Barcelona Supercomputing Center
Centro Nacional de Supercomputación

SUMMARY 2019

WHO WE ARE

Barcelona Supercomputing Center – Centro Nacional de Supercomputación (BSC) is the leader in supercomputing in Spain and a reference center in this field internationally.

BSC Missions

Supercomputing services to Spanish and EU researchers

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

PhD programme, technology transfer, public engagement

We are a Severo Ochoa Center of Excellence, a Singular Scientific and Technical Infrastructure (ICTS), first-level members of the PRACE (Partnership for Advanced Computing in Europe) European research infrastructure, we manage the Spanish Supercomputing Network (Red Española de Supercomputación) and we support the biomedical community coordinating the Elixir and INB-ISCIII Infrastructure.

Created in 2005 from the experience of successful collaborations between public administrations and private companies, such as CEPBA and CIRI, we have grown rapidly from 60 to nearly 700 staff members, thanks to our ability to attract competitive funds from institutions and companies, and we have already installed four consecutive versions of the MareNostrum supercomputer.

BSC is a public consortium made up of:

Generalitat de Catalunya
Departament d'Empresa i Coneixement

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

UPC
UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Where **we are,** where **we are going**

Mateo Valero and **Josep M. Martorell**
Director and associate director of BSC

The first word that springs to mind when we look back at 2019 is "Europe". This year our desire to promote high-performance computing at European level, which was envisaged in the EuroHPC declaration signed in Rome in 2017 in which BSC played a very important role, has taken concrete form. The results have been a cause of great satisfaction for our center.

On the one hand, in June the European Commission announced that BSC had been chosen as the hosting entity for one of the three large European pre-exascale supercomputers. This great success, which was enabled by the excellent work of our teams and the enthusiastic support of our patrons and the governments of Portugal, Croatia and Turkey, will make it possible to start installing the new MareNostrum 5 by the end of 2020 with €217 million in investment by 2025, including the acquisition, delivery, installation and maintenance of the supercomputer.

On the other hand, this European alliance is also intended to promote technology designed on our continent to contribute to enhanced European technological sovereignty. BSC is an important leader in the great initiative to develop European chips (the European Processor Initiative), a project that is going hand-in-hand with our efforts regarding open architecture. This work has already led to the manufacture of Lagarto, the first open-hardware based processor developed in Spain. Accordingly, in November we announced the setting up of a new European Open Architecture Laboratory (LOCA), which will be the basis for public-private collaboration on these technologies. In addition, the future MareNostrum 5 will play a role in taking us down this path: it will be the only supercomputer promoted by the EC that will host an experimental platform to develop and test European technologies, the MareNostrum Experimental Exascale Platform (MEEP).

All of this magnificent news emanates from a center that is continuing to grow in size and budget: we have ended the year with more than 650 people and a current budget in excess of €39 million. We remain the Catalan institution that attracts the most funds from the European research program (now more than €75 million from Horizon 2020) and the third in Spain. Our scientific productivity has grown spectacularly: the number of articles published in first-quartile journals has doubled in just 4 years. Our collaboration with companies and governments is continuing to expand and we now have 4 spin-offs in full operation. There is also growth in infrastructure: at the end of this year we have installed the two latest MareNostrum 4 clusters, which consolidates our commitment to managing a system with a diversity of architectures that can meet the needs of as many user communities as possible with minimum energy cost.

The quality of our research has received new recognition again this year, notably including the arrival of three new ICREA researchers, three new ERC projects, the national award for public-private collaboration with Repsol and important awards to some BSC researchers. Companies such as ARM, Asepeyo and Aigües de Barcelona have signed new joint research agreements with BSC and as many as 31 of our doctoral students have defended their theses.

When you read through the following pages, it is easy to see that artificial intelligence is one of the areas in which our center is seeing the greatest growth. Its influence is multiplying in all knowledge areas and will undoubtedly be a major theme in all kinds of areas in the coming years, as recommended in the Strategic Guidelines approved by our Board of Trustees in 2019, based on the recommendations made by our Scientific Advisory Board (SAP) at the end of 2018.

We cannot overlook the many different activities with an impact in terms of public communication and education. Many people with public responsibilities visited us in 2019 (notably including Pedro Sánchez, Roberto Viola, Pedro Duque and Àngels Chacón) and almost 20,000 people came to see MareNostrum 4, a new record in BSC's history. These visitors include 5,500 primary school students as part of the "We Are Researchers" program and the first year of "Crazy about Supercomputing".

All in all, this has been a very intense year filled with great news, a prelude to a 2020 from which we expect great new things: our new headquarters and our new supercomputer await us!

The ceremony for the signing of the contracts to host the new European supercomputers was the culmination of a year of intensive work aimed at ensuring that Barcelona Supercomputing Center would play a major role in this new HPC roadmap. The event was presided over by Commissioner Mariya Gabriel, one of the key people in securing a joint HPC project for the European Union.

BSC in numbers

RESEARCH

51 Research groups

6 ERC Projects in progress

8 ICREA researchers

80 Researchers with personal scholarships

€75.8 M HORIZON 2020

#3 Third Spanish entity receiving funds from the H2020 Program

119 H2020 projects in progress

36 H2020 projects begun in 2019

14 H2020 projects coordinated by BSC

Presence in the 10 European Centres of Excellence for HPC applications:

