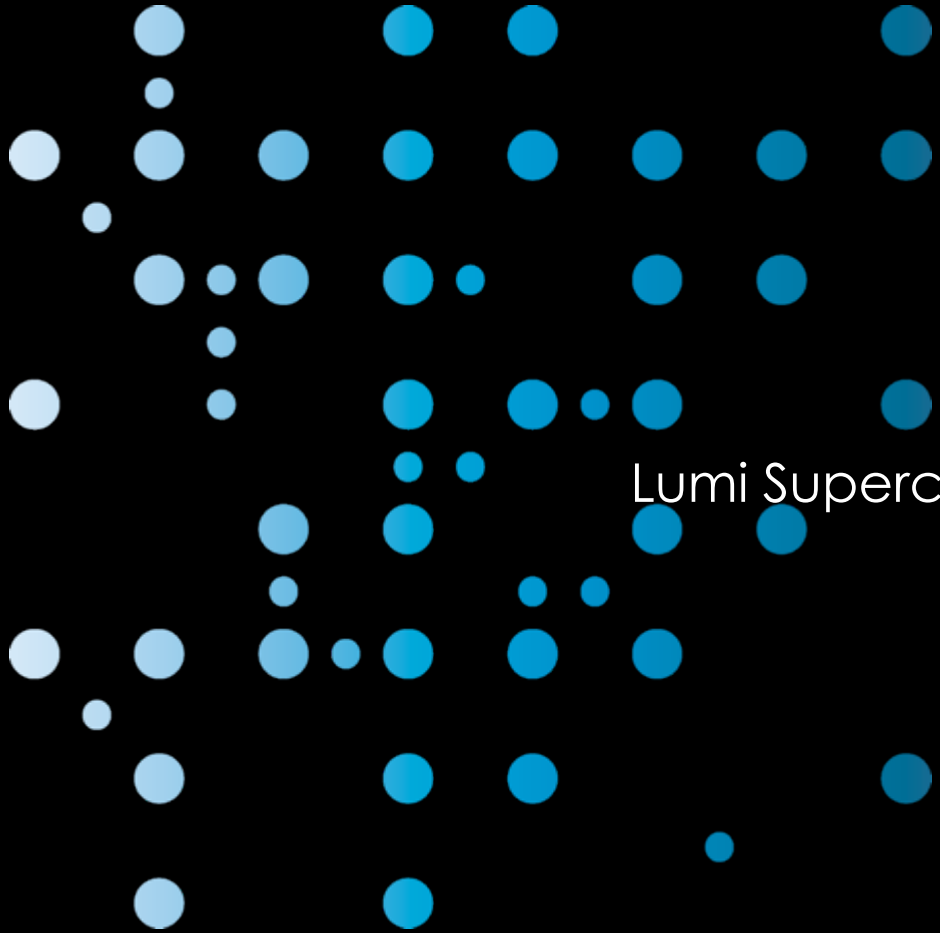
Towards quantum-accelerated AI supercomputing with LUMI-AI

Dr. Pekka Manninen
Director, LUMI AI Factory
CSC – IT Center for Science, Finland

Outline

- Briefly on EuroHPC Lumi supercomputer
- AI supercomputing?
- LUMI AI Factory & LUMI-AI supercomputer (“Lumi-Next”)



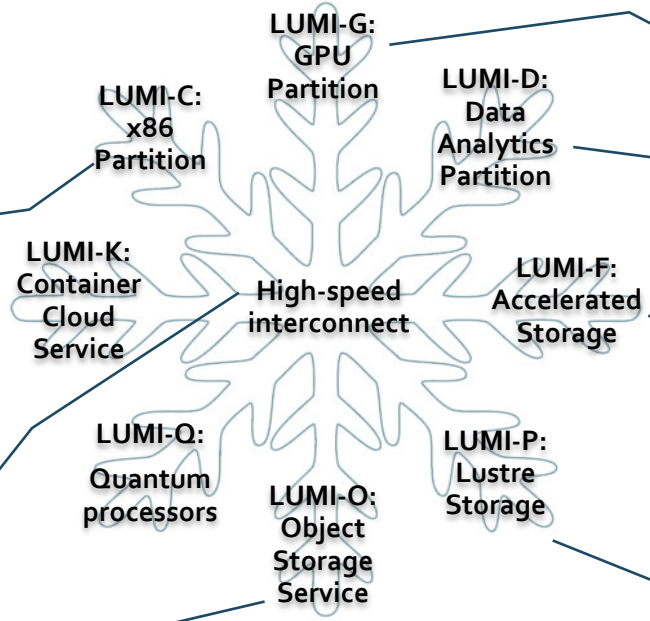
LUMI

Lumi Supercomputer


RENFORSIN RANTA
RATA
LUMI
KONE

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, 262,000 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: 11,900 AMD Instinct MI250X GPUs

