Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

HPC evolution for weather and climate models

Mario C. Acosta

27/09/2022

EWGLAM 2022, Brussels

Who we are

- The refactoring of Earth System Models
  - Computational performance analysis and new optimizations for our numerical models.
  - Studying new algorithms for the new generation of high performance platforms (path to exascale).
- We are collaborating with several institutions on different projects working together in the same direction

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

immerse
IMPROVING OCEAN MODELS FOR THE COPERNICUS PROGRAMME

is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

High Resolution Hirlam Limited Area Model

MERCATOR OCEAN INTERNATIONAL

ETP 4 HPC EsDs

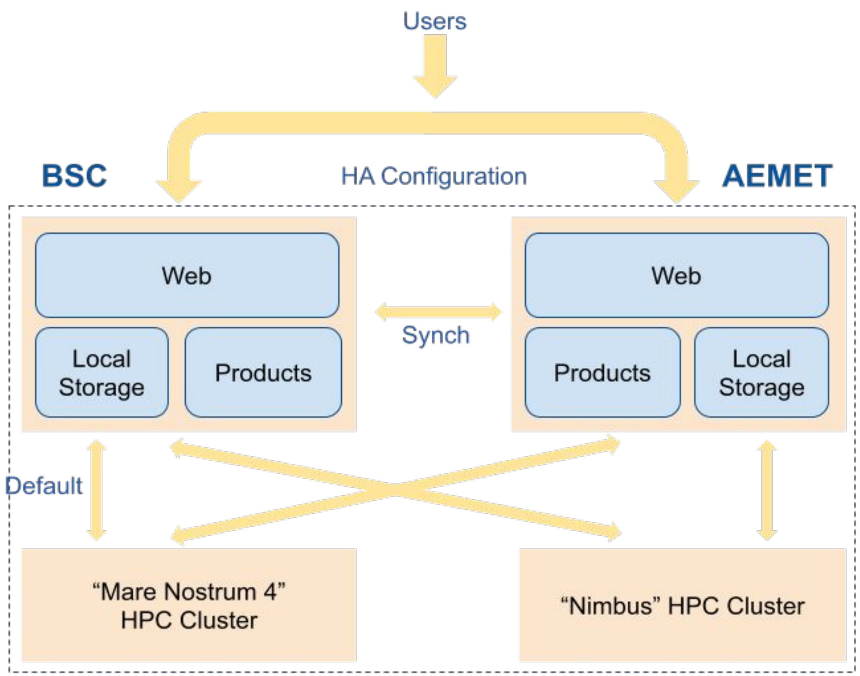
ESCAPE 2

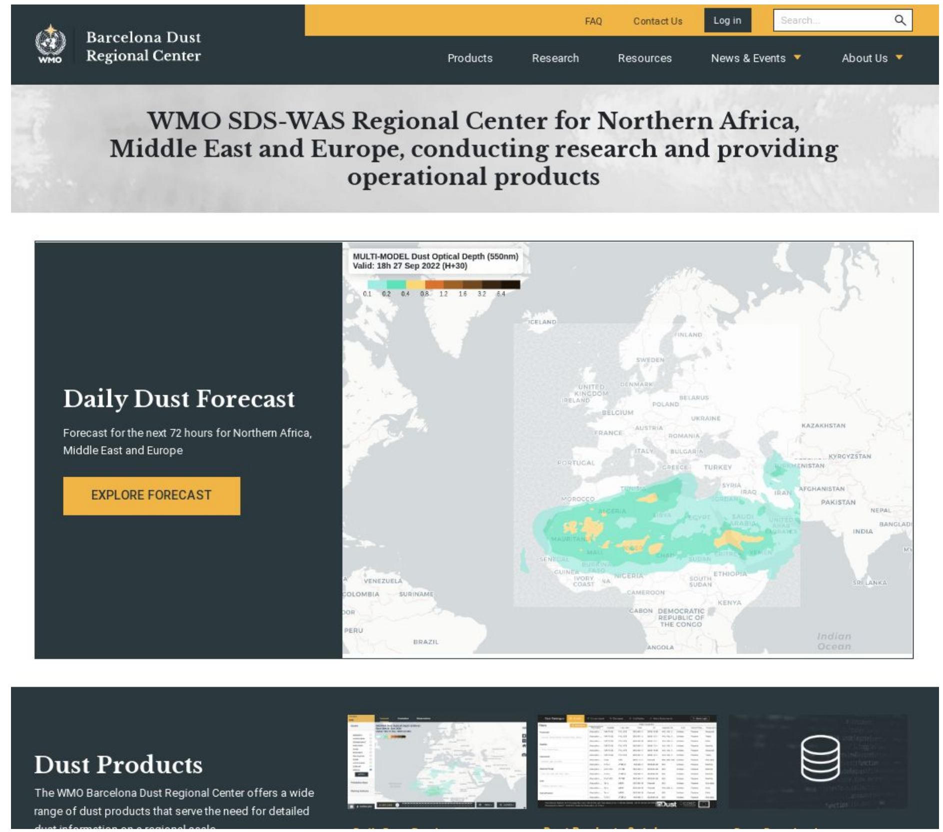
OEMES RETOS-COLABORACIÓN

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

PRACE

Barcelona Dust Regional Center





Consortium operated by **BSC** and **AEMET** under the **WMO** umbrella

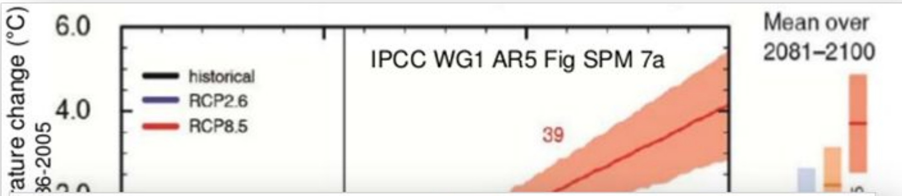
Operational 72 hours (3-hourly) forecast of 14 external NWP models + 1 in house

Interactive visualization dashboard, multi-model products, skill scores, numerical data download

