interview with

Mateo Valero

Director
Barcelona Supercomputing Center – Centro Nacional
de Supercomputación (BSC-CNS)
Barcelona - Spain

Born in Alfamen (Zaragoza, Spain), Mateo Varela already excelled at maths at a very young age. He completed a degree in telecommunications engineering in Madrid in 1974 and went on to study a PhD at the Polytechnic University of Catalonia (UPC), where he has been professor of Computer Architecture since 1983.

Mateo's «school» of hardware design has pioneered multi-processor system development in Spain, and he has received numerous national and international accolades, including the Eckert-Mauchly Award in 2007 -which is regarded as the Nobel Prize of computers- and the Association of Computing Machinery's Distinguished Service Award in 2012.

The year 1984 marked a turning point for the UPC, as it was then that the Computer Science department acquired a 64-processor transputer. This device can be used to develop software for parallel machines working with multiple processors, such as supercomputers.

This research has been carried out not only at the UPC but also at other facilities that Mateo has helped to create, including the European Parallelism Centre in Barcelona (CEPBA) and the Barcelona Supercomputing Centre (BSC), establishing respectively in 1995 and 2005. The BSC, which coordinates the Spanish Super-computing Network, has led the development of supercomputing in Spain. It combines supercomputing services to scientists across the country and Europe with research in the fields of computer science and the application of parallel computing to other disciplines, such as bioinformatics, engineering and earth sciences.

Mateo has been ranked among the 25 most influential Spanish researchers every year from 2007 to 2014, as listed by El Mundo newspaper.

www.bsc.es/cv-mateo

«Nobody expected Spain to build such a powerful machine»

Mateo Varela is the director of the BSC, Spain's leading institution in the area of supercomputing, which specializes in high performance computing (HPC). His job involves coordinating 300 experts and R&D professionals organized into four main research fields: Computer Science, Life Science, Earth Science and Computer Applications in Science and Engineering. The BSC was established with the declared purpose to develop and manage technology that facilitates scientific progress. Since its inception, it has also engaged in partnerships with industry and the private sector.

How did the Barcelona Supercomputing Center project come into being? Who is behind it? What similar initiatives have been undertaken in other countries?

It was started by a group of passionate computer architects from the Catalonia Polytechnic University School of Computer Science, including myself, who began researching parallel computing in the early 1980s. By 1990 we had set up the first parallel architecture research centre, the CEPBA (Barcelona Parallelism Centre), together with the central government department and the Catalan regional government. This was joined by another centre in 1995, the CIRI (CEPBA-IBM Research Institute), where IBM became a partner.

After almost 15 years of research and experiments with different architectures and programming models, in 2004 we had a very ambitious project, a dream actually: to build and operate one of the fastest supercomputers in the world, something that required government help and support. That was how the Barcelona Supercomputing Center was created. The whole world was staggered because our supercomputer became number one in Europe and number four globally. Nobody expected Spain to build such a powerful machine. When we did, that propelled us straight into the world's first division of supercomputing. And the quality of our research projects has made sure we stayed there.

Nobody expected Spain to build such a powerful machine. When we did, that propelled us straight into the world's first division of supercomputing

What is the BSC's basic infrastructure?

In terms of infrastructure -meaning machinery- the MareNostrum supercomputer remains our central pillar. Its third version was installed last year. Our data repositories are also currently a key element of machinery. But the BSC would not be the BSC if, in addition to infrastructure, it did not have a staff of more than 300 researchers from a range of disciplines, who use supercomputing to do science and generate wealth.

What are the main features of MareNostrum that enable it to remain one of the world's top-ranking supercomputers?
MareNostrum has nearly 50,000 processors, 96 terabytes of primary storage and 2 petabytes of hard disk storage space. The key to its operation lies in making all these elements work together and communicate quickly between them. To do that, we have an optic fibre network almost 80 km in length that links all motherboards and components together.

