



EXPLORING THE FRONTIERS OF DIGITAL INTELLIGENCE



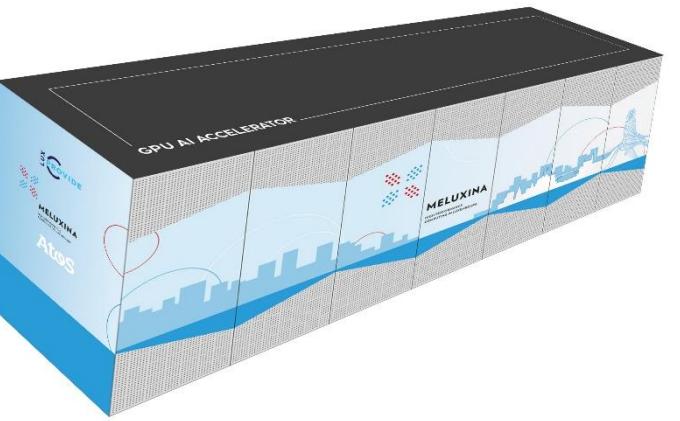


AI on MeluXina

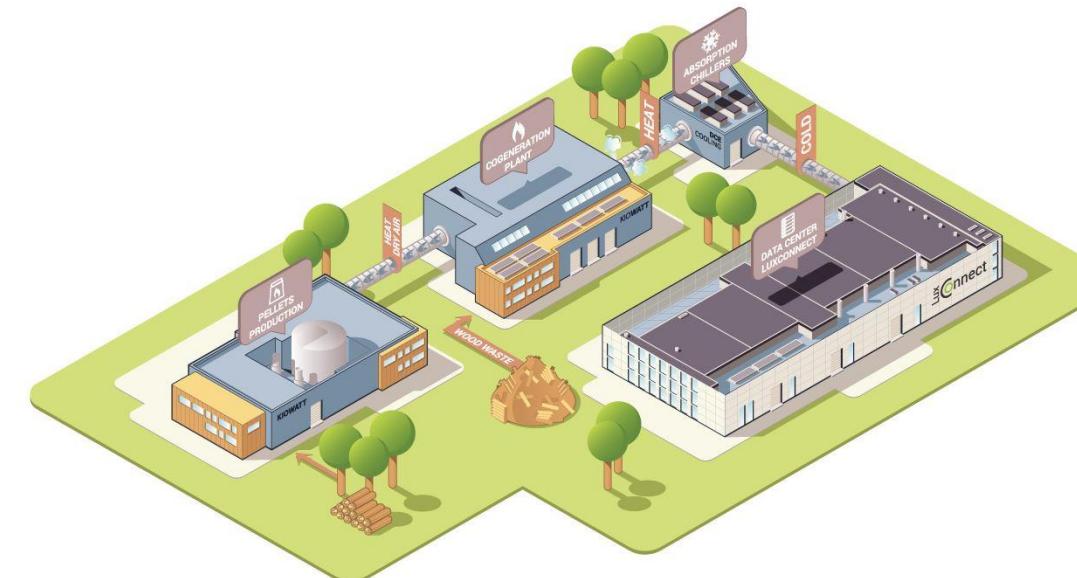
Dr. Francesco Bongiovanni



ARCHITECTURE



FACILITY



HARDWARE SOFTWARE

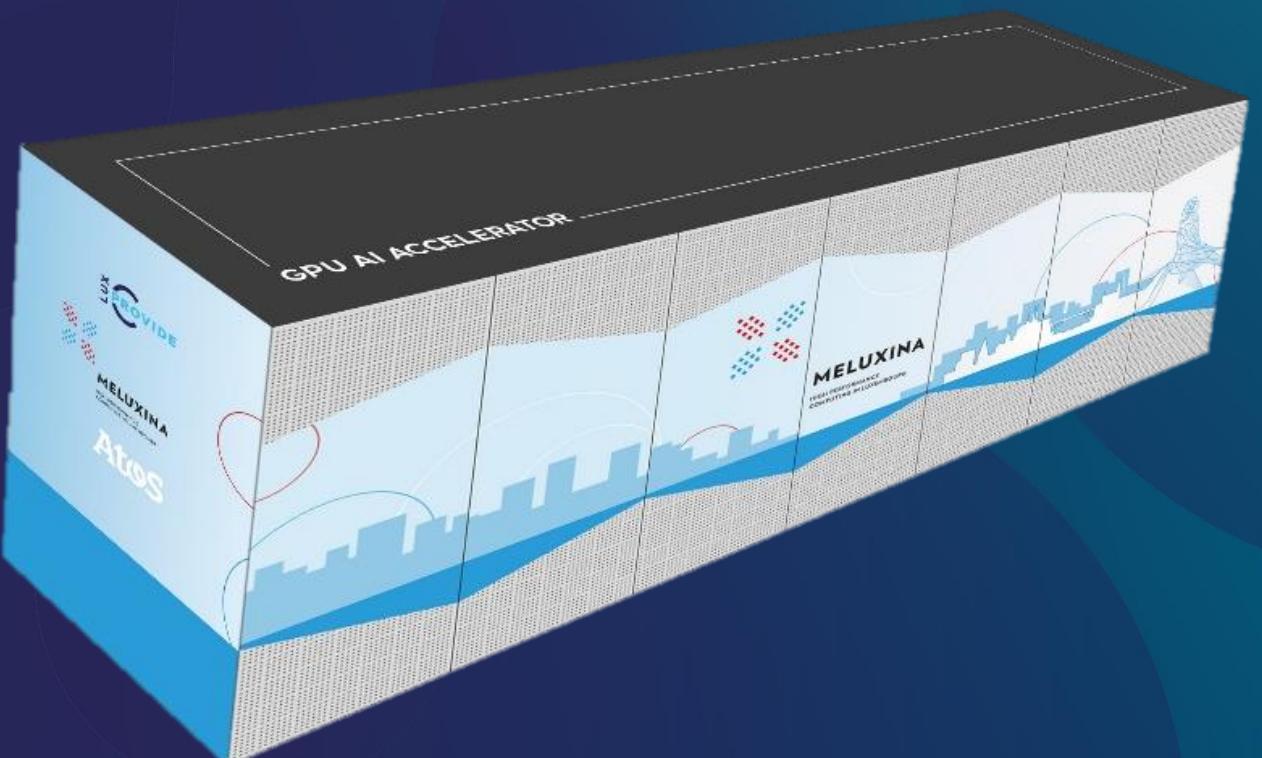


AI USE CASES





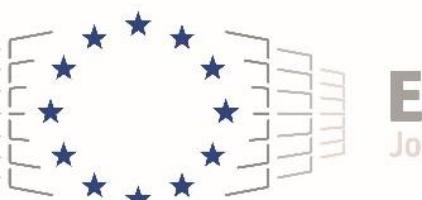
ARCHITECTURE



LuxProvide offers a unique platform that combines data science and supercomputing resources to help organizations increase the ROI of their most challenging innovation projects



ICT Business Partner
of the Year



EuroHPC
Joint Undertaking



European High Performance Computing Joint Undertaking

Over 8 billion Euro for cutting-edge supercomputing and quantum computing ecosystem in Europe



