



## CSC Services for new HPC Computing Era

High Performance Computing with New Opportunities – Allied ICT  
Finland Webinar – 22nd April 2020

Dr. Juhani Huttunen,

Customer Solution Manager, Services for Computational Research



# Contents

- CSC – non-profit organization with special tasks
- High Performance Computing – Motivation for Large Infrastructures
- Infrastructures for scientific computing in Finland
- National infrastructure (DL2021)
  - Real-world research examples deploying CSC computational services
- Other CSC services
- EuroHPC/LUMI
- CSC Services Offering in a nutshell
  - CSC's Approach to AIF collaborative network

Non-profit state  
organization with  
special tasks



Turn over  
in 2019  
**51** M€



Headquarters in  
Espoo,  
datacenter in  
Kajaani



Owned by state **(70%)**  
and all Finnish higher education  
institutions **(30%)**



Circa  
**430**  
employees

# CSC's solutions



Computing and software



Data management and analytics for research



Support and training for research



Research administration



Solutions for managing and organizing education



Solutions for learners and teachers



Solutions for educational and teaching cooperation



Hosting services tailored to customers' needs



Identity and authorisation



Management and use of data

**ICT platforms, FUNET network and data center functions are the base for our solutions**

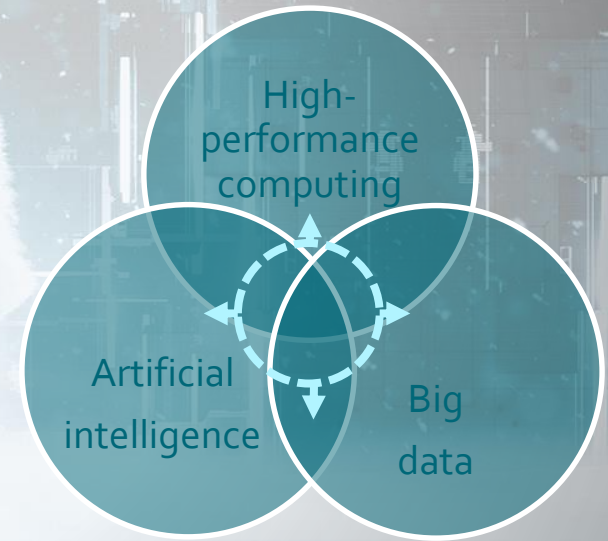


# In the era of data-intensive computing

We need **supercomputers**  
**more than ever:**

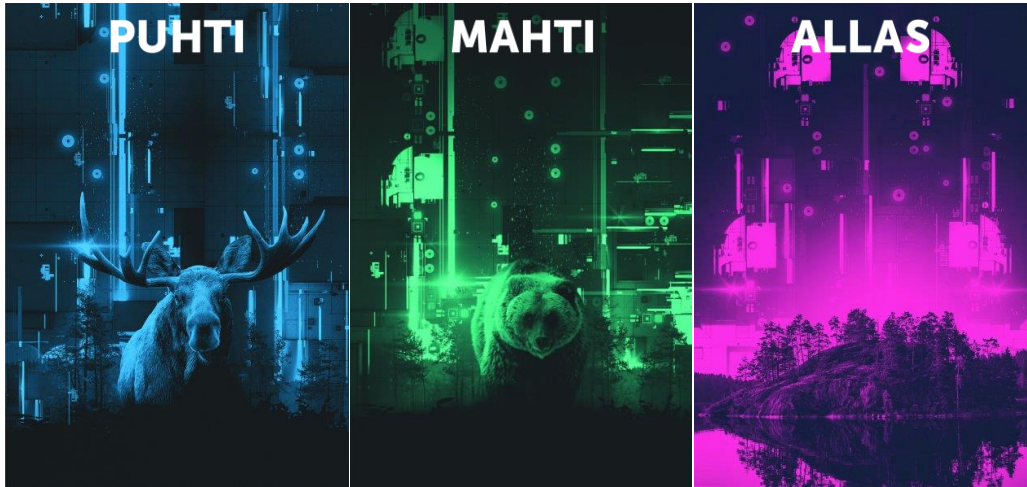
- high demand for computing capability
- interconnect performance
- I/O bandwidth

at the same time.