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

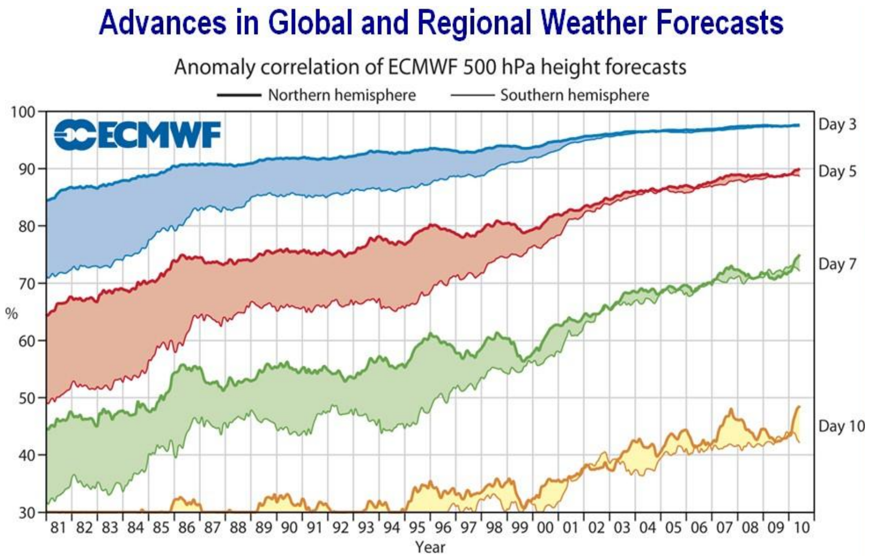
Introduction: Climate overview

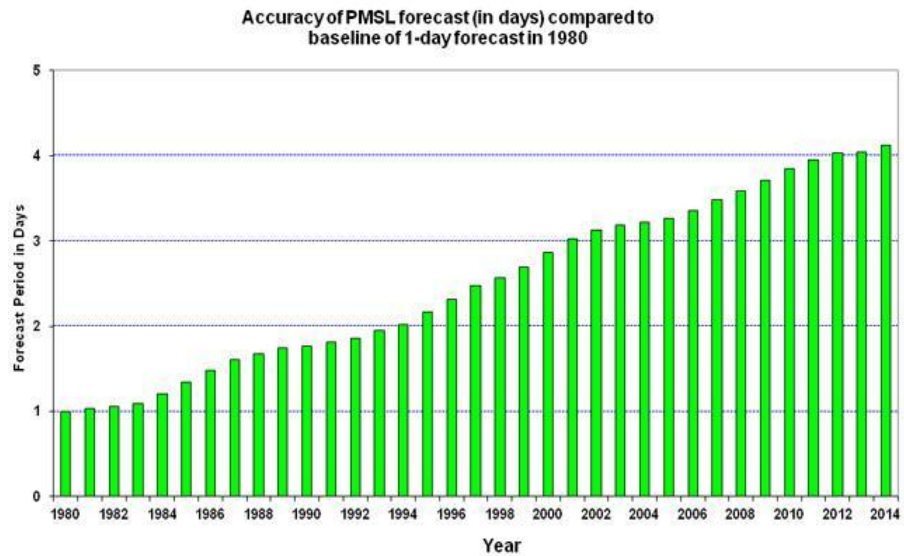
- Projections
- Impact analysis
- Adaptation to climate change.

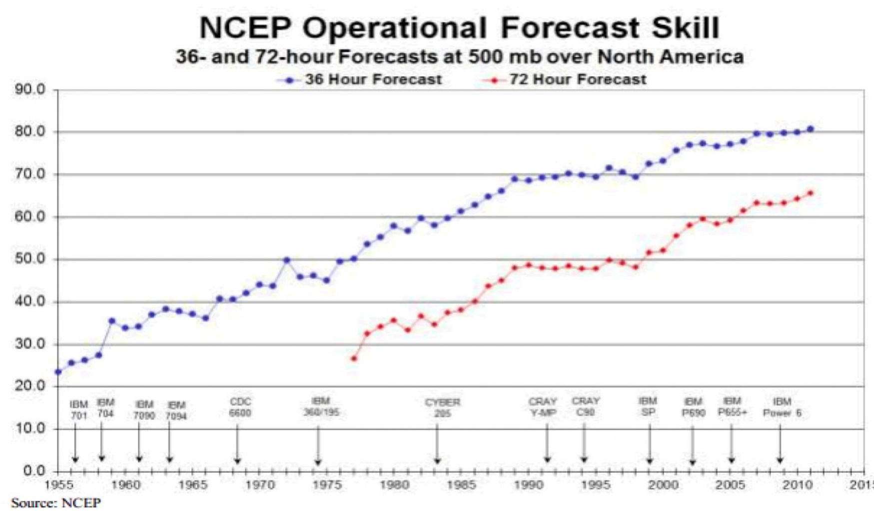


Currently, **only computational models have the potential** to provide geographically and physically consistent estimates.