9/10 Participating in

2/10 Leading

223 Articles in peer-reviewed journals

172 In Q1

101 Articles in conference proceedings

49 In A and A* conferences

31 Doctoral theses

111 Open access publications

TECHNOLOGY TRANSFER

Licensed or transferred technologies

4 Patents

9 Software

4 Spin-offs created

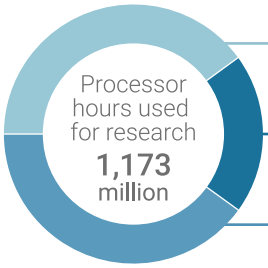
Bilateral collaboration with companies

49 In progress

5 Started in 2019

27 Patents approved or pending approval

SUPERCOMPUTING



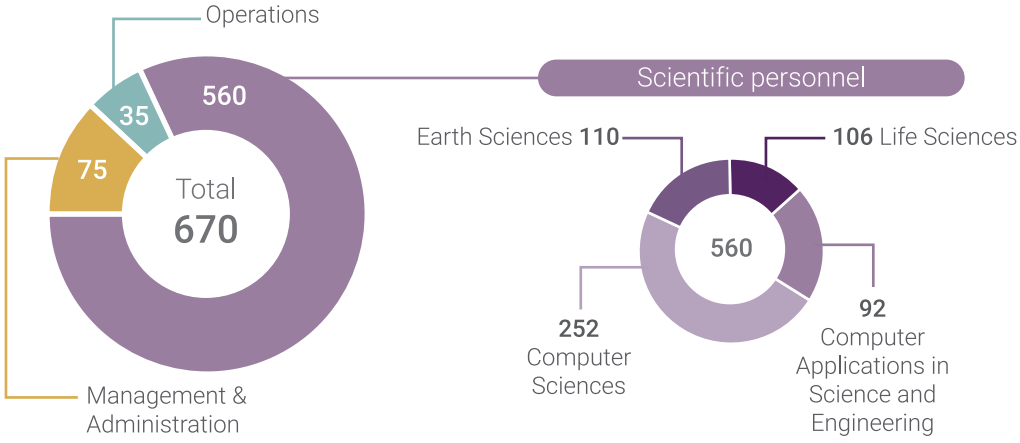
40% **RES** - 285 activities

20% **BSC** - 373 researchers

40% **PRACE** - 14 projects

PEOPLE

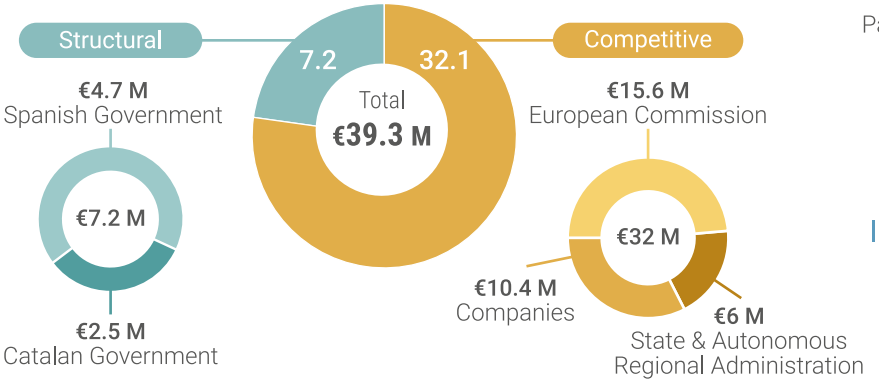
Data as of December 31, 2019



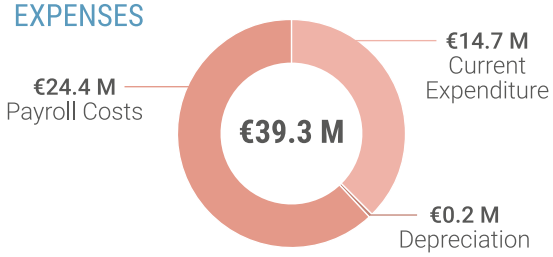
RESOURCES

2019 executed budget

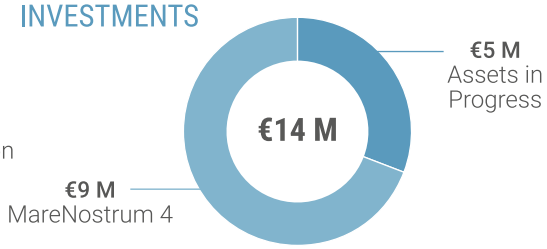
REVENUE



EXPENSES



INVESTMENTS



UPC contributes, in terms of assigned staff and room space, the equivalent of 10% structural funds.

Revenue and expenses according to finance criteria. Investments according to budgetary criteria.

New **BSC facilities** map

In 2019, BSC focused its efforts on the final sprint to make the center's new facilities map a reality in 2020. This new map will make it possible for BSC's nearly 700 workers to finally no longer be spread among six buildings and come together in three buildings in the gardens of Torre Girona. This concentration has been long-awaited because it will increase communication and then create synergies among the diffe-

BSC-Repsol Building

The ground floor and the four upper floors of the new corporate building will house 520 of BSC's workers, including the center's management. The building will also have an auditorium, 38 meeting rooms, 8 office rooms, a nursing and breastfeeding room, microwave areas and a bar/restaurant on the ground floor.

A new machine room and a visitable area

The new computer room will be 1,000 m^2 in size and will house MareNostrum 5 and other HPC clusters that are currently spread around various buildings. It will be suitable for educational visits about supercomputing and BSC's activities.

Torre Girona Chapel

The Operations Department will continue to work in the upper floors and its activity will extend into the north wing of Torre Girona.

The foyer and the chapel will continue to be the center's most iconic areas and will be used for educational visits and corporate and institutional events.

