EURO²

HPC Teaching Kit

By EuroCC Belgium

What is HPC?

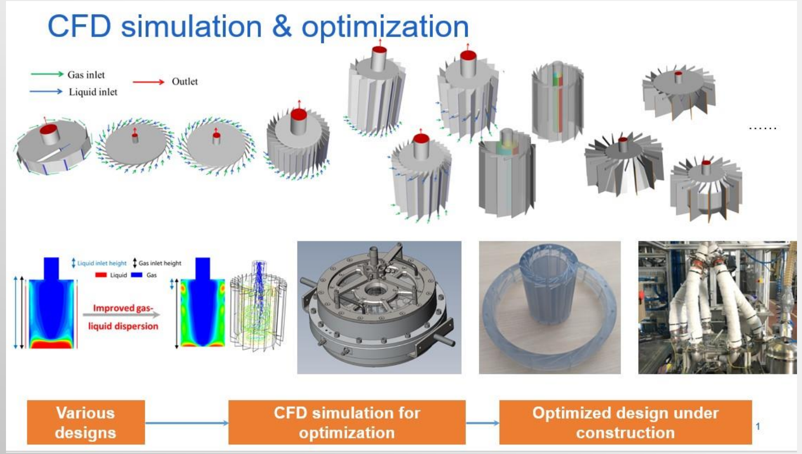
- Aggregation of computer power (in the form of supercomputer or *clusters*) to deliver high computational performances, in order to solve large problems.
- *Clusters* are linked computers (= nodes) that works together, so that, for many aspects, they look like one single computer (*more on that latter*).
- The work is distributed across workers (*more on that latter*).

Exemples of HPC uses

Have a look at www.enccb.be/stories

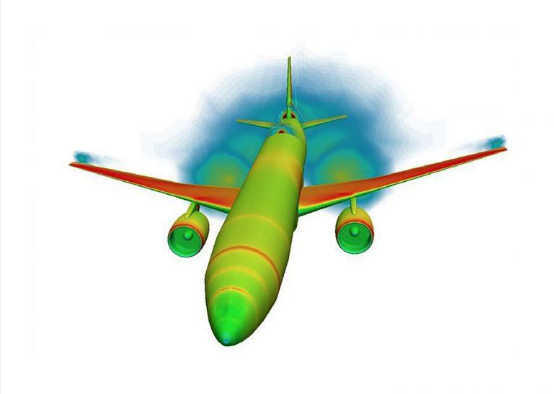
Examples: fluid dynamics

Design and optimization of a vortex unit for CO_2 capture



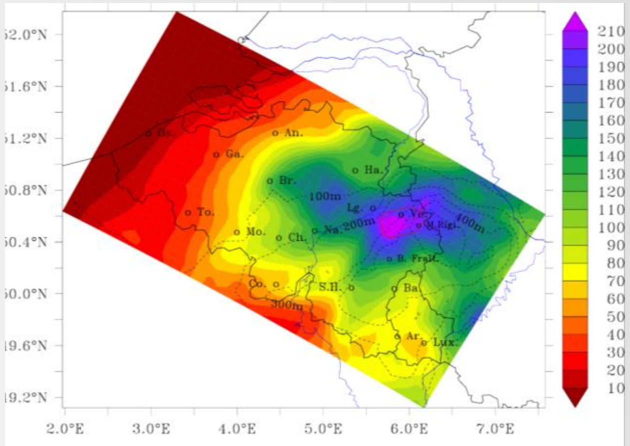
https://www.enccb.be/usvortexunit

Aircraft design



https://prace-ri.eu/automating-aircraft-design-and-optimisation/

Predicting the climate

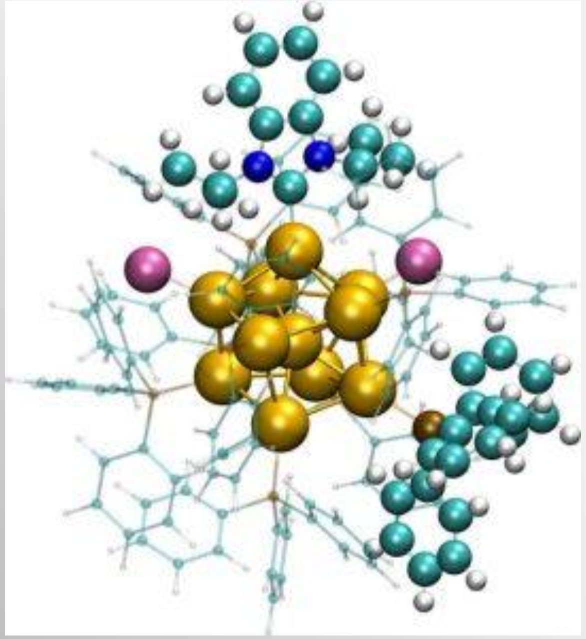


Total precipitation over 15-17 August 2033

https://www.enccb.be/usxavierfettweis

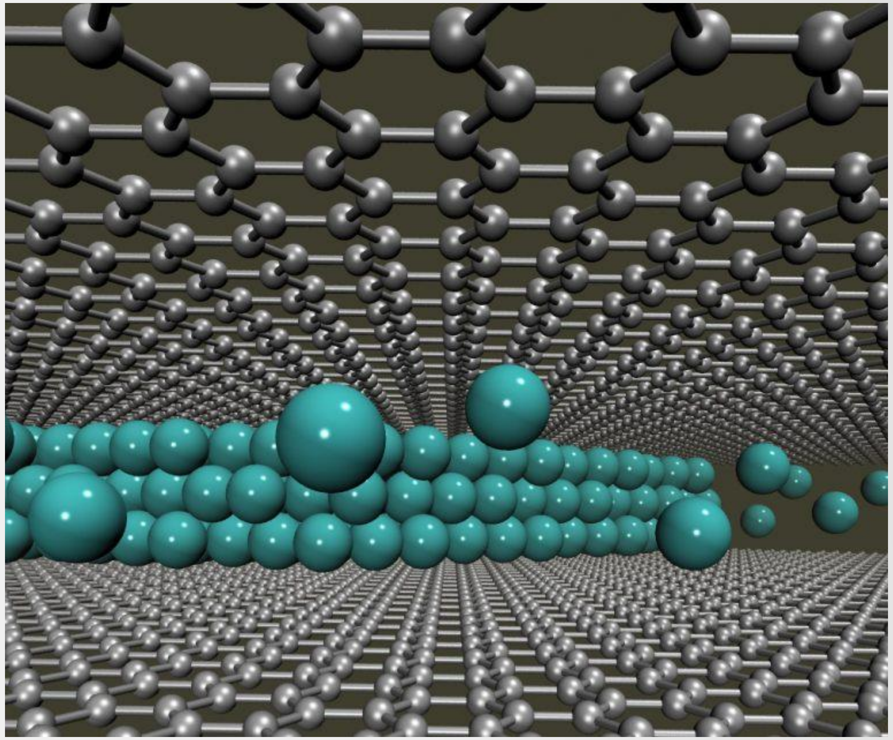
Examples: materials

Predicting the structure of gold cluster that chops carbon dioxide



https://prace-ri.eu/computer-simulations-gold-cluster/

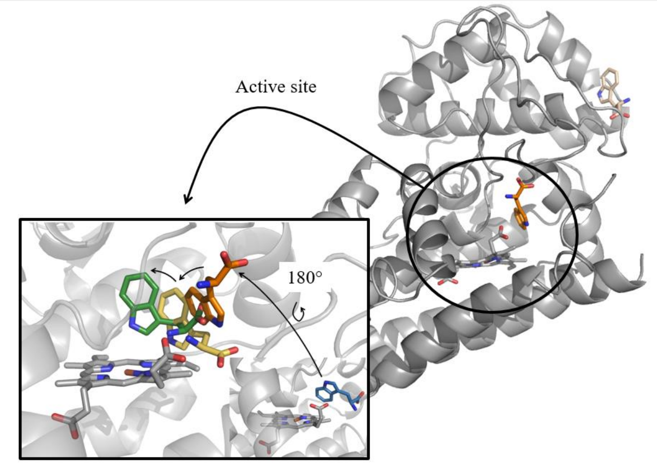
Tailoring the properties of 2D materials



