

22 e 23
OUTUBRO

encontro de computação avanzada 2025



Universidade de Aveiro,
Edifício da Reitoria

FCCN



CNCA



universidade de aveiro
UNIVERSIDADE DE AVEIRO

Parceiros:

comudel



Dias de Sousa
INSTITUTO DE INVESTIGAÇÃO EM ANÁLISE E GESTÃO DE SISTEMAS

Apoio:



EUROCC PORTUGAL



EuroHPC
EUROPEAN HPC JOURNAL



PRR
PORTUGAL RESEARCH RESEARCH



Financiado pela
União Europeia
NextGenerationEU



October 2025



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



EXCELENCIA
SEVERO
OCHOA



MareNostrum 5 System and Support

David Vicente

HPC and AI user support manager

david.vicente@bsc.es

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 programme, as well as the Participating States Spain, Portugal and Türkiye



Hosting Consortium:



Spain



Portugal



Türkiye





BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

MareNostrum 5

Total peak performance: **315.2 Pflops**

General Purpose Partition: 46.4 Pflops (29-04-2024)

Accelerated Partition: 260 Pflops (29-04-2024)

Next Generation GPP: 2.82 Pflops (WiP)

Next Generation ACC: 6 Pflops (TBA)



MareNostrum 1

2004 – 42.3 Tflops

1st Europe / 4th World

New technologies

MareNostrum 2

2006 – 94.2 Tflops

1st Europe / 5th World

New technologies

MareNostrum 3

2012 – 1.1 Pflops

12th Europe / 36th

World

MareNostrum 4

2017 – 11.1 Pflops

2nd Europe / 13th

World

New technologies

MareNostrum 5

2022

260 + 46.4 Pflops

8th and 19th World

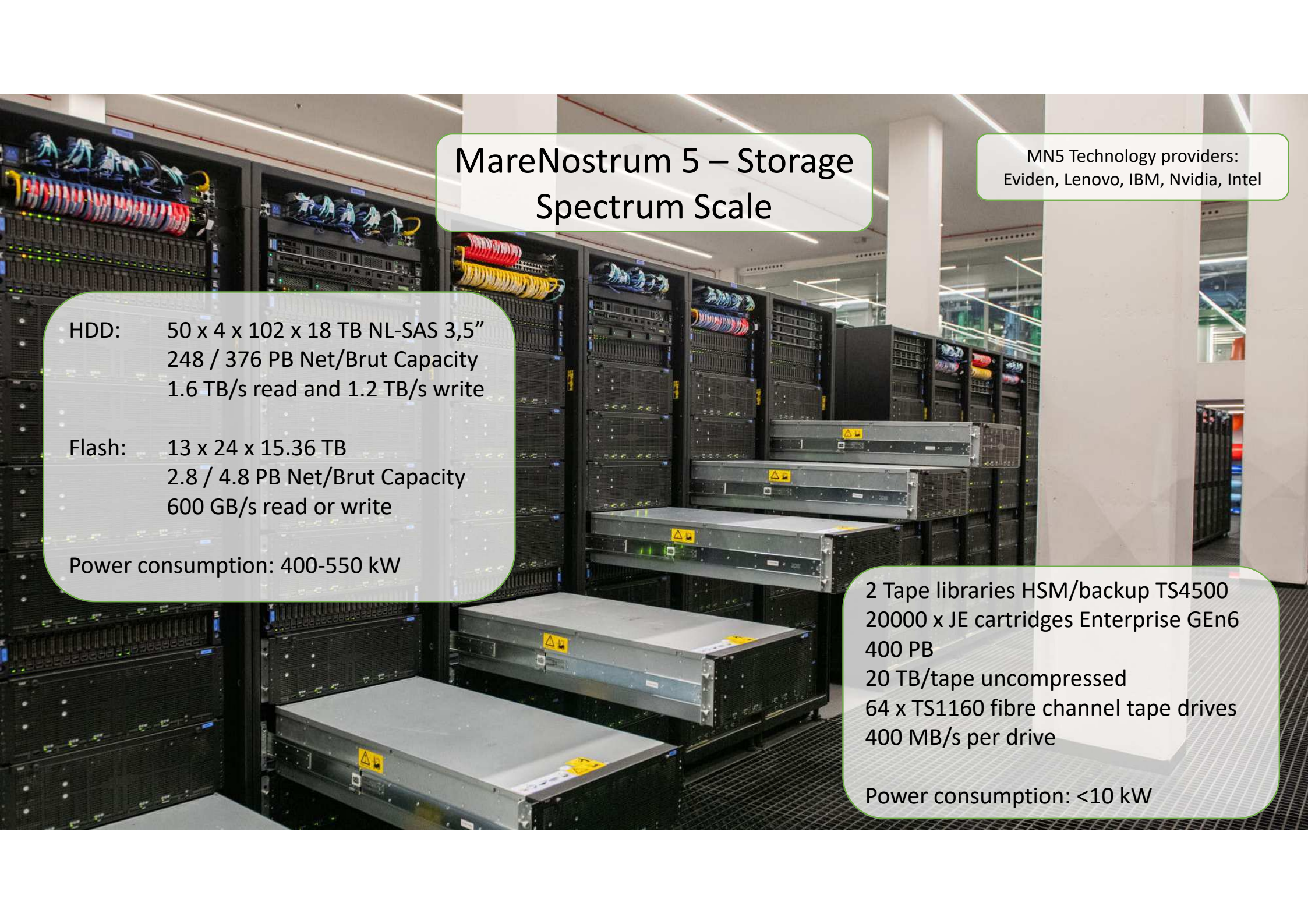
3rd and 7th Europe

MareNostrum5



Top 23, June 2024

Rank	Name	Country	Cores	Accelerators	Rmax [TFlop/s]	Rpeak [TFlop/s]	GFlops/Watts	Computer	Site
1	Frontier	United States	8.699.904	8.138.240	1.206.000	1.714.814	52,93	HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	DOE/SC/Oak Ridge National Laboratory
2	Aurora	United States	9.264.128	8.159.232	1.012.000	1.980.006	26.15	HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11	DOE/SC/Oak Ridge National Laboratory
3	Eagle	United States	2.073.600	1.900.800	561.200	846.835		Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR	Microsoft Azure
4	Fugaku	Japan	7.630.848		442.010	537.212	14,78	Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D	RIKEN Center for Computational Science
5	LUMI	Finland	2.752.704	2.566.080	379.700	531.505	53,43	HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	EuroHPC/CSC
6	Alps	Switzerland	1.305.600	844.800	270.000	353.748	51,98	HPE Cray EX235a, AMD Optimized 3 rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	CSCS
