Deploying and Managing LUMI Supercomputer, Sustainably

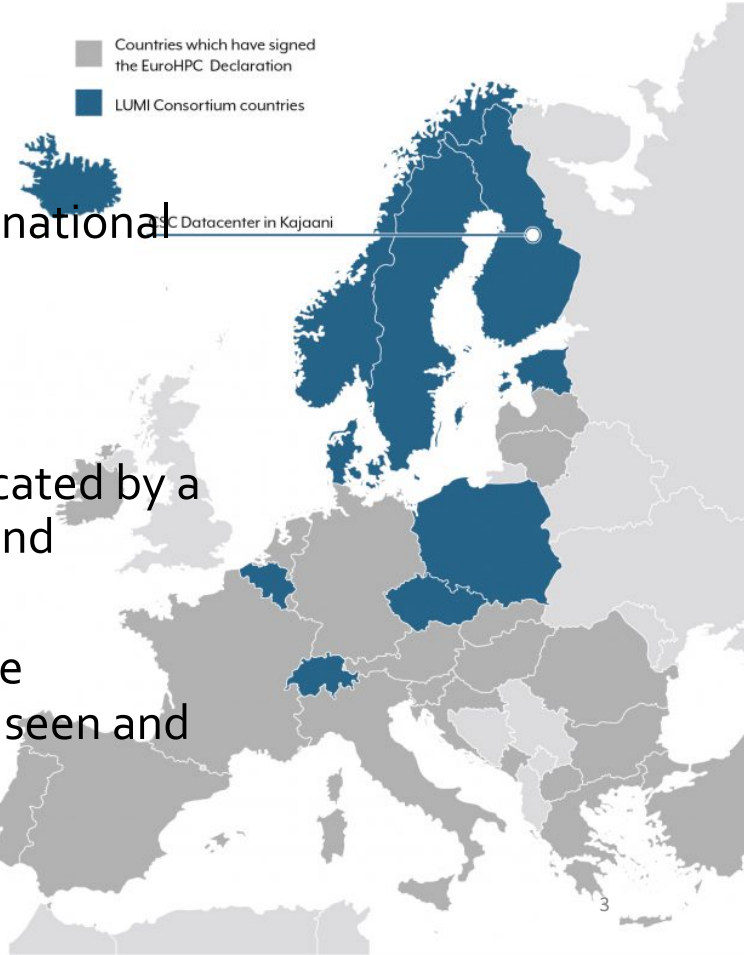
Dr. Pekka Manninen
Director, LUMI
CSC – IT Center for Science, Finland
Adjunct Professor, University of Helsinki

Outline

- Overview of LUMI & LUMI DC
- Monitoring LUMI and future ambitions
- On environmental sustainability of operating HPC systems

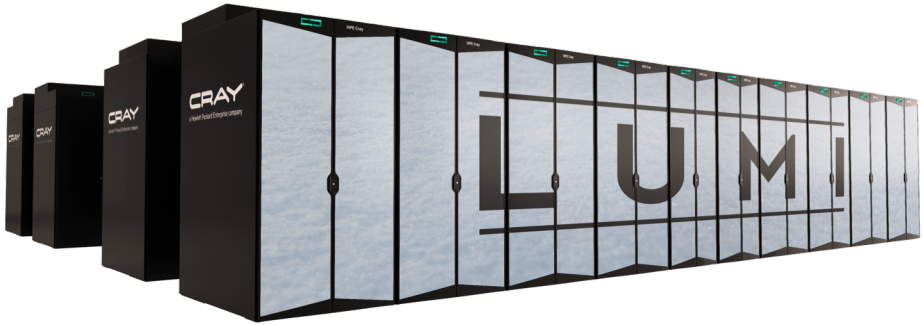
LUMI Consortium

- 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

LUMI: one of the fastest supercomputers in the world



- LUMI will be an **HPE Cray EX** supercomputer manufactured by **Hewlett Packard Enterprise**
- HPL performance over **375 petaflop/s** makes the system one of the world's fastest
  - Partial system listed 05/22 with 152 Pflop/s, #3 Top500
  - #3 also in Green500 and HPCG