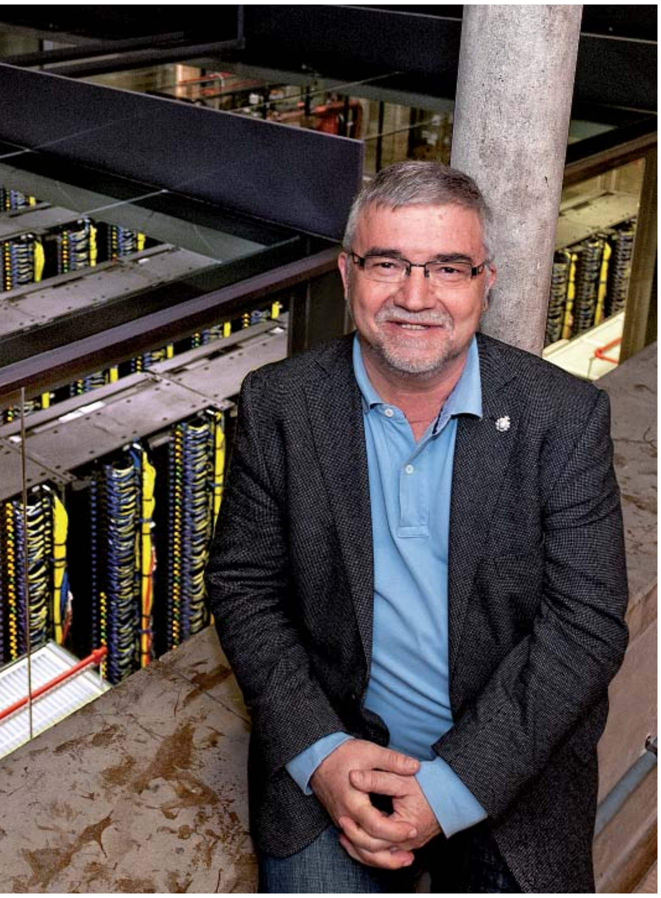
If we had the fastest supercomputer in the world but it was unable to meet the needs of our users, we would not be doing our job properly, which is to serve science

In 2004, MareNostrum reached number four among the world's 500 top supercomputers, and currently it is one of the leading 20 or 30. What does it take to rank with the top few?
Basically, it has to do with money. But, as with other things in life, money is not everything there is to supercomputing. MareNostrum3 ranked 29th in the world when it was installed. With the same investment, we could have built a higher-ranking supercomputer, possibly among the top 10. Supercomputers are ranked by running a very specific software called Linpack on them. The machine that runs it in less time is the fastest and given an accordingly higher ranking. But Linpack is very different from the programs researchers use. MareNostrum performs assignments for scientists across Europe, who work in widely different disciplines and organisations. This means we need to have what we call a «general architecture», in other words a supercomputer that can cater for users with a wide variety of software to run. If we had the fastest supercomputer in the world but it was unable to meet the needs of our users, we would not be doing our job properly, which is to serve science.

All the experts who control and manage MareNostrum are part of the Barcelona Supercomputing Center. This means you have a vast infrastructure at your disposal. Are you also using outsourced cloud storage or cloud computing systems?
No. The Barcelona Supercomputing Center has its own data storage system. In fact, we provide data services to other research centres and international projects.

Talking about cloud computing, how secure is this technology?
The cloud has some obvious benefits. For instance, it enables making continuous backups of your files, which you can then access from any computer or mobile device. It also makes it very easy to share content with other users and to expand your computing power. Some aspects are really good. But it is also the case that privacy is never one hundred per cent assured in the cloud. So you need to look at your individual needs and decide whether you want to have a cloud backup of certain files or whether you prefer to have a one hundred per cent guarantee of their privacy.

In what way is the BSC dependent on its optic fibre infrastructure?
Internally, optic fibre is what enables fast communications between a supercomputer's components so that they can all work together. Externally, it makes it possible to share data remotely, and allows an infrastructure's users to have access to it. Universities and research institutions are interconnected by optic fibre networks.



