Jyrki Savolainen | CSC – IT Center for Science Ltd.

Outline

LUMI

1. Why supercomputing and LUMI?
2. Getting started with LUMI: Working principles & available tools
3. CSC's expert support

What is supercomputing / HPC?

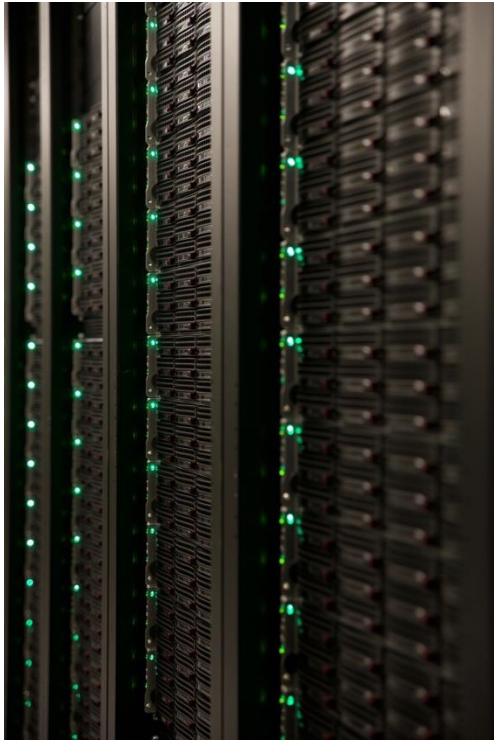
LUMI

A large cluster of computers (”nodes”) working together to solve a numerical task

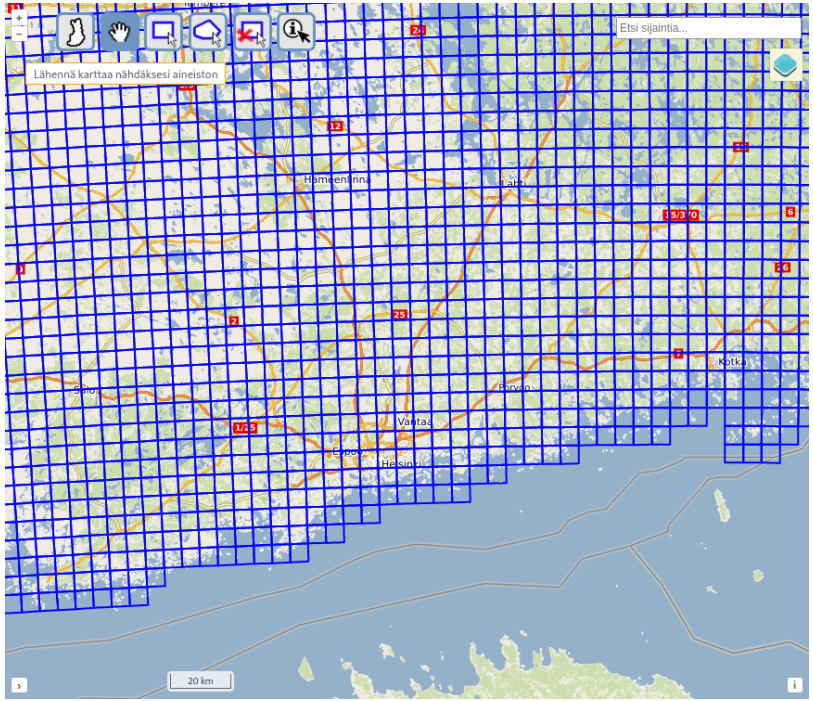
1. Central and Graphics Processing Units (CPU/GPU)
2. Large memory and storage space
3. Fast connection between nodes

A heavy computational task is divided into smaller subtasks – allows one to leverage *parallelism* to reach a solution faster

