EuroCC@Greece

https://eurocc-greece.gr/newsletter/
https://www.linkedin.com/company/eurocc-greece
https://www.youtube.com/@euroccgreece9501
https://x.com/EuroCC_Greece

EuroCC@Greece

The overall objective of the Greek National Competence Center is to enable the efficient uptake of HPC technologies with the 3-fold goal to:

- advance competitiveness in **research**
- improve the effectiveness of **government** services and
- promote innovation in **industry**

EuroCC@Greece

Services

- Technological Support & Consulting
  - High-Performance Computing,
  - Artificial Intelligence, and
  - High-Performance Data Analytics
- Training and Skills Development
- Access to computational resources

https://eurocc-greece.gr/

Fields of Applications

- Artificial Intelligence
- Machine Learning
- Computer Vision
- Large Language Models
- Finite Elements Analyses
- Computational Fluid Dynamics
- Molecular Simulations
- Atmospheric & Oceanic Sciences

EuroCC@Greece

Consortium

The Greek National Competence Center “EuroCC@Greece”, is run by a consortium of 5 institutions, namely

1. National Infrastructures for Research and Technology (coordinator) - **GRNET**
2. National Center for Scientific Research - **Demokritos**
3. Institute of Communication and Computer Systems - **NTUA**
4. Aristotle University of Thessaloniki - **AUTH**
5. Foundation for Research and Technology Hellas - **FORTH**

The European High Performance Computing Joint Undertaking (EuroHPC JU)

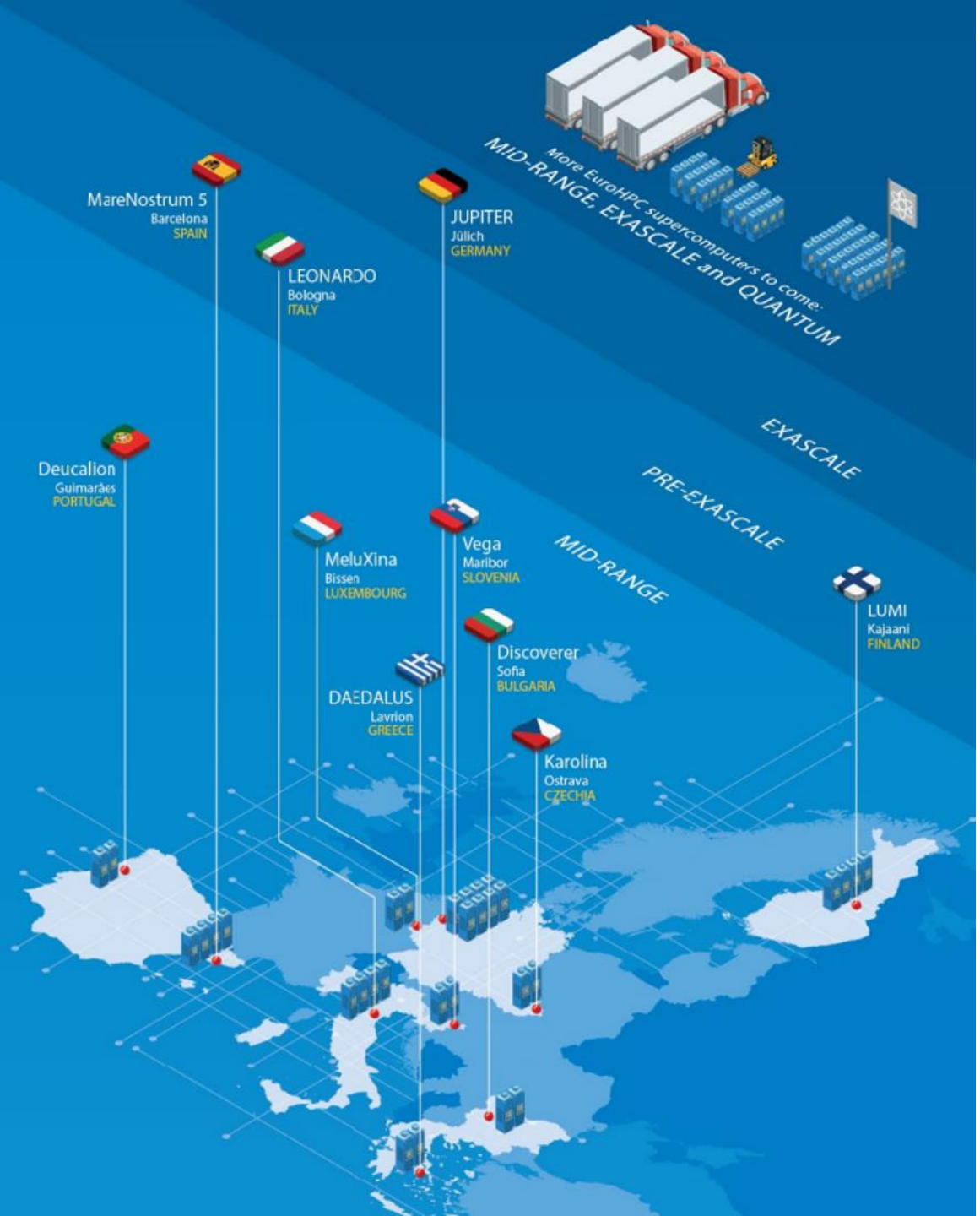
is a joint initiative between the EU, European countries and private partners to develop a World Class Supercomputing Ecosystem.