7	Leonardo	Italy	1.824.768	1.714.176	241.200	306.311	32,19	BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, 4x NVIDIA 100	EuroHPC/CINECA
8	MareNostrum 5 ACC	Spain	663.040	591.360	175.300	249.435	42,15	BullSequana XH3000, Xeon Platinum 8460Y+ 40C 2.3GHz, NVIDIA H100 64GB, Infiniband NDR200	EuroHPC/BSC
9	Summit	United States	2.414.592	2.211.840	148.600	200.795	14,72	IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	DOE/SC/Oak Ridge National Laboratory
10	Eos NVIDIA DGX SuperPOD	United States	485.888	439.296	121.400	188.645		NVIDIA DGX H100, Xeon Platinum 8480C 56C 3.8GHz, NVIDIA H100, Infiniband NDR400	NVIDIA Corporation
11	Venado	United States	481.440	311.520	98.510	130.44	58,29	HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11	DOE/NNSA/LANL
12	Sierra	United States	1.572.480	1.382.400	94.640	125.712	12,72	IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	DOE/NNSA/LLNL
13	Sunway TaihuLight	China	10.649.600		93.015	125.436	6,05	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	National Supercomputing Center in Wuxi
14	Perlmutter	United States	888.832	774.144	79.230	112.998	26,90	HPE Cray EX 235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-11	DOE/SC/LBNL/NERSC
15	Selene	United States	555.520	483.840	63.460	79.215	23,98	NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband	NVIDIA Corporation
16	Tianhe-2A	China	4.981.760	4.554.752	61.444	100.679	3,32	TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000	National Super Computer Center in Guangzhou
17	CEA-HE	France	389.232	251.856	57.110	112.560	47,32	BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail Bxi v2	Commissariat a l'Energie Atomique (CEA)
18	Explorer-WUS3	United States	445.440	422.400	53.960	86.987		ND96_amsr_MI200_v4, AMD EPYC 7V12 48C 2.45GHz, AMD Instinct MI250X, Infiniband HDR	West US3
19	ISEG	Netherlands	218.880	200.640	46.540	86.792	35,26	Gigabyte G593-SD0, Xeon Platinum 8468 48C 2.1GHz, NVIDIA H100 SXM5 80 GB, Infiniband NDR400	Nebius
20	Adastr	France	319.072	297.440	46.100	61.608	50,03	HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11	GENCI-CINES
21	JUWELS Booster Module	Germany	449.280	404.352	44.120	70.980	25,01	Bull Sequana XH2000 , AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite	Forschungszentrum Juelich (FZJ)
22	MareNostrum 5 GPP	Spain	725.760		40.102	46.371	6,97	ThinkSystem SD650 v3, Xeon Platinum 03H-LC 56C 1.7GHz, Infiniband NDR200	EuroHPC/BSC
23	Shaheen III - CPU	Saudi Arabia	877.824		35.658	39.607	6,73	HPE Cray EX, AMD EPYC 9654 96C 2.4GHz, Slingshot-11	King Abdullah University of Science and Technology