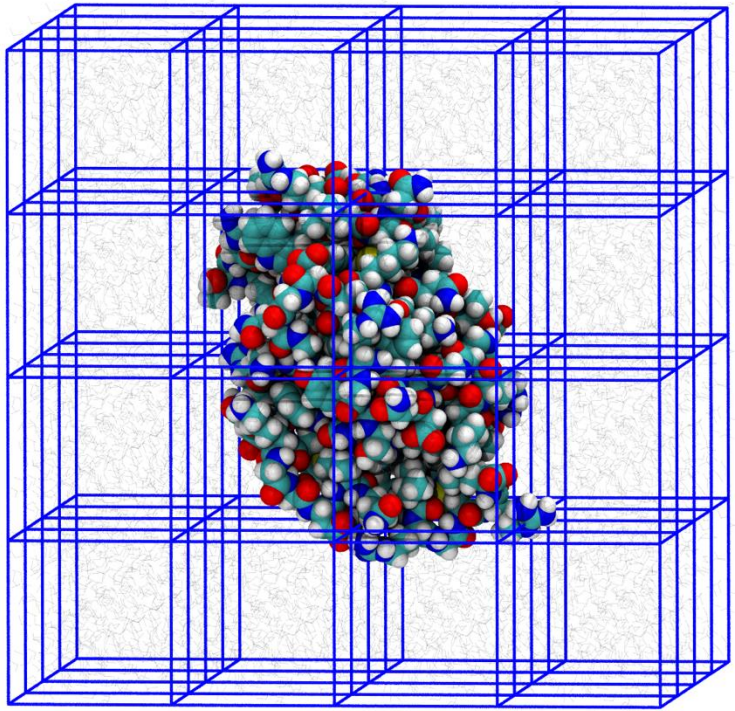
- Memory- and storage-intensive tasks are also common HPC use cases



Examples of parallelism



Analyzing spatial data (2D)



Molecular dynamics simulations (3D)

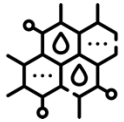
Examples of research areas and software* powered by supercomputers

LUMI

Artificial intelligence

PyTorch, TensorFlow

- Large language models
- Computer vision



Chemistry and materials science

CP2K, LAMMPS

- Catalysts for green hydrogen
- Solar cell and battery materials

Engineering

Elmer, OpenFOAM

- Computational fluid dynamics
- Structural mechanics

Life sciences

GROMACS, OpenMM

- Drug discovery
- Atomistic properties of biosystems

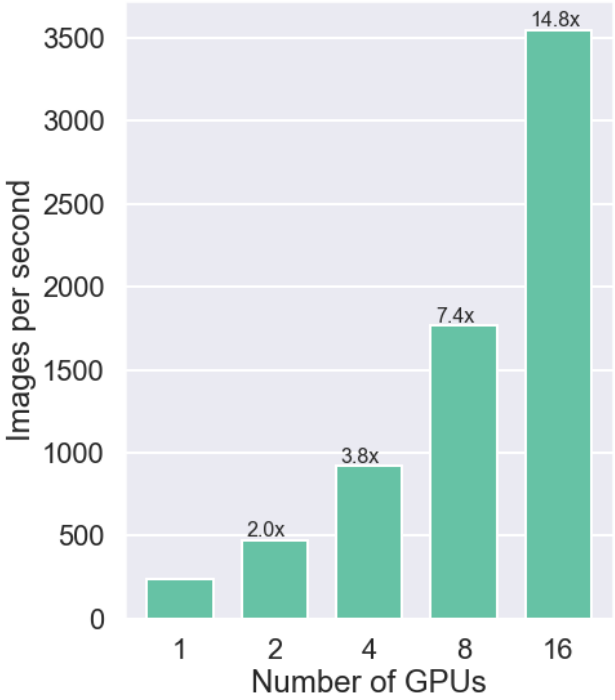
**All available for LUMI!*

Supercomputers for deep learning

LUMI

All our supercomputers have GPU acceleration

- Resources for parallel training of (Deep Learning) models and their inference
- Multi-GPU and multi-node computing support
- Radeon Open Compute (ROCm) versions of standard AI/ML frameworks available on LUMI
  - **PyTorch**, **TensorFlow**, **JAX**, etc.



A computational task does not have to be massive to use a supercomputer

LUMI

Perhaps you just have a large number of smaller experiments

- Run them all at the same time!
- Example of *trivial parallelism:* **pre-processing of images**

  ➔ No limit for the scalability

Other reasons why to use a supercomputer:

- **"Outsourcing"** heavy/specialized computations to a remote server
- Prebuilt environment (software and libraries)
- Possibility to collaborate (public-private co-innovation)
- **CSC's expert support**

Getting started with LUMI: Working principles & available tools

Programming skills?

