



EUROPEAN
TECHNOLOGY
PLATFORM
FOR HIGH
PERFORMANCE
COMPUTING

ETP 4
HPC



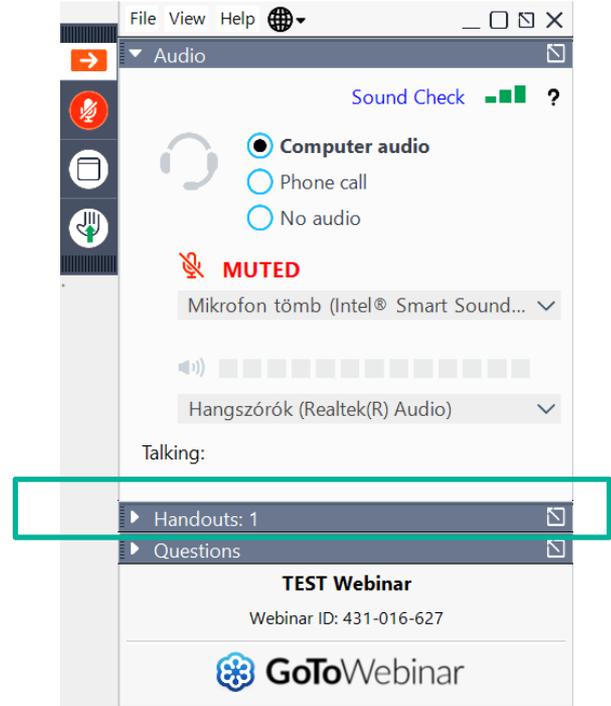
ETP4HPC webinar #9

MeluXina, the EuroHPC system in Luxembourg

28 May 2021

Before we start

- This webinar is recorded, you will receive the link on Monday
- This webinar is in listen-only mode
- Use the GoToWebinar control panel displayed on the right of your screen to download handouts



Questions

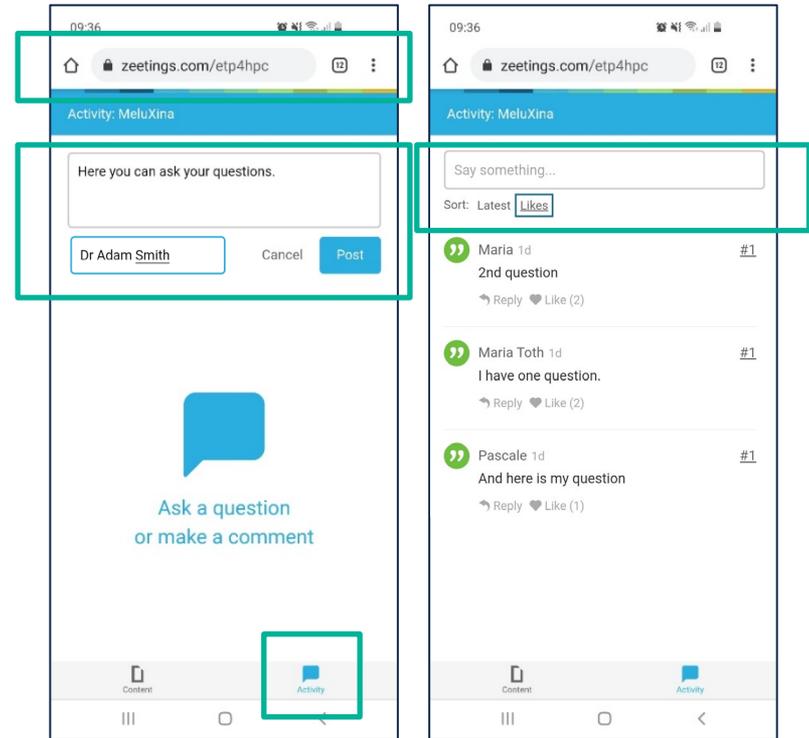
- You can ask questions to the speakers via your phone, tablet or computer typing into your browser:

www.zeetings.com/etp4hpc

- by clicking on the „Activity’ tab
- or on this icon:



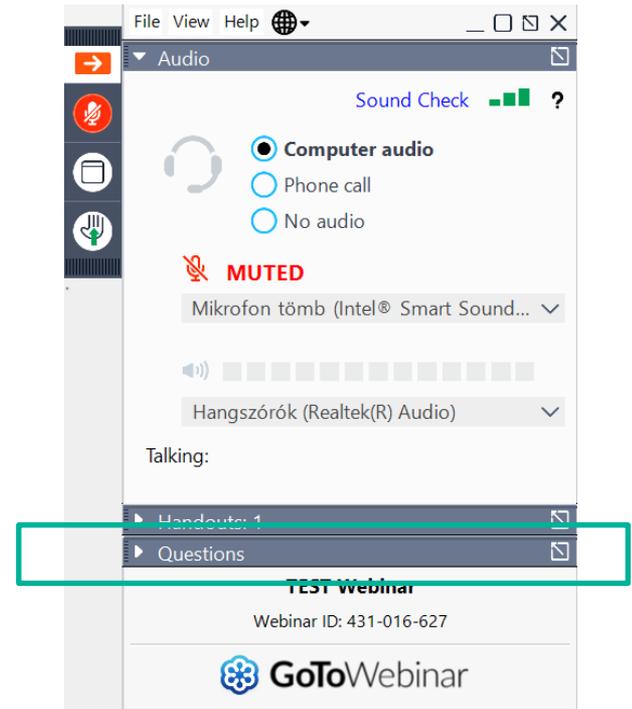
- Please indicate your name



Questions

- If you have technical problems accessing www.zeetings.com/etp4hpc

You can send your questions via the Questions pane of your Go2Webinar control panel.



Agenda

11:00	Introduction	Pascale Bernier-Bruna, ETP4HPC Office Hugo Falter, ParTec
11:05	MeluXina – a new generation supercomputer	Valentin Plugaru, LuxProvide
11:30	ParaStation Modulo and the MSA	Thomas Moschny, ParTec
11:45	Q&A	Moderated by Hugo Falter, ParTec
12:00	End	



Scaling Innovation

ParTec and MeluXina

ETP4HPC Webinar –
MeluXina, the Luxemburg EuroHPC System

May 28, 2021

Hugo Falter, ParTec

■ Hosting Site Procurement

- *Call for expression of interest for the selection of the Hosting Entities for Petascale Supercomputer under REF: EUROHPC-2019-CEI-PT-01*
- *Date of LuxProvide's application for the Petascale Supercomputers with Modular System Architecture: 15th April 2019*
- *Date of decision by EuroHPC: 4th June 2019*
- *Project Management ParTec by Ina Schmitz*
- *Technical Support ParTec by Thomas Moschny*

■ Vendor Procurement

- *Start in August 2019 with the preparation of the documents*
- *Award of the contract in July 2020*
- *Support by ParTec Project Management: Ina Schmitz*

■ Installation of the MeluXina supercomputer

- *Supported by ParTec Technical team (Jacopo de Amicis, Thomas Moschny and others)*



Speakers



Valentin Plugaru is the Chief Technology Officer of LuxProvide. With over a decade of passion for all things HPC, he has worked as part of national and European HPC initiatives, helping shape the roadmap for the European HPC ecosystem and create Luxembourg's first national HPC center and supercomputing platform. Today he is time-sharing between boosting LuxProvide's HPC, Data and AI capabilities, growing highly expert groups, and supporting the development of competencies networks, in particular through the EuroCC/National Competence Centers and CASTIEL projects.



Thomas Moschny has been working at ParTec since 2008, from 2013 as the Chief Technology Officer (CTO). He has a Diploma in Theoretical Particle Physics from the University of Wuppertal. From 2000 to 2008 he was a member of the group of Prof. Tichy at the CS Department of University of Karlsruhe (now Karlsruhe Institute of Technology, KIT), working on high performance communication software and parallel programming environments. In addition to the CTO responsibilities his main focus at ParTec now is on programming models and workload management topics of the Modular Supercomputing Architecture (MSA).



HIGH PERFORMANCE

Computing & Data Analytics

Valentin Plugaru, Chief Technology Officer

Our Mission

Leading-edge national supercomputing and data infrastructure

Provide HPC, HPDA, Big Data & AI commercial services

Empower Luxembourg digital ecosystem



Co-funded by the Government of Luxembourg



EuroHPC
Joint Undertaking

Hosting entity for EuroHPC Joint Undertaking



Part of the European Competence Centers initiative



A brief re-introduction of Supercomputing

Supercomputing for citizens

... **key role in medicine**: for discovering new drugs, developing and targeting medical therapies

... crucial to **understand** the generation and evolution of **epidemics and diseases**

... critical importance to **anticipate severe weather conditions** [...] key to **activate early warning systems** to **save human lives** and **reduce damages** to our properties and public infrastructures

... key to **monitor the effects of the climate change**

... **improving our knowledge** of geophysical processes, **monitoring earth resource** evolution, **reducing the environmental footprint** of industry and society or **supporting sustainable agriculture** through numerical simulations of plant growth.

... vital for national security, defence and sovereignty, as they are used to **increase cybersecurity** and **in the fight against cyber-criminality**, in particular for the protection of critical infrastructures

Supercomputing for industry



... **enables industrial sectors** like automotive, aerospace, renewable energy and health to innovate, become **more productive** and to **scale up to higher value products and services**

... growing impact on industries and businesses by **significantly reducing product design** and production cycles, **accelerating the design** of new materials, **minimising costs**, **increasing resource efficiency** and shortening and optimising decision processes

... paves the way to **novel industrial applications**: from safer and greener vehicles to more efficient photovoltaics, sustainable buildings and optimised turbines for electricity production

... will **make it easier for SMEs** without the financial means to invest in in-house skills **to develop and produce better products and services**

Supercomputing for science

... at the heart of the digital transformation of science, it **enables deeper scientific understanding and breakthroughs in nearly every scientific field**

... applications [...] are countless: from **fundamental physics** (advancing the frontiers of knowledge of matter or exploring the universe) to **material sciences** (designing new critical components for the pharmaceutical or energy sectors) and **earth science** (modelling the atmospheric and oceanic phenomena at planetary level)

Many recent breakthroughs would not have been possible without access to the most advanced supercomputers.

