Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Designing HPC Architectures at the Barcelona Super-computing Center

Miquel Moretó

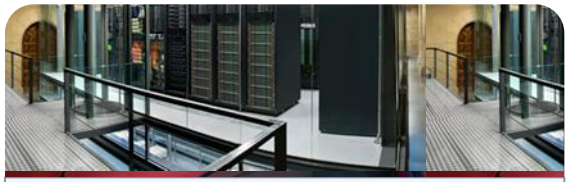
Senior Researcher

28/Nov/2018

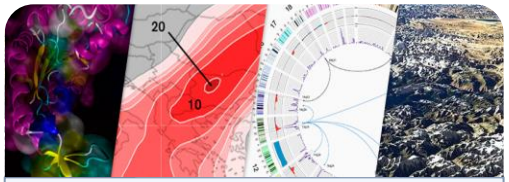
Universidad Complutense de Madrid

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

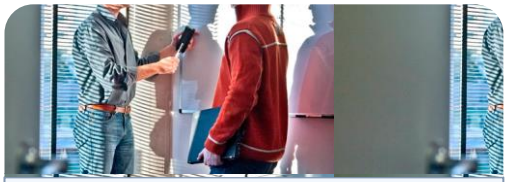
BSC-CNS objectives



Supercomputing services to Spanish and EU researchers



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



PhD programme, technology transfer, public engagement

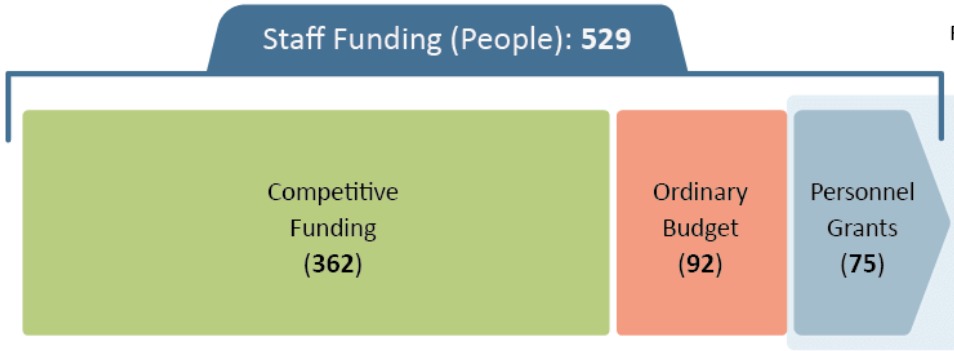
BSC-CNS is a consortium that includes

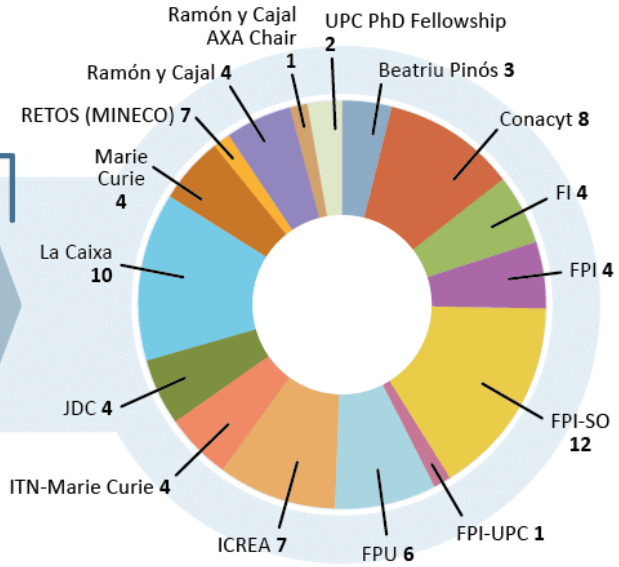
Spanish Government	**60%**	GOBIERNO DE ESPAÑA · MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
Catalan Government	**30%**	Generalitat de Catalunya Departament d'Empresa i Coneixement
Univ. Politècnica de Catalunya (UPC)	**10%**	UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

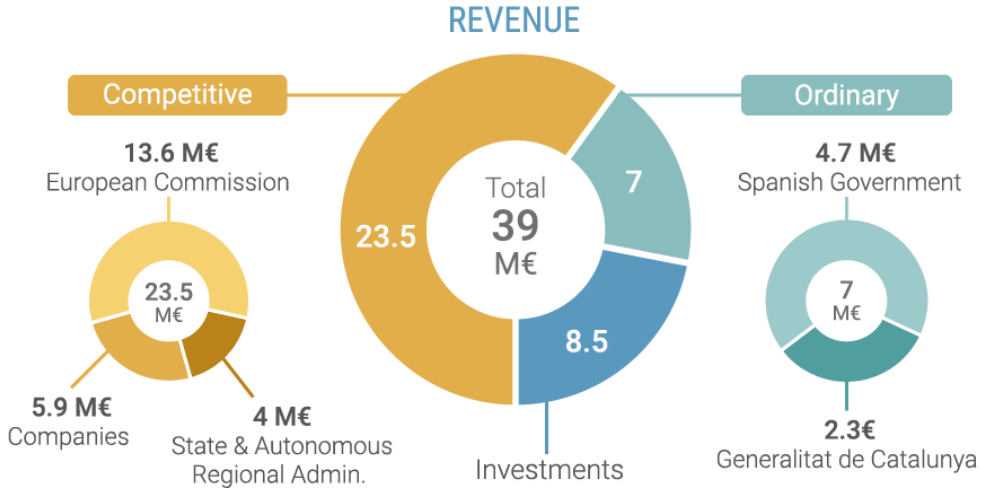
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

