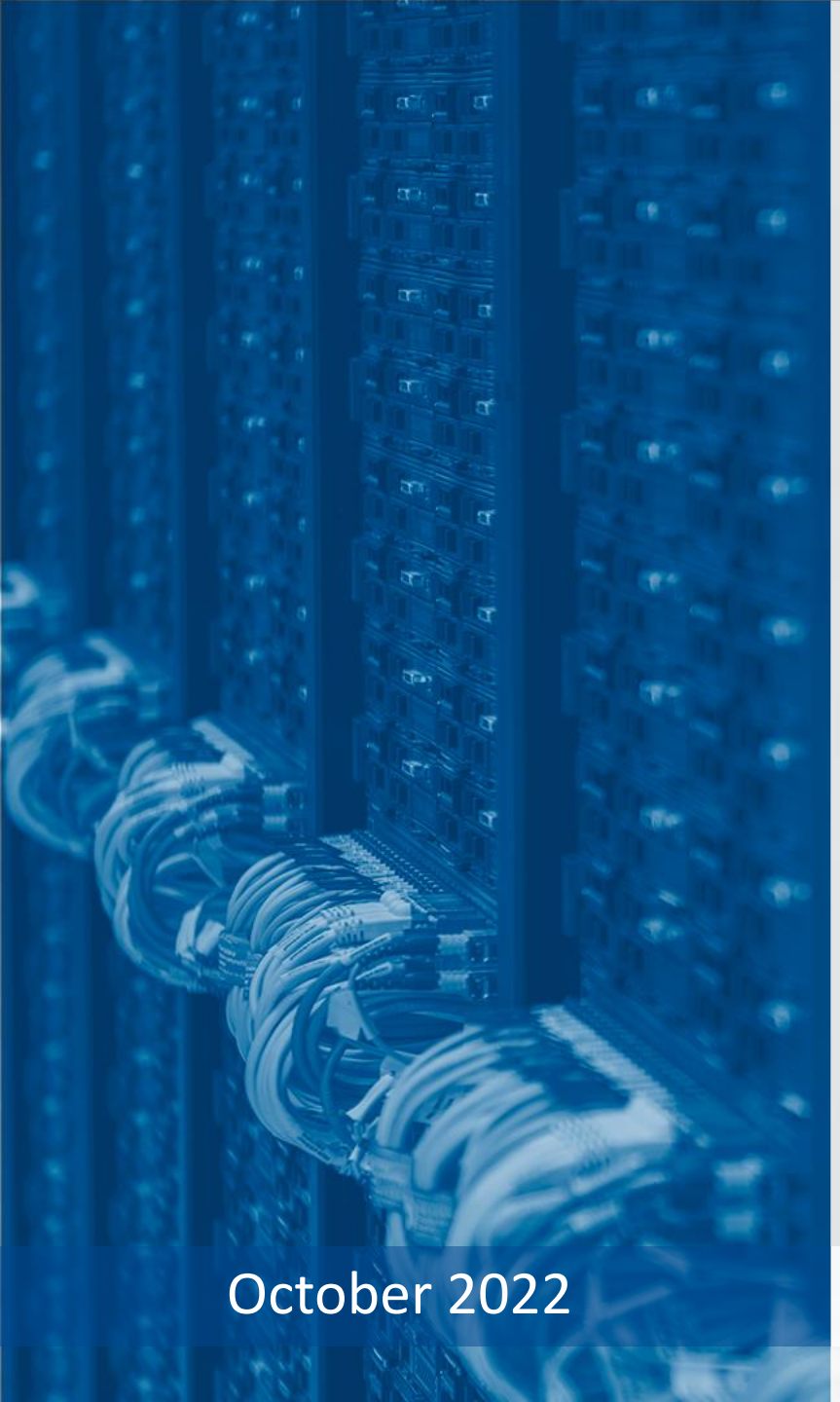

BSC
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

# MareNostrum 5 and its Data Center

Dr. Sergi Girona
Operations Director

Big Science Business Forum 2022

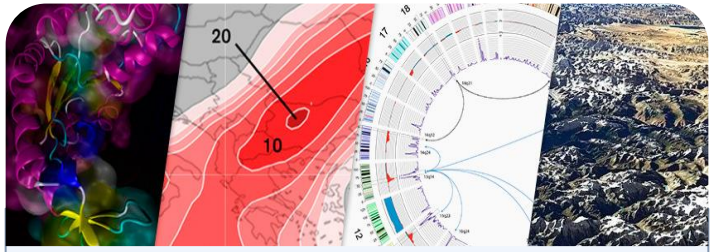
October 2022

BSBF 2022, Granada

# Barcelona Supercomputing Center
# Centro Nacional de Supercomputación

## BSC-CNS objectives

Supercomputing services
to Spanish and EU researchers




R&D in Computer, Life, Earth and
Engineering Sciences

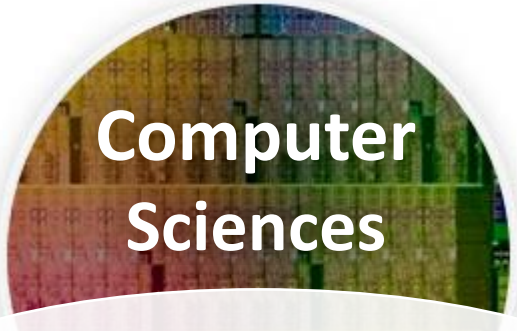
PhD programme, technology
transfer,  public engagement

**BSC-CNS is a consortium that includes**

| | | |
|---|---|---|
| **Spanish Government** | **60%** | Gobierno de España · Ministerio de Ciencia, Innovación y Universidades |
| **Catalan Government** | **30%** | Generalitat de Catalunya · Departament d'Empresa i Coneixement |