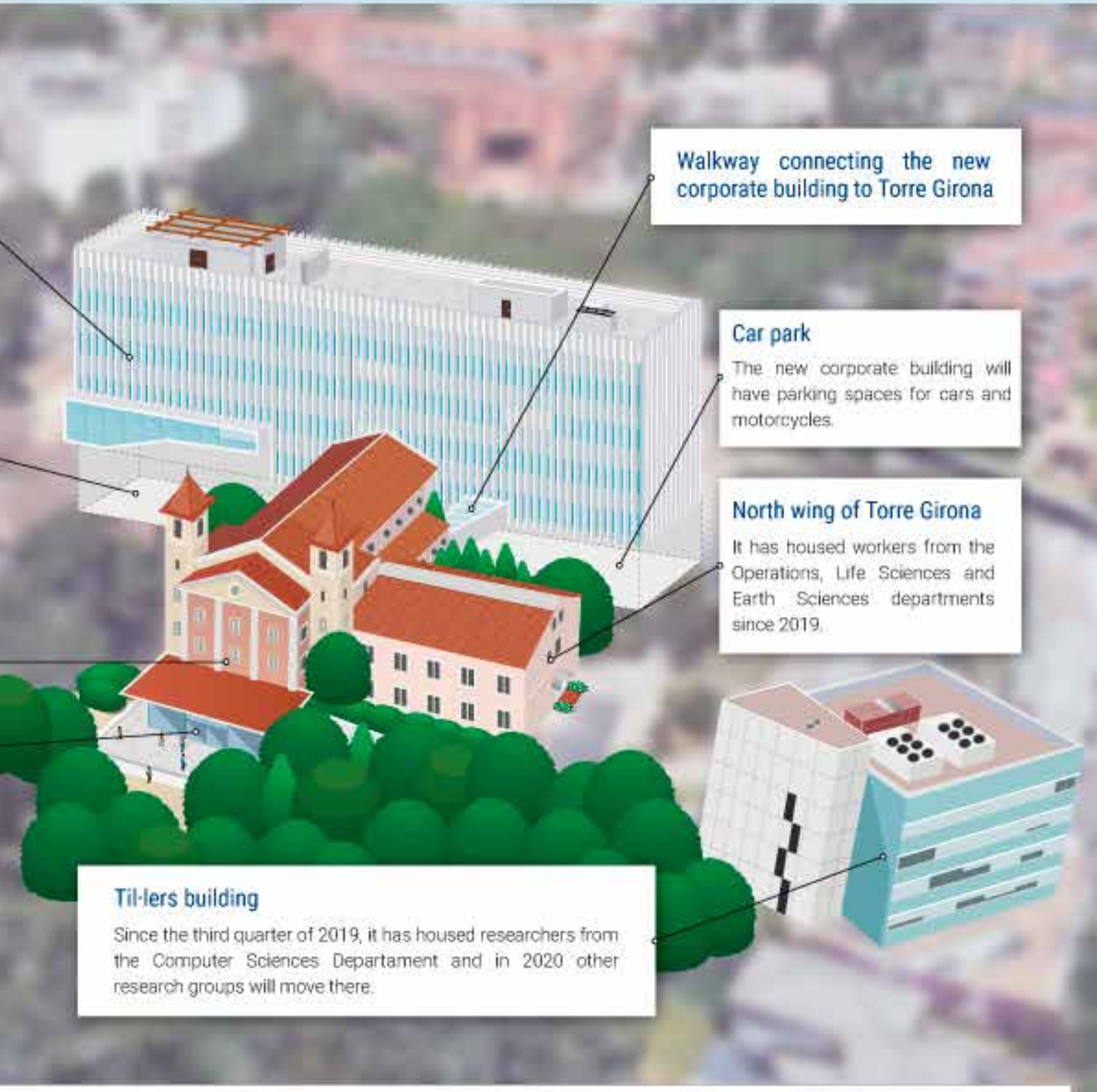
New electrical substation

A new basement electrical substation will supply the necessary power for the center's supercomputing infrastructure.

REPSOL Fundación

BSC's new corporate building has received additional funding from Fundación Repsol and the ICTS-Actions Program for the promotion of Large Scientific and

rent research teams. This change will be made possible by the UPC collaboration, the completion of the corporate building and the occupation of Til·lers building and the north wing of Torre Girona.
In addition to offices, the corporate building will house BSC's new computer room, where a large part of the MareNostrum 5 supercomputer will be installed.



Walkway connecting the new corporate building to Torre Girona

Car park

The new corporate building will have parking spaces for cars and motorcycles.

North wing of Torre Girona

It has housed workers from the Operations, Life Sciences and Earth Sciences departments since 2019.

Til·lers building

Since the third quarter of 2019, it has housed researchers from the Computer Sciences Departament and in 2020 other research groups will move there.

Technological infrastructures, with the participation of the Government of Catalonia and the European Regional Development Fund.

Generalitat de Catalunya

Unió Europea
Fons Europeu de Desenvolupament Regional

From MareNostrum 4...

The installation of the MareNostrum 4 supercomputer was completed in 2019 with the arrival of the final two emerging technology clusters. The cutting-edge technologies of these clusters, used to build exascale supercomputers that will soon come into operation in the United States and Japan, puts us, once again, in the vanguard of high-performance computing.

MareNostrum 4 supercomputer

13.9 petaflops: around 14 quadrillion floating point operations per second

40%

PRACE

Access: prace-ri.eu/hpc-access

40%

RES

RED ESPAÑOLA DE SUPERCOMPUTACIÓN

Access: bsc.es/res-intranet

20%

BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

New emerging technology clusters

MN4 CTE AMD:

Peak performance: 0.51 Pflops.

Architecture: 34 nodes with an AMD Rome, 64c and two AMD Radeon Instinct MI50. 34 terabytes of RAM.

The same technology as will be used in the American Frontier 1.5 exaflop supercomputer.

MN4 CTE ARM:

Peak performance: 0.65 Pflops.

Architecture: 192 nodes with an ARMv8. 6 petabytes of RAM and a Direct Liquid Cooling system.

The same technology as will be used in the Japanese Fugaku 0.5 exaflop supercomputer.

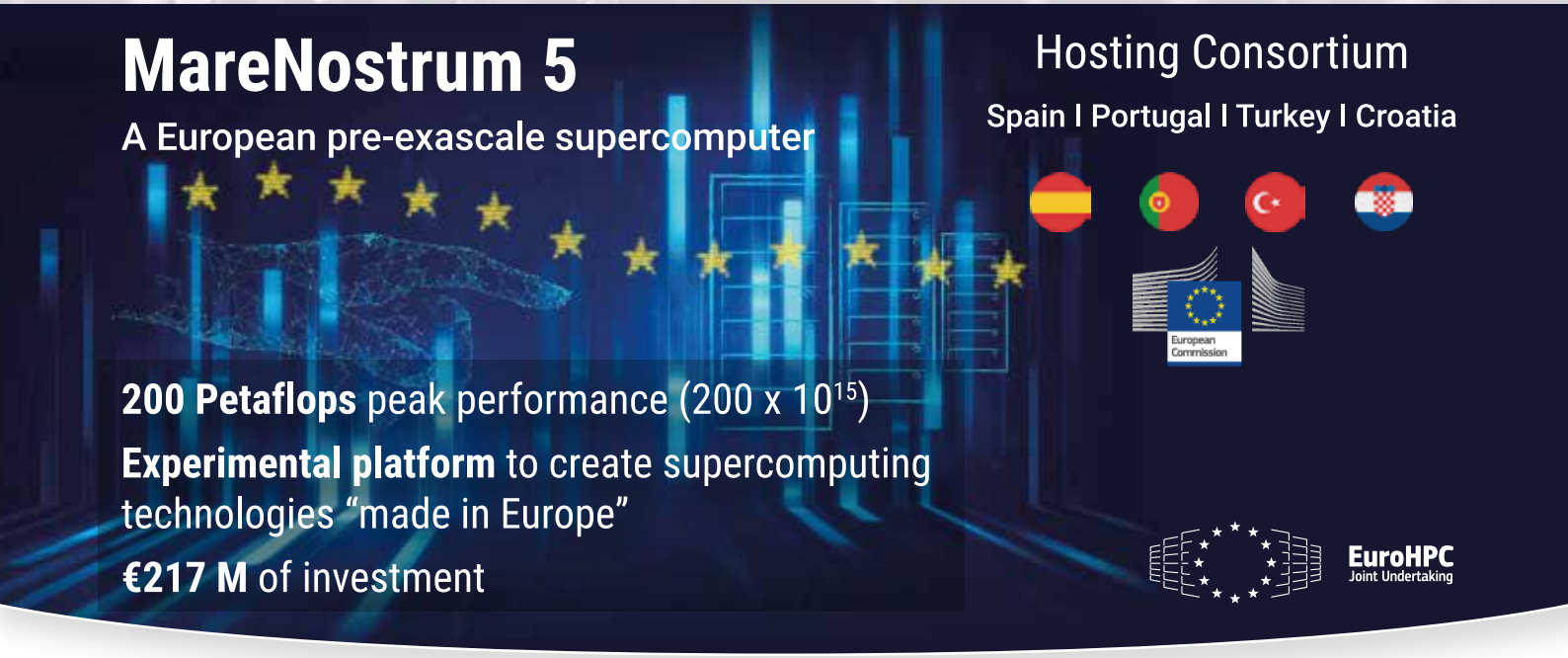
MareNostrum 4 is co-financed by the Smart Growth Operational Program 2014-2020 of the European Regional Development Fund (ERDF)