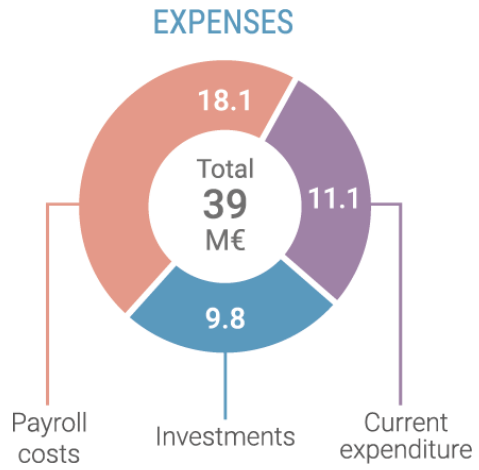
People and Resources

Data as of 31st of December 2017

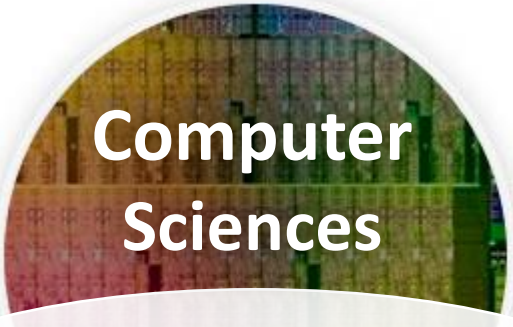




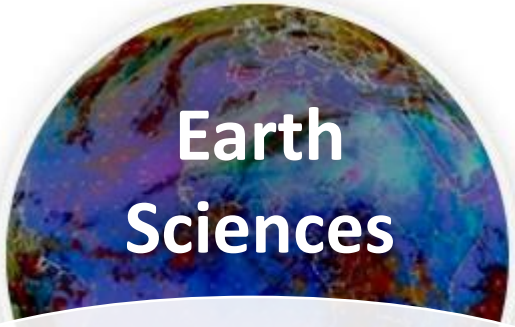




Mission of BSC Scientific Departments



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

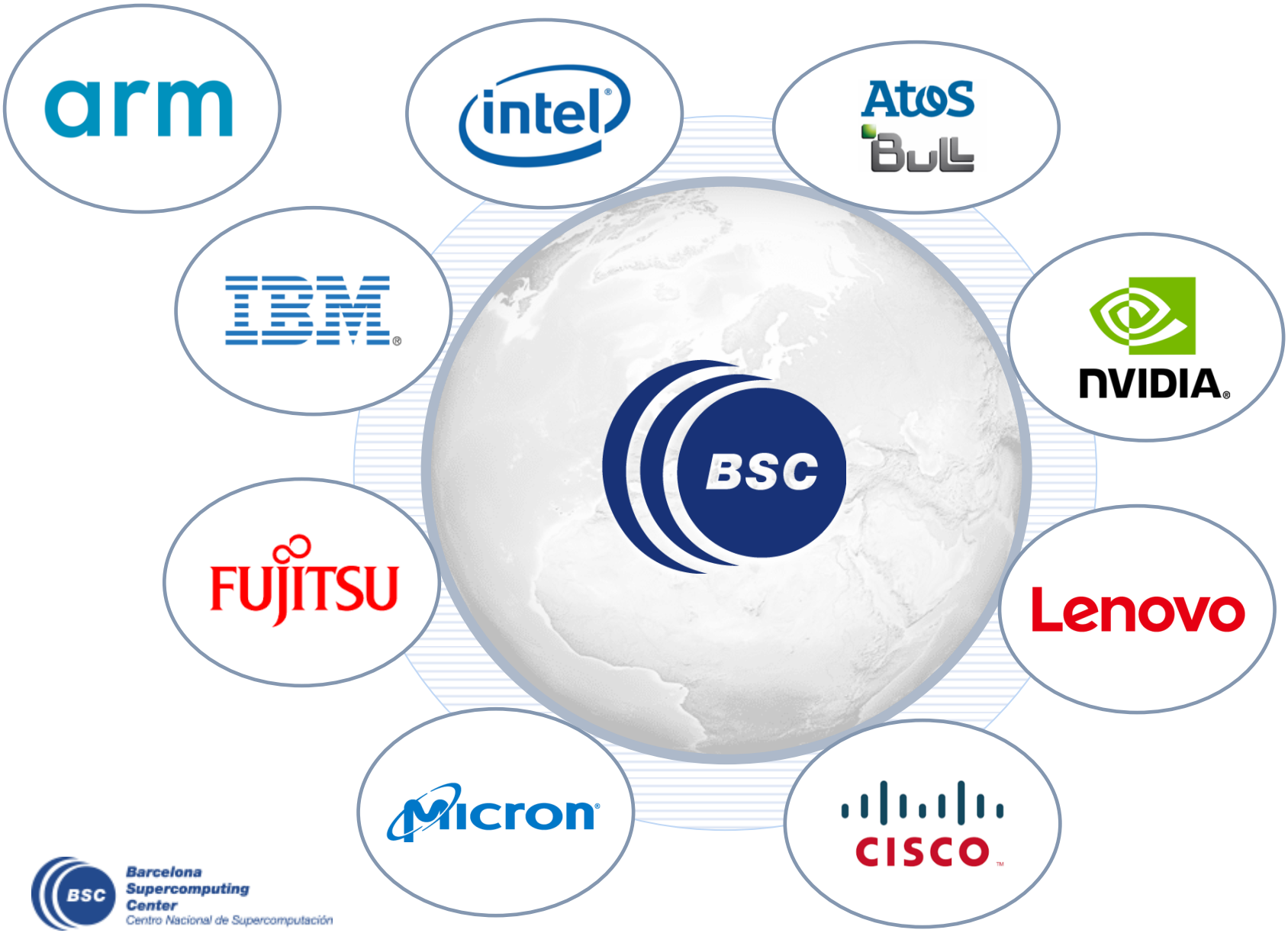


To develop and implement global and regional state-of-the-art models for short-term air quality forecast and long-term climate applications

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

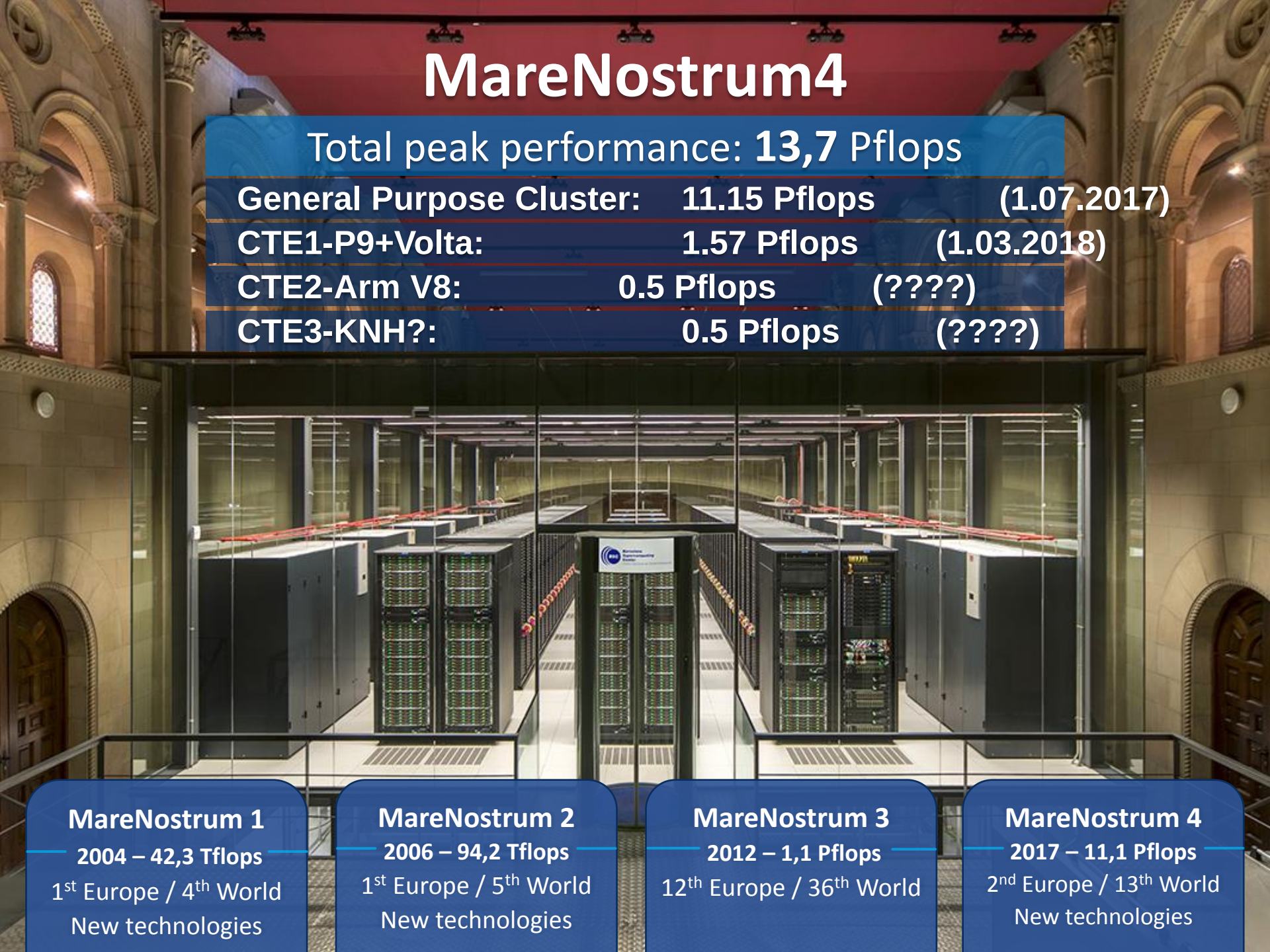
To develop scientific and engineering software to efficiently exploit super-computing capabilities (biomedical, geophysics, atmospheric, energy, social and economic simulations)

BSC & The Global IT Industry 2018



MareNostrum4

Total peak performance: **13,7** Pflops		
General Purpose Cluster:	**11.15 Pflops**	**(1.07.2017)**
CTE1-P9+Volta:	**1.57 Pflops**	**(1.03.2018)**
CTE2-Arm V8:	**0.5 Pflops**	**(????)**
CTE3-KNH?:	**0.5 Pflops**	**(????)**



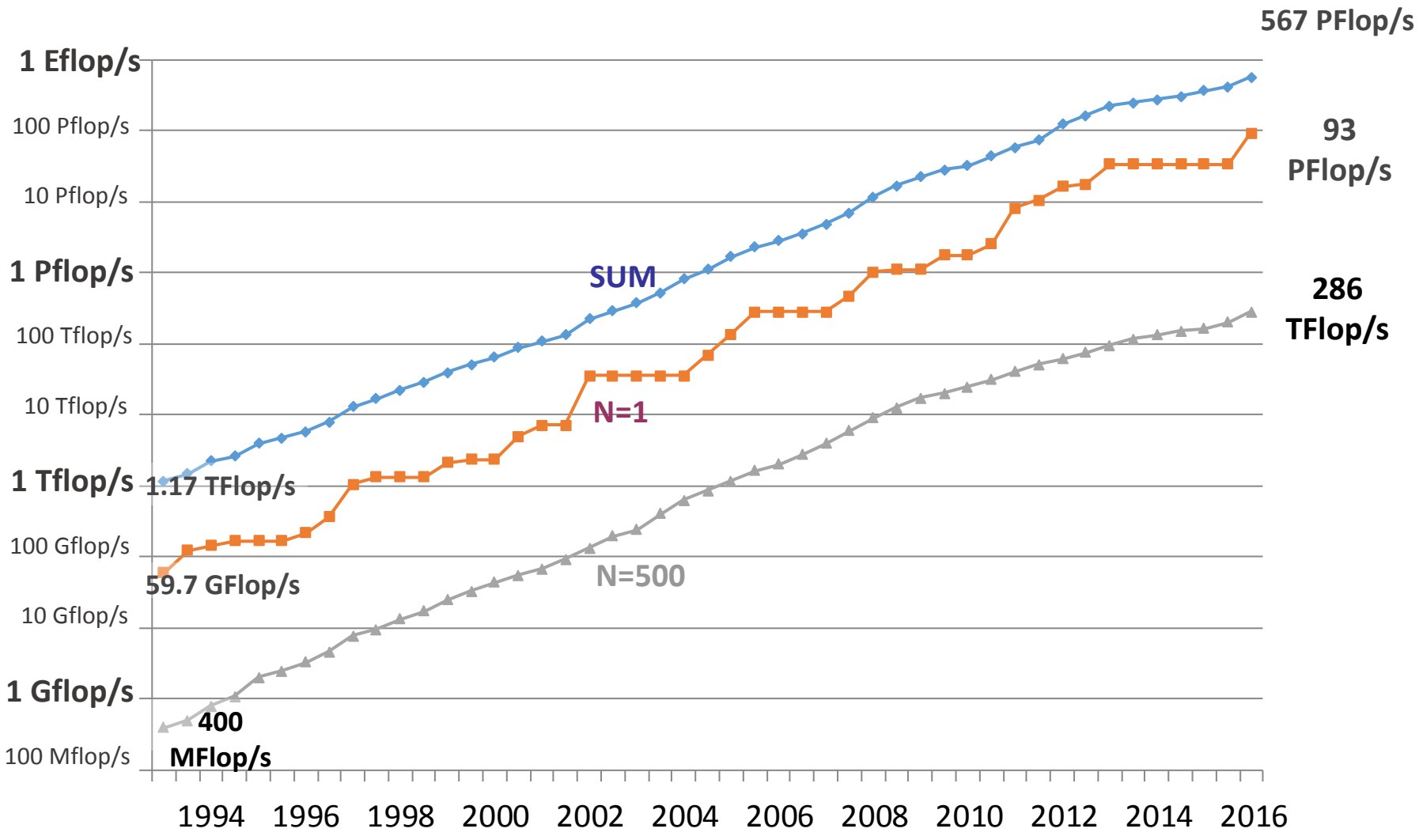
MareNostrum 1
2004 – 42,3 Tflops
1st Europe / 4th World
New technologies