https://eurohpc-ju.europa.eu/index_en

BDV BIG DATA VALUE ASSOCIATION

QIC European Quantum Industry Consortium

ETP 4 HPC EUROPEAN TECHNOLOGY PLATFORM FOR HIGH PERFORMANCE COMPUTING



8 operational systems, all ranking among the world's most powerful supercomputers:

1. LUMI in Finland #5
2. LEONARDO in Italy #6
3. MARENOSTRUM in Spain
4. VEGA in Slovenia
5. MELUXINA in Luxembourg
6. KAROLINA in Czechia
7. DEUCALION in Portugal
8. DISCOVERER in Bulgaria

Underway:

JUPITER in Germany

DAEDALUS in Greece

1	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE
2	**Aurora** - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11, Intel
3	**Eagle** - Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR, Microsoft
4	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu
5	**LUMI** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE
6	**Leonardo** - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband, EVIDEN
7	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM

LUMI
FINLAND

LEONARDO
ITALY

MELUXINA
LUXEMBOURG

KAROLINA
CHECH REPUBLIC

DISCOVERER
BULGARIA

VEGA
SLOVENIA

DEUCALIO
PORTUGAL

MARENOSTRUM 5
SPAIN

1. **LUMI (CSC, Finland)**
   - LUMI-C: 1536 nodes, 128 cores/node, 256-1024 GB RAM/node
   - GPU: 2560 nodes, 64 cores/node, 4 GPUs, 128 GB GPU-RAM
   - Visualization: 64 nodes, 1 GPU, 48 GB GPU-RAM
   - Peak Performance: 550 petaflops
   - URL: https://www.lumi-supercomputer.eu/lumis-full-system-architecture-revealed/
2. **Leonardo (Cineca, Italy)**
   - Booster Module: 3456 nodes, 32 cores/node, 512 GB RAM/node, 4 GPUs, 64 GB GPU-RAM
   - Data Centric Module: 1536 nodes, 112 cores/node, 512 GB RAM/node
   - Peak Performance: 323.4 petaflops
   - URL: https://leonardo-supercomputer.cineca.eu/hpc-system/
3. **MareNostrum 5 (Barcelona Supercomputing Center, Spain)**
   - General Purpose Partition: 6408 nodes, 112 cores/node, 256 GB RAM/node
   - Accelerated Partition: 1120 nodes, 64 cores/node, 512 GB RAM/node, 4 GPUs, 64 GB GPU-RAM
   - Peak Performance: 314 petaflops
   - URL: https://www.bsc.es/innovation-and-services/marenostrum/marenostrum-5

4. **MeluXina (LuxProvide, Luxembourg)**
   - Cluster: 573 nodes, 128 cores/node, 512 GB RAM/node
   - Accelerator-GPU: 200 nodes, 64 cores/node, 512 GB RAM/node, 4 GPUs, 40 GB GPU-RAM
   - Large memory: 20 nodes, 128 cores/node, 4096 GB RAM/node
   - Peak Performance: 18.29 petaflops
   - URL: https://docs.lxp.lu/system/overview/
5. **Karolina (IT4I, Czech Republic)**
   - CPU: 828 nodes, 128 cores/node, 256-24000 GB RAM/node
   - GPU: 72 nodes, 8 GPUs, 40 GB GPU-RAM
   - Peak Performance: 15.69 petaflops
   - URL: https://www.it4i.cz/en/infrastructure/karolina

6. **Vega (IZUM, Slovenia)**
   - GPU partition: 60 nodes, 128 cores/node, 512 GB RAM/node, 4 GPUs, 40 GB GPU-RAM
   - CPU node Standard: 768 nodes, 128 cores/node, 256 GB RAM/node
   - CPU node Large Memory: 192 nodes, 128 cores/node, 1000 GB RAM/node
   - Peak Performance: 10.05 petaflops
   - URL: https://doc.vega.izum.si/architecture/
7. **Deucalion (Guimarães, Portugal)**
   - ARM cluster: 1632 nodes, 48 cores/node
   - X86 cluster: 500 nodes, 48+ cores/node
   - Accelerated partition: 33 nodes
   - Peak Performance: 10 petaflops
   - URL: https://macc.fccn.pt/resources#deucalion
8. **Discoverer (Sofia Tech Park, Bulgaria)**
   - CPU: 1128 nodes, 128 cores/node, 256 GB RAM/node
   - CPU-Fat: 18 nodes, 128 cores/node, 1000 GB RAM/node
   - Peak Performance: 5.94 petaflops
   - URL: https://docs.discoverer.bg/resource_overview.html

ARIS – HPC Infrastructure in Greece Compute Nodes

The ARIS infrastructure consists of a total of five computing system nodes based on Intel x86 architecture, interconnected into a single InfiniBand FDR14 network offering multiple options and processing architectures. More specifically, the infrastructure consists of:

- **Thin Nodes:** 426 IBM NeXtScale nodes, Intel Xeon E5-2680v2, 8,520 cores.
- **Fat Nodes:** 44 Dell PowerEdge R820, 4 Intel Xeon E5-4650v2, 512 GB memory per node.
- **GPU Nodes:** 44 Dell PowerEdge R730, 2 Intel Xeon E5-2660v3, 64 GB memory, 2 NVIDIA K40 GPUs per node.
- **Xeon Phi Nodes:** 18 Dell PowerEdge R730, 2 Intel Xeon E5-2660v3, 64 GB memory, 2 Xeon Phi 7120P co-processors per node.
- **ML Node:** 1 server, 2 Intel E5-2698v4, 512 GB memory, 8 NVIDIA V100 GPUs.