Introduction: Weather overview

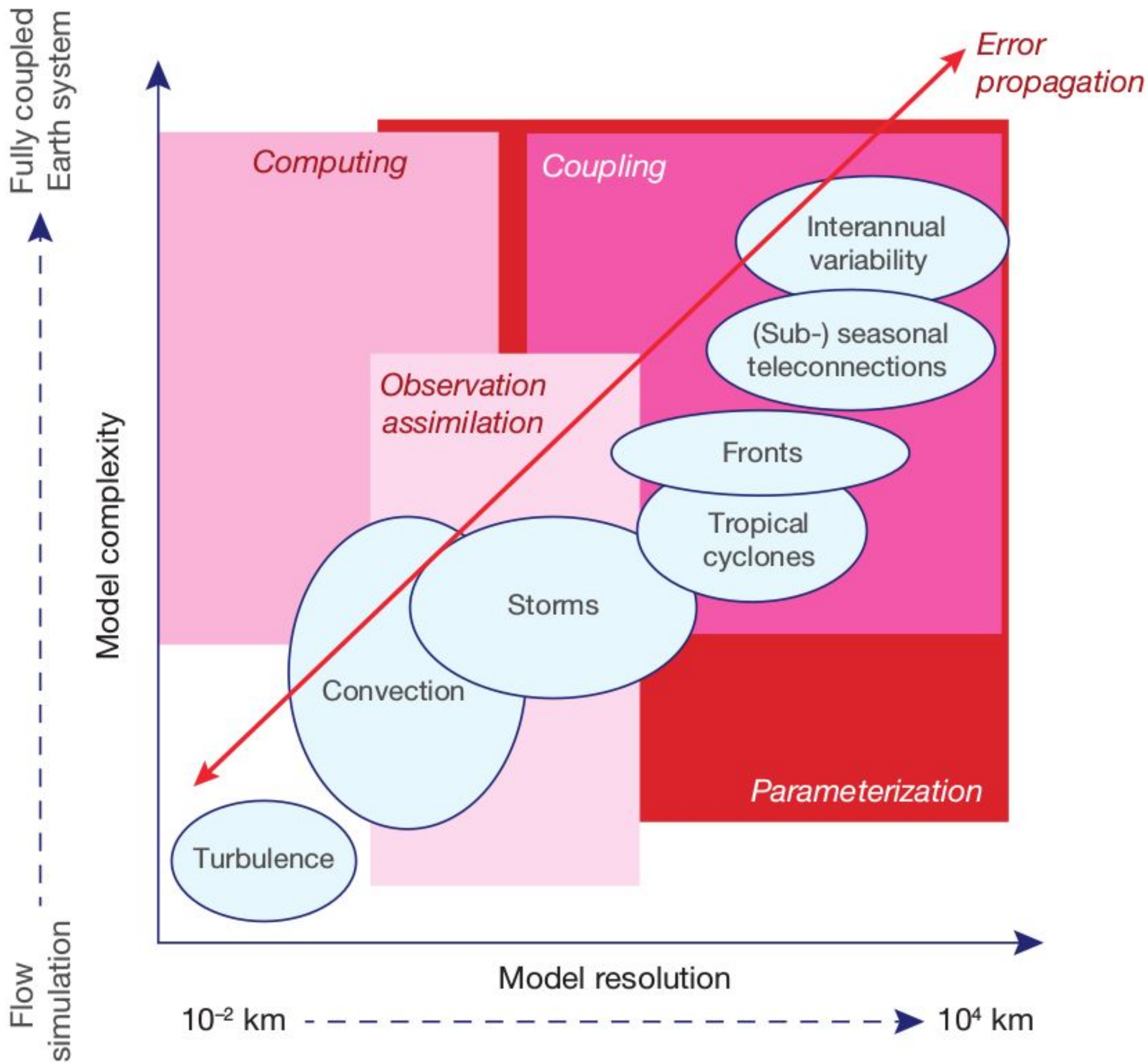






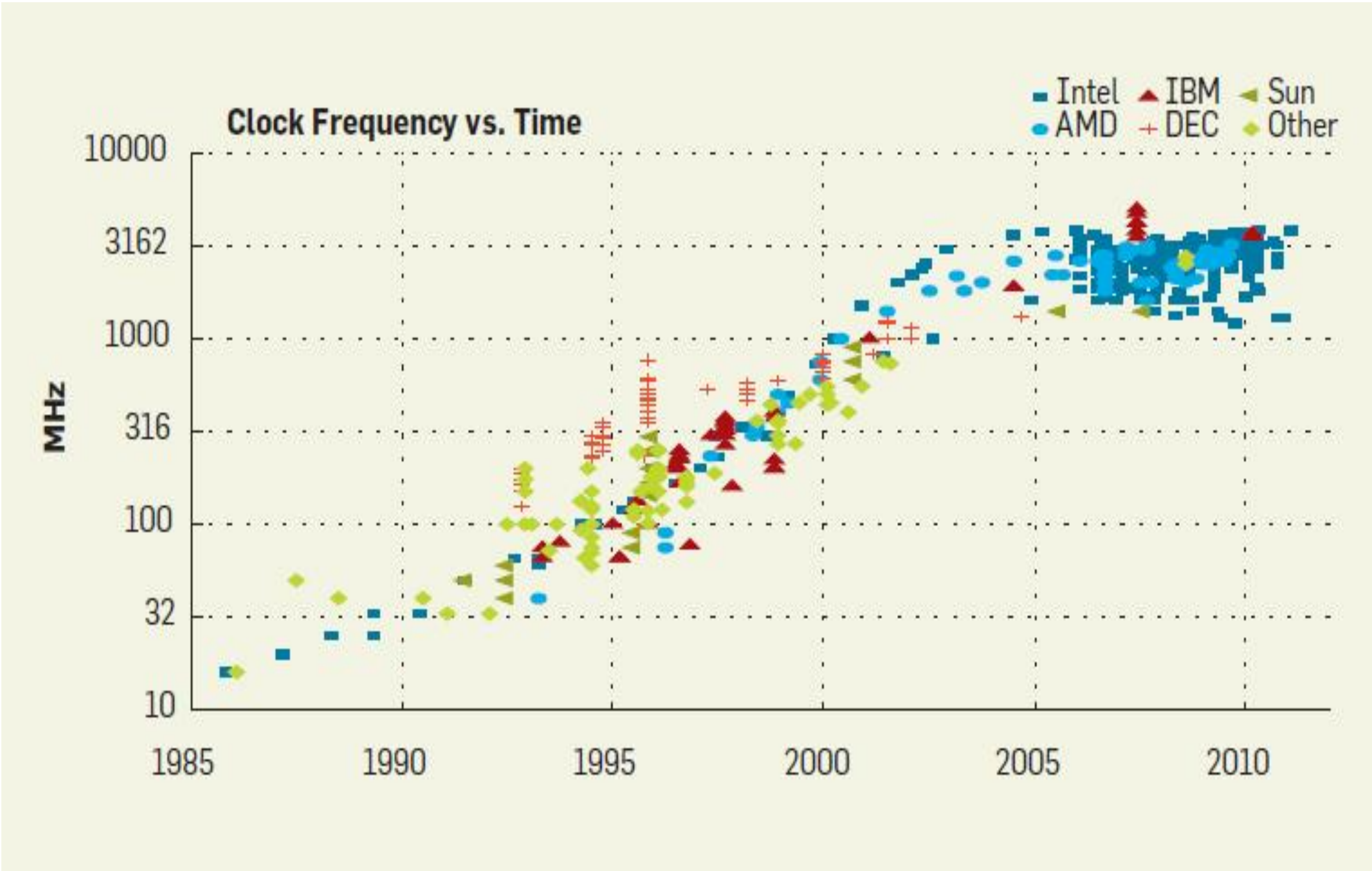
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Introduction