https://prace-ri.eu/simulations-help-to-tailor-the-properties-of-2d-materials/

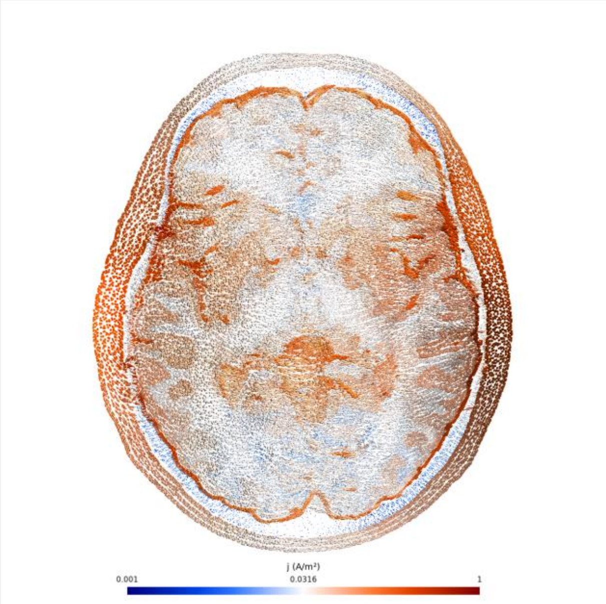
Examples: biology

Unraveling the behavior of the hIDO1 protein



https://www.enccb.be/usmanonmirgaux

Understanding how radio waves propagate in the head with Shamo



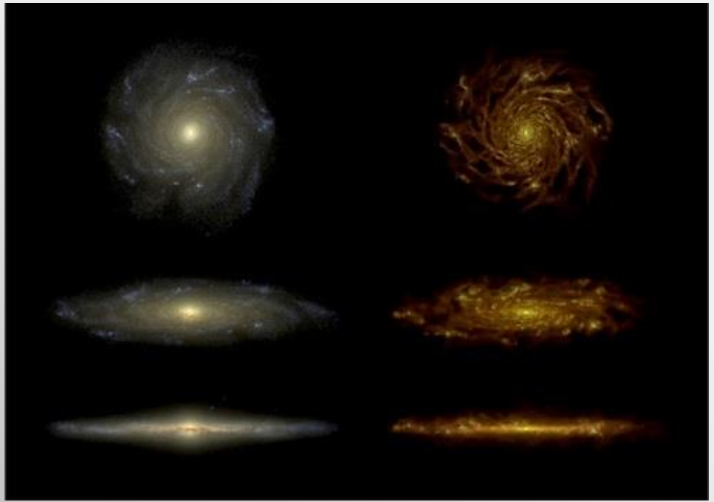
https://www.enccb.be/usshamo

Improving chocolate with supercomputing

https://www.enccb.be/uschocolate

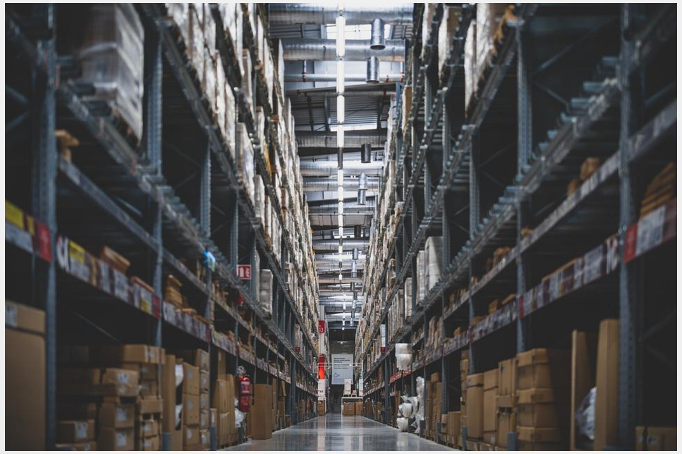
Examples: physics, mathematics, engineering,…

Simulating Galaxies



https://www.enccb.be/usuniverse

Outsmarting NP-hardness



https://www.enccb.be/uslogisticsnphardness

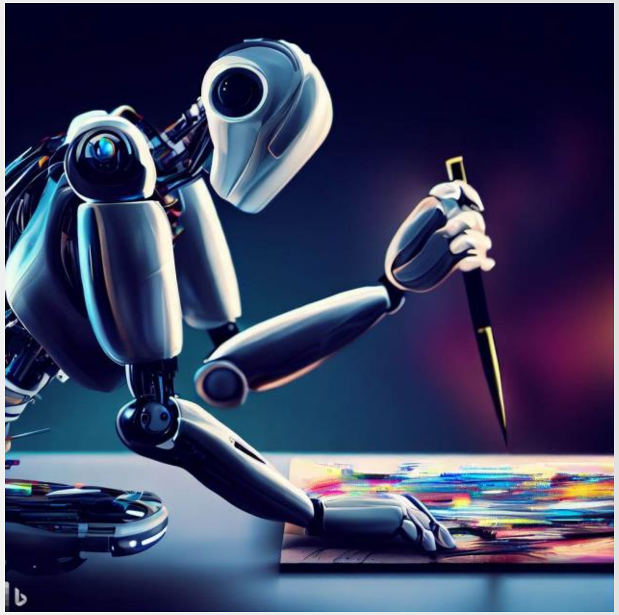
Optimizing particle processes through simulation with MPacts

https://www.enccb.be/usmpacts

Examples

But also:

- Artificial intelligence
- Machine learning
- Data analysis
- …



What is EuroCC?

- EuroHPC Join Undertaking is a joint initiative between the EU (European countries) and private partners to develop a world class supercomputing system in Europe.
- EuroCC is a project of EuroHPC-JU, which tasks each participating countries to create a national competence center (NCC) in the area of HPC. They coordinate activities in the HPC/HPDA/AI fields and serve as contact point.
  → More information on https://www.enccb.be/.
- Different missions, including contact with industry, raising awareness and organizing trainings.

Outline of this presentation

EURO²

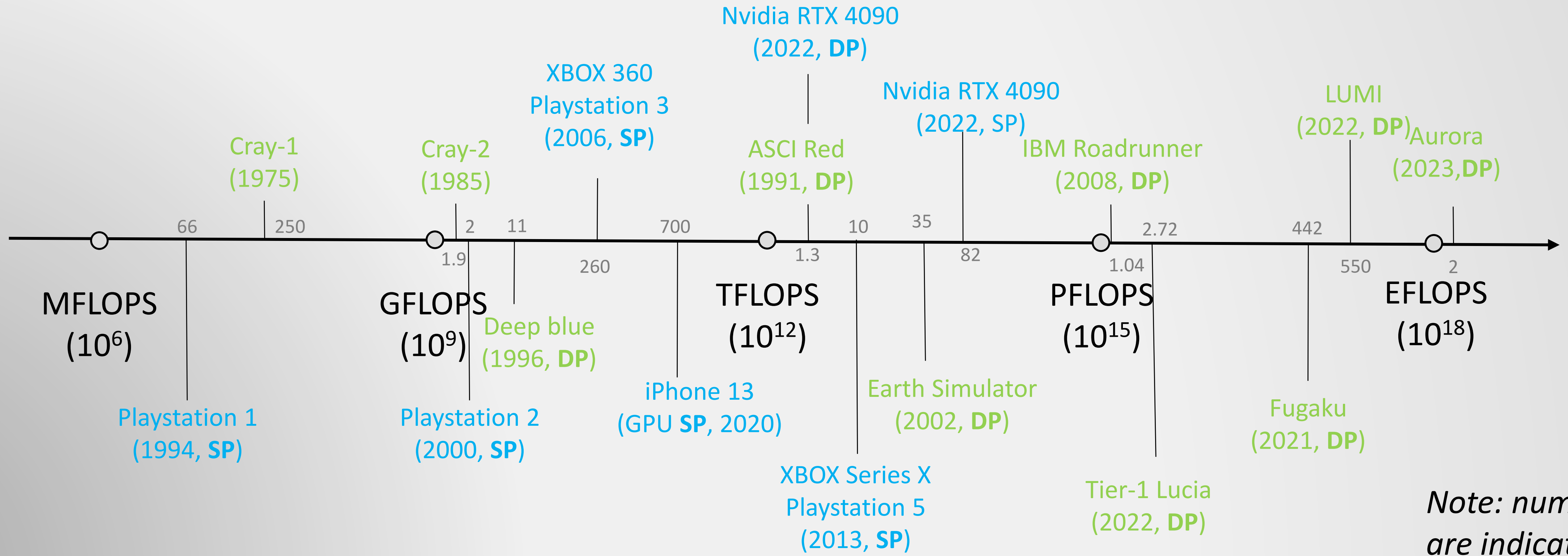
- Current status of the supercomputing infrastructures
  - Performance and the TOP500 list
  - Supercomputers in Europe
  - Supercomputers in Belgium
- Understand how a supercomputer works:
  - Architecture
  - Components
  - Interactions
- Understand how program can use such large resources, and what are the issues that needs to be overcome:
  - Parallelism
  - Parallelisation issues

Where are we today?