... for the Chemistry Nobel Prize winners in 2013, supercomputers were used to develop powerful computing programs and software, to **understand and predict complex chemical processes**

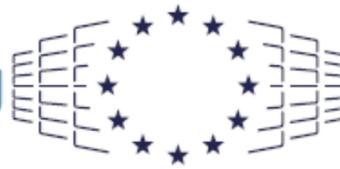
... for the Physics Nobel Prize in 2017 supercomputers helped to **make complex calculations to detect hitherto theoretical gravitational waves**

European context

EuroHPC JU

2019-2020: PUTTING EUROPE IN THE LEAD

The EuroHPC Joint Undertaking



EuroHPC
Joint Undertaking

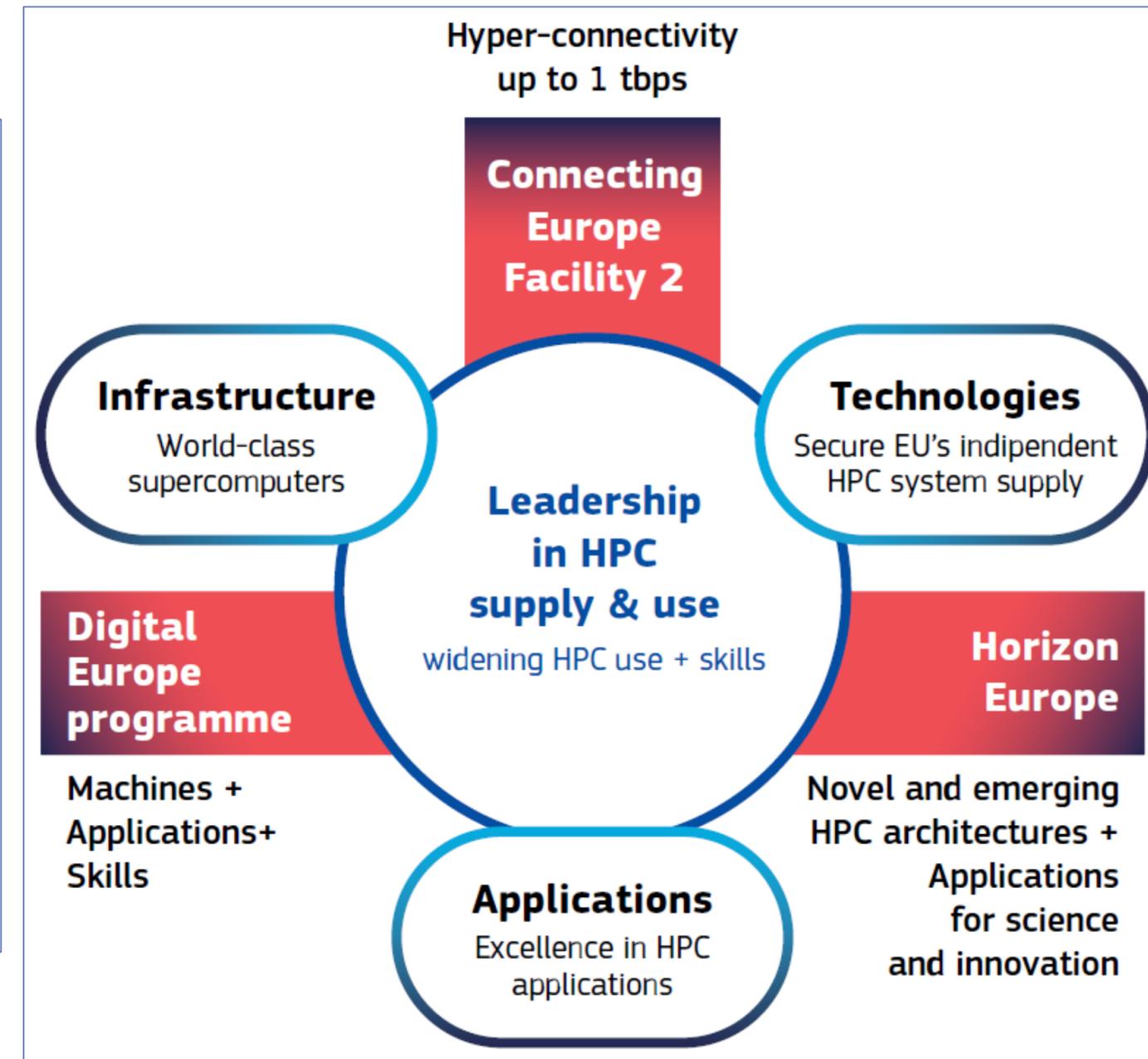
Legal and funding body created by the Council in September 2018



Mission: Pool European resources to establish an integrated world-class supercomputing & data infrastructure and support a highly competitive and innovative HPC and Big Data ecosystem.



- ➔ 29 Participating States + EU
- ➔ Site: Luxembourg
- ➔ Budget (2019-2020): €1.5 Billion
- ➔ Operational: 11/2018 to 2026



EuroHPC JU – Supercomputing infrastructure

INFRASTRUCTURE TO BRING EUROPE TO THE TOP 3 WORLDWIDE

- 3 EuroHPC supercomputers in top 5 in 2020.
- EuroHPC supercomputers will provide 4 to 5 times more computing power than the current systems available at European level.
- Several will include European technology.

RESEARCH & INNOVATION

MORE THAN €350 MILLION WILL BE INVESTED FOR:

- HPC and Big Data application platforms
- Extreme scale technologies and Industrial software codes
- National HPC Competence Centres
- Support to SMEs
- Support to the European Processor Initiative (EPI)



European Commission @EU_Commission

Replying to @EU_Commission

We want to invest 8 billion euros in the next generation of supercomputers. These advanced systems are capable of dealing with complex matters in health, energy, engineering, climate research and much more. #DigitalEU #SOTEU #EuroHPC

#SOTEU

- Monitor and mitigate the effects of climate change
- Search for new treatments
- Design safer and greener cars and aircrafts

2:55 pm · 18 Sep 2020 · Twitter Web App

EuroHPC_JU @EuroHPC_JU · 29 Sep

Today our Director met @Xavier_Bettel @FranzFayot & @luxprovide to celebrate the signature of the contract of #MeluXina, new EuroHPC #supercomputer ⚡

By 2021 MeluXina will support European #researchers & #industry, wherever they are in Europe 🇪🇺

eurohpc-ju.europa.eu/news/meluxina-...

EuroHPC_JU @EuroHPC_JU · 28 Sep

👉 Two years ago, @EUCouncil regulation 2018/1488 establishing the #EuroHPC Joint Undertaking was adopted!

The aim was to develop a world-class #HPC ecosystem in Europe ⚡ In this 1st phase, achievements were already made, stay tuned for the next steps

eurohpc-ju.europa.eu/sites/default/...

EuroHPC_JU @EuroHPC_JU · 25 Sep

#Fridayreads

For @ThierryBreton #Supercomputers ⚡ will be at the core of our European #DigitalDecade: from simulating drug molecules against the #pandemic, to monitoring the effects of #globalwarming & designing innovative products

More ec.europa.eu/commission/com...

#HPC

DigitalSingleMarket @DSMeu · 20 Nov

Supercomputers + #ArtificialIntelligence = enormous potential to:

- Develop new personalised medical treatments
- Reducing time to market for innovative products
- Predicting & managing effects of #climatechange

