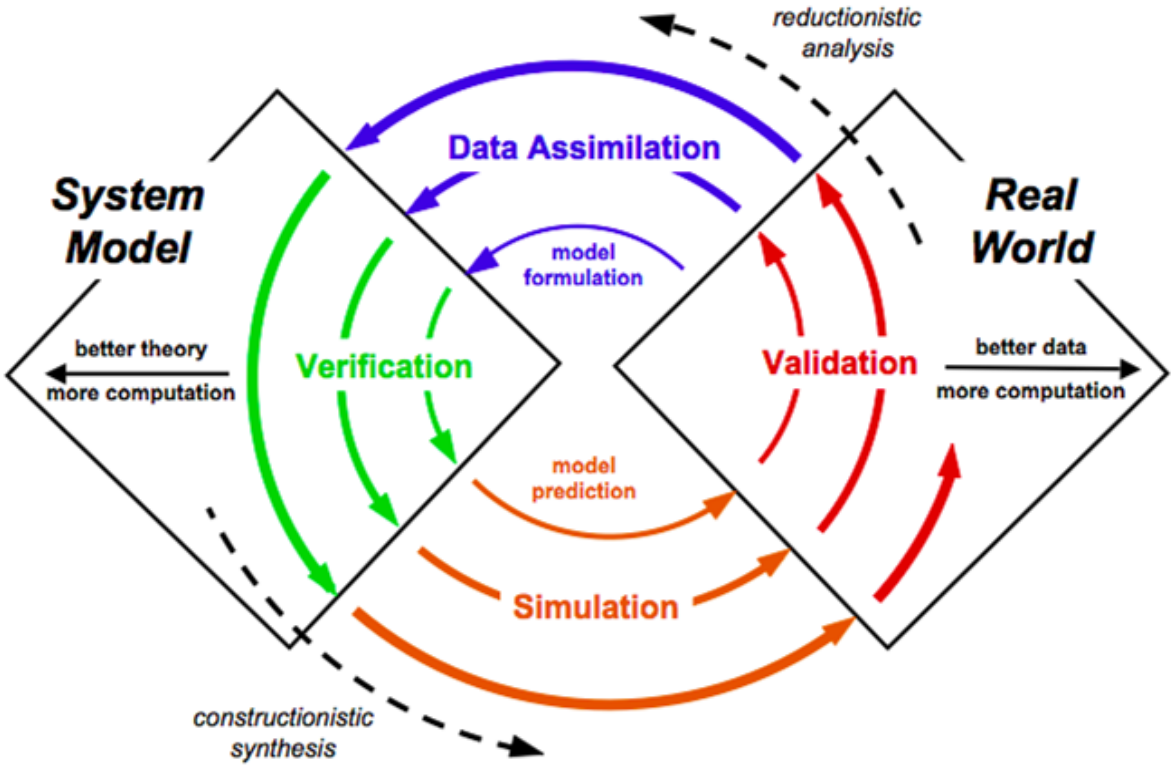
- See chapter one section 1.2 of reference 1
- Look around on the internet..



Research is changing..

- Inference Spiral of System Science

  As models become more complex and new data bring in more information, we require ever increasing computational power



Data are flooding us..

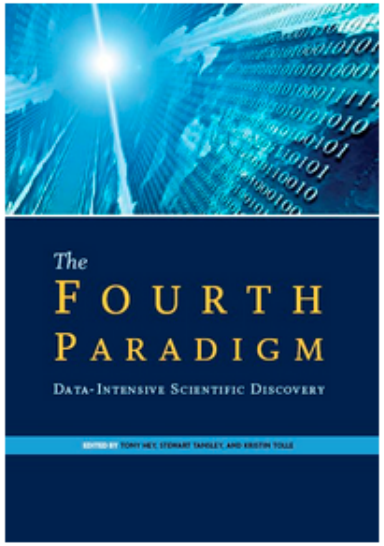
In today's world, larger and larger amounts of data are constantly being generated, from 33 zettabytes globally in 2018 to an expected 175 zettabytes in 2025). As a result, the nature of computing is changing, with an increasing number of data-intensive critical applications. is key to processing and analysing this growing volume of data, and to making the most of it for the benefit of citizens, businesses, researchers and public administrations.

