LUX PROVIDE

HIGH PERFORMANCE
Computing & Data Analytics

Valentin Plugaru, Chief Technology Officer
Robert Mijaković, HPC System Software Architect

6th EasyBuild User Meeting, Jan 28th 2021

Our Mission

Leading-edge national supercomputing and data infrastructure

Provide HPC, HPDA, Big Data & AI commercial services

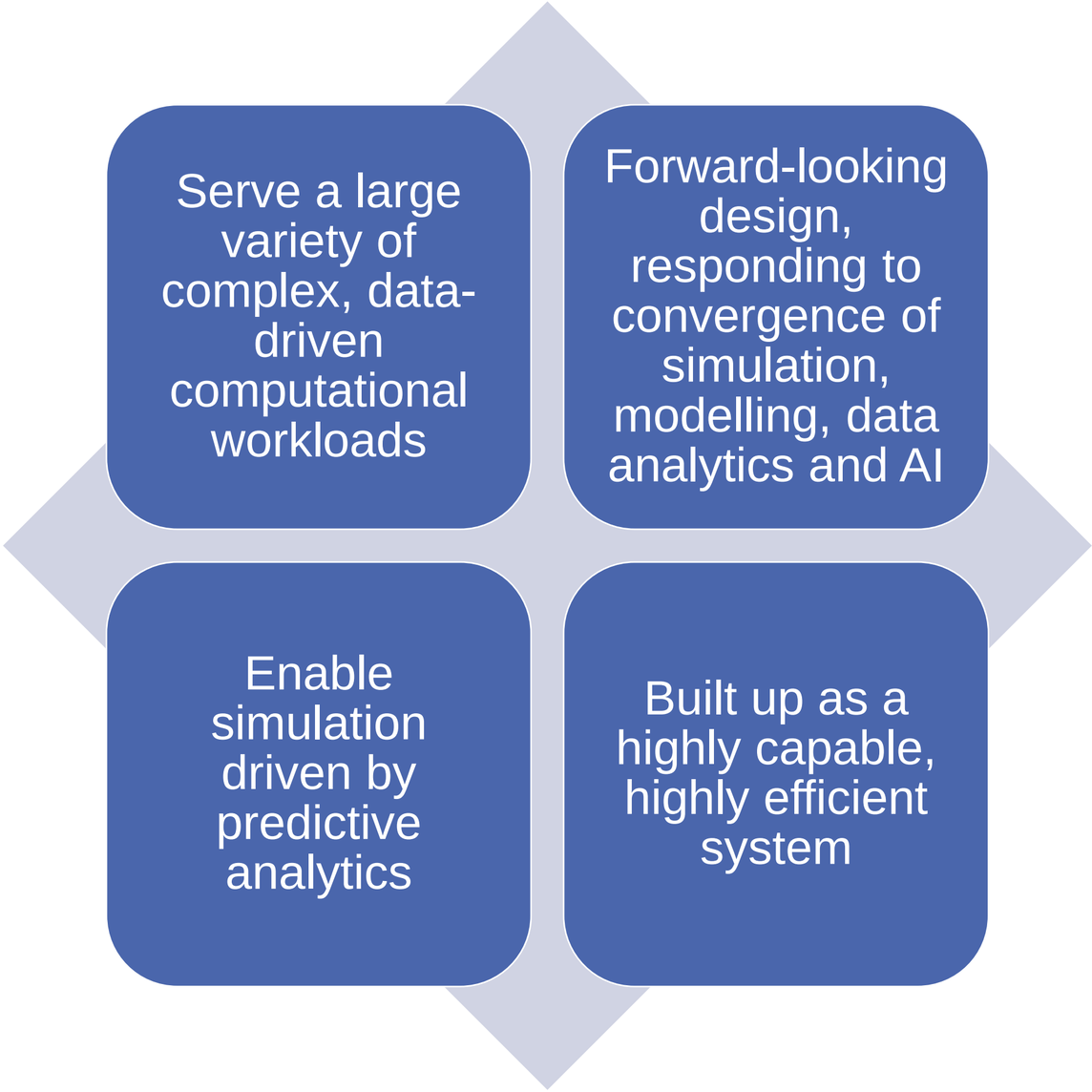
Empower Luxembourg digital ecosystem

Co-funded by the Government of Luxembourg

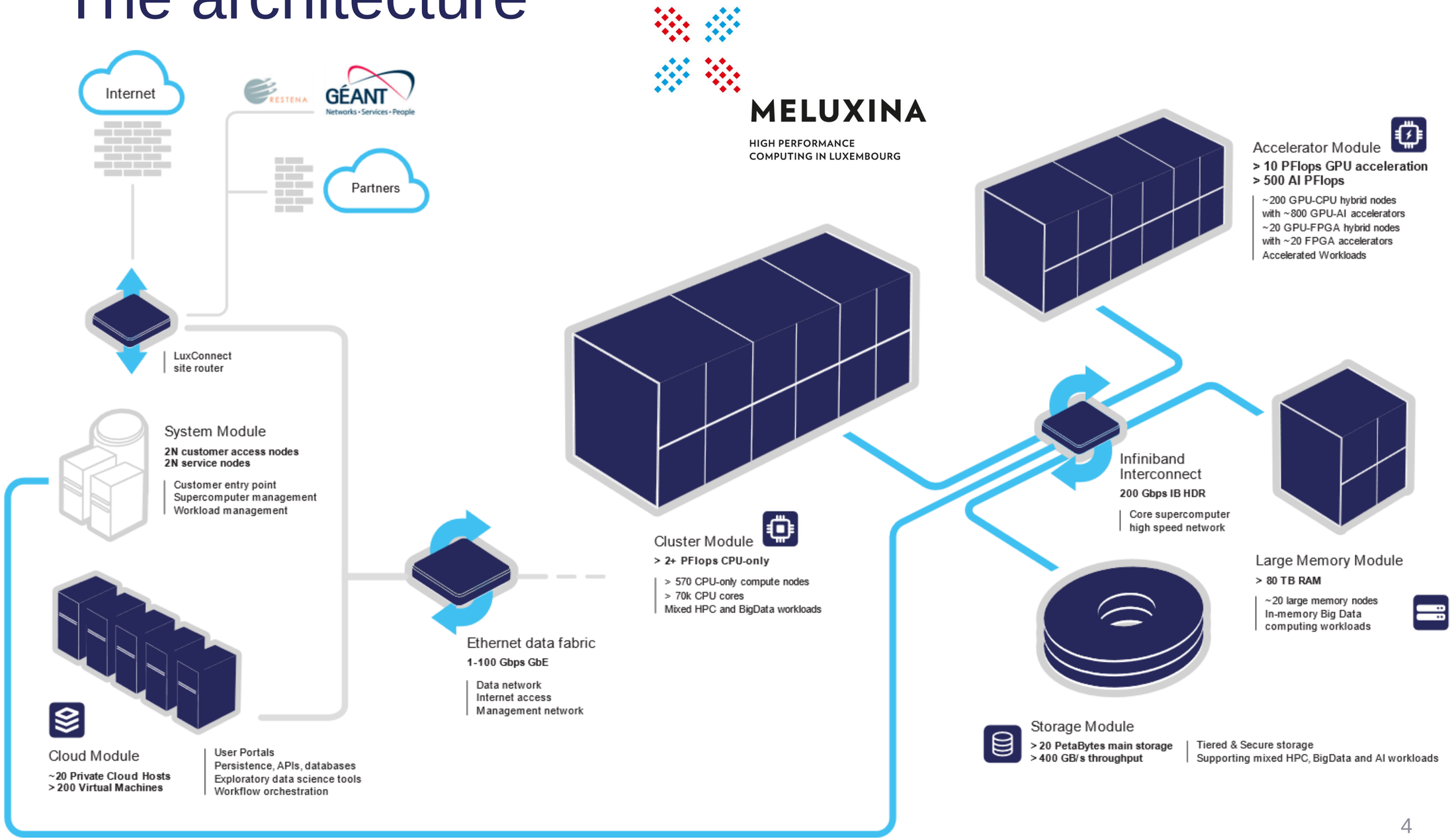
Hosting entity for EuroHPC Joint Undertaking

Part of the European Competence Centers initiative

MeluXina supercomputer – the vision



The architecture



The numbers

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	Backup	Long-term
Very intensive IO	All-purpose		

Interconnect
HDR 200G

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

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

LUX PROVIDE

The technologies

CPU nodes

CPU & RAM

- AMD EPYC 7H12
- 2x 64C @ 2.6GHz
- 512 GB (~ 4 GB / core)

Fabric

- HDR200

System type

- DLC: BullSequana XH2000

GPU nodes

CPU & RAM

- AMD EPYC 7452
- 2x 32C @ 2.35GHz
- 512 GB (~ 8 GB / core)

Accelerators

- NVIDIA Ampere
- 4 x A100 (40GB HBM2)

Local storage

- 1.92 TB SSD

Fabric

- 2x HDR200

System type

- DLC: BullSequana XH2000

Large Memory

CPU & RAM

- AMD EPYC 7H12
- 2x 64C @ 2.6GHz
- 4096 GB (~ 32 GB / core)