& more bit.ly/3nCT4rV

#frAlday #DeepLearning

AI

Artificial Intelligence & Supercomputing

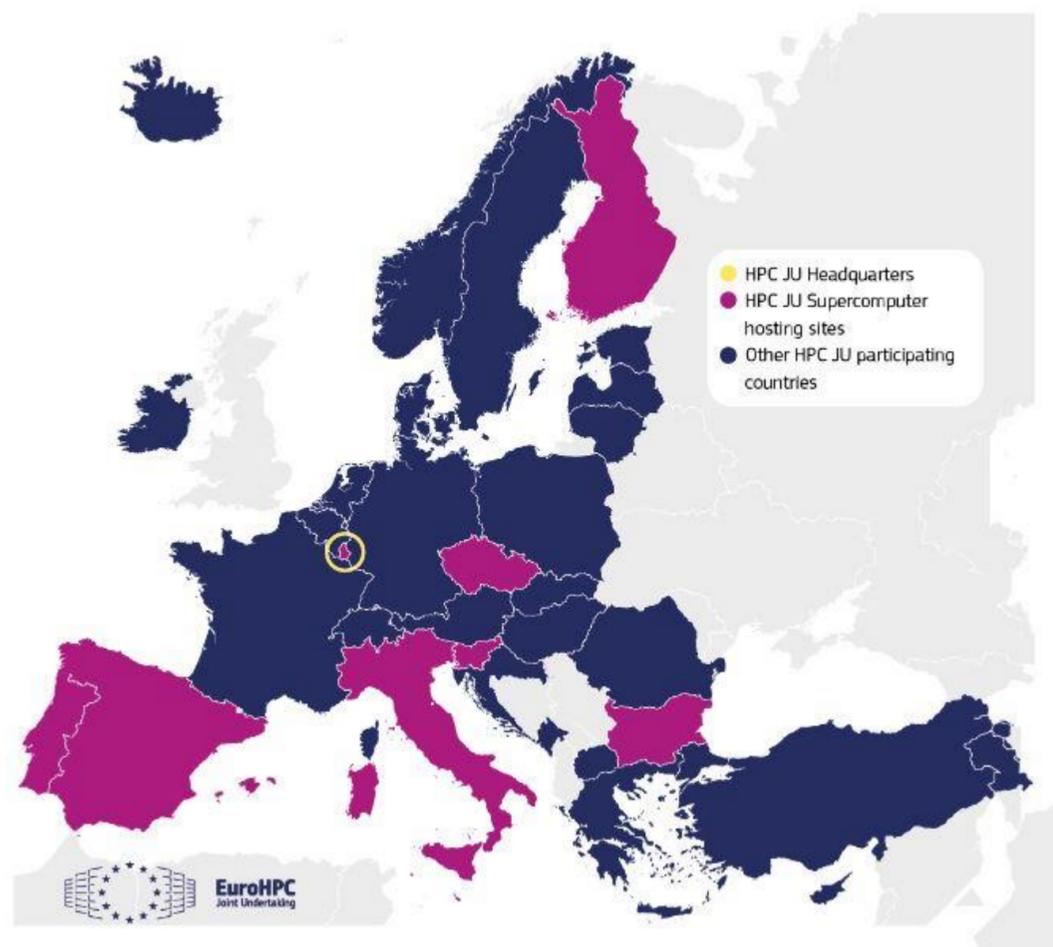
fosters the rapid development of **new HPC applications & services** across multiple sectors

#frAlday

MeluXina, part of the European #DigitalDecade

“Building on Europe's success in next-generation high-performance computing, **supercomputing** will play a key role in Europe's path towards recovery. It has been identified as a strategic investment priority, **and will underpin the entire digital strategy**, from big data analytics and artificial intelligence to cloud technologies and cybersecurity.”

State of the Union: Commission sets out new ambitious mission to lead on supercomputing
ec.europa.eu/commission/presscorner/detail/en/ip_20_1592



“... the MeluXina supercomputer will boost European innovation and competitiveness, and support European researchers and industry, wherever they are in Europe.”

MeluXina: a new EuroHPC world-class supercomputer in Luxembourg
eurohpc-ju.europa.eu/news/meluxina-new-eurohpc-world-class-supercomputer-luxembourg

EuroCC & CASTIEL Projects



MORE THAN INFRASTRUCTURE...

- EuroCC project about kickstarting **National (HPC, HPDA & AI) Competence Centers**
- 1 NCC per country, 33 participating countries
- **Luxembourg NCC:** LuxInnovation, the University of Luxembourg and LuxProvide

... ALL ABOUT EXPERTISE

- NCCs will provide **broad service portfolio** tailored to national needs
- ... of industry, academia & public administration
- LuxProvide leading work on **Technology Transfer, Business Development & Collaboration with Industry**
- ... will also organize trainings, workshops

... AND COORDINATION

- CASTIEL project will promote interaction & exchanges across NCCs

Find out more about EuroCC Luxembourg: eurocc-luxembourg.lu

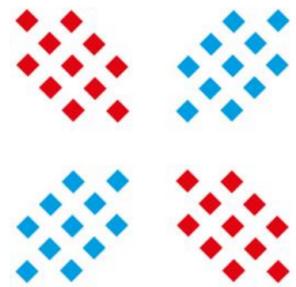


[linkedin.com/company/eurocc-luxembourg](https://www.linkedin.com/company/eurocc-luxembourg)



twitter.com/EuroCC_Lux

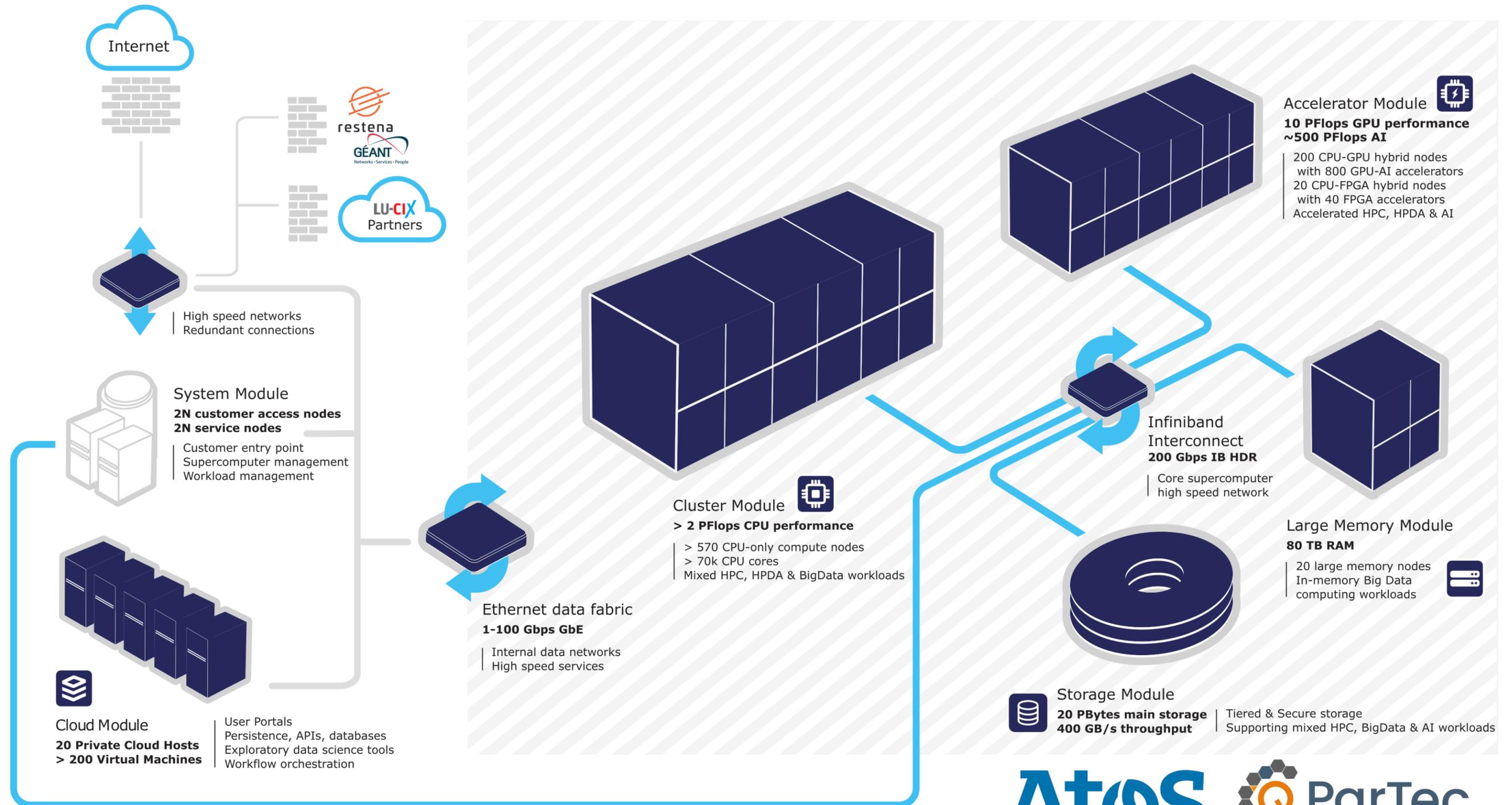




MELUXINA – a modular supercomputer

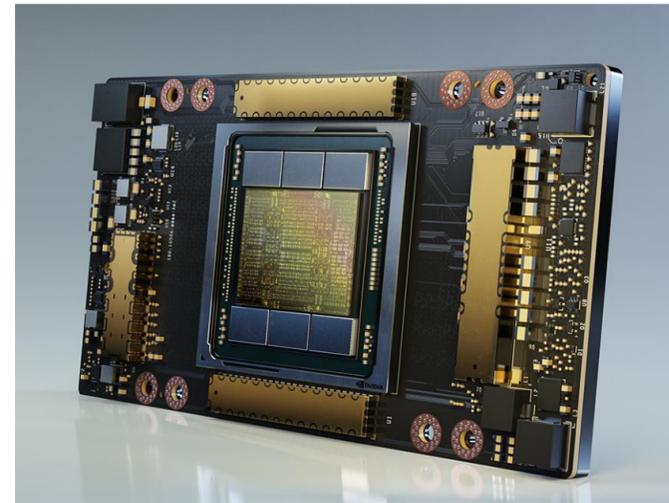
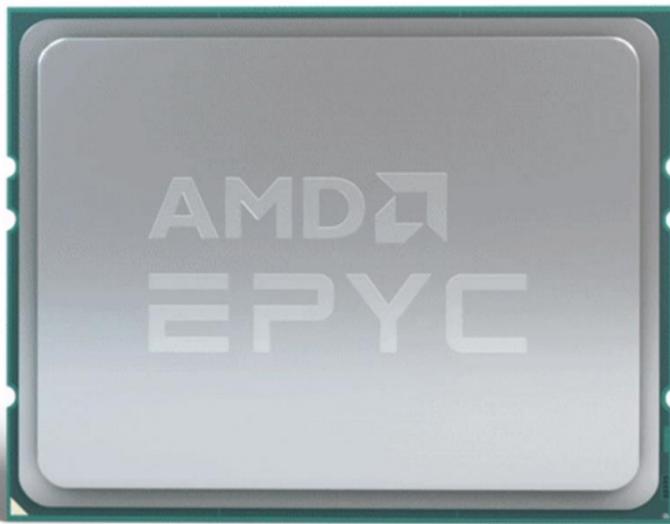
HIGH PERFORMANCE
COMPUTING IN LUXEMBOURG

