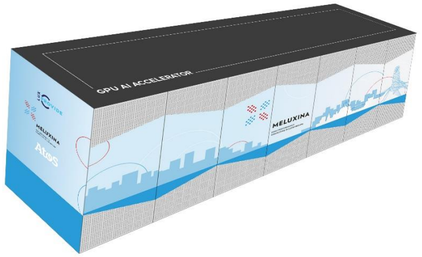
LUX PROVIDE

# Making environmental science environmentally friendly

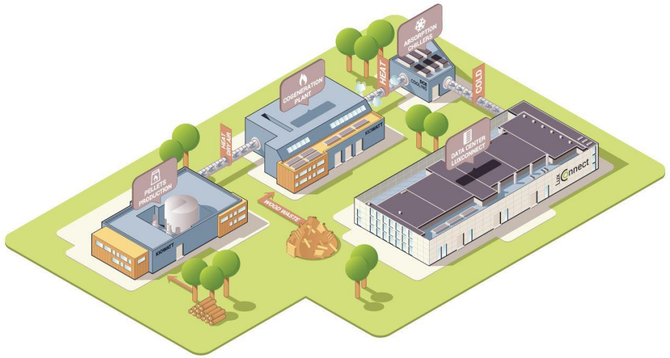
## Follow the merm(AI)d

**Dr. Francesco Bongiovanni**

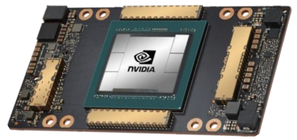
## ARCHITECTURE




## FACILITY




## HARDWARE
## SOFTWARE



OpenCL™

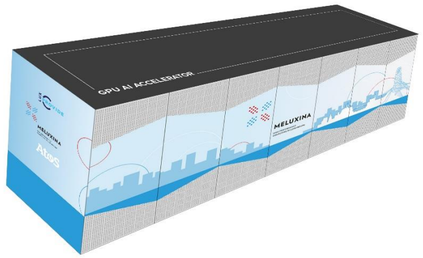
oneAPI

NVIDIA CUDA

## ESG IMPACT

LUX PROVIDE

# 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

LUX PROVIDE high-performance computing center

INFORMATION & COMMUNICATION TECHNOLOGIES LUXEMBOURG AWARDS 2022

ICT Business Partner of the Year

ISO 27001 CERTI TRUST

EuroHPC Joint Undertaking

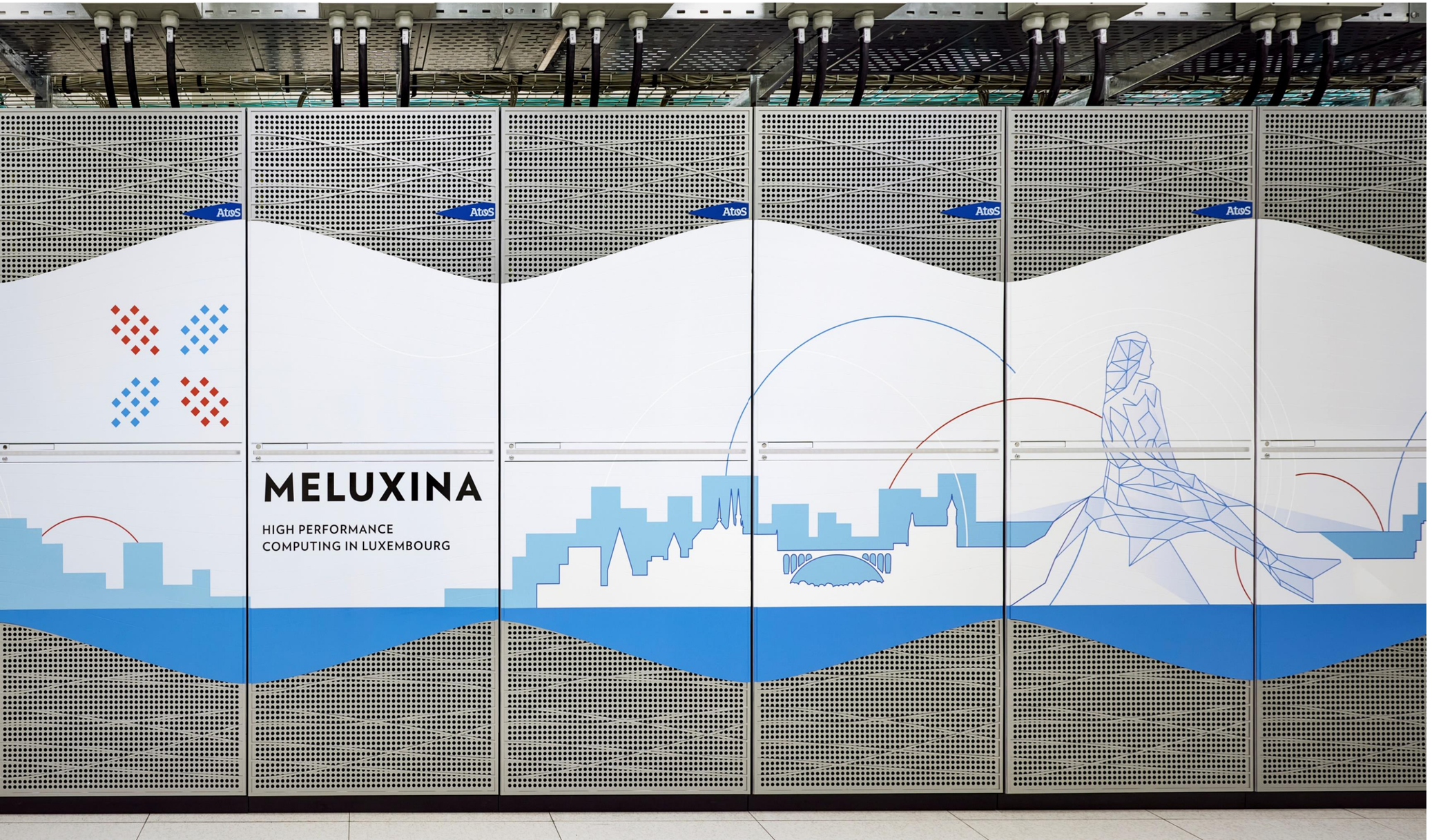
# The EUROHPC Supercomputers

- **5 operational systems** in Luxembourg, Slovenia, Czechia, Bulgaria & Finland;
- **3 systems underway** in Italy, Portugal & Spain;
- New hosting entities selected in June 2022, to host new mid-range systems and the **first European exascale supercomputer.**