EuroCC@Greece

Publications

Greek supercomputer ARIS

https://www.hpc.grnet.gr/en/publications/

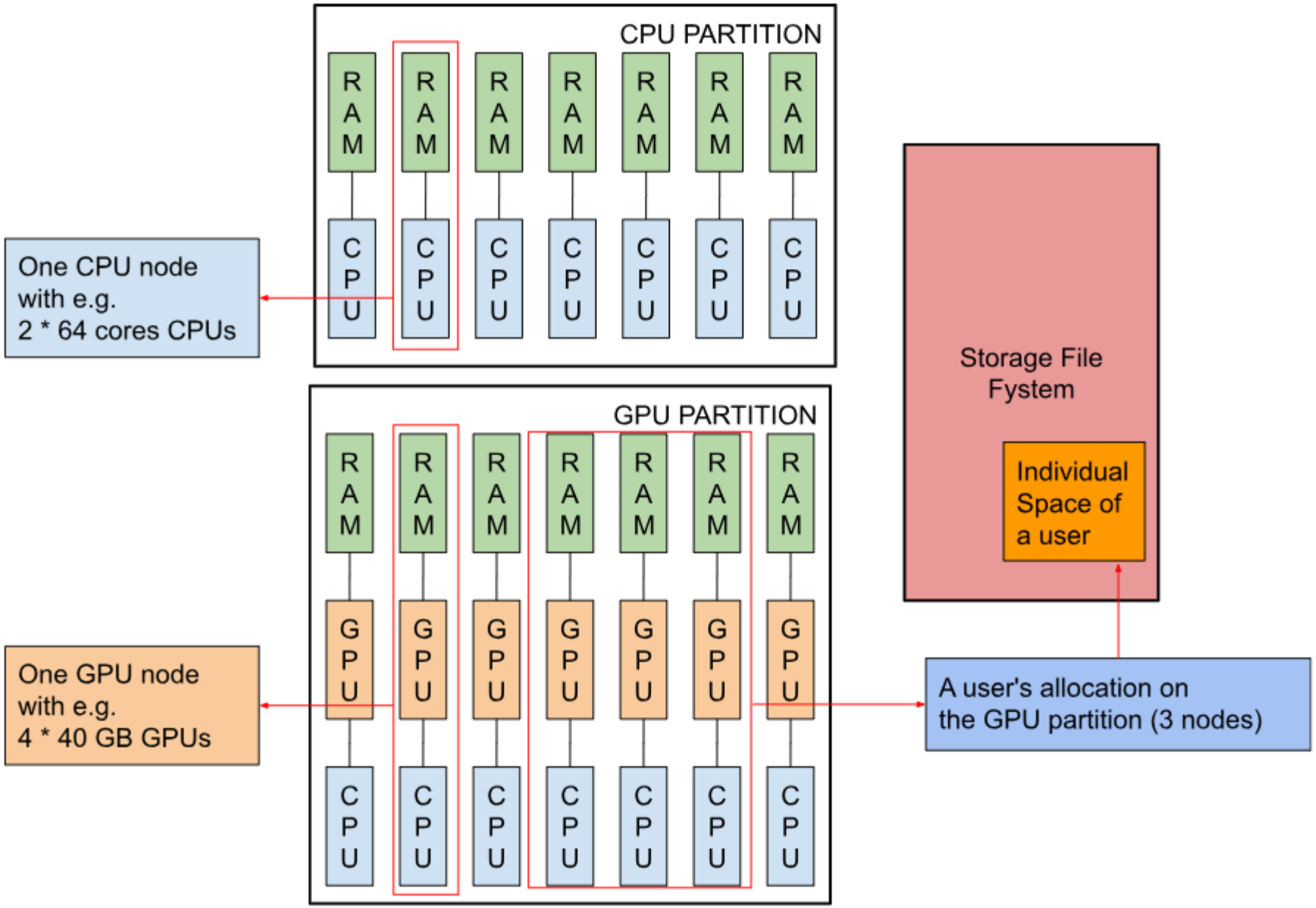
- In **fluid dynamics**, HPC powers deep learning models for super-resolution imaging and **turbulent flow reconstruction**, along with complex **multi-phase flow simulations**.
- **Materials science** benefits from HPC-driven **machine learning** and **molecular dynamics**, refining interatomic potentials for **pharmaceuticals**, and investigating **polymer mechanics**.
- **Computational chemistry** leverages quantum mechanical calculations for **thermoelectric materials**, **electronic structures**, and **drug binding** studies.
- **Astrophysics** research utilizes HPC for **modeling neutron star** thermoelectric effects and pulsar equations.
- **Atmospheric and oceanic sciences** apply HPC to turbulence modeling in **air pollution studies**, **weather forecasting**, and sea surface simulations.
- Additionally, HPC enhances radiation modeling for space applications and **Monte Carlo-based dosimetry** calculations, underscoring its vital role in advancing interdisciplinary research.

The way is open to building a EuroHPC world-class supercomputer in Greece

- A hosting agreement has been signed between the EuroHPC Joint Undertaking and the National Infrastructures for Research and Technology (**GRNET**) in **Greece**, where **DAEDALUS**, a new **EuroHPC supercomputer**, will be located.

