Smilei) Workshop

Super-computing landscape

ELI Beamlines – November 2023
Arnaud Beck

There are limitations to what personal computers can do

- 3D simulation
- 1000^3 cells
- 8 particles per cell
- 10^7 iterations
- 25 ns per particle per iteration on a single modern processor

There are limitations to what personal computers can do

- 3D simulation
- 1000^3 cells
- 8 particles per cell
- 10^6 iterations
- 25 ns per particle per iteration on a single modern processor

More than 6 years to make this simulation on a good desktop computer !

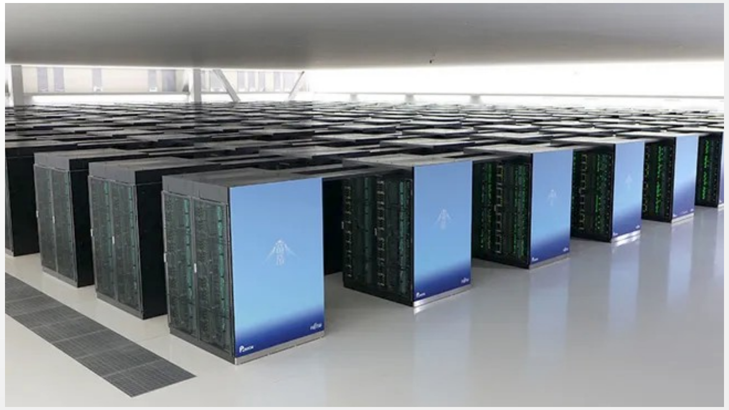
(Assuming you have a few TB of RAM and another few in disk space to store the results)

There are limitations to what personal computers can do

- 3D simulation
- 1000^3 cells
- 8 particles per cell
- 10^6 iterations
- 25 ns per particle per iteration on a single modern processor

More than 6 years to make this simulation on a good desktop computer !

(Assuming you have a few TB of RAM and another few in disk space to store the results)



If we really want to do this, we need to use a « Super » computer.

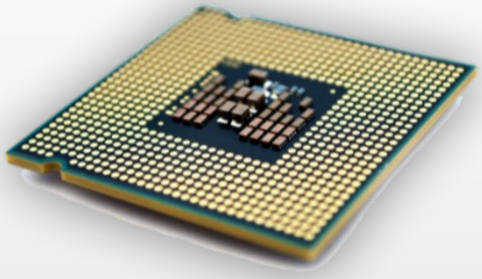
Basic elements of a computer

- The memory is a storage space for the data
- The processor is the central computing unit (CPU)
- The GPU (graphical processing units) assists the CPU for graphical tasks

Volatile memory (RAM) ⟷ Processor (CPU) ⟷ Accelerator (GPU)

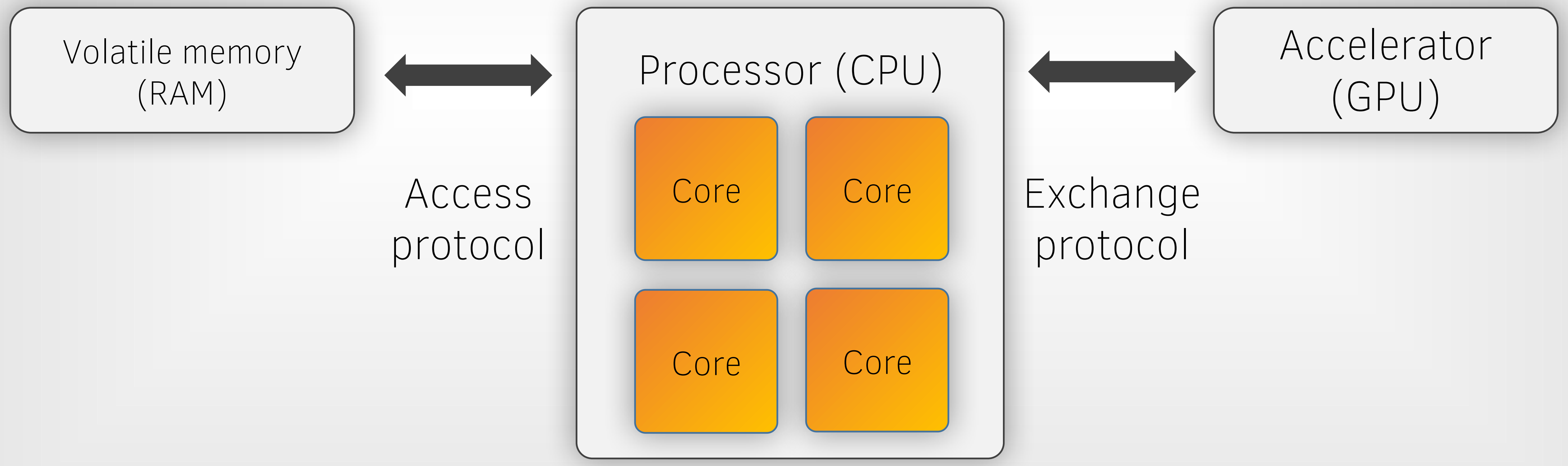
Access protocol

Exchange protocol



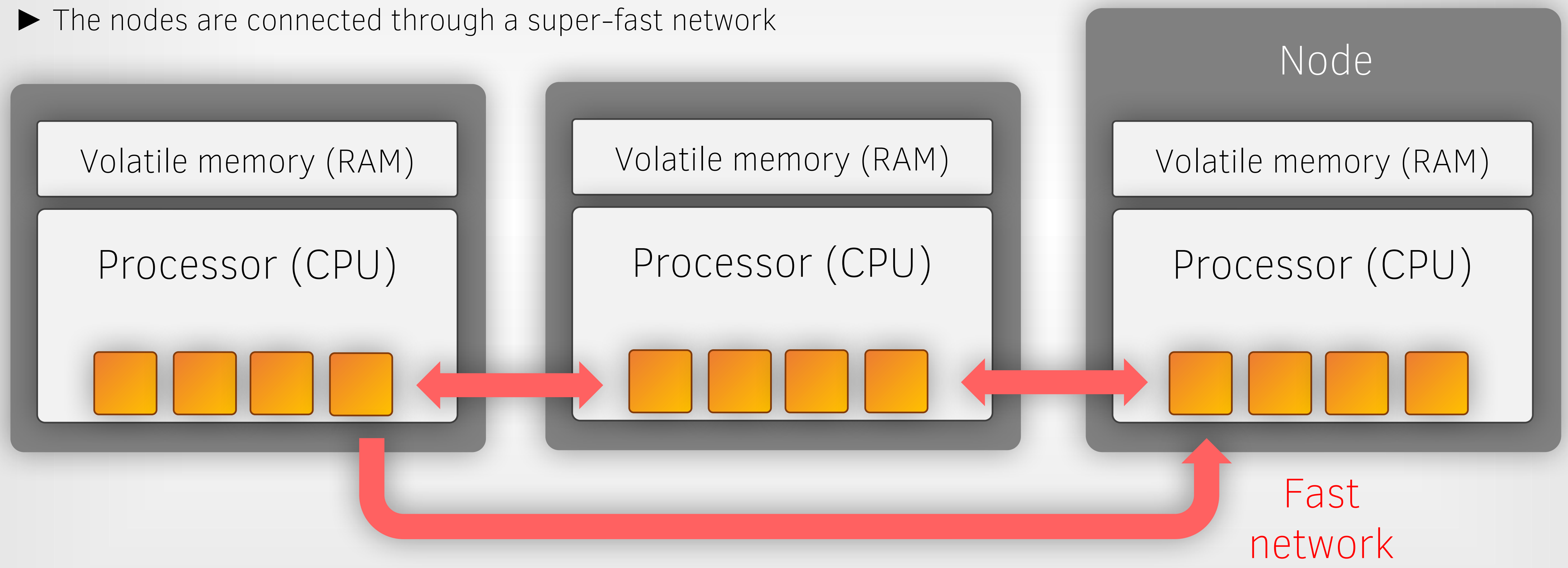
Basic elements of a computer

- A modern processor (CPU) is composed of several cores
- The most-recent CPUs have up to 128 cores



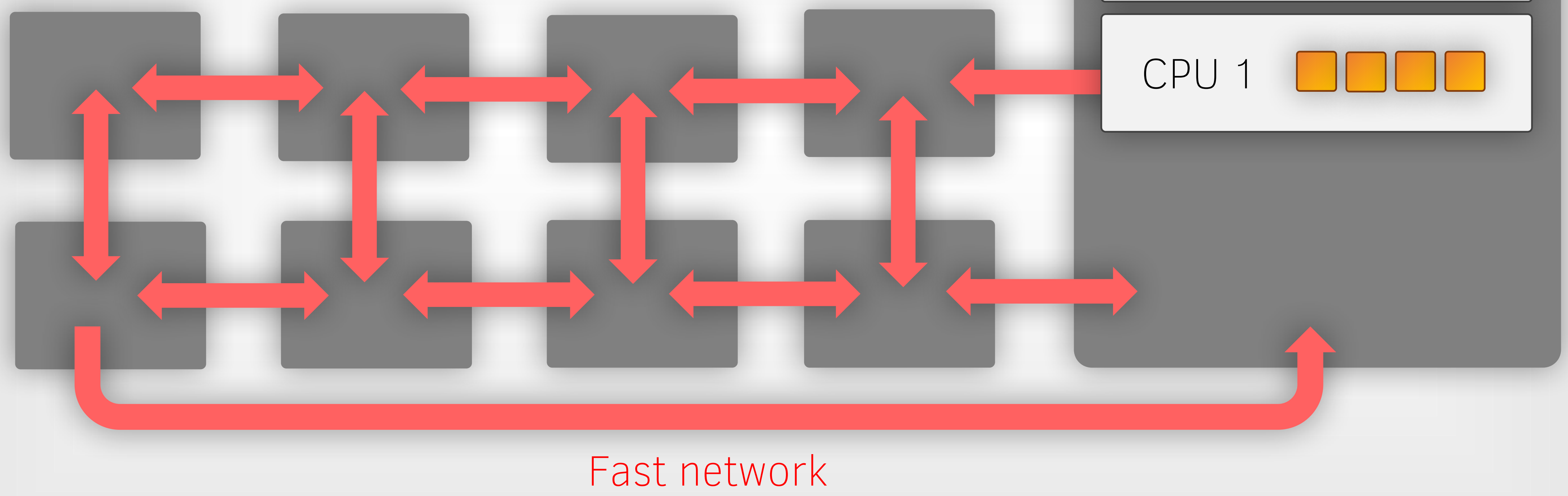
A super-computer is a cluster of many connected powerful computers

- Each computer is called a node
- The nodes are connected through a super-fast network