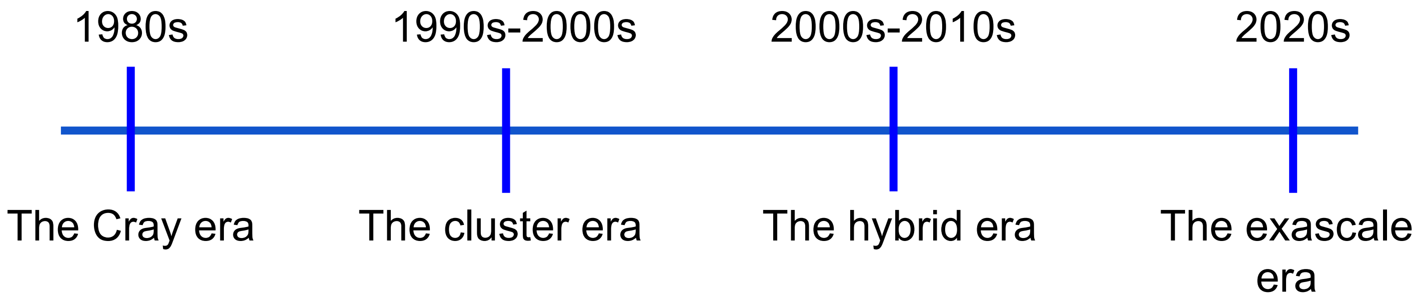
Introduction

Introduction



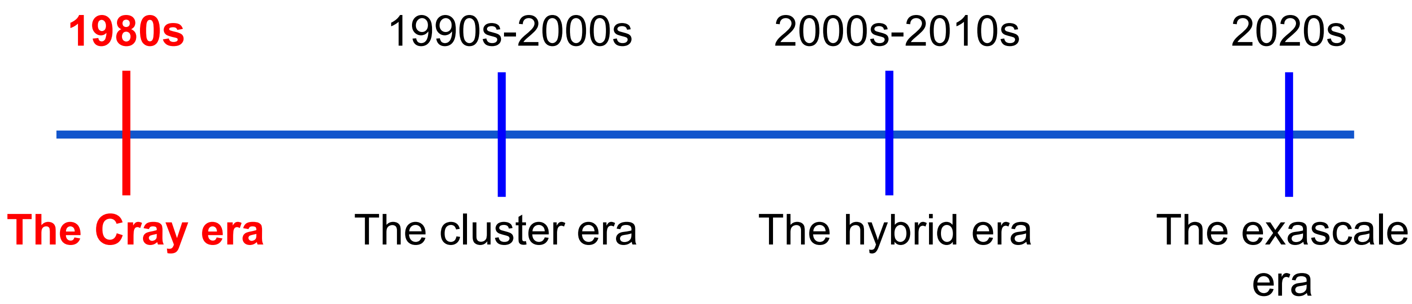
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

HPC history



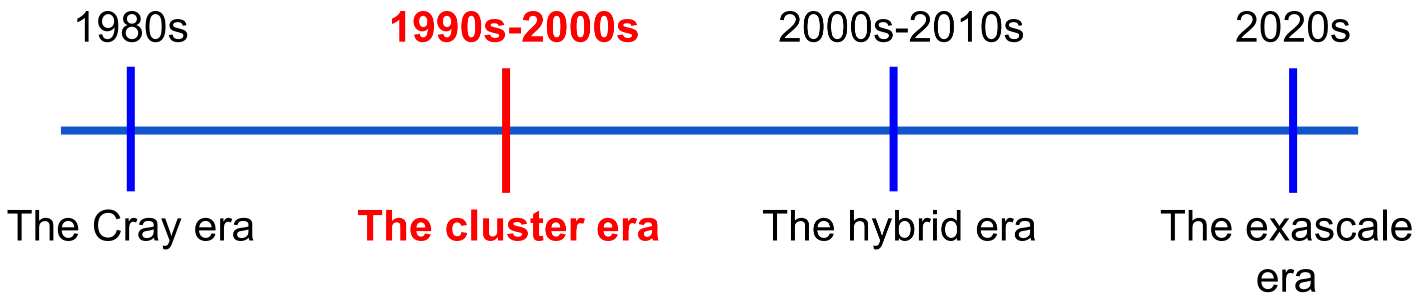
- High performance computing has always pushed the limits and designs of computing technology.
- HPC architectures have gone through rapid changes to facilitate the increasing computational demand of scientific applications in many areas such as Earth Science.

HPC history



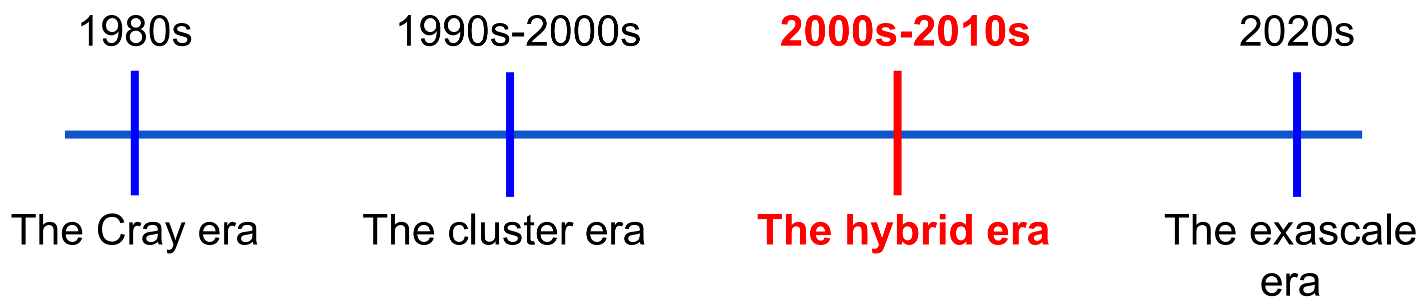
- Supercomputers composed of vector processors such as CRAY-1 and -2 dominated computing fields
- These computations focused on floating point operations per second (FLOPS)
- The supercomputers had central processing units (CPU) that implemented an instruction set designed to operate efficiently and effectively on large one-dimensional arrays of data called vectors.

HPC history



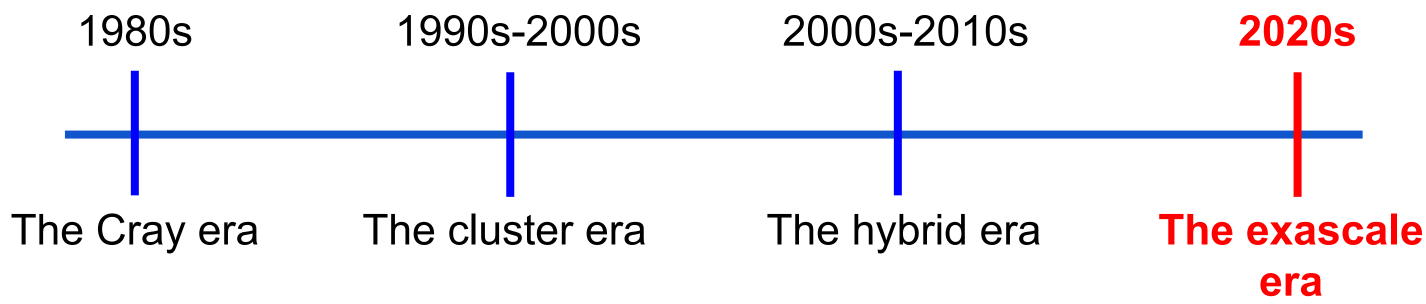
- Multi- and many-core CPUs approach.
- Break the technological limitation of a higher operating clock frequency and a higher density of large integrated circuits to align with Moore's Law.
- Solve other issues as the gap between processor and memory speed or the instruction level parallelism.
- This led to clusters with a large number of processors, each with a small number of core sharing RAM and some cache space.

HPC history



- As the numbers of cores increased, Graphics Processing Units were integrated into HPC clusters to accelerate the performance for many applications.
- These systems contain large numbers (hundreds to thousands) of small efficient cores (many-cores) that worked in unison.
- This approach essentially “offloaded” certain types of operations from the CPU to the GPU.

