

LUMI Computing Environment and Support

Maciej Szpindler

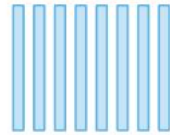
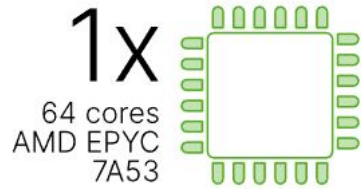
LUMI User Support Team

Compute environment

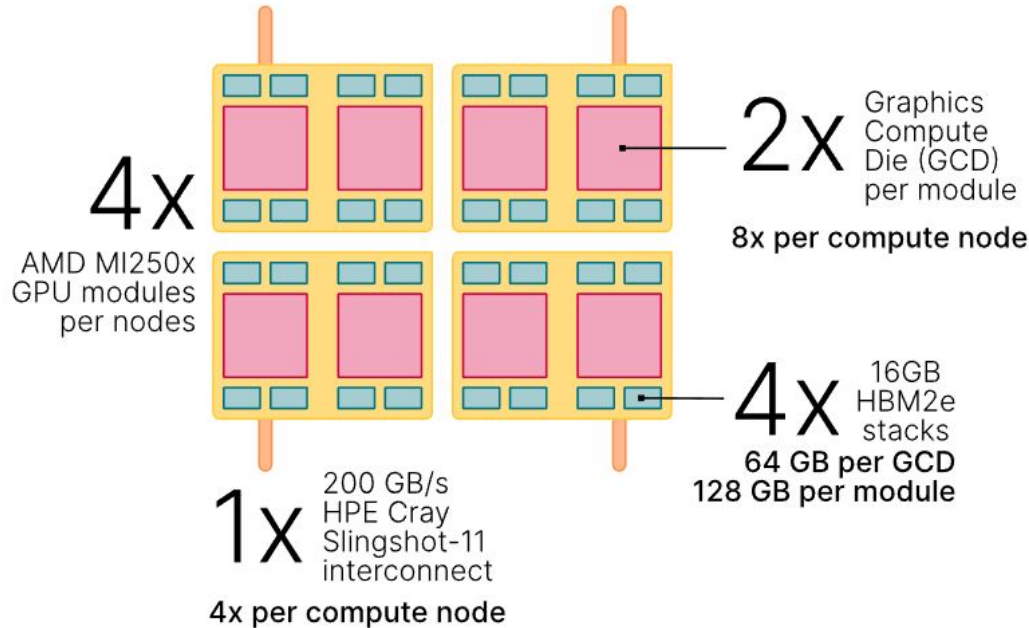
- Hardware overview
- User environment
- Software on LUMI
- Development
- Storage