Unión Europea
Fondo Europeo de Desarrollo Regional
"Una manera de hacer Europa"

GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

to MareNostrum 5

A particular highlight this year was the magnificent news that BSC will be one of the three European centers to house pre-exascale supercomputers promoted and co-funded by the European Commission. We submitted our application in April and in November the agreement to house the future MareNostrum 5 supercomputer by the end of 2020 was signed.



MareNostrum 5 will be a pre-exascale supercomputer with a peak performance of approximately 200 petaflops (200 quadrillion operations per second), which will start to operate by the end of 2020. It will cost 217 million euros, 50% of which will be funded by the European Commission and the rest by a consortium made up of the governments of Spain, Croatia, Portugal and Turkey. The Catalan Government will also contribute.

MareNostrum 5 ensures that Barcelona will continue to hold a leading position in the European high-performance computing map. Apart from its relevance as infrastructure made available to researchers, it is of great interest in research into the supercomputers of the future, as around 5% of the budget will be used to incorporate an experimental platform to develop new HPC technologies "made in Europe" for future generations of supercomputers.

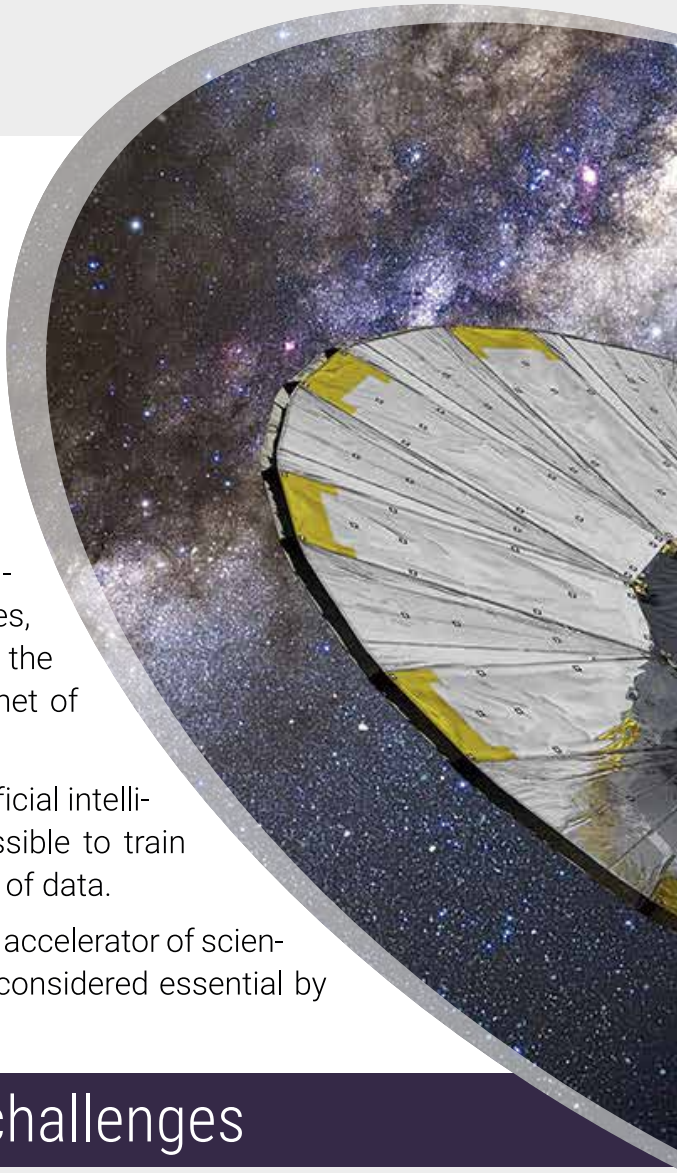
'The EuroHPC Joint Undertaking is deploying the first European supercomputers, as part of its mission to build a world class supercomputing and data infrastructure in Europe, that will support public and private users in developing leading scientific and industrial applications in a wide range of areas, and to make Europe a top supercomputing region globally'

At the service of science

From the handcrafted telescopes used by Galileo to the particle accelerator in Geneva, scientific instruments have undergone huge technological advances.

As a consequence, the types of questions that science tries to answer have also changed. We want answers to questions about phenomena that can no longer be observed by a simple view nor with the help of optical instruments.

What is supercomputing?



Supercomputing enables scientific experiments by simulating *in silico* the behaviour of the object under study. Bringing together all the knowledge about the object in a computer simulation and experimenting with it can reduce costs, avoid suffering and enable experiments to be conducted that could not be performed in the real world, because they would be too expensive, too dangerous or simply impossible.

Supercomputers are also needed to analyze large amounts of data, such as those provided by modern scientific instruments (e.g. particle accelerators, large telescopes, interferometers or genome sequencing platforms) and the ever-growing number of devices that make up the Internet of Things.

In recent years, it has proven to be an essential ally of artificial intelligence, since its large processing capacity makes it possible to train algorithms and draw conclusions from massive amounts of data.

High performance computing has already become a great accelerator of science and engineering. It is used, and is increasingly being considered essential by the majority of scientific disciplines.

Higher performance for bigger challenges

A supercomputer is a computer with a high level of performance compared to a personal or business computer. The performance of a supercomputer is commonly measured in floating-point operations per second (flops). MareNostrum 5 will have a peak performance of around 200 Petaflops ($200x10^{15}$ or 200,000,000,000,000,000 flops). The supercomputing sector is working hard to create the first exascale supercomputers, which will perform 10^{18} floating-point operations per second, which is expected to usher in a new era.