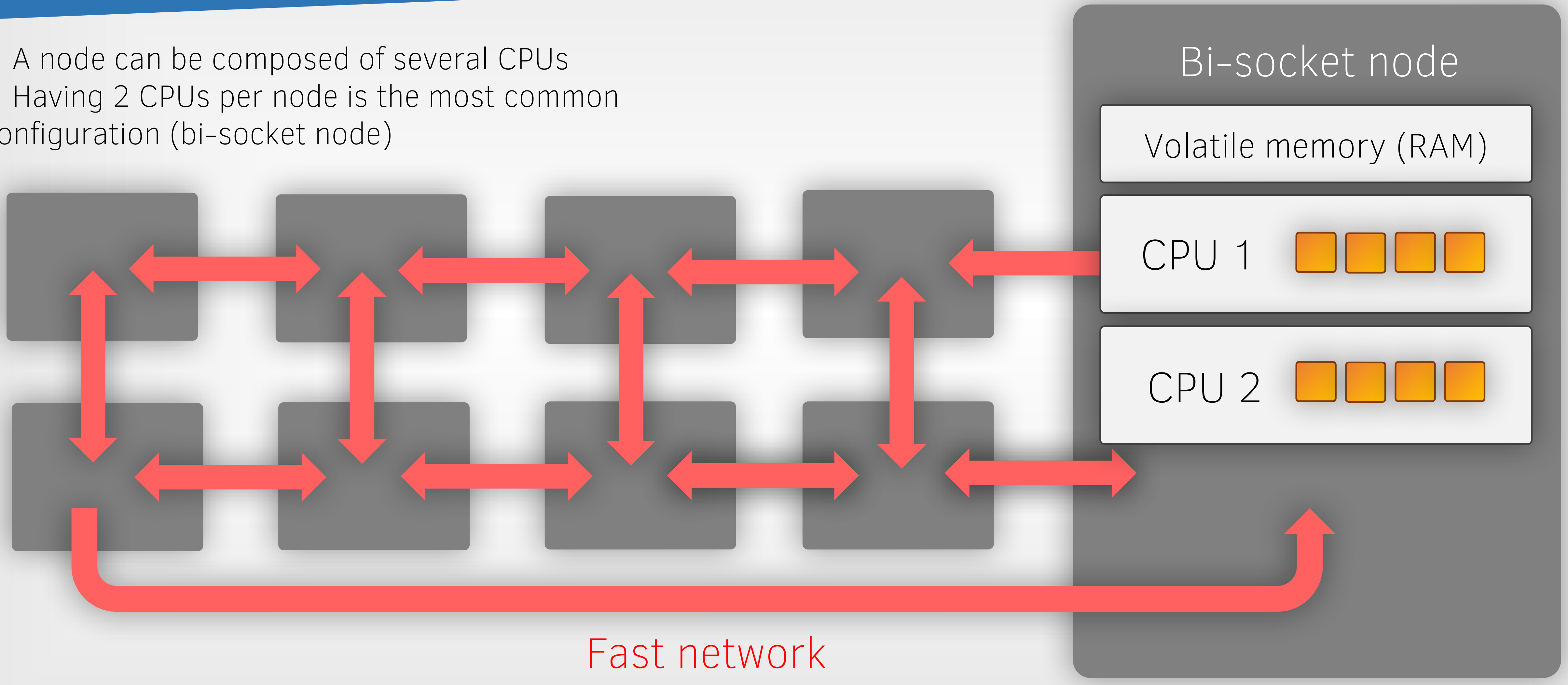
A super-computer usually has many nodes

- The largest super-computers can have more than 100 000 nodes



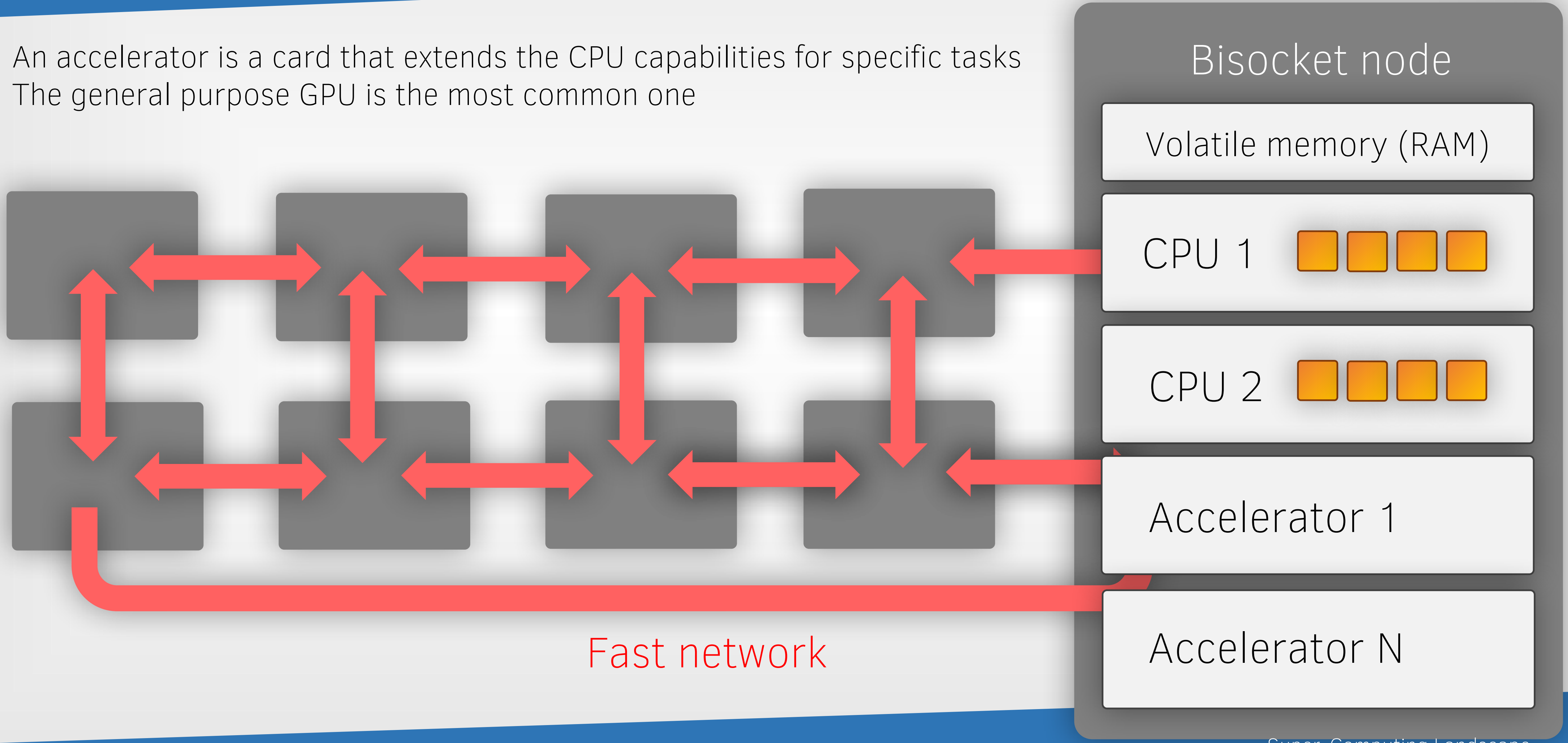
A node can have several CPUs

- A node can be composed of several CPUs
- Having 2 CPUs per node is the most common configuration (bi-socket node)



The node can be accelerated with so-called accelerators

- An accelerator is a card that extends the CPU capabilities for specific tasks
- The general purpose GPU is the most common one



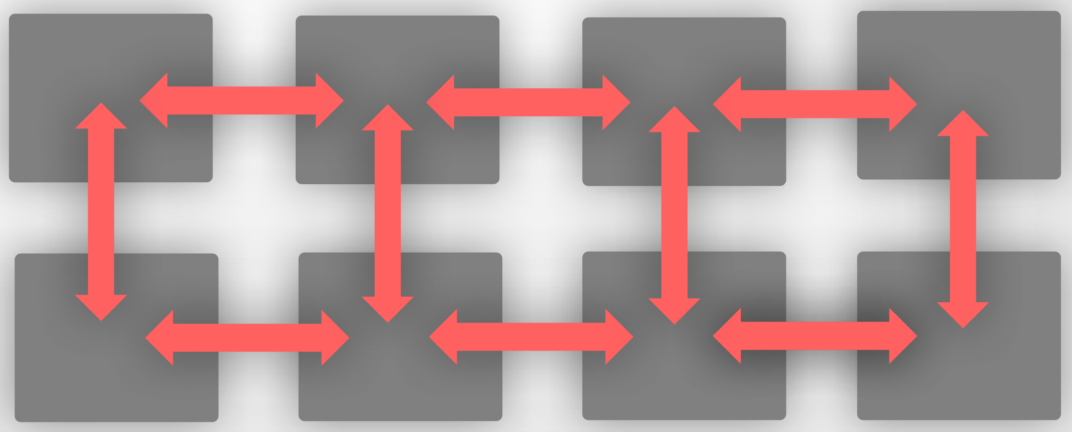
Vocabulary recap

- ▶ The total computational power is the sum of each **computing unit** (node/core/accelerator...)