| **Univ. Politècnica de Catalunya (UPC)** | **10%** | Universitat Politècnica de Catalunya · BarcelonaTech |

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

# Mission of BSC Scientific Departments

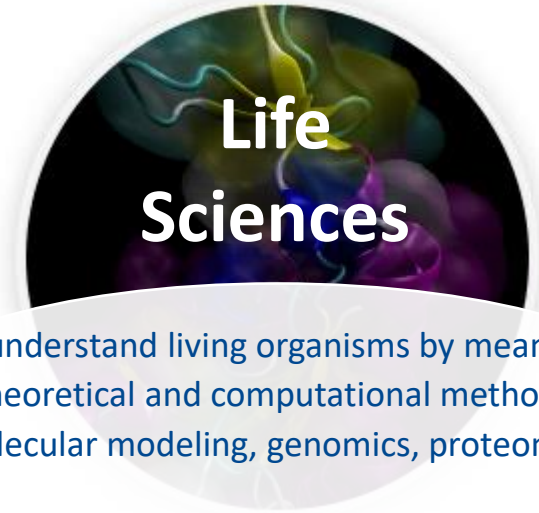


## Computer Sciences

To influence the way machines are built, programmed and used: programming models, performance tools, Big Data, Artificial Intelligence , computer architecture, energy efficiency

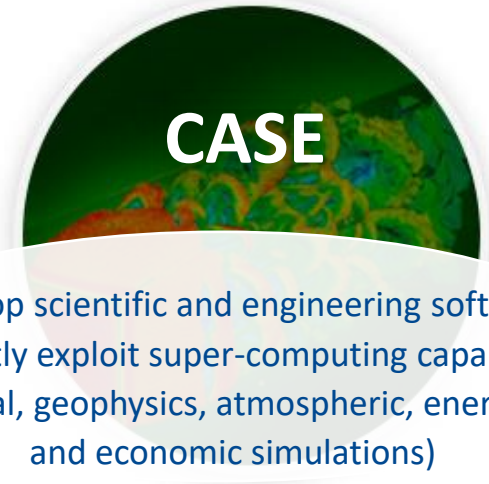
## Earth Sciences

To develop and implement global and regional state-of-the-art models for short-term air quality forecast and long-term climate applications



## Life Sciences

To understand living organisms by means of theoretical and computational methods (molecular modeling, genomics, proteomics)



