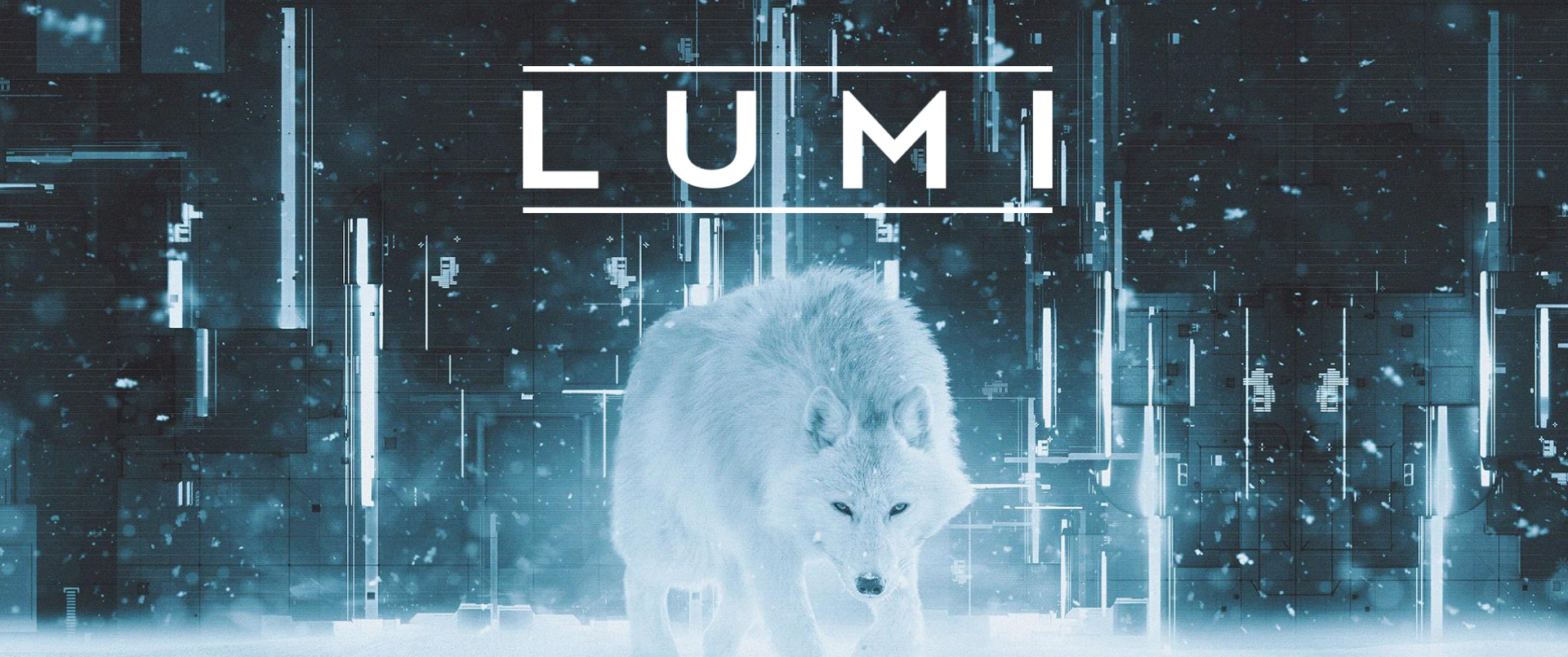


The pan-European supercomputer of the North

Dr. Pekka Manninen
Director, LUMI Leadership Computing Facility
CSC – IT Center for Science, Finland

Adjunct Professor, University of Helsinki

Outline

- The long story of LUMI and EuroHPC (made short)
- The opportunities and benefits for research, development and innovation offered by LUMI
- How to prepare for LUMI?
- Q&A

The EuroHPC Initiative

- The **EuroHPC Joint Undertaking** will pool EU and national resources in high-performance computing (HPC)
  - **acquiring and providing a world-class supercomputing and data infrastructure** for Europe's scientific, industrial and public users
  - supporting an ambitious **research and innovation agenda**
- The EuroHPC declaration has been signed by **32 European countries**
- The first generation of EuroHPC systems announced in June 2019
  - 3 pre-exascale systems to Finland, Italy and Spain
  - 5 petascale systems to Czech Republic, Bulgaria, Luxembourg, Portugal and Slovenia
- Next generations of systems planned for 2023-2024 and 2026-2027