# Infrastructures for scientific computing in Finland

## *National*



- Large (Tier-1) scale simulations
- High throughput computing
- High-performance data analytics
- Data streams

## *EuroHPC/LUMI*

- Most resource-intensive projects (Tier-0)
- Science and innovation policy priorities
- Artificial intelligence
- Support for large research infras
- Collaboration with industry
- International collaboration

# Who can use these resources and how?

**National** (\*) Usage policy: <https://research.csc.fi/free-of-charge-use-cases>



- The funding by Ministry of Education and Culture
- enables the use of CSC's computing, cloud and storage services

**free-of-charge(\*)**

- for the purpose of academic research and related education
- based on user's affiliation to a **Finnish higher education institution, or state research institute.**

- **Companies** can deploy the resources under similar conditions
- by joining an open research project
- that produces open publications or open data sets.
- **Direct sales** to companies is limited with a tough yearly cap

## EuroHPC/LUMI

25% of LUMI resources for Finland national sharing, of which

- 80% for open research
- 20% directly to company R&D

**Sharing principles are TBD.**

### LUMI total resources:

50% for all-European usage

50% for LUMI consortium members,

of which 50% for Finland

# Real-world examples – medium scale simulations

- Spreading of aerosol particles in air
  - Computational fluid dynamics modelling of airborne transmission of coronavirus
  - Medium scale simulations with few hundred CPU cores
  - PI Ville Vuorinen, Aalto University





# Real-world examples – artificial intelligence

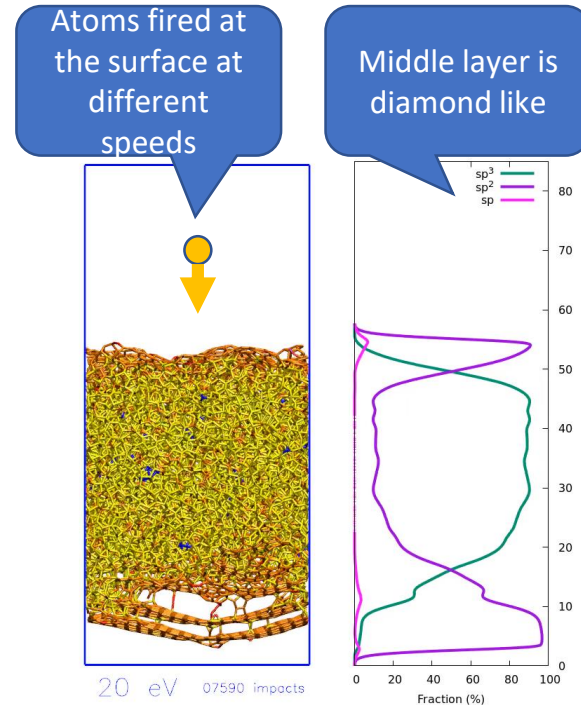
- Prediction of molecular phenotypes for cancer diagnosis
  - Molecular phenotyping can make diagnosis of prostate cancer more precise and comprehensive
  - Deep neural networks are utilized for classification of image data of biopsies
  - Large data volumes and optimized computational pipeline using both CPU and GPU computation
  - PI: Pekka Ruusuvuori, Tampere University