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

HPC Performance Evolution from the Top500



Top10 in Top500 list

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

According to HPL

Rank	Name	Country	Rmax	Rpeak	% Efficiency
1	Sunway TaihuLight	China	93,015	125,436	74.15 %
2	Tianhe-2	China	33,863	54,902	61.68%
3	Piz Daint				
4	Gyoukou				
5	Titan				
6	Sequoia				
7	Trinity	US	14,137	43,902	32.20%
8	Cori	US	14,015	27,881	50.27%
9	Oakforest-PACS	Japan	13,555	24,913	54.41%
10	K Computer	Japan	10,510	11,280	93.17%
16	Mare nostrum	Spain	6,471	10,296	62,85%

According to HPCG benchmark

Rank	Name	Rmax	Rpeak	% Efficiency
1	K Computer	603	11,280	5.34%
2	Tianhe-2	580	54,902	1.06%
3	Trinity	546	43,902	1.24%
				1.92%
				0.38%
				1.54%
7	Cori	355	27,881	1.27%
8	Sequoia	330	20,133	1.64%
9	Titan	322	27,113	1.19%
10	Mira	167	10,066	1.66%
15	Mare Nostrum	122	10,296	1,18%

Current HPC systems suffer from significant performance inefficiencies

MN3 (2012) vs MN4 (2017): Are applications really faster with new processors?

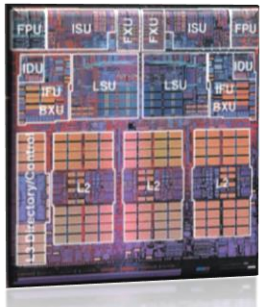
Application	Cores	Performance
WRF	256	1.37
	128	1.06
GROMACS	1024	
		1.19
NAMD		1.31
		1.17
		1.25
		1.20
VASP	64	2.2
	32	2.0
HPL	96	2.24
	48	2.21

Performance of new processors marginally increases

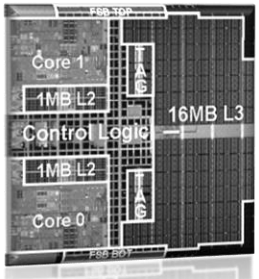
The MultiCore Era

Moore's Law + Memory Wall + Power Wall

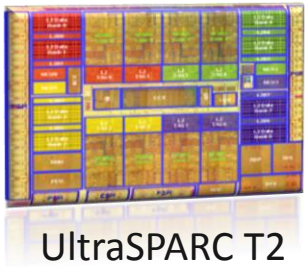
Chip MultiProcessors (CMPs)



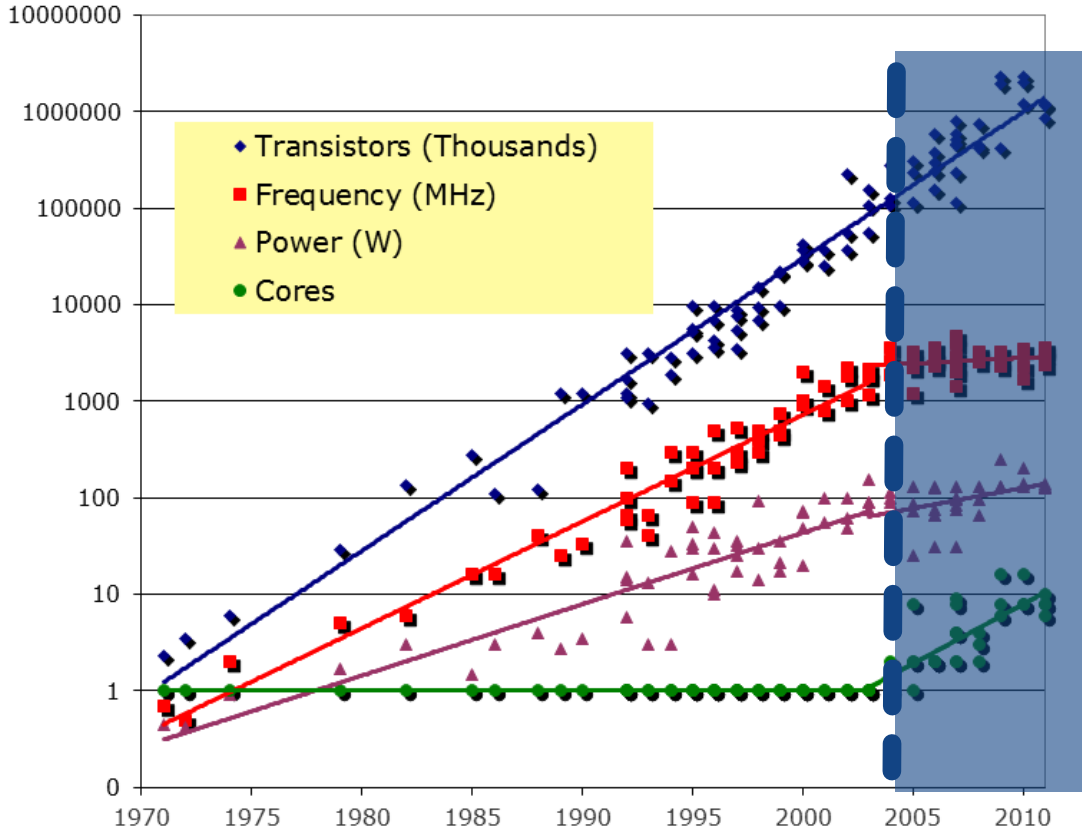
POWER4
(2001)



Intel Xeon
7100 (2006)



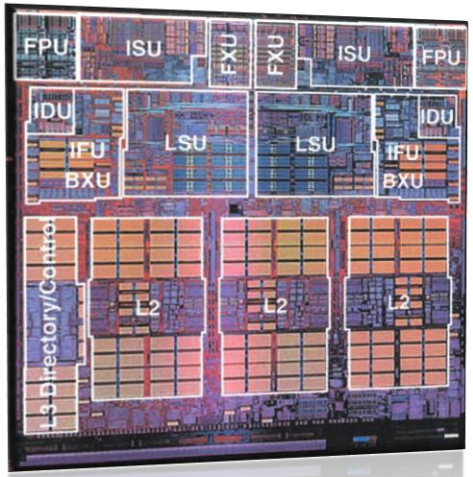
UltraSPARC T2
(2007)



How Multicores Were Designed at the Beginning?

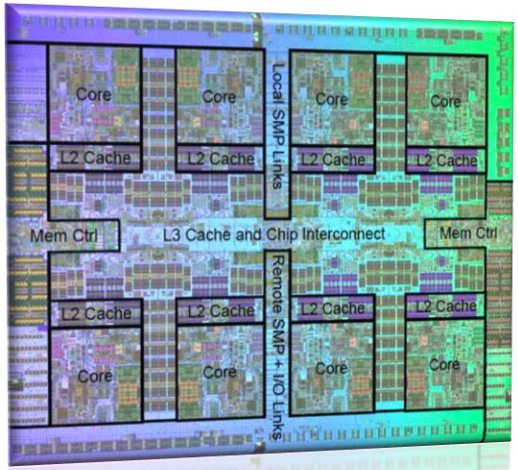
IBM Power4 (2001)

- 2 cores, ST
- 0.7 MB/core L2, 16MB/core L3 (off-chip)
- 115W TDP
- 10GB/s mem BW



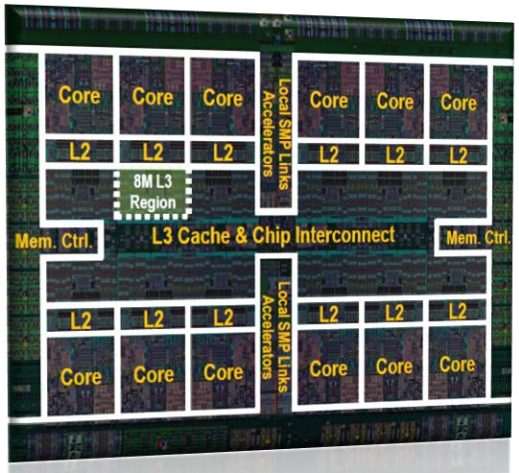
IBM Power7 (2010)

- 8 cores, SMT4
- 256 KB/core L2 16MB/core L3 (on-chip)
- 170W TDP
- 100GB/s mem BW



IBM Power8 (2014)

- 12 cores, SMT8
- 512 KB/core L2 8MB/core L3 (on-chip)
- 250W TDP
- 410GB/s mem BW