European Network of Competence Centers in HPC

United network of HPC actors in 33 European countries to foster adoption of HPC



European Master for HPC

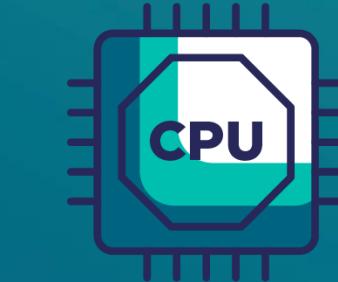
Pan-European Master's program by leading educational institutions focused on HPC





FLOPS/SECOND PEAK COMPUTE PERFORMANCE

If all the humans on the planet were to make a calculation in one second, it would be necessary to multiply the population of the Earth by more than two million to reach MeluXina's capacity



90.000

HPC CPU cores

800

GPU-AI accelerators

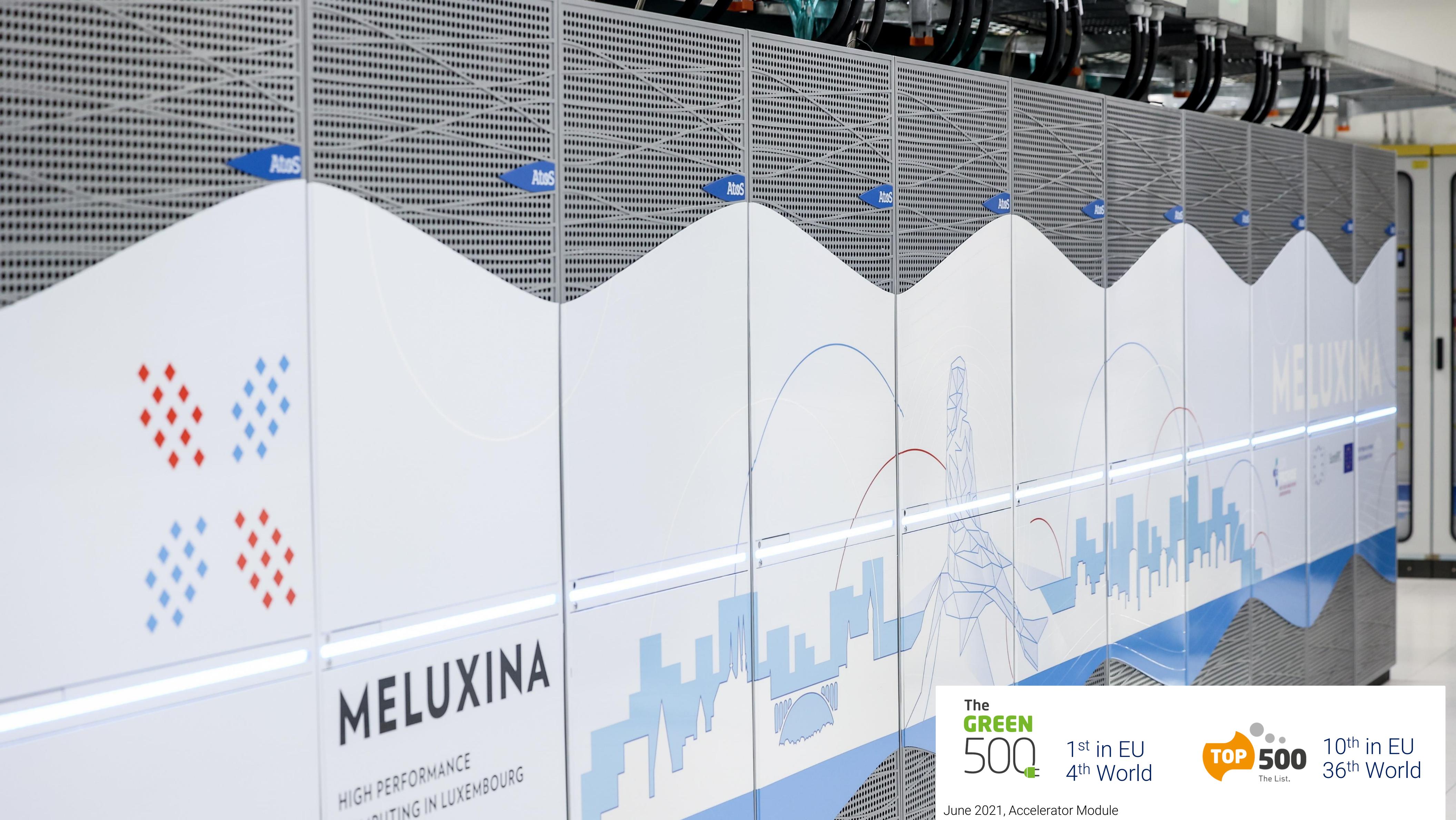


20

Petabytes high-speed storage

300+

Software packages



MELUXINA
HIGH PERFORMANCE
COMPUTING IN LUXEMBOURG

The
GREEN
500

1st in EU
4th World

TOP 500
The List.

10th in EU
36th World



Infrastructure essentials

573x CPU NODES

- 2x AMD Rome (64-core, 2.6 GHz)
- 512 GB RAM
- HDR 200Gb/s Infiniband

200x GPU NODES

- 2x AMD Rome (32-core, 2.35 GHz)
- 512 GB RAM
- 4x NVIDIA A100 (40GB HBM, Nvlink)
- 1.92 TB SSD local storage
- 2x HDR 200Gb/s Infiniband

20x LARGE MEMORY NODES

- 2x AMD Rome (64-core, 2.6 GHz)
- 4 TB RAM
- 1.92 TB NVMe local storage
- HDR 200Gb/s Infiniband