Given the importance of this supercomputer, maximum security measures are needed to protect it from any acts of sabotage, attacks and natural disasters. In addition to these, are there any specific contingency plans in place against potential threats?
We have security systems to prevent vandalism and attacks against the infrastructure, surveillance to prevent intrusions and misuse, and agreements with other institutions to continue providing certain services in disaster scenarios.

IBM was MareNostrum's manufacturer of choice from the start, and the BSC continues to use IBM's technology. What is the nature of BSC's relation with IBM in today's innovative, highly sophisticated environment?
IBM built the supercomputer and Intel provided the microchips. We have worked on joint research projects with both these companies since the early days of the BSC (and also with Microsoft, Nvidia, Repsol, Iberdrola, and others). Developing new wealth-creating technologies in partnership with business entities is a defining trait of the BSC. Our partnership with IBM goes back to 1998. We always have joint research projects ongoing. At some points we have had up to 40 researchers working together.

Where the operating system is concerned, you seem to favour Linux , which is free software. What is the reason for this choice? Could you mention a few of its benefits?
MareNostrum was the first supercomputer to use Linux. For us, the chief benefit of using it is that we can provide service to the largest number of users.

If we were to connect thousands of computers over the internet, we would have the same number of processors and the same storage capacity as MareNostrum. Therefore, by making them all work together, we would be able to process the same number of operations as MareNostrum. To what extent is this an accurate depiction of reality?
Fairly accurate, except for the fact that the run time of the operations would be much longer. Supercomputers have fast connections linking all their components together, and that is a crucial aspect. The internet and home connections have a much lower latency, so they simply cannot provide that speed of execution.
It is important to bear in mind that processors are constantly sending and receiving data from one another when they work together. Slowing down those transmissions significantly lengthens run time because processors have to remain idle until they receive the input for the next calculation.

So you need to look at your individual needs and decide whether you want to have a cloud backup of certain files or whether you prefer to have a one hundred per cent guarantee of their privacy

Do you think it will become possible to tap into the power of the thousands of processors inside mobile phones?
Mobile telephone technology is certainly appealing. It is more energy-efficient and less costly. We are currently working on projects to gauge the extent to which it can be of value to build supercomputers in the future. Project Mont-Blanc, which is led by the BSC, is a good example. It has garnered a good deal of international interest.

The MareNostrum is currently made up of several thousand processors, with a processing speed in excess of 1 petaflop/s,

Making personalised medicine a reality is definitely the most exciting project. Doctors will no longer treat patients based on the names of the illnesses they suffer, but on the specific changes taking place in their bodies

i.e. more than a trillion operations per second. What kind of projects and domains are best suited to use its full potential?
We aim to have a machine that caters efficiently to the needs of most users. That is why we chose SandyBridge chips to build our computer over others which may perform faster in certain tasks but are less efficient for other users. MareNostrum performs tasks from all fields of science, from biology to astrophysics and meteorology, as well as a broad range of engineering-related assignments.

Which are the most innovative projects now ongoing in the medical area? When do you think they will be in actual use in public and private hospitals?
Making personalised medicine a reality is definitely the most exciting project. Doctors will no longer treat patients based on the names of the illnesses they suffer, but on the specific changes taking place in their bodies. Cancer is the paradigm of this. Today we know there are as many different types of cancer as there are cancer patients. This realisation paves the way for the end of blanket treatments -which are highly aggressive and not always satisfactory-and the beginning of personalised treatments. This kind of medicine is largely based on genomic analysis and therefore heavily dependent on technology, including hardware and data processing. We are putting a lot of effort into this area and, as far as technology is concerned, I believe much of the road towards personalised medicine has already been covered.