[Taken again from https://ec.europa.eu/digital-single-market/en/high-performance-computing]

Data intensive science

The Fourth Paradigm: Data-Intensive Scientific Discovery

Presenting the first broad look at the rapidly emerging field of data-intensive science



Increasingly, scientific breakthroughs will be powered by advanced computing capabilities that help researchers manipulate and explore massive datasets.

The speed at which any given scientific discipline advances will depend on how well its researchers collaborate with one another, and with technologists, in areas of eScience such as databases, workflow management, visualization, and cloud computing technologies.

In *The Fourth Paradigm: Data-Intensive Scientific Discovery*, the collection of essays expands on the vision of pioneering computer scientist Jim Gray for a new, fourth paradigm of discovery based on data-intensive science and offers insights into how it can be fully realized.

Critical praise for *The Fourth Paradigm*

Download

- Full text, low resolution (6 MB)
- Full text, high resolution (93 MB)
- By chapter and essay

Purchase from Amazon.com

- Paperback
- Kindle version

In the news

- Sailing on an Ocean of 0s and 1s *(Science* Magazine)
- A Deluge of Data Shapes a New Era in Computing *(New York Times)*
- A Guide to the Day of Big Data *(Nature)*

Big data challenge: from HPC to HPDA through AI

- Organizations are expanding their definitions of high-performance computing (HPC) to include workloads such as artificial intelligence (AI) and high-performance data analytics (HPDA) in addition to traditional HPC simulation and modeling workloads.

From https://insidebigdata.com/2019/07/22/converged-hpc-clusters/

Agenda

- ~~Prologue: why and where HPC ?~~ 
- What is HPC ?
- Performance and metrics ?
- Supercomputers and TOP500
- Parallel Computers

HPC: a ~~first~~ second definition

High Performance Computing (HPC) is the use of servers, clusters, and supercomputers – plus associated software, tools, components, storage, and services – for scientific, engineering, or analytical tasks that are particularly intensive in computation, memory usage, or data management

HPC is used by scientists and engineers both in research and in production across industry, government and academia.

[to be continued]

Elements of the HPC ecosystem..

- use of servers, clusters, and supercomputers

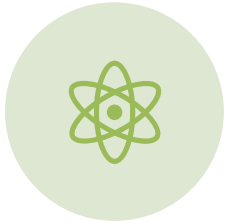
  → HARDWARE
- associated software, tools, components, storage, and services

  → SOFTWARE
- scientific, engineering, or analytical tasks

  → PROBLEMS TO BE SOLVED..

A list of HPC items

COMPUTATIONAL SERVERS



ACCELERATORS

HIGH SPEED NETWORKS

HIGH END PARALLEL STORAGE

IS ALL THIS ENOUGH ?