Vision in the Programming Revolution

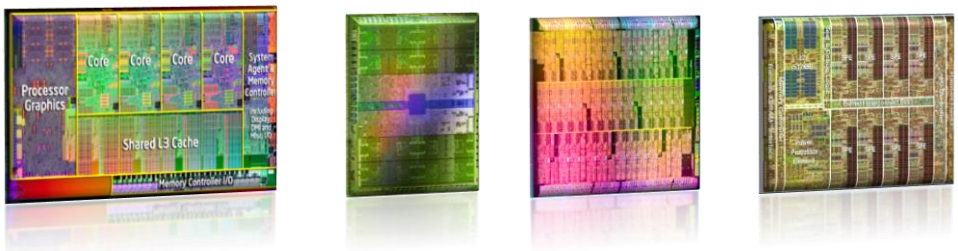
Need to decouple again

Applications

PM: High-level, clean, abstract interface

Power to the runtime

ISA / API



Application logic

Arch. independent

General purpose

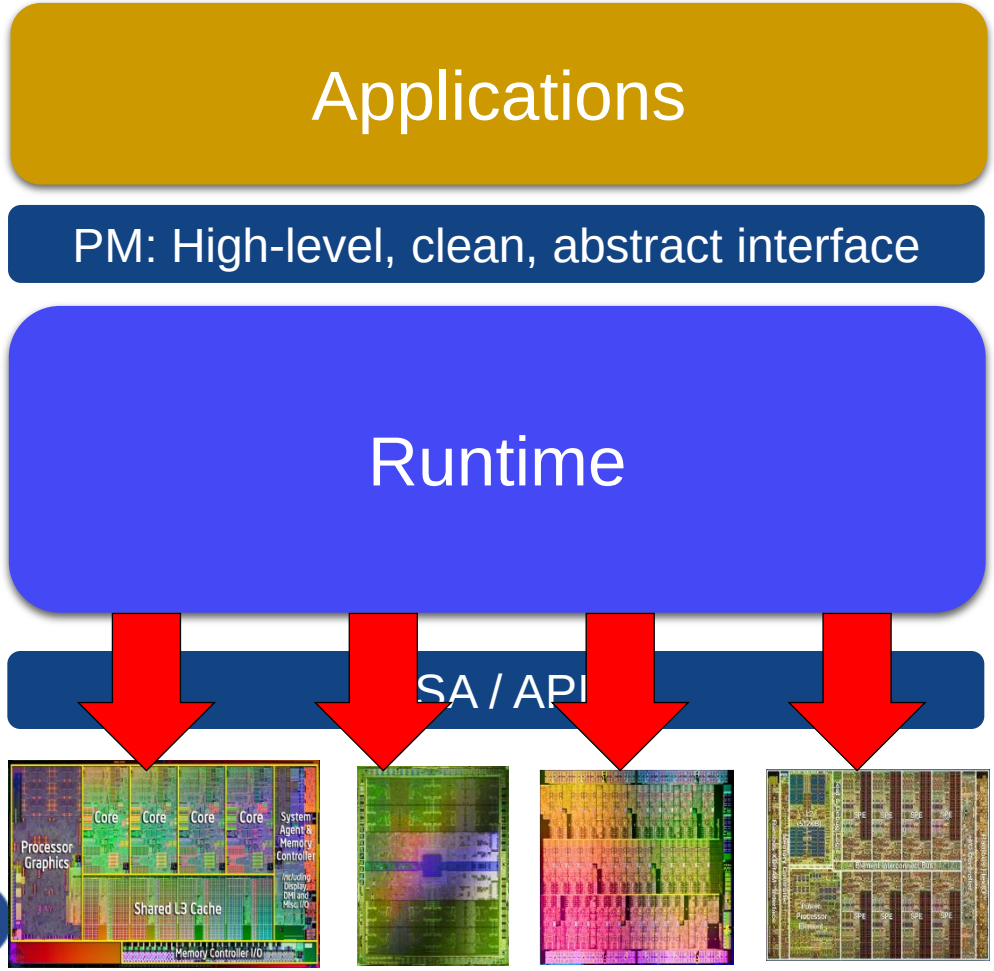
Single address space

The efforts are focused on **efficiently using** the underlying hardware

Runtime Aware Architectures

RoMoL Project

- The runtime **drives** the hardware design
  - Tight collaboration between software and hardware layers



Task based PM annotated by the user

Data dependencies detected at runtime

Dynamic scheduling

"Reuse" architectural ideas under new constraints

BSC

Research Lines (2013 – 2018)

European Research Council
Established by the European Commission

- Management of hybrid memory hierarchies with scratchpad memories (ISCA'15, PACT'15) and stacked DRAMs (ICS'18)
- Runtime Exploitation of Data Locality (PACT'16, TPDS'18, SC'18)
- Exploiting the Task Dependency Graph (TDG) to reduce data movements (ICS'18)
- Architectural Support for Task-dependence Management (IPDPS'17, HPCA'18)
- Criticality-aware task scheduling (ICS'15) and acceleration (IPDPS'16)
- Approximate Task Memoization (IPDPS'17)
- Vector Extensions to Optimize DBMS (Micro´12, HPCA´15, ISCA´16)
- Dealing with Variation due to Hardware Manufacturing (ICS'16)

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

The International Exascale Challenge

- Exascale computing: 10^{18} FLOPS
  - 150x faster than MareNostrum 4
- Sustained real life application performances, not just Linpack…
- Exascale will not just allow present solutions to run faster, but will enable new solutions not affordable with today HPC technology
- From simulation to high predictability for precise medicine, energy, climate change, autonomous driven vehicles…
- The International context (US, China, Japan and EU…)
- The European HPC programme
- The European Processor Initiative (EPI)
- BSC role

Worldwide HPC roadmaps

The White House
Office of the Press Secretary

For Immediate Release July 29, 2015

Executive Order -- Creating a National Strategic Computing Initiative

EXECUTIVE ORDER

CREATING A NATIONAL STRATEGIC COMPUTING INITIATIVE

By the authority vested in me as President by the Constitution and the laws of the United States of America, and to maximize benefits of high-performance computing (HPC) research, development, and deployment, it is hereby ordered as follows:

with domestic technology.

From K computer…

… to Post K

with domestic technology.

From Tianhe-2..

…to Tianhe-2A

with domestic technology.

IPCEI on HPC

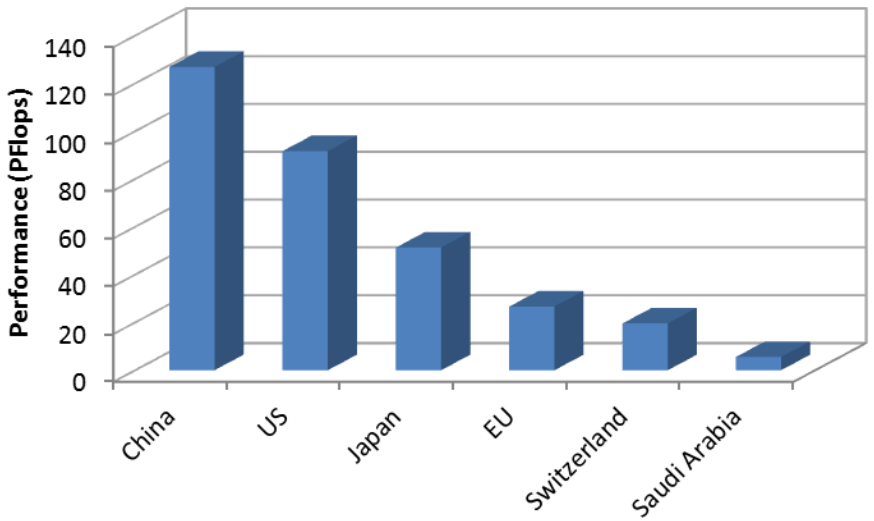
From the PPP for HPC…

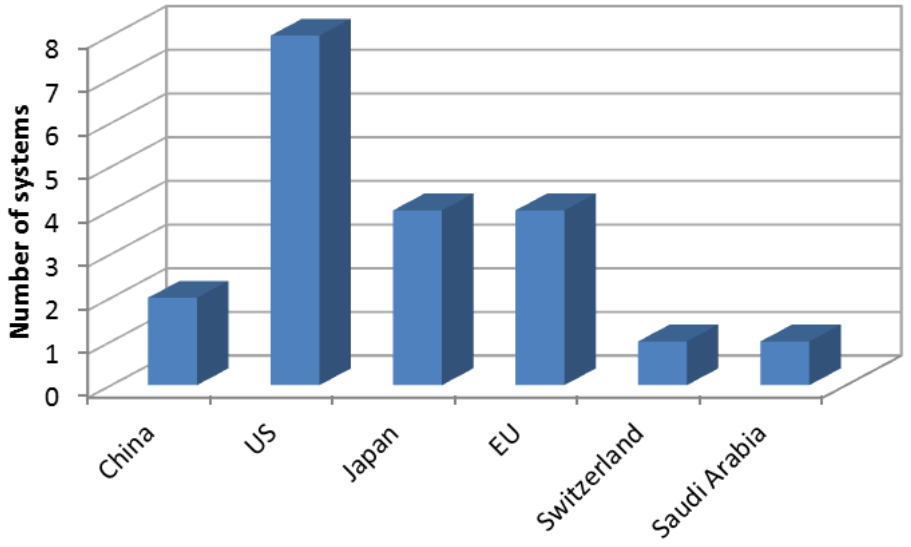
to future PRACE systems…

…with domestic technology ?