- ▶ Need to use all the resources in **parallel** => **High Performance Computing** (HPC)

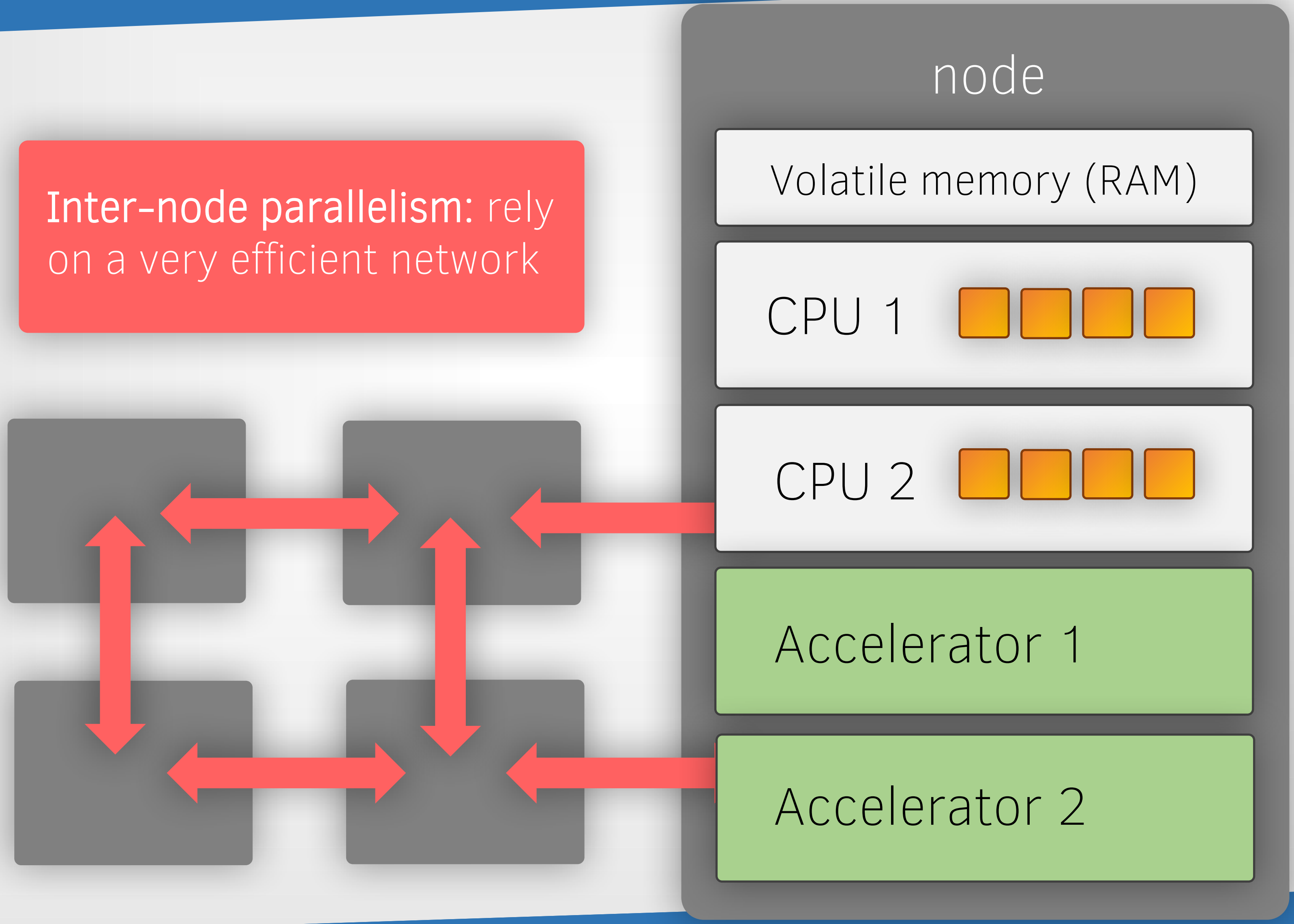
- ▶ This is the reason why we often call a super-computer :
  - A parallel machine
  - A cluster

- ▶ The computation must be **distributed** between the nodes/cores/accelerators



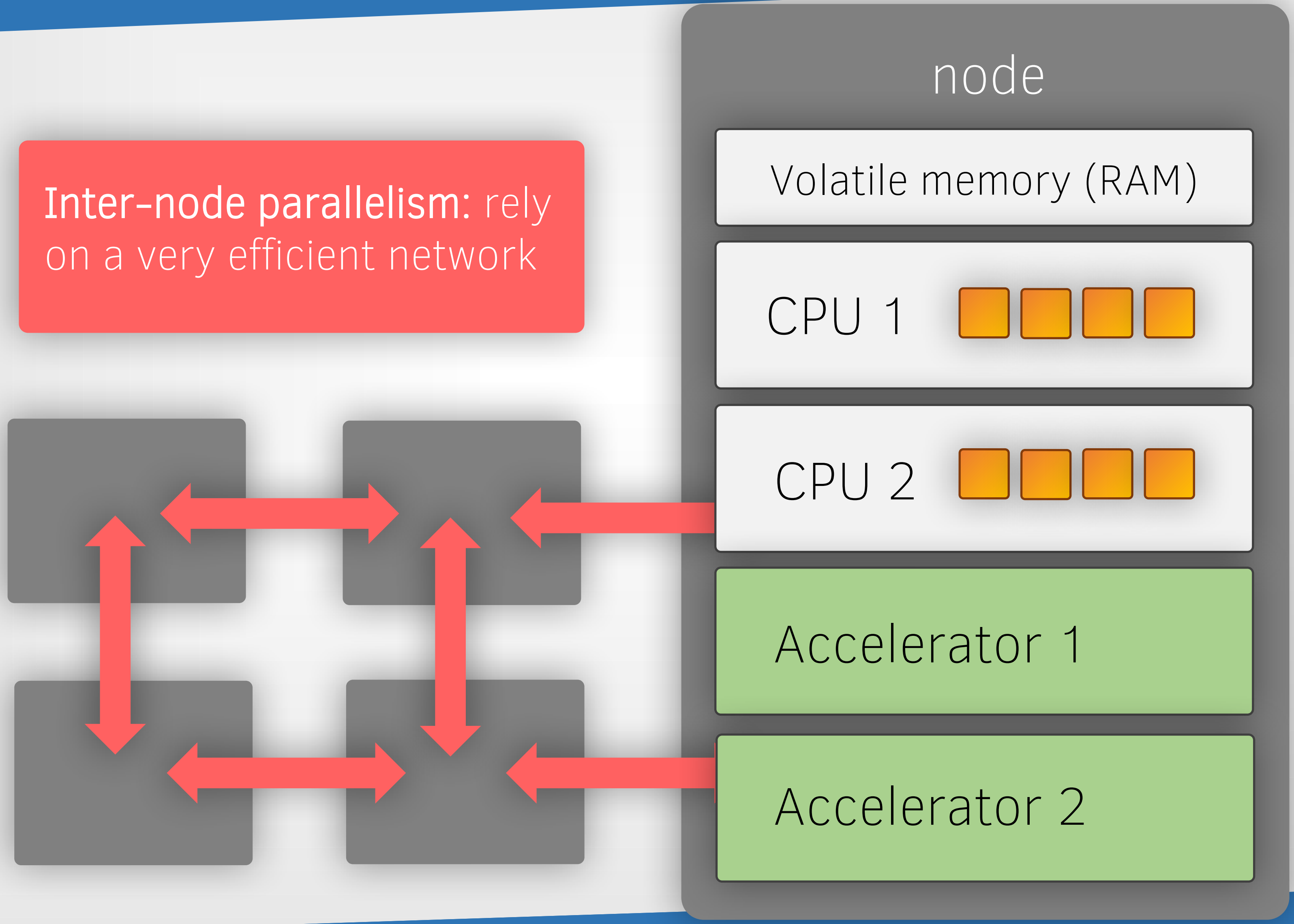
Different parallelism levels to handle

Inter-node parallelism: rely on a very efficient network



Different parallelism levels to handle

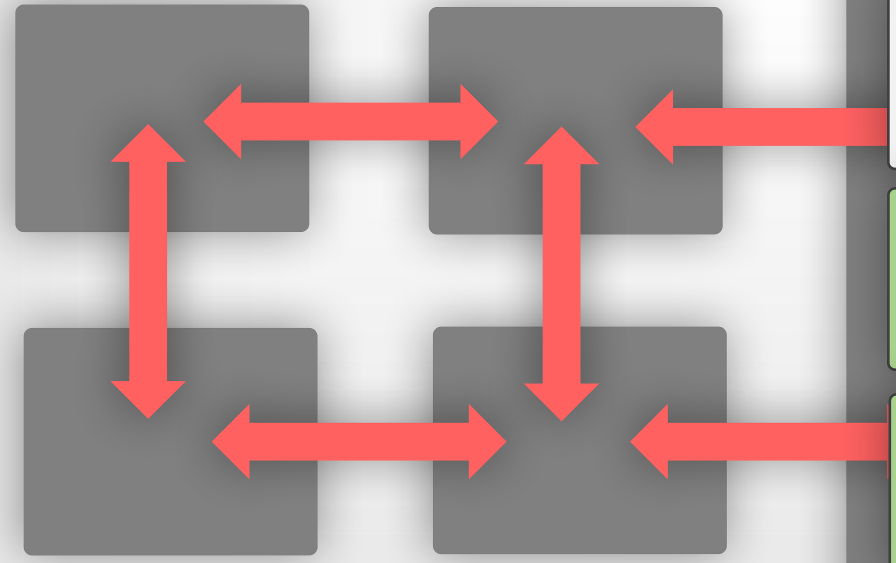
Inter-node parallelism: rely on a very efficient network