# Machine learned atomic simulation tool

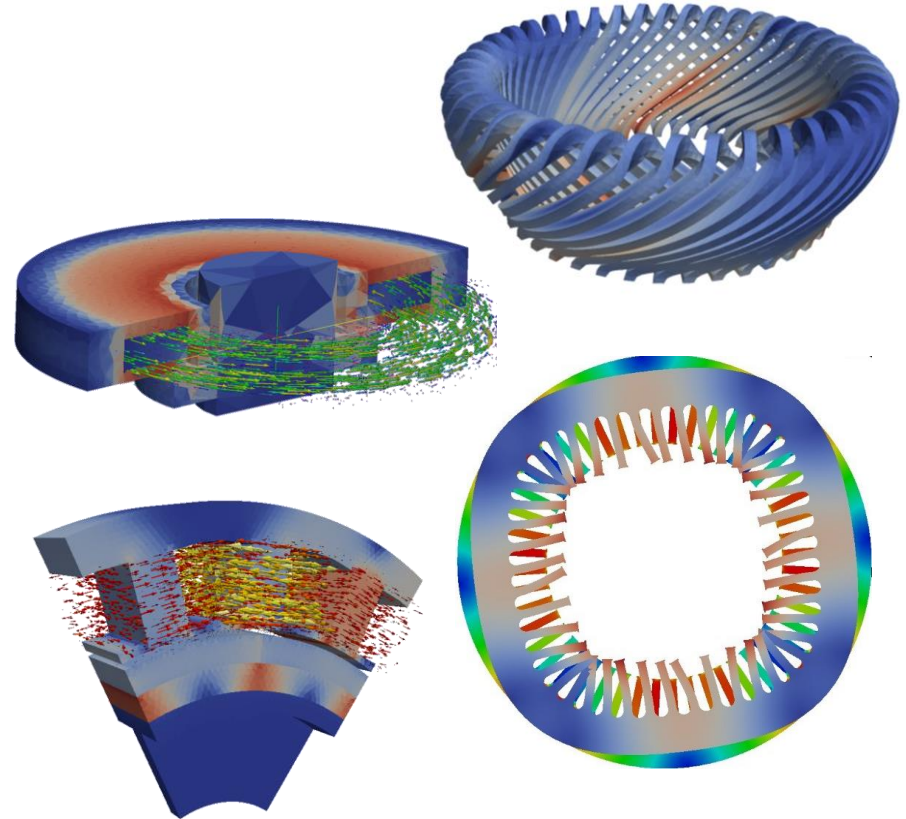
Miguel Caro (Aalto), Volker Deringer (Oxford), Anja Aarva (Aalto), et al.

- a-Carbon is a high tech film, but growth mechanism not known
- Simulation too inaccurate or expensive
- High quality quantum chemistry calculations used to create training data (HPC, 2M CPUh)
- Used to train a ML model, which is cheap to run and solved the mechanism\*
- Complemented with other methods has been used to characterize sensor surfaces\*\*



# Elmer: Developments in electromechanics

- Elmer is a finite element software mainly developed at CSC
- Most important industrial application area at the moment is electromagnetics
- In SEMTEC and Arrowhead Tools projects Elmer has been developed as a tool for heavy electromagnetics computations.
  - Collaboration with CSC, VTT, Aalto Univ., TUT, LUT, ABB, Kone, Konecranes, Sulzer, Ingersoll-Rand, Trafotek, Scanveir
  - Existing commercial solutions provided unsatisfactory scalability



# Example: Digital twin of permanent magnet motor

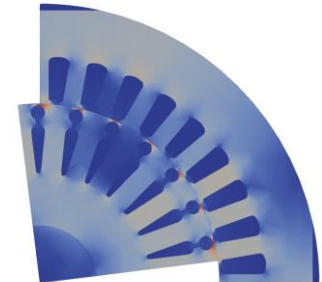
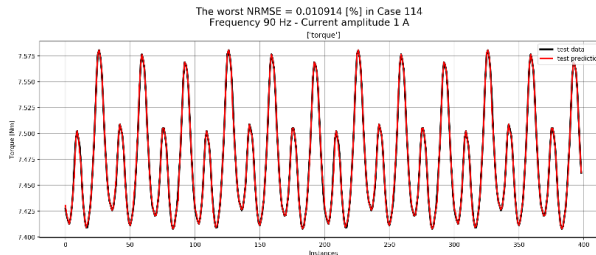
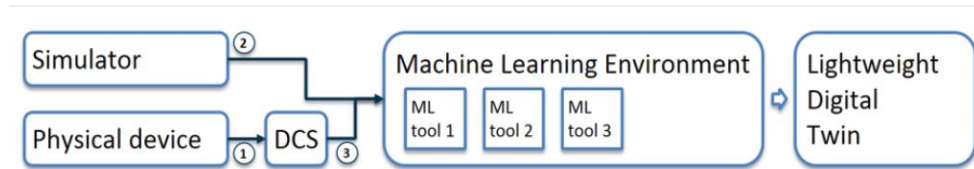
- Deep learning technology is used to create a digital twin for electrical motor
- Large number of 2D finite element simulations are used for training
- Resulting model is more lightweight and may be used in real-time.
- Novel combination of HPC computation and deep learning
- Still under development