Interactive partition with graphics GPUs for data analytics and visualization

9 PB Flash-based storage layer with read bandwidth of 3 TB/s and extreme IOPS capability. Cray ClusterStor E1000.

80 PB parallel file system

Lumi storyline

Decision to host the system in Kajaani 06/2019

2020

2021

2022

2023

2024

Q2-3/2022 Gradual deployment of GPU partition

Object storage

In customer use 12/04/23

System procurement: 11/2019 – 08/2020

Data center preparation ready 03/2021

1st phase 11/2021

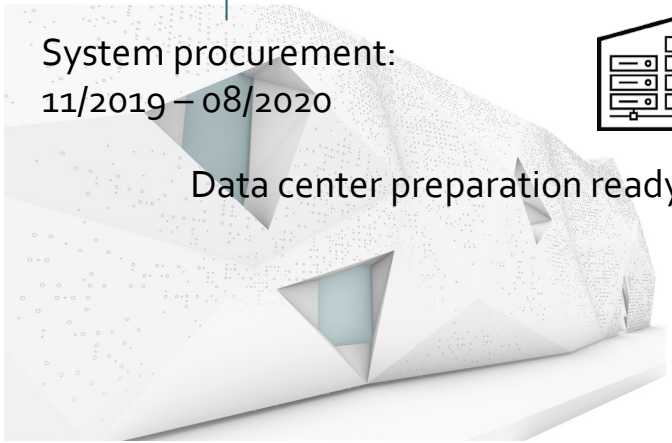
In customer use 01/01/2022

2nd phase 11/2022

In customer use 19/12/2022

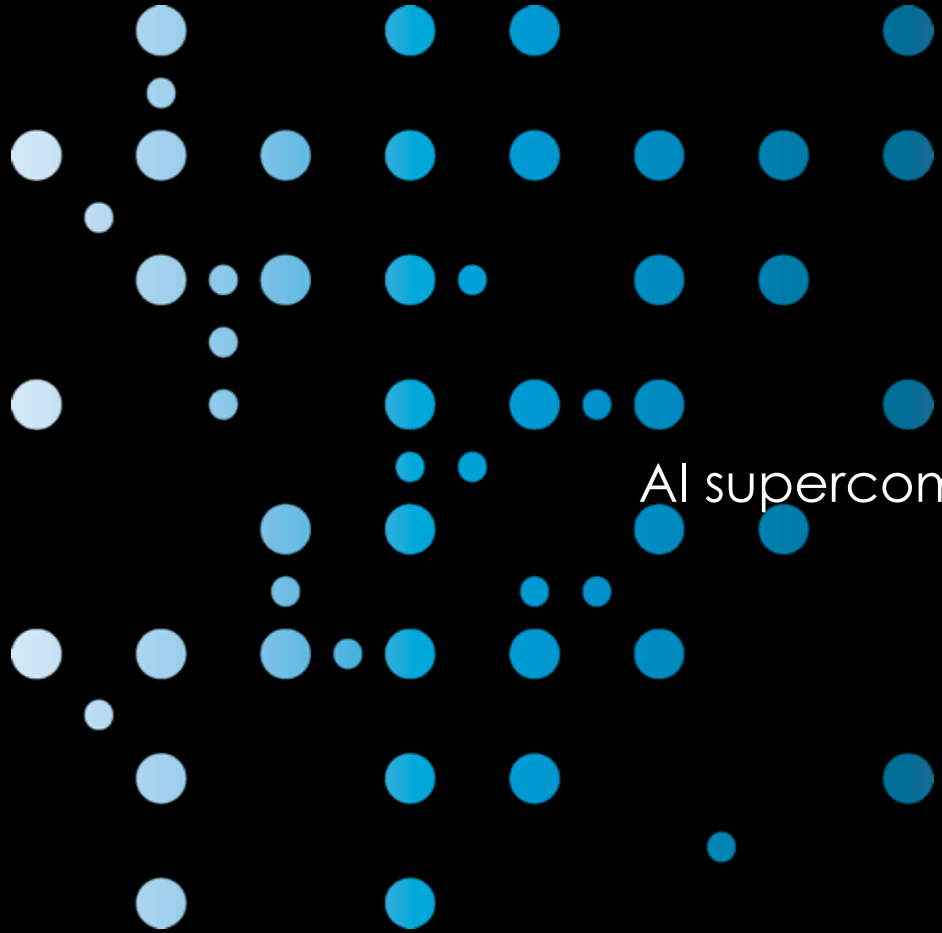
3rd phase 08-10/2023

In customer use 30/08/2023 and 05/11/2023



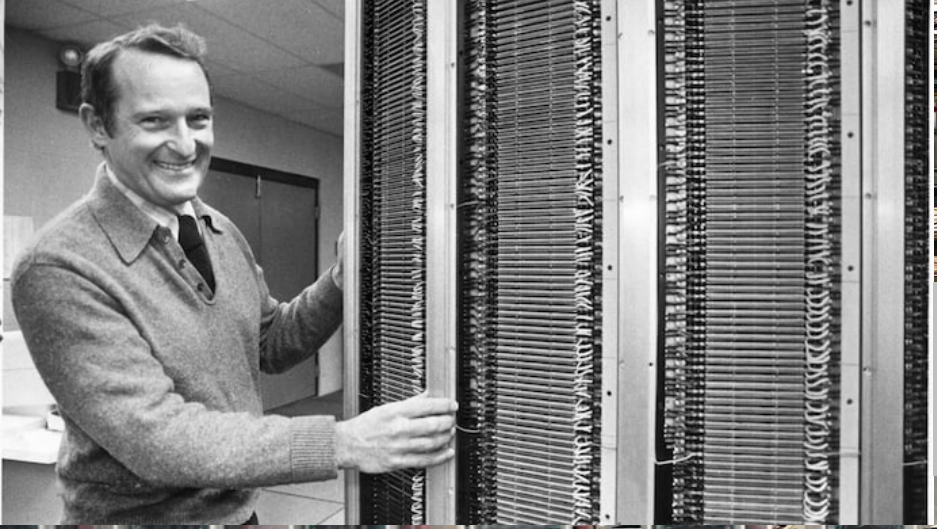
What makes Lumi unique?

- **Modern system architecture** – catering the needs of the scientific computing and data-driven discovery of the 2020's
- **Sustainability** – by far the greenest large supercomputer installation in the world
  - 100% CO_2-free renewable energy
  - Waste heat utilization in the Kajaani district heating, or alternatively 100% free cooling (PUE 1.03)
  - Very energy-efficient (performance / Watt) architecture
- **European collaboration** – first truly joint investment and collaborative hosting effort by 11 countries in European supercomputing
- **Societal impact** – focus on solving societal grand challenges
  - Understand the impact of climate change and help in related decision-making
  - Democratize generative AI
  - Fusion reactor modelling
  - Develop new materials
  - Modeling of biodiversity and biodiversity loss



LUMI

AI supercomputing?