EuroHPC
Joint Undertaking

# Luxembourg's national supercomputer Meluxina

LUX PROVIDE high-performance computing center




**90.000**
HPC CPU Cores

**5.120**
Cloud vCPUs

**800**
GPU-AI accelerators

Infiniband HDR
**200**
Gb/s

**20**
Petabytes high performance storage

**300+**
Tailored software packages

# World-class supercomputing

## Performance

Compute

**18 FP64** PetaFlops

**500 AI** PetaFlops

Data

**500+** GB/s Scratch storage

**190+** GB/s Project storage

Interconnect

**200** Gb/s on Cluster nodes

**400** Gb/s on Accelerator and large memory nodes

MELUXINA
HIGH PERFORMANCE COMPUTING IN LUXEMBOURG

The GREEN 500 — 1st in EU, 4th World

TOP 500 The List. — 10th in EU, 36th World

June 2021, Accelerator Module

# MeluXina's Architecture

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

**CLOUD MODULE**
GATEWAY TO HPC

Purpose:
User Portals
APIs
Workflow and
Orchestration tools

**SYSTEM MODULE**
Purpose:
User entry point
Supercomputer management
Workload management

**ACCELERATOR MODULE**
FPGA PARTITION
Purpose:
Specialised HPC, HPDA
& AI workloads

**LARGE MEMORY MODULE**
Purpose:
In-memory workloads

**EXTERNAL CONNECTIVITY**

Public Internet

restena

GÉANT

LU-CIX

Dedicated private links

EXAMPLE

**BETTEMBOURG DC 1.2**

**STORAGE MODULE**
ARCHIVE TIER
Purpose:
Long-term storage for projects
requiring archiving

**BISSEN DC 2**

**INTERNAL HIGH SPEED NETWORK - INFINIBAND**

**DIRECT LIQUID COOLED MODULES**

**STORAGE MODULE**
PROJECT TIER
Purpose:
User home and project directories
Intensive HPC, HPDA and
AI IO workloads
Software stack

**STORAGE MODULE**
SCRATCH TIER
Purpose:
Highly intensive, short term HPC,
HPDA and AI IO workloads

**STORAGE MODULE**
BACKUP TIER
Purpose:
Data backups for projects
requiring redundancy

**STORAGE MODULE**
DATA GATEWAYS
Purpose:
Data moving and export
Cloud services and Object Storage

**DATA PLATFORM**

**GPU AI ACCELERATOR**
Purpose:
Accelerated HPC, HPDA
& AI workloads

**CLUSTER MODULE**
Purpose:
Mixed HPC, HPDA
& BigData workloads

# MeluXina User Software Environment

- Compilers, Languages & Performance Eng.
- Numerical & data libraries
- Parallelization tools, MPI suites & acceleration libraries
- Frameworks, runtime & platform tools
- End user applications

## Singularity-CE

*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

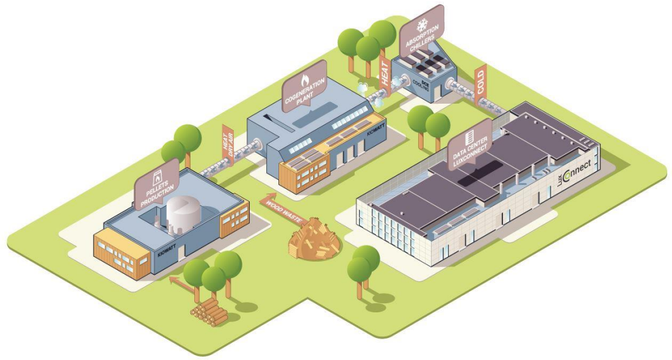
## Bring-your-own software stack

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

PGI AMD NVIDIA CUDA intel GCC OpenBLAS ParaStation MPI TensorFlow FFTW Apache Spark PyTorch Keras HOROVOD

LUX PROVIDE high-performance computing center

# FACILITY

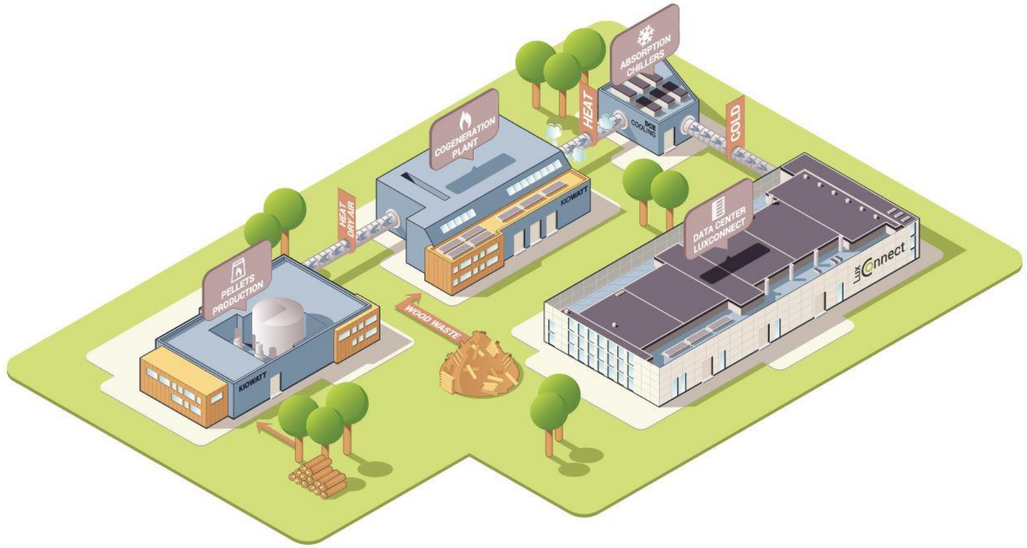



GREEN Power

LUX PROVIDE

# Green TierIV Data Center

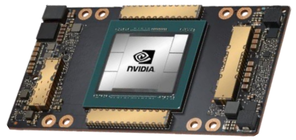
- **PUE** of the Data Centers constantly measured and monitored
- Continuous improvement plan for **Energy Efficiency** supervised by the government agency Klima-Agence
- **Waste heat** from servers used to heat office space and preheat diesel generators




- Certified 100% **Green Electricity** supplied from hydroelectric power sources
- Optimised use of **Free Cooling**
- **Biomass** recycling representing a yearly reduction in $CO_2$ in excess of 27,000 metric tons