Which supercomputers for Belgium

Performances

- One of the measurement is FLOPS = floating point operations per second (≃ speed)
- Note that it depends on single (SP, 32 bits) or double (DP, 64 bits) precision!

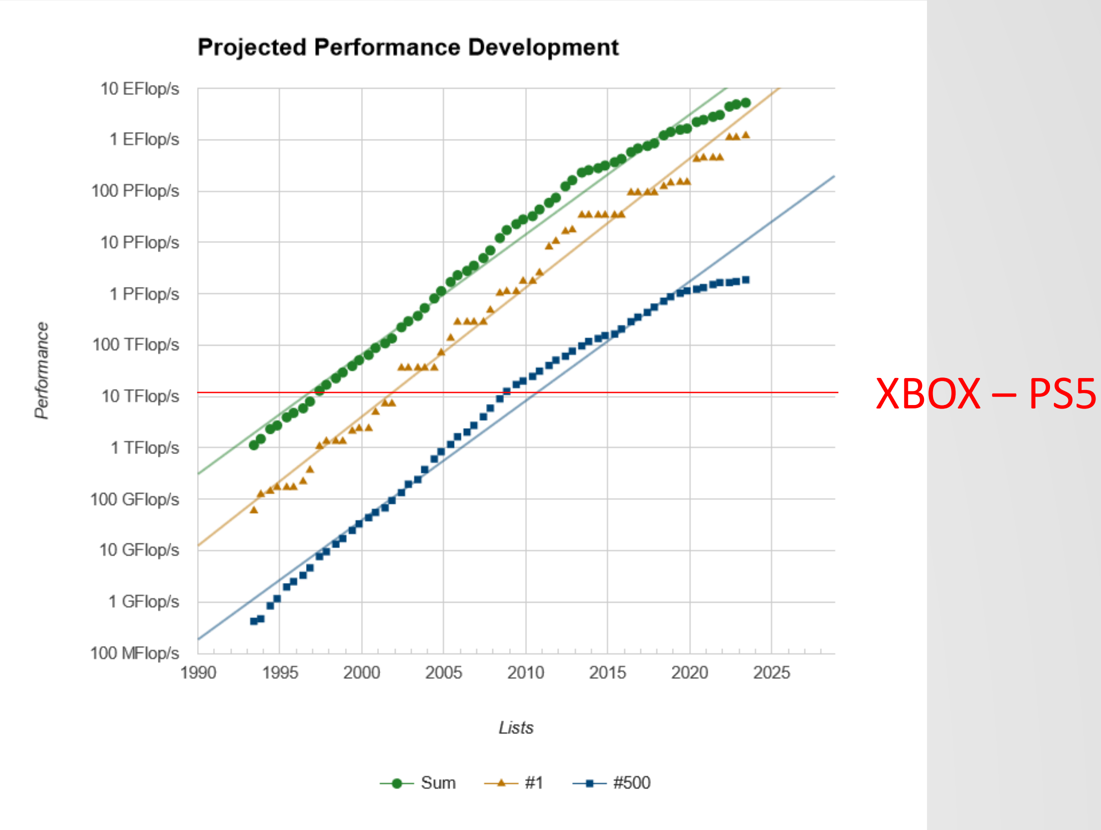


Note: numbers are indicative

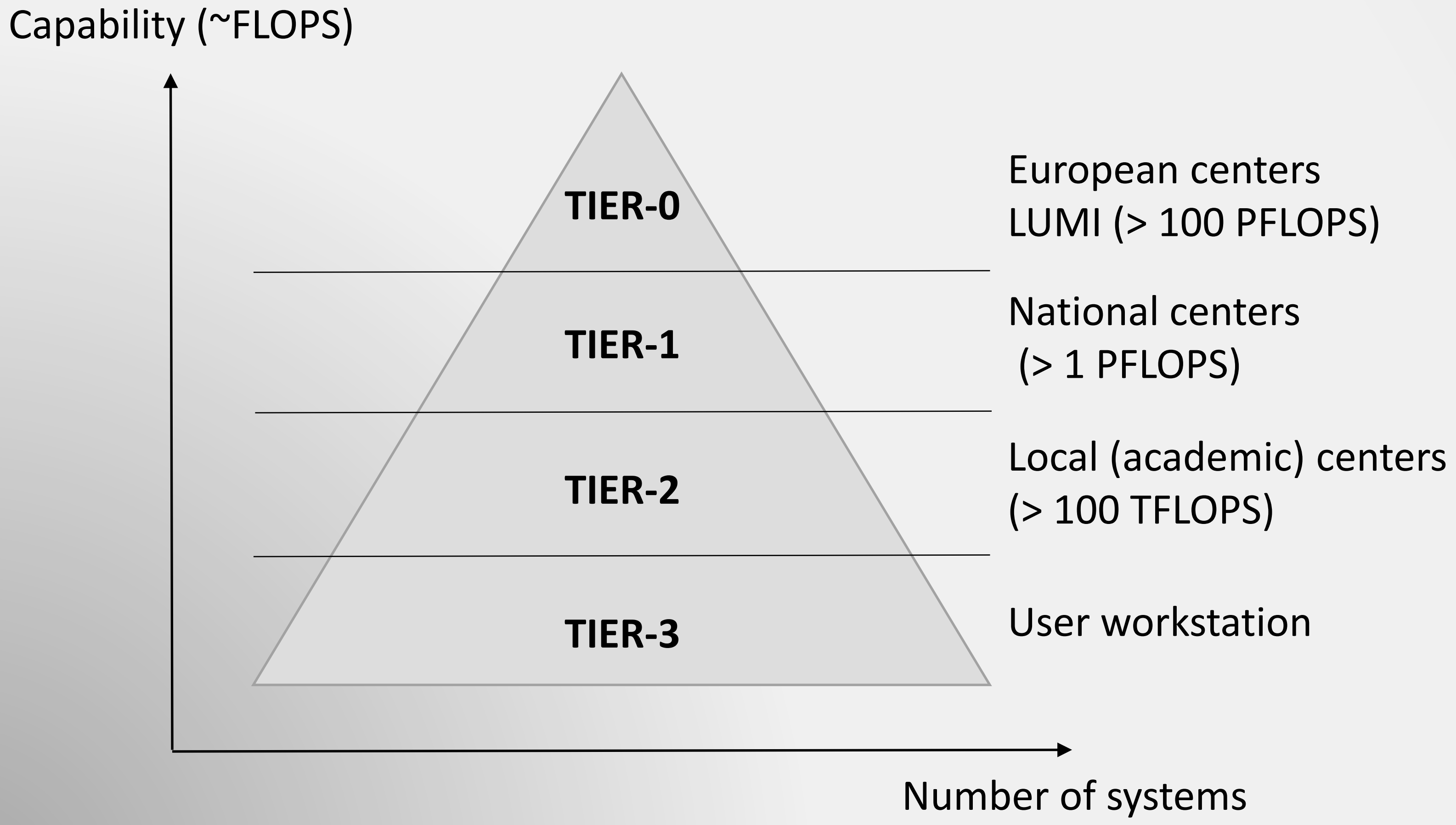
The TOP500

The TOP500 is a list of the fastest supercomputer in the world

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	8,699,904	1,194.00	1,679.82	22,703
2	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
3	**LUMI** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,220,288	309.10	428.70	6,016
4	**Leonardo** - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband, Atos EuroHPC/CINECA Italy	1,824,768	238.70	304.47	7,404
5	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148.60	200.79	10,096
6	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94.64	125.71	7,438