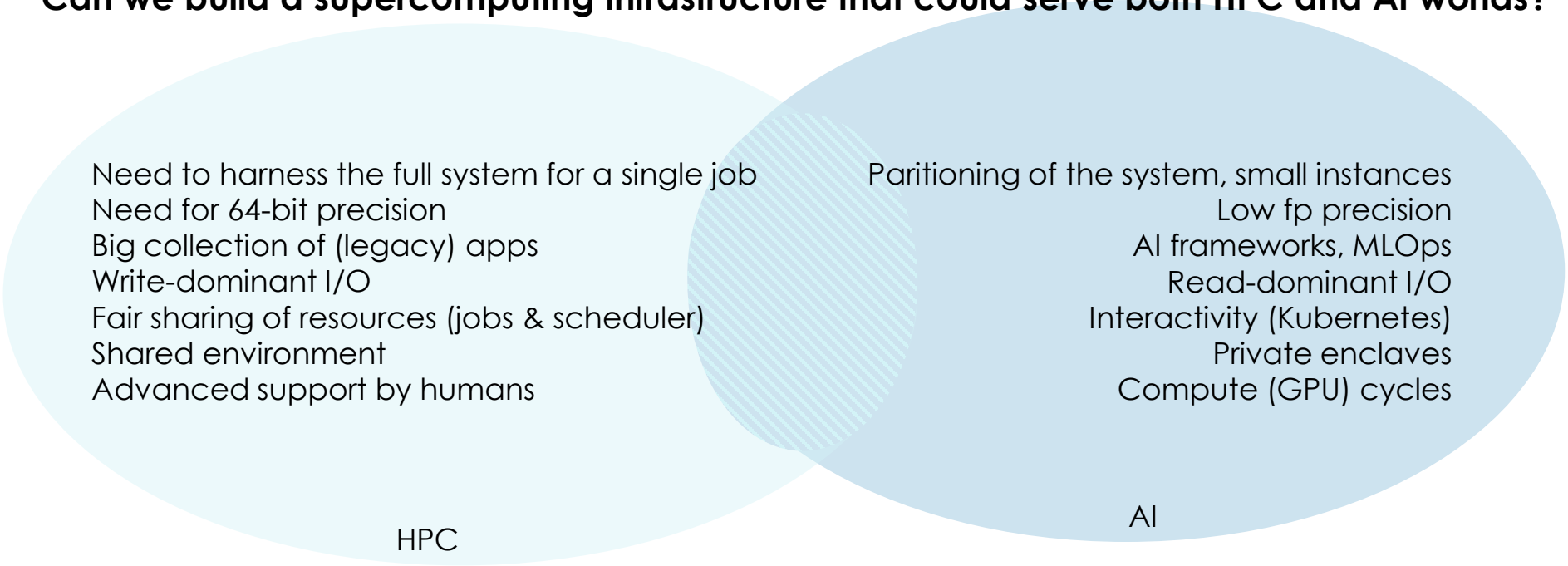


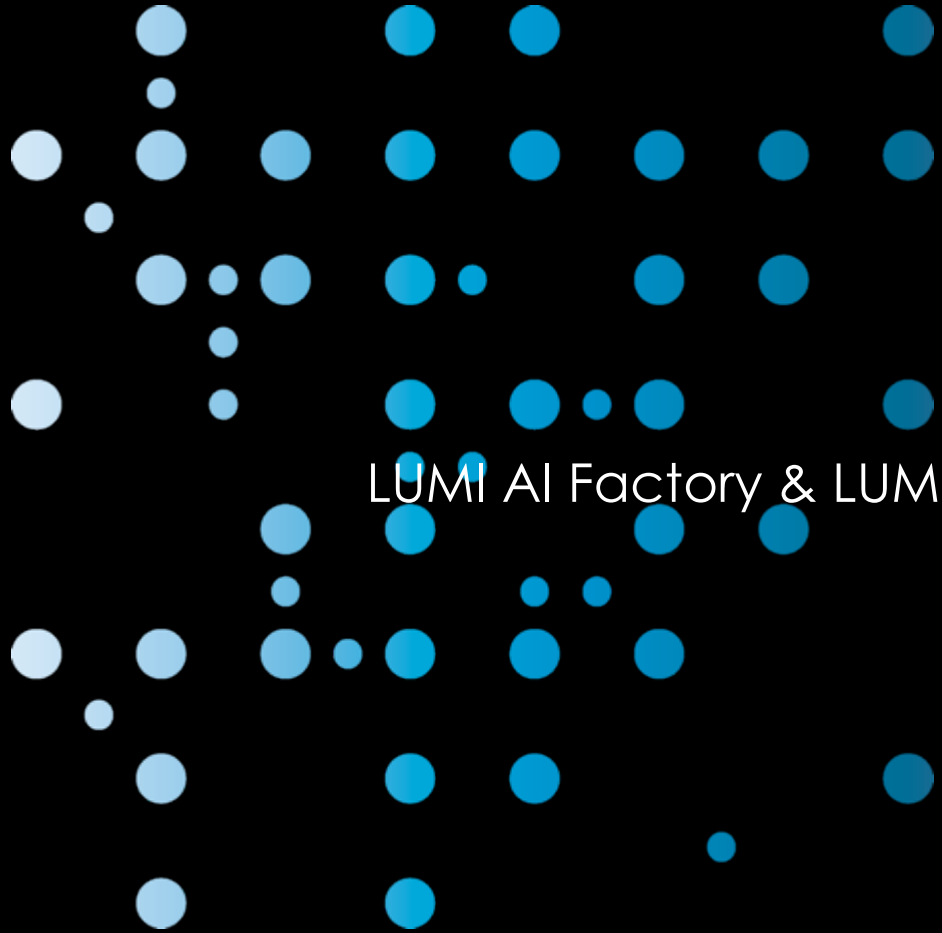
"Anyone can build a fast CPU. The trick is to build a fast system."
Seymour Cray

Anyone can pile up servers into the same room. The trick is to build a fast system.

AI supercomputing

Can we build a supercomputing infrastructure that could serve both HPC and AI worlds?