HPC history



- Exascale machines are more purpose-built and rated using a specific application performance rather than the general Top500 High Performance Linpack benchmark.
- Extending multi-core/many-core clusters to the Exascale range is hampered by the disconnect between hardware and software.
- Heterogeneous computing and co-desing as solution
  - Fixed Accelerators provide order(s) of magnitude more specialized performance
  - The main problem could be complexity and programmability

The pre- and exascale path

- What we can gain
  - New computing elements (GPUs, FPGAs, AI, Quantum, RISCV)
  - More parallelism (Million threads)
  - Much more computing (Exaflops)
  - Data streaming
  - More data (Exabytes)
  - More complexity in our models (increased resolution, more parameters or components, more ensembles)
  - Larger Datasets

The pre- and exascale path

- What we can gain
  - New computing elements (GPUs, FPGAs, AI, Quantum, RISCV)
  - More paralleism (Million threads)
  - Much more computing (Exaflops)
  - Data streaming
  - More data (Exabytes)
  - More complexity in our models (increased resolution, more parameters or components, more ensembles)
  - Larger Datasets

Modeling and simulation + **Machine Learning** + **Big Data Analytics**

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

The pre- and exascale path

- What we can gain
  - New computing elements (GPUs, FPGAs, AI, Quantum, RISCV)
  - More paralleism (Million threads)
  - Much more computing (Exaflops)
  - Data streaming
  - More data (Exabytes)
  - More complexity in our models (increased resolution, more parameters or components, more ensembles)
  - Larger Datasets

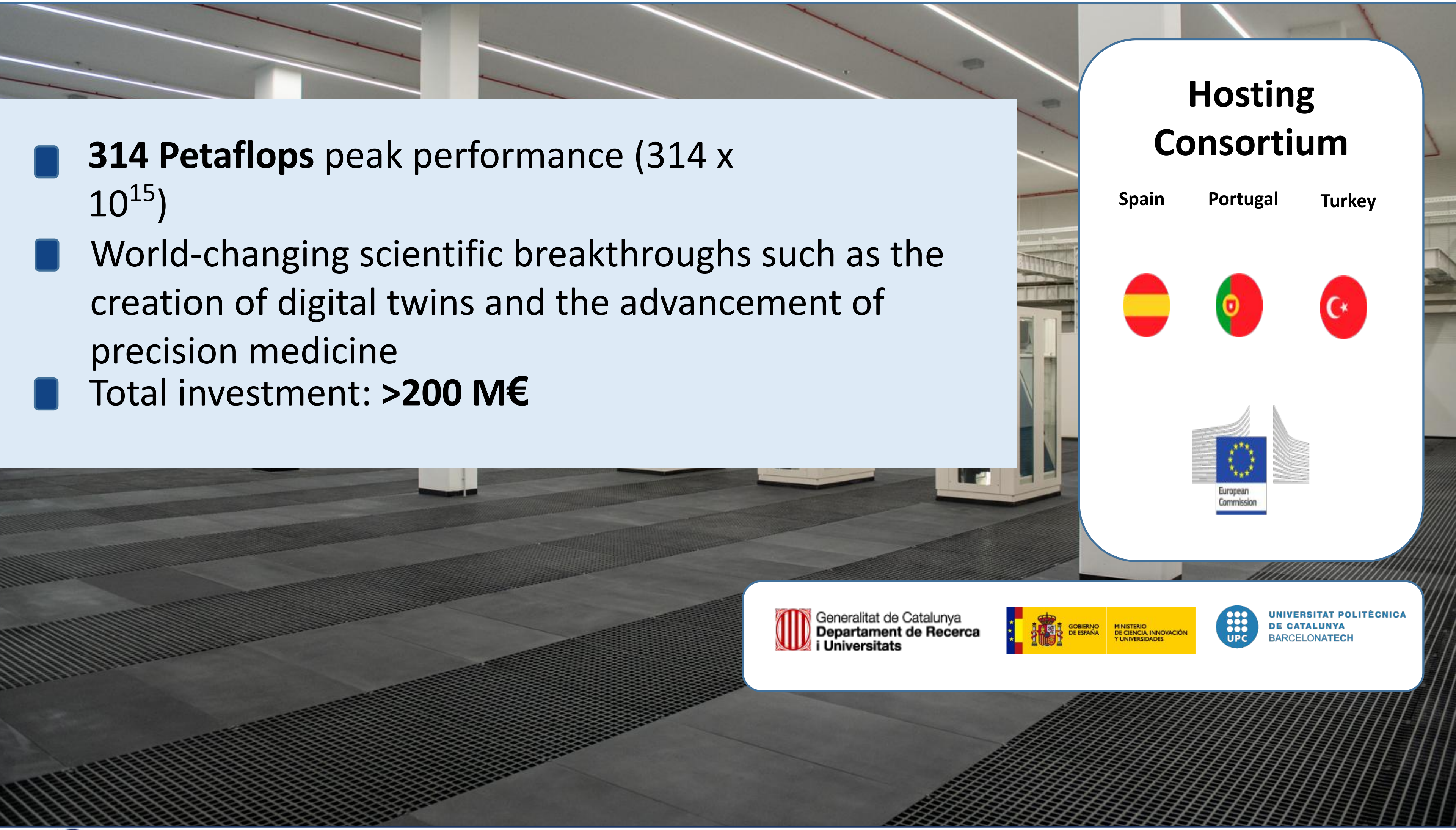
Modeling and simulation + **Machine Learning** + **Big Data Analytics**

The pre- and exascale path

- In short time
  - New pre-exascale machines (LUMI, LEONARDO, Marenostrum5).
  - High-resolution “Digital Twins” using EuroHPC hardware
  - European Projects pursuing HPC improvements (ESiWACE3)
- In medium/long time
  - RISCV, Quantum, FPGAs
  - Clouding
  - Hardware and software acceleration (IA, GPUs)
  - …

Modeling and simulation

MareNostrum 5. A European pre-exascale supercomputer



- **314 Petaflops** peak performance (314 x 10^{15})
- World-changing scientific breakthroughs such as the creation of digital twins and the advancement of precision medicine
- Total investment: **>200 M€**