## CASE

To develop scientific and engineering software to efficiently exploit super-computing capabilities (biomedical, geophysics, atmospheric, energy, social and economic simulations)

# Spanish Supercomputing Network (RES), since 2006

RED ESPAÑOLA DE SUPERCOMPUTACIÓN

Infraestructuras Científicas y Técnicas Singulares


www.res.es

*Membership update: September 2022*

*HPC and data management resources for the scientific community*

- 14 institutions
  - 16 supercomputers
  - 9 data management centres
- +22 PFlop/s combined capacity
- +20 PB storage in 2022 (and growing)
- +800 million CPU hours/year 2022
- +1.000 regular users
- +200 scientific papers annually

- 3 HPC calls per year
- 1 Data call per year
- Applications Support Teams

- Member of Spanish Unique Scientific and Technical Infrastructure network (**ICTS**)
- Access Committee and Users Commitee
- EuroHPC National Competence Centre
- Coordinated by **BSC-CNS**

# MareNostrum 4

| Total peak performance: **13.9** Pflops | | |
|---|---|---|
| General Purpose Cluster: | 11.15 Pflops | (1-07-2017) |
| CTE1-P9+Volta: | 1.57 Pflops | (1-03-2018) |
| CTE2-Arm V8: | 0.65 Pflops | (12-2019) |
| CTE3-AMD: | 0.52 Pflops | (12-2019) |

**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

# EuroHPC: towards European HPC technologies

## EuroHPC-Ju members:

Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Montenegro, the Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland and Turkey

*"A new legal and funding structure – the EuroHPC Joint Undertaking – shall acquire, build and deploy across Europe a world-class High-Performance Computing (HPC) infrastructure.*

*It will also support a research and innovation programme to develop the technologies and machines (hardware) as well as the applications (software) that would run on these supercomputers."*

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EuroHPC Joint Undertaking

*February 2020*

# MareNostrum 5. A European pre-exascale supercomputer

- **200 Petaflops** peak performance (200 x $10^{15}$)*
- **Experimental platform** to create supercomputing technologies “made in Europe”
- **217 M€** of investment

EuroHPC Joint Undertaking

Hosting Consortium:

**Spain** **Portugal** **Turkey** **Croatia**

European Commission

** At the time of call for HE, peak performance expected of 200 Petaflops*
** At the time of tender publications, minimum aggregated sustained HPL of 205 Petaflops*