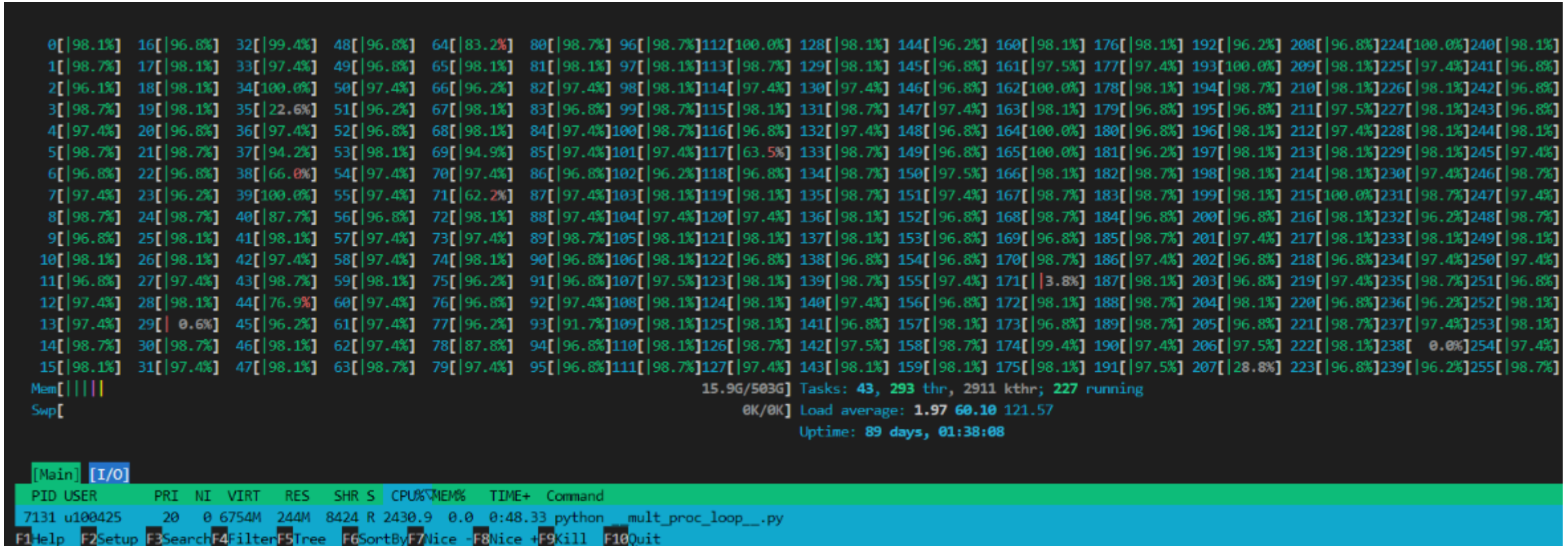
- **60 petaflops** or 60 million billion calculations per second https://grnet.gr/en/business-directory/grant-for-the-development-of-a-new-national-hpc-system-daedalus/
- **Lavrion** Technological and Cultural Park (TCPL) https://eurohpc-ju.europa.eu/way-open-building-eurohpc-world-class-supercomputer-greece-2022-11-28_en
- June 11, 2024: GRNET S.A. conducts a Public Consultation on the Open Tender Announcement Issue https://grnet.gr/2024/06/11/public-consultation-lavrio-daedalus/