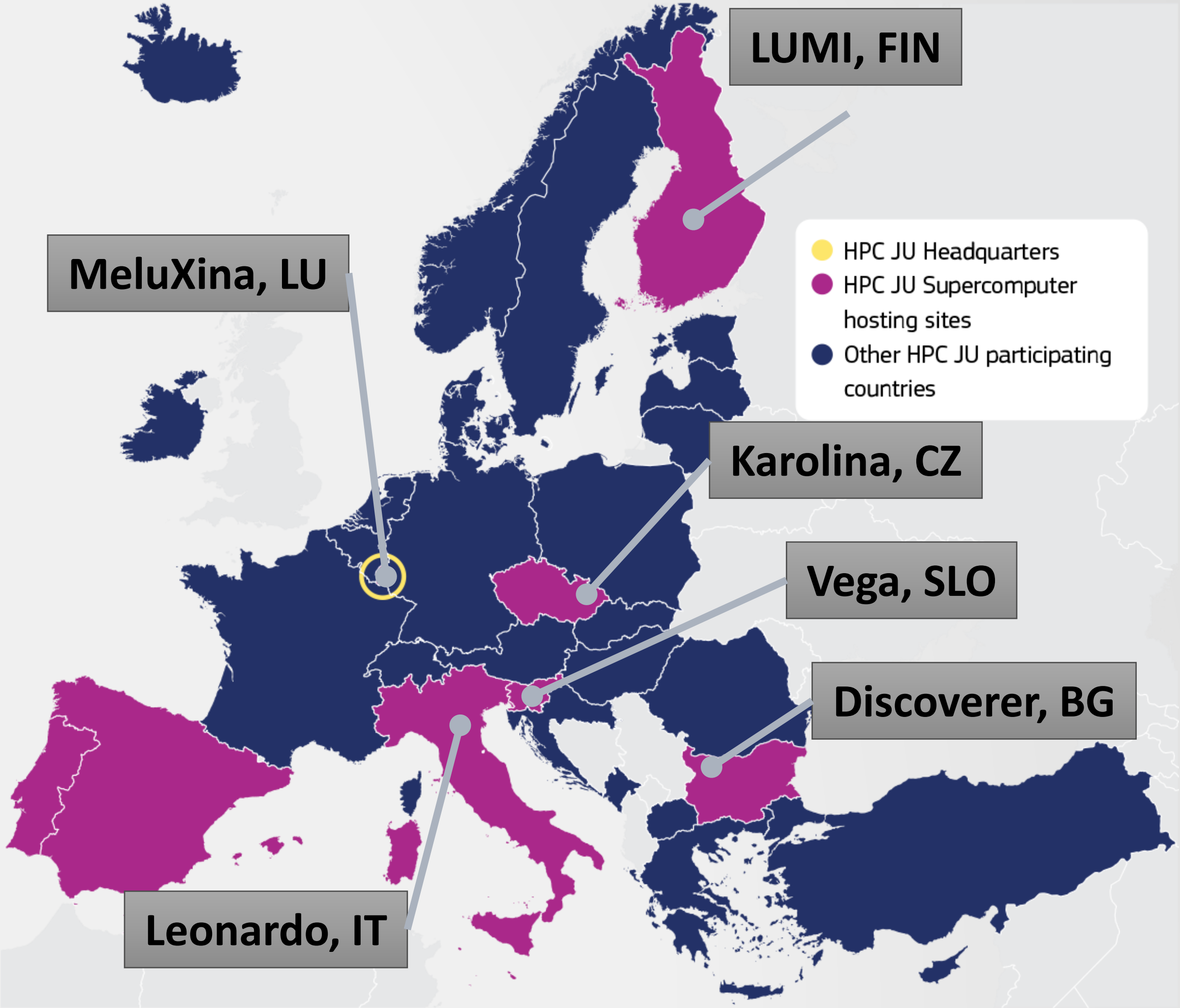


Supercomputers in Europe



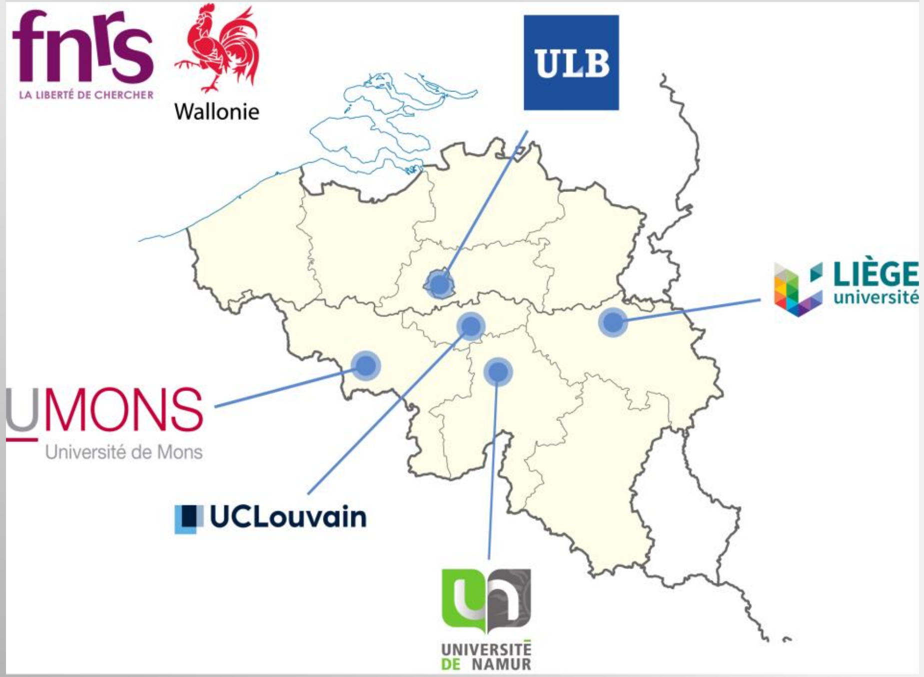
Note: getting access to the top TIER requires to demonstrate scalability on the previous ones

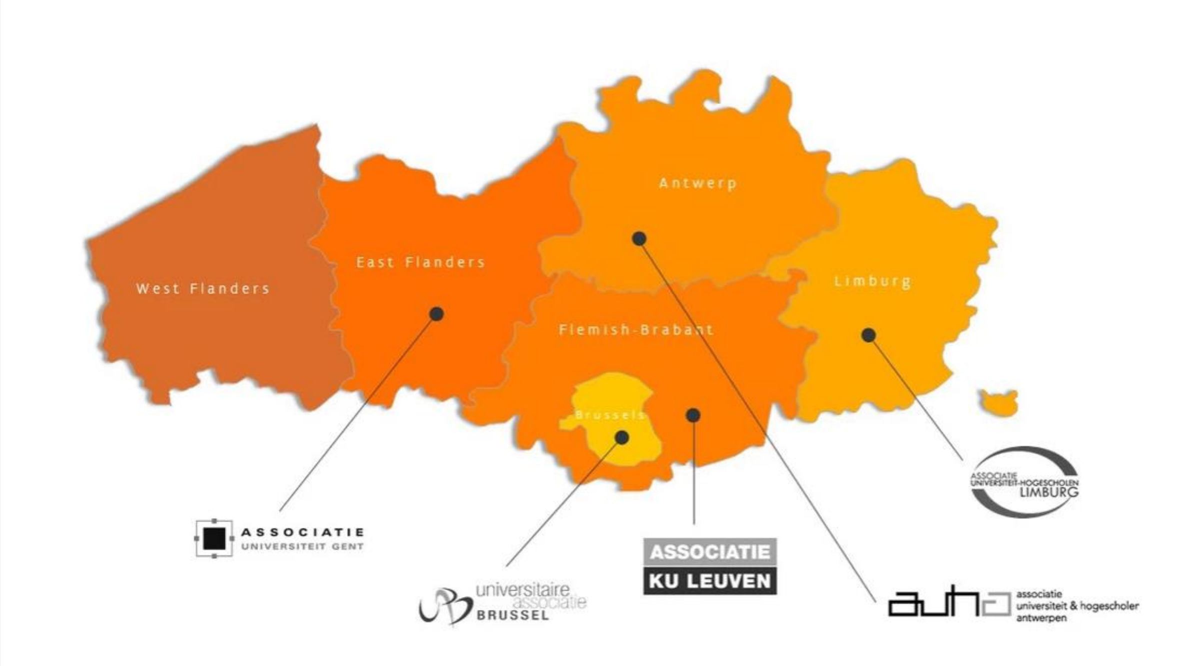
European infrastructure



TIER-2 in Belgium (academic level)