MeluXina – architecture



MeluXina – building blocks

**OVER 1600 HPC CPUs
90K CPU CORES**



**800 HPC & AI ACCELERATORS
5.5M CUDA & 345K AI CORES**

**LIVE MEMORY
OVER 476 TB RAM**



**20 PB HPC & AI
TIERED DATA STORAGE
+5 PB LONG-TERM LIBRARY**

**INFINIBAND HDR
DRAGONFLY+ TOPOLOGY**



**HPC, HPDA & AI
SOFTWARE ECOSYSTEM**

MeluXina – performance

Compute 18 PFlops

Peak aggregated performance
Modular supercomputer architecture

Cluster	Accel. GPU	Accel. FPGA	Large Memory
573 nodes	200 nodes	20 nodes	20 nodes
2.25 PF HPL (CPU)	10 PF HPL (GPU)	40 FPGAs	80 TB RAM
All-purpose	Accelerated workloads	Specialized workloads	In-memory workloads

Data 25 PBytes

Aggregated capacity
Multi-tier storage architecture

Tier0	Tier1	Tier2	Tier3
0.5 PB all-flash	12.5 PB	7 PB	5 PB
400 GB/s	190 GB/s		
Very intensive IO	All-purpose	Backup	Long-term

Interconnect HDR 200G

Single or Dual-rail HDR in DragonFly+
High speed internal/external networks

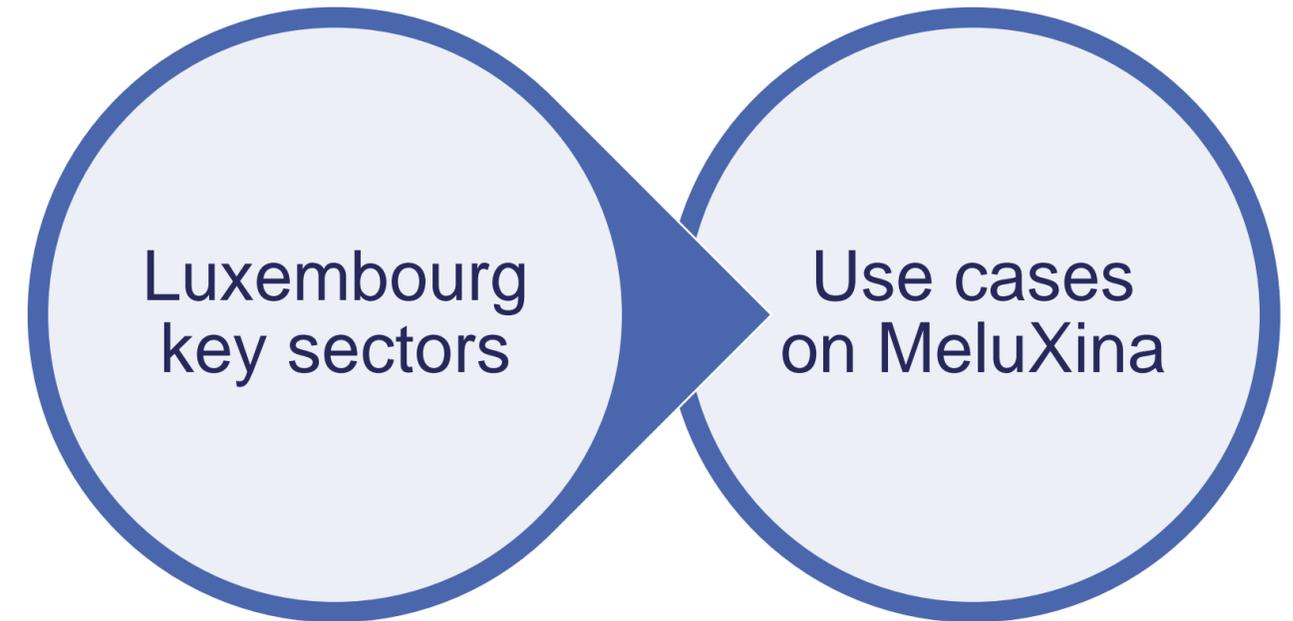
Internal	GEANT (RESTENA)	ISPs	Partners
Very high bandwidth	Very high bandwidth	High bandwidth	High bandwidth
	EuroHPC sites		
	Public sector		
IB HDR Eth.100G	Research & start-ups	All-purpose	Dedicated connections
			15

MeluXina – hosting facility

- Hosted in **LuxConnect** data centers, main MeluXina Modules in Bissen DC2
- **UPTIME TIER IV** and **ISO 27001** certified
- Powered with 100% **GREEN ELECTRICITY** from hydroelectric power sources