Generalitat de Catalunya
Departament d'Empresa i Coneixement

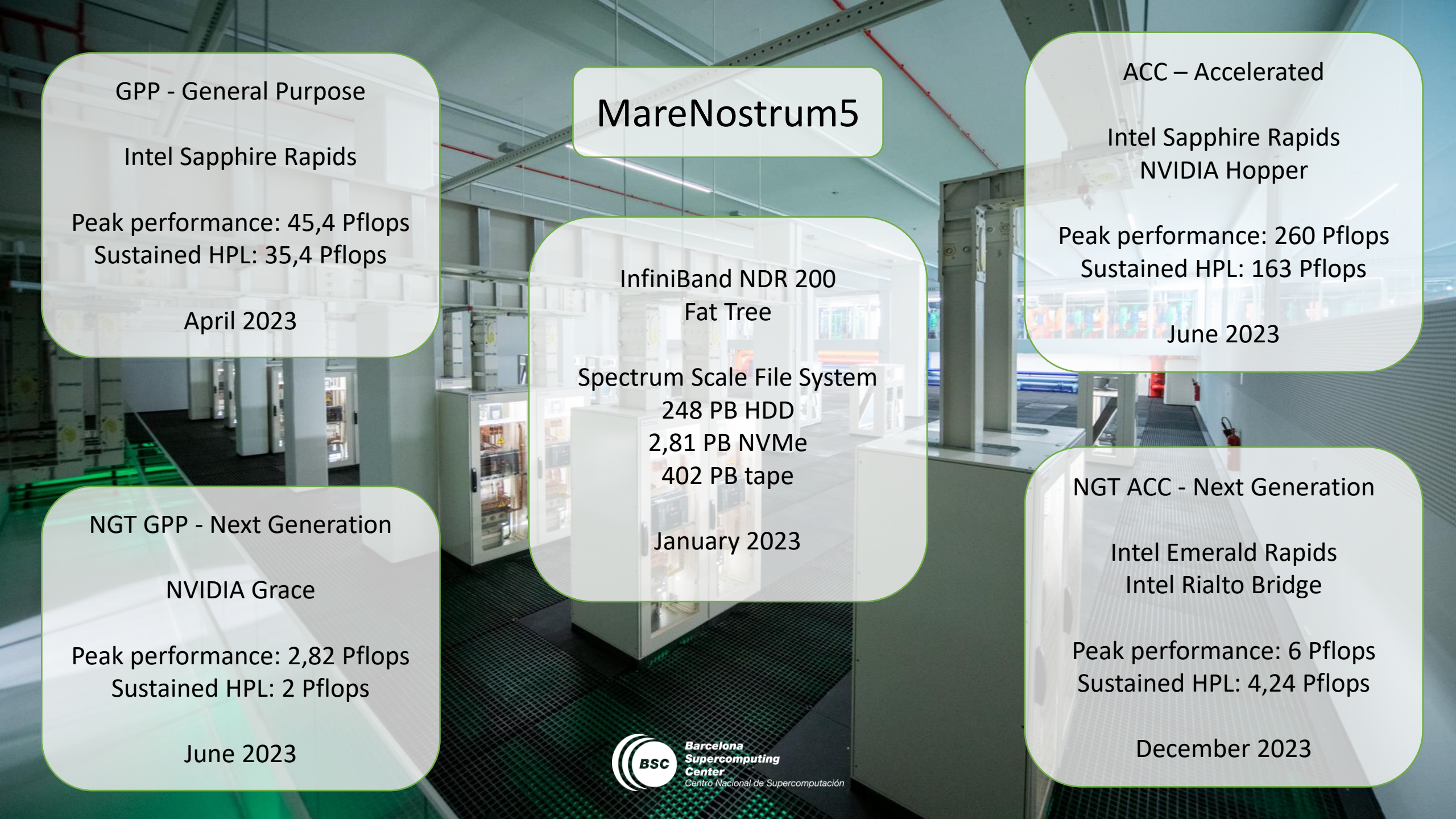
Gobierno de España
Ministerio de Ciencia, Innovación y Universidades