LUMI Consortium

- Unique consortium of 10 countries with strong national HPC centers
- The resources of LUMI will be allocated per the investments
- The share of the EuroHPC JU (50%) will be allocated by a peer-review process (cf. PRACE Tier-0 access) and available for all European researchers
- The shares of the LUMI partner countries will be allocated by local considerations and policies – seen and handled as extensions to national resources

LUMI Datacenter in Kajaani

100% hydroelectric energy up to 200 MW

Very reliable power grid: Only one 2 min outage in 38 years

100% free cooling available, PUE 1.03

Waste heat reuse: effective energy price 35 €/MWh,
negative CO_2 footprint: 13500 tons reduced every year

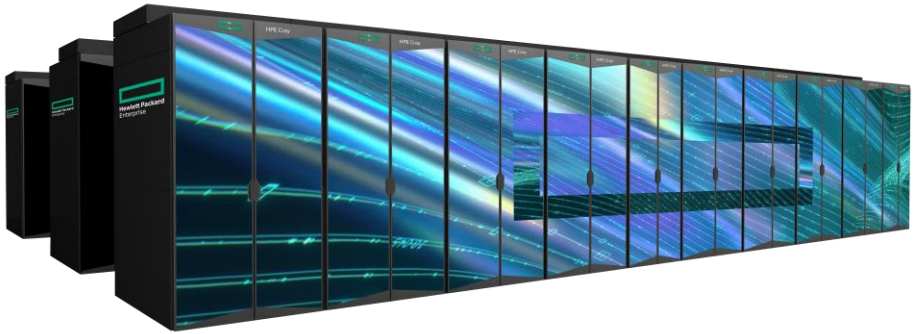
Extreme connectivity: Kajaani DC is a direct part of the Nordic backbone.
4x100 Gbit/s to GÉANT in place, can be easily scaled up to multi-terabit level

Elevated security standards guaranteed by ISO27001 compliancy

Benefits and opportunities for R&I by LUMI

LUMI: one of the fastest supercomputers in the world

LUMI



- LUMI will be an **HPE Cray EX** supercomputer manufactured by **Hewlett Packard Enterprise**
- Peak performance over **550 petaflop/s** makes the system one of the world's fastest
  - Fastest today is Fugaku supercomputer in Japan with 513 petaflop/s, second fastest Summit in USA with 200 petaflop/s)

1 system

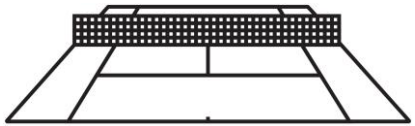
550 Pflop/s

Peak Performance

Computing power equivalent to

1 500 000

Modern laptop computers



Size of a tennis court

Modern platform for

High-performance computing, Artificial intelligence, Data analytics

Based on GPU technology

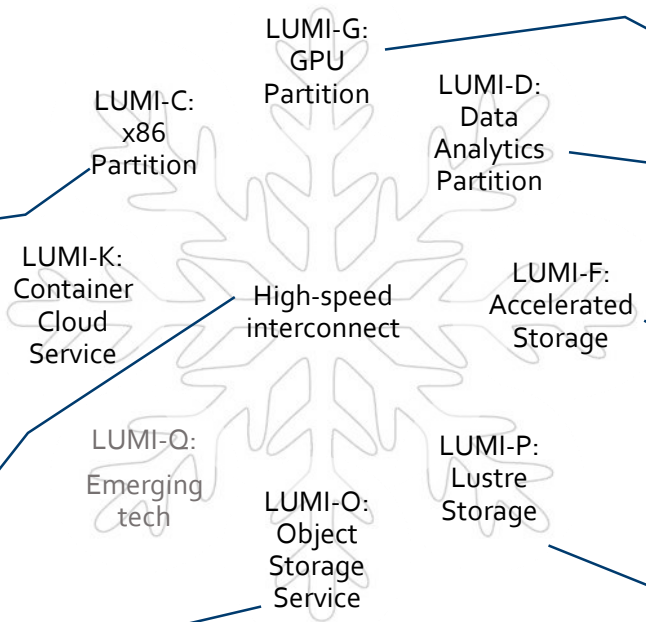
LUMI, the Queen of the North

LUMI is a Tier-0 ***GPU-accelerated supercomputer*** *that enables the convergence of* ***high-performance computing***, ***artificial intelligence***, *and* ***high-performance data analytics.***

- Supplementary CPU partition
- ~200,000 AMD EPYC CPU cores

Possibility for combining different resources within a single run. HPE Slingshot technology.