GPU nodes - LUMI-G

2978x compute nodes



8x 64 GB
DDR4
memory
512 GB total



LUMI-G compute node

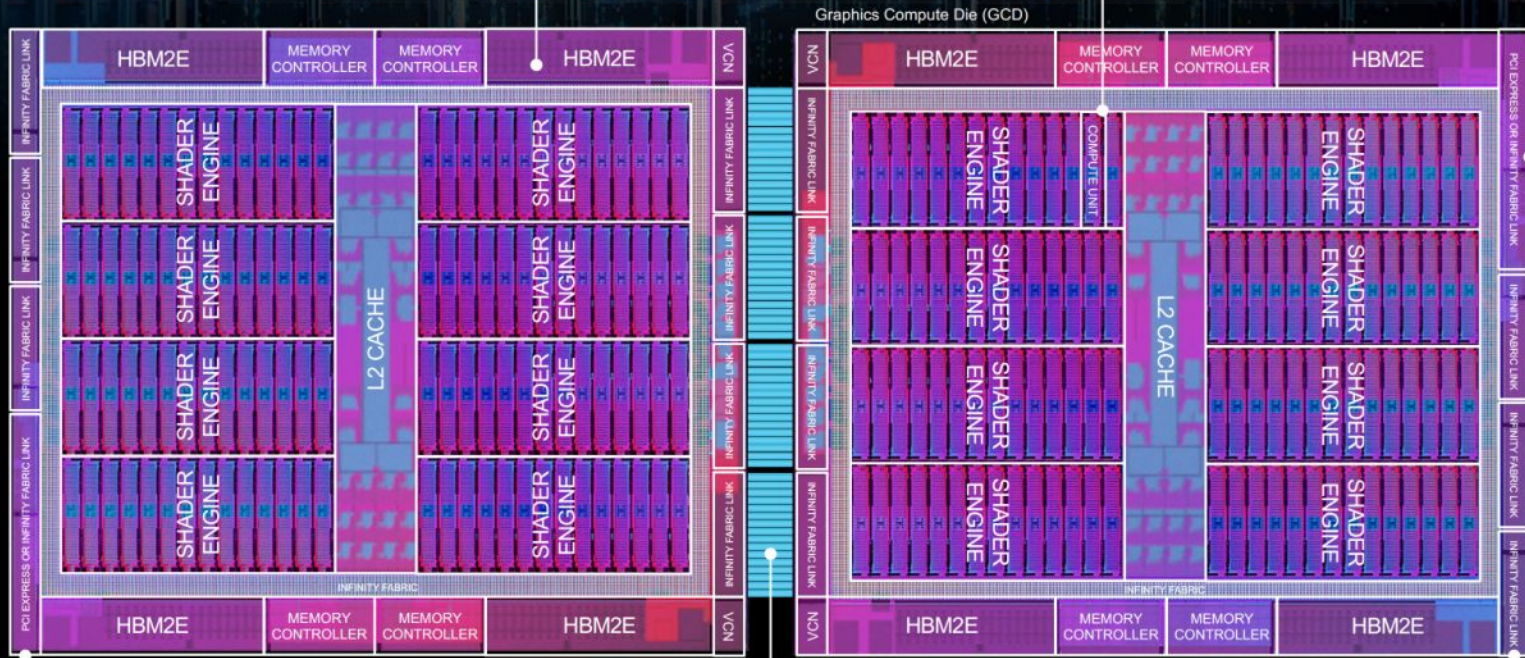
- **4 AMD MI250X GPUs** based on the 2nd Gen AMD CDNA architecture
 - MI250x GPU is a multi-chip module with **two Graphics Compute Dies (GCD)**
 - **64 GB slice of HBM** memory per **110 compute units (CU)**, 220 CUs and 128 GB total memory per module
 - LUMI-G nodes can be considered as **8 GPUs** nodes
- **64-core AMD EPYC 7A53 “Trento” CPU**
 - “Zen 3” compute cores AVX2 256-bit instructions, maximum throughput of 16 double precision FLOP/s
 - 32 KiB of private L1 cache, a 32 KiB instruction cache, and 512 KiB of L2 cache per core
 - 32 MiB L3 cache shared between the group of eight cores
 - 256 MiB of L3 cache per processor.
 - CPU is configured as **4 NUMA nodes (NPS₄)**
- 128 GiB of DDR4 memory per NUMA node for a total of **512 GiB CPU memory**