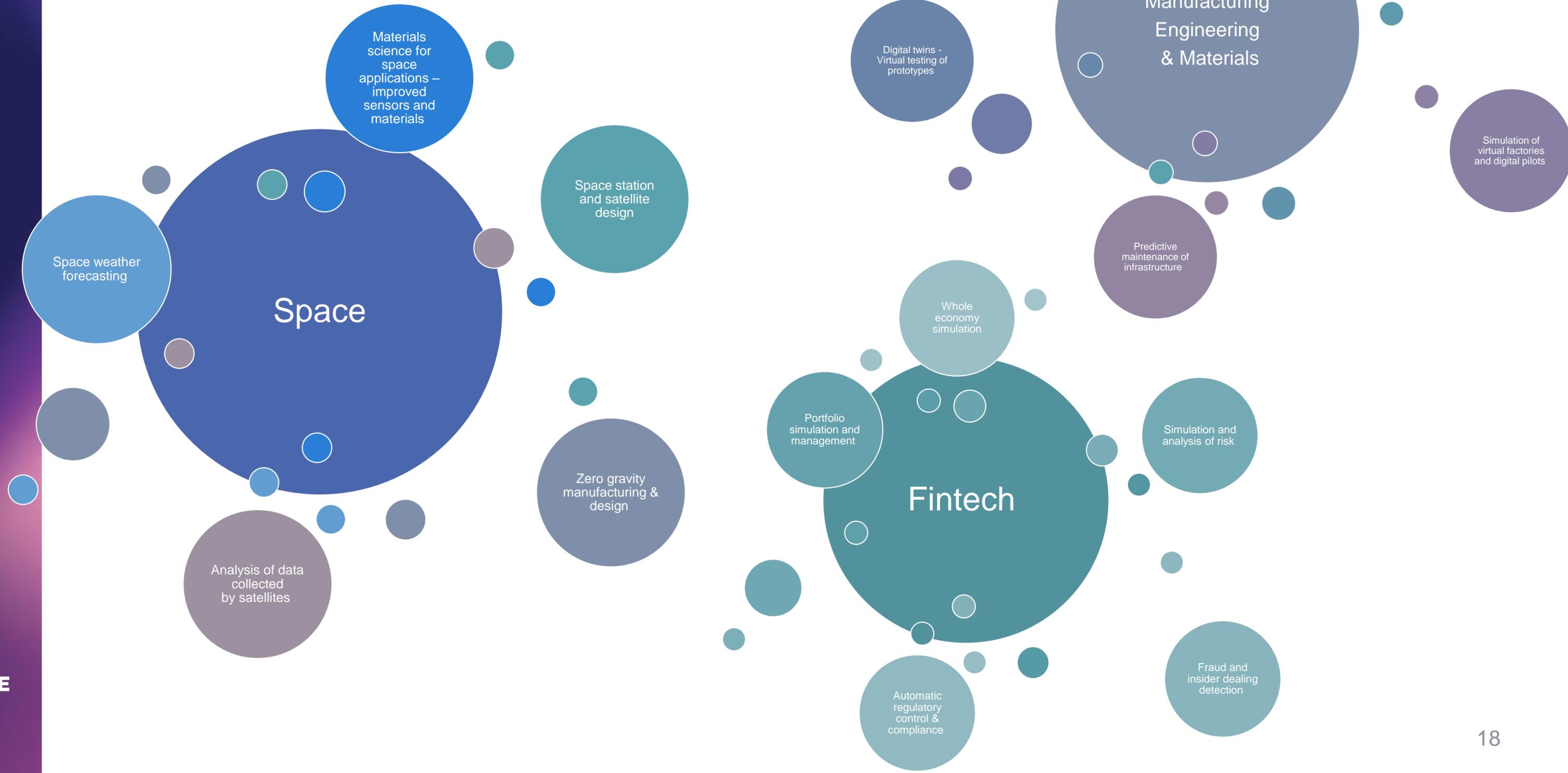


Vision for MeluXina & target applications

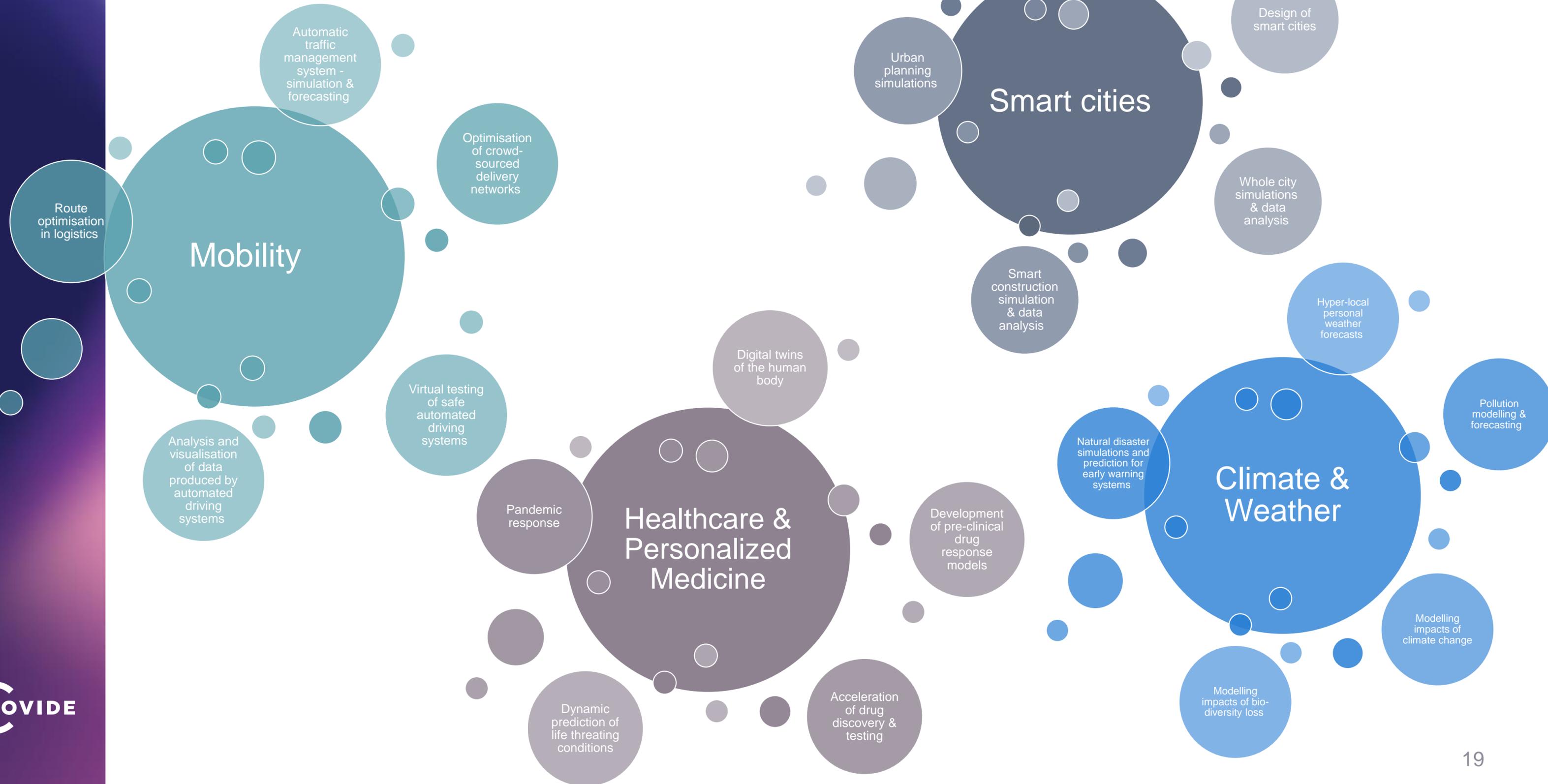


- Smart Mobility
- Financial Services
- Healthtech
- Industry 4.0
- Logistics
- Space
- Traditional **HPC modelling & simulation**
- **HPDA data-driven workloads**
- **AI data-driven workloads & HPC with AI-in-the-loop**

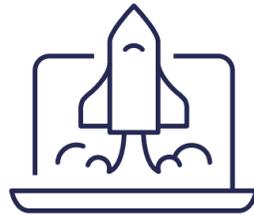
Where can MeluXina help?



Where can MeluXina help?



Our HPC Value Proposition



HPC facility

- **Leading-edge** computing and data platforms
- Low latency **interconnect**
- Specialized software ecosystem for **large-scale workloads**
- Technical support



High value-added services

- **Solution** engineering by HPC & domain experts
- **Software engineering** by experienced engineers
- Including optimization, profiling and parallelization



Focus on customer

- **Tailored packages:** connectivity, compute, storage and expertise
- Guidance throughout the whole process, **dedicated team**

Our HPDA & AI Value Proposition



Dedicated Cloud Module

- **Automation** of complex processes
- Data workflows, data services
- Dedicated **support**



Specialized Data Analytics & AI hardware

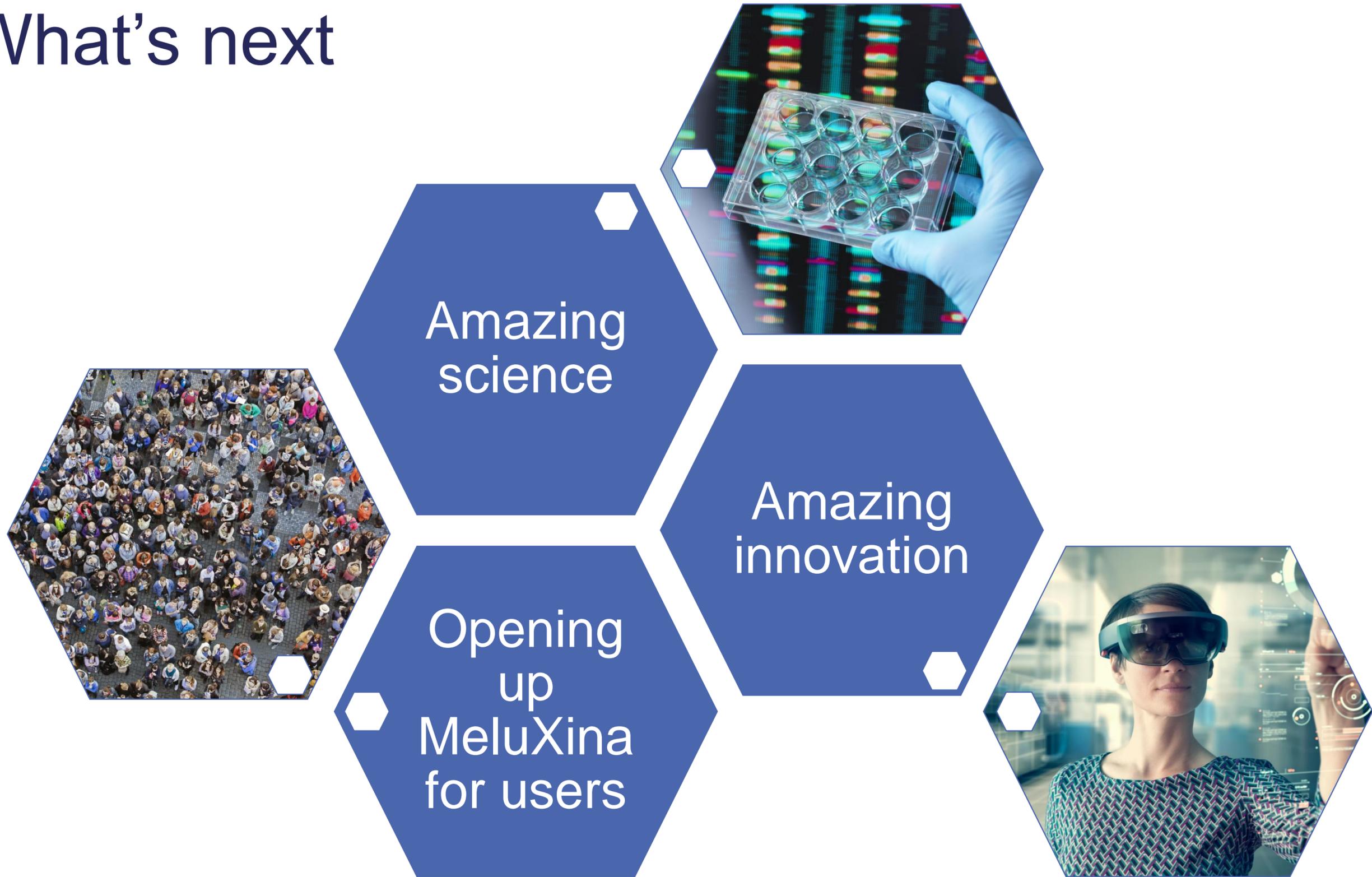
- State-of-the-art GPU-AI accelerators and large memory systems
- FPGAs for customized **data processing**
- Fast storage and low latency network for efficient communication



Solution Engineering

- Design solutions to **extract value and generate insights** from data
- Define workflows for data generation and retention
- Help to unveil **new business cases**

What's next



MeluXina implementation in pictures





MeluXina core compute - 1st power-up

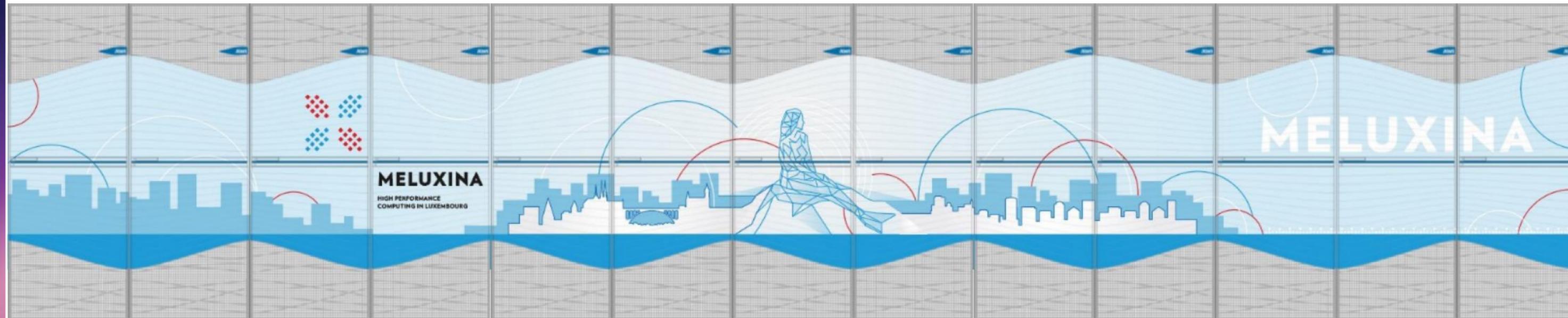
Thank you for your kind attention!