Research with MareNostrum 4

MareNostrum 4 provided **1,173 million processing hours** in 2019. 80% of these hours were used by researchers who gained access to the supercomputer through the **PRACE** (Partnership for Advanced Computing in Europe) infrastructure network or the Spanish Supercomputing Network (**RES**). The computing hours allocated to PRACE and RES are valued at **8.5 million euros**.

Also in 2019, BSC's Board of Trustees engaged its high performance computing services in three strategic projects: **Gaia**, which is creating the most precise map of the Milky Way, the European Genome-Phenome Archive (**EGA**) and the Spanish contribution to data processing for the **Large Hadron Collider** at CERN.

Below we set out a selection of the wide variety of projects from outside of BSC that have used MareNostrum 4 this year and their principal investigator:

esa

Gaia is the ambitious mission to create a three-dimensional map of our galaxy, the Milky Way, and, in the process, reveal its composition, formation and evolution. It will provide unprecedented radial speed and positional measures with the necessary precision to produce a stereoscopic and cinematic census of approximately one billion stars.

MareNostrum 4 is producing project simulations and BSC is storing the data.

TÉCNICO LISBOA

Luis Silva

"*PULSAR: Plasma physics of ultra high fields in neutron stars*".

INSTITUTO DE ASTROFISICA DE CANARIAS · IAC

Javier Trujillo

"*The forward scattering polarization of the Ca I 422.7 nm resonance line as a tool to probe the solar chromosphere*".

DIPC

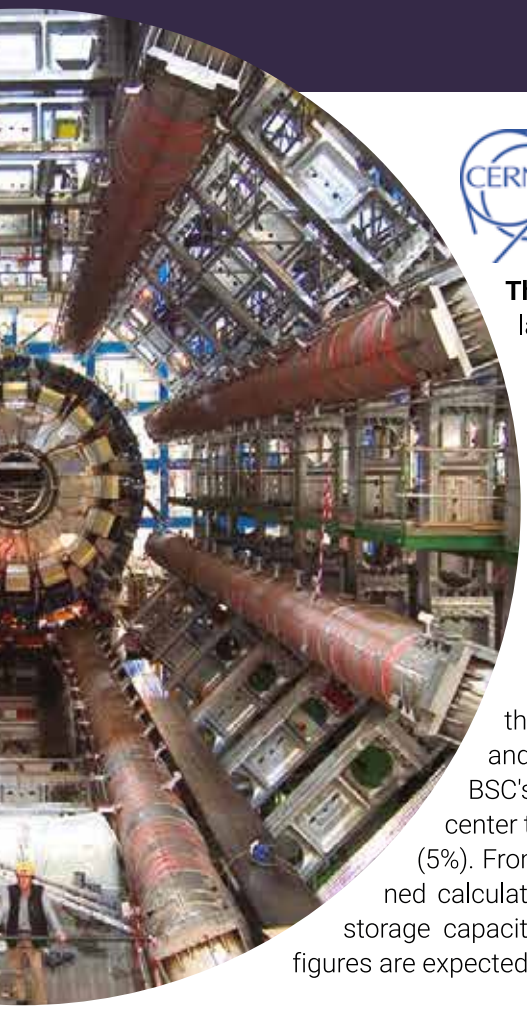
Jens Stücker

"*The first reliable simulations of Warm Dark Matter*".

UAM Universidad Autónoma de Madrid

Gustavo Yepes

"*The MareNostrum numerical cosmology project: grand challenge simulations of structure formation in the Universe*".



The Large Hadron Collider is the largest particle accelerator in the world and is located at the European Laboratory for Particle Physics (CERN). Its main objective is to expand the frontiers of knowledge of physics by making protons collide at speeds close to the speed of light and extract information from this.

Forty-two countries are providing the necessary resources to process and analyze the data generated. BSC's Board of Trustees has asked the center to handle the Spanish contribution (5%). From 2020, BSC will provide a sustained calculation capacity of 500 teraflopsand storage capacity of 7 petabytes. Both of these figures are expected to double in the 2023-25 period.

UNIVERSIDAD D SEVILLA 1505 Linda A Zotti

"Incorporating single proteins in solid-state devices".

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH Manel Soria

"On the aeroacoustic noise radiated by a cavity at low Mach number ".

UNIVERSITY OF CAMBRIDGE Michele Vendruscolo

"Targeting alzheimer's-associated amyloid-beta using small molecules".

Technische Universität München Ville R. I. Kaila

"Functional dynamics of the mammalian respiratory complex I".

CRG Centre for Genomic Regulation EMBL-EBI

The **European Genome-Phenome Archive** (**EGA**) is the largest service in the world for the management of biological data from biomedical research made available to authorized scientists. It has 1.8 million files that add up to 7.2 petabytes of data. Its computational needs are growing exponentially. BSC has been commissioned with providing the services of storage (currently 10 petabytes and 24 petabytes in 2024), computing, big data technologies, cloud computing and broadband to transfer data through the internet.

UIB Universitat de les Illes Balears Victor Homar

"Predictability of Mediterranean severe weather: Optimizing the benefits from ensemble data assimilation of EUMETSAT RS-AMVs in coastal extreme events".

upf. Universitat Pompeu Fabra Barcelona Marta Reynal-Querol

"Computing pixel based socioeconomic measures to analyze economic development".

Research News

Rosa Maria Badia receives the Dona TIC and Euro-Par awards

Rosa M. Badia, the leader of the Workflows and distributed computing group, has been awarded the Dona TIC Prize from the Catalan Government in the Academic/Researcher category. Badia was also awarded the Euro-Par Achievement Award 2019, which is given to researchers who have had an extraordinary impact on parallel processing.

Asun Lera is appointed as a member of the Horizon Europe mission board

The senior researcher in the Earth system services group, Asun Lera St.Clair, has been chosen by the European Commission to be a member of the mission board for "Adaptation to Climate Change including Societal Transformation". Mission boards are high-level initiatives to find solutions to our society's major challenges and only three Spanish researchers are members.

Mateo Valero joins the Academia de Ingeniería de México and receives the CénitS Prize

BSC's director, Mateo Valero, has been appointed as a member of the Academia de Ingeniería de México due to his intense academic collaboration with that country. Valero has also received the research excellence prize from CénitS (Centro Extremeño de iNvestigación, Innovación Tecnológica y Supercomputación).

Carlos Pérez García-Pando is appointed as an ICREA professor

Carlos Pérez García-Pando, the head of the Atmospheric composition group and AXA Chair on Sand and Dust Storms, has been appointed as an ICREA professor. The researcher currently has a European Research Council (ERC) Consolidator Grant for his project FRAGMENT: "Frontiers in Dust Mineralogical Composition and its Effects upon Climate".