MI250X MCM

58B Transistors in 6nm

128 GB HBM2e
3.2 TB/s

220 Compute Units
880 Matrix Cores



Coherent CPU-GPU Memory
3RD Gen Infinity Architecture
144 GB/s

Scale Out

PCIe Gen4 ESM
100 GB/s



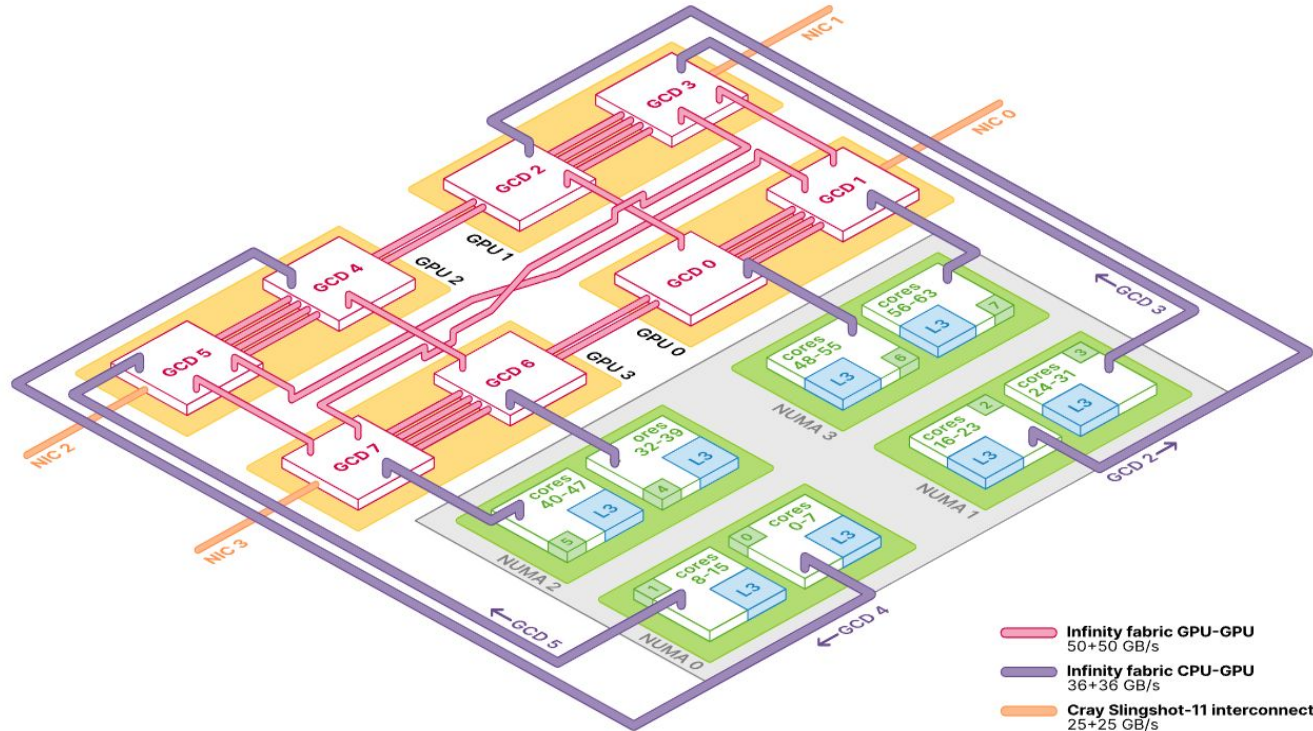
In-package
Infinity Fabric
400 GB/s



Scale Up

External Infinity Fabric
500 GB/s

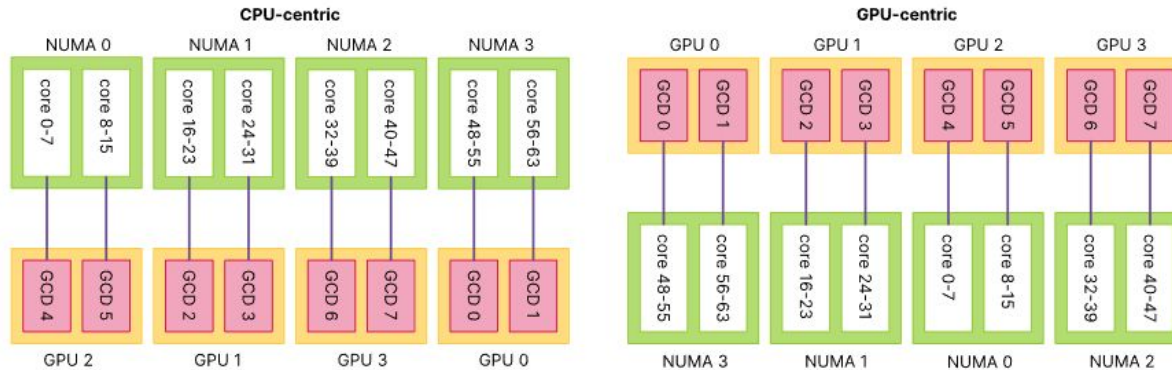
GPU-centric HW design



- 5 links per GCD
- Two GCDs in the module use in-package “Infinity Fabric” interface with theoretical peak bidirectional bandwidth of up to **400 GB/s**
- GCDs on different modules are linked with either a single or double “Infinity Fabric” link, with a peak bidirectional bandwidth of **100 GB/s** and **200 GB/s**, respectively
- MI250x module is directly connected to the slingshot 11 network providing up to up to **25+25 GB/s** peak bandwidth

NUMA binding

Proper binding the NUMA node to the GPU is crucial for achieving optimal performance



Source <https://docs.lumi-supercomputer.eu/hardware/lumig/>

CPU nodes - LUMI-C

- **2x AMD EPYC 7763 CPUs** with 64 cores each running at 2.45 GHz
 - total of **128 cores per node**
 - 2-way simultaneous multithreading (SMT) , up to 256 threads per node
 - "Zen 3" compute cores, AVX2 256-bit vector instructions for a maximum throughput of 16 double precision FLOP/clock (AVX2 FMA operations)