World Top 20 machines (status November 2017)

EU not in HPC world leaders

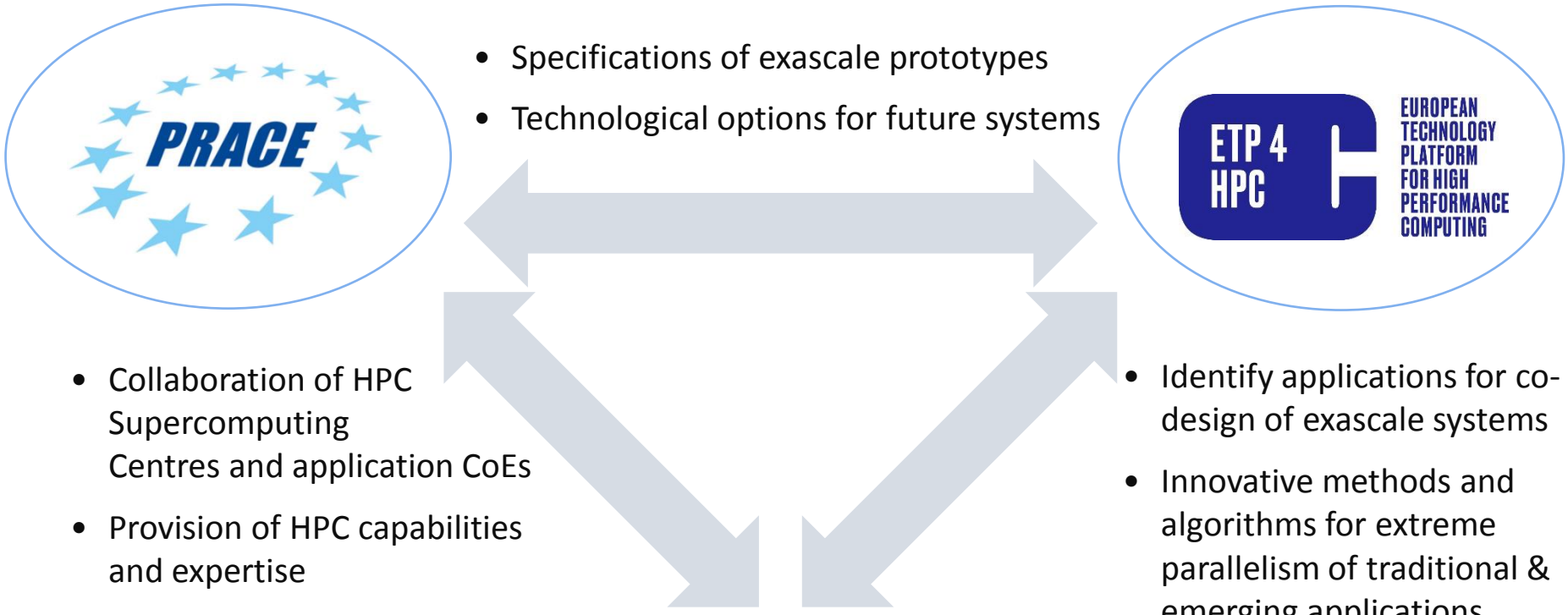




Europe has only 4 machines in world top 20

- Italy (CINECA) – Nr 14
- UK (Meteorological office) – Nr 15
- Spain (BSC, Barcelona) – Nr 16
- Germany (HLRS, Stuttgart) – Nr 19

EU HPC Ecosystem



- Specifications of exascale prototypes
- Technological options for future systems

- Collaboration of HPC Supercomputing Centres and application CoEs
- Provision of HPC capabilities and expertise

- Identify applications for co-design of exascale systems
- Innovative methods and algorithms for extreme parallelism of traditional & emerging applications

Centers of Excellence in **HPC applications**

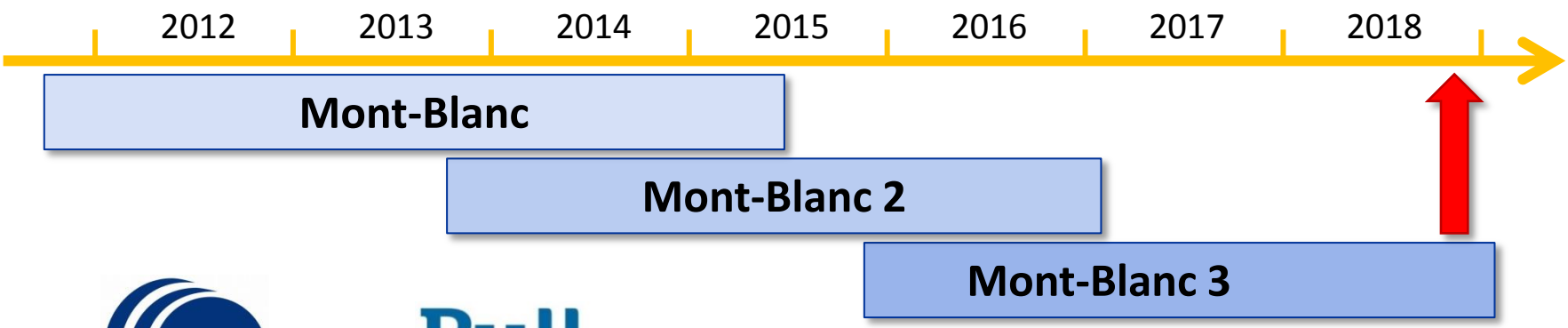
bioexcel
NOMAD NOVEL MATERIALS DISCOVERY
E CAM
POP
EoCoE
MaX
esiwace CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE
CompBioMed
CoeGSS Centre of excellence for GLOBAL SYSTEMS SCIENCE

BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación


SAGE
ExaHYPE
Memory & storage
Mathematics
NEXTGenIO
NLAFET
ExaFLOW
ExCAPE
ComPat
ExaNEST
Algorithms
READEX
ESCAPE
ALLScale
Interconnect
Programming tools
INTERTWINE
HPC
Ecosystem
EXDCI
Eurolab-4-HPC
ECOSCALE
ANTAREX
ExaNoDe
Compute
Data-intensive real-time
Mont-Blanc 3
Mango
EXTRA
GreenFLASH
Centres of Excellence
BioExcel
Biomolecular
COEGSS
Global systems
CompBioMed
Computational Biomedicine
EoCoE
Energy
E-CAM
Simulation Modelling
ESiWACE
Weather Climate
MAX
Materials
NOMAD
PoP
Performance optimisation

Mont-Blanc projects in a glance

Vision: to leverage the fast growing market of mobile technology for scientific computation, HPC and non-HPC workloads



MB2020 started!

BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

BSC contributions in Mont-Blanc

ARM-based prototypes
- Mobile technology
- Server technology
- Custom design

System software
- Scientific libraries
- Performance analysis tools
- Support for runtimes
- Power monitor

Scientific applications
- Porting and benchmarking of mini-apps and full scale applications
- Scalability study on real ARM-based platforms

Resiliency
- Application based fault tolerance
- Fault tolerance support in the runtime
- Reliability study of the Mont-Blanc prototype

Next-generation studies
- big.LITTLE studies
- Limitation analysis
- Performance projections

Mont-Blanc HPC Stack for Arm

Industrial applications

Pharmacelera, MUREX, DASSAULT AVIATION, AVL, Termo Fluids, Cenaero, ROLLS ROYCE

Applications

GENCI, CINECA, UNI GRAZ, HLRS, University of BRISTOL

System software

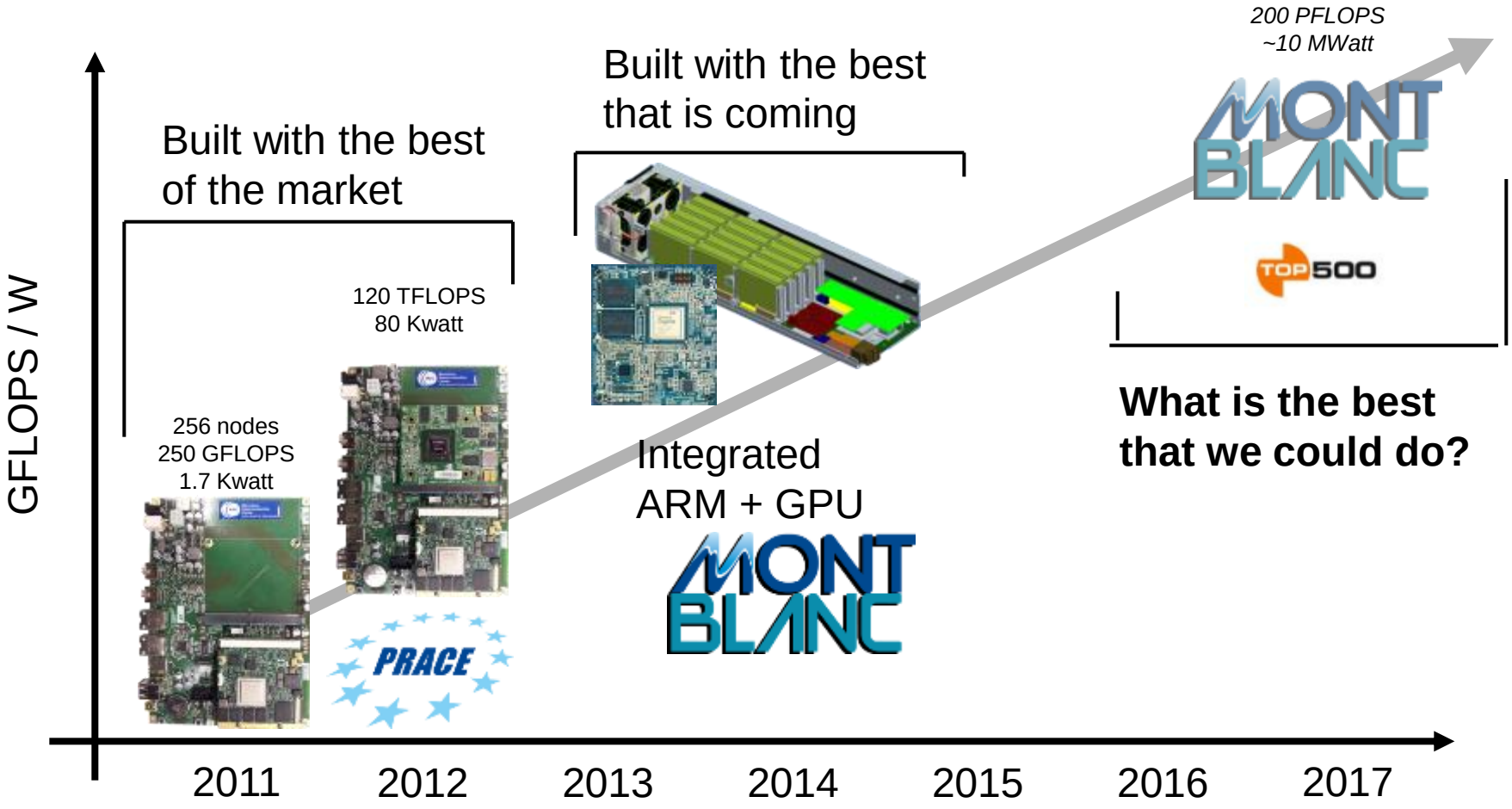
ARM, ETH zürich, Inria, BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación, JÜLICH FORSCHUNGSZENTRUM, Bull atos technologies