30 PB encrypted object storage (Ceph) for storing, sharing and staging data



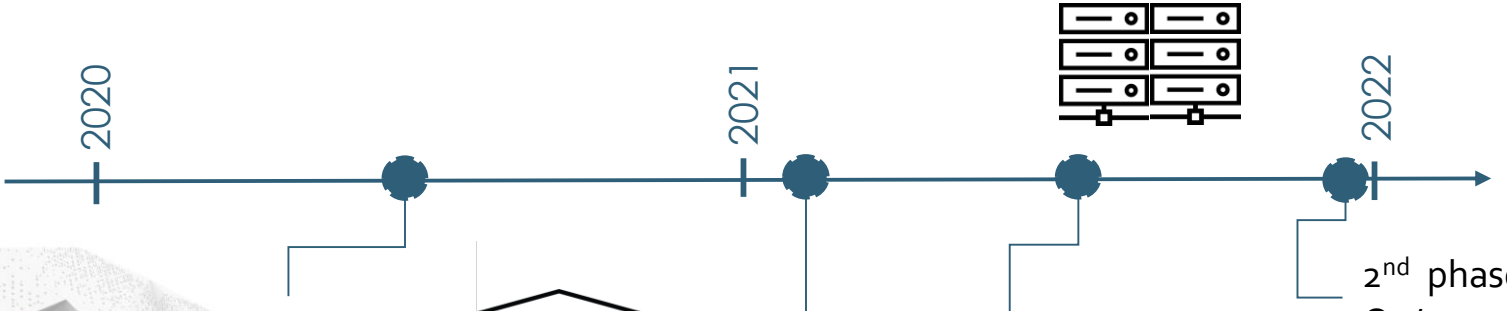
Tier-0 GPU partition: over 550 Pflop/s powered by AMD Instinct GPUs

Interactive partition with 32 TB of memory and graphics GPUs for data analytics and visualization

7 PB Flash-based storage layer with extreme I/O bandwidth of 2 TB/s and IOPS capability. Cray ClusterStor E1000.

80 PB parallel file system

LUMI Timeline



LUMI

Enhanced user experience

- In addition to traditional CLI, we wish to support high-level interfaces on LUMI, i.e. seamlessly integrate Jupyter Notebooks, Rstudio and such to back-end to LUMI
  - Ultimately the LUMI resources being an extension to your laptop
- Large software budget will enable a rich stack of pre-installed software
- Datasets as a Service: curated large reference datasets available and maintained
- Support for handling sensitive (GDPR subjected, IP-closed, etc) data

LUMI user support

- LUMI user support and a centralized help-desk by the distributed LUMI User Support Team
  - The model is based on a network of **dedicated LUMI experts**: each partner will provide one full-time person for the task
  - User Support Team will also provide end-user training, maintain the software portfolio and user documentation of the system
- ”Level 3” support (e.g. application enabling, methodology support) via local centers as well as the EuroHPC Competence Centers

How to prepare for LUMI?

LUMI capacities, a brief summary

- Extreme computing capacity based on LUMI-G and LUMI-C partitions
  - LUMI queue policies will support jobs from single CPU core or a GPU to 50% of the nodes, even 100% with special arrangements
  - Jobs can combine resources from both sides within a workflow, even within the same executable
- Interactive use (visualization, data analysis, pre/post processing,..) on LUMI-D
- Broad stack of pre-installed scientific software, databases and datasets, both commercial and community
- Sharing datasets over LUMI-O service
- Running microservices on LUMI-K
- Exploring the quantum computing world with LUMI-Q

Getting LUMI resources

- LUMI resources are allocated in terms of GPU-hours, CPU-core-hours, and storage hours
  - Each project applies and gets a combination of this
  - No dedicated hardware - all users can access the whole system within the batch job policies
  - All countries receive shares of these pools per their share of the TCO – 4% for Sweden
- Resources brokered in terms of
  - Preparatory access projects (XS) – single-PI
  - Development access projects (S) – single-PI
  - General access (Tier-1) projects (M) – single-PI
  - Extreme scale (Tier-0) projects (L) – single-PI, should be mostly GPU hours
  - (Strategic Community Access projects (XL) – multi-PI, multi-year)
- Researchers affiliated to Swedish institutions can apply from the EuroHPC allocation or from Sweden's allocation