 - 32 KiB of private L1 cache, a 32 KiB instruction cache, and 512 KiB of L2 cache
 - 8x core complex dies (CCDs) 8 cores each
 - L3 cache is shared between the eight cores of a CCD and has a capacity of 32 MiB for a total 256 MiB of L3 cache per processor

CPU node

- The LUMI-C compute nodes are configured with 4 NUMA zones ("quadrant mode") with 2 CCDs per quadrant

Core Complex Die (CCD)

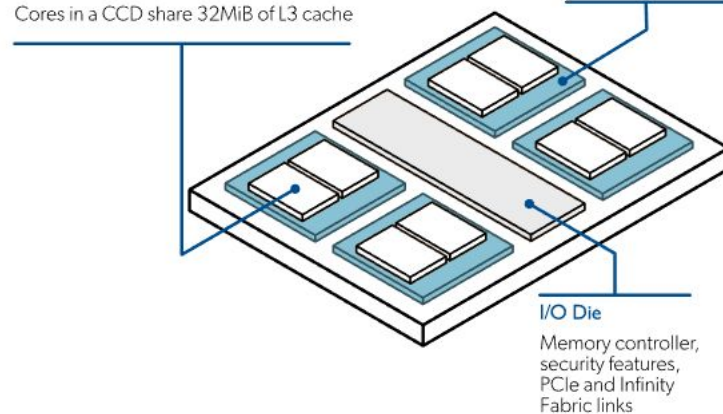
8 CPU cores with two-way SMT

- 32KiB of L1 cache
- 512KiB of L2 cache

Cores in a CCD share 32MiB of L3 cache

NUMA node

4 NUMA nodes per socket, 2 CCDs per NUMA node



		NUMA nodes CPU 1				NUMA nodes CPU 2			
		0	1	2	3	4	5	6	7
NUMA nodes CPU 1	0	10	12	12	12	32	32	32	32
	1	12	10	12	12	32	32	32	32
	2	12	12	10	12	32	32	32	32
	3	12	12	12	10	32	32	32	32
NUMA nodes CPU 2	4	32	32	32	32	10	12	12	12
	5	32	32	32	32	12	10	12	12
	6	32	32	32	32	12	12	10	12
	7	32	32	32	32	12	12	12	10

Network and interconnect

- All LUMI compute nodes uses HPE Cray Slingshot-11 200 Gbps network
- LUMI-C (CPU nodes) are equipped with a single endpoints (NIC)
- LUMI-G nodes (GPU nodes) have 4 endpoints - one for each AMD MI250x GPU module
- Each endpoints provide up to 50 GB/s of bidirectional bandwidth
- The HPE Cray Slingshot NIC features high performance RDMA and hardware acceleration for MPI and SHMEM based software.