LUMI

- A common **misconception**:

~~*"Advanced programming skills are required to use a supercomputer"*~~

- Can be useful, but **required only if you intend to write your own code!**

```c
#include <stdio.h>
#include <mpi.h>

int main(int argc, char *argv[]){
    int myrank, nprocs;

    MPI_Init(&argc, &argv);
    MPI_Comm_size(MPI_COMM_WORLD, &nprocs);
    MPI_Comm_rank(MPI_COMM_WORLD, &myrank);

    printf("This is process %d out of %d\n", myrank, nprocs);

    MPI_Finalize();
    return 0;
}
```

What skills are *actually* needed?

LUMI

- Supercomputers run Linux

  → **Basic Linux commands and shell scripting**

- A supercomputer is a remote server that one needs to connect to

  → **SSH via terminal or clients like PuTTY for Windows**

  → **Also possible via www-interface!**

- Supercomputers are shared among hundreds of simultaneous users with different software and resource needs

  → **Understanding batch jobs and resource management**

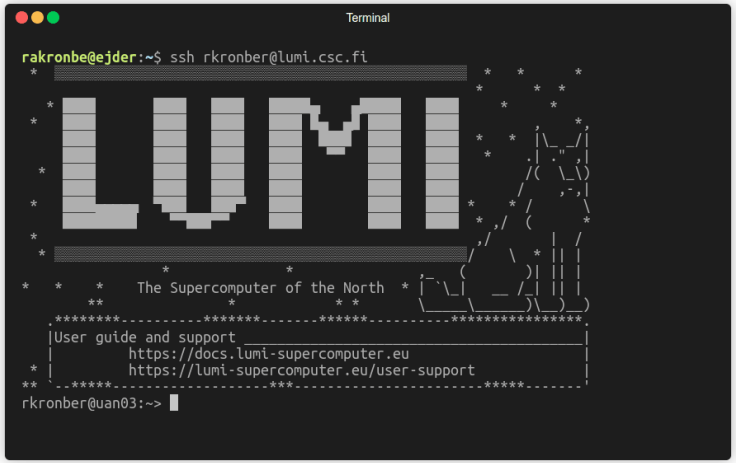
Connecting to LUMI

Logging in to LUMI using an SSH terminal client and SSH key pair:

1. User generates an SSH key pair on their workstation and adds the public key to the my.csc.fi customer portal
2. A connection to LUMI is then opened by running `ssh` -command in a terminal program:

```
$ ssh <username>@lumi.csc.fi
```

Alternatively, LUMI can also be accessed simply using a browser at https://www.lumi.csc.fi



Migrating your data to LUMI

Most simple option is to use the LUMI web interface (drag&drop)

- GUI, but max. size of individual file upload is 10 GiB

For larger files (> 10 GiB) command-line tools are more flexible and powerful

- `rsync`, `scp`

Available storage areas

- **LUMI-P:**
  - User home (20 GiB, 100k files)
  - Project persistent (50 GiB, 100k files)
  - Project scratch (50 TiB, 2000k files)
- **LUMI-F:**
  - Project fast (2 TiB, 100 files)
- **LUMI-O:**
  - Object storage (150 TiB)
- Only project members can access project directories

Software on LUMI are available as modules

The module system allows tailoring the LUMI computing environment for different user requirements

- Typically multiple versions of software available
- Ability to swap easily between programming environments (Cray, GNU)
- Example: List available CP2K(*) modules and load a specific one:

```
$ module use /appl/local/csc/modulefiles
$ module avail cp2k

------------------ /appl/local/csc/modulefiles ------------------
cp2k/2023.1        cp2k/2023.2            cp2k/2024.1 (D)
cp2k/2023.1-gpu    cp2k/2023.2-gpu        cp2k/2024.1-gpu

$ module load cp2k/2024.1-gpu
```

(*) = molecular dynamics software package

Computations on LUMI are run through a batch job system

LUMI

LUMI is a shared research instrument, Slurm batch job system ensures a fair allocation of resources to users

- A batch script ("recipe") is prepared and submitted to a queue
  - Specifies required resources, modules and program launch commands
  - Queuing time depends on requested resources and priority
- **You're only billed for the resources you reserve!**
  - No "hidden" costs or accidental use of CPUh / GPUh

Batch jobs are well-suited for R&D workloads, that are non-time-critical tasks

What does it all look like?