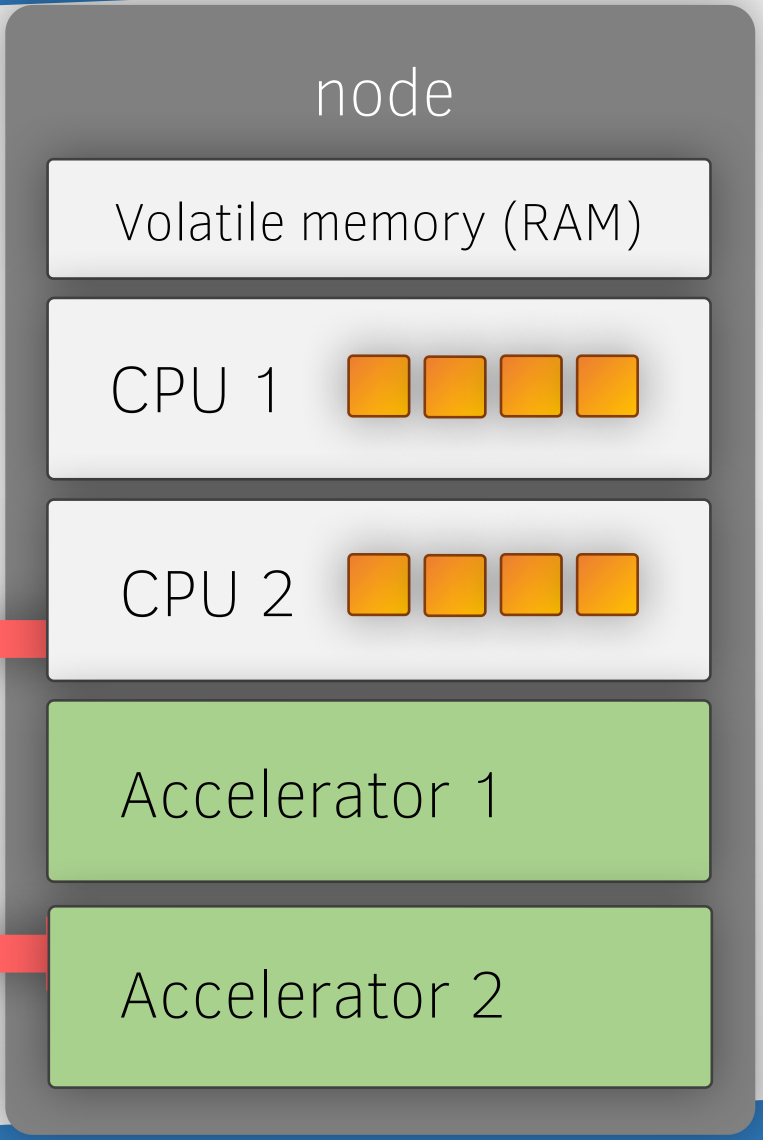


Intra-node parallelism: how to deal with all the cores efficiently

Different parallelism levels to handle

Inter-node parallelism: rely on a very efficient network

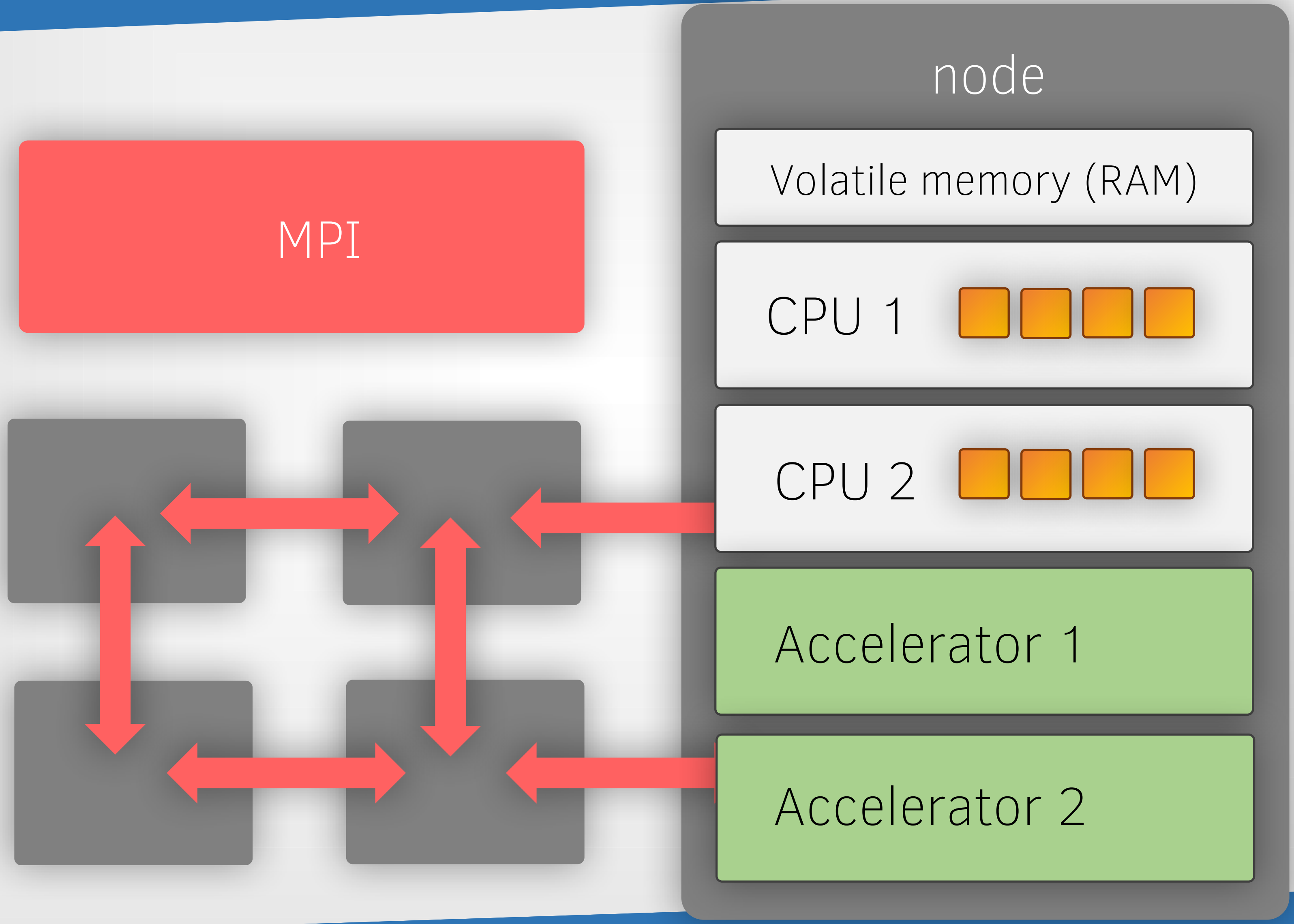




Intra-node parallelism: how to deal with all the cores efficiently

Heterogeneity: nodes with different types of computing units

Many software technologies adapted to each level



Smilers choices

OpenMP

OpenMP, OpenACC, CUDA, HIP

Programming challenges for HPC applications

- Developing efficiently for a super-computer is more difficult than for a simple desktop computer:
  - Communication and synchronization through the network between the nodes
  - Load balancing between the nodes
  - Work share within the node (between cores and/or accelerators)
  - Node heterogeneity (CPU/GPU)
  - Memory usage
  - Architecture-specific optimizations (Memory affinity/hierarchy, Vectorization...)
  - Etc
- Typical HPC applications use only a fraction of the total theoretical peak computational power.
- Efficiency on a given hardware also strongly depends on the type of algorithm and the physical case.

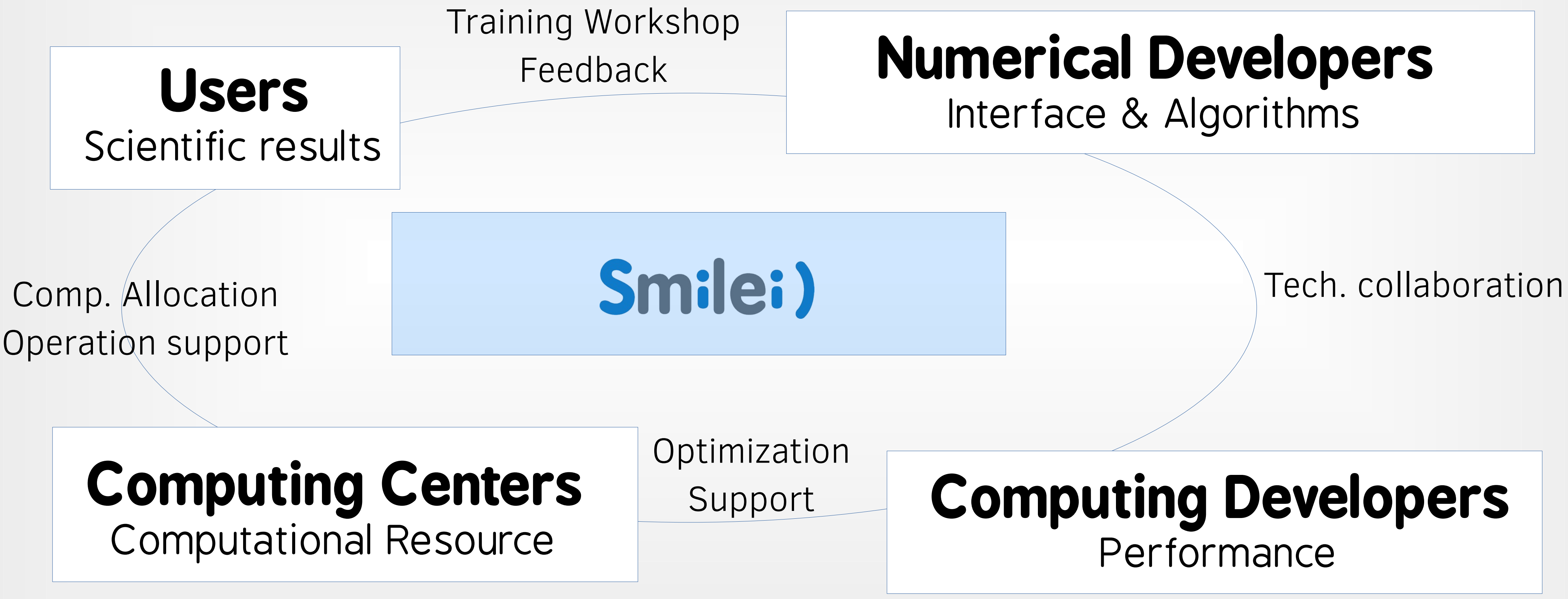
Physicists and computer scientists are needed

- Computer architectures and programming models complexity makes efficient code programming more difficult than before: need for advanced knowledge.

- Collaboration between scientists and HPC research-engineers is extremely valuable and even necessary for the largest frameworks.

- This is the case of Smilei)

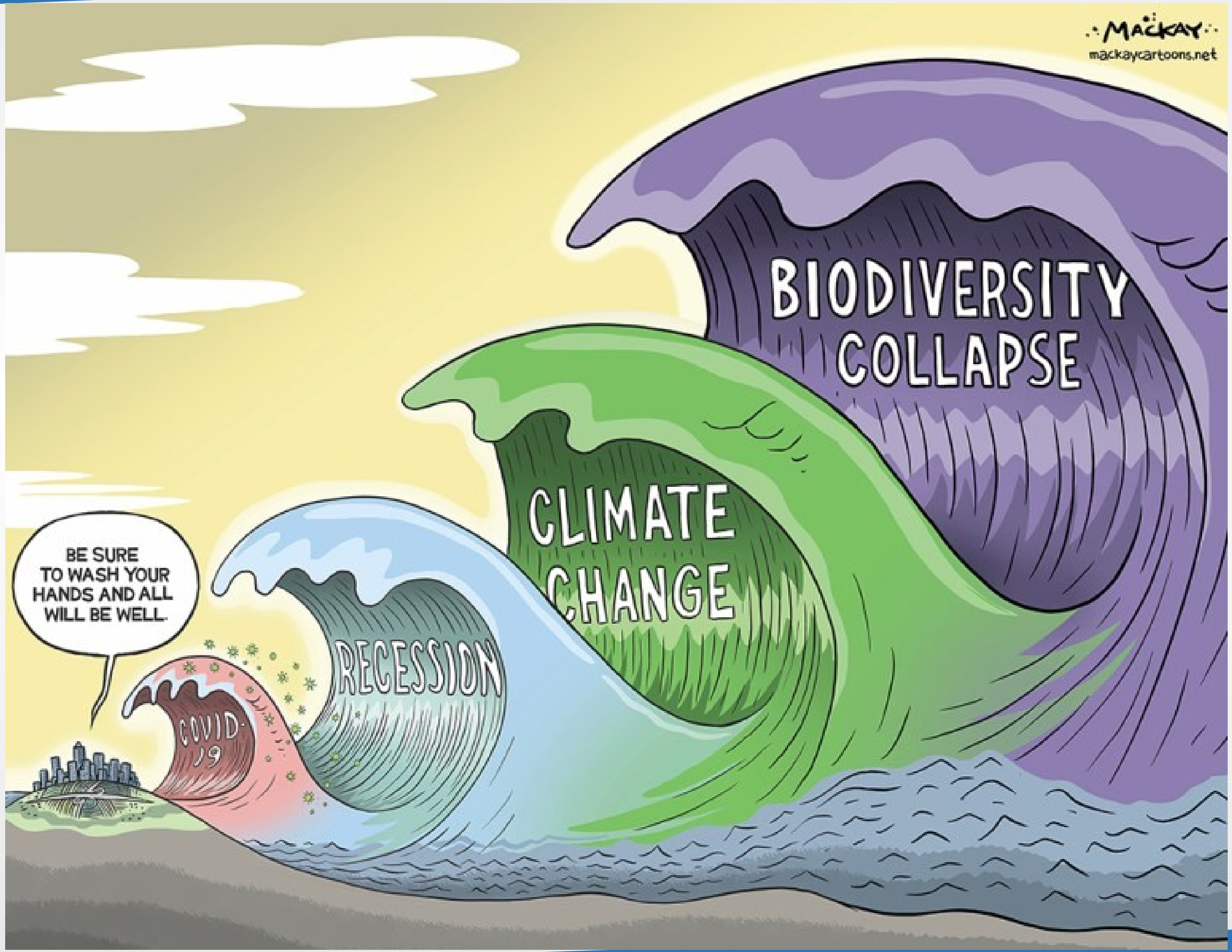
The Ecosystem of an HPC application



What are the limiting factors of super-computers ?