Hosting Consortium

Spain Portugal Turkey

European Commission

Generalitat de Catalunya
Departament de Recerca i Universitats

Gobierno de España
Ministerio de Ciencia, Innovación y Universidades

Universitat Politècnica de Catalunya
BarcelonaTech

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, and Turkey

EuroHPC Joint Undertaking

MareNostrum 5

GPP - General Purpose

Intel Sapphire Rapids

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

April 2023

NGT GPP - Next Generation

NVIDIA Grace

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

June 2023

InfiniBand NDR 200
Fat Tree

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

January 2023

ACC – Accelerated

Intel Sapphire Rapids
NVIDIA Hopper

Peak performance: 260 Pflops
Sustained HPL: 163 Pflops

June 2023

NGT ACC - Next Generation

Intel Emerald Rapids
Intel Rialto Bridge

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

December 2023

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
	General Purpose	DLC +RDHX	>200	Lenovo	2x Intel Sapphire R.		>8GB/core 1024GB DDR5	35.43	>205	960GB NVMe	1x NDR200 Shared by 2 nodes
	General Purpose	DLC +RDHX	>50	Lenovo	2x Intel Sapphire R. HBM		> 0.5GB HBM/core 128GB HBM + 32GB DDR5	0.34	>205	960GB NVMe	1x NDR200 Shared by 2 nodes
	Accelerated	DLC	> 1000	Atos	2x Intel Sapphire R.		512GB	163	>205	480GB NVMe	4x NDR200
	Accelerated	DLC	> 1000	Atos	4x Nvidia Hopper 64GB HBM		512GB	163	>205	480GB NVMe	4x NDR200
Next Gen	General Purpose	AC +RDHX	> 400	Atos	Nvidia Grace	144c @ > 2.4GHz	240GB LPDDR5	2	>205	128GB NVMe	1x NDR200
Next Gen	Accelerated	DLC +RDHX		Lenovo	2x Intel Emerald R.		512GB DDR5	4.24	>205	960GB NVMe	2x NDR
Next Gen	Accelerated	DLC +RDHX		Lenovo	4x Intel Rialto Bridge ≧128GB HBM			4.24	>205	960GB NVMe	2x NDR

Modeling and simulation

Digital Twins

- High-resolution “Digital Twins” using EuroHPC hardware
  - Provide tools, computational analysis and optimizations during the development of the new weather and climate models.
  - Ensure that the novel EuroHPC machines are used efficiently to reach the desired throughput.
  - The creation of approaches and tests which should provide measurable indicators about the quality of the solution: reproducibility, computational efficiency and code quality

  - Accelerate the new models through kernels and data structures suitable for the new GPU partitions.
    - Computationally intensive parts will be optimised for running on GPUs
    - Optimize from the computational point the new kernels to take into account the particularities of the new EuroHPC machines.

Modeling and simulation

Digital Twins

- High-resolution “Digital Twins” using EuroHPC hardware
  - Provide tools, computational analysis and optimizations during the development of the new weather and climate models.
  - Ensure that the novel EuroHPC machines are used efficiently to reach the desired throughput.
  - The creation of approaches and tests which should provide measurable indicators about the quality of the solution: reproducibility, computational efficiency and code quality

  - Accelerate the new models through kernels and data structures suitable for the new GPU partitions.
    - Computationally intensive parts will be optimised for running on GPUs
    - Optimize from the computational point the new kernels to take into account the particularities of the new EuroHPC machines.

Modeling and simulation

Digital Twins

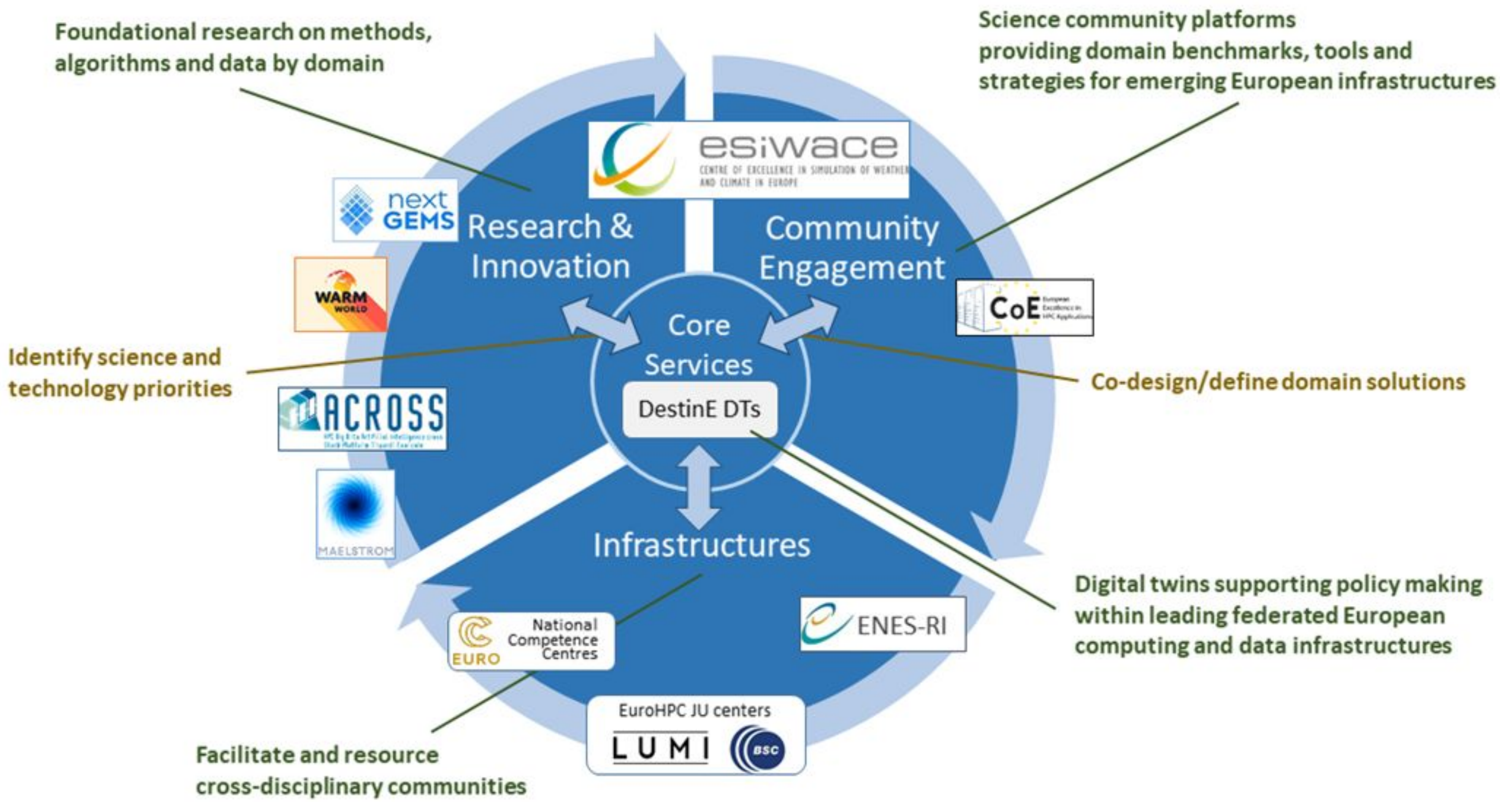
- High-resolution “Digital Twins” using EuroHPC hardware
  - Provide tools, computational analysis and optimizations during the development of the new weather and climate models.
  - Ensure that the novel EuroHPC machines are used efficiently to reach the desired throughput.
  - The creation of approaches and tests which should provide measurable indicators about the quality of the solution: reproducibility, computational efficiency and code quality
  - Accelerate the new models through kernels and data structures suitable for the new GPU partitions.
    - Computationally intensive parts will be optimised for running on GPUs
    - Optimize from the computational point the new kernels to take into account the particularities of the new EuroHPC machines.