User Environment

- **Operating System:** Linux (Cray Operating System)
- **Remote access:** SSH, web interface (remote desktop), cloud interface (kubernetes)
- **Resource management:** SLURM
- **Software:** module environment, software stacks
- **User applications:** build via EasyBuild (and Spack)
- **HPC Containers:** singularity, bring your own container (limited support)
- **Development:** Cray Programming Environment

Software stacks

- **CrayEnv** offers the Cray Programming Environment (PE) and allows one to use it completely in the way intended by HPE-Cray. The environment also offers a limited selection of additional tools, often in updated versions compared to what SUSE Linux, the basis of the Cray Linux environment, offers. If you need a richer environment, you should use our other software stacks.
- **LUMI/yy.mm** is an extensible software stack that is mostly managed through EasyBuild. Each version of the LUMI software stack is based on the version of the Cray Programming Environment with the same version number.
- **Spack**

Software library

LUMI | LUMI Software Library

Q Search

a b c d e f g h i j k l m n o p q r s t u v w x y z home

[package list]

GROMACS

user-installable

License information

GROMACS is Free Software, available under the GNU Lesser General Public License (LGPL) version 2.1. You can redistribute it and/or modify it under the terms of the LGPL as published by the Free Software Foundation, either version 2.1 of the License, or (at your option) any later version.

See the "About" page on the GROMACS web site.

User documentation

A note about the GPU versions.

There exist two different versions of GROMACS for AMD GPUs.

- The authors of GROMACS use SVCL for an implementation for AMD GPUs. These versions are or will be part of the official GROMACS distribution channels.
- AMD has made a big part of the CUDA version. It is not clear though to what extent that branch will see further development as GROMACS evolves.

When testing in early 2023, the HP part offered significantly better performance than the SVCL part.

Check the technical documentation of the EasyConfig to find out on which branch of GROMACS the recipes are based. That documentation is more towards the bottom of the page.

A note about the CPU versions with PLUMED after the March/April 2023 system maintenance/update

After the March/April 2023 system update building PLUMED broke so those easyconfigs were replaced with two versions, one without Python support and one with a different way of enabling support for the `stew-system` modules. Therefore the corresponding EasyConfig of GROMACS have also been replaced.

It is done this way because it is also unclear if the Python support is needed when used with GROMACS. Both versions with and without Python support for PLUMED build.

See also the page on PLUMED.

User-installable modules (and EasyConfigs)

Install with the EasyBuild user module:

```
eb easyconfig --
```

To access module help after installation and get reminded for which stacks and partitions the module is installed, use `module spider GROMACS` or `module search`.

EasyConfig

- EasyConfig GROMACS-2021.4-cpu-Cray-22.08-PLUMED-2.7.4-cpu-python-3.9.12.1-CPU-els, will build GROMACS/2021.4-cpu-Cray-22.08-PLUMED-2.7.4-cpu-python-3.9.12.1-CPU
- EasyConfig GROMACS-2021.4-cpu-Cray-22.08-PLUMED-2.7.4-cpu-python-CPU-els, will build GROMACS/2021.4-cpu-Cray-22.08-PLUMED-2.7.4-cpu-python-CPU
- EasyConfig GROMACS-2021.4-cpu-Cray-22.08-PLUMED-2.7.4-cpu-python-3.9.12.1-CPU-els, will build GROMACS/2021.4-cpu-Cray-22.08-PLUMED-2.7.4-cpu-python-3.9.12.1-CPU
- EasyConfig GROMACS-2021.4-cpu-Cray-22.08-PLUMED-2.8.0-cpu-python-CPU-els, will build GROMACS/2021.4-cpu-Cray-22.08-PLUMED-2.8.0-cpu-python-CPU
- EasyConfig GROMACS-2021.4-cpu-GNU-22.08-PLUMED-2.7.4-cpu-python-3.9.12.1-CPU-els, will build GROMACS/2021.4-cpu-GNU-22.08-PLUMED-2.7.4-cpu-python-3.9.12.1-CPU
- EasyConfig GROMACS-2021.4-cpu-GNU-22.08-PLUMED-2.7.4-cpu-python-CPU-els, will build GROMACS/2021.4-cpu-GNU-22.08-PLUMED-2.7.4-cpu-python-CPU
- EasyConfig GROMACS-2021.4-cpu-GNU-22.08-PLUMED-2.8.0-cpu-python-3.9.12.1-CPU-els, will build GROMACS/2021.4-cpu-GNU-22.08-PLUMED-2.8.0-cpu-python-3.9.12.1-CPU
- EasyConfig GROMACS-2021.4-cpu-GNU-22.08-PLUMED-2.8.0-cpu-python-CPU-els, will build GROMACS/2021.4-cpu-GNU-22.08-PLUMED-2.8.0-cpu-python-CPU
- EasyConfig GROMACS-2021.6-cpu-Cray-22.08-CPU-els, will build GROMACS/2021.6-cpu-Cray-22.08-CPU
- EasyConfig GROMACS-2021.6-cpu-GNU-22.08-CPU-els, will build GROMACS/2021.6-cpu-GNU-22.08-CPU
- EasyConfig GROMACS-2023-dev-cpu-GNU-22.08-NPI-GPU-els, will build GROMACS/2023-dev-cpu-GNU-22.08-NPI-GPU

