Energy Aware Simulations

Benjamin Czaja
HPC Advisor SURF
April 2024

https://ondemand.snellius.surf.nl/

https://github.com/sara-nl/energy-efficient-computing

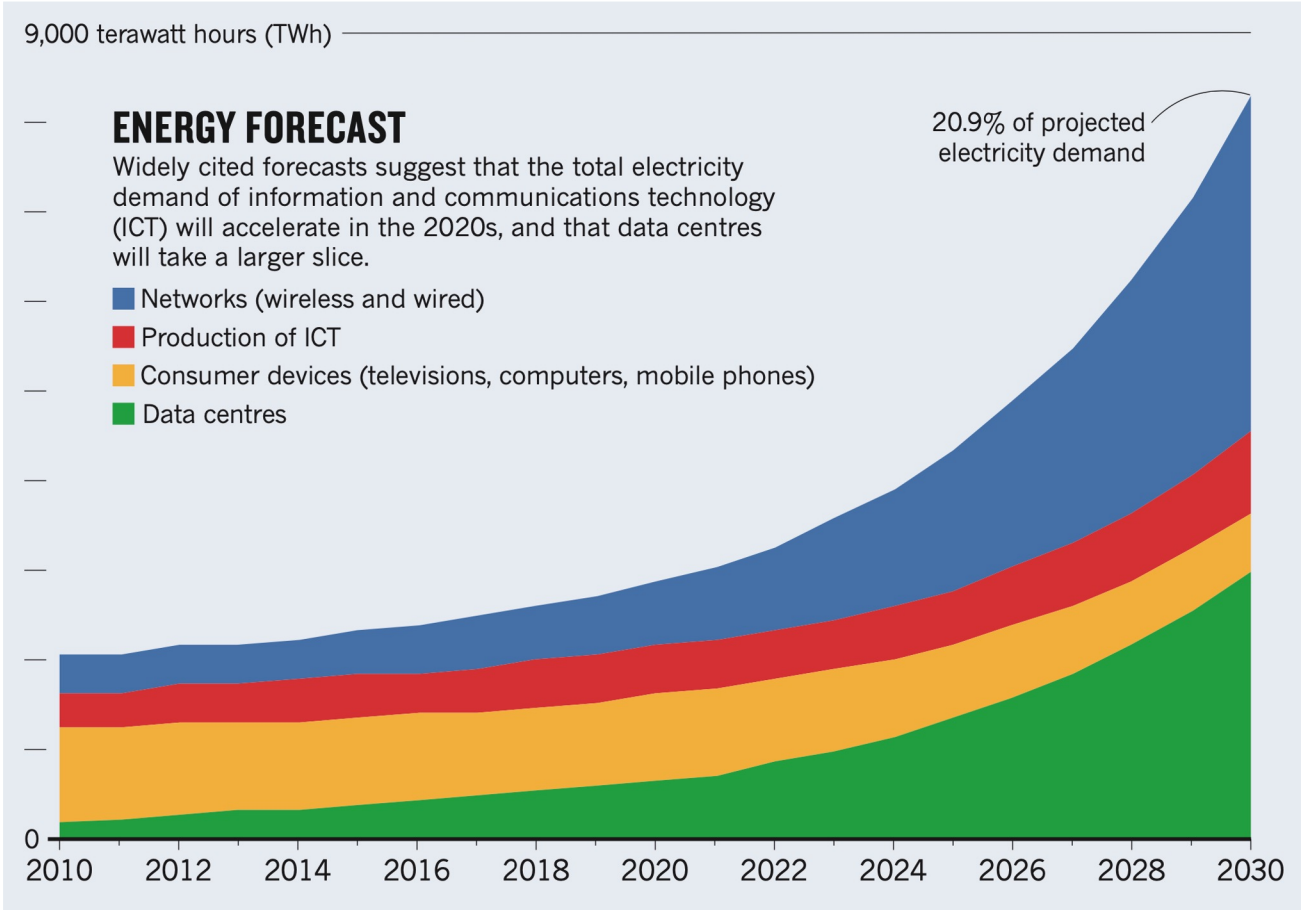
login: scur0311

Reservation:
jupyterhub_course_jhssrf009_2024-04-18

scur0311

ICT/Data center energy forecast

- Data centre energy usage:
  - ~ 200 TWh for data centres in 2018
  - ~ 3000 TWh in 2030
- Dutch Data Center Usage (2019)
  - 1300 MW installed capacity
  - 0.3 % electricity usage of the Netherlands (CBS)
  - 3x as much power than the NS



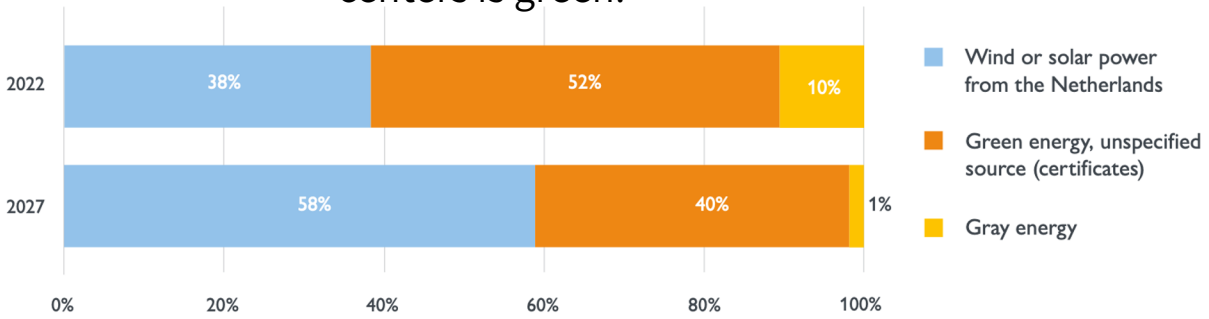
The chart above is an 'expected case' projection from Anders Andrae, a specialist in sustainable ICT. In his 'best case' scenario, ICT grows to only 8% of total electricity demand by 2030, rather than to 21%.