MIDDLEWARE

SCIENTIFIC/TECHNICAL/ DATA ANALYSIS SOFTWARE

RESEARCH/TECHNICAL DATA

PROBLEMS TO BE SOLVED

Last but not least: people

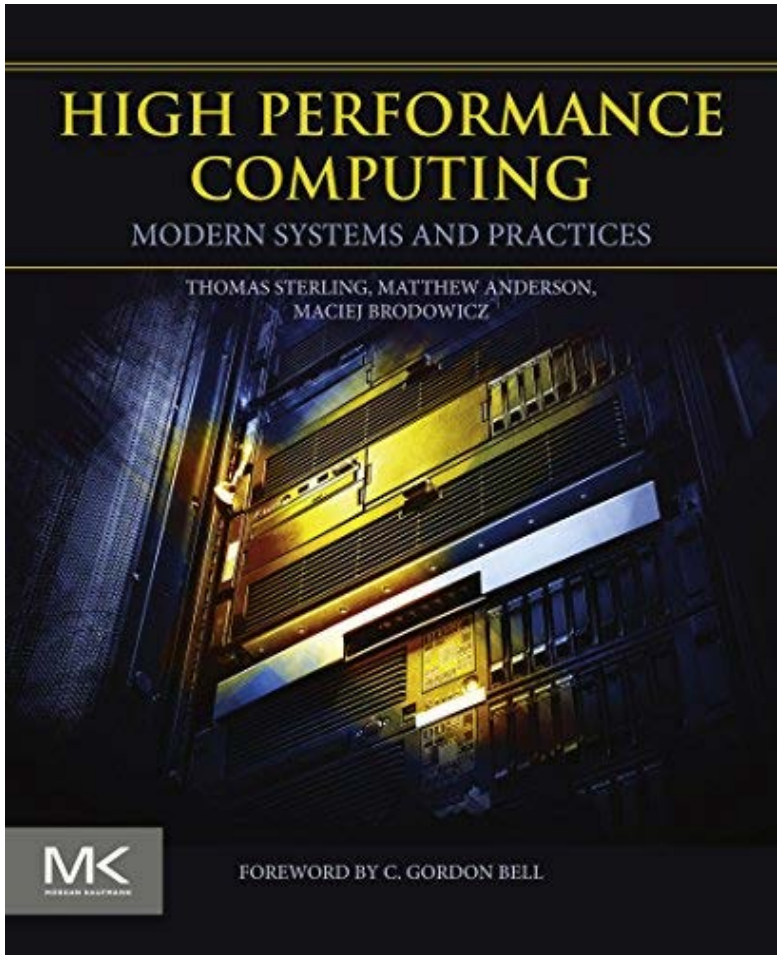
- Human capital is by far the most important aspect
- Two important roles:
  - HPC providers
    - plan/install/manage HPC resources
  - HPC user :
    - use at best HPC resource

MIXING/INTERPLAYING ROLES INCREASES COMPETENCE LEVELS

Yet another definition

- HPC incorporates all facets of three disciplines:
  - Technology
  - Methodology
  - Application

- The main defining property and value provided by HPC is delivering performance for end-user application



Agenda

- ~~Prologue: why and where HPC ?~~ ✓
- ~~What is HPC ?~~ ✓
- Performance and metrics
- Supercomputers and TOP500
- Parallel Computers

It is all about Performance

- It is difficult to define Performance properly

  "speed" / "how fast" are vague terms
- Performance as a measure again ambiguous and not clearly defined and in its interpretation
- In any case performance it is at core to HPC as a discipline
- Let discuss it in some details

Does P stand just for Performance ?

- Performance is not always what matters..

to reflect a greater focus on the productivity, rather than just the performance, of large-scale computing systems, many believe that HPC should now stand for High Productivity Computing. [from wikipedia]

- P should also stand for PROFITABILITY

Performance vs Productivity

- A possible definition:
  - Productivity = (application performance) / (application programming effort)
- Example:
  - To speed up a code by a factor of two it takes 6 months work
  - does this deserve to be done ?

- people in HPC arena have different goals in mind thus different expectations and different definitions of productivity.

- Suggestion: Understand which kind of productivity are you interested in

How do measure (basic) performance of HPC systems

- How fast can I crunch numbers on my CPUs ?
- How fast can I move data around ?
  - from CPUs to memory
  - from CPUs to disk
  - from CPUs on different machines
- How much data can I store ?

Number crunching on CPU: what do we count ?

- Rate of [million/billions of] floating point operations per second ([M|G]flops) FLOPs/S
- Theoretical peak performance:
  - determined by counting the number of floating-point additions and multiplications that can be completed during a period of time, usually the cycle time of the machine

FLOPS=clock_rate*Number_of_FP_operation*Number_of_cores

Sustained (peak) performance

- Real (sustained) performance: a measure

FLOPS= (total number of floating point operations done by a program) / (time the program takes to run in second)

- Number_of_floating_point_operations not easy to be defined for real application
- benchmarks are available for that..
- Top500 list uses HPL Linpack:
  - Sustained peak performance is what's matter in TOP500

Agenda

- ~~Prologue: why and where HPC ?~~ ✓
- ~~What is HPC ?~~ ✓
- ~~Performance and metrics~~ ✓
- Supercomputers and TOP500
- Parallel Computers

TOP 500 List

- The TOP500 list www.top500.org
- published twice a year from 1993
  - ISC conference in Europe (June)
  - Supercomputing conference in USA (November)
- List the most powerful computers in the world
- yardstick: Linpack benchmark (HPL)