20x FPGA NODES

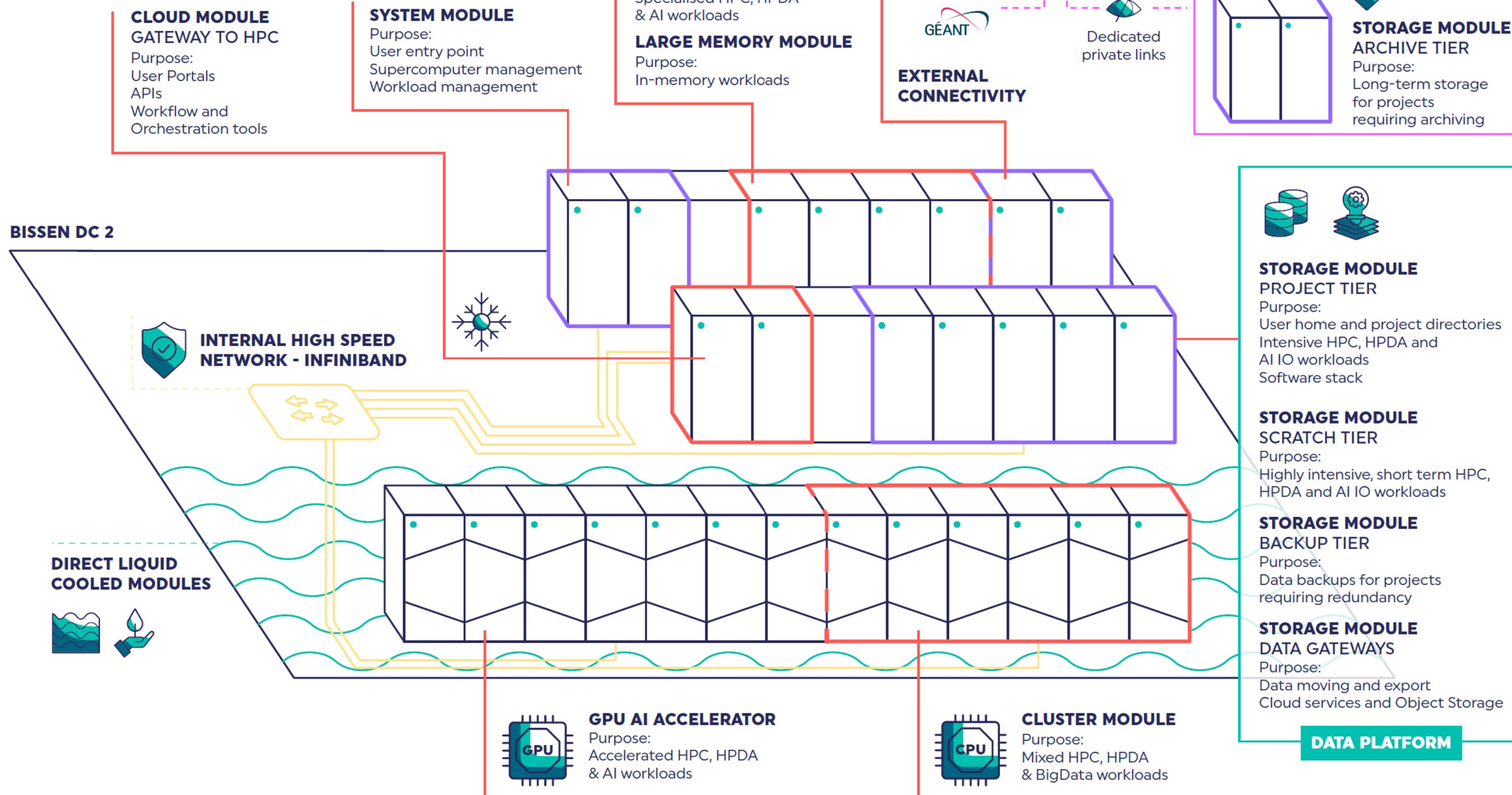
- 2x AMD Rome (32-core, 2.35 GHz)
- 512 GB RAM
- 2x BittWare520N-MX (Intel Stratix 10MX 16GB chip)
- 1.92 TB SSD local storage
- 2x HDR 200Gb/s Infiniband

DATA STORAGE

- Tier1 –Scratch
 - 0.6 PB, 400 GB/s, Lustre PFS
 - Self-Encrypting NVMe drives
- Tier2 -Home/Project
 - 13 PB, 190 GB/s, Lustre PFS
 - Self-Encrypting NVMe & HDD drives
- Tier3 –Backup
 - 7 PB, 30 GB/s, Lustre PFS
 - Self-Encrypting NVMe & HDD drives
- Tier4 –Archive
 - 5 PB
 - Tape-based Archival system

MELUXINA ARCHITECTURE

- + Modular design -> versatility
- + Liquid cooled -> efficiency
- + External connectivity -> high speed accessibility
- + Optimized software stack & containers
- + Tier IV Data Center -> highest availability



MeluXina User Software Environment

PGI AMD



intel



[Op]
[en]
[BL]
[AS]



ParaStation
MPI

TensorFlow

FFTW

APACHE
Spark

PyTorch

K Keras



Compilers, Languages & Performance Eng.

Numerical & data libraries

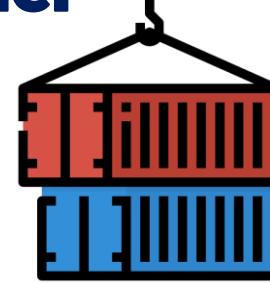
Parallelization tools,
MPI suites & acceleration libraries

Frameworks, runtime & platform tools

End user applications



Singularity-CE / Apptainer



Container system on MeluXina

- Docker & OCI compatible
- Non-privileged mode for improved security
- Support for GPU accelerated applications
- Support for creating and running encrypted containers
- Support for trusted containers: PGP signed & verified

Bring-your-own software stack



- Enable users to easily control complete stack --- REPRODUCIBILITY
- Users create tooling and pipeline on their infrastructure --- RUN PIPELINE ON MELUXINA