Modeling and simulation

ESiWACE3

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

- ESiWACE Centre of Excellence (COE) have pushed the resolution of global Earth system models to unprecedented levels.
- ESiWACE3 support the weather and climate modelling community to reach a higher level regarding exascale supercomputing.



Modeling and simulation

ESiWACE3

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

- Design tools to close technology gaps for high performance computing
  - Automated tools to allow for the use of reduced numerical precision
  - Domain-specific languages to port weather and climate models to heterogeneous supercomputing hardware.
    - **PsyClone and GridTools** as real options to separate the frontend used by domain scientists to develop code from the backend that facilitates to port the code to new hardware
  - Tools to perform automated performance, energy, and accuracy testing

The pre- and exascale path

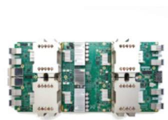
- In short time
  - New pre-exascale machines (LUMI, LEONARDO, Marenostrum5).
  - High-resolution "Digital Twins" using EuroHPC hardware
  - European Projects pursuing HPC improvements (ESiWACE3)
- In medium/long time
  - RISCV, Quantum, FPGAs
  - Clouding
  - Hardware and software acceleration (IA, GPUs)
  - …

Modeling and simulation

Software and hardware acceleration

Data Driven

GPU



TPU

- A deep learning model trained with a large amount of data.
- Developed with less effort than a physical-based model.

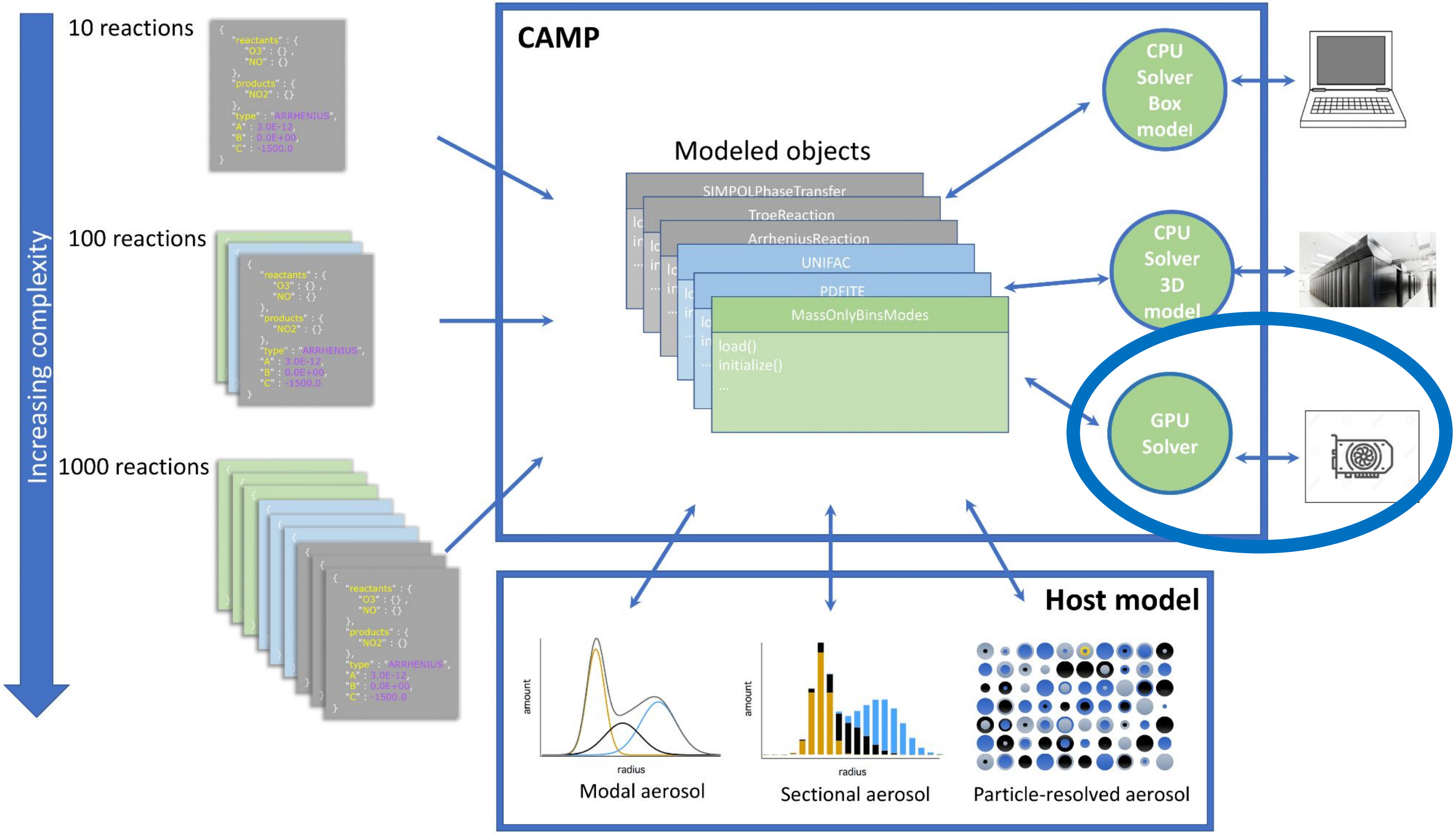
Physical model accelerated

CPU

GPU

- Traditional approach accelerating intensive parts
- GPU porting from CPU
- Progress in science “ensured” through better physical processes simulated