- Energy consumption => environmental impacts and financial cost
- Memory capability
- Network performance
- Core performance
- File system performance
- Building and maintenance cost
- And more

Environmental limitations



Computing environmental and financial limitations

- ▶ A super-computer has significant environmental impacts: greenhouse gas, pollution, human rights, water, etc.

Cobalt mine in Congo (2016)

Toxic lake in China

- ▶ Powering a super computer is not cheap.
  - Today an exascale 21MW system power bill is ~ 30 Million €/year.
  - Electricity price is very volatile.

- ▶ Today it already happens that computing centers are asked to shut down for periods of time for energy savings.

Memory capability is another limitation to the node-level parallelism

- Computational performance has increased faster than the memory capability in both **size** and **access speed**
- Available memory per core has even slightly decreased
- Many algorithm implementations limited by the memory bandwidth and/or the memory size and not the computation power (memory bound algorithms)

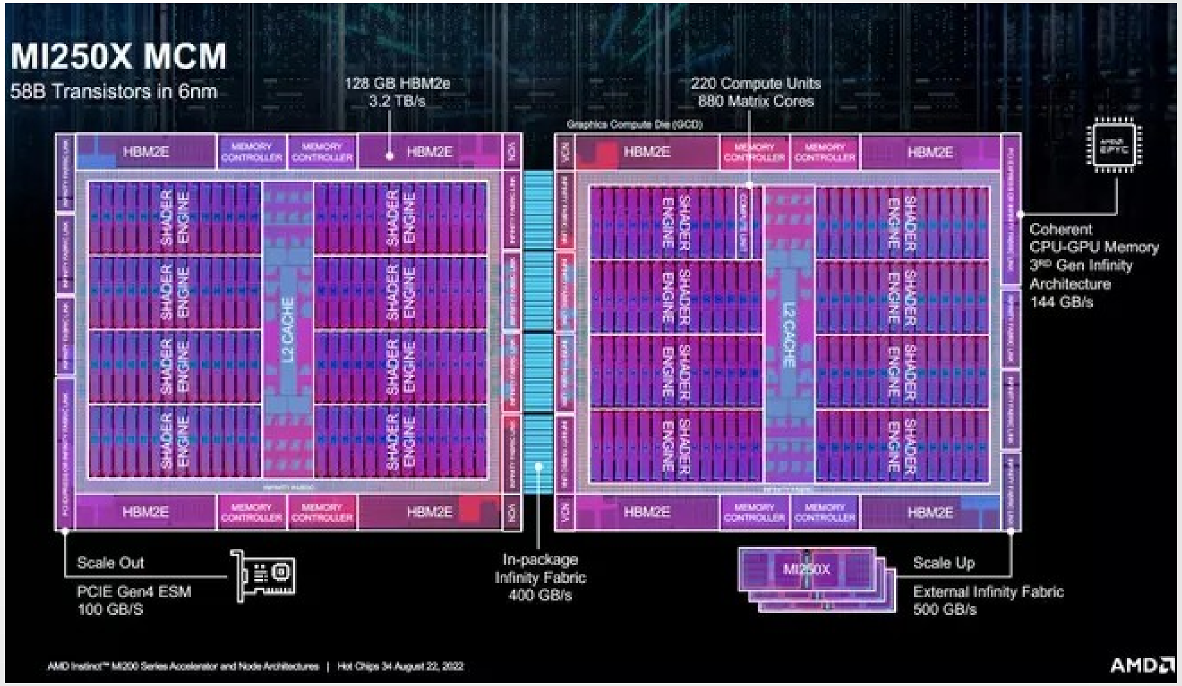
Network performance is another strong issue to unlimited parallelism

- Network technologies have also evolved less rapidly than the computational power
- Nodes have to communicate and synchronize through the network
- More cores and more nodes lead to more network usage and pressure

- Same problem as the road interconnection network

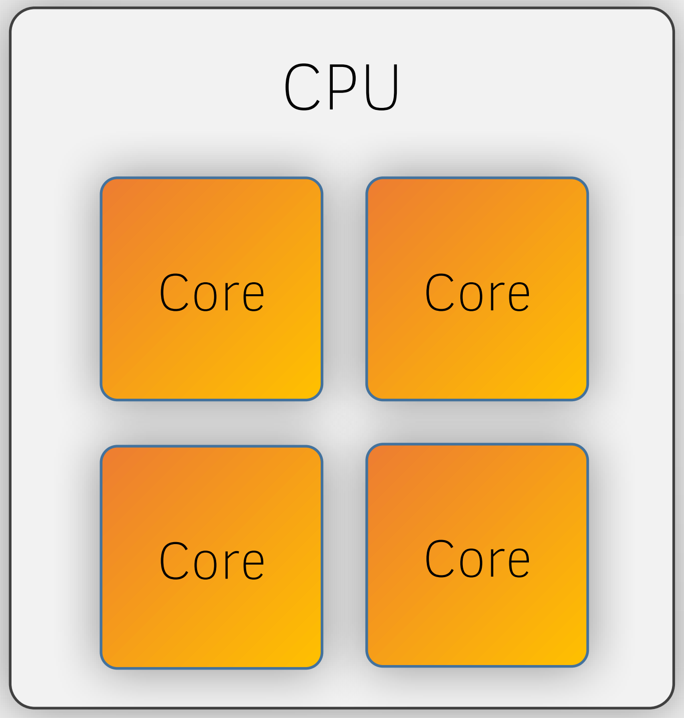
Trying to overcome the limitations with GPUs

- Hardware dedicated to computing to make it more energy efficient.
- More compact computing units for less network communication.
- High memory bandwidth.

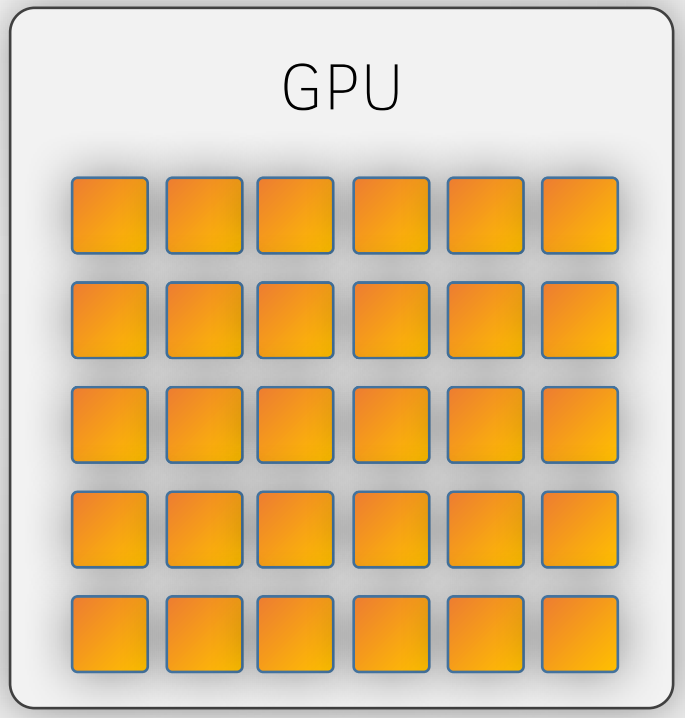


GPU basic description

Both share the same DNA but:

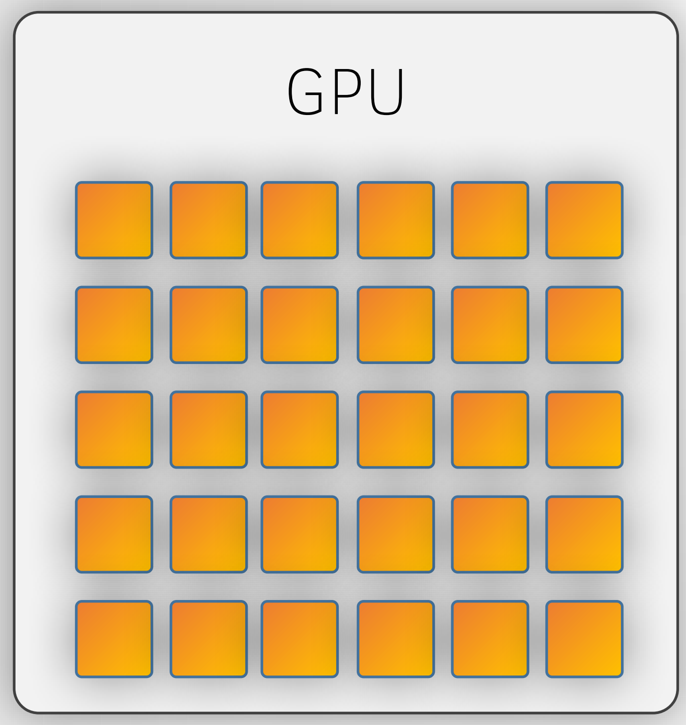


CPU have few complex cores



GPU have a lot of simple cores

GPU main characteristics



- GPU architecture is efficient for large and simple problems (solving linear algebra problems for instance) contrary to the CPU that is more versatile.
- More difficult to program than CPUs.
- Less libraries and applications.
- They are an accelerator and need a CPU for system and IO tasks.