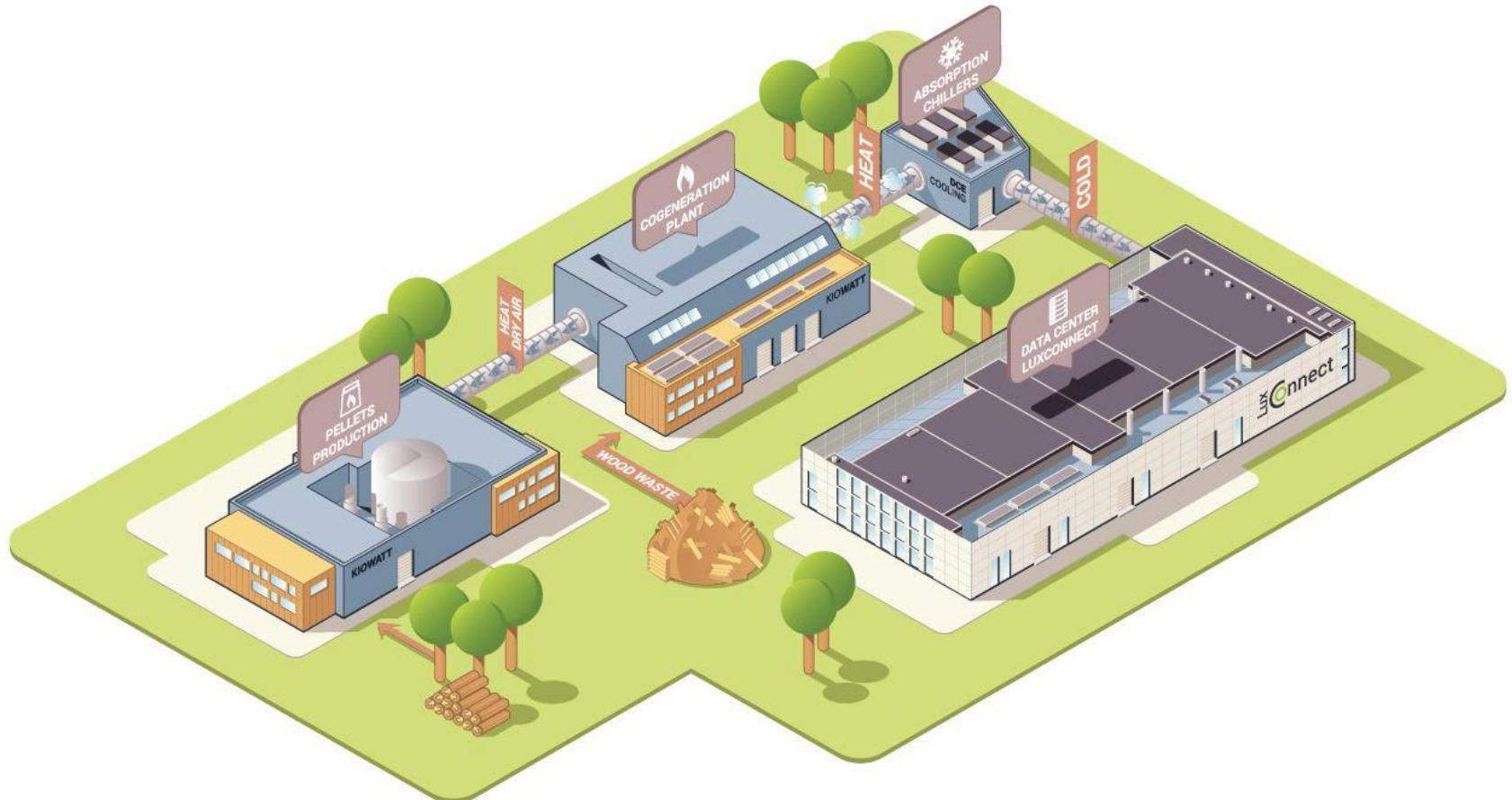


FACILITY

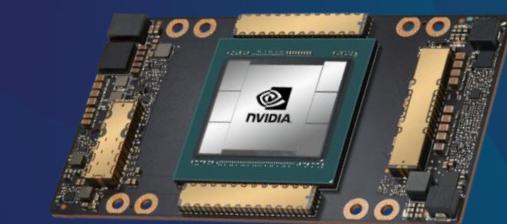


Green TierIV Data Center

- PUE of the Data Centers constantly measured and monitored
- Continuous improvement plan for Energy Efficiency supervised by the government agency Klima-Agence
- Waste heat from servers used to heat office space and preheat diesel generators
- Certified 100% Green Electricity supplied from hydroelectric power sources
- Optimised use of Free Cooling
- Biomass recycling representing a yearly reduction in CO₂ in excess of 27,000 metric tons



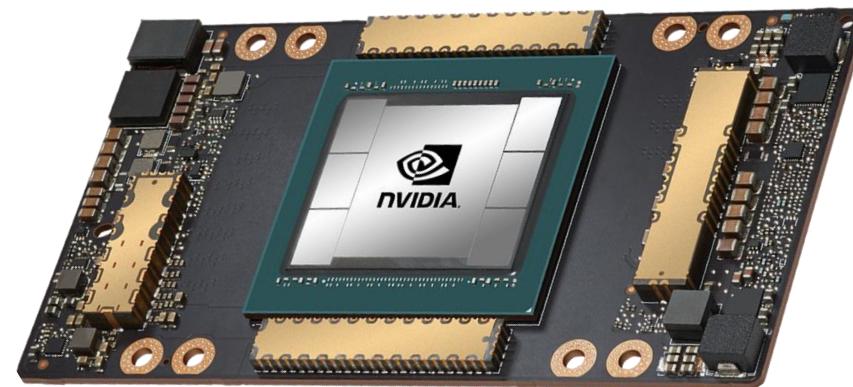
HARDWARE & SOFTWARE



Know your hardware



7H12	
Cores	64
Base clock	2.6 Ghz
Max. Boost clock	3.3 Ghz
L3 cache	256 MB
Memory	DDR4
Memory channel	8
Memory Bandwidth	204.8 GB/s per socket
TDP	280W

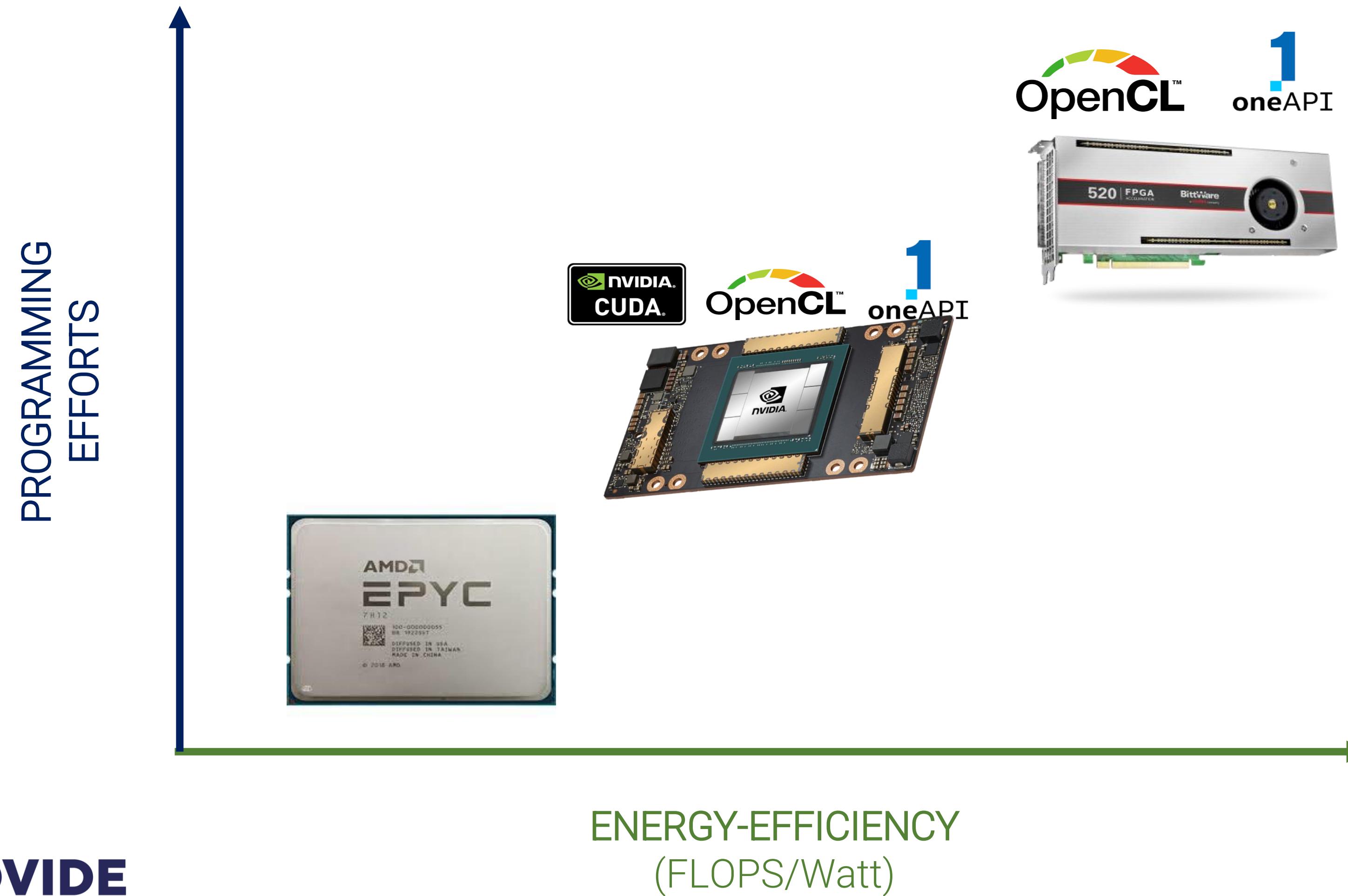


A100 SXM	
FP64	9.7 TFLOPS
FP64 Tensor Core	19.5 TFLOPS
FP32	19.5 TFLOPS
TF32	156 TFLOPS
BFLOAT16	312 TFLOPS
FP16	312 TFLOPS
INT8	624 TOPS
Memory	40GB HBM2
Memory Bandwidth	1,555 GB/s
Interconnect	600GB/s
TDP	250W