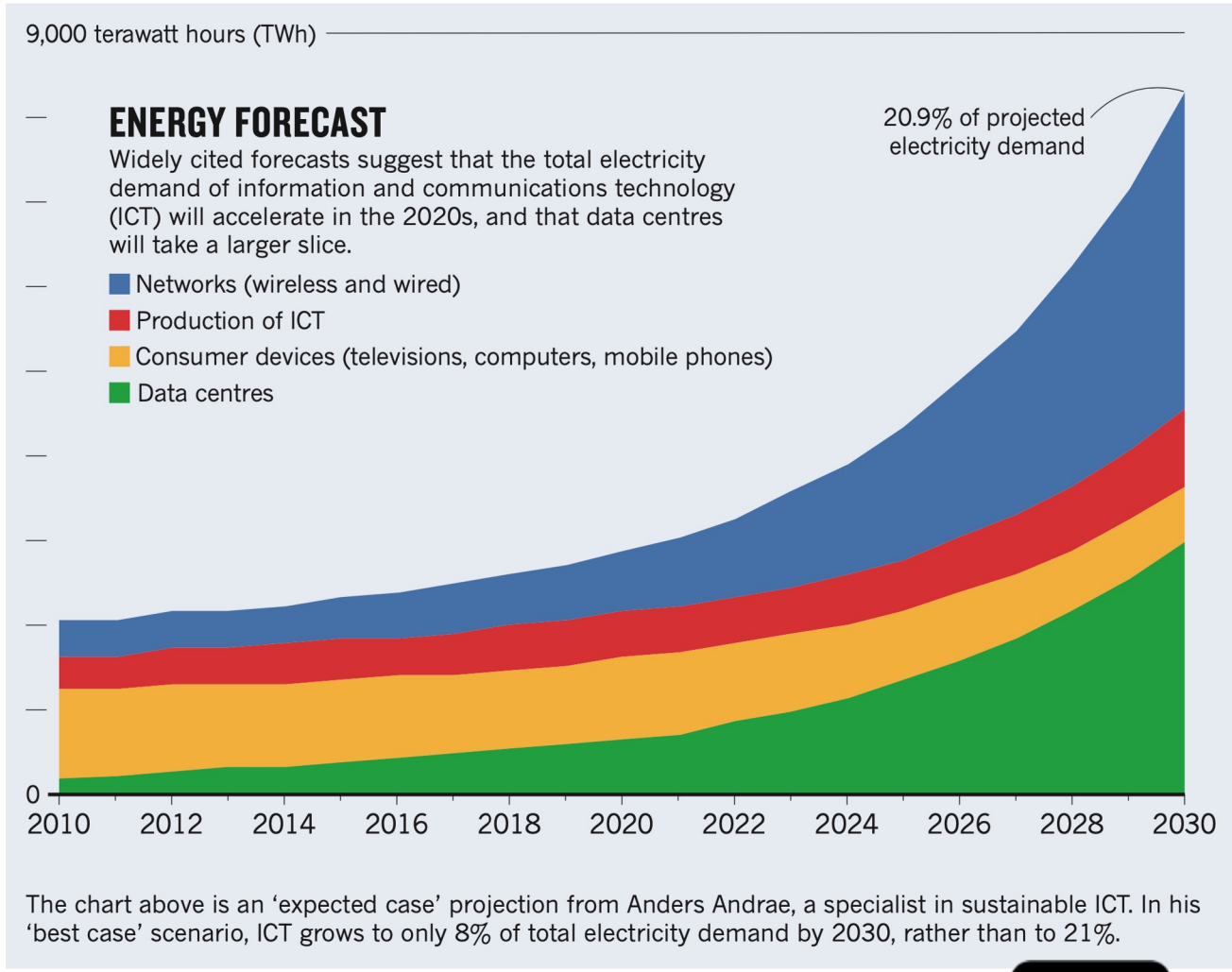
- Jones, Nicola. "How to stop data centres from gobbling up the world's electricity." *Nature* 561.7722 (2018): 163-166.
- Andrae, Anders SG, and Tomas Edler. "On global electricity usage of communication technology: trends to 2030." *Challenges* 6.1 (2015): 117-157.
- Bakkeren, Hanno. "Datacenters verbruiken drie keer zoveel stroom als de NS". NRC, 14 May 2019
- State of the Dutch Data Centers, The Dutch Data Center Report, 2022, https://www.dutchdatacenters.nl/

ICT/Data center energy forecast

- Data centre energy usage:
  - ~ 200 TWh for data centres in 2018
  - ~ 3000 TWh in 2030
- Dutch Data Center Usage (2019)
  - 1300 MW installed capacity
  - 0.3 % electricity usage of the Netherlands (CBS)
  - 3x as much power than the NS

- Dutch Data center energy consumption (2022)
  - 90% of all energy consumed by colocation data centers is green.