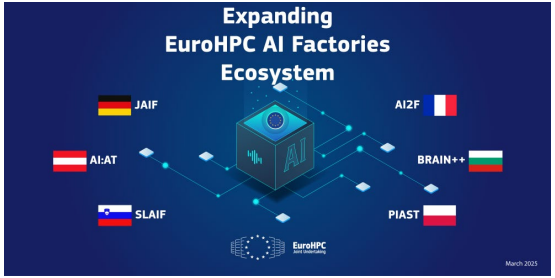




LUMI AI Factory & LUMI-AI supercomputer

EU promotes AI innovation

- European Commission launched the AI Innovation Package in January 2024 to support European startups, and SMEs in the development of trustworthy AI
- A series of competitive calls for AI Factory proposals – first seven announced in December 2024 and further six in March 2025
- AI Factory = **compute** + **data** + **talent**
- AI Factories focus on certain **AI ecosystems and communities in** alignment with **national AI strategies**

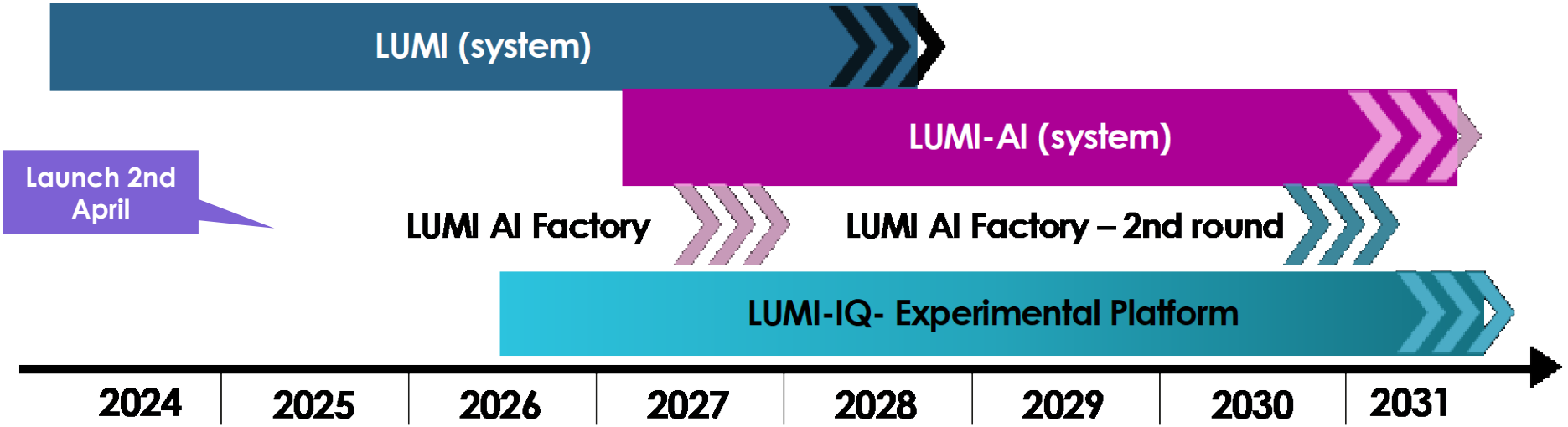


The EuroHPC AI Factories initiative *– a one-stop shop to offer AI startups, SMEs, and researchers comprehensive support, including access to AI-optimised high-performance computing (HPC) resources, training, and technical expertise.*

LUMI AI Factory

- **LUMI AI Factory**
  - AI-optimised supercomputer **LUMI-AI**
  - AI Factory **service center**
  - Experimental quantum computing platform **LUMI-IQ**
- **CSC (Finland)** coordinates the consortium with participation from **Czechia**, **Denmark, Estonia, Norway and Poland**
  - with Finnish partners **Finnish Center for AI** (Aalto University, University of Helsinki) and **AI Finland** (Technology industries)
- Total budget over 612 million euros – largest public computing infrastructure investment in Finland, and largest EU AI Factory investment
- Significant investment in **talent and competence development**

LUMI AI Factory timeline



Services – people and talent

Training

- Structured training paths for AI and HPC, tailored training for different domains

- Deep collaboration with AI centers

Co-working spaces and student facilities to nurture talent

- Main hub on grounds of Aalto University together with ELLIS Institute

- A distributed network connected to partner countries

Consultation (company focus)

- Getting started, feasibility analysis

- Support in applying for large resources

- Trustworthy AI: regulatory consultation, compliance, sandboxes

- Ecosystem development

Focus on competence development

- Support both scientific researchers and industrial innovators to adopt AI methods on a large scale.