Local storage

- 1.92 TB SSD

Fabric

- 2x HDR200

System type

- AirCooled: BullSequana

The technologies – cont’d

FPGA nodes

CPU & RAM
- AMD EPYC 7H12
- 2x 64C @ 2.6GHz
- 512 GB (~ 4 GB / core)

Accelerators
- Intel Stratix FPGA
- 2x Stratix 10MX (16GB HBM)

Local storage
- 1.92 TB SSD

Fabric
- HDR200

System type
- DLC: BullSequana XH2000

Cloud nodes

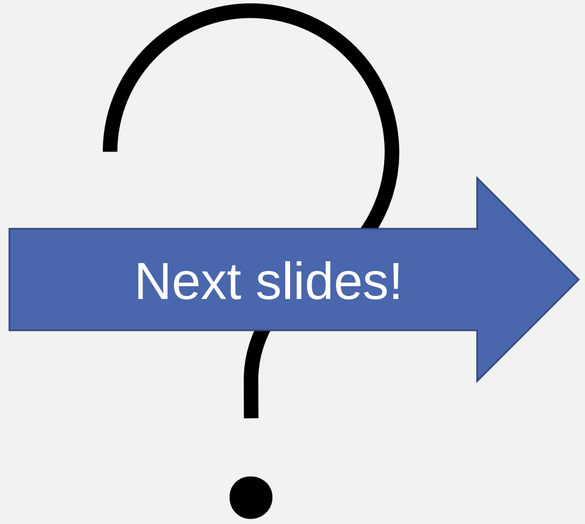
CPU & RAM
- AMD EPYC 7H12
- 2 x 64C @ 2.6GHz
- 512 GB (~ 4 GB / core)

Local storage
- 1.92 TB SSD

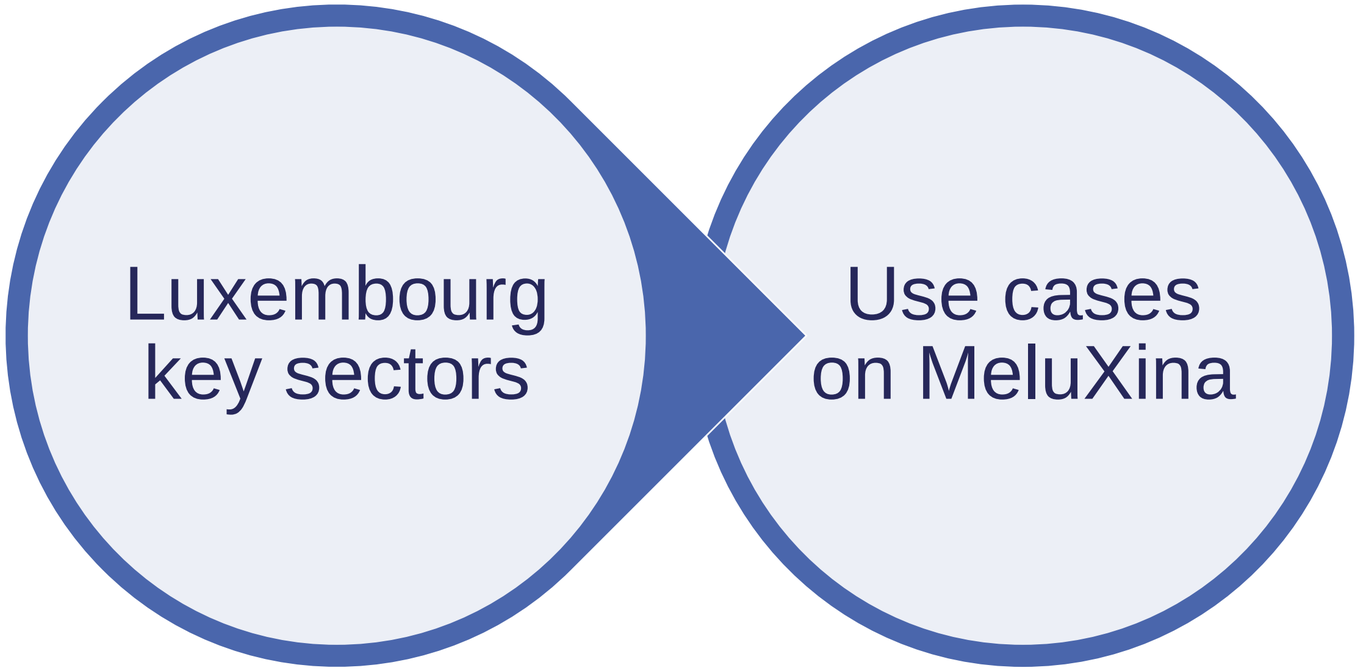
Fabric
- 2x HDR200

System type
- AirCooled: BullSequana

Software stack



The workloads



- Circularity, Smart Mobility
- Financial Services
- Healthtech
- Industry 4.0
- Logistics
- Space