Hardware

ARM, BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación, Bull atos technologies, cea, UC UNIVERSIDAD DE CANTABRIA

BSC Barcelona Supercomputing Center Centro Nacional de Supercomputación

A big challenge, and a huge opportunity for Europe



- Extend current mobile chips with the needed HPC features
  - Explore the use **vector architectures** in mobile accelerators (vector processor ARM-based, 15+ Teraflops chip, 150 watts)… unique opportunity for Europe
  - One design for all market segments: mobile, data centers, supercomputers

Mont-Blanc 3: Dibona test platform

Power supply units

48 computes nodes:
16 blades
96 CPUs
3000 cores
12000 threads

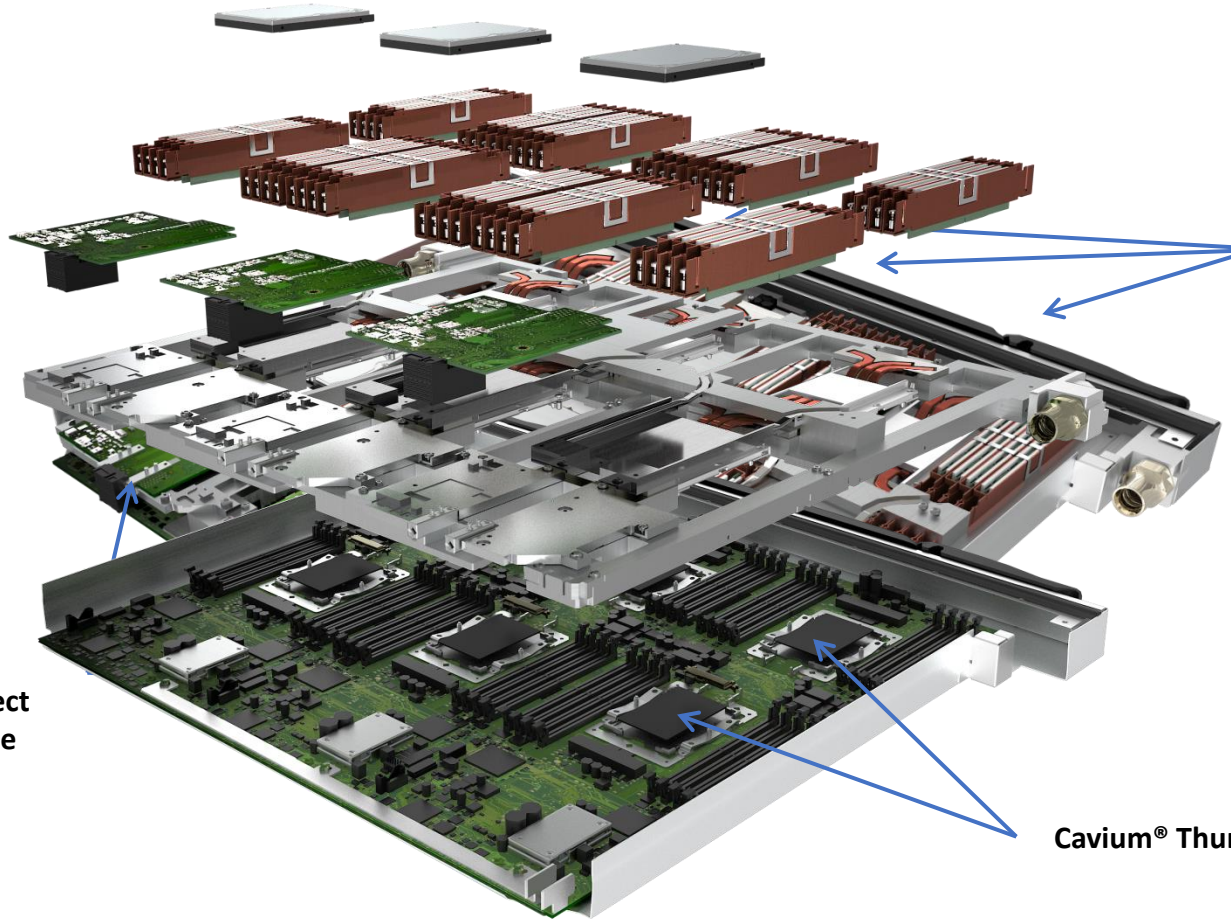
Hydraulics for Direct Liquid Cooling (ultra-energy efficient with hot water cooling)