Stratix 10 MX 2100	
FP32	6.3 TFLOPS
Memory	16GB HBM2
Memory Bandwidth	512 GB/s
TDP	225W

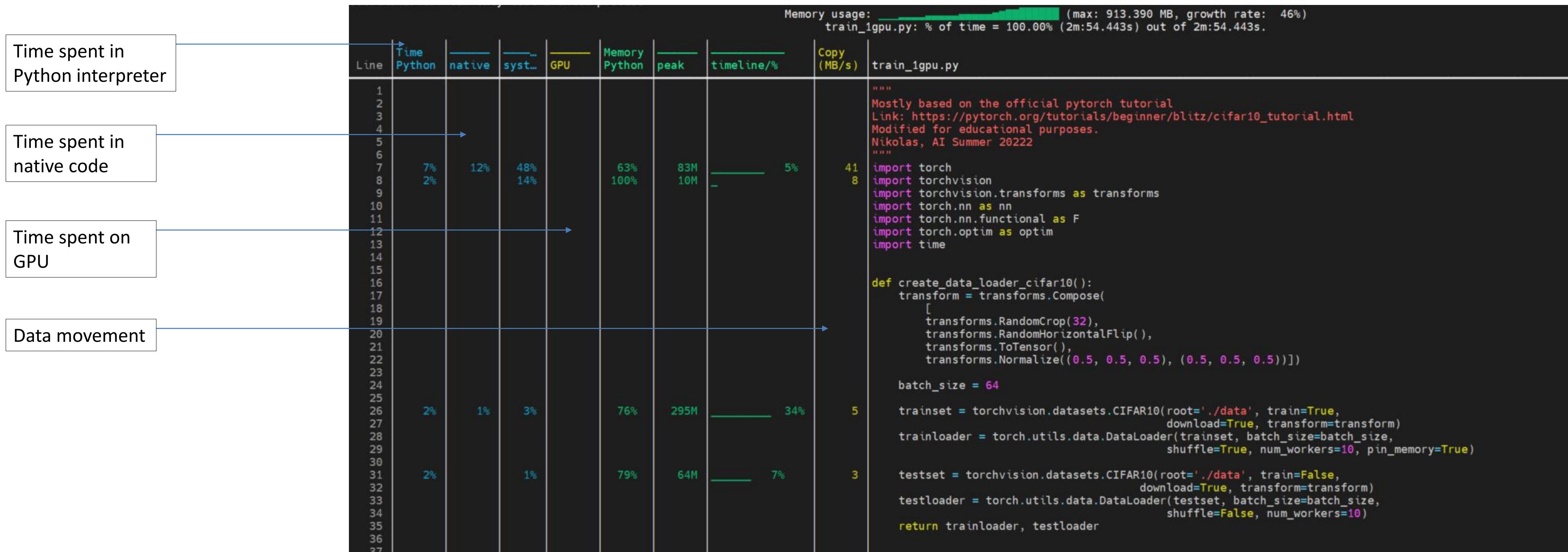
Know your hardware



Know your software

AI-oriented profilers: Scalene

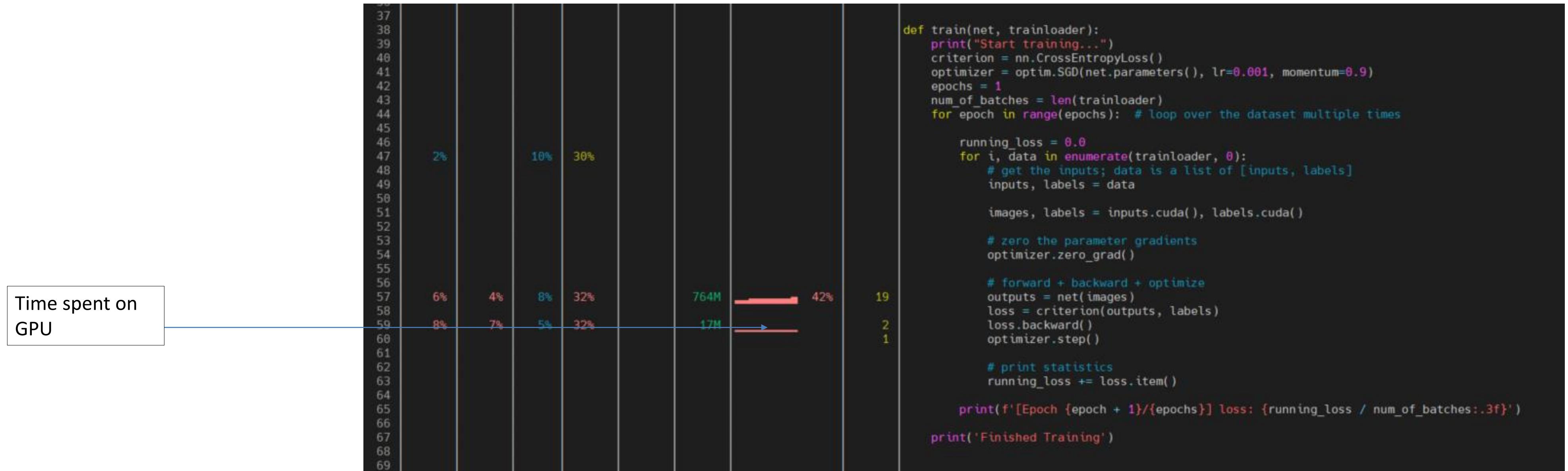
scalene python3 train_1gpu.py



Know your software

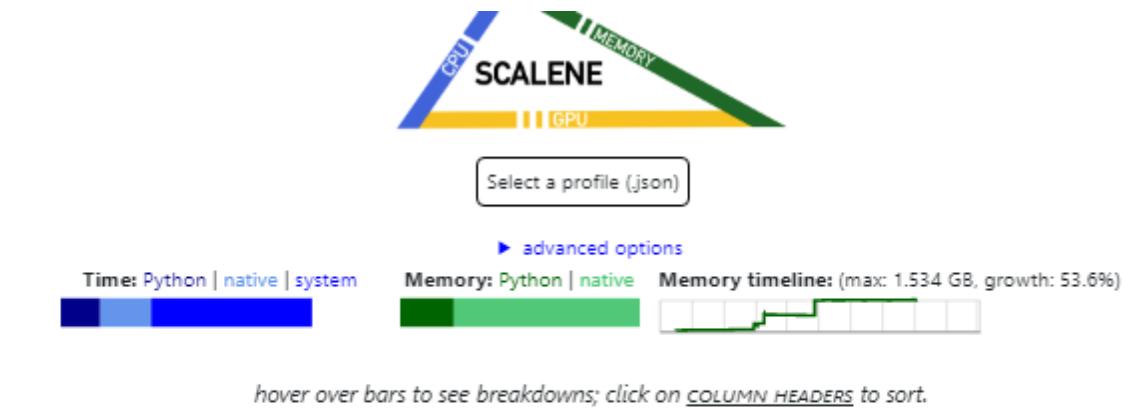
AI-oriented profilers: Scalene

Zooming on GPU usage



Know your software

AI-oriented profilers: Scalene



scalene --json --outfile profile_scalene.json python3 train_1gpu.py

Using the scalene GUI (<http://plasma-mass.org/scalene-gui/>), You can view the .json profile and ask ChatGPT for some optimizations 😊