UPC Universitat Politècnica de Catalunya BarcelonaTech

BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

*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, Croatia, and Turkey*

EuroHPC Joint Undertaking



# MareNostrum5

GPP - General Purpose

Intel Sapphire Rapids

Peak performance: 45,4 Pflops
Sustained HPL: 35,4 Pflops

April 2023

ACC – Accelerated

Intel Sapphire Rapids
NVIDIA Hopper

Peak performance: 260 Pflops
Sustained HPL: 163 Pflops

June 2023

InfiniBand NDR 200
Fat Tree

Spectrum Scale File System
248 PB HDD
2,81 PB NVMe
402 PB tape

January 2023

NGT GPP - Next Generation

NVIDIA Grace

Peak performance: 2,82 Pflops
Sustained HPL: 2 Pflops

June 2023

NGT ACC - Next Generation

Intel Emerald Rapids
Intel Rialto Bridge

Peak performance: 6 Pflops
Sustained HPL: 4,24 Pflops

December 2023

BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

# Compute partitions overview

| | | Cooling | Nodes Total | Technology | Processor/Accelerator | | Memory | PFlops (HPL) | | Local Drive | High-Perf. Network |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | General Purpose | DLC +RDHX | >6000 | Lenovo | 2x Intel Sapphire R. | | >2GB/core 256GB DDR5 | 35.43 | >205 | 960GB NVMe | 1x NDR200 Shared by 2 nodes |
| | | | >200 | | | | >8GB/core 1024GB DDR5 | | | | |
| | | | >50 | | 2x Intel Sapphire R. HBM | | > 0.5GB HBM/core 128GB HBM + 32GB DDR5 | 0.34 | | | |
| | Accelerated | DLC | > 1000 | Atos | 2x Intel Sapphire R. | | 512GB | 163 | | 480GB NVMe | 4x NDR200 |
| | | | | | 4x Nvidia Hopper 64GB HBM | | | | | | |
| Next Gen | General Purpose | AC +RDHX | > 400 | Atos | Nvidia Grace | 144c @ > 2.4GHz | 240GB LPDDR5 | 2 | | 128GB NVMe | 1x NDR200 |