MareNostrum 5 – Storage Spectrum Scale

MN5 Technology providers:
Eviden, Lenovo, IBM, Nvidia, Intel

HDD: 50 x 4 x 102 x 18 TB NL-SAS 3,5"
248 / 376 PB Net/Brut Capacity
1.6 TB/s read and 1.2 TB/s write

Flash: 13 x 24 x 15.36 TB
2.8 / 4.8 PB Net/Brut Capacity
600 GB/s read or write

Power consumption: 400-550 kW

2 Tape libraries HSM/backup TS4500
20000 x JE cartridges Enterprise GEn6
400 PB
20 TB/tape uncompressed
64 x TS1160 fibre channel tape drives
400 MB/s per drive

Power consumption: <10 kW



MareNostrum5 GPP



Racks	Cooling	Nodes		Processor/Accelerator		Memory	PFlops (HPL)	Local Drive	High-Perf. Network
		Total	per rack						
89	DLC + RDHX	6192	72 (6x6x2)	2x Intel Sapphire R. 8480+	56c @ 2GHz	>2GB/core 256GB DDR5	40.10	960GB NVMe	1x NDR200 Shared by 2 nodes
		216				>8GB/core 1024GB DDR5			
1		72		2x Intel Sapphire R. 9480	56c @ 1.9GHz	> 0.5GB HBM/core 128GB HBM + 32GB DDR5	0.34		