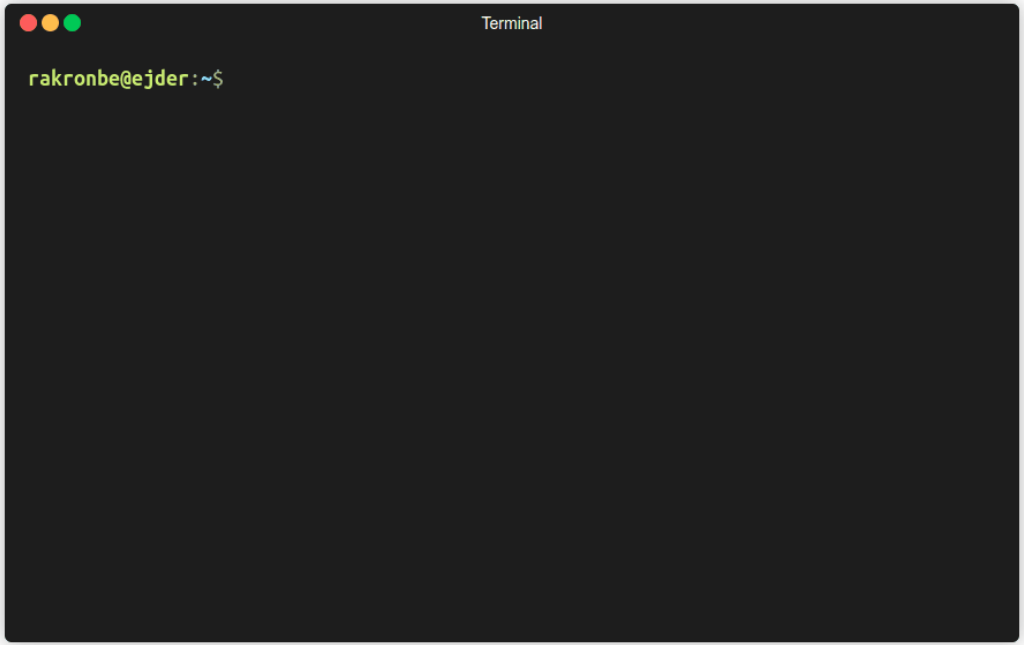
LUMI

Connect to a login node

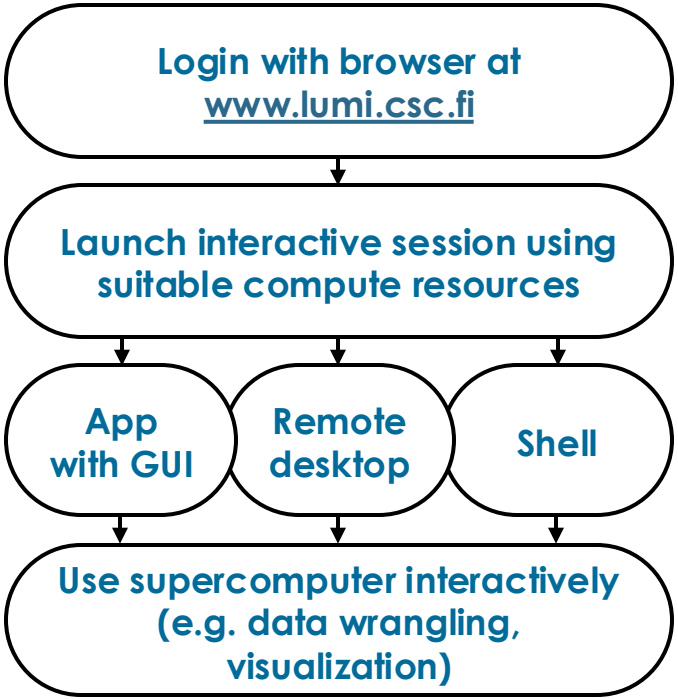
Customize batch job template (specify resources, load modules, run program)