Services – data and computing

Computing capacity

- Massive GPU, AI inference, multi-tenant environment, fast storage

- Quantum computing capacity for AI through the LUMI-IQ experimental platform

- Fast lane for ambitious AI startups

Data access

- Cloud-like data environment

- Public and restricted data, dynamic data

- Datasets-as-a-service dataset hosting

- Data spaces and other infrastructures

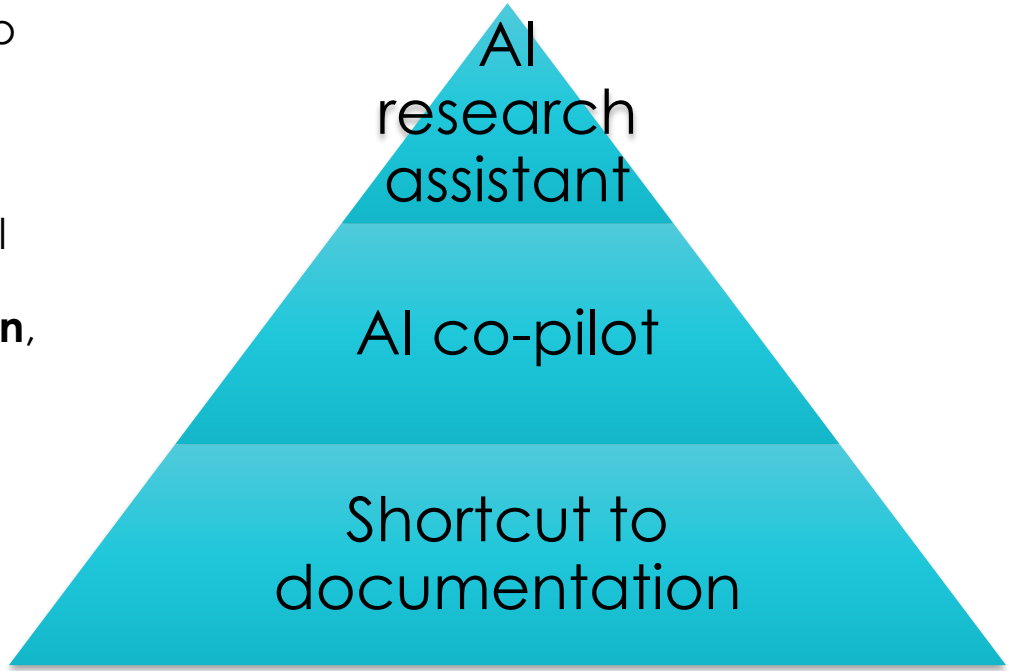
Data Support

- Data wrangling, MLops