Marta Melé receives a L'Oréal-Unesco research prize

The leader of the Transcriptomics and functional genomics group, Marta Melé, has received a prize from the L'Oréal-Unesco Program for Women in Science for her work in analysing variations in genes between individuals and their implications for diseases such as breast cancer. The research looks for specific biomarkers and therapeutic targets to prevent this cancer.

Nataša Pržulj is appointed as a member of the Serbian Academy

Nataša Pržulj, the leader of the Biology integrating computational networks group, has been selected as a member of the Serbian Royal Academy of Scientist and Artists. Pržulj, who joined BSC at the beginning of 2019, is an ICREA professor and has been awarded a European Research Council (ERC) Consolidator Grant for her project "Integrated Connectedness for a New Representation of Biology".

Toni Gabaldón is the vice-president of the new Spanish Youth Academy

The leader of the Comparative genomics group, Toni Gabaldón, has been appointed as the vice-president of the Spanish Youth Academy, which was created in May 2019 to give visibility to young scientists in the experimental sciences. Gabaldón, who was also one of the founding members of this Academy, joined BSC in 2019 and is an ICREA professor.

A major leap in research into open-source hardware

BSC has always been greatly involved in research into hardware proposals for HPC that eliminate European dependence on technology created on other continents. In 2011, the Mont-Blanc project was the first to develop ARM-based clusters. EPI (European Processor Initiative) project coordinates and develops accelerators based on RISC-V. Miguel Eduardo Litvin was brought in to manage this task in 2019. Now, two new major projects have been created to further develop open-source hardware architectures based on RISC-V, MIPS and Power 9: MEEP and LOCA. A leap in scale.

MEEP

The MareNostrum 5 experimental platform

The MareNostrum Exascale Experimental Platform (MEEP) is a three-year project (2020-2023) involving MareNostrum 5 to create a new FPGA emulation platform for next-generation European exascale supercomputers.

It is the only pre-exascale deployment with an experimental component. It has a budget of 10.3 million euros, partners in Croatia and Turkey and will be have a team of 35 people managed by Peter Hsu with a PhD in Computer Science from the University of Illinois at Urbana-Champaign, who joined BSC in 2019.

Peter Hsu, Mateo Valero and John D. Davis

The platform will be used to perform full stack research and codesign hardware and software for traditional and emerging HPC applications. It will be capable of emulating portions of exascale platforms and enable software development.

LOCA

An open-source hardware laboratory

The European Laboratory for Open Computer Architecture (LOCA) is an initiative to develop European chips based on open-source hardware and software suitable for HPC, the Internet of Things and other computing fields. Its first mission is to create components for exascale supercomputers.

LOCA is a global center for research, training, education and collaboration between industry and academia. It will provide the necessary infrastructure for a completely open-source ecosystem, from applications to CPU and accelerators. It will be supported by industrial partners and funding and will be managed by John D. Davis, who holds a PhD in Electronic Engineering (Computer Architecture) from Stanford University and who joined BSC in 2019.

Lagarto, the first chip with an open-source instruction set architecture developed in Spain

Marco Antonio Martínez Salinas, CIC-IPN's Director, and Mateo Valero

Lagarto, built with 65-nanometer TSMC transistors, is the first chip with an open-source instruction set architecture (ISA) developed at BSC. Its initial design was produced by the Centro de Investigación en Computación at IPN in Mexico. Subsequently, a team of 30 researchers from BSC, CIC-IPN, CSIC's Centro Nacional de Microelectrónica and the Universitat Politècnica de Catalunya developed an open-source RISC-V chip that worked at the first attempt. The project will be used as the basis for new developments with open-source ISAs, such as those that will be produced by MEEP, LOCA and the DRAC project.

Featured publications

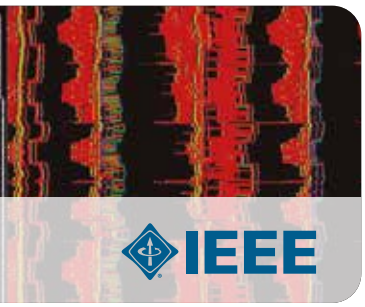
The multidisciplinary nature of BSC is evident in the wide range of disciplines in which its scientists publish articles. Below is a selection of the most outstanding publications in 2019.

A Hardware Runtime for Task-Based Programming Models

This paper presents Picos++, a hardware runtime to accelerate critical runtime functions in task-based programming models such as OpenMP and OmpSs, task dependence analysis, nested task support, and heterogeneous task scheduling. As a proof-of-concept, Picos++ has been implemented in a FPGA SoC achieving up to 7.6x speedup and 90 percent energy savings, when using 4 threads and up to 4 HwAccs. Xubin Tan, Jaume Bosch, Carlos Álvarez, Daniel Jiménez-González, Eduard Ayguadé and Mateo Valero. IEEE Transactions on Parallel and Distributed Systems, vol. 30, no. 9, pp.1932-1946 (2019).

Increasing the Reliability of Software Timing Analysis for Cache-Based Processors

This paper presents mechanisms to intercept relevant Cache Conflict Placements (CCP) when analyzing cache memories in high-performance real-time systems. The proposals are exhaustively evaluated on well-known benchmarks and a railway case study, on top of an accurate simulator and a concrete RTL implementation. Suzana Milutinovic, Enrico Mezzetti, Jaume Abella and Francisco J. Cazorla. IEEE Transactions on Computers, vol. 68 no. 6, pp.836-851 (2019).

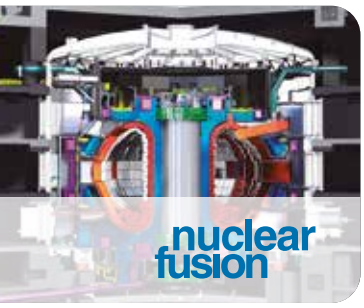


Sampled Simulation of Task-Based Programs

This paper presents TaskPoint, a sampled simulation technique for dynamically scheduled task-based programs. TaskPoint is the first technique combining sampled simulation and analytical modeling and provides a new way to trade off simulation speed and accuracy. Compared to detailed simulation, TaskPoint accelerates architectural simulation with 8 simulated threads by an average factor of 220x at an average error of 0.5% and a maximum error of 7.9%. Thomas Grass, Trevor E. Carlson, Alejandro Rico, Germán Ceballos, Eduard Ayguadé, Marc Casas and Miquel Moretó. IEEE Transactions on Computers, vol. 68, no. 2, pp.255-269 (2019)

Volcanic ash forecast using ensemble-based data assimilation: the Ensemble Transform Kalman Filter coupled with FALL3D-7.2 model (ETKF-FALL3D, version 1.0)

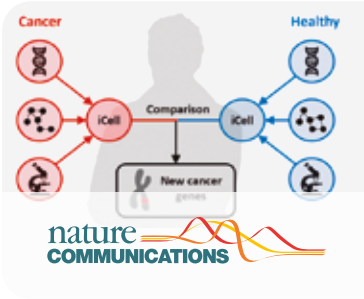
The authors present an ensemble-based data assimilation and forecast system for volcanic ash dispersal and deposition aimed at reducing uncertainties related to eruption source parameters. The FALL3D atmospheric dispersal model is used to drive ensemble members trough an Ensemble Transform Kalman Filter (ETKF) data assimilation technique. The process gives both better estimations of ash concentration and time-dependent optimized values of eruption source parameters and leads to a better analysis and forecast of the 3D ash concentrations. Soledad Osores, Juan Ruiz, Arnau Folch and Estela Collini. Geoscientific Model Development.