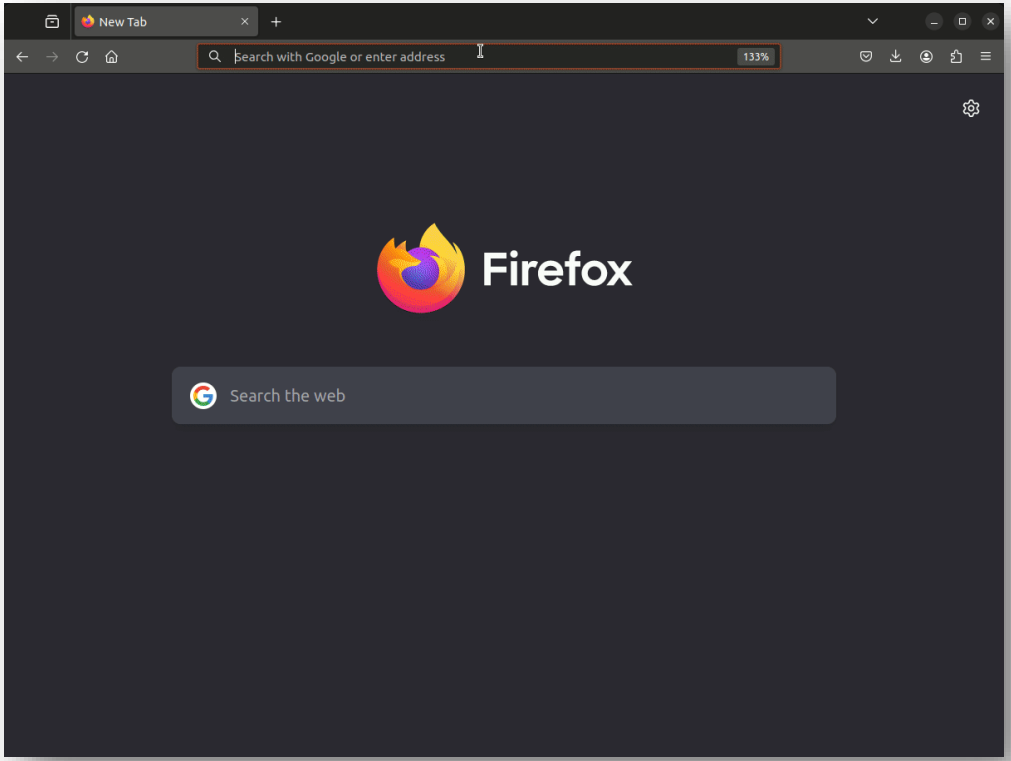
Submit job from the login node to a compute node