| | Accelerated | DLC +RDHX | | Lenovo | 2x Intel Emerald R. | | 512GB DDR5 | 4.24 | | 960GB NVMe | 2x NDR |
| | | | | | 4x Intel Rialto Bridge ≧128GB HBM | | | | | | |

# Compute racks Infrastructure

| | General Purpose | Accelerated | Next Gen General Purpose | Next Gen Accelerated |
|---|---|---|---|---|
| Size of the rack ( H x W x D ) (in cm) | 201 x 60 x 160 | 225 x 90 x 135 (no doors) | ~ 202 x 60 x 127 | 201 x 60 x 160 |
| Weight ( kg / m2) | 1637 kg/m² | < 2415 kg/m2 | ~ 1000 kg/m2 | 675 kg/m2 |
| Compute nodes per rack | 72 | 32 | 68 | 24 |
| KW per rack (average) HPL | > 65 kW | > 110 kW | > 50 kW | > 70 kW |
| Cooling mechanism | Direct-To-Node Warm Water Cooling and Rear Door Heat Exchangers | Direct Warm Water Cooling | Rear Door Heat Exchangers | Direct-To-Node Warm Water Cooling and Rear Door Heat Exchangers |
| Residual heat to ambient (kW) | Room neutral | 3.86kW | Room neutral | None |

# MN5 Tender some requirements Infrastructure

- Not exceed: 12 MW (under HPL) and PUE of 1,08
- Per rack requirements
  - Power, weight, recommended dimensions, dissipation, …
  - Remove doors
  - Cabling, colours, visibility …
- Each rack must dissipate minimum of 95% of heat generated
- Cold-water 18ºC up to 1MW
- Warm-water 35ºC  up to 12 MW
- MN5 site as Exhibition center (L2-I10 very high)

# MN5 Site preparation

- Public tender: CONOBR02019010OP
  - Awarded on 01/08/2019
  - Awarded Prize: 12.557.990 € (excluding VAT)
    - Including: project, construction and maintenance
  - Awardee: Climava SL
  - Formalisation on 26/11/2019

- Climava SL

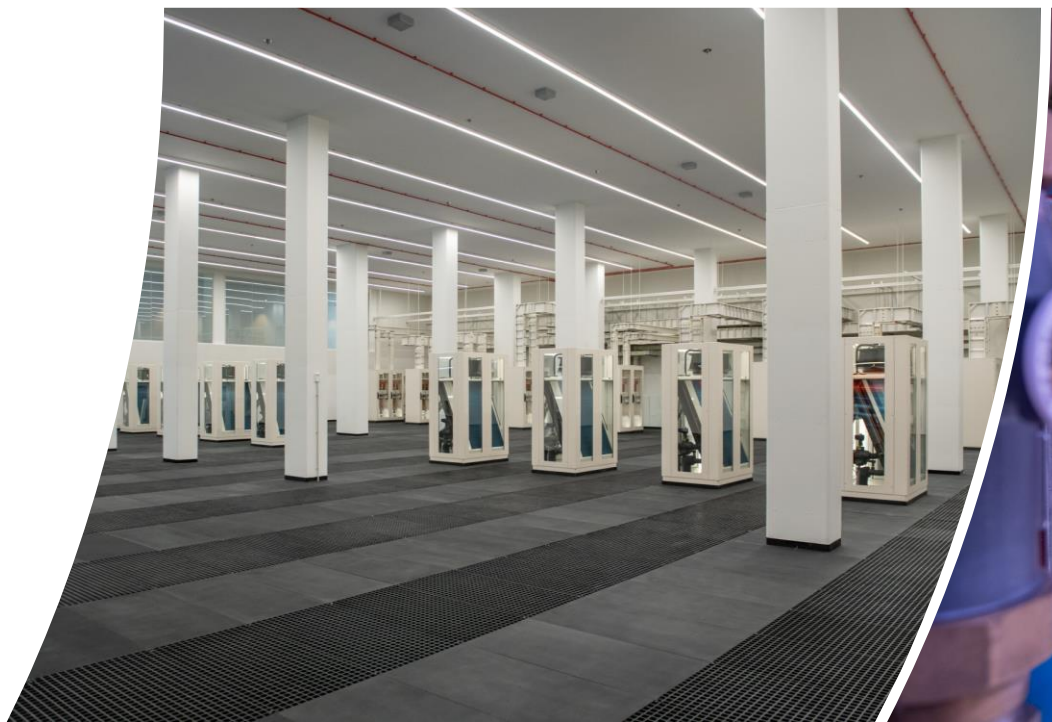
  - Gisela Valderrama, Jaume Villa
  - https://www.climava.com