HARDWARE

SOFTWARE



OpenCL

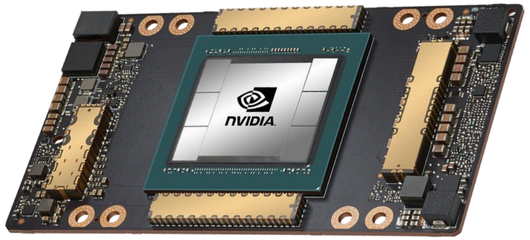
oneAPI

NVIDIA CUDA

LUX PROVIDE

# 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 |

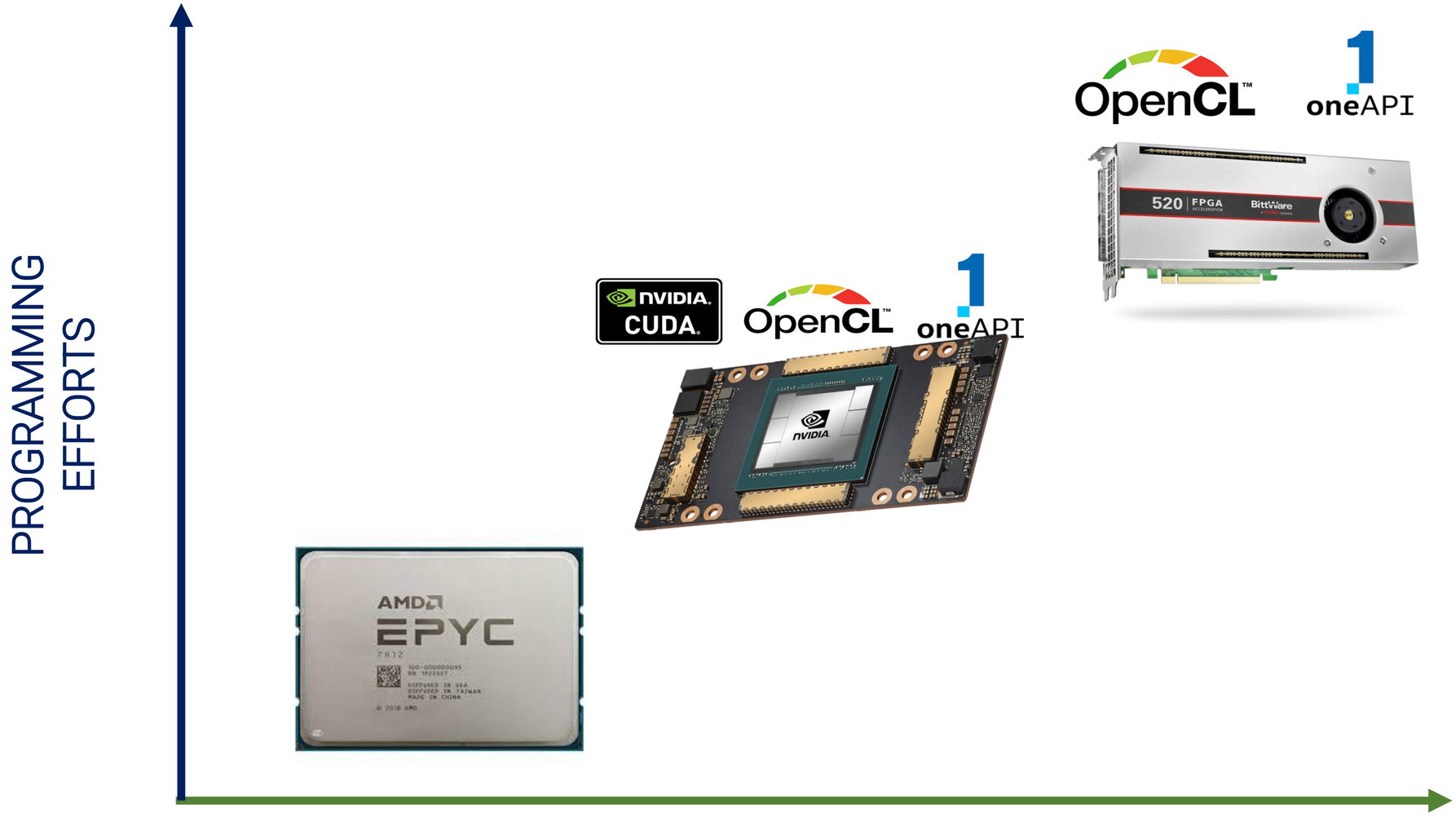



| A100 SXM | |
|---|---|
| FP64 | 9.7 TFLOPS |
| FP64 Tensor Core | 19.5 TFLOPS |
| FP32 | 19.5 TFLOPS |
| TF32 | 156 TFLOPS |
| BFLOAT16 | 312 TFLOPS |
| FP16 | 312 TFLOPS |
| INT8 | 624 TOPS |
| Memory | 40GB HBM2 |
| Memory Bandwidth | 1,555 GB/s |
| Interconnect | 600GB/s |
| TDP | 250W |