Example of an HPC cluster

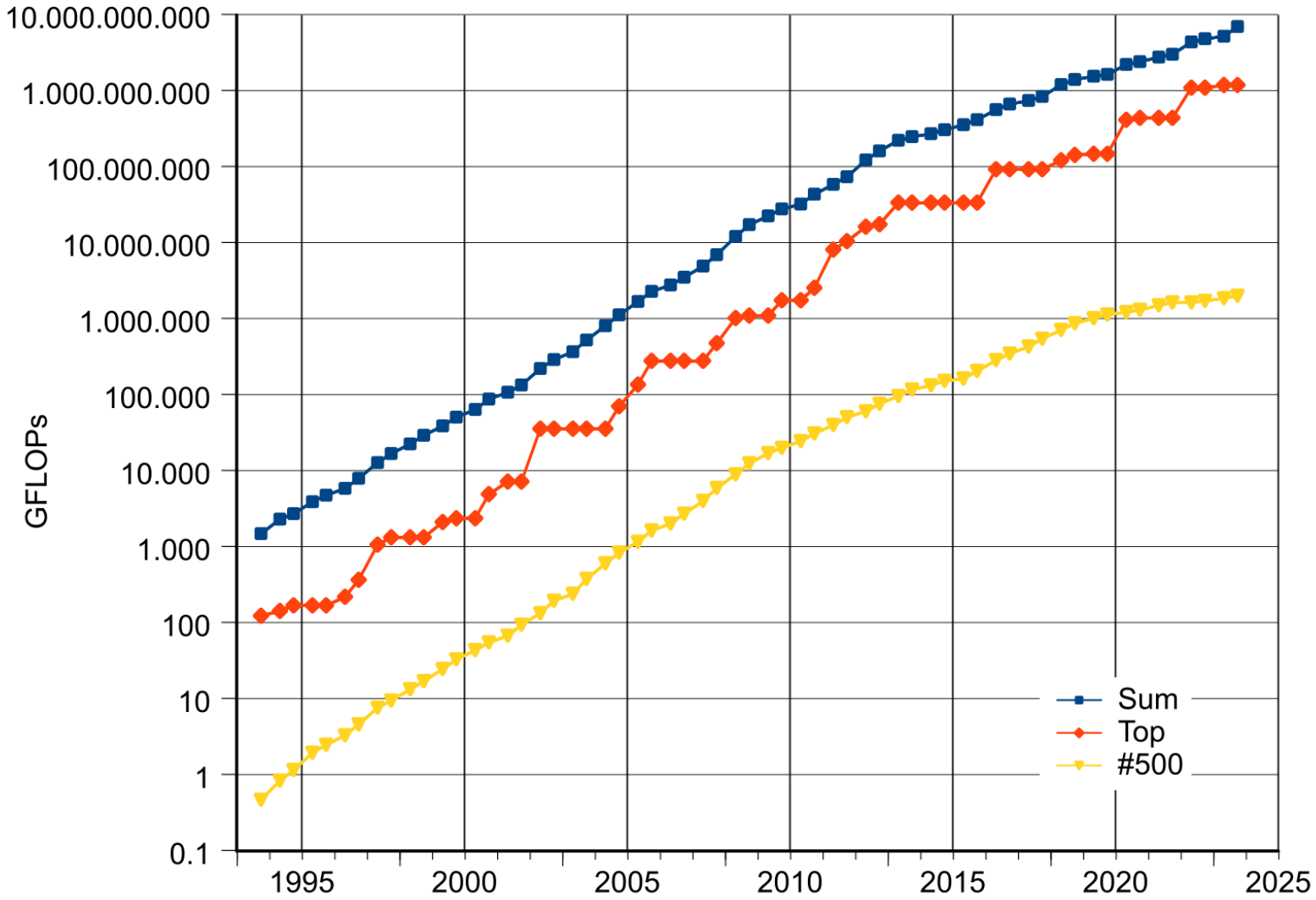


Threads on 1 node of MeluXina supercomputer

https://docs.lxp.lu/system/overview/



Growth of HPC systems



The new **El Capitan** system at the **Lawrence Livermore National Laboratory** in California, U.S.A., has debuted as the most powerful system on the list with an HPL score of 1.742 EFlop/s.
https://top500.org/