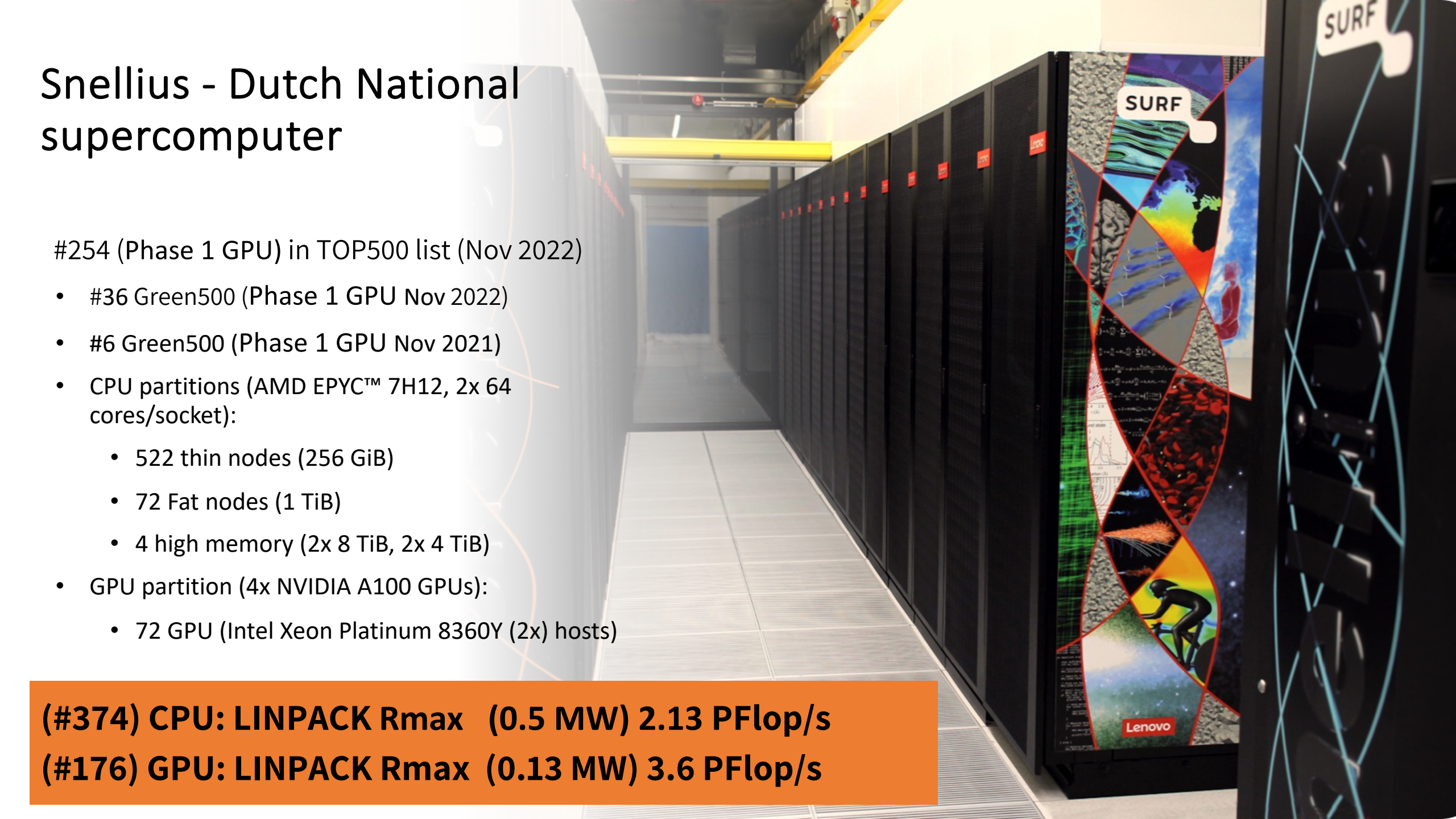




The chart above is an 'expected case' projection from Anders Andrae, a specialist in sustainable ICT. In his 'best case' scenario, ICT grows to only 8% of total electricity demand by 2030, rather than to 21%.

- Jones, Nicola. "How to stop data centres from gobbling up the world's electricity." *Nature* 561.7722 (2018): 163-166.
- Andrae, Anders SG, and Tomas Edler. "On global electricity usage of communication technology: trends to 2030." *Challenges* 6.1 (2015): 117-157.
- Bakkeren, Hanno. "Datacenters verbruiken drie keer zoveel stroom als de NS". NRC, 14 May 2019
- State of the Dutch Data Centers, The Dutch Data Center Report, 2022, https://www.dutchdatacenters.nl/

Snellius - Dutch National supercomputer



#254 (Phase 1 GPU) in TOP500 list (Nov 2022)

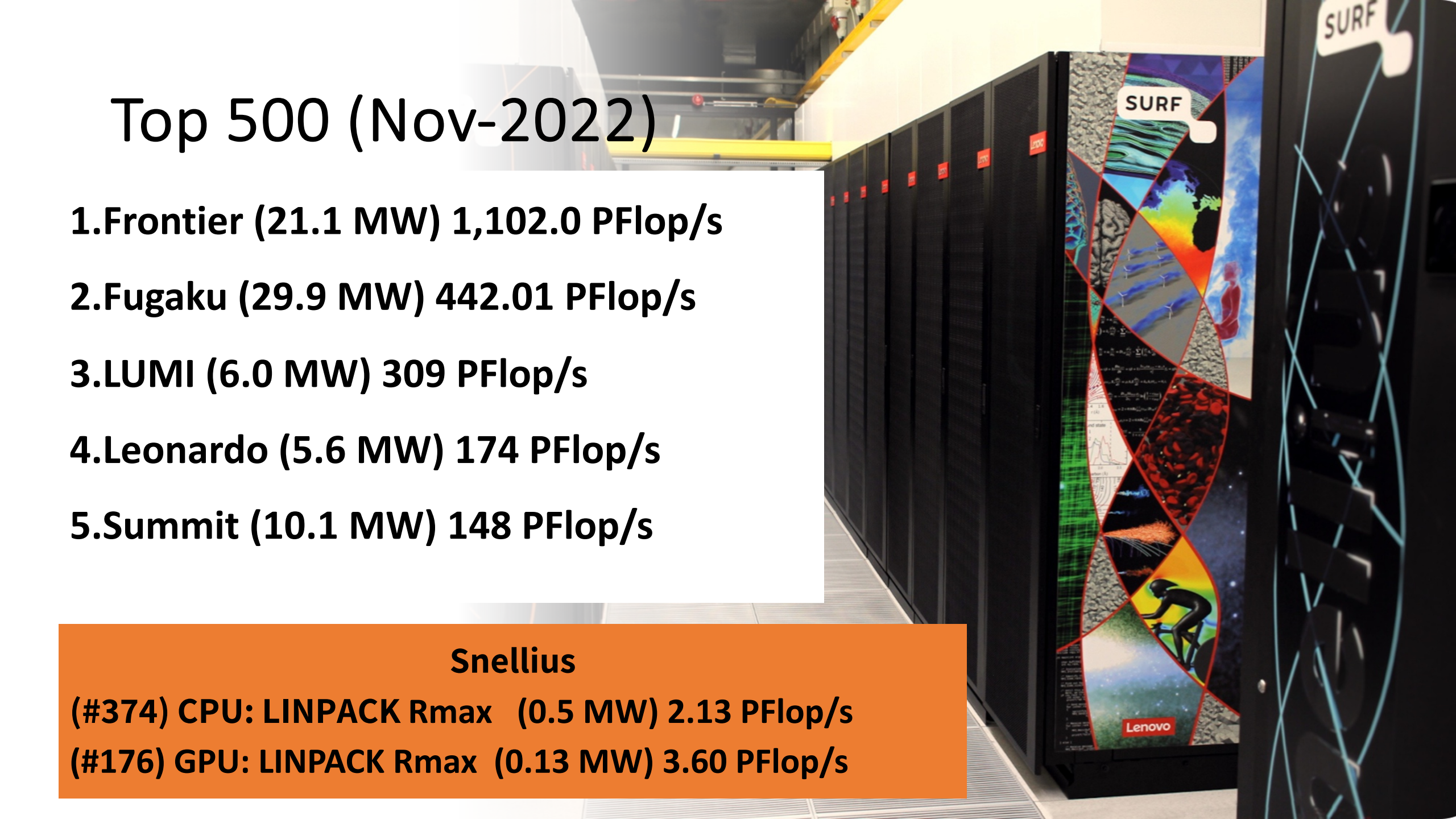
- #36 Green500 (Phase 1 GPU Nov 2022)
- #6 Green500 (Phase 1 GPU Nov 2021)
- CPU partitions (AMD EPYC™ 7H12, 2x 64 cores/socket):
  - 522 thin nodes (256 GiB)
  - 72 Fat nodes (1 TiB)
  - 4 high memory (2x 8 TiB, 2x 4 TiB)
- GPU partition (4x NVIDIA A100 GPUs):
  - 72 GPU (Intel Xeon Platinum 8360Y (2x) hosts)

(#374) CPU: LINPACK Rmax (0.5 MW) 2.13 PFlop/s

(#176) GPU: LINPACK Rmax (0.13 MW) 3.6 PFlop/s

Top 500 (Nov-2022)

1. Frontier (21.1 MW) 1,102.0 PFlop/s
2. Fugaku (29.9 MW) 442.01 PFlop/s
3. LUMI (6.0 MW) 309 PFlop/s
4. Leonardo (5.6 MW) 174 PFlop/s
5. Summit (10.1 MW) 148 PFlop/s