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

LUX PROVIDE high-performance computing center

# Know your hardware




LUX PROVIDE high-performance computing center

# Know your software

## AI-oriented profilers: *Scalene*

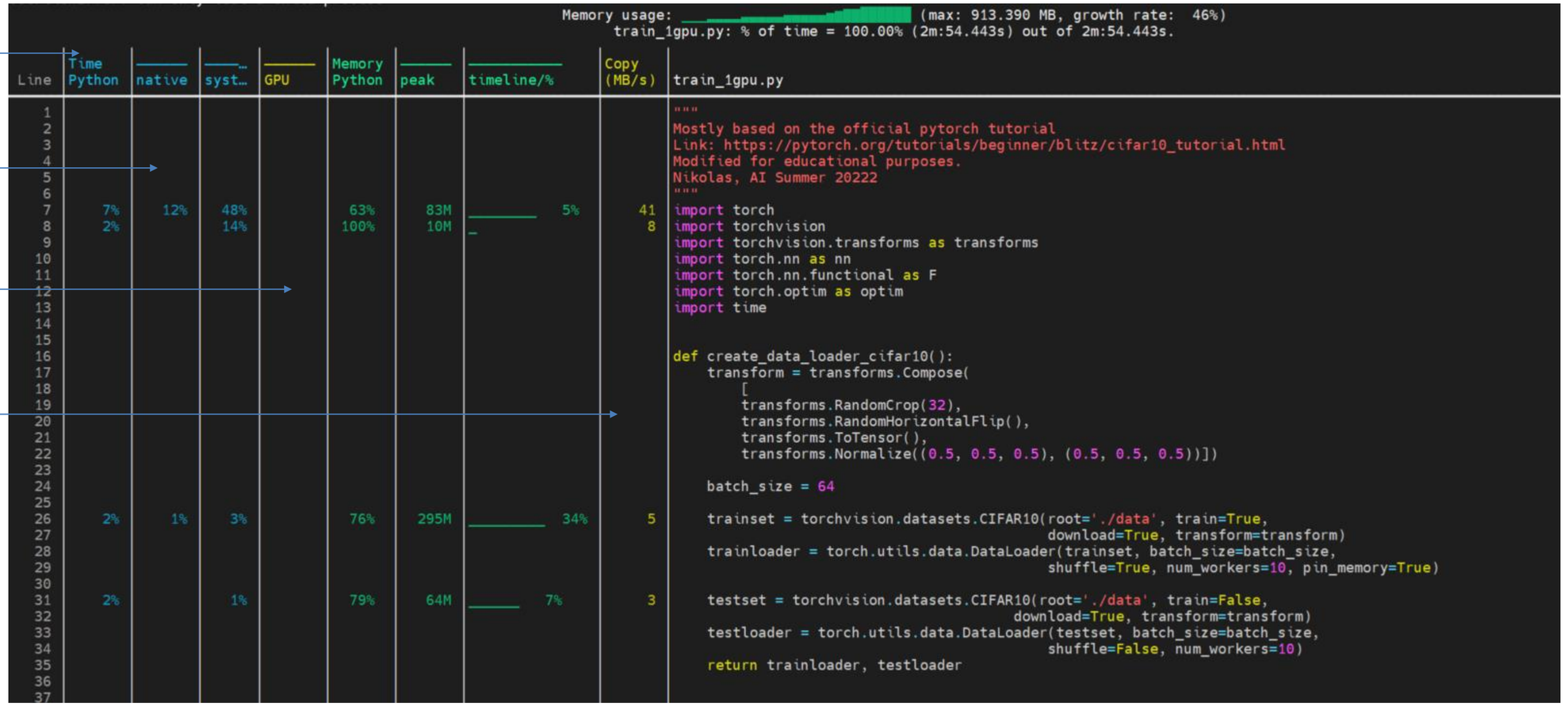
scalene python3 train_1gpu.py

Time spent in Python interpreter

Time spent in native code

Time spent on GPU

Data movement

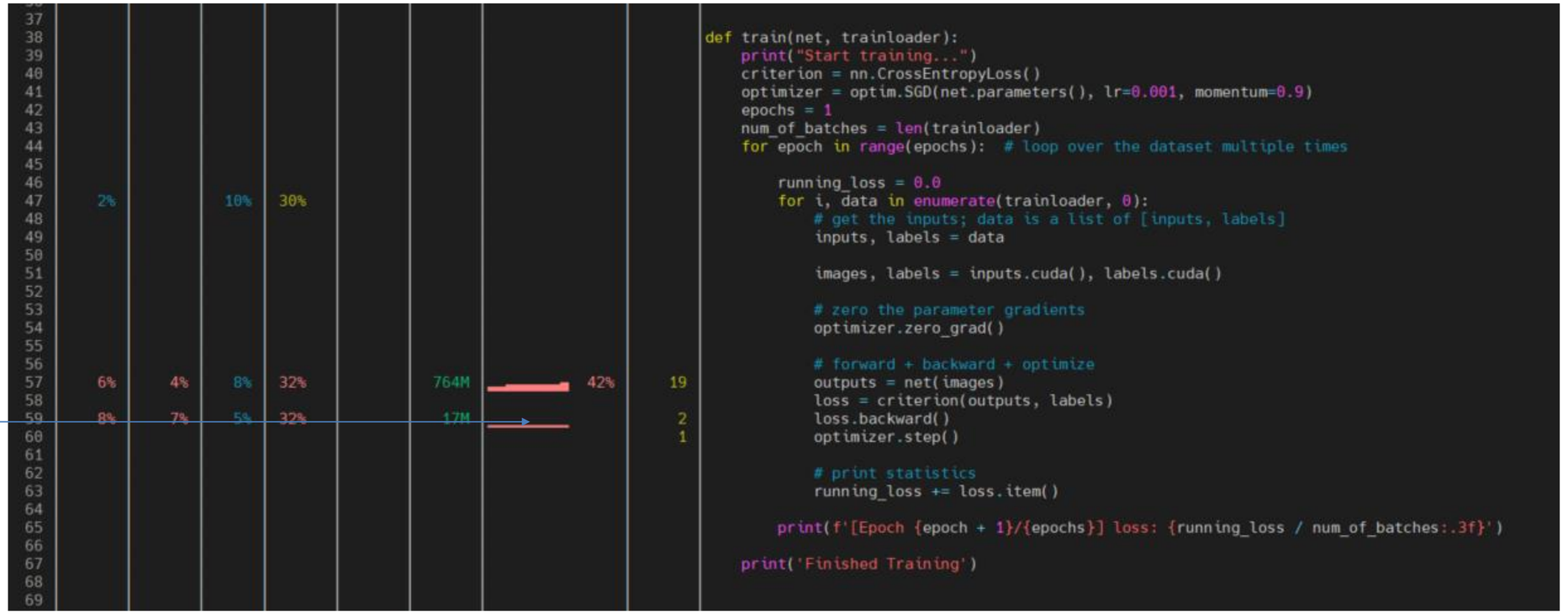


# Know your software

## AI-oriented profilers: *Scalene*

Zooming on GPU usage

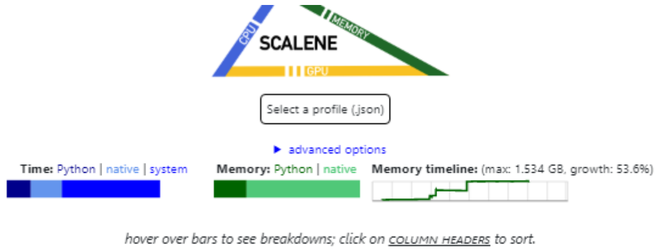
Time spent on GPU



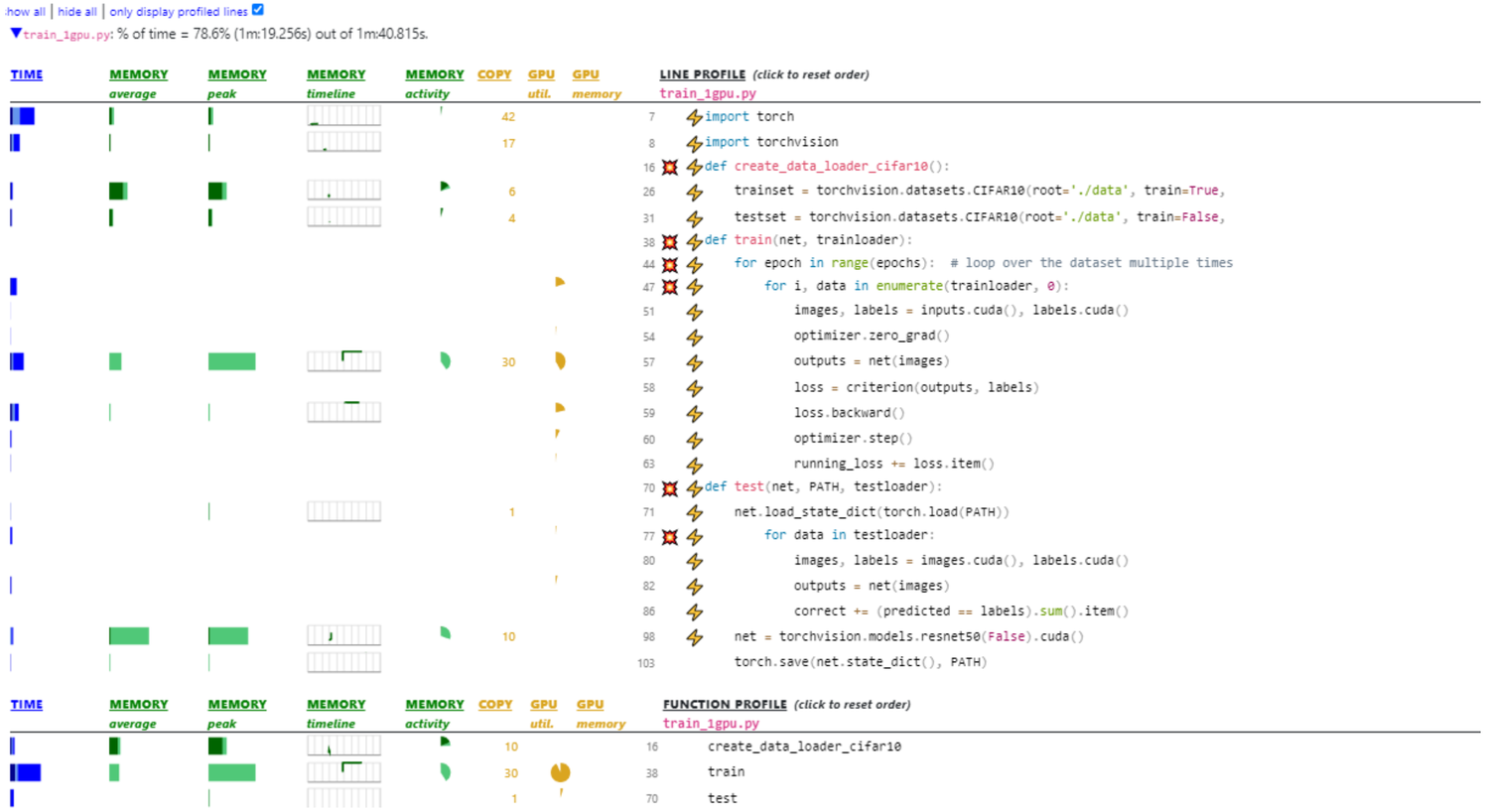
# Know your software

## 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.

```python
if __name__ == '__main__':
    start = time.time()

    import torchvision

    PATH = './cifar_net.pth'
    trainloader, testloader = create_data_loader_cifar10()
    net = torchvision.models.resnet50(False).cuda()
    nvtx.init(enable_function_stack=True)
    start_train = time.time()
    with torch.autograd.profiler.emit_nvtx():
      train(net, trainloader)
    end_train = time.time()
    # save
    torch.save(net.state_dict(), PATH)
    # test
    test(net, PATH, testloader)

    end = time.time()
    seconds = (end - start)