Valentin Plugaru

ETP4HPC Webinar, 2021-05-28

Sparked your interest?

For more information, reach us at info@lpx.lu



Acknowledgements & Disclaimer

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 program, as well as the Grand Duché du Luxembourg.

This publication only reflects the authors' view and the EuroHPC Joint Undertaking is not responsible for any use that may be made of the information it contains.

Names and logos used in this presentation are property of their respective owners.

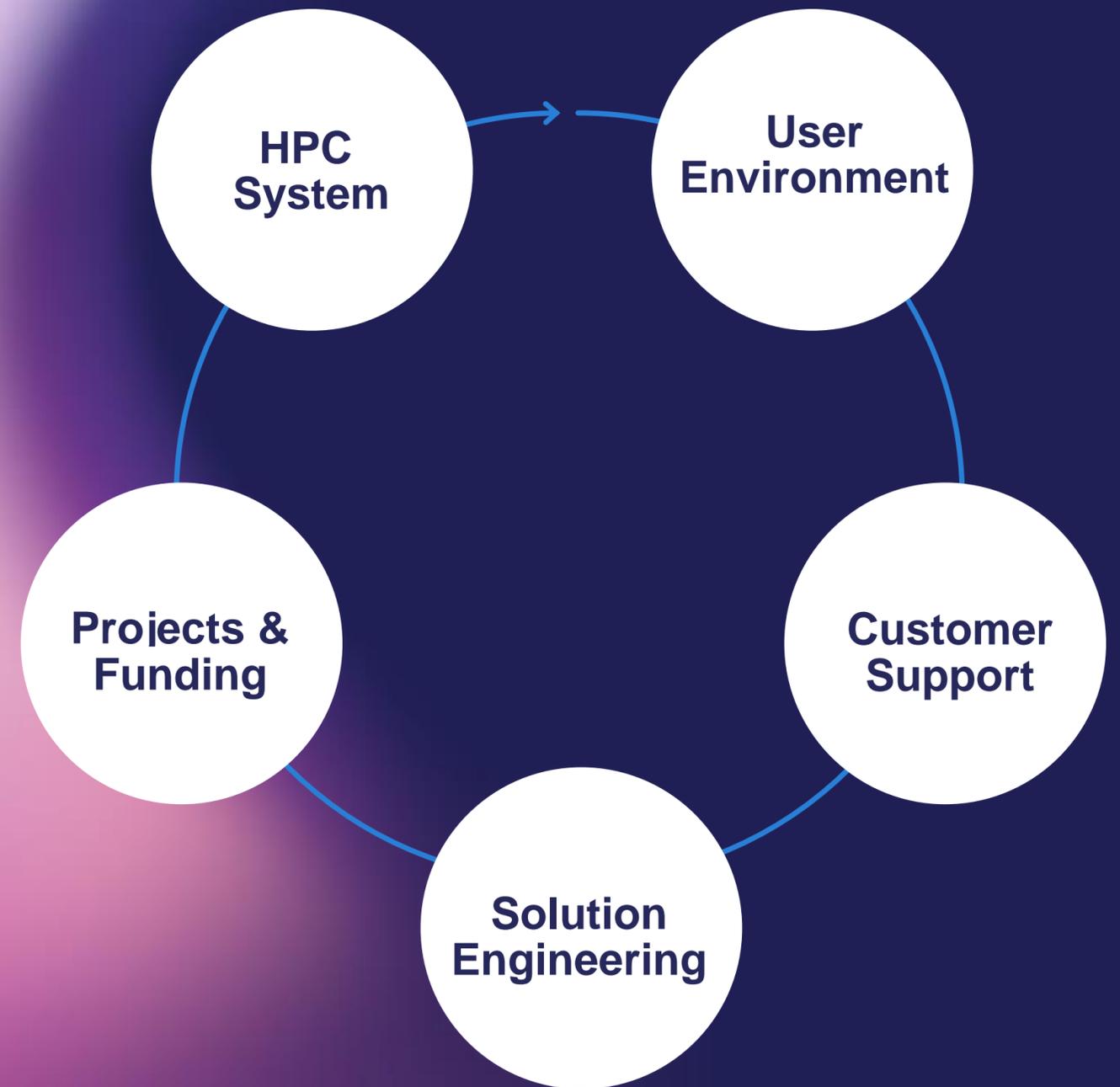


EuroHPC
Joint Undertaking





Luxembourg's one-stop shop for
high performance
computing and data analytics



Speakers



Valentin Plugaru is the Chief Technology Officer of LuxProvide. With over a decade of passion for all things HPC, he has worked as part of national and European HPC initiatives, helping shape the roadmap for the European HPC ecosystem and create Luxembourg's first national HPC center and supercomputing platform. Today he is time-sharing between boosting LuxProvide's HPC, Data and AI capabilities, growing highly expert groups, and supporting the development of competencies networks, in particular through the EuroCC/National Competence Centers and CASTIEL projects.



Thomas Moschny has been working at ParTec since 2008, from 2013 as the Chief Technology Officer (CTO). He has a Diploma in Theoretical Particle Physics from the University of Wuppertal. From 2000 to 2008 he was a member of the group of Prof. Tichy at the CS Department of University of Karlsruhe (now Karlsruhe Institute of Technology, KIT), working on high performance communication software and parallel programming environments. In addition to the CTO responsibilities his main focus at ParTec now is on programming models and workload management topics of the Modular Supercomputing Architecture (MSA).

ParaStation Modulo and the MSA

ETP4HPC Webinar –
MeluXina, the Luxemburg EuroHPC System

May 28, 2021

Thomas Moschny, ParTec

- Strong general purpose cluster specialist for more than two decades
 - *ParaStation research project: 1995 (Univ. of Karlsruhe)*
 - *ParTec founded as a spin-off in 1999*
 - *HPC full service provider since 2004*
- Cooperation with Jülich Supercomputing Centre since 2004
 - *ParaStation Consortium founded in 2005*
- Engagement and active development in projects towards Exascale
 - *Since 2010: DEEP Projects*
- Pioneering and enabling Modular Supercomputing
 - *ParaStation Modulo Software Suite: 2017*
- Technology partner for LuxProvide in the MeluXina project

- **ParaStation Modulo Software Suite**
 - *Software for HPC Systems developed since >20 years*
 - *Pioneering the Modular Supercomputing Architecture since >10 years*
 - *Extensively used in production environments*
 - *Platform for research projects*
- **ParTec Support: on-site (or remote) system operations**
 - *System setup and installation*
 - *System maintenance and administration*
 - *General 1st and 2nd level support*
- **Co-design and co-development**
 - *Transferring results from research projects into production*
 - *Enhancing production systems over their lifetime*
- **Significant contributions in European research projects**
 - *Exascale-related: DEEP-Projects, *-SEA Projects, EUPEX*
 - *Quantum- and AI-related: HPCQS, CoE RAISE*

ParaStation
MODULO



- ParaStation ClusterTools
 - *Tools for provisioning and management*
- ParaStation HealthChecker & TicketSuite
 - *Automated error detection & error handling*
 - *Ensuring integrity of the computing environment*
 - *Keeping track of issues*
 - *Powerful analysis tools*
- ParaStation MPI & Process Management
 - *Runtime environment specifically tuned to the largest distributed memory supercomputers*

ParaStation
MODULO

**Maximize job throughput -
Minimize administration effort**

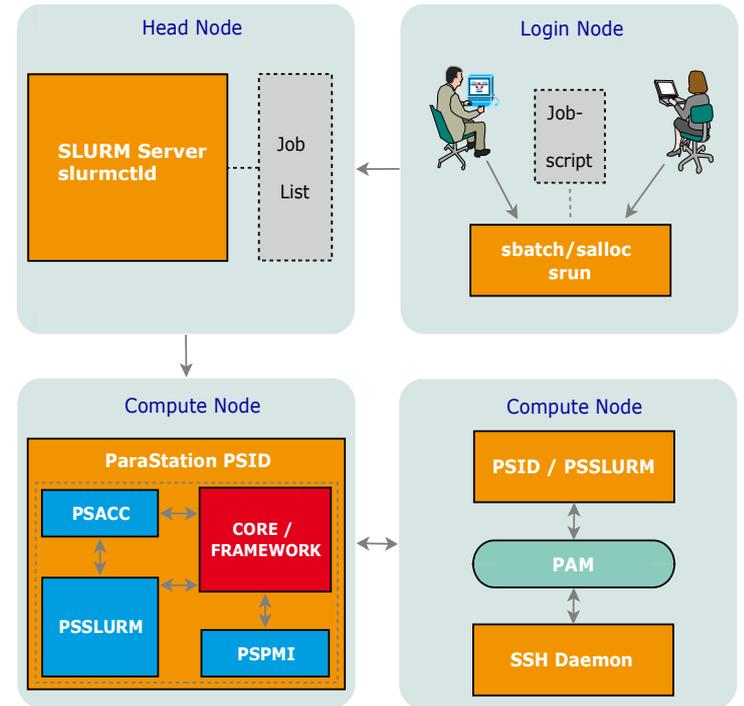
- Install / update all types of nodes; on-disk and diskless
- Image-based, local modifications after cloning supported
 - *Supports multiple images per node*
 - *Rolling update supported*
 - *Synchronization for configuration changes*
- Automatic PXE boot of nodes for system install or system diagnostics
- Set of standard CLI tools for managing the system
 - *Configuration and image handling*
 - *Console redirection and other IPMI functionality*
 - *Provides uniform interface, hides away HW differences*
- Tight integration with HealthChecker and TicketSuite
 - *High degree of automation*