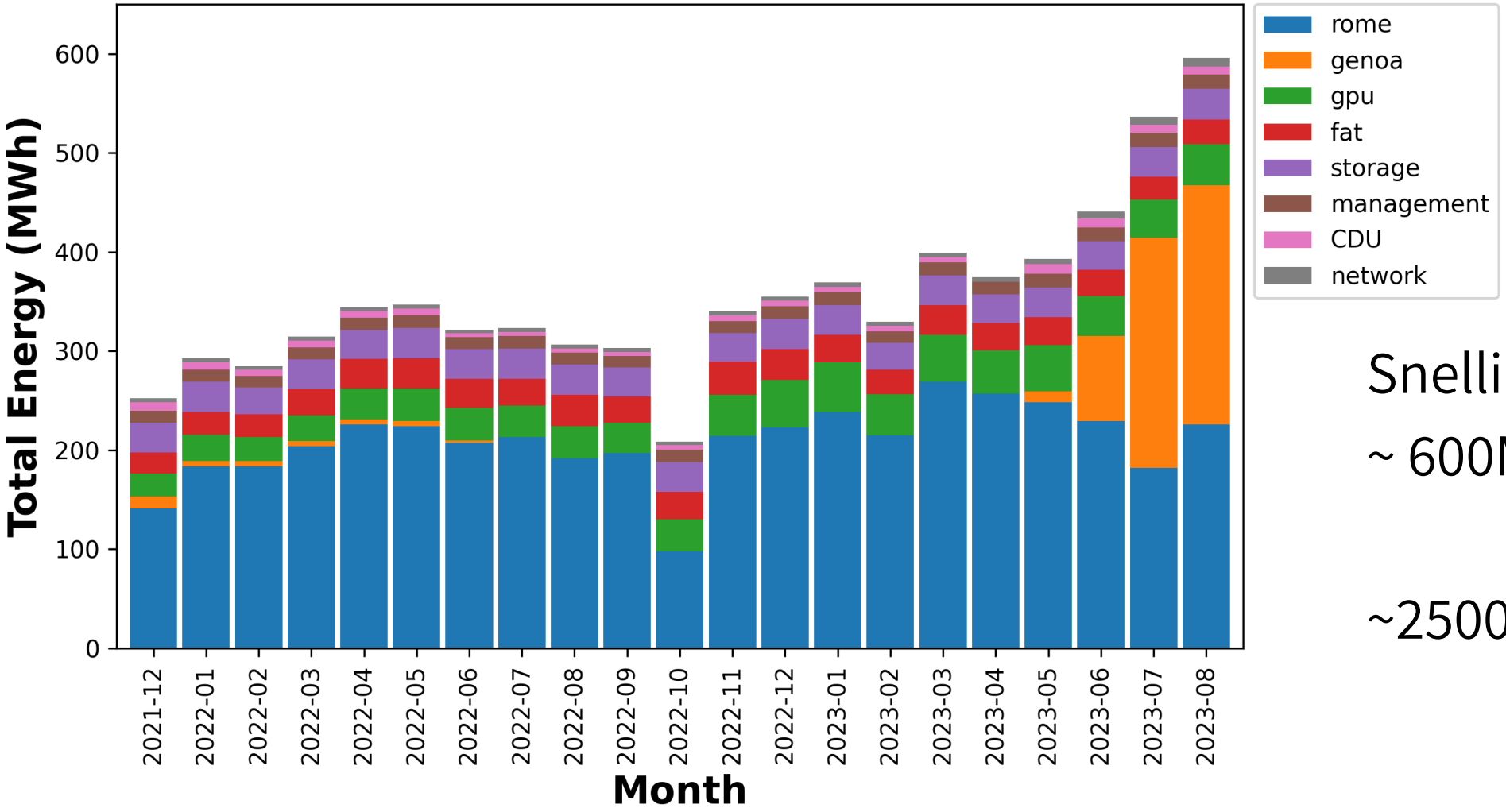


Snellius

(#374) CPU: LINPACK Rmax (0.5 MW) 2.13 PFlop/s

(#176) GPU: LINPACK Rmax (0.13 MW) 3.60 PFlop/s

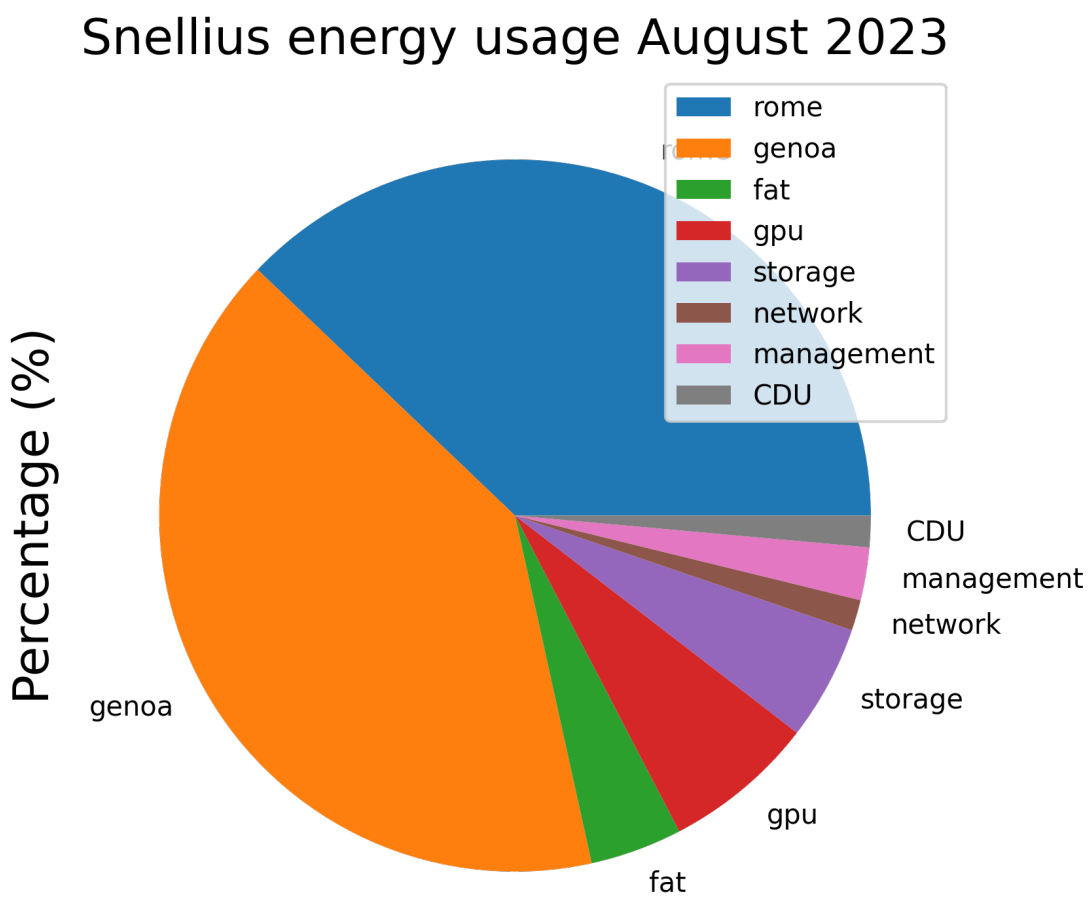
Snellius - Total Energy Usage



Snellius in August used ~ 600MWh of energy.