X 200,000

0.05 km/h

1 km/h

100 km/h

1,000 km/h

10,000 km/h

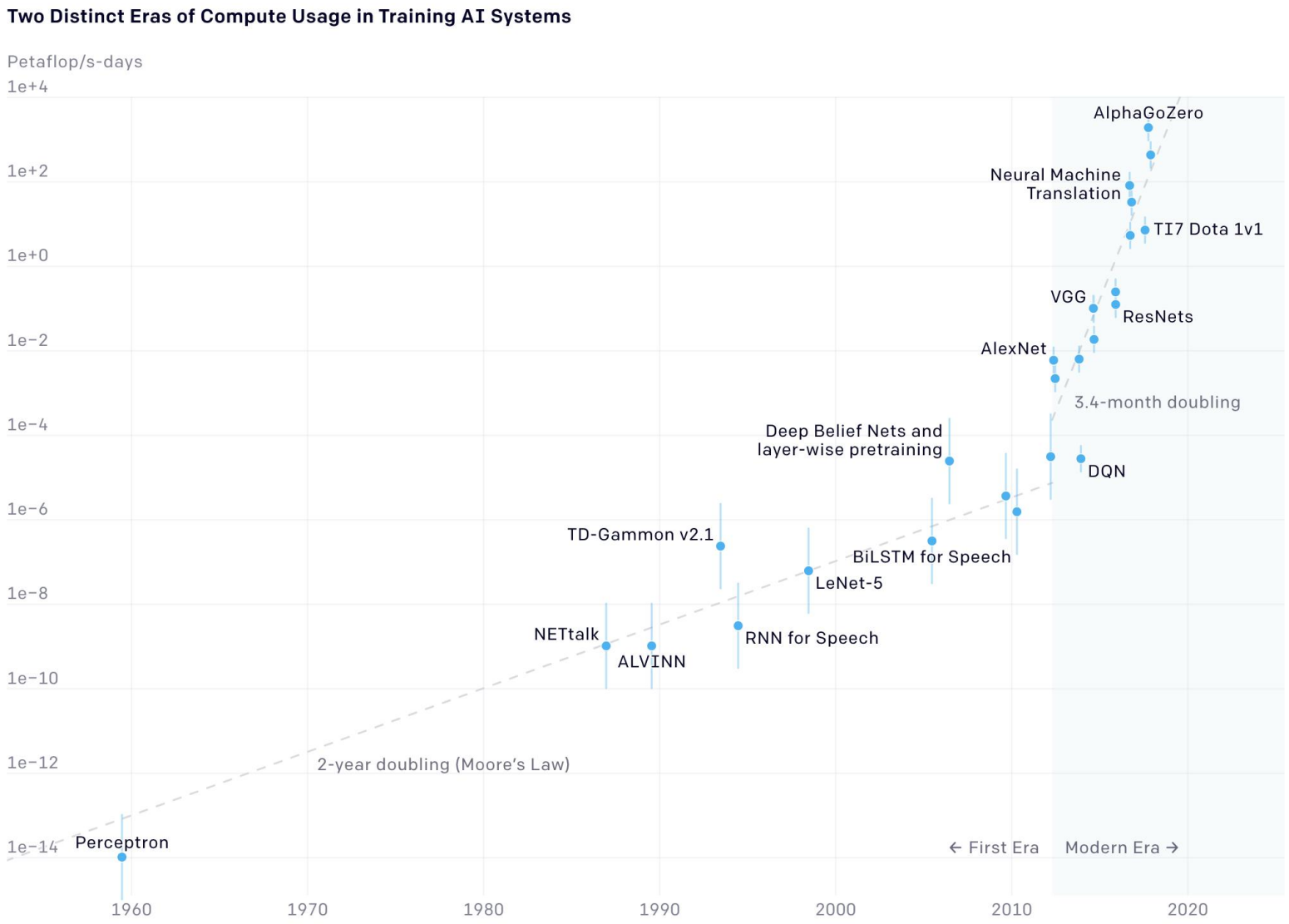
1,000,000,000 km/h

1 TFlop

X 1,000,000

1 EFlop

Since 2102 we observe a 3.4-month doubling in computing power used to train AI models.



https://openai.com/research/ai-and-compute

Large Language Models on HPC

Estimated GPU Hours for Training:

1.Small LLM (~8B):

- ~1.3M GPU hours (LLAMA 3 8B).

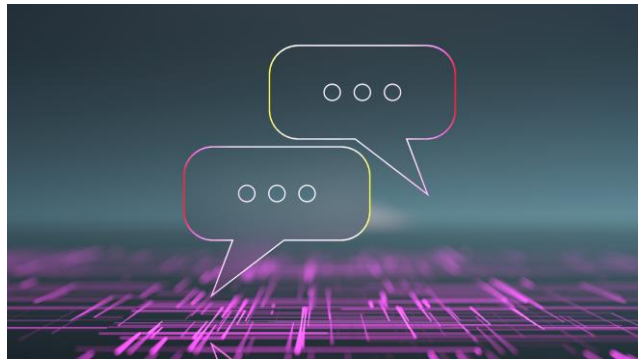
2.Medium LLM (~70B):

- ~6.4M-7.0M GPU hours (LLAMA 3/3.1 70B).

3.Large LLM (~405B):

- ~30.84M GPU hours (LLAMA 3.1 405B).
- Falcon 180B (slightly smaller): ~7M GPU hours.

https://huggingface.co/meta-llama/Meta-Llama-3-8B-Instruct
https://huggingface.co/meta-llama/Llama-3.3-70B-Instruct
https://huggingface.co/meta-llama/Llama-3.1-405B-Instruct
https://docs.lxp.lu/howto/llama3-vllm/
https://huggingface.co/blog/falcon-180b



Estimated GPU Requirements for Inference:

- **Small LLM (~8B):** ~80GB GPU RAM (LLAMA 3.1 8B).
- **Medium LLM (~70B):** ~320GB GPU RAM (GPTQ/int4 on Falcon).

Large LLM (~405B): ~800GB GPU RAM (FP8 on LLAMA 3.1).

EuroHPC Access Modes

EuroHPC JU Call for Proposals – Extreme Scale Access Mode
For applications with high-impact, high-gain innovative research

EuroHPC JU Call for Proposals – Regular Access Mode
The expected impact in the application's domain should justify the need for large allocations

EuroHPC JU Call for Proposals – AI and Data-Intensive Applications Access Mode
To support ethical artificial intelligence & machine learning

EuroHPC JU Call for Proposals – Development Access Modes
To develop, test and optimise applications

EuroHPC JU Call for Proposals – Benchmark Access Modes
To test or benchmark applications

https://eurohpc-ju.europa.eu/access-our-supercomputers/access-policy-and-faq_en

https://access.eurohpc-ju.europa.eu/

2025 Cut off dates for EuroHPC Access Calls

BENCHMARK ACCESS:

- 1st day of each month

DEVELOPMENT ACCESS:

- 1st day of each month

AI AND DATA INTENSIVE APPLICATIONS ACCESS:

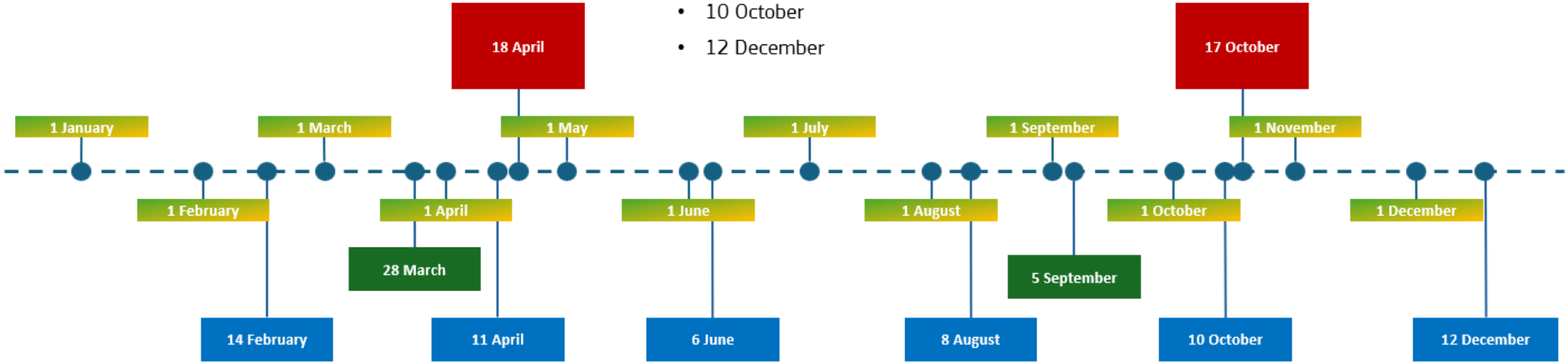
- 14 February
- 11 April
- 6 June
- 8 August
- 10 October
- 12 December