LUMI

Plans for batch job queue policies

Queue	Max Wall-Time (hh:mm)	Max Resources	Remarks
test	00:30	2% GPUs / 2% CPU cores	Testing and debugging queue. Max 1 job per user, high priority
short	06:00	2% GPUs / 2% cores	
default	48:00	50% GPUs / 50% cores	Standard queue for production work
longrun	168:00	2% GPUs / 2% cores	Maximum one long job per user, lower priority
large	24:00	99% GPUs / 99% cores	By special arrangement only. One job per user at a time. Draining of the system for large queue jobs during weekends.
interactive	06:00	4 GPUs / 128 cores	for interactive batch jobs, max 1 job per user

Subject to change per user feedback and observed usage

How to prepare for LUMI?

- Thinking projects and use cases for Lumi
  - Cases for Tier-0 grand challenges
  - Combining simulation and AI methods within the same workflow
- There is a vast pool of GPU-enabled community codes
  - See if your favorite software suite already has been enabled - if not, consider moving to a competing package that is
  - Perhaps only part of the application needs to be GPU-enabled, rest running on the CPU nodes
- Modernizing applications and GPU-enabling them
  - "even if it works, fix it"

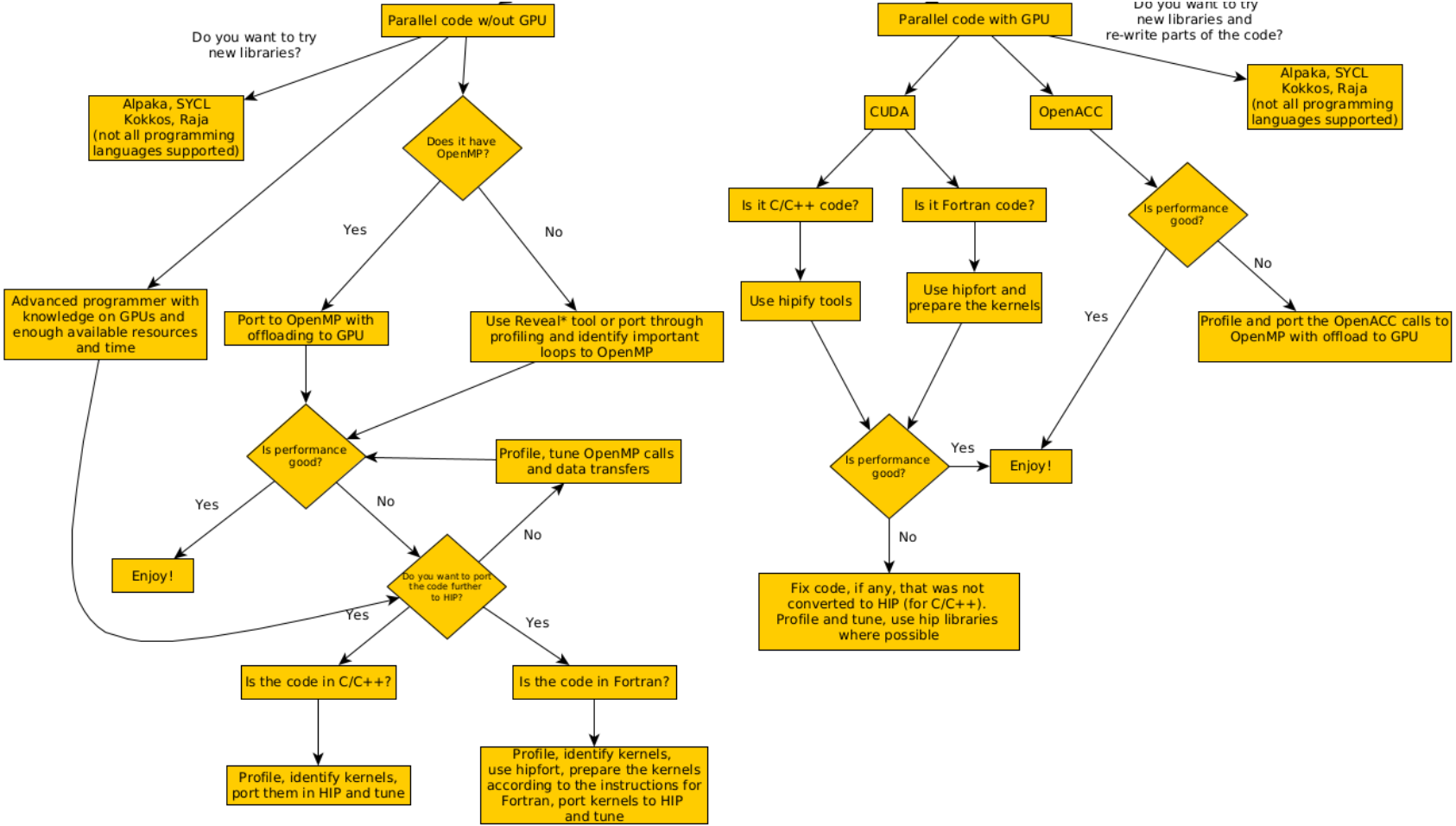
LUMI programming environment

- ROCm (Radeon Open Compute)
  - Usual set of accelerated scientific libraries (BLAS, FFT etc)
  - Usual machine leaning frameworks and libraries (Tensorflow, PyTorch etc)
  - Compilers for the GPUs
- Cray Programming Environment (CPE) stack
  - Cray Compiling Environment, LibSci libraries, CrayPAT, Reveal, debuggers,…
  - CPE Deep Learning Plugin
- More information:
  https://www.lumi-supercomputer.eu/may-we-introduce-lumi/

Preparing applications and workflows for LUMI

LUMI

- Remember the possibility of combining CPU and GPU nodes within one job – perhaps only part of the application needs to be GPU-enabled
- Consider writing your application on top of modern frameworks and libraries
  - Kokkos, Alpaka etc, or domain-specific frameworks
- Convert CUDA codes to HIP, OpenACC codes to OpenMP5
  - HIPify tools can automatize the effort
- LUMI 1st phase will come with a code porting platform (MI100 GPUs)
  - HIP porting can be done already now on Nvidia GPU platforms


Parallel code w/out GPU
Do you want to try new libraries?
Alpaka, SYCL Kokkos, Raja (not all programming languages supported)
Does it have OpenMP?
Yes
No