HPL: some details

- From http://icl.cs.utk.edu/hpl/index.html:
  - The code solves a uniformly random system of linear equations and reports time and floating-point execution rate using a standard formulafor operation count.
  - Number_of_floating_point_operations = $2/3n^3 + 2n^2$ (n=size of the system)

```
================================================================================
T/V                N    NB     P     Q               Time                 Gflops
--------------------------------------------------------------------------------
WR03R2L2       86000  1024     2     1             191.06              2.219e+03
--------------------------------------------------------------------------------
||Ax-b||_oo/(eps*(||A||_oo*||x||_oo+||b||_oo)*N)=        0.0043644 ...... PASSED
```

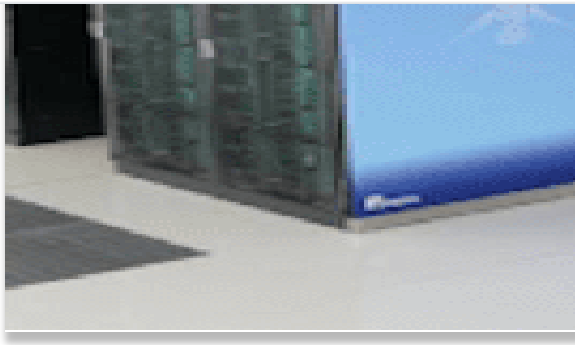
HPL&TOP 500 List

- For each machine the following numbers are reported using HPL:
  - Rmax: the performance in GFLOPS for the largest problem run on a machine.
  - Rpeak: the theoretical peak performance GFLOPS for the machine.
  - The measure of the power required to run the benchmark

And the winner is..

SUPERCOMPUTER FUGAKU - SUPERCOMPUTER FUGAKU, A64FX 48C 2.2GHZ, TOFU INTERCONNECT D

Site:	RIKEN Center for Computational Science
System URL:	https://www.r-ccs.riken.jp/en/fugaku/project
Manufacturer:	Fujitsu
Cores:	[illegible]
Memory:	
Processor:	