~2500 house holds worth

Snellius - Energy usage per partition

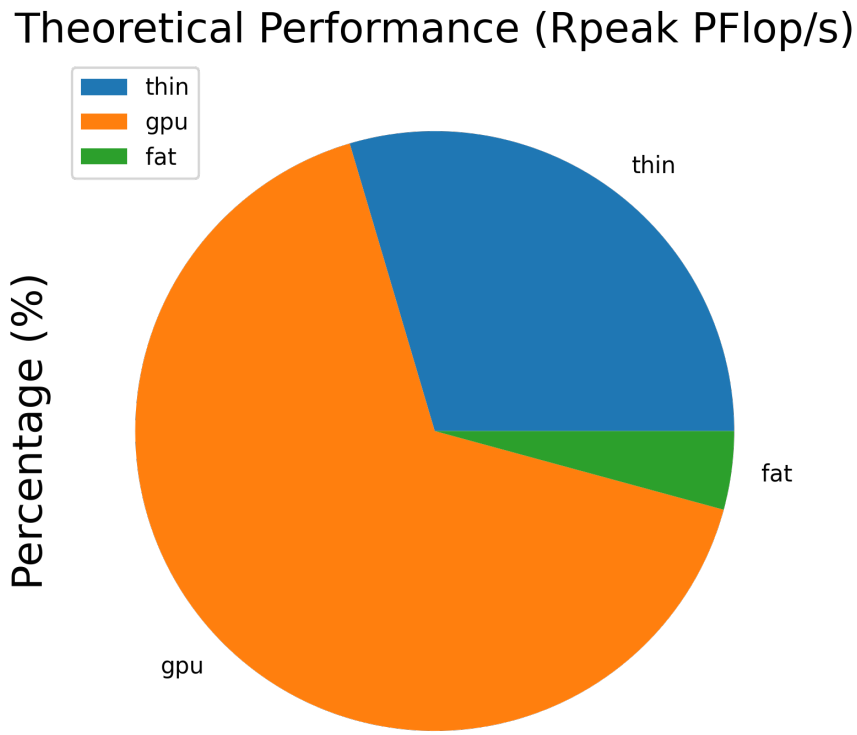


- 90% Compute
  - 80% CPU
    - 39% Rome
    - 41% Genoa
  - 7% Gpu
  - 3% Fat

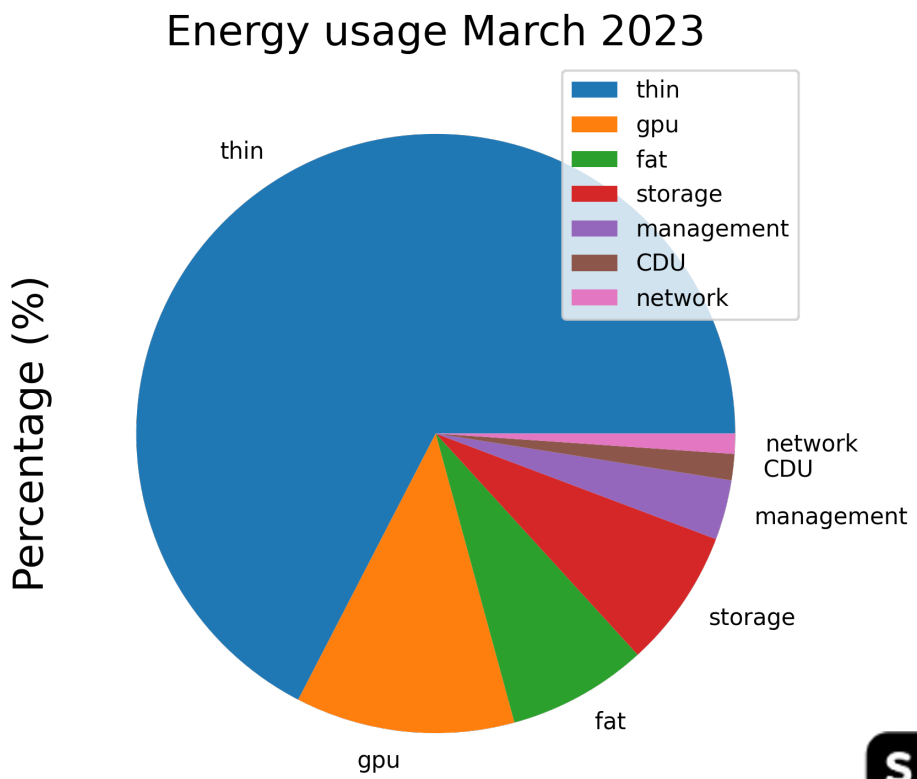
- 10% Other
  - 5.0% Storage
  - 2.4% Management
  - 1.5% Cooling
  - 1.1% Network

Performance vs Energy

- **Thin** (Rpeak) **– 2.86 PFlops**
- **Fat** (Rpeak) **– 0.38 PFlops**
- **GPU** (Rpeak) **– 6.01 Pflops**