Today's landscape: The TOP500 list

- A global ranking of the 500 most powerful super-computers
- Databases available for statistics
- Updated in November (SC event) and June (ISC event) every year
- https://www.top500.org/

How the performance is measured

- A performance metric is the number of operations per second called “flops”.
- TOP500 ranking is based on the LINPACK benchmark that consists on solving a dense system of linear equations.
- The LINPACK measured performance is always less than the theoretical peak performance.
- The LINPACK measured performance is always much higher than the performance measured with “real” applications.

TOP 500 The List.

A few powerful HPC systems today

Summit (rank 5)

- USA
- IBM CPUs + NVIDIA V100 GPUs
- 149 Pflops
- 10 MW

Fugaku SC (rank 2)

- Japan
- ARM CPUs
- 442 Pflops
- 29,8 MW

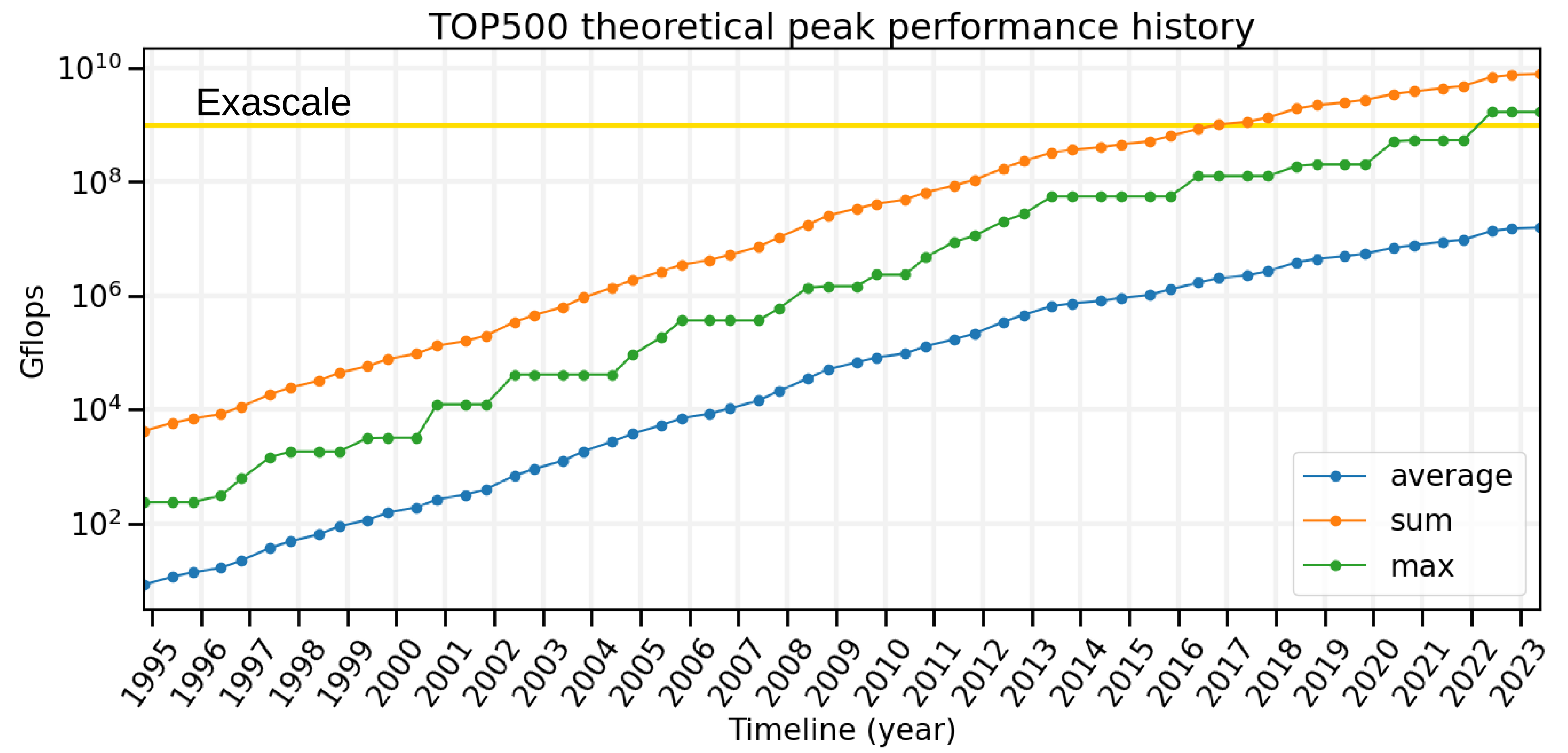
Frontier (rank 1)

- USA
- AMD CPU+GPU
- 1194 Pflops
- 22,7 MW

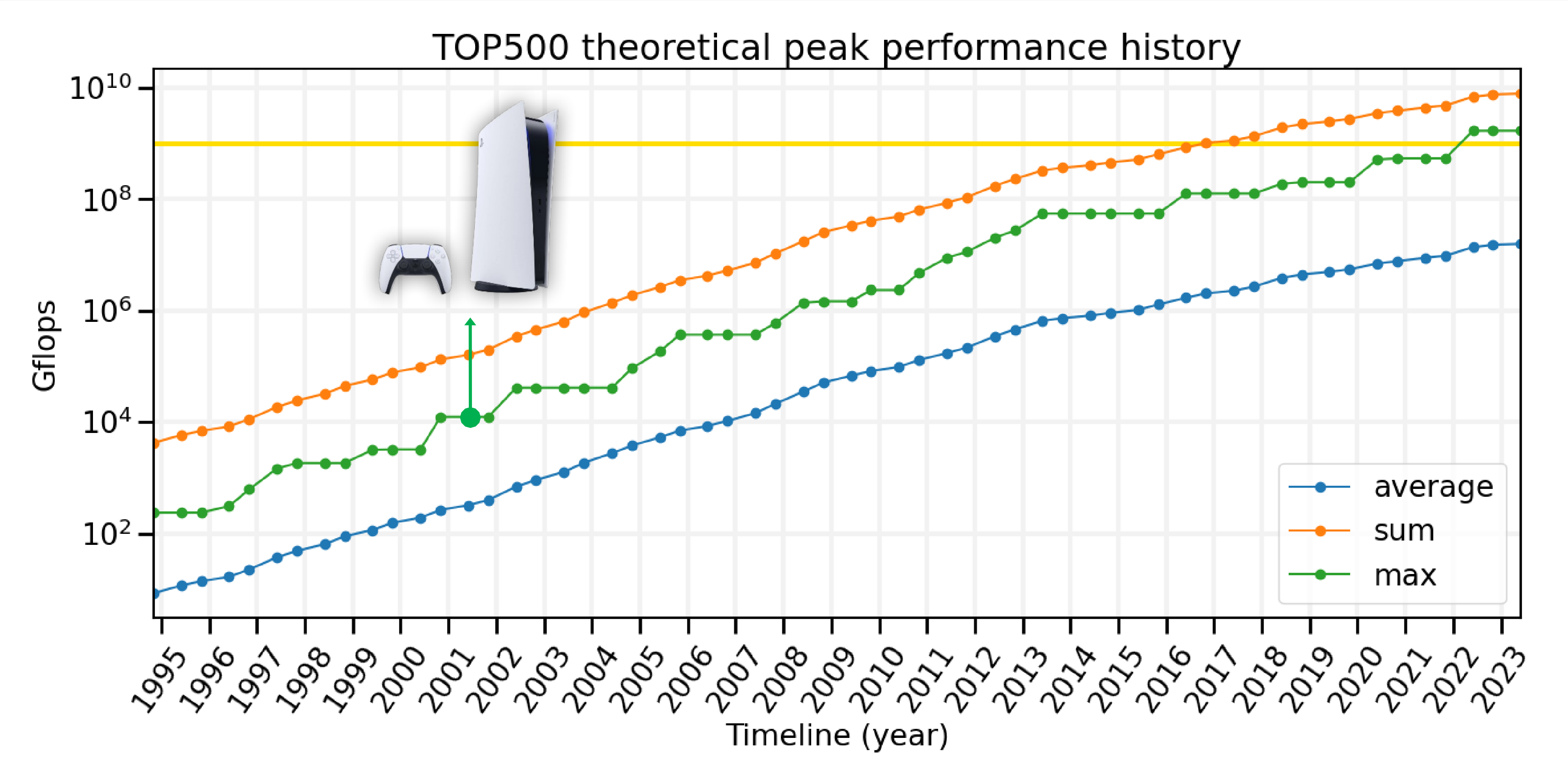
Adastra (rank 12)