Technical documentation

Table of contents

License information

User documentation

A note about the GPU versions

A note about the CPU versions with PLUMED after the March/April 2023 system maintenance/update

User-installable modules (and EasyConfig)

Technical documentation

GROMACS and PLUMED

EasyBuild

Version 2020.6 for CPE 21.08

Version 2021.3 for CPE 21.08

Version 2020.6 with PLUMED 2.6.4 for CPE 21.08

Version 2021.5 for CPE 21.12

Version 2021.6 with PLUMED 2.7.4 and 3.0 for CPE 21.12 and later

GROMACS-2022-dev-cpu-GPU

GPU

Archived EasyConfig

<https://lumi-supercomputer.github.io/LUMI-EasyBuild-docs/>

Installing software

- **EasyBuild** - main channel for software provisioning, central installation (Software stack), custom installation (Contributed stack), recipes available in the software library (and GitHub)
- **Spack** - alternative tool for building entire environments
- **Containers** - last resort for complicated, complex environments (such as Conda)

Code development

- Cray Programming Environment (CPE) stack
 - Compiler suites (currently supported): Cray (CCE), GCC, AMD (ROCm)
 - Low-level GPU programming: **HIP** (ROCm), SYCL
 - High-level GPU programming: **OpenMP**, OpenACC (limited support)
 - Compiler wrappers provide unified compiling environment
 - Libraries: CrayMPICH, Cray LibSci (numerical primitives), AMD GPU libraries
- Productivity tools
 - Parallel debuggers
 - Performance Analysis tools

GPU programming

- Code enabling scenarios for GPUs:
 - Translation from CUDA: “hipification”
 - OpenMP GPU Offload (recommended)
 - ML/DL frameworks: PyTorch, TensorFlow, already supported
 - Use portable frameworks: Kokkos, Raja, Alpaka, others (support for AMD GPU is extending)