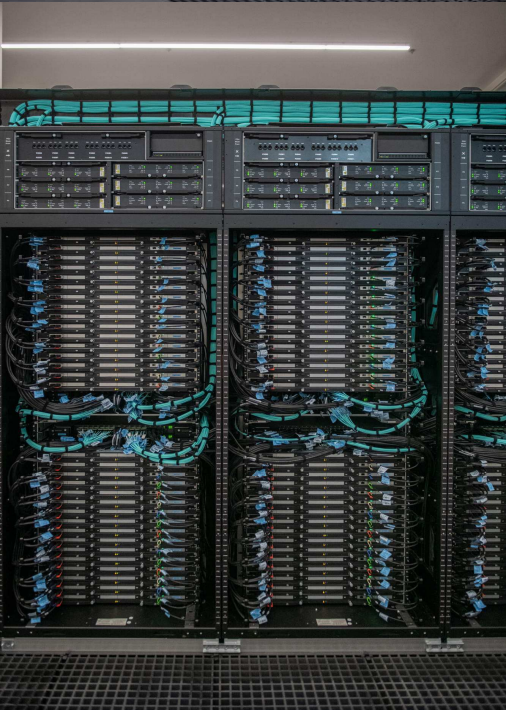
November 2023

HPL: #19 , #1 x86
HPCG: #24
Green500: #81
5.7 MW under HPL

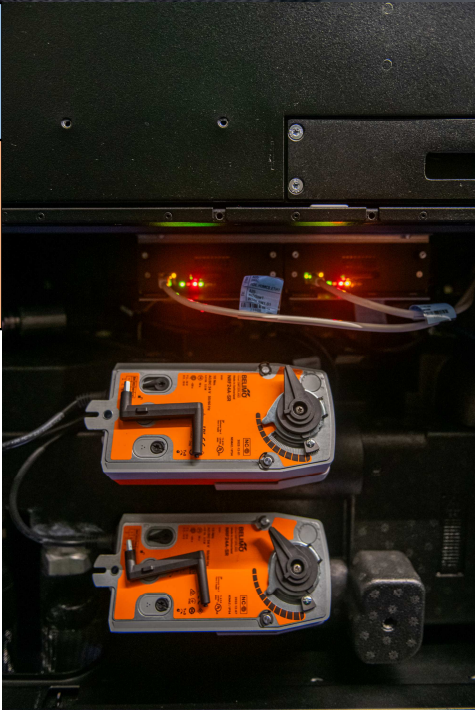
HPL: 40.10 PFlops
HPCG: 484.36 TFlops
Green500: 6.97 Gflops/watt



MareNostrum5
ACC



Racks	Cooling	Nodes		Processor/Accelerator		Memory	PFlops (HPL)	Local Drive	High-Perf. Network
		Total	per rack						
35	DLC	1120	32	2x Intel Sapphire R. 8460Y+	40c @ 2GHz	512GB	175.3	480GB NVMe	4x NDR200
				4x Nvidia Hopper 64GB HBM					



November 2023
HPL: #8
Green500: #5

HPL: 175.30 Pflops
HPCG: 1.146 PFlops
Green500: 42.15 Gflops/watt
Graph500: 15.73e+12 GTEPS (BFS)
4.1 MW under HPL (175 PF/s)

Nvidia Hopper BF16: 120 TFLOPs, Tensorcore: 1 PFLOPs
MN5 Total performance BF16: 536 PFLOPs, Tensorcore: 4.48 EFLOPs