- Traditional **HPC modelling & simulation**
- **HPDA data-driven workloads**
- **AI data-driven workloads** & also **HPC with AI-in-the-loop**

The *MUSE*: MeluXina User Software Environment

Conceptual goals

- Rich software environment – enable wide variety of workloads
  FOSS & commercial packages
- Keep software close to recent versions
- Maintain production software sets for a ‘good’ amount of time

Challenges

- Complex software environment
  MeluXina - modular system with multiple architectures: Zen2, Ampere, Stratix
- Keep software close to recent versions
  Validate functionality & performance
- Maintain production software sets for a ‘good’ amount of time

Meluxina User Software Env. - cont'd

Software deployment

- Main provisioning toolkit: **EasyBuild**
  - Also enable easy integration for Spack users
  - Keep an eye on EESSI project evolution & lessons
- Lmod modules system
  - Hierarchical naming scheme from EB

EasyConfigs,Blocks & Spack packages from community

- Adapt, upgrade & contribute back as much as possible

1st production environment on MeluXina

- EB toolchains: FOSS-2020b, AMD
- Initial set of common HPC compilers, libraries, frameworks & real apps.
- Extend based on observed user need

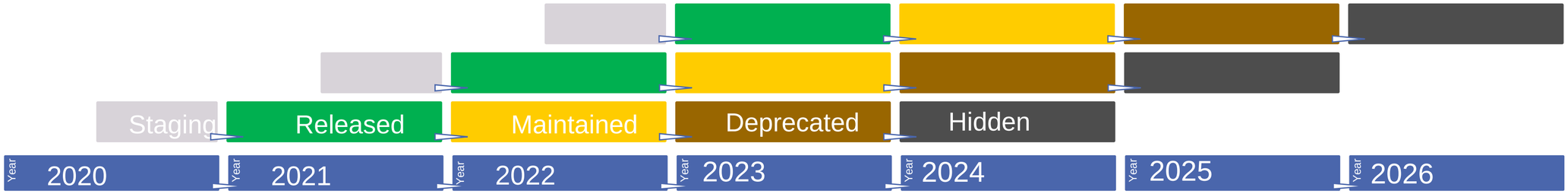
Meluxina User Software Env. - cont'd

Optimizations for MeluXina

- AOCC 2.3 & GCC 10.2: support for *znver2*
- CUDA toolkit 11+: support for *Ampere*
- Intel FPGA OpenCL SDK & OneAPI: support for the *Stratix* FPGAs

Software Environment planning

- Provide stable *production* releases
  - maintained after a while – no updates but serviceable
- Provide *staging* stack for advanced (adventurous) users
  - used also for internal testing



Meluxina User Software Env. - cont'd

Compilers

- AOCC, GCC, Intel, NVIDIA HPC SDK (incl. PGI)

MPI suites

- OpenMPI, Intel MPI, ParaStationMPI

Numerical & data libraries

- BLIS, Intel MKL, FFTW, OpenBLAS, ScalaPACK
- cuBLAS, cuFFT, cuDNN, TensorRT, HDF5, netCDF

Frameworks

- PyTorch, TensorFlow, Horovod, Keras, Apache Spark

Others

- Visualisation
- Build tools
- Performance and debugging tools
- Last but not least: applications!

Testing Meluxina User Software Env. - cont'd

HPC software stack challenges:

- Very complex, particular application/library version combinations work well, others don’t
- Sensitive to changes
- Influenced by changes in the system configuration, OS-level drivers, libraries

Testing is essential & needs to be:

- Consistent
- Maintainable
- Automated

***Periodic* and *on-demand* testing:**

- Report on performance ‘jitter’
- Check new software/version behavior, regressions
- Also: verify stack reproducibility

Use ReFrame for:

- Functionality tests
- Performance tests

Software usage tracking, to understand:

- System usage in view of optimization
- (Rapidly evolving) user needs

MeluXina & containerized workloads

Bring-your-software-stack-in-a-container

- One alternative to global software environment
- Users control complete stack, good for reproducibility

Challenges:

- Integration with MPI suites (where applicable)
- Integration with ‘outer’ software stack
- Integration with host drivers for optimal performance
- Security

Container system on MeluXina

- Singularity

Thank you for your kind attention

Q&A

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.

LUX PROVIDE

Luxembourg’s one-stop shop for

high performance

computing and data analytics