Storage Environment

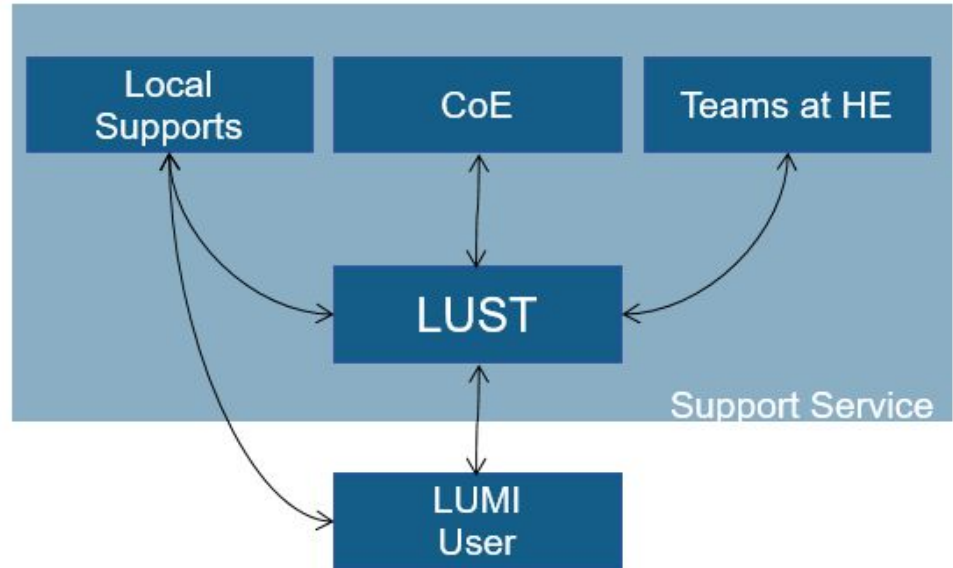
- **LUMI-P:** main storage
 - 4 independent Lustre file systems (20PB each)
 - Aggregate bandwidth of 240 GB/s
 - User and project directories
 - Scratch space
- **LUMI-F:** fast storage
 - Lustre file system with a storage capacity of 7 PB
 - Aggregate bandwidth of 1 740 GB/s
- **LUMI-O:** object storage
 - offers a total of 30 PB storage for storing, sharing, and staging of data

Getting support

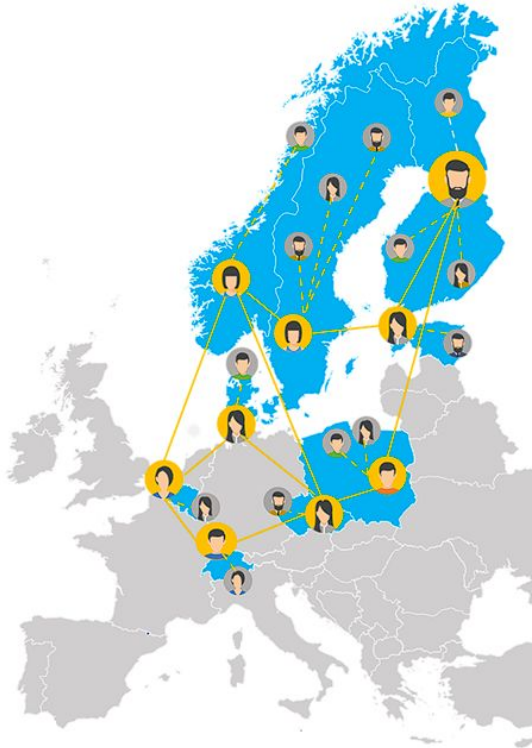
- Help desk
- LUMI user documentation
- Support web pages
- Software stack
- Porting and optimization
- User training
- Application Readiness Program

Support overview

- LUMI user support and a centralized help-desk by the distributed LUMI User Support Team



The LUMI User Support Team



- LUMI User Support Team (LUST) of about 10 HPC experts
- Each LUMI consortium member has at least one representative
- Your local point of contact
- Coordinates support from local experts, LUST and CoE members (AMD and Cray)

Project's support Options

- LUMI Help-desk
- **"Level 3"** support (e.g. application enabling, methodology support) via local centers as well as the EuroHPC Competence Centers
- **"Code Porting"** Call (2024, TBA)
- **"Epicure"** EuroHPC funded application support project

Communication

- Support web pages
 - <https://lumi-supercomputer.eu/user-support/>
- Newsletter
 - <https://lumi-supercomputer.eu/newsletter/>
- FAQ
 - <https://lumi-supercomputer.eu/faq/>
- Help desk
 - <https://lumi-supercomputer.eu/user-support/need-help/>
- LUMI Coffee breaks
 - every last Wednesday of the month on Zoom
 - Ask anything