REGULAR ACCESS:

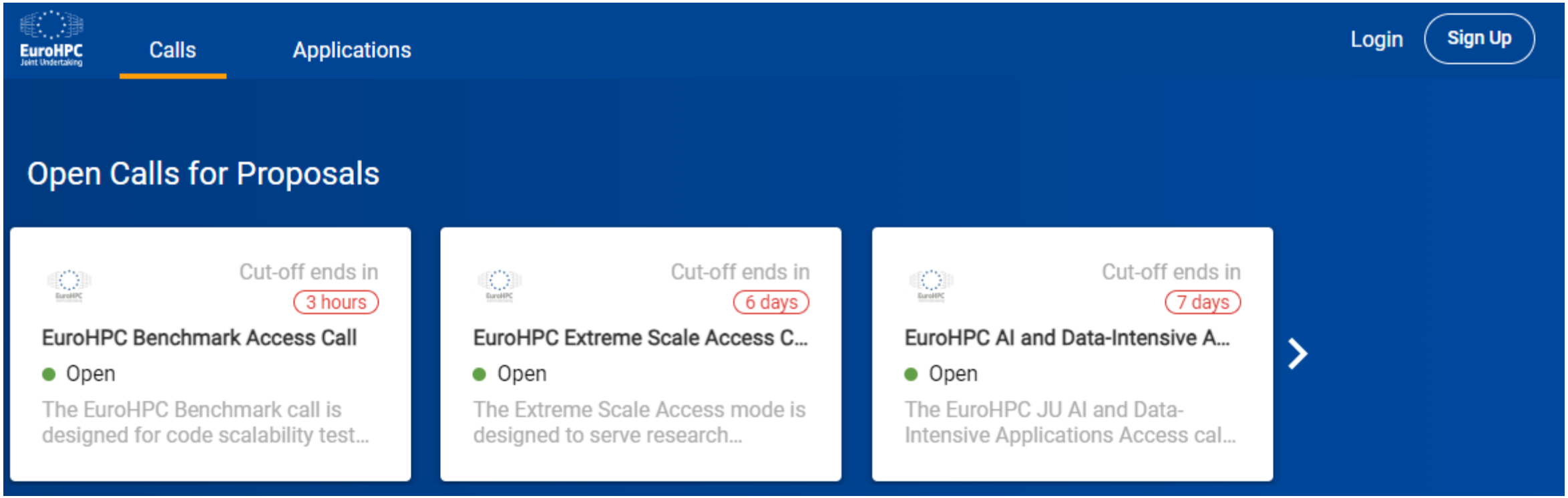
- 28 March
- 5 September

EXTREME SCALE ACCESS:

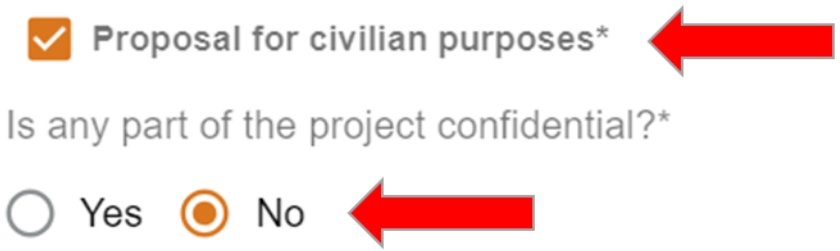
- 18 April
- 17 October

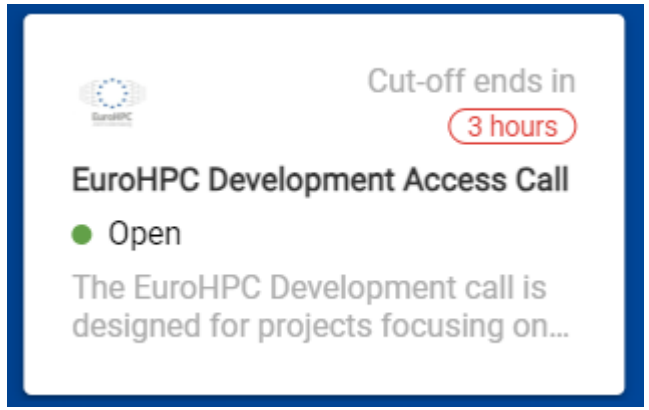


https://eurohpc-ju.europa.eu/access-our-supercomputers/access-policy-and-faq_en



https://access.eurohpc-ju.europa.eu/





Research fields

Research field title*

PE6 Computer Science and Informatics

Research field sub-title*

PE6_7 Artificial intelligence, intelligent systems, natural language processing

Research field share (%)*

50

The sum of all research fields should not exceed the total of 100%

https://access.eurohpc-ju.europa.eu/

Research fields #2

Research field title*

PE6 Computer Science and Informatics

Research field sub-title*

PE6_11 Machine learning, statistical data processing and applications using signal

Research field share (%)*

50

The sum of all research fields should not exceed the total of 100%

Remove + Research fields

AI set of technologies selection

Machine Learning | Natural Language Processing | Deep Learning

If applicable, please select used AI technologies. This is a multi-select field so you are able to choose more than one option.