*Wallonia: the **CÉCI** (http://www.ceci-hpc.be/)*



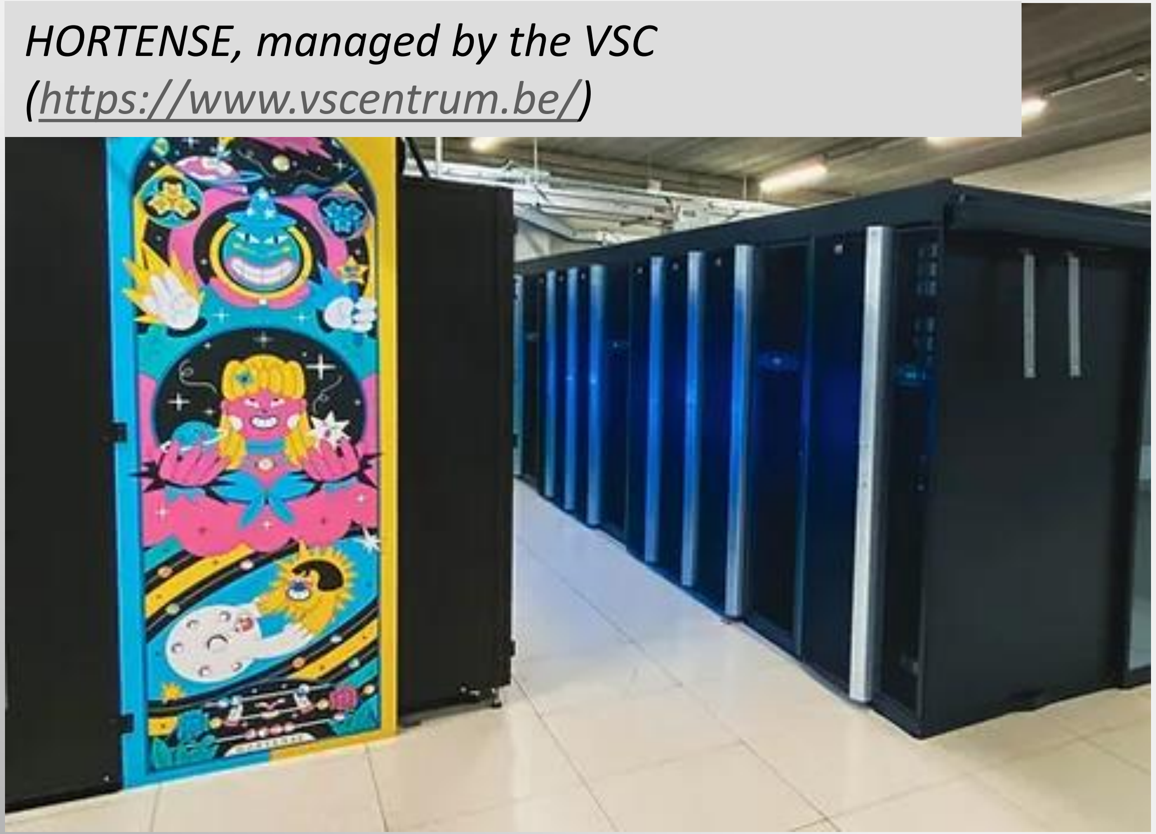


*Flanders: the **VSC** (https://www.vscentrum.be/)*

Accessible for every researchers of the corresponding universities. **Trainings** are also provided

TIER-1 in Belgium

HORTENSE, managed by the VSC
(https://www.vscentrum.be/)



Lucia, managed by Cenaero
(https://www.cenaero.be/)

Provides **access** (through calls) and **support**.

TIER-0 (LUMI)

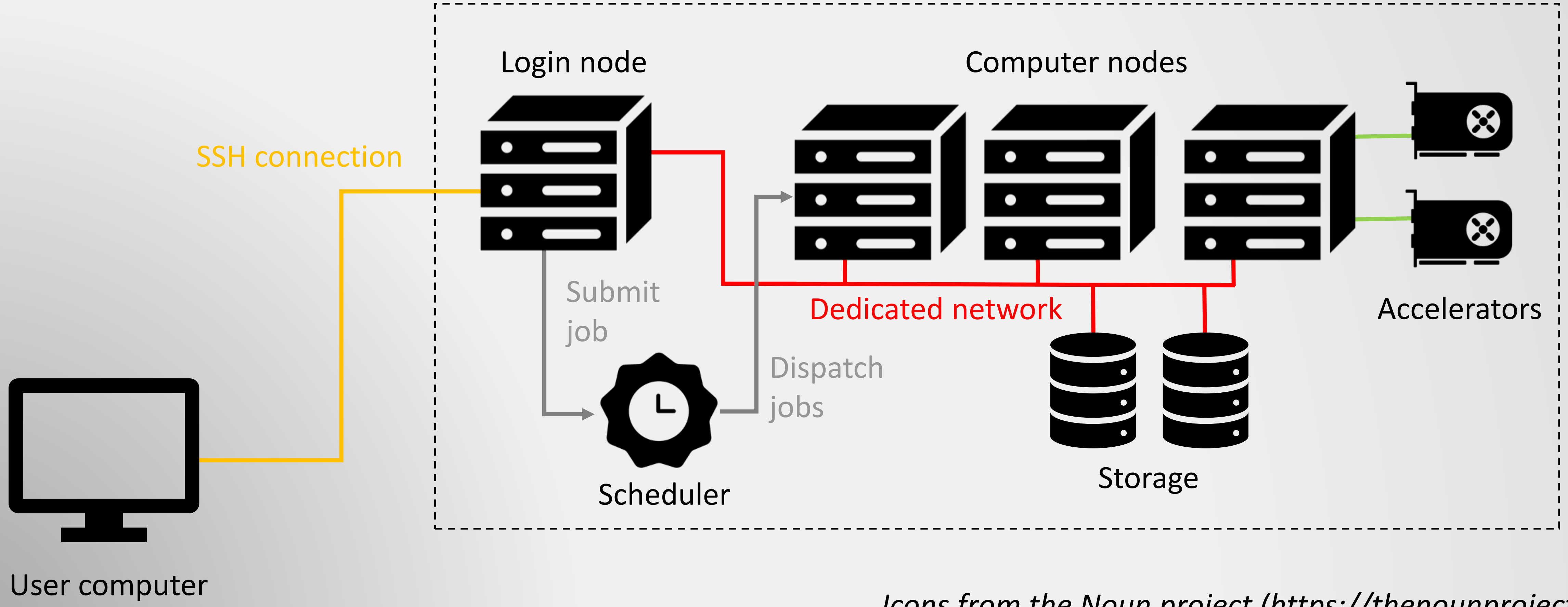
https://www.lumi-supercomputer.eu/
https://www.enccb.be/LUMI/

- First european pre-exascale supercomputer
- 100% hydropowered energy
- Regular training in Belgium

What are supercomputers?

How they work

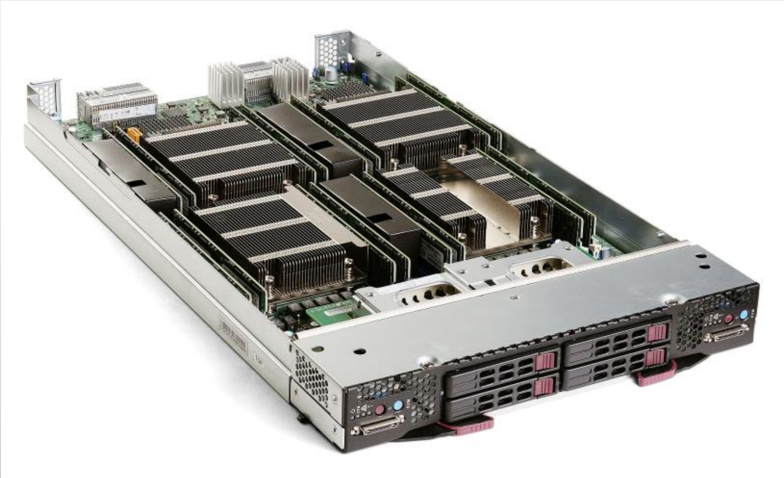
Anatomy of a cluster



Icons from the Noun project (https://thenounproject.com/)

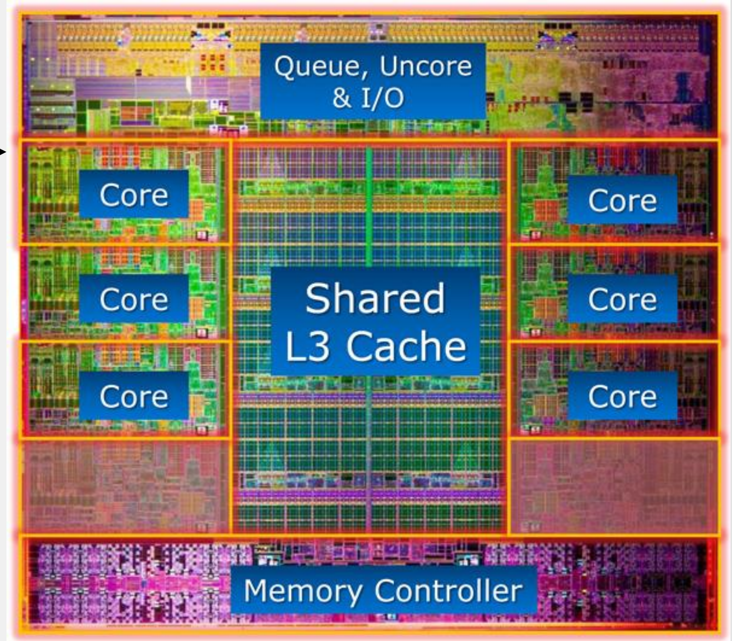
Components

EURO²



Computer node (Wikipedia)
Acts basically like a computer

- Socket (CPU)
- RAM
- Networking
- Cooling
- Local storage
- ...