  - Global Technia Consulting
    - Lluis Gironella
    - https://www.b-global.tech

| | |
|---|---|
| Expected date before covid19: | September 2020 |
| Acceptance date: | April 2022 |

# Compute room

- Space: 900 sqm
- >6 meters height
- 140 cm false floor
  - Electricity
  - Water
- 2500 kg/sqm
- Electrical switch boards
- Fire detection & extinction
- 3 circuits water distribution
- 9 Crahs



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

# Transformers Low voltage switchboard room

- 5 x  TRANSFORMADOR 4150KVA VACUUM CAST FILLED DRY
  - 4150 kVA
  - Primary: 25 kV, Secondary: 420 V
  - Frequency: 50Hz
  - 3 phases
- Emergency line: 5 MW
- Foreseen expansion to 40 MW
- Main switchboard with max. intensity of 6300 A
- UPS 1MW, 2N. Lithium batteries, 10 minutes durations
- Power distribution with BlindosBarra, double path
- Estimated PUE on nominal workload below 1,08
- 100% Green energy

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

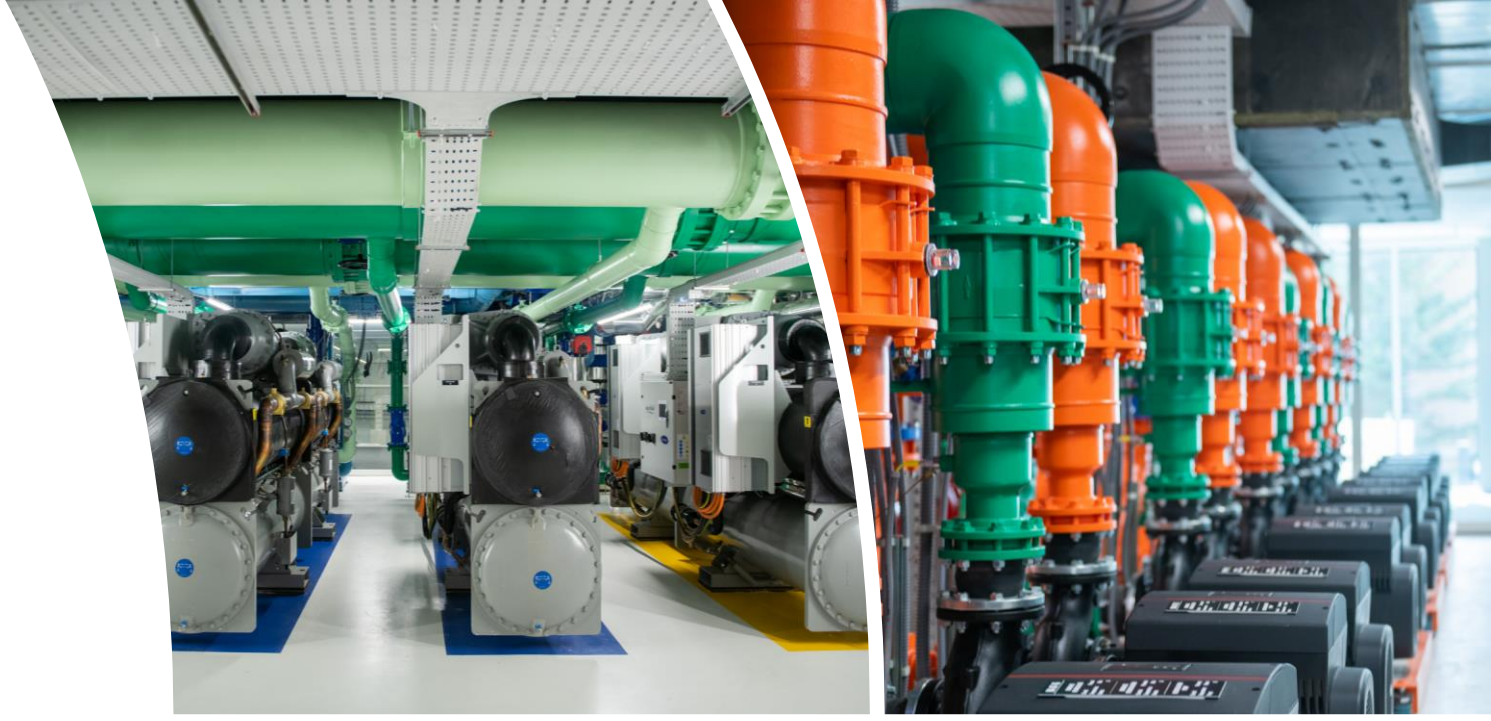
# Fire Detection and Extinction

- VESDA system
- 4 programmable levels.
- 9 independent areas for fire detection and extinction
- Fire protection using water mist system which uses very fine water sprays (i.e. water mist). The small water droplets allow the water mist to control, suppress or extinguish fires by:
  - cooling both the flame and surrounding gases by evaporation
  - displacing oxygen by evaporation
  - attenuating radiant heat by the small droplets themselves
- Integrated in BMS and building fire system

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

# Heat exchanger, chiller and pumps room

- 6 (4+2) Heat exchanger T25-PFM
  - Water flow: 1170 $m^3/h$
  - Temperatures
    - To tower: outlet: 28,1ºC , Inlet: 38,1ºC
    - To rack: outlet: 30ºC , Inlet: 40ºC
  - Total dissipation power: 13500 kW
- 5 (2 MT + 1 LT+ 2) Chillers
  - Water flow: 302 $m^3/h$ + 151 $m^3/h$
  - Temperatures, separate loops
    - 16ºC – 26ºC
    - 8ºC – 14ºC
    - To rack outlet: 30ºC , Inlet: 40ºC
- Redundancy:  N+2 in chillers and heat exchangers
- Heat reuse