Know your software

AI-oriented profilers: *Nvidia DLProf*

Profiling with Nvidia DLProf requires some light modifications to the source code

1. Import the library

```
import nvidia_dlprof_pytorch_nvtx as nvtx
```

2. Modify the main function

add the following initialization code :

```
nvtx.init(enable_function_stack=True)
```

Wrap the train function with this code:

```
with torch.autograd.profiler.emit_nvtx():
```

3. Once you have done the modifications, just run:

```
dlprof python3 train_1gpu.py
```

This will create a couple of files, two sqlite files and one nsys-rep.

```
92 if __name__ == '__main__':
93     start = time.time()
94
95     import torchvision
96
97     PATH = './cifar_net.pth'
98     trainloader, testloader = create_data_loader_cifar10()
99     net = torchvision.models.resnet50(False).cuda()
100    nvtx.init(enable_function_stack=True)
101    start_train = time.time()
102    with torch.autograd.profiler.emit_nvtx():
103        train(net, trainloader)
104    end_train = time.time()
105    # save
106    torch.save(net.state_dict(), PATH)
107    # test
108    test(net, PATH, testloader)
109
110    end = time.time()
111    seconds = (end - start)
112    seconds_train = (end_train - start_train)
113    print(f"Total elapsed time: {seconds:.2f} seconds, \\"
```

```
[fbongiovanni@login02 pytorch-ddp]$ ls -lah *.sqlite
-rw-r-----. 1 fbongiovanni lxp 101M Mar 22 15:15 dlprof_dlldb.sqlite
-rw-r--r--. 1 fbongiovanni lxp 650M Mar 22 15:12 nsys_profile.sqlite
[fbongiovanni@login02 pytorch-ddp]$ ls -lah *.nsys-rep
-rw-rw-r--. 1 fbongiovanni lxp 54M Mar 22 15:11 nsys_profile.nsys-rep
[fbongiovanni@login02 pytorch-ddp]$ █
```

Know your software

- User reporting that his workload (face recognition software) is taking 2x longer than expected
- First investigations point to an issue with how user is launching the software with SLURM.
- Constant communication with end user
- Further investigations show that with minimal changes to code (1 line), and optimized pinning/binding, the user could get 8.8x performance improvement.
- Further details regarding pinning/binding can be found [here](#)

In the face_detector_deepsparse.py file, I modified again on line 34 the following code

with self._compiled_model = Engine(model=onnx_filepath, batch_size=batch_size)

self._compiled_model = Engine(model=onnx_filepath, batch_size=batch_size*4,num_cores=256,scheduler="multi_stream", num_streams=8)

I basically increase the batch size by 4, increase the num_cores to match the hyperthreaded cpu cores and fix the number of concurrent streams to match the numa domains, 8

and then with the following command to run it

NM_BIND_THREADS_TO_CORES=1 numactl --physcpubind=0-255 python3 pipeline_manager.py --mp4_dir ..//mp4_25fps/ --saving_dir ..//results/

I manage to get it from **8min24 sec to 57 secs** on my 4 videos dataset.



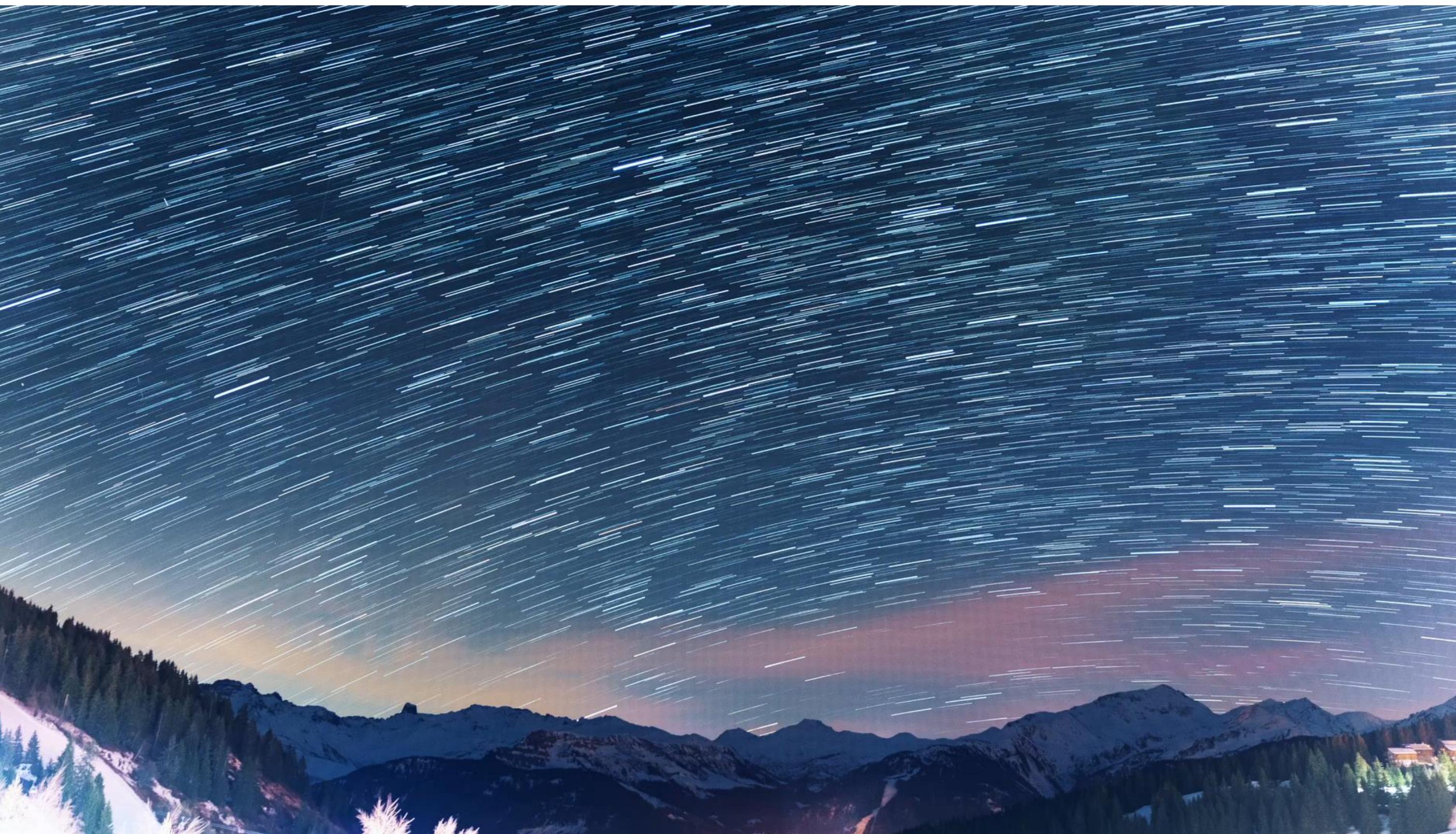
AI USE CASES



SES[▲]