1 system

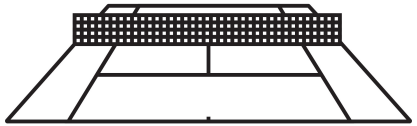
375 Pflop/s

Sustained performance

Computing power equivalent to

1 500 000

Modern laptop computers



Size of two tennis courts

Modern platform for

High-performance computing, Artificial intelligence, Data analytics

Based on GPU technology

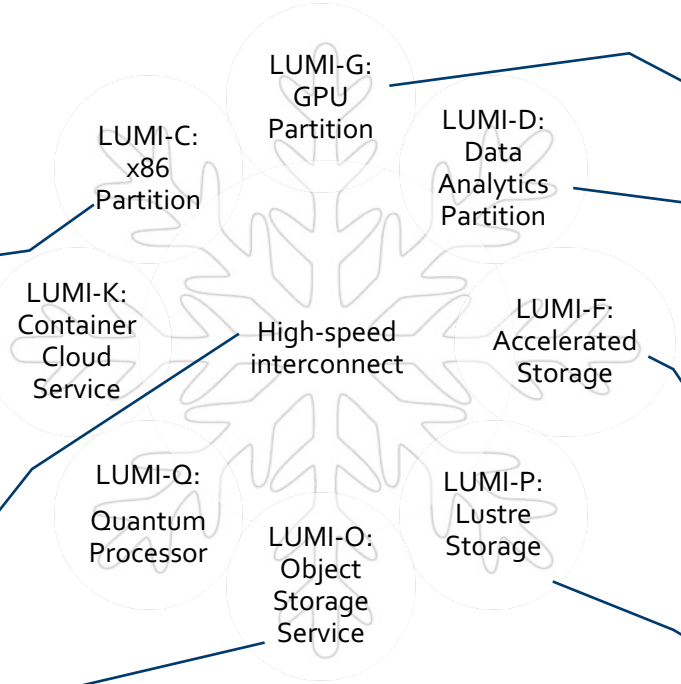
LUMI, the Queen of the North

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
- 196,000 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 10,240 AMD Instinct MI250X GPUs