https://access.eurohpc-ju.europa.eu/

Partitions

Partition name*

MeluXina CPU

Code(s) used*

XGBoost MPI Horovod Pytorch

This field is a multi-text field, for adding another code separate it with a comma

Requested amount of resources (node hours)*

4 000

Average number of processes/threads*

128

Average job memory (total usage over all nodes in GB)*

400

Maximum amount of memory per process/thread (MB)*

10 000

Total amount of data to transfer to/from (GB)*

100

https://access.eurohpc-ju.europa.eu/

Partitions #2

Partition name*

MeluXina GPU

Code(s) used*

Llama Falcon Mistral

This field is a multi-text field, for adding another code separate it with a comma

Requested amount of resources (node hours)*

800

Average number of processes/threads*

64

Average job memory (total usage over all nodes in GB)*

800

Maximum amount of memory per process/thread (MB)*

12 500

https://access.eurohpc-ju.europa.eu/

Frequently Asked Questions (FAQ)

https://eurohpc-ju.europa.eu/access-our-supercomputers/access-policy-and-faq_en

- **How can I gain access to computation time on EuroHPC machines?**
  - You will need to **apply** to one of the open **access calls** that **EuroHPC** provides. The list of available calls can be found here.
- **Which organisations are eligible for access to EuroHPC machines?**
  - Any European organisation is eligible for access to perform Open Science research (the **results** of the work are made available for **open access**). This includes **public and private** academic and research institutions, public sector organisations, industrial enterprises and SMEs
- **What is the cost?**
  - Currently access is **free of charge**.
- **What are the participation conditions?**
  - Participation conditions depend on the specific access call that a research group has applied. In general users of EuroHPC systems commit to: **acknowledge** the use of the **resources** in their related publications, contribute to **dissemination** events, produce and submit a **report** after completion of a resource allocation. More information on participation conditions can be found in the call's Documents section.

Our Training Events https://eurocc-greece.gr/events-2/

MARCH 29 | 09:45 EET | ONLINE

APRIL 19, 2024 | 10:00 EET | ONLINE

APRIL 24, 2024 | 10:00 EET | ONLINE

MAY 29, 2024 | 10:00 EET | HYBRID

JULY 11-14, 2024 | 10:00 - 16:00 EET | ON-SITE

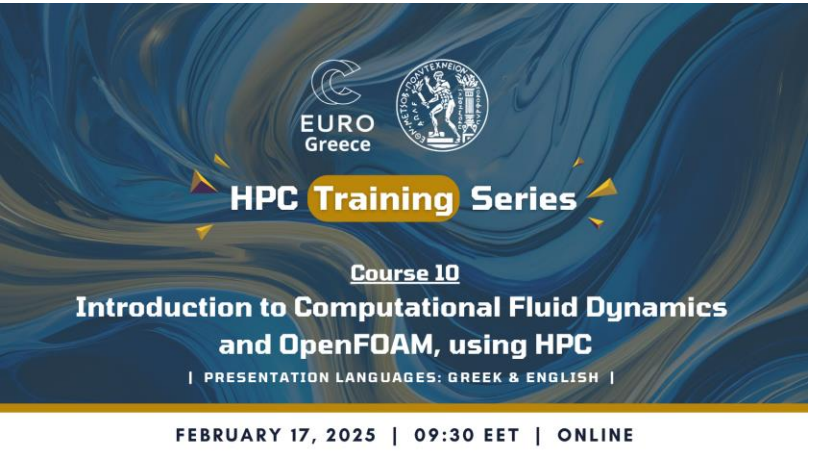
JUNE 14, 2024 | 09:30 EET | ONLINE

SEPTEMBER 2-3, 2024 | 09:00 - 17:00 EET | ATHENS, GREECE

OCTOBER 4, 2024 | 10:00 EET | ONLINE

NOVEMBER 1, 2024 | 10:00 EET | ONLINE

Our Training Events https://eurocc-greece.gr/events-2/



HPC Training Series

Course 12

Introduction to Accelerators: GPUs / CUDA

10:00 → 10:10 **Introduction to EuroCC & the training events**
Speaker: Dr Nikolaos Bakas (GRNET)

10:10 → 10:20 **How to access the EuroHPC-JU supercomputers**
Speaker: Dr Nikolaos Bakas (GRNET)

10:20 → 10:30 **How to access the Greek HPC Infrastructure ARIS**
Speaker: Mr Nikolaos Triantafyllis (GRNET)

10:30 → 12:30 **Introduction to modern graphics processing units (GPU) architecture and programming in CUDA**
Speaker: Dr Xenophon Trompoukis (NTUA)

12:30 → 13:15 **Performance optimizations in CUDA**
Speaker: Prof. Ioannis Venetis (UniPi)

13:15 → 14:00 **Usage of high-performance libraries for GPUs**
Speaker: Prof. Ioannis Venetis (UniPi)

14:00 → 14:30 **SCALE: a Cross-Vendor extension of the CUDA Programming Model for GPUs**
Speaker: Dr Manos Pavlidakis (Scale)

14:30 → 14:40 **Q&A**
Speaker: Dr Nikolaos Bakas (GRNET)

EuroCC@Greece

https://eurocc-greece.gr/newsletter/
https://www.linkedin.com/company/eurocc-greece
https://www.youtube.com/@euroccgreece9501
https://twitter.com/EuroCC_Greece

This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 951732. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, United Kingdom, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Switzerland, Turkey, Republic of North Macedonia, Iceland, Montenegro