RSS-Hydro



**DESTINATION
EARTH
(DestinE)**



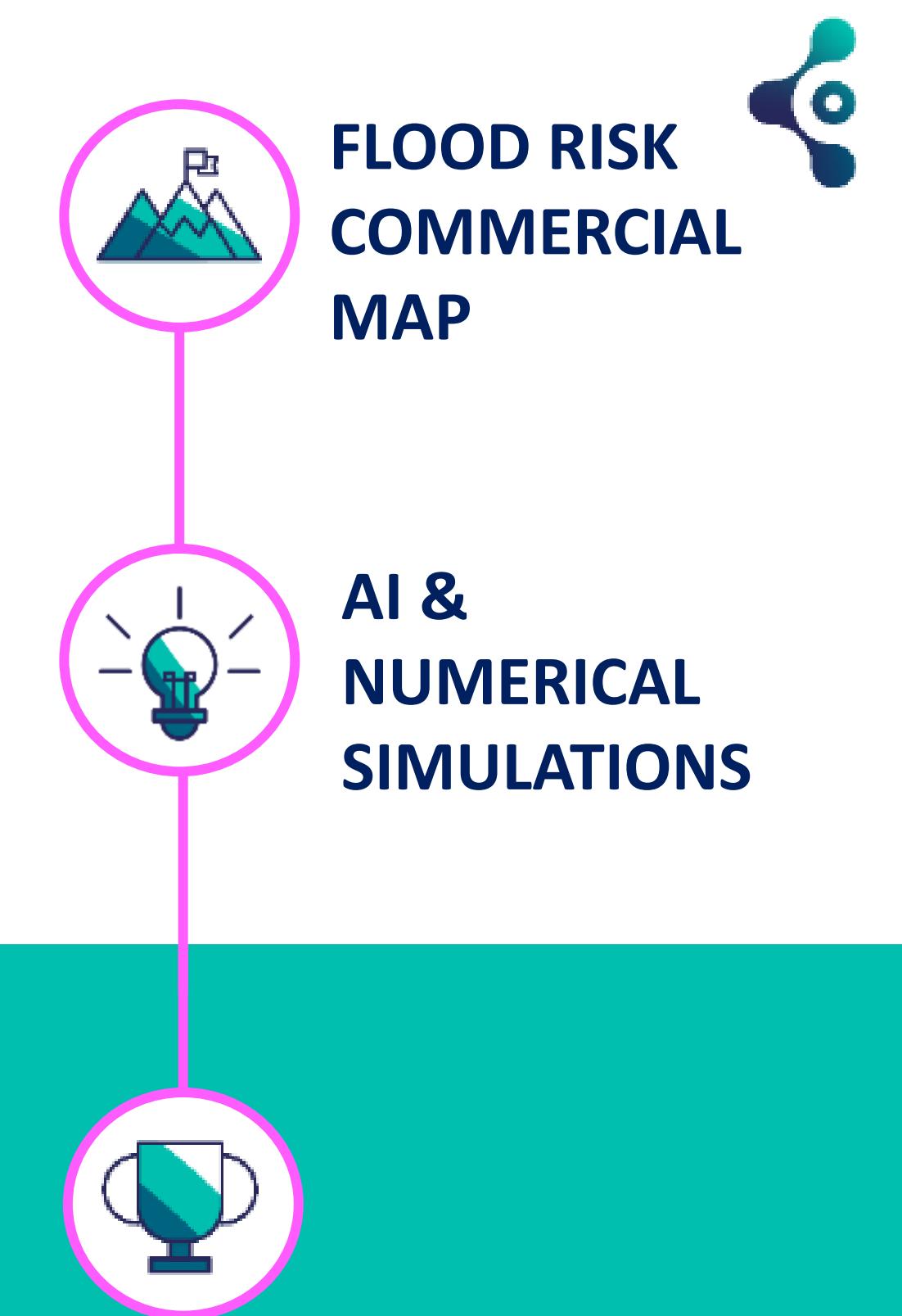
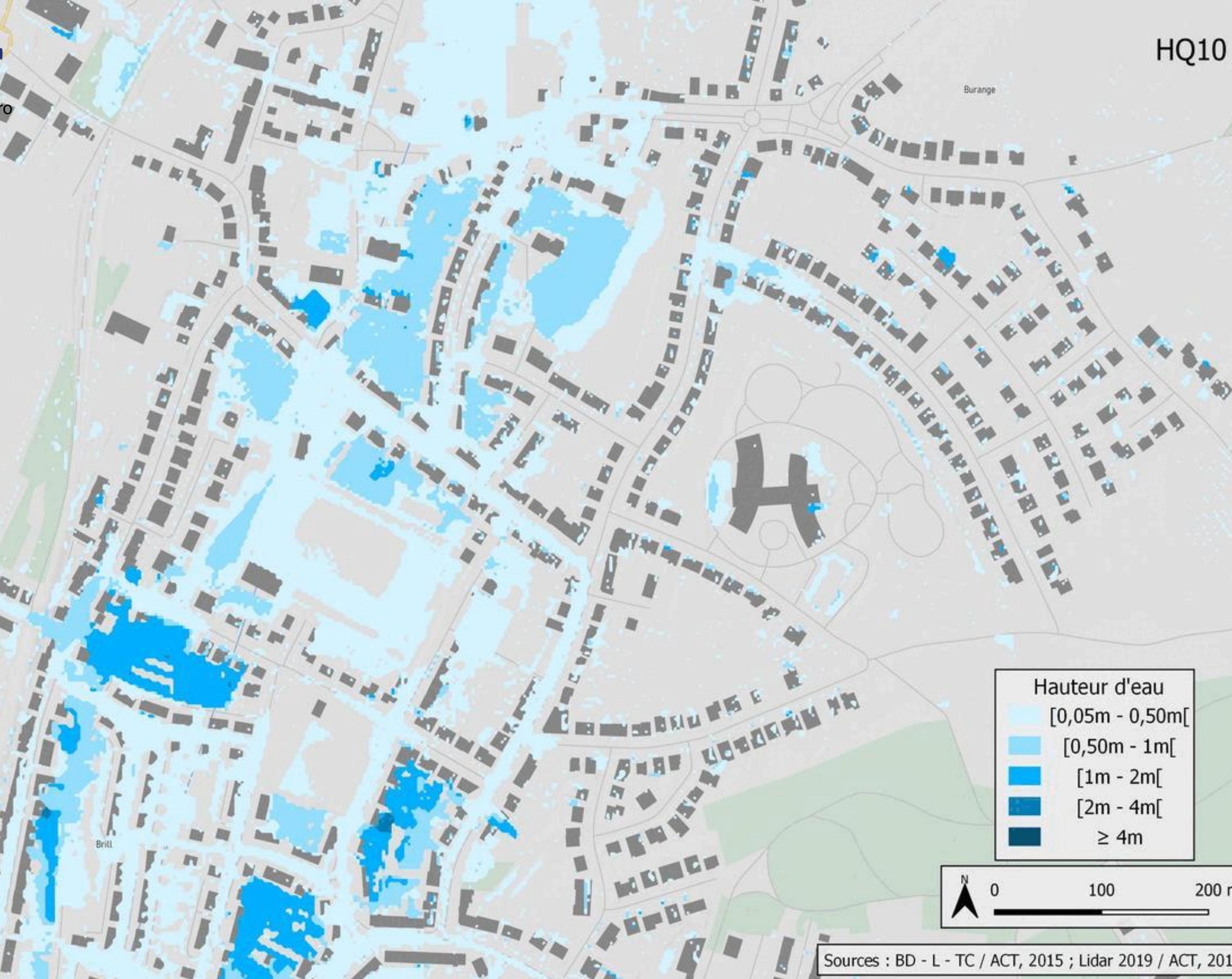
**AI
& NUMERICAL
SIMULATIONS**