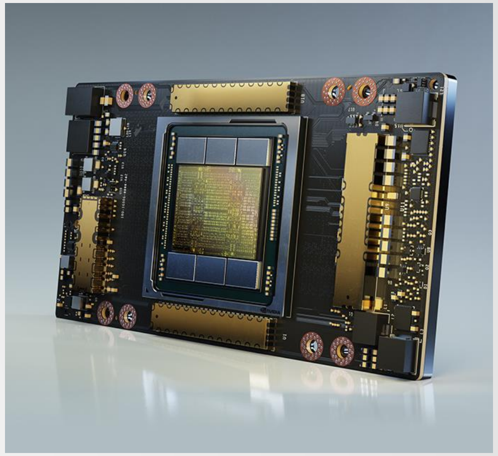


Intel sandybridge (die)
(https://www.anandtech.com/)

Components: GPU / Accelerators

- Instead of a few powerful cores, many less powerful cores
- Consumer grade GPU: provide good FLOPS for single precision operation, not for double precision
- Dedicated class of cluster grade GPU (e.g., NVIDIA Ampere or AMD Instinct)
- Future of HPC?

AMD Instinct (amd.com)



Ampere (nvidia.com)

Interaction with the supercomputer

- Generally, through command line
- Users submit jobs on the cluster, and wait for the results
- For each "job", the scheduler (e.g., SLURM) requires to know time/memory/number of processor and node and tries to fit the job when a slot is available.
- Different clusters have different purposes, e.g.,
  - High-memory application,
  - Small and fast jobs,
  - Nodes with accelerators, ...

How to use such large resources efficiently?

Solutions... and challenges

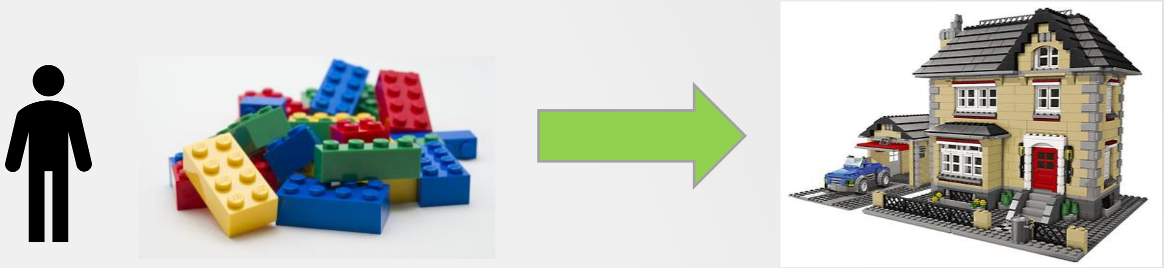
Particularities of supercomputing

- Advantage of supercomputing
  - Larger size: some problem requires large amount of memory
  - More speed: some problem requires long time to be solved
- Solution … And issues:
  - More memory (but storage hierarchy)
  - Parallelism (… has inherent difficulties)

Tackle speed: parallel computing

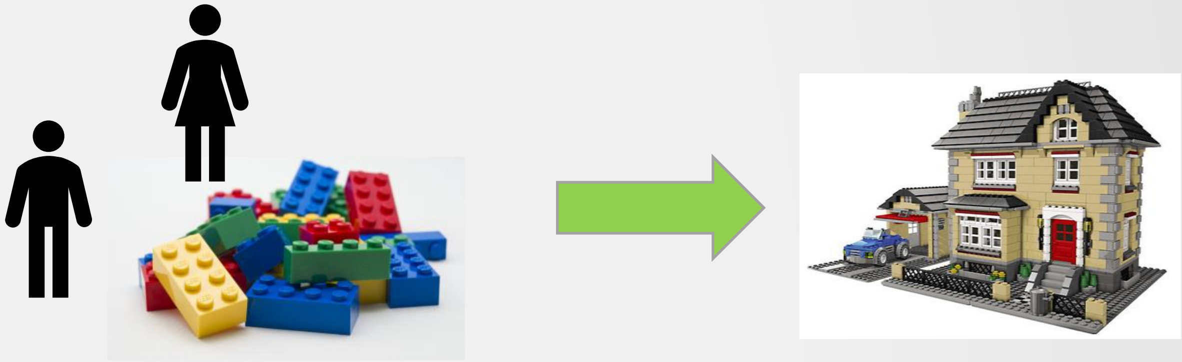
- To achieve such performances, the main idea is to rely on **parallel computing :** executing many operations in a single instance of time.
- But the program needs to be adapted for such purpose!

Serial world: 1 worker (person) to build



Images from LEGO (lego.com)

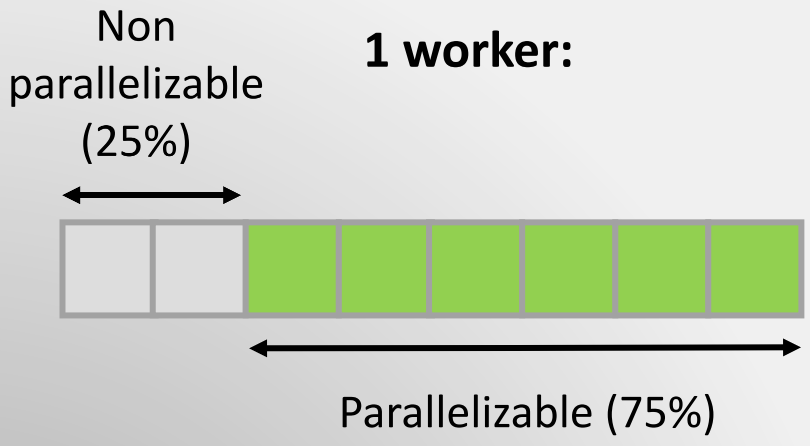
Parallel world: 2 workers to build → about 2 times faster



Speedup is never what you expect...

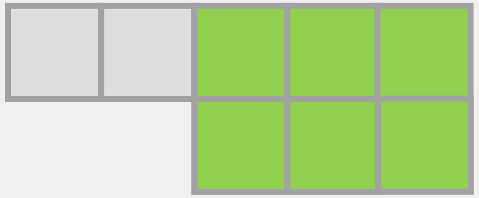
In most cases, a problem is never fully parallelizable (i.e., **embarrassingly** parallel)

1 worker:



$T_{tot,1} = 8$

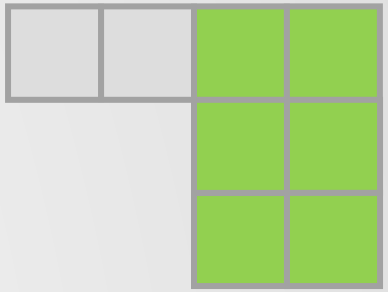
2 workers:



$T_{tot,2} = 5$

Speedup = $T_{tot,1} / T_{tot,2} = 1.6$

3 workers:



$T_{tot,3} = 4$

Speedup = $T_{tot,1} / T_{tot,3} = 2$

→ No matter how fast the parallel portion, we will always be limited by the **serial** part.

Speedup is never what you expect...