The genomic revolution has brought on an uncontainable avalanche of biological data which is putting storage capacity to the test, despite recent advances. How has the ELIXIR programme structured European cooperation in this area? How does the BSC contribute to the project as a member of the consortium?
We have the largest genotype and phenotype database in Europe (El European Genome Archive, EGA) hosted at the BSC. That currently includes data from one hundred thousand patients, and it is expected to grow extremely rapidly over the coming years. Our role involves providing host services, some secure protocols and making sure scientists can access the databases quickly and securely.

Turning to the issue of seafloor scanning for hydrocarbon exploration (the Repsol project), have any new reservoirs been identified in areas not previously considered likely to contain oil or gas?
The aim of Project Kaleidoscope, which was carried out in partnership with Repsol, was to facilitate hydrocarbon exploration in particularly challenging areas, such as the Gulf of Mexico, where the earth's crust is separated from the sea water by salt layers up to two kilometres thick, posing a great hindrance to exploration operations. The technology we developed proved a success. It significantly improved the company's accuracy ratio, gave them a competitive edge, increased their operating profit and opened up new markets in countries that were previously closed to them. It was also a success for us. We are already working on a new, even more finely honed version, and we have already signed up other contacts. All this is proof of the BSC's interest in being a technology partner for the industry.

Studying climate change involves using climate models and processing large volumes of data. What role has the BSC played in this major scientific challenge of the 21st century?
We work with climate models of great complexity and we are world leaders in several fields of research - the study of desert sand movements, for instance. We have conducted studies of future climatic scenarios for government agencies. Unfortunately, the key to preventing climate change will not lie in scientific research, but in how resolved societies -and particularly governments- are to address the issue.

Actuarial science, which is a vital part of the insurance industry, involves using observed and projected statistics to price a range of insurance products, including Life, Motor, Home and Health, to name but the most common categories. Have you ever considered working with the insurance and reinsurance industry?
So far we have not had the opportunity to work with any companies in this area, though it is certainly something we would like to do. Insurance companies are obviously potential or actual users of supercomputing, to the extent

that large calculations help them foresee individual and collective risks, as well as the impact of disasters.

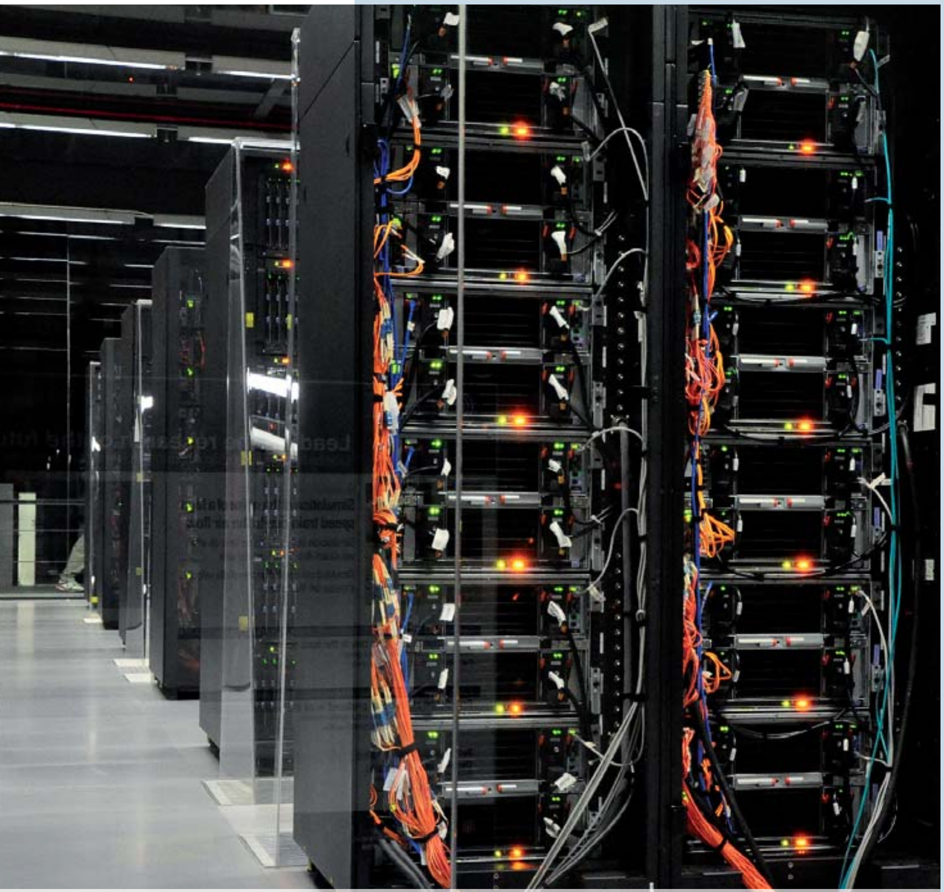
Natural disaster models bring together actuaries, earth science specialists, mathematicians and finance experts in handling the stochastic models used to estimate the likelihood of earthquakes and tropical cyclones -to name the costliest- and their impact in terms of insurance and reinsurance. Do you work with any commercial modelling companies or international researchers in this area?