Shear Alfvén wave continuum spectrum with bifurcated helical core equilibria

The authors present a modelling system of the radial structure of the continuum spectrum of shear Alfvén and Alfvén-acoustic waves in the beta-induced Alfvén eigenmode (BAE) frequency range for tokamak plasmas. The modelling (including 3D effects) correctly reproduces the phenomenon of continuum frequency splitting and provides a possible solution for the differences of few kHz in frequency splitting, which remained unexplained with the 2D kinetic calculations and is in agreement with the experimentally observed low-frequency modes. Allah Rakha, Philipp Lauber, Mervi Johanna Mantsinen and Donald A Spong. Nuclear Fusion. 9 Sept 2019.

Towards a data-integrated cell

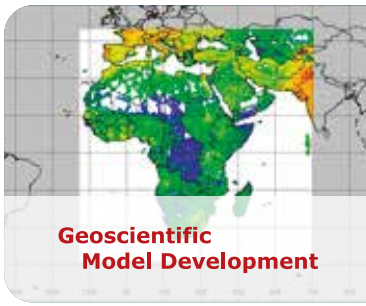


The paper presents a new artificial intelligence-based computational method that accelerates the identification of new genes related to cancer. The method is based on the machine learning technique Non-negative matrix tri-factorization. It fuses large amounts of omic data from three tissue-specific molecular interaction networks (protein-protein interaction, gene co-expression, and genetic interaction networks) and recreates them in a computational prototype: the data-Integrated Cell. Noël Malod-Dognin, Julia Petschnigg, Sam F. L. Windels, Janez Povh, Harry Hemingway, Robin Ketteler and Natasa Przulj. Nature Communications. 18 Feb 2019.

Evolutionary and functional patterns of shared gene neighbourhood in fungi

This paper sheds light on how genomes of eukaryotic organisms organize groups of genes in accordance with their functions. The scientists developed an algorithm capable of identifying genes that remain close to each other during evolution. The study, based on large public databases, provided evidence that most shared gene blocks in distant species emerged from independent re-assembly events rather than from genetic transfer from one species to another (horizontal transfer). Marina Marcet-Houben and Toni Gabaldón. Nature Microbiology 4. 16 Sept 2019.

HERMESv3, a stand-alone multi-scale atmospheric emission modelling framework – Part 1: global and regional module



Atmospheric emission inventories, which describe the amounts of pollutants released into the air by different sources and for specific regions, are an essential input to numerical models that estimate air quality. This work presents the High-Elective Resolution Modelling Emission System version 3 (HERMESv3), a Python-based and parallel modelling framework that allows adapting existing global and regional emission inventories to the input requirements of air quality models in a flexible and transparent way. Marc Guevara, Carles Tena, Manuel Porquet, Oriol Jorba and Carlos Pérez García-Pando. Geoscientific Model Development. 14 May 2019.

The Weather Roulette: a game to communicate the usefulness of probabilistic climate predictions

The paper applies a game to communicate the usefulness of climate predictions to users. The game uses the Weather Roulette conceptual framework to present skill metrics commonly used by the scientific community with metrics more easily understood by the business sector, showing that in skillful areas economic benefits are obtained in the long term. This approach could provide the basis for a better integration of knowledge about climate anomalies into operational and managerial processes. Marta Terrado, Llorenç Lledó, Dragana Bojovic, Asun Lera St. Clair and Albert Soret. American Metereological Society. 24 Oct 2019.

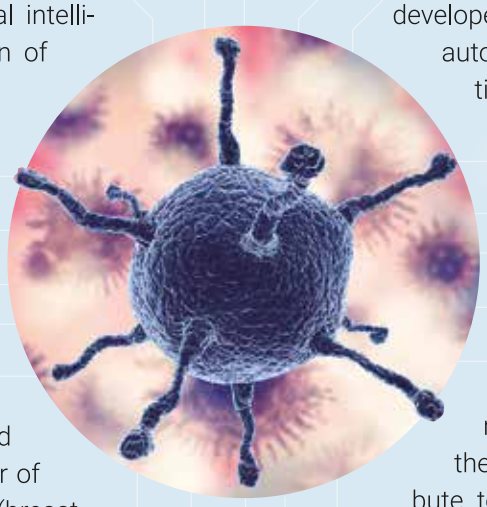
Applied Research & Innovation

Applied artificial intelligence

A method based on machine learning techniques makes it possible to identify new genes connected with cancer

The researcher Nataša Pržulj has led the creation of a new computational method based on artificial intelligence that accelerates the identification of new genes connected with cancer.

Pržulj used machine learning techniques to relate large quantities of omics data and recreated them in a computational prototype. Her method merged three specific molecular interaction networks in tissues: protein-protein interaction, gene coexpression and genetic interaction networks. The authors of the article published in *Nature Communications* used this method to reconstruct cells from four of the most common types of cancer (breast, prostate, lung and colon) and in all of them it proved to be useful in locating new genes connected with these diseases. The method pinpointed 63 genes and biological validation process confirmed that at least 36 of them contribute to irregular cell growth.



A prototype based on natural language processing makes it possible to automatically classify tumor mutations

The Natural language processing group has developed an application prototype to automatically classify genetic mutations from tumoral processes based on clinical reports. The prototype has been automatically trained with clinical articles from the Memorial Sloan Kettering Cancer Center in New York.

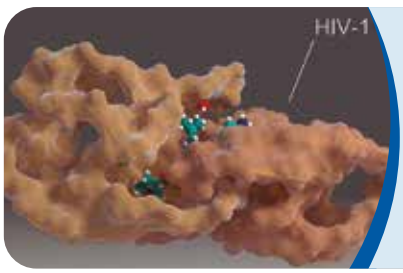
Its purpose is to help pathologists distinguish among the many different genetic mutations their patients have that may contribute to tumor growth. This task has hitherto been conducted manually and is very time-consuming, since the genetic sequencing of a cancer tumor usually presents thousands of mutations and the vast majority are neutral in terms of the disease.