Redundant management server including storage

IB EDR switches

Internal Ethernet management network

Dibona Processor: Cavium ThunderX2



Interconnect Mezzanine

Cavium® ThunderX2™

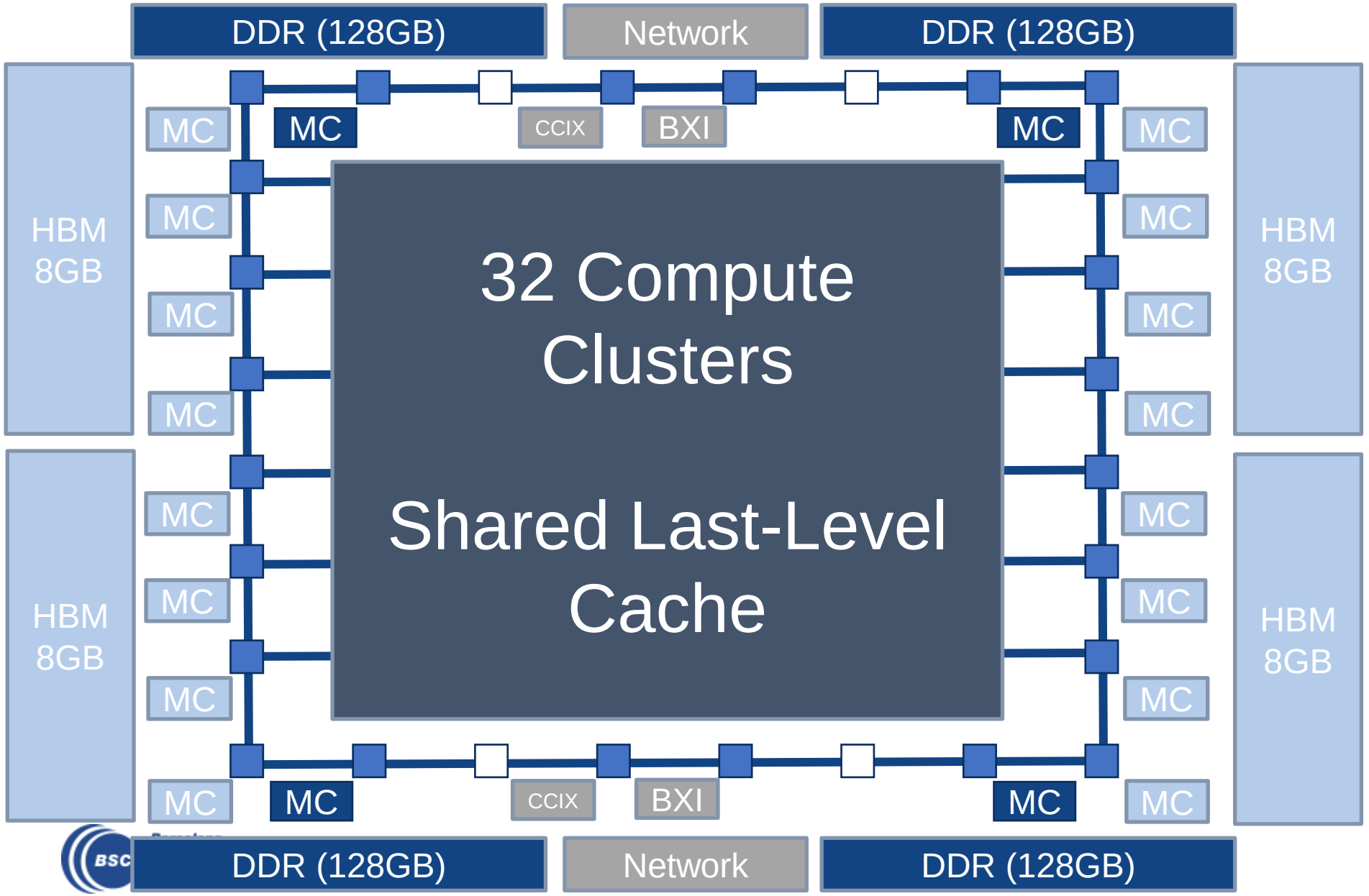
1U form factor

Direct liquid cooling

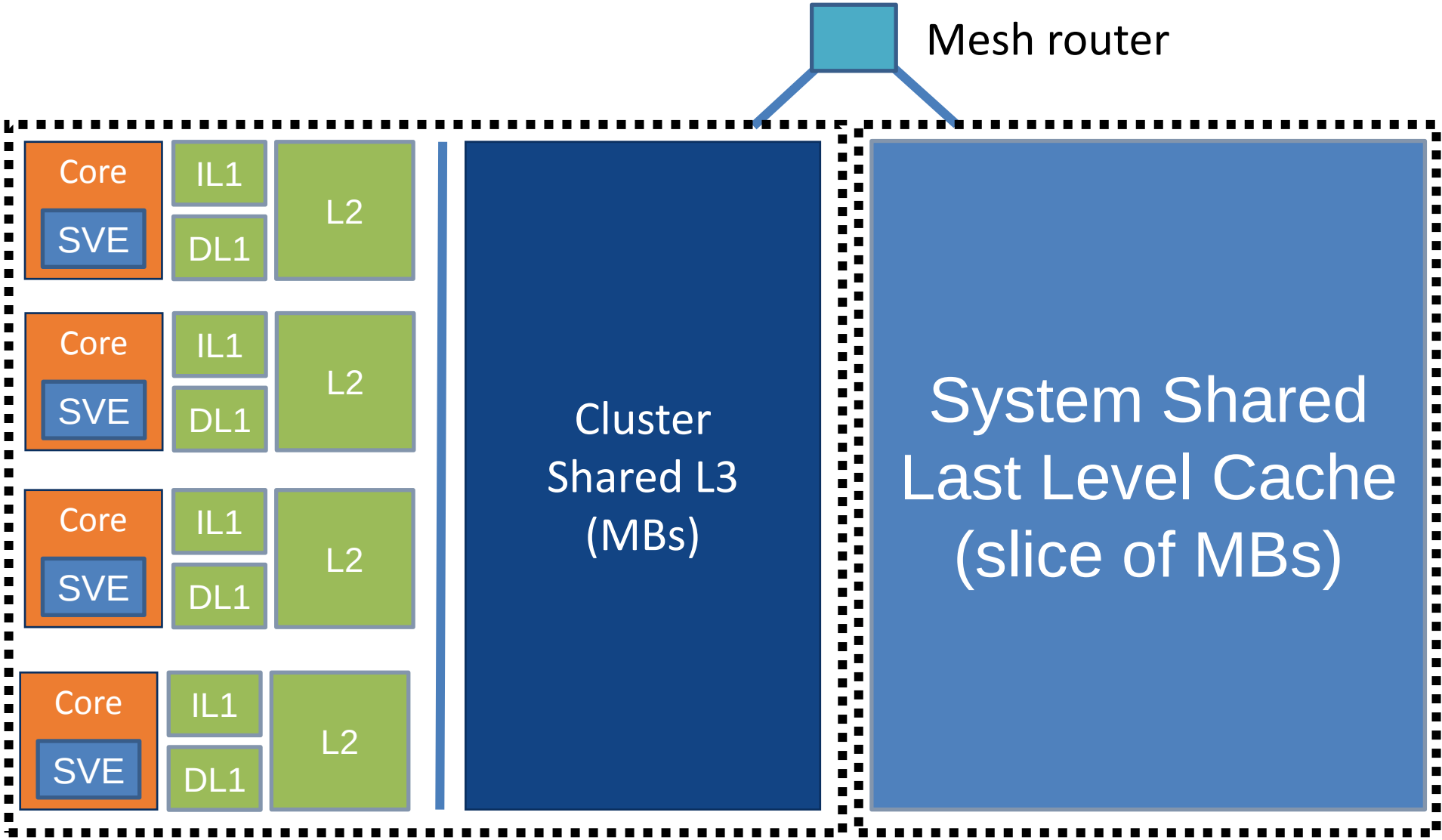
3 compute nodes per blade with:

- 2 Cavium ThunderX2 processors (up to 32 Armv8 cores per CPU, 4 threads per core, up to 2.5GHz)
- 16 DDR4 DIMM slots
- 1 Interconnect mezzanine board (EDR)

Mont-Blanc 2020: Architectural Vision



Designing Mont-Blanc 2020 Compute Clusters



BSC and the European Commission

Final plenary panel at ICT - Innovate, Connect, Transform conference, 22 October 2015 Lisbon, Portugal.

The transformational impact of excellent science in research and innovation

"Europe can develop an exascale machine with ARM technology. Maybe we need an consortium for HPC and Big Data".

Seymour Cray Award Ceremony Nov. 2015
Mateo Valero

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

The European Commission and HPC

European Commission President Jean-Claude Juncker

"Our ambition is for Europe to become one of the top 3 world leaders in high-performance computing by 2020"

Paris, 27 October 2015

Vice-President Andrus Ansip

"I encourage even more EU countries to engage in this ambitious endeavour"

- Ministers from seven MS (France, Germany, Italy, Luxembourg, Netherlands, Portugal and Spain) sign a declaration to support the next generation of computing and data infrastructures

Digital Day Rome, 23 March 2017

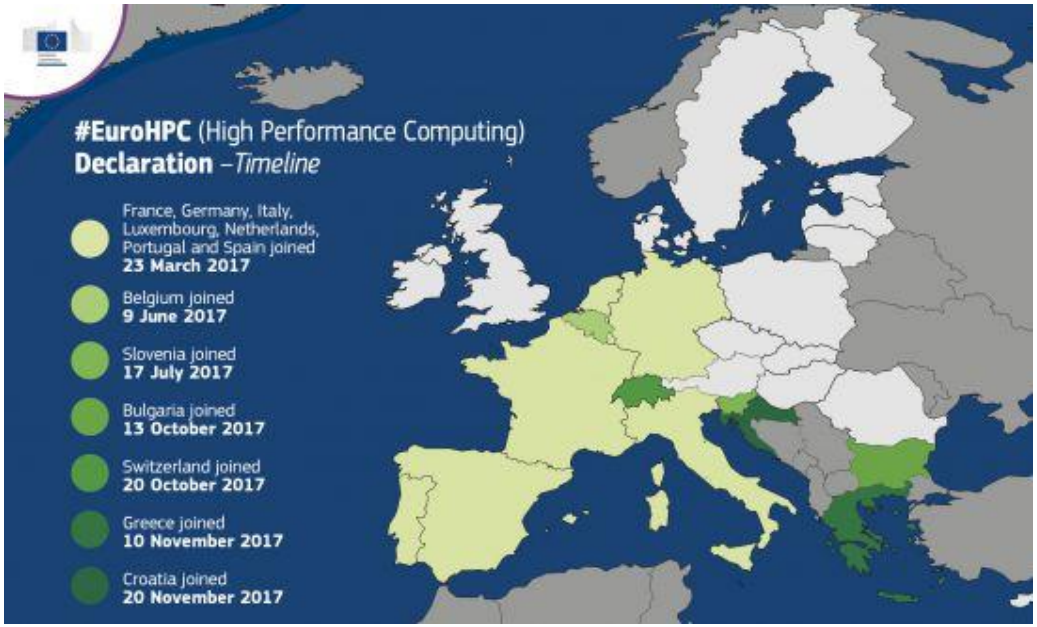
The EuroHPC Declaration

Declaration signed in Rome, March 23rd, 2017 by:

France | Germany | Italy | Luxembourg | Netherlands | Portugal | Spain

Six more countries signed the Declaration:

Belgium | Slovenia | Bulgaria | Switzerland | Greece | Croatia



Agree to work towards the establishment of a **cooperation framework** - EuroHPC - for **acquiring and deploying an integrated exascale supercomputing infrastructure** that will be **available across the EU** for scientific communities as well as public and private partners

EuroHPC news:

European Processor Initiative

- Europa portal: (January 2018)

http://europa.eu/rapid/press-release_IP-18-64_en.htm

European Commission - Press release

Commission proposes to invest EUR 1 billion in world-class European supercomputers

Brussels, 11 January 2018

The European Commission unveiled today its plans to invest jointly with the Member States in building a world-class European supercomputers infrastructure.

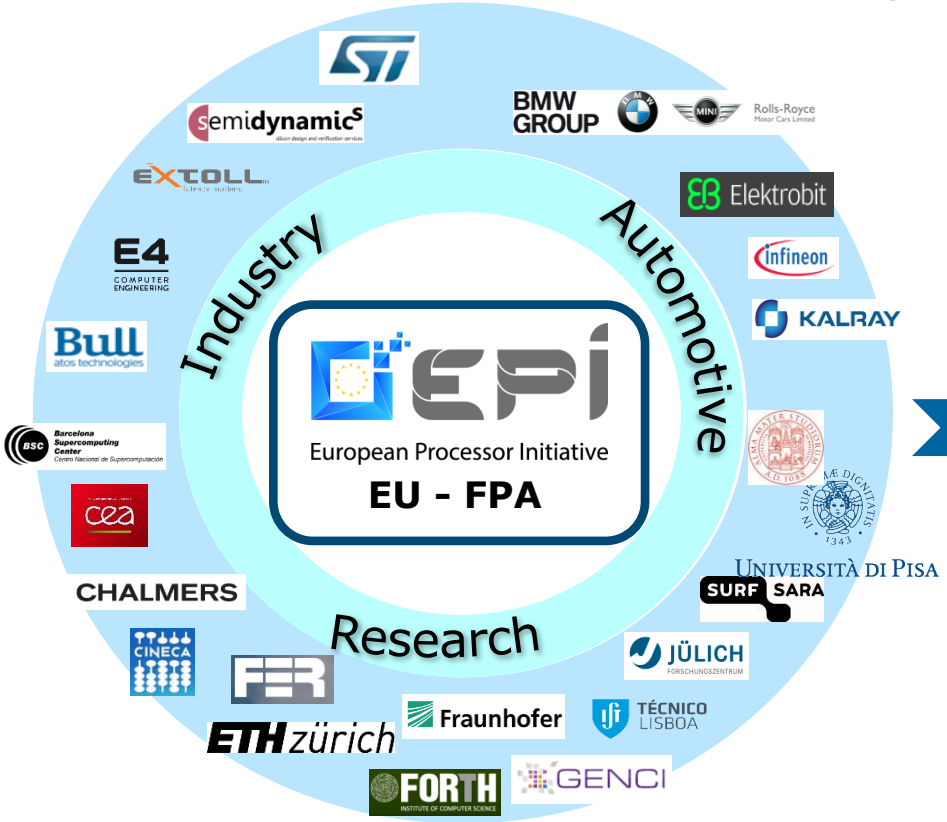
Supercomputers are needed to process ever larger amounts of data and bring benefits tothe society in many areas from health care and renewable energy to car safety and cybersecurity.