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

8 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

Pilot use
Extended beta testing
LUMI-C regular access
LUMI-G development access

Q3/2022 **Final Configuration**
- LUMI-O, -K
General availability

2020

2021

2022

System procurement:
November 2019 – August 2020

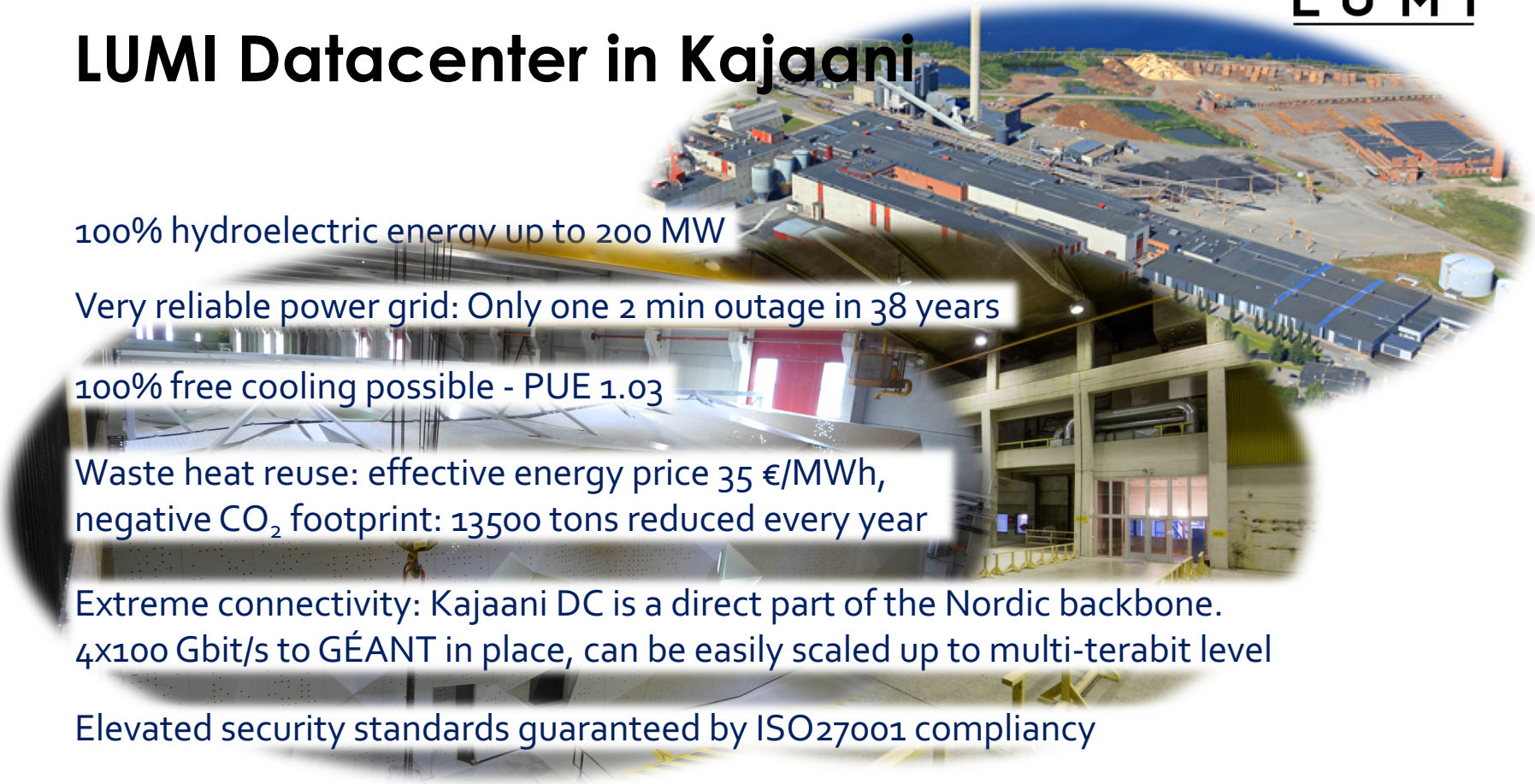
Data center preparation ready Q1/2020

1st phase Q4/2021
- LUMI-C, P, F, D
- Early Access Platform

Q2/2022
- Gradual deployment of LUMI-G
- LUMI-C NIC upgrade

In customer use 01/01/22

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 possible - 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


RENFORSIN RANTA
RATA
LUMI
KONE

LUMI power monitoring and building management system

LUMI

- LUMI data center infrastructure monitoring and management is provided by Fidelix which is a pioneer in smart building automation. More than 8000 points connected to BMS system.
- Power is monitored by ABB, Schneider electric and Carlo Gavazzi power monitoring products.
  - Power monitoring accuracy is between ±0.5-1% with max. sample rate of 4096 samples/s @ 50Hz
- All monitoring data is collected to building management system database

Future ambitions with operational data analytics

LUMI

	Building infrastructure	System hardware	System software	Applications
Prescriptive	• Switching between types of cooling, tuning of cooling machinery • Responding to anomalies	• Cooling optimization at system level • CPU frequency tuning • Tuning of hardware knobs	• Intelligent placement of tasks and threads • Plan-based scheduling • Power and KPI-aware scheduling	• Auto-tuning of HPC applications • Code improvement and recommendations
Predictive	• Predicting DC KPIs • Predicting cooling demands • Modeling cooling performance	• Forecasting hardware sensors • Component failure prediction • Predicting CPU instruction mixes	• Simulating HPC systems and schedulers • Predicting HPC workloads	• Predicting job durations and resource use • Predicting performance profiles of code regions
Diagostic	• Fingerprinting DC crises • Anomaly detection • Infrastructure stress testing	• Node-level anomaly detection • System-level root-cause analysis • Network contention issues	• Diagnosing data locality issues • Detection of software anomalies • Identifying sources of OS noise	• Application fingerprinting • Performance patterns • Diagnosing code-level issues
Descriptive	• PUE calculation • Facility data processing • Dashboards	• ITUE calculation • System performance indicators • Dashboards	• Slowdown calculation • Scheduler dashboards	• Job performance • Job data processing • Job-level dashboards