    seconds_train = (end_train - start_train)
    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_dldb.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]$
```

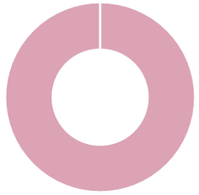
# Know your software

## AI-oriented profilers: *NVidia DLProf*

NVIDIA DLProf Viewer  View  Dashboard  Help

**GPU Utilization (%)**

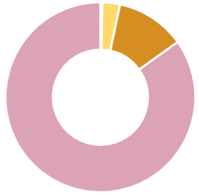
- Using GPU
- Idle GPU



More...

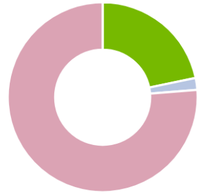
**Resource Usage Breakdown**

- Using Tensor Cores
- Not Using Tensor Cores
- Memory
- Dataloader
- I/O
- CPU
- Other



**Total Kernel GPU Time (ms)**

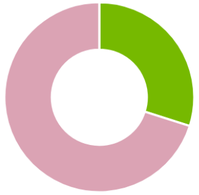
- Using Tensor Cores
- Memory Kernels
- All Other Kernels



More...

**Tensor Core Kernel Efficiency (%)**

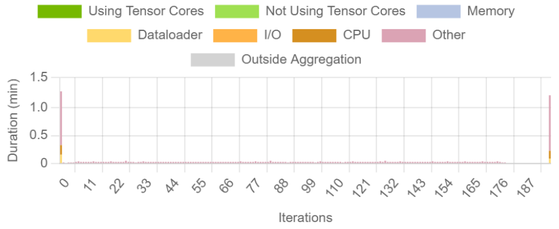
- Using Tensor Cores
- Not Using Tensor Cores



**Performance Summary**

| | |
|---|---|
| Wall Clock Time (min) | 8.48 |
| Tensor Core Kernel Efficiency (%) | 29.9 |
| GPU Utilization (%) | 0.3 |
| Total Iterations | 197 |
| Profiled Iterations | 197 |
| Start Iteration | 0 |
| Stop Iteration | 196 |
| Average Iteration Time (s) | 2.58 |

**Iteration Summary**




**Zoom: enabled (1x), tooltip: enabled**

Reset zoom  Toggle Tooltip

More...