Collaboration with Almirall to find treatments for dermatological diseases

BSC has established collaborations with the pharmaceutical company Almirall to identify new therapeutic targets to treat dermatological diseases and use the center's technologies to find initial drugs to treat them. Targets are identified by analyzing the vast amounts of omics data and drug research is conducted with the PELE software.

Supercomputing to help with research into an HIV vaccine with Irsicaixa

BSC is collaborating with IrsiCaixa on a project to design an effective vaccine for HIV. BSC uses MareNostrum 4 supercomputer and the PELE software for computerized generation and testing of antibodies that can help combat the virus and that subsequently, in the laboratory, will help identify the most vulnerable parts.



Applied artificial intelligence

FC Barcelona and BSC create an intelligent system to manage the Espai Barça facilities

FC Barcelona and BSC are working on a project that applies the Internet of Things and artificial intelligence to create an intelligent system that will analyze and predict the movement of people inside and outside the football club's facilities (Camp Nou, Palau Blaugrana, Estadi Johan Cruyff, Campus Barça, Museum and Barça Store, etc.) and will provide real-time information to manage the club's activities and the Espai Barça facilities.

The project is intended to model people's movements through its facilities, especially Camp Nou and Espai Barça, on both match and non-match days. The intelligent system will monitor the facilities, gathering data from sensors spread around the area, and will create a computerized simulation that will reproduce people's usual movements. This simulation will be used to see what would be the result of different responses the club could give to particular situations, whether planned changes or unforeseen incidents that take place in real time.

The results of the simulations will allow FC Barcelona to decide on the response it should take with a high level of relevant information. In order to obtain the results, different cutting-edge technologies will be combined, such as artificial intelligence, the Internet of Things, edge computing and supercomputing.

Copernicus commissions BSC to code data and create products for decade-long predictions

Copernicus, the European Union's Earth Observation Programme, has started up a project to provide the infrastructure, insurance, agriculture and energy sectors with climate predictions ten years away that may be useful in planning their medium-term future. BSC will lead the coding and user products side and will carry out a case study on its application in the agricultural sector.

High performance computing for agriculture and livestock

BSC is collaborating on the CYBELE innovation project, which has received 14 million euros in funding from the EC, and is intended to demonstrate how the convergence of HPC, Big Data, cloud computing and the Internet of Things can revolutionize agriculture, reducing the scarcity of staples and increasing food production, thus providing social, economic and environmental benefits.

Applied artificial intelligence

BSC is collaborating on a European Space Agency project to use generative artificial intelligence in space

BSC will research the use of generative artificial intelligence to generate and augment synthetic datasets from observations in space. In the DeepLIM project, funded by the European Space Agency (ESA), BSC will contribute the emerging technologies in the MareNostrum 4 supercomputer and optimize the performance of deep learning libraries involved in the generation of simulated data.

A project to train machines to describe items of European cultural heritage with machine learning

BSC is generating descriptions of items of European cultural heritage with deep learning methods and other semantic and statistical techniques to generate automatic explanations of cultural images with natural language. The project is called "Sant Jordi en moto" [St George on a bike], due to the response a machine gave when it saw an image of St George on a horse killing the dragon, which it confused with a dog.

New Arm-BSC Center of Excellence

BSC and Arm Research have signed an agreement to set up the Arm-BSC Center of Excellence, which will expand the interaction and collaboration between the two parties in fields ranging from HPC to built-in computing in cars, and from scientific computing to real-time systems. The collaboration will cover computer architecture, runtime systems, programming models for performance portability and other areas.

Prize for the BSC-Repsol joint center

The Catalan Government and the Catalan Foundation for Research and Innovation (FCRI) have awarded the Repsol-BSC Research Center the national prize for public-private partnership in R&D. This recognition highlights the long-standing collaboration between the two centers in research into subsoil visualization technologies.

Meet us in person

A program to teach the public about supercomputing

"Computing with you" is BSC's new educational program to teach the public about supercomputers and show them what they could be capable of doing by using this scientific infrastructure that is of great value for generating knowledge.

The program has different action areas (guided tours, visits to primary and secondary schools, courses for secondary and baccalaureate students, open doors days, teacher training, participation in science education activities, etc.). In all of them, BSC seeks to collaborate on inspiring new scientific vocations and narrowing the gender gap in technological areas.

The first completed year of Crazy about Supercomputing

Twenty-five baccalaureate students complete the first promotion of Crazy about Supercomputing program jointly organized by BSC and the Fundació Catalunya La Pedrera. The students have attended 19 sessions held by BSC researchers, which included subjects as varied as introduction to supercomputing, parallel programming, simulations, supercomputing mathematics, artificial intelligence, smart cities, personalized medicine, climate change and quantum computing.

5,500 primary school students enjoy "We are young women researchers"

More than 5,500 primary school students have enjoyed the "We are young women researchers" project, co-funded by Barcelona City Council, during 2019. This program introduces children to the world of supercomputing and computational thought with activities designed to be especially attractive to girls.

New record of visitors to MareNostrum

2019 closed with an absolute record number of both visits and visitors to the MareNostrum supercomputer surpassing 19,800 people, which represents an increase of more than 5,700 compared to the previous year.

MareNostrum collaborates on a symphonic work inspired by artificial intelligence

MareNostrum supercomputer has collaborated in *Human Brother*, a piece of music for soprano and orchestra by the Barcelona composer Ferran Cruixent, which Orquestra Simfònica de Barcelona i Nacional de Catalunya (OBC) premiered at the Barcelona Auditorium. Cruixent explores the role of technology in creation, and on this occasion he examines the relationship between AI and music. Artur García Sáez, a BSC researcher with a PhD in physics, was his ally and guide on this adventure. Together they provided MareNostrum with audios of many different symphonic compositions by Cruixent so that the supercomputer could use them as a reference and create its own works with a machine learning algorithm.

BSC gratefully acknowledges the support of:

BSC is a public consortium made up of:

Generalitat de Catalunya
Departament d'Empresa i Coneixement

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

EUROPEAN UNION
European Regional Development Fund

EXCELENCIA SEVERO OCHOA

Infraestructuras Científicas y Técnicas Singulares

HR EXCELLENCE IN RESEARCH

Jordi Girona, 31
Torre Girona Building
08034 Barcelona (Spain)

info@bsc.es
www.bsc.es