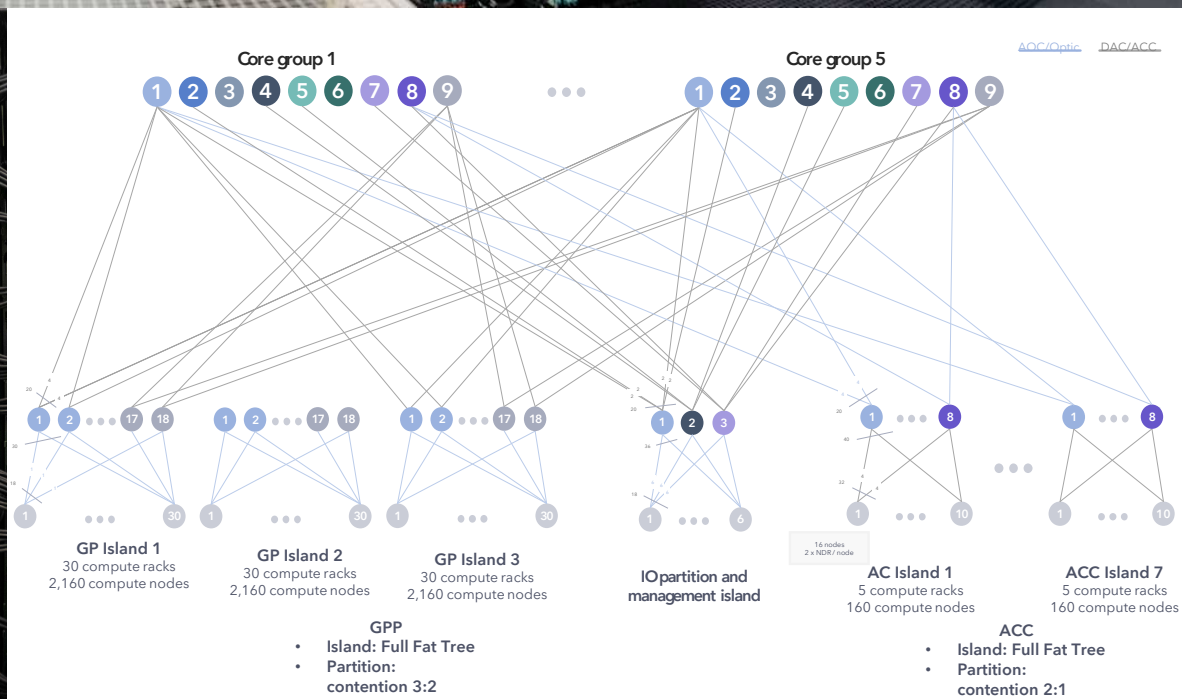
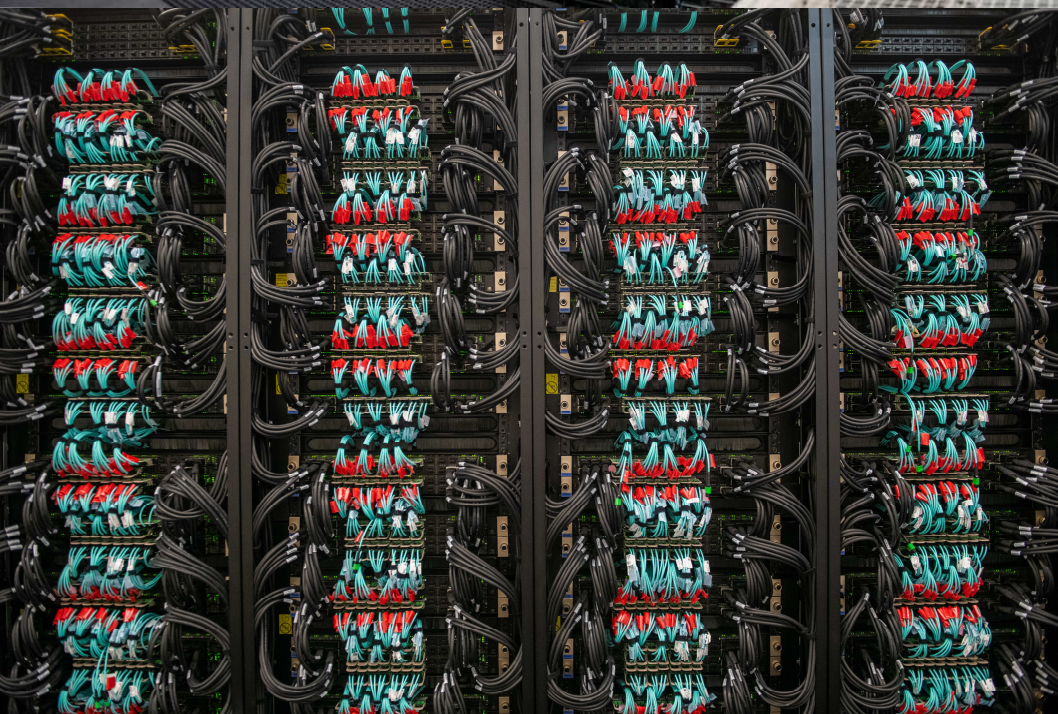
MareNostrum5
NG GP



Racks	Cooling	Nodes		Processor/Accelerator		Memory	PFlops (HPL)	Local Drive	High-Perf. Network
		Total	per rack						
7	RDHX	408	6x60 1x48	2x NVidia Grace	72c @ 2GHz	240 GB LPDDR5	2	745GB NVMe	1x NDR200



MareNostrum 5 – High Speed network IB NDR200





**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación



EXCELENCIA
SEVERO
OCHOA



MareNostrum 5 Users and Support

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 programme, as well as the Participating States Spain, Portugal and Türkiye



Hosting Consortium:



Spain



Portugal



Türkiye



MareNostrum5 – specific features

- Ideal for highly scalable applications, being an exceptionally large General-Purpose Processor (GPP) machine with over 6400 nodes and 717,000 cores.
- Well-suited for Large Language Models and Artificial Intelligence, thanks to its over 4400 NVIDIA H100 GPUs (each with 64GB HBM2e memory).
- Excellent for heterogeneous executions, with three partitions (GPP, HBM, and ACC) utilizing the same batch system, allowing mixed jobs.
- Perfect for large-scale data applications, equipped with the IBM Spectrum Scale file system supporting up to 240 petabytes with 1.2 TB/s write and 1.6TB/s read bandwidth and 400 PB on tapes.