- France
- AMD CPUs + GPU
- 46 Pflops
- 0,9 MW

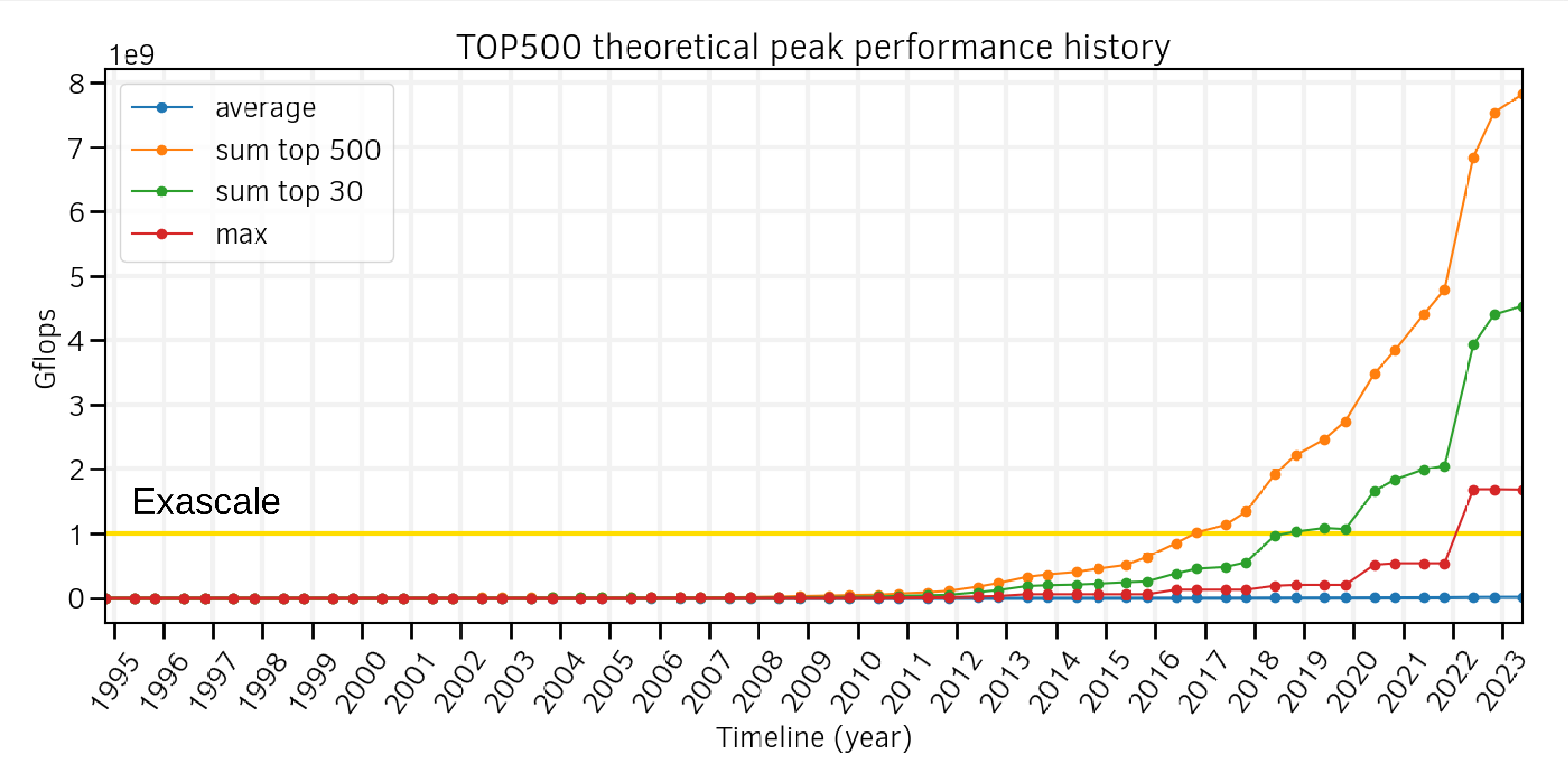
Evolution of computing power



Evolution of computing power

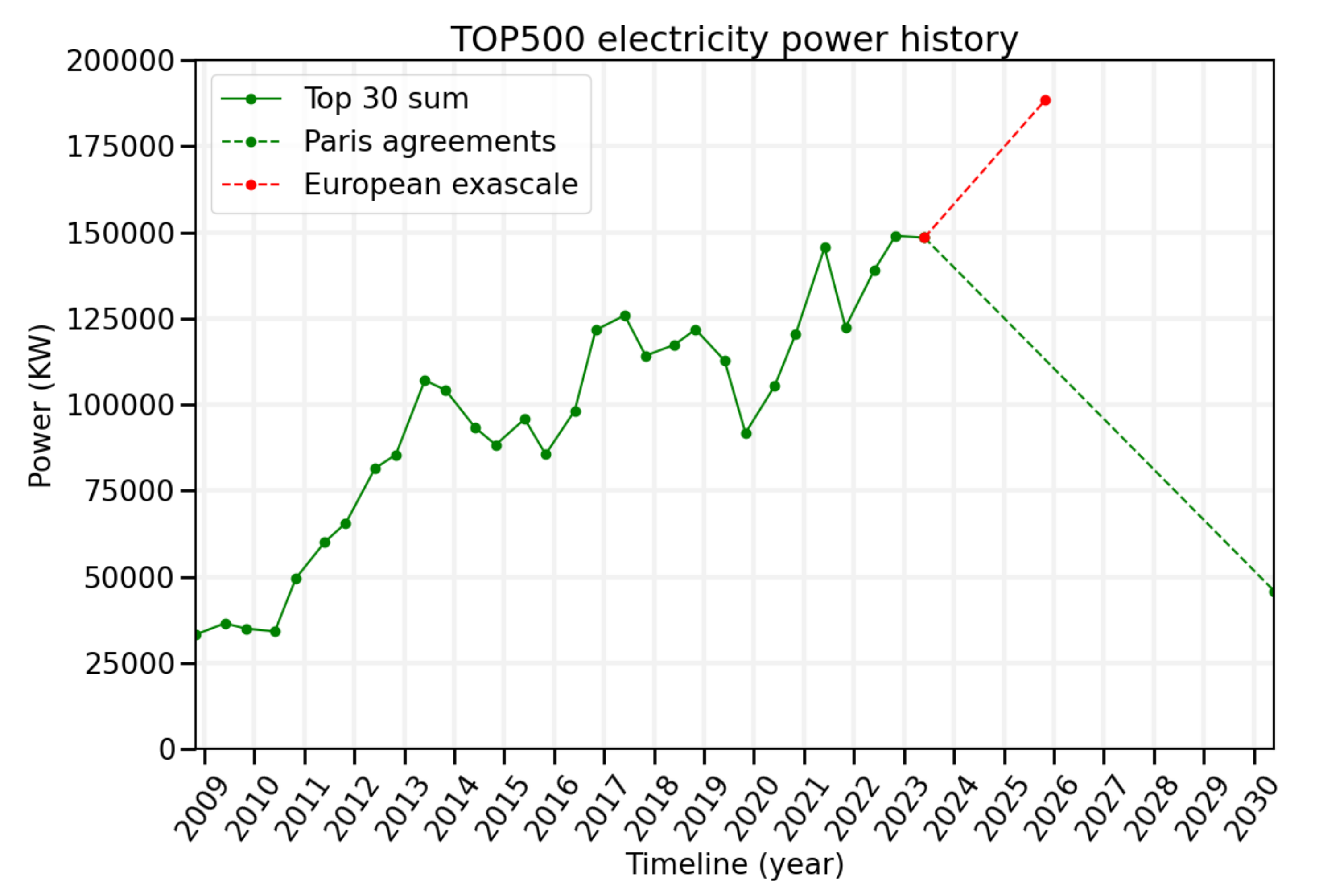


Evolution of computing power



Focus on the TOP500 electricity consumption

- Power increase is slower than performance.
- Important gain of efficiency.
- Europe has announced 2 exascale systems before 2026.
- This does not account for the increase of the number of HPC systems.



Technological focus on energy efficiency

Energy efficiency is a major driver of technological development

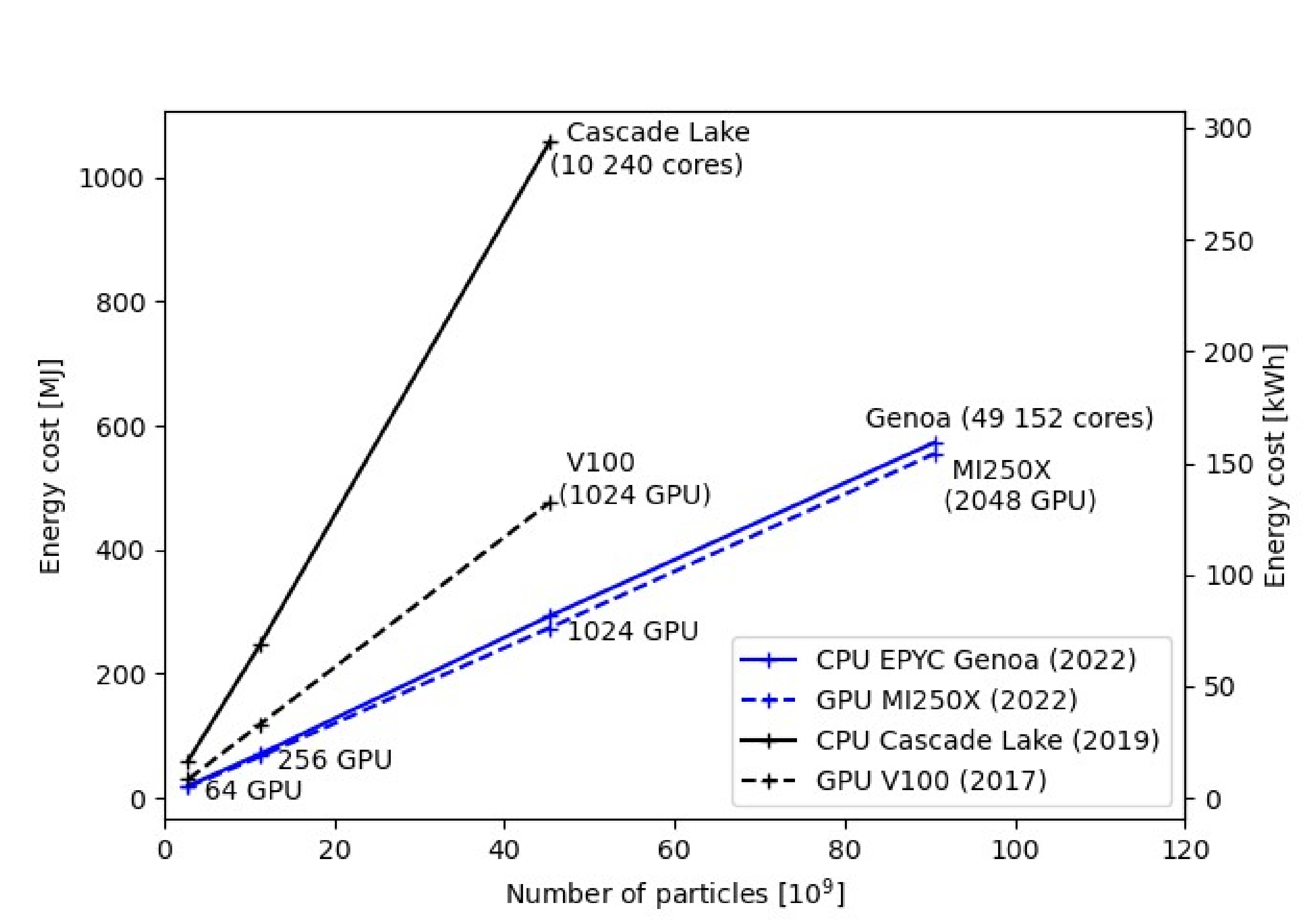
With 10 year old technologies, an exascale system would need a power of 1 GW

With 5 year old technologies, an exascale system would need a power of 120 MW

At present technologies, an exascale system consumes 20 MW

(And yet energy consumption keeps increasing !!)

Energy: the proper metric for software performance too



► Weak scaling: the resources scale with the problem size.

► The configuration is optimized for each system.

► Results may differ with another physical case.

► The energy cost depends linearly on the size.

► Be aware of the “Rebound effect”.

Software aspects in the coming related presentations

A lot of efforts is put into maintaining the code efficiency on new architectures and to provide scientific results with less computation.

How we focus on these challenges will be presented in session 2 and 3:

- ▶ Parallelization – Francesco Massimo (how to distribute the work)
- ▶ Achieving performance – Charles Prouveur (how to compute efficiently)
- ▶ Advanced techniques – Guillaume Bouchard (how to compute less)

Thanks & Keep Smileing!)