**Top 10 GPU Ops**

| GPU Time (ns) | Op Name | Direction | Op Type | Calls | TC Eligible | Using TC |
|---|---|---|---|---|---|---|
| 324,433,581 | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Sequential::forward/Bottleneck::_call_impl/Bottleneck::forward/Conv2d::_call_impl[7]/Conv2d::forward/Conv2d::_conv_forward/conv2d | bprop | conv2d | 196 | ✔ | ✔ |
| 324,310,584 | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Sequential::forward/Bottleneck::_call_impl/Bottleneck::forward/Conv2d::_call_impl[7]/Conv2d::forward/Conv2d::_conv_forward/conv2d | bprop | conv2d | 196 | ✔ | ✔ |
| 323,863,592 | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Sequential::forward/Bottleneck::_call_impl/Bottleneck::forward/Conv2d::_call_impl[7]/Conv2d::forward/Conv2d::_conv_forward/conv2d | bprop | conv2d | 196 | ✔ | ✔ |
| 317,166,280 | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Sequential::forward/Bottleneck::_call_impl/Bottleneck::forward/Conv2d::_call_impl/Conv2d::forward/Conv2d::_conv_forward/conv2d | bprop | conv2d | 196 | ✔ | ✔ |
| 317,123,764 | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Sequential::forward/Bottleneck::_call_impl/Bottleneck::forward/Conv2d::_call_impl/Conv2d::forward/Conv2d::_conv_forward/conv2d | bprop | conv2d | 196 | ✔ | ✔ |
| 229,397,613 | /module/train/backward | bprop | backward | 196 | ✔ | ✔ |
| 208,715,128 | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Sequential::forward/Bottleneck::_call_impl/Bottleneck::forward/Conv2d::_call_impl/Conv2d::forward/Conv2d::_conv_forward/conv2d | bprop | conv2d | 196 | ✔ | ✔ |
| 98,507,540 | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[7]/Sequential::forward/Bottleneck::_call_impl/Bottleneck::forward/Conv2d::_call_impl/Conv2d::forward/Conv2d::_conv_forward/conv2d | bprop | conv2d | 196 | ✔ | ✔ |
| 98,330,453 | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[7]/Sequential::forward/Bottleneck::_call_impl/Bottleneck::forward/Conv2d::_call_impl/Conv2d::forward/Conv2d::_conv_forward/conv2d | bprop | conv2d | 196 | ✔ | ✔ |
| 98,317,123 | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[7]/Sequential::forward/Bottleneck::_call_impl/Bottleneck::forward/Conv2d::_call_impl/Conv2d::forward/Conv2d::_conv_forward/conv2d | bprop | conv2d | 196 | ✔ | ✔ |

LUX PROVIDE high-performance computing center

# Know your software

## AI-oriented profilers: *NVidia DLProf*

### System Config

| | |
|---|---|
| GPU Count | 4 |
| GPU Name(s) | NVIDIA A100-SXM4-40GB<br>NVIDIA A100-SXM4-40GB<br>NVIDIA A100-SXM4-40GB<br>NVIDIA A100-SXM4-40GB |
| CPU Model | AMD EPYC 7452 32-Core Processor |
| GPU Driver Version | 515.65.01 |
| Framework | PyTorch 1.12.0 |
| CUDA Version | 11.7 |
| NSys Version | 2022.4.1.21-0db2c85 |
| DLProf CLI Version | v1.8.0 |
| DLProf DB Version | 1.8.0 |
| DLProf Viewer Version | 1.8.0 |

### Expert Systems

| | Problem | Recommendation |
|---|---|---|
| » | 2269 ops were eligible to use tensor cores but none are using FP16 | Try enabling AMP (Automatic Mixed Precision). For more information: https://developer.nvidia.com/automatic-mixed-precision |
| | The following GPU device IDs are not being used: 1 2 3 | Make sure to pass the necessary options to docker, the framework, and the model to use all available GPUs |
| | 12.1% of the aggregated run was spent in the dataloader while not simultaneously running on the GPU | Focus on reducing time spent in the training data input process. This could be time spent in file reading, preprocessing and augmentation or file transfer. Set num_workers > 0 in the dataloader to enable asynchronous data loading. Consider using NVIDIA DALI, a library that is a high performance alternative to built-in data loaders and data iterators. Learn more here: https://developer.nvidia.com/DALI |
| » | The aggregated iteration range of 0 to 196 contains a lot of variation | Try limiting the iteration range to a steady range by rerunning with the --database option and setting --iter_start=27 --iter_stop=83 |
| | Convolution operations were detected but torch.backends.cudnn.benchmark was not enabled. | Try setting torch.backends.cudnn.benchmark = True in your network. For best performance, the input shapes should be relatively stable. |
| | GPU Memory is underutilized: Only 12% of GPU Memory is used | Try increasing batch size by 4x to increase data throughput |

### Guidance

Understanding GPU utilization and timing details of the operations is the first step in profiling your model.

- To learn more about Tensor cores and Mixed Precision training, visit this site:https://developer.nvidia.com/tensor_cores
- You will find resources on how to train networks with mixed precision and make full use of Tensor cores for Tensorflow models here: https://docs.nvidia.com/deeplearning/sdk/mixed-precision-training/index.html#training_tensorflow
- Note that if there are multiple kernels being observed on single op, these are likely performing data transposes to prepare the data for efficient use by tensorcores. Such transposes themselves would not use tensor cores.

## Ops and Kernel Summaries

### Ops

Click an op to see its kernels below

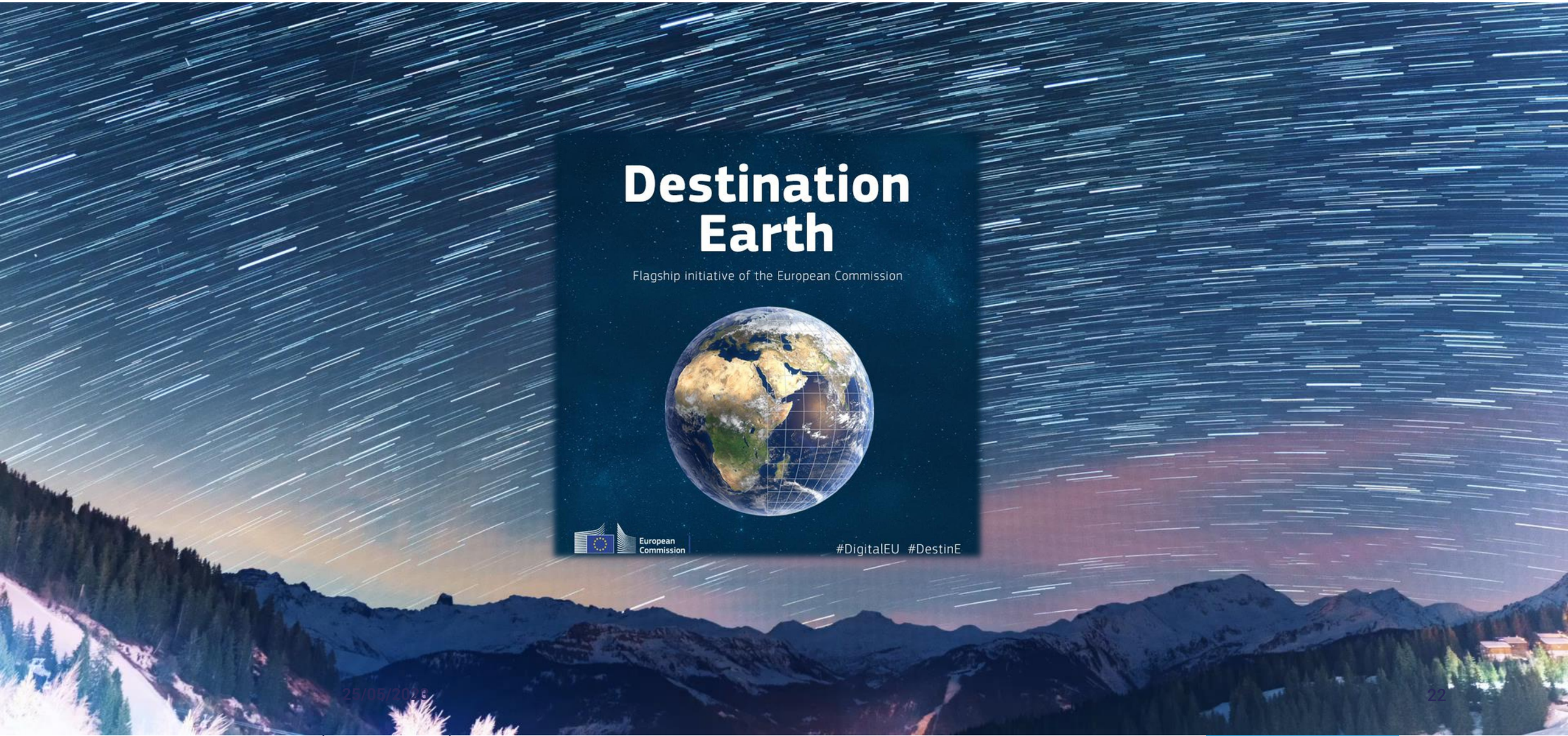
Show 10 entries  Export to: Excel PDF Clipboard CSV JSON  Search:

| GPU Time (ns) | CPU Time (ns) | Op ID | Op Name | Direction | Op Type | Calls | TC Eligible | Using TC | Kernel Calls | Data Type | Stack Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Search GPU Time | Search CPU Time (r | Search Op ID | Search Op Name | Search Direction | Search Op Type | Search Calls | 0 or 1 | 0 or 1 | Search Kernel Ca | Search Data Typ | Search Stack Trace |
| 324,433,581 | 3,109,669,488 | CONV2D_53_BPROP | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Seq See more | bprop | conv2d | 196 | ✔ | ✔ | 977 | float32 | /mnt/tier2/project/lxp/fbongiovanni/python_projects/pytorch-ddp/train_1gpu_prof.py:103 See more |
| 324,310,584 | 31,528,107 | CONV2D_46_BPROP | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Seq See more | bprop | conv2d | 196 | ✔ | ✔ | 977 | float32 | /mnt/tier2/project/lxp/fbongiovanni/python_projects/pytorch-ddp/train_1gpu_prof.py:103 See more |
| 323,863,592 | 31,719,988 | CONV2D_50_BPROP | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Seq See more | bprop | conv2d | 196 | ✔ | ✔ | 977 | float32 | /mnt/tier2/project/lxp/fbongiovanni/python_projects/pytorch-ddp/train_1gpu_prof.py:103 See more |
| 317,166,280 | 31,151,502 | CONV2D_48_BPROP | /module/train/ResNet::_call_impl/ResNet::forward/ResNet::_forward_impl/Sequential::_call_impl[8]/Seq See more | bprop | conv2d | 196 | ✔ | ✔ | 979 | float32 | /mnt/tier2/project/lxp/fbongiovanni/python_projects/pytorch-ddp/train_1gpu_prof.py:103 See more |

LUX PROVIDE high-performance computing center

# ESG IMPACT

LUX PROVIDE


Destination Earth
Flagship initiative of the European Commission
European Commission
#DigitalEU #DestinE

# Destination Earth (DestinE)

DestinE will provide unique digital modeling capabilities of the Earth to enhance the EU's ability to monitor and model environmental changes, predict extreme events, and adapt EU actions and policies to climate-related challenges.

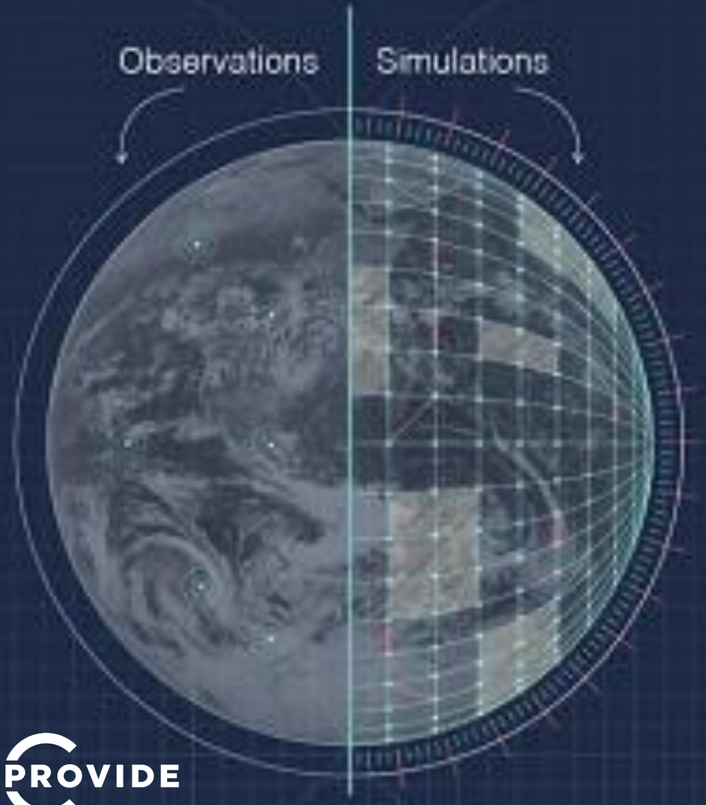
Destination Earth factsheet
digital-strategy.ec.europa.eu/en/library/destination-earth-factsheet

# UNDERSTAND THE PAST, PREDICT THE FUTURE

Fed by real-world observations, these digital twins let us understand what has happened on Earth – and what will happen in the decades ahead.