Projects examples – ALIA Large language model

- ALIA, Europe's first public, open and multilingual AI infrastructure (January 2025)
- The project, coordinated by BSC, provides open and transparent language models to promote the use of Spanish and co-official languages in the development and deployment of AI
- ALIA-40B, the most advanced public multilingual foundational model in Europe, trained on the MareNostrum 5 supercomputer, emerges in this context



LA INFRAESTRUCTURA PÚBLICA
DE IA ABIERTA Y MULTILINGÜE



ALIA-40B - 40 billion parameters (40,000 million, equivalent to 40×10^9), which has been trained for more than 8 months on MareNostrum 5 with 6.9 billion tokens (words or fragments of words used in these systems) in 35 European languages.

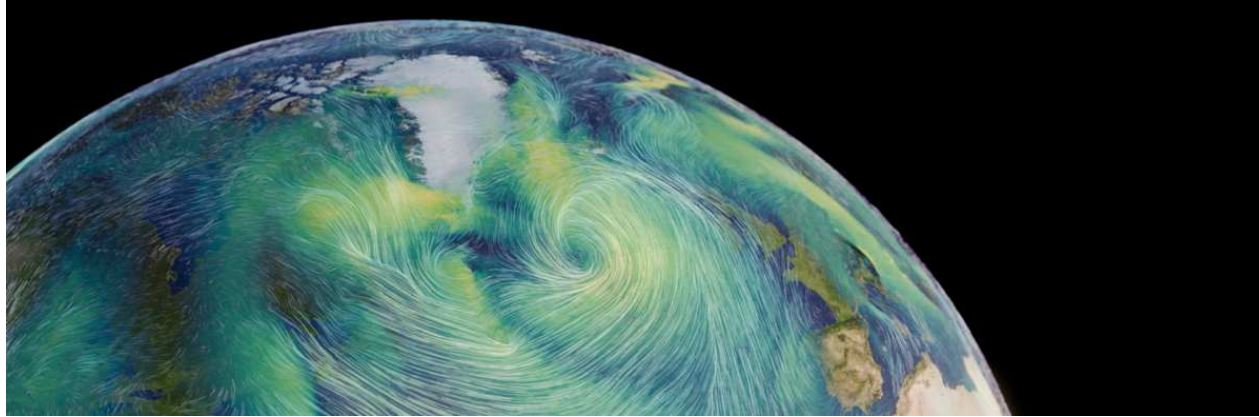
Projects examples – EuroLLM Large language model



- The EuroLLM project includes Unbabel, Instituto Superior Técnico, the University of Edinburgh, Instituto de Telecomunicações, Université Paris-Saclay, Aveni, Sorbonne University, Naver Labs, and the University of Amsterdam.
- Together they created EuroLLM-9B, a multilingual AI model supporting all 24 official EU languages. Developed with support from **Horizon Europe, the European Research Council, and EuroHPC**, this open-source LLM aims to enhance Europe's digital sovereignty and foster AI innovation.
- Trained on the **MareNostrum 5** supercomputer, EuroLLM outperforms similar-sized models. It is fully open source and available via Hugging Face.

For the EuroLLM-1.7B model pre-training, it used **256 Nvidia H100 GPUs** of the Marenostrum 5 supercomputer, training the model with a constant batch size of 3,072 sequences, which corresponds to approximately 12 million tokens, using the Adam optimizer, and BF16 precision.