View and analyze results on supercomputer or locally



Alternatively: Web interface





CSC's expert support

Support channels

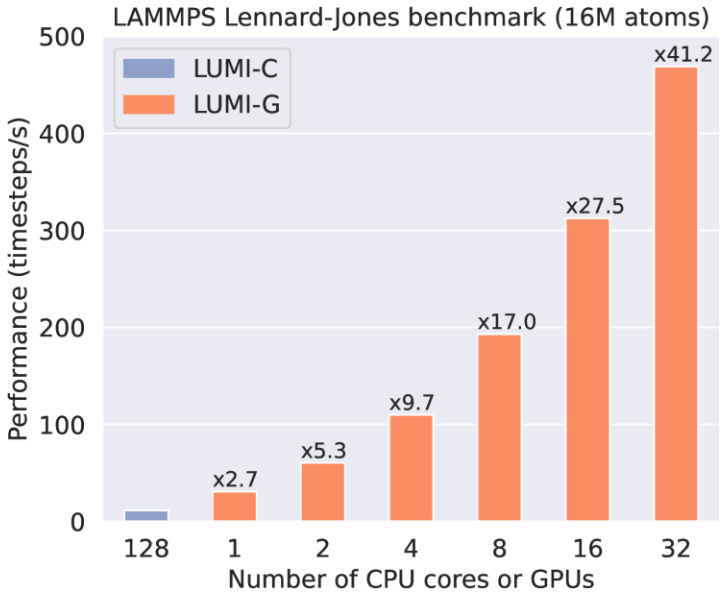
LUMI

- **LUMI Documentation** (docs.lumi-supercomputer.eu)
  - How to use LUMI
  - Available base software stack, instructions for installing own software, batch script examples, etc.
- **CSC Documentation** (docs.csc.fi)
  - Additional local software stack documentation, examples and tutorials
- **Science and user specific help pages** (research.csc.fi)
  - research.csc.fi/geocomputing
  - research.csc.fi/commercial-use
- **High-quality training and expert support provided by CSC and the LUMI User Support Team**

CSC provides expert support in using LUMI

CSC supports users in selecting and applying HPC simulation software and designing efficient workflows for LUMI

- *Solving technical and domain-specific issues*
- *Software installation support*
- *Code and batch script optimization*
- *Parallelization support*
- ***Try&Buy onboarding support***



Thinking big – how to scale up efficiently?

Contact us at servicedesk@csc.fi

- Or join our **weekly user support sessions** every Wednesday at 14 (online)

High-quality customer training

LUMI

CSC and LUMI User Support Team organize regular user training events

- On-site courses, webinars, hackathons, self-learning
- Collaboration with vendors and other European HPC centers and projects (e.g. HPE, AMD, EuroCC, BioExcel, ...)
- csc.fi/en/trainings & lumi-supercomputer.eu/events

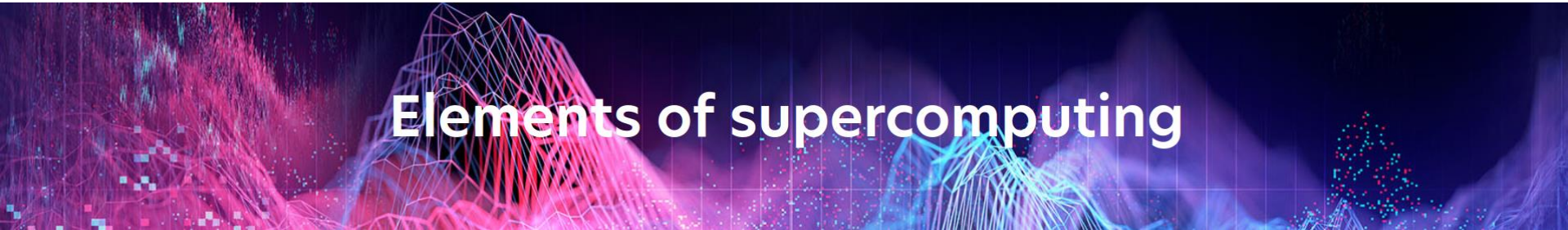
Wide range of topics covering methods and different aspects of HPC and cloud computing, including data management

- Domain-specific courses, using the HPC environment, parallel/GPU programming, application porting and optimization for GPU systems

Tailored trainings to match specific needs and competence gaps!

Customer training and events

- **Elements of Supercomputing** (self-learning, collaboration with KAMK)
  - edukamu.fi/elements-of-supercomputing
- **Courses and registration information organized by CSC and LUMI:**
  - csc.fi/en/trainings/training-calendar
  - www.lumi-supercomputer.eu/events
- **Training and events organized by EuroCC Finland and other NCC countries:**
  - https://www.eurocc-finland.fi/



Summary

LUMI

Supercomputers are clusters of powerful processing units that solve numerically heavy tasks in parallel and/or using a lot of memory

- Parallelism enables a faster time-to-solution
- LUMI has a large number of AMD GPUs suitable for e.g. AI and simulation workloads
- Using a supercomputer does not require advanced programming skills, although Linux experience is helpful
- Extensive documentation, expert support and training available for a swift onboarding!

More about Supercomputing for Business

CSC / Industry web pages

- https://csc.fi/en/about-us/customers/high-performance-computing-for-companies/

EuroCC Finland

- **Web pages** (events, training, success stories): https://www.eurocc-finland.fi/
- **Follow us on LinkedIN:** www.linkedin.com/company/eurocc-finland/