**DIGITAL TWIN OF THE
EARTH**

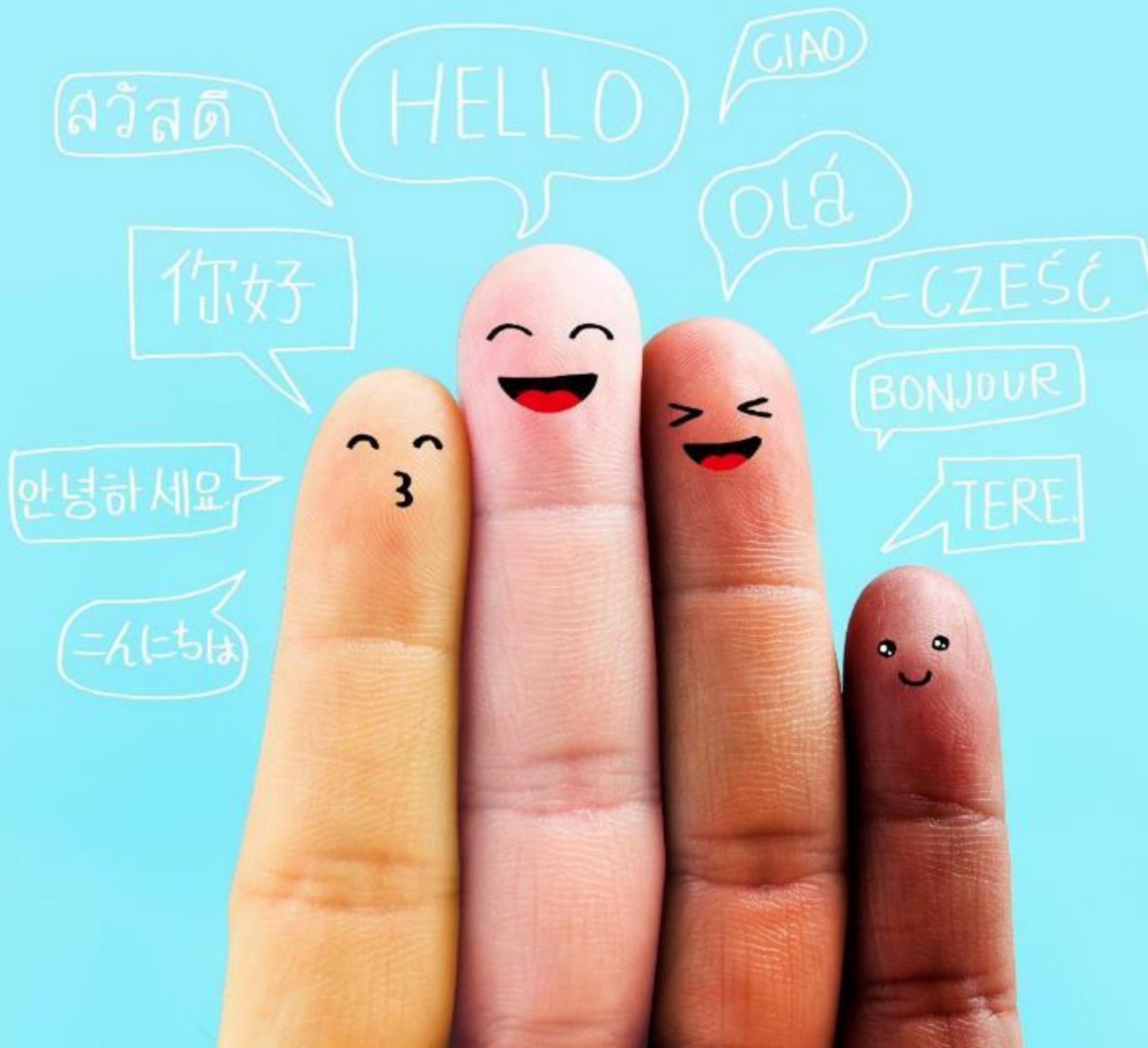
**UNDERSTAND THE PAST,
PREDICT THE FUTURE**





10X Speed-up on
MeluXina GPU nodes





eTRANSLATION

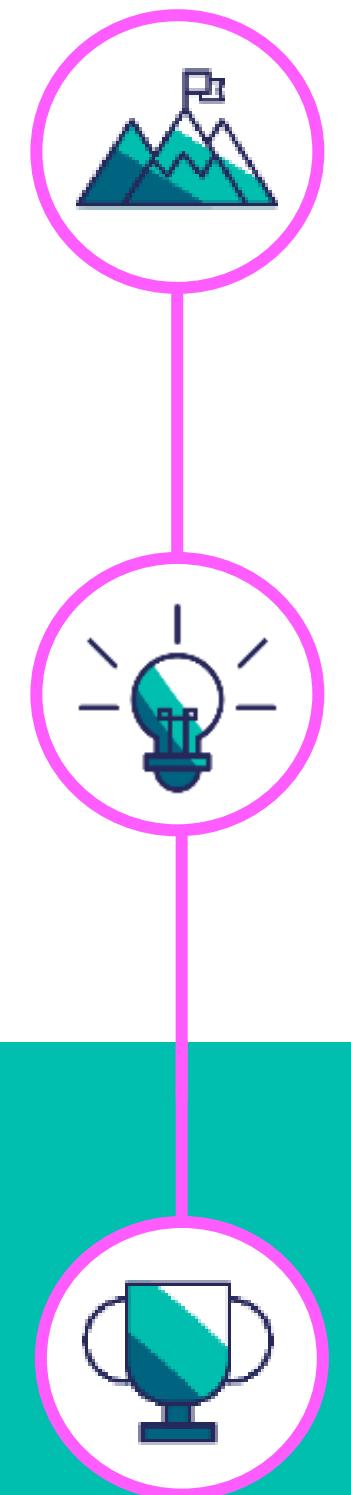
**AI | NATURAL
LANGUAGE
PROCESSING**



**TRANSLATE FROM AND
INTO ANY EU LANGUAGE**

(+)

**OFFERS SEVERAL STYLES
(DOMAINS) OF
TRANSLATION**





CoeFont



SPEECH
GENERATION



TTS / STT



GENERATION OF VOICES
FOR MANGAS/ANIMES

+

AUTO-TRANSCRIPTION
INTO TEXT FROM AUDIO



EXPLORING THE FRONTIERS OF DIGITAL INTELLIGENCE

Our mission is to explore the frontiers of digital intelligence fusing data science and supercomputing capabilities.