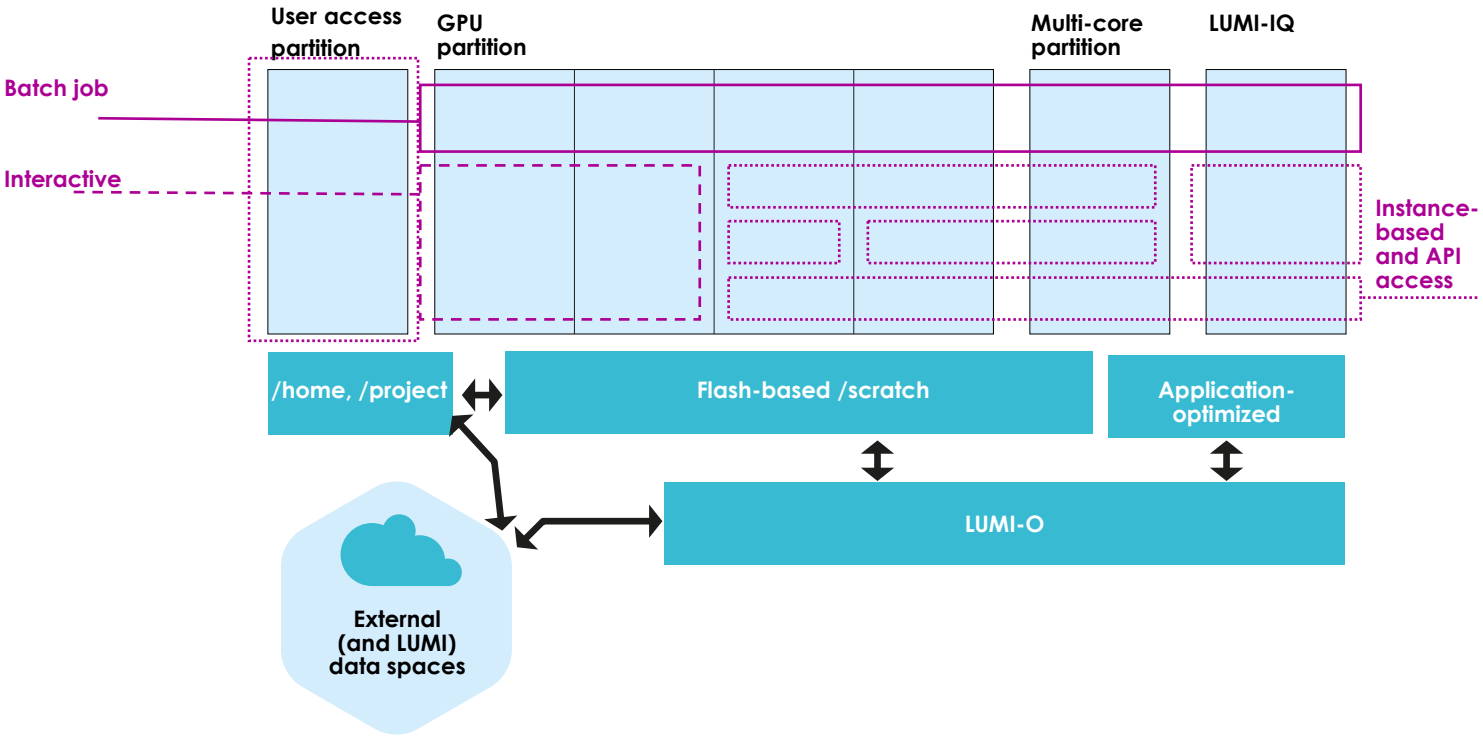
- Deep support for AI methods, scalability

AI assistant for supercomputers

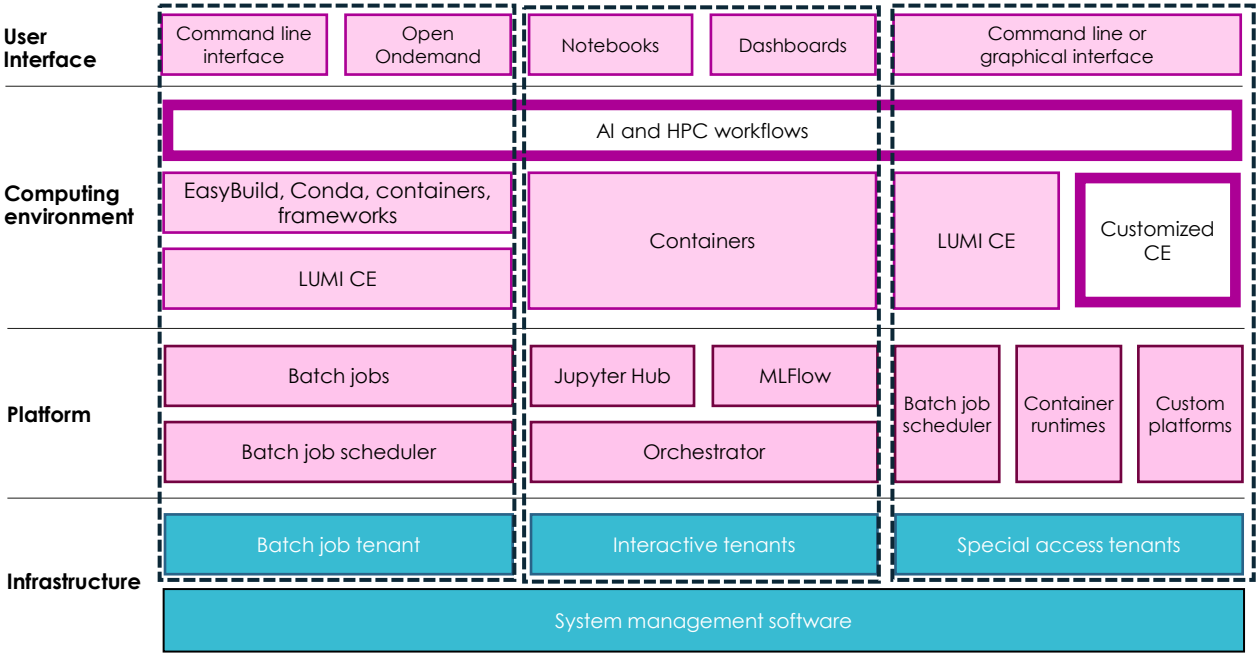
- An LLM-based AI assistant to help users in using a supercomputer system
- **Replacing and complementing support** received from traditional sources like **user documentation**, **helpdesk** and **expert consultation**, to increase researcher's productivity
- Can be actualized with various ambition levels