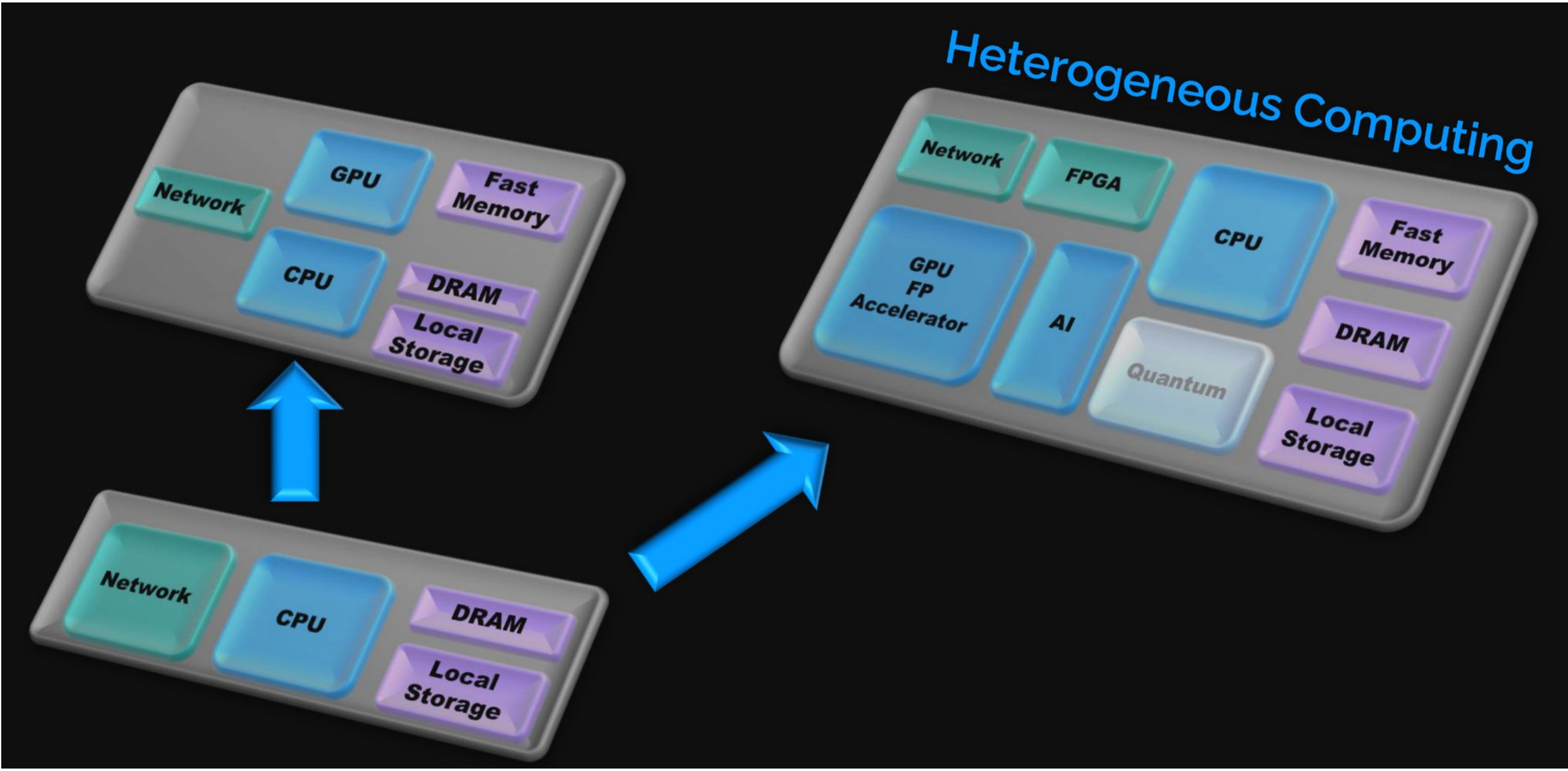
CAMP: Chemistry Across Multiple Phases



Dawson, Guzman, Curtis, Acosta, et. al., Chemistry Across Multiple Phases (CAMP) version 1.0: An integrated multi-phase chemistry mode, 2022

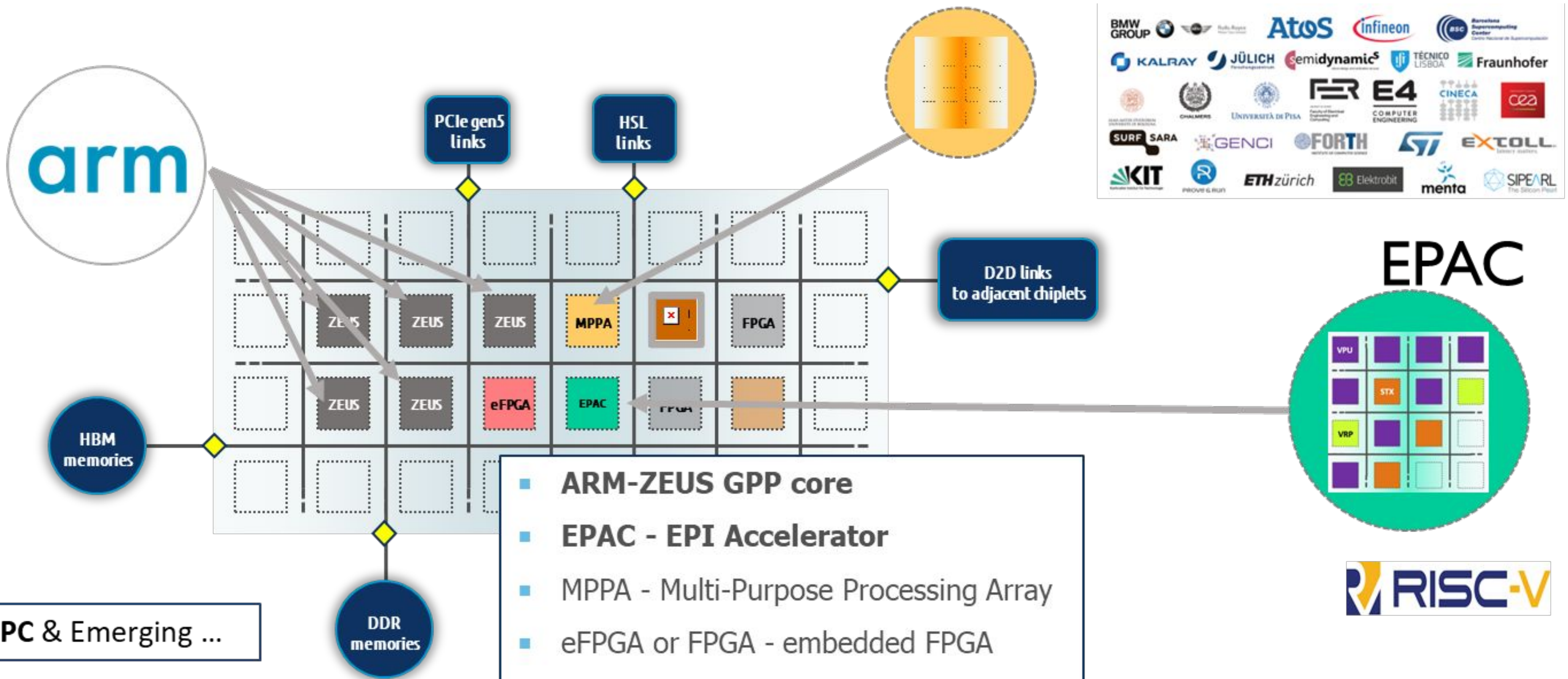
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Hardware acceleration



An approach to architectural design for exascale systems, 19th Workshop on HPC in meteorology, 2021, ATOS.

EPAC within EPI



- **ARM-ZEUS GPP core**
- **EPAC - EPI Accelerator**
- MPPA - Multi-Purpose Processing Array
- eFPGA or FPGA - embedded FPGA
- Cryptographic ASIC (EU Sovereignty)

HPC & Emerging …

The RISC-V "accelerator" in EPI, Jesús Labarta (BSC), ACM Summer Schools

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Thank you

mario.acosta@bsc.es