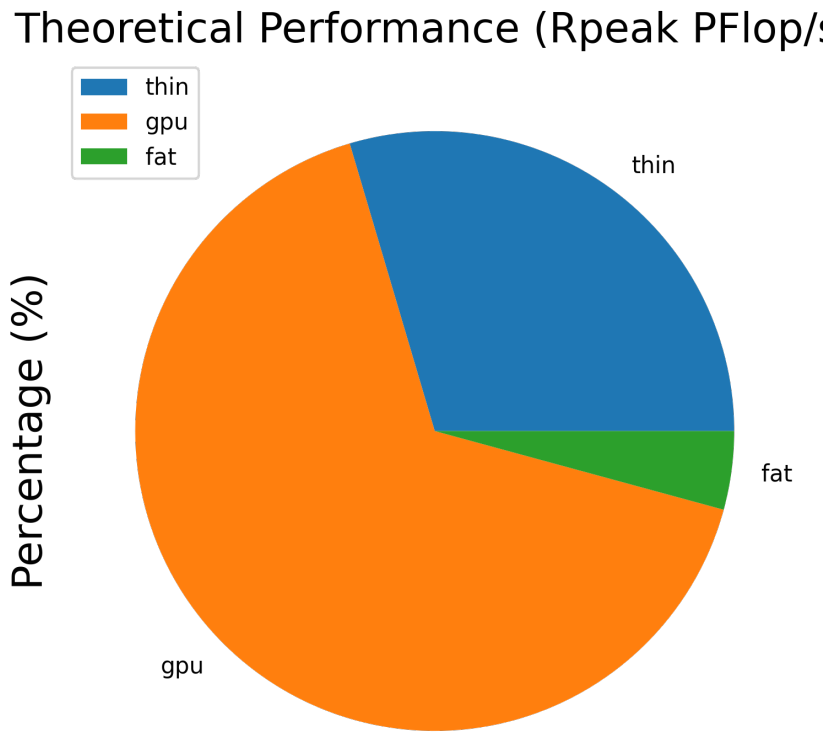


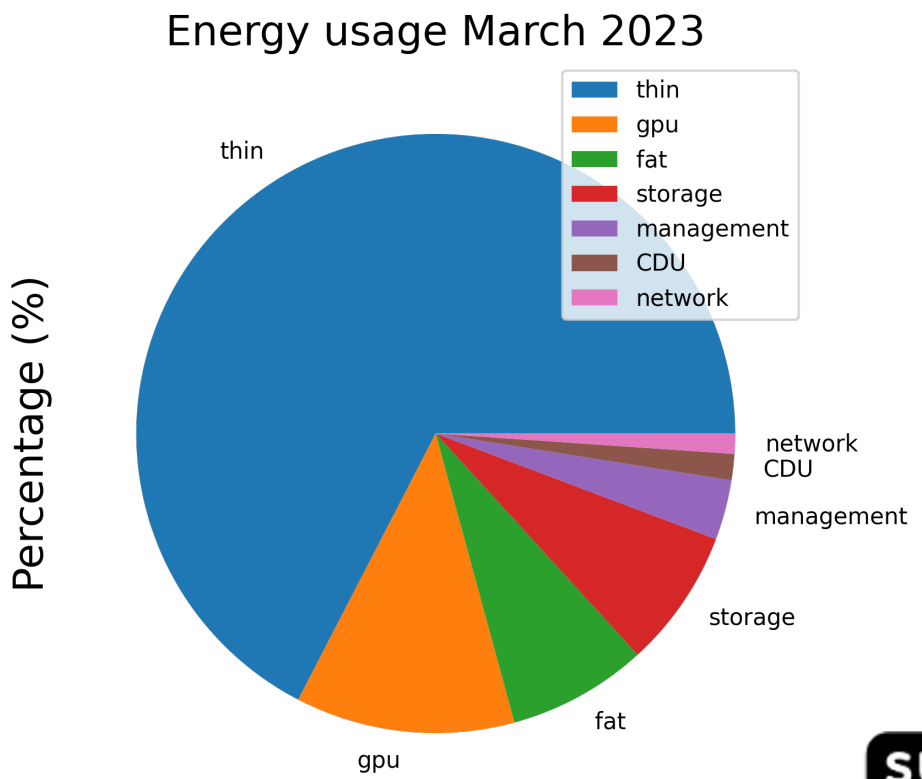
- **Thin – 269 MWh**
- **Fat – 30 MWh**
- **GPU – 47 MWh**



Performance vs Energy

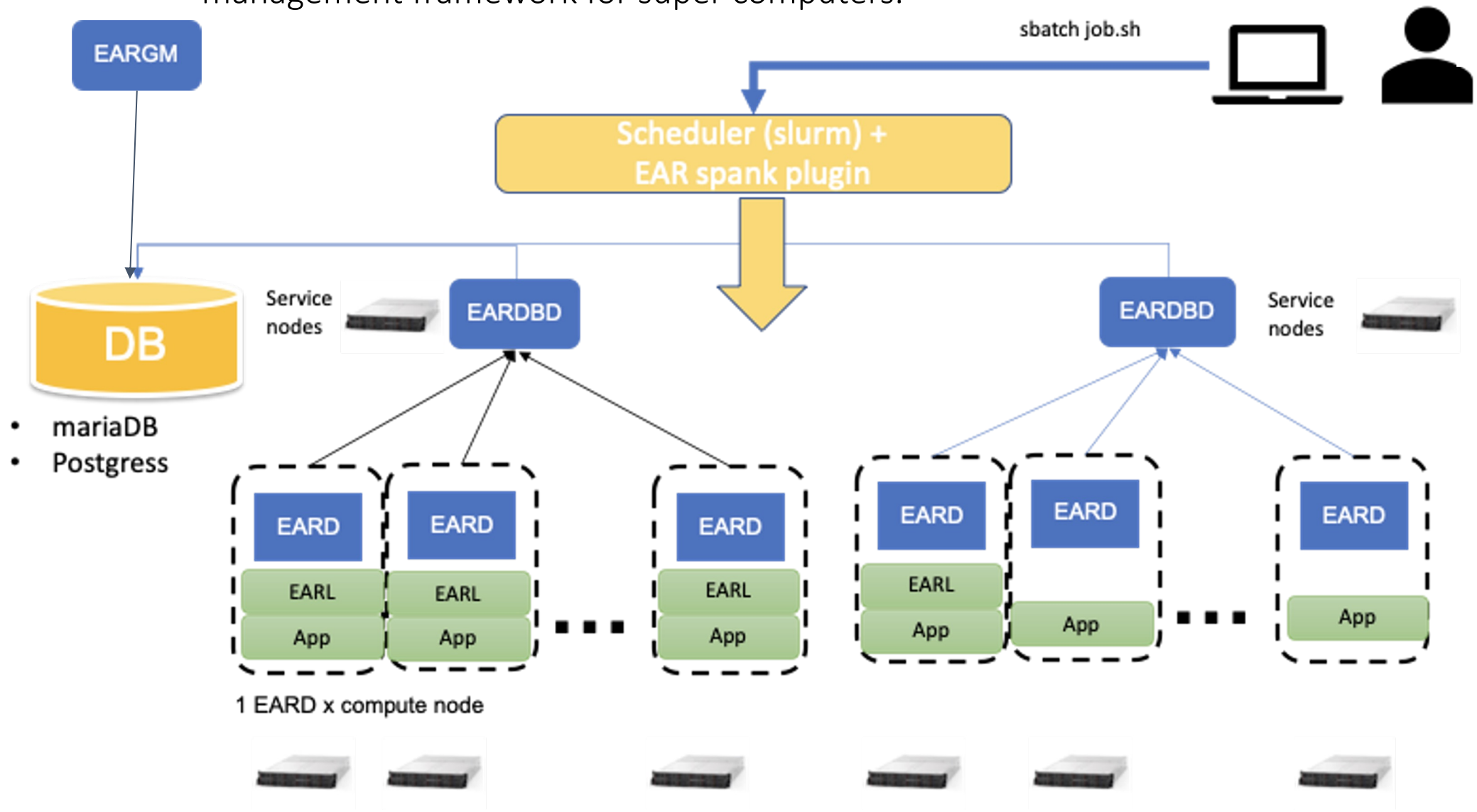
GPU ~ same compute for 1/5 the energy (March 2023)





Energy Aware Runtime (EAR)

Energy Aware Runtime (EAR) package provides an energy management framework for super computers.



Energy Aware Runtime (EAR)

Energy Aware Runtime (EAR) package provides an energy management framework for super computers.