Projects Examples – Destination Earth's Gordon Bell



- Destination Earth's Climate Digital Twin shortlisted for prestigious ACM Gordon Bell Prize for Climate Modelling
- The submission to the Gordon Bell Prize was prepared by a consortium including the European Centre for Medium-Range Weather Forecasts (ECMWF), Max Planck Institute for Meteorology (MPI-M), CSC–IT Center for Science, the German Climate Computing Center (DKRZ), and BSC, with strategic access to Europe's pre-exascale supercomputers LUMI (Finland) and **MareNostrum 5 ACC and GPP partition (Spain)**.

MareNostrum 5 – Important points

Topic of interest	description
Internet access from login nodes	No outgoing connection from any compute or login node This limitation affects the creation of python envs The solution currently implemented is just ask support@bsc.es to install the env for you (as user or in a generic python module)
Containers options available	Only Singularity containers are available, and without fakeroot, any NGC container from NVIDIA version 23.7 or higher can be installed and run. Previous versions may cause issues with GLIBC.
LLM models	Good scalability with up to 64 nodes (256 GPUs), for larger runs the filesystem can become a limitation, but 128 nodes has still a quite good efficiency but starts degradation. We are studying ways to improve scalability further.
Energy Efficiency and Power Management	We provide EAR monitoring in all the nodes, ACC and GPP



MareNostrum5 – Important points

Topic of interest	description
Data Transfer nodes for large data movements	200GB/s connection shared between 4 data transfer nodes. SSH protocol and Grid FTP available (GridFTP only under petition)
Network GPP and ACC	GPP – 100GB/s per node ACC -- 800GB/s per node (4 devices of 200GB/s, 1 per GPU) <u>contention:</u> island GPP 2160 nodes (241.920 cores), then 3:2 island ACC 160 nodes (640 GPUs), then 2:1
Userportal	Job status on real time and post-mortem Data from GPUs available and power usage per node



UserPortal



Please, have a look at these very useful features:

CPU and Disk accounting with usage alarms: <https://userportal.bsc.es/accounting>

Job status alarms: https://www.bsc.es/user-support/hpc_portal.php#jobnotifications

- bsc099349

MareNostrum 5

Updated on 17/10/2024 09:59:03



EXPORT TO CSV

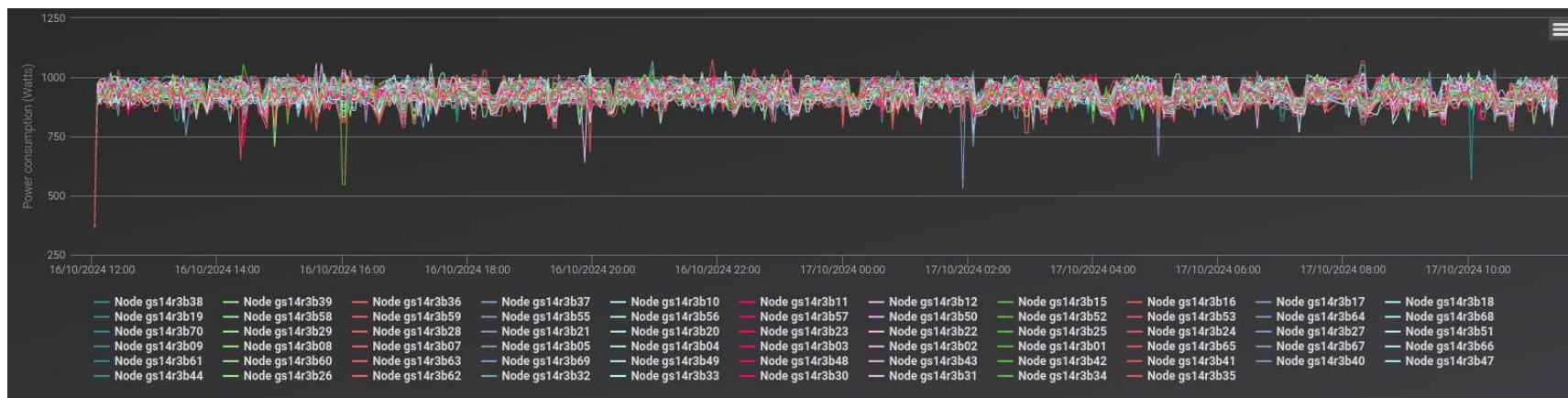
ID	Name	Status	User	Machine	QOS	Submit time	Start	Wallclock	Nodes	Tasks	CPU ①	Memory ①		
9858717	run_script	Completed	bsc099349	MareNostrum 5	acc_bench	09/10/2024 12:48:30	09/10/2024 12:48:55	02-23:59	1	80	N/A	N/A	PREVIEW	VIEW
9858524	run_script	Completed	bsc099349	MareNostrum 5	acc_bench	09/10/2024 12:42:16	09/10/2024 12:42:19	02-23:59	1	20	N/A	N/A	PREVIEW	VIEW
9818239	interactiv...	Completed	bsc099349	MareNostrum 5	gp_bench	08/10/2024 08:59:32	08/10/2024 08:59:35	Unlimited	1	1	N/A	N/A	PREVIEW	VIEW
9818230	interactiv...	Completed	bsc099349	MareNostrum 5	gp_bench	08/10/2024 08:57:40	08/10/2024 08:57:45	Unlimited	1	1	N/A	N/A	PREVIEW	VIEW
9789079	interactiv...	Completed	bsc099349	MareNostrum 5	gp_bench	07/10/2024 14:33:07	07/10/2024 14:33:08	Unlimited	1	1	N/A	N/A	PREVIEW	VIEW
9788839	interactiv...	Completed	bsc099349	MareNostrum 5	gp_bench	07/10/2024 14:26:31	07/10/2024 14:26:37	Unlimited	1	1	N/A	N/A	PREVIEW	VIEW