Adapted from Bates et al MODA 2021

Part II: On the environmental sustainability of HPC installations

Considerations for a HPC system's carbon footprint

LUMI

- Data center level choices
  - Power: used electricity, power-usage efficiency
  - Waste heat reuse
    - District heating, sorption cooling, water preheating, desalination, biomass processing, greenhouses,…
  - Construction/retrofitting of the data center
  - Intelligent operations via advanced monitoring
- System level choices
  - ICT manufacturing
  - Eco-efficiency of the hardware and software ("science per watt")

Benefits of the brownfield solution

We assume having **reduced the CO2 footprint of LUMI data center construction by over 80%** with the brownfield solution vs. constructing an all-new building for LUMI

Materials - building shell 5,700 ft² (530 m²) office facility	Tonnes of CO_2	Percentage of total
Foundation (concrete)	4.7	4%
Flooring (concrete slab, insulation)	39.9	31%
Ceilings (plaster board)	2.3	2%
Structure (steel beams)	15.4	12%
External walls (brick, insulation)	32.1	25%
Internal walls (wood frame and plasterboard)	8.7	7%
Stairs (concrete)	1.1	1%
Windows (glass and frame)	0.59	0.4%
Internal doors (particle board)*	-0.4	-0.3%
External doors (plastic)	0.6	0.5%
Roof (wood, concrete, insulation)	23.4	18%
TOTAL	128.3	100%

For a 1 MW DC, source: Schneider-Electric white paper 66

Waste heat utilization

- **95% of LUMI's waste heat can be re-used in the district heating system of Kajaani**
  - Energy costs go down by 37% as local energy company pays for the waste heat.
  - As an alternative, 100% free cooling available, PUE 1.03
- With LUMI's heat, the local energy company can reduce the use of oil that corresponds in CO_2 emissions to removal of 3000 cars from traffic

100 %

Carbon-neutral energy

CO_2 footprint

-13 500T

Co_2 eq/emissions

LUMI produces

20 %

of Kajaani's district heating needs

DRY AIR COOLING FOR BACK UP ~ 9 MW

DISTRICT HEATING NETWORKS ~ 10 MW
-Renforsin Ranta Business Park
-CITY of Kajaani

LUMI

+31 °C

+45 °C

+60 °C

+80 °C

Maximum heat production approximately 10 MW

HEAT EXCHANGERS

HEAT PUMPS

Electricity to heat pumps is CO_2 free.

+32 °C

+44 °C

+32 °C

+44 °C

HPC load Max. 9 MW

In addition of DLC there is approximately 1 MW of capacity for the air-cooled servers (e.g. storage). This IT is free-cooled.

LUMI system level choices

- ICT equipment life cycle
  - Responsibility and sustainability required and rewareded in the CfT
- Operations and energy efficiency
  - In the procurment, performance figures normalized with energy
  - LUMI will be at top of Green500 (#3 05/22) over multiple lists
- Other considerations
  - Benchmark applications included cases important for green transition (climate models, materials science)
  - LUMI is strongly positioned as an instrument for climate research, especially EU's Destination Earth programme

Concluding remarks

- Green-ness of a HPC installation starts on the data-center level choices, especially contracted source of energy
- Carbon-neutral (even negative) HPC operations possible already today
- Advanced facility and system monitoring vital for adaptive optimization of the operations

LUMI

Dr Pekka Manninen

Director of LUMI RI

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

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

REGIONAL COUNCIL OF KAINUU

EURO