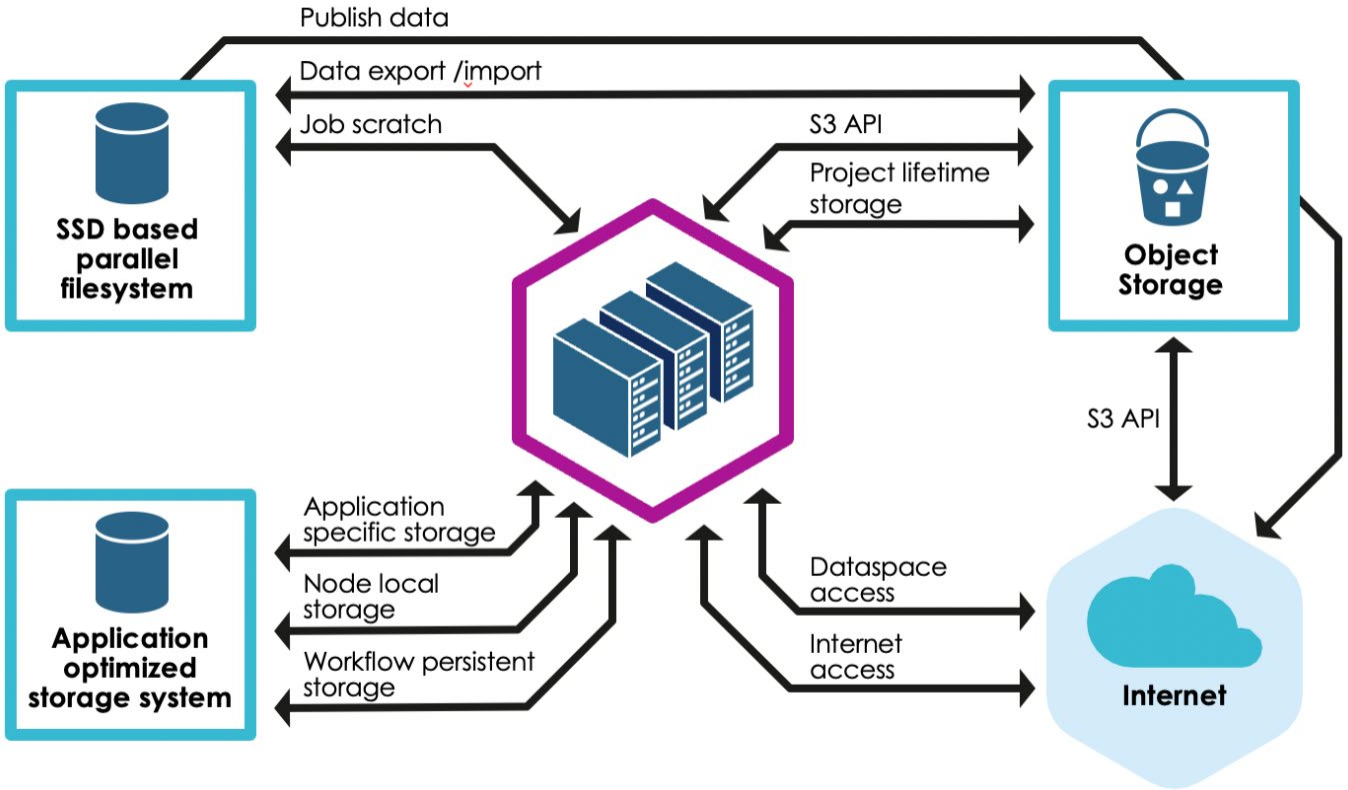
LUMI-AI architecture: HPC + cloud computing converged



LUMI-AI system architecture: Software-defined clusters



LUMI-AI system architecture: Storage solution



Concluding remarks

- We *can* serve the AI community, including enterprises, with HPC systems: "AI supercomputing". The same platform can cater for the scientific simulation workflows of this and the next decade.
- **LUMI AI Factory** will be empowering Europe's AI, scientific computing and "AI for science" ecosystem with one of the most advanced and powerful HPC+AI+QC platforms in the world (**LUMI-AI**)