ParaStation
CLUSTER**TOOLS**

- Test framework with generic, fast, and non-destructive tests
 - Ensures node/system is “ready to run a job” – also covers **degraded** components
 - Checks hardware and software conditions on trigger (before/after a job, on reboot, manual stress-test...); or periodically
- Node Local Checking (NLC)
 - Runs on each node, checks health of this node (local and remote parameters)
 - One independent instance per node → infinite scalability
- Cluster Global Checking (CGC)
 - Runs on some node (usually an admin node)
 - Checks health of other components
- Actions aiming at ...
 - ... keeping the system usable: Disable failing/degraded components; reschedule jobs
 - ... fixing problems: Create tickets; send notifications

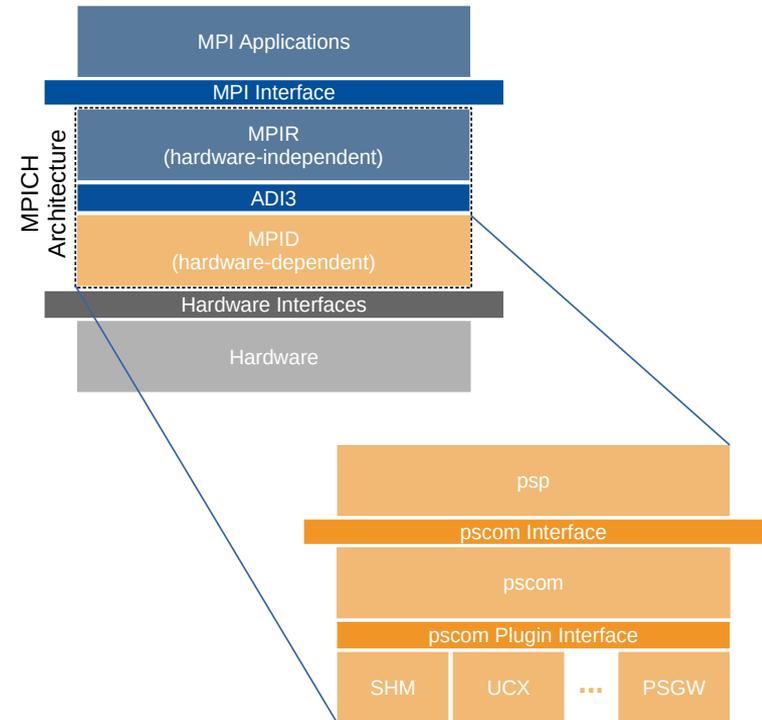
ParaStation
HEALTH**CHECKER**

- Scalable network of MPI process management daemons
 - *Running on the computational nodes*
 - *Process startup and control, I/O forwarding, ...*
 - *Precise resource monitoring*
 - *Proper cleanup after jobs*
- **PSSLURM: Full integration for Slurm**
 - *Plugins to the ParaStation Management daemons*
 - *Replace node-local Slurm daemons (also reduces number of daemons)*
 - *Enforces resource limits*



ParaStation MPI

- Based on MPICH 3.3.2 (MPI-3.1 compliant)
 - Maintains MPICH ABI compatibility
 - Supports MPICH tools (tracing, debugging, ...)
 - MPICH layers beneath ADI3 are replaced by ParaStation PSP Device
 - Powered by pscom low-level communication library: non-blocking p2p semantics
- Support for various transports and protocols via pscom plugins
 - Support for InfiniBand, Omni-Path, Extoll, ...
 - Applications may use multiple transports / plugins at the same time
 - Gateway capability via PSGW plugin to bridge transparently between different networks
 - CUDA awareness for all transports / CUDA optimization via GPUDirect for UCX, and Extoll
- Proven to scale up to ~3,500 nodes and ~140,000 processes per job



ParaStation
MPI

- Project Series

- DEEP Cluster-Booster
- DEEP-ER I/O, resiliency
- DEEP-EST Modular Supercomputing Architecture

- Co-Design

- Hardware
- Software
- Applications

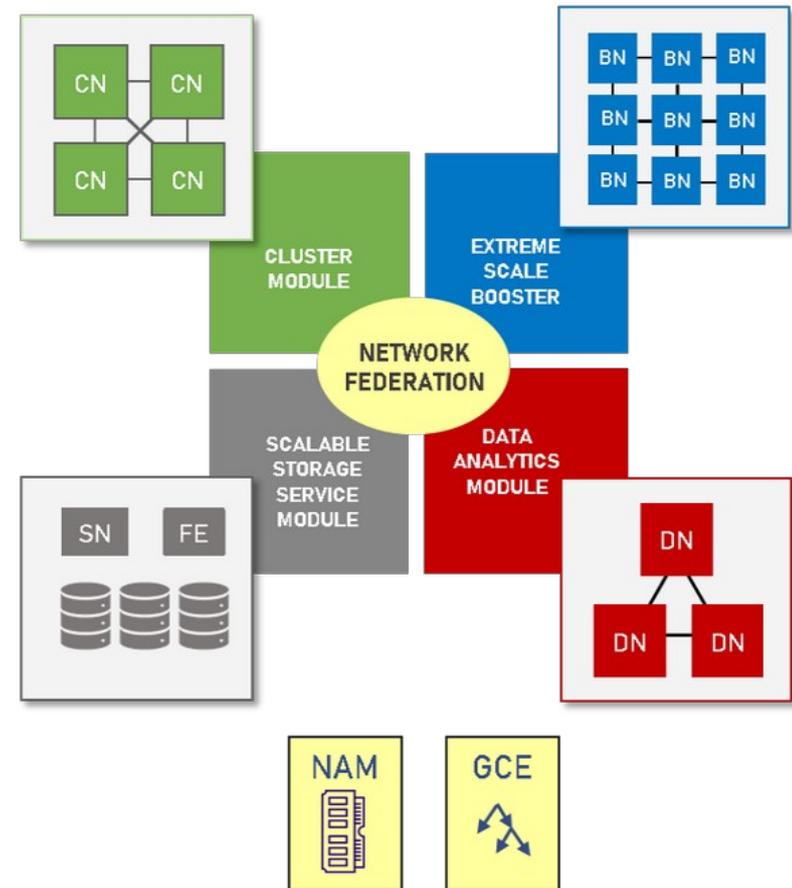
- Time frame: 2011-2021

- Total funding: 30 M€

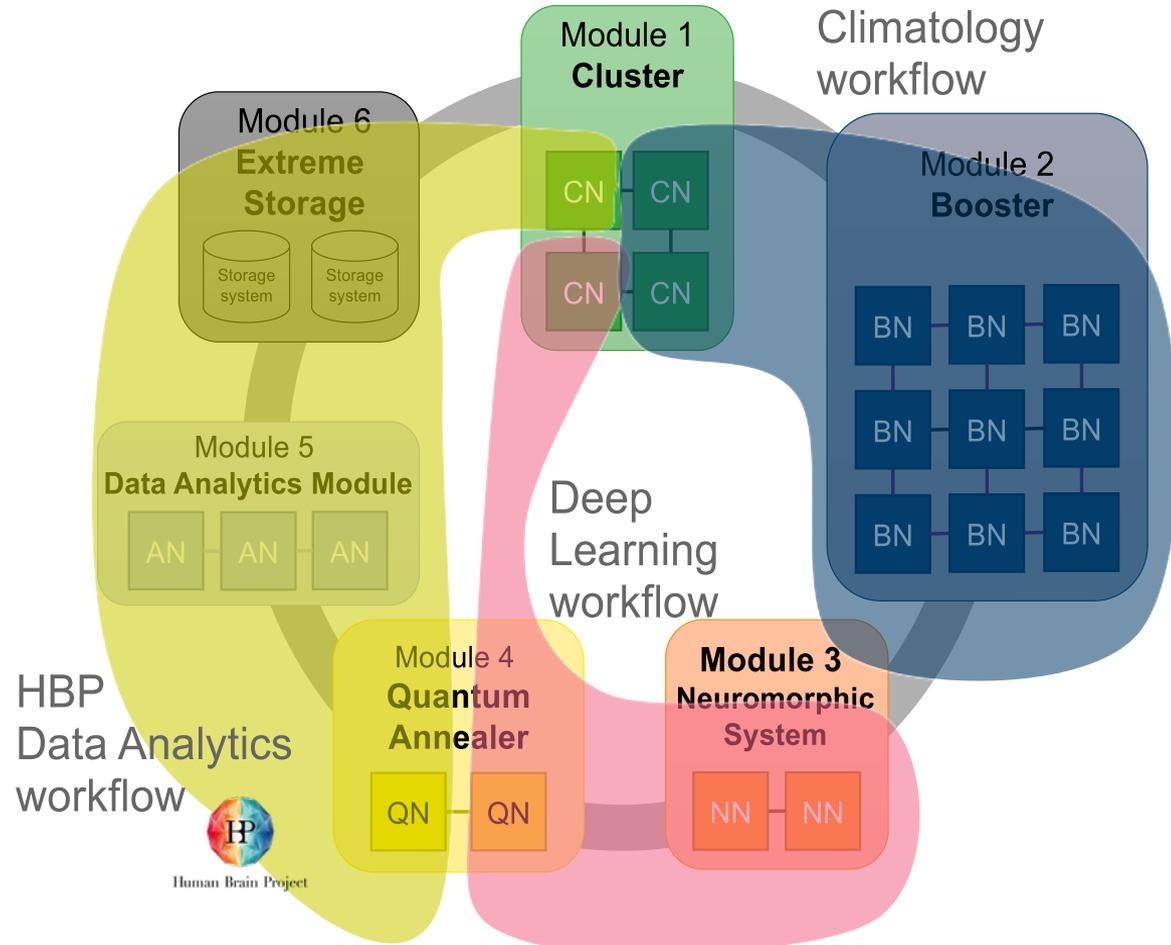


- Generalization of the Cluster-Booster Concept
- Address diverse user requirements in one system
 - “Data intensive”
 - “Highly scalable parallel computing”
 - “High-throughput computing”
- Any number of specialized modules possible
- Achieve leading scalability & energy efficiency → Exascale
- Unified SW environment to run applications across all modules
 - *Slurm & ParaStation Modulo*