European Processor Initiative (EPI)

- Dates: end 2018 – end 2022
- Budget: **<u>120M€</u>** (+ external funding to be defined)
- EPI objectives aligned with EuroHPC initiative
- Consortium:
  - ATOS/Bull (coordinator), BSC, Infineon, BMW, ST, ExToll, E4, SemiDynamics, Juelich, ETHZ, CEA, FORTH, Kalray, …
- Define, design and manufacture:
  - European **general purpose processor** → probably based on Arm ISA
  - European **accelerator** → based on RISC-V ISA
- Suitable for multiple markets: HPC, Big Data, automotive, Internet-of-Things, etc.
- <u>Goal</u>: deploy pre-Exascale (2021) and Exascale systems (2023) based on EPI technology in selected European Institutions

European Processor Initiative

EPI 23 partners, from research to industry from consortium to EU high tech fabless



Semiconductor

Fabless company
Industrial hand of EPI
Incorporated by a couple EPI members and external investors

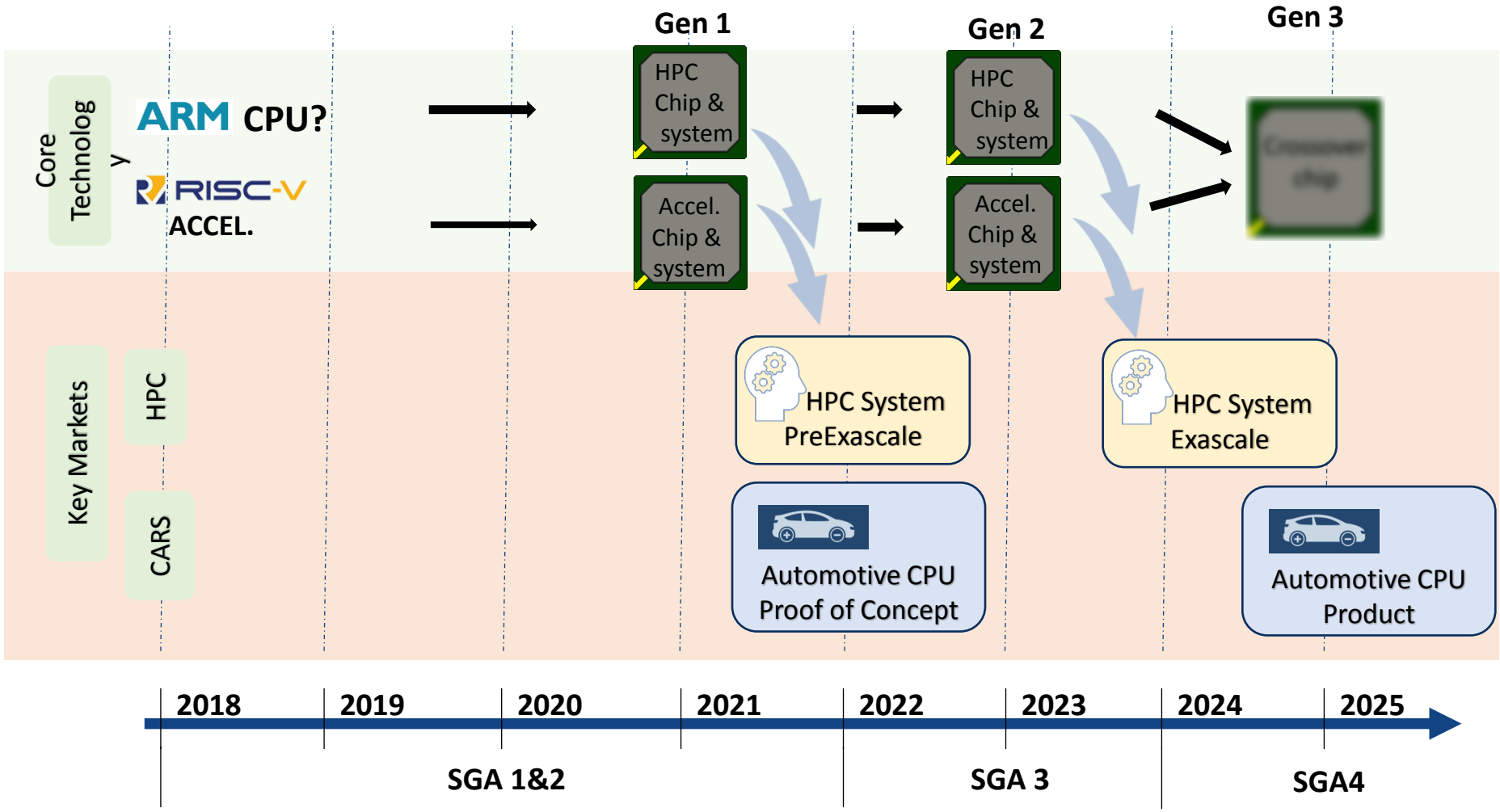
EPI Common Platform

1st EPI production

Three streams

- **General purpose and Common Platform**
  - Arm SVE or other candidates…
  - Bull: System integrator → chip integrator
  - Mont-Blanc 2020 Architectural Vision
- **Accelerator**
  - RISC-V
  - EU design: BSC, CEA, Chalmers, ETHZ, EXTOLL, E4, FORTH, Fraunhofer, IST, UNIBO, UNIZG, Semidynamics
- **Automotive**
  - Infineon, BMW…

EPI ROADMAP



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

European Processor Initiative

RISC-V accelerator vision @ EPI

- **High throughput devices**
  - **Long Vectors** (a la Cray? A la Cyber205? ...)
    - Decouple Front end - Back end engines
    - Optimize memory throughput ([Command vector, 98])
    - Explicit locality management (long register file)
  - **ISA is important**
    - Decouple/hide again hardware details, reuse SW technologies (compilers, OS,…),
    - Specific instructions?
  - **"limited" number of control flows**
- **Low power**: ~ low voltage x ~ low frequency
- **Hierarchical Acceleration**
  - **Nesting**
- **MPI+OpenMP**
  - Task based, throughput oriented programming approach
  - Malleability in application + Dynamic resource (cores, power, BW) management
  - Intelligent runtimes & Runtime Aware Architectures
    - Architectural support for the runtime
- **Accelerator for ML**
  - Specialized "non Von-Neumann" compute and data motion engines (neural/stencil)
  - Tuned numerical precision

BSC is Hiring

European Processor Initiative

BSC is looking for talented and motivated students and professionals with expertise in the design and verification of IPs to be integrated into top-level HPC SoC designs. The immediate responsibilities of this group will be related to The European Processor Initiative.

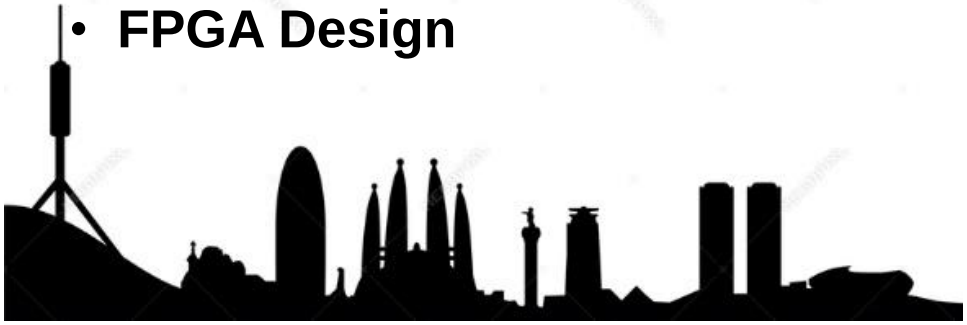
PhD students and experienced professionals (Engineers and/or PhD holders) wanted for:

- **RTL/Microarchitecture**
- **Verification**
- **FPGA Design**

Find out more:

https://www.bsc.es/join-us/job-opportunities/103csrre

Or contact: rrhh@bsc.es



HR EXCELLENCE IN RESEARCH

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

MareNostrum 5 in 2023…

Conclusions

- BSC is designing the next generation of supercomputers
  - Runtime-Aware Architectures
  - Mont-Blanc European project series
  - European Processor Initiative (EPI)

- **Goal**: Ensure technological leadership of Europe for the next decade in High Performance Computing (HPC)

- We need help from all of you to achieve such an ambitious goal ☺


Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación
BSC
EXCELENCIA
SEVERO
OCHOA
THANK YOU!
Please, contact me at:
miquel.moreto@bsc.es
www.bsc.es