Advanced programmer with knowledge on GPUs and enough available resources and time
Port to OpenMP with offloading to GPU
Use Reveal* tool or port through profiling and identify important loops to OpenMP
Is performance good?
Profile, tune OpenMP calls and data transfers
Yes
No
Enjoy!
Do you want to port the code further to HIP?
No
Yes
Yes
Is the code in C/C++?
Is the code in Fortran?
Profile, identify kernels, port them in HIP and tune
Profile, identify kernels, use hipfort, prepare the kernels according to the instructions for Fortran, port kernels to HIP and tune
Parallel code with GPU
Do you want to try new libraries and re-write parts of the code?
Alpaka, SYCL Kokkos, Raja (not all programming languages supported)
CUDA
OpenACC
Is it C/C++ code?
Is it Fortran code?
Is performance good?
No
Yes
Use hipify tools
Use hipfort and prepare the kernels
Profile and port the OpenACC calls to OpenMP with offload to GPU
Is performance good?
Yes
Enjoy!
No
Fix code, if any, that was not converted to HIP (for C/C++). Profile and tune, use hip libraries where possible

Concluding remarks

- **EuroHPC era: Unprecendent amount of computational resources and capabilities** available for European research & innovation
  - Complemented by competence building and user support activities
- **LUMI, the Queen of the North**: leadership-class resource designed for a broad range of user communities and workloads, with an enhanced user experience
  - **LUMI will be a GPU system,** which needs some preparatory work – but it will be a robust production system, and not experimental or esoteric in any manner
- **Modernizing HPC applications** for harnessing the largest systems is not trivial, and needs a lot of focused effort – but it will pay off
  - It is time already to start preparing for the LUMI era

LUMI

Dr Pekka Manninen

Director

LUMI Leadership Computing Facility

CSC – IT Center for Science Ltd

pekka.manninen@csc.fi

tel. +358 50 3812 831

Follow us

Twitter: @LUMIhpc

LinkedIn: LUMI supercomputer

YouTube: LUMI supercomputer

www.lumi-supercomputer.eu

contact@lumi-supercomputer.eu

EuroHPC Joint Undertaking

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC Joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the of Participating States FI, BE, CH, CZ, DK, EE, IS, NO, PL, SE.

Leverage from the EU 2014–2020

European Union European Regional Development Fund

Kainuun liitto