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

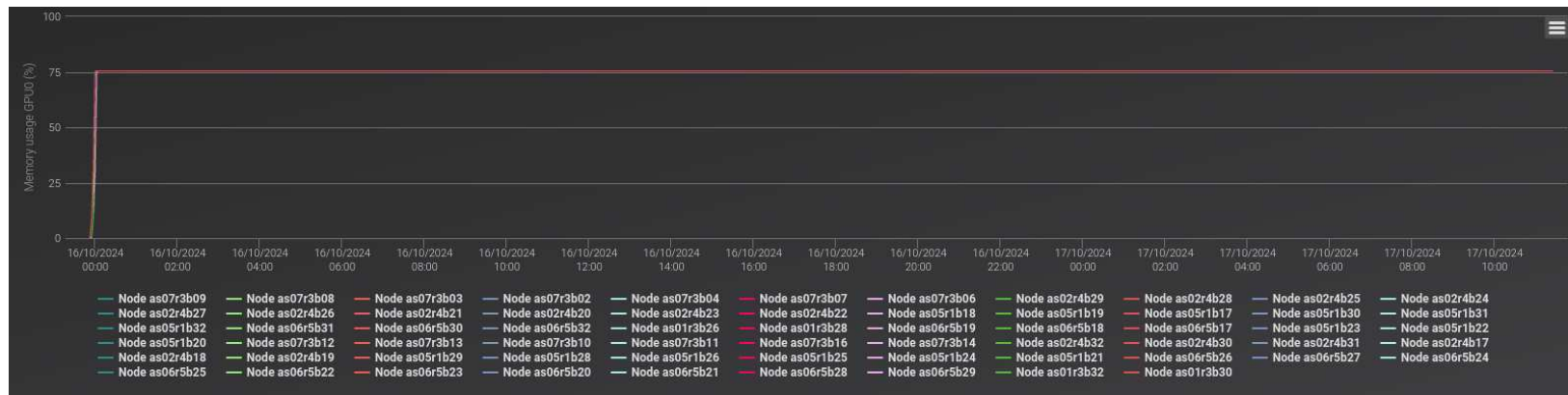
UserPortal node compute usage and power consumption



UserPortal node compute usage and power consumption



UserPortal GPU0 memory and compute usage





MareNostrum5 what is next?

- Quantum partition(s)
 - Quantum Spain and EuroHPC
 - Semiconductors based
- Next Generation Accelerated partition
- AI factories partition
- MareNostrum 6

Thank you

david.vicente@bsc.es