Parameter sweeps with

- input: frequency and current magnitude
- output: torque

Machine learning for surrogate model

- Fit to test data to FEM computations
- Test with KNN, RF, SVR, GBR, XGBoost, NN methods



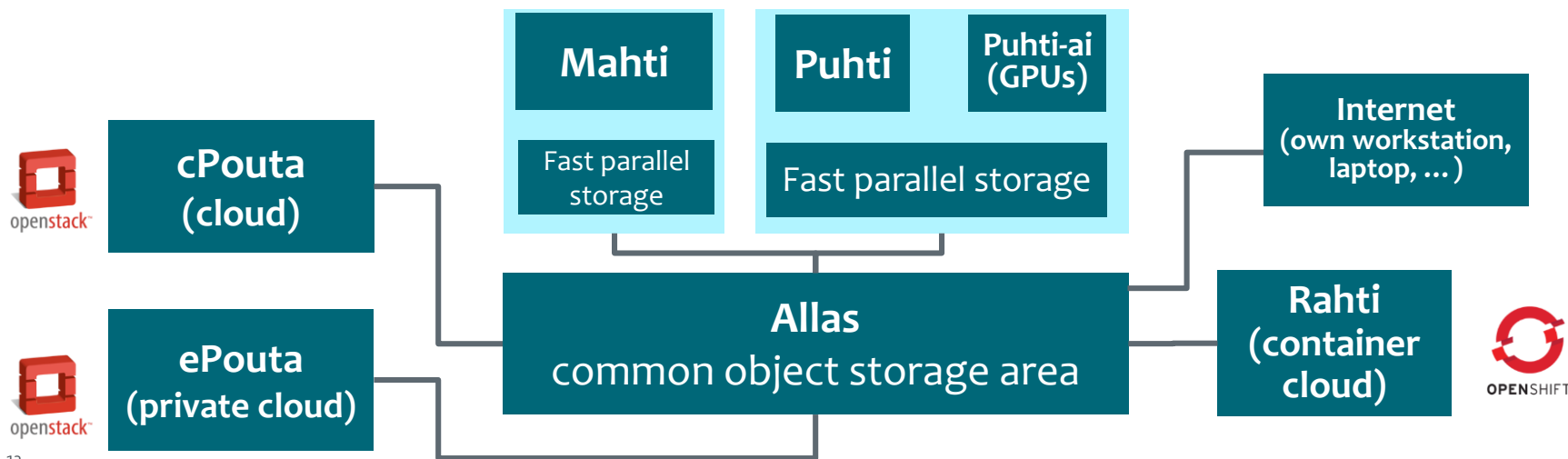
Simulations by VTT, 2020

# The new Finnish research infrastructure for data management and computing



## Balanced HPC ecosystem for various workloads

Heterogeneous, workload-optimized node architecture, support for complex workflows, datasets-as-a-service and containerization



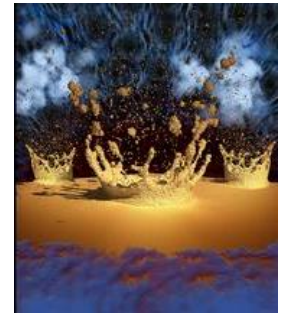


# Other CSC Services



# Expert Support and Training in Sciences and Methods

- Help customers to select the most suitable and effective tools and software for each research project
- Advise in the effective use of supercomputers and cloud computation
- Support in optimisation and parallelisation of your own applications
- Wide collection of scientific software
- In-house scientific software development
  - Elmer, Chipster
- Scientific visualization
- Data analytics



# Data analytics services

- Running big data frameworks in CSC cloud environment
  - Hadoop, Spark
- Specialized computing infrastructure and integration into international infrastructures (EUDAT, ELIXIR, CLARIN)
- Expert support for tools and methods
  - R, Python, Tensorflow, statistics, machine learning, data science
- Solutions for biomedical, language and spatial data
- Notebooks, web based environment for education and data analytics,  
<https://notebooks.csc.fi>

# LUMI - a EuroHPC pre-exascale system

- One of the 3 precursors to exascale supercomputers of the **EuroHPC Joint Undertaking** consisting of 31 European member countries
- For European users from **academia, industry and the public sector**
- Consortium lead by **CSC**: Belgium, Czech Republic, Denmark, Estonia, Norway, Poland, Sweden, Switzerland, and Iceland (under negotiations)
- LUMI is in use 2021-2026