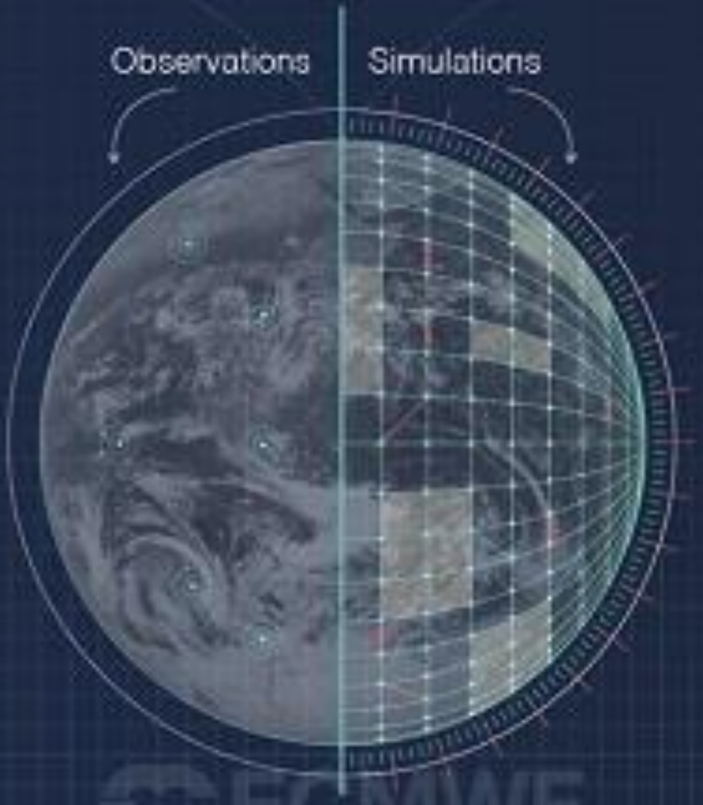
## PAST

Revisit the past and understand change



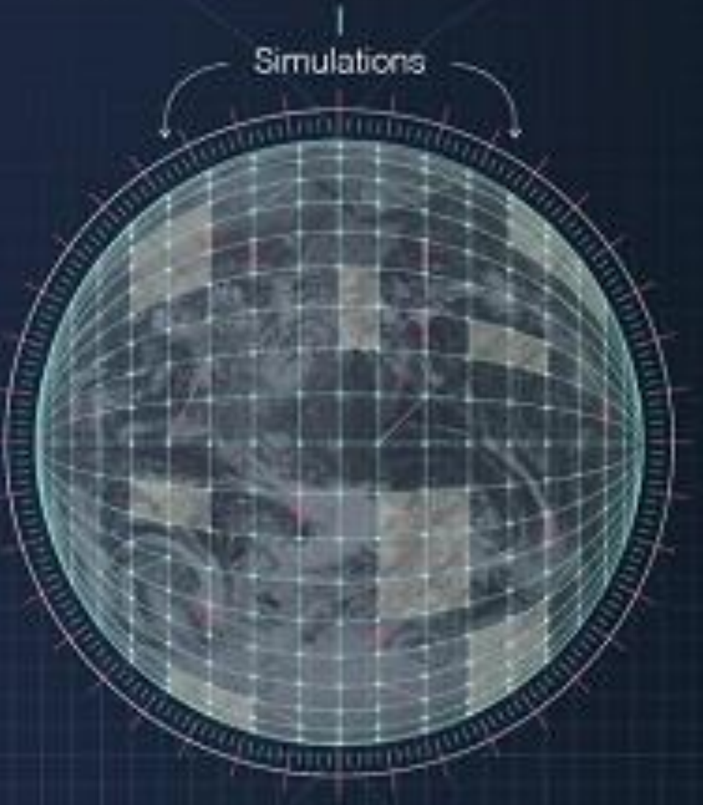

## PRESENT

Assess and monitor our current situation




## FUTURE

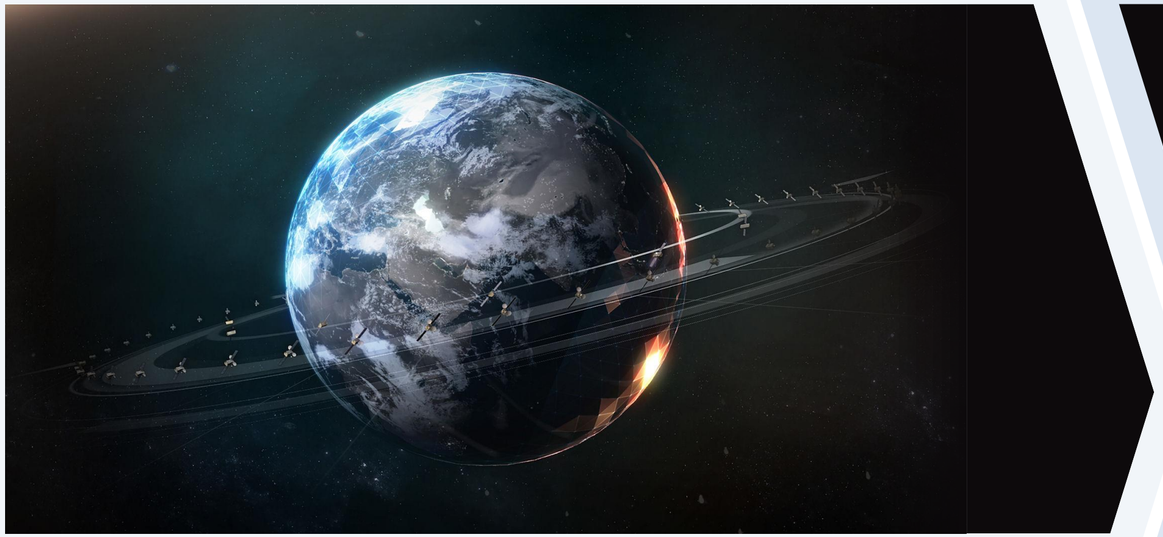
Predict future changes and test how we might intervene



# O3b mPOWER: Reimagining Satellites

## Terabit-level System for Cloud Everywhere on Earth

### O3b MEO Constellation



- 20-satellites
- 10 user beams per satellite
- Up to 1.6 Gbps per beam
- 9 SES data gateways globally

### O3b mPOWER Constellation

- 11-satellites
- Thousands of beams per satellite
- From tens of Mbps to multiple Gbps per beam
- Flexible data gateways

LUX PROVIDE high-performance computing center

SES

SES

O3b mPOWER

# Running Flood Depth Maps at X10 speeds


**Climate change-aware risk analysis**

**Flood risk commercial maps** from:

- 3rd party Earth Observation data
- 2-D flood modeling (+ HPC)

Showing flooding risk likelihood of **today**
+ including **climate change effects**:

- Likelihood 1:10
- Likelihood 1:100
- Likelihood 1:1000
- Customized

**10x speed** with CPU model version on MeluXina GPU nodes

LUX PROVIDE high-performance computing center

RSS-Hydro

LUX PROVIDE

# Thanks for your attention

luxprovide.lu

Follow us in