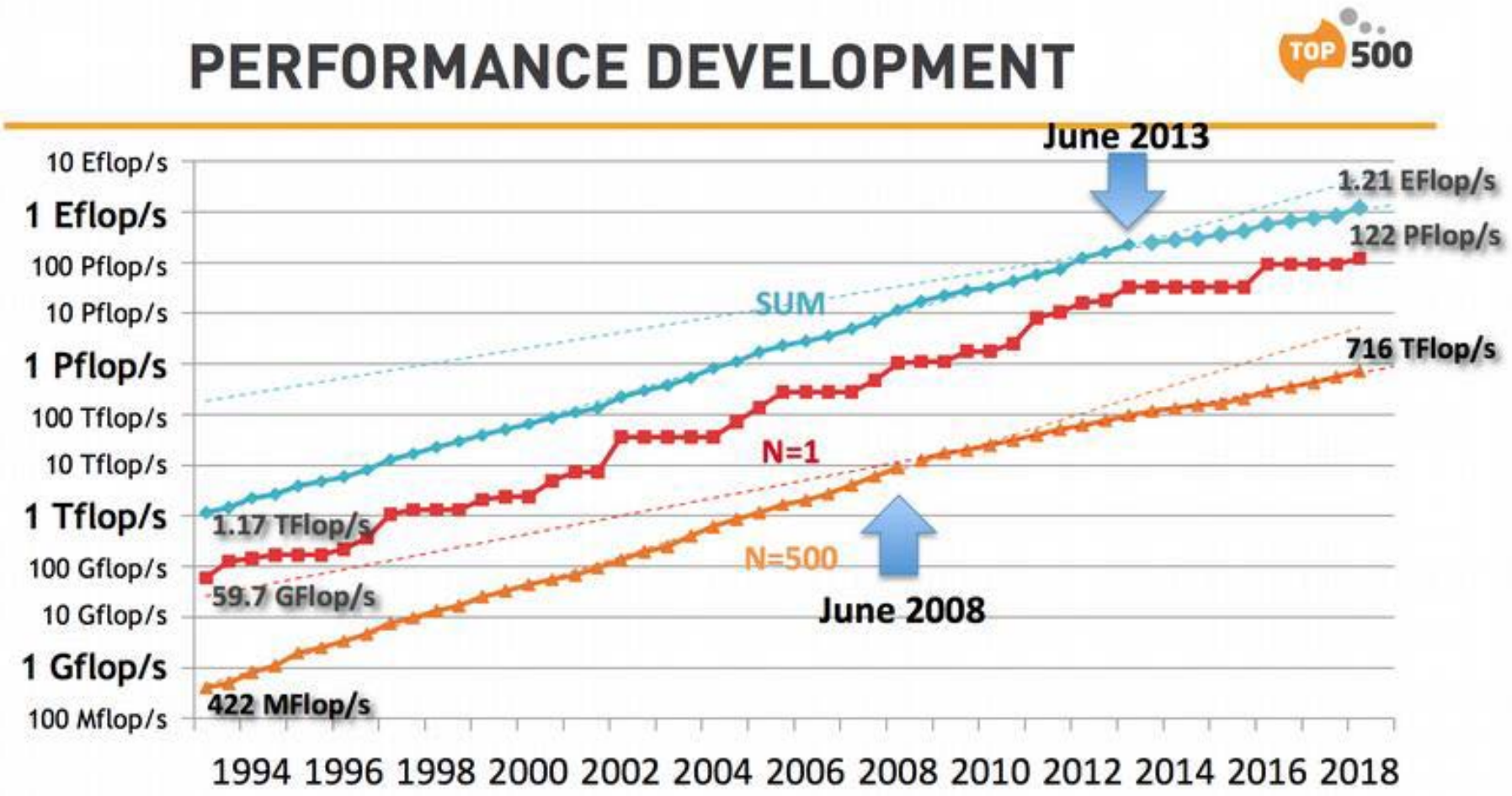


Performance	
Linpack Performance (Rmax)	415,530 TFlop/s
Theoretical Peak (Rpeak)	513,855 TFlop/s
Nmax	20,459,520
HPCG [TFlop/s]	13,366.4
Power Consumption	
Power:	28,334.50 kW (Submitted)

Highlights (from www.top500.org)

- After a few editions with very little new entries in the Top10 we have a new #1 and a total of 4 new systems:
- Supercomputer Fugaku, a system based on Fujitsu's custom ARM A64FX processor is the new #1. It is installed at the RIKEN Center for Computational Science (R-CCS) in Kobe, Japan, the location of the former K-Computer. It was developed by Fujitsu in close collaboration with Riken and uses Fujitsu's Tofu D interconnect to transfer data between nodes.
- It achieved a spectacular 416 Pflop/s on the HPL benchmark easily exceeding the old #1 Summit by 2.8x. In single or half precision (16-bit), which are often used in machine learning and AI applications, it's peak performance is actually above 1,000 PFlop/s (= 1 Exaflop/s) and because of this, it is often introduced as the first 'Exascale' supercomputer.
- HPC5 at #6 is the second new system. It is a PowerEdge system build by Dell installed by the Italian company Eni S.p.A. and is now the most powerful system in Europe. It achieves its performance of 35.5 Pflop/s by using NVIDIA Tesla V100 as accelerators and Mellanox HDR Infiniband as network.
- Selene at #7 is an NVIDIA DGX A100 SuperPOD installed in house at NVIDIA in the USA. Naturally it uses the new NVIDIA Ampere A100 for acceleration and a Mellanox HDR Infiniband as network as well to achieve it's 27.6 Pflop/s.
- Marconi-100 at #9 is the second Italian system in the Top10. It is an IBM Power System AC922 based on IBM POWER9 processors, Nvidia Volta V100 accelerators, and a Dual-rail Mellanox EDR Infiniband. The system is installed at the Italian research center CINECA and achieved 21.6 Pflop/s.

PERFORMANCE DEVELOPMENT



Agenda

- ~~Prologue: why and where HPC ?~~ 
- ~~What is HPC ?~~ 
- ~~Performance and metrics~~ 
- ~~Supercomputers and TOP500~~ 
- Parallel Computers

- To be continued