1 system

**Over 200  
Pflop/s**

Peak Performance

Computing power  
equivalent to

**600 000**

MacBook Pro Computers

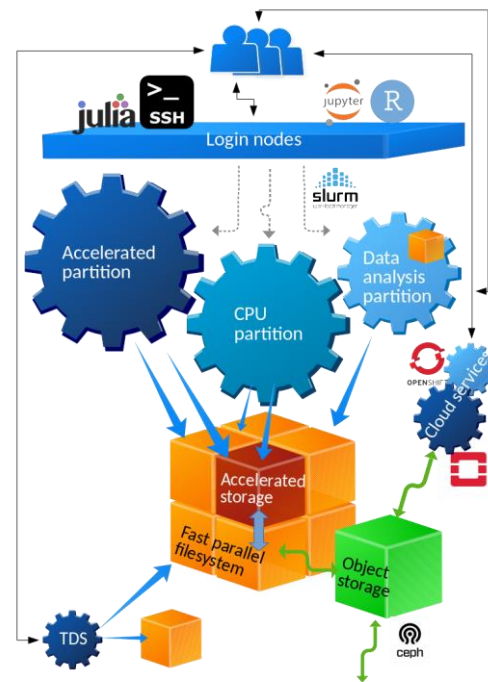
Data reading speed from  
disk to memory  
corresponds to the  
simultaneous operation of

**32 000**

Ultra-HD Streams

# LUMI - a EuroHPC pre-exascale system

- A GPU-accelerated supercomputer
- A highly capable parallel filesystem with anticipated volume of at least 60 PB
- An accelerated I/O (flash) layer providing more than 1 TB/s sustained bandwidth and an extreme IOPS capability
- **Object storage** service for project-time (5 years) storage





# How to prepare for Lumi?

- LUMI is both huge possibility and huge challenge
  - Order of magnitude larger performance increase than normally with new CSC supercomputers
    - Mahti ~ 4 x Sisuu, LUMI ~ 30 x Mahti
  - Paradigm shift: fully accelerated system
- Thinking projects and use cases for Lumi
  - Cases for Tier-0 grand challenges
  - Existing GPU enabled applications
  - Combining simulation and AI methods within the same workflow
- Modernizing HPC applications and GPU-enabling them
  - "even if it works, fix it"

# CSC Services Offering in a nutshell



# Opportunities with new computational and data management resources - Computing



- HPC combining heavy simulation and AI methods within the same workflow
  - Scalable GPU in Puhti and LUMI
  - CSC environment includes the most important AI libraries
  - HPC batch job system in use, costs will be created only when job is running
- Cloud services (virtual servers and containers) for specific use
  - Flexible configuration of the computing environment
  - Interactive use
  - GPUs also available (limited amount)

# Opportunities with new computational and data management resources - Data



- Fairdata.fi services enable verifiable and reproducible science and secure preservation of digital research outputs
- Research data will be efficiently available for CSC's computing services
  - Incl. Databanks (GIS, genome data, language bank)
- Companies will be guided how to open their data for open research usage
  - Datasets-as-a-service
- CSC develops services to manage and elaborate the data and gives expertise support for the methods used (e.g., automatic workflows)

## CSC's Approach to Allied ICT Finland collaborative network

- CSC offers the platform (DL2021 and LUMI) together with
- CSC competences in HPC and AI for
- joint programs of commercial companies and higher education institutions and research institutes
  - With EuroHPC/LUMI environment part of the capacity will be dedicated to companies!





## Juhani Huttunen (Ph.D.)

Customer Solution Manager,  
Services for Computational  
Research, CSC

Keilaranta 14 (PL 405),  
02101 Espoo  
+358 40 581 1138  
juhani.huttunen@csc.fi



[facebook.com/CSCfi](https://facebook.com/CSCfi)



[twitter.com/CSCfi](https://twitter.com/CSCfi)



[youtube.com/CSCfi](https://youtube.com/CSCfi)



[linkedin.com/company/csc---it-center-for-science](https://linkedin.com/company/csc---it-center-for-science)



[github.com/CSCfi](https://github.com/CSCfi)