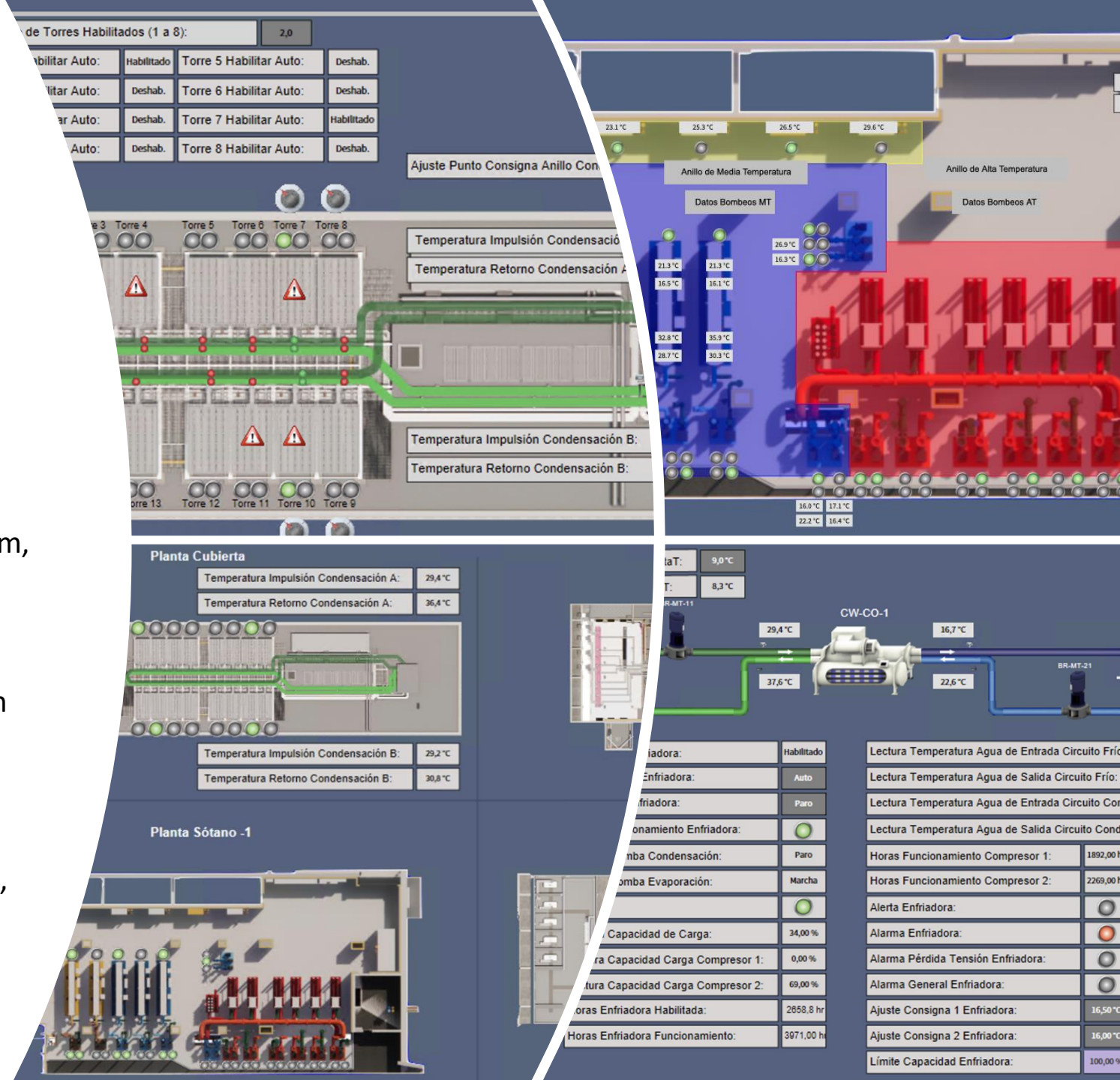
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

# Cooling towers

- 14+2 Torraval CTFP-2436(SB)
- Water flow: 1500 m3/h
- Outlet: 28,1 C
- Inlet: 38,1 C
- Wet bulb temperature: 25C
- Total dissipation power: 17300 kW
- Water source
  - Underground/phreatic water
  - Industrial water

# BMS: Building Monitoring System

- Redundant Ethernet/TCP communications ring, with redundant Master Controllers.
- Fully bistable system, in case of loss of communications or failure of the management system, the infrastructure remains operational without any alteration.
- Option of operation in manual mode remotely controlled by an operator or 100% local manual from the plant itself.
- Management of alarms and warnings via SNMP (bidirectional).
- Storage of historical events, alarms and logs in event, alarm and log databases in SQL databases




Barcelona Supercomputing Center
Centro Nacional de Supercomputación

# Next projects?

- On-going
  - System installation
  - System and facility validation
  - Water quality, control and treatment
- On construction or procurement
  - Osmosis Facility
  - Installation of quantum systems
- At legal/economical validation
  - Utilization of phreatic water
  - Power station
- On background preparation
  - MareNostrum VI

Big Science Business Forum 2022

# Thank you

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

sergi.girona@bsc.es