- **EAR Node Manager (EARD) “Daemon”**
  - Energy metrics via the RAPL (Running Average Power Limit) function
  - Global energy limits or just offer global cluster monitoring
- **EARL** is a library that is loaded next to the application
  - Offers application metrics monitoring
  - Can select the frequencies based on the application behavior on the fly.
  - Integrated with SLURM on Snellius.
  - Intercepts the MPI symbols through the PMPI interface to provide “traces” of MPI applications.

```
#!/bin/bash

#SBATCH --ntasks=256
#SBATCH --time=24:0:0
#SBATCH -p thin --exclusive

#SBATCH --ear=on
#SBATCH --ear-policy=monitoring

module load 2021
module load foss/2021a
module load CMake/3.20.1-GCCcore-10.3.0
module load VASP6/6.2.1-foss-2021a-CUDA-11.3.1

srun vasp_std
```

SURF

Energy Aware Runtime (EAR)

Energy Aware Runtime (EAR) package provides an energy management framework for super computers.

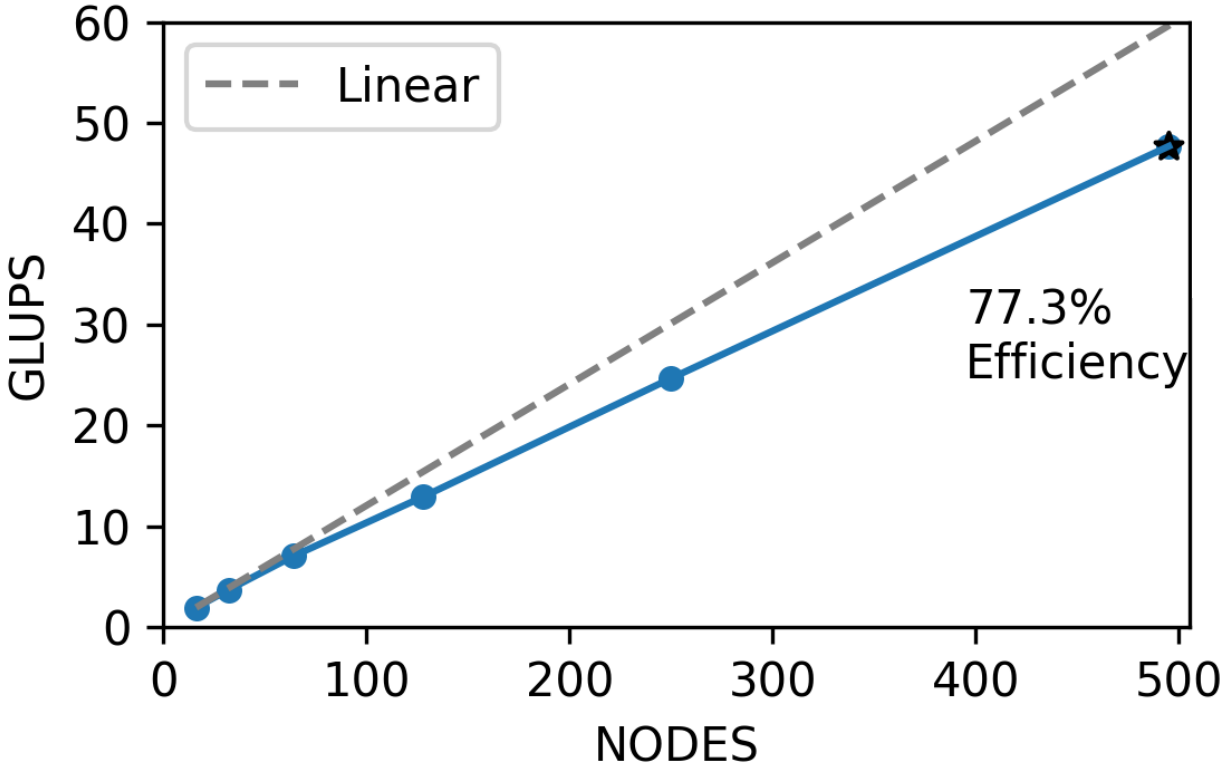
- **EAR Node Manager (EARD) “Daemon”**
  - Energy metrics via the RAPL (Running Average Power Limit) function
  - Global energy limits or just offer global cluster monitoring
- **EARL** is a library that is loaded next to the application
  - Offers application metrics monitoring
  - Can select the frequencies based on the application behavior on the fly.
  - Integrated with SLURM on Snellius.
  - Intercepts the MPI symbols through the PMPI interface to provide “traces” of MPI applications.

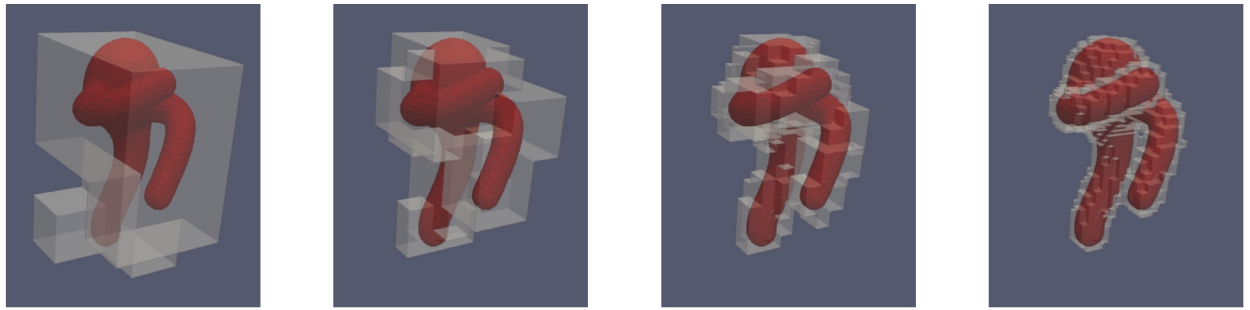
```
#!/bin/bash

#SBATCH --ntasks=256
#SBATCH --time=24:0:0
#SBATCH -p thin --exclusive

#SBATCH --ear=on
#SBATCH --ear-policy=monitoring

module load 2021
module load foss/2021a
module load CMake/3.20.1-GCCcore-10.3.0
module load VASP6/6.2.1-foss-2021a-CUDA-11.3.1

srun vasp_std
```