Amdahl's Law:

$$S = \frac{1}{s + \frac{P}{N}}$$

S: actual speedup

s: serial portion of the code (in %)

P: parallel portion (in %)

N: number of processors

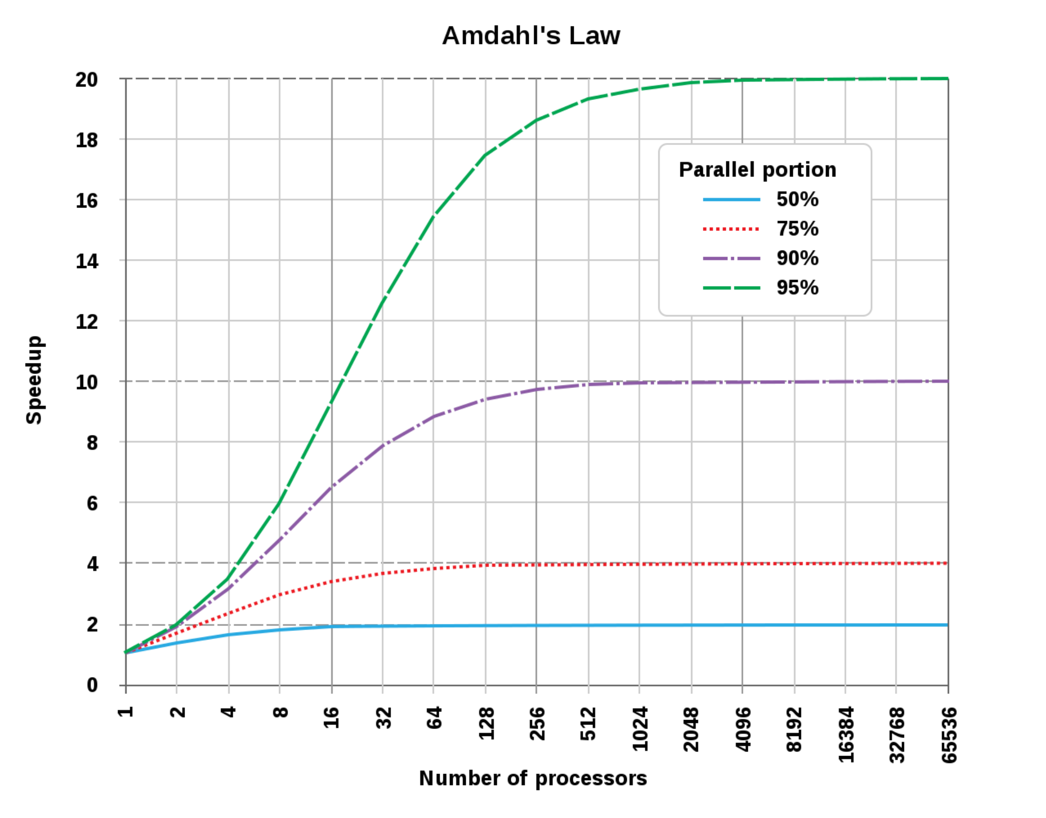


Image from Wikipedia (wikipedia.org)

This address the question "how much processors can I use for a given problem?"

On the bright side: weak scaling

If we increase the size of the problem when more processors are added,

$$S = N - s(N-1)$$

S: actual speedup

N: number of processors

S: serial fraction (in %)

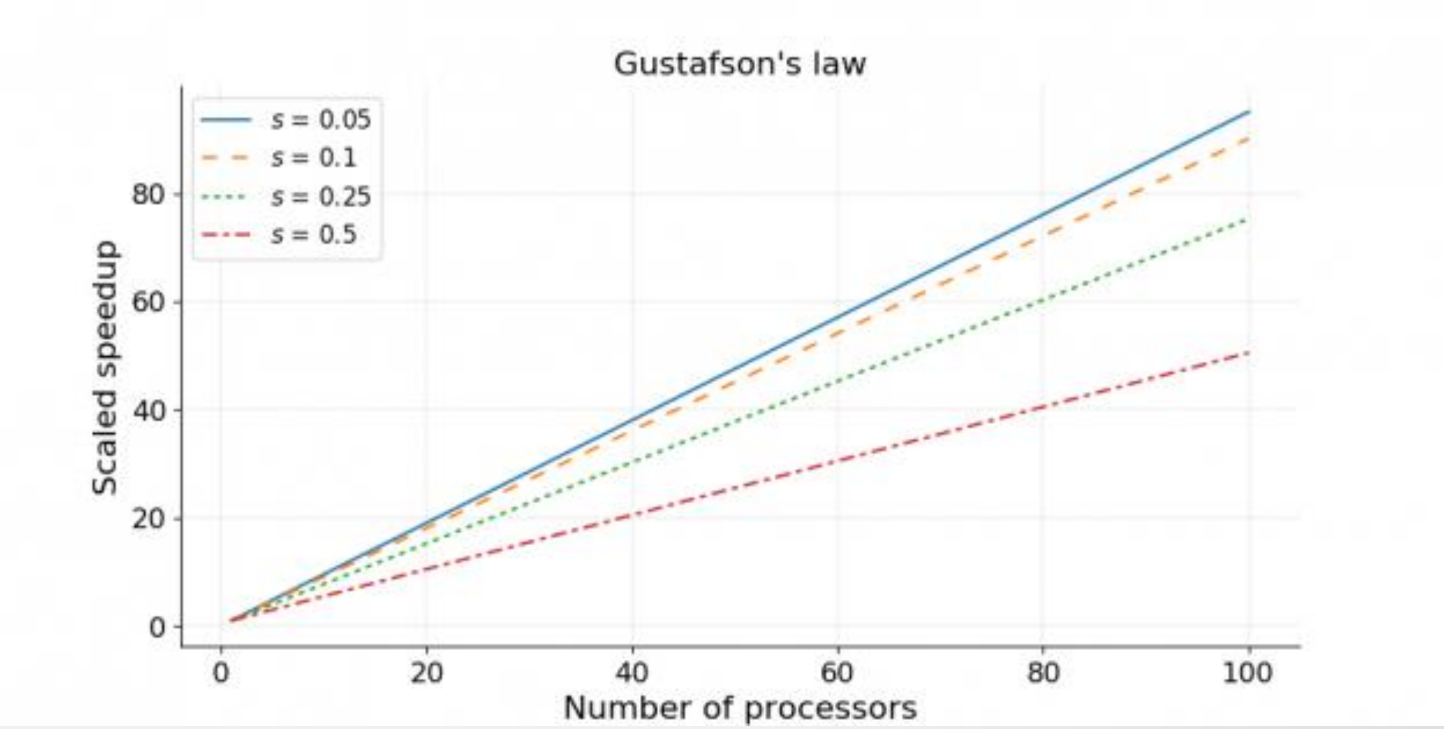
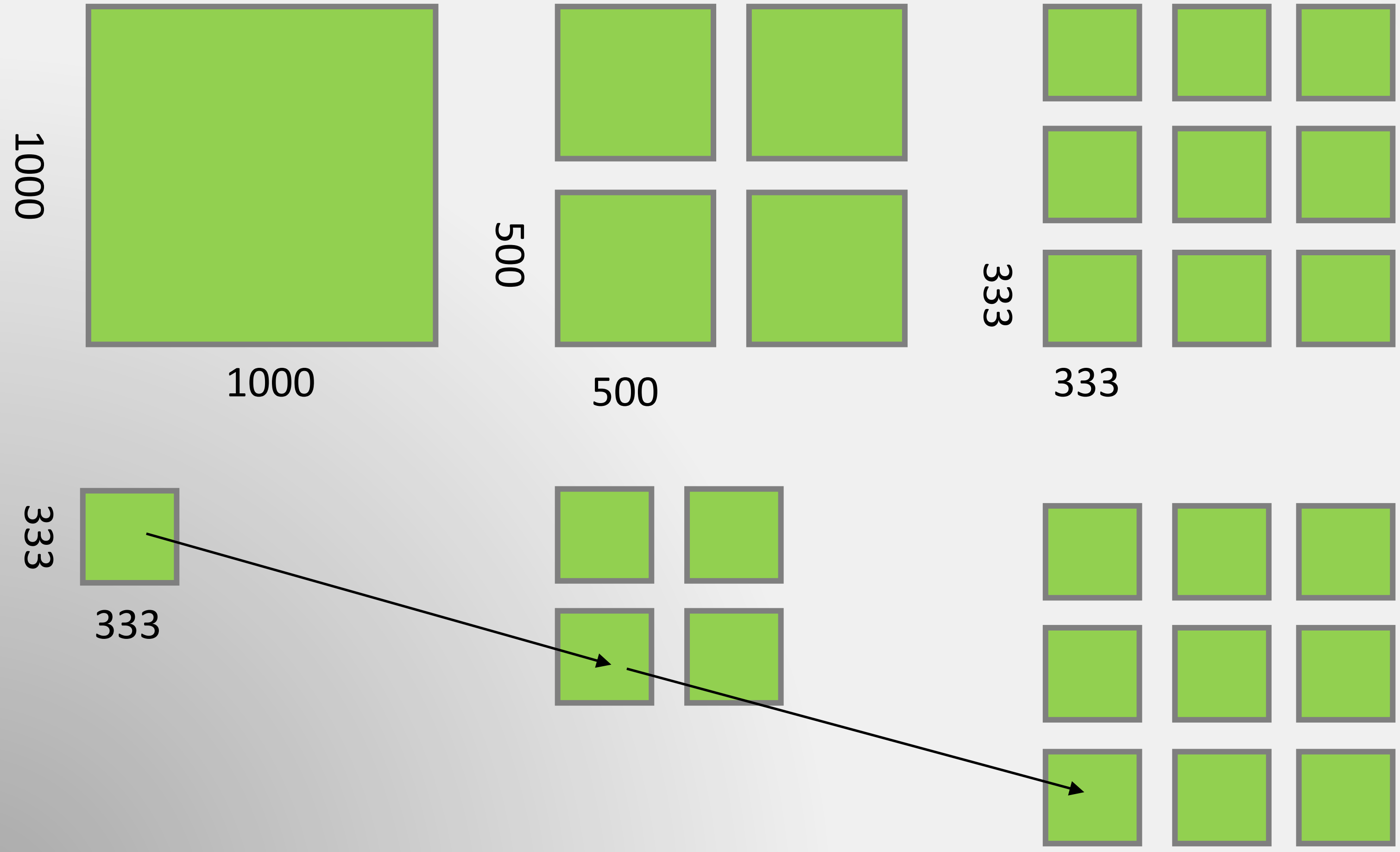