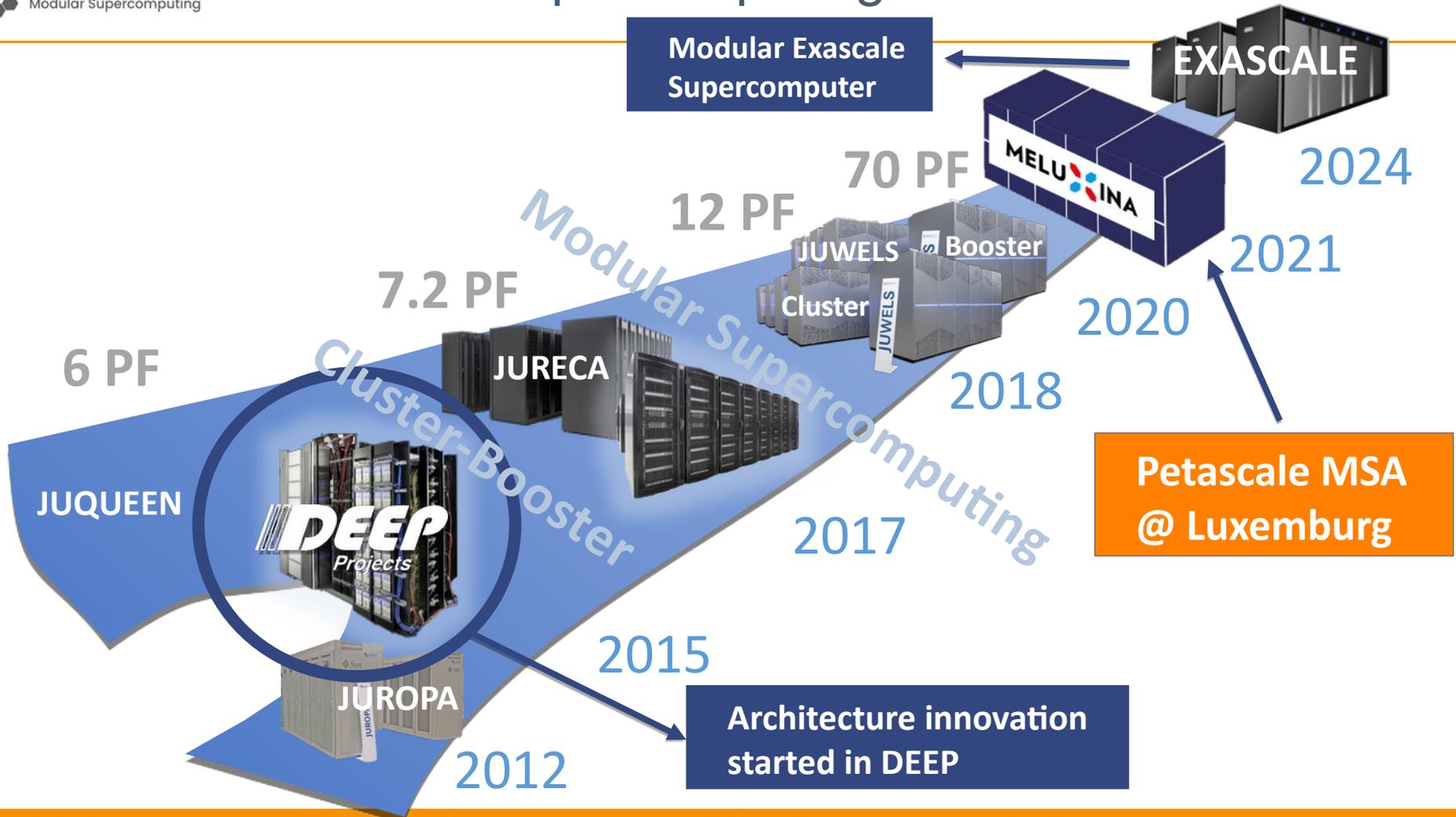
Integrate **diverse technologies**, combining the best of each



Generalization of the Cluster-Booster Concept



Modular Supercomputing



Operated as a single, heterogeneous system using ParaStation Modulo
#29 in Top500 11/2017



- Cluster: Production started 10/2015
- 1872 compute + 64 fat nodes (T-Platforms V-Class / SuperMicro)
- Intel Xeon E5-2680v3 (Haswell) / 128+ GB
- Mellanox Infiniband EDR
- 1.8 (CPU) / 0.44 (GPU) PFlop/s peak
- #49 of Top500 list (Nov 2015)

- Booster: Production started 11/2017
- 1640 Booster nodes (Intel SSF / Dell PowerEdge C6230P)
- Intel Xeon PHI 7250-F
- Intel Omni-Path (OPA)
- 5 PFlop/s peak
- Intel/Dell/JSC/ParTec Cooperation

JURECA Cluster-Booster



- JURECA DC replacing Cluster Module
- Production start 05/2021
- 768 compute nodes (Atos/Bull XH2000)
- Dual-socket AMD EPYC Rome
- 192 nodes acc. with 4×A100 Nvidia GPU
- Mellanox InfiniBand HDR, Dragonfly+
- 18.52 PFlop/s peak



- Booster: Production started 11/2017
- 1640 Booster nodes (Intel SSF / Dell PowerEdge C6230P)
- Intel Xeon PHI 7250-F
- Intel Omni-Path (OPA)
- 5 PFlop/s peak
- Intel/Dell/JSC/ParTec Cooperation

Operated as one Modular System with ParaStation Modulo and Slurm

JUWELS – A Modular Supercomputer

Cluster Module



Booster Module



JÜLICH
SUPERCOMPUTING
CENTRE

- 12 PFlop/s peak
- #23 on Top500 list (June 2018)
- 2575 nodes (Bull Sequana X1000)
- Intel Xeon Platinum 8168 / Gold 6148
- Mellanox EDR, ParaStation MPI



- GPU-accelerated module, 70 PFlop/s peak
- #7 on Top500, #3 on Green500 (Nov. 2020)
- 936 nodes (Bull Sequana XH2000)
- 4x NVIDIA A100 GPUs per node
- Quad-rail Mellanox HDR200, ParaStation MPI

Operated as one Modular System with ParaStation Modulo and Slurm

Thank you!

Questions

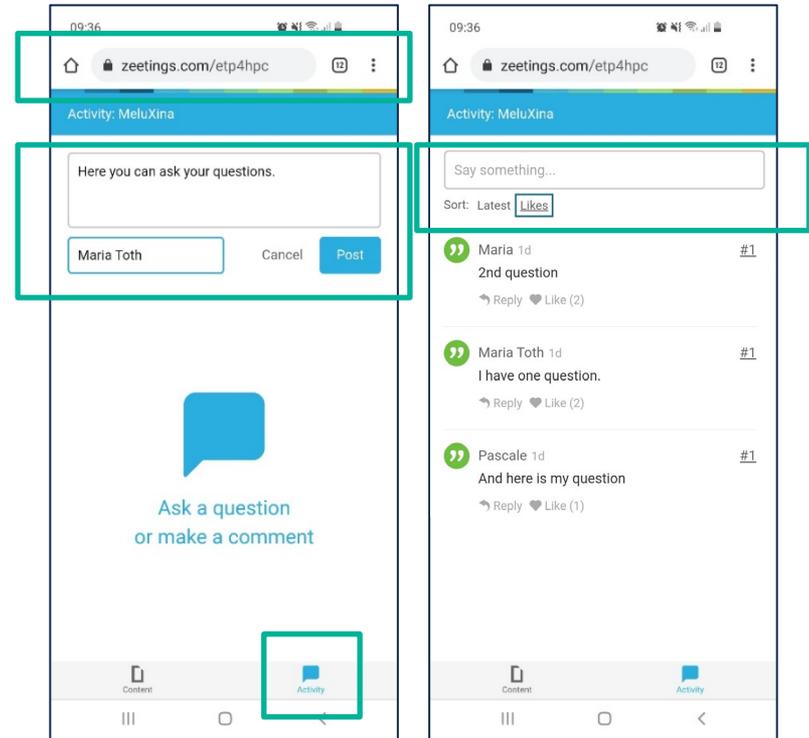
- You can ask questions to the speakers via your phone, tablet or computer typing into your browser:

www.zeetings.com/etp4hpc

- by clicking on the „Activity’ tab
- or on this icon:



- Please indicate your name





**EUROPEAN TECHNOLOGY
PLATFORM FOR HIGH
PERFORMANCE COMPUTING**

THANKS!

Please fill in the survey to help us improve our webinars!

Next webinar:

- 18 June, Friday, 11-12am CET
[Industrial Use of HPC: the standpoint of SMEs](#)

You can find us at:

[@etp4hpc](#)

office@etp4hpc.eu

www.etp4hpc.eu