We do work with businesses specialising in risk prediction. Specifically, we had a successful partnership with a consultancy firm called Amphos21. The project involved modelling what would become the first nuclear waste dump to be built in Europe -in Sweden, actually- to predict leaks over a time horizon of one hundred thousand million years. This is no doubt an interesting field of study, and one that requires supercomputing to ensure maximum simulation accuracy, because large quantities of data are needed.

At the moment, we are looking for clients to start using a seismic wave propagation simulator, accurate to within metres, which has a number of potential applications. For instance, we could have a clear map of the areas and buildings most and least affected by an earthquake within two or three hours of its detection by a seismological observatory. That map would be extremely valuable to insurance companies, because it would allow them to quickly plan how best to deploy their appraisers on the ground.

In terms of preventive action, it could also be used to examine several seismic scenarios and assign individual risk levels to different areas. If there is a known likelihood that a geological fault will break in a given area, we would be able to foresee the consequences of the break happening at different spots. This is a valuable input for urban planning in towns at risk of having earthquakes, as well as for insurance companies.

At the moment, we are looking for clients to start using a seismic wave propagation simulator, accurate to within metres, which has a number of potential applications



Detail of the MareNostrum supercomputer

What is special about the BSC headquarters?
It is under construction! We have grown a lot faster than expected and we are currently scattered across several buildings in the north campus, but the new building is now under way. I have to thank Repsol for their generous donation, which will allow the works to be completed.

How many people are currently employed there?
Almost 400 people from a wide range of disciplines, all of them seriously committed to innovation.

What professional backgrounds prevail among the researchers and engineers?
They come from extremely varied backgrounds. From PhD students to senior professionals with impressive qualifications and experience. From biologists to nuclear physicists. The common denominator is computing.

What is the BSC's annual budget?
In 2013, we had 24 million euros.

How is BSC funded?
Our recurrent budget is about six million euros, which is government provided. The rest comes from competitive revenue from our involvement in research projects and business partnerships.

How can the BSC be contacted to appraise a project's feasibility?
If it is an ongoing scientific project, the researchers should contact one of the access committees of the Spanish Supercomputing Network or PRACE, which are the institutions that control access to public supercomputer services. If it is a company or an organisation interested in undertaking research in partnership with the BSC, they can contact us through any of our research departments: Computer Science, Life Science, Earth Science and Science and Engineering Applications. All contact details can be found on www.bsc.es.

How much does an hour of MareNostrum calculation cost?
MareNostrum is a computer devoted to science and research. Anyone wishing to use it has to submit a scientific project, which is assessed by an expert committee, or else have a joint research agreement with the BSC.

How do you see the BSC in 10 years' time?
Strongly focused on research on hardware and software, personalised medicine and energy, and working with business to create wealth.