Image from Wikipedia (wikipedia.org)

This address the question "how much can I increase the size of my problem such that the execution time is the same as if I ran the problem with only one process?"

Recap of strong vs weak scaling

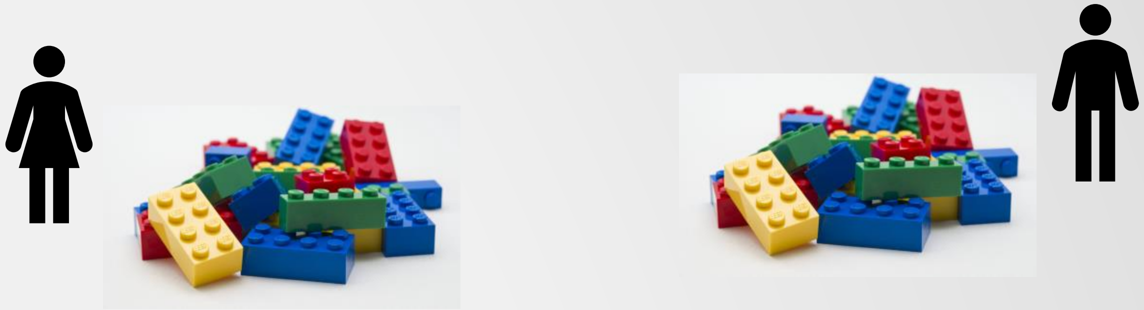


Strong scaling: size of the problem is constant and split across additional processors

Weak scaling: the size of the problem is the same for each processor.

… and there is parallel overhead!

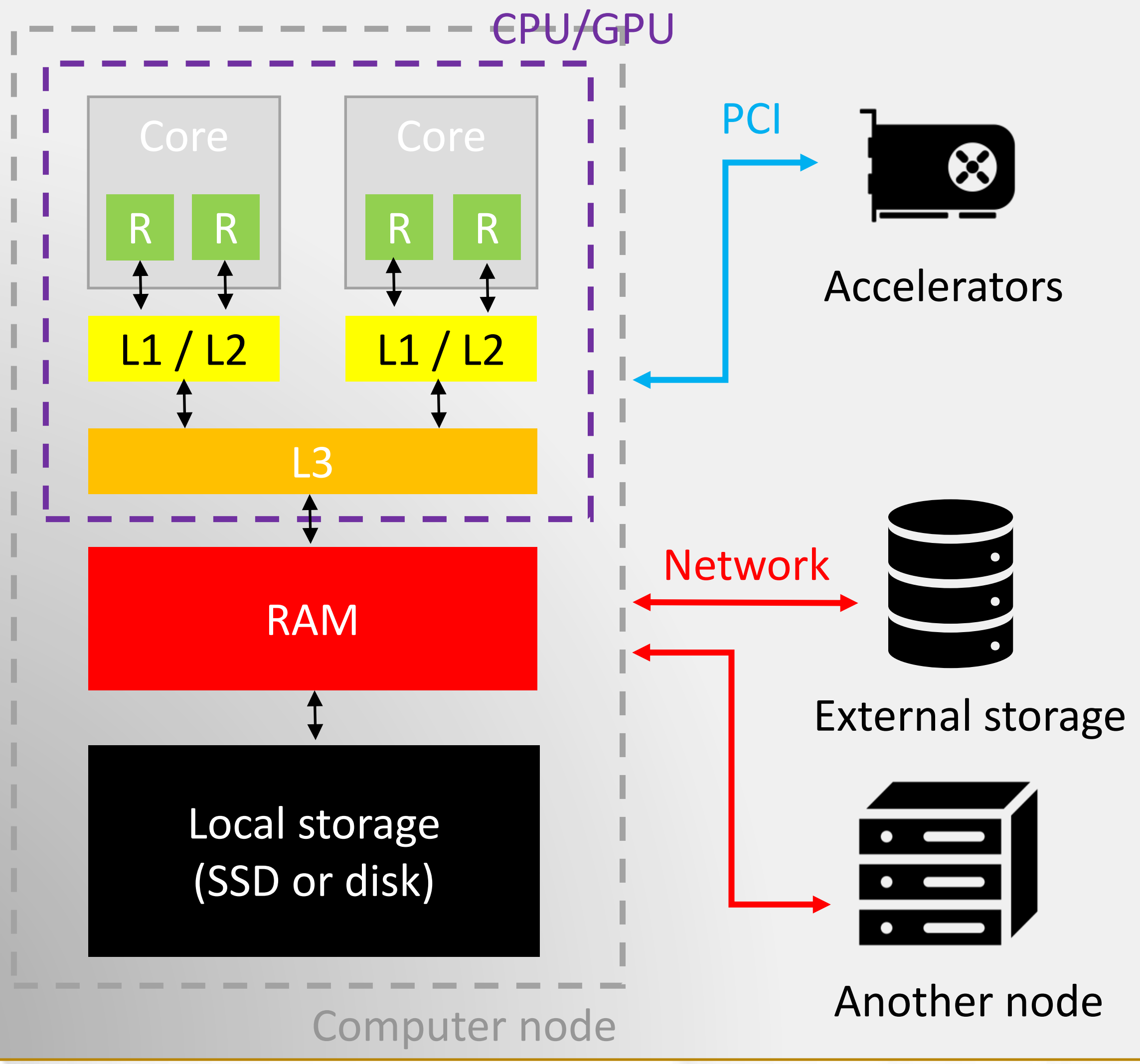
Shared memory model: all workers (or, here, person) are working on the same pool of data (or, here, lego pieces). Small overhead due to **synchronisation** (all person cannot work on the same part of the model at the same time).



Distributed memory model: each worker (or, here, person) is working on its own set of data (or, here, lego pieces). Generally, more efficient (no collaboration during work), but overhead due to **communication** (or, here, distribute pieces in the beginning and assembling the result at the end).

Generally, the two are mixed. It also requires a good **load balancing** (i.e., every person has the same amount of work to do). It is not that easy to achieve.

Also important: storage hierarchy



- The further away from the core, the slowest (but, generally, the more capacity)
- I/O may not be parallel
- Communication is a bottleneck when using multiple nodes
- Efficient movement of data to and from an accelerators

Icons from the Noun project (https://thenounproject.com/)

Tools of the trade

Parallel programming:

- Vectorization (core level)
- Threading / OpenMP (node level)
- CUDA / HIP / OpenCL / OpenACC / OpenMP (accelerator level)
- Socket / MPI / PGAS (cluster level)

Optimized libraries:

- BLAS / LAPACK / MKL (linear algebra)
- FFTW (fast-Fourier transform)
- HDF5 / netCDF (parallel I/O)

To conclude

Conclusions

EURO²

- HPC is important for actual and future research
- HPC is an active field in Belgium
- There are challenge to overcome, but tools are developed
- This was an introduction, don't be afraid to reach us if you want to know more!
  → More information on https://www.enccb.be/.

Thanks!

Funded by
the European Union

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia under grant agreement No 101101903.