SURF

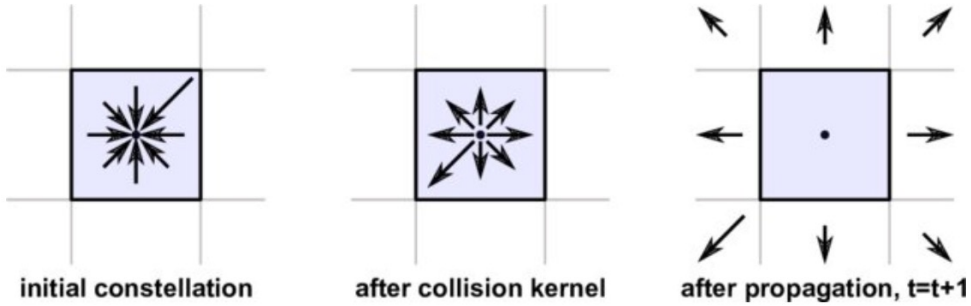
Palabos: Lattice-Boltzmann Solver

Strong Scaling Benchmark





J. Latt, O. Malaspinas, D. Kontaxakis et al. / Computers and Mathematics with Applications 81 (2021) 334–350



SURF

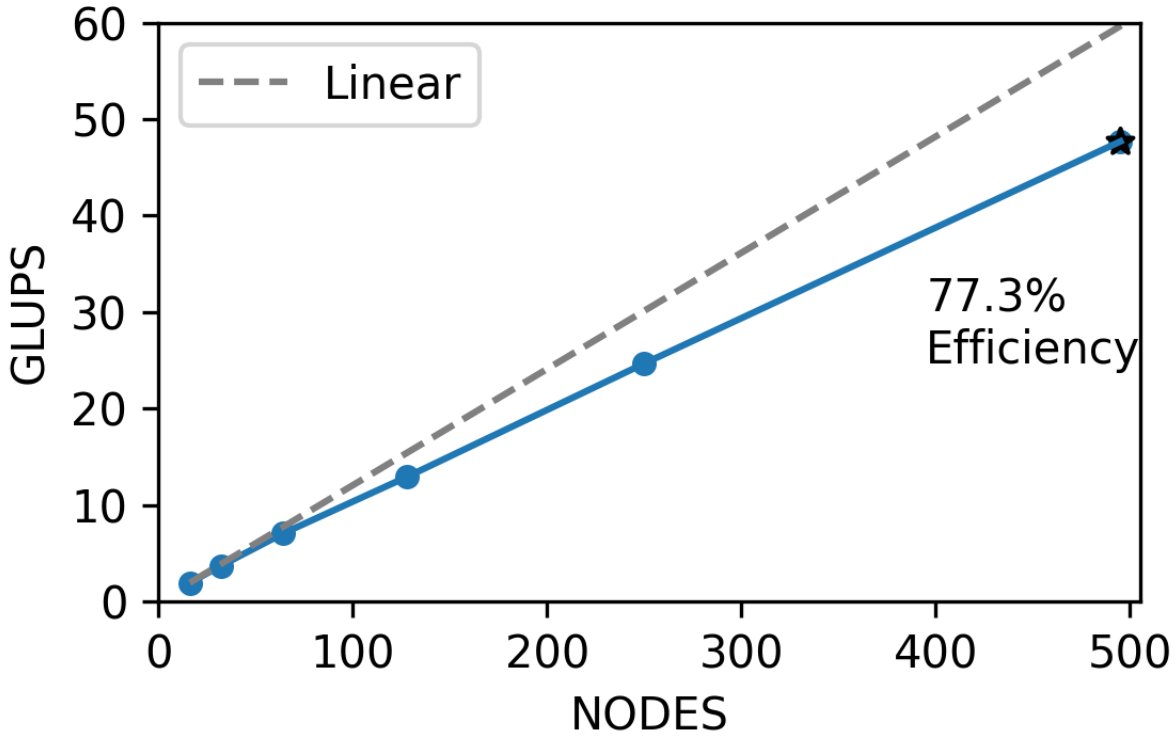
CompBioMed

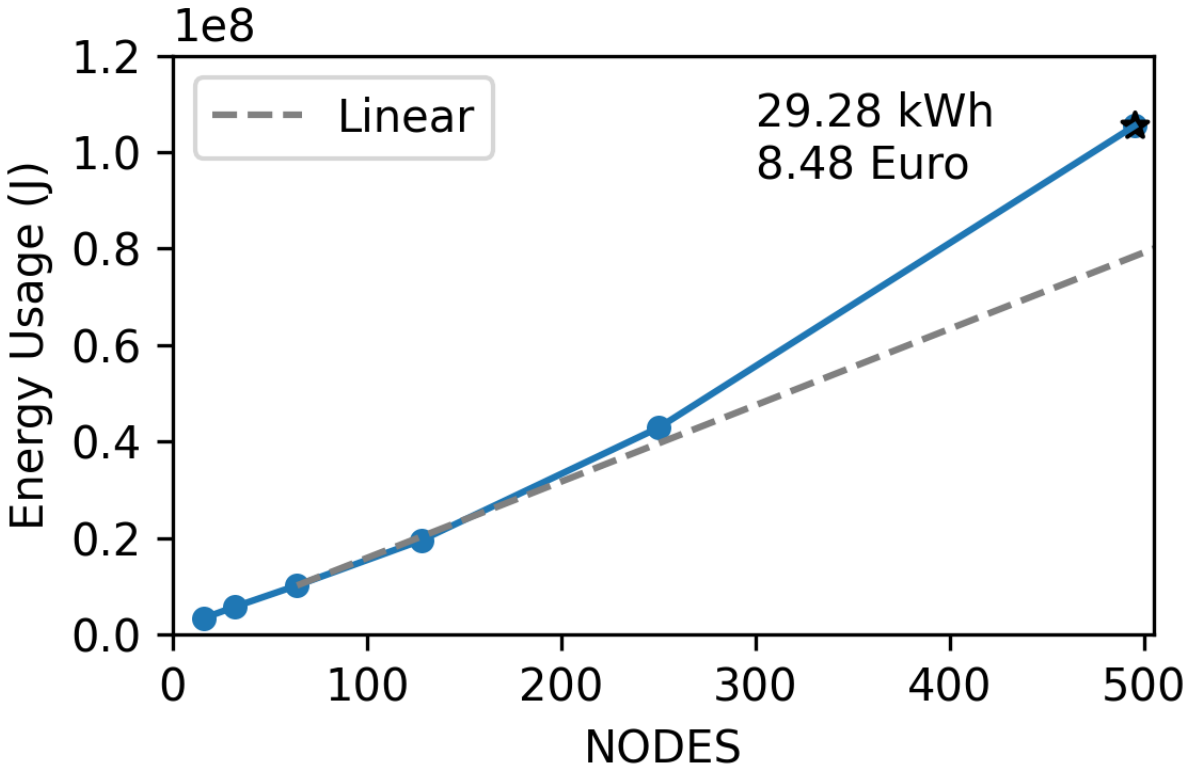
UNIVERSITÉ DE GENÈVE

PALABOS

Palabos: Lattice-Boltzmann Solver

Strong Scaling Benchmark





- An average electric car consumes about 0.2 kwh/km
- 495 node case (which ran for 9 minutes) just drove a car to Leeuwarden (149km from SP)!

Palabos:

Strong Scaling Benchmark

- Per-node, Per-iteration ”traces”

- Node Power
- Avg CPU Freq (node)
- Main memory BW (GB/s)
- CPI (Cycles per instruction)
- MPI% (percentage spent in MPI calls)
- I/O (network communication)