Thanks for supporting this event

eli beamlines

PLAS@PAR Plasmas à Paris

GdR APPEL Groupement de Recherche Accélérateurs Plasma PompÉs par Lasers

Contributing labs, institutions & funding agencies

MAISON DE LA SIMULATION

PLAS@PAR Plasmas à Paris

P2IO Physique des 2 Infinis et des Origines

PALM Laboratoire d'Excellence Physique : Atomes Lumière Matière

LPP

irap astrophysique & planétologie

iris

intel

CELIA

GdR APPEL Groupement de Recherche Accélérateurs Plasma PompÉs par Lasers

Physique des Infinis Une Initiative Sorbonne Université

AIDAS AI Data Analytics and Scalable Simulations

cnrs

cea

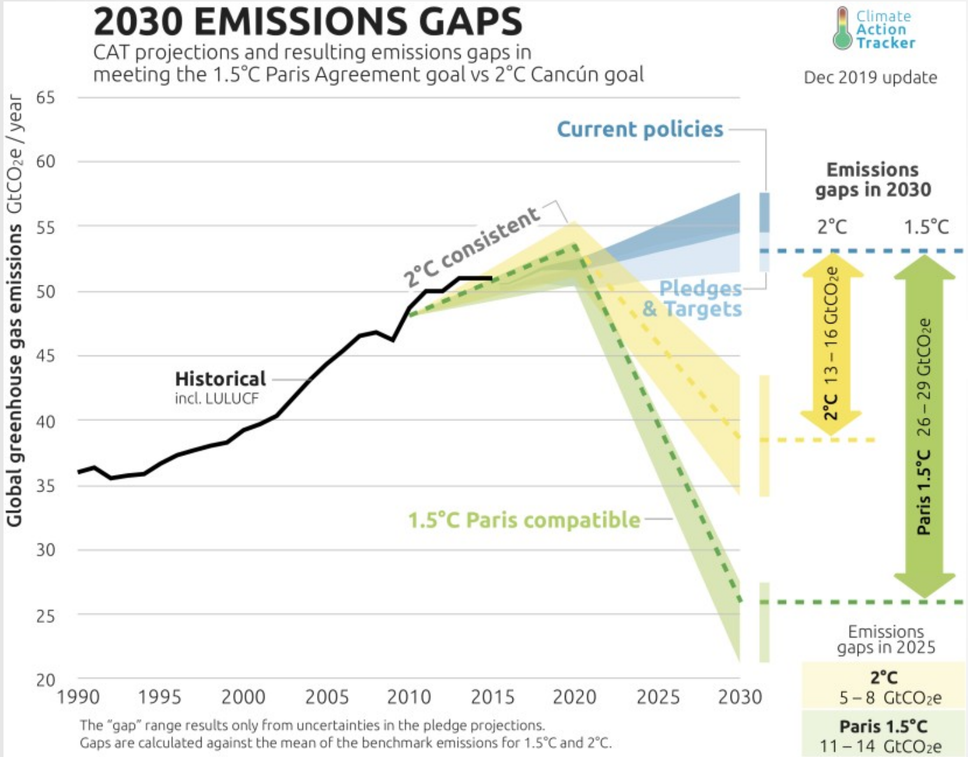
SORBONNE UNIVERSITÉ

université PARIS-SACLAY

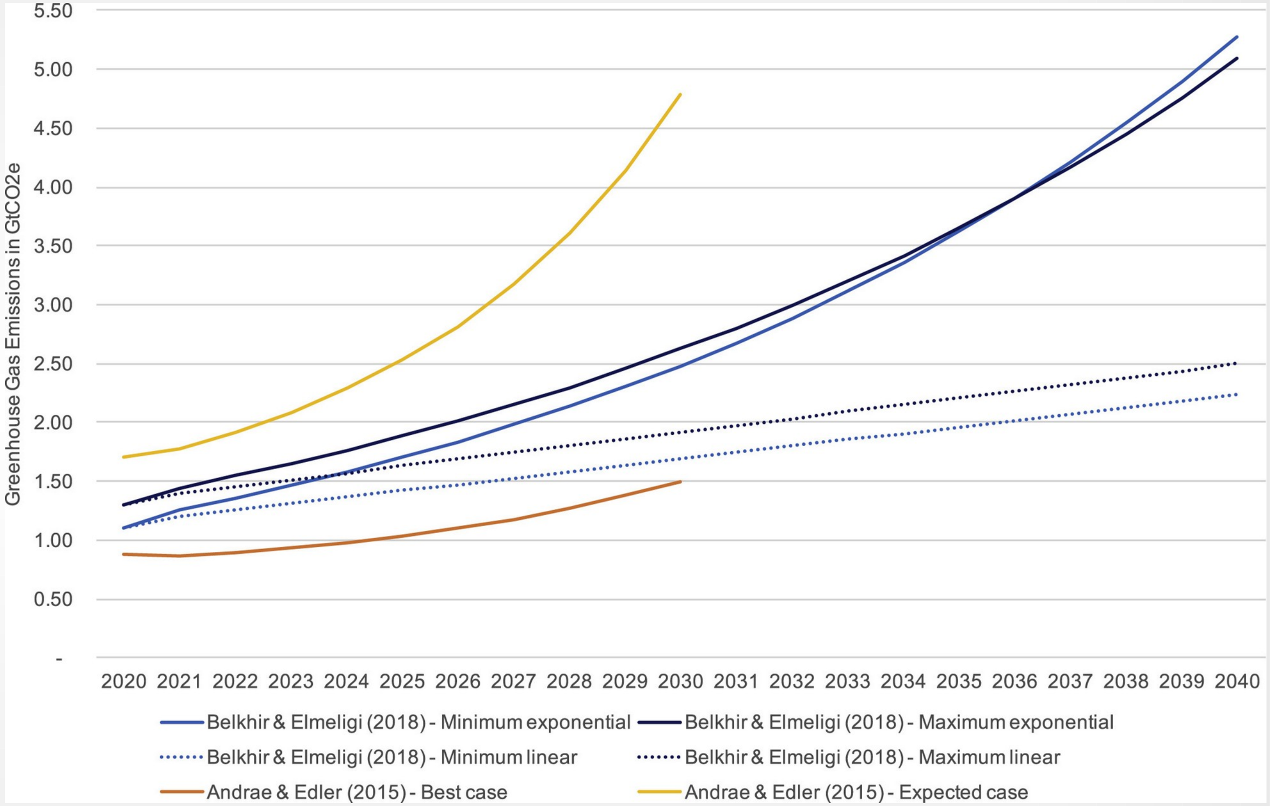
l'X ÉCOLE POLYTECHNIQUE

INSTITUT POLYTECHNIQUE DE PARIS

Environmental limits



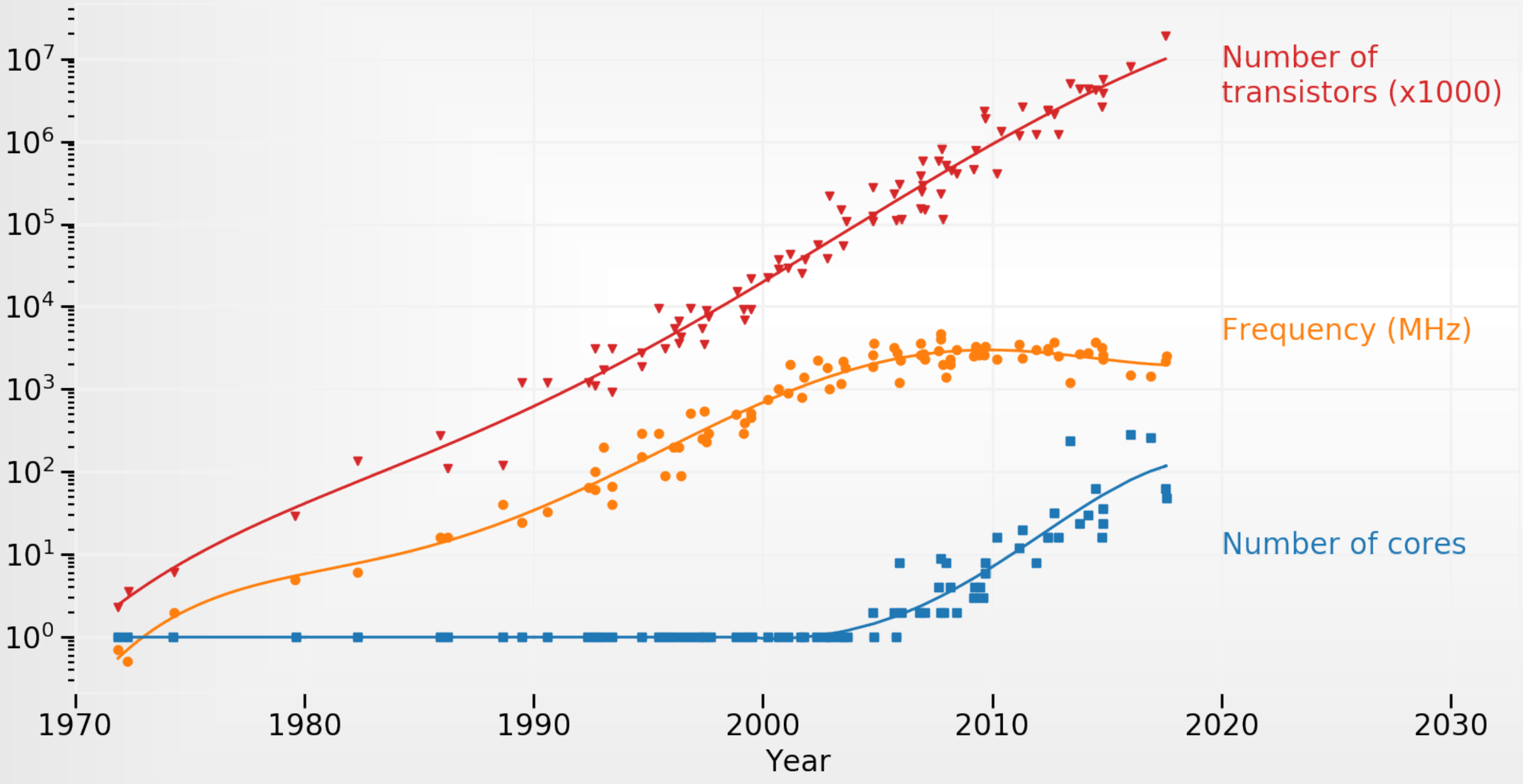
Paris agreement in 2015 (196 countries)



Freitag et al, 2021, «The climate impact of ICT: A review of estimates, trends and regulations» (open access).

ICT is the infrastructure and components that enable modern computing.

CPU micro-architecture trend



Computational power increase:

- Decrease of the manufacturing process
- Increase of the number of transistors per socket
- Increase of the number of cores per socket
- Similar frequency